

**Universidad de Castilla-La Mancha**  
**Facultad Ciencias del Medio Ambiente**  
**Estadística. Tercer Curso**  
**Examen de convocatoria ordinaria, 25-01-2005**

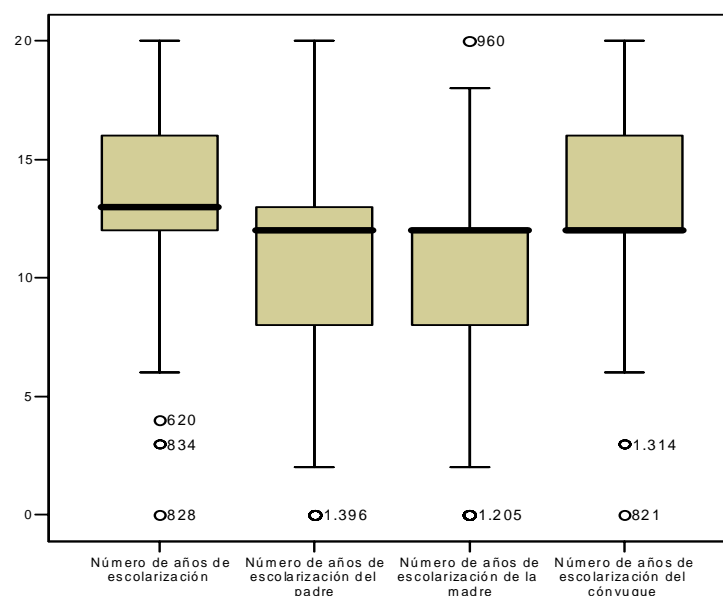
**INSTRUCCIONES**

- Escribe claramente tus apellidos y nombre, en todas y cada una de las hojas de examen.
- No se admitirá más de una hoja por problema. No escribas partes de problemas diferentes en una misma hoja. Los problemas se han de entregar por separado.
- En función de si uno tiene el examen parcial de evaluación suspenso, la duración, los ejercicios y su puntuación, quedan distribuidos del siguiente modo:

	Ejercicios	Puntuación	Duración
Con parcial aprobado	3,4,5,7 (8 opcional)	2'5 ptos cada uno (hasta 1 pto)	3h.
Sin parcial aprobado	1,2,3,4,6,7,8	1,1'5,1'5,2,1,1'5,1'5 ptos. resp.	3h.

**EJERCICIOS**

1. En una encuesta realizada sobre la población de los EE.UU. de Norte America, se han estudiado entre otras, cuatro variables de interés relacionadas con el nivel de estudios. Éstas son: Número de años de escolarización de la persona encuestada, de su padre, de su madre y de su cónyuge. Se han obtenido los diagramas de cajas correspondientes a las cuatro variables así como unos estadísticos descriptivos con el fin de caracterizar estas variables, y se han obtenido los siguientes resultados (téngase en cuenta que no todos los encuestados respondieron a todas las preguntas):



Estadísticos descriptivos

	Número de años de escolarización	Número de años de escolarización del padre	Número de años de escolarización del cónyuge	Número de años de escolarización de la madre	N válido (según lista)
N	1510	1069	790	1233	535
Rango	20	20	20	20	
Mínimo	0	0	0	0	
Máximo	20	20	20	20	
Suma	19455	11632	10184	13298	
Media	12.88	10.88	12.89	10.79	
Desv. típ.	2.984	4.129	3.059	3.463	
Varianza	8.904	17.045	9.359	11.990	
Asimetría	-.168	-.184	-.237	-.646	
	.063	.075	.087	.070	
Curtosis	.710	-.094	1.098	.956	
	.126	.149	.174	.139	

A la vista de los resultados responda a las siguientes preguntas:

- Valore las diferencias estadísticas más significativas entre los niveles educativos de la generación anterior y la que respondía a la encuesta.
- Dónde se produce mayor cambio generacional en los parámetros de posición o en los de dispersión. Justifique la respuesta.
- Valore la simetría de las distribuciones. Justifique el resultado
- Valore el apuntamiento de las distribuciones. Justifique la re-

spuesta.

- (e) Hay alguna información en el rango. ¿Por qué? ¿Hay alguna información en el rango intercuartílico? ¿Por qué?
  - (f) ¿Diría que las distribuciones son normales? Justifique la respuesta.
2. En tres grandes poblaciones  $A$ ,  $B$  y  $C$  la proporción de individuos infectados por un determinado virus es del 30, 60 y 10% respectivamente. Se toma al azar una de las 3 poblaciones (las tres son igual de probables), y de ella elegimos a 10 individuos al azar, resultando que 2 de ellos están infectados. ¿A qué población es más probable que pertenezcan?
  3. El número de partículas en suspensión que había en Madrid el día 22 de Febrero estaba distribuido de manera uniforme entre 20000 y 30000 partículas por decímetro cúbico ( $\text{dm}^3$ ). Sabiendo que en cada inspiración introducimos  $0.5 \text{ dm}^3$  de aire en nuestros pulmones y que realizamos 10 inspiraciones por minuto, calcular la probabilidad de que el número de partículas inspirado en una hora sea superior a 7510000 (Utilizar el Teorema Central del Límite).
  4. Un investigador está estudiando un parámetro  $\theta$  que representa la duración media (en días) de vida de una rata desde que se le suministra cierto tipo de veneno. Se sabe que esta duración viene dada por una variable aleatoria de distribución  $N(\theta, 0.2)$ . Les suministra el veneno a 100 ratas resultando que la suma de los tiempos que sobreviven dichas ratas es de 30 días. En base a esto construir un intervalo de confianza del 95% para el parámetro  $\theta$ . Interpretar convenientemente el resultado obtenido.
  5. Una población sigue una distribución de probabilidad cuya función de densidad de masa viene dada por

$$P_{\xi}^*(x) = p(1-p)^{x-1} \quad x = 1, 2, \dots$$

siendo  $p$  cierto parámetro del intervalo  $(0, 1]$ . Se pide determinar el estimador de  $p$  mediante el método de máxima verosimilitud ¿Cuál sería la estimación correspondiente si la realización de  $\bar{\xi}$  es  $\bar{x} = 6.5$ ?

6. Lanzamos un dardo sobre el plano  $\mathbb{R}^2$  al objeto de acertar en una diana de radio  $\sqrt[2]{1'39}$  y centro en el punto de coordenadas  $(3, 2)$ . Se supone que los lanzamientos tienen carácter aleatorio y que cada impacto  $(\xi_1, \xi_2)$  es un vector aleatorio, de suerte que las componentes  $\xi_1$  y  $\xi_2$  son independientes y además  $\xi_1 \sim N(3, 1)$  y  $\xi_2 \sim N(2, 1)$ . Calcular la probabilidad

de que el dardo impacte dentro de la diana (Utilizar la distribución de una  $\mathcal{X}^2$ ).

7. Sean,  $\xi_1$  la v.a. que representa a los gastos por impagados en millones de euros de una determinada empresa, y  $\xi_2$  la v.a. de los ingresos netos. Se sabe que la función de densidad del vector  $(\xi_1, \xi_2)$  es

$$f_{(\xi_1, \xi_2)}(x, y) = \begin{cases} 5(1+x)e^{20(2-y)} & \text{si } x \in (0, 2) \text{ e } y > 2 \\ 0 & \text{en el resto} \end{cases}$$

- (a) Calcular la probabilidad de que los gastos por impagados sean menores a un millón de euros.
  - (b) ¿Son independientes los gastos por impagados y los ingresos netos?
  - (c) ¿Qué cantidad de ingresos netos se esperan obtener si los impagados ascendieron hasta 1 millón de euros?
8. Una empresa desea determinar la proporción de clientes dispuestos a adquirir uno de sus productos. Estima que dicha proporción es 0.45 ó 0.5. Decidir en base al Principio de Máxima Verosimilitud, una estimación de dicha proporción si después de realizar una muestra aleatoria simple de tamaño 15 entre sus clientes potenciales, 6 de ellos afirmaron estar dispuestos a la adquirir y los 9 restantes no estaban dispuestos a optar por dicho producto.