Tarea 3
Inferencia Estadística

Armenta Trejo Luis Felipe
Cruz Martínez Ricardo
Robles Leong Daniel
Sánchez Galván Ximena

9 de junio de 2021

1. Considerando la siguiente densidad conjunta, calcula lo que se pide

$$f(x,y) = \frac{1}{8y}e^{-(\frac{x}{2y}+\frac{y}{4})} \quad x,y>0$$

Encuentra la densidad marginal $f_Y(y)$ y la densidad condicional $f_{X|Y}(x|y)$, finalmente encuentre $E(Y)$ y $Var(Y)$, así como $E(X|Y)$ y $Var(X|Y)$

Solución. Procederemos a calcular la densidad marginal de Y, pero antes, calcularemos la integral indefinida

$$\int \frac{1}{8y}e^{-(\frac{x}{2y}+\frac{y}{4})}dx = -\frac{1}{4}\int e^u du = -\frac{1}{4}e^u = -\frac{1}{4}e^{\frac{-2x-y^2}{4y}}$$

Donde la primera igualdad se da haciendo el cambio de variable $u = \frac{-2x-y^2}{4y}$.
Por lo que finalmente, la densidad marginal de Y está dada por

$$f_Y(y) = \int_0^\infty \frac{1}{8y}e^{-(\frac{x}{2y}+\frac{y}{4})}dx = -\frac{1}{4}e^{\frac{-2x-y^2}{4y}}|_{x=0}^{\infty} = \frac{1}{4}e^{-\frac{y}{4}}$$

Por otra parte, la densidad condicional $f_{X|Y}(x|y)$ está dada por

$$\frac{f(x,y)}{f_Y(y)} = \frac{\frac{1}{8y}e^{-(\frac{x}{2y}+\frac{y}{4})}}{\frac{1}{4}e^{-\frac{y}{4}}} = \frac{\frac{1}{8y}e^{-\frac{x}{2y}}}{\frac{1}{4}} = \frac{e^{-\frac{x}{2y}}}{2y} \quad x,y>0$$

Ahora, $E(Y)$ está dada por

$$\int_0^\infty y\frac{1}{4}e^{-\frac{y}{4}}dy = ye^{-\frac{y}{4}}|_0^\infty + \int_0^\infty e^{-\frac{y}{4}}dy = -4e^{-\frac{y}{4}}|_0^\infty = 4$$

Donde la primera igualdad de da por la fórmula de integración por partes, tomado $f(y) = y$ y $g'(y) = \frac{1}{4}e^{-\frac{y}{4}}$, por lo tanto, $E(Y) = 4$.
Para calcular la varianza, primero encontraremos cúanto vale el segundo momento de la

variable aleatoria Y, utilizando integración por partes dos veces, obtenemos lo siguiente

$$\begin{aligned}\int_0^\infty y^2 \frac{1}{4} e^{-\frac{y}{4}} dy &= -y^2 e^{-\frac{y^2}{4}}|_0^\infty + \int_0^\infty 2ye^{-\frac{y}{4}} dy \\ &= -8ye^{-\frac{y}{4}}|_0^\infty + \int_0^\infty ye^{-\frac{y}{4}} dy \\ &= -32e^{-\frac{y}{4}}|_0^\infty \\ &= 32\end{aligned}$$

De esta última cuenta se sigue que la varianza de Y está dada por $Var(Y) = E(Y^2) - E^2(Y) = 32 - 4^2 = 32 - 16 = 16$.
Ahora calcularemos $E(X|Y)$

$$\begin{aligned}E(X|Y=y) &= \int_0^\infty x \frac{e^{-\frac{x}{2y}}}{2y} dx \\ &= -e^{-\frac{x}{2y}}|_0^\infty + \int_0^\infty e^{-\frac{x}{2y}} dx \\ &= -2ye^{-\frac{x}{2y}}|_0^\infty = 2y\end{aligned}$$

Donde la segunda igualdad se da usando el método de integración por partes, tomando $f(x) = \frac{1}{2y}x$ y $g'(x) = e^{-\frac{x}{2y}}$.
Por lo que, abusando de la notación, tenemos que $E(X|Y) = 2Y$.
Finalmente, procederemos a calcular $E(X^2|Y)$, esto para poder encontrar la varianza condicional de X dado Y, para esto, usaremos el método de integración por partes dos veces

$$\begin{aligned}\int_0^\infty \frac{x^2}{2y} e^{-\frac{x}{2y}} dx &= x^2 e^{-\frac{x}{2y}}|_0^\infty + \int_0^\infty 2xe^{-\frac{x}{2y}} dx \\ &= \int_0^\infty 2xe^{-\frac{x}{2y}} dx \\ &= -4y \cdot 2ye^{-\frac{x}{2y}}|_0^\infty \\ &= -8y^2 e^{-\frac{x}{2y}}|_0^\infty \\ &= 8y^2\end{aligned}$$

Así, tenemos que $E(X^2|Y) = 8Y^2$ y por ende, $Var(X|Y) = 8Y^2 - 4Y^2 = 4Y^2$

2. Sea X_1, X_2, X_3, X_4, X_5 muestra aleatoria de tamaño $n = 5$ de una población con densidad dada por

$$f_X(x;\theta) = \frac{\theta^x}{x!} e^{-\theta}; \quad \theta > 0; \quad x \in \{0, 1, 2, ...\}$$

Considere las siguientes funciones de la muestra

$$\hat{\theta}_1 = X_1 + X_2 + X_3 + X_4 + X_5$$
$$\hat{\theta}_2 = X_1 + \theta X_2$$
$$\hat{\theta}_3 = \frac{X_1+X_2+X_3+X_4+X_5}{5}$$
$$\hat{\theta}_4 = \frac{2X_1+3X_2}{5}$$

   a) Indique cuáles funciones de la muestra pueden ser considerados estimadores para θ y cuáles no.

Respuesta. Claramente $\hat{\theta_2}$ no puede ser un estimador para θ, pues ni siquiera es una estadística, pues depende de un parámetro desconocido (θ).
Por otra parte, es claro también que $\hat{\theta_1}$ y $\hat{\theta_3}$ son estimadores para θ, al igua que $\hat{\theta_4}$.

b) Indique cuáles estimadores son insesgados y cuáles sesgados.

Solución. Veamos primero calcularemos la esperanza de X_1

$$\begin{aligned} E(X_1) &= \sum_{x=0}^{\infty} x\frac{\theta^x}{x!}e^{-\theta} \\ &= e^{-\theta}\theta\sum_{x=0}^{\infty}\frac{\theta^{x-1}}{(x-1)!} \\ &= e^{-\theta}\theta e^{\theta} \\ &= \theta \end{aligned}$$

Por lo tanto, $E(X_1) = \theta$, y dado que estamos tomando en cuenta una muestra aleatoria, entonces, la esperanza de $X_2, ..., X_5$ es la misma.
Ahora, veamos si $\hat{\theta_1}$ es insesgado, para esto, por linealidad de la esperanza tenemos que $E(X_1 + X_2 + X_3 + X_4 + X_5)$

$$E(\hat{\theta_1}) = E(X_1 + X_2 + X_3 + X_4 + X_5) = \sum_{i=1}^{5} E(X_i) = \sum_{i=1}^{5}\theta = 5\theta$$

Por lo tanto, concluimos que $\hat{\theta_1}$ es sesgado.
De la cuenta anterior, notamos que

$$E(\hat{\theta_3}) \cdot 5 = E(\hat{\theta_1}) = 5\theta$$

Por lo tanto, fácilmente tenemos que $E(\hat{\theta_3}) = \theta$ y por ende, $\hat{\theta_3}$ es un estimador insesgado.
Finalmente, veamos si $\hat{\theta_4}$ es un estimador insesgado

$$E(\hat{\theta_4}) = E\left(\frac{2X_1 + 3X_2}{5}\right) = \frac{1}{5}E(2X_1 + 3X_2) = \frac{2}{5}E(X_1) + \frac{3}{5}E(X_2) = \frac{2}{5}\theta + \frac{3}{5}\theta = \theta$$

De esto último, cocluimos que $\hat{\theta_3}$ es también un estimador insesgado.

3. Sea X_1, X_2, X_3 una muestra aleatoria de tamaño 3 de una población con distribución $F_X(x)$ tal que $E(X_i) = \mu$ y $Var(X_i) = 1$. Calcular el error cuadrático medio para los siguientes estimadores de μ

$$\hat{\mu_1} = X_1 \qquad ; \qquad \hat{\mu_2} = \frac{3X_1 - 2X_2 + X_3}{6}$$

¿Para qué valores de μ se tiene que $\hat{\mu_2}$ es mejor estimador que $\hat{\mu_1}$?

Solución. Calcularemos el error cuadrático medio de $\hat{\mu_1}$, para esto, sabemos que

$$E.C.M(\hat{\mu_1}) = Var(\hat{\mu_1}) + (E(\hat{\mu_1}) - \mu)^2 = Var(X_1) + (E(X_1) - \mu)^2 = 1 + (\mu - \mu)^2 = 1$$

Por otra parte, calcularemos el error cuadrático medio para $\hat{\mu_2}$

$$
\begin{aligned}
E.C.M(\hat{\mu_2}) &= Var(\hat{\mu_2}) + (E(\hat{\mu_2}) - \mu)^2 \\
&= Var\left(\frac{3X_1 - 2X_2 + X_3}{6}\right) + \left(E\left(\frac{3X_1 - 2X_2 + X_3}{6}\right) - \mu\right)^2 \\
&= \frac{9}{36}Var(X_1) - \frac{4}{36}Var(X_2) + \frac{1}{36}Var(X_3) + \left(\frac{3}{6}E(X_1) + \frac{2}{6}E(X_2) + \frac{1}{6}E(X_3) - \mu\right)^2 \\
&= \frac{9}{36} - \frac{4}{36} + \frac{1}{36} + \left(\frac{3}{6}\mu + \frac{2}{6}\mu + \frac{1}{6}\mu - \mu\right)^2 \\
&= \frac{6}{36} \\
&= \frac{1}{6}
\end{aligned}
$$

Por lo que respondiendo a la pregunta, dado que el error cuadrático medio de ambos estimadores resultó ser constante, podemos concluir que $\hat{\mu_2}$ es mejor estimador que $\hat{\mu_1}$ para cualquier valor de μ.

4. Sea $X_1, ..., X_n$ muestra aleatoria de tamaño n de una población con distribución $F_X(x)$ tal que $E(X_i) = \mu$ y $Var(X_i) = \sigma^2$. Sea $\alpha_1, ..., \alpha_n \in \mathbb{R}$ tales que $\sum_{i=1}^n \alpha_i = 1$. Defina el siguiente estimador para μ

$$\hat{\mu} = \sum_{i=1}^n \alpha_i X_i$$

Pruebe que

a) $\hat{\mu}$ es un estimador insesgado

Prueba.

$$
\begin{aligned}
E(\hat{\mu}) &= E\left(\sum_{i=1}^n \alpha_i X_i\right) \\
&= \sum_{i=1}^n E(\alpha_i X_i) \\
&= \sum_{i=1}^n \alpha_i E(X_i) \\
&= \sum_{i=1}^n \alpha_i \mu \\
&= \mu \sum_{i=1}^n \alpha_i \\
&= \mu
\end{aligned}
$$

Por lo tanto, concluimos que $\hat{\mu}$ es un estimador insesgado.

b) La varianza de $\hat{\mu}$ se minimiza cuando $\alpha_i = \frac{1}{n}$ para toda $i \in \{1, ..., n\}$. (Este inciso muestra entonces que $\bar{X}$ es el mejor estimador lineal). Cuado ocurre esto, se dice que el estimador es BLUE (Best Linear Unbised Estimator).

Prueba. Calculemos la varianza del estimador

$$\begin{aligned} Var(\hat{\mu}) &= Var\left(\sum_{i=1}^{n}\alpha_i X_i\right) \\ &= \sum_{i=1}^{n} Var(\alpha_i X_i) \\ &= \sum_{i=1}^{n} \alpha_i^2 Var(X_i) \\ &= \sum_{i=1}^{n} \alpha_i^2 \sigma^2 \\ &= \sigma^2 \sum_{i=1}^{n} \alpha_i^2 \end{aligned}$$

Ahora, definimos la funcion $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, dada por

$$f(\alpha_1, ..., \alpha_n) = \sigma^2 \sum_{i=1}^{n} \alpha_i^2$$

Y consideramos el siguiente conjunto

$$S = \{\alpha \in \mathbb{R}^n : \sum_{i=1}^{n} \alpha_i = 1\}$$

de este conjunto, es intuitivo definir la función $g : \mathbb{R}^n \longrightarrow \mathbb{R}$, dada por

$$g(\alpha) = \sum_{i=1}^{n} \alpha_i - 1$$

Es fácil ver que S es el conjunto de nivel cero de la función g, ahora, supongamos que $f|_S$ alcanza un minimo local (no puede ser máximo local, pues podemos tomar alguna α_i arbitrariamente grande compensando las demás), entonces, por el *Teorema de los Multiplicadores de Lagrange* (visto en Cálculo 3 y Análisis 1), $\exists \lambda \in \mathbb{R}$ tal que

$$\lambda(1, ..., 1) = \lambda \nabla g(\alpha) = \nabla f(\alpha) = (2\sigma^2\alpha_1, ..., 2\sigma^2\alpha_n)$$

Por lo que nuestro problema queda resuelto si logramos resolver el siguiente sistema de ecuaciones

$$\begin{aligned} 2\sigma^2\alpha_1 &= \lambda \\ 2\sigma^2\alpha_2 &= \lambda \\ &\vdots \\ 2\sigma^2\alpha_n &= \lambda \end{aligned}$$

Despejando α_i, tenemos que

$$\alpha_i = \frac{\lambda}{2\sigma^2} \qquad \forall i \in \{1, ..., n\} \tag{1}$$

De esto, tenemos que

$$1 = \sum_{i=1}^{n} \alpha_i = \frac{n\lambda}{2\sigma^2}$$

Por lo que es claro que $\lambda = \frac{2\sigma^2}{n}$, así, sustituyendo el valor de λ en la ecuación (1), tenemos que

$$\alpha_i = \frac{2\sigma^2}{n}\frac{1}{2\sigma^2} = \frac{1}{n} \qquad \forall i \in \{1, ..., n\}$$

Por ende, tenemos que la varianza de $\hat{\mu}$ se minimiza cuando $\alpha_i = \frac{1}{n} \quad \forall i \in \{1, ..., n\}$