Estatísticas de ordem

As estatísticas de ordem, assim como os momentos amostrais, desempenham um papel importante na inferência estatística.

DEFINIÇÃO: Dada uma amostra aleatória $X_1, \dots, X_n$ com função de distribuição comum F. Coloque a amostra em ordem crescente

$$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)},$$ temos que

$$X_{(1)} = \min(X_1, \dots, X_n),$$

$$X_{(n)} = \max(X_1, \dots, X_n),$$

$X_{(i)}$ = i-é sima estatística de ordem.

Comentário: Note que $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ são variáveis aleatórias, mas não são independentes. Por exemplo,

$$P(X_{(1)} \leq y \mid X_{(n)} \leq y) = 1.$$

<u>DEFINIÇÃO</u>: Dada uma amostra aleatória $X_1, \dots, X_n$, a <u>mediana amostral</u> é dada por

$$M_o = \begin{cases} X_{\left(\frac{n+1}{2}\right)} & \text{se } n \text{ é ímpar} \\ \frac{1}{2}\left[X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n+2}{2}\right)} \right. & \text{se } n \text{ é par} \end{cases}$$

<u>Teorema</u>: Dada uma amostra aleatória $X_1, \dots, X_n$ com função de distribuição comum F.

Então,

$$F_{X_{(n)}}(t) = P(X_{(n)} \leq t) = [F(t)]^n$$

e

$$F_{X_{(1)}}(t) = P(X_{(1)} \leq t) = 1 - [1 - F(t)]^n.$$

<u>Demonstração</u>:

$$F_{X_{(n)}}(t) = P(X_{(n)} \leq t) = P(X_1 \leq t, \dots, X_n \leq t)$$

independência

$$= P(X_1 \leq t) \dots P(X_n \leq t) = [F(t)]^n.$$

$F_{X_{(1)}}(t) = P(X_{(1)} \leq t) = 1 - P(X_{(1)} > t) =$

EXERCÍCIO

E se não forem identicamente distribuídas?

$1 - P(X_1 > t, \ldots, X_n > t) \overset{\text{independência}}{=}$

$1 - P(X_1 > t) \ldots P(X_n > t) = 1 - [1 - F(t)]^n.$

Consequência:

Se as variáveis aleatórias $X_1, \ldots, X_n$ são contínuas e tem fdp f, temos que

$$f_{X_{(n)}}(t) = n\,[F(t)]^{n-1} f(t).$$

$$f_{X_{(1)}}(t) = n\,[1 - F(t)]^{n-1} f(t).$$

Teorema Se $X_1, \ldots, X_n$ é uma amostra aleatória com fdp $f(\cdot)$ e função de distribuição $F(\cdot)$. Sejam $\underset{X_{(1)}}{Y_1} \leq Y_\alpha \leq \ldots \leq \underset{X_{(n)}}{Y_n}$ as estatísticas de ordem, então

$$f_{Y_\alpha}(y) = \frac{n!}{(\alpha-1)!\,(n-\alpha)!} [F(y)]^{\alpha-1} [1 - F(y)]^{n-\alpha} f(y);$$

$$f_{Y_\alpha, Y_\beta}(x,y) = \frac{n!}{(\alpha-1)!\,(\beta-\alpha-1)!\,(n-\beta)!} [F(x)]^{\alpha-1} [F(y)-F(x)]^{\beta-\alpha-1}$$

$$[1-F(y)]^{n-\beta} f(x) f(y) \; I_{(x \le y \le \infty)} ;$$

$$f_{Y_1,\dots,Y_n}(y_1,\dots,y_n) = \begin{cases} n!\, f(y_1)\cdots f(y_n) & \text{para } y_1 < y_2 < \dots < y_n \\ 0 & \text{cc} \end{cases}$$

qualquer conjunto de densidades marginais podemos obter a partir da densidade conjunta

$f_{Y_1,\dots,Y_n}(y_1,\dots,y_n)$ por simples integração.

PROVAS DOS TEOREMAS: Mood / Barry James e Probability and Statistical Inference pág 319 (Rosset and Magdalena)

Exemplos na lista de exercícios !!

Exemplo:

$X_1, \ldots, X_n$ a.a. $X \sim U(0,1)$

qual a distribuição de $X_{(n)}$?

$X_{(n)} = \max\{X_1, \ldots, X_n\}$

$$f_{X_{(n)}}(y) = n\left[y\right]^{n-1} = n y^{n-1}, \quad 0 < y < 1.$$

Beta $(n, 1)$, pois

$$\beta(n,1) = \int_0^1 x^{n-1}\, dx = \frac{x^n}{n}\Big|_0^1 = \frac{1}{n}.$$

→ Importância das Estatísticas de ordem

→ distribuição conjunta: obtenção da distribuição da amplitude amostral.

Inferência não paramétrica!!

(POSTOS - RANKS)

→ distribuição dos estimadores!!

↳ $X_{(n)}$ no caso de $U(0, \theta)$, etc...