Relações entre noções de convergência

Gabriel de Freitas Curti

2024-11-07

1. Introdução: Teoria assintótica e convergência

A análise assintótica é um método fundamental em estatística e teoria das probabilidades, usado para aproximar comportamentos de distribuições à medida que certos parâmetros, como o tamanho da amostra, tendem ao infinito. Essa abordagem permite estudar limites adequados para sequências, sendo a mais comum a análise do comportamento das distribuições amostrais quando o tamanho da amostra cresce indefinidamente (Hansen, 2022).

Formalmente, uma sequência (x_n) tem um limite L, denotado como $\lim_{n\to\infty} x_n = L$ ou simplesmente $x_n \to L$ conforme $n \to \infty$, se, para todo $\epsilon > 0$, existe um índice N tal que, para todo $n \geq N$, a diferença $|x_n - L| < \epsilon$. Em termos intuitivos, isso significa que os elementos da sequência ficam cada vez mais próximos de L conforme n aumenta. Vale destacar que, se uma sequência possui um limite, ele é único.

Entretanto, quando se trata de sequências de variáveis aleatórias, a definição de convergência requer uma análise mais cuidadosa. Por exemplo, considere a média amostral $\bar{X}_n$. À medida que o tamanho da amostra n aumenta, a distribuição de $\bar{X}_n$ também se modifica. Assim, a questão central é: em que sentido podemos definir e descrever o limite de $\bar{X}_n$? E como caracterizamos essa convergência? (Hansen, 2022).

A teoria assintótica, portanto, investiga como as distribuições de funções de variáveis aleatórias se comportam à medida que o número de variáveis aumenta (DasGupta, 2011). Embora o conceito de uma amostra infinita seja uma idealização teórica, ele fornece aproximações valiosas para amostras de tamanho finito, ajudando a desenvolver resultados práticos em estatística (Cassella & Berger, 2001).

2. Tipos de Convergência

Convergência em Probabilidade

X_n converge para X em probabilidade, denotado por $X_n \xrightarrow{p} X$, se, para todo $\epsilon > 0$,

$$P(|X_n - X| > \epsilon) \to 0$$

à medida que $n \to \infty$.

Convergência em Distribuição

X_n converge para X em distribuição, denotado por $X_n \xrightarrow{d} X$, se, para todo t no qual a função de distribuição cumulativa (cdf) F de X é contínua,

$$\lim_{n\to\infty} F_n(t) = F(t).$$

Convergência para uma Massa Pontual

Uma variável aleatória X é considerada uma massa pontual em c se $P(X=c)=1$. Isso significa que X assume o valor c com certeza total. A função de distribuição (FDA) de X, denotada por $F(x)$, é definida como:

- $F(x)=0$ para todo $x<c$.
- $F(x)=1$ para todo $x\geq c$.

Convergência Quase-Certa

Dizemos que X_n converge quase certamente para X, escrito $X_n \xrightarrow{\text{q.c.}} X$, se

$$P\left(\{s: X_n(s) \to X(s)\}\right)=1.$$

Convergência em Momentos

Dizemos que $(X_n)_{n=1}^{\infty}$ converge em momentos $\mathcal{M}_r$, se

$$\lim_{n\to\infty} \mathbb{E}(|X_n - X|^r)=0.$$

Iremos denotar Convergência em momentos como $X_n \xrightarrow{\mathcal{M}_r} X$. Os casos mais importantes de convergência em momentos são:

- Para $r=1$, é dito que converge na média:

$$\lim_{n\to\infty} \mathbb{E}(|X_n - X|)=0.$$

- Para $r=2$, é dito que converge na média quadrática:

$$\lim_{n\to\infty} \mathbb{E}(|X_n - X|^2)=0.$$

3. Relações entre Noções de Convergência

Teorema: As seguintes relações são válidas:

- a. Se $X_n \xrightarrow{\text{a.s.}} X$, então $X_n \xrightarrow{p} X$.
- b. Se $X_n \xrightarrow{L^2} X$, então $X_n \xrightarrow{L^1} X$.
- c. Se $X_n \xrightarrow{L^1} X$, então $X_n \xrightarrow{p} X$.
- d. Se $X_n \xrightarrow{L^2} X$, então $X_n \xrightarrow{p} X$.
- e. Se $X_n \xrightarrow{p} X$, então $X_n \xrightarrow{d} X$.
- f. Se $X_n \xrightarrow{d} X$ e se $P(X=c)=1$ para algum número real c, então $X_n \xrightarrow{p} X$.

Em geral, nenhuma das implicações reversas é válida, exceto o caso especial em (f).

4. Provas

Prova para (a)

Convergência quase-certa implica convergência em probabilidade

Prova: Seja $A_n^\epsilon = \{z : |X_n(z) - X(z)| > \epsilon\}$ uma família de eventos. Tome ϵ arbitrário. Segue da definição que

$$\lim_{n\to\infty} P(|X_n - X| > \epsilon) = 0.$$

Logo, para $\epsilon > 0$, temos que

$$P\left(\limsup_{n\to\infty} A_n^\epsilon\right) = 0.$$

Pelo Lema de Fatou,

$$\limsup_{n\to\infty} P(A_n^\epsilon) \leq P\left(\limsup_{n\to\infty} A_n^\epsilon\right) = 0.$$

Logo,

$$\lim_{n\to\infty} P(A_n^\epsilon) = \lim_{n\to\infty} P\{z : |X_n(z) - X(z)| > \epsilon\} = 0,$$

que é a própria definição de convergência em probabilidade. Portanto, $X_n \overset{p}{\to} X$.

Convergência em probabilidade não implica convergência quase-certa

Contraexemplo: Suponha que temos um espaço amostral $S = [0, 1]$, com distribuição uniforme, onde sorteamos $s \sim U[0, 1]$ e definimos $X(s) = s$. Definimos a sequência como:

$$X_1(s) = s + I_{[0,1]}(s), \quad X_2(s) = s + I_{[0,1/2]}(s), \quad X_3(s) = s + I_{[1/2,1]}(s),$$

$$X_4(s) = s + I_{[0,1/3]}(s), \quad X_5(s) = s + I_{[1/3,2/3]}(s), \quad X_6(s) = s + I_{[2/3,1]}(s).$$

Agora, pode-se verificar que essa sequência converge em probabilidade, mas não quase certamente. Aproximadamente, o “pico de $1 + s$” torna-se menos frequente ao longo da sequência (permitindo a convergência em probabilidade), mas o limite não é bem definido. Para qualquer s, $X_n(s)$ alterna entre s e $1 + s$.

Prova para (b)

Convergência em média quadrática implica em convergência em média

Prova: Usando o fato de que, pela inequação de Cauchy-Schwarz, se X e $Y \in L^1$, então $XY \in L^1$ e

$$\mathbb{E}[|XY|] \leq \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}.$$

Podemos escrever que

$$\mathbb{E}[|X_n - X|] \leq \sqrt{\mathbb{E}[(X_n - X)^2]}.$$

Então, para $\mathbb{E}[|X_n - X|] \to 0$, é preciso que $\mathbb{E}[(X_n - X)^2] \to 0$.

Convergência em média não implica convergência em média quadrática

Contraexemplo: Considere o espaço de probabilidade $([0,1], \mathcal{B}([0,1]), P)$, onde P é a distribuição uniforme. A sequência

$$X_n = \sqrt{n} 1_{(0, \frac{1}{n})}$$

converge para zero quase certamente e em L^1, uma vez que $\mathbb{E}[X_n] = \frac{1}{\sqrt{n}}$, mas não em média quadrática, porque $\mathbb{E}[X_n^2] = 1$ para todo n.

Prova para (c)

Convergência em média implica convergência em probabilidade

Prova: Usando o fato de que, pela inequação de Chebyshev, seja $X \geq 0$. Seja $g > 0 \in \mathbb{R}^+$, uma função crescente, então para todo $a > 0$,

$$P(X \geq a) \leq \frac{\mathbb{E}[g(X)]}{g(a)}.$$

Para todo $\epsilon > 0$,

$$P(|X_n - X| \geq \epsilon) \leq \frac{1}{\epsilon} \mathbb{E}[|X_n - X|].$$

Então, $\mathbb{E}[|X_n - X|] \to 0$ implica que $P(|X_n - X| \geq \epsilon) \to 0$.

Convergência em probabilidade não implica convergência em média

Contraexemplo: Considere o espaço de probabilidade $([0,1], \mathcal{B}([0,1]), P)$, onde P é a distribuição uniforme. A sequência

$$X_n = n 1_{\left(0, \frac{1}{n}\right)}$$

converge para zero quase certamente e, portanto, em probabilidade e em distribuição, mas não em L^1, já que $\mathbb{E}[X_n] = nP\{X_n = n\} = 1$ para cada n.

Prova para (d)

Convergência em média quadrática implica convergência em probabilidade

Prova: Como consequência imediata da desigualdade de Markov, temos:

$$P(|X_n - X| > \epsilon) \leq \frac{\mathbb{E}[(X_n - X)^2]}{\epsilon^2},$$

em que o lado direito da equação converge para zero.

Convergência em probabilidade não implica convergência em média quadrática

Contraexemplo: Considere o espaço de probabilidade $([0,1], \mathcal{B}([0,1]), P)$, onde P é a distribuição uniforme. A sequência

$$X_n = \sqrt{n} 1_{\left(0, \frac{1}{n}\right)}$$

converge para zero quase certamente e em L^1, uma vez que $\mathbb{E}[X_n] = \frac{1}{\sqrt{n}}$, mas não em média quadrática, porque $\mathbb{E}[X_n^2] = 1$ para todo n.

Prova para (e)

Convergência em probabilidade implica convergência em distribuição

Prova: Seja t um ponto de continuidade de $F_X(t)$. Então, para $\epsilon > 0$ e para todo $n \in \mathbb{N}^+$,

$$F_{X_n}(t) = P(X_n \le t) = P(X_n \le t, |X_n - X| \le \epsilon) + P(X_n \le t, |X_n - X| > \epsilon).$$
$$= P(X \le t + \epsilon) + P(|X_n - X| > \epsilon).$$

Tomando $n \to \infty$ e $\epsilon \downarrow 0$, usando o resultado de convergência em probabilidade e que F_X é contínua em t, temos que $F_X \le \liminf F_{X_n}$. Conclui-se assim que $F_{X_n}(t) = F_X(t)$.

Convergência em distribuição não implica convergência em probabilidade

Contraexemplo: Seja X uma variável aleatória tal que $X \sim N(0,1)$. Defina $X_n = -X$ para $n = 1, 2, 3, \ldots$; logo, $X_n \sim N(0,1)$. X_n possui a mesma função de distribuição que X para todo n, portanto, trivialmente,

$$\lim_{n\to\infty} F_n(x) = F(x) \quad \text{para todo } x.$$

Portanto, $X_n \xrightarrow{d} X$. Mas

$$P(|X_n - X| > \epsilon) = P(|-X - X| > \epsilon) = P(2|X| > \epsilon) = 2P\left(X > \frac{\epsilon}{2}\right) \neq 0.$$

Assim, X_n não converge para X em probabilidade.

Prova para (f)

Se $X_n \xrightarrow{d} X$ e X é uma variável aleatória degenerada em c (isto é, $P(X = c) = 1$), então $X_n \xrightarrow{p} X$.

Prova: Seja, para $\epsilon > 0$ e para todo $n \in \mathbb{N}^+$,

$$P(|X_n - X| > \epsilon) = P(X_n < c - \epsilon) + P(X_n > c + \epsilon)$$
$$= F_{X_n}(c - \epsilon) + (1 - F_{X_n}(c + \epsilon))$$
$$\to F_c(c - \epsilon) + (1 - F_c(c + \epsilon))$$
$$= 0,$$

desde que $c - \epsilon$ e $c + \epsilon$ são pontos de continuidade de F.

5. Referências

- HANSEN, B. *Probability and Statistics for Economists*. Princeton: Princeton University Press, 2022.
- WASSERMAN, L. *All of Statistics: A Concise Course in Statistical Inference*. New York: Springer Science & Business Media, 2004.
- DASGUPTA, A. *Probability for Statistics and Machine Learning: Fundamentals and Advanced Topics*. New York: Springer Science & Business Media, 2011.
- RESNICK, S. *A Probability Path*. New York: Springer Science & Business Media, 2003.
- **Convergence Concepts**. Disponível em: https://cran.r-project.org/web/packages/ConvergenceConcepts/index.html (https://cran.r-project.org/web/packages/ConvergenceConcepts/index.html). Acesso em: [data de acesso: 04/13/2024].
- KARR, A.F. (1993). *Probability*. In: *Probability. Springer Texts in Statistics*. Springer, New York, NY. https://doi.org/10.1007/978-1-4612-0891-4_2 (https://doi.org/10.1007/978-1-4612-0891-4_2).
- CASELLA, G.; BERGER, R.L. 2002. *Statistical Inference*. Australia; Pacific Grove, CA: Thomson Learning.