

CONVERGÊNCIAS E TEOREMA DO LIMITE CENTRAL.

Vanderlei da Costa Bueno

Instituto de Matemática e Estatística

Universidade de São Paulo, SP. Brasil

Outubro de 2020

Convergência em média quadrática.

Os espaços vetoriais L^p , $p \geq 1$ são espaços de variáveis aleatórias X tais que $E[|X|^p] < \infty$, sem distinção entre variáveis aleatórias X e Y que são iguais quase certamente, isto é $P(X = Y) = 1$. Sob a norma $\|X\|_p = (E[|X|^p])^{\frac{1}{p}}$ é um espaço completo em que toda sequência de Cauchy converge.

Pela desigualdade de Jensen, se $1 \leq p \leq q$, $\|X\|_p \leq \|X\|_q$ e $L^q \subseteq L^p$. Em nossas notas estudamos o caso particular em que $p = 2$.

Convergência em média quadrática.o

Definição 1

Seja $(X_n)_{n \geq 1}$ sequência de v.a.(s) tal que $E[|X_n|^2] < \infty, n \geq 1$.

Dizemos que a sequência converge em média quadrática para uma variável aleatória X , com $E[|X|^2] < \infty$, e denotamos por

$X_n \rightarrow^{mq} X$, se

$$\lim_{n \rightarrow \infty} E[|X_n - X|^2] = 0.$$

Convergência em média quadrática.

A convergência em média quadrática tem um papel técnico na Teoria da Probabilidade, em situações quando a informação sobre a sequência de variáveis consiste das médias $E[X_n]$, $n \geq 1$ e das covariâncias $E[(X_n - \mu_{X_n})(X_m - \mu_{X_m})]$. Em tais casos não conseguimos, em geral, trabalhar com a função característica $E[e^{itX_n}]$ ou avaliar quantidades tais como $P(X_n \geq \varepsilon)$ que são ferramentas para analisar convergência em distribuição e convergência quase certa, respectivamente.

Convergência em média quadrática.

Ao contrário, os momentos de segunda ordem são suficientes para determinar se $(X_n)_{n \geq 1}$ converge em média quadrática ou não. Enfim, podemos utilizar o critério de Cauchy para provar a convergência

Teorema 1

Uma sequência $(X_n)_{n \geq 1}$ sequência de v.a.(s) tal que $E[|X|^2] < \infty$ converge em média quadrática para uma variável aleatória X se, e somente se

$$\lim_{n,m \rightarrow \infty} E[|X_n - X_m|^2] = 0.$$

Convergência em média quadrática.

Prova Observe o desenvolvimento quadrático:

$$|X_n - X_m|^2 = |X_n - X|^2 + |X_m - X|^2 + 2|X_n - X||X_m - X|.$$

Pela desigualdade de Schwarz, se $X_n \xrightarrow{mq} X$ temos

$$\lim_{n,m \rightarrow \infty} E[|X_n - X||X_m - X|] \leq \lim_{n,m \rightarrow \infty} \sqrt{E[|X_n - X|^2]E[|X_m - X|^2]} = 0.$$

Assim, sob tal hipótese

$$\lim_{n,m \rightarrow \infty} E[|X_n - X_m|^2] = \lim_{n,m \rightarrow \infty} E[|X_n - X|^2 + |X_m - X|^2 + 2|X_n - X||X_m - X|] = 0.$$

Por outro lado, se aceitamos que $\lim_{n,m \rightarrow \infty} E[|X_n - X_m|^2] = 0$ concluímos também, pelo desenvolvimento quadrático acima, que $\lim_{n \rightarrow \infty} E[|X_n - X|^2] = 0$.

Convergência em média quadrática.

Observação 1

Procede da Desigualdade de Markov que

$$P(|X_n - X| > \varepsilon) \leq \frac{E[|X_n - X|^2]}{\varepsilon^2}.$$

e concluímos que a convergência em média quadrática implica em convergência em probabilidade, por sua vez, implica em convergência em distribuição. Como indica o exemplo seguinte a convergência quase certa não implica em convergência em probabilidade.

Convergência em média quadrática.

Exemplo 1

Seja uma sequência $(X_n)_{n \geq 1}$ com $P(X_n = 2^n) = \frac{1}{n(n+1)}$ e $P(X_n = 0) = 1 - \frac{1}{n(n+1)}$ para todo n .

Temos que

$$\sum_{n=1}^{\infty} P(|X_n| > \frac{1}{m}) = \sum_{n=1}^{\infty} P(X_n = 2^n) = \sum_{n=1}^{\infty} \frac{1}{n(n+1)} < \infty$$

e portanto $X_n \rightarrow^{qc} 0$.

Contudo

$$\lim_{n \rightarrow \infty} E[|X_n - 0|^2] = \lim_{n \rightarrow \infty} E[X_n^2] = \lim_{n \rightarrow \infty} \frac{2^{2n}}{n(n+1)} = \infty$$

e $X_n \not\rightarrow^{mq} 0$.

Convergência em média quadrática.

Apesar do exemplo acima, existem situações em que a implicação é verdadeira:

Teorema 2

Se $(X_n)_{n \geq 1}$ sequência de v.a.(s) e X é uma variável aleatória tal que $X_n \xrightarrow{qc} X$ e $|X_n| \leq Y$, com $E[Y^2] < \infty$, então $X_n \xrightarrow{mq} X$

Convergência em média quadrática.

Prova Se $X_n \xrightarrow{qc} X$, $|X_n - X|^2 \xrightarrow{qc} 0$. Se $|X_n| \leq Y$, q.c. para todo $n \geq 1$ temos também que $|X| \leq Y$, q.c. e então $|X_n - X|^2 \leq 4Y^2$, q.c.

Portanto, pelo teorema a convergência dominada temos

$$\lim_{n \rightarrow \infty} E[|X_n - X|^2] = E[\lim_{n \rightarrow \infty} |X_n - X|^2] = 0.$$

O teorema do limite central.

Consideremos uma sequência de variáveis aleatórias independentes $(X_n)_{n \geq 1}$ definidas no mesmo espaço de probabilidade $(\Omega, \mathfrak{F}, P)$ e seja $(S_n)_{n \geq 1}$ a sequência das somas parciais definidas por $S_n = \sum_{i=1}^n X_i$.

O problema do limite central trata da convergência em distribuição das somas parciais normalizadas

$$\frac{S_n - E[S_n]}{\sqrt{\text{Var}(S_n)}} \rightarrow^D N(0, 1).$$

No capítulo anterior analisamos uma solução quando as variáveis aleatórias na sequência $(X_n)_{n \geq 1}$ são independentes e identicamente distribuídas, com média μ e variância σ^2 finita:

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \rightarrow^D N(0, 1).$$

O teorema do limite central.

Enunciamos o Teorema do limite Central de Lindeberg que dá condições gerais para a validade da convergência (para a prova, ver Barry James).

Teorema de Lindeberg

Seja $(X_n)_{n \geq 1}$ uma sequência de variáveis aleatórias independentes com $E[X_n] = \mu_n$ e $Var(X_n) = \sigma_n^2 < \infty$ e pelo menos um $\sigma_n^2 > 0$.
Sejam $S_n = \sum_{i=1}^n X_i$ e $s_n^2 = Var(S_n) = \sum_{i=1}^n \sigma_i^2$. Se a condição de Lindeberg, isto é,

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

é satisfeita, então

$$\frac{S_n - E[S_n]}{\sqrt{Var(S_n)}} \rightarrow^D N(0, 1).$$

O teorema do limite central.

Observação

A notação $\int_{\{|x-\mu_k|>\varepsilon s_n\}}$ significa que o cálculo da integral é feito na região $(-\infty, \mu_k - \varepsilon s_n) \cup (\mu_k + \varepsilon s_n, \infty)$. Se X_k for discreta com função de probabilidade $p_k(x_i)$, então

$$\int_{\{|x-\mu_k|>\varepsilon s_n\}} (x - \mu_k)^2 dF_k(x) = \sum_{\{i:|x_i-\mu_k|>\varepsilon s_n\}} (x_i - \mu_k)^2 p_k(x_i).$$

O teorema do limite central.

Por outro lado se X_k for absolutamente contínua com função densidade de probabilidade $f_k(x)$, temos

$$\int_{\{|x-\mu_k|>\varepsilon s_n\}} (x - \mu_k)^2 dF_k(x) = \\ \int_{-\infty}^{\mu_k - \varepsilon s_n} (x - \mu_k)^2 f_k(x) dx + \int_{\mu_k + \varepsilon s_n}^{\infty} (x - \mu_k)^2 f_k(x) dx.$$

O teorema do limite central.

Observe que

$$\sigma_k^2 = \int_{\{|x-\mu_k|>\varepsilon s_n\}} (x-\mu_k)^2 dF_k(x) + \int_{\{|x-\mu_k|\leq\varepsilon s_n\}} (x-\mu_k)^2 dF_k(x),$$

de modo que a condição de Lindeberg pode ser escrita da seguinte forma

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

O teorema do limite central.

A condição de Lindeberg significa, basicamente, que as parcelas $\frac{X_k - \mu_k}{s_n}$ da soma $\frac{S_n - E[S_n]}{s_n}$ são uniformemente pequenas para n grande. Por exemplo, a condição de Lindeberg implica

$$\lim_{n \rightarrow \infty} \max_{1 \leq k \leq n} \frac{\sigma_k^2}{s_n^2} = 0,$$

ou seja, para n grande, as variâncias das parcelas são uniformemente pequenas em relação à variância da soma.

O teorema do limite central.

Para ver isto, observe que Para todo k ,

$$\frac{\sigma_k^2}{s_n^2} = \frac{1}{s_n^2} \int_{\{|x-\mu_k|>\varepsilon s_n\}} (x - \mu_k)^2 dF_k(x) +$$
$$\frac{1}{s_n^2} \int_{\{|x-\mu_k|\leq\varepsilon s_n\}} (x - \mu_k)^2 dF_k(x) \leq$$

$$\frac{1}{s_n^2} \sum_{j=1}^n \int_{\{|x-\mu_j|>\varepsilon s_n\}} (x - \mu_j)^2 dF_j(x) + \frac{1}{s_n^2} \int_{\{|x-\mu_k|\leq\varepsilon s_n\}} \varepsilon^2 s_n^2 dF_k(x) \leq$$

$$\frac{1}{s_n^2} \sum_{j=1}^n \int_{\{|x-\mu_j|>\varepsilon s_n\}} (x - \mu_j)^2 dF_j(x) + \frac{1}{s_n^2} \int_{-\infty}^{\infty} \varepsilon^2 s_n^2 dF_k(x).$$

O teorema do limite central.

O último termo não depende de k pois a última parcela é igual a ε^2 . Portanto temos

$$\max_{1 \leq k \leq n} \frac{\sigma_k^2}{s_n^2} \leq \varepsilon^2 + \frac{1}{s_n^2} \sum_{j=1}^n \int_{\{|x - \mu_j| > \varepsilon s_n\}} (x - \mu_j)^2 dF_j(x)$$

que converge para ε^2 , pela condição de Lindeberg. Como vale para todo $\varepsilon > 0$, temos o resultado

$$\lim_{n \rightarrow \infty} \max_{1 \leq k \leq n} \frac{\sigma_k^2}{s_n^2} = 0.$$

O teorema do limite central.

Do ponto de vista intuitivo, isso serve para justificar a afirmação: a soma de um grande número de pequenas quantidades independentes e de médias zero tem aproximadamente a distribuição normal.

Observemos que a condição de Lindeberg é formalmente mais forte que a condição sobre o máximo das variâncias. Como $s_n^2 = \sum_{k=1}^n (x - \mu_k)^2 dF_k(x)$, a condição diz que, quando n é grande, é pequena a parte da variância da soma devida às "caudas" das X_k situadas a mais de ε desvios-padrão s_n das suas respectivas médias μ_k .

O teorema do limite central.

É interessante, porém, que na presença da condição sobre o máximo, a condição de Lindeberg torna-se necessária para a validade do teorema do limite central:

Teorema de Feller

Seja $(X_n)_{n \geq 1}$ uma sequência de variáveis aleatórias independentes com $E[X_n] = \mu_n$ e $Var(X_n) = \sigma_n^2 < \infty$ e pelo menos um $\sigma_n^2 > 0$.

Sejam $S_n = \sum_{i=1}^n X_i$ e $s_n^2 = Var(S_n) = \sum_{i=1}^n \sigma_i^2$.

Se

$$\lim_{n \rightarrow \infty} \max_{1 \leq k \leq n} \frac{\sigma_k^2}{s_n^2} = 0$$

e

$$\frac{S_n - E[S_n]}{\sqrt{Var(S_n)}} \rightarrow^D N(0, 1),$$

então

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

O teorema do limite central.

O Lema seguinte será utilizado nos exemplos e exercícios: **Lema 1**

Para $\lambda > 0$,

$$\lim_{n \rightarrow \infty} \frac{1}{n^{\lambda+1}} \sum_{k=1}^n k^{\lambda} = \frac{1}{\lambda + 1},$$

de maneira que $\sum_{k=1}^n k^{\lambda}$ é da ordem de $n^{\lambda+1}$.

O teorema do limite central.

Exemplo

Seja $(X_n)_{n \geq 1}$ uma sequência de variáveis aleatórias independentes com $X_n \sim U(-n, n)$. Verifiquemos a condição de Lindeberg:

Observa-se facilmente que $E[X_k] = 0$ e $Var(X_k) = \frac{k^2}{3}$ para todo k .

Consideremos a parcela

$$\int_{|x| > \varepsilon s_n} x^2 dF_k(x) = \frac{1}{2k} \int_{-k}^k x^2 \mathbf{1}_{\{|x| > \varepsilon s_n\}}(x) dx$$

e esta última integral é nula se $n < s_n$, pois neste caso, o integrando toma o valor zero em $(-k, k)$. Pelo lema acima temos

O teorema do limite central.

$s_n^2 = \sum_{k=1}^n \frac{k^2}{3}$ é tal que

$$\frac{1}{3} \frac{1}{n^3} \sum_{k=1}^n k^2 = \frac{1}{3} \frac{1}{3} = \frac{1}{9},$$

isto é, quando n é grande, $\frac{s_n^2}{n^3} = \frac{s_n^2}{n^2} \frac{1}{n} \approx \frac{1}{9}$ e concluímos que $\lim_{n \rightarrow \infty} \frac{s_n}{n} = \infty$.

Portanto,

$$\forall M > 0, \exists n_0 \mid \text{ se } n \geq n_0 \rightarrow \frac{s_n}{n} > M.$$

Basta tomar $M = \frac{1}{\varepsilon}$ e temos $\varepsilon s_n > n$.

O teorema do limite central.

O teorema do limite central para seqüências de variáveis aleatórias independentes e identicamente distribuídas, analisado no capítulo anterior, segue como um corolário do teorema de Lindeberg

Corolário 1

Seja $(X_n)_{n \geq 1}$ uma seqüência de variáveis aleatórias independentes e identicamente distribuídas com média $E[X_1] = \mu$ e $Var(X_1) = \sigma^2 < \infty$, então

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \rightarrow^D N(0, 1).$$

O teorema do limite central.

Prova

Observe que $s_n^2 = n\sigma^2$.

Verifiquemos a condição de Lindeberg, para todo $\varepsilon > 0$

$$\frac{1}{n\sigma^2} \sum_{k=1}^n \int_{\{|x-\mu| \leq \sigma\sqrt{n\varepsilon}\}} (x - \mu)^2 dF_k(x) =$$

$$\frac{1}{\sigma^2} \int_{\{|x-\mu| \leq \sigma\sqrt{n\varepsilon}\}} (x - \mu)^2 dF_1(x) \rightarrow_{n \rightarrow \infty}$$

$$\frac{1}{\sigma^2} \int_{-\infty}^{\infty} (x - \mu)^2 dF_1(x) = \frac{\sigma^2}{\sigma^2} = 1.$$

O teorema do limite central.

O Teorema de Liapunov, a seguir, é muito útil quando as variáveis aleatórias X_n possuem momentos finitos de ordem maior que 2. O teorema afirma que a convergência normal vale se as somas dos momentos centrais absolutos de ordem $2 + \delta$ é assintoticamente pequena em relação a $s_n^{2+\delta}$.

O teorema do limite central.

Teorema de Liapunov

Seja $(X_n)_{n \geq 1}$ uma sequência de variáveis aleatórias independentes tais que $E[X_n] = \mu_n$ e $Var(X_n) = \sigma_n^2 < \infty$, com pelo menos um $\sigma_n^2 > 0$. Seja $s_n^2 = \sum_{k=1}^n \sigma_k^2$. Se existir $\delta > 0$ tal que

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^{2+\delta}} \sum_{k=1}^n E[|X_k - \mu_k|^{2+\delta}] = 0,$$

então

$$\frac{S_n - E[S_n]}{\sqrt{Var(S_n)}} \rightarrow^D N(0, 1).$$

O teorema do limite central.

Exemplo

Seja $(X_n)_{n \geq 1}$ uma sequência de variáveis aleatórias independentes, $X_n \sim U(-n, n)$. Verifiquemos a condição de Liapunov:

Verifica-se facilmente que $E[X_k] = 0$ e $Var(X_k) = \frac{k^2}{3}$ para todo k .

Assim $s_n^2 = \sum_{k=1}^n \frac{k^2}{3}$ é tal que

$$\frac{1}{n^3} \sum_{k=1}^n \frac{k^2}{3} = \frac{11}{33} = \frac{1}{9},$$

temos que para n grande, $\frac{s_n^2}{n^3} = \frac{s_n^2}{n^2} \frac{1}{n} = \frac{1}{9}$. s_n é da ordem de $n^{\frac{3}{2}}$ e s_n^3 é da ordem de $n^{\frac{9}{2}}$.

O teorema do limite central.

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

e

$$\lim_{n \rightarrow \infty} \frac{1}{n^4} \sum_{k=1}^n \frac{k^3}{4} = \frac{1}{16}.$$

isto é, $\sum_{k=1}^n E[|X_k|^3]$ é da ordem de n^4 . Portanto

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^3} \sum_{k=1}^n E[|X_k|^3] =$$

$$\lim_{n \rightarrow \infty} \frac{n^{\frac{9}{2}}}{s_n^3} \frac{\sum_{k=1}^n E[|X_k|^3]}{n^4} \frac{n^4}{n^{\frac{9}{2}}} = \frac{27}{16} \cdot \lim_{n \rightarrow \infty} \frac{1}{n^{\frac{1}{2}}} = 0.$$