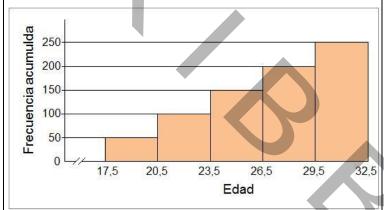
INTRODUCCIÓN AL ANÁLISIS DE DATOS FEBRERO 2011 Código asignatura: 62011037 EXAMEN MODELO C

SOLUCIONES

Situación 1. La gráfica muestra la distribución de la edad (X) de los 250 sujetos de una investigación. En el eje horizontal, se recogen los límites exactos de los intervalos de X y en el eje vertical la frecuencia absoluta acumulada (n_a).



Situación 2. En una investigación para estudiar la relación entre siesta y mejora del aprendizaje, todos los sujetos realizaron una tarea de aprendizaje por la mañana. Después de comer, la mitad de los sujetos se echó la siesta. Finalmente, todos los sujetos volvieron a realizar la misma tarea de aprendizaje por la tarde. Los datos fueron los siguientes:

		Mejora del		
		Sí	No	
Siesta	Sí	170	30	200
	No	70	130	200
		240	160	400

- El número de sujetos de una muestra que realizaron correctamente una tarea de discriminación en un experimento psicofísico es 80, lo que representa el 40% de la muestra. ¿Cuál el número de sujetos de la muestra? A) 200; B) 320; C) 500
- 2. La variable *mejora del aprendizaje* medida en la situación 2 es: A) dicotómica; B) cuantitativa discreta; C) cuasicuantitativa
- En la situación 1, el número de sujetos con edades comprendidas entre 20,5 y 23,5 es A) 22; B) 50;
 C) 100
- 4. En la situación 1, la distribución de la edad de los sujetos: A) no tiene moda; B) tiene una moda; C) tiene dos modas
- 5. En la situación 1, el 80% de los sujetos tiene una edad menor que: A) 26,5; B) 28; C) 29,5
- 6. En la situación 1, la edad media de los sujetos es: A) 25; B) 50; C) 150
- 7. En la situación 1, la varianza de las edades de los sujetos es: A) 18; B) 22; C) 24
- 8. En la situación 1, el rango o amplitud total (A_T) del conjunto de las edades de los sujetos es: A) 3; B) 15; C) 50
- 9. Tenemos 10 puntuaciones cuya media es 15, si sumamos un 5 a cada una de las puntuaciones, la media de las nuevas puntuaciones es: A) 15; B) 20; C) 75
- 10. En la situación 2, el valor del estadístico X² para cuantificar el grado de asociación entre siesta y mejora del aprendizaje está entre: A) 30 y 40; B) 50 y 60; C) 100 y 110
- 11. En la situación 2, el valor del coeficiente de contingencia C para cuantificar el grado de asociación entre siesta y mejora del aprendizaje está entre: A) 0,26 y 0,30; B) 0,33 y 0,36; C) 0,43 y 0,46
- 12. La recta de regresión que permite pronosticar el riesgo de padecer una enfermedad coronaria (Y) en función de la hostilidad (X) es Y' = 1,1 + 0,9X, ¿cuál es el riesgo de padecer una enfermedad

- coronaria de una persona que ha obtenido una puntuación X = 8 en hostilidad?: A) 0,9; B) 1,1; C) 8,3
- 13. La propiedad 0 ≤ P(A) ≤ 1 es válida: A) sólo para la definición clásica de la probabilidad; B) sólo para la definición estadística de la probabilidad; C) para las definiciones clásica y estadística de la probabilidad
- 14. En una clase, la mitad son chicos y la otra mitad son chicas. La mitad de los chicos y la mitad de las chicas han elegido inglés como optativa. Si elegimos una persona al azar de esta clase ¿cuál es la probabilidad de que sea chica y estudie inglés?: A) 0,25; B) 0,50; C) 0,75
- 15. Con los datos del ejercicio 14, ¿podemos decir que los sucesos "ser chica" y "estudiar inglés" son independientes: A) no; B) sí; C) no se puede saber con los datos disponibles
- 16. Una variable aleatoria discreta X toma los valores 0, 1, 2 y 3. Si sabemos que P(X > 2) = 0.125 ¿cuál es la probabilidad de que X sea igual a 3?: A) 0,125; B) 0,25; C) 0,875
- 17. Si lanzamos al aire 10 veces una moneda (no trucada), la varianza de la variable aleatoria "número de caras" es: A) 2,5; B) 4; C) 10
- 18. Lanzamos al aire 10 veces una moneda (no trucada), ¿cuál es la probabilidad de que salgan 4 caras? : A) 0,2051; B) 0,3770; C) 0,5000
- 19. Lanzamos al aire 100 veces una moneda (no trucada), ¿cuál es la probabilidad de que salgan 60 caras?: A) 0,0015; B) 0,0108; C) 0,0225
- 20. En una población de 2500 personas, las puntuaciones de un test de inteligencia siguen una distribución normal con media 100 y desviación típica 15. ¿Cuántas personas tienen en el test una puntuación superior a 130?: A) 15; B) 25; C) 57
- 21. En una distribución F de Snedecor con 7 grados de libertad en el numerador y 30 en el denominador, el percentil 5 es: A) 0,296; B) 2,334; C) 3,376
- 22. La probabilidad de que una variable que sigue una distribución t de Student con 10 grados de libertad tome el valor -0,7 o uno menor es: A) 0,25; B) 0,50; C) 0,75
- 23. Para realizar inferencias sobre el parámetro μ con cierta probabilidad: A) necesitamos el error típico de la media; B) debemos conocer la desviación típica de la población; C) podemos aplicar el muestreo casual
- 24. Sabemos que el error típico de la media vale 1,5 y el tamaño de la muestra es 100, ¿cuál es la desviación típica de la población?: A) 1,5; B) 15; C) 150
- 25. La amplitud deseada de un intervalo de confianza para la media es 4 para un nivel de confianza igual a 0,95, ¿cuál debe ser el tamaño de la muestra si la desviación típica de la población es 10? A) 24; B) 67; C) 96

SOLUCIONES:

1. A

$$n_i = 80$$
 $p_i = 0,40$
 $pi = n_i / n \rightarrow n = n_i / p_i$
 $n = 80 / 0,40 = 200$

2. A

La variable mejora del aprendizaje es una variable cualitativa con dos categorías: sí y no.

3. B

Límites exactos	ni	na
29,5 - 32,5	50	250
26,5 – 29,5	50	200
23,5 – 26,5	50	150
20,5 – 23,5	50	100
17,5 – 20,5	50	50

4. A

La distribución no tiene moda (es amodal) dado que todos los intervalos tienen la misma frecuencia absoluta.

5. C

En la gráfica de la situación 1 se observa que el 80% de los sujetos tiene una edad menor que 29,5. Obtendríamos el mismo resultado aplicando la siguiente fórmula:

Límites exactos	ni	na
29,5 – 32,5	50	250
26,5 – 29,5	50	200
23,5 – 26,5	50	150
20,5 – 23,5	50	100
17,5 - 20,5	50	50

$$P_{80} = L_{i} + \left(\frac{\frac{n \cdot k}{100} - n_{d}}{n_{c}}\right) \cdot I = 26.5 + \left(\frac{\frac{250 \cdot 80}{100} - 150}{50}\right) \cdot 3 = 29.5$$

6. A

X	ni	Xi	n _i X _i
29,5 - 32,5	50	31	1550
26,5 - 29,5	50	28	1400
23,5 - 26,5	50	25	1250
20,5 - 23,5	50	22	1100
17,5 – 20,5	50	19	950
	250		6250

$$\overline{X} = \frac{\sum n_i X_i}{n} = \frac{6250}{250} = 25$$

7. A

X	n _i	X_{i}	$n_i X_i^2$
29,5 – 32,5	50	31	48050
26,5 - 29,5	50	28	39200
23,5 - 26,5	50	25	31250
20,5 - 23,5	50	22	24200
17,5 – 20,5	50	19	18050
_	•	•	160750

$$S_X^2 = \frac{\sum n_i X_i^2}{n} - \overline{X}^2 = \frac{160750}{250} - 25^2 = 18$$

8. B

Para una variable continua, la amplitud del intervalo es la diferencia entre el límite exacto superior y el límite exacto inferior: $A_T = 32,5 - 17,5 = 15$

9. B

$$\overline{X} = 15$$
 $a = 5$ $\rightarrow \overline{Y} = 15 + 5 = 20$

10.C

			Mejora del		
			Sí	No	
		Sí	170	30	200
c	Siesta		(120)	(80)	•
	Oioota	No	70	130	200
			(120)	(80)	
			240	160	400

$$x^{2} = \frac{(170 - 120)^{2}}{120} + \frac{(30 - 80)^{2}}{80} + \frac{(70 - 120)^{2}}{120} + \frac{(130 - 80)^{2}}{80} = 104,17$$
11. C

$$C = \sqrt{\frac{X^2}{X^2 + n}} = \sqrt{\frac{104,17}{104,17 + 400}} = 0,45$$

12. C

$$Y' = 1.1 + 0.9(8) = 8.3$$

13.C

El axioma $0 \le P(A) \le 1$ es válido tanto para las definiciones clásica y estadística.

14.A

$$P(chica) = 0.5$$
 $P(inglés/chica) = 0.5$

P(chica
$$\cap$$
 inglés) = P(chica) P(inglés/chica) = 0,5 · 0,5 = 0,25

15.B

$$P(inglés) = P(chica \cap inglés) + P(chico \cap inglés) = 0.25 + 0.25 = 0.50$$

Los sucesos "ser chica" y "estudiar inglés" son independientes dado que:

P(inglés/chica) = P(inglés)

 $P(chica \cap inglés) = P(chica) P(inglés)$

16.A

$$P(X > 2) = P(X = 3) = 0.125$$

17.A

Binomial con n = 10 y p = 0,5
$$\rightarrow$$
 varianza = npq = 10 · 0,5 · 0,5 = 2,5

18.A

Binomial con
$$n = 10$$
, $p = 0.5$ y $x = 4$ (Tabla I)

19.B

Binomial con n=100, p=0,5 y x=60

Aproximación de la binomial a la normal:

Media = np = 1000,5 = 50 Desviación típica =
$$\sqrt{npq}$$
 = $\sqrt{1000,50,5}$ = $\sqrt{25}$ = 5

$$P\left\lceil \frac{\left(60-0.5\right)-50}{5} \leq Z \leq \frac{\left(60+0.5\right)-50}{5} \right\rceil = P\left(1.9 \leq Z \leq 2.1\right) = 0.9821 - 0.9713 = 0.0108 \quad \text{(Tabla IV)}$$

$$P(X = 60) = P(1.9 \le Z \le 2.1) = 0.0108$$
 por aproximación de la binomial a la normal

20.C

$$z = \frac{\left(130 - 100\right)}{15} = 2$$

$$P(Z > 2) = 1 - P(Z \le 2) = 1 - 0.9772 = 0.0228$$
 (Tala IV)

Nº de personas con inteligencia mayor que 130 → $2500 \cdot 0,0228 = 57$

21. A

$$F_{0.05:7.30} = 1/F_{0.95:30.7} = 1/3,376 = 0,296$$
 (Tabla VII)

22. A

$$P(X \le -0.7) = 1 - P(X \le 0.7) = 1 - 0.75 = 0.25$$
 (Tabla VI con 10 g.l)

23.A

Necesitamos conocer la desviación típica de la distribución muestral de la media para realizar inferencias.

24.B

$$\sigma_{\overline{x}} = 1.5$$
 $n = 100$

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} \quad \to \quad \sigma = \sqrt{n} \cdot \sigma_{\overline{X}}$$

$$\sigma = \sqrt{100} \cdot 1,5 = 10 \cdot 1,5 = 15$$

25. C

n.c. =
$$0.95 \rightarrow z_{1-\alpha/2} = z_{1-0.05/2} = z_{0.975} = 1.96$$
 (Tabla IV)

$$E_{m\acute{a}x}=4/2=2$$

$$n = \frac{z_{_{1-\alpha/2}}^2 \sigma^2}{E_{m\acute{a}x}^2} = \frac{1,96^2.10^2}{2^2} = 96,04 \rightarrow 96$$

