Tema 4

Muestreo y Estimación

4.1 Sea $X_1, \ldots, X_n$ una m.a.s. procedente de una población que se distribuye de acuerdo a una ley de Poisson de parámetro λ. ¿Es $T(X_1, \ldots, X_n) = \frac{1}{n^2} \sum_{i=1}^{n} X_i^2$ un estimador insesgado de λ^2? En caso negativo calcular el sesgo.

4.2 Sea $X_1, \ldots, X_n$ una m.a.s. procedente de una distribución continua con función de densidad:

$$f(x;\theta) = \begin{cases} 1 - \theta + \frac{\theta}{2\sqrt{x}} & \text{si} \quad x \in (0,1) \\ 0 & \text{en otro caso} \end{cases}$$

donde $\theta \in [0,1]$. Demostrar que $T(X_1, \ldots, X_n) = \sum_{i=1}^{n} \frac{X_i}{n}$ no es un estimador insesgado de θ y encontrar el sesgo.

4.3 En muestras de tamaño $n = 3$ de una v.a. normal de media μ y varianza conocida $\sigma^2 = 1$, se consideran los estimadores:

$$U_1 = \frac{X_1 + X_2 + X_3}{3} \qquad U_2 = \frac{X_1 + 2X_2 + X_3}{4} \qquad U_3 = \frac{X_1 + 3X_2 + 4X_3}{8}$$

Comprobar que son estimadores insesgados de μ y comparar su eficiencia. Calcular la estimación de μ para la muestra: $x_1 = 2.6,\ \ x_2 = 3.1,\ \ x_3 = 1.8$.

4.4 Sea X una variable aleatoria uniformemente distribuida en $(0, \beta)$, donde β es desconocido, y sea $X_1, \ldots, X_n$ una m.a.s. de X. Consideremos el estimador $U = max\{X_1, \ldots, X_n\}$ cuya función de densidad viene dada por

$$f_U(x) = \begin{cases} \frac{nx^{n-1}}{\beta^n} & \text{si} \quad x \in (0,\beta) \\ 0 & \text{si} \quad x \notin (0,\beta) \end{cases}$$

(a) Comprobar que es un estimador sesgado de β.

(b) Estimar β, utilizando un estimador insesgado, a partir de la muestra:

0.53 0.73 1.54 2.48 1.3 0.2

(c) Demostrar que el estimador U es consistente para β.

4.5 Un ocular micrométrico está adaptado a un microscopio para medir partes de una célula. Los errores de medida cometidos siguen una distribución $N(0, \sigma^2)$. Se pide:

(a) Probar, en muestras de tamaño 2, que los siguientes estimadores son insesgados para σ^2 y comparar su eficiencia:

$$U_1 = \frac{X_1^2 + 2X_2^2}{3} \quad U_2 = \frac{2X_1^2 + 3X_2^2}{5}$$

(b) Se han comprobado los siguientes errores de medida (en micras): 2.1, -1.8. Hallar la estimación más eficiente de σ^2.

4.6 Sea X_1, X_2, X_3, X_4 una muestra aleatoria de tamaño cuatro de una población cuya distribución es exponencial de parámetro θ desconocido. De los siguientes estadísticos, ¿cuáles son insesgados de θ^{-1}?

$$T_1 = \tfrac{1}{6}(X_1 + X_2) + \tfrac{1}{3}(X_3 + X_4) \quad T_2 = \tfrac{X_1+2X_2+3X_3+4X_4}{5} \quad T_3 = \tfrac{X_1+X_2+X_3+X_4}{4}$$

¿Cuál tiene varianza mínima?

4.7 Sea $X_1, \ldots, X_n$ una m.a.s. procedente de una población que se distribuye según una ley Geométrica de parámetro p y sea $T = 1 - \frac{1}{\overline{X}}$. ¿Es T un estimador suficiente para $q = 1 - p$?

4.8 Sea $X_1, \ldots, X_n$ una muestra aleatoria simple de una variable aleatoria X con función de densidad:

$$f(x) = \frac{x^3}{6\theta^4}\, e^{-\frac{x}{\theta}} \quad \text{si} \quad x > 0$$

donde $\theta > 0$ es un parámetro desconocido. Se sabe que $E[X] = 4\theta$, $var(X) = 4\theta^2$.

Sean $T_1 = \frac{\overline{X}}{4}$ y $T_2 = \frac{\sum_{i=1}^n X_i^2}{n}$:

(a) Demostrar que T_1 es un estimador suficiente para θ.
(b) Demostrar que T_1 es un estimador consistente para θ.
(c) Demostrar que T_2 no es un estimador insesgado de θ^2 y determinar el sesgo. Obtener a partir de T_2 un estimador insegado de θ^2.

4.9 Se ha llevado a cabo un análisis de nicotina en una marca de cigarrillos y se ha encontrado en una muestra de 100 cigarrillos una cantidad media $\overline{X}$=26 mg. Si se sabe que la desviación típica de la marca es 8 mg., calcular un intervalo de confianza del 99% para el contenido medio en nicotina de la marca. (Suponer normalidad)

4.10 Determinar un intervalo a un nivel de confianza del 95% para la media de una distribución normal con varianza 9, usando una muestra de tamaño $n = 100$ y media $\overline{X}$=5.

4.11 Una muestra de 200 personas tiene una estatura media de 1.72 m. y varianza 0.52 m. Suponiendo que proceden de una población normal, obtener un intervalo de confianza para la media al 98%.

4.12 Se sabe que la longitud media de una determinada especie de lagartija es de 12 cm. Para estimar la varianza de la longitud de la especie se toma una muestra de 30 lagartijas para las cuales se obtuvo una varianza de 3.4. Obtener un intervalo de confianza para la varianza poblacional al 96%. (Suponer normalidad)

4.13 Los resultados que se muestran a continuación son relativos a las temperaturas de funcionamiento de dos procesadores:

	tamaño muestral	media	varianza
procesador 1	10	44.67	2.87
procesador 2	9	47.24	8.94

Supuesto que las características bajo estudio siguen distribuciones normales independientes con distintas varianzas, calcular un intervalo de confianza al 90% para la diferencia de temperaturas medias de ambos procesadores.

4.14 Se quiere probar la efectividad de un antitérmico (reductor de la fiebre). Con tal fin se tomó la temperatura a 10 niños afectados de gripe, antes y después de la administración del antitérmico resultando una media de las diferencias de temperaturas (antes-después) de $\overline{D} = 1.58$ y una cuasidesviación típica de las diferencias $S_{c_d} = 0.71$. Suponiendo que la característica bajo estudio sigue una distribución normal, calcular un intervalo de confianza al 90% para la diferencia de medias.

4.15 En un estudio antropométrico sobre la longitud, en centímetros, del fémur se han tomado dos muestras de tamaño 20 y 30 respectivamente, y se han realizado mediciones de dicha longitud en los individuos que formaban parte de ambas muestras. Supuesto que la característica bajo estudio sigue distribuciones normales en ambas poblaciones calcular un intervalo de confianza, al 95%, para el cociente de varianzas poblacionales, sabiendo que las varianzas de las muestras son $S_X^2 = 57$ y $S_Y^2 = 45$.

4.16 Se desea estudiar el tiempo que tarda en hacer efecto dos tratamientos A y B. Una muestra compuesta por diez pacientes sobre la que se aplicó el tratamiento A dio una media de 12.3 minutos y una varianza de 6.5, mientras que los resultados para otra muestra de tamaño 10 con el tratamiento B fueron 10.1 minutos de media y varianza 4.1. Supuesto que el tiempo que tarda en hacer efecto el tratamiento sigue una distribución normal, hallar un intervalo de confianza al 90% para σ_A/σ_B.

4.17 El Ayuntamiento de una ciudad se encuentra interesado en estimar la cantidad promedio de dinero que la población se gasta, al cabo del mes, en espectáculos. Se realizó una encuesta a 16 personas elegidas al azar obteniéndose (en euros) los siguientes datos: 15.00, 17.50, 16.30, 14.80, 14.20, 18.90, 13.50, 17.40, 16.80, 15.20, 15.80, 18.40, 13.40, 14.60, 15.50, 16.30. Si se supone que la cantidad de dinero gastada en un mes es una variable aleatoria que sigue una distribución normal, obtener los intervalos de confianza para la cantidad promedio real al 90%, 95% y 98%. Obtener un intervalo de confianza para la varianza con nivel de confianza del 98%.

4.18 En una muestra aleatoria simple de 50 programadores se ha observado una media de la diferencia entre el sueldo más alto y el más bajo por hora pagado a un mismo programador de 90 euros. Si se supone normalidad, siendo la varianza de la diferencia $\sigma_D^2 = 0.1$ euros^2, calcular un intervalo de confianza para la diferencia de sueldos al 95%.

4.19 Dos universidades tienen métodos diferentes para matricular a sus alumnos. En cada universidad se anotaron los tiempos de inscripción para 40 alumnos seleccionados al azar. Las medias y desviaciones típicas observadas fueron las siguientes: $\overline{X_1} = 50.2$, $\overline{X_2} = 52.9$, $S_1 = 4.8$ y $S_2 = 5.4$. Si se supone que el muestreo se llevó a cabo sobre dos poblaciones normales, obtener un intervalo de confianza al 90% para la diferencia de medias supuesto que las varianzas de los tiempos de matriculación son iguales en ambas universidades.

4.20 Se utilizan dos máquinas para llenar botellas de plástico con lavavajillas. Se sabe que las desviaciones típicas del volumen de llenado son $\sigma_1 = 11$ ml y $\sigma_2 = 15$ ml, respectivamente. Se toman dos muestras aleatorias, 30 botellas de la máquina 1 y 45 botellas de la máquina 2. Los volúmenes promedio de llenado son $\overline{X_1} = 900$ ml y $\overline{X_2} = 830$ ml. Construye un intervalo de confianza al 90% para la diferencia entre las medias del volumen de llenado.