

UNIVERSIDAD DE BUENOS AIRES  
FACULTAD DE PSICOLOGÍA

ESTADÍSTICA - CÁTEDRA II  
PROF. HORACIO F. ATTORRESI

GUÍA DE TRABAJOS PRÁCTICOS

MARÍA SILVIA GALIBERT  
HORACIO F. ATTORRESI  
FACUNDO J. P. ABAL  
MIRTA S. MAURO  
LIDIA I. NICOLAI



FACULTAD DE PSICOLOGIA

# PRÁCTICA I

## EJERCICIO MODELO I

Una comisión integrada por profesionales de la salud, de la educación y padres, de una pequeña localidad del interior está interesada en conocer el nivel de consumo de alcohol en sus adolescentes. Para ello llevaron a cabo una encuesta anónima a 10 varones y 10 mujeres elegidos al azar de distintas zonas de la localidad. Se los interrogaba sobre cantidad y tipo de bebida alcohólica que consumían aproximadamente por semana durante el verano. Luego se convertían a cantidad neta de alcohol etílico.

A continuación se exhibe la lista donde se registra para cada sujeto el número de encuestado, sexo y cantidad aproximada de alcohol consumida semanalmente durante el verano.

Código	Sexo	Consumo semanal de alcohol en litros
1	F	0,200
2	F	0,060
3	M	0,480
4	F	0,000
5	M	0,075
6	F	0,140
7	F	0,100
8	M	0,440
9	M	0,420
10	M	0,700
11	F	0,040
12	F	0,000
13	F	0,100
14	M	0,300
15	M	0,500
16	F	0,050
17	M	0,480
18	F	0,025
19	M	0,000
20	M	0,450

- ¿Cuál es la población de individuos y cuál la de observaciones?
- ¿Quiénes constituyen la muestra de individuos y quiénes la de observaciones?
- Mencione la variable de interés y clasifíquela.
- Mencione qué nivel de medición se utilizó e indique los valores de la escala.
- Indique la posible fuente sistemática de variación que provoque variaciones previsibles en la cantidad de alcohol consumida por los adolescentes.
- Ordene la información dada de manera que aparezca exhibida dicha fuente sistemática.
- Enuncie posibles fuentes fortuitas de variación y distinga entre ellas las que son susceptibles de ser estudiadas como sistemáticas.

## RESPUESTAS

a) La población objeto de estudio es el conjunto de los adolescentes de la localidad donde se llevó a cabo la experiencia.

La población de observaciones es el conjunto de números que representa a las cantidades de alcohol medido en litros correspondientes al consumo semanal aproximado de cada adolescente durante el verano.

b) La muestra de individuos está formada por los 20 adolescentes que fueron escogidos para la encuesta y la de observaciones, por sus correspondientes consumos de alcohol.

c) La variable es la cantidad de alcohol que consume por semana durante el verano un adolescente de esta localidad del interior. Es una variable cuantitativa continua.

d) El nivel de medición es de razón ya que las diferencias entre los números obtenidos por las mediciones tienen un significado: diferencias en el consumo de alcohol y el cero no es convencional, es absoluto porque significa ausencia del consumo (en cualquier unidad que se tome). El rango de valores de la escala son todos los números no negativos.

e) La posible fuente de variación sistemática es el sexo; podría suponerse que los consumos difieren entre los sexos (o, más aún, que los consumos del sexo femenino – F – tiendan a ser inferiores que los masculinos – M – ) Esto puede constituir una hipótesis por probar; si el experimentador lo probara afirmaría efectivamente que el sexo es una fuente sistemática de variación.

f) Ordenando los consumos de menor a mayor e indicando a cuántas mujeres o varones corresponden dichos consumos puede tenerse una primera idea acerca de la verosimilitud de la hipótesis de que el sexo es una fuente de variación.

Consumo semanal de alcohol en litros	Sexo	
	Mujeres	Varones
0,000	2	1
0,025	1	
0,040	1	
0,050	1	
0,060	1	
0,075		1
0,100	2	
0,140	1	
0,200	1	
0,300		1
0,420		1
0,440		1
0,450		1
0,480		2
0,500		1
0,700		1

Los mayores consumos corresponden a los varones. Se observa una “tendencia” de la variable “cantidad de alcohol consumida” a adoptar valores más altos dentro del conjunto de los varones que de las mujeres, lo cual no quiere decir que todas las mujeres consuman menos alcohol que todos los varones: la encuestada 1 consume 0,2 litros mientras que el encuestado 19 no consume alcohol. Esto demuestra que no toda

la variabilidad en el consumo está asociada al sexo hay otras fuentes de variación no registradas en esta experiencia: las fuentes fortuitas. Ellas hacen necesario utilizar una serie de observaciones dentro de cada grupo (varón / mujer) a fin de poder detectar el efecto de la fuente sistemática.

g) Posibles fuentes fortuitas de variación son: edad, clase social, predisposición genética, nivel de autoestima, antecedentes familiares, nivel de receptibilidad de la publicidad, etc. Tal vez podría imaginarse (hágalo) experimentos para probar si tales factores son o no fuentes de variación. Si se prueba, serían fuentes sistemáticas. Posiblemente sea difícil imaginar un experimento para medir en cada persona el nivel de receptibilidad de la publicidad; si no se pudiera medir sería siempre considerada como una posible fuente fortuita de variación.

## EJERCICIO MODELO II

El presente ejercicio se basa en una investigación cuya referencia es: Simón, T., Ruiz Gallego-Largo, T. y Suengas, A. (2009). Memoria y envejecimiento: recuerdo, reconocimiento y sesgo positivo. *Psicothema*, 21, 3, 409-415.

El artículo completo se halla en [www.psicothema.com/pdf/3646.pdf](http://www.psicothema.com/pdf/3646.pdf).

A continuación se extractan algunos párrafos correspondientes a los temas por trabajar en esta unidad indicando si son fragmentados, sintetizados o adaptados. Léalos atentamente e intente resolver la tarea.

### TEXTO

Resumen (fragmento)

El objetivo del estudio es analizar el declive y mantenimiento de distintos aspectos del rendimiento mnémico al envejecer (...). Para ello, se comparó el rendimiento en tareas de recuerdo y reconocimiento de dos grupos de distinta edad (joven, mayor). (...)

Introducción (fragmentada)

El envejecimiento afecta a distintos aspectos de la memoria episódica y bastantes estudios han mostrado que, por ejemplo, el deterioro es superior en el recuerdo que en el reconocimiento, que en diversas circunstancias parece incluso no verse alterado (por ejemplo, Sekuler, McLaughlin, Kahana, Wingfield y Yotsumoto, 2006). También se ha sugerido que uno de los cambios mnémicos al envejecer consiste en un aumento de la probabilidad del recuerdo de la información positiva frente a la negativa (por ejemplo, Singer, Rexhaj y Baddeley, 2007). El estudio que presentamos analiza el rendimiento diferencial en función de la edad al recordar y reconocer información a la que se atribuye distinta carga emocional (...)

El estudio que presentamos investiga una reinterpretación del sesgo positivo, en el sentido de que al envejecer se aprecia más lo que se recuerda. Las valoraciones de unos estímulos neutros nos van a permitir determinar si todos los participantes o sólo los mayores consideran más agradables las imágenes que recuerdan que las que no recuerdan. De acuerdo con la misma argumentación que propone un aumento del control emocional por las ventajas que conlleva de satisfacción personal, podríamos argüir que las personas mayores desarrollan un aumento en el aprecio que conceden a lograr recordar algo con el incremento consecuente del agrado que produce lo recordado, quizá como contrapartida frente a las experiencias de olvidos frecuentes (...)

## Método

### *Participantes* (sintetizado)

En el estudio participaron voluntariamente 44 personas, de las cuales 22 eran adultos jóvenes con edades comprendidas entre 17 y 33 años y 22 adultos mayores con edades comprendidas entre los 65 y 91 años. En cada muestra había sólo un varón.

Los jóvenes eran alumnos del primer curso de Logopedia de la Universidad Complutense de Madrid. Los adultos mayores procedían de la Asociación Altamira y de la Residencia de Nuestra Señora de Montserrat de Madrid y fueron seleccionados aleatoriamente entre quienes los profesionales sanitarios de los centros indicaron que no padecían enfermedades ni deterioro cognitivo patológico, reflejado en una puntuación superior a 26 en la prueba Mini-Mental (Folstein, Folstein y McHugh, 1975; Giménez-Roldán, Novillo, Navarro, Dobato y Jiménez-Zuccarelli, 1997) y en la participación activa de todos ellos en los cursos académicos y culturales organizados por las respectivas instituciones.

### *Materiales* (sintetizado)

Los estímulos consistieron en 100 fotografías en color de elementos representativos de distintas categorías (ropa, animales, alimentos, utensilios cotidianos, medios de transporte y muebles). Las imágenes seleccionadas cumplían el requisito de que tres jueces independientes las consideraran categorialmente representativas, claramente reconocibles, fácilmente etiquetables y puntuaban con 4, en una escala de 1 (muy desagradable) a 7 (muy agradable); de ahí que consideremos que todas ellas eran neutras y sin carga emocional. En la fase inicial se presentaron 50 imágenes y el total de las 100 en la tarea de reconocimiento; en ambos casos se contrabalancearon para presentar un orden diferente a cada participante.

### *Procedimiento*

Las personas mayores realizaron las pruebas de manera individual y los jóvenes en grupos reducidos. En la primera fase presentamos una serie de 50 imágenes, a una velocidad de una cada diez segundos. Para el grupo joven, las imágenes se proyectaron mediante ordenador sobre una pantalla, en tanto que para el grupo mayor, componían las hojas de un cuaderno de tamaño A4. Las instrucciones que recibieron los participantes consistieron, simplemente, en que observaran atentamente las imágenes, sin indicaciones específicas para recordarlas, ni aviso sobre la posterior prueba de recuerdo, ya que se ha demostrado que el conocimiento anticipado de la tarea mnémica mejora el rendimiento de los jóvenes, pero empeora el de las personas mayores (por ejemplo, Barrett, Crucian, Wingard, Graybeal y Heilman, 2003). A continuación, realizaron la prueba de recuerdo libre, en que se pidió a los participantes que nombraran todas las imágenes que recordaran de la serie que acababan de ver. Posteriormente, se realizó la prueba de reconocimiento mediante la presentación de 100 imágenes (50 inicialmente presentadas y 50 distractoras) frente a las que los participantes tenían que indicar si las habían visto previamente o no. Por último, se les volvieron a presentar las 50 imágenes que habían visto inicialmente para que evaluaran el agrado que cada una de ellas les suscitaba en una escala tipo Likert de 1 (muy desagradable) a 7 (muy agradable). Los jóvenes realizaron las tres pruebas (recuerdo libre, reconocimiento y valoración de las imágenes) por escrito, y los mayores oralmente. Las diferencias en el procedimiento utilizado para los dos grupos respondieron al intento de minimizar la fatiga en los participantes de mayor edad.

### *Análisis de datos* (Sintetizado y adaptado)

#### *Recuerdo libre*

El efecto de la edad sobre el recuerdo fue medido como el número de imágenes recordadas, con la edad como un factor con dos niveles (joven, mayor).

Para analizar el efecto mnémico del orden de presentación de las imágenes, comparamos la proporción de recuerdo de los 5 primeros elementos (posición inicial), con la de los 10 centrales (posición central) y los 5 finales (posición final).

#### *Reconocimiento (adaptado)*

Una de las medidas que se tomó en cuenta para evaluar la habilidad de los participantes para discriminar entre las imágenes presentadas y las distractoras se denomina *probabilidad de reconocimiento* y es la razón entre la diferencia de proporciones de aciertos ( $p(A)$ ) con falsas alarmas ( $p(FA)$ ) y la proporción de respuestas que no son falsas alarmas ( $1-p(FA)$ ). Por falsa alarma se entiende que el individuo indica haber visto una lámina en la serie anterior cuando en realidad ésta no formaba parte de la misma. En símbolos, esta medida es:  $[p(A) - p(FA)] / [1 - p(FA)]$ .

#### *Referencias*

- Barett, A., Crucian, G., Wingard, E., Graybeal, L., y Heilman, K. (2003). Age-dependent recall bias for material of internal versus external origin. *Cognitive & Behavioral Neurology*, 16, 160-169.
- Folstein, M.F., Folstein, S., y McHugh, P. (1975). «Mini-Mental State»: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189-198.
- Giménez-Roldán, S., Novillo, M., Navarro, E., Dobato, J., y Jiménez-Zuccarelli, M. (1997). Examen del estado Mini-Mental: propuesta de una normativa para su aplicación. *Revista de Neurología*, 25, 576-583.
- Sekuler, R., McLaughlin, C., Kahana, M., Wingfield, A., y Yotsumoto, Y. (2006). Short-term visual recognition and temporal order memory are both well-preserved in aging. *Psychology and Aging*, 21, 632-637.
- Simón, T., Ruiz Gallego-Largo, T. y Suengas, A. (2009). Memoria y envejecimiento: recuerdo, reconocimiento y sesgo positivo. *Psicothema*, 21, 3, 409-415.
- Singer, J., Rexhaj, B., y Baddeley, J. (2007). Older, wiser, and happier? Comparing older adults' and college students' self-defining memories. *Memory*, 15, 886-898.

### **TAREA**

- a) i) Mencione la hipotética fuente sistemática de variación más directamente relacionada con el objetivo de este estudio y la variable estadística que la representa. Indique qué tipo de variable es y su nivel de medición.  
ii) ¿Cuál es la población de individuos y cuál la de observaciones de esa variable?  
iii) ¿Quiénes constituyen la muestra de individuos y quiénes las de observaciones?
- b) i) Mencione las variables vinculadas con los distintos aspectos del rendimiento mnémico al que alude el objetivo y sus correspondientes representaciones como variables estadísticas.  
ii) Clasifique dichas variables e indique su nivel de medición.
- c) Considere el párrafo de *Materiales*.  
i) ¿Qué fuentes de variación se intentaron evitar en el modo en que se seleccionaron los estímulos?  
ii) ¿Qué hipotética fuente de variación se intentó controlar en el modo en que se administraron los estímulos?

- d) Considere el párrafo de *Procedimiento*.
- i) Mencione la fuente sistemática de variación para la memoria que según Barret et al (2003) interactúa con la edad. ¿De qué manera se tuvo en cuenta esta fuente en el procedimiento?
  - ii) ¿Qué otros factores podrían constituir fuentes sistemáticas de variación que interactúan con la edad y que diferenciaron el procedimiento para ambos grupos?
- e) Enuncie posibles fuentes fortuitas de variación para el rendimiento en las tareas de recordación y distinga entre ellas las que son susceptibles de ser estudiadas como sistemáticas.

## RESPUESTAS

a) i) Todo el estudio focaliza sobre el efecto de la edad en la recordación; por lo que la principal fuente sistemática de variación es la edad. La variable estadística que la representa es la edad dicotomizada en *Mayores* y *Jóvenes*; por lo que es una variable cuasi cuantitativa de nivel ordinal.

ii) Si bien el objetivo del estudio apunta a las personas jóvenes y mayores en general, la población de individuos sobre la que *en rigor* son aplicables las inferencias de este estudio son los jóvenes entre 17 y 33 años del primer curso de Logopedia de la Universidad Complutense de Madrid y los mayores entre 65 y 91 años de la Asociación Altamira y de la Residencia de Nuestra Señora de Montserrat que no padecen enfermedades ni deterioro cognitivo patológico, en su mayoría mujeres en ambos casos. La población de observaciones es el conjunto de valores “*mayor*” o “*joven*” que la edad toma sobre cada uno de los individuos de la mencionada población.

iii) La muestra de individuos está integrada por las 44 personas que participaron de la experiencia y la de observaciones por los 44 respectivos valores “*mayor*” o “*joven*”.

*Nota:* Hemos resaltado la expresión *en rigor* porque es usual en este tipo de investigaciones considerar que las conclusiones pueden extenderse a un universo más amplio del que en rigor está seleccionada la muestra en la medida en que se piense que las limitaciones del mismo no tienen efecto relevante sobre las variables objeto de estudio. Por ejemplo, si se leen las conclusiones de este estudio se verá que en ningún caso está mencionado el hecho de que la mayoría de los participantes sean mujeres o que los jóvenes pertenezcan al primer año de Logopedia. Esto es porque está implícito que estos factores (sexo y nivel y tipo de estudio) no tienen incidencia sobre la Memoria. Sin embargo, siempre que sea posible, conviene delimitar con precisión la población objetivo y tomar una muestra aleatoria de la misma con un diseño que favorezca su representatividad. El modo conveniente de tomar las muestras según los objetivos de la medición corresponde a una subdisciplina de la Estadística denominada Teoría de Muestreo.

b) i) Las variables son las tres siguientes:

*Memoria libre.* Representada por la *cantidad de imágenes recordadas* de una lista de 50 previamente presentada.

*Reconocimiento.* Consiste en reconocer las 50 imágenes ya vistas en una nueva presentación de 100 imágenes donde 50 son distractoras. Esta variable está representada por la *probabilidad de reconocimiento*, calculada como la razón  $[p(A) - p(FA)] / [1 - p(FA)]$  conforme se explica en el apartado de análisis de datos.

*Valoración de las 50 imágenes presentadas inicialmente.* Medida según una escala Likert desde 1 (muy desagradable) hasta 7 (muy agradable).

ii) La *cantidad de imágenes recordadas* es una variable cuantitativa discreta de nivel de razón.

La *probabilidad de reconocimiento* es una variable cuantitativa continua de nivel de razón. Obsérvese que, por cómo está construida esta medida, probabilidad nula significa que no hay diferencia entre las proporciones de verdadero y de falso reconocimiento.

La *valoración* de las imágenes, al hacerse en una escala Likert, es típicamente una variable cuasicuantitativa, de nivel ordinal, ya que no hay una unidad objetiva de medición.

*Nota:* En el análisis estadístico es habitual tratar a los datos provenientes de escalas Likert como si éstas fueran de nivel intervalar porque está implícito que las distancias entre los valores son aproximadamente iguales, aunque no están objetivadas. En otras palabras, se usa el nivel intervalar como una aproximación, ya que en este nivel es posible llevar a cabo análisis estadísticos más potentes y sencillos que en el nivel ordinal. Para que esta aproximación no sea muy grosera se debe tomar la precaución de considerar una cantidad suficiente de categorías para la escala; usualmente se trabaja con al menos 5 categorías. Sin embargo, se han desarrollado procedimientos para poder objetivar las distancias en escalas Likert y así posibilitar un tratamiento intervalar genuino de las mismas. Cañadas Osinski y Alfonso Sánchez Bruno (1991)

c) i) Las fuentes de variación que podrían haber afectado a los resultados confundiendo los efectos que se pretendían estudiar son: la *no representatividad* de las imágenes, la *dificultad en reconocerlas*, la *dificultad en mencionarlas* (etiquetarlas) y la *carga emocional*.

ii) Se intentó controlar el efecto del *orden de presentación de las imágenes*, de modo que éste fuera parejo para todas las imágenes, al rotarlas.

d) i) La fuente de variación a la que se refiere es el *conocimiento anticipado de la tarea mnémica*; ya que puede afectar la recordación a favor de los jóvenes. Este factor se tuvo en cuenta al no dar indicaciones específicas para recordar las imágenes ni dar aviso sobre la posterior prueba de recuerdo.

ii) Otras fuentes de variación que interactúan con la edad podrían ser: 1) La *familiaridad con el medio por el que se presentan las imágenes*. Por eso se presentaron por ordenador para los jóvenes y sobre papel para los mayores. 2) La *fatiga en la resolución de la tarea*. Para evitarla se llevó a cabo de modo oral para los mayores.

e) En este estudio, fuentes fortuitas de variación para el rendimiento en las tareas propuestas pueden ser:

- *Características genéticas* propias de cada persona en su relación con la memoria. En efecto, además de la edad cuyo efecto se estudia en la presente investigación, las personas difieren en memoria también por características genéticas.

- *Disposición* hacia la resolución de este tipo de tareas. La falta de interés o la mala predisposición a ser sometido a un proceso de evaluación puede incidir en el resultado al no poner el debido empeño en la resolución de la tarea.

- *Nivel de actividad intelectual o tipo de profesión/empleo*. Las personas que han llevado una vida intelectual activa, o cierta clase de empleos probablemente tengan su memoria mejor entrenada.

- *Experiencias personales particulares*. Aunque las imágenes se hayan seleccionado aparentemente neutras, las experiencias particulares de cada persona podrían estar asociadas más o menos a algunas de las imágenes presentadas, con lo cual éstas resultarían con distinta carga emocional para cada persona, lo que podría incidir en su recordación.

Podría ser estudiada como una hipotética fuente sistemática de variación el *nivel o tipo de actividad intelectual*. Para ello debería llevarse a cabo la experiencia relatada



en el artículo con personas de la misma edad elegidas de modo de representar distintas actividades: las que completaron estudios universitarios, las que desarrollan oficios que involucran especialmente la memoria (ej. Mozo en Argentina) y las que a lo largo de su vida han desarrollado tareas no tan directamente vinculadas con el estudio o el ejercicio específicamente intelectual (ej. deportistas, empleadas domésticas, entre otras).

## EJERCITACIÓN

### EJERCICIO 1

Enuncie algunas fuentes de variación que puedan afectar la calificación en un parcial de estadística y distinga las que son susceptibles de ser estudiadas como sistemáticas. Explique cómo exhibiría la información si quisiera mostrar que la asistencia a los teóricos es una fuente sistemática de variación.

### EJERCICIO 2

Un equipo de psicopedagogos desea estudiar las habilidades adquiridas en el aprendizaje escolar de la lectoescritura mediante una prueba de rendimiento (en la escala usual de 1 a 10) y el nivel de agresividad (bajo / medio / alto) de niños provenientes de hogares con carencias económicas del Gran Buenos Aires.

Determine la población, las unidades, las características a estudiar y su nivel de medición, las variables estadísticas, su clasificación y sus valores.

### EJERCICIO 3

Clasifique las siguientes variables e indique el nivel de medición de las escalas adoptadas.

- a) Calificación de 8 individuos de un grupo según su posición en rendimiento: 1 al de mayor rendimiento, 8 al de menor rendimiento.
- b) Psicodiagnóstico correspondiente a un paciente según los cuadros clínicos (neurosis, psicosis, etc.)
- c) Nacionalidad de un sujeto.
- d) Cantidad de palabras correctamente leídas por un disléxico en un minuto.
- e) Cociente intelectual (Binet-Stern).
- f) Tiempo que tarda un alumno de Psicología en concluir su carrera.
- g) Temperatura en grados centígrados en Ciudad de Buenos Aires a las 0 hs.
- h) Orden de mérito obtenido por un aspirante en un concurso.
- i) Edad de una persona.

### EJERCICIO 4

Una Consultora de Recursos Humanos está interesada en estudiar el clima laboral de una organización. Con este fin se efectúa una evaluación objetiva al contar la cantidad de “quejas” diarias que brindan 30 de sus 150 empleados.

- a) ¿Cuál es la característica objeto de estudio y cuál es la variable estadística con la que se la operacionaliza?
- b) Clasifique y determine el nivel de medición de la variable.
- c) Determine la población de individuos y muestra de individuos
- d) Determine la población de observaciones y la muestra de observaciones.

## EJERCICIO 5

El Centro de estudiantes de la Facultad de Psicología de la UBA está interesado en conocer el grado de acuerdo de distintos actores de dicha facultad, respecto de la inclusión de la materia Matemática en el CBC de la carrera. Para ello al finalizar el primer cuatrimestre de 2005, tomó una muestra aleatoria de 1500 estudiantes y 400 docentes de la carrera. A continuación se presentan los resultados (no reales):

<i>Grado de acuerdo</i>	<i>Alumnos</i>	<i>Docentes</i>
De acuerdo	20%	40%
Indiferente	10%	30%
En desacuerdo	70%	30%
Total	100%	100%

- Determine quiénes constituyen la población y quiénes la muestra de individuos.
- Determine la población y la muestra de observaciones
- Mencione la variable, clasifíquela e indique la escala de medición.
- Indique la posible fuente sistemática que provoque variaciones previsibles en el grado de acuerdo con la inclusión de la materia Matemática en el CBC.

## EJERCICIO 6

Para comparar la eficiencia de un método de autoinstrucción con el tradicional en la enseñanza del tema “Prueba de hipótesis” (Estadística Inferencial) se han elegido al azar 40 alumnos de la cátedra I de Estadística del primer cuatrimestre de 2009. Fueron asignados aleatoriamente 20 a cada modalidad de enseñanza. Los resultados de su posterior evaluación son:

Calificación	Número de alumnos	
	Método tradicional	Método de autoinstrucción
3	1	
4	5	1
5	7	1
6	4	2
7	3	7
8		5
9		3
10		1

- Lea la información de la tabla.
- Determine quiénes constituyen la población y quiénes la muestra de individuos y de observaciones.
- Mencione la variable estadística y clasifíquela.
- ¿Observa alguna tendencia en las observaciones que den cuenta de la existencia de una fuente sistemática de variación? Si su respuesta es afirmativa indique cuál.
- ¿Todas las calificaciones obtenidas bajo el método de autoinstrucción fueron mayores que las obtenidas bajo el método tradicional? ¿A qué atribuye esto?
- Mencione posibles fuentes fortuitas de variación que afecten la calificación.
- Explique, a partir de d) y e) por qué se acudió a una serie de observaciones en cada grupo para compararlos en lugar de usar un solo representante de cada grupo.

## EJERCICIO 7

Los datos de este ejercicio fueron obtenidos de un largo estudio por Mr. Joseph Raffaele, de la universidad de Pittsburgh, para analizar la comprensión lectora (C) y la rapidez lectora (R), usando los puntajes obtenidos en los subtests del IOWA TEST OF BASIC SKILLS.

Después de seleccionar aleatoriamente 30 estudiantes y dividirlos al azar en 6 submuestras de tamaño 5, los grupos fueron asignados aleatoriamente a dos tratamientos (clases abreviadas y clases no abreviadas) y tres maestros. Los datos son:

		Clases abreviadas		Clases no abreviadas	
		R	C	R	C
Maestro 1		10	21	9	14
		12	22	8	15
		9	19	11	16
		10	21	9	17
		14	23	9	17
Maestro 2		11	23	11	15
		14	27	12	18
		13	24	10	16
		15	26	9	17
		14	24	9	18
Maestro 3		8	17	9	22
		7	15	8	18
		10	18	10	17
		8	17	9	19
		7	19	8	19

- Indique cuáles son las posibles fuentes sistemáticas de variación.
- Mencione las características a estudiar, su nivel de medición, las variables estadísticas y su clasificación.
- Mencione posibles fuentes fortuitas de variación.

## EJERCICIO 8 (Para reflexionar y discutir en conjunto)

En Attorresi et al. (2008) se presenta un cuestionario para medir la actitud altruista. Ésta consiste de 18 ítems con dos opciones: una corresponde a una actitud altruista y la otra no. El puntaje que se asigna al individuo corresponde a la cantidad de respuestas que denotan una actitud altruista.

Suponiendo que Enrique, Luis, y Matías hayan obtenido 12 puntos, 16 puntos y 8 puntos respectivamente:

- Responda justificando las respuestas. ¿Afirmaría que...
  - ... la actitud altruista de Enrique difiere de la de Luis en la misma medida que difiere de la actitud altruista de Matías?
  - ... Luis es el doble de altruista que Matías?
- Mencione la característica objeto de estudio, su nivel de medición, el instrumento de medición y la variable estadística.

### EJERCICIO 9

Al registrarse el puntaje de 5 sujetos en una prueba de razonamiento lógico y el tiempo empleado para resolverla se obtuvieron las siguientes observaciones apareadas:

Puntaje X (sobre un total de 20):	15	13	20	15	18
Tiempo Y (en minutos):	60	55	68	50	65

Verifique si son iguales o no las siguientes expresiones. En caso de ser iguales decidir si lo son para cualquier conjunto de observaciones.

a)	$\sum_{i=1}^5 x_i^2$	y	$\left(\sum_{i=1}^5 x_i\right)^2$
b)	$\sum_{i=1}^5 2y_i$	y	$2\sum_{i=1}^5 y_i$
c)	$\sum_{i=1}^5 x_i \cdot y_i$	y	$\sum_{i=1}^5 x_i \cdot \sum_{i=1}^5 y_i$

### RESPUESTAS

### EJERCICIO 1

Algunas fuentes de variación podrían ser: la capacidad intelectual de cada alumno, el método de estudio, el tiempo dedicado al estudio, el nivel de ansiedad en el momento del examen, el nivel de cansancio con el que el alumno llega al momento del examen, la dificultad propia del examen (en el caso que se tome un examen distinto para cada banda horaria), circunstancias individuales previas al examen, etc. Todas, salvo la última pueden ser estudiadas como sistemáticas.

La asistencia a los teóricos como fuente sistemática de variación debería exhibirse de la siguiente manera:

Calificación	Asistencia a los teóricos	
	Asiste	No asiste
10		
9		
8		
7		
6		
5		
4		
3		
2		
1		

## EJERCICIO 2

La población: los niños escolarizados provenientes de hogares con carencias económicas del Gran Buenos Aires.

Las unidades: cada uno de dichos niños.

Las características a estudiar: “habilidades adquiridas en el aprendizaje de la lectoescritura” y “nivel de agresividad”.

El nivel de medición: en la primera, puede considerarse de razón si la habilidad se representa por el rendimiento entendido como cantidad de respuestas correctas en la prueba. El nivel de medición de la “agresividad” es claramente ordinal por la escala adoptada.

Las variables estadísticas: “calificación en la prueba de habilidades en lectoescritura” y “nivel de agresividad”. La calificación en la prueba es una variable cuantitativa discreta cuyos valores son los números del 1 al 10 (enteros o no según la precisión adoptada) aunque la naturaleza de la característica sea continua. El “nivel de agresividad” es cuasicuantitativa y sus valores son “bajo, medio y alto”.

## EJERCICIO 3

- Cuasicuantitativa, nivel ordinal.
- Cualitativa, nivel nominal.
- Cualitativa, nivel nominal.
- Cuantitativa discreta, nivel de razón.
- Cuantitativa continua, nivel intervalar o Cuasicuantitativa, nivel ordinal (es discutible).
- Cuantitativa continua, nivel de razón (el cero de la escala no tiene por qué ser un valor posible para la variable).
- Cuantitativa continua, nivel intervalar.
- Cuasicuantitativa, nivel ordinal.
- Cuantitativa discreta (pues se considera en años o meses enteros), nivel de razón. El cero es una convención de la definición de “edad” pero no una convención de la escala. En nuestra cultura la edad se define como cantidad de años enteros (o meses para los pequeños) que transcurren a partir del nacimiento.

## EJERCICIO 4

- La característica: clima laboral percibido por los empleados. La variable estadística: cantidad de quejas diarias de los empleados.
- Cuantitativa discreta, de cociente o razón.
- La población de individuos está constituida por el conjunto de empleados de la organización (que en este caso son 150). La muestra de individuos está constituida por los 30 empleados de la organización seleccionados.
- La población de observaciones está constituida por el conjunto de números que representan a la cantidad de quejas diarias de los 150 empleados de la organización. La muestra de observaciones está constituida por el conjunto de números que representan a la cantidad de quejas diarias de los 30 empleados de la organización seleccionados.

## EJERCICIO 5

- Hay dos poblaciones de individuos: una está constituida por el conjunto de estudiantes de la carrera de Psicología de la UBA, y la otra por el conjunto de docentes de la carrera de Psicología de la UBA. Hay dos muestras de individuos: una está formada por los 1500 estudiantes seleccionados de la carrera de Psicología de la UBA y la otra por los 400 docentes seleccionados de la misma carrera.

- b) Hay dos poblaciones de observaciones: una está formada por el conjunto de valores que representan el grado de acuerdo con la inclusión de Matemática en el CBC para la carrera de Psicología entre los alumnos y la otra por el conjunto valores que representan el grado de acuerdo con la inclusión de Matemática en el CBC para la carrera de Psicología entre los docentes. Hay dos muestras de observaciones: una está formada por el conjunto de valores que representan el grado de acuerdo con la inclusión de Matemática en el CBC para la carrera de Psicología entre los 1500 alumnos seleccionados y la otra, por el conjunto valores que representan el grado de acuerdo con la inclusión de Matemática en el CBC para la carrera de Psicología entre los 400 docentes seleccionados.
- c) Variable: Grado de acuerdo con la inclusión de Matemática en el CBC para la carrera de Psicología. Cuasicuantitativa. Nivel de medición Ordinal.
- d) Fuente sistemática de variación: tipo de actor - alumno o docente -.

## EJERCICIO 6

- a) Hay un alumno instruido con el método tradicional que obtuvo una calificación de 3 puntos, cinco alumnos instruidos con el método tradicional obtuvieron una calificación de 4, siete alumnos obtuvieron 7 con el método de autoinstrucción, etc.
- b) La población de individuos: está integrada por la totalidad de los alumnos de la Cátedra I de Estadística que cursaron la materia en el primer cuatrimestre de 2009.  
Muestra de individuos: hay dos muestras de individuos, cada una de 20 alumnos, que son los que participaron de la experiencia.  
Población de observaciones: hay dos poblaciones hipotéticas de observaciones, una está formada por las calificaciones que habrían obtenido los alumnos de la Cátedra I de Estadística del primer cuatrimestre de 2009 en la evaluación sobre prueba de hipótesis, si éstos hubieran sido instruidos por el método tradicional, y la otra población hipotética la integran las calificaciones que habrían obtenido bajo el método de autoinstrucción dichos alumnos.  
Muestra de observaciones: hay dos muestras de observaciones, una, las calificaciones bajo el método tradicional obtenidas por los 20 alumnos de este grupo, y la otra, las calificaciones bajo el método de autoinstrucción obtenidas por los otros 20 alumnos.
- c) La variable estadística es la calificación obtenida por los alumnos en la evaluación acerca del tema “prueba de hipótesis”. Es una variable cuantitativa discreta.
- d) Sí. Las calificaciones tienden a ser mayores para el grupo correspondiente al método de autoinstrucción. Aparentemente el método de instrucción es una fuente sistemática de variación para la calificación. Decimos “aparentemente” porque para llegar a la conclusión utilizamos la simple inspección ocular (más adelante veremos métodos adecuados de decisión estadística).
- e) No, debido a las fuentes fortuitas de variación.
- f) Capacidad individual, preferencia del sujeto a recibir un determinado tipo de instrucción, el docente a cargo de cada uno de los métodos, etc.
- g) Si no existieran fuentes fortuitas de variación y toda la variabilidad en la calificación se debiera al método, entonces todas las calificaciones coincidirían dentro del mismo método y bastaría tomar una observación de cada uno para compararlos. Pero al existir otras fuentes de variación es necesario hacer una “serie de observaciones” para que el efecto de las fuentes fortuitas se vea compensado en el conjunto y deje ver la tendencia en el comportamiento de la variable. Por ejemplo: si se hubiera tomado una sola observación de cada grupo podría haber ocurrido que se eligiera el alumno que sacó 7 con el método tradicional y el que sacó 4 con el de autoinstrucción; no se sabría entonces si estas calificaciones son debidas a que el método tradicional es mejor o a que el azar hizo que el alumno más inteligente recibiera el método tradicional. Pero cuando a cada método se asigna al azar un

conjunto de individuos, es poco probable que todos los más inteligentes queden en uno de ellos y el efecto de la inteligencia se verá compensado en dicho conjunto sin ocultar el efecto de la fuente sistemática.

### EJERCICIO 7

- a) Las posibles fuentes sistemáticas de variación son la modalidad de la clase y el maestro.
- b) Las características por estudiar: “rapidez lectora” y “comprensión lectora”.  
El nivel de medición: de razón para la primera y discutible para la segunda, puede ser ordinal o intervalar (el tratamiento estadístico posterior que de ella hace el autor corresponde al nivel intervalar).  
Las variables estadísticas: “calificación en rapidez lectora” – cuantitativa discreta, aunque la característica sea continua – y “calificación en comprensión lectora” – cuasicuantitativa o cuantitativa discreta según el nivel de medición adoptado - .
- c) Posibles fuentes fortuitas de variación: la capacidad individual, el entrenamiento extra-escolar, el nivel cultural de los estudiantes, etc.

### EJERCICIO 8

- a) 1) y 2) No parece realista cuantificar la actitud altruista con tal precisión que permita dar una métrica de la diferencia en altruismo entre uno y otro individuo.
- b) La característica es la actitud altruista.  
Nivel de medición: más propiamente ordinal, pues tiene sentido pensar que una persona es más o menos altruista que otra, inclusive podría decirse “mucho más” o “mucho menos” altruista; es decir, podría pensarse en la proximidad o distancia entre las personas en cuanto a su actitud altruista en términos “ordinales” pero cuantificarla con la precisión que supone el nivel intervalar parece muy pretencioso. Parece más razonable interpretar los puntajes como rangos que como cantidades.  
El instrumento de medición: es el cuestionario.  
La variable estadística: es la “cantidad de respuestas que denotan la actitud altruista”. Obsérvese que si se tomara esta variable como objeto de estudio en sí misma sin remitirnos al significado en términos de la característica que representa, correspondería al nivel de razón, ya que es una variable intrínsecamente cuantitativa (cantidad de respuestas...) y el cero no es arbitrario, significa que el individuo no dio ninguna respuesta que denotara una actitud altruista. Si a la variable se le diera un tratamiento estadístico correspondiente al nivel de razón habría que tener en cuenta que todo lo que se concluye se referirá al puntaje y podría interpretarse en términos de actitud altruista sólo en tanto y en cuanto se considere que dicho puntaje refleja con suficiente precisión a la característica en cuestión (actitud altruista).

### EJERCICIO 9

- (a) y (c) Son distintas.
- (b) Siempre son iguales. Por propiedad distributiva.

### REFERENCIAS

- Attorresi, H., Abal, F., Lozzia, G., Picón Janeiro, J., Aguerri, M. y Galibert, M. (2008). Construcción de una escala para la medición del Altruismo. *Investigando en Psicología*, 10, 7 – 24.
- Cañadas Osinski y Alfonso Sánchez Bruno (1998). Categorías de respuesta en escalas tipo Likert. *Psicothema*, 10, 3, 623 – 631.

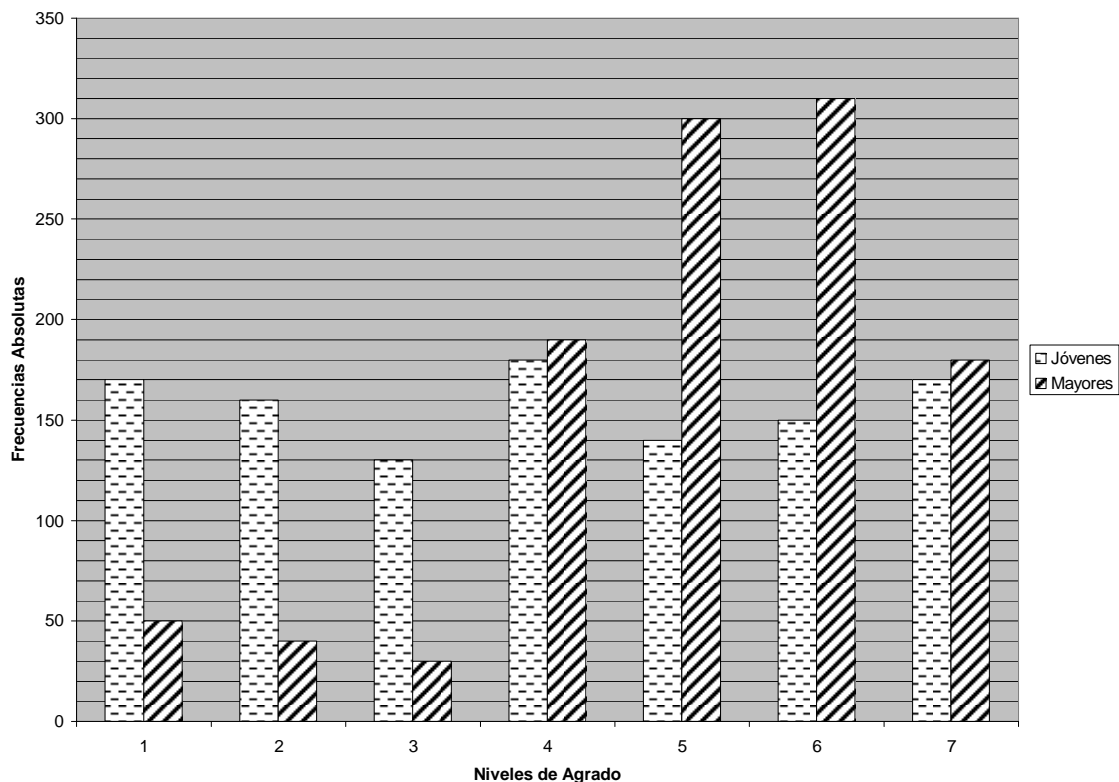
# PRÁCTICA II

## EJERCICIO MODELO

Para analizar si jóvenes y mayores utilizaban diferencialmente los valores de 1 (muy desagradable) a 7 (muy agradable) al juzgar el agrado suscitado por las imágenes, Simón et al (2009) muestran la distribución de frecuencias con las que fueron elegidos estos valores para cada grupo. Tenga en cuenta que cada persona evaluó 50 imágenes y cada grupo consistió de 22 personas; por lo que el total de valoraciones por grupo fue de 1100.

Los resultados se exhiben en el siguiente gráfico:

Distribución de la Valoración según la Edad



- Mencione la fuente sistemática de variación y la variable cuya distribución de frecuencias se grafica ¿Cuántas distribuciones de frecuencias están representadas?
- ¿Qué tipo de gráfico se ha empleado para representar la información?
- Compare ambas distribuciones y extraiga alguna conclusión acerca de los resultados.
- A partir del gráfico exhiba para ambos grupos en una misma tabla la distribución de frecuencias absolutas, relativas, porcentuales y sus correspondientes acumuladas.

### RESPUESTAS

- La fuente sistemática de variación es la edad, dicotomizada en jóvenes y mayores. La variable cuyas frecuencias se grafican es el Nivel de Agrado o Valoración de las Imágenes. Como lo que se grafica es la distribución de frecuencias del Nivel de Agrado según la edad, en realidad hay representadas dos



distribuciones (llamadas condicionales): Nivel de Agrado de los Jóvenes y Nivel de Agrado de los Mayores.

- b) Se ha utilizado un diagrama de rectángulos adyacentes.  
 c) En el diagrama se observa una tendencia de las personas mayores a emitir valoraciones más frecuentemente altas que bajas mientras que los jóvenes han emitido valoraciones con similar frecuencia en todos los niveles; es decir que, en general las personas mayores juzgaron las imágenes más agradables que las personas jóvenes.

*Nota:* Los autores atribuyen esto al bienestar que les produce a las personas mayores el haber podido recordar las imágenes; "...quizá los mayores juzguen más agradables las imágenes que de hecho ya habían conseguido recordar que las que habían olvidado. Por tanto, una reinterpretación del sesgo positivo consistiría en afirmar que no se recuerda lo positivo, sino que como se recuerda, se considera positivo."

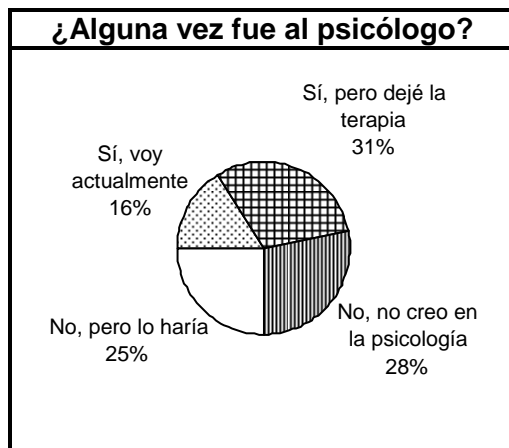
d)

Valo- ración	Frecuencias						Frecuencias Acumuladas					
	Absoluta ( $n_i$ )		Relativa ( $p_i$ )		Porcentual ( $P_i\%$ )		Absoluta ( $n_a$ )		Relativa ( $p_a$ )		Porcentual ( $P_a\%$ )	
	Jóv.	May.	Jóv.	May.	Jóv.	May.	Jóv.	May.	Jóv.	May.	Jóv.	May.
1	170	50	0,1545	0,0455	15,45	4,55	170	50	0,1545	0,0455	15,45	4,55
2	160	40	0,1455	0,0364	14,55	3,64	330	90	0,3000	0,0819	30,00	8,19
3	130	30	0,1182	0,0273	11,82	2,73	460	120	0,4182	0,1092	41,82	10,92
4	180	190	0,1636	0,1727	16,36	17,27	640	310	0,5818	0,2819	58,18	28,19
5	140	300	0,1273	0,2727	12,73	27,27	780	610	0,7091	0,5546	70,91	55,46
6	150	310	0,1364	0,2818	13,64	28,18	930	920	0,8455	0,8364	84,55	83,64
7	170	180	0,1545	0,1636	15,45	16,36	1100	1100	1	1	100	100
	1100	1100	1	1	100	100						

## EJERCITACIÓN

### EJERCICIO 1

El siguiente diagrama circular muestra los resultados de una encuesta de opinión llevada a cabo sobre 30.646 personas de Ciudad de Buenos Aires y Gran Buenos Aires, publicada en el diario Clarín el 26 de octubre de 2005.



- a) Presente la información en una tabla donde se mencione la variable, se muestren sus valores, los ángulos de cada sector y la cantidad de personas que sustenta cada opinión.  
 b) Mencione el nivel de medición de la variable.  
 c) Indique si la escala está bien diseñada en términos de las propiedades que deben caracterizarla. Fundamente la respuesta.

## EJERCICIO 2

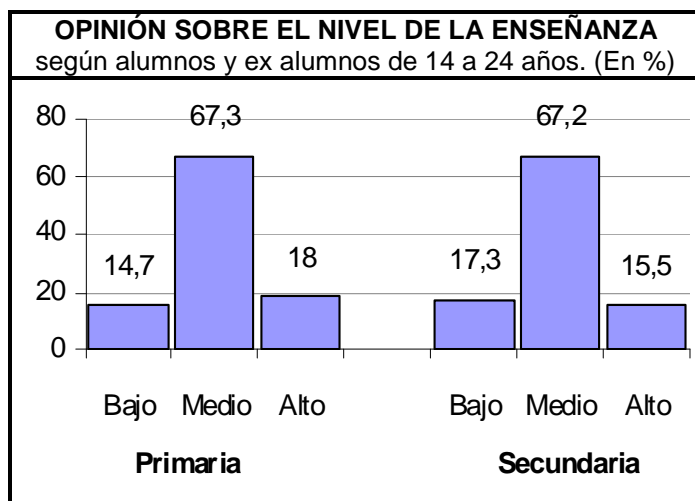
El siguiente cuadro resume los resultados obtenidos en una investigación de Inhelder y Vinh-Bang (Piaget & Inhelder, 1970, pág.16), en la que se estudió la conservación de la sustancia, peso y volumen de acuerdo con cómo son percibidos por niños de 5 a 11 años. En la siguiente tabla se muestran los resultados expresados en porcentajes:

		Edad						
		5	6	7	8	9	10	11
Sustancia	No conservación	84	68	64	24	12	-	-
	Intermedio	0	16	4	4	4	-	-
	Conservación	16	16	32	72	84	-	-
Peso	No conservación	100	84	76	40	16	16	0
	Intermedio	0	4	0	8	12	8	4
	Conservación	0	12	24	52	72	76	96
Volumen	No conservación	100	100	88	44	56	24	16
	Intermedio	0	0	0	28	12	20	4
	Conservación	0	0	12	28	32	56	80

- Mencione la/s variable/s y la fuente sistemática de variación.
- Calcule los ángulos de un diagrama circular correspondiente a la distribución de frecuencias de la variable nivel de conservación del volumen en niños de 10 años.
- Indique el gráfico adecuado para comparar a los niños de 7 y 8 años en la percepción del nivel de conservación de las sustancias.

## EJERCICIO 3

A continuación se exhibe una parte de los resultados de una encuesta de opinión sobre el nivel de la enseñanza primaria y secundaria. (Clarín, 27 / 02 /1994).



Ficha Técnica
- <b>Tipo de encuesta:</b> cuestionario estructurado con preguntas cerradas y abiertas. Las preguntas de este estudio forman parte de un cuestionario más amplio que abarca temas sociales y económicos.
- <b>Tamaño de la muestra:</b> 1019 jóvenes (14 a 24 años) y 1002 adultos (25 años y más)
- <b>Tipo de muestra:</b> probabilística y por cuotas según sexo, edad y ocupación.
- <b>Área geográfica:</b> Ciudad de Buenos Aires y Gran Buenos Aires
- <b>Trabajo de campo:</b> agosto 1992
Fuente: Demoskopia.

- ¿Cuántas distribuciones están representadas? Mencione las variables correspondientes, indique el nivel de medición de la escala utilizada y sus valores.
- Lea la ficha técnica. Vuelque los datos correspondientes a las opiniones sobre la primaria en tablas de distribución de frecuencias absolutas, relativas y porcentuales con sus correspondientes acumuladas. Si lo cree necesario para su ejercitación repítalo para la otra distribución.
- ¿Cuántos encuestados consideran que el nivel de enseñanza secundaria es, a lo sumo, medio?
- ¿Qué porcentaje considera que la enseñanza primaria es, por lo menos, medio?

#### EJERCICIO 4

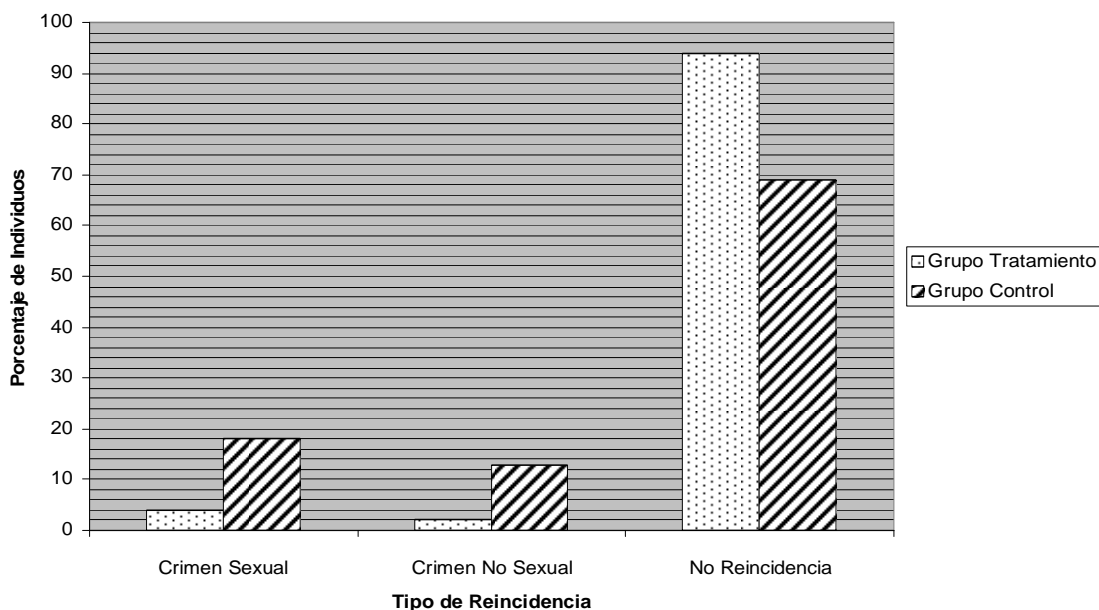
Estos son los resultados (no reales) de una encuesta efectuada a graduados de nuestra facultad que respondieron cuál ha sido la asignatura más compleja del primer año de la carrera. Observe la siguiente distribución de frecuencias relativas e identifique los ERRORES cometidos en su confección.

Asignatura	$p_i$	$p_a$
Psicoanálisis	1,2	1,2
Neurofisiología	0,1	1,3
Psicología Genética	- 0,8	0,5
Estadística	0,5	1
Total	1	

#### EJERCICIO 5

En Redondo y Garrido (2008) se evalúa la eficacia de cierto tratamiento de delinquentes sexuales que se hallaban en la prisión de Brians, Barcelona. Para ello se seleccionaron dos grupos mediante criterios que permitieron considerarlos equivalentes; el grupo al que se le aplicó el tratamiento consistía de 49 sujetos y el otro, llamado grupo control, de 74. Finalizado el período de tratamiento y una vez cumplida la sentencia, se llevó a cabo un seguimiento de los mismos durante aproximadamente 4 años y se registró si tuvieron reincidencias en delitos sexuales, otros delitos o no tuvieron reincidencias. Los resultados redondeados se exhiben en el siguiente gráfico.

**Distribución del Tipo de Reincidencia según el Tratamiento**



- Elabore las distribuciones de frecuencias, absolutas, relativas y porcentuales, a partir de la información que presenta el gráfico.
- ¿Qué tipo de gráfico se ha empleado para representar la información?
- Mencione la fuente sistemática de variación.
- Compare ambas distribuciones y extraiga alguna conclusión acerca de los resultados.

El artículo completo se halla en <http://www.psicothema.com/pdf/3422.pdf>

## EJERCICIO 6

La siguiente tabla de distribución de frecuencias corresponde a un estudio de habilidades numéricas para el cual se ha tomado la cantidad de problemas resueltos por un grupo de niños:

$X_i$	$n_i$	$n_a$	$p_i$	$p_a$	$P_i\%$	$P_a\%$
16 – 20		150				
11 – 15				0,7		
6 – 10	30					
1 – 5					40	
Total		-		-		-

- Complete la tabla a partir de los datos presentados.
- Obtenga los puntos medios y los límites exactos de cada uno de los intervalos

## EJERCICIO 7

Las siguientes calificaciones han sido obtenidas por 30 adultos en una prueba que evalúa su inteligencia:

100	95	86	97	104	105	110	85	96	102
103	112	98	106	101	104	95	114	89	91
92	115	120	117	99	97	95	100	90	112

- Presente la información en una distribución de frecuencias (absolutas, relativas, porcentuales con sus correspondientes acumuladas) con intervalos de amplitud 6. Redondee las frecuencias relativas a 4 decimales.
- Determine los puntos medios y límites exactos de cada uno de los intervalos.
- Grafique esta distribución con un histograma y un polígono de frecuencias absolutas.
- Confeccione una nueva tabla que contenga la información en 8 intervalos de amplitud 5, pero presente sólo las frecuencias absolutas.
- Exhiba en un diagrama de tallo y hojas las calificaciones de los 30 adultos.

## EJERCICIO 8

Un grupo de 20 niños de una escuela pública de Santiago del Estero han sido evaluados con una prueba de ortografía. Estos son los resultados:

2,5	5	6,5	9	4,5
5,5	8	6	5,5	10
3,5	8,5	7	5	4
4,5	9,5	6	7,5	6,5

- Obtenga la distribución de frecuencias completa con 4 intervalos.
- Determine los puntos medios y límites exactos de cada uno de los intervalos.
- Grafique esta distribución con un polígono de frecuencias absolutas.
- Dibuje un polígono de frecuencias porcentuales acumuladas.

## EJERCICIO 9

Qué gráfico utilizaría para representar la distribución de frecuencias de:

- La cantidad de materias aprobadas por los alumnos de una comisión.
- La orientación política de los habitantes de la provincia de Río Negro.
- Los resultados obtenidos por un alumno en una batería de tests.
- La opinión acerca del grado de satisfacción con el empleo actual de los argentinos.
- Tiempo de reacción frente a un estímulo visual de un grupo de 60 niños en etapa escolar.
- La opinión que tienen los graduados y los estudiantes sobre el plan de estudios vigente. Comparar ambas opiniones.

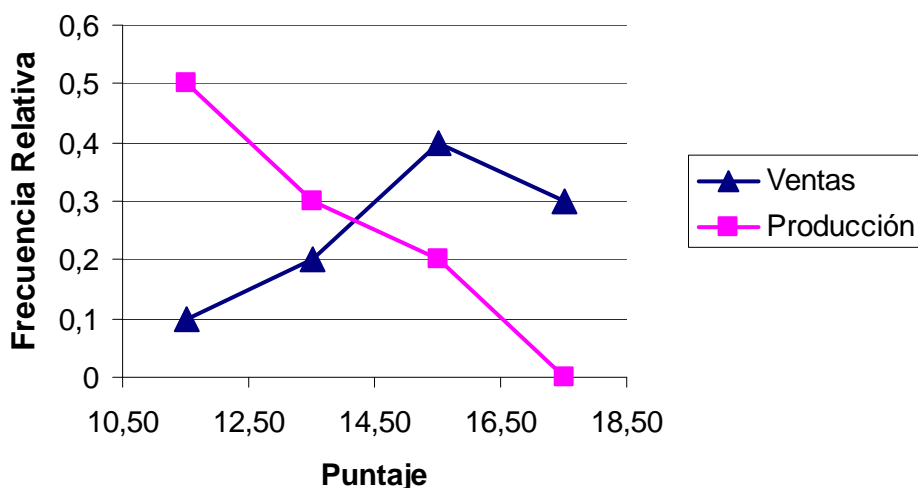
## EJERCICIO 10

Con el fin de estudiar la relación existente entre el género y el índice de religiosidad un investigador ha seleccionado aleatoriamente a 80 mujeres y 20 varones. Aquí reproducimos los resultados obtenidos por ambos grupos. Dibuje un polígono de frecuencias comparando ambas distribuciones y comente los resultados.

Puntaje	Varones	Mujeres
13 - 15	2	36
10 - 12	4	28
7 - 9	8	4
4 - 6	6	12

## EJERCICIO 11

El siguiente diagrama corresponde a los resultados de una prueba administrada a 16 empleados de ventas y 52 empleados del área de producción de una empresa que han respondido a la escala de Extroversión de un test de personalidad; mayores puntajes significan mayor extroversión.



- Diseñe una tabla de frecuencias absolutas y porcentuales con sus correspondientes frecuencias acumuladas. Incluya los límites exactos y aparentes.
- Cómo se denomina el gráfico empleado.
  - Por qué se aplicó dicho gráfico.
  - Por qué se ha utilizado la frecuencia relativa para su confección.
- ¿Qué conclusión podría extraer de los resultados obtenidos?

## RESPUESTAS

### EJERCICIO 1

a)

Variable: "Hábito o Disposición de los habitantes de Ciudad de Buenos Aires y Gran Buenos Aires para frecuentar al psicólogo".	Ángulos	Cant. de personas
Sí; voy actualmente.	57° 36'	4903
Sí pero dejé la terapia.	111° 36'	9500
No; no creo en la Psicología.	100° 48'	8581
No pero lo haría.	90° 00'	7662

b) El nivel de medición es nominal.

c) Las categorías cumplen la condición de ser mutuamente excluyentes y, con respecto a la información de hecho obtenida y en aras de una síntesis, también pueden considerarse colectivamente exhaustivas. Sin embargo, a priori, también quizás se hayan considerado las categorías NS/NC (No Sabe/No Contesta) y Otras; esta última suele incluirse para lograr la exhaustividad ante la posibilidad o el hecho de que se presenten respuestas con baja frecuencia; por ejemplo: "No iría porque no soy perseverante para las terapias".

### EJERCICIO 2

a) El cuadro contiene la información acerca de tres variables: "Percepción del nivel de conservación de sustancia", "Percepción del nivel de conservación de peso" y "Percepción del nivel de conservación de volumen". La fuente sistemática de variación es la "edad" de los niños.

b) Los ángulos son: 86° 24' (No conservación), 72° (Intermedio), 201° 36' (Conservación)

c) Una posibilidad es el diagrama de rectángulos adyacentes.

### EJERCICIO 3

a) Están representadas en el mismo diagrama dos distribuciones correspondientes a las variables "opinión de los jóvenes de 14 a 24 años respecto del nivel de la enseñanza primaria" y "opinión de los jóvenes de 14 a 24 años respecto del nivel de la enseñanza secundaria". El nivel de medición es ordinal y los valores son bajo, medio y alto.

b)

Opinión	Enseñanza Primaria						Enseñanza Secundaria					
	$n_i$	$n_a$	$p_i$	$p_a$	$P_i\%$	$P_a\%$	$n_i$	$n_a$	$p_i$	$p_a$	$P_i\%$	$P_a\%$
Alto	183	1019	0,18	1	18	100	158	1019	0,155	1	15,5	100
Medio	686	836	0,673	0,82	67,3	82	685	861	0,672	0,845	67,2	84,5
Bajo	150	150	0,147	0,147	14,7	14,7	176	176	0,173	0,173	17,3	17,3
	1019		1		100		1019		1		100	

c) 861 personas.

d) 85,3%.

#### EJERCICIO 4

ERROR 1) Carece de sentido acumular las frecuencias (absolutas, relativas o porcentuales) de los valores de una variable que han sido medidos en escala nominal.

ERROR 2) Las frecuencias relativas son índices que oscilan entre 0 y 1. Por ende, es imposible encontrar valores de  $p_i$  como 1,2 y  $-0,8$ .

#### EJERCICIO 5

a)

Tipo de Reincidencia	Grupo Tratamiento			Grupo Control		
	$n_i$	$p_i$	$P_i\%$	$n_i$	$p_i$	$P_i\%$
Crimen Sexual	2	0,04	4	13	0,18	18
Crimen No sexual	1	0,02	2	10	0,13	13
No Reincidencia	46	0,94	94	51	0,69	69
	49	1	100	74	1	100

c) Se ha empleado un diagrama de rectángulos adyacentes.

d) La fuente sistemática es el tipo de tratamiento: haberlo recibido o no.

e) El tratamiento parece ser eficaz dado que en esta muestra se observó un porcentaje considerablemente mayor ( $25\% = 94\% - 69\%$ ) de no reincidencia para el grupo que lo recibió.

#### EJERCICIO 6

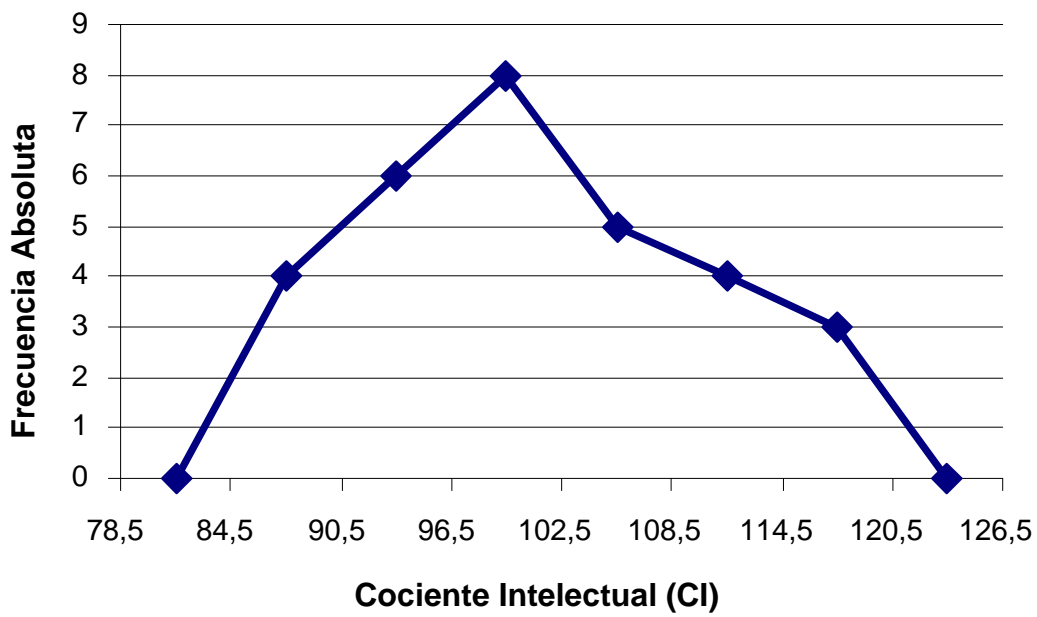
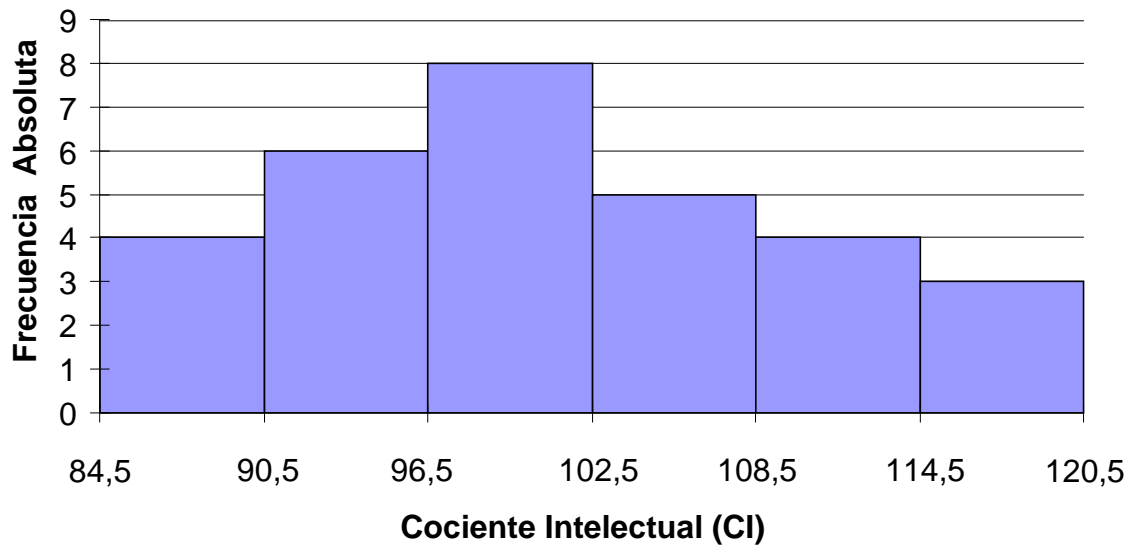
$X_i$	$n_i$	$n_a$	$p_i$	$p_a$	$P_i\%$	$P_a\%$	Puntos Medios	Límites Exactos
16 - 20	45	150	0,3	1	30	100	18	15,5 - 20,5
11 - 15	15	105	0,1	0,7	10	70	13	10,5 - 15,5
6 - 10	30	90	0,2	0,6	20	60	8	5,5 - 10,5
1 - 5	60	60	0,4	0,4	40	40	3	0,5 - 5,5
Total	150	-	1	-	100	-		

#### EJERCICIO 7

a) y b)

$X_i$	$n_i$	$n_a$	$p_i$	$p_a$	$P_i\%$	$P_a\%$	Puntos medios	Límites Exactos
115 - 120	3	30	0,1000	1	10	1	117,5	114,5 - 120,5
109 - 114	4	27	0,1333	0,9000	13,33	90	111,5	108,5 - 114,5
103 - 108	5	23	0,1667	0,7667	16,67	76,67	105,5	102,5 - 108,5
97 - 102	8	18	0,2667	0,6000	26,67	60	99,5	96,5 - 102,5
91 - 96	6	10	0,2000	0,3333	20	33,33	93,5	90,5 - 96,5
85 - 90	4	4	0,1333	0,1333	13,33	13,33	87,5	84,5 - 90,5
	30	-	1	-	100	-		

c) Histograma y polígono



d)

$X_i$	$n_i$
118 - 122	1
113 - 117	3
108 - 112	3
103 - 107	4
98 - 102	7
93 - 97	6
88 - 92	4
83 - 87	2

e)

8	569
9	01255567789
10	001234456
11	022457
12	0

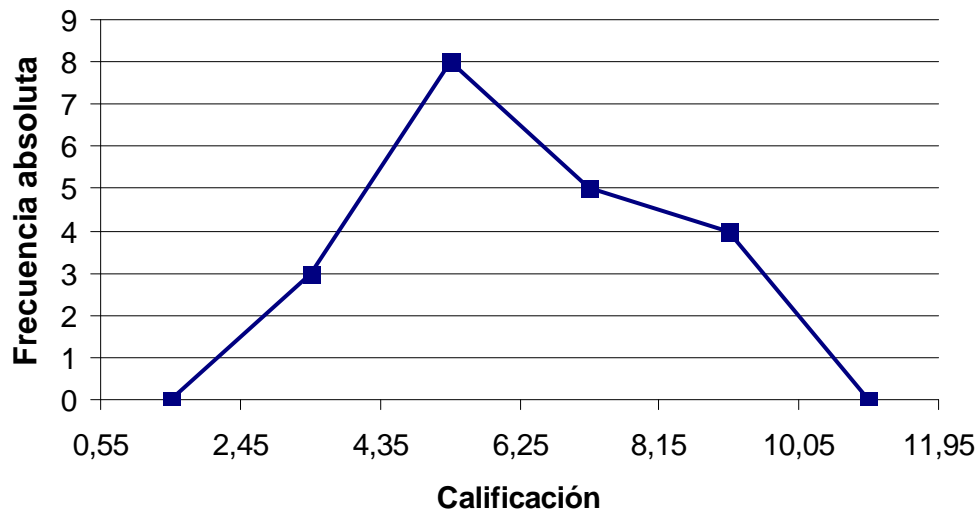
**EJERCICIO 8**



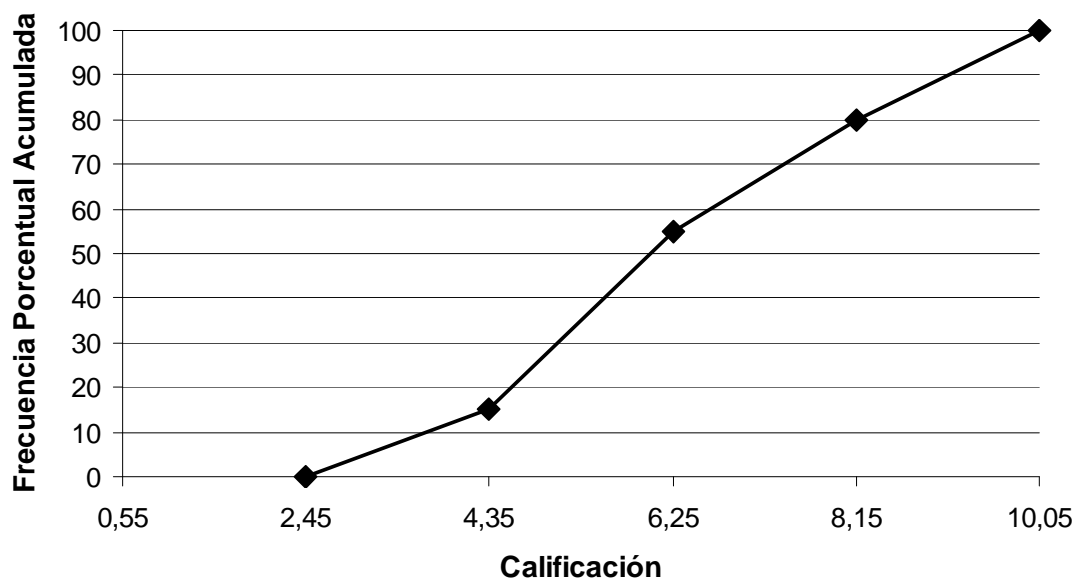
a) y b)

$X_i$	$n_i$	$n_a$	$p_i$	$p_a$	$P_i\%$	$P_a\%$	Puntos medios	Límites Exactos
8,2 – 10	4	20	0,20	1,00	20	100	9,1	8,15 – 10,05
6,3 – 8,1	5	16	0,25	0,80	25	80	7,2	6,25 – 8,15
4,4 – 6,2	8	11	0,40	0,55	40	55	5,3	4,35 – 6,25
2,5 – 4,3	3	3	0,15	0,15	15	15	3,4	2,45 – 4,35
	20	-	1	-	100	-		

c) Polígono de frecuencias absolutas



d) Polígono de frecuencias porcentuales acumuladas

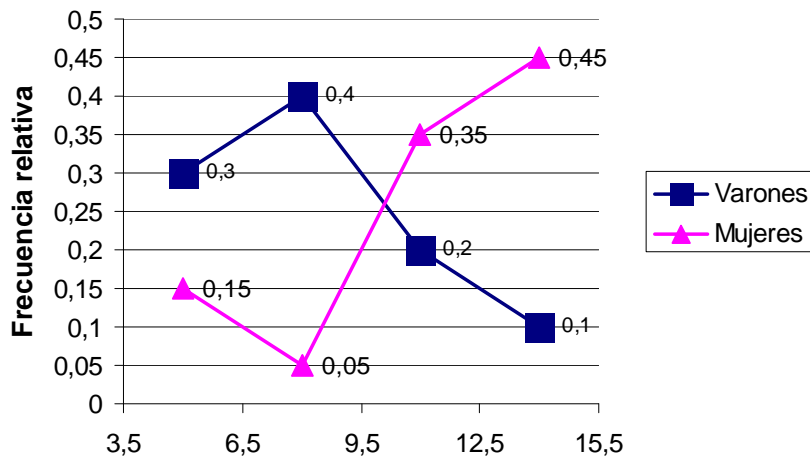


## EJERCICIO 9

- a) Diagrama de barras.
- b) Pictograma / Diagrama de rectángulos.
- c) Perfil ortogonal.
- d) Diagrama de rectángulos.
- e) Histograma / Polígono de frecuencias.
- f) Diagrama de rectángulos adyacentes.

**EJERCICIO 10**

Se trata de muestras de tamaños diferentes, es por este motivo que el gráfico no puede ser realizado con frecuencias absolutas. Estos datos muestran una tendencia de las mujeres a obtener índices de religiosidad más elevados que los varones.



**EJERCICIO 11**

X <sub>i</sub>	Ventas				Producción				Puntos medios	Límites Exactos
	n <sub>i</sub>	n <sub>a</sub>	P <sub>i</sub> %	P <sub>a</sub> %	n <sub>i</sub>	n <sub>a</sub>	P <sub>i</sub> %	P <sub>a</sub> %		
11 – 12	2	2	10	10	26	26	50	50	11,5	10,5 – 12,5
13 – 14	3	5	20	30	16	42	30	80	13,5	12,5 – 14,5
15 – 16	6	11	40	70	10	52	20	100	15,5	14,5 – 16,5
17 – 18	5	16	30	100	0	52	0	100	17,5	16,5 – 18,5
	16		100		52		100			

- b) Se ha empleado un polígono de frecuencias ya que se pretendía comparar distribuciones referidas a una variable cuantitativa continua. Se ha utilizado la frecuencia relativa porque las muestras son de diferente tamaño.
- c) En el área de Ventas son más frecuentes los puntajes altos de extroversión que en el de Producción; por lo que se concluye que para la muestra considerada los empleados de la sección Ventas son en general más extrovertidos que los del área de Producción.

**REFERENCIAS**

Piaget, J. y Inhelder, B. (1970). *El Desarrollo de las Cantidades en el Niño*. Introducción. Editorial Nova Terra. Barcelona.

Redondo, S. y Garrido, V. (2008). Efficacy of a psychological treatment for sex offenders. *Psicothema*, 20, 1, 4 - 9.

Simón, T., Ruiz Gallego-Largo, T. y Suengas, A. (2009). Memoria y envejecimiento: recuerdo, reconocimiento y sesgo positivo. *Psicothema*, 21, 3, 409 - 415.

## PRÁCTICA III

*Cuando era joven y practicaba en el juzgado perdí muchos casos que debería haber ganado. Pero con el tiempo gané muchos casos que debería haber pedido; así la justicia quedó compensada.*

*Lord Justice Matthews*

### EJERCICIO MODELO

#### PARTE I.

Utilice las tablas halladas en la parte d) del Ejercicio Modelo de la Práctica II para responder los siguientes ítems.

- Considerando el nivel de medición de la Valoración de las Imágenes ¿qué medidas de tendencia central, *en rigor*, podrían tenerse en cuenta para comparar ambas distribuciones?
- Calcule la moda para cada distribución. ¿En cuál de las dos distribuciones la moda es más representativa y por qué?
- ¿Cuál de las dos distribuciones tiene mayor entropía?
- Si quisiera comparar ambas distribuciones sobre la base de un representante de cada una ¿qué resumen estadístico utilizaría y por qué?
- Halle los cuartiles para ambas distribuciones y saque una conclusión sobre la base de su comparación.

#### PARTE II.

Para la variable Recordación, como número de imágenes recordadas entre 50, suponga que se obtuvieron los siguientes datos<sup>1</sup>:

Jóvenes	20	28	22	31	28	19
	25	16	29	27	21	25
	25	26	18	26	24	23
	27	24	28	33		

Mayores	13	16	17	23	19	20
	16	14	14	12	12	16
	19	11	9	15	23	17
	17	21	19	10		

- Considerando el nivel de medición de la Recordación de las Imágenes ¿qué medidas de tendencia central, en principio, podrían tenerse en cuenta para comparar a los jóvenes con los mayores en cuanto a su facilidad de recordación? ¿Si tuviera que elegir una de ellas cuál elegiría en este caso y por qué? Compare los grupos sobre la base de la medida elegida.
- Compare ambas distribuciones respecto de su variabilidad. ¿Cuál de las dos distribuciones tiene más variabilidad relativa a la media?
- La capacidad de recordación podría concebirse en un continuo que ha sido discretizado por el modo de medición (cantidad de palabras recordadas). Rescatando este carácter continuo de la variable, para cada distribución
  - agrupe los datos en intervalos de clase de amplitud 3.
  - obtenga los polígonos de frecuencias y compárelos respecto de su asimetría y curtosis.
  - halle los cuartiles y compárelos.

<sup>1</sup> Al no disponer de los datos originales de la investigación, para el propósito del presente ejercicio los datos se simularon de modo que sus resúmenes estadísticos coincidieran aproximadamente con los que informaron los autores en los resultados.

### PARTE III.

Sobre la base de lo analizado en las Partes I y II sintetice una conclusión general.

### RESPUESTAS

#### Parte I.

- Siendo de nivel ordinal *en rigor*, corresponde tomar la moda o la mediana.<sup>2</sup>
- Para los jóvenes es 4 y para los mayores es 6. Es más representativa para el grupo de los mayores, ya que hay mayor concentración de la frecuencia en la moda y en su valor adyacente (5), mientras que en el grupo de los jóvenes la frecuencia de la moda no difiere mayormente de la frecuencia de los otros valores, inclusive los más distantes a ella como son el 1 o el 7. En efecto, para este último grupo la distribución es casi uniforme.
- Claramente hay mayor entropía en el grupo de los jóvenes, precisamente porque la distribución de frecuencias es más uniforme.
- Utilizaría la mediana por el nivel de medición ordinal y porque toma en cuenta más información de los datos que la moda, como es su posición relativa; además de que en este caso la moda no es representativa para el grupo de los jóvenes.
- 

Cuartiles	Jóvenes	Mayores
Q <sub>1</sub>	2	4
Q <sub>2</sub> (Mediana)	4	5
Q <sub>3</sub>	6	6

El hecho de que el primer y segundo cuartil son mayores en el grupo de las personas mayores muestra que hay una tendencia de este grupo a experimentar mayor agrado por las imágenes.

#### PARTE II.

- Dado que la variable es de nivel de razón y, por tanto, reúne las propiedades de los niveles de medición anteriores; tiene sentido obtener la moda, la mediana y la media. La moda es de poco valor en este caso a menos que se agrupen los datos, ya que ningún valor aislado concentra una parte importante de la frecuencia total. Para representar cada grupo y compararlos sería conveniente, entonces, elegir la mediana o la media. Pero al no haber datos fuertemente atípicos que distorsionen la media, se prefiere este resumen porque utiliza más información que la mediana; no sólo la posición relativa de los valores sino también su magnitud. En conclusión, elegiría a la media. Los jóvenes recordaron en promedio 24,773 palabras mientras que los mayores 16,045; es decir que los jóvenes recordaron en general más palabras que los mayores.
- La variabilidad del número de palabras recordadas, en términos absolutos, es algo mayor en el grupo de los jóvenes. La mayor diferencia se observa en la amplitud total, lo que da cuenta de que la variable toma valores en un recorrido más amplio. El rango semi intercuartílico es el mismo para ambos grupos y tanto la desviación media como la estándar son ligeramente superiores para el grupo de los jóvenes. Sin embargo, en relación al valor medio, que es sensiblemente menor en el grupo de adultos, la variabilidad resulta mayor para éstos, como expresa el coeficiente de variación. Estas medidas se muestran en la siguiente tabla:

<sup>2</sup> Sin embargo es habitual que los investigadores hagan el supuesto de que los distintos niveles son equidistantes y hagan un tratamiento intervalar de este tipo de escalas y resuman los datos con la media y la desviación estándar.

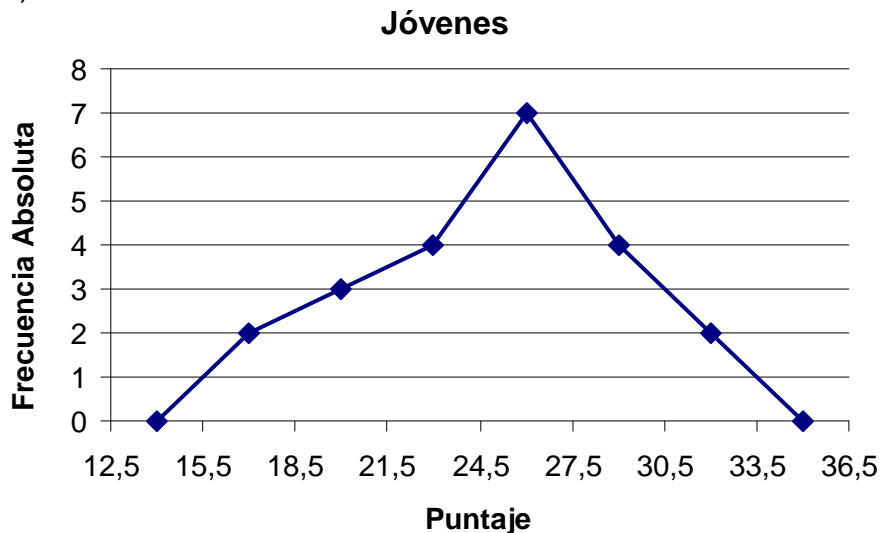
Medidas de Variabilidad	Jóvenes	Mayores
Amplitud Total	17	14
Amplitud Semi intercuartílica	3	3
Desviación Media	3,27	3,14
Desviación Estándar	4,12	3,87
Cuasi Desviación Estándar	4,22	3,96
Coefficiente de Variación <sup>3</sup>	17,03	24,67

c) i)

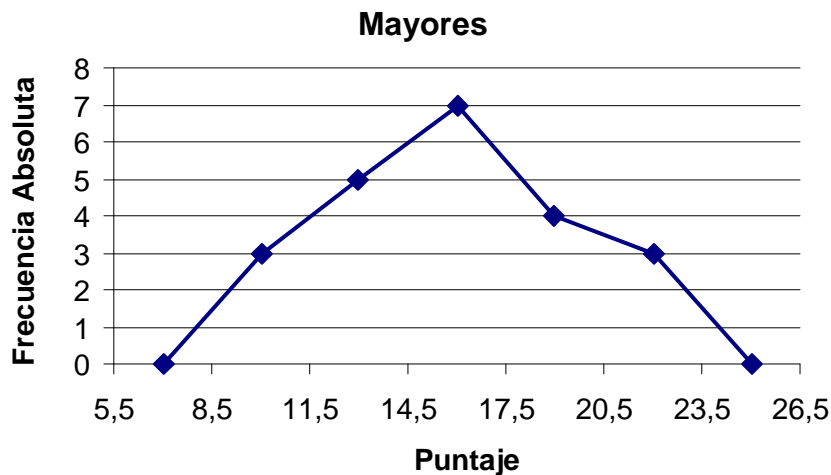
Datos agrupados en intervalos de clase para el grupo "Jóvenes"	Límites Aparentes	Límites Exactos	Punto Medio	$n_i$	$n_a$
	16 – 18	15,5 – 18,5	17	2	2
	19 – 21	18,5 – 21,5	20	3	5
	22 – 24	21,5 – 24,5	23	4	9
	25 – 27	24,5 – 27,5	26	7	16
	28 – 30	27,5 – 30,5	29	4	20
	31 – 33	30,5 – 33,5	32	2	22

Datos agrupados en intervalos de clase para el grupo "Mayores"	Límites Aparentes	Límites Exactos	Punto Medio	$n_i$	$n_a$
	9 – 11	8,5 – 11,5	10	3	3
	12 – 14	11,5 – 14,5	13	5	8
	15 – 17	14,5 – 17,5	16	7	15
	18 – 20	17,5 – 20,5	19	4	19
	21 – 23	20,5 – 23,5	22	3	22

ii)



<sup>3</sup> Está calculado sobre la base de la Cuasi Desviación Estándar. Para comparar ambas distribuciones es redundante calcular desviación estándar y cuasi desviación estándar. Para fines meramente descriptivos puede usarse la desviación estándar; para fines inferenciales como se verá en la segunda parte de la materia, se utiliza la cuasi desviación estándar.



La distribución de los puntajes para los mayores, aunque presenta una ligera asimetría positiva, es más simétrica que la de los jóvenes. En efecto, obsérvese que el valor donde alcanza el máximo es más equidistante de los extremos en la distribución de los mayores que en la de los jóvenes, la cual presenta una cierta asimetría negativa y es algo más cúrtica.

iii)

Cuartiles	Jóvenes	Mayores
Q <sub>1</sub>	21,875	13
Q <sub>2</sub> (Mdn)	25,357	15,786
Q <sub>3</sub>	27,875	18,625

Comparando los cuartiles, siendo todos ellos mayores en el grupo de los jóvenes, se llega a la misma conclusión que en a) habiendo utilizado la media; es decir que los jóvenes tuvieron, en general, mejor recordación que los mayores.

### PARTE III.

Los datos muestrales dan indicios de que la edad es una fuente sistemática de variación tanto para la Valoración de las Imágenes como para la facilidad de Recordación, siendo los mayores los que tienden a experimentar más agrado por las imágenes y también a recordar menos imágenes.<sup>4</sup>

## EJERCITACIÓN

### EJERCICIO 1

Un investigador pretende analizar la posible relación entre soledad y uso de Internet. Para tal fin interrogó a 35 personas que viven solas y a 25 que viven con sus familias acerca de la cantidad de minutos que suelen estar conectadas en una sesión típica durante el fin de semana. Ambas muestras fueron homogeneizadas en la edad y el nivel socioeconómico. Los datos son los siguientes:

<sup>4</sup> Hasta el momento esto puede afirmarse sólo para los individuos involucrados en estas muestras. En Estadística Inferencial se verán los criterios como para saber si esta conclusión también puede extenderse sobre la población en general.

Tiempo (en minutos)	Personas que viven solas	Personas que viven con su familia
60 - 74	3	1
45 - 59	11	3
30 - 44	10	8
15 - 29	6	7
0 - 14	5	6

Calcule el  $D_5$  en cada grupo. ¿Qué indica este resultado respecto de la relación entre la soledad y el uso de Internet?

## EJERCICIO 2

Replicando un clásico estudio empírico de la Psicología Social llevado a cabo por Rosenhan (1988), un investigador decidió infiltrar a “falsos pacientes” en distintas clínicas psiquiátricas de EEUU. Su objetivo fue evaluar la reacción del personal al contacto verbal iniciado por el paciente. Los pseudo-pacientes abordaron a psiquiatras y enfermeros de las respectivas instituciones en un encuentro que parecía casual y les formularon un máximo de seis preguntas secuenciadas referidas a su “diagnóstico” y su “pronóstico”. El investigador consideró como variable estadística a la cantidad de preguntas respondidas por los profesionales en ese encuentro. Estos son los resultados (datos ficticios):

Cantidad de preguntas respondidas	Enfermeros	Psiquiatras
0	1	2
1	3	4
2	6	2
3	8	1
4	9	0
5	7	1
6	1	0

La hipótesis del investigador sostiene que el contacto de los psiquiatras con los pacientes es más frágil que el de los enfermeros. Calcule el segundo cuartil de ambas distribuciones y extraiga conclusiones al respecto.

## EJERCICIO 3

Un profesor considera que deben promocionar la asignatura los alumnos que obtuvieron un promedio de notas en ambos parciales superior al centil 90. El profesor utiliza la puntuación 35 para determinar la categoría de “aprobado”. La siguiente es la distribución de frecuencias de los promedios de las notas de dicho profesor:

$X_i$	$n_i$
81 - 100	6
61 - 80	25
41 - 60	38
21 - 40	16
1 - 20	5

En función de los datos de la tabla:

- Calcule el porcentaje de alumnos que desaprobaron la materia.
- Calcule cuál será la puntuación mínima para promocionar.

#### EJERCICIO 4

Una institución que asiste a drogadependientes se interesa por conocer el nivel de ansiedad motora de sus pacientes. Los profesionales deciden administrar la prueba ISRA de Miguel Tobal y Cano Vindel (1998). La misma utiliza el primer cuartil para calificar “ansiedad baja”; el tercer cuartil para calificar “ansiedad alta” y el centil 99 para la “ansiedad extrema”. Calcule las puntuaciones que delimitan estas categorías.

Puntaje	$n_i$
106 - 119	10
92 - 105	45
78 - 91	36
64 - 77	18
50 - 63	20
36 - 49	15

#### EJERCICIO 5

Según un test que evalúa la variable Depresión, los pacientes que obtienen una puntuación superior a 40 presentan riesgo suicida. La siguiente es la distribución de frecuencias de los puntajes obtenidos en el test por 45 mujeres y 60 varones diagnosticados como depresivos, que concurren a un Centro de Salud del Gran Buenos Aires.

Puntaje	Mujeres	Varones
43,5 – 50,5	4	5
36,5 – 43,5	10	2
29,5 – 36,5	14	11
22,5 – 29,5	7	24
15,5 – 22,5	10	18

¿Cuál de los dos grupos presenta el mayor porcentaje de pacientes con riesgo suicida?

#### EJERCICIO 6

El gerente de una empresa se pregunta si los años de experiencia de un obrero en su puesto favorecen el cumplimiento de las normas de seguridad y mantenimiento de las maquinarias. Para ello decide:

1) *Seleccionar dos grupos bien diferenciados.* El grupo de los novatos que estará compuesto por aquellos operarios cuyos años de experiencia no superen el centil 10 y el grupo experimentado que estará compuesto por los que superen el centil 90. Sabiendo que la empresa emplea a 200 operarios ¿cuántas personas integrarán cada uno de los grupos?

2) *Evaluar a los grupos y analizar los resultados.* Ambos grupos respondieron una prueba destinada a evaluar el conocimiento y cumplimiento de las normas de seguridad y mantenimiento de las maquinarias. Los resultados obtenidos son los siguientes:

Puntaje	Grupo Novato	Grupo Experimentado
17 – 20	2	1
13 – 16		3
9 – 12	8	
5 – 8	5	6
1 – 4	2	2
Total		



- a) Complete la tabla a partir de los datos obtenidos en 1).
- b) Obtenga los cuartiles del grupo novato.
- c) Si los cuartiles para el grupo experimentado son los siguientes:  $Q_1= 6,5$ ;  $Q_2=9,5$ ;  $Q_3=12$ , ¿cómo interpretaría estos resultados comparando ambos grupos para responder a la pregunta del gerente?

**EJERCICIO 7**

La distribución de frecuencias que se presenta muestra los resultados obtenidos en una prueba que mide el nivel de Escrupulosidad de los postulantes para un empleo de vigilancia privada. Los encargados de la selección decidieron optar por aquellos aspirantes que presentaran puntajes intermedios de Escrupulosidad. Esto significa que sólo quedará preseleccionado el 50% de los candidatos que obtienen las puntuaciones centrales.

$X_i$	$n_i$
13 - 15	25
10 - 12	23
7 - 9	14
4 - 6	6
1 - 3	2

¿Qué puntuaciones como mínimo y como máximo será necesario obtener para quedar preseleccionado?

**EJERCICIO 8**

El director de un geriátrico decide analizar la cantidad de visitas semanales que recibe cada uno de los ancianos internados con el fin de tomar precauciones contra los síntomas depresivos que puede producir el sentimiento de abandono. Él considera que debe despreocuparse de aquellos ancianos que reciban más de 4 visitas, mientras que deberá evaluar psicológicamente a aquellos cuya cantidad de visitas no supere el decil 3. Los siguientes datos corresponden a la cantidad de visitas recibidas por los 30 ancianos del establecimiento.

0	5	2	1	7	6	5	0	2	3	5	1	2	2	5
7	6	1	3	2	4	1	4	0	2	5	6	3	4	1

- a) Presente la información en una tabla de frecuencias absolutas (sin agrupar los datos en intervalos de clase).
- b) ¿Qué porcentaje de casos no preocupan al director?
- c) ¿Qué cantidad máxima de visitas debe tener el anciano para ser evaluado psicológicamente?
- d) Si el director desea seleccionar a aquellos ancianos cuya cantidad de visitas ha superado el centil 80 o no ha superado al centil 20 para realizarles un test de personalidad ¿cuántos ancianos debe elegir?

**EJERCICIO 9**

Treinta ratas fueron separadas aleatoriamente en dos grupos iguales con la finalidad de comparar el efecto de la anestesia en las intervenciones quirúrgicas. El Grupo Tratamiento recibió un simulacro de operación que consistió en anestesiarse al animal y efectuarle un corte inocuo, mientras que el Grupo Control no fue intervenido.

Un día después se registró el tiempo que cada una necesitó para recorrer un laberinto. Esta tabla resume los resultados de ambos grupos:

Tiempo (seg.)	Grupo Tratamiento	Grupo Control
55 - 59	1	-
50 - 54	2	4
45 - 49	5	3
40 - 44	4	6
35 - 39	3	2

- Calcule la media, la mediana y la moda aproximada<sup>5</sup> de ambas distribuciones.
- ¿Observa diferencias en el rendimiento de las ratas entre las anestesiadas y las no anestesiadas? Conteste a partir del estadístico más adecuado.

### EJERCICIO 10

Una vez tomados todos los tiempos del Ejercicio anterior los investigadores descubrieron que un error en el instrumento de medición agregó 10 segundos a cada una de las marcaciones. Indique cuáles son los valores corregidos de la media, la mediana y la moda.

### EJERCICIO 11

Se ha realizado un estudio a fin de conocer los niveles de estrés laboral percibidos por los habitantes de una ciudad argentina y se ha pensado en la profesión como una posible fuente sistemática de variación. Aquí reproducimos los puntajes obtenidos en la prueba que mide el estrés laboral para cada una de las profesiones:

Grupo de maestros	45 – 32 – 29 – 49 – 40 – 44 – 33 – 46 – 45 – 48
Grupo de bomberos	41 – 36 – 31 – 36 – 39 – 42 – 35 – 36 – 40
Grupo de abogados	38 – 25 – 46 – 41 – 37 – 50 – 43
Grupo de médicos	43 – 37 – 50 – 53 – 40 – 35 – 56 – 49 – 55 – 45

Calcule la mediana de cada uno de los grupos y responda, en función de este estadístico, cuál parece ser el grupo más afectado por el estrés.

### EJERCICIO 12

Con el objetivo de analizar la influencia del contexto en la resolución de problemas de matemática un grupo de investigadores evaluó el desempeño de 90 niños que ayudan a sus padres en la atención de los puestos de una feria. La experiencia constaba de: 1) una prueba informal realizada en la feria en donde un entrevistador proponía una situación problemática de compra-venta de atención al público y 2) una prueba formal que consistía en la resolución de operaciones aritméticas equivalentes a las de la prueba informal pero sin la situación de compra-venta de atención al público. Se eligieron 45 niños al azar entre los 90 para administrarle la prueba informal y al resto se les administró la prueba formal.

<sup>5</sup> Existe una fórmula para la moda de datos agrupados en intervalos de clase que toma en cuenta la marca de clase del intervalo de mayor frecuencia y también la frecuencia de los dos intervalos adyacentes. Por simplicidad no utilizaremos esa fórmula y aproximaremos con el punto medio del intervalo modal.

Cantidad de problemas resueltos correctamente	Prueba Informal	Prueba Formal
54 – 62	7	3
45 – 53	15	5
36 – 44	9	3
27 – 35	7	9
18 – 26	6	14
9 – 17	0	9
0 – 8	1	2

- a) Calcule media, mediana y moda aproximada de ambos grupos.  
b) Extraiga conclusiones a partir de los resultados obtenidos.

*Nota:* El contenido de este ejercicio está inspirado en una investigación de Caraher, Carraher y Schlieman (1998). Remitimos al alumno interesado en los hallazgos y conclusiones de este estudio a consultar la publicación citada en la bibliografía.

### EJERCICIO 13

Un investigador supone que la cantidad de público presente en un auditorio provoca reacciones psicofisiológicas que afectan al orador. Para probar esto, se trabajó con 26 conferencistas que debían pronunciar un discurso de igual complejidad y que fueron repartidos aleatoriamente en dos grupos: uno con el auditorio completo de público y otro con el auditorio con escaso público. Mientras tanto, un equipo registraba las pulsaciones por minuto ofreciendo, al final de la experiencia, un valor promedio de las mismas para cada uno de los conferencistas. A continuación se reproducen los datos:

Cantidad de pulsaciones por minuto	Auditorio completo De público	Auditorio con escaso público
101 – 109	2	1
92 – 100	6	1
83 – 91	5	3
74 – 82	2	4
65 – 73	1	1

- a) Calcule la cantidad de pulsaciones promedio por minuto de cada grupo.  
b) Extraiga conclusiones a partir de los resultados de la experiencia.

### EJERCICIO 14

En cierta comisión de Estadística la media de las calificaciones del primer parcial fue 6 para las mujeres y 5 para los varones. Calcule la media de todas las calificaciones de dicha comisión sabiendo que corresponden a 20 mujeres y 10 varones.

### EJERCICIO 15

En una empresa el personal está categorizado según se indica a continuación.

Categorías	Salario	Nº de empleados
1	\$ 1.300	12
2	\$ 1.500	31
3	\$ 1.700	27
4	\$ 2.200	7
5	\$ 3.000	4
6	\$ 5.000	2
7	\$ 7.000	3

- a) Calcule las medidas de tendencia central adecuadas para la variable "Categoría".
- b) Calcule las medidas de tendencia central adecuadas para la variable "Salario".
- c) Si quisiera representar la escala salarial de la empresa cuál de las medidas calculadas en b) utilizaría en caso de ser:
  - i. Sindicalista.
  - ii. Directivo.
  - iii. Economista objetivo.

### EJERCICIO 16

Sobre la base de la información brindada en este cuadro responda las siguientes preguntas (los datos no son reales):

¿Qué diario compra habitualmente?				
Barrio	Clarín	La Nación	Página 12	Ninguno
Belgrano	39 %	51 %	5 %	5 %
Flores	40 %	20 %	20 %	20 %
Mataderos	47 %	16 %	11 %	26 %

- a) ¿Cuál es el diario comprado más frecuentemente en cada barrio?
- b) ¿Con qué medida compararía las distribuciones de los barrios con respecto a su variabilidad en el consumo de diarios?
- c) ¿Qué distribución presenta mayor entropía: la del barrio de Belgrano o la del barrio de Flores?

### EJERCICIO 17

¿Está de acuerdo con la despenalización de la marihuana?		Ficha técnica (Datos ficticios)
Menores de 25	Mayores de 25	
		<p>Encuesta realizada en la página web de un periódico porteño por 420 sujetos; 70% menores de 25 años y 30% mayores de 25, en el primer semestre de 2009.</p>

Teniendo en cuenta la información gráfica, responda lo siguiente:

- a) ¿Cuál es la opinión más frecuente en cada grupo etéreo?
- b) ¿En cuál de los grupos la opinión de la población está mejor representada por la moda?
- c) ¿Qué distribución de opiniones presenta mayor entropía?
- d) Indique con qué otro tipo de gráfico podría representar los datos si quisiera comparar los diferentes grupos etéreos.

### EJERCICIO 18

Un investigador desea estudiar si el hecho de realizar una actividad o de permanecer ocioso durante un determinado lapso influye en la estimación subjetiva del tiempo transcurrido en dicho período. Efectuó una prueba piloto con 10 sujetos, la mitad permaneció inactivo (grupo inactivo) durante 25 minutos, la otra mitad realizó una tarea (grupo activo) que consistió en tomar nota del dictado de un texto. Ninguno de los dos grupos tenía la ayuda de un reloj. Al finalizar los 25 minutos los 10 sujetos debieron estimar el lapso transcurrido y resultó:

Grupo Inactivo	30	23	28	19	26
Grupo Activo	23	24	26	18	29

- Calcule la amplitud total, la media y el desvío típico de ambos grupos.
- ¿Qué indicios dan estos resultados respecto de la hipótesis de la investigación?
- ¿Qué grupo ha resultado más variable en su estimación del tiempo?

### EJERCICIO 19

¿Cómo habrían resultado los resúmenes estadísticos calculados en el Ejercicio 18 si...

- todos los datos hubieran estado incrementados en 2 minutos;
- todos los tiempos hubieran sido percibidos en un 50% más largos?

### EJERCICIO 20

Los 28 docentes de una asignatura fueron divididos aleatoriamente en dos grupos para calificar un parcial. Mientras que el Grupo 1 no poseía referencia alguna sobre la calificación obtenida por el alumno en el primer parcial, el Grupo 2 sí la poseía.

Calificación	Grupo 1	Grupo 2
10	1	1
9	2	5
8	3	4
7	4	3
6	2	1
5	2	0

- Calcule la media, la varianza y el desvío típico de ambos grupos.
- Si a las calificaciones originales se las multiplica por 10, ¿cómo se modifican los resultados de la media, la varianza y el desvío típico?

### EJERCICIO 21

Para medir el tiempo reacción ante ciertos estímulos visuales se dispone del test A y para estímulos sonoros del test B. En la tabla siguiente se muestran los tiempos de reacción de Karina a 10 estímulos de cada tipo. ¿En cuál test ha tenido un desempeño más homogéneo?

. Tiempo (seg.)	Test A	Test B
33 – 39	3	-
26 – 32	3	1
19 – 25	2	2
12 – 18	1	3
5 – 11	1	4

## EJERCICIO 22

En un centro de psicopedagogía clínica 20 niños disléxicos, fueron divididos en dos grupos de 10 y recibieron el mismo tratamiento pero con diferentes terapeutas: un grupo con el terapeuta A y el otro con el terapeuta B. Al finalizar el tratamiento se les administró un test cuyos mayores puntajes están vinculados a mayor cantidad de palabras incorrectamente leídas. Los resultados obtenidos por ambos terapeutas son resumidos a continuación:

Terapeuta A	
Puntaje	ni
3	4
4	1
5	3
6	1
7	1

Terapeuta B
$\bar{x} = 6$ Mdn = 5,5 Mo = 7 s = 1,44 CV = 24

Basándose sobre los resúmenes adecuados responda

- ¿Qué grupo obtuvo mejores resultados en el test?
- ¿Puede afirmarse que el terapeuta A es más eficiente que el terapeuta B? Justifique.

## EJERCICIO 23

La estatura promedio de una muestra de 150 adultos es de 1,72 m con un desvío estándar de 0,07m.

- Si se sabe que el promedio de peso de los sujetos de esta muestra es de 75 kg. con un desvío de 8 kg. ¿Es posible afirmar que los pesos son más variables que las estaturas porque les corresponde desvío estándar mayor?
- Si se sabe que la estatura promedio de los hijos de los sujetos de la muestra es de 0,83 m con un desvío de 0,05 m. ¿Es posible afirmar que los padres varían más en la estatura que los hijos?

## EJERCICIO 24

Los alumnos de una comisión de Escuela Francesa obtuvieron las calificaciones que se muestran en las siguientes tablas:

Calificaciones	Frecuencias	
	1º Parcial	2º Parcial
3	2	1
4	4	1
5	7	2
6	6	4
7	3	6
8	2	8
9	1	3
10	1	1

Grafique ambas distribuciones de modo que puedan ser comparadas y responda sin hacer cálculos:

- ¿En qué parcial la comisión obtuvo mejor rendimiento?

- b) ¿Qué parcial discriminó mejor a los alumnos de mejor rendimiento? ¿Y a los de peor rendimiento?
- c) ¿Cómo debería ser la forma de la distribución para discriminar por igual a los alumnos de mejor y peor rendimiento?
- d) Explique la vinculación entre el tipo de asimetría y el efecto discriminatorio de la prueba.

**EJERCICIO 25**

En su artículo “Elementos de estadística al servicio de la evaluación del rendimiento”, la Lic. Floralba Cano de Becerra (1971) aconseja tomar pruebas que produzcan puntajes con asimetría positiva al inicio del curso y asimetría negativa en los exámenes finales. ¿Sobre qué criterio cree Ud. que se apoya esta sugerencia?

**EJERCICIO 26 (Parcialmente resuelto)**

Con los datos del Ejercicio 7 de la Práctica I realice un análisis exploratorio a fin de determinar si la muestra da indicios de que el maestro y la modalidad de la clase son fuentes sistemáticas de variación para la rapidez y la comprensión lectora. (Use calculadora).

Resolveremos, a manera de ejemplo, el ejercicio para la variable R = Rapidez lectora y dejamos como ejercicio para el alumno el análisis para la variable C = Comprensión lectora.

Comenzaremos resumiendo la información. Dado que el nivel de medición lo permite, podemos representar los datos de las celdas por su media y resumir su variabilidad con el desvío estándar (usamos el que se obtiene de dividir por n). Codificaremos con CA y CN las modalidades de clases abreviadas y no abreviadas respectivamente, y con M1, M2 y M3 a los maestros 1, 2 y 3.

La tabla con los resúmenes queda así:

	CA		CN		Total	
	R	C	R	C	R	C
M1	$\bar{X} = 11$ $s = 1,79$	$\bar{X} =$ $s =$	$\bar{X} = 9,2$ $s = 0,98$	$\bar{X} =$ $s =$	$\bar{X} = 10,1$ $s = 1,7$	$\bar{X} =$ $s =$
M2	$\bar{X} = 13,4$ $s = 1,36$	$\bar{X} =$ $s =$	$\bar{X} = 10,2$ $s = 1,17$	$\bar{X} =$ $s =$	$\bar{X} = 11,8$ $s = 2,04$	$\bar{X} =$ $s =$
M3	$\bar{X} = 8$ $s = 1,09$	$\bar{X} =$ $s =$	$\bar{X} = 8,8$ $s = 0,75$	$\bar{X} =$ $s =$	$\bar{X} = 8,4$ $s = 1,02$	$\bar{X} =$ $s =$
Total	$\bar{X} = 10,8$ $s = 2,64$	$\bar{X} =$ $s =$	$\bar{X} = 9,4$ $s = 1,14$	$\bar{X} =$ $s =$	$\bar{X} = 10,1$ $s = 2,15$	$\bar{X} =$ $s =$

Antes de continuar responda las siguientes preguntas:

Pregunta 1:

¿Qué representan las medias y los desvíos del margen derecho de la tabla? ¿Y los de los últimos renglones?

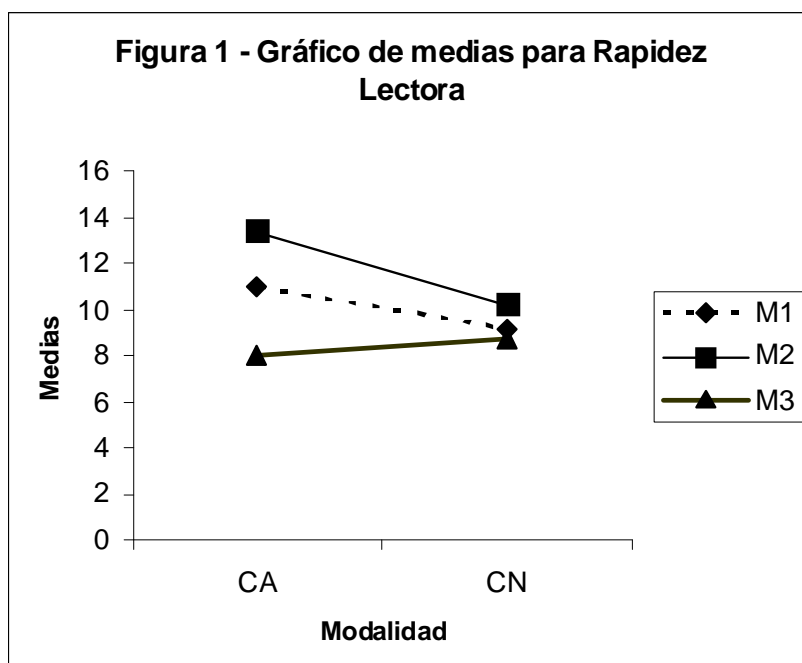
Pregunta 2:

Las medias marginales de las filas ¿coinciden con el promedio de las medias de sus respectivas filas? ¿Y las de las columnas? ¿Por qué?

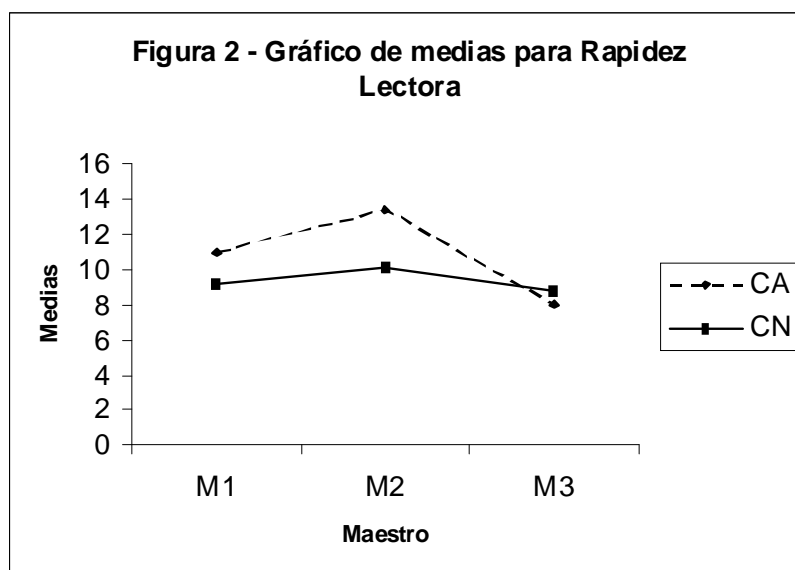
Pregunta 3:

Los desvíos estándar marginales de las filas ¿coinciden con las medias de los desvíos de las correspondientes filas? ¿Por qué?

Comparar los resultados entre maestros o entre modalidades de clase se reduce, en una simplificación del problema, a comparar las medias que los representan. Puede ser muy útil apoyarse en un gráfico de las mismas.



Los resultados son mejores para el maestro 2, luego para el maestro 1 y en último lugar para el maestro 3. Esto sucede en ambas modalidades de clase aunque las diferencias son notoriamente más pronunciadas en CA que en CN. Luego, según la muestra, parece haber un efecto del factor “maestro” dentro de cada modalidad. Recíprocamente se observa un efecto de la modalidad dentro de cada maestro, más pronunciado para los maestros 1 y 2 que para el maestro 3. Esto se aprecia por las pendientes de cada segmento (las primeras son decrecientes y la otra creciente) o bien graficando las medias de las modalidades para cada maestro a saber:





Este gráfico muestra más claramente que con clases abreviadas los maestros 1 y 2 obtuvieron mejores resultados que con las no abreviadas. El maestro 3 obtuvo una pequeña diferencia a favor de CN.

Dado que las diferencias entre los maestros no son iguales dentro de cada modalidad ni recíprocamente, se puede decir que sobre la variable R hay un efecto de la “combinación” de ambos factores: *maestro y modalidad de la clase*. Cuando esto ocurre se dice que hay “interacción” de los factores.

Como una manera de constatar si entendió estos conceptos le proponemos que responda a las siguientes preguntas: ¿Cómo sería la figura 1 si...

Pregunta 4: ... no hubiera efecto de la interacción pero si del maestro y de la modalidad de la clase?

Pregunta 5: ...hubiera efecto del factor maestro pero no de la modalidad de la clase?

Pregunta 6: ...no hubiera efecto del maestro pero sí de la modalidad?

Pregunta 7: ... no hubiera efecto de ninguno de ellos?

Resumiendo...

A partir de la inspección “ocular” de las medias podemos concluir sobre esta muestra que:

- 1) El maestro 2 obtiene mejores resultados siguiéndole el maestro 1 y luego el maestro 3, en ambas modalidades.
- 2) Los resultados entre los maestros difieren menos en las clases no abreviadas que en las abreviadas por lo que hay un efecto de la interacción maestro – modalidad.
- 3) Con las clases abreviadas se obtienen mejores resultados, excepto para el maestro 3.

Estas conclusiones ¿serán aplicables más allá de esta muestra? En otras palabras ¿se obtendrán resultados parecidos con otros estudiantes de similares características o, aun con los mismos, si éstos fueran asignados de otro modo a los distintos grupos? (Recordar que la asignación fue aleatoria). Si la respuesta fuera afirmativa podría concluirse, por ejemplo, que el maestro 2 se desempeña mejor que el 3 en la modalidad de clases abreviadas. La pregunta podría reformularse así: las diferencias observadas ¿son atribuibles a los factores considerados en el experimento o son atribuibles al azar; es decir, particulares de la muestra seleccionada? Para responder a esta pregunta hay criterios que proporcionan las técnicas de inferencia estadística algunas de las cuales, como anticipamos en el comentario del ejercicio 1, veremos al final de este curso. Por ahora sólo podemos tener un cierto indicio mirando las diferencias entre las medias y teniendo en cuenta el desvío estándar: cuanto mayor sea la diferencia y menor la variabilidad se podrá estar más seguro de que dicha diferencia es atribuible a una fuente sistemática. Por ejemplo: para CA se puede sospechar que efectivamente el M2 es mejor que el M3 porque hay una diferencia “amplia” de 5,4 unidades con un desvío “pequeño” de poco más de una unidad en ambos casos. La diferencia entre M1 y M2 bien podría deberse al azar puesto que es de 2,4 unidades con un desvío cercano a 2 para M1. Sin embargo, no hay indicios tan fuertes de que los maestros difieran en modalidad cuando sus clases no son abreviadas (CN); ya que las diferencias son mucho menos pronunciadas y los desvíos no tanto más pequeños, por tanto podrían deberse al azar y no al maestro.

Es claro que este comentario tiene ciertas imprecisiones ¿cuándo puede una diferencia de medias considerarse “amplia” o un desvío “pequeño”? Estas imprecisiones se superan con técnicas adecuadas desarrolladas por la teoría matemática de la inferencia estadística. Sin embargo es importante tener un “ojo entrenado” al abordar un problema aunque no se debería fiar sólo de él

## RESPUESTAS

### EJERCICIO 1

- Personas que viven solos:  $D_5 = 39,25$  min.
- Personas que viven con su familia:  $D_5 = 28,43$  min.

La diferencia entre los valores del quinto decil indica una tendencia de las personas que viven solas a estar más conectadas a Internet que aquellos que viven con sus familias. Esta afirmación en principio es válida sólo para los individuos de la muestra. Sin embargo, es claro que el objetivo de este tipo de estudios no es concluir sobre los individuos en particular que participaron de la investigación sino en general sobre una población más amplia. Habría que ver si la diferencia entre los deciles se explica, al menos parcialmente, por vivir o no con una familia o es casual (por azar); es decir: cualquier otra muestra de individuos ¿arrojaría resultados semejantes? Esto se vincula directamente con el tema de la variabilidad: si la diferencia entre los deciles variara poco de muestra en muestra, con la mayoría de las muestras se afirmarían lo mismo: que las personas que viven solas tienden a pasar más tiempo conectadas a Internet; con lo cual, vivir solo o con la familia sería una fuente sistemática de variación; pero si los deciles variaran mucho, muestras distintas podrían arrojar resultados contradictorios y la afirmación perdería validez general.

### EJERCICIO 2

Grupo Enfermeros:  $Q_2 = 3$  preguntas respondidas.

Grupo Psiquiatras:  $Q_2 = 1$  pregunta respondida.

Los resultados no difieren mucho de los que se hubieran obtenido de haber tratado a la variable como continua según lo explican Botella, León y San Martín (1993, pág 87) – 3,44 y 1,25 - pero son más exactos ya que se usa toda la información original. Cuando los datos se agrupan en intervalos se pierde información. Esto era útil para facilitar los cálculos cuando no se disponía de medios computacionales como los de hoy en día, donde ya no se justifica.

La diferencia encontrada al comparar los segundos cuartiles está en la misma dirección de la hipótesis formulada por el investigador. Sin embargo, se debe ser prudente en tanto que esta afirmación tiene un alcance limitado en su posibilidad de generalización. Al igual que en la respuesta del Ejercicio 1, sólo describen los datos pertenecientes a la muestra tomada por el investigador.

### EJERCICIO 3

- a) 18,44%      b)  $C_{90} = 78,1$

### EJERCICIO 4

Ansiedad Leve $Q_1 = 64,28$	Ansiedad Moderada $Q_3 = 97,41$	Ansiedad Extrema $C_{99} = 117,48$
--------------------------------	------------------------------------	---------------------------------------

### EJERCICIO 5

En esta muestra, las mujeres tienen un mayor porcentaje de riesgo suicida.  
Mujeres: 20% - Varones: 10%

### EJERCICIO 6

- 1) Serán seleccionados 20 operarios por grupo.

2)

- b) Grupo novato:  $Q_1 = 6,9$ ;  $Q_2 = 10$ ;  $Q_3 = 12,5$
- c) En los dos grupos la mitad de los valores se ubican a izquierda y derecha de medidas “aproximadas” lo cual no da evidencias de que los años de experiencia de un obrero en su puesto favorecen el cumplimiento de las normas de seguridad y mantenimiento de las maquinarias.

*Nota:* ¿Cuándo dos medidas son “aproximadas”? ¿Por qué una diferencia de 0,5 en este caso puede considerarse “pequeña”? Nuevamente, como aclaramos en la respuesta del ejercicio 1, habría que tener en cuenta la variabilidad. Una variabilidad pequeña podría ser el producto del azar y no de la incidencia de los años de experiencia. Es necesario precisar cuándo una diferencia es “suficientemente pequeña” o “demasiado grande” para no ser explicada por el azar. Estas precisiones se obtienen mediante el empleo de técnicas estadísticas adecuadas, algunas de las cuales veremos hacia el final de este curso.

### EJERCICIO 7

Puntaje mínimo:  $Q_1 = 8,54$     Puntaje máximo:  $Q_3 = 13,4$ .

### EJERCICIO 8

a)

Cantidad de visitas	$n_i$
0	3
1	5
2	6
3	3
4	3
5	5
6	3
7	2

- b) Se despreocupará del 33,33% de los casos.
- c) Debe ser asistido psicológicamente todo anciano que tenga a lo sumo 2 visitas semanales.
- d) Serán seleccionados 12 ancianos.

### EJERCICIO 9

a)

Grupo Tratamiento	Media = 45 seg	Mediana = 45 seg	Moda aprox= 47 seg
Grupo Control	Media = 45 seg	Mediana = 44,08 seg	Moda aprox= 42 seg

b) Elegimos la media por ser una variable cuantitativa que no ha tomado valores extremos no compensados que pudieran distorsionarla. Al ser igual en ambos grupos concluimos que el uso de anestesia en las operaciones parece no influir en el rendimiento.

### EJERCICIO 10

Grupo Tratamiento	Media = 35 seg	Mediana = 35 seg	Moda = 37 seg
Grupo Control	Media = 35 seg	Mediana = 34,08 seg	Moda = 32 seg

### EJERCICIO 11

Maestros = 45	Bomberos = 36	Abogados = 41	Médicos = 47
---------------	---------------	---------------	--------------

En función de los datos analizados, el grupo de médicos percibe un mayor nivel de estrés laboral.

## EJERCICIO 12

a)

Prueba informal	Media = 41,2	Mediana = 44	Moda = 49
Prueba Formal	Media = 27,8	Mediana = 24,9	Moda = 22

b) Las medidas de tendencia central resultaron mucho mayores para la prueba informal que para la formal; es decir, la cantidad de problemas bien resueltos resultó en general mayor cuando éstos se plantearon en contexto real.

## EJERCICIO 13

a)

Alta densidad de público	Media = 90,375 pulsaciones/min	Baja densidad de público	Media = 84,3 pulsaciones/min
--------------------------	--------------------------------	--------------------------	------------------------------

b) La densidad de público parece ser una fuente sistemática de variación, siendo en promedio mayor el número de pulsaciones por minuto ante alta densidad de público.

## EJERCICIO 14

Media =  $170/30 = 5,67$

## EJERCICIO 15

- a) La categoría modal es 2 y la categoría mediana: es el par (2,3)  
b) Salario modal = 1.500, Salario mediano = \$1.600 (lo tomamos como representante del intervalo (1.500,1.700)). Salario medio = \$1.934,88  
c) i) La moda. ii) La media. iii) La mediana.

## EJERCICIO 16

- a) Las modas son: “La Nación” para el barrio de Belgrano, el diario “Clarín” para las distribuciones de Flores y Mataderos.  
b) Con la entropía.  
c) El barrio de Flores.

## EJERCICIO 17

- a) “Acuerdo” en ambos casos.  
b) En el caso de los menores de 25, pues más de la mitad de la población tiene esa opinión. Hay más concentración de opiniones en la moda que para los mayores. Cuando el número de clases es pequeño como en este caso, una gran concentración en la moda va acompañada de una menor entropía (queda poco por repartir entre las clases restantes).  
c) La correspondiente a los mayores de 25.  
d) Diagrama de rectángulos adyacentes.

## EJERCICIO 18

a)

Grupo Inactivo	AT = 11	Media = 25,2	S = 3,87
Grupo Activo	AT = 11	Media = 24	S = 3,63

- b) Los resultados van en la dirección de la hipótesis de que el nivel de actividad influye en la estimación subjetiva del tiempo transcurrido, en el sentido de que los inactivos tienden a percibirlo como mayor. Sin embargo para inferir esto más allá de la muestra en cuestión hay que completar el análisis con un contraste de hipótesis.
- c) En esta muestra el grupo inactivo ha resultado levemente más variable tanto en términos absolutos (desviación estándar) como relativos a la media; ya que los coeficientes de variación son 15,36 y 15,12 respectivamente. Para determinar si esta diferencia aparentemente leve es estadísticamente significativa también se debería llevar a cabo el correspondiente contraste de hipótesis.

### EJERCICIO 19

- a) La transformación aplicada no altera los resultados de AT ni de S pero sí de la media, la cual queda incrementada en dos unidades: 27,2 y 26.
- b) Incrementar los valores en un 50% es lo mismo que aplicar la transformación  $Y=1,5X$ ; por lo que los resúmenes estadísticos resultarían:

Grupo Inactivo	AT = 16,5	Media = 37,8	S = 5,805
Grupo Activo	AT = 16,5	Media = 36	S = 5,445

### EJERCICIO 20

	Media		Varianza		Desviación Estándar	
	a)	b)	a)	b)	a)	b)
Grupo 1	7,29	72,9	2,06	206	1,44	28,8
Grupo 2	8,14	81,4	1,12	112	1,06	21,2

### EJERCICIO 21

Test A	Media = 26,2	S = 8,96	CV = 34,22
Test B	Media = 15	S = 7	CV = 46,67

Al comparar los coeficientes de variación se concluye que el desempeño de Karina ha sido más homogéneo en el test de estímulos visuales.

### EJERCICIO 22

- a) Obtuvo mejor resultado el grupo del terapeuta A ya que su media es 4,4; o sea menor que la del grupo del terapeuta B.
- b) No puede afirmarse esto porque:
- No se sabe si ambos grupos tenían el mismo nivel de dificultad al momento de comenzar el tratamiento ya que no se hizo un pretest. Tampoco fueron aleatorizados como para suponer una distribución homogénea en los grupos.
  - Aunque los dos grupos estuvieran en condiciones semejantes de partida habría que investigar si la diferencia entre las medias es suficientemente grande en relación a su variabilidad de muestra en muestra.

### EJERCICIO 23

- a) No, porque los desvíos estándar no son directamente comparables por estar en unidades diferentes. Sin embargo, puesto que ambas variables corresponden al

nivel de razón, se puede comparar la variabilidad en relación a la media mediante el coeficiente de variación, que es adimensional. Éste es 4,1 para la estatura y 10,67 para el peso; luego son más variables los pesos que las estaturas, en relación a la media.

- b) No es lo más adecuado comparar los desvíos en términos absolutos sino en relación a la media cuando el nivel de medición lo permite. Como la variable “estatura” corresponde al nivel de razón, podemos calcular el coeficiente de variación.

Para los padres	CV = 4,1	La variabilidad relativa es mayor en los niños.
Para los hijos	CV = 6	

## EJERCICIO 24

- a) En el segundo parcial se obtuvo mejor rendimiento que en el primero pues, visualmente, se observa que la distribución es asimétrica negativa para el segundo parcial y asimétrica positiva para el primero.
- b) El primer parcial discriminó más a los alumnos de mejor rendimiento, el segundo discriminó más a los de peor rendimiento.
- c) Para discriminar a los de mejor y peor rendimiento por igual la forma de la distribución debería ser simétrica.
- d) Los exámenes cuyos puntajes toman valores a lo largo de casi toda la escala de evaluación y tienen una distribución asimétrica negativa son más fáciles que los que tienen una distribución asimétrica positiva; por tanto los primeros discriminan mejor a los individuos de bajo rendimiento y los segundos a los de alto rendimiento.

## EJERCICIO 25

Si al inicio del curso los puntajes se distribuyen con asimetría positiva son pocos los alumnos que obtendrán los mayores puntajes y no tan pocos los que podrían sacar una nota baja. La intención sería incentivar al alumno desde el inicio del curso para que se esfuerce en no reprobado o en merecer notas excelentes. En cambio en un examen final interesa que las calificaciones bajas de reprobación sean infrecuentes y no tanto las excelentes; es decir que sea más difícil la reprobación que el premio de una nota alta; para que si alguien reprueba se tenga la tranquilidad de que es porque realmente lo merece al destacarse por su mal rendimiento.

## EJERCICIO 26

	CA		CN		Total	
	R	C	R	C	R	C
M1	$\bar{X} = 11$ s = 1,79	$\bar{X} = 21,2$ s = 1,33	$\bar{X} = 9,2$ s = 0,98	$\bar{X} = 15,8$ s = 1,17	$\bar{X} = 10,1$ s = 1,7	$\bar{X} = 18,5$ s = 2,97
M2	$\bar{X} = 13,4$ s = 1,36	$\bar{X} = 24,8$ s = 1,47	$\bar{X} = 10,2$ s = 1,17	$\bar{X} = 16,8$ s = 1,17	$\bar{X} = 11,8$ s = 2,04	$\bar{X} = 20,8$ s = 4,21
M3	$\bar{X} = 8$ s = 1,09	$\bar{X} = 17,2$ s = 1,33	$\bar{X} = 8,8$ s = 0,75	$\bar{X} = 19$ s = 1,67	$\bar{X} = 8,4$ s = 1,02	$\bar{X} = 18,1$ s = 1,78
Total	$\bar{X} = 10,8$ s = 2,64	$\bar{X} = 21,1$ s = 3,39	$\bar{X} = 9,4$ s = 1,14	$\bar{X} = 17,2$ s = 1,90	$\bar{X} = 10,1$ s = 2,15	$\bar{X} = 19,1$ s = 3,35

Pregunta 1: Las medias del margen derecho son las medias de los puntajes correspondientes a todos los alumnos de cada maestro, tomados globalmente sin discriminar la modalidad de clase. El desvío estándar corresponde al desvío de los puntajes de dichos alumnos para cada maestro. Análogamente, las medias y desvíos por columnas lo son de los alumnos correspondientes a cada modalidad sin tener en cuenta el maestro. La media y el desvío del último cuadro son las que se obtienen a partir de los 30 puntajes.

Pregunta 2: Las medias marginales pueden obtenerse promediando las medias por fila y por columna ya que cada celda corresponde a la misma cantidad de estudiantes: 5.

Pregunta 3: Los desvíos estándar no coinciden con los promedios de los desvíos por fila y por columna por razones de álgebra elemental; entre otras, la falta de distributividad de los cuadrados y de las raíces para la suma.

Pregunta 4: Los segmentos deberían ser paralelos no coincidentes y no horizontales.

Pregunta 5: Los segmentos deberían ser paralelos no coincidentes y horizontales.

Pregunta 6: Los segmentos deberían ser coincidentes no horizontales.

Pregunta 7: Los segmentos deberían ser coincidentes horizontales.

## REFERENCIAS

- Botella, J., León, O. y San Martín, R. (1993). *Análisis de Datos en Psicología I*. Madrid: Ediciones Pirámide.
- Cano de Becerra, F. (1971). Elementos de estadística al servicio de la evaluación del rendimiento. *Revista de Psicología*, 16, 61 - 77.
- Caraher, T., Carraher, D. y Schlieman, A. (1998). *En la vida diez, en la escuela cero: los contextos culturales del aprendizaje de las matemáticas*. México: Siglo XXI.
- Miguel-Tobal J. J. y Cano-Vindel A. R. (1998). *Inventario de Situaciones y Respuestas de Ansiedad (ISRA)*. Manual (2da edición). Madrid: TEA.
- Rosenhan, D. (1988). Acerca de estar sano en un medio enfermo. En Watzlawick, P. (comp.). *La realidad inventada ¿cómo sabemos lo que creemos saber?* (pp. 99-120). España: Gedisa.

# PRÁCTICA DE INTEGRACIÓN ESTADÍSTICA DESCRIPTIVA

## EJERCICIO 1

Un fabricante de calzado deportivo está interesado en imponer su marca en el mercado a través de una efectiva campaña publicitaria dirigida a los jóvenes. Para ello le encarga a un psicólogo especializado en grupos motivacionales testear dos propuestas publicitarias de su marca de manera tal de elegir aquella con mayor impacto.

El psicólogo convoca a 30 jóvenes de 18 a 25 años de ambos sexos, residentes en Ciudad de Buenos Aires y los divide al azar en dos grupos de 15. Al grupo 1 le presenta tres publicidades, dos de las cuales pertenecen a la competencia, y la tercera la propuesta publicitaria a testear que llamaremos A. Al grupo 2 le presenta las dos mismas publicidades de la competencia y la propuesta publicitaria a testear que llamaremos B. Luego solicita a cada grupo que le asigne una posición en el ranking a las publicidades testeadas con relación a las publicidades de la competencia. Los resultados obtenidos fueron:

Ranking	Grupo 1	Grupo 2
	Propuesta A	Propuesta B
1	9	5
2	4	6
3	2	4

- a) i) Indique quiénes constituyen la población de individuos y quiénes la muestra de individuos  
 ii) Determine quiénes constituyen las poblaciones de observaciones (¿son reales o hipotéticas?) y quiénes las muestras de observaciones.
- b) Mencione la variable estadística, clasifíquela e indique el nivel de medición utilizado.
- c) ¿Observa alguna tendencia en las observaciones que den cuenta de la existencia de una fuente sistemática de variación? En caso afirmativo explique dicha tendencia y mencione la fuente sistemática de variación.
- d) Represente con un diagrama adecuado la comparación entre las dos propuestas.

## EJERCICIO 2

Los siguientes datos corresponden a la cantidad de errores cometidos por un niño disléxico en la lectura de 15 párrafos breves de diferentes textos.

3	4	7	3	4	5	3	4	6	5	2	1	2	4	5
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

- a) Presente la información en una tabla de frecuencias absolutas y calcule las medidas de tendencia central, los cuartiles, el desvío estándar.
- b) ¿En qué porcentaje de ocasiones el niño cometió como máximo cuatro errores?

## EJERCICIO 3

Un investigador está realizando un estudio sobre la producción académica de los profesores pertenecientes a Universidades privadas y públicas de la ciudad de Buenos Aires. Uno de los indicadores que toma para evaluar la producción es la número de participaciones en Congresos nacionales e internacionales durante el último año. Para ello toma una muestra aleatoria de 30 profesores de universidades privadas y 15 de universidades públicas y obtiene los resultados que siguen:.



Número de participaciones en Congresos	Número de Profesores	
	Universidades Privadas	Universidades Públicas
0	6	-
1	10	4
2	7	2
3	4	3
4	2	3
5	-	2
6	-	1
7	1	-

- a) i) Indique quiénes constituyen la población de individuos y quiénes la muestra de individuos  
 ii) Determine quiénes constituyen las poblaciones de observaciones (¿son reales o hipotéticas?) y quiénes las muestras de observaciones.
- b) Mencione la variable estadística, clasifíquela e indique el nivel de medición utilizado.
- c) Explique qué tendencia siguen las observaciones que da cuenta de la existencia de una fuente sistemática de variación y mencione dicha fuente. ¿Todas las observaciones responden a lo previsto? ¿A qué atribuye esto? Explique.
- d) Indique cuál es el gráfico adecuado para representar la distribución de frecuencias para cada tipo de universidad de modo que puedan ser comparadas.
- e) ¿Qué porcentaje de profesores participó por lo menos de dos Congresos en el último año?

#### EJERCICIO 4

A continuación presentamos una tabla con la distribución de frecuencias de las calificaciones finales en Lengua de niños de 7° grado del turno mañana y también algunos resúmenes estadísticos correspondientes al turno tarde.

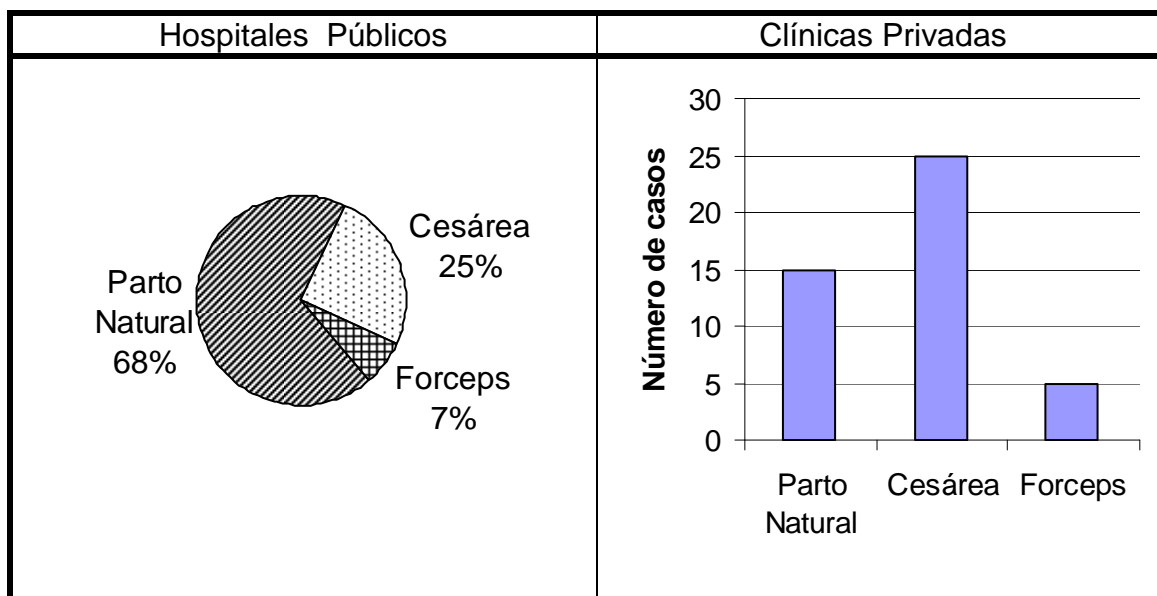
Turno Mañana	
Calificación	$n_i$
8,5 – 10,5	2
6,5 – 8,5	7
4,5 – 6,5	8
2,5 – 4,5	2
0,5 – 2,5	1

Turno Tarde
$\bar{X} = 5,01$ $Md = 5,4$ $s = 1,86$ $n = 25$

- a) ¿Qué grupo parece haber tenido mejor rendimiento? ¿Por qué?
- b) ¿En qué grupo hay mayor variabilidad relativa a la media?

#### EJERCICIO 5

Se realizó una investigación con el objetivo de conocer las modalidades de partos que se dan en hospitales públicos y clínicas privadas. Para ello se seleccionó una muestra aleatoria de 100 madres que dieron a luz en hospitales y clínicas emplazadas en jurisdicción del Gobierno de la Ciudad de Buenos Aires durante el mes de setiembre de 2006. A continuación se presenta la información obtenida:



- Determine la población y la muestra de individuos.
- Mencione la variable de interés y clasifíquela. Mencione qué nivel de medición se utilizó e indique los valores de la escala.
- Mencione la posible fuente sistemática de variación.
- ¿Qué tipo de gráficos se han utilizado para representar la información de ambas distribuciones?

### EJERCICIO 6

Con la misma información que se presenta en el Ejercicio 5:

- Determine qué cantidad de mujeres corresponde a la muestra de hospitales públicos y clínicas privadas, respectivamente, observando ambos gráficos.
- Presente en dos tablas cada una de las distribuciones de frecuencias, especificando en cada caso las absolutas, relativas y porcentuales.
- ¿Con qué medidas compararía las dos distribuciones con respecto a su variabilidad y a su tendencia central? Aplíquelas y extraiga conclusiones.

### EJERCICIO 7

Un grupo de investigadores estudia la incidencia de la cafeína sobre la memoria auditiva. Para ello reunieron dos grupos, uno denominado “experimental” conformado por 15 personas, a las que se les suministró 600 miligramos de cafeína. Otro, denominado “de control”, conformado por 26 personas, a las que no se les suministró sustancia alguna. Ambos grupos fueron sometidos a una prueba de memoria auditiva, los resultados se presentan a continuación:

Grupo Experimental	Cantidad de palabras olvidadas	$n_i$
	10	2
	9	6
	8	5
	7	0
	6	2

Grupo Control	$\bar{X} = 6$
	$Md = 6,5$
	$Mo = 6$
	$s^2 = 1,4$
	$n = 26$

Mediante el cálculo de los estadísticos adecuados responda y justifique:

- ¿Parece influir el consumo de cafeína en la memoria auditiva a juzgar por los resultados muestrales?
- ¿Qué grupo tiene menor variabilidad relativa?
- Calcule el tercer cuartil de la distribución del grupo de experimental.
- Los investigadores desean seleccionar a aquellos individuos del grupo experimental que no hayan superado el primer cuartil. ¿Cuántas personas cumplen esta condición?

### EJERCICIO 8

Un establecimiento educativo utiliza el cociente intelectual de Binet – Stern para ubicar a los alumnos aspirantes al ingreso en tres grupos. El grupo 1 estará integrado por aquéllos que no superen el centil 15 y el grupo 3 tiene como mínimo puntaje un CI de 129. A continuación se presenta la distribución de frecuencias correspondiente:

CI	$n_i$
136 - 150	8
121 - 135	24
106 - 120	36
91 - 105	11
76 - 90	1

- Calcule el porcentaje de alumnos que ingresarán al grupo 3.
- ¿Cuál será la puntuación mínima para ingresar al grupo 2?
- Si los directivos desean seleccionar a quienes superen el centil 95 o no superan el centil 5 ¿Cuántos alumnos serán elegidos?

### EJERCICIO 9

Las autoridades de nuestra facultad están interesadas en conocer el tiempo que tardan los alumnos en llegar al establecimiento ya sea desde sus hogares o lugares de trabajo. Se solicita al alumno que considere el lugar más frecuente desde el cual se traslada para llegar a la facultad y que mida el tiempo requerido. Para tal fin, fueron seleccionados al azar 20 alumnos de los tres turnos que concurren a la sede de HY y 20 alumnos de los tres turnos que concurren a la sede de Independencia. A continuación se presenta la distribución de frecuencias de una sede y algunos estadísticos correspondientes a la otra:

Sede Hipólito Yrigoyen	Tiempo (minutos)	
		$n_i$
	60 – 74	1
	45 – 59	6
	30 – 44	3
	15 – 29	8
0 – 14	2	

Sede Independencia
$\bar{X} = 35$ $s^2 = 81$ $n = 20$

- Mencione la población y muestra de individuos.
- Mencione la población y muestra de observaciones.
- Mencione la variable de interés y clasifíquela. Indique qué nivel de medición se utilizó.

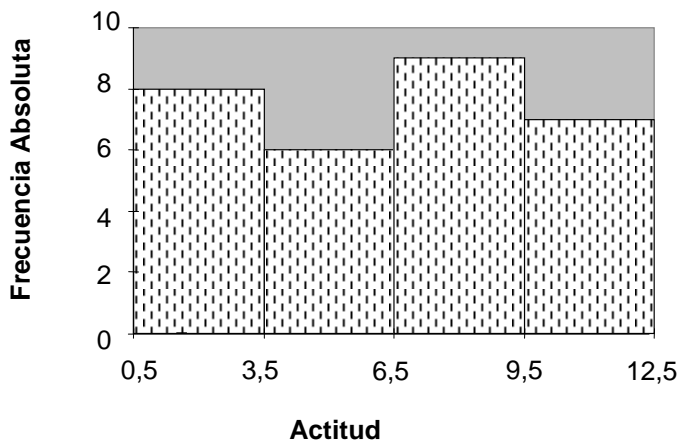
## EJERCICIO 10

A partir de la información suministrada en el Ejercicio 9 determine:

- Para qué sede el tiempo que tardan los alumnos en llegar es mayor. Justifique.
- Cuál de las dos medias es más representativa. Justifique la respuesta.

## EJERCICIO 11

Un grupo de 30 mujeres jubiladas elegidas aleatoriamente de los Centros de Jubilados del barrio de Flores respondió a una escala de actitud hacia el consumo de medicamentos no prescritos por un médico. Esta escala permite discriminar a los sujetos propensos a consumir de manera abusiva si superan los 11 puntos. El gráfico resume los resultados obtenidos.



- Mencione la población y la muestra de individuos.
- Mencione la población y muestra de observaciones.
- Mencione la variable, clasifíquela conforme al tratamiento estadístico que se le ha dado e indique el nivel de medición.

## EJERCICIO 12

A partir de la información del Ejercicio 11,

- Construya la tabla de distribución de frecuencias que incluya: frecuencias absolutas, relativas, porcentuales y acumuladas (absolutas, relativas y porcentuales).
- Indique qué otro u otros gráficos podría utilizar para representar la distribución de frecuencias de esta variable.

## EJERCICIO 13

A partir de la información de los Ejercicios 11 y 12 responda:

- ¿Cuántas mujeres de la muestra resultaron ser propensas a consumir medicamentos de manera abusiva?, teniendo en cuenta que esta escala permite discriminar a los sujetos propensos a consumir de manera abusiva si superan los 11 puntos.
- Si una muestra de 30 mujeres que concurren a Centros de Jubilados del barrio de Balvanera presentan una media en la escala de actitud hacia el consumo de medicamentos no prescritos por un médico de 4,3, indique cuál de las dos muestras presenta una actitud más proclive al consumo de medicamentos no prescritos. Justifique la respuesta.

## EJERCICIO 14

Un investigador ha relevado la cantidad de horas que miran televisión los fines de semana los niños residentes en Ciudad de Buenos Aires, para lo cual tomó una muestra aleatoria de 16 niños. Otro investigador realizó el mismo relevamiento pero sobre una muestra de niños del Gran Buenos Aires. A continuación se presentan los datos de ambos investigadores:

Ciudad de Buenos Aires	Cantidad de horas	
	12 – 14	2
	9 – 11	4
	6 – 8	5
	3 – 5	3
	0 – 2	2

Gran Buenos Aires	
	$\bar{X} = 6,3\text{hs.}$
	$s = 2,2\text{hs.}$
	$Mo = 6,8\text{hs}$
	$Md = 6,2\text{hs}$
	$n = 20$

- ¿Qué grupo mira en general mayor cantidad de horas de televisión? ¿Por qué?
- Si se comparan las dos muestras ¿cuál presenta mayor variabilidad relativa a la media?

## RESPUESTAS

### EJERCICIO 1

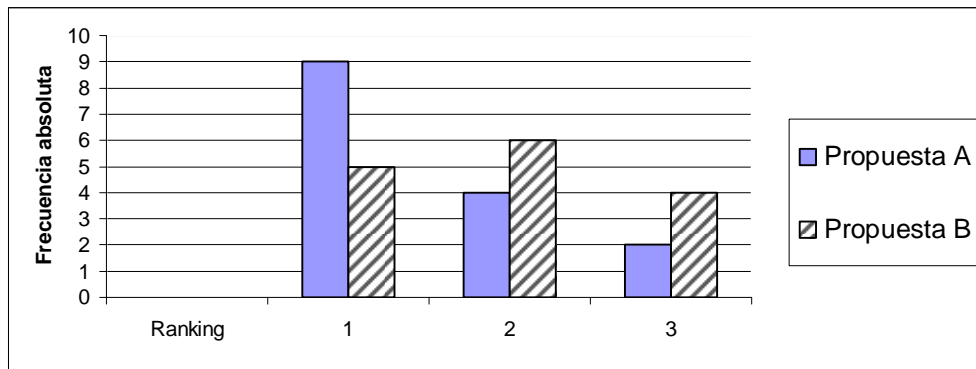
a) i. Población de individuos: Totalidad de jóvenes de 18 a 25 años de ambos sexos residentes en Ciudad de Buenos Aires. Muestra de individuos: los 30 jóvenes de 18 a 25 años de ambos sexos residentes en Ciudad de Buenos Aires seleccionados para la experiencia.

ii) Poblaciones de observaciones: una población es el conjunto de todos los rankings que los jóvenes de 18 a 25 años de ambos sexos de Ciudad de Buenos Aires le asignarían a la publicidad de haber sido expuestos a la propuesta A. La otra población es la de los rankings que dichos jóvenes le asignarían a la propuesta publicitaria B. Por cómo están definidas estas dos poblaciones, son claramente hipotéticas. Muestra de observaciones: son también dos y corresponden a cada población hipotética. Están constituidas por los rangos asignados efectivamente por los jóvenes de 18 a 25 años de ambos sexos de Ciudad de Buenos Aires seleccionados para la experiencia que vieron la publicidad tipo A y por los rangos asignados por los jóvenes que vieron la publicidad de tipo B.

b) La variable estadística es "Ranking asignado a la publicidad entre las dos marcas competidoras". Es cuasicuantitativa y el nivel de medición utilizado es ordinal.

c) Más de la mitad de los jóvenes que recibieron la propuesta A le asignaron el primer lugar a la publicidad. El grupo de jóvenes que evaluó la propuesta B presenta opiniones más heterogéneas, no observándose ninguna posición privilegiada para esta propuesta. De modo que la tendencia es posicionar mejor la publicidad de la propuesta A. Por tanto el tipo de propuesta publicitaria parece ser una fuente sistemática de variación.

d)



## EJERCICIO 2

a)

$X_i$	1	2	3	4	5	6	7
$n_i$	1	2	3	4	3	1	1

$\bar{X} = 3,87$	$Mo = 4$	$s = 1,54$
$Md = 4$	$Q_1 = 3$	$Q_3 = 5$

b) 66,67%

## EJERCICIO 3

a)

i. Población de individuos: La totalidad de profesores de universidades privadas y públicas de la ciudad de Buenos Aires. Muestra de individuos: los 45 profesores de universidades privadas y públicas de la Ciudad de Buenos Aires.

ii. Las poblaciones de observaciones: son dos, una es la formada por todos los números que indican cuántas veces participó de un Congreso en el último año cada profesor de Universidad privada de la Ciudad de Buenos Aires. La otra población está definida de la misma manera para los profesores de universidades públicas. Son poblaciones reales. Hay dos muestras de observaciones que se corresponde con cada una de las dos poblaciones de observaciones: una es la de los 30 números que representan cuántas veces participó de un Congreso en el último año cada uno de los profesores de Universidades privadas y la otra es la definida de manera análoga sobre los 15 profesores de Universidades públicas.

b) La variable estadística es "Número de participaciones en Congresos nacionales e internacionales realizados por los profesores en el último año". Es una variable cuantitativa discreta y el nivel de medición es de razón.

c) Para poder comparar ambos grupos es necesario usar las frecuencias porcentuales o relativas, ya que son dos muestras de diferente tamaño. Puede observarse que las frecuencias porcentuales tienden a acumularse en los valores más bajos de la variable para el caso de los profesores de Universidades privadas y hacia los valores más altos en el otro grupo. Por lo que el tipo de Universidad parece ser una fuente sistemática de variación, los profesores de las Universidades privadas tienden a participar menos que los de la Universidad pública en Congresos. No todas las observaciones responden a esta tendencia, ya que vemos que el único profesor que más ha participado entre todos (7 presentaciones en Congresos) es de la Universidad privada. Esto puede atribuirse a fuentes fortuitas de variación.

d) Hay que realizar un diagrama de barras adyacentes usando frecuencias porcentuales.

e) El 55,56%.

#### EJERCICIO 4

- El turno mañana tiene mejor rendimiento ya que, resumiendo sus puntajes con las medidas de tendencia central, vemos que son mayores que en el turno tarde: Media = 6,2 y Md = 6,25.
- La mayor variabilidad relativa a la media se da en el turno tarde, pues el coeficiente de variación es 37% mientras que en el turno mañana es de 31%.

#### EJERCICIO 5

- Población de Individuos: todas las madres que han dado a luz en hospitales públicos y clínicas privadas emplazadas en jurisdicción del Gobierno de la Ciudad de Buenos Aires, durante el mes de setiembre de 2006. Muestra de Individuos: 100 madres que han dado a luz en hospitales públicos y clínicas privadas emplazadas en jurisdicción del Gobierno de la Ciudad de Buenos Aires durante el mes de setiembre de 2006.
- La variable es "Modalidad de parto", es de tipo cualitativa. Se ha empleado una escala nominal cuyos valores son: "Parto Natural", "Cesárea" y "Fórceps".
- Una posible fuente sistemática de variación es: Tipo de establecimiento donde se realizó el parto: hospitales públicos y clínicas privadas.
- Para representar la distribución del Hospital Público se ha utilizado un pictograma en tanto para la de la Clínica Privada un diagrama de rectángulos.

#### EJERCICIO 6

- Hospitales públicos: 55 mujeres; Clínicas Privadas: 45 mujeres
- 

$X_i$	Hospitales Públicos			Clínicas Privadas		
	$n_i$	$p_i$	$P_i\%$	$n_i$	$p_i$	$P_i\%$
Parto Natural	37	0,68	68	15	0,33	33,3
Cesárea	14	0,25	25	25	0,56	55,6
Fórceps	4	0,07	7	5	0,11	11,1
	55	1	100	45	1	100

- Como se trata de una variable medida en escala nominal la única medida de tendencia central que es posible emplear es la Moda. Para la distribución de los Hospitales Públicos la Moda es "Parto Natural" y para la distribución de las Clínicas Privadas es "Cesárea". La variabilidad de las distribuciones puede ser comparada con la Entropía. Como es posible observar, la moda de la distribución de los Hospitales Públicos concentra un mayor número de casos (68%) que la moda de las Clínicas Privadas (55,6%), esto nos indica que esta última es la distribución con mayor entropía durante el mes de setiembre de 2006.

#### EJERCICIO 7

- Al comparar las medias obtenidas por ambos grupos se encuentra que el Grupo Experimental (media = 8,4) ha tenido un peor rendimiento que el Grupo Control (media = 6). Si bien no nos es posible determinar aún si esta diferencia es significativa, los datos muestrales apuntan en la dirección de la hipótesis de que el consumo de cafeína afecta negativamente el rendimiento de la memoria.
- El Grupo Experimental presenta menor variabilidad relativa ya que su CV es 13,6%, en tanto que el Grupo Control tiene un CV de 19,7%.
- $Q_3 = 9$  palabras olvidadas.
- $Q_1 = 8$  y, en consecuencia, serán 7 las personas las seleccionas (2 + 5 personas no superan las 8 palabras olvidadas).

## EJERCICIO 8

- a) El 23% de los alumnos ingresarán al grupo 3.
- b) El puntaje mínimo para ingresar al grupo 2 es un CI de 105,5.
- c) Deben seleccionar a 8 niños.

## EJERCICIO 9

- a) Población de Individuos: dos poblaciones de individuos, todos los estudiantes que asisten a la sede de HY y todos los estudiantes que asisten a la sede de Independencia de la Facultad de Psicología de la UBA. Muestra de individuos: dos muestras de individuos, 20 estudiantes que asisten a la sede de HY y 20 estudiantes que asisten a la sede de Independencia de la Facultad de Psicología de la UBA.
- b) Población de observaciones: hay dos, el conjunto de números que representan al tiempo expresado en minutos que tardan en llegar (desde sus casas o lugares de trabajo) los estudiantes que asisten a la sede de HY y el conjunto de números que representan al tiempo expresado en minutos que tardan en llegar (desde sus casas o lugares de trabajo) los estudiantes que concurren a la sede de Independencia, de la Facultad de Psicología, UBA.  
Muestra de observaciones: hay dos, el conjunto de números que representan al tiempo expresado en minutos que tardan en llegar (desde sus casas o lugares de trabajo) los 20 alumnos que concurren a la sede de HY y el conjunto de números que representan al tiempo expresado en minutos que tardan en llegar (desde sus casas o lugares de trabajo) los 20 alumnos que concurren a la sede de Independencia, Facultad de Psicología, UBA.
- c) Tiempo que tarda un estudiante en llegar a la Facultad de Psicología (sede HY o Independencia) desde su hogar o lugar de trabajo. Variable cuantitativa continua. Nivel de medición de cociente o razón.

## EJERCICIO 10

- a)  $\bar{X} = 34$  en HY. Los alumnos tardan más en llegar a la sede de Independencia pues el tiempo medio correspondiente a esta sede es mayor, 35 minutos.
- b) En HY: el  $s = 16,84$  y el  $CV = 49,52\%$ . En Independencia:  $s = 9$  y el  $CV = 25,7\%$ . La media correspondiente a la sede de Independencia es más representativa porque existe menor variabilidad relativa.

## EJERCICIO 11

- a) Población individuos: las mujeres jubiladas que concurren a los Centros de Jubilados del barrio de Flores. Muestra de individuos: 30 mujeres jubiladas que concurren a los Centros de Jubilados del barrio de Flores.
- b) Población de observaciones: el conjunto de puntajes que podrían obtener en esta escala todas las mujeres jubiladas que concurren a los Centros de Jubilados del barrio de Flores. Muestra de observaciones: el conjunto de puntajes obtenidos en esta escala por las 30 mujeres jubiladas que concurren a los Centros de Jubilados del barrio de Flores.
- c) La variable es: "Puntaje en la actitud hacia el consumo de medicamentos no prescritos por un médico". Se le ha dado un tratamiento de variable continua medida en un nivel de medición utilizado es al menos intervalar; al no explicarse cómo se ha contruido la escala para saber si el origen es convencional o no, no se sabe si puede ser de razón.



## EJERCICIO 12

a)

$X_i$	$N_i$	$n_a$	$p_i$	$p_a$	$P_i \%$	$P_a \%$
10 - 12	7	30	0.23	1	23	100
7 - 9	9	23	0.3	0.77	30	77
4 - 6	6	14	0.2	0.47	20	47
1 - 3	8	8	0.27	0.27	27	27
Total	30	-	1	-	100	-

b) Polígono de frecuencias y Ojiva de Galton (polígono de frecuencias acumuladas)

## EJERCICIO 13

- a) Calculo el rango percentilar correspondiente a 11 puntos: 88%. Luego,  $100\% - 88\% = 12\%$ . El 12% de 30 mujeres es igual a  $3.6 \sim 4$ . Hay 4 mujeres de la muestra propensas a consumir medicamentos no prescritos por un médico de manera abusiva.
- b) El promedio de puntaje en la escala entre las mujeres jubiladas del barrio de Balvanera es 4.3, mientras que entre las jubiladas del barrio de Flores es de 6.5. Estas últimas presentan una actitud más favorable hacia el consumo de medicamentos no prescritos.

## EJERCICIO 14

- a)  $\bar{X} = 7,19$  (Ciudad de Buenos Aires). En la muestra de Ciudad de Buenos Aires se mira mayor cantidad de horas televisión dado que la media en esta localidad es mayor que en Gran Buenos Aires (6.3).
- b) En Ciudad de Buenos Aires el  $s = 3.59$  y el  $CV = 49.98\%$ . En Gran Buenos Aires:  $s = 2.2$  y el  $CV = 34.92\%$ . En la distribución correspondiente a la Ciudad de Buenos Aires existe mayor variabilidad relativa a la media.

# PRÁCTICA IV

## EJERCICIO 1

Complete la siguiente tabla a partir de la información exhibida.

	TEST A			TEST B			TEST C		
	Puntaje	Media	Desv. Est.	Puntaje	Media	Desv. Est.	Puntajes	Media	Desv. Est.
Bruto	65	45	8	40		12	36	18	
Z				-1				-2	
T							60		
CI									

## EJERCICIO 2

Indique cuáles de las siguientes afirmaciones son verdaderas y cuáles no.

- En una prueba deportiva de velocidad se registra la distancia en metros recorrida por ciertos atletas en un minuto. En esta prueba Andrés obtuvo un puntaje  $Z = -0,2$ . Posteriormente se exhibió la distancia recorrida por cada atleta en centímetros. Entonces, en esta escala, el puntaje  $Z$  de Andrés resultó ser  $-20$ .
- Los puntajes  $T$  siempre se hallan comprendidos entre 20 y 80.
- La valoración de las puntuaciones tipificadas, en principio, sólo es posible en términos relativos al grupo de referencia y no en términos absolutos.
- Los puntajes  $Z$  siempre se hallan comprendidos entre  $-1$  y  $1$ .
- Al medir en segundos el tiempo de reacción a un estímulo, Catalina obtiene un puntaje  $Z = 0,5$ . Eso significa que tardó en reaccionar medio segundo más que la media del grupo normativo.
- Las puntuaciones tipificadas son adimensionales y por eso permanecen invariantes si se cambian las unidades de la escala.

## EJERCICIO 3

Analía tiene un puntaje  $Z = 2$  en Perfeccionismo y  $T = 35$  en Afabilidad, dos escalas del Cuestionario Factorial de Personalidad 16PF. Los puntajes  $Z$  y  $T$  de Gonzalo en dichas escalas son respectivamente  $-1,5$  y  $20$ . En estas pruebas, puntajes más altos representan mayores niveles del rasgo medido. Indique cuáles de todas las siguientes afirmaciones son verdaderas y cuáles falsas.

- El puntaje directo de Analía en Perfeccionismo está dos puntos por encima de la media del grupo normativo.
- Gonzalo es menos perfeccionista y afable que la media del grupo normativo.
- El puntaje directo de Gonzalo es más alto en Afabilidad que en Perfeccionismo.
- Analía es tan afable como Gonzalo perfeccionista.
- No es posible comparar dos rasgos tan diferentes como son el perfeccionismo y la afabilidad.
- Gonzalo es más perfeccionista que afable.
- El puntaje bruto de Gonzalo en Perfeccionismo se halla a una desviación estándar y medio por debajo de la media del grupo normativo.

# RESPUESTAS

## EJERCICIO 1

	TEST A			TEST B			TEST C			
	Puntaje	Media	Desv. Est.	Puntaje	Media	Desv. Est.	Puntajes		Media	Desv. Est.
Bruto	<b>65</b>	<b>45</b>	<b>8</b>	<b>40</b>	52	<b>12</b>	<b>36</b>	<b>18</b>	30	6
Z	2,5	0	1	<b>-1</b>	0	1	1	<b>-2</b>	0	1
T	75	50	10	40	50	10	<b>60</b>	30	50	10
CI	137,5	100	15	85	100	15	115	70	100	15

## EJERCICIO 2

Solamente son verdaderas c) y f).

## EJERCICIO 3

Solamente son verdaderas b), d), f) y g).

# PRÁCTICA V

## EJERCICIO 1

Indicar cuáles de las siguientes afirmaciones sobre el coeficiente de correlación lineal  $r$  de Pearson son siempre verdaderas y cuáles son siempre falsas y cuáles contingentes.

- a) Está comprendido entre  $-1$  y  $1$ .
- b) Es adimensional.
- c) Si es  $0$ , las variables son independientes.
- d) Si  $r = 0,07$  para las variables peso y longitud de un objeto expresados en kg y cm respectivamente el valor del coeficiente pasa a ser  $r = 0,7$  para las mismas variables expresadas en gramos y metros.
- e) La correlación lineal entre el tiempo empleado en completar una tarea y el número de errores cometidos en la misma es  $r = -0,6$ . De ello se deduce que una de las causas que contribuyen importantemente en el número de errores cometidos en la tarea es el tiempo que se emplea en resolverla.

## EJERCICIO 2

En la siguiente tabla se informan los valores de  $r$  y las características de los diagramas de dispersión en cinco situaciones. Establezca la correspondencia correcta entre ambas columnas.

Valores de $r$	Los puntos del diagrama de dispersión...
a) 1	1) tienden a ubicarse en torno a una recta creciente.
b) -1	2) están distribuidos en el plano sin ninguna estructura particular.
c) 0	3) tienden a ubicarse en torno a una recta decreciente.
d) 0,9	4) están perfectamente alineados en una recta creciente.
e) -0,7	5) están perfectamente alineados en una recta decreciente.

### EJERCICIO 3

En la siguiente tabla se informan los valores de  $r$  y los pares de variables estudiadas en cinco situaciones. Establezca la correspondencia entre ambas columnas que podría esperarse suponiendo que los diseños de investigación fueron adecuados.

Valores de $r$	Variables estudiadas
a) - 0,4	1) Puntaje en ansiedad medido con el test A y con el test B.
b) 0,8	2) Estatura de ciertos sujetos y su puntaje en razonamiento verbal.
c) 0,9	3) Puntaje en un ítem dicotómico de un test de habilidades matemáticas y el puntaje en el resto del test.
d) 0,02	4) Puntaje en un test de Bienestar y puntaje en un test de Estrés.
e) 0,3	5) Puntaje en los ítems pares y puntaje en los ítems impares de un test de analogías verbales.

### RESPUESTAS

#### EJERCICIO 1

a) Siempre verdadera.      b) Siempre verdadera.      c) Contingente (pueden ser independientes o puede haber una relación que no se aproxima a la linealidad).

d) Siempre falsa.      e) Contingente (si hay un buen diseño de la experiencia puede darse una interpretación causal).

#### EJERCICIO 2

a)	4)
b)	5)
c)	2)
d)	1)
e)	3)

#### EJERCICIO 3

a)	4)
b)	1)
c)	5)
d)	2)
e)	3)

# PRÁCTICA VI

## EJERCICIO MODELO I

Un profesor evalúa a sus alumnos con una prueba de elección múltiple de 10 ítems de Verdadero / Falso. El docente considera que debe promocionar a un sujeto cuando ha respondido acertadamente a, por lo menos, 7 de los 10 ítems. Así también, considera “desaprobado” cuando responde correctamente a menos de 4 ítems.

I. ¿Cuál es la probabilidad de que un alumno que responde azarosamente

- a) promocióne la materia?
- b) conteste incorrectamente a alguna pregunta?
- c) apruebe la materia sin promocionar?

II. En una población hipotética de alumnos que contestara todo por azar; qué porcentaje se espera que

- a) promocióne.
- b) apruebe.

III. ¿Cuál es el número promedio de respuestas correctas que se espera obtener por azar?

IV. ¿Qué cantidad de respuestas correctas debería exigir el profesor para aprobar el examen de modo que a lo sumo el 6% de los alumnos que contestara por azar aprobara?

V. ¿Cómo cree que podría lograrse una menor probabilidad de aprobación por azar manteniendo la longitud de la prueba y el punto de corte original; esto es, 4.

VI. Recalcule las probabilidades anteriores suponiendo que cada ítem, en lugar de ser de Verdadero/Falso, tiene 5 opciones con una sola correcta.

### RESOLUCIÓN

Sea  $X$  = Cantidad de respuestas correctas habiendo contestado los 10 ítems al azar.

$X \sim B(10; 0,5)$

I. a) Promocionar es equivalente a que la cantidad de respuestas correctas sea por lo menos 7; en símbolos:  $X \geq 7$ .

$$P(X \geq 7) = 0,117 + 0,044 + 0,010 + 0,001 = \underline{0,172}$$

b) Contestar incorrectamente alguna pregunta es lo opuesto a contestar todo bien. Contestar todo bien, en símbolos, es  $X = 10$ . Luego,

$$P(\text{tener alguna respuesta incorrecta}) = 1 - P(X = 10) = 1 - 0,001 = \underline{0,999}$$

c) Aprobar la materia sin promocionar es obtener como mínimo 4 y no más de 6. En símbolos:  $4 \leq X \leq 6$ . Luego,  $P(4 \leq X \leq 6) = 0,205 + 0,246 + 0,205 = \underline{0,656}$

II. a)  $P(X \geq 7) = 0,172 \Rightarrow$  El 17,2% de los alumnos que contestaran todo por azar promocionaría.

b) Aprobar es sacarse por lo menos 4; es decir aprobar sin promocionar o con promoción directa; por lo que se pueden sumar las probabilidades a) y c) de la parte I y resulta:  $P(X \geq 4) = 0,172 + 0,656 = 0,828 \Rightarrow$  El 82,8% de tales alumnos aprobaría.

III. Como  $p = 0,5$ , el número promedio de respuestas por azar es la mitad; esto es  $10 \times 0,5 = 5$ .

*Nota:* El promedio de los valores de una variable Binomial de parámetros  $n$  y  $p$  es  $n.p$  y la varianza  $n.p.(1-p)$ .

IV. Llamaremos  $r$  al punto de corte para aprobar. Hay que hallar  $r$  de modo tal que  $P(X \geq r) \leq 0,06$  o, lo que es equivalente, para que lo contrario, que es desaprobado, ocurra con 0,94 de probabilidad; en símbolos:  $P(X < r) \geq 0,94$  (Medite en esta equivalencia). Para ello se busca en la tabla el primer valor de la variable donde, sumando las probabilidades desde 0 hasta ese valor, se excede 0,96. Eso ocurre en el valor 7. Por tanto la respuesta es que habría que exigir 8 respuestas correctas.

V. La probabilidad de aprobación por azar puede disminuirse si a su vez se disminuye la probabilidad de acierto por azar en cada ítem; lo que puede lograrse aumentando el número de opciones.

VI. Ahora  $X \sim B(10; 0,2)$  ya que la probabilidad de acertar por azar es  $1/5 = 0,2$ .

I. a)  $P(X \geq 7) = 0,001 + 0^+ + 0^+ + 0^+ = \underline{0,001}$

b)  $1 - P(X = 10) = 1 - 0^+ = \underline{1}$

c)  $P(4 \leq X \leq 6) = 0,088 + 0,026 + 0,006 = \underline{0,12}$

II. a)  $P(X \geq 7) = 0,001 \Rightarrow 0,1\%$

b)  $P(X \geq 4) = 0,121 \Rightarrow 12,1\%$

III.  $10 \times 1/5 = 2$

IV. Hallar  $r$  tal que  $P(X < r) \geq 0,94 \Rightarrow r = 5$

## EJERCICIO MODELO II

### Aplicación de la distribución binomial en una investigación real

José Luis Marcos Malmierca y Alfonso Barca Lozano (2009). Efecto del nivel de consciencia de la presencia del estímulo sobre el aprendizaje de expectativa *Psicothema*, 21,3,397-402.

El artículo completo se halla en [www.psicothema.com/pdf/3646.pdf](http://www.psicothema.com/pdf/3646.pdf). A continuación se extraen algunos párrafos, se describe parte de la experiencia y se plantea un problema con su resolución.

#### *Participantes* (Textual)

La muestra inicial estaba compuesta por 92 estudiantes de las titulaciones de Logopedia y Terapia Ocupacional, que participaron de modo voluntario tras haber sido convenientemente informados sobre el experimento.

#### *Procedimiento* (Textual)

c) Fase de evaluación de la consciencia. Para evaluar el grado de consciencia del estímulo enmascarado E1 se efectuaron 80 ensayos de una tarea de identificación de

elección forzada, bajo las mismas condiciones visuales y de duración que en las fases anteriores. El estímulo enmascarado E1 consistía en la letra «o» en la mitad de los ensayos y en la letra «x» en la otra mitad. Inmediatamente después de la presentación de cada secuencia el participante debía indicar qué letra había sido presentada («o» o «x»), pulsando las correspondiente teclas «O» y «X» del teclado del ordenador.

## PROBLEMA

Los autores proceden luego a clasificar a los participantes teniendo en cuenta el nivel de conciencia del estímulo enmascarado de acuerdo con el desempeño que tuvieron en la tarea de elección forzosa. El criterio que adoptan para ello es clasificar como semiconscientes o concientes a los sujetos cuya cantidad de identificaciones correctas superan cierto umbral que es poco probable (menos de 0,05), de alcanzar por azar; el resto es clasificado como no conciente.

¿Cuál es el umbral (o punto de corte) para dicha clasificación?

## RESOLUCIÓN

$X$  = Número de identificaciones correctas por azar entre las 80 presentaciones.

$X \sim B(80; 0,5)$

Debemos hallar  $r$  tal que  $P(X > r) < 0,05$ , lo cual es equivalente buscar el primer valor de  $r$  donde la probabilidad acumulada supera a 0,95. Ese valor es  $r = 48$ .

Veamos ahora lo que dicen los autores

### *Medida y análisis (Textual)*

Los participantes que en la tarea de detección de elección forzada mostraban unos resultados de detección correcta inferior al 60% eran asignados al grupo no-consciente, ya que, según una distribución binomial, con  $N = 80$  (ensayos) y  $p$  (identificación correcta)=0.5, la probabilidad de que ocurran 48 (60%) o más de 48 identificaciones correctas es de 0.046. Aplicando ese criterio, fueron asignados a este grupo no-consciente 32 sujetos. Los participantes con identificaciones correctas que oscilaban entre 48 (60%) y 64 (80%) fueron asignados al grupo semiconsciente, que quedó constituido por otros 32 sujetos. Finalmente, otros 28 participantes que habían logrado superar 64 detecciones correctas (+80%) conformaron el grupo consciente.

## EJERCICIO MODELO III

Al estudiar el efecto de la edad sobre el recuerdo, medido como el número de imágenes recordadas, Simón et al (2009) informan que la media y la desviación estándar de los puntajes para los jóvenes son respectivamente 24,77 y 4,32 mientras que para los mayores estas medidas son 16,05 y 3,9. Suponiendo que el puntaje en recordación se distribuye normalmente en cada población con los parámetros mencionados calcule

- La probabilidad de que un joven elegido al azar tenga un puntaje inferior al de la mediana de los mayores.
- El porcentaje de personas mayores que tienen un puntaje superior a 20,5.
- El puntaje mínimo que debe tener una persona mayor para pertenecer al 25% del grupo con mejores puntajes.
- La probabilidad de que en una muestra de 6 personas mayores elegidas al azar al menos una tenga un puntaje superior a 21,042.

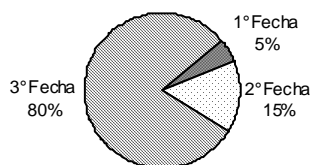


## RESOLUCIÓN

- a)  $X_J$  = Puntaje en recordación de un joven elegido al azar.  $X_J \sim N(24,77; 4,32)$   
Como los puntajes de los mayores son normales la mediana es igual a la media; en este caso 16,05.  
 $P(X_J < 16,05) = P[(X_J - 24,77)/4,32 < (16,05 - 24,77)/4,32] = P(Z_J < -2,02) = \underline{0,0218}$ .
- b)  $X_M$  = Puntaje en recordación de una persona mayor elegida al azar.  
 $X_M \sim N(16,05; 3,9)$   
 $P(X_M > 20,5) = P[(X_M - 16,05)/3,9 > (20,5 - 16,05)/3,9] = P(Z_M > 1,14) = \underline{0,1271}$ .
- c) Dicho puntaje mínimo corresponde al tercer cuartil de la distribución. En la distribución normal estándar, éste es  $z_{0,75} = 0,67$ . Desestandarizando resulta  
 $x_{0,75} = 16,05 + 0,67 \cdot 3,9 = 18,66$
- d) Sea  $Y$  = Número de personas en la muestra de 6 que tiene un puntaje superior a 21,042. Entonces  $Y \sim B(6; p)$  donde  $p = P(X_M > 21,042) = P(Z_M > 1,28) = 0,10$ .  
 $P(Y \geq 1) = 1 - P(Y = 0) = 1 - 0,5314 = \underline{0,4686}$ .

## EJERCICIO 1

Estadísticas recientes de nuestra facultad afirman que de cada 10 inscriptos para rendir finales en diciembre sólo 3 se presentan. El siguiente gráfico reproduce cómo se distribuyen los alumnos que se presentan en las tres fechas tentativas:



A partir de la información brindada

I. calcule la probabilidad de que de una muestra aleatoria de 12 alumnos inscriptos

- a) más de 4 y menos de 8 no se presenten a rendir.  
b) por lo menos 4 se presenten.

II. calcule la probabilidad de que de una muestra aleatoria de 6 alumnos que se presentan

- a) a lo sumo 1 no se presente en la tercera fecha.  
b) ninguno se presente en la primera fecha.

## EJERCICIO 2

Un grupo de 4 estudiantes que han preparado juntos el final de Psicoanálisis Freud II decidieron analizar la posibilidad que tienen de aprobarlo. Recurrieron a la encuesta que publica la cátedra en su página de Internet y hallaron los siguientes datos:

En diciembre	Se presentaron	Aprobaron	Desaprobaron
1998	654	591	63
1999	699	643	56
2000	729	670	59
2001	619	578	41

Ellos están interesados en saber cuál es la probabilidad de que

- a) apruebe todo el grupo.  
b) alguno desaprobe.

Estime estas probabilidades para responder a sus inquietudes.

### EJERCICIO 3

Una prueba de percepción visual consiste en responder a cuál de los dos matices, previamente presentados, corresponde el color verde que aparece sucesivamente en 12 estampados diferentes. El resultado de la prueba se considera “muy satisfactorio” cuando se ha respondido acertadamente en, por lo menos, 9 de las 12 presentaciones.

En prácticas previas un sujeto ha logrado acertar el 80% de las representaciones. ¿Cuál es la probabilidad de que al realizar la prueba

- obtenga el resultado “muy satisfactorio”?
- no obtenga un resultado “muy satisfactorio” suponiendo que contestó completamente al azar en todos los casos?
- Si se hicieran sólo 8 presentaciones, ¿cuál debería ser el punto de corte para clasificar un resultado como “muy satisfactorio” si se desea que la probabilidad de obtenerlo contestando por azar sea pequeña, digamos inferior a 0,05?

### EJERCICIO 4 (Datos ficticios)

En cierta región de una provincia argentina Ignacio (sociólogo) y Lucía (maestra) están estudiando el fenómeno de la deserción escolar a nivel primario. De una encuesta del último año a varias escuelas obtuvieron la siguiente información:

Causas de deserción	Nro. de casos
Necesidad de trabajar	180
Distancia a la escuela	40
Enfermedad	55
Dificultades intelectuales	20
Otros	5

Habiendo iniciado el año lectivo en dichas escuelas un total de 6000 alumnos.

Ignacio se pregunta qué probabilidad tendrá de que entre 12 niños que abandonen la escuela el año entrante, más de la tercera parte lo haga debido a la necesidad de trabajar. Lucía se pregunta con qué probabilidad concluirá el año conservando a sus 15 alumnos.

Suponiendo que el año entrante tenga un comportamiento similar al anterior con respecto al fenómeno de la deserción, responda a

- Ignacio.
- Lucía.

### EJERCICIO 5 (Datos ficticios)

Resultados de la selección para los integrantes de la Nueva Policía Bonaerense

**2 de cada 5 postulantes para integrar la Policía tienen trastornos psicológicos**

Se trata de 1500 aspirantes que perdieron la posibilidad de integrarse a las fuerzas policiales por presentar un nivel de agresividad muy elevado y falta de sensibilidad social. La mayoría son hombres y casi la mitad pertenece a clase baja. El 80% afirmó que ser policía es su verdadera vocación. A su vez, 450 de ellos declararon que su ingreso promedio es inferior al que ganarían como policía.

A partir de la información brindada

I. calcule la probabilidad de que de una muestra aleatoria de 6 aspirantes a integrar la nueva policía bonaerense

- a) por lo menos 5 tengan trastorno psicológicos.
- b) más de la mitad no hayan presentado trastorno psicológicos.

II. calcule la probabilidad de que de una muestra aleatoria de 5 postulantes con trastornos psicológicos

- a) a lo sumo 3 no elijan esta profesión por vocación.
- b) ninguno posea un ingreso inferior al que aspira.

## EJERCICIO 6

Gastón sabe que el 95% de las personas logran responder correctamente una pregunta de aritmética de un test de inteligencia. Pero interroga a sus colegas acerca de la cantidad de tiempo que tardaron aquellos pacientes que acertaron con su respuesta

Respuesta Correcta	Consulta 1	Consulta 2	Consulta 3
Antes de 30seg.	34	18	44
Después de 30seg.	83	42	98

I. Suponiendo que Gastón administra el test a 14 personas calcule la probabilidad de que a) alguno responda incorrectamente.

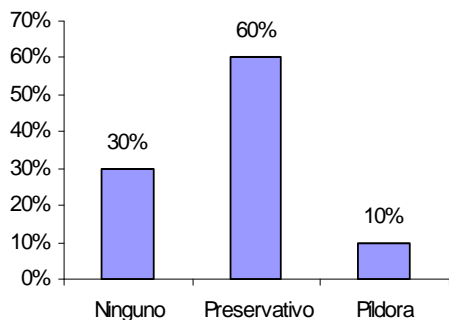
- b) por lo menos 13 respondan correctamente.

II. Suponiendo que 16 personas respondan correctamente calcule la probabilidad de que a) más del 75% de ellos lo hagan después de 30 segundos.

- b) ninguno lo haga antes de 30 segundos.

## EJERCICIO 7 (Datos ficticios)

El caso de un joven de 14 años que se convirtió en padre motivó la publicación de una estadística aseguró que 1 de cada 5 adolescentes se inician sexualmente antes de cumplir 15 años. Consultados por el método anticonceptivo empleado por ellos se ha podido confeccionar el gráfico.



A partir de la información brindada

I. calcule la probabilidad de que de una muestra aleatoria de 9 adolescentes

- a) menos de 6 hayan iniciado sexualmente antes de los 15 años.
- b) la mayoría no se haya iniciado sexualmente antes de cumplir los 15 años.

II. calcule la probabilidad de que de una muestra aleatoria de 5 jóvenes que se hayan iniciado antes de los 15 años

- a) a lo sumo 4 hayan usado algún método anticonceptivo.
- b) el 40% haya usado preservativo.

## EJERCICIO 8

Suponiendo que los puntajes en una prueba de razonamiento verbal se distribuyen normalmente con media 12 y desvío estándar 3 en un grupo normativo, calcule:

- El rango percentilar de un sujeto que obtiene 14 puntos.
- El puntaje correspondiente al sujeto que es superado por el 90% del grupo.
- El centil 95.
- La probabilidad de que en una muestra aleatoria de 10 sujetos, la mayoría tenga un puntaje superior a 14,52.

## EJERCICIO 9

Para ingresar a cierta universidad se debe rendir un examen de ingreso y sólo hay cupo para el 20% de los aspirantes. Suponiendo que los puntajes se distribuyen normalmente en el grupo de aspirantes con media 60 y desviación estándar 15 .

- ¿Qué nota mínima de aprobación debe fijarse?
- ¿Cuál es la probabilidad de que en una muestra aleatoria de 15 aspirantes más de la tercera parte aprueben el examen de ingreso?

## EJERCICIO 10

La medicina tradicional considera normal dormir 7,5 horas con un desvío de 0,5 horas para mujeres de 40 años. Una cantidad de horas inferior al centil 15 y superior al centil 95 se considera riesgosa para este grupo poblacional por considerarse el insomnio y la hipersomnia dos indicadores claros de trastornos depresivos. Sobre una base de 160 mujeres estudiadas responder:

- ¿Cuáles son las cantidades máxima y mínima de horas de sueño que debe dormir una mujer para pertenecer al grupo de riesgo?
- Si una paciente duerme 8 horas ¿a cuántas personas supera?
- ¿Cuántas horas duerme una mujer ubicada en el último centil?
- ¿Cuál es la probabilidad de que una mujer duerma entre 7 y 8 horas?

## EJERCICIO 11

Los puntajes de un test que mide el grado de expresión del enojo de un sujeto se distribuye normalmente con media 25 y desvío 2 en una población argentina. La teoría califica como "irritable" a las personas que están a menos de un desvío y medio por encima de la media y como "violento" a las personas que exceden dos desvíos y medio por encima de la media. Sobre una base de 120 personas evaluadas responda:

- ¿Qué porcentaje de sujetos cabe esperar en el grupo "violento"?
- ¿Cuántos sujetos cabe esperar en el grupo "irritable"?
- ¿Cuál es el puntaje correspondiente al tercer decil?

## EJERCICIO 12

El estrés ocupacional percibido por los 200 empleados que realizan atención de cliente en una empresa fue medido con el Inventario de Estrés Ocupacional (OSI) de Osipow y Spokane. Esta variable se distribuye normalmente y utilizan los puntajes T para establecer el nivel de estrés.

Nivel de Estrés	Referencias
Fuerte	$T > 69,5$
Leve	$59,5 < T < 69,5$
Normal	$39,5 < T < 59,5$
Ausencia	$T < 39,5$

Calcular:

- ¿En cuántos empleados cabe esperar que estén ausentes los indicadores de estrés?
- ¿Qué porcentaje de sujetos percibirán un leve nivel de estrés?
- ¿Qué puntuaciones como mínimo y como máximo será necesario establecer si se desea seleccionar al 50% de los sujetos que obtienen las puntuaciones centrales?
- ¿Qué nivel de estrés percibe un sujeto que es superado por el 60% de la población?

### EJERCICIO 13

De los 200 alumnos que rindieron un examen para obtener una beca, sólo 44 han podido recibirla por superar los 65 puntos. Si las calificaciones de estos alumnos están distribuidas normalmente con una desviación típica de 6 puntos ¿cuál es el promedio de esta población?

### EJERCICIO 14

El tiempo necesario para concluir con el armado de un rompecabezas de 100 piezas se distribuye normalmente con media 210 minutos. Sabiendo que el 45% de los sujetos tarda como máximo 185 minutos ¿cuál es la desviación estándar de esta distribución?

### EJERCICIO 15 (Resuelto)

Una prueba de rendimiento consiste en 50 ítems de 4 opciones de las cuales sólo una es verdadera. El requerimiento para aprobar es contestar correctamente por lo menos el 40% de la prueba, esto es, 20 ítems.

En este tipo de pruebas interesa conocer qué probabilidad tiene de aprobar un alumno que contesta todo por azar.

Si definimos la variable  $X$  = Cantidad de respuestas acertadas entre las 50 habiendo contestado todas por azar

- ¿Qué modelo de distribución de probabilidades es adecuado para la variable  $X$ ? Justifique.
- Calcule la probabilidad de aprobar por azar.
- Halle la nota mínima de aprobación para que apruebe por azar a lo sumo el 2% de los alumnos.

a) Por la manera en que la variable está definida y por las características de la experiencia el modelo más adecuado es el binomial de parámetros  $n = 50$  y  $p = 0,25$ . En efecto, contestar todos los ítems de la prueba por azar puede pensarse como repetir  $n = 50$  veces un ensayo de Bernoulli, ya que es contestar un ítem donde sólo caben dos posibilidades: responderlo acertadamente (éxito) o incorrectamente

(fracaso). Al responder completamente al azar los ítemes resultan independientes, ya que todos tienen la misma cantidad de opciones (cuatro) y sólo una es verdadera. Por tanto, se cumplen las dos condiciones de éxito constante en todos los ensayos.

b) Como se aprueba con 20 ítemes o más respuestas correctas debemos calcular  $P(X \geq 20)$ . La probabilidad exacta es la suma de las probabilidades binomiales desde 20 hasta 50:

$$P(X \geq 20) = P(X = 20) + P(X = 21) + P(X = 22) + P(X = 23) + \dots + P(X = 50)$$

Sumar tantos términos es tedioso y actualmente hay calculadoras que dan directamente el resultado de tales sumas. También hay tablas de probabilidades “acumuladas” que dan esta información. Pero para quien no tenga estos recursos puede apelar al uso de la distribución normal, la cual da una buena aproximación de las probabilidades binomiales acumuladas cuando se cumplen ciertas condiciones de simetría que no detallaremos aquí. Llamaremos  $X_N$  a la variable normalmente distribuida que usaremos para aproximar las probabilidades de la variable binomial  $X$ . Sus parámetros serán la media y desviación estándar de  $X$ , o sea:

$$\text{Media de } X = n.p = 50 \cdot \frac{1}{4} = 12,5 = \text{Media de } X_N$$

$$\text{Desviación estándar de } X = \sqrt{n.p.(1-p)} = \sqrt{50 \cdot \frac{1}{4} \cdot \frac{3}{4}} = 3,06 = \text{Desvío estándar de } X_N$$

Resumiendo:

$$X_N \sim N(12,5 ; 3,06)$$

Para aproximar la probabilidad binomial por la normal debe hacerse una corrección que consiste en considerar media unidad más; es decir que

$$P(X \geq 20) = P(X_N \geq 19,5)$$

Esto se llama corrección por continuidad y es necesaria cuando a una variable que es discreta se la trata como continua. (El lector interesado en una mayor explicación puede remitirse al cuadro 13.4, pág. 325 de Botella)

$$P(X_N \geq 19,5) = P\left[\frac{X_N - 12,5}{3,06} \geq \frac{19,5 - 12,5}{3,06}\right] = P(Z \geq 2,29) = 0,011 \text{ donde } Z \sim N(0;1)$$

Este resultado puede interpretarse así: de cada 100 alumnos que resuelvan toda la prueba por azar, aproximadamente uno solo aprobará.

c) Debemos obtener un puntaje  $x$  de manera tal que:

$$P(X_N \geq x) < 0,02$$

Estandarizando tenemos que:

$$P(X_N \geq x) = P\left[Z \geq \frac{x-12,5}{3,06}\right] < 0,02$$

Buscando en la tabla normal estándar el valor  $z$  que acumula a izquierda una probabilidad de 0,98 hallamos que, aproximadamente:

$$Z > 2,06 \quad \text{o sea que} \quad \left[Z \geq \frac{x-12,5}{3,06}\right] > 2,06$$

y despejando obtenemos

$$x > 3,06 \cdot 2,06 + 12,5 > 18,8036$$

y como la cantidad mínima para aprobar de respuestas correctas es un número entero, habrá que fijar este mínimo en

$$\underline{x = 19}$$

## RESPUESTAS

### EJERCICIO 1

I. a) 0,267    b) 0,507

II. a) 0,655    b) 0,735

### EJERCICIO 2

a) 0,656    b) 0,344

### EJERCICIO 3

a) 0,794    b) 0,927    c) 7

### EJERCICIO 4

a) 0,944    b) 0,463

### EJERCICIO 5

I. a) 0,041    b) 0,545

II. a) 0,994    b) 0,168

### EJERCICIO 6

I. a) 0,512    b) 0,847

II. a) 0,246    b) 0,003

### EJERCICIO 7

I. a) 0,997    b) 0,980

II. a) 0,832    b) 0,230

### EJERCICIO 8

a) 74,86    b) 8,16

c) 16,94    d) 0,007

### EJERCICIO 9

a) 72,6    b) 0,061

### EJERCICIO 10

a) 6,98 y 8,32    b) 134

c) 8,67    d) 0,6826

### EJERCICIO 11

a) 0,62%    b) 52

c) 23,96

### EJERCICIO 12

a) 29,    b) 14,55

c) 43,3 y 56,7    d) 47,5

### EJERCICIO 13

60,368

### EJERCICIO 14

192,31



# PRÁCTICA VII

## EJERCICIO MODELO I

Un psicopedagogo que trabaja en una escuela de arte lleva registros históricos de los puntajes de los alumnos en un test de actitudes hacia la matemática. Según sus registros, la media para los estudiantes que completaron el ciclo básico es de 43 puntos con una desviación estándar de 9. Se propone llevar a cabo un programa de actividades tendientes a mejorar dicha actitud. Pone a prueba su programa con 36 alumnos y al final del período correspondiente les administra el test. Obtiene una media de 46 puntos. a) ¿Puede afirmarse al 5% que el programa respondió a sus objetivos? b) En caso afirmativo evalúe el tamaño del efecto. c) Calcule el nivel crítico de probabilidad o “valor p” e indique qué decisión habría tomado de haber trabajado con un nivel de significación del 1%. d) ¿Cómo mínimo qué valor de la media muestral conduciría a atribuir su diferencia respecto de 43 a que la media poblacional es en realidad mayor que 43 (el programa tuvo efecto) más que al azar; en otras palabras, cuál es el punto crítico a partir del cual la diferencia es estadísticamente significativa?

### RESOLUCIÓN

a) Se sugiere seguir los siguientes pasos:

**1) Mencionar la/s variable/s de interés en la población respecto de la cual se desea hacer inferencia estadística y explicitar los supuestos mínimo necesarios (si los hubiere) para llevar a cabo el contraste.**

X = Puntaje que obtendrían en el test de actitudes hacia la matemática los estudiantes de la escuela de arte donde trabaja el psicopedagogo, una vez completado el ciclo básico, si se implementara el programa. (Nótese que se trata de una población hipotética.)

**2) Plantear las hipótesis por contrastar.**

$H_0: \mu = 43$  El efecto del programa es nulo.  
 $H_1: \mu > 43$  El programa tiene un efecto positivo.

**3) Elegir un nivel de significación.**

$\alpha = 0,05$

**4) Indicar el estadístico de contraste adecuado y explicitar su distribución bajo la hipótesis nula.**

$$E = \frac{\bar{X} - 43}{\frac{9}{\sqrt{36}}} \approx N(0,1)$$

por el Teorema Central del Límite ya que la muestra es grande.

Bajo  $H_0$

**Nota.** Bajo  $H_0$  el efecto del programa es nulo; por lo que también la desviación estándar de X es la misma que la de los puntajes históricos.

### 5) Determinar la zona de rechazo de $H_0$ o “regla de decisión”.

Siendo un contraste unilateral a derecha, con un nivel de significación de 0,05 y la distribución del estadístico de contraste normal, la zona de rechazo está dada por todos los valores mayores que el percentil 95 de la distribución normal estándar; esto es  $z_{0,95} = 1,645$ .

**Regla de Decisión:** Se rechaza  $H_0$  si y sólo si el valor observado del estadístico de contraste es mayor que 1,645.

### 6) Calcular el valor observado del estadístico de contraste.

$$e = \frac{46 - 43}{\frac{9}{\sqrt{36}}} = 2$$

### 7) Tomar la decisión y dar una conclusión en términos del problema planteado.

Como  $2 > 1,645$  se rechaza  $H_0$ ; por lo que se concluye que el programa sería eficaz en el sentido de que los puntajes de actitud hacia la matemática aumentarían en promedio si se implementara el programa.

### b) Tamaño del efecto

$$d = \frac{46 - 43}{9} = 0,33, \text{ lo que se interpreta como un efecto más bien pequeño.}$$

**Nota.** Conviene valorar el tamaño del efecto en términos relativos a la información acopiada en torno al problema estudiado. A falta de tal información, Cohen (1988) sugiere el siguiente criterio orientativo: Valores de  $d$  entre 0,2 y 0,3 indican un efecto pequeño, alrededor de 0,5 un efecto mediano y mayores que 0,8 un efecto alto.

### c) Nivel crítico

$\text{Valor } p = P(E > 2 / \mu = 43) = 1 - 0,9772 = 0,0228 > 0,01$ ; por lo que a ese nivel no se habría rechazado la hipótesis nula.

### d) Punto crítico

En la regla de decisión del paso 5) se estableció el rechazo de  $H_0$  para

$$e > 1,645. \text{ Como } e = \frac{\bar{x} - 43}{\frac{9}{\sqrt{36}}}, \text{ se tiene que se rechaza } H_0 \text{ si y sólo si}$$

$$\frac{\bar{x} - 43}{\frac{9}{\sqrt{36}}} > 1,645. \text{ Despejando,}$$

$$\bar{x} > 43 + 1,645 \cdot \frac{9}{\sqrt{36}} = 45,4675 \Rightarrow \text{valor crítico de la media.}$$

Por tanto la diferencia es estadísticamente significativa al 5% para cualquier valor de la media muestral superior a 45,4675.

## EJERCICIO MODELO II

El artículo completo sobre el que se basa el presente ejercicio puede obtenerse a través del vínculo <http://www.psicothema.com/psicothema.asp?id=3637>.

Para estudiar la eficacia de la terapia de aceptación y compromiso en la mejora del rendimiento ajedrecístico, Ruiz, Francisco y Luciano, Carmen (2009) seleccionaron a 8 jugadores de la Federación Española de Ajedrez a los que se les aplicó un protocolo de intervención durante 5 días de concentración. Para integrar el grupo control seleccionaron doce ajedrecistas de la Federación Internacional de Ajedrez (FIDE) de acuerdo con su similitud con los participantes de la condición experimental. Los participantes del grupo control no recibieron ningún tipo de intervención ni tuvieron contacto con los experimentadores.

Para cada jugador se calculó el puntaje *performance*, según el archivo de la FIDE, a lo largo de 9 torneos jugados contra una computadora antes de la intervención del grupo experimental y 9 torneos posteriores a la misma, siempre que el jugador hubiera ganado al menos 4 de ellos. Finalmente quedaron 7 jugadores en cada grupo.

Los autores afirman textualmente que “se estableció como criterio de cambio una diferencia de más de 40 puntos de *performance* entre el período anterior y posterior a la intervención. La justificación del criterio seleccionado se encuentra en que la obtención de 40 puntos más de *performance* en el cómputo global de un torneo implica, normalmente, que el jugador ascendería una cantidad relevante de puestos en la clasificación final. Además, numerosos ajedrecistas y entrenadores cualificados consultados lo consideraron un criterio de cambio bastante alto”.

Los resultados de los participantes de la condición experimental y control en los 9 torneos previos y posteriores a la intervención se exhiben en la siguiente tabla:

	<b>EXP PRE</b>	<b>EXP POST</b>	<b>CTRL PRE</b>	<b>CTRL POST</b>
P1	2354	2414	2361	2371
P2	2314	2414	2299	2313
P3	2317	2427	2260	2272
P4	2364	2375	2377	2371
P5	2126	2142	2122	2135
P6	2076	2208	2209	2170
P7	2216	2300	2242	2268

- Verifique al 10% que los participantes de ambos grupos están en las mismas condiciones iniciales; es decir, que al comenzar el seguimiento, no difieren en su desempeño ajedrecístico.
- ¿Puede afirmarse al 10% que hubo un aumento importante de la *performance* después de la intervención para el grupo experimental?
- ¿Hubo un aumento estadísticamente significativo al 10% en la *performance* del grupo control?
- ¿Qué puede concluir respecto del programa de intervención?

### RESOLUCIÓN

- Es necesario llevar a cabo un contraste de hipótesis para la diferencia de medias de dos poblaciones a partir de muestras independientes.

### 1) Variables y supuestos mínimo necesarios para llevar a cabo el contraste

$X_1$ : Puntaje en performance de los ajedrecistas de la población teórica representada por la muestra de los que formarán parte del grupo experimental.

$X_2$ : Puntaje en performance de los ajedrecistas de la población teórica representada por la muestra de los que formarán parte del grupo control.

Supuestos:  $X_1$  y  $X_2$  se distribuyen normalmente con igual varianza y son independientes.

### 2) Hipótesis Estadísticas

$H_0: \mu_1 - \mu_2 = 0$  Los sujetos de ambas poblaciones no difieren en promedio en su performance.

$H_1: \mu_1 - \mu_2 \neq 0$  Hay diferencia en la performance promedio entre ambas poblaciones.

### 3) Nivel de significación: $\alpha = 0,10$

**Nota:** La elección de este nivel de significación, más alto que el que generalmente se toma (5%) se justifica en el hecho de que se trata de un contraste para validar una afirmación "No hay diferencias entre ambos grupos", afirmación que se valida si no se rechaza la hipótesis nula. De ser así, se estaría en riesgo de cometer error de tipo II y este nivel permite tener un menor riesgo de cometerlo (mejor potencia).

### 4) Estadístico de contraste y su distribución bajo $H_0$

$$E = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{6S_1^2 + 6S_2^2}{7+7-2} \left(\frac{1}{7} + \frac{1}{7}\right)}} \sim t_{12} \text{ bajo } H_0$$

### 5) Regla de decisión

Al ser un contraste bilateral, se rechaza  $H_0$  si y sólo si  $e > t_{10;0,95}$  o  $e < -t_{10;0,95}$

Buscando en la tabla,  $t_{12;0,95} = 1,812$ .

Luego, se rechaza  $H_0$  si y sólo si el valor observado del estadístico de contraste es mayor que 0.1782 o menor que -0.1782.

### 6) Cálculo del valor observado del estadístico de contraste

$$e = \frac{2252,43 - 2267,14}{\sqrt{\frac{6.13193,95 + 6.7837,14}{7+7-2} \left(\frac{1}{7} + \frac{1}{7}\right)}} = -0,245$$

### 7) Conclusión

Como  $-1,812 < -0,245 < 1,812$  no se rechaza  $H_0$ ; es decir, no hay diferencias estadísticamente significativas entre las medias de los grupos experimental y control; lo que se interpreta como que la diferencia observada entre las medias de ambos grupos es atribuible al azar.

b) Se debe hacer un contraste sobre diferencia de medias de datos pareados para el grupo experimental tomando en cuenta el umbral de 40 puntos.

### 1) Variables y supuestos mínimo necesarios para llevar a cabo el contraste

D: Diferencia de puntaje después y antes de la intervención de los sujetos de la población hipotética representada por la muestra del grupo experimental.

Supuesto: D tiene distribución normal.

### 2) Hipótesis Estadísticas

$$H_0: \mu_D \leq 40$$

La intervención no tiene un efecto relevante.

$$H_1: \mu_D > 40$$

La intervención es eficaz para mejorar substancialmente la performance.

### 3) Nivel de significación: $\alpha = 0,10$

**Nota:** La elección de este nivel de significación, menos exigente que lo usual para probar la hipótesis científica de que la intervención es eficaz para mejorar substancialmente la performance, es para compensar el hecho de que el umbral elegido de 40 puntos es ya de por sí muy exigente y se desea cubrirse contra el riesgo de no detectar el caso en que efectivamente se pasa este umbral (error de tipo II).

### 4) Estadístico de contraste y su distribución bajo $H_0$

$$E = \frac{\bar{D} - 40}{\frac{S'_D}{\sqrt{7}}} \sim t_6 \text{ bajo } H_0$$

### 5) Regla de decisión

Al ser un contraste unilateral a derecha, se rechaza  $H_0$  si y sólo si  $e > t_{6,0,90}$

Buscando en la tabla,  $t_{6,0,90} = 1,44$ .

Luego, se rechaza  $H_0$  si y sólo si el valor observado del estadístico de contraste es mayor que 1,44.

### 6) Cálculo del valor observado del estadístico de contraste

Los valores muestrales de la variable D en el grupo experimental son: 60,100,110,11,16,132 y 84.

$$e = \frac{73,29 - 40}{\frac{46,48}{\sqrt{7}}} = 1,895$$

### 7) Conclusión

Como  $1,895 > 1,44$  se rechaza  $H_0$ ; es decir, la performance media de los ajedrecistas aumentaría en más de 40 puntos con el programa de intervención.

c) Se debe hacer un contraste sobre la diferencia de medias de datos pareados para el grupo control.

### 1) Variables y supuestos mínimo necesarios para llevar a cabo el contraste

D: Diferencia de puntaje después y antes de la intervención de los sujetos de la población hipotética representada por la muestra del grupo control.

Supuesto: D tiene distribución normal.

### 2) Hipótesis Estadísticas

$H_0: \mu_D = 0$  La performance en la población de ajedrecistas prometedores es en promedio la misma a lo largo de dos períodos si no hay intervención.

$H_1: \mu_D > 0$  La performance en la población de ajedrecistas prometedores aumenta en promedio a lo largo de dos períodos aunque no reciban intervención.

### 3) Nivel de significación: $\alpha = 0,10$

**Nota:** La elección de este nivel de significación, más alto que el que generalmente se toma (5%) se justifica en el hecho de que se trata de un contraste para validar una afirmación “Entre los períodos considerados no hay progreso sin la intervención”, afirmación que se valida si no se rechaza la hipótesis nula. De ser así, se estaría en riesgo de cometer error de tipo II y este nivel permite tener un menor riesgo de cometerlo (mejor potencia).

### 4) Estadístico de contraste y su distribución bajo $H_0$

$$E = \frac{\bar{D}}{\frac{S_D}{\sqrt{7}}} \sim t_6 \text{ bajo } H_0$$

### 5) Regla de decisión

Al ser un contraste unilateral a derecha, se rechaza  $H_0$  si y sólo si  $e > t_{6;0,90}$

Buscando en la tabla,  $t_{6;0,90} = 1,44$ . Luego, se rechaza  $H_0$  si y sólo si el valor observado del estadístico de contraste es mayor que 1,44.

### 6) Cálculo del valor observado del estadístico de contraste

Los valores muestrales de D en el grupo control son: 10, 14, 12, -6, 13, -39 y 26.

$$e = \frac{4,29}{\frac{21,27}{\sqrt{7}}} = 0,534$$

### 7) Conclusión

Como  $0,534 < 1,44$  no se rechaza  $H_0$ ; es decir, sin la intervención, la performance media de los ajedrecistas no aumentaría a lo largo de dos períodos consecutivos.

d) Cuando se completa el período de seguimiento se observa un aumento de la performance que en promedio excede el umbral de 40 puntos para el grupo experimental mientras que no hubo diferencias estadísticamente significativas para el grupo control. Como ambos grupos estaban en las mismas condiciones iniciales (ya que se probó que no había diferencias significativas entre ambos al comenzar), el cambio puede atribuirse a la eficacia del programa de intervención.

### EJERCICIO MODELO III

Jiménez, M. et al (2009) administraron cuestionarios sobre satisfacción a 673 personas que trabajaban como voluntarios en diversas organizaciones del ámbito socioasistencial, de los cuales 171 abandonaron dentro del transcurso de un año y el resto permaneció. Obtuvieron los datos que se muestran en la siguiente tabla. El artículo completo se halla en <http://www.psicothema.com/pdf/3603.pdf>

¿Puede afirmarse al 5% que los voluntarios que permanecen más de un año están más satisfechos en sus motivaciones? Si es así, estime el tamaño del efecto.

		Media	Desv. Est.
Satisfacción gestión	Abandonaron	5,30	1,15
	Continúan	5,55	0,99
Satisfacción tareas	Abandonaron	5,42	0,90
	Continúan	5,66	0,80
Satisfacción motivaciones	Abandonaron	5,21	0,83
	Continúan	5,37	0,86

#### RESOLUCIÓN

Es necesario llevar a cabo un contraste de hipótesis para la diferencia de medias de dos poblaciones a partir de muestras independientes.

#### 1) Variables y supuestos mínimo necesarios para llevar a cabo el contraste

$X_1$ : Puntaje en el cuestionario de satisfacción motivacional de los voluntarios de organizaciones socioasistenciales que permanecen por lo menos un año en la institución.

$X_2$ : Puntaje en el cuestionario de satisfacción motivacional de los voluntarios de organizaciones socioasistenciales que permanecen menos un año en la institución.

Supuestos:  $X_1$  y  $X_2$  son independientes.

#### 2) Hipótesis Estadísticas

$H_0: \mu_1 - \mu_2 = 0$  Los sujetos de ambas poblaciones no difieren en promedio en su satisfacción motivacional.

$H_1: \mu_1 - \mu_2 > 0$  Los sujetos que no abandonan antes del año están en promedio más satisfechos en su motivación.

#### 3) Nivel de significación: $\alpha = 0,05$

#### 4) Estadístico de contraste y su distribución bajo $H_0$

$$E = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{502} + \frac{S_2^2}{171}}} \approx N(0,1)$$

por un teorema similar al Teorema Central del Límite ya que cada muestra es grande.

Bajo  $H_0$

## 5) Regla de decisión

Al ser un contraste unilateral a derecha, se rechaza  $H_0$  si y sólo si  $e > z_{0,95} = 1,645$ .

## 6) Cálculo del valor observado del estadístico de contraste

$$e = \frac{5,37 - 5,21}{\sqrt{\frac{0,86^2}{502} + \frac{0,83^2}{171}}} = 2,16$$

## 7) Conclusión

Como  $2,16 > 1,645$  se rechaza  $H_0$ ; es decir, que los voluntarios que perseveran tienen en promedio una mayor satisfacción motivacional.

## Tamaño del Efecto:

$$d = \frac{5,37 - 5,21}{\sqrt{\frac{501 \cdot 0,86^2 + 170 \cdot 0,83^2}{502 + 171 - 2}}} = 0,19$$
 Por lo que el tamaño del efecto es muy pequeño y

no parece tener mucha relevancia en términos de significación psicológica.

## EJERCICIO MODELO IV

En una investigación sobre violencia contra la mujer en estudiantes universitarias, Vázquez et al (2010) observaron que 159 jóvenes, entre las 1043 que fueron elegidas al azar para el estudio, experimentaron algún tipo de maltrato en algún momento de sus vidas por parte de su pareja o de alguna persona del sexo opuesto diferente de su pareja. El muestreo fue estratificado por ciclo (1ero., 2do. y 3ero.) y por área de conocimiento (ciencias humanas y sociales, biosanitarias y experimentales) para las mujeres matriculadas en las universidades gallegas de Santiago de Compostela, Coruña y Vigo. ¿Puede afirmarse con un nivel de significación del 1% que el porcentaje de jóvenes universitarias de la correspondiente población que ha recibido este tipo de maltrato no llega al a) 16%? b) 20%? Para responder a estas preguntas calcule y utilice el nivel crítico (valor p).

Artículo completo en <http://www.psicothema.com/pdf/3715.pdf> .

## RESOLUCIÓN

a) Se trata de un contraste sobre una proporción poblacional.

### 1) Variables y supuestos mínimo necesarios para llevar a cabo el contraste

X: Condición de haber recibido o no maltrato por parte de su pareja u otra persona del sexo opuesto, en la población de estudiantes universitarias gallegas (Santiago de Compostela, Coruña y Vigo) de los tres primeros ciclos de las carreras de ciencias humanas y sociales, biosanitarias y experimentales.

Supuestos. Las observaciones muestrales son independientes una de otra y con la misma probabilidad de haber recibido maltrato. Es decir que; para cada elección al azar, X es una variable Bernoulli con el mismo parámetro  $\pi$  sobre el cual se desea hacer inferencias.



## 2) Hipótesis estadísticas

$H_0: \pi \geq 0,16$  Al menos el 16% de las universitarias en cuestión ha recibido maltrato.

$H_1: \pi < 0,16$  Menos del 16% de las universitarias en cuestión ha recibido maltrato.

3) Nivel de significación:  $\alpha = 0,01$

## 4) Estadístico de contraste y su distribución bajo $H_0$

por el Teorema Central del Límite ya que la muestra es grande y  $n \cdot \pi_0 = 1043 \cdot 0,16 = 167 > 5$  y  $n(1-\pi_0) = 1043 \cdot 0,84 = 876 > 5$ .

$$E = \frac{p - 0,16}{\sqrt{\frac{0,16 \cdot 0,84}{1043}}} \approx N(0,1)$$

Bajo  $H_0$

## 5) Cálculo del valor observado del estadístico de contraste

$$e = \frac{0,152 - 0,16}{\sqrt{\frac{0,16 \cdot 0,84}{1043}}} \approx -0,71$$

## 6) Cálculo del nivel crítico o valor p

El contraste es unilateral a izquierda; por lo que el valor p es  $P(E < -0,71 \text{ bajo } H_0) = P(Z < -0,71) = 0,2389$ .

## 7) Conclusión

Como  $0,2389 > 0,01$  no se rechaza  $H_0$ ; es decir que no puede afirmarse al 1% (y ni siquiera al 10%) que el porcentaje de mujeres de la población considerada que ha recibido maltrato por parte de sus parejas u otros varones sea inferior al 16%.

b) Ídem a) con  $\pi_0 = 0,20$

## 1) Variables y supuestos mínimo necesarios para llevar a cabo el contraste

Ídem a)

## 2) Hipótesis estadísticas

$H_0: \pi \geq 0,20$  Al menos el 20% de las universitarias en cuestión ha recibido maltrato.

$H_1: \pi < 0,20$  Menos del 20% de las universitarias en cuestión ha recibido maltrato.

3) Nivel de significación: Ídem a)  $\alpha = 0,01$

## 4) Estadístico de contraste y su distribución bajo $H_0$

por el Teorema Central del Límite ya que la muestra es grande y  $n \cdot \pi_0 = 1043 \cdot 0,20 = 209 > 5$  y  $n(1-\pi_0) = 1043 \cdot 0,80 = 834 > 5$ .

$$E = \frac{p - 0,20}{\sqrt{\frac{0,20 \cdot 0,80}{1043}}} \approx N(0,1)$$

Bajo  $H_0$

## 5) Cálculo del valor observado del estadístico de contraste

$$e = \frac{0,152 - 0,20}{\sqrt{\frac{0,20 \cdot 0,80}{1043}}} \approx -3,88$$

## 6) Cálculo del nivel crítico o valor p

El contraste es unilateral a izquierda; por lo que el valor p es  $P(E < -3,88 \text{ bajo } H_0) = P(Z < -3,88) = 0,0001$

## 7) Conclusión

Como  $0,0001 < 0,01$  se rechaza  $H_0$ ; es decir que hay evidencia al 1% (e inclusive al 1‰) que menos de la quinta parte de las mujeres de la población considerada ha recibido maltrato por parte de sus parejas u otros varones.

# EJERCITACIÓN

## EJERCICIO 1

En una investigación para validar la versión en español del Test Conductual de Memoria de Rivermead (RBMT) para población mayor de 70 años, Alonso, M. y Prieto, P. (2004) informan una media de 16,1 con una desviación estándar de 4,2 para las personas entre 70 y 75 años de un grupo normativo de Tenerife. Si en una muestra aleatoria de 81 personas del mismo rango de edad de la ciudad A se observara una media de 16,8 puntos

- ¿habría suficiente evidencia al 5% de que el rendimiento en el test de las personas entre 70 y 75 años de la ciudad A es en promedio superior al del grupo normativo de Tenerife?
- ¿Qué valores de la media muestral harían la diferencia estadísticamente significativa en la dirección de la hipótesis planteada?

**Nota.** Los valores de la media y desviación estándar están redondeados a un decimal. El artículo completo puede hallarse en <http://www.psicothema.com/pdf/1201.pdf>

## EJERCICIO 2

La impuntualidad de 16 empleados elegidos al azar de cierta dependencia de la administración pública fue de 11 minutos. ¿Son estos datos consistentes con la hipótesis de que la impuntualidad se distribuye normalmente con media 10 y desvío 4 minutos? Usar un nivel de significación del 10%.

## EJERCICIO 3

Algunas investigaciones sobre el tiempo de reacción de automovilistas medido como el tiempo en que tardan en frenar desde que perciben un obstáculo inesperado, informan una media de 1,5 segundos con una desviación estándar de 0,2 segundos aproximadamente. Para decidir al 1% si el tiempo de reacción promedio de los porteños es inferior, se tomará una muestra de 64 automovilistas.

- ¿Para qué valores de la media muestral podría considerarse que el tiempo de reacción promedio de los porteños es inferior al reportado por las investigaciones?
- Si de la muestra se obtuvo una media de 1,36 segundos, ¿qué concluiría en este caso?
- Halle el nivel crítico de probabilidad o valor p.

#### EJERCICIO 4

El gerente de cierta agencia de publicidad desea tomarte a prueba para que dictes un curso de capacitación a sus vendedores asegurándote que sólo cobrarás si el entrenamiento resulta eficaz. Él afirma que hasta el momento el volumen de ventas mensual de sus empleados es en promedio aproximadamente \$2150. Antes de aceptar el trabajo decides validar esta afirmación interrogando a 9 vendedores elegidos al azar y obtienes un volumen medio mensual de ventas de \$2080 con un desvío estándar de \$80.

- ¿Son estos datos consistentes al 5% con lo informado por el gerente? ¿Aceptarías el trabajo? ¿Por qué?
- Supongamos que aceptas el trabajo y que entrenas a una muestra aleatoria 16 vendedores, los cuales en los meses subsiguientes del entrenamiento logran un volumen medio mensual de ventas de \$2190 con un desvío estándar de \$60. Usando el mismo nivel de significación, ¿podrás aspirar al cobro de tu trabajo? Explicita los supuestos necesarios para la resolución de a) y de b).

#### EJERCICIO 5

La edad en meses de la menarca de las niñas de cierto distrito del Gran Buenos Aires se distribuye normalmente con una media de 140,5. Un relevamiento reciente sobre las últimas 25 consultas recibidas por el Servicio de Pediatría del hospital zonal dio una edad promedio de 132,2 meses con un desvío de 21,5 meses. Un equipo interdisciplinario pretende hallar la respuesta a este suceso pero necesitan saber si la disminución en la edad de la menarca registrada es estadísticamente significativa al 10% ¿podrías ayudarlos?

#### EJERCICIO 6 (Datos reales extraídos de una investigación de Ayala et al, 1990).

Para probar la eficacia de un programa de estimulación temprana (PET) se administró un test de C.I. a un grupo control y a un grupo experimental antes de iniciar el programa y después de un año de aplicación del mismo sobre el grupo experimental. Ambos grupos quedaron determinados por asignación aleatoria de niños de 2 a 4 años que asistían al Hogar Infantil “Copetín” de la ciudad de Pamplona, Colombia.

- Pruebe que no hay diferencias estadísticamente significativas al 10% entre el grupo control y el grupo experimental antes de la implementación del programa.
- ¿Puede concluirse al 1% que el C.I. aumenta, en promedio, en los niños que reciben el PET? Calcule y valore el tamaño del efecto.
- No hay evidencias al 1% de que aumente el C.I. medio de los niños que no reciben el PET.
- ¿Por qué es necesario usar un grupo control? ¿Qué conclusión general puede sacar?

DATOS															
Grupo Control															
Sujeto	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Pre-test	86	135	70	98	111	109	96	92	117	142	107	103	105	116	80
Post-test	110	112	104	106	118	103	110	129	109	117	118	102	124	121	131
Grupo Experimental															
Sujeto	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Pre-test	108	125	105	111	126	105	92	90	104	87	104	104	127	102	100
Post-test	157	147	134	138	154	127	127	129	124	132	124	133	128	144	157

## EJERCICIO 7

Los Trastornos de la Conducta Alimentaria (TCA) han despertado un creciente interés durante los últimos años entre los profesionales de la salud debido al incremento de su incidencia y prevalencia<sup>6</sup>. Los TCA tienen una etiología policausal, interactuando de una forma compleja factores predisponentes, factores desencadenantes y mantenedores (Garner y Garfinkel, 1980; Toro y Vilardell, 1987). Dentro de estos factores se ha postulado la insatisfacción con la propia figura (IC) como uno de los factores claves y se encontraron importantes relaciones entre IC y la instauración de TCA.

Un grupo de investigadores realiza un estudio seleccionando dos muestras aleatorias e independientes de 29 chicas y 24 chicos de alrededor de 15 años que cursan el 3º año en una escuela privada de cierto distrito. Les administran el Eating Disorders Inventory (EDI), test original de Garner, Olmsted y Polivy (1983), es un autoinforme que evalúa las características conductuales y cognitivas de la anorexia y la bulimia nerviosa. Consta de 64 ítems agrupados en 8 subescalas. Hay una 2º versión (EDI-2) con 3 subescalas más quedando con 91 ítems. Una de las subescalas mide IC. En esta escala IC los investigadores obtuvieron los siguientes puntajes:

Chicas	13	11	13	10	16	13	13	5	17	17	21	9	8	19	16	2	11	15	10	25	13	22	13	15
	9	15	6	15	10																			
Chicos	8	10	7	3	0	8	16	7	3	12	6	11	4	5	2	4	10	4	10	8	7	8	13	15

¿Es la diferencia observada en la IC entre chicas y chicos estadísticamente significativa al 5%? ¿Es el sexo una FSV? Explícite los supuestos que necesitó hacer para llevar a cabo el contraste. ¿De ser así, cuál es el tamaño del efecto del sexo sobre la IC?

**Nota.** Este ejercicio, con datos ficticios, está basado sobre la investigación de Baile et al (2003). Artículo completo en [http://www.um.es/analesps/v19/v19\\_2/02-19\\_2.pdf](http://www.um.es/analesps/v19/v19_2/02-19_2.pdf)

## EJERCICIO 8

Un investigador desea comprobar si, como afirma la teoría, los sujetos que realizan un dibujo pequeño en el test de la Figura Humana son más introvertidos. El psicólogo posee una muestra de 150 sujetos a los que les administró la escala Introversión Social del Inventario de personalidad MMPI-2. Los puntajes en dicha escala se distribuyen normalmente con media 50 y desviación típica 10. Conformó dos grupos de la siguiente manera: el grupo de introvertidos, compuesto por los sujetos que no superaron los 40 puntos y el grupo de extrovertidos, integrado por los que superaron los 60 puntos. A ambos grupos les administró la técnica de la Figura Humana y midió el tamaño de la representación realizada. La media y desviación estándar para cada grupo, expresadas en centímetros, se muestran en la tabla de abajo. a) Calcule el tamaño de muestra para cada grupo. (Redondeé al entero más cercano).

b) ¿Ha podido el investigador comprobar su hipótesis al 5%? Explícite los supuestos necesarios para llevar a cabo el contraste correspondiente. Si su respuesta es afirmativa estime el tamaño del efecto.

	$\bar{x}$	s
Grupo "Introvertido"	10,3	2,3
Grupo "Extrovertido"	12,2	2,5

<sup>6</sup> Términos utilizados en Epidemiología: *Prevalencia* se denomina a la proporción de individuos de un grupo o una población que presentan una característica o evento determinado en un momento o en un período determinado. Refiere a todos los individuos afectados. *Incidencia* es una medida del número de casos nuevos de una enfermedad en un período de tiempo determinado.

## EJERCICIO 9

Efectúe el mismo análisis del Ejercicio Modelo III para el puntaje en satisfacción en las tareas.

## EJERCICIO 10

Se desea investigar si el nivel de estrés difiere entre las personas que desempeñan tareas ejecutivas y aquellas que trabajan en relación de dependencia. Para ello se tomaron dos muestras aleatorias de 16 ejecutivos y 16 empleados y se les administró un test de ansiedad. Por la manera en que se construyó la escala, los psicólogos consideran que para que una diferencia de puntajes en el nivel de estrés sea considerada relevante, ésta debe ser superior a dos puntos. En otras palabras, el umbral es de dos puntos.

Los resultados muestrales dieron un nivel de estrés medio de 55 puntos para los ejecutivos con un desvío estándar de 6 puntos y una media de 50 con un desvío de 7 puntos para los empleados. Suponiendo igualdad de varianzas responda:

- ¿Es la diferencia observada en el nivel de estrés entre ambos grupos estadísticamente significativa al 5%?
- ¿Puede deducirse de a) la significación psicológica; esto es, que el tipo de actividad tiene un efecto importante sobre el nivel de estrés?
- ¿Qué hipótesis estadística debería plantear para poder deducir la significación psicológica a partir de la estadística? Haga la prueba correspondiente.
- Formule una conclusión general.

## EJERCICIO 11

En el artículo sobre memoria y envejecimiento de Simon et al (2009) se exhibe la siguiente tabla cuyos resultados corresponden a muestras de 22 adultos jóvenes y 22 adultos mayores.

Valoración del agrado suscitado por las imágenes: Media (desviación típica)				
Edad	Imágenes			
	Recordadas	No recordadas	Reconocidas	No reconocidas
Jóvenes	3,998 (0,547)	3,991 (0,619)	3,943 (0,518)	3,398 (1,282)
Mayores	5,168 (0,760)	4,982 (0,606)	5,029 (0,613)	4,521 (1,032)

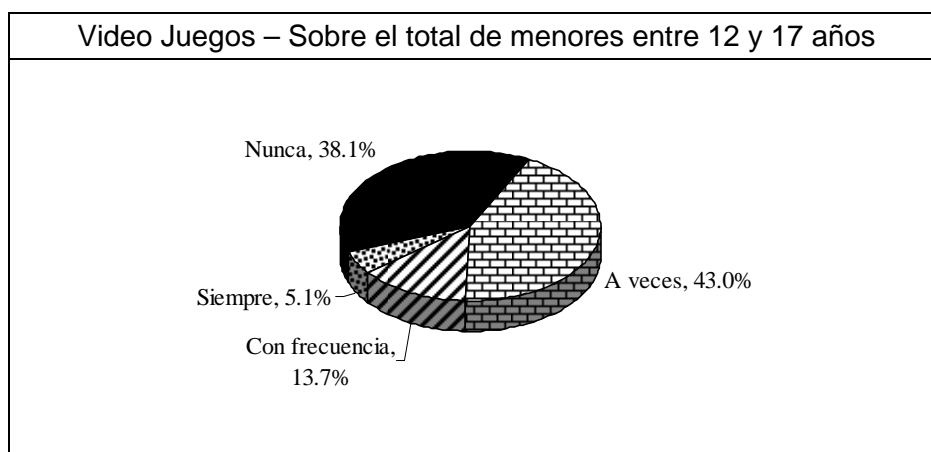
Nota: el agrado se valoró en una escala de 1 (muy desagradable) a 7 (muy agradable)

Resumiendo la valoración de las imágenes en las cuatro condiciones mediante el promedio se obtiene una media y desviación estándar de 3,8325 (0,8053) para los jóvenes y 4,925 (0,7723) para los mayores. Según estos resúmenes ¿puede afirmarse al 1% que la edad es una fuente de variación para la valoración de las imágenes, en qué sentido? De ser así calcule el tamaño del efecto.

## EJERCICIO 12

En una investigación sobre conductas indicadoras de posibles problemas de adicción a las nuevas tecnologías, Labrador, F. Villadangos, S. (2010) informan la frecuencia con la que menores escolarizados de la comunidad de Madrid, entre 12 y 17 años, manifiestan tener inclinaciones de dependencia hacia la televisión, Internet, telefonía móvil y video juegos. Para ello tuvieron que responder un cuestionario con escalas Likert de valores Nunca, A veces, Con frecuencia y Siempre. Parte de esos

resultados se exhiben a continuación. El artículo completo puede hallarse en <http://www.psicothema.com/pdf/3713.pdf>.



TV	12 años	13 años	14 años	15 años	16 años	17 años
Nunca	90	130	124	174	68	29
A veces	118	120	149	194	56	14
Con frecuencia	50	61	57	68	27	7
Siempre	32	39	40	35	18	10

Se considera que los menores perciben como problemática a una tecnología si sus respuestas son “Con frecuencia” o “Siempre”. A partir de dicha información pruebe las siguientes hipótesis al 1%

- Menos de la cuarta parte de los menores percibe como problemáticos a los video juegos.
- Más de la quinta parte de los menores percibe como problemática a la TV.
- La cuarta parte de los que tienen 14 años percibe como problemática a la TV.

### EJERCICIO 13

Una firma hace un estudio de mercado con el fin de decidir si es necesario o no aumentar la publicidad el siguiente año. Como un aumento de la publicidad implicaría un importante incremento del presupuesto para ese fin, la empresa quiere evitar la posibilidad de lanzar una campaña innecesariamente. La campaña será lanzada sólo si menos del 45% de la población consume el producto de la firma. De 400 personas encuestadas, seleccionadas al azar de distintos distritos de Capital Federal, 168 afirmaron consumir el producto. ¿Qué decisión tomará la firma al nivel del 5%?

### RESPUESTAS

Nota. En las pruebas unilaterales para la diferencia de medias de datos pareados o independientes, tanto el signo en la hipótesis alternativa  $>$  (derecha) o  $<$  (izquierda) como el signo del valor observado del estadístico de contraste (positivo o negativo), dependen del sentido en que se haya planteado la diferencia, el cual es arbitrario. Para las respuestas a este tipo de problemas se eligió una de las dos posibles formulaciones pero lo importante es que todo esté en coherencia.

### EJERCICIO 1

Prueba unilateral a derecha sobre la media de una variable con desvío conocido.  
a) No;  $z_{0,95} = 1,645$ . y  $e = 1,5$ . No se rechaza  $H_0$ . b) Valores mayores que 16,87.

## EJERCICIO 2

Prueba bilateral sobre una media de variable normal con desvío conocido.

$z_{0,05} = -1,645$  y  $z_{0,95} = 1,645$ .  $e = 1$ . No se rechaza  $H_0$ . Los datos son consistentes con dicha hipótesis.

## EJERCICIO 3

Prueba unilateral a izquierda sobre la media de una variable con desvío conocido (suponiendo que la desviación estándar de la variable para los porteños es aproximadamente la misma que la reportada en la literatura) a)  $z_{0,01} = -2,33$ . Valores menores que 1,4417. b) El tiempo medio de los porteños es inferior al reportado.

c)  $e = -5,6$ . Valor  $p = 0^+$ .

## EJERCICIO 4

a) Prueba bilateral sobre la media de una variable con desvío desconocido.

Supuesto: normalidad del volumen mensual de ventas.  $t_{8;0,025} = -2,306$  y  $t_{8;0,975} = 2,306$ .

$e = -2,625$ . Se rechaza  $H_0$ . No aceptaría el trabajo porque el gerente está sobreestimando los actuales salarios y eso podría conducir a no detectar si hubiera un verdadero efecto positivo del entrenamiento. b) Test unilateral a derecha sobre la media de una variable con desvío desconocido. Supuesto: normalidad del volumen mensual de ventas que tendrían los vendedores si se entrenaran.  $t_{15;0,95} = 1,753$  y  $e = 2,67$ . Se rechaza  $H_0$ . Cobraría el trabajo.

## EJERCICIO 5

Prueba unilateral a izquierda sobre la media de una variable normal con desvío desconocido.  $t_{24;0,10} = -1,318$  y  $e = -1,93$ . Se rechaza  $H_0$ . La disminución de la edad media es estadísticamente significativa al 10%.

## EJERCICIO 6

a) Prueba bilateral sobre una diferencia de medias de variables normales independientes con la misma varianza.  $t_{28;0,05} = -1,701$  y  $t_{28;0,95} = 1,701$ .  $e = -0,2617$ . No se rechaza  $H_0$ , por lo que no hay diferencia significativa entre ambos grupos.

b) Prueba unilateral a derecha para datos pareados. Supuesto: La diferencia de puntuaciones después y antes del PET es normal.  $t_{14;0,99} = 2,624$  y  $e = 8,59$ . Se rechaza  $H_0$ . Hubo un aumento estadísticamente significativo en los niños del grupo experimental. El tamaño del efecto es 2,22; muy alto.

c) Prueba unilateral a derecha para datos pareados. Supuesto: Normalidad de la diferencia de puntajes.  $t_{14;0,99} = 2,624$  y  $e = 1,78$ . No se rechaza  $H_0$ ; por lo que no hay evidencias al 1% de que aumentaría el CI de los niños que no recibieran el PET.

d) Es necesario usar un grupo control para poder atribuir los cambios observados (en este caso el aumento del CI) al factor que define al grupo experimental: haber recibido el PET. Al haberse probado en a) que ambos grupos estaban en similares condiciones de partida y en b) que aumentaría el CI de quienes recibieran el PET pero en c) que no aumentaría el de quienes no lo recibieran, puede concluirse que el PET es eficaz para aumentar el CI.

## Comentario

Para formular la conclusión d), además del contraste de a) que era necesario para probar que ambos grupos partían de condiciones similares, se han llevado a cabo dos contrastes por separado, b) y c), cada uno con su nivel de significación; por lo que la afirmación simultánea de ambas cosas no obedece a un nivel de significación global conocido. Siempre que sea posible, deben tratarse todos los datos del experimento

conjuntamente a fin de que la decisión final obedezca a un único nivel de significación. Por otro lado, los niños que no reciban el PET podrían aumentar su CI en alguna medida por la estimulación ordinaria del hogar infantil. De hecho, el lector puede verificar que hay un aumento estadísticamente significativo al 5%. Esto no hace menos cierto que el PET sea eficaz, pues la eficacia del mismo no radica en que los niños que reciban el PET sean los únicos que aumentan su CI sino en que lo aumentan más que si hubieran recibido el programa usual. Teniendo en cuenta esto, sería más acertado plantear que la diferencia entre los CI del post-test y pre-test es, en promedio, mayor para la población hipotética representada por el grupo experimental que para la representada por el grupo control. En símbolos:

$H_0: \mu_{DE} = \mu_{DC}$  En ambas poblaciones la diferencia media de CI del pre al post test es igual.  
 $H_1: \mu_{DE} > \mu_{DC}$  El aumento, en promedio, del CI de quienes reciben el PET es mayor que el de quienes no lo reciben.

o, equivalentemente,

$H_0: \mu_{DE} - \mu_{DC} = 0$   
 $H_1: \mu_{DE} - \mu_{DC} > 0$

donde  $\mu_{DE}$  es la media de la diferencia para la población representada por el grupo experimental y  $\mu_{DC}$  la correspondiente para el grupo control.

### EJERCICIO 7

Prueba bilateral para la diferencia de medias de variables normales independientes con la misma varianza.  $t_{51;0,975} = 2,008$  y  $e = 4,39$ . Se rechaza  $H_0$ . El sexo es una fuente de variación, ya que las mujeres, en promedio, presentan mayor insatisfacción con la propia figura. El tamaño del efecto es 1,21; lo cual se considera alto.

### EJERCICIO 8

a) 24.  
 b) Prueba unilateral a izquierda para la diferencia de medias de dos variables normales independientes y con la misma varianza.  $t_{46;0,05} = -1,679$  y  $e = -2,74$ . Se rechaza  $H_0$ . El tamaño del efecto es 0,79, medio alto. El investigador ha podido comprobar su hipótesis.

### EJERCICIO 9

Prueba unilateral a derecha para la diferencia de medias de dos variables independientes con varianzas desconocidas pero estimadas con muestras grandes.  $z_{0,95} = 1,645$  y  $e = 3,095$ . Se rechaza  $H_0$ . El tamaño del efecto es 0,29; es decir pequeño. Los voluntarios que no abandonan muestran en promedio un poco más de satisfacción en las tareas que los que abandonan antes del año.

### EJERCICIO 10

a) Prueba bilateral para la diferencia de medias de dos variables normales e independientes con la misma varianza.  $t_{30;0,025} = -2,042$  y  $t_{30;0,975} = 2,042$ .  $e = 2,17$ . La diferencia es estadísticamente significativa, mayor para los ejecutivos.  
 b) No; para que la diferencia tenga una significación psicológica debe superar el umbral.  
 c)  $H_0: \mu_{EJ} = \mu_{EM} + 2$



$$H_1: \mu_{EJ} > \mu_{EM} + 2 \quad \text{o, equivalentemente,}$$

$$H_0: \mu_{EJ} - \mu_{EM} = 2$$

$$H_1: \mu_{EJ} - \mu_{EM} > 2$$

donde  $\mu_{EJ}$  es La media de estrés en la población de ejecutivos y  $\mu_{EM}$  la correspondiente en la población de empleados.

$t_{30;0,95} = 1,697$  y  $e = 1,3$ . No se rechaza  $H_0$ . No hay evidencias de que el puntaje medio de estrés en los ejecutivos supere en más de dos puntos al de los empleados.

d) Aunque la diferencia entre los puntajes de ambos grupos es estadísticamente significativa, siendo mayor el puntaje medio de estrés de los ejecutivos; ésta no es psicológicamente significativa porque no hay evidencias de que supere el umbral.

### EJERCICIO 11

Prueba bilateral para la diferencia de medias de variables normales independientes con igual varianza.  $t_{42;0,005} = -2,698$  y  $t_{42;0,995} = 2,698$ .  $e = 4,593$ . Se rechaza  $H_0$ , la valoración de las imágenes en los mayores es en promedio más alta; por lo que la edad es una fuente de variación. El tamaño del efecto es 1,38; que es alto.

### EJERCICIO 12

a) Prueba unilateral a izquierda para una proporción.  $z_{0,01} = -2,33$  y  $e = -5,92$ . Se rechaza  $H_0$ ; menos de la cuarta parte percibe como problemáticos los videos juegos.

b) Prueba unilateral a derecha para una proporción.  $z_{0,99} = 2,33$  y  $e = 6,17$ . Se rechaza  $H_0$ ; más de la quinta parte percibe como problemática a la TV.

c) Prueba bilateral.  $z_{0,005} = -2,575$  y  $z_{0,995} = 2,575$ .  $e = 0,54$ . No se rechaza  $H_0$ ; se valida la afirmación de que la cuarta parte de los jóvenes de 14 años perciben como problemática a la TV.

### EJERCICIO 13

Con el fin de controlar el error de hacer publicidad innecesariamente deben plantearse las hipótesis

$$H_0: \pi \geq 0,45$$

$$H_1: \pi < 0,45$$

El procedimiento es el mismo que si las hipótesis fueran

$$H_0: \pi = 0,45$$

$$H_1: \pi < 0,45$$

$z_{0,05} = -1,645$  y  $e = -1,21$ . No se rechaza  $H_0$ . La empresa decidirá no hacer publicidad por no haber suficiente evidencia de su necesidad.

### REFERENCIAS

Alonso, M. y Prieto, P. (2004). Validación de la versión en español del Test Conductual de Memoria de Rivermead (RBMT) para población mayor de 70 años. *Psicothema*, 16, 2, 325-328.

Ayala, H., Sabrera de Coronado, H., Coronado, M. (1990) . Estimulación temprana: su influencia sobre el desarrollo intelectual de niños de 2 a 4 años. *Avances en Psicología Clínica Latinoamericana*, 8, 115-124.

Baile, J.; Raich, R. y Garrido, E. (2003). Evaluación de Insatisfacción Corporal en adolescentes: Efecto de la forma de administración de una escala. *Anales de Psicología*, 19, 2, 187-192.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates.

- Garner, D. y Garfinkel, P. (1980). Cultural Expectations of Thinness in Woman. *Psychological reports*, 47, 483-491.
- Garner, D., Olmsted, M. y Polivy, J. (1983). Development and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. *International Journal of Eating disorders*, 2, 15-34.
- Jacob Cohen (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates.
- Jiménez, M.; Chacón Fuertes, F. y Sueiro Abad, M. (2009). Satisfacción en el voluntariado: estructura interna y relación con la permanencia en las organizaciones. *Psicothema*, 21, 1, 112-117.
- Labrador, F. y Villadangos, S. (2010). Menores y nuevas tecnologías: conductas indicadoras de posible problema de adicción. *Psicothema*, 22, 2, 180-188.
- Ruiz, F. y Luciano, C. (2009). Eficacia de la terapia de aceptación y compromiso (ACT) en la mejora del rendimiento ajedrecístico de jóvenes promesas. *Psichotema*, 21, 3, 347-352.
- Simón, Teresa, Ruiz Gallego-Largo, Trinidad y Suengas, Aurora (2009). Memoria y envejecimiento: recuerdo, reconocimiento y sesgo positivo. *Psicothema*, 21,3,409-415.
- Toro, J. y Vilardell, E. (1987). *Anorexia Nerviosa*. Barcelona: Martínez Roca.
- Vázquez, F., Torres, A., Otero, P., Blanco, V y López, M. (2010). Prevalencia y factores de riesgo de la violencia contra la mujer en estudiantes universitarias españolas. *Psicothema*, 22, 2, 196-201.

# PRÁCTICA DE INTEGRACIÓN

## PROBABILIDAD Y ESTADÍSTICA INFERENCIAL

### EJERCICIO 1

Los puntajes en un test de creatividad se distribuyen normalmente con media 65 y desvío estándar 10 puntos en un grupo normativo de 600 personas. Los sujetos se consideran según los siguientes criterios: “*Poco creativos*” si exceden dos desvíos por debajo de la media, “*Algo creativos*” si están a menos de dos desvíos por debajo de la media, “*Creativos*” si están a menos de dos desvíos por encima de la media y “*Muy creativos*” si distan más de dos desvíos por encima de la media.

Calcule

- la cantidad de individuos del grupo normativo correspondiente a cada categoría.
- el puntaje mínimo de los individuos que pertenecen al 2% más creativo.
- cómo es clasificado un sujeto que obtuvo 80 puntos.
- el puntaje correspondiente al tercer centil.

### EJERCICIO 2

Sabiendo por experiencia que sólo 4 de cada 10 personas que comienzan un determinado tratamiento perseveran en él, ¿cuál es la probabilidad de que en una muestra aleatoria de 8 personas que comienzan el tratamiento

- persevere más de la mitad?
- persevere por lo menos uno?
- no persevere a lo sumo la mitad?

### EJERCICIO 3

En una población determinada se incluye en la “clase activa” a todas aquellas personas que superan los 18 años y que no superan los 70 años. Los que han quedado por fuera de este intervalo son incluidos en la categoría “clase pasiva”. Sabiendo que el 68,26% de la población (en torno a la media) oscila entre 30 y 62 años y que la distribución de las edades es normal calcular:

- La media y el desvío poblacional de la distribución.
- La probabilidad de que una persona elegida al azar pertenezca a la clase pasiva.
- El porcentaje de personas con más de 23 años.
- La edad mínima necesaria para acceder a una asignación especial que el gobierno dispone entregar por única vez al 20% de las personas de más edad.
- La probabilidad de que en una muestra aleatoria de 10 personas a lo sumo 7 no estén en condiciones de acceder a la asignación especial de d).

### EJERCICIO 4. Artículo completo en <http://www.psicothema.com/pdf/3496.pdf>

Orgilés, M., Espada, J. y Méndez, F. (2008) administraron el Cuestionario de Ansiedad por Separación Infantil, Forma Niños, CASI-N (Espada et al, 2006) a 95 niños cuyos padres se encontraban separados o divorciados y a 95 niños que convivían con ambos progenitores por no haberse producido una ruptura de pareja. La media y desviación estándar para los primeros fue 29,66 y 11,49 mientras que para los segundos 26,99 y 10,83. ¿Hay diferencias estadísticamente significativas al 10% entre los niños de ambas condiciones? De ser así estime el tamaño del efecto e indique en qué sentido se da tal diferencia.

## EJERCICIO 5

Los puntajes en una prueba objetiva de rendimiento escolar siguen la distribución normal con media 15 y varianza 11,25 bajo el supuesto de responder todos los ítems al azar. Un docente toma la prueba a un curso de 36 alumnos y obtiene un puntaje medio de 16,5 puntos. ¿Puede descartar, a un nivel del 1%, la posibilidad de que sus alumnos hayan respondido al azar?

## EJERCICIO 6

El siguiente cuadro muestra la cantidad de alumnos que han elegido cada una de las opciones de post-grado que la universidad brinda. Estos resultados permiten inferir la demanda de cursos que tendrán los alumnos que hoy cursan el último año de la carrera. Suponiendo que el año entrante tenga un comportamiento similar al anterior responda:

Especialidad	Cant
Área Clínica	150
Área Jurídica	30
Área Educacional	60
Área Comunitaria	25
Área Laboral	35
Total	300

I. Con qué probabilidad en una muestra de 7 alumnos que cursarán el posgrado en esta universidad

- a lo sumo 5 no elijan el área Jurídica.
- todos elijan el área Clínica.
- ninguno elija el área Educacional.
- la mayoría elija Laboral o Comunitaria.
- por lo menos 2 no elijan el área Educacional.

II. ¿Puede afirmarse con un nivel de significación del 1% que para el año entrante menos de la quinta parte de los alumnos elegirán el área laboral?

## EJERCICIO 7

Remitiéndose a los datos del ejercicio 6 de la Práctica VII pruebe, al 5%, que los CI de la población correspondiente al grupo experimental aumentaron más, en promedio, que los CI de la del grupo control, según se sugiere en el comentario de la respuesta.

## EJERCICIO 8. Artículo completo en <http://www.psicothema.com/pdf/3523.pdf>

Rosario et al (2008) estudian la ansiedad en los exámenes en relación al sexo, entre otras variables personales y familiares. La ansiedad se evaluó con el Cuestionario de Ansiedad ante los Exámenes (CAE), de Rosario y Soares (2004), el cual está destinado a los alumnos entre 11 y 15 años que cursan la Enseñanza Secundaria Obligatoria (ESO). El cuestionario fue administrado a una muestra compuesta por 258 mujeres y 275 varones estudiantes de la ESO. Los autores exhiben la siguiente tabla:

Sexo	Media	Desvío típico	n
Varones	24,26	7,14	275
Mujeres	27,02	6,96	258

¿Puede afirmarse, con un nivel de significación del 5%, que hay diferencias estadísticamente significativas entre mujeres y varones? Si es así, calcule el tamaño del efecto y concluya si tal diferencia es psicológicamente significativa.

## EJERCICIO 9

Gastón se recibió recientemente de psicólogo y va a abrir un consultorio pero no se siente muy seguro en la atención de casos con síndrome de depresión. Consultando entre sus colegas obtuvo la siguiente información:

	Consulta A	Consulta B	Consulta C	Consulta D
Casos síndrome de depresión	3	5	15	7
Total de pacientes	30	55	135	80

En el hipotético caso de que al consultorio de Gastón acudieran 9 pacientes ¿qué probabilidad le asignaría al hecho de que

- alguno tenga síndrome de depresión?
- más de la mitad no tenga síndrome de depresión?

## EJERCICIO 10

En una investigación donde se analizan los efectos del video juego *Tradislexia* en la conciencia fonológica y reconocimiento de palabras en niños disléxicos, la hipótesis es que el entrenamiento multimedia contribuye a mejorar los procesos fonológicos. Para poner a prueba esta hipótesis se seleccionaron 32 niños con problemas de dislexia entre 9 y 12 años del segundo y tercer ciclo de educación primaria, y se asignaron al azar a un grupo experimental y a un grupo control, de igual tamaño (16). Al grupo experimental se lo entrenó con el video juego Tradislexia, 30 minutos diarios, en un total de 20 sesiones mientras que el grupo control recibía la instrucción común de la escuela. Al cabo del período de entrenamiento se les administró la subescala de lectura de palabras (correcta identificación de 30 palabras) de la batería de Evaluación de los Procesos Lectores de los Niños de Educación Primaria PROLEC (Cuetos et al., 1996). Se obtuvieron los siguientes datos:

Grupo Experimental:

21 – 15 – 18 – 13 – 17 – 21 – 20 – 14 – 18 – 16 – 24 – 10 – 23 – 11 – 14 – 17

Grupo Control:

11 – 8 – 5 – 17 – 3 – 19 – 10 – 13 – 8 – 14 – 15 – 11 – 7 – 12 – 9 – 14

- Enuncie la hipótesis científica y su formulación en términos estadísticos.
- Lleve a cabo el contraste de hipótesis correspondiente con un nivel de significación del 10% y exprese una conclusión. Calcule el tamaño del efecto, si lo hay.
- ¿Qué supuestos sobre la distribución de los puntajes en las correspondientes poblaciones son necesarios para validar el contraste?

**Nota:** Este ejercicio es una adaptación, con datos ficticios, de la investigación de Jiménez, J. y Rojas, E. (2008). El artículo completo puede consultarse en <http://www.psicothema.com/pdf/3491.pdf>

## EJERCICIO 11

Un peso inferior al kilo o superior a los cuatro kilos al nacer puede considerarse riesgoso en una población determinada. Suponiendo que los pesos se distribuyan normalmente en los recién nacidos con media 2,8 y desvío 0,6 kilogramos. Calcule la probabilidad de que un recién nacido

- esté dentro del grupo de riesgo.
- pese entre 2 y 3 kilogramos.

## EJERCICIO 12

Una prueba de oído musical consiste en reproducir el sonido de la nota “la” (que corresponde a una frecuencia de 440) luego de oírla. Un grupo de 10 personas elegidas aleatoria e independientemente entre profesionales que se dedican a la afinación de instrumentos musicales realizó la prueba y de las 10 emisiones se obtuvo una frecuencia media de 435 con un desvío estándar de 11. ¿Puede afirmarse que la frecuencia media de las notas que emitirían los profesionales de la población muestreada difiere de 440 al nivel de 10%? Suponga la distribución normal de las frecuencias emitidas.

**EJERCICIO 13.** Artículo completo en <http://www.psicothema.com/pdf/3593.pdf>

Errasti Pérez et al (2009) estudian la eficacia de un programa para la prevención del consumo de drogas. Lo pusieron a prueba sobre 378 adolescentes de escuelas secundarias de Asturias a los que administraron, entre otras escalas, un cuestionario de consumo de drogas antes y después de la implementación del programa. Obtuvieron una diferencia media entre los puntajes del post-test menos los del pre-test de -0,4067 con una desviación estándar de 2,8961. ¿Puede afirmarse con un nivel de significación del 1% que el programa es eficaz? En caso afirmativo dé una medida del efecto.

## RESPUESTAS

### EJERCICIO 1

- a) Hay 14, 286, 286 y 14 personas en los respectivos grupos.
- b) 85,5.
- c) Creativo pues está a un desvío y medio por encima de la media.
- d) 46,2.

### EJERCICIO 2

- a) 0,174      b) 0,983      c) 0,406

### EJERCICIO 3

- a)  $\mu = 46$  y  $\sigma = 16$     b) 0,1069    c) 92,51    d) edad  $\geq 59,44$ ; es decir, 60    e) 0,3222

### EJERCICIO 4

Prueba bilateral.  $z_{0,05} = -1,645$  y  $z_{0,95} = 1,645$ .  $e = 1,648$ . Se rechaza  $H_0$ .  $d = 0,24$ , pequeño. La separación de los padres tiene efecto, aunque pequeño, sobre el nivel de ansiedad por separación de los hijos, aumentándolo.

### EJERCICIO 5

Prueba bilateral.  $z_{0,005} = -2,575$  y  $z_{0,005} = 2,575$ .  $e = 2,68$ . Se rechaza  $H_0$ . Se descarta la posibilidad de que los alumnos hayan contestado por azar.

### EJERCICIO 6

- I. a) 0,1497    b) 0,0078    c) 0,2097    d) 0,0333    e) 0,9996
- II. Prueba unilateral a izquierda.  $z_{0,01} = -2,33$ .  $e = -3,61$ . Se rechaza  $H_0$ . Sí, será inferior.

### EJERCICIO 7

Prueba unilateral a derecha.  $t_{28;0,95} = 1,701$ .  $e = 3,33$ . Se rechaza  $H_0$ . El programa es eficaz porque los niños que lo recibieran aumentarían su CI más que aquellos que sólo tuvieran la estimulación ordinaria.

### EJERCICIO 8

Prueba bilateral.  $z_{0,025} = -1,96$  y  $z_{0,975} = 1,96$ .  $e = 4,518$ . Se rechaza  $H_0$ , por lo que la diferencia es estadísticamente significativa. El tamaño del efecto es 0,39, apenas moderado, siendo las mujeres las que presentan un nivel de ansiedad en promedio algo mayor.

### EJERCICIO 9

a) 0,613      b) 0,999

### EJERCICIO 10

a) Hipótesis científica: el entrenamiento multimedia contribuye a mejorar los procesos fonológicos.

Hipótesis estadística:  $\mu_1 - \mu_2 > 0$  donde  $\mu_1$  es la media en la subescala PROLEC de los niños disléxicos españoles si recibieran el entrenamiento multimedia para el grupo experimental y control respectivamente y  $\mu_2$  si sólo recibieran el entrenamiento escolar.

b) Prueba unilateral a derecha.  $t_{30;0,90} = 2,457$ .  $e = 4,04$ . Se rechaza  $H_0$  con tamaño de efecto 1,43 que es alto. El entrenamiento multimedia contribuye a mejorar los procesos fonológicos.

c) Supuestos:  $X_1$  y  $X_2$  son variables aleatorias independientes, normalmente distribuidas y con igual varianza. Donde  $X_1$  es el puntaje en la subescala PROLEC de los niños disléxicos españoles si recibieran el entrenamiento multimedia y  $X_2$  el correspondiente si sólo recibieran el escolar.

### EJERCICIO 11

a) 0,0241      b) 0,5375

### EJERCICIO 12

Prueba bilateral.  $t_{9;0,05} = -1,833$  y  $t_{9;0,95} = 1,833$ .  $e = -1,44$ . No se rechaza  $H_0$ . La diferencia observada puede atribuirse a razones fortuitas y no necesariamente a que los profesionales emitan en promedio una frecuencia diferente de la correcta para la nota "la".

### EJERCICIO 13

Prueba unilateral a izquierda para diferencia de medias de datos pareados.

$t_{377;0,01} = -2,33$ .  $e = -2,73$ . Se rechaza  $H_0$ . El tamaño del efecto es 0,14, pequeño. El programa tiene eficacia pero con un efecto pequeño.

## REFERENCIAS

- Cuetos, F., Rodríguez, B., y Ruano, E. (1996). *Batería de evaluación de los procesos lectores de los niños de Educación Primaria (PROLEC)*. Madrid: T.E.A., Ediciones.
- Errasti Pérez, J.; Al-Halabí Díaz, S.; Secades Villa, R.; Fernández-Hermida, J.; Carballo, J. y García-Rodríguez, O. (2009). Prevención familiar del consumo de drogas: el programa «Familias que funcionan». *Psicothema*, 21, 1, 45-50.
- Espada, J., Méndez, F., Orgilés, M. e Hidalgo, M. (2006). Cuestionario de ansiedad por separación infantil, Forma Niños. Documento policopiado. Alicante: Universidad Miguel Hernández.
- Jiménez, J. y Rojas, E. (2008). Efectos del videojuego Tradislexia en la conciencia fonológica y reconocimiento de palabras en niños disléxicos. *Psicothema*, 20, 3, 347 – 353.
- Orgilés Amorós, M., Espada Sánchez, J. y Méndez Carrillo, X. (2008). Trastorno de ansiedad por separación en hijos de padres divorciados. *Psicothema*, 20, 3, 383-388.
- Rosario, P., y Soares, S. (2004). Questionário de Ansiedade face aos Testes (QAT). En Leandro, A., Simões, M., Machado, C., y Gonçalves, M. (Eds.): *Avaliação Psicológica: Instrumentos validados para a população portuguesa* (vol. II, pp. 39-51). Coímbra: Quarteto Editora.
- Rosario, P.; Núñez, J.; Salgado, A.; González-Pienda, J.; Valle, A.; Joly, C. y Bernardo, A. (2008). Ansiedad ante los exámenes: relación con variables personales y familiares. *Psicothema*, 20, 4, 563-570.