INSTRUCCIONES

1. El examen consta de cuestiones, que se responden sobre la hoja de codificación proporcionada, y problemas, que se responden en papel aparte.
2. Para escoger una respuesta, basta efectuar una marca **rellenando debidamente el rectángulo sobre el que está la letra escogida** en la hoja de codificación. Piénsalo antes; aunque puedes borrar si escribes con lápiz (número 2 o similar), marcas que no estén perfectamente borradas pueden ser leídas. Te aconsejamos que señales sobre el formulario de examen las respuestas que te parezcan adecuadas, y emplees los últimos diez minutos del tiempo asignado en transcribirlas a la hoja de codificación.
3. Hay siempre, en las preguntas de elección múltiple, una **única** respuesta correcta. Todas las cuestiones correctamente resueltas valen 1 punto mientras que las fallidas suponen una penalización de 0.2 puntos. Las preguntas no contestadas no suponen penalización.
4. Cada uno de los problemas, A, B y C debe responderse en una hoja de papel diferente. La recogida se producirá escalonadamente, en los momentos que constarán en la pizarra; primero, la hoja de codificación, y luego los problemas A, B, y C en este orden.
5. El formulario de examen tiene seis hojas numeradas correlativamente al pie (del 0.1 al 0.7). Cerciórate de recibirlas todas, y reclama si tu formulario fuera incompleto. Hay distintos tipos de examen. Este es del tipo 0; marca un 0 en la columna I de tu hoja de codificación, como en el ejemplo.
6. Los puntos obtenibles en cuestiones y problemas son 30 y 30 respectivamente. Son precisos 15 y 15 para superar el examen. Cuestionarios con puntuaciones iguales o superiores a 14 pueden en algún caso ser compensados por una buena nota en los problemas.
7. Rellena tus datos en la hoja de codificación y pliegos de papel suministrados. En "Convocatorias" (columna II) pondrás el número de convocatorias consumidas *incluyendo ésta.*

Ejemplo:

$\underbrace{\text{12545}}$ PEREZ, Ernesto Examen tipo $\underbrace{\text{0}}$ Convocatorias

CUESTIONES (Duración: 1 hora 30 minutos)

1. PREGUNTA-REGALO. La capital de España es:

   (A) París (B) Sebastopol (C) Madrid (D) Londres (E) Pekín

Las cuestiones 2 a 5 hacen referencia al siguiente enunciado:

La probabilidad de que un conductor tenga un accidente en la autopista que va desde Bilbao a Burgos es 0.10. Se supone independencia entre los distintos conductores que realizan este recorrido.

2. Si 15 conductores realizan este recorrido, la probabilidad de que como mucho tres de ellos tengan un accidente es:

   (A) 0.1285 (B) 0.9873 (C) 0.8159 (D) 0.9444 (E) 0.0556

3. La probabilidad de que, entre los 15 conductores que realizan este recorrido, exactamente nueve conductores **no** tengan un accidente es:

   (A) 0.1413 (B) 1 (C) 0.0019 (D) 0.3431 (E) 0.2059

4. Si ahora son 60 los conductores que realizan este recorrido, la probabilidad aproximada de que al menos cuatro conductores tengan un accidente es:

   (A) 0.8488 (B) 0.7149 (C) 0.2851 (D) 0.1339 (E) 0.1512

5. Si ahora son 360 los conductores que realizan este recorrido, la probabilidad aproximada de que exactamente 34 conductores tengan un accidente es:

   (A) 0.0674 (B) 0.6026 (C) 0.6700 (D) 0.9326 (E) 0.1434

Las cuestiones 6 y 7 hacen referencia al siguiente enunciado:

El número de accidentes laborales **diarios** en una determinada región sigue una distribución de Poisson de media 3. Se asume independencia entre los accidentes laborales de diferentes días en dicha región.

6. La probabilidad de que en **dos días** ocurran exactamente seis accidentes laborales en dicha región es:

   (A) 0.3937 (B) 0.6063 (C) 0.1606 (D) 0.4457 (E) 0.1377

7. La probabilidad de que en **10 días** ocurran menos de 25 accidentes laborales en dicha región es, aproximadamente:

   (A) 0.8212 (B) 0.1587 (C) 0.8413 (D) 0.1788 (E) 0.2061

8. Sea $\{X_n\}_{n\in\mathcal{N}}$ una sucesión de variables aleatorias con distribución $N(0, 1/n)$. Si se sabe que la función característica de una v.a. normal $N(m, \sigma^2)$ es $\psi_n(u) = e^{ium-\frac{\sigma^2 u^2}{2}}$, la sucesión convergerá:

   (A) Sólo en distribución a $X = 0$

   (B) Sólo en distribución a $X = 1$

   (C) En distribución y probabilidad a $X = 1$

   (D) En distribución y probabilidad a $X = 0$

   (E) Todo falso

9. Sea $\{X_n\}_{n\in\mathcal{N}}$ una sucesión de variables aleatorias con la siguiente función de cuantía:

$$P_n(x) = \begin{cases} \frac{1}{2} - \frac{1}{2n}, & \text{si } x = -\frac{1}{n} \\ \frac{1}{2n}, & \text{si } x = 0 \\ \frac{1}{2}, & \text{si } x = \frac{1}{n} \end{cases}$$

La sucesión convergerá:

(A) En distribución y probabilidad a $X = \frac{1}{2}$

(B) En distribución y probabilidad a $X = 0$

(C) Sólo en distribución a $X = 0$

(D) Sólo en distribución a $X = \frac{1}{2}$

(E) Todo falso

Las cuestiones 10 y 11 hacen referencia al siguiente enunciado:

Sean $X_1, X_2, \cdots, X_{75}$ v.a. independientes entre sí y todas con distribución uniforme $U(-1,1)$. **Ayuda:** Nota que $\mathrm{E}(X) = 0$ y que $\mathrm{Var}(X) = \frac{1}{3}$.

10. Si se define la v.a. $Y = X_1 + X_2 + \cdots + X_{75}$, entonces $P(Y \geq 6)$ es, aproximadamente:

(A) 0.1587 (B) 0.8413 (C) 0.1151 (D) 0.8849 (E) 0

11. Si se define la v.a. $Z = (X_1 + X_2 + \cdots + X_{75})/75$, entonces $P(0 \leq Z \leq 0.10)$ es, aproximadamente:

(A) 0.50 (B) 1 (C) 0.9332 (D) 0 (E) 0.4332

12. Si W es una v.a. una v.a. tal que $W \in \gamma(1/2, 5)$, entonces $P(W \leq 4.87)$ es:

(A) 0.10 (B) 0.05 (C) 0.25 (D) 0.95 (E) 0.90

13. Si Y es una v.a. tal que $Y \in \mathcal{F}_{12,10}$, entonces $P(Y \leq 0.3636)$ es:

(A) 0.95 (B) 0.10 (C) 0.05 (D) 0.90 (E) 0.25

14. Si Z es una v.a. tal que $Z \in t_{15}$, entonces $P(-1.75 \leq Z \leq 0.691)$ es:

(A) 0.40 (B) 0.10 (C) 0.30 (D) 0.70 (E) 0.25

15. Sea X una variable aleatoria con función de densidad de probabilidad:

$$f(x,\theta) = \frac{1}{4^\theta \Gamma(\theta)} \, x^{\theta-1} \, e^{-x/4}, \; x > 0, \; \theta > 0$$

Para estimar el parámetro θ se ha tomado una m.a.s. de tamaño n, $X_1, \ldots, X_n$. Un estadístico suficiente para el parámetro θ es:

(A) $\sum_{i=1}^n X_i^2$ (B) $\prod_{i=1}^n X_i$ (C) $\sum_{i=1}^n X_i$ (D) $\prod_{i=1}^n \left(\frac{1}{X_i}\right)$ (E) $\prod_{i=1}^n \ln X_i$

Las cuestiones 16 y 17 hacen referencia al siguiente enunciado:

Sea X una v.a. con función de cuantía dada por:

$$P(X=0)=2\theta; \quad P(X=1)=\frac{1}{2}-\theta; \quad P(X=-1)=\frac{1}{2}-\theta.$$

Para estimar el parámetro θ se ha tomado una m.a.s. de tamaño n, en la que han salido tres ceros.

16. La estimación de θ por el método de máxima verosimilitud es:

(A) $\frac{n-3}{2n}$ (B) $\frac{3}{2n}$ (C) $\frac{1}{n}$ (D) $\frac{n-3}{n}$ (E) $\frac{3}{n}$

17. La estimación de θ por el método de momentos es:

(A) $\frac{3}{n}$ (B) $\frac{3}{2n}$ (C) $\frac{n-3}{2n}$ (D) $\frac{n-3}{n}$ (E) $\frac{1}{n}$

Las cuestiones 18 y 19 hacen referencia al siguiente enunciado:

Sea X una v.a. con función de densidad de probabilidad dada por:

$$f(x)=\begin{cases} \frac{x}{\theta}e^{\frac{-x^2}{2\theta}}, & x>0, \quad \theta>0 \\ 0, & \text{en otro caso} \end{cases},$$

para la que se verifica que $\mathrm{E}(X)=\sqrt{\frac{\theta\pi}{2}}$. Para estimar el parámetro θ se ha tomado una m.a.s. de tamaño n.

18. El estimador de θ por el método de máxima verosimilitud es:

(A) $2n\overline{X}^2$ (B) Todo falso (C) $2\sum_{=1}^{n}X_i^2$ (D) $\frac{\sum_{i=1}^{n}X_i^2}{2n}$ (E) $\frac{\overline{X}^2}{2n}$

19. El estimador de θ por el método de momentos es:

(A) $\frac{\overline{X}^2}{\pi}$ (B) $\frac{2\overline{X}^2}{\pi}$ (C) $\pi\overline{X}^2$ (D) $\frac{\pi\overline{X}^2}{2}$ (E) Todo falso

Las cuestiones 20 y 21 hacen referencia al siguiente enunciado:

Sea $X_1,\dots,X_n$ una m.a.s. de una población $N(m,\sigma^2)$. Para estimar el parámetro m, se propone el estimador:

$$\hat{m}=\frac{X_1+X_2+\cdots+X_{n-1}}{n}$$

20. El sesgo del estimador $\hat{m}$ es:

(A) $-\frac{m}{n}$ (B) $\frac{1}{n}$ (C) 0 (D) $\frac{2m}{n}$ (E) nm

21. ¿Es el estimador $\hat{m}$ consistente?

(A) No (B) No se puede determinar (C) Sí (D) - (E) -

22. Sea X una v.a. con distribución exponencial de parámetro $1/\theta$. Si se toma una m.a.s. de tamaño $n=3$ y se propone $\hat{\theta}=(X_1+X_2+X_3)/2k$ como estimador del parámetro θ, el valor de k que hace que $\hat{\theta}$ sea un estimador insesgado de θ es:

(A) $\frac{2}{3}$ (B) Todo falso (C) $\frac{1}{3}$ (D) $\frac{3}{2}$ (E) 3

Las cuestiones 23 y 24 hacen referencia al siguiente enunciado:

Se quiere contrastar la hipótesis nula de que la v.a. X tiene una función de densidad $f(x) = 2x,\ x \in (0,1)$ frente a la hipótesis alternativa de que la función de densidad es $f(x) = 2(1-x),\ x \in (0,1)$. Para ello, se toma una m.a.s. de tamaño $n = 1$.

23. El contraste más potente para un nivel de significación del 5% nos dice que debemos rechazar la hipótesis nula si:

(A) $X \leq 2.24$ (B) $X \leq 0.224$ (C) $X \geq 0.224$ (D) $X \in (-0,224, 0.224)^c$ (E) Todo falso

24. La potencia de dicho contraste es:

(A) 0.224 (B) 0.398 (C) Todo falso (D) 0.95 (E) 0.603

Las cuestiones 25 y 26 hacen referencia al siguiente enunciado:

La variable aleatoria X sigue una distribución con la siguiente función de cuantía:

$$P(X=0) = 2\theta^2; \qquad P(X=1) = 3\theta^2; \qquad P(X=2) = 1 - 5\theta^2$$

Para contrastar la hipótesis nula de que $\theta = 0.05$ frente a la alternativa de que $\theta = 0.40$ se toma una m.a.s. de tamaño $n = 1$ y se decide rechazar la hipótesis nula si los valores obtenidos son $X = 0$ ó $X = 1$.

25. El nivel de significación de la prueba es:

(A) 0.25 (B) 0.75 (C) 0.0125 (D) 0.125 (E) 0.9875

26. La probabilidad de error tipo II de la prueba es:

(A) 0.20 (B) 0.95 (C) 0.75 (D) 0.25 (E) 0.80

Las cuestiones 27 a 30 hacen referencia al siguiente enunciado:

Un individuo está interesado en comprar una Play Station 2. Antes de hacerlo pregunta el precio en 11 comercios, obteniendo un precio medio muestral de 125 euros con una desviación típica muestral de 15 euros. Se supone normalidad.

27. Con una confianza del 90% se puede afirmar que el precio medio de la Play Station 2 se encuentra en el intervalo:

(A) (116.41, 133.59) (B) (118.50, 131.50) (C) (114.42, 135.58)
(D) (122.29, 127.72) (E) (120.36, 129.64)

28. Con una confianza del 95% se puede afirmar que la varianza del precio de la Play Station 2 se encuentra en el intervalo:

(A) (8.05, 50.77) (B) (135.25, 628.17) (C) (150.33, 892.44)
(D) (9.02, 41.88) (E) (120.73, 761.54)

29. Si al nivel de significación del 10% se desea contrastar la hipótesis nula de que el precio medio de la Play Station es $m = 130$, el resultado del contraste será:

(A) No rechazar la hipótesis nula (B) - (C) Rechazar la hipótesis nula (D) - (E) -

30. Si al nivel de significación del 5% se desea contrastar la hipótesis nula de que la varianza del precio de la Play Station es $\sigma^2 = 800$, el resultado del contraste será:

(A) Rechazar la hipótesis nula (B) - (C) - (D) - (E) No rechazar la hipótesis nula

PROBLEMAS (Duración: 75 minutos)

A. (10 puntos, 25 minutos)

Sean $X_1, X_2, \ldots, X_n$ v.a. independientes y tales que, para cada i, X_i sigue una distribución de Poisson de parámetro $k_i\lambda$. Es decir, $X_i \in \mathcal{P}(k_i\lambda),\ i = 1, \ldots, n$, donde las k_i's son constantes positivas conocidas y $\lambda > 0$.

a) Obtén **detalladamente** el estimador máximo verosímil de λ.

b) ¿Es insesgado? ¿Consistente? **Ayuda**: Puedes asumir que $\lim_{n\to\infty} \sum_{i=1}^{n} k_i = +\infty$.

B. (10 puntos, 25 minutos)

El profesor de una asignatura pretende determinar si el número de alumnos que asisten a clase se distribuye de forma equitativa durante los cinco días de clase de la semana. En base a una muestra aleatoria de cuatro semanas completas de clase observó los siguientes datos

Día	Lunes	Martes	Miércoles	Jueves	Viernes
Nº de alumnos	49	35	32	39	45

Al nivel de significación del 5%, ¿existe alguna razón para creer que el número de alumnos que asisten a clase no se distribuye uniformemente a lo largo de los cinco días de clase de la semana?

C. (10 puntos, 25 minutos)

Se desea contrastar la hipótesis nula $H_0 : \lambda = \frac{1}{2}$ frente a la alternativa $H_1 : \lambda = 2$ en una distribución exponencial de parámetro λ, para lo cual se toma una m.a.s. de tamaño $n = 10$, $X_1, \ldots, X_{10}$. A un nivel de significación del 5%, obtén **detalladamente** la región crítica más potente, si el estadístico de contraste utilizado es $Z = \sum_{i=1}^{10} X_i$. **Ayuda**: Si $Y \in \gamma(1/2, n/2) \Longrightarrow Y \in \chi^2_n$.

SOLUCIONES DEL CUESTIONARIO

1: C	11: E	21: C
2: D	12: A	22: D
3: C	13: C	23: B
4: A	14: D	24: B
5: A	15: B	25: C
6: C	16: B	26: A
7: B	17: B	27: A
8: D	18: D	28: E
9: B	19: B	29: A
10: C	20: A	30: A

SOLUCIONES DE LOS PROBLEMAS

A) Como $X_i \in \mathcal{P}(k_i\lambda)$, entonces tenemos que

$$P(x_i) = \frac{e^{-k_i\lambda}(k_i\lambda)^{x_i}}{x_i!}, \quad x_i = 0, 1, 2, \cdots, k_i, \quad \lambda > 0.$$

a) Así, la función de verosimilitud estará dada por

$$L(\lambda) = P(x_1; k_1\lambda) \cdots P(x_2; k_2\lambda) \cdots P(x_n; k_n\lambda)$$

$$L(\lambda) = \frac{e^{-k_1\lambda}(k_1\lambda)^{x_1}}{x_1!} \cdots \frac{e^{-k_n\lambda}(k_n\lambda)^{x_n}}{x_n!}$$

$$L(\lambda) = \frac{e^{-\lambda\sum_{i=1}^n k_i}\lambda^{\sum_{i=1}^n x_i}\Pi_{i=1}^n k_i^{x_i}}{\Pi_{i=1}^n x_i!}$$

El estimador máximo verosímil de λ es el valor que maximiza la función de verosimilitud o, equivalentemente, su logaritmo neperiano:

$$\ln L(\lambda) = -\lambda\sum_{i=1}^n k_i + \sum_{i=1}^n x_i \ln\lambda + \sum_{i=1}^n x_i \ln k_i - \sum_{i=1}^n \ln x_i!$$

Derivando respecto de λ, tendremos que:

$$\frac{\partial \ln L(\lambda)}{\partial\lambda} = 0$$

Por lo tanto,

$$-\sum_{i=1}^n k_i + \sum_{i=1}^n x_i \frac{1}{\lambda} = 0,$$

de donde

$$\hat{\lambda}_{MV} = \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n k_i}$$

b) El estimador será insesgado si $\mathrm{E}(\hat{\lambda}_{MV}) = \lambda$.

$$\mathrm{E}(\hat{\lambda}_{MV}) = \frac{1}{\sum_{i=1}^n k_i}\mathrm{E}\left(\sum_{i=1}^n X_i\right) = \frac{1}{\sum_{i=1}^n k_i}\sum_{i=1}^n \mathrm{E}(X_i)$$

$$\mathrm{E}(\hat{\lambda}_{MV}) = \frac{1}{\sum_{i=1}^n k_i}\sum_{i=1}^n k_i\lambda = \lambda$$

Por lo tanto $\hat{\lambda}_{MV}$ es un estimador insesgado para λ. Para comprobar si el estimador $\hat{\lambda}_{MV}$ es consistente podemos verificar si se cumplen las dos condiciones suficientes siguientes:

i) $\lim\limits_{n\to\infty} \mathrm{E}(\hat{\lambda}_{MV}) = \lambda$

ii) $\lim\limits_{n\to\infty} \mathrm{Var}(\hat{\lambda}_{MV}) = 0$

Como $\hat{\lambda}_{MV}$ es un estimador insesgado de λ, la condición i) se cumple. Por otro lado,

$$\lim_{n\to\infty} \mathrm{Var}(\hat{\lambda}_{MV}) = \lim_{n\to\infty} \frac{1}{\left(\sum_{i=1}^{n} k_i\right)^2} \sum_{i=1}^{n} \mathrm{Var}(X_i)$$

$$\lim_{n\to\infty} \mathrm{Var}(\hat{\lambda}_{MV}) = \lim_{n\to\infty} \frac{1}{\left(\sum_{i=1}^{n} k_i\right)^2} \sum_{i=1}^{n} k_i\lambda = \lim_{n\to\infty} \frac{\left(\sum_{i=1}^{n} k_i\right)}{\left(\sum_{i=1}^{n} k_i\right)^2}\ \lambda = \lim_{n\to\infty} \frac{\lambda}{\sum_{i=1}^{n} k_i} = 0$$

Por lo tanto, se cumplen las dos condiciones suficientes y, en consecuencia, $\hat{\lambda}_{MV}$ es un estimador consistente de λ.

B) Tenemos un **contraste de bondad de ajuste a una distribución totalmente especificada**. Los datos son los que aparecen a continuación:

Día	Lunes	Martes	Miércoles	Jueves	Viernes
Nº de alumnos	49	35	32	39	45

Bajo la hipótesis nula de que el número de alumnos que asisten a clase se distribuye uniformemente en los cinco días de la semana (i.e., es el mismo), tenemos que $p_i = 1/5 = 0.20,\ i = 1, 2, 3, 4, 5$. Dado que tenemos cinco días de la semana o cinco clases, $k = 5$. Con estos datos y para realizar el contraste, construimos la tabla:

Día	n_i	p_i	np_i	$\frac{(n_i-np_i)^2}{np_i}$
Lunes	49	0.20	40	2.025
Martes	35	0.20	40	0.625
Miércoles	32	0.20	40	1.600
Jueves	39	0.20	40	0.025
Viernes	45	0.20	40	0.625
	$n = 200$	1	$n = 200$	$z = 4.90$

Bajo la hipótesis nula de que el número de alumnos que asisten a clases se distribuye uniformemente en los cinco días de la semana, el estadístico $\sum_i \frac{(n_i-np_i)^2}{np_i} \sim \chi^2_{k-1}$, donde k es el número de días de la semana.

La regla de desisión es rechazar la hipótesis nula al nivel de significación aproximado del $\alpha = 5\%$ si:

$$z > \chi^2_{k-1\,,\,0.05} = \chi^2_{4,0.05}$$

En este caso:

$$z = 4.90 < 9.49 = \chi^2_{4,0.05},$$

por lo que, a un nivel de significación aproximado del $\alpha = 5\%$, no se rechaza la hipótesis nula.

C) Tenemos una distribución exponencial de parámetro λ (i.e., $X \in \exp(\lambda)$), y queremos contrastar la hipótesis nula $H_0 : \lambda = \lambda_0 = \frac{1}{2}$ frente a la hipótesis alternativa $H_1 : \lambda = \lambda_1 = 2$. Dado que tenemos una m.a.s. de tamaño $n = 10$, por el teorema de Neyman-Pearson, tenemos que la región más potente para este contraste estará dada por la que verifique, para una constante $k > 0$ la siguiente desigualdad entre las funciones de verosimilitud $L(\cdot)$ bajo las hipótesis nula y alternativa:

$$\frac{L(\lambda_0)}{L(\lambda_1)} \leq k$$

En este caso, tendremos que:

$$\frac{\lambda_0 e^{-\lambda_0 x_1} \cdots \lambda_0 e^{-\lambda_0 x_n}}{\lambda_1 e^{-\lambda_1 x_1} \cdots \lambda_1 e^{-\lambda_1 x_n}} \leq k \Longrightarrow \left(\frac{\lambda_0}{\lambda_1}\right)^n e^{(\lambda_1 - \lambda_0)\sum_{i=1}^n x_i} \leq k$$

Dado que $\lambda_1, \lambda_0 > 0$:

$$\Longrightarrow e^{(\lambda_1 - \lambda_0)\sum_{i=1}^n x_i} \leq k_1, \quad (k_1 > 0)$$

Tomando logaritmos neperianos y sabiendo que $\lambda_1 > \lambda_0$, tendremos que:

$$(\lambda_1 - \lambda_0)\sum_{i=1}^n x_i \leq k_2, \quad (k_2 > 0) \Longrightarrow \sum_{i=1}^n x_i \leq C, \quad (C > 0)$$

Por tanto, rechazaremos la hipótesis nula siempre que $\sum_{i=1}^n X_i \leq C$. Bajo la hipótesis nula, tenemos que $\lambda = \frac{1}{2}$, por lo que $X_i \in \exp(\frac{1}{2}) \equiv \gamma(\frac{1}{2}, 1)$. Así, tendremos que $\sum_{i=1}^n X_i \in \gamma(\frac{1}{2}, n) \equiv \chi^2_{2n}$. De esta forma, en el caso particular en que $n = 10$, tendremos que $\sum_{i=1}^{10} X_i \in \gamma(\frac{1}{2}, 10) \equiv \chi^2_{20}$. Así, al nivel de significación del 5%, tendremos que:

$$\alpha = 0.05 = P\left(\sum_{i=1}^{10} X_i \leq C \,\middle|\, \lambda = \frac{1}{2}\right) \Longrightarrow 0.05 = P\left(\chi^2_{20} \leq C\right) \Longrightarrow C = \chi^2_{20,0.95} = 10.9,$$

de donde, la región crítica más potente para este contraste será $\text{RC} = (0, 10.9]$.