Universidade de São Paulo
Instituto de Matemática e Estatística
Aluno: Guilherme Dias Vianna

Professor: Luis Fantozzi
Monitor: -
Disciplina: Verão de Estatística

Lista 3 (Convergência Estocástica)

(1) Considere o espaço de probabilidade $(\{0,1\}, 2^{\{0,1\}}, \mathbb{P})$, com $\mathbb{P}[\{0\}] = 1/2$. Considere a sequência de variáveis aleatórias $X_n = \mathbf{1}_{\{1\}}$ para todo $n \in \mathbb{N}$ e $X = \mathbf{1}_{\{0\}}$. Mostre que $X_n \xrightarrow{d} X$, mas que não há convergência em probabilidade de X_n a X.

Resolução:

Para todo $n \in \mathbb{N}$, temos $X_n = \mathbf{1}_{\{1\}}$ e $X = \mathbf{1}_{\{0\}}$. Logo,

$$X_n(0) = 0, \quad X_n(1) = 1, \qquad \text{e} \qquad X(0) = 1, \quad X(1) = 0.$$

Primeiro, calculemos as funções de distribuição. Para $c \in \mathbb{R}$,

$$F_{X_n}(c) = \mathbb{P}(X_n \leqslant c).$$

Como X_n assume apenas os valores 0 e 1, segue que

$$F_{X_n}(c) = \begin{cases} 0, & c < 0, \\ \mathbb{P}(X_n = 0) = \mathbb{P}(\{0\}) = \dfrac{1}{2}, & 0 \leqslant c < 1, \\ 1, & c \geqslant 1. \end{cases}$$

De modo análogo, como X também assume apenas os valores 0 e 1, com $\mathbb{P}(X = 0) = \mathbb{P}(\{1\}) = \frac{1}{2}$ e $\mathbb{P}(X = 1) = \mathbb{P}(\{0\}) = \frac{1}{2}$, obtemos

$$F_X(c) = \begin{cases} 0, & c < 0, \\ \dfrac{1}{2}, & 0 \leqslant c < 1, \\ 1, & c \geqslant 1. \end{cases}$$

Em particular, $F_{X_n}(c) = F_X(c)$ para todo $c \in \mathbb{R}$ e todo n, portanto:

$$\lim_{n\to\infty} F_{X_n}(c) = F_X(c),$$

logo $X_n \xrightarrow{d} X$.

Agora, verifiquemos a convergência em probabilidade. Para qualquer $\omega \in \Omega$,

$$|X_n(\omega) - X(\omega)| = \begin{cases} |0 - 1| = 1, & \omega = 0, \\ |1 - 0| = 1, & \omega = 1, \end{cases}$$

portanto $|X_n - X| = 1$ em todo Ω. Assim, para todo ϵ tal que $0 < \epsilon < 1$,

$$\{\omega : |X_n(\omega) - X(\omega)| > \epsilon\} = \Omega \quad \Longrightarrow \quad \mathbb{P}(|X_n - X| > \epsilon) = \mathbb{P}(\Omega) = 1$$

para todo n.

Em particular,

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \epsilon) = 1 \neq 0,$$

o que contradiz a definição de convergência em probabilidade. Logo, X_n não converge em probabilidade para X.

(2) Considere o espaço de probabilidade $([0,1], \mathcal{B}[0,1], \lambda)$, com λ a medida uniforme. Considere a sequência de variáveis aleatórias:

$$\begin{aligned} X_1 &= \mathbf{1}_{[0,1]} \\ X_2 &= \mathbf{1}_{[0,1]} + \mathbf{1}_{[0,1/2]} \\ X_3 &= \mathbf{1}_{[0,1]} + \mathbf{1}_{[1/2,1]} \\ X_4 &= \mathbf{1}_{[0,1]} + \mathbf{1}_{[0,1/3]} \\ X_5 &= \mathbf{1}_{[0,1]} + \mathbf{1}_{[1/3,2/3]} \\ X_6 &= \mathbf{1}_{[0,1]} + \mathbf{1}_{[2/3,1]} \\ X_7 &= \mathbf{1}_{[0,1]} + \mathbf{1}_{[0,1/4]} \\ &\vdots \end{aligned}$$

(a) Mostre que, para todo $\omega \in [0,1]$, a sequência $(X_n(\omega))_{n\in\mathbb{N}}$ **não converge**. Conclua que não X_n não converge quase certamente.

Resolução:

$$X_1 = \mathbf{1}_{[0,1]}.$$

Para cada $m \geqslant 2$ e $j \in \{1, \dots, m\}$, defina

$$I_{m,j} := \left[\frac{j-1}{m}, \frac{j}{m}\right].$$

Pelo padrão indicado na definição da sequência, os termos após X_1 percorrem, para cada m, todos os intervalos $I_{m,1}, \dots, I_{m,m}$. Em particular, se

$$N_m := 1 + \sum_{k=2}^{m} k = \frac{m(m+1)}{2},$$

então, para $m \geqslant 2$ e $j \in \{1, \dots, m\}$,

$$X_{N_{m-1}+j} = \mathbf{1}_{[0,1]} + \mathbf{1}_{I_{m,j}}.$$

Fixe $\omega \in [0,1]$. Para cada $m \geqslant 2$, existe ao menos um índice $j_m \in \{1, \dots, m\}$ tal que $\omega \in I_{m,j_m}$, pois

$$[0,1] = \bigcup_{j=1}^{m} I_{m,j}.$$

Logo,

$$X_{N_{m-1}+j_m}(\omega) = \mathbf{1}_{[0,1]}(\omega) + \mathbf{1}_{I_{m,j_m}}(\omega) = 1 + 1 = 2,$$

o que mostra que 2 ocorre infinitas vezes ao longo da sequência $(X_n(\omega))_n$.
Além disso, para cada $m \geqslant 3$, o ponto ω pertence a no máximo dois dos intervalos $I_{m,1}, \dots, I_{m,m}$ (apenas no caso em que ω é um ponto de fronteira $\omega = k/m$), de modo que existe algum $i_m \in \{1, \dots, m\}$ com $\omega \notin I_{m,i_m}$. Para esse i_m,

$$X_{N_{m-1}+i_m}(\omega) = \mathbf{1}_{[0,1]}(\omega) + \mathbf{1}_{I_{m,i_m}}(\omega) = 1 + 0 = 1,$$

isto é, 1 também ocorre infinitas vezes.
Portanto, existem duas subsequências $(X_{n_m}(\omega))_m$ e $\left(X_{n'_m}(\omega)\right)_m$ tais que

$$X_{n_m}(\omega) = 2 \quad \text{para todo } m, \qquad \text{e} \qquad X_{n'_m}(\omega) = 1 \quad \text{para todo } m.$$

Se $(X_n(\omega))_n$ tivesse limite, então todas as subsequências teriam o mesmo limite, o que é impossível pois as duas subsequências acima são constantes com valores distintos. Logo, $(X_n(\omega))_n$ não converge para nenhum $\omega \in [0,1]$.
Assim,
$$\{\omega \in [0,1] : X_n(\omega) \text{ não converge}\} = [0,1],$$
e como $\lambda([0,1]) = 1$, segue que X_n não converge quase certamente.

(b) Mostre que $X_n \xrightarrow{p} X_1$.

Resolução:

Note que, para $n \geqslant 2$, existe um par (m,j) com $m \geqslant 2$ tal que $n = N_{m-1} + j$ e
$$X_n = X_1 + \mathbf{1}_{I_{m,j}}.$$
Logo,
$$|X_n - X_1| = \mathbf{1}_{I_{m,j}}.$$
Fixe $\epsilon > 0$. Se $\epsilon \geqslant 1$, então $|X_n - X_1| \leqslant 1$ implica
$$\lambda\left(\{\omega : |X_n(\omega) - X_1(\omega)| > \epsilon\}\right) = 0.$$
Se $0 < \epsilon < 1$, então
$$\{\omega : |X_n(\omega) - X_1(\omega)| > \epsilon\} = \left\{\omega : \mathbf{1}_{I_{m,j}}(\omega) = 1\right\} = I_{m,j},$$
e portanto
$$\lambda\left(\{\omega : |X_n(\omega) - X_1(\omega)| > \epsilon\}\right) = \lambda(I_{m,j}) = \frac{1}{m}.$$
Para concluir o limite, dado $\delta > 0$, escolha $M \in \mathbb{N}$ tal que $M \geqslant 2$ e $\frac{1}{M} < \delta$. Se $n > N_M = \frac{M(M+1)}{2}$, então o índice n já está em algum bloco com $m \geqslant M+1$, de modo que
$$\lambda\left(\{\omega : |X_n(\omega) - X_1(\omega)| > \epsilon\}\right) = \frac{1}{m} \leqslant \frac{1}{M+1} < \delta.$$
Como $\delta > 0$ é arbitrário, obtemos
$$\lim_{n\to\infty} \lambda\left(\{\omega : |X_n(\omega) - X_1(\omega)| > \epsilon\}\right) = 0 \quad \text{para todo } \epsilon > 0,$$
isto é,
$$X_n \xrightarrow{p} X_1.$$

3 Seja X_n uma sequência de variáveis aelatórias, que converge em distribuição a uma variável aleatória X. Mostre que, se a função distribuição F_X de X é contínua, então:
$$\lim_{n\to\infty} \sup_{s\in(-\infty,\infty)} |F_{X_n}(s) - F_X(s)| = 0\,.$$

Dica: considere $k \in \mathbb{N}$ e use continuidade para encontrar pontos $s_1, s_2, \ldots, s_k$ tais que $F_X(s_j) = \frac{j}{k+1}$. Use esses pontos para limitar por cima o supremo.

Resolução:

$$F_n(s) := F_{X_n}(s) = \mathbb{P}(X_n \leqslant s), \qquad F(s) := F_X(s) = \mathbb{P}(X \leqslant s).$$

Como $X_n \xrightarrow{d} X$ e F é contínua, temos

$$\lim_{n\to\infty} F_n(s) = F(s) \quad \text{para todo } s \in \mathbb{R}.$$

Fixe $k \in \mathbb{N}$. Para $j = 1, \dots, k$, defina

$$s_j := \inf\left\{ s \in \mathbb{R} : \ F(s) \geqslant \frac{j}{k+1} \right\}.$$

Pelas propriedades básicas de função de distribuição, F é não-decrescente e satisfaz $\lim_{s\to-\infty} F(s) = 0$ e $\lim_{s\to+\infty} F(s) = 1$, logo o conjunto acima é não-vazio e limitado inferiormente, de modo que $s_j \in \mathbb{R}$ está bem-definido. Além disso, pela continuidade de F,

$$F(s_j) = \frac{j}{k+1} \qquad \text{para todo } j = 1, \dots, k.$$

De fato, por definição de ínfimo e monotonicidade, $F(s_j) \geqslant \frac{j}{k+1}$. Se $F(s_j) > \frac{j}{k+1}$, pela continuidade existiria $t < s_j$ com $F(t) > \frac{j}{k+1}$, contrariando a minimalidade de s_j. Portanto vale a igualdade. Em particular, como $\frac{j}{k+1} < \frac{j+1}{k+1}$ e F é não-decrescente, segue que

$$s_1 < s_2 < \cdots < s_k.$$

Defina ainda

$$M_n := \max_{1\leqslant j\leqslant k} |F_n(s_j) - F(s_j)|.$$

Como $F_n(s_j) \to F(s_j)$ para cada j e o máximo é sobre um conjunto finito, temos $M_n \to 0$.
Agora mostremos a estimativa uniforme. Fixe $s \in \mathbb{R}$.
Se $s < s_1$, então $F(s) \leqslant F(s_1) = \frac{1}{k+1}$ e, por monotonicidade de F_n, $F_n(s) \leqslant F_n(s_1)$. Logo,

$$F_n(s) - F(s) \leqslant F_n(s_1) = F(s_1) + (F_n(s_1) - F(s_1)) \leqslant \frac{1}{k+1} + M_n,$$

e também $F(s) - F_n(s) \leqslant F(s) \leqslant \frac{1}{k+1}$. Portanto,

$$|F_n(s) - F(s)| \leqslant \frac{1}{k+1} + M_n.$$

Se $s \geqslant s_k$, então $F(s) \geqslant F(s_k) = \frac{k}{k+1}$, logo $1 - F(s) \leqslant \frac{1}{k+1}$, e

$$F_n(s) - F(s) \leqslant 1 - F(s) \leqslant \frac{1}{k+1}.$$

Além disso, por monotonicidade, $F_n(s) \geqslant F_n(s_k)$, então

$$F(s) - F_n(s) \leqslant F(s) - F_n(s_k) \leqslant 1 - F_n(s_k) = 1 - F(s_k) + (F(s_k) - F_n(s_k)) \leqslant \frac{1}{k+1} + M_n.$$

Logo, novamente,

$$|F_n(s) - F(s)| \leqslant \frac{1}{k+1} + M_n.$$

Por fim, se $s \in [s_j, s_{j+1})$ para algum $j \in \{1, \dots, k-1\}$, então, por monotonicidade,

$$F_n(s_j) \leqslant F_n(s) \leqslant F_n(s_{j+1}) \quad \text{e} \quad F(s_j) \leqslant F(s) \leqslant F(s_{j+1}).$$

Assim,

$$F_n(s) - F(s) \leqslant F_n(s_{j+1}) - F(s) \leqslant (F_n(s_{j+1}) - F(s_{j+1})) + (F(s_{j+1}) - F(s)) \leqslant M_n + \frac{1}{k+1},$$

pois $F(s_{j+1}) - F(s) \leqslant F(s_{j+1}) - F(s_j) = \frac{1}{k+1}$.
Analogamente,

$$F(s) - F_n(s) \leqslant F(s) - F_n(s_j) \leqslant (F(s) - F(s_j)) + (F(s_j) - F_n(s_j)) \leqslant \frac{1}{k+1} + M_n.$$

Concluímos que, para todo $s \in \mathbb{R}$,

$$|F_n(s) - F(s)| \leqslant \frac{1}{k+1} + M_n,$$

e portanto

$$\sup_{s\in\mathbb{R}} |F_n(s) - F(s)| \leqslant \frac{1}{k+1} + M_n.$$

Agora, dado $\epsilon > 0$, escolha $k \in \mathbb{N}$ tal que $\frac{1}{k+1} < \frac{\epsilon}{2}$. Como $M_n \to 0$, existe $N \in \mathbb{N}$ tal que, para todo $n \geqslant N$, $M_n < \frac{\epsilon}{2}$. Logo, para todo $n \geqslant N$,

$$\sup_{s\in\mathbb{R}} |F_n(s) - F(s)| \leqslant \frac{1}{k+1} + M_n < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Como $\epsilon > 0$ é arbitrário,

$$\lim_{n\to\infty} \sup_{s\in\mathbb{R}} |F_{X_n}(s) - F_X(s)| = 0.$$

Comentário:

Acabamos de provar uma versão mais fraca do **Teorema de Glivenko-Cantelli**:
Este teorema afirma que, se $Y_1, Y_2, \ldots$ são iid com função de distribuição F e

$$\widehat{F}_n(s) := \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{Y_i \leqslant s\}},$$

então $\sup_{s\in\mathbb{R}} |\widehat{F}_n(s) - F(s)| \to 0$ quase certamente. O resultado acima tem a mesma forma de conclusão (convergência uniforme em norma sup), mas é um enunciado determinístico: se uma sequência de funções de distribuição converge pontualmente para uma F contínua, então a convergência é automaticamente uniforme. Em muitas demonstrações de Glivenko–Cantelli no caso em que F é contínua, aparece exatamente este argumento de discretização por quantis $\big(F(s_j) = \frac{j}{k+1}\big)$ mais monotonicidade para elevar convergência pontual a convergência uniforme; o conteúdo probabilístico de Glivenko–Cantelli está em garantir a convergência (quase certa) da função empírica $\widehat{F}_n$ para F. Além disso, Glivenko–Cantelli vale para F possivelmente descontínua, enquanto aqui a continuidade de F é essencial: com apenas convergência em distribuição (isto é, convergência pontual apenas nos pontos de continuidade), a convergência uniforme pode falhar quando F tem saltos.

④ No que segue, sejam X_n e Y_n duas sequências de variáveis aleatórias definidas num mesmo espaço de probabilidade, e a_n e b_n duas sequências de números reais. Mostre que:

1.) $X_n = O_P(1)$ se e somente para qualquer sequência de números reais $\epsilon_n \uparrow \infty$ monotônica tais que $\lim_{n\to\infty} \mathbb{P}[|X_n| > \epsilon_n] = 0$.

Resolução:

($\Rightarrow$)

Pela definição, $X_n = O_{\mathbb{P}}(1)$ significa que, para todo $\eta > 0$, existem $M > 0$ e $N \in \mathbb{N}$ tais que

$$\mathbb{P}\left(|X_n| > M\right) \leqslant \eta, \qquad \forall\, n \geqslant N.$$

Seja $(\varepsilon_n)_{n\in\mathbb{N}}$ uma sequência monótona com $\varepsilon_n \uparrow \infty$. Fixe $\eta > 0$ e tome M, N como acima. Como $\varepsilon_n \to \infty$, existe $N' \in \mathbb{N}$ tal que

$$\varepsilon_n \geqslant M, \qquad \forall\, n \geqslant N'.$$

Então, para todo $n \geqslant \max\{N, N'\}$,

$$\mathbb{P}\left(|X_n| > \varepsilon_n\right) \leqslant \mathbb{P}\left(|X_n| > M\right) \leqslant \eta.$$

Como $\eta > 0$ é arbitrário, conclui-se que $\mathbb{P}(|X_n| > \varepsilon_n) \to 0$.

($\Leftarrow$

Suponha que, para toda sequência monótona $\varepsilon_n \uparrow \infty$, vale

$$\mathbb{P}\left(|X_n| > \varepsilon_n\right) \xrightarrow[n\to\infty]{} 0.$$

Vamos provar que $X_n = O_{\mathbb{P}}(1)$ por contradição. Se X_n não fosse $O_{\mathbb{P}}(1)$, então existiria $\eta_0 > 0$ tal que

$$\forall\, M > 0,\ \forall\, N \in \mathbb{N},\ \exists\, n \geqslant N :\ \mathbb{P}\left(|X_n| > M\right) \geqslant \eta_0.$$

Escolha $n_1 \in \mathbb{N}$ tal que $\mathbb{P}(|X_{n_1}| > 1) \geqslant \eta_0$. Recursivamente, dado n_k, aplique a propriedade acima com $M = k+1$ e $N = n_k + 1$ para obter $n_{k+1} > n_k$ tal que

$$\mathbb{P}\left(|X_{n_{k+1}}| > k+1\right) \geqslant \eta_0.$$

Defina então $(\varepsilon_n)_n$ por blocos:

$$\varepsilon_n \coloneqq k \quad \text{se} \quad n_k \leqslant n < n_{k+1}.$$

Esta sequência é monótona não-decrescente e satisfaz $\varepsilon_n \uparrow \infty$. Além disso, para todo k,

$$\mathbb{P}\left(|X_{n_k}| > \varepsilon_{n_k}\right) = \mathbb{P}\left(|X_{n_k}| > k\right) \geqslant \eta_0,$$

logo $\mathbb{P}(|X_n| > \varepsilon_n)$ não pode convergir para 0, contradizendo a hipótese.
Portanto, $X_n = O_{\mathbb{P}}(1)$.

2.) Se $X_n = O_P(a_n)$ e $a_n \downarrow 0$, então $X_n = o_P(1)$.

Resolução:

Se $X_n = O_{\mathbb{P}}(a_n)$, então $\frac{X_n}{a_n} = O_{\mathbb{P}}(1)$. Fixe $\epsilon > 0$ e $\eta > 0$. Pela propriedade $O_{\mathbb{P}}(1)$, existem $M > 0$ e $K_1 \in \mathbb{N}$ tais que

$$\mathbb{P}\left(\left|\frac{X_n}{a_n}\right| > M\right) \leqslant \eta, \qquad \forall n \geqslant K_1.$$

Como $a_n \to 0$, existe $K_2 \in \mathbb{N}$ tal que $M|a_n| \leqslant \epsilon$ para todo $n \geqslant K_2$. Logo, para $n \geqslant \max\{K_1, K_2\}$,

$$\{|X_n| > \epsilon\} = \left\{\left|\frac{X_n}{a_n}\right| > \frac{\epsilon}{|a_n|}\right\} \subseteq \left\{\left|\frac{X_n}{a_n}\right| > M\right\},$$

e portanto

$$\mathbb{P}(|X_n| > \epsilon) \leqslant \mathbb{P}\left(\left|\frac{X_n}{a_n}\right| > M\right) \leqslant \eta.$$

Como $\eta > 0$ é arbitrário, conclui-se que $\mathbb{P}(|X_n| > \epsilon) \to 0$ para todo $\epsilon > 0$, isto é, $X_n = o_{\mathbb{P}}(1)$.

3.) Se $X_n = o_P(a_n)$, então $X_n = O_P(a_n)$.

Resolução:

Se $X_n = o_{\mathbb{P}}(a_n)$, então $\frac{X_n}{a_n} \xrightarrow{p} 0$. Fixe $\eta > 0$ e tome $\epsilon = 1$ na definição de convergência em probabilidade. Existe $K \in \mathbb{N}$ tal que, para todo $n \geqslant K$,

$$\mathbb{P}\left(\left|\frac{X_n}{a_n}\right| > 1\right) \leqslant \eta,$$

ou seja,

$$\mathbb{P}(|X_n| > 1 \cdot |a_n|) \leqslant \eta, \qquad \forall n \geqslant K.$$

Isso é exatamente $X_n = O_{\mathbb{P}}(a_n)$ (com $M = 1$).

4.) Se $X_n = O_P(a_n)$ e $Y_n = O_P(b_n)$ então $X_nY_n = O_P(a_nb_n)$. e $X_n + Y_n = O_P(\max\{a_n, b_n\})$.

Resolução:

Primeiro, provemos $X_nY_n = O_{\mathbb{P}}(a_nb_n)$. Fixe $\eta > 0$. Como $X_n = O_{\mathbb{P}}(a_n)$, existem $M_1 > 0$ e K_1 tais que

$$\mathbb{P}(|X_n| > M_1|a_n|) \leqslant \frac{\eta}{2}, \qquad \forall n \geqslant K_1.$$

Como $Y_n = O_{\mathbb{P}}(b_n)$, existem $M_2 > 0$ e K_2 tais que

$$\mathbb{P}(|Y_n| > M_2|b_n|) \leqslant \frac{\eta}{2}, \qquad \forall n \geqslant K_2.$$

Para $n \geqslant K := \max\{K_1, K_2\}$, vale a inclusão

$$\{|X_nY_n| > M_1M_2|a_nb_n|\} \subseteq \{|X_n| > M_1|a_n|\} \cup \{|Y_n| > M_2|b_n|\},$$

pois, se $|X_n| \leqslant M_1|a_n|$ e $|Y_n| \leqslant M_2|b_n|$, então $|X_nY_n| \leqslant M_1M_2|a_nb_n|$. Logo, pela sub-aditividade

$$\mathbb{P}(|X_nY_n| > M_1M_2|a_nb_n|) \leqslant \mathbb{P}(|X_n| > M_1|a_n|) + \mathbb{P}(|Y_n| > M_2|b_n|) \leqslant \eta,$$

para todo $n \geqslant K$. Portanto $X_nY_n = O_{\mathbb{P}}(a_nb_n)$.
Segundo, defina $c_n := \max\{|a_n|, |b_n|\}$. Para todo $n \geqslant K$, pela desigualdade triangular,

$$\{|X_n + Y_n| > (M_1 + M_2)c_n\} \subseteq \{|X_n| > M_1c_n\} \cup \{|Y_n| > M_2c_n\},$$

pois, se $|X_n| \leqslant M_1c_n$ e $|Y_n| \leqslant M_2c_n$, então $|X_n + Y_n| \leqslant |X_n| + |Y_n| \leqslant (M_1 + M_2)c_n$. Assim,

$$\mathbb{P}(|X_n + Y_n| > (M_1 + M_2)c_n) \leqslant \mathbb{P}(|X_n| > M_1c_n) + \mathbb{P}(|Y_n| > M_2c_n).$$

Como $c_n \geqslant |a_n|$ e $c_n \geqslant |b_n|$, tem-se $M_1c_n \geqslant M_1|a_n|$ e $M_2c_n \geqslant M_2|b_n|$, logo

$$\{|X_n| > M_1c_n\} \subseteq \{|X_n| > M_1|a_n|\}, \qquad \{|Y_n| > M_2c_n\} \subseteq \{|Y_n| > M_2|b_n|\}.$$

Portanto, para todo $n \geqslant K$,

$$\mathbb{P}(|X_n + Y_n| > (M_1 + M_2)c_n) \leqslant \mathbb{P}(|X_n| > M_1|a_n|) + \mathbb{P}(|Y_n| > M_2|b_n|) \leqslant \eta.$$

Como $\eta > 0$ é arbitrário, segue que

$$X_n + Y_n = O_{\mathbb{P}}(c_n) = O_{\mathbb{P}}(\max\{|a_n|, |b_n|\}).$$

5.) Se, para $r > 0$, $\mathbb{E}[|X_n|^r] < \infty$ para todo $n \in \mathbb{N}$ então $X_n = O_P(\|X_n\|_r)$. *Dica:* use a desigualdade de Markov.

Resolução:

Seja $r > 0$ e suponha $\mathbb{E}\left[|X_n|^r\right] < \infty$. Denote

$$\|X_n\|_r := \left(\mathbb{E}\left[|X_n|^r\right]\right)^{1/r}.$$

Pela desigualdade de Markov, para todo $t > 0$,

$$\mathbb{P}\left(|X_n| > t\right) \leqslant \frac{\mathbb{E}\left[|X_n|^r\right]}{t^r}.$$

Tomando $t = M\|X_n\|_r$ com $M > 0$, obtemos

$$\mathbb{P}\left(|X_n| > M\|X_n\|_r\right) \leqslant \frac{\mathbb{E}\left[|X_n|^r\right]}{M^r\|X_n\|_r^r} = \frac{1}{M^r}.$$

Dado $\eta > 0$, escolha $M \geqslant \eta^{-1/r}$. Então $\frac{1}{M^r} \leqslant \eta$ e, portanto,

$$\mathbb{P}\left(|X_n| > M\|X_n\|_r\right) \leqslant \eta \qquad \text{para todo } n \in \mathbb{N}.$$

Isto é exatamente $X_n = O_{\mathbb{P}}\left(\|X_n\|_r\right)$ (com $K = 1$).

5 Sejam $X_1, X_2, \ldots, X_n$ duas variáveis aleatórias definidas num mesmo espaço de probabilidade. Mostre que essas variáveis aleatórias são independentes se, e somente se, definindo o vetor aleatório $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)'$, a função caractéristica de $\boldsymbol{X}$ satisfaz:

$$\phi_{\boldsymbol{X}}(s) = \prod_{i=1}^{n} \phi_{X_i}(s_i), \quad \forall s \in \mathbb{R}^n.$$

Resolução:

$$\boldsymbol{X} := (X_1, \ldots, X_n), \qquad \boldsymbol{s} = (s_1, \ldots, s_n) \in \mathbb{R}^n, \qquad \phi_{\boldsymbol{X}}(\boldsymbol{s}) := \mathbb{E}\left[e^{i\boldsymbol{s}^\top \boldsymbol{X}}\right] = \mathbb{E}\left[e^{i\sum_{j=1}^{n} s_j X_j}\right].$$

Para cada j, denote $\phi_{X_j}(t) := \mathbb{E}\left[e^{itX_j}\right]$, $t \in \mathbb{R}$.

($\Rightarrow$) Suponha $X_1, \ldots, X_n$ independentes. Como independência é definida via σ-álgebras geradas, para cada j temos

$$\sigma\left(e^{is_jX_j}\right) \subseteq \sigma(X_j),$$

logo as variáveis complexas $e^{is_1X_1}, \ldots, e^{is_nX_n}$ são independentes. Em particular, como $|e^{is_jX_j}| = 1$, todas são integráveis e

$$\mathbb{E}\left[\prod_{j=1}^{n} e^{is_jX_j}\right] = \prod_{j=1}^{n} \mathbb{E}\left[e^{is_jX_j}\right].$$

Portanto,

$$\phi_{\boldsymbol{X}}(\boldsymbol{s}) = \mathbb{E}\left[e^{i\sum_{j=1}^{n} s_jX_j}\right] = \mathbb{E}\left[\prod_{j=1}^{n} e^{is_jX_j}\right] = \prod_{j=1}^{n} \phi_{X_j}(s_j), \qquad \forall\, \boldsymbol{s} \in \mathbb{R}^n.$$

($\Leftarrow$) Suponha agora que

$$\phi_{\boldsymbol{X}}(\boldsymbol{s}) = \prod_{j=1}^{n} \phi_{X_j}(s_j), \qquad \forall\, \boldsymbol{s} \in \mathbb{R}^n.$$

Seja μ_j a lei de X_j em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ e considere a medida produto

$$\boldsymbol{\mu} := \mu_1 \otimes \cdots \otimes \mu_n \quad \text{em} \quad (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)).$$

No espaço $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \boldsymbol{\mu})$, defina o vetor aleatório identidade

$$\boldsymbol{Y}(\boldsymbol{\omega}) := \boldsymbol{\omega}, \qquad \boldsymbol{\omega} \in \mathbb{R}^n.$$

Então $\boldsymbol{Y}$ tem lei $\boldsymbol{\mu}$, e as coordenadas $Y_1, \ldots, Y_n$ são independentes com leis marginais $\mu_1, \ldots, \mu_n$. Pelo sentido já provado,

$$\phi_{\boldsymbol{Y}}(\boldsymbol{s}) = \prod_{j=1}^{n} \phi_{Y_j}(s_j) = \prod_{j=1}^{n} \phi_{X_j}(s_j) = \phi_{\boldsymbol{X}}(\boldsymbol{s}), \qquad \forall\, \boldsymbol{s} \in \mathbb{R}^n.$$

Pela propriedade de identificação via função característica (nos slides: dois vetores aleatórios têm a mesma distribuição se, e somente se, suas funções características coincidem), conclui-se que $\boldsymbol{X}$ e $\boldsymbol{Y}$ têm a mesma lei, isto é,

$$\mathbb{P}_{\boldsymbol{X}} = \mathbb{P}_{\boldsymbol{Y}} = \boldsymbol{\mu}.$$

Em particular, para quaisquer $A_1, \ldots, A_n \in \mathcal{B}(\mathbb{R})$,

$$\mathbb{P}\left(X_1 \in A_1, \ldots, X_n \in A_n\right) = \mathbb{P}_{\boldsymbol{X}}(A_1 \times \cdots \times A_n) = \boldsymbol{\mu}(A_1 \times \cdots \times A_n) = \prod_{j=1}^{n} \mu_j(A_j) = \prod_{j=1}^{n} \mathbb{P}\left(X_j \in A_j\right),$$

logo $X_1, \ldots, X_n$ são independentes.

6. Sejam $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)'$ um vetor de variáveis aleatórias definidas no mesmo espaço de probabilidade. Dizemos que $\boldsymbol{X}$ segue distribuição normal com média $\boldsymbol{\mu} \in \mathbb{R}^n$ e matriz de variância-covariância $\Sigma \in \{A \in \mathbb{R}^{n \times n} : A \text{ positiva semidefinida}\}$, denotado por $\boldsymbol{X} \sim \mathcal{N}(\mu, \boldsymbol{\Sigma})$ se a medida induzida $\mathbb{P}_{\boldsymbol{X}}$ admite densidade (com respeito à medida de Lebesgue em $\mathbb{R}^n$) dada por:

$$f(\boldsymbol{s}) = \frac{1}{(2\pi)^{n/2}\sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2}(\boldsymbol{s} - \boldsymbol{\mu})'\Sigma(\boldsymbol{s} - \boldsymbol{\mu})\right).$$

Nesse caso, é possível mostrar que $\mathbb{E}[X_j] = \boldsymbol{\mu}_j$ e $\Sigma_{i,j} = \mathbb{C}(\boldsymbol{X}_i, \boldsymbol{X}_j)$ para todo $i, j = 1, \ldots, n$.

(a) Mostre que $X_1, \ldots, X_n$ são independentes se, e somente se, $\mathbb{C}(\boldsymbol{X}_i, \boldsymbol{X}_j) = 0$ para $i \neq j$. *Dica:* a função característica de $\boldsymbol{X}$ é $\boldsymbol{s} \mapsto \exp(i\boldsymbol{\mu}'\boldsymbol{s} - (1/2)\boldsymbol{s}'\Sigma\boldsymbol{s})$.

Resolução:

Como $\boldsymbol{X} \sim \boldsymbol{\mathcal{N}}(\boldsymbol{\mu}, \Sigma)$, temos $X_i \in L^2$

($\Rightarrow$) Se $X_1, \ldots, X_n$ são independentes, então, para $i \neq j$,

$$\mathbb{C}(X_i, X_j) = \mathbb{E}\left[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])\right] = \mathbb{E}[X_i - \mathbb{E}[X_i]]\, \mathbb{E}[X_j - \mathbb{E}[X_j]] = 0.$$

Logo $\Sigma_{ij} = 0$ para $i \neq j$.

($\Leftarrow$) Suponha agora que $\mathbb{C}(X_i, X_j) = 0$ para $i \neq j$, isto é, Σ é diagonal. Pela dica, a função característica conjunta de X é

$$\phi_{\boldsymbol{X}}(\boldsymbol{s}) = \mathbb{E}\left[e^{i\boldsymbol{s}^\top}\right] = \exp\left(i\boldsymbol{\mu}^\top\boldsymbol{s} - \frac{1}{2}\boldsymbol{s}^\top\Sigma\boldsymbol{s}\right), \qquad \boldsymbol{s} = (s_1, \ldots, s_n) \in \mathbb{R}^n,$$

Como Σ é diagonal,

$$\boldsymbol{s}^\top\Sigma\boldsymbol{s} = \sum_{j=1}^{n} \Sigma_{jj}s_j^2 \quad \Longrightarrow \quad \phi_{\boldsymbol{X}}(\boldsymbol{s}) = \exp\left(\sum_{j=1}^{n}\left(i\mu_j s_j - \frac{1}{2}\Sigma_{jj}s_j^2\right)\right) = \prod_{j=1}^{n} \exp\left(i\mu_j s_j - \frac{1}{2}\Sigma_{jj}s_j^2\right).$$

Por outro lado, para cada j, tomando $t \in \mathbb{R}$ e e_j o j-ésimo vetor da base canônica,

$$\phi_{X_j}(t) = \mathbb{E}\left[e^{itX_j}\right] = \mathbb{E}\left[e^{i(t\boldsymbol{e}_j)^\top \boldsymbol{X}}\right] = \phi_{\boldsymbol{X}}(t\boldsymbol{e}_j) = \exp\left(i\mu_j t - \frac{1}{2}\Sigma_{jj}t^2\right).$$

Portanto,

$$\phi_{\boldsymbol{X}}(\boldsymbol{s}) = \prod_{j=1}^{n} \phi_{X_j}(s_j), \qquad \forall s \in \mathbb{R}^n.$$

Pela caracterização de independência via fatorização da função característica (exercício anterior), conclui-se que $X_1, \ldots, X_n$ são independentes.

(b) A equivalência anterior também é verdadeira pra qualquer distribuição não normal?

Resolução:

Não. Em geral, $\mathbb{C}(X, Y) = 0$ não implica independência.
Por exemplo, seja Z com

$$\mathbb{P}(Z = 1) = \mathbb{P}(Z = -1) = \frac{1}{4}, \qquad \mathbb{P}(Z = 0) = \frac{1}{2},$$

e defina $X := Z$ e $Y := Z^2$. Então

$$\mathbb{E}[X] = \mathbb{E}[Z] = 1 \cdot \frac{1}{4} + (-1) \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} = 0, \qquad \mathbb{E}[Y] = \mathbb{E}[Z^2] = 1 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} = \frac{1}{2},$$

e

$$\mathbb{E}[XY] = \mathbb{E}[Z^3] = 1 \cdot \frac{1}{4} + (-1) \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} = 0,$$

logo

$$\mathbb{C}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0.$$

Mas X e Y não são independentes, pois

$$\mathbb{P}(X = 1, Y = 1) = \mathbb{P}(Z = 1) = \frac{1}{4} \neq \mathbb{P}(X = 1)\mathbb{P}(Y = 1) = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}.$$

Assim, a equivalência do item (a) é uma propriedade especial da família normal (mais precisamente, da normalidade conjunta).

7 Seja $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ uma sequência de vetores aleatórios k-dimensionais iid com $\mathbb{V}[\boldsymbol{X}_{1,l}] < \infty$, para $l = 1 \ldots, k$. Denote por $\boldsymbol{\mu} \in \mathbb{R}^k$ o vetor com j-ésima entrada dada por $\boldsymbol{\mu}_j = \mathbb{E}[\boldsymbol{X}_{1,j}]$ e Σ a matriz $k \times k$ com entrada (i, j) dada por $\Sigma_{i,j} = \mathbb{C}(\boldsymbol{X}_{1,i}, \boldsymbol{X}_{1,j})$ Use o dispositivo de Crámer-Wold para mostrar que, quando $n \to \infty$:

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{n} (\boldsymbol{X}_j - \boldsymbol{\mu}) \xrightarrow{d} \mathcal{N}(\boldsymbol{0}, \Sigma)$$

Resolução:

$$\boldsymbol{S}_n := \frac{1}{\sqrt{n}} \sum_{j=1}^{n} (\boldsymbol{X}_j - \boldsymbol{\mu}) \in \mathbb{R}^k.$$

Pelo dispositivo de Crámer–Wold, basta mostrar que, para todo $\boldsymbol{t} \in \mathbb{R}^k$,

$$\boldsymbol{t}^\top \boldsymbol{S}_n \xrightarrow{d} \boldsymbol{t}^\top \boldsymbol{Z},$$

onde $\boldsymbol{Z} \sim \mathcal{N}(0, \Sigma)$.
Fixe $\boldsymbol{t} = (t_1, \ldots, t_k) \in \mathbb{R}^k$ e defina as variáveis escalares

$$Y_j := \boldsymbol{t}'(\boldsymbol{X}_j - \boldsymbol{\mu}) = \sum_{\ell=1}^{k} t_\ell \, (X_{j,\ell} - \mu_\ell), \qquad j \in \mathbb{N}.$$

Como $(\boldsymbol{X}_j)_j$ é iid, $(Y_j)_j$ é iid. Além disso,

$$\mathbb{E}[Y_1] = \boldsymbol{t}^\top (\mathbb{E}[\boldsymbol{X}_1] - \boldsymbol{\mu}) = 0.$$

Para a variância, note primeiro que, para todo i, ℓ,

$$|\mathbb{C}(X_{1,i}, X_{1,\ell})| \leqslant \sqrt{\mathbb{V}[X_{1,i}]} \, \sqrt{\mathbb{V}[X_{1,\ell}]} < \infty,$$

logo Σ tem entradas finitas. Então

$$\mathbb{V}[Y_1] = \mathbb{E}\left[\left(\sum_{i=1}^{k} t_i (X_{1,i} - \mu_i)\right)\left(\sum_{\ell=1}^{k} t_\ell (X_{1,\ell} - \mu_\ell)\right)\right] = \sum_{i=1}^{k}\sum_{\ell=1}^{k} t_i t_\ell \, \mathbb{C}(X_{1,i}, X_{1,\ell}) = \boldsymbol{t}^\top \Sigma \boldsymbol{t}.$$

Em particular, $\mathbb{V}[Y_1] < \infty$. Pelo Teorema do Limite Central unidimensional,

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{n} Y_j \xrightarrow{d} \mathcal{N}\left(0, \boldsymbol{t}^\top \Sigma \boldsymbol{t}\right).$$

Como $\frac{1}{\sqrt{n}} \sum_{j=1}^{n} Y_j = \boldsymbol{t}^\top \boldsymbol{S}_n$, concluímos que

$$\boldsymbol{t}^\top \boldsymbol{S}_n \xrightarrow{d} \mathcal{N}\left(0, \boldsymbol{t}^\top \Sigma \boldsymbol{t}\right).$$

Por outro lado, se $\boldsymbol{Z} \sim \boldsymbol{\mathcal{N}}(0, \Sigma)$, então sua função característica é

$$\phi_{\boldsymbol{Z}(s)} = \exp\left(-\frac{1}{2}\, \boldsymbol{s}^\top \Sigma \boldsymbol{s}\right), \qquad \boldsymbol{s} \in \mathbb{R}^k.$$

Logo, para $u \in \mathbb{R}$,

$$\phi_{\boldsymbol{t}^\top \boldsymbol{Z}}(u) = \mathbb{E}\left[e^{iu\,\boldsymbol{t}^\top \boldsymbol{Z}}\right] = \mathbb{E}\left[e^{i(u\boldsymbol{t})^\top \boldsymbol{Z}}\right] = \phi_{\boldsymbol{Z}}(u\boldsymbol{t}) = \exp\left(-\frac{1}{2}\, u^2\, \boldsymbol{t}^\top \Sigma \boldsymbol{t}\right),$$

que é a função característica de $\mathcal{N}(0, \boldsymbol{t}^\top \Sigma \boldsymbol{t})$. Portanto,

$$\boldsymbol{t}^\top \boldsymbol{Z} \sim \mathcal{N}\left(0, \boldsymbol{t}^\top \Sigma \boldsymbol{t}\right).$$

Assim, para todo $\boldsymbol{t} \in \mathbb{R}^k$,

$$\boldsymbol{t}^\top \boldsymbol{S}_n \xrightarrow{d} \boldsymbol{t}^\top \boldsymbol{Z}.$$

Pelo dispositivo de Crámer–Wold, segue que

$$\boldsymbol{S}_n \xrightarrow{d} \boldsymbol{Z} \sim \boldsymbol{\mathcal{N}}(0, \Sigma),$$

isto é,

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{n} (\boldsymbol{X}_j - \boldsymbol{\mu}) \xrightarrow{d} \boldsymbol{\mathcal{N}}(0, \Sigma).$$

(8) Seja $(X_n)_n$ uma sequência de variáveis aleatórias. Suponha que $X_n \xrightarrow{d} c$, onde $c \in \mathbb{R}^k$ é uma **constante**. Mostre que $X_n \xrightarrow{p} c$.

Resolução:

$$\boldsymbol{X}_n \xrightarrow{d} \boldsymbol{c} \in \mathbb{R}^k \quad \text{(constante)}.$$

Considere a função

$$g : \mathbb{R}^k \to \mathbb{R}, \qquad g(\boldsymbol{x}) := \|\boldsymbol{x} - \boldsymbol{c}\|,$$

que é contínua em $\mathbb{R}^k$. Pelo **Teorema do Mapa Contínuo** (para convergência em distribuição),

$$g(\boldsymbol{X}_n) \xrightarrow{d} g(\boldsymbol{c}) = 0.$$

Defina $Y_n := g(\boldsymbol{X}_n) = \|\boldsymbol{X}_n - \boldsymbol{c}\|$. Então $Y_n \xrightarrow{d} 0$, onde 0 denota a variável aleatória constante e igual a 0.
Seja $\epsilon > 0$. Como $Y_n \geqslant 0$,

$$\mathbb{P}\left(\|\boldsymbol{X}_n - \boldsymbol{c}\| > \epsilon\right) = \mathbb{P}\left(Y_n > \epsilon\right) = 1 - \mathbb{P}\left(Y_n \leqslant \epsilon\right).$$

Denote por $F_{Y_n}(t) = \mathbb{P}(Y_n \leqslant t)$ a função de distribuição de Y_n, e por F_0 a função de distribuição da constante 0:

$$F_0(t) = \begin{cases} 0, & t < 0, \\ 1, & t \geqslant 0. \end{cases}$$

Pela definição de convergência em distribuição, para todo ponto de continuidade de F_0,

$$F_{Y_n}(t) \to F_0(t).$$

Como $\epsilon > 0$, o ponto $t = \epsilon$ é ponto de continuidade de F_0 e $F_0(\epsilon) = 1$. Logo,

$$\mathbb{P}\left(Y_n \leqslant \epsilon\right) = F_{Y_n}(\epsilon) \xrightarrow[n\to\infty]{} 1.$$

Consequentemente,

$$\mathbb{P}\left(\|\boldsymbol{X}_n - \boldsymbol{c}\| > \epsilon\right) = 1 - \mathbb{P}\left(Y_n \leqslant \epsilon\right) \xrightarrow[n\to\infty]{} 0.$$

Como $\epsilon > 0$ é arbitrário, concluímos que $\boldsymbol{X}_n \xrightarrow{p} \boldsymbol{c}$.