EVALUACIONES DE SERIES TEMPORALES

ÍNDICE TEMÁTICO

1. DESCOMPOSICIÓN CLÁSICA

2. MODELIZACIÓN CON VARIABLES CATEGÓRICAS

3. AUTOCORRELACIÓN

4. SUAVIZADO EXPONENCIAL

1 EVALUACIONES PROPUESTAS

Respuesta correcta +1; incorrecta –0,2

- **1.** El modelo de tendencia ha sido T = 76,23 + 0,54 t – 0,02 t^2. Los respectivos niveles de significación de los términos t y t^2 han sido 0,002 y 0,423. El modelo definitivo es

  76,23+0,54t–0,02t^2 ☐ 76,23+0,54t ☐ Hay que recalcularlo ☐ ………………. ☐

13.5.98

- **1** Los valores disponibles de una serie temporal son: 11,2; 13,4; 9,9; 11,9; 14,2; 11,0; 13,1; 14,8; 12,2; 14,1; 16,3; .... Se trata de un modelo:

multiplicativo	☐	tendencia rectilínea	☐	estacionalidad de p=2	☐
aditivo	☐	tendencia parabólica	☐	estacionalidad de p=3	☐
. . . .	☐	ninguna tendencia	☐	estacionalidad de p=4	☐
		. . . .	☐	. . . .	☐

- **2** Por los datos anteriores, el valor de la tercera media móvil es:
  11,75 ☐ 12 ☐ 12,15 ☐ 13,36 ☐ . . . . ☐

- **3** Los primeros datos de una serie multiplicativa p = 4 son: 32; 26; 22; 45; 52; 42; 29; ... El valor de la media móvil asociada a t = 4 es:
  31,25 ☐ 36,25 ☐ 38,25 ☐ 40,25 ☐ . . . . ☐

- **4** En una serie multiplicativa de p = 4, $E_1^* = 43.4$ $E_2^* = 37.9$ $E_3^* = 52.5$ $E_4^* = 66.2$; ¿cuál es el valor de E_3?
  2.5 ☐ 44.6 ☐ 52.5 ☐ 105 ☐ . . . . ☐

- **5.** Sobre 106 valores, la tendencia estimada es 254,9 + 0,25 t ; los índices estacionales son E_1 = 35,5; E_2 = 72,8; E_3 = –60,7 y E_4 = –47,6 y el último coeficiente de autocorrelación significativo es ρ_3. El valor más alejado que se puede prever de la serie es:
  317,65 ☐ 282,15 ☐ 221,45 ☐ 194,95 ☐ . . . . ☐

- **6.** Se dispone de los datos cronológicos: Y_1 = 45,74; Y_2 = 47,95; Y_3 = 49,23; Y_4 = 51,47; ...
  Para un valor λ = 0,8, cuál es el cuarto valor de la serie suavizada (S_4)?
  48,89 ☐ 51,37 ☐ 41,18 ☐ 50,95 ☐ . . . . ☐

- **7.** Un modelo aditivo de período 3, ha dado los siguientes índices estacionales: E_1 = 10; E_2 = 20 y E_3 = –30. Los coeficientes β_2 y β_3 del modelo en variables categóricas se estiman como:
  20 y 30 ☐ 10 y –40 ☐ 25 y 45 ☐ –10 y 10 ☐ . . . . ☐

- **8.** La modelización de una serie aditiva con variables categóricas ha dado
  $\hat{Y}$ = 104,8 –0,5 t –8,2 Q_2 + 15,4 Q_3. El valor previsto para t = 50 es:
  71,6 ☐ 87 ☐ 95,2 ☐ 79,8 ☐ . . . . ☐

- **9.** En la serie de la pregunta anterior, el último valor observado ha sido y = 81,5 para t = 49. ¿Qué valor tiene el residuo?
  –13,2 ☐ 0 ☐ 1,2 ☐ 9,4 ☐ . . . . ☐

- **10** Con 252 datos se han obtenido los coeficientes de autocorrelación: r_1= 0,983; r_2= 0,537; r_3= 0,684; r_4= 0,322; ... ¿En qué intervalo de valores se puede considerar nulo ρ_3?
  ±0,266 ☐ ±0,236 ☐ ±0,299 ☐ ±0,225 ☐ . . . . ☐

3.5.99

● ● ● Se dispone de 100 valores de una serie siendo los 6 últimos 53,0; 89,3; 66,6; 29,1; 194,8 y 61,2. Se detecta que tiene una estacionalidad de periodo 5 y que es de tipo multiplicativo.

● **1.** El valor de la última media móvil es:

74,02 ☐ 86,56 ☐ 88,2 ☐ 87,38 ☐ .... ☐

● **2.** Se ha obtenido $E_1 = 108{,}3$; $E_2 = 75{,}1$; $E_4 = 220{,}6$ y $E_5 = 65{,}6$. ¿Qué valor tiene E_3?

25,8 ☐ 30,4 ☐ –469,6 ☐ 220,6 ☐ .... ☐

● **3.** Por la tendencia se han probado los modelos rectilíneo y parabólico, obteniendo

Modelo		**T = a + b t**		**T = a + b t + c t^2**	
Coeficientes	a = 65,24	b = 0,79	a = 65,62	b = 0,68	c= 0,0050
p–value	–	0,0000	–	0,0221	0,6943
R^2		0,893		0,900	

¿Cuál es la previsión para t = 104?

420,07 ☐ 97,21 ☐ 126,05 ☐ 325,16 ☐ .... ☐

● ● ● Una serie de la que tenemos 92 valores se ha modelizado con variables categóricas obteniéndose $\hat{Y} = 250{,}83 + 1{,}27t - 0{,}006t^2 + 5{,}35Q_2 - 8{,}27Q_3 - 10{,}2Q_4 + 15{,}60Q_5$

● **4.** ¿Cuál es la longitud de la estacionalidad (p)?

3 ☐ 4 ☐ 5 ☐ 6 ☐ no se sabe ☐ ....☐

● **5.** Siendo $Y_{92} = 320$, ¿qué valor tiene su residuo?

–2,236 ☐ 13,154 ☐ 11,224 ☐ –6,137 ☐ ....☐

● **6.** ¿Qué valor tiene el índice estacional E_2?

no se sabe ☐ –0,496 ☐ 4,854 ☐ –8,766 ☐ ....☐

● **7.** En una serie de 100 datos, los coeficientes de autocorrelación calculados son $r_1 = 0{,}952$ $r_2 = 0{,}741$ $r_3 = 0{,}583$ $r_4 = 0{,}492$. ρ_4 será considerado nulo si r_4, en valor absoluto, es menor que

0,2792 ☐ 0,4050 ☐ 0,4285 ☐ 0,5412 ☐ ....☐

● ● Los valores de una serie son 40,22; 54,89; 63,51; ....

● **8.** En un suavizado exponencial con $\lambda = 0{,}4$, ¿cuál es el valor de S_3?

58,338 ☐ 53,0568 ☐ 49,0220 ☐ 52,1252 ☐ .... ☐

● **9.** Según el método de Brown, ¿cuál es el valor modelado para t = 3 ($\hat{Y}_3$) ?

63,790 ☐ 56,614 ☐ 51,956 ☐ 40,220 ☐ .... ☐

● **10.** Los valores de una serie son 67,38; 56,09; 75,11; 55,90 y 61,25 y los estimados según el modelo resultante del análisis han sido 56,44; 62,29; 72,13; 59,60; y 65,45.¿Cuál es el valor del error cuadrático medio (MSE)?

42,931 ☐ 40,697 ☐ 40,374 ☐ 39,667 ☐ .... ☐

23.6.99

● ● ● ● ● Los primeros valores de una serie, de la que se dispone de 141 observaciones, son: 225; 219; 196; 197; 235; 208; 191; 212; 216; .... Se trata de un modelo aditivo con estacionalidad de período 4. Por el sistema clásico se ha obtenido como tendencia $T_t = 200 + 0{,}10\ t$ y como índices estacionales $E_1 = 0{,}73$; $E_2 = 0{,}87$ y $E_3 = -0{,}4$.

● **1**. ¿Cuál es el valor de la primera media móvil?
207,500 ☐ 208,250 ☐ 208,375 ☐ 209,625 ☐ 210,500 ☐ .... ☐

● **2**. ¿Y cuál el del residuo para t = 8?
12,20 ☐ 12,30 ☐ 12,32 ☐ 12,40 ☐ 12,42 ☐ .... ☐

● **3**. ¿Cuáles son los valores de les variables categóricas asociadas a t = 10?
(0; 0; 0) ☐ (1; 0; 0) ☐ (0; 1; 0) ☐ (0; 0; 1) ☐ (1; 1; 1) ☐ .... ☐

● **4**. Si se hubiese modelado con variables categóricas, ¿cuál habría sido el valor de la constante α_0?
200,63 ☐ 200,65 ☐ 200,73 ☐ 200,75 ☐ 200,83 ☐ ....☐

● **5**. En una ponderación exponencial simple ha resultado $S_3 = 211{,}96$. ¿Cuál es el valor de λ?
2,63 ☐ 3,20 ☐ 3,30 ☐ 4,50 ☐ 5,43 ☐ ....☐

● **6**. En la modelización de una serie multiplicativa de p = 3, se ha obtenido $T_t = 50 + 0{,}2\ t + 0{,}1\ t^2$; $E_1 = 150$; $E_2 = 50$ y para t=3 el residuo ha sido $R_3 = 0{,}8$. ¿Cuál es el valor de Y_3?
72,10 ☐ 61,80 ☐ 55,75 ☐ 46,35 ☐ 52,30 ☐ .... ☐

● **7**. Con los 50 valores de una serie se ha obtenido $\sum_{i=1}^{50}(y_i - \bar{y})^2 = 4$; $\sum_{i=1}^{49}(y_i - \bar{y})\ (y_{i+1} - \bar{y}) = 3{,}6$; $\sum_{i=1}^{48}(y_i - \bar{y})\ (y_{i+2} - \bar{y}) = -\ 3{,}2$ y $\sum_{i=1}^{47}(y_i - \bar{y})\ (y_{i+3} - \bar{y}) = -2{,}8$ ¿Qué valor tiene la variancia de r_3?

0,0600 ☐ 0,0652 ☐ 0,0712 ☐ 0,0754 ☐ 0,0780 ☐ .... ☐

● **8**. Los valores de una serie sin estacionalidad y con tendencia rectilínea son 7,3; 7,8; 8,1; 8,5; 8,8; 9,0; .... Con $\lambda = 0{,}4$, ¿cuál es el valor modelizado para t=3?
7,700 ☐ 7,380 ☐ 7,004 ☐ 7,540 ☐ 7,860 ☐ .... ☐

12.1.00

● ● Unos datos cronológicos trimestrales han dado lugar a una tendencia $T=120+1{,}4\,t-0{,}2\,t^2$ y a una estacionalidad $E_1 = -10$; $E_2 = -8$; $E_3 = 15$ y $E_4 = 3$.

● **1.** ¿Qué diferencia existirá entre los valores estimados del primer trimestre del primer año y el segundo del año siguiente?
–2 ☐ –25 ☐ –13 ☐ –4 ☐ 18 ☐ ……… ☐

● **2.** El último dato disponible es el de t = 47. ¿Cuál es el valor previsto para t = 50?
–310 ☐ –348 ☐ –378 ☐ –318 ☐ –345 ☐ ……… ☐

● **3.** En una serie aditiva de p= 7, los pares de valores (t, Y_t) son (1; 15), (2; 19), (3; 17),…, (6;25), (7; 28), (8; 32), (9; 35), ... La media móvil para t = 4 es igual a 26. ¿Qué vale la de t = 5?
faltan datos ☐ 28,86 ☐ 28,43 ☐ 29,52 ☐ ……… ☐

● ● Un modelo en variables categóricas, con ordenada en el origen igual a 500, ajustado sobre una serie de período p=3, ha evidenciado que la serie crece 0,5 unidades por unidad de tiempo y que la segunda estación supera a la primera en 20 unidades, mientras que la tercera está 30 unidades por debajo de la segunda.

● **4.** El valor del coeficiente Q_3 es igual a
–30 ☐ –35 ☐ –5 ☐ –10 ☐ –15 ☐ ……… ☐

● **5.** La previsión para t = 53 es
528 ☐ 529,5 ☐ 546,5 ☐ 548 ☐ 549,5 ☐ ……… ☐

● **6.** En una serie de 100 valores se ha obtenido $\sum_{i=1}^{100} y_i = 0$; $\sum_{i=1}^{100} y_i^2 = 125$ y $\sum_{i=1}^{97} y_i \; y_{i+3} = 120$.

¿Qué vale r_3?
faltan datos ☐ 0 ☐ 0,80 ☐ 0,96 ☐ 1 ☐ ……… . ☐

● **7.** En una serie con 80 datos se ha obtenido $r_1 = 0{,}90$; $r_2 = 0{,}80$; $r_3 = 0{,}70$; $r_4 = 0{,}60$. ¿Cuál es el valor absoluto límite de r_5 para ser considerado distinto de cero?
0,43 ☐ 0,50 ☐ 0,53 ☐ 0,61 ☐ 0,64 ☐ ……… . ☐

● **8.** Los valores de una serie son 16,4; 16,9; 18,1; 18,5; 19,3; 19,8;… en un suavizado exponencial con $\lambda = 0{,}6$. ¿Cuál es el error de previsión para t = 4?
0,805 ☐ 0,925 ☐ 0,960 ☐ 1,115 ☐ 1,300 ☐ ……… . ☐

● **9.** En la misma serie del apartado anterior y con igual factor de ponderación, ¿cuál sería el valor estimado para t = 4 $(\hat{Y}_4)$ utilizando el método de Brown?
17,920 ☐ 19,076 ☐ 18,672 ☐ 19,137 ☐ …………… ☐

17.5.00

● ● En la descomposición clásica de una serie aditiva de período estacional p = 7, se ha obtenido como tendencia T=223,82 + 0,63 t. La previsión para t = 102 ha sido 187,25.

● **1.** ¿Qué vale la previsión para t = 109?
faltan datos ☐ 292,49 ☐ 191,66 ☐ 182,66 ☐ .... ☐

● **2.** Siendo E_i la i-ésima estacionalidad, y sabiendo que $E_6 = E_4 - 27{,}16$. ¿Cuál es la previsión para t = 104?
faltan datos ☐ 262,18 ☐ 305,18 ☐ 161,35 ☐ 195,35 ☐ .... ☐

● ● ● Unos datos bimensuales se modelizan como
$Y = 187{,}52 + 0{,}42\,t + 10\,Q_2 + 12\,Q_3 + 16\,Q_4 - 8\,Q_5 - 2\,Q_6$

● **3.** ¿En qué cantidad se diferencian el segundo y el sexto bimestre de un mismo año? (6° – 2°)
–12 ☐ –10,32 ☐ –18 ☐ –16,74 ☐ .... ☐

● **4.** Si el último valor disponible es $Y_{106} = 250{,}27$, ¿ qué vale el residuo de este punto?
–9,77 ☐ –30,27 ☐ 5,73 ☐ 2,23 ☐ .... ☐

● **5.** ¿Cuál es la previsión para t = 107?
232,88 ☐ 250,69 ☐ 230,61 ☐ 224,46 ☐ ....☐

● ● Con 100 datos se ha obtenido $\sum_{i=1}^{94}(y_i - \bar{y})(y_{i+6} - \bar{y}) = -483{,}22$ y $\sum_{i=1}^{100}(y_i - \bar{y})^2 = 793{,}42$

● **6.** ¿Qué vale r_6?
hay un error ☐ –0,371 ☐ –0,609 ☐ –0,684 ☐ ....☐

● **7.** ¿Cuál es el intervalo de no significación para r_6 si $r_1 = -0{,}95$; $r_2 = 0{,}32$; $r_3 = -0{,}84$; $r_4 = 0{,}60$ y $r_5 = 0{,}90$?
±0,464 ☐ ±0,179 ☐ ±0,520 ☐ ±1,323 ☐ ....☐

● **8.** Se dispone de los valores 23,87; 15,22; 42,75; 54,23 y 50,80. En una ponderación exponencial simple con λ = 0,8, ¿qué vale el error cuadrático medio?
410,17 ☐ 350,72 ☐ 254,34 ☐ 180,69 ☐ .... ☐

● **9.** En un suavizado exponencial por Brown, con λ = 0,7, sobre 50 datos, ha resultado $Y_{50} = 55{,}87$; $S_{50} = 49{,}32$; $S_{50}^{(2)} = 47{,}54$. ¿Cuál es la previsión para t = 52?
70,25 ☐ 59,41 ☐ 40,23 ☐ 36,44 ☐ .... ☐

● **10.** En una serie multiplicativa de período p = 3, se ha obtenido $E_1^* = 15{,}25$; $E_2^* = 30{,}50$ y $E_3^* = 45{,}75$. ¿Cuál es el valor del primer índice estacional?
25 ☐ 50 ☐ 100 ☐ 150 ☐ 200 ☐ ....☐

2 EVALUACIONES RESUELTAS

Respuesta correcta +1; incorrecta −0,2

● **1.** En un análisis de componentes principales los valores propios, de la matriz de correlaciones, son {2,78; 2; 0,16; 0,05; 0,01} y g_{13} = 0,768. ¿Qué vale r_{13} ?.

0,143 ☐ 0,527 ☐ 0,12 ☐ 0,3072 ■ .... ☐

Puesto que $\Sigma\, d_i = 5$ es un valor entero, coincidente con el número de valores propios, necesariamente se trabaja con variables estandarizadas y se ha diagonalizado la matriz de correlaciones. Entonces,

$$r_{13} = g_{13}\sqrt{d_3} = 0{,}768\sqrt{0{,}16} = 0{,}3072$$

17.3.99

En una tabla de correspondencias la 3ª columna es 13; 23; 17 y 20, y los totales de les columnas son 100; 97; 73; 133 y 152.

● **1.** ¿Cuántos valores propios no triviales hay?

2 ☐ 3 ■ 4 ☐ 5 ☐ .... ☐

Dado que hay p = 4 files y q = 5 columnas, resulta min (p–1, q–1) = 3.

● **2.** ¿Cuál es la masa de la 2ª columna?

0,314 ☐ 0,240 ☐ 0,175 ■ 0,711 ☐ .... ☐

Las masas de las columnas son

$$f_{\bullet j} = \frac{n_{\bullet j}}{n} = \frac{\sum_{i=1}^{p} n_{ij}}{\sum_{i=1}^{p} n_{\bullet j}}$$

que, para j=2, es

$$f_{\bullet 2} = \frac{n_{\bullet 2}}{\sum n_{\bullet j}} = \frac{97}{555} = 0{,}175$$

● **3.** ¿Cuál es el valor del perfil medio de las filas asociado a la 3ª columna?

0,714 ☐ 0,132 ■ 0,312 ☐ 0,511 ☐ .... ☐

El perfil medio de las filas coincide con las masas de las columnas

$$f_{\bullet 3} = \frac{n_{\bullet 3}}{n} = \frac{73}{555} = 0{,}132$$

● **4.** ¿Qué vale la masa total?

1 ■ 0,13 ☐ 0,312 ☐ 0,811 ☐ .... ☐

La masa total es, obviamente,

$$\sum_{i=1}^{p} f_{i\bullet} = \sum_{j=1}^{q} f_{\bullet j} = \frac{n}{n} = 1$$

En un análisis de componentes principales los valores propios de la matriz de correlaciones son {2,78; 2; 0,16; 0,05; 0,01} y g_{13} = 0,768.

● **5**. ¿De qué dimensión es el vectorio aleatorio?

4 ☐ 5 ■ 6 ☐ 7 ☐ .... ☐

La dimensión del vector aleatorio **X**, coincide con el número de valores propios. En este caso p = 5.

● **6**. ¿Qué vale r_{13} ?.

0,143 ☐ 0,527 ☐ 0,12 ☐ 0,3072 ■ .... ☐

Dado que $\Sigma\ d_i = 5$, un valor entero coincidente con el nombre de valores propios, necesariamente se trabaja con variables estandardizadas y se ha diagonalizado la matriz de correlaciones. Entonces

$$r_{13} = g_{13}\sqrt{d_3} = 0{,}768\ \sqrt{0{,}16} = 0{,}3072$$

● **7**. ¿Cuántos componentes principales se utilizarían?

1 ☐ 2 ■ 3 ☐ 4 ☐ .... ☐

La proporción acumulada que representan los valores propios (variancias de los componentes principales) con relación al total es: 2,78/5 = 0,556 (2,78 + 2)/5 = 0,956 etc. Entonces los dos primeros ya son suficientes ya que explican el 95,6% del total.

● **8.** Al estudiar los componentes principales ha resultado ${}^tg_1$ = {0,48 0,32 0,47 0,48 0,46}, ${}^tg_2$= {0,40 0,21 0,8 –0,28 0,26} y Q = diag{4 9 6,25 7,75 8}. ¿Qué vale $\sum_{i=1}^{p} d_i$?

No se sabe ☐ –3,14 ☐ 4 ☐ 35 ■ .... ☐

En las se indican Q = diag $(s_1^2\ s_2^2 \ldots s_p^2)$ y $\sum_{i=1}^{p} d_i = \sum_{i=1}^{p} s_i^2$ resultando

$$\sum_{i=1}^{p} d_i = 35$$

19.4.99

En una tabla de correspondencias les 3ª y 4ª filas son {47; 65; 78; 35} y {82; 42; 76; 23}; Además, las masas de las filas son {0,134; 0,268; 0,225; 0,223; 0,150}

● **1.** ¿Cuál es la suma total, n?

225 ☐ 223 ☐ 777 ☐ 1000 ■ .... ☐

Resulta $n = \dfrac{n_{3\bullet}}{f_{3\bullet}} = \dfrac{225}{0{,}225} = 1000$

● **2.** Si la masa de la 3ª columna es 0,232, ¿qué vale X_{33}?

1 ☐ 0,2141 ☐ 0,7197 ■ 0,0682 ☐ .... ☐

Por definición $X_{3\,3} = \dfrac{f_{33}}{f_{3\bullet}\sqrt{f_{\bullet 3}}} = \dfrac{0{,}078}{0{,}225\sqrt{0{,}232}} = 0{,}7197$

Al estudiar los componentes principales ha resultado ${}^t g_1$ = {0,47 0,32 0,48 0,46 0,48}, ${}^t g_2$ = {0,40 –0,28 0,8 0,21 0,26} y Q = diag{4 9 6,25 7,75 3,8}

● **3.** Si r_{12} = 0,632, ¿qué vale d_2?

22,14 ☐ 36,48 ☐ –25,78 ☐ 9,99 ■ .... ☐

Las variancias, expuestas en la diagonal de la matriz Q, son razonablemente homogéneas, indicando que se ha diagonalizado la matriz S, y teniendo en cuenta que

$$r_{12} = \frac{g_{12}\sqrt{d_2}}{s_1} \quad \text{resulta} \quad d_2 = \left(\frac{r_{12}\ s_1}{g_{12}}\right)^2 = \frac{0{,}632^2 \times 4}{0{,}40^2} = 9{,}99$$

● **4.** Si los dos primeros componentes expliquen un 95% de la variabilidad total, ¿qué vale d_1+d_2?

27 ☐ 32 ☐ 64 ☐ 25,65 ■ .... ☐

Dado que $\sum_{i=1}^{p} d_i = \sum_{i=1}^{p} s_i^2 = 27$ y que $\dfrac{d_1 + d_2}{\sum_i d_i} = 0{,}95$ resulta

$$d_1+d_2 = 0{,}95\times 27 = 25{,}65$$

● **5.** Si la primera fila de la matriz R_{XY} es (0,942 0,265 0,202 –0,011 0,004), ¿qué porcentaje de la variabilidad de X_1 es explicado por los tres primeros componentes?.

1 ☐ 0,5236 ☐ 0,9984 ■ 0,9763 ☐ .... ☐

La explicación es $\sum_{j=1}^{3} r_{ij}^2 = 0{,}942^2 + 0{,}265^2 + 0{,}202^2 = 0{,}9984$

5.11.99

● **1.** Habiendo diagonalizado la matriz de variancias-covariancias, ¿qué vale $\sum_{i=1}^{p} r_{i2}^2$?

d_2 ☐ 1 ☐ 0,9 ☐ Es una función cuadrática de g_{i2} ■ ……………………… ☐

Si se ha diagonalizado la matriz variancias-covariancias

$$R_{XY} = Q^{-1/2}\ G\ D^{1/2}$$

según aparece en la Pág. 6 de los apuntes, por lo que el producto escalar es

$$^tR_{XY}\ R_{XY} = D^{1/2}\ {}^tG\ Q^{-1}\ G\ D^{1/2}$$

y siendo Q = diag (S_1^2 . . . S_p^2), resulta

$$\sum_{i=1}^{p} r_{i2}^2 = \sum_{i=1}^{p} \frac{g_{i2}^2}{S_i^2}\ d_2$$

función cuadrática de g_{i2}.

● **2.** Si U_{23}=33, U_{32}=44, $\bar{U}_2 = 10$, $\bar{U}_3 = 30$, S_2=5, S_3=10 y se estandariza, ¿qué vale X_{23}?

0,3 ■ 4,6 ☐ 6,8 ☐ 1,4 ☐ ……………………………………………… ☐

$$X_{23} = \frac{U_{23} - \bar{U}_3}{S_3} = \frac{33 - 30}{10} = 0,3$$

● **3.** Si el mayor valor absoluto de la matriz de correlaciones es 0,307, ¿qué procede?

Calcular los componentes principales ☐ Estandarizar ☐ Factorizar ☐
Analizar las variables directas ■ ……………………………………………… ☐

Si máx |ρ| = 0,307, las correlaciones entre las variables son muy reducidas, la información redundante es prácticamente nula y se requeriría un número muy elevado de componentes para explicar razonablemente la variabilidad total. Por todo ello los componentes principales son inútiles.

$$\text{Si } D = \text{diag}(3{,}24\ \ 0{,}7\ \ 0{,}045\ \ 0{,}015) \text{ y } G = 0{,}5 \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \end{pmatrix}$$

● **4.** ¿Qué proporción de X_3 es explicada por Y_2?

17,5% ■ 92,3% ☐ 1,125% ☐ No se sabe ☐ ……………………………… ☐

Puesto que la suma de los valores d_i es 4, el orden de la matriz, se ha diagonalizado la matriz de correlaciones y la proporción de X_3 explicada por Y_2 es

$$r_{32}^2 = g_{32}^2\ d_2 = (-0{,}5)^2 \times 0{,}7 = 0{,}175$$

● **5.** ¿Qué vale la correlación experimental entre X_3 e Y_1?

–1,2 ☐ –0,9 ■ –0,1061 ☐ 3,4 ☐ ....☐

Teniendo en cuenta que la matriz D, de valores propios, así como la matriz G, de vectores propios, son de orden 4×4, se trata de un análisis de p=4 variables y como la suma de los valores d_i es cuatro, se ha diagonalizado la matriz de correlaciones y

$$r_{31} = g_{31} \sqrt{d_1} = (-0{,}5) \sqrt{3{,}24} = -0{,}9$$

● **6.** El primer componente principal es

un factor de tamaño ☐ un contraste ■ una media ☐ no se sabe ☐ .... ☐

Un contraste de X_1 y X_2 con X_3 y X_4 ya que los correspondientes coeficientes g_{i1} cambian de signo.

● **7.** ¿Cuál es la variancia experimental de Y_5?

No existe ■ 4 ☐ 3 ☐ 0,25 ☐ .... ☐

Dado que las matrices D y G son de orden 4×4, sólo hay cuatro variables y, por tanto, el número de Componentes Principales es, también, 4 y no existe Y_5.

● **8** ¿Cuál es la medida relativa de la información compartida por dos variables?

La razón de medias ☐ la razón de variancias ☐ la covariancia ☐ la correlación ■
.... ☐

La correlación es la medida adimensional relativa de la información compartida (covariancia) por dos variables aleatorias.

● **9** La homogeneidad entre modalidades de un criterio de clasificación es medida por

los factores comunes ☐ el factor específico ☐
las correspondencias ■ la covariancia ☐
.... ☐

Las correspondencias, comparando los perfiles mediante la distancia de χ^2.

ATENCIÓN, MARCAR LA ÚNICA RESPUESTA INCORRECTA

● **10** Los componentes principales:
reducen la masa de datos ☐ conservan la información ☐
eliminan información redundante ☐ reducen el nº de individuos ■
reducen el número de variables ☐

Reducen el número de individuos es incorrecto, ya que disminuyen el número de variables.

20.3.00

● **1.** Una fila de una tabla de correspondencias es {13 26 39 22}, ¿cuál es el tercer elemento de su perfil?

Falta n □ 0,39 ■ 39 □ 1 □ .... □

El tercer elemento de su perfil es

$$\frac{f_{i3}}{f_{i\bullet}} = \frac{n_{i3}}{n_{i\bullet}} = \frac{n_{i3}}{\sum_j n_{ij}} = \frac{39}{13+26+39+22} = 0{,}39$$

● **2.** Si el perfil de la 3ª fila es {0,31 0,60 0,74 0,26} y $n_{3\bullet} = 500$, ¿qué vale el tercer elemento de esa fila?

Hay un error ■ 370 □ 0,025 □ 0,01 □ .... □

Hay un error ya que si fuese un perfil la suma de sus elementos sería 1 y aquí, obviamente, no se cumple este requisito.

● **3.** Si hay 14 puntos fila y 23 puntos columna, ¿cuántos valores propios nulos hay en total?

1 □ 0 □ 10 ■ 13 □ .... □

Los valores propios no triviales (distintos de cero) son mín(p–1, q–1) = 13, por lo que los nulos son

máx(p; q) – mín(p–1; q–1) = 23 – 13 = 10.

● **4.** Si $n_{13} = 24$, $n_{1\bullet} = 100$, $n_{\bullet 3} = 90$ y $n = 900$, ¿qué vale el elemento correspondiente de la matriz *Z* para el estudio de las distancias de χ^2 entre las columnas?

0,99 □ 0,95 □ 0,05 □ 0,8 ■ .... □

El elemento de la matriz *Z* será

$$\frac{f_{13}}{\sqrt{f_{1\bullet}\ f_{\bullet 3}}} = \frac{n_{13}\ /\ n}{\sqrt{n_{1\bullet}/n\ \ (n_{\bullet 3}/n)}} = \frac{n_{13}\ \sqrt{n}}{\sqrt{n_{1\bullet}\ \ n_{\bullet 3}}} = \frac{24\sqrt{900}}{\sqrt{100\ \times 90}} = 0{,}8$$

● **5.** Si una columna es { 0,15 0,18 0,22 0,45 }, ¿es un perfil o es de frecuencias?

Hay un error □ Faltan datos □ Es un perfil ■ Son frecuencias □ .... □

Se reconoce que es un perfil si suma 1. Dado que la columna en cuestión cumple dicha condición se trata, efectivamente, de un perfil.

● **6.** Con Q = diag(2 4 1600 725), D = diag(3,5 0,4 0,07 0,03), si r_{12} = 0,87, ¿cuál es la parte de $V(X_1)$ explicada por el segundo componente?

Falta g_{12} ☐ 0,87 ☐ 0,4 ☐ 0,7569 ■ .... ☐

La matriz Q muestra que las variancias S_i^2 son harto heterogéneas, por lo que se ha estandardizado (Opción B), circunstancia corroborada por el hecho de que traza D = p = 4, y la parte de $V(X_1) = 1$ explicada por el segundo componente principal coincide con la proporción, es decir

$$r_{12}^{\ 2} = 0{,}87^2 = 0{,}7569.$$

● **7.** Si Q = diag(16 9 25 4) y $\mathrm{cov}(X_1, X_3) = -18$, ¿qué vale $r_{X_3X_1}$?

Falta n ☐ –1,8 ☐ –0,9 ■ 0,361 ☐ .... ☐

Resulta

$$r_{X_3X_1} = r_{X_1X_3} = \frac{\mathrm{cov}(X_1,\ X_3)}{S_1\ \ S_3} = -\frac{18}{\sqrt{16\ \times\ 25}} = -0{,}9$$

● **8.** Si los puntos de dos variables se oponen, representa que

se trata de un error ☐ son no correlacionadas ☐

su correlación es alta e inversa ■ su correlación es alta y directa ☐

.... ☐

Si los puntos se oponen están altamente correlacionados de forma inversa.

● **9.** $U_{32} = 24$, $U_{23} = 32$, $\bar{\mathbf{U}} = {}^t(9\ \ 12\ \ 16\ \ 8\ \ 14)$ y Q = diag(12 14 16 8 13), ¿qué vale X_{32}?

4 ☐ 6 ☐ 12 ■ 16 ☐ .... ☐

Teniendo en cuenta que, como muestra la matriz Q, las variancias son del mismo orden de magnitud, sólo se requiere centrar y

$$X_{32} = U_{32} - \bar{U}_2 = 24 - 12 = 12$$

● **10.** Si $r_{31} = 1$, ¿qué vale r_{32}?

No se sabe ☐ 0 ■ 1 ☐ –1 ☐ .... ☐

Dado que siempre $\sum_{j=1}^{p} r_{ij}^2 = 1$, si $r_{31} = 1$, necesariamente todos los demás r_{3j}, j > 1, han de ser nulos.