

Inferência Estatística II

Prof. Fernando de Souza Bastos
fernando.bastos@ufv.br

Departamento de Estatística
Programa de Pós-Graduação em Estatística Aplicada e Biometria
Universidade Federal de Viçosa
Campus UFV - Viçosa



Sumário

1 Consistência

2 Convergência em Distribuição

Definição 1

Seja X uma variável aleatória com função de distribuição acumulada $F(x, \theta)$, $\theta \in \mathcal{A} \subseteq \Omega$. Seja X_1, \dots, X_n uma amostra da distribuição de X e seja T_n uma estatística ($T_n = T(X_1, \dots, X_n)$). Dizemos que T_n é um estimador consistente para θ se $T_n \xrightarrow{P} \theta$.

Exemplo

Sejam X_1, \dots, X_n, \dots uma sequência de variáveis aleatórias iid de uma distribuição com média finita μ e variância $\sigma^2 < +\infty$, então, pela Lei

Fraca dos Grandes Números, temos que, $\bar{X}_n = \frac{\sum_{i=1}^n X_i}{n} \xrightarrow{P} \mu$. Ou seja, \bar{X}_n é um estimador consistente de μ .

Exemplo

Sejam X_1, \dots, X_n uma amostra aleatória de uma distribuição com média μ e variância $\sigma^2 < +\infty$. Suponha que $E[X_1^4] < +\infty$, de tal forma que $\text{Var}(S^2) < +\infty$.

$$\begin{aligned} S_n^2 &= \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \\ &= \frac{1}{n-1} \left(\sum_{i=1}^n X_i^2 - n\bar{X}_n^2 \right) \\ &= \frac{n}{n-1} \left(\frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}_n^2 \right) \\ &\xrightarrow{P} 1 \cdot [E(X_1^2) - \mu^2] = \sigma^2. \end{aligned}$$

Portanto, a variância da amostra é um estimador consistente de σ^2 . A partir da discussão acima, temos imediatamente que $S_n \xrightarrow{P} \sigma$; ou seja, o desvio padrão da amostra é um estimador consistente do desvio padrão populacional. Vejam estes exemplos anteriores na prática, [Clique aqui!](#).

Exemplo

Considere $X_i \stackrel{\text{iid}}{\sim} U(0, \theta)$, $i = 1, 2, \dots, n$, e $Y_n = \max\{X_1, \dots, X_n\}$.
Seja $\varepsilon > 0$, segue que:

$$\begin{aligned} P(|Y_n - \theta| \geq \varepsilon) &= P(\theta - Y_n \geq \varepsilon) \\ &= P(Y_n \leq \theta - \varepsilon). \end{aligned}$$

Exemplo

Considere $X_i \stackrel{\text{iid}}{\sim} U(0, \theta)$, $i = 1, 2, \dots, n$, e $Y_n = \max\{X_1, \dots, X_n\}$.
Seja $\varepsilon > 0$, segue que:

$$\begin{aligned} P(|Y_n - \theta| \geq \varepsilon) &= P(\theta - Y_n \geq \varepsilon) \\ &= P(Y_n \leq \theta - \varepsilon). \end{aligned}$$

Se $\theta - \varepsilon \leq 0$, então $P(Y_n \leq \theta - \varepsilon) = 0$, pois $0 \leq Y_n \leq \theta$, com $P(0 \leq Y_n \leq \theta) = 1$.

Exemplo

Considere $X_i \stackrel{\text{iid}}{\sim} U(0, \theta)$, $i = 1, 2, \dots, n$, e $Y_n = \max\{X_1, \dots, X_n\}$.
Seja $\varepsilon > 0$, segue que:

$$\begin{aligned} P(|Y_n - \theta| \geq \varepsilon) &= P(\theta - Y_n \geq \varepsilon) \\ &= P(Y_n \leq \theta - \varepsilon). \end{aligned}$$

Se $\theta - \varepsilon \leq 0$, então $P(Y_n \leq \theta - \varepsilon) = 0$, pois $0 \leq Y_n \leq \theta$, com $P(0 \leq Y_n \leq \theta) = 1$.

Se $0 < \varepsilon < \theta$ então,

$$P(|Y_n - \theta| \geq \varepsilon) = P(Y_n \leq \theta - \varepsilon) = F_{Y_n}(\theta - \varepsilon) = \left(\frac{\theta - \varepsilon}{\theta}\right)^n \xrightarrow{n \rightarrow \infty} 0$$

Ou seja, $Y_n \xrightarrow{P} \theta$. Logo, Y_n é um estimador consistente para θ .

Para

Exercícios 2.8.18, 5.1.2, 5.1.3, 5.1.7 e 5.1.9

Convergência em Distribuição

Definição 2

Seja $\{X_n\}_{n \geq 1}$ uma sequência de variáveis aleatórias com função de distribuição F_{X_n} , $n \geq 1$. Seja X uma variável aleatória com função de distribuição F_X . Seja $C(F_X)$ o conjunto de todos os pontos de continuidade de F_X . Dizemos que X_n converge em distribuição para X se,

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x), \quad \forall x \in C(F_X).$$

Denotamos essa convergência por $X_n \xrightarrow{D} X$ ou $X_n \xrightarrow{\mathcal{L}} X$.

Convergência em Distribuição

Definição 2

Seja $\{X_n\}_{n \geq 1}$ uma sequência de variáveis aleatórias com função de distribuição F_{X_n} , $n \geq 1$. Seja X uma variável aleatória com função de distribuição F_X . Seja $C(F_X)$ o conjunto de todos os pontos de continuidade de F_X . Dizemos que X_n converge em distribuição para X se,

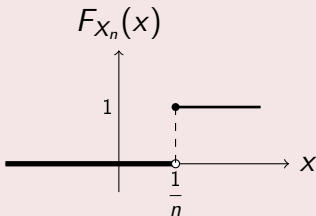
$$\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x), \quad \forall x \in C(F_X).$$

Denotamos essa convergência por $X_n \xrightarrow{D} X$ ou $X_n \xrightarrow{\mathcal{L}} X$.

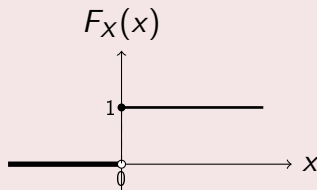
À medida que o tamanho da amostra aumenta, a distribuição das médias amostrais se aproxima da distribuição normal, veja isso acontecendo na prática, [Clique aqui!](#).

Exemplo

$$P\left(X_n = \frac{1}{n}\right) = 1, \forall n \geq 1, \quad P(X = 0) = 1, \quad \mathcal{C}(F_X(x)) = \{x \in \mathbb{R}; x \neq 0\}$$



(a) $P\left(X_n = \frac{1}{n}\right) = 1$



(b) $P(X = 0) = 1$

Figura: $\lim_{n \rightarrow +\infty} F_{X_n}(x) = F_X(x), \forall x \neq 0$, ou seja $X_n \xrightarrow{D} X$

Exemplo 2 (Convergência em Distribuição não Implica Convergência em Probabilidade)

Seja X uma variável aleatória contínua simétrica em torno do zero (ou seja, se f denota sua densidade, então $f(x) = f(-x), \forall x \in \mathbb{R}$. Neste caso, X e $-X$ tem a mesma distribuição (Verifiquem!). Defina a sequência de variáveis aleatórias X_n como:
$$X_n = \begin{cases} X, & \text{se } n \text{ é par} \\ -X, & \text{se } n \text{ é ímpar} \end{cases}$$

Exemplo 2 (Convergência em Distribuição não Implica Convergência em Probabilidade)

Seja X uma variável aleatória contínua simétrica em torno do zero (ou seja, se f denota sua densidade, então $f(x) = f(-x), \forall x \in \mathbb{R}$. Neste caso, X e $-X$ tem a mesma distribuição (Verifiquem!). Defina a sequência de variáveis aleatórias X_n como:
$$X_n = \begin{cases} X, & \text{se } n \text{ é par} \\ -X, & \text{se } n \text{ é ímpar} \end{cases}$$

É fácil ver que $F_{X_n}(x) = F_X(x)$. Logo, $X_n \xrightarrow{D} X$. Porém, $X_n \not\xrightarrow{P} X$, pois
$$P(|X_n - X| \geq \varepsilon) = \begin{cases} 0, & \text{se } n \text{ é par} \\ P(2|X| \geq \varepsilon), & \text{se } n \text{ é ímpar} \end{cases}$$

Exemplo 3

Seja T_n uma variável aleatória com distribuição t-Student com n graus de liberdade, ou seja, a densidade de T_n é dada por:

$$f_{T_n}(y) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)} \frac{1}{\left(1 + \frac{y^2}{n}\right)^{\frac{n+1}{2}}}, \quad y \in \mathbb{R}$$

Exemplo 3

Seja T_n uma variável aleatória com distribuição t-Student com n graus de liberdade, ou seja, a densidade de T_n é dada por:

$$f_{T_n}(y) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)} \frac{1}{\left(1 + \frac{y^2}{n}\right)^{\frac{n+1}{2}}}, \quad y \in \mathbb{R}$$

Temos que,

$$\lim_{n \rightarrow +\infty} F_{T_n}(t) = \lim_{n \rightarrow +\infty} \int_{-\infty}^t \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)} \frac{1}{\left(1 + \frac{y^2}{n}\right)^{\frac{n+1}{2}}} dy$$

$$\lim_{n \rightarrow +\infty} F_{T_n}(t) = \lim_{n \rightarrow +\infty} \int_{-\infty}^t \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)} \frac{1}{\left(1 + \frac{y^2}{n}\right)^{\frac{n+1}{2}}} dy$$

$$= \int_{-\infty}^t \lim_{n \rightarrow +\infty} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)} \frac{1}{\left(1 + \frac{y^2}{n}\right)^{\frac{n+1}{2}}} dy$$

Teorema da
Convergência Dominada

= ★★

Considere a seguinte aproximação de Stirling (Conhecida como fórmula de Stirling):

$$\Gamma(t+1) \approx \sqrt{2\pi t} \left(\frac{t}{e}\right)^t$$

Ou seja,

$$\lim_{t \rightarrow +\infty} \frac{\Gamma(t+1)}{\sqrt{2t\pi} \left(\frac{t}{e}\right)^t} = 1$$

Considere a seguinte aproximação de Stirling (Conhecida como fórmula de Stirling):

$$\Gamma(t+1) \approx \sqrt{2\pi t} \left(\frac{t}{e}\right)^t$$

Ou seja,

$$\lim_{t \rightarrow +\infty} \frac{\Gamma(t+1)}{\sqrt{2t\pi} \left(\frac{t}{e}\right)^t} = 1$$

Logo,

$$\begin{aligned} \lim_{n \rightarrow +\infty} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)} \frac{1}{\left(1 + \frac{y^2}{n}\right)^{\frac{n+1}{2}}} &= \lim_{n \rightarrow +\infty} \frac{\sqrt{2\pi} \left(\frac{n-1}{2}\right)^{\frac{n-1}{2} + \frac{1}{2}} e^{-\left(\frac{n-1}{2}\right)}}{\sqrt{n}\sqrt{2\pi} \left(\frac{n-2}{2}\right)^{\frac{n-2}{2} + \frac{1}{2}} e^{-\left(\frac{n-2}{2}\right)}} \frac{1}{\left(1 + \frac{y^2}{n}\right)^{\frac{n+1}{2}}} \\ &= \star \left(\text{t da fórmula de Stirling será } \frac{n-1}{2} \right) \end{aligned}$$

$$\begin{aligned}
 \star &= \frac{e^{-\frac{1}{2}}}{\sqrt{2\pi}} \lim_{n \rightarrow \infty} \frac{(n-1)^{\frac{n}{2}}}{(n-2)^{\frac{n}{2}}(n-2)^{-\frac{1}{2}}} \frac{1}{\sqrt{n}} e^{-\frac{y^2}{2}} \\
 &= \frac{e^{-\frac{1}{2}}}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n-2}\right)^{\frac{n}{2}} \\
 &= \frac{e^{-\frac{1}{2}}}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} e^{\frac{1}{2}} = \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}
 \end{aligned}$$

$$\begin{aligned}
 \star &= \frac{e^{-\frac{1}{2}}}{\sqrt{2\pi}} \lim_{n \rightarrow \infty} \frac{(n-1)^{\frac{n}{2}}}{(n-2)^{\frac{n}{2}} (n-2)^{-\frac{1}{2}}} \frac{1}{\sqrt{n}} e^{-\frac{y^2}{2}} \\
 &= \frac{e^{-\frac{1}{2}}}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n-2}\right)^{\frac{n}{2}} \\
 &= \frac{e^{-\frac{1}{2}}}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} e^{\frac{1}{2}} = \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}
 \end{aligned}$$

Portanto, substituindo em $\star\star$, temos

$$\lim_{n \rightarrow +\infty} F_{T_n}(t) = \int_{-\infty}^t \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy$$

Logo, $T_n \xrightarrow{D} N(0, 1)$.

Referências I

HOGG, RV; MCKEAN, J; CRAIG, AT. **Introduction to Mathematical Statistics**. Eighth Edition. [S.l.]: Pearson, 2019.