Probabilidade (PPGECD000000001)

Programa de Pós-Graduação em Estatística e Ciência de Dados (PGECD)

Sessão 12

Raydonal Ospina

Departamento de Estatística
Universidade Federal da Bahia
Salvador/BA

Motivação

Consideremos uma sequência de variáveis aleatórias independentes, $X_1, X_2, \ldots,$ definidas no mesmo espaço de probabilidade $(\Omega, \mathcal{A}, P)$, e seja $S_1, S_2, \ldots$ a sequência de somas parciais, definidas por $S_n = X_1 + X_2 + \ldots + X_n$. A Lei dos Grandes Números trata da convergência de $\frac{1}{n}(S_n - ES_n)$ para 0, quando $n \to \infty$, supondo que as variáveis aleatórias X_i's sejam integráveis. Quando a sequência obedece à lei dos grandes números, existe uma tendência da variável aleatória $\frac{S_n}{n}$, a média amostral no caso de variáveis aleatórias independentes e identicamente distribuídas, para concentrar-se em torno de sua média. O Teorema Central do Limite prova que sob certas hipóteses gerais, a distribuição da média amostral padronizada tende à normal. O problema consiste em achar condições sob as quais

$$\frac{S_n - ES_n}{\sqrt{VarS_n}} \to^D N(0,1).$$

Resumidamente, estas condições exigem que cada parcela da soma contribua com um valor sem importância para a variação da soma, ou seja é muito improvável que qualquer parcela isolada dê uma contribuição muito grande para a soma.

Motivação

O Teorema Central do Limite dá apoio ao uso da normal como distribuição de erros, pois em muitas situações reais é possível interpretar o erro de uma observação como resultante de muitos erros pequenos e independentes. Há também outras situações que o Teorema Central do Limite pode justificar o uso da normal. Por exemplo, a distribuição de alturas de homens adultos de certa idade pode ser considerada aproximadamente normal, pois a altura pode ser pensada como soma de muitos efeitos pequenos e independentes.

Teoremas e provas

Existem vários Teoremas Centrais do Limite que variam de acordo com as hipóteses sobre as distribuições das variáveis aleatórias X_i's na sequência. Como teoremas centrais do limite tratam de convergência em distribuição e como, pelo Teorema da Continuidade de Levy, sabe-se que uma sequência de variáveis aleatórias $Y_n \to^D Y$ se, e somente se, $\phi_{Y_n} \to \phi_Y$, a idéia será provar que a função característica de $\frac{S_n - ES_n}{\sqrt{VarS_n}}$ converge para $e^{\frac{-t^2}{2}}$ que é a função característica da $N(0, 1)$. Nós iremos agora enunciar e provar alguns desses teoremas, começando pelo caso de variáveis aleatórias independentes e identicamente distribuídas.

Teorema 1

Sejam $X_1, X_2, \ldots$ variáveis aleatórias iid com $E(X_n) = \mu$ e $Var(X_n) = \sigma^2$. Suponha que N é uma variável aleatória com distribuição $N(0, 1)$. Se $S_n = X_1 + X_2 + \ldots + X_n$, então

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \to^D N.$$

Prova do TCL para V.A. i.i.d.

Sem perda de generalidade, seja $E(X_n) = 0$ e $E(X_n^2) = 1$ (caso este não seja o caso, pode-se provar o resultado para

$$X_i^* = \frac{X_i - \mu}{\sigma},$$

já que $E(X_i^*) = 0$ e $E(X_i^*)^2 = 1$).

Seja $\phi_n(t) = E(e^{it\frac{S_n}{\sqrt{n}}})$ e $\phi(t) = E(e^{itX_1})$. Como a função característica de uma soma de variáveis aleatórias independentes é igual ao produto das funções características das variáveis aleatórias, tem-se que

$$\phi_n(t) = (E(e^{it\frac{X_1}{\sqrt{n}}}))^n = \phi^n(t/\sqrt{n}).$$

Prova do TCL para V.A. i.i.d.

Como os dois primeiros momentos existem, ϕ possui duas derivadas contínuas. Então, utilizando a expansão de Taylor de ϕ e o fato que $\phi^{(k)}(0) = i^k E(X_1^k)$, temos que

$$\phi(t) = 1 + t\phi'(0) + \frac{t^2}{2}\phi''(\theta(t)),$$

onde $|\theta(t)| \leq |t|$. Logo, como ϕ'' é contínua em 0, temos que $\phi''(\theta(t)) - \phi''(0) \to 0$ quando $t \to 0$. Então, tem-se

$$\phi(t) = 1 - \frac{t^2}{2} + \frac{t^2}{2}e(t),$$

onde $e(t) = \phi''(\theta(t)) + 1$ e $\lim_{t\to 0} e(t) = 0$.

Prova do TCL para V.A. i.i.d.

Então, para t fixo

$$\phi^n(\frac{t}{\sqrt{n}}) = [1 - \frac{t^2}{2n} + \frac{t^2}{2n} e(\frac{t}{\sqrt{n}})]^n$$
$$= [1 + \frac{-t^2}{2n}[1 - e(\frac{t}{\sqrt{n}})]]^n \to e^{\frac{-t^2}{2}},$$

quando $n \to \infty$, pois $[1 - e(\frac{t}{\sqrt{n}})] \to 1$ e para números complexos $c_n \to c \Rightarrow (1 + \frac{c_n}{n})^n \to e^c$ (Esse limite é conhecido como limite de Euler e sua prova será omitida).

TCL de De Moivre e Laplace

Um caso especial do Teorema Central do Limite para variáveis aleatórias independentes e identicamente distribuídas é quando estas variáveis são distribuídas de acordo com a distribuição de Bernoulli, este caso é conhecido como Teorema Central do Limite de De Moivre e Laplace.

Corolário 1

Seja $X_1, X_2, \ldots$ *uma sequência de variáveis aleatórias independentes e distribuídas de acordo com a distribuição de Bernoulli com parâmetro p, ou seja,* $P(X_i = 1) = p = 1 - P(X_i = 0)$ *para* $0 < p < 1$*. Então, se* $S_n = X_1 + \ldots X_n$*,*

$$\frac{S_n - np}{\sqrt{np(1-p)}} \to^{D} N(0, 1).$$

Demonstração.

É imediata dado o teorema anterior, já que $E(X_i) = p$ e $E(X_i^2) = p$. □

Exemplo

Exemplo 1

Suponha que temos algumas voltagens de ruídos independentes, por exemplo $V_i, i = 1, 2, \ldots, n$, as quais são recebidas naquilo que se denomina um "somador". Seja V a soma das voltagens recebidas. Suponha também que cada variável aleatória V_i seja uniformemente distribuída sobre o intervalo [0,10]. Daí, $EV_i = 5$ volts e $VarV_i = \frac{100}{12}$. De acordo com o Teorema Central do Limite, se n for suficientemente grande, a variável aleatória

$$S = \frac{(V - 5n)\sqrt{12}}{10\sqrt{n}}$$

terá aproximadamente a distribuição $N(0, 1)$. Portanto, se $n = 20$, podemos calcular que a probabilidade de que a voltagem total na entrada exceda 105 volts da seguinte maneira:

$$P(V > 105) = P\left(\frac{(V - 100)\sqrt{12}}{10\sqrt{20}} > \frac{(105 - 100)\sqrt{12}}{10\sqrt{20}}\right)$$
$$\simeq 1 - \Phi(0, 388) = 0, 352.$$

TCL de Lindeberg

Agora analisaremos um resultado mais forte que dá condições gerais que garantem convergência da média amostral padronizada para normal: o Teorema Central do Limite de Lindeberg.

Teorema 2

Sejam $X_1, X_2, \ldots$ variáveis aleatórias independentes tais que $E(X_n) = \mu_n$ e $Var(X_n) = \sigma_n^2 < \infty$, onde pelo menos um $\sigma_i^2 > 0$. Sejam $S_n = X_1 + \ldots + X_n$ e $s_n = \sqrt{Var(S_n)} = \sqrt{\sigma_1^2 + \ldots + \sigma_n^2}$. Considere a seguinte condição, conhecida como condição de Lindeberg,

$$\forall \epsilon > 0, \lim_{n \to \infty} \frac{1}{s_n^2} \sum_{k=1}^{n} \int_{|x-\mu_k| > \epsilon s_n} (x - \mu_k)^2 dF_k(x) = 0.$$

Então, se a condição de Lindeberg é satisfeita

$$\frac{S_n - ES_n}{s_n} \to^D N(0,1).$$

TCL de Lindeberg

Antes de provarmos este teorema, vamos primeiro dar alguma intuição sobre a condição de Lindeberg. Esta condição diz que, para n grande, a parcela da variância devida às caudas das X_k é desprezível.
A condição de Lindeberg implica que as parcelas X_k da soma têm variâncias uniformemente pequenas para n grande, em outras palavras nenhuma parcela tem muito peso na soma. Formalmente, a condição de Lindeberg implica que $\max_{1 \leq k \leq n} \frac{\sigma_k^2}{s_n^2} \rightarrow 0$ quando $n \rightarrow \infty$.

TCL de Lindeberg

Para ver isto, observe que para todo k,

$$\begin{aligned}
\frac{\sigma_k^2}{s_n^2} &= \frac{1}{s_n^2}\int_{|x-\mu_k|\le\epsilon' s_n}(x-\mu_k)^2 dF_k(x)\\
&+\frac{1}{s_n^2}\int_{|x-\mu_k|>\epsilon' s_n}(x-\mu_k)^2 dF_k(x)\\
&\le \frac{1}{s_n^2}\int_{|x-\mu_k|\le\epsilon' s_n}(\epsilon' s_n)^2 dF_k(x)\\
&+\frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|x-\mu_j|>\epsilon' s_n}(x-\mu_j)^2 dF_j(x)\\
&\le \frac{1}{s_n^2}\int_{-\infty}^{\infty}(\epsilon' s_n)^2 dF_k(x)\\
&+\frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|x-\mu_j|>\epsilon' s_n}(x-\mu_j)^2 dF_j(x).
\end{aligned}$$

TCL de Lindeberg

Este último termo não depende de k, pois a primeira parcela é igual a $(\epsilon')^2$. Portanto, temos

$$\max_{1\leq k\leq n}\frac{\sigma_k^2}{s_n^2}\leq(\epsilon')^2+\frac{1}{s_n^2}\sum_{k=1}^{n}\int_{|x-\mu_k|>\epsilon' s_n}(x-\mu_k)^2dF_k(x),$$

que converge para $(\epsilon')^2$, pela condição de Lindeberg. Como isto vale para todo ϵ', temos $\max_{1\leq k\leq n}\frac{\sigma_k^2}{s_n^2}\to 0$.

Portanto, o Teorema Central do Limite de Lindeberg pode ser aplicado para justificar o seguinte raciocínio: a soma de um grande número de pequenas quantidades independentes tem aproximadamente uma distribuição normal.

Exemplo

Exemplo 2

Vamos verificar neste exemplo que uma sequência $X_1, X_2, \ldots$ de variáveis aleatórias i.i.d. com $EX_i = \mu$ e $VarX_i = \sigma^2$ satisfaz a condição de Lindeberg. Note que $s_n = \sqrt{VarS_n} = \sigma\sqrt{n}$. Então para $\epsilon > 0$, e F a distribuição comum das variáveis aleatórias:

$$
\begin{aligned}
&\frac{1}{s_n^2}\sum_{k=1}^{n}\int_{|x-\mu_k|>\epsilon s_n}(x-\mu_k)^2 dF_k(x) \\
&= \frac{1}{n\sigma^2}\sum_{k=1}^{n}\int_{|x-\mu|>\epsilon\sigma\sqrt{n}}(x-\mu)^2 dF(x) \\
&= \frac{1}{n\sigma^2}n\int_{|x-\mu|>\epsilon\sigma\sqrt{n}}(x-\mu)^2 dF(x).
\end{aligned}
$$

Então, finalmente,

$$\lim_n \frac{1}{\sigma^2}\int_{|x-\mu|>\epsilon\sigma\sqrt{n}}(x-\mu)^2 dF(x) = 0.$$

Prova do TCL de Lindeberg

Assim como no caso de variáveis aleatórias i.i.d., mostraremos que a função característica de $\frac{S_n - ES_n}{s_n}$ converge para $e^{\frac{-t^2}{2}}$.
Para tanto, fixemos $t \in R$. Usaremos duas versões da fórmula de Taylor aplicada à função $g(x) = e^{itx}$:

$$e^{itx} = 1 + itx + \theta_1(x)\frac{t^2x^2}{2}, \text{ onde } |\theta_1(x)| \leq 1$$

e

$$e^{itx} = 1 + itx - \frac{t^2x^2}{2} + \theta_2(x)\frac{t^3x^3}{6}, \text{ onde } |\theta_2(x)| \leq 1.$$

Prova do TCL de Lindeberg

Seja $\epsilon > 0$. Usando a primeira fórmula para $|x| > \epsilon$ e a segunda para $|x| \leq \epsilon$, podemos escrever e^{itx} da seguinte forma geral:

$$e^{itx} = 1 + itx - \frac{t^2x^2}{2} + r_\epsilon(x),$$

onde

$$r_\epsilon(x) = \begin{cases} (1 + \theta_1(x))\frac{t^2x^2}{2} & \text{se } |x| > \epsilon, \\ \theta_2(x)\frac{t^3x^3}{6} & \text{se } |x| \leq \epsilon. \end{cases}$$

Prova do TCL de Lindeberg

Portanto,

$$
\begin{aligned}
E(e^{it\frac{X_k-\mu_k}{s_n}}) &= \int e^{it\frac{x-\mu_k}{s_n}}\,dF_k(x) \\
&= \int (1 + it\frac{x-\mu_k}{s_n} - \frac{t^2(\frac{x-\mu_k}{s_n})^2}{2} + r_\epsilon(\frac{x-\mu_k}{s_n}))dF_k(x) \\
&= 1 + itE(\frac{X_k-\mu_k}{s_n}) - \frac{t^2}{2}E((\frac{X_k-\mu_k}{s_n})^2) + \\
&+ \frac{t^2}{2}\int_{|x-\mu_k|>\epsilon s_n}(1+\theta_1(\frac{x-\mu_k}{s_n}))(\frac{x-\mu_k}{s_n})^2 dF_k(x) + \\
&\frac{t^3}{6}\int_{|x-\mu_k|\le\epsilon s_n}\theta_2(\frac{x-\mu_k}{s_n})(\frac{x-\mu_k}{s_n})^3 dF_k(x).
\end{aligned}
$$

Prova do TCL de Lindeberg

Como $EX_k = \mu_k$ e $Var(X_k) = \sigma_k^2$, temos

$$E(e^{it\frac{X_k - \mu_k}{s_n}}) = 1 - \frac{t^2 \sigma_k^2}{2s_n^2} + e_{n,k},$$

onde o resto $e_{n,k}$ satisfaz

$$\begin{aligned}
|e_{n,k}| &\leq t^2 \int_{|x-\mu_k|>\epsilon s_n} (\frac{x-\mu_k}{s_n})^2 dF_k(x) \\
&+ \frac{|t^3|}{6} \int_{|x-\mu_k|\leq\epsilon s_n} \epsilon(\frac{x-\mu_k}{s_n})^2 dF_k(x) \\
&\leq \frac{t^2}{s_n^2} \int_{|x-\mu_k|>\epsilon s_n} (x-\mu_k)^2 dF_k(x) \\
&+ \frac{\epsilon|t^3|}{6s_n^2} \int_{-\infty}^{\infty} (x-\mu_k)^2 dF_k(x).
\end{aligned}$$

Prova do TCL de Lindeberg

Temos então,

$$\sum_{k=1}^{n} |e_{n,k}| \leq \frac{t^2}{s_n^2} \sum_{k=1}^{n} \int_{|x-\mu_k|>\epsilon s_n} (x-\mu_k)^2 dF_k(x) + \frac{\epsilon |t^3|}{6}.$$

Pela condição de Lindeberg, a primeira parcela do termo à direita tende a zero quando $n \to \infty$. Logo, para n suficientemente grande,

$$\sum_{k=1}^{n} |e_{n,k}| \leq \frac{\epsilon |t|^3}{3}.$$

Prova do TCL de Lindeberg

Vamos então escolher uma sequência de ϵ's que converge para zero. Para $\epsilon = \frac{1}{m}$, existe n_m tal que para $n \geq n_m$,

$$\sum_{k=1}^{n} |e_{n,k}| \leq \frac{|t^3|}{3m}, \tag{1}$$

onde os restos $e_{n,k}$ são os determinados pela fórmula baseada em $\epsilon = \frac{1}{m}$. Portanto, existe uma sequência de inteiros positivos $n_1 < n_2 < \ldots$ tal que (1) é satisfeita para $n_m \leq n < n_{m+1}$, onde para estes valores de n os restos são baseados em $\epsilon = \frac{1}{m}$. É importante lembrar durante o restante da prova que o valor de ϵ que determina o resto $e_{n,k}$ depende da posição de n em relação aos n_m. Temos, então,

$$\sum_{k=1}^{n} |e_{n,k}| \to 0 \text{ quando } n \to \infty.$$

Prova do TCL de Lindeberg

Como X_i's são independentes,

$$\phi_{\frac{S_n - ES_n}{s_n}}(t) = \prod_{k=1}^{n} E(e^{it\frac{X_k - \mu_k}{s_n}}) = \prod_{k=1}^{n}(1 - \frac{t^2\sigma_k^2}{2s_n^2} + e_{n,k}).$$

Para provar que o termo à direita converge para $e^{\frac{-t^2}{2}}$, usaremos o seguinte Lema sobre números complexos.

Prova do TCL de Lindeberg

Lema 1

Sejam $c_{n,k}$ números complexos tais que $\sum_{k=1}^{n} c_{n,k} \to c$ quando $n \to \infty$. Se

$$\max_{1 \leq k \leq n} |c_{n,k}| \to 0 \text{ quando } n \to \infty$$

e

$$\sum_{k=1}^{n} |c_{n,k}| \leq M < \infty,$$

onde M é uma constante que não depende de n, então

$$\prod_{k=1}^{n} (1 + c_{n,k}) \to e^{c} \text{ quando } n \to \infty.$$

Demonstração.

Nós omitimos a prova deste lema que pode ser encontrada no livro do Chung seção 7.1. □

Prova do TCL de Lindeberg

Em nosso caso, sejam $c_{n,k} = -\frac{t^2\sigma_k^2}{2s_n^2} + e_{n,k}$ e $c = \frac{-t^2}{2}$. Temos que

$$\sum_{k=1}^{n} |c_{n,k}| \leq \frac{t^2}{2} + \sum_{k=1}^{n} |e_{n,k}| \to \frac{t^2}{2},$$

logo existe $M < \infty$ tal que $\forall n$, $\sum_{k=1}^{n} |c_{n,k}| < M$. Para aplicar o lema resta verificar a condição sobre o máximo

$$\max_{1\leq k\leq n} |c_{n,k}| \leq \max_{1\leq k\leq n} \frac{t^2\sigma_k^2}{2s_n^2} + \max_{1\leq k\leq n} |e_{n,k}|$$
$$\leq \max_{1\leq k\leq n} \frac{t^2}{2}\frac{\sigma_k^2}{s_n^2} + \sum_{k=1}^{n} |e_{n,k}|.$$

Como já provamos que os dois termos acima tendem a zero, a prova está terminada.

Exemplo do TCL de Lindeberg

Seja $\{X_n : n \geq 1\}$ uma sequência de variáveis i.i.d. com média 0 e variância 1. Também, seja $\{Y_n : n \geq 1\}$ uma sequência de variáveis independentes com

$$P(Y_n = \pm n) = \frac{1}{2n^2} \text{ e } P(Y_n = 0) = 1 - \frac{1}{n^2}, n \geq 1.$$

Sendo X_n e Y_n independentes para $n \geq 1$, temos $\frac{1}{\sqrt{n}} \sum_{k=1}^{n} (X_k + Y_k) \xrightarrow{D} N(0, 1)$, mas a condição de Lindeberg não está satisfeita.

Exemplo do TCL de Lindeberg

Solução: Pelo TCL para variáveis i.i.d., temos que $\frac{1}{\sqrt{n}}\sum_{k=1}^n X_k \xrightarrow{D} N(0,1)$, vamos provar que $\frac{1}{\sqrt{n}}\sum_{k=1}^n Y_k \xrightarrow{P} 0$. Deste modo o resultado segue por Slutsky. Pela desigualdade de Markov, temos

$$P(|\frac{1}{\sqrt{n}}\sum_{k=1}^n Y_k| > \epsilon) \leq \frac{E|\sum_{k=1}^n Y_k|}{\epsilon\sqrt{n}} \leq \frac{\sum_{k=1}^n E|Y_k|}{\epsilon\sqrt{n}} = \frac{\sum_{k=1}^n 1/k}{\epsilon\sqrt{n}} \xrightarrow{n\to\infty} 0,$$

(onde o último limite pode ser visto pelo fato de que usando o teste da integral para séries pode-se provar que $\frac{1}{\log n}\sum_{k=1}^n 1/k \xrightarrow{n\to\infty} 1$). Logo, $\frac{1}{\sqrt{n}}\sum_{k=1}^n Y_k \xrightarrow{P} 0$. Como $Var(X_k + Y_k) = Var(X_k) + Var(Y_k) = 2$, temos que se a condição de Lindeberg fosse satisfeita, teríamos $\frac{1}{\sqrt{n}}\sum_{k=1}^n (X_k + Y_k) \xrightarrow{D} N(0,2)$. Logo, a condição de Lindeberg não é satisfeita, caso contrário teríamos uma contradição.

TCL de Liapunov

Teorema 3

Teorema Central do Limite de Liapunov. *Sejam* $X_1, X_2, \ldots$ *variáveis aleatórias independentes tais que* $EX_n = \mu_n$ *e* $VarX_n = \sigma_n^2 < \infty$ *com pelo menos um* $\sigma_j^2 > 0$. *Seja* $S_n = X_1 + \ldots + X_n$ *e* $s_n^2 = VarS_n$. *Se existir* $m > 0$ *tal que*

$$\frac{1}{s_n^{2+m}} \sum_{k=1}^{n} E(|X_k - \mu_k|^{2+m}) \to 0 \text{ quando } n \to \infty,$$

então,

$$\frac{S_n - ES_n}{s_n} \to^D N(0,1).$$

Prova do TCL de Liapunov

Para provar este teorema, é suficiente verificar que as condições do Teorema de Liapunov implicam as condições do Teorema de Lindeberg. A condição de Lindeberg estabelece uma integral na região $|x-\mu_k| > \epsilon s_n, \epsilon > 0$. Nessa região, temos que $\frac{|x-\mu_k|}{\epsilon s_n} > 1$, o que por sua vez implica $\frac{|x-\mu_k|^m}{\epsilon^m s_n^m} > 1$. Desse modo, temos que:

$$
\begin{aligned}
&\frac{1}{s_n^2}\sum_{k=1}^{n}\int_{|x-\mu_k|>\epsilon s_n}(x-\mu_k)^2 dF_k(x)\\
&\leq \frac{1}{s_n^2}\sum_{k=1}^{n}\int_{|x-\mu_k|>\epsilon s_n}(x-\mu_k)^2\frac{|x-\mu_k|^m}{\epsilon^m s_n^m} dF_k(x)\\
&= \frac{1}{\epsilon^m s_n^{2+m}}\sum_{k=1}^{n}\int_{|x-\mu_k|>\epsilon s_n}|x-\mu_k|^{2+m} dF_k(x)\\
&\leq \frac{1}{\epsilon^m s_n^{2+m}}\sum_{k=1}^{n}\int_{-\infty}^{\infty}|x-\mu_k|^{2+m} dF_k(x)\\
&= \frac{1}{\epsilon^m s_n^{2+m}}\sum_{k=1}^{n}E|X_k-\mu_k|^{2+m}.
\end{aligned}
$$

Prova do TCL de Liapunov

Mas a condição de Liapunov implica que o último termo tende a zero quando $n \to \infty$. Portanto, a condição de Lindeberg está satisfeita.

Resultado Auxiliar

Antes de verificarmos um exemplo do Teorema Central do Limite de Liapunov, vamos considerar o seguinte Lema.

Lema 2

Para $\lambda > 0$*,*

$$\frac{1}{n^{\lambda+1}} \sum_{k=1}^{n} k^{\lambda} \to \frac{1}{\lambda+1},$$

quando $n \to \infty$*, de maneira que* $\sum_{k=1}^{n} k^{\lambda}$ *é da ordem de* $n^{\lambda+1}$*.*

Prova do Lema

Como $x^\lambda \leq k^\lambda$ se $k-1 \leq x \leq k$, e $k^\lambda \leq x^\lambda$ se $k \leq x \leq k+1$, segue-se que

$$\begin{aligned}\int_{k-1}^{k} x^\lambda dx &\leq \int_{k-1}^{k} k^\lambda dx = k^\lambda \\ &= \int_{k}^{k+1} k^\lambda dx \leq \int_{k}^{k+1} x^\lambda dx,\end{aligned}$$

somando-se em k de 1 até n, temos

$$\int_0^n x^\lambda dx \leq \sum_{k=1}^{n} k^\lambda \leq \int_1^{n+1} x^\lambda dx.$$

Prova do Lema

Logo,

$$\frac{n^{\lambda+1}}{\lambda+1} \leq \sum_{k=1}^{n} k^{\lambda} \leq \frac{(n+1)^{\lambda+1}-1}{\lambda+1} \leq \frac{(n+1)^{\lambda+1}}{\lambda+1},$$

o que é equivalente a

$$\frac{1}{\lambda+1} \leq \frac{1}{n^{\lambda+1}} \sum_{k=1}^{n} k^{\lambda} \leq \frac{1}{\lambda+1} \cdot \left(\frac{n+1}{n}\right)^{\lambda+1}.$$

Como $(\frac{n+1}{n})^{\lambda+1} \to 1$ quando $n \to \infty$, o lema está provado.

Exemplo 1

Sejam $X_1, X_2, \ldots,$ independentes, $X_n \sim U[-n, n]$. Prove que $\frac{S_n - ES_n}{s_n} \to^D N(0,1)$.
Solução: Vamos verificar a condição de Liapunov para $\delta = 1$. Temos

$$E|X_k - \mu_k|^3 = E|X_k|^3 = \frac{1}{2k}\int_{-k}^{k} |x|^3 dx$$
$$= \frac{1}{k}\int_0^k x^3 dx = \frac{k^3}{4}.$$

Logo, o Lema anterior implica que $\sum_{k=1}^n E|X_k - \mu_k|^3$ é da ordem de n^4. Vamos determinar a ordem de s_n^3. Como $\mu_k = EX_k = 0$ e

$$\sigma_k^2 = Var(X_k) = EX_k^2 = \frac{1}{2k}\int_{-k}^{k} x^2 dx = \frac{k^2}{3}, \text{ temos}$$

$$s_n^2 = \sum_{k=1}^n \frac{k^2}{3}.$$

Exemplo 1

Portanto, aplicando o resultado do Lema, temos:

$$\frac{s_n^2}{n^3} \to \frac{1}{9}.$$

Então,

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{s_n^3} \sum_{k=1}^{n} E|X_k - \mu_k|^3 \\
&= \lim_{n\to\infty} \left(\frac{n^{9/2}}{s_n^3} \cdot \frac{\sum_{k=1}^{n} E|X_k - \mu_k|^3}{n^4} \cdot \frac{1}{n^{1/2}} \right) \\
&= 9^{3/2} \cdot \frac{1}{16} \cdot \lim_{n\to\infty} \frac{1}{n^{1/2}} = 0.
\end{aligned}$$

Exemplo 2

Sejam X_n, $n \geq 1$, variáveis independentes com

$$P(X_n = \pm 2^n) = 2^{-n-1} \text{ e } P(X_n = \pm 1) = \frac{1}{2}(1 - 2^{-n}), n \geq 1.$$

Verifique que $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \xrightarrow{D} N(0,1)$.

Exemplo 2

Solução: Defina $Y_n = X_n I_{[|X_n| \leq n]}$. Deste modo, $P(Y_n = \pm 1) = \frac{1}{2}(1 - 2^{-n})$ e $P(Y_n = 0) = 2^{-n}$. Vamos verificar que Y_n satisfaz a condição de Liapunov para $m = 1$. Temos que $EY_n = 0$, $Var(Y_n) = EY_n^2 = (1 - 2^{-n})$, e $E|Y_n|^3 = (1 - 2^{-n}) = Var(Y_n)$. Logo, $s_n^2 = \sum_{k=1}^n Var(Y_k) = \sum_{k=1}^n (1 - 2^{-k}) = n - \frac{\frac{1}{2} - (\frac{1}{2})^{n+1}}{\frac{1}{2}}$. Portanto,

$$\frac{1}{s_n^3} \sum_{k=1}^n E|Y_k|^3 = \frac{1}{s_n^3} \sum_{k=1}^n Var(Y_k) = \frac{1}{s_n} = \frac{1}{n - \frac{\frac{1}{2} - (\frac{1}{2})^{n+1}}{\frac{1}{2}}} \xrightarrow{n \to \infty} 0.$$

O Teorema Central do Limite de Liapunov implica que,

$$\frac{1}{\sqrt{n - \frac{\frac{1}{2} - (\frac{1}{2})^{n+1}}{\frac{1}{2}}}} \sum_{k=1}^n Y_k \xrightarrow{D} N(0, 1).$$

Portanto,

$$\frac{\sqrt{n}}{\sqrt{n - \frac{\frac{1}{2} - (\frac{1}{2})^{n+1}}{\frac{1}{2}}}} \frac{1}{\sqrt{n}} \sum_{k=1}^n Y_k \xrightarrow{D} N(0, 1).$$

Exemplo 2

Como $\frac{\sqrt{n}}{\sqrt{n - \frac{\frac{1}{2} - (\frac{1}{2})^{n+1}}{\frac{1}{2}}}} \xrightarrow{n\to\infty} 1$, temos que

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} Y_k \xrightarrow{D} N(0,1).$$

Seja $Z_n = X_n - Y_n$. Então,

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} Y_k + \frac{1}{\sqrt{n}} \sum_{k=1}^{n} Z_k.$$

Se conseguirmos provar que $\frac{1}{\sqrt{n}} \sum_{k=1}^{n} Z_k \xrightarrow{P} 0$, então o resultado segue por Slustky.
Mas $P(Z_n = \pm 2^n) = 2^{-n-1}$ e $P(Z_n = 0) = 1 - 2^{-n}$. Como
$P(|Z_n| > \frac{1}{k}) = P(|Z_n| = 2^n) = 2^{-n}$, temos que

$$\sum_{n=1}^{\infty} P(|Z_n| > \frac{1}{k}) = \sum_{n=1}^{\infty} 2^{-n} < \infty, \forall k \geq 1.$$

Portanto, $Z_n \to 0$ cp1, ou seja, $P(\{w \in \Omega : \lim_{n\to\infty} Z_n(w) = 0\}) = 1$.

Exemplo 2

Como

$$
\begin{aligned}
&\lim_{n\to\infty} Z_n(w) = 0 \Leftrightarrow \forall\epsilon > 0, \exists N \text{ tal que } |Z_n(w)| < \epsilon, \forall n \geq N\\
&\Rightarrow \exists N \text{ tal que } |Z_n(w)| < 1, \forall n \geq N\\
&\Rightarrow \exists N \text{ tal que } |Z_n(w)| = 0, \forall n \geq N\\
&\Rightarrow |\sum_{i=1}^{\infty} Z_i(w)| < \infty\\
&\Rightarrow \lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i(w) = 0,
\end{aligned}
$$

temos que $\{w \in \Omega : \lim_{n\to\infty} Z_n(w) = 0\} \subseteq \{w \in \Omega : \lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i(w) = 0\}$. Logo, $P(\{w \in \Omega : \lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i(w) = 0\}) = 1$, o que por sua vez implica que, $\frac{1}{\sqrt{n}} \sum_{k=1}^{n} Z_k \xrightarrow{P} 0$.

TCL Multivariado

Concluímos dizendo que o Teorema Central do Limite também pode ser estendido ao caso de vetores aleatórios. Neste caso, tem-se que a distribuição da média amostral centrada converge para uma distriuição normal multivariada. A seguir, nós enunciamos formalmente o teorema sem prová-lo.

Teorema 4

Seja $\vec{X}_1, \vec{X}_2, \ldots$ uma sequência de vetores aleatórios k-dimensionais, independentes e identicamente distribuídos. Suponha que $\vec{X}_1$ tenha variância finita, e sejam $\vec{\mu}$ a média e Σ a matriz de covariância de $\vec{X}_1$. Seja $\overline{X}_n$ a média amostral, definida como a média aritmética dos vetores $\vec{X}_1, \ldots, \vec{X}_n$. Então,

$$\sqrt{n}(\overline{X}_n - \vec{\mu}) \to^{D} N(\vec{0}, \Sigma), \text{ quando } n \to \infty.$$

Método Delta

O método Delta é um resultado que aumenta significativamente a relevância do Teorema Central do Limite. Antes de enunciarmos o teorema, vamos provar dois lemas. Dizemos que uma sequência de variáveis aleatórias $\{Y_n\}$ é *limitada em probabilidade* se para todo $\epsilon > 0$, existir K e n_0 tal que $P(|Y_n| \leq K) > 1 - \epsilon$ para todo $n > n_0$.

Lema 3

Se $\{Y_n\}$ converge em distribuição para uma variável aleatória com função de distribuição H, então a sequência é limitada em probabilidade.

Prova do Lema 1

Fixemos K_1 e $-K_2$ pontos de continuidade de H tal que $H(K_1) > 1 - \epsilon/4$ e $H(-K_2) < \epsilon/4$. Escolhamos n_0 tal que, $\forall n > n_0$,

$$H_n(K_1) > H(K_1) - \epsilon/4 > 1 - \epsilon/2$$

e

$$H_n(-K_2) < H(-K_2) + \epsilon/4 < \epsilon/2.$$

Então,

$$P(-K_2 \leq Y_n \leq K_1) \geq H_n(K_1) - H_n(K_2) > 1 - \epsilon.$$

O resultado está provado se escolhermos

$$K = \max(|K_1|, |K_2|).$$

Lema 2

Lema 4

Se $\{Y_n\}$ é limitada em probabilidade e $X_n = o(Y_n)$, então $X_n \to^P 0$.

Demonstração.

Dados quaisquer $\epsilon > 0$ e $\delta > 0$, precisamos mostrar que existe N tal que $P(|X_n| > \epsilon) < \delta$ para todo $n \geq N$. Como $\{Y_n\}$ é limitada em probabilidade, existe K e n_1 tal que $P(|Y_n| \leq K) > 1 - \delta$ para todo $n \geq n_1$. Como $X_n = o(Y_n)$, sabemos que existe n_2 tal que $\frac{|X_n|}{|Y_n|} < \frac{\epsilon}{K}$ para todo $n \geq n_2$. Façamos $N = \max(n_1, n_2)$, então para $n \geq N$, $|X_n| > \epsilon \Rightarrow |Y_n| > K$. Logo

$$P(|X_n| > \epsilon) \leq P(|Y_n| > K) < \delta.$$

$\square$

Método Delta

Teorema 5

Se $\sqrt{n}(T_n - \theta) \to^D N(0, \tau^2)$, *então*

$$\sqrt{n}[f(T_n) - f(\theta)] \to^D N(0, \tau^2[f'(\theta)]^2), \qquad (2)$$

desde que $f'(\theta)$ *exista e não seja zero.*

Prova do Método Delta

Utilizaremos a versão da série de Taylor em torno de $T_n = \theta$ que diz que:

$$f(T_n) = f(\theta) + (T_n - \theta)f'(\theta) + o(T_n - \theta),$$

e então

$$\sqrt{n}[f(T_n) - f(\theta)] = \sqrt{n}(T_n - \theta)f'(\theta) + o(\sqrt{n}(T_n - \theta)).$$

O primeiro termo do lado direito converge em distribuição para $N(0, \tau^2[f'(\theta)]^2)$. Por outro lado, como $\sqrt{n}(T_n - \theta)$ converge em distribuição, pelo Lema 3, temos que $\sqrt{n}(T_n - \theta)$ é limitada em probabilidade. Então pelo Lema 4, $o(\sqrt{n}(T_n - \theta))$ converge para zero em probabilidade. O resultado portanto é uma consequência do Teorema de Slutsky.

Observação

Este teorema pode parecer uma surpresa, já que se X é distribuído normalmente, a distribuição de $f(X)$, por exemplo, $1/X$, $\log X$, ou e^X não será tipicamente normal. A explicação para este paradoxo aparente pode ser encontrada na prova. Como $o(T_n - \theta) \to^P 0$, nós estamos quase certos que quando n for grande, $f(T_n)$ é aproximadamente linear, e uma função linear de uma variável normal é também normal. O processo de aproximar a diferença $f(T_n) - f(\theta)$ pela função linear $(T_n - \theta)f'(\theta)$ e o limite em (2) é chamado de *método delta*.

Exemplo

Para estimar p^2, suponha que temos a escolha entre

(a) n ensaios bernoulli com probabilidade p^2 de sucesso; ou

(b) n ensaios bernoulli com probabilidade p de sucesso.

Sejam X e Y o número de sucessos no primeiro e segundo tipo de ensaios, e suponha que como estimadores de p^2 nos dois casos, nós usaríamos X/n e $(Y/n)^2$, respectivamente. Então nós temos:

$$\sqrt{n}(\frac{X}{n} - p^2) \to^D N(0, p^2(1-p^2))$$

e

$$\sqrt{n}((\frac{Y}{n})^2 - p^2) \to^D N(0, p(1-p)4p^2).$$

Exemplo

Então, pelo menos para n grande, X/n será mais acurado que $(Y/n)^2$, desde que

$$p^2(1-p^2) < p(1-p)4p^2.$$

Dividindo ambos os lados por $p^2(1-p)$, podemos ver que $\frac{X}{n}$ ou $\frac{Y^2}{n^2}$ é preferível se $p > 1/3$ ou $p < 1/3$, respectivamente.

Transformações que Estabilizam Variância

O método delta proporciona a base para derivar transformações que estabilizam a variância, ou seja, transformações que levem a uma variância assintótica que é independente do parâmetro. Suponha, por exemplo, que $X_1, \ldots, X_n$ são variáveis Poisson com parâmetro λ. Segue do Teorema Central do Limite que

$$\sqrt{n}(\overline{X} - \lambda) \to N(0, \lambda).$$

Para problemas de inferência que se referem a λ, é quase sempre inconveniente que λ ocorre não somente na esperança mas também na variância da distribuição limite.

Transformações que Estabilizam Variância

É portanto de interesse achar uma função f para a qual $\sqrt{n}[f(\overline{X}) - f(\lambda)]$ tende em distribuição para $N(0, c^2)$, onde c^2 não depende de λ. Em geral, suponha que $\sqrt{n}(T_n - \theta) \to^D N(0, \tau^2(\theta))$. Então, pelo método delta:

$$\sqrt{n}[f(T_n) - f(\theta)] \to^D N(0, \tau^2(\theta)(f')^2(\theta)),$$

desde que a derivada de f exista em θ e seja diferente de 0. A distribuição limite do lado direito terá portanto variância constante c^2 se $f'(\theta) = \frac{c}{\tau(\theta)}$. A transformação resultante é dita ser estabilizadora de variância.

Exemplo - Poisson

Exemplo 3

No caso de Poisson, temos $\theta = \lambda$ e $\tau(\theta) = \sqrt{\lambda}$. Logo,

$$f'(\lambda) = \frac{c}{\sqrt{\lambda}} \text{ ou } f(\lambda) = 2c\sqrt{\lambda}.$$

Fazendo $c = 1$, temos que

$$2\sqrt{n}(\sqrt{\overline{X}} - \sqrt{\lambda}) \to^{D} N(0,1).$$

Exemplo - Chi-quadrado

Exemplo 4

Chi-Quadrado. Seja $Y_i = X_i^2$, onde as X_i's são i.i.d. $N(0, \sigma^2)$. Então, $EY_i = \sigma^2$ e $VarY_i = 2\sigma^4$ e pelo Teorema Central do Limite, temos

$$\sqrt{n}(\overline{Y} - \sigma^2) \to^D N(0, 2\sigma^4),$$

ou seja, $T_n = \overline{Y}$, $\theta = \sigma^2$, e $\tau^2(\theta) = 2\theta^2$. Logo,

$$f'(\theta) = \frac{c}{\sqrt{2\theta}} \text{ ou } f(\theta) = \frac{c}{\sqrt{2}} \log \theta.$$

Fazendo $c = 1$, vemos que

$$\sqrt{\frac{n}{2}} \log(\frac{\overline{Y}}{\sigma^2}) \to^D N(0, 1).$$