

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 Teorema Central do Limite

2 Método Delta

O Teorema Central do Limite (TCL) é um pilar fundamental na Estatística, Probabilidade e Ciência de Dados, pois estabelece que a soma de um grande número de variáveis aleatórias independentes e identicamente distribuídas tende a se aproximar de uma distribuição normal, independentemente da forma original da distribuição dessas variáveis. Este teorema não apenas justifica a aplicação de métodos estatísticos que assumem normalidade, mas também permite a utilização de técnicas analíticas e inferenciais robustas para uma ampla gama de problemas práticos.

Introdução

Em Estatística, o TCL fundamenta a construção de intervalos de confiança e a realização de testes de hipóteses, possibilitando a generalização de resultados e a tomada de decisões com base em amostras. Na Ciência de Dados, ele é crucial para a modelagem e previsão, facilitando a análise de grandes volumes de dados e a implementação de algoritmos que dependem de pressupostos normais. Em suma, o Teorema Central do Limite não só proporciona uma base teórica sólida, mas também une teoria e prática, permitindo a solução eficaz de problemas complexos e a interpretação confiável de resultados em diversas disciplinas.

Teorema Central do Limite

Seja $\{X_n\}_{n \geq 1}$ uma sequência de variáveis aleatórias iid com média μ e variância $\sigma^2 < \infty$. Então,

$$\frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{D} N(0, 1)$$

Demonstração

Assuma, sem perda de generalidade, $\mu = 0$ e $\sigma^2 = 1$.

$$\begin{aligned}M_{\frac{\sum X_i - n\mu}{\sqrt{n}\sigma}}(t) &= M_{\frac{\sum X_i}{\sqrt{n}}}(t) = E\left(e^{\frac{t \sum X_i}{\sqrt{n}}}\right) = E\left(\prod_{i=1}^n e^{\frac{tX_i}{\sqrt{n}}}\right) \\&= \prod_{i=1}^n E\left(e^{\frac{tX_i}{\sqrt{n}}}\right) = \prod_{i=1}^n M_{X_i}\left(\frac{t}{\sqrt{n}}\right) = \left(M_{X_i}\left(\frac{t}{\sqrt{n}}\right)\right)^n\end{aligned}$$

Demonstração

Assuma, sem perda de generalidade, $\mu = 0$ e $\sigma^2 = 1$.

$$\begin{aligned} M_{\frac{\sum X_i - n\mu}{\sqrt{n}\sigma}}(t) &= M_{\frac{\sum X_i}{\sqrt{n}}}(t) = E\left(e^{\frac{t \sum X_i}{\sqrt{n}}}\right) = E\left(\prod_{i=1}^n e^{\frac{tX_i}{\sqrt{n}}}\right) \\ &= \prod_{i=1}^n E\left(e^{\frac{tX_i}{\sqrt{n}}}\right) = \prod_{i=1}^n M_{X_i}\left(\frac{t}{\sqrt{n}}\right) = \left(M_{X_i}\left(\frac{t}{\sqrt{n}}\right)\right)^n \end{aligned}$$

Notemos que, $\ln M_{\frac{\sum X_i}{\sqrt{n}}}(t) = n \ln M_{X_i}\left(\frac{t}{\sqrt{n}}\right) = \frac{\ln M_{X_i}\left(\frac{t}{\sqrt{n}}\right)}{\frac{1}{n}}$. Além disso, sabemos que $\lim_{n \rightarrow \infty} \ln[f(n)] = \ln \lim_{n \rightarrow \infty} f(n)$, sempre que $\lim_{n \rightarrow \infty} f(n) > 0$ e que para n grande $\frac{t}{\sqrt{n}} \approx 0$ e $M_{X_i}(0) = 1$.

Aplicando L'Hôpital, temos:

$$\begin{aligned}\lim_{n \rightarrow \infty} \frac{\ln M_{X_i}\left(\frac{t}{\sqrt{n}}\right)}{\frac{1}{n}} &= \lim_{n \rightarrow \infty} \frac{\frac{M'_{X_i}\left(\frac{t}{\sqrt{n}}\right)}{M_{X_i}\left(\frac{t}{\sqrt{n}}\right)} t \left(-\frac{1}{2}\right) n^{-\frac{3}{2}}}{-\frac{1}{n^2}} \\ &= \frac{t}{2} \lim_{n \rightarrow \infty} \sqrt{n} \frac{M'_{X_i}\left(\frac{t}{\sqrt{n}}\right)}{M_{X_i}\left(\frac{t}{\sqrt{n}}\right)} \\ &= \star\end{aligned}$$

Sabe-se que o k -ésimo momento de X_i é dado por $M_{X_i}^{(k)}(0) = E(X_i^k)$, ou seja,

$$M_{X_i}^{(')}(0) = E(X_i) = \mu = 0 \text{ e } M_{X_i}^{('')} (0) = E(X_i^2) = \sigma^2 = 1 \text{ (por hipótese)}$$

Logo,

$$\begin{aligned}\star &= \frac{t}{2} \lim_{n \rightarrow \infty} \sqrt{n} \frac{M'_{X_i}\left(\frac{t}{\sqrt{n}}\right)}{M_{X_i}\left(\frac{t}{\sqrt{n}}\right)} \\ &= \frac{t}{2} \lim_{n \rightarrow \infty} \frac{\frac{M'_{X_i}\left(\frac{t}{\sqrt{n}}\right)}{M_{X_i}\left(\frac{t}{\sqrt{n}}\right)}}{\frac{1}{\sqrt{n}}}\end{aligned}$$

Logo,

$$\begin{aligned}\star &= \frac{t}{2} \lim_{n \rightarrow \infty} \sqrt{n} \frac{M'_{X_i}\left(\frac{t}{\sqrt{n}}\right)}{M_{X_i}\left(\frac{t}{\sqrt{n}}\right)} \\ &= \frac{t}{2} \lim_{n \rightarrow \infty} \frac{\frac{M'_{X_i}\left(\frac{t}{\sqrt{n}}\right)}{M_{X_i}\left(\frac{t}{\sqrt{n}}\right)}}{\frac{1}{\sqrt{n}}}\end{aligned}$$

aplicando L'Hôpital novamente, temos:

$$\star = \frac{t}{2} \lim_{n \rightarrow \infty} \frac{\frac{M_{X_i}\left(\frac{t}{\sqrt{n}}\right)M''_{X_i}\left(\frac{t}{\sqrt{n}}\right)\left(-\frac{1}{2}tn^{-\frac{3}{2}}\right) - \left(M'_{X_i}\left(\frac{t}{\sqrt{n}}\right)\right)^2\left(\frac{1}{2}tn^{-\frac{3}{2}}\right)}{\left(M_{X_i}\left(\frac{t}{\sqrt{n}}\right)\right)^2 - \left(\frac{1}{2}\right)n^{-\frac{3}{2}}}}$$

Simplificando os termos $-(\frac{1}{2})n^{-\frac{3}{2}}$ e observando que para n suficientemente grande, $\frac{t}{\sqrt{n}} \approx 0 \Rightarrow M_{X_i}(0) = 1, M_{X_i}''(0) = 0$ e $M_{X_i}'''(0) = 1$, temos:

$$\frac{t}{2} \lim_{n \rightarrow \infty} \frac{\frac{M_{X_i}(\frac{t}{\sqrt{n}})M_{X_i}''(\frac{t}{\sqrt{n}})(-\frac{1}{2}tn^{-\frac{3}{2}}) - (M_{X_i}'(\frac{t}{\sqrt{n}}))^2(\frac{1}{2}tn^{-\frac{3}{2}})}{\left(M_{X_i}(\frac{t}{\sqrt{n}})\right)^2 - (\frac{1}{2})n^{-\frac{3}{2}}}}{ } = \frac{t^2}{2}$$

Simplificando os termos $-(\frac{1}{2})n^{-\frac{3}{2}}$ e observando que para n suficientemente grande, $\frac{t}{\sqrt{n}} \approx 0 \Rightarrow M_{X_i}(0) = 1, M_{X_i}^{(')}(0) = 0$ e $M_{X_i}^{('')} (0) = 1$, temos:

$$\frac{t}{2} \lim_{n \rightarrow \infty} \frac{\frac{M_{X_i}(\frac{t}{\sqrt{n}})M_{X_i}^{('')}(\frac{t}{\sqrt{n}})(-\frac{1}{2}tn^{-\frac{3}{2}}) - (M_{X_i}'(\frac{t}{\sqrt{n}}))^2(\frac{1}{2}tn^{-\frac{3}{2}})}{\left(M_{X_i}(\frac{t}{\sqrt{n}})\right)^2 - (\frac{1}{2})n^{-\frac{3}{2}}}} = \frac{t^2}{2}$$

Portanto, $\lim_{n \rightarrow \infty} M_{\frac{\sum X_i}{\sqrt{n}}}(t) = \exp \lim_{n \rightarrow \infty} \ln M_{\frac{\sum X_i}{\sqrt{n}}}(t) = \exp\left(\frac{t^2}{2}\right)$, ou seja,

$$\frac{\sum X_i}{\sqrt{n}} \xrightarrow{D} X \sim N(0, 1).$$

Método Delta

Suponha que conhecemos a distribuição de uma variável aleatória, mas que queremos determinar a distribuição de uma função dela. Isso também é verdade na teoria assintótica, o teorema de Slutsky's e o teorema visto em aula, imediatamente anterior a ele, são ilustrações disso. Outro resultado desse tipo é chamado de método delta. Vejamos o próximo slide!

Teorema 1

Seja $\{X_n\}_{n \geq 1}$ uma sequência de variáveis aleatórias satisfazendo a seguinte convergência em distribuição:

$$\sqrt{n}(X_n - \theta) \xrightarrow{D} N(0, \sigma^2).$$

Se g é uma função diferenciável em θ e $g'(\theta) \neq 0$, então

$$\sqrt{n}(g(X_n) - g(\theta)) \xrightarrow{D} N(0, [g'(\theta)]^2 \sigma^2).$$

Demonstração

Vamos mostrar primeiro que $X_n \xrightarrow{P} \theta$. Seja $\varepsilon > 0$ e $m > 0$ inteiro ($m \in \mathbb{N}^*$) fixado.

$$\begin{aligned} P(|X_n - \theta| < \varepsilon) &= P(|\sqrt{n}(X_n - \theta)| < \varepsilon\sqrt{n}) \\ &\geq P(|\sqrt{n}(X_n - \theta)| < m) \end{aligned}$$

$$\text{para } n \geq \left(\frac{m}{\varepsilon}\right)^2$$

Segue que,

$$\begin{aligned}\liminf_{n \rightarrow \infty} P(|X_n - \theta| < \varepsilon) &\geq \liminf_{n \rightarrow \infty} P(|\sqrt{n}(X_n - \theta)| < m) \\ &= \lim_{n \rightarrow \infty} P(|\sqrt{n}(X_n - \theta)| < m) \\ &= P\left(|Z| < \frac{m}{\sigma}\right), \quad Z \sim N(0, 1)\end{aligned}$$

Usando o fato de que
 $\sqrt{n}(X_n - \theta) \xrightarrow{D} N(0, \sigma^2)$

Segue que,

$$\begin{aligned}\liminf_{n \rightarrow \infty} P(|X_n - \theta| < \varepsilon) &\geq \liminf_{n \rightarrow \infty} P(|\sqrt{n}(X_n - \theta)| < m) \\ &= \lim_{n \rightarrow \infty} P(|\sqrt{n}(X_n - \theta)| < m) \\ &= P\left(|Z| < \frac{m}{\sigma}\right), \quad Z \sim N(0, 1)\end{aligned}$$

Usando o fato de que
 $\sqrt{n}(X_n - \theta) \xrightarrow{D} N(0, \sigma^2)$

Assim,

$$\begin{aligned}\liminf_{n \rightarrow \infty} P(|X_n - \theta| < \varepsilon) &\geq P\left(|Z| < \frac{m}{\sigma}\right) = \Phi\left(\frac{m}{\sigma}\right) - \Phi\left(-\frac{m}{\sigma}\right) \\ &= 2\Phi\left(\frac{m}{\sigma}\right) - 1\end{aligned}$$

Em resumo,

$$\liminf_{n \rightarrow \infty} P(|X_n - \theta| < \varepsilon) \geq 2\Phi\left(\frac{m}{\sigma}\right) - 1, \quad \forall m \in \mathbb{Z}^*$$

$$\begin{aligned} \liminf_{n \rightarrow \infty} P(|X_n - \theta| < \varepsilon) &\geq \lim_{m \rightarrow +\infty} \left[2\Phi\left(\frac{m}{\sigma}\right) - 1 \right] \\ &= 1 \end{aligned}$$

Em resumo,

$$\begin{aligned}\liminf_{n \rightarrow \infty} P(|X_n - \theta| < \varepsilon) &\geq 2\Phi\left(\frac{m}{\sigma}\right) - 1, \quad \forall m \in \mathbb{Z}^* \\ \liminf_{n \rightarrow \infty} P(|X_n - \theta| < \varepsilon) &\geq \lim_{m \rightarrow +\infty} \left[2\Phi\left(\frac{m}{\sigma}\right) - 1 \right] \\ &= 1\end{aligned}$$

Como $\limsup \geq \liminf$, temos que

$$\limsup P(|X_n - \theta| < \varepsilon) = 1 \Rightarrow \lim_{n \rightarrow \infty} P(|X_n - \theta| < \varepsilon) = 1 \quad (1)$$

Em resumo,

$$\begin{aligned}\liminf_{n \rightarrow \infty} P(|X_n - \theta| < \varepsilon) &\geq 2\Phi\left(\frac{m}{\sigma}\right) - 1, \quad \forall m \in \mathbb{Z}^* \\ \liminf_{n \rightarrow \infty} P(|X_n - \theta| < \varepsilon) &\geq \lim_{m \rightarrow +\infty} \left[2\Phi\left(\frac{m}{\sigma}\right) - 1 \right] \\ &= 1\end{aligned}$$

Como $\limsup \geq \liminf$, temos que

$$\limsup P(|X_n - \theta| < \varepsilon) = 1 \Rightarrow \lim_{n \rightarrow \infty} P(|X_n - \theta| < \varepsilon) = 1 \quad (1)$$

Ou seja, $X_n \xrightarrow{P} \theta$.cqdd ■

Série de Taylor

Agora, lembre-se que uma série de Taylor é a série de funções da forma:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x - a)^n,$$

Neste caso, a série acima é dita ser a série de Taylor de $f(x)$ em torno do ponto $x = a$. Em que $f^{(n)}(a)$ representa a n -ésima derivada de f calculada em a .

Série de Taylor

Agora, lembre-se que uma série de Taylor é a série de funções da forma:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x - a)^n,$$

Neste caso, a série acima é dita ser a série de Taylor de $f(x)$ em torno do ponto $x = a$. Em que $f^{(n)}(a)$ representa a n -ésima derivada de f calculada em a .

Expandindo g em série de Taylor até a primeira ordem em torno de θ , temos que

$$g(x) = g(\theta) + g'(\theta)(x - \theta) + \mathcal{C}(x)(x - \theta),$$

em que $\lim_{x \rightarrow \theta} \mathcal{C}(x) = 0$.

Segue que,

$$\begin{aligned}g(x) - g(\theta) &= g'(\theta)(x - \theta) + \mathcal{C}(x)(x - \theta) \\ \sqrt{n}(g(X_n) - g(\theta)) &= \sqrt{n}(X_n - \theta)(g'(\theta) + \mathcal{C}(X_n)).\end{aligned}$$

Para mostrar o resultado desejado, devemos mostrar que $\mathcal{C}(X_n) \xrightarrow{P} 0$. Com isso, utilizando o teorema de Slutsky, o resultado é obtido. Como $\lim_{x \rightarrow \theta} \mathcal{C}(x) = 0$, para $\varepsilon > 0$, $\exists \delta > 0$ tal que $|x - \theta| < \delta \Rightarrow |\mathcal{C}(x)| < \varepsilon$.

Ora,

- $\{|X_n - \theta| < \delta\}$: Este evento ocorre quando a variável aleatória X_n está suficientemente próxima de θ , ou seja, dentro de uma faixa de tamanho δ ao redor de θ .
- $\{|\mathcal{C}(X_n)| < \varepsilon\}$: Este evento ocorre quando o valor absoluto de $\mathcal{C}(X_n)$ está dentro de uma faixa de tamanho ε .

Como a função $\mathcal{C}(x)$ tende a 0 quando x tende a θ ($\lim_{x \rightarrow \theta} \mathcal{C}(x) = 0$), para qualquer $\varepsilon > 0$, existe um $\delta > 0$ tal que, se x estiver suficientemente próximo de θ (ou seja, se $|x - \theta| < \delta$), então $|\mathcal{C}(x)| < \varepsilon$.

Dessa forma, se x pertence ao conjunto $\{|x - \theta| < \delta\}$, isso implica que x também pertence ao conjunto $\{|\mathcal{C}(x)| < \varepsilon\}$, ou seja:

$$\{|X_n - \theta| < \delta\} \subseteq \{|\mathcal{C}(X_n)| < \varepsilon\}$$

Daí,

$$\{|X_n - \theta| < \delta\} \subset \{|\mathcal{C}(X_n)| < \varepsilon\} \Rightarrow P(|X_n - \theta| < \delta) \leq P(|\mathcal{C}(X_n)| < \varepsilon)$$

Esse resultado, juntamente com (1) implica que,

$$\lim_{n \rightarrow \infty} P(|\mathcal{C}(X_n)| < \varepsilon) = 1, \text{ ou seja, } \mathcal{C}(X_n) \xrightarrow{P} 0.$$

Daí,

$$\{|X_n - \theta| < \delta\} \subset \{|\mathcal{C}(X_n)| < \varepsilon\} \Rightarrow P(|X_n - \theta| < \delta) \leq P(|\mathcal{C}(X_n)| < \varepsilon)$$

Esse resultado, juntamente com (1) implica que,

$$\lim_{n \rightarrow \infty} P(|\mathcal{C}(X_n)| < \varepsilon) = 1, \text{ ou seja, } \mathcal{C}(X_n) \xrightarrow{P} 0.$$

Portanto, como por hipótese, $\sqrt{n}(X_n - \theta) \xrightarrow{D} N(0, \sigma^2)$, temos que

$$\sqrt{n}(g(X_n) - g(\theta)) = \sqrt{n}(X_n - \theta)g'(\theta) \xrightarrow{D} N(0, [g'(\theta)]^2 \sigma^2)$$

Para 

Seja $X_i \stackrel{\text{iid}}{\sim} \text{Binomial}(1, p)$, $i = 1, \dots, n$. Mostre que

$$\sqrt{n}(\arcsen\sqrt{\bar{X}} - \arcsen\sqrt{p}) \xrightarrow{D} N(0, \frac{1}{4}), \text{ em que } \bar{X} = \sum_{i=1}^n \frac{X_i}{n}$$

Para

Exercícios 5.3.1 à 5.3.8, 5.3.11, 5.3.12

Referências I



Casella, George e Roger L Berger (2021). *Statistical inference*. Cengage Learning.



Hogg, RV, J McKean e AT Craig (2019). *Introduction to Mathematical Statistics*.