Colección de Problemas

1. Muestreo y distribuciones muestrales

1. Sea X una variable aleatoria con distribución de Poisson de parámetro λ. Consideremos una muestra aleatoria simple de tamaño n.

   a) Encuentre la distribución conjunta de las variables aleatorias muestrales.

   b) ¿Cuáles de las siguientes funciones son estadísticos?

$$\begin{aligned} T_1 &= \min\{X_1,\ldots,X_n\} \\ T_2 &= X_1 + X_2 + X_3 \\ T_3 &= \textstyle\sum_{i=1}^n X_i \\ T_4 &= X_1 \times X_2 \times \ldots \times X_n - \lambda \\ T_5 &= \textstyle\prod_{i=1}^n X_i \end{aligned}$$

2. Sea X una variable aleatoria con distribución exponencial de parámetro α. Consideremos una muestra aleatoria simple de tamaño n.

   a) Encuentre la distribución conjunta de las variables aleatorias muestrales.

   b) ¿Cuáles de las siguientes funciones son estadísticos?

$$\begin{aligned} T_1 &= \max\{X_1,\ldots,X_n\} \\ T_2 &= 5 \\ T_3 &= \textstyle\prod_{i=1}^n X_i - \alpha \\ T_4 &= \alpha \\ T_5 &= e^{\sum_{i=1}^n X_i} \end{aligned}$$

3. Sea X una variable aleatoria con distribución normal $N(\mu,\sigma)$. Consideremos una muestra aleatoria simple de tamaño n.

   a) Encuentre la distribución conjunta de las variables aleatorias muestrales.

   b) Encuentre la distribución de $\overline{X}_n$ (media muestral).

   c) Encuentre la distribución de S_n^2 (varianza muestral).

   d) ¿Cuál es la distribución conjunta de $(\overline{X}_n, S_n^2)$.

   e) Halle $\overline{X}_n$ y S_n^2 para los siguientes valores muestrales correspondientes a una muestra de tamaño $n = 10$

   1,74 0,45 2,52 1,19 1,24 2,68 3,51 1,83 1,00 0,87

4. Hallar la distribución exacta del estadístico $\sum_{i=1}^n X_i$ correspondiente a una muestra aleatoria simple de tamaño n de una variable aleatoria X con distribución:

a) Bernouilli de parámetro p, $Be(p)$.
b) Exponencial de parámetro α, $Exp(\alpha)$.
c) Gamma de parámetros α y k, $G(\alpha, k)$.
d) Poisson de parámetro λ, $P(\lambda)$.

5. Sea X una variable aleatoria con distribución normal $N(0,1)$. Consideremos una muestra aleatoria simple de tamaño n. Halle la distribución del estadístico

$$U_n = X_1^2 + X_2^2 + \ldots + X_n^2$$

6. Sean Z y U dos estadísticos tales que $Z \sim N(0,1)$, $U \sim \chi_n^2$ y además son independientes en sentido estocástico. Halle la función de densidad del estadístico $T = \dfrac{Z}{\sqrt{U/n}}$. Dicha distribución se conoce como distribución t de Student con n grados de libertad.

7. Sean U y V dos estadísticos tales que $U \sim \chi_m^2$, $V \sim \chi_n^2$ y además son independientes en sentido estocástico. Halle la función de densidad del estadístico $Y = \dfrac{U/m}{V/n}$. Dicha distribución se conoce como distribución F de Fisher con m y n grados de libertad, F_n^m. ¿Cuál es la distribución de $1/Y$?

8. Sea X una variable aleatoria a valores reales con función de distribución de probabilidad igual a F_X y sean $X_1, \ldots, X_n$ las variables aleatorias muestrales correspondientes a una muestra aleatoria simple de tamaño n de X. Halle la función de distribución de los estadísticos $X_{(1)}$ y $X_{(n)}$.

9. Sea F la función de distribución de una variable aleatoria absolutamente continua y que posee función inversa, F^{-1}. Demuestre que si $X \sim U(0,1)$ entonces la función de distribución de la variable aleatoria $Y = F^{-1}(X)$ es precisamente F. ¿Para qué puede servir esta propiedad?

10. Utilice el resultado del problema anterior para generar con una hoja de cálculo (Excel, Calc, ...) números pseudoaleatorios con distribución exponencial de esperanza α y normal de esperanza μ y desviación típica σ.

11. Supongamos que el tiempo en que un fármaco es activo sigue una distribución normal $N(1200, 40)$ y hemos de preparar un lote con n muestras de dicho fármaco. Calcule n de forma que la media muestral del tiempo en que el fármaco es activo sea superior a 1180 con probabilidad igual a 0,95.

12. Supongamos que la duración de un determinado tipo de bombillas de bajo consumo es una variable aleatoria cuya distribución es normal con esperanza $\mu = 50$ meses y desviación estándar $\sigma = 2$ meses.

   a) Calcule la probabilidad de que la media muestral de un lote de 25 bombillas sea superior a 49 meses.

b) ¿De qué tamaño muestral debería ser el lote para garantizar que la diferencia entre la media muestral y la poblacional (esperanza) fuera inferior a 0,5 meses con probabilidad del 95 %?

13. Supongamos que el número de clientes atendidos cada 15 minutos en una caja cualquiera de un supermercado es una variable aleatoria con distribución de Poisson de parámetro $\lambda = 6$.

    a) ¿Cuál es la probabilidad de que en muestras de 12 cajas en promedio se atiendan más de 4 clientes cada 15 minutos?

    b) ¿Cuál es la probabilidad de que en una hora una caja atienda a más de 10 clientes?

14. Supongamos que la duración en horas de una batería eléctrica es una variable aleatoria con distribución exponencial de parámetro $\alpha = 0{,}6$. Consideremos lotes de 10 baterías ¿cuál es la probabilidad de que la suma de la duración de todas las baterías sea superior a 25 horas?

15. De un determinado producto manufacturado la proporción de piezas defectuosas que se averían durante el tiempo de garantía es igual al 1 %. En muestras de 50 piezas, ¿cuál es la probabilidad de que la proporción muestral de piezas defectuosas sea superior al 2,5 %?

16. Sea X una variable aleatoria discreta cuya función de densidad de probabilidad viene dada por:

$$f(x) = \begin{cases} 0{,}7 & \text{si } x = 1 \\ 0{,}1 & \text{si } x = 2 \\ 0{,}2 & \text{si } x = 3 \\ 0 & \text{en caso contrario} \end{cases}$$

    En muestras de tamaño 25, calcule de forma aproximada la probabilidad de que la media muestral tome valores superiores a 1.

17. Supongamos que tomamos nuestras de tamaño $n_1 = 5$ de una población estadística con distribución normal de esperanza $\mu_1 = 50$ y varianza $\sigma_1^2 = 9$ y una segunda muestra de tamaño $n_2 = 4$ de una población estadística también normal pero de esperanza $\mu_2 = 40$ y varianza $\sigma_2^2 = 4$, independiente de la primera. Calcule la probabilidad de que las medias muestrales difieran en menos de 8,2.

18. Sea X una variable aleatoria con distribución normal de parámetros $\mu = 0{,}5$ y $\sigma = 0{,}1$. Tomamos una muestra aleatoria simple correspondiente a dicha variable de la que observamos 21 valores. Calcule la probabilidad de que la varianza muestral de la presente muestra sea mayor que 0,02.

19. Supongamos que el intervalo de tiempo, X, entre dos llamadas telefónicas que llegan a una centralita sigue una distribución exponencial de esperanza 20 segundos. Se toma una muestra aleatoria simple de 100 intervalos de dichos tiempos entre llamadas llegadas a la citada centralita. Calcule:

a) ¿Cuál es la probabilidad de que la media muestral $\bar{X}_n$ esté comprendida entre 15 y 25 segundos?

b) ¿Cuál es la probabilidad de que el intervalo más grande entre dos llamadas consecutivas sea superior a un minuto?

c) ¿Cuál es la probabilidad de que el intervalo más corto entre dos llamadas consecutivas no sea superior a un segundo?

20. Supongamos que disponemos de una muestra aleatoria simple de tamaño $n = 40$ de una distribución exponencial de esperanza 3. ¿Cuál es la probabilidad de que los valores de la función de distribución muestral y la poblacional en el punto $x = 1$ difieran en menos de 0,01? ¿Cuál debería ser el tamaño muestral para que dicha probabilidad fuera aproximadamente igual a 0,98?

21. Supongamos que disponemos de una muestra aleatoria simple de tamaño $n = 50$ de una distribución de Poisson de esperanza 2. ¿Cuál es la probabilidad de que los valores de la función de distribución muestral y la poblacional en el punto $x = 2$ difieran en menos de 0,03? ¿Cuál debería ser el tamaño muestral para que dicha probabilidad fuera aproximadamente igual a 0,99?

Algunas indicaciones y soluciones de los ejercicios propuestos

1. Para el apartado *a)* hay que tener en cuenta que $X_1, \ldots, X_n$ son variables aleatorias independientes e idénticamente distribuidas. En cuanto al apartado *b)*, son estadísticos T_1, T_2, T_3 y T_5.

2. Para el apartado *a)* hay que tener en cuenta que $X_1, \ldots, X_n$ son variables aleatorias independientes e idénticamente distribuidas. En cuanto al apartado *b)*, son estadísticos T_1, T_2 y T_5.

3. Para este ejercicio, recordemos el resultado conocido como Teorema de Fisher que establece que si $X_1, \ldots, X_n$ son las variables aleatorias muestrales de una muestra aleatoria simple de tamaño n correspondiente a una distribución normal $N(\mu, \sigma)$ entonces la media muestral, $\overline{X}_n$, sigue una distribución $N(\mu, \sigma/\sqrt{n})$ y la varianza muestral multiplicada por el tamaño muestral y dividida por la varianza poblacional, nS_n^2/σ^2, sigue una distribución ji-cuadrado con $n-1$ grados de libertad, y ambas variables aleatorias son independientes en sentido estocástico. En cuanto al apartado *e)* los resultados son $\overline{X}_n = 1,70300$, $S_{10}^2 = 0,81404$.

4. *a)* $\sum_{i=1}^{n} X_i$ seguirá una distribución Binomial, $B(n, p)$.

   b) $\sum_{i=1}^{n} X_i$ seguirá una distribución Gamma, $G(\alpha, n)$.

   c) $\sum_{i=1}^{n} X_i$ seguirá una distribución Gamma, $G(\alpha, nk)$.

   d) $\sum_{i=1}^{n} X_i$ seguirá una distribución de Poisson, $P(n\lambda)$.

5. U_n seguirá una distribución ji-cuadrado con n grados de libertad.

6. A partir de la función de densidad conjunta del par (Z,U) se estudia la densidad conjunta del par (T,W) con W, por ejemplo, definida como $W = U$. Seguidamente se halla la densidad marginal de T.

7. A partir de la función de densidad conjunta del par (U,V) se estudia la densidad conjunta del par (Y,W) con W, por ejemplo, definida como $W = V$. Seguidamente se halla la densidad marginal de Y. La distribución de Y también será una F de Fisher pero con n y m (grados de libertad cambiados): $1/Y \sim F_m^n$.

8. Si designamos por F_X la función de distribución de X, entonces la función de distribución del máximo es $F_{X_n}(u) = F_X(u)^n$, mientras que la función de distribución del mínimo es $F_{X_1}(u) = 1-(1-F_X(u))^n$.

9. Basta probar que la función de distribución de Y, verifica $F_Y(y) = F(y)$.

11. $n = 11$ es el menor entero que garantiza que la probabilidad de que la media muestral sea superior a 1180 es mayor que 0,95.

12. Obtenemos el valor 0,99977.

13. *a)* Mediante una aproximación normal obtendremos 0,99766, con un cálculo más exacto, basado en la distribución de Poisson correspondiente, obtendremos 0,99827.

    b) Mediante una aproximación normal obtendremos 0,99787, con un cálculo más exacto, basado en la distribución de Poisson correspondiente, obtendremos 0,99892.

14. *a)* Mediante una aproximación normal obtendremos 0,05692 y con un cálculo más exacto basado en la distribución Gamma $G(\alpha,n)$, con $\alpha = 0,6$ y $n = 10$, obtendremos 0,06985.

15. Con un cálculo exacto basado en una binomial obtenemos el valor 0,08944. En este caso la aproximación normal no es buena (obtendríamos el valor 0,14131), mientras que la aproximación de una binomial por una Poisson, de parámetro $\lambda = np$ sería razonable y nos proporcionaría el valor 0,09020.

16. $E(X) = 1,5$, $std(X) = 0,80623$. Si utilizamos una aproximación normal, basada en el Teorema del Límite Central, obtendremos que la probabilidad buscada es 0,99904.

17. Observar que la varianza de la diferencia de variables aleatorias independientes es igual a la suma de varianzas. El resultado es igual a 0,14103.

18. Por el Teorema de Fisher sabemos que nS_n^2/σ^2 sigue una distribución ji-cuadrado con $n-1$ grados de libertad. A partir de este resultado, a partir de la función de distribución ji-cuadrado con 20 grados de libertad, obtenemos 0,00277.

19. *a)* Si efectuamos una aproximación normal obtenemos 0,98758. Un cálculo basado en la distribución Gamma proporciona el valor 0,98727.

b) Un cálculo basado en la distribución del máximo de una exponencial proporciona el valor 0,99395.

c) Un cálculo basado en la distribución del mínimo de una exponencial proporciona el valor 0,99326.

20. Sea F_n la función de distribución empírica basada en una muestra aleatoria simple de tamaño n, i.e.:

$$F_n(x) = \sum_{i=1}^{n} \frac{1}{n} \mathbb{1}_{(-\infty,x]}(X_i)$$

La función de distribución poblacional es:

$$F(x) = (1 - e^{-x/3}) \mathbb{1}_{(0,\infty)}(x)$$

Nos piden:

$$P(|F_n(1) - F(1)| < 0,01)$$

observemos que $nF_n(1)$ sigue una distribución binomial $B(n, F(1))$. En el presente caso, $n = 40$ y $F(1) = 0,28347$, por tanto, si efectuamos una aproximación normal, obtenemos que la probabilidad buscada es igual a 0,11162, mientras que si efectuamos un cálculo más exacto basado en la distribución binomial, la probabilidad buscada es igual a 0,13910 (igual a la probabilidad de que en una muestra aleatoria simple de tamaño 40, de una población exponencial de esperanza 3, haya 11 resultados menores o iguales que 1).

21. Sea $F_n(x)$ la función de distribución empírica en el punto x. La función de distribución poblacional es:

$$F(x) = \sum_{k=0}^{\infty} e^{-2} \frac{2^k}{k!} \mathbb{1}_{(-\infty,x]}(k)$$

Nos piden hallar n para que:

$$P(|F_n(2) - F(2)| < 0,03) \simeq 0,99$$

observemos que $nF_n(2)$ sigue una distribución binomial $B(n, F(2))$. En el presente caso $F(2) = 0,28347$, por tanto, si efectuamos una aproximación normal, tendremos que

$$\sqrt{n} \, \frac{0,03}{0,46775} = \sqrt{n} \, 0,06414 = 2,57583$$

y por tanto $n \simeq 1613$.