INTRODUCCIÓN AL ANÁLISIS DE DATOS
FEBRERO 2018 Código asignatura: 62011037
EXAMEN TIPO TEST MODELO A DURACION: 2 HORAS

Material: **Addenda (Formulario y Tablas) y calculadora (cualquier modelo)**

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

No debe entregar los enunciados

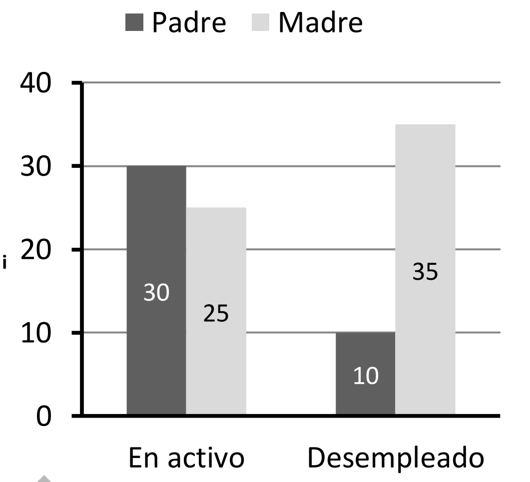
Tabla 1. Resultados obtenidos en un test de creatividad infantil (X) en una clase de 30 alumnos de primaria.

X	Alumnos
20-22	5
17-19	9
14-16	7
11-13	5
8-10	4

Tabla 2. Se realiza un estudio sobre percepción emocional y ajuste psicosocial en 1650 adolescentes.
En la siguiente tabla se ofrece información descriptiva de ambas variables y su relación.

	Media	Desviación típica	Correlación
Percepción emocional (X)	15,20	3,94	$r_{XY} = 0{,}34$
Ajuste psicosocial (Y)	18,70	1,6	

Gráfica 1. Información obtenida en un estudio sobre la situación laboral de padres (n = 40) y madres (n = 60) que tienen hijos con autismo.



1. En la Estadística Inferencial, el *parámetro* es una propiedad descriptiva de: A) una muestra; B) **una población**; C) un estadístico.
2. Para los datos recogidos en la Tabla 1 la representación gráfica más habitual es un: A) **histograma**; B) diagrama de dispersión; C) diagrama de sectores.
3. Con los datos de la Tabla 1, la media en creatividad infantil para los alumnos de la clase es: A) 11,5; B) 16,2; C) **15,6.**
4. Siguiendo con la información recogida en la Tabla 1, ¿qué percentil corresponde a un alumno que tiene una puntuación de 18? A) P_{80}; B) P_{90}; C) $\mathbf{P_{68}}$.
5. El valor de la varianza de la variable creatividad, cuyos datos se representan en la Tabla 1 se encuentra entre: A) 10 y 12; B) **14 y 16**; C) 18 y 20.
6. Siguiendo con los datos de la Tabla 1, la distribución es: A) **asimétrica negativa**; B) simétrica; C) asimétrica positiva.
7. Según los datos de la Gráfica 1, la distribución de la situación laboral condicionada a ser madre es: A) 75% y 41,75%; B) **41,67% y 58,3%;** C) 25% y 58,3%.
8. Con los datos de la Gráfica 1, el valor de la asociación entre las variables, evaluada a través de χ^2, está comprendido entre: A) **9 y 12**; B) 13 y 16; C) 18 y 20.
9. Para que exista independencia entre dos variables cualitativas el índice χ^2 debe ser: A) mayor que 0; B) **igual a 0**; C) χ^2 no evalúa si existe o no independencia.
10. En el estudio sobre percepción emocional y ajuste psicosocial en adolescentes, cuya información aparece en la Tabla 2, la covarianza entre ambas variables es: A) 1,24; B) **2,14**; C) 4,12.

11. Con los datos de la Tabla 2, la ordenada en el origen y la pendiente de la ecuación de la recta de regresión para pronosticar el ajuste psicosocial (Y) a partir de la percepción emocional (X) es: A) 12,47 y 3,25; B) 10,33 y 3,94; C) **16,57 y 0,14.**
12. Con la recta de regresión obtenida en la pregunta anterior, ¿qué puntuación en ajuste psicosocial pronosticaremos a un adolescente que ha obtenido una puntuación de 12 en percepción emocional?: A) 16,27; B) 18,60; C) **18,25**.
13. Indica qué afirmación NO es correcta sobre el teorema del producto: A) se aplica a situaciones en las que se quiere calcular la probabilidad de que aparezcan dos sucesos de forma simultánea; B) sirve para calcular la probabilidad de intersección entre dos sucesos; C) **sirve para calcular la probabilidad de unión entre dos sucesos.**
14. En una escuela infantil 1/3 de los niños de 5 años está aprendiendo a leer con un método de lectura global (G) y el resto de niños con lectura fonológica y silábica (FS). Tras la aplicación de los dos métodos a final de curso, presenta una mejoría (M) significativa en la lectura el 70% de los niños que han recibido metodología (G) y el 85% de los niños con metodología (FS). Se elige un niño al azar y se observa que ha presentado mejoría (M), ¿cuál es la probabilidad de que haya recibido el método de lectura global (G)? A) **0,29**; B) 0,33; C) 0,70.
15. Se ha diseñado una nueva prueba para diagnosticar a personas con sordera moderada. Para ello, se ha pasado la prueba a 150 personas ya diagnosticadas previamente con ese tipo de sordera. Al aplicar la nueva prueba 25 personas han dado negativo en el diagnóstico. ¿Cuál es la sensibilidad de la prueba? A) **0,8333**; B) 0,7990; C) 0,8990.
16. La función de probabilidad de una variable *X* es: *f* (1) = 0,30; *f* (2) = 0,45 y *f* (3) = 0,25. ¿Entre qué valores se encuentra la varianza? A) 2 y 1,58; B) 2,50 y 1; C) **0,80 y 0,30.**
17. Se sabe que el 15% de la población de alumnos de un instituto tiene dificultades de aprendizaje. Si elegimos al azar una muestra de 10 estudiantes del centro, ¿cuál es la probabilidad de que tres estudiantes tengan dificultades de aprendizaje? A) 0,9500**;** B) **0,1298**; C) 0,5443.
18. Sabiendo que las puntuaciones en el test de percepción emocional de la Tabla 2 se distribuyen normalmente, ¿cuál es la proporción de sujetos con puntuaciones entre 13 y 17? A) **0,3895**; B) 0,6772; C) 0,2877.
19. Sea X una variable que se distribuye según una distribución *t* con 12 grados de libertad. ¿Cuál es la probabilidad de obtener valores menores o iguales a 2,681? A) 0,950; B) **0,990**; C) 0,995.
20. ¿Qué tipo de muestreo considera que todas las muestras posibles son equiprobables? A) Por cuotas; B) Intencional; C) **Aleatorio simple.**
21. En una población formada por los números del 1 al 10, ¿qué probabilidad hay de obtener una muestra de tamaño 3 en un muestreo aleatorio con reposición? A) **0,001**; B) 0,010; C) 0,030.
22. Sea X una variable con distribución normal en la población con N(μ; 4). Se extraen muestras de tamaño 15, ¿entre que valores de χ^2 se encuentra el 80% central de las varianzas muestrales? A) **(7,7895 y 21,0641)**; B) (5,6287 y 21,0641); C) (4,6604 y 23,6848).
23. La variable autoestima en la población de adolescentes se distribuye normalmente, con una desviación típica de 3,5. Se quiere estimar la media en la población y para ello se ha seleccionado mediante m. a. s. una muestra de 300 adolescentes, obteniéndose una media para la muestra de 18. Trabajando con un nivel de confianza del 95%, ¿cuáles son los límites del intervalo de confianza? A) (15,25 y 19,46); B) **(17,60 y 18,40)**; C) (17,33 y 19,14).
24. Un investigador trabaja con un nivel de confianza del 95%. Dada una variable X, se conocen los límites inferior y superior (5,70 y 8,30) del intervalo confidencial de la media poblacional. Se ha seleccionado una muestra de 130 personas, obteniéndose una media muestral de 7. ¿Cuál es la desviación típica de la variable X en la población? A**) 7,56;** B) 8,03; C) 7,05.
25. Se quiere estimar con un nivel de confianza del 99% y con un error máximo de estimación de 0,1 la proporción de niños de Educación Secundaria que tienen ordenador en casa. Estudios previos señalan que esa proporción es del 0,3. ¿Cuál es el tamaño de la muestra necesario para realizar dicha estimación asumiendo una población infinita? A) 100; B) 120; C) **140.**

SOLUCIONES:

1. B
2. A
3. C

X	X_i	n_i	$n_i\ X_i$
20-22	21	5	105
17-19	18	9	162
14-16	15	7	105
11-13	12	5	60
8-10	9	4	36
$\sum$		30	468

$$\overline{X} = \frac{\sum n_i X_i}{n} = \frac{468}{30} = 15{,}6$$

4. C

X	X_i	n_i	n_a
20-22	21	5	30
17-19	18	9	25
14-16	15	7	16
11-13	12	5	9
8-10	9	4	4
$\sum$		30	

La puntuación X = 18 está en el intervalo [17-19] que va a ser, por tanto, el intervalo crítico.
Por lo tanto, P_k = 18; L_i = 16,5; n_c = 9, I = 3 y n_d = 16. Se aplica la fórmula y se obtiene lo siguiente:

$$k = \left[\frac{\dfrac{(P_k - L_i)\cdot n_c}{I} + n_d}{n}\right]\cdot 100 = \left[\frac{\dfrac{(18-16{,}5)\times 9}{3} + 16}{30}\right]\times 100 = \left[\frac{20{,}5}{30}\right]\times 100 =$$

$$= 0{,}683 \times 100 = 68{,}33 \approx 68$$

5. B

De la pregunta 3 sabemos que la media es 15,6

X	X_i	n_i	$(X_i - \overline{X})$	$(X_i - \overline{X})^2$	$n_i(X_i - \overline{X})^2$
20-22	21	5	5,4	29,16	145,80
17-19	18	9	2,4	5,76	51,84
14-16	15	7	-0,6	0,36	2,52
11-13	12	5	-3,6	12,96	64,80
8-10	9	4	-6,6	43,56	174,24
$\sum$		30			439,20

$$S_X^2 = \frac{\sum n_i (X_i - \overline{X})^2}{n} = \frac{439,2}{30} = 14,64$$

6. A

Conocemos los siguientes datos:

$\overline{X} = 15,6\,;\ Mo = 18,$ ya que es el punto medio del intervalo con mayor frecuencia absoluta [17-19]

$$S_X = \sqrt{14,64} = 3,826$$

$$A_P = \frac{\overline{X} - Mo}{S_X} = \frac{15,6 - 18}{3,826} = -0,627$$

Como A_p es negativo, la distribución presenta una asimetría negativa.

7. B

Se trata de una distribución de Y (situación laboral) condicionada a uno de los valores de X (madre). Por tanto, se considera únicamente la fila en la que aparecen las frecuencias de las madres. Para calcular estos porcentajes, cada frecuencia de la fila de las madres se divide por el total de madres (60) y se multiplica por 100.

	En activo	Desempleado	
Padre	30	10	40
Madre	25 (41,67%)	35 (58,33%)	60
	55	45	Total 100

8. A

Primero hay que calcular las frecuencias teóricas (entre paréntesis en la tabla):

$$n_{t11} = \frac{40 \times 55}{100} = 22 \qquad n_{t12} = \frac{40 \times 45}{100} = 18$$

$$n_{t21} = \frac{60 \times 55}{100} = 33 \qquad n_{t22} = \frac{60 \times 45}{100} = 27$$

	Empleado	Desempleado	
Padre	30 (22)	10 (18)	40
Madre	25 (33)	35 (27)	60
	55	45	Total 100

$$\chi^2 = \frac{(30-22)^2}{22} + \frac{(10-18)^2}{18} + \frac{(25-33)^2}{33} + \frac{(35-27)^2}{27} = 2,91 + 3,56 + 1,94 + 2,37 = 10,78$$

9. B

10. B

$$r_{XY} = \frac{S_{XY}}{S_X S_Y} \rightarrow r_{XY} \times S_X S_Y = S_{XY} \rightarrow S_{XY} = 0{,}34 \times 3{,}94 \times 1{,}60 = 2{,}14$$

11. C

$$b = r_{XY} \frac{S_Y}{S_X} = 0{,}34 \times \frac{1{,}60}{3{,}94} = 0{,}1380 \approx 0{,}14$$

$$a = \overline{Y} - b\overline{X} = 18{,}70 - 0{,}14 \times 15{,}20 = 16{,}572 \approx 16{,}57$$

12. C

$$Y_i' = a + bX_i = 16{,}57 + 0{,}14 \times 12 = 16{,}57 + 1{,}68 = 18{,}25$$

13. C (Página 269)

14. A

La probabilidad de que el niño escogido que presenta mejoría en la lectura (*M)*, haya recibido la intervención en la lectura con el método (G) se obtiene mediante el teorema de Bayes.

Lo que se pide es:

$$P(G|M) = \frac{P(G) \cdot P(M|G)}{P(M)}$$

$$P(G) = 0{,}33$$

$$P(FS) = 0{,}67$$

$$P(M|G) = 0{,}70$$

$$P(M|FS) = 0{,}85$$

$$P(G/M) = \frac{P(G) \cdot P(M|G)}{P(G) \cdot P(M|G) + P(FS) \cdot P(M/FS)} = \frac{0{,}33 \times 0{,}70}{0{,}33 \times 0{,}70 + 0{,}67 \times 0{,}85} =$$

$$= \frac{0{,}231}{0{,}8} = 0{,}29$$

15. A

Sensibilidad = $P(+|T)$ Es decir, discrimina a los verdaderos positivos. De las 150 personas con sordera moderada ha detectado a 125. Por tanto:

$$P(+|T) = \frac{125}{150} = 0{,}8333$$

16. A

x	f(x)	$x \cdot f(x)$	$(x-\mu)$	$(x-\mu)^2$	$(x-\mu)^2 \cdot f(x)$
1	0,30	0,30	-0,95	0,9025	0,2708
2	0,45	0,90	0,05	0,0025	0,0011
3	0,25	0,75	1,05	1,1025	0,2756
		1,95			0,5475

$$\sigma^2 = \sum (x-\mu)^2 \cdot f(x) = 0{,}5475 \approx 0{,}55$$

$$\mu = \sum x \cdot f(x) = 1{,}95$$

17. B

Debemos calcular $P(X = 3)$ con n = 10 y p = 0,15; x = 3

Se resuelve por la función de probabilidad binomial, por lo que el valor se localiza en la Tabla I → 0,1298

18. A

Se pide:

$P(13 \leq X \leq 17) = P(X \leq 17) - P(X \leq 13)$

Para calcular estas proporciones, transformamos las puntuaciones directas a puntuaciones típicas y buscamos en las tablas de la distribución normal (Tablas III y IV).

$$z = \frac{X - \overline{X}}{S_X} = \frac{13 - 15,20}{3,94} = -0,558 \approx -0,56\,;\; P(z \leq -0,56) = 0,2877$$

$$z = \frac{X - \overline{X}}{S_X} = \frac{17 - 15,20}{3,94} = 0,4568 \approx 0,46\,;\; P(z \leq 0,46) = 0,6772$$

$$P(13 \leq X \leq 17) = P(X \leq 17) - P(X \leq 13) = 0,6772 - 0,2877 = 0,3895$$

19. B $t_{12} \leq 2,681) = 0,99$ (ver Tabla VI)

20. C (Página 379)

21. A

La probabilidad de una muestra concreta de tamaño 3 es: $\frac{1}{N^n} = \frac{1}{10^3} = 0,001$

22. A

Se trata de la distribución muestral de la varianza. Sabemos que dicha distribución es χ^2_{n-1}

$P(\chi^2 \leq \chi^2_{n-1}) = 0,10$

Acudiendo a la Tabla V con 14 grados de libertad, los valores de χ^2 correspondientes son:

$\chi^2_{14;0,10} = 7,7895$

$\mathrm{X}^2_{14;0,90} = 21,0641$

23. B

Variable con distribución normal y varianza poblacional conocida, por tanto, la distribución muestral de la media es: $N\left(\mu;\, \sigma/_{\sqrt{n}}\right)$

El intervalo confidencial viene definido por:

$$\boxed{P\left(\overline{X} - \left|Z\alpha/_2\right| \sigma/_{\sqrt{n}} \leq \mu \leq \overline{X} + \left|Z\alpha/_2\right| \sigma/_{\sqrt{n}}\right) = 1 - \alpha}$$

Siendo los límites:

$L_{inf} = \overline{X} - \left|Z\alpha/_2\right| \sigma/_{\sqrt{n}}$	$L_{sup} = \overline{X} + \left|Z\alpha/_2\right| \sigma/_{\sqrt{n}}$

Sustituyendo:

n.c. = $1 - \alpha$ = 0,95 → $\alpha = 0,05$ → $\left|\alpha/_2\right| = |1,96|$ (Tabla III y IV)

$$L_{inf} = 18 - 1{,}96\frac{3{,}5}{\sqrt{300}} = 18 - 0{,}40 = 17{,}60$$
$$L_{sup} = 18 + 1{,}96\frac{3{,}5}{\sqrt{300}} = 18 + 0{,}40 = 18{,}40$$

Los límites buscados son: (17,60; 18,40)

24. A

n.c. $= 0{,}95 \rightarrow Z\alpha_{/2} = Z_{0{,}975} = 1{,}96$ (Tabla III y IV)

$$L_{sup} = \bar{X} + \left|Z\alpha_{/2}\right| {}^{\sigma}/_{\sqrt{n}} \rightarrow 8{,}30 = 7 + 1{,}96 \times {}^{\sigma}/_{\sqrt{130}} \rightarrow 1{,}30 = 1{,}96 \times {}^{\sigma}/_{11{,}40}$$

$$\sigma = \frac{1{,}30 \times 11{,}40}{1{,}96} = \frac{14{,}82}{1{,}96} = 7{,}56$$

o, utilizando el límite inferior:

$$L_{inf} = \bar{X} - \left|Z\alpha_{/2}\right| {}^{\sigma}/_{\sqrt{n}} \rightarrow 5{,}70 = 7 - 1{,}96 \times {}^{\sigma}/_{\sqrt{130}} \rightarrow -1{,}30 = -1{,}96 \times {}^{\sigma}/_{11{,}40}$$

$$\sigma = \frac{1{,}30 \times 11{,}40}{1{,}96} = \frac{14{,}82}{1{,}96} = 7{,}56$$

25. C

Sabemos que p = 0,3, por lo que: (1- p) = 0,7.

$$n = \frac{Z^2_{\alpha/2} P(1-P)}{E^2_{max}} = \frac{2{,}58^2 \times 0{,}3 \times 0{,}7}{0{,}1^2} = \frac{1{,}398}{0{,}01} = 139{,}8 \approx 140$$