Probabilidad

Grado en Matemáticas

Curso 2018-2019

Hoja de Problemas 8

Fecha de Entrega: **Miércoles, 10 de Abril**

1. Enunciar la **desigualdad de Markov** para una variable aleatoria X no negativa, y demostrarla para el caso en el que X es **discreta**.

   (**Examen de Mayo de 2018**)

   Solución: La desigualdad de Markov establece que si X una variable aleatoria X no negativa con esperanza finita ($E(X) < \infty$), entonces, para cualquier $\epsilon > 0$ se verifica

$$P(X \geq \epsilon) \leq \frac{E(X)}{\epsilon}.$$

   Demostración para el caso discreto: Llamemos μ al valor esperado de la variable no negativa X, es decir,

$$\mu = E(X),$$

   que evidentemente verifica

$$0 \leq \mu < \infty.$$

   La esperanza de la variable discreta X responde a la fórmula

$$\mu = E(X) = \sum_{k \in S_X} k \cdot P(X = k),$$

   que puede descomponerse en dos sumandos

$$\mu = \sum_{\{k \in S_X : k < \epsilon\}} k \cdot P(X = k) + \sum_{\{k \in S_X : k \geq \epsilon\}} k \cdot P(X = k).$$

   Puesto que, por ser $X \geq 0$,

$$\sum_{\{k \in S_X : k < \epsilon\}} k \cdot P(X = k) \geq 0$$

   se tiene que

$$\begin{aligned} \mu &\geq \sum_{\{k \in S_X : k \geq \epsilon\}} k \cdot P(X = k) \\ &\geq \sum_{\{k \in S_X : k \geq \epsilon\}} \epsilon \cdot P(X = k) \\ &= \epsilon \sum_{\{k \in S_X : k \geq \epsilon\}} P(X = k) \\ &= \epsilon \cdot P(\, X \geq \epsilon), \end{aligned}$$

   y despejando de esta desigualdad se obtiene

$$P(\, X \geq \epsilon) \leq \frac{\mu}{\epsilon} = \frac{E(X)}{\epsilon}.$$

   c.q.d.

2. A día de hoy, y tras los incendios desatados en los últimos días, el consumo diario de de agua de los habitantes de Pedrógao Grande (Portugal) es una variable aleatoria A con distribución normal de media 4 litros y varianza 9 litros^2.

   Determinar la media y la mediana de A, y calcular los cuantiles $\varphi_{0.242}(A)$ y $\varphi_{0.998}(A)$.

   (**Examen de Junio de 2017**)

   Solución: La función de densidad de la distribución normal con respecto a μ y por tanto su media y su mediana (o segundo cuartil) coinciden:

$$E(A) = q_2(A) = \mathbf{4} \text{ litros.}$$

   El cuantil $\varphi_{0.242} = \varphi_{0.242}(A)$ es el valor que verifica

$$F_A(\varphi_{0.242}) = P(A \leq \varphi_{0.242}) = 0.242,$$

   y por tanto

$$P\left(\frac{A-4}{3} < \frac{\varphi_{0.242}-4}{3}\right) = P\left(Z < \frac{\varphi_{0.242}-4}{3}\right) = 0.242,$$

   donde Z denota la distribución normal estándar, $Z \sim N(0,1)$. En la tabla de la normal observamos que

$$P(Z \leq 0.7) = 0.758,$$

   luego

$$P(Z \geq 0.7) = 0.242,$$

   y por la simetría de Z,

$$P(Z \leq -0.7) = 0.242,$$

   De aquí se deduce que

$$\frac{\varphi_{0.242}-4}{3} = -0.7,$$

   y despejando obtenemos que

$$\varphi_{0.242}(A) = 4 - 0.7 \cdot 3 = \mathbf{1.9} \text{ litros.}$$

   Del mismo modo se deduce que

$$\varphi_{0.998}(A) = 4 + 2.88 \cdot 3 = \mathbf{12.64} \text{ litros.}$$

3. Demostrar que si X una variable aleatoria continua con función de densidad f_X, y $g : \mathbb{R} \longrightarrow \mathbb{R}$, es una función que verifica $g'(x) > 0 \quad$ para todo $x \in S_X$, entonces, la variable aleatoria

$$Y = g(X)$$

   tiene función de densidad dada por

$$f_Y(t) = \begin{cases} f_X(h(t)) \cdot |h'(t)| & \text{si } t \in (\alpha, \beta), \\ 0 & \text{si } t \notin (\alpha, \beta), \end{cases}$$

   donde

$$h(y) = g^{-1}(y),$$

$$\alpha = \min\{\, g(z) \;:\; z \in S_X \},$$

$$\beta = \max\{\, g(z) \;:\; z \in S_X \}.$$

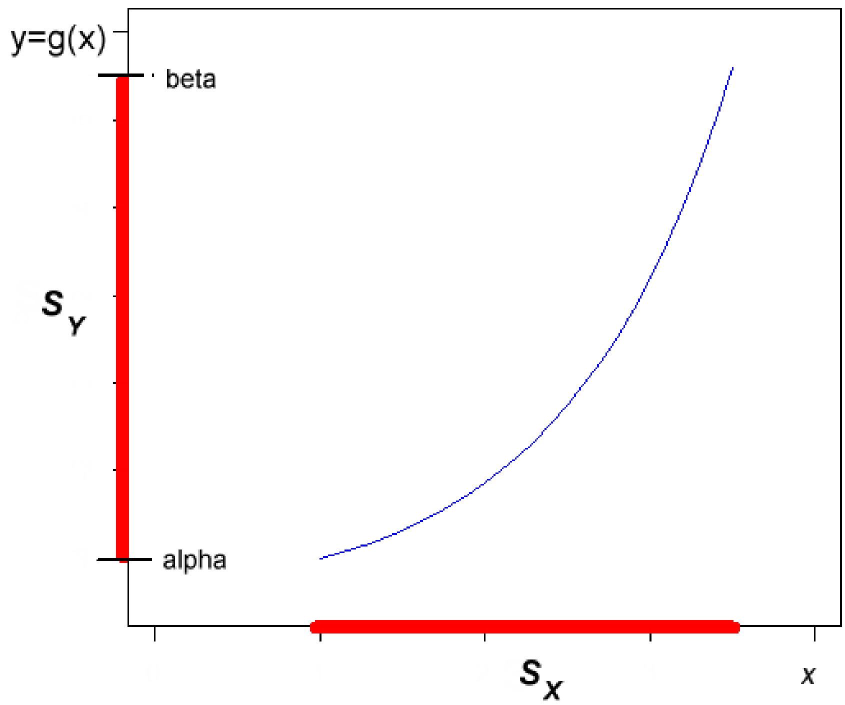
Solución: Consideremos una variable aleatoria continua X , con función de densidad f_X, y sea F_X su función de distribución, que viene dada por

$$F_X(z) = P[X \leq z] = \int_{-\infty}^{z} f_X(x)dx.$$

Sea $g : \mathbb{R} \longrightarrow \mathbb{R}$, una función que verifica $g'(x) > 0$ para todo $x \in S_X$.

Esto supone que g es **estrictamente creciente** en el soporte de X, S_X, y por tanto es una función biyectiva (y por tanto **invertible**) en este rango de valores, es decir, para $t \in S_Y = g(S_X)$.

De forma esquemática, y sin pérdida de generalidad, podría representarse $g(X)$ como



Para conocer la distribución de probabilidades de la variable aleatoria $Y = g(X)$ comenzaremos por buscar su funcion de distribución (F_Y) y a partir de ésta determinaremos su función de densidad (f_Y).

Sean

$$\alpha = \min\{\, g(z) \;:\; z \in S_X\},$$

$$\beta = \max\{\, g(z) \;:\; z \in S_X\}.$$

Obviamente el soporte de la variable aleatoria $Y = g(X)$ viene dado por $S_Y = (\alpha, \beta)$, y por tanto se tiene que

$$F_Y(t) = 0 \quad \text{para } t \leq \alpha,$$

y

$$F_Y(t) = 1 \quad \text{para } t \geq \beta.$$

Por otra parte, para $t \in (\alpha, \beta)$ se tiene (por ser g creciente) que

$$F_Y(t) = P[Y \leq t] = P[g(X) \leq t] = P\left[g^{-1}\left[(g(X)\right] \leq g^{-1}(t)\right] = P\left[X \leq g^{-1}(t)\right].$$

Llamemos

$$h(t) = g^{-1}(t).$$

Nótese que $h(t)$ es una función estrictamente creciente en (α, β), y por tanto su derivada verifica

$$h'(t) \geq 0 \quad \text{para todo } t \in (\alpha, \beta).$$

Con esta notación se tiene que

$$F_Y(t) = P[Y \leq t] = P[g(X) \leq t] = P\left[X \leq g^{-1}(t)\right] = P[X \leq h(t)] = P[X \leq h(t)] = F_X(h(t)).$$

Juntando sus tres tramos se deduce que la expresión completa para la función de distribución de la variable transformada Y, F_Y, es

$$F_Y(t) = P[Y \leq t] = \begin{cases} 0 & \text{si } t \leq \alpha, \\ F_X(h(t)) & \text{si } t \in (\alpha, \beta), \\ 1 & \text{si } t \geq \beta. \end{cases}$$

Derivando esta expresión se obtiene que la función de densidad de Y para $t \in (\alpha, \beta)$ es:

$$f_Y(t) = F_Y'(t) = F_X'(h(t)) = f_X(h(t))\ h'(t).$$

Puesto que

$$h'(t) \geq 0 \quad \text{para todo } t,$$

se tiene que

$$|h'(t)| = h'(t),$$

por lo que f_Y puede expresarse como

$$f_Y(t) = \begin{cases} f_X(h(t)) \cdot |h'(t)| & \text{si } t \in (\alpha, \beta), \\ 0 & \text{si } t \notin (\alpha, \beta), \end{cases}$$

c.q.d.

Observación: En determinados casos puede ocurrir que $\alpha = -\infty$ y/o $\beta = \infty$, lo cual no hace perder generalidad a la demostración.

4. El soporte de la variable aleatoria H es

$$S_H = \{-3, -1, 1, 3\},$$

y su función de masa de probabilidad viene dada por la expresión

$$f_H(x) = c\, x^2 \qquad \text{para } x \in S_H,$$

siendo $c \in \mathbb{R}$ una constante.

(a) Determinar el valor de c para que f_H sea una función de masa de probabilidad.

(b) Determinar los cuartiles de H y los cuantiles $\varphi_{0.1}(H)$, $\varphi_{0.4}(H)$, $\varphi_{0.9}(H)$, y $\varphi_{0.99}(H)$.

Solución:

(a) Para que f_H sea una función de masa debe verificarse

$$\sum_{x \in S_H} f_H(x) = (-3)^2 c + (-1)^2 c + 1^2 c + 3^2 c = 20 \times c = 1 \quad \Rightarrow \quad \boxed{c = \mathbf{\frac{1}{20}}}$$

Por tanto, expresado de forma abreviada, el soporte S_H y la función de masa de f_H son

$$H \equiv \begin{pmatrix} -3 & -1 & 1 & 3 \\ \\ \dfrac{9}{20} & \dfrac{1}{20} & \dfrac{1}{20} & \dfrac{9}{20} \end{pmatrix}$$

o, lo que es lo mismo,

$$H \equiv \begin{pmatrix} -3 & -1 & 1 & 3 \\ \\ 0.45 & 0.05 & 0.05 & 0.45 \end{pmatrix}$$

(b) Por tratarse de una variable aleatoria discreta, cada cuantil de orden p es el menor valor en el que la variable H acumula una probabilidad mayor o igual que p:

$$\varphi_p(H) = F_H^{-1}(p) = \inf\{u \in \mathbb{R} : F_H(u) \geq p\}$$

La función de distribución de H es,

$$F_H(x) = \begin{cases} 0 & \text{si } x < -3, \\ 0.45 & \text{si } -3 \leq x < -1, \\ 0.5 & \text{si } -1 \leq x < 1, \\ 0.55 & \text{si } 1 \leq x < 3, \\ 1 & \text{si } x \geq 3. \end{cases}$$

Por tanto,

$$\begin{aligned} q_1(H) = \varphi_{0.25}(H) &= -3 \\ q_2(H) = \varphi_{0.50}(H) &= -1 \\ q_3(H) = \varphi_{0.75}(H) &= 3 \end{aligned}$$

$$\begin{aligned} \varphi_{0.1}(H) &= -3, \\ \varphi_{0.4}(H) &= -3, \\ \varphi_{0.9}(H) &= 3 \\ \varphi_{0.99}(H) &= 3. \end{aligned}$$

5. Sea X una variable aleatoria con esperanza $\mu = 80$ y desviación típica σ desconocida. Determinar un valor de σ para el cual puede asegurarse que se verifica

$$P(75 < X < 85) \geq 0.9$$

Solución: La **desigualdad de Chebyshev** implica que si X es una variable aleatoria con esperanza μ y varianza $\sigma^2 < \infty$, entonces, para cualquier $\epsilon > 0$ se verifica

$$P(\,|X - \mu| < \epsilon) \geq 1 - \frac{\sigma^2}{\epsilon^2}.$$

En nuestro caso, esto conlleva que

$$P(75 < X < 85) = P(75 - 80 < X - 80 < 85 - 80) = P(-5 < X - \mu < 5) = P(\,|X - \mu| < 5) \geq 1 - \frac{\sigma^2}{25}.$$

Queremos asegurar que se cumple

$$P(75 < X < 85) \geq 0.9,$$

y para ello debemos garantizar que ocurre

$$1 - \frac{\sigma^2}{25} \geq 0.9.$$

Despejando σ^2 de esta desigualdad, deducimos que esta varianza debe verificar

$$\sigma^2 \leq 2.5,$$

y en consecuencia la desviación típica tiene que cumplir

$$\sigma \leq \sqrt{2.5},$$

es decir,

$$\boxed{\sigma \leq 1.5811}$$