



FACULTAD DE CIENCIAS HUMANAS Y SOCIALES

**DESARROLLO DEL RAZONAMIENTO ESTADÍSTICO  
EN ESTUDIANTES DE ENFERMERÍA.**

Autora: Soledad Ferreras Mencía

Directora: Belén Urosa Sanz

Madrid  
Septiembre, 2014



## AGRADECIMIENTOS

Deseo expresar mi gratitud a la directora de esta tesis, Dña. Belén Urosa Sanz, por su confianza en mi trabajo, por dejarme hacer y por el interés y la cordialidad con la que lo ha dirigido y me ha guiado. Le agradezco profundamente sus excelentes consejos y sus acertadas opiniones compartiendo conmigo cada idea que pudiera hacer de esta tesis un mejor trabajo. Valoro especialmente sus conocimientos y su experiencia y, sobre todo, la capacidad de comprensión de mis particulares ritmos y dificultades. También quiero agradecer a Belén algo que, generosamente, me ha dado sin tenerlo: su tiempo.

Mi agradecimiento, desde el fondo de mi corazón, a mis alumnos que han dotado de contenido, significado y valor a este trabajo. Ellos son el motor de mi vida profesional y por los únicos que inicié y finalicé esta tesis. Me siento agradecida a todos aquellos que me sorprendieron con una frase, una reflexión o un comentario del que aprendí algo y que consiguieron, con sus “chispas” de conocimiento, prender la luz que me mostró el camino de esta tesis. Gracias queridos alumnos porque, aunque la estadística no esté entre vuestros intereses, siempre habéis mostrado que confiabais en aprender conmigo y esto me ha motivado para indagar otras maneras de enseñar de la mejor forma posible. Gracias porque, aunque yo me dedico a la enseñanza, nunca he dejado de aprender con vosotros.

Quiero agradecer a la gran Institución y familia hospitalaria de San Juan de Dios, donde tengo la suerte y el privilegio de contar con excelentes amigos y que ha sido determinante en mi vida, los valores que me ha transmitido y que han sido el soporte que ha hecho posible mi crecimiento personal y profesional. Estos valores han inspirado mi labor docente y mi compromiso con la formación de los futuros profesionales de Enfermería. Quiero dejar constancia en estos agradecimientos de cómo cambió mi vida el día que atravesé las puertas de este Centro en el que se encuentra la Escuela de Enfermería y Fisioterapia en la que trabajo como docente y donde he encontrado compañeros y amigos que durante estos años demostraron que su apoyo hacia mí era sincero.

Un enorme agradecimiento a mis padres porque, un día hace muchos años y a pesar de las dificultades, me dejaron venir a estudiar a Madrid y trabajaron duro para que finalizara mis estudios. Especialmente a mi madre que, durante los últimos veranos, ha visto preocupada cómo dedicaba mis días de descanso a la redacción de esta tesis. Y a mis hijos. Mi esperanza es que al verme trabajar y dedicar tantas horas a este objetivo haya sido una pequeña lección de vida que les ayude a entender, en los peores momentos, que todo esfuerzo al

final, de una forma o de otra, tiene su recompensa. Espero y deseo que les ilumine en ese camino de futuro que acaban de emprender.

Muchos son los compañeros que, a lo largo de tantos años de docencia, me han acompañado en mi devenir profesional. No puedo dejar de mencionar aquí a Rosa Fernández, compañera de tantas fatigas como éxitos en la consecución de nuestros objetivos; Rosa Rodríguez, con quien he compartido tantos momentos inolvidables de mi vida profesional y personal; Luis Alberto Catalán, compañero leal, fiel y firme defensor de nuestros principios con quien he conseguido construir un centro que, aunque en la actualidad está pasando por momentos difíciles, estoy segura que conseguiremos que siga obteniendo los fantásticos resultados de siempre; Juan Manuel Arribas, cuya tenacidad ha sido un ejemplo para mí a lo largo de todos estos años, y con quien he compartido temporalmente parte del camino en la realización de nuestras respectivas tesis. Le deseo lo mejor en su etapa final.

No puedo dejar de mencionar a dos compañeros muy especiales en mi vida. Juanma Morillo, quien fue mi compañero del alma en los momentos más importantes de mi vida y, César Lama, con quien he aprendido que la amistad se basa en pequeños gestos y en la humildad de no pretender grandes protagonismos en la vida de las personas a las que queremos y para las que, de verdad, somos importantes.

Finalmente, la música, ese escenario mágico que nunca he querido abandonar a lo largo de este camino y que me ha acompañado con personas, emociones, experiencias estéticas y personales que me han enriquecido y me han dado fuerza en los momentos de flaqueza.

Nada de lo que conseguimos y es reconocido como un logro personal se debe únicamente a nuestro trabajo y empeño. Cuando llegamos a un destino, muchas personas han puesto algo más o menos importante. A todas ellas, muchas gracias.

*A mis alumnos*

**Our students must understand that, when we try to grow, we sometimes fail at first, and that failure is all right. The trick is to help students understand that failure holds the seeds of later success.**

Stiggins, R.J. "Assessment, Student Confidence, and School Success." *Phi Delta Kappan* (November) (1999): 191-198.



## ÍNDICE

---

**PLANTEAMIENTO GENERAL DE LA INVESTIGACIÓN**


---

<b>INTRODUCCIÓN.....</b>	<b>21</b>
<b>IDENTIFICACIÓN DEL PROBLEMA .....</b>	<b>26</b>
<b>FINALIDAD Y OBJETIVOS DE LA TESIS. ....</b>	<b>27</b>
<b>ESTRUCTURA GENERAL DEL TRABAJO .....</b>	<b>27</b>

---

**CONTEXTUALIZACIÓN DEL PROBLEMA DE INVESTIGACIÓN**


---

<b>CAPÍTULO 1. LA ESTADÍSTICA EN LOS ESTUDIOS DE GRADO DE ENFERMERÍA.....</b>	<b>33</b>
1.1 El aprendizaje basado en competencias.....	33
1.2. La Enfermería en el Espacio Europeo de Educación Superior .....	35
1.3. Competencia estadística y de investigación en Enfermería .....	36
1.4. Desarrollo de la competencia estadística y de investigación en los documentos normativos de los estudios de grado de Enfermería en España.....	39
1.5. Competencia estadística y de investigación en las especialidades profesionales.....	42
1.6. Estudio comparativo del abordaje de las competencias estadísticas y de investigación en diferentes Universidades españolas.....	43
1.7. La accesibilidad estadística como medición de la eficacia de la formación estadística .	50
<b>CAPÍTULO 2. ENSEÑANZA Y APRENDIZAJE DE LA ESTADÍSTICA .....</b>	<b>59</b>
2.1. El constructivismo.....	59
2.1.1. Evolución del constructivismo y tendencias principales.....	60
2.1.2. Constructivismo y educación .....	61
2.1.3. El aprendizaje significativo.....	62
2.1.3.1. La construcción de significados. ....	63
2.1.3.2. El aprendizaje significativo desde diferentes perspectivas constructivistas ....	65

2.1.4. Estrategias de Aprendizaje en el ámbito universitario.....	68
2.1.5. La enseñanza de la estadística desde una perspectiva constructivista.....	72
2.2. Evolución histórica de la enseñanza de la estadística.....	77
2.2.1. Antecedentes históricos.....	77
2.2.2. Conceptos emergentes en la didáctica de la estadística.....	79
2.2.3. El cambio en la enseñanza de la estadística.....	82
2.2.3.1. Aprendizaje activo.....	83
2.2.3.2. Utilización de datos reales.....	84
2.2.3.3. Generar sus propios datos y trabajar con ellos.....	85
2.2.3.4. Actividades y simulaciones con ordenador y casos prácticos.....	85
2.2.3.5. Aprendizaje cooperativo.....	86
2.2.3.6. Cambio conceptual, trabajo de errores.....	87
2.3. Directrices para el aprendizaje de la Estadística. Guidelines for Assessment and Instruction in Statistics Education (GAISE).....	88
2.3.1. Desarrollar la cultura y el pensamiento estadístico.....	89
2.3.2. Usar datos reales.....	89
2.3.3. Importancia de la comprensión conceptual, más que el mero conocimiento de los procedimientos.....	89
2.3.4. Fomentar el aprendizaje activo en el aula.....	90
2.3.5. Usar la tecnología para el desarrollo de la comprensión conceptual y el análisis de datos.....	90
2.3.6. Utilizar las evaluaciones para mejorar el aprendizaje del estudiante y no sólo para valorarlo.....	91
<b>CAPÍTULO 3. EL RAZONAMIENTO ESTADÍSTICO Y SU EVALUACIÓN.....</b>	<b>93</b>
3.1. Desarrollar el razonamiento estadístico.....	93
3.2. La literatura acerca del razonamiento estadístico de los estudiantes.....	96
3.3. Evaluación como aprendizaje y para el aprendizaje.....	103
3.4. La revolución de la evaluación de la estadística.....	105
3.5. Evaluación para conseguir un estudio razonado de la estadística.....	106
3.6. Instrumentos para la evaluación del razonamiento estadístico.....	110
3.6.1. Descripción de los principales instrumentos revisados.....	111
3.6.2. Valoración de los instrumentos disponibles.....	115



<b>CAPÍTULO 4. METODOLOGÍA .....</b>	<b>119</b>
4.1. Objetivos de la investigación .....	119
4.2. Objetivos instrumentales.....	120
4.3. Hipótesis .....	120
4.4. Diseño general del estudio .....	121
4.5. Participantes .....	122
4.6. Variables e instrumentos .....	126
4.6.1. Características del alumno. ....	127
4.6.2. Percepción de la competencia estadística. ....	127
4.6.3. Variables relacionadas con el aprendizaje. ....	134
4.7. Intervención.....	135
<b>CAPÍTULO 5. CONSTRUCCIÓN DE INSTRUMENTOS DE MEDICIÓN DEL RAZONAMIENTO ESTADÍSTICO .....</b>	<b>137</b>
5.1. Antecedentes y justificación .....	137
5.2. Metodología de la construcción de los cuestionarios de evaluación de razonamiento estadístico.....	143
5.3. Cuestionario para la Evaluación del Razonamiento Estadístico (CERES). ....	146
5.3.1. Tabla de especificaciones de los resultados de aprendizaje que evalúa CERES.....	149
5.3.2. Alineación del cuestionario CERES / GAISE (Guidelines for Assessment and Instruction in Statistics Education GAISE 2010). ....	152
5.4. Validez del Cuestionario de Evaluación del Razonamiento Estadístico (CERES).....	154
5.4.1. Validez de contenido.....	155
5.4.1.1. Opinión de expertos. ....	156
5.4.2. Prueba piloto del cuestionario. Validez discriminante.....	165
5.4.2.1. Análisis descriptivo del cuestionario piloto .....	166
5.4.2.2. Análisis de ítems del cuestionario piloto .....	168
5.4.3. Validez de constructo.....	187
5.4.4. Validez convergente.....	189
5.4.4.1. Prueba de evaluación del razonamiento estadístico de respuesta abierta parcialmente restringida.....	190
5.4.4.2. Correlación con las calificaciones de la prueba de resolución de problemas.	191
5.5. Análisis de Fiabilidad.....	191

<b>CAPÍTULO 6. INTERVENCIÓN Y RECOGIDA DE DATOS .....</b>	<b>193</b>
6.1. Descripción de la intervención.....	193
6.2. Pruebas parciales de razonamiento estadístico (PaRES). ....	200
6.2.1. Resultados de aprendizaje de las pruebas parciales de razonamiento estadístico (PaRES). ....	203
6.2.2. Alineación de los cuestionarios parciales PaRES / GAISE (Guidelines for Assessment and Instruction in Statistics Education GAISE 2010). ....	209
6.2.3. Pruebas piloto de los cuestionarios parciales. ....	211
6.2.3.1. Análisis de ítems de PaRES 1.....	211
6.2.3.2. Análisis de ítems de PaRES 2.....	215
6.2.3.3. Análisis de ítems de PaRES 3.....	219
6.2.3.4. Análisis de ítems de PaRES 4.....	223
6.3 Recogida de Datos .....	226
<b>CAPÍTULO 7. RESULTADOS DE LA INTERVENCIÓN DE EVALUACIÓN .....</b>	<b>231</b>
7.1. Análisis descriptivos .....	231
7.1.1. Análisis descriptivo del Cuestionario de evaluación del razonamiento estadístico CERES.....	231
7.1.2. Prueba de razonamiento estadístico de respuesta abierta.....	233
7.1.3. Examen práctico de resolución de problemas .....	234
7.1.4. Escala de percepción de la competencia estadística.....	236
7.2. Contrastes de hipótesis.....	237
7.2.1. Hipótesis 1.....	237
7.2.1.1. Comparativa de las respuestas correctas en CERES1 y CERES2.....	242
7.2.1.2. Diferencias en el razonamiento estadístico, medido con la prueba de respuesta abierta, entre el grupo de intervención y el grupo control.....	248
7.2.2. Hipótesis 2.....	249
7.2.3 Hipótesis 3.....	250
7.2.4. Hipótesis 4.....	253
7.2.5. Hipótesis 5.....	254
7.2.6. Hipótesis 6.....	254
7.2.7. Hipótesis 7.....	254

---

**CONCLUSIONES DE LA INVESTIGACIÓN**


---

<b>CAPÍTULO 8. CONCLUSIONES .....</b>	<b>259</b>
8.1. Objetivos y conclusiones .....	259
8.2. Limitaciones .....	263
8.3. Reflexiones, aportaciones e implicaciones prácticas.....	265

---

**REFERENCIAS BIBLIOGRÁFICAS**


---

<b>REFERENCIAS BIBLIOGRÁFICAS .....</b>	<b>277</b>
---	------------

---

**ANEXOS**


---

<b>ANEXO 1. CUESTIONARIO DE EVALUACIÓN DE RAZONAMIENTO ESTADÍSTICO (CERES).....</b>	<b>307</b>
<b>ANEXO 2. CUESTINARIO PARCIAL DE RAZONAMIENTO ESTADÍSTICO (PaRES1). ESTADÍSTICA DESCRIPTIVA .....</b>	<b>335</b>
<b>ANEXO 3. CUESTINARIO PARCIAL DE RAZONAMIENTO ESTADÍSTICO (PaRES2) . DISTRIBUCIÓN NORMAL E INTERVALOS DE CONFIANZA.....</b>	<b>349</b>
<b>ANEXO 4. CUESTINARIO PARCIAL DE RAZONAMIENTO ESTADÍSTICO (PaRES3). PRUEBAS DE SIGNIFICACIÓN.....</b>	<b>357</b>
<b>ANEXO 5. CUESTINARIO PARCIAL DE RAZONAMIENTO ESTADÍSTICO (PaRES4) . REGRESIÓN LINEAL Y ANÁLISIS DE VARIANZA.....</b>	<b>369</b>
<b>ANEXO 6. PRUEBA DE RAZONAMIENTO ESTADÍSTICO CON RESPUESTA ABIERTA PARCIALMENTE RESTRINGIDA .....</b>	<b>377</b>
<b>ANEXO 7. CUESTIONARIO ALUMNOS ESTADÍSTICA (ON-LINE).....</b>	<b>387</b>



## INDICE DE FIGURAS

Figura 1. Menciones a la formación estadística en enfermería, recogida en los documentos fundamentales de la American Association of Colleges of Nursing (AACN). .....	38
Figura 2. Competencias de investigación que aparecen en los documentos normativos de los planes de estudio de Enfermería. ....	41
Figura 3. Formación investigadora en los programas de las especialidades de Enfermería. ....	42
Figura 4. Clasificación de las categorías de análisis.....	55
Figura 5. Accesibilidad estadística.....	57
Figura 6. Perspectivas constructivistas del aprendizaje significativo.....	68
Figura 7. Elementos constructivistas en la enseñanza de la estadística. ....	76
Figura 8. Tareas que pueden distinguir los tres dominios de instrucción.....	81
Figura 9. Enfoques formativos en la enseñanza de la estadística. ....	96
Figura 10. Revisión bibliográfica. Estudios que tratan sobre aspectos del razonamiento estadístico. ....	103
Figura 11. Edad de los componentes de la muestra. ....	122
Figura 12. Distribución por sexo de los componentes de la muestra. ....	123
Figura 13. Estudios de acceso de los alumnos de la muestra. (N=109).....	123
Figura 14. Comparativa de la frecuencia de edad en los grupos A y B. ....	124
Figura 15. Frecuencia de hombres y mujeres en los grupos de comparación. ....	124
Figura 16. Frecuencias de los estudio de acceso en los grupos de comparación. ....	125
Figura 17. Instrumentos utilizados para la obtención de las medidas de las variables de estudio. ....	126
Figura 18. Formulación de los ítems que componen la Escala de Percepción de la competencia Estadística (EpCE). ....	131
Figura 19. Gráfico de sedimentación del análisis factorial de la escala de percepción de la competencia estadística (EPCE). ....	133
Figura 20. Ventajas de las pruebas objetivas para medir el razonamiento estadístico. ....	139

Figura 21. Instrumentos de Evaluación del Razonamiento Estadístico.....	140
Figura 22. Ítems de CAOS que conforman el cuestionario CERES piloto.....	149
Figura 23. Contenidos y resultados de aprendizaje del Cuestionario para Evaluar el Razonamientos Estadístico (CERES).....	152
Figura 24. Alineación de los ítems de CERES con las recomendaciones GAISE.....	154
Figura 25. Gráfico de barras de la frecuencia del número de respuestas correctas de CERES piloto.....	166
Figura 26. Gráfico de cajas del número de respuestas correctas de CERES piloto.....	167
Figura 27. Gráfico de cajas de la comparación de grupos extremos de CERES piloto.....	168
Figura 28. Descripción y análisis de normalidad de los índices de dificultad del Cuestionario de Evaluación de Razonamiento Estadístico piloto (CERES piloto).....	175
Figura 29. Estudio de normalidad de los índices de discriminación (DC2) de los ítems de CERES piloto.....	182
Figura 30. Gráfico de sedimentación del análisis factorial de CERES piloto.....	188
Figura 31. Ejemplos de estudios de investigación educativa con intervención realizados con grupos pequeños.....	199
Figura 32. Ítems de las escalas ARTIST que conforman PaRES1.....	200
Figura 33. Ítems de las escalas ARTIST que conforman PaRES2.....	201
Figura 34. Ítems de las escalas ARTIST que conforman PaRES3.....	202
Figura 35. Ítems de las escalas ARTIST que conforman PaRES4.....	202
Figura 36. Resultados de aprendizaje que evalúa el cuestionario parcial de razonamiento estadístico PaRES1.....	204
Figura 37. Resultados de aprendizaje que evalúa el cuestionario parcial de razonamiento estadístico PaRES2.....	206
Figura 38. Resultados de aprendizaje que evalúa el cuestionario parcial de razonamiento estadístico PaRES3.....	207
Figura 39. Resultados de aprendizaje que evalúa el cuestionario parcial de razonamiento estadístico PaRES4.....	208
Figura 40. Alineación de los cuestionarios parciales (PaRES) 1 – 4 con GAISE 2010.....	211
Figura 41. Resultados del análisis de ítems de PaRES1.....	215

---

Figura 42. Resultados del análisis de ítems de PaRES2. ....	218
Figura 43. Resultados del análisis de ítems de PaRES3. ....	222
Figura 44. Resultados del análisis de ítems de PaRES4. ....	225
Figura 45. Cronograma del proceso de recogida de datos.....	227
Figura 46. Descripción del número de respuestas correctas al cuestionario de razonamiento estadístico (CERES) en el pretest y en el postest.....	232
Figura 47. Descriptivo de las puntuaciones de la prueba de razonamiento estadístico de respuesta abierta (Prueba RA).....	233
Figura 48. Descriptivo de las calificaciones del examen práctico de resolución de problemas. ....	235
Figura 49. Puntuaciones de la Escala de percepción de la competencia estadística. ....	236
Figura 50. Gráfico de perfil de las medias marginales estimadas de CERES de los grupos de intervención y grupo control.....	241
Figura 51. Gráfico de perfil de las medias marginales estimadas de EPCE de los grupos de intervención y grupo control.....	252





## ÍNDICE DE TABLAS

---

Tabla 1 Artículos relacionados con la enseñanza de la estadística desde 2007.....	24
Tabla 2. Estadística e Investigación presentes en los planes de estudio del grado de Enfermería de las universidades españolas. ....	43
Tabla 3. Matriz de componentes rotados de la escala de percepción de la competencia estadística (EPCE). ....	134
Tabla 4. Descriptivos de la valoración realizada por el grupo de expertos acerca de los resultados de aprendizaje del cuestionario de evaluación de razonamiento estadístico. ....	159
Tabla 5. Relevancia asignada por los expertos a los resultados de aprendizaje. ....	161
Tabla 6. Análisis de ítems de CERES piloto. ....	171
Tabla 7. Prueba de bondad de ajuste para la normalidad de los índices de dificultad del Cuestionario de Evaluación de Razonamiento Estadístico piloto (CERES piloto).....	175
Tabla 8. Distribución de ítems de CERES piloto según su nivel de dificultad. ....	180
Tabla 9. Índices de discriminación de los ítems de CERES piloto .....	181
Tabla 10. Frecuencias de los Índices de discriminación (DC2) de los ítems de CERES piloto. ....	181
Tabla 11 . Correspondencia de numeración y modificaciones entre las dos versiones del cuestionario.....	186
Tabla 12. Extracción inicial de factores del cuestionario CERES piloto. ....	187
Tabla 13. Matriz de componentes rotados del análisis factorial del cuestionario CERES piloto..	189
Tabla 14. Análisis de ítems de PaRES 1. ....	211
Tabla 15. Análisis de ítems de PaRES 2. ....	215
Tabla 16. Análisis de ítems de PaRES 3. ....	219
Tabla 17. Análisis de ítems de PaRES 4. ....	223
Tabla 18. Prueba de Levene de igualdad de varianzas y estadísticos descriptivos de la comparación de CERES en el pretest y en el postest. ....	238
Tabla 19. Análisis de varianza de medidas repetidas. Diferencias en las puntuaciones de CERES1 y CERES2 entre el grupo control y el grupo de intervención. ....	239

Tabla 20. Pruebas de efectos intra-sujetos. Análisis de varianza de medidas repetidas. Diferencias en las puntuaciones de CERES1 y CERES2 entre el grupo control y el grupo de intervención. ....	239
Tabla 21. Medias marginales del cuestionario de razonamiento estadístico (CERES) pre y postest. ....	240
Tabla 22. Pruebas de los efectos inter-sujetos. Análisis de varianza de medidas repetidas. Diferencias en las puntuaciones de CERES1 y CERES2 entre el grupo control y el grupo de intervención. ....	241
Tabla 23. Ítems del cuestionario CERES con un porcentaje de respuestas correctas superior al 60% en el pretest y en el postest.....	243
Tabla 24. Ítems del cuestionario CERES con un porcentaje de respuestas correctas en el postest superior al 50%.....	244
Tabla 25. Ítems del cuestionario CERES con un porcentaje de respuestas correctas en el postest inferior al 50%. ....	246
Tabla 26. Ítems del cuestionario CERES con un aumento en los errores en el postest.....	247
Tabla 27. Prueba t de comparación de las medias de los resultados de la prueba de razonamiento estadístico de respuesta abierta entre el grupo de intervención y el grupo control. ....	248
Tabla 28. Prueba t de comparación de medias de los resultados de la prueba práctica de resolución de problemas entre el grupo de intervención y el grupo control.....	249
Tabla 29. Prueba de Levene de igualdad de varianzas y estadísticos descriptivos de la comparación de la escala de percepción de la competencia estadística (EpCE) en el pretest y en el postest. ....	250
Tabla 30. Análisis de varianza de medidas repetidas. Diferencias en las puntuaciones de EPCE1 y EPCE2 entre el grupo control y el grupo de intervención. ....	251
Tabla 31. Pruebas de efectos intra-sujetos. Análisis de varianza de medidas repetidas. Diferencias en las puntuaciones de EPCE1 y EPCE2 entre el grupo control y el grupo de intervención.....	251
Tabla 32. Pruebas de los efectos inter-sujetos. Análisis de varianza de medidas repetidas. Diferencias en las puntuaciones de EpCE1 y EpCE2 entre el grupo control y el grupo de intervención. ....	253

## **PLANTEAMIENTO GENERAL DE LA INVESTIGACIÓN**

---



## INTRODUCCIÓN

Tras más de veinte años de experiencia docente nuestra responsabilidad hacia el proceso de aprendizaje de los alumnos, no ha hecho más que consolidarse con el paso del tiempo. Esta investigación tiene en estos momentos más sentido del que hubiera tenido hace años. Todo este tiempo dedicado a la docencia nos ha proporcionado la experiencia y el conocimiento necesarios para abordar un trabajo de estas características. A lo largo de los años hemos podido comprobar cómo las políticas educativas han ido modificando los planes de estudios de los diferentes niveles educativos y hemos de decir que, en nuestra opinión, estos cambios no siempre han propiciado mejoras en la enseñanza. Es por este motivo por el que, a pesar de las circunstancias, los docentes debemos velar para que nuestra enseñanza, nutrida por el conocimiento que la experiencia nos proporciona, tenga cada vez más calidad y sea innovadora, porque si un peligro acecha la labor de los docentes, este es, sin duda, la acomodación.

Los profesores comenzamos enseñando de la misma manera que aprendimos, y qué diferente es la forma en la que aprendimos estadística en la Universidad, memorizando fórmulas, desarrollos matemáticos y realizando interminables cálculos, a cómo entendemos hoy la enseñanza de esta materia. Muchos han sido los motivos para este cambio. Uno de ellos fue el avance en tecnología informática y su aplicación a la enseñanza, motivo por el que los cálculos matemáticos dejaron de ser tan importantes. Otro factor ha sido el desarrollo de la comunicación por Internet que nos da acceso a la investigación que se realiza en todo el mundo. Con tanta investigación disponible, la estadística se convierte en una herramienta imprescindible para interpretar y enjuiciar los resultados de investigación, pero también es necesaria para la vida diaria, multitud de cuestiones sociales, políticas y económicas se nos muestran en términos estadísticos.

Los profesores de estadística debemos hacer que los alumnos incorporen esta herramienta y la utilicen en su vida profesional, y para ello los conceptos y su aplicación han de ser razonadamente entendidos.

El presente trabajo ha representado una experiencia educativa, que ha aportado elementos de interés en el desarrollo de nuestra labor docente, permitiendo el análisis de estrategias para el desarrollo del razonamiento estadístico del alumno y su integración en los recursos y competencias que ha de adquirir el futuro enfermero.

Durante todos estos años de experiencia docente, hemos percibido que los alumnos poseen un razonamiento estadístico muy frágil, que se basa en concepciones intuitivas que resultan un obstáculo para el aprendizaje, situación que se agrava por el exceso de aprendizaje memorístico y repetitivo incluso en asignaturas de matemáticas, en las que se aprende por repetición de problemas y ejercicios con la misma formulación sin una base conceptual sólida. Poner el énfasis en los modelos matemáticos, no hace más que empeorar las cosas si no se asocia a una situación empírica que debe guardar la mayor similitud con las situaciones reales y cercanas al alumno.

Muchos han sido los estudios publicados acerca de la docencia de la estadística, sus peculiaridades, las dificultades y la forma de mejorar el aprendizaje. Según unos de los más importantes investigadores en este área (Bjornsdottir, A., Garfield, J. 2008; 2009a y b; 2010; 2012a y b y 2013), las contribuciones científicas publicadas desde 2007 en las revistas internacionales de matemáticas y estadística más prestigiosas superan los 400 artículos relacionados con la enseñanza de la materia en diferentes niveles académicos.

Las publicaciones sobre las que Bjornsdottir y Garfield realizan estas revisiones son las siguientes:

- *Educational Studies in Mathematics Educ Stud Math*

Publicada por Springer. ISSN (edición impresa): 0013-1954. ISSN (edición electrónica): 1573-0816.

*Educational Studies in Mathematics* presenta nuevas ideas y avances para las personas que trabajan en el campo de la educación matemática. Refleja tanto la variedad de problemas de investigación dentro de esta disciplina, como la variedad de métodos utilizados para su estudio. Se ocupa de temas metodológicos y pedagógicos.

- *International Journal for Technology in Mathematics Education Int J Tech Math Educ*

Publicada por Research Information Ltd. ISSN: 1744-2710.

La revista publica propuestas sobre cualquier aspecto del uso de la tecnología en la educación matemática en los diferentes niveles de enseñanza desde la escuela hasta la universidad.

- *Journal for Research in Mathematics Education J Res Math Educ*

Publicada por el National Council of Teachers of Mathematics. ISSN: 0021-8251.

*Journal for Research in Mathematics Education* es una revista oficial del

National Council of Teachers of Mathematics dedicada a los intereses de los profesores de matemáticas y la educación matemática en todos los niveles académicos.

- *Mathematical Thinking and Learning Math Think Learn*

Publicada por Lawrence Erlbaum Associates, Inc. ISSN (edición impresa): 1098-6065. ISSN (edición electrónica): 1532-7833.

Esta revista recoge artículos que abordan temas relacionados con estudios interdisciplinarios sobre el aprendizaje matemático, la diversidad de procesos de razonamiento matemático y nuevos enfoques sobre la evolución de la comprensión matemática a lo largo de las diferentes etapas.

- *Mathematics Teacher Math Teach*

Publicada por el National Council of Teachers of Mathematics. ISSN: 0025-5769.

Es una revista oficial del Consejo Nacional de Profesores de Matemáticas que pretende la mejora de la enseñanza de las matemáticas y el apoyo a los programas de formación del profesorado. Proporciona un foro para compartir actividades y estrategias pedagógicas, para la profundización de la comprensión de las ideas matemáticas, y la vinculación de la investigación en educación matemática con la práctica docente.

- *Mathematics Teaching in the Middle School Math Teach Middle Sch*

Publicada por el National Council of Teachers of Mathematics. ISSN: 1072-0839.

Es una revista oficial del Consejo Nacional de Profesores de Matemáticas y pretende ser un recurso para los estudiantes de secundaria, maestros y formadores de docentes. El objetivo de la revista se centra en investigaciones intuitivas y exploratorias que utilizan el razonamiento informal para ayudar a los estudiantes a desarrollar una sólida base conceptual que lleva a una mayor abstracción matemática.

- *Model Assisted Statistics and Applications*

Publicada por IOS Press. ISSN: 1574-1699.

Publicación que recoge artículos sobre teoría del muestreo, econometría, series temporales, diseño de experimentos y análisis multivariado.

- *Statistics Education Research Journal Stat Educ Res J*

Publicada por la International Association for Statistical Education. ISSN: 1570-1824.

SERJ es una revista electrónica de la Asociación Internacional para la

Educación Estadística (IASE) y el Instituto Internacional de Estadística (ISI) cuyo objetivo es mejorar la enseñanza, el aprendizaje y la comprensión de la estadística en todos los niveles educativos y en contextos formales e informales.

- *Teaching Statistics Teach Stat*

Publicada por John Wiley and Sons. ISSN (edición impresa): 0141-982X. ISSN (edición electrónica): 1467-9639.

Teaching Statistics está dirigida a profesores y estudiantes con el fin de resolver problemas que se plantean en el aula.

- *Technology Innovations in Statistics Education (TISE) Tech Innovat Stat Educ*

Publicada por California Digital Library. ISSN: 1933-4214.

TISE publica documentos que permitan avanzar en la comprensión de cómo enseñar mejor la estadística con el apoyo de la tecnología y cómo enseñar mejor tecnología para futuros profesionales que utilicen la estadística.

- *The American Statistician*

Publicada por Taylor and Francis Group. ISSN (edición impresa): 0003-1305. ISSN (edición electrónica): 1537-2731.

The American Statistician (TAS), publicada trimestralmente por la Asociación Americana de Estadística (TAS), contiene artículos organizados en las siguientes secciones: uso de la Estadística, enseñanza de la Estadística, Historia, informática, gráficos estadísticos y reseñas de libros y materiales didácticos.

La Tabla 1 recoge el número de artículos referidos a la enseñanza de la estadística localizados en las revistas comentadas anteriormente. Estas ocho revisiones, realizadas de los periodos indicados, han sido publicadas en la revista Journal of Statistics Education.

**Tabla 1 Artículos relacionados con la enseñanza de la estadística desde 2007**

Periodo	nº de artículos
2007	150
2008	66
Enero – Noviembre 2009	61
Noviembre 2009 – Junio 2010	34
Enero – Junio 2011	28
Noviembre 2011 – Enero 2012	18
Enero – Noviembre 2012	46
Noviembre 2012 – Enero 2013	30



Con la lectura de algunos de estos interesantes trabajos, comenzamos a considerar que el investigar sobre la metodología específica de la enseñanza de la estadística podía mejorar nuestro trabajo docente, y esto fue lo que dio origen a esta Tesis Doctoral.

En la revisión de la bibliografía, una línea de investigación nos llamó especialmente la atención. Esta línea de investigación fue aquella que analizaba el pensamiento y el razonamiento estadístico, tal como lo define Garfield, (2002), como la forma en la que los alumnos razonan con ideas estadísticas y dan sentido a la información estadística. En el capítulo 3 de este trabajo se analiza y enuncia la definición operativa del concepto de razonamiento estadístico, sobre la que se centra el objetivo fundamental de este estudio.

La asignatura de estadística incluida en el plan de estudios de Enfermería, al igual que la que se imparte en otras muchas titulaciones de grado, ha de considerarse que se estructura como un primer curso de iniciación o curso básico de esta materia seguido, en algunas titulaciones, con cursos superiores de mayor especialización. La mayor parte de los alumnos tienen, en su etapa universitaria, su primer contacto con la estadística, si bien es cierto que se contemplan contenidos de estadística en primaria, enseñanza secundaria y bachillerato.

Hay muchos aspectos importantes en estos cursos introductorios en los que es necesario que el alumno realice una comprensión conceptual y un razonamiento que afiancen el conocimiento para posteriormente aplicarlo. Entre otros ejemplos, el razonamiento sobre la asociación, o razonamiento covariacional, juega un papel importante en un curso de estas características porque conocer e interpretar la relación entre dos variables implica procesos de traducción entre los datos brutos numéricos, las representaciones gráficas y las repercusiones que tiene el que esa asociación sea o no causal. Conceptos relacionados con datos de dos variables, tanto categóricas como cuantitativas, se encuentran en los contenidos de esta asignatura básica e incluyen diagramas de dispersión, tablas de contingencia, correlación, regresión y chi-cuadrado.

Los profesores sabemos la dificultad que tiene transmitir al alumno algunos conceptos estadísticos como los relacionados con las distribuciones de probabilidad, la significación estadística o la variabilidad, entre otros muchos. Ponemos empeño en que lo entienda tratando de guiarle hacia un enfoque profundo en su aprendizaje, y le pedimos que desarrolle habilidades de razonamiento, porque de otra manera el aprendizaje de calidad no está garantizado. Y después, ¿evaluamos si de verdad es capaz de razonar con ideas y conceptos estadísticos?, o ¿simplemente le pedimos que resuelva una serie de problemas de cálculo, ayudado o no por programas informáticos de análisis

estadístico?. De esta última manera los alumnos, que actúan de la forma más eficiente para superar la asignatura, aprenden lo suficiente para resolver los problemas sin entender los conceptos, y se muestran más preocupados por manejar el programa informático que por saber qué es lo que están haciendo y por qué lo hacen. Desde luego no es fácil hacer preguntas diferentes a las que encontramos en los libros de texto, preguntas que pongan a prueba que efectivamente los estudiantes comprenden la estadística, aunque es necesario ese esfuerzo para que exista coherencia entre cómo enseñamos y cómo queremos que aprendan nuestros alumnos.

## IDENTIFICACIÓN DEL PROBLEMA

Como responsable de la formación en estadística de la titulación de Enfermería, consciente de las dificultades que entraña la enseñanza de esta materia respecto a la adquisición por parte de los alumnos de algunos conceptos y, sobre todo, preocupados porque sea una herramienta útil en el ámbito social y profesional se nos planteaban una serie de interrogantes que son los que guían este trabajo:

1º ¿Qué saben los alumnos universitarios de Enfermería de primer curso al finalizar la asignatura de estadística?

2º ¿Entienden los conceptos más importantes y usan los conocimientos estadísticos básicos para leer y enjuiciar la información científica y del mundo que les rodea?

3º ¿Qué queremos en realidad que aprendan?

4º ¿De qué manera hemos de enseñarlo?

Las dos primeras preguntas se enmarcan fundamentalmente en el ámbito de la evaluación de los resultados de aprendizaje que en esta investigación, son resultados fundamentados en la capacidad de razonamiento estadístico. Se intentará darles respuesta con la elaboración de un instrumento que pueda proporcionar a los profesores un modelo de evaluación diseñado para valorar en los estudiantes el desarrollo del razonamiento estadístico y la conexión de ideas estadísticas en el contexto de un problema.

Las preguntas tercera y cuarta nos conducen a la elección de unas opciones metodológicas entre las que se incluye, en el proceso de enseñanza y aprendizaje, un sistema de evaluación que ayude a desarrollar en el alumno el

razonamiento estadístico y la capacidad de aplicar este conocimiento en el marco de los problemas de investigación en el ámbito sanitario.

## **FINALIDAD Y OBJETIVOS DE LA TESIS.**

La finalidad última de esta investigación es mejorar en nuestros alumnos el proceso de enseñanza-aprendizaje de la estadística, desarrollar el razonamiento estadístico, y mejorar el rendimiento académico.

La finalidad fundamental de esta investigación es realizar una acción educativa fundamentada en la evaluación, utilizando unas pruebas específicas para evaluar el razonamiento estadístico. Se trata de una acción continuada de evaluación que ofrecerá al alumno información de retorno que le permitirá procesar los errores, las contradicciones e incoherencias. Todo ello generará un proceso de autorregulación que le ayudará a valorar la relación entre lo que sabe y el éxito en la tarea que se le propone. Le facilitará la comprensión de conceptos y la interpretación de resultados, le entrenará en el razonamiento y le facilitará el logro de un aprendizaje más autónomo.

En esta línea los objetivos que se plantean para este trabajo se pueden concretar de la siguiente forma:

**Objetivo 1:** Realizar una búsqueda y análisis de instrumentos adecuados para medir el razonamiento estadístico.

**Objetivo 2:** Desarrollar una serie de nuevos cuestionarios del razonamiento estadístico sobre los conceptos fundamentales que los alumnos universitarios de Enfermería han de adquirir en un curso de introducción a la estadística.

**Objetivo 3:** Identificar cuáles son los logros del aprendizaje y de la comprensión conceptual de la estadística de los estudiantes a los que se aplica una acción formativa basada en la evaluación que tenga como objetivo fundamental el desarrollo del razonamiento estadístico.

## **ESTRUCTURA GENERAL DEL TRABAJO**

Expondremos a continuación la estructura del trabajo que comprende dos grandes apartados. El primero de ellos se dedica a contextualizar el

problema de investigación mientras que en el segundo se desarrolla el estudio empírico.

En el capítulo 1 se describe el panorama de la enseñanza de la estadística en los estudios de enfermería, desde el punto de vista del aprendizaje basado en competencias, teniendo presente que esta materia ha de ser un instrumento que aumente la competencia investigadora de la enfermería. La capacidad de poder mejorar la práctica enfermera incorporando los resultados de investigación necesita de una formación que dote a los profesionales, entre otras cosas, de una “accesibilidad estadística” que le permita hacer juicios de la mejor evidencia disponible.

El capítulo 2 comienza con una revisión de las teorías constructivistas del aprendizaje, pues es en ese marco en el que podemos entender los elementos que desarrollan el razonamiento estadístico en los alumnos y que han de incorporarse a la enseñanza de esta materia. Alfabetización, pensamiento y razonamiento estadístico, son términos que han sido definidos y conceptualizados de forma diferente. Los matices diferenciadores entre ellos son revisados en este capítulo. Estos conceptos nacen en el seno de un cambio que se produce en la forma de entender la enseñanza tradicional de la estadística. Aparecen estudios acerca de cómo pueden los estudiantes mejorar su aprendizaje, se analizan los errores en el razonamiento estadístico, se comprueban los resultados de diferentes experiencias educativas y nuevas formas de evaluar. Toda esta literatura es revisada y resumida en este capítulo.

El capítulo 3 se centra en la evaluación del razonamiento estadístico, en la revisión de los distintos instrumentos encontrados, analizándolos desde el punto de vista de su posible utilidad en nuestra investigación. Se expone un planteamiento de la evaluación “para” el aprendizaje y “como” aprendizaje, más allá de su función calificadora.

En el capítulo 4, que introduce la parte empírica presentando la metodología de la investigación, los objetivos y las variables del estudio, se enuncian las principales hipótesis y se describe de forma general el procedimiento y los análisis realizados. En este mismo capítulo se presentan los distintos instrumentos que se han utilizado en este trabajo. Se describe el tipo de instrumento, los criterios y los pasos que han estado presentes en su elaboración y cuál es su finalidad en nuestra investigación y se describen las características psicométricas y los análisis de fiabilidad y validez realizados en el estudio piloto de los instrumentos.

---

Se describen los cambios en los instrumentos al tiempo que se expone el proceso llevado a cabo para la recogida de datos.

El capítulo 5 está dedicado a la exposición de los resultados de los análisis realizados para verificar las hipótesis del estudio.

Por último, el capítulo 6 relata las conclusiones y consideraciones más importantes extraídas de la revisión teórica y de los análisis realizados en la parte empírica de la tesis. Expone también las limitaciones encontradas y las posibles líneas de trabajo que se vislumbran en el ámbito del desarrollo y evaluación del razonamiento estadístico en los alumnos.



## **CONTEXTUALIZACIÓN DEL PROBLEMA DE INVESTIGACIÓN**





## **CAPÍTULO 1. LA ESTADÍSTICA EN LOS ESTUDIOS DE GRADO DE ENFERMERÍA.**

### **1.1 El aprendizaje basado en competencias**

La Unión Europea ha propiciado la convergencia en distintos ámbitos entre los que se encuentra el educativo. En este aspecto se ha impulsado un importante movimiento encaminado al desarrollo de un Espacio Europeo de Educación Superior que permite un reconocimiento más fácil de las titulaciones al tiempo que intenta asegurar una mejor formación de los estudiantes y su integración en un mercado laboral unificado y sin fronteras. Entre los factores que han contribuido a ello están los programas de movilidad de estudiantes Sócrates y Erasmus, que han determinado la necesidad de encontrar un sistema adecuado de equivalencias y reconocimiento de estudios; este ha sido el origen del Sistema Europeo de Transferencia de Créditos o ECTS (European Credit Transfer System). Dicho sistema ha sido la clave para la transferencia y el reconocimiento de los períodos cursados en otros países, junto con un método de trabajo basado en la utilización de principios compartidos y documentos con formatos normalizados.

Esta nueva dimensión en el panorama educativo europeo se inicia con las Declaraciones de la Sorbona (1998), Bolonia (1999) y Praga (2001) y las posteriores de los Ministerios de Educación de los países de la Unión europea. Con la Declaración de la Sorbona, se propone, desde algunos países europeos, la necesidad de promover la convergencia entre los sistemas nacionales de educación superior. En 1999, los Ministros de Educación de cada país miembro de la Unión Europea refrendaron, con la firma de la Declaración de Bolonia, la importancia de un desarrollo armónico de un Espacio Europeo de Educación Superior. Europa y sus universidades se enfrentaron a un cambio social y cultural y al reto de trabajar seriamente para realizar su labor docente e investigadora en un nuevo marco conjunto.

Aunque es evidente que la estructura de la formación universitaria ha tenido que transformarse para poder armonizarse con el resto de países, la verdadera transformación ha estado en la concepción pedagógica de la formación. El objetivo es desarrollar perfiles profesionales equiparables y por tanto la adquisición de competencias profesionales.

En este momento más de 40 países de la Unión Europea y área de influencia, han firmado el acuerdo de incorporación al Espacio Europeo de Educación Superior. En estos últimos años también se han desarrollado acciones de colaboración concreta con los países de América Latina y el Caribe.

En este nuevo marco, el objetivo principal del proceso de aprendizaje consiste no sólo en la adquisición de conocimientos por parte del estudiante, sino que queda supeditado al desarrollo de una serie de competencias. El término competencia, toma como referente los resultados de aprendizaje, es decir las capacidades y destrezas, lo que el alumno es capaz de hacer al finalizar el proceso educativo, y lo que le permitirá seguir aprendiendo de forma autónoma a lo largo de su vida. El objetivo principal es algo más complejo que el simple dominio o transmisión de conocimientos, como ha sucedido tradicionalmente, se ha pasado de una educación centrada en la enseñanza a una educación centrada en el aprendizaje. Este nuevo escenario nos está llevando, a los docentes, a una manera de enfocar la docencia y el aprendizaje que tiene consecuencias en lo que hacemos y en lo que pedimos hacer a nuestros alumnos (Morales 2012). Este trabajo de Morales cita a Dysthe (2007), quien constata a través de una amplia muestra de profesores noruegos cómo la Declaración de Bolonia ha cambiado las tareas encomendadas a los alumnos. El alumno realiza más trabajo de forma autónoma, se utiliza más la metodología activa en clase, se persigue una relación más estrecha entre métodos de enseñanza y evaluación que produce un mayor feedback con el alumno y se hace un uso mayor de las tecnologías.

Los objetivos de las titulaciones y de las distintas asignaturas que las conforman se basan, no solamente en la simple acumulación de conocimientos, sino también en las habilidades profesionales y las actitudes personales que permiten a los estudiantes ocupar el lugar correspondiente en el mundo laboral y en la sociedad.

El Proyecto Tuning Educational Structures in Europe (<http://www.unideusto.org/tuningeu/>) define la competencia como “una combinación dinámica de atributos, en relación a conocimientos, habilidades, actitudes y responsabilidades, que describen los resultados del aprendizaje de un programa educativo o lo que los alumnos son capaces de demostrar al final de un proceso educativo” (González y Wagenaar 2003). La OCDE en el Informe DeSeCo (Descripción y Selección de Competencias clave) define las competencias como “la capacidad para responder a las demandas y llevar a cabo tareas de forma adecuada”. Cada competencia se construye a través de la combinación de

habilidades cognitivas y prácticas, conocimiento, motivación, valores, actitudes, emociones y otros componentes sociales y conductuales” (Symone y Hersh 2006). En resumen, las competencias son lo que se espera que los graduados conozcan, comprendan o hagan, al tiempo que posibilitan un lenguaje común entre académicos y profesionales para expresar el perfil que hay que desarrollar a través de la formación y para el empleo.

Los objetivos por tanto de cualquier plan de estudios es el desarrollo de un conjunto de competencias profesionales en los alumnos. Estas competencias se clasifican en genéricas y específicas.

Se definen como competencias genéricas del título aquellas que aportan las habilidades necesarias para el empleo y para la vida como ciudadanos responsables y son importantes para todos los alumnos independientemente de la disciplina que están estudiando. Las competencias específicas de la titulación están vinculadas al perfil profesional aportándole su identidad y consistencia. Las competencias específicas de área y asignatura, se relacionan con los conocimientos, habilidades y actitudes que debe desarrollar la asignatura o el conjunto de asignaturas de una materia. La adquisición de las competencias genéricas y algunas de las específicas se produce de manera progresiva a lo largo de la titulación y presenta diferentes niveles de desarrollo dependiendo del curso en el que estén ubicadas.

## **1.2. La Enfermería en el Espacio Europeo de Educación Superior**

La estructura de las enseñanzas y títulos universitarios oficiales dentro del proceso de convergencia con los principios de la construcción del Espacio Europeo de Educación Superior (R.D. 1393/2007), ha supuesto un cambio sustancial en los estudios de Enfermería. En el pasado, el Diplomado Universitario de Enfermería realizaba una formación académica durante tres años que le capacitaba profesionalmente para la prestación de cuidados, pero sin acceso a niveles académicos superiores dentro de la misma titulación. A diferencia de lo ocurrido con otras diplomaturas, nunca ha existido una licenciatura a cuyo segundo ciclo se pudiera acceder como continuación natural de los estudios de Enfermería. Muchos profesionales de esta especialidad han cursado otras licenciaturas que, aun siendo muy interesantes, e incluso estando relacionadas con la profesión (Humanidades) parecían ser un escollo que habían de superar para poder continuar con los estudios de doctorado.

La actual estructura de la enseñanza universitaria de Enfermería ha aportado cambios sustanciales entre los que se encuentran, la obtención del título de grado con idénticas competencias y la ampliación de la titulación a cuatro años y la posibilidad de continuar con los niveles de Máster y Doctorado. Esta estructura supone la equiparación básica de estos estudios con el resto de titulados, ya que las antiguas licenciaturas y diplomaturas han pasado todas a ser titulaciones de grado que se desarrollan en cuatro años académicos.

En la actualidad, los estudios superiores de Enfermería, al igual que el resto de titulaciones, se organiza en tres ciclos: Grado, Máster y Doctorado. El Grado, tiene como finalidad la obtención por parte del estudiante de una formación general orientada al ejercicio profesional. Las enseñanzas de Máster aportan una formación avanzada, de carácter especializado o multidisciplinar, orientada a la especialización científica o profesional y a promover la iniciación en tareas investigadoras. El Doctorado tiene como objeto la formación avanzada en técnicas de investigación.

Otro objetivo fundamental de la nueva organización de las enseñanzas es establecer vínculos dentro del Espacio Europeo de Educación y del Espacio Europeo de Investigación. Con este fin se establece una mayor apertura en la organización de los estudios de doctorado. Dichos estudios de doctorado, han sido acogidos desde la disciplina enfermera con especial interés ya que posibilitan que la Enfermería se integre, como miembro de pleno derecho, en las acciones investigadoras de las Ciencias de la Salud.

Las nuevas enseñanzas no sólo responden a un cambio estructural sino que, como ya hemos comentado, han propiciado innovaciones en la metodología docente. En este nuevo marco, el objetivo principal de los planes de estudio es la adquisición de competencias por parte de los estudiantes y se pone el énfasis en los métodos de aprendizaje y en la evaluación de dichas competencias.

### **1.3. Competencia estadística y de investigación en Enfermería**

Sin duda, las competencias estadísticas y de investigación son fundamentales para la competencia profesional a la que se refiere Bunk (1994) como la capacitación real para resolver problemas contando con conocimientos, destrezas, aptitudes y la flexibilidad y autonomía social. En nuestra opinión, las competencias estadísticas contribuyen a ampliar la participación en el entorno

profesional, así como en la organización del trabajo y en las actividades de planificación.

Existen distintos modelos educativos para desarrollar las competencias estadísticas y de investigación. En algunas universidades americanas tienen como requisito previo que los estudiantes que acceden al Máster y al Doctorado hayan realizado previamente un curso de Estadística de al menos tres créditos, con la recomendación de que los contenidos en Estadística sean aplicados a la práctica enfermera (Burke et al. 2005).

La Universidad de Pittsburgh sigue un modelo de educación en el que la Enfermería basada en la evidencia se encuentra introducida en los distintos niveles académicos, desde el grado al doctorado. Entre las competencias identificadas para los cuatro cursos del grado, el mismo artículo de Burke señala que el alumno, tras haber completado un curso de Estadística básica en investigación, debe ser capaz de evaluar críticamente las evidencias de investigación para aplicar las conclusiones a la práctica clínica en el segundo curso. En esta universidad se desarrollan, en el primer curso del grado, habilidades de búsqueda de la literatura, acceso y aplicación de los resultados de investigación, y en el segundo curso los alumnos adquieren los conocimientos estadísticos que les facilitan la evaluación crítica de la mejor evidencia.

En relación con la investigación en Enfermería, la American Association of Colleges of Nursing (AACN) formula que el objetivo final de la formación en investigación en Enfermería, en todos los niveles, es el fortalecimiento de la profesión, la contribución a la mejora de la salud y la atención a los individuos y a las poblaciones. Define las expectativas y competencias en cada nivel educativo de manera que, sirviéndonos de referencia en el grado, deberíamos contemplar programas que preparen a los enfermeros para comprender los procesos básicos de investigación y para aplicar la mejor evidencia a la práctica diaria, identificar problemas y colaborar en equipos de investigación. Los programas de Máster han de preparar a estos profesionales para evaluar resultados de investigación y para desarrollar habilidades de liderazgo, a fin de que puedan formar y dirigir equipos dentro de sus organizaciones y grupos profesionales. Y en el programa doctoral, se alcanzará el más alto nivel en los conocimientos y técnicas para evaluar la aplicación de la investigación en la práctica, para planificar y poner en marcha investigaciones independientes, y para la formación y supervisión de investigadores.

En las publicaciones de la AACN no aparece ninguna orientación con respecto al desarrollo del currículo, acerca del contenido del curso, o de los

objetivos de aprendizaje de la estadística (Hayat, 2014). Sin embargo la estadística es considerada lo suficientemente importante para que la mayoría de los programas, de todos los niveles académicos, de enfermería incluyan uno o varios cursos sobre esta materia (Hayat, 2014).

Se recogen en la Figura 1 las menciones a la estadística, breves, generales y sin consistencia, que se recogen en las publicaciones de la American Association of Colleges of Nursing (AACN).

Source	Statement
BSN <i>Essentials</i> (AACN, 2008, p. 17)	The phrase "basic applied statistics" is included in a list of topics that should be covered in order to adequately prepare baccalaureate graduates.
MSN <i>Essentials</i> (AACN, 2011, p. 6)	The statement "...course work should provide graduates with the knowledge and skills to: ...use computer hardware and appropriate software, and to understand statistics and research methods" is included in the Research section.
DNP <i>Essentials</i> (AACN, 2006, p. 16)	The statement "The DNP program prepares the graduate to: 1. Analyze epidemiological, biostatistical, environmental, and other appropriate scientific data related to individual, aggregate, and population health" is included in the section titled "Essential VII."
PhD <i>Essentials</i> (AACN, 2010, p. 5)	The phrase "Advanced research design and statistical methods" is included in Table 1 in the section "Expected Outcomes and Curricular Elements in the Core Curricular Elements."

*Note.* BSN = baccalaureate nursing; MSN = Master of Science in Nursing; DNP = Doctor of Nursing Practice; PhD = Doctor of Philosophy.

Figura 1. Menciones a la formación estadística en enfermería, recogida en los documentos fundamentales de la American Association of Colleges of Nursing (AACN).

En la conferencia anual conjunta de la American Statistical Association (ASA) y otras organizaciones internacionales de 2012, se presentó un panel de expertos cerca de la enseñanza de la estadística en estudiantes de enfermería (Hayat, Eckardt, Higgins, Kim y Schmiede, 2013) con un acuerdo referente a la necesidad de establecer las competencias estadísticas en la formación enfermera. En opinión de estos expertos y a la vista de las pautas recomendadas por la AACN, los estudiantes de grado, deben adquirir la capacidad de leer y comprender las publicaciones y revistas de enfermería y de ciencias de la salud, para lo que necesitarán una enseñanza que les forme en la alfabetización y el razonamiento estadístico (Hayat et al. 2013).

#### **1.4. Desarrollo de la competencia estadística y de investigación en los documentos normativos de los estudios de grado de Enfermería en España.**

La necesidad del conocimiento estadístico como herramienta de investigación en Enfermería se puso ya de manifiesto en documentos de trabajo realizados previamente a la implantación de los nuevos planes de estudio

Tal es el caso del libro blanco de Enfermería, que fue elaborado por diversas comisiones y grupos de trabajo de distintas universidades españolas, y la colaboración de profesionales de instituciones sanitarias, sociosanitarias y organizaciones profesionales y científicas. Todo el proyecto fue coordinado por la Conferencia Nacional de Directores de Escuelas de Enfermería y refrendado por el Consejo General de Enfermería. El objetivo de este trabajo fue realizar un estudio previo que adaptara el Título de Grado en Enfermería al Espacio Europeo de Educación Superior.

En este documento, en el que se valoran las competencias genéricas siguiendo las propuestas del grupo Tuning, se puso de manifiesto que la competencia “Habilidades de investigación” era valorada como importante por el grupo de docentes encuestados, aunque los profesionales asistenciales la consideraron de menor importancia para el perfil profesional de la Enfermería. En este mismo trabajo, se valoraron y clasificaron también las competencias específicas del grado de Enfermería, algunas de ellas relacionadas con la investigación: “Capacidad de hacer valer los juicios clínicos para asegurar que se alcanzan los estándares de calidad y que la práctica esté basada en la evidencia” y “Conocimientos relevantes de y capacidad para aplicar principios de investigación e información”. Posteriormente el Real Decreto 55/2005 de 21 de enero por el que se establece la estructura de las enseñanzas universitarias y se regulan los estudios universitarios oficiales de grado, dio lugar a unas primeras directrices para el título de grado de enfermería. Estas directrices, recogidas en la llamada ficha técnica y remitidas por el Consejo de Coordinación Universitaria, ya no hacían mención expresa a la metodología de investigación como materia, pese a haber estado presente en borradores previos.

En la Orden CIN/2134/2008 tampoco aparece explícitamente la investigación como una competencia que los estudiantes deben adquirir, aunque sí hay un especial interés en cuestiones relativas a búsqueda, análisis y aplicación práctica de los resultados de la investigación: “Basar las intervenciones de la Enfermería en la evidencia científica y en los medios disponibles”.

Por otro lado la Directiva 2005/36/CE del Parlamento Europeo y del Consejo, de 7 de septiembre de 2005, relativa al reconocimiento de cualificaciones profesionales, establece un régimen de reconocimiento de las cualificaciones profesionales, con objeto de favorecer la flexibilidad de los mercados laborales, conseguir una mayor liberalización de la prestación de los servicios, fomentar una mayor automaticidad en el reconocimiento de las cualificaciones y simplificar los procedimientos administrativos.

Esta Directiva establece, sobre la base de la coordinación de las condiciones mínimas de formación, el reconocimiento automático de algunos títulos entre los que se encuentra el de enfermero responsable de cuidados generales. En el apartado 5.2.1. del anexo V.2. de esta directiva se expone el programa de estudios necesarios para obtener el título de enfermero responsable de cuidados generales y se enumeran las materias que, como mínimo, ha de incluir esta formación. La formación en investigación, no aparece entre las materias correspondientes a la formación teórica y práctica de este título.

Si tenemos en cuenta las competencias relacionadas con la investigación que recogen los distintos documentos que configuran los ejes normativos de los planes de estudio de Enfermería (Figura 2), aquellas relacionadas con la investigación nos muestra un Graduado en Enfermería con capacidad para buscar el conocimiento científico más actual, valorar su relevancia y su validez y aplicarlo a la prestación de cuidados y a las intervenciones profesionales. Las competencias que permitan al enfermero colaborar en proyectos de investigación o liderarlos se desarrollarán en niveles académicos superiores. Por tanto, en el grado de Enfermería encontramos un perfil básicamente “consumidor”, por lo que las competencias estadísticas han de estar encaminadas a conocer, comprender, diferenciar e interpretar los análisis estadísticos fundamentales, desde un conocimiento teórico, sin profundizar excesivamente en los aspectos matemáticos. Estas competencias estadísticas se complementan con las competencias básicas de investigación, conocimiento de bases de datos bibliográficas, recogida y tratamiento de datos, elaboración y exposición de informes y proyectos científicos, etc. Competencias que son evaluadas en el trabajo de fin de Grado.



<b>COMPETENCIAS DE INVESTIGACIÓN EN LA TITULACIÓN DE GRADO DE ENFERMERÍA</b>
<b>Libro blanco</b> Título de grado de Enfermería (Agencia Nacional de Evaluación de la Calidad y Acreditación)
Capacidad de hacer valer los juicios clínicos para asegurar que se alcanzan los estándares de calidad y que <b>la práctica esté basada en la evidencia.</b> Conocimiento relevante y capacidad para <b>aplicar principios de investigación e información.</b>
<b>R.D. 1393/2007</b> Por el que se establece la ordenación de las enseñanzas universitarias oficiales.
Poseer y comprender conocimientos <b>procedentes de la vanguardia</b> de su campo de estudio. <b>Capacidad de reunir e interpretar datos relevantes</b> para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ético.
<b>Orden CIN/2134/2008</b> Por la que se establecen los requisitos para la verificación de los títulos universitarios oficiales que habiliten para el ejercicio de la profesión de enfermero.
Ser capaz, en el ámbito de la enfermería, de prestar una atención técnica y profesional adecuada a las necesidades de salud de las personas que atiende <b>de acuerdo con el estado de desarrollo de los conocimientos científicos de cada momento</b> y con los niveles de calidad. Basar las intervenciones de la Enfermería en la <b>evidencia científica</b> y en los medios disponibles.

Figura 2. **Competencias de investigación que aparecen en los documentos normativos de los planes de estudio de Enfermería.**

Por otro lado el Real Decreto 1027/2011, de 15 de julio, por el que se establece el Marco Español de Cualificaciones para la Educación Superior (MECES). Lo establecido en este documento permite la nivelación de todas las cualificaciones de la educación superior, para, de esta forma, facilitar la movilidad de las personas en el espacio europeo tanto en el ámbito académico como laboral.

El Real Decreto recoge una descripción de los distintos niveles académicos. Entre los resultados de aprendizaje que definen las características de las cualificaciones del nivel de Grado, aparecen en el artículo 6 los siguientes relacionados con la investigación:

“Tener la capacidad de recopilar e interpretar datos e informaciones sobre las que fundamentar sus conclusiones incluyendo, cuando sea preciso y pertinente, la reflexión sobre asuntos de índole social, científica o ética en el ámbito de su campo de estudio”.

“Ser capaces de desenvolverse en situaciones complejas o que requieran el desarrollo de nuevas soluciones tanto en el ámbito académico como laboral o profesional dentro de su campo de estudio”.

### 1.5. Competencia estadística y de investigación en las especialidades profesionales.

Además de los tres niveles académicos en los que se estructuran los estudios de Enfermería, el Real decreto 450/2005 de 22 de abril sobre especialidades de Enfermería, recoge el desarrollo de 7 especialidades: Enfermería familiar y comunitaria, Enfermería gerontológica, Enfermería pediátrica, Enfermería en salud mental, Enfermería del trabajo y Enfermería obstétrico-ginecológica. La obtención del título de Enfermero Especialista supone una cualificación puramente profesional sin reconocimiento académico. Dicha formación se realiza, tras superar una prueba selectiva de acceso que tienen carácter nacional, por el sistema de residencia en unidades docentes acreditadas para la formación especializada por el Ministerio de Sanidad, Servicios Sociales e Igualdad.

En los programas formativos de estas especialidades aparecen contenidos de investigación, análisis estadísticos, metodología de análisis crítico para basar la práctica en la mejor evidencia científica y la realización de un proyecto de investigación. Se recoge en la Figura 3 la presencia de estos contenidos en prácticamente todas las especialidades.

Especialidad	Contenidos de investigación	Estadística y análisis de datos cualitativos	Prácticas basadas en la evidencia	Proyecto de Investigación
Enfermería Familiar y Comunitaria	X	X	X	X
Enfermería Gerontológica	X		X	X
Enfermería Geriátrica	X		X	X
Enfermería en Salud Mental	X	X	X	X
Enfermería del Trabajo	X	X	X	X
Enfermería Obstétrico-Ginecológica (Matrona).	X	X	X	X

Figura 3. Formación investigadora en los programas de las especialidades de Enfermería.

Por otro lado, se ha establecido la carrera profesional implantada en casi todas las comunidades autónomas, como un sistema de reconocimiento del desarrollo profesional en cuanto a conocimientos y experiencia en las tareas asistenciales e investigadoras. Mediante un sistema de evaluación y acreditando los méritos exigibles, entre los que se encuentra el desarrollo de la actividad de investigación, se establecen unas retribuciones económicas complementarias acordes a cada uno de los niveles adquiridos.

### 1.6. Estudio comparativo del abordaje de las competencias estadísticas y de investigación en diferentes Universidades españolas.

Se han analizado los contenidos de estadística y de investigación recogidos en los actuales planes de estudio del grado de Enfermería de las 52 universidades españolas en las que se imparte el grado de Enfermería, así como nuestro propio plan de estudios de la Escuela de Enfermería San Juan de Dios de la Universidad Pontificia Comillas de Madrid.

Todos los planes de Estudio del grado de Enfermería en las universidades españolas tienen una asignatura con contenidos de estadística, a excepción de la Universidad de la Rioja y la Universidad de Girona.

El tratamiento que tienen los contenidos de estadística en las distintas escuelas analizadas no es homogéneo. En algunos casos aparece la Estadística como asignatura independiente pero también hemos encontrado estos contenidos unidos a las Tecnologías de la Información y Comunicación Sanitaria, y a Metodología de Investigación. Se recoge en la Tabla 2 las distintas asignaturas de estadística y de investigación encontradas en los planes de estudio de Enfermería en nuestro país.

**Tabla 2. Estadística e Investigación presentes en los planes de estudio del grado de Enfermería de las universidades españolas.**

UNIVERSIDAD	ASIGNATURAS DE ESTADÍSTICA E INVESTIGACIÓN	Nº DE CRÉDITOS	CURSO
Universidad de Almería	Análisis de datos estadísticos y TICS en salud. (Enfoque epidemiológico)	6	1º
	Enfermería basada en la evidencia	6	3º
Universidad de Cádiz	bioestadística, metodología de la investigación y TICS en cuidados de salud	9	1º
Universidad de Córdoba	Epidemiología, bioestadística y TICS aplicados a problemas de salud	6	1º

UNIVERSIDAD	ASIGNATURAS DE ESTADÍSTICA E INVESTIGACIÓN	Nº DE CRÉDITOS	CURSO
	Introducción a la metodología de la investigación en ciencias de la salud	3	Optativa 3º
	Investigación cualitativa	3	Optativa 3º
Universidad de Granada	TICS en cuidados de salud y metodología de la investigación	6	1º
	Estadística	6	2º
Universidad de Huelva	Bioestadística y tecnologías de la información y comunicación en cuidados	6	2º
	Introducción al conocimiento científico y evidencia en cuidados	3	2º
Universidad de Jaén	Sistemas de información y análisis de datos en los cuidados de salud	6	1º
Universidad de Málaga	Análisis de datos estadísticos, TICS en cuidados de salud	6	1º
Universidad de Sevilla	Estadística y tecnología de la información y comunicación	6	1º
Universidad de Zaragoza	Estadística aplicada a ciencias de la salud	6	1º
	Metodología de investigación	6	1º
Universidad de San Jorge. Zaragoza	Investigación básica y fuentes documentales	6	1º
	Bioestadística	6	2º
Universidad de La Laguna	Investigación y enfermería basada en la evidencia. (con contenidos de Estadística)	9	1º
Universidad de las Palmas	Estadística	6	1º
Universidad de Cantabria	Bioestadística	6	1º
Universidad de Castilla la Mancha	Estadística en ciencias de la salud	6	1º
	Metodología de investigación	6	Optativa 3º
Universidad de Burgos	Estadística (enfoque epidemiológico)	6	1º
	Bases para la interpretación del conocimiento científico	6	2º
Universidad Católica de Ávila	Bioestadística, documentación y TICS	6	1º

UNIVERSIDAD	ASIGNATURAS DE ESTADÍSTICA E INVESTIGACIÓN	Nº DE CRÉDITOS	CURSO
Universidad de León	Bioestadística y metodología científica	6	1º
Universidad Pontificia de Salamanca	Estadística	6	1º
Universidad de Salamanca	Estadística e investigación operativa	6	1º
Universidad de Valladolid *	Estadística, sistemas de información y nuevas tecnologías	6	1º
	Metodología de investigación	6	2º
Universidad de Girona	Sistemas de información y lenguaje científico (no tiene estadística)	6	1º
Universidad Autónoma de Barcelona	Metodología científica y estadística	6	1º
	Lectura crítica y metodología de investigación	3	Optativa 4º
Universidad de Barcelona *	Iniciación a la metodología científica documentación y TICS	6	1º
Universidad de Lleida	Bioestadística	6	1º
Universidad Ramón Llull	Estadística	6	2º
Universidad Rovira i Virgili	Bioestadística	6	1º
Universidad de Vic	Estadística y sistemas de información en salud	6	1º
	Elaboración de proyectos	3	2º
Universidad Alfonso X el Sabio	Estadística	6	1º
	Metodología de investigación	4	Optativa 4º
Universidad de Alcalá de Henares	Estadística y metodología de investigación	6	1º
Universidad Nebrija	Estadística y tecnología de la información aplicada	6	1º
Universidad Autónoma de Madrid	Bioestadística y TICS	6	1º
	Metodología de la investigación	6	3º
Universidad Camilo José Cela	Bioestadística, metodología de la investigación y TIC'S en cuidados de salud	9	1º

UNIVERSIDAD	ASIGNATURAS DE ESTADÍSTICA E INVESTIGACIÓN	Nº DE CRÉDITOS	CURSO
Universidad Complutense	Bases científicas de la Enfermería	6	Optativa 3º
	Salud pública (con contenidos de epidemiología y Estadística)		
Universidad Rey Juan Carlos	Epidemiología (contenidos de Estadística e investigación )	8.5	1º
Universidad Europea*	Estadística e investigación cuantitativa	6	1º
	Investigación cualitativa	6	2º
Universidad Francisco de Vitoria	Métodos estadísticos en Enfermería I	3	2º
	Métodos estadísticos en Enfermería II	3	3º
Universidad San Pablo Ceu	Estadística	6	1º
Universidad Pontificia Comillas de Madrid	Estadística y herramientas de investigación en ciencias de la salud	9	1º
Universidad de Navarra*	Estadística	6	2º
	Bases de la metodología científica en ciencias de la salud	3	2º
Universidad Pública de Navarra	Estadística aplicada a las ciencias de la salud	6	1º
	Elaboración de proyectos en el campo sanitario	3	Optativa 3º
	Fuentes de datos documentales en el campo sanitario	3	Optativa 1º
Universidad de Alicante	Bioestadística	6	1º
	Sistemas informacionales, investigación e innovación en Enfermería	6	1º
	Investigación cualitativa y desarrollo de la Enfermería	6	Optativa 3º
Universidad Jaume I	Estadística	6	1º
Universidad de Valencia	Bioestadística y TICS aplicadas a las ciencias de la salud	6	1º
	Introducción a la investigación en la disciplina enfermera	6	4º
Universidad Católica de Valencia	Bioestadística y metodología de investigación	6	1º

UNIVERSIDAD	ASIGNATURAS DE ESTADÍSTICA E INVESTIGACIÓN	Nº DE CRÉDITOS	CURSO
Universidad de Extremadura	Metodología de investigación aplicada a la Enfermería	6	1º
	Métodos estadísticos aplicado en Enfermería	6	Optativa 4º
Universidad de la Coruña	Bases para la interpretación del conocimiento científico	6	1º
Universidad Santiago de Compostela	Estadística y metodología de investigación	6	1º
Universidad de las Islas Baleares	Fundamentos de Estadística aplicada a las ciencias de la salud	6	1º
	Introducción a la investigación cualitativa en salud	3	Optativa 4º
Universidad de la Rioja	No tiene estadística ni investigación		
Universidad del País Vasco	Sistemas de información y análisis de datos	6	1º
Universidad de Oviedo	Epidemiología y bioestadística	6	1º
	Metodología de la Investigación	6	4º
Universidad Católica de San Antonio	Instrumentos para la investigación (con contenidos de Estadística)	6	1º
Universidad de Murcia	Bioestadística	6	1º

Nota. \* No se han encontrado las guías de las asignaturas.

En 23 de los planes de estudio analizados, aparecen los contenidos estadísticos en una asignatura denominada directamente estadística, o en algunos casos bioestadística. Resulta curioso el hecho de que la estadística aparezca junto a la epidemiología o con un enfoque epidemiológico en universidades en las que el grado de Enfermería se encuentra en la facultad de medicina o junto al grado en medicina en un departamento de ciencias de salud. En este caso la estadística es tratada con una perspectiva epidemiológica como un conjunto de métodos para evaluar los datos obtenidos en poblaciones humanas.

En 13 de los planes de estudio analizados los contenidos de estadística aparecen junto a los sistemas de información sanitarios y a las tecnologías de la información y la comunicación. La Estadística y la metodología de investigación las hemos encontrado unidas en una sola asignatura en los planes de estudio de

Enfermería de 5 universidades españolas. También se encuentran unidos en una asignatura los contenidos de estadística, metodología de investigación y tecnologías de la información, de esta manera aparecen en 3 de los planes de estudio de Enfermería revisados.

En la casi totalidad de los casos, la Estadística se incluye en una asignatura de 6 créditos que se imparte en primer curso.

En general, la estructura de contenidos de estadística encontrada en los diferentes planes de estudio analizados es muy coincidente en lo que llamaremos curso introductorio o curso básico de Estadística con los siguientes contenidos:

- **Estadística descriptiva.** Tipos de variables estadísticas, así como su forma de representarlas a través de gráficos y tablas y la interpretación de las medidas más usuales de centralización y dispersión.
- **Distribuciones de probabilidad. Distribución normal.** Conceptos básicos de probabilidad y variable aleatoria. Características y aplicación de la distribución normal en el cálculo de probabilidades y en la estadística inferencial.
- **Inferencia estadística.** Fundamentos básicos sobre muestreo, estimadores, intervalos de confianza y contrastes de hipótesis, las principales propiedades y usos de los mismos.
  - **Estimación de parámetros.**
  - **Pruebas de significación. Contraste de hipótesis**
    - Contraste de medias.
    - Cuantificación de las diferencias. Tamaño de efecto.
    - Pruebas  $\chi^2$
  - **Diseño experimental. Análisis de la varianza.**
  - **Regresión lineal simple y correlación.**
- **Introducción a la Estadística no paramétrica.** Conocimiento teórico de las pruebas no paramétricas básicas.

Respecto a la formación en metodología de investigación, hay que decir que, de una u otra forma se encuentran en la mayoría de los planes de estudio de Enfermería. En algunos casos con un enfoque de investigación de enfermería basada en la evidencia, en otros con contenidos de metodología de investigación más clásicos. A veces se aborda de forma práctica para la realización del trabajo



fin de grado e incluso están presentes asignaturas más específicas como la investigación cualitativa.

En la mayoría de las universidades españolas el plan de estudios del grado de Enfermería contempla la formación en investigación. Si bien es cierto que aparecen asignaturas muy diferentes en contenidos, con nivel básico o con un nivel que podríamos considerar muy avanzado, en algún caso como asignatura obligatoria y otras veces como optativa, lo cierto es que las competencias de investigación están presentes en los graduados de Enfermería.

La Guía Docente de la Escuela de Enfermería y Fisioterapia de la Universidad Pontificia Comillas de Madrid, define de la siguiente forma lo que la Asignatura de “Estadística y Herramientas para la Investigación en Ciencias de la Salud” aporta al perfil profesional de la titulación:

*Esta asignatura pretende que el alumno comprenda básicamente el proceso de investigación, sea capaz de identificar problemas y de aplicar la mejor evidencia a la práctica diaria, así como colaborar en equipos de investigación. Así mismo, aportará al alumno los conocimientos para que valore la necesidad del conocimiento estadístico como herramienta de investigación en enfermería, conozca el fundamento de las distintas pruebas estadísticas básicas, su aplicación en ciencias de la salud y sea capaz de enfrentarse a un estudio estadístico sencillo desde el planteamiento del problema hasta la exposición de resultados. Proporcionará a los estudiantes los conocimientos que le aporten competencias para utilizar herramientas informáticas durante su periodo formativo y en su posterior desarrollo profesional, tanto en los sistemas de información como de otras tecnologías del ámbito sanitario y recursos para el tratamiento estadístico de los datos y para la búsqueda de documentación científica.*

Tal como se recoge en esta guía, la Estadística es un recurso necesario para basar la práctica de la Enfermería en la evidencia científica. Los profesionales de Enfermería se enfrentan en la actualidad a una gran variedad de técnicas estadísticas y el grado de conocimientos de dichas técnicas determinará el nivel con que dichos lectores podrán valorar críticamente los resultados de investigación. En resumen, no podrán aplicar las mejores evidencias sin tener, entre otras competencias, acceso estadístico a la investigación.

### **1.7. La accesibilidad estadística como medición de la eficacia de la formación estadística**

Como se ha comentado anteriormente, en estudios presentados por expertos (Hayat et al. 2013 y Hayat 2014) y desde el análisis de las competencias que subyacen en los documentos normativos del grado de Enfermería en nuestros planes de estudio, la formación estadística en el grado no debe centrarse en la habilidad para aplicar métodos estadísticos o en el desarrollo de análisis de datos. Lo que necesitan los futuros graduados es una base sólida para la lectura de las publicaciones científicas de enfermería y de ciencias de la salud.

Cuando comenzamos a abordar el tema principal de esta tesis, la enseñanza de la estadística en el ámbito de enfermería, nos interesamos en analizar qué conocimientos de Estadística necesitaría un enfermero para poder tener acceso a la publicación de resultados de investigación, requisito fundamental para valorar y aplicar la mejor evidencia científica a la práctica profesional. Nos interesó también conocer si la formación estadística que reciben los alumnos de Enfermería, les hace competentes para leer, enjuiciar y aplicar a la práctica los resultados de la investigación enfermera, y qué resultados de aprendizaje debemos tener en cuenta en los distintos niveles académicos.

Realizamos para ello un análisis de artículos originales de varias revistas de Enfermería de ámbito nacional e internacional, registrando y clasificando las técnicas estadísticas que contenían como reflejo, por un lado, del nivel de la investigación enfermera y por otro, de las necesidades de conocimiento estadístico de los alumnos de grado para poder entenderlos y aplicar la mejor evidencia a la práctica, tal como se especifica en la orden CIN 2134/2008 como una de las competencias que los estudiantes de grado de Enfermería deben adquirir.

El análisis del nivel que mostraba la investigación de enfermería nos dio también pistas acerca de las competencias estadísticas, como herramienta de investigación, que han de desarrollarse a lo largo de los tres ciclos en los que se estructuran las enseñanzas de Enfermería.

En la revisión bibliográfica se encontraron estudios bibliométricos que tratan de analizar el desarrollo de la investigación enfermera, para detectar qué aspectos presentan carencias que deben ser corregidas (Torrá 1996, Gálvez et al. 2004, Serrano et al. 2005). En otros estudios de este tipo consultados, se

pone de manifiesto que la producción científica de la Enfermería en nuestro país, con respecto a las publicaciones de Enfermería con factor de impacto internacional, tiene marcadas diferencias con respecto a disciplinas afines como son la Medicina, de carácter más experimental y la Psicología más social (Icart et al. 1999).

Los trabajos que se han interesado por los análisis estadísticos que contienen los artículos, en algunos casos se remiten a realizar un listado de las pruebas utilizadas, agrupándolas de forma simple y anotando la frecuencia con que aparecen en los artículos (Icart et al. 1999, Lizarbe 2007). Otros utilizan categorías de análisis basadas en las establecidas por Emerson y Colditz en su trabajo sobre la accesibilidad estadística de la revista *The New England Journal of Medicine* (Emerson y Colditz 1983). Esta clasificación ha sido frecuentemente utilizada al igual que el índice de accesibilidad definido también por estos autores.

La accesibilidad estadística es analizada en investigaciones médicas españolas (Mora et al. 1995a y b, 1996; Carré et al. 1996; García 2000, González de Dios 2001, Fernández et al. 2003, Ortiz 2003, Granda et al. 2002), en revistas médicas internacionales (Reznick et al 1987, Aylward y Verhulst 1991, Horton y Switzer 2005; Hellems 2007), pero no se encontró ningún trabajo que analizara la accesibilidad estadística en revistas de Enfermería.

El término accesibilidad estadística fue introducido por Emerson y Colditz con el siguiente significado: “El acceso estadístico a un artículo es la cuantificación del repertorio estadístico que necesita el lector para comprender todos los análisis estadísticos empleados en dicho artículo” (Emerson y Colditz 1983). La accesibilidad estadística de una revista se puede expresar como el número de artículos de la revista, expresado en porcentaje acumulado respecto al total de artículos, en los que se utiliza un análisis estadístico de un determinado nivel de complejidad, según un listado de categorías establecidas. Este índice evalúa el número de artículos originales estadísticamente accesibles para un lector, con unos conocimientos determinados de Estadística.

Para asignar la accesibilidad estadística de cada artículo, se necesita en primer lugar contar con un listado de categorías de análisis. La primera y más utilizada fue la propuesta por Emerson y Colditz (1983) y sobre la que algunos autores (Mora et al. 1996, Fernández et al. 2003) señalaban la discutible ordenación jerárquica de los diferentes test estadísticos que proponen, argumentando que no coincide con el grado de dificultad para el aprendizaje o con su presentación en los libros de texto. Pese a ello, se ha constituido en un

“estándar” de evaluación y ha sido utilizada en múltiples estudios, permitiendo comparar los resultados obtenidos con los comunicados por otros autores.

En nuestro trabajo utilizamos una clasificación, (Figura 4), que varía en muchos aspectos de la propuesta por Emerson y Colditz y que elaboramos de acuerdo a la progresión lógica para el aprendizaje de la metodología estadística, intentando analizar la sucesión de complejidad necesaria para ir entendiendo el fundamento y aplicación de cada uno de los análisis. Se aplicó un criterio pedagógico como si se tratara de ordenar los contenidos de la asignatura de Estadística a la hora de exponerlos a los alumnos. Cada una de las 15 categorías establecidas, recogen técnicas estadísticas generales que comparten características comunes de análisis. Las pruebas más específicas incluidas en la clasificación, son sólo aquellas que fueron apareciendo en los artículos analizados.

Categoría	Descripción
1. Sólo estadística descriptiva	Se calculan y presentan los estadísticos más adecuados para describir las variables. Frecuencias, porcentajes, mediana y cuartiles para variables cualitativas ordinales. Para variables cuantitativas los parámetros centrales (media, mediana y moda) junto con los de dispersión (rango, varianza, desviación típica y coeficiente de variación) y de posición (percentiles). Representaciones gráficas de las variables cualitativas utilizando diagrama de barras y sectores y las variables cuantitativas con histograma y diagrama de cajas.
2. Correlación y regresión simple	Análisis de la relación y asociación lineal entre dos variables cuantitativas, permitiendo predecir un valor promedio de la variable dependiente. Gráficos de dispersión. Pruebas paramétricas: correlación lineal de Pearson, regresión mediante mínimos cuadrados. Test-retest para el análisis de fiabilidad en psicometría
3. Pruebas chi cuadrado, prueba exacta de Fisher y prueba de McNemar.	Pruebas chi cuadrado para valorar asociaciones o diferencias entre dos variables cualitativas en las que sólo una de ellas sea ordinal. Pruebas chi cuadrado para comparar proporciones entre dos o más grupos independientes, pareados y para valorar tendencias.

<p>Calculo de odss ratio (OR) y riesgo relativo (RR). Prueba de Mantel-Haenszel</p>	<p>Corrección de Yates cuando las frecuencias esperadas es menor de 5. Prueba exacta de Fisher cuando alguna frecuencia esperada es inferior a 1 o el tamaño muestral es menor a 100. Prueba de McNemar para comparar sólo una proporción entre dos grupos pareados</p>
<p>4. Correlación no paramétrica</p>	<p>Pruebas no paramétricas: correlación lineal de Spearman o Tau-b de Kendall, regresión no paramétrica, pruebas de tendencia. Comparación de métodos cualitativos de medida (entre observadores) Medidas de concordancia, índice de Kappa (Kappa-Cohen)</p>
<p>5. Prueba t de Student y pruebas Z</p>	<p>Para una o dos muestras, tanto de muestras independientes como de muestras relacionadas (datos pareados).</p>
<p>6. Pruebas no paramétricas</p>	<p>Prueba U de Mann-Whitney para comparar dos medias independientes extraídas de la misma población (equivalente no paramétrica de la t Student no pareada)</p>
<p>7. Valoración de la magnitud de las diferencias.</p>	<p>Prueba T de Wilcoxon para comparar dos medias extraídas de la misma muestra en dos situaciones distintas. (equivalente no paramétrica de la t Student pareada) Prueba de los signos</p>
<p>8. Análisis de la varianza y de la covarianza</p>	<p>Se acompaña de los correspondientes intervalos de confianza para la valoración de la magnitud de las diferencias. Se calcula el Tamaño del efecto</p>
<p>9. Comparaciones múltiples</p>	<p>Comparación de dos o más medias independientes. Análisis de covarianza Pruebas F Prueba de Kruskal-Wallis</p>
<p>10. Estimación por intervalo</p>	<p>Comparaciones múltiples con igualdad de varianza: Bonferroni, Scheffé, Duncan, Tukey y Newman-Keuls En caso de desigualdad de varianzas: T2 de Tamhane, T3 de Dunnett, Game-Howell, etc.)</p>
<p>Utilización de intervalos de confianza para las</p>	<p>Utilización de intervalos de confianza para las</p>

<p>Análisis de Supervivencia Univariante: método de Kaplan-Meier, y sus intervalos de confianza por el método de Greenwood.</p> <p>Análisis de Supervivencia Bivariante: prueba del Log-Rank también conocida como Test de Savage, Test de Cox-Mantel o Test de Mantel-Haenszel</p> <p><b>Estadística multivariante</b></p> <p>11. Análisis discriminante: cuando la variable dependiente tiene dos o más categorías</p> <p>Regresión múltiple: cuando la variable dependiente es cuantitativa.</p> <p>12. Regresión logística: cuando la variable dependiente es dicotómica</p> <p>Regresión de Poisson: cuando el fenómeno de estudio es raro o poco probable</p> <p>Regresión de Cox: cuando la variable dependiente es la variable tiempo.</p> <p>13. Coeficientes de regresión se convierten en medidas de significado más útil, como tasas, medias, riesgos relativos y odds ratios.</p> <p>14. Transformación de variables</p> <p>Modelo de Rasch</p>	<p>estimaciones de parámetros poblacionales (proporciones, medias, varianzas)</p> <p>Relación matemática entre una o más variables dependientes y varias covariables independientes o predictoras, permitiendo predecir el valor de una en función de otra, ajustadas todas las demás.</p> <p>Prueba de Hosmer-Lemeshow Modelos lineales generalizados y ecuaciones de estimación generalizada Regresión ZIP</p> <p>Análisis de Supervivencia Multivariante: Modelo de Regresión de Cox ( Modelo de Riesgos Proporcionales)</p> <p>Transformación de variables con distribución no normal (asimétrica) para su análisis paramétrico. Transformación logarítmica. Transformación logit en regresión, cuando intervienen proporciones.</p>
--	--

<p>15. Otros análisis</p> <ul style="list-style-type: none"> <li>Análisis cluster.</li> <li>Modelos causales.</li> <li>Escalamiento multidimensional</li> <li>Comparación de métodos cuantitativos de medida. Medidas de concordancia. Análisis de relación estructural</li> </ul>	
--	--

Figura 4. Clasificación de las categorías de análisis.

De acuerdo a este listado de categorías se calculó el índice de accesibilidad estadística de 60 artículos de revistas nacionales y otros 60 de revistas internacionales, estudiando cuál sería el nivel de accesibilidad de un estudiante de Enfermería, tomando como referencia los contenidos de un curso básico o introductorio de estadística.

Las revistas objeto de este análisis pertenecientes al ámbito español fueron: *Index de Enfermería*, *Enfermería Clínica* y *Metas de Enfermería*. Se eligieron por ser las revistas españolas con más altos indicadores de repercusión inmediata e impacto entre las principales revistas de Enfermería del Espacio Científico Iberoamericano (Gálvez et al. 2004, Gálvez et al. 2007). Las tres fueron seleccionadas y clasificadas, según criterios internacionales de calidad editorial previamente probados y convenidos por el Sistema Latindex (<http://www.latindex.unam.mx/>). En esta elección también se valoró su presencia en las bases de datos nacionales e internacionales (Sobrino et al. 2005) y su nivel de difusión entre los profesionales de Enfermería.

Las revistas internacionales elegidas fueron *Nursing Research*, *Research Nursing & Health* y *Journal of Advanced Nursing*.

Se valoró el índice de impacto recogido en Journal Citation Reports ([http://www.thomsonreuters.com/business\\_units/scientific/](http://www.thomsonreuters.com/business_units/scientific/)), y en el caso de *Journal of Advanced Nursing*, además fue elegida por ser un buen exponente de las publicaciones europeas de esta disciplina. Las tres recogen las dos líneas principales que guían el desarrollo de la Enfermería internacional, es decir tanto la visión más científico-técnica del cuidado, como el aspecto más humanista. La revista *Journal of Advanced Nursing* incluye un importante número de trabajos de carácter cualitativo bajo una perspectiva de análisis sociocultural, mientras que *Nursing Research* está publicando fundamentalmente estudios que se vinculan directamente con las ciencias básicas.

Se realizó un registro en el que a cada artículo se le asignó el número correspondiente a la categoría de análisis de cada una de las técnicas estadísticas empleadas en la investigación, considerando que en cada artículo podía identificarse más de una categoría, exceptuando la categoría número 1 que se utilizó exclusivamente para los artículos que sólo incluyen análisis descriptivos. Para calcular la accesibilidad estadística se registró sólo la categoría de análisis de mayor complejidad recogida en cada artículo, utilizando las frecuencias relativas acumuladas (%) de estas categorías más altas, ordenadas de menor a mayor complejidad.

De todos los resultados de este estudio, lo que nos interesa resaltar aquí es la información referente a la accesibilidad estadística de un profesional de Enfermería que haya recibido formación básica de Estadística.

Observamos en la Figura 5 que un lector que sólo tenga conocimientos de estadística descriptiva podría entender el 20% de los artículos de las revistas nacionales pero sólo tendría acceso al 1,7% de los artículos de revistas internacionales, es decir prácticamente ninguno.

La accesibilidad estadística en publicaciones internacionales aumenta hasta un 56,7% si consideramos que dicho lector conoce las técnicas multivariantes.

Un graduado en Enfermería, que ha recibido una formación básica en estadística tal como muestran los diferentes planes de estudio analizados en este mismo capítulo, tendría acceso al 21,7% de los resultados de investigación publicados a nivel internacional, que tomamos como referencia frente a las publicaciones nacionales, por ser artículos con un nivel de complejidad estadística mayor y por tanto una menor accesibilidad.



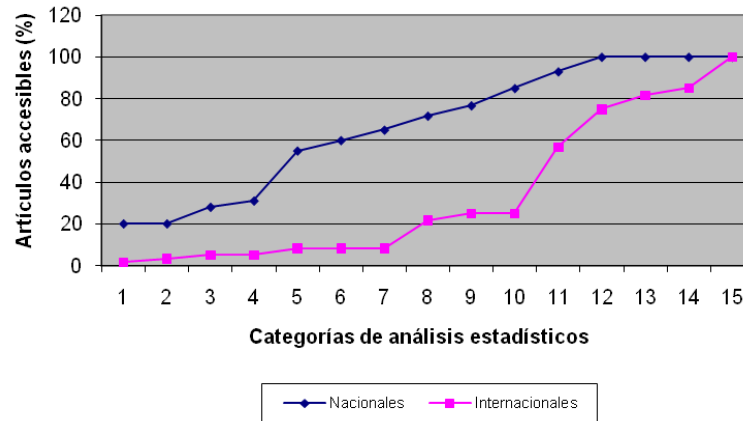


Figura 5. Accesibilidad estadística.

Las revistas internacionales están, por tanto, más lejos de la accesibilidad de un lector estándar. Para tener acceso al 21,7% de los artículos, no sólo se han de tener conocimientos de Estadística básica, sino haber profundizado en ellos, dada la diversidad de pruebas y aplicaciones específicas que nos encontramos incluso dentro de las primeras categorías. Igual, pero en mayor medida, ocurre con los artículos que contienen estadística más compleja, en los que es necesario algo más que un conocimiento somero de estos análisis para entender correctamente la metodología y el propósito de las investigaciones.

Todo lo anteriormente expuesto, es otro de los motivos que justifica nuestro interés por la enseñanza de la estadística en el grado de Enfermería. Estamos convencidos de que las competencias adquiridas a través de su aprendizaje, suponen una valiosa herramienta para el desempeño de la labor de estos profesionales en el escenario de la investigación enfermera y en la valoración del nivel de evidencia de dichas investigaciones.



## CAPÍTULO 2. ENSEÑANZA Y APRENDIZAJE DE LA ESTADÍSTICA

Para contextualizar el problema de investigación, comenzamos este capítulo delimitando el marco teórico y metodológico que orienta el proceso de enseñanza-aprendizaje de la Estadística. Con este fin realizaremos una revisión de las teorías constructivistas comprobando cómo este paradigma ha permeado en la educación. Para finalizar, se expondrán diversos escenarios utilizados para el aprendizaje de la Estadística, así como las directrices más importantes para implementar el desarrollo del razonamiento estadístico.

### 2.1. El constructivismo

Tal y como dijo Michael R. Matthews (1994) el constructivismo es como una vivienda que contiene múltiples estancias. Tanto es así que encontramos una amplia variedad de perspectivas y puntos de vista acerca de esta teoría. Es necesario tener en cuenta que existen teorías constructivistas en el campo de la epistemología, de la psicología, de la sociología y en la didáctica de las ciencias. Es por este motivo por lo que resulta imprescindible reconocer de qué constructivismo se está hablando para no mezclar ideas en marcos teóricos inapropiados. En este sentido, Coll (1996) aconseja distinguir entre constructivismo, teorías constructivistas del desarrollo y del aprendizaje, planteamientos constructivistas en educación y explicaciones constructivistas de los procesos de enseñanza y aprendizaje.

Las teorías constructivistas son, ante todo, teorías epistemológicas; es decir, son teorías que nos proveen de una explicación de cómo se produce el conocimiento, y de cuáles son las condiciones para que esta producción tenga lugar. Existen muchas corrientes epistemológicas que reclaman el apelativo de “constructivistas” y, en general, han tenido una fuerte influencia en la educación matemática en todo el mundo (Waldegg, 1998).

El constructivismo se gesta desde una perspectiva epistemológica con el fin de identificar en las personas los problemas de aprendizaje. El constructivismo plantea dicho aprendizaje situándose en el interior del sujeto actuando sobre su contexto y circunstancias al tiempo que los transforma.

Los principios teóricos que resumen los rasgos más destacados del constructivismo son (Araya, Alfaro, Andonegui, 2007):

- Principio de interacción del hombre con el medio.
- Principio de la experiencia previa como condicionadora del conocimiento a construir.

- Principio de elaboración de “sentido” en el mundo de la experiencia.
- Principio de organización activa.
- Principio de adaptación funcional entre el conocimiento y la realidad.

### 2.1.1. Evolución del constructivismo y tendencias principales

En la década de los sesenta y setenta los países anglosajones desarrollaron iniciativas curriculares con una base constructivista siguiendo los postulados de Piaget. El objetivo era mejorar la calidad de la enseñanza de los contenidos científicos de la población escolar y con el objetivo de implementar el desarrollo tecnológico de los años futuros (Aliberas, Gutiérrez, Izquierdo, 1989). A partir de la segunda mitad de los setenta, disminuye la influencia de los planteamientos constructivistas de Piaget debido a la aparición de otras alternativas empíricas (Driver y Easley, 1978; Novak, 1982), manteniéndose los planteamientos de Piaget en sectores minoritarios de investigadores pero permaneciendo hasta la actualidad con una significativa presencia.

Las referencias del constructivismo piagetiano se encuentran en la epistemología genética, que tiene como idea fundamental que el conocimiento es un fenómeno adaptativo del organismo humano al medio. Para Piaget<sup>1</sup>, el conocimiento es consecuencia de la interacción entre el sujeto y la realidad en la que se desenvuelve, realidad sobre la que actúa el individuo construyendo las propiedades de ésta, al tiempo que estructura su propia mente. Las estructuras generales del conocimiento científico son la clave de la teoría constructivista de Piaget (sujeto epistémico), sin embargo, adolece de no tener en cuenta ni el carácter multifacético ni contextual del conocimiento, lo que representaría una limitación del constructivismo piagetiano (Martí, 1996). No se debe obviar que los conceptos piagetianos de asimilación y acomodación son muy relevantes en la teoría constructivista, a pesar de que la contribución fundamental de Piaget haya estado ligada a la perspectiva del desarrollo cognitivo (Sebastiá, 1989). Tanto es así, que los factores que influyen en el cambio de las estructuras cognitivas según Piaget son: la maduración, la experiencia, el equilibrio y la transmisión social (Calzadilla, 2002).

Coincidiendo con la aparición del artículo de Driver y Easley (1978) surgen una gran cantidad de estudios empíricos desarrollados muchos de ellos sin una base teórica. Esta nueva manera de hacer constructivismo se empezó a percibir como una nueva tendencia, constructivismo social, gracias a la labor aglutinadora que desarrolla Solomon (1994). Esta teoría se crea, principalmente, a partir de los problemas que presenta el alumnado para comprender las ciencias y defiende que el conocimiento lejos de encontrarse de forma exclusiva en la

<sup>1</sup> Piaget, J. (1970). Genetic epistemology. New York. W.W. Norton and Company.

mente de las personas o en su entorno, reside en las formas de relación social e intercambio simbólico. Esta corriente pretende el máximo y multifacético desarrollo posible de las facultades e intereses de los estudiantes. Esta concepción constructivista tiene en los postulados de Bruner y Vygotski, las bases conceptuales en la que se sustenta. Los seguidores de esta corriente defienden que el conocimiento es una experiencia compartida en la que la interacción entre el sujeto y el medio procura nuevos caracteres y rasgos, estableciéndose una compleja y recíproca relación entre ambos.

En el paso de la década de los setenta a los ochenta, Novak (1988), partiendo principalmente de la propuesta en psicología educativa de Ausubel<sup>2</sup>, va a desarrollar lo que posteriormente se denominaría constructivismo humano. Esta corriente tendrá una acusada influencia en el desarrollo de la didáctica de las ciencias y deriva de la visión particular de Ausubel de la psicología cognitiva, matizada con concepciones epistemológicas y algunas incursiones en estudios de neurobiología cerebral.

A comienzos de los años noventa, la historia constructivista para la didáctica de las ciencias adquiere mayor complejidad con la aparición del constructivismo radical, asociado especialmente a su principal exponente von Glasersfeld (Matthews, 1994), que se basa, sobre todo, en los trabajos de Piaget y defiende la idea de que el conocimiento es activamente construido por el sujeto, por lo que no se pueden transmitir significados o ideas al alumno ya que es éste quien finalmente lo construye (Marín, Solano, Jiménez, 1999). Para el constructivismo radical, que posee una base más filosófica, el conocimiento se adquiere de manera subjetiva. Este es el motivo por el que defiende que es imposible la interpretación objetiva de la realidad.

### 2.1.2. Constructivismo y educación

En educación, se suele equiparar el constructivismo con la psicología genética de Jean Piaget, identificada tradicionalmente como la teoría constructivista por excelencia. Pero resulta imprescindible recordar que sus postulados, con una base epistemológica, tratan de responder a la pregunta formulada por el propio Piaget acerca de cómo aumentar el nivel de conocimiento en los sujetos pasando de un estado de menor a mayor conocimiento. Por este motivo, no podemos obviar la influencia piagetiana en la educación y en los objetivos planteados por esta, en la concepción del alumno como agente activo y autónomo, en una concepción no autoritaria del profesor y en las metodologías didácticas participativas.

Sólo mediante aprendizajes significativos el alumno va a poder desarrollar su crecimiento personal al mejorar su conocimiento del entorno. A su vez, la

---

<sup>2</sup> Ausubel, D. (1968). *Educational Psychology: A Cognitive View*. New York and Toronto: Holt, Rinehart and Winston.

finalidad de la labor del docente es hacer al alumno capaz de desarrollar aprendizajes significativos por sí solo en cualquier circunstancia de su proceso formativo: aprender a aprender (Coll, 1998). La concepción constructivista se organiza en torno a tres ideas fundamentales (Coll, 1990):

- La responsabilidad del aprendizaje recae sobre el alumno.
- La actividad mental constructiva del alumno se aplica a contenidos que poseen ya un grado considerable de elaboración.
- El docente debe engranar los procesos de construcción del conocimiento del alumno con el saber colectivo culturalmente originado en su entorno.

Desde la década de los noventa los enfoques constructivistas han otorgado al profesor un papel central: el profesor es quien debe generar un entorno didáctico adecuado que permita al alumno utilizar sus conocimientos y experiencias previas (Waldegg, 1998).

En el caso de la enseñanza de las matemáticas, Bruner (1960), por ejemplo, percibe las matemáticas no como un producto o resultado sino más bien como un proceso que los estudiantes deberían experimentar en la clase de matemáticas. Los alumnos en los que está pensando Bruner (Gadanidis, 1994), participan en situaciones matemáticas problemáticas, buscando y posteriormente descubriendo patrones y estructuras unificadoras, haciendo del aprendizaje de las matemáticas una “montaña rusa” de sucesivos equilibrios y desequilibrios que terminan en el logro o descubrimiento del estado cognitivo deseado (Shulman, 1970).

### 2.1.3. El aprendizaje significativo

Desde diferentes concepciones y enfoques psicopedagógicos se afirma que sólo los aprendizajes significativos tienen la capacidad de fomentar el desarrollo personal del alumno. Los estudios realizados en torno a los procesos de aprendizaje por descubrimiento que tuvieron lugar una vez comenzada la segunda mitad del siglo XX (aunque previamente habían existido ya algunas iniciativas al respecto) nos sitúan en los antecedentes modernos del aprendizaje significativo. Dichas investigaciones tenían como hipótesis la convicción de que los alumnos debían adquirir el conocimiento “mediante el uso de su propia mente” (Bruner, 1961). En esta misma idea convergían las iniciativas pedagógicas que se habían inspirado en las propuestas constructivistas de Piaget (1974), que él mismo sintetizaba como el principio fundamental de los métodos activos: comprender es inventar o reconstruir por reinención.

Lejos de ser una fórmula mágica, el aprendizaje significativo, conceptualmente, posee un extraordinario potencial para el análisis, la reflexión

y la intervención psicopedagógica y ofrece una valiosa perspectiva que puede resultar muy útil a la hora de estudiar todos los aspectos relacionados con los procesos tanto de enseñanza como de aprendizaje.

### **2.1.3.1. La construcción de significados.**

Uno de los objetivos principales del aprendizaje significativo es que, a la hora de desarrollar los procesos de enseñanza y aprendizaje, se estimule el proceso de construcción de significados. Sólo cuando es capaz de atribuirle un significado podemos decir que el alumno ha aprendido un contenido. Estos contenidos también se pueden aprender de manera diferente, que es lo que ocurre al aprender algo de memoria sin comprenderlo. En muchas ocasiones lo que ocurre es que el alumno tan sólo es capaz de atribuir significados parciales. Cuando ocurre esto, lo que está pasando es que lo que se ha aprendido tiene un significado diferente para quien lo ha enseñado que para quien lo ha aprendido. Esta situación nos lleva, en nuestro caso, a reflexionar acerca de la significatividad del aprendizaje de la Estadística ante la dificultad de transmitir algunos conceptos por parte del profesor y de la comprensión de éstos por los alumnos. Esta realidad nos debe situar en el camino que nos conduzca a desarrollar el aprendizaje de forma gradual hasta estar completamente seguros de que lo enseñado y lo aprendido mantienen un equilibrio razonable.

Pero ¿qué quiere decir exactamente construir significados? Para Ausubel y sus colaboradores (Ausubel, Novak, Hanesian, 1983) construir significados consiste en adquirir la capacidad de fijar relaciones con fundamento entre aquello que pretendemos aprender y lo que pertenece a nuestro conocimiento adquirido anteriormente.

Los estudios realizados por David Ausubel (1918-2008), a partir de la segunda mitad del siglo XX, en relación con el desarrollo del intelecto en la escuela marcaron los discursos de la psicología de la educación. En opinión de Ausubel, para que se realice el aprendizaje, el alumno debe ser el encargado de reestructurar activamente tanto aquello que percibe, como las ideas, los conceptos y esquemas que ha adquirido previamente. En este sentido, su postura no sólo es que el alumno debe ser el encargado de reestructurar activamente tanto aquello que percibe, como las ideas, los conceptos y esquemas que ha adquirido previamente al interrelacionarse e interactuar los materiales de estudio y la información exterior con los esquemas previos de conocimientos y las características del propio alumno (Díaz Barriga, 1989). Ausubel es consciente de que no es posible adquirir todo el aprendizaje significativo por descubrimiento (a pesar de la importancia que tiene para él este método). Por este motivo, propone que en los estudios medios y superiores se utilice el aprendizaje verbal significativo de manera complementaria para adquirir un dominio suficiente de los contenidos pertenecientes a los diferentes currículos de cada nivel. Otro aspecto fundamental del aprendizaje significativo

reside en que, si aquél que debe aprender no presenta una actitud positiva y activa hacia el aprendizaje, no podrá relacionar con sus conocimientos previamente adquiridos los significados que se encuentran en los materiales didácticos.

Una enseñanza basada en la interacción profesor-alumno que enfatice el intercambio de preguntas, potencia el aprendizaje significativo. Si un alumno enuncia una pregunta relevante y acertada, está utilizando su conocimiento previo de manera no arbitraria y no literal. Eso es aprendizaje significativo. La utilización de materiales seleccionados con cuidado y la manera en la que el profesor presenta la tarea, son también factores que facilitan el aprendizaje significativo (Moreira, 2005).

En el aprendizaje significativo el profesor debe ayudar a construir el conocimiento del alumnado siendo una guía, promoviendo la participación y proponiendo acciones con el fin de construir significados cada vez más próximos a los del currículo<sup>3</sup> (Coll, 1988).

**Para Ausubel**, principal representante de los postulados del aprendizaje significativo, éste, es la manera en que una nueva información o conocimiento se relaciona de forma *no arbitraria y sustantiva* (no al pie de la letra) con lo que el alumno ya sabe. Por lo tanto, este proceso se convierte en el mecanismo principal con el que adquirir información en todos los ámbitos de conocimiento (Ausubel, 1963). La no arbitrariedad y la sustantividad son las características básicas del aprendizaje significativo, lo que significa que las ideas se relacionan con aspectos específicos y relevantes (imagen, símbolos, conceptos) que ya se encuentran en la estructura cognitiva del alumno.

Para Ausubel, el conocimiento previo es la variable crucial para que se produzca un aprendizaje significativo de tal manera que el alumno pueda establecer una relación con aquello que debe aprender.

---

<sup>3</sup> Desde el punto de vista de una enseñanza centrada en el aprendizaje, la idea más acertada de currículo quedaría definida como el conjunto de experiencias de aprendizaje que llevan a los alumnos a convertirse en aprendices autónomos, autorregulados y con recursos suficientes para aprender por sí mismos. Navarro, L. (2004). La alineación constructiva en el aprendizaje universitario. En J.C. Torre Puente (editor) y E. Gil Coria (editor), *Hacia una enseñanza universitaria centrada en el aprendizaje* (pp.111-142). Madrid: Universidad Pontificia Comillas.



### 2.1.3.2. El aprendizaje significativo desde diferentes perspectivas constructivistas

El concepto de aprendizaje significativo se sitúa en el marco del constructivismo y es originalmente propuesto por Ausubel en su teoría del aprendizaje.

“El concepto de aprendizaje significativo es compatible con otras teorías constructivistas pero su mayor potencial, en la perspectiva de la instrucción, está en la teoría original de Ausubel, complementada por Novak y Gowin” (Moreira, Caballero y Rodríguez, 199).

A continuación, exponemos algunas de las principales teorías constructivistas en las que subyace el concepto de aprendizaje significativo basándonos en el estudio en el que Moreira, Caballero y Rodríguez (1997) pretenden rescatar los significados originales del concepto.

**El aprendizaje significativo desde una perspectiva piagetiana.** La teoría de Piaget se basa en la asimilación, la acomodación, la adaptación y el equilibrio. La **asimilación** mental consiste en la interiorización de un evento o de un objeto en el comportamiento o en el conocimiento previamente adquirido a través de esquemas. Efectivamente, los esquemas para Piaget representan aquello que puede repetirse en una acción. Al principio se trata de movimientos reflejos y posteriormente incluyen movimientos voluntarios para convertirse finalmente en operaciones mentales. La **acomodación** permite al sujeto ajustarse a las condiciones externas y es necesaria para coordinar los múltiples esquemas de asimilación. La **adaptación** permite al alumno acercarse y conseguir ajustarse de manera dinámica con el medio. Por último, el **equilibrio**, sería el mecanismo que rige la relación entre la asimilación y la acomodación y se establecería en tres niveles: entre los esquemas del sujeto y los acontecimientos externos, entre los propios esquemas del sujeto y en una integración jerárquica de esquemas diferenciados.

**El aprendizaje significativo desde una perspectiva kellyana.** En opinión de George Kelly (1905-1967), “el hombre crea sus propias maneras de ver el mundo en el que vive. El mundo no las crea por él” (Kelly, 1955). En 1955 George Kelly presentó su Teoría de los Constructos Personales<sup>4</sup>. Según esta teoría, cuando miramos el mundo "exterior", lo hacemos a través de nuestras propias creencias o de lo que él llama constructos personales. Estos, serían mecanismos con los que dotamos de significado a nuestra experiencia, ordenando y estructurando la recepción de estímulos externos en base a los esquemas que se han ido configurando a lo largo del tiempo. Esta organización es dicotómica y clasifica las experiencias por categorías opuestas, como por ejemplo: "simpático-antipático",

---

<sup>4</sup> Kelly, GA (1955) The Psychology of Personal Constructs Volume 1: A Theory of Personality Volume 2: Clinical Diagnosis and Psychotherapy, New York, US: Norton.

"dulce-desagradable", "amigable-huraño". Estas polaridades son construidas de forma particular y distinta por cada persona. En esta actividad constructiva forman parte del proceso, la atención, la percepción, la memoria, la motivación, el aprendizaje y la conducta.

**Aprendizaje significativo desde un enfoque vigotskiano.** Para Lev Vigotsky (1896-1934)<sup>5</sup>, el aprendizaje se basa en las experiencias socio-culturales del individuo y, por tanto, en el entorno en el que vive. Para Vigotsky los mecanismos psicológicos siempre son cambiantes y dependen del contexto que rodea a los sujetos, siendo la asimilación de las actividades sociales y culturales la clave del desarrollo humano.

**Aprendizaje significativo en la perspectiva de Johnson-Laird.** Para Johnson-Laird (1936 - ), las personas usan modelos mentales para razonar y para representar el mundo con el que interactúan. Estos, conforman una representación de la realidad y, al mismo tiempo, crean imágenes que están relacionadas con modelos mentales particulares que están formados por sub-modelos y se pueden representar visual y analógicamente y sobre experimentos del pensamiento. Estaríamos hablando de representaciones internas que conectan la mente del sujeto con el mundo en el que se encuentra. Esta conexión o puente es necesaria porque, en opinión de Laird, no es posible entenderlo todo de forma directa.

**Aprendizaje significativo según la teoría de Novak.** Joseph Novak (1932 - ) diseñó los mapas conceptuales con el fin de ayudar al sujeto a relacionar la información que tienen que asimilar con la que ya poseen. Los mapas conceptuales representan gráficamente los conceptos en base a la selección, la jerarquización y el impacto visual. Novak, además, concibe los mapas conceptuales con el fin de ayudar al sujeto a relacionar la información que tienen que asimilar con la que ya poseen ya que las personas piensan y sienten y, por tanto, no entiende el aprendizaje como un acto sólo cognitivo. A lo largo del proceso de aprendizaje, las experiencias afectivas serán positivas cuando funciona la comprensión por parte de la persona que aprende y, por el contrario, negativas cuando el sujeto percibe que no está asimilando el nuevo conocimiento.

**Aprendizaje significativo según Gowin.** La teoría de educación de Dixie Bob Gowin (1925) presentada en su obra *Educating* (1981) se basa en la relación entre profesor, materiales educativos y alumno. Gowin considera la educación como un proceso de cambio debido a que todo proceso de aprendizaje está enfocado a conseguir algún tipo de cambio en el alumno en función de los contenidos a aprender. Todo lo aprendido, si lo hacemos de manera significativa,

---

<sup>5</sup> Las obras de Vygotsky han sido reeditadas años después de su muerte, algunas de ellas en varias ocasiones como es su obra más importante: Vygotsky (1962). *Thought and language*. Cambridge, MA:MIT Press.

podremos aplicarlo a contextos diferentes y diversos con lo que estaremos dotando de nuevos significados a la experiencia y, por ende, al concepto. A pesar de que el profesor sea el principal responsable de la enseñanza, sin embargo, no toda la responsabilidad recae sobre él ya que depende del alumno la elección de lo que desea aprender lo que tiene como consecuencia que el aprendizaje verdadero se produzca por la acción del que aprende y no por la del que enseña. El profesor, eso sí, deberá verificar que el aprendizaje del alumno es el adecuado y el alumno deberá saber si aprendió lo que el profesor pretendía enseñar. Este es un aspecto que señala la importancia de la evaluación, y de la recepción por parte del alumno de estos resultados de evaluación que permitan su autorregulación.

**El aprendizaje significativo por descubrimiento.** Desarrollada por Jerome Bruner (1915-) (Bruner, 1960), esta teoría se basa en un tipo de aprendizaje en el que el alumno tiene una gran participación. El docente debe dirigir al alumno para que alcance una meta propuesta por el propio profesor. También tiene que servir como mediador y guía aunque debe ser el alumno quien recorra el camino para alcanzar los objetivos propuestos. Por tanto, podemos decir que el aprendizaje por descubrimiento se produce cuando el alumno descubre por sí mismo lo que tiene que aprender gracias a que el profesor le ha presentado los recursos necesarios para ello.

**La teoría del aprendizaje social o teoría social cognitiva.** En 1977, Bandura sistematiza teóricamente la importancia que tienen los demás en el aprendizaje: el lenguaje, la relación profesor alumno, la publicidad, es decir, el entorno en general. Al igual que el constructivismo, la teoría del aprendizaje social de Bandura (1977), basa el aprendizaje en la interactividad y el medio social con la cognición como elemento clave. También reconoce la importancia de la observación y los modelados<sup>6</sup> para promover el aprendizaje. Los modelados son esenciales para aprender en la teoría del aprendizaje social. Éstos son descritos como un proceso que se produce en cuatro niveles: la atención, la retención, la reproducción y la motivación. La atención debe ser adquirida y mantenida para que el aprendizaje se produzca. Una vez aprendida, a continuación, la información o las habilidades deben ser conservadas o almacenadas en el nivel de retención. El nivel de reproducción es en el que las habilidades aprendidas se practican y mejoran. El último nivel, la motivación, se ve afectada por el refuerzo positivo y/o castigo. En esencia, el aprendizaje sustentado en los modelados se basa en la forma en que éstos se refuerzan en el entorno. La Teoría del Aprendizaje social contiene aspectos de prácticas educativas que pueden ser usadas en el desarrollo de la formación profesional y la formación práctica (Hinshaw, Burden, Shriner, 2012).

Se resumen a continuación, en la Figura 6, las diferentes perspectivas constructivistas del aprendizaje significativo expuestas anteriormente.

---

<sup>6</sup> También denominados imitación, aprendizaje observacional o aprendizaje vicario.

Marco constructivista	Aprendizaje significativo
<b>David P. Ausubel</b> (1918 - 2008), psicólogo y pedagogo estadounidense.	Relación de un nuevo conocimiento de forma <b>sustantiva y no arbitraria</b> con la estructura cognitiva del que aprende. El <b>conocimiento previo</b> es la variable crucial.
<b>Jean William Fritz Piaget</b> (1896-1980) epistemólogo, psicólogo y biólogo suizo.	El conocimiento aumenta cuando el esquema de <b>asimilación</b> sufre <b>acomodación</b> .
<b>George Kelly</b> (1905-1967) psicólogo estadounidense.	Teoría de los <b>constructos personales</b> . Ordenamos y estructuramos la recepción de estímulos externos en base a los esquemas que se han ido configurando a lo largo del tiempo
<b>Lev Vigotsky</b> (1896-1934) psicólogo ruso.	El desarrollo cognitivo depende del <b>contexto social, histórico y cultural</b> en el que se produce.
<b>Philip Johnson-Laird</b> (1936 -) psicólogo inglés.	Las personas usan <b>modelos mentales</b> para razonar en lugar de una lógica mental.
<b>Joseph D. Novak</b> (1932 -) educador e investigador estadounidense.	La <b>experiencia afectiva</b> es intelectualmente constructiva cuando funciona la comprensión por parte de la persona que aprende y la afectividad es negativa cuando el aprendiz siente que no está aprendiendo.
<b>Dixie Bob Gowin</b> (1925 -) profesor estadounidense.	Su teoría educativa se basa en la <b>relación entre profesor, materiales educativos y alumno</b> .
<b>Jerome S. Bruner</b> (1915) psicólogo y biólogo inglés.	<b>Aprendizaje por descubrimiento</b> : el docente debe presentar al alumno los recursos necesarios para que descubra lo que tiene que aprender.
<b>Albert Bandura</b> (1925 -) psicólogo y pedagogo canadiense.	Sistematiza la importancia que tiene el entorno en el aprendizaje: <b>teoría social cognitiva</b> .

Figura 6. Perspectivas constructivistas del aprendizaje significativo.

#### 2.1.4. Estrategias de Aprendizaje en el ámbito universitario

Existen tres líneas principales de investigación que se han centrado en el papel del alumno y su perfil actitudinal a la hora de definir y clasificar los procesos de aprendizaje en el ámbito universitario (Labatut, 2004). Cada una de ellas surgió en diferentes contextos tanto desde una perspectiva social como cultural y, por ende, también educativa: Australia, Suecia y Reino Unido. A pesar de ello, muchos son los investigadores que han hecho interesantes propuestas al respecto. En este apartado, queremos definir los enfoques o estrategias de aprendizaje en alumnos del ámbito universitario a través de las aportaciones de

diferentes investigaciones y, también, analizar las tres tendencias principales de investigación comentadas anteriormente.

Quizá, para comenzar, sea inevitable mencionar las aportaciones realizadas por John Biggs (1987a y b, 1988, 1993, 1999) con el fin de legitimar las investigaciones realizadas en torno al aprendizaje de los estudiantes que se encuentran en niveles universitarios (Macfarlane, Grant, 2012). Dichas aportaciones contribuyeron a definir el concepto de “enfoque de aprendizaje” aplicado al ámbito universitario. Este término, fue propuesto por Marton y Säljö (1976) con el fin de establecer las diferentes estrategias que han de ser desarrolladas por los estudiantes en esta etapa de su formación para abordar las diferentes tareas a las que deben enfrentarse. Para Biggs (1987a y b) las percepciones que reciben los estudiantes de las tareas que han de desempeñar en su proceso formativo, están influenciadas por sus circunstancias generando un proceso de aprendizaje concreto o, lo que es lo mismo, un enfoque de aprendizaje. Dicho enfoque supone que las circunstancias de cada sujeto y su actitud en el proceso de aprendizaje se encuentren relacionadas. Esto quiere decir que, aunque las características personales de los individuos les predisponen a afrontar el aprendizaje dentro de los parámetros actitudinales que les hagan dirigir la atención o el interés hacia un asunto o problema desde unos supuestos previos, no se puede obviar el hecho de que, además, pueden darse circunstancias concretas que provoquen, faciliten o impidan asumir el desempeño de determinados enfoques (Valle, González, Núñez, Suárez, Piñeiro, Rodríguez, 2000). Este es el motivo por el que los enfoques de aprendizaje determinan en los estudiantes tanto la manera de afrontar de manera consistente los diferentes quehaceres relacionados con el aprendizaje, como su actitud personal en determinadas situaciones (Biggs, 1991).

Cuando un estudiante se encuentra frente a un contexto de aprendizaje tiene que asumir dos retos bien diferenciados (Biggs, 1988): cuáles son los objetivos a alcanzar y qué debe hacer para conseguirlos. Por tanto, según Biggs (1988, 1993), los enfoques de aprendizaje están planteados en base a un “qué” y a un “cómo”. La combinación de ambas cuestiones nos llevará a desarrollar un proceso de aprendizaje en el que el conocimiento, la concienciación, el control y la naturaleza de los procesos de aprendizaje se verán implicados. Estaremos hablando, por tanto, de un proceso metacognitivo. Cabe destacar, por tanto, que los enfoques de aprendizaje se enmarcan en teorías de aprendizaje que parten de la perspectiva del alumno. De tal manera, nos encontramos una vez más frente a la evidencia de que el sujeto que aprende es el objetivo de la educación y no el docente o los contenidos curriculares.

A continuación analizaremos brevemente las tres líneas de investigación más significativas desarrolladas en este ámbito.

**John Biggs.** Este investigador australiano distingue dos ámbitos especialmente favorecedores de la calidad de la enseñanza: el contexto y el aula. Para demostrar su hipótesis, Biggs realizó una investigación en la Universidad de Australia (1979) por la que estableció una clasificación de los distintos enfoques de aprendizaje:

- Enfoque *profundo*: relacionado con la motivación y el interés mostrados por el estudiante hacia la materia. La reflexión, la lectura consciente y el debate son algunas de las estrategias que utiliza para abordar el proceso de aprendizaje.
- Enfoque *superficial*: condicionado por agentes externos que le inducen al miedo al fracaso. Utiliza estrategias básicamente memorísticas.
- Enfoque del *logro*: motivado por el deseo de obtener las mejores calificaciones. Desarrolla diferentes habilidades relacionadas en su mayoría con la organización para bordar el estudio y obtener resultados altamente exitosos.

**Ference Marton**, junto con su grupo de investigación de la Universidad de Gotemburgo (Suecia), hizo visible el motivo por el que los alumnos optan por diferentes caminos para comprender los conceptos. Además de interpretar el contexto de aprendizaje debe construir el significado del contenido del mismo. Esto ocurre como consecuencia de su propia experiencia del acto de aprender. Para Marton, el descubrimiento debe ser la estrategia principal de aprendizaje y, éste está estrechamente relacionado con los posibles y diferentes niveles de comprensión. El equipo de Marton realizó una investigación entrevistando a los sujetos para saber cómo abordaban la lectura de materiales sobre los que, posteriormente, tendrían que responder a una serie de cuestiones relacionadas con el contenido de dichas lecturas.

Fruto de su investigación, Marton y Säljö (1976b) establecen la siguiente clasificación:

- La conclusión orientada: el aprendizaje se basa en la relación entre el argumento y la evidencia. Se intenta relacionar los nuevos conceptos con los conocimientos adquiridos previamente por el sujeto.
- La descripción orientada: el aprendizaje se basa en la adopción de las ideas principales de la información contenida en el artículo.

Distingue entre motivación intrínseca (interés del sujeto por aprender), basado en el encuentro profesor-alumno y motivación extrínseca (memorización realizada por el sujeto) que actuará para el agente externo (profesor) y no para sí, sintiéndose a menudo bajo la sensación de la amenaza y experimentando frecuentemente ansiedad.

**Noel Entwistle**, en el Reino Unido, puso de manifiesto con sus investigaciones realizadas con el llamado grupo de Lancaster (1988) que el alumno aprende a partir de sí mismo. Previamente había basado su teoría en otras investigaciones (Entwistle, Ramsden, 1983 y Entwistle, Kozéki, 1985) realizadas con el fin de determinar las estrategias y resultados de aprendizaje de los alumnos. Define tres maneras de afrontar el estudio:

- Enfoque de profundidad: alto grado de implicación e intento de profundización en su comprensión. Existe la intención de relacionar los conceptos nuevos con adquiridos previamente y las experiencias personales.
- Enfoque superficial: se aborda la tarea con preocupación por memorizar la información cuyo recuerdo, según creen los alumnos, les valdrá en la futura evaluación.
- Enfoque estratégico: se caracteriza por intentar obtener el máximo rendimiento posible mediante la planificación de las actividades, los materiales y el tiempo disponible.

Por tanto, una misma tarea presentada de igual manera a los alumnos, podrá tener diferentes enfoques según la actitud que desarrolla cada uno frente al proceso de aprendizaje: conectar lo previo y lo nuevo, memorizar elementos discretos de información o rentabilizar de la mejor manera posible el tiempo y esfuerzo disponibles.

Según la motivación del alumnado, la tendencia mostrada en el estudio señalaba que, quienes enfocaban la tarea en profundidad, mostraban mayor interés por el contenido y su relevancia; quienes lo hacían desde el enfoque superficial, estaban condicionados por el afán de éxito y, los que optaron por un enfoque estratégico, experimentaban un mayor temor al fracaso.

Todo lo anterior nos hace consciente que cada uno de nuestros alumnos, por sus características personales, tenderá a fijarse, en su proceso de formación, unos tipos de metas y unas estrategias para alcanzarlas. Estos enfoques para abordar el aprendizaje pueden ser orientados y modificados por el profesor hacia

procesos de estudio que consigan un aprendizaje significativo y una mayor profundidad en la comprensión de los conceptos estadísticos, que es el objetivo del presente trabajo.

### **2.1.5. La enseñanza de la estadística desde una perspectiva constructivista.**

Muchos son los avances pedagógicos que se han logrado gracias a la ayuda que el constructivismo nos ha proporcionado ayudándonos a interpretar el conocimiento y a establecer múltiples y diferentes maneras de afrontar los procesos de aprendizaje relacionando éstos con el contexto, las circunstancias personales y las experiencias previas de los sujetos. Si pensamos en cómo hemos adquirido las competencias propias relacionadas con aquello a lo que nos dedicamos, en muchos casos, convendremos en que la mayoría de los conocimientos no teóricos que poseemos los hemos ido construyendo a través de la acción y del apoyo de numerosos agentes tanto personales (Bandura, 2001), como materiales que hemos ido aprendiendo, paulatinamente, a utilizar con autonomía. El hecho de que el constructivismo nos ayude a entender el aprendizaje nos debería hacer más proclives a innovar en nuestro entorno pedagógico y favorecer el desarrollo de estrategias diferentes a las que hemos seguido hasta ahora, en nuestro caso, en la enseñanza de la estadística.

El constructivismo nos hace asumir la responsabilidad de pensar no sólo en lo que debemos hacer como docentes para favorecer el aprendizaje de nuestros alumnos si no, también en cómo se deben establecer las relaciones entre los sujetos tanto docentes como discentes y con la multitud de instrumentos que nos acompañarán en el paulatino proceso de desarrollo de la comprensión para razonar estadísticamente con autonomía (Ordóñez, 2004). Algunos autores se han preocupado por remarcar factores a tener en cuenta, desde una perspectiva cognitiva y constructivista del aprendizaje, en la enseñanza de la estadística (véase, por ejemplo Garfield, 1995 y Lovett y Greenhouse, 2000)

Si un concepto es necesario a la hora de aprender a razonar estadísticamente, este es el de flexibilidad. Tal y como asegura Perkins (1997), “la comprensión no necesariamente está ligada a las representaciones mentales, pero sí se traduce siempre en flexibilidad de desempeño”. Para él, “el verdadero aprendizaje se traduce en poder pensar y actuar de manera flexible, en contextos diferentes, con aquello que se va aprendiendo” (Perkins, 1992 y 1997).



Aunque el aprendizaje siempre se construye individualmente, no debemos obviar que los mecanismos que se activan para obtener los resultados de aprendizaje, sólo lo hacen gracias a la relación e interacción con otros agentes.

El aspecto que nos interesa de los postulados de Perkins de cara a la enseñanza de la estadística es su concepción de la comprensión. Para este autor, es necesario saber utilizar de manera flexible los conocimientos que ya se poseen de manera que las acciones que se realicen también lo sean y se puedan utilizar en diferentes escenarios. Para Perkins, el resultado de aprendizaje más importante y la constatación de que los conocimientos se han adquirido adecuadamente es contar con la capacidad de utilizar lo aprendido en diferentes circunstancias. De cara al aprendizaje de la estadística, se trataría de adquirir una serie de mecanismos que ayuden a razonar estadísticamente en diversos contextos pero de forma lógica y ordenada. De manera gradual el alumno adquirirá la habilidad para extrapolar tanto los conocimientos adquiridos como los que vaya adquiriendo y aplicarlos en contextos diferentes (Perkins, 1977). Es necesario desarrollar la capacidad cognitiva y científica para que el alumno aprenda a decidir de manera razonada. Dicho conocimiento vendrá dado cuando el estudiante sepa relacionar sus circunstancias con el entorno y con los conocimientos adquiridos y realizar acciones para que esto suceda. Esto es lo que necesitamos, precisamente, para desarrollar nuestra capacidad de razonamiento estadístico (Perkins, 1999).

Tenemos que ocuparnos de que el que alumno adquiera la capacidad de fundamentar las relaciones entre lo que le enseñamos en estadística y otros conocimientos de su formación enfermera teórica y práctica así como de otros aspectos de su vida, tal como postula Ausubel (1968) en sus teorías educativas. Extraer información y conocimientos que posee el alumno y hacerlos relevantes de manera que ayudemos al alumno a ir “anclando”, en ellos los nuevos conceptos estadísticos. Hay que procurar también que la relación profesor-alumno facilite el intercambio de información y preguntas, que las tareas propuestas sean amenas e indiquen el camino para que el alumno reestructure y construya su aprendizaje. Que el alumno podrá realizar esa construcción si los contenidos, tal como sostiene Piaget, poseen ya un elevado grado de elaboración, que en nuestro caso tiene que ver con la profundidad de la comprensión de los conceptos estadísticos. También es importante que tengan una vivencia positiva de su experiencia educativa que facilite el aprendizaje significativo (Novak, 1977).

Los aprendizajes adquiridos a través de la experiencia pueden resultar muy valiosos durante períodos prolongados en el tiempo. Su vigencia puede

peligrar, sin embargo, cuando dichos aprendizajes choquen de manera conflictiva con experiencias nuevas. Según Gardner (1991), el conocimiento se desarrolla en el sujeto sumando a lo que ya sabe lo nuevo que va aprendiendo. Generalmente, cuando aprendemos acerca de algo es habitual que la persona cuente ya con ciertos conocimientos previos. En muchos casos, ese conocimiento ya “adquirido” puede ser erróneo, sesgado o insuficiente. Esto suele ser muy habitual en relación con algunos conceptos estadísticos erróneos con los que los alumnos llegan al aula.

Boix-Mansilla y Gardner (1997) plantean un análisis de cuatro factores para el manejo de la comprensión, que nos resulta de interés cuando estamos hablando de la enseñanza de la estadística fundamentada en una comprensión de fondo de los conceptos. Dichos factores son:

El **conocimiento** se trata de un factor transversal a los tres restantes y sólo se habrá adquirido fiablemente cuando se pueda expresar poniendo de manifiesto los contenidos de todos los factores que nos permitirán evaluar la comprensión de los contenidos de aprendizaje.

Los **propósitos** o la flexibilidad para usar y aplicar lo aprendido en diversos escenarios resulta indispensable para entender y ser capaz de transmitir los resultados acerca de las cuestiones planteadas.

Los **métodos** han de basarse en la creencia de que el conocimiento es consecuencia de rigurosos mecanismos de indagación.

Las **formas** están relacionadas con la adquisición de la capacidad para transmitir, utilizando diferentes vías en escenarios diversos, lo que se ha comprendido. La enseñanza de la estadística debe trasladarse a escenarios de aprendizaje en los que los estudiantes puedan experimentar con los conocimientos adquiridos y moldearlos para poder razonar estadísticamente cada vez con mayor complejidad, profundidad y autonomía.

Desde este punto de vista señalamos la importancia del uso de la evaluación para la construcción del aprendizaje. Las evaluaciones formativas son un material de apoyo que han de estar pensadas para ayudar a los alumnos a comprender los conceptos y para verificar que el alumno ha aprendido lo que le hemos enseñado de forma significativa. Al mismo tiempo los estudiantes tendrán un elemento para su autorregulación al recibir información de sus resultados de evaluación. La evaluación como el resto de materiales educativos (Gowin, 1981) son los vehículos por el que el profesor y los alumnos comparten significados con

respecto a conocimientos (enfoque de Vigotski). En la enseñanza de la estadística, la secuencia de contenidos, la elección de ejemplos y de datos con los que trabajar con los alumnos, los programas informáticos, el uso de simulaciones, gráficos, etc. ha de estar pensado para que tenga lugar de forma efectiva ese encuentro entre los que el profesor quiere enseñar y que el alumno de verdad pueda aprenderlo.

El constructivismo ofrece a los docentes una manera de reflexionar acerca de cómo piensan los alumnos y cómo llegan al conocimiento. La pedagogía constructivista requiere que los profesores tomen en consideración lo que los alumnos saben, lo que necesitan saber y cómo deben avanzar hacia el conocimiento deseado. Dado que, lo que los alumnos necesitan aprender puede cambiar y porque las experiencias necesarias para que exploren conceptos pueden variar, los docentes se encuentran en muchos casos en situaciones imprevisibles. La enseñanza constructivista requiere una receptividad y capacidad de respuesta por parte tanto de los estudiantes como de los docentes (Draper, 2002).

Tal y como hemos visto, la idea central de la teoría constructivista se basa en que la adquisición del conocimiento debe ser un proceso activo basado en el conocimiento existente y la experiencia del alumno. El objetivo de la enseñanza de la estadística debería ser que los estudiantes adquieran la capacidad de dar sentido a los contenidos de la asignatura por sí mismos contando con la guía y apoyo del profesor a lo largo del proceso. Al mismo tiempo, la colaboración entre los alumnos puede ayudar a que el aprendizaje resulte más efectivo.

Tanto el conocimiento existente del alumno como su propia experiencia, se constituyen en la base para la construcción de nuevos conceptos en el aprendizaje. Para decidir de manera razonable desde qué punto comenzar la enseñanza de la estadística, el profesor deberá realizar las acciones necesarias para la colaboración entre los alumnos puede ayudar a que el aprendizaje resulte más efectivo con los que el estudiante llega a la clase de estadística. Sólo de esta manera, el profesor podrá diseñar un plan de enseñanza adecuado a sus alumnos.

Durante las clases, el profesor debería usar métodos efectivos que permitan a los alumnos llegar a ser el centro de su propia enseñanza motivando así su interés por el aprendizaje. De acuerdo, una vez más, con la teoría constructivista, la mejor manera para que un alumno comprenda el conocimiento que ha de adquirir, es ponerle en una situación real para permitirle descubrir y resolver el problema por sí mismo (Yahong, 2011). Afortunadamente,

la estadística es una disciplina que nos ofrece múltiples oportunidades en este sentido.

Una última alusión a la teoría constructivista nos lleva a reflexionar acerca de que el aprendizaje de los estudiantes es reflexivo. El proceso de llegar a entender un concepto tiene lugar en la mente del alumno. El paso final de lograr la plena comprensión de un concepto es muy íntimo, casi místico. Se escapa a cualquier persona ajena, como puede ser el profesor, saber si el proceso de aprendizaje se ha llevado a cabo en su totalidad (Lerman, 1989). Sólo a través de un instrumento de medida, tanto los estudiantes como el profesor, pueden corroborar que se han alcanzado los objetivos de aprendizaje. En la parte empírica de esta investigación, proponemos un instrumento que nos ayudará a tal fin.

<b>Elementos constructivistas en la enseñanza de la estadística</b>
<ul style="list-style-type: none"> <li>• La adquisición del conocimiento es un proceso activo que se basa en el ya existente y la experiencia del alumno. <ul style="list-style-type: none"> <li>Conocer los conocimientos previos de los estudiantes.</li> <li>Ayudar a fundamentar relaciones entre lo que le enseñamos en estadística y otros conocimientos de su formación teórica y práctica y de otros aspectos de su vida.</li> </ul> </li> <li>• El aprendizaje significativo se construye con contenidos que tienen un importante grado de profundidad. <ul style="list-style-type: none"> <li>Enseñanza de Conceptos estadísticos razonados.</li> </ul> </li> <li>• El proceso de enseñanza-aprendizaje se caracteriza por compartir significados entre alumno y profesor “vehiculados” por los materiales educativos. <ul style="list-style-type: none"> <li>Selección de materiales curriculares, bases de datos, aplicaciones informáticas, tareas, etc.</li> </ul> </li> <li>• El conocimiento adquirido ha de ser flexible y utilizable en contextos diversos. <ul style="list-style-type: none"> <li>La enseñanza de la estadística debe trasladarse a escenarios de aprendizaje que fomenten esta flexibilidad.</li> </ul> </li> <li>• El conocimiento es consecuencia de rigurosos mecanismos de indagación. <ul style="list-style-type: none"> <li>Utilizar situaciones reales para permitir al alumno descubrir y resolver el problema por sí mismo.</li> </ul> </li> <li>• La relación profesor-alumno debe facilitar el intercambio de información y preguntas. <ul style="list-style-type: none"> <li>Actividades participativas</li> </ul> </li> <li>• La colaboración entre los alumnos puede ayudar a que el aprendizaje resulte más efectivo. <ul style="list-style-type: none"> <li>Actividades cooperativas.</li> </ul> </li> <li>• Una vivencia positiva de su experiencia educativa facilita el aprendizaje significativo.</li> <li>• Es necesario un instrumento de medida que corrobore que se han alcanzado los objetivos de aprendizaje.</li> </ul>

Figura 7. Elementos constructivistas en la enseñanza de la estadística.

## 2.2. Evolución histórica de la enseñanza de la estadística

Las implicaciones del enfoque constructivista sobre la enseñanza de la estadística se puso de manifiesto en la aparición de múltiples estudios interesados en los procesos cognitivos y las estrategias adecuadas para que los estudiantes aprendan mejor la estadística. Como consecuencia de estas investigaciones y del interés por aplicar los principios teóricos para la consecución de un aprendizaje significativo, surgió un movimiento orientado a conseguir que los estudiantes construyan su conocimiento mediante actividades que les faciliten pensar y razonar estadísticamente. Esta reforma ha cambiado sustancialmente la enseñanza de esta materia.

### 2.2.1. Antecedentes históricos

En la década de 1980 surge desde la psicología el interés por los errores que se cometen en el razonamiento acerca de los conceptos estadísticos, influenciados por el trabajo pionero de Kahneman, Slovic y Tversky (1982) que, interesados por los procesos cognitivos, tratan de entender cómo razonan estadísticamente las personas y cómo esto influye en la toma de decisiones. Estas investigaciones ayudaron a identificar muchos malentendidos y errores en el razonamiento que se han generalizado y cuyo conocimiento es interesante para los profesores de estadística (Garfield and Ahlgren, 1988; Shaughnessy, 1992; Garfield, 1995).

Otras investigaciones se centraron en los estudiantes de estadística, en unos casos tratando de probar diferentes modelos de razonamiento estadístico que explicaran los errores de los estudiantes (Konold, 1989; Pollatsek, Konold, Well, and Lima 1984) o intentando observar diferencias en el razonamiento de éstos respecto a la población general (Fong, Krantz y Nisbett, 1986). En estos primeros estudios no se aportan sugerencias para mejorar la enseñanza de la estadística.

En los años noventa se produce un cambio en el objeto de la investigación sobre la enseñanza de la estadística hacia el estudio de estrategias para que los alumnos aprendan a razonar estadísticamente. Se inicia un movimiento para reformar la enseñanza de la estadística en los cursos más básicos o de introducción de esta materia (Cobb, 1992; Hogg, 1992). Esta reforma basó sus argumentos en los estudios que se realizaron sobre la dificultad de los alumnos en la comprensión de conceptos básicos. Los movimientos de reforma se centraron, desde los contenidos y la pedagogía, en cambiar el enfoque de una

estadística de cálculos y procedimientos, a otra que pone el énfasis en el razonamiento y el pensamiento estadístico (Moore, 1997).

Este cambio vino de la mano del desarrollo de programas informáticos que no sólo han facilitado los cálculos sino que han pasado a formar parte del proceso de aprendizaje (Lovett, 2001).

Zieffler A., Garfield J., Alt S., Dupuis D., Holleque K., and Chang B. (2008) estudian cuáles son las ideas estadísticas mal entendidas por los estudiantes desde el punto de vista de los patrones de pensamiento, analizando también otras razones que expliquen esos errores en el razonamiento, y sugiere métodos para ayudar a los alumnos a comprender los conceptos.

Las investigaciones sobre la enseñanza de la estadística de los últimos tiempos tratan de establecer una unión entre las teorías sobre cómo aprenden los estudiantes y la demostración de la eficacia de las distintas estrategias de aprendizaje adecuadas para el desarrollo del pensamiento estadístico (Lovett, 2001, Garfield J., Ben-Zvi D. 2009)

En la disciplina en la que se enmarca este trabajo, la Enfermería, la asignatura de estadística se ha ido incorporando paulatinamente en los últimos 20 años en los planes de estudio. Estos cursos de estadística están diseñados para aportar a los estudiantes una comprensión básica de los conceptos estadísticos. Otras disciplinas también de ciencias de la salud han incorporado una formación estadística en el currículo, siendo en estos momentos una asignatura presente prácticamente en todas las titulaciones de Ciencias de la Salud, Ciencias Sociales y Ciencias Técnicas. La asignatura de estadística aparece asociada a la aplicación del método científico para la investigación justificada en el contexto universitario. La concepción de la Estadística como un campo de las matemáticas ha derivado hacia una disciplina cuya enseñanza se centra en proporcionar ocasiones de plantear preguntas, definir problemas, generar hipótesis y definiciones operativas, diseñar experimentos y encuestas, recoger datos y manejar errores de medida, resumir y analizar datos, comunicar resultados y diseñar futuros trabajos (Hogg, 1991 citado en Blanco, 2004).

La coincidencia en el tiempo con los avances informáticos ha llevado a que el planteamiento de un curso básico de estadística haya pasado de una concepción basada principalmente en enseñar a los alumnos a realizar cálculos, a otra en la que los cálculos no son tan importantes, y que tiene como objetivo aprender a razonar estadísticamente. Este cambio en el aula provoca también un

cambio en el interés de la investigación en cuanto a las temáticas tratadas en los mismos. El foco de atención se centra entonces, en el estudio de las dificultades que tienen los alumnos para desarrollar un adecuado razonamiento estadístico.

Algunos estudios han tratado sobre la existencia de una mayor conciencia de la incapacidad de los estudiantes a pensar o razonar estadísticamente, (Sotos, A. E. C., Vanhoof, S., Van den Noortgate, W., & Onghena, P. 2007); la dificultad para desarrollar la comprensión profunda necesarias para integrar los conceptos estadísticos y aplicarlos en sus razonamientos, incluso a pesar de tener un buen rendimiento académico (delMas R., Garfield J., y Chance B. 1999; Hirsch y O'Donnell, 2001 ); las dificultades de los estudiantes para entender ideas estadísticas como la covariación, incluso después de un curso introductorio de estadística ( Zieffler A., Garfield J., 2009) y la existencia de aprendizajes erróneos, entendimientos incompletos o superficiales acerca de la variabilidad que se detecta en estudiantes universitarios (Garfield J., Ben-Zvi D. 2005). Otros estudios analizan los problemas que surgen cuando los profesores tratan de enseñar conceptos abstractos (Ben-Zvi, Garfiel 2004).

### 2.2.2. Conceptos emergentes en la didáctica de la estadística

Del contexto anteriormente expuesto surgen los términos razonamiento estadístico, pensamiento estadístico y cultura estadística y comienza a tener especial importancia analizar qué es lo que se entiende por cada uno de ellos.

En gran parte de la literatura, estos conceptos se establecen de forma jerárquica en términos de metas de aprendizaje, siendo la meta básica la cultura o alfabetización estadística y el pensamiento estadístico la meta más avanzada. Ben-Zvi y Garfield (2004) plantean una categorización que actualmente se encuentra entre las más aceptadas por los profesores de estadística. En ella se diferencia entre Alfabetización Estadística, Razonamiento Estadístico y Pensamiento Estadístico, según estos autores y otros que utilizan esta clasificación.

- La **alfabetización o cultura estadística** implica la comprensión y la utilización del lenguaje y las herramientas estadísticas básicas: saber lo que significan los términos estadísticos, la comprensión y el uso de los símbolos estadísticos, y reconocer y ser capaz de interpretar las representaciones de datos (Rumsey 2002). La alfabetización estadística abarca conocimientos básicos e importantes como ser capaz de organizar

y mostrar datos de manera adecuada y trabajar con diferentes representaciones de datos, así como la comprensión de conceptos, vocabulario y símbolos (Ben-Zvi y Garfield 2004). También incluye la capacidad de interpretar, evaluar, argumentar y comunicar la información estadística.

Wallman (1993), citado por Blanco (2004), citado por Wade y Goodfellow (2009) define la alfabetización estadística como “ la capacidad de comprender y evaluar críticamente los resultados estadísticos que impregnan nuestra vida cotidiana, junto con la capacidad de apreciar las contribuciones que el pensamiento estadístico puede hacer en las decisiones públicas y privadas, profesionales y personales”. Hace hincapié en su importancia en la sociedad civil.

Para Schield (1999) “la alfabetización estadística es la capacidad de leer e interpretar los datos: la capacidad de utilizar la estadística como prueba en los argumentos. Alfabetización estadística es una competencia: la capacidad de pensar críticamente acerca de las estadísticas” (p.1.)

- **El razonamiento estadístico** es la forma en la que los alumnos razonan con ideas estadísticas y dan sentido a la información estadística (Garfield, 2002). Requiere la interpretación de resúmenes o representaciones de datos y supone la conexión de un concepto a otro o la combinación de ideas, entender y ser capaz de explicar los procesos estadísticos e interpretar los resultados. (Ben-Zvi y Garfield 2004).
- **El pensamiento estadístico** implica una comprensión de por qué y cómo se llevó a cabo la investigación estadística. Esto incluye el reconocimiento y la comprensión del proceso de investigación, desde la pregunta de investigación, la recopilación de datos, la elección de los análisis y la interpretación de resultados. Abarca también la comprensión de los modelos de simulación de fenómenos aleatorios, y cómo y cuándo han de utilizarse las herramientas de inferencia, así como la capacidad de utilizar el contexto de un problema para planificar la investigación, evaluarla y sacar conclusiones (Wild and Pfannkuch 1999). El pensamiento estadístico consiste en mantener una relación constante con el contexto del problema y ser capaz de interpretar las conclusiones en términos no estadísticos, así como ser capaz de ver el proceso completo con constantes revisiones de cada componente (Ben-Zvi y Garfield 2004, Chance 2002).

El pensamiento estadístico aplicado a la vida real es el argumento que Snee en 1993 plantea ante la necesidad de cambiar el contenido y la manera de



educar estadísticamente para permitir a los estudiantes experimentar el uso del pensamiento y los métodos de estadística para hacer frente a los problemas y cuestiones reales. Defiende que estas experiencias producen una actitud más favorable hacia la disciplina y un mayor deseo de utilizar el pensamiento estadístico de forma cotidiana.

DelMas (2002) señala que algunos investigadores consideran que estos conceptos se solapan entre sí aunque tienen aspectos que los diferencian. Otros consideran que el razonamiento y el pensamiento estadístico son objetivos en el desarrollo de todo lo que abarca lo que se ha llamado cultura estadística.

Robert DelMas (2002) propone una categorización teniendo más en cuenta las tareas que los contenidos estadísticos, lo que determina en cada caso los resultados de aprendizaje asociados a estos tres conceptos, Figura 8.

Alfabetización estadística	Razonamiento	Pensamiento
Identificar	Por qué	Aplicar
Describir	Cómo	Criticar
Expresar de otro modo	Explicar	Evaluar
Trasladar	(el procedimiento)	Generalizar
Interpretar		
Leer		

Figura 8. Tareas que pueden distinguir los tres dominios de instrucción.

Fuente: del Mas, R. (2002). *Statistical Literacy, Reasoning, and Learning: A commentary*. *Journal of Statistics Education*. Volumen 10, nº3.

Consultado: [http://www.amstat.org/publications/jse/v10n3/delmas\\_discussion.html](http://www.amstat.org/publications/jse/v10n3/delmas_discussion.html)

Pese al esfuerzo de definir y delimitar estos conceptos, los últimos estudios centran su atención en el desarrollo del razonamiento estadístico y, aunque no hay un acuerdo en la manera de definirlo, la mayoría de los investigadores consideran que incluye el uso de conceptos, y herramientas estadísticas, para resumir, hacer predicciones y sacar conclusiones acerca de los datos (Lovett 2001). La capacidad de hacer interpretaciones correctas de los datos, es fundamental en cualquier definición de razonamiento estadístico, e implica que los estudiantes se involucren. Es decir, que sean capaces de explicar los procesos estadísticos (Ben-Zvi y Garfield, 2004).

Teniendo en cuenta que el presente trabajo tiene por objeto de interés el razonamiento estadístico, enunciaremos nuestra propia definición del concepto, aunque compartimos con DelMas (2002) que hay aspectos que se solapan con el denominado pensamiento estadístico.

Por otro lado y de acuerdo con las anteriores definiciones de ambos conceptos, consideramos que el pensamiento estadístico es la meta más alta de aprendizaje y, por tanto, acoge de forma inclusiva los procesos de razonamiento. Es por ello que también se harán referencias y serán base de nuestro trabajo, estudios sobre el pensamiento estadístico.

Se enuncia a continuación la definición operacional de razonamiento estadístico sobre la que se fundamenta el presente trabajo.

*Razonamiento estadístico: es la forma en la que los alumnos ordenan, combinan las ideas y relacionan los conceptos con coherencia, para entender, interpretar y dar sentido a una información estadística.*

### 2.2.3. El cambio en la enseñanza de la estadística

La enseñanza de habilidades estadísticas y procedimientos ha dominado, durante años, el paisaje de la educación de los cursos básicos de estadística. Sin embargo, el método tradicional de teórica de estadística fundamentada en las matemáticas, ha demostrado ser un método ineficaz para formar estudiantes competentes en estadística. Después de finalizar un curso de iniciación de estadística impartido con dicho método, muchos estudiantes universitarios siguen teniendo malentendidos en las ideas y conceptos básicos de estadística y sus habilidades de razonamiento sobre conceptos estadísticos fundamentales siguen siendo insuficientes.

Como ya se ha comentado, esta situación, de la mano de las teorías constructivistas del conocimiento, hizo posible que se iniciara una importante reforma que en los últimos tiempos ha producido múltiples llamadas de atención para promover el cambio en la enseñanza de la Estadística. Este cambio se basa en el desarrollo del razonamiento, apartándose de la memorización, de las definiciones y de los procedimientos (Ben-Zvi y Garfield 2004), y recomendando, como ya lo hicieron anteriormente otros autores (Gnanadesikan, Scheaffer, Watkins y Witmer, 1997; Moore, 1997; Shaughnessy, 1977), el uso de técnicas de aprendizaje activo en sustitución de las clases magistrales tradicionales. Este aprendizaje activo se basa en la exploración libre por parte de los estudiantes, dentro de un marco o de una estructura dada, en la que puedan desarrollar actividades centradas en sus habilidades y consolidar así su aprendizaje (Ormrod, 2003). Debemos alejarnos de métodos en los que

los alumnos leen textos, escuchan al profesor y hacen problemas, por modelos en los que los estudiantes aprenden mediante la comparación de datos con una teoría o conjetura, (Joiner, 1990).

La cuestión de cómo los estudiantes aprenden mejor la estadística ha sido considerada en diversos trabajos (Chance, 2005; Gal & Garfield, 1997; Garfield, 1995; Lovett & Greenhouse, 2000) que se han centrado, principalmente, en el contenido educativo y en los procedimientos.

Se han definido diversos escenarios para el aprendizaje de la estadística en los que el estudiante tiene un mayor grado de implicación. Escenarios que transforman la estructura didáctica tradicional centrada en el profesor y en una comunicación unidireccional en otra en la que el alumno, el aprendizaje autónomo y cooperativo son los protagonistas.

### **2.2.3.1. Aprendizaje activo**

Garfield (1995) sugiere que los estudiantes desarrollan mejor la construcción del conocimiento si se convierten en participantes activos del proceso de aprendizaje.

El aprendizaje, por tanto, tiene mejores resultados cuando los alumnos realizan actividades que les proporcionan oportunidades para pensar, razonar y reflexionar y cuando tienen la oportunidad de manejar por sí solos los datos. Aprenden, no porque se les explique lo que sucede, sino porque lo experimentan, haciendo sus propias inferencias, descubrimientos y conclusiones, construyendo sobre el conocimiento que ya poseen. Aparecen propuestas educativas para desarrollar el razonamiento estadístico basadas en la teoría constructivista del aprendizaje, en la que además el alumno maneja datos reales, y trabaja de forma cooperativa (Garfield J., Ben-Zvi D 2009).

El aprendizaje activo no debe reducirse a las prácticas con ordenador con un paquete de software estadístico. En la mayoría de los casos estas clases se convierten en la realización de ejercicios siguiendo instrucciones para su resolución. Otros textos han propuesto como técnicas de aprendizaje activo una serie de actividades con objetos sencillos, como son las monedas, que pueden ser utilizadas en cualquier aula, (Scheaffer, Watkins, Gnanadesikan, y Witmer, 1996 ). Las características de estas actividades proporcionan a los estudiantes la posibilidad de descubrir conceptos (Keeler y Steinhorst y 1995); Moore, 1997).

La aplicación en enfermería de metodologías de aprendizaje activo de la estadística las encontramos en Schwartz, (2014), con las llamadas aulas Flipped. En esta experiencia aplicada en cursos de doctorado, se invierte el orden tradicional de transmisión de conocimientos, de manera que el alumno trabaja con los materiales y realiza actividades con anterioridad a la clase presencial con el profesor. Esta clase se destina a la realización de actividades guiadas, discusión de casos, permitiendo que los alumnos interactúen con el grupo y con el profesor que orienta la actividad. Este aprendizaje activo, muestra una mejora en el aprendizaje y una aceptación positiva de la metodología por parte de los estudiantes.

En alumnos de enfermería también aplica Hayat, (2014) múltiples opciones metodológicas para conseguir un aprendizaje activo: utilización de datos reales, trabajo cooperativo, utilización de programas informáticos, etc.

De igual modo Shillam, Ho y Commodore-Mensah (2014), diseñan un curso de estadística online para estudios de máster en enfermería, centrado en el aprendizaje activo del estudiante. Los módulos de dicho curso tienen el siguiente formato: los contenidos se transmiten a través de conferencias grabadas, realización de actividades interactivas y la puesta en marcha de un panel de discusión.

### **2.2.3.2. Utilización de datos reales**

La utilización de datos obtenidos de la realidad y, si es posible, que susciten en el alumno un interés por su análisis contribuirá a facilitar la comprensión de los conceptos y procesos estadísticos. Nelson (2009), siguiendo esta recomendación de American Statistical Association (ASA), utilizó datos sobre la concentración atmosférica de clorofluorocarbonos obtenidos de un archivo público y embarcó a sus alumnos en la tarea de evaluar la eficacia de los acuerdos de Montreal. Una historia convincente con la que aprendían la regresión lineal.

Un método para generar datos reales, utilizado en grupos numerosos de estudiantes de enfermería, fue administrar un cuestionario a los alumnos el primer día de clase. La encuesta contenía preguntas destinadas a recabar datos con variables continuas y discretas, con las que posteriormente se trabaja en clase (Hayat, 2014)

### 2.2.3.3. Generar sus propios datos y trabajar con ellos

Hogg (1991) describe el problema de la enseñanza tradicional de la estadística en términos de diseño del aprendizaje: los estudiantes son aprendices pasivos y no entran en contacto directamente con los muchos problemas que se presentan en la recogida de datos y análisis. Sugiere que los estudiantes deben generar sus propios datos en lugar de utilizar un conjunto de datos de un libro de texto o los facilitados por el profesor. Al trabajar con sus propios datos, los estudiantes tienen oportunidades para definir problemas, formular hipótesis, diseñar experimentos, y analizar y sintetizar datos reales.

El uso de datos reales tiene el potencial de incrementar el compromiso y aprendizaje en los estudiantes universitarios de estadística. Neumann, Neumann y Hood (2010) en un curso universitario de introducción a la estadística, desarrollaron un trabajo en el que recopilaron datos de los propios alumnos usándolos de manera creativa a lo largo del curso para aprender conceptos relevantes y métodos estadísticos. Esta experiencia fue percibida por parte de los alumnos como un enfoque diferente y contribuyó a que mostraran un mayor interés en la comprensión y apreciación de la relevancia de la estadística.

Snee (1993) hizo hincapié en que, debido a que la recopilación de datos es el núcleo del análisis estadístico, el aprendizaje que se centra en el análisis de datos reales recogidos por los estudiantes los conecta con el sentido práctico de la estadística. Este es uno de los aspectos que la American Statistical Association (ASA, 2005), recomienda para la enseñanza de la estadística en la Universidad. Los alumnos que participan de forma activa en las tareas en las que pueden manipular y analizar datos reales recogidos por ellos mismo, demuestran una comprensión mucho más profunda de los conceptos estadísticos (Ben-Zvi, 2004; Mackisack, 1994)

### 2.2.3.4. Actividades y simulaciones con ordenador y casos prácticos

Hay muchas formas de incorporar el aprendizaje activo y los casos prácticos en la enseñanza de la estadística: formación asistida por ordenador (Ragasa, 2008), simulaciones por ordenador (delMas, Garfield y Chance, 1999), clases prácticas en el laboratorio (Bradstreet, 1996); o actividades en clase (Dietz, 1993; Gnanadesikan, Scheaffer, Watkins, y Witmer, 1997).

Weinberg, Wiesner y Pfaff (2010), desarrollaron una actividad de clase, que permitió a los alumnos reunir datos y explorar ideas de manera informal,

para llegar a un razonamiento formal de la inferencia estadística. En esta actividad los alumnos jugaron a una simulación de la lotería, para entender los conceptos subyacentes en las pruebas de inferencia estadística, modelos de probabilidad, variabilidad de las muestras, distribución muestral y errores.

El trabajo mediante proyectos (Hunter, 1977; Chance, 1997; Fillebrown, 1994; Ledolter, 1995; Mackisack, 1994), que ha sido utilizado con mucha frecuencia, posee las siguientes ventajas: permite situar a la estadística en un contexto determinado, facilita el trabajo en grupo y la interpretación de los resultados, da importancia a su utilidad y aplicación y motiva al alumno en el proceso de resolución de cuestiones más o menos interesantes.

Debido a la naturaleza abstracta de los conceptos estadísticos, los alumnos, a menudo, son inducidos a error y, en consecuencia, son llevados a transitar por el camino de las ideas falsas. Chan e Ismail (2012) indican, mediante una revisión de la investigación, que la integración de la tecnología tiene un impacto positivo a la hora de abordar los errores y mejorar el razonamiento estadístico, siempre que se use adecuadamente. El uso de herramientas informáticas, que facilitan los aspectos del cálculo matemático, ha de efectuarse de manera que la actividad se centre en otras tareas de orden superior. La utilización de herramientas tecnológicas en el proceso de enseñanza-aprendizaje facilita el diseño de escenarios de investigación para el abordaje de problemas globales, y posibilita indagar sobre los datos con diversas alternativas. (Blanco, A. 2004).

Morris, Joiner, Scanlon (2002), en un estudio de investigación sobre el empleo de actividades con ordenador para la enseñanza de la estadística realizado con alumnos de psicología, pusieron de manifiesto que las actividades que implicaban la manipulación directa de los datos contribuyeron a la comprensión, por parte de los estudiantes, de las medidas de tendencia central pero, por el contrario, no comprobaron que contribuyeran a la mejor comprensión de las correlaciones.

#### **2.2.3.5. Aprendizaje cooperativo**

El aprendizaje cooperativo se sustenta en la teoría social del aprendizaje de Bandura (1977) y la concepción de Perkins (1997) acerca de la cognición socialmente compartida como generadora de aprendizaje significativo. Respecto a esto, muchos investigadores han publicado estudios en los que han puesto de

manifiesto importantes logros al incorporar experiencias de aprendizaje cooperativo a las clases de introducción a la estadística (Dietz, 1993; Keeler y Steinhorst, 1995; Shaughnessy, 1977; Magel, 1998).

Los métodos de aprendizaje activo promueven el aprendizaje cooperativo, que permite a los alumnos aprender unos de otros. Con el aprendizaje activo los estudiantes descubren, construyen y entienden los conceptos estadísticos. En resumen, desarrollan el pensamiento estadístico al tiempo que las actividades que involucran a los alumnos en su aprendizaje hace que el aprendizaje sea más divertido (ASA, 2005).

### **2.2.3.6. Cambio conceptual, trabajo de errores**

Para lograr un cambio en la comprensión conceptual de los alumnos, no es suficiente la incorporación de técnicas de aprendizaje activo (Chance, Delmas y Garfield, 2004). Una teoría que tiene una sólida base de investigación en ciencias de la educación es la teoría del cambio conceptual. Esta teoría postula que los estudiantes aprenden cuando se cuestionan una concepción que ya poseen. Para ello, el primer objetivo es descubrir las preconcepciones que tienen los estudiantes sobre un tema o fenómeno y usar, a continuación, las técnicas adecuadas para ayudar a que los estudiantes cambien su marco conceptual (Davis, 2001). Este modelo propone que los estudiantes que poseen conceptos erróneos o mal entendidos necesitan que se les presente un problema o pruebas en contra con el fin de cambiar sus concepciones actuales. (Chance, Delmas y Garfield, 1999).

Chance et al. (2004) encontraron que estas actividades, diseñadas a partir de un marco conceptual diferente, mejoraron significativamente el razonamiento de los estudiantes sobre las distribuciones del muestro.

Aunque existen muchas investigaciones que indican la eficacia de técnicas de enseñanza alternativa, no siempre dejan claro cómo puede el profesor ponerlas en práctica (Garfield, 1995). Apenas existen libros de texto al margen de lo tradicional que recojan nuevas ideas para la enseñanza de la estadística (Bryce 2005). Pero lo cierto es que la experiencia del profesor de estadística hace que vaya adquiriendo la habilidad necesaria para detectar los aspectos que deben tenerse en cuenta con el objetivo de mejorar el aprendizaje. Estos aspectos, sin duda, coinciden con las Directrices de la American Statistical Association (ASA, 2005), que se tratan en el siguiente apartado.

### 2.3. Directrices para el aprendizaje de la Estadística. Guidelines for Assessment and Instruction in Statistics Education (GAISE)

Toda la investigación y las experiencias metodológicas que surgieron encaminadas a mejorar la enseñanza de la estadística y promover el razonamiento, tuvieron como consecuencia la aparición de modelos de enseñanza más eficientes que los modelos tradicionales. El modelo que se describe en este apartado es una guía para la que se recogen una serie de recomendaciones que fueron elaboradas por un grupo de trabajo constituido desde la American Statistical Association (ASA) y presidido por Joan Garfield, con el fin de marcar las directrices para el aprendizaje y la evaluación de la estadística en alumnos universitarios (Guidelines for Assessment and Instruction in Statistics Education GAISE 2010).

Editado por Andrew Zieffler y Stacy Karl (2010), este documento es una guía para el desarrollo de los procesos de enseñanza y aprendizaje de la estadística. Este informe ofrece un resumen en el que se recogen las recomendaciones para el cambio en la enseñanza de la estadística como resultado del movimiento de reforma en la enseñanza de esta materia.

Tal como se expuso anteriormente la no existencia de directrices acerca de las competencias estadísticas en la formación de enfermería, el informe GAISE, supone una herramienta de orientación, destinada directamente a los profesores, muy válida para enfermería.

*El movimiento de reforma de la educación estadística ocurrido en las dos últimas décadas, se ha centrado en que la estadística sea más accesible y menos intimidante. El informe GAISE (ASA, 2005) y sus seis recomendaciones son un buen punto de partida para cambiar la enseñanza de la estadística en enfermería hacia un enfoque de cultura, razonamiento y pensamiento estadístico. (Hayat, 2014. p. 197)*

El estudio que ha dado lugar a esta tesis, anterior a la publicación del artículo de Hayat (2014), ha tenido en cuenta el documento *Guidelines for Assessment and Instruction in Statistics Education* (GAISE, 2010), del que se habla en este apartado, para la metodología docente utilizada en el estudio empírico, así como para la elección de la intervención y de los instrumentos de evaluación utilizados.

Este documento (GAISE, 2010) basa sus orientaciones para el cambio de la enseñanza de la estadística en 6 recomendaciones:



### **2.3.1. Desarrollar la cultura y el pensamiento estadístico**

En esta primera recomendación se plantean tanto un modelo de formación como un esquema teórico del proceso de enseñanza-aprendizaje, que podemos denominar Modelo de pensamiento estadístico.

### **2.3.2. Usar datos reales**

Esta guía aconseja el uso de datos reales que permitan analizar el origen y el motivo de la recolección de los mismos, estimular el interés del estudiante y relacionar el análisis con el contexto real de un problema así como generar nuevas preguntas. Utilizar un conjunto de datos reales estimula el interés del alumno, le lleva a plantearse preguntas adecuadas y le motiva a responderlas teniendo en cuenta el proceso por el que se han obtenido dichos datos.

Recomienda, si es posible, utilizar datos procedentes de investigaciones científicas, así como datos de los propios alumnos u obtenidos por ellos. Estos datos pueden obtenerse en clase o de forma electrónica mediante cuestionarios que recojan información sobre aspectos diversos, siempre teniendo en cuenta la privacidad de la información.

El uso de un conjunto de datos reales que tenga interés para los estudiantes es una buena forma de hacerles pensar tanto en dichos datos como sobre conceptos estadísticos relevantes. Hacer preguntas sobre una base de datos que no tiene interés para el alumno no es recomendable para un curso introductorio de Estadística.

### **2.3.3. Importancia de la comprensión conceptual, más que el mero conocimiento de los procedimientos**

Lo más importante en un curso básico de estadística es que los conceptos importantes sean entendidos sólidamente, prestando más atención a los propios conceptos que a los procedimientos. Para esta comprensión conceptual puede utilizarse alguna técnica estadística, aunque se ha de evitar enseñar multitud de técnicas sin un mínimo de ideas subyacentes.

Si los estudiantes consiguen una buena base conceptual, tendrán más facilidad para entender pruebas estadísticas de mayor complejidad en su posterior formación académica o profesional. El contenido de un curso de introducción de estadística no debe ser ambicioso en su extensión, por el

contrario, debe centrarse en un núcleo reducido de conceptos que han de ser abordados con mayor profundidad.

La tecnología nos facilita los cálculos y permite que el esfuerzo se dirija a la interpretación de los resultados. El lenguaje y las fórmulas matemáticas deben utilizarse para mejorar la comprensión de los conceptos, evitando que los cálculos se presenten disociados de la comprensión.

#### **2.3.4. Fomentar el aprendizaje activo en el aula**

Utilizar métodos activos de aprendizaje en el aula es una manera de promover el aprendizaje colaborativo, permitiendo a los estudiantes aprender los unos de los otros. El aprendizaje activo permite a los estudiantes descubrir, construir y entender las ideas estadísticas importantes y configurar el pensamiento estadístico. Estas actividades tienen un beneficio adicional pues hacen que el estudiante se involucre en su aprendizaje y que este proceso resulte más ameno e, incluso, divertido.

En este tipo de práctica los alumnos aprenden a trabajar en equipo y se comunican con un lenguaje estadístico. La resolución de problemas en grupo, discusiones y debates, prácticas informáticas, simulación con el ordenador, o la recopilación anónima de datos entre los alumnos son algunas de las actividades que involucran a los estudiantes y hacen que el proceso de aprendizaje resulte divertido. Es importante que estas actividades sean evaluadas.

#### **2.3.5. Usar la tecnología para el desarrollo de la comprensión conceptual y el análisis de datos**

La tecnología ha cambiado la forma de trabajar de los profesionales de la estadística y en el ámbito educativo ha cambiado el qué y el cómo enseñar. Se ha mencionado anteriormente que la tecnología debe utilizarse para analizar los datos, permitiendo a los estudiantes centrarse en la interpretación de los resultados. El manejo de los programas informáticos como uso de la tecnología no es tan importante si tenemos en cuenta, además, que ésta cambia tan rápidamente. La tecnología debe ser utilizada para ayudar a los estudiantes a visualizar los conceptos y desarrollar la comprensión de ideas abstractas mediante simulaciones. No sólo como una herramienta para el cálculo, sino como una forma de explorar conceptualmente ideas y mejorar el aprendizaje del estudiante.

### **2.3.6. Utilizar las evaluaciones para mejorar el aprendizaje del estudiante y no sólo para valorarlo**

La información que recibe el alumno a través de su evaluación conduce el aprendizaje. Esto es debido a que, por un lado, los estudiantes sólo dan valor a lo que se evalúa y, por otro, a que la información que recibe de los logros o deficiencias de su proceso de aprendizaje permite su autorregulación.

La evaluación integrada como un componente esencial del curso, debe centrarse en la comprensión de las ideas clave, y no solo en las habilidades de realizar cálculos o procedimientos.

La evaluación pretende aumentar la capacidad de asociar ideas, ordenarlas y conectar conceptos con coherencia. Este objetivo se puede alcanzar realizando actividades que evalúen aspectos básicos relacionados con la alfabetización estadística, tales como la interpretación de gráficos o artículos de prensa o evaluando el pensamiento estadístico mediante proyectos o tareas relacionadas con la investigación.

Everson M., Zieffler A., Garfiel J.(2008) proporcionan dos ejemplos de cursos de introducción a la estadística que fueron diseñados aplicando estas directrices. Las propuestas educativas para desarrollar en los estudiantes el razonamiento estadístico de Garfield J., Ben-Zvi D.(2009) tienen también en cuenta las recomendaciones GAISE.



## CAPÍTULO 3. EL RAZONAMIENTO ESTADÍSTICO Y SU EVALUACIÓN

En este capítulo se analizan los aspectos y objetivos de la evaluación que tienen mayor interés para promover el aprendizaje razonado de la estadística. Se revisan y recogen los instrumentos existentes que evalúan el razonamiento estadístico y que han sido utilizados en nuestro estudio.

### 3.1. Desarrollar el razonamiento estadístico

Muchos investigadores (Moore, 2004; delMas, 2004; Pfannkuch y Wild, 2004; Pfannkuch, 2008; Casey, 2008, 2010) han puesto de manifiesto que un correcto proceso de instrucción de la estadística debe tener en cuenta los aspectos específicos del pensamiento y del razonamiento estadístico<sup>7</sup>.

Pfannkuch y Wild (2004, p. 21) indican que “El pensamiento estadístico impregna la forma de operar y funcionar en la vida cotidiana”, y desde este punto de vista el docente cobra máxima importancia en el desarrollo de razonamiento estadístico en los alumnos. Dicho razonamiento no se desarrolla de forma espontánea, tal como indica el análisis de los errores que comenten los alumnos en cuestiones estadísticas básicas. No podemos esperar que se desarrolle el razonamiento estadístico sin un adecuado proceso de enseñanza-aprendizaje. Si consideramos que el pensamiento-razonamiento estadístico es la piedra angular de la enseñanza de la estadística (Wild y Pfannkuch, 1999; Pfannkuch y Wild, 2004; Franklin et al. 2007), la función del docente debe situarse bajo la premisa de contribuir al desarrollo de esta capacidad en el alumnado.

Moore (1997) presenta la siguiente lista de elementos del pensamiento estadístico aprobada por el Consejo de la Asociación Americana de Estadística (ASA) y la Asociación Matemática Americana: la necesidad de los datos, la importancia de la producción de datos, la omnipresencia de la variabilidad, la medición y modelización de la variabilidad. Sin embargo, esto es sólo una parte de lo que los estadísticos entienden por "pensar estadísticamente".

---

<sup>7</sup> Hay que tener en cuenta, tal como se ha visto anteriormente, que para algunos autores ambos conceptos se solapan mientras que, para otros, el pensamiento estadístico, situado en un nivel mayor de competencias, engloba el razonamiento estadístico. Por ello, entendemos que aquellos estudios sobre pensamiento estadístico son también de interés para el presente trabajo.

Snee (1993) y Wild C.J., M. Pfannkuch (1999), fundamentan el desarrollo del pensamiento estadístico en los siguientes aspectos que deben ser adquiridos y desarrollados por los alumnos:

- La necesidad de utilizar datos para abordar la solución de problemas y que el aprendizaje se realice en el ámbito de un contexto, así como la importancia del proceso de generación de datos. El proceso de pensamiento estadístico ocurre ya en el proceso de obtención de datos, y la forma de analizarlos está íntimamente ligada a cómo han sido obtenidos.
  - No se debe comenzar un curso de estadística partiendo ya de unos datos. En la obtención de datos hay un proceso de pensamiento estadístico y obtenerlos supone un avance en la solución del problema. No obstante, el aprendizaje es mucho más que la recopilación de información. Se trata de sintetizar las nuevas ideas con las preexistentes para conseguir una mejora de la comprensión.
  - Tanto la recopilación como el análisis de datos se encuentran en el corazón del pensamiento estadístico. La recopilación de datos promueve el aprendizaje por la experiencia y conecta el proceso de aprendizaje con la realidad, Joiner (1990).
- Tener en cuenta la variabilidad y la incertidumbre como elementos clave. Percibir de manera natural la omnipresencia de la variabilidad e incluirla como un componente importante en el entendimiento de la realidad. Cuando recogemos datos de un sistema, la variabilidad real se complementa con la variabilidad añadida en diferentes sentidos por el propio proceso de recolección de datos.
- Involucrar en la lógica de pensamiento las ideas de validez externa (representatividad) y de validez interna (control de factores de confusión). Comprender y apropiarse de los argumentos que justifican la utilidad de la estimación por medio de muestras, a pesar de que los resultados pueden variar de una muestra a otra.
- Desarrollar la capacidad de abordar problemas faltos de estructura, como lo son la mayoría de las situaciones reales. Realizar preguntas inteligentes para colocar una situación en un contexto que la convierta en un problema estadístico.

- Comprender que la significación estadística está relacionada con la medida de la variabilidad del azar, pero que no está en relación con la significación práctica.
- Valorar la utilidad de la estadística, su alcance y sus limitaciones para comparar, predecir, estimar y valorar el impacto de un factor sobre la variabilidad de otros; para construir indicadores y decidir entre diferentes opciones.
- Desarrollar la habilidad de comunicar los resultados, teniendo en cuenta la capacidad y las condiciones para poder generalizarlos.

Derry, S., Levin, J. y Schauble, L. (1995) realizan la propuesta de un curso cuyo objetivo es el desarrollo del razonamiento estadístico en el que los conocimientos aprendidos sirvan para razonar estadísticamente en escenarios auténticos, argumentando que este aprendizaje debe hacerse en un contexto de la vida real y que las capacidades de los estudiantes pueden mejorar gracias a:

Una instrucción que utilice ejemplos, ilustraciones, analogías, debates y demostraciones que sean relevantes en las culturas a las que los estudiantes pertenecen o esperan pertenecer (relevancia cultural).

Una participación tutelada en un contexto social y colaborativo de solución de problemas, con ayuda de mediadores como grupos de discusión, debate, juego de roles y el descubrimiento guiado (actividad social). (p. 54)

El núcleo de la propuesta de estos autores es un modelo de enseñanza que corresponde con los niveles 5 y 6 de los distintos enfoques de enseñanza que ellos estructuran y que resumimos en la Figura 9.

ENFOQUES FORMATIVOS EN LA ENSEÑANZA DE LA ESTADÍSTICA
<p><b>1. Enseñanza descontextualizada</b> Instrucción centrada en el profesor quien, básicamente, transmite las reglas y fórmulas para el cálculo estadístico. Los ejemplos que emplea son irrelevantes culturalmente y los alumnos manifiestan una pasividad social (receptividad) que se asocia con este enfoque tradicional, donde se suelen proporcionar lecturas abstractas y descontextualizadas (el manual de fórmulas y procedimientos estadísticos).</p> <p><b>2. Análisis colaborativo de datos inventados</b> El alumno “hace”, en vez de sólo “recibir”. Realiza ejercicios donde se aplican fórmulas o se trabaja con paquetes estadísticos informáticos sobre datos hipotéticos, se analizan preguntas de investigación o se decide sobre la pertinencia de pruebas estadísticas. El contenido y los datos son ajenos a los intereses de los alumnos.</p>

- 3. Clases con ejemplos relevantes**

El profesor refuerza los conceptos estadísticos abstractos y procedimientos, proporcionando contenidos significativos de actualidad con los que los estudiantes pueden relacionar dichos conceptos y procedimientos estadísticos más importantes.
- 4. Análisis colaborativo de datos relevantes**

Se centra en el estudiante y en la vida real. Busca inducir el razonamiento estadístico mediante la discusión crítica.
- 5. Simulaciones de la realidad**

Los alumnos se involucran colaborativamente en la resolución de problemas simulados o casos tomados de la vida real (investigación médica, encuestas de opinión, investigación social) con la intención de que desarrollen el razonamiento sobre los conceptos estadísticos importantes.
- 6. Aprendizaje in situ**

Modelo cognitivo in situ que adopta el modelo de aprendizaje que busca desarrollar habilidades y conocimientos propios de la profesión, así como la participación en la solución de problemas sociales o de la comunidad a la que pertenece. Enfatiza la utilidad o funcionalidad de lo aprendido y el aprendizaje en escenarios reales.

Figura 9. Enfoques formativos en la enseñanza de la estadística.

Fuente: Adaptado de Derry, S., Levin, J. y Schauble, L. (1995). Stimulating statistical thinking through situated simulations. *Teaching of Psychology*. Vol. 22, nº1, (1995), p. 51-57.

El modelo propuesto involucra de forma activa a los estudiantes en su aprendizaje mediante la utilización de problemas reales que el alumno ha de resolver. Pone especial énfasis en la utilidad de lo aprendido en el ámbito de la profesión o del entorno social al que pertenece.

### 3.2. La literatura acerca del razonamiento estadístico de los estudiantes

En este apartado se presenta una revisión de las investigaciones que han estudiado el razonamiento y la comprensión conceptual de los alumnos sobre aspectos concretos de la Estadística. Estos estudios nos parecen relevantes para entender cómo razonan los estudiantes.

Se han seleccionado, fundamentalmente, estudios realizados con alumnos universitarios que realizan un primer curso básico de estadística. También se han incluido algunos trabajos realizados con estudiantes de otros niveles educativos que nos parecen interesantes para nuestro estudio.

Se han revisado las principales revistas internacionales de educación y enseñanza de la estadística con alto índice de impacto:



- **International Statistical Review / Revue Internationale de Statistique *Int Stat Rev***

Publicada por John Wiley and Sons. ISSN: 0306-7734.

International Statistical Review (ISR) es la revista insignia del International Statistical Institute (Bernoulli Society for Mathematical Statistics and Probability, International Association for Official Statistics, International Association for Statistical Computing, International Association for Statistical Education, International Association of Survey Statisticians y la International Society for Business and Industrial Statistics). Su objetivo principal es la publicación de artículos originales, revisiones y tutoriales.

- **Journal for Research in Mathematics Education *J Res Math Educ***

Publicada por el National Council of Teachers of Mathematics. ISSN: 0021-8251.

Journal for Research in Mathematics Education es una revista oficial del Consejo Nacional de Profesores de Matemáticas dedicada a los intereses de los profesores de matemáticas y la educación matemática en todos los niveles, desde preescolar hasta adultos. JRME es un foro para la investigación disciplinada en la enseñanza y el aprendizaje de las matemáticas. Los editores animan a la presentación de artículos variados tales como informes de investigación (incluidos los experimentos), estudios de caso, encuestas, estudios de filosofía y estudios históricos; artículos sobre investigación, incluida la revisión de la literatura y análisis teóricos; breves informes de investigación; críticas de artículos y libros, y breves comentarios sobre temas relacionados con la investigación. JRME se publica cinco veces al año.

- **Journal of Statistics Education *J Stat Educ***

Publicada por la American Statistical Association. ISSN: 1069-1898.

Publica artículos sobre temas relacionados con la enseñanza de la estadística, incluyendo estudios de casos, análisis de las nuevas tecnologías y artículos de opinión.

- **Statistics Education Research Journal *Stat Educ Res J***

Publicada por la International Association for Statistical Education. ISSN: 1570-1824.

SERJ es una revista electrónica de la Asociación Internacional para la Educación Estadística (IASE) y el Instituto Internacional de Estadística (ISI). SERJ se publica dos veces al año y es gratis. Pretende avanzar en el conocimiento basado en la investigación que puede ayudar a mejorar la enseñanza, el aprendizaje y la comprensión de la estadística en todos los niveles educativos y en contextos formales (en el aula) e informales (fuera del aula). Este tipo de investigación puede examinar, por ejemplo,

factores y procesos cognitivos, motivacionales, actitudinales, curriculares, relacionados con la enseñanza, con la tecnología, de organización o de la sociedad relacionados con el desarrollo y la comprensión del conocimiento estocástico. Además, las investigaciones pueden estar centradas en cómo las personas usan o aplican la información y las ideas estadísticas y probabilísticas. Esta publicación alienta, por una parte, la presentación de trabajos de calidad relacionados con los objetivos anteriores y, por otra, los informes de investigaciones originales (tanto cuantitativos como cualitativos), revisiones integradoras y críticos de la literatura de investigación, análisis de los modelos teóricos y metodológicos basados en la investigación, y otros tipos de documentos se descritos en detalle en la Guía para Autores. Todos los trabajos son evaluados internamente por el editor asociado o por el editor, y se someten al menos a dos revisiones externas. Se recomiendan las propuestas en Inglés, aunque también se consideran en francés y español.

- **Teaching Statistics *Teach Stat***

Publicada por John Wiley and Sons. ISSN (edición impresa): 0141-982X. ISSN (edición electrónica): 1467-9639.

Teaching Statistics es una publicación que está dirigida a profesores y estudiantes de edades comprendidas entre los 9 y los 19 años. Esta revista pone el énfasis en la enseñanza de la asignatura y en cómo hacer frente a los problemas que se plantean en el aula. La revista tiene como objetivo apoyar no sólo a los profesores especializados en estadística, sino también los de otras disciplinas, como la biología y la geografía económica que hacen amplio uso de la estadística en su enseñanza. Teaching Statistics busca informar, aclarar, estimular, entretener y alentar de manera correcta. Las propuestas deben ser de fácil lectura.

Como consecuencia de esta revisión, se observa que aparecen numerosas publicaciones sobre razonamiento estadístico con la llegada del nuevo siglo. Estas publicaciones están realizadas por investigadores de la Universidad de Minnesota junto con algunos autores de otros países. Las citas revelan la presencia de un “colegio invisible”, término acuñado por Price en 1973 y que se refiere, tal y como ocurre en este caso, a un grupo de investigadores que intercambian información, dedicándose durante un espacio de tiempo a producir y difundir conocimiento científico (Kogan 2002), en este caso sobre el razonamiento estadístico en los estudiantes. La mayor parte de la producción científica centrada en este tema, aparece bajo la rúbrica de estos autores que, en

muchos casos, firman de forma conjunta formando una red en la que destacan los investigadores Joan Garfield J. y Robert C. DelMas.

En España, un grupo de investigación sobre la educación estadística con sede en la Universidad de Granada, realiza diversas publicaciones y se une al grupo de Minnesota. Aunque la línea de investigación del grupo de Minnesota se centra en el razonamiento estadístico en el ámbito universitario, el grupo de la Universidad de Granada focaliza su atención, fundamentalmente, en alumnos de enseñanza primaria y secundaria, así como en el profesorado de estos niveles educativos.

Al margen de esta red de autores mencionada anteriormente, son muy pocos los artículos y trabajos de investigación publicados sobre el razonamiento estadístico de los estudiantes. Durante una década, el interés por el aprendizaje de la estadística dio lugar a artículos de investigación con los que se sentaron las bases conceptuales del razonamiento estadístico, se investigó sobre el razonamiento de los estudiantes, se experimentaron propuestas metodológicas para desarrollar el razonamiento estadístico en los alumnos y se elaboraron instrumentos de evaluación muy interesantes. En los últimos años, sin embargo, la búsqueda bibliográfica pone de manifiesto que el razonamiento estadístico ha perdido relevancia en el panorama de la investigación educativa.

En la Figura 10, se recogen y clasifican los principales trabajos centrados en aspectos concretos del desarrollo del razonamiento estadístico en estudiantes universitarios.

Los temas que se han considerado de mayor interés en el desarrollo de razonamiento estadístico y que han sido estudiados son:

- comprensión de la variabilidad
- asociación de dos variables
- significación estadística
- interpretación de los datos representados gráficamente
- parámetros centrales
- representatividad de las muestras
- distribuciones de datos
- inferencia estadística
- probabilidad
- intervalos de confianza.

Estos aspectos abarcan, prácticamente, todos los conceptos básicos que aborda un curso de introducción de Estadística. Conceptos fundamentales para asentar el conocimiento de la Estadística basado en el desarrollo de la capacidad de entender, interpretar y dar sentido a una información estadística.

Muchos de estos trabajos han utilizado una prueba de evaluación sobre los aspectos concretos que investigan para analizar el razonamiento estadístico que los estudiantes poseen antes de comenzar un primer curso de Estadística. Estas investigaciones analizan los logros alcanzados y los errores que permanecen una vez que el curso ha finalizado aun cuando, en muchos de ellos, se aplican propuestas metodológicas encaminadas a desarrollar el razonamiento estadístico.

ESTUDIOS SOBRE RAZONAMIENTO ESTADÍSTICO	
Razonamiento estadístico	Olani, A., Hoekstra, R., Harskamp, E., & Van Der Werf, G. (2011). Statistical Reasoning Ability, Self-Efficacy, and Value Beliefs in a Reform Based University Statistics Course. <i>Electronic Journal of Research in Educational Psychology</i> , 9(1), 49-72.
Razonamiento informal acerca de la distribución	Bakker, A., Gravemeijer, K. P. (2004) Learning to reason about distribution. En Dani Ben-Zvi y Joan Garfield (ed.). <i>The challenge of developing statistical literacy, reasoning, and thinkingmore</i> . (pp. 147-168). Dordrecht. Kluwer Academic Publishers.
Razonamiento estadístico acerca de la distribución univariante	Gravemeijer, K. P. E. (Abril, 2000). <i>A rationale for an instructional sequence for analyzing one- and two-dimensional data sets</i> . Artículo presentado en el encuentro anual de la American Educational Research Association, Montreal, Canada.
Pensamiento y razonamiento estadístico en relación con el análisis de datos gráficos, y elaboración de gráficos (utiliza un programa informático)	Fitzallen, N. (2006). A model of students' statistical thinking and reasoning about graphs in an ICT environment. <i>Identities, cultures and learning spaces</i> , 203-210.  Fitzallen, N. (2008). Validation of an assessment instrument developed for eliciting student prior learning in graphing and data analysis. In <i>Proceedings of the 31st Annual Conference of the Mathematics Education Research Group of Australasia</i> . Brisbane: MERGA.

Razonamiento sobre la interpretación de los diagramas de cajas	Maxine Pfannkuch. (2007). Year 11 students' informal inferential reasoning: a case study about the interpretation of box plots. <i>International Electronic Journal</i> , 2(3), 149-167.
Razonamiento acerca de las medidas centrales	Konold, C., Pollatsek, A. (2004) Conceptualizing an average as a stable feature of a noisy process. En Dani Ben-Zvi and Joan Garfield (ed.). <i>The challenge of developing statistical literacy, reasoning, and thinking more</i> . (pp. 169-199). Dordrecht. Kluwer Academic Publishers
Comprensión de la variación	Watson, J. M., & Kelly, B. A. (2007). Assessment of students' understanding of variation. <i>Teaching Statistics</i> , 29(3), 80-88.
Razonamiento sobre la variabilidad	Slauson, L. V. (2007). <i>Students' conceptual understanding of variability</i> . Tesis doctoral. Universidad del Estado de Ohio.
Razonamiento acerca de la variabilidad	Garfield J., Ben-Zvi D-. (2005) A framework for teaching and assessing reasoning about variability. <i>Statistics Education Research Journal</i> , 4(1), 92-99.
Razonamiento y comprensión de la desviación estándar	delMas, R. O. B. E. R. T., & Liu, Y. (2005). Exploring students' conceptions of the standard deviation. <i>Statistics Education Research Journal</i> , 4(1), 55-82.
Razonamiento sobre asociación de dos variables. Covarianza.	Zieffler, A. S. (2006). <i>A longitudinal investigation of the development of college students' reasoning about bivariate data during an introductory statistics course</i> . Tesis doctoral. Universidad de Minnesota.
Razonamiento estadístico sobre datos bivariados (asociación entre dos variables)	Zieffler, A. S., & Garfield, J. B. (2009). Modeling the growth of students' covariational reasoning during an introductory statistics course. <i>Statistics Education Research Journal</i> , 8(1), 7-31.
Razonamiento sobre distribuciones de muestreo.	delMas, R. C., Garfield, J., & Chance, B. (1999). A model of classroom research in action: Developing simulation activities to improve students' statistical reasoning. <i>Journal of Statistics Education</i> , 7(3). <a href="http://www.amstat.org/publications/jse/secure/v7n3/delmas.cfm">http://www.amstat.org/publications/jse/secure/v7n3/delmas.cfm</a>
Uso de representaciones gráficas para desarrollar el razonamiento acerca de la inferencia estadística, la muestra, la población, la distribución, el muestreo y la	

<p>variabilidad en el muestreo</p> <p>Razonamiento sobre inferencia estadística</p> <p>Razonamiento sobre la inferencia estadística</p> <p>Razonamiento inferencial informal</p>	<p>Zieffler, A., Garfield, J., Delmas, R., &amp; Reading, C. (2008). A framework to support research on informal inferential reasoning. <i>Statistics Education Research Journal</i>, 7(2), 40-58.</p> <p>Castro Sotos, A. E., Vanhoof, S., Van den Noortgate, W., &amp; Onghena, P. (2007). Students' misconceptions of statistical inference: A review of the empirical evidence from research on statistics education. <i>Educational Research Review</i>, 2(2), 98-113.</p> <p>Rossmann, A. (2008). Reasoning about informal statistical inference: One statistician's view. <i>Statistics Education Research Journal</i>, 7(2), 5-19.</p>
<p>Sobre conceptos erróneos en el razonamiento de la probabilidad</p> <p>Ideas erróneas en los estudiantes acerca de la probabilidad</p> <p>Ideas erróneas acerca de la probabilidad</p>	<p>Garfield, J., &amp; Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. <i>Journal for research in Mathematics Education</i>, 44-63.</p> <p>Hirsch, L. S., &amp; O'Donnell, A. M. (2001). Representativeness in statistical reasoning: Identifying and assessing misconceptions. <i>Journal of Statistics Education</i>, 9(2), 1-22.</p> <p>Shaughnessy, J. M. (1992). Research in Probability and Statistics: Reflections and Directions, en <i>Handbook of Research on Mathematics Teaching and Learning</i>, ed. D. Grouws, New York: McMillan, pp. 465-494.</p>
<p>Razonamiento sobre la significación estadística y el p-valor</p> <p>Razonamiento estadístico sobre los intervalos de confianza</p> <p>Razonamiento sobre distribuciones muestrales</p>	<p>Lane-Getaz, S. J. (2007). <i>Development and Validation of a Research-based Assessment: Reasoning about P-values and Statistical Significance</i> (Doctoral dissertation, University of Minnesota).</p> <p>Lane-Getaz, S. J. (2013). Development of a reliable measure of students' inferential reasoning ability. <i>Statistics Association for Statistical Education</i>, 12(1), 20-47.</p> <p>Cumming, G. (2007). Inference by eye: Pictures of confidence intervals and thinking about levels of confidence. <i>Teaching Statistics</i>, 29(3), 89-93.</p> <p>Chance, B., Mas, R., &amp; Garfield, J. (2005). Reasoning about Sampling Distributions. <i>The challenge of</i></p>

	<i>developing statistical literacy, reasoning and thinking</i> , 295-323.
Conceptos erróneos que tienen los estudiantes sobre la media, la mediana y la variabilidad de datos representados en un histograma o en un gráfico de tallo y hojas	Cooper, L., & Shore, F. (2008). Students' misconceptions in interpreting center and variability of data represented via histograms and stem-and-leaf plots. <i>Journal of Statistics Education</i> , 16(2), n2. <a href="http://www.amstat.org/publications/jse/V16n2/cooper.pdf">http://www.amstat.org/publications/jse/V16n2/cooper.pdf</a>
Mejora de los resultados de aprendizaje de la estadística en alumnos universitarios	Garfield, J. (1995). How students learn statistics. <i>International Statistical Review/Revue Internationale de Statistique</i> , 25-34.
Pruebas de comprensión de la potencia estadística mediante gráficos	Garfield, J., and delMas, R. (1994). Students' Informal and Formal Understanding of Statistical Power. Artículo presentado en la 4ª Conferencia Internacional sobre la Enseñanza de la Estadística (ICOTS4), celebrada en Julio de 1994, en Marrakech, Marruecos.

Figura 10. Revisión bibliográfica. Estudios que tratan sobre aspectos del razonamiento estadístico.

Esta revisión puede ser de utilidad para el profesor de estadística de cara a la selección de actividades formativas con sus alumnos en un curso básico de estadística.

### 3.3. Evaluación como aprendizaje y para el aprendizaje

“Evaluación **del** aprendizaje, evaluación **para** el aprendizaje y evaluación **como** aprendizaje” (Wester and Northern Canadian Protocol for Colaboration in Education 2006, citado en Garfield y col. 2011).

De esta acertada frase que sintetiza tres objetivos de la evaluación, nos interesan, en el marco de este trabajo, los dos últimos: evaluación para el aprendizaje y evaluación como aprendizaje. Black y Wiliam (1998) analizaron más de 250 estudios que relacionaban la evaluación y el aprendizaje. Tras dicho análisis concluyeron que el uso intencional de la evaluación en el aula para promover el aprendizaje ayuda a mejorar el rendimiento académico. La evaluación debe servir para que el profesor tome conciencia de los conocimientos y habilidades de los estudiantes en un contexto de aprendizaje determinado. Posteriormente, deberá utilizar esta información como punto de partida para su labor de formación.

La evaluación **para** el aprendizaje está diseñada con el fin de ofrecer a los profesores información útil. Esta información les servirá para modificar y programar las actividades de enseñanza y aprendizaje. Los profesores también pueden utilizar dicha información para agilizar y mejorar el proceso formativo y los recursos que van a participar en él. Los alumnos, a través de la evaluación, también reciben información que les ayudará a avanzar en su aprendizaje.

La evaluación **para** el aprendizaje se desarrolla paralelamente al proceso de aprendizaje. Está diseñada para poder “visualizar” la comprensión de cada alumno. De esta manera, los profesores estarán en disposición de decidir lo que pueden hacer para ayudar a los estudiantes a progresar. Los estudiantes aprenden de forma individual e idiosincrásica, pero, al mismo tiempo, hay patrones predecibles de conexiones e ideas preconcebidas que algunos estudiantes pueden experimentar a medida que avanzan en el proceso formativo. En la evaluación **para** el aprendizaje, los profesores utilizan la evaluación como una herramienta de investigación para averiguar todo lo posible acerca de lo que sus estudiantes saben y pueden hacer, y las confusiones, prejuicios, o lagunas que pudieran tener.

La evaluación **como** aprendizaje es un proceso de desarrollo y apoyo metacognitivo de los estudiantes y se centra en el papel del alumno como conector crítico entre la evaluación y el aprendizaje. Cuando los estudiantes son activos, comprometidos y poseen sentido crítico, dan sentido a la información, la relacionan con sus conocimientos previamente adquiridos, y la utilizan para sus nuevos aprendizajes. Este proceso que regula la metacognición se produce cuando los alumnos “vigilan” su propio aprendizaje. La retroalimentación de este proceso les sirve para hacer ajustes, adaptaciones, e incluso cambios importantes a favor de una mejor comprensión de los conceptos.

La evaluación **como** aprendizaje emerge de la idea de que los procesos de aprendizaje no son sólo una cuestión de transferencia de ideas de alguien que posee conocimiento a otro que no los tiene, sino que se trata de un proceso activo de reestructuración cognitiva que ocurre cuando los individuos interactúan con nuevas ideas. Los estudiantes deben aprender a ser evaluadores críticos que dan sentido a la información, la relacionan con su conocimiento previo y la usan para el nuevo aprendizaje.

La evaluación **como** aprendizaje amplía el papel de los profesores que se convierten en los responsables del diseño de un proceso formativo y una evaluación que permita a los estudiantes autorregular su aprendizaje. Esta es una habilidad compleja y difícil de adquirir que no se desarrolla rápidamente o de forma espontánea y requiere el modelado y la enseñanza por parte del profesor y la práctica por parte del estudiante.

La evaluación **como** aprendizaje se basa en la convicción de que los estudiantes son capaces de ser adaptables, flexibles e independientes durante el



proceso de aprendizaje y la toma de decisiones. Cuando los profesores involucran a los estudiantes y promueven su independencia, les están proporcionando herramientas para llevar a cabo su propio aprendizaje de forma acertada.

A veces pensamos en la evaluación como punto final del proceso de enseñanza-aprendizaje cuando, en realidad, también es parte del mismo proceso (Morales 2010). En opinión de Morales (2010), con la evaluación se consigue algo más que calificar a un alumno, se consiguen objetivos didácticos y el más importante que es cambiar qué y cómo estudia el alumno, ejercitando capacidades que permanecen aunque lo aprendido pueda olvidarse (Morales 1995).

La evaluación es una actividad de aprendizaje aunque no se pretenda que sea así (Morales, 2012).

El alumno, cuando es evaluado, asigna un nivel de dificultad a las cuestiones a las que se enfrenta y autoevalúa su capacidad de unir ideas, asociar conceptos y hacer un ejercicio de razonamiento. La evaluación contribuye, por tanto, a la autorregulación del alumno. Le permite saber qué ha aprendido y de qué forma lo ha aprendido, induciendo a un cambio de su sistema de aprendizaje. En nuestro caso, se pretende que incluya estrategias para un estudio razonado de la estadística y transforme sus aptitudes mentales en competencias académicas (Nuñez y col. 2006).

En McDowell, Wakelin, Montgomery, y King, 2011, citados en Morales 2012, se muestra que un enfoque de evaluación formativa (assessment for learning) con un adecuado feedback, consigue, entre otras cosas, que el alumno se habitúe con más facilidad a un enfoque profundo del estudio. De esta manera, adquirirá la habilidad de verificar, practicar y corregir y se autorregulará mejor.

“Si queremos formar alumnos que estudien bien, que sepan relacionar lo que estudian con otras cosas, que se sientan capaces de seguir aprendiendo, parece obvio que nos interesa habituar a nuestros alumnos a un enfoque profundo de aprendizaje”. (Morales, 2012 p. 7).

### **3.4. La revolución de la evaluación de la estadística**

Próximo al comienzo del nuevo siglo Garfield y Gal (1999), hacen referencia a la denominada "revolución de la evaluación". Dicha "revolución" afectaba a todas las áreas y asignaturas del escenario educativo y surgió como consecuencia de la insatisfacción general existente en el mundo educativo con la forma tradicional de evaluar. Se alegaba que, algunos tipos de evaluación, como por ejemplo los exámenes de opción múltiple, presentaban limitaciones a la hora de proporcionar información suficiente sobre el aprendizaje del estudiante.

Las formas tradicionales de evaluación comenzaron a ser vistas como no alineadas con los objetivos educativos y curriculares del momento y, las puntuaciones numéricas individuales o calificaciones asignadas a los estudiantes, como una manera inadecuada para evaluar la comprensión del estudiante y de promover el aprendizaje.

Parte de ese cambio tenía que ver con pasar de darle importancia a la "prueba" a poner el énfasis en la "evaluación" y en la integración de la evaluación con la enseñanza (Birenbaum y Dochy, 1996).

### **3.5. Evaluación para conseguir un estudio razonado de la estadística.**

La ampliación de los objetivos de aprendizaje en la enseñanza de la Estadística ha provocado que se explore y se desarrollen alternativas de evaluación para ser utilizadas en clase y en la investigación educativa. Estas alternativas estarían basadas en dos principios:

- La evaluación debe reflejar el contenido estadístico más relevante para el aprendizaje de los alumnos.
- La evaluación debe mejorar el aprendizaje de la Estadística y apoyar las buenas prácticas de la enseñanza.

En nuestra opinión, la evaluación es el elemento clave que va a orientar al alumno hacia un estudio de la estadística en el que los conceptos fundamentales han de ser entendidos en sí mismos y en relación a otros conceptos, han de asociarse al fundamento y objetivo de las distintas pruebas estadísticas, han de hacer al alumno pensar, unir ideas y aprender de manera consciente.

La recomendación número 6 del informe GAISE8 propone el uso de la evaluación para mejorar el aprendizaje de la Estadística en el estudiante, además de para valorar dicho aprendizaje. Entendemos que en Estadística es preciso realizar un tipo de evaluación que no sólo mida el progreso continuo durante el curso, sino que ayude a que los alumnos mantengan en el tiempo lo que han aprendido y que sean capaces de aplicarlo en su vida cotidiana y profesional.

La mayor parte de las pruebas de evaluación estadística que tradicionalmente se han utilizado se centran en los procedimientos, las definiciones y las habilidades, en lugar de valorar la comprensión conceptual. En Estadística se siguen utilizando pruebas y exámenes tradicionales, similares a los problemas que aparecen en los libros de texto, que dependen en gran medida de los cálculos numéricos y de la capacidad de recordar información. Y este tipo de evaluación, en el que se repite un modelo de problema, no revela

---

<sup>8</sup> Guidelines for Assessment and Instruction in Statistics Education, 2010.

información acerca de si los estudiantes comprenden, razonan o aplican sus conocimientos estadísticos para resolverlos.

La evaluación del aprendizaje en Estadística basada fundamentalmente en ítems que miden el dominio de habilidades, procedimientos y vocabulario aislados del contexto de un problema, rara vez son capaces de probar si los estudiantes entienden los conceptos estadísticos, si tienen la capacidad de integrar los conocimientos estadísticos para resolver un problema, o para comunicarse utilizando eficazmente el lenguaje estadístico.

Son necesarias nuevas formas de evaluar, no sólo para proporcionar información a los alumnos y profesores, sino también para investigar sobre la enseñanza y el aprendizaje de la estadística, para evaluar la eficacia de los diferentes planes de estudio o los nuevos enfoques pedagógicos y para explorar el desarrollo del razonamiento estadístico. Algunos de estos procedimientos alternativos, que muestran más posibilidades de percibir el rendimiento intelectual complejo del alumno son el portfolio, los proyectos de los estudiantes, y las actividades de laboratorio (Darling-Hammond et al., 1995).

La literatura describe múltiples alternativas a los métodos tradicionales de evaluación en el ámbito de la estadística, Garfield y Chance (2000) sugieren algunas técnicas para evaluar el razonamiento estadístico de los estudiantes:

- Los proyectos individuales y de grupo, que implican presentar un problema, diseñar el estudio, la elección de la muestra, la recogida y análisis de datos y la interpretación de los resultados.
- Los estudios de casos: problemas detallados basados en un contexto real que pueden revelar las estrategias de los estudiantes para resolver el problema y su interpretación. Permiten a los estudiantes reflexionar sobre ejemplos reales de la práctica estadística.
- Mapas conceptuales: representaciones visuales de las conexiones entre los conceptos que los estudiantes pueden construir o completar por la selección de palabras de una lista que puede o no incluir distractores
- Las críticas sobre la información estadística ofrecida en artículos de investigación. los estudiantes razonan y discuten sobre la información proporcionada en un artículo, guiados por preguntas específicas las siguientes: ¿Cuál cree usted que es el propósito de la investigación descrita en este artículo?, ¿Qué método o métodos se utilizaron para responder a la pregunta de investigación?, ¿Qué preguntas le gustaría hacer a los investigadores a entender mejor el estudio?, ¿Hay algún aspecto del estudio que podrían hacer que usted se cuestiona la conclusiones presentadas en el artículo ?. etc.
- Preguntas que contengan gráficos presentados en los medios de comunicación y la literatura de investigación

- Minute papers: Al finalizar la clase, los estudiantes realizan por escrito y de forma anónima, observaciones, comparaciones de conceptos o breves explicaciones de lo que han aprendido.
- Portafolio de trabajo del estudiante. Selección de trabajos del alumno, que se realiza junto con el profesor y que recoge además reflexiones sobre lo que se ha aprendido.
- Examen de opción múltiple diseñado para evaluar el razonamiento correcto e incorrecto de los estudiantes
- Otros métodos informales para utilizar en el aula, como interpretación de datos, explicación de conceptos, interpretación conjunta de dos representaciones gráficas de los mismos datos, etc.

No es suficiente que los alumnos hayan recibido una enseñanza cuyo objetivo sea el desarrollo del razonamiento estadístico, es necesario examinar con cuidado, sobre todo en contextos aplicados, que los estudiantes son capaces de razonar correctamente acerca de la información estadística. De otra forma podrían tener sólo un primer nivel de razonamiento y no ser capaces de tener una comprensión integrada que les permita hacer juicios e interpretaciones correctas (Garfield y Gal, 1999).

Si lo que se pretende es fomentar el razonamiento estadístico, las pruebas de evaluación deben diseñarse desde esta perspectiva. Deben proporcionar la información que permita valorar el conocimiento que debe perdurar en los alumnos, primando la comprensión de los conceptos sobre la memorización de las definiciones y, desde luego, no deben contener preguntas que requieran cálculos o aplicación de fórmulas (Berenson, M. et al. 2008).

Berenson y col. (2008) analizan, aportando razones específicas, los contenidos que han de componer un curso de introducción a la Estadística y cuál es la evaluación más adecuada de los mismos. De acuerdo a las ideas propuestas por Berenson, exponemos algunas cuestiones que consideramos importantes a la hora de diseñar pruebas de evaluación de la Estadística:

- En algunas ocasiones puede que resulte necesario realizar un desarrollo matemático hasta llegar a una fórmula para ayudar al estudiante a entender un concepto, pero no es interesante para la evaluación. Lo importante y lo que hay que evaluar es que entienda el concepto.
- En otros casos, el contenido es fundamental para entender los temas posteriores. Por ejemplo, la probabilidad como base para entender la inferencia. Ante una situación de estas características, no es necesario que el estudiante retenga información sobre todos los detalles acerca de la probabilidad. Basta con que evaluemos la capacidad que tiene de aplicarlo en los temas posteriores.
- Si decidimos evaluar la capacidad del estudiante para aplicar un contenido estadístico en contextos que encontrará en la vida real, habrá que darle argumentos acerca de la importancia de la aplicación de la

Estadística en situaciones reales. Es más importante entender el proceso, identificar y aplicar las técnicas apropiadas a los problemas, en lugar de memorizar los detalles de cada procedimiento.

- Respecto a los conocimientos que consideramos que los estudiantes van a necesitar en cursos posteriores, el alumno debe saber dónde son aplicables. Por ejemplo, las pruebas chi-cuadrado en estudios epidemiológicos, aunque tendríamos que asegurarnos en la evaluación que entienda lo que realmente persigue la prueba. De esta manera, será capaz en un futuro de identificarla bajo una aplicación específica.

Garfield y col. (2011), plantean tres elementos que conforman un triángulo fundamental en la evaluación: el conocimiento, la observación y la interpretación, al tiempo que indican que la mayoría de los docentes se centran en el segundo de estos elementos. Si ponemos estos aspectos frente al concepto de razonamiento estadístico, una evaluación de este razonamiento debería contener tanto la observación, desde el punto de vista de examinar con atención y llegar a comprender la información estadística, como la interpretación con la que dar sentido a la misma.

El profesor debe plantearse de qué manera esta evaluación puede proporcionar al alumno una información útil para mejorar su aprendizaje. Por otro lado, los profesores debemos ofrecer a los estudiantes una información de estas evaluaciones cuya naturaleza y calidad determinarán que esté asegurado el feedback con el alumno. Esta información fortalecerá su motivación e interés para que realice el trabajo que necesita para aprender. También le ayudará a identificar sus fortalezas y debilidades sin que resulte tan importante calificarle. De esta forma, la evaluación será una herramienta efectiva para promover el aprendizaje.

En el estudio de Garfield y col. (2011), se hace referencia a una revisión realizada por varios autores sobre la evaluación formativa en la que concluyen que una acción intencional en la evaluación formativa provoca efectos positivos sin precedentes en el rendimiento de los estudiantes y que estos efectos son más importantes en los alumnos de bajo rendimiento. Esta evaluación facilita al profesor una información que puede serle útil para hacer cambios en la enseñanza sobre la marcha. Jordan (2004) y Mosteller (1989), citados en Garfield (2011) aportan algunos ejemplos de evaluación formativa en la clase de Estadística.

El uso potencial de la evaluación, cuando se incluye como parte del proceso de aprendizaje, hace participar a los estudiantes y la convierte en una evaluación eficaz. Esto requiere una relación coherente y explícita entre los objetivos, los contenidos, las actividades de enseñanza y aprendizaje y la evaluación, tal como expone Biggs en su modelo de alineación constructiva (Prieto, 2004). Enseñanza, aprendizaje y evaluación han de integrarse de manera adecuada con un enfoque en el aprendizaje del estudiante (Holmes 2002).

Es importante que se haga uso de herramientas informáticas para calificar a los estudiantes utilizando los datos de sus propias puntuaciones con el fin de aplicar pruebas estadísticas como herramienta de trabajo en el aula. De esta manera se puede conseguir que la evaluación sea un elemento motivador del aprendizaje que a su vez haga posible que las calificaciones sean más precisas (Kaplan citado en Garfiel 2011). También es una forma de plantear un problema con datos reales que tiene la capacidad de mostrar la importancia práctica de la Estadística y que ayudará al alumno en el proceso de la comprensión y asimilación de los conceptos.

La evaluación de los conceptos estadísticos clave y la forma de evaluarlos es el camino por el que podemos conseguir que nuestros estudiantes sean buenos consumidores y usuarios de la Estadística. Además, estaremos realizando una importante contribución a la formación de ciudadanos informados y de profesionales que tomarán buenas decisiones.

### 3.6. Instrumentos para la evaluación del razonamiento estadístico

Los investigadores en educación estadística han utilizado una gran variedad de métodos para estudiar y evaluar el razonamiento estadístico de los estudiantes universitarios. En este estudio hemos analizado, fundamentalmente, aquellas pruebas que con un enfoque cuantitativo en la medida, evalúan el razonamiento estadístico. Las más relevantes pertenecen a estudios realizados dentro del proyecto (ARTIST) Assessment Resource Tools for Improving Statistical Thinking, proyecto de la National Science Foundation para mejorar la evaluación de la estadística en la educación superior (Garfield J. and DelMas R. 2006, 2010)<sup>9</sup>.

Se trata de instrumentos conocidos y utilizados en diversas investigaciones (por ejemplo Zieffler, 2006, Slauson, 2008), que se encuentran disponibles en la Web<sup>10</sup> de ARTIST Project<sup>11</sup> y también en la página del International Statistical Literacy Project (ISLP)<sup>12</sup>.

En la página web de ARTIST Project se pueden consultar recursos de evaluación para cursos de introducción a la Estadística. Entre dichos recursos se encuentra un “generador de evaluación” que incluye más de 1000 ítems. Su uso

<sup>9</sup> Descripción del proyecto ARTIS y sus objetivos. delMas, R., Ooms, A., Garfield, J., & Chance, B. (2006, July). Assessing students' statistical reasoning. In *Proceedings of the seventh international conference on teaching statistics* (Vol. 2006). University of Minnesota

<sup>10</sup> Descripción de la Web del proyecto ARTIST. Garfield, J., & DelMas, R. (2010). A Web Site That Provides Resources for Assessing Students' Statistical Literacy, Reasoning and Thinking. *Teaching Statistics*, 32(1), 2-7.

<sup>11</sup> <https://apps3.cehd.umn.edu/artist/index.html>

<sup>12</sup> <http://www.stat.auckland.ac.nz/~iase/islp/assess>

online es gratuito y ofrece información acerca del uso de las pruebas al tiempo que recoge y almacena datos con fines de investigación y evaluación.

Las pruebas ARTIST incluyen ítems con alternativas de respuesta, con respuesta abierta e ítems de interpretación. Las áreas de contenido de estas pruebas son: tipos de datos, recopilación de datos, representación de datos, medidas centrales y de dispersión, medidas de posición, comparación de grupos, distribución normal, datos bivariados, regresión lineal, datos categóricos, chi-cuadrado, probabilidad, muestras y muestreo, pruebas de significación, distribuciones, intervalos de confianza y ANOVA de una vía, con ítems que valoran la alfabetización, el razonamiento y el pensamiento estadístico e ítems de cálculo.

No se ofrecen datos concretos de las propiedades psicométricas debido a que cada usuario compone la evaluación con diferentes elementos lo que da lugar a múltiples pruebas por combinación de ítems, no existiendo, por tanto, un único instrumento.

### 3.6.1. Descripción de los principales instrumentos revisados

***Assessment Resource Tools for Improving Statistical Thinking (ARTIST).*** Está compuesto por 11 escalas temáticas que constan de 7 a 15 ítems con opciones múltiples de respuesta que evalúan el razonamiento estadístico para temas específicos. Se cumplimenta online y los profesores pueden obtener informes sobre los resultados de las pruebas de sus alumnos e identificar lo que se ha aprendido con éxito y dónde tienen los estudiantes dificultades para razonar. (delMas, Garfield, Ooms and Chance 2006).

***Comprehensive Assessment of Outcomes in a First Statistics course, (CAOS).*** Prueba diseñada para medir la comprensión conceptual de los estudiantes de un primer curso universitario de estadística. Define las áreas en las que los estudiantes, después de un curso, mostraron mejora en el razonamiento y las que no, así como las ideas erróneas sobre determinados conceptos estadísticos. (DelMas, Garfield, Ooms and Chance, 2007). El test CAOS fue desarrollado durante tres años a través de un proceso en el que se incorporaron ítems existentes de profesores, se crearon otros nuevos para las áreas no cubiertas hasta el momento por los ya existentes, se revisaron ítems, se obtuvieron propuestas por parte de diversos asesores y se realizaron dos extensas evaluaciones de validez de contenido. Los 40 ítems del cuestionario CAOS fueron respondidos por un total de 1470 alumnos de 33 instituciones de enseñanza superior de 21 estados de Norteamérica. El análisis de consistencia interna del instrumento definitivo produjo un alfa de Cronbach de 0.82. Ha sido usado también en otras investigaciones (por ejemplo, Zieffler, 2006; Slauson, 2008).

**Statistical Reasoning Assessment (SRA)** es una prueba para la evaluación de razonamiento estadístico con opciones múltiples de respuesta que consta de 20 ítems. Fue desarrollado por Konold (1990) y Garfield (1991) como parte de un proyecto para evaluar la eficacia sobre el aprendizaje de la estadística de un nuevo plan de estudios en escuelas secundarias de EE.UU. Posteriormente fue modificado y utilizado en diversos estudios. ( Garfield 1996, 1998 2003; Gal and Garfield 1997; Tempelaar, Gijsselaers, and Schim van der Loeff 2006; Olani, A., Hoekstra, R., Harskamp, E., & Van Der Werf, G. 2011).

Los 20 ítems del SRA están diseñados para evaluar tanto el razonamiento estadístico correcto como el incorrecto. Las áreas de razonamiento estadístico correcto que evalúa este instrumento son:

- Interpretar correctamente probabilidades.
- Entender cómo se selecciona un promedio.
- Calcular correctamente una probabilidad, entendiendo las probabilidades como proporciones y utilizando el razonamiento combinatorio.
- Entender la independencia.
- Entender la variabilidad del muestro.
- Distinguir entre correlación y causalidad.
- Interpretar correctamente tablas de doble entrada.
- Entender la importancia de las grandes muestras.

Las áreas de errores estadísticos que evalúa el SRA son:

- Malentendidos con la media (la media es el número más común, no tomar en cuenta los valores extremos al calcular la media, confundir la media con la mediana y comparar grupos en función de sus medias).
- Ideas erróneas en la orientación de los resultados.
- Las muestras buenas tienen que representar un alto porcentaje de la población.
- La ley de los pequeños números.
- Error en la representatividad.
- La correlación implica causalidad.
- La equiprobabilidad del sesgo.
- Los grupos sólo pueden ser comparados si son del mismo tamaño.

La prueba presentó una baja consistencia interna, las correlaciones test-retest fueron 0.70 para las respuestas correctas de razonamiento y 0.75 para los errores.

**Reasoning about p-values and Statistical Significance. (RPASS).** Se han empleado sucesivas versiones de este instrumento que fue desarrollado para



facilitar la investigación y la evaluación en la enseñanza y el aprendizaje de la comprensión conceptual y los malentendidos en relación al p-valor y a la significación estadística (Lane-Getaz 2007).

En su última versión el RPASS-7 (Lane-Getaz 2013), mejora el instrumento como herramienta de investigación para evaluar los efectos de los métodos de enseñanza sobre el razonamiento inferencial de los estudiantes. Consta de 34 ítems y el análisis de fiabilidad resultante con una muestra de 150 alumnos universitarios de estadística, obtuvo un  $\alpha$  de Cronbach de 0.76.

**Quantitative Reasoning Quotient (QRQ).** Se trata de una versión revisada del SRA (Garfiel 1998). Es un instrumento de 40 ítems que se diseñó para evaluar el razonamiento y los errores relacionados con éste mediante un cociente entre el razonamiento correcto y las ideas falsas (Sundre 2003).

Las áreas de razonamiento que evalúa son:

- Interpretar correctamente probabilidades.
- Interpretar correctamente las medidas de tendencia central.
- Calcular probabilidades correctamente.
- Entender la independencia.
- Entender la variabilidad en el muestreo.
- Distinguir entre correlación y causalidad.
- Interpretar correctamente tablas de doble entrada.
- Entender la importancia de las grandes muestras.
- Comprender las fuentes de sesgo y de error.
- Reconocer las características de un buen diseño experimental.

Los conceptos erróneos que detecta el QRQ son:

- Conceptos erróneos relacionados con la media.
- Las buenas muestras tienen que representar un alto porcentaje de la población.
- La ley de los pequeños números.
- Error en la representatividad.
- La correlación implica causalidad.
- Equiprobabilidad del sesgo.
- Los grupos sólo pueden ser comparados si son del mismo tamaño.
- No distinguir la diferencia entre una muestra y una población.
- No considerar y evaluar todos los datos.
- Incapacidad para crear y evaluar porcentajes.
- Sólo efectos de gran tamaño pueden ser considerados significativos.
- Falta de reconocimiento de las fuentes potenciales de sesgo y error.
- Suponer que más decimales indican una mayor precisión.
- Incapacidad para interpretar probabilidades.

El instrumento fue analizado con 1083 estudiantes universitarios de primer curso. Presenta una consistencia interna con un alfa de Cronbach de 0.55.

Entre otros instrumentos de medida del razonamiento estadístico encontrados, citamos el utilizado por Hirsch y O'Donnell (2001). Se trata de un instrumento creado para identificar a los estudiantes que tienen ideas erróneas acerca de la probabilidad y obtener información acerca de cuáles son los errores que se cometen. Consta de 16 ítems de opciones múltiples y algunas preguntas abiertas. De los 16 ítems, 14 de ellos son de opción múltiple y, además, incorpora una pregunta de seguimiento, también de opción múltiple, en la que se pide a los estudiantes que justifiquen su respuesta. Las opciones de respuesta de seguimiento incluyen una explicación del razonamiento correcto así como varias explicaciones que representan los errores más comunes que tienen los estudiantes. Sus respuestas a las dos partes de la pregunta deben ser congruentes. Deben seleccionar la justificación que corresponda a la respuesta que eligieron en la primera parte del ítem.

El instrumento de prueba se evaluó en un estudio de intervención educativa en el que participaron 263 estudiantes de postgrado ( $n = 61$ ) y universitarios ( $n = 202$ ). El alfa de Cronbach fue de 0.84.

***Cuestionario sobre la “comprensión del intervalo de confianza” (Olivo 2008).***

Se trata de un cuestionario aplicado a 252 estudiantes de ingeniería del curso de Probabilidad y Estadística, una vez finalizado el tema de intervalos de confianza. El cuestionario está compuesto por 18 ítems. 12 de opción múltiple y seis de respuesta abierta. El análisis de fiabilidad ofreció un  $\alpha = 0.654$

***Cuestionario acerca de la construcción del significado de la distribución normal (Tauber, 2001).***

En el estudio participan 117 alumnos universitarios de dos cursos consecutivos. En el segundo ellos se había proporcionado a los alumnos unos apuntes complementarios, entre otros, contenido sobre el uso de Statgraphics. El cuestionario cuenta con 21 ítems y una fiabilidad de  $\alpha = 0.766$

***Statistical thinking and reasoning about graphs in an ICT environment (Fitzallen, 2006).***

Se trata de un instrumento para medir el razonamiento estadístico de alumnos de secundaria y primaria acerca de la interpretación de los gráficos. Se incluyen ítems de este instrumento, que a su vez son recopilaciones de trabajos de otros autores. Se considera de interés a pesar de no tratarse de un estudio con alumnos universitarios.

***Students' misconceptions in interpreting center and variability of data represented via histograms and stem-and-leaf plots (Cooper y Shore, 2008).***

Este instrumento se construye con el fin de detectar los errores que tienen los alumnos con respecto a la interpretación de la media, la mediana y la varianza de datos representados en un histograma en un diagrama de tallos y hojas.

La muestra consistió en 186 estudiantes universitarios. El instrumento consta de 4 ítems, dos de opción múltiple en los que se pide a los estudiantes que interpreten los histogramas de datos agrupados y, los dos restantes, en los que se les pide que calculen la media y la mediana a partir de gráficos. Posteriormente, se realizaron entrevistas con los estudiantes para detectar las dificultades y la justificación de los errores cometidos, material con el que se elaboró un patrón de errores.

### **3.6.2. Valoración de los instrumentos disponibles.**

La mayor parte de los instrumentos de evaluación del razonamiento estadístico y los que más se han utilizado en la investigación tienen el formato de pruebas objetivas. Algunos de los instrumentos revisados incluyen también preguntas abiertas, pero fundamentalmente son pruebas constituidas por ítems de opción múltiple o de verdadero y falso.

En algunos casos estos instrumentos incluyen ítems que podrían ser considerados meramente de cálculo más que de evaluación conceptual, aunque resulta difícil considerar si pueden también contribuir o no a la evaluación del razonamiento estadístico.

En general todas estas pruebas tienen en ocasiones coincidencia en incluir ítems para evaluar los mismos aspectos del razonamiento estadístico, y en algunos casos coinciden en la forma de hacerlo. Todas ellas han sido validadas y utilizadas en investigaciones sobre el razonamiento estadístico y no hemos localizado muchos más instrumentos que merecieran ser tenidos en cuenta para la elección de los que se utilizarán en esta investigación.

El hecho de no disponer de instrumentos de medida en castellano, unido a la utilidad de contar con un cuestionario ajustado al contexto de ciencias de salud, nos llevó a orientar el trabajo hacia la consecución del mismo, siendo esta una de las partes fundamentales que conforma la investigación y su objetivo instrumental: Diseñar y validar un instrumento para evaluar el razonamiento estadístico.

Decidimos, seleccionar de entre los instrumentos descritos aquellos que podrían ser utilizados como base para la construcción o adaptación de los instrumentos que serían utilizados en este trabajo.

Valoramos que la pruebas ARTIS<sup>13</sup>, por tratarse de pruebas sobre temas específicos podría ser idóneas para ser utilizadas como evaluación formativa al finalizar los diferentes bloques de contenidos de la asignatura. Las once pruebas temáticas contemplan todos los aspectos de un curso básico de estadística y tienen un número reducido de ítems que evalúan los aspectos más importantes del razonamiento estadístico. Son pruebas cortas que permiten realizar la evaluación en poco tiempo, integrándose dentro de las sesiones lectivas de la asignatura de estadística.

Por otro lado encontramos que CAOS<sup>14</sup> era la prueba de carácter más general en la evaluación de la comprensión de los conceptos estadísticos de todos los instrumentos analizados. Esta prueba de evaluación tiene como referencia los contenidos de un primer curso universitario de estadística de características similares al nuestro. Es además el instrumento más utilizado en la investigación, y cuenta con sucesivas versiones mejoradas. Todo ello nos llevó a elegirlo para ser la base del instrumento que sería utilizado en este estudio para medir, inicialmente y al finalizar la asignatura, el razonamiento estadístico de nuestros alumnos.

Los instrumentos seleccionados se han utilizado eficazmente para evaluar la comprensión estadística en grandes grupos, para comprender y diagnosticar las habilidades de razonamiento estadístico de los estudiantes, y para aportar información de cómo se desarrolla el razonamiento estadístico.

---

<sup>13</sup> Assessment Resource Tools for Improving Statistical Thinking.

<sup>14</sup> Comprehensive Assessment of Outcomes in a First Statistics course.

## ESTUDIO EMPÍRICO

---



## CAPÍTULO 4. METODOLOGÍA

El estudio empírico de esta investigación se inicia en este capítulo enunciando los objetivos y las hipótesis que han guiado este trabajo. A continuación se describe los instrumentos empleados para la obtención de datos, la intervención, las muestras y el procedimiento.

### 4.1. Objetivos de la investigación

Destacamos la pretensión de constatar, las diferencias en el razonamiento estadístico entre dos grupos de alumnos del grado de Enfermería, diferenciados por una intervención que consiste en la realización de pruebas de evaluación pensadas para medir esta capacidad de razonamiento.

El objetivo del estudio empírico es:

**Comprobar si hay un aumento significativo en la comprensión conceptual de la estadística, por parte de estudiantes universitarios de Enfermería, cuando se aplica un sistema de evaluación utilizando instrumentos específicos que miden el razonamiento estadístico.**

Para conseguir este objetivo se proponen los siguientes objetivos específicos:

#### Objetivo 1.

Analizar las mejoras en el razonamiento que muestran los alumnos al finalizar la asignatura, así como también los errores que permanecen.

#### Objetivo 2.

Estudiar si un aumento en la comprensión conceptual de los estudiantes tiene como resultado una mayor capacidad de resolución de problemas.

#### Objetivo 3.

Indagar si la percepción de la competencia estadística de los alumnos sufre variaciones como resultado del proceso de enseñanza-aprendizaje y si también se pueden constatar diferencias significativas en esa percepción cuando se aplica un sistema de evaluación con instrumentos específicos que miden el razonamiento estadístico.

#### Objetivo 4.

Comprobar las relaciones existentes entre las puntuaciones de razonamiento estadístico y otras variables, como la capacidad de resolución de problemas y la percepción de la competencia estadística.

#### 4.2. Objetivos instrumentales

Seleccionar y diseñar instrumentos fiables y válidos que midan los constructos:

- Razonamiento estadístico.
- Percepción de la competencia estadística

Tal como se ha puesto de manifiesto anteriormente en la revisión de instrumentos de evaluación del razonamiento estadístico, todos los que han sido utilizados en investigaciones relevantes están en inglés, y ninguno de ellos se ajusta a nuestro contexto de ciencias de la salud.

Por este motivo se fija como objetivo elaborar unos cuestionarios adaptados al contexto sanitario y a los contenidos de la asignatura de estadística de los planes de estudio del grado de Enfermería, válidos para medir los resultados de aprendizaje desde el punto de vista conceptual y razonado de la estadística.

#### 4.3. Hipótesis

De acuerdo con los objetivos formulados, la hipótesis general que guía esta investigación es:

**La comprensión conceptual estadística y el rendimiento académico de los alumnos se incrementa cuando se utiliza en el proceso de enseñanza-aprendizaje de esta materia, una estrategia de evaluación diseñada para desarrollar el razonamiento estadístico.**

A continuación se enuncian las hipótesis que serán contrastadas.

##### Hipótesis 1.

Los alumnos a los que se les aplica una evaluación formativa con pruebas de evaluación del razonamiento estadístico, muestran una mayor comprensión conceptual de la estadística que el grupo de alumnos que no realiza ese tipo de evaluación.



**Hipótesis 2.**

Los alumnos a los que se les aplica una evaluación formativa con pruebas de evaluación del razonamiento estadístico obtienen mayores calificaciones en la prueba práctica de resolución de problemas que el grupo que no realiza pruebas de razonamiento estadístico.

**Hipótesis 3.**

La percepción de la competencia estadística en los alumnos que han realizado pruebas de evaluación de razonamiento estadístico es mayor que la percepción que tienen los alumnos que no han realizado ese tipo de pruebas.

**Hipótesis 4.**

Existe una correlación significativa entre los resultados del cuestionario de evaluación de razonamiento estadístico y los de la prueba práctica de resolución de problemas.

**Hipótesis 5.**

Las puntuaciones del cuestionario de evaluación del razonamiento estadístico correlacionan de forma significativa con la percepción que tienen los alumnos de su competencia estadística.

**Hipótesis 6.**

La percepción de la competencia estadística que manifiestan los alumnos, correlaciona de forma significativa con las calificaciones del examen práctico de resolución de problemas.

**Hipótesis 7.**

La percepción de la competencia estadística que manifiestan los alumnos, correlaciona significativamente con las puntuaciones de las pruebas de razonamiento estadístico de respuesta abierta.

**4.4. Diseño general del estudio**

Con el fin de alcanzar los objetivos propuestos y contrastar las hipótesis formuladas se desarrolla una investigación cuasi experimental<sup>15</sup> que permita medir la efectividad o eficacia de la intervención.

---

<sup>15</sup> La investigación cuasi experimental proviene del ámbito educativo, donde la investigación de ciertos fenómenos no puede llevarse a cabo siguiendo los procedimientos experimentales (Campbell y Stanley, 1963, 1966).

Se trata de un diseño pretest-postest con grupo control no equivalente. Dada la naturaleza del diseño, se tendrán en cuenta las posibles amenazas a la validez interna, tales como las diferencias iniciales en los grupos.

#### 4.5. Participantes

La muestra para el estudio con los instrumentos definitivos y posterior comprobación de hipótesis estuvo compuesta por 109 alumnos de primer curso del grado de Enfermería de la Escuela Universitaria de Enfermería y Fisioterapia "San Juan de Dios" de la Universidad Pontificia Comillas de Madrid, del curso 2012- 2013.

El número de participantes ( N=109) corresponden a los alumnos de los dos grupos de primer curso, en los que la asignación de los alumnos se realiza por orden alfabético.

Es una muestra de conveniencia en el que dada la dificultad de la intervención, el criterio de accesibilidad nos decidió a utilizar solamente estos sujetos. Es un muestreo común en el área de educación (McMillan & Schumacher, 2001), que el profesor utilice a sus estudiantes para llevar a cabo una investigación educativa (Lodico, Spaulding y Voegtler, 2006). No pueden generalizarse los resultados porque la muestra no es representativa pero sin embargo la muestra provee información útil para responder a preguntas y comprobar hipótesis.

La edad de los alumnos de la muestra está comprendida entre 17 y 33 años. Hay un número alto de alumnos muy jóvenes con 17 y 18 años, siendo la mediana 18 años.

La media de edad de los estudiantes de la muestra es de 19.6 años con una desviación típica de 3.

Son muy pocos los alumnos de mayor edad, siendo 5 de ellos mayores de 25 años.

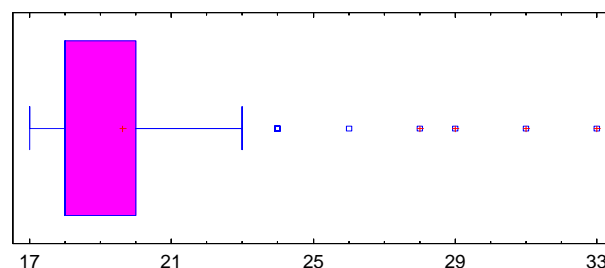


Figura 11. Edad de los componentes de la muestra.

La muestra está constituida mayoritariamente por mujeres que representan un 85,3% de los participantes.

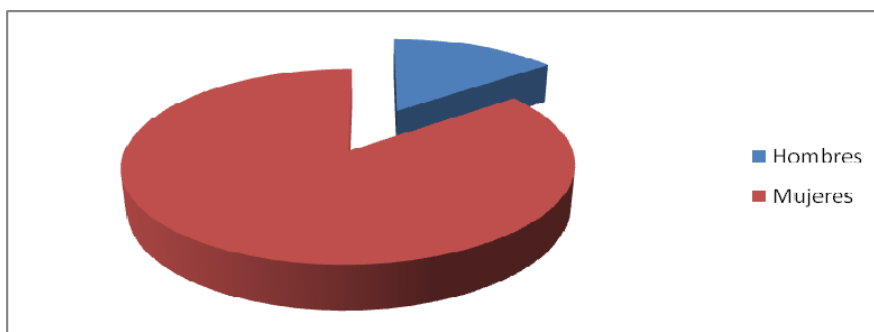


Figura 12. Distribución por sexo de los componentes de la muestra.

En cuanto al procedimiento de acceso a la universidad, la gran mayoría procede de Bachillerato (75,2%) y en segundo lugar los ciclos formativos de grado superior (22%), tan sólo 2,7% provienen de otros títulos universitarios o de las pruebas de acceso para mayores de 25 años.

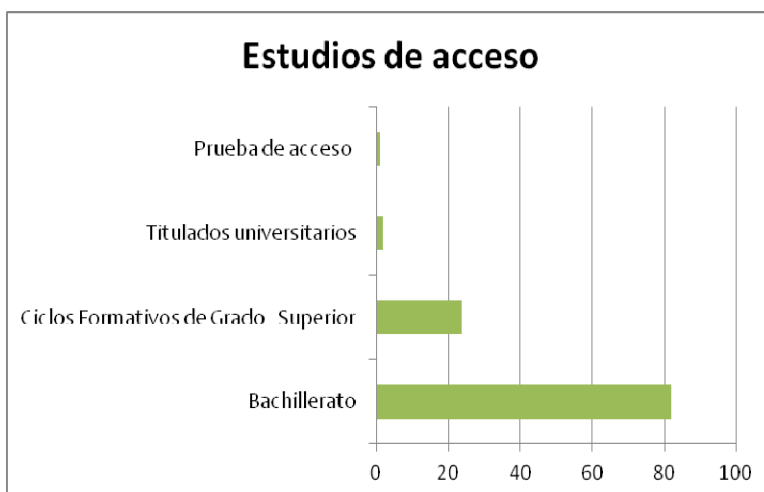


Figura 13. Estudios de acceso de los alumnos de la muestra. (N=109)

La mayor parte de los alumnos (87,2%) cursaron el bachillerato en la modalidad de ciencias y tecnología y sólo un 12,8% lo hizo en humanidades y ciencias sociales.

La nota media de bachillerato de los alumnos de la muestra fue de 6,69 y una desviación típica de 0.80 y la estimación de los propios alumnos de su

calificación en matemáticas en su trayectoria de estudiante tubo una media de 5,89 y una desviación típica de 0,84.

La nota media de los estudios de acceso de los alumnos procedentes de Ciclos Formativos, todos ellos de la familia sanitaria, fue de 6,83 con una desviación típica de 0,89, y la media de la estimación de las calificaciones en la asignatura de matemáticas fue de 5,89 con una desviación típica de 0,93.

En cuanto a las asignaturas de estadística que habían cursado anteriormente, el 58,7 % refiere no haber cursado nunca estadística, el 33%, que corresponderá en su mayor parte con los alumnos procedentes de los ciclos formativos de grado superior, había cursado una asignatura con contenidos de estadística, y sólo tres alumnos tenían esta formación por haber sido cursada en estudios universitarios previos.

Los alumnos de la muestra estaban distribuidos en dos grupos, grupo A y grupo B, con un criterio alfabético respecto al primer apellido, antes del comienzo del curso. La distribución de alumnos con este criterio dio como resultado dos grupos con similares características en cuanto a la edad (Figura 14), sexo (Figura 15), nota de acceso y estudios de procedencia (Figura 16).

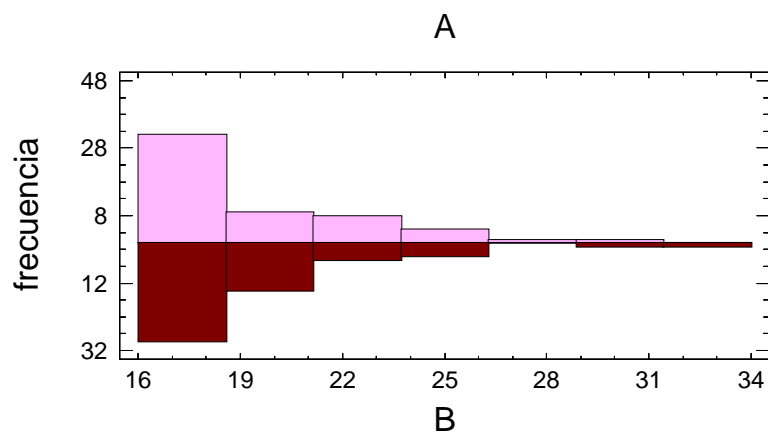


Figura 14. Comparativa de la frecuencia de edad en los grupo A y B.

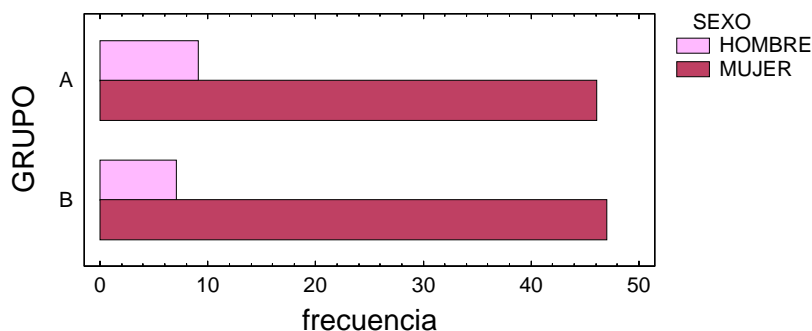


Figura 15. Frecuencia de hombres y mujeres en los grupos de comparación.

El porcentaje de mujeres en ambos grupos es similar (83,64% y 87,04%) y muy superior al de los alumnos de sexo masculino (16,36% en el grupo A y 12,96% en el grupo B).

La nota media de los estudios de acceso de los alumnos del grupo A fue de 6,8 con una desviación típica de 0,81. En el grupo B la nota media de acceso fue de 6,5 y la desviación típica 0,80.

Los estudios de procedencia mayoritarios, bachillerato y ciclos formativos, se reparte con porcentajes similares en los alumnos de ambos (Figura 16)

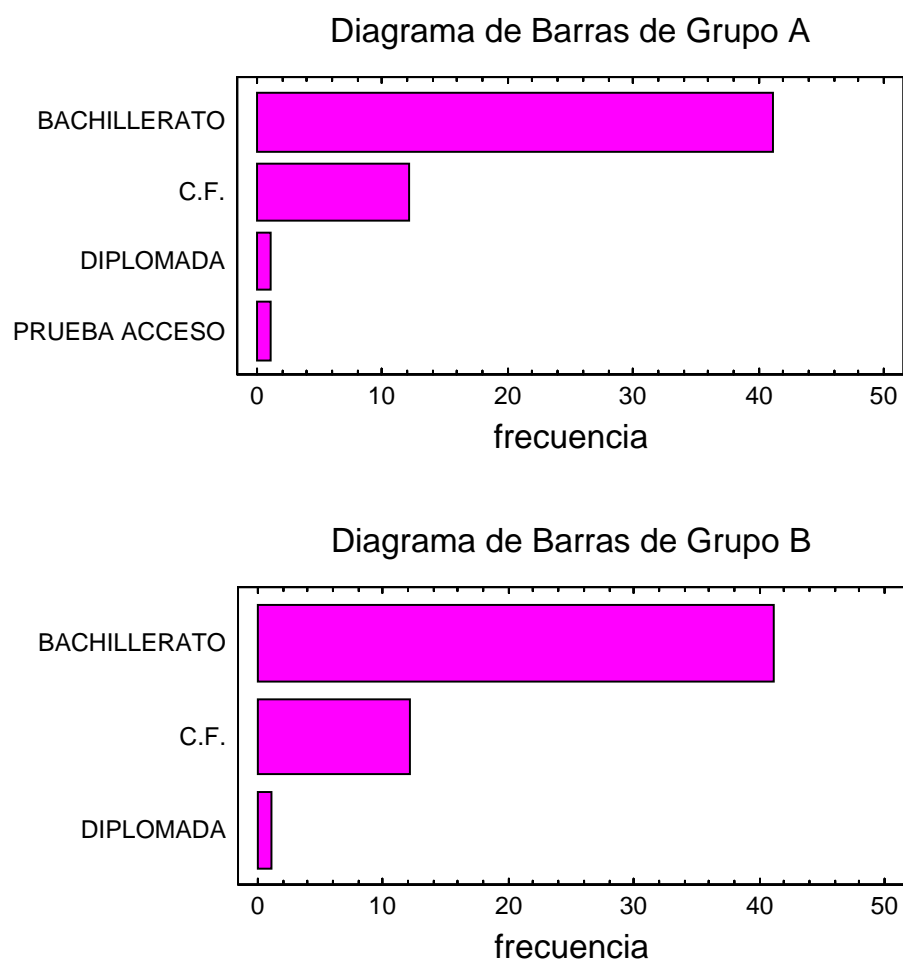


Figura 16. Frecuencias de los estudio de acceso en los grupos de comparación.

La asignación de los alumnos a los grupos de clase se realiza en secretaría por orden alfabético de intervención y control se hizo de forma aleatoria antes de conocer el listado de cada uno de ellos, resultando:

Grupo A: Grupo de intervención

Grupo B: Grupo Control

#### 4.6. Variables e instrumentos

La Figura 17 facilita una visión general de las variables en estudio y los instrumentos empleados para su medida, que posteriormente se expondrán en detalle.

Variables	Instrumentos
Características del alumno	. <b>Cuestionario</b>
Percepción de la competencia estadística	. <b>Escala de percepción de la competencia estadística (EpCE)</b> Factor “competencia percibida hacia la Estadística” de la Escala de Actitudes hacia la Estadística de Muñoz (2002).
Razonamiento estadístico	. <b>Cuestionario para la evaluación del razonamiento estadístico (CERES)</b>  . <b>Pruebas parciales de evaluación del razonamiento estadístico.</b> <b>PaRES 1</b> <b>PaRES 2</b> <b>PaRES 3</b> <b>PaRES 4</b>  . <b>Prueba de evaluación del razonamiento estadístico de respuesta abierta parcialmente restringida</b>
Capacidad de resolución de problemas	. <b>Prueba de resolución de problemas prácticos con un programa informático.</b>

Figura 17. Instrumentos utilizados para la obtención de las medidas de las variables de estudio.

#### 4.6.1. Características del alumno.

Algunas de estas variables, como edad o sexo que sirven para caracterizar la muestra, pueden ser útiles también para separar grupos de comparación si se estima conveniente. Se han incluido también variables que recogen información del tipo de acceso a los estudios universitarios y la nota de acceso.

Las variables acerca de la calificación en la asignatura de matemáticas, así como el haber cursado previamente alguna asignatura de estadística, son incluidas también por su posible utilidad para establecer diferencias entre grupos de alumnos así como su relación con la percepción de la competencia estadística e incluso con los resultados de las pruebas de razonamiento estadístico.

Las variables de este grupo son:

- **Edad**
- **Sexo**
- **Acceso a la Universidad:** Bachillerato (modalidad), Ciclos Formativos (Rama), otros estudios universitarios (titulación) y prueba de acceso.
- **Nota de los estudios de acceso:** nota media de Bachillerato, Ciclos Formativos, otros estudios Universitarios o prueba de acceso.
- **Calificación en matemáticas:** Estimación del alumno de sus calificaciones en matemáticas en su trayectoria de estudiante (de 0 a 10).
- **Experiencia previa estadística:** número de cursos previos de estadística.

Para la recogida de datos se empleó un cuestionario con una serie de preguntas que se administraron junto con la escala de percepción de la competencia estadística. Todas las preguntas eran de respuesta abiertas excepto la correspondiente al sexo, y al acceso a la universidad en el que se facilitan las opciones de respuesta.

#### 4.6.2. Percepción de la competencia estadística.

Esta variable puede aportar datos acerca de si el desarrollo de la comprensión conceptual de la estadística en el alumno, tiene que ver con la percepción de su competencia estadística. Algunos autores lo han relacionado también con el rendimiento académico (Muñoz, 2002).

Para su medición, se ha empleado la Escala de percepción de la competencia estadística (EpCE). Se trata de una escala tipo Likert con siete opciones de respuesta. Este instrumento está compuesto de 12 ítems correspondientes al factor “Competencia percibida hacia la Estadística”, de la Escala de Actitudes hacia la Estadística, de Muñoz (2002). Estos ítems valoran la confianza que tiene el alumno para enfrentarse con la estadística, la dificultad que tiene para él la asignatura, el grado de éxito que percibe que tendrá y la habilidad que creen tener cuando se enfrentan a ella.

La Escala de Actitudes hacia la Estadística (Muñoz, 2002), está compuesta por 5 subescalas, encaminadas a medir, además de la percepción de la competencia estadística otros aspectos como el gusto por la estadística, la utilidad para su trabajo profesional y las dificultades ante la asignatura. Esta escala fue utilizada y validada con alumnos de ciencias sociales (N=269). La fiabilidad del instrumento fue de 0,96. El coeficiente de fiabilidad para la subescala *competencia percibida hacia la estadística* fue de 0,92.

Gran parte de los instrumentos más utilizados para medir actitudes hacia la estadística incluyen una dimensión que valora la competencia cognitiva percibida

Entre estos instrumentos destaca, como uno de los más empleados para evaluar actitudes de los estudiantes a través de diversos escenarios educativos, intervenciones y enfoques de instrucción (Carnell, 2008; Dempster y McCorry, 2009; Carlson y Winqvist, 2011; Posner, 2011; Bond, Perkins y Ramirez 20012; Griffith, 2012), el Survey of Attitudes Toward Statistics (SATS) desarrollado por Schau et al. (1995) y Schau (2003). En la construcción de esta escala Likert con cuatro dimensiones se incluye *la competencia cognitiva* como las actitudes que tiene el alumno hacia los componentes intelectuales asociados a la Estadística, factor que en esta escala consta de 6 ítems.

Otro de los instrumentos que tienen en cuenta ítems relativos a la percepción de la competencia estadística es el Statistics Attitude Survey (SAS) desarrollado por Roberts y Bilderback (1980). Este instrumento “fue construido teniendo en cuenta varios conceptos tales como la competencia percibida en estadística y la utilidad de la estadística, aunque no se supuso que era factorialmente complejo” (Roberts y Reese, 1987, p. 759, citado por Carmona 2004).



La competencia cognitiva se incluye también en estudios que ponen a prueba un modelo de relación entre las actitudes hacia la estadística y el rendimiento (Hood, Creed y Neumann, 2012), siendo la competencia cognitiva la que realiza la mayor contribución indirecta al modelo.

La percepción de la competencia para entender la estadística se encuentra como dimensión en las siguientes escalas de actitudes ante la estadística (Nolan, Beran y Hecker, 2012):

- Attitudes Toward Statistics (ATS), (Cashin y Elmore, 2005)
- Multi-factorial Scale of Attitudes Toward Statistics (MSAS), (Auzmendi, 1991)
- Students' Attitudes Toward Statistics (STATS-A), (Sutarso, 1992)
- Survey of Attitudes Toward Statistics-28 (SATS-28), (Schau, Stevens, Dauphinee y Del Vecchio, 1995).
- Survey of Attitudes Toward Statistics-36 (SATS-36), (VanHoof, Kuppens, Sotos, Verschaffel y Onghena, 2011)
- Bad Attitudes Toward Statistics (BATS), (Berk y Nanda, 1998)
- Students' Attitudes Toward Statistics and Technology Scale (SASTSc), (Anastasiadou, 2011)

Blanco (2008) en una revisión crítica de la investigación sobre actitudes hacia la estadística encuentra que una de las dimensiones más contempladas en los instrumentos que miden actitudes hacia la estadística es la autopercepción de competencia o confianza en relación con el aprendizaje de la estadística. Pone de manifiesto también que la mayor parte de los estudios han analizado estas actitudes con el interés de observar su relación con el rendimiento en la materia (Gal, Ginsburg y Schau, 1997; Cashin y Elmore, 1997; Carmona, 2004, citados en Blanco 2008) . En concreto algunos estudios ( Geske, Mickelson, Bandalos, Jonson y Smith, 2000; citados en Blanco 2008) encuentran correlaciones entre la nota final del examen y la competencia cognitiva de 0,33.

Varias son las investigaciones que han demostrado que las actitudes hacia la estadística se relacionan positivamente con el rendimiento de los estudiantes en la asignatura, lo que indica que las actitudes más positivas se relacionan con mayores logros ( Chiesi y Primi, 2009; Dempster y McCorry, 2009; Limpscomb, Hotard, Shelley y Baldwin, 2002; Sorge y Schau, 2002; citados en Emmioglu y Capa-Aydin 2012)

Emmioglu y Capa-Aydin (2012) en un metaanálisis sobre las actitudes y rendimiento en estadística, analizan en 17 estudios de investigación, las

correlaciones entre las distintas dimensiones actitudinales que contempla la escala SATS y el rendimiento en estadística, encontrando una correlación positiva (desde  $r=0,12$  a  $r=0,60$ ) en 15 de estos estudios. En general, los estudios en los que los participantes mostraron mayor competencia cognitiva de la estadística, mostraron también mayor rendimiento en esta materia.

Por lo anteriormente expuesto, decidimos utilizar la escala de “Competencia percibida hacia la Estadística”, factor, como hemos comentado, de la Escala de Actitudes hacia la Estadística, de Muñoz (2002). La finalidad de esta escala es analizar su posible relación con el Cuestionario de Evaluación del Razonamiento Estadístico (CERES), partiendo del supuesto que ambas medidas han de tener cierta correlación y probar si esta percepción aumenta en los estudiantes a medida que avanzan en el camino del razonamiento estadístico. En este estudio de Muñoz se evidencia una correlación significativa entre el rendimiento (nota final de la asignatura) con la percepción de la competencia estadística.

Por otro lado, en la medida en que el alumno pueda comprobar que realiza con éxito, de forma repetida, actividades de cierta dificultad como el hecho de enfrentarse a un estudio y a unas pruebas de evaluación que le obliguen a pensar y razonar, puede modificar la percepción de su competencia estadística. Se obtendrán datos de esta percepción, a través de esta escala, al comienzo y al finalizar el curso para observar estos posibles cambios, comparando el grupo de alumnos que realiza pruebas parciales de evaluación de razonamiento estadístico y el grupo que no las realiza.

La Escala de percepción de la competencia estadística es una escala Likert con 7 opciones de respuesta para expresar el grado de acuerdo con las afirmaciones planteadas en los ítems que se recogen en la Figura 18. Las opciones son las siguientes:

Totalmente de acuerdo	Muy de acuerdo	De acuerdo	Indiferente	En desacuerdo	Muy en desacuerdo	Totalmente en desacuerdo
-----------------------	----------------	------------	-------------	---------------	-------------------	--------------------------

<b>Escala de Percepción de la Competencia Estadística (EpCE)</b>
1. Cuando tengo que resolver problemas empiezo a ponerme nervioso/a.
2. Nunca entiendo lo que se explica en clase de matemáticas o estadística y eso me pone nervioso/a.
3. Considero la asignatura de estadística como un muro difícil de franquear.
4. Considero que mis resultados en esta asignatura van a ser bastante buenos.
5. Tengo facilidad para resolver problemas matemático-estadísticos.
6. Me cuesta mucho resolver los problemas matemático-estadísticos aunque me los expliquen.
7. No se me dan nada bien las materias relacionadas con las matemáticas como lo es la estadística.
8. Siempre he tenido mucha dificultad con las matemáticas y la estadística.
9. Necesito que me expliquen varias veces los conceptos y métodos matemático-estadísticos para que los entienda.
10. No me cuesta trabajo comprender materias matemático-estadísticas.
11. Normalmente entiendo sin problemas los apuntes cogidos en matemáticas o estadística sin que nadie tenga que explicármelos de nuevo.
12. En los trabajos de clase siempre evito realizar cualquier tipo de análisis matemático-estadístico.

Figura 18. **Formulación de los ítems que componen la Escala de Percepción de la competencia Estadística (EpCE).**

Esta escala fue respondida de forma online, elaborándose mediante la herramienta “Docs” de Google e insertando el enlace de acceso vía web en la plataforma de la Universidad.

La utilización de esta escala tuvo como objetivos:

- Estudiar el funcionamiento de la escala y usarla posteriormente para recoger datos acerca de los niveles de percepción de la Competencia Estadística de los alumnos de Enfermería de la muestra.

- Estudiar su correlación con los instrumentos de evaluación diseñados.
- Observar en los alumnos posibles cambios de esta percepción.

Para el análisis de esta escala se contó con una muestra piloto de 83 alumnos correspondientes al curso 2011-2012. La escala fue cumplimentada por los alumnos al finalizar la asignatura.

La Escala de Percepción de la Competencia Estadística (EPCE), para la muestra piloto tuvo una fiabilidad medida como consistencia interna de  $\alpha = 0,930$ .

Se obtuvo la correlación de cada ítem con la suma de puntuaciones de los restantes ítems de la EPCE y el coeficiente de fiabilidad (alpha de Cronbach) si fuera eliminado el ítem. Resultando todas las correlaciones significativas y mayores a 0,55 a excepción del ítem 12 ítems (en los trabajos de clase siempre evito realizar cualquier tipo de análisis matemático-estadístico). Este ítems en el análisis de la escala que realiza Muñoz (2002) presenta pesos importantes en varios de los factores.

Respecto al análisis factorial de la EPCE, las pruebas de ajuste muestran el cumplimiento de requisitos previos en los datos para la realización del análisis factorial (KMO = 0,900 y esfericidad Chi-cuadrado = 721,198;  $p < 0,001$ ).

Se utilizó el método de componentes principales para la extracción de factores y la rotación Varimax con Kaiser.

Son dos los factores con un autovalor superior a 1 seleccionados con el procedimiento Guttman-Kaiser. El porcentaje de varianza total explicada por ambos fue de 66,527%.

El primer factor explica el 48,892% y el segundo el 17,635% de la varianza total.

En el siguiente gráfico de sedimentación (Figura 19) se aprecia que el factor 1 es el que explica un porcentaje mayor de la varianza total.

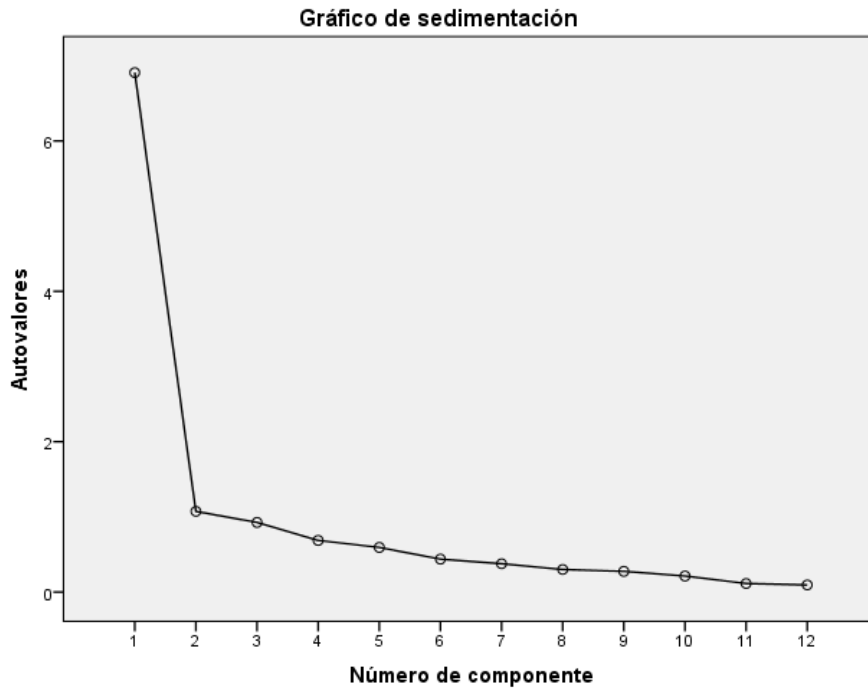


Figura 19. Gráfico de sedimentación del análisis factorial de la escala de percepción de la competencia estadística (EPCE).

A partir de la matriz de componentes rotados (Tabla 3), se observa que los ítems se agrupan básicamente en un factor. El ítem 12 que no correlacionaba con el resto, parece hacerlo con el ítem 4 aunque conceptualmente parece que difiere del resto de ítems.

Los análisis con esta muestra piloto concluyen que disponemos de una escala para medir la percepción de la competencia estadística que consta fundamentalmente de un solo factor.

Tabla 3. Matriz de componentes rotados de la escala de percepción de la competencia estadística (EPCE).

	Componente		% varianza	$\alpha$
	1	2		
ITEM9	,872		48,892%	0,936
ITEM6	,841			
ITEM7	,841			
ITEM8	,808			
ITEM5	,796			
ITEM11	,721	,446		
ITEM10	,707			
ITEM1	,681			
ITEM2	,624	,539		
ITEM3	,573	,483		
ITEM12		,848	17,635%	0,516
ITEM4	,447	,644		
Total			66,527%	0,930

En el curso 2012-13 con la muestra definitiva, respondieron a esta escala 109 alumnos y lo hicieron en dos ocasiones, al comienzo de la asignatura y al finalizar la misma. Se registraron ambos datos de forma asociada, para obtener dos muestras relacionadas y estudiar los posibles cambios en la percepción de esta competencia.

#### 4.6.3. Variables relacionadas con el aprendizaje.

Estas son las variables de mayor interés en este estudio, para las que se elaboran instrumentos específicos y sobre las que se establecen los objetivos fundamentales de la investigación:

- **Razonamiento estadístico:** Comprensión y utilización del lenguaje estadístico. Interpretación de resúmenes y representaciones de datos conectando conceptos y combinando ideas para entender y explicar los

procesos estadísticos e interpretar los resultados. Tal como se exponía anteriormente, el razonamiento estadístico se medirá mediante:

- Cuestionario para la evaluación del razonamiento estadístico (CERES)
- Pruebas parciales de evaluación del razonamiento estadístico: PaRES 1, PaRES 2, PaRES 3 y PaRES 4
- Prueba de evaluación del razonamiento estadístico de respuesta abierta parcialmente restringida

Dada la complejidad del diseño de estos instrumentos, se dedicará todo un capítulo a la exposición del procedimiento de construcción de los mismos.

- **Capacidad de resolución de problemas:** calificación de la prueba de evaluación práctica realizada al finalizar la asignatura. Para esta prueba se facilita al estudiante una base real de datos sanitarios, en la que se omite la información referente a datos de carácter personal, que recoge información de múltiples variables. En relación a esta base de datos se plantean una serie de preguntas con aspectos que van desde información descriptiva a otras cuestiones sobre la posible asociación de variables, sobre diferencias entre grupos, etc. Se pide al alumno que resuelva estas cuestiones utilizando un programa informático.

#### 4.7. Intervención

La intervención consiste fundamentalmente en desarrollar una actividad de evaluación en el transcurso de la asignatura utilizando unas pruebas parciales de razonamiento estadístico (PaRES) construidas para tal fin.

Para la intervención se dispone de cuatro pruebas de razonamiento estadístico, cada una de ellas evalúa unos contenidos específicos por bloques temáticos.

PaRES1. Estadística descriptiva.

PaRES2. Distribución normal.

PaRES3. Pruebas de significación.

PaRES4. Análisis de varianza y regresión.

Los alumnos del grupo de intervención realizaron estas pruebas al finalizar cada uno de los bloques de contenidos que evalúa cada una de ellas. Fueron calificados y recibieron información acerca de sus resultados. Cada una

de las cuatro pruebas realizadas fueron corregidas con los alumnos en una sesión de clase en la que se explicó y fundamentó la respuesta correcta, por el interés de detectar aquellas cuestiones que no habían sido entendidas, devolver esta información al grupo y poder aclarar los errores de razonamiento cometidos por los estudiantes.

Dado que esta tesis se fundamenta principalmente en una intervención de evaluación, tanto la descripción detallada de la misma como la exposición y análisis de las pruebas de razonamiento estadístico utilizadas en dicha intervención se exponen en un capítulo aparte. (capítulo 6.)



## CAPÍTULO 5. CONSTRUCCIÓN DE INSTRUMENTOS DE MEDICIÓN DEL RAZONAMIENTO ESTADÍSTICO

En este capítulo se exponen cuáles son los instrumentos utilizados para el estudio y los pasos que se han seguido para la construcción de los mismos.

### 5.1. Antecedentes y justificación

“A pesar de que los alumnos pueden ser evaluados por medio de trabajos, tales como proyectos, o casos prácticos, existe una necesidad práctica de disponer también de un instrumento fácilmente puntuable que valore el pensamiento, razonamiento y la aplicación del conocimiento, en lugar de demostrarnos lo que recuerdan o que pueden realizar cálculos y procedimientos correctamente” (Garfield, 2003) p. 23.

Uno de los objetivos de esta investigación es elaborar unos cuestionarios que permitan evaluar el razonamiento estadístico en alumnos de Enfermería asegurando su validez y fiabilidad.

Como ya se ha comentado la mayor parte de los instrumentos existentes para la evaluación del razonamiento estadístico y los que más se han utilizado en la investigación tienen el formato de pruebas objetivas.

Por otro lado desde nuestra definición de Razonamiento estadístico, como la forma en la que los alumnos ordenan, combinan las ideas y relacionan los conceptos con coherencia, para entender, interpretar y dar sentido a una información estadística, encontramos válidas este tipo de pruebas que tiene su utilidad (Morales 2010) para comprobar:

- La **comprensión** de conceptos, interpretar e identificar ejemplos.
- La capacidad **de análisis**, la capacidad de diferenciar e identificar relaciones.
- La capacidad de **aplicación** y de resolución de problemas.

El alumno que se enfrenta a una prueba objetiva, al no tener que construir una respuesta, tiene más tiempo para leer la pregunta, pensar sobre ella, poner en funcionamiento una serie de capacidades superiores y entrenar “hábitos de la mente” para el desarrollo del razonamiento estadístico. De esta

forma, el enunciado de las preguntas puede servir para iniciar ya un primer nivel en ese razonamiento buscado.

Al tener las opciones de respuesta, el aspecto memorístico no es tan importante como la oportunidad que le damos de pensar y permite usar más recursos que favorecen el desarrollo del razonamiento estadístico.

Este tipo de pruebas permite reconocer errores generalizados en el razonamiento de los estudiantes y pueden ser cuantificados. Con estas pruebas también podemos valorar si el alumno tiene capacidad de resolver problemas sencillos.

Las preguntas de opción múltiple o los ítems que requieren que los estudiantes hagan coincidir las preguntas con conceptos o explicaciones apropiadas, pueden ser útiles para detectar el razonamiento y la comprensión conceptual de los estudiantes (Garfield y Gal 1999). Los ítems de este tipo de cuestionarios pueden evaluar diferentes objetivos y tienen un formato flexible y fácil de usar, al tiempo que los alumnos suelen estar familiarizados con este tipo de pruebas.

Desde un punto de vista métrico, cumplen lo que expone Muñiz y Fonseca-Pedrero (2009): “los instrumentos que se utilicen para la evaluación han de ser objetivos, claros, comprensibles por las partes, preferiblemente cuantitativos, fiables y válidos” (p.15)

Los ítems de elección múltiple son utilizados en numerosos contextos de evaluación educativa y su formato resulta objetivo, fácil, cómodo de corregir, y su versatilidad permite adaptarse a todo tipo de contextos y contenidos (Moreno, Martínez y Muñiz, 2004).

Respecto a las ventajas que expone Morales (2009), seleccionamos aquellas que nos proporcionan argumentos en la elección de pruebas objetivas para evaluar el razonamiento estadístico, se recogen en la Figura 20.

ELECCIÓN DE PRUEBAS OBJETIVAS PARA LA EVALUACIÓN DEL RAZONAMIENTO ESTADÍSTICO	
<b>Ventajas de las pruebas objetivas según (Morales, 2009, p. 11)</b>	<b>Justificación de la elección de pruebas objetivas para medir el “razonamiento estadístico”</b>
La fiabilidad es mayor: con otras pruebas <i>similares</i> los alumnos hubieran quedado <i>ordenados</i> de manera similar, <b>diferencian adecuadamente niveles de competencia</b> entre los alumnos. No hay lugar para ambigüedades en la respuesta o para que el alumno responda a lo que no se pregunta; la <b>corrección es muy objetiva</b> .	Necesidad de una prueba objetiva y fiable para medir un constructo complejo: “razonamiento estadístico”
Tienen gran variedad de aplicaciones y se pueden comprobar con nitidez objetivos muy distintos: no sólo de memoria, sino también objetivos de <b>comprensión, interpretación, análisis</b> , etc.	Se pretende medir la percepción del significado de los conceptos, el análisis y la síntesis de la información, la utilización de la información estadística y la capacidad de formular un juicio sobre situaciones concretas.
Se presta a <b>análisis estadísticos</b> muy útiles que también pueden hacerse con programas de ordenador.	Se comprobarán hipótesis con los datos del cuestionario.
Tienen un valor diagnóstico de fácil interpretación, <b>identifican con claridad puntos no sabidos o no entendidos</b> , etc.	Se pretende comprobar malentendidos y errores en el razonamiento estadístico.
<b>Pueden condicionar un estudio inteligente</b> si las preguntas son las adecuadas y los alumnos conocen el tipo de pregunta para el que deben estudiar.	Se utiliza la evaluación como intervención para promover el aprendizaje razonado de la estadística.

Figura 20. **Ventajas de las pruebas objetivas para medir el razonamiento estadístico.**

En una primera fase, se revisaron los instrumentos existentes para evaluar el razonamiento estadístico que han sido usados en la investigación y se seleccionaron aquellos que habían mostrado ser válidos para este fin.

De entre los instrumentos para la evaluación del razonamiento estadístico ya existentes, y tal como se comentó en el capítulo anterior, fueron elegidos los siguientes:

- *Comprehensive Assessment of Outcomes in a First Statistics course*, (CAOS).
- *Assessment Resource Tools for Improving Statistical Thinking* (ARTIST).

Estos instrumentos han sido la base para la construcción de los nuevos cuestionarios, por lo que fueron recopilados, traducidos sus ítems y adaptados al contexto de las ciencias de la salud.

En todo lo que se refiere a la elaboración de pruebas objetivas, formulación y análisis de ítems hemos seguido los principios y recomendaciones de Nunnally y Berstein, (1995), Martínez Arias, (1995); Haladyna, T.M., Downing, S. M., Rodríguez, M. C. (2002); Moreno, Martínez y Muñiz, 2006, Muñiz; Muñiz y Fonseca-Pedrero, (2004, 2009) y Morales, (2006, 2012, 2013).

Se han tenido en cuenta las recomendaciones de la American Statistical Association (ASA 2005)<sup>16</sup> para el aprendizaje y la evaluación de la estadística en alumnos universitarios, la literatura existente referente a las dificultades que tienen los estudiantes para comprender y aplicar ciertos conceptos estadísticos, y la propia experiencia docente con el fin de elaborar una tabla en la que quedan especificados los resultados de aprendizaje que evalúan nuestros cuestionarios.

Se elaboran para este trabajo tres tipos de instrumento para evaluar el razonamiento estadístico con diferentes objetivos:

Instrumentos de Evaluación del Razonamiento Estadístico	Finalidad del instrumento
Cuestionario para la evaluación del razonamiento estadístico (CERES).	<b>EVALUACIÓN</b> Detectar el conocimiento y la capacidad de razonamiento inicial y los cambios producidos como resultado de su aprendizaje
Pruebas parciales de evaluación del razonamiento estadístico (PaRES.)	<b>INTERVENCIÓN</b> Evaluación formativa utilizando estas pruebas
Prueba de evaluación del razonamiento estadístico de respuesta abierta parcialmente restringida.	<b>VALIDEZ CONVERGENTE</b> Aportar datos acerca de la validez de CERES

Figura 21. Instrumentos de Evaluación del Razonamiento Estadístico.

- ***Cuestionario para la evaluación del razonamiento estadístico (CERES).***

El cuestionario para la evaluación del razonamiento estadístico (CERES), (ANEXO 1), se elaboró específicamente para fines de esta investigación y nos permitió valorar el razonamiento estadístico de los alumnos.

Con este cuestionario se pretendía disponer de un instrumento de evaluación que tuviera la capacidad de hacer pensar y razonar a los estudiantes y

<sup>16</sup> Ver capítulo 3. (Guidelines for Assessment and Instruction in Statistics Education GAISE 2010).

nos permitiera recoger datos que nos aportaran información acerca del razonamiento estadístico mostrado por el grupo de alumnos de primer curso del Grado de Enfermería de la Universidad Pontificia Comillas de Madrid. Este instrumento de medida también podría ser aplicado a otras poblaciones de alumnos que cursan Estadística en un primer curso introductorio en otras titulaciones de Ciencias de la Salud.

Este instrumento permite diferenciar a los alumnos que muestran un buen nivel en el razonamiento estadístico y a los que no lo tienen, al tiempo que permite identificar errores particulares de los sujetos o de los grupos. Esta prueba se administró en dos ocasiones, antes de comenzar la asignatura de estadística y al finalizar el curso, lo que nos permitió detectar el conocimiento y la capacidad de razonamiento con la que llegaron los estudiantes, así como los cambios que se produjeron como resultado de su aprendizaje.

Este cuestionario se ha elaborado mediante una adaptación del *Comprehensive Assessment of Outcomes in a First Statistics Course*, (CAOS), (DelMas, Garfield, Ooms and Chance, 2007). A partir de este instrumento se realizó la adaptación de los ítems al contexto de ciencias de la salud, de manera que fueran reflejo de la aplicación de la estadística a la disciplina enfermera, y se redactaron también nuevos ítems.

En su versión definitiva CERES está formado por 40 ítems que tienen de 2 a 5 opciones de respuestas, en las que sólo una de ellas es correcta. Los contenidos del cuestionario atienden a conceptos básicos de un curso de iniciación a la estadística:

- Representatividad de las muestras.
- Interpretación de datos contenidos en tablas y gráficos.
- Uso e interpretación de los parámetros centrales.
- Concepto e interpretación de la desviación típica.
- Variabilidad.
- Distribución normal.
- Intervalos de confianza.
- Pruebas de hipótesis y significación estadística.
- Relación entre dos variables.

- ***Pruebas parciales de evaluación del razonamiento estadístico (PaRES.)***

Se han elaborado cuatro cuestionarios de evaluación del razonamiento estadístico (ANEXOS 2, 3, 4, 5) acerca de contenidos concretos :

- PaRES1. Estadística descriptiva (25 ítems).
- PaRES2. Distribución normal (15 ítems).
- PaRES3. Pruebas de significación (23 ítems).
- PaRES4. Análisis de varianza y Regresión (15 ítems).

Estas pruebas han sido pensadas para la evaluación formativa que forma parte de la intervención que será expuesta en un apartado posterior. Estas pruebas han sido administradas durante el curso con el objetivo de orientar al alumno hacia un estudio razonado y devolverle la información necesaria acerca de sus logros que le permita la autorregulación del aprendizaje y contribuir al mismo.

Las pruebas contienen un número diferente de ítems seleccionados de las pruebas ARTIST (*Assessment Resource Tools for Improving Statistical Thinking*). Estos ítems fueron todos adaptados al contexto sanitario y se incluyeron también nuevos ítems formulados para cubrir algunos objetivos de evaluación.

- ***Pruebas de evaluación del razonamiento estadístico de respuesta abierta parcialmente restringida***

El objetivo de esta prueba fue utilizarla para comprobar su correlación con los resultados de CERES, cuestionario construido para este trabajo, y aportar un elemento más a su validez. Además, se trata de una prueba de evaluación útil que aporta información sobre el aprendizaje de los alumnos.

La prueba (ANEXO 6) consta de 5 preguntas acerca de aspectos razonados de la estadística correspondientes a contenidos básicos. las respuestas son abiertas pero orientando concreta y sintéticamente las respuestas. La calificación se realiza mediante una clave de corrección.

La prueba y sus resultados se exponen en el apartado en el que se trata la validez convergente del cuestionario CERES.

## 5.2. Metodología de la construcción de los cuestionarios de evaluación de razonamiento estadístico.

Para la elaboración de los cuestionarios de evaluación, establecimos una serie de pasos:

1. Delimitar las principales áreas de contenido. Para ello nos servimos del análisis de los diferentes documentos normativos que configuran los ejes de los planes de estudios de enfermería, del análisis de las competencias y del análisis de los contenidos de los actuales planes de estudios de diversas universidades españolas, (capítulo 1).

2. Realizar un análisis de las distintas concepciones sobre el constructo “razonamiento estadístico”, analizando los elementos de su significado y delimitando y definiendo nuestro propio concepto para la presente investigación, (capítulo 3).

3. Realizar una revisión y síntesis de las investigaciones previas relacionadas con la evaluación del razonamiento estadístico. Para cumplir este objetivo, se procedió a la traducción de los instrumentos de evaluación encontrados. La mayor parte de ellos, como ya se ha comentado, se componen de pruebas objetivas, aunque también se revisaron otros constituidos por preguntas abiertas, (capítulo 3).

4. Elegir, entre los instrumentos de evaluación revisados, aquellos que se ajustan a los objetivos del presente trabajo, y seleccionar los ítems para su adaptación e inclusión en los nuevos cuestionarios.

5. Adaptar todos los ítems al contexto de las ciencias de la salud, procurando utilizar ejemplos reales de la práctica enfermera, teniendo en cuenta las recomendaciones de la American Statistical Association (ASA 2005), las concepciones erróneas de los alumnos recogidas en la literatura y en la propia experiencia docente.

6. Definir los resultados de aprendizaje que evalúan cada uno de los ítems seleccionados y someterlos al juicio de un grupo de expertos.

7. Llevar a cabo una consulta a expertos con el fin de recoger las opiniones de otros profesores de estadística acerca de si los resultados de aprendizaje que evalúa el cuestionario son acertados en un instrumento para evaluar el razonamiento estadístico de los estudiantes.

8. Realizar una prueba piloto con los cuestionarios diseñados.

Todo este proceso se llevó a cabo con la finalidad de que el conjunto de ítems representara adecuadamente el objetivo: evaluar el razonamiento estadístico.

En cuanto a los contenidos que van a evaluar los cuestionarios han de ser coincidentes con los incluidos en la asignatura de estadística correspondiente al plan de estudios del Grado de Enfermería de la Universidad Pontificia Comillas de Madrid, de dónde proceden las muestras para el estudio. Se han seleccionado aquellos contenidos que se consideran más importantes para la adquisición de las competencias que esta disciplina debe aportar a estos futuros profesionales de la salud y que tienen que ver con la explotación y la interpretación de la información estadística que resulta de los registros enfermeros y con la capacidad de lectura crítica de resultados de investigación y su aplicación a la práctica de enfermería. Se comprobó que estos contenidos estaban contemplados en los instrumentos seleccionados y fueron tenidos en cuenta en las posteriores adaptaciones que se llevaron a cabo con sus ítems.

En la búsqueda de algunas recomendaciones acerca de qué contenidos concretos deben estar presentes en los ítems de un cuestionario que evalúe el razonamiento estadístico, encontramos tan sólo algunos autores que realizan propuestas al respecto y que nos parece interesante recoger. Cobb (1998) propone algunas cuestiones a tener en cuenta en el diseño de pruebas objetivas para evaluar el pensamiento estadístico, de las que señalamos: la utilidad de las tablas 2x2 para hacer pensar sobre la asociación y las relaciones entre pares de variables, y las preguntas en las que los estudiantes tengan que asociar la forma de un histograma con la información que recibe del contexto del problema.

Varios son los autores que han considerado importante el razonamiento estadístico acerca de la asociación y su evaluación, (por ejemplo Garfield, 2003). Otros autores entienden que el razonamiento sobre la covariación juega un papel importante en un curso de introducción a la estadística, razonar e interpretar la relación entre dos variables “involucra comúnmente procesos de traducción entre los datos numéricos, las representaciones gráficas y las afirmaciones sobre covariación estadística y asociación causal” (Moritz, 2004, p.228). En un curso de introducción a la estadística ha de evaluarse el razonamiento sobre los conceptos relacionados con datos de dos variables, tanto cuantitativas como cualitativas, incluyendo, entre otros, diagramas de dispersión, tablas de contingencia, correlación, regresión y pruebas chi-cuadrado.



Se puso especial cuidado en asegurar que la formulación de cada ítem fuera congruente con el objetivo del cuestionario. Es decir, que evaluara el razonamiento estadístico del estudiante y de esa manera lograr la validez del cuestionario.

Para la selección de los ítems, se ha tenido en cuenta la forma y el momento de administrarlo. El Cuestionario para la Evaluación del Razonamiento Estadístico (CERES) fue administrado de forma grupal antes de iniciar las clases de estadística, por lo que nos encontramos alumnos de muy diversa formación matemática y estadística, cuyos conocimientos iniciales nos interesa detectar.

Con esta prueba “pretest” se pretendía poder cuantificar los cambios en el razonamiento de los alumnos, ya que se volvería a administrar al finalizar la asignatura como medida “postest”.

Además de recoger información sobre los cambios ocurridos en la capacidad de razonar de los estudiantes, nos interesaba detectar aquellos conceptos que ya poseían de forma correcta, así como los errores en el razonamiento y en la forma de interpretar información estadística. Por este motivo el cuestionario incluyó algunos ítems muy básicos, relativos a contenidos que han sido adquiridos en niveles de estudio anteriores y para los que se requiere un nivel inicial de razonamiento.

Los ítems de opción múltiple se han elaborado con un enunciado sencillo que presente la cuestión y deje claro qué se está pidiendo al alumno, una sola alternativa de respuesta correcta y varios “distractores” u opciones de respuesta incorrectas. Para el diseño de los distractores se han tenido en cuenta los errores más frecuentes que se recogen en la literatura revisada y los razonamientos equivocados que detectamos de forma recurrente en nuestros alumnos.

Los ítems se han elaborado con un número de entre tres y cinco distractores. Pese a que sobre este asunto hay distintas y argumentadas opiniones, en ningún caso se ha forzado una alternativa más cuando de forma natural no resulta necesaria. También se incluyen preguntas con respuestas de verdadero o falso. Estos cuestionarios pese a que los ítems tienen diferente número de opciones de respuesta, hay que señalar que la puntuación de cada ítems solo puede codificarse como “correcta” o “incorrecta”, por lo que se puede calcular el  $\alpha$  de Cronbach. También hay que decir que en el análisis de ítems no se aplicó la corrección de la adivinación.

Estos cuestionarios con preguntas objetivas nos permiten asignar puntuaciones a los sujetos e identificar patrones de errores más frecuentes. Por

otro lado este tipo de ítems es bastante flexible, se adapta a diversos tipos de objetivos educativos, y es de fácil aplicación.

Estos instrumentos fueron traducidos del inglés y cada uno de sus ítems fue analizado respecto a los contenidos y resultados de aprendizaje que evalúan. Algunos de los ítems tan sólo se tradujeron y se incluyeron sin modificación. En otros casos se realizaron algunos cambios en la redacción y, en todos ellos, se hizo una adaptación al contexto de ciencias de la salud, cambiando los ejemplos utilizados en los enunciados por otros del ámbito sanitario. Por último, también se enunciaron nuevos ítems, para evaluar algunos aspectos para los que no se encontró ninguna pregunta adecuada en los instrumentos de evaluación publicados.

### **5.3. Cuestionario para la Evaluación del Razonamiento Estadístico (CERES).**

Para el cuestionario CERES se seleccionaron del *Comprehensive Assessment of Outcomes in a First Statistics Course*, (CAOS), 37 ítems que compondrían el primer cuestionario piloto. De esta prueba que tiene 40 ítems, se eliminaron tres de ellos que contenían preguntas en las que los resultados de aprendizaje evaluados están relacionados con los conceptos y cálculo de probabilidad. Consideramos, después de revisar la literatura y a través de nuestra propia experiencia no recogerlos en el cuestionario. Los conceptos de probabilidad tienen unas dificultades para los estudiantes que son peculiares y específicas. El razonamiento acerca de la probabilidad ha sido tratado de forma particular en los estudios de investigación y por otro lado no hay contenidos sobre probabilidad, aparte de nociones básicas, recogidos en el programa de nuestra asignatura.

Una vez seleccionados estos 37 ítems, fueron traducidos y adaptados al contexto de las ciencias de la salud mediante la redacción de enunciados y opciones de respuesta nuevas, pero manteniendo los resultados de aprendizaje evaluados y el patrón con el que estaban planteados los ítems. Algunos de los ítems se mantuvieron en su redacción original traducida.

En la Figura 22 se recoge la composición del cuestionario CERES piloto, indicando los contenidos que evalúan y las modificaciones y adaptaciones que se han realizado respecto a los ítems originales de CAOS.

Contenidos	CERES piloto	Ítems de CAOS que conforman el cuestionario CERES piloto
REPRESENTACIONES GRÁFICAS	ÍTEM 1	Adaptación del ítem 1 de CAOS
	ÍTEM 2	Adaptación del ítem 2 de CAOS
	ÍTEMS 3 - 5	Adaptación de los ítems 3 – 5 de CAOS
	ÍTEM 6	Ítem nuevo basado en los gráficos de los ítems 3 – 5
DISTRIBUCIÓN DE DATOS REPRESENTACIÓN GRÁFICA DE DATOS VARIABILIDAD	ÍTEM 7	Adaptación del ítem 6 de CAOS
EFFECTOS DE LA ASIGNACIÓN AL AZAR	ÍTEM 8	Tomado del cuestionario CAOS (ítem 7)
VARIABILIDAD Y DISPERSIÓN REPRESENTACIÓN GRÁFICA DE DATOS GRÁFICO DE CAJAS Y PATILLAS	ÍTEMS 9 y 10	Adaptación de los ítems 8 – 9 de CAOS
PARÁMETROS CENTRALES <ul style="list-style-type: none"> <li>• Representación gráfica de datos</li> <li>• Gráfico de cajas y patillas</li> </ul>	ÍTEM 11	Adaptación del ítem 10 de CAOS
PARÁMETROS CENTRALES <ul style="list-style-type: none"> <li>• Distribución de datos</li> <li>• Valores extremos</li> </ul>	ÍTEMS 12 y 13	Tomados del cuestionario CAOS (ítems 11 - 12)
MUESTREO <ul style="list-style-type: none"> <li>• Tamaño muestral</li> <li>• Comparación de muestras</li> </ul>	ÍTEM 14	Tomado del cuestionario CAOS (ítem 13)
DISTRIBUCIÓN DE DATOS REPRESENTACIÓN GRÁFICA DE DATOS VARIABILIDAD	ÍTEMS 15 y 16	Tomados del cuestionario CAOS (ítems 14 y 15)
VARIABILIDAD EN EL MUESTREO	ÍTEM 17	Adaptación del ítems 16 de CAOS
VARIABILIDAD	ÍTEM 18	Tomado del cuestionario CAOS (ítem 18)

<p>INFERENCIA ESTADÍSTICA</p> <ul style="list-style-type: none"> <li>• Análisis de significación</li> <li>• Nivel de significación</li> <li>• p-valor</li> </ul>	ÍTEM 19	Adaptación del ítem 19 de CAOS
ANÁLISIS BIVARIADO	ÍTEM 20	Adaptación del ítem 20 de CAOS
	ÍTEM 21	Adaptación del ítem 21 de CAOS
	ÍTEM 22	Adaptación del ítem 22 de CAOS
<p>INFERENCIA ESTADÍSTICA</p> <ul style="list-style-type: none"> <li>• Análisis de significación</li> <li>• Nivel de significación</li> <li>• p-valor</li> </ul>	ÍTEM 23	Adaptación del ítem 23 de CAOS
<p>INFERENCIA ESTADÍSTICA</p> <ul style="list-style-type: none"> <li>• Diseño experimental</li> <li>• Causalidad</li> </ul>	ÍTEM 24	Adaptación del ítem 24 de CAOS
<p>INFERENCIA ESTADÍSTICA</p> <ul style="list-style-type: none"> <li>• Análisis de significación</li> <li>• Nivel de significación</li> <li>• p-valor</li> </ul>	ÍTEMS 25 - 27	Adaptación de los ítems 25 - 27 de CAOS
INTERVALOS DE CONFIANZA	ÍTEMS 28 - 31	Adaptación de los ítems 28 - 31 de CAOS
<p>DISTRIBUCIÓN DE DATOS</p> <p>REPRESENTACIÓN GRÁFICA DE DATOS</p> <p>VARIABILIDAD</p>	ÍTEM 32	Adaptación del ítem 33 de CAOS
<p>MUESTREO</p> <ul style="list-style-type: none"> <li>• Representatividad de la muestra</li> </ul>	ÍTEM 33	Adaptación del ítem 34 CAOS
DISTRIBUCIÓN DE MUESTREO	ÍTEM 34	Adaptación del ítem 35 de CAOS
<p>MUESTREO</p> <ul style="list-style-type: none"> <li>• Representatividad de la muestra</li> </ul>	ÍTEM 35	Adaptación del ítem 38
ANÁLISIS BIVARIADO	ÍTEM 36	Adaptación del ítem 39 de CAOS

INFERENCIA ESTADÍSTICA <ul style="list-style-type: none"> <li>• Análisis de significación</li> <li>• Nivel de significación</li> <li>• p-valor</li> </ul>	ÍTEM 37	Adaptación del ítem 40 de CAOS
---	---------	--------------------------------

Figura 22. Ítems de CAOS que conforman el cuestionario CERES piloto.

### 5.3.1. Tabla de especificaciones de los resultados de aprendizaje que evalúa CERES.

Los objetivos que pueden plantearse para evaluar el razonamiento estadístico pueden ser de distinto tipo, pero fundamentalmente tienen que ver con la percepción del significado de los conceptos, el análisis y la síntesis de la información, la utilización de la información estadística y la capacidad de formular un juicio sobre situaciones concretas (Bloom, 1956).

Una vez seleccionados los ítems de los cuestionarios se analizó y definió cuál es el resultado de aprendizaje que se pretende evaluar con cada uno de ellos. La Figura 23 recoge los resultados de aprendizaje correspondientes a los ítems que conforman los instrumentos diseñados.

CERES PILOTO		
CONTENIDOS	RESULTADOS DE APRENDIZAJE	ÍTEMS
REPRESENTACIONES GRÁFICAS	Describir e interpretar en el contexto de los datos, la distribución de una variable representada en un histograma.	1
	Reconocer dos formas de representar gráficamente el mismo conjunto de datos.	2
	Asociar las representaciones gráficas presentadas con el tipo de variable descrita	3, 4, 5 Y 6
DISTRIBUCIÓN DE DATOS REPRESENTACIÓN GRÁFICA DE DATOS VARIABILIDAD	Asociar a un gráfico de frecuencias la información contenida en una tabla de datos respecto al centro y dispersión de la variable.	7
	Capacidad de estimar correctamente y comparar las desviaciones estándar de diferentes histogramas	15 Y 16
	Entender que una distribución de muestreo con una media mayor que la mediana es más probable que se desvíe a la derecha.	32

VARIABILIDAD Y DISPERSIÓN REPRESENTACIÓN GRÁFICA DE DATOS GRÁFICO DE CAJAS Y PATILLAS	Entender el concepto de desviación típica como un parámetro que informa de la dispersión, y la capacidad de interpretarlo en el contexto de un gráfico de cajas y patillas.	9
	Entender que los gráficos de cajas no proporcionan estimaciones precisas de los porcentajes de los datos por encima o por debajo de los valores a excepción de los cuartiles.	10
PARÁMETROS CENTRALES <ul style="list-style-type: none"> <li>• Representación gráfica de datos</li> <li>• Gráfico de cajas y patillas</li> </ul>	Comprensión del concepto de mediana y su interpretación en el contexto de un gráfico de cajas y patillas.	11
PARÁMETROS CENTRALES <ul style="list-style-type: none"> <li>• Distribución de datos</li> <li>• Valores extremos</li> </ul>	Capacidad para comparar grupos, teniendo en cuenta, en las distribuciones, dónde se centran la mayoría de los datos al margen del comportamiento de algunos individuos de la muestra.	12
	Capacidad para comparar grupos, mediante la comparación de los promedios, extrayendo esta información de la representación de la distribución de datos.	13
MUESTREO <ul style="list-style-type: none"> <li>• Representatividad de la muestra</li> </ul>	Comprensión de los factores que permiten que los resultados obtenidos sobre una muestra sean generalizados a la población.	35
	Comprensión de la representatividad de la población de una muestra grande aleatoria. Ley de los grandes números.	33
MUESTREO <ul style="list-style-type: none"> <li>• Tamaño muestral</li> <li>• Comparación de muestras</li> </ul>	Entender que la comparación de dos grupos no requiere que las muestras tengan igual tamaño, especialmente si los dos conjuntos de datos son grandes.	14
VARIABILIDAD EN EL MUESTREO	Entender que las muestras pequeñas tienen mayor variabilidad que las muestras grandes.	17
VARIABILIDAD	Comprensión del significado de la variabilidad en un conjunto de datos obtenido como resultado de medidas repetidas y en un contexto en el que se desea que esa variabilidad sea pequeña.	18

DISTRIBUCIÓN EN EL MUESTREO	Comprensión de que la distribución de medias de un conjunto de muestras tiene una distribución normal de media próxima a la media poblacional.	34
INFERENCIA ESTADÍSTICA <ul style="list-style-type: none"> <li>• Análisis de significación</li> <li>• Nivel de significación</li> <li>• p-valor</li> </ul>	Capacidad de reconocer una interpretación correcta, incorrecta y errónea acerca de un p-valor	25, 26 Y 27
	Comprensión del concepto de nivel de significación y la capacidad de asociarlo a la interpretación del p-valor, conociendo qué valores de p-valor son deseables en los estudios de investigación.	19
	Entender que los resultados de un análisis de significación para detectar diferencias entre dos grupos, depende del tamaño muestral.	23
	Interpretación de los resultados de una prueba de significación cuando es rechazada la hipótesis nula.	37
INFERENCIA ESTADÍSTICA <ul style="list-style-type: none"> <li>• Diseño experimental</li> <li>• Causalidad</li> </ul>	Capacidad de entender que un diseño experimental con asignación aleatoria apoya la inferencia causal.	24
INTERVALOS DE CONFIANZA	Detectar una correcta interpretación del intervalo de confianza, como los valores entre los que se encuentran la variable en cada uno de los individuos de la muestra.	28
	Detectar una incorrecta interpretación del nivel de confianza, como el porcentaje de individuos de la población cuyo valor de una variable se encuentra dentro de los límites del intervalo de confianza.	29
	Detectar una incorrecta interpretación del nivel de confianza como el porcentaje de todas las posibles muestras que tienen una media dentro de los límites del intervalo de confianza	30
	Interpretar correctamente un intervalo de confianza. Seguridad de que el intervalo de confianza contenga la verdadera media de la población.	31

ANÁLISIS BIVARIADO	Asociar la descripción de la relación existente entre dos variables con la representación de dichas variables en un diagrama de dispersión.	20
	Describir la relación entre dos variables que se muestran en un gráfico de dispersión.	21
	Entender que la correlación no implica causalidad.	22
	Comprensión de que no es prudente extrapolar un modelo de regresión.	36

Figura 23. **Contenidos y resultados de aprendizaje del Cuestionario para Evaluar el Razonamientos Estadístico (CERES).**

A la vista de la información recogida en la Figura 23 respecto a los contenidos que valora el cuestionario CERES podemos decir que todos ellos han sido objeto de interés en los estudios acerca del razonamiento estadístico. Hacemos notar el paralelismo existente entre los contenidos de los que tratan los estudios publicados en los artículos de investigación acerca del razonamiento estadístico que han sido revisados y que se encuentran recogidos en la Figura 10 y los que evalúan nuestro cuestionario.

### 5.3.2. Alineación del cuestionario CERES / GAISE (Guidelines for Assessment and Instruction in Statistics Education GAISE 2010).

Ya hemos mostrado en capítulos anteriores el interés por tener presente en este trabajo las directrices para la enseñanza y evaluación de la estadística de American Statistical Association (ASA, 2005). En el estudio de Zieffler, Garfield, delMas y Björnsdóttir (2010), examinan mediante un consenso de expertos la coincidencia de los ítems de una versión de CAOS, con los resultados de aprendizaje documentados en el informe GAISE. En nuestro trabajo, como otra aportación a la validez de contenido, hemos querido también analizar esta alineación con los ítems de nuestro cuestionario y el resultado se resume en la siguiente Figura.



ALINEACIÓN CERES – GAISE	
GAISE	CERES
<b>Los estudiantes deben entender por qué:</b>	
Los datos han de superar las anécdotas	Ítem 19
La variabilidad es natural, previsible y cuantificable	Ítems 13 y 23
Un muestreo aleatorio permite que los resultados de las encuestas y experimentos, puedan extenderse a la población de la cual se ha tomado la muestra	Ítem 18
La asignación aleatoria en experimentos comparativos permite extraer conclusiones respecto a la causa-efecto	Ítem 31
La asociación no es causalidad	Ítem 39
La significación estadística no implica necesariamente una importancia práctica, especialmente para los estudios con muestras de gran tamaño	---
Al no encontrar una diferencia estadísticamente significativa o relación, no significa necesariamente que no hay diferencia o que no haya relación, especialmente para los estudios con muestras pequeñas	Ítem 30
<b>Los estudiantes deben reconocer:</b>	
Las fuentes comunes de sesgo en los estudios y experimentos	Ítem 18
Cómo determinar la población a la que se pueden extrapolar los resultados de la inferencia estadística, sobre la base de cómo se recogieron los datos	Ítem 18
Cómo determinar una inferencia de causa y efecto de una asociación, basada en cómo se recogieron los datos (diseño del estudio)	Ítem 39
Que las palabras tales como " normal ", "azar ", y "correlación " tienen significados específicos en estadística diferentes al que tienen en un uso común	Ítems 2, 12, 13, 19, 20, 21, 25 y 39
<b>Los estudiantes deben entender las partes de los siguientes procesos estadísticos:</b>	
Cómo obtener o generar datos	Ítem 18*
Cómo representar gráficamente los datos como un primer análisis , y cómo saber cuándo es suficiente para responder a la cuestión de interés	Ítems 2, 3, 4, 8, 9, 10, 11, 12 y 37

Cómo interpretar resúmenes numéricos y representaciones gráficas de datos, tanto para responder preguntas como para comprobar las condiciones (uso correcto de procedimientos estadísticos)	Ítems 1, 2, 4, 5, 6, 12, 15, 16, 17, 13, 14, 19, 20, 21 y 38
Cómo hacer un uso adecuado de la inferencia estadística	Ítems 29, 26, 27, 28, 33, 34, 35 y 36
Cómo comunicar los resultados de un análisis estadístico	
<b>Los estudiantes deben entender las ideas básicas de la inferencia estadística, incluyendo:</b>	
El concepto de distribución de la muestra y cómo se aplica para hacer inferencias de parámetros poblacionales, basadas en una muestras (error estándar)	Ítems 22, 24 y 25
El concepto de significación estadística, niveles de significación y p-valor	Ítems 29, 26, 27 y 28
El concepto de intervalo de confianza, incluyendo la interpretación de nivel de confianza y margen de error	Ítems 33, 34, 35 y 36
<b>Finalmente, los estudiantes deben saber:</b>	
Cómo interpretar los resultados estadísticos en un contexto	Ítems 32* y 7
Como criticar noticias y artículos de revistas que incluyen información estadística, incluyendo la presentación, identificando lo que falta y los errores en el estudios o en los métodos utilizados para generar la información	

Figura 24. Alineación de los ítems de CERES con las recomendaciones GAISE.

\* Ítems relacionados tangencialmente con el resultado de aprendizaje

Comprobamos que el cuestionario CERES está constituido por ítems que cubren prácticamente la totalidad de los resultados de aprendizaje que recoge el informe GAISE.

#### 5.4. Validez del Cuestionario de Evaluación del Razonamiento Estadístico (CERES)

En esta sección del trabajo se aportan evidencias acerca de la validez del instrumento en varios aspectos:

### 5.4.1. Validez de contenido

Este cuestionario trata de comprobar, desde un punto de vista razonado, los conocimientos respecto a la materia de estadística. Por ello, su validez, por un lado, dependerá del grado con el que el cuestionario represente los contenidos y los resultados de aprendizaje de un curso básico de estadística y lo haga de forma equilibrada, aspecto analizado que ha sido anteriormente expuesto.

Los contenidos del cuestionario son coincidentes con:

- La estructura de contenidos reflejada en el capítulo 1 como resultado de la revisión de los planes de estudio del grado de Enfermería de las diferentes universidades.
- Los contenidos de la asignatura de estadística de la Escuela de Enfermería con cuyos alumnos se realiza este estudio.
- Los contenidos encontrados en las investigaciones acerca del razonamiento estadístico.
- Las recomendaciones de la American Statistical Association (ASA) para la enseñanza y evaluación de la estadística orientada a conseguir el razonamiento estadístico de los alumnos. En este aspecto, se ha presentado en el apartado anterior, la alineación del cuestionario CERES/GAISE.

Por otro lado, también es una aportación a la validez, comprobar que la formulación de los ítems vaya encaminada a que el alumno lleve a cabo procesos de pensamiento más que aspectos memorísticos. Para asegurarnos de ello se han extraído los resultados de aprendizaje que, respecto a los contenidos, son evaluados en el cuestionario y éstos han sido valorados por un grupo de expertos, resultados que se presentan a continuación.

Que el cuestionario represente el alcance del constructo *razonamiento estadístico* está de alguna manera garantizado también por la procedencia de la mayoría de los ítems que componen el instrumento. Nos referimos al test CAOS, de probada utilidad para medir el razonamiento de los estudiantes en diversas investigaciones, y que forma parte de un programa en el que se han recogidos datos de una muestra muy extensa y variada de estudiantes universitarios durante varios años.

La adaptación de la formulación de los ítems al contexto sanitario que se ha realizado en el presente trabajo, ha sido cuidadosa y ha mantenido con rigor el objetivo y los resultados de aprendizaje que pretenden evaluar cada uno de los ítems del cuestionario.

El análisis de ítems realizado, estudio de los índices de dificultad y de los distractores, proporciona también otro tipo de evidencia de la validez de contenido.

#### **5.4.1.1. Opinión de expertos.**

Uno de los criterios que se utiliza para la selección de ítems en la construcción de un instrumento es el generado por un juicio de expertos (Osterlind, 1989; Sireci, 1998; Wilson 2005; Muñiz y Fonseca-Pedrero 2009). De esta forma se asegura que la selección de ítems representa al constructo y es relevante para el uso que se dará a las puntuaciones generadas por el instrumento.

En nuestro caso, los nuevos cuestionarios están integrados por preguntas que han sido adaptadas de ítems tomados de instrumentos de investigaciones previas. Dichos ítems han sido ya valorados como adecuados para medir el razonamiento estadístico.

Para construir nuestros instrumentos hemos adaptado los ítems al ámbito de la enfermería manteniendo los resultados de aprendizaje que pretende evaluar cada uno de ellos, por lo que contamos con cierta confianza acerca de la validez de contenido. Por ello no se plantea un juicio de expertos como tal, con la rigurosidad metodológica y estadística de cuando se aplica como única prueba de la validez de contenido, sino una consulta a profesores de estadística del grado de enfermería a los que se les pide una opinión acerca de los resultados de aprendizaje del cuestionario CERES, que nos aporten datos para reforzar la validación de contenido. Motivo por el que hablamos de opinión de expertos.

Por otro lado, el resultado final de la adaptación de las preguntas, tanto de CERES como de las pruebas parciales de razonamiento estadístico usadas en la intervención de este trabajo, pretenden ser una de las aportaciones de esta tesis. Con estos cuestionarios adaptados, constituidos en su conjunto por 118 ítems, se ha pretendido contar con unas pruebas de evaluación que permitan a los estudiantes de enfermería establecer relaciones significativas de los

conocimientos estadísticos con otros conocimientos del ámbito de su formación en ciencias de la salud y con escenarios reales de investigación de enfermería.

De acuerdo al propósito y al ámbito de aplicación de este instrumento, consideramos que los expertos debían ser profesores de estadísticas en Ciencias de la Salud (Millman y Greene, 1989). En este caso, entendemos que este grupo de docentes puede aportarnos información muy valiosa (Zimmerman, Sudweeks, Shelley y Wood, 1990).

Esta consulta de opinión se planteó de forma anónima y con la intención de contar con un gran grupo de profesores que expresaran su opinión, asegurándonos que representasen una diversidad relevante de puntos de vista. Se preparó un cuestionario para ser completado vía online con una carta de presentación, solicitando su colaboración y explicándoles las características de la investigación. Esta carta se remitió en el cuerpo del correo electrónico junto con el enlace que da acceso al cuestionario. También se les solicitó que agregasen otros posibles resultados de aprendizaje que no estuvieran contemplados en el cuestionario y que consideraran importantes para la finalidad del instrumento. El cuestionario fue enviado a 42 profesores de estadística de la titulación de Enfermería de otras tantas universidades españolas.

Se les solicitó que emitieran un juicio acerca de la importancia que tenían para la finalidad del instrumento los resultados de aprendizaje evaluados en cada bloque de contenidos. Es decir, si en su opinión eran relevantes para evaluar el razonamiento estadístico de los alumnos.

Se solicitó que cada uno de estos resultados fuera puntuando de 0 a 5 con el criterio de **0= Nada relevante , 1,2, 3, 4, 5=Muy relevante**. Se utilizó una escala de valoración en lugar de expresar su acuerdo o desacuerdo para obtener una información más matizada, en el supuesto de que el número de expertos con el que pudiéramos contar fuera pequeño.

Finalmente fueron 8 los profesores que contestaron el cuestionario. En la Tabla 4 se presentan la mediana y desviación típica de las puntuaciones asignadas por el grupo de expertos para cada uno de los resultados de aprendizaje del cuestionario.

Se obtuvieron puntuaciones, en el valor de la mediana, por encima de 4 en 25 de los 28 resultados de aprendizaje propuestos. El más valorado, con 5 puntos y el acuerdo total de todos los expertos, es el siguiente: *“compresión de los factores que permiten que los resultados obtenidos sobre una muestra sean generalizados a la población”*.

Los dos resultados en los que se encontró menor grado de acuerdo entre expertos son:

*“reconocer dos formas de representar gráficamente el mismo conjunto de datos. Asociar la distribución de una variable representada en un histograma a su representación en un gráfico de cajas y patillas”.*

*“Detectar una incorrecta interpretación del nivel de confianza como el porcentaje de todas las posibles muestras que tienen una media dentro de los límites del intervalo de confianza”*

En ambos casos, tres de los expertos que respondieron el cuestionario lo consideraron poco relevante. Se revisaron las preguntas correspondientes y se decidió mantenerlos en el cuestionario, aplicando nuestro criterio pese a los resultados de la consulta. Son dos preguntas de interés en una prueba de evaluación que constituirá, no hay que olvidarlo, el examen con el que se calificará oficialmente a los alumnos con los que realizamos esta investigación.

La primera de ellas evalúa la interpretación de la información recogida en dos representaciones gráficas que son básicas y fundamentales (ver ítem 3 del cuestionario CERES definitivo, (ANEXO 1) . Hacer esta interpretación de forma conjunta requiere un razonamiento para el que se ha de utilizar múltiples conceptos estadísticos, por lo que creemos que es un ítem adecuado para esta prueba. En el segundo caso, se trata del resultado de aprendizaje de una pregunta que pertenece a una serie de 4 ítems en los que se plantean todas las posibles interpretaciones erróneas del nivel de confianza de un intervalo calculado para la media, y la interpretación correcta, con respuesta verdadero o falso (ítems 33-36, ANEXO 1). Eliminar una de las interpretaciones equivocadas, altera el objetivo de estas preguntas que es presentar al alumno todas las posibilidades de error que puede cometer para que pueda discriminar la interpretación válida.

**Tabla 4. Descriptivos de la valoración realizada por el grupo de expertos acerca de los resultados de aprendizaje del cuestionario de evaluación de razonamiento estadístico.**

Resultado de aprendizaje	Media	Mediana	Desviación típica
1 Reconocer dos formas de representar gráficamente el mismo conjunto de datos. Asociar la distribución de una variable representada en un histograma a su representación en un gráfico de cajas y patillas.	2,87	3,5	1,95
2 Asociar a un gráfico de frecuencias la información respecto al centro y dispersión de una variable.	4,25	4,5	0,88
3 Asociar las representaciones gráficas presentadas con el tipo de variable descrita.	4,62	5	0,74
4 Capacidad de estimar correctamente y comparar las desviaciones estándar de diferentes histogramas.	3,75	4	1,28
5 Entender que una distribución de muestreo con una media mayor que la mediana es más probable que se desvíe a la derecha.	3,37	3,5	1,18
6 Entender el concepto de desviación típica como un parámetro que informa de la dispersión, y la capacidad de interpretarlo en el contexto de un gráfico de cajas y patillas.	3,75	4	1,66
7 Entender que los gráficos de caja y patillas no proporcionan estimaciones precisas de los porcentajes de los datos por encima o por debajo de los valores a excepción de los cuartiles.	3,75	3,5	0,88
8 Comprensión del concepto de mediana y su interpretación en el contexto de un gráfico de cajas y patillas.	3,75	4	1,66
9 Comprensión de los factores que permiten que los resultados obtenidos sobre una muestra sean generalizados a la población.	5	5	0
10 Capacidad para comparar grupos, teniendo en cuenta, en las distribuciones, dónde se centran la mayoría de los datos al margen del comportamiento de algunos individuos de la muestra.	4,37	5	0,91
11 Capacidad para comparar grupos, mediante la comparación de las medias, extrayendo esta información de la representación de la distribución de datos.	4,37	4,5	0,74
12 Entender que la comparación de dos grupos no requiere que las muestras tengan igual tamaño, especialmente si los dos conjuntos de datos son grandes.	4,12	4	0,64

13 Entender que las muestras pequeñas tienen mayor variabilidad que las muestras grandes.	3,87	5	1,88
14 Comprensión del significado de la variabilidad en un conjunto de datos obtenido como resultado de medidas repetidas y en un contexto en el que se desea que esa variabilidad sea pequeña.	4,12	4,5	0,99
15 Comprensión de la representatividad de la población de una muestra grande aleatoria. Ley de los grandes números.	4,5	4,5	0,53
16 Comprensión de que la distribución de medias de un conjunto de muestras tiene una distribución normal de media próxima a la media poblacional.	4,5	5	0,75
17 Capacidad de reconocer una interpretación correcta, incorrecta y errónea acerca de un p-valor.	4,75	5	0,46
18 Comprensión del concepto de nivel de significación y la capacidad de asociarlo a la interpretación del p-valor, conociendo qué valores de p son deseables en los estudios de investigación.	4,87	5	0,35
19 Entender que los resultados de un análisis de significación para detectar diferencias entre dos grupos, dependen del tamaño muestral.	4,5	5	0,75
20 Capacidad de entender que un diseño experimental con asignación aleatoria apoya la inferencia causal.	4,62	5	0,74
21 Interpretación de los resultados de una prueba de significación cuando es rechazada la hipótesis nula.	4,87	5	0,35
22 Detectar una incorrecta interpretación del intervalo de confianza, como los valores entre los que se encuentran la variable en cada uno de los individuos de la muestra.	4,5	4,5	0,53
23 Detectar una incorrecta interpretación del nivel de confianza, como el porcentaje de individuos de la población cuyo valor de la variable se encuentra dentro de los límites del intervalo de confianza.	3,87	4	1,35
24 Detectar una incorrecta interpretación del nivel de confianza como el porcentaje de todas las posibles muestras que tienen una media dentro de los límites del intervalo de confianza.	3,87	4	1,35
25 Interpretar correctamente un intervalo de confianza. Seguridad de que el intervalo de confianza contenga la verdadera media de la población.	4,75	5	0,46



26 Asociar la descripción de la relación existente entre dos variables con la representación de dichas variables en un diagrama de dispersión.	4,35	4,5	1,03
27 Entender que la correlación no implica causalidad.	4,62	5	0,74
28 Comprensión de que no es prudente extrapolar un modelo de regresión	4,25	4,5	1,03

En la tabla 5 se observa que prácticamente todos los resultados de aprendizaje han sido considerados relevantes por los profesores consultados, con algunas discrepancias. En general las puntuaciones han sido mayoritariamente altas para todos los resultados de aprendizaje que son considerados adecuados para medir evaluar el razonamiento estadístico.

Tabla 5. Relevancia asignada por los expertos a los resultados de aprendizaje.

Resultado de Aprendizaje nº	Nº de expertos	
	Poco relevante	Relevante-muy relevante
1	3	5
2		8
3		8
4	2	6
5	1	3
6	1	7
7		8
8	1	7
9		8
10		8
11		8
12		8
13	2	6
14		8
15		8
16		8
17		8
18		8
19		8
20		8
21		8
22		8
23	1	7
24	3	5
25		8
26	1	7
27		8
28	1	7

Por otro lado, algunos de los profesores consultados consideraron que había algunos aspectos que no se habían tenido en cuenta y que en su opinión deberían estar recogidos en el instrumento. De sus sugerencias extraemos los resultados de aprendizaje siguientes:

- Interpretar resúmenes de datos y tablas de frecuencias
- Interpretar la influencia del sesgo sobre los valores de las medidas de tendencia central
- Interpretar el rango intercuartílico como medida de variabilidad
- Interpretación del concepto de percentil en las situaciones en las que se utiliza en Enfermería

Estas indicaciones nos llevaron a incluir cuatro ítems que valoraran estos aspectos que corresponden respectivamente a los ítems 1, 4, 5 y 7 del cuestionario definitivo y que se recogen a continuación:

1. La siguiente tabla de frecuencias recoge los datos de tensión arterial sistólica (mmHg) de los 300 trabajadores de una empresa.

Tabla de Frecuencias para TA Sistólica \_mmHg\_

Clase	Límite Inferior	Límite Superior	Marca	Frecuencia	Frecuencia Relativa	Frecuencia Acumulativa	Frecuencia Acum.Re
menor o igual		60,0		0	0,0000	0	0,00
1	60,0	75,0	67,5	2	0,0067	2	0,00
2	75,0	90,0	82,5	30	0,1000	32	0,10
3	90,0	105,0	97,5	84	0,2800	116	0,38
4	105,0	120,0	112,5	110	0,3667	226	0,75
5	120,0	135,0	127,5	59	0,1967	285	0,95
6	135,0	150,0	142,5	11	0,0367	296	0,98
7	150,0	165,0	157,5	2	0,0067	298	0,99
8	165,0	180,0	172,5	2	0,0067	300	1,00
9	180,0	195,0	187,5	0	0,0000	300	1,00
10	195,0	210,0	202,5	0	0,0000	300	1,00
mayor	210,0			0	0,0000	300	1,00

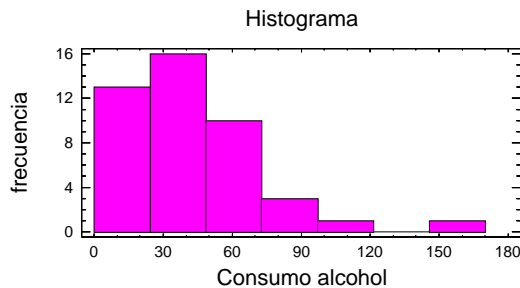
De las siguientes afirmaciones acerca de la información que nos proporciona esta tabla ¿cuál es cierta?

- a. La tensión arterial sistólica más frecuente en este grupo de trabajadores es 110 mmHg.
- b. Hay 226 trabajadores que tienen una tensión arterial sistólica inferior a 120 mmHg.**
- c. Hay 2 trabajadores que tienen una tensión arterial sistólica de 172,5 mmHg.
- d. El 95% de los trabajadores tienen unos valores de tensión arterial entre 120 y 135 mmHg.

**Resultado de aprendizaje 1:**  
**Capacidad de interpretar el resumen de datos recogido en una tabla de frecuencias**

Los ítems 4 y 5 se refieren a la siguiente situación:

El siguiente gráfico representa la distribución del consumo semanal de alcohol (en gramos) de una muestra aleatoria de estudiantes universitarios.



4. ¿Qué valores para la media y la mediana parecen más acertados?

- a. media= 40,8 y mediana= 52,3
- b. media= 31,4 y mediana= 38,5
- c. media= 38,5 y mediana= 31,4**
- d. media= 32,5 y mediana= 32,5

**Resultado de aprendizaje 4:**  
**Reconocer en un histograma la asimetría y el sesgo de una distribución de datos y su influencia sobre los valores de las medidas de tendencia central**

Item 5. ¿Cuál es la mejor medida de la variabilidad de este conjunto de datos?

- a. El rango, porque te informa de la extensión total de los datos.
- b. La desviación estándar, porque está basada en la información de todo el conjunto de datos.
- c. La desviación estándar porque es la medida de variabilidad más utilizada.
- d. **El rango intercuartílico, porque es resistente a los valores extremos.**

**Resultado de aprendizaje 5:**

**Reconocer en un histograma con puntuaciones atípicas la conveniencia de utilizar el rango intercuartílico como medida de variabilidad del conjunto de datos representado, frente a la desviación estándar y el rango, parámetros afectados por los valores extremos.**

Ítem 7. La madre de un niño de cuatro años que ha sido pesado y medido en la consulta de enfermería, pregunta a un estudiante de enfermería que realiza allí sus prácticas. ¿Qué significa que mi hijo esté en el percentil 65º de peso?

¿Tú qué respondes a esa pregunta?

- a. Significa que la media de peso del 65% de los niños de cuatro años coincide con el peso de su hijo.
- b. **Significa que el 65% de los niños de cuatro años pesan menos que su hijo.**
- c. Significa que el 65% de los niños de cuatro años pesan más que su hijo.
- d. Significa que el peso de su hijo varía en un 65% respecto al resto de los niños de cuatro años.

**Resultado de aprendizaje 7:**

**Comprensión del concepto de percentil y su interpretación en un contexto específico**

La valoración de expertos, es un elemento más que aporta información sobre la validez. Nos ha aportado datos sobre la idoneidad de los resultados de aprendizaje para evaluar la dimensión del razonamiento. Y por último, nos ha permitido incorporar cuatro nuevas preguntas sobre aspectos que nos han parecido muy acertados.

### 5.4.2. Prueba piloto del cuestionario. Validez discriminante

Este primer cuestionario con el que se realizó la prueba piloto se elaboró, como se ha comentado anteriormente, seleccionando 37 ítems del Comprehensive Assessment of Outcomes in a First Statistics Course, (DelMas, Garfield, Ooms and Chance, 2007).

Este cuestionario fue administrado en junio de 2012 a una **muestra piloto de 124 alumnos** del primer curso del grado de Enfermería que habían concluido la asignatura de estadística y realizado, también, los exámenes de la convocatoria ordinaria, aunque aún no se habían hecho públicas las calificaciones. Este grupo de alumnos sobre el que se realizó la prueba piloto se presupone con características similares a los sujetos que conformarán la posterior muestra del estudio y que serán los alumnos de primero del grado de Enfermería del siguiente curso académico.

En la aplicación de la prueba, se les animó a contestar razonada y cuidadosamente el cuestionario, a no dejar ningún ítem sin responder recurriendo al conocimiento que habían adquirido, a su intuición, y a todos los recursos de los que disponían y se les aseguró que se contestaría a las dudas que surgieran en el transcurso de la prueba.

Esta prueba se planteó con el fin de conseguir información objetiva y específica sobre los ítems. Durante su realización se respondió a los alumnos todas las dudas que surgieron a la hora de entender los enunciados de los ítems, acerca de los conceptos que no conocían y sobre las confusiones creadas con las alternativas de respuesta. Toda esta información fue anotada para su posterior análisis y posibles mejoras del cuestionario.

Los datos que se obtuvieron de la prueba piloto fueron analizados mediante el programa estadístico confeccionado por los profesores Pedro Morales y Guillermo Rodríguez-Izquierdo en 1992 para la corrección de exámenes de pruebas objetivas de la Universidad Pontificia Comillas de Madrid. El programa proporciona información sobre los niveles de dificultad y discriminación de cada ítem, análisis de distractores, distribución de frecuencias del grupo y las puntuaciones individuales de los sujetos: número de respuestas correctas, puntuación corregida por adivinación y puntuación típica del sujeto. Además informa sobre los valores descriptivos de la prueba, media y desviación típica, así como de su consistencia interna  $\alpha$  de Cronbach.

Se ordenó a los sujetos según su puntuación total en el cuestionario y se seleccionaron los dos grupos extremos correspondientes al 25% de alumnos con

puntuaciones más altas y al 25% con puntuaciones más bajas. Se realizó el análisis de cada uno de los ítems del cuestionario, en primer lugar, tabulando el número de sujetos de cada grupo que eligieron cada una de las alternativas de respuesta del ítem en concreto. Se calculó también el índice de dificultad y los índices de discriminación. Se realizaron análisis sobre la puntuación total del cuestionario y sobre la fiabilidad y la validez.

Esta información sirvió para revisar los ítems con el fin de disminuir el error de la medida de puntuación del cuestionario, se mejoró la redacción de los ítems modificando aquellos cuyo enunciado no resultó claro a los alumnos, se cambiaron alternativas de respuesta por otras más adecuadas, se eliminaron algunos ítems y se incluyeron otros.

#### 5.4.2.1. Análisis descriptivo del cuestionario piloto

En primer lugar, se analiza con la muestra piloto la puntuación total del cuestionario calculándose los estadísticos descriptivos. El cuestionario piloto con 37 ítems tiene una media en el número de respuestas correctas de 18,73 con una desviación típica de 4,23. Se representa a continuación la frecuencia de alumnos en función del número de ítems contestados correctamente (Figuras 25 y 26).

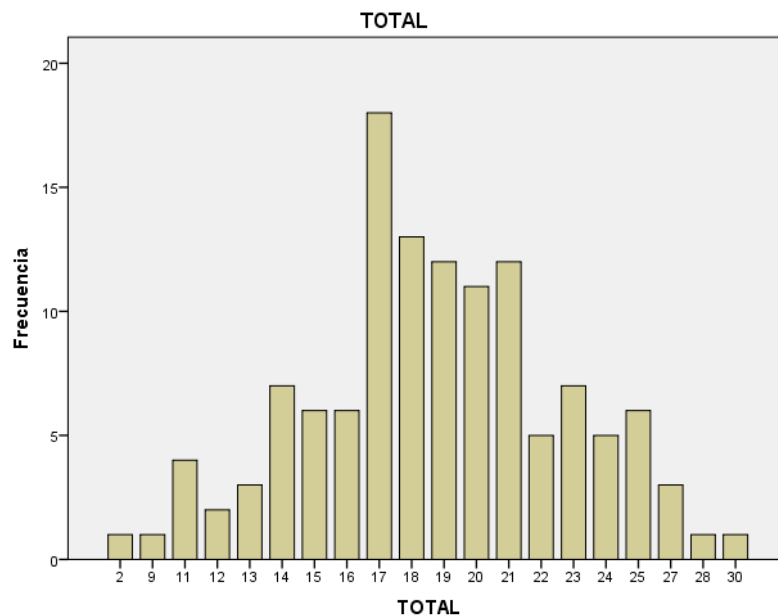


Figura 25. Gráfico de barras de la frecuencia del número de respuestas correctas de CERES piloto.

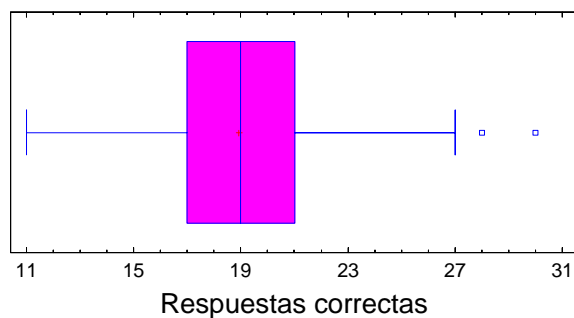


Figura 26. Gráfico de cajas del número de respuestas correctas de CERES piloto.

El intervalo de confianza para la media del número de respuestas correctas por alumno en esta prueba fue  $[18,2395; 19,6315]$ , con una desviación típica dentro del intervalo  $[3,48111; 4,47414]$ , ambos con un 95% de nivel de confianza. El número mínimo de respuestas correctas por estudiante en este cuestionario piloto de 37 preguntas fue de 11 y el máximo, tan sólo en el caso de un alumno, de 30.

Para un primer acercamiento al estudio de la validez discriminante del cuestionario, se dividió a los estudiantes, respecto a su puntuación total en la prueba, en cuatro grupos correspondientes a los cuartiles, seleccionando para el análisis los dos grupos extremos. Es decir, los estudiantes con puntuaciones más bajas, por debajo de 17 (grupo 1) y, el grupo con puntuaciones más altas, superiores a 21 (grupo 2). Cabe esperar que los alumnos con puntuación total más alta tengan mayor facilidad para contestar cada uno de los ítems y con el grupo de puntuaciones más bajas ocurrirá el caso contrario.

Los resultados comparativos se representan en la figura 27.

El grupo 1 formado por 48 alumnos tuvo una media de 14,75 respuestas correctas con una desviación típica de 2,84, mientras que el grupo 2 de 40 alumnos obtuvo una media de 23,3 preguntas correctas con una desviación típica de 2,26. La mediana de los grupos 1 y 2 fue respectivamente 15,5 y 23,0.

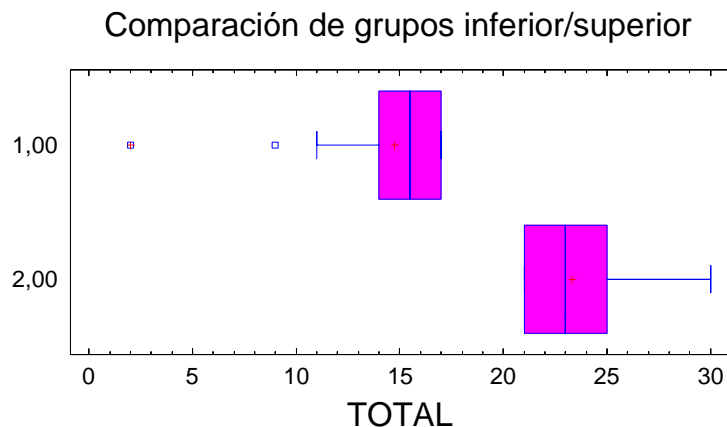


Figura 27. Gráfico de cajas de la comparación de grupos extremos de CERES piloto.

En cuanto al número medio de respuestas correctas en el cuestionario, el grupo 2, correspondiente a los alumnos de puntuaciones más altas, supera en 9 puntos al grupo inferior y el error típico de la media de este grupo es algo menor.

El contraste t para la diferencia de medias, junto con la prueba de Levene para contrastar la homogeneidad de las varianzas, ponen de manifiesto diferencias significativas en el total de la escala ( $t = -15,68$  y  $p\text{-valor} = 0,00$ ) y en el 62% de los ítems, lo que indica una adecuada capacidad de discriminación del total del cuestionario y de gran parte de los ítems.

#### 5.4.2.2. Análisis de ítems del cuestionario piloto

Se presentan en la tabla 6 los resultados obtenidos para cada ítem, mostrando un resumen de los porcentajes de respuestas correctas, así como los índices de dificultad y discriminación y la correlación de cada ítem con el total del cuestionario.

Cabe destacar, también, el análisis de los distractores para el que se fijaron los siguientes criterios en cuanto al número de respuestas:

- Distractor que predomina sobre la respuesta correcta y sobre el resto de distractores.
- Distractor que domina sólo respecto al resto de distractores.
- Distractor que no es elegido nunca
- Distractor elegido en mayor porcentaje por los alumnos del grupo de puntuaciones más bajas.



- Distractor elegido en mayor porcentaje por los alumnos del grupo de puntuaciones más altas.

En general, la dificultad de los ítems es aceptable, aunque algunos de ellos resultaron para los alumnos demasiado fáciles o demasiado difíciles.

El índice de dificultad<sup>17</sup> se calcula dividiendo el número de alumnos que contestan correctamente entre el número de personas que realizó la prueba. Multiplicado por cien indica la proporción de aciertos en la muestra utilizada. Se calcula con los grupos superior e inferior, es decir con el 25% de sujetos de puntuaciones más bajas y el 25% de sujetos de puntuaciones más altas.

A medida que este índice se aproxima a uno o a cero significa que el ítem no tiene capacidad para aportar información acerca de las diferencias entre los individuos que realizan la prueba, ya que se trataría de un ítems que todos contestan correctamente (muy fácil) o por el contrario que nadie contesta correctamente (muy difícil).

Las preguntas de dificultad media (DF = 0.5) son las que tienen más capacidad de discriminar.

$$DF = \frac{\text{nº de acertantes del grupo inferior} + \text{nº de acertantes del grupo superior}}{\text{Número de sujetos de los dos grupos}}$$

Este índice se utiliza junto con los índices de discriminación con los que se valora la capacidad del ítem para distinguir a los sujetos, señalando las diferencias en la ejecución de la prueba entre el grupo superior y el grupo inferior.

El cálculo del índice de discriminación<sup>18</sup> más utilizado, al que denominamos DC1 se realiza con la diferencia de los acertantes del grupo superior menos los del grupo inferior.

<sup>17</sup> Morales 2012.

<sup>18</sup> Morales 2012.

$$DC1 = \frac{\text{nº de acertantes del grupo superior} - \text{nº de acertantes del grupo inferior}}{\text{Número de sujetos de los dos grupos}}$$

Este índice toma el valor de cero, tanto para ítem muy fáciles que todos responden correctamente, como en los muy difíciles en los que todos se equivocan. Toma el valor + 0,5 si todos los de grupo superior, y ninguno del inferior contesta correctamente al ítem y toma el valor – 0,5 si sólo y todos los del grupo inferior aciertan la pregunta, por lo que debería ser revisada.

Lo habitual es utilizar este índice de forma comparativa entre todos los ítems de una prueba.

El segundo índice de discriminación utilizado<sup>19</sup> (DC2) nos indica la proporción de aciertos del grupo superior respecto al total de aciertos de los dos grupos.

$$DC2 = \frac{\text{Nº de aciertos en el grupo superior}}{\text{Nº total de acertantes}}$$

El índice de discriminación DC2 indica la proporción de aciertos en el grupo superior con respecto al número total de aciertos, es independiente del grado de dificultad de la pregunta y puede considerarse satisfactorio si es superior a 0,5 (Morales 2012).

Se recogen a continuación (Tabla 6) los valores de estos índices, junto con la correlación de cada ítem con el total del cuestionario, que han resultado de la prueba piloto.

Las casillas que aparecen en color indican que el valor del índice en cuestión no está dentro de los valores satisfactorios que ha de tener dicho índice.

<sup>19</sup> Morales 2012.

Tabla 6. Análisis de ítems de CERES piloto.

ANÁLISIS DE ÍTEMS DE CERES PILOTO						
ÍTEM	DF	DC1	DC2	ANÁLISIS DISTRACTORES	% RESPUESTAS CORRECTAS	CORRELACIÓN ÍTEM-TOTAL
1	0,47	0,23	0,62	Domina el distractor C sobre el resto de distractores.	45,96	,093
2	0,95	0,10	0,53	No se eligen los distractores.	96,77	,296
3	0,53	0,61	0,79	Domina el distractor B en el grupo de alumnos de puntuaciones más bajas.	56,45	,366
4	0,76	0,29	0,60		75	,263
5	0,52	0,52	0,75		48,38	,352
6	0,61	0,39	0,66		51,61	,258
7	0,16	- 0,0	0,40	Domina el distractor D sobre la respuesta correcta y el resto de distractores.	17,74	-,085
8	0,15	0,10	0,67	Domina el distractor A sobre la respuesta correcta y el resto de distractores.	10,48	,100
9	0,50	0,10	0,55		51,61	-,010
10	0,03	0,06	1,00	Domina el distractor B sobre la respuesta correcta y el resto de distractores.	3,22	,089
11	0,29	0,19	0,67	Domina el distractor B sobre la respuesta correcta y el resto de distractores.	26,61	,052
12 CON RESPUESTAS VERDADERO/ FALSO	0,87	0,19	0,56		91,12	,237
13 CON RESPUESTAS VERDADERO/ FALSO	0,69	0,32	0,61	Distractor con alto porcentaje de respuestas en el grupo de alumnos de puntuaciones más bajas.	76,61	,225
14 CON RESPUESTAS VERDADERO/ FALSO	0,60	0,42	0,68	Domina el distractor en el grupo de alumnos de puntuaciones más bajas.	57,25	,218

ÍTEM	DF	DC1	DC2	ANÁLISIS DISTRACTORES	% RESPUESTAS CORRECTAS	CORRELACIÓN ÍTEM-TOTAL
15	0,45	0,39	0,71	Domina el distractor C en el grupo de alumnos de puntuaciones más bajas. No elijen el distractor E.	48,38	,199
16	0,32	0,32	0,75	Domina el distractor C sobre el resto de distractores. No elijen el distractor E.	26,61	,100
17	0,15	0,16	0,78	No elijen el distractor E.	15,32	,049
18	0,71	0,45	0,66		66,12	,311
19	0,60	0,35	0,65	Domina el distractor B en el grupo de alumnos de puntuaciones más bajas.	62,09	,207
20	0,92	0,16	0,54		93,54	,280
21	0,29	0,32	0,78	Domina el distractor B sobre la respuesta correcta y el resto de distractores.	30,64	,192
22	0,26	0,19	0,69	Domina el distractor C sobre la respuesta correcta y el resto de distractores.	17,74	,055
23	0,68	0,52	0,69		70,96	,331
24	0,61	0,45	0,68	Domina el distractor C en el grupo de alumnos de puntuaciones más bajas.	60,48	,166
25 CON RESPUESTAS VERDADERO/ FALSO	0,37	0,10	0,57	Domina el distractor.	32,25	,016
26 CON RESPUESTAS VERDADERO/ FALSO	0,50	-0,0	0,48	Domina el distractor.	54,83	-,106
27 CON RESPUESTAS VERDADERO/ FALSO	0,55	0,19	0,59	Distractor con alto porcentaje de respuestas.	41,93	,055
28 CON RESPUESTAS VERDADERO/ FALSO	0,61	0,32	0,63	Domina el distractor en el grupo de alumnos de puntuaciones más bajas.	63,70	,244

ÍTEM	DF	DC1	DC2	ANÁLISIS DISTRACTORES	% RESPUESTAS CORRECTAS	CORRELACIÓN ÍTEM-TOTAL
29 CON RESPUESTAS VERDADERO/ FALSO	0,79	0,16	0,55		78,22	,148
30 CON RESPUESTAS VERDADERO/ FALSO	0,61	0,32	0,63	Domina el distractor en el grupo de alumnos de puntuaciones más bajas.	55,64	,103
31	0,81	0,32	0,60	Distractor con alto porcentaje de respuestas en el grupo de alumnos de puntuaciones más bajas.	79,03	,265
32	0,39	0,52	0,83	Domina el distractor C sobre el resto de distractores.	35,48	,294
33	0,61	0,26	0,61		67,74	,096
34	0,48	0,58	0,80		47,58	,309
35	0,37	0,29	0,70	Domina el distractor D sobre el resto de distractores.	45,16	,132
36	0,18	- 0,1	0,36	Domina el distractor D sobre la respuesta correcta y el resto de distractores.	12,90	-,047
37	0,71	0,19	0,57	Domina el distractor A sobre el resto de distractores.	78,22	,168

Los índices de dificultad de los ítems son muy diversos, están comprendidos entre 0,03 y 0,95. La dificultad media para el cuestionario, calculada con los valores de este índice es de 0,51 y la mediana 0,53.

Lo idóneo en una prueba es utilizar ítems de dificultad media con buenos índices de discriminación de cara a maximizar la fiabilidad de la prueba aunque siempre hay que tener en cuenta que en los exámenes, así como en el resto de instrumentos, es necesario priorizar la validez y que el conjunto de ítems represente los conocimientos que deben adquirir los alumnos, aunque algunos de ellos hayan sido adquiridos por todos los estudiantes.

El cuestionario incluye algunos ítems sobre los que, de antemano, teníamos la certeza que los alumnos mostrarían un porcentaje alto de aciertos antes de comenzar su primer curso universitario. Estos ítems tienen como objetivo poner de manifiesto el razonamiento y los conceptos aprendidos en

niveles de enseñanza previos a la universidad así como aquellos que forman parte de la alfabetización estadística que todo ciudadano posee como consecuencia de recibir información estadística de forma habitual y por distintas vías.

Por otro lado, como ya se ha comentado anteriormente, esta prueba además de servir como instrumento de medida para la presente investigación, tuvo como objetivo obtener una calificación de los alumnos que constituirá su nota oficial de la asignatura de estadística. Al elaborar este cuestionario no podemos dejar de pensar como un profesor que tiene esta responsabilidad y desde ese punto de vista la prueba debía contener preguntas que resultarán fáciles al estudiante porque evalúan conocimientos mínimos y preguntas con gran dificultad para detectar y recompensar el esfuerzo de los alumnos excelentes.

Estas preguntas muy fáciles o muy difíciles no discriminan y por tanto no contribuyen a la fiabilidad pero en palabras de Morales (2012), no quiere decir que necesariamente sean malas preguntas. En algunos casos, la facilidad del ítems responde al objetivo de medir el conocimiento previo que tiene el alumno y su capacidad de razonamiento en cuestiones sencillas por lo que han de estar en el cuestionario. Los más difíciles, es cierto que recogen aspectos que tienen que ver con conceptos complejos o malentendidos que parecen ser difíciles de corregir, a juzgar por los resultados de este cuestionario, y teniendo en cuenta que el grupo de sujetos que lo contestaron habían finalizado la asignatura de estadística. Y por otro lado entendemos que la formulación se ha realizado de forma cuidadosa por lo que los datos del análisis nos aportan información pero no es determinante para la eliminación de un ítem (Burton, 2001, citado en Morales 2012).

En la observación de los índices de dificultad en conjunto, el valor del coeficiente de asimetría estandarizado de los índices de dificultad de este cuestionario (-0.424), está dentro del rango esperado para una distribución normal, al igual que ocurre con el coeficiente de curtosis estandarizado con un valor de -0.684.

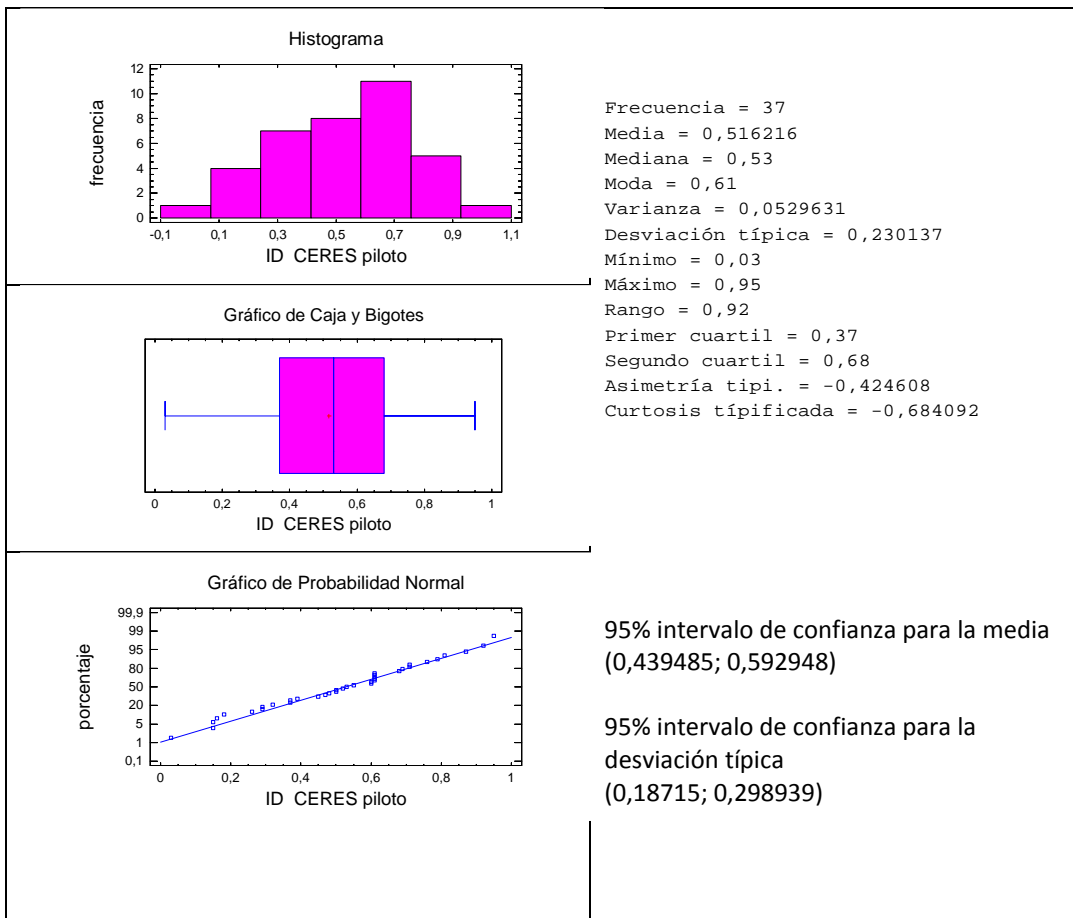
Para verificar el supuesto de normalidad de los índices de dificultad, se realizó una prueba de bondad de ajuste, Tabla 7, resultando adecuado aceptar la normalidad de los datos, al menos con un 90% de confianza.

**Tabla 7. Prueba de bondad de ajuste para la normalidad de los índices de dificultad del Cuestionario de Evaluación de Razonamiento Estadístico piloto (CERES piloto).**

Estadístico	EDF	Valor	Forma Modificada	P-Valor
Kolmogorov-Smirnov	D	0,101555	0,630909	>=0.10*
Anderson-Darling	A^2	0,26465	0,27045	0,6763*

\* significación estadística

En la Figura 28 se presentan estos resultados.



**Figura 28. Descripción y análisis de normalidad de los índices de dificultad del Cuestionario de Evaluación de Razonamiento Estadístico piloto (CERES piloto).**

Los ítems con índice de dificultad por encima de 0,9 son el ítem 2 y el 20. El ítem 2 evalúa la capacidad de reconocer el gráfico de cajas y patillas que representa a un conjunto de datos y prácticamente la totalidad de los alumnos respondieron correctamente, por lo que su índice de discriminación es bajo. El ítem 20 evalúa la capacidad de asociar la descripción de la relación entre dos

variables con su representación en un diagrama de dispersión. Éstas son dos cuestiones que resultan sencillas para los estudiantes con un alto porcentaje de respuestas correctas y con unas medias de estos ítems para toda la muestra de 0,94 y 0,93 respectivamente.

Otros ítems que mostraron ser fáciles ( $DF > 0,70$ ) son los número 4, 12, 18, 20, 29, 31 y 37. Todos ellos con índice de dificultad superior a 0,7.

El ítem 4 forma parte de un grupo de preguntas que valoran la capacidad de asociar un histograma a la descripción de una variable. En este caso la variable es el número de abortos previos, representado en un diagrama de barras. Los alumnos en este caso reconocieron sin dificultad una variable discreta y el gráfico más adecuado para ser representada.

El ítem 12 evalúa la capacidad para comparar grupos mediante una representación gráfica en la que utilizará, para la comparación, la zona en la que se centran la mayor parte de los datos, al margen de la existencia de valores extremos. Se trata de un ítem con respuesta verdadero/falso.

El ítem 18 pretende que los alumnos demuestren que comprenden el significado de la variabilidad en un conjunto de datos obtenido como resultado de medidas repetidas. Este ítem, debido a la facilidad que mostraron los alumnos para resolverlo correctamente, es un claro ejemplo de que la variabilidad, como hecho natural, es percibida por el alumno, que tiene además la capacidad de interpretarla.

Los ítems 29 y 31, también son dos de los 10 ítems del cuestionario planteados con respuesta de verdadero o falso. En el caso del ítem 29 se trata de detectar una afirmación errónea acerca de la interpretación del intervalo de confianza, mientras que el 31 plantea una interpretación correcta del mismo.

El ítem 37, que ha mostrado también facilidad en su resolución, requiere que los estudiantes interpreten los resultados de una prueba de significación cuando es rechazada la hipótesis nula.

A continuación se analizan los ítems que han resultado difíciles para los alumnos ( $DF < 0,30$ ), son el 7, 8, 10, 11, 17, 21, 22 y 36.

El ítem 7 evalúa la capacidad de asociar la información que ofrece una tabla que muestra los porcentajes de una variable cualitativa, al gráfico que los representa. Este ítem, que evalúa un aspecto sencillo y fundamental de la



interpretación de gráficos, fue corregido al detectar un error en la clave de corrección.

El ítem 8 resultó ser difícil para los alumnos, que en su mayoría eligieron un distractor con un planteamiento muy general, que aporta un elemento de confusión. El ítem evalúa los efectos de la asignación al azar del tratamiento en un diseño experimental. Se analizaron las dudas que los alumnos plantearon a la hora de entender la pregunta y las posibles opciones de respuesta. El análisis permitió comprobar que, este ítem, no resultaba adecuado para el cuestionario y se optó por la eliminación del mismo en el CERES definitivo.

El ítem 10, valora la interpretación de los gráficos de cajas y patillas al entender que no proporcionan estimaciones precisas de los porcentajes de los datos por encima o por debajo de un valor, a excepción de los cuartiles. Se puso de manifiesto que este es un aspecto no bien entendido por la mayor parte de los alumnos. Este ítem tiene el coeficiente de dificultad más bajo del cuestionario (0,03), y sólo dos alumnos pertenecientes al grupo más alto, lo respondieron correctamente.

Así mismo el ítem 11 referido a los mismos gráficos y que evalúa la comprensión de la mediana también presentó cierta dificultad en la elección de la respuesta correcta.

El ítem 17 mostró tener dificultad para los alumnos en entender que las muestras aleatorias pequeñas tienen mayor variabilidad que las muestras aleatorias grandes.

El ítem 21 planteado para recoger información acerca de la capacidad del alumno para describir la relación entre dos variables que se muestran en un gráfico de dispersión, revela problemas a la hora de interpretar la no existencia de relación lineal, eligiendo la mayoría de los alumnos el distractor que propone una moderada relación lineal.

Los resultados del análisis del ítem 22 ponen de manifiesto un error frecuente que es entender que la correlación implica causalidad. El distractor que recoge esta asociación de conceptos tiene un alto porcentaje de respuestas. Y en el ítem 36 el distractor que propone dos métodos para extrapolar un modelo de regresión en una situación en la que no es apropiado hacerlo, tiene el mayor porcentaje de respuestas sobre la respuesta correcta y sobre el resto de distractores.

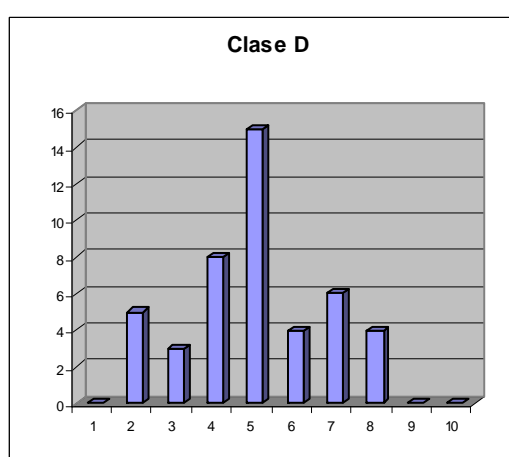
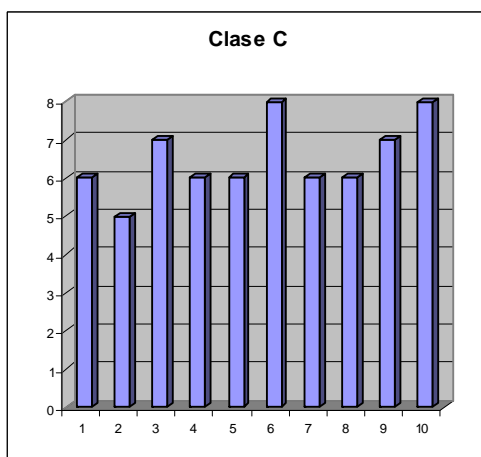
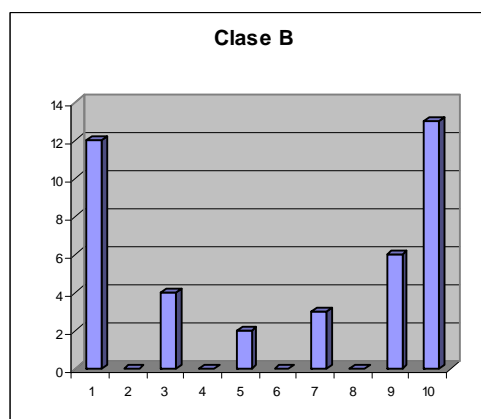
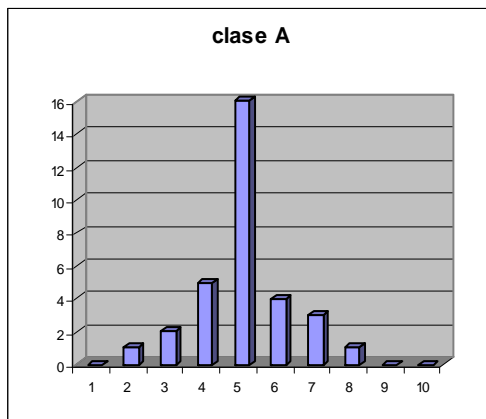
Los ítems 15, 16 y 17, son los únicos del cuestionario que tienen cinco alternativas de respuesta. En todos ellos nunca se elige el distractor E. Esta circunstancia haría pensar en eliminar este último distractor que posiblemente esté forzado y no aporte nada a la pregunta.

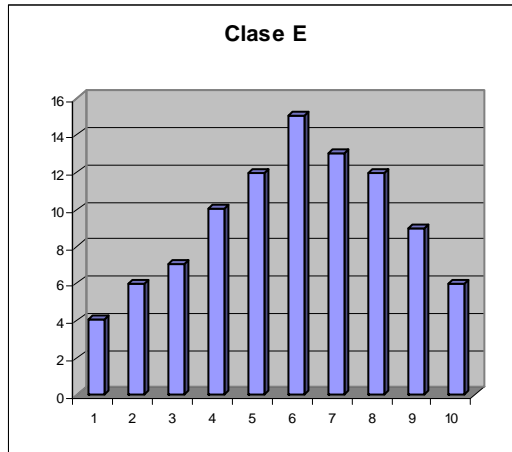
A continuación se transcribe estos tres ítems en los que se una vez analizados estos distractores no elegidos por ningún alumno, se valoró que constituyen una opción válida y se decidió mantenerlos por los siguientes motivos:

En el caso de los ítems 15 y 16, en los que se presentan cinco gráficos sobre los que se pregunta por la magnitud de su desviación típica, es idóneo formular un distractor por cada una de las cinco opciones gráficas planteadas.

### Los ítems 15 y 16 se refieren a la siguiente situación:

Se presentan a continuación cinco gráficos, cada uno de ellos muestra el resultado en una escala de 0 a 10 de los exámenes de cinco clases de estadística.





15. ¿Cuál de las clases se espera que tenga una menor desviación típica, y por qué?

- a. **La clase A, ya que tiene la mayoría de los valores cercanos a la media.**
- b. La clase B, porque tiene un menor número de resultados distintos.
- c. La clase C, porque no hay variaciones en las puntuaciones.
- d. Las Clases A y D porque los dos tienen el menor rango.
- e. La clase E porque parece el que más se ajusta a la normalidad.

16. ¿Cuál de las clases se espera que tenga una mayor desviación típica, y por qué?

- a. La clase A, ya que cuenta con la mayor diferencia entre las alturas de las barras.
- b. **La clase B, ya que varios de sus resultados están lejos de la media.**
- c. La clase C, porque tiene el mayor número de puntuaciones diferentes.
- d. La clase D, porque la distribución es muy desigual e irregular.
- e. La clase E porque tiene mayor rango y se ajusta a la normalidad.

En el caso del ítem 17 el quinto distractor lejos de ser una opción de relleno, nos sigue pareciendo una opción válida pues completa las opciones de respuesta posible y por ello se mantiene en el cuestionario.

Ítem 17. Una empresa fabrica mascarillas de polipropileno para protección sanitaria en cinco colores de las que el 50% son de color blanco. Las comercializa en cajas de 50 y en cajas de 500 que se rellenan al azar.

¿Qué caja es más probable que contenga más del 70% de mascarillas de color blanco?

- La de 500 porque tiene más mascarillas, por lo que puede contener más de color blanco.
- La de 500, porque hay mayor variabilidad en la proporción de mascarillas blancas en muestras grandes.
- La de 50, ya que hay mayor variabilidad en la proporción de mascarillas blancas en muestras pequeñas.**
- La de 50, porque la mayor parte de las cajas pequeñas tienen más del 50% de mascarillas blancas.
- Las dos tienen la misma probabilidad porque son muestras al azar.

En la tabla 8 se presenta una distribución de las frecuencias de los índices de dificultad de los ítems del cuestionario. Se han establecido unos intervalos para la interpretación de este índice.

Tabla 8. Distribución de ítems de CERES piloto según su nivel de dificultad.

DISTRIBUCIÓN DE ÍTEMS SEGÚN SU NIVEL DE DIFICULTAD		
Índice de dificultad	Porcentaje de ítems del cuestionario	Dificultad
< 0,2	5 13.5%	Difíciles
0,2 – 0,4	7 19 %	Mediana dificultad
0,4 – 0,6	10 27%	Dificultad media
0,6 – 0,8	11 29,7%	Fáciles
0,8 -1	4 10,8%	Muy fáciles

Respecto a la diferencia de la proporción de aciertos entre el grupo superior e inferior, que corresponde al índice de discriminación (DC1) recogido en la Tabla 9, se observa que 13 ítems tienen un DC1 entre 0,2 y 0,4 por lo que podríamos decir que, aunque se haya considerado un intervalo un poco amplio

para la interpretación del mismo, estos ítems poseen capacidad de discriminar<sup>20</sup> y 8 de ellos lo hacen de forma excelente.

En 16 ítems apenas hay diferencias en el número de aciertos entre los alumnos del grupo superior y del inferior, lo que quiere decir que no discriminan adecuadamente.

Tabla 9. Índices de discriminación de los ítems de CERES piloto

INDICE DE DISCRIMINACIÓN DE CERES PILOTO		
Índice de discriminación DC1	Frecuencia de ítems	Porcentaje
< 0,2	16	43,2%
0,2-0,4	13	35,1%
0,4-0,6	8	21,6%

El cuestionario tiene un índice medio de discriminación entre [0,606995; 0,687059] lo que puede ser considerado como aceptable. Si observamos la distribución de frecuencias de este índice en el total del cuestionario (tabla 10) observamos que la mayor parte de los ítems tienen suficiente capacidad discriminatoria (DC2 > 0,5). Se ha probado también la normalidad de este índice (Figura 29).

Tabla 10. Frecuencias de los Índices de discriminación (DC2) de los ítems de CERES piloto.

ÍNDICE DE DISCRIMINACIÓN DE CERES PILOTO		
Índice de discriminación DC2	Frecuencia de ítems	Porcentaje
0,35 – 0,44	2	5,4%
0,44 – 0,53	2	5,4%
0,53 – 0,62	12	32,4%
0,62 – 0,72	13	35%
0,72 – 0,81	6	16,2%
0,81 – 0,90	1	2,7%
0,90 – 1,00	1	2,7%

95,0% intervalo de confianza para la media: [0,606995;0,687059]  
 95,0% intervalo de confianza para la desviación típica: [0,0976392;0,155961]

<sup>20</sup> Ebel y Frisbiel 1986. DC1> 0,29 discriminación y DC1> 0,39 excelente discriminación.

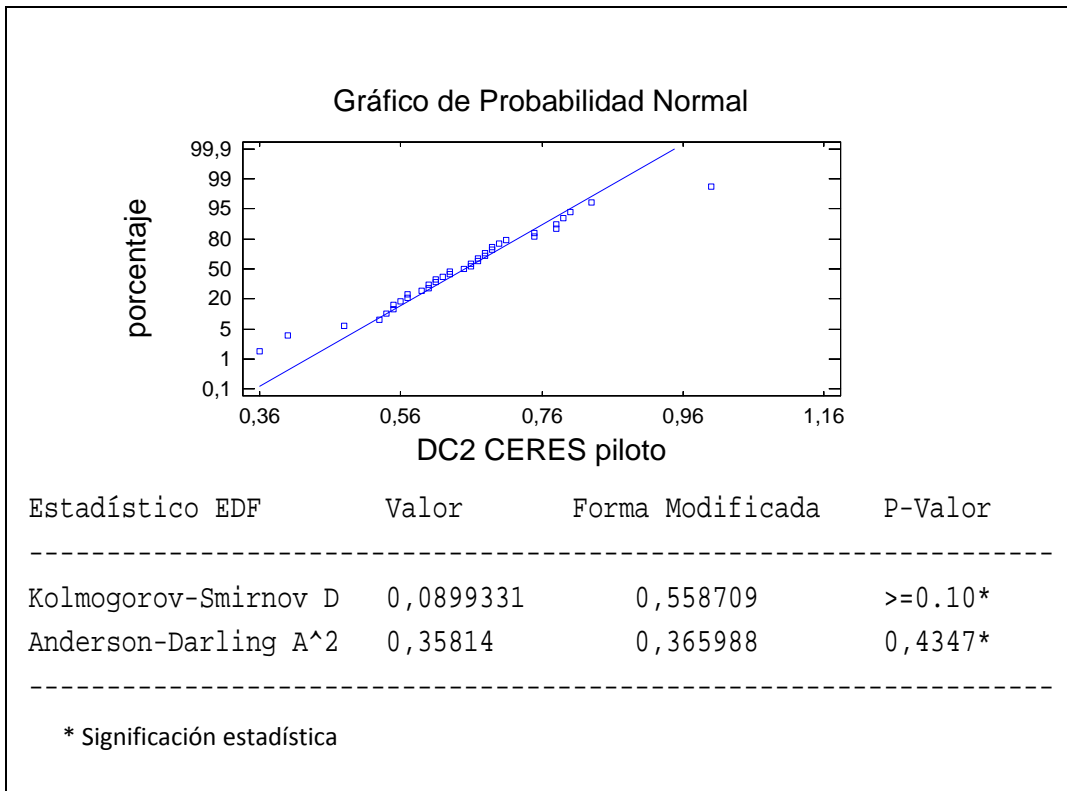


Figura 29. Estudio de normalidad de los índices de discriminación (DC2) de los ítems de CERES piloto.

Se analizaron las correlaciones de los ítems con el total del cuestionario que aportan información, también, acerca de la discriminación y de la homogeneidad del cuestionario y se basa en los datos de todos los sujetos de la muestra piloto.

Los valores de estas correlaciones resultaron ser significativas en 16 de los ítems, en los que el coeficiente fue superior a 0,2.

Algunos ítems mostraron correlaciones negativas pero cercanas a cero (Tabla 6). Estos ítems son el 7, 9, 26 y 36 y fueron de nuevo revisados. Uno de ellos, el ítem 7, contenía un error en la clave de corrección como se ha comentado anteriormente. El resto se analizan a continuación:

En el análisis del ítem 36 (ítem-total = -0,47) se observa que domina el distractor D sobre el resto de distractores y sobre la respuesta correcta.

	A	B	C	D	E	F	OMIT.	ERR.	DF	DC1	DC2
ITEM 36 S	2	4	4*	21	-	-	0	27			
	---	---	---	---	---	---	---	---	0,19	-0,1	0,33
I	9	2	8*	12	-	-	0	23			
R036T	-0,22	0,11	-0,05	0,18							

Es una cuestión que no demuestran conocer los alumnos. Consideramos que el ítem 36, está bien formulado y en las opiniones recogidas ningún alumno mostró duda acerca de lo que se le preguntaba, por lo que hay que pensar que estamos detectando un error en la comprensión de las situaciones en las que no puede ser extrapolado un modelo de regresión.

Ítem 36. Un estudio en el ámbito universitario utilizó datos de los últimos 15 años para relacionar los costes de los proyectos de investigación con los tiempos de ejecución en días. Se obtuvo la siguiente ecuación de regresión para los datos disponibles en los que el proyecto de mayor duración fue de 3,5 años.

$$\text{Coste (€)} = 2142 + 35 \cdot \text{Tiempo (días)}$$

¿Cuál es el mejor método para predecir el coste de un proyecto de 5 años de duración?

- Sustituir el valor de 1825 días (5 años) en el tiempo en la ecuación de regresión y calcular el coste previsto.
- Trazar la línea de regresión en el diagrama de dispersión, localizar 1825 en el eje horizontal, y leer el valor correspondiente al coste en el eje vertical.
- Ningún método basado en estos datos es apropiado para hacer una predicción para un proyecto de 5 años.**
- Ambos métodos son apropiados para hacer una predicción del coste basado en estos datos.

El ítems 26 (ítem-total= -0,1) pertenece a un grupo de preguntas de respuesta verdadero o falso en la que se presentan tres interpretaciones del p-valor. En el caso del ítem 26 se valora la capacidad de reconocer una interpretación errónea del p-valor. En estos tres ítems (26, 26 y 27), planteados sobre una misma cuestión, los coeficientes de correlación son próximos a cero y en el ítem 26 además es negativo..

Un artículo de investigación ha publicado los resultados de un estudio sobre los efectos para combatir la infección por el virus del papiloma humano (HPV) en mujeres VIH-positivas, de una combinación de medicamentos contra el VIH. El artículo recoge un p-valor de 0,03 en el análisis estadístico. Los ítems 25, 26 y 27 presentan tres diferentes interpretaciones de este p-valor. Indique en cada una de ellas si es o no válida.

Ítem 26. Es la probabilidad de que ocurran los resultados obtenidos si la hipótesis nula es verdadera.

- a. **Válido.**
- b. No válido.

Ítem 27. Es la probabilidad de que el tratamiento no sea eficaz

- a. Válido.
- b. No válido.**

Ítem 28. Es la probabilidad de que el tratamiento sea eficaz

- a. Válido.
- b. No válido.**

El ítem 9 (ítem-total= -0,01), tiene un índice de dificultad medio (DF=0,5), el 51, 61% de los alumnos responden correctamente y hay alumnos de ambos grupos que conocen la respuesta correcta.

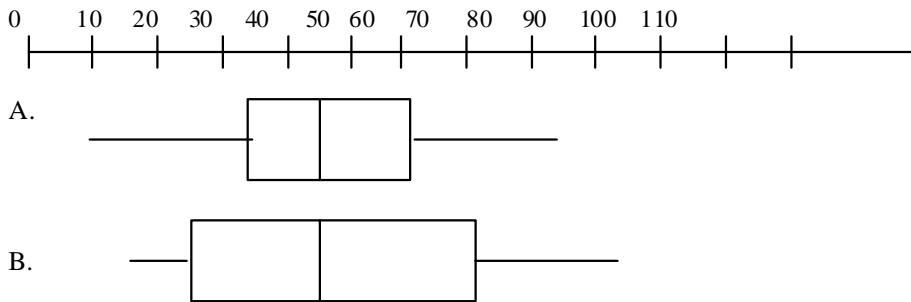
	A	B	C	D	E	F	OMIT.	ERR.	DF	DC1	DC2
ITEM 009 S	1	17*	8	5	-	-	0	14			
	---	---	---	---	---	---	---	---	0,50	0,10	0,55
I	8	14*	5	2	-	-	2	17			
R009T	-0,17	-0,01	0,14	0,03							

Se ha revisado el contenido de la pregunta, las opciones de respuesta y la clave de corrección, entendiendo que todo ello es correcto. También consideramos que es adecuado para un cuestionario de razonamiento estadística el resultado de aprendizaje que evalúa este ítem: entender el concepto de



desviación típica como un parámetro que informa de la dispersión, así como la capacidad de interpretarlo en el contexto de un diagrama de cajas y patillas.

Los dos siguientes diagramas de cajas muestran los resultados de un estudio sobre la amnesia postraumática tras una lesión en la cabeza. La variable estudiada es el tiempo de hospitalización en días. Recogiéndose en el siguiente gráfico los datos de dos grupos que se diferencian en el tipo de lesión que sufrieron en la cabeza.



Ítem 9. ¿Qué grupo se espera que tenga una mayor desviación típica respecto al tiempo de hospitalización?

- a. Grupo A.
- b. Grupo B.**
- c. Ambos grupos son casi iguales.
- d. Es imposible saberlo.

Como resultado de las pruebas con el cuestionario piloto se realizaron las siguientes modificaciones:

- Se corrigieron los errores en la clave de corrección del ítem 7.
- Se eliminó el ítem 8
- Se incluyeron 4 ítem nuevos, formulados de acuerdo a las sugerencias de los profesores de estadística consultados como expertos.
- Se realizaron pequeños cambios que mejoraron la redacción de alguno de los ítems sin alterar su contenido.
- Se cambió el orden de los ítems, de acuerdo a un criterio de complejidad y al orden en el que los contenidos han sido expuestos en las sesiones de clase. La correspondencia de numeración entre las dos versiones se recoge en la tabla 11.

**Tabla 11 . Correspondencia de numeración y modificaciones entre las dos versiones del cuestionario.**

CERES DEFINITIVO	CERES PILOTO
Ítem 1 es un ítem nuevo	
Ítem 2	Ítem 1
Ítem 3	Ítem 2
Ítem 4 es un ítem nuevo	
Ítem 5 es un ítem nuevo	
Ítem 6	Ítem 7
Ítem 7 es un ítem nuevo	
Ítems 8 – 11	Ítems 3 – 6
Ítems 12 y 13	Ítems 15 y 16
Ítem 14	Ítem 32
Ítems 15 – 17	Ítems 9- 11
Ítem 18	Ítem 35
Ítems 19 – 21	Ítems 12 - 14
Ítem 22	Ítem 17
Ítem 23	Ítem 18
Ítems 24 y 25	Ítems 33 y 34
Ítems 26 – 28	Ítems 25 - 27
Ítem 29	Ítem 19
Ítems 30 y 31	Ítems 23 y 24
Ítem 32	Ítem 37
Ítems 33 – 36	Ítems 28 -31
Ítem 37	Ítem 20
Ítem 38	Ítem 21
Ítem 39	Ítem 22
Ítem 40	Ítem 36
	Ítem 8 (se elimina)

Como conclusión del estudio del cuestionario piloto podemos indicar la idoneidad de los ítems y de los distractores para la evaluación del razonamiento

estadístico. Se ha tenido en cuenta la búsqueda de un cuestionario que resultara con una dificultad media y una máxima discriminación, pero no se han eliminado ítems que se han mostrado fáciles o muy difíciles por la necesidad de asignar calificaciones a los estudiantes.

### 5.4.3. Validez de constructo

Para analizar la validez de constructo y comprobar si el instrumento mide el concepto teórico de razonamiento estadístico se ha utilizado un análisis factorial exploratorio del conjunto de las puntuaciones de los ítems del cuestionario. Con este análisis se pone de manifiesto la existencia de un constructo subyacente que permite clarificar la estructura del instrumento.

Respecto al análisis factorial del cuestionario se comprueba en primer lugar el ajuste del análisis factorial. La prueba de esfericidad de Bartlett muestra valores significativos con  $p < 0,01$ , aunque con un  $KMO < 0,5$ , que indica que los ítems no están correlacionados.

Se realizó la extracción de factores mediante el método de componentes principales y rotación Varimax, conservando las comunalidades y la suma de la varianza explicada por los factores. Se examinaron las correlaciones entre factores y la matriz de correlaciones entre los ítems para completar el análisis de los mismos.

La extracción inicial (Tabla 12) muestra 14 factores que explican el 68% de la varianza total, aunque son 5 los factores fundamentales tal como se observa en el gráfico de sedimentación (Figura 30).

Tabla 12. Extracción inicial de factores del cuestionario CERES piloto.

Factor	ítems	% de varianza explicada
Factor 1	11, 26, 27	6,174
Factor 2	13, 14, 19	5,541
Factor 3	3, 23, 33, 34	5,457
Factor 4	30, 31	5,384
Factor 5	4, 5, 6, 21	5,225
Factor 6	20, 36, 37	4,928
Factor 7	32	4,806
Factor 8	22, 35	4,794

Factor	ítems	% de varianza explicada
Factor 9	8, 10, 15	4,612
Factor 10	7, 9, 16	4,482
Factor 11	2, 18, 24, 28	4,427
Factor 12	1, 12	4,410
Factor 13	25	4,141
Factor 14	29	3,715
Varianza total explicada		68,094%

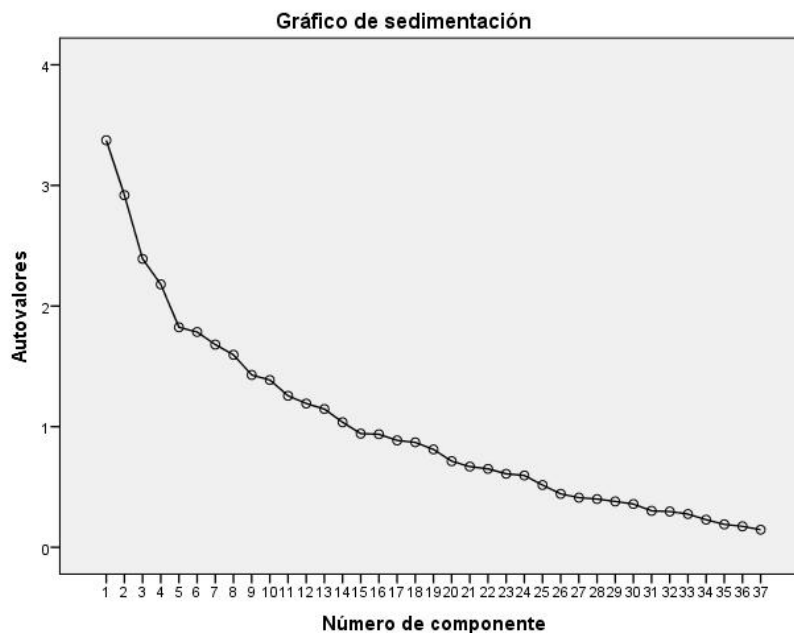


Figura 30. Gráfico de sedimentación del análisis factorial de CERES piloto.

La rotación ortogonal Varimax deja más clara la estructura del cuestionario (Tabla 13), quedando agrupados en el mismo factor los ítems que tienen mayores correlaciones entre sí que con los demás. En este caso sólo serían 4 los grupos compuestos por al menos tres ítems con pesos por encima de 0,40, el resto serían factores constituidos por uno o dos ítems.

**Tabla 13. Matriz de componentes rotados del análisis factorial del cuestionario CERES piloto.**

**Matriz de componentes rotados<sup>a</sup>**

	Componente													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
ITEM27	,879													
ITEM26	-,874													
ITEM11	,510							,439						
ITEM14		,747												
ITEM13		,640												
ITEM19		,461							-,310			,314		
ITEM34			,746											
ITEM3			,743											
ITEM33			,618			,368								
ITEM23			,423			,365								
ITEM30				,809										
ITEM31				,713										
ITEM17		-,342		,342										
ITEM6					,793									
ITEM5					,686									
ITEM4					,421		,331			,372				-,388
ITEM21					,357							,335		,319
ITEM37						,784								
ITEM36				-,345		-,506								
ITEM20						,506	,300				,315			
ITEM32							,767							
ITEM35								,709						
ITEM22								-,628						
ITEM8									,686					
ITEM10									,682					
ITEM15									,561					
ITEM9										,644				
ITEM7										-,612				
ITEM16										,590				
ITEM2											,734			
ITEM24		,377									,580			
ITEM18							,378				,425			-,327
ITEM28							,356	,340			,388			
ITEM12												,818		
ITEM1		-,485										,616		
ITEM25													,793	
ITEM29														,837

Método de extracción: Análisis de componentes principales.

Método de rotación: Normalización Varimax con Kaiser.

a. La rotación ha convergido en 21 iteraciones.

El análisis de esta matriz, y de los ítems que componen cada factor entre los que no se encuentra una clara relación conceptual que nos permita definir estos factores y en qué aspectos contribuyen a medir el razonamiento estadístico, junto con otros datos del análisis factorial, nos lleva a pensar que el cuestionario no es unidimensional para el constructo medido. Por el contrario, está compuesto de ítems distintos y no suficientemente correlacionados, hecho que es de esperar en una prueba de evaluación de estas características.

#### 5.4.4. Validez convergente.

Se verificó la existencia de correlación entre el cuestionario CERES con otra prueba que ha sido elaborada con objeto de medir también la capacidad de razonar con los conceptos estadísticos. Se trata de una prueba de evaluación del razonamiento estadístico de respuesta abierta parcialmente restringida. De esta

manera se aportan datos para comprobar la validez del constructo que pretende medir el Cuestionario.

#### **5.4.4.1. Prueba de evaluación del razonamiento estadístico de respuesta abierta parcialmente restringida**

Para elaborar esta prueba se seleccionaron algunos aspectos básicos que implican un cierto razonamiento estadístico y se elaboraron cinco preguntas muy concretas, formuladas con referencia a situaciones del ámbito sanitario. Los resultados de aprendizaje que evalúa esta prueba están relacionados con:

- Reconocimiento del análisis estadístico que puede dar respuesta a una pregunta de investigación.
- Correcto planteamiento de las hipótesis.
- Representación de una variable que se distribuye de forma normal e interpretación de la puntuación típica de un sujeto con respecto al grupo.
- Interpretación de pruebas de significación.

La calificación de la prueba se realizó de acuerdo a una clave de corrección que recoge una serie de criterios para la ponderación de las preguntas.

La corrección de esta prueba tiene un enfoque analítico (Morales 1995). Se dispone de una clave de corrección en la que se especifican los diversos aspectos que deben aparecer en la respuesta para considerarla correcta. De esta manera se intenta reducir los aspectos subjetivos de la corrección y facilita la calificación matizándola y haciendo que sean más fácilmente comparables los resultados

Esta prueba fue realizada con la muestra piloto, sólo en una ocasión, el mismo día que realizaron el examen práctico, como parte de la evaluación final de la asignatura de estadística de la convocatoria ordinaria. Se les indicó que deberían contestar exclusivamente en los espacios destinados a las respuestas, forzando así a los alumnos a realizar un proceso de síntesis y concreción de las mismas.

Se incluye en los anexos esta prueba y su clave de corrección (ANEXO 6), para ilustrar mejor su estructura y características.

El análisis de correlación resultó significativo ( $r=0,301$  ;  $p < 0,01$ ), tal como se esperaba ya que ambas pruebas han sido construidas con el mismo objetivo: medir el razonamiento estadístico.

#### **5.4.4.2. Correlación con las calificaciones de la prueba de resolución de problemas**

Por otro lado se comprueba también la existencia de correlación con las calificaciones de una prueba práctica de resolución de problemas. Al finalizar la asignatura, los alumnos tienen que demostrar sus conocimientos aplicados a la resolución de problemas estadísticos utilizando un programa informático. Se trata de un examen de carácter práctico que realizan en el aula de informática. El formato de esta prueba consiste en contestar una serie de preguntas para las que necesitan extraer la información contenida en una base de datos sanitaria, aplicar las pruebas estadísticas adecuadas, e interpretar los resultados.

Para aportar más información en la confirmación de la validez del instrumento se realiza con la muestra piloto un análisis de correlación entre las puntuaciones del Cuestionario de Evaluación del Razonamiento Estadístico (CERES) y las calificaciones del examen práctico. Se obtuvo una  $r= 0.313$  con un valor de  $p < 0.001$  lo que nos indica que existe una buena asociación entre esta prueba y la prueba construida de razonamiento estadístico que nos aporta información sobre la adecuada validez del instrumento, así como de su utilidad práctica.

#### **5.5. Análisis de Fiabilidad**

Se analizó la consistencia interna del instrumento calculando el coeficiente Alfa de Cronbach como apoyo a la interpretación de que todos los ítems miden el mismo constructo y al grado de confianza que podemos tener en las puntuaciones como estimadores de las puntuaciones verdaderas de los sujetos.

El valor obtenido de este coeficiente es de 0,614 que dada la complejidad que tiene el constructo, y las indicaciones acerca del límite de este coeficiente para pruebas de evaluación<sup>21</sup>, puede considerarse aceptable. Resultados como este, de un valor moderado, son frecuentes en cuestionarios que evalúan conocimientos complejos y muy amplios.

La prueba T-cuadrado de Hotelling sobre la hipótesis nula de que todos los ítems del cuestionario tienen la misma media, nos pone de manifiesto que existen diferencias significativas entre las medias de los ítems con un valor de  $T=2673.135$  y un valor de  $p < 0.001$ , lo que parece corresponder con la variedad de índices de dificultad encontrado entre los ítems de esta prueba.

El coeficiente de correlación intraclase es significativo tanto para las medidas individuales ( $r=0.041$ ,  $p < 0.001$ ), como para los valores promedio ( $r=0.614$ ,  $p < 0.001$ ). El alto valor de éste último es indicador de la fiabilidad de una sola medida. El intervalo de confianza para el índice de fiabilidad sería (0,510; 0,705).

---

<sup>21</sup> No hay un valor mínimo adecuado para aceptar el coeficiente de fiabilidad, Morales (2007, p. 15-16). En este mismo autor se citan las distintas propuestas clásicas que van desde 0,50, Guilford (1954:388-389) hasta 0,70, Nunnally (1978).

Santisteban (1990) indica para cuestionarios de evaluación un valor del coeficiente de 0,5 como límite general aceptable.

Es importante señalar que en la valoración de este coeficiente ha de tenerse en cuenta la complejidad del constructo. En nuestro instrumento, que se requieren ítems diferenciados que valoren diversos y múltiples resultados de aprendizaje, los ítems estarán menos relacionados y la consistencia interna será menor.



## CAPÍTULO 6. INTERVENCIÓN Y RECOGIDA DE DATOS

### 6.1. Descripción de la intervención

La decisión acerca de la naturaleza de la intervención tiene su origen en las directrices que marca la American Statistical Association (ASA) para el aprendizaje y la evaluación de la estadística en alumnos universitarios (Guidelines for Assessment and Instruction in Statistics Education GAISE 2010).

El curso de estadística sobre el que se realiza este estudio se estructura en su metodología docente sobre las seis recomendaciones que recogen el documento anteriormente mencionado (GAISE, 2012) pero una de estas recomendaciones constituirá la intervención sobre un grupo de alumnos: El uso de la evaluación no sólo para valorar el aprendizaje sino para mejorarlo (GAISE, 2012).

Aunque no se pretenda la evaluación es una actividad de aprendizaje (Morales, 2012) pero en nuestro caso hemos utilizado la evaluación de forma intencional como una actividad formativa con la intención de promover el aprendizaje y mejorar el rendimiento académico (Black y Wiliam 1998). Además, siguiendo otra de las recomendaciones de la ASA, en nuestra intervención nos interesa evaluar la comprensión conceptual, el razonamiento de la estadística. Las pruebas de evaluación que se utilizan en la intervención recogen aspectos propios de las ciencias de la salud y de la profesión enfermera. Los ítem de estos cuestionarios están adaptados para proporcionar al estudiante un escenario más próximo que refuerce los conceptos estadísticos abstractos y los procedimientos, con contenidos significativos con los que pueda relacionarlos y percibirlo como una adquisición de habilidades y conocimiento propios de su profesión (Derry,S., Levin, J. y Schauble, L. 1995).

La evaluación es uno de los aspectos de mayor relevancia en los cambios producidos en la manera de concebir la enseñanza de la estadística, utilizándola como una herramienta para mejorar el aprendizaje de los estudiantes. Por este motivo se han utilizado a lo largo del curso estas pruebas de razonamiento estadístico como parte de una evaluación formativa con el objeto de recoger información acerca de los progresos en el aprendizaje y para orientar al alumno a un estudio razonado. Se ha utilizado la actividad de evaluación como una oportunidad para apoyar el aprendizaje a la vez que promueve la autorregulación del mismo.

Como ya se dijo anteriormente, la intervención consistió en que uno de los grupos de alumnos realizó a lo largo del curso cuatro pruebas parciales (bloques diferentes de contenidos), de evaluación del razonamiento estadístico:

PaRES1. Estadística descriptiva.

PaRES2. Distribución normal.

PaRES3. Pruebas de significación.

PaRES4. Análisis de varianza y regresión.

Los dos grupos de alumnos de primer curso de enfermería, grupo de intervención (grupo A) y grupo control (grupo B), aunque con distinto horario, eran coincidentes en el cronograma semanal de la asignatura, de manera que fue posible abordar en clase, los mismos contenidos cada una de las semanas que constituyeron el curso.

La realización de estas pruebas parciales de razonamiento estadístico (PaRES) por parte del grupo de intervención, tuvo lugar siempre en la sesión de clase siguiente a la finalización de la exposición teórica de dichos contenidos y a la realización de ejercicios prácticos en el aula de informática. Los alumnos fueron informados a comienzo de curso de la realización de estas pruebas de evaluación formativa. Se les advirtió acerca de la obligatoriedad de realizarlas para poder superar la asignatura, a pesar de que su calificación no formaría parte de la nota final de estadística.

El objetivo de la utilización de pruebas de evaluación del razonamiento estadístico como intervención persigue:

- Que los alumnos **mantengan un estudio continuo** de la asignatura al introducir en el proceso de enseñanza-aprendizaje un elemento de atención a través de la evaluación que se realiza periódicamente.
- Que los alumnos **reciban información de sus logros y sus lagunas en el aprendizaje**, al tiempo que refuerzan los conocimientos de la asignatura. Posteriormente a la realización de las pruebas se mostró a los alumnos el resultado de aprendizaje pretendido en cada una de las preguntas, se fundamentó la respuesta correcta y se analizaron los errores habituales que comenten los alumnos recogidos en los distractores de los ítems de las pruebas.

- Que los alumnos realicen un **entrenamiento de su capacidad de razonar y relacionar conceptos** al enfrentarse a este tipo de pruebas de evaluación.
- Que los alumnos **orienten su estudio hacia el razonamiento**, evitando que adquieran los conocimientos exclusivamente de forma memorística o por repetición.
- Que los alumnos **mejoren su aprendizaje**.

Para realizar una enseñanza de la estadística que consiga el desarrollo del razonamiento en los alumnos, ha de tenerse en cuenta, por un lado, la elección de una metodología docente, cuidada y pensada para ayudar y, en algunos casos, forzar al alumno a pensar. Por otro lado, la evaluación utilizada debe medir su razonamiento y comprensión de la estadística, dejando los aspectos memorísticos o de cálculo en segundo plano.

Las opciones metodológicas en las que hemos fundamentado la enseñanza de la estadística de los alumnos que constituyen la muestra del presente trabajo se basan, como se ha referido previamente, en las directrices para la enseñanza y evaluación de la estadística (GAISE, 2010). Su idoneidad y conveniencia de hacerlas extensibles a la enseñanza de la estadística en todos los niveles formativos de Enfermería ha sido señalada recientemente: “El informe GAISE es un excelente recurso para integrar el conocimiento y los hallazgos de las investigaciones realizadas sobre la enseñanza de la estadística en la formación de los estudiantes de enfermería”. (Hayat, 2014).

GAISE (2010) se erige por tanto como un glosario estructurado de las conclusiones resultantes del movimiento de reforma de la enseñanza de la estadística que intenta fomentar, en los profesores, la renovación de los cursos introductorios de estadística para hacerlos más participativos y dotarles de una mayor identidad (Zieffler, Park, Garfield, delMas, & Audbjorg, 2012).

Las opciones metodológicas que hemos seleccionado para alcanzar nuestro objetivo de desarrollar el razonamiento estadístico y que han sido empleadas con todos los alumnos, tanto del grupo experimental como del grupo control se exponen a continuación.

- Hacer evidente a los alumnos de enfermería, desde el primer momento de su formación estadística, la **relevancia y aplicabilidad** que tiene para el ejercicio de su profesión.

- Se han utilizado **bases de datos reales** procedentes de la actividad sanitaria, datos procedentes de investigaciones en el ámbito de la salud y datos que se recogen directamente del grupo de alumnos.

Usar datos reales para la realización de ejercicios prácticos y de la exposición teórica de conceptos, aumenta el interés del estudiante y mantiene su atención en el proceso de aprendizaje.

Con datos reales los estudiantes pueden, de primera mano, plantearse cómo y por qué han sido recogidos los datos y qué análisis han de utilizarse para responder la pregunta de investigación.

Una estrategia es obtener una base de datos de los propios estudiantes. Para ello se informó a los alumnos, el primer día de clase, acerca de esta cuestión y se les facilitó el enlace para acceder y responder a un cuestionario online con 35 preguntas, solicitándoles que lo cumplimentaran en los días posteriores. Dicho cuestionario tenía carácter anónimo y su objetivo fue generar datos con diferentes tipos de medida: cualitativa y cuantitativa discreta y continua.

Las preguntas del cuestionario recogen información sobre aspectos relacionados con el estudiante tales como datos antropométricos, otros referentes a sus gustos y aficiones o a la anotación del resultado de un suceso aleatorio como el lanzamiento de un dado o una moneda.

El alumno responde a las preguntas eligiendo una de las opciones facilitadas o con respuesta abierta.

Algunas de las preguntas fueron ¿Cuál es la red social que más utilizas? ¿Cuánto tiempo empleas en las redes sociales?, ¿te gusta el videoclip al que puedes acceder a través del siguiente enlace?, se le pregunta acerca de algunos rasgos genéticos, ¿eres capaz de rular la lengua?, etc. Responder al cuestionario fue sencillo para los alumnos y su utilización para la realización de pruebas estadísticas resultó motivador y ameno al despertarles curiosidad por conocer algunas características de su propio grupo.

Este cuestionario se recoge en el ANEXO 7 ya que puede servir como un ejemplo útil para otros docentes.

También se utilizan datos provenientes de trabajos de investigación de enfermería, algunos de ellos pertenecientes a estudios realizados por nosotros y otros que fueron cedidos por equipos de investigación

de centros sanitarios. Una ventaja del uso de estos datos es que nos permite explicar a los estudiantes el proceso de obtención de la muestra y el objetivo de la investigación, al tiempo que se les aporta información sobre el ámbito real de la investigación enfermera, qué y cómo se investiga en esta disciplina.

- Tal como se desprende del anterior apartado se utilizó **el trabajo con problemas y casos en escenarios auténticos** que promueven la curiosidad y por ello la búsqueda de información.
- El **desarrollo de las capacidades de análisis, y el aprendizaje comprensivo**. Para ello es importante el desarrollo de procesos reflexivos que se consiguen estructurando la información que se facilita al alumno, ayudándole en el análisis de sus elementos y de las relaciones existentes entre ellos y con situaciones precedentes.
- La utilización, como herramienta, de un **programa estadístico informático** por lo que las técnicas de cálculo que deben aprender son las estrictamente necesarias para la comprensión de los conceptos y del fundamento de las pruebas estadísticas.

Estos programas informáticos de estadística por su rapidez en la generación de análisis y múltiples gráficos se utilizan también de forma muy eficaz como ayuda para visualizar conceptos, y comprender ideas abstractas.

- Utilizar **métodos de aprendizaje activos** que involucren a los estudiantes. La **comunicación de información a partir de tablas y gráficos**, como aspecto importante para el desarrollo del razonamiento estadístico. La **elaboración de informes**, para los que debe organizar la información y comunicarla, promoviendo el desarrollo de las habilidades de pensamiento crítico y la discusión. Se les pide, por ejemplo, que elijan un artículo de investigación de enfermería y presenten un resumen previamente discutido por un grupo de alumnos.

La metodología utilizada, que hemos expuesto en los párrafos anteriores, se ajusta a las cinco recomendaciones del informe GAISE. La sexta recomendación, referida al **uso de evaluaciones para mejorar el aprendizaje** y no sólo para valorarlo, corresponde a la intervención que se realizó exclusivamente con el grupo experimental.

Tras determinar el diseño de este estudio, y optar por una intervención en uno de los grupos de la asignatura de estadística, era de importancia capital

asegurarse de que la forma en la que se organizara el trabajo en las dos aulas y las condiciones principales en las que había de producirse el aprendizaje fueran idénticas en los dos grupos. Esta circunstancia hace muy difícil poder involucrar a otros docentes que aporten datos de sus alumnos y aumentar así el tamaño de la muestra. Este fue uno de los motivos por los cuales el estudio se realizó con los grupos de alumnos de la asignatura de estadística de un solo curso académico. La otra razón es que se presuponía de antemano que la utilización de pruebas periódicas de razonamiento estadístico facilitarían en los estudiantes la comprensión de los conceptos por lo que no se consideró la posibilidad de no utilizarlas en otro grupo del siguiente curso.

Cuando investigamos en un entorno natural, como es el aula universitaria, nos encontramos con las limitaciones que dificultan un estudio experimental ya que es complicado disponer de un grupo de control equivalente<sup>22</sup> o de hacer una asignación aleatoria, más allá de la que hemos realizado en este trabajo. Es decir designar aleatoriamente cuál de los dos grupos de los que se dispone será el grupo de intervención y cuál el grupo control. No es frecuente encontrar estudios empíricos con un diseño experimental cuando se realiza con grupos naturales.

Lo que sí es común es que los docentes realicen investigaciones acerca del proceso de enseñanza-aprendizaje cuando se introduce algún tipo de innovación en uno de sus cursos. Es frecuente que se trate de dos grupos en los que imparte clase el mismo profesor y hay múltiples trabajos en los que el tamaño de la muestra no es muy grande, tal y como ocurre en el presente estudio.

Morales (2013) recoge algunos ejemplos de comparación entre dos grupos distintos, uno de ellos sobre el que se ha aplicado una innovación didáctica o de evaluación (grupo experimental) y otro grupo de las mismas características en el que se ha seguido el método más tradicional (grupo control). Este planteamiento común en educación, en ocasiones se aplica sobre grupos de reducido tamaño, utilizando un grupo natural y de conveniencia que son los alumnos del profesor investigador. Los ejemplos de estudios de estas

---

<sup>22</sup> No se trata de un verdadero grupo control que no recibe el tratamiento. En los estudios educativos con grupos naturales, el grupo control recibe también enseñanza, por lo que se comparan dos tipos diferentes de enseñanza, en lugar de comparar un tratamiento con la ausencia de este, (Kember, 2003).

características que se encuentran en la literatura son muchos pues se trata de un diseño que se utiliza con frecuencia en educación (Figura 31).

Estudio de Investigación educativa	Grupo control	Grupo experimental
Connor-Greene P. (2000). Assessing and Promoting Student Learning: Blurring the Line Between Teaching and Testing. <i>Teaching of Psychology</i> , Vol. 27, nº 2, 84-88.	Sistema convencional de evaluación N=25	Sistema innovador de evaluación N=30
Myers, C. y Myers S. (2007). Assessing Assessment: The Effects of Two Exam Formats on Course Achievement and Evaluation. <i>Innovative Higher Education</i> . 31: 227-236	Sistema convencional de evaluación N=63	Sistema innovador de evaluación N=68
Payne, E. y Brown G. (2011). Communication and practice with examination criteria. Does this influence performance in examinations? <i>Assessment &amp; Evaluation in Investigación experimental, diseños y contraste de medias Higher Education</i> Vol. 36, No. 6, 619-626	Evaluación No se conocen los criterios de corrección N=32	Evaluación Criterios de corrección conocidos N=54
Rabin, L. y Nutter-Upham, K. (2010). Introduction of a Journal Excerpt Activity Improves Undergraduate Students' Performance in Statistics. <i>College Teaching</i> . Vol. 58 Issue 4, p156-160	Actividad didáctica tradicional N= 35	Actividad didáctica innovadora N= 32
Roscoe, L. y Strapp, C. (2009). Increasing Psychology Students' Satisfaction With Preparedness Through a Professional Issues Course. <i>Teaching of Psychology</i> , 36: 18-23	Satisfacción con los estudios N=52	Satisfacción con los estudios Con asignatura para la inserción laboral N=19
Schuetzes, P. (2004). Evaluation of a Brief Homework Assignment Designed to Reduce Citation Problems Teaching of Psychology, Vol. 31 Issue 4, 257-262.	Evaluación sobre problemas al realizar citas N=40	Evaluación sobre problemas al realizar citas Con ejercicios para reconocer el plagio N=35
Saville, B.; Zinn, T.; Brown, A. y Marchuk, K. (2010) Syllabus Detail and Students' Perceptions of Teacher Effectiveness <i>Teaching of Psychology</i> , 37: 186-189	Percepción que se tiene del profesor Entrega de un programa de la asignatura sencillo N= 47	Percepción que se tiene del profesor Entrega de un programa de la asignatura detallado N= 50
Stickels, J. y Dobbs, R.(2007). Helping Alleviate Statistical Anxiety with Computer Aided Statistical Classes. <i>The Journal of Scholarship of Teaching and Learning</i> , Vol. 7, No. 1, pp. 1-15.	Actitudes Metodología tradicional (calculadora) N=40	Actitudes Metodología innovadora (uso del programa SPSS) N=48
Winniger, S. (2005). Using Your Tests to Teach: Formative Summative Assessment. <i>Teaching of Psychology</i> , Vol. 32, (3), 164-166	Resultados del examen repetido por segunda vez N=57	Resultados del examen repetido por segunda vez feedback  pormenorizado de los resultados de un examen N=37

Figura 31. Ejemplos de estudios de investigación educativa con intervención realizados con grupos pequeños.

Fuente: Adaptado de Morales, P. (2013). Investigación experimental, diseños y contraste de medias. <http://web.upcomillas.es/personal/peter/investigacion/Dise%1osMedias.pdf>

## 6.2. Pruebas parciales de razonamiento estadístico (PaRES).

A continuación se describen las pruebas de evaluación de razonamiento estadístico utilizadas en la intervención.

Para componer estas pruebas parciales se ha llevado a cabo una selección de ítems extraídos de las once pruebas ARTIST. Han sido traducidos y adaptados al contexto de Ciencias de la Salud y se han incluido también ítems nuevos. El paralelismo de los ítems de las distintas pruebas parciales con los cuestionarios ARTIST originales se recoge en las siguientes Figuras 32, 33, 34 y 35.

Contenidos	PaRES1	Ítems de las escalas ARTIST que conforman PaRES1
Recopilación de datos	ÍTEM 1	Adaptación del ítem 1 de la escala ARTIST.
	ÍTEM 2	Adaptación del ítem 2 de la escala ARTIST.
	ÍTEM 3	Adaptación del ítem 8 de la escala ARTIST.
Medidas de centro	ÍTEM 4	Adaptación del ítem 1 de la escala ARTIST.
	ÍTEM 5	Adaptación del ítem 2 de la escala ARTIST.
	ÍTEM 6	Adaptación del ítem 3 de la escala ARTIST.
	ÍTEM 7	Adaptación del ítem 4 de la escala ARTIST.
	ÍTEMS 8 y 9	Adaptación de los ítems 5 y 6 de la escala ARTIST.
Medidas de dispersión	ÍTEM 10	Adaptación del ítem 1 de la escala ARTIST.
	ÍTEMS 11 y 12	Adaptación de los ítems 3 y 4 de la escala ARTIST.
	ÍTEM 13	Adaptación del ítem 5 de la escala ARTIST.
	ÍTEM 14	Adaptación del ítem 6 de la escala ARTIST.
	ÍTEMS 15 y 16	Adaptación de los ítems 7 y 8 de la escala ARTIST.
	ÍTEM 17	Adaptación del ítem 9 de la escala ARTIST.
	ÍTEM 18	Adaptación del ítem 10 de la escala ARTIST.
Resumen de datos- Tablas de frecuencias	ÍTEM 19	Ítem nuevo creado para el cuestionario.
Medidas de dispersión	ÍTEM 20	Adaptación del ítem 13 de la escala ARTIST.
	ÍTEM 21	Adaptación del ítem 14 de la escala ARTIST.
Representación de datos	ÍTEM 22	Adaptación del ítem 3 de la escala ARTIST:
	ÍTEM 23	Adaptación del ítem 8 de la escala ARTIST.
	ÍTEMS 24 y 25	Adaptación de los ítems 7 y 13 de la escala ARTIST.

Figura 32. Ítems de las escalas ARTIST que conforman PaRES1.



Contenidos	PaRES2	Ítems de las escalas ARTIST que conforman PaRES2
Distribución normal y medidas de posición	ÍTEM 1	Adaptación del ítem 1 de la escala ARTIST.
	ÍTEM 2	Adaptación del ítem 2 de la escala ARTIST.
	ÍTEMS 3 – 5	Adaptación de los ítems 3 y 4 de la escala ARTIST e ítem nuevo basado en los anteriores (ítem 5).
	ÍTEM 6	Adaptación del ítem 5 de la escala ARTIST.
	ÍTEM 7	Adaptación del ítem 6 de la escala ARTIST.
	ÍTEM 8	Adaptación del ítem 7 de la escala ARTIST.
	ÍTEM 9	Adaptación del ítem 8 de la escala ARTIST.
Intervalos de confianza de una muestra	ÍTEM 10	Adaptación del ítem 1 de la escala ARTIST.
	ÍTEM 11	Adaptación del ítem 2 de la escala ARTIST.
	ÍTEM 12	Adaptación del ítem 3 de la escala ARTIST.
	ÍTEM 13	Adaptación del ítem 5 de la escala ARTIST.
	ÍTEM 14	Adaptación del ítem 6 de la escala ARTIST.
	ÍTEM 15	Adaptación del ítem 8 de la escala ARTIST.

Figura 33. Ítems de las escalas ARTIST que conforman PaRES2.

Contenidos	PaRES3	Ítems de las escalas ARTIST que conforman PaRES3
Test de significación	ÍTEM 1	Adaptación del ítem 1 de la escala ARTIST.
	ÍTEM 2	Adaptación del ítem 3 de la escala ARTIST.
	ÍTEMS 3 – 5	Adaptación de los ítems 5 y 6 de la escala ARTIST e ítem nuevo basado en los anteriores (ítem 5)
	ÍTEM 6	Ítem nuevo creado para el cuestionario
	ÍTEM 7	Adaptación del ítem 7 de la escala ARTIST.
	ÍTEM 8	Adaptación del ítem 8 de la escala ARTIST.
	ÍTEM 9	Adaptación del ítem 9 de la escala ARTIST.
	ÍTEM 10	Adaptación del ítem 10 de la escala ARTIST.
	ÍTEMS 11 y 12	Ítems nuevos creados para el cuestionario.
Tamaño del efecto	ÍTEM 13	Ítem nuevo creado para el cuestionario.
	ÍTEM 14	Ítem nuevo creado para el cuestionario.
Dos variables de datos cualitativos	ÍTEM 15	Adaptación del ítem 1 de la escala ARTIST.
	ÍTEM 16	Adaptación del ítem 2 de la escala ARTIST.
	ÍTEM 17	Adaptación del ítem 3 de la escala ARTIST.
	ÍTEM 18	Adaptación del ítem 4 de la escala ARTIST.

Contenidos	PaRES3	Ítems de las escalas ARTIST que conforman PaRES3
	ÍTEM 19	Adaptación del ítem 5 de la escala ARTIST.
	ÍTEM 20	Adaptación del ítem 6 de la escala ARTIST.
	ÍTEM 21 y 22	Adaptación de los ítems 8 y 9 de la escala ARTIST.
	ÍTEM 23	Ítem nuevo creado para el cuestionario.

Figura 34. Ítems de las escalas ARTIST que conforman PaRES3.

Contenidos	PaRES4	Ítems de las escalas ARTIST que conforman PaRES4
Datos bivariados, cuantitativa	ÍTEM 1	Adaptación del ítem 1 de la escala ARTIST.
	ÍTEM 2	Tomado de la escala ARTIST: (ítem 2).
	ÍTEM 3	Adaptación del ítem 3 de la escala ARTIST.
	ÍTEMS 4 – 6	Tomados de la escala ARTIST: (ítems 4, 5 y 6).
	ÍTEMS 7 y 8	Tomados de la escala ARTIST: (ítems 7 y 8).
	ÍTEM 9	Adaptación del ítem 9 de la escala ARTIST.
	ÍTEMS 10 y 11	Tomados de la escala ARTIST: (ítems 10 y 11).
Análisis de varianza	ÍTEM 12	Adaptación del ítem 13 de la escala ARTIST.
	ÍTEMS 13 y 14	Ítems nuevos creados para el cuestionario.
	ÍTEM 15	Ítem nuevo creado para el cuestionario.

Figura 35. Ítems de las escalas ARTIST que conforman PaRES4.

De los 78 ítems que componen los cuatro cuestionarios parciales (PaRES), 7 de ellos son de nueva creación, 4 son tomados de las escalas ARTIST sin ninguna modificación sobre el ítem original, y los 67 restantes son ítems que han sido modificados y adaptados en su enunciado y en la situación e información estadística que exponen, al ámbito sanitario.

Para la ordenación de los ítems de las pruebas parciales (PaRES), se ha tenido en cuenta la progresión lógica del aprendizaje de la estadística, respetando la sucesión de complejidad de los conceptos, y se ha usado un criterio pedagógico respecto al orden en que se suceden los contenidos de la asignatura de Estadística a la hora de exponerlos a los alumnos.

Los contenidos que se evalúan en cada uno de los cuestionarios tienen también en cuenta los criterios anteriormente expuestos agrupándose en cada uno de estos cuestionarios, ítems relacionados con conceptos y pruebas

estadísticas que comparten características comunes en los fundamentos de los análisis.

### 6.2.1. Resultados de aprendizaje de las pruebas parciales de razonamiento estadístico (PaRES).

En las Figuras 36, 37, 38 y 39 se enuncian los resultados de aprendizaje que evalúan los ítems que componen los cuestionarios parciales de razonamiento estadístico. De esta manera es posible comprobar que la selección de ítems de las cuatro pruebas cubre los contenidos de la asignatura de estadística y son adecuados para el desarrollo del razonamiento estadístico.

Consideramos que estas pruebas contienen ítems que alcanzan de forma más amplia, que el cuestionario CERES, todos los contenidos de un curso básico de estadística.

PRIMER CUESTIONARIO PARCIAL DE DE RAZONAMIENTO ESTADÍSTICO (PaRES1)		
ESTADÍSTICA DESCRIPTIVA		
CONTENIDOS	RESULTADOS DE APRENDIZAJE	ÍTEMS
VARIABLES	Capacidad de reconocer que una variable es cualitativa a pesar de presentarse codificada numéricamente	1 y 2
MUESTREO	Capacidad para seleccionar el muestro más apropiado para una investigación y entendimiento de que un muestreo aleatorio es más representativo	3
MEDIDAS DE TENDENCIA CENTRAL	Compresión del concepto de media aritmética y sus propiedades	4
	Comparar y reconocer las medidas de tendencia central de una distribución de datos que no es simétrica. Entender el concepto de sesgo y asociarlo al valor de la mediana	5
	Reconocer en un histograma la asimetría y el sesgo de una distribución de datos a la hora de asignar los valores correspondientes a las medidas de tendencia central, media y mediana	6
	Comprender el concepto de mediana de una distribución de datos	7 y 8
	Comprender la diferencia entre las dos medidas de tendencia central más importantes, la media y la mediana, respecto a verse afectadas por los valores extremos de la distribución	9
	Aplicar el concepto y cálculo de la mediana a un conjunto de datos en el que todos ellos distan de la media una sola unidad	15
	Entender el valor máximo que puede tener la mediana utilizando los valores máximo y mínimo de un conjunto de datos	16

CONTENIDOS	RESULTADOS DE APRENDIZAJE	ÍTEMS
MEDIDAS DE DISPERSIÓN. VARIABILIDAD. <ul style="list-style-type: none"> <li>• Rango de la variable</li> <li>• Desviación típica</li> <li>• Rango intercuatílico</li> </ul>	Comprender el concepto de desviación estándar y entender que cuando la desviación típica tiene un valor de cero, todos los datos de la muestra tienen el mismo valor	10
	Asociar la desviación estándar al grado de homogeneidad de los datos y la proximidad de los mismos al valor de la media, reconociendo la mayor o menor variabilidad de dos conjuntos de datos representados en un histograma	11
	Reconocer que dos muestras tienen la misma desviación estándar aunque se representen en histogramas que aparentemente tienen diferentes dispersión de los datos	12
	Comprender el concepto y el cálculo de la desviación típica reconociendo que no puede tener valores negativos	13
	Entender la desviación estándar como medida del grado de homogeneidad de una población respecto a una variable	14
	Reconocer al rango como un indicador débil de la dispersión ya que sólo utiliza para su cálculo dos puntuaciones que pueden ser atípicas	18
	Comprender que la introducción o eliminación de valores extremos en un conjunto de datos no afecta al valor del rango intercuatílico	21
REPRESENTACIONES GRÁFICAS	Reconocer sobre un histograma con puntuaciones atípicas, la conveniencia de utilizar la mediana y el rango intercuatílico para describir el centro y la dispersión de la distribución, frente a la media y desviación típica que afectadas por valores extremos pueden no representar al grupo	17
	Comprensión de la representación de un conjunto de datos en un gráfico de cajas y capacidad para interpretar comparativamente el rango de varios conjuntos de datos	20
	Capacidad de seleccionar el gráfico más adecuado para representar una variable cualitativa	22
	Capacidad para interpretar información relativa a los valores de frecuencias de una variable representada en un histograma	23
	Reconocer las variables representadas en cada uno de los ejes de un histograma	24
	Capacidad de reconocer un conjunto de datos en dos representaciones gráficas diferentes	25

Figura 36. Resultados de aprendizaje que evalúa el cuestionario parcial de razonamiento estadístico PaRES1.

SEGUNDO CUESTIONARIO PARCIAL DE RAZONAMIENTO ESTADÍSTICO (PaRES2)		
DISTRIBUCIÓN NORMAL E INTERVALOS DE CONFIANZA		
CONTENIDOS	RESULTADOS DE APRENDIZAJE	ÍTEMS
PUNTUACIONES TÍPICAS	Entender el concepto de puntuación típica	1
	Comprensión del concepto de puntuación típica y aplicarlo para situar a un sujeto en el conjunto de datos, respecto a la media y al posible valor de la variable para dicho sujeto	6
	Interpretación del valor que tiene en un sujeto una variable que se distribuye de forma normal respecto al grupo estudiado	7
	Comprensión del concepto de puntuación típica que utiliza la desviación típica como unidad de medida y permite comparar resultados de procedencia distinta	8
	Entender la desviación típica como unidad de medida que permite situar la puntuación de un sujeto respecto al resto del grupo	9
CARACTERÍSTICAS Y PROPIEDADES DE LA DISTRIBUCIÓN NORMAL	Comprensión de las propiedades de la distribución normal y su simetría. Conocer que en muestras grandes la amplitud o recorrido de una variable comprende unas seis desviaciones típicas	2
DISTRIBUCIÓN NORMAL Y PROBABILIDAD	Comprensión de las propiedades de la distribución normal: Entender que si conocemos en cuántas desviaciones típicas se distancia de la media un sujeto, podemos conocer la probabilidad y las proporciones de casos esperados por encima o por debajo de dicho sujeto	3
	Asociar la puntuación de un sujeto respecto a una variable que se distribuye de forma normal con el percentil correspondiente a dicha puntuación, conociendo las proporciones de individuos esperados por encima de dicho sujeto	5
DISTRIBUCIÓN NORMAL Y DESVIACIÓN TÍPICA	Comprensión de las propiedades de la distribución normal en cuanto a la proporción de sujetos que se encuentran a distancia de una, dos o tres desviaciones típicas	4
ERROR TÍPICO INTERVALO DE CONFIANZA	Comprensión del concepto de error típico de la media y entender que los intervalos de confianza para la media calculados en muestras grandes son más precisos que en muestras pequeñas	10
	Entender que un intervalo de confianza para estimar la media poblacional a partir de una muestra, no asegura al 100% que la media poblacional se encuentre en dicho intervalo	11
	Comprender que si extraemos varias muestras del mismo tamaño y calculamos un intervalo de confianza para cada muestra, el 95% de todos los intervalos van a incluir a la vida media poblacional	12

	Interpretar la información que nos aporta un intervalo de confianza del 90% para la estimación de la media poblacional. Entendiendo que con un 90% de seguridad la verdadera media de la población se encontrará dentro del intervalo	13
	Comprender que a mayor nivel de seguridad con el que se construye un intervalo de confianza, los límites del intervalo se alejan, haciendo más ancho el intervalo de estimación	14
	Entender las limitaciones de la estimación por intervalos a partir de muestras no aleatorias	15

Figura 37. Resultados de aprendizaje que evalúa el cuestionario parcial de razonamiento estadístico PaRES2.

TERCER CUESTIONARIO PARCIAL DE RAZONAMIENTO ESTADÍSTICO (PaRES3)			
PRUEBAS DE SIGNIFICACIÓN			
CONTENIDOS	RESULTADOS DE APRENDIZAJE	ÍTEMS	
FORMULACIÓN DE HIPÓTESIS	Comprender la forma correcta de plantear la hipótesis nula y alternativa de un estudio, como hipótesis enfrentadas. Reconociendo en un ejemplo concreto la hipótesis alternativa como aquella que expresa la situación que se pretende comprobar en el proceso de investigación	1	
SIGNIFICACIÓN ESTADÍSTICA	Reconocer cuándo la magnitud de un p-valor pone de manifiesto una significación estadística	2	
	Identificar en el contexto de una prueba de hipótesis el tipo de error asociado a la aceptación o rechazo de la hipótesis nula	3 y 4	
	Reconocer el error de tipo I como el más importante desde el punto de vista de las pruebas de hipótesis en el contexto de un ejemplo que ayuda a entender la gravedad de rechazar la hipótesis nula siendo cierta	5	
	<ul style="list-style-type: none"> <li>Errores en las pruebas de significación</li> </ul>	Entender que un p-valor significativo representa el hecho de que lo que se observa en la muestra no es producto de la variabilidad natural debida al muestreo	6
	<ul style="list-style-type: none"> <li>Estadísticos de contraste</li> </ul>	Reconocer que una diferencia puede ser significativa y sin embargo puede que no sea relevante en la práctica	7
	<ul style="list-style-type: none"> <li>Nivel de significación</li> </ul>	Reconocer situaciones en las que no se cumplen las condiciones necesarias para realizar una prueba de significación	8
		Entender que si no podemos probar la hipótesis alternativa no significa que no sea cierta, sobre todo en muestras pequeñas	9
		Interpretar la información que aporta un p-valor muy pequeño en el contexto de un ejemplo	10
	Interpretar todas las conclusiones que se derivan de no rechazar la hipótesis nula en una comparación de muestras	11	

CONTENIDOS	RESULTADOS DE APRENDIZAJE	ÍTEMS
TAMAÑO DEL EFECTO	Entender que una diferencia que no es estadísticamente significativa puede ser importante y sobre todo en muestras pequeñas en las que puede que existan diferencias no apreciadas en las pruebas de significación es importante aplicar un método que nos permita calcular la magnitud de las diferencias existentes	12
	Interpretar el tamaño del efecto en un contexto de comparación de medias siendo capaz de reconocer la distribución de datos en relación a la distancia en desviaciones típicas que las separan	13
	Comprender que al expresar la diferencia entre medias de forma tipificada pueden ser comparables datos recogidos con instrumentos de medición diferentes	14
ANÁLISIS BIVARIADO. VARIABLES.  TABLAS DE CONTINGENCIA	Reconocer una prueba de independencia de caracteres e interpretar los resultados en términos de asociación entre las variables	15
	Entender el concepto de nivel de significación como referencia para la interpretación de los valores de p. Reconocer en un contraste de proporciones significativo, que existe asociación y no causalidad entre las variables	16
	Detectar errores en una tabla de contingencia que recoge los datos para un análisis de asociación entre variables cualitativas. Entendiendo que las observaciones tienen que ser independientes	17
	Interpretar, para un ejemplo dado, un resultado significativo en una prueba chi-cuadrado de independencia de caracteres	18
	Reconocer que en una prueba chí-cuadrado para examinar relación entre dos variables se tratan datos que corresponden a categorías en las que se clasifican los sujetos	19
	Distinguir entre los dos tipos de frecuencias que se comparan en una prueba de independencia chi-cuadrado	20
	Interpretar las frecuencias que se recogen en una tabla de contingencia en la que se relacionan dos variables, calculando, las proporciones respecto al total de la muestra, de la ocurrencia conjunta de dos sucesos	21
	Interpretar las frecuencias que se recogen en una tabla de contingencia en la que se relacionan dos variables, calculando, las proporciones respecto a una categoría, de la ocurrencia de un suceso	22
	Interpretar para un ejemplo dado, un resultado no significativo en una prueba chi-cuadrado de homogeneidad de muestras	23

Figura 38. Resultados de aprendizaje que evalúa el cuestionario parcial de razonamiento estadístico PaRES3.

CUARTO CUESTIONARIO PARCIAL DE RAZONAMIENTO ESTADÍSTICO (PaRES4)		
REGRESIÓN LINEAL Y ANÁLISIS DE VARIANZA		
CONTENIDOS	RESULTADOS DE APRENDIZAJE	ÍTEMS
<b>ANÁLISIS DE REGRESIÓN Y CORRELACIÓN</b> <ul style="list-style-type: none"> <li>• Asociación de dos variables</li> <li>• Gráficos de dispersión</li> <li>• Estimación de la recta de regresión</li> <li>• Coeficientes de regresión y determinación</li> </ul>	Entender que el coeficiente de correlación es independiente de las unidades de medida de las variables	1
	Relacionar el coeficiente de correlación con el coeficiente de determinación y con el porcentaje de variabilidad común entre dos variables	2
	Interpretar un coeficiente de correlación muy pequeño y distinguir que no existe relación lineal entre las variables pero puede existir una relación de otro tipo entre ellas	3
	Reconocer un gráfico que dispersión que representa dos variables que no están relacionadas linealmente	4
	Seleccionar, entre varios ejemplos, el diagrama de dispersión que muestra la representación que más se aproxima a un valor dado de correlación entre las variables	5
	Identificar, entre varios ejemplos, el diagrama de dispersión que muestra la relación más fuerte entre dos variables	6
	Indicar qué está representado en cada uno de los ejes de un gráfico de dispersión	7 y 8
	Interpretar la magnitud de un coeficiente de correlación positivo, reconociendo que los sujetos tienen simultáneamente valores, altos o bajos en las dos variables	9
	Extraer información de un diagrama de dispersión acerca del grado de asociación entre dos variables	10
	Considerar que la eliminación de un solo sujeto, especialmente si tiene puntuaciones extremas, puede modificar el coeficiente de correlación	11
	Utilizar la ecuación y la recta de regresión para predecir el valor de una variable	12
	Seleccionar la interpretación correcta de un análisis de varianza, utilizando los datos recogidos en una tabla de resultados y relacionándolo con las hipótesis de este tipo de análisis	13
	<b>ANÁLISIS DE VARIANZA</b> <ul style="list-style-type: none"> <li>• Fuentes de variación</li> </ul>	Reconocer e interpretar las fuentes de variabilidad de un análisis de varianza en el contexto de un ejemplo concreto
Distinguir las respuestas que puede aportar un análisis de varianza en el contexto de un ejemplo concreto		15

Figura 39. Resultados de aprendizaje que evalúa el cuestionario parcial de razonamiento estadístico PaRES4.



### 6.2.2. Alineación de los cuestionarios parciales PaRES / GAISE (Guidelines for Assessment and Instruction in Statistics Education GAISE 2010).

Como se ha comentado en capítulos precedentes, estos cuestionarios son los instrumentos con los que se realiza la intervención que es objeto de estudio en esta investigación. Esta evaluación formativa, destinada a orientar el estudio del alumno hacia la comprensión conceptual e integradora de la estadística, constituye también para los estudiantes un entrenamiento en la consecución de unos resultados de aprendizaje basados en el razonamiento estadístico.

Se expone a continuación un acercamiento de alineación con las directrices marcadas por American Statistical Association (ASA), para el desarrollo y evaluación del razonamiento estadístico.

ALINEACIÓN PaRES – GAISE				
GAISE	PaRES (CUESTIONARIOS PARCIALES)			
Los estudiantes deben entender por qué:	PaRES1	PaRES 2	PaRES3	PaRES4
Los datos han de superar las anécdotas				
La variabilidad es natural, previsible y cuantificable				
Un muestreo aleatorio permite que los resultados de las encuestas y experimentos, puedan extenderse a la población de la cual se ha tomado la muestra	Ítem 3			
La asignación aleatoria en experimentos comparativos permite extraer conclusiones respecto a la causa-efecto				
La asociación no es causalidad			Ítem 16	
La significación estadística no implica necesariamente una importancia práctica, especialmente para los estudios con muestras de gran tamaño			Ítem 7	
Al no encontrar una diferencia estadísticamente significativa o relación no significa necesariamente que no hay diferencia o que no haya relación, especialmente para los estudios con muestras pequeñas			Ítem 9, 11. 12	

<b>Los estudiantes deben reconocer:</b>	<b>PaRES 1</b>	<b>PaRES2</b>	<b>PaRES 3</b>	<b>PaRES 4</b>
Las fuentes comunes de sesgo en los estudios y experimentos				
Cómo determinar la población a la que se pueden extrapolar los resultados de la inferencia estadística, sobre la base de cómo se recogieron los datos		Ítem 15		
Cómo determinar una inferencia de causa y efecto de una asociación, basada en cómo se recogieron los datos (diseño del estudio)			Ítem 16	
Que las palabras tales como " normal ", "azar ", y "correlación " tienen significados específicos en estadística diferentes al que tienen en un uso común				
<b>Los estudiantes deben entender las partes de los siguientes procesos estadísticos:</b>	<b>PaRES 1</b>	<b>PaRES 2</b>	<b>PaRES 3</b>	<b>PaRES 4</b>
Cómo obtener o generar datos				
Cómo representar gráficamente los datos como un primer análisis , y cómo saber cuándo es suficiente para responder a la cuestión de interés	Ítems 6, 11, 12, 17, 18, 22, 23, 24, 25			Ítem 4, 5, 6, 10
Cómo interpretar resúmenes numéricos y representaciones gráficas de datos, tanto para responder preguntas y para comprobar las condiciones (uso correcto de procedimientos estadísticos)	Ítem 6, 11, 12, 17, 18, 19, 20, 23		Ítem 17, 20, 21, 22	Ítem 4, 5, 6, 10,
Cómo hacer un uso adecuado de la inferencia estadística			Ítem 5, 8	
Cómo comunicar los resultados de un análisis estadístico				
<b>Los estudiantes deben entender las ideas básicas de la inferencia estadística, incluyendo:</b>	<b>PaRES 1</b>	<b>PaRES 2</b>	<b>PaRES 3</b>	<b>PaRES 4</b>
El concepto de distribución de la muestra y cómo se aplica para hacer inferencias de parámetros poblacionales, basadas en una muestras (error estándar)		Ítem 10, 12		
El concepto de significación estadística, niveles de significación y p-valor			Ítem 2, 3, 4, 6, 10, 18	
El concepto de intervalo de confianza, incluyendo la interpretación de nivel de confianza y margen de error		Ítem 11, 12, 13, 14		

Los estudiantes deben saber:	PaRES 1	PaRES 2	PaRES3	PaRES 4
Cómo interpretar los resultados estadísticos en un contexto		Ítem 13, 15,	Ítem 10, 13, 18, 23	
Como criticar noticias y artículos de revistas que incluyen información estadística, incluyendo la presentación, identificando lo que falta y los errores en el estudios o en los métodos utilizados para generar la información				

Figura 40. Alineación de los cuestionarios parciales (PaRES) 1 – 4 con GAISE 2010.

### 6.2.3. Pruebas piloto de los cuestionarios parciales.

Los cuatro cuestionarios parciales, fueron administrados a los alumnos de la muestra piloto del curso 2011-2012, al finalizar los distintos bloques de contenidos a lo largo de curso.

Se llevó a cabo el análisis de ítems de estas pruebas, analizándose también la fiabilidad y validez de los cuestionarios.

#### 6.2.3.1. Análisis de ítems de PaRES 1

Se dispuso de una muestra de 123 alumnos que cumplimentaron esta primera prueba parcial. En la Tabla 14 se recogen los datos del análisis de ítems.

Tabla 14. Análisis de ítems de PaRES 1.

ANÁLISIS DE ÍTEMES DE PaRES 1 (PILOTO)						
ÍTEM	DF	DC1	DC2	ANÁLISIS DISTRACTORES	% RESPUESTAS CORRECTAS	CORRELACIÓN ÍTEM-TOTAL
1	<b>0,79</b>	0,35	0,61	Domina el distractor B en el grupo de alumnos de puntuaciones más bajas.	78,04	0,193
2	0,48	0,06	0,53	Domina el distractor B sobre el resto de distractores.	50,40	- 0,015
3	0,52	0,39	0,69	Domina el distractor D en el grupo de alumnos de puntuaciones más bajas.	49,59	0,201
4	<b>0,16</b>	0,13	0,70	Dominan los distractores.	17,88	0,077
5	<b>0,29</b>	0,26	0,72	Domina el distractor C	30,89	0,132

ÍTEM	DF	DC1	DC2	ANÁLISIS DISTRACTORES	% RESPUESTAS CORRECTAS	CORRELACIÓN ÍTEM-TOTAL
6	0,42	0,32	0,69		37,39	0,068
7	0,44	0,61	0,85	Domina el distractor C en el grupo de alumnos de puntuaciones más bajas.	43,90	0,294
8	<b>0,79</b>	0,35	0,61		76,42	0,266
9	0,44	0,61	0,85	Domina el distractor B en el grupo de alumnos de puntuaciones más bajas.	47,96	0,239
10	0,55	0,52	0,74		47,96	0,239
11	<b>0,13</b>	- 0,1	0,25	Domina el distractor B	14,63	- 0,182
12	<b>0,77</b>	0,32	0,60		74,79	0,082
13	0,40	0,42	0,76	Domina el distractor C en el grupo de alumnos de puntuaciones más bajas.	30,89	0,266
14	0,61	0,45	0,68		60,16	0,232
15	0,47	0,61	0,83	Domina el distractor B en el grupo de alumnos de puntuaciones más bajas.	45,52	0,303
16	0,48	0,19	0,60		43,08	0,007
17	<b>0,29</b>	0,39	0,83	Domina el distractor D	26,01	0,219
18	0,58	0,00	0,50		52,84	- 0,159
19	<b>0,27</b>	0,23	0,71	Domina el distractor B	35,52	0,050
20	<b>0,29</b>	0,06	0,56	Domina el distractor D	30,89	- 0,067
21	0,37	0,35	0,74	Domina el distractor D en el grupo de alumnos de puntuaciones más bajas.	39,02	0,151
22	0,42	0,39	0,73	Domina el distractor C en el grupo de alumnos de puntuaciones más bajas.	42,27	0,131
23	0,39	0,39	0,75	Domina el distractor C en el grupo de alumnos de puntuaciones más bajas.	43,90	0,130
24	0,56	0,48	0,71		63,41	0,192
25	0,58	0,45	0,69		61,78	0,260

En la Figura 41 se recoge un resumen de los resultados de este análisis del cuestionario.

**RESULTADOS DEL ANÁLISIS DE PaRES1 (PILOTO)****DESCRIPTIVOS**

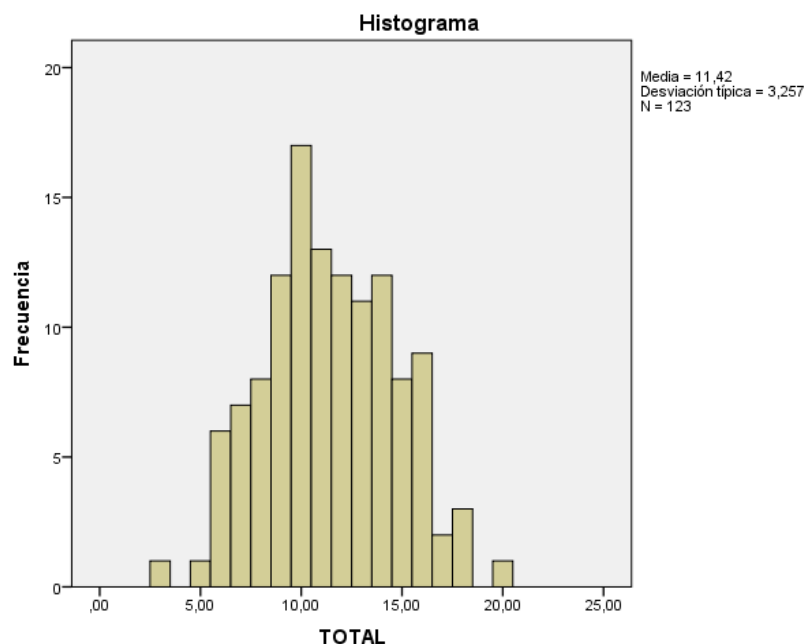
Nº de ítems = 25

N=123

Puntuación media del cuestionario: 11,423

Mediana: 11,00

Desviación típica: 3,256

**ÍNDICE DE DIFICULTAD**

- Los valores del índice de dificultad encontrados en la muestra están comprendidos entre 0,13 y 0,79.
- El intervalo de confianza para la media del índice de dificultad es [0,387605; 0,531595] con un 95% de confianza.
- Un 24% de los ítems muestran un nivel elevado de dificultad y en todos ellos dominan los distractores sobre la respuesta correcta. Estos ítems evalúan los siguientes resultados de aprendizaje:
  - Comprensión del concepto de media aritmética y sus propiedades.
  - Comparar y reconocer las medidas de tendencia central de una distribución de datos que no es simétrica. Entender el concepto de sesgo y asociarlo al valor de la mediana.
  - Asociar la desviación estándar al grado de homogeneidad de los datos y la proximidad de los mismos al valor de la media, reconociendo la mayor o menor variabilidad de dos conjuntos de datos representados en un histograma.
  - Reconocer sobre un histograma con puntuaciones atípicas, la conveniencia de utilizar la mediana y el rango intercuartílico para describir el centro y la dispersión de la distribución, frente a la media y desviación típica que afectadas por valores extremos pueden no representar al grupo.
  - Capacidad de interpretar el resumen de datos recogidos en una tabla de frecuencias
  - Comprensión de la representación de un conjunto de datos en un gráfico de cajas y capacidad para interpretar comparativamente el rango de varios conjuntos de datos.

**ÍNDICE DE DISCRIMINACIÓN**

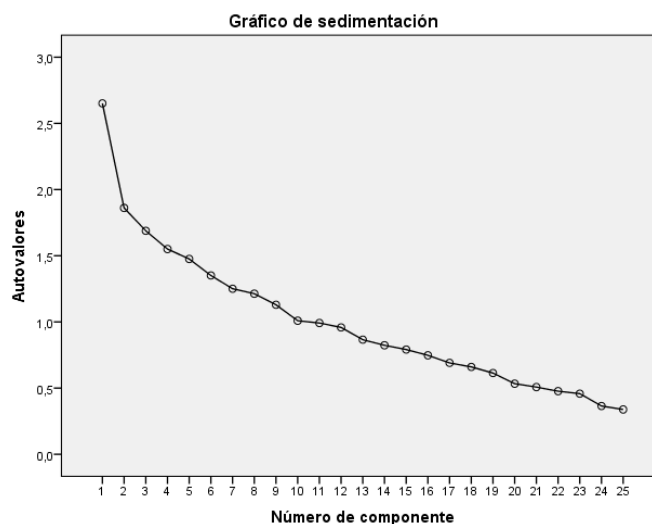
- Los índices de discriminación DC2 de los ítems se encuentran entre 0,5 y 0,85.
- El intervalo para la media de este índice es [0,623611; 0,730789].
- Algunos ítems presenta correlación ítem-total negativas. Los ítems que además de mostrar correlación negativa, tienen los otros índices (DC1 y DC2) indicando poca capacidad de discriminación son el 4, 11 y 20. En los tres casos resultaron difíciles para los alumnos ( $DF < 0,3$ ) pese a que sus resultados de aprendizaje son básicos y la redacción de los enunciados y de las opciones de respuesta son adecuadas para evaluarlos.

**FIABILIDAD**

$\alpha$  de Cronbach = 0,495

**VALIDEZ**

- Existen diferencias significativas entre las medias de los dos grupos extremos para la puntuación total del cuestionario.
- $KMO=0,541$ . Prueba de esfericidad de Bartlett significativa ( $p < 0,01$ )
- El análisis factorial extrajo 10 factores que explican el 60,72% de la varianza total con un primer factor que explica el 10,6% de la varianza, mientras que el segundo explica el 7,4% .
- El primer factor es el único que está compuesto por tres ítems, el resto de factores tiene sólo dos ítems.



Matriz de componentes rotados <sup>a</sup>										
	Componente									
	1	2	3	4	5	6	7	8	9	10
ITEM9	,759									
ITEM7	,590			,376						
ITEM15	,549				,340					
ITEM13		,716								
ITEM8		,647		,310						
ITEM1		,401		,312	,326					
ITEM16			,741							
ITEM3			,605							
ITEM23				,737						
ITEM25				,587						
ITEM6			,371		,578					
ITEM5					,578					
ITEM11					-556					-406
ITEM19						,767				
ITEM17			,345			,590				
ITEM20	-,381					,491				
ITEM21							,623			
ITEM4							,569			
ITEM18			,311				-,495	-,454		
ITEM24			,306				,462			,309
ITEM2								,850		
ITEM10		,439	,305					,462		
ITEM12									,869	
ITEM14									,389	
ITEM22										,818

Método de extracción: Análisis de componentes principales.  
Método de rotación: Normalización Varimax con Kaiser.  
a. La rotación ha convergido en 40 iteraciones.

Figura 41. Resultados del análisis de ítems de PaRES1.

### 6.2.3.2. Análisis de ítems de PaRES 2

La Tabla 15 muestra los datos del análisis de ítems de PaRES 2 y la Figura 42 los resultados del análisis de dichos datos.

Tabla 15. Análisis de ítems de PaRES 2.

ANÁLISIS DE ÍTEM DE PaRES 2 (PILOTO)						
ÍTEM	DF	DC1	DC2	ANÁLISIS DISTRACTORES	% RESPUESTAS CORRECTAS	CORRELACIÓN ÍTEM-TOTAL
1	<b>0,78</b>	0,31	0,60		72,41	0,121
2	<b>0,05</b>	- 0,1	0,0	Domina el distractor A	9,48	- 0,219
3	0,57	0,72	0,82	Domina el distractor A en el grupo de alumnos de puntuaciones más bajas.	48,27	0,308
4	0,62	0,62	0,75		56,03	0,235
5	0,67	0,52	0,69		75,86	0,254
6	0,64	0,72	0,78	Domina el distractor A en el grupo de alumnos de puntuaciones más bajas.	74,13	0,409

ÍTEM	DF	DC1	DC2	ANÁLISIS DISTRACTORES	% RESPUESTAS CORRECTAS	CORRELACIÓN ÍTEM-TOTAL
7	0,45	0,69	0,88	Domina el distractor D en el grupo de alumnos de puntuaciones más bajas.	37,93	0,335
8	<b>0,22</b>	- 0,1	0,31	Dominan los distractores.	24,13	- 0,261
9	<b>0,71</b>	0,59	0,71		70,68	0,318
10	0,38	- 0,1	0,41	Domina el distractor C en el grupo de alumnos de puntuaciones más altas.	46,55	- 0,117
11	<b>0,28</b>	0,14	0,63	Domina el distractor D	32,75	- 0,122
12	0,47	0,79	0,93	Dominan los distractores A y B en el grupo con puntuaciones más bajas.	42,24	0,349
13	0,57	0,31	0,64	Domina el distractor B en el grupo con puntuaciones más bajas.	53,44	0,022
14	0,64	0,66	0,76		68,10	0,338
15	0,38	- 0,0	0,45	Domina el distractor B en el grupo con puntuaciones más altas.	48,27	- 0,214



**RESULTADOS DEL ANÁLISIS DE ÍTEMS DE PaRES 2 (PILOTO)****DESCRIPTIVOS**

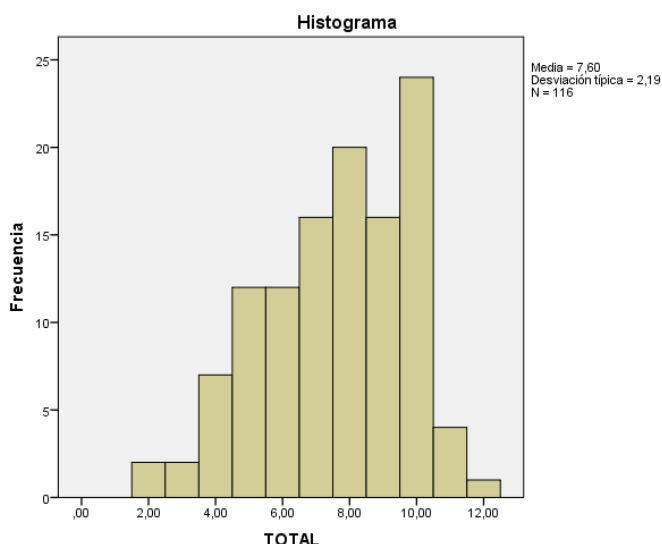
Nº de ítems = 15

N=116

Puntuación media del cuestionario: 7,603

Mediana: 8,00

Desviación típica: 2,190

**ÍNDICE DE DIFICULTAD**

- Los valores del índice de dificultad encontrados en la muestra están comprendidos entre 0,05 y 0,78.
- El intervalo de confianza para la media del índice de dificultad es [0,383008; 0,607658] con un 95% de confianza.
- Un 20% de los ítems muestran un nivel elevado de dificultad con índices por debajo de 0,3. Estos ítems evalúan los siguientes resultados de aprendizaje:
  - Comprensión del concepto de puntuación típica que utiliza la desviación típica como unidad de medida y permite comparar resultados de procedencia distinta
  - Entender que un intervalo de confianza para estimar la media poblacional a partir de una muestra, no asegura al 100% que la media poblacional se encuentre en dicho intervalo.
- El ítem 2 con tan sólo un 9,48% de respuestas correctas, es el que presenta mayor dificultad. Este ítem ha sido revisado por tener una correlación negativa con el total del cuestionario. Su formulación y su clave de corrección son correctas y el resultado de aprendizaje que evalúa es:
  - Conocer que en muestras grandes, la amplitud o recorrido de una variable comprende unas seis desviaciones típicas.
- En dos de los ítems (10 y 15) el distractor domina sólo para el grupo con puntuaciones más altas por lo que habría que pensar que existe un error de comprensión con respecto a estos resultados de aprendizaje.
  - Comprensión del concepto de error típico de la media y entender que los intervalos de confianza para la media calculados en muestras grandes son más precisos que en muestras pequeñas.
  - Entender las limitaciones de la estimación por intervalos a partir de muestras no aleatorias.

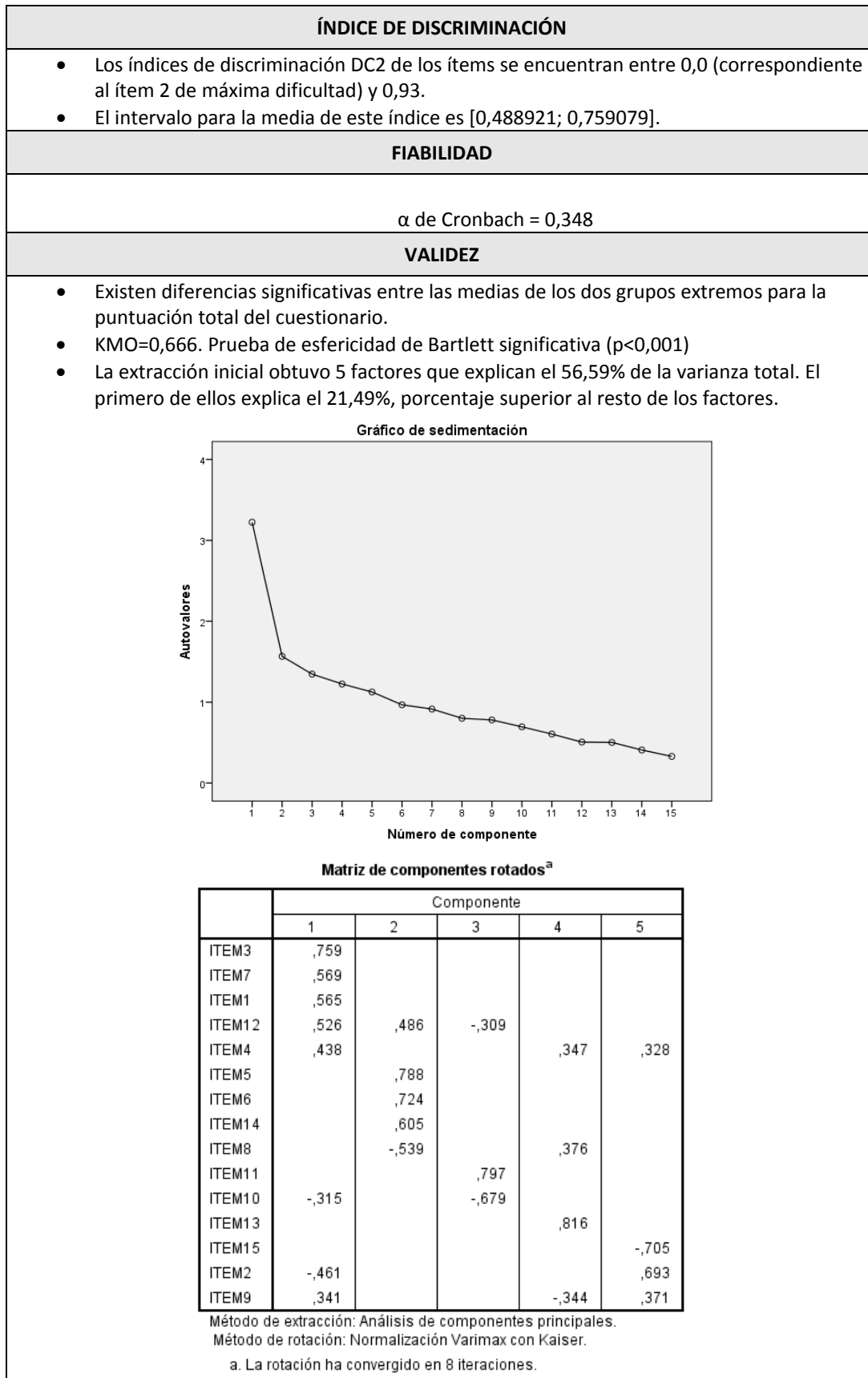


Figura 42. Resultados del análisis de ítems de PaRES2.

### 6.2.3.3. Análisis de ítems de PaRES 3

En la Tablas 16 y en la Figura 43 encontramos los datos correspondientes al análisis de ítems del cuestionario PaRES 3 compuesto por 23 preguntas de respuesta múltiple.

Tabla 16. Análisis de ítems de PaRES 3.

ANÁLISIS DE ÍTEMS DE PaRES3 (PILOTO)						
ÍTEM	DF	DC1	DC2	ANÁLISIS DISTRACTORES	% RESPUESTAS CORRECTAS	CORRELACIÓN ÍTEM-TOTAL
1	0,43	- 0,3	0,28	Domina el distractor B	38,26	- 0,348
2	<b>0,78</b>	0,45	0,64		75,65	0,185
3	0,66	0,69	0,76	Domina el distractor A en el grupo de alumnos de notas más bajas.	72,17	0,449
4	0,66	0,69	0,76	Domina el distractor B en el grupo de alumnos de notas más bajas.	73,04	0,469
5	0,60	0,66	0,77	Dominan los distractores el grupo de alumnos de notas más bajas.	65,21	0,431
6	0,69	0,48	0,68		75,65	0,190
7	<b>0,17</b>	0,07	0,60	Dominan los distractores	14,78	0,001
8	0,62	0,55	0,72	Domina el distractor A en el grupo de alumnos de puntuaciones más bajas.	67,82	0,379
9	<b>0,24</b>	- 0,0	0,43	Dominan los distractores	35,65	- 0,177
10	0,60	0,66	0,77	Domina el distractor D en el grupo de alumnos de puntuaciones más bajas.	62,21	0,417
11	0,48	- 0,2	0,36	Domina el distractor A en el grupo de alumnos de puntuaciones más altas.	53,04	- 0,278
12	0,59	0,55	0,74	Domina el distractor C en el grupo de alumnos de puntuaciones más bajas.	58,26	0,216
13	0,47	0,72	0,89	Domina el distractor B en el grupo de alumnos de puntuaciones más bajas.	42,60	0,440
14	0,59	0,69	0,79	Domina el distractor B en el grupo de alumnos de puntuaciones más bajas.	62,60	0,411
15	<b>0,78</b>	0,31	0,60		75,65	0,083
16	<b>0,90</b>	0,14	0,54		84,34	0,036

ÍTEM	DF	DC1	DC2	ANÁLISIS DISTRACTORES	% RESPUESTAS CORRECTAS	CORRELACIÓN ÍTEM-TOTAL
17	<b>0,28</b>	- 0,0	0,44	Dominan los distractores.	30,43	- 0,169
18	<b>0,09</b>	0,03	0,60	Domina el distractor D.	17,39	- 0,057
19	0,66	0,62	0,74		56,52	0,217
20	0,55	0,48	0,72		46,08	0,190
21	0,41	0,34	0,71	Dominan los distractores en el grupo de alumnos de puntuaciones más bajas.	37,39	0,234
22	0,36	0,17	0,62	Dominan los distractores.	32,17	0,002
23	0,52	0,41	0,70		46,08	0,130

### RESULTADOS DEL ANÁLISIS DE ÍTEMS DE PaRES 3 (PILOTO)

#### DESCRIPTIVOS

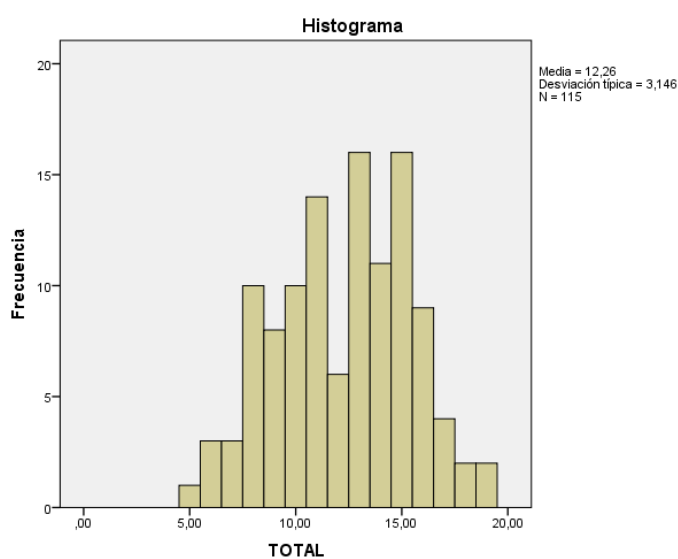
Nº ítems = 23

N= 115

Puntuación media del cuestionario: 12,26

Mediana: 13,00

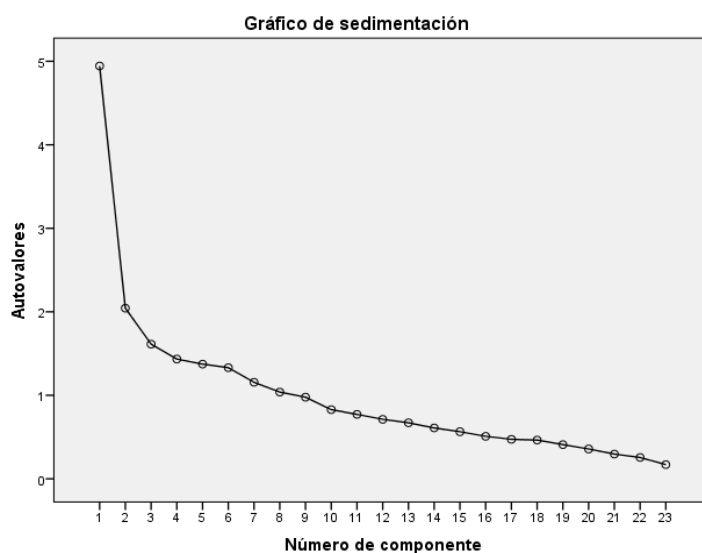
Desviación típica: 3,145



<b>ÍNDICE DE DIFICULTAD</b>
<ul style="list-style-type: none"> <li>• Los valores del índice de dificultad encontrados en la muestra están comprendidos entre 0,09 y 0,90.</li> <li>• El intervalo de confianza para la media del índice de dificultad es [0,440276; 0,614507] con un 95% de confianza.</li> <li>• Hay cuatro ítems con un elevado nivel de dificultad por debajo de 0,3. En todos ellos dominan los distractores sobre la respuesta correcta. Los resultados de aprendizaje que evalúan estos ítems son: <ul style="list-style-type: none"> <li>▪ Reconocer que una diferencia puede ser significativa y sin embargo puede que no sea relevante en la práctica.</li> <li>▪ Entender que si no puede ser probada la hipótesis alternativa, no significa que no sea cierta, sobre todo en muestras pequeñas.</li> <li>▪ Detectar errores en una tabla de contingencia que recoge los datos para un análisis de asociación entre variables cualitativas, entendiendo que las observaciones tienen que ser independientes.</li> <li>▪ Interpretar para un ejemplo dado un resultado significativo en una prueba chi-cuadrado de independencia de caracteres.</li> </ul> </li> <li>• El ítem 1, que evalúa el correcto planteamiento de la hipótesis nula y alternativa, tiene una correlación negativa considerable con el resto del cuestionario. En este ítem domina un distractor sobre la respuesta correcta, lo que revela un error de comprensión en la forma correcta de plantear la hipótesis nula y alternativa de un estudio.</li> <li>• El ítem 9 que presenta una correlación negativa con el resto del cuestionario y con índices que muestran no tener capacidad de discriminación, resultó una cuestión difícil, dominando los distractores sobre la respuesta correcta. La pregunta valora el resultado de aprendizaje: “Entender que si no podemos probar la hipótesis alternativa no significa que no sea cierta”, sobre todo en muestras pequeñas y tiene una redacción del enunciado y de las opciones de respuesta claro y sencillo, no encontrando explicación para los datos resultantes de su análisis.</li> <li>• Muestra también correlación negativa el ítem 11, pese a ser contestado correctamente por el 53% de los estudiantes y presentar de forma clara todas las posibles conclusiones de no rechazar la hipótesis nula en una comparación de muestras. Domina un distractor en el grupo con puntuaciones más altas.</li> <li>• Los ítems 17 y 18 con correlaciones negativas, resultados difíciles para los alumnos y sin capacidad de discriminar. Evalúan resultados de aprendizaje relacionados con la interpretación correcta de los datos recogidos en una tabla de contingencia.</li> </ul>
<b>ÍNDICE DE DISCRIMINACIÓN</b>
<ul style="list-style-type: none"> <li>• Los índices de discriminación DC2 de los ítems se encuentran entre 0,28 y 0,89.</li> <li>• El intervalo de confianza para la media de este índice es [0,581094; 0,711079].</li> </ul>
<b>FIABILIDAD</b>
$\alpha$ de Cronbach = 0,527

**VALIDEZ**

- Existen diferencias significativas entre las medias de los dos grupos extremos para la puntuación total del cuestionario.
- KMO=0,724. Prueba de esfericidad de Bartlett significativa ( $p < 0,001$ )
- El análisis factorial en una primera extracción muestra 8 factores que explican el 64,92% de la varianza total, correspondiéndole al primer factor el 21,49% de dicha varianza.



**Matriz de componentes rotados<sup>a</sup>**

	Componente							
	1	2	3	4	5	6	7	8
ITEM4	,897							
ITEM3	,882							
ITEM5	,614	,424						
ITEM10	,548			,429	,313			
ITEM12	,478	,350						
ITEM1	-,421							,351
ITEM13	,390	,342		,353	,311			
ITEM9		-,790						
ITEM20		,570						
ITEM8		,559	,374			,362		
ITEM11	-,329	-,363						
ITEM21			,837					
ITEM22			,652	-,440				
ITEM17		-,309	,516			-,364		
ITEM23				,698				
ITEM2				,666				
ITEM16					,830			
ITEM14	,336				,616	,430		
ITEM18						,751		
ITEM7							,806	
ITEM19		,460		,303			,467	
ITEM15		,304						,727
ITEM6		,386						-,605

Método de extracción: Análisis de componentes principales.

Método de rotación: Normalización Varimax con Kaiser.

a. La rotación ha convergido en 11 iteraciones.

Figura 43. Resultados del análisis de ítems de PaRES3.

#### 6.2.3.4. Análisis de ítems de PaRES 4

Este apartado recoge los datos relativos a PaRES 4. Este cuestionario compuesto de 15 ítems fue la última prueba parcial realizada por los alumnos durante el curso. Los resultados quedan recogidos en las Tabla 17 y Figura 44.

Tabla 17. Análisis de ítems de PaRES 4

ANÁLISIS DE ÍTEMS DE PaRES 4 (PILOTO)						
ÍTEM	DF	DC1	DC2	ANÁLISIS DISTRACTORES	% RESPUESTAS CORRECTAS	CORRELACIÓN ÍTEM-TOTAL
1	0,62	0,70	0,78	Domina el distractor B en el grupo con puntuaciones más bajas.	64,40	0,327
2	0,33	0,33	0,75	Domina el distractor A en el grupo con puntuaciones más bajas	33,05	0,109
3	0,53	0,47	0,72	Domina el distractor C en el grupo con puntuaciones más bajas	47,45	0,156
4	<b>0,78</b>	0,30	0,60		67,79	- 0,010
5	0,62	0,70	0,78	Domina el distractor D en el grupo con puntuaciones más bajas	72,03	0,457
6	0,30	- 0,2	0,33	Domina el distractor E en el grupo con puntuaciones más altas.	35,59	- 0,261
7	0,60	0,67	0,78		61,86	0,350
8	<b>0,70</b>	0,60	0,71		73,72	0,374
9	<b>0,90</b>	0,20	0,56	No se eligen los distractores C y D.	92,37	0,273
10	<b>0,82</b>	0,23	0,57		85,59	0,109
11	<b>0,12</b>	- 0,0	0,43	Domina el distractor C	12,71	- 0,052
12	0,32	0,10	0,58	Domina el distractor B	33,89	- 0,092
13	0,52	0,63	0,81	Domina el distractor C en el grupo con puntuaciones más bajas.	61,01	0,378
14	0,53	0,67	0,81	Domina el distractor B en el grupo con puntuaciones más bajas.	50	0,318
15	0,63	0,67	0,76	Domina el distractor A en el grupo con puntuaciones más bajas.	72,88	0,325

**RESULTADOS DEL ANÁLISIS DE ÍTEMS DE PaRES 4 (PILOTO)****DESCRIPTIVOS**

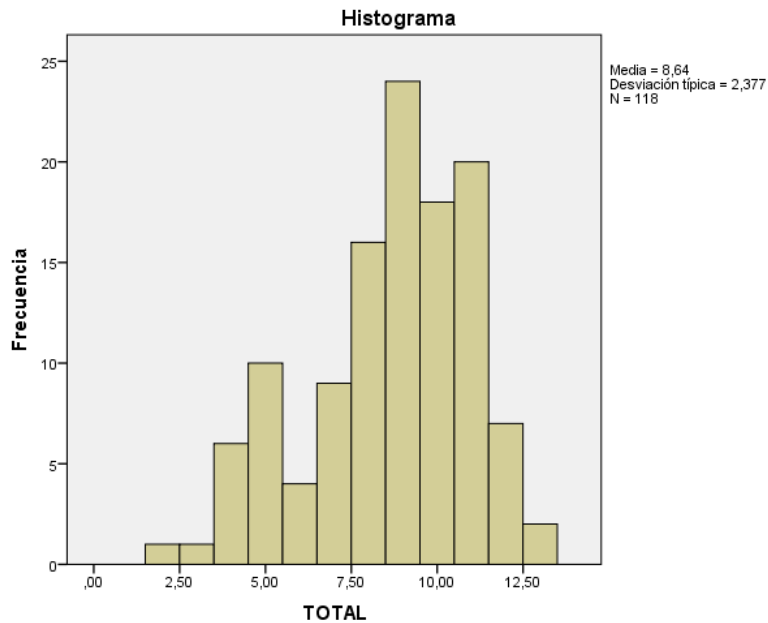
Nº ítems = 15

N= 118

Puntuación media del cuestionario: 8,644

Mediana: 9,00

Desviación típica: 2,377

**ÍNDICE DE DIFICULTAD**

- Los valores del índice de dificultad encontrados en la muestra están comprendidos entre 0,12 y 0,90.
- El intervalo de confianza para la media del índice de dificultad es [0,436302; 0,673031] con un 95% de confianza.
- Sólo el ítem 11 muestra un nivel de dificultad alto. El resultado de aprendizaje que evalúa este ítem es: "Considerar que la eliminación de un solo sujeto, especialmente si tiene puntuaciones extremas, puede modificar el coeficiente de correlación". Este ítem tiene una correlación ítem-total negativa
- En general, los ítems del cuestionario muestran una dificultad media-baja.
- En el ítem 6, con una correlación negativa con el resto del cuestionario, domina un distractor en el grupo con puntuaciones más altas. Una vez revisada la formulación del distractor se observa que induce a error, por lo que dicho distractor será eliminado del cuestionario.

**ÍNDICE DE DISCRIMINACIÓN**

- Los índices de discriminación DC2 de los ítems se encuentran entre 0,33 y 0,81.
- El intervalo de confianza para la media de este índice es [0,583335; 0,745998].



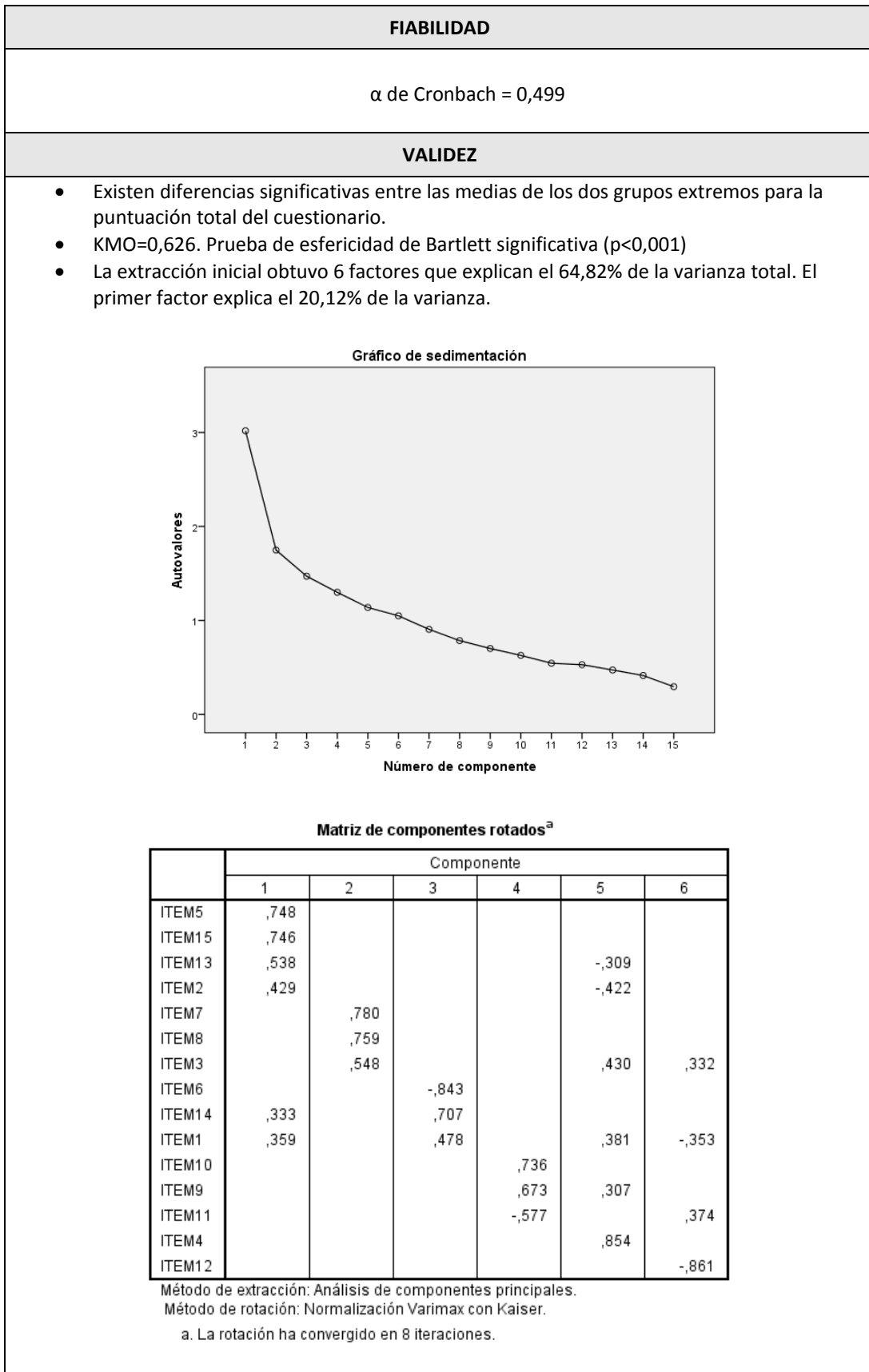


Figura 44. Resultados del análisis de ítems de PaRES4.

El análisis de ítems expuesto anteriormente se realizó para estudiar en profundidad estos cuestionarios y revisar, a la vista de los resultados, la redacción de distractores y de los enunciados de las preguntas.

Estas pruebas fueron destinadas a entrenar en los estudiantes su capacidad de razonar al tiempo que les ayuda a fijar conceptos cuando fueron corregidas y argumentadas por la profesora todas las posibles respuestas correctas e incorrectas. Por este motivo todos los ítems, de mucha o poca dificultad tienen su función en el cuestionario, en algunos casos por evaluar conceptos básicos y en otros por presentar una dificultad que supondrá para el estudiante un esfuerzo para pensar y razonar la respuesta.

La construcción de estas pruebas fue exigente en la elección de los enunciados adaptados a las ciencias de la salud, rigurosa en reflejar los resultados de aprendizaje marcados, así como en utilizar para la redacción de distractores los errores más frecuentes frente a la respuesta correcta.

Estas pruebas no han constituido un instrumento de evaluación para discriminar, ordenar y calificar a los alumnos. No era un objetivo obtener unos instrumentos de dificultad media y máxima discriminación, sí nos preocupaba que los ítems fueran correctos y adecuados para ser útiles en una evaluación formativa, forzara a los alumnos a razonar y sirvieran de ejemplo al profesor para corregir errores frecuentes en la comprensión de los conceptos estadísticos.

### **6.3 Recogida de Datos**

Al tiempo que se implementa una metodología docente para promover en el alumno la capacidad de razonar estadísticamente, existe la necesidad de evaluar este razonamiento. Esta evaluación además de ser una medida de los resultados y las calificaciones de los alumnos, analiza también si las metas del curso se están cumpliendo y si los cambios metodológicos están produciendo el resultado deseado. Un objetivo importante de este trabajo ha sido medir el progreso en el razonamiento y en la comprensión de la estadística a lo largo del curso, considerando que estos cambios son un objeto de investigación relevante en la investigación educativa.

Los instrumentos para medir las variables de investigación se aplicaron en distintos momentos del curso:

- Recogida de datos al comienzo de curso: Los alumnos de ambos grupos realizaron el primer día de clase el cuestionario de razonamiento estadístico (CERES1), junto con la escala de percepción de la competencia estadística (EpCE1) y un pequeño cuestionario que recoge información sobre algunas variables como la edad, sexo, nota de acceso a los estudios universitarios,..etc
- Recogida de datos al finalizar la asignatura: Los alumnos de ambos grupos realizaron, de nuevo, al finalizar la asignatura, el cuestionario de razonamiento estadístico (CERES2). También volvieron a cumplimentar la escala de percepción de la competencia estadística (EpCE2). Posteriormente los alumnos realizaron las pruebas de evaluación de la asignatura consistentes en un examen práctico junto con la prueba de respuestas abiertas parcialmente restringidas diseñada para este estudio.

	Al inicio de la asignatura Septiembre 2012 <b>Pretest</b>	Durante el curso	Al finalizar la asignatura Mayo 2012 <b>Postest</b>
Cuestionario de razonamiento estadístico. (CERES)	Grupo A y B		Grupo A y B
4 Pruebas parciales de razonamiento estadístico. (PaRES)		Grupo A PaRES1 PaRES2 PaRES3 PaRES4	
Escala de percepción de la competencia estadística EpCE	Grupo A y B		Grupo A y B
Prueba de respuesta abierta parcialmente restringida.			Grupo A y B
Examen práctico			Grupo A y B

Figura 45. Cronograma del proceso de recogida de datos.

El cuestionario de evaluación de razonamiento estadístico (CERES), fue utilizado con todos los alumnos de la **muestra de estudio (N=109)** en dos ocasiones, como pretest (CERES1) antes de comenzar la asignatura y como postest (CERES2) al finalizar la misma.

La aplicación del cuestionario en el pretest (CERES1), se realizó como una evaluación diagnóstica inicial para ver cuántos de los aspectos del razonamiento

estadístico poseen ya los alumnos para posteriormente ser contrastados con los resultados de aprendizaje adquiridos en la asignatura. Algunos estudios han examinado los cambios en las habilidades de razonamiento de los estudiantes durante su formación estadística y han puesto de manifiesto mejoras en el nivel de estas habilidades al final de un curso de estadística, (Delmas, Garfield, Ooms y Chance 2007; Derry, Levin, Osana, Jones y Peterson, 2000; Ragasa, 2008; Olani, A., Hoekstra, R., Harskamp, E., y VanDerWerf, G., 2011).

Se realizó un análisis de los ítems con un porcentaje alto de respuestas correctas en el cuestionario realizado como pretest (CERES1), como información del conocimiento razonado que ya poseían los alumnos.

Para analizar las mejoras en el razonamiento que muestran los alumnos al finalizar la asignatura, así como también los errores que permanecen, se realiza un análisis comparativo entre el pretest (CERES1) y el postest (CERES2) de los porcentajes de respuestas correctas para cada uno de los ítems. Dicho análisis se realizó tanto en el grupo de intervención como en el grupo control y se centró en varios aspectos:

- Ítems en los que aumentaba el porcentaje de respuestas correctas en el postest.
- Ítems que permanecían con el mismo porcentaje de respuestas correctas en el pretest y en el postest.
- Ítems que tenían menos respuestas correctas en el postest en comparación con el pretest.

En primer lugar se realizó un análisis comparativo de las puntuaciones medias del cuestionario pretest entre los grupos de intervención y el grupo control. De este modo se pone también de manifiesto el hecho de que los grupos inicialmente son similares en este aspecto antes de la intervención.

Las diferencias en el razonamiento estadístico entre los dos grupos de alumnos, diferenciados por la intervención, se estudiaron comparando las medias de las puntuaciones individuales del cuestionario postest (CERES2) entre los dos grupos. Se indagó sobre si esas diferencias se manifestaban también en la calificación de los alumnos en el examen práctico que conforma la prueba de evaluación de la convocatoria final de esta asignatura. Este mismo objetivo, que pretende constatar diferencias entre el grupo control y el grupo de intervención, se aplicó también a las puntuaciones en la prueba de razonamiento con respuestas abierta parcialmente restringida.

Con las distintas variables se elaboró una base de datos en Excel, en la que fueron eliminados los nombres de los alumnos, manteniendo la clave de estudiante, para constatar la paridad de los datos de cada sujeto.



## CAPÍTULO 7. RESULTADOS DE LA INTERVENCIÓN DE EVALUACIÓN

En este capítulo se pretende, mediante análisis estadísticos, contestar a algunas de las preguntas que motivaron este estudio. Se analizan los cambios producidos en el razonamiento estadístico a lo largo del proceso de enseñanza-aprendizaje, se identifican los logros así como los errores que se mantienen después de cursar la asignatura. Mediante comparaciones entre las muestras, se pondrá de manifiesto si la realización de una evaluación formativa con pruebas diseñadas para que el alumno desarrolle su capacidad de pensar usando los conceptos estadísticos, está asociado con una mejor comprensión conceptual y un mayor rendimiento académico.

Se indaga también sobre los posibles cambios en la percepción que los alumnos tienen de su competencia estadística a lo largo del proceso formativo.

### 7.1. Análisis descriptivos

En este apartado se describe cómo se distribuyen las variables de este estudio, sus dimensiones y frecuencias a fin de especificar sus características para pasar posteriormente a considerar y exponer los resultados acerca de las hipótesis.

El estudio descriptivo se realizó con los datos resultantes de los cuestionarios que cumplimentaron los alumnos de los dos grupos, grupo control y grupo de intervención (N=109).

#### 7.1.1. Análisis descriptivo del Cuestionario de evaluación del razonamiento estadístico CERES.

En primer lugar, se presenta el estudio acerca del cuestionario de evaluación de razonamiento estadístico (CERES) (Figura 46). Se comparan, descriptivamente, las frecuencias del número de respuestas correctas a este cuestionario al comienzo del curso (CERES1) y al finalizar la asignatura (CERES2).

Al comienzo del curso, la media de respuestas correctas del grupo fue de 15,8 con una desviación típica de 3,25, mientras que al finalizar la asignatura el número medio de repuestas correctas aumentó a 21,4 con una desviación típica de 4.

La mitad de los estudiantes contestó correctamente al 50% de las preguntas aunque el número máximo de preguntas contestadas correctamente fue 29 en la prueba del postest, que es un valor no muy alto sobre el total de 40 ítems de los que está constituido CERES. Más adelante se analizará cuáles son los ítems en los que permanecen errores de razonamiento después de cursar la asignatura de estadística.

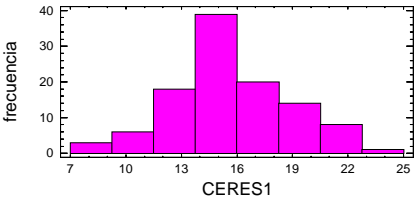
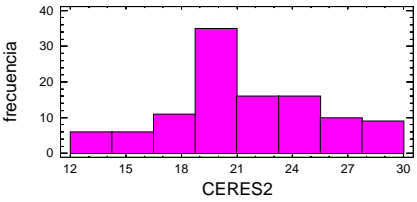
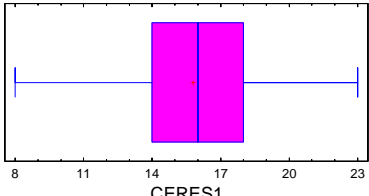
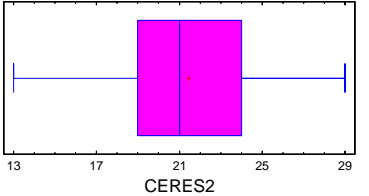
CERES 1 (pre) N° alumnos=109 Número de respuestas correctas	CERES 2 (post) N° alumnos=109 Número de respuestas correctas
	
	
<p>Media = 15,789            Mediana = 16,0            Varianza = 10,594            Desviación típica = 3,25484            Mínimo = 8,0            Máximo = 23,0            Primer cuartil = 14,0            Tercer cuartil = 18,0            Rango intercuar. = 4,0            Asimetría tipi. = 0,315149            Curtosis típificada = -0,795344</p>	<p>Media = 21,4404            Mediana = 21,0            Varianza = 16,1747            Desviación típica = 4,02177            Mínimo = 13,0            Máximo = 29,0            Primer cuartil = 19,0            Tercer cuartil = 24,0            Rango intercuar. = 5,0            Asimetría tipi. = -0,00438166            Curtosis típificada = -1,16732</p>
<p>95,0% intervalo de confianza para la media:            15,789 +/- 0,617957 [15,171;16,4069]</p> <p>95,0% intervalo de confianza para la desviación típica: [2,8726;3,75541]</p>	<p>95,0% intervalo de confianza para la media:            21,4404 +/- 0,763566 [20,6768;22,2039]</p> <p>95,0% intervalo de confianza para la desviación típica: [3,54947;4,64029]</p>

Figura 46. Descripción del número de respuestas correctas al cuestionario de razonamiento estadístico (CERES) en el pretest y en el postest.



### 7.1.2. Prueba de razonamiento estadístico de respuesta abierta

La prueba de razonamiento estadístico de respuesta abierta parcialmente restringida (Prueba RA) fue administrada a todos los alumnos, en una sola ocasión, al finalizar la asignatura como parte de sus exámenes de evaluación y junto con el cuestionario CERES2 (postest).

Esta prueba de respuesta abierta resultó tener cierta dificultad para los estudiantes con una media de 2,9 sobre los 6 puntos máximos sobre los que se calificaba, y una desviación típica de 1,2.

La frecuencia de alumnos con puntuaciones muy bajas fue considerable, y muy pocos estudiantes consiguieron una buena puntuación en esta prueba.

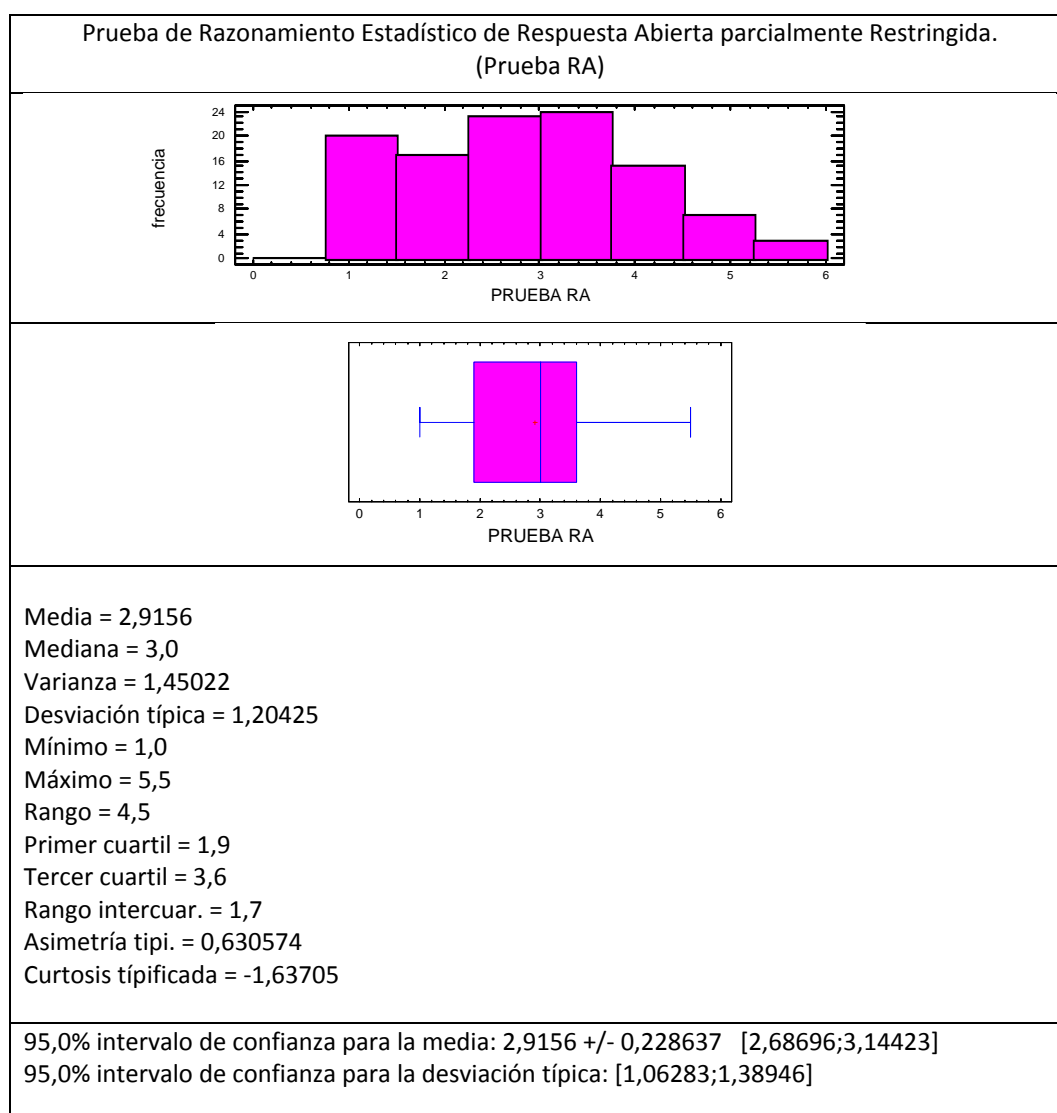


Figura 47. Descriptivo de las puntuaciones de la prueba de razonamiento estadístico de respuesta abierta (Prueba RA).

La prueba de respuesta abierta fue diseñada con el propósito de probar su correlación con el Cuestionario de Razonamiento Estadístico (CERES) y reforzar la validez de contenido del mismo. Utilizando los datos de la muestra de estudio se volvió a comprobar que ambas pruebas están correlacionadas ( $r=0,37$  ;  $p<0,01$ ) lo que nos informa que tienen coincidencia en sus objetos de medida.

### **7.1.3. Examen práctico de resolución de problemas**

La descripción de los resultados del examen de resolución de problemas, que los estudiantes realizan utilizando un programa informático, se exponen en la Figura 48.

Esta prueba se valora sobre una calificación máxima de 10 y en este caso la nota media del grupo fue de 5,8 con una desviación típica de 2,65.

Respecto a las frecuencias más llamativas comentamos que un grupo grande de alumnos contestó correctamente sólo a una de las preguntas, por lo que la calificación fue cercana a 2. Esta pregunta se refería a aspectos descriptivos básicos de una variable. Hay que tener en cuenta que esta prueba de evaluación al igual que el cuestionario de razonamiento estadístico se realizaron en el periodo de exámenes correspondientes a la convocatoria ordinaria de mayo-junio. Algunos alumnos ante la proximidad de la siguiente convocatoria, nos referimos a la convocatoria extraordinaria que se realiza aproximadamente un mes después, organizan su estudio priorizando la preparación de los exámenes para una u otra convocatoria. Este hecho puede ser la explicación de la frecuencia tan alta de alumnos con esas calificaciones tan bajas, demostrando no haber profundizado en el estudio de la asignatura aunque posteriormente mejoraron sus resultados y superaron la asignatura en la siguiente convocatoria.

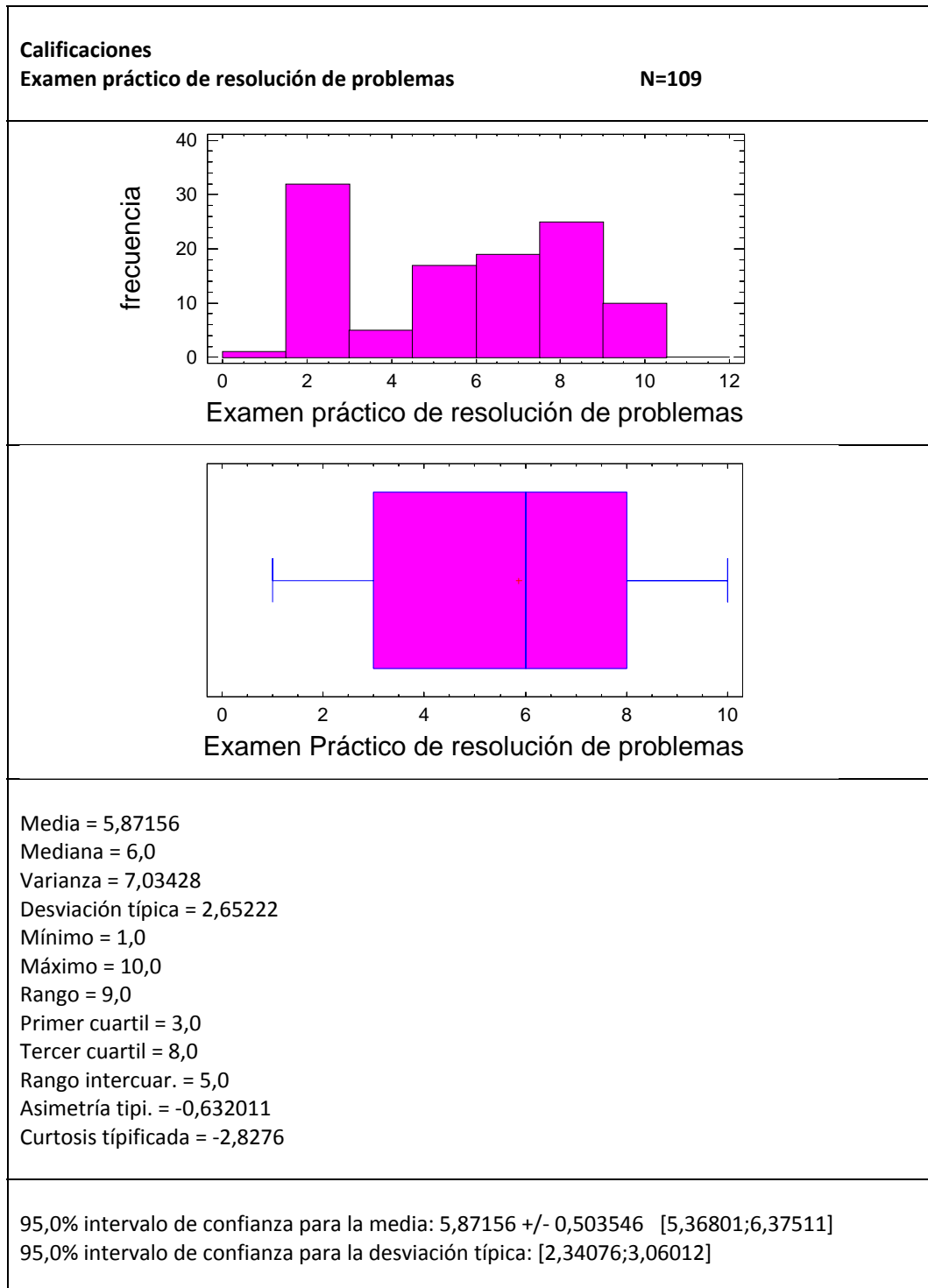


Figura 48. Descriptivo de las calificaciones del examen práctico de resolución de problemas.

El análisis de correlación entre las puntuaciones del cuestionario de evaluación del razonamiento estadístico (CERES) y las calificaciones del examen práctico, nos aporta más información acerca de la validez del instrumento. Se obtuvo, en este caso con la muestra de estudio, una  $r=0,263$  ( $p<0,01$ ).

#### 7.1.4. Escala de percepción de la competencia estadística

Esta escala, como ya se ha expuesto, se corresponde íntegramente con uno de los factores de la escala de Actitudes hacia la Estadística de Muñoz (2002), quién realizó su validación en un grupo de alumnos de ciencias sociales.

En nuestro estudio esta escala fue cumplimentada por los alumnos en dos ocasiones al comienzo de la asignatura y al finalizar la misma. Se recogen en la Figura 49 los datos descriptivos de los resultados de su aplicación en la muestra de estudio.

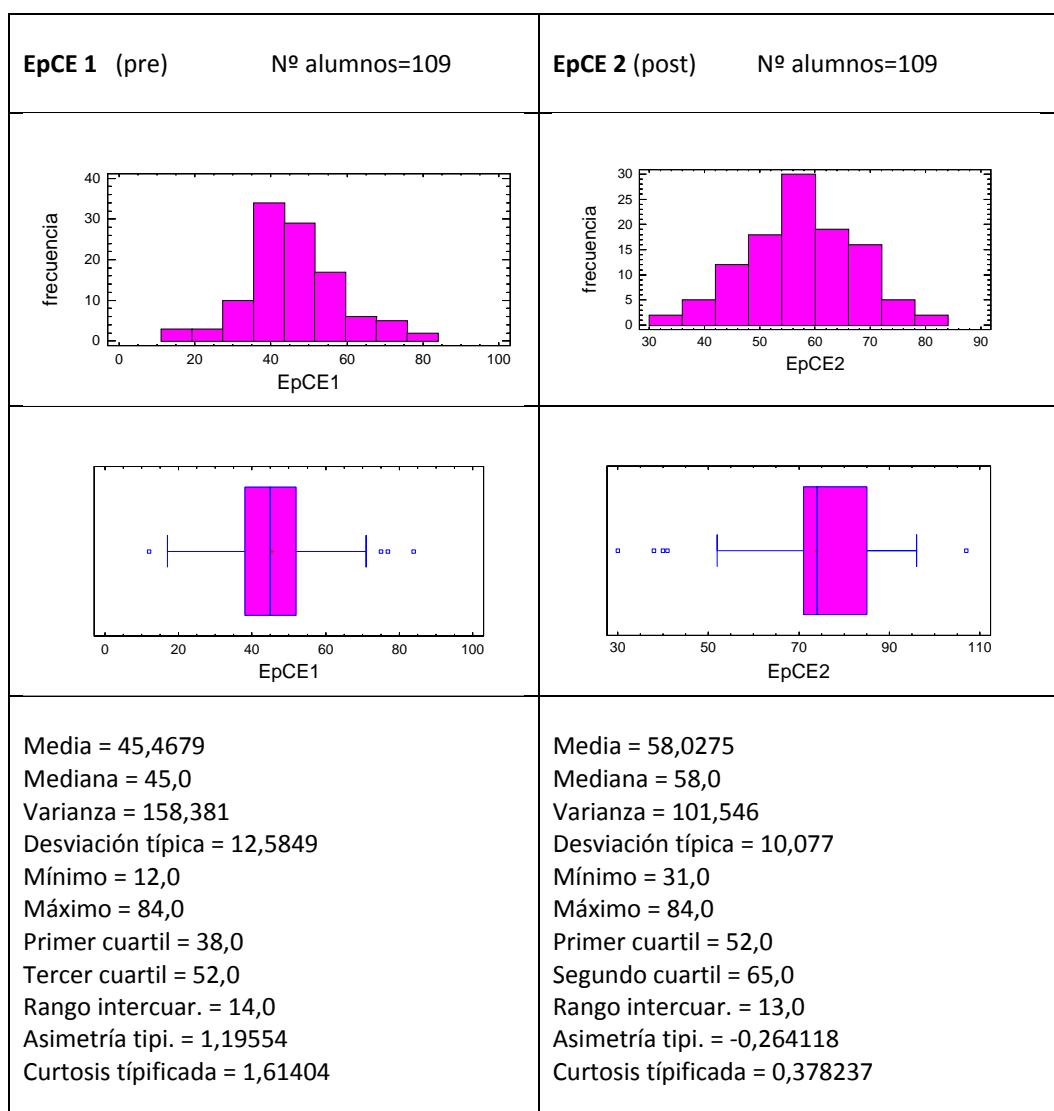


Figura 49. Puntuaciones de la Escala de percepción de la competencia estadística.

El estudio de fiabilidad se realizó con las respuestas recogidas al administrar la escala al final del curso (N =109) y alcanzó un coeficiente  $\alpha = 0,913$ . La fiabilidad encontrada por Muñoz (2002) para este mismo factor de su escala, sobre los datos recogidos en el mes de abril, tuvo un valor similar ( $\alpha = 0,92$  ; N = 269).

En cuanto a las medias de las puntuaciones de la escala de percepción de la competencia estadística, observamos que al comienzo de curso la media es de 45,47 con una desviación típica de 12,58 y al finalizar la asignatura había aumentado a 58,03 con una desviación típica de 10,08.

A modo descriptivo, aunque más adelante se comprobará si existen diferencias significativas, la percepción de la competencia estadística aumenta en los alumnos cuando finalizan su curso de estadística.

## 7.2. Contrastes de hipótesis

A continuación se exponen los resultados de los análisis realizados para comprobar cada una de las hipótesis de este estudio.

### 7.2.1. Hipótesis 1.

***Los alumnos a los que se les aplica una evaluación formativa con pruebas de evaluación del razonamiento estadístico, muestran una mayor comprensión conceptual de la estadística que el grupo de alumnos que no realiza ese tipo de evaluación.***

Se utiliza un análisis de varianza de medidas repetidas para indagar sobre las diferencias en las puntuaciones de CERES1 y CERES2 entre el grupo control y el grupo de intervención.

La media de las puntuaciones de CERES1 en el grupo asignado para la intervención es de 3,955 con una desviación típica de 0,828 y en el grupo que se utilizará como control la media es de 3,878 y la desviación típica 0,820. Recordamos que el cuestionario de razonamiento estadístico CERES1 se pasó a los alumnos al comienzo del curso antes de comenzar la asignatura y de realizarse la intervención. La igualdad de estos parámetros en las dos muestras aportan información respecto a la igualdad entre estos grupos, en el razonamiento estadístico que muestran inicialmente ambos y la idoneidad de utilizarlos como grupos de comparación.

Al finalizar el curso las puntuaciones de este mismo cuestionario, que denominamos CERES2 tiene una media de 5,716 y una desviación típica de 0,880 en el grupo de intervención, mientras que en el grupo control la media es de 5,130 y la desviación 1,038.

El cuestionario de razonamiento estadístico en su prueba inicial (CERES1) con toda la muestra tubo una puntuación media de 3,9165 con una DT = 0,821, mientras que la prueba con este mismo cuestionario al finalizar el curso (CERES2) obtuvo una puntuación media de 5,426 y una DT= 1,00.

**Tabla 18. Prueba de Levene de igualdad de varianzas y estadísticos descriptivos de la comparación de CERES en el pretest y en el postest.**

CERES / GRUPO	Media	Desviación típica	N	Contraste de Levene sobre la igualdad de las varianzas
CERES1				
Grupo intervención	3,9545	,82817	55	F = 0,015; p=0,902
Grupo Control	3,8778	,82088	54	
Total	3,9165	,82164	109	
CERES2				
Grupo Intervención	5,7164	,88082	55	F=0,450 ; p= 0,504
Grupo Control	5,1296	1,03880	54	
Total	5,4257	1,00216	109	

Se valida el supuesto de igualdad de las varianzas para los dos grupos, tanto para CERES1 (F = 0,015;  $p > 0,1$ ) como para CERES2 (F=0,450 ;  $p > 0,1$ ).

La prueba de Box acepta también el supuesto de igualdad de varianzas (M de Box= 1,503; F=0,491;  $p=0,689$ ), en este caso de las matrices de covarianza generadas en cada grupo (intervención y control).

los resultados de los análisis multivariados para indagar sobre las diferencias en las puntuaciones de CERES1 y CERES2 entre el grupo control y el grupo de intervención (Tabla 19) nos llevan a rechazar la hipótesis de igualdad de medias y concluir que la puntuación media de CERES 2 es significativamente mayor que la puntuación en la prueba inicial CERES1, ( F (1, 107)= 193,753 ;  $p < 0,001$ ; Lambda de Wilks =0,356,  $\eta^2= 0.644$ . ).

**Tabla 19. Análisis de varianza de medidas repetidas. Diferencias en las puntuaciones de CERES1 y CERES2 entre el grupo control y el grupo de intervención.**

Lambda de Wilks	Valor	F	GL de las hipótesis	Gl del error	Sig.	Eta al cuadrado o parcial
CERES1/CERES2	0,356	193,753	1			
CERES GRUPO A- GRUPOB	0,951	5,548	1			

Se encuentran también diferencias estadísticamente significativa en las puntuaciones del cuestionario de razonamiento estadístico entre el grupo de intervención y el grupo control,  $F(1,107)=5.548$  ;  $p < 0.05$ ; Lambda de Wilks =0.913, eta cuadrado parcial = 0.049.

Examinando la significación de los efectos intrasujetos (Tabla 20), observamos que resulta significativo el efecto de las pruebas pre y post ( $p < 0,001$ ). Por lo que, de nuevo, concluimos que el número de respuestas correctas difiere entre las dos pruebas ( $F=193,75$ ) con un tamaño del efecto alto  $\eta^2 = 0,644$

Es significativo también el efecto de la interacción ( $F=5,548$  ;  $p < 0,05$ ) por lo que puede afirmarse que el efecto de las pruebas pre y post no es la misma en los dos grupos de alumnos.

**Tabla 20. Pruebas de efectos intra-sujetos. Análisis de varianza de medidas repetidas. Diferencias en las puntuaciones de CERES1 y CERES2 entre el grupo control y el grupo de intervención.**

Efectos intrasujetos	Suma de cuadrados tipo III	Gl.	Media cuadrática	F	Sig.	Eta al cuadrado parcial
CERES1/CERES2	123,735	1	123,735	193,75	0,00	0,644
CERES GRUPO A-GRUPOB	3,543	1	3,543	5,548	0,02	0,049
$\alpha = 0,05$						

En función de la validación de contrastes intrasujetos podemos concluir que:

- Es significativa la tendencia lineal, ( $F=193,753$  ;  $p<0,001$ ) por lo que podemos decir que en la prueba de razonamiento estadístico realizada al finalizar la asignatura tuvo un aumento en el número de respuestas correctas. El conjunto de alumnos aumentó su razonamiento estadístico a lo largo del curso.
- Es significativo también ( $F=5,548$  ;  $p< 0,05$ ) el contraste de interacción de las pruebas de razonamiento estadístico (CERES) en relación a los grupos de intervención (GRUPO A) y de control (GRUPO B). Resultando diferencias significativas entre los grupos. Estas diferencias se pueden observar en la Figura 50 en la que observamos que el grupo de intervención muestra una tendencia lineal más marcada que la del grupo control, más cercana a la horizontal.

Tabla 21. **Medias marginales del cuestionario de razonamiento estadístico (CERES) pre y postest.**

	Media	Error típico	Intervalo de confianza 95%
CERES 1	3,916	0,79	( 3,760 ; 4,073 )
CERES 2	5,423	0,92	( 5,240 ; 5.606 )
<b>Grupo Intervención</b>			
CERES 1	3,955	0,111	( 3,734 ; 4,175 )
CERES 2	5,716	0,130	( 5,459 ; 5,974 )
<b>Grupo Control</b>			
CERES 1	3,878	0,112	( 3,655 ; 4,100 )
CERES 2	5,130	0,131	( 4,870 ; 5,389 )



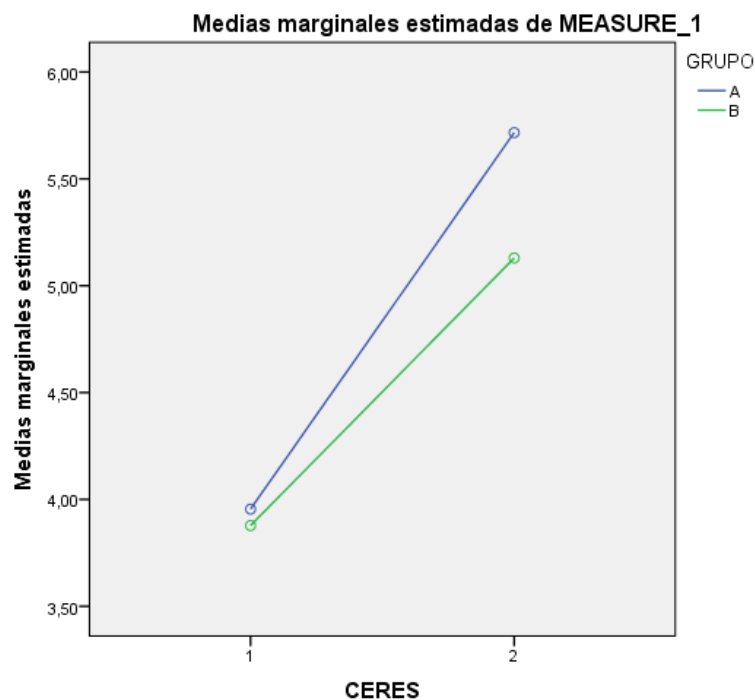


Figura 50. Gráfico de perfil de las medias marginales estimadas de CERES de los grupos de intervención y grupo control.

De nuevo, con las pruebas de los efectos iter-sujetos (tabla 22), se concluye que existen diferencias significativas en el número de respuestas correctas en el cuestionario de razonamiento estadístico CERES en función del grupo (grupo de intervención y grupo control ( $F=6,200$ ,  $p < 0,05$ ). Podemos decir que el grupo de intervención (grupo A) ha obtenido un promedio de respuestas correctas en CERES2 significativamente mayor al grupo control (grupo B)

Tabla 22. Pruebas de los efectos inter-sujetos. Análisis de varianza de medidas repetidas. Diferencias en las puntuaciones de CERES1 y CERES2 entre el grupo control y el grupo de intervención.

Efectos inter-sujetos	Suma de cuadrados tipo III	Gl.	Media cuadrática	F	Sig.	Eta al cuadrado parcial
intersección	4753,083	1	4753,083	4913,566	0,00	0,979
CERES1/CERES2	5,998	1	5,998	6,200	0,014	0,055
$\alpha = 0,05$						

Los resultados expuestos nos llevan a verificar la hipótesis, y concluir que la realización en el transcurso de la asignatura, de cuatro pruebas de razonamiento estadístico que abarcaron, entre las cuatro, los contenidos del curso, ha dado como resultado una mayor comprensión conceptual de la estadística en comparación con el grupo que no realizó estas pruebas.

#### **7.2.1.1. Comparativa de las respuestas correctas en CERES1 y CERES2**

Una vez que hemos comprobado que se ha producido un aumento significativo de la comprensión conceptual estadística en la totalidad de la muestra, nos interesa saber qué conocimiento tienen los alumnos cuando llegan a su primer curso universitario, en cuáles de ellos mejoran y cuáles permanecen con errores de razonamiento. El objetivo de este análisis es recoger información de utilidad para próximos cursos que nos permitirá saber cuál es el punto de partida respecto a los conocimientos de nuestros alumnos y en cuáles hay que poner el énfasis pues ni siquiera al finalizar la asignatura han quedado claros para la mayoría de los alumnos. Concretamente sobre esto último ha de centrarse nuestro mayor interés, tenerlo presente en cursos próximos e idear estrategias para conseguir que ciertos conceptos queden perfectamente entendidos.

La Tabla 23 recoge aquellos ítems que ya en CERES1 mostraron un porcentaje de respuestas correctas superior al 60%, lo que quiere decir que los estudiantes fueron capaces de comprender y resolver correctamente la pregunta que se les solicitaba.

Todos estos ítems, que inicialmente en el pretest fueron contestados por más de la mitad de los alumnos, utilizan representaciones gráficas para la formulación del problema y de la pregunta. Las cuestiones planteadas en ellos tiene que ver con la asociación de gráficos a una Tabla de datos y la comparación entre grupos.

Tabla 23. Ítems del cuestionario CERES con un porcentaje de respuestas correctas superior al 60% en el pretest y en el postest.

ÍTEMES CON UN PORCENTAJE DE RESPUESTAS CORRECTAS SUPERIOR A 60% TANTO EN CERES 1 COMO EN CERES 2			
Nº ITEM	% Repuestas correctas en el CERES1	% Repuestas correctas en el CERES2	Resultados de aprendizaje
3	87,5	60,36	Capacidad de reconocer dos formas de representar gráficamente el mismo conjunto de datos
6	77,67	83,78	Entender la necesidad de representar gráficamente las frecuencias para describir adecuadamente la distribución de una variable cuantitativa. Asociar el gráfico a la información que la tabla nos ofrece respecto al centro de la variable y su dispersión
9	60,71	81,08	Capacidad de asociar las distintas representaciones gráficas con el tipo de variable, representada y asociar el comportamiento de la gráfica a la variable concreta
19	70,53	85,58	Capacidad para comparar grupos, teniendo en cuenta, en las distribuciones, dónde se centran la mayoría de los datos al margen del comportamiento de algunos individuos de la muestra
20	78,57	69,36	Capacidad para comparar grupos, mediante la comparación de los promedios, extrayendo esta información de la representación de la distribución de datos

De los cinco ítems anteriores (Tabla 23), el número 3 (capacidad de reconocer dos formas de representar gráficamente el mismo conjunto de datos) y el 20 (Capacidad para comparar grupos, mediante la comparación de los promedios, extrayendo esta información de la representación de la distribución de datos) fueron contestados correctamente por un número de alumnos menor a los que lo hicieron en el pretest. El resto aumentó el número de respuestas correctas al finalizar la asignatura.

En 26 de los ítems las respuestas correctas en CERES2 supera el 50% (Tabla 24). Representan el conocimiento razonado de la estadística que los alumnos han adquirido a lo largo del proceso de aprendizaje del curso de introducción de esta materia.

Tabla 24. Ítems del cuestionario CERES con un porcentaje de respuestas correctas en el postest superior al 50%.

ÍTEMES CON UN % DE RESPUESTAS CORRECTAS SUPERIOR AL 50% EN CERES 2			
Nº ITEM	% Repuestas correctas en el CERES1	% Repuestas correctas en el CERES2	Resultados de aprendizaje
1	29,46	51,35	Capacidad de interpretar el resumen de datos recogido en una tabla de frecuencias
2	33,92	67,56	Capacidad para describir e interpretar en el contexto de los datos, la distribución de una variable representada en un histograma
3	87,5	60,36	Capacidad de reconocer dos formas de representar gráficamente el mismo conjunto de datos
4	41,96	52,25	Reconocer en un histograma la asimetría y el sesgo de una distribución de datos y su influencia sobre los valores de las medidas de tendencia central
6	77,67	83,78	Entender la necesidad de representar gráficamente las frecuencias para describir adecuadamente la distribución de una variable cuantitativa. Asociar el gráfico a la información que la tabla nos ofrece respecto al centro de la variable y su dispersión
7	6,25	88,28	Compresión del concepto de percentil y su interpretación en un contexto específico
8	31,25	57,65	Capacidad de asociar un histograma a la descripción de una variable. En este caso el de una variable que se distribuye de forma normal. Reconocer este tipo de variables y la distribución de datos con un gráfico simétrico, campana de Gauss
9	60,71	81,08	Capacidad de asociar las distintas representaciones gráficas con el tipo de variable, representada y asociar el comportamiento de la gráfica a la variable concreta
10	26,78	50,45	Capacidad de reconocer el comportamiento y la distribución de frecuencias de una variable aleatoria, con igual probabilidad en todos los posibles resultados
11	31,25	52,25	Reconocer la distribución de frecuencias de una variable sesgada negativamente y asociarla al comportamiento real de una variable.
13	48,21	63,06	Capacidad de estimar correctamente y comparar las desviaciones estándar para diferentes histogramas. Entender que la mayor desviación estándar se asocia por lo general a un gráfico con una mayor distancia de los datos al valor central de la distribución
19	70,53	85,58	Capacidad para comparar grupos, teniendo en cuenta, en las distribuciones, dónde se centran la mayoría de los datos al margen del comportamiento de algunos individuos de la muestra
20	78,57	69,36	Capacidad para comparar grupos, mediante la comparación de los promedios, extrayendo esta información de la representación de la distribución de datos

Nº ITEM	% Repuestas correctas en el CERES1	% Repuestas correctas en el CERES2	Resultados de aprendizaje
21	52,67	63,06	Entender que la comparación de dos grupos no requiere que las muestras tengan igual tamaño, especialmente si los dos conjuntos de datos son grandes
23	67,85	58,55	Comprensión del significado de la variabilidad en un conjunto de datos obtenido como resultado de medidas repetidas y en un contexto en el que se desea que esa variabilidad sea pequeña
24	43,75	57,65	Comprensión de la representatividad de la población de una muestra grande aleatoria. Ley de los grandes números
27	44,64	54,05	Capacidad de reconocer una interpretación incorrecta de un p-valor
28	50	54,95	Capacidad de reconocer una interpretación errónea de un p-valor
29	21,42	65,76	Comprensión del concepto de nivel de significación y la capacidad de asociarlo a la interpretación del p-valor, conociendo qué valores de p-valor son deseables en los estudios de investigación. Manejar como iguales dos formas de expresar el mismo concepto: valor de contraste y p-valor
30	58,92	63,96	Entender que aunque un análisis de significación no ponga de manifiesto diferencias significativas no quiere decir que el tratamiento, en este caso la dieta, no tenga efecto
31	53,57	57,65	Capacidad de entender que un diseño experimental con asignación aleatoria apoya la inferencia causal
32	24,10	54,05	Interpretación de los resultados de una prueba de significación cuando es rechazada la hipótesis nula
34	53,57	67,56	Capacidad para detectar una incorrecta interpretación del nivel de confianza, como el porcentaje de datos de la población cuyo contenido en nitrito potásico se encuentra dentro de los límites del intervalo de confianza
35	33,03	59,45	Capacidad para detectar una incorrecta interpretación del nivel de confianza como el porcentaje de todas las posibles muestras que tienen una media dentro de los límites del intervalo de confianza
36	49,10	64,86	Capacidad de interpretar correctamente un intervalo de confianza. Seguridad de que el intervalo de confianza contenga la verdadera media de la población
37	51,78	72,07	Capacidad de asociar la descripción de la relación de dos variables con la representación en un diagrama de dispersión

Los objetivos conseguidos sólo en parte por los estudiantes, respecto a los resultados de aprendizaje esperados, están representados por aquellos ítems cuyos porcentajes de respuestas no llegó al 50% en el cuestionario CERES2. Se recogen en la Tabla 25 y tienen que ver con la variabilidad, la correcta interpretación del p-valor y de los intervalos de confianza, con el error en la interpretación de la correlación como evidencia de la causa de un fenómeno, así como las características que hacen extrapolable un modelo de regresión.

Tabla 25. Ítems del cuestionario CERES con un porcentaje de respuestas correctas en el postest inferior al 50%.

ÍTEMS CON UN % DE RESPUESTAS CORRECTAS INFERIOR AL 50% EN CERES 2			
Nº ITEM	% Repuestas correctas en el CERES1	% Repuestas correctas en el CERES2	Resultados de aprendizaje
5	7,14	27,92	Reconocer en un histograma con puntuaciones atípicas la conveniencia de utilizar el rango intercuartílico como medida de variabilidad del conjunto de datos representado, frente a la desviación estándar y el rango, parámetros afectados por los valores extremos
12	11,60	29,72	Capacidad de estimar correctamente y comparar las desviaciones estándar para diferentes histogramas. Entender que la menor desviación estándar se asocia, por lo general, a un gráfico con una menor distancia de los datos al valor central de la distribución
14	26,78	41,44	Entender que una distribución de muestreo con una media mayor que la mediana es más probable que se desvíe a la derecha
15	44,64	42,34	Entender el concepto de desviación típica como un parámetro que informa de la dispersión, y la capacidad de interpretarlo en el contexto de un diagrama de cajas y patillas
16	8,03	12,61	Entender que los diagramas de caja no proporcionan estimaciones precisas de los porcentajes de los datos por encima o por debajo de los valores a excepción de los cuartiles
17	18,75	36,03	Comprensión del concepto de mediana y su interpretación en el contexto de un gráfico de cajas y patillas
18	25	36,03	Comprensión de los factores que permiten que los resultados obtenidos sobre una muestra sean generalizados a la población
22	16,07	20,72	Entender que las muestras pequeñas tienen mayor variabilidad que las muestras grandes
25	26,78	30,73	Comprensión de que la distribución de medias de un conjunto de muestras tiene una distribución normal de media próxima a la media poblacional

Nº ITEM	% Repuestas correctas en el CERES1	% Repuestas correctas en el CERES2	Resultados de aprendizaje
26	38,39	45,04	Capacidad de reconocer una interpretación correcta de un p-valor
33	44,64	38,73	Capacidad para detectar una incorrecta interpretación del intervalo de confianza, como los valores entre los que se encuentra la variable en cada individuo o elemento de la población, con una seguridad del 95%
38	23,21	45,04	Capacidad de describir la relación entre dos variables que se muestran en un gráfico de dispersión
39	23,21	45,04	Entender que la correlación no implica causalidad
40	8,03	16,21	Comprensión de que no es prudente extrapolar un modelo de regresión

Y por último, se detectan las cuestiones que se han puesto de manifiesto como errores en los conceptos y en su interpretación (Tabla 26), teniendo en cuenta que los porcentajes de respuestas correctas al comenzar la asignatura eran superiores a los encontrados en CERES2. En alguno de estos ítems las diferencias en los porcentajes no son muy grandes pero en cualquier caso consideramos que no se ha producido el esperado aumento de este porcentaje como consecuencia del aprendizaje. Entendemos, por ello, que los alumnos tienen deficiencias en la interpretación de un gráfico de cajas y patillas, en la interpretación de un intervalo de confianza, y en aspectos concretos sobre la variabilidad de los datos.

Tabla 26. ítems del cuestionario CERES con un aumento en los errores en el postest.

ÍTEMS CON AUMENTO DE ERRORES EN CERES 2			
Nº ITEM	% Repuestas correctas en el CERES1	% Repuestas correctas en el CERES2	Resultados de aprendizaje
3	87,5	60,36	Capacidad de reconocer dos formas de representar gráficamente el mismo conjunto de datos, el histograma y el gráfico de cajas.
15	44,64	42,34	Entender el concepto de desviación típica como un parámetros que informa de la dispersión, y la capacidad de interpretarlo en el contexto de un diagrama de cajas y patillas
20	78,57	69,36	Capacidad para comparar grupos, mediante la comparación de los promedios, extrayendo esta información de la representación de la distribución de datos

Nº ITEM	% Repuestas correctas en el CERES1	% Repuestas correctas en el CERES2	Resultados de aprendizaje
23	67,85	58,55	Comprensión del significado de la variabilidad en un conjunto de datos obtenido como resultado de medidas repetidas y en un contexto en el que se desea que esa variabilidad sea pequeña
33	44,64	38,73	Capacidad para detectar una incorrecta interpretación del intervalo de confianza, como los valores entre los que se encuentra la variable en cada individuo o elemento de la población, con una seguridad del 95%

### 7.2.1.2. Diferencias en el razonamiento estadístico, medido con la prueba de respuesta abierta, entre el grupo de intervención y el grupo control.

Las diferencias en el razonamiento estadístico entre el grupo control y el grupo de intervención evidenciadas con los resultados del cuestionario CERES, también se ponen de manifiesto cuando utilizamos para evaluarlo la prueba de razonamiento estadístico de respuesta abierta.

La Tabla 27 recoge los resultados de la prueba t de student realizada para comprobar si los resultados en la prueba de razonamiento estadístico de respuesta abierta a final de curso fueron distintos entre ambos grupos.

Tabla 27. Prueba t de comparación de las medias de los resultados de la prueba de razonamiento estadístico de respuesta abierta entre el grupo de intervención y el grupo control.

	N	X	DT	Error típico	t	p	d (Cohen)
Grupo intervención	N=55	3,11	1,23	0,331	1,76	< 0,05	0,33
Grupo control	N=54	2,71	1,16	0,316			

El grupo de intervención obtuvo una media en las calificaciones de la prueba de razonamiento estadístico con respuesta abierta, significativamente mayor que el grupo control, al igual que ocurre con los resultados de razonamiento estadístico medidos con el cuestionario CERES. Como ya se ha visto anteriormente, las dos pruebas de evaluación (CERES y la prueba de respuesta abierta), fueron construidas para evaluar el razonamiento estadístico y



entre ellas existe una correlación positiva. Ambas pruebas tienen coincidencia en su objeto de medida y con sus resultados se pone de manifiesto la influencia de la intervención, al mostrar diferencias entre el grupo de intervención y el grupo control.

### 7.2.2. Hipótesis 2.

***Los alumnos a los que se les aplica una evaluación formativa con pruebas de evaluación del razonamiento estadístico obtienen mayores calificaciones en la prueba práctica de resolución de problemas que el grupo que no realiza pruebas de razonamiento estadístico.***

Se compara el rendimiento del examen de resolución de problemas entre el grupo de intervención y el grupo control, mediante una prueba t de student para muestras independientes. Los resultados pueden verse en la Tabla 28.

**Tabla 28. Prueba t de comparación de medias de los resultados de la prueba práctica de resolución de problemas entre el grupo de intervención y el grupo control.**

	N	X	DT	Error típico	t	p	d (Cohen)
Grupo intervención	N=55	6,95	2,38	0,32	4,703	<0,001	0,9
Grupo control	N=54	4,77	2,47	0,34			

Podemos concluir que el grupo de intervención tiene un mayor rendimiento en la prueba de evaluación de resolución de problemas que el grupo control. Esto, junto con las diferencias encontradas en los resultados de la evaluación con las dos pruebas de razonamiento estadístico (CERES2 y la prueba de respuesta abierta) entre el grupo de intervención y el grupo control, dan información para la comprobación de la hipótesis fundamental de este estudio:

***La comprensión conceptual estadística y el rendimiento académico de los alumnos se incrementa cuando se utiliza en la enseñanza de esta materia una estrategia de evaluación diseñada para desarrollar en el razonamiento estadístico.***

### 7.2.3 Hipótesis 3.

***La percepción de la competencia estadística en los alumnos que han realizado pruebas de evaluación de razonamiento estadístico es mayor que la percepción que tienen los alumnos que no han realizado ese tipo de pruebas.***

Se utiliza un análisis de varianza de medidas repetidas para indagar sobre las diferencias entre los dos grupos de alumnos, en las puntuaciones de la escala de percepción de la competencia estadística pre (EpCE1) y post (EpCE2).

Tabla 29. Prueba de Levene de igualdad de varianzas y estadísticos descriptivos de la comparación de la escala de percepción de la competencia estadística (EpCE) en el pretest y en el posttest.

EpCE / GRUPO	Media	Desviación típica	N	Contraste de Levene sobre la igualdad de las varianzas
EpCE1				
Grupo intervención	50,42	6,408	55	F = 0,859; p=0,356
Grupo Control	48,83	5,589	54	
Total	49,63	6,041	109	
EpCE2				
Grupo Intervención	53,55	4,921	55	F = 0,459; p=0,500
Grupo Control	53,00	5,607	54	
Total	53,28	5,254	109	

La media de las puntuaciones de EPCE1 en el grupo asignado para la intervención es de 50,42 con una DT= 6,408 y en el grupo que se utilizará como control la media es de 48,83 y la desviación típica 5,589.

Al finalizar el curso las puntuaciones de esta misma escala (EPCE2) tiene una media de 53,55 , DT= 4,921 en el grupo de intervención, mientras que en el grupo control la media es de 53,00 y la desviación 5,607.

Se valida el supuesto de igualdad de las varianzas para los dos grupos, tanto para EPCE1 (F = 0,859; p=0,356) como para EPCE2 (F = 0,459; p=0,500).

La prueba de Box acepta también el supuesto de igualdad de varianzas (M de Box= 3,346; F=1,093; p=0,351), en este caso de las matrices de covarianza generadas en cada grupo (intervención y control).

**Tabla 30. Análisis de varianza de medidas repetidas. Diferencias en las puntuaciones de EPCE1 y EPCE2 entre el grupo control y el grupo de intervención.**

Lambda de Wilks	Valor	F	GL de las hipótesis	Gl del error	Sig.	Eta al cuadrado o parcial
EpCE1/EpCE2	0,756	34,595	1	107	0,000	0,244
EpCE GRUPO A- GRUPOB	0,993	0,703	1	107	0,404	0,007

A la vista de los resultados podemos rechazar la hipótesis de igualdad de medias y concluir que la puntuación media de EPCE2 es significativamente mayor que la puntuación de EPCE1, ( $F(1, 107) = 34,595$ ;  $p < 0,001$ ; Lambda de Wilks = 0,756, eta cuadrado parcial = 0.244 ).

Sin embargo no se encuentran diferencias significativas en las puntuaciones de la escala EPCE entre el grupo de intervención y el grupo control,  $F(1,107) = 0,703$ ;  $p > 0,05$ ; Lambda de Wilks = 0.993, eta cuadrado parcial = 0.007.

**Tabla 31. Pruebas de efectos intra-sujetos. Análisis de varianza de medidas repetidas. Diferencias en las puntuaciones de EPCE1 y EPCE2 entre el grupo control y el grupo de intervención.**

Efectos intrasujetos	Suma de cuadrados tipo III	Gl.	Media cuadrática	F	Sig.	Eta al cuadrado parcial
CERES1/CERES2	724,810	1	724,810	34,595	0,00	0,244
CERES GRUPO A-GRUPOB	14,718	1	14,718	0,703	0,404	0,007
$\alpha = 0,05$						

Examinando la significación de los efectos intrasujetos (Tabla 31), observamos que resulta significativo el efecto de las pruebas pre y post ( $p < 0,001$ ). Por lo que, de nuevo, concluimos que el número de respuestas correctas difiere entre las dos pruebas ( $F = 34,595$ ) con un tamaño del efecto alto  $\eta^2 = 0,244$

No es significativo el efecto de la interacción ( $F = 0,703$ ;  $p > 0,05$ ) por lo que no parece haber influencia de los grupos sobre las puntuaciones de la EPCE pre y post.

En función de la validación de contrastes intrasujetos podemos concluir que:

- Es significativa la tendencia lineal, ( $F=34,595$  ;  $p<0,001$ ) por lo que podemos decir que en la percepción de la competencia estadística aumentó al finalizar la asignatura.

- No es significativa ( $F=0,703$  ;  $p>0,05$ ) la interacción de la percepción de la competencia estadística (EPCE) en relación a los grupos de intervención (GRUPO A) y de control (GRUPO B). De forma gráfica se observa en el gráfico de las medias marginales (Figura 51)

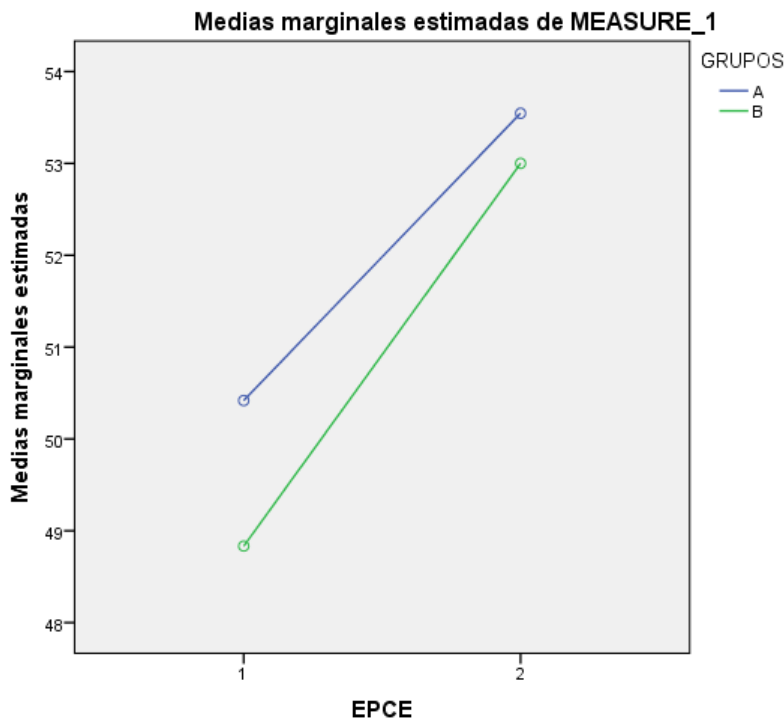


Figura 51. Gráfico de perfil de las medias marginales estimadas de EPCE de los grupos de intervención y grupo control.

La prueba de los efectos inter-sujetos (Tabla 42) concluye que no se han encontrado diferencias significativas ( $F= 1,437$ ;  $p>0,05$ ) en la puntuación de la escala EPCE en función del grupo (grupo de intervención y grupo control), la percepción de la competencia estadística es la misma en los dos grupos.

Tabla 32. Pruebas de los efectos inter-sujetos. Análisis de varianza de medidas repetidas. Diferencias en las puntuaciones de EpCE1 y EpCE2 entre el grupo control y el grupo de intervención.

Efectos inter-sujetos	Suma de cuadrados tipo III	Gl.	Media cuadrática	F	Sig.	Eta al cuadrado parcial
intersección	577002,782	1	577002,782	13407,847	0,00	0,992
CERES1/CERES2	61,828	1	61,828	1,437	0,233	0,013
$\alpha = 0,05$						

Respecto a la hipótesis que ha dado lugar al análisis de este apartado concluimos que: con los resultados expuestos anteriormente no podemos verificar la hipótesis y concluimos que la realización en el transcurso de la asignatura de las pruebas de razonamiento estadístico que constituyeron la intervención no ha evidenciado como resultado el aumento de la percepción de la competencia estadística en el grupo experimental. Los alumnos del grupo control y del grupo experimental muestran valores similares en la percepción de su competencia estadística.

A continuación se comprueban las hipótesis relativas a las posibles correlaciones entre variables. Algunas de estas correlaciones ya han sido expuestas en el capítulo de validación de instrumentos, como evidencia de validez. Volvemos a comentar ahora estos análisis desde el punto de vista de las repercusiones que este tipo de evaluación pueda tener sobre la capacidad del alumno para afrontar otro tipo de pruebas como es el caso de pruebas basadas en la resolución de problemas prácticos.

#### 7.2.4. Hipótesis 4.

***Existe una correlación significativa entre los resultados del cuestionario de evaluación de razonamiento estadístico y los de la prueba práctica de resolución de problemas.***

La correlación positiva y significativa ( $r = 0,263$ ;  $r^2 = 6,92$ ;  $p < 0,01$ ) entre el cuestionario que mide el razonamiento estadístico y los resultados del examen práctico nos lleva a pensar en los posibles beneficios que el estudio razonado de la materia puede tener sobre los resultados de aprendizaje asociados a la resolución de problemas: la capacidad de resolver preguntas sobre casos prácticos que están basados en datos reales de la práctica sanitaria, y de interpretar los resultados en ese mismo contexto.

### 7.2.5. Hipótesis 5.

***Las puntuaciones del cuestionario de evaluación del razonamiento estadístico correlacionan de forma significativa con la percepción que tienen los alumnos de su competencia estadística.***

No hemos podido comprobar la existencia de una posible correlación entre las puntuaciones del cuestionario de razonamiento estadístico y la percepción que tienen los estudiantes sobre su competencia estadística.

Al principio de curso ambas variables tenían un coeficiente de correlación de Pearson  $r = -0,147$  ( $p > 0,05$ ), y ese mismo coeficiente fue de  $r = 0,124$  ( $p > 0,05$ ) al final de curso. Este resultado es contradictorio con lo constatado por Muñoz (2002).

Sin duda, son muchas las variables que tienen que ver con la percepción de la competencia estadística que tienen los alumnos, la influencia del profesor, la experiencia previa con las matemáticas, algunas actividades como la resolución práctica de problemas estadísticos utilizando programas informáticos, etc., por lo que podría justificarse esta ausencia de correlación con un solo instrumento de evaluación.

### 7.2.6. Hipótesis 6.

***La percepción de la competencia estadística que manifiestan los alumnos, correlaciona de forma significativa con las calificaciones del examen práctico de resolución de problemas.***

No hemos podido demostrar la existencia de relación significativa entre la percepción de la competencia estadística y el examen práctico de resolución de problemas ( $r = 0,065$ ;  $p > 0,05$ ). Este resultado es contrario a lo obtenido en otros estudios (Muñoz 2002).

### 7.2.7. Hipótesis 7.

***La percepción de la competencia estadística que manifiestan los alumnos, correlaciona significativamente con las puntuaciones de las pruebas de razonamiento estadístico de respuesta abierta.***

Sin embargo sí parece existir una correlación estadísticamente significativa entre la percepción de competencia estadística y la prueba de razonamiento estadístico de respuesta abierta ( $r=0,219$ ;  $r^2=6,76$ ;  $p<0,05$ ).

Para concluir este capítulo resumimos, respecto a la comprobación de hipótesis, los siguientes resultados:

La utilización de pruebas de razonamiento estadístico a lo largo del curso, como parte de una evaluación formativa, aumenta en los alumnos la comprensión conceptual de la estadística. Este hecho se pone de manifiesto en las mejores calificaciones que obtienen los alumnos tanto en pruebas de razonamiento estadístico como de resolución de problemas prácticos. Hemos comprobado de forma comparativa con un grupo control, que el rendimiento académico de los alumnos sobre los que se ha llevado a cabo esta experiencia de evaluación con pruebas de razonamiento estadístico, es mayor.

La percepción de los alumnos de su competencia estadística aumenta al finalizar la asignatura, lo que muestra que el aprendizaje consigue que se sientan algo más competentes. Sin embargo a diferencia de lo que esperábamos, esa percepción no es mayor en los alumnos que se han enfrentado durante el curso a pruebas de razonamiento estadístico.

Esta percepción no parece tampoco estar relacionada con el rendimiento y las calificaciones que obtiene el alumno en las dos pruebas que han constituido la evaluación para la obtención de la calificación final de la asignatura (prueba de razonamiento estadístico con formato de prueba objetiva y prueba práctica de resolución de problemas). Es decir no tiene mayor percepción de su competencia estadística el estudiante que obtiene mejores resultados, y al contrario.





## **CONCLUSIONES DE LA INVESTIGACIÓN**

---



## CAPÍTULO 8. CONCLUSIONES

De la presente investigación se desprenden una serie de conclusiones de interés acerca del razonamiento estadístico como núcleo central de la enseñanza de la estadística. En concreto, de la utilidad de la evaluación de este razonamiento para conseguir en los alumnos una mayor comprensión de la materia. En este capítulo se resumen las principales aportaciones de los resultados de investigación del estudio al campo general de la educación estadística.

Se centrará, en una primera parte, en concluir dando respuesta a los objetivos de la investigación. Serán expuestas, también, otras conclusiones acerca de las aportaciones que esta tesis ha supuesto en nuestra labor docente y la oportunidad que, la realización de este trabajo, nos ha ofrecido para revisar el grado de coherencia entre la metodología y la evaluación que utilizamos y los objetivos que nos planteamos conseguir en la enseñanza de la estadística. En esta segunda parte, se expondrán unas reflexiones respecto a las aportaciones e implicaciones prácticas que el presente trabajo puede tener para nosotros y para otros profesores de estadística.

### 8.1. Objetivos y conclusiones

**Objetivo 1:** Realizar una búsqueda y análisis de instrumentos adecuados para medir el razonamiento estadístico.

Se ha encontrado un número reducido de instrumentos para evaluar el razonamiento estadístico que han sido validados y utilizados ampliamente en investigaciones y que se adaptan a los contenidos de un curso básico universitario. Las pruebas revisadas que cumplen estos criterios, tienen el formato de pruebas objetivas y coinciden en gran parte en los aspectos del razonamiento estadístico que evalúan así como en la forma de hacerlo.

No se ha localizado ningún instrumento para medir el razonamiento estadístico redactado en castellano y tampoco ninguno específico para las ciencias de la salud.

**Objetivo 2:** Desarrollar una serie de nuevos cuestionarios del razonamiento estadístico sobre los conceptos fundamentales que los alumnos universitarios de Enfermería han de adquirir en un curso de introducción a la estadística.

Se han construido, mediante la adaptación de cuestionarios que fueron seleccionados entre los ya existentes, una serie de instrumentos que evalúan el razonamiento estadístico y que están específicamente destinados a hacerlo en estudiantes de ciencias de la salud de un curso básico de estadística.

El análisis de ítems en la muestra piloto y la opinión de expertos llevó a la obtención de un instrumento de medición representativo de los contenidos que se consideraron ligados a la comprensión y razonamiento de la estadística de un curso básico universitario.

Los ítems probados han mostrado, con sus fortalezas y debilidades, una dificultad e índices de discriminación aceptables.

Todos Los contenidos y resultados de aprendizaje que evalúa el cuestionario de razonamiento estadístico han sido objeto de interés en los estudios publicados en artículos de investigación que tratan sobre el razonamiento estadístico y cumplen prácticamente con la totalidad de las directrices para la enseñanza y evaluación de la estadística de la American Statistical Association (ASA, 2005).

El instrumento construido es una adaptación del test CAOS, de probada utilidad, para medir el razonamiento de los estudiantes en diversas investigaciones con muestras extensas y variadas de estudiantes universitarios. La adaptación de la formulación de los ítems al contexto sanitario, utilizando situaciones reales, es una de las aportaciones más valiosas de este trabajo, así como la definición de los resultados de aprendizaje que, respecto a los contenidos, son evaluados en el cuestionario.

El cuestionario de evaluación del razonamiento estadístico (CERES), es un instrumento fiable que alcanza la representación del constructo "*razonamiento estadístico*". Respecto a su estructura está compuesto por ítems no suficientemente correlacionados, como es de esperar en un cuestionario que evalúa múltiples contenidos.

Este cuestionario (CERES) correlaciona con la otra prueba de razonamiento estadístico construida con un formato de respuesta abierta parcialmente restringida por lo que se concluye que, ambas pruebas, coinciden en su objeto de medida.

El cuestionario (CERES) correlaciona con las calificaciones del examen práctico de resolución de problemas. Los estudiantes que muestran un mayor

razonamiento estadístico tienen al mismo tiempo mejores calificaciones en la prueba de resolución de problemas.

Las pruebas parciales de razonamiento estadístico han mostrado también su utilidad, en la intervención realizada en el estudio empírico, como instrumentos de evaluación que han propiciado el desarrollo del razonamiento en los estudiantes.

**Objetivo 3:** Identificar cuáles son los logros del aprendizaje y de la comprensión conceptual de la estadística de los estudiantes a los que se aplica una acción formativa basada en la evaluación que tenga como objetivo fundamental el desarrollo del razonamiento estadístico.

Se comprueba que se produce un aumento significativo de la comprensión conceptual de la estadística al finalizar el curso en la totalidad de la muestra y en los grupos de intervención y grupo control por separado, aunque hay errores que permanecen. Estudios que se centran en la evaluación del razonamiento de los estudiantes de estadística han revelado que, incluso después de la instrucción, muchos siguen sin poder razonar correctamente sobre ideas y conceptos estadísticos importantes (Konold, Pollatsek, Bueno, Lohmeier y Lipson, 1993 ; Schau y Mattern , 1997 ; Hirsch y O'Donnell, 2001 ; Hirsch y O'Donnell, 2001 ; Garfield, 2003).

Se comprueba que los alumnos a los que se les aplica a lo largo del curso una evaluación formativa utilizando pruebas de evaluación del razonamiento estadístico, muestran una mayor comprensión conceptual de la estadística que el grupo de alumnos que no realiza ese tipo de evaluación.

Estas diferencias en el razonamiento estadístico entre el grupo control y el grupo de intervención han podido ser comprobadas cuando se utiliza como instrumento de medida la prueba de razonamiento estadístico de respuesta abierta.

La intervención evaluativa, utilizando pruebas de razonamiento estadístico, parece mejorar los resultados de los estudiantes en las pruebas prácticas de resolución de problemas.

Los alumnos a los que se les aplica una evaluación formativa con pruebas de evaluación del razonamiento estadístico obtienen mayores calificaciones en la prueba práctica de resolución de problemas que el grupo que no realiza pruebas de razonamiento estadístico.

La percepción de la competencia estadística de los alumnos aumenta al finalizar el curso como resultado del proceso de enseñanza-aprendizaje. Según diversos autores, los aspectos estadísticos que habrán desarrollado los estudiantes para mejorar su actitud hacia la estadística serán, por ejemplo, las habilidades para comunicarse estadísticamente, las habilidades de procedimientos, entender la lógica de los procesos estadísticos, el análisis exploratorio de datos y la lógica de las investigaciones estadísticas (Gal & Garfield, 1997; Batanero, 2001, Escalante 2010).

Una de las hipótesis planteadas en esta investigación exponía, como resultado esperado, que los alumnos que se habían enfrentado periódicamente a pruebas de evaluación de razonamiento estadístico y en los que los resultados confirman haber adquirido un mayor nivel en su capacidad para razonar, mostraran una mayor percepción de su competencia estadística. Sin embargo, no se ha podido constatar diferencias significativas en el aumento de esta percepción cuando se aplica un sistema de evaluación con instrumentos específicos para medir el razonamiento estadístico. De esta manera, no se pudo verificar la hipótesis en la que se esperaba que el aumento del razonamiento estadístico constatado en el grupo de intervención llevaría consigo también un aumento en su percepción de la competencia estadística, tal y como se apuntaba anteriormente.

El hecho de que en los resultados de este estudio no se encontraran diferencias significativas en la percepción de la competencia estadística entre el grupo control y el grupo de intervención difiere, por ejemplo, de lo afirmado por Escalante (2010). Este autor encuentra diferencias significativas ( $p < 0,05$ ) en la subescala que mide la competencia cognitiva cuando examina los cambios de actitudes en los alumnos de posgrado que participan en un módulo de estadística aplicado a la investigación. Este autor investigó si los alumnos desarrollaban actitudes más positivas hacia la estadística después del curso que implicó una estrategia didáctica específica, encontrando, como se ha mencionado, un aumento en la percepción de la competencia estadística.

En el presente estudio, las puntuaciones del cuestionario de evaluación del razonamiento estadístico no correlacionan de forma significativa con la percepción que tienen los alumnos de su competencia estadística. Tampoco se ha podido comprobar que la percepción de la competencia estadística que manifiestan los alumnos, correlaciona de forma significativa con las calificaciones del examen práctico de resolución de problemas.

Estos resultados se muestran contradictorios con lo constatado por Muñoz (2002) y Geske, Mickelson, Bandalos, Jonson y Smith, (2000), citados en Blanco (2008), en los que se encontró correlación de 0,33 entre la nota final del examen y la competencia cognitiva.

A la hora de justificar el hecho de que no se haya podido constatar una relación entre la percepción de la competencia cognitiva y el rendimiento en nuestros alumnos debería tenerse en cuenta que la percepción de la competencia estadística es considerada como un factor entre otros muchos de la actitud del alumno hacia la estadística. Cuando se trata de medir estas actitudes y su relación con el rendimiento en la asignatura, aparecen diferencias debidas, por ejemplo, a las muestras (alumnos de grado o postgrado), a la medida con la que se ha valorado el rendimiento, al componente actitudinal que ha sido medido (autopercepción de la competencia, gusto o agrado por la estadística, dificultad, etc.) o al instrumento usado para medir las actitudes (Blanco 2008). De forma importante, también influye sobre los resultados de esta relación, el momento del curso en el que se realiza la medición. Diversos estudios obtuvieron resultados de correlaciones entre las escalas del SATS (entre las que se encuentra la competencia cognitiva) y el rendimiento, que oscilaron entre .26 y .36 en las medidas pretest al comenzar el curso y entre .30 y .45 en la medida de las actitudes al término del curso cuando esta medida se realizaba junto con la evaluación del rendimiento (Sorge y Schau, 2002, Cashin y Elmore, 2005, citados en Blanco, 2008). Similares resultados aparecen también en el trabajo de Vanhoof et al. (2006).

Por último, se señala que no se han identificado hipótesis justificadas teóricamente sobre el patrón de relaciones de los distintos componentes de las actitudes con el rendimiento en la materia, Blanco (2008).

## 8.2. Limitaciones

Muchas son las limitaciones que pueden argumentarse a un estudio como este en el que se pretende atribuir los diferentes resultados observados a la intervención realizada en uno de los grupos de comparación. Estas limitaciones se refieren a la igualdad de los grupos, ya que en un entorno natural académico es muy complejo reorganizar alumnos para hacer una asignación aleatoria, limitaciones en el control de variables extrañas, las referentes al efecto sobre el aprendizaje del grupo control cuando recibe información del otro grupo o por el hecho de saber que están siendo estudiados.

Otra de las limitaciones se refiere al tamaño de la muestra en los dos grupos de comparación, que al no ser muy grande aporta el riesgo de no detectar las interacciones entre las variables. Respecto al número de alumnos hay cuestiones éticas, que se tuvieron en cuenta, derivadas de las posibles desventajas de alguno de los grupos por tener diferentes tratamientos. Especialmente el grupo control que no realiza una actividad de evaluación formativa que se prevé beneficiosa para su proceso de aprendizaje. La dificultad de justificar de forma convincente la posible desventaja y convencer de ello a los alumnos que participan son motivos por los cuales se optó por no aumentar el tamaño de los grupos de comparación con alumnos del curso siguiente. Además mantener constantes las variables extrañas cuando el tiempo que dura el estudio es largo, es difícil. Y se hace aún mucho más complicado el control de estas variables cuando se cuenta con otro profesor, de otro centro docente, opción que desechamos para asegurarnos que se mantenían constantes en los dos grupos las condiciones educativas, tanto del entorno como de la metodología empleada.

Sin embargo, y pese a las limitaciones de los estudios experimentales en el ámbito educativo que algunos autores han analizado (Kember, 2003), optamos por realizar una intervención con la que observar las mejoras del proceso de enseñanza- aprendizaje de nuestros alumnos durante un curso, siendo especialmente cuidadosos y teniendo siempre en cuenta los aspectos que limitan un estudio de este tipo.

A efectos de la comparación y valoración de los cambios producidos, podemos considerar que ambos grupos partían de unos niveles iniciales de razonamiento similares y característicos de alumnos que se matriculan por primera vez en un curso de estadística.

Una de las amenazas de la validez interna es la sensibilización por el pretest. En este caso, dicha sensibilización, colabora con los objetivos buscados pues supone ya de por sí un aprendizaje. La falta de aleatoriedad, frecuente en la investigación en educación, impide generalizar los resultados pero nos permite aplicar una acción educativa y observarla en sus escenarios naturales, nuestros alumnos, nuestras aulas y nuestra actividad docente. Este tipo de diseño tiene también un control parcial de otras variables por lo que han de tenerse en cuenta otras posibles explicaciones a los resultados.

Otra amenaza de la validez interna ha sido la posible contaminación entre los sujetos de ambos grupos. Los estudiantes del grupo de intervención y del



grupo control han coincidido en espacios y en jornadas lectivas por lo que la intervención no se ha producido de forma aislada. Por el contrario, ha sido conocida por todos los alumnos la aplicación de un sistema de evaluación sólo en uno de los grupos. Ciertamente ha existido una salida de información hacia el grupo control, aunque lo importante es que los estudiantes de este grupo no se han enfrentado periódicamente a unas pruebas que requieren un ejercicio de razonamiento, aunque hayan tenido información indirecta sobre ellas. Por tanto, ha sido mínimo el efecto que estas circunstancias hayan podido tener sobre los resultados.

Por otro lado, si bien se hizo todo lo posible por garantizar la igualdad en la instrucción, exceptuando la evaluación que constituye la intervención, en el grupo control y en el grupo experimental, inevitablemente puede haber existido alguna diferencia que haya sido origen de variabilidad.

Por tanto, como conclusión final con respecto del diseño y metodología de esta investigación, el presente estudio, sin pretensión de generalizar los resultados pero con una razonable validez interna, puede ser una fuente de información para nosotros y para otros docentes de la estadística. Pese a las limitaciones, los resultados nos sugieren unas importantes implicaciones prácticas para los profesores de estadística.

### **8.3. Reflexiones, aportaciones e implicaciones prácticas.**

Finalizada la exposición de las conclusiones que justifican los objetivos y valoran las hipótesis se pasa a desarrollar una serie de reflexiones, aportaciones e implicaciones prácticas acerca de los resultados de esta investigación.

Con respecto al marco teórico se destacan las ideas principales que, en concordancia con otros autores, han servido de guía a esta investigación:

En la concepción de este trabajo se ha considerado el enfoque constructivista sobre la enseñanza de la estadística, y la aplicación de los principios teóricos para conseguir un aprendizaje significativo. En este marco hemos pretendido justificar unas estrategias que consideramos adecuadas para que los estudiantes aprendan mejor la estadística y construyan su conocimiento mediante actividades que les ayuden a pensar y razonar.

Este marco nos ha hecho tener presente las múltiples formas con las que se afronta el proceso de aprendizaje, relacionado con el contexto, las circunstancias personales y las experiencias previas de los alumnos. Entender así

el aprendizaje nos lleva a asumir la responsabilidad de favorecer acciones educativas diferentes a las utilizadas tradicionalmente en la enseñanza de la estadística que estimulen el proceso de construcción de significados. El estudiante aprenderá un contenido sólo cuando sea capaz de atribuirle un significado, fijando relaciones fundamentadas entre aquello que pretende aprender y lo que pertenece a su conocimiento adquirido (Ausubel, Novak, Hanesian, 1983).

Por ello es importante obtener información de los conocimientos que posee el alumno y hacerlos relevantes para construir sobre ellos los nuevos conceptos estadísticos. En esa construcción será muy útil relacionarlo con el ámbito de su formación enfermera y aplicarlo a contextos diferentes que aporten nuevos significados y mayor grado de elaboración de los conceptos. Relacionar los conceptos nuevos con los adquiridos previamente o con experiencias personales, también consigue un mayor compromiso del alumno y una mayor profundidad en su proceso de aprendizaje. Para que los alumnos muestren un mayor interés por lo enseñado, se habrá de favorecer su implicación promoviendo el aprendizaje activo.

Hay que procurar también que la relación profesor-alumno facilite el intercambio de información y preguntas, que las tareas propuestas sean amenas y que tengan una vivencia positiva de su experiencia educativa que facilite el aprendizaje significativo

El pensamiento-razonamiento estadístico es la piedra angular de la enseñanza de la estadística y la función del docente debe situarse bajo la premisa de contribuir al desarrollo de esta capacidad en el alumnado (Wild y Pfannkuch, 1999; Pfannkuch y Wild, 2004; Moore, 2004; delMas, 2004; Franklin et al. 2007; Pfannkuch, 2008; Casey, 2008, 2010).

Una acción intencional en la evaluación formativa provoca efectos positivos sin precedentes en el rendimiento académico. Estos efectos son más importantes en los alumnos de bajo rendimiento (Garfield y col. 2011). Esta evaluación, además, facilita al profesor una información que puede serle útil para hacer cambios, sobre la marcha, en la enseñanza (Jordan, 2004; Mosteller, 1989, citados en Garfield, 2011).

Se deben tener en cuenta las recomendaciones GAISE en las propuestas educativas para desarrollar en los estudiantes el razonamiento estadístico (Garfield J., Ben-Zvi D., 2009). Estas pautas “son un buen punto de partida para

cambiar la enseñanza de la estadística en enfermería hacia un enfoque de cultura, razonamiento y pensamiento estadístico”. (Hayat, 2014. p. 197)

A continuación, volviendo a la génesis de este estudio, es necesario que nos remitamos a los documentos normativos para el grado de Enfermería con el fin de señalar aspectos importantes que han servido para motivar esta investigación. Dichos documentos indican que los titulados han de adquirir los recursos que les permitan basar la práctica profesional en la evidencia científica. La Orden CIN/2134/2008 recoge como competencias: *“Ser capaz, en el ámbito de la enfermería, de prestar una atención técnica y profesional adecuada a las necesidades de salud de las personas que atiende de acuerdo con el estado de desarrollo de los conocimientos científicos de cada momento y con los niveles de calidad...”,* y *“Basar las intervenciones de la enfermería en la evidencia científica y en los medios disponibles”*.

El problema radica en que, en la actualidad, existe una gran variedad de técnicas estadísticas a las que los profesionales de Enfermería han de enfrentarse para entender la investigación publicada. El grado de conocimiento de dichas técnicas determinará el nivel con que dichos profesionales podrán valorar críticamente los resultados de investigación. En resumen, no se podrán aplicar las mejores evidencias sin tener, entre otras cosas, acceso estadístico a la investigación.

Una de las principales consecuencias de esta situación está relacionada con la accesibilidad estadística. En un estudio que realizamos sobre este tema se concluyó que la accesibilidad estadística de un supuesto graduado en Enfermería que sólo haya cursado una asignatura básica de estadística, (tal como muestran los contenidos de los diferentes planes de estudios analizados), es limitada. Evidencia, además, un reducido acceso a la comprensión estadística de los artículos de investigación publicados.

Según dicho estudio un graduado en Enfermería podría entender los aspectos estadísticos del 85% de los artículos publicados en nuestro país. Para cubrir las necesidades del enfermero que se enfrenta a la lectura de las publicaciones de investigación de enfermería nacionales sólo necesitaría conocer determinados análisis como el análisis discriminante y la regresión múltiple. Sin embargo y considerando que en estas revistas no se publican los avances de vanguardia en investigación enfermera y que los alumnos de Enfermería de nuestras Universidades tienen conocimientos básicos que podemos encuadrar

en las técnicas bivariantes, sólo serían capaces de entender el 21,7% de artículos de enfermería internacionales de alto índice de impacto. Para la lectura crítica de estos artículos no sólo se ha de tener conocimientos de Estadística básica, sino haber profundizado en ellos, dada la diversidad de pruebas y aplicaciones específicas que se encuentran en las publicaciones.

La situación real en nuestras Escuelas y Facultades de Enfermería, es que los contenidos de Estadística que se imparten no son homogéneos en todas ellas. Incluso, en algunas, es inferior a los contenidos que conformarían un curso básico. Por tanto, el índice de accesibilidad estadística de la Enfermería en nuestro país es bastante limitado y se ha ido reduciendo aún más en los últimos años. Esto se ha debido, sobre todo, a que el nivel de la investigación de Enfermería internacional, en manos de enfermeros con niveles académicos que no existían hasta ahora en nuestro país, está muy por encima de la formación de nuestros profesionales. No obstante, los artículos nacionales han aumentado también el nivel estadístico por el impulso de la investigación enfermera en los últimos años y la facilidad de acceso a la producción científica internacional a través de Internet. Todo ello viene a remarcar lo que estamos exponiendo respecto a la baja accesibilidad estadística en los profesionales de Enfermería, tanto de las publicaciones nacionales, aunque presenten un menor nivel de complejidad, como, y sobre todo, de las internacionales.

Por último, parece interesante incluir una reflexión respecto a las publicaciones de los resultados de investigación. Con el paso del tiempo, las publicaciones se van consolidando e incluyendo artículos con métodos más complejos y análisis nuevos (Horton y Switzer 2005). Este creciente avance que, sin duda, está ocurriendo en nuestro país como consecuencia de los nuevos estudios de posgrado de Enfermería, pondrá, una vez más, en evidencia la necesidad de mejorar el proceso de enseñanza-aprendizaje de la Estadística, objeto principal de nuestro trabajo.

En relación con lo anteriormente expuesto sería muy interesante poder medir la capacidad que tienen los graduados en Enfermería de aplicar las mejores evidencias en mejora de su práctica, con instrumentos diseñados para ello, en los que estarían incluidos aspectos estadísticos, y de metodología de investigación. Los cuestionarios o instrumentos que se han encontrado en la revisión de trabajos de este tipo sobre residentes de medicina (Reznick et al. 1987; Novak et al. 2006) incluyen, como es de esperar, un gran número de ítems sobre el conocimiento de métodos básicos de Estadística y análisis de datos, (Sherudan y Pignone 2002). Es una preocupación en la formación médica y

consideran los conocimientos de estadística fundamentales para la práctica basada en la evidencia (Slawson y Shaughnessy 2005).

No hay estudios acerca de los errores en la información estadística que puedan localizarse en la literatura de enfermería, (Hayat et al. 2013), información que podría ser útil para establecer lo que habría que enseñar a los estudiantes de enfermería.

Hayat, Schmiege y Cook, (2014), es un equipo que trabajan como consultores estadísticos de profesores y estudiantes de enfermería. Publican en su artículo varias cuestiones que han detectado en su actividad consultora, de años de experiencia, por la que sugieren que existe una necesidad de hacer cambios en la formación de los estudiantes de enfermería en los distintos niveles académicos. De ellas recogemos sólo las correspondientes al grado:

- Incluir más preparación y práctica en la manipulación de bases de datos, análisis de los datos y procedimientos para mejorar la calidad de los datos.
- Resaltar la importancia de respetar la ley de protección de datos de carácter personal que se encuentran en las bases de información de ciencias de la salud.
- Enseñar el manejo de diferentes software estadísticos pero enseñando el fundamento de los análisis. Enseñar el “por qué” además del “Cómo”
- Proporcionar técnicas para el tratamiento de datos perdidos.

Como se vio en el capítulo 1 incluimos en este estudio un análisis de la presencia de los contenidos de estadística y de investigación en los planes de estudio de las universidades españolas en estos momentos en los que es relativamente reciente la implantación del grado. Pero sería interesante analizar estos contenidos en profundidad, así como la metodología utilizada en la enseñanza de la estadística y la formación de los docentes y observar la repercusión que pueda tener, a largo plazo, en los graduados de enfermería (Hayat et al. 2013).

Hay que cuidar que la estadística, en los programas de grado de enfermería, sea un componente curricular mejor adaptado a las necesidades actuales de los graduados en enfermería. La formación estadística tradicional resulta ineficaz tal como demuestran las investigaciones educativas. El manejo de fórmulas y cálculos no proporciona a los estudiantes habilidades de razonamiento estadístico para la lectura crítica de las publicaciones de ciencias de la salud.

En la actualidad, el rápido y fácil acceso a la investigación mundial que proporciona internet y la gran cantidad de trabajos de investigación que podemos disponer, convierte a la estadística en una herramienta imprescindible para seleccionar y enjuiciar aquellos resultados que puedan ser aplicados para mejorar los cuidados de enfermería. Los alumnos han de aprender esta materia entendiéndola conceptual y razonadamente para que pueda serle útil en su vida profesional. Para ello, los profesores de estadística, debemos seleccionar cuáles son los aspectos más importantes y las metodologías más adecuadas que pueden afianzar el conocimiento. El desarrollo de una formación estadística sobre una base sólida de razonamiento estadístico permite a los profesionales de enfermería pensar de manera crítica sobre la evidencia que puede mejorar su práctica profesional (Hayat et al. 2013).

*“El desarrollo eficaz de los conocimientos estadísticos aumentará el rigor de la investigación enfermera, y mejorará la aplicación de los conocimientos científicos a la práctica y conducirá a mejorar las oportunidades de financiación competitiva de los investigadores de enfermería. Desarrollará enfermeras bien formadas para colaborar de manera efectiva en un contexto multidisciplinar que contribuirá al desarrollo de un sistema sanitario más eficaz.” ( Hayat, Schmiege y Cook, 2014. p. 191)*

En esta tesis se ha puesto de manifiesto que, una acción continuada de evaluación con instrumentos pensados para valorar el razonamiento estadístico, puede ser una opción metodológica válida para mejorar el proceso de enseñanza aprendizaje de la estadística. Devuelve al estudiante información que le permite corregir errores, afianzar conceptos y orientarle hacia un estudio razonado mejorará, además, su capacidad de razonamiento en general.

A través de esta investigación se ha puesto de manifiesto que la utilización de la evaluación con instrumentos para medir el razonamiento estadístico ha sido exitosa. Ha logrado que el grupo experimental incremente significativamente su comprensión razonada de la estadística. También ha mejorado el rendimiento específico en la resolución de problemas prácticos por parte de los alumnos que han recibido este entrenamiento. Los resultados de los análisis estadísticos realizados parecen confirmar la eficacia que tiene, para el aprendizaje de la estadística, la utilización de pruebas específicas que evalúen el razonamiento. Estos resultados apuntan a la consecución del fin fundamental de esta tesis:

*La finalidad última de esta investigación es mejorar en nuestros alumnos el proceso de enseñanza-aprendizaje de la estadística, desarrollar el razonamiento estadístico, y mejorar el rendimiento académico.*

Por ello, los resultados que confirman que hemos mejorado la formación de los estudiantes de Enfermería, son para nosotros los más relevantes. Desde nuestra Escuela de Enfermería y Fisioterapia San Juan de Dios de la Universidad Pontificia Comillas de Madrid son múltiples las iniciativas que hemos desarrollado para dotar a los alumnos de Enfermería de herramientas para la investigación, estímulo para hacerlo y percepción de la importancia que tiene la investigación en la mejora de los cuidados de salud. Todo ello se hizo en esta Escuela mucho antes del cambio en los planes de estudios que han llevado a la Enfermería al nivel de grado y a la posibilidad de poder contar con doctores en esta disciplina. Se ha contado desde el año 1981 con un Certamen Nacional de Enfermería que ha premiado trabajos de investigación sobre cuidados, gestión y organización y docencia de la Enfermería. Cuando la mayor parte de las escuelas no incluían la estadística y la metodología de investigación, nos esforzamos para que nuestro plan de estudios acogiera esas asignaturas.

La investigación en el campo de la Enfermería tiene un potencial que se verá desarrollado, en un futuro próximo, por los alumnos que tenemos ahora en nuestras aulas. Es el momento de plantearse cómo facilitar la adquisición de las competencias investigadoras que hagan posible desarrollar la esencia de la Enfermería de forma plena.

Al revisar la bibliografía acerca del razonamiento estadístico hemos podido comprobar cómo nuestra experiencia docente nos lleva a compartir muchas de las cuestiones planteadas en las investigaciones acerca de la dificultad de la enseñanza de la estadística. Nos dimos cuenta que llevamos a cabo una enseñanza de la estadística que camina en paralelo con las directrices para el aprendizaje y la evaluación de la estadística en alumnos universitarios que enuncia la American Statistical Association, recomendaciones que son citadas habitualmente por investigadores del ámbito de la educación preocupados por el aprendizaje de la estadística.

Tomamos conciencia de la medida en que en nuestras evaluaciones nos preocupa medir la comprensión y el conocimiento que, por ser razonado, ha de perdurar. Es decir, nos dimos cuenta que evaluamos con cuestiones en las que el alumno ha de razonar por encima de otros aspectos de cálculo o de memoria. En realidad, la revisión que hemos realizado de los estudios sobre razonamiento estadístico, nos ha ofrecido una oportunidad para reflexionar sobre lo que

hemos hecho hasta ahora para que nuestros alumnos aprendan estadística y nos ha reforzado aquello que nuestra experiencia evidenciaba como fuente de buenos resultados.

Con esta investigación, hemos tratado también de manifestar que poseer un conocimiento razonado, al menos de los aspectos básicos de la Estadística, hará a nuestros alumnos portadores de un valor que les permitirá un mejor desempeño de su actividad.

Hemos intentado despertar en los alumnos la inquietud por desarrollar la esencia de la Enfermería de forma más plena, haciéndoles conscientes de la importancia de la investigación y, por ello, de la necesidad de contar con las herramientas que la hagan posible. Hemos trabajado por conseguir aumentar su percepción acerca de su capacidad para abordar la información estadística con seguridad y competencia, de manera que les convierta en consumidores de la literatura que contribuye al aumento del conocimiento de la disciplina enfermera. Esto permitirá que puedan seguir formándose y actualizándose en su área de trabajo, y les permitirá también asentar su práctica profesional en las bases del conocimiento científico contribuyendo a la mejorar de los cuidados de la salud.

De cara a otros docentes de la estadística en ciencias de la salud, especialmente, hemos aportado una recopilación de interesantes fuentes de información acerca del razonamiento estadístico y su evaluación. Se han creado unos instrumentos que pueden ser utilizados o ser fuente de inspiración para plantear cuestiones para evaluar la estadística. Estos mismos instrumentos pueden ser usados en clase como ilustración de conceptos en los que los profesores encontramos dificultades para su comprensión. En resumen, este trabajo aporta materiales docentes nuevos y útiles, que abren diversas posibilidades de seguir utilizándolos para la investigación.

Como acciones futuras de investigación, apuntamos la posibilidad de crear una plataforma en internet, similar al proyecto ARTIST de la Universidad de Minnesota, que permita acceder a responder en línea los distintos cuestionarios de razonamiento estadístico. Estos instrumentos de evaluación estarían disponibles para profesores que quieran utilizarlos para evaluar el razonamiento estadístico de sus alumnos, o como material para conseguir una mayor comprensión de algunos conceptos estadísticos.



El profesor obtendría los resultados de la evaluación y podría identificar a sus alumnos mediante unas claves con las que cada estudiante habría realizado su prueba. Por otro lado la plataforma nos permitiría acumular multitud de datos como resultado de la evaluación con las distintas pruebas que servirían para realizar futuros estudios acerca de los propios instrumentos y sobre el razonamiento estadístico, identificando aspectos concretos en los que los estudiantes tienen mayores dificultades de aprendizaje y favoreciendo el desarrollo de estrategias consensuadas para su mejor abordaje.

Cuando planteábamos, en el primer capítulo, el problema de investigación que íbamos a tratar, expusimos nuestra preocupación acerca de que la estadística que enseñamos sea una herramienta útil en el ámbito social y profesional. Por otro lado, en diversos apartados de este trabajo se ha dado importancia a la enseñanza de la estadística como una herramienta que aumenta la competencia investigadora de la enfermería. Este es el aspecto que mueve nuestro interés para realizar investigaciones futuras: conocer la competencia investigadora de nuestros egresados y lo que la estadística ha aportado a dicha competencia. Pensamos que sería interesante que el estudio y la valoración de la competencia en investigación de los titulados en enfermería se hiciera vinculándolo, de alguna forma a la realización del trabajo fin de grado. Dado que en la mayor parte de centros docentes, dicho trabajo se está centrando en el diseño de proyectos, ya sean profesionales, educativos o de investigación, sería el momento ideal para evaluar el grado en el que los alumnos han interiorizado su aprendizaje de conceptos estadísticos de los primeros cursos y lo han integrado con el resto del saber enfermero para dar respuesta a los planteamientos de investigación que pueden surgir en su práctica profesional.

Y por último, como toda investigación, nos ha abierto un camino y nos ha señalado otros muchos por los que seguir aportando conocimientos para innovar y aumentar la calidad de la docencia.



## REFERENCIAS BIBLIOGRÁFICAS

---



## REFERENCIAS BIBLIOGRÁFICAS

- Aliaga, M., Cuff, C., Garfield, J. B., Lock, R., Utts, J., & Witmer, J. (2010). *Guidelines for assessment and instruction in statistics education* (GAISE College Report American Statistical association).
- Aliberas, J., Gutiérrez, R., & Izquierdo, M. (1989). La didáctica de las ciencias: Una empresa racional. *Enseñanza De Las Ciencias*, 7(3), 277-284.
- American Statistical Association. (2005). *GAISE college report*. December 12, 2013. Retrieved from <http://www.amstat.org/education/gaise/GAISECollege.htm>
- American Statistical Association. (2005). *GAISE endorsement*. December 15, 2013. Retrieved from <http://www.amstat.org/education/gaise/ASAEndorse.htm>
- Anastasiadou, S. D. (2011). Reliability and validity testing of a new scale for measuring attitudes toward learning statistics with technology. *Acta Didactica Napocensia*, 4(1), 1-10.
- Andrich, D. (1989). Statistical reasoning in psychometric models and educational measurement. *Journal of Educational Measurement*, 26(1), 81-90.
- Araya, V., Alfaro, M., & Andonegui, M. (2007). Constructivismo: Orígenes y perspectivas. *Revista Laurus*, 13(24), 76-92.
- Ausubel, D. P. (1963). *The psychology of meaningful verbal learning*. New York: Grune and Stratton.
- Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. New York: Holt, Rinehart & Winston.
- Ausubel, D. P., Novak, J. D., & Hanesian, H. (1983). *Psicología educativa [Educational Psychology]*. México: Trillas.
- Auzmendi, E. (1991). Factors related to attitudes toward statistics: A study with a spanish sample. *Paper presented at the annual meeting of the american educational research association*. Chicago.
- Aylward, G. P., & Verhulst, S. J. (1991). Data analysis techniques in behavioral pediatrics: Some practical advice. *Journal of Developmental and Behavioral Pediatrics*, 12(6), 370-377.

- Bakker, A., & Gravemeijer, K. (2004). Learning to reason about distribution. In D. Ben-Zvi, & J. B. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 147-168). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, New Jersey: Prentice Hall.
- Bandura, A. (1998). Personal and collective efficacy in human adaptation and change. In J. G. Adair, D. Belanger & K. L. Dion (Eds.), *Advances in psychological science: Vol.1. Personal, social and cultural aspects* (pp. 51-71). Hove, UK: Psychology Press.
- Bandura, A. (1999). Exercise of agency in personal and social change. In E. Sanavio (Ed.), *Behavior and cognitive therapy today: Essays in honor of Hans J. Eysenck* (pp. 1-29). Oxford: Anonima Romana.
- Bandura, A. (2001). Social Cognitive Theory: An agentic perspective. *Annual Review of Psychology*, 52, 1-26.
- Batanero, C. (2002). Discussion: The role of models in understanding and improving statistical literacy. *International Statistical Review*, 70(1), 37-40.
- Behar Gutierrez, R., & Grima Cintas, P. (2004). La estadística en la educación superior ¿Formamos pensamiento estadístico? *Ingeniería y Competitividad*, 5(2), 84-90.
- Ben-Zvi, D. (2004). Reasoning about data analysis. In D. Ben-Zvi, & J. B. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 121-146). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Ben-Zvi, D., & Garfield, J. B. (2004). Statistical literacy, reasoning, and thinking: Goals, definitions, and challenges. *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 3-15). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Berenson, M. L., Utts, J., Kinard, K. A., Rumsey, D. J., Jones, A., & Gaines, L. M. (2008). Assessing student retention of essential statistical ideas: Perspectives priorities and possibilities. *The American Statistician*, 62(1), 54-61.
- Berk, R. A., & Nanda, J. P. (1998). Effects of jocular instructional methods on attitudes, anxiety, and achievement in statistics courses. *Humor - International Journal of Humor Research*, 11(4), 383-410.
- Biehler, R. (1997). Students difficulties in practicing computer supported data analysis. In J. B. Garfield, & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics: 1996 proceedings of the 1996 IASE round table conference* (pp. 169-190). Voorburg, The Netherlands: International Statistical Institute.

- Biggs, J. B. (1979). Individual differences in study processes and the quality of learning outcomes. *Highereducation*, 8(4), 381-394.
- Biggs, J. B. (1984). Learning strategies, student motivation patterns, and subjectively perceived success. In J. R. Kirby (Ed.), *Cognitive strategies and educational performance* (pp. 111-134). Orlando, Florida: Academic Press.
- Biggs, J. B. (1985). The role of metalearning in study processes. *British Journal of Educational Psychology*, 55(3), 185-212.
- Biggs, J. B. (1987a). *Student approaches to learning and studying*. Hawthorn, Victoria: Australian Council for Educational Research.
- Biggs, J. B. (1987b). *Study process questionnaire (SPQ)*. Hawthorn, Victoria: Australian Council for Educational Research.
- Biggs, J. B. (1988). Assessing study approaches to learning. *Australian Psychologist*, 23(2), 197-206.
- Biggs, J. B. (1991). Approaches to learning in secondary and tertiary students in Hong Kong: Some comparative studies. *Educational Research Journal*, 6, 27-39.
- Biggs, J. B. (1993). What do inventories of students' learning processes really measure? A theoretical review and clarification. *British Journal of Educational Psychology*, 63(1), 3-19.
- Biggs, J. B. (1999). *Teaching for quality learning at university*. Buckingham, England: Open University Press.
- Bjornsdottir, A., & Garfield, J. B. (2008). Teaching bits: Statistics education articles from 2007. *Journal of Statistics Education*, 16(3), May 12, 2014. Retrieved from <http://www.amstat.org/publications/jse/v16n3/garfield.html>
- Bjornsdottir, A., & Garfield, J. B. (2009a). Teaching bits: Statistics education articles from 2008. *Journal of Statistics Education*, 17(1), December 2, 2013. Retrieved from <http://www.amstat.org/publications/jse/v17n1/garfield.pdf>
- Bjornsdottir, A., & Garfield, J. B. (2009b). Teaching bits: Statistics education articles from 2009. *Journal of Statistics Education*, 17(3), November 11, 2013. Retrieved from <http://www.amstat.org/publications/jse/v17n3/garfieldTB.html>
- Bjornsdottir, A., & Garfield, J. B. (2010). Teaching bits: Statistics education articles from 2009 and 2010. *Journal of Statistics Education*, 18(2), 1-5. January 27, 2014. Retrieved from <http://www.amstat.org/publications/jse/v18n2/garfieldtb.pdf>

- Bjornsdottir, A., & Garfield, J. B. (2012a). Teaching bits: Statistics education articles from 2011 and 2012. *Journal of Statistics Education*, 20(1), 1-5. May 12, 2014. Retrieved from <http://www.amstat.org/publications/jse/v20n1/garfieldtb.pdf>
- Bjornsdottir, A., & Garfield, J. B. (2012b). Teaching bits: Statistics education articles from 2012. *Journal of Statistics Education*, 20(3), 1-5. January 11, 2014. Retrieved from <http://www.amstat.org/publications/jse/v17n1/garfield.pdf>
- Bjornsdottir, A., & Garfield, J. B. (2013). Teaching bits: Statistics education articles from 2012 and 2013. *Journal of Statistics Education*, 21(1), 1-4. February 7, 2014. Retrieved from <http://www.amstat.org/publications/jse/v21n1/garfieldtb.pdf>
- Blanco, A. (2004). Hacia una enseñanza universitaria centrada en el aprendizaje In J. C. Torre Puente, & E. Gil Coria (Eds.), *Enseñar y aprender estadística en las titulaciones universitarias de ciencias sociales* (pp. 143-190). Madrid: Universidad Pontificia Comillas.
- Blanco, A. (2008). Una revisión crítica de la investigación sobre las actitudes de los estudiantes universitarios hacia la estadística. *Revista Complutense De Educación*, 19(2), 311-330. December 13, 2013. Retrieved from <http://revistas.ucm.es/index.php/RCED/article/viewFile/RCED0808220311A/15466>
- Bloom, B. S., Englehart, M. B., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives, the classification of educational goals – handbook I: Cognitive domain*. New York: McKay.
- Bloom, B. S. (1979). *Taxonomía de los objetivos de la educación: Ámbito del conocimiento*. Alcoy, Alicante: Marfil.
- Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18(3), 355-387.
- Boix-Mansilla, V., & Gardner, H. (1997). What are the qualities of understanding In M. S. Wiske (Ed.), *Teaching for understanding: Linking research with practice* (pp. 161-196). San Francisco: Jossey-Bass Publishers.
- Bond, M. E., Perkins, S. N., & Ramirez, C. (2012). Students' perceptions of statistics: An exploration of attitudes, conceptualizations, and content knowledge of statistics. *Statistics Education Research Journal*, 11(2), 6-25. December 16, 2013. Retrieved from [http://iase-web.org/documents/SERJ/SERJ11\(2\)\\_Bond.pdf](http://iase-web.org/documents/SERJ/SERJ11(2)_Bond.pdf)
- Bradstreet, T. E. (1996). Teaching introductory statistics courses so that nonstatisticians experience statistical reasoning. *The American Statistician*, 50(1), 69-78.



- Brase, G. L. (2009). Pictorial representations in statistical reasoning. *Applied Cognitive Psychology*, 23(3), 369-381.
- Bruner, J. S. (1960). *The process of education*. New York: Vintage books.
- Bruner, J. S. (1961). The art of discovery. *Harvard Educational Review*, (31), 21-32.
- Bryce, G. R. (2003). Data driven experiences in an introductory statistics course for engineers using student collected data. *1992 Proceedings of the Section on Statistical Education*, pp. 155-160.
- Bryce, G. R. (2005). Developing tomorrow's statisticians. *Journal of Statistics Education*, 13(1), March 16, 2012. Retrieved from <http://www.amstat.org/publications/jse/v13n1/bryce.html>
- Budgett, S., & Pfannkuch, M. (2010). Assessing students' statistical literacy: An international perspective . In P. Bidgood, N. Hunt & F. Jolliffe (Eds.), *Assessment methods in statistical education* (pp. 103-121). West Sussex, United Kingdom: John Wiley & Sons, Ltd.
- Calzadilla, M. E. (2002). Aprendizaje colaborativo y tecnologías de la información y la comunicación. *Revista Iberoamericana De Educación*, 1(10), 1-10.
- Campbell, D. T., & Stanley, J. C. (1963). Diseños experimentales y cuasi-experimentales para la investigación sobre la enseñanza. In N. L. Gagé (Ed.), *Handbook for research on teaching*. Chicago: Rand McNally.
- Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research* [Diseños experimentales y cuasiexperimentales en la investigación social. Buenos Aires: Amorrortu. 1973]. Chicago: Rand McNally.
- Carlin, J. B., & Doyle, L. W. (2000). 3: Basic concepts of statistical reasoning: Standard errors and confidence intervals. *Journal of Paediatrics and Child Health*, 36(5), 502-505.
- Carlin, J. B., & Doyle, L. W. (2001). Basic concepts of statistical reasoning: Hypothesis tests and the t-test. *Journal of Paediatrics and Child Health*, 37(1), 72-77.
- Carlson, K. A., & Winqvist, J. R. (2011). Evaluating an active learning approach to teaching introductory statistics: A classroom workbook approach. *Journal of Statistics Education*, 19(1), February 6, 2014. Retrieved from <http://www.amstat.org/publications/jse/v19n1/carlson.pdf>

- Carmona Márquez, J. (2004). Una revisión de las evidencias de fiabilidad y validez de los cuestionarios de actitudes y ansiedad hacia la estadística. *Statistics Education Research Journal*, 3(1), 5-28. April 1, 2014. Retrieved from [https://www.stat.auckland.ac.nz/~iase/serj/SERJ7\(1\).pdf](https://www.stat.auckland.ac.nz/~iase/serj/SERJ7(1).pdf)
- Carré Llopis, M. C., Jiménez Villa, J., Martín Mateo, M., & Jané Carreña, F. (1996). La estadística en la investigación clínica de medicamentos: Estudio de artículos originales procedentes de centros españoles. *Medicina Clínica*, 106(16), 611-616.
- Cashin, S. E., & Elmore, P. B. (2005). The survey of attitudes toward statistics scale: A construct validity study. *Educational and Psychological Measurement*, 65(3), 509-524.
- Chan, S. W., & Ismail, Z. (2012). The role of information technology in developing students' statistical reasoning. *Procedia - Social and Behavioral Sciences*, 46, 3660-3664.
- Chance, B. L. (1997). Experiences with authentic assessment techniques in an introductory statistics course. *Journal of Statistics Education*, 5(3), March 29, 2012. Retrieved from <http://www.amstat.org/publications/jse/v5n3/chance.html>
- Chance, B. L. (2002). Components of statistical thinking and implications for instruction and assessment. *Journal of Statistics Education*, 10(3), October 16, 2007. Retrieved from <http://www.amstat.org/publications/jse/v10n3/chance.html>
- Chance, B. L. (2005). Integrating pedagogies to teach statistics. In J. B. Garfield (Ed.), *Innovations in teaching statistics* (pp. 101-112). Washington, DC: Mathematical Association of America.
- Chance, B. L., Delmas, R. C., & Garfield, J. B. (2004). Reasoning about sampling distribution. In D. Ben-Zvi, & J. B. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 295-323). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Chance, B. L., Delmas, R. C., & Garfield, J. B. (1999). A model of classroom research in action: Developing simulation activities to improve students' statistical reasoning. *Journal of Statistics Education*, 7(3), March 15, 2012. Retrieved from <http://www.amstat.org/publications/jse/secure/v7n3/delmas.cfm>
- Chance, B. L., & Garfield, J. B. (2002). New approaches to gathering data on student learning for research in statistics education. *Statistics Education Research Journal*, 1(2), 38-41. May 5, 2014. Retrieved from [https://www.stat.auckland.ac.nz/~iase/serj/SERJ1\(2\).pdf](https://www.stat.auckland.ac.nz/~iase/serj/SERJ1(2).pdf)

- Chance, B.L., Delmas, R.C., & Garfield, J. B. (2005). Reasoning about sampling distributions. In D. Ben-Zvi, & J. B. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 295-323). Dordrecht, The Netherlands: kluwer Academic Publishers.
- Cheng, P. W., Holyoak, K. J., Nisbett, R. E., & Oliver, L. M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology*, 18(3), 293-328.
- Cobb, G. (1992). Teaching statistics. In L. A. Steen (Ed.), *Heeding the call for change: Suggestions for curricular action* (pp. 3-43). Washington, DC: The Mathematical Association of America.
- Cobb, G. W. (1998). The objective-format question in statistics: Dead horse, old bath water, or overlooked baby? *Annual meeting of the american educational research association*. San Diego, California.
- Cobb, P. (1994). Where is the mind? constructivist and sociocultural perspectives on mathematical development. *Educational Researcher*, 23(7), 13-20.
- Cobb, P., McClain, K., & Gravemeijer, K. (2003). Learning about statistical covariation. *Cognition and Instruction*, 21(1), 1-78.
- Coll, C. (1988). Significado y sentido en el aprendizaje escolar. Reflexiones en torno al concepto de aprendizaje significativo. *Infancia y Aprendizaje*, (41), 131-142.
- Coll, C. (1996). Constructivismo y educación escolar: Ni hablamos siempre de lo mismo ni lo hacemos siempre desde la misma perspectiva epistemológica. *Anuario De Psicología*, 69, 153-178.
- Coll, C., Martín, E., Mauri, T., Miras, M., Onrubia, J., Solé, I., et al. (1990). *El constructivismo en el aula*. Barcelona: Graó.
- Connor-Greene, P. A. (2000). Assessing and promoting student learning: Blurring the line between teaching and testing. *Teaching of Psychology*, 27(2), 84-88.
- Cooper, L. L., & Shore, F. S. (2008). Students' misconceptions in interpreting center and variability of data represented via histograms and stem-and-leaf plots. *Journal of Statistics Education*, 16(2), March 24, 2014. Retrieved from <http://www.amstat.org/publications/jse/v16n2/cooper.html>
- Costa, A. L. (1989). Redirecting Assessment: Reassessing assessment. *Educational Leadership*, 46(7), January 9, 2014. Retrieved from [http://www.ascd.org/ASCD/pdf/journals/ed\\_lead/el\\_198904\\_costa.pdf](http://www.ascd.org/ASCD/pdf/journals/ed_lead/el_198904_costa.pdf)

- Cravero, M., Redondo, Y., & Santellán, S. (2011). Competencias en educación estadística. De una alfabetización estadística hacia una alfabetización científica. *XIII Conferencia Interamericana de Educación Matemática CIAEM*, Recife. Brasil.
- Cumming, G. (2007). Inference by eye: Pictures of confidence intervals and thinking about level of confidence. *Teaching Statistics*, 29(3), 89-93.
- Daher, A. M., & Amin, F. (2010). Assessing the perceptions of a biostatistics and epidemiology module: Views of year 2 medical students from a Malaysian university. A cross-sectional survey. *Bmc Medical Education*, 10, 34-34.
- Davis, J. (2001). Conceptual change. In M. Orey (Ed.), *Emerging perspectives on learning, teaching, and technology*. December 8, 2013. Retrieved from <http://projects.coe.uga.edu/epltt/>
- De Solla Price, Derek J. (1963). *Little science, big science [Hacia una ciencia de la ciencia]* (J. M. López Piñero Trans.). New York: Columbia University Press.
- Delmas, R. C., Garfield, J. B., & Chance, B. L. (1998). Assessing the effects of a computer microworld on statistical reasoning. In L. Pereira-Mendoza (Ed.), *Proceedings of the fifth international conference on teaching statistics* (pp. 1083-1090). Voorburg, The Netherlands: International Statistical Institute.
- Delmas, R. C., Garfield, J. B., Ooms, A., & Chance, B. L. (2007). Assessing student's conceptual understanding after a first course in statistics. *Statistics Education Research Journal*, 6(2), 28-58. May 7, 2014. Retrieved from [https://www.stat.auckland.ac.nz/~iase/serj/SERJ6\(2\)\\_delMas.pdf](https://www.stat.auckland.ac.nz/~iase/serj/SERJ6(2)_delMas.pdf)
- Delmas, R. C., & Liu, Y. (2005). Exploring student's conceptions of the standard deviation. *Statistics Education Research Journal*, 4(1), 55-82. May 6, 2014. Retrieved from [http://www.stat.auckland.ac.nz/~iase/serj/SERJ4\(1\)\\_delMas\\_Liu.pdf](http://www.stat.auckland.ac.nz/~iase/serj/SERJ4(1)_delMas_Liu.pdf)
- Delmas, R. C., Ooms, A., Garfield, J. B., & Chance, B. L. (2006). Seventh international conference on teaching statistics. *Assessing Students' Statistical Reasoning*, Salvador, Bahia, Brazil.
- Delmas, R.C., Ooms, A., Garfield, J. B., & Chance, B. L. (2006). Assessing students' statistical reasoning. *Proceedings of the Seventh International Conference on Teaching Statistics*, University of Minnesota, 2006.

- Delmas, R. C. (2002). Statistical literacy, reasoning, and learning. *Journal of Statistics Education*, 10(3), March 16, 2012. Retrieved from [http://www.amstat.org/publications/jse/v10n3/delmas\\_intro.html](http://www.amstat.org/publications/jse/v10n3/delmas_intro.html)
- Delmas, R. C., Garfield, J. B., & Chance, B. L. (1999). A model of classroom research in action: Developing simulation activities to improve students' statistical reasoning. *Journal of Statistics Education*, 7(3), March 29, 2012. Retrieved from <http://www.amstat.org/publications/jse/secure/v7n3/delmas.cfm>
- Dempster, M., & McCorry, N. K. (2009). The role of previous experience and attitudes toward statistics in statistics assessment outcomes among undergraduate psychology students. *Journal of Statistics Education*, 17(2), May 14, 2014. Retrieved from <http://www.amstat.org/publications/jse/v17n2/dempster.html>
- Derry, S. J., Levin, J. R., Osana, H. P., Jones, M. S., & Peterson, M. (2000). Fostering students' statistical and scientific thinking: Lessons learned from an innovative college course. *American Educational Research Journal*, 37(3), 747-773.
- Derry, S., Levin, J., & Schauble, L. (1995). Stimulating statistical thinking through situated simulations. *Teaching of Psychology*, 22(1), 51-57.
- Díaz Barriga, A. (1989). Tendencias e innovaciones curriculares en la educación superior. *Revista de la Educación Superior*, 18(71)
- Dietz, E. J. (1993). A cooperative learning activity on methods of selecting a sample. *The American Statistician*, 47(2), 104-108.
- Draper, R. J. (2002). School mathematics reform, constructivism, and literacy: A case for literacy instruction in the reform-oriented math classroom. *Journal of Adolescent & Adult Literacy*, 45(6), 520-529.
- Driver, R., & Easley, J. (1978). Pupils and paradigms: A review of literature related to concept development in adolescent science students. *Studies in Science Education*, 5(1), 61-84.
- Ebel, R. L., & Frisbee, D. A. (1986). *Essentials of educational measurement* (4th ed.). Toronto, Canada: Prentice Hall.
- Emerson, J. D., & Colditz, G. A. (1983). Use of statistical analysis in the New England journal of medicine. *New England Journal of Medicine*, 309(709), 713.
- Emmioglu, E., & Capa-Aydin, Y. (2012). Attitudes and achievement in statistics: A meta-analysis study *Statistics Education Research Journal*, 11(2), 95-102. February 3, 2014. Retrieved from [http://iase-web.org/documents/SERJ/SERJ11\(2\)\\_Emmioglu.pdf](http://iase-web.org/documents/SERJ/SERJ11(2)_Emmioglu.pdf)

- Entwistle, N. J. (1987). *Understanding classroom learning*. Londres: Hodder and Stoughton.
- Entwistle, N. J., & Ramsden, P. (1983). *Understanding student learning*. Londres: Croom Helm.
- Entwistle, N. J. (1988). *La comprensión del aprendizaje en el aula*. Barcelona: Paidós.
- Entwistle, N. J., & Kozéki, B. (1985). Relationships between school motivation, approaches to studying and attainment among british and hungarian adolescents. *British Journal of Educational Psychology*, 55(2), 124-137.
- Ernst, M. D. (2009). Teaching inference for randomized experiments. *Journal of Statistics Education*, 17(1), March 12, 2013. Retrieved from <http://www.amstat.org/publications/jse/v17n1/ernst.html>
- Escalante Gómez, E. (2010). Actitudes de alumnos de posgrado hacia la estadística aplicada a la investigación *Encuentro*, 42(85), 27-38. January 12, 2014. Retrieved from <http://www.lamjol.info/index.php/ENCUENTRO/article/view/57/55>
- Everson, M., Zieffler, A., & Garfield, J. (2008). Implementing new reform guidelines in teaching introductory college statistics courses. *Teaching Statistics*, 30(3), 66-70.
- Fernández Aparicio, T., Miñana López, B., Guzmán Martínez-Valls, P., & Hita Villaplana, G. (2003). Uso de los análisis estadísticos en los artículos originales de actas urológicas españolas. accesibilidad para el lector. *Actas Urológicas Españolas*, 27(2), 103-109.
- Fillebrown, S. (1994). Using projects in an elementary statistics course for non-science majors. *Journal of Statistics Education*, 2(2), March 30, 2012. Retrieved from <http://www.amstat.org/publications/jse/v2n2/fillebrown.html>
- Finney, S., Barron, K. & Horst, S. J. (2013). *Assessment of statistical literacy*. July 15, 2013. Retrieved from <https://www.stat.auckland.ac.nz/~iase/islp/assess>
- Fitzallen, N. (2006). A model of students statistical thinking and reasoning about graphs in an ICT environment. In P. Grootenboer, R. Zevenbergen & M. Chinnappan (Eds.), *Identities, cultures and learning spaces (proceedings of 29 th annual conference of the mathematics education research group of australasia)* (pp. 203-210). Camberra, Australia: MERGA. November 26, 2013. Retrieved from <http://www.merga.net.au/documents/RP212006.pdf>

- Fitzallen, N. (2008). Validation of an assessment instrument developed for eliciting student prior learning in graphing and data analysis. In M. Goos, R. Brown & K. Makar (Eds.), *Navigating currents and charting directions (proceedings of the 31st annual conference of the mathematics education research group of australasia)* (pp. 203-209). Brisbane, Australia: MERGA. November 26, 2013. Retrieved from <http://www.merga.net.au/documents/RP212008.pdf>
- Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology*, 18(3), 253-292.
- Gadanidis, G. (1994). Deconstructing constructivism. *The Mathematics Teacher*, 87(2), 91-97.
- Gal, I. (1995). Statistical tools and statistical literacy: The case of the average. *Teaching Statistics*, 17(3), 97-99.
- Gal, I. (2002). Adult's statistical literacy: Meaning, components, responsibilities. *International Statistical Review*, 70(1), 1-25.
- Gal, I., & Garfield, J. B. (1997). *The assessment challenge in statistics education*. Amsterdam: IOS Press and the International Statistical Institute.
- Gálvez-Toro, A., Amescua, M., & Salido-Moreno, M. & Hueso-Montoro, C. (2006). Repercusión e impacto de las revistas de enfermería del espacio científico iberoamericano. *Index De Enfermería*, 57, 83-86.
- Gálvez-Toro, A., Hueso-Montoro, C., & Amescua, M. (2004). Indicadores CUIDEN de repercusión de las revistas de enfermería del área lingüística del español y del portugués. *Index De Enfermería*, 46, 76-80.
- García-López, J. A. (2000). Métodos estadísticos empleados en los artículos originales publicados sobre tabaquismo en cuatro revistas médicas españolas (1985-1996). *Revista Española de Salud Pública*, 74(1), 33-43.
- Gardner, H. (1991). *The unschooled mind: How children think and how schools should teach*. New York: Basic Books.
- Garfield, J. B. (1994). Beyond testing and grading: Using assessment to improve student learning. *Journal of Statistics Education*, 2(1), March 29, 2012. Retrieved from <http://www.amstat.org/publications/jse/v2n1/garfield.html>
- Garfield, J. B. (1995). How students learn statistics. *International Statistical Review*, 63(1), 25-34.

- Garfield, J. B. (1996). Assessing student learning in the context of evaluating a chance course. part A: Theory and methods. *Communications in statistics* (pp. 2863-2873)
- Garfield, J. B. (1998). Challenges in assessing statistical reasoning. *Paper Presented at the American Educational Research Association Annual Meeting*, San Diego, CA.
- Garfield, J. B. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education*, 10(3), March 18, 2014. Retrieved from <http://www.amstat.org/publications/jse/v10n3/garfield.html>
- Garfield, J. B. (2003). Assessing statistical reasoning. *Statistics Education Research Journal*, 2(1), 22-38. November 8, 2013. Retrieved from [http://www.stat.auckland.ac.nz/~iase/serj/SERJ2\(1\).pdf](http://www.stat.auckland.ac.nz/~iase/serj/SERJ2(1).pdf)
- Garfield, J. B., & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education*, 19(1), 44-63.
- Garfield, J. B., Aliaga, M., Cobb, G., Cuff, C., Gould, R., Lock, R., et al. (2005). *Guidelines for assessment and instruction in statistics education* (GAISE college report). Alexandria, VA: American Statistical Association.
- Garfield, J.B., & Ben-Zvi, D. (2005). A framework for teaching and assessing reasoning about variability. *Statistics Education Research Journal*, 4(1), 92-99. February 7, 2014. Retrieved from [http://iase-web.org/documents/SERJ/SERJ4\(1\)\\_Garfield\\_BenZvi.pdf](http://iase-web.org/documents/SERJ/SERJ4(1)_Garfield_BenZvi.pdf)
- Garfield, J. B., & Ben-Zvi, D. (2008). Assessment in statistics education. *Developing students' statistical reasoning: Connecting research and teaching practice* (pp. 65-90) Springer.
- Garfield, J. B., & Ben-Zvi, D. (2009). Helping students develop statistical reasoning: Implementing a statistical reasoning learning environment. *Teaching Statistics*, 31(3), 72-77.
- Garfield, J. B., & Delmas, R. C. (1991). Students' conceptions of probability. In D. Vere-Jones, S. Carlyle & B. P. Dawkins (Eds.), *Proceedings of the third international conference on teaching statistics* (pp. 340-349)
- Garfield, J.B., & Delmas, R. C. (1994). International conference on teaching statistics (ICOTS4). *Students' Informal and Formal Understanding of Statistical Power*, Marrakech, Morocco.



- Garfield, J. B., & Delmas, R. C. (2010). A web site that provides resources for assessing students' statistical literacy, reasoning and thinking. *Teaching Statistics*, 32(1), 2-7.
- Garfield, J. B., Delmas, R. C., & Chance, B. L. (2003). Assessment statistical reasoning to enhance educational quality of AERA annual meeting. *The Web Based ARTIST: Assessment Resource for Improving Statistical Thinking*, Chicago.
- Garfield, J.B., & Everson, M. (2009). Preparing teachers of statistics: A graduate course for future teachers. *Journal of Statistics Education*, 17(2), April 16, 2013. Retrieved from <http://www.amstat.org/publications/jse/v17n2/garfield.html>
- Garfield, J. B., & Gal, I. (1999). Assessment and statistics education: Current challenges and directions. *International Statistical Review*, 67(1), 1-12.
- Garfield, J.B., Zieffler, A., Kaplan, D., Cobb, G. W., Chance, B. L., & Holcomb, J. P. (2011). Rethinking assessment of student learning in statistics courses. *American Statistician*, 65(1), 1-10.
- Gnanadesikan, M., Scheaffer, R. L., Watkins, A. E., & Witmer, J. A. (1997). An activity-based statistics course. *Journal of Statistics Education*, 5(2), April 25, 2004. Retrieved from <http://www.amstat.org/publications/jse/v5n2/gnanadesikan.html>
- González-de-Dios, J. (2001). Valoración del nivel de calidad de la evidencia científica de anales españoles de pediatría. *Anales de Pediatría*, 54, 380-389.
- Gowin, D. B. (1981). *Educating*. Ithaca, New York: Cornell University Press.
- Granda-Orive, J. I., García-Río, F., Gutiérrez-Jiménez, J., Escobar-Sacristán, J., Gallego-Rodríguez, V., & Sáez-Valls, R. (2002). Uso y accesibilidad al análisis estadístico en la revista archivos de bronconeumología (1970-1999). *Archivos de Bronconeumología*, 38, 356-361.
- Gravemeijer, K. (2000). *A rationale for an instructional sequence for analyzing one- and two-dimensional data sets. Paper presented at the annual meeting of the american educational research association*. Montreal, Canada.
- Griffith, J. D., Adams, L. T., Gu, L. L., Hart, C. L., & Whitehead, P. N. (2012). Students' attitudes toward statistics across the disciplines: A mixed-methods approach. *Statistics Education Research Journal*, 11(2), 45-56.
- Haack, D. G. (1979). Teaching statistical literacy. *Teaching Statistics*, 1(3), 74-76.
- Haack, D. G. (1980). A note on? teaching statistical literacy? *Teaching Statistics*, 2(1), 22-23.

- Haladyna, T. M., Downing, S. M., & Rodríguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education, 15*(3), 309-333.
- Hampson, P. J., & Morris, P. E. (1996). *Understanding cognition*. Cambridge: MA Blackwell Publishers Inc.
- Hayat, M. J. (2014). Guidelines for assessment and instruction in statistics education (GAISE) : Extending GAISE into nurse education. *Journal of Nursing Education, 53*(4), 192-198.
- Hayat, M. J., Eckardt, P., Higgins, M., Kim, M., & Schmiede, S. J. (2013). Teaching statistics to nursing students: An expert panel consensus. *Journal of Nursing Education, 52*(6), 330-334.
- Hayat, M. J., Schmiede, S. J., & Cook, P. F. (2014). Perspectives on statistics education: Observations from statistical consulting in an academic nursing environment. *Journal of Nursing Education, 53*(4), 185-191.
- Hellems, M. A., Gurka, M. J., & Hayden, G. F. (2007). Statistical literacy for readers of pediatrics: A moving target. *Pediatrics, 119*(6), 1083-1088.
- Hinshaw, R. E., Burden, R., & Shriner, M. (2012). Supporting post-graduates' skill acquisition using components of constructivism and social learning theory. *Creative Education, 3*, 874-877.
- Hirsch, L. S., & O'Donnell, A. M. (2001). Representativeness in statistical reasoning: Identifying and assessing misconceptions. *Journal of Statistics Education, 9*(2). June 05, 2014. Retrieved from <http://www.amstat.org/publications/jse/v9n2/hirsch.html>
- Hogg, R. V. (1991). Statistical education: Improvements are badly needed. *The American Statistician, 45*(4), 342-343.
- Holmes, P. (2002). Teaching, learning and assessment: Complementary or conflicting categories for school statistics. In B. Phillips (Ed.), *Proceedings of the sixth international conference on teaching statistics*. Voorburg: International Statistical Institute.
- Hood, W., Creed, P. A., & Neumann, D. L. (2012). Using the expectancy value model of motivation to understand the relationship between student attitudes and achievement in statistics. *Statistics Education Research Journal, 11*(2), 72-85.

- Hoog, R. V. (1992). Workshop on statistics education. In L. A. Steen (Ed.), *Heeding the call for change: Suggestions for curricular action* (pp. 34-43). Washington: Mathematical Association of America.
- Huberty, C. J. (2000). Assessment of student performance in statistics. *Teaching Statistics*, 22(2), 44-48.
- Hunt, N. (2010). Individualised assessment in statistics. In P. Bidgood, N. Hunt & F. Jolliffe (Eds.), *Assessment methods in statistical education: An international perspective* (pp. 203-210). The Atrium southern gate, Chichester, west Sussex, United Kingdom: John Wiley & Sons, Ltd.
- Hunter, W. G. (1977). Some ideas about teaching design of experiments, with 25 examples of experiments conducted by students. *The American Statistician*, 31(1), 12-17.
- Icart-Isern, M. T., Gracia-García, S., Pulpón-Segura, A. M., Fuentelsanz, C., López, S., & Pedreny, R. (1999). Introducción y análisis estadístico de los originales publicados en dos revistas de enfermería y en dos de medicina (1995-1996). *Enfermería Clínica*, 9, 55-59.
- International Association for Statistical Education. (2005). Reasoning about distribution. *SRTL-4 report*. June 05, 2014. Retrieved from [http://srtl.fos.auckland.ac.nz/?page\\_id=123](http://srtl.fos.auckland.ac.nz/?page_id=123)
- Johnson- Laird, P. (1983). *Mental models. towards a cognitive science of language, inference, and consciousness*. Cambridge: Harvard University Press.
- Jordan, J. (2004). The use of orally recorded exam feedback as a supplement to written comments. *Journal of Statistics Education*, 12(1).
- Kahnemann, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.
- Keeler, C. M., & Steinhorst, R. K. (1995). Developing material for introductory statistics courses from a conceptual, active learning viewpoint. *Journal of Statistics Education*, 3(3), October 20, 2007. Retrieved from <http://www.amstat.org/publications/jse/v3n3/steinhorst.html>
- Keeler, C. M., & Steinhorst, R. K. (1995). Using small groups to promote active learning in the introductory statistics course: A report from the field. *Journal of Statistics Education*, 3(2), March 29, 2012. Retrieved from <http://www.amstat.org/publications/jse/v3n2/keeler.html>

- Kelly, G. A. (1955). *The psychology of personal constructs (vols 1-2)*. New York: W.W. Norton.
- Kelly, G. A. (1963). *A theory of personality - the psychology of personal constructs*. New York: W.W. Norton.
- Kogan, M. (2002). Education communities and academic identity. *Higher Education Quarterly*, 54(3), 207-2016.
- Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction*, 6(1), 59-98.
- Konold, C. (1995). Confessions of a coin flipper and would-be instructor. *The American Statistician*, 49(2), 203-209.
- Konold, C., & Higgins, T. L. (2003). Reasoning about data. In J. Kilpatrick, W. G. Martin & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 193-215). Reston, VA: NCTM.
- Konold, C., & Pollatsek, A. (2004). Conceptualizing an average as a stable feature of a noisy process. In D. Ben-Zvi, & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking more* (pp. 169-199). Dordrecht: Kluwer Academic Publishers.
- Konold, C., Pollatsek, A., Well, A. D., Lohmeier, J., & Lipson, A. (1993). Inconsistencies in students' reasoning about probability. *Journal for Research in Mathematics Education*, 24(5), 392-414.
- Labatut Portilho, E. M. (2004). Aprendizaje universitario: Un enfoque metacognitivo. Universidad Complutense de Madrid: Facultad de Educación. Departamento de Psicología Evolutiva y de la Educación.
- Lancaster, G. A., Chellaswamy, H., Taylor, S., Lyon, D., & Dowrick, C. (2007). Design of a clustered observational study to predict emergency admissions in the elderly: Statistical reasoning in clinical practice. *Journal of Evaluation in Clinical Practice*, 13(2), 169-178.

- Lane-Getaz, S. J. (2007). Toward the development and validation of the reasoning about  $p$ -values and statistical significance scale. In B. Phillips, & L. Weldon (Eds.), *Proceedings of the international statistical institute (ISI) / International association of statistical education (IASE) satellite conference on assessing student learning in statistics*. Voorburg, The Netherlands: International Statistical Institute. June 05, 2014. Retrieved from <http://www.stat.auckland.ac.nz/~iase/publications/sat07/Lane-Getaz.pdf>
- Lane-Getaz, S. J. (2007). Development and validation of a research-based assessment: Reasoning about P-values and statistical significance. Universidad de Minnesota.
- Lane-Getaz, S. J. (2008). Introductory and intermediate student's understanding and misunderstanding of  $p$ -values and statistical significance. Paper presented at the *Proceedings of the 11th International Congress on Mathematical Education (ICME)*. Monterrey, Mexico.
- Lane-Getaz, S. J. (2013). Development of a reliable measure of students' inferential reasoning ability. *International Association for Statistical Education*, 12(1), 20-47.
- Lawton, L. (2009). An exercise for illustrating the logic of hypothesis testing. *Journal of Statistics Education*, 17(2), May 4, 2013. Retrieved from <http://www.amstat.org/publications/jse/v17n2/lawton.html>
- Ledolter, J. (1995). Projects in introductory statistics courses. *The American Statistician*, 49(4), 364-367.
- Lerman, S. (1898). Constructivism, mathematics and mathematics education. *Educational Studies in Mathematics*, 20(2), 211-223.
- Lizarbe-Chocarro, M. (2007). La enfermería de cuidados intensivos y coronarios en España: Análisis bibliométrico. *Metas de Enfermería*, 10, 9-15.
- Lovett, M. C. (2001). A collaborative convergence on studying reasoning processes: A case study in statistics. In S. M. Carver, & D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp. 347-384). Mahwah, NJ: Lawrence Erlbaum.
- Lovett, M. C., & Greenhouse, J. B. (2000). Applying cognitive theory to statistics instruction. *The American Statistician*, 54(3), 196-206.
- MacDougall, M. (2008). Ten tips for promoting autonomous learning and effective engagement in the teaching of statistics to undergraduate medical students involved in short-term research projects. *Journal of Applied Quantitative Methods*, 3(3), 223-240.

- Macfarlane, B., & Grant, B. (2012). The growth of higher education studies: From forerunners to pathtakers. *Higher Education Research & Development*, 31(5), 621-624.
- Mackisack, M. (1994). What is the use of experiments conducted by statistics students? *Journal of Statistics Education*, 2(1), April 25, 2004. Retrieved from <http://www.amstat.org/publications/jse/v2n1/mackisack.html>
- Magel, R. (1998). Using cooperative learning in a large introductory statistics class. *Journal of Statistics Education*, 6(3)
- Marín Martínez, N., Solano Martínez, I., & Jiménez Gómez, E. (1999). Tirando del hilo de la madeja constructivista. *Enseñanza De Las Ciencias*, 17(3), 479-492.
- Martí, E. (1996). El constructivismo y sus sombras. *Anuario De Psicología*, 69, 3-18.
- Martínez Arias, R. (1995). *Psicometría: Teoría de los tests psicológicos y educativos*. Madrid: Síntesis.
- Marton, F., & Säljö, R. (1976). On qualitative differences in learning. I. outcome and process. *British Journal of Educational Psychology*, 46(1), 4-11.
- Marton, F., & Säljö, R. (1976). On qualitative differences in learning. II. outcome as a function of the learner's conception of the task. *British Journal of Educational Psychology*, 46(2), 115-127.
- Matthews, M. R. (1994). Vino viejo en botellas nuevas: Un problema con la epistemología constructivista. *Enseñanza De Las Ciencias*, 12(1), 79-88.
- McNiece, R. (2010). An assessment strategy to promote judgement and understanding of statistics in medical applications. In P. Bidgood, N. Hunt & F. Jolliffe (Eds.), *Assessment methods in statistical education* (pp. 123-131). The Atrium southern gate, Chichester, west Sussex, United Kingdom: John Wiley & Sons, Ltd.
- Metz, M. L. (2010). Using GAISE and NCTM standards as frameworks for teaching probability and statistics to pre-service elementary and middle school mathematics teachers. *Journal of Statistics Education*, 18(3), July 5, 2013. Retrieved from <http://www.amstat.org/publications/jse/v18n3/metz.pdf>
- Millman, J., & Greene, J. (1989). The specification and development of test of achievement and ability. In R. L. Linn (Ed.), *Educational measurement* (pp. 335-366). London: Macmillan.

- Moore, D. S. (1997). New pedagogy and new content: The case of statistics. *International Statistical Review*, 65, 123-165.
- Morales Vallejo, P. (2006). *Las pruebas objetivas: Normas, modalidades y cuestiones discutidas*. 23 Julio, 2013. Recuperado de <http://www.upcomillas.es/personal/peter/otrosdocumentos/PruebasObjetivas.pdf>
- Morales Vallejo, P. (2009). *Las pruebas objetivas*. Bilbao: Universidad de Deusto, Instituto de Ciencias de la Educación.
- Morales Vallejo, P. (2011). *Corrección de las pruebas objetivas teniendo en cuenta el nivel de seguridad en las respuestas*. 12 Julio, 2013. Retrieved from <http://www.upcomillas.es/personal/peter/otrosdocumentos/NivelSeguridad.pdf>
- Morales Vallejo, P. (2012). *Análisis de ítems en las pruebas objetivas*. Agosto 24, 2013. Retrieved from <http://www.upcomillas.es/personal/peter/otrosdocumentos/analisisitemspruebasobjetivas.pdf>
- Morales Vallejo, P. (2013). *Investigación experimental, diseños y contraste de medias*. Unpublished manuscript.
- Mora-Ripoll, R., Ascao-Terrén, C., & Sentís-Vilalta, J. (1995a). Tendencias actuales en la utilización de la estadística en medicina. Estudio de los artículos originales publicados en medicina clínica (1991-1992). *Medicina Clínica*, 104, 444-447.
- Mora-Ripoll, R., Ascao-Terrén, C., & Sentís-Vilalta, J. (1995b). Uso y presentación de la metodología estadística en los artículos originales publicados en medicina clínica durante 1993. *Medicina Clínica*, 105, 9-12.
- Mora-Ripoll, R., Ascao-Terrén, C., & Sentís-Vilalta, J. (1996). Uso actual de la estadística en investigación biomédica: Una comparación entre revistas de medicina general. *Medicina Clínica*, 106, 451-456.
- Moreira, M. A. (2005). Aprendizaje significativo crítico. *Indivisa: Boletín De Estudios e Investigación*, (6), 83-102.
- Moreno, R., Martínez, R. J., & Muñiz, J. (2004). Directrices para la construcción de ítems de elección múltiple. *Psicothema*, 16(3), 490-497.
- Moreno, R., Martínez, R. J., & Muñiz, J. (2006). New guidelines for developing multiple-choice items. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 2(2), 65-72.

- Morris, E. J., Joiner, R., & Scanlon, E. (2002). The contribution of computer-based activities to understanding statistics. *Journal of Computer Assisted Learning*, 18(2), 116-126.
- Muñiz, J., & Fonseca-Pedrero, E. (2008). Construcción de instrumentos de medida para la evaluación universitaria. *Revista De Investigación En Educación*, (5), 13-25.
- Muñoz San Roque, I. (2002). *Actitudes hacia la estadística y su relación con otras variables en alumnos universitarios del área de las ciencias sociales*. (Tesis doctoral). Universidad Pontificia Comillas de Madrid.
- Myers, C. B., & Myers, S. M. (2007). Assessing assessment: The effects of two exam formats on course achievement and evaluation. *Innovative Higher Education*, 31(4), 227-236.
- Navarro, L. (2004). Hacia una enseñanza universitaria centrada en el aprendizaje In J. C. Torre Puente, & E. Gil Coria (Eds.), *La alineación constructiva en el aprendizaje universitario* (pp. 111-142). Madrid: Universidad Pontificia Comillas.
- Nelson, D. (2009). Using simple linear regression to assess the success of the montreal protocol in reducing atmospheric chlorofluorocarbons. *Journal of Statistics Education*, 17(2). June 05, 2014. Retrieved from <http://www.amstat.org/publications/jse/v17n2/nelson.html>
- Neumann, D. L., Neumann, M. M., & Hood, M. (2010). The development and evaluation of a survey that makes use of student data to teach statistics. *Journal of Statistics Education*, 18(1).
- Nikiforidou, Z., Lekka, A., & Pange, J. (2010). Statistical literacy at university level: The current trends. *Procedia - Social and Behavioral Sciences*, 9, 795-799.
- Nolan, W. M., Beran, T., & Hecker, K. G. (2012). Surveys assessing students' attitudes toward statistics: A systematic review of validity and reliability. *Statistics Education Research Journal*, 11(2), 103-123.
- Novak, J. D. (1977). *A theory of education*. Ithaca, New York: Cornell University Press.
- Novak, J. D. (1982). *Teoría y práctica de la educación*. Madrid: Alianza Universitaria.
- Novak, J. D. (1988). Constructivismo humano: Un consenso emergente. *Enseñanza De Las Ciencias*, 6(3), 213-223.
- Nunnally, J. C., & Bernstein, I. H. (1995). *Teoría psicométrica* (3ª ed.). México: McGraw-Hill.



- Olani, A., Hoekstra, R. M., Harskamp, E., & Van Der Werf, G. (2011). Statistical reasoning ability, self-efficacy, and value beliefs in a reform based university statistics course. *Electronic Journal of Research in Educational Psychology, 9*(1), 49-72.
- Orden SAS/1349/2009, de 6 de mayo, por la que se aprueba y publica el programa formativo de la especialidad de enfermería obstétrico-ginecológica (matrona), (2009).
- Orden SPI/1356/2011, de 11 de mayo, por la que se aprueba y publica el programa formativo de la especialidad de enfermería de salud mental, (2011).
- Orden SAS/1348/2009, de 6 de mayo, por la que se aprueba y publica el programa formativo de la especialidad de enfermería del trabajo, (2009).
- Orden SAS/1730/2010, de 17 de junio, por la que se aprueba y publica el programa formativo de la especialidad de enfermería pediátrica, (2010).
- Orden SAS/3225/2009, de 13 de noviembre, por la que se aprueba y publica el programa formativo de la especialidad de enfermería geriátrica, (2009).
- Orden SAS/1729/2010, de 17 de junio, por la que se aprueba y publica el programa formativo de la especialidad de enfermería familiar y comunitaria, (2010).
- Ordoñez, C. L. (2004). Pensar pedagógicamente desde el constructivismo. de las concepciones a las prácticas pedagógicas *Revista De Estudios Sociales, 19*, 7-12.
- Ormrod, J. E. (2003). *Educational psychology: Developing learners*. (4<sup>th</sup> ed.). Old Tappan, NJ: Prentice-Hall.
- Ortiz-Sagrasta, J. C. (2003). Uso de las técnicas de bioestadística en revistas de anestesiología. Facultad de Medicina. Universidad de Barcelona.
- Osterlind, S. J. (1989). *Constructing test items*. Boston: Kluwer Academic Publishers.
- Panadero, E., & Alonso-Tapia, J. (2014). ¿Cómo autorregulan nuestros alumnos? revisión del modelo cíclico de Zimmerman sobre autorregulación del aprendizaje. *Anales de Psicología, 30*(2), 450-462.
- Payne, E., & Brown, G. (2011). Communication and practice with examination criteria. Does this influence performance in examinations? *Assessment & Evaluation in Higher Education, 36*(6), 619-626.
- Pérez, O. (2006). ¿Cómo diseñar el sistema de evaluación del aprendizaje en la enseñanza de las matemáticas? *Revista Latinoamericana De Investigación En Matemática Educativa, 9*(2), 267-297.

- Perkins, D. (1992). *Smart schools*. New York: The Free Press.
- Perkins, D. (1995). *Outsmarting IQ: The emerging science of learnable intelligence*. New York: Free Press.
- Perkins, D. (1997). What is understanding? In M. S. Wiske (Ed.), *Teaching for understanding* (pp. 39-57). San Francisco: Jossey-Bass Publishers.
- Pfannkuch, M. (2007). Year 11 students' informal inferential reasoning: A case study about the interpretation of box plots. *International Electronic Journal of Mathematics Education*, 2(3), March 20, 2013. Retrieved from <http://www.iejme.com/032007/d3.pdf>
- Piaget, J. (1970). Carmichael's manual of child psychology (vol. 1). In P. H. Mussen (Ed.), *Piaget's theory* (3<sup>a</sup> ed., pp. 703-732). New York: Wiley.
- Piaget, J. (1970). *Genetic epistemology*. New York: W.W. Norton and Company.
- Piaget, J. (1998). *Success and understanding*. Cambridge, Mass: Harvard University Press.
- Pollatsek, A., Konold, C. E., Well, A. D., & Lima, S. D. (1984). Beliefs underlying random sampling. *Memory and Cognition*, 12(4), 395-401.
- Posner, M. A. (2011). The impact of a proficiency-based assessment and reassessment of learning outcomes system on student achievement and attitudes. *Statistics Education Research Journal*, 10(1), 3-14. June 05, 2014. Retrieved from [https://www.stat.auckland.ac.nz/~iase/serj/SERJ10%281%29\\_Posner.pdf](https://www.stat.auckland.ac.nz/~iase/serj/SERJ10%281%29_Posner.pdf)
- Pruzek, R., & Helmreich, E. (2009). Enhancing dependent sample analyses with graphics. *Journal of Statistics Education*, 17(1), May 14, 2013. Retrieved from <http://www.amstat.org/publications/jse/v17n1/helmreich.html>
- Rabina, L. A., & Nutter-Uphama, K. E. (2010). Introduction of a journal excerpt activity improves undergraduate students' performance in statistics. *College Teaching*, 58(4), 156-160.
- Ragasa, C. Y. (2008). A comparison of computer-assisted instruction and the traditional method of teaching basic statistics. *Journal of Statistics Education*, 16(1), 1-10.
- Ramirez, C., Candace, S., & Emmioğlu, E. (2012). The importance of attitudes in statistics education5. *Statistics Education Research Journal*, 11(2), 57-71.

- Reading, C., & Shaughnessy, C. (2004). Reasoning about variation. In D. Ben-Zvi, & J. B. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 201-226). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- REAL DECRETO 450/2005, de 22 de abril, sobre especialidades de enfermería, (2005).
- Reid, L., & Reading, C. (2008). Measuring the development of student's consideration of variation. *Statistics Education Research Journal*, 7(1), 40-59. June 05, 2014. Retrieved from [https://www.stat.auckland.ac.nz/~iase/serj/SERJ7\(1\)\\_Reid\\_Reading.pdf](https://www.stat.auckland.ac.nz/~iase/serj/SERJ7(1)_Reid_Reading.pdf)
- Reznick, R. K., Dawson-Saunders, E., & Folse, J. R. (1987). A rationale for the teaching of statistics to surgical residents. *Surgery*, 101, 611-617.
- Roberts, D. M., & Bilderback, E. W. (1980). Reliability and validity of a statistics attitude survey. *Educational and Psychological Measurement*, 40(1), 235-238.
- Rogers, C. R. (1967). *Becoming a person. A therapist's view of psychotherapy*. London: Constable.
- Roscoe, L. J., & Strapp, C. M. (2009). Increasing psychology students' satisfaction with preparedness through a professional issues course. *Teaching of Psychology*, 36(1), 18-23.
- Rosebery, A. S., & Rubin, A. (1990). Teaching statistical reasoning with computers. *Teaching Statistics*, 12(2), 38-42.
- Rossman, A. (2008). Reasoning about informal statistical inference: One statistician's view. *Statistics Education Research Journal*, 7(2), 5-19.
- Rubin, A., Bruce, B., & Tenney, Y. (1990). Learning about sampling: Trouble at the core. *Paper Presented at the Annual Meeting of the American Educational Research Association*, Boston, MA.
- Rumsey, D. J. (2002). Discussion: Statistical literacy: Implications for teaching, research, and practice. *International Statistical Review*, 70(1), 32-36.
- Saville, B. K., Zinn, T. E., Brown, A. R., & Marchuk, K. A. (2010). Syllabus detail and students' perceptions of teacher effectiveness. *Teaching of Psychology*, 37(3), 186-189.
- Schau, C., Stevens, J., Dauphinee, T. L., & Del Vecchio, A. (1995). The development and validation of the survey of attitudes toward statistics. *Educational and Psychological Measurement*, 55(5), 868-875.

- Schau, C., Stevens, J., Dauphinee, T. L., & Del Vecchio, A. (1995). The development and validation of the survey attitudes toward statistics. *Educational and Psychological Measurement, 55*(5), 868-875.
- Scheaffer, R. L., Watkins, A., Gnanadesikan, M., & Witmer, J. A. (1996). *Activity-based statistics: Instructor resources*. New York: Springer-Verlag.
- Schild, M. (1999). Statistical literacy: Thinking critically about statistics. *Of Significance, 1*(1), 15-20.
- Schild, M. (2010). Assessing statistical literacy: Take CARE. In P. Bidgood, N. Hunt & F. Jolliffe (Eds.), *Assessment methods in statistical education: An international perspective* (pp. 133-152) John Wiley & Sons, Ltd.
- Schneiter, K. (2008). Two applets for teaching hypothesis testing. *Journal of Statistics Education, 16*(3). March 21, 2012. Retrieved from <http://www.amstat.org/publications/jse/v16n3/schneiter.html>
- Schuetze, P. (2004). Evaluation of a brief homework assignment designed to reduce citation problems. *Teaching of Psychology, 31*(4), 257-262.
- Schulte, O. (2008). Reliable reasoning: Induction and statistical learning theory by HARMAN, G. and KULKARNI, S. *Biometrics, 64*(3), 992-993.
- Schwartz, T. A. (2014). Flipping the statistics classroom in nursing education. *Journal of Nursing Education, 53*(4), 199-206.
- Sebastiá, J. M. (1989). El constructivismo: Un marco teórico problemático. *Enseñanza De Las Ciencias, 7*(2), 158-161.
- Serrano-Gallardo, P., Jiménez-Maroto, A., & Arroyo Gordo, P. (2005). Análisis de la producción científica publicada en la revista metas de enfermería. *Index de Enfermería, 14*(48-49), 78-82.
- Shah, P. (2001). Improving statistical reasoning: Theoretical models and practical implications. *Applied Cognitive Psychology, 15*(6), 696-697.
- Shaughnessy, J. M. (1977). Misconceptions of probability: An experiment with a small-group, activity-based, model building approach to introductory probability at the college level. *Educational Studies in Mathematics, 8*(3), 295-316.
- Shaughnessy, J. M. (1992). Researches in probability and statistics: Reflections and directions. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 465-494). New York: Michigan Publishing Company.

- Shillam, C. R., Ho, G., & Commodore-Mensah, Y. (2014). Online biostatistics: Evidence-based curriculum for Master's nursing education. *Journal of Nursing Education*, 53(4), 229-232.
- Shulman, L. S. (1970). Psychology and mathematics education. In E. G. Begle (Ed.), *National society for the study of education* (pp. 23-71). Chicago: University of Chicago Press.
- Sireci, S. G. (1998). Gathering and analyzing content validity data. *Educational Assessment*, 5(4), 299-321.
- Slauson, L. V. (2008). *Students' conceptual understanding of variability*. Unpublished University of Ohio, EE.UU.
- Snee, R. D. (1993). What's missing in statistical education? *The American Statistician*, 47(2), 149-154.
- Sobrinho-Prieto, M., Sobrinho-Prieto, N., González-Gutián, C., Pichel-Guerrero, M. J., García-Sánchez, M. M., & Prieto-Díaz, A. (2005). Revistas españolas de enfermería en bases de datos nacionales e internacionales. *Index De Enfermería*, 14(48-49), 74-77.
- Solomon, J. (1994). The rise and fall of constructivism. *Studies in Science Education*, 23(1), 1-19.
- Sotos, A. E. C., Vanhoof, S., Van den Noortgate, W., & Onghena, P. (2007). Students' misconceptions of statistical inference: A review of the empirical evidence from research on statistics education. *Educational Research Review*, 2(2), 98-113.
- Sovak, M. M. (2010). *The effect of student-driven projects on the development of statistical reasoning*. Unpublished University of Pittsburgh.
- Stemberger, J. P., & MacWhinney, B. (1986). Form-oriented inflectional errors in language processing. *Cognitive Psychology*, 18(3), 329-354.
- Stickels, J. W., & Dobbs, R. R. (2007). Helping alleviate statistical anxiety with computer aided statistical classes. *The Journal of Scholarship of Teaching and Learning*, 7(1), 1-15.
- Stiggins, R. J. (1999). Assessment, student confidence, and school success. *Phi Delta Kappan*, 81(3), 191-198.

- Sundre, D. L. (2003). *Assessment of quantitative reasoning to enhance educational quality*. Paper presented at the *Meeting of the American Educational Research Association*, Chicago. June 05, 2014. Retrieved from <http://www.stat.auckland.ac.nz/~iase/cblumberg/sundreqpaper.doc>
- Sutarso, T. (1992). Students' attitudes toward statistics (STATS). *Paper presented at the annual meeting of the mid-south educational research association*. Knoxville, TN.
- Tauber Liliana, M. (2001). *La construcción del significado de la distribución normal a partir de análisis de datos*. Unpublished Universidad de Sevilla, *A taxonomy for learning, teaching, and assessing: A revision of bloom's taxonomy of educational objectives* (2001). In Anderson L. W., Krathwohl D. R. (Eds.). New York: Longman.
- Tempelaar, D. T., Gijsselaers, W. J., & van der Loeff, S. S. (2006). Puzzles in statistical reasoning. *Journal of Statistics Education*, 14(1). June 05, 2014. Retrieved from <http://www.amstat.org/publications/jse/v14n1/tempelaar.html>
- Tempelaar, D. T., Van Der Loeff, S. S., & Gijsselaers, W. H. (2007). A structural equation model analyzing the relationship of students' attitudes toward statistics, prior reasoning abilities and course performance. *Statistics Education Research Journal*, 6(2), 78-102.
- Torra-i-Bou, J. E. (1996). Estudio bibliométrico de los artículos publicados en tres revistas españolas durante los años 1991 al 1993. *Revista Rol de Enfermería*, 218, 33-40.
- Valle Arias, A., González Cabanach, R., Núñez Pérez, J. C., Suárez Riveiro, J. M., Piñeiro Aguin, I., & Rodríguez Martínez, S. (2000). Enfoques de aprendizaje en estudiantes universitarios. *Psicothema*, 12(3), 368-375.
- VanHoof, S., Kuppens, S., Sotos, A. E. C., Verschaffel, L., & Onghena, P. (2011). Measuring statistics attitudes: Structure of the survey of attitudes toward statistics (SATS-36). *Statistics Education Research Journal*, 10(1), 35-51.
- Vaughn, B. (2009). An empirical consideration of a balanced amalgamation of learning strategies in graduate introductory statistics classes. *Statistics Education Research Journal*, 8(1), 106-130. June 05, 2014. Retrieved from [http://www.stat.auckland.ac.nz/~iase/serj/SERJ8\(1\) Vaughn.pdf](http://www.stat.auckland.ac.nz/~iase/serj/SERJ8(1) Vaughn.pdf)
- Vygotsky, L. (1978). *Mind in society*. Cambridge, Mass: Harvard University Press.
- Vygotsky, L. S. (1988). *A formação social da mente* (2ª ed. brasileira). São Paulo: Martins Fontes.

- Vygotsky, L. S. (1987). *Pensamento e linguagem* (1º ed. brasileira). São Paulo: Martins Fontes.
- Wade, B., & Goodfellow, M. (2009). Confronting statistical literacy in the undergraduate social science curriculum. *Sociological Viewpoints*, , 75-90.
- Waldegg, G. (1998). Principios constructivistas para la educación matemática. *Revista EMA*, 4(1), 16-31.
- Wallman, K. K. (1993). Enhancing statistical literacy: Enriching our society. *Journal of the American Statistical Association*, 88(421), 1-8.
- Watson, J. M., & Ben, A. K. (2007). Assessment of students' understanding of variation. *Teaching Statistics*, 29(3), 80-88.
- Watson, J., & Callingham, R. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal*, 2(2), 3-46. June 05, 2014. Retrieved from [http://www.stat.auckland.ac.nz/~iase/serj/SERJ2\(2\)\\_Watson\\_Callingham.pdf](http://www.stat.auckland.ac.nz/~iase/serj/SERJ2(2)_Watson_Callingham.pdf)
- Weinberg, A., Wiesner, E., & Pfaff, T. J. (2010). Using informal inferential reasoning to develop formal concepts: Analyzing an activity. *Journal of Estatistics Education*, 18(2). March 20, 2012. Retrieved from <http://www.amstat.org/publications/jse/v18n2/weinberg.pdf>
- Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 223-265.
- Wilson, M. (2005). *Constructing measures: An item response modeling approach* . Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Winniger, S. R. (2005). Using your tests to teach: Formative summative assessment. *Teaching of Psychology*, 32(3), 164-166.
- Witmer, J., & et. al. (2005). *Guidelines for assessment and instruction in statistics education* (College Report. Alexandria, VA: American Statistical Association. October 16, 2007. Retrieved from [http://www.amstat.org/education/gaise/GaiseCollege\\_Full.pdf](http://www.amstat.org/education/gaise/GaiseCollege_Full.pdf)
- Yahong, H., Weihong, W., & Li, J. (2011). 6th international conference computer science & education (ICCSE). Teaching Discrete Mathematics with the Constructivism Learning Theory, Singapore.
- Zieffler, A. S. (2006). *A longitudinal investigation of the development of college student's reasoning about bivariate data during an introductory statistics course*. Unpublished University of Minnesota, EE.UU.

- Zieffler, A. S., Garfiel, J. B., Delmas, R. C., & Reading, C. (2008). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal*, 7(2), 40-58. June 06, 2014. Retrieved from [http://www.stat.auckland.ac.nz/~iase/serj/SERJ7\(2\)\\_Zieffler.pdf](http://www.stat.auckland.ac.nz/~iase/serj/SERJ7(2)_Zieffler.pdf)
- Zieffler, A. S., Garfield, J. B., Alt, S., Dupuis, D., Holleque, K., & Chang, B. (2008). What does research suggest about the teaching and learning of introductory statistics at the college level? A review of the literature. *Journal of Estatistics Education*, 16(2), March 19, 2012. Retrieved from <http://www.amstat.org/publications/jse/v16n2/zieffler.html>
- Zieffler, A., Garfield, J., Delmas, R., & Björnsdóttir, A. (2010). Development of an instrument to assess statistical thinking . Paper presented at the *Data and Context in Statistics Education: Towards an Evidence-Based Society. Proceedings of the Eighth International Conference on Teaching Statistics ICOTS8*, Ljubljana, Slovenia. March 19, 2012. Retrieved from [http://iase-web.org/documents/papers/icots8/ICOTS8\\_5F2\\_ZIEFFLER.pdf](http://iase-web.org/documents/papers/icots8/ICOTS8_5F2_ZIEFFLER.pdf)
- Zieffler, A., Park, J., Garfield, J., Delmas, R., & Audbjorg, B. (2012). The statistics teaching inventory: A survey on statistics teachers' classroom practices and beliefs. *Journal of Statistics Education*, 20(1). March 19, 2012. Retrieved from <http://www.amstat.org/publications/jse/v20n1/zieffler.pdf>
- Zieffler, A. S., & Garfield, J. B. (2009). Modeling the growth of students' covariational reasoning during an introductory statistics course. *Statistics Education Research Journal*, 8(1), 7-31. March 19, 2012. Retrieved from [http://www.stat.auckland.ac.nz/~iase/serj/SERJ8\(1\)\\_Zieffler\\_Garfield.pdf](http://www.stat.auckland.ac.nz/~iase/serj/SERJ8(1)_Zieffler_Garfield.pdf)
- Zimmerman, B. B., Sudweeks, R. R., Shelley, M. F., & Wood, B. (1990). *How to prepare better tests: Guidelines for university faculty*. Provo, Utah: Brigham Young University Testing Services and The department for instructional science. March 19, 2012. Retrieved from <http://testing.byu.edu/info/handbooks/betteritems.pdf>



**ANEXOS**

---



## ANEXO 1. CUESTIONARIO DE EVALUACIÓN DE RAZONAMIENTO ESTADÍSTICO (CERES)

1. La siguiente tabla de frecuencias recoge los datos de tensión arterial sistólica (mmHg) de los 300 trabajadores de una empresa.

Tabla de Frecuencias para TA Sistólica \_mmHg\_

Clase	Límite Inferior	Límite Superior	Marca	Frecuencia	Frecuencia Relativa	Frecuencia Acumulativa	Frecuencia Acum.Rel.
menor o igual		60,0		0	0,0000	0	0,0000
1	60,0	75,0	67,5	2	0,0067	2	0,0067
2	75,0	90,0	82,5	30	0,1000	32	0,1067
3	90,0	105,0	97,5	84	0,2800	116	0,3867
4	105,0	120,0	112,5	110	0,3667	226	0,7533
5	120,0	135,0	127,5	59	0,1967	285	0,9500
6	135,0	150,0	142,5	11	0,0367	296	0,9867
7	150,0	165,0	157,5	2	0,0067	298	0,9933
8	165,0	180,0	172,5	2	0,0067	300	1,0000
9	180,0	195,0	187,5	0	0,0000	300	1,0000
10	195,0	210,0	202,5	0	0,0000	300	1,0000
mayor	210,0			0	0,0000	300	1,0000

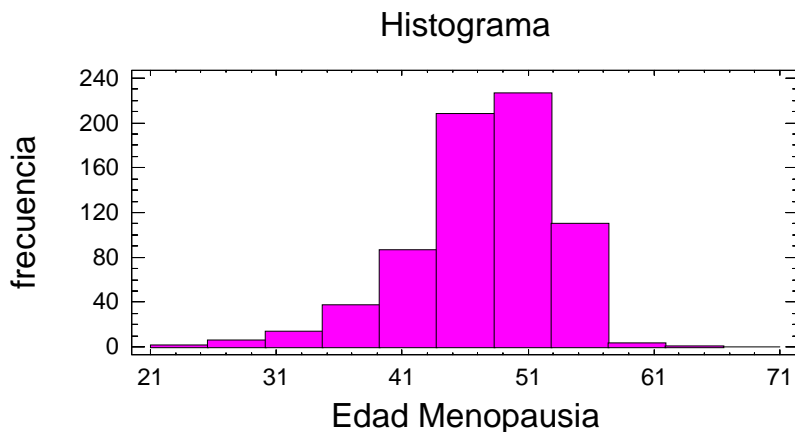
De las siguientes afirmaciones acerca de la información que nos proporciona esta tabla ¿cuál es cierta?

- La tensión arterial sistólica más frecuente en este grupo de trabajadores es 110 mmHg.
- Hay 226 trabajadores que tienen una tensión arterial sistólica inferior a 120 mmHg.**
- Hay 2 trabajadores que tienen una tensión arterial sistólica de 172,5 mmHg.
- El 95% de los trabajadores tienen unos valores de tensión arterial entre 120 y 135 mmHg.

### Resultado de aprendizaje 1:

Interpretar el resumen de datos recogido en una tabla de frecuencias.

2. El siguiente gráfico muestra la distribución de la edad en la que se produjo la menopausia en un grupo de 1000 mujeres de la consulta de ginecología de un Hospital.



Seleccione la opción que da la descripción más completa de la gráfica, de manera que demuestre una comprensión de cómo describir e interpretar estadísticamente la distribución de una variable.

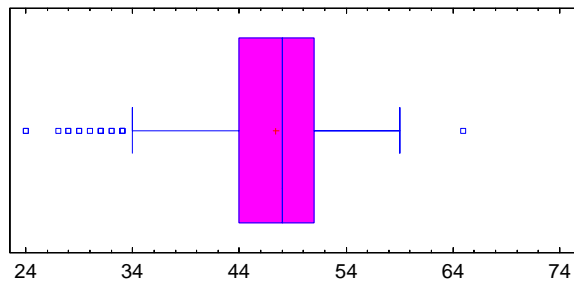
- Las barras 21 a 51 van aumentando en altura hasta la 51 que es la más alta, y luego disminuyen hasta la 70.
- La distribución es normal, con una media alrededor de 51 y una desviación típica alrededor de 10.
- La mayor parte de las mujeres tienen la menopausia alrededor de los 50 años, aunque algunas dejan de tener su ciclo menstrual muy pronto y otras lo mantienen hasta edad muy avanzada.
- La distribución de la edad en la que apareció la menopausia es algo simétrica aunque tiene un sesgo negativo. La edad habitual de la menopausia es aproximadamente 47 años y el rango es 50.**

**Resultado de aprendizaje 2:**

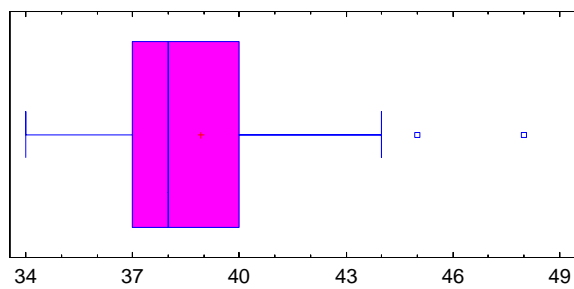
**Describir e interpretar en el contexto de los datos, la distribución de una variable representada en un histograma.**

3. Cuál de los siguientes diagramas de cajas, representan los datos del gráfico de la pregunta 2.

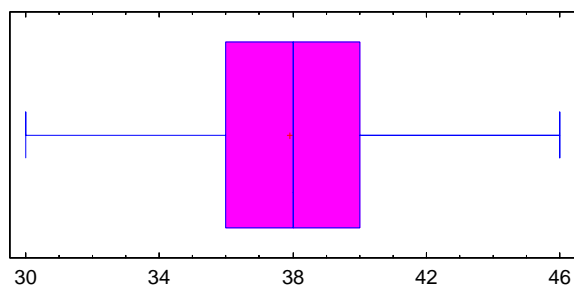
A.



B.



C.



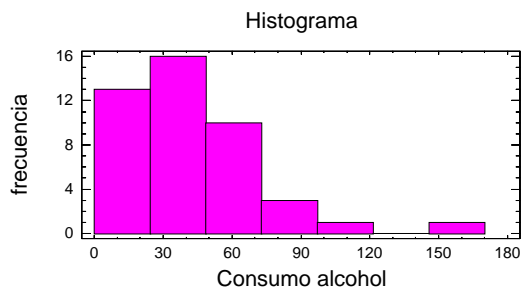
- a. Diagrama A
- b. Diagrama B
- c. Diagrama C

**Resultado de aprendizaje 3:**

Reconocer dos formas de representar gráficamente el mismo conjunto de datos.

Los ítems 4 y 5 se refieren a la siguiente situación:

El siguiente gráfico representa la distribución del consumo semanal de alcohol (en gramos) de una muestra aleatoria de estudiantes universitarios.



4. ¿Qué valores para la media y la mediana parecen más acertados?

- media= 40,8 y mediana= 52,3
- media= 31,4 y mediana= 38,5
- media= 38,5 y mediana= 31,4**
- media= 32,5 y mediana= 32,5

**Resultado de aprendizaje 4:**

Reconocer en un histograma la asimetría y el sesgo de una distribución de datos y su influencia sobre los valores de las medidas de tendencia central.

5. ¿Cuál es la mejor medida de la variabilidad de este conjunto de datos?

- El rango, porque te informa de la extensión total de los datos.
- La desviación estándar, porque está basada en la información de todo el conjunto de datos.
- La desviación estándar porque es la medida de variabilidad más utilizada.
- El rango intercuartílico, porque es resistente a los valores extremos.**

**Resultado de aprendizaje 5:**

Reconocer en un histograma con puntuaciones atípicas la conveniencia de utilizar el rango intercuartílico como medida de variabilidad del conjunto de datos representado, frente a la desviación estándar y el rango, parámetros afectados por los valores extremos.

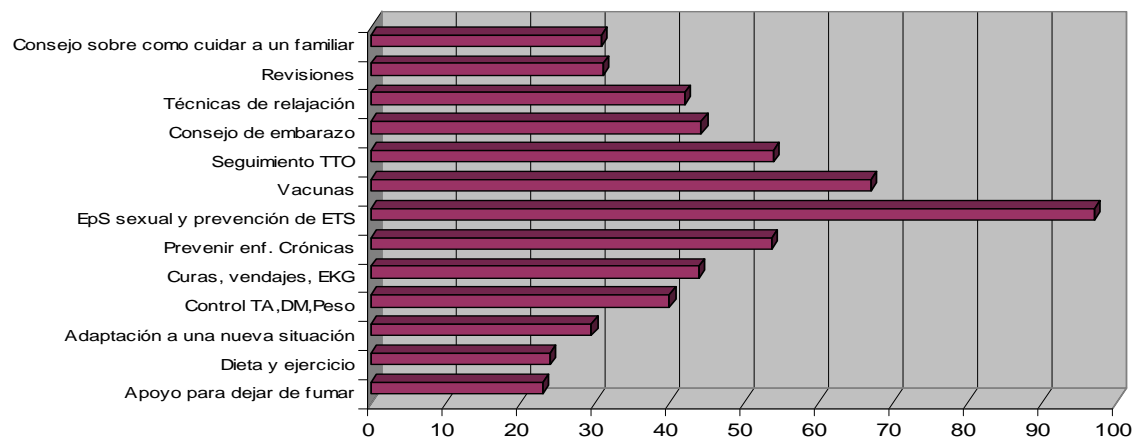
6. Un Centro de Salud suele publicar periódicamente información estadística sobre la actividad que realiza. La siguiente tabla muestra la proporción con la que los usuarios del centro de salud utilizan los distintos servicios de la consulta de enfermería.

<i>Servicios de Enfermería</i>	Consejo sobre cómo cuidar a un familiar	Revisiones	Técnicas de relajación	Consejo de embarazo	Seguimiento del tratamiento	Vacunas	Educación sexual y prevención de Enfermedades de Transmisión sexual (ETS)
<i>Proporción asistencia</i>	13,92	60,67	7,17	15,18	43,45	80,59	10,12

Prevenir enfermedades Crónicas	Curas, vendajes Electrocardiograma	Control tensión arterial, diabetes y Peso	Adaptación a una nueva situación	Dieta y ejercicio	Apoyo para dejar de fumar
20,67	61,18	53,16	8,01	24,89	2,53

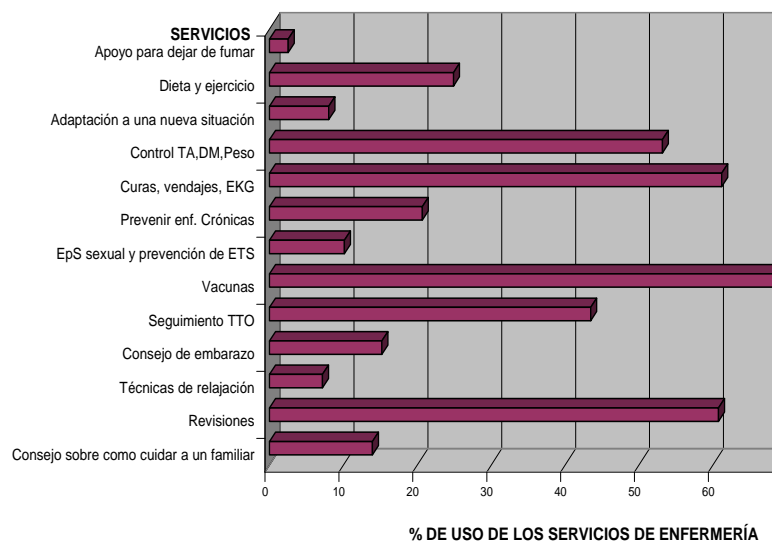
¿Cuál de los siguientes gráficos ofrece la mejor visualización de la distribución de la proporción del uso de los servicios de enfermería y permite al usuario describir la forma, el centro y la dispersión de la variable “proporción de asistencia a los distintos servicios de enfermería”?

A.

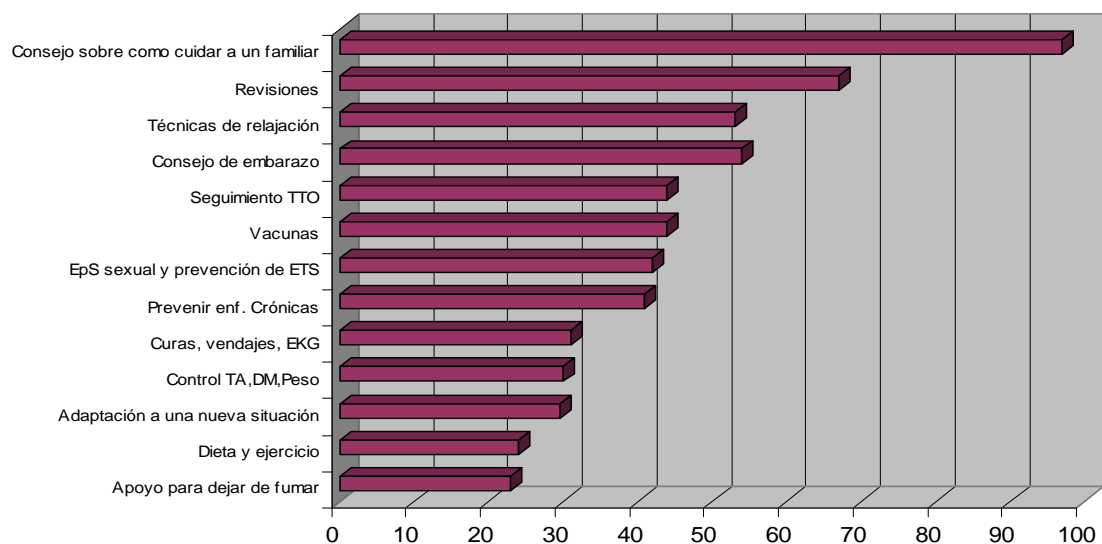




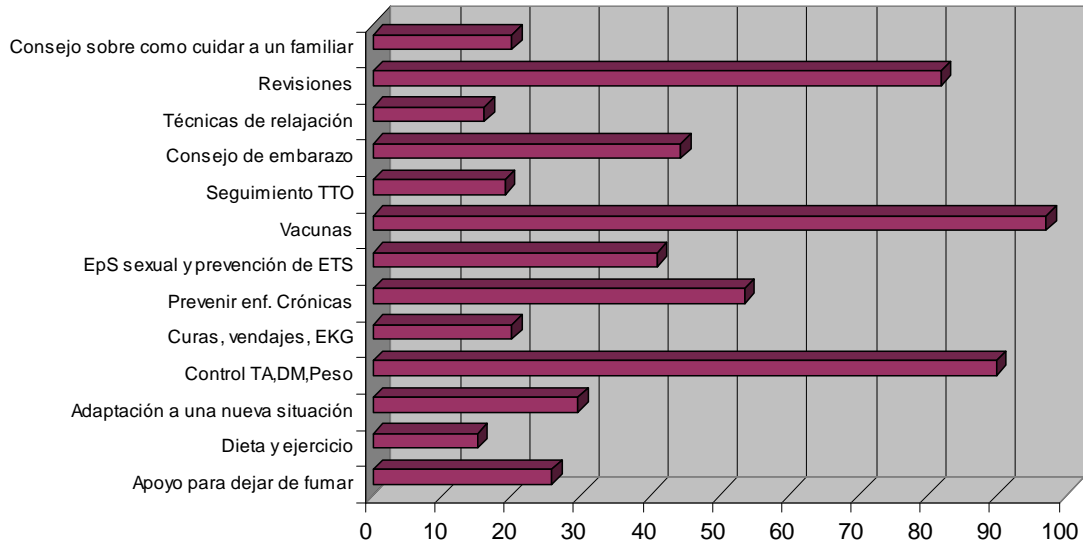
B.



C.



D.

**Resultado de aprendizaje 6:**

Entender la necesidad de representar gráficamente las frecuencias para describir adecuadamente la distribución de una variable cuantitativa. Asociar el gráfico a la información que la tabla de frecuencias nos ofrece respecto al centro de la variable y su dispersión.

7. La madre de un niño de cuatro años que ha sido pesado y medido en la consulta de enfermería, pregunta a un estudiante de enfermería que realiza allí sus prácticas. ¿Qué significa que mi hijo esté en el percentil 65% de peso?

¿Tú qué respondes a esa pregunta?

- Significa que la media de peso del 65% de los niños de cuatro años coincide con el peso de su hijo.
- Significa que el 65% de los niños de cuatro años pesan menos que su hijo.**
- Significa que el 65% de los niños de cuatro años pesan más que su hijo.
- Significa que el peso de su hijo varía en un 65% respecto al resto de los niños de cuatro años.

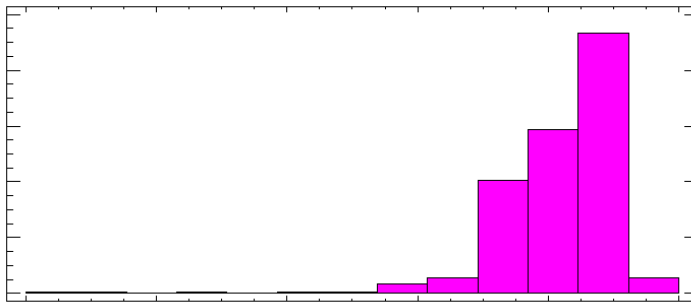
**Resultado de aprendizaje 7:**  
**Comprender el concepto de percentil y su interpretación en un contexto específico.**

Los ítems 8 al 11 se refieren a la siguiente situación:

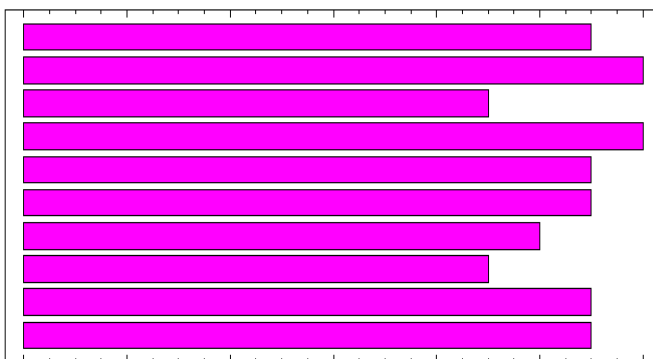
Se trata de un estudio neonatológico con datos de 2140 madres y de sus recién nacidos.

Los ítems 8, 9, 10 y 11 describen una situación que coincide con una de las siguientes representaciones gráficas.

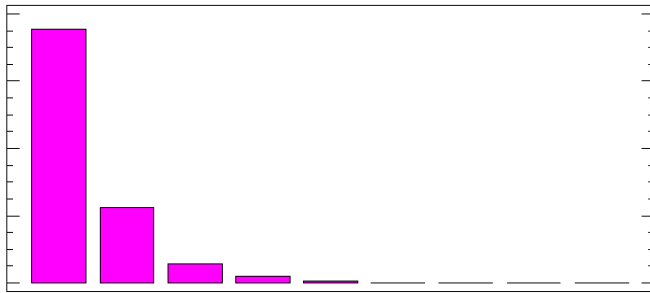
A.



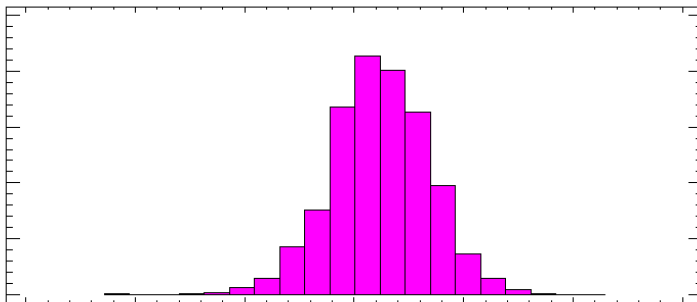
B.



C.



D.



8. Representa la distribución del peso en gramos de los recién nacidos.

- a. Gráfico A.
- b. Gráfico B.
- c. Gráfico C.
- d. **Gráfico D.**

**Resultado de aprendizaje 8:**

Asociar un histograma a la descripción de una variable. En este caso es una variable que se distribuye de forma normal. Reconocer este tipo de variables y la distribución de datos con un gráfico simétrico, campana de Gauss.

9. Representa el número de abortos previos.

- a. Gráfico A.
- b. Gráfico B.
- c. **Gráfico C.**
- d. Gráfico D.

**Resultado de aprendizaje 9:**

Asociar las distintas representaciones gráficas con el tipo de variable representada y asociar el comportamiento de la gráfica a la variable concreta. En este caso el número de abortos previos, variable discreta que aparece representada en un diagrama de barras, con una mayor frecuencia en la barra correspondiente a 0 y 1 abortos.

10. Representa la frecuencia de la última cifra de un grupo de historias clínicas seleccionadas al azar.

- a. Gráfico A.
- b. Gráfico B.**
- c. Gráfico C.
- d. Gráfico D.

**Resultado de aprendizaje 10:**

Reconocer el comportamiento y la distribución de frecuencias de una variable aleatoria, con igual probabilidad en todos los posibles resultados. En este caso el último número de la historia clínica.

11. Representa las semanas de gestación.

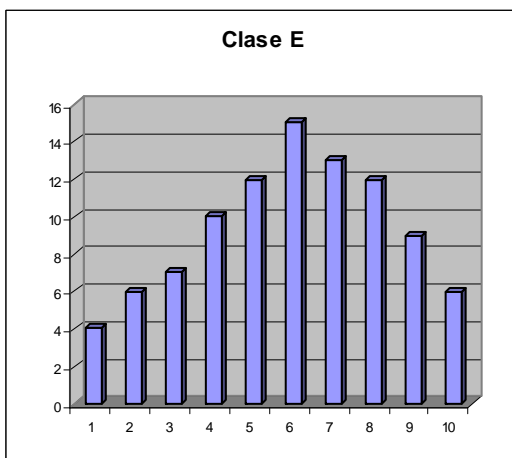
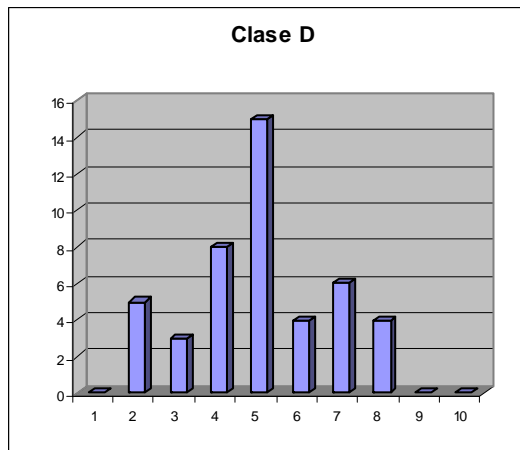
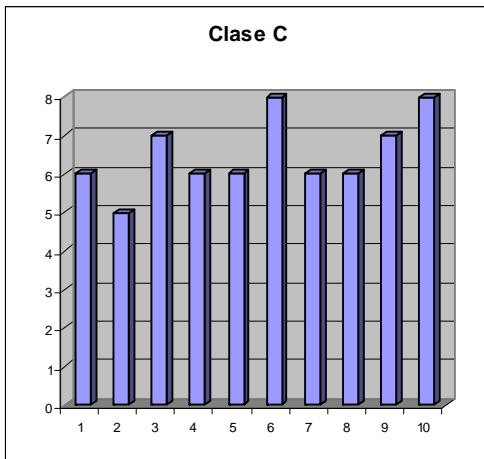
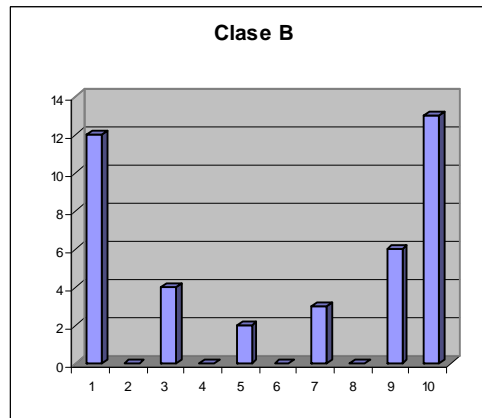
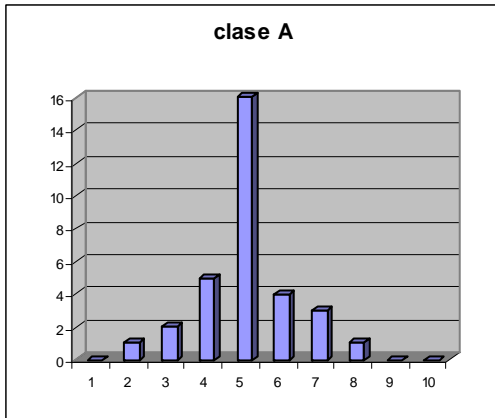
- a. Gráfico A.**
- b. Gráfico B.
- c. Gráfico C.
- d. Gráfico D.

**Resultado de aprendizaje 11:**

Reconocer la distribución de frecuencias de una variable sesgada negativamente y asociarla al comportamiento real de una variable. En este caso las semanas de gestación.

Los ítems 12 y 13 se refieren a la siguiente situación:

Se presentan a continuación cinco gráficos, cada uno de ellos muestra el resultado en una escala de 0 a 10 de los exámenes de cinco clases de estadística.



12. ¿Cuál de las clases se espera que tenga una menor desviación típica, y por qué?

- a. **La clase A ya que tiene la mayoría de los valores cercanos a la media.**
- b. La clase B porque tiene un menor número de resultados distintos.
- c. La clase C porque no hay variaciones en las puntuaciones.
- d. Las Clases A y D porque las dos tienen el menor rango.
- e. La clase E porque parece que es el que más se ajusta a la normalidad.

**Resultado de aprendizaje 12:**

**Estimar correctamente y comparar las desviaciones estándar para diferentes gráficos de barras. Entender que la menor desviación estándar se asocia, por lo general, a un gráfico con una menor distancia de los datos al valor central de la distribución.**

13. ¿Cuál de las clases se espera que tenga una mayor desviación típica, y por qué?

- a. La clase A, ya que cuenta con la mayor diferencia entre las alturas de las barras.
- b. **La clase B, ya que varios de sus resultados están lejos de la media.**
- c. La clase C, porque tiene el mayor número de puntuaciones diferentes.
- d. La clase D, porque la distribución es muy desigual e irregular.
- e. La clase E porque tiene mayor rango y se ajusta a la normalidad.

**Resultado de aprendizaje 13:**

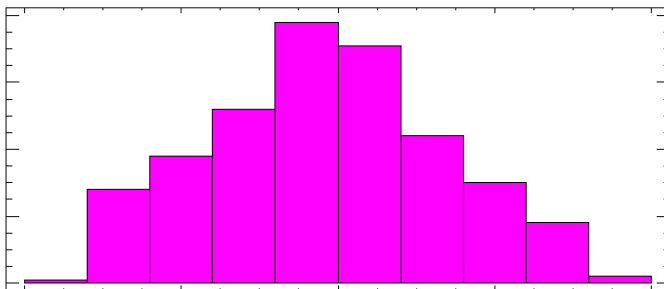
**Estimar correctamente y comparar las desviaciones estándar para diferentes gráficos de barras. Entender que la mayor desviación estándar se asocia por lo general a un gráfico con una mayor distancia de los datos al valor central de la distribución.**

14. En el análisis de la concentración de ozono troposférico de una determinada zona, se miden los valores de ozono diarios en un periodo de tres meses. El análisis estadístico presenta los siguiente parámetros.

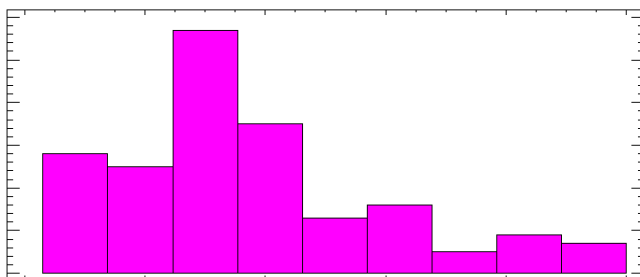
Media	Mediana	Desviación típica	Valor mínimo	Valor máximo
74.98 $\mu\text{g}/\text{m}^3$	66.00 $\mu\text{g}/\text{m}^3$	45.23 $\mu\text{g}/\text{m}^3$	6.0 $\mu\text{g}/\text{m}^3$	200.0 $\mu\text{g}/\text{m}^3$

¿Cuál de los siguientes histogramas es más probable que sea el que representa estos datos?

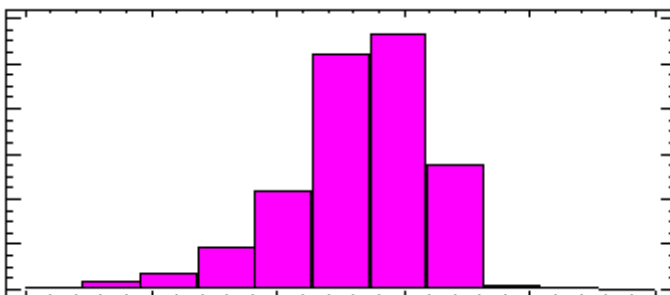
A.



B.



C.





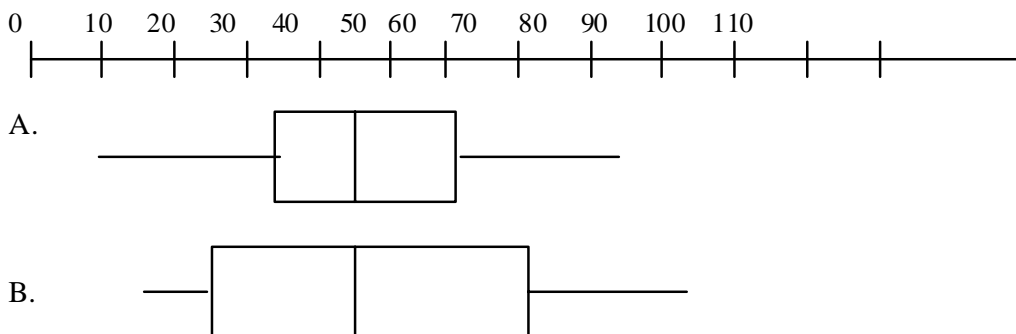
- a) El histograma A
- b) El histograma B**
- c) El histograma C

**Resultado de aprendizaje 14:**

Entender que una distribución de datos con una media mayor que la mediana es más probable que tenga un sesgo a la derecha.

Los ítems del 15 al 17 se refieren a la siguiente situación:

Los dos siguientes diagramas de cajas muestran los resultados de un estudio sobre la amnesia postraumática ocurrida tras una lesión en la cabeza. La variable estudiada es el tiempo de hospitalización en días. Se recogen en el siguiente gráfico los datos de dos grupos que se diferencian en el tipo de lesión que sufrieron en la cabeza.



15. ¿Qué grupo se espera que tenga una mayor desviación típica respecto al tiempo de hospitalización?

- a. Grupo A.
- b. Grupo B.**
- c. Ambos grupos son casi iguales.
- d. Es imposible saberlo.

**Resultado de aprendizaje 15:**

Entender el concepto de desviación típica como un parámetro que informa de la dispersión, así como la capacidad de interpretarlo en el contexto de un diagrama de cajas y patillas.

16. ¿Qué grupo tiene un mayor porcentaje de pacientes con un tiempo de hospitalización mayor a 20 días?

- a. Grupo A.
- b. Grupo B.
- c. Ambos grupos son casi iguales.
- d. **Es imposible saberlo.**

**Resultado de aprendizaje 16:**

Entender que los diagramas de caja no proporcionan estimaciones precisas de los porcentajes de los datos por encima o por debajo de los valores a excepción de los cuartiles.

17. ¿Qué grupo tiene un mayor porcentaje de pacientes con un tiempo de hospitalización superior a 50 días?

- a. Sección A.
- b. Sección B.
- c. **Ambos grupos son casi iguales.**

**Resultado de aprendizaje 17:**

Comprender el concepto de mediana y su interpretación en el contexto de un gráfico de cajas y patillas.

18. Se quiere realizar un estudio sobre el estado de dependencia de la población anciana que vive actualmente en las residencias de un municipio. Se realiza un muestreo entre los aproximadamente 7000 residentes de los Centros de Mayores.

¿Cuál de las siguientes situaciones NO afecta a la capacidad de generalizar los resultados de la encuesta a todos los ancianos que viven en residencias?

- a. **Se utilizó para el estudio una muestra aleatoria de 600 ancianos de los 7000 que viven actualmente en Centros de Mayores.**
- b. Los datos se recogieron sólo de residencias públicas.
- c. De los 600 ancianos incluidos en la muestra, sólo se pudo obtener datos acerca del grado de dependencia de 150 de ellos.
- d. Todo lo anterior presenta problemas para la generalización de los resultados.

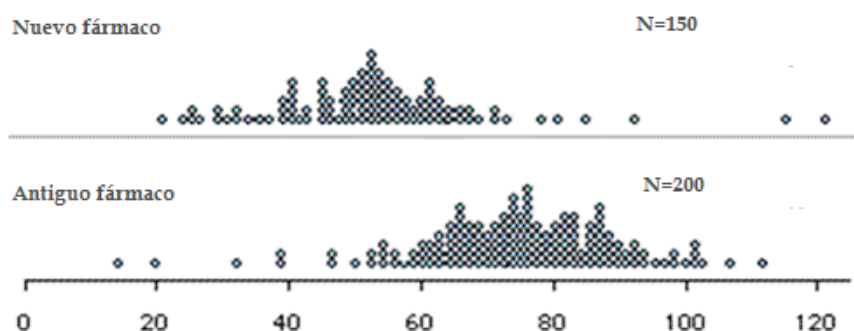
**Resultado de aprendizaje 18:**

Comprender los factores que permiten que los resultados obtenidos sobre una muestra sean generalizados a la población.

Los ítems del 19 al 21 se refieren a la siguiente situación:

Una compañía farmacéutica desarrolló unas mejoras sobre un medicamento para la jaqueca. Para probar su eficacia se realizó un ensayo clínico para el que se seleccionaron al azar 350 personas de una población grande de pacientes con dolores de cabeza. A 150 de ellos se le asignó al azar el nuevo fármaco, y los otros 200 recibieron el medicamento con la fórmula anterior. Se registró el tiempo en minutos, que transcurrió hasta desaparecer el dolor de cabeza. Los resultados se muestran a continuación.

Los ítems 19, 20, 21 presentan tres afirmaciones sobre las que debe indicar su opinión acerca de la validez o no de la misma.



19. El medicamento antiguo funciona mejor. Dos personas que lo tomaron sintieron alivio en menos de 15 minutos. Además, el peor resultado, cerca de 120 minutos, fue con el nuevo medicamento.

- a. Válida.
- b. No válida.

**Resultado de aprendizaje 19:**

Comparar grupos, teniendo en cuenta en las distribuciones, dónde se centran la mayoría de los datos al margen del comportamiento de algunos individuos de la muestra.

20. El tiempo medio para aliviar el dolor de cabeza con el nuevo medicamento es menor que el tiempo medio con el medicamento antiguo. La conclusión es que la gente que toma la fórmula nueva tiende a sentir alivio unos 25 minutos antes que los que toman la antigua fórmula.

- a. **Válida.**
- b. No válida.

**Resultado de aprendizaje 20:**

**Comparar grupos, mediante la comparación de los promedios, extrayendo esta información de la representación de la distribución de datos.**

21. No se puede concluir nada con estos datos. El número de pacientes en los dos grupos no es el mismo, por lo que no se pueden comparar los dos medicamentos.

- a. Válida.
- b. No válida.**

**Resultado de aprendizaje 21:**

**Entender que la comparación de dos grupos no requiere que las muestras tengan igual tamaño, especialmente si los dos conjuntos de datos son grandes.**

22. Una empresa fabrica mascarillas de polipropileno para protección sanitaria en cinco colores de las que el 50% son de color blanco. Se comercializan en cajas de 50 y en cajas de 500 que se rellenan al azar.

¿Qué caja es más probable que contenga más del 70% de mascarillas de color blanco?

- a. La de 500 porque contiene más mascarillas, por lo que puede tener más de color blanco.
- b. La de 500, porque hay mayor variabilidad en la proporción de mascarillas blancas en muestras grandes.
- c. La de 50, ya que hay mayor variabilidad en la proporción de mascarillas blancas en muestras pequeñas.**
- d. La de 50, porque la mayor parte de las cajas pequeñas tienen más del 50% de mascarillas blancas.
- e. Las dos tienen la misma probabilidad porque son muestras al azar.

**Resultado de aprendizaje 22:**

**Entender que las muestras pequeñas tienen mayor variabilidad que las muestras grandes.**

23. Irene vive a unos 10 kilómetros de la universidad donde va a cursar sus estudios. Hay dos rutas principales para llegar a la Universidad, una por la ciudad que es más corta en kilómetros pero tiene más semáforos, y otra por la circunvalación, más larga pero con menos semáforos. Irene realiza un experimento aleatorio en el que cada día lanza una moneda para decidir qué camino tomar ese día y registra los siguientes datos relativos a los tiempos de cinco días de viajes en cada ruta.

Ruta por la circunvalación      17, 15, 17, 16, 18

Ruta de la ciudad                      18, 13, 20, 10, 16

Irene quiere ser puntual a sus clases, pero no quiere llegar demasiado pronto, ya que aumentaría la tarifa del estacionamiento. Basándose en los datos recogidos, ¿Qué ruta le aconsejas que elija?

- a. **La circunvalación, porque los tiempos son siempre entre 15 y 18 minutos.**
- b. La ruta de la ciudad, porque puede llegar en 10 minutos si se le da bien, y el tiempo medio es menor que el de la circunvalación.
- c. Debido a que los tiempos de las dos rutas se solapan mucho, no es una ruta mejor que la otra. Podría elegir el camino lanzando una moneda.

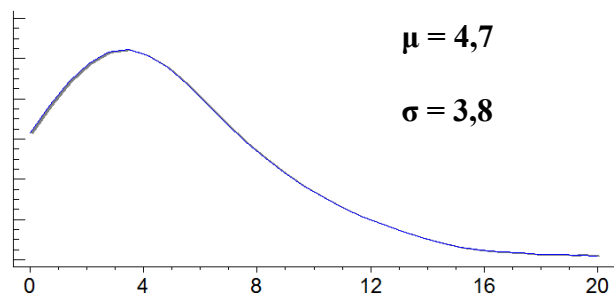
**Resultado de aprendizaje 23:**

**Comprender el significado de la variabilidad en un conjunto de datos obtenido como resultado de medidas repetidas y en un contexto en el que se desea que esa variabilidad sea pequeña.**

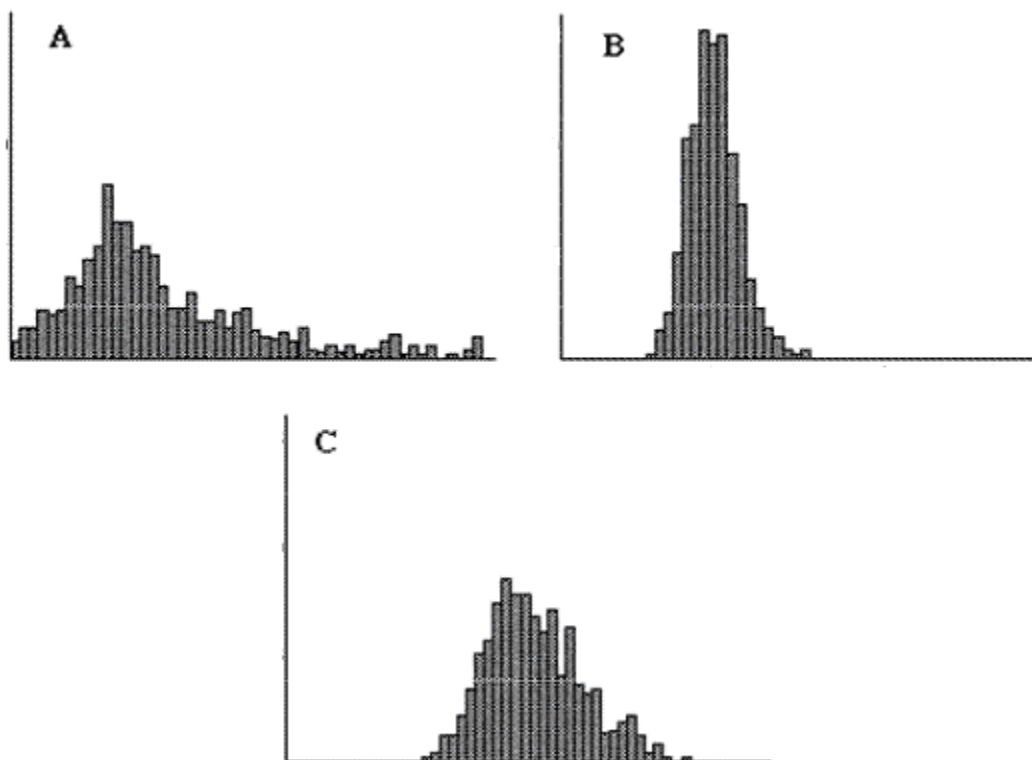
**Los ítems 24 y 25 se refieren a la siguiente situación:**

El siguiente gráfico representa la distribución de las puntuaciones de las encuestas de satisfacción de todos los usuarios de un sistema sanitario público. La puntuación media es de 4,7 y la desviación estándar es de 3,8.

Población



A continuación se muestran tres gráficos



24. ¿Qué gráfico crees que representa una sola muestra aleatoria de las puntuaciones de 500 usuarios de esa población?

- Gráfico A.
- Gráfico B.
- Gráfico C.

**Resultado de aprendizaje 24:**

Reconocer la representatividad de la población de una muestra grande aleatoria.

25. ¿Qué gráfico crees que representa la distribución de las medias de 500 muestras de esa población tomadas al azar, cada una de ellas de tamaño 10?
- Gráfico A
  - Gráfico B**
  - Gráfico C

**Resultado de aprendizaje 25:**

Comprender que la distribución de las medias de un conjunto de muestras extraídas aleatoriamente de una población, tiene una distribución normal de media próxima a la media poblacional.

**Los ítems 26 al 28 se refieren a la siguiente situación**

Un artículo de investigación ha publicado los resultados de un estudio sobre los efectos de una combinación de medicamentos contra el VIH, para combatir la infección por el virus del papiloma humano (HPV) en mujeres VIH-positivas. El artículo recoge un p-valor de 0,03 en el análisis estadístico. Los ítems 25, 26 y 27 presentan tres diferentes interpretaciones de este p-valor. Indique en cada una de ellas si es o no válida.

26. Es la probabilidad de que ocurran los resultados obtenidos si la hipótesis nula es verdadera.
- Válido.**
  - No válido.

**Resultado de aprendizaje 26:**

Reconocer una interpretación correcta de un p-valor.

27. Es la probabilidad de que el tratamiento no sea eficaz.
- Válido.
  - No válido.**

**Resultado de aprendizaje 27:**

Reconocer una interpretación incorrecta de un p-valor.

28. Es la probabilidad de que el tratamiento sea eficaz

- a. Válido.
- b. **No válido.**

**Resultado de aprendizaje 28:**  
**Reconocer una interpretación errónea de un p-valor.**

29. Un alumno de posgrado diseña un estudio de investigación en el que espera poder demostrar que los resultados de un experimento son estadísticamente significativos con un 99% de confianza en los resultados. ¿Cuál de los siguientes valores de contraste querría obtener?

- a. Un p-valor de 0.03.
- b. Un p-valor de 1.
- c. **Un p-valor de 0.003.**
- d. La magnitud de un p-valor no tiene ningún efecto en la significación estadística.

**Resultado de aprendizaje 29:**  
**Comprender el concepto de nivel de significación y la capacidad de asociarlo a la interpretación del p-valor, conociendo qué valores de p-valor son deseables en los estudios de investigación. Manejar como iguales dos formas de expresar el mismo concepto: valor de contraste y p-valor.**

**Los ítems 30 y 31 se refieren a la siguiente situación:**

Para un estudio del efecto del tratamiento con melatonina sobre la concentración de la lipoproteína HDL, se cuenta con 80 mujeres sanas entre 40 y 50 años, asignadas aleatoriamente al grupo de intervención (administración de 1mg/día de melatonina) y al grupo control. Las mujeres del grupo de tratamiento mostraron unos niveles mayores de HDL que las del grupo control.

30. Supongamos que se llevó a cabo correctamente una prueba de significación y no mostró diferencia estadísticamente significativa entre la media de los niveles de HDL del grupo experimental y del grupo control. ¿Qué conclusión puede extraerse de estos resultados?



- a. El investigador no ha interpretado los resultados correctamente, debe haber una diferencia significativa.
- b. Puede que el tamaño de la muestra sea demasiado pequeño para detectar diferencias estadísticamente significativas.**
- c. Debe ser cierto que el tratamiento con melatonina no produce niveles más altos de la lipoproteína HDL.

**Resultado de aprendizaje 30:**

**Entender que aunque un análisis de significación no ponga de manifiesto diferencias significativas entre dos grupos de comparación, puede que dichas diferencias existan y no hayan podido ser probadas con las muestras disponibles.**

31. Supongamos que se realiza correctamente una prueba de significación y muestra diferencias estadísticamente significativas entre las medias del nivel de HDL de las mujeres del grupo de tratamiento y las del grupo control. ¿Qué conclusión puede extraerse de estos resultados?

- a. Hay evidencia de que existe asociación, pero no de un efecto causal de la melatonina sobre los niveles de la lipoproteína HDL.
- b. El tamaño de la muestra es demasiado pequeño para sacar una conclusión válida.
- c. Se ha demostrado que la melatonina produce niveles más altos de la lipoproteína HDL.
- d. Existe evidencia de que la melatonina produce niveles más altos de la lipoproteína HDL en estas mujeres.**

**Resultado de aprendizaje 31:**

**Entender que un diseño experimental con asignación aleatoria apoya la inferencia causal.**

32. En una epidemia se sospecha que el número de afectados jóvenes es menor que el de adultos. La hipótesis nula es que no hay diferencia entre el número de afectados jóvenes frente a los adultos. Si se realiza una prueba estadística que rechaza la hipótesis nula, ¿Cuál de las siguientes afirmaciones es verdadera?

- a. Definitivamente la enfermedad afecta más a los adultos.
- b. Se acepta que la enfermedad afecta más a los adultos pero puede que no sea así.**
- c. Definitivamente la enfermedad afecta por igual a los jóvenes y a los adultos.
- d. Lo más probable que es afecte por igual a los jóvenes y a los adultos pero podría ser que fueran los adultos los más afectados.

**Resultado de aprendizaje 32:**

**Interpretar los resultados de una prueba de significación cuando es rechazada la hipótesis nula.**

**Los ítem 33 a 36 se refieren a la siguiente situación:**

Se quiere conocer el contenido medio de nitrito potásico (mg/kg) utilizado como conservante en una marca de embutido. Se recoge una muestra aleatoria de barras de este embutido y se analiza el contenido de nitrito en cada una de las barras, calculando el intervalo de confianza al 95% para la media de nitritos por barra en (32,5 a 35,7). Los ítems 33, 34, 35 y 36 presentan cuatro interpretaciones diferentes de estos resultados. Indique si cada interpretación es válida o no.

33. Estamos seguros en un 95%, que cada barra de embutido de esta marca tiene aproximadamente de 32.5 a 35,7 mg/kg de nitrito potásico

- a. Válido.
- b. No válido.**

**Resultado de aprendizaje 33:**

**Detectar una incorrecta interpretación del intervalo de confianza, como los valores entre los que se encuentra la variable en cada individuo o elemento de la población, con una seguridad del 95%.**

34. Esperamos que el 95% de las barras de embutido de esta marca tengan de 32.5 a 35,7 mg/kg de nitrito potásico.

- a. Válido.
- b. No válido.**

**Resultado de aprendizaje 34:**

Detectar una incorrecta interpretación del nivel de confianza, como el porcentaje de elementos de la población cuyo contenido en nitrito potásico se encuentra dentro de los límites del intervalo de confianza.

35. Cabe esperar que aproximadamente el 95% de las posibles muestras de estos embutidos, tengan una media de nitritos entre 32.5 a 35,7 mg/kg.

- a. Válido.
- b. **No válido.**

**Resultado de aprendizaje 35:**

Detectar una incorrecta interpretación del nivel de confianza como el porcentaje de todas las posibles muestras que tienen una media dentro de los límites del intervalo de confianza.

36. Estamos seguros al 95% que el intervalo de confianza de 32.5 a 35,7 mg/kg, incluye la verdadera media del contenido de nitritos de esta marca de embutidos.

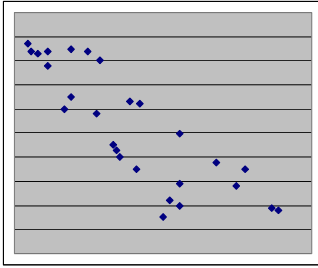
- a. **Válido.**
- b. No válido.

**Resultado de aprendizaje 36:**

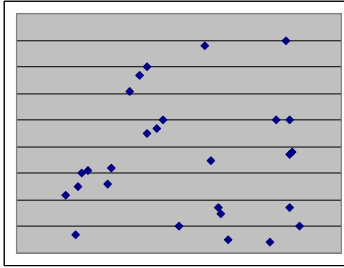
Interpretar correctamente un intervalo de confianza. Seguridad de que el intervalo de confianza contenga la verdadera media de la población.

37. Ciertos productos de nutrición enteral almacenados a bajas temperaturas durante largos periodos de tiempo sufren una pérdida de peso en el contenido total del producto. ¿Cuál de los siguientes gráficos muestra esta situación?

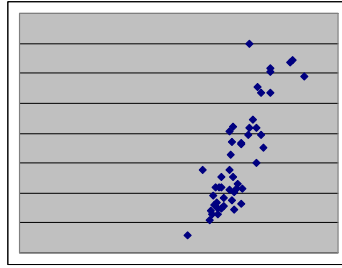
A.



B.



C.

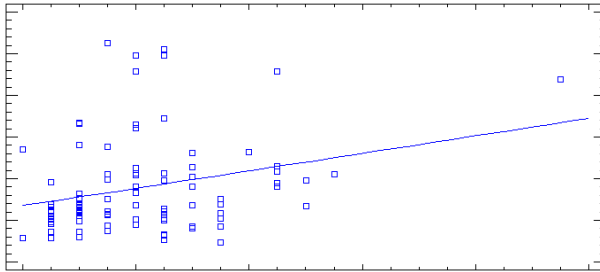


- a. Gráfico A.
- b. Gráfico B.
- c. Gráfico C.

**Resultado de aprendizaje 37:**

Asociar la descripción de la relación de dos variables con la representación en un diagrama de dispersión.

38. El siguiente gráfico de dispersión muestra la relación entre la medición del perímetro abdominal y la presión intraabdominal en pacientes graves.



Elige la mejor interpretación de la relación entre el perímetro abdominal y la presión intraabdominal basada en el diagrama de dispersión.

- a. Este gráfico muestra una fuerte relación lineal negativa entre el perímetro abdominal y la presión intraabdominal en pacientes graves.
- b. Este gráfico muestra una moderada relación lineal entre el perímetro abdominal y la presión intraabdominal en pacientes graves.
- c. **Este gráfico muestra muy poca o ninguna relación lineal entre el perímetro abdominal y la presión intraabdominal en pacientes graves.**

**Resultado de aprendizaje 38:****Describir la relación entre dos variables que se muestran en un gráfico de dispersión.**

39. Se ha estudiado una muestra aleatoria de 1000 individuos en los que se ha encontrado una fuerte correlación positiva, estadísticamente significativa, entre la capacidad auditiva y el tiempo (en horas) que se utilizan reproductores portátiles de música (mp3).

Cuál de las siguientes opciones es la mejor interpretación de estos resultados.

- a. **No podemos concluir que el uso de reproductores portátiles de música aumenten la capacidad auditiva, porque este tipo de diseño no nos permite inferir causalidad.**
- b. Esta muestra es demasiado pequeña para sacar conclusiones sobre la relación entre el uso de reproductores portátiles y la capacidad auditiva.
- c. Este resultado indica que las personas que utilizan habitualmente reproductores portátiles tienen mayor capacidad auditiva.

**Resultado de aprendizaje 39:****Entender que la correlación no implica causalidad.**

40. Un estudio en el ámbito universitario utilizó datos de los últimos 15 años para relacionar los costes de los proyectos de investigación con los tiempos de ejecución en días. Se obtuvo la siguiente ecuación de regresión para los datos disponibles en los que el proyecto de mayor duración fue de 3,5 años.

$$\text{Coste (€)} = 2142 + 35 \cdot \text{Tiempo (días)}$$

¿Cuál es el mejor método para predecir el coste de un proyecto de 5 años de duración?

- a. Sustituir el valor de 1825 días (5 años) en el tiempo en la ecuación de regresión y calcular el coste previsto.
- b. Trazar la línea de regresión en el diagrama de dispersión, localizar 1825 en el eje horizontal, y leer el valor correspondiente al coste en el eje vertical.
- c. **Ningún método basado en estos datos es apropiado para hacer una predicción para un proyecto de 5 años.**
- d. Ambos métodos son apropiados para hacer una predicción del coste basado en estos datos.

**Resultado de aprendizaje 40:**

**Comprender que no es prudente extrapolar un modelo de regresión.**

## ANEXO 2. CUESTINARIO PARCIAL DE RAZONAMIENTO ESTADÍSTICO (PARES1). ESTADÍSTICA DESCRIPTIVA

1. En una encuesta a los alumnos se les pregunta ¿Cuál es la red social que usas mayoritariamente? Los datos recogidos en respuesta a esta pregunta ¿Qué tipo de datos son?

- a. cualitativos
- b. cuantitativos
- c. continuos

**Resultado de aprendizaje 1:**  
**Reconocer variables de tipo cualitativo.**

2. En los registros sobre la localización del catéter venoso se recogen los datos de la siguiente manera: 1=antebrazo, 2= dorso de la mano, 3= flexura del codo, 4= brazo, 5= subclavia, 6= yugular, 7= otras.

¿Qué tipo de variable es esta?

- a. cualitativa
- b. cuantitativa
- c. continua

**Resultado de aprendizaje 2:**  
**Reconocer que una variable es cualitativa a pesar de presentarse codificada numéricamente.**

3. El responsable de un Centro de Atención Primaria quiere conocer la opinión de los usuarios para mejorar el servicio y las instalaciones del Centro. ¿Cuál de los siguientes planes de recogida de datos proporcionaría la mejor representación de las opiniones de los usuarios de ese Centro?

- a. Encuestar diariamente al décimo usuario que entra en el Centro de Salud entre las 9:00H y las 15:00H hasta completar una muestra de 100 usuarios.
- b. Tomar una muestra al azar de 50 números de tarjeta sanitaria de entre todos los usuarios del Centro de Salud y enviarles una encuesta.**
- c. Colocar un anuncio en la revista del Centro de Salud, invitando a los usuarios a completar una encuesta por Internet. Recoger los datos de las primeras 200 encuestas que respondan.
- d. Todas las anteriores serían igualmente eficaces.

**Resultado de aprendizaje 3:**

**Seleccionar el muestro mas apropiado para una investigación y entender de que un muestreo aleatorio es más representativo.**

4. En el mismo Centro de Salud querían conocer la media de consultas por usuario, que se realizan en un año. Para ello dividieron el número total de consultas realizadas en ese año entre 500, que es el total de usuarios del Centro de Salud.

¿Cuál de las siguientes afirmaciones es cierta si la media anual de consultas realizadas por paciente es de 2,2?

- a. La mitad de los usuarios del Centro de Salud. realizan más de dos consultas al año.
- b. Ese año se realizaron un total de 11000 consultas.**
- c. El número más común de consultas por usuarios es de 2,2.
- d. Ninguna de las anteriores es cierta.

**Resultado de aprendizaje 4:**

**Comprender el concepto de media aritmética y sus propiedades.**

5. La distribución de la tensión arterial sistólica (TA), en mmHg., de los trabajadores de una empresa está fuertemente sesgada a la derecha. Las dos medidas centrales de TA son 150 y 116 . ¿Qué número representa la media y cuál la mediana de la TA de los trabajadores de esa empresa?

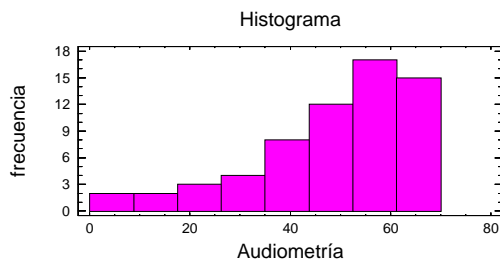
- a. 116 es la media y 150 es la mediana.
- b. 116 es la mediana y 150 es la media.**
- c. No hay suficiente información para decir cuál es cada una.



**Resultado de aprendizaje 5:**

**Comparar y reconocer las medidas de tendencia central de una distribución de datos que no es simétrica. Entender el concepto de sesgo y asociarlo al valor de la mediana.**

6. Por la información de este gráfico que representa los valores de audiometrías (pérdida en decibelios del oído derecho en la frecuencia 3000Hz) de un grupo de trabajadores de la empresa anterior, qué valores para la media y la mediana parecen más acertados.



- mediana= 35.0 y media= 28.0
- mediana= 47.0 y media= 49.3
- mediana= 52.0 y media= 47.0**
- mediana= 52.3 y media= 52.2

**Resultado de aprendizaje 6:**

**Reconocer en un histograma la asimetría y el sesgo de una distribución de datos a la hora de asignar los valores correspondientes a las medidas de tendencia central, media y mediana.**

7. La profesora de estadística calculó la mediana de las puntuaciones (número de respuestas correctas) de un examen test realizado a todos sus alumnos. Más tarde decidió premiar con 3 puntos a los 10 estudiantes que tuvieron mejor puntuación.

La mediana de la nueva distribución de calificaciones será:

- Más baja que la mediana de la distribución de puntuaciones inicial.
- Igual que la mediana de la distribución inicial de puntuaciones.**
- Más alta que la mediana de la distribución inicial de puntuaciones.
- En función de la asimetría, mayor o menor que la mediana de la distribución inicial de puntuaciones.

**Resultado de aprendizaje 7:**

Comprender el concepto de mediana como medida central que no se ve afectada por valores extremos.

Las preguntas 8 y 9 se refieren a la siguiente situación:

Unos estudiantes de Enfermería realizaron un estudio sobre el perímetro craneal de los niños (varones) de 0 a 1 mes, analizando los datos de todos los niños nacidos en un determinado hospital en un año. El análisis dio como resultado un valor para la media de 34,8 cm y una mediana de 33,74 cm.

8. Un estudiante de enfermería afirma que la mediana indica que la mayoría de los niños varones de 0 a 1 mes tiene un perímetro craneal alrededor de 33,74 cm. ¿Tú qué respondes a eso?

- a. Estoy de acuerdo, la mediana es un promedio y esa es la información que proporciona un promedio.
- b. Estoy de acuerdo, un perímetro craneal de 33,74 cm es representativo de la población de niños de menos de un mes.
- c. Estoy en desacuerdo, la mayoría de los niños tienen perímetros craneales mayores de 33,74.
- d. **Estoy en desacuerdo, la mediana te dice que la mitad de los niños de menos de un mes tiene un perímetro craneal menor de 33,74 cm.**

**Resultado de aprendizaje 8:**

Comprender el concepto de mediana de una distribución de datos.

9. Los alumnos del grupo comprobaron que habían cometido un error y que un valor introducido como 138 debía haberse introducido como 38. Se rectificó el dato y se calcularon de nuevo todos los parámetros estadísticos. Para la nueva situación ¿Cuál de las siguientes afirmaciones es verdadera?

- a. El valor de la mediana disminuye y el valor de la media sigue siendo el mismo.
- b. Ambos valores de la mediana y de la media disminuyen.
- c. **El valor de la mediana se mantiene igual y el valor de la media disminuye.**

**Resultado de aprendizaje 9:**

Comprender la diferencia entre las dos medidas de tendencia central más importantes, la media y la mediana, respecto a verse afectadas por los valores extremos de la distribución.

10. Un grupo de alumnos de estadística responde a un examen test de 20 ítems, en la que cada ítem vale 1 punto. La desviación estándar para la distribución de las puntuaciones resultantes del examen es 0.

Es cierto que:

- Alrededor de la mitad de las puntuaciones estaban por encima de la media.
- Debe haberse cometido un error de cálculo.
- Todo el mundo respondió correctamente el mismo número de ítems.**
- La media, mediana, y moda deben ser todas 0.

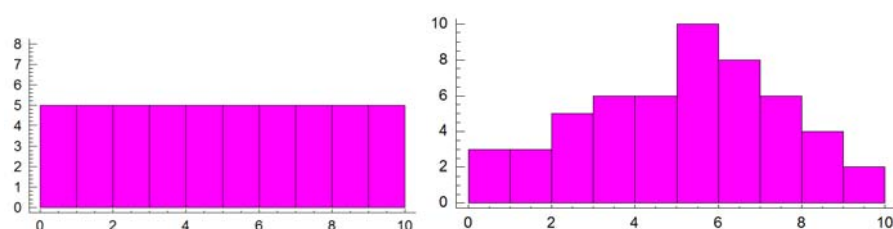
**Resultado de aprendizaje 10:**

Comprender el concepto de desviación estándar y entender que cuando la desviación típica tiene un valor de cero, todos los datos de la muestra tienen el mismo valor.

Las preguntas 11 y 12 se refieren a la siguiente situación:

Para cada par de gráficos, determina cual tiene la desviación estándar más alta (no es necesario hacer ningún cálculo para responder a esta pregunta.).

11.

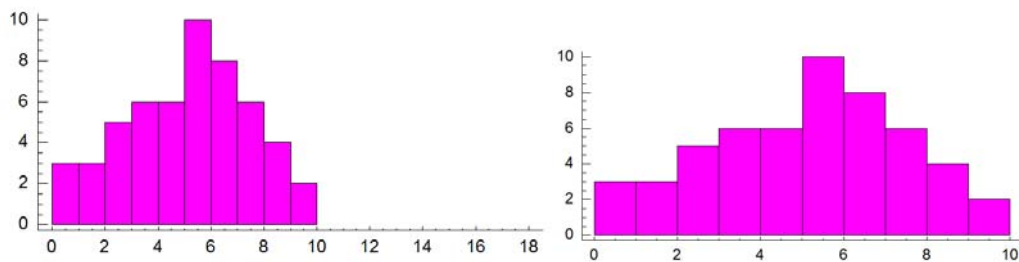


- A tiene una desviación estándar mayor que B.**
- B tiene una desviación estándar mayor que A.
- Ambos gráficos tiene la misma desviación estándar.

**Resultado de aprendizaje 11:**

Asociar la desviación estándar al grado de homogeneidad de los datos y la proximidad de los mismos al valor de la media, reconociendo la mayor o menor variabilidad de dos conjuntos de datos representados en un histograma.

12.



- A tiene una desviación estándar mayor que B.
- B tiene una desviación estándar mayor que A.
- Ambos gráficos tiene la misma desviación estándar.**

**Resultado de aprendizaje 12:**

Reconocer que dos muestras tienen la misma desviación estándar aunque se representen en histogramas que aparentemente tienen diferente dispersión de los datos.

13. Un profesor pasa a los alumnos un test de anatomía de 50 preguntas. El estudiante recibe un punto por cada respuesta correcta; 0 puntos por no responder y pierde un punto por cada respuesta incorrecta. La puntuación total del test puede oscilar entre los + 50 puntos y los -50. El profesor calcula la desviación estándar (desviación típica) de los resultados del examen para la clase que es  $-2'30$ . ¿Qué información nos da ese dato?

- La desviación estándar se calculó incorrectamente.**
- La mayoría de los estudiantes obtuvieron puntuaciones negativas.
- La mayoría de los estudiantes obtuvo calificaciones por debajo de la media.
- Ninguna de las anteriores.

**Resultado de aprendizaje 13:**

Comprender el concepto y el cálculo de la desviación típica reconociendo que no puede tener valores negativos.

14. Consideremos dos poblaciones del mismo tamaño (10.000). La población 1 consiste en todos los estudiantes de una Universidad. La Población 2 está formada por todos los residentes de una pequeña ciudad. Considere la variable Edad, ¿Qué población tendría, con mayor probabilidad, la desviación estándar más grande?

- a. La Población 1 es más probable que tenga una desviación estándar mayor que la Población 2.
- b. La Población 2 es más probable que tenga una desviación estándar mayor que la Población 1.**
- c. Tendrían probablemente la misma desviación estándar para la edad, porque son poblaciones del mismo tamaño.
- d. No hay suficiente información para saberlo.

**Resultado de aprendizaje 14:**

Entender la desviación estándar como medida del grado de homogeneidad de una población respecto a una variable.

**Las preguntas 15 y 16 se refieren a la siguiente situación:**

Se presentan a continuación dos listados con datos de presión arterial sistólica (mmHg) que corresponden a dos grupos diferentes de pacientes (Lista A y Lista B). Seleccione la mejor estimación para la desviación estándar sabiendo que la media de cada listado es 150. No se requieren cálculos para responder esta pregunta.

15. LIST A: 149, 151, 149, 151, 149, 151, 149, 151, 149, 151

- a. 1**
- b. 2
- c. 5

**Resultado de aprendizaje 15:**

Aplicar el concepto y cálculo de la mediana a un conjunto de datos en el que todos ellos distan de la media una sola unidad.

16. LIST B: 131, 136, 148, 150, 150, 153, 154, 156, 160, 162

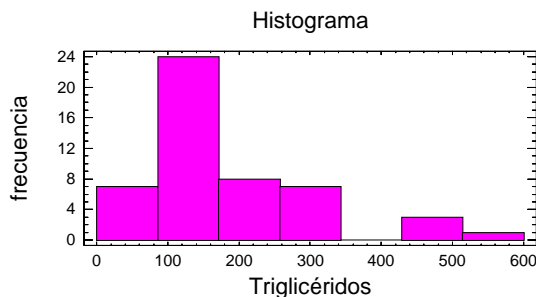
- a. 1
- b. 3
- c. 8
- d. 20

**Resultado de aprendizaje 16:**

Entender el valor máximo que puede tener la mediana utilizando los valores máximo y mínimo de un conjunto de datos.

17. Se midieron los niveles de Triglicéridos (mg/dl) de un grupo de 50 personas. El histograma representa los resultados.

¿Qué dos medidas serían más apropiadas para describir el centro y la dispersión de esta distribución?

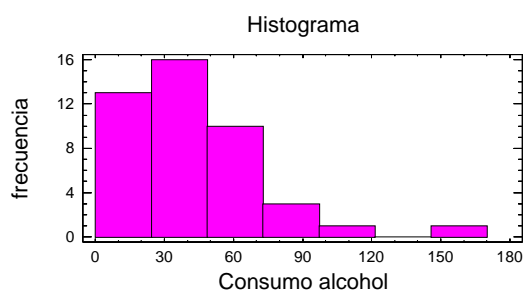


- a. Rango y media
- b. Media y mediana
- c. **Mediana y rango intercuartílico**
- d. Media y desviación estándar

**Resultado de aprendizaje 17:**

Reconocer sobre un histograma con puntuaciones atípicas, la conveniencia de utilizar la mediana y el rango intercuartílico para describir el centro y la dispersión de la distribución, frente a la media y desviación típica que afectadas por valores extremos pueden no representar al grupo.

18. El siguiente gráfico representa la distribución del consumo semanal de alcohol (g.) de una muestra aleatoria de estudiantes universitarios.



El rango para esta distribución es 156,7 g. Indica la mejor opción que expresa tu acuerdo o desacuerdo con la siguiente afirmación: “El rango no es un parámetro útil para valorar la variabilidad del consumo de alcohol en este grupo de datos”

- Estoy de acuerdo. Es demasiado vago.
- Estoy de acuerdo. Está afectado por los valores extremos.**
- Estoy de acuerdo, no usa información de los datos centrales
- Estoy en desacuerdo, un rango de 156,7 g. es una buena medida de variabilidad. porque los estudiantes pueden consumir cualquier cantidad de alcohol entre 0 y 156,7 g.

**Resultado de aprendizaje 18:**

**Reconocer al rango como un indicador débil de la dispersión ya que sólo utiliza para su cálculo dos puntuaciones que pueden ser atípicas.**

19. Se ha realizado la determinación del contenido de Calcio en la sangre de 25 pacientes.

Los resultados han sido los siguientes:

CLASES	Marca de clase $X_i$	Frecuencia absoluta $f_i$	Frecuencia acumulada $F_i$	Frecuencia relative $f_r$	Frecuencia r. acumulada $Fr$	%
8.25 – 8.65	8.45	2	2	0.08	0.08	8
8.65 – 9.05	8.85	5	7	0.2	0.28	20
9.05 – 9.45	9.25	10	17	0.4	0.68	40
9.45 – 9.85	9.65	6	23	0.24	0.92	24
9.85 – 10.25	10.05	2	25	0.08	1	8

¿Qué información nos aporta el valor marcado en gris?

- Es el valor de calcio más frecuente entre los 25 pacientes.
- Es el número de pacientes que tienen un valor de calcio entre 9.05 y 9.45.
- Hay 17 pacientes con un valor de calcio superior a 9.45.
- Hay 17 pacientes con un valor de calcio inferior a 9.45.**

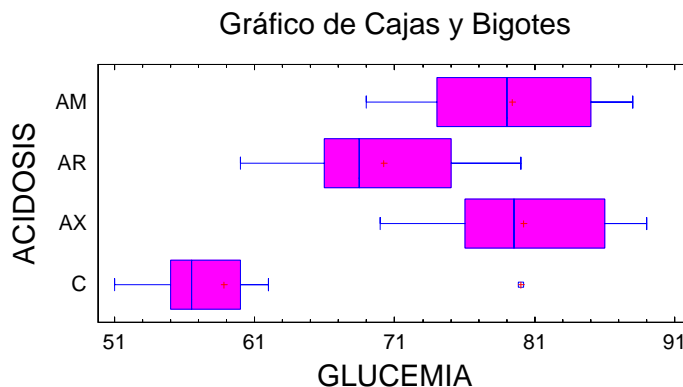
**Resultado de aprendizaje 19:**

**Interpretar el resumen de datos recogido en una tabla de frecuencias.**

20. Para evaluar la influencia del tipo de acidosis, medida en el cordón umbilical, del recién nacido sobre los niveles de glucemia, se obtuvieron los datos de la siguiente tabla:

	Niveles de glucemia									
Controles (niños sin acidosis)	51	56	58	60	62	63	65	68	72	73
Acidosis Respiratoria	60	65	66	68	68	69	73	75	78	80
Acidosis Metabólica	69	73	74	78	79	79	82	85	87	88
Acidosis Mixta	70	75	76	77	79	80	82	86	88	89

El siguiente gráfico muestra la distribución de los niveles de glucemia para cada grupo de recién nacidos





¿En qué tipo de acidosis los niveles de glucemia tienen el rango intercuartílico menor?

- a. Acidosis respiratoria
- b. Acidosis metabólica
- c. Acidosis mixta
- d. **Grupo Control**

**Resultado de aprendizaje 20:**

Comprender la representación de un conjunto de datos en un gráfico de cajas y capacidad para interpretar comparativamente el rango intercuartílico de varios conjuntos de datos.

21. Si fuera eliminado el valor extremo (outlier) del grupo control ¿Cuál de los siguientes parámetros se vería menos afectado?

- a. El rango.
- b. **El rango intercuartílico**
- c. La desviación estándar.
- d. Ninguno de los anteriores.

**Resultado de aprendizaje 21:**

Comprender que la introducción o eliminación de valores extremos en un conjunto de datos no afecta al valor del rango intercuartílico.

22. ¿Qué gráfico es más apropiado para mostrar los resultados de un estudio sobre los diferentes fármacos antidepresivos utilizados por los residentes de un centro de mayores?

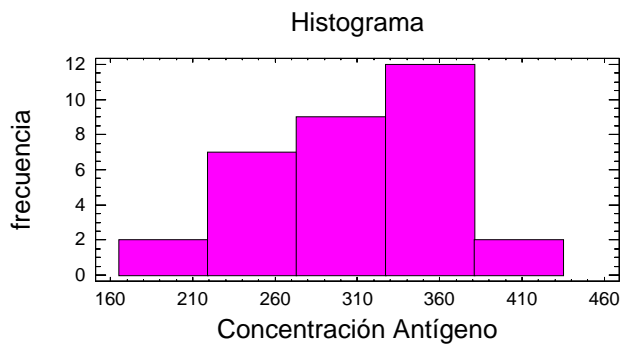
- a. **Gráfico de sectores.**
- b. Histograma.
- c. Cualquiera de los dos, histograma o gráfico de sectores.
- d. Ninguno de los anteriores.

**Resultado de aprendizaje 22:**

Seleccionar el gráfico más adecuado para representar una variable cualitativa.

23. Una investigación publicada en la revista *Clinical Immunology* se refiere a las anomalías inmunológicas en niños autistas. Se tomaron mediciones de la concentración sérica de un antígeno en una muestra de niños autistas menores de 10 años. Las mediciones en unidades por milímetro de suero se representan a continuación.

¿Cuántos niños tenían niveles del antígeno por encima de 310?



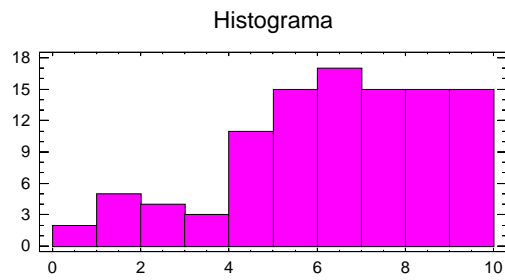
- a. 17
- b. 9
- c. 24
- d. 8

**Resultado de aprendizaje 23:**

**Interpretar información relativa a los valores de frecuencias de una variable representada en un histograma.**

**Las preguntas 24 y 25 se refieren a la siguiente situación:**

Al comenzar el curso se preguntó a un grupo de estudiantes de estadística, cuál era su nivel de ansiedad o angustia con el que abordaban la asignatura en una valoración de 0 a 10, siendo 0= muy baja y 10= muy alta. Los resultados de la encuesta se representan en el siguiente gráfico:



24. ¿Qué significan los números representados en el eje vertical?

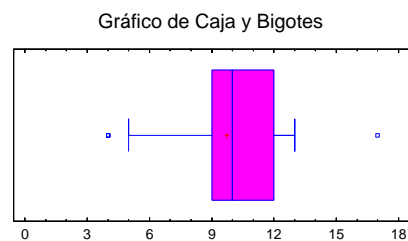
- a. La variable independiente.
- b. Las puntuaciones de la encuesta.
- c. La variable dependiente.
- d. El número de estudiantes.**

**Resultado de aprendizaje 24:**

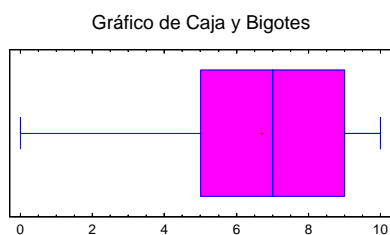
**Reconocer las variables representadas en cada uno de los ejes de un histograma.**

25. ¿Cuál de los siguientes diagramas de cajas representa el mismo conjunto de datos que muestra el anterior histograma?

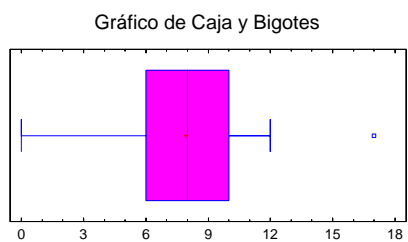
a.



b.



C.

**Resultado de aprendizaje 25:****Reconocer un conjunto de datos en dos representaciones gráficas diferentes.**

### ANEXO 3. CUESTINARIO PARCIAL DE RAZONAMIENTO ESTADÍSTICO (PARES2) . DISTRIBUCIÓN NORMAL E INTERVALOS DE CONFIANZA

1. Las puntuaciones de un examen test se distribuyen de forma normal con una media de 100 y una desviación estándar de 20. Si estas puntuaciones se convierten en puntuaciones típicas, ¿Cual de las siguientes afirmaciones sería la correcta?

- a. Las puntuaciones tanto de la media como de la mediana serán igual a 0.
- b. La media será igual a 0, pero la mediana no puede ser determinada.
- c. La media de las puntuaciones típicas será igual a 100.
- d. La media de las puntuaciones típicas será igual a 5.

**Resultado de aprendizaje 1:**  
Entender el concepto de puntuación típica.

2. Una enfermera registra una muestra grande de medidas del ritmo cardíaco que siguen aproximadamente una distribución normal de la que sólo conocemos tres parámetros: la media = 100 pulsaciones por minuto, el valor mínimo = 65 pulsaciones por minuto, y el máximo = 155 pulsaciones por minuto. ¿Cuál de los siguientes resultados tiene más posibilidad de ser la desviación estándar de la distribución?

- a. 5
- b. 15
- c. 35
- d. 90

**Resultado de aprendizaje 2:**  
Comprender las propiedades de la distribución normal y su simetría. Conocer que en muestras grandes la amplitud o recorrido de una variable comprende unas seis desviaciones típicas.

Los ítems 3, 4 y 5 se refieren a la siguiente situación:

3. Entre los diabéticos el nivel de glucosa en sangre en ayunas, puede suponerse que sigue una distribución normal, con media 106 mg por cada cien ml. y desviación típica 8 mg por cien ml. María tiene una glucemia de 114. ¿En qué proporción tiene María la glucemia más alta que la media de la población diabética?

- a. 0.32
- b. 0.49
- c. 0.68
- d. **0.84**

**Resultado de aprendizaje 3:**

Comprender las propiedades de la distribución normal: Entender que si conocemos en cuántas desviaciones típicas se distancia de la media un sujeto, podemos conocer la probabilidad y las proporciones de casos esperados por encima o por debajo de dicho sujeto.

4. El 2.5% de los diabéticos tiene niveles de glucosa más altos que Miguel. ¿Cuál será el nivel de glucosa de Miguel?

- a. 106
- b. 114
- c. **122**
- d. 130

**Resultado de aprendizaje 4:**

Comprender las propiedades de la distribución normal en cuanto a la proporción de sujetos que se encuentran a distancia de una, dos o tres desviaciones típicas.

5. En qué percentil es más probable que se encuentre Miguel, respecto a su nivel de glucosa.

- a.  $P_{25}$
- b.  $P_{75}$
- c.  **$P_{97}$**
- d.  $P_1$

**Resultado de aprendizaje 5:**

**Asociar la puntuación de un sujeto respecto a una variable que se distribuye de forma normal con el percentil correspondiente a dicha puntuación, conociendo las proporciones de individuos esperados por encima de dicho sujeto.**

6. La distribución de la altura de un hombre adulto es aproximadamente normal con una media de 175.2 cm y una desviación estándar de 50.8 centímetros. La altura de David tiene una puntuación típica de -0.5 cuando se compara con todos los hombres adultos. ¿Cuál de las siguientes afirmaciones es verdadera?

- a. David mide menos de 175.2 cm.
- b. La altura de David está a la mitad de una desviación estándar por debajo de la media.
- c. David mide 149.8 cm.
- d. **Todas las anteriores son ciertas.**

**Resultado de aprendizaje 6:**

**Comprender el concepto de puntuación típica y aplicarlo para situar a un sujeto en el conjunto de datos, respecto a la media y al posible valor de la variable para dicho sujeto.**

7. Una compañía de seguros de salud está interesada en los niveles de colesterol en personas de 40 años o mayores de esta edad. Se tomó una muestra al azar de la población diana de 100 personas. Se obtuvo la siguiente información de la muestra: Media= 158 mg., mediana= 159 mg., desviación estándar = 20mg. Una persona tiene el nivel de colesterol en 175 mg. Basándonos tan sólo en estos parámetros ¿Es esta una medida inusual y elevada de colesterol para alguien de esta población? ¿Por qué o por qué no?

- a. Sí, porque 175 es 17 mg. más alto que la media de los niveles de colesterol.
- b. Sí, porque es mejor tener un nivel de colesterol bajo.
- c. **No, porque 175 está a menos de una desviación estándar por encima de la media.**
- d. No, porque no está por encima de 200 mg. que es el nivel máximo de colesterol recomendado para un adulto.

**Resultado de aprendizaje 7:**

**Interpretar el valor que tiene en un sujeto una variable que se distribuye de forma normal respecto al grupo estudiado.**

8. En la asignatura de estadística los alumnos realizan una serie de pruebas parciales de razonamiento estadístico. Cada test evalúa unos contenidos diferentes y tiene un número diferente de ítems y diferentes medias y desviaciones estándar. Una de las estudiantes, María, quiere saber en cuál de esas pruebas ha obtenido mejores resultados. ¿Cuál de los siguientes métodos debería usar?

- Comparar sus puntuaciones totales en cada test.
- Comparar su porcentaje de respuestas correctas en cada test.
- Comparar su puntuación típica en cada test.**
- Todas las anteriores deberían llevarnos a la misma conclusión.

**Resultado de aprendizaje 8:**

**Comprender el concepto de puntuación típica que utiliza la desviación estándar como unidad de medida y permite comparar resultados de procedencia distinta.**

9. Cristina está matriculada en la asignatura de estadística, y accedió a la Universidad por prueba de acceso de mayores de 25 años y ha obtenido una puntuación de 26 en el examen. El profesor diferencia a los alumnos en dos grupos, los que accedieron por Bachillerato y los que lo hicieron por otro acceso y separa las puntuaciones del examen para esos dos grupos. Ambos grupos de puntuación están distribuidos aproximadamente de forma normal y tienen la misma media, pero las puntuaciones de los que accedieron por bachillerato tiene una desviación estándar menor. Se calcula la puntuación típica para el examen de Cristina en ambos casos (para el grupo de bachillerato y el de otro acceso). Dado que la puntuación de Cristina es muy superior a la media, ¿Cuál de las siguientes afirmaciones sería verdadera acerca de estas dos puntuaciones típicas de Cristina?

- La puntuación típica de Cristina basada en las puntuaciones de los alumnos que accedieron por bachillerato será más alta.**
- La puntuación típica de Cristina basada en las puntuaciones del grupo de "otro acceso" sería más alta.
- Las dos puntuaciones típicas serían iguales.

**Resultado de aprendizaje 9:**

**Entender la desviación típica como unidad de medida que permite situar la puntuación de un sujeto respecto al resto del grupo.**



10. Se toman dos muestras diferentes del mismo producto fabricado por una empresa de nutrición parenteral, del que se desconoce la media y la desviación típica de su contenido en grasa. La primera muestra tendrá 40 datos y la segunda tendrá 100 datos. Se construirá un intervalo de confianza del 95% para cada muestra con el fin de estimar la media del contenido en grasa de la población (producción del producto). ¿Qué intervalo de confianza tendrá una mayor precisión para estimar la media de la población?

- a. **El intervalo de confianza basado en la muestra de 100 datos es más ajustado y más preciso.**
- b. Ambos intervalos de confianza tendrán la misma precisión.
- c. El intervalo de confianza basado en la muestra de 40 datos es más ajustado y más preciso.

**Resultado de aprendizaje 10:**

**Comprender el concepto de error típico de la media y entender que los intervalos de confianza para la media calculados en muestras grandes son más precisos que en muestras pequeñas.**

11. Se calcula un intervalo de confianza del 95% para estimar la media de ingresos hospitalarios por día en una determinada ciudad. ¿Cuál de los siguientes valores estará, sin duda, dentro de los límites de este intervalo de confianza?

- a. La media de población.
- b. La media de la muestra.**
- c. La desviación estándar de la media de la muestra.
- d. Ninguna de las anteriores.

**Resultado de aprendizaje 11:**

**Entender que un intervalo de confianza para estimar la media poblacional a partir de una muestra, no asegura al 100% que la media poblacional se encuentre en dicho intervalo.**

12. Cada uno de los 60 estudiantes de la clase de estadística selecciona una muestra diferente al azar de 35 alumnos de la Universidad, de una población de 5000 Universitarios. Usando sus datos, cada estudiante construye un intervalo de confianza del 90 % para la edad media de los 5000 estudiantes. ¿Cuál de las siguientes conclusiones es correcta?

- a. Alrededor del 10% de las medias muestrales no serán incluidas en el intervalo de confianza.
- b. Alrededor del 90% de los intervalos de confianza calculados sobre las muestras contendrán el valor de la media de la población.**
- c. Es probable que el 90% de los intervalos de confianza sean idénticos.
- d. Alrededor del 10% de las edades de las muestras no se encontrarán en estos intervalos de confianza.

**Resultado de aprendizaje 12:**

**Comprender que si extraemos varias muestras del mismo tamaño y calculamos un intervalo de confianza del 90% para cada muestra, el 95% de todos los intervalos van a incluir la media poblacional.**

13. Para estudiar la presión intraocular en individuos de edad avanzada, se tomó una muestra aleatoria de 100 personas mayores y se calculó un intervalo de confianza del 95% para la media de presión intraocular (mm Hg). ¿Qué nos dice en este caso un intervalo de confianza del 95% para la presión intraocular? Selecciona la mejor respuesta.

- a. Estoy seguro al 95% de confianza de que este intervalo incluye la media muestral de la presión intraocular.
- b. Estoy seguro que el 95% de las personas mayores tendrá una presión intraocular dentro de este intervalo.
- c. Estoy seguro al 95% de que la mayoría de las personas de edad avanzada tienen una presión intraocular en este intervalo.
- d. Estoy seguro al 95% que este intervalo incluye la media de la presión intraocular de la población de personas mayores.**

**Resultado de aprendizaje 13:**

**Interpretar la información que nos aporta un intervalo de confianza del 95% para la estimación de la media poblacional. Entendiendo que con un 95% de seguridad la verdadera media de la población se encontrará dentro del intervalo.**

14. Supongamos que a una muestra aleatoria de 41 estudiantes de la Universidad se les pide que mida la longitud de su pie derecho en centímetros. Un intervalo de confianza del 95% para la media de la longitud del pie de los estudiantes de esta universidad resulta ser (21.709 ; 25.091). Si se calculara en su lugar un intervalo de confianza del 90%, ¿Cómo diferiría del intervalo de confianza del 95%?

- a. **El intervalo de confianza del 90% sería más estrecho.**
- b. El intervalo de confianza del 90% sería más ancho.
- c. El intervalo de confianza del 90% sería el mismo que el intervalo de confianza del 95%.

**Resultado de aprendizaje 14:**

**Comprender que a mayor nivel de seguridad con el que se construye un intervalo de confianza, los límites del intervalo se alejan, haciendo más ancho el intervalo de estimación.**

15. El artículo de un periódico asegura con el 95% de confianza que entre el 55% y el 65% de todos los alumnos de secundaria de España consume alcohol. Este intervalo de confianza está basado en una encuesta realizada a 2000 estudiantes de secundaria de Madrid. ¿Cómo interpretarías el intervalo de confianza de este artículo del periódico?

- a. El 95% de las grandes ciudades de España tienen entre un 55% y un 65% de estudiantes de secundaria que consume alcohol.
- b. Si tomamos numerosas muestras de alumnos de secundaria de diferentes ciudades, el 95% de las muestras tendrían entre un 55% y un 65% de alumnos de secundaria que consumen alcohol.
- c. **No puedes usar ese intervalo de confianza para generalizar a todos los adolescentes de España por la forma de tomar la muestra.**
- d. Podemos estar seguros al 95% de que entre el 55% y el 65% de los estudiantes de secundaria de España consumen alcohol.

**Resultado de aprendizaje 15:**

**Entender las limitaciones de la estimación por intervalos a partir de muestras no aleatorias.**



## ANEXO 4. CUESTINARIO PARCIAL DE RAZONAMIENTO ESTADÍSTICO (PARES3). PRUEBAS DE SIGNIFICACIÓN

1. Un laboratorio farmacéutico que fabrica cierto tipo de suero sabe que la presencia de impurezas está relacionado con la temperatura a la que se realiza la mezcla de sus componentes, pero tiene controladas las condiciones de producción de sueros y no contienen impurezas. Para comprobar si los sueros están exentos de impurezas. ¿Cuál de las siguientes opciones corresponde al planteamiento correcto de la hipótesis nula de la prueba?

- a. El control del sistema de producción evita la presencia de impurezas en los sueros.
- b. El control del sistema de producción no evita la presencia de impurezas en los sueros.
- c. Los sueros contienen impurezas.

### **Resultados de aprendizaje 1:**

Comprender la forma correcta de plantear la hipótesis nula y alternativa de un estudio, como hipótesis enfrentadas. Reconociendo en un ejemplo concreto la hipótesis alternativa como aquella que expresa la situación que se pretende comprobar en el proceso de investigación.

2. Si un investigador tiene la esperanza de demostrar que los resultados de un experimento fueron estadísticamente significativos ¿Qué prefiere?

- a. un p-valor alto
- b. un p-valor bajo
- c. el p-valor no está relacionado con la significación estadística.

### **Resultados de aprendizaje 2:**

Reconocer cuándo la magnitud de un p-valor pone de manifiesto una significación estadística.

**Las preguntas 3, 4 y 5 se refieren a la siguiente situación:**

En el control de un nuevo análisis para determinar la presencia de un determinado virus, se realizó una prueba de hipótesis donde:

$H_0$  = El análisis detecta el virus

$H_a$  = El análisis no detecta el virus

Identificar cada una de las siguientes afirmaciones como error de Tipo I o error de Tipo II

3. El análisis detecta el virus, pero en realidad el virus no está presente

- a. Error Tipo I
- b. Error Tipo II**
- c. No es un error

**Resultados de aprendizaje 3:**

Identificar en el contexto de una prueba de hipótesis el tipo de error asociado a la aceptación de la hipótesis nula.

4. El análisis no detecta el virus, pero en realidad el virus está presente

- a. Error Tipo I**
- b. Error Tipo II
- c. No es un error

**Resultados de aprendizaje 4:**

Identificar en el contexto de una prueba de hipótesis el tipo de error asociado a rechazar la hipótesis nula.

5. ¿Cuál de los dos errores consideras más importante?

- a. Error Tipo I**
- b. Error Tipo II
- c. Los dos errores son igualmente importantes

**Resultados de aprendizaje 5:**

Reconocer el error de tipo I como el más importante desde el punto de vista de las pruebas de hipótesis en el contexto de un ejemplo que ayuda a entender la gravedad de rechazar la hipótesis nula siendo cierta.

6. Un artículo de investigación da un p-valor de 0.001 para un análisis estadístico. ¿Qué interpretación del p-valor es la más exacta?

- a. No es probable que el resultado observado vuelva a ocurrir.
- b. Es improbable que el resultado observado sea debido únicamente al azar.**
- c. La variabilidad debida al muestreo es suficiente para explicar el resultado observado.

**Resultados de aprendizaje 6:**

**Entender que un p-valor significativo representa el hecho de que lo que se observa en la muestra no es producto de la variabilidad natural debida al muestreo.**

7. Un artículo de periódico afirma que la media de edad de las personas que utilizan comedores de ayuda social es de 40 años. Usted piensa que la media de edad es mayor y para comprobarlo, toma una muestra al azar de 100 personas que utilizan esta ayuda. La media de edad en esta muestra es de 41.2 años, con un p-valor  $< 0.05$ , comprobando que esta edad es significativamente mayor a los 40 años que afirmaba el artículo.

¿Cuál sería una adecuada interpretación de este resultado?

- a. El resultado estadísticamente significativo indica que la mayoría de las personas que utilizan los comedores sociales es mayor de 40 años.
- b. Aunque el resultado es estadísticamente significativo la diferencia de edad no es de importancia práctica.**
- c. Es un error. Esta diferencia es demasiado pequeña para ser estadísticamente significativa.

**Resultados de aprendizaje 7:**

**Reconocer que una diferencia puede ser significativa y sin embargo puede que no sea relevante en la práctica.**

8. Un artículo de periódico declaró que el Ministerio de sanidad recibió 812 cartas de todo el país expresando su opinión sobre la posibilidad de establecer el “copago” sanitario (pago de una tasa por receta).

De estas 812 cartas, 800 expresaron la opinión de que establecer el “copago” sanitario mejoraría la asistencia sanitaria. Se le plantea que con esta información, realice una prueba de significación estadística que ponga de manifiesto que el 95% de los españoles están a favor del “copago” sanitario.

¿Qué le contestaría a esta petición?

- a. Esta es una muestra lo suficientemente grande como para proporcionar una estimación precisa de la opinión pública española sobre el tema.
- b. No se cumplen las condiciones necesarias para una prueba de significación, por lo que la prueba no puede realizarse.**
- c. Con un gran número de personas a favor del “copago” sanitario no hay necesidad de una prueba estadística.

**Resultados de aprendizaje 8:**

**Reconocer situaciones en las que no se cumplen las condiciones necesarias para realizar una prueba de significación**

9. Un investigador realiza un experimento para comprobar que la asistencia a una sala de ludoterapia, afecta al proceso de recuperación de niños hospitalizados. Utiliza 20 niños que participan en el estudio, se realiza el análisis de los resultados y se obtiene un p-valor de 0.17

¿Cuál de las siguientes es una interpretación razonable de los resultados?

- a. Esto demuestra que su tratamiento experimental no tiene ningún efecto sobre la recuperación de los niños hospitalizados.
- b. Que podría haber un efecto del tratamiento, pero el tamaño de la muestra era demasiado pequeño para poder detectarlo.**
- c. Se debe rechazar la hipótesis nula.
- d. Hay evidencia de un efecto pequeño del tratamiento experimental sobre la recuperación de los niños.

**Resultados de aprendizaje 9:**

**Entender que si no podemos probar la hipótesis alternativa no significa que no sea cierta, sobre todo en muestras pequeñas.**

10. Los datos del Sistema Nacional de Salud informan que el tiempo medio de espera en España para las cirugías de oftalmología es de 64 días.

En la Comunidad de Madrid el mismo indicador es de 60.9 días. Se piensa que esta cifra es significativamente menor que la del país. Después de realizar la prueba estadística apropiada se obtiene un p-valor de 0.0025

¿Cuál de las siguientes afirmaciones es la mejor interpretación del p-valor?



- a. Un valor de  $p$  de 0.0025 proporciona una fuerte evidencia del tiempo medio de espera para las cirugías de oftalmología en la Comunidad de Madrid es menor que el tiempo de espera en todo el país.
- b. Un valor de  $p$  de 0.0025 indica que hay una pequeña posibilidad de que la Comunidad de Madrid tenga un tiempo de espera para las cirugías de oftalmología menor que todo el país.
- c. Un valor de  $p$  de 0.0025 proporciona evidencia de que la Comunidad de Madrid gestiona eficazmente sus listas quirúrgicas de oftalmología.
- d. Ninguna de las anteriores.

**Resultados de aprendizaje 10:**

Interpretar la información que aporta un  $p$ -valor muy pequeño en el contexto de un ejemplo.

Las preguntas 11 y 12 se refieren a la siguiente situación:

Se comparan dos métodos diferentes para recoger sangre para estudios de coagulación, analizando el tiempo parcial de tromboplastina activada (APTT) en dos muestras aleatorias de 24 pacientes a los que se les aplica cada uno de los dos métodos.

11. Si el contraste de hipótesis no encuentra diferencias significativas, significa que
  - a. Que el valor de  $t$  no sea estadísticamente significativo quiere decir que la diferencia no es extrapolable a las poblaciones representadas por esas dos muestras.
  - b. Puede que existan diferencias significativas entre los dos métodos de recolección de sangre, pero con una muestra pequeña estas diferencias puede que no se pongan de manifiesto.
  - c. Si contáramos con más sujetos la diferencia sería probablemente significativa.
  - d. Todas las anteriores son ciertas.

**Resultados de aprendizaje 11:**

Interpretar todas las conclusiones que se derivan de no rechazar la hipótesis nula en una comparación de muestras.

12. En el caso anterior. ¿Sería conveniente calcular el tamaño del efecto entre las dos muestras?

- a. Es innecesario, ya que no existen diferencias significativas entre los dos métodos de recogida de sangre.
- b. Es necesario, al tratarse de muestras de pequeño tamaño.**
- c. Podemos calcularlo, pero en este caso resultará un tamaño del efecto pequeño.

**Resultados de aprendizaje 12:**

**Entender que una diferencia que no es estadística significativa puede ser importante y sobre todo en muestras pequeñas en las que puede que existan diferencias no apreciadas en las pruebas de significación es importante aplicar un método que nos permita calcular la magnitud de las diferencias existentes.**

13. La cotinina es un metabolito de la nicotina que se forma en el organismo, y que permanece durante más tiempo en el cuerpo.

La comparación de los niveles de cotinina entre un grupo de fumadores con una media de 90,9 ng/mL y un grupo de no fumadores de media 7.9 ng/mL, dio como resultado un valor del tamaño del efecto igual a 2. ¿Cuál de las siguientes afirmaciones interpreta correctamente este dato?

- a. La media de cotinina del grupo de fumadores es mayor que los valores de cotinina del 50 % de los individuos del grupo de no fumadores.
- b. La media de cotinina del grupo de fumadores es mayor que los valores de cotinina de todos los individuos del grupo de no fumadores.
- c. La media de cotinina del grupo de fumadores es mayor que los valores de cotinina del 97% de los individuos del grupo de no fumadores.**
- d. La media de cotinina del grupo de fumadores está 2 ng/mL por encima que la media de cotinina del grupo de no fumadores.

**Resultados de aprendizaje 13:**

**Interpretar el tamaño del efecto en un contexto de comparación de medias siendo capaz de reconocer la distribución de datos en relación a la distancia en desviaciones típicas que las separan.**

14. La pulsioximetría es una técnica no invasiva que mide la saturación de oxígeno de la hemoglobina en sangre. Se quiere comparar los resultados de las medidas de saturación de oxígeno en dos grupos de pacientes. En uno de los grupos se ha utilizado para realizar las mediciones un pulsioxímetro fijo incorporado en un sistema de monitorización complejo y en el otro grupo de paciente se ha utilizado un pulsioxímetro portátil de pequeño tamaño. ¿Qué procedimiento será más adecuado para realizar esta comparación?

- a. **Al ser dos métodos distintos de medición, con distintos aparatos de medida lo mejor es compararlo con la magnitud de las diferencias o tamaño del efecto.**
- b. Test de significación estadística, mediante un contraste de hipótesis comparando las medias de los dos grupos.
- c. No es posible la comparación por ser dos métodos distintos de medición, con distintos aparatos de medida.

**Resultados de aprendizaje 14:**

**Comprender que al expresar la diferencias entre medias de forma tipificada pueden ser comparables datos recogidos con instrumentos de medición diferentes.**

15. Se diseña un estudio entre mujeres infectadas por VIH para probar si hay relación o no en que aparezcan infectadas también con el papilomavirus humano (PVH)

Si se obtiene un valor significativo de  $\chi^2$ , quiere decir que:

- a. Hay una diferencia entre la media de mujeres infectadas por VIH y la media de infección por el papilomavirus humano.
- b. La infección del papilomavirus humano (PVH) está determinada por padecer la infección por VIH.
- c. **Hay una asociación entre padecer infección por VIH y estar infectado por el papilomavirus humano (PVH).**

**Resultados de aprendizaje 15:**

**Reconocer una prueba de independencia de caracteres e interpretar los resultados en términos de asociación entre las variables.**

16. Un grupo de 1.000 adultos seleccionados al azar están clasificados en 4 grupos respecto al hábito de fumar: no fumadores, fumadores light, fumadores moderados y fumadores empedernidos. Al final de un período de diez años, se determinó en cada grupo la proporción existente de aquellos que habían desarrollado un cáncer de pulmón. La hipótesis nula dice que los 4 grupos tendrán la misma tasa de incidencia de cáncer de pulmón. Se obtiene un p-valor menor que el nivel de significación.

¿Cuál es la conclusión más apropiada?

- a. Que un adulto desarrolle o no cáncer de pulmón no está relacionado con su hábito de fumar.
- b. Hay una asociación entre la incidencia de cáncer y el nivel del consumo de tabaco.**
- c. Fumar produce cáncer.

**Resultados de aprendizaje 16:**

Entender el concepto de nivel de significación como referencia para la interpretación de los valores de p. Reconocer en un contraste de proporciones significativo, que existe asociación y no causalidad entre las variables.

17. Un centro sanitario acaba de despedir a 300 empleados. El análisis de género y la clasificación profesional son mostrados en la tabla siguiente. La compañía decide realizar un test de chi-cuadrado usando esta tabla de doble entrada. ¿Es este un procedimiento apropiado para determinar si hay relación entre el sexo, el tipo de trabajo y si un empleado es despedido o no?

	NO DESPEDIDO	DESPEDIDO
HOMBRE	300	200
MUJER	400	100
MÉDICO	420	80
ENFERMERÍA	280	220

- a. Sí, porque los datos están organizados en una tabla de doble entrada.
- b. Sí, un test de chi-cuadrado es apropiado porque las variables son categóricas.
- c. No, porque los datos se presentan como frecuencias y no como porcentajes.
- d. No, porque los individuos son contados en más de una celda de la tabla.**

**Resultados de aprendizaje 17:**

Detectar errores en una tabla de contingencia que recoge los datos para un análisis de asociación entre variables cualitativas, entendiendo que las observaciones tienen que ser independientes.

18. Un nuevo medicamento está siendo evaluado para determinar su efectividad en el tratamiento del eccema, una inflamación de la piel. Se seleccionaron trescientos pacientes con eccema para participar en el estudio. Los pacientes fueron divididos aleatoriamente en dos grupos. Un centenar de pacientes en el grupo experimental recibieron la medicación, mientras que doscientos pacientes del grupo control recibieron un placebo. Los resultados tras dos meses de estudio se muestran en la tabla. Se encontró una relación estadística significativa (valor-p = 0.0075). ¿Qué sugieren los datos acerca de la efectividad de la medicación?

	Grupo Experimental Medicación	Grupo Control Placebo
Mejora	40	50
No mejora	60	150

- Eficaz, porque el valor-p demuestra que hay una diferencia significativa entre el grupo experimental y el grupo de control.
- Eficaz, porque en el grupo experimental mejoró un porcentaje más alto que en el grupo de control.**
- No es efectivo, porque en el grupo experimental mejoraron 40 personas, pero en el grupo de control mejoraron 50.
- Es imposible saber la efectividad porque hay diferente número de personas en el grupo experimental y en el grupo control.

**Resultados de aprendizaje 18:**

**Interpretar, para un ejemplo dado, un resultado significativo en una prueba chi-cuadrado de independencia de caracteres.**

19. La asociación de alumnos de una facultad de Ciencias de la Salud, ha reunido una importante base de datos de antiguos graduados. Algunas de las variables de la base de datos son sexo, titulación, calificación media y primer salario. Están interesados en encontrar una relación entre estas variables. ¿Para cuál de los siguientes pares de variables sería apropiado un test de chi-cuadrado para examinar si hay relación o no entre las dos variables?

- Sexo y titulación.**
- Calificación media y sexo.
- Titulación y primer salario.
- Calificación media y primer salario.

**Resultados de aprendizaje 19:**

**Reconocer que en una prueba chí-cuadrado para examinar relación entre dos variables se tratan datos que corresponden a categorías en las que se clasifican los sujetos.**

20. Para comprobar que la realización de programas de preparación al parto tiene relación con la valoración de la experiencia del parto, se realizó un estudio entre 360 mujeres. Las frecuencias esperadas para este estudio se recogen en la siguiente tabla de contingencia:

Valoración del parto	Realización de preparación al parto	No realización de preparación al parto
Experiencia positiva	77	125
Experiencia negativa	56	90

La frecuencia esperada para las mujeres que realizaron preparación al parto y que tienen una valoración positiva del mismo es 77. ¿Qué representa este número?

- a. El número de mujeres que realizaron preparación al parto y que su valoración del mismo fue positiva.
- b. **El valor esperado del número de mujeres que realizaron preparación al parto, con valoración positiva del mismo si no hubiera relación entre la valoración del parto y el haber realizado el curso de preparación.**
- a. c. El porcentaje esperado de mujeres que realizaron preparación al parto, que tienen valoración positiva, si la hipótesis nula es cierta.
- c. Todo lo anterior es cierto.

**Resultados de aprendizaje 20:**

**Distinguir entre los dos tipos de frecuencias que se comparan en una prueba de independencia chí-cuadrado.**

Los ítems 21 y 22 se refieren a la siguiente situación:

La siguiente tabla muestra los resultados de una encuesta a 2979 usuarios del Sistema Sanitario Nacional para evaluar su opinión acerca de la atención recibida en los Centros de Atención Primaria, clasificándolos también en relación a sus estudios.

Nivel de estudios	OPINIÓN			totales
	Buena	Regular	Mala	
Ninguno	800	144	32	976
Primarios	905	312	67	1284
Bachiller	287	157	44	488
Medios	95	48	11	154
Superiores	38	32	7	77
Totales	2125	693	161	2979

21. Basándonos en la tabla, ¿Cuál es la proporción de usuarios que tienen estudios primarios y una buena opinión de la atención recibida en su Centro de Salud?

- a. **905/2979**
- b. 905/1284
- c. 905/2125

**Resultados de aprendizaje 21:**

Interpretar las frecuencias que se recogen en una tabla de contingencia en la que se relacionan dos variables, calculando, las proporciones respecto al total de la muestra, de la ocurrencia conjunta de dos sucesos.

22. Basándonos en la tabla, ¿Cuál es la proporción de los usuarios con estudios primarios que tienen una buena opinión de la atención recibida en su Centro de Salud?

- a. 905/2979
- b. **905/1284**
- c. 905/2125

**Resultados de aprendizaje 22:**

Interpretar las frecuencias que se recogen en una tabla de contingencia en la que se relacionan dos variables, calculando, las proporciones respecto a una categoría, de la ocurrencia de un suceso.

23. Se ha realizado un estudio sobre caries dental en niños de seis ciudades con diferentes cantidades de fluor en el suministro de agua potable.

Se seleccionó una muestra aleatoria de 125 niños de cada ciudad con los siguientes resultados:

CIUDAD	niños sin caries	niños con caries	TOTAL
A	38	87	125
B	8	117	125
C	30	95	125
D	44	81	125
E	64	61	125
F	32	93	125
TOTAL	216	534	750

Queremos averiguar si existen diferencias significativas en la variable niños con caries/ sin caries, entre las seis poblaciones de las que se han extraído las muestras aleatorias.

Si se obtiene un valor no significativo de  $\chi^2$ , quiere decir que:

- La variable niños con/sin caries tiene la misma distribución de probabilidad en las seis ciudades.
- Las muestras de las seis ciudades son homogéneas respecto a la variable niños con/sin caries.
- Las diferencias observadas en el número de niños con o sin caries en las seis ciudades es debido al azar al obtener las muestras.
- Todo lo anterior es cierto.**

**Resultados de aprendizaje 23:**

Interpretar para un ejemplo dado, un resultado no significativo en una prueba chi-cuadrado de homogeneidad de muestras.



## ANEXO 5. CUESTINARIO PARCIAL DE RAZONAMIENTO ESTADÍSTICO (PARES4) . REGRESIÓN LINEAL Y ANÁLISIS DE VARIANZA

1. Dos alumnos lleva a cabo un estudio para la clase de estadística en la que analizan la correlación existente entre la estatura de los alumnos de su grupo de clase y el perímetro de su cráneo. Alberto toma la medida de la estatura y del perímetro craneal en centímetros y María mide la estatura en metros y el perímetro craneal en centímetros.

Alberto encuentra una correlación de 0,78 entre la estatura y el perímetro del cráneo. ¿Cuál debería ser la correlación encontrada por María?

- La correlación de María debería ser 1, porque encajará exactamente con la de Alberto.
- La correlación de María debería ser  $0.78 / 100 = 0.078$  porque es necesario cambiar las unidades de metro a centímetro. 1 metro=100 cm.
- La correlación de María debería ser 0.78, la misma que la de Alberto.**

### **Resultado de aprendizaje 1:**

**Entender que el coeficiente de correlación es independiente de las unidades de medida de las variables.**

2. La correlación entre el peso corporal y la circunferencia de la cintura en un grupo numeroso de varones adultos trabajadores de una empresa es de 0.75. Qué porcentaje de variabilidad en la circunferencia de la cintura no es explicado por el peso?.

- $1 - 0.75 = 0.25$  ó 25%
- $(0.75)^2 = 0.5625$  ó 56.25%
- $1 - (0.75)^2 = 0.4375$  ó 43.75%**
- $(1 - 0.75)^2 = 0.0625$  ó 6.25%

### **Resultado de aprendizaje 2:**

**Relacionar el coeficiente de correlación con el coeficiente de determinación y con el porcentaje de variabilidad común entre dos variables.**

3. Un estudiante realizó un estudio para la clase de estadística acerca de la relación entre las horas de estudio semanales de cada estudiante y lo que gasta semanalmente en entretenimiento. Basándose en una muestra de 300, calculó un coeficiente de correlación  $r = -0.023$  para estas dos variables.

¿Cuál de las siguientes afirmaciones es una interpretación apropiada e los resultados?

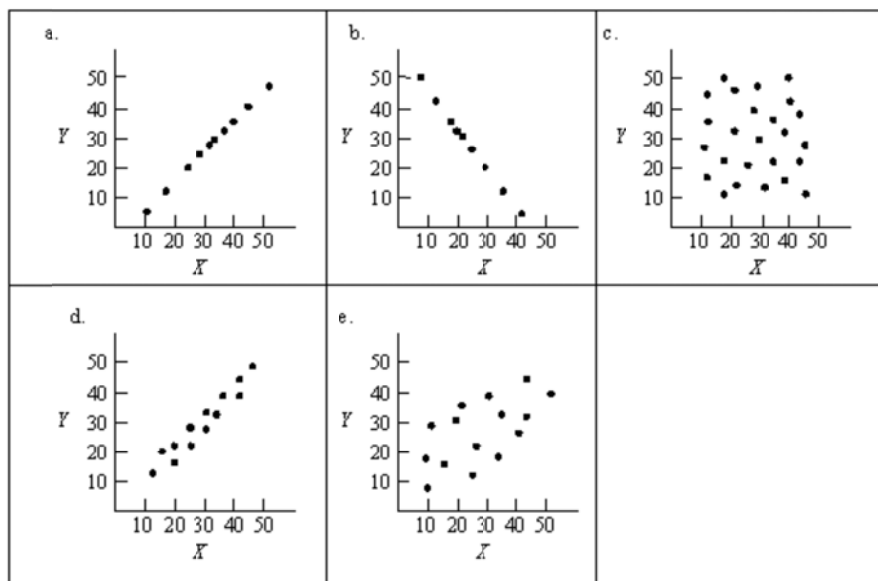
- Esta baja correlación de  $-0.023$  indica que no hay relación entre las horas que dedica al estudio y el gasto en entretenimiento.
- No hay relación lineal pero podría existir una relación no lineal.**
- Hay una relación lineal entre las dos variables, de forma que cuantas más horas dedica al estudio, gasta menos en entretenimiento.

**Resultado de aprendizaje 3:**

**Interpretar un coeficiente de correlación muy pequeño y distinguir que no existe relación lineal entre las variables pero puede existir una relación de otro tipo entre ellas.**

Las preguntas 4, 5 y 6 se refieren a la siguiente situación:

Considera los cinco diagramas de dispersión que se muestran más abajo:



4. Selecciona el diagrama de dispersión que muestra una correlación de cero.

- a) a
- b) b
- c) c
- d) d
- e) e

**Resultado de aprendizaje 4:**

Reconocer un gráfico de dispersión que representa dos variables que no están relacionadas linealmente.

5. Selecciona el diagrama de dispersión que muestra una correlación aproximada de 0.60.

- a) a
- b) b
- c) c
- d) d
- e) e

**Resultado de aprendizaje 5:**

Seleccionar entre varios ejemplos, el diagrama de dispersión que muestra la representación que más se aproxima a un valor dado de correlación entre las variables.

6. Selecciona el diagrama de dispersión que muestra la relación más fuerte entre las variables X e Y.

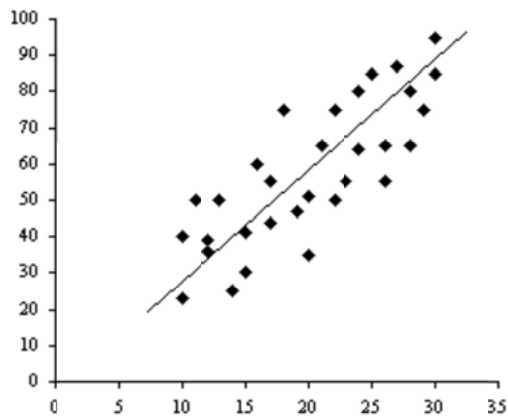
- a) a
- b) b
- c) a y b
- d) a y d
- e) a, b y d

**Resultado de aprendizaje 6:**

Identificar entre varios ejemplos, el diagrama de dispersión que muestra la relación más fuerte entre dos variables.

Las preguntas 7 y 8 se refieren a la siguiente situación:

Un profesor de estadística realiza el siguiente diagrama de dispersión y recta de regresión para ver si las puntuaciones de los exámenes de sus alumnos pueden predecirse por sus puntuaciones en un test básico de habilidad matemática.



7. ¿Qué representan los números en el eje horizontal?

- a. Puntuaciones del examen de estadística.
- b. El número de personas que obtienen cada puntuación de examen.
- c. La variable respuesta.
- d. **Las puntuaciones de habilidad matemática.**

**Resultado de aprendizaje 7:**

Indicar qué está representado en cada uno de los ejes de un gráfico de dispersión.

8. ¿Qué representan los números en el eje vertical?

- a. **Puntuación de los exámenes de estadística.**
- b. El número de personas que obtienen cada puntuación de examen.
- c. Las puntuaciones de habilidad matemática.

**Resultado de aprendizaje 8:**

Indicar qué está representado en cada uno de los ejes de un gráfico de dispersión.

9. En un estudio se recogen los valores de consumo de metilmercurio ( $\mu\text{g Hg/día}$ ) y la concentración total de mercurio en la sangre ( $\text{ng/g}$ ) de un grupo de individuos expuestos al metilmercurio por consumir peces contaminados. Se encontró un coeficiente de correlación de 0.80 entre el consumo de metilmercurio y la concentración de mercurio en sangre. ¿Cual de las siguientes afirmaciones es la mejor interpretación de esta correlación?

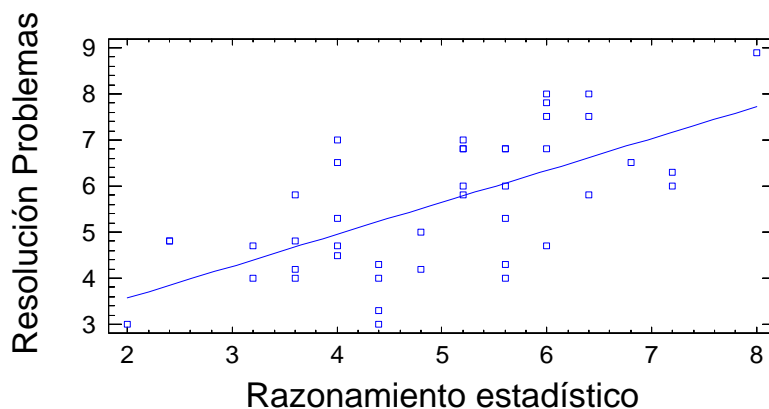
- Las personas que consumen más peces contaminados, tienen mayores valores de mercurio en sangre.**
- No hay una relación lineal significativa entre el consumo de peces contaminados y la cantidad de mercurio en sangre.
- La cantidad de mercurio en sangre no puede ser explicada por el consumo de peces contaminados.
- La cantidad de mercurio en sangre disminuye en las personas que consumen peces contaminados porque es capturado por el metilmercurio.

**Resultado de aprendizaje 9:**

Interpretar la magnitud de un coeficiente de correlación positivo, reconociendo que los sujetos tienen simultáneamente valores altos o bajos en las dos variables.

**Las preguntas 10 y 11 se refieren a la siguiente situación:**

La profesora de estadística evaluó a sus estudiantes con una prueba de razonamiento estadístico y otra de resolución práctica de problemas con un programa informático. La profesora está interesada en determinar si hay una relación entre las puntuaciones del ambos exámenes, de manera que analizó los datos obtenidos, construyó el siguiente diagrama de dispersión y calculó el coeficiente de correlación.



10. ¿Cuál de las siguientes afirmaciones es la mejor interpretación del diagrama de dispersión?

- a. **Hay una correlación positiva moderada entre las puntuaciones del examen de razonamiento estadístico y las del ejercicio de resolución de problemas.**
- b. No hay correlación entre las puntuaciones del examen de razonamiento estadístico y las del ejercicio de resolución de problemas.
- c. Las puntuaciones de todos los alumnos en el examen de resolución de problemas fue mayor que en la prueba de razonamiento estadístico.

**Resultado de aprendizaje 10:**

**Extraer información de un diagrama de dispersión acerca del grado de asociación entre dos variables.**

11. Localiza el punto que muestra una puntuación de 9 en el examen de resolución de problemas. Este punto representa la puntuación de Antonio y es incorrecta. Si la puntuación de Antonio se borra de la base de datos, ¿cómo se vería afectado el coeficiente de correlación?

- a. **El valor de la correlación disminuiría.**
- b. El valor de la correlación aumentaría.
- c. El valor de la correlación se mantendría igual.

**Resultado de aprendizaje 11:**

**Considerar que la eliminación de un solo sujeto, especialmente si tiene puntuaciones extremas, puede modificar el coeficiente de correlación.**

12. Una enfermera de atención primaria quiere realizar un estudio sobre consumo alimenticio y su relación con otros factores. Analizó el colesterol de los usuarios de su consulta y reunió datos de consumo de grasa de origen animal, encontrando una relación lineal con el siguiente modelo de regresión:

$$COL = 83,3312 + 5,50671 * GRAANI$$

Un compañero le preguntó acerca de la posibilidad de predecir el colesterol de un paciente que tiene un consumo elevado de 100 g de grasa de origen animal usando esta ecuación. ¿Cuál sería de las siguientes, la respuesta más acertada?

- Traza la recta de regresión, encuentra el 100 en el eje horizontal y comprueba el valor correspondiente en el eje y.
- Sustituye por 100 la Grasa animal en la ecuación y calcula el Colesterol.
- Ambos métodos son correctos.**
- Ninguno de estos métodos es correcto.

**Resultado de aprendizaje 12:**

Utilizar la ecuación y la recta de regresión para predecir el valor de una variable.

**Las preguntas 13 y 14 se refieren a la siguiente situación:**

Se analiza la influencia del tipo de parto, (parto eutócico, forceps, ventosas y espátulas) sobre el ph arterial medido en el cordón umbilical del recién nacido.

Los resultados del análisis son los siguientes:

Tabla ANOVA para Ph Art\_Umbilical según Tipo Parto

Análisis de la Varianza					
Fuente	Sumas de cuad.	Gl	Cuadrado Medio	Cociente-F	P-Valor
Entre grupos	2,10401E7	3	7,01335E6	16,96	0,0000
Intra grupos	2,03499E8	492	413616,0		
Total (Corr.)	2,24539E8	495			

13. ¿Qué indican estos resultados?

- Que el pH arterial varía en todos los tipos de parto.
- Que en algún tipo de parto el pH arterial es diferente al resto.**
- Que el pH es igual en los cuatro tipo de partos.

**Resultado de aprendizaje 13:**

Seleccionar la interpretación correcta de un análisis de varianza, utilizando los datos recogidos en una tabla de resultados y relacionándolo con las hipótesis de este tipo de análisis.

14. La fuente de variabilidad “intra grupos” analiza:

- Las diferencias en el pH entre los 4 tipos de partos.
- La variabilidad en el pH explicada por el tipo de parto.
- La variabilidad en el pH no explicada por el tipo de parto.**
- Todo lo anterior es correcto.

**Resultado de aprendizaje 14:**

**Reconocer e interpretar las fuentes de variabilidad de un análisis de varianza en el contexto de un ejemplo concreto.**

15. Mediante la aplicación de tres tipos de radiaciones se consiguen los siguientes resultados en la reducción de la masa en tumores de ciertos pacientes oncológicos.

Radiaciones	Reducción de masa (mg)									
A	210	100	150	125	100	110	210	150	125	
B	145	247	125	250	260	245	247	125	250	
C	177	300	206	150	160	177	200	106	150	

¿Cuál de las siguientes preguntas pueden responderse con un análisis de varianza?

- ¿Hay diferencias significativas entre los tres tipos de radiaciones, en cuanto a su eficacia al disminuir la masa del tumor?
- ¿Cuál de las radiaciones disminuye más la masa tumoral?
- ¿Alguna de las radiaciones tiene un efecto mayor sobre la disminución de la masa tumoral?
- Todas las anteriores.**
- Ninguna de las anteriores.

**Resultado de aprendizaje 15:**

**Distinguir las respuestas que puede aportar un análisis de varianza en el contexto de un ejemplo concreto.**



## ANEXO 6. PRUEBA DE RAZONAMIENTO ESTADÍSTICO CON RESPUESTA ABIERTA PARCIALMENTE RESTRINGIDA

**Indicación para el alumno:** La siguiente prueba ha de contestarse en los espacios destinados a la respuesta

1º Un grupo de investigadores quieren saber si la administración de anestesia epidural modifica el pH de la arteria del cordón umbilical del recién nacido. Utilizan para el estudio un grupo de recién nacidos a cuyas madres se les administro anestesia epidural y otro grupo de recién nacidos de madres a las que no se les administró anestesia epidural.

Las medias de pH en ambos grupos son:

CON ANESTESIA  
EPIDURAL  
pH medio= 7,249  
Desv. típica= 0,0697

SIN ANESTESIA  
EPIDURAL  
pH medio = 7,283  
Desv. típica= 0,0709

a) ¿Qué tipo de análisis te parece adecuado para el estudio que pretende este grupo de investigadores?

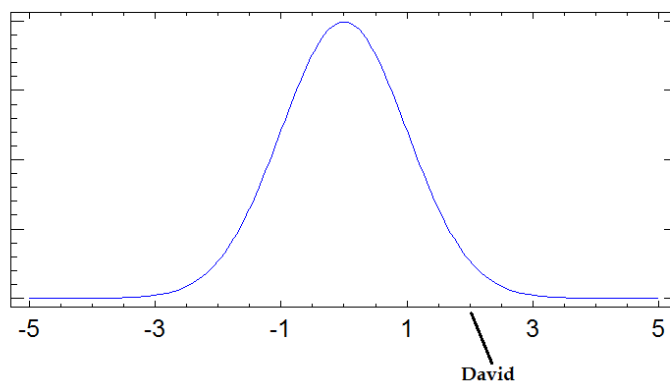
b) ¿Cuáles son las hipótesis? Plantea las hipótesis y describe lo que se está contrastando con cada una de ellas.

2° Si el **pH de la arteria** del cordón umbilical es una variable de **distribución normal**.

Representa el gráfico de la distribución de datos de ambos grupos señalando los parámetros que las caracterizan

3° Cuál de las anteriores gráficas es más apuntada y qué significa en el contexto del estudio que se ha descrito.

4° David es un recién nacido a cuya madre NO le administraron anestesia epidural



MUESTRA Sin anestesia epidural

Con la información del gráfico, escribe 5 conclusiones/interpretaciones sobre el pH de David en relación a su propio grupo (niños con madre sin anestesia epidural)

1.

2.

3.

4.

5.

5° Entre la información que ha recogido este grupo de investigadores hemos encontrado los siguientes datos:

	Parto Eutócico	Parto con Fórceps	Parto con Ventosa	Parto con Espátula	
Sin anestesia epidural	386	4	13	0	403
Con anestesia epidural	470	62	41	23	596
	856	66	54	23	999

Han aplicado un análisis con el que han obtenido un p-valor = 0.00 Enumera cinco conclusiones/interpretaciones de los resultados de dicho análisis.

1.

2.

3.

4.

5.

## CLAVE DE CORRECCIÓN DE LA PRUEBA DE RAZONAMIENTO ESTADÍSTICO DE RESPUESTA ABIERTA PARCIALMENTE RESTRINGIDA

Calificación máxima de la prueba: 6 puntos.	
<b>Pregunta 1 a</b>	<b>Resultado de aprendizaje:</b> Elección de una prueba adecuada a un problema estadístico planteado. <b>Calificación:</b> 1 punto
<b>Respuesta correcta:</b>	
<b>Opciones válidas</b>	
	Contraste de hipótesis para la diferencia de medias (t-student). Muestras independientes.
	Contraste de hipótesis para la comparación de dos medias (t-student). Muestras independientes.
	Comparación de dos medias (t-student ). Muestras independientes.
<b>Pregunta 1 b</b>	<b>Resultado de aprendizaje:</b> Enunciar correctamente las hipótesis de la comparación de dos medias en el contexto planteado. <b>Calificación:</b> 1 punto
<b>Respuesta correcta:</b>	
	<b>Plantea las hipótesis</b>
	$H_0 \quad \bar{X}_{\text{epidural}} = \bar{X}_{\text{sin epidural}}$
	$H_a \quad \bar{X}_{\text{epidural}} < \bar{X}_{\text{sin epidural}}$
	<b>Describe las hipótesis</b>
	$H_0$ En el enunciado de esta hipótesis debe expresar que <b>no existen diferencias</b> entre las medias de los grupos, haciendo alusión al concepto <b>significativo</b> o a que las <b>diferencias observadas pueden ser asumidas por el azar (por factores aleatorios)</b> .
	$H_a$ deberá expresar que <b>existen diferencias</b> entre las medias de los grupos,

especificando la dirección de las diferencias, haciendo alusión al concepto **significativo** o a que las **diferencias observadas no pueden ser explicadas por el azar (por factores aleatorios)**.

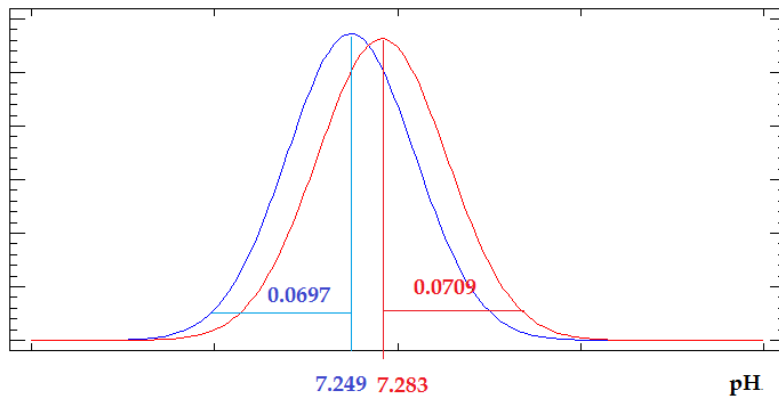
### Pregunta 2

**Resultado de aprendizaje:** Representar gráficamente una variable normal y localizar en ella los parámetros que la caracterizan.

#### Respuesta correcta:

Dibuja las dos curvas normales en el mismo eje de la variable pH y localiza la media y la desviación típica.

**Calificación** = 1 punto



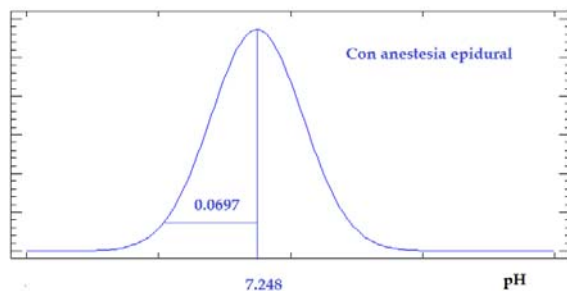
Con anestesia epidural

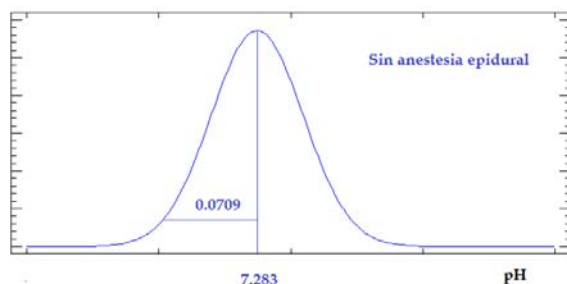
Sin anestesia epidural

#### Respuesta correcta:

Dibuja en dos gráficos separados las curvas correspondientes a cada muestra, y localiza la media y la desviación típica en cada una de ellas.

**Calificación:** 0,5 puntos





### Pregunta 3

**Resultado de aprendizaje:** Reconocer la distribución de datos que tiene menor variabilidad o dispersión identificando las características de la curva normal y asociarla con el valor de la desviación típica.

**Respuesta correcta a la primera parte de la pregunta:** La curva más apuntada corresponde a los valores de pH arterial del cordón umbilical de los niños de madres que recibieron anestesia epidural.

**Calificación:** 0,5.

**Respuesta correcta a la segunda parte de la pregunta:**

**Calificación:** 0,5

#### Cualquiera de las siguientes opciones

La curva es más apuntada porque es menor la desviación típica de los valores de pH de esta muestra.

Los valores de Ph se parecen más a la media que los de la otra muestra.

Los valores de Ph están menos dispersos que los de la otra muestra.

Los valores de pH tienen menor variabilidad que los de la otra muestra.

### Pregunta 4

**Resultado de aprendizaje:** Comprensión del concepto de puntuación típica y su interpretación utilizando las características de la distribución de una variable normal.

**Calificación máxima:** 1 punto

Respuestas correctas	Calificación
David se encuentra a una distancia de dos desviaciones típicas del valor medio de pH de su grupo.	0.2
La puntuación típica del pH de David es 2.0.	0.2
El 97.72% de los niños de su grupo tienen valores de pH	0.2

inferiores al suyo. La probabilidad de que un niño de su grupo elegido al azar tenga un pH inferior al suyo es de 0,9772	
El 2.28% de los niños de su grupo tienen valores de pH por encima del suyo. La probabilidad de que un niño de su grupo elegido al azar tenga un pH superior al suyo es de 0,228.	0.2
El valor de pH de David es 7,424.	0.2
En relación al pH, David se encuentra cercano al percentil 97 de su grupo.	0.2

**Pregunta 5**

**Resultados de aprendizaje:** Reconocer un análisis chí cuadrado de independencia de caracteres y conocer las hipótesis de dicho análisis.

Interpretar el p-valor.

Extraer información de la tabla de contingencia respecto a la relación entre las dos variables.

Utilizar las frecuencias relativas para comparar los distintos grupos en los que se dividen las variables.

**Calificación máxima:** 1 punto

Respuestas correctas	Calificación
Está relacionada la aplicación o no aplicación de anestesia epidural con el tipo de parto que se realiza.	0.2
Hay más partos con fórceps en el grupo en el que a las madres se les administra anestesia epidural. (O el planteamiento inverso respecto al otro grupo).	0.2
Hay más partos con ventosas en el grupo en el que a las madres se les administra anestesia epidural. (O el planteamiento inverso respecto al otro grupo).	0.2
Hay más partos con espátulas en el grupo en el que a las madres se les administra anestesia epidural. (O el planteamiento inverso respecto al otro grupo).	0.2
Hay más partos eutócicos (normales) en el grupo en el que a las madres NO se les administra anestesia epidural. (O el planteamiento inverso respecto al otro grupo).	0.2
El número de mujeres que precisaron algún tipo de ayuda	



---

durante el parto (fórceps, ventosa, espátulas) es significativamente mayor en el grupo de mujeres a las que se les administró anestesia epidural.	0.2
La mayoría de las madres dieron a luz de manera natural, independientemente de que se les hubiera administrado anestesia epidural o no.	0.2



**ANEXO 7. CUESTIONARIO ALUMNOS ESTADÍSTICA (ON-LINE)**

Los datos que se recojan de este cuestionario se utilizarán para la realización de ejercicios en las clases de estadística, trabajando con datos reales sobre información del propio grupo. Es absolutamente anónimo, no es posible conocer la identidad de la persona que ha cumplimentado el cuestionario. Os pido la máxima colaboración. Muchas gracias

\*Obligatorio

Edad numérico en años

Peso numérico en Kg, sin decimales.

Circunferencia de la cintura mídete la cintura a la altura del ombligo y regístralo numéricamente en centímetros

número de hermanos

0   1   2   3   4   5   6   7   8   9   10  
                             

Sexo

- mujer
- hombre

Nota media de bachillerato sin decimales, redondeamos a la cifra más próxima.

Lanza una moneda. Por favor hazlo de verdad no inventes el resultado, han de ser datos de un proceso aleatorio

- cara
- cruz

Ultima cifra de tu número de teléfono Sólo el último número

Número de pie que calzas se refiere al número de talla de zapatos que usas. Lo registramos numéricamente. Por ejemplo, 38,39,40..

Talla de pantalones que usas Lo registramos numéricamente. Por ejemplo 36,38,40...

Padres Zurdos La variable la registramos de esta forma: 0= ninguno es zurdo, 1= uno de ellos es zurdo, 2= los dos son zurdos

- 0
- 1
- 2

¿Eres diestro o zurdo? D= diestro, Z= zurdo

- D
- Z

¿Cómo cruzas los brazos? D= Al cruzar los brazos dejas tu mano izquierda visible, Z=al cruzar los brazos dejas tu mano derecha visible.

- D
- Z

Rular la lengua. Se trata de saber si puedes doblar la lengua haciendo un rulo en longitud. 1= puedo hacerlo, 0=no puedo hacerlo

- 0
- 1

Padres Profesionales de Enfermería 0=ninguno es Enfermero, 1= uno de ellos es Enfermero, 2= los dos son Enfermeros

- 0
- 1
- 2

¿Tienes alguna alteración visual? Si tienes varias señala la que más afecta a tu visión o la de mayor graduación

- ninguna
- miopía
- astigmatismo
- hipermetropía
- otras

¿Tienes perro?

- Sí
- No

Número de libros que lees al año. Es una cifra aproximada. lo registramos numéricamente. Por ejemplo: 0, 5, 12...

Videojuegos en los que eres experto. Es una cifra aproximada. lo registramos numéricamente. Por ejemplo: 0, 2, 3,...

Número de redes sociales en las que tienes cuenta.

- 0   1   2   3   4   5   6   7   8   9   10
-

### Red social que más utilizas

- Tuenti
- Facebook
- Twitter
- MySpace
- Otra

Si en la pregunta anterior señalaste otra red social, escribe su nombre utilizando todas las letras mayúsculas.

Tiempo semanal utilizado en las redes sociales. Es una estimación lo registramos en horas, sin fracciones. Por ejemplo, 0,1,2,3...

### Tipo de bebida alcohólica que consumes más

- no consumo alcohol
- cerveza
- vino
- combinados de licores de alta graduación

Cálculo de gramos de alcohol consumidos un día de fin de semana. Expresamos numéricamente el resultado del siguiente cálculo. sin decimales, redondeando a la cifra más próxima. \* Multiplicamos la graduación de alcohol de la bebida consumida por el número 8, sea cual sea la bebida, y el resultado se multiplica por los gramos consumidos expresados en litros. Ejemplos: un vaso de 250ml de cerveza (4,5 grados)  $4.5 \times 8 = 36$  y  $36 \times 0.2 = 9$ . Una copa de vino tiene 100ml y 13 grados y un licor fuerte tiene un promedio de 30 grados y un combinado tiene unos 100ml. Suma todos los gramos calculados y anótalo en la casilla.

¿Eres fumador? 0=no, 1=sí

- 0
- 1

Busca en Youtube el siguiente vídeo: Ara Malikian Gipsy song expresa tu opinión sobre la música de este video. 0= no me gusta, 1= me gusta

- 0
- 1

Dificultad percibida sobre el bloque de contenidos de estadística Os pido que valoréis la dificultad que pensáis que va a tener para vosotros el bloque de estadística (sólo el módulo que imparte la profesora Soledad Ferreras). 0= poca a 10=mucha

0   1   2   3   4   5   6   7   8   9   10  
                             

¿cuántas horas duermes diariamente?

1   2   3   4   5   6   7   8   9   10  
                          

¿Con cuál equipo de futbol simpatizas?

- R. Madrid
- At. Madrid
- Barcelona
- Otro
- Ninguno

Qué grado de riesgo para la salud le otorgas a un consumo intensivo de alcohol en el fin de semana

- Alto
- Moderado
- Bajo

Qué grado de riesgo para la salud le otorgas a un consumo semanal de cannabis

- Alto
- Moderado
- Bajo

Qué calificación te otorgas como estudiante. Ten en cuenta tu constancia, tu organización, tus apuntes, tus resultados académicos...

0 1 2 3 4 5 6 7 8 9 10

Nivel de ansiedad o angustia con el que abordas el bloque de contenidos de Estadística

0 1 2 3 4 5 6 7 8 9 10

En una escala de 0 a 10 señala el nivel de conocimientos que tienes de matemáticas

0 1 2 3 4 5 6 7 8 9 10



---

De 0 a 10 señala cuál es tu motivación para obtener una buena calificación en esta asignatura

0 1 2 3 4 5 6 7 8 9 10

Enviar

**ACLARACIONES PARA COMPLETAR CORRECTAMENTE LA BASE DE DATOS**

- Edad: en años
- Peso: en gramos
- Circunferencia de la cintura
- Mídete la cintura a la altura del ombligo y regístralo centímetros
- Número de hermanos. 1,2,3,4...etc
- Sexo. V y M con mayúsculas  
V varón  
M Mujer
- Lugar entre hermanos: es el lugar que ocupas entre tus hermanos.  
1,2,3,4....etc
- Nota media de bachillerato: Los decimales se registran con un punto.  
Ejemplo 5.6
- Lanza una moneda. Por favor hazlo de verdad no inventes el resultado,  
han de ser datos de un proceso aleatorio.
- C (cara)  
X (cruz)
- Última cifra de tu número de teléfono. Sólo el último número
- Nº de calzado. Es el número de zapatos que usas
- Talla de ropa. La registramos con número 38, 40, 42...etc
- Padres zurdos. La variable la registramos de esta forma:  
0 si ninguno es zurzo  
1 si uno de ellos es zurzo  
2 si los dos son zurzos
- Diestro o zurzo: D y Z mayúsculas.  
D (si eres diestro)  
Z (si eres zurzo)
- Cruzas los brazos a derechas o izquierda



- D Al cruzar los brazos dejas tu mano izquierda visible
- Z Al cruzar los brazos dejas tu mano derecha visible

-Rular la lengua. Se trata de saber si puedes doblar la lengua haciendo un rulo en longitud como indica la foto.



- 1 puedes hacerlo
- 0 no puedes hacerlo

-Padres Enfermeros.

- 0 si ninguno de tus padres se dedica a la Enfermería
- 1 si uno de ellos tiene como profesión enfermero
- 2 si los dos son de enfermeros

-Alteración visual  
ninguna alteración  
miopía  
astigmatismo  
hipermetropía  
otros

- Tienes perro/os
- 0 no tienes
  - 1 si tienes

-Nº de libros leídos al año. Es una cifra aproximada....

- Videojuegos en los que eres experto.  
0, 1 , 2, 3 ....etc..

- Nº de redes sociales en las que tienes cuenta y utilizas( Facebook, tuenti...etc)  
Se registrará con numeros 0, 1, 2, 3...etc.

- Red social que más utilizas. Escríbela como se indica a continuación pero con letras mayúsculas.

**Facebook, Orkut, Identi.ca, Twitter**  
**Viadeo, Xing , Linked In.**  
**Wipley, Minube Dogster, Last.FM , Moterus.**  
**Yuglo, Unience, PideCita, 11870**  
**Koornk, Dopplr, Youare , Tuenti**  
**Scribd, Flickr, Bebo, Friendster, Dipity, StumbleUpon, FileRide.**  
**Latitud, Brighkite, Fire Eagle, Skout**  
**Rejaw, Blogger, Kwippy, Plaxo, Bitacoras.com, Plurk**  
**MySpace, Friendfeed, Hi5**

- Tiempo semanal utilizado en las redes sociales y tiempo de estudio lo registraremos en minutos.

- tipo de bebida alcohólica más consumida  
 no consumo alcohol  
 cerveza  
 vino  
 combinados de licores de alta graduación (ron, ginebra..)

- Cálculo de los gramos de alcohol consumidos

Multiplicar la graduación alcohólica de la bebida que has consumido por el número 8, sea cual sea la bebida. El resultado que obtienes, lo multiplicas por la cantidad que has consumido expresada en litros. Veamos ejemplos puntuales.

**Un vaso de 250 ml de cerveza (4,5 grados)**

$4,5 \times 8 = 36$  y  $36 \times 0,25 = 9$ .

**vino. Un vino, generalmente tiene unos 13 grados de alcohol.**

**Una copa de 200 ml**

$13 \times 8 = 104$  y  $104 \times 0,2 = 20,8$ .

**licor fuerte, cuya graduación alcohólica promedio es de 30 grados.**

**Un vaso pequeño, de unos 100 ml.**

$30 \times 8 = 240$  y  $240 \times 0,1 = 24$ .

Suma los gramos y anótalo en la casilla

- Fumador  
 0 no fumador  
 1 fumador

- Videos  
 ¿Te gustan la música de estos videos? Solo valoramos la música, no el video.  
 Regístralo en su casilla correspondiente, hay una casilla para cada uno de los cuatro videos.

Lo registramos como  
 0 no me gusta  
 1 si me gusta

**Nº 1** [http://www.youtube.com/watch?v=ma-whEy\\_Hwg](http://www.youtube.com/watch?v=ma-whEy_Hwg)

**Nº 2** <http://www.youtube.com/watch?v=IAhHNCfA7NI&feature=fvw>

**Nº 3** <http://www.youtube.com/watch?v=mME-URRbak8>

**Nº 4** <http://www.youtube.com/watch?v=D7o7BrIbaDs>

-Dificultad de la estadística.  
1= poca y 10= mucha

Os pido que valoréis La dificultad que va a tener para vosotros el bloque de estadística (sólo el módulo que imparte La profesora Soledad Ferreras)

- Cuantas horas duermes diariamente

- Cuántas horas duermes un sábado

- Con qué equipo de futbol simpatizas

RMadrid  
AtMadrid  
Barcelona  
Otro

-Intención de voto en las próximas elecciones generales

PP  
PSOE  
IU  
UPID  
otros

- Que grado de riesgo para la salud le otorgas a un consumo intensivo de alcohol en el fin de semana

Alto  
Moderado  
Bajo

- Qué grado de riesgo para la salud le otorgas a un consumo semanal de cannabis en adolescentes

Alto  
Moderado  
Bajo

- Qué calificación te otorgas como estudiante. Ten en cuenta tu constancia, tu organización, tus apuntes, tus resultados académicos....