Material Auxiliar - Unidade 2

Convergência Estocástica e Resultados Limite

Explicações Detalhadas e Didáticas

Curso de Inferência Estatística

Outubro 2025

Sumário

1 Introdução

Este material auxiliar complementa as notas de aula da Unidade 2, fornecendo explicações mais detalhadas e didáticas dos principais conceitos abordados. O objetivo é facilitar a compreensão dos teoremas de convergência e suas aplicações práticas.

2 Notação O(·) e o(·) - Big O e Little o

2.1 Motivação e Intuição

A notação $O(\cdot)$ e $o(\cdot)$ é fundamental para descrever o comportamento assintótico de sequências e funções. Intuitivamente:

- $a_n = O(b_n)$: "a_n cresce *no máximo* tão rápido quanto b_n"
- $a_n = o(b_n)$: "a_n cresce *mais devagar* que b_n"

2.2 Definições Formais

Definição 2.1 (Big O para sequências)**.** Sejam $\{a_n, n \geq 1\}$ e $\{b_n, n \geq 1\}$ sequências de números reais. Dizemos que

$$a_n = O(b_n) \quad \text{se e somente se} \quad \exists\, k > 0,\, n_0 \in \mathbb{N} : \left|\frac{a_n}{b_n}\right| \leq k, \quad \forall n \geq n_0$$

Isto é, a razão $|a_n/b_n|$ é limitada para n suficientemente grande.

Definição 2.2 (Little o para sequências)**.** Dizemos que

$$a_n = o(b_n) \quad \text{se e somente se} \quad \lim_{n\to\infty} \frac{a_n}{b_n} = 0$$

Isto é, a_n é desprezível comparado a b_n quando n é grande.

Exemplo 2.1 (Comparações Comuns)**.**
1. $10n^2+n = O(n^2)$ porque $\frac{10n^2+n}{n^2} = 10+\frac{1}{n} \leq 11$ para $n \geq 1$
2. $n = o(n^2)$ porque $\lim_{n\to\infty} \frac{n}{n^2} = \lim_{n\to\infty} \frac{1}{n} = 0$
3. $\log(n) = o(n)$ porque $\lim_{n\to\infty} \frac{\log(n)}{n} = 0$
4. $n^{1/2} = O(n)$ mas $n \neq O(n^{1/2})$

2.3 Propriedades Importantes

Observação 2.1 (Álgebra de O e o)**.**
1. *Se $a_n = o(b_n)$, então $a_n = O(b_n)$ (mas a recíproca é falsa)*
2. *Se $a_n = O(b_n)$ e $c_n = O(d_n)$, então:*
   - $a_n \cdot c_n = O(b_n \cdot d_n)$
   - $a_n + c_n = O(\max\{|b_n|, |d_n|\})$
3. *$O(1)$ significa limitado: $|a_n| \leq k$ para algum $k > 0$ e n grande*
4. *$o(1)$ significa que $a_n \to 0$*

2.4 Aplicação em Séries de Taylor

A notação $O(\cdot)$ é essencial para expressar aproximações via série de Taylor:

Exemplo 2.2 (Série de Taylor)**.** Para uma função $F(x)$ derivável até ordem n em torno de x_0:

$$F(x) = \sum_{k=0}^{n} \frac{F^{(k)}(x_0)}{k!}(x - x_0)^k + o\left((x - x_0)^n\right)$$

quando $x \to x_0$.

Por exemplo:

- $e^x = 1 + x + \frac{x^2}{2} + O(x^3)$ quando $x \to 0$
- $\log(1 + x) = x - \frac{x^2}{2} + O(x^3)$ quando $x \to 0$

3 Convergência em Probabilidade

3.1 Intuição e Definição

A convergência em probabilidade expressa a ideia de que, à medida que n cresce, a probabilidade de U_n estar "longe"de u torna-se arbitrariamente pequena.

Definição 3.1 (Convergência em Probabilidade)**.** Uma sequência de variáveis aleatórias $\{U_n, n \geq 1\}$ converge em probabilidade para um número u se

$$P\left(|U_n - u| \geq \varepsilon\right) \xrightarrow{n\to\infty} 0, \quad \forall\, \varepsilon > 0$$

Notação: $U_n \xrightarrow{P} u$

3.2 Interpretação Prática

Pense em U_n como uma estimativa de u baseada em n observações. Convergência em probabilidade significa que:

- Com n grande, é *altamente improvável* que U_n esteja longe de u
- Para qualquer margem de erro $\varepsilon > 0$ que você escolha, a probabilidade de erro pode ser tornada arbitrariamente pequena aumentando n

Exemplo 3.1 (Média Amostral)**.** Se $X_1, X_2, \ldots$ são v.a.'s i.i.d. com $E[X_i] = \mu$ e $\text{Var}(X_i) = \sigma^2 < \infty$, então

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu$$

Isso significa que a média amostral converge para a média populacional.

3.3 Métodos para Provar Convergência em Probabilidade

1. **Desigualdade de Chebyshev:** Se $E[U_n] \to u$ e $\text{Var}(U_n) \to 0$, então $U_n \xrightarrow{P} u$
2. **Convergência de momentos:** Se $E[|U_n - u|^r] \to 0$ para algum $r > 0$, então $U_n \xrightarrow{P} u$
3. **Função geradora de momentos:** Se $M_{U_n}(t) \to e^{tu}$ para todo t, então $U_n \xrightarrow{P} u$

3.4 Propriedades Algébricas

Observação 3.1 (Álgebra da Convergência em Probabilidade)**.** *Se* $U_n \xrightarrow{P} u$ *e* $V_n \xrightarrow{P} v$*, então:*

1. $U_n + V_n \xrightarrow{P} u + v$
2. $U_n \cdot V_n \xrightarrow{P} u \cdot v$
3. $\frac{U_n}{V_n} \xrightarrow{P} \frac{u}{v}$ *(se* $P(V_n = 0) = 0$ *e* $v \neq 0$*)*
4. *Se* $g(\cdot)$ *é contínua, então* $g(U_n) \xrightarrow{P} g(u)$

4 Convergência em Distribuição

4.1 Definição e Diferenças

A convergência em distribuição é um conceito mais fraco que convergência em probabilidade.

Definição 4.1 (Convergência em Distribuição)**.** Uma sequência $\{U_n, n \geq 1\}$ com f.d.a. $F_n(u)$ converge em distribuição para uma v.a. U com f.d.a. $F(u)$ se

$$F_n(u) \xrightarrow{n\to\infty} F(u)$$

em todos os pontos de continuidade de $F(\cdot)$.

Notação: $U_n \xrightarrow{D} U$

4.2 Diferenças entre Convergências

- **Convergência em Probabilidade** $\Rightarrow$ **Convergência em Distribuição**
- **Convergência em Distribuição** $\not\Rightarrow$ **Convergência em Probabilidade** (em geral)
- **Exceção:** Se $U_n \xrightarrow{D} c$ (constante), então $U_n \xrightarrow{P} c$

Exemplo 4.1 (Distinção Importante)**.** Considere $X \sim N(0,1)$ e defina $U_n = X$ para todo n. Então:

- $U_n \xrightarrow{D} X$ (trivialmente, pois $F_n = F$ para todo n)
- $U_n \overset{P}{\not\to} X$ (não faz sentido: $U_n - X = 0$ sempre!)

Agora considere $U_n = (-1)^n X$:

- $U_n \xrightarrow{D} X$ (ambos têm distribuição $N(0,1)$)
- $U_n \overset{P}{\not\to} X$ (pois $|U_n - X|$ não vai para zero)

4.3 Método da Função Geradora de Momentos

Um método poderoso para provar convergência em distribuição:

Observação 4.1 (Teorema de Continuidade de Lévy)**.** *Se $M_{U_n}(t) \to M_U(t)$ para todo t em uma vizinhança de zero, então $U_n \xrightarrow{D} U$.*

Este método é frequentemente usado nas provas do TCL.

5 Lei Fraca dos Grandes Números

5.1 Versões e Interpretação

A Lei Fraca dos Grandes Números (LFGN) é um dos resultados fundamentais da probabilidade.

Observação 5.1 (LFGN - Versão Simples)**.** *Se $X_1, \dots, X_n$ são v.a.'s i.i.d. com $E[X_i] = \mu < \infty$ e $Var(X_i) = \sigma^2 < \infty$, então*

$$\bar{X}_n \xrightarrow{P} \mu$$

Observação 5.2 (LFGN de Khinchin)**.** *A condição de variância finita pode ser relaxada: basta $E[X_i] = \mu < \infty$.*

5.2 Interpretação Prática

- A média amostral é um estimador *consistente* da média populacional
- Quanto maior a amostra, mais confiável é a estimativa
- Justifica a "Lei dos Grandes Números"empírica: frequências relativas convergem para probabilidades

Exemplo 5.1 (Lançamento de Moedas)**.** Se $X_i = 1$ (cara) ou $X_i = 0$ (coroa) com $P(X_i = 1) = p$, então

$$\frac{\text{número de caras em } n \text{ lançamentos}}{n} = \bar{X}_n \xrightarrow{P} p$$

5.3 Aplicações

1. **Estimação de parâmetros:** $\bar{X}_n$ estima μ, S_n^2 estima σ^2
2. **Simulação Monte Carlo:** Aproximar $E[g(X)]$ por $\frac{1}{n}\sum_{i=1}^{n} g(X_i)$
3. **Testes de hipóteses:** Proporções amostrais convergem para proporções populacionais

6 Teorema Central do Limite

6.1 Enunciado e Importância

O Teorema Central do Limite (TCL) é possivelmente o teorema mais importante da estatística.

Observação 6.1 (TCL - Versão Clássica). *Se $X_1, \ldots, X_n$ são v.a.'s i.i.d. com $E[X_i] = \mu$ e $Var(X_i) = \sigma^2 < \infty$, então*

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

6.2 Por Que é Tão Importante?

1. **Universalidade:** Funciona para *qualquer* distribuição com variância finita
2. **Base para inferência:** Justifica o uso da distribuição normal em intervalos de confiança e testes
3. **Aproximação prática:** Com n moderadamente grande ($n \geq 30$), $\bar{X}_n$ tem distribuição aproximadamente normal

6.3 Interpretação Geométrica

O TCL diz que:

- A distribuição de $\bar{X}_n$ fica mais concentrada em torno de μ (taxa $1/\sqrt{n}$)
- A forma da distribuição de $\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$ converge para a curva normal
- Não importa a distribuição original dos X_i!

Exemplo 6.1 (Distribuição Uniforme). Se $X_i \sim U(0,1)$ (distribuição uniforme), então $E[X_i] = 1/2$ e $\text{Var}(X_i) = 1/12$.

$$\frac{\bar{X}_n - 1/2}{\sqrt{1/(12n)}} = \sqrt{12n}\left(\bar{X}_n - \frac{1}{2}\right) \xrightarrow{D} N(0,1)$$

Embora X_i seja uniforme (nada parecido com normal), $\bar{X}_n$ tem distribuição aproximadamente $N(1/2, 1/(12n))$ para n grande.

6.4 Versões Padronizadas

- **σ conhecido:** $Z_n = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{D} N(0,1)$
- **σ desconhecido:** $T_n = \frac{\sqrt{n}(\bar{X}_n - \mu)}{S_n} \xrightarrow{D} N(0,1)$
  (onde S_n é o desvio padrão amostral)

7 Teorema de Slutsky

7.1 Enunciado e Utilidade

O Teorema de Slutsky permite combinar convergências de tipos diferentes.

Observação 7.1 (Teorema de Slutsky). *Se $U_n \xrightarrow{D} U$ e $V_n \xrightarrow{P} c$ (constante), então:*

1. $U_n + V_n \xrightarrow{D} U + c$
2. $U_n \cdot V_n \xrightarrow{D} c \cdot U$
3. $\frac{U_n}{V_n} \xrightarrow{D} \frac{U}{c}$ *(se $c \neq 0$)*

7.2 Por Que é Útil?

O teorema de Slutsky é crucial quando:

- Temos uma convergência em distribuição mas precisamos fazer operações algébricas
- Queremos substituir parâmetros desconhecidos por estimadores consistentes
- Provamos distribuições assintóticas de estatísticas de teste

Exemplo 7.1 (Substituição do Desvio Padrão). Pelo TCL: $\frac{\sqrt{n}(\bar{X}_n-\mu)}{\sigma} \xrightarrow{D} N(0,1)$

Como $S_n \xrightarrow{P} \sigma$, pelo Slutsky:

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{S_n} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \cdot \frac{\sigma}{S_n} \xrightarrow{D} N(0,1) \cdot 1 = N(0,1)$$

Isso justifica usar S_n quando σ é desconhecido!

7.3 Aplicação em Testes de Hipóteses

O teorema de Slutsky permite construir estatísticas de teste quando parâmetros são desconhecidos, substituindo-os por estimadores consistentes sem alterar a distribuição assintótica.

8 Teorema de Mann-Wald (Método Delta)

8.1 Enunciado

O Método Delta é uma ferramenta para encontrar a distribuição assintótica de transformações de estimadores.

Observação 8.1 (Teorema de Mann-Wald). *Se $\sqrt{n}(T_n - \theta) \xrightarrow{D} N(0, \sigma^2(\theta))$ e $g(\cdot)$ é uma função diferenciável com $g'(\theta) \neq 0$, então*

$$\sqrt{n}\,[g(T_n) - g(\theta)] \xrightarrow{D} N\left(0, \sigma^2(\theta) \cdot [g'(\theta)]^2\right)$$

8.2 Interpretação

O método delta diz que:

- Se T_n é aproximadamente normal com taxa $1/\sqrt{n}$
- Então $g(T_n)$ também é aproximadamente normal com taxa $1/\sqrt{n}$
- A variância é "inflada"por $[g'(\theta)]^2$

8.3 Ideia da Prova

A prova usa aproximação de Taylor de primeira ordem:

$$g(T_n) \approx g(\theta) + g'(\theta)(T_n - \theta)$$

Multiplicando por $\sqrt{n}$:

$$\sqrt{n}[g(T_n) - g(\theta)] \approx g'(\theta) \cdot \sqrt{n}(T_n - \theta)$$

Como $\sqrt{n}(T_n - \theta) \xrightarrow{D} N(0, \sigma^2)$, o resultado segue.

Exemplo 8.1 (Transformação Logarítmica)**.** Se $\bar{X}_n$ estima $\mu > 0$ e queremos estimar $\log(\mu)$, tome $g(x) = \log(x)$.

Como $g'(x) = 1/x$, temos:

$$\sqrt{n}\left[\log(\bar{X}_n) - \log(\mu)\right] \xrightarrow{D} N\left(0, \frac{\sigma^2}{\mu^2}\right)$$

Exemplo 8.2 (Transformação de Variância)**.** Para estimar a variância σ^2, usamos S_n^2. Se queremos estimar o desvio padrão $\sigma = \sqrt{\sigma^2}$, usamos $g(x) = \sqrt{x}$ com $g'(x) = \frac{1}{2\sqrt{x}}$.

8.4 Aplicações Práticas

1. **Transformações estabilizadoras de variância**
2. **Intervalos de confiança para funções de parâmetros**
3. **Testes de hipóteses sobre transformações**
4. **Modelos não-lineares**

9 Teorema Central do Limite para Variância Amostral

9.1 Motivação

Enquanto o TCL clássico trata da distribuição assintótica de $\bar{X}_n$, é natural perguntar: qual a distribuição assintótica de S_n^2 (a variância amostral)?

Observação 9.1 (TCL para S_n^2)**.** *Se $X_1, \ldots, X_n$ são v.a.'s i.i.d. com $E[X_i] = \mu$, $Var(X_i) = \sigma^2$, e $\mu_4 = E[(X_i - \mu)^4] < \infty$, então*

$$\sqrt{n}(S_n^2 - \sigma^2) \xrightarrow[n\to\infty]{d} N(0, \mu_4 - \sigma^4)$$

9.2 Interpretação

- A variância assintótica é $\mu_4 - \sigma^4$, que depende do quarto momento central
- Para distribuições simétricas, μ_4 mede o "peso nas caudas"
- Distribuições com caudas pesadas têm μ_4 maior, logo maior variabilidade em S_n^2

9.3 Comparação com Normalidade

Para $X_i \sim N(\mu, \sigma^2)$:

- $\mu_4 = 3\sigma^4$, logo a variância assintótica é $3\sigma^4 - \sigma^4 = 2\sigma^4$
- Isto coincide com a variância exata de $\frac{(n-1)S_n^2}{\sigma^2} \sim \chi^2_{n-1}$

9.4 Ideia da Prova

A prova combina o TCL clássico com o Teorema de Slutsky:

1. Defina $W_n = \frac{n-1}{n} S_n^2$ e $Y_i = (X_i - \mu)^2$
2. Mostre que $W_n = \bar{Y}_n - (\bar{X}_n - \mu)^2$
3. Aplique TCL a $\bar{Y}_n$: $\sqrt{n}(\bar{Y}_n - \sigma^2) \xrightarrow{d} N(0, \mu_4 - \sigma^4)$
4. Note que $\sqrt{n}(\bar{X}_n - \mu)^2 \xrightarrow{P} 0$ (é de ordem $O_P(1/\sqrt{n})$)
5. Use Slutsky para concluir sobre W_n, depois relate a S_n^2

9.5 Aplicação Prática

Este teorema permite construir intervalos de confiança assintóticos para σ^2:

$$IC_{1-\alpha}(\sigma^2) = S_n^2 \pm z_{\alpha/2} \cdot \frac{\sqrt{\mu_4 - \sigma^4}}{\sqrt{n}}$$

onde μ_4 pode ser estimado por $\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^4$.

10 Estimadores Consistentes

10.1 Definição e Intuição

Consistência é uma propriedade fundamental de estimadores que garante convergência para o parâmetro verdadeiro.

Definição 10.1 (Estimador Consistente)**.** Um estimador $T_n = T_n(X_1, \ldots, X_n)$ de $\tau(\theta)$ é **consistente** (no sentido fraco) se

$$T_n \xrightarrow[n\to\infty]{P} \tau(\theta), \quad \forall\, \theta \in \Theta$$

10.2 Interpretação Prática

Um estimador consistente significa que:

- Com amostras grandes, a probabilidade de erro significativo torna-se desprezível
- É um requisito mínimo para que um estimador seja considerado "bom"
- Diferente de não-viesamento (propriedade de amostra finita), consistência é assintótica

10.3 Relação entre Viés, Variância e Consistência

Observação 10.1 (Consistência via EQM). *Se $EQM_\theta[T_n] = E_\theta[(T_n - \tau(\theta))^2] \to 0$, então T_n é consistente.*

Como $EQM = Viés^2 + Variância$, temos consistência quando:

$$B_\theta^2[T_n] \to 0 \quad e \quad Var_\theta[T_n] \to 0$$

Exemplo 10.1 (Estimadores Clássicos).

1. $\bar{X}_n$ é consistente para μ (pela LFGN)
2. $S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2$ é consistente para σ^2
3. O EMV é geralmente consistente sob condições de regularidade

10.4 Exemplo Detalhado: Máximo da Uniforme (Questão 3.23)

Exemplo 10.2 (Consistência de $X_{n:n}$ para $U(0,\theta)$). Se $X_1, \ldots, X_n \sim U(0,\theta)$, o EMV de θ é $\hat{\theta} = X_{n:n} = \max\{X_1, \ldots, X_n\}$.

Demonstração completa da consistência:

Passo 1: A f.d.a. de $X_{n:n}$ é:

$$F_{T_n}(t) = P_\theta(T_n \le t) = P_\theta\left\{\bigcap_{i=1}^n X_i \le t\right\} = [F_{X_1}(t)]^n = \begin{cases} 0, & t < 0 \\ \left(\frac{t}{\theta}\right)^n, & 0 \le t \le \theta \\ 1, & t > \theta \end{cases}$$

Passo 2: Para $\varepsilon > 0$:

$$\begin{aligned} P_\theta\{|X_{n:n} - \theta| < \varepsilon\} &= P_\theta\{\theta - \varepsilon < X_{n:n} < \theta + \varepsilon\} && (1) \\ &= P_\theta\{\theta - \varepsilon < X_{n:n} < \theta\} \quad \text{(pois } X_{n:n} \le \theta \text{ sempre)} && (2) \\ &= F_{X_{n:n}}(\theta) - F_{X_{n:n}}(\theta - \varepsilon) && (3) \\ &= \begin{cases} 1, & \varepsilon \ge \theta \\ 1 - \left(\frac{\theta-\varepsilon}{\theta}\right)^n = 1 - \left(1 - \frac{\varepsilon}{\theta}\right)^n, & \varepsilon < \theta \end{cases} && (4) \end{aligned}$$

Passo 3: Quando $n \to \infty$:

$$\lim_{n\to\infty} P_\theta\{|X_{n:n} - \theta| < \varepsilon\} = \lim_{n\to\infty}\left[1 - \left(1 - \frac{\varepsilon}{\theta}\right)^n\right] = 1 - 0 = 1$$

pois $0 < 1 - \frac{\varepsilon}{\theta} < 1$ quando $\varepsilon < \theta$.

Conclusão: Logo, $X_{n:n} \xrightarrow{P} \theta$. $\square$

Observação 10.2 (Observações Importantes sobre Consistência)**.** ***Obs 1:*** *Dados* $\varepsilon > 0$ *e* $\delta \in (0, 1)$*, existe* $n_0 = n_0(\varepsilon, \delta, \theta)$ *tal que:*

$$P_\theta\{|T_n - \theta| > \varepsilon\} \leq \delta \quad \iff \quad P_\theta\{|T_n - \theta| \leq \varepsilon\} \geq 1 - \delta, \quad \forall n \geq n_0$$

Obs 2: T_n *é consistente se, e só se:*

$$\lim_{n\to\infty} P_\theta\{|T_n - \theta| > \varepsilon\} = 0 \quad \text{ou equivalentemente} \quad \lim_{n\to\infty} P_\theta\{|T_n - \theta| \leq \varepsilon\} = 1$$

Obs 3 (Consistência via EQM): $T_n \xrightarrow{P} \theta$ *se* $EQM_\theta[T_n] \xrightarrow{n\to\infty} 0$.
Isto pode ser verificado pela desigualdade de Chebyshev. Para qualquer $\varepsilon > 0$ *e* $\theta \in \Theta$*:*

$$P_\theta\left(|T_n - \theta| > \varepsilon\right) \leq \frac{E_\theta\left[(T_n - \theta)^2\right]}{\varepsilon^2} = \frac{EQM_\theta[T_n]}{\varepsilon^2} \xrightarrow{n\to\infty} 0$$

Como $EQM = \text{Viés}^2 + \text{Variância}$*, se* T_n *é centrado (não-viesado), basta checar:*

$$Var_\theta[T_n] \xrightarrow{n\to\infty} 0$$

10.5 Tamanho Amostral Mínimo

Um aspecto prático interessante: dado $\varepsilon > 0$ e $\delta \in (0, 1)$, qual o tamanho amostral mínimo n_0 para garantir

$$P(|X_{n:n} - \theta| < \varepsilon) \geq 1 - \delta?$$

Solução: De $1 - (1 - \varepsilon/\theta)^n \geq 1 - \delta$, obtemos

$$n \geq \frac{\log \delta}{\log(1 - \varepsilon/\theta)}$$

Assim:

$$n_0 = \left\lceil \frac{\log \delta}{\log(1 - \varepsilon/\theta)} \right\rceil + 1$$

Exemplo 10.3 (Cálculo Numérico)**.** Para $\theta = 1$, $\varepsilon = 0.1$, $\delta = 0.05$:

$$n_0 = \left\lceil \frac{\log(0.05)}{\log(0.9)} \right\rceil + 1 = \left\lceil \frac{-2.996}{-0.105} \right\rceil + 1 = 29$$

Com 29 observações, temos 95% de chance de $X_{n:n}$ estar a menos de 0.1 de $\theta = 1$.

11 Propriedades Assintóticas dos Estimadores de Máxima Verossimilhança

11.1 Introdução

Os EMVs possuem propriedades assintóticas excepcionais que justificam sua popularidade na prática estatística.

11.2 Eficiência Relativa Assintótica

Definição 11.1 (Eficiência Relativa Assintótica)**.** Se dois estimadores $T_n^{(1)}$ e $T_n^{(2)}$ para $g(\theta)$ são assintoticamente normais:

$$\sqrt{n}(T_n^{(i)} - g(\theta)) \xrightarrow[n\to\infty]{d} N(0, \sigma_i^2(\theta)), \quad i = 1, 2$$

a eficiência relativa assintótica (ERA) de $T_n^{(2)}$ em relação a $T_n^{(1)}$ é:

$$\text{ERA} = \frac{\sigma_1^2(\theta)}{\sigma_2^2(\theta)}$$

11.3 Interpretação

- ERA > 1: $T_n^{(2)}$ é mais eficiente (menor variância assintótica)
- ERA $= 1$: Ambos são igualmente eficientes
- ERA < 1: $T_n^{(1)}$ é mais eficiente

11.4 Teorema Central do Limite para EMVs

Observação 11.1 (TCL para EMVs)**.** *Sob condições de regularidade, se $\hat{\theta}_n$ é o EMV de θ, então:*

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow[n\to\infty]{d} N(0, I_X^{-1}(\theta))$$

onde $I_X(\theta)$ é a informação de Fisher:

$$I_X(\theta) = E_\theta\left[\left(\frac{\partial \log f(X;\theta)}{\partial\theta}\right)^2\right]$$

11.5 Condições de Regularidade (Detalhadas)

As condições necessárias para o Teorema 3.8.1 (caso univariado) incluem:

1. **(A1) Diferenciabilidade:** $\theta \mapsto f(x;\theta)$ é três vezes diferenciável sobre Θ, $\forall x \in \mathbb{X}$.

2. **(A2) Troca de derivação e integração:** É válido trocar $\frac{\partial}{\partial\theta}$ com $\int$:

$$\int_{\mathbb{X}} \frac{\partial}{\partial\theta} f(x;\theta)\,dx = \frac{\partial}{\partial\theta}\int_{\mathbb{X}} f(x;\theta)\,dx = 0$$

   e

$$\int_{\mathbb{X}} \frac{\partial^2}{\partial\theta^2} f(x;\theta)\,dx = \frac{\partial^2}{\partial\theta^2}\int_{\mathbb{X}} f(x;\theta)\,dx = 0$$

3. **(A3) Informação de Fisher finita e positiva:**

$$0 < I_X(\theta) \triangleq E_\theta\left[\left(\frac{\partial \log f(x;\theta)}{\partial\theta}\right)^2\right] < \infty, \quad \forall\theta \in \Theta$$

4. **(A4) Dominação local:** Para cada $\theta_0 \in \Theta$, existe $\varepsilon = \varepsilon(\theta_0) > 0$ tal que

$$\left| \frac{\partial^3 \log f(x;\theta)}{\partial \theta^3} \right| \leq g(x), \quad \forall \theta \in [\theta_0 - \varepsilon, \theta_0 + \varepsilon],$$

   em que $\int_X g(x) f(x;\theta)\, dx < \infty$.

5. **(A5) Existência de solução consistente:** A equação de verossimilhança

$$\frac{\partial l(\theta)}{\partial \theta} = 0 \quad \Leftrightarrow \quad \sum_{i=1}^{n} \frac{\partial \log f(x_i;\theta)}{\partial \theta} = 0$$

   tem uma solução consistente $\hat{\theta}_n \xrightarrow{P} \theta$.

Observação 11.2 (Caso Multivariado - Teorema 3.9.2)**.** *Para o caso multivariado ($\theta \in \mathbb{R}^p$), as condições análogas são:*
(B1) *$\mathcal{C} = \{x : f(x;\theta) > 0\}$ independe de θ.*
(B2) *$f(x;\theta)$ é três vezes diferenciável em θ para todo $x \in \mathcal{C}$.*
(B3) *Troca de derivação e integração válida:*

$$\frac{\partial}{\partial \theta_r} \int_{\mathcal{C}} f(x;\theta) dx = \int_{\mathcal{C}} \frac{\partial}{\partial \theta_r} f(x;\theta) dx = 0$$

e

$$\int_{\mathcal{C}} \frac{\partial^2}{\partial \theta_r \partial \theta_s} f(x;\theta) dx = 0, \quad r, s = 1, \ldots, p$$

(B4) *A matriz de informação de Fisher*

$$I_{X_i}(\theta) = \left[E \left(\frac{\partial \log f(X_i;\theta)}{\partial \theta_r} \frac{\partial \log f(X_i;\theta)}{\partial \theta_s} \right) \right]_{r,s=1}^{p}$$

é finita e não singular para todo $\theta \in \Theta$.
(B5) *Condições de dominação para derivadas de terceira ordem.*
Então, existe uma sequência $\hat{\theta}_n$ tal que:

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N_p\left(0, I^{-1}(\theta_0)\right)$$

onde N_p denota a distribuição normal multivariada de dimensão p.

11.6 Propriedades dos EMVs

Sob as condições de regularidade, os EMVs são:

1. **Consistentes:** $\hat{\theta}_n \xrightarrow{P} \theta$

2. **Assintoticamente não-viesados:** $\lim_{n\to\infty} E[\hat{\theta}_n] = \theta$

3. **Assintoticamente normais:** A distribuição converge para normal

4. **Assintoticamente eficientes:** Atingem o limite inferior de Cramér-Rao assintótico

11.7 Limite Inferior de Cramér-Rao Assintótico

Para qualquer estimador não-viesado T_n de θ:

$$\text{Var}(T_n) \geq \frac{1}{n \cdot I_X(\theta)}$$

Os EMVs atingem este limite assintoticamente!

11.8 Aplicação Prática

Exemplo 11.1 (Intervalo de Confiança via EMV)**.** Se $\hat{\theta}_n$ é o EMV, um IC assintótico de nível $1-\alpha$ para θ é:

$$IC_{1-\alpha}(\theta) = \hat{\theta}_n \pm z_{\alpha/2} \cdot \frac{1}{\sqrt{n \cdot I_X(\hat{\theta}_n)}}$$

onde $I_X(\hat{\theta}_n)$ é a informação de Fisher avaliada em $\hat{\theta}_n$.

Exemplo 11.2 (Distribuição Normal)**.** Para $X_i \sim N(\mu, \sigma^2)$ com σ^2 conhecido:

- EMV: $\hat{\mu}_n = \bar{X}_n$
- Informação de Fisher: $I_X(\mu) = 1/\sigma^2$
- Distribuição assintótica: $\sqrt{n}(\bar{X}_n - \mu) \sim N(0, \sigma^2)$

Neste caso, a distribuição assintótica coincide com a exata!

11.9 Vantagens e Limitações

Vantagens:

- Propriedades ótimas assintoticamente
- Princípio unificado para construir estimadores
- Aproximação normal facilita inferência

Limitações:

- Requer condições de regularidade
- Propriedades são assintóticas (podem não valer para n pequeno)
- Computação pode ser complexa (requer otimização numérica)

12 Resumo e Conexões

12.1 Hierarquia das Convergências

Convergência quase certa $\Rightarrow$ Convergência em Probabilidade $\Rightarrow$ Convergência em Distribuição

12.2 Teoremas Principais e Suas Relações

1. **LFGN:** $\bar{X}_n \xrightarrow{P} \mu$ (onde está o valor)
2. **TCL clássico:** $\frac{\sqrt{n}(\bar{X}_n-\mu)}{\sigma} \xrightarrow{D} N(0,1)$ (quão rápido chega lá e qual a forma da distribuição)
3. **TCL para S_n^2:** $\sqrt{n}(S_n^2 - \sigma^2) \xrightarrow{d} N(0, \mu_4 - \sigma^4)$ (distribuição assintótica da variância amostral)
4. **Slutsky:** Permite combinações algébricas de convergências diferentes
5. **Método Delta:** Estende para transformações não-lineares
6. **Consistência:** Propriedade fundamental de estimadores ($T_n \xrightarrow{P} \theta$)
7. **TCL para EMVs:** Propriedades assintóticas ótimas dos estimadores de máxima verossimilhança

12.3 Estratégia de Resolução de Problemas

1. **Identificar o tipo de problema:**
   - Convergência pontual? → Usar LFGN ou consistência
   - Distribuição assintótica? → Usar TCL
   - Variância/segunda ordem? → TCL para S_n^2
2. **Verificar condições:**
   - Variáveis i.i.d.?
   - Momentos necessários existem?
   - Condições de regularidade satisfeitas?
3. **Aplicar teoremas apropriados:**
   - Para médias: LFGN ou TCL clássico
   - Para variâncias: TCL para S_n^2
   - Para EMVs: TCL para EMVs
4. **Lidar com complicações:**
   - Parâmetros desconhecidos? → Slutsky
   - Transformações não-lineares? → Método Delta
   - Múltiplas convergências? → Algebra de convergências
5. **Construir inferência:**
   - Intervalos de confiança assintóticos
   - Testes de hipóteses assintóticos
   - Regiões de confiança

12.4 Quadro Sinótico: Quando Usar Cada Teorema

Objetivo	**Teorema**	**Condições**
Estimador consistente?	LFGN ou EQM $\to 0$	Momentos finitos
Distribuição de $\bar{X}_n$?	TCL clássico	$E[X_i^2] < \infty$
Distribuição de S_n^2?	TCL para S_n^2	$E[X_i^4] < \infty$
σ desconhecido?	Slutsky	$S_n \xrightarrow{P} \sigma$
Função $g(\bar{X}_n)$?	Método Delta	$g'(\mu) \neq 0$
EMV?	TCL para EMVs	Regularidade

12.5 Mensagem Final - Unidade 2

Este material auxiliar complementa as notas de aula, fornecendo:

- Explicações detalhadas dos conceitos principais
- Interpretações práticas dos resultados teóricos
- Exemplos numéricos e aplicações
- Estratégias para resolução de problemas

Os teoremas de convergência formam a base da inferência estatística moderna. Compreendê-los profundamente permite:

- Justificar procedimentos estatísticos comuns
- Desenvolver novos métodos para problemas específicos
- Avaliar propriedades de estimadores e testes
- Construir aproximações úteis para cálculos práticos

13 Capítulo 4 - Testes de Hipóteses

13.1 Introdução aos Testes de Hipóteses

13.1.1 Definições Fundamentais

Definição 13.1 (Hipótese Estatística (4.1.1))**.** Uma hipótese é uma afirmação sobre o parâmetro desconhecido θ. Por exemplo:

$$H : \mu = \mu_0, \quad H : \sigma^2 > \sigma_0^2, \quad H : \lambda \neq \alpha$$

Definição 13.2 (Formulação de Neyman-Pearson)**.** Neyman e Pearson formularam o problema de testar hipóteses como segue. Considere que se deseja escolher entre:

$$\begin{cases} H_0 : \theta \in \Theta_0, \\ H_1 : \theta \in \Theta_1, \end{cases} \tag{5}$$

tal que $\Theta = \Theta_0 \cup \Theta_1$ e $\Theta_0 \cap \Theta_1 = \varnothing$.

Então, baseando-se numa amostra aleatória $X_1, \ldots, X_n$ de X, deve-se tomar a decisão de rejeitar H_0 ou não rejeitar H_0.

13.1.2 Tipos de Hipóteses

- **Hipótese Simples:** Especifica completamente a distribuição (ex: $H_0 : \theta = \theta_0$)
- **Hipótese Composta:** Não especifica completamente (ex: $H_0 : \theta > \theta_0$)
- **Unilateral:** $H_1 : \theta > \theta_0$ ou $H_1 : \theta < \theta_0$
- **Bilateral:** $H_1 : \theta \neq \theta_0$

13.2 Seção 4.2 - Probabilidade de Erro e Função Poder

Definição 13.3 (Teste de Hipóteses (4.2.1))**.** Sejam $X_1, \ldots, X_n$ uma amostra de X com fdp (ou fmp) $f(x; \theta)$. Um teste de hipóteses é definido por uma região crítica R_C e sua complementar R_C^c tais que:

$$R_C \cup R_C^c = \mathbb{R}^n \quad \text{e} \quad R_C \cap R_C^c = \emptyset$$

Se $x \in R_C$, rejeitamos H_0; caso contrário, não rejeitamos H_0.

13.2.1 Erros em Testes de Hipóteses

Definição 13.4 (Erro Tipo I)**.**

$$\alpha = P(\text{Erro tipo I}) = P_{H_0}(\text{Rejeitar } H_0) = P_{H_0}(X \in R_C)$$

É a probabilidade de rejeitar H_0 quando H_0 é verdadeira.

Definição 13.5 (Erro Tipo II)**.**

$$\beta = P(\text{Erro tipo II}) = P_{H_1}(\text{Não rejeitar } H_0) = P_{H_1}(X \notin R_C)$$

É a probabilidade de não rejeitar H_0 quando H_1 é verdadeira.

Definição 13.6 (Função Poder (4.2.2)). O poder ou função poder de um teste γ, denotada como $Q_\gamma(\theta)$, é a probabilidade de rejeitar H_0 quando a verdade é θ:

$$Q_\gamma(\theta) = P_\theta(X \in R_C), \quad \forall \theta \in \Theta$$

Observação 13.1 (Interpretação da Função Poder).
- *Para $\theta \in \Theta_0$: $Q_\gamma(\theta) = \alpha$ (probabilidade de erro tipo I)*
- *Para $\theta \in \Theta_1$: $Q_\gamma(\theta) = 1 - \beta$ (poder do teste)*
- *Queremos: α pequeno e $Q_\gamma(\theta)$ grande para $\theta \in \Theta_1$*

13.2.2 Função Crítica e Testes Aleatorizados

Definição 13.7 (Função Crítica (4.2.3)). A função $\psi : \mathbb{R}^n \to [0, 1]$ é chamada de função crítica ou função de teste se $\psi(x)$ representa a probabilidade com a qual H_0 é rejeitada quando $X = x$ é observada.

Definição 13.8 (Tipos de Teste (4.2.4)). Um teste γ para a hipótese H_0 pode ser:

- **Não aleatorizado:** $\psi(x) \in \{0, 1\}$ (decisão determinística)
- **Aleatorizado:** $\psi(x) \in (0, 1)$ em alguma fronteira (usado em distribuições discretas)

Definição 13.9 (Tamanho e Nível do Teste (4.2.5)). Um teste γ é chamado de:

- **Tamanho α:** se $\sup_{\theta \in \Theta_0} Q_\gamma(\theta) = \alpha$
- **Nível α:** se $\sup_{\theta \in \Theta_0} Q_\gamma(\theta) \leq \alpha$

Equivalentemente usando função crítica:

$$\sup_{\theta \in \Theta_0} E_\theta[\psi(X)] = \alpha \quad \text{(tamanho)} \quad \text{ou} \quad \leq \alpha \quad \text{(nível)}$$

13.2.3 Melhor Teste e Teste UMP

Definição 13.10 (Teste Uniformemente Mais Poderoso (4.2.6)). Considere uma classe $\mathcal{C}$ de todos os testes de nível α para H_0 vs H_1. Um teste $\gamma \in \mathcal{C}$ com função poder $Q_\gamma(\theta)$ é o **melhor teste de nível α** ou o **teste uniformemente mais poderoso (UMP)** de nível α se, e só se:

$$Q_\gamma(\theta) \geq Q_{\gamma^*}(\theta), \quad \forall \theta \in \Theta_1, \quad \forall \gamma^* \in \mathcal{C}$$

13.3 Seção 4.3 - Lema de Neyman-Pearson

Observação 13.2 (Lema de Neyman-Pearson (4.3.1)). *Para testar hipóteses simples*

$$H_0 : \theta = \theta_0 \quad vs. \quad H_1 : \theta = \theta_1,$$

seja $\psi(x)$ uma função crítica que satisfaz:
(1) Para $k \geq 0$:

$$\psi(x) = \begin{cases} 1, & se\ L(\theta_1; x) > k\, L(\theta_0; x) \\ 0, & se\ L(\theta_1; x) < k\, L(\theta_0; x) \end{cases}$$

(2) k é determinado por:

$$E_{\theta_0}[\psi(x)] = \alpha$$

Então, qualquer teste satisfazendo (1) e (2) é um teste MP de nível α.

13.3.1 Estrutura da Prova do LNP

Ideia central: Mostrar que

$$[\psi_\gamma(x) - \psi_{\gamma^*}(x)]\,[L(\theta_1, x) - kL(\theta_0, x)] \geq 0$$

para todo x, onde γ é o teste LNP e γ^* é qualquer outro teste de nível α.
Passos principais:

1. Verificar a desigualdade em três casos: $\psi_\gamma = 1$, $\psi_\gamma = 0$, $\psi_\gamma \in (0, 1)$
2. Integrar ambos os lados sobre o espaço amostral
3. Usar que ambos os testes têm nível α: $Q_\gamma(\theta_0) = Q_{\gamma^*}(\theta_0) = \alpha$
4. Concluir que $Q_\gamma(\theta_1) \geq Q_{\gamma^*}(\theta_1)$

13.4 Seção 4.4 - Teste para Hipótese Composta Unilateral

13.4.1 Teste UMP via Lema de Neyman-Pearson

Para hipóteses do tipo:

$$H_0 : \theta = \theta_0 \quad \text{vs.} \quad H_1 : \theta > \theta_0 \quad (\text{ou } \theta < \theta_0)$$

Estratégia:

1. Fixar um valor arbitrário $\theta_1 \in \Theta_1$ (ex: $\theta_1 > \theta_0$)
2. Aplicar o LNP para obter teste MP para $H_0 : \theta = \theta_0$ vs $H_1 : \theta = \theta_1$
3. Se o teste resultante não depende da escolha específica de θ_1, então ele é UMP

13.4.2 Razão de Verossimilhança Monótona (RVM)

Definição 13.11 (RVM)**.** Uma família $\{f(x; \theta) : \theta \in \Theta\}$ tem **razão de verossimilhança monótona (RVM)** em $T(x)$ se, para $\theta_1 > \theta_0$, a razão

$$\frac{L(\theta_1; x)}{L(\theta_0; x)}$$

é uma função não decrescente de $T(x)$.

13.5 Procedimentos Operacionais de Testes

13.5.1 Teste Z (Normal com σ^2 conhecido)

Contexto: $X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$ com $\sigma^2 > 0$ conhecido.
Hipóteses típicas:

- Unilateral: $H_0 : \mu = \mu_0$ vs $H_1 : \mu > \mu_0$ (ou $\mu < \mu_0$)
- Bilateral: $H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$

Estatística de teste:

$$Z(X) = \sqrt{n}\frac{\bar{X}_n - \mu_0}{\sigma} \overset{H_0}{\sim} N(0,1)$$

Regras de decisão:

- *Unilateral à direita:* Rejeitar H_0 se $Z_{cal} > z_\alpha$
- *Unilateral à esquerda:* Rejeitar H_0 se $Z_{cal} < -z_\alpha$
- *Bilateral:* Rejeitar H_0 se $|Z_{cal}| > z_{\alpha/2}$

Valor-p:

- Unilateral (direita): $\hat{\alpha} = P(Z > Z_{cal})$
- Unilateral (esquerda): $\hat{\alpha} = P(Z < Z_{cal})$
- Bilateral: $\hat{\alpha} = 2 \cdot P(Z > |Z_{cal}|)$

13.5.2 Teste χ^2 (Exponencial)

Contexto: $X_1, \ldots, X_n \overset{i.i.d.}{\sim} \text{Exp}(\theta)$
Hipóteses:

$$H_0 : \theta = \theta_0 \quad \text{vs.} \quad H_1 : \theta = \theta_1 \quad (\theta_1 > \theta_0)$$

Estatística de teste:

$$Q(X) = \frac{2}{\theta_0}\sum_{i=1}^{n} X_i \sim \chi^2_{2n} \quad \text{sob } H_0$$

Justificativa: Se $X_i \sim \text{Exp}(\theta)$, então $\frac{2X_i}{\theta} \sim \chi^2_2$, logo $\sum_{i=1}^{n} \frac{2X_i}{\theta} \sim \chi^2_{2n}$.
Regra de decisão:

- Rejeitar H_0 se $Q_{cal} > \chi^2_{2n,\alpha}$ (quantil superior)
- Valor-p: $\hat{\alpha} = P(\chi^2_{2n} > Q_{cal})$

13.5.3 Teste para Bernoulli/Proporções

Contexto: $X_1, \ldots, X_n \overset{i.i.d.}{\sim} \text{Bernoulli}(p)$ para $p \in (0,1)$ desconhecido.
Hipóteses:

$$H_0 : p = p_0 \quad \text{vs.} \quad H_1 : p > p_0 \quad (\text{ou } p = p_1 > p_0)$$

Estatística suficiente: $T = \sum_{i=1}^{n} X_i \sim \text{Binomial}(n, p)$
Via LNP: A razão de verossimilhanças é crescente em T, logo o teste MP rejeita H_0 se $T > k_1$.
Determinação de k_1: Escolher o menor inteiro k_1 tal que:

$$P_{p_0}(T > k_1) \leq \alpha$$

Se necessário, usar aleatorização na fronteira:

$$\delta = \frac{\alpha - P_{p_0}(T > k_1)}{P_{p_0}(T = k_1)}$$

13.6 Exemplos Práticos de Testes

Exemplo 13.1 (Exemplo Q(4.1): Comparação de Testes para Normal)**.** Sejam $X_1, \ldots, X_n$ uma amostra de $X \sim N(\theta, 1)$ com $\theta \in \mathbb{R}$ desconhecido. Para testar $H_0 : \theta = 5.5$ vs $H_1 : \theta = 8$, considere quatro testes:

- **Teste #1:** Rejeitar H_0 se $x_1 > 7$
- **Teste #2:** Rejeitar H_0 se $\frac{x_1+x_2}{2} > 7$
- **Teste #3:** Rejeitar H_0 se $\bar{X}_n > 6$
- **Teste #4:** Rejeitar H_0 se $\bar{X}_n > 7.5$

Análise:

- Teste #1 usa apenas x_1 (desperdiça informação)
- Teste #2 usa apenas duas observações
- Teste #3 usa toda a amostra mas tem α muito alto
- Teste #4 usa toda a amostra e tem α controlado - é o teste MP!

Exemplo 13.2 (Cálculo de α e β)**.** Para Teste #1 ($n = 1$, rejeitar se $x_1 > 7$):
$\alpha = P_{H_0}(x_1 > 7) = P(Z > (7 - 5.5)/1) = P(Z > 1.5) = 0.0668$
$\beta = P_{H_1}(x_1 \leq 7) = P(Z \leq (7 - 8)/1) = P(Z \leq -1) = 0.1587$
Logo, poder $= 1 - \beta = 0.8413$.

13.7 Resumo de Conceitos-Chave

Conceito	**Definição/Fórmula**
Hipótese Nula	$H_0 : \theta \in \Theta_0$ (a ser testada)
Hipótese Alternativa	$H_1 : \theta \in \Theta_1$ (oposta a H_0)
Região Crítica	R_C: região onde rejeitamos H_0
Erro Tipo I	$\alpha = P_{H_0}(\text{rejeitar } H_0)$
Erro Tipo II	$\beta = P_{H_1}(\text{não rejeitar } H_0)$
Poder	$1 - \beta = P_{H_1}(\text{rejeitar } H_0)$
Função Poder	$Q_\gamma(\theta) = P_\theta(X \in R_C)$
Teste MP	Maximiza poder para hipóteses simples vs simples
Teste UMP	Maximiza poder uniformemente para $\theta \in \Theta_1$
LNP	Teste MP baseado em $L(\theta_1)/L(\theta_0) > k$
RVM	Razão de verossimilhança monótona em estatística T

13.8 Estratégias para Resolver Problemas de Testes

1. **Identificar o tipo de teste:**
   - Simples vs Simples? $\rightarrow$ LNP
   - Simples vs Composta Unilateral? $\rightarrow$ Verificar RVM + Karlin-Rubin

    - Bilateral? $\rightarrow$ Teste RV (não existe UMP geralmente)

2. **Encontrar a estatística apropriada:**
    - Usar estatística suficiente quando possível
    - Para Normal: $\bar{X}_n$ ou S_n^2
    - Para Exponencial: $\sum X_i$ (transformar para χ^2)
    - Para Poisson/Bernoulli: $\sum X_i$

3. **Determinar distribuição sob H_0:**
    - Padronizar para distribuições conhecidas (Z, t, χ^2, F)
    - Usar transformações quando necessário

4. **Construir região crítica:**
    - Via LNP para casos simples
    - Via quantis da distribuição sob H_0
    - Garantir nível α especificado

5. **Calcular valor-p quando aplicável:**
    - Valor-p = probabilidade sob H_0 de obter evidência tão ou mais extrema
    - Comparar com α para decisão

13.9 Conexão entre Unidades 2 e 4

Os testes de hipóteses (Unidade 4) dependem fundamentalmente dos resultados de convergência (Unidade 2):

- **TCL** justifica o uso da distribuição normal em testes Z
- **Slutsky** permite substituir σ por S_n (teste t)
- **Consistência** garante que estimadores usados em testes convergem
- **Propriedades assintóticas** validam testes para amostras grandes

Estude com atenção, pratique muito, e boa sorte!