Maestría en Computo Estadístico
Inferencia Estadística
Tarea 3
20 de septiembre de 2020
Enrique Santibáñez Cortés
Repositorio de Git: Tarea 3, IE.

Escriba de manera concisa y clara sus resultados, justificando los pasos necesarios. Serán descontados puntos de los ejercicios mal escritos y que contenga ecuaciones sin una estructura gramatical adecuada. Las conclusiones deben escribirse en el contexto del problema. Todos los programas y simulaciones tienen que realizarse en R.

1. Resuelva lo siguiente:

a) Sea $X \sim Exponencial(\beta)$. Encuentre $\mathbb{P}(|X-\mu_X| \geq k\sigma_X)$ para $k>1$. Compare esta probabilidad con la que obtiene de la desigualdad de Chebyshev.

RESPUESTA

Cómo $X \sim Exponencial(\beta)$ entonces $\mu_X=\beta$ y $\sigma_X=\sqrt{\beta^2}=\beta$. Por lo que:

$$\mathbb{P}(|X-\mu_X| \geq k\sigma_X) = \mathbb{P}(|X-\beta| \geq k\beta) = \mathbb{P}(X-\beta \leq -k\beta) + \mathbb{P}(X-\beta \geq k\beta)$$

$$= \mathbb{P}(X \leq -k\beta+\beta) + \mathbb{P}(X \geq k\beta+\beta),$$

ahora como $k>1$ eso implica que $-k\beta+\beta<0$, como X es una variable con distribución exponencial podemos concluir que $\mathbb{P}(X \leq -k\beta+\beta)=0$. Continuando simplificando y sabiendo que X se distribuye exponencial:

$$\begin{aligned}\mathbb{P}(|X-\mu_X| \geq k\sigma_X) &= \mathbb{P}(X \geq k\beta+\beta)\\ &= F(k\beta+\beta) \qquad \text{(fda exponencial)}\\ &= e^{-\frac{k\beta+\beta}{\beta}}\\ &= e^{-k-1}.\end{aligned}$$

Ahora, recordemos la desigualdad de Chebyshev.

Teorema: 1 *(Desigualdad de Chebyshev) Sea X una v.a.,$\mu=\mathbb{E}(X)$ y $\sigma^2=V(X)$. Entonces, si $t>0$*

$$\mathbb{P}(|X-\mu| \geq t) \leq \frac{\sigma^2}{t^2} \tag{1}$$

Entonces, considerando el teorema anterior y haciendo a $t=k\sigma_X$ (note que se sigue cumpliendo que $t>0$) tenemos:

$$\mathbb{P}(|X-\mu_X| \geq k\sigma_X) \leq \frac{\sigma_X^2}{(k\sigma_X)^2} = \frac{1}{k^2}.$$

Comparando la probabilidad obtenida con la cota de la desigualdad de Chebyshev:

$$e^{-k-1} \leq \frac{1}{k^2} \quad \blacksquare.$$

b) Sean $X_1, \cdots, X_n \sim Bernoulli(p)$ independientes y $\bar{X} = n^{-1}\sum_{i=1}^{n} X_i$. Usando las desigualdades de Chebyshev y Hoeffding, acote $\mathbb{P}(|\bar{X} - p| > \epsilon)$. Demuestre que para n grande la cota de Hoeffding es más pequeña que la cota de Chebyshev. ¿En qué beneficia esto?

RESPUESTA

Para calcula la cota de Chebyshev, primero calculemos $E[\bar{X}]$ y $Var[\bar{X}]$. Como las X_i son independientes y debido a que X_i sin $Bernoulli(p)$:

$$E[\bar{X}] = \frac{1}{n}E[\sum_{i=1}^{n} X_i] = \frac{1}{n}\sum_{i=1}^{n} E[X] = \frac{np}{n} = p, \quad \text{y}$$

$$Var[\bar{X}] = \frac{1}{n^2}Var[\sum_{i=1}^{n} X_i] = \frac{1}{n^2}\sum_{i=1}^{n} Var[X] = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}.$$

Utilizando (1) y haciendo $\epsilon = t$ tenemos:

$$\mathbb{P}(|\bar{X} - p| > \epsilon) \leq \mathbb{P}(|\bar{X} - p| \geq \epsilon) \leq \frac{\frac{p(1-p)}{n}}{\epsilon^2} = \frac{p(1-p)}{n\epsilon^2},$$

y como $0 \leq p \leq 1$ podemos ver que $p(1-p) \leq \frac{1}{4}$ y por lo tanto la cota de Chebyshev es:

$$\mathbb{P}(|\bar{X} - p| > \epsilon) \leq \frac{p(1-p)}{n\epsilon^2} \leq \frac{1}{4n\epsilon^2}. \tag{2}$$

Ahora, recordemos la desigualdad de Hoeffding.

Teorema: 2 *(Desigualdad de Hoeffding) Sea $Y_1,\ Y_2\ , \cdots, \mathbf{Y}_n$ v.a independientes tales que $E(Y_i) = 0$, $a_i \leq Y_i \leq b_i$. Sea $\epsilon > 0$. Entonces, para cualquier $t > 0$*

$$\mathbb{P}\left(\sum_{i=1}^{n} Y_i \geq \epsilon\right) \leq e^{-t\epsilon}\prod_{i=1}^{n} e^{t^2(b_i - a_i)^2/8}. \tag{3}$$

Para determinar las cota de la probabilidad solicitada observemos que

$$|\bar{X} - p| > \epsilon = (\bar{X} - p > \epsilon) \cup (p - \bar{X} > \epsilon). \tag{4}$$

Además,

$$\bar{X} - p = \frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} p = \sum_{i=1}^{n}\frac{X_i - p}{n}. \tag{5}$$

Entonces se reduce a encontrar la cota de Hoeffiding para:

$$\mathbb{P}\Big((\bar{X} - p > \epsilon) \cup (p - \bar{X} > \epsilon)\Big) = \mathbb{P}(\bar{X} - p > \epsilon) \cup \mathbb{P}(p - \bar{X} > \epsilon).$$

Primero encontremos la cota para $\mathbb{P}(\bar{X} - p > \epsilon)$. Para ello denotemos a $Y_i = \frac{X_i - p}{n},\ i = 1, 2, \cdots, n$. Ahora, como $X_i \sim Bernoulli(p)$ esto implica que

$$\mathbb{E}(Y_i) = \mathbb{E}\left(\frac{X_i - p}{n}\right) = \frac{\mathbb{E}(X_i) - p}{n} = \frac{p - p}{n} = 0.$$

Notemos que Y_i tiene una cota inferior haciendo $X = 0$ y una cota superior $X = 1$ las cuales son:

$$-\frac{p}{n} \leq Y_i \leq \frac{1-p}{n}.$$

Cómo ya cumple todos los supuestos, podemos ocupar la desigualdad de Hoeffding en las Y_i. Sea $\epsilon > 0$, y para cualquier $t > 0$

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{n} Y_i > \epsilon\right) = \mathbb{P}\left(\sum_{i=1}^{n} \frac{X_i - p}{n} > \epsilon\right) \leq \mathbb{P}\left(\sum_{i=1}^{n} \frac{X_i - p}{n} \geq \epsilon\right) &\leq e^{-t\epsilon} \prod_{i=1}^{n} e^{t^2(\frac{1-p}{n}+\frac{p}{n})^2/8} \\
&= e^{-t\epsilon} \prod_{i=1}^{n} e^{\frac{t^2}{8n^2}} \\
&= e^{-t\epsilon} e^{\frac{nt^2}{8n^2}} \\
&= e^{-t\epsilon + \frac{t^2}{8n}}.
\end{aligned}$$

Ya encontramos una cota utilizando la desigualdad de Hoeffding, pero no es una cota mínima. Para encontrar la mínima encontremos el valor de t que minimiza el exponente de e, es decir, encontremos un mínimo para $f(t) = -t\epsilon + \frac{t^2}{8n}$. Para ello utilicemos el criterio de primera y segunda derivada, derivamos $f(t)$:

$$f'(t) = -\epsilon + \frac{t}{4n}.$$

Igualamos a cero y despejamos t:

$$\begin{aligned}
-\epsilon + \frac{t}{4n} &= 0 \\
t &= 4n\epsilon.
\end{aligned}$$

Calculemos la segunda derivada de $f(t)$:

$$f''(t) = \frac{1}{4n}.$$

Como $f''(t) > 0$ implica que $t = 4n\epsilon$ sea un mínimo para $f(t)$, y la cota mínima de Hoeffding para este problema es cuanto $t = 4n\epsilon$. Sustituyendo en la desigualdad calculada:

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{n} \frac{X_i - p}{n} > \epsilon\right) &\leq e^{-(4n\epsilon)\epsilon + \frac{(4n\epsilon)^2}{8n}} \\
&= e^{-4n\epsilon^2 + \frac{16n^2\epsilon^2}{8n}} \\
&= e^{-2n\epsilon^2}.
\end{aligned}$$

Y por lo tanto (5) podemos concluir que:

$$\mathbb{P}\left(\bar{X} - p > \epsilon\right) \leq e^{-2n\epsilon^2}.$$

Realizando un razonamiento análogo determinemos la cota para $\mathbb{P}\left(p - \bar{X} > \epsilon\right)$. Para ello denotemos $Z_i = \frac{p - X_i}{n},\ i = 1, 2, \cdots, n$. Ahora, como $X_i \sim Bernoulli(p)$ esto implica que

$$\mathbb{E}(Z_i) = \mathbb{E}\left(\frac{p - X_i}{n}\right) = \frac{p - \mathbb{E}(X_i)}{n} = \frac{p - p}{n} = 0.$$

Notemos que Z_i tiene una cota inferior haciendo $X = 1$ y una cota superior $X = 0$ las cuales son:

$$\frac{p-1}{n} \leq Z_i \leq \frac{p}{n}.$$

Cómo ya cumple todos los supuestos, podemos ocupar la desigualdad de Hoeffding en las Z_i. Sea $\epsilon > 0$, y para cualquier $t > 0$

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{n} Z_i > \epsilon\right) = \mathbb{P}\left(\sum_{i=1}^{n} \frac{p - X_i}{n} > \epsilon\right) \leq \mathbb{P}\left(\sum_{i=1}^{n} \frac{p - X_i}{n} \geq \epsilon\right) &\leq e^{-t\epsilon} \prod_{i=1}^{n} e^{t^2(\frac{p}{n} - \frac{p-1}{n})^2/8} \\
&= e^{-t\epsilon} \prod_{i=1}^{n} e^{\frac{t^2}{8n^2}} \\
&= e^{-t\epsilon} e^{\frac{nt^2}{8n^2}} \\
&= e^{-t\epsilon + \frac{t^2}{8n}}.
\end{aligned}$$

Ya encontramos una cota utilizando la desigualdad de Hoeffding, pero no es una cota mínima. Para encontrar la mínima encontremos el valor de t que minimiza el exponente de e, es decir, encontremos un mínimo para $f(t) = -t\epsilon + \frac{t^2}{8n}$. Para ello utilicemos el criterio de primera y segunda derivada, derivamos $f(t)$:

$$f'(t) = -\epsilon + \frac{t}{4n}.$$

Igualamos a cero y despejamos t:

$$\begin{aligned}
-\epsilon + \frac{t}{4n} &= 0 \\
t &= 4n\epsilon.
\end{aligned}$$

Calculemos la segunda derivada de $f(t)$:

$$f''(t) = \frac{1}{4n}.$$

Como $f''(t) > 0$ implica que $t = 4n\epsilon$ sea un mínimo para $f(t)$, y la cota mínima de Hoeffding para este problema es cuanto $t = 4n\epsilon$. Sustituyendo en la desigualdad calculada:

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{n} \frac{p - X_i}{n} > \epsilon\right) &\leq e^{-(4n\epsilon)\epsilon + \frac{(4n\epsilon)^2}{8n}} \\
&= e^{-4n\epsilon^2 + \frac{16n^2\epsilon^2}{8n}} \\
&= e^{-2n\epsilon^2}.
\end{aligned}$$

Entonces:

$$\mathbb{P}\left(p - \bar{X} > \epsilon\right) \leq e^{-2n\epsilon^2}.$$

Por lo tanto, ocupando (4) podemos concluir que la cota de Hoeffding es:

$$\mathbb{P}\left(|\bar{X} - p| > \epsilon\right) \leq 2e^{-2n\epsilon^2}.$$

Entonces para mostrar que cuando n es grande la cota de Hoeffing es más pequeña que la cota de Chebyshev (1), veamos la primera derivada de cada cota con respecto a n. Sea $f(n)$ el denominador de la cota de Hoeffing y $g(n)$ el denominador la cota de Chebyshev:

$$\begin{aligned}
f'(n) &= 2\epsilon^2 e^{2n\epsilon^2}/2 = \epsilon^2 e^{2n\epsilon^2}. \\
g'(n) &= 4\epsilon^2
\end{aligned}$$

De aquí podemos observar que el crecimiento de la cota de Chebyshev es lineal no depende de n y en la cota de Hoeffing depende de n. Entonces de lo anterior implica que para que

$$g'(n) < f'(n)$$

se tiene que cumplir que $4 < e^{2n\epsilon^2}$, entonces es sencillo ver que para cualquier $\epsilon > 0$ se puede encontrar un n lo suficientemente grande para que se cumpla la igualdad. Por lo tanto, para algún n grande cumple que el dominador del cota de Hoeffindg es más grande que la cota de Chebyshev y por lo tanto que la cota de Hoeffing es más pequeña que la cota de Chebyshev para algún n grande. $\blacksquare$.

2. Sean $X_1, \cdots, X_n \sim Bernoulli(p)$.

a) Sea $\alpha > 0$ fijo y defina

$$\epsilon_n = \sqrt{\frac{1}{2n} \log\left(\frac{2}{\alpha}\right)}.$$

Sea $\hat{p} = n^{-1} \sum_{i=1}^{n} X_i$. Defina $C_n = (\hat{p}_n - \epsilon_n, \hat{p}_n + \epsilon_n)$. Use la desigualdad de Hoeffding para demostrar que

$$\mathbb{P}(C_n \text{contiene a } p) \geq 1 - \alpha$$

. Diremos que C_n es un $(1-\alpha)-$intervalo de confianza para p. En la practica, se trunca el intervalo de tal manera de que no vaya debajo del 0 o arriba del 1.

RESPUESTA

Observemos que el evento de que C_n contiene a p es igual al evento que:

$$\{p \notin C\} = |\hat{p} - p| > \epsilon_n.$$

Entonces, como $X_i \sim Bernoulli(p)$ independientes y utilizando la desigualdad de Hoeffding encontrada en el ejercicio anterior tenemos que:

$$\mathbb{P}(p \notin C) = \mathbb{P}(|\hat{p} - p| > \epsilon_n) \leq 2e^{-2n\epsilon_n^2},$$

sustituyendo el valor de ϵ_n:

$$\begin{aligned}
\mathbb{P}(p \notin C) &\leq 2e^{-2n\left(\sqrt{\frac{1}{2n}\log\left(\frac{2}{\alpha}\right)}\right)^2} \\
&= 2e^{-2n\left(\frac{1}{2n}\log\left(\frac{2}{\alpha}\right)\right)} \\
&= 2e^{-\log\left(\frac{2}{\alpha}\right)} \\
&= \frac{2\alpha}{2} = \alpha.
\end{aligned}$$

Es decir, la probabilidad de que el intervalo C_n no contenga a p es menor que α:

$$\mathbb{P}(p \notin C) \leq \alpha.$$

Recordando la propiedad de probabilidad $\mathbb{P}(A) + \mathbb{P}(A^c) = 1$. Entonces, podemos ver que

$$\begin{aligned}
\mathbb{P}(p \notin C) &\leq \alpha \\
-\mathbb{P}(p \notin C) &\geq -\alpha \\
1 - \mathbb{P}(p \notin C) &\geq 1 - \alpha \\
\mathbb{P}(p \in C) &\geq 1 - \alpha.
\end{aligned}$$

Es decir, queda probado que

$$\mathbb{P}(C_n \text{contiene a } p) \geq 1 - \alpha. \quad \blacksquare.$$

3. Una partícula se encuentra inicialmente en el origen de la recta real y se mueve en saltos de una unidad. Para cada salto, la probabilidad de que la partícula salte una unidad a la izquierda es p y la probabilidad de que salte una unidad a la derecha es $1-p$. Denotemos por X_n a la posición de la partícula después de n unidades. Encuentre $\mathbb{E}(X_n)$ y $\mathrm{Var}(X_n)$. Esto se conoce como una caminata aleatoria en una dimensión.

RESPUESTA

Este proceso cambia de un estado a otro en dos tiempos consecutivos de acuerdo con probabilidad de transición descritas en el problemas. Estas probabilidades se pueden escribir de la forma siguiente:

$$\mathbb{P}(X_{n+1} = j | X_n = i) = \begin{cases} p & \text{si } j = i+1, \\ 1-p & \text{si} j = i-1, \\ 0 & \text{otro caso.} \end{cases}$$

Esa caminata aleatoria puede también definirse de la siguiente forma: $\xi_1,\ \xi_2, \cdots$ una sucesión de variables aleatorias independientes e idénticamente distribuidas. Por la idéntica distribución denotemos a cualquiera de ellas mediante la letra ξ sin subíndice. Ahora, si suponemos que $\mathbb{P}(\xi = +1) = p$ y $\mathbb{P}(\xi = -1) = 1-p$. Entonces para $n \geq 1$ se define

$$X_n := X_0 + \xi_1 + \xi_2 + \cdots + \xi_n.$$

donde X_0 en este caso suponemos que empieza en 0, es decir, el estado inicial de la partícula es cero. Entonces, a partir de la expresión anterior implica que la esperanza es:

$$\mathbb{E}(X_n) = \mathbb{E}(X_0) + \sum_{i=1}^{n} \mathbb{E}(\xi_i) = n\mathbb{E}(\xi) = n(p-1+p) = n(2p-1).$$

Ahora, como $\mathbb{E}(\xi^2) = p(1)^2 + (1-p)(-1)^2 = 1$ y $\mathbb{E}(\xi) = 2p-1$, se tiene que $Var(\xi) = 1-(2p-1)^2 = 1-4p^2+4p-1 = 4p(1-p)$. Y por lo tanto la varianza de X_n es:

$$Var(X_n) = \sum_{i=1}^{n} Var(\xi_i) = nVar(\xi) = 4np(1-p). \quad \blacksquare.$$

7. Demuestre que la fórmula de la densidad de la Beta integra 1.

RESPUESTA

Sea $X \sim Beta(\alpha, \beta)$, entonces su función de densidad esta definida como:

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \qquad 0 < x < 1.$$

Entonces, la integral de la densidad sería:

$$\int_{-\infty}^{\infty} f(x)dx = \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}dx = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx. \tag{6}$$

Ahora recordemos una identidad con la función Gamma, la cuál es:

$$\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

Por lo anterior tenemos en (6):

$$\int_{-\infty}^{\infty} f(x)dx = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} = 1.$$

Por lo tanto, queda demostrado que la densidad de v.a con distribución Beta integra 1. $\blacksquare$.

Honors problems

1.

a) Sea X una v.a. discreta con media finita y que toma valores en el conjunto $0, 1, 2 \cdots$. Demuestre que

$$\mathbb{E}(X) = \sum_{k=1}^{\infty} \mathbb{P}(X \geq k).$$

RESPUESTA
Definamos la siguiente notación para hacer más entendible la demostración:

$$p_x = \mathbb{P}(X = x), \quad x = 0, 1, \cdots,$$

y

$$q_x = \mathbb{P}(X > x) = \sum_{k=x+1}^{\infty} p_k \qquad x = 0, 1, \cdots.$$

Entonces debemos probar que:

$$\mathbb{E}(x) = \sum_{k=1}^{\infty} \mathbb{P}(X \geq k) = \sum_{x=0}^{\infty} q_x.$$

Como sabemos que se cumple que $\mathbb{P}(X \leq x) + \mathbb{P}(x > x) = 1$, observemos que:

$$\begin{aligned}
\sum_{x=1}^{N} x p_x &= p_1 + 2p_2 + 3p_3 + \cdots + Np_N \\
&= (p_1 + p_2 + p_3 + \cdots + p_N) + (p_2 + p_3 + \cdots + p_N) + \cdots + (p_{N-1} + p_N) + p_N \\
&= \sum_{x=0}^{N-1} q_x - Nq_N.
\end{aligned}$$

Es decir,

$$\sum_{x=1}^{N} x p_x = \sum_{x=0}^{N-1} q_x - Nq_N. \tag{7}$$

Utilizando que X tiene media finita, es decir, como $E(X) = \sum_{x=0}^{\infty} x p_x$ podemos decir que la serie $\sum_{x=0}^{\infty} x p_x$ es convergente. Ahora observemos que se cumple la siguiente desigualdad:

$$0 \leq N \sum_{x=N+1}^{\infty} p_x \leq \sum_{x=N+1}^{\infty} x p_x. \tag{8}$$

La justificación de la desigualdad, se debe a que $n > N$ por como se definieron los límites de la suma y de lado derecho a que N y p_x son positivos.

Ahora, ocupemos una propiedad conocida de series convergentes:

Teorema: 3 *(Condición necesaria de convergencia) Si la serie $\sum_{n=1}^{\infty} a_n$ es convergente, entoncess*

$$\lim_{n\to\infty} a_n = 0.$$

Ocupando la propiedad anterior en (8):

$$\begin{aligned}\lim_{N\to\infty} 0 \leq \lim_{N\to\infty} N \sum_{x=N+1}^{\infty} p_x \leq \lim_{N\to\infty} \sum_{x=N+1}^{\infty} xp_x. \\ = \sum_{x=N+1}^{\infty} \lim_{x\to\infty} xp_x \\ = 0.\end{aligned}$$

Es decir, podemos decir que

$$\lim_{N\to\infty} N \sum_{x=N+1}^{\infty} p_x = 0.$$

Entonces haciendo un límite en (7) tenemos que:

$$\begin{aligned}\lim_{N\to\infty} \sum_{x=1}^{N} xp_x &= \lim_{N\to\infty} \left(\sum_{x=0}^{N-1} q_x - Nq_N \right) \\ &= \lim_{N\to\infty} \left(\sum_{x=0}^{N-1} q_x \right) - \lim_{N\to\infty} (Nq_N) \\ &= \lim_{N\to\infty} \left(\sum_{x=0}^{N-1} q_x \right) - \lim_{N\to\infty} \left(N \sum_{x=N+1}^{\infty} p_x \right) \\ &= \sum_{x=0}^{\infty} q_x = \sum_{k=1}^{\infty} \mathbb{P}(X \geq k).\end{aligned}$$

Es decir, queda demostrado que

$$\mathbb{E}(X) = \sum_{x=1}^{\infty} xp_x = \sum_{k=1}^{\infty} \mathbb{P}(X \geq k). \quad \blacksquare.$$

b) Sea X una v.a. continua no-negativa con media finita, función de densidad f y función de distribución F. Demuestre que

$$\mathbb{E}(X) = \int_0^{\infty} (1 - F(t))dt.$$

RESPUESTA

Utilizaremos algunas definiciones vistas en clase de confiabilidad y su relación con función de densidad. $R(x)$ se le conoce como la confiabilidad de X, la cual se define como:

$$R(x) = 1 - F(x).$$

Es claro que se cumple que:

$$R(0) = 1 \;\text{ y }\; \lim_{x\to\infty} R(x) = 0.$$

Derivando de ambos lados observamos que:

$$R'(x) = -F'(x) = -f(x).$$

Ocupando lo anterior y la definición de esperanza de una variable continua tenemos que:

$$\begin{aligned}
\mathbb{E}(X) &= \int_{-\infty}^{\infty} xf(x)dx && \text{definición de esperanza}\\
&= \int_{0}^{\infty} -xR'(x)dx && \text{relación confiabilidad-densidad}\\
&= \int_{0}^{\infty} -xR'(x) - R(x) + R(x)dx && \text{sumamos un cero}\\
&= \int_{0}^{\infty} -(xR(x))' + R(x)dx && \text{definición de derivada}\\
&= -xR(x)|_{x=0}^{\infty} + \int_{0}^{\infty} R(x)dx\\
&= \int_{0}^{\infty} R(x)dx\\
&= \int_{0}^{\infty}(1 - Fx(x))dx \quad \blacksquare.
\end{aligned}$$

c) ¿Cómo cambia la fórmula del caso anterior cuando el soporte de X es todo $\mathbb{R}$?

RESPUESTA

Cuando el soporte de X es todo $\mathbb{R}$ no tiene sentido la función de confiabilidad. Por lo que

2. Sea X una v.a. continua con primer momento finito. Demuestre que la función $G(c) = E(|X - c|).c \in \mathbb{R}$, se minimiza en $c = M(X)$ para $M(X)$ la mediana de X.

RESPUESTA

Sea $f(x)$ la función de densidad de X. Por propiedades de la esperanza tenemos que:

$$G(c) = E(|X - c|) = \int_{-\infty}^{\infty} |x - c|f(x)dx = \int_{-\infty}^{c} (c - x)f(x)dx + \int_{c}^{\infty} (x - c)f(x)dx$$

$$= c\int_{-\infty}^{c} f(x)dx - \int_{-\infty}^{c} xf(x)dx + \int_{c}^{\infty} xf(x)dx - c\int_{c}^{\infty} f(x)dx.$$

Ahora diferenciamos con respecto a c e igualamos a cero la expresión anterior:

$$G'(c) = \frac{d}{dc}\left(c\int_{-\infty}^{c} f(x)dx - \int_{-\infty}^{c} xf(x)dx + \int_{c}^{\infty} xf(x)dx - c\int_{c}^{\infty} f(x)dx\right)$$

$$= \int_{-\infty}^{c} f(x)dx + c\frac{d}{dc}\int_{-\infty}^{c} f(x)dx - \frac{d}{dc}\int_{-\infty}^{c} xf(x)dx + \frac{d}{dc}\int_{c}^{\infty} xf(x)dx - \int_{c}^{\infty} f(x)dx - c\frac{d}{dc}\int_{c}^{\infty} f(x)dx$$

Ahora por el Teorema Fundamental del Calculo.

$$\begin{aligned}
G'(c) &= \int_{-\infty}^{c} f(x)dx + cf(x)|_{x=c} - xf(x)|_{x=c} + xf(x)|_{x=c} - \int_{c}^{\infty} f(x)dx - cf(x)|_{x=c}\\
&= \int_{-\infty}^{c} f(x)dx - \int_{c}^{\infty} f(x)dx.
\end{aligned}$$

Ahora igualamos a cero:

$$\begin{aligned} G'(c) = \int_{-\infty}^{c} f(x)dx - \int_{c}^{\infty} f(x)dx = 0 \\ \int_{-\infty}^{c} f(x)dx = \int_{c}^{\infty} f(x)dx. \end{aligned}$$

Es decir,

$$\mathbb{P}(X \leq c) = \mathbb{P}(X > c).$$

Por definición de probabilidad sabemos que se cumple que $\mathbb{P}(X \leq c) + \mathbb{P}(X > c) = 1$. Lo que implica que:

$$\mathbb{P}(X \leq c) = \mathbb{P}(X > c) = \frac{1}{2}.$$

Y por lo tanto, cuando c es la mediana (por definición) de X minimiza la función $G(c)$ $\blacksquare$.