

# Bioestadística: preguntas y problemas resueltos

Javier Manzano

2023-10-13



# Índice general

<b>1. Introducción</b>	<b>9</b>
<b>2. Planteamiento de una investigación: Anatomía y Fisiología de la investigación</b>	<b>11</b>
2.1. Pregunta test . . . . .	11
2.2. Pregunta test . . . . .	12
2.3. Pregunta test . . . . .	12
2.4. Problema . . . . .	12
2.5. Pregunta test . . . . .	15
2.6. Pregunta test . . . . .	15
2.7. Pregunta test . . . . .	16
2.8. Problema . . . . .	16
2.9. Pregunta test . . . . .	17
2.10. Pregunta test . . . . .	17
2.11. Pregunta test . . . . .	17
2.12. Pregunta test . . . . .	18
2.13. Pregunta test . . . . .	18
2.14. Pregunta test . . . . .	18
2.15. Pregunta test . . . . .	19
2.16. Problema . . . . .	19
2.17. Pregunta test . . . . .	20
2.18. Pregunta test . . . . .	21
2.19. Pregunta test . . . . .	21
2.20. Pregunta test . . . . .	21
2.21. Problema . . . . .	22
2.22. Pregunta test . . . . .	23
2.23. Pregunta test . . . . .	23
2.24. Pregunta test . . . . .	23
2.25. Problema . . . . .	24
2.26. Pregunta test . . . . .	26
2.27. Pregunta test . . . . .	27
2.28. Pregunta test . . . . .	27
2.29. Pregunta test . . . . .	27

2.30. Pregunta test . . . . .	28
2.31. Problema . . . . .	28
2.32. Pregunta test . . . . .	30
2.33. Pregunta test . . . . .	30
2.34. Pregunta test . . . . .	31
2.35. Pregunta test . . . . .	31
2.36. Pregunta test . . . . .	31
2.37. Pregunta test . . . . .	32
2.38. Pregunta test . . . . .	32
2.39. Pregunta test . . . . .	32
2.40. Pregunta test . . . . .	33
2.41. Pregunta test . . . . .	33
2.42. Pregunta test . . . . .	33
2.43. Pregunta test . . . . .	34
2.44. Pregunta test . . . . .	34
2.45. Pregunta test . . . . .	34
2.46. Pregunta test . . . . .	35
2.47. Problema . . . . .	35
2.48. Pregunta test . . . . .	37
2.49. Pregunta test . . . . .	38
2.50. Pregunta test . . . . .	38
2.51. Pregunta test . . . . .	38
2.52. Pregunta test . . . . .	39
2.53. Pregunta test . . . . .	39
2.54. Pregunta test . . . . .	39
2.55. Pregunta test . . . . .	40
2.56. Pregunta test . . . . .	40
2.57. Pregunta test . . . . .	40
<b>3. Análisis Descriptivo y Gráfico de datos cuantitativos</b>	<b>41</b>
3.1. Pregunta test . . . . .	41
3.2. Pregunta test . . . . .	42
3.3. Problema . . . . .	42
3.4. Pregunta test . . . . .	44
3.5. Pregunta test . . . . .	44
3.6. Pregunta test . . . . .	44
3.7. Pregunta test . . . . .	45
3.8. Problema . . . . .	45
3.9. Pregunta test . . . . .	47
3.10. Pregunta test . . . . .	47
3.11. Pregunta test . . . . .	49
3.12. Problema . . . . .	49
3.13. Pregunta test . . . . .	50
3.14. Pregunta test . . . . .	50
3.15. Problema . . . . .	51
3.16. Pregunta test . . . . .	52

3.17. Pregunta test . . . . .	52
3.18. Pregunta test . . . . .	53
3.19. Pregunta test . . . . .	53
3.20. Pregunta test . . . . .	53
3.21. Pregunta test . . . . .	54
3.22. Pregunta test . . . . .	54
3.23. Pregunta test . . . . .	54
3.24. Pregunta test . . . . .	55
3.25. Pregunta test . . . . .	55
3.26. Pregunta test . . . . .	55
3.27. Pregunta test . . . . .	56
3.28. Pregunta test . . . . .	56
3.29. Pregunta test . . . . .	56
3.30. Pregunta test . . . . .	57
3.31. Pregunta test . . . . .	57
3.32. Pregunta test . . . . .	57
3.33. Pregunta test . . . . .	58
3.34. Pregunta test . . . . .	58
3.35. Pregunta test . . . . .	58
3.36. Pregunta test . . . . .	59
3.37. Pregunta test . . . . .	59
3.38. Pregunta test . . . . .	59
3.39. Pregunta test . . . . .	60
3.40. Pregunta test . . . . .	60
3.41. Pregunta test . . . . .	60
3.42. Pregunta test . . . . .	61
3.43. Pregunta test . . . . .	61
3.44. Pregunta test . . . . .	61
3.45. Pregunta test . . . . .	62
3.46. Pregunta test . . . . .	62
3.47. Pregunta test . . . . .	62
3.48. Pregunta test . . . . .	62
3.49. Pregunta test . . . . .	63
<b>4. Análisis Inferencial. Aplicaciones.</b>	<b>65</b>
4.1. Pregunta test . . . . .	65
4.2. Pregunta test . . . . .	65
4.3. Pregunta test . . . . .	66
4.4. Problema . . . . .	66
4.5. Pregunta test . . . . .	72
4.6. Pregunta test . . . . .	73
4.7. Pregunta test . . . . .	73
4.8. Pregunta test . . . . .	74
4.9. Problema . . . . .	74
4.10. Pregunta test . . . . .	80
4.11. Pregunta test . . . . .	81

4.12. Pregunta test . . . . .	81
4.13. Pregunta test . . . . .	81
4.14. Problema . . . . .	82
4.15. Pregunta test . . . . .	83
4.16. Pregunta test . . . . .	83
4.17. Pregunta test . . . . .	83
4.18. Pregunta test . . . . .	85
4.19. Pregunta test . . . . .	86
4.20. Pregunta test . . . . .	87
4.21. Pregunta test . . . . .	88
4.22. Pregunta test . . . . .	88
4.23. Pregunta test . . . . .	88
4.24. Problema . . . . .	89
4.25. Pregunta test . . . . .	89
4.26. Pregunta test . . . . .	90
4.27. Pregunta test . . . . .	90
4.28. Pregunta test . . . . .	90
4.29. Pregunta test . . . . .	91
4.30. Pregunta test . . . . .	91
4.31. Pregunta test . . . . .	92
<b>5. Regresión y correlación.</b>	<b>93</b>
5.1. Pregunta test . . . . .	93
5.2. Pregunta test . . . . .	94
5.3. Pregunta test . . . . .	94
5.4. Pregunta test . . . . .	95
5.5. Pregunta test . . . . .	97
5.6. Pregunta test . . . . .	98
5.7. Pregunta test . . . . .	98
5.8. Pregunta test . . . . .	98
5.9. Pregunta test . . . . .	99
5.10. Pregunta test . . . . .	99
5.11. Pregunta test . . . . .	100
5.12. Pregunta test . . . . .	100
5.13. Pregunta test . . . . .	100
5.14. Pregunta test . . . . .	101
5.15. Pregunta test . . . . .	102
5.16. Pregunta test . . . . .	102
5.17. Pregunta test . . . . .	102
5.18. Pregunta test . . . . .	103
5.19. Pregunta test . . . . .	103
5.20. Pregunta test . . . . .	103
5.21. Pregunta test . . . . .	104
5.22. Pregunta test . . . . .	104
5.23. Pregunta test . . . . .	104
5.24. Pregunta test . . . . .	105

5.25. Pregunta test . . . . .	105
5.26. Pregunta test . . . . .	105
5.27. Pregunta test . . . . .	106
5.28. Pregunta test . . . . .	106
5.29. Pregunta test . . . . .	107
<b>6. Tablas de contingencia.</b>	<b>109</b>
6.1. Problema . . . . .	109
6.2. Pregunta test . . . . .	112
6.3. Pregunta test . . . . .	112
6.4. Problema . . . . .	112
6.5. Pregunta test . . . . .	114
6.6. Pregunta test . . . . .	115
6.7. Pregunta test . . . . .	115
6.8. Pregunta test . . . . .	116
<b>7. Medidas de importancia clínica.</b>	<b>117</b>
7.1. Pregunta test . . . . .	117
7.2. Pregunta test . . . . .	118
7.3. Pregunta test . . . . .	118
7.4. Problema . . . . .	118
7.5. Pregunta test . . . . .	119
7.6. Pregunta test . . . . .	119
7.7. Problema . . . . .	120
7.8. Pregunta test . . . . .	121
7.9. Pregunta test . . . . .	121
7.10. Problema . . . . .	121
7.11. Pregunta test . . . . .	123
7.12. Pregunta test . . . . .	123
7.13. Pregunta test . . . . .	123
7.14. Pregunta test . . . . .	124
7.15. Pregunta test . . . . .	124
7.16. Pregunta test . . . . .	124
7.17. Pregunta test . . . . .	125
7.18. Pregunta test . . . . .	125
7.19. Pregunta test . . . . .	125
7.20. Pregunta test . . . . .	126
7.21. Pregunta test . . . . .	126
7.22. Pregunta test . . . . .	126
7.23. Pregunta test . . . . .	127
7.24. Problema . . . . .	127
7.25. Pregunta test . . . . .	127
7.26. Pregunta test . . . . .	128
7.27. Pregunta test . . . . .	128
7.28. Problema . . . . .	128
7.29. Pregunta test . . . . .	129

7.30. Pregunta test . . . . .	129
7.31. Pregunta test . . . . .	130
7.32. Problema . . . . .	130
7.33. Pregunta test . . . . .	131
7.34. Problema . . . . .	131
7.35. Pregunta test . . . . .	132
<b>8. Referencias</b>	<b>133</b>



# Capítulo 1

## Introducción

En estas páginas encontrarás preguntas tipo test y problemas de exámenes resueltos además de prácticas para la asignatura de Bioestadística en Grados de Ciencias de la Salud (Enfermería, Fisioterapia, Farmacia, etc.) y de exámenes BIR/FIR/QIR con las explicaciones correspondientes.

Para un eficaz, efectivo y eficiente uso del contenido de estas páginas, se recomienda el uso de *papel y bolígrafo*.

Estas páginas son un complemento del Curso de Bioestadística que incluye prácticas con Excel®.

En temario sobre el que basamos esta colección de problemas es el de la asignatura de la Universidad de Salamanca que incluye los siguientes bloques temáticos:

- Planteamiento de una investigación: Anatomía y Fisiología de la investigación
- Análisis Descriptivo y Gráfico de datos cuantitativos
- Análisis Inferencial. Aplicaciones.
- Regresión y correlación.
- Tablas de contingencia.
- Medidas de importancia clínica.

Después de cada pregunta tipo test hay un botón para ver la respuesta correcta así como un enlace a la Explicación correspondiente.

Respuesta correcta

Respuesta correcta



## Capítulo 2

# Planteamiento de una investigación: Anatomía y Fisiología de la investigación

En este capítulo se resolverán problemas relativos a:

- Diseño de una investigación
- Métodos de muestreo
- Métodos de recolección de datos
- Variables y Escalas de Medida
- Errores en la Investigación

### 2.1. Pregunta test

La proporción de individuos de una población que tiene una enfermedad en un momento dado de tiempo se denomina:

- a) Incidencia acumulada
- b) Densidad de incidencia
- c) Fracción atribuible
- d) Prevalencia
- e) Odds ratio

Respuesta correcta

Explicación

## 2.2. Pregunta test

Es una Variable Cualitativa Ordinal:

- a) Sexo (M, F).
- b) Bebe (no, poco, mucho).
- c) N° de muelas cariadas.
- d) Temperatura corporal.
- e) Raza (blanca, negra, amarilla).

Respuesta correcta

Explicación

## 2.3. Pregunta test

¿Cuál es el mejor tipo de estudio epidemiológico para estimar la prevalencia de una enfermedad crónica en una comunidad determinada?

- a) Estudio ecológico
- b) Ensayo clínico aleatorio
- c) Estudio transversal de la población
- d) Estudio de casos y controles
- e) Estudio de cohortes

Respuesta correcta

Explicación

## 2.4. Problema

1. El Ministerio de Sanidad, Servicios Sociales e Igualdad ha publicado recientemente el Barómetro Sanitario nacional, correspondiente a 2011.

**a)** Según los resultados del estudio, de los 7757 encuestados, 2487 afirman ser fumadores en la actualidad, 5262 afirman ser no fumadores y el resto no contesta. Construye la tabla de frecuencias correspondiente y realiza una gráfica representativa de la distribución de la muestra según el hábito de fumar.

**b)** A aquellos que habían declarado ser fumadores se les preguntó a continuación como habían influido en su consumo las medidas de la nueva Ley del tabaco. A continuación se muestra una tabla incompleta que resume las respuestas. Completa la tabla y realiza una gráfica representativa.

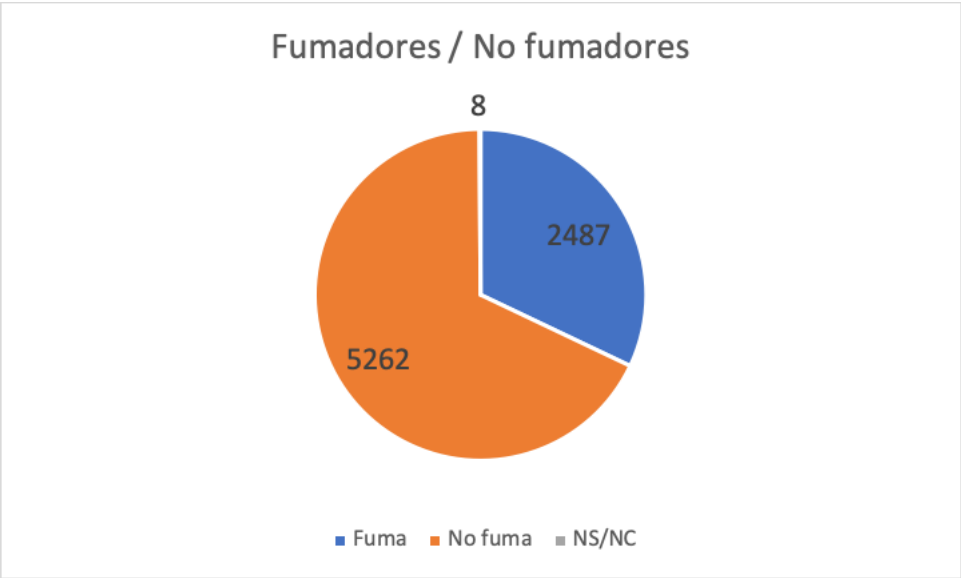
Influencia de la Ley del tabaco	$n_i$	$f_i$
Fuma menos que antes	125	0.68275
Fuma más que antes		
No ha influido en el consumo		
No contesta	28	

#### 2.4.1. Solución

a) Construimos la tabla de frecuencias con Excel© según esta Práctica del Curso de Bioestadística:

x	f	h	F	H
Fuma	2487	32,1%	2487	32,1%
No fuma	5262	67,8%	7749	99,9%
NS/NC	8	0,1%	7757	100,0%
	7757			

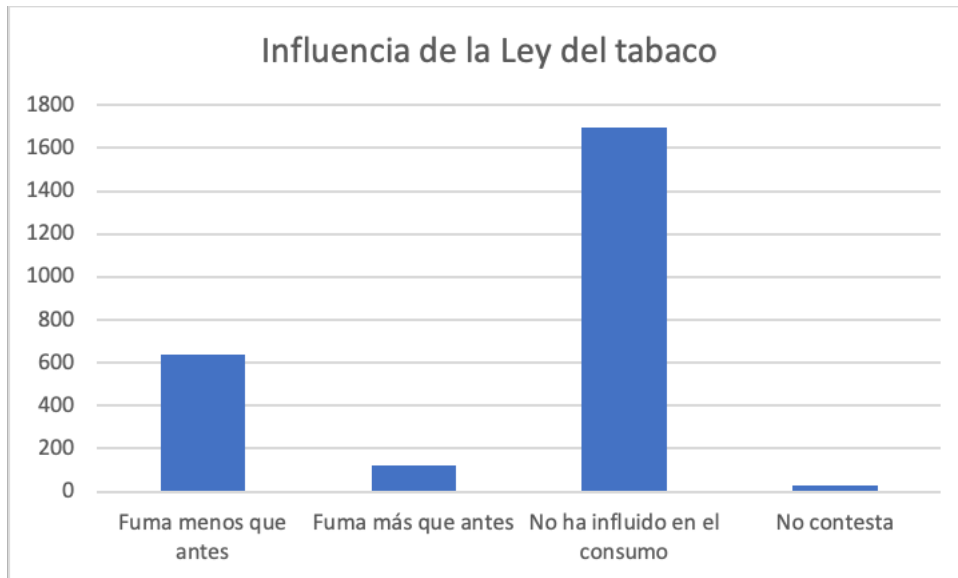
Con los datos de la tabla, hacemos un diagrama de sectores según esta Práctica del Curso de Bioestadística



b) En este caso, se pregunta a los que habían declarado ser fumadores por lo que  $N = 2487$ . Como la frecuencia relativa de la opción “No ha influido en el consumo” es 0,68275, la refuencia absoluta es  $0,68275 \cdot 2487 = 1697,99925 \approx 1698$  (no puede haber decimales). No tenemos más que calcular la frecuencia absoluta de la opción “Fuma menos que antes”,  $2487 - 28 - 1698 - 125 = 636$  y completar la tabla con las frecuencias relativas:

Influencia de la Ley del tabaco	$n_i$	$f_i$
Fuma menos que antes	636	0,2557301
Fuma más que antes	125	0,05026136
No ha influido en el consumo	1698	0,68275
No contesta	28	0,01125854
	2487	

Finalmente, obtenemos un diagrama de barras de los datos:



## 2.5. Pregunta test

Se realiza una auditoría de historias clínicas tomando una primera historia al azar y después sucesivamente, la que ocupa la vigésima posición detrás de la anterior. Este procedimiento de muestreo se denomina:

- a) Por conglomerados.
- b) Sistemático.
- c) Correlativo.
- d) Consecutivo.
- e) Equidistante.

Respuesta correcta

Explicación

## 2.6. Pregunta test

En una muestra de pacientes, el número de varones dividido entre el total de pacientes es:

- a) Una frecuencia relativa.
- b) Una frecuencia absoluta.
- c) Una variable cuantitativa.
- d) Una variable cualitativa.

- e) Un valor de la variable.

Respuesta correcta

Explicación

## 2.7. Pregunta test

Señale cuál de las siguientes afirmaciones es falsa:

- a) La aparición o no de bacterias en un cultivo es una variable dicotómica
- b) La estatura de un individuo es una variable cuantitativa discreta.
- c) El lugar que ocupa una persona entre sus hermanos (de menor a mayor edad) es una variable ordinal.
- d) El estado civil es una variable cualitativa.
- e) La glucemia es continua.

Respuesta correcta

Explicación

## 2.8. Problema

En base a la siguiente distribución de frecuencias relativas acumuladas de la variable  $X$  = “Número de contratos conseguidos en el mes de enero” obtenida de la observación de la actividad de 50 teleoperadores de una compañía de telefonía móvil, indique el número mínimo de contratos que tiene que haber conseguido un teleoperador para estar entre los 5 que han destacado más:

$X_i$	58	60	62	65	68	70	71
$H_i$	0.06	0.2	0.4	0.64	0.8	0.92	1

### 2.8.1. Solución

Al haber 50 teleoperadores, si tiene que estar entre los 5 que han destacado mas, debe dejar a 45 por detrás. Como  $\frac{45}{50} = 0.9$ , deberá superar al 90 %, es decir, estar por encima del 0.9 en la frecuencia relativa acumulada.

En la tabla vemos que para el valor 70 se alcanza la frecuencia relativa acumulada de 0.92 por lo que **para estar entre los 5 que más han destacado, deberá haber firmado, al menos, 70 contratos.**



## 2.9. Pregunta test

¿A qué fase del proceso de investigación pertenece la recogida, análisis e interpretación de los resultados?

- a) Fase conceptual.
- b) Fase Metodológica.
- c) Fase Empírica.
- d) Fase de análisis e interpretación de los datos.

Respuesta correcta

Explicación

## 2.10. Pregunta test

La incidencia de una enfermedad es:

- a) La relación entre enfermos y fallecidos
- b) La prevalencia multiplicada por la morbilidad
- c) Lo mismo que la prevalencia
- d) El nº de casos nuevos de esa enfermedad
- e) Ninguna de las anteriores

Respuesta correcta

Explicación

## 2.11. Pregunta test

En el caso de una variable ordinal, el número  $n$  de datos válidos es:

- a) La suma de las frecuencias absolutas.
- b) La frecuencia absoluta acumulada de la categoría más frecuente.
- c) La suma de las frecuencias relativas.
- d) La frecuencia relativa acumulada en la última categoría.
- e) La (a) y la (d) son ciertas.

Respuesta correcta

Explicación

### 2.12. Pregunta test

El nº de casos nuevos de una enfermedad que se desarrolla en una población en un periodo de tiempo determinado se conoce como:

- a) Densidad de incidencia
- b) Incidencia acumulada
- c) Prevalencia
- d) Fracción atribuible
- e) Riesgo relativo

Respuesta correcta

Explicación

### 2.13. Pregunta test

Se realiza un estudio con objeto de determinar el tiempo de supervivencia en pacientes con cáncer. Para ello de los dos hospitales existentes en una ciudad, se selecciona aleatoriamente uno de ellos, y se elige una muestra aleatoria de pacientes, atendiendo al tipo de cáncer: El muestreo realizado es:

- a) Sistemático.
- b) Aleatorio.
- c) Por conglomerados.
- d) Estratificado.
- e) Por conglomerados y estratificado.

Respuesta correcta

### 2.14. Pregunta test

En un estudio sobre problemas cervicales preguntamos a los pacientes acerca del tipo de almohada que usan. Las respuestas deberían ser consideradas como una variable:

- a) Cualitativa nominal
- b) Numérica
- c) Discreta
- d) Continua.
- e) Ordinal

Respuesta correcta

Explicación

## 2.15. Pregunta test

Al inicio de un estudio de cohortes ¿cómo está la población a estudiar?

- a) Todos los efectos del proceso que se estudia
- b) Todos sanos
- c) La cohorte expuesta sana y la no expuesta enferma
- d) La cohorte expuesta enferma y la no expuesta sana
- e) Ninguna de las anteriores

Respuesta correcta

Explicación

## 2.16. Problema

De la distribución de la variable  $X = \text{'Peso (en Kg)'} de un colectivo de adolescentes agrupada en 4 intervalos con límites superiores 60, 65, 70 y 75 se sabe que:$

- la mitad del colectivo pesa entre 65 y 70 kg
- una cuarta parte pesa como máximo 65 kg
- 9 adolescentes tiene un peso máximo de 60 kg
- 18 pesan entre 70 y 75 kg.

Calcula

- a) El número  $n$  de adolescentes entrevistados
- b) El porcentaje de adolescentes que pesan entre 55 y 60 kg
- c) El peso mínimo de la mitad de adolescentes con mayor peso
- d) Cuántos alumnos pesan como máximo, 65 kg

### 2.16.1. Solución

Vemos que tenemos mucha información que conviene organizar en forma de tabla. Empezamos escribiendo una tabla con los datos que tenemos:

Intervalo	$f_i$	$h_i$	$F_i$	$H_i$
[55, 60)	9		9	
[60, 65)				0.25
[65, 70)		0.50		
[70, 75)	18			1

A partir de estos datos, vamos a completar el resto.

Como el 25 % pesan menos de 65 y el 50 % entre 65 y 70, entonces el 75 % pesarán menos de 70 kg y el 25 % pesarán más de 70 hg.

Intervalo	$f_i$	$h_i$	$F_i$	$H_i$
[55, 60)	9		9	
[60, 65)				0.25
[65, 70)		0.50		0.75
[70, 75)	18	0.25		1

Así, el 25 % (la cuarta parte) del número  $n$  de adolescentes entrevistados es 18 por lo que  $n = 18 \cdot 4 = 72$ . El 50 % de 72 es 36 y, como hay 9 adolescentes entre 55 y 60 kg y como  $72 - 9 - 36 - 18 = 9$ , tendremos

Intervalo	$f_i$	$h_i$	$F_i$	$H_i$
[55, 60)	9	0.125	9	0.125
[60, 65)	9	0.125	18	0.25
[65, 70)	36	0.50	54	0.75
[70, 75)	18	0.25	72	1

Y a la vista de la tabla, podemos responder a las preguntas:

- a) Se entrevistaron a 72 adolescentes
- b) El 25 % de adolescentes pesa entre 55 y 60 kg
- c) El 50 % de los adolescentes con mayor peso están en los intervalos [65, 70) y [70, 75) y, como no podemos saber exactamente cuál es el peso menor de ese 50 %, **el peso mínimo de la mitad de adolescentes con mayor peso es de, al menos, 65 kg.**
- d) 18 alumnos pesan como máximo 65 kg

## 2.17. Pregunta test

¿Cuál de las siguientes características pertenece al paradigma naturalista?

- a) Pretende buscar la objetividad.
- b) El investigador interactúa con los sujetos investigados y los resultados se crean de esa interacción.
- c) Utilización de procesos deductivos.
- d) Importancia en el análisis estadístico.

Respuesta correcta

Explicación

## 2.18. Pregunta test

En un estudio sobre la enfermedad coronaria en la población española, se selecciona una muestra de individuos hipertensos y un grupo de control de no hipertensos. Se les sigue durante 5 años y se compara la incidencia de la enfermedad de ambos grupos ¿A qué tipo de diseño corresponde el estudio?

- a) Estudio de cohortes
- b) Estudio de casos y controles
- c) Estudio transversal
- d) Ensayo clínico
- e) Estudio ecológico

Respuesta correcta

Explicación

## 2.19. Pregunta test

Se desea estimar confidencialmente el número medio de veces que asiste a un servicio de salud los individuos de una población. Para ello se toman muestras aleatorias entre los individuos que asisten regularmente a los mismos. Esta técnica de muestreo es:

- a) Un muestreo aleatorio simple.
- b) Un muestreo aleatorio estratificado.
- c) Un muestreo aleatorio por conglomerados.
- d) Incorrecta.
- e) Ninguna de las anteriores.

Respuesta correcta

## 2.20. Pregunta test

El estudio estadístico en el que se pretenden extrapolar los datos de una muestra a la población se denomina:

- a) Estadística descriptiva.
- b) Estadística inferencial.
- c) Medidas de tendencia central.

d) Medidas de posición.

Respuesta correcta

Explicación

## 2.21. Problema

Con el objetivo de programar las actividades en un consultorio se obtiene información del número de consultas realizadas el año anterior:

	Mujeres	Hombres
Intervalo	n	n
1 - 3	18	22
4 - 6	39	31
7 - 9	53	46
8 - 10	45	40
11 - 13	53	35
14 - 16	39	29
17 - 20	18	26
Total	265	229

a) Indique el (o los) nombre(s) de las(s) variables(s) de la tabla e identifique sus categorías.

b) Indique el tipo de escala de las(s) variables (s) de la tabla.

c) ¿Qué porcentaje de pacientes realiza, al menos, 8 consultas?

### 2.21.1. Solución

a) La variable estudiada es **número de consultas realizadas el año anterior con 2 categorías, Mujeres y Hombres**.

b) Es una variable cualitativa discreta de escala ordinal con resultados agrupados en intervalos.

c) Para calcular el porcentaje pedido, vemos que:

- Mujeres con, al menos 8 consultas:  $45 + 53 + 39 + 18 = 155$
- Hombres con, al menos 8 consultas:  $40 + 35 + 29 + 26 = 130$
- Pacientes con, al menos 8 consultas:  $155 + 130 = 285$
- Total de pacientes:  $265 + 229 = 494$

Como  $\frac{285}{494} \approx 0.577$ , entonces **el 57.7 % de pacientes realiza, al menos, 8 consultas.**

## 2.22. Pregunta test

Elija la afirmación correcta sobre variables observadas en individuos:

- a) Poseer vivienda propia es una variable numérica.
- b) Poseer animales de compañía es una variable cualitativa.
- c) La nacionalidad es una variable ordinal.
- d) El tipo de almohada que usa es variable ordinal.
- e) La longitud de la cama donde duerme es variable discreta.

Respuesta correcta

Explicación

## 2.23. Pregunta test

La estadística en Ciencias de la Salud se utiliza para obtener información sobre situaciones de carácter:

- a) Determinista.
- b) Sistemático.
- c) Exhaustivo.
- d) Aleatorio.
- e) Excluyente.

Respuesta correcta

Explicación

## 2.24. Pregunta test

Elija la afirmación que pueda considerarse admisible al leer un estudio estadístico:

- a) Se estudió a una muestra en vez de a la población, para mayor precisión.
- b) Se estudió a la población para obtener información sobre la muestra.
- c) Se estudió a una muestra representativa de la población.
- d) Se estudiaron todas las variables de la población.
- e) Se observó a un individuo de cada variable.

Respuesta correcta

Explicación

## 2.25. Problema

En un estudio sobre supervivencia tras un tratamiento con quimioterapia para cierto tipo de cáncer ha sido registrado el tiempo transcurrido desde el inicio del tratamiento hasta el fallecimiento de los individuos. Los tiempos registrados se resumen en la tabla adjunta, agrupados por intervalos de 6 meses de amplitud:

Tiempo en meses	Nº de sujetos
0 - 6	28
6 - 12	14
12 - 18	10
18 - 24	8
24 - 30	8
30 - 36	5
Total	73

- Calcule las frecuencias relativas y porcentajes de los distintos intervalos.
- Calcule los puntos medios de los intervalos.
- Calcule las frecuencias absolutas y porcentajes acumulados
- Construya el histograma, polígono de frecuencias y polígono acumulativo

### 2.25.1. Solución

a), b) y c)

- f: frecuencias absolutas
- h: frecuencias relativas (porcentajes)
- F: frecuencias absolutas acumuladas
- H: frecuencias relativas acumuladas (en porcentaje)



Tiempo (meses)	pto. Medio	f	h (%)	F
[0, 6)	3	28	38,4%	28
[6, 12)	9	14	19,2%	42
[12, 18)	15	10	13,7%	52
[18, 24)	21	8	11,0%	60
[24, 30)	27	8	11,0%	68
[30, 36)	33	5	6,8%	73
		73		

Tabla construida siguiendo esta Práctica con Excel© del Curso de Bioestadística

d)

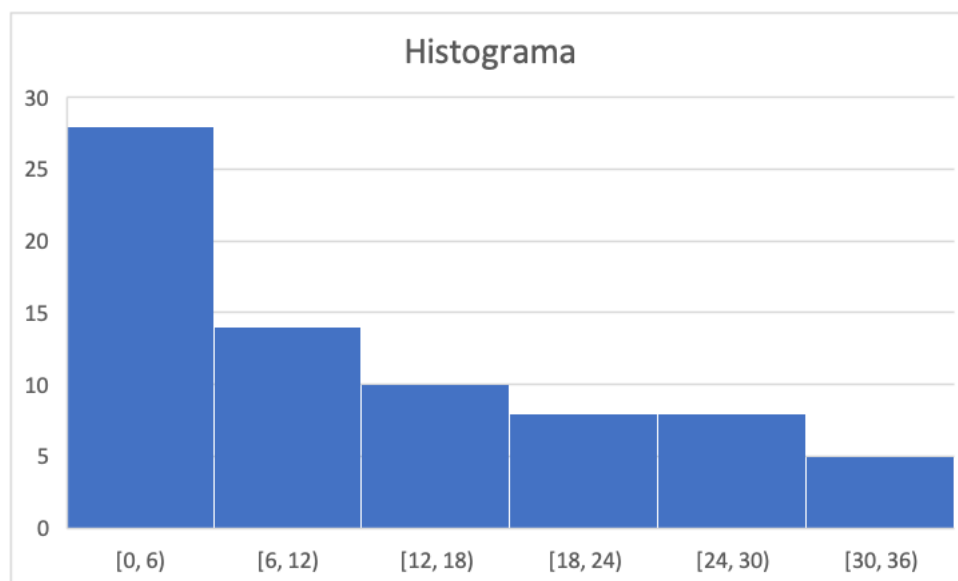
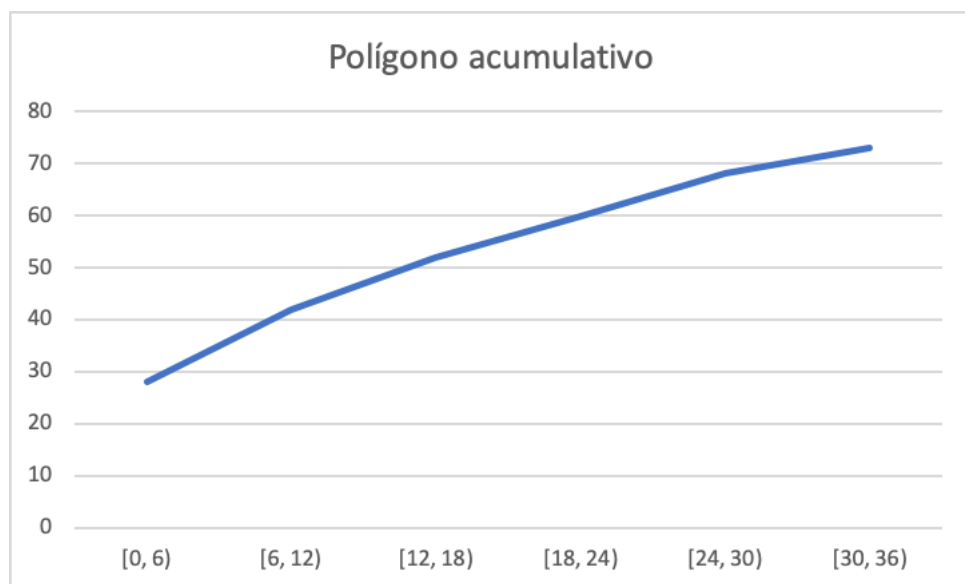
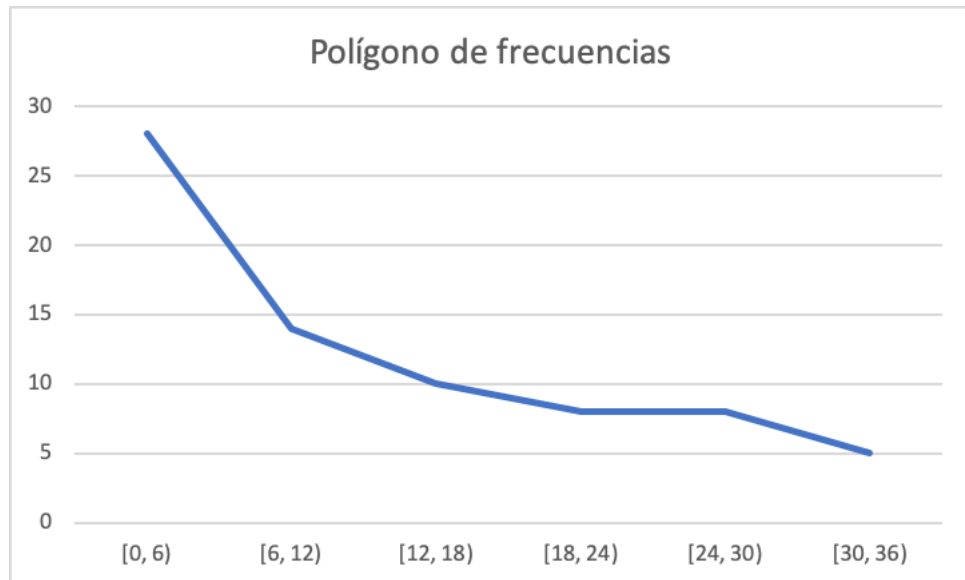


Gráfico construido siguiendo esta Práctica con Excel© del Curso de Bioestadística



Gráficos contruidos siguiendo esta Práctica con Excel© del Curso de Bioestadística

## 2.26. Pregunta test

Elija la afirmación correcta:

- a) Los valores de cualquier variable deben ser agrupados en intervalos.
- b) Las variables deben ofrecer valores que no se repitan en los diferentes individuos.
- c) Las modalidades de una variable deben poder ser observadas en todos los individuos.
- d) Los individuos pueden poseer diferentes modalidades de la misma variable.
- e) Todo lo anterior es falso.

Respuesta correcta

Explicación

## 2.27. Pregunta test

Elija la opción correcta.

- a) Un parámetro es algo calculado sobre cada individuo.
- b) Un parámetro es calculado sobre la muestra.
- c) Una variable se calcula sobre los parámetros de una población.
- d) Un estadístico se calcula sobre la población.
- e) Nada de lo anterior es correcto.

Respuesta correcta

Explicación

## 2.28. Pregunta test

Disponemos de la distribución de edades de los individuos de una población. El número de ellos que no es mayor de edad, es:

- a) Una frecuencia relativa.
- b) Una frecuencia absoluta.
- c) Una frecuencia acumulada.
- d) Una variable numérica.
- e) Una variable cualitativa.

Respuesta correcta

Explicación

## 2.29. Pregunta test

¿Cuál de las siguientes no es una característica de los estudios de cohortes?

- a) Son estudios observacionales
- b) El criterio de selección de los sujetos es la presencia o no de enfermedad
- c) Son estudios longitudinales
- d) Pueden ser prospectivos o retrospectivos
- 5) Tienen direccionalidad hacia delante

Respuesta correcta

Explicación

### 2.30. Pregunta test

Para un estudio epidemiológico sobre dolencias de suelo pélvico en mujeres en la provincia de Albacete, se decide seguir la siguiente estrategia de muestreo: Se elige aleatoriamente 10 poblaciones de la provincia, y en cada una de ellas se elige aleatoriamente 10 calles. Allí se elige aleatoriamente 5 números de la calle y se estudia a las mujeres que aceptan participar. El muestreo es:

- a) Aleatorio simple
- b) Por conglomerados.
- c) Estratificado.
- d) Sistemático.
- e) Estratificado y por conglomerados.

Respuesta correcta

Explicación

### 2.31. Problema

Los datos corresponden a las medidas de tensión arterial sistólica (en mm/Hg) registradas sobre 20 individuos fumadores de más de una cajetilla de cigarrillos diaria:

145, 185, 120, 160, 165, 160, 175, 145, 145, 175, 130, 130, 120, 110, 145, 150, 155, 160, 145, 135

- a) Construya la tabla de distribución de frecuencias para los datos originales.
- b) Construya la tabla de distribución de frecuencias por intervalos de amplitud 10 mm/Hg.
- c) Grafique la distribución de la variable.

#Solución

a)

Contando *con papel y boli*:

Tensión	110	120	130	135	145	150	155	160	165	175	185
f	1	2	2	1	5	1	1	3	1	2	1

b)

Se pide ahora considerar intervalos de amplitud 10 mm/Hg. Como el mínimo es 110 y el máximo 185, establecemos 8 intervalos:

Intervalo	f
[110, 120)	1
[120, 130)	2
[130, 140)	3
[140, 150)	5
[150, 160)	2
[160, 170)	4
[170, 180)	2
[180, 190)	1

Tabla construida siguiendo esta Práctica con Excel© del Curso de Bioestadística

c) Al tener los datos en intervalos, utilizamos un histograma:

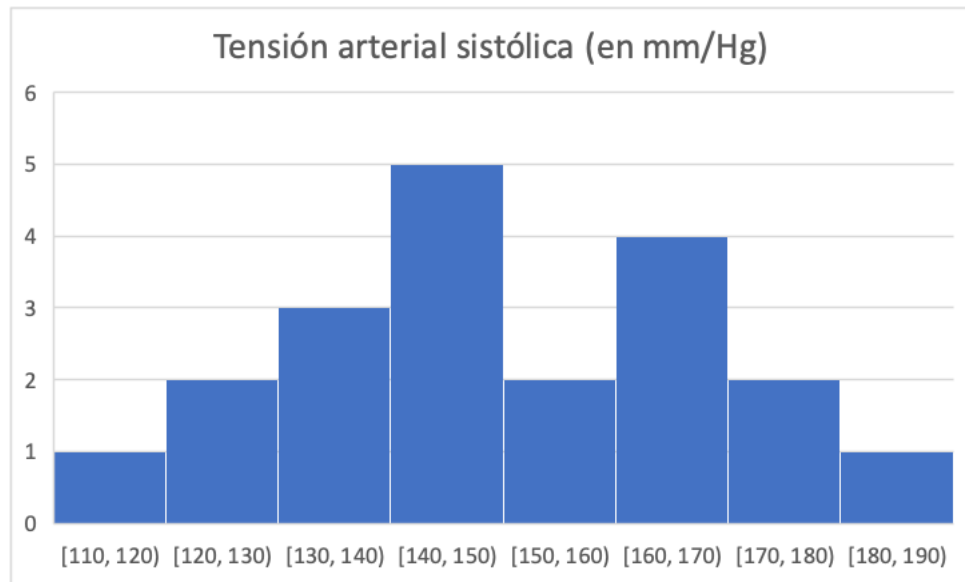


Gráfico construido siguiendo esta Práctica con Excel© del Curso de Bioestadística

### 2.32. Pregunta test

Conocemos la distribución de estudiantes entre las distintas facultades del campus Viriato. El número de estudiantes de Enfermería es:

- a) Una frecuencia relativa.
- b) Una frecuencia absoluta.
- c) Una frecuencia acumulada.
- d) Un porcentaje.
- e) Una variable cualitativa.

Respuesta correcta

Explicación

### 2.33. Pregunta test

Se llama parámetro a:

- a) Una función de valor numérico definida sobre alguna característica observable en los individuos de una población.
- b) Una función definida sobre los valores numéricos de una muestra.
- c) Cualquier variable observable de una población

- d) Las variables numéricas de la muestra
- e) Cualquier función sobre las variables observadas

Respuesta correcta

Explicación

## 2.34. Pregunta test

Respecto a los estudios de casos y controles es cierto que:

- a) Se analizan comparando la incidencia de una enfermedad o proceso en el grupo de casos respecto al grupo de controles
- b) Pueden escogerse varios controles para cada caso
- c) Una de las medidas de asociación que puede calcularse directamente en su análisis es el riesgo relativo
- d) Se denominan también estudios de prevalencia
- e) Es preferible seleccionar casos prevalentes en vez de casos incidentes de la enfermedad o proceso en estudio

Respuesta correcta

Explicación

## 2.35. Pregunta test

Si se realiza un estudio para evaluar el rendimiento académico de los estudiantes en una universidad y se selecciona una muestra de estudiantes de manera que cada estudiante tenga la misma probabilidad de ser elegido, ¿qué tipo de muestreo se está utilizando?

- a) Muestreo por conglomerados o por grupos
- b) Muestreo sistemático
- c) Muestreo aleatorio simple
- d) Muestreo no probabilístico
- e) Muestreo estratificado

Respuesta correcta

## 2.36. Pregunta test

El grado de satisfacción (poco/regular/mucho) con la política española la trataría como:

- a) una variable cualitativa nominal.

- b) una variable cuantitativa discreta.
- c) una variable cualitativa ordinal.
- d) una variable numérica continua.
- e) ninguna de las anteriores es correcta.

Respuesta correcta

Explicación

### 2.37. Pregunta test

Con respecto a la modalidades de una variable cualquiera:

- a) Pueden siempre agruparse en clases.
- b) Deben formar un sistema exhaustivo.
- c) No pueden agruparse en intervalos.
- d) No tienen porqué formar un sistema excluyente.
- e) Solo dos son correctas.

Respuesta correcta

Explicación

### 2.38. Pregunta test

¿Cuál es el estudio de elección para evaluar la eficacia de un nuevo tratamiento?

- a) Estudio de casos y controles
- b) Ensayo clínico aleatorio
- c) Estudio transversal
- d) Estudio de morbilidad
- e) Estudio de cohortes

Respuesta correcta

Explicación

### 2.39. Pregunta test

Cuando hablamos de número de cumpleaños que ha tenido una persona estamos ante:

- a) Una variable cualitativa ordinal.
- b) Una variable cualitativa nominal.
- c) Una variable cuantitativa discreta.



- d) Una variable cuantitativa continua.
- e) El número de cumpleaños no es una variable.

Respuesta correcta

Explicación

## 2.40. Pregunta test

Las pruebas de cribado son una actividad de:

- a) Prevención primaria
- b) Prevención secundaria
- c) Prevención terciaria
- d) Promoción de la salud
- e) Prevención primaria y promoción de la salud

Respuesta correcta

Explicación

## 2.41. Pregunta test

Las frecuencias acumuladas tienen sentido para:

- a) Variables ordinales
- b) Variables numéricas
- c) Variables nominales
- d) Todas son correctas.
- e) Las opciones a) y b) son correctas.

Respuesta correcta

Explicación

## 2.42. Pregunta test

Señale la respuesta INCORRECTA respecto a los estudios de cohortes:

- a) Pueden ser prospectivos y retrospectivos
- b) Son estudios observacionales y descriptivos
- c) Permiten establecer con claridad la secuencia temporal de los eventos de interés
- d) Permiten medir la incidencia de la enfermedad
- e) Permite medir los efectos de exposiciones infrecuentes en la población

Respuesta correcta

Explicación

### 2.43. Pregunta test

Disponemos de la distribución de edades de los individuos de una población. El número de ellos que tiene dos o menos hijos es:

- a) Una variable cualitativa.
- b) Una variable numérica.
- c) Una frecuencia acumulada.
- d) Son correctas a) y b)
- e) Ninguna es correcta.

Respuesta correcta

Explicación

### 2.44. Pregunta test

Deseamos conocer la opinión de los ciudadanos de Málaga sobre el sistema de salud pública. Para ello elegimos una muestra aleatoria de entre los abonados a telefónica. Entonces:

- a) La población de estudio es la de los ciudadanos de Málaga.
- b) La población de estudio es la de los abonados a telefónica.
- c) La población objetivo es la de los abonados a telefónica.
- d) El conjunto de abonados a telefónica son la muestra.
- e) Nada de lo anterior es cierto.

Respuesta correcta

Explicación

### 2.45. Pregunta test

Los principales objetivos de la estadística descriptiva son:

- a) Sintetizar la información contenida en los datos.
- b) Aportar resúmenes significativos de las distribuciones.
- c) Contribuye a la realización de los posteriores análisis estadísticos.
- d) Todos son correctos.

Respuesta correcta

Explicación

## 2.46. Pregunta test

Se diseña un estudio para evaluar el efecto sobre la salud de la exposición a los teléfonos móviles en el que durante 10 años se sigue a una población inicialmente sana ¿Qué tipo de diseño tiene ese estudio?

- a) Estudio casos y controles
- b) Estudio de cohortes
- c) Estudio transversal
- d) Estudio de casos
- e) Ensayo controlado

Respuesta correcta

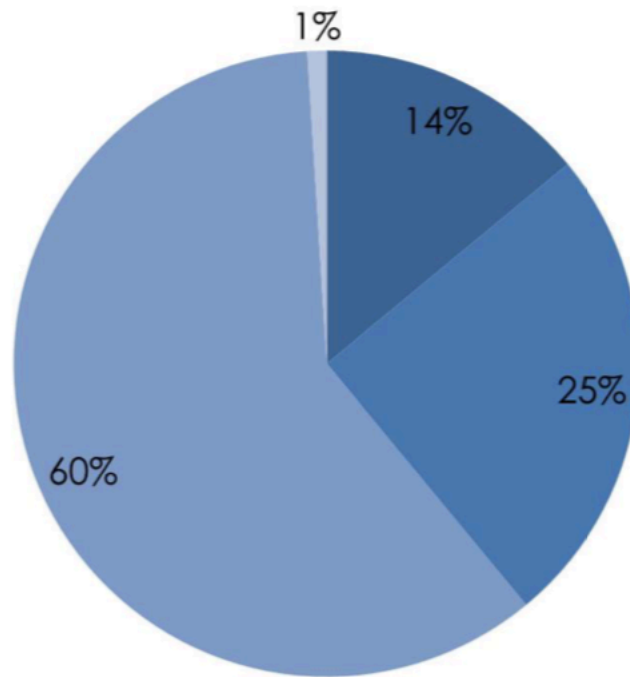
Explicación

## 2.47. Problema

Un estudio llevado a cabo por el Pew Research Center's Internet & American Life Project (<http://www.pewinternet.org>) tiene como objetivo analizar la actitud de los jóvenes en EEUU ante las redes sociales y su configuración de la privacidad. Para ello se ha llevado a cabo una encuesta entre usuarios de Facebook. A continuación se muestran dos gráficas con datos de dicho estudio.

### Configuración de la privacidad del perfil

■ Público ■ Parcialmente privado  
■ Privado (solo amigos) ■ No sabe



a) ¿Qué graficas aparecen representadas? ¿4A qué tipo de variables hacen referencia?

b) ¿Cuál es el tamaño muestral? ¿Cuántos encuestados tienen su perfil parcialmente privado? ¿Qué porcentaje de usuarios comparte la misma información con sus padres y amigos?

c) Supongamos que les preguntamos a los encuestados cuántos amigos tienen en Facebook. ¿Qué tipo de gráfico crees que deberías utilizar para resumir esa información?

**2.47.1. Solución**

a) El gráfico de la izquierda es un gráfico de sectores de la variable cualitativa “Configuración de la privacidad del perfil” con valores *Público*, *Paccialmente privado*, *Privado (solo amigos)* y *No sabe*.

El gráfico de la derecha es una diagrama de barras de la variable cualitativa “¿Compartes la misma información con tus padres y amigos?” con valores *Limita lo que pueden ver los padres*, *Padres y amigos ven lo mismo*, *Los padres no usan Facebook* y *No contesta*.

b) El tamaño muestral es la suma de las frecuencias absolutas, es decir,  $29 + 500 + 53 + 6 = 588$ .

Hay un 25% de encuestados que tienen su perfil parcialmente privado, es decir  $0.25 \cdot 588 = 147$ .

El porcentaje de usuarios que comparte la misma información con sus padres y amigos es  $\frac{500}{588} \approx 0.8503 = 85\%$ .

c) La variable “Número de amigos en Facebook” es cuantitativa discreta. Sin embargo, aparecerán muchísimos valores por lo que convendría agrupar los datos en intervalos, entre 0 y 100, entre 100 y 500, entre 500 y 1000 y más de 1000, por ejemplo. De este modo, habría que considerar a la variable como cuantitativa continua y para resumir esa información, usar un histograma.

**2.48. Pregunta test**

En un estudio sobre las causas del cáncer de pulmón se compararon los antecedentes de tabaquismo en los pacientes que habían desarrollado esta enfermedad con los de un grupo de personas y en la enfermedad ¿de qué tipo de estudio epidemiológico se trata?

- a) Estudio de casos y controles
- b) Estudio de cohortes
- c) Ensayo clínico aleatorio
- d) Estudio ecológico
- e) Estudio transversal

Respuesta correcta

Explicación

### 2.49. Pregunta test

El tipo de variable cualitativa que sus valores o categorías no pueden ser ordenados, se denomina:

- a) Variable ordinal.
- b) Variable discreta.
- c) Variable nominal.
- d) Variable continua.

Respuesta correcta

Explicación

### 2.50. Pregunta test

Se quiere hacer un estudio sobre el tabaquismo en la Comunidad de Cantabria. Queremos asegurarnos tener cierto número de individuos de la zona litoral, la capital y del interior, pues creemos que en cada una de esas zonas la incidencia es diferente. Haremos un muestreo:

- a) Aleatorio simple.
- b) Estratificado.
- c) Sistemático.
- d) Por grupos.
- e) No probabilístico.

Respuesta correcta

### 2.51. Pregunta test

¿A qué fase del proceso de investigación pertenece la relación de los objetivos e hipótesis de la investigación?

- a) Fase conceptual.
- b) Fase Metodológica.
- c) Fase Empírica.
- d) Fase de análisis e interpretación de los datos.

Respuesta correcta

Explicación

## 2.52. Pregunta test

Cuál es el estudio de elección para evaluar si existe una relación causa-efecto entre un factor y una enfermedad poco frecuente?

- a) Transversal
- b) Casos y controles
- c) Cohortes
- d) Serie de casos clínicos
- e) Correlaciones temporales

Respuesta correcta

Explicación

## 2.53. Pregunta test

Para tratar de establecer una relación causal entre el consumo de benzodiacepinas durante el embarazo y el riesgo de fisura palatina en el recién nacido se seleccionan madres de recién nacidos con fisura palatina y se compararon con madres de recién nacidos sanos en cuanto a los antecedentes de toma de benzodiacepinas ¿cuál es el tipo de diseño de estudio empleado?

- a) Casos y controles
- b) Estudio de cohortes
- c) Ensayo clínico aleatorizado
- d) Estudio ecológico
- e) Ensayo clínico cruzado

Respuesta correcta

Explicación

## 2.54. Pregunta test

Cuando la población objetivo y de estudio en un muestreo difieren mucho, entonces:

- a) Debe usarse el método de respuestas aleatorizadas.
- b) Pueden existir sesgos.
- c) No pueden seleccionarse unidades de muestreo.
- d) Se debe usar un muestreo no probabilístico.
- e) Nada de lo anterior es correcto.

Respuesta correcta

Explicación

### 2.55. Pregunta test

Un estudio en el que los participantes se asignan al azar para recibir un nuevo tratamiento o placebo se denomina:

- a) Cohortes
- b) Casos y controles
- c) Transversal
- d) Ensayo clínico
- e) Serie de casos clínicos

Respuesta correcta

Explicación

### 2.56. Pregunta test

¿Cómo se denomina el ensayo clínico en el que los pacientes, los investigadores y los profesionales sanitarios implicados sanitarios implicados en la atención de los pacientes desconocen el tratamiento asignado?

- a) Enmascaramiento
- b) Triple ciego
- c) Abierto
- d) Simple ciego
- e) Doble ciego

Respuesta correcta

Explicación

### 2.57. Pregunta test

¿A qué se debe el sesgo de selección?

- a) A falta de sinceridad en los individuos de la muestra.
- b) A las diferencia existente entre diversas muestras.
- c) A la diferencia entre la población de estudio y la población objetivo.
- d) A no usar la técnica de respuesta aleatorizada.
- e) A nada de lo anterior.

Respuesta correcta

Explicación



## Capítulo 3

# Análisis Descriptivo y Gráfico de datos cuantitativos

En este capítulo se resolverán problemas relativos a:

- Medidas de tendencia central: Media, Moda, Mediana.
- Medidas de dispersión: Recorrido, Varianza, Desviación típica, Coeficiente de variación, Recorrido intercuartílico. Error estándar.
- Representaciones gráficas: Diagrama de barras, Pictogramas, Cartogramas,

### 3.1. Pregunta test

Cuál de las siguientes medidas define mejor la tendencia central de los datos: 5, 4, 42, 4, 6

- a) La mediana.
- b) La media.
- c) El sesgo
- d) El rango.
- e) La proporción.

Respuesta correcta

Explicación

### 3.2. Pregunta test

Los diagramas de sectores son muy útiles para comparar:

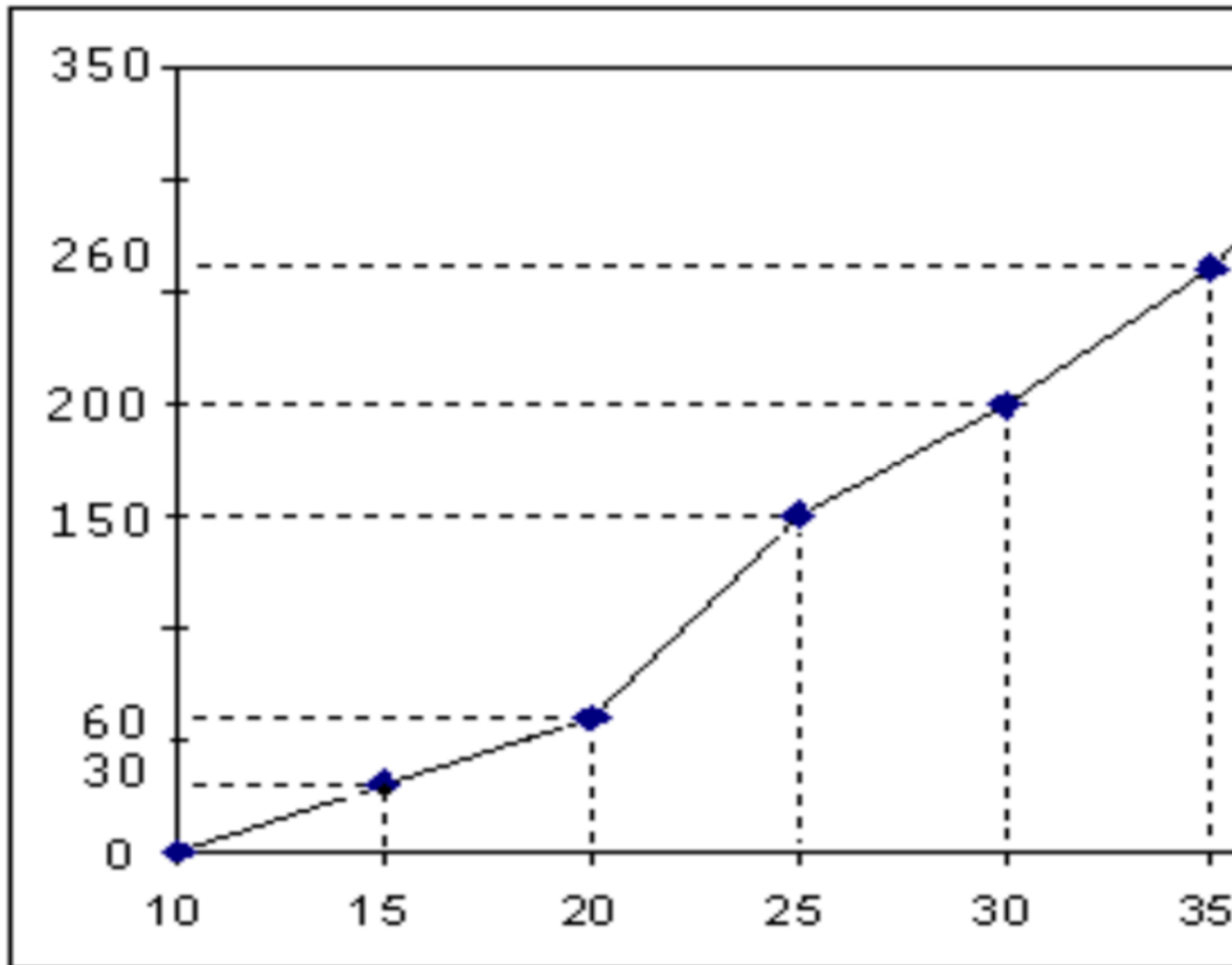
- a) Dos variables cualitativas en una población.
- b) Dos variables cuantitativas en una población.
- c) Una variable cualitativa en dos poblaciones.
- d) Una variable cuantitativa en dos poblaciones.
- e) Una variable cuantitativa con otra cualitativa.

Respuesta correcta

Explicación

### 3.3. Problema

El siguiente polígono de frecuencias absolutas acumuladas corresponde a la distribución de frecuencias de la variable  $X$  = “Duración en minutos de una consulta médica especializada”.



- a) ¿Qué porcentaje de consultas han durado como máximo 30 minutos?  
b) ¿Qué porcentaje de consultas han durado entre 25 y 30 minutos?

### 3.3.1. Solución

Al ser un polígono de frecuencias absolutas acumuladas, vemos que se han contabilizado un total de 350 consultas.

- a) Vemos que hay 200 consultas que han durado como máximo 30 minutos. Como  $\frac{200}{350} \approx 0.57$ , entonces **un 57% de las consultas han durado entre 25**

y **30 minutos**.

b) Entre 25 y 30 minutos, han habido  $200 - 150 = 50$  consultas. Como  $\frac{50}{350} \approx 0.14$ ,  
**un 14 % de las consultas han durado entre 25 y 30 minutos**

### 3.4. Pregunta test

En cuanto a la presentación ordenada del estudio de una variable aislada:

- a) Lo más informativo es mostrar las medidas de tendencia central.
- b) Lo más informativo es mostrar las medidas de dispersión.
- c) Se deben presentar todos los valores observados de la variable, uno a uno, de menor a mayor.
- d) Las representaciones gráficas dan más información que las tablas de frecuencia.
- e) A veces no tiene sentido usar frecuencias acumuladas.

Respuesta correcta

Explicación

### 3.5. Pregunta test

La mediana es una medida de tendencia central que se usa cuando:

- a) Los datos son impares.
- b) La muestra es asimétrica.
- c) La muestra es heterogénea.
- d) La muestra es simétrica.
- e) La muestra es homogénea.

Respuesta correcta

Explicación

### 3.6. Pregunta test

El Coeficiente de Variación se calcula:

- a) Multiplicando la Varianza por la Media.
- b) Dividiendo la Desviación Típica por la Media.
- c) Dividiendo la Media por la Desviación Típica.
- d) Dividiendo la Media por la Varianza.
- e) Multiplicando la Desviación Típica por la Media.

Respuesta correcta

Explicación

### 3.7. Pregunta test

En las representaciones gráficas de variables cualitativas, la regla fundamental a tener en cuenta es:

- a) Las alturas en cada modalidad son proporcionales al valor de la variable.
- b) Las áreas para cada modalidad son proporcionales al valor de la variable.
- c) Las áreas para cada modalidad son proporcionales a las frecuencias acumuladas.
- d) Las áreas para cada modalidad son proporcionales a las frecuencias absolutas o relativas.
- e) Las alturas para cada modalidad son proporcionales a las frecuencias acumuladas.

Respuesta correcta

Explicación

### 3.8. Problema

En un estudio para evaluar la eficacia de cierto programa educativo sobre salud bucodental, se preguntó a los asistentes con qué frecuencia acudían al dentista por razones preventivas. Dos años después del programa educativo se volvió a preguntar a los asistentes al programa la misma pregunta. En la tabla adjunta se describen los resultados obtenidos:

Nº de veces en el último año	Nº de sujetos antes del programa	Nº de sujetos después del programa
0	20	9
1	10	21
2	3	7
3	4	2
4	3	1
Total	40	40

Construya los diagramas de barras representando gráficamente las distribuciones del número de veces que fueron al dentista en el último año, antes y después del programa educativo. Compare los resultados.

### 3.8.1. Solución

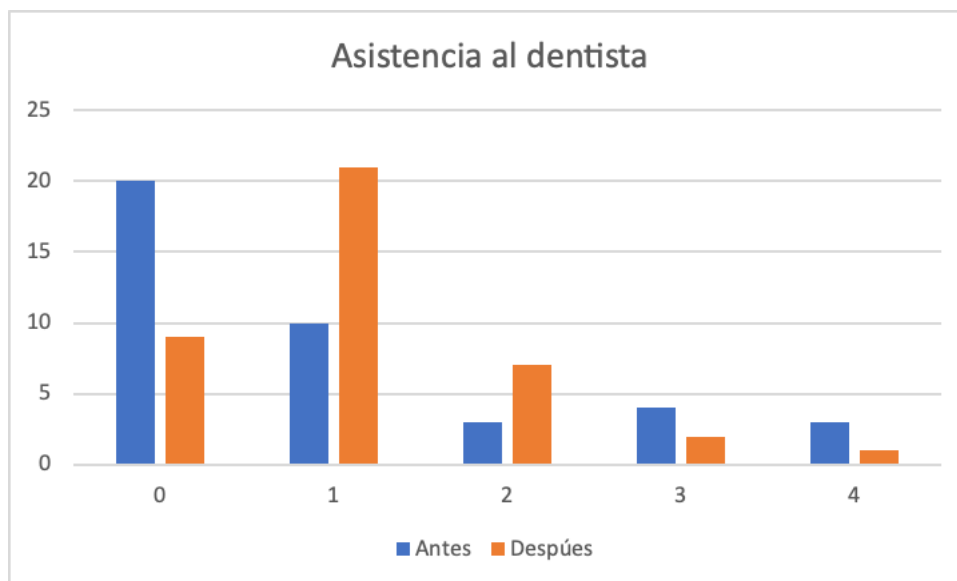


Gráfico construido siguiendo esta Práctica con Excel© del Curso de Bioestadística

Vemos que, antes de participar en el programa, la mayoría de participantes o no acudía al dentista o lo hacía una vez al año. Tras el programa, vemos que ha disminuido drásticamente el número de personas que no acuden al dentista pasando a ser la mayoría los que acuden 1 o 2 veces.

Se observa que, tras el programa, bajan los participantes que acudían 3 o 4 veces.

### 3.9. Pregunta test

Entre las representaciones gráficas para variables cualitativas tenemos:

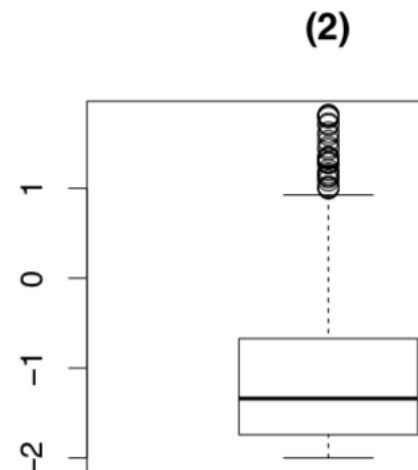
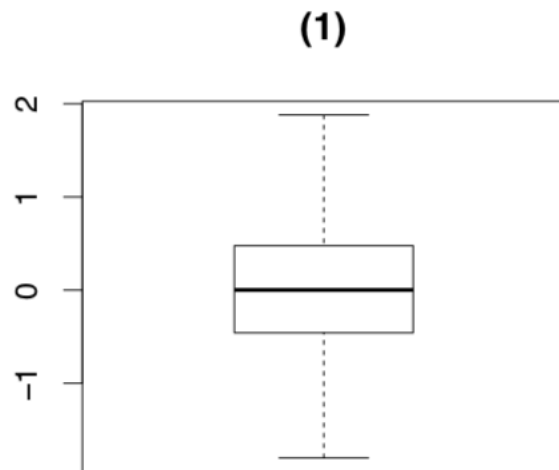
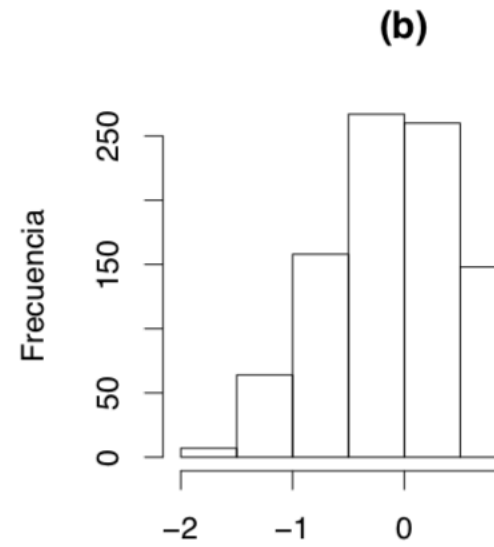
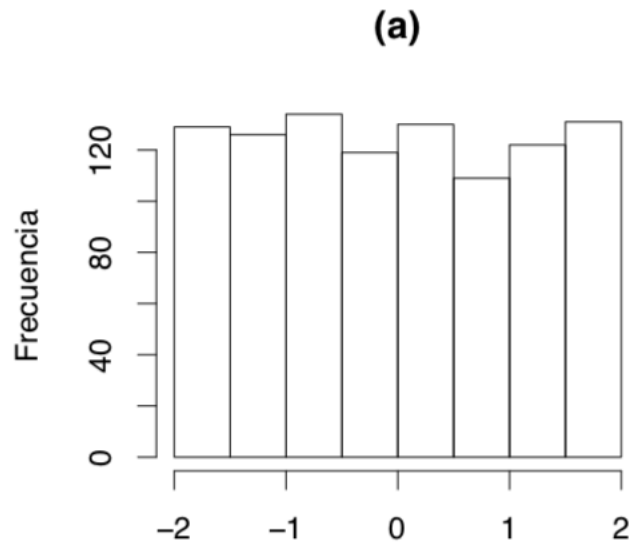
- a) Histogramas.
- b) Diagramas integrales.
- c) Diagramas diferenciales.
- d) Diagramas de cajas y bigotes.
- e) Nada de lo anterior.

Respuesta correcta

- d) Explicación

### 3.10. Pregunta test

A continuación aparecen representados los histogramas y diagramas de cajas de tres conjuntos de datos distintos. Empareja cada histograma con el diagrama de cajas que le corresponde.



- a) a-1, b-2, c-3
- b) a-3, b-2, c-1
- c) a-3, b-1, c-2
- d) a-1, b-3, c-2
- e) a-2, b-1, c-3

Respuesta correcta

Explicación



### 3.11. Pregunta test

De los siguientes conceptos indique el que no tenga sentido:

- a) Diagrama de barras para la variable “Grupo sanguíneo”
- b) Pictograma para la variable “Altura”
- c) Diagrama integral para la variable “Nivel de colesterol”
- d) Diagrama de sectores para la variable “Sexo”
- e) Histograma para la variable “Peso”

Respuesta correcta

Explicación

### 3.12. Problema

Investigadores de un centro hospitalario planificaron un estudio para determinar la eficacia de cierto complemento dietético en el tratamiento de la artritis reumatoide. El estudio se realizó sobre 50 pacientes con esta enfermedad, administrando a la mitad el complemento dietético y al resto un placebo durante veinte semanas. De los 25 pacientes que recibieron en complemento dietético, 18 presentaron mejoría, mientras que esto ocurrió en 10 de los que recibieron el placebo. Estructure los datos en una tabla de distribución de frecuencias conjuntas y calcule e interprete los porcentajes por filas y columnas.

#### 3.12.1. Solución

Comenzamos escribiendo la tabla de doble entrada con las variables y los datos proporcionados:

	Mejora	No mejora	Total
Complemento	18		25
Placebo	10		25
Total			50

Completamos los datos y añadimos los porcentajes por filas y columnas.

	Mejora	No mejora	Total	
<b>Complemento</b>	<b>18</b>	<b>7</b>	<b>25</b>	<b>5</b>
<b>Placebo</b>	<b>10</b>	<b>15</b>	<b>25</b>	<b>5</b>
<b>Total</b>	<b>28</b>	<b>22</b>	<b>50</b>	
	<b>56%</b>	<b>44%</b>		

A la vista de los datos por filas, vemos que se suministró a la mitad de los participantes el complemento dietético y a la otra mitad el placebo (50 % vs 50 %). Por columnas, vemos que un 56 % de participantes mejoran frente a un 44 % que no lo hacen. Esta diferencia es debida, sobre todo, al grupo que recibió el complemento donde la mejoría es bastante mayor que en el grupo que recibió el placebo (18 vs 10).

### 3.13. Pregunta test

Si queremos representar gráficamente los porcentajes de una variable cuantitativa continua debemos usar:

- a) Pictogramas
- b) Diagrama de barras
- c) Diagrama diferencial acumulado
- d) Histograma
- e) No existe gráfica posible

Respuesta correcta

Explicación

### 3.14. Pregunta test

Los gráficos indicados para variables cualitativas son:

- a) Los diagramas de barras y los histogramas
- b) Los diagramas de barras, los de sectores y los pictogramas
- c) Los histogramas y pictogramas
- d) Sólo los diagramas de barras
- e) Los diagramas integrales

Respuesta correcta

Explicación

### 3.15. Problema

El volumen corpuscular medio (VCM) es uno de los parámetros calculados en un examen de conteo sanguíneo completo. El VCM indica el tamaño de los glóbulos rojos y se mide en fentolitros. A continuación se muestran los valores de VCM de 19 pacientes que se han sometido a un examen de conteo sanguíneo completo.

83; 77; 82; 84; 85; 92; 92; 93; 91; 86; 89; 109; 81; 79; 81; 88; 110; 90; 80;

a) Construye la tabla de frecuencias y representa el histograma correspondiente.

b) Dibuja el boxplot de estos datos.

a) Al pedirnos el histograma, tenemos que agrupar los datos en intervalos. Vemos que el valor mínimo es 77 y el máximo 110. Así, vamos a considerar intervalos de amplitud 5 comenzando por el intervalos [75, 80). En el último intervalos, incluimos al extremo superior 110:

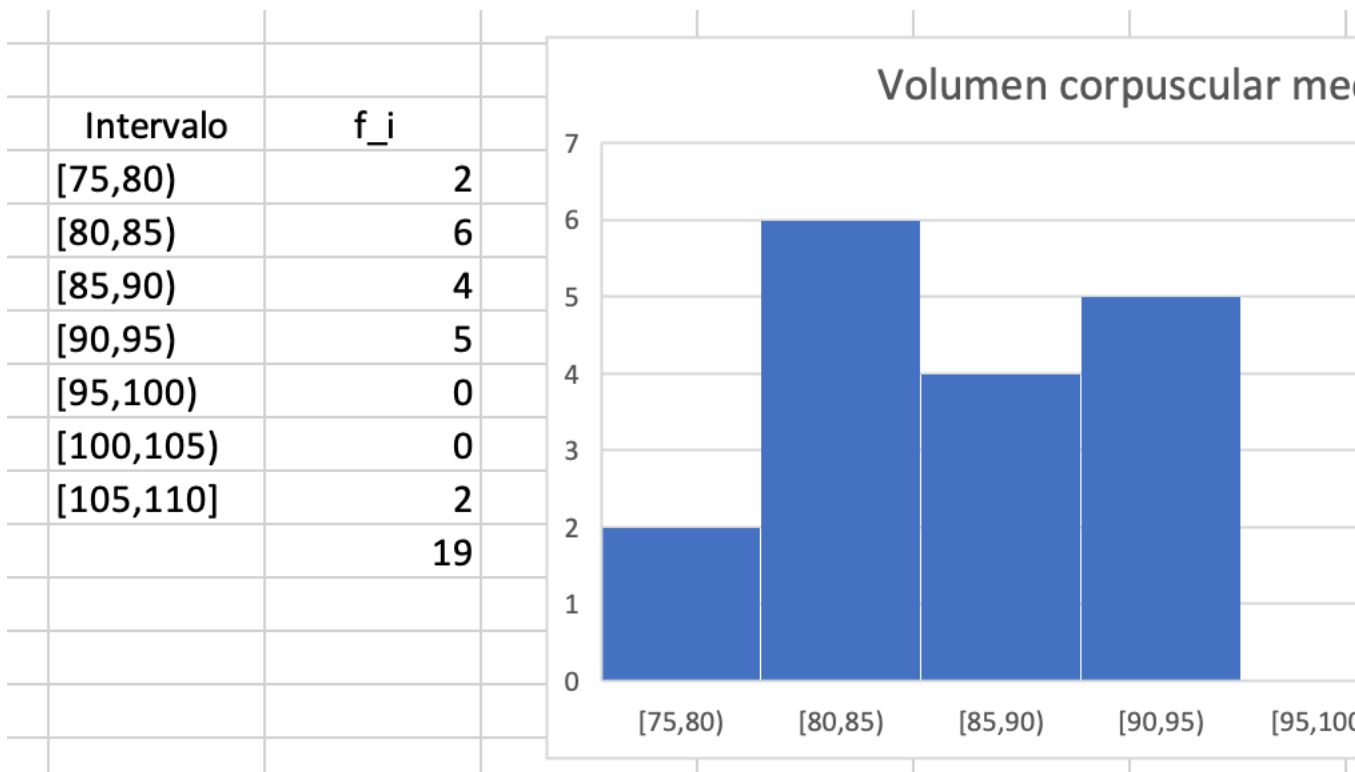
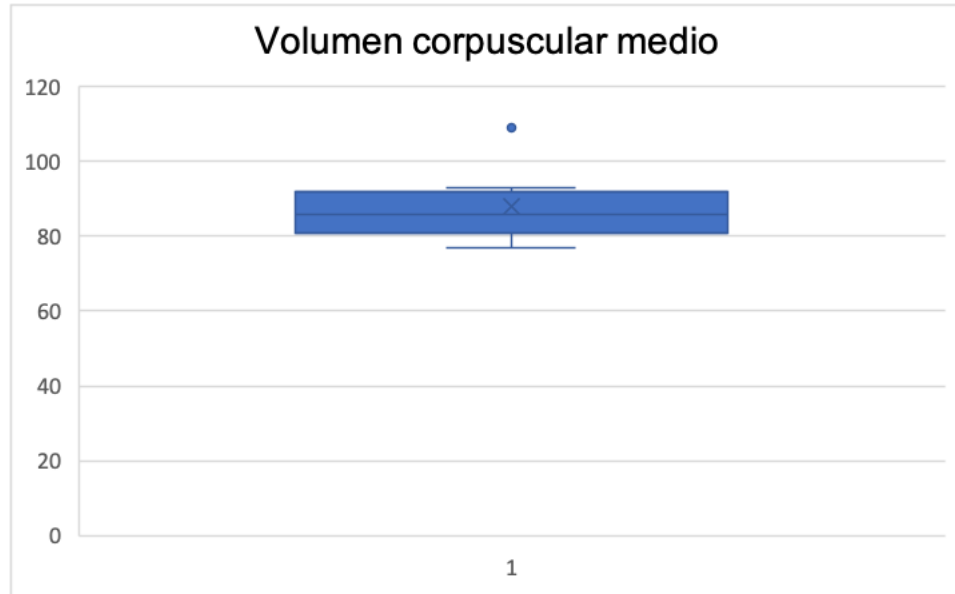


Gráfico obtenido siguiendo esta Práctica del Curso de Bioestadística

b) Obtenemos el boxplot (diagrama de cajas y bigotes) con Excel©



### 3.16. Pregunta test

¿Qué gráfico elegirías para representar una las respuestas a una encuesta sobre el número de hijos que tiene la población?

- a) Histograma
- b) Diagrama de sectores
- c) Pictograma
- d) Diagrama de Barras
- e) Ninguna de las anteriores

Respuesta correcta

Explicación

### 3.17. Pregunta test

Para comparar la variabilidad relativa de la tensión arterial diastólica y el nivel de colesterol en sangre de una serie de individuos, utilizamos

- a) Las desviaciones típicas.
- b) Los rangos.
- c) Los coeficientes de variación.

- d) La diferencia de las medias.
- e) La diferencia de las varianzas.

Respuesta correcta

Explicación

### 3.18. Pregunta test

La media aritmética de una variable cuantitativa:

- a) Es siempre un valor de la variable.
- b) No tiene sentido calcularla para variables discretas.
- c) Es el valor más representativo de una modalidad.
- d) Si la variable es discreta, puede no ser única.
- e) Existe siempre.

Respuesta correcta

Explicación

### 3.19. Pregunta test

Las siguientes medidas son todas de centralización, excepto:

- a) La media.
- b) La moda.
- c) La mediana.
- d) Rango intercuartílico.
- e) El percentil 50.

Respuesta correcta

Explicación

### 3.20. Pregunta test

En un estudio descriptivo se obtiene una que el peso tiene una media de 60 kg y una desviación típica de 20 kg., mientras que la media de las edades es 15 años, con una desviación típica de 5 años. Entonces:

- a) Hay más dispersión en pesos que en edades.
- b) Hay más dispersión en edades que en pesos.
- c) Peso y edad están dispersos de modo equivalente.
- d) No tiene sentido compararlos al no coincidir las unidades de medida.

- e) Para comparar ambas dispersiones debemos usar la covarianza.

Respuesta correcta

Explicación

### 3.21. Pregunta test

¿Cuál de las siguientes características no se corresponde con el concepto de mediana?

- a) Es el centro de gravedad de la distribución.
- b) No se ve afectada por los valores extremos.
- c) Deja por debajo el mismo número de datos que por encima.
- d) Es el segundo cuartil.
- e) Todo lo anterior se corresponde con la mediana.

Respuesta correcta

Explicación

### 3.22. Pregunta test

Señale cuál de las siguientes afirmaciones es falsa:

- a) La media aritmética es siempre el centro de gravedad de la distribución.
- b) En una distribución continua simétrica, media y mediana coinciden.
- c) La media aritmética cambia cuando cambia algún dato.
- d) La mediana no siempre cambia cuando lo hace algún dato.
- e) En las distribuciones continuas simétricas todas las medidas de centralización coinciden.

Respuesta correcta

Explicación

### 3.23. Pregunta test

El coeficiente de variación:

- a) Permite comparar la dispersión de dos poblaciones.
- b) Es menor que la media.
- c) Es menor que la desviación típica.
- d) No depende de la media ni la desviación típica.
- e) Depende de la escala que se use al medir la variable.

Respuesta correcta

Explicación

### 3.24. Pregunta test

Se pide a unos enfermos que valoren su grado de mejoría tras un tratamiento en una escala de 1 a 5. De la siguiente colección de posibilidades, cuál cree que resume mejor los mismos:

- a) Media, Mediana y Moda.
- b) Percentil 25, Percentil 50, Percentil 75.
- c) Media y desviación típica.
- d) Mediana y desviación típica.
- e) Rango

Respuesta correcta

Explicación

### 3.25. Pregunta test

De las siguientes medidas, cuáles podría utilizar para argumentar en favor o en contra de la asimetría de la variable edad:

- a) Percentil 25 y percentil 75.
- b) Media y Percentil 60.
- c) Media y mediana
- d) Media y desviación típica.
- e) Ninguna de las anteriores.

Respuesta correcta

Explicación

### 3.26. Pregunta test

La pregunta: ¿qué nivel de colesterol sólo es superado por el 5% de los individuos?, tiene por respuesta:

- a) El percentil 95.
- b) El percentil 5.
- c) Los percentiles 2,5 y 97,5
- d) 95%.
- e) Nada de lo anterior.

Respuesta correcta

Explicación

### 3.27. Pregunta test

Qué peso no llega a alcanzar el 40 % de los individuos de una población:

- a) El 40 %.
- b) El 60 %.
- c) El percentil 60.
- d) El percentil 40.
- e) Los percentiles 20 y 60.

Respuesta correcta

Explicación

### 3.28. Pregunta test

La media aritmética de una variable discreta:

- a) Puede ser un valor de la variable.
- b) No debería ser utilizada como medida de centralización.
- c) Es lo mismo que el percentil 50.
- d) Puede no ser única.
- e) Todo lo anterior es falso.

Respuesta correcta

Explicación

### 3.29. Pregunta test

Se pregunta a los individuos su opinión sobre una cuestión, pudiendo valorar estos su respuesta en términos de: en contra, en parte a favor, muy a favor, totalmente de acuerdo. Elija la afirmación correcta:

- a) Podemos calcular la media.
- b) Podemos calcular el coeficiente de variación.
- c) La variable es de tipo ordinal
- d) La variable es de tipo cualitativo nominal.
- e) Nada de lo anterior es cierto.



Respuesta correcta

Explicación

### 3.30. Pregunta test

En una población, el 70 % de las alturas consideradas “más normales” se encuentran:

- a) Por encima del percentil 70.
- b) Por debajo del cuantil 0,30
- c) Entre el percentil 30 y el 70
- d) Entre el percentil 15 y el 85.
- e) Entre la media y la mediana.

Respuesta correcta

Explicación

### 3.31. Pregunta test

Las medidas de centralización, en cuanto a la información que ofrecen sobre una variable numérica, preferimos (por orden, de peor a mejor):

- a) media, mediana, moda
- b) moda, media, mediana
- c) media, moda, mediana.
- d) No se puede en general recomendar una como mejor que las otras.
- e) Todo lo anterior es falso.

Respuesta correcta

Explicación

### 3.32. Pregunta test

Si una muestra posee valores anómalos, de las siguientes cuál usarías como medida de dispersión:

- a) Varianza.
- b) Desviación típica.
- c) Rango intercuartílico.
- d) Rango.
- e) Máximo y coeficiente de variación.

Respuesta correcta

Explicación

### 3.33. Pregunta test

Si queremos saber cómo de disperso está una variable relativamente con respecto a la magnitud de los valores centrales de la misma, usaremos:

- a) Varianza.
- b) Desviación típica.
- c) Rango intercuartílico.
- d) Rango.
- e) Coeficiente de variación.

Respuesta correcta

Explicación

### 3.34. Pregunta test

Si el coeficiente de asimetría en una población presenta el valor 0,99 entonces:

- a) La distribución presenta una cola a la derecha.
- b) La distribución presenta una cola a la izquierda.
- c) La distribución es más apuntada que la normal.
- d) La distribución es menos apuntada que la normal.
- e) La distribución es prácticamente simétrica.

Respuesta correcta

Explicación

### 3.35. Pregunta test

Si la media del peso en una población es 60 kg. y la mediana 65kg., entonces afirmamos que la distribución del peso en la población es:

- a) Platicúrtica.
- b) Mesocúrtica.
- c) Leptocúrtica.
- d) Asimétrica.
- e) Unimodal.

Respuesta correcta

Explicación

### 3.36. Pregunta test

Si el coeficiente de asimetría en una población presenta el valor -5,22 entonces:

- a) La distribución presenta una cola a la derecha.
- b) La distribución presenta una cola a la izquierda.
- c) La distribución es más apuntada que la normal.
- d) La distribución es menos apuntada que la normal.
- e) Ese valor de asimetría es imposible.

Respuesta correcta

Explicación

### 3.37. Pregunta test

Medimos el número de glóbulos rojos y el de blancos en cada individuo de una población. Se observa determinada variabilidad en esas cantidades. Queremos saber de qué tipo de célula se presenta mayor variabilidad

- a) Compararemos las desviaciones típicas.
- b) Compararemos los rangos.
- c) Estudiaremos la covarianza.
- d) Estudiaremos el coeficiente de correlación lineal de Pearson.
- e) Compararemos los coeficientes de variación.

Respuesta correcta

- e) Explicación

### 3.38. Pregunta test

En una muestra de 1000 mujeres se estudia su número de hijos. Si quiero tener el máximo de información sobre la variable del estudio, preferimos:

- a) Media, Mediana y Moda.
- b) Percentil 25, Percentil 50, Percentil 75.
- c) Media y desviación típica.
- d) Media, mediana, cuartiles, asimetría, curtosis y desviación típica.
- e) Distribución de frecuencias

Respuesta correcta

Explicación

### 3.39. Pregunta test

El 3 % de los individuos tiene una altura superior a 190cm. El 5 % mide menos de 150cm. Conocemos:

- a) El percentil 3
- b) El cuantil 0,06
- c) El percentil 95
- d) El percentil 97
- e) Nada de lo anterior.

Respuesta correcta

Explicación

### 3.40. Pregunta test

Respecto a las medidas de centralización:

- a) La media no debe usarse en distribuciones muy asimétricas.
- b) La moda puede no ser única.
- c) En distribuciones simétricas media, mediana y moda coinciden.
- d) Las tres anteriores son correctas.
- e) Sólo la a) y la b) son correctas

Respuesta correcta

Explicación1 y Explicación2

### 3.41. Pregunta test

El coeficiente de asimetría en una población vale 3. Elija la afirmación correcta:

- a) La distribución presenta una cola a la derecha.
- b) La distribución presenta una cola a la izquierda.
- c) La distribución es simétrica.
- d) La distribución es más apuntada que la normal
- e) La media es igual a la mediana.

Respuesta correcta

Explicación

**3.42. Pregunta test**

¿Cuál de las siguientes medidas define mejor la tendencia central de los datos:  
1, 2, 4, 5, 9, 1, 3, 9, 400?

- a) Media.
- b) Cuantil 0,5.
- c) Moda
- d) Desviación típica.
- e) Ninguna de las anteriores.

Respuesta correcta

Explicación

**3.43. Pregunta test**

De las siguientes variables ¿con cuáles NO puedo calcular la media?

- a) temperatura corporal
- b) pH del estómago
- c) grupo sanguíneo
- d) número de glóbulos rojos
- e) edad

Respuesta correcta

- c) Explicación

**3.44. Pregunta test**

De las siguientes variables con cuál sería menos adecuado un diagrama de barras?

- a) Número de hijos
- b) Número de coches que posee la familia
- c) Número de cigarros fumados al día
- d) Número de glóbulos rojos
- e) Número de mascotas.

Respuesta correcta

Explicación

### 3.45. Pregunta test

Cuál es la mediana de los siguientes datos 22, 5, 9, 11, 10, 14, 7

- a) 5
- b) 9
- c) 11
- d) 10
- e) 14

Respuesta correcta

Explicación

### 3.46. Pregunta test

Si el cuantil 0,9 del peso es 70 kilogramos, quiere decir esto:

- a) Que una frecuencia del 70 % individuos pesa más de 70 kilogramos.
- b) Que una frecuencia del 90 % de individuos pesa más de 70 kilogramos.
- c) Que una frecuencia del 90 % individuos pesa menos de 70 kilogramos.
- d) Que una frecuencia de 70 % de individuos pesa menos de 90 kilogramos.
- e) Todas son falsas.

Respuesta correcta

Explicación

### 3.47. Pregunta test

En una distribución:  $P_{25} = 40$ ,  $P_{50} = 60$  y  $P_{75} = 70$ .

- a) La distribución es simétrica
- b) La distribución sugiere asimetría negativa
- c) La distribución sugiere asimetría positiva
- d) La distribución es leptocúrtica
- e) Las opciones a) y d) son ciertas

Respuesta correcta

Explicación

### 3.48. Pregunta test

En una distribución la mediana es 20 y la media es 26:

- a) Con seguridad hay asimetría negativa
- b) Con seguridad hay asimetría positiva
- c) Hay colas hacia la derecha y hacia la izquierda.
- d) Los datos son simétricos.
- e) Los datos sugieren una cola hacia la derecha. Habría que estudiarlo con más detalle

Respuesta correcta

Explicación

### 3.49. Pregunta test

El Rango Intercuartílico:

- a) Es sensible a los datos extremos.
- b) Es la distancia ente el primer y segundo cuartil.
- c) Es la raíz cuadrada de la varianza
- d) Sus unidades son el cuadrado de las variables.
- e) Mide el grado de dispersión de los datos, independientemente de su causa.

Respuesta correcta

- e) Explicación





## Capítulo 4

# Análisis Inferencial. Aplicaciones.

En este capítulo se resolverán problemas relativos a:

- Objetivos del estudio, hipótesis de trabajo e hipótesis estadísticas
- Importancia de las distribuciones de probabilidad en el trabajo práctico
- Estimación puntual y por intervalo
- Verificación de las hipótesis de trabajo: contraste de hipótesis

### 4.1. Pregunta test

En relación a las técnicas de inferencia estadística, elija la afirmación correcta:

- a) La media poblacional es una estimación puntual.
- b) La media muestral es un parámetro.
- c) Sólo se rechaza una hipótesis nula si esta es falsa.
- d) Un intervalo de confianza es una estimación confidencial de un parámetro.
- e) Todo lo anterior es falso

Respuesta correcta

Explicación

### 4.2. Pregunta test

Un estudio sobre la efectividad de un fármaco llega a la conclusión de que éste es mejor que el placebo con  $p < 0,05$  ¿Cuál es la interpretación correcta de este resultado?

- a) Con toda seguridad, el tratamiento es mejor que el placebo.
- b) La probabilidad de que el nuevo tratamiento sea mejor que el placebo es superior al 95 %.
- c) El tratamiento es un 95 % más efectivo que el placebo.
- d) La probabilidad de que el placebo sea mejor que el nuevo fármaco es menor de 5 %.
- e) Si el tratamiento no fuese efectivo, existe menos del 5 % de probabilidad de observar unas muestras tan contrarias a dicha hipótesis como las obtenidas.

Respuesta correcta

Explicación

### 4.3. Pregunta test

En la Distribución Normal:

- a) El intervalo  $\mu \pm \sigma$  abarca el 68 % del área total.
- b) El intervalo  $\mu \pm 1.96\sigma$  abarca el 95 % del área.
- c) El intervalo  $\mu \pm 2.6\sigma$  abarca el 99 % del área.
- d) El intervalo  $\mu \pm 2.6\sigma$  NO abarca el 5 % del área.
- e) Todas son ciertas.

Respuesta correcta

Explicación1

Explicación2

### 4.4. Problema

Sea  $Z$  una variable aleatoria normal estándar. Calcula:

- a) El área encerrada por la función de densidad entre  $z = 0$  y  $z = 1.35$
- b)  $P(Z \leq 2)$
- c)  $P(-0.5 \leq Z \leq 2.65)$
- d) El valor de  $z$  ( $z > 0$ ) de manera que el área encerrada entre 0 y  $z$  sea 0.2.
- e) El valor de  $z$  tal que la probabilidad de obtener un valor mayor que  $z$  sea 0.1.

**4.4.1. Solución**

Para resolver estas cuestiones, se puede hacer uso de la tabla de la Normal Tipificada, o programas como Excel© y la app Probability Distributions.

En este caso, vamos a utilizar el recurso Geogebra: Calculadora de Probabilidad. No tenemos más que introducir los datos y obtendremos las soluciones pedidas. Como el problema hace referencia a la distribución  $Z \equiv N(0,1)$ , usamos la distribución Normal, como parámetro  $\mu$  pondremos 0 y como parámetro  $\sigma$  pondremos 1.

**a)** Solución:  $P(0 \leq Z \leq 1.35) = 0.4115$

GeoGebra Probabilidad

Distribución  
Normal

Acumulada

Intervalos

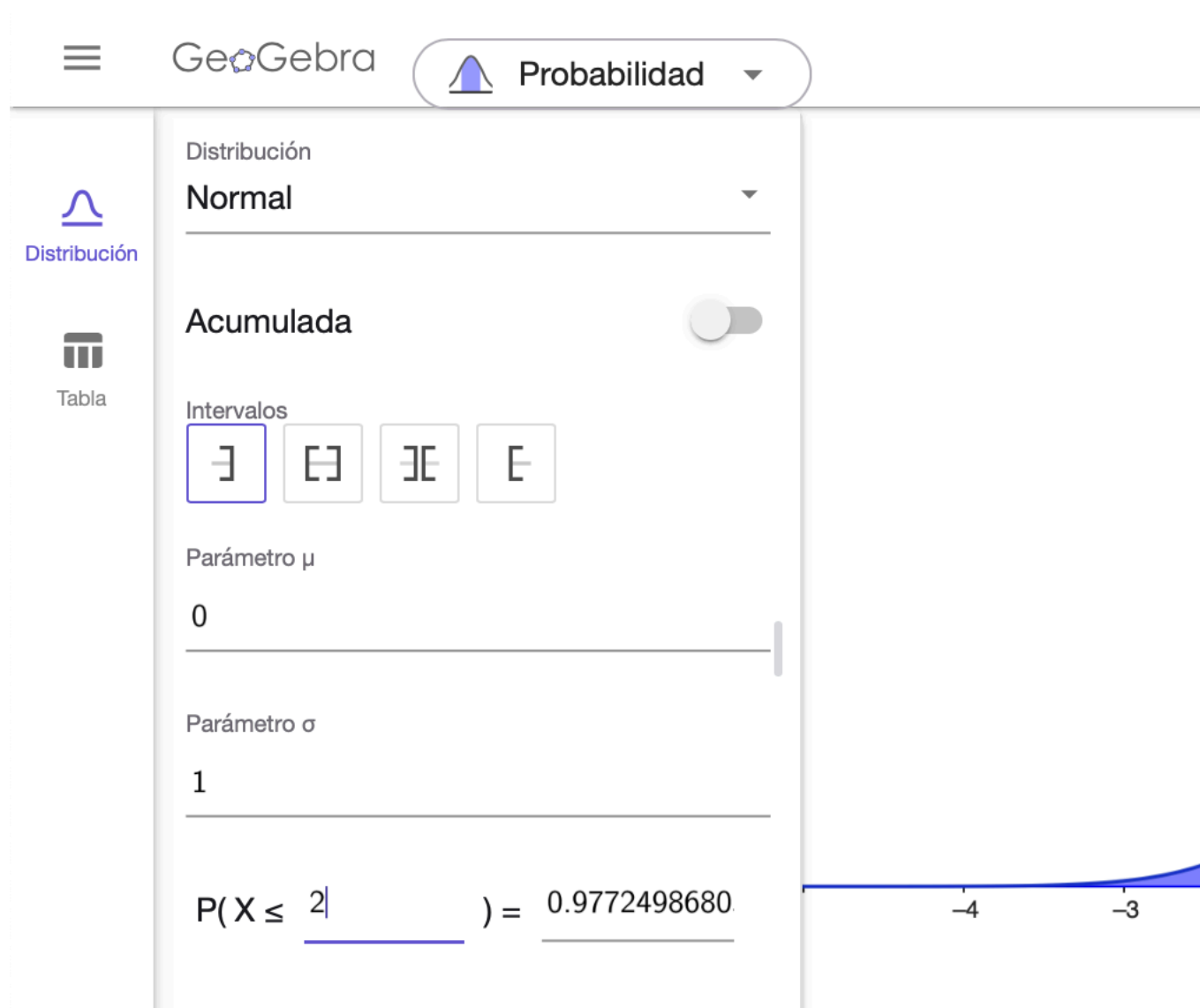
Parámetro  $\mu$   
0

Parámetro  $\sigma$   
1

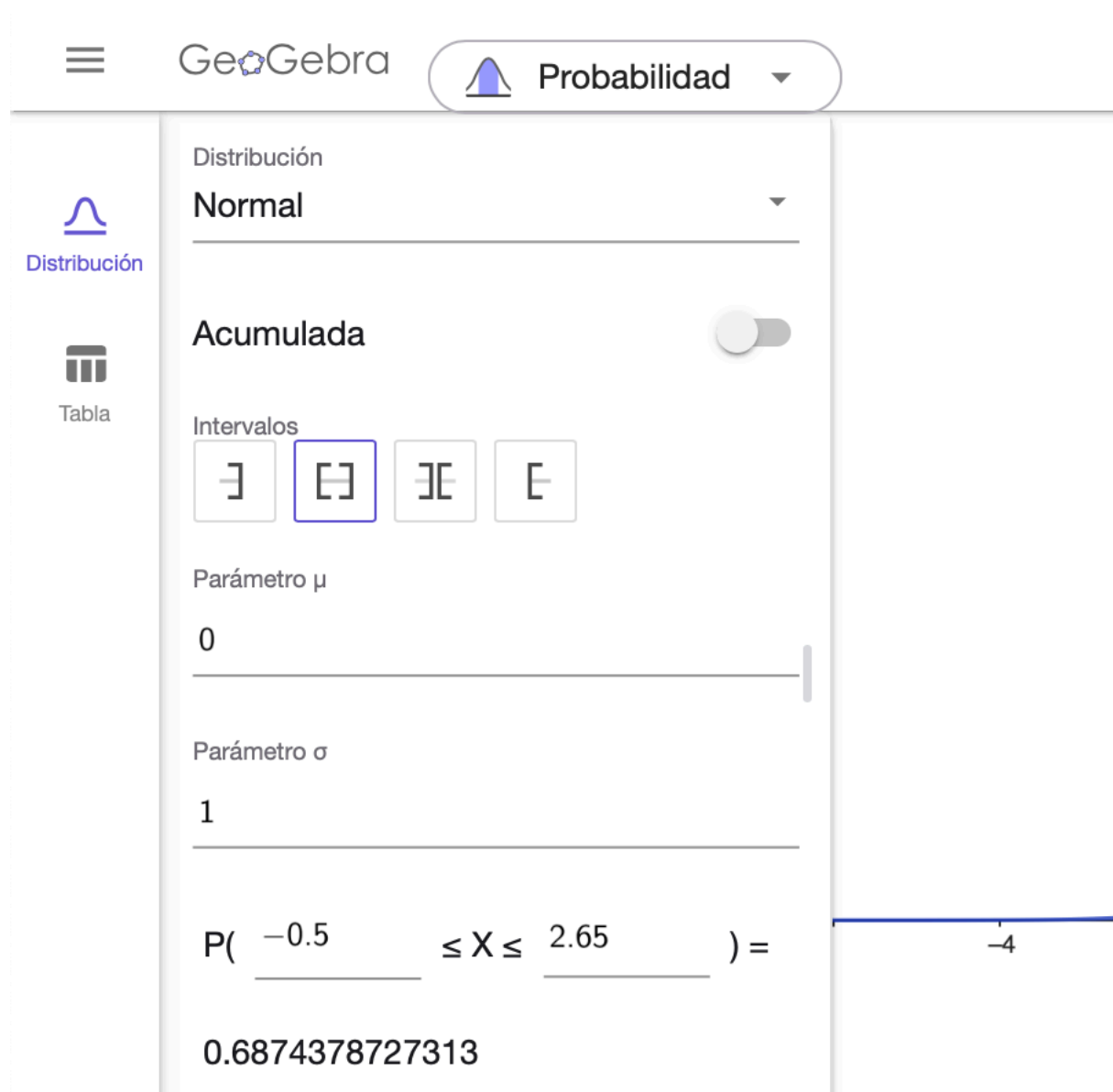
$P(0 \leq X \leq 1.35) =$

0.4114920085626

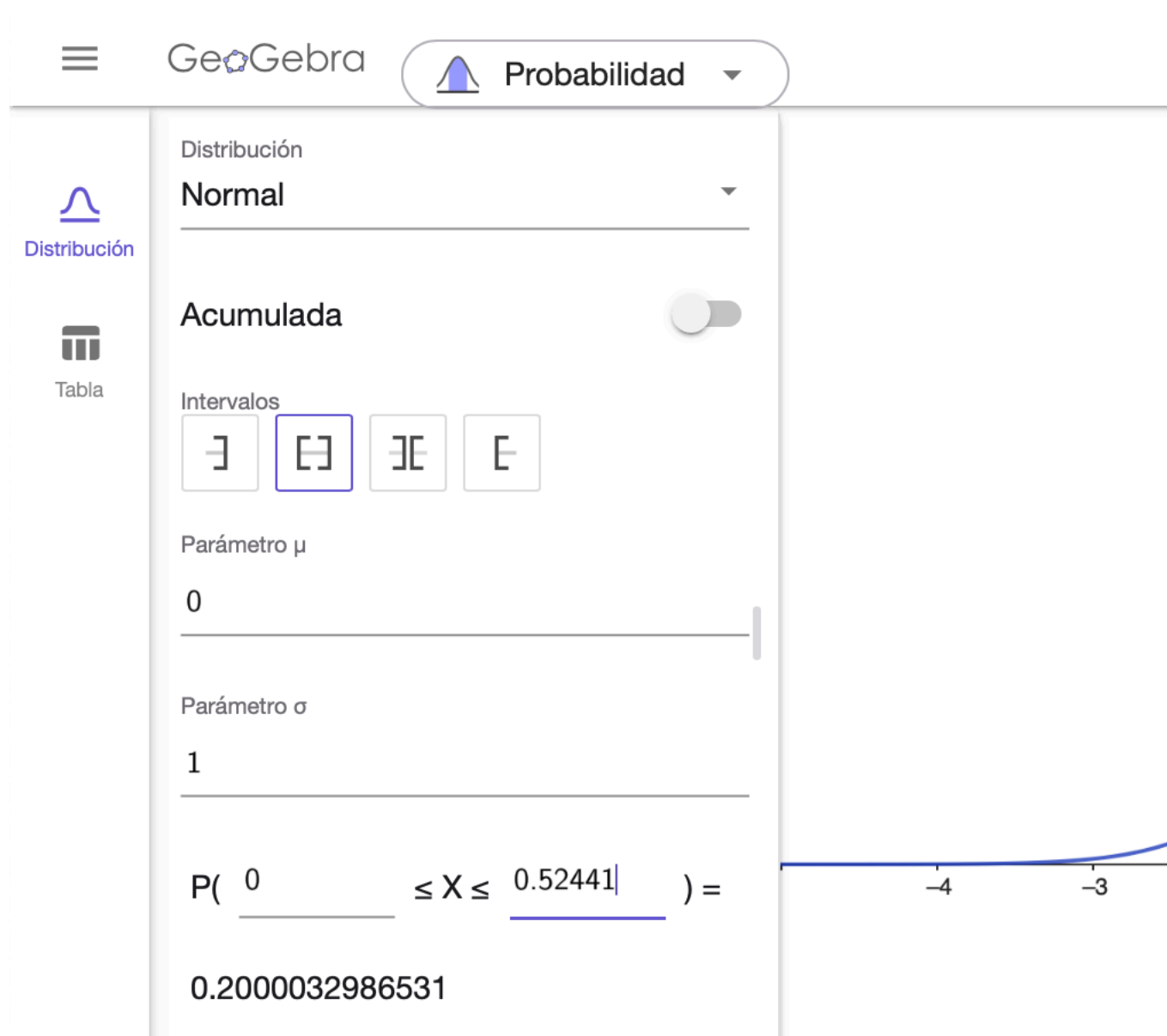
b) Solución:  $P(Z \leq 2) = 0.9772$



c) Solución:  $P(-0.5 \leq Z \leq 2.65) = 0.6874$



d) Solución:  $P(0 \leq Z \leq 0.5244) = 0.2$



En este caso, hemos ido aproximando a 4 decimales hasta obtener el valor más próximo a 0.2.

e) Solución:  $P(Z \geq 1.2816) = 0.1$

GeoGebra Probabilidad

Distribución

Normal

Acumulada

Intervalos

Parámetro  $\mu$

0

Parámetro  $\sigma$

1

$P(1.28155156 \leq X) = 0.1$

#### 4.5. Pregunta test

Una estimación confidencial para un nivel de confianza fijado, da por respuesta:

- a) Una aproximación de la media.
- b) Una aproximación de una proporción.
- c) Una probabilidad.
- d) Un intervalo.
- e) Un nivel de significación.



Respuesta correcta

Explicación

## 4.6. Pregunta test

En un contraste de hipótesis la cantidad  $p$  es:

- a) Un número pequeño.
- b) Fijada antes de realizar el contraste.
- c) La probabilidad de rechazar la hipótesis nula.
- d) La probabilidad de error al rechazar la hipótesis alternativa.
- e) Conocida al extraer la muestra y calcular el estadístico experimental.

Respuesta correcta

Explicación

## 4.7. Pregunta test

El perímetro torácico en un grupo de militares presenta distribución gaussiana con 95 cm de media y 5 cm de desviación típica. Elegimos a una muestra de 100 individuos y calculamos la media de la misma. Elija la afirmación correcta:

- a) La media de la muestra valdrá 95 cm.
- b) La media de la muestra sería un valor comprendido entre 90 y 100 cm con confianza del 68 %.
- c) La media de la muestra será un valor comprendido entre 95 y 100 cm con confianza del 95 %.
- d) La media de la muestra será un valor comprendido entre 94 y 96 cm con confianza del 95 %.
- e) Todo lo anterior es falso.

Respuesta correcta

### Explicación

Consideramos la variable aleatoria

$X$  = perímetro torácico en un grupo de militares  $X \equiv N(95, 5)$

El intervalo de confianza para la media muestral  $\bar{x}$  al 95 % será:

$$\mu \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Para una confianza del 95 %,  $z_{\alpha/2} = 1.96$  (ver vídeo) de donde:

$$\mu \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 95 = 95 \pm 1.95 \cdot \frac{5}{\sqrt{100}} = (94.02002, 95.97998) \approx (94, 96)$$

Se puede comprobar el resultado utilizando el applet Probability Distributions: Statistical Inference for  $\mu$  introduciendo los valores  $n = 100$ ,  $\bar{x} = 95$  y  $\sigma = 5$  con 95 % de CI (confidence Interval) y activando la opción *Show equations*.

## 4.8. Pregunta test

El consumo diario de Calorías se distribuye en una población de forma normal, con media 2500 y desviación típica 100. Si elijo una muestra de tamaño 100, entre qué valores espero encontrar su media (con una probabilidad del 95 % de acertar):

- a) Entre 2400 y 2600.
- b) Entre 2300 y 2700.
- c) Entre 2490 y 2510.
- d) Entre 2480 y 2520.
- e) Entre 2498 y 2502.

Respuesta correcta

Explicación

## 4.9. Problema

El nivel de colesterol en la sangre se mide de acuerdo a un índice llamado LDL. Para el caso de personas adultas, la distribución del colesterol en la sangre es aproximadamente normal y en el caso de los hombres tiene una media de 4.8 unidades LDL con una desviación estándar igual a 0.6 unidades. El nivel normal (o riesgo normal) de colesterol se considera aquel que queda entre los límites  $\mu \pm \sigma$  en unidades LDL. Una persona con más de  $\mu + \sigma$  pero menos de  $\mu + 2\sigma$  unidades LDL tiene un nivel de riesgo moderado. Si tiene un nivel de  $\mu + 2\sigma$  o superior se considera de alto riesgo y se hace propenso a sufrir un infarto. Por otra parte, si el nivel de colesterol en la sangre de un adulto está por debajo de  $\mu - \sigma$  unidades, se considera de riesgo bajo.

a) ¿Cuáles son los porcentajes de población de hombres adultos que están incluidos en cada uno de los 4 niveles de riesgo descritos?

b) ¿A partir de qué nivel de colesterol se encuentra el 10 % de la población de hombres adultos con mayor riesgo?

a) Los niveles son:

Riesgo	Nivel de LDL
Bajo	$< 4.2$
Normal	$4.2 - 5.4$

Riesgo	Nivel de LDL
Moderado	5.4 – 6
Alto	> 6

Utilizando la calculadora de Probabilidad de Geogebra obtenemos los porcentajes de poblacion:

Riesgo	Nivel de LDL	% población
Bajo	< 4.2	15.9 %
Normal	4.2 – 5.4	68.3 %
Moderado	5.4 – 6	13.6 %
Alto	> 6	2.2 %

GeoGebra

Probabilidad

Distribución

Normal

Acumulada

Intervalos

$\leq$   $<$   $\geq$   $>$

Parámetro  $\mu$

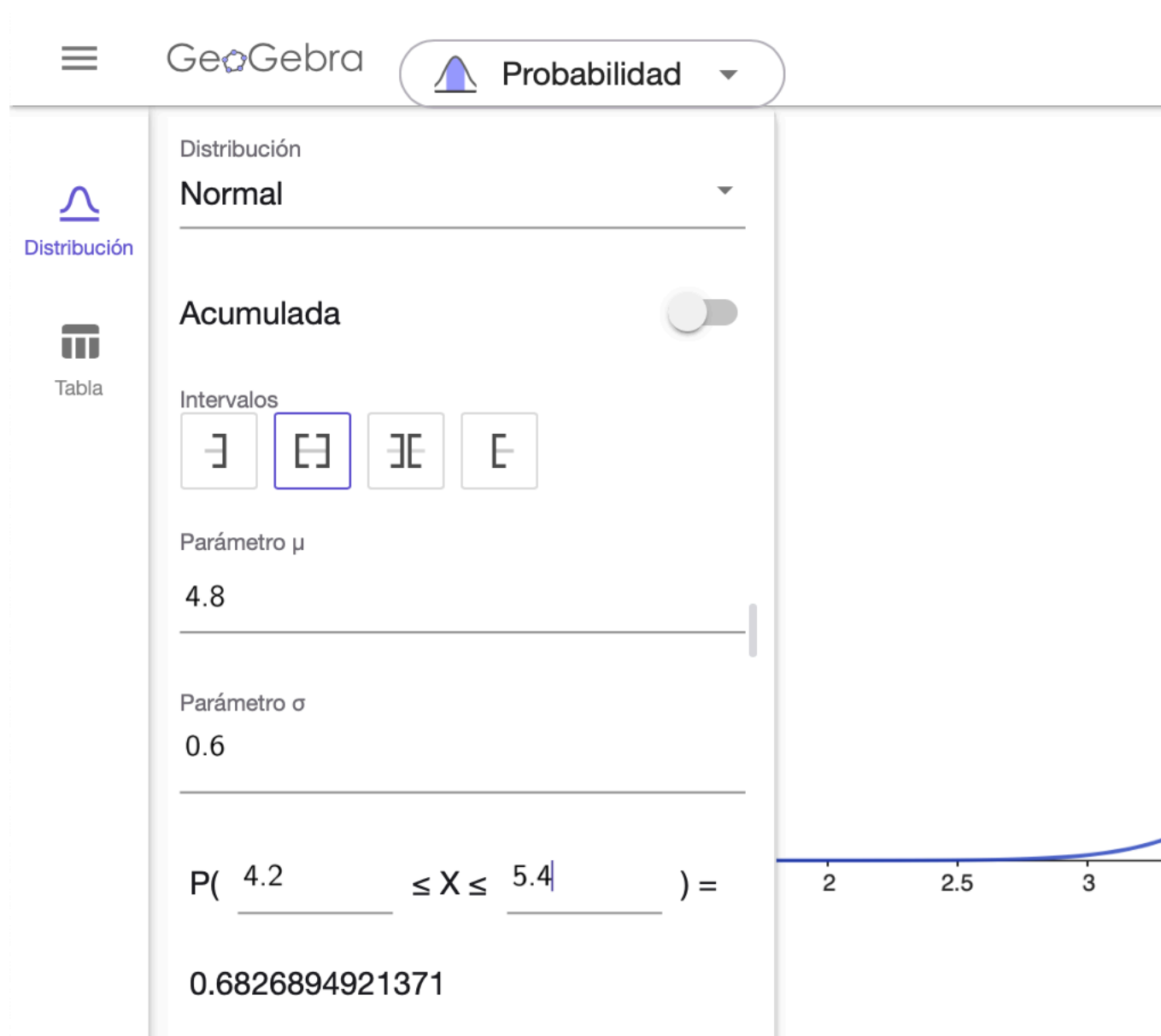
4.8

Parámetro  $\sigma$

0.6

$P(X \leq 4.2) = 0.1586552539$

2 2.5



GeoGebra

Probabilidad

Distribución

Normal

Acumulada

Intervalos

Parámetro  $\mu$

4.8

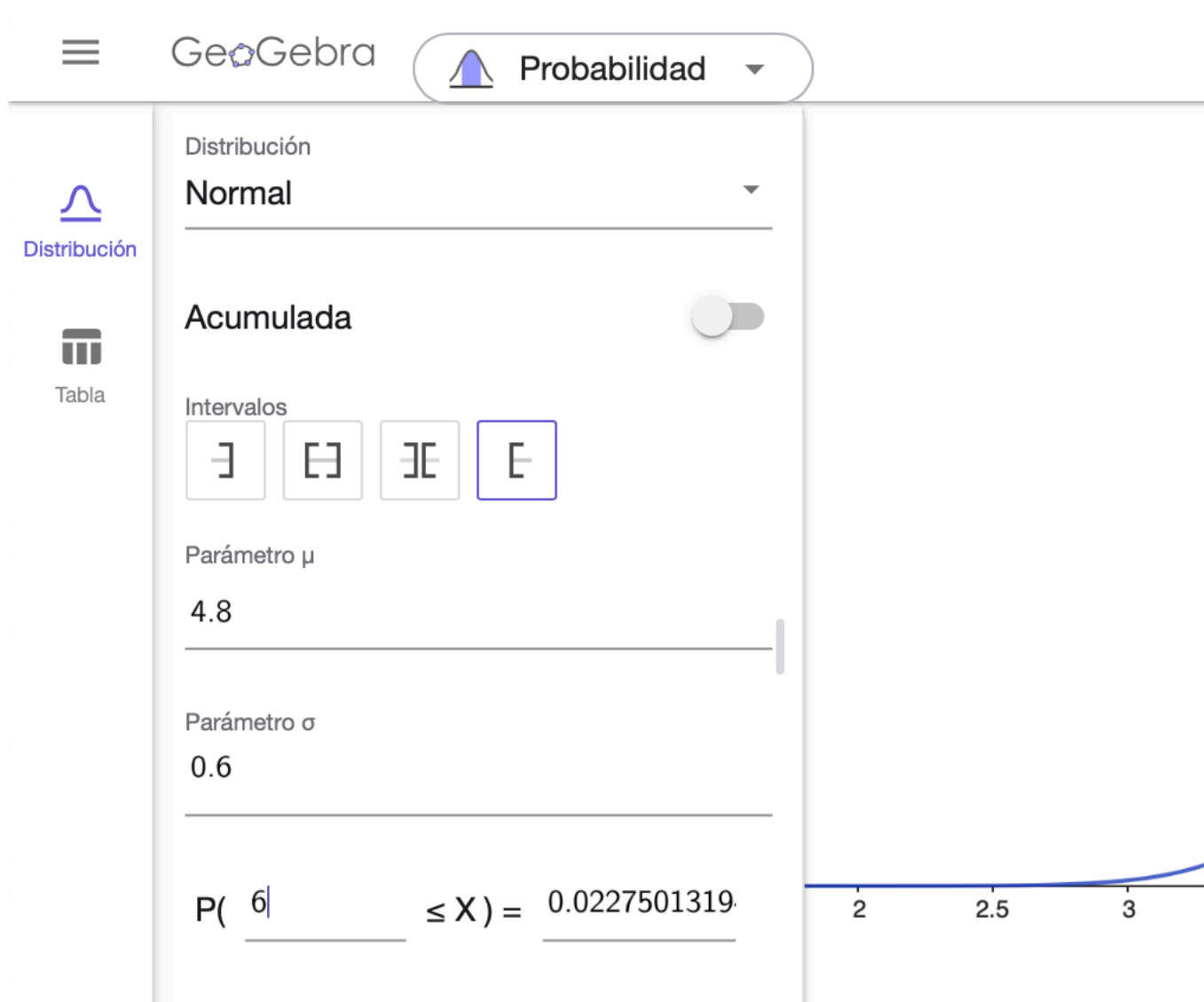
Parámetro  $\sigma$

0.6

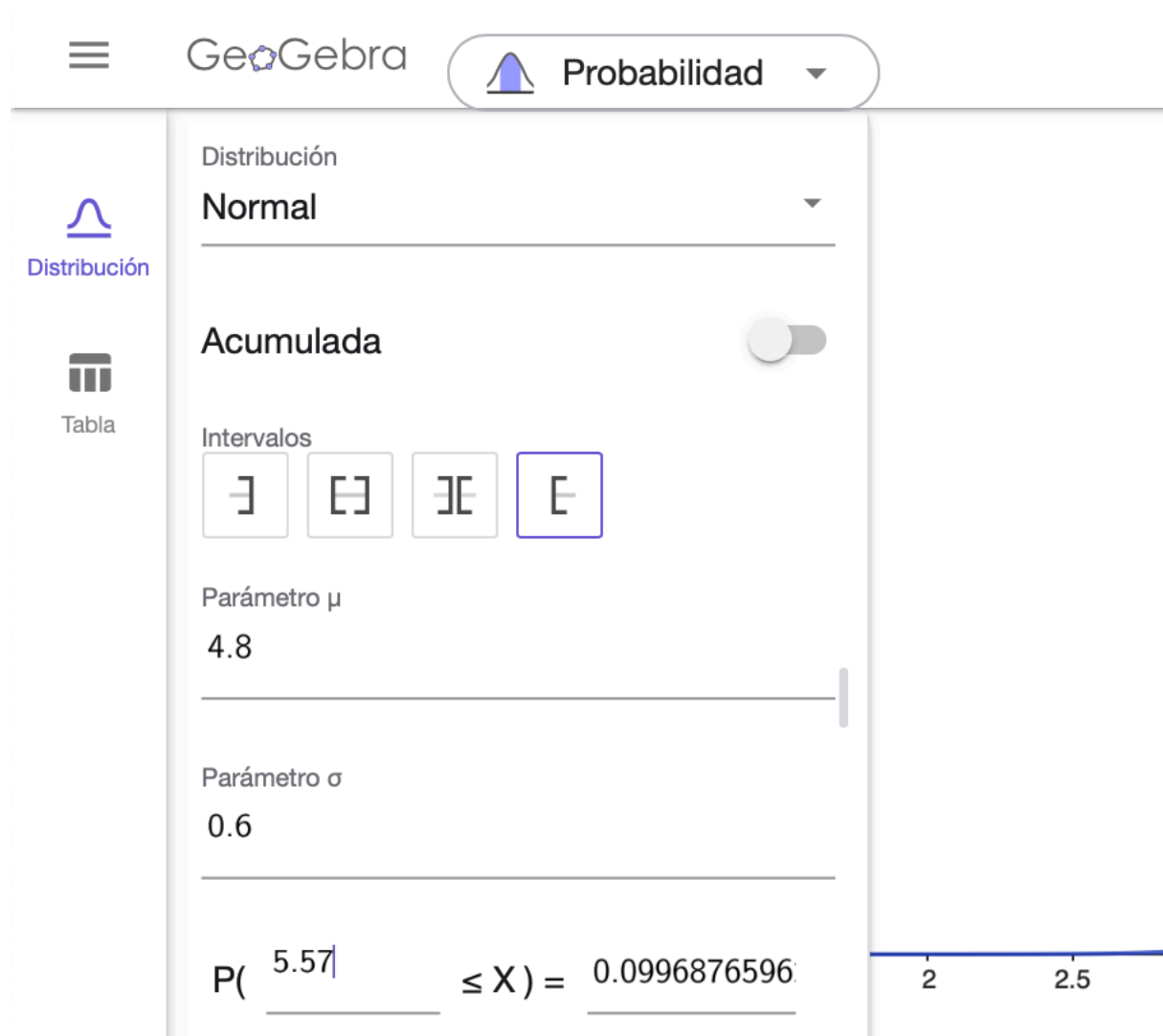
$P(5.4 \leq X \leq 6) =$

0.1359051219833

2 2.5



b) Siendo  $X$  = nivel de colesterol en la sangre, tenemos que calcular el valor de  $x$  para el que  $P(X > x) = 0.10$  (normal inversa). Con Geogebra, obtenemos que ese valor es 5.57.



#### 4.10. Pregunta test

Un contraste de hipótesis se considera significativo si:

- a) Una muestra aleatoria es coherente con la hipótesis nula.
- b) Una muestra aleatoria no es coherente con la hipótesis nula.
- c) La hipótesis alternativa es más probable que la nula.
- d) Todo lo anterior es cierto.
- e) Son ciertas (b) y (c).



Respuesta correcta

Explicación

### 4.11. Pregunta test

En una muestra aleatoria de 100 individuos se obtiene una media muestral de 50, la desviación típica es 20. Elija la afirmación correcta:

- a) El 68 % de los individuos de la muestra tiene sus valores comprendidos entre 48 y 52.
- b) El 95 % de los individuos de la muestra tiene sus valores comprendidos entre 46 y 54.
- c) Hay una probabilidad del 68 % de que la media de la población esté comprendida entre 30 y 70.
- d) Hay una probabilidad del 95 % de que la media de la población esté entre 46 y 54.
- e) Todo lo anterior es falso.

Respuesta correcta

Explicación

### 4.12. Pregunta test

Un contraste de hipótesis se considera no significativo si:

- a) Una muestra aleatoria es coherente con la hipótesis nula.
- b) Una muestra aleatoria no es coherente con la hipótesis nula.
- c) La hipótesis nula es más probable que la alternativa.
- d) Todo lo anterior es cierto.
- e) Son ciertas (a) y (c).

Respuesta correcta

Explicación

### 4.13. Pregunta test

Una muestra aleatoria de 64 pacientes refleja que el presión arterial diastólica media es 150 (DT 16), con distribución aproximadamente normal. Elija la afirmación correcta.

- a) La media de la población está con confianza del 95 % entre 134 y 166
- b) La media de la población está con confianza del 68 % entre 142 y 158

- c) La media de la población está con confianza del 95 % entre 148 y 152
- d) La media de la población está con confianza del 95 % entre 146 y 154
- e) El error típico es de 1 punto.

Respuesta correcta

Explicación

## 4.14. Problema

Para ayudar a la evaluación del pronóstico de pacientes con una determinada enfermedad pulmonar se calculan dos índices, independientes entre sí. Se asume que el primero de los índices se distribuye según una normal  $N(120, 10)$  y que el segundo se distribuye según una normal  $N(15, 3)$ . Se consideran susceptibles de una revisión más profunda aquellos pacientes que en el primer índice superen el valor 142. También son susceptibles de una revisión más profunda aquellos pacientes que en el segundo índice presenten un valor inferior a 8. ¿Qué porcentaje de pacientes son susceptibles de una revisión más profunda?

### 4.14.1. Solución

Definimos los sucesos y calculamos sus probabilidades con Probability Distributions: Normal Distribution:

- $A$  = valor superior a 142 en el primer índice  $P(A) = 0.0139$
- $B$  = valor inferior a 8 en el segundo índice  $P(B) = 0.00982$

De este modo, los pacientes susceptible de una revisión más profunda serán los del suceso  $A \cup B$ .

Sabemos que  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ . Así, nos falta saber el valor de  $P(A \cap B)$  (probabilidad de ser susceptible de una revisión más profunda por los 2 índices).

Nos dicen que los índices son independientes. Entonces

$P(A \cap B) = P(A) \cdot P(B) = 0.0139 \cdot 0.00982 = 0.000136498$ , probabilidad prácticamente nula.

Así,  $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.0139 + 0.00982 - 0.000136498 = 0.0235835$ .

Entonces, **el 2.36 % de pacientes son susceptibles de una revisión más profunda**

### 4.15. Pregunta test

Qué propiedad o propiedades caracterizan a una distribución normal tipificada frente a una distribución normal cualquiera:

- a) El área bajo su función de densidad es igual a 1.
- b) Su media es 1 y su desviación típica es 0.
- c) Su rango de valores oscila entre 0 y 3.
- d) Su media es 0 y su desviación típica es 1.
- e) Son ciertas (c) y (d)

Respuesta correcta

Explicación

### 4.16. Pregunta test

En relación con los contrastes de hipótesis, elija la afirmación correcta:

- a) La hipótesis nula es la correcta.
- b) La hipótesis nula es la falsa.
- c) Si la hipótesis alternativa es cierta, seguro que se rechaza la nula.
- d) El contraste es significativo cuando los datos muestrales no son los esperados si la hipótesis nula fuese cierta,
- e) Si es más probable que sea cierta la hipótesis alternativa que la nula, el contraste es significativo.

Respuesta correcta

Explicación

### 4.17. Pregunta test

En una población, el peso tiene media 60 kg y desviación típica 6 Kg. La altura tiene de media 170 cm y desviación 6 cm. Cierta individuo tiene un peso de 70 Kg y altura 180 cm.

- a) La altura tiene un valor más extremo que el peso.
- b) El peso es menos extremo que la altura.
- c) Peso y altura son valores igualmente extremos.
- d) El peso es más extremo que la altura.
- e) La altura es menos extrema que el peso.

Respuesta correcta

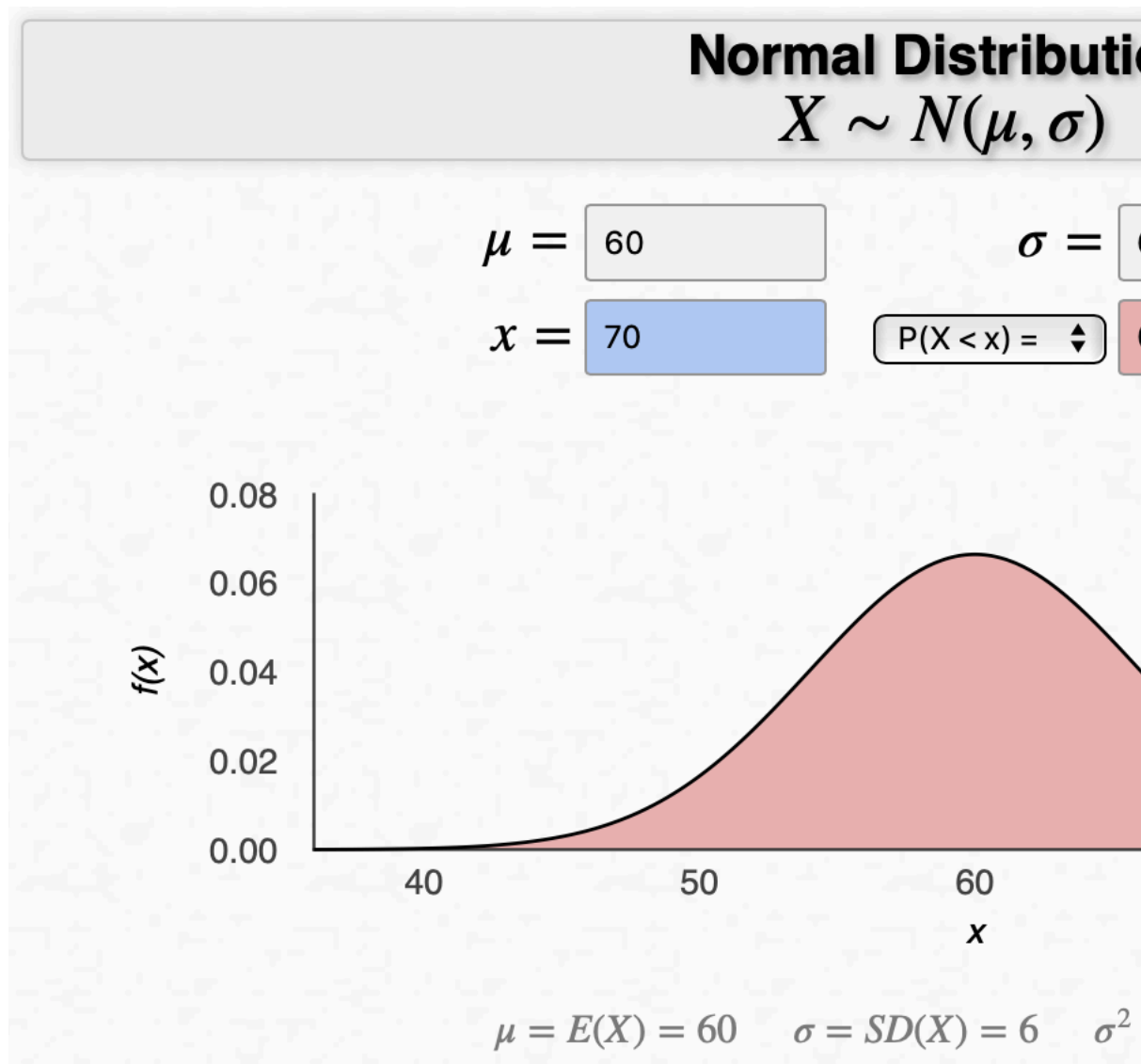
**Explicación**

En este caso debemos comparar valores en 2 variables normales (o gaussianas):

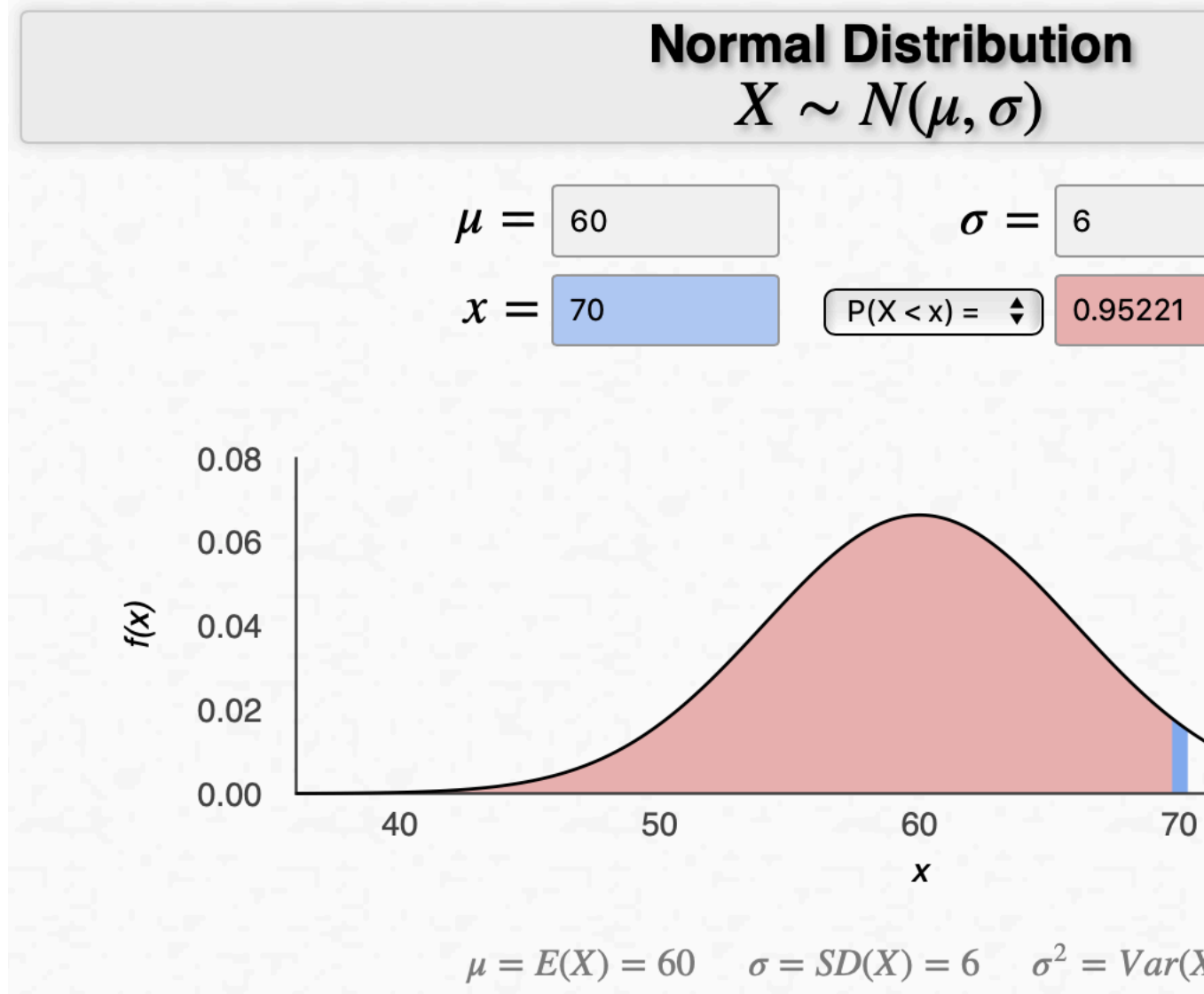
- Peso  $\equiv N(60, 6)$  para un valor de 70 kg.
- Altura  $\equiv N(170, 6)$  para un valor de 180 cm.

Podemos hacerlo con una tabla (ver vídeo) o usando Excel® o una app específica. (ver vídeo).

Utilizando la aplicación Probability Distributions: Normal Distribution, obtenemos la probabilidad de pesar menos de 70 kg:



y la probabilidad de medir menos de 180 cm:



En ambos casos, la probabilidad es 0.95221

#### 4.18. Pregunta test

El nivel medio de glucemia en una población tiene un comportamiento gausiano con media 150mg/dl, y un coeficiente de variación del 10%. Entre qué valores

se situa el 95 % de los individuos de la población.

- a) Entre 140 y 160.
- b) Entre 130 y 170.
- c) Entre 120 y 180.
- d) Entre 110 y 190.
- e) Entre 100 y 200.

Respuesta correcta

### Explicación

En este caso, nos dan el coeficiente de variación,  $CV = \frac{\sigma}{\bar{x}}$ . Entonces:

$$0.10 = \frac{\sigma}{150} \Rightarrow \sigma = 15$$

Así la variable  $X = \text{nivel medio de glucemia} \equiv N(150, 15)$ .

En una distribución normal, el 95 % de los datos están en el intervalo  $\mu \pm 2\sigma = 150 \pm 30 = (120, 180)$

## 4.19. Pregunta test

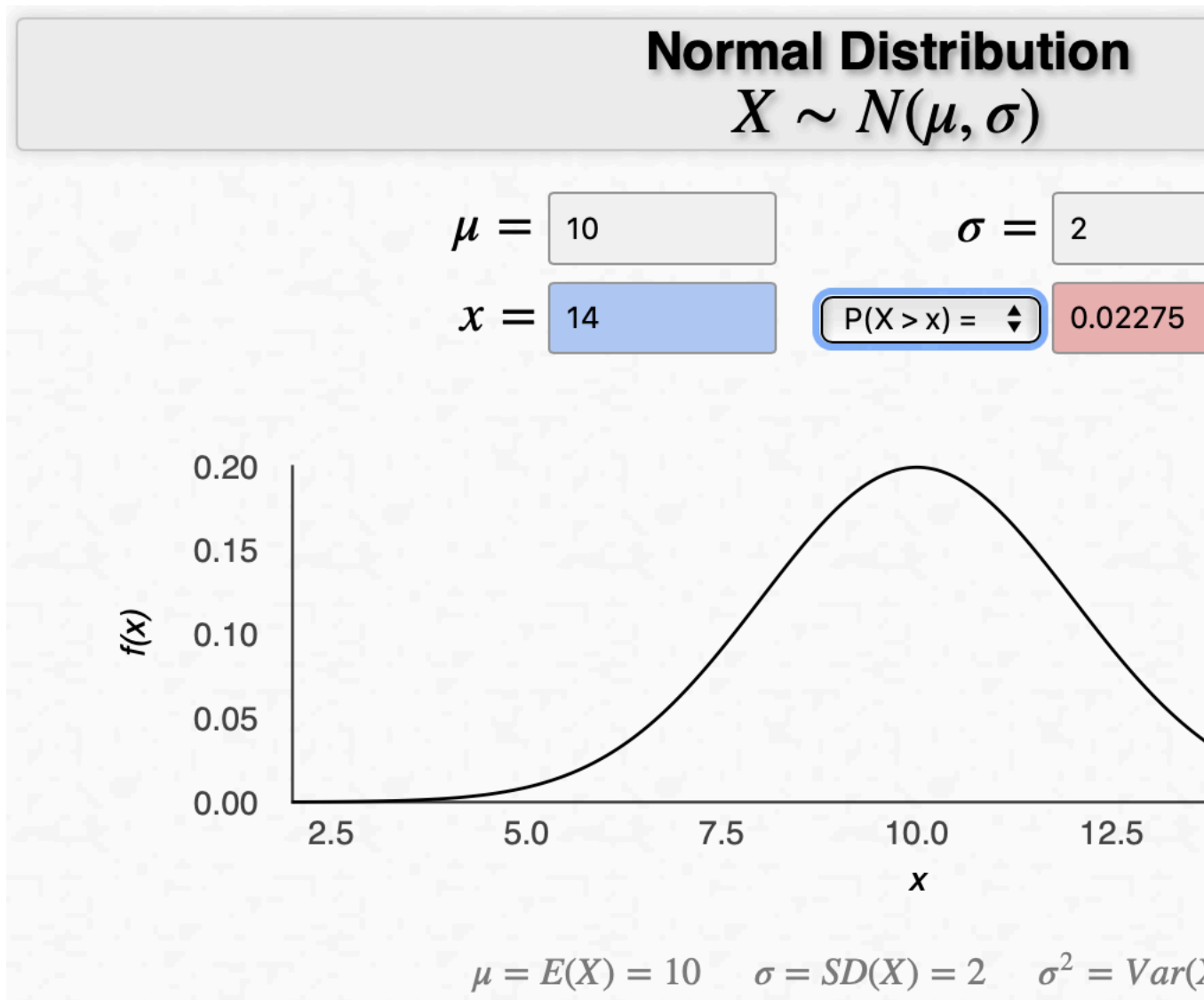
La concentración de calcio se comporta en los mamíferos como una distribución normal de media 10 y desviación típica 2. ¿Con qué frecuencia se encuentran mamíferos con una concentración superior a 14?

- a) 95 %
- b) 68 %
- c) 50 %
- d) 5 %
- e) 2,5 %

Respuesta correcta

### Explicación

Con Probability Distributions: Normal Distribution tenemos:



#### 4.20. Pregunta test

Se realiza un estudio para saber si dos tratamientos de quimioterapia presentan diferencias en cuanto a la supervivencia de los pacientes. No se encontró diferencia estadísticamente significativa. ¿Cuál de las siguientes razones podrían ser causantes del resultado?

- a) Los tratamientos ofrecen tiempos de supervivencia muy diferentes.

- b) El nivel de significación es demasiado alto.
- c) Las muestras son demasiado numerosas.
- d) Las muestras son demasiado pequeñas.
- e) Nada de lo anterior.

Respuesta correcta

Explicación

#### 4.21. Pregunta test

Se realiza un experimento donde nos basaremos en un contraste de hipótesis para tomar una decisión con un nivel de significación del 1 %. De las siguientes cuál no es un resultado posible de un contraste de hipótesis:

- a) El experimento no es concluyente.
- b) El experimento permite obtener conclusiones.
- c) Se rechaza la hipótesis nula.
- d) Se rechaza la hipótesis alternativa.
- e) Se acepta la hipótesis alternativa.

Respuesta correcta

Explicación

#### 4.22. Pregunta test

El IMC se distribuye en una población de forma normal. El 95 % central de los individuos tiene un IMC comprendido entre 20 y 24. Entonces:

- a) La media es 22.
- b) La desviación típica es 1.
- c) La curtosis es cero.
- d) Todas las anteriores son correctas.
- e) Sólo dos de las anteriores son correctas.

Respuesta correcta

Explicación

#### 4.23. Pregunta test

Elija la afirmación **falsa**:

- a) El nivel de significación es normalmente un valor pequeño.



- b) La significación de un contraste es conocida tras analizar los datos.
- c) El nivel de significación de un contraste debe ser fijado antes de analizar los datos.
- d) Un contraste debe ser declarado significativo antes de recoger los datos.
- e) Un contraste es declarado significativo si se obtiene una muestra que discrepa mucho de la hipótesis nula.

Respuesta correcta

Explicación

## 4.24. Problema

El tiempo que los empleados de una oficina tardan en completar una tarea específica sigue una distribución normal con una media de 45 minutos y una desviación típica de 5 minutos. ¿Cuál es el tiempo que separa aproximadamente al 16 % más rápido de los empleados del resto?

### 4.24.1. Explicación

Este es un problema de distribución normal inversa Ver vídeo.

Si utilizamos la aplicación Probability Distributions: Normal Distribution no tenemos más que introducir los valores  $\mu = 45$ ,  $\sigma = 5$  y  $P(X > x) = 0.16$  para obtener un valor de  $x = 49.97229 \approx 50$ .

Entonces, **l tiempo que separa aproximadamente al 16 % más rápido de los empleados del resto es de 50 minutos.**

## 4.25. Pregunta test

Los ingresos de un grupo de personas siguen una distribución normal con una media de 50,000 € y una desviación típica de 10,000 €. ¿Cuál es el rango aproximado de ingresos que abarca al 95 % de las personas en este grupo?

- a Entre 30,000€ y 70,000€ b Entre 20,000€ y 80,000€ c Entre 40,000€ y 60,000€  
d Entre 10,000€ y 90,000€ e Entre 50,000€ y 100,000€

Respuesta correcta

Explicación

### 4.26. Pregunta test

La cantidad de tiempo que los estudiantes tardan en completar un examen sigue una distribución normal con una media de 120 minutos y una desviación típica de 20 minutos. ¿Qué porcentaje de estudiantes termina el examen en menos de 100 minutos?

- a) 2.5 %
- b) 5 %
- c) 16 %
- d) 34 %
- e) 50 %

Respuesta correcta

Explicación

### 4.27. Pregunta test

Solamente una de las siguientes frases podría alguna vez encontrarse como conclusión de un estudio científico. ¿Cuál es?

- a) El tratamiento produjo un efecto significativamente mayor que el placebo ( $p=0.75$ )
- b) El tratamiento produjo un efecto significativamente menor que el placebo ( $p=0.25$ )
- c) El tratamiento no produjo un resultado diferente al placebo ( $p<0.001$ )
- d) El tratamiento produjo un efecto significativamente menor que el placebo ( $p=0.99$ )
- e) Se apreciaban diferencias significativas entre el placebo y el tratamiento ( $p<0.001$ )

Respuesta correcta

Explicación

### 4.28. Pregunta test

En un estudio que investiga la eficacia de una nueva terapia de rehabilitación para pacientes que han sufrido un accidente cerebrovascular, solo una de las siguientes afirmaciones cumple con los criterios de significación estadística. ¿Cuál es?

- a) La nueva terapia no mostró diferencias significativas en la recuperación en comparación con la terapia estándar ( $p=0.90$ )

- b) La nueva terapia fue significativamente menos efectiva que la terapia estándar ( $p=0.14$ )
- c) La nueva terapia fue significativamente más efectiva que la terapia estándar ( $p<0.001$ )
- d) La nueva terapia fue igual de efectiva que la terapia estándar ( $p=0.51$ )
- e) La terapia estándar fue significativamente más efectiva que la nueva terapia ( $p=0.68$ )

Respuesta correcta

Explicación

## 4.29. Pregunta test

La duración de las llamadas en un centro de atención al cliente sigue una distribución normal con una media de 8 minutos y una desviación típica de 2 minutos. ¿Cuál es el tiempo que separa aproximadamente al 2.5 % de las llamadas más cortas del resto?

- a) 2 minutos
- b) 4 minutos
- c) 6 minutos
- d) 10 minutos
- e) 12 minutos

Respuesta correcta

Explicación

## 4.30. Pregunta test

En un estudio que evalúa la efectividad de la terapia cognitivo-conductual (TCC) en comparación con la terapia de apoyo en el tratamiento de la ansiedad, solo una de las siguientes afirmaciones es correcta según los criterios de significación estadística. ¿Cuál es?

- a) La TCC no mostró diferencias significativas en la reducción de la ansiedad en comparación con la terapia de apoyo ( $p=0.15$ )
- b) La TCC fue significativamente más efectiva que la terapia de apoyo en la reducción de la ansiedad ( $p<0.001$ )
- c) La terapia de apoyo fue significativamente más efectiva que la TCC en la reducción de la ansiedad ( $p=0.91$ )
- d) Ambas terapias fueron igualmente efectivas en la reducción de la ansiedad ( $p=0.06$ )

- e) La TCC fue significativamente menos efectiva que la terapia de apoyo en la reducción de la ansiedad ( $p=0.56$ )

Respuesta correcta

### 4.31. Pregunta test

En una población de estudiantes, las calificaciones en matemáticas tienen una media de 75 y una desviación típica de 10, mientras que las calificaciones en ciencias tienen una media de 80 y una desviación típica de 5. Un estudiante obtiene una calificación de 90 en matemáticas y 85 en ciencias. ¿Cuál de las siguientes afirmaciones es correcta?

- a) La calificación de matemáticas es más extrema que la de ciencias.
- b) La calificación de ciencias es más extrema que la de matemáticas.
- c) Las calificaciones de matemáticas y ciencias son igualmente extremas.
- d) La calificación de matemáticas es menos extrema que la de ciencias.
- e) No se pueden comparar ambas calificaciones.

Respuesta correcta

Explicación

## Capítulo 5

# Regresión y correlación.

En este capítulo se resolverán problemas relativos a:

- Introducción a la regresión y correlación
- Estudio de la representatividad de la recta de regresión
- Otros modelos de regresión
- Correlación

### 5.1. Pregunta test

Si al calcular el coeficiente de correlación de dos variables X e Y, se tiene  $r = -0.20$  ocurre que:

- a) La pendiente de la recta de regresión es pequeña.
- b) La pendiente de la recta de regresión es grande.
- c) X e Y están poco relacionadas, aunque cuando X decrece, Y tiene tendencia a crecer.
- d) El modelo lineal de regresión explica el 20 % de la varianza de una variable cualquiera en función de la otra.
- e) El modelo lineal de regresión explica el 80 % de la varianza de una variable cualquiera en función de la otra.

Respuesta correcta

Explicación

## 5.2. Pregunta test

La recta de regresión de  $Y$  sobre  $X$  se muestra como un buen modelo para explicar la relación entre dos variables numéricas. Entonces:

- a)  $Y$  se puede calcular exactamente como una función matemática de  $X$ .
- b)  $Y$  es independiente de  $X$ .
- c) La covarianza de  $X$  e  $Y$  no es nula.
- d) La media de  $X$  coincide con la media de  $Y$ .
- e) Sólo dos de las afirmaciones anteriores son correctas.

Respuesta correcta

Explicación

## 5.3. Pregunta test

En una población se obtiene con una bondad de ajuste de 0,9 que la relación entre nivel de glucemia ( $Y$ ) y nivel de colesterol ( $X$ ) es de  $Y = 20 + \frac{X}{4}$ . Entonces:

- a) Todos los individuos con un valor de colesterol 100, presentan glucemia 45.
- b) Existe tendencia a que a mayor nivel de glucemia, mayor nivel de colesterol.
- c) Hay mas individuos con colesterol alto que con glucemia baja.
- d) Las observaciones se muestran como una nube de puntos creciente.
- e) Sólo dos de las afirmaciones anteriores son correctas.

Respuesta correcta

### Explicación

Tenemos que  $r = 0.9$  lo que indica una correlacion lineal positiva y fuerte entre las 2 variables.

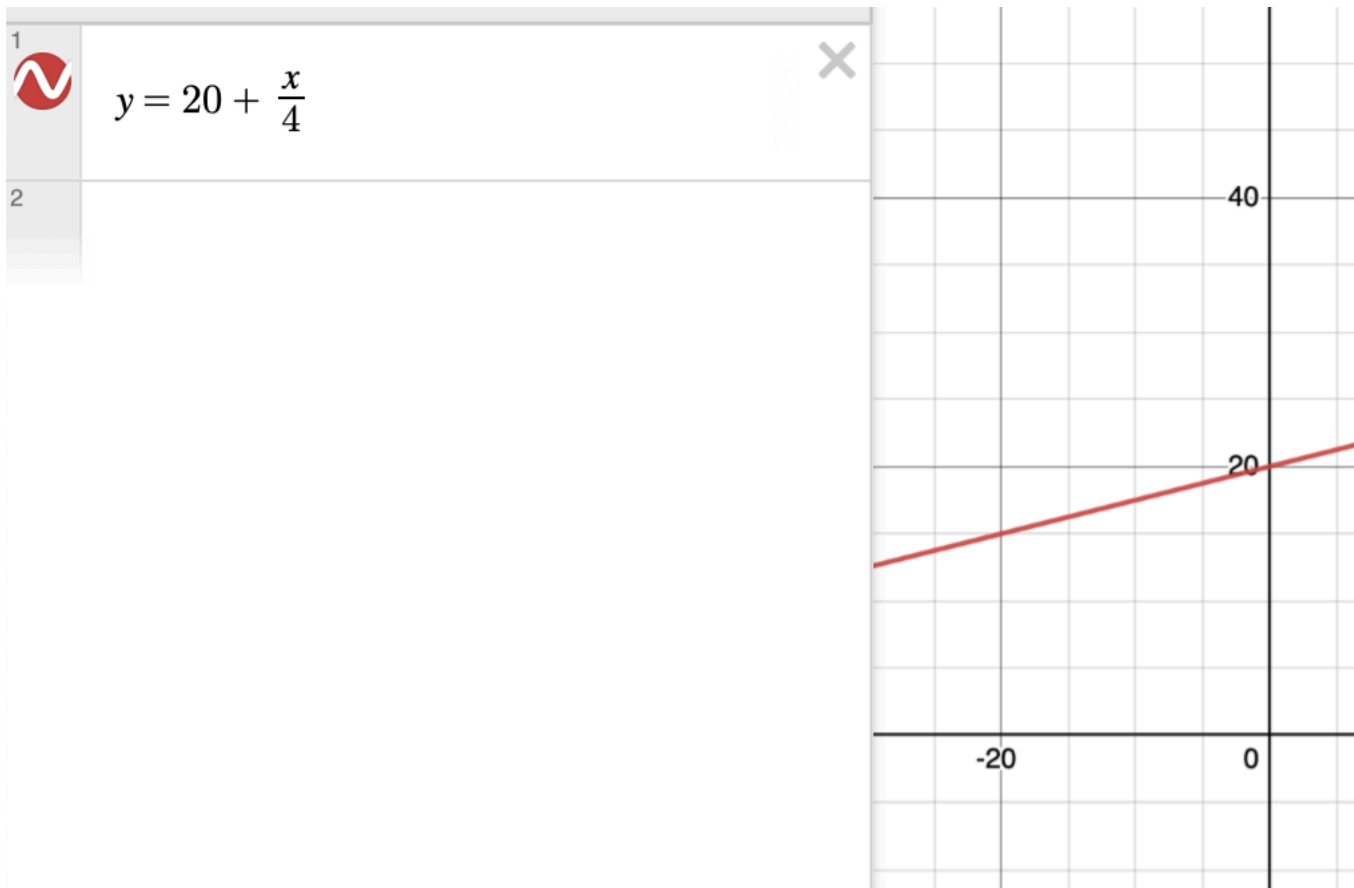
Además, tenemos la ecuación de la recta de regresión de  $Y$  (nivel de glucemia) sobre  $X$  (nivel de colesterol),  $Y = 20 + \frac{X}{4}$

Veamos cada apartado:

- a) Efectivamente, para un valor de colesterol 100, como  $Y = 20 + \frac{100}{4} = 20 + 25 = 45$ , se estima que el valor de glucemia es de 45. Pero eso no implica que *todos los individuos* con un valor de colesterol 100, presentan glucemia 45.
- b) Al haber correlación lineal positiva y fuerte ( $r^2 = 0.9 \Rightarrow r \approx 0.95$ ), se cumple que existe tendencia a que a mayor nivel de glucemia, mayor nivel de colesterol.

c) Con los datos del problema, no podemos saber si hay mas individuos con colesterol alto que con glucemia baja.

d) Como la nube de puntos se ajusta bastante bien a la recta de regresión cuya gráfica (obtenida con Desmos) es:



se cumple que las observaciones se muestran como una nube de puntos creciente.

## 5.4. Pregunta test

Dos variables numéricas son incorreladas. Entonces:

- a)  $r = 0$
- b) El modelo lineal de regresión sólo propone un valor como predicción de  $Y$ .
- c) La nube de puntos no presenta aspecto creciente.
- d) La varianza residual en el modelo de regresión de  $Y$  sobre  $X$  es igual a la varianza de  $Y$ .

e) Todo lo anterior es cierto.

Respuesta correcta

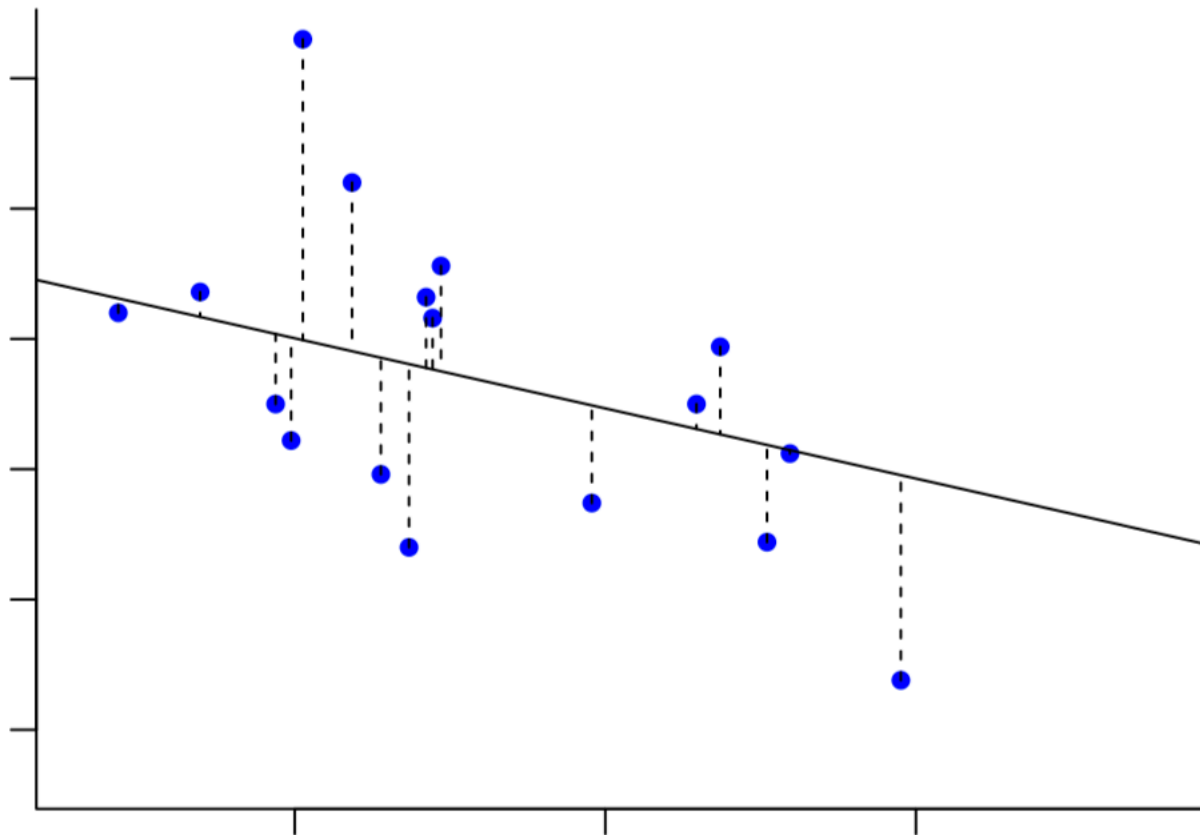
### Explicación

a) Si no existe correlación entre las variables (incorreladas),  $r = 0$ . Al mínimo de correlación que haya,  $r \neq 0$ , tanto positiva como negativa.

b) Siempre se cumple que el modelo lineal de regresión sólo propone un valor como predicción de  $Y$ . Al obtener la ecuación de la recta de regresión (modelo), para un valor de  $X$  únicamente tenemos un valor de  $Y$ .

c) Al no haber correlación, la nube de puntos no presenta tendencia alguna por lo que la nube de puntos no presenta aspecto creciente (ni decreciente).

d) Recordamos que, modelo de regresión de  $Y$  sobre  $X$ , los residuos son las direfencias (en vertical) de los puntos a la recta de regresión, es decir las medidas del los segmentos de la imagen:

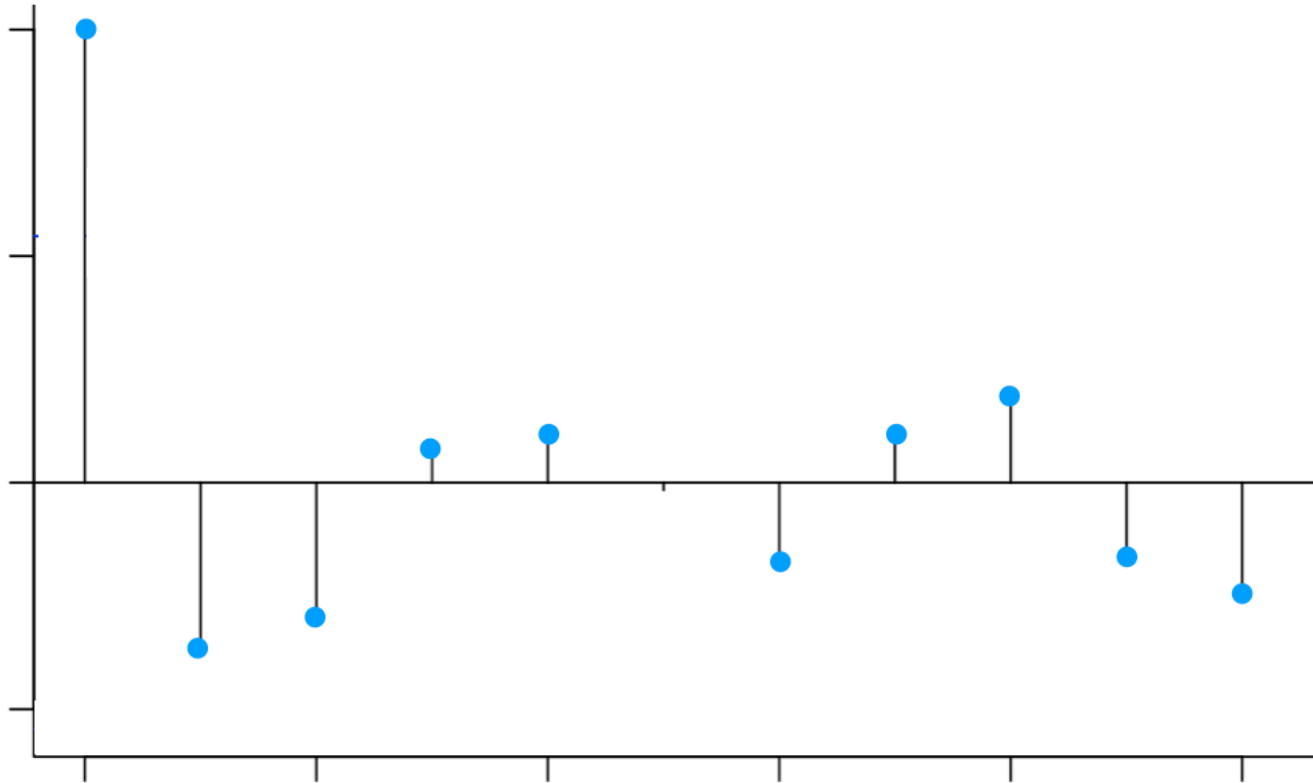


Así, la varianza residual en el modelo de regresión de  $Y$  sobre  $X$  es la media



de los cuadrados de esos valores (residuos). Si la recta es oblicua, esos residuos dependen de esa inclinación, esto es, dependen de  $X$  e  $Y$ .

En nuestro caso, las variables son incorreladas por lo que la recta de regresión del modelo de regresión de  $Y$  sobre  $X$  es horizontal:



De este modo, los residuos dependen únicamente de  $Y$ . Al pasar la recta de regresión por el punto  $(\bar{x}, \bar{y})$ , la media de los cuadrados de los residuos coincide con la varianza de  $Y$  por lo que, cuando las variables son incorreladas, la varianza residual en el modelo de regresión de  $Y$  sobre  $X$  es igual a la varianza de  $Y$ .

## 5.5. Pregunta test

De las siguientes parejas de variables, en cuáles crees que puede ser útil un análisis de regresión lineal:

- a) La presión sanguínea y el grupo sanguíneo.
- b) El nivel de colesterol y la concentración de bilirrubina.
- c) El grupos sanguíneo y el factor Rh.

- d) El género y la edad.
- e) Poseer ideología racista y el factor RH.

Respuesta correcta

Explicación

### 5.6. Pregunta test

Si el coeficiente de correlación lineal de Pearson entre dos variables es  $-0,8$  podemos decir:

- a) La covarianza es negativa.
- b) La relación entre las variables es directa.
- c) Hay poca relación lineal entre las variables.
- d) Hay un error de cálculo.
- e) El 80 % de las predicciones son correctas.

Respuesta correcta

Explicación

### 5.7. Pregunta test

En un estudio de regresión lineal, donde el peso se estudie conjuntamente con otras variables, en qué casos lo usarías como variable dependiente:

- a) Al estudiarlo con la altura.
- b) Al estudiarlo con el nivel del colesterol.
- c) Al estudiarlo con la presión sanguínea.
- d) Al estudiarlo con el grupo sanguíneo.
- e) Nada de lo anterior.

Respuesta correcta

Explicación

### 5.8. Pregunta test

En una población formada por unidades familiares, la altura media del padre en la familia se comporta como una distribución normal de media 170 cm con desviación típica 5 cm. La altura del primer hijo varón es otra variable con distribución similar. Con estos datos podemos afirmar:

- a) No hay relación entre ambas variables.

- b) Hay relación inversa entre las variables.
- c) No debemos intentar predecir la altura del hijo de un padre que mide 140 cm.
- d) Hay relación directa entre las variables.
- e) Nada de lo anterior.

Respuesta correcta

Explicación

## 5.9. Pregunta test

Se observa que al aumentar el consumo de estanol, disminuye el nivel de colesterol en sangre. Se utiliza un modelo de regresión lineal donde el nivel de colesterol es la variable independiente y el consumo de estanol es la dependiente. Se calcula una bondad de ajuste para el modelo del 25 %. Entonces:

- a) El 25 % de las predicciones del modelo son correctas.
- b)  $r = 0.5$
- c)  $r = 0.25$
- d)  $r = -0.25$
- e)  $r = -0.5$

Respuesta correcta

Explicación

## 5.10. Pregunta test

Si el coeficiente de correlación lineal de Pearson entre dos variables es -0,1 podemos decir:

- a) La covarianza es pequeña.
- b) Hay fuerte relación inversa entre las variables.
- c) Hay poca relación lineal entre las variables.
- d) Hay un error de cálculo.
- e) El 10 % de las predicciones son correctas.

Respuesta correcta

Explicación

### 5.11. Pregunta test

Se estudia la asociación lineal entre dos variables numéricas. El coeficiente de determinación vale 0,95.

- a) Hay poca asociación.
- b) Hay asociación directa.
- c) Hay asociación inversa.
- d) Hay una buena asociación
- e) Nada de lo anterior.

Respuesta correcta

Explicacion

### 5.12. Pregunta test

Se observa que al disminuir el consumo de comida rápida, disminuye el nivel de colesterol en sangre. Se usa un modelo de regresión entre ambas que ofrece una bondad de ajuste del 36 %. Entonces:

- a) El 36 % de las predicciones del modelo son correctas.
- b)  $r = +0.60$
- c)  $r = +0.36$
- d)  $r = -0.60$
- e)  $r = -0.36$

Respuesta correcta

Explicacion

### 5.13. Pregunta test

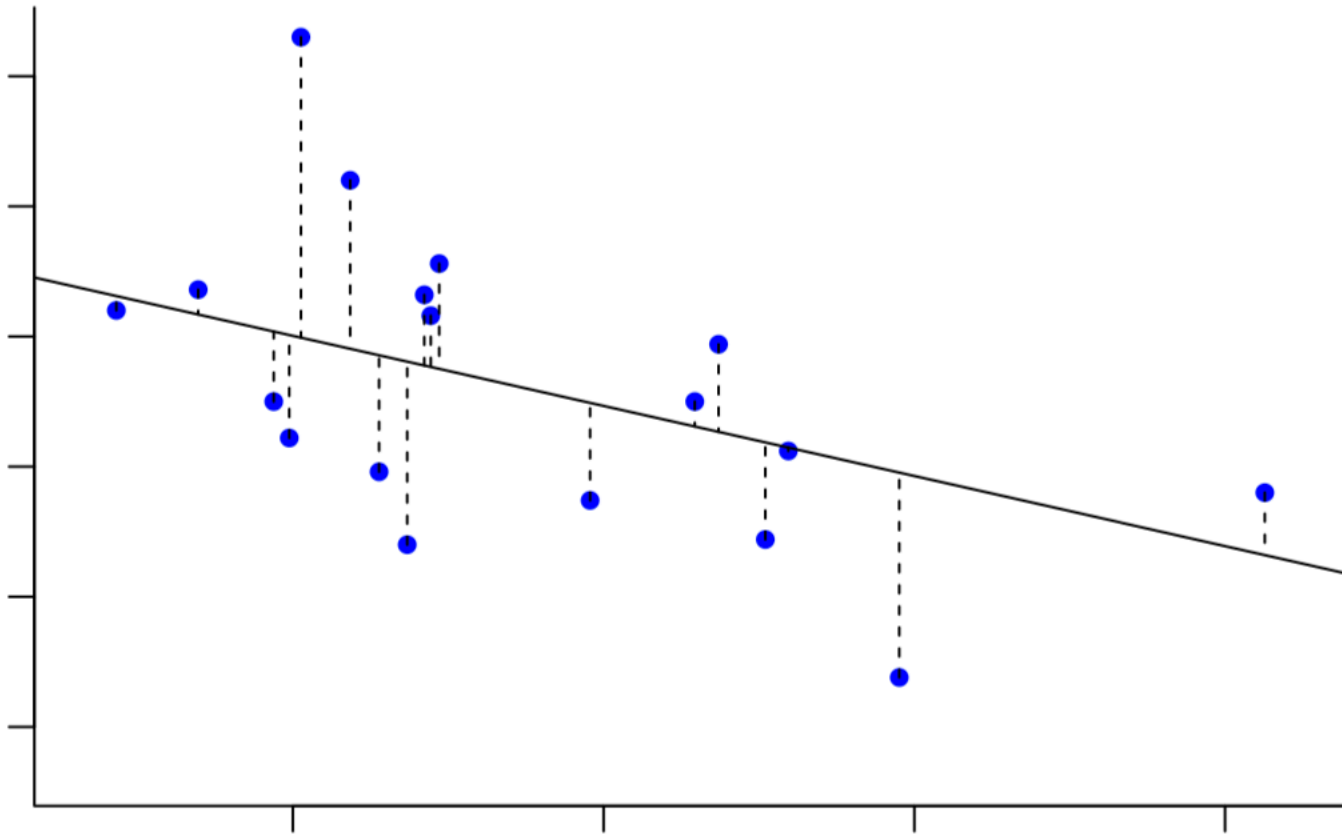
Un modelo de regresión lineal para calcular la glucemia (sangre) a partir de la de la orina (glucosuria) es "glucemia =  $20 + 0.5$  glucosuria". Si dos personas se diferencian en 10 unidades de glucosuria, cual es la mejor estimación que puede hacer para la diferencia en glucemia:

- a) 5
- b) 10
- c) 15
- d) 20
- e) 25

Respuesta correcta

**Explicacion**

El modelo establece que “glucemia =  $20 + 0.5$  glucosuria”. Así, la pendiente (o inclinación) de la recta de regresión es  $0.5 = \frac{1}{2}$



De este modo, si dos personas se diferencian en 10 unidades de glucosuria (en horizontal), la variación en glucemia (en vertical según el modelo) será la mitad, es decir, 5.

### 5.14. Pregunta test

Qué afirmación sobre la covarianza es falsa:

- a) La covarianza es una medida de la variabilidad conjunta de dos variables numéricas.
- b) Si la covarianza es positiva implica una relación creciente entre las variables.
- c) A partir de ella se obtiene el coeficiente de correlación lineal de Pearson.

- d) Posee dimensiones.
- e) Si es 0 podemos afirmar que no existe relación posible entre las variables.

Respuesta correcta

Explicacion

### 5.15. Pregunta test

La pendiente de una recta de una función de regresión lineal es  $Y = b_0 + b_1 \cdot X$

- a) Representa el incremento de  $Y$  por cada unidad de incremento de  $X$ .
- b) Tiene el mismo signo que la covarianza.
- c) Es el valor de la variable  $Y$  cuando  $X = 0$ .
- d) Todas las anteriores son correctas.
- e) Sólo la a) y la b) son correctas.

Respuesta correcta

Explicacion

### 5.16. Pregunta test

De los siguiente estudios de relación entre variables, en cuál crees que no sería oportuno usar la técnica de regresión lineal.

- a) La presión sanguínea y la acidez (ph).
- b) El número de glóbulos rojos y el grupo sanguíneo
- c) La altura y las horas de sueño.
- d) La edad y el conteo de plaquetas.
- e) El nivel de colesterol y la concentración de bilirrubina.

Respuesta correcta

Explicacion

### 5.17. Pregunta test

Después de estudiar la relación existente entre la flexión y la extensión de cuello de los alumnos de la USAL, obtenemos que el valor de la covarianza es  $-0,57$ . ¿El valor de  $r$  saldrá positivo o negativo?

- a) Saldrá positivo porque la relación es inversa.
- b) Saldrá negativo también porque el signo de la covarianza y del coeficiente de correlación lineal de Pearson siempre coinciden.

- c) No podemos saber el signo de  $r$  sabiendo la covarianza porque no están relacionados.
- d) Todas las anteriores son falsas.
- e) Necesitamos conocer  $R^2$  para saber el signo de  $r$

Respuesta correcta

Explicacion

### 5.18. Pregunta test

Si el coeficiente de correlación lineal de Pearson entre dos variables es  $-0,9$  podemos decir que:

- a) La covarianza será positiva.
- b) La relación lineal es buena.
- c) Al ser inferior a 1, la relación lineal es pequeña.
- d) Tenemos una relación lineal inversa, pero no buena.
- e) Sólo dos son correctas.

Respuesta correcta

Explicacion

### 5.19. Pregunta test

Cual de las siguientes propiedades de  $r$  son correctas:

- a) Es adimensional
- b) Cuanto más cerca esté  $r$  de  $+1$  o  $-1$  mejor será el grado de relación lineal
- c) Las variables son incorreladas cuando  $r=0$
- d) Todas las anteriores son correctas
- e) Son todas incorrectas

Respuesta correcta

Explicacion

### 5.20. Pregunta test

Se dice que la relación entre dos variables es directa cuando:

- a) La covarianza es igual a cero
- b) La covarianza es negativa

- c) La covarianza es mayor que cero
- d) El coeficiente de correlación lineal es positivo
- e) Las respuestas c) y d) son correctas

Respuesta correcta

Explicacion

### 5.21. Pregunta test

Sabiendo que  $r = +0.7$  elija la afirmación falsa.

- a) La covarianza es positiva
- b) Hay cierta relación lineal entre las variables
- c) La bondad de ajuste es 0.14
- d) La nube de puntos es creciente
- e) Existe una relación directa

Respuesta correcta

Explicacion

### 5.22. Pregunta test

En un estudio de regresión, ¿cuándo coincidirán los valores de la variable dependiente con los propuestos por el modelo lineal de regresión?

- a) Cuando  $r$  tenga un valor positivo
- b) Cuando  $r$  sea igual a cero
- c) Nunca, aunque el modelo sea perfecto
- d) Cuando  $r$  valga 1 ó -1
- e) Las opciones c) y d) son correctas

Respuesta correcta

Explicacion

### 5.23. Pregunta test

Si el coeficiente de correlación lineal de Pearson entre dos variables es  $-0,82$ , podemos afirmar que:

- a) la relación entre las dos variables es casi nula
- b) la relación que hay entre las variables es muy buena y directa
- c) la covarianza es positiva



- d) la relación que hay entre las variables es muy buena e inversa
- e) sólo dos de las afirmaciones anteriores son correctas

Respuesta correcta

Explicacion

## 5.24. Pregunta test

¿Qué otro nombre reciben los diagramas de dispersión?

- a) Diagrama de regresión
- b) Nube de puntos
- c) Diagrama lineal
- d) Diagrama de relación inversa
- e) Diagrama simple

Respuesta correcta

Explicacion

## 5.25. Pregunta test

Un modelo de regresión lineal para calcular “Fatty liver Index” ( $FLI$ ) a partir del consumo de *aceite* de oliva es “ $FLI = 70 - 4 \cdot \text{aceite}$ ”. Si dos personas se diferencian en 5 unidades de consumo de aceite, cual es la mejor estimación que puede hacer para la diferencia en  $FLI$ :

- a) 5
- b) 10
- c) 15
- d) 20
- e) 60

Respuesta correcta

Explicacion

## 5.26. Pregunta test

El porcentaje de variabilidad explicada por un modelo lineal de regresión es 3 %.

- a) El modelo lineal de regresión es insuficiente para explicar la variable dependiente.
- b) Las variables son incorreladas.

- c) El error cometido por el modelo lineal de regresión es pequeño, por tanto el ajuste lineal es bueno.
- d) Hay una relación creciente entre las variables.
- e) Todo lo anterior es falso.

Respuesta correcta

Explicacion

### 5.27. Pregunta test

Si en un experimento realizado sobre estudiantes voluntarios a los que se coloca en situación de estrés se observa que los cambios en ritmo cardíaco  $RC$  (latidos por minuto) se asocian a cambios en la frecuencia de la voz  $FV$  (Hz) con una bondad de ajuste del 15 % según el modelo  $FV = -5 + 3 \cdot RC$ , marque la afirmación verdadera.

- a) La relación entre las variables es inversa.
- b) Las variables no presentan ninguna relación.
- c) La voz disminuye su frecuencia en 5Hz.
- d) Por cada aumento de 1 latido por minuto cardíaco, se produce un aumento de 3 Hz en frecuencia de la voz,
- e) Todo lo anterior es falso.

Respuesta correcta

Explicacion

### 5.28. Pregunta test

Un modelo de regresión lineal que relaciona el tiempo de estudio con el promedio obtenido en una prueba tiene una bondad de ajuste del 36 %. ¿Qué significa esto en términos de la relación entre el tiempo de estudio y el promedio obtenido?

- a) La relación es muy fuerte y positiva
- b) La relación es muy débil y negativa
- c) La relación es muy fuerte y negativa
- d) La relación es muy débil y positiva
- e) No se puede inferir la relación entre ambas variables a partir de la bondad de ajuste.

Respuesta correcta

Explicacion

### 5.29. Pregunta test

¿Cuál de las siguientes medidas estadísticas es la más adecuada para medir la relación entre el peso y la estatura de un grupo de personas?

- a) Rango intercuartílico
- b) Desviación estándar
- c) Coeficiente de variación
- d) Coeficiente de correlación de Pearson
- e) Coeficiente de asimetría

Respuesta correcta

Explicacion



## Capítulo 6

# Tablas de contingencia.

En este capítulo se resolverán problemas relativos a:

- Contrastes de asociación y homogeneidad en tablas bifactoriales
- Coeficientes de asociación

### 6.1. Problema

Un estudio transversal para conocer la prevalencia de osteoporosis y su relación con algunos factores de riesgo potenciales incluyó a 400 mujeres con edades entre 50 y 54 años. A cada una se le realizó una densitometría de columna y en cada caso se completó un cuestionario de antecedentes. Se pretende determinar si existe una asociación significativa entre la prevalencia de osteoporosis y antecedentes de dieta pobre en calcio. De las 80 pacientes que presentaban osteoporosis 58 presentaban antecedentes de dieta pobre en calcio, en tanto que entre las 320 que no tenían osteoporosis, el número de mujeres con este antecedente era de 62.

**a)** Construye la tabla de contingencia correspondiente y determina, para una nivel de significación del 1 %, si existe una asociación significativa entre la prevalencia de osteoporosis y antecedentes de dieta pobre en calcio.

**b)** Calcula el estadístico Chi-cuadrado corregido (corrección de Yates) y determina en base a ese estadístico si, para un nivel de significación del 5 %, existe una asociación significativa entre la prevalencia de osteoporosis y antecedentes de dieta pobre en calcio.

**c)** Calcula el riesgo relativo.

### 6.1.1. Solución

a) Incluimos los datos en una tabla de contingencia y la completamos:

	Osteoporosis	No osteoporosis	
Dieta pobre en calcio	58	62	120
Dieta rica en calcio	22	258	280
	80	320	400

Establecemos la hipótesis nula y la alternativa:

- $H_0$  : No hay asociación entre las variables (son independientes)
- $H_1$  : Sí hay asociación entre las variables

Ahora, necesitamos calcular el valor del estadístico chi-cuadrado y el valor p asociado para aceptar o rechazar la hipótesis nula. Para ello, podemos utilizar Excel© y seguir las instrucciones de esta práctica del Curso de Bioestadística

En este caso, vamos a utilizar la calculadora online Chi-Square Calculator que nos da los siguientes resultados:

	Ost	No ost
<b>bajo calcio</b>	58 (24) [48.17]	62 (96) [12.5]
<b>rica calcio</b>	22 (56) [20.64]	258 (224) [5.0]
<b>Marginal Column Totals</b>	80	320

The chi-square statistic is 86.0119. The  $p$ -value is  $< 0.00001$ . Significant

The chi-square statistic with Yates correction is 83.5007. The  $p$ -value

Así, el valor del estadístico Chi-cuadrado es 86.01 con un valor p asociado  $<$

0.00001. Se rechaza la hipótesis de independencia al 1 %: hay relación significativa entre padecer osteoporosis y tener antecedentes de dieta pobre en calcio.

b) Nos piden ahora el valor del estadístico Chi-cuadrado corregido (corrección de Yates) para un nivel de significación del 5 % (0.05). Obtenemos:

	Ost	No ost
<b>bajo calcio</b>	58 (24) [48.17]	62 (96) [12.04]
<b>rica calcio</b>	22 (56) [20.64]	258 (224) [5.16]
<b>Marginal Column Totals</b>	80	320

The chi-square statistic is 86.0119. The  $p$ -value is  $< 0.00001$ . Significant at  $p < 0.00001$ .

The chi-square statistic with Yates correction is 83.5007. The  $p$ -value is  $< 0.00001$ .

Ahora la corrección de Yates es 83.5. con valor  $p < 0.00001$ . Se rechaza la hipótesis de independencia al 5 % (hay relación significativa)

c) Recordemos que el Riesgo Relativo es

$$\frac{\text{incidencia acumulada en expuestos}}{\text{incidencia acumulada en no expuestos}}$$

En nuestro caso:

$$RR = \frac{\frac{58}{58+62}}{\frac{22}{22+258}} = 6.15.$$

Así, una mujer con antecedentes de dieta baja en calcio tiene 6 veces más posibilidades de padecer osteoporosis que las que no tienen antecedentes de dieta baja en calcio.

## 6.2. Pregunta test

Supongamos que se quiere estudiar la posible asociación entre el hecho de que una gestante fume durante el embarazo y que el niño presente bajo peso al nacer. Para responder a esta pregunta se realiza un estudio de seguimiento sobre una cohorte de 2000 gestantes, a las que se interroga sobre su hábito tabáquico durante la gestación y se determina además el peso del recién nacido. Los resultados de este estudio se muestran en la siguiente tabla:

	<b>Recién nacido de bajo peso</b>	
<b>Gestante</b>	Sí	No
Fumadora	43	204
No fumadora	105	1645

¿Se puede concluir que, con una confianza del 99 %, existe una relación estadísticamente significativa entre el hecho de que una gestante fume durante el embarazo y que el niño presente bajo peso al nacer?

- a) Sí
- b) No

Respuesta correcta

Explicación

## 6.3. Pregunta test

## 6.4. Problema

La siguiente tabla muestra la clasificación de 1343 niños según el grado de cumplimiento de su calendario vacunal y el nivel socio-cultural de sus padres. Determina si existe una asociación significativa entre el grado de cumplimiento del calendario vacunal de los niños y el nivel socio- cultural de sus padres.



	<b>Cumplimiento calendario</b>	
<b>Nivel socio-cultural</b>	Bajo	Medio
Bajo	114	229
Medio bajo	7	134
Medio alto	7	63
Alto	2	38

#### 6.4.1. Solución

En este caso necesitamos la Chi-Square Calculator for 5 x 5 (or less) Contingency Table. Cumplimentando la tabla y obteniendo los resultados con una confianza del 95 %:

## Chi-Square Calculator

Success! The contingency table below provides the following information (the expected cell totals) and [the chi-square statistic for each cell].

The chi-square statistic,  $p$ -value and statement of significance appear below. The chi-square statistic is 163.5085. The  $p$ -value is  $< 0.00001$ . The results are significant; red, independent.

Results				
	Bajo	Medio	Alto	
Bajo	114 (55.27) [62.40]	229 (197.28) [5.10]	228 (318.45) [25.69]	
Medio bajo	7 (40.46) [27.67]	134 (144.42) [0.75]	277 (233.12) [8.26]	
Medio alto	7 (21.30) [9.60]	63 (76.01) [2.23]	150 (122.70) [6.08]	
Alto	2 (12.97) [9.28]	38 (46.30) [1.49]	94 (74.73) [4.97]	
<b>Column Totals</b>	130	464	749	

The chi-square statistic is 163.5085. The  $p$ -value is  $< 0.00001$ . The results are significant; red, independent.

Así, obtenemos un valor  $p < 0.05$  por lo que existe una asociación significativa entre el grado de cumplimiento del calendario vacunal de los niños y el nivel socio- cultural de sus padres.

### 6.5. Pregunta test

Para evaluar el efecto de la exposición a asbesto sobre el riesgo de fallecer por cáncer de pulmón, un estudio comparó un grupo de 6.245 trabajadores expuestos a este agente con otro grupo de 7.895 trabajadores sin exposición a este factor. A lo largo de 22 años de seguimiento, en el primer grupo se presentaron 76 defunciones por cáncer en el aparato respiratorio, en tanto que en el grupo no expuesto el número de defunciones por esta causa fue 28. ¿Existe una asociación significativa entre la exposición a asbesto y el riesgo de fallecer por cáncer de pulmón?

- a) Sí

b) No

Respuesta correcta

Explicación

## 6.6. Pregunta test

En un estudio sobre VIH se pretende determinar si existe asociación significativa entre la edad del paciente y el nivel de linfocitos CD4. Para ello se determina el nivel de linfocitos CD4 ( $<200$ ,  $200-500$ ,  $>500$ ) en pacientes de 3 grupos de edad. ¿Se puede concluir que existe una relación estadísticamente significativa entre el nivel de linfocitos y la edad del paciente?

a) Sí

b) No

Respuesta correcta

Explicación

## 6.7. Pregunta test

Se quiere estudiar la posible asociación entre la presencia de infección postoperatoria (IPO) y la diabetes (DIAB) en una población de operados. En una muestra de 1337 personas de edad  $< 65$  años y en otra de 892 de edad  $\geq 65$  años se obtuvieron los siguientes resultados:

	<b><math>&lt; 65</math> años</b>	
<b>DIAB</b>	IPO	NO IPO
Sí	15	29
No	190	1103

<b>DIAB</b>	
Sí	
No	

¿Existe asociación significativa entre IPO y diabetes en cada grupo de edad?

- a) Hay asociación significativa en los dos grupos
- b) No hay asociación significativa en ningún grupo
- c) la hay para  $< 65$  y no para  $\geq 65$

d) no la hay para  $< 65$  y sí para  $\geq 65$

Respuesta correcta

Explicación

## 6.8. Pregunta test

Se realizó un estudio de seguimiento para detectar la posible asociación entre enfermedades cardiovasculares y el exceso de peso. Se eligieron 1990 hombres con edades entre 55 y 59 años de estatura similar. Tras 5 años de seguimiento se observaron los datos resumidos en la tabla.

	<b>Peso</b>		
<b>Infarto</b>	55 – 64 kg.	65 – 74 kg.	75 – 84 kg.
Sí	8	18	48
No	290	680	550

¿Se puede admitir que el exceso de peso se asocia con el infarto de miocardio?

- a) Sí
- b) No

Respuesta correcta

Explicación

## Capítulo 7

# Medidas de importancia clínica.

En este capítulo se resolverán problemas relativos a:

- Diferencias entre Proporción, Tasa, Razón, odds.
- Medidas de asociación en tablas 2x2. Riesgo Relativo. Riesgo Absolutos. Odds-Ratio.
- Indicadores estadísticos básicos para evaluar el desempeño de un procedimiento diagnóstico: Sensibilidad y Especificidad. Probabilidades pre y post prueba.

### 7.1. Pregunta test

¿Cómo se denomina la medida de epidemiología que indica la probabilidad de que una enfermedad se desarrolle en un grupo de individuos expuesto a un factor de riesgo comparada con la del grupo no expuesto?

- a) Incidencia acumulada
- b) Densidad de incidencia
- c) Fracción atribuible
- d) Prevalencia
- e) Riesgo relativo

Respuesta correcta

Explicación

## 7.2. Pregunta test

En un estudio de casos y controles si se constata asociación estadística mediante la chi- cuadrado clásica ¿Cómo mediremos la magnitud de la asociación?

- a) Riesgo relativo
- b) Riesgo atribuible
- c) Fracción de riesgo atribuible
- d) Odds ratio
- e) Densidad de incidencia

Respuesta correcta

Explicación

## 7.3. Pregunta test

A la capacidad de una prueba diagnóstica para identificar correctamente a los que no padecen la enfermedad se le denomina:

- a) Sensibilidad
- b) Especificidad
- c) Valor predictivo positivo
- d) Valor predictivo negativo
- e) Razón de probabilidad

Respuesta correcta

Explicación

## 7.4. Problema

Si una prueba diagnóstica tiene una sensibilidad del 75 % y una especificidad del 85 %, calcule la razón de probabilidad positiva.

### 7.4.1. Solución

La razón de probabilidad positiva, también conocido como *cociente de probabilidad positivo* (CPP) o *likelihood ratio* (LR), nos indica cuánto más probable es tener un positivo en la prueba en un enfermo que en un sano. El uso del LR constituye una herramienta de gran utilidad para la toma de decisiones clínicas frente a la solicitud de algún test diagnóstico, porque son valores inherentes a éste e independientes de la prevalencia de la enfermedad.

Matemáticamente, sería el resultado del cociente:

$$\frac{\text{positivos en enfermos}}{\text{positivos en sanos}}$$

Sabemos que:

- la proporción de positivos en los enfermos es la Sensibilidad
- la proporción de los positivos en sanos son los Falsos Positivos que serían aquellos sanos que no dan negativo o, lo que es lo mismo,  $1 - \text{Especificidad}$ .

Entonces, en nuestro problema:

$$CPP = \frac{S}{1-E} = \frac{0.75}{1-0.85} = \frac{0.75}{0.15} = 5$$

Así, **la razón de probabilidad positiva es 5** lo que indica que en esta prueba diagnóstica, un enfermo tiene 5 veces más posibilidades de dar positivo que un sano.

Explicación

## 7.5. Pregunta test

La proporción de enfermos que han dado un resultado negativo en la prueba diagnóstica dividido entre la proporción de sanos que también han dado negativo en dicha prueba se denomina:

- a) Razón de probabilidad positiva
- b) Valor predictivo negativo
- c) Valor predictivo positivo
- d) Razón de probabilidad negativa
- e) Especificidad

Respuesta correcta

Explicación

## 7.6. Pregunta test

Si al aplicar una prueba diagnóstica se observa un 10 % de falsos positivos cuál de las siguientes afirmaciones es cierta:

- a) La sensibilidad es el 90 %
- b) La especificidad es el 90 %
- c) El valor predictivo positivo es el 90 %
- d) El cociente de probabilidad positivo es del 90 %
- e) La sensibilidad es el 10 %

Respuesta correcta

Explicación

## 7.7. Problema

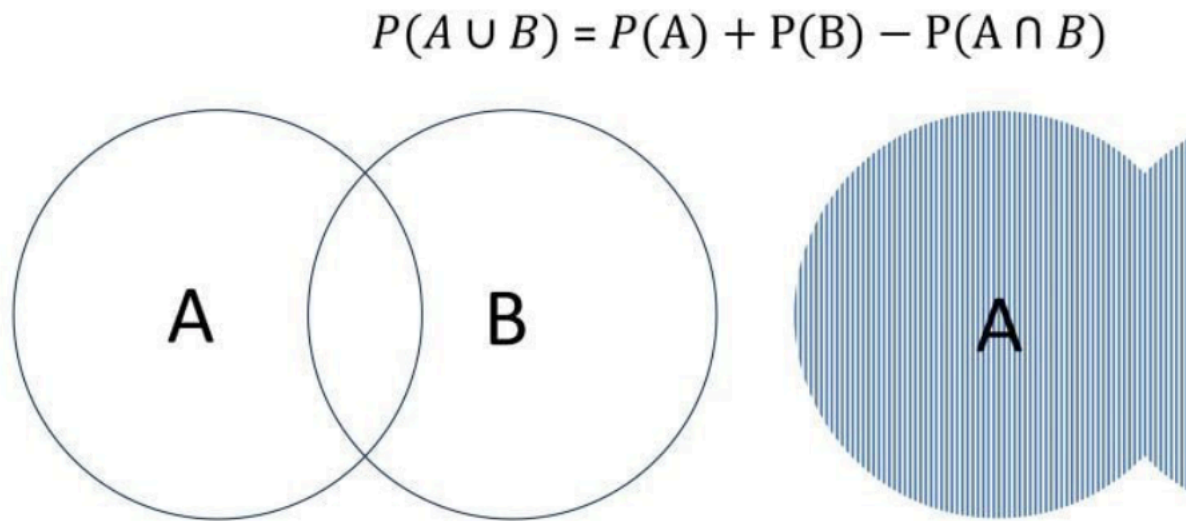
Si la probabilidad de tener la enfermedad  $A$  es del 5 %, la de tener la enfermedad  $B$  es del 10 % y la de tener al menos una de las dos es del 13 %, ¿cuál es la probabilidad de tener las dos?

### 7.7.1. Solución

Éste es un típico problema de Probabilidad con sucesos de Bachillerato que se resuelve muy fácilmente recordando la operaciones con sucesos y los Diagramas de Venn:

- $A$  = tener la enfermedad  $A$   $P(A) = 0.05$
- $B$  = tener la enfermedad  $B$   $P(B) = 0.10$
- $A \cap B$  = tener las dos enfermedades
- $A \cup B$  = tener, al menos, una de las dos enfermedades  $P(A \cup B) = 0.13$

Con un diagrama de Venn:



Así,  $P(A \cup B) = P(A) + P(B) - P(A \cap B) \Rightarrow 0.13 = 0.05 + 0.10 - P(A \cap B)$

de donde

$$P(A \cap B) = 0.02$$

Entonces, la probabilidad de tener las dos enfermedades es del 2 %.



## 7.8. Pregunta test

En una prueba diagnóstica es importante determinar:

- a) La sensibilidad
- b) La especificidad
- c) El valor predictivo positivo
- d) Todas son ciertas

Respuesta correcta

Explicación

## 7.9. Pregunta test

Si una prueba diagnóstica se aplica a un grupo de población en el que la prevalencia de la enfermedad es superior a la de la población general aumentará su:

- a) Sensibilidad
- b) Especificidad
- c) Valor predictivo positivo
- d) Razón de probabilidad positivo
- e) Sensibilidad y especificidad

Respuesta correcta

Explicación

## 7.10. Problema

Si una prueba diagnóstica que tiene una sensibilidad del 90 % y una especificidad también del 90 % se aplica a una población de 200 individuos con una prevalencia de enfermedad del 50 % ¿Cuál será el valor predictivo positivo?

### 7.10.1. Solución

Tenemos que recordar que

$$VPP = \frac{\text{Verdaderos positivos}}{\text{Total positivos}} = \frac{\text{Verdaderos positivos}}{\text{Verdaderos positivos} + \text{Falsos positivos}}$$

Vamos a hacer una tabla con los datos del problema. Sabemos que:

	Enfermo	Sano
Positivo (+)		

	Enfermo	Sano	
Negativo (-)	100	100	200

ya que se aplica a una población de 200 individuos con una prevalencia de enfermedad del 50 % . Ahora:

- la Sensibilidad es la proporción de enfermos que son diagnosticados como positivos (proporción de verdaderos positivos) por lo que:

$0.90 = \hat{P}(+/E) = \frac{\text{verdaderos positivos}}{100}$  de donde el numero de verdaderos positivos es 90.

- la Especificidad es la proporción de sanos diagnosticados como negativos (proporción de verdaderos negativos) por lo que:

$0.90 = \hat{P}(-/S) = \frac{\text{verdaderos negativos}}{100}$  de donde el numero de verdaderos negativos es 90.

En la tabla:

	Enfermo	Sano	
Positivo (+)	90		
Negativo (-)		90	
	100	100	200

Terminando de cumplimentar los datos que faltan, tenemos:

	Enfermo	Sano	
Positivo (+)	90	10	100
Negativo (-)	10	90	100
	100	100	200

Nos piden el VPP:

$$VPP = \frac{\text{Verdaderos positivos}}{\text{Total positivos}} = \frac{90}{100} = 0.90$$

por lo que **el VPP de la prueba es del 90 %**.

Explicación

### 7.11. Pregunta test

Cierto tests diagnóstico acierta sobre el 100 % de los individuos enfermos y el 50 % de los sanos. Cierta persona pasa el test con resultado negativo. Entonces:

- a) Esta sana.
- b) Esta enferma.
- c) Existe una probabilidad del 50 % de que esté sana.
- d) Existe una probabilidad del 75 % de que esté sana.
- e) Existe una probabilidad del 75 % de que esté enferma.

Respuesta correcta

Explicación

### 7.12. Pregunta test

Si aplicamos una prueba de laboratorio para el diagnóstico de una determinada enfermedad que es 2 veces más frecuente en hombres que en mujeres ¿Cuál de los siguientes parámetros será más elevado en la población femenina que en la masculina?

- a) La prevalencia de la enfermedad
- b) La sensibilidad de la prueba
- c) La especificidad de la prueba
- d) El valor predictivo positivo de la prueba
- e) El valor predictivo negativo de la prueba

Respuesta correcta

Explicación

### 7.13. Pregunta test

¿Cómo se calcula la sensibilidad de un test diagnóstico?

- a) Contabilizando el número de tests positivos en una muestra aleatoria de individuos.
- b) Contabilizando el número de tests negativos en una muestra aleatoria de individuos.
- c) Contabilizando el número de tests positivos en una muestra aleatoria de enfermos.
- d) Contabilizando el número de tests negativos en una muestra aleatoria de sanos.
- e) Ninguna de las anteriores es cierta.

Respuesta correcta

Explicación

### 7.14. Pregunta test

¿Cuál es la proporción de enfermos sobre el total?

- a) La prevalencia
- b) La incidencia
- c) El total de positivos para el test
- d) El valor predictivo
- e) El total de las personas estudiadas

Respuesta correcta

Explicación

### 7.15. Pregunta test

Cierto test diagnóstico acierta sobre el 100 % de los individuos sanos y el 0 % de los individuos enfermos. Elegida una persona al azar:

- a) Hay una probabilidad del 50 % de que esté enferma.
- b) Hay una probabilidad del 0 % de que esté enferma.
- c) Hay una probabilidad del 100 % de que esté enferma.
- d) El test será negativo.
- e) Ninguna de las anteriores es cierta.

Respuesta correcta

Explicación

### 7.16. Pregunta test

El parámetro que mide la fuerza de asociación entre la exposición y la enfermedad se denomina:

- a) Factor de riesgo
- b) Riesgo atribuible
- c) Factor protector
- d) Riesgo relativo

Respuesta correcta

Explicación

### 7.17. Pregunta test

En una población, hay tantos hombres como mujeres, el 20 % son varones y fumadores y el 20 % de las mujeres fuman. Entonces:

- a) Fuman tantos hombres como mujeres.
- b) Por cada mujer fumadora hay dos hombres fumadores.
- c) Por cada hombre fumador hay dos mujeres fumadoras.
- d) Hay un 40 % de fumadores en la población.
- e) Nada de lo anterior es cierto.

Respuesta correcta

#### Explicación

Con los datos del problema, tenemos:

	Hombre	Mujer	
Fuma (+)	20	10	
No fuma (-)	50	50	100

(El 20 % de las mujeres fuman, el 20 % del 50 % es 10)

### 7.18. Pregunta test

Para conocer el exceso de riesgo en los individuos expuestos comparando con los no expuestos utilizaremos:

- a) Riesgo relativo
- b) Diferencia de incidencias
- c) Odds ratio
- d) Incidencia acumulada
- e) Riesgo atribuible

Respuesta correcta

Explicación

### 7.19. Pregunta test

Para estudiar la efectividad de un test diagnóstico ante una enfermedad se toma un grupo de 200 personas enfermas y 200 que no la padecen, y se observan los resultados. ¿Qué podemos estimar directamente de ellos?

- a) La sensibilidad y especificidad del test.
- b) La incidencia de la enfermedad en la población.
- c) El índice predictivo de verdaderos positivos.
- d) Son correctas (a) y (c).
- e) Todo lo anterior.

Respuesta correcta

Explicación

### 7.20. Pregunta test

¿Cuál de las siguientes medidas utilizaría para cuantificar el impacto potencial de un programa preventivo en la población?

- a) Riesgo relativo
- b) Odds ratio
- c) Razón de prevalencia
- d) Disminución de la prevalencia
- e) Riesgo atribuible

Respuesta correcta

Explicación

### 7.21. Pregunta test

El porcentaje de individuos fumadores o con bronquitis se puede interpretar como una probabilidad:

- a) De un suceso intersección
- b) Condicionada.
- c) De un suceso unión.
- d) A posteriori.
- e) De un suceso complementario.

Respuesta correcta

Explicación

### 7.22. Pregunta test

El porcentaje de individuos con bronquitis entre los fumadores se puede interpretar como una probabilidad:

- a) De un suceso intersección
- b) Condicionada.
- c) De un suceso unión.
- d) A posteriori.
- e) De un suceso complementario.

Respuesta correcta

Explicación

### 7.23. Pregunta test

El porcentaje de individuos con bronquitis que además son fumadores se puede interpretar como una probabilidad:

- a) De un suceso intersección
- b) Condicionada.
- c) De un suceso unión.
- d) A posteriori.
- e) De un suceso complementario.

Respuesta correcta

Explicación

### 7.24. Problema

El 12 % de los individuos de una población padece osteoporosis. EL 25 % de ellos lo sabe. ¿Qué tasa de individuos tiene osteoporosis y lo desconoce?

#### 7.24.1. Solución

Si el 25 % de los que tienen osteoporosis lo sabe, el 75 % de los que tienen osteoporosis lo desconoce. Por lo tanto, la tasa de individuos que tiene osteoporosis y lo desconoce será el 75 % del 12 %, es decir,  $0.75 \cdot 0.12 = 0.09$ . Entonces, **la tasa de individuos que tiene osteoporosis y lo desconoce es del 9 %**.

### 7.25. Pregunta test

¿Cómo se denomina la proporción de enfermos que presentan un resultado positivo de un método diagnóstico?

- a) Valor predictivo positivo del método

- b) Valor predictivo negativo del método
- c) Especificidad del método
- d) Razón de probabilidad positiva
- e) Ninguna de las anteriores

Respuesta correcta

Explicación

### 7.26. Pregunta test

Una prueba con alta sensibilidad:

- a) Presenta pocos falsos negativos
- b) Presenta pocos falsos positivos
- c) Tiene una  $p < 0.05$
- d) Necesariamente tiene una especificidad alta
- e) Presenta muchos falsos positivos

Respuesta correcta

Explicación

### 7.27. Pregunta test

La probabilidad de que un individuo tomado aleatoriamente en una serie de sujetos de estudio tenga un resultado negativo en las pruebas diagnosticas si realmente no tiene la enfermedad se denomina:

- a) Sensibilidad
- b) Especificidad
- c) Proporción de falsos negativos
- d) Proporción de falsos positivos
- e) Valor predictivo negativo

Respuesta correcta

Explicación

### 7.28. Problema

La osteoporosis afecta 4 veces más a mujeres que a hombres. El 8% de las mujeres padece osteoporosis en una población donde hay tantos hombres como mujeres. ¿Cuál es la prevalencia de la osteoporosis en la población?



**7.28.1. Solución**

Hagamos una tabla con los datos:

	Hombres	Mujeres	
Osteoporosis (+)	1	4	
No Osteoporosis (-)	50	50	100

ya que el 50 % son mujeres y el 8 % padece osteoporosis ( $0.08 \cdot 0.50 = 0.04$ ) y los hombres que padecen de osteoporosis son la cuarta parte que las mujeres.

Si completamos la tabla:

	Hombres	Mujeres	
Osteoporosis (+)	1	4	5
No Osteoporosis (-)	49	46	95
	50	50	100

por lo que la **prevalencia de la osteoporosis en la población es del 5 %**

**7.29. Pregunta test**

Para conocer los índices (valores) predictivos en un test diagnóstico para una enfermedad que tiene un 1 % de afectados en la población, será necesario conocer:

- a) Sensibilidad y verdaderos positivos
- b) Prevalencia.
- c) Verdaderos positivos y prevalencia.
- d) Especificidad y verdaderos negativos
- e) Falsos positivos y verdaderos positivos.

Respuesta correcta

Explicación

**7.30. Pregunta test**

Elija la afirmación correcta relativa a pruebas diagnósticas:

- a) La sensibilidad se obtiene usando la noción subjetiva de probabilidad.

- b) El índice predictivo positivo se obtiene directamente de la noción frecuentista de probabilidad.
- c) La tasa de verdaderos positivos se obtiene directamente de la noción frecuentista de probabilidad.
- d) La prevalencia de la enfermedad se obtiene a partir del teorema de Bayes.
- e) nada de lo anterior es cierto.

Respuesta correcta

Explicación

### 7.31. Pregunta test

El 2 % de la población padece diabetes. Si de ellos, el 30 % no está diagnosticado, esta cantidad puede entenderse como una probabilidad...

- a) De un suceso intersección
- b) Condicionada.
- c) De un suceso unión.
- d) A posteriori.
- e) De un suceso complementario.

Respuesta correcta

Explicación

### 7.32. Problema

En una población, el 5 % son enfermos diagnosticados de una enfermedad, la cual padece el 10 % de la población. Calcular la probabilidad de estar diagnosticado para un individuo enfermo.

#### 7.32.1. Solución

Éste es un problema de probabilidad condicionada. Recordamos:

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

En nuestro caso:

- $E$  = estar enfermo  $P(E) = 0.10$
- $D$  = estar diagnosticado
- $E \cap D$  = estar enfermo y diagnosticado  $P(E \cap D) = 0.05$

Se pide calcular la probabilidad de estar diagnosticado para un individuo enfermo, es decir,  $P(D/E)$ . Entonces:

$$P(D/E) = \frac{P(D \cap E)}{P(E)} = \frac{0.05}{0.10} = 0.50$$

Así, la probabilidad de estar diagnosticado para un individuo enfermo es del 50 %.

### 7.33. Pregunta test

Una prueba diagnóstica de cierta enfermedad, tiene una tasa de aciertos del 90 % tanto sobre enfermos como sanos. La incidencia de la enfermedad en la población es del 50 %. Si se pasa el test a una persona y sale positivo, la probabilidad de que realmente esté enferma es:

- a) 45 %
- b) 50 %
- c) 75 %
- d) 90 %
- e) 100 %

Respuesta correcta

Explicación

### 7.34. Problema

Una enfermedad tiene una incidencia del 50 % en la población. Un test para detectarla posee una tasa de verdaderos positivos del 80 %, y de falsos positivos del 20 %. Si un individuo resulta ser positivo, calcular la probabilidad de que esté enfermo.

#### 7.34.1. Solución

En este caso, vamos a hacer una tabla de contingencia con los datos para 200 individuos:

	Enfermo (E)	Sano (S)	
Positivo (+)	80	20	
Negativo (-)			
	100	100	200

Si completamos la tabla:

	Enfermo (E)	Sano (S)	
Positivo (+)	80	20	100
Negativo (-)	20	80	100
	100	100	200

Se pide que, si un individuo resulta ser positivo, calcular la probabilidad de que esté enfermo, es decir  $P(E/+) = \frac{80}{100} = 0.80$ .

Entonces **Si un individuo resulta ser positivo, la probabilidad de que esté enfermo es del 80 %.**

### 7.35. Pregunta test

En una población el 30 % son hombres de los cuales son deportistas el 20 %, frente al 25 % de las mujeres. Escogida una persona al azar es deportista. La probabilidad de que sea mujer es (aproximadamente):

- a) 0,235
- b) 0,60
- c) 0,74
- d) 0,25
- e) No puede calcularse con esos datos.

Respuesta correcta

## Capítulo 8

# Referencias

- Apuntes y vídeos de Bioestadística, Francisco Javier Barón López
- Matemática Aplicada y Estadística (Grado en Farmacia) 2011/2012
- Dr. Eduardo Buesa. Jefe emérito del Servicio de Pediatría. Hospital General de Castellón
- Exámenes de Estadística
- EMEI. Epidemiología Molecular de Enfermedades Infecciosas. Test de Bioestadística