Estadígrafos de orden

Inferencia Estadística

24 de septiembre de 2024

Notación

X	población; variable aleatoria $X\colon \Omega \longrightarrow \mathbb{R}$
$X_1, \ldots, X_n$	muestra aleatoria simple = variables aleatorias independientes e idénticamente distribuidas como X
$F(\cdot)$	función de distribución de X
$f(\cdot)$	función de densidad de X
T	estadígrafo; variable aleatoria $T\colon X(\Omega)^n \longrightarrow \mathbb{R}$
$F_T(\cdot)$	función de distribución de T
$f_T(\cdot)$	función de densidad de T

Muestra ordenada

$$\begin{array}{rccc} \vec{X}_{(\cdot)}: & X(\Omega)^n & \longrightarrow & \mathbb{R}^n \\ & \vec{X}(\omega) = (x_1, ..., x_n) & \longmapsto & (x_{(1)}, ..., x_{(n)}) \end{array}$$

donde $(x_{(1)}, ..., x_{(n)})$ es una permutación de $(x_1, ..., x_n)$ tal que

$$\min_{1 \le i \le n} x_i = x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)} = \max_{1 \le i \le n} x_i$$

k-ésimo estadígrafo de orden

$$\begin{array}{rccc} X_{(k)}: & X(\Omega)^n & \longrightarrow & \mathbb{R} \\ & \vec{X}(\omega) = (x_1, ..., x_n) & \longmapsto & x_{(k)} \end{array}$$

Distribución de estadígrafo de orden

- $F_{X_{(n)}}(x) = F(x)^n$
- $F_{X_{(1)}}(x) = 1 - [1 - F(x)]^n$
- $F_{X_{(k)}}(x) = \sum_{i=k}^{n} \binom{n}{i} F(x)^i [1 - F(x)]^{n-i}$

Para el último, considérese

$$\sum_{i=1}^{n} \mathbb{1}_{(-\infty,x]}(X_i) \sim \mathcal{B}(n, F(x))$$

Densidad de estadígrafo de orden

Si X es continua:

$$f_{X_{(k)}}(x) = n! f(x) \frac{F(x)^{k-1}}{(k-1)!} \frac{[1 - F(x)]^{n-k}}{(n-k)!}$$

Densidad de estadígrafo de orden

$$f_{X_{(k)}}(x) = n!\, f(x) \frac{F(x)^{k-1}}{(k-1)!} \frac{[1-F(x)]^{n-k}}{(n-k)!}$$

$k = n$

$$f_{X_{(n)}}(x) = \frac{\mathrm{d}F_{X_{(n)}}(x)}{\mathrm{d}x} = \frac{\mathrm{d}F(x)^n}{\mathrm{d}x} = n\ f(x)F(x)^{n-1}$$

Densidad de estadígrafo de orden

$$f_{X_{(k)}}(x) = n!\, f(x) \frac{F(x)^{k-1}}{(k-1)!} \frac{[1-F(x)]^{n-k}}{(n-k)!}$$

$k = 1$

$$f_{X_{(1)}}(x) = \frac{\mathrm{d}F_{X_{(1)}}(x)}{\mathrm{d}x} = \frac{\mathrm{d}\{1-[1-F(x)]^n\}}{\mathrm{d}x} = n\ f(x)[1-F(x)]^{n-1}$$

Densidad de estadígrafo de orden

$$f_{X_{(k)}}(x) = n!\, f(x) \frac{F(x)^{k-1}}{(k-1)!} \frac{[1-F(x)]^{n-k}}{(n-k)!}$$

$1 < k < n$

$$\begin{aligned}
f_{X_{(k)}}(x) &= \frac{\mathrm{d}F_{X_{(k)}}(x)}{\mathrm{d}x} = \frac{\mathrm{d}\left\{\sum_{i=k}^{n} \binom{n}{i} F(x)^i [1-F(x)]^{n-i}\right\}}{\mathrm{d}x} \\
&= \frac{\mathrm{d}\left\{F(x)^n + \sum_{i=k}^{n-1} \binom{n}{i} F(x)^i [1-F(x)]^{n-i}\right\}}{\mathrm{d}x} \\
&= n\, f(x) F(x)^{n-1} + \sum_{i=k}^{n-1} \binom{n}{i} i\, f(x) F(x)^{i-1} [1-F(x)]^{n-i} \\
&- \sum_{i=k}^{n-1} \binom{n}{i} F(x)^i (n-i)\, f(x) [1-F(x)]^{n-i-1}
\end{aligned}$$

Densidad de estadígrafo de orden

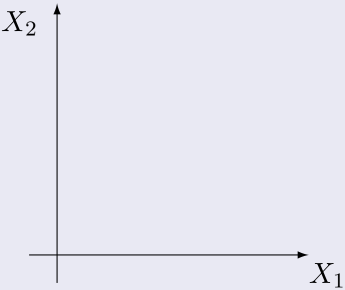
$$f_{X_{(k)}}(x) = n!\, f(x) \frac{F(x)^{k-1}}{(k-1)!} \frac{[1-F(x)]^{n-k}}{(n-k)!}$$

$1 < k < n$

$$\begin{aligned}
f_{X_{(k)}}(x) &= \sum_{i=k}^{n-1} \binom{n}{i} i\, f(x) F(x)^{i-1} [1-F(x)]^{n-i} + n\, f(x) F(x)^{n-1} \\
&- \sum_{i=k}^{n-1} \binom{n}{i} F(x)^{i} (n-i)\, f(x) [1-F(x)]^{n-i-1} \\
&= \sum_{i=k}^{n} \frac{n!}{(i-1)!\,(n-i)!} f(x) F(x)^{i-1} [1-F(x)]^{n-i} \\
&- \sum_{i=k+1}^{n} \frac{n!}{(i-1)!\,(n-i)!} f(x) F(x)^{i-1} [1-F(x)]^{n-i}
\end{aligned}$$

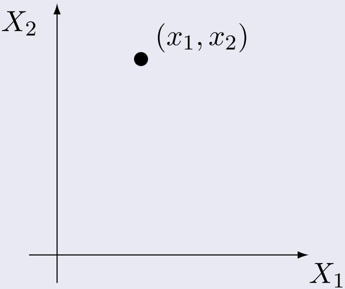
Densidad de la muestra ordenada

$$f_{\vec{X}_{(\cdot)}}(x_1, ..., x_n) = n! \prod_{i=1}^{n} f(x_i) \, \mathbb{1}(x_1 \leq ... \leq x_n)$$



Densidad de la muestra ordenada

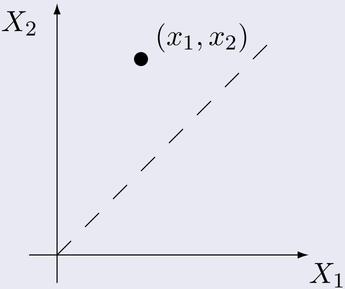
$$f_{\vec{X}_{(\cdot)}}(x_1, ..., x_n) = n! \prod_{i=1}^{n} f(x_i)\, \mathbb{1}(x_1 \leq ... \leq x_n)$$



Densidad de la muestra ordenada

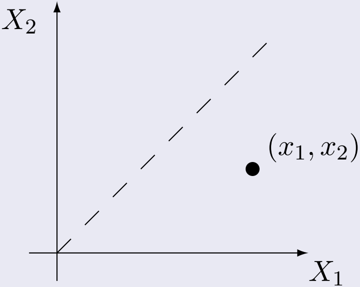
$$f_{\vec{X}_{(\cdot)}}(x_1, ..., x_n) = n! \prod_{i=1}^{n} f(x_i)\, \mathbb{1}(x_1 \leq ... \leq x_n)$$

$$x_{(1)} = x_1$$
$$x_{(2)} = x_2$$



Densidad de la muestra ordenada

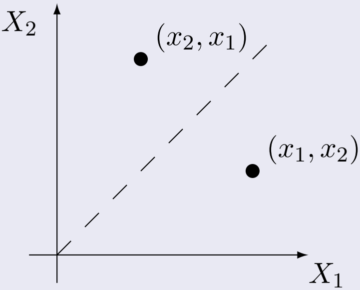
$$f_{\vec{X}_{(\cdot)}}(x_1, \ldots, x_n) = n! \prod_{i=1}^{n} f(x_i)\, \mathbb{1}(x_1 \leq \ldots \leq x_n)$$



Densidad de la muestra ordenada

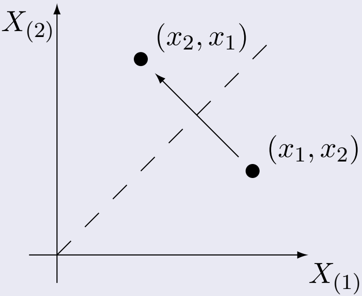
$$f_{\vec{X}_{(\cdot)}}(x_1, \dots, x_n) = n! \prod_{i=1}^{n} f(x_i)\, \mathbb{1}(x_1 \leq \dots \leq x_n)$$

$$x_{(1)} = x_2$$
$$x_{(2)} = x_1$$



Densidad de la muestra ordenada

$$f_{\vec{X}_{(\cdot)}}(x_1, \dots, x_n) = n! \prod_{i=1}^{n} f(x_i)\, \mathbb{1}(x_1 \leq \dots \leq x_n)$$



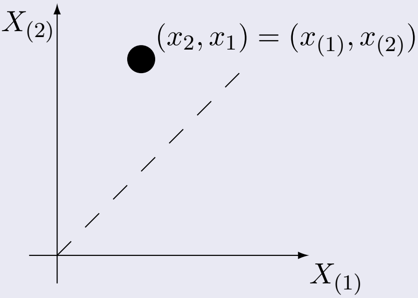
Densidad de la muestra ordenada

$$f_{\vec{X}_{(\cdot)}}(x_1, ..., x_n) = n! \prod_{i=1}^{n} f(x_i)\, \mathbb{1}(x_1 \leq ... \leq x_n)$$

$$f_{X_{(1)},X_{(2)}}$$

$$||$$

$$2 \cdot f_{X_1,X_2}$$



Densidad de la muestra ordenada

$$f_{\vec{X}_{(\cdot)}}(x_1, ..., x_n) = n! \prod_{i=1}^{n} f(x_i)\, \mathbb{1}(x_1 \leq ... \leq x_n)$$

Demostración

- Cada muestra ordenada proviene de $n!$ muestras.
- Alternativas:
  - usar Teorema de Cambio de Variable
  - usar función de distribución y derivar

Densidad de la muestra ordenada

$$f_{\vec{X}_{(\cdot)}}(x_1, ..., x_n) = n! \prod_{i=1}^{n} f(x_i) \, \mathbb{1}(x_1 \leq ... \leq x_n)$$

Teorema de Cambio de Variable

Sean

- X continua en $S = \bigcup_{i \in I} A_i$ con I numerable, $A_i \cap A_j = \emptyset$ si $i \neq j$
- g es diferenciable y tiene inversa g_i^{-1} en cada A_i

Entonces $Y = g(X)$ tiene densidad

$$f_Y(y) = \sum_{i \in I} f_X\left(g_i^{-1}(y)\right) \left|(g_i^{-1})'(y)\right| \qquad y \in g(S)$$

Densidad de la muestra ordenada

$$f_{\vec{X}_{(\cdot)}}(x_1, ..., x_n) = n! \prod_{i=1}^{n} f(x_i) \mathbb{1}(x_1 \leq ... \leq x_n)$$

Demostración por T.C.V.

Considerar los $n!$ subespacios de $\mathbb{R}^n$

$$\begin{aligned} A_1 &= \{x \in \mathbb{R}^n \mid x_1 < x_2 < \cdots < x_{n-1} < x_n\} \\ A_2 &= \{x \in \mathbb{R}^n \mid x_2 < x_1 < \cdots < x_{n-1} < x_n\} \\ &\vdots \\ A_{n!} &= \{x \in \mathbb{R}^n \mid x_n < x_{n-1} < \cdots < x_2 < x_1\} \end{aligned}$$

En cada A_i la muestra ordenada es una permutación con jacobiano ± 1

$$f_{X_{(\cdot)}}(x_1, ..., x_n) = \sum_{i=1}^{n!} \prod_{i=1}^{n} f(x_i)|\pm 1|$$

Densidad de la muestra ordenada

$$f_{\vec{X}_{(\cdot)}}(x_1, ..., x_n) = n! \prod_{i=1}^{n} f(x_i)\, \mathbb{1}(x_1 \leq ... \leq x_n)$$

Demostración directa para $n = 2$

$$F_{X_{(1)},X_{(2)}}(a, b) = \Pr\left[X_{(1)} \leq a \cap X_{(2)} \leq b\right]$$

$$= \Pr[X_1 \leq a \cap X_2 \leq b \cap X_1 \leq X_2] + \Pr[X_2 \leq a \cap X_1 \leq b \cap X_2 < X_1]$$

$$= \int_{-\infty}^{a} \int_{x_1}^{b} f(x_2)\,\mathrm{d}x_2\, f(x_1)\,\mathrm{d}x_1 + \int_{-\infty}^{a} \int_{x_2}^{b} f(x_1)\,\mathrm{d}x_1\, f(x_2)\,\mathrm{d}x_2$$

$$= \int_{-\infty}^{a} [F(b) - F(x_1)] f(x_1)\,\mathrm{d}x_1 + \int_{-\infty}^{a} [F(b) - F(x_2)] f(x_2)\,\mathrm{d}x_2$$

$$= \left| -\frac{1}{2}[F(b) - F(x_1)]^2 \right|_{x_1=-\infty}^{a} + \left| -\frac{1}{2}[F(b) - F(x_2)]^2 \right|_{x_2=-\infty}^{a}$$

Densidad de la muestra ordenada

$$f_{\vec{X}_{(\cdot)}}(x_1, ..., x_n) = n! \prod_{i=1}^{n} f(x_i)\, \mathbb{1}(x_1 \leq ... \leq x_n)$$

Demostración directa para $n = 2$

$$F_{X_{(1)},X_{(2)}}(a, b) = \Pr\big[X_{(1)} \leq a \cap X_{(2)} \leq b\big]$$

$$= F(b)^2 - [F(b) - F(a)]^2$$

$$\implies f_{X_{(1)},X_{(2)}}(a, b) = \frac{\partial F_{X_{(1)},X_{(2)}}(a, b)}{\partial a\, \partial b} = 2\, f(a) f(b)$$

Demostración directa para $n = 3$

$$F_{X_{(1)},X_{(2)},X_{(3)}}(a,b,c) = \Pr\left[X_{(1)} \le a \cap X_{(2)} \le b \cap X_{(3)} \le c\right]$$

$$= 3!\,\Pr[X_1 \le a \cap X_2 \le b \cap X_3 \le c \cap X_1 \le X_2 \le X_3]$$

$$= 6 \int_{-\infty}^{a} \int_{x_1}^{b} \int_{x_2}^{c} f(x_3)\,\mathrm{d}x_3\, f(x_2)\,\mathrm{d}x_2\, f(x_1)\,\mathrm{d}x_1$$

$$= 6 \int_{-\infty}^{a} \int_{x_1}^{b} [F(c) - F(x_2)] f(x_2)\,\mathrm{d}x_2\, f(x_1)\,\mathrm{d}x_1$$

$$= 6 \int_{-\infty}^{a} \left| -\frac{1}{2}[F(c) - F(x_2)]^2 \right|_{x_2=x_1}^{x_2=b} f(x_1)\,\mathrm{d}x_1$$

$$= 3 \int_{-\infty}^{a} \left\{ [F(c) - F(x_1)]^2 - [F(c) - F(b)]^2 \right\} f(x_1)\,\mathrm{d}x_1$$

$$= \left| -[F(c) - F(x_1)]^3 \right|_{x_1=-\infty}^{a} - 3[F(c) - F(b)]^2 F(a)$$

Demostración directa para $n = 3$

$$F_{X_{(1)},X_{(2)},X_{(3)}}(a,b,c) = \Pr\left[X_{(1)} \leq a \cap X_{(2)} \leq b \cap X_{(3)} \leq c\right]$$

$$= \left| -[F(c) - F(x_1)]^3 \right|_{x_1=-\infty}^{a} - 3[F(c) - F(b)]^2 F(a)$$

$$= F(c)^3 - [F(c) - F(a)]^3 - 3[F(c) - F(b)]^2 F(a)$$

$$\Longrightarrow f_{X_{(1)},X_{(2)},X_{(3)}}(a,b,c) = \frac{\partial F_{X_{(1)},X_{(2)},X_{(3)}}(a,b,c)}{\partial a\,\partial b\,\partial c} = 6\,f(a)f(b)f(c)$$