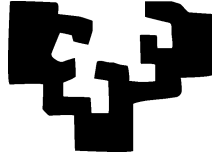


eman ta zabal zazu



Universidad
del País Vasco

Euskal Herriko
Unibertsitatea

INSTRUCCIONES

1. Las cuestiones respondidas correctamente valen un punto. Hay una única respuesta correcta para cada cuestión. Las respuestas fallidas tienen una penalización de -0.25 puntos, por tanto, es mejor dejar una cuestión sin responder que dar una respuesta errónea.
2. Nuestro objetivo es valorar hasta qué punto has asimilado y comprendido la materia impartida durante el curso. Sin embargo, en un examen de elección múltiple hay que prestar mucha atención a los detalles. Es bastante habitual que estudiantes preparados malgasten sus oportunidades de obtener una buena nota por no dedicar suficiente atención al enunciado de las cuestiones.

En tu propio beneficio, lee cuidadosamente la pregunta antes de contestarla!

3. Probablemente te sea de ayuda descartar primero aquellas respuestas que sean claramente incorrectas.
4. Aquellos estudiantes que consigan 33 puntos o más en este examen no tendrán que examinarse de la parte de Estadística Descriptiva en el final de Junio. Incluso si no consigues los 33 puntos, no pierdes nada por realizar este examen: en ningún caso el resultado de esta prueba puede ir en detrimento de la nota que obtengas en Junio.
5. El tiempo asignado para realizar este examen es 2h 50'.

No dar la vuelta a esta hoja hasta que se avise!

EPE

Examen parcial, 11 Abril 2008, Tipo: A

Apellidos: _____

Nombre: _____

DNI: _____

Grupo: _____

Profesor : _____

Sección 1. Preguntas de elección múltiple

1. ¿Cuál de las siguientes expresiones corresponde a la covarianza entre dos variables estadísticas X e Y ?
 - (a) $\sum_i \sum_j (x_i - \bar{x})(y_j - \bar{y})n_{ij}$
 - (b) $\frac{1}{N} \sum_i \sum_j (x_i - \bar{x})^2 (y_j - \bar{y})^2 n_{ij}$
 - (c) $\frac{1}{N} \sum_i x_i y_i - \bar{x} \frac{1}{N} \sum_i y_i - \bar{y} \frac{1}{N} \sum_i x_i + \bar{x} \bar{y}$
 - (d) $m_{11} - \bar{x} \bar{y}$
 - (e) Todas las respuestas son falsas

COMIENZO DE UN BLOQUE DE PREGUNTAS

2. El precio del producto A en el año 2008 es 85, tras un incremento del 7% entre los años 2005 y 2008. El precio del producto en 2005 era aproximadamente:
 - (a) 80.75
 - (b) 79.44
 - (c) 64
 - (d) Todas las respuestas son falsas
 - (e) 68
3. La tasa de crecimiento medio anual acumulativa del precio de A entre 2005 y 2008 ha sido aproximadamente:
 - (a) 0.0233
 - (b) 0.774
 - (c) 0.0172
 - (d) 0.0228
 - (e) Todas las respuestas son falsas

4. Suponiendo que esta tasa de crecimiento anual del precio de A se mantuviese constante desde 2008 hasta 2012, el precio de A en 2012 sería aproximadamente:
- (a) 96.825
 - (b) 91.35
 - (c) 89
 - (d) 93.02
 - (e) Todas las respuestas son falsas
5. Si el Índice de Precios al Consumo ha subido un 5% entre 2005 y 2008, el precio real de A en 2008 expresado en euros del 2005 es aproximadamente:
- (a) 84.5
 - (b) 73.91
 - (c) 80
 - (d) 80.95
 - (e) Todas las respuestas son falsas
6. Hay otro producto, B, cuyo precio entre 2005 y 2008 pasó de 125 a 130. Las cantidades consumidas de A y B fueron 25 y 50, respectivamente, el el año 2005. Entonces, el índice de Laspeyres de precios de los dos productos para el año 2008 (Base=2005) es aproximadamente:
- (a) 102
 - (b) 150
 - (c) 111.27
 - (d) 104.72
 - (e) Todas las respuestas son falsas

FINAL DE UN BLOQUE DE PREGUNTAS

7. ¿Cuál de los siguientes estadísticos permanece invariante ante transformaciones $Y = aX$?
- (a) La media
 - (b) El coeficiente de variación
 - (c) La moda
 - (d) El recorrido intercuartílico
 - (e) La desviación típica
8. ¿Cuál de los siguientes estadísticos se diferencia *necesariamente* del primer cuartil (q_1 ó $q_{0.25}$, dependiendo de la notación usada) *menos* que un recorrido intercuartílico?
- (a) La media
 - (b) La mediana
 - (c) La moda
 - (d) No se puede decir

COMIENZO DE UN BLOQUE DE PREGUNTAS

Las preguntas situadas antes de la próxima línea horizontal se refieren a la siguiente tabla.

Descargar como:



Indice de Precios de Consumo
Base 2006. Series desde enero de 2002

Indices nacionales de clases
 Unidades: Base 2006=100

	Indice			
	Pan y cereales	Carnes	Tabaco	Electricidad
2007M01	102,173	102,483	108,475	103,139
2007M02	103,361	102,161	108,608	103,139
2007M03	103,722	102,673	108,637	103,139
2007M04	104,104	103,886	108,749	103,139
2007M05	104,389	104,345	108,766	103,139
2007M06	104,362	104,058	108,835	103,139
2007M07	104,411	104,358	108,850	103,139
2007M08	104,813	104,576	108,858	103,139
2007M09	106,447	105,703	108,858	103,139
2007M10	109,853	106,636	108,853	103,139
2007M11	110,551	106,855	108,860	103,139
2007M12	111,290	107,377	108,861	103,139
2008M01	112,559	107,080	111,492	106,482

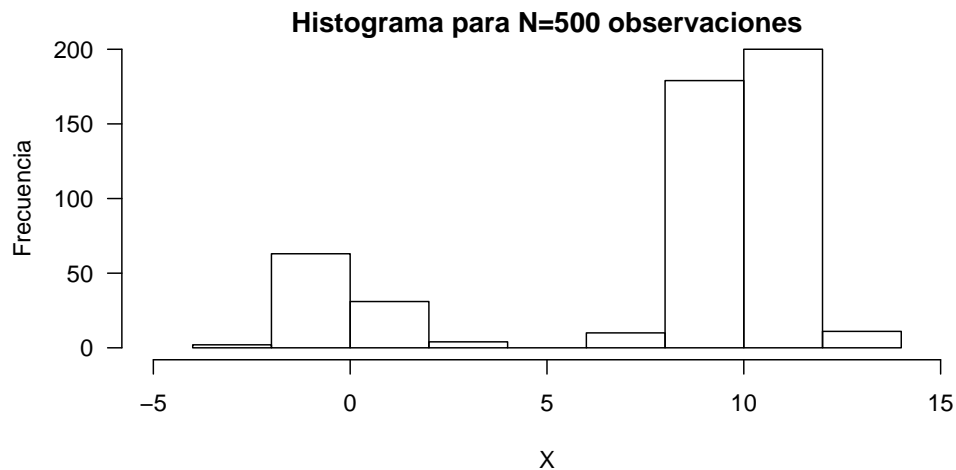
9. De las cuatro variables consideradas, la que ha experimentado un mayor incremento en el precio de Enero 2007 a Enero 2008 es:
- (a) Pan y cereales
 - (b) Carnes
 - (c) Electricidad
 - (d) Tabaco
 - (e) La tabla sólo proporciona el *índice*, por tanto no puede calcularse el incremento del precio

10. De las cuatro variables consideradas, la que ha experimentado un mayor incremento en el precio desde el año 2006 es:
- (a) Pan y cereales
 - (b) Carnes
 - (c) Electricidad
 - (d) Tabaco
 - (e) La tabla sólo proporciona datos desde Enero 2007, por tanto no puede obtenerse la evolución de los precios desde el año 2006
11. En el periodo base 2006, una persona gastó toda su renta en los cuatro productos considerados, en las siguientes proporciones: 25 % Pan, 30 % Carne, 10 % Tabaco, 35 % Electricidad. En Diciembre 2007 consume exactamente las mismas cantidades que en el año 2006. Entonces, necesita gastar
- (a) 5.1 % más
 - (b) 8.7 % más
 - (c) 12.2 % más
 - (d) 13.4 % más
 - (e) No se puede decir

FINAL DE UN BLOQUE DE PREGUNTAS

COMIENZO DE UN BLOQUE DE PREGUNTAS

Las preguntas situadas antes de la próxima línea horizontal se refieren a la variable X cuyo histograma aparece a continuación.



12. Es completamente obvio que:
- (a) La mediana está alrededor de 5
 - (b) La moda está alrededor de 5
 - (c) La mediana es mayor que la moda
 - (d) La mediana estará a la derecha de 5
 - (e) La moda estará a la izquierda de 5

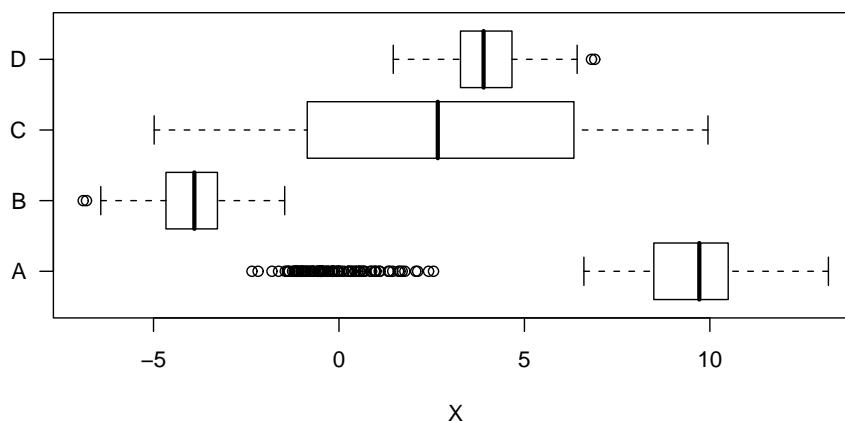
13. El máximo valor que toma la variable X es:
- (a) Cercano a 200
 - (b) No se puede decir con la información dada
 - (c) Menor que 15
 - (d) $500/15$
 - (e) Mayor que $25 = 500/(15 - (-5))$

14. Los valores de la variable X han sido tabulados como se indica a continuación:

X					
Clase	n_i	Clase	n_i	Clase	n_i
$(-4, -2]$	2	$(2, 4]$	4	$(8, 10]$	179
$(-2, 0]$	63	$(4, 6]$	0	$(10, 12]$	200
$(0, 2]$	31	$(6, 8]$	10	$(12, 14]$	11

El número total de observaciones es $N = 500$. A partir de esta tabla, podemos decir que el intervalo modal y el intervalo mediano son, respectivamente:

- (a) $(8, 10]$ y $(6, 8]$
 - (b) $(10, 12]$ y $(6, 8]$
 - (c) $(8, 10]$ y $(8, 10]$
 - (d) $(8, 10]$ y $(10, 12]$
 - (e) $(10, 12]$ y $(8, 10]$
15. Mira atentamente el siguiente gráfico.



¿Cuál de los diagramas de caja, en caso de que haya alguno, podría corresponder a las mismas observaciones que las del histograma mostrado al comienzo de este bloque de preguntas?

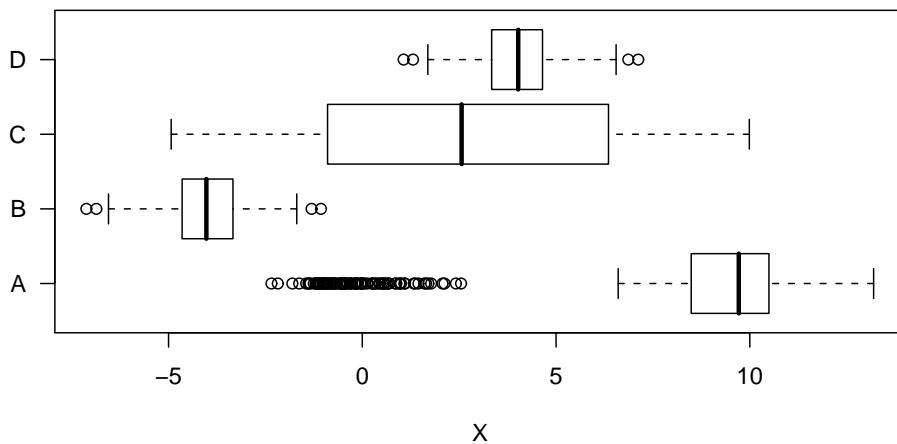
- (a) Diagrama de caja C
- (b) Diagrama de caja B
- (c) Diagrama de caja A
- (d) Diagrama de caja D
- (e) Ninguno

FINAL DE UN BLOQUE DE PREGUNTAS

16. Si transformamos una variable estadística X de la siguiente manera: $Y = aX + b$
- (a) $S_y^2 = S_x^2$; la varianza no varía ante transformaciones lineales
 - (b) $S_y^2 = a^2 S_x^2$
 - (c) $S_y^2 = a^2 S_x^2 + b$
 - (d) $S_y^2 = a^2 S_x^2 + b^2$
 - (e) Todas las respuestas son falsas
17. La capital de España es:
- (a) París
 - (b) Sebastopol
 - (c) Madrid
 - (d) Kuala Lumpur
18. ¿Cuál de los siguientes estadísticos permanece invariante ante transformaciones lineales $Y = X + b$?
- (a) Todos los cuartiles
 - (b) Todas las respuestas son falsas
 - (c) La mediana
 - (d) La varianza
 - (e) La media

COMIENZO DE UN BLOQUE DE PREGUNTAS

Fíjate en el siguiente gráfico:



19. ¿Cuál de los diagramas de caja, en caso de que haya alguno, representa una distribución con un recorrido intercuartílico mayor que 5?
- (a) Diagrama de caja D
 - (b) Diagrama de caja B
 - (c) Diagrama de caja C
 - (d) Diagrama de caja A
 - (e) Todas las respuestas son falsas

20. ¿Cuál de los diagramas de caja, en caso de que haya alguno, representa una distribución con una mediana mayor que 5?
- (a) Diagrama de caja D
 - (b) Diagrama de caja B
 - (c) Diagrama de caja A
 - (d) Diagrama de caja C
 - (e) Todas las respuestas son falsas
21. ¿Cuál de los diagramas de caja, en caso de que haya alguno, representa una distribución con el 25 % de las observaciones más grandes por encima de 10?
- (a) Diagrama de caja D
 - (b) Diagrama de caja B
 - (c) Diagrama de caja A
 - (d) Todas las respuestas son falsas

FINAL DE UN BLOQUE DE PREGUNTAS

COMIENZO DE UN BLOQUE DE PREGUNTAS

22. Del *Boletín Estadístico* del Banco de España hemos extraído datos mensuales del rendimiento anual de las Letras del Tesoro, X , y del índice IBEX35, Y , desde Enero de 1988 hasta Enero de 2008. Con los datos de estas dos variables hemos obtenido los siguientes estadísticos:

Variable		Estadísticos							
		Media	Mediana	Moda	S^2	g_2	g_1	Min	Max
Letras del Tesoro	X	6.84	4.93	14.49	17.68	-1.25	0.51	1.84	14.52
IBEX35	Y	11.48	12.26	—	552.54	-0.02	0.35	-38.21	88.21

A partir de estos datos es claro que la mitad de los valores de X son menores o iguales que:

- (a) 6.84
 - (b) 17.68
 - (c) 14.49
 - (d) 4.93
 - (e) 4.2
23. El coeficiente de variación de Y es igual a:
- (a) 0.488
 - (b) 0.350
 - (c) -0.02
 - (d) 2.04
 - (e) 552.54

24. ¿Cuál de las dos medias es más representativa?
- (a) La media de Y , ya que ésta variable no tiene moda
 - (b) La media de X , ya que $g_2(X) < g_2(Y)$
 - (c) La media de Y , ya que $g_0(Y) < g_0(X)$
 - (d) La media de X , ya que $g_0(X) < g_0(Y)$
 - (e) La media de Y , ya que se parece más a su mediana
25. En un determinado mes el rendimiento de las Letras fue 11 y el del IBEX35 19, ¿cuál de estos dos valores es más grande en términos relativos, es decir, destaca más en relación al resto de observaciones?
- (a) IBEX35, ya que está 8 puntos por encima del rendimiento de las Letras
 - (b) IBEX35, por corresponderle un mayor valor centrado
 - (c) Letras del Tesoro, por corresponderle un mayor valor tipificado
 - (d) Ambos valores son iguales en términos relativos
 - (e) Todas las respuestas son falsas
26. Considera una nueva variable $Z = X - Y$. La media de Z es:
- (a) 4.64
 - (b) 18.32
 - (c) Todas las respuestas son falsas
 - (d) -4.64
 - (e) Es imposible obtener la media de Z con la información disponible
27. Si $\text{cov}(X, Y) = -24.74$, entonces la varianza de Z debe ser:
- (a) 619.7
 - (b) 520.74
 - (c) 534.86
 - (d) -534.86
 - (e) 570.22

FINAL DE UN BLOQUE DE PREGUNTAS

28. Dada la definición de un percentil, q_α , como el valor que deja a su izquierda una proporción α de observaciones, es completamente obvio que q_α :
- (a) No decrece con α
 - (b) Decrece con α
 - (c) Es independiente de α
 - (d) Depende de la desviación típica de la variable
29. Consideremos dos países, A y B, tales que el índice de Gini de la distribución de la renta de A es mayor que el de B. Esto significa que:
- (a) Los ricos son más ricos en A que en B
 - (b) Los pobres son más pobres en A que en B
 - (c) En A los ricos son más ricos y los pobres son más pobres que en B
 - (d) La distribución de la renta es más equitativa en A
 - (e) La distribución de la renta es más equitativa en B

COMIENZO DE UN BLOQUE DE PREGUNTAS

Las preguntas situadas antes de la próxima línea horizontal se refieren a las tablas que aparecen en la siguiente cuestión.

30. Para cierto colectivo de 100 estudiantes se han analizado las variables estadísticas: X = “Tiempo de estudio al día, en horas”, e Y = “Número de problemas no resueltos en cierto examen que consta de dos problemas”. Sabemos que:

X	Y			Total
	0	1	2	
1				60
2			4	
Total	36			

Y	$f_{Y X=1}$
0	0.10
1	0.25
2	0.65

Usando esta información, podemos decir que el número de estudiantes que dejaron un problema sin resolver y estudiaron dos horas al día es:

- (a) 4
 - (b) 6
 - (c) 9
 - (d) Todas las respuestas son falsas
 - (e) 15
31. De los estudiantes que han resuelto dos problemas, el porcentaje que estudió dos horas al día es:
- (a) 30 %
 - (b) 0.06 %
 - (c) 15 %
 - (d) Todas las respuestas son falsas
 - (e) 83 %
32. A la vista de la información que tenemos, podemos decir que las variables X e Y :
- (a) Son independientes
 - (b) No son independientes, tienen una relación lineal inversa
 - (c) No son independientes, tienen una relación lineal directa
 - (d) Todas las respuestas son falsas
 - (e) Son incorrelacionadas

FINAL DE UN BLOQUE DE PREGUNTAS

33. Considera una distribución con $S_x = 10.3$ y coeficiente de variación $g_0 = 1.2$. El valor absoluto de la media, $|\bar{x}|$, debe ser:
- (a) $10.3 \times 1.2 = 12.36$
 - (b) Mayor que 10.3
 - (c) Menor que 10.3
 - (d) No se puede decir

COMIENZO DE UN BLOQUE DE PREGUNTAS

Las preguntas situadas antes de la próxima línea horizontal se refieren a la tabla de contingencia que aparece en la siguiente cuestión.

34. Para valorar la resistencia a una plaga de dos variedades diferentes de maíz (“Maíz1” y “Maíz2”) se han examinado 300 plantas, después de 100 días de siembra, con los resultados mostrados en la tabla siguiente:

Tipo de maíz	Resistencia a la plaga		Total
	Infestado	No infestado	
Maíz1	50	50	100
Maíz2	60	140	200
Total	110	190	300

A partir de la información proporcionada, la distribución de “Infestado” y “No infestado” condicionada a “Maíz1”, en términos de frecuencias relativas, es aproximadamente:

- (a) (0.5, 0.5)
 - (b) Todas las respuestas son falsas
 - (c) (0.4545, 0.5455)
 - (d) (0.3333, 0.6667)
 - (e) (0.3571, 0.6429)
35. La distribución de las variedades de maíz condicionadas a estar infestadas es, en términos de frecuencias relativas, aproximadamente:
- (a) (0.3412, 0.6588)
 - (b) Todas las respuestas son falsas
 - (c) (0.4545, 0.5455)
 - (d) (0.3333, 0.6667)
 - (e) (0.3571, 0.6429)
36. El maíz del tipo “Maíz2” parece más vulnerable a la plaga.
- (a) Verdadero
 - (b) Falso
 - (c) No se puede decir con la información disponible
37. La distribución marginal de frecuencias relativas de Infestado/No infestado es aproximadamente:
- (a) (0.366, 0.633)
 - (b) (0.333, 0.667)
 - (c) (0.357, 0.643)
 - (d) No se puede decir con la información disponible

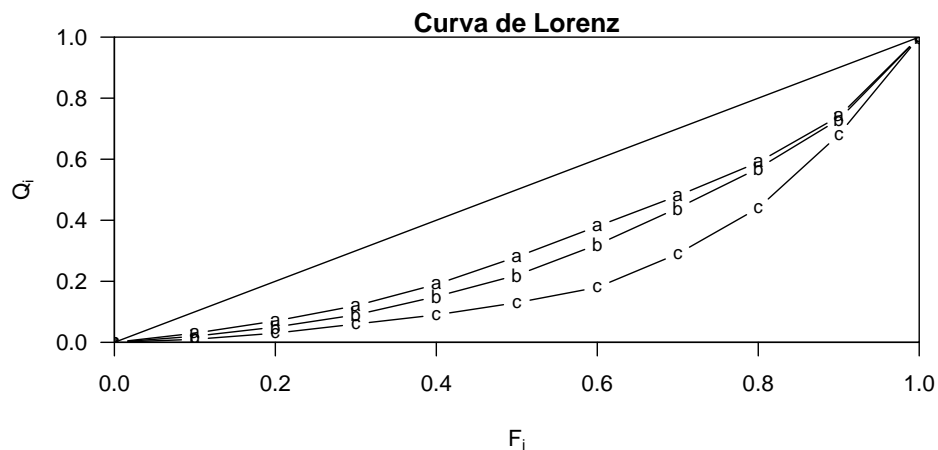
38. La frecuencia relativa conjunta de “Infestado” y “Maíz1” es aproximadamente:
- (a) 0.166
 - (b) 0.4545
 - (c) 0.50
 - (d) No se puede decir con la información disponible
39. De las 300 plantas testadas, un mayor porcentaje de la variedad “Maíz1” resiste a la plaga.
- (a) Falso
 - (b) Verdadero
 - (c) No se puede decir con la información disponible

FINAL DE UN BLOQUE DE PREGUNTAS

COMIENZO DE UN BLOQUE DE PREGUNTAS

Todas las preguntas situadas antes de la próxima línea horizontal se refieren al siguiente gráfico.

40. Considera las siguientes curvas de Lorenz superpuestas en el mismo gráfico, las cuales corresponden a la distribución de la renta en tres países, “a”, “b” y “c”. Llamemos $I_G(a)$, $I_G(b)$ y $I_G(c)$ a sus respectivos índices de Gini.



Debe ocurrir que:

- (a) $I_G(a) \geq I_G(b) \geq I_G(c)$
- (b) $I_G(b) \geq I_G(c) \geq I_G(a)$
- (c) $I_G(c) \geq I_G(b) \geq I_G(a)$
- (d) No se puede decir, no hay información suficiente

41. Mirando las curvas de Lorenz podemos concluir que:

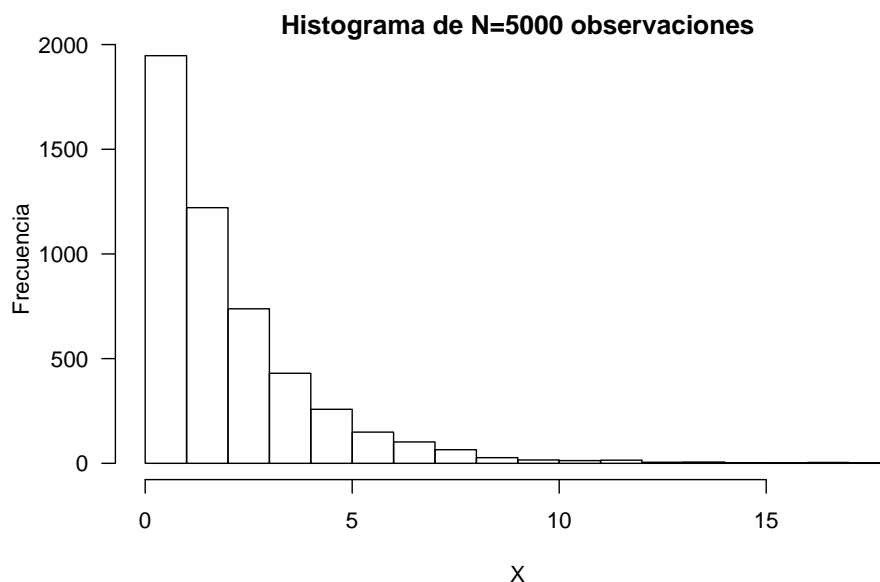
- (a) El 40 % de la población más pobre en “a” tiene una renta media mayor que el 40 % de la población más pobre en “c”
- (b) El 10 % de la población más rica en “a” tiene una renta media mayor que el 10 % de la población más rica en “b”
- (c) En ninguno de los tres países el 40 % de la población más pobre alcanza la proporción del 30 % de la renta total.
- (d) La renta total de “a” es mayor que la renta total de “b”, la cual es además mayor que la renta total en “c”

FINAL DE UN BLOQUE DE PREGUNTAS

COMIENZO DE UN BLOQUE DE PREGUNTAS

Las preguntas situadas antes de la próxima línea horizontal se refieren a la variable X cuyo histograma aparece en la siguiente cuestión.

42. La variable X tiene una distribución de frecuencias representada en el siguiente histograma.



Es claramente un caso de distribución:

- (a) Simétrica
 - (b) Asimétrica negativa
 - (c) Asimétrica positiva
 - (d) Multimodal
43. El porcentaje de observaciones entre 1 y 2 es:
- (a) Entre 20 % y 30 %
 - (b) Inferior al 10 %
 - (c) Mayor que el porcentaje de observaciones entre 0 y 1
 - (d) No se puede decir

44. El momento de tercer orden $m_3 = (1/N) \sum_i (x_i - \bar{x})^3$ será claramente:
- (a) Positivo
 - (b) Negativo
 - (c) Cero
 - (d) No se puede decir
45. Claramente la mediana de X está:
- (a) A la derecha de la moda
 - (b) A la izquierda de la moda
 - (c) Coincide con la moda
 - (d) No se puede decir
46. Date cuenta de que este histograma es de $N = 5000$ observaciones. Simplemente mirando el histograma, uno puede darse cuenta de que el primer cuartil q_1 (o $q_{0.25}$ usando una notación alternativa) es:
- (a) Menor que 1
 - (b) Mayor que 1 pero no mayor que 2
 - (c) Mayor que 5
 - (d) No se puede decir; para responder necesitaríamos los valores de q_2 y q_3 ($q_{0.50}$ y $q_{0.75}$ usando una notación alternativa)
47. En relación al *segundo* cuartil q_2 (o $q_{0.50}$ usando una notación alternativa), podemos afirmar que es:
- (a) Menor que 1
 - (b) Mayor que 1 pero no mayor que 2
 - (c) Mayor que 5
 - (d) No se puede decir; para responder necesitaríamos los valores de q_1 y q_3 ($q_{0.25}$ y $q_{0.75}$ usando una notación alternativa)
48. El *recorrido* de la distribución es:
- (a) Mayor que 10
 - (b) Mayor que 10 pero no mayor que 11
 - (c) Menor que 10
 - (d) No se puede decir
49. El *recorrido intercuartílico* es:
- (a) Mayor que el recorrido
 - (b) Menor que el recorrido
 - (c) Igual al recorrido
 - (d) A la derecha del recorrido
 - (e) No se puede decir

FINAL DE UN BLOQUE DE PREGUNTAS

50. Tienes una población con $N = 1238$ individuos, y recoges valores de una variable estadística X para cada uno de ellos. Existe la sospecha de que una de las observaciones x_i se ha recogido con un error muy grande. Si fuese verdad, ¿qué medida de posición se vería más afectada?
- (a) La media \bar{x}
 - (b) La mediana, M_e
 - (c) La moda, suponiendo que sólo hay una
 - (d) No se puede decir