EYP 1027 Métodos Probabilísticos

Clase 17

Profesor: Reinaldo B. Arellano-Valle

Departamento de Estadística
Pontificia Universidad Católica de Chile

Segundo Semestre 2020

Contenido I

Distribución de funciones de vectores aleatorios

Estadísticos de orden

Para n variables aleatorias $X_1, \ldots, X_n$, todas ellas definidas sobre un mismo espacio de probabilidad, defina

$$X_{(k)} = \ k-\text{ésimo menor de } X_1, \ldots, X_n \Longrightarrow X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}.$$

Definición 1.1

El vector aleatorio $(X_{(1)}, \ldots, X_{(n)})$ se llama estadístico de orden, donde $X_{(k)}$ es el k-ésimo orden, y los ordenes extremos son

$$X_{(1)} = \min\{X_1, \ldots, X_n\} \quad \text{y} \quad X_{(n)} = \max\{X_1, \ldots, X_n\}.$$

Nota: recuerde que hay $n!$ maneras de reordenar n valores $x_1, \ldots, x_n$.

Distribución de funciones de vectores aleatorios

Caso iid F

A continuación se estudia el caso en que $X_1, \ldots, X_n$ son variables aleatorias iid con fda F, y con dfp $f = F'$ en el caso continuo.

Teorema 1.1

Sean $X_{(1)} = \min\{X_1, \ldots, X_n\}$ y $X_{(n)} = \max\{X_1, \ldots, X_n\}$, donde $X_1, \ldots, X_n$ son variables aleatorias iid con fda F. Entonces,

$$F_{X_{(1)}}(x) = 1 - (1 - F(x))^n \quad \text{y} \quad F_{X_{(n)}}(x) = (F(x))^n .$$

Si además, la fda F es continua, entonces,

$$\begin{aligned} f_{X_{(1)}}(x) &= \frac{dF_{X_{(1)}}(x)}{dx} = n\,(1 - F(x))^{n-1} f(x), \\ f_{X_{(n)}}(x) &= \frac{dF_{X_{(n)}}(x)}{dx} = n\,(F(x))^{n-1} f(x). \end{aligned}$$

Distribución de funciones de vectores aleatorios

Demostración 1.1

Idea de la demostración:

$$
\begin{aligned}
F_{X_{(1)}}(x) &= P(X_{(1)} \leq x) \\
&= P(\text{mín}\{X_1, \ldots, X_n\} \leq x) \\
&= 1 - P(\text{mín}\{X_1, \ldots, X_n\} > x) \\
&= 1 - P(X_1 > x, \ldots, X_n > x) \\
&= 1 - P(\cap_{i=1}^n \{X_i > x\}) \\
&= 1 - \prod_{i=1}^n P(X_i > x) \quad \text{(por independencia)} \\
&= 1 - \prod_{i=1}^n (1 - P(X_i \leq x)) \\
&= 1 - \prod_{i=1}^n (1 - F(x)) \quad (X_i \sim F \ \forall \ i) \\
&= 1 - (1 - F(x))^n .
\end{aligned}
$$

Similarmente,

$$\begin{aligned}
F_{X_{(n)}}(x) &= P(X_{(n)} \leq x) \\
&= P(\max\{X_1, \ldots, X_n\} \leq x) \\
&= P(X_1 \leq x, \ldots, X_n \leq x) \\
&= P(\cap_{i=1}^{n}\{X_i \leq x\}) \\
&= \prod_{i=1}^{n} P(X_i \leq x) \quad \text{(por independencia)} \\
&= \prod_{i=1}^{n} F(x) \\
&= (F(x))^n .
\end{aligned}$$

Ejemplo 1.1

Sean $X_1, \ldots, X_n \overset{iid}{\sim} U(0, \theta)$ $(\theta > 0)$; es decir,

$$F(x) = \begin{cases} 0, & \text{si } x < 0, \\ \dfrac{x}{\theta}, & \text{si } 0 \leq x < \theta, \\ 1, & \text{si } x \geq \theta, \end{cases}$$

$$f(x) = \begin{cases} \dfrac{1}{\theta}, & \text{si } 0 \leq x < \theta, \\ 0, & \text{si no.} \end{cases}$$

Luego,

$$f_{X_{(1)}}(x) = \begin{cases} \dfrac{n}{\theta}\left(1 - \dfrac{x}{\theta}\right)^{n-1}, & \text{si } 0 \leq x < \theta, \\ 0, & \text{si no,} \end{cases}$$

$$f_{X_{(n)}}(x) = \begin{cases} \dfrac{n}{\theta}\left(\dfrac{x}{\theta}\right)^{n-1}, & \text{si } 0 \leq x < \theta, \\ 0, & \text{si no.} \end{cases}$$

Tarea:

1) En el ejemplo anterior, encuentre $\mathsf{E}(X_{(j)})$ y $\mathsf{Var}(X_{(j)})$ para $j = 1, n$.

2) Sean $R_n = X_{(n)} - X_{(1)}$ y $M_n = (X_{(n)} + X_{(1)})/2$ el rango y el punto medio de la muestra, respectivamente. Determine el rango esperado y el punto medio esperado. Qué necesita para estudiar la distribuciones conjunta y marginales de R_n y M_n?

Ejemplo 1.2

Sean $X_1, \ldots, X_n \overset{iid}{\sim} \exp(1)$; es decir,

$$F(x) = \begin{cases} 0, & \text{si } x < 0, \\ 1 - e^{-x}, & \text{si } x \geq 0, \end{cases}$$

$$f(x) = \begin{cases} e^{-x}, & \text{si } x \geq 0, \\ 0, & \text{si no.} \end{cases}$$

Entonces,

$$f_{X_{(1)}}(x) = \begin{cases} ne^{-nx}, & \text{si } x \geq 0, \\ 0, & \text{si no,} \end{cases}$$

$$f_{X_{(n)}}(x) = \begin{cases} n\left(1 - e^{-x}\right)^{n-1} e^{-x}, & \text{si } x \geq 0, \\ 0, & \text{si no.} \end{cases}$$

En este caso, es claro que $X_{(1)} \sim \exp(n)$, de modo que $\mathsf{E}(X_{(1)}) = 1/n$ y $\mathsf{Var}(X_{(1)}) = 1/n^2$.

Tarea:

1) Determine $\mathsf{E}(X_{(n)})$ y $\mathsf{Var}(X_{(n)})$.

2) Determine el rango esperado y el punto medio esperado en este caso. Qué necesita para estudiar la distribuciones conjunta y marginales de R_n y M_n?

Distribución de funciones de vectores aleatorios

Teorema 1.2

Sean $X_1, \ldots, X_n \overset{\text{iid}}{\sim} F$. Entonces, para $k = 1, \ldots, n$, se tiene que,

$$\begin{aligned} F_{X_{(k)}}(x) &= \sum_{m=1}^{k} \binom{n}{m} (F(x))^m (1 - F(x))^{n-m} \quad (*) \\ &= \frac{\Gamma(n+1)}{\Gamma(k)\Gamma(n+1-k)} \int_0^{F(x)} y^{k-1}(1-y)^{n-k} dy \quad (**), \end{aligned}$$

ya que,

$$\sum_{m=1}^{k} \binom{n}{m} z^m (1-z)^{n-m} = \frac{\Gamma(n+1)}{\Gamma(k)\Gamma(n+1-k)} \int_0^{z} y^{k-1}(1-y)^{n-k} dy$$

para todo $0 \leq z \leq 1$ y $k = 1, \ldots, n$.

Distribución de funciones de vectores aleatorios continuos

De $(**)$ se tiene que $F_{X_{(k)}}(x) = F_Z(x)$, donde $Z \sim Beta(k, n+1-k)$.

En particular, si $X_1, \ldots, X_n \overset{\text{iid}}{\sim} U(0,1)$, entonces el k-ésimo menor se distribuye como,

$$X_{(k)} \sim Beta(k, n+1-k), \quad k = 1, \ldots, n.$$

Demostración 1.2

Defina,

$$Z(x) = \text{ exactamente } m \text{ de los } X_j\text{'s son } \leq x.$$

Luego,

$$Z(x) \sim Bin(n, F(x)), \quad \text{con } \; F(x) = P(X_j \leq x) \; \forall \; j$$

Ahora, use que

$$\{X_{(k)} \leq x\} \Longleftrightarrow \{Z(x) \geq k\} = \cup_{m=k}^{n} \{Z(x) = m\}.$$

Distribución de funciones de vectores aleatorios continuos

Es decir,

$$P(X_{(k)} \leq x) = P(Z(x) \geq k) = P(\cup_{m=k}^{n}\{Z(x) = m\}) = \sum_{m=k}^{n} P(Z(x) = m),$$

de donde se obtiene $(*)$. Para la demostración de $(**)$, vea el libro de Gut, pg 105.

En particular, si F es una fda continua, entonces la fdp del k-ésimo menor esta dada por,

$$f_{X_{(k)}}(x) = \underbrace{\frac{\Gamma(n+1)}{\Gamma(k)\Gamma(n+1-k)}}_{\left(\begin{array}{c} n+1 \\ k \end{array}\right)} (F(x))^{k-1} (1 - F(x))^{n-k} f(x).$$

Note que $\quad f_{X_{(k)}}(x) = f_Z(F(x))f(x),\;$ donde $\;Z \sim Beta(k, n+1-k)$.

Teorema 1.3

Si $X_1, \ldots, X_n \overset{\text{iid}}{\sim} F$, entonces la distribución conjunta de $X_{(1)}$ y $X_{(n)}$ esta dada por,

$$\begin{aligned} F_{X_{(1)},X_{(n)}}(x,y) &= F_{X_{(n}}(y) - P(X_{(1)} > x, X_{(n)} \leq y) \\ &= \begin{cases} (F(y))^n - (F(y) - F(x))^n, & \text{si } x < y, \\ (F(y))^n, & \text{si } x \geq y. \end{cases} \end{aligned}$$

En particular, si F es continua, entonces $X_{(1)}$ y $X_{(n)}$ tienen fdp conjunta dada por,

$$\begin{aligned} & f_{X_{(1)},X_{(n)}}(x,y) \\ & \quad = \begin{cases} n(n-1)\left(F(y) - F(x)\right)^{n-2} f(x)f(y), & \text{si } x < y, \\ 0, & \text{si no.} \end{cases} \end{aligned}$$

Demostración 1.3

Para cualquier par de variables aleatorias X e Y con fda conjunta $F_{X,Y}$, se tiene que

$$\underbrace{P(X \leq x, Y \leq y)}_{F_{X,Y}(x,y)} + P(X > x, Y \leq y) = \underbrace{P(Y \leq Y)}_{F_Y(y)}.$$

Luego,

$$F_{X,Y}(x,y) = F_Y(y) - P(X > x, Y \leq y).$$

Sean $X = X_{(1)}$ e $Y = X_{(n)}$, el máximo y el mínimo, respectivamente. Entonces,

$$F_{X_{(1)},X_{(n)}}(x,y) = F_{X_{(n)}}(y) - P(X_{(1)} > x, X_{(n)} \leq y),$$

Distribución de funciones de vectores aleatorios

donde $F_{X_{(n)}}(y) = (F(y))^n$ y

$$\begin{aligned} P(X_{(1)} > x, X_{(n)} \leq y) &= P(x < X_i \leq y,\ \text{para } i = 1, \ldots, n) \\ &= P(\cap_{i=1}^n \{x < X_i \leq y\}) \\ &= \prod_{i=1}^n P(x < X_i \leq y) \quad \text{(por independencia)} \\ &= \begin{cases} (F(y) - F(x))^n, & \text{si } x < y, \\ 0, & \text{si no,} \end{cases} \end{aligned}$$

de donde sigue el primer resultado. En el caso continuo, el resultado para la fdp conjunta se obtiene derivando la fda conjunta.

Distribución de funciones de vectores aleatorios

A partir de estos resultados, también se pueden obtener las distribuciones conjunta y marginales de las variables aleatorias

$$R_n = X_{(n)} - X_{(1)} \quad \text{y} \quad M_n = \frac{X_{(n)} + X_{(1)}}{2}.$$

Por ejemplo, para el caso continuo se tiene que,

$$f_{R_n}(x) = \begin{cases} n(n-1)\int_{-\infty}^{\infty} (F(x+y) - F(x))^{n-2} f(x+y)f(x)dy, & \text{si } x > 0, \\ 0, & \text{si no.} \end{cases}$$

Ejemplo 1.3

Si $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \exp(1)$, entonces

$$f_{X_{(1)},X_{(n)}}(x,y) = \begin{cases} n(n-1)\left(e^{-x}-e^{-y}\right)^{n-2} e^{-x}e^{-y}, & \text{si } 0<x<y<\infty, \\ 0, & \text{si no.} \end{cases}$$

y

$$f_{R_n}(r) = \begin{cases} n(n-1)\int_0^\infty \left(e^{-x}-e^{-x-y}\right)^{n-2} e^{-x-y}e^{-x}dy, & \text{si } 0<x<\infty, \\ 0, & \text{si no,} \end{cases}$$

$$= \begin{cases} ne^{-nx}, & \text{si } 0<x<\infty, \\ 0, & \text{si no.} \end{cases}$$

$$\Longrightarrow R_n = X_{(n)} - X_{(1)} \sim \exp(n).$$

1) Obtenga la distribución del rango muestral $R_n = X_{(n)} - X_{(1)}$ cuando $X_1, \ldots, X_n \overset{\text{iid}}{\sim} U(0,1)$

2) Sean $Y_1 = X_{(1)}$ e $Y_i = X_{(i)} - X_{(i-1)}$ para $i = 2, \ldots, n$. Asumiendo que $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \exp(1)$, determine la distribución conjunta de $Y_1, \ldots, Y_n$, y discuta si estas variables son independientes.

3) Para $n = 2$, con $X_1, X_2 \overset{\text{iid}}{\sim} Geom(p)$, pruebe que el rango y el mínimo son independientes. Compare con la misma situación en el caso $\exp(1)$.

4) Sean $X_1, \ldots, X_n$ variables aleatorias continuas iid con fdp f. Pruebe que,

$$f_{X_{(1)},\ldots,X_{(n)}}(y_1, \ldots, y_n) = \begin{cases} n! \prod_{i=1}^{n} f(y_i), & \text{si } y_1 < y_2 < \ldots < y_n, \\ 0, & \text{si no.} \end{cases}$$