Código asignatura	Nombre asignatura
62011037	**INTRODUCCIÓN AL ANÁLISIS DE DATOS**
Fecha alta y origen	Convocatoria
15/02/2022 **Curso virtual**	**FEBRERO 2022 – TIPO B**

INTRODUCCIÓN AL ANÁLISIS DE DATOS
FEBRERO 2022 Código asignatura: 62011037
EXAMEN TIPO TEST MODELO B DURACION: 2 HORAS

Material: **Addenda (Formulario y Tablas) y calculadora (no programable)**

Calificación= (0,4 x Aciertos) - (0,2 x Errores)

No debe entregar los enunciados

Gráfico 1. Resultados obtenidos por un grupo de 50 varones en una prueba de orientación espacial (X).

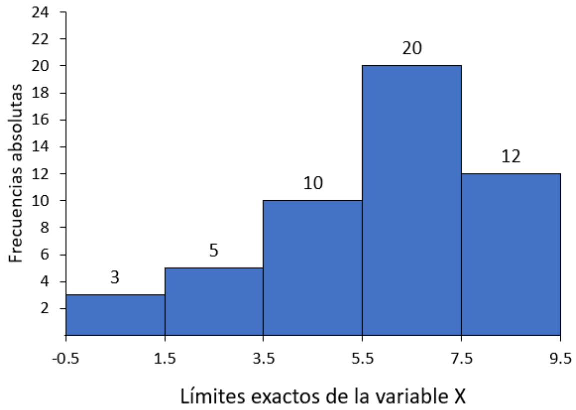

Tabla 1. Datos de 5 alumnos de Secundaria en un estudio para pronosticar la nota en Matemáticas (Y) a partir de las puntuaciones en un test de aptitud numérica (X).

Alumno	**Test Aptitud Numérica (X)**	**Nota en Matemáticas (Y)**
1	30	4
2	40	5
3	80	9
4	60	7
5	50	6

Tabla 2. Datos de 100 mujeres empresarias en un estudio para pronosticar el éxito profesional (Y) a partir de las puntuaciones en un test de habilidades sociales (X).

	Media	Desviación típica	Pendiente de la recta de regresión de Y sobre X
X	7	2	b= 1,35
Y	8	3	

Tabla 3. Variables y características de su distribución.

Variable	**Distribución**	
X	N(40,2)	Normal con media 40 y desviación típica 2
Y	t_{18}	t de Student con 18 grados de libertad
V	$F_{20,6}$	F con 20 grados de libertad en el numerador y 6 grados de libertad en el denominador

1. La moda es: A) uno de los índices estadísticos apropiados para las variables cualitativas; **B) el único índice apropiado para variables cualitativas**; C) un índice de dispersión.
2. La variable X del Gráfico 1 sería: A) de razón; **B) de intervalos**; C) discreta.
3. Con los datos Gráfico 1, la moda para los varones es: A) 20; B) 2,5; **C) 6,5**.
4. Siguiendo el Gráfico 1, si un varón obtiene una puntuación de 7 en orientación espacial, el percentil que le corresponde es: **A) 66**; B) 54; C) 32.
5. Con los datos del Gráfico 1, el Índice de Asimetría de Pearson calculado para la distribución de varones es, aproximadamente: A) -0,545; B) 0,302; **C) -0,302.**
6. La puntuación típica para un varón que ha obtenido una puntuación directa en orientación espacial de 8, con los datos del Gráfico 1 es: **A) 0,97**; B) 1,95; C) -1,95.
7. El estudio de la relación entre dos variables cualitativas se puede llevar a cabo observando: A) las frecuencias marginales de cada una de las dos variables por separado; B) las frecuencias conjuntas de ambas variables; **C) las distribuciones condicionadas de una de las dos variables agrupadas en función de los valores de la otra.**
8. El coeficiente de correlación por rangos de Spearman: A) es apropiado para variables nominales; B) es apropiado para variables cuantitativas que tengan una distribución normal; **C) oscila entre -1 y +1.**

9. Según los datos de la Tabla 1, el valor de la asociación entre las variables, evaluada a través del coeficiente de correlación de Spearman, es: A) 0; **B) 1**; C) 0,95.
10. Según la Tabla 2, la covarianza entre las variables X e Y es**: A) 5,4**; B) 0,90; C) 4,5.
11. Con los datos de la Tabla 2, el coeficiente de correlación de Pearson entre X e Y es: A) 0,85; **B) 0,90**; C) 0,95.
12. Siguiendo los datos de la Tabla 2, ¿qué puntuación en éxito profesional le pronosticaremos a una mujer empresaria que ha obtenido una puntuación de 4 en habilidades sociales? A) 4,55; B) 2,45; **C) 3,95**.
13. Si P(A U B) = P(A) + P(B) diremos que los sucesos A y B son: **A) mutuamente excluyentes**; B) compatibles; **C) dependientes.** * Dado que las alternativas A y C son correctas, esta pregunta se dará por válida a todos los estudiantes, la hayan respondido o no.
14. Tráfico ha estimado que la probabilidad de multa por “exceso de velocidad”, P(ev), es 0,50, por “saltarse un semáforo”, P(ss), es de 0,10 y por el “resto de infracciones”, P(r), es de 0,40. También se conoce que la probabilidad de “accidente por exceso de velocidad” es de 0,80; la de “accidente por saltarse un semáforo” es de 0,70 y la probabilidad de “accidente por otras infracciones” es de 0,30. La probabilidad total de tener un accidente por las tres causas es: A) 0,40; **B) 0,59**; C) 0,12.
15. Siguiendo con la pregunta anterior, ha ocurrido un accidente, la probabilidad de que haya sido por saltarse un semáforo es: **A) 0,1186**; B) 0,8814; C) 0,6780.
16. En Psicología de la Salud, se entiende por “prevalencia”: **A) la proporción de personas que tienen una enfermedad en un momento determinado respecto de la población**; B) la proporción de casos nuevos de una enfermedad en una población, durante un período determinado; C) la “incidencia” del trastorno.
17. Si la función de probabilidad de una variable discreta X es f(1) = 0,25; f(2) = 0,35; f(3) = 0,20; f(4) = 0,20. La media de X es: A) 4,22; B) 6,45; **C) 2,35**.
18. Se sabe que el 20% de la población de estudiantes sufre ansiedad previa ante los exámenes. Si elegimos al azar una muestra de 8 estudiantes de un centro, ¿cuál es la probabilidad de que más de cuatro estudiantes sufran de ansiedad previa a los exámenes? A) 0,9896; B) 0,0459; **C) 0,0104**.
19. Con los datos de la Tabla 3, la probabilidad de que X sea menor o igual que 42 es aproximadamente**: A) 0,8413**; B) 0,1587; C) 0,9772.
20. Teniendo en cuenta la Tabla 3, para la variable Y el valor de probabilidad 0,688 corresponde al percentil: A) 25; **B) 75**; C) 50. *Dado que hay una errata en el enunciado (sobra la palabra probabilidad), esta pregunta se dará por válida a todos los estudiantes, la hayan respondido o no.
21. Atendiendo a los datos de la Tabla 3, ¿cuál es el percentil 10 de la variable V? A) 2,091; B) 0,353; **C) 0,478.**
22. ¿Cuál de estas afirmaciones es la correcta? A) La media muestral es un estimador insesgado y consistente, pero no suficiente de la media poblacional; B) La proporción muestral no se considera un buen estimador de la proporción poblacional; **C) La cuasivarianza muestral es un estimador insesgado de la varianza poblacional.**
23. Se sabe que el 22% de la población adulta española fuma. Si se selecciona una muestra aleatoria de 250 adultos españoles, ¿cuál es la probabilidad de que menos del 20% de la muestra sean fumadores? **A) 0,2005**; B) 0,7995; C) 0,2018.
24. En una muestra seleccionada al azar de 400 adolescentes se observó que la mitad jugaba a diario con consolas de videojuegos. ¿Cuál es el intervalo de confianza al 99% relativo a la proporción de adolescentes que juegan a diario con este tipo de consolas? A) 0,36 y 0,48; B) 0,44 y 0,65; **C) 0,44 y 0,56**.

25. Sabiendo que P= 0,20, ¿cuál es el tamaño muestral necesario para que el $E_{máx}$ asociado al intervalo de confianza al 95% sea 0,08? Se asume población infinita. **A) 96**; B) 45; C) 84.

SOLUCIONES:

1. B (p.75)
2. B
3. A

   Mo (varones) = 6,5 (media del intervalo más frecuente, que es (6-7), con n_i = 20).

4. A (p. 83)

X	n_i	n_a
8-9	12	50
6-7	20	38
4-5	10	18
2-3	5	8
0-1	3	3
Σ	50	

Si X_i= 7, entonces el IC es (6-7) y la I= 7,5 – 5,5 = 2.

$$k = \left[\frac{\frac{(P_k - L_i)n_i}{I} + n_d}{n}\right]100 = \left[\frac{\frac{(7-5,5)20}{2} + 18}{50}\right]100 = 66$$

5. C (p. 117)

X	X_i	n (varones)	nX_i	X^2_i	nX^2_i
8-9	8,5	12	102	72,25	867
6-7	6,5	20	130	42,25	845
4-5	4,5	10	45	20,25	202,5
2-3	2,5	5	12,5	6,25	31,25
0-1	0,5	3	1,5	0,25	0,75
Σ		50	291		1946,5

$\bar{X} = \frac{\sum n_i X_i}{\sum n_i} = \frac{\sum n_i X_i}{n}$ = 291/50= 5,82

$S_x^2 = \frac{\sum n_i X_i^2}{\sum n_i} - \bar{X}^2 = \frac{\sum n_i X_i^2}{n} - \bar{X}^2$ = 1946,5/50 – $(5,82)^2$= 5,0576; S_x= $\sqrt{5,0576}$= 2,25

$$Ap = \frac{\bar{X} - Mo}{Sx} = \frac{5,82 - 6,5}{2,25} = -0,302$$

6. A (p. 128)

$$z = \frac{X - \bar{X}}{S_X} = \frac{8 - 5{,}82}{2{,}25} = 0{,}97$$

7. C

8. C (170, 171)

9. B

Rangos(X)	Rangos (Y)	d	d_i^2
1	1	0	0
2	2	0	0
5	5	0	0
4	4	0	0
3	3	0	0

Spearman (p. 171) → $r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$ = 1 - 0 = 1

Existe una correlación perfecta entre ambas variables. A mayor aptitud numérica, mayor nota en Matemáticas.

10. A (pp. 197, 198, 220)

$r_{xy} = b\frac{S_x}{S_y}$; $r_{xy} = 1{,}35\frac{2}{3} = 0{,}90$

$$r_{xy} = \frac{S_{xy}}{S_x S_y}$$; $S_{xy} = r_{xy} S_x S_y = 0{,}90 \times 2 \times 3 = 5{,}4$

11. B

12. C (p. 221)

a= $\bar{Y}$– b($\bar{X}$)= 8- 1,35(7)= -1,45

Y´= a + bX= -1,45+ 1,35(4)= 3,95

.

13. A y C (p. 265) * Dado que las alternativas A y C son correctas, esta pregunta se dará por válida a todos los estudiantes, la hayan respondido o no

14. B

Teorema de la Probabilidad Total: $P(B) = \sum_{i=1}^{k} P(A_i) \cdot P(B|A_i)$

P(ev)= 0,50; P(A|ev)= 0,80

P(ss)= 0,10; P(A|ss)= 0,70

P(r)= 0,40; P(A|r)= 0,30

P(A)= 0,50 x 0,80 + 0,10 x 0,70 + 0,40 x 0,30= 0,59

15. A

Teorema de Bayes: $P(A_i|B) = \frac{P(A_i) \cdot P(B|A_i)}{P(B)}$

P(ss|A)= (P(ss)P(A|ss))/P(A)= (0,10 x 0,70)/0,59= 0,1186

16. A
17. C

f(1)= 0,25; f(2)= 0,35; f(3)= 0,20; f(4)= 0,20

μ = E(X) = ∑xf(x) = 1(0,25) + 2(0,35) + 3(0,20) + 4(0,20)= 2,35

18. C

Tabla II de la Binomial con valores: p= 0,20; n=8.
$P(X> 4)= 1 – P(X\leq 4) = 1 - F(4) = 1- 0,9896= 0,0104$

19. A

Z= (42-40)/2= 1
$P(X\leq 42)= P(Z\leq 1)= 0,8413$ (Tabla IV)

20. B *Dado que hay una errata en el enunciado (sobra la palabra probabilidad), esta pregunta se dará por válida a todos los estudiantes, la hayan respondido o no

Tabla VI (*t* de Student)
Para gl=18, el valor 0,688 se corresponde con p=0,75
P(Y≤ 0,688) = 0,75. Por tanto, Y= 0,688 es el percentil 75.

21. C

$_{0,10}F_{20,6}=1 / {}_{0,90}F_{6,20}= 1/ 2,091= 0,478$

22. C

23. A

El 20% de 250 es 50 [(20 x 250)/100], por lo que en esta pregunta hay que calcular la probabilidad de que menos de 50 personas sean fumadoras: P(X < 50)

Como el tamaño de la muestra es grande (n > 20), podemos utilizar la aproximación de la binomial a la distribución normal para solucionar este ejercicio. Para ello, primero necesitamos hallar la media, μ, y la desviación típica, σ: (ver fórmula en la página 353 del manual):

μ = *np* = 250 x 0,22 = 55

$\sigma = \sqrt{npq} = \sqrt{250 \ x \ 0,22 \ x \ 0,78} = 6,55$

Para hallar la probabilidad que se pide aproximando la distribución binominal a la distribución normal, tenemos que aplicar una pequeña corrección, restando 0,5 a nuestro valor: P(X < 49,5)

Ahora se procede a calcular la puntuación Z correspondiente, utilizando para ello la media y desviación típica que hemos calculado previamente:

P(X < 49,5) = P(Z ≤ [(49,5 - 55)/6,55]) = P(Z ≤ -0,84)

En la Tabla III podemos ver que la probabilidad que queda por debajo de una puntuación Z de -0,84 es 0,2005: P(X < 49,5) = P(Z ≤ [(49,5-55)/6,55]) = P(Z ≤ -0,84) = 0,2005

24. C
Intervalo de confianza la para el parámetro π (aproximación a la Normal); α= 0,01

P= 200/400= 0,5; $z\alpha/_2 = z_{0,005} = -2{,}58$ y $z_{1-\alpha/_2} = z_{0{,}995}$= 2,58

L_i= P- $z_{1-\alpha/2}\sqrt{\frac{P(1-P)}{n}}$= 0,5-2,58 $\sqrt{\frac{0{,}5(0{,}5)}{400}}$= 0,44

L_s= P+ $z_{1-\alpha/2}\sqrt{\frac{P(1-P)}{n}}$= 0,5+2,58 $\sqrt{\frac{0{,}5(0{,}5)}{400}}$= 0,56

25. A (p. 465)
Tamaño muestral para π ; $Z_{0,025}$= -1,96

$$n= \frac{Z^2_{\alpha/_2} P(1-P)}{E^2_{max}} = \frac{(-1{,}96)^2(0{,}20)(0{,}80)}{(0{,}08)^2} \approx 96$$