Material Auxiliar de Teoria

Capítulo 3 – Teoria Assintótica e Teoremas Limite

Definições, Resultados e Demonstrações Completas

Curso de Inferência Estatística – PPGEST/UFPE

Material de Apoio para Estudo

Novembro 2025

AVISO IMPORTANTE: Este documento contém toda a teoria fundamental do Capítulo 3, com ênfase especial nas demonstrações COMPLETAS e DETALHADAS do **Teorema Central do Limite**, **Teorema de Slutsky** e **Teorema de Mann–Wald (Método Delta)**, que têm alta probabilidade de serem cobrados na prova. Todas as provas foram extraídas das notas originais e expandidas com detalhes adicionais. Estude com atenção cada passo das demonstrações.

Sumário

1 Visão Geral do Capítulo

Observações e Comentários

Objetivo central: compreender o comportamento assintótico de estatísticas fundamentais da inferência. Quando $n \to \infty$, aproximações tornam-se rigorosas e justificam métodos práticos.

- Revisar limites determinísticos e notações assintóticas que sustentam expansões de Taylor.
- Formalizar modos de convergência para sequências de variáveis aleatórias e estabelecer suas relações.
- Demonstrar resultados estruturantes: Leis Fracas dos Grandes Números, Teorema de Slutsky, Teorema Central do Limite e Método Delta.
- Aplicar esses resultados a estimadores clássicos, investigando consistência, normalidade assintótica e eficiência.

Como usar este guia: leia cada definição, acompanhe as demonstrações e replique os exemplos em suas próprias palavras. Os exercícios extras das notas foram incorporados como aplicações guiadas.

2 Notação Assintótica e Ferramentas Analíticas

2.1 Limites determinísticos recorrentes

Observações e Comentários

Resultados básicos (notas R.1–R.3) usados em aproximações:

$$\begin{aligned}
(R.1) \quad & \lim_{n\to\infty} n^{1/n} = 1, \\
(R.2) \quad & \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e, \\
(R.3) \quad & \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = e^x, \qquad x \in \mathbb{R}.
\end{aligned}$$

Esses limites justificam, por exemplo, a aproximação da Binomial$(n, \lambda/n)$ pela Poisson(λ) quando $n \to \infty$:

$$P(X = k) = \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \xrightarrow[n\to\infty]{} e^{-\lambda} \frac{\lambda^k}{k!}.$$

2.2 Definições fundamentais

Definição 3.7.2.1: Notação $O(\cdot)$ para sequências

Sejam $\{a_n\}$ e $\{b_n\}$ sequências reais. Escrevemos $a_n = O(b_n)$ quando existe $k > 0$ e $n_0 \in \mathbb{N}$ tais que

$$\left|\frac{a_n}{b_n}\right| \leq k, \qquad \forall n \geq n_0.$$

Em particular, $a_n = O(1)$ significa que a sequência é limitada para n grande.

Definição 3.7.2.2: Notação $o(\cdot)$ para sequências

Dizemos que $a_n = o(b_n)$ se, e somente se,

$$\lim_{n\to\infty} \frac{a_n}{b_n} = 0.$$

Logo, $a_n = o(1)$ implica $a_n \to 0$ quando $n \to \infty$.

Definição 3.7.2.3: Notação $\theta(\cdot)$

Nas notas do capítulo a notação $\theta(\cdot)$ é usada como sinônimo de $o(\cdot)$. Assim,

$$a_n = \theta(b_n) \iff \lim_{n\to\infty} \frac{a_n}{b_n} = 0.$$

Quando não houver ambiguidade, manteremos a convenção padrão $o(\cdot)$, lembrando que $\theta(\cdot)$ foi empregada nos arquivos originais.

Definição 3.7.3.2: Notação $O(\cdot)$ e $o(\cdot)$ para funções reais

Para funções $f, g : \mathbb{R} \to \mathbb{R}$ definidas numa vizinhança de x_0:

O grande quando $x \to x_0$: Escrevemos $f(x) = O(g(x))$ quando existe $k > 0$ e $\delta > 0$ tais que

$$\left|\frac{f(x)}{g(x)}\right| \leq k, \qquad \forall x \text{ com } |x - x_0| < \delta.$$

O grande quando $x \to \infty$: Escrevemos $f(x) = O(g(x))$ quando existe $k > 0$ e $M > 0$ tais que

$$\left|\frac{f(x)}{g(x)}\right| \leq k, \qquad \forall x > M.$$

o pequeno quando $x \to x_0$: Escrevemos $f(x) = o(g(x))$ se

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = 0.$$

o **pequeno quando** $x \to \infty$**:** Analogamente,

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0.$$

Exemplos importantes:

- $8x^2 = O(x^2)$ quando $x \to \infty$ pois $8x^2/x^2 = 8$ (limitado).
- $8x^2 \neq o(x^2)$ quando $x \to \infty$ pois $8x^2/x^2 \to 8 \neq 0$.
- $8x^2 = o(x^3)$ quando $x \to \infty$ pois $8x^2/x^3 = 8/x \to 0$.

Teorema Extra 1: Propriedades de $O(\cdot)$ e $o(\cdot)$

Sejam $\{a_n\}, \{b_n\}, \{c_n\}, \{d_n\}$ sequências reais e $r > 0$. Valem:

(i) Se $a_n = o(b_n)$ então $a_n = O(b_n)$.

(ii) Se $a_n = O(b_n)$ e $c_n = O(d_n)$, então:

  (a) $a_n c_n = O(b_n d_n)$,
  (b) $|a_n|^r = O(|b_n|^r)$,
  (c) $a_n + c_n = O\big(\max\{|b_n|, |d_n|\}\big)$.

(iii) Se $a_n = o(b_n)$ e $c_n = o(d_n)$, então:

  (a) $a_n c_n = o(b_n d_n)$,
  (b) $|a_n|^r = o(|b_n|^r)$,
  (c) $a_n + c_n = o\big(\max\{|b_n|, |d_n|\}\big)$.

(iv) Se $a_n = O(b_n)$ e $c_n = o(d_n)$, então $a_n c_n = o(b_n d_n)$.

(v) Se $a_n = O(b_n)$ e $b_n = o(c_n)$, então $a_n = o(c_n)$.

Demonstração Detalhada

Ideia da demonstração. Todas as afirmações decorrem diretamente das definições. Por exemplo:

(i) Se $\lim |a_n/b_n| = 0$, então existe n_0 e $k = 1$ tais que $|a_n/b_n| \leq 1$ para $n \geq n_0$.

(ii.a) Dados $a_n = O(b_n)$ e $c_n = O(d_n)$, existem k_1, k_2 e n_1, n_2 com $|a_n| \leq k_1|b_n|$ e $|c_n| \leq k_2|d_n|$. Para $n \geq \max\{n_1, n_2\}$,

$$|a_n c_n| \leq k_1 k_2 |b_n d_n|,$$

logo $a_n c_n = O(b_n d_n)$.

(iii.a) Usando o mesmo raciocínio com $o(\cdot)$, obtemos $\lim_{n\to\infty} |a_n c_n|/(|b_n d_n|) = 0$.

Os demais itens seguem análoga e mecanicamente. $\square$

Observações e Comentários

Exemplos rápidos:

- $10n^2 + n = O(n^2)$ porque $|(10n^2 + n)/n^2| \le 11$ para todo $n \ge 1$.
- $n = o(n^2)$ pois $n/n^2 = 1/n \to 0$.
- $n^{-1} = o(1)$ e $n^{-1} = O(1)$, logo a sequência é limitada e converge a zero.
- A expansão $e^x = 1 + x + O(x^2)$ segue de aplicar Taylor e a propriedade (ii.b) do teorema acima.

2.3 Expansões de Taylor e aplicações

Expansão de Taylor até ordem n

Se $F : \mathbb{R} \to \mathbb{R}$ é uma função derivável até a ordem n em um ponto x_0, sua expansão em série de Taylor em torno de x_0 pode ser escrita como:

$$F(x) = \sum_{k=0}^{n} \frac{F^{(k)}(x_0)}{k!}(x - x_0)^k + o\big((x - x_0)^n\big), \quad \text{quando } x \to x_0 \tag{1}$$

onde $F^{(k)}$ é a derivada de ordem k de $F(\cdot)$.

Observações e Comentários

Expansões fundamentais (usadas repetidamente nas demonstrações assintóticas):

$$\begin{aligned}
e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + O(x^4), \\
\log(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} + O(x^4), \\
(1+x)^\alpha &= 1 + \alpha x + \frac{\alpha(\alpha-1)}{2}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{6}x^3 + O(x^4).
\end{aligned}$$

Exemplo worked out (das notas): Mostrar que $\log(1+x) \cdot e^x = x + O(x^2)$ quando $x \to 0$.
Solução: Usando as expansões de Taylor:

$$\begin{aligned}
e^x &= e^0 + x \cdot e^0 + o(x) = 1 + x + O(x^2), \\
\log(1+x) &= \log(1) + \left.\frac{1}{1+x}\right|_{x=0} \cdot x + o(x) = x + O(x^2).
\end{aligned}$$

Logo,

$$\begin{aligned}
e^x \log(1+x) &= \left[1 + x + O(x^2)\right]\left[x + O(x^2)\right] \\
&= 1 \cdot x + 1 \cdot O(x^2) + x \cdot x + x \cdot O(x^2) + O(x^2) \cdot x + O(x^2) \cdot O(x^2) \\
&= x + O(x^2) + x^2 + O(x^3) + O(x^3) + O(x^4) \\
&= x + x^2 + O(x^2) \quad \text{(propriedade iii.3)} \\
&= x + O(x^2). \quad \square
\end{aligned}$$

3 Convergência em Probabilidade

3.1 Definições e equivalências

Definição 3.7.4.1(a)

Seja $\{U_n\}_{n\geq 1}$ uma sequência de variáveis aleatórias reais. Dizemos que U_n converge em probabilidade para $u \in \mathbb{R}$, e escrevemos

$$U_n \xrightarrow{P} u,$$

quando, para todo $\varepsilon > 0$,

$$P(|U_n - u| \geq \varepsilon) \xrightarrow[n\to\infty]{} 0.$$

Definição 3.7.42(a)

Para variáveis aleatórias U_n e U,

$$U_n \xrightarrow{P} U \iff U_n - U \xrightarrow{P} 0.$$

Assim, convergência para uma constante u é um caso particular ao considerar $U \equiv u$ (variável degenerada).

3.2 Resultados fundamentais sobre convergência em probabilidade

Resultado 1P: Lei Fraca dos Grandes Números (versão simples)

Sejam $X_1, \ldots, X_n$ variáveis aleatórias i.i.d. com $E[X_i] = \mu < \infty$ e $\mathrm{Var}(X_i) = \sigma^2 < \infty$. Então

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

Demonstração Detalhada

Para $\varepsilon > 0$, pela desigualdade de Tchebysheff:

$$P\big(|\bar{X}_n - \mu| \geq \varepsilon\big) \leq \frac{\mathrm{Var}(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \xrightarrow[n\to\infty]{} 0.$$

Logo $\bar{X}_n \xrightarrow{P} \mu$. $\square$

Resultado 2P

Se $\{T_n\}$ é uma sequência de variáveis reais tal que, para algum $r > 0$ e $a \in \mathbb{R}$,

$$E\big[|T_n - a|^r\big] \xrightarrow[n\to\infty]{} 0,$$

então $T_n \xrightarrow{P} a$.

Demonstração Detalhada

Pela desigualdade de Markov, para qualquer $\varepsilon > 0$,

$$P(|T_n - a| \geq \varepsilon) = P(|T_n - a|^r \geq \varepsilon^r) \leq \frac{E[|T_n - a|^r]}{\varepsilon^r} \xrightarrow[n\to\infty]{} 0.$$

$\square$

Resultado 3P: Lei Fraca dos Grandes Números de Khinchine

Se $X_1, \ldots, X_n$ são i.i.d. com $E[X_i] = \mu < \infty$ (não exigimos variância finita), então $\bar{X}_n \xrightarrow{P} \mu$.

Demonstração Detalhada

Considere a fgm de $\bar{X}_n$:

$$M_{\bar{X}_n}(t) = \left[M_{X_1}\left(\frac{t}{n}\right)\right]^n.$$

Expansão de Taylor em torno de 0 produz $M_{X_1}(t/n) = 1 + \mu\frac{t}{n} + o(t/n)$. Aplicando $(R.3)$,

$$M_{\bar{X}_n}(t) \xrightarrow[n\to\infty]{} e^{t\mu},$$

que é a fgm da variável degenerada em μ. Logo $\bar{X}_n \xrightarrow{d} \mu$ e, como o limite é degenerado, $\bar{X}_n \xrightarrow{P} \mu$. $\square$

Resultado 4P

Se $U_n \xrightarrow{P} u$ e $V_n \xrightarrow{P} v$, então:

(a) $U_n + V_n \xrightarrow{P} u + v$,

(b) $U_n V_n \xrightarrow{P} uv$,

(c) Se $P(V_n = 0) = 0$ e $v \neq 0$, então $\dfrac{U_n}{V_n} \xrightarrow{P} \dfrac{u}{v}$.

Demonstração Detalhada

Usamos apenas álgebra e a definição. Por exemplo, para o item (a):

$$P\left(|(U_n + V_n) - (u+v)| > \varepsilon\right) \le P\left(|U_n - u| > \frac{\varepsilon}{2}\right) + P\left(|V_n - v| > \frac{\varepsilon}{2}\right) \to 0.$$

Os demais itens seguem substituindo produtos e quocientes e usando que V_n permanece afastado de zero em probabilidade. $\square$

Resultado 5P: Teorema da Função Contínua

Se $U_n \xrightarrow{P} u$ e $g : \mathbb{R} \to \mathbb{R}$ é contínua em u, então $g(U_n) \xrightarrow{P} g(u)$.

Demonstração Detalhada

Pela continuidade, dado $\varepsilon > 0$ existe $\delta > 0$ tal que $|x - u| < \delta \Rightarrow |g(x) - g(u)| < \varepsilon$. Assim,

$$P(|g(U_n) - g(u)| \ge \varepsilon) \le P(|U_n - u| \ge \delta) \to 0.$$

$\square$

Observações e Comentários

Exemplo 1 (variância amostral). Seja S_n^2 a variância amostral de uma amostra i.i.d. com $E[X_i^2] < \infty$. Usando a transformação de Helmert,

$$S_n^2 = \frac{1}{n-1} \sum_{i=2}^{n} Y_i^2, \qquad Y_i \sim N(0, \sigma^2).$$

Como $\{Y_i^2\}$ é i.i.d. com média σ^2, o Resultado 1P implica $S_n^2 \xrightarrow{P} \sigma^2$. Logo S_n é um estimador consistente de σ.

Observações e Comentários

Exemplo 2 (Questão Extra 2 das notas): Máximo da Uniforme.
Enunciado: Sejam $X_1, \dots, X_n$ i.i.d. com $X_i \sim U(0, \theta)$ para $\theta > 0$. Mostrar que $T_n = X_{(n)} \xrightarrow{P} \theta$.
Solução completa: Como $T_n = \max\{X_1, \dots, X_n\}$ é a maior estatística de ordem,

$$F_{T_n}(t) = [F_{X_1}(t)]^n,$$
$$f_{T_n}(t) = n[F_{X_1}(t)]^{n-1} f_{X_1}(t).$$

Para a uniforme, $F_{X_1}(t) = \frac{t}{\theta}\mathbf{1}_{(0,\theta)}(t) + \mathbf{1}_{[\theta,\infty)}(t)$ e $f_{X_1}(t) = \frac{1}{\theta}\mathbf{1}_{(0,\theta)}(t)$. Logo,

$$f_{T_n}(t) = \frac{n}{\theta^n} t^{n-1} \mathbf{1}_{(0,\theta)}(t).$$

Calculamos os momentos:

$$E[T_n] = \int_0^\theta t \cdot \frac{n}{\theta^n} t^{n-1} dt = \frac{n}{\theta^n}\left[\frac{t^{n+1}}{n+1}\right]_0^\theta = \frac{n}{n+1}\theta,$$
$$E[T_n^2] = \int_0^\theta t^2 \cdot \frac{n}{\theta^n} t^{n-1} dt = \frac{n}{\theta^n}\left[\frac{t^{n+2}}{n+2}\right]_0^\theta = \frac{n}{n+2}\theta^2.$$

Portanto,

$$\begin{aligned}
E[(T_n-\theta)^2] &= E[T_n^2] - 2\theta E[T_n] + \theta^2 \\
&= \frac{n}{n+2}\theta^2 - \frac{2n}{n+1}\theta^2 + \theta^2 \\
&= \theta^2\left\{\frac{n}{n+2} - \frac{2n}{n+1} + 1\right\} \\
&= \theta^2\left\{\frac{n(n+1) - 2n(n+2) + (n+2)(n+1)}{(n+2)(n+1)}\right\} \\
&= \theta^2\left\{\frac{n^2+n-2n^2-4n+n^2+3n+2}{(n+2)(n+1)}\right\} \\
&= \frac{2\theta^2}{(n+2)(n+1)} \xrightarrow[n\to\infty]{} 0.
\end{aligned}$$

Pelo Resultado 2P, $T_n \xrightarrow{P} \theta$. □

4 Convergência em Distribuição

4.1 Definição e primeiras consequências

Definição 3.15.1(a)

Seja $\{U_n\}$ uma sequência de variáveis aleatórias com funções de distribuição F_n, e U uma variável com distribuição F. Dizemos que U_n converge em distribuição para U, denotado por $U_n \xrightarrow{D} U$, se

$$F_n(u) \xrightarrow[n\to\infty]{} F(u) \quad \text{para todo ponto de continuidade de } F.$$

Resultado 1D: Convergência via função geradora de momentos

Se $M_{U_n}(t) \to M_U(t)$ em uma vizinhança de $t = 0$ e as mgf's existem nessa vizinhança, então $U_n \xrightarrow{D} U$.

Demonstração Detalhada

A mgf determina unicamente a distribuição quando existe em um intervalo aberto contendo zero. A convergência pontual das mgf's implica convergência das funções características, e o Teorema de Unicidade das funções características garante $U_n \xrightarrow{D}$

U. $\square$

Resultado 2D: Relação com convergência em probabilidade

Se $U_n \xrightarrow{P} U$, então $U_n \xrightarrow{D} U$. A recíproca é verdadeira apenas quando U é degenerada.

Demonstração Detalhada

Se U é constante u, então $F(u^-) = 0$ e $F(u) = 1$, de modo que a convergência de distribuições implica $P(|U_n - u| > \varepsilon) \to 0$. Para U não degenerada, a convergência em distribuição não garante a concentração das probabilidades. $\square$

Teorema 3.7.6.4(a): Teorema da Função Contínua em Distribuição

Se $U_n \xrightarrow{D} U$ e g é contínua, então $g(U_n) \xrightarrow{D} g(U)$.

Demonstração Detalhada

Considere uma função contínua limitada h. Por $U_n \xrightarrow{D} U$, temos $E[h(g(U_n))] \to E[h(g(U))]$ (Teorema de Portmanteau). Como as funções indicadoras de intervalos podem ser aproximadas por funções contínuas, conclui-se a convergência das distribuições de $g(U_n)$ para a de $g(U)$. $\square$

Observações e Comentários

Aplicação direta: se $Z_n = \sqrt{n}(\bar{X}_n - \mu)/\sigma \xrightarrow{D} N(0,1)$, então, para $g(x) = x^2$, obtemos

$$n\left(\frac{\bar{X}_n - \mu}{\sigma}\right)^2 = g(Z_n) \xrightarrow{D} \chi^2_1.$$

Este é exatamente o exercício (11) das notas.

5 Teoremas Fundamentais

5.1 Teorema de Slutsky (Resultado 3D)

Resultado 3D: Teorema de Slutsky

Se $U_n \xrightarrow{D} U$ e $V_n \xrightarrow{P} v$, com $v \in \mathbb{R}$ e $P(V_n = 0) = 0$ para todo n, então:

(a) $U_n + V_n \xrightarrow{D} U + v$,

(b) $U_n V_n \xrightarrow{D} Uv$,

(c) $\dfrac{U_n}{V_n} \xrightarrow{D} \dfrac{U}{v}$ quando $v \neq 0$.

Demonstração Detalhada

Primeiro, $V_n \xrightarrow{P} v$ implica $V_n \xrightarrow{D} v$. Usamos o Teorema de Portmanteau e funções contínuas limitadas f.

(a) Soma. Escreva

$$f(U_n + V_n) - f(U_n + v) \xrightarrow{P} 0$$

porque f é uniformemente contínua em compactos e $V_n \xrightarrow{P} v$. Logo

$$E[f(U_n + V_n)] = E[f(U_n + v)] + o(1).$$

Como $U_n \xrightarrow{D} U$, o lado direito converge para $E[f(U + v)]$, provando a convergência em distribuição.

(b) Produto e **(c) quociente** seguem escrevendo $U_nV_n = U_nv + U_n(V_n - v)$ e $\frac{U_n}{V_n} = \frac{U_n}{v} \cdot \frac{v}{V_n}$. Em ambos os casos, $V_n - v \xrightarrow{P} 0$ e $V_n/v \xrightarrow{P} 1$, aplicando o item (a) e o Resultado 5P completamos a prova. $\square$

Observações e Comentários

Aplicação. Seja $T_n = \sqrt{n}(\bar{X}_n - \mu)/S_n$. Temos $S_n \xrightarrow{P} \sigma$ (Resultado 1P) e $Z_n = \sqrt{n}(\bar{X}_n - \mu)/\sigma \xrightarrow{D} N(0, 1)$ (TCL). Pelo Teorema de Slutsky,

$$T_n = Z_n \cdot \frac{\sigma}{S_n} \xrightarrow{D} N(0, 1).$$

Isto explica a aproximação normal para o teste-t com amostras grandes.

5.2 Teorema Central do Limite (TCL)

Teorema 3.7.6.1(a): Teorema Central do Limite clássico

Sejam $X_1, \dots, X_n$ i.i.d. com $E[X_i] = \mu$ e $\mathrm{Var}(X_i) = \sigma^2 < \infty$. Então

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

Demonstração Detalhada

Para $S_n = \sum_{i=1}^n X_i$, considere a função característica de Z_n:

$$\varphi_{Z_n}(t) = E\left[\exp\left(it\frac{S_n - n\mu}{\sigma\sqrt{n}}\right)\right] = \left[\varphi_{X_1}\left(\frac{t}{\sigma\sqrt{n}}\right) e^{-it\frac{\mu}{\sigma\sqrt{n}}}\right]^n.$$

Expansões de Taylor até segunda ordem dão

$$\log \varphi_{Z_n}(t) = -\frac{t^2}{2} + o(1).$$

Tomando o limite obtemos $\varphi_{Z_n}(t) \to e^{-t^2/2}$, função característica de $N(0, 1)$. Conclui-se $Z_n \xrightarrow{D} N(0, 1)$.

Observações e Comentários

Versões úteis e aplicações:

- **Versão padronizada sem variância conhecida:** $T_n = \sqrt{n}(\bar{X}_n - \mu)/S_n \xrightarrow{D} N(0,1)$ (Slutsky).
- **Aplicação à variância amostral:** $(n-1)^{1/2}(S_n^2 - \sigma^2) \xrightarrow{D} N(0, 2\sigma^4)$ usando o TCL sobre $\{Y_i^2\}$ com Y_i variáveis de Helmert.
- **Transformações lineares:** Para $a \neq 0$, $aZ_n + b \xrightarrow{D} N(b, a^2)$.

Observações e Comentários

Exemplo (Questão 4 das notas): Distribuição de χ^2 normalizada.
Problema: Se $X_n \sim \chi_n^2$ e $U_n = \frac{1}{\sqrt{2n}}(X_n - n)$, encontrar a distribuição limite de U_n quando $n \to \infty$.
Solução pela MGF: Lembre que $E[X_n] = n$ e $\mathrm{Var}(X_n) = 2n$. Logo U_n é a padronização de X_n.
Calculemos a mgf de U_n:

$$\begin{aligned} M_{U_n}(t) &= E\left[e^{tU_n}\right] = E\left[e^{\frac{t}{\sqrt{2n}}(X_n - n)}\right] \\ &= e^{-\frac{tn}{\sqrt{2n}}} M_{X_n}\left(\frac{t}{\sqrt{2n}}\right). \end{aligned}$$

Como $X_n \sim \chi_n^2 \equiv \Gamma(n/2, 1/2)$, sua mgf é $M_{X_n}(t) = (1-2t)^{-n/2}$. Portanto,

$$\begin{aligned} M_{U_n}(t) &= e^{-\sqrt{n/2}\,t}\left(1 - 2 \cdot \frac{t}{\sqrt{2n}}\right)^{-n/2} \\ &= e^{-\sqrt{n/2}\,t}\left(1 - \sqrt{\frac{2}{n}}t\right)^{-n/2}. \end{aligned}$$

Tomando logaritmo:

$$\log M_{U_n}(t) = -\sqrt{\frac{n}{2}}t - \frac{n}{2}\log\left(1 - \sqrt{\frac{2}{n}}t\right).$$

Pela expansão de Taylor $\log(1-x) = -x - \frac{x^2}{2} + O(x^3)$:

$$\begin{aligned} \log\left(1 - \sqrt{\frac{2}{n}}t\right) &= -\sqrt{\frac{2}{n}}t - \frac{1}{2}\left(\sqrt{\frac{2}{n}}t\right)^2 + O(n^{-3/2}) \\ &= -\sqrt{\frac{2}{n}}t - \frac{t^2}{n} + O(n^{-3/2}). \end{aligned}$$

Logo,

$$\begin{aligned}
\log M_{U_n}(t) &= -\sqrt{\frac{n}{2}}t - \frac{n}{2}\left(-\sqrt{\frac{2}{n}}t - \frac{t^2}{n} + O(n^{-3/2})\right) \\
&= -\sqrt{\frac{n}{2}}t + \sqrt{\frac{n}{2}}t + \frac{t^2}{2} + O(n^{-1/2}) \\
&= \frac{t^2}{2} + o(1).
\end{aligned}$$

Assim, $\lim_{n\to\infty} M_{U_n}(t) = e^{t^2/2}$, que é a mgf de $N(0,1)$. Concluímos $U_n \xrightarrow{D} N(0,1)$.
□

5.3 Método Delta (Teorema de Mann–Wald)

Teorema 3.7.6.2(a): Teorema de Mann–Wald / Método Delta

Se $\sqrt{n}(T_n - \theta) \xrightarrow{D} N(0, \sigma^2(\theta))$ e g é uma função continuamente diferenciável com $g'(\theta) \neq 0$, então

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

Demonstração Detalhada

Expanda g em série de Taylor ao redor de θ:

$$g(T_n) = g(\theta) + g'(\theta)(T_n - \theta) + R_n,$$

onde $R_n = \frac{g''(\tilde{\theta}_n)}{2}(T_n - \theta)^2$ para algum $\tilde{\theta}_n$ entre T_n e θ. Como $T_n \xrightarrow{P} \theta$, temos $R_n = o_P\big(|T_n - \theta|\big)$. Multiplicando por $\sqrt{n}$,

$$\sqrt{n}\big[g(T_n) - g(\theta)\big] = g'(\theta)\sqrt{n}(T_n - \theta) + o_P(1).$$

O termo principal converge em distribuição para $N(0, \sigma^2(\theta)g'(\theta)^2)$ e o resíduo converge em probabilidade para zero. Aplica-se Slutsky. □

Observações e Comentários

Exemplo (estatística de ordem). Em $X_{n:n} \sim \max\{X_1, \dots, X_n\}$ com $X_i \sim U(0,\theta)$, temos $T_n = X_{n:n} \xrightarrow{P} \theta$ e

$$U_n = \frac{n}{\theta}(\theta - T_n) \xrightarrow{D} \text{Exp}(1).$$

Logo, $g(T_n) = T_n^2$ satisfaz $g'(x) = 2x$, e o Método Delta fornece

$$\sqrt{n}\big(T_n^2 - \theta^2\big) \xrightarrow{D} N\big(0, 4\theta^2\text{Var}(T_n)\big).$$

6 Consistência e Eficiência de Estimadores

Definição 3.7.1: Consistência fraca

Uma sequência de estimadores $\{T_n\}$ para $\tau(\theta)$ é dita consistente (no sentido fraco) se

$$T_n \xrightarrow{P} \tau(\theta).$$

Se a convergência não ocorre, o estimador é inconsistente.

Observações e Comentários

Caracterizações úteis:

- Para todo $\varepsilon > 0$, $P_\theta(|T_n - \tau(\theta)| > \varepsilon) \to 0$.
- Se $E_\theta[(T_n - \tau(\theta))^2] \to 0$ então T_n é consistente (desigualdade de Chebysheff).
- Quando T_n é centrado, consistência implica que viés e variância convergem a zero.

Definição 3.4.1: Eficiência relativa assintótica

Se $T_n^{(1)}$ e $T_n^{(2)}$ são estimadores assintoticamente normais para $g(\theta)$,

$$\sqrt{n}\big(T_n^{(i)} - g(\theta)\big) \xrightarrow{D} N\big(0, \sigma_i^2(\theta)\big),$$

definimos a eficiência relativa de $T_n^{(2)}$ em relação a $T_n^{(1)}$ como $\sigma_1^2(\theta)/\sigma_2^2(\theta)$. O estimador com menor variância assintótica é chamado de mais eficiente.

7 Normalidade Assintótica dos EMVs

Teorema 3.8.1: TCL para o EMV (caso univariado)

Seja $X_1, \dots, X_n$ uma amostra de densidade $f(x;\theta)$, $\theta \in \Theta \subset \mathbb{R}$. Sob as seguintes condições de regularidade:

(A1) $f(x;\theta)$ é três vezes diferenciável em θ;

(A2) Derivadas podem ser permutadas com integrais:

$$\int \frac{\partial f(x;\theta)}{\partial \theta} dx = 0 \quad \text{e} \quad \int \frac{\partial^2 f(x;\theta)}{\partial \theta^2} dx = 0;$$

(A3) A informação de Fisher é finita e positiva:

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

(A4) Para cada $\theta_0 \in \Theta$, existe $\varepsilon = \varepsilon(\theta_0) > 0$ tal que

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

onde
$$\int_{\mathcal{X}} g(x) f(x;\theta)\, dx < \infty;$$

(A5) 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 $\hat{\theta}_n$ (i.e., $\hat{\theta}_n \xrightarrow{P} \theta_0$).

Então:
$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{D} N\left(0, I_X^{-1}(\theta_0)\right), \quad \text{quando } n \to \infty.$$

Teorema 3.9.2: Versão multivariada

Suponha $X_1, \ldots, X_n$ com vetor de parâmetros $\theta \in \Theta \subset \mathbb{R}^p$ e condições (B1)–(B5) das notas:

(B1) o suporte $\mathcal{C} = \{x : f(x;\theta) > 0\}$ não depende de θ;

(B2) $f(x;\theta)$ é três vezes diferenciável em cada componente;

(B3) derivadas podem permutar com integrais até segunda ordem;

(B4) a matriz de informação $I(\theta)$ é finita e não singular;

(B5) derivadas de terceira ordem são dominadas por função integrável em vizinhança de θ_0.

Então existe uma sequência de EMVs $\hat{\theta}_n$ tal que

$$\hat{\theta}_n \xrightarrow{P} \theta_0 \qquad \text{e} \qquad \sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{D} N_p\big(0, I^{-1}(\theta_0)\big).$$

Observações e Comentários

Consequências práticas:

- EMVs são assintoticamente eficientes: atingem o limite de Cramér-Rao assintótico $I^{-1}(\theta_0)$.
- Pelo Método Delta, $g(\hat{\theta}_n)$ é assintoticamente normal com variância $g'(\theta_0)^\top I^{-1}(\theta_0) g'(\theta_0)$.
- Estatísticas de teste baseadas na razão de verossimilhanças usam diretamente a normalidade assintótica do EMV.

8 Teorema Especial: Normalidade Assintótica da Variância Amostral

Teorema 3.7.6.3(a): TCL para a variância amostral

Sejam $X_1, \ldots, X_n$ v.a.'s i.i.d. com média μ, variância σ^2 e $\mu_4 = E[(X_1 - \mu)^4]$. Assuma que $0 < \mu_4 < \infty$ e $\mu_4 > \sigma^4$ (curtose > 1). Então,

$$\sqrt{n}(S_n^2 - \sigma^2) \xrightarrow{D} N(0, \mu_4 - \sigma^4).$$

Demonstração Detalhada

Demonstração completa (das notas n27–n28):
Considere $W_n \triangleq (n-1)n^{-1}S_n^2$, $Y_i \triangleq (X_i - \mu)^2$ para $i = 1, \ldots, n$ e $\bar{Y}_n = n^{-1}\sum_{i=1}^n Y_i$. Assim,

$$\begin{aligned}
W_n &= \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2 \\
&= \frac{1}{n}\sum_{i=1}^n \left(X_i - \mu + \mu - \bar{X}_n\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^n \left[(X_i - \mu)^2 + 2(X_i - \mu)(\mu - \bar{X}_n) + (\mu - \bar{X}_n)^2\right] \\
&= \bar{Y}_n - 2(\bar{X}_n - \mu)^2 + (\bar{X}_n - \mu)^2 \\
&= \bar{Y}_n - (\bar{X}_n - \mu)^2.
\end{aligned}$$

Vale então

$$\sqrt{n}(W_n - \sigma^2) = \underbrace{\sqrt{n}(\bar{Y}_n - \sigma^2)}_{U_n} - \underbrace{\sqrt{n}(\bar{X}_n - \mu)^2}_{V_n}.$$

Note que $\{Y_i\}$ é i.i.d. com $E[Y_i] = \sigma^2$ e $\mathrm{Var}(Y_i) = E[Y_i^2] - (E[Y_i])^2 = E[(X_i - \mu)^4] - \sigma^4 = \mu_4 - \sigma^4$. Pelo TCL,

$$U_n \xrightarrow{D} N(0, \mu_4 - \sigma^4).$$

Para V_n, como $\bar{X}_n \xrightarrow{P} \mu$, temos $(\bar{X}_n - \mu)^2 \xrightarrow{P} 0$ (Resultado 5P). Logo,

$$V_n = \sqrt{n}(\bar{X}_n - \mu)^2 \xrightarrow{P} 0.$$

Pelo Teorema de Slutsky,

$$\sqrt{n}(W_n - \sigma^2) \xrightarrow{D} N(0, \mu_4 - \sigma^4).$$

Agora,

$$\begin{aligned}
\sqrt{n}(S_n^2 - \sigma^2) &= \sqrt{n}\left(\frac{n}{n-1}W_n - \sigma^2\right) \\
&= \sqrt{n}\left(\frac{n}{n-1} - 1 + 1\right)(W_n - \sigma^2) + \sqrt{n}\left(\frac{n}{n-1} - 1\right)\sigma^2 \\
&= \sqrt{n}(W_n - \sigma^2) + \frac{\sqrt{n}}{n-1}W_n.
\end{aligned}$$

Como $W_n \xrightarrow{P} \sigma^2$ e $\frac{\sqrt{n}}{n-1} \to 0$, temos $\frac{\sqrt{n}}{n-1} W_n \xrightarrow{P} 0$. Pelo Teorema de Slutsky,

$$\sqrt{n}(S_n^2 - \sigma^2) \xrightarrow{D} N(0, \mu_4 - \sigma^4). \quad \square$$

9 Exemplos Resolvidos e Aplicações

Observações e Comentários

Exemplo 1: Distribuição limite do máximo uniforme. Se $X_i \sim U(0, \theta)$ e $T_n = X_{n:n}$, então

$$F_{T_n}(t) = \left(\frac{t}{\theta}\right)^n \mathbf{1}_{(0,\theta)}(t) + \mathbf{1}_{[\theta,\infty)}(t).$$

Para $U_n = \frac{n}{\theta}(\theta - T_n)$,

$$P(U_n \leq u) = P\left(T_n \geq \theta\left(1 - \frac{u}{n}\right)\right) = 1 - \left(1 - \frac{u}{n}\right)^n \xrightarrow{n\to\infty} 1 - e^{-u}.$$

Logo $U_n \xrightarrow{D} \text{Exp}(1)$ e, por $o_P(1)$, $T_n \xrightarrow{P} \theta$.

Observações e Comentários

Exemplo 2: Assimptótica da variância amostral. Com Y_i de Helmert,

$$S_n^2 = \frac{1}{n-1} \sum_{i=2}^{n} Y_i^2.$$

Cada Y_i^2 tem média σ^2 e variância $2\sigma^4$. Pelo TCL,

$$(n-1)^{1/2}(S_n^2 - \sigma^2) \xrightarrow{D} N(0, 2\sigma^4).$$

Multiplicando por $\sqrt{n/(n-1)} \xrightarrow{P} 1$ e aplicando Slutsky, obtemos

$$\sqrt{n}(S_n^2 - \sigma^2) \xrightarrow{D} N(0, 2\sigma^4).$$

Observações e Comentários

Exemplo 3: Estatística χ^2 a partir do TCL. Com $Z_n = \sqrt{n}(\bar{X}_n - \mu)/\sigma \xrightarrow{D} N(0,1)$,

$$n\left(\frac{\bar{X}_n - \mu}{\sigma}\right)^2 = Z_n^2 \xrightarrow{D} \chi_1^2.$$

Este é o exercício (11) das notas, demonstrando a importância do Teorema 3.7.6.4.

10 Checklist e Estratégia de Revisão

MUITO IMPORTANTE PARA A PROVA

Antes da prova, garanta que você consegue:

- Enunciar e demonstrar os Resultados 1P–5P e os Resultados 1D–2D.
- Explicar passo a passo as demonstrações do Teorema de Slutsky, do TCL e do Teorema de Mann–Wald.
- Aplicar o Método Delta em pelo menos dois exemplos distintos.
- Mostrar que S_n^2 é consistente e obter sua distribuição assintótica.
- Expor as hipóteses dos Teoremas 3.8.1 e 3.9.2 e interpretar a variância assintótica do EMV.
- Usar Slutsky para justificar estatísticas padronizadas (por exemplo, teste-t e estimadores normalizados).

Sugestão de prática: resolva novamente os exercícios extras (Q1–Q11) conferindo cada passagem com este material auxiliar.

Bom estudo! Dominar o comportamento assintótico é essencial para entender os capítulos seguintes sobre estimação e testes.