

Convergência e o lema de Borel - Cantelli

Cristian Coletti

Primeiro Quadrimestre de 2026

Convergência Quase Certa e o Lema de Borel-Cantelli

Lembrando a f.d. das leis exponenciais,

$$P(X_n \geq c \log n) = e^{-\lambda c \log n} = \frac{1}{n^{\lambda c}},$$

que é o termo geral de uma série convergente se e somente se $c > \frac{1}{\lambda}$. Logo, o limite superior (3.3) tem probabilidade 0 se $c > \frac{1}{\lambda}$ e probabilidade 1 se $c \leq \frac{1}{\lambda}$.

O cálculo acima fornece um exemplo de sequência convergindo em probabilidade, mas não q.c.: a sequência $(\frac{1}{\log n} X_n)_n$ tende a zero em L^p , e portanto também em probabilidade, pois para todo $p > 0$,

$$\lim_{n \rightarrow \infty} E\left[\left(\frac{X_n}{\log n}\right)^p\right] = \lim_{n \rightarrow \infty} \frac{1}{(\log n)^p} E(X_1^p) = 0$$

(uma v.a. exponencial tem momentos finitos de todas as ordens).

No entanto, convergência q.c. não ocorre: como visto acima, com probabilidade 1

$$\frac{X_n}{\log n} \geq \epsilon$$

infinitas vezes assim que $\epsilon \leq \frac{1}{\lambda}$, de modo que convergência q.c. não pode ocorrer.

Podemos agora dar condições práticas que asseguram convergência q.c.

Proposição 3.6

Seja $(X_n)_n$ uma sequência de v.a.'s com valores em um espaço métrico (E, d) . Então $\lim_{n \rightarrow \infty} X_n = X$ q.c. se e somente se

$$P\left(\overline{\lim_{n \rightarrow \infty}} \{d(X_n, X) > \delta\}\right) = 0 \quad \text{para todo } \delta > 0. \quad (3.4)$$

Demonstração: Se $\lim_{n \rightarrow \infty} X_n = X$ q.c., então, com probabilidade 1, $d(X_n, X)$ pode ser maior que $\delta > 0$ para um número finito de índices no máximo, logo (3.4).

Reciprocamente, se (3.4) vale, então com probabilidade 1 $d(X_n, X) > \delta$ apenas para um número finito de índices, de modo que $\lim_{n \rightarrow \infty} d(X_n, X) \leq \delta$ e o resultado segue graças à arbitrariedade de δ . Juntamente com a Proposição 3.6, o Lema de Borel-Cantelli fornece um critério para convergência q.c.:

Observação 3.7

Se para todo $\delta > 0$ a série $\sum_{n=1}^{\infty} P(d(X_n, X) > \delta)$ converge (sem suposições de independência), então (3.4) vale e $X_n \xrightarrow[n \rightarrow \infty]{a.s.} X$.

Note que, em comparação, apenas $\lim_{n \rightarrow \infty} P(d(X_n, X) > \delta) = 0$ para todo $\delta > 0$ é necessário para ter convergência em probabilidade.

Na sequência, usaremos frequentemente o seguinte fato elementar muito útil.

Critério 3.8 (Critério da Sub-subsequência)

Seja $(x_n)_n$ uma sequência no espaço métrico (E, d) . Então $\lim_{n \rightarrow \infty} x_n = x$ se e somente se de toda subsequência $(x_{n_k})_k$ puder ser extraída uma nova subsequência convergindo para x .

Proposição 3.9

(a) Se $(X_n)_n$ converge para X em probabilidade, então existe uma subsequência $(X_{n_k})_k$ tal que $X_{n_k} \xrightarrow{k \rightarrow \infty} X$ q.c.

Proposição 3.9 (continuação)

(b) $(X_n)_n$ converge para X em probabilidade se e somente se toda subsequência $(X_{n_k})_k$ admite uma nova subsequência convergindo para X q.c.

Demonstração (a): Pela definição de convergência em probabilidade temos, para todo inteiro positivo k ,

$$\lim_{n \rightarrow \infty} P(d(X_n, X) > 2^{-k}) = 0.$$

Seja, para todo k , n_k um inteiro tal que $P(d(X_n, X) > 2^{-k}) \leq 2^{-k}$ para todo $n \geq n_k$.

Podemos assumir a sequência $(n_k)_k$ crescente.

Para $\delta > 0$ seja k_0 um inteiro tal que $2^{-k} \leq \delta$ para $k > k_0$.

Então, para $k > k_0$,

$$P(d(X_{n_k}, X) > \delta) \leq P(d(X_{n_k}, X) > 2^{-k}) \leq 2^{-k}$$

e a série $\sum_{k=1}^{\infty} P(d(X_{n_k}, X) > \delta)$ é somável pois $P(d(X_{n_k}, X) > \delta) < 2^{-k}$ eventualmente.

Pelo lema de Borel-Cantelli, $P(\overline{\lim}_{k \rightarrow \infty} [d(X_{n_k}, X) > \delta]) = 0$ e $X_{n_k} \rightarrow_{k \rightarrow \infty} X$ q.c. pela Proposição 3.6.

Convergência Quase Certa e o Lema de Borel-Cantelli

(b) A parte "somente se" segue de (a).

Reciprocamente, aproveitemos o Critério 3.8: vamos provar que de toda subsequência de $(P(d(X_n, X) \geq \delta))_n$ podemos extrair uma nova subsequência convergindo para 0.

Mas, por suposição, de toda subsequência $(X_{n_k})_k$ podemos extrair uma nova subsequência $(X_{n_{k_h}})_h$ tal que $X_{n_{k_h}} \xrightarrow[h \rightarrow \infty]{\text{a.s.}} X$, logo também $\lim_{h \rightarrow \infty} P(d(X_{n_{k_h}}, X) \geq \delta) = 0$, pois convergência q.c. implica convergência em probabilidade.

A Proposição 3.9, junto com o Critério 3.8, permite obter alguns insights valiosos sobre convergência em probabilidade.

- Para convergência em probabilidade, muitas propriedades válidas para convergência q.c. se mantêm.

Em particular, se $X_n \xrightarrow{P_{n \rightarrow \infty}} X$ e $\Phi : E \rightarrow G$ é uma função contínua (G outro espaço métrico), então também $\Phi(X_n) \xrightarrow{P_{n \rightarrow \infty}} \Phi(X)$.

De fato, de toda subsequência de $(X_n)_n$ pode ser extraída uma nova subsequência, digamos $(X_{n_k})_k$, convergindo para X q.c. e claro que $\Phi(X_{n_k}) \xrightarrow{\text{a.s.}_{k \rightarrow \infty}} \Phi(X)$.

Logo, para toda subsequência de $(\Phi(X_n))_n$ pode ser extraída uma nova subsequência convergindo q.c. para $\Phi(X)$ e a afirmação segue da Proposição 3.9.

De maneira bem similar, outras propriedades úteis da convergência em probabilidade podem ser obtidas.

Por exemplo, se $X_n \xrightarrow{P_{n \rightarrow \infty}} X$ e $Y_n \xrightarrow{P_{n \rightarrow \infty}} Y$, então também $X_n + Y_n \xrightarrow{P_{n \rightarrow \infty}} X + Y$.

- O limite q.c. é obviamente único: se Y e Z são dois limites q.c. da mesma sequência $(X_n)_n$, então $Y = Z$ q.c.

Vamos provar unicidade também para o limite em probabilidade, o que é menos imediato.

Assuma que $X_n \xrightarrow{P_{n \rightarrow \infty}} Y$ e $X_n \xrightarrow{P_{n \rightarrow \infty}} Z$.

Pela Proposição 3.9(a) podemos encontrar uma subsequência de $(X_n)_n$ convergindo q.c. para Y .

Esta subsequência obviamente ainda converge para Z em probabilidade e dela podemos extrair uma nova subsequência convergindo q.c. para Z .

Esta sub-subsequência converge q.c. para ambos Y e Z e portanto $Y = Z$ q.c.

- Os limites q.c. e em probabilidade coincidem: se $X_n \xrightarrow[n \rightarrow \infty]{\text{a.s.}} Y$ e $X_n \xrightarrow[n \rightarrow \infty]{P} Z$ então $Y = Z$ q.c.

- Convergência em L^p implica convergência q.c. para uma subsequência.

Proposição 3.10 (Sequências de Cauchy em Probabilidade)

Seja $(X_n)_n$ uma sequência de v.a.'s com valores no espaço métrico completo E e tal que para todo $\delta, \epsilon > 0$ existe um n_0 tal que

$$P(d(X_n, X_m) > \delta) \leq \epsilon \quad \text{para todo } n, m \geq n_0. \quad (3.5)$$

Então $(X_n)_n$ converge em probabilidade para alguma v.a. X com valores em E .

Demonstração: Para todo $k > 0$ seja n_k um índice tal que, para todo $m \geq n_k$,

$$P(d(X_{n_k}, X_m) \geq 2^{-k}) \leq 2^{-k}.$$

A sequência $(n_k)_k$ pode ser escolhida crescente, portanto

$$\sum_{k=1}^{\infty} P(d(X_{n_k}, X_{n_{k+1}}) \geq 2^{-k}) < +\infty,$$

e, pelo Lema de Borel-Cantelli, o evento $\overline{\lim_{k \rightarrow \infty}} \{d(X_{n_k}, X_{n_{k+1}}) \geq 2^{-k}\}$ tem probabilidade 0.

Fora de N temos $d(X_{n_k}, X_{n_{k+1}}) < 2^{-k}$ para todo k maior que algum k_0 e, para $\omega \in N^c$, $k \geq k_0$ e $m > k$,

$$d(X_{n_k}, X_{n_m}) \leq \sum_{i=k}^m 2^{-i} \leq 2 \cdot 2^{-k}.$$

Portanto, para $\omega \in N^c$, $(X_{n_k}(\omega))_k$ é uma sequência de Cauchy em E e converge para algum limite $X(\omega) \in E$.

Logo, a sequência $(X_{n_k})_k$ converge q.c. para alguma v.a. X .

Vamos deduzir que $X_n \xrightarrow{P}_{n \rightarrow \infty} X$: escolha primeiro um índice n_k como acima e grande o suficiente para que

$$\begin{aligned} P(d(X_n, X_{n_k}) \geq \frac{\delta}{2}) &\leq \frac{\epsilon}{2} && \text{para todo } n \geq n_k \\ P(d(X, X_{n_k}) \geq \frac{\delta}{2}) &\leq \frac{\epsilon}{2}. \end{aligned} \quad (3.6)$$

Um índice n_k com estas propriedades existe graças a (3.5) e pois

$$X_{n_k} \xrightarrow{P}_{n \rightarrow \infty} X.$$

Assim, para todo $n \geq n_k$,

$$P(d(X_n, X) \geq \delta) \leq P(d(X_n, X_{n_k}) \geq \frac{\delta}{2}) + P(d(X, X_{n_k}) \geq \frac{\delta}{2}) \leq \epsilon.$$

Leis Fortes dos Grandes Números

Nesta seção vemos que, sob suposições bastante fracas, se $(X_n)_n$ é uma sequência de v.a.'s independentes (ou ao menos não correlacionadas) e tendo esperança matemática finita b , então suas médias empíricas

$$\overline{X}_n := \frac{1}{n}(X_1 + \cdots + X_n)$$

convergem q.c. para b . Este tipo de resultado é uma **lei forte dos Grandes Números**, em oposição às leis fracas, que tratam de convergência em L^p ou em probabilidade.

Note que podemos assumir $b = 0$: caso contrário, se $Y_n = X_n - b$ as v.a.'s Y_n têm média 0 e, como $\overline{Y}_n = \overline{X}_n - b$, provar que $\overline{X}_n \xrightarrow[n \rightarrow \infty]{a.s.} b$ ou que $\overline{Y}_n \xrightarrow[n \rightarrow \infty]{a.s.} 0$ é a mesma coisa.

Teorema 3.11 (Lei Forte de Rajchman)

Seja $(X_n)_n$ uma sequência de v.a.'s duas a duas não correlacionadas com média comum b e variância finita e suponha que

$$\sup_{n \geq 1} \text{Var}(X_n) M < +\infty. \quad (3.8)$$

Então $\overline{X}_n \rightarrow_{n \rightarrow \infty} b$ q.c.

Demonstração: Seja $S_n X_1 + \cdots + X_n$ e assumamos $b = 0$.

Para todo $\delta > 0$, pela desigualdade de Chebyshev,

$$P(|\bar{X}_{n^2}| > \delta) \leq \frac{1}{\delta^2} \text{Var}(\bar{X}_{n^2}) = \frac{1}{\delta^2 n^4} \sum_{k=1}^{n^2} \text{Var}(X_k) \leq \frac{M}{\delta^2} \frac{1}{n^2}.$$

Como a série $\sum_{n=1}^{\infty} \frac{1}{n^2}$ é somável, pela Observação 3.7 a subsequência $(\bar{X}_{n^2})_n$ converge para 0 q.c.

Agora precisamos investigar o comportamento de $\bar{X}_n$ entre dois inteiros consecutivos da forma n^2 .

Com este objetivo, seja

$$D_n = \sup_{n^2 \leq k < (n+1)^2} |S_k - S_{n^2}|$$

(lembre que $S_k = X_1 + \cdots + X_k$), de modo que se $n^2 \leq k < (n+1)^2$

$$|\bar{X}_k| = \frac{|S_k|}{k} \leq \frac{|S_{n^2}| + D_n}{k} \leq \frac{1}{n^2} (|S_{n^2}| + D_n) = |\bar{X}_{n^2}| + \frac{1}{n^2} D_n.$$

Resta provar que $\frac{1}{n^2} D_n \rightarrow_{n \rightarrow \infty} 0$ q.c.

Isso seguirá assim que mostrarmos que o termo $n \mapsto P(|\frac{1}{n^2} D_n| > \delta)$ é somável e, para fazer isso, pensando na desigualdade de Markov, buscaremos estimativas do momento de segunda ordem de D_n .

Temos $D_n^2 \leq \sum_{n^2 \leq k < (n+1)^2} (S_k - S_{n^2})^2$, portanto

$$E(D_n^2) \leq \sum_{n^2 \leq k < (n+1)^2} E[(S_k - S_{n^2})^2]. \quad (3.9)$$

Como as X_n são centradas e não correlacionadas, para $n^2 \leq k < (n+1)^2$,

$$\begin{aligned} E[(S_k - S_{n^2})^2] &= E[(X_{n^2+1} + \cdots + X_k)^2] = \text{Var}(X_{n^2+1} + \cdots + X_k) \\ &= \sum_{i=n^2+1}^k \text{Var}(X_i) \leq [(n+1)^2 - n^2 - 1] \cdot M = 2nM \end{aligned}$$

e juntamente com (3.9)

$$E(D_n^2) \leq [(n+1)^2 - n^2 - 1] \cdot 2nM = 4n^2M$$

de modo que, para todo $\delta > 0$, pela desigualdade de Markov (2.28)

$$P\left(\left|\frac{1}{n^2}D_n\right| > \delta\right) \leq \frac{1}{\delta^2 n^4} E(D_n^2) \leq \frac{4M}{\delta^2} \frac{1}{n^2},$$

que é somável, completando a demonstração.

Note que, sob as hipóteses do Teorema 3.11 de Rajchman, pela desigualdade de Chebyshev,

$$\begin{aligned} P(|\bar{X}_n - b| \geq \delta) &\leq \frac{1}{\delta^2} \text{Var}(\bar{X}_n) \\ &= \frac{1}{\delta^2 n^2} (\text{Var}(X_1) + \cdots + \text{Var}(X_n)) \leq \frac{M}{\delta^2 n} \xrightarrow{n \rightarrow \infty} 0, \end{aligned}$$

de modo que a lei fraca, $\bar{X}_n \xrightarrow{P_{n \rightarrow \infty}} b$, é imediata e muito mais fácil de provar que a lei forte.

Enunciamos finalmente, sem demonstração, a mais célebre Lei dos Grandes Números.

Ela requer que as v.a.'s sejam independentes e identicamente distribuídas, mas as hipóteses de existência de momentos são mais fracas (as variâncias podem ser infinitas) e o enunciado é muito mais preciso.

Teorema 3.12 (Lei Forte de Kolmogorov)

Seja $(X_n)_n$ uma sequência de v.a.'s reais i.i.d. Então

(a) se as X_n são integráveis, então $\overline{X}_n \rightarrow_{n \rightarrow \infty} b = E(X)$ q.c.;
(continua)