Problemas de Inferencia Estadítica

Julio Muñoz

Departamento de Matemáticas, Universidad de Castilla-La Mancha
E-mail address: `julio.munoz@uclm.es`

Resumen. Se trata de una colección de unos 100 problemas de Estadítica.

Índice general

Preámbulo

Enunciados, indicaciones y resolución de ejercicios propuestos.

Parte 1

Problemas propuestos

CAPíTULO 1

Problemas del tema 1

1. Enunciados

1. La distribución por pesos de 70 alumnos de Ingeniería es la siguiente:

kgs.	frecuencia
$[50, 60]$	8
$[60, 70]$	15
$[70, 80]$	21
$[80, 90]$	18
$[90, 100]$	8

   Realizar un análisis descriptivo.
2. Se ha realizado la determinación del contenido de Potasio en la bioquímica automatizada de la sangre de 25 pacientes embarazadas. Los resultados han sido 2.7, 3.3,4.1, 3.2, 3.1, 3.3, 3.4, 2.7, 2.8, 2.7, 3.2, 2.3, 4.2, 3.5, 3.6, 3.7, 3.2, 3.3, 2.8, 3.5, .3,8 3.1, 3.2, 3.6, 2.4. Llevar a cabo un análisis descriptivo de esta muestra.
3. Demostrar las siguientes relaciones entre los momentos:
   a) $m_3 = \alpha_3 - 3\alpha_1\alpha_2 + 2(\alpha_1)^3$.
   b) $m_4 = \alpha_4 - 4\alpha_3\alpha_1 + 6(\alpha_1)^2\alpha_2 - 3(\alpha_1)^4$.
4. Sean A, B y C sucesos. Se pide
   a) Calcular $P(A \cup B \cup C)$
   b) Demostrar que
   $$P(A \cup B | C) = P(A | C) + P(B | C) - P(A \cap B | C)$$
   supuesto que $P(C) > 0$.
   c) Demostrar las desigualdades
   $$P(A \cap B) \leq P(A) \leq P(A \cup B) \leq P(A) + P(B).$$
5. Demostrar que si A y B son sucesos independientes entonces también lo son A y $\overline{B}$. ¿$\overline{A}$ y $\overline{B}$?. ¿Es verdad que $P(A \cup B) = 1 - P(\overline{A})P(\overline{B})$?
6. Cuando la concentración en sangre de una determinada sustancia en un individuo sobrepasa cierto valor se dice que el individuo pertenece a cierto grupo de riesgo. Supongamos que el 5 % de los habitantes de una población pertenecen al grupo de riesgo. Determinar la probabilidad de que de tres individuos elegidos al azar, uno esté en el grupo de riesgo.

7. Determinar la probabilidad del suceso consistente en extraer una bola blanca de una urna que contiene 5 bolas de dos colores, blanco y negro. Se supone que las distintas composiciones de la urna son igualmente probables.
8. Un dispositivo de control consta de dos aparatos. La probabilidad de que falle el k-ésimo es $1 - \alpha_k$. Estimar la probabilidad de que ambos funcionen bien si:
   a) si el funcionamiento de dichos aparatos es independiente,
   b) si no se sabe si hay o no independencia entre el funcionamiento de los dos aparatos.
9. En una universidad el 4 % de los chicos y el 1 % de las chicas tienen una altura superior a 180 cm. El 60 % de los alumnos son chicas. Se toma un alumno al azar y se comprueba que mide más de 180 cm. Hallar la probabilidad de que tal alumno sea chico.
10. El volumen de ventas en un concesionario de coches es de 500 unidades (al año) para el deportivo, 1000 para el familiar y 2000 para el utilitario. Se sabe que el porcentaje de coches defectuosos es de un 2 % para el primero, un 1 % para el segundo y de un 3 % para el tercero. Se pide calcular la probabilidad del siguiente suceso:
    a) Elegido uno al azar, que éste no sea defectuoso.
    b) Habiendo elegido un automóvil defectuoso, que dicho automóvil pertnezca al grupo de los familiares.
11. Se disponen dos urnas U_1 y U_2. La urna U_1 contiene el 70 % de bolas blancas y el 30 % de negras, y en la urna U_2 hay un 30 % de blancas y un 70 % de negras. Se selecciona una urna al azar (se supone que ambas tienen la misma probabilidad de ser elegidas) y se toman 10 bolas una tras otra con reemplazamiento. El resultado fue el suceso $S = bnbbbbnbbb$, siend b bola blanca y n bola negra. ¿Cuál es la probabilidad de que el suceso S proceda de la urna U_1?
12. Una empresa proyecta introducir un nuevo artículo y conjetura que con una probabilidad del 50 % será adquirido por un segmento del 5 % de sus clientes habituales, con una probabilidad de 25 % por el 10 % de sus clientes habituales, con un 20 % de probabilidad por un 15 % de sus clientes y con probabilidad del 5 % por el 20 %. Para asegurarse de ello recogen información preguntando a 20 de esos clientes, siendo 3 de éstos los que estarían dispuestos a comprar el nuevo artículo. ¿Cuál sería la probabilidad de que el 15 % comprara el mencionado artículo?

13. Sean 10 urnas. Nueve de ellas contienen 2 bolas blancas y dos negras y una contiene cinco blancas y una negra. Se elige una urna de forma aleatoria y sea extrae una bola resultando ser una blanca. ¿Cuál es la probabilidad de que la urna elegida sea la que posee 4 blancas y una negra?
14. La probabilidad de que tres jugadores de dardos hagan diana es 4/5, 3/4 y 2/3. Supongamos que tiran los tres y se producen dos aciertos sobre la diana. ¿Cuál es la probabilidad de que haya fallado el tercero?
15. Sea $\Omega = (0, +\infty)$. Demostar que si definimos P sobre todos los intervalos $(a, b) \subset \Omega$ con ayuda de la fórmula

$$P((a,b)) = \int_a^b e^{-x} dx,$$

entonces P cumple todos los axiomas de la probabilidad.

CAPíTULO 2

Problemas del tema 2

1. Enunciados

1. Comprobar que si $F_\xi(x)$ es la función de distribución de la v.a. ξ entonces
   a) $P(a < \xi < b) = F_\xi(b) - F_\xi(a) - P_\xi^*(b)$
   b) $P(a \leq \xi < b) = F_\xi(b) - F_\xi(a) - P_\xi^*(b) + P_\xi^*(a)$
2. En las mismas condiciones que en el ejercicio precedente demostrar que para todo $x \in \mathbf{R}$
$$P_\xi^*(x) = 0$$
supuesto que ξ es una v.a. de tipo continuo.
3. Sea ξ una v.a. tal que se cumple la siguiente tabla de probabilidades:

x_i	1	2	3	4	5
$P_\xi^*(x_i)$	2/8	1/8	2/8	2/8	1/8

Se pide:
   a) Probar que $P_\xi^*(x_i)$ es una función de densidad de masa.
   b) Dibujar la función de distribución de ξ.
   c) Calcular la probabilidad
$$P(1.1 < \xi < 3.3).$$
4. Dada la v.a. ξ cuya densidad de probabilidad es
$$f_\xi(x) = \begin{cases} k\sin(x) & \text{si } x \in \left[0, \frac{\pi}{2}\right] \\ 0 & \text{en el resto,} \end{cases}$$
donde k es una constante, se pide:
   a) Determinar la constante k para que $f_\xi(x)$ sea efectivamente una densidad de probabilidad.
   b) Determinar la función de distribución asociada.
   c) Calcular $P_\xi^*((\pi/4, \pi/2))$.
5. Sea la función $f_\xi(x) = \frac{C}{3}$ si $x \in [0,3]$, e $= 0$ si $x \notin [0,3]$. Se pide
   a) Determianr C para que f_ξ sea una función de densidad de probabilidad.
   b) Determinar su función de distribución.
   c) Calcular
$$P(1.5 < \xi < 3.5).$$
6. Determinar el parámetro C para que las siguientes funciones sean funciones de densidad de probabilidad

a)

$$f_\xi(x) = \begin{cases} Cx^2 & \text{si } x \in [0{,}1] \\ 0 & \text{en el resto.} \end{cases}$$

b)

$$f_\xi(x) = \begin{cases} C\exp(-Cx) & \text{si } x \in [0{,}\infty] \\ 0 & \text{en el resto.} \end{cases}$$

c)

$$f_\xi(x) = \frac{C}{1+x^2}, \quad x \in \mathbf{R}$$

7. Determinar el parámetro C para que las siguientes funciones sean funciones de densidad de masa
   a) $P_\xi^*(x) = C\frac{x-1}{n}$, $\quad x = 2, ..., n$ y $P_\xi^*(x) = 0$ para cualquier otro valor de x.
   b) $P_\xi^*(x) = C\frac{m^x}{x!}$, $\quad x = 0, 1, 2, ..., n, ...$ y $P_\xi^*(x) = 0$ para cualquier otro valor de x.[1]
   c) $P_\xi^*(x) = C(\frac{1}{3})^x$, $\quad x = 0, 1, 2, ..., n, ...$ y $P_\xi^*(x) = 0$ para cualquier otro valor de x.
8. Calcular la esperanza y desviación típica de cada uno de los apartados de los ejercicios 6 y 7.
9. Sea ξ v.a. de tipo continuo cuya función de densidad de probabilidad es

$$f_\xi(x) = \begin{cases} 2e^{-2x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}.$$

   a) Hallar la función de distribución de ξ
   b) Calcular la siguiente probabilidad condicionada

$$P\left\{ \xi > 2|_{\xi<4} \right\}.$$

10. Sea la variable aleatoria ξ cuya función de densidad de probabilidad es

$$f_\xi(x) = \begin{cases} x & \text{si } x \in [a,b] \\ 0 & \text{en el resto.} \end{cases}$$

   Se supone que a y b son números tales que $0 < a < b$.
   a) Calcular el valor de a y de b sabiendo que $P\{a \leq \xi \leq 2a\} = \frac{3}{8}$.
   b) Determinar la función de distribución de ξ.
   c) Calcular la probabilidad condicionada $P\left(T|_Q \right)$ siendo T y Q los sucesos $\left\{ \frac{a}{2} \leq \xi \leq \frac{3a}{2} \right\}$ y $\{a \leq \xi \leq 2a\}$ respectivamente.
11. Se supone que la hospitalización semanal de enfermos en un gran centro de salud sigue una distribución cuya densidad es

$$f_\xi(x) = \begin{cases} \frac{2}{9}(3x - x^2) & \text{si } x \in [0,3] \\ 0 & \text{en el resto,} \end{cases}$$

[1] Téngase en cuenta que $\sum_{r=0}^{\infty} \frac{m^r}{r!} = e^m$.

con x expresado en centenas. ¿Qué cantidad de camas libres debe tener el centro al comienzo de cada semana, si se pretende hospitalizar a todos los enfermos que lo precisen durante dicha semana, con una probabilidad de 0.5?

12. Se supone que la variable aleatoria ξ, que expresa en miles de años el tiempo que permanece activo cierto tipo de deshecho radioactivo, está distribuida con arreglo a la función de densidad de probabilidad
$$f_\xi(x) = \begin{cases} \lambda(1-x) & \text{si } x \in (0,1) \\ 0 & \text{resto} \end{cases}.$$
    a) Determinar λ para que f_ξ sea efectivamente una densidad de probabilidad.
    b) Calcular la probabilidad de que la actividad de un deshecho dure entre 300 y 500 años.
    c) Si para un determinado deshecho se sabe que su actividad dura más de 600 años, ¿cuál será la probabilidad de que dure entre 400 y 950 años?
    d) Calcular el tiempo esperado de actividad para el mencionado tipo de deshechos radioactivos.
13. Sea la variable aleatoria η_n que representa la suma de los puntos que salen al lanzar n veces un dado.
    a) Calcular $E[\eta_n]$ y $V[\eta_n]$.
    b) Utilizar la desigualdad de Tchevichev para encontrar un natural $n \in \mathbf{N}$ de modo que
$$P\left\{\left|\frac{\eta_n}{n} - 3.5\right| \geq 0.1\right\} \leq 0.1.$$
14. Sea la función
$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \sqrt[2]{x} & \text{si } x \in [0,1] \\ 1 & \text{si } x > 1. \end{cases}$$
Comprobar que es una función de distribución de probabilidad. Determinar su densidad asociada. Calcular, si es posible, la esperanza y varianza.
15. Idem con
    a)
$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 2/3 & \text{si } x \in [0,5] \\ 1 & \text{si } x > 1. \end{cases}$$
    b)
$$F(x) = \begin{cases} 1 - \frac{1}{1+x} & \text{si } x \leq 0 \\ 0 & \text{si } x \geq 0 \end{cases}$$
16. Sea la v.a. de tipo discreto ξ cuya función de densidad de masa es
$$P_\xi^*(x) = p(1-p)^{x-1}, \quad x = 1, 2, 3, ...$$
donde $p \in (0,1)$. Determinar $E[\xi]$ y $V[\xi]$.
17. Idem si
    a) $P_\xi^*(0) = 1/3$ y $P_\xi^*(1) = 2/3$.

$b)$ $P_{\xi}^{*}(x) = \left(\begin{array}{c} n \\ x \end{array} \right) p^{x}(1-p)^{n-x}, \quad x = 0, 1, 2, 3, ...n,$ donde $p \in (0,1)$.

CAPíTULO 3

Problemas del tema 3

1. Enunciados

1. La aplicación de determinado producto químico en el tratamiento de aguas residuales produce cierta mejoría en el 70 % de los casos. Se aplica tal producto en diez plantas de tratamiento de aguas. Se pide:
   a) La probabilidad de que mejore el agua en 4 de las plantas.
   b) La probabilidad de que al menos mejore en 3 de las plantas.
2. Una partícula se mueve en un fluido de manera rectilínea, digamos que lo hace a lo largo del eje coordenado OX del plano $\mathbb{R}^2$. Sobre este movimiento se sabe lo siguiente: al cabo de un segundo la partícula se mueve a la derecha 0.001 mm con una probabilidad de $\frac{1}{60}$ y -0.001 mm a la izquierda con una probabilidad $1-\frac{1}{60}$. Se supone además que cada movimiento en un segundo es independiente del movimiento realizado en segundos anteriores. Se pide:
   a) Calcular la probabilidad de que la partícula realice 15 movimientos a la derecha al cabo de 360 segundos.
   b) Calcular la probabilidad de que la partícula, al cabo de 360 segundos, esté en el punto $(-0.3, 0)$ si se supone que en el instante inicial (justamente antes de empezar a contabilizar los 360 segundos) se encontraba en el origen de coordenadas $(0,0)$.
   c) Encontrar la posición esperada de la partícula al cabo de los 360 segundos.
3. La probabilidad de que se produzca un choque entre una partícula de un fluido y las paredes del recinto en el que esta confinado es 0.002. Sea ψ la v.a. que representa el número de choques en un grupo de 1200 partículas. Se pide:
   a) $P(\psi \leq 5)$
   b) $P(\psi = 7)$
4. Una jaula de laboratorio contine 25 ratones, de los que 8 son blancos y el resto pardos. Un ayudante de laboratorio con los ojos vendados extrae sin reemplazaminto 4 ratones de la jaula. Calcular la probabilidad que sólamente uno de los ratones sea blanco.
5. Si X es una v.a. con distribución $N(2,3)$,calcúlese:
   a) $P(X \leq 6.32)$ y $P(6.15 \leq X \leq 7.35)$.
   b) $P(X^2 \geq 3.15)$.
6. Si X es una variable aleatoria con distribución $N(3, 0.5)$, calcúlese:
   a) $P(X \leq 3.32)$ y $P(2.15 \leq X \leq 3.35)$.
   b) $P(X^2 \geq 9)$.
7. Supongamos que el número de llamadas que recibe una centralita en 5 minutos viene dado por una v.a. ψ tal que $\psi \sim \mathcal{P}(5)$. Se piden las siguientes probabilidades:

a) De tener 6 llamadas en 5 miniutos
b) De tener 3 en 10 minutos.

8. Una fábrica produce fusibles eléctricos, resultando defectuosos el 3 %. Los fusibles se empaquetan en cajas de 24 unidades. Se pide:
   a) Calcular la probabilidad de que una caja elegida al azar contenga al menos un fusible defectuoso.
   b) Si selecionamos 5 cajas al azar, ¿cuál es la probabilidad de que exactamente en 2 de ellas no haya ningún fusible defectuoso?
   c) Vamos selecionando cajas al azar y comprobando si sus fusibles son defectuosos o no. ¿Cuál es el número medio de cajas que tendremos que inspeccionar hasta encontar algún fusible defectuoso?
9. Las velocidades de dos partículas dentro de un fluido vienen dadas por lasvariables aleatorias ξ_1y ξ_2 respectivamente. Se supone que la función de densidad de probabilidad para ξ_1 es
$$f_{\xi_1}(x) = \begin{cases} 2 & \text{si } x \in (0, 0.5) \\ 0 & \text{en el resto,} \end{cases}$$
y que para todo $x \in (0, 0{,}5)$ la densidad de probabilidad de ξ_2 condicionada por la primera es
$$f\left(y|\, x\right) = \begin{cases} \frac{1}{2} + x\left(y-1\right) & \text{si } y \in (0,2) \\ 0 & \text{en el resto.} \end{cases}$$
Se pide:
   a) Determinar la función de densidad de probabilidad conjunta del vector aleatorio (ξ_1, ξ_2).
   b) Hallar las densidades marginales del vector aleatorio (ξ_1, ξ_2).
   c) ¿Qué velocidad se espera para la segunda partícula cuando la velocidad de la primera es de 0. 75 unidades?
10. Dado el vector aleatorio $\xi = (\xi_1, \xi_2)$ cuya ley de probabilidades viene dada a través de la función de densidad de probabilidad
$$f_\xi\left(x, y\right) = \begin{cases} x + y & \text{si } x \in (0,1) \text{ e } y \in (0,1) \\ 0 & \text{en el resto,} \end{cases}$$
se pide:
   a) Calcular la probabilidad del suceso $\{\xi_1 \leq 0.5, \xi_2 \leq 0.2\}$.
   b) ¿Son ξ_1 y ξ_2 independientes?
   c) Calcular $E\left[\left.\xi_1\right|_{\xi_2=0.5}\right]$.
11. Sean, ξ_1 la v.a. que representa a los gastos por impagados en millones de euros de una determinada empresa, y ξ_2 la v.a. de los ingresos netos. Se sabe que la función de densidad del vector (ξ_1, ξ_2) es
$$f_{(\xi_1,\xi_2)}\left(x, y\right) = \begin{cases} 5(1+x)e^{20(2-y)} & \text{si } x \in (0,2) \text{ e } y > 2 \\ 0 & \text{en el resto} \end{cases}$$

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?

12. Se consideran las variables aleatorias siguientes relativas a la incidencia de un fuego en dos áreas determinadas, a saber:

$$\xi_1 = \text{intensidad del fuego}$$
$$\xi_2 = \text{zona en la que se inicia el fuego}$$

Se supone que ξ_1 toma los valores 1, cuando el fuego es menor, 2 cuando causa serios problemas a la fauna y flora del lugar, y 3 cuando es devastador. La variable ξ_2 toma los valores 0 y 1 según sea la zona de rastrojo o de arboleda en la que se inicia el fuego. Teniendo en cuenta la tabla de probabilidades relativa a estas dos variables aleatorias es

$\xi_2\backslash\xi_1$	$x_1 = 1$	$x_2 = 2$	$x_3 = 3$
$y_1 = 0$	0.1	0.3	0.2
$y_2 = 1$	0.06	0.18	0.16

contestar de manera razonada a las siguientes cuestiones:
a) ¿Son independientes la variable intensidad del fuego ξ_1 y la variable zona la que se inicia ξ_2?
b) Encontrar la esperanza de la variable ξ_1.
c) Hallar la esperanza condicionada de $\xi_1 | \xi_2 = 0$.

13. La duración, en segundos, precisa para que se produzca una reacción entre dos compuestos sigue una distribución $N(\mu, \sigma)$. El 60 % de las veces dura más de 40 segundos y el 55 % dura menos de 50. Hallar μ y σ.
14. Se sabe que la distribución ξ de los coeficientes intelectuales de los alumnos de un colegio sigue la ley nomal. Se sabe que $P(\xi \geq 1{,}4) = 0.1056$ y que $P(\xi > 1) = 0.4013$. Calcular los parámetros μ y σ de la citada distribución.
15. Se sabe que $\xi_1 \sim N(1, 0.1)$, que $\xi_2 \sim N(\mu, \sigma)$ que son dos v.a. independientes entre sí y que $\xi_1 + \xi_2 \sim N(1.5, 0.3)$. Determinar μ y σ
16. La duración media sin rotura de las sábanas de un Sanatorio de 300 camas es de 50 días con una desviación típica de 8. Si suponemos que la duración de las sábanas sigue una distribución normal, calcular:
a) ¿Cuántas sábanas se habrán tenido que reponer antes de los 35 días?
b) ¿Cuántas sábanas habrá que reponer pasados 60 días?

CAPíTULO 4

Problemas del tema 4

1. Enunciados

1. Calcular los siguientes valores de la variable que sigue una $\mathcal{F}$ de Fisher-Snedecor: $\mathcal{F}_{5,21,0.05}$, $\mathcal{F}_{24,12,0.01}$, $\mathcal{F}_{8,4,0.05}$.
2. Calcular las siguientes probabilidades:
   a) $P(\mathcal{X}^2_{10} \leq 20.4)$
   b) $P(\mathcal{X}^2_{14} > 23.2)$
   c) $P(11{,}9 \leq \mathcal{X}^2_{18} \leq 32.4)$
3. Calcular $\mathcal{X}^2_{30,0.025}$ y $\mathcal{X}^2_{24,0.95}$
4. Dada una población que sigue una distribución normal con media conocida y desvicaión típica desconocida, se extrae una muesra de tamaño 16. Calcular probabilidad $P(0.5 < \dfrac{S^2}{\sigma^2} < 1.8)$.
5. Calcular $t_{15,0.1}$, $t_{25,0.2}$ y $t_{8,0.05}$.
6. La producción diaria de un determinado artículo se supone de tipo uniforme y oscila entre 6.000 y 10.000 unidades. Determinar la probabilidad de que la producción media supere las 8.100 unidades, habiéndose realizado observaciones durante 320 días, y supuesto que el número de unidades producidas en un día es independiente de los restantes.
7. Sea la variable aleatoria η_n que representa la suma de los puntos que salen al lanzar n veces un dado.
   a) Calcular $E[\eta_n]$ y $V[\eta_n]$.
   b) Utilizar el Teorema Central del Límite para encontrar un natural n, lo más pequeño posible, de modo que
   $$P\left\{\left|\frac{\eta_n}{n} - 3.5\right| \geq 0.1\right\} \leq 0.1.$$
8. Encuéntrese un número k tal que
   $$P(490 < \xi < k) = 0.5$$
   si ξ = número de caras obtenido al lanzar una moneda 1000 veces.
9. Sean $X_1, ..., X_{50}$ m.a.s. de distribución $U(0,1)$ (uniforme en $(0,1)$). Consideramos el estadístico $\overline{X} = \frac{1}{50}\sum_{i=1}^{50} X_i$. Enunciar con todo detalle el Teorema Central del Límite y aplicarlo al cálculo de
   $$P\left\{\overline{X} < 0.4\right\}.$$

10. La longitud en milímetros ξ que se puede estirar un filamento de nylon sin ruptura sigue una distribución de probabilidad cuya densidad de probabilidad es
$$f_\xi(x) = \begin{cases} 5e^{-5x} & \text{si } x \geq 0 \\ 0 & \text{en el resto} \end{cases}$$
Emplear el Teorema Central del Límite para determinar de manera aproximada, la probabilidad de que la longitud promedio de 100 filamentos esté comprendida entre 0. 18 y 0. 22 (Ayuda: $\alpha_1 = 1/5$ y $\alpha_2 = \frac{2}{25}$).
11. Se sabe que en un banco la probabilidad de recibir un cheque sin fondos a lo largo de una semana es 0. 15. Si durante una semana se espera recibir 1000 cheques, se pide:
    a) Calcular de manera aproximada, utilizando el Teorema Central del Límite para ello, la probabilidad de que en tal semana se reciban como máximo 125 cheques sin fondos.
    b) Si suponemos que por cada cheque sin fondos la entidad bancaria se beneficia en 200 pesetas, determinar el número de cheques que deben recibir en una semana para que con una probabilidad de 0. 9, el banco, por ese concepto, se beneficie en 10000 pesetas.
12. De una población $N(\mu, \sigma)$ y una muestra aleatoria simple $\xi_1, ..., \xi_n$, se pretende estimar su varianza σ^2 a través del estadístico
$$\xi^* = \frac{\sum_{i=1}^n \xi_i^2}{n}.$$
Se pide:
    a) Calcular la esperanza de ξ^*.
    b) ¿En qué condiciones ξ^* es insesgado?
    c) ¿Cuándo es ξ^* eficiente?
    (Ayuda: usar que $E[\chi_1^2] = 1$ y que $V[\chi_1^2] = 2$)
13. Una comapañía aérea sabe por experiencia que el 12 % de las reservas telefónicas de plazas no se llevan a efecto, de modo que reserva más plazas de las que dispone. Si en un vuelo hay 150 plazas, ¿cuántas reservas puede hacer la compañía para que la probabilidad de cubrir al menos 145 plazas sea del 99 %? Si la compañía reserva 160 plazas, ¿cuál es la probabilidad de que, al menos un pasajero no tenga plaza disponible a la hora de embarcar?
14. Sea una v.a. ξ que representa el número de veces que encontramos infección en un determinado arbusto, al cabo de n inspecciones y supuesto que la probabilidad de infección es $p = 0.65$. Si suponemos que el número de inspecciones, n, es suficientemente grande, determinar un n para que se cumpla $P\{\xi > n/2\} \geq 0.9$.

CAPíTULO 5

Problemas del tema 5

1. Enunciados

1. La distancia ξ entre un árbol cualquiera y el árbol más próximo a él en un bosque sigue una distribución con función de densidad de probabilidad
$$f_\xi(x) = \begin{cases} 2\theta x \exp\left(-\theta x^2\right) & \text{si } x \geq 0 \\ 0 & \text{en el resto} \end{cases}$$
($\theta > 0$). Obtener el estimador de máxima verosimilitud de θ, supuesto que se realiza una m. a. s. de tamaño n.
2. Idem para
$$f_\xi(x) = \begin{cases} \frac{x^2}{2\theta^3} \exp\left(-x/\theta\right) & \text{si } x \geq 0 \\ 0 & \text{en el resto} \end{cases}$$
($\theta > 0$)
3. El coseno ξ del ángulo con el que se emiten los electrones en un proceso radiactivo es una variable aleatoria con densidad de probabilidad
$$f_\xi(x) = \begin{cases} \frac{1+\theta x}{2} & \text{si } x \in [-1,1] \\ 0 & \text{en el resto} \end{cases}$$
donde $\theta \in [-1,1]$. Dada una m.a.s. $\xi_1, ..., \xi_n$, encontrar el estimador de θ por el método de los momentos.
4. Una variable aleatoria ξ tiene definida su densidad según las igualdades
$$P_\xi^*(-1) = \frac{1-\theta_1}{2},\ P_\xi^*(0) = \frac{\theta_1+\theta_2}{2},\ P_\xi^*(1) = \frac{1-\theta_2}{2},$$
con $\theta_1,\ \theta_2 \in (0,1)$. Estimar el parámetro $\theta = (\theta_1, \theta_2)$ por el método de los momentos.
5. Idem para
$$f_\xi(x) = \begin{cases} \theta x^{\theta-1} & \text{si } x \in [0,1] \\ 0 & \text{en el resto} \end{cases}$$
con $\theta > 0$.
6. Sea $\xi_1, ..., \xi_n$ una m.a.s. de tamaño n y cuya distribución viene definida como
$$P_{\xi_i}^*(x) = p(1-p)^{x-1}, \quad x = 0,1,2,...$$
para $i = 1,2,...,n$ ($0 < p \leq 1$). Se pide
   a) Estimador de p por el método de los momentos.

b) Estimador de p por el método de la máxima verosimilitud. (Ayuda: $E[\xi_i] = \frac{1}{p}$.)

7. Idem con el caso en que $\xi_1, ..., \xi_n$ una m.a.s. de tamaño n y cuya distribución viene definida por la fucnión de densidad

$$f_\xi(x) = \begin{cases} e^{\theta - x} & \text{si } x \in [\theta, +\infty] \\ 0 & \text{en el resto} \end{cases}$$

(se supone $\theta > 0$).

8. Sea la variable aleatoria ξ cuya función de densidad de probabilidad es

$$f_\xi(x) = \begin{cases} \dfrac{\theta}{x^{\theta+1}} & \text{si } x \geq 1 \\ 0 & \text{en el resto,} \end{cases}$$

donde se supone que $\theta > 1$. Dada una muestra aleatoria simple de ξ y tamaño n, $\xi_1, ...\xi_n$, se pide:

a) Determinar un estimador del parámetro θ por el método de los momentos.

b) Obtener un estimador para θ por el método de máxima verosimilitud.

9. Sea ξ una variable aleatoria que modeliza a cierto fenómeno. Se sabe que

$$P\{\xi = -1\} = \frac{1-\theta}{2},\ P\{\xi = 0\} = \frac{1}{2},\ P\{\xi = 1\} = \frac{\theta}{2},$$

donde θ es cierto parámetro del que sólo se sabe que $\theta \in (0, 1)$.

a) Dada una m.a.s. $\xi_1, ..., \xi_n$ de la v.a. ξ, determinar el estimador de máxima verosimilitud del parámetro θ.

b) Si se toma $n = 50$ y se realiza la muestra obteniéndose los resultados,

x_i :	-1	0	1
n_i :	10	25	15

donde n_i es la frecuencia del dato x_i (número de veces que aparece x_i), realizar una estimación del parámetro θ usando para ello el método de la máxima verosimilitud.

10. Sea una población modelizada con ayuda de la variable aleatoria ξ, de la que se sabe que su densidad de probabilidad es

$$f_\xi(x) = \begin{cases} x^\theta(\theta + 1) & \text{si } x \in [0, 1] \\ 0 & \text{en el resto,} \end{cases}$$

con θ un parámetro tal que $\theta > -1$. Dada una muestra aleatoria simple, hallar un estimador para θ empleando para ello el principio de máxima verosimilitud.

11. Deducir de manera razonada un intervalo de confianza para la varianza de una población que se distribuye normalmente y donde se supone que la esperanza es desconocida. Aplica el resultado obtenido al siguiente ejemplo: un metalúrgico ha hecho 4 determinaciones del punto de fusión del manganeso, 1269, 1271, 1263 y 1265 grados C. Tras valorar los resultados y los posibles errores decidió que la varianza habría de ser 1 ó 2. ¿ Es posible esta afirmación con los datos experimentales obtenidos, suponiendo normalidad y a un nivel de confianza del 95 % ?

12. Sobre la base de la realización de una muestra aleatoria simple de 81 observaciones, los expertos de seguridad estimaron que el tiempo de reacción de los camioneros ante una luz roja era, en media, 2 segundos, con una cuasivarianza muestral de $(0.60)^2$, esto es
$$\frac{\sum_{i=1}^{144} x_i}{81} = 2 \quad \text{y} \quad \frac{\sum_{i=1}^{144} (x_i - 2)^2}{80} = (0.60)^2 .$$
Se desea calcular un intervalo de confianza del 99 % para el tiempo medio de reacción:
    a) Haciendo la hipótesis de normalidad y tomando como varianza poblacional la cuasivarianza muestral.
    b) Haciendo hipótesis de normalidad y suponiendo desconocida la desviación típica poblacional.
13. Se sostiene la hipótesis de que la demanda por hora (en barriles) de petróleo por parte de la Comunidad Económica Europea (CEE) a los países árabes, posee una desviación típica de 20, mientras que la de EEUU a este mismo grupo de países tiene como desviación típica 10. Se ha realizado una muestra aleatoria simple de tamaño 125 (125 horas) para la demanda de la CEE, resultando que la demanda media ha sido de 300 barriles. También se ha realizado un muestreo aleatorio simple de tamaño 100 para la demanda de EEUU, la media ha sido de 250. Suponiendo que la demanda diaria de petróleo (tanto de la CEE como de EEUU) está regida por una ley de probabilidad normal, elaborar con todo detalle un intervalo de confianza del 95 % para la diferencia de las demandas medias. Nota: se supone que las demandas de petróleo por parte de EEUU y la CEE son independientes.
14. Una población está representada por una variable aleatoria ξ, cuya función de distribución es
$$F_\xi(x) = \begin{cases} 1 - \left(\dfrac{1}{x}\right)^\theta & \text{si } x > 1, \\ 0 & \text{en el resto,} \end{cases}$$
y donde θ es un parámetro real positivo. Sea una muestra aleatoria simple $\xi_1, ..., \xi_n$ perteneciente a dicha población. Se pide:
    a) Determinar la función de densidad de probabilidad asociada a la variable aleatoria ξ.
    b) Encontar un estimador del parámetro θ usando para ello el principio de máxima verosimilitud.
    c) Idem usando el método de los momentos.
15. Sea ξ una variable aleatoria cuya función de densidad es
$$f_\xi(x) = \begin{cases} \frac{2a}{1-a} x^{\frac{2a}{1-a}-1} & \text{si } x \in [0,1], \\ 0 & \text{en el resto.} \end{cases}$$
donde a es un parámetro del cual se sabe que es.positivo. Sea $\xi_{1,}\xi_2, ..., \xi_n$ una muestra aleatoria simple de ξ. Se pide:
    a) Hallar un estimador de máxima verosimilitud para el parámetro a.
    b) Obtener por el método de los momentos un estimador para a.

16. Una empresa desea determinar la proporción de clientes dispuestos a adquirir uno de sus productos. Estima que dicha proporción es 0.4 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.
17. Sea la v.a. ξ cuya función de densidad de probabilidad es
$$f_{\xi}(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}} & \text{si } x \geq 0 \\ 0 & \text{en el resto.} \end{cases}$$
    a) Dada una m.a.s. $\xi_1, ..., \xi_n$ de la v.a. ξ, determinar por el método de máxima verosimilitud un estimador del parámetro θ.
    b) ¿Es insesgado el estimador obtenido en el apartado anterior?
    c) ¿Cuál es el riesgo para este estimador?
18. Sea una v.a. ξ cuya densidad de probabilidad es
$$f(x) = \begin{cases} (\theta + 1) x^{\theta} & \text{si } x \in (0, 1) \\ 0 & \text{en el resto,} \end{cases}$$
con θ cierto parámetro desconocido del que sólo se sabe que $\theta > -1$. Dada una muestra aleatoria simple $\xi_1, \xi_2, ..., \xi_n$ se pide:
    a) Un estimador de θ por el método de la máxima verosimilitud.
    b) Un estimador de θ por el método de los momentos.

CAPíTULO 6

Problemas del tema 6

1. Enunciados

1. Se toma una muestra de tamaño 1 en una población distribuida según una $\mathcal{P}(\lambda)$. Para contrastar $H_0 : \lambda = 1$ contra $H_1 : \lambda = 2$ se considera el siguiente contrate no aleatorizado:
$$\varphi(x) = \begin{cases} 1 & \text{si } x > 3 \\ 0 & \text{si } x \leq 3 \end{cases}$$
Hallar las probabilidades para los errores de tipo I y II y la potencia del contraste. ¿Cómo deberíamos reformular el contraste si queremos un **nivel** 0.05?
2. Sea una muestra de tamaño n tomada de una población $N(\mu, 1)$. Se quiere contrastar $H_0 : \mu \leq \mu_0$ contra $H_1 : \mu > \mu_0$ mediante el contraste
$$\varphi(x_1, ..., x_n) = \begin{cases} 1 & \text{si } \overline{x} > \mu_0 + \frac{Z_\alpha}{\sqrt[2]{n}} \\ 0 & \text{si } \overline{x} \leq \mu_0 + \frac{Z_\alpha}{\sqrt[2]{n}} \end{cases}$$
Hallar la potencia del contraste.
3. Una caja contiene 10 trozos de mármol de igual tamaño, de los cuales M son blancos y $10 - M$ son negros. Se quiere contrastar $M = 5$ frente a $M = 6$, y para ello se extrae una muestra de tamaño 3 sin reemplazamiento. La hipótesis nula se rechaza si la muestra contiene 2 ó 3 mármoles blancos, y en otro caso caso se acepta. Hallar el nivel de significación y la potencia de este contraste.
4. Utiliza el Lema de Neyman-Pearson para llevar a cabo el contrate, a nivel α, de $H_0 : \lambda = 1$ contra $H_1 : \lambda = \lambda_0 (> 1)$ basada en una muestra de tamaño 1 si la densidad de la población es
$$f_\xi(x) = \begin{cases} \lambda x^{\lambda - 1} & \text{si } 0 < x < 1 \\ 0 & \text{en el resto.} \end{cases}$$
5. Utiliza el Lema de Neyman-Pearson para llevar a cabo el contrate, a nivel α, de $H_0 : \theta \leq \theta_0$ contra $H_1 : \theta > \theta_0 (> 1)$ basada en una muestra de tamaño n si la densidad de la población es
$$f_\xi(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta} & \text{si } x > 0 \\ 0 & \text{en el resto.} \end{cases}$$
(se supone $\theta > 0$).
6. Se sabe que cuando se usa en determinado tipo de circuito el tiempo que tarda fallar un cebador electrónico es una v.a. de tipo exponencial con parámetro λ. Se someten a prueba 1000 de estos cebadores y se da que la suma de los tiempos en que fallan es de

109.652 horas. En base a estas observaciones llevar a cabo un contraste de $H_0 : \lambda \leq 0.008$ contra $H_1 : \lambda > 0.008$ a un nivel de significación $\alpha = 0.05$.

7. Tomamos 10 muestras de agua y determinamos en miligramos la cantidad de un determinado contaminante, resultando de estas mediciones la siguiente tabla:

muestra	1	2	3	4	5	6	7	8	9	10
contenido en mg.	5.2	4.9	5	5.1	5.2	4.8	4.9	5.3	4.6	5.4

Se supone que la desviación típica de la población es 0.10 miligramos y se pretende saber si los valores observados son compatibles con la media aritmética $\mu = 5$, todo ello suponiendo que la población se distribuye según una normal.

8. La publicidad de un determinado producto farmacéutico indica que reduce el peso. Doce individuos decidieron tomar dicho producto sin alterar el resto de la dieta alimenticia que estaban siguiendo. Los cambios en el peso que sufrieron los 12 individuos están reflejados en la tabla adjunta:

Individuo	1	2	3	4	5	6	7	8	9	10	11	12
variación peso	0.2	0	1	0.6	-0.5	-0.6	-1	0.6	1	0.5	-0.4	-0.5

Teniendo en cuenta estos datos decidir si se puede admitir como verdadera la indicación de tal anuncio. Se supone que la variación del peso sigue una distribución normal.

9. El nivel medio de protombina en una población noral es conocido y resulta ser aproximadamente de 20 mg/100ml de plasma. Se toma una muestra de 40 pacientes que se saben tienen deficiencia de vitamina B. Los resultados fueron: $\overline{x} = 18.5$ mg/100ml y $s = 4$ mg/100ml. ¿Es la muestra comparable con la población, con un nivel de significación del 0.05?

10. En una granja avícola se usan dos tipos de piensos, A y B. Se quiere saber si la media de engorde de éstos es idéntica, para un nivel de significación del 5 %. Para ello se alimentan a 20 aves con A obteniendo una incrememto de peso medio de 0.4 kgs, con una desviación típica $s = 0.2$ kgs. Se hace lo mismo con otras 20 aves y con el pienso B, obteniéndose un engorde medio de 0.5 kgs, con una desviación típica de $s = 0.3$ kgs. Se supone que las variables engorde, con cada uno de los piensos, son normales y con la misma varianza σ^2.

11. Un análisis de determinación de plomo en comestibles fue aplicado a un mismo producto con dos orígenes diferentes. Los datos observados en mg/Kgr fueron:

A	12.5	14	13	12.5	15	14.5	13		13.5	16
B	9.5	13	13	10.5	11	14	13.5	15		14.5

   a) ¿Es igualmente variable el contenido del plomo?
   b) ¿Hay diferencia significativa en el contenido medio de plomo?

12. Se quiere comprobar la efectividad de una vacuna contra una enfermedad. Para esto se suministró la vacuna a 100 plantas de una determinada especie y se las comparó con otro grupo testigo de otras 100. A las 200 plantas se les contagió la enfermedad. De las vacunada murieron 8 y de las del grupo testigo 20. ¿Se puede conluir que la vacuna es eficaz a un nivel de significación del 0.05?

13. Ante una epidemia en una especie, un botánico sostiene que el número de individuos jóvenes muertos es menor que el de los individuos adultos. Para confirmarlo, aisla a 200 individuos jóvenes de los que al cabo de un determinado tiempo mueren 58. Hace lo mismo con 150 adultos y mueren 36. ¿Qué conlcusión ha de sacar el botánico a un nivel de significación del 5 %?
14. Los valores observados para la tensión de ruptura de una muestra de tamaño 14 de una clase de fibra sintética son

    9.9 5 5.2 7. 11.8 10.3 8.2 7.5 6.6 12.6 16.8 12.3 9.8 10.3

    a) Determinar los intervalos de confianza para la tensión media de ruptura, así como para la varianza poblacional de la variable tensión.
    b) De acuerdo con los datos, ¿es aceptable la afirmación según la cual dicha fibra soporta por término medio una tensión de ruptura al menos igual a 12?

Parte 2

Indicaciones y soluciones

CAPíTULO 7

Soluciones Tema 1

1. Resolución en prácticas
2. Resolución en prácticas
3. Para demostrar estas relaciones ntre los momentos basta con hacer el desarrollo del Binomio de Newton: en particular, para el primer apartado tenemos

$$(x-\mu)^3 = x^3 - 3x^2\mu + 3x\mu^2 - \mu^3,$$

luego

$$\begin{aligned}
\frac{\sum_{i=1}^{N}(x_i-\alpha_1)^3}{N} &= \frac{\sum_{i=1}^{N}\left(x_i^3 - 3x_i^2\alpha_1 + 3x_1\alpha_1^2 - \alpha_1^3\right)}{N} \\
&= \frac{\sum_{i=1}^{N} x_i^3}{N} \\
&\quad -3\alpha_1\frac{\sum_{i=1}^{N} x_i^2}{N} \\
&\quad +3\alpha_1^2\frac{\sum_{i=1}^{N} x_1}{N} \\
&\quad -\alpha_1^3\frac{\sum_{i=1}^{N} 1}{N} \\
&= \alpha_3 - 3\alpha_1\alpha_2 + 3\alpha_1^3 - \alpha_1^3
\end{aligned}$$

Para el apartado segundo hay que hacer el desarrollo de la potencia cuarta y sumar por partes como antes.

4.

   a) Si hacemos $D = B \cup C$ y usamos dos veces la fórmula

$$P(a\cup b) = P(a) + P(b) - P(a\cap b),$$

entonces

$$
\begin{aligned}
P(A\cup B\cup C) &= P(A)+P(B)+P(C)-P(B\cap C)-P(A\cap(B\cup C)) \\
&= P(A)+P(B)+P(C)-P(B\cap C)-P((A\cap B)\cup(A\cap C)) \\
&= P(A)+P(B)+P(C)-P(B\cap C) \\
&\quad -P((A\cap B))-P((A\cap C))+P((A\cap B)\cap(A\cap C)) \\
&= P(A)+P(B)+P(C)-P(B\cap C) \\
&\quad -P((A\cap B))-P((A\cap C))+P(A\cap B\cap C)
\end{aligned}
$$

b) Si usamos la fórmula

$$P((A\cup B)\cap C)=P(A\cap C)+P(B\cap C)-P(A\cap B\cap C)$$

y luego dividimos por $P(C)$ obtenemos la igualdad pedida.

c) $P(A\cap B)\leq P(A)\leq P(A\cup B)\leq P(A)+P(B)$Es obvio que

$$P(A\cup B)=P(A)+P(B)-P(A\cap B)\geq P(A)+P(B)$$

y que como $A\subset A\cup B$ entonces $P(A)\leq P(A\cup B)$. Por otro lado es trivial verificar $P(A\cap B)\leq P(A)$ ya que $A\cap B\subset A$.

5. Que si A y B sean sucesos independientes significa que $P(A\cap B)=P(A)P(B)$. Que lo san A y $\overline{B}$ lo sean equivale a verificar

$$P\left(A\cap\overline{B}\right)=P(A)P\left(\overline{B}\right)$$

pero $P\left(\overline{B}\right)=1-P(B)$, luego la identidad de arriba, la que queremos verificar, es

$$P\left(A\cap\overline{B}\right)=P(A)(1-P(B))=P(A)-P(A)P(B)$$

y como A y B son independientes entonces esto equivale a

$$P\left(A\cap\overline{B}\right)=P(A)(1-P(B))=P(A)-P(A\cap B)$$

pero esto es lo mismo que escribir

$$P(A)=P\left(A\cap\overline{B}\right)+P(A\cap B)$$

lo cual es cierto pues $A=\left(A\cap\overline{B}\right)\cup(A\cap B)$ (unión disjunta).
Como A y $\overline{B}$ lo son entonces también$\overline{A}$ y $\overline{B}$. Para ver $P(A\cup B)=1-P(\overline{A})P(\overline{B})$ sustituimos que $P\left(\overline{B}\right)=1-P(B)$ y $P\left(\overline{A}\right)=1-P(A)$ y usamos que $P(A\cap B)=P(A)P(B)$.

6. Escribimos la probabilidad de que pertenezca al grupo de riesgo así:

$$P(R)=0.05$$

Ahora, el suceso S descrito como que de tres individuos elegidos al azar, uno esté en el grupo de riesgo se escribe así

$$\left(R1\cap\overline{R2}\cap\overline{R3}\right)\cup\left(\overline{R1}\cap R2\cap\overline{R3}\right)\cup\left(\overline{R1}\cap\overline{R2}\cap R3\right)$$

como la unión de tres sucesos disjuntos, donde, por ejemplo, $R1 \cap \overline{R2} \cap \overline{R3}$ significa que el primero está en grupo de riesgo y ni el segundo ni el tercero están. Así

$$P(S) = P\left(R1 \cap \overline{R2} \cap \overline{R3}\right) + P\left(\overline{R1} \cap R2 \cap \overline{R3}\right) + P\left(\overline{R1} \cap \overline{R2} \cap R3\right)$$

Cada uno de los tres sucesos en los que se descompone S es la interesección de otros tres sucesos, y éstos son independientes. Por ello tenemos, por ejemplo, que

$$\begin{aligned} P\left(\overline{R1} \cap R2 \cap \overline{R3}\right) &= P\left(\overline{R1}\right) P(R2)\, P\left(\overline{R3}\right) \\ &= (0.95)(0.05)(0.95) \end{aligned}$$

Por tanto $P(S) = 3(0.95)(0.05)(0.95)$

7. Las distintas composiciones para la urna son:

$$U_1,\ U_2,\ U_3,\ U_4$$

donde U_i es la urna que tiene i bolas blancas (y $5-i$ negras). Por tanto, extraer blanca, B, se puede dar en cualquiera de estos cuatro escenarios con igual probabilidad. Así

$$\begin{aligned} P(B) &= P(B \cap U_1) + P(B \cap U_2) + P(B \cap U_3) + +P(B \cap U_4) \\ &= P(B|\, U_1)\, P(U) + P(B|\, U_2)\, P(U_2) + P(B|\, U_3)\, P(U_3) + P(B|\, U_4)\, P(U_4) \\ &= \frac{1}{4}\left(\frac{1}{5} + \frac{2}{5} + \frac{3}{5} + \frac{4}{5}\right) = 0.5 \end{aligned}$$

8. $P(R_i)$ denota la probabilidad de que funcion el $i-$ésimo.
   a) $P(R_1 \cap R_2) = P(R_1)\, P(R_2) = \alpha_1\alpha_2$
   b)

$$\text{máx}\{\alpha_1, \alpha_2\} \leq P(R_1 \cap R_2) \leq P(R_1) + P(R_2) = \alpha_1 + \alpha_2$$

9. Como el 40 % son chicos y el 60 % chicas

$$P(V) = 0.4 \text{ y } P(F) = 0.6$$

y como entre los chicos un 4 % tienen una altura superior a 180 cm

$$P(H|\, V) = 0.04$$

donde H es el suceso que consiste en medir más de 180 cms. Para la chicas, según los datos del problema

$$P(H|\, F) = 0.01$$

Así

$$\begin{aligned} P(H) &= P(H \cap V) + P(H \cap F) \\ &= P(H|\, V)\, P(V) + P(H|\, F)\, p(F) \\ &= (0.04)(0.4) + (0.01)(0.6) \end{aligned}$$

10. Como el volumen de ventas total es 3500 las probabilidades que podemos otorgar para los distintos tipos son:

$$\begin{aligned} P(D) &= \frac{500}{3500} = 0.142\,86 \\ P(F) &= \frac{1000}{3500} = 0.285\,71 \\ P(U) &= \frac{2000}{3500} = 0.571\,43 \end{aligned}$$

Que el porcentaje de coches defectuosos es de un 2 % para los deportivos se traduce en

$$P(De|\,D) = 0.02$$

que los sea un 1 % para el segundo escomo escribir

$$P(De|\,F) = 0.01$$

y de un 3 % para el tercero es

$$P(De|\,U) = 0.03$$

a) La probabilidad de que no sea defectuoso, probabilidad de NDe ó de $\overline{De}$, es

$$\begin{aligned} P(NDe) &= P(NDe \cap D) + P(NDe \cap F) + P(NDe \cap U) \\ &= P(NDe|\,D)\,P(D) + P(NDe|\,F)\,P(F) + P(NDe|\,U)\,P(U) \\ &= (1 - P(De|\,D))\,\frac{500}{3500} + (1 - 0{,}01)\,\frac{1000}{3500} + (1 - 0{,}03)\,\frac{2000}{3500} \\ &= 0.977\,14 \end{aligned}$$

b) Si hemos elegido un automóvil defectuoso, que dicho automóvil pertnezca al grupo de los familiares tiene probabilidad $P(F|\,De)$, se calcula con ayuda del Teorema de Bayes y con el resultado del apartado anterior:

$$\begin{aligned} P(F|\,De) &= \frac{P(De|\,F)\,P(F)}{P(De)} = \frac{P(De|\,F)\,P(F)}{1 - P(NDe)} \\ &= \frac{(0.01)\left(\frac{1000}{3500}\right)}{1 - 0.977\,14} = 0.124\,98 \end{aligned}$$

11. Nos están pidiendo $P(U_1|\,S)$ y para ello usamos que

$$\begin{aligned} P(U_1|\,S) &= \frac{P(S|\,U_1)\,P(U_1)}{P(S)} \\ &= \frac{P(S|\,U_1)\,P(U_1)}{P(S \cap U_1) + P(S \cap U_2)} \\ & \quad \frac{P(S|\,U_1)\,P(U_1)}{P(S|\,U_1)\,P(U_1) + P(S|\,U_2)\,P(U_2)} \end{aligned}$$

Los cálculos son sencillos:

$$\begin{aligned} P\left(S\middle|\, U_1\right) &= (0.7)^8\,(0.3)^2 \\ P\left(S\middle|\, U_2\right) &= (0.3)^8\,(0.7)^2 \end{aligned}$$

12. De manera esquemática:

$$\begin{array}{l} P(p=0{,}05)=1/2 \\ P(p=0{,}1)=1/4, \\ P(p=0{,}15)=1/5,\text{ y} \\ P(p=0{,}2)=1/20. \end{array}$$

Entonces

$$\begin{gathered} P(p=15/100|_{S=3}) = \\ = \frac{P(S=3|\,p=0{,}15)P(p=0{,}15)}{P(S=3)} = \\ \frac{P(S=3|\,p=0{,}15)P(p=0{,}15)}{P(S=3|\,p=0{,}05)P(0{,}05)+...+P(S=3|\,p=0{,}2)P(0{,}2)} \\ = \frac{\binom{20}{3}(0{,}15)^3(0{,}85)^{17}\frac{20}{100}}{\binom{20}{3}(0{,}05)^3(0{,}95)^{17}\frac{1}{2}+...+\binom{20}{3}(0{,}2)^3(0{,}8)^{17}\frac{1}{20}} = 0.35. \end{gathered}$$

13. Hay dos tipos de urna: de tipo 1 ($T1$) con 2 B y 2 N, y de tipo 2 (T2) con 4 B y 1 N. Hemos de calcular

$$\begin{aligned} P\left(T2\middle|\, B\right) &= \frac{P\left(B\middle|\, T2\right)P\left(T2\right)}{P\left(B\right)} \\ &= \frac{P\left(B\middle|\, T2\right)P\left(T2\right)}{P\left(B\middle|\, T1\right)P\left(T1\right)+P\left(B\middle|\, T2\right)P\left(T2\right)} \end{aligned}$$

14. Las probabilidades de acierto para cada jugador son$P(A_1) = 4/5$, $P(A_2) = 3/4$ y $P\left(A_3\right) = 2/3$. Ha ocurrido un fallo, esto esto es

$$S = 1 \text{ fallo}$$

y nos preguntan por $P\left(F_3|\,S\right)$. Empleamos la fórmula de la probabilidad condicionada y dividimos el suceso S según sea el jugador que ha fallado:

$$\begin{aligned}
P\left(F_3|\,S\right) &= \frac{P\left(F_3\cap S\right)}{P\left(S\right)}\\
&= \frac{P\left(F_3\cap S\right)}{P\left(F_1\cap S\right)+P\left(F_2\cap S\right)+P\left(F_3\cap S\right)}\\
&= \frac{P\left(A_1\cap A_2\cap F_3\right)}{P\left(A_2\cap A_3\cap F_1\right)+P\left(A_1\cap A_3\cap F_2\right)+P\left(A_1\cap A_2\cap F_3\right)}\\
&= \frac{P\left(A_1\right)P\left(A_2\right)\left(1-P\left(A_3\right)\right)}{P\left(A_2\right)P\left(A_3\right)\left(1-P\left(A_1\right)\right)+P\left(A_1\right)P\left(A_3\right)\left(1-P\left(A_2\right)\right)+P\left(A_1\right)P\left(A_2\right)\left(1-P\left(A_3\right)\right)}\\
&= \frac{(4/5)\,(3/4)\,(1-(2/3))}{(3/4)\,(2/3)\,(1-(4/5))+(4/5)\,(2/3)\,(1-(3/4))+(4/5)\,(3/4)\,(1-(2/3))}\\
&= 0{,}461\,54
\end{aligned}$$

15. Sea $\Omega=(0,+\infty)$. Demostar que si definimos P sobre todos los intervalos $(a,b)\subset\Omega$ con ayuda de la fórmula

$$P\left((a,b)\right)=\int_a^b e^{-x}dx,$$

entonces P cumple todos los axiomas de la probabilidad.

(i) $P\left(\Omega\right)=1$, se traduce en que $P\left((0,+\infty)\right)=1$, y así es

$$P\left((0,+\infty)\right)=\int_0^{+\infty}e^{-x}dx=-e^{-x}\Big|_0^{+\infty}=1$$

(ii) $P(S)\geq 0\ \forall S\in\mathcal{A}$, ahora se convierte en $P\left((a,b)\right)=\int_a^b e^{-x}dx\geq 0$ para todo $(a,b)\subset\Omega$. Esto es verdad pues el integrando, e^{-x} es siempre positivo y la integral de una función positiva siempre es no negativa.

(iii)Verificar $P\left(\cup_{j=1}^{\infty}S_j\right)=\sum_{j=1}^{\infty}P\left(S_j\right)$, con S_j disjuntos es directo pues por las propiedades de las integrales

$$P\left((a,b)\cup(c,d)\right)=\int_a^b e^{-x}dx+\int_c^d e^{-x}dx$$

CAPíTULO 8

Soluciones Tema 2

1. Debemos tener en cuenta que
$$F_\xi(x) = P^*_\xi((-\infty, x]) = P(\omega \in \Omega : \xi(\omega) \in (-\infty, x])$$
y que por ello
$$F_\xi(b) - F_\xi(a) = P^*_\xi((a, b]) = P(\omega \in \Omega : \xi(\omega) \in (a, b])$$
   a) La notación $P(a < \xi < b)$ es
$$P(a < \xi < b) = P^*_\xi((a, b)) = P(\omega \in \Omega : \xi(\omega) \in (a, b))$$
   Bien, como $(a, b) \cup \{b\} = (a, b]$ entonces
$$P^*_\xi((a, b)) + P^*_\xi(\{b\}) = P^*_\xi((a, b]) = F_\xi(b) - F_\xi(a)$$
   y de esto se llega a que $P(a < \xi < b) = F_\xi(b) - F_\xi(a) - P^*_\xi(b)$.

   b) Lo mismo ocurre con $P(a \leq \xi < b)$
$$[a, b) = (a, b) \cup \{a\}$$
   por lo que
$$P^*_\xi([a, b)) = P^*_\xi((a, b)) + P^*_\xi(\{a\})$$
   Basta emplear la fórmula del apartado anterior para concluir que
$$P^*_\xi([a, b)) = F_\xi(b) - F_\xi(a) - P^*_\xi(b) + P^*_\xi(a)$$

2. Hagamos un esbozo de la prueba de lo que nos están pididiendo: puesto que ξ es una v.a. de tipo continuo
$$\lim_{h \to 0} F_\xi(x - h) = F_\xi(x)$$
Y como
$$\{x\} \approx (x - h, x]$$
con h pequeño, $h \to 0$, entonces
$$P^*_\xi(x) = \lim_{h \to 0} P^*_\xi((x - h, x]) = \lim_{h \to 0} (F_\xi(x) - F_\xi(x - h)) = 0.$$

3. Sea ξ una v.a. tal que se cumple la siguiente tabla de probabilidades:

x_i	1	2	3	4	5
$P^*_\xi(x_i)$	2/8	1/8	2/8	2/8	1/8

Se pide:

a) Basta ver que todos los $P_\xi^*(x_i)$ son positivos y que sus suma es uno.es una función de densidad de masa.

b)

$$\begin{aligned} F_\xi(x) &= P_\xi^*((-\infty, x]) = \\ &= P(\omega \in \Omega : \xi(\omega) \in (-\infty, x]) \\ &= \begin{cases} 0 & \text{si } x < 1 \\ 2/8 & \text{si } x \in [1,2) \\ 3/8 & \text{si } x \in [2,3) \\ 5/8 & \text{si } x \in [3,4) \\ 7/8 & \text{si } x \in [4,5) \\ 1 & \text{si } x \geq 5 \end{cases} \end{aligned}$$

c)

$$\begin{aligned} P(1{,}1 &< \xi < 3{,}3) = F_\xi(3{,}3) - F_\xi(1{,}1) - P_\xi^*(\{3{,}3\}) \\ &= \frac{5}{8} - \frac{2}{8} = \frac{3}{8} \end{aligned}$$

4. Dada la v.a. ξ cuya densidad de probabilidad es

$$f_\xi(x) = \begin{cases} k\sin(x) & \text{si } x \in \left[0, \frac{\pi}{2}\right] \\ 0 & \text{en el resto,} \end{cases}$$

donde k es una constante, se pide:

a) Que $f_\xi(x)$ sea efectivamente una densidad de probabilidad equivale a pedir que

$$\int_{\mathbb{R}} f_\xi(x)dx = 1$$

y que $f_\xi(x) \geq 0$. La segunda imposición se cumpla siempre que $k \geq 0$ ya que $\sin(x) \geq 0$ en $\left[0, \frac{\pi}{2}\right]$. Imponer la primera es

$$1 = \int_{-\infty}^{\infty} f_\xi(x)dx = \int_0^{\frac{\pi}{2}} k\sin(x)dx = -k\cos(x)|_0^{\pi/2} = k$$

Por ello basta con poner $k = 1$ para que f_ξ sea densidad de probabilidad.

b)

$$F_\xi(x) = P_\xi^*((-\infty, x]) = \begin{cases} 0 & \text{si } x < 0 \\ \int_0^x \sin(s)ds = -\cos x + 1 & \text{si } x \in [0, \pi/2] \\ 1 & \text{si } x \geq \pi/2 \end{cases}$$

c)

$$\begin{aligned} P_\xi^*((\pi/4, \pi/2)) &= F_\xi(\pi/2) - F_\xi(\pi/4) \\ &= (-\cos\pi/2 + 1) - (-\cos\pi/4 + 1) \\ &= \cos\pi/4 \end{aligned}$$

5. Sea la función $f_\xi(x) = \frac{C}{3}$ si $x \in [0,3]$, e $= 0$ si $x \notin [0,3]$. Se pide

a) Para que f_ξ sea una función de densidad de probabilidad se tiene que cumplir que $C \geq 0$ y que $\int_{\mathbb{R}} f_\xi(x)dx = 1$. Pedir esto último es

$$1 = \int_{-\infty}^{\infty} f_\xi(x)dx = \int_0^3 \frac{C}{3}dx = \frac{C}{3}x\bigg|_0^3 = C$$

Tomando $C = 1$ se cumple por tanto que f_ξ es densidad de probabilidad

b)

$$F_\xi\left(x\right) = P_\xi^*\left(\left(-\infty, x\right]\right) = \begin{cases} 0 & \text{si } x < 0 \\ \int_0^x \frac{1}{3}ds = \frac{x}{3} & \text{si } x \in [0,3] \\ 1 & \text{si } x \geq 3 \end{cases}$$

c)

$$P(1{,}5 < \xi < 3{,}5) = F_\xi\left(3.5\right) - F_\xi\left(1.5\right) = 1 - \frac{1.5}{3} = \frac{1}{2}$$

6. Determinar el parámetro C:
   a) Es inmediato que lo que hay que exigir es que $C \geq 0$ y que

   $$1 = \int_0^1 Cx^2dx = \frac{1}{3}C$$

   Bsta tomar $C = 3$.
   b) Igualmente C ha de ser mayor o igual que cero y

   $$1 = \int_0^\infty C\exp(-Cx)dx = -\exp(-Cx)|_0^\infty = 1$$

   Se puede tomar cualquier $C > 0$ (no se toma $C = 0$ porque en tal caso la integral no da 1)
   c) Para

   $$f_\xi\left(x\right) = \frac{C}{1+x^2}, \quad x \in \mathbf{R}$$

   se procede igual: $C \geq 0$ y $1 = \int_{-\infty}^{+\infty} \frac{C}{1+x^2}dx = C\arctan(x)|_{-\infty}^{\infty} = C(\pi-(-\pi)))2\pi C$. Tómese $C = \frac{1}{2\pi}$.
7. Determinar el parámetro C
   a) $P_\xi^*(x) = C\frac{x-1}{n}, \quad x = 2, ..., n$ y $P_\xi^*(x) = 0$ para cualquier otro valor de x. Para que todas las masas sean positivas tomaremos $C > 0$. Como queremos que

   $$1 = \sum_{x=2}^{n} P_\xi^*(x)$$

entonces exigimos

$$\begin{aligned} 1 &= C\frac{1}{n} + C\frac{3-1}{n} + ... + C\frac{n-1}{n} \\ &= C\left(\frac{1}{n} + \frac{2}{n} + ... + \frac{n-1}{n}\right)^{1} \\ &= C\left(\frac{n(n-1)}{2}\right) \end{aligned}$$

Por ende, la elección de C es

$$C = \frac{2}{n(n-1)}.$$

b) $P_{\xi}^{*}(x) = C\frac{m^x}{x!}$, $\quad x = 0, 1, 2, ..., n, ...$ y $P_{\xi}^{*}(x) = 0$ para cualquier otro valor de x.[2] El valor de C ha de ser forzosamente positivo. Puesto que

$$1 = \sum_{x=0}^{\infty} P_{\xi}^{*}(x)$$

entonces

$$1 = \sum_{x=0}^{\infty} C\frac{m^x}{x!} = Ce^m$$

Por tanto $C = e^{-m}$

c) $P_{\xi}^{*}(x) = C(\frac{1}{3})^x$, $\quad x = 0, 1, 2, ..., n, ...$ y $P_{\xi}^{*}(x) = 0$ para cualquier otro valor de x. $C > 0$ y

$$\begin{aligned} 1 &= \sum_{x=0}^{\infty} P_{\xi}^{*}(x) = C\sum_{x=0}^{\infty}(\frac{1}{3})^x \\ &= C\frac{1}{1-\frac{1}{3}} = C\frac{3}{2} \end{aligned}$$

Por consiguiente $C = \frac{2}{3}$.

8. Calcular la esperanza y desviación típica de cada uno de los apartados de los ejercicios 6 y 7.

a)

$$\begin{aligned} \mu &= \int_{-\infty}^{\infty} x f_{\xi}(x)\, dx = \int_{0}^{1} 3x^3 dx = \frac{3}{4}, \\ \alpha_2 &= \int_{-\infty}^{\infty} x^2 f_{\xi}(x)\, dx = \int_{0}^{1} 3x^4 dx = \frac{3}{5}, \\ \sigma^2 &= \frac{3}{5} - \left(\frac{3}{4}\right)^2 = \frac{3}{80}. \end{aligned}$$

[2]Téngase en cuenta que $\sum_{r=0}^{\infty} \frac{m^r}{r!} = e^m$.

b)

$$
\begin{aligned}
\mu &= \int_{-\infty}^{\infty} x f_{\xi}(x)\,dx = \int_{0}^{\infty} Cx\exp(-Cx)dx^{3} \\
&= \left(-x\exp(-Cx) + \int_{0}^{\infty}\exp(-Cx)dx\right)\Bigg|_{0}^{\infty} = \\
&= -x\exp(-Cx)|_{0}^{\infty} - \frac{1}{C}\exp(-Cx)\Bigg|_{0}^{\infty} \\
&= \frac{1}{C}, \\
\alpha_2 &= \int_{0}^{\infty} Cx^2\exp(-Cx)dx = \frac{2}{C^2}, \\
\sigma^2 &= \frac{2}{C^2} - \left(\frac{1}{C}\right)^2 = \frac{1}{C^2}.
\end{aligned}
$$

Para calcular α_2 hemos de integra dos veces por partes.

c) Se observa que el momento de orden 1 no existe ya que la expresión

$$\int_{-\infty}^{\infty} x f_{\xi}(x) = \int_{-\infty}^{\infty} x\frac{\frac{1}{2\pi}}{1+x^2}dx = \left|\frac{1}{4\pi}\log(1+x^2)\right|_{-\infty}^{\infty}$$

no está definida.

d) La esperanza para una variable aleatoria discreta es

$$\mu = \sum_{x=2}^{n} x P_{\xi}^{*}(x)$$

Para nuestro caso es

$$
\begin{aligned}
\mu &= C\frac{1}{n}2 + C\frac{3-1}{n}3 + ... + C\frac{n-1}{n}n \\
&= C\left(\frac{2}{n} + \frac{2\times 3}{n} + ... + \frac{(n-1)\,n}{n}\right)^{4} \\
&= C\sum_{k=2}^{n}\left(\frac{k(k-1)}{2}\right) \\
&= \frac{C}{2}\left(\sum_{k=2}^{n}k^2 - \sum_{k=2}^{n}k\right).
\end{aligned}
$$

Está claro que $\sum_{k=2}^{n} k = \left(\sum_{k=1}^{n} k\right) - 1 = \frac{n(n+1)}{2} - 1$. Falta calcular $\sum_{k=2}^{n} k^2$. Sabemos

$$L = \sum_{k=1}^{n} k^2 = \frac{n\,(n+1)\,(2n+1)}{6}$$

por tanto

$$\mu = \frac{C}{2}\left(\frac{n\,(n+1)\,(2n+1)}{6} - 1 - \frac{n\,(n+1)}{2} - 1\right)$$

e) Sabemos que $\sum_{x=0}^{\infty} \frac{m^x}{x!} = e^m$ y por ello

$$\begin{aligned} \mu &= e^{-m}\sum_{x=0}^{\infty} x\frac{m^x}{x!} \\ &= e^{-m}\sum_{x=0}^{\infty} m\frac{m^{x-1}}{(x-1)!} \\ &= me^{-m}\sum_{x=1}^{\infty} \frac{m^{x-1}}{(x-1)!} \\ &= me^{-m}e^m = m \end{aligned}$$

Del mismo modo

$$\begin{aligned} \alpha_2 &= e^{-m}\sum_{x=0}^{\infty} x^2\frac{m^x}{x!} = me^{-m}\sum_{x=1}^{\infty}(x-1+1)\,\frac{m^{x-1}}{(x-1)!} \\ &= me^{-m}\sum_{x=1}^{\infty}(x-1)\,\frac{m^{x-1}}{(x-1)!} + me^{-m}\sum_{x=1}^{\infty}\frac{m^{x-1}}{(x-1)!} \\ &= me^{-m}\sum_{x=0}^{\infty} x\frac{m^x}{x!} + me^{-m}\sum_{x=1}^{\infty}\frac{m^{x-1}}{(x-1)!} \\ &= m^2 + m \end{aligned}$$

Así $V[\xi] = m^2 + m - m^2 = m$

f) Con $P_{\xi}^{*}(x) = C(\frac{1}{3})^x, \quad x = 0, 1, 2, ...$tenemos

$$\begin{aligned} \mu &= \frac{2}{3}\sum_{x=0}^{\infty} x(\frac{1}{3})^x \\ &= \frac{2}{3}\sum_{x=1}^{\infty}(x-1)\,(\frac{1}{3})^{x-1} \\ &= \frac{2}{3}3\sum_{x=1}^{\infty} x(\frac{1}{3})^x - \frac{2}{3}3\sum_{x=1}^{\infty}(\frac{1}{3})^x \\ &= 2\mu - 2\frac{1/3}{1-1/3} = \mu \end{aligned}$$

De esto se sigue que necesariamente ha de cumplirse

$$2\mu - 2\frac{1/3}{1-1/3} = \mu$$

y por tnato que $\mu = 1$.
Para calcular

$$\alpha_2 = \frac{2}{3}\sum_{x=0}^{\infty} x^2 (\frac{1}{3})^x$$

se procede haciendo algo parecido al caso del cálculo de μ :

$$\begin{aligned}
\alpha_2 &= \frac{2}{3}\sum_{x=1}^{\infty} (x-1)^2 (\frac{1}{3})^{x-1} \\
&= \frac{2}{3}\sum_{x=1}^{\infty} x^2 (\frac{1}{3})^{x-1} + \frac{2}{3}\sum_{x=1}^{\infty} 2x(\frac{1}{3})^{x-1} + \frac{2}{3}\sum_{x=1}^{\infty} (\frac{1}{3})^{x-1}
\end{aligned}$$

Se omiten los detalles.

9. Sea ξ v.a. de tipo continuo cuya función de densidad de probabilidad es

$$f_\xi(x) = \begin{cases} 2e^{-2x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases} .$$

a)

$$F_\xi(x) = P_\xi^*((-\infty, x]) = \begin{cases} 0 & \text{si } x \leq 0 \\ \int_0^x 2e^{-2t}dt = -e^{-2x} + 1 & \text{si } x \geq 0 \end{cases}$$

b) Usamos la fórmula de la probabilidad condicionada:

$$\begin{aligned}
P\left\{\xi > 2|_{\xi<4}\right\} &= \frac{P\{(\xi < 4) \cap (\xi > 2)\}}{P(\xi < 4)} \\
&= \frac{P(2 < \xi < 4)}{P(\xi < 4)} = \frac{P_\xi^*((2,4))}{P_\xi^*((-\infty, 4])} \\
&= \frac{F_\xi(4) - F_\xi(2)}{F_\xi(4)}
\end{aligned}$$

10. La variable aleatoria ξ tiene función de densidad de probabilidad

$$f_\xi(x) = \begin{cases} x & \text{si } x \in [a,b] \\ 0 & \text{en el resto.} \end{cases}$$

Se supone que a y b son números tales que $0 < a < b$.

a) Usamos $P\{a \leq \xi \leq 2a\} = \frac{3}{8}$ y también que $P\{-\infty \leq \xi \leq \infty\} = 1$ para ajustar los parámetros a y b :
la primera dice

$$\frac{3}{8} = \int_a^{2a} x dx = \frac{3}{2}a^2$$

(se sabe que $2a \leq b$) ; y la segunda da

$$1 = \int_a^b x dx = \frac{1}{2}b^2 - \frac{1}{2}a^2$$

Resolvemos el sistema

$$\begin{cases} \frac{3}{2}a^2 = \frac{3}{8} \\ \frac{1}{2}b^2 - \frac{1}{2}a^2 = 1 \end{cases}$$

Las soluciones posibles son: $\left(a = \frac{1}{2}, b = \frac{3}{2}\right)$, $\left(a = \frac{1}{2}, b = -\frac{3}{2}\right)$, $\left(a = -\frac{1}{2}, b = \frac{3}{2}\right)$ y $\left(a = -\frac{1}{2}, b = -\frac{3}{2}\right)$. La única válida es $a = \frac{1}{2}$, $b = \frac{3}{2}$.

11. Si m es el número de camas y ξ es el número de enfermos, entonces para poder hospitalizarlos es preciso que $\xi \leq m$. Por ello piden m para que

$$P(\xi \leq m) = 0{,}5$$

Ahora bien

$$P(\xi \leq m) = \int_{-\infty}^{m} f_\xi(x)\,dx = \int_0^m \frac{2}{9}\left(3x - x^2\right)dx = 0.5$$

supuesto que $m \leq 3$. La identidad de arriba se escribe

$$-4m^3 + 18m^2 - 27 = 0$$

La única raíz admisible es $m = 1.5$.
Nótese que si dibujamos la gráfica dela función $y = \frac{2}{9}\left(3x - x^2\right)$ no es difícil verificar la recta $x = 1{,}5$ es una recta de simetría para función y que por tanto hay el mismo área entre 0 y 1.5 que entre 1.5 y 3.

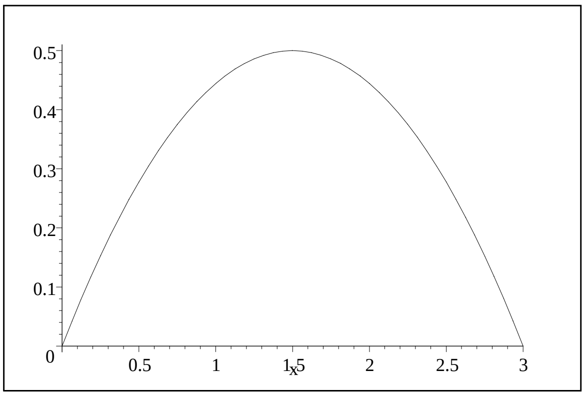

Gráfico de la curva $y = \frac{2}{9}\left(3x - x^2\right)$

12. Nota: el tiempo esperado es el valor esperado y espereanza de la v.a.
13. Sean

$$\xi_i = \text{número de puntos que salen el la tirada } i.$$

Es obvio que

$$\eta_n = \sum_{i=1}^{n} \xi_i$$

Tampoco es difícil ver que la esperanza y la varianza de las ξ_i es 3.5 y $n\,()$:

$$\begin{aligned} E[\xi_i] &= 1\frac{1}{6}+2\frac{1}{6}+...+6\frac{1}{6}=\frac{1}{6}(1+2+...+6)=\frac{7}{2} \\ E[\xi_i^2] &= 1^2\frac{1}{6}+2^2\frac{1}{6}+...+6^2\frac{1}{6}=\frac{1}{6}\left(1+2^2+...+6^2\right)=\frac{1}{6}\left(\frac{42\,(13)}{6}\right)=\frac{45}{3} \\ V[\xi_i] &= \frac{45}{3}-\frac{49}{4}=\frac{11}{4} \end{aligned}$$

a) Por tanto

$$\begin{aligned} E[\eta_n] &= E\left[\sum_{i=1}^{n}\xi_i\right]=\sum_{i=1}^{n}E[\xi_i] \\ &= n\frac{7}{2} \\ V[\eta_n] &= V\left[\sum_{i=1}^{n}\xi_i\right]=\sum_{i=1}^{n}V[\xi_i] \\ &= n\frac{11}{4} \end{aligned}$$

Por consiguiente

$$\begin{aligned} E[\frac{\eta_n}{n}] &= \frac{7}{2}=3.5 \\ V[\frac{\eta_n}{n}] &= \frac{1}{n^2}n\frac{11}{4}=\frac{11}{4n} \end{aligned}$$

b)

$$\begin{aligned} & P\left\{\left|\frac{\eta_n}{n}-3.5\right|\geq 0.1\right\} \\ =\ & P\left\{\left(\frac{\eta_n}{n}-3.5\right)^2\geq (0.1)^2\right\}\leq\frac{V[\eta_n]}{(0.1)^2} \\ =\ & 100\frac{11}{4n}=\frac{275}{n} \end{aligned}$$

Imponemos

$$\frac{275}{n}\leq 0{,}1$$

y obtenemos así que el n a elegir ha de ser mayor o igual que 2750 : $n\geq 2750$.

14. Problema estándar.
15. Idem
16. Sea la v.a. de tipo discreto ξ cuya función de densidad de masa es

$$P_\xi^*(x)=p(1-p)^{x-1},\quad x=1,2,3,...$$

donde $p \in (0,1)$.

$$\begin{aligned}
\mu &= E[\xi] = \sum_{x=1}^{\infty} xp(1-p)^{x-1} \\
&= \sum_{x=0}^{\infty} (x+1)\, p(1-p)^{x} \\
&= (1-p) \sum_{x=0}^{\infty} (x)\, p(1-p)^{x-1} + \sum_{x=0}^{\infty} p(1-p)^{x} \\
&= (1-p)\mu + 1 = \mu
\end{aligned}$$

Así pues, $(1-p)\mu + 1 = \mu$ y por tnato $\mu = \frac{1}{p}$.
Para el momento de orden 2 se hace algo similar:

$$\begin{aligned}
\alpha_2 &= \sum_{x=1}^{\infty} x^2 p(1-p)^{x-1} \\
&= \sum_{x=0}^{\infty} (x+1)^2\, p(1-p)^{x}
\end{aligned}$$

17.

a) Si $P_\xi^*(0) = 1/3$ y $P_\xi^*(1) = 2/3$ entonces

$$\begin{aligned}
\mu &= 0\frac{1}{3} + 1\frac{2}{3} = \frac{2}{3} \\
\alpha_2 &= 0^2\frac{1}{3} + 1^2\frac{2}{3} = \frac{2}{3} \\
\sigma^2 &= (2/3) - (2/3)^2
\end{aligned}$$

b) Como $P_\xi^*(x) = \begin{pmatrix} n \\ x \end{pmatrix} p^x(1-p)^{n-x}, \quad x = 0,1,2,3,...n$, entonces, cualquiera que sea $n \geq 1$ se tiene

$$\sum_{x=0}^{n} \begin{pmatrix} n \\ x \end{pmatrix} p^x(1-p)^{n-x} = 1.$$

Empleemos esto:

$$\begin{aligned}
\mu &= \sum_{x=0}^{n} x \frac{n!}{x!\,(n-x)!} p^x (1-p)^{n-x} \\
&= \sum_{x=0}^{n-1} (x+1) \frac{n!}{(x+1)!\,(n-(x+1))!} p^{x+1} (1-p)^{n-(x+1)} \\
&= np \sum_{x=0}^{n-1} \frac{(n-1)!}{(x)!\,(n-1-x)!} p^x (1-p)^{n-1-x} \\
&= np.
\end{aligned}$$

Del mismo modo se puede calcular el momento de orden 2:

$$\begin{aligned}
\alpha_2 &= \sum_{x=0}^{n} x^2 \frac{n!}{x!\,(n-x)!} p^x (1-p)^{n-x} \\
&= \sum_{x=0}^{n-1} (x+1)^2 \frac{n!}{(x+1)!\,(n-(x+1))!} p^{x+1} (1-p)^{n-(x+1)}
\end{aligned}$$

y desarrollando $(x+1)^2$ llegaremos al resultado con cierta facilidad (se omiten los detalles).

CAPíTULO 9

Soluciones Tema 3

1. Identificamos el suceso producir cierta mejoría com M. Por ello $P(M) = 0.7$. Se realizan 10 observaciones y consideramos la variable aleatoria ξ que da el número de plantas en las que ha habido mejoría. Se cumple que $\xi \sim B(10, 0.7)$

   a) La probabilidad de que mejore el agua en 4 de las plantas: es $P(\xi = 4)$, esto es

$$\begin{aligned} P(\xi = 4) &= \binom{10}{4} (0.7)^4 (1 - 0.7)^6 \\ &= \frac{10!}{4!6!} (0.7)^4 (1 - 0.7)^6 = 3.6757 \times 10^{-2} \end{aligned}$$

   b) Que al menos mejore en 3 de las plantas es equivalente a escribir $\xi \geq 3$, por ello lo que nos piden es

$$\begin{aligned} P(\xi \geq 3) &= 1 - P(\xi = 0) - P(\xi = 1) - P(\xi = 2) \\ &= 1 - (0.7)^0 (1 - 0.7)^{10} \\ &\quad - \frac{10!}{1!9!} (0.7)^1 (1 - 0.7)^9 \\ &\quad - \frac{10!}{2!8!} (0.7)^2 (1 - 0.7)^8 \\ &= 0.99841 \end{aligned}$$

2. En cada segundo la partícula se mueve a la derecha 0.001 mm con una probabilidad de $\frac{1}{60}$ y $-0{,}001$ mm a la izquierda con una probabilidad $1 - \frac{1}{60}$. Esto se interpreta diciendo que $P(D) = \frac{1}{60}$. Como se analiza el movimiento de 360 segundos podemos definir la v.a. $\xi =$ númerode movimientos a la derecha. Se tiene $\xi \sim B\left(360, \frac{1}{60}\right)$ y en tal caso $B\left(360, \frac{1}{60}\right) \approx \mathcal{P}(\lambda)$ (una distribución de Poisson de parámetro λ) con $\lambda = 360 \times \frac{1}{60} = 6{,}0$ $(n \geq 30,\ p \leq 0{,}1)$.

   a) La probabilidad de que la partícula realice 15 movimientos a la derecha al cabo de 360 segundos es

$$P(\xi = 15) = e^{-6} \frac{(6)^{15}}{15!} = 8.9126 \times 10^{-4}.$$

   b) Que esté en el punto $(-0.3, 0)$ significa qus si m son los desplazameintos a la derecha y n a la izquierda entonces

$$\begin{aligned} m + n &= 360 \\ m(0.001) + n(-0.001) &= -0.3 \end{aligned}$$

Entonces $n = 330$ y $m = 30$ y lo que nos están pidienod es

$$P\left(\xi = 30\right) = e^{-6}\frac{\left(6\right)^{30}}{30!}.$$

c) La posición esperada de la partícula al cabo de los 360 segundos se determina calculando la esperanza de ξ, que es el número de desplazamientos a la derecha que se esperan de la partícula. Al tratarse de una Poisson de parámetro $\lambda = 6$ la esperanza es

$$E[\xi] = 6$$

y por ende la posición esperada es

$$6\left(0.001\right) + \left(354\right)\left(-0.001\right) = -0.348.$$

3. La probabilidad de choque es 0.002, digamos que $P\left(I\right) = 0.002$. y se llevan a cabo 1200 observaciones. Se trata de una binomial $B(1200, 0.002)$ la cual puede aproximarse por una distribución de Poisson $\mathcal{P}\left(1200 \times 0.002 = 2.4\right)$.

a)

$$\begin{aligned} P(\psi &\leq 5) = P(\psi = 0) + P(\psi = 1) + P(\psi = 2) + P(\psi = 3) + P(\psi = 4) + P(\psi = 5) \\ &= \sum_{j=0}^{5} e^{-2.4}\frac{(2.4)^j}{j!} \end{aligned}$$

b) $P(\psi = 7) = e^{-2.4}\frac{(2.4)^7}{7!}$.

4. Al tratarse de una muestra de tamaño 4 sobre 25 individuos la probabilidad de cada muestra es

$$p = \frac{1}{\binom{25}{4}}$$

Que en una muestra de tamaño 4 haya sólamente uno blanca signica que hay 3 pardos. El número de posibilidades es

$$n = \binom{8}{1}\binom{17}{3}.$$

Así pués la probabilidad pedida es np.

5. Si $X \sim N(2,3)$ entonces $Y = \frac{X-2}{3} \sim N(0,1)$:

a)

$$\begin{aligned} P(X &\leq 6.32) = P\left(Y \leq \frac{6.32 - 2}{3}\right) = P\left(Y \leq 1.44\right) \\ &= F_\eta\left(1.44\right) \\ &= \end{aligned}$$

La otra probabilidad se calcula igual:

$$\begin{aligned} P(6.15 &\leq X \leq 7.35) = P(\frac{6.15-2}{3} \leq Y \leq \frac{7.35-2}{3}) \\ &= P\left(1.3833 \leq Y \leq 1.7833\right) \\ &= F_\eta\left(1.7833\right) - F_\eta\left(1.3833\right) \\ &= \end{aligned}$$

Téngase en cuenta n las últimas igualdades se emplean las tablas de la normal $N(0,1)$.

b) Para esta probabilidad se toman raíces cudradas:

$$\begin{aligned} P(X^2 &\geq 3.15) = P(|X| \geq \sqrt[2]{3.15}) \\ &= P(\left\{X \geq \sqrt[2]{3.15}\right\} \cup \left\{X \leq -\sqrt[2]{3.15}\right\}) \\ &= p_1 + p_2 \end{aligned}$$

Ahora bien,.

$$\begin{aligned} p_1 &= P(X \geq \sqrt[2]{3.15}) = 1 - P(X \leq \sqrt[2]{3.15}) = 1 - P(Y \leq \frac{\sqrt[2]{3.15}-2}{3}) \\ &= 1 - P(Y \leq x_0) \end{aligned}$$

pero el número $x_0 = -7.5059 \times 10^{-2}$ es negativo y por ello

$$\begin{aligned} P\left(Y \leq x_0\right) &= P\left(Y \geq -x_0\right) = 1 - P\left(Y \leq -x_0\right) \\ &= 1 - P\left(Y \leq 7.5059 \times 10^{-2}\right) \end{aligned}$$

Entonces

$$\begin{aligned} p_1 &= 1 - P(Y \leq x_0) = 1 - \left(1 - P\left(Y \leq 7.5059 \times 10^{-2}\right)\right) \\ &= P\left(Y \leq 7.5059 \times 10^{-2}\right) \end{aligned}$$

Por otra parte

$$\begin{aligned} p_2 &= P\left(Y \leq \frac{-\sqrt[2]{3.15}-2}{3} = -1.2583\right) \\ &= 1 - P\left(Y \leq 1.2583\right) \end{aligned}$$

Resta por buscar en las tablas de la normal.

6. Se hace igual que el número 5.
7. Sabemos que la v.a. es $\psi =$ número de llamadas que recibe una centralita en 5 minutos, y que $\psi \sim \mathcal{P}\left(5\right)$.

   a) La probabilida de tener 6 llamadas en 5 miniutos es

$$P\left(\psi = 6\right) = e^{-5}\frac{5^6}{6!}$$

b) La probabilidad de tener 3 en 10 minutos se descompone según 2 intervalos de 5 minutos: 3 en 10 minutos es el suceso que denotamos por S, así

$$S = \{0,3\} \cup \{1,2\} \cup \{2,1\} \cup \{3,0\}$$

donde $\{i,j\}$ significa recibir i llamadas en los primeros cinco minutos y j en los 5 restantes. Puesto que la unión anterior es de disjuntos entonces la probabilidad pedida es

$$p = P(\{0,3\}) + P(\{1,2\}) + P(\{2,1\}) + P(\{3,0\})$$

El cálculo de la parobabilidad de $\{i,j\}$ se efectúa como sigue:

$$\begin{aligned} P(\{i,j\}) &= P(\text{recibir } i \text{ llamadas})\, P(\text{recibir } j \text{ llamadas}) \\ &= \left(e^{-5}\frac{5^i}{i!}\right)\left(e^{-5}\frac{5^j}{j!}\right). \end{aligned}$$

Los detalles del cálculo se omiten.

8. Como la proporción de defectuosos es el 3 % escribiremos $P(D) = 0.03$. Los fusibles se empaquetan en cajas de 24 unidades. Se pide:

a) Al empaquetarse de 24 en 24 estamos considerando una muesra de tamaño 24. Si en ésta nos fijamos en el número de defectuosos lo que hacemos es construir una v.a.ξ, de distribución binomial $B(n,p)$ con $n = 24$ y $p = 0.03$. En este apartado nos están pidiendo $P(\xi \geq 1)$. Calculamos pasando al complementario:

$$\begin{aligned} P(\xi \geq 1) &= 1 - P(\xi = 0) \\ &= 1 - \binom{24}{0} 0.03^0 (1-0.03)^{24} \\ &= 1 - (1-0.03)^{24} \end{aligned}$$

b) Sea ahora p la probabilidad de que no haya ningún fusible defectuoso en una caja de 25, i.e. $p = (1-0.03)^{24} = 0.481\,42$. Si selecionamos 5 cajas consideramos una muestra de tamaño 5 y una v.a. η que da el número de cajas que no tienen fusible defectuoso. Hecho esto, lo que nos piden es $P(\eta = 2)$. Pero como $\eta \sim B(5, 0.481\,42)$ entonces

$$P(\eta = 2) = \binom{5}{2} (0.481\,42)^2 (1-0.481\,42)^3$$

c) Ahora se trata de una v.a. ψ que se distribuye según una geomética de parámetro p, con p la probabilidad de que 1 caja contenga algún fusible defectuoso (estamos identificando el suceso que una caja contenga fusible defectuoso con éxito). Esta probabilidad ya la sabemos del primer apartado, $p = 1 - 0.481\,42 = 0.518\,58$. ψ es el número de observaciones hasta obtener el primer éxito, i.e. $\psi \sim G(0.518\,58)$. Como $E[\psi] = \frac{1}{p} = \frac{1}{0.518\,58} = 1.928\,3$, diremos pues que el número de cajas hasta obtener una con fusibles defectuosos es 2.

9. Se trata de un vector aleatorio de tipo continuo. De éste sabemos una marginal y la condicionada.

a) La función de densidad de probabilidad conjunta se calcula mediante la fórmula

$$f_\xi(x,y) = f(y|\, x)\, f_{\xi_1}(x)$$

y esto para nuestro caso es

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

b) La margina de ξ_1 ya está dada, la de ξ_2 es

$$\begin{aligned} f_{\xi_2}(y) &= \int_{\mathbb{R}} f_\xi(x,y)\, dx \\ &= \begin{cases} \int_0^{0.\,5} \left(\frac{1}{2} + x(y-1)\right) dx & \text{si } y \in (0,2) \\ 0 & \text{en el resto.} \end{cases} \\ &= \begin{cases} 0.\,125 + 0.\,125\, y & \text{si } y \in (0,2) \\ 0 & \text{en el resto.} \end{cases} \end{aligned}$$

c) La velocidad esperada para la segunda partícula cuando la velocidad de la primera es de $0.\,75$ unidades viene dado por

$$\begin{aligned} E[\xi_2|\, \xi_1 &= 0.\,75] = \int_{\mathbb{R}} y f_\xi(y/0.\,75)\, dy \\ &= \int_0^2 y\left(\frac{1}{2} + (0.\,75)(y-1)\right) dy = 1.\,5. \end{aligned}$$

10. La función de densidad de probabilidad

$$f_\xi(x,y) = \begin{cases} x+y & \text{si } x \in (0,1) \text{ e } y \in (0,1) \\ 0 & \text{en el resto,} \end{cases}$$

se pide:
a)

$$\begin{aligned} P\{\xi_1 \le 0.\,5, \xi_2 \le 0.\,2\} &= \int_0^{0.\,5} \left(\int_0^{0.\,2} f_\xi(x,y)\, dy\right) dx \\ &= \int_0^{0.\,5} \left(\int_0^{0.\,2} (x+y)\, dy\right) dx \\ &= \int_0^{0.\,5} (0.\,2x + 0.\,02)\, dx \\ &= 0.\,035 \end{aligned}$$

b) Para saber si ξ_1 y ξ_2 son independientes calculamos sus marginales y vemos si su producto da la densidad conjunta.

La marginal de ξ_1 es

$$\begin{aligned} f_{\xi_1}(x) &= \begin{cases} \displaystyle\int_0^1 (x+y)\,dy & \text{si } x \in (0,1) \\ 0 & \text{en el resto,} \end{cases} \\ &= \begin{cases} x + \frac{1}{2} & \text{si } x \in (0,1) \\ 0 & \text{en el resto,} \end{cases} \end{aligned}$$

y la de ξ_2 igual:

$$\begin{aligned} f_{\xi_2}(y) &= \begin{cases} \displaystyle\int_0^1 (x+y)\,dx & \text{si } y \in (0,1) \\ 0 & \text{en el resto,} \end{cases} \\ &= \begin{cases} y + \frac{1}{2} & \text{si } y \in (0,1) \\ 0 & \text{en el resto,} \end{cases}, \end{aligned}$$

Como

$$f_{\xi_1}(x)\, f_{\xi_2}(y) \neq f_{\xi}(x,y)$$

entonces ξ_1 y ξ_2 no son independientes.

c) Se sabe

$$E\left[\xi_1|_{\xi_2=0.5}\right] = \int_{\mathbb{R}} x f_{\xi_1}(x/0.5)\,dx,$$

falta determinar la densidad condicionada $f_{\xi}(x/0.5)$. Es fácil,

$$f_{\xi}(x/0.5) = \frac{f_{\xi}(x,0.5)}{f_{\xi_1}(0.5)} = \frac{x+0.5}{0.5+\frac{1}{2}} = x + 0.5$$

Así

$$E\left[\xi_1|_{\xi_2=0.5}\right] = \int_0^1 x\,(x+0.5)\,dx = 0.583\,33$$

:

11. Se hace igual que los dos ejecicios anteriores.
12. Tenemos la tabla

$\xi_2 \backslash \xi_1$	$x_1 = 1$	$x_2 = 2$	$x_3 = 3$
$y_1 = 0$	0.1	0.3	0.2
$y_2 = 1$	0.06	0.18	0.16

a) Para saber si son independientes las variables ξ_1 y ξ_2 calculamos las marginales:

$$P_{\xi_1}(x) = \sum_{j=1}^{2} P_{\xi}(x, y_j) = \begin{cases} 0.1 + 0.06 = 0.16 & \text{si } x = 1 \\ 0.48 & \text{si } x = 2 \\ 0.36 & \text{si } x = 3 \end{cases}$$

y

$$P_{\xi_2}(y) = \sum_{i=1}^{2} P_{\xi}(x_i, y) = \begin{cases} 0.1 + 0.3 + 0.2 = 0{,}6 & \text{si } y = 0 \\ 0.4 & \text{si } y = 1 \end{cases}$$

Puesto que

$$\begin{aligned} 0.1 &= P_{\xi}^{*}(1,0) = \\ P(\xi_1 = 1, \xi_2 = 0) &\neq P(\xi_1 = 1) P(\xi_2 = 0) \\ &= P_{\xi_1}^{*}(1) P_{\xi_2}^{*}(0) \\ &= (0.6)(0.16) = 0{,}096 \end{aligned}$$

entonces no son independientes.

b) $E[\xi_1] = \sum x_i P_{\xi_1}^{*}(x_i) = 1(0.16) + 2(0.48) + 3(0.36) = 2.2$

c) Para hallar la esperanza condicionada de $\xi_1 | \xi_2 = 0$ hemos de saber quién es la distribución de esa v.a. Sabemos que

$$\begin{aligned} & P(\xi_1 = x | \xi_2 = 0) \\ = & \frac{P(\xi_1 = x, \xi_2 = 0)}{P(\xi_2 = 0)} \\ = & \begin{cases} \frac{0.1}{0.6} = 0.16667 & \text{si } x = 1 \\ \frac{0.3}{0.6} = 0.5 & \text{si } x = 2 \\ \frac{0.2}{0.6} = 0.33333 & \text{si } x = 3 \end{cases} \end{aligned}$$

Así

$$\begin{aligned} E[\xi_1 | \xi_2 &= 0] = \sum x_i P(\xi_1 = x_i | \xi_2 = 0) \\ &= 1(0.16667) + 2(0.5) + 3(0.33333) = 2.1667 \end{aligned}$$

13. Se sabe que ξ = duración, en segundos, precisa para que se produzca una reacción entre dos compuestos entonces $\xi \sim N(\mu, \sigma)$ y además

$$\begin{aligned} P(\xi \geq 40) &= 0.6, \\ P(\xi \leq 50) &= 0.55 \end{aligned}$$

Tipificando se llega a que $X = \frac{\xi - \mu}{\sigma}$ sigue una $N(0,1)$ y que usando las tablas de la normal las identidades

$$\begin{aligned} P\left(\frac{\xi - \mu}{\sigma} \geq \frac{40 - \mu}{\sigma}\right) &= 0.6, \\ P\left(\frac{\xi - \mu}{\sigma} \leq \frac{50 - \mu}{\sigma}\right) &= 0.55 \end{aligned}$$

se ha de cumplir

$$\begin{aligned} -\frac{40 - \mu}{\sigma} &= 0.255 \\ \frac{50 - \mu}{\sigma} &= 0.125 \end{aligned}$$

lo que da $\sigma = 26.316, \mu = 46.711$

14. Se hace igual que el anterior.

15. Se sabe que $\xi_1 \sim N(1, 0.1)$, que $\xi_2 \sim N(\mu, \sigma)$ entonces $\xi = \xi_1 + \xi_2 \sim N(1+\mu, \sigma_0)$. Al ser ξ_1 y ξ_2 v.a. independientes
$$\sigma_0^2 = (0.1)^2 + \sigma^2$$
Pero como sabemos que $\xi \sim N(1.5, 0.3)$ entonces
$$\begin{aligned} 1.5 &= 1+\mu \\ (0.3)^2 &= (0{,}1)^2 + \sigma^2 \end{aligned}$$
Resolvemos el sistema y obtenemos como única solución admisible $\mu = 0.5$ y $\sigma = 0.28284$.

16. Sea ξ la duración de una sábana. Esta v.a. sigue una distribución normal, de media $\mu = 50$ y desviación típica $\sigma = 8$.
   a) Calculamos la probabilidad de que una sábana dure 35 días sin romperse. Esto es lo mismo que
   $$\begin{aligned} P(\xi \leq 35) &= P\left(\frac{\xi - 50}{8} \leq \frac{35-50}{8} = -1.875\right) \\ &= P(\eta \leq -1.875) = 1 - P(\eta \leq 1.875) \\ &= 1 - 0.969 = 0.031 \end{aligned}$$
   (se ha tipificado y se han usado las tablas de la normal). Podemos decir que poco más del 3 % se rompen. Como hay 300, entonces se romperán $300 \times 0.031 = 9.3$, es decir, entre 9 y 10 sábanas.
   b) El cálculo es el mismo, hemos de evaluar $P(\xi \leq 60)$ y para ello hemos de tipifica ξ y emplear las tablas de la normal

CAPíTULO 10

Soluciones Tema 4

1. Las tablas de la $\mathcal{F}$ de Fisher-Snedecor da lugar a:

$$\begin{aligned} \mathcal{F}_{5,21,0,05} &= 4.68 \\ \mathcal{F}_{24,12,0,01} &= 3.78 \\ \mathcal{F}_{8,4,0,05} &= 21.35 \end{aligned}$$

2. Usamos las tablas de la $\mathcal{X}^2$
   a) $P(\mathcal{X}^2_{10} \leq 20.4) = 1 - 0.025$
   b) $P(\mathcal{X}^2_{14} > 23.2) = 0.05$
   c) $P(11.9 \leq \mathcal{X}^2_{18} \leq 32.4) = (1 - 0.02) - (1 - 0.85) = 0.83$
3. $\mathcal{X}^2_{30,0,025} = 46.979$ y $\mathcal{X}^2_{24,0,95} = 13.848$.
4. Se sabe que $\dfrac{(n-1)\,S^2}{\sigma^2} \sim \mathcal{X}^2_{n-1}$. Entonces

$$\begin{aligned} P(0.5 &< \frac{S^2}{\sigma^2} < 1.8) \\ &= P((0.5)\,15 < \frac{15S^2}{\sigma^2} < (1.8)\,15) \\ &= P(7.5 < \frac{15S^2}{\sigma^2} < 27) \\ &= P(7.5 < \mathcal{X}^2_{15} < 27) \\ &= (1 - 0.025) - (1 - 0.95) \end{aligned}$$

5. $t_{15,0,1} = 1.341$, $t_{25,0,2} = 0.856$ y $t_{8,0.05} = 1.86$.
6. Del enunciado se deduce que si ξ_i es la producción del i-ésimo día, entonces $\xi_i \sim U(6,10)$, expresado en miles de unidades. Así, para todo i

$$E[\xi_i] = \frac{6+10}{2} = 8$$

y

$$V[\xi_i] = \frac{(10-6)^2}{12} = \frac{4}{3}$$

Por otro lado la producción media es $\overline{\xi} = \frac{\sum_{i=1}^{n} \xi_i}{n}$, donde $n = 320$. Nos están pidiendo

$$P\left(\overline{\xi} \geq 8.1\right).$$

Empleando el Teorema Central del Límite tendremos:

$$\begin{aligned} P\left(\overline{\xi} \geq 8.1\right) &= P\left(\frac{\overline{\xi}-8}{\frac{4}{3\sqrt[2]{320}}} \geq \frac{(8.1-8)}{\frac{4}{3\sqrt[2]{320}}}\right) \\ &\overset{TCL}{\approx} P\left(\eta \geq 1.3416\right) \\ &= 1-P\left(\eta \leq 1.3416\right) \\ &= 1-0.909 = 0.091 \end{aligned}$$

7. La v.a. η_n se puede expresar como suma de n v.a. ξ_i, donde ξ_i es el número de puntos del lanzamiento i-ésimo. Así $\eta_n = \sum_{i=1}^n \xi_i$. Las ξ_i son independientes e idénticamente distribuidas. Se sabe que ξ_i es discreta con masa $\frac{1}{6}$ en cada punto 1, 2, 3, 4, 5 y 6. Es fácil ver que

$$\begin{aligned} E[\xi_i] &= 1\frac{1}{6}+2\frac{1}{6}+...+6\frac{1}{6} = \frac{1}{6}(1+2+...+6) = \frac{7}{2}, \\ E[\xi_i^2] &= 1^2\frac{1}{6}+2^2\frac{1}{6}+...+6^2\frac{1}{6} = \frac{1}{6}\left(1+2^2+...+6^2\right) = \frac{1}{6}\left(\frac{42\,(13)}{6}\right) = \frac{45}{3}, \\ V[\xi_i] &= \frac{45}{3}-\frac{49}{4} = \frac{11}{4}. \end{aligned}$$

Sabiendo esto espondamos a las cuestiones:

a)

$$\begin{aligned} E\left[\eta_n\right] &= E\left[\sum_{i=1}^n \xi_i\right] = \sum_{i=1}^n E\left[\xi_i\right] \\ &= n\frac{7}{2} \end{aligned}$$

$$\begin{aligned} V\left[\eta_n\right] &= V\left[\sum_{i=1}^n \xi_i\right] = \sum_{i=1}^n V\left[\xi_i\right] \\ &= n\frac{11}{4} \end{aligned}$$

b) La v.a. $\psi = \frac{\eta_n}{n}$ es la media muestral luego

$$E[\frac{\eta_n}{n}] = \frac{7}{2}$$

y

$$\begin{aligned} V\left[\frac{\eta_n}{n}\right] &= \frac{1}{n^2}V\left[\eta_n\right] = \frac{1}{n^2}n\frac{11}{4} \\ &= \frac{11}{4n} \end{aligned}$$

Gracias al Teorema Central del Límite

$$P\left\{\left|\frac{\eta_n}{n}-3.5\right|\geq 0.1\right\}=1-P\left\{\left|\frac{\eta_n}{n}-3.5\right|\leq 0.1\right\}$$

y

$$\begin{aligned} P\left\{\left|\frac{\eta_n}{n}-3.5\right|\leq 0.1\right\} &= P\left\{\frac{\left|\frac{\eta_n}{n}-3.5\right|}{\sqrt[2]{\frac{11}{4n}}}\leq\frac{0.1}{\sqrt[2]{\frac{11}{4n}}}\right\} \\ &\approx P\left\{\eta\leq\frac{0.1}{\sqrt[2]{\frac{11}{4n}}}\right\} \\ &= P\left(\eta\leq 6.030\,2\times 10^{-2}\sqrt{n}\right) \end{aligned}$$

con $\eta\sim N(0,1)$. Así

$$P\left\{\left|\frac{\eta_n}{n}-3.5\right|\geq 0.1\right\}\approx 1-P\left(\eta\leq 6.030\,2\times 10^{-2}\sqrt{n}\right)=0.1$$

y por ende

$$P\left(\eta\leq s\right)=0.9$$

con $s=6.030\,2\times 10^{-2}\sqrt{n}$. Buscamos en las tablas de la normal para asegurar que $s=1.285$. Por lo tanto

$$6.030\,2\times 10^{-2}\sqrt{n}=1.285$$

y así $n=454.09$.

8. Previamente escribiremos $\xi=\sum_{i=1}^{1000}\xi_i$ y usaremos el hecho de que la esperanza de las ξ_i $\frac{1}{2}$, y varianza $\frac{1}{4}$. Con esto y el Teorema Central del Límite tenemos

$$\begin{aligned} P\Bigg(\frac{\frac{490}{1000}-\frac{1}{2}}{\frac{\sqrt[2]{\frac{1}{4}}}{\sqrt[2]{1000}}} &< \frac{\frac{\sum_{i=1}^{1000}\xi_i}{1000}-\frac{1}{2}}{\frac{\sqrt[2]{\frac{1}{4}}}{\sqrt[2]{1000}}}<\frac{\frac{k}{1000}-\frac{1}{2}}{\frac{\sqrt[2]{\frac{1}{4}}}{\sqrt[2]{1000}}}\Bigg) \\ \overset{TCL}{\approx} P\left(-\frac{1}{5}\sqrt{10}\right. &< \left.\eta<\frac{1}{50}\left(k-500\right)\sqrt{10}\right) \\ &= F_\eta\left(s\right)-F_\eta\left(-\frac{1}{5}\sqrt{10}\right) \end{aligned}$$

con $\eta\sim N(0,1)$ y $s=\frac{1}{50}\left(k-500\right)\sqrt{10}$. Buscamos en las tablas de la normal para cerciorarnos de que

$$\begin{aligned} F_\eta\left(-\frac{1}{5}\sqrt{10}\right) &= 1-F\left(0.632\right) \\ &= 1-0.7357=0.264\,3, \end{aligned}$$

y que por tanto

$$F_\eta\left(s\right)=0.7+F_\eta\left(-\frac{1}{5}\sqrt{10}\right)=0.764\,3$$

Según las tablas lo anterior implica que $s = 0.72$, i.e.

$$0.72 = \frac{1}{50}(k - 500)\sqrt{10}$$

De la última expresión obtenemos $k = 511.38$.

9. Se procede igual que antes pero teniendo en cuenta que ahora son v.a. uniformes $(0,1)$ y que por tanto, $\mu = \frac{1}{2}$ y $\sigma^2 = \frac{1}{12}$. Así, con $n = 50$, empleando el Teorema Central del Límite y buscando en las tablas, se tiene que

$$\begin{aligned} P\left\{\overline{X} < 0.4\right\} &= P\left\{\frac{\overline{X} - \frac{1}{2}}{\frac{1}{\sqrt[2]{50}\sqrt[2]{12}}} < \frac{0.4 - \frac{1}{2}}{\frac{1}{\sqrt[2]{50}\sqrt[2]{12}}}\right\} \\ &\overset{TCL}{\approx} P(\eta \leq -2.4495) \\ &= 1 - P(\eta \leq 2.4495) \\ &= 1 - 0.992 = 0.008. \end{aligned}$$

10. $\alpha_1 = 1/5$ y $\alpha_2 = \frac{2}{25}$ permiten calcular la varianza:

$$\sigma^2 = V[\xi] = \frac{2}{25} - \left(\frac{1}{5}\right) = \frac{1}{25}.$$

La longitud media de 100 filamentos es

$$\overline{\xi} = \frac{\sum_{i=1}^{100} \xi_i}{100}$$

donde las ξ_i son las v.a. que dan las mediciones de cada filamento. Nos están pidiedo

$$P\left(0.18 \leq \overline{\xi} \leq 0.22\right).$$

Esta expresión es idéntica a

$$P\left(\frac{0.18 - 1/5}{\frac{1}{50}} \leq \frac{\overline{\xi} - 1/5}{\frac{1}{50}} \leq \frac{0.22 - 1/5}{\frac{1}{50}}\right)$$

y ésta, según el Teorema Central del Límite, es aproximadamente

$$P(-1 \leq \eta \leq 1)$$

con η una normal $(0,1)$. Resta buscar en las tablas de la normal.

11. La probabilidad de recibir un cheque sin fondos a lo largo de una semana es 0.15, esto es $P(SF) = 0.15$. Sea una muestra de tamaño $n = 1000$.
   a) Sea $\eta_{1000} = \sum_{i=1}^{1000} \xi_i$ donde $\xi_i = 1$ si es SF el i-ésimo cheque e $= 0$ en otro caso. $\xi_i \sim B(1, 0.15)$, $E[\xi_i] = 0.15$ y $V[\xi_i] = 0.15(1 - 0.15)$. La probabilidad de que

en tal semana se reciban como máximo 125 cheques sin fondos es $P(\eta_n \leq 125)$ y

$$\begin{aligned} & P(\eta_{1000} \leq 125) \\ = \; & P\left(\frac{\frac{\eta_{1000}}{1000} - 0.15}{\frac{\sqrt[2]{0.15(1-0.15)}}{\sqrt[2]{1000}}} \leq \frac{\frac{125}{1000} - 0.15}{\frac{\sqrt[2]{0.15(1-0.15)}}{\sqrt[2]{1000}}}\right) \\ & \overset{TCL}{\approx} P(\eta \leq -2.214) \\ = \; & 1 - P(\eta \leq 2.214) = \\ = \; & 1 - 0.983 = 0.017. \end{aligned}$$

b) Para tal beneficio se necitan 50 cheques. Sea $\eta_n = \sum_{i=1}^{n} \xi_i$. Nos están pidiendo n tal que

$$P(\eta_n \geq 50) = 0.9$$

Usamos de nuevo el TCL:

$$\begin{aligned} P(\eta_n \geq 50) &= P\left(\frac{\frac{\eta_n}{n} - 0.15}{\frac{\sqrt[2]{0.15(1-0.15)}}{\sqrt[2]{n}}} \leq \frac{\frac{50}{n} - 0.15}{\frac{\sqrt[2]{0.15(1-0.15)}}{\sqrt[2]{n}}}\right) \\ &= P(\eta \leq s) \end{aligned}$$

con $s = \frac{\frac{50}{n}-0.15}{\frac{\sqrt[2]{0.15(1-0.15)}}{\sqrt[2]{n}}}$. Como la probabilidad anterior queremos que sea 0.9, entonces, según las tablas de la normal, $s = 1.285$, luego

$$\frac{\frac{50}{n} - 0.15}{\frac{\sqrt[2]{0.15(1-0.15)}}{\sqrt[2]{n}}} = 1.285$$

de donde se sigue que $n = 281.97$. Esto es, han de recibirse unos 282 cheques sin fondos.

12. De una población $N(\mu, \sigma)$ y una muestra aleatoria simple $\xi_1, ..., \xi_n$, se pretende estimar su varianza σ^2 a través del estadístico

$$\xi^* = \frac{\sum_{i=1}^{n} \xi_i^2}{n}.$$

Se pide:

a)

$$\begin{aligned} E[\xi^*] &= E[\frac{\sum_{i=1}^{n} \xi_i^2}{n}] = \frac{1}{n} E[\sum_{i=1}^{n} \xi_i^2] \\ &= \frac{1}{n} \sum_{i=1}^{n} E[\xi_i^2] = \frac{1}{n} n E[\xi_i^2] \\ &= E[\xi_i^2] \end{aligned}$$

y como $E[\xi_i^2]$ es elmomento de orden dos de una normal (μ, σ) entonces $E[\xi_i^2] = \sigma^2 + \mu^2$. Por esto

$$E[\xi^*] = \sigma^2 + \mu^2$$

b) Es insesgado cuando $E[\xi^*] = \sigma^2$, pero como $E[\xi^*] = \sigma^2 + \mu^2$, sólo será insegado cuando $\mu = 0$.

c) ξ^* es eficiente cuando es insesgado y su varianza es mínima. Nos limitaremos a ver cuando es insesgado y a calcular su varianza. Como queremos que sea centrado hemos de imponer $\mu = 0$; su varianza es

$$V[\xi^*] = V[\frac{\sum_{i=1}^n \xi_i^2}{n}] = \frac{1}{n} V[\xi_i^2].$$

Ahora bien, como $\mu = 0$ entonces $\xi_i \sim N(0, \sigma)$, luego $\frac{\xi_i}{\sigma} \sim N(0,1)$ y así $\left(\frac{\xi_i}{\sigma}\right)^2 \sim \chi_1^2$. Entonces

$$V[\xi_i^2] = V[\sigma^2 \left(\frac{\xi_i}{\sigma}\right)^2] = \sigma^4 V[\left(\frac{\xi_i}{\sigma}\right)^2] = \sigma^4 V\left[\chi_1^2\right] = 2\sigma^4$$

Por tanto $V[\xi^*] = \frac{2\sigma^4}{n}$(Ayuda: usar que $E\left[\chi_1^2\right] = 1$ y que $V\left[\chi_1^2\right] = 2$)

13. Si la compañía reserva 160 plazas, ¿cuál es la probabilidad de que, al menos un pasajero no tenga plaza disponible a la hora de embarcar?

$$P\left\{\sum_{i=1}^n \xi_i \geq 145\right\} = P\left\{\eta = \frac{\sum_{i=1}^n \xi_i - n\,(0.88)}{\sqrt{n\,(0.88)\,(0.12)}} > \frac{145 - n\,(0.88)}{\sqrt{n\,(0{,}88)\,(0.12)}}\right\} = 0.99.$$

Empleando el TCL tendremos la aporximación

$$P\left\{\sum_{i=1}^n \xi_i \geq 145\right\} \approx P\left\{N(0,1) > c\right\} = 0.99$$

con $c = \frac{145 - n(0.88)}{\sqrt{n(0.88)(0.12)}}$. Pero esta estimación implica (mirar tablas de la normal)

$$c = \frac{145 - n\,(0.88)}{\sqrt{n\,(0.88)\,(0.12)}} = -2.325,$$

cuya solución aproximada es $n = 177$. Basta que el número de reservas sea mayor o igual que 177 para cubrir 145 de las plazas.

La segunda cuestión se resuelve de forma parecida:

$$\begin{aligned} P\left\{\sum_{i=1}^{160} \xi_i \geq 151\right\} &= P\left\{\eta \geq \frac{151 - (0.88)\,(160)}{\sqrt{0.1056\,(160)}} = 2.4815\right\} \\ &\approx 1 - P\left\{\eta < 2.4815\right\} = 1 - 0.9934 \\ &= 0.0066 \end{aligned}$$

14. Una ξ sigue la misma distribución que la suma de n v.a. $\xi_1, ..., \xi_n$ de tipo $B(1,p)$. Por tanto en lugar de escribir ξ tomaremos $\sum_{i=1}^n \xi_i$ y así

$$\begin{aligned}
P\{\xi > n/2\} &= P\left\{\sum_{i=1}^n \xi_i > n/2\right\} \\
&= P\left\{\frac{\sum_{i=1}^n \xi_i - np}{\sqrt{V\left[\sum_{i=1}^n \xi_i\right]}} > \frac{n/2 - np}{\sqrt{V\left[\sum_{i=1}^n \xi_i\right]}}\right\} \\
&= P\left\{N(0,1) > \frac{n/2 - np}{\sqrt{np(1-p)}}\right\} \geq 0.9
\end{aligned}$$

si y sólo si $\frac{n/2-np}{\sqrt{np(1-p)}} \leq -1.285$. Por ejemplo con $p = 0.65 =$ queda

$$-\frac{3}{20}n = n/2 - n\,(13/20) \leq -1.285\sqrt{n\,(91/400)}$$

i.e.

$$(3n/20)^2 \geq \left(1.285\sqrt{n\,(91/400)}\right)^2 \approx (0.375)\,n$$

con lo cual bastará con elegir

$$n \geq \frac{400}{9}(0.375) > 17.$$

CAPíTULO 11

Soluciones Tema 5

1. Para una m. a. s. de tamaño n la función de verosimilitud es
$$\begin{aligned} L &= f_{\xi_1}(x_1)\ldots f_{\xi_n}(x_n) \\ &= \begin{cases} 2^n\theta^n\Pi_{i=1}^n x_i \exp\left(-\theta\sum_{i=1}^n x_i^2\right) & \text{si } x_i \geq 0 \text{ para todo } i \\ 0 & \text{si algún } x_i < 0. \end{cases} \end{aligned}$$
La segunda función de verosimilitud (en la zona donde se puede definir ésta) es
$$\begin{aligned} l &= \log(L) = \log\left(2^n\theta^n\Pi_{i=1}^n x_i \exp\left(-\theta\sum_{i=1}^n x_i^2\right)\right) \\ &= n\log 2 + n\log\theta + \sum_{i=1}^n \log x_i - \theta\sum_{i=1}^n x_i^2 \end{aligned}$$
El máximo se alcanzará en alguno de sus puntos crítiticos. Resolvemos la ecuación
$$\frac{dl}{d\theta} = 0$$
i.e.
$$\frac{n}{\theta} - \sum_{i=1}^n x_i^2 = 0$$
Se sigue que en $\theta = \frac{n}{\sum_{i=1}^n x_i^2}$ hay un máximo absoluto[1] y por ende que
$$\theta^* = \frac{n}{\sum_{i=1}^n \xi_i^2}$$
es el EMV.
2. Se cumple que
$$L = \begin{cases} \frac{\Pi_{i=1}^n x_i^2}{2^n\theta^{3n}} \exp\left(-\sum_{i=1}^n x_i/\theta\right) & \text{si } x_i \geq 0 \text{ para todo } i \\ 0 & \text{en el resto} \end{cases}$$
y
$$l = -n\log 2 - 3n\log\theta + \sum_{i=1}^n \left(\log x_i^2\right) - \sum_{i=1}^n (x_i/\theta)$$

[1] Que es un punto de máximo absoluto se cumple gracias a que l es cóncava. Esto se puede verificar comprobando que $l'' < 0$.

Por tanto $l' = 0$ es lo mismo que

$$\frac{-3n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} x_i = 0$$

De aquí se obtiene que en

$$\theta = \frac{\sum_{i=1}^{n} x_i}{3n}$$

se alcanza el máximo absoluto (¿por qué?) y que el EMV es

$$\theta^* = \frac{\sum_{i=1}^{n} \xi_i}{3n}.$$

3. Hemos de estimar un parámetro en un modelo cuya esperanza es

$$E[\xi] = \int_{-1}^{1} x\left(\frac{1+\theta x}{2}\right) dx = \frac{1}{3}\theta$$

La ecuación con la que realizamos la estimación por elmétodo de los momentos es

$$\alpha_1 = M_1 = \overline{\xi}$$

De aquí deducimos que

$$\theta^* = 3\overline{\xi}$$

es el estimador pedido.

4. Calculamos los momentos de orden uno y dos:

$$\begin{aligned}\alpha_1 &= (-1)\frac{1-\theta_1}{2} + \frac{1-\theta_2}{2}\\ \alpha_2 &= (-1)^2\frac{1-\theta_1}{2} + \frac{1-\theta_2}{2}\end{aligned}$$

Las estimaciones de los parámetros θ_1 y θ_2 se deducen del sistema

$$\begin{aligned}-\frac{1-\theta_1}{2} + \frac{1-\theta_2}{2} &= \overline{\xi}\\ \frac{1-\theta_1}{2} + \frac{1-\theta_2}{2} &= M_2\end{aligned}$$

donde $M_2 = \frac{\sum_{i=1}^{n} x_i^2}{n}$. La solución es $\theta_1 = 1 + \overline{\xi} - M_2, \theta_2 = 1 - \overline{\xi} - M_2$.

5. Se sabe que para este caso la ecuación a resolver es

$$\alpha_1 = \overline{\xi},$$

y

$$\alpha_1 = \int_0^1 x\theta x^{\theta-1} dx = \frac{\theta}{1+\theta} = s$$

Despejamos θ y se obtiene el estimador

$$\theta = \frac{\overline{\xi}}{1-\overline{\xi}}$$

6.
   a) Por el método de los momentos hemos de plantear
   $$\alpha_1 = \frac{1}{p} = \overline{\xi}$$
   de donde resulta que el estimador pedido es
   $$p^* = \frac{1}{\overline{\xi}}$$
   b) Hecho en clase.
7. Por el método de los momentos: al ahaber un solo parámetro basta con identificar la media muestral con la poblacional;
$$\alpha_1 = M_1 = \overline{\xi}$$
y
$$\begin{aligned}\alpha_1 &= \int_\theta^\infty x e^{\theta - x} dx = -x e^{\theta - x} - e^{\theta - x}\Big|_\theta^\infty \\ &= \lim_{R\to\infty} \left(-R\, e^{\theta - R} - e^{\theta - R}\right) - (-\theta - 1) = \theta + 1\end{aligned}$$
Así
$$\theta + 1 = \overline{\xi}$$
con lo que
$$\theta^* = \overline{\xi} - 1$$
es el estimador por el método de los momentos (EMM)
Para hallar el EMV calculamos la función de verosimilitud:
$$L = f(x_1)\ldots f(x_n) = \begin{cases} \exp\left\{n\theta - \sum_{i=1}^n x_i\right\} & \text{si } x_i \in [\theta, +\infty) \text{ para todo } i \\ 0 & \text{en el resto}\end{cases}$$
y puesto que queremos hacer máxima la expresión $e^{n\theta - \sum x_i}$ habremos de tomar θ la más grande posible. Para ello obsevamos que $x_i \in [\theta, +\infty)$ para todo i, lo cual supone que $\min x_i \geq \theta$. De esto se desprende que el valor más grande que puede ser es $\min x_i$. El EMV es $\theta^* = \min \xi_i$.
8. Sea la variable aleatoria ξ cuya función de densidad de probabilidad es
$$f_\xi(x) = \begin{cases} \dfrac{\theta}{x^{\theta+1}} & \text{si } x \geq 1 \\ 0 & \text{en el resto,}\end{cases}$$
donde se supone que $\theta > 1$. Dada una muestra aleatoria simple de ξ y tamaño n, $\xi_1, \ldots \xi_n$, se pide:
   a) Determinar un estimador del parámetro θ por el método de los momentos.
   b) Obtener un estimador para θ por el método de máxima verosimilitud.

EMM:

$$\begin{aligned}\alpha_1 &= \int_1^{\infty} x\frac{\theta}{x^{\theta+1}}dx = \int_1^{\infty} \theta x^{-\theta}dx = \left.\frac{\theta x^{-\theta}}{-\theta+1}\right|_1^{+\infty} \\ &= \lim_{x\to\infty}\left(\frac{\theta x^{-\theta}}{-\theta+1}\right) - \left(\frac{\theta}{-\theta+1}\right) \\ &= \frac{\theta}{\theta-1}\end{aligned}$$

pues $\lim_{x\to+\infty}\left(\frac{\theta x^{-\theta}}{-\theta+1}\right) = 0$
EMV:

$$L = \frac{\theta}{x_1^{\theta+1}}\cdots\frac{\theta}{x_n^{\theta+1}}$$

si todoos los si $x_i \geq 1$, y vale cero en toro caso. Así, si todoos los si $x_i \geq 1$,

$$L = \frac{\theta^n}{\Pi x_i^{\theta+1}}$$

y la segund funciónde verosimilitud es

$$\begin{aligned}L &= \log L = \log\frac{\theta^n}{\Pi x_i^{\theta+1}} = n\theta - \log \Pi x_i^{\theta+1} \\ &= n\log\theta - (\theta+1)\sum \log x_i\end{aligned}$$

que derivado e igualado a cero proporciona el punto crítico:

$$\frac{n}{\theta} - \sum \log x_i = 0$$

y

$$\theta^* = \frac{n}{\sum \log \xi_i}$$

9. Sea ξ una variable aleatoria que modeliza a cierto fenómeno. Se sabe que

$$P\{\xi=-1\} = \frac{1-\theta}{2},\ P\{\xi=0\} = \frac{1}{2},\ P\{\xi=1\} = \frac{\theta}{2},$$

donde θ es cierto parámetro del que sólo se sabe que $\theta \in (0,1)$.

a) Dada una m.a.s. $\xi_1, ..., \xi_n$ de la v.a. ξ, determinar el estimador de máxima verosimilitud del parámetro θ.

b) Si se toma $n = 50$ y se realiza la muestra obteniéndose los resultados,

x_i :	-1	0	1
n_i :	10	25	15

donde n_i es la frecuencia del dato x_i (número de veces que aparece x_i), realizar una estimación del parámetro θ usando para ello el método de la máxima verosimilitud.

La función de verosimilitud está en función del número de veces que ocurren cada uno de los puntos de masa, -1, 0 y 1. Teniendo en cuenta que el tamaño de la muestra es n queda

$$\begin{aligned} L &= P^*(x_1)...P^*(x_n) \\ &= (P^*(-1))^s (P^*(0))^t (P^*(0))^{n-s-t} \\ &= \left(\frac{1-\theta}{2}\right)^s \left(\frac{1}{2}\right)^t \left(\frac{\theta}{2}\right)^{n-s-t} \end{aligned}$$

Entonces

$$L' = -s\left(\frac{1-\theta}{2}\right)^{s-1} \left(\frac{1}{2}\right)^{t+1} \left(\frac{\theta}{2}\right)^{n-s-t} + (n-s-t)\left(\frac{1-\theta}{2}\right)^{s} \left(\frac{1}{2}\right)^{t+1} \left(\frac{\theta}{2}\right)^{n-s-t-1}$$

y por ende los puntos críticos han de cumplir

$$s\left(\frac{1-\theta}{2}\right)^{s-1} \left(\frac{1}{2}\right)^{t+1} \left(\frac{\theta}{2}\right)^{n-s-t} = (n-s-t)\left(\frac{1-\theta}{2}\right)^{s} \left(\frac{1}{2}\right)^{t+1} \left(\frac{\theta}{2}\right)^{n-s-t-1}$$

que simplificado da lugar a

$$s\left(\frac{\theta}{2}\right) = (n-s-t)\left(\frac{1-\theta}{2}\right)$$

y así la EMV es

$$\theta^* = \frac{-n+s+t}{-n+t}.$$

Para contestar al apartado b basta sustituir los datos $n = 50$, $s = 10$, $t = 25$, quedando

$$\theta^* = \frac{-50+35}{-50+25} = \frac{3}{5}$$

10. El EMV se obtiene maximizando la función

$$l = \log(L) = \theta\left(\sum_{i=1}^{n} \log x_i\right) + n\log(\theta+1).$$

Como

$$l' = \left(\sum_{i=1}^{n} \log x_i\right) v + \frac{n}{\theta+1}$$

entonces

$$\theta = -\frac{\sum_{i=1}^{n} \log x_i + n}{\sum_{i=1}^{n} \log x_i}$$

es la estimación máximo verosimil y

$$\theta^* = -\frac{\sum_{i=1}^{n} \log \xi_i + n}{\sum_{i=1}^{n} \log \xi_i}$$

es el EMV.

11. Primero vamos a deducir de manera razonada un intervalo de confianza $1-\alpha$ para la varianza de una distribución normal $N(\mu,\sigma)$. Se toma como cantidad pivotal al estadístico $Q\left(\mu,\xi_1,...,\xi_n\right)=\frac{n-1}{\sigma^2}S^2\sim\chi^2_{n-1}$. Buscamos q_1 y q_2 tales que

$$P\left(q_1<\frac{n-1}{\sigma^2}S^2<q_2\right)=1-\alpha$$

esto es,

$$P\left(\frac{n-1}{q_2}S^2<\sigma^2<\frac{n-1}{q_1}S^2\right)=1-\alpha.$$

La distancia entre los extremos ha de minimizarse, y como esta es proporcional a q_2-q_1 entonces los puntos q_1 y q_2 han de tomarse lo más próximos posible. Por la forma que tiene la densidad de una χ^2_{n-1} la elección debe ser

$$q_1=\chi^2_{n-1,1-\frac{\alpha}{2}} \text{ y } q_2=\chi^2_{n-1,\frac{\alpha}{2}}.$$

El intervalo de confianza (1-$\alpha=0'95$, $\alpha=0'05$) pedido es

$$I=\left[\frac{n-1}{\chi^2_{n-1,\frac{\alpha}{2}}}S^2,\frac{n-1}{\chi^2_{n-1,1-\frac{\alpha}{2}}}S^2\right].$$

Aplicamos esto al ejemplo expuesto: realizamos el estadístico S^2 con ayuda de la muestra dada

$$\begin{aligned}
S^2 &= \frac{1}{n-1}\sum_{i=1}^{4}\left(\xi_i-\overline{\xi}\right)^2\\
&= \frac{1}{3}\left[2^2+4^2+(-4)^2+(-2)^2\right]\\
&= \frac{40}{3}.
\end{aligned}$$

$$\begin{aligned}
q_1 &= \chi^2_{3,0.\,025}=9.\,35,\\
q_2 &= \chi^2_{3,0.\,975}=0.\,216.
\end{aligned}$$

Por tanto el intervalo para este ejemplo es

$$I=\left[\frac{40}{9.\,35},\frac{40}{0.\,216}\right]=\left[4.\,278,181.\,18\right].$$

Como $1,2\notin I$ entonces no podremos afirmar que σ^2 es 1 ó 2 a un nivel de confianza 0.95.

12. Se desea calcular un intervalo de confianza del 99 % para el tiempo medio de reacción: Se emplea la cantidad pivotal

$$Q=\frac{\overline{\xi}-\mu}{\frac{\sigma}{\sqrt{n}}}\sim N(0,1).$$

$\alpha = 0{,}01$, $n = 100$. El intervalo pedido es

$$\begin{aligned} I &= \left[\overline{\xi} - Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{81}}, \overline{\xi} + Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{81}}\right] \\ &= \left[2 - Z_{\frac{0{,}01}{2}}\frac{0.6}{9}, 2 + Z_{\frac{0{,}01}{2}}\frac{0.6}{9}\right]. \end{aligned}$$

Puesto que $Z_{0{,}005} = 2.57$

$$I = \left[2 - 2.57\frac{0.6}{9}, 2 + 2.57\frac{0.6}{9}\right] = [1.8287, 2.1713].$$

En cuanto al apartado b) decir que se ha de emplear la cantidad pivotal

$$Q = \frac{\overline{\xi} - \mu}{\frac{S}{\sqrt{n}}} \sim \mathrm{t}_{n-1}.$$

El intervalo pedido es

$$\begin{aligned} I &= \left[\overline{\xi} - \mathrm{t}_{81-1,\frac{\alpha}{2}}\frac{S}{\sqrt{81}}, \overline{\xi} + \mathrm{t}_{81-1}\frac{S}{\sqrt{81}}\right] \\ &= \left[2 - \mathrm{t}_{81-1,\frac{0{,}01}{2}}\frac{0{,}6}{9}, 2 + \mathrm{t}_{81-1,\frac{0{,}01}{2}}\frac{0{,}6}{9}\right]. \end{aligned}$$

Puesto según las tablas sabemos que $\mathrm{t}_{80,\frac{0{,}01}{2}} = 2{,}639$, el intervalo pedido es

$$I = \left[2 - 2.639\frac{0.6}{9}, 2 + 2.639\frac{0.6}{9}\right] = [1.8241, 2.1759].$$

13. En forma resumida: sean las m.a.s $X_1, ..., X_{125}$, de distribución $N(\mu_1, 20)$ y que representan ventas a la CEE. Igualmente sean $Y_1, ..., Y_{100}$, de distribución $N(\mu_2, 10)$ las v.a. que modelan las ventas a EEUU. La cantidad pivotal a emplear en este caso es

$$\begin{aligned} Q &= \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \\ &= \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{20^2}{125} + \frac{10^2}{100}}}. \end{aligned}$$

Se deben elegir t_1 y t_2, situados a una distancia mínima (uno del otro) tal que

$$P\{t_1 \leq Q \leq t_2\} = 0.95 \quad (\alpha = 0.05),$$

De esto se deduce que

$$I = \left[\left(\overline{X} - \overline{Y}\right) - t_2\frac{1}{5}\sqrt{105}, \left(\overline{X} - \overline{Y}\right) - t_1\frac{1}{5}\sqrt{105}\right].$$

Recordamos que los extremos t_j empleados para que la longitud del intervalo sea mínima son $t_1 = -t_2$ y $t_2 = Z_{\alpha/2}$. Empleando los datos de la realización de la muestra el intervalo pedido es

$$\begin{aligned} I &= \left[50 - Z_{0.025}\left(\frac{1}{5}\sqrt{105}\right), 50 + Z_{0.025}\left(\frac{1}{5}\sqrt{105}\right)\right] \\ &= \left[50 - 1.96\left(\frac{1}{5}\sqrt{105}\right), 50 + 1.96\left(\frac{1}{5}\sqrt{105}\right)\right] \\ &= [45.983, 54.017]. \end{aligned}$$

14. Una población está representada por una variable aleatoria ξ, cuya función de distribución es

$$F_\xi(x) = \begin{cases} 1 - \left(\frac{1}{x}\right)^\theta & \text{si } x > 1, \\ 0 & \text{en el resto,} \end{cases}$$

y donde θ es un parámetro real positivo. Sea una muestra aleatoria simple $\xi_1, ..., \xi_n$ perteneciente a dicha población. Se pide:

a) Como $f_\xi(x) = \frac{dF_\xi(x)}{dx}$ entonces

$$f_\xi(x) = \begin{cases} x^{-1-\theta}\theta & \text{si } x > 1, \\ 0 & \text{en el resto,} \end{cases}$$

b)

$$L = \Pi_{i=1}^n f_{\xi_i}(x_i) = \begin{cases} \theta^n \Pi_{i=1}^n (x_i)^{-\theta-1} & \text{si } x_i > 1 \text{ para todo } i, \\ 0 & \text{en el resto,} \end{cases}$$

y

$$l = \log(L) = n \log\theta - (\theta + 1)\sum_{i=1}^n \log(x_i).$$

$$l' = \frac{n}{\theta} - \sum_{i=1}^n \log(x_i) =$$

implica que en

$$\theta = \frac{n}{\sum_{i=1}^n \log(x_i)}$$

se alcanza el máximo absoluto (es una función cóncava). Así pues, el EMV es

$$\theta^* = \frac{n}{\sum_{i=1}^n \log(\xi_i)}$$

Encontar un estimador del parámetro θ usando para ello el principio de máxima verosimilitud.

c) Basta tener en cuenta que

$$\begin{aligned}\alpha_1 &= \int_1^{+\infty} x\theta\,(x)^{-\theta-1}\,dx \\ &= \theta\int_1^{+\infty}(x)^{-\theta}\,dx = \frac{\theta}{-\theta+1}\,(x)^{-\theta+1}\Big|_1^{\infty} \\ &= \frac{\theta}{-\theta+1}\end{aligned}$$

supuesto que $-\theta+1<0$. Hecha esta suposición resta por despejar θ de la ecuación

$$\frac{\theta}{-\theta+1}=\overline{\xi}.$$

15. Sea

$$\begin{aligned}L &= f_{\xi_1}(x_1)f_{\xi_2}(x_2)...f_{\xi_n}(x_n) \\ &= \begin{cases}\left(\frac{2a}{1-a}\right)^n \Pi_{i=1}^n\,(x_i)^{\frac{2a}{1-a}-1} & \text{si todos los } x_i\in[0,1], \\ 0 & \text{en el resto.}\end{cases}\end{aligned}$$

y consideremos

$$\begin{aligned}l &= \log L \\ &= n\log\left(\frac{2a}{1-a}\right)+\left(\frac{2a}{1-a}-1\right)\sum_{i=1}^n \log x_i.\end{aligned}$$

Entonces

$$\begin{aligned}l' &= \frac{d}{da}\left(n\log\left(\frac{2a}{1-a}\right)+\left(\frac{2a}{1-a}-1\right)v\right) \\ &= \frac{-n+na-2a\sum_{i=1}^n\log x_i}{(-1+a)^2\,a}\end{aligned}$$

que igualado a cero da

$$-n+na-2av=0$$

esto es,

$$a=\frac{1}{1-2\frac{\sum_{i=1}^n\log x_i}{n}}$$

Por tanto el EMV es

$$a^*=\frac{1}{1-2\frac{\sum_{i=1}^n\log \xi_i}{n}}$$

Por el método de los momentos:

$$\begin{aligned}
\overline{\xi} &= \alpha_1 = \int_{\mathbb{R}} x f_\xi(x)dx \\
&= \frac{2a}{1-a}\int_0^1 x^{\frac{2a}{1-a}}dx \\
&= \frac{2a}{1-a}\left.\frac{x^{\frac{2a}{1-a}+1}}{\frac{2a}{1-a}+1}\right|_0^1 \\
&= \frac{2a}{1-a}\left.\frac{x^{\frac{a+1}{1-a}}}{\frac{2a}{1-a}+1}\right|_0^1 \\
&= \frac{2a}{a+1}
\end{aligned}$$

De esta expresión última despejamos a.

16. θ es 0,4 ó 0,5. Habremos de evaluar $P_\xi^*(6)$ cuando $\xi \sim B(15,\theta)$:

	$P_\xi^*(6)$
$\theta = 0{,}4$	$\binom{15}{6}(0{,}4)^6(0{,}6)^9 = 4.127\,8\times 10^{-5}$
$\theta = 0{,}5$	$\binom{15}{6}(0{,}5)^6(0{,}5)^9 = 3.051\,8\times 10^{-5}$

El valor mayor se corresponde con $\theta = 0{,}45$, luego ésta sería la estimación máximo verosimil del parámetro θ.

17. La función de verosimiliud es

$$L = \frac{1}{\theta}e^{-\frac{x_1}{\theta}}...\frac{1}{\theta}e^{-\frac{x_n}{\theta}} = \frac{1}{\theta^n}e^{-\frac{\sum_{i=1}^n x_1}{\theta}}$$

y la segunda función de verosimilitud es

$$l = \log L = -n\log\theta - \frac{\sum_{i=1}^n x_1}{\theta}$$

Maximizamos l y para ello derivalmos l e igualamos a cero al objeto de encontrar el máximo absoluto:

$$l' = \frac{-n\theta}{\theta^2} + \frac{\sum_{i=1}^n x_1}{\theta^2} = 0$$

i.e. $-n\theta + \sum_{i=1}^n x_1 = 0$. Por tanto el EMV es

$$\theta^* = \overline{\xi}.$$

Como $E[\theta^*] = E\left[\overline{\xi}\right] = E[\xi]$, bastará con calcular la integral

$$\begin{aligned}
E[\xi] &= \int_0^\infty x\frac{1}{\theta}e^{-\frac{x}{\theta}}dx = \lim_{R\to+\infty} -e^{-\frac{x}{\theta}}(x+\theta)\Big|_0^R \\
&= \theta
\end{aligned}$$

ya que $-e^{-\frac{x}{\theta}}(x+\theta) = \frac{-(x+\theta)}{e^{x/\theta}} = -\frac{\infty}{\infty}$ cuando $x \to +\infty$ y aplicando L'Hôpital llegamos a que su límite es cero. Por tanto el EMV es insegado.
Al ser insesgado su error coincide con la varianza:

$$\begin{aligned} R &= E\left[(\theta^* - \theta)^2\right] = V[\theta^*] = V\left[\overline{\xi}\right] \\ &= \frac{1}{n} V[\xi] \end{aligned}$$

Resta por tanto calcular $V[\xi]$:

$$V[\xi] = \alpha_2 - \alpha_1^2 = \alpha_2 - \theta^2$$

y

$$\begin{aligned} \alpha_2 \int_0^\infty &= \int x^2 \frac{1}{\theta} e^{-\frac{x}{\theta}} dx \\ &= -e^{-\frac{x}{\theta}} \left(x^2 + 2x\theta + 2\theta^2\right)\Big|_0^\infty \\ &= 2\theta^2 \end{aligned}$$

Por consiguiente el error o riesgo de θ^* es

$$E = \theta^2.$$

CAPíTULO 12

Soluciones Tema 6

1. Se definen los errores de tipo I y II como rechazar H_0 siendo verdadera y aceptar H_0 siendo falsa, respectivamente. Por tanto
$$\begin{aligned} P(E_1) &= P(\text{rechazar } H_0 | \, H_0 \text{ verdadera}) \\ &= P(x > 3 | \, \lambda = 1) \\ &= 1 - P(x = 2 | \, \lambda = 1) \\ &\quad - P(x = 1 | \, \lambda = 1) - P(x = 0 | \, \lambda = 1) \end{aligned}$$
y
$$P(x = x_0 | \, \lambda = 1) = e^{-1} \frac{1^{x_0}}{x_0!}$$
Por tanto
$$P(E_1) = 1 - \frac{e^{-1}}{0!} - \frac{e^{-1}}{1!} - \frac{e^{-1}}{2!} = 8.\,030\,1 \times 10^{-2}$$
De modo similar se calcular la probabilidad de cometer el error de tipo II:
$$\begin{aligned} P(E_2) &= P(\text{aceptar } H_0 | \, H_0 \text{ falsa}) \\ &= P(x \leq 3 | \, \lambda = 2) \\ &= e^{-2} \frac{2^3}{3!} + e^{-2} \frac{2^2}{2!} + e^{-2} \frac{2^1}{1!} + e^{-2} \frac{2^0}{0!} \\ &= 0.\,85712. \end{aligned}$$
¿Cómo deberíamos reformular el contraste si queremos un **nivel** de significación 0.05? Sabemos que tal nivel es α si $P(E_1) \leq \alpha$ y este número para el contraste elegido es $0.\,08030\,1$. Ahora bien si se pretende minorar más este ínidice habra que considerar x un criterio más exigente habrá que considera $x > 4$ ó $x > 5$. Si la región de rechazo es $x > 4$ entonces se obtiene
$$P(E_1) = 1 - \frac{e^{-1}}{0!} - \frac{e^{-1}}{1!} - \frac{e^{-1}}{2!} - \frac{e^{-1}}{3!} = 1.\,898\,8 \times 10^{-2}$$
y si la la región de rechazo está definida por $x > 5$
$$P(E_1) = 1 - \frac{e^{-1}}{0!} - \frac{e^{-1}}{1!} - \frac{e^{-1}}{2!} - \frac{e^{-1}}{3!} - \frac{e^{-1}}{4!} = 3.\,659\,8 \times 10^{-3}.$$
Como se ve es suficiente tomar como cotraste a
$$\varphi(x) = \begin{cases} 1 & \text{si } x > 4 \\ 0 & \text{si } x \leq 4 \end{cases}$$

para alcanzar un nivel de significación 0.05.

2. Sea una muestra de tamaño n tomada de una población $N(\mu,1)$. Se quiere contrastar $H_0 : \mu \leq \mu_0$ contra $H_1 : \mu > \mu_0$ mediante el contraste
$$\varphi(x_1,...,x_n) = \begin{cases} 1 & \text{si } \overline{x} > \mu_0 + \frac{Z_\alpha}{\sqrt[2]{n}} \\ 0 & \text{si } \overline{x} \leq \mu_0 + \frac{Z_\alpha}{\sqrt[2]{n}} \end{cases}$$
Hallar la potencia del contraste.
Tengamos en cuenta que $\Theta_0 \equiv \mu \leq \mu_0$ y $\Theta_1 \equiv \mu > \mu_0$ y que la potencia del contraste se define como
$$\beta_\varphi(\mu) = \begin{cases} P(E_1) & \text{si } \mu \in \Theta_0 \\ 1 - P(E_2) & \text{si } \mu \in \Theta_1 \end{cases}$$
Calculamos la potencia del contraste si $\theta \in \Theta_0$:
$$\begin{aligned} \beta_\varphi(\mu) &= P(E_1) = P(\text{rechazar } H_0 | \, H_0 \text{ verdadera}) \\ &= P\left(\overline{x} > \mu_0 + \frac{Z_\alpha}{\sqrt[2]{n}} \middle| \mu \leq \mu_0\right) \\ &= P\left(\frac{\overline{x}-\mu}{1/\sqrt[2]{n}} > \frac{\mu_0-\mu}{1/\sqrt[2]{n}} + Z_\alpha \middle| \mu \leq \mu_0\right) \\ &= P\left(N(0,1) > \frac{\mu_0-\mu}{1/\sqrt[2]{n}} + Z_\alpha\right) \end{aligned}$$
y como $\mu \leq \mu_0$ entonces $x_0(\mu) \doteq \frac{\mu_0-\mu}{1/\sqrt[2]{n}} + Z_\alpha$ es un número que a medida que μ crece se va haciendo cada vez más pequeño y por ello $P(N(0,1) > x_0(\mu))$ es no decreciente.
La potencia del contraste si $\theta \in \Theta_0$ es
$$\begin{aligned} \beta_\varphi(\mu) &= 1 - P(E_2) = 1 - P(\text{Aceptar } H_0 | \, H_1 \text{ verdadera}) \\ &= 1 - P\left(\overline{x} \leq \mu_0 + \frac{Z_\alpha}{\sqrt[2]{n}} \middle| \mu > \mu_0\right) \\ &= P\left(\frac{\overline{x}-\mu}{1/\sqrt[2]{n}} \geq \frac{\mu_0-\mu}{1/\sqrt[2]{n}} + Z_\alpha \middle| \mu > \mu_0\right) \\ &= P(N(0,1) \geq x_0(\mu)) \end{aligned}$$
con $x_0(\mu)$definido como antes. Al ser $\mu > \mu_0$ este $x_0(\mu)$ es negativo y al hacer crecer a μ resulta que $x_0(\mu)$ decrece y por ende $P(N(0,1) \geq x_0(\mu))$ crece.

Por lo que acabamos de analizar, la función potencia es una función decreciente como función delparámetro μ.

Si tuviéramos que evaluar el nivel de significación de este contraste tendríamos que determinar $\alpha = \sup_{\mu\in\Theta_0} \beta_\varphi(\mu)$, y esto gracias a que $\beta_\varphi(\mu)$ es no decreciente
$$\sup_{\mu\leq\mu_0} P(N(0,1) > x_0(\mu)) = P(N(0,1) > x_0(\mu_0)) = \alpha$$

3. El contraste según el enunciado ha quedado definido como

$$\varphi(x) = \begin{cases} 1 & \text{si } s = 2 \text{ ó } 3 \\ 0 & \text{si } s = 1 \end{cases}$$

siendo s el número de mármoles blancos obtenidos en la muestra.
Lo primero que calculamos el $P(E_1)$:

$$\begin{aligned} P(E_1) &= P(\text{rechazar } H_0 | H_0 \text{ verdadera}) \\ &= P(s = 2 \text{ ó } 3 | M = 5) \end{aligned}$$

Calculamos esta probabilidad por partes:

$$\begin{aligned} P(s = 2 | M = 5) &= \frac{1}{\binom{10}{3}} \binom{5}{2} \binom{5}{1} \\ &= \frac{3!7!}{(10)!} \frac{5!}{2!3!} 5 = 0.41667 \end{aligned}$$

Igualmente

$$\begin{aligned} P(s = 3 | M = 5) &= \frac{1}{\binom{10}{3}} \binom{5}{3} \binom{5}{0} \\ &= \frac{3!7!}{(10)!} \frac{5!}{2!3!} = 8.3333 \times 10^{-2} \end{aligned}$$

Por tanto

$$P(E_1) = 0{,}41667 + 8.3333 \times 10^{-2} = 0.5$$

Por otro lado

$$\begin{aligned} P(E_2) &= P(\text{Aceptar } H_0 | H_1 \text{ verdadera}) \\ &= P(s = 1 | M = 6) \\ &= \frac{1}{\binom{10}{3}} \binom{6}{1} \binom{4}{2} \\ &= \frac{3!7!}{(10)!} 6 \frac{4!}{2!2!} = 0.3 \end{aligned}$$

Con estos cálculos hechos el resto es inemdiato.

4. Utiliza el Lema de Neyman-Pearson para llevar a cabo el contrate, a nivel α, de $H_0 : \lambda = 1$ contra $H_1 : \lambda = \lambda_0 (> 1)$ basada en una muestra de tamaño 1 si la densidad de la población es

$$f_\xi(x) = \begin{cases} \lambda x^{\lambda - 1} & \text{si } 0 < x < 1 \\ 0 & \text{en el resto.} \end{cases}$$

Recordamos el Lema: sea el problema

$$(\alpha, \Theta_0 = \{\lambda = 1\}, \Theta_1 = \{\lambda = \lambda_0\}),$$

entonces:

a) Existe un contraste φ y una constante λ tales que
   1) $\beta_\varphi(\theta_0) = E_{\theta_0}[\varphi] = \alpha$
   2)
$$\varphi(x_1, ..., x_n) = \begin{cases} 1 & \text{si } f_{\theta_1}(x_1, ..., x_n) > l f_{\theta_0}(x_1, ..., x_n) \\ 0 & \text{si } f_{\theta_1}(x_1, ..., x_n) < l f_{\theta_0}(x_1, ..., x_n) \end{cases}$$
   y $\varphi(x_1, ..., x_n) = \gamma$ (una constrante) si $f_{\theta_1}(x_1, ..., x_n) = l f_{\theta_0}(x_1, ..., x_n)$

b) Todo contraste φ que satisface el punto $4a$, para cierto l, es de máxima potencia y nivel α.

c) Si φ es un contraste de máxima potencia y nivel $\alpha < 1$ entonces φ tiene que ser como en el punto $4a$.

Consideramos entonces

$$\begin{aligned} r(x_1, ..., x_n) &= \frac{f_{\theta_1}(x_1, ..., x_n)}{f_{\theta_0}(x_1, ..., x_n)} \\ &= \frac{\lambda_0^n x_1^{\lambda_0 - 1} x_2^{\lambda_0 - 1} ... x_n^{\lambda_0 - 1}}{1} \end{aligned}$$

por tanto

$$r \geq l$$

equivale a que

$$\log r \geq l'$$

para $\lambda' = \log \lambda$. Pero esto equivale a

$$n \log \lambda_0 + (\lambda_0 - 1) \sum_{i=1}^n \log x_i \geq l'$$

esto es

$$\sum_{i=1}^n (-\log x_i) \leq \frac{-l' + n \log \lambda_0}{\lambda_0 - 1}$$

Ahora bien, se puede demostrar que si ξ sigue la distribución dada por la función de densidad f_ξ dada arriba, entonces $\psi = -\log \xi$ sigue una distribución exponencial de parámetro λ, esto es, tiene como función de densidad a

$$f_\psi(s) = \begin{cases} \lambda \exp(-\lambda s) & \text{si } s > 0 \\ 0 & \text{en el resto.} \end{cases}$$

O sea, que $\psi \sim \Gamma(p = 1, a = \lambda)$, y por tanto $\sum_{i=1}^n (-\log \xi_i) \sim \Gamma(p = n, a = \lambda)$. Entonces, como la definición de la región de rechazo viene dada por $E_{\theta_0}[\varphi] = \alpha$ entonces

$$P\left(\Gamma(n, 1) \leq l^* \doteq \frac{-l' + n \log \lambda_0}{\lambda_0 - 1}\right) = \alpha.$$

Sabemos que el tamaño de la muestra es $n = 1$, lo cual significa que la identidad previa se escribe

$$P\left(\Gamma(1,1) \le l^* \doteq \frac{-l' + \log\lambda_0}{\lambda_0 - 1}\right) = \alpha.$$

determina de manera unívoca el valor de l^* y por tanto el de l' y l. Veámoslo:

$$\int_0^{l^*} \int e^{-s} ds = \alpha$$

es lo mismo que

$$1 - e^{-l^*} = \alpha$$

esto es, $l^* = -\ln(1-\alpha)$ y por ello

$$\frac{-l' + \log\lambda_0}{\lambda_0 - 1} = -\ln(1-\alpha)$$

Como consecuencia

$$l' = \ln\lambda_0 + \lambda_0 \ln(1-\alpha) - \ln(1-\alpha)$$

5. Utiliza el Lema de Neyman-Pearson para llevar a cabo el contrate, a nivel α, de $H_0 : \theta \le \theta_0$ contra $H_1 : \theta > \theta_0 (> 1)$ basada en una muestra de tamaño n si la densidad de la población es

$$f_\xi(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta} & \text{si } x > 0 \\ 0 & \text{en el resto.} \end{cases}$$

(se supone $\theta > 0$). Ahora

$$\begin{aligned} r(x_1, ..., x_n) &= \frac{f_{\theta_1}(x_1, ..., x_n)}{f_{\theta_0}(x_1, ..., x_n)} \\ &= \frac{\frac{1}{\theta_1} e^{-x_1/\theta_1} \frac{1}{\theta_1} e^{-x_2/\theta_1} ... \frac{1}{\theta_1} e^{-x_n/\theta_1}}{\frac{1}{\theta_0} e^{-x_1/\theta_0} \frac{1}{\theta_0} e^{-x_2/\theta_0} ... \frac{1}{\theta_0} e^{-x_n/\theta_0}} \\ &= \frac{\theta_0^n}{\theta_1^n} \exp\left(\left(\frac{1}{\theta_0} - \frac{1}{\theta_1}\right) \sum x_i\right) \end{aligned}$$

Como $\theta_1 > \theta_0$ entonces la imposición acerca del nivel de significación se reescribe así

$$\alpha = P(r(\xi_1, ..., \xi_n) > \lambda) = P\left(\sum_{i=1}^n \xi_i > \lambda'\right)$$

Ahora bien, se sabe que si $\xi \sim \Gamma(1, 1/\theta)$ entonces $\sum_{i=1}^n \xi_i \sim \Gamma(n, 1/\theta)$ y por tanto λ' queda determinado así como el contraste.

6. El hecho de que el tamaño de la muestra sea grande permite establecer que el contraste de esta manera: que $r \geq l$ equivale a

$$\begin{aligned} r(x_1, ..., x_n) &= \frac{f_{\theta_1}(x_1, ..., x_n)}{f_{\theta_0}(x_1, ..., x_n)} \\ &= \frac{\theta_1^n}{\theta_0^n} \exp\left((-\theta_1 + \theta_0) \sum x_i\right) \geq l \end{aligned}$$

y como $\theta_1 > \theta_0$ lo anterior equivale a

$$\sum x_i \leq l'$$

La probabilidad $\alpha = P(r(\xi_1, ..., \xi_n) \leq l)$ se reescribe

$$\begin{aligned} & P\left(\sum \xi_i \leq l'\right) \\ &= P\left(\frac{\frac{\sum \xi_i}{n} - \mu}{\sigma/\sqrt[2]{n}} \leq \frac{\frac{l'}{n} - \mu}{\sigma/\sqrt{n}} = l^*\right) \end{aligned}$$

con

$$\begin{aligned} \mu &= \theta_0 \int_0^\infty x^2 e^{-\theta_0 x} dx = \frac{1}{\theta_0} \\ \sigma^2 &= \frac{2}{\theta_0^2} - \left(\frac{1}{\theta_0}\right)^2 = \frac{1}{\theta_0^2} \end{aligned}$$

$\left(\alpha_2 = \theta_0 \int_0^\infty x^2 e^{-\theta_0 x} dx = \frac{2}{\theta_0^2}\right)$. Por ello en Θ_0 y gracias al TCL se tiene que la expresión de arriba es aproximadamente

$$P\left(N(0,1) \leq l^* = \frac{\frac{l'}{n} - \frac{1}{\theta_0}}{1/(\theta_0\sqrt{n})} = \frac{l'\theta_0 - n}{\sqrt{n}}\right)$$

lo que implica $l^* = Z_\alpha$. Ahora bien, lo que se persigue es

$$\alpha = \sup_{\theta \in \Theta_0} \beta_\varphi(\theta) = \sup_{\theta \in \Theta_0} P\left(N(0,1) \leq \frac{l'\theta - n}{\sqrt{n}}\right)$$

y como $P\left(N(0,1) \leq \frac{l'\theta_0 - n}{\sqrt{n}}\right)$ es decreciente en θ, el máximo se alcanzará en aquél valor más pequeño de θ; como estamos en $\Theta_0 = \lambda \leq 0.008$, entonces tomaremos $\theta = 0.008$ y el $\alpha-$valor es

$$\frac{l'\theta_0 - n}{\sqrt{n}} = \frac{l'(0.008) - n}{\sqrt{n}}.$$

De manera que como $\alpha = P\left(N(0,1) \geq \frac{l'(0.008) - n}{\sqrt{n}}\right)$ entonces $\frac{l'(0.008) - n}{\sqrt{n}} = Z_\alpha$,, de donde se sigue

$$l' = n\left(125 + 125\frac{Z_\alpha}{\sqrt{n}}\right)$$

De resultas, contraste es

$$\varphi(x_1,...,x_n)=\begin{cases}1 & \text{si } \frac{\sum_{i=1}^n x_i}{n}>125+125\frac{Z_\alpha}{\sqrt{n}}\\ 0 & \text{si } \frac{\sum_{i=1}^n x_i}{n}<125+125\frac{Z_\alpha}{\sqrt{n}}\end{cases}$$

Como $\alpha=0.05$ y $n=1000$, $125+125\frac{2.575}{10\sqrt{10}}=135.18$ y quedará

$$\varphi(x_1,...,x_n)=\begin{cases}1 & \text{si } \frac{\sum_{i=1}^n x_i}{n}>135.18\\ 0 & \text{si } \frac{\sum_{i=1}^n x_i}{n}<135.18\end{cases}$$

pues $Z_{0.05}=2.575$. En base a la realización de la muestra $\varphi(x_1,...,x_n)=0$ y por ende se acepta H_0.

7. Se trata de una población normal en la que el parámetro a contrastar es la esperanza y la desviación típica se supone conocida. La hipótesis nula es $H_0:\mu=5$. La realización de la muestra es

$$x=[5.2,4.9,5,5.1,5.2,4.8,4.9,5.3,4.6,5.4]$$

y la de la media muestral es

$$\overline{x}=\frac{\sum_{i=1}^{10}x_i}{10}=5.04.$$

El contraste para H_0 está definido por

$$|\overline{x}-\mu_0|\geq\frac{\sigma}{\sqrt{n}}z_{\alpha/2}$$

esto es

$$\varphi(x_1,...,x_n)=\begin{cases}1 & \text{si } \overline{x}\geq 5+\frac{0.10}{\sqrt{10}}z_{\alpha/2} \text{ ó } \overline{x}\leq 5-\frac{0.10}{\sqrt{10}}z_{\alpha/2}\\ 0 & \text{en el resto}\end{cases}$$

donde $Z_{\alpha/2}=Z_{0.025}=2.81$. Como

$$\begin{aligned}5+\frac{0.10}{\sqrt{10}}2.81 &= 5.0889,\\ 5-\frac{0.10}{\sqrt{10}}2.81 &= 4.9111\end{aligned}$$

y $\overline{x}=5.04$ entonces aceptamos la hipótesis nula.

8. La v.a. ξ del modelo es la variación del peso y $\xi\sim N(\mu,\sigma)$. La hipótesis nula es $H_0:\mu=0$. El contraste está dado mediante

$$|\overline{x}|\geq\frac{s}{\sqrt{n}}t_{n-1,\alpha/2}$$

Puesto que $n=12$, $\alpha=0.05$, $\overline{x}=\frac{\sum_{i=1}^{12}x_i}{12}=0.075$, y

$$s^2=\frac{\sum_{i=1}^{12}(x_i-0.075)^2}{11}=0.45114.$$

Además $\mathrm{t}_{11,0.025} = 2.201$. Entonces

$$\frac{s}{\sqrt{n}}\mathrm{t}_{n-1,\alpha/2} = \frac{\sqrt[2]{0.45114}}{\sqrt[2]{12}}2.201 = 0{,}42676$$

Como $0.075 \ngeq 0.42676$ entonces aceptamos la hipótesis nula y valoramos como falsa la indicación dada en la publicidad.

9. Se tiene que $\xi \sim N(\mu_0, \sigma)$, siendo ξ la cantidad de protombina con $\mu_0 = 20$. Se lleva a cabo un muestreo para contrastar $H_0 : \mu = \mu_0$. En caso de que rechazaramos lo que deduciríamos es que la población que se ha tomado en la muestra no es representativa para la población total. De nuevo el contrate está dictado por

$$|\overline{x} - \mu_0| \geq \frac{s}{\sqrt{n}}\mathrm{t}_{n-1,\alpha/2},$$

esto es

$$|\overline{x} - 20| \geq \frac{4}{\sqrt{40}}\mathrm{t}_{39,0.025} = \frac{4}{\sqrt{40}}2.02 = 1.2776,$$

y $\overline{x} = 18.5$. Por tanto

$$|18.5 - 20| = 1.5 \geq 1.2776,$$

y rechazamos la hipótesis nula, lo cual, según el planteamiento del problema quiere decir que la población compuesta por los individuos con deficiencia en vitamina B no es representativa de toda la población a efectos del estudio de la protombina.

10. La hipótesis nula es $H_0 : \mu_1 - \mu_2 = 0$ y el contraste adecuado para ésta viene dado por la región de rechazo

$$|\overline{x} - \overline{y}| \geq z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}$$

para el caso en el que la varianzas son conocidas y por

$$|\overline{x} - \overline{y}| \geq \mathrm{t}_{m+n-2,\frac{\alpha}{2}}s_p\sqrt{\frac{1}{m} + \frac{1}{n}}$$

siendo $s_p = \frac{(m-1)s_1^2+(n-1)s_2^2}{m+n-2}$, cuando las varianzas sean iguales pero desconocidas. Se trata del segundo contraste con los datos $m = n = 20$, $s_1 = 0.2$ y $s_2 = 0.3$. De esto se sigue que habremos de comporbar si

$$|0.4 - 0.5| \geq (2.02)\sqrt[2]{0.065}\sqrt{\frac{1}{20} + \frac{1}{20}}$$

pues $s_p^2 = \frac{19(0.2)^2+19(0.3)^2}{38} = 0.065$. Como $(2.02)\sqrt[2]{0.065}\sqrt{\frac{1}{20} + \frac{1}{20}} = 0.16286$ no se produce rechazo sobre la igualdad de medias, se acepta H_0.

11. Se supone que el modelo es normal. A partir de los datos de la muestra calculamos s_1 y s_2. Si

$$\begin{aligned} x &= (12{,}5, 14, 13, 12{,}5, 15, 14{,}5, 13, 13{,}5, 16), \\ y &= (9{,}5, 13, 13, 10{,}5, 11, 14, 13{,}5, 15, 14{,}5) \end{aligned}$$

entonces

$$\overline{x} = \frac{1}{9}\sum_{i=1}^{9} x_i = 13.778,$$
$$\overline{y} = \frac{1}{9}\sum_{i=1}^{9} y_i = 12.667,$$
$$s_1^2 = \frac{1}{8}\sum_{i=1}^{9} (x_i - 13.778)^2 = 1.4444,$$

y

$$s_2^2 = \frac{1}{8}\sum_{i=1}^{9} (y_i - 12.667)^2 = 3.625$$

a) Nos están preguntando si las varianzas son iguales. Para contestar usamos el contraste dado mediante

$$\begin{cases} \frac{s_1^2}{s_2^2} \geq F_{m-1,n-1,\frac{\alpha}{2}} \text{ si } s_1^2 \geq s_2^2 \\ \text{ó} \\ \frac{s_2^2}{s_1^2} \geq F_{m-1,n-1,\frac{\alpha}{2}} \text{ si } s_1^2 \leq s_2^2 \end{cases}$$

Empleando los datos vemos que $s_1^2 \leq s_2^2$ y que

$$\frac{3.625}{1.4444} = 2.5097 \ngeq F_{8,8,0.025} = 3.44$$

y por tanto se acepta la hipótesis de que las varianzas son iguales.

b) Para ver si hay diferencia significativa entre las medias consideramos la región crítica

$$|\overline{x} - \overline{y}| \geq t_{m+n-2,\frac{\alpha}{2}} s_p \sqrt{\frac{1}{m} + \frac{1}{n}}$$

Como

$$1.111 = |13.778 - 12.667| \ngeq (2.12)\sqrt[2]{(2.5347)}\sqrt{\frac{1}{9} + \frac{1}{9}} = 1.5911$$

pues $s_p^2 = \frac{8(1.4444)+8(3.625)}{16} = 2.5347$ y $t_{m+n-2,\frac{\alpha}{2}} = 2.12$, entonces aceptamos la hipótesis de igualdad de esperanzas.

12. Aunque en esta situación no tenemos una ley normal disponemos de un tamaño de muestra suficientemente grande. El planteamiento del problema nos conduce a emitir como hipótesis nula $H_0 : p_1 \leq p_2$, donde las p_i son las proporciones. Para esta situación disponemos de un contraste cuya región crítica de rechazo es

H_0	Región de rechazo
$p_1 \leq p_2$	$\overline{x} - \overline{y} > Z_\alpha \sqrt{p(1-p)\left(\frac{1}{m} + \frac{1}{n}\right)}$

donde
$$p = \frac{m\overline{x} + n\overline{y}}{m+n}.$$
Empleamos los datos: $n = m = 100$, $\overline{x} = \frac{8}{100}$, $\overline{y} = \frac{8}{100}$ y $Z_{0.05}$ =1.64. Así, tenemos que
$$\frac{8}{100} - \frac{20}{100} \not> (1.64) \sqrt[2]{\frac{28}{200}\left(1 - \frac{28}{200}\right)\left(\frac{2}{100}\right)}$$
luego aceptamos H_0.

13. Después de llevar a cabo un contraste análogo al del ejercicio anterior se llega a la conclusión de que el número de aves jóvenes muertas es inferior que el de aves adultas. Se omiten los detalles.
14. Los valores observados para la tensión de ruptura de una muestra de tamaño 14 de una clase de fibra sintética son

    9.9 5 5.2 7. 11.8 10.3 8.2 7.5 6.6 12.6 16.8 12.3 9.8 10.3

    a) Determinar los intervalos de confianza para la tensión media de ruptura, así como para la varianza poblacional de la variable tensión.
    b) De acuerdo con los datos, ¿es aceptable la afirmación según la cual dicha fibra soporta por término medio una tensión de ruptura al menos igual a 12?