# Probabilidade I

Prova Final

June 28, 2023

Observações:

- A prova tera duração de 4 horas.
- Questões deverão ser respondidas de maneira matematicamente rigorosa e clara.
- Todos os resultados provados em aula (exceto exercícios) poderão ser utilizados, salvo quando a questão pedir que algum destes seja provado. Nesse caso, todos os resultados anteriores poderão ser utilizados.
- Caso uma questão seja composta de vários itens, cada um poderá ser resolvido independentemente, supondo a validade dos anteriores.

## Exercício I

Consideramos $Y_n$ uma sequência de variáveis Gaussians com distribuição $\mathcal{N}(m_n, \sigma_n^2)$ aonde $(m_n)_{n\geqslant 0}$ e $(\sigma_n)_{n\geqslant 0}$ são sequências de numeros reais (reais positivos no caso de $\sigma_n$).

(i) *(1 pt)* Provar que $Y_n$ converge em distribuição se e somente se as sequências $m_n$ e $\sigma_n$ convergem em $\mathbb{R}$.

(ii) *(1 pt)* Assumindo que $m = \lim_{n\to\infty} m_n$ e $\sigma = \lim_{n\to\infty} \sigma_n$ $(m, \sigma \in \mathbb{R})$, identificar a distribuição limite de $Y_n$.

*Usando o Teorema de Lévy, se $Y_n$ converge em distribuição para $Y$ então temos convergência das funções características de $Y_n$ para a de $Y$. Como*

$$\Phi_{Y_n}(\xi) = E\left[e^{i\xi Y_n}\right] = e^{i\xi m_n - \frac{\xi^2\sigma_n^2}{2}}.$$

*Daí, temos convergência pontual de $\Phi_{Y_n}(\xi)$ para uma função contínua somente se $\sigma_n^2$ e $m_n$ convergem para limites finitos (se $\sigma_n \to \infty$ temos convergência para $\mathbf{1}_0$ que não é uma função característica). Isso prova que convergência em distribuição implica convergência de $m_n$ e $\sigma_n$. Para outra implicação, chamando $m$ e $\sigma$ os limites respetivos de $m_n$ e $\sigma$ temos*

$$\lim_{n\to\infty} \Phi_{Y_n}(\xi) = e^{i\xi m - \frac{\xi^2\sigma^2}{2}} = \Phi_Y(\xi)$$

*aonde $Y \sim \mathcal{N}(m, \sigma^2)$. Usando Teorema de Lévy de novo isso implica a convergência em distribuição e permite identificar o limite.*

(iii) *(1 pt)* Usando respostas anteriores, mostrar que se uma sequência de variáveis Gaussiana $Y_n$ converge para $Y$ em probabilidade então $Y_n$ converge para $Y$ em $L^p$ para todo $p \in [1, \infty)$.

*Para provar que $Y_n$ converge em $L_p$ vamos provar que $\sup_{n \geqslant 1} E[|Y_n|^q] < \infty$ para $q > p$ (usando um resultado visto em aula sobre integrabilidade uniforme). Como $Y_n$ converge para $Y$ em probabilidade, também converge em distribuição. Por (i) isso implica convergência de $\sigma_n$ e $m_n$. Como temos para $k \geqslant 1$*

$$E[|Y_n|^2 k] \leqslant (2k)! E[\cosh(Y_n)] = \cosh\left(m_n + \frac{\sigma_n^2}{2}\right)$$

*isso implica que a sequência $E[|Y_n|^2 k]$ é limitada e permite concluir.*

(iv) *(1 pt)* Nas hipotèses do item anterior, podemos affirmar que $Y_n$ converge quase certamente para $Y$? Prove ou dê um contra-exemplo.

*Consideramos $\sigma_n = \frac{1}{(\log(n+2)^{1/4}}$, e $Y_n$ uma sequência de variáveis Gaussianas independentes de media 0 e variança $\sigma_n^2$. A sequência $Y_n$ converge em distribuição eis probabilidade para 0 (do item (i)) mas temos*

$$P[Y_n \geqslant 1] \geqslant P[Y_n \in [1, 1+\sigma_n]] \geqslant \frac{1}{\sqrt{2\pi}\sigma_n} \int_1^{1+\sigma_n} e^{-\frac{u}{2\sigma_n^2}} \, \mathrm{d}u \geqslant \frac{1}{\sqrt{2\pi}} e^{e^{-\frac{(1+\sigma_n)^2}{2\sigma_n^2}}} \geqslant e^{-\sqrt{\log n}},$$

*aonde a última desigualdade vale para $n$ grande sufficiente. Usando o Lemma de Borel-Cantelli (no caso de eventos independentes) obtemos que $Y_n \geqslant 1$ infinitas vezes com probabilidade 1 e então que $Y_n$ não converge quase certamente para 0.*

## Exercício II

Consideramos $(X_n)_{n \geqslant 0}$ uma sequência IID de variáveis exponenciais de esperança 1. Dado $a > 0$, definimos $T_a := \inf\{ n \ : \ X_n \geqslant a\}$.

(i) *(1 pt)* Calcular para todo $k \in \mathbb{N}$, $P[T_a \geqslant k]$, deduzir que $T_a < \infty$ quase certamente.

*O evento $\{T_a \geqslant k\}$ coincide com $\{\max(X_1, \dots, X_{k-1}) < a\}$. Por independência temos*

$$P[X_1 < a; \dots ; X_{k-1} < a] = P[X_1 < a]^{k-1} = (1 - e^{-a})^{k-1}.$$

*Temos (usando continuidade por cima de $P$) $P[T_a = \infty] = P[\bigcup_{k \geqslant 1} \{T_a \geqslant k\}] = \lim_{k \to \infty} (1 - e^{-a})^{k-1} = 0$.*

(ii) *(2 pt)* Definimos $Y_a = X_{T_a} - a$. Calcular a probabilidade conditional $P[Y_a \in A \mid T_a = k]$ para $k \in \mathbb{N}$ e $A \in \mathcal{B}(\mathbb{R})$. Deduzir que $Y_a$ e $T_a$ são independentes.

*Temos*

$$\begin{aligned} &P[\{Y_a \in A\} \cap \{T_a = k\}] = P[X_k \in (a + A); X_1 < a; \dots ; X_{k-1} < a] \\ &= P[X_k \in (a+A)] P[X_1 < a]^{k-1} = \left(\int_{A+a} e^{-t} \, \mathrm{d}t\right) (1-e^{-a})^{k-1} = \left(\int_A e^{-u} \, \mathrm{d}u\right) e^{-a} (1-e^{-a})^{k-1} \end{aligned} \quad (1)$$

*Por outro lado do item (i) temos*

$$P[T_a = k] = P[T_a \geqslant k] - P[T_a \geqslant k+1] = e^{-a}(1-e^{-a})^{k-1},$$

*eis*

$$P[Y_a \in A \mid T_a = k] = \int_A e^{-u}\,\mathrm{d}u = P[X_1 \in A].$$

*Daí concluemos que* $P[Y_a \in A \cap T_a \in B] = P[X_1 \in A]P[T_a \in B]$ *para qualquer* $B \subset \mathbb{N}$ *e que* $Y_a$ *e* $T_a$ *são independentes.*

## Exercício III

Sejam $Y_1^{(N)}, Y_2^{(N)}, \dots Y_N^{(N)}$, $N$ variáveis aleatórias IID uniforme em $[0, N+1]$.

(i) *(1 pt)* Definimos $X_N := \min_{i\in\{1,\dots,N\}} Y_i^{(N)}$. Calcular $P[X_N \geqslant t]$ para $t \geqslant 0$.

*Temos por independência*

$$P[X_N \geqslant t] = P[Y_1^{(N)} \geqslant t, \dots, Y_N^{(N)} \geqslant t] = P[Y_1^{(N)} \geqslant t]^N = (1 - \frac{t}{N+1})^N$$

(ii) *(1 pt)* Mostrar que $X_N$ converge em distribuição e identificar o limite.

*Temos que*

$$\lim_{N\to\infty} P[X_N \geqslant t] = \lim_{N\to\infty} (1 - \frac{t}{N+1})^N = e^{-t}.$$

*Isso implica que* $\lim_{N\to\infty} P[X_N \leqslant t] = 1 - e^{-t}$. *A função acumulada de distribuição de* $X_N$ *converge para a de uma variável exponencial (em todo ponto eis em todo ponto de continuidade). Eis,* $X_N$ *converge em distibuição para uma variável exponencial de parametro* 1.

(iii) *(★)* Consideramos $X_N^{(1)} \leqslant X_N^{(2)} \leqslant \dots \leqslant X_N^{(N)}$ os valores ordenados de $Y_1^{(N)}, Y_2^{(N)}, \dots Y_N^{(N)}$. Mostrar que para $k \in \mathbb{N}$ fixo, a sequência $(X_N^{(k)})_{N \geqslant k}$ converge em distribuição e identificar a distribuição limite.

*Para* $k \geqslant 2$ *e observando que os* $Y_i^{(N)}$ *assumem valores distintos com proababilidade* 1, *temos*

$$\begin{aligned} P[X_N^{(k)} \geqslant t] &= \sum_{I \subset [\![0,N]\!]\ :\ |I|=k-1} P\left[\forall i \in I,\ Y_i^{(N)} < t, \forall j \in I^{\complement},\ Y_j^{(N)} \geqslant t\right] \\ &= \binom{N}{k-1} \left(\frac{t}{N+1}\right)^{k-1} \left(1 - \frac{t}{N+1}\right)^{N-k}. \end{aligned} \quad (2)$$

*Considerando o limite quando* $N \to \infty$ *obtemos*

$$\lim_{N\to\infty} P[X_N^{(k)} \geqslant t] \leqslant \frac{t^{k-1}}{(k-1)!} e^{-t}.$$

## Exercício IV

Sejam $X_1$ e $X_2$ duas variáveis exponencias independentes de parametro 1. Definimos $\xi := X_1$ e $\zeta = X_1 + X_2$.

(i) *(1,5 pt)* Provar que a lei de distribuição do vetor $(\xi, \zeta)$ tem uma densidade em $\mathbb{R}^2$ calculando essa densidade.

*Seja $\varphi$ uma função mensurável limitada. Temos usando a independência de $X_1$ e $X_2$ e as densidade respetivas.*

$$E[\varphi(\xi,\zeta)] = \int_{\mathbb{R}_+^2} \varphi(u, u+v) e^{-u} e^{-v} \, \mathrm{d}u \, \mathrm{d}v$$

*Fazendo a troca de variável $(u, v) \to (u, t)$ com $t = u + v$ temos*

$$E[\varphi(\xi,\zeta)] = \int_0^\infty \int_u^\infty \varphi(u,t) e^{-t} \, \mathrm{d}u \, \mathrm{d}t = \int_{\mathbb{R}_+^2} \varphi(u,t) \mathbf{1}_{\{t \geqslant u\}} e^{-t} \, \mathrm{d}u \, \mathrm{d}t.$$

*Daí (considerando funções indicadoras) podemos concluir que a distribuição conjunta de $(\xi, \zeta)$ tem densidade dada por $\mathbf{1}_{\{t \geqslant u\}} e^{-t}$.*

(ii) *(1 pt)* Identificar (via um cálculo) um núcleo de transição $K_1$ tal que para todo $A \in \mathcal{B}(\mathbb{R})$

$$E[\mathbf{1}_{\xi \in A} \mid \zeta] = K_1(\zeta, A)$$

quase certamente.

*Usando os resultados vistos em aula, um tal núcleo de transição pode ser obtido usando a fórmula*

$$K_1(t, A) = \frac{\int_A p(u,t) \, \mathrm{d}u}{\int_{\mathbb{R}_+} p(u,t) \, \mathrm{d}u}$$

*aonde $p$ é a densidade de distribuição de $(\xi, \zeta)$. Usando a resposta da pergunta anterior obtemos*

$$K_1(t, A) = \frac{1}{t} \lambda(A \cap [0, t])$$

*aonde $\lambda$ é a medida de Lebesgue. (Com o condicionamento $\xi$ é uma variável uniforme no intervalo $[0, \zeta]$)*

(iii) *(1 pt)* Identificar (via um cálculo) um núcleo de transição $K_2$ tal que tal que para todo $B \in \mathcal{B}(\mathbb{R})$

$$E[\mathbf{1}_{\zeta \in B} \mid \xi] = K_2(\xi, B)$$

quase certamente.

*Do mesmo jeito temos*

$$K_1(t, A) = \frac{\int_B p(u,t) \, \mathrm{d}t}{\int_{\mathbb{R}_+} p(u,t) \, \mathrm{d}t} = \int_B \mathbf{1}_{t \geqslant u} e^{u-t} \, \mathrm{d}t = \int_{\mathbb{R}_+} \mathbf{1}_{s \in (B-u) \cap \mathbb{R}_+} e^{-s} \, \mathrm{d}s.$$

*(Com o condicionamento, $\zeta$ é simplesmente $\xi$ mais uma variável exponencial )*

## Exercício V

Sejam $(X_k)_{k\geqslant 1}$ variáveis IID tal que $P(X_k = 1) = P(X_k = -1) = 1/2$. Definimos

$$Z_N := \frac{1}{\sqrt{\log N}} \sum_{k=1}^{N} \frac{1}{\sqrt{k}} X_k.$$

(i) *(1 pt)* Calcular a variança de $Z_N$ e o valor de $\lim_{N\to\infty} \mathbb{E}[Z_N^2]$

*Usando independência, a variança da soma vale a soma das varianças e temos* $\mathrm{Var}[Z_N] = \mathbb{E}[Z_N^2] = \frac{1}{\log N} \sum_{k=1}^{N} \frac{1}{k}$*. Em particular temos* $\lim_{N\to\infty} \mathbb{E}[Z_N^2] = 1$*.*

(ii) *(2 pt)* Usando o Teorema de Lévy, mostrar que $Z_N$ converge em distribuição e identificar a distribuição limite.

*Usando independência,*

$$\Phi_{Z_N}(\xi) = \prod_{k=1}^{N} \Phi_{1/(\sqrt{k \log N})}(\xi) = \prod_{k=1}^{N} \cos\left(\frac{\xi}{\sqrt{k \log N}}\right).$$

*Agora temos para $\xi$ fixo e $N \geqslant N_0(\xi)$, temos* $\left|\frac{\xi}{\sqrt{k\log N}}\right| \leqslant \pi/3$ *eis* $\cos\left(\frac{\xi}{\sqrt{k\log N}}\right) \geqslant 1/2$*. Ademais, por Taylor temos para todo $k$ e $N \geqslant N_0(\xi)$*

$$\left|\log\cos\left(\frac{\xi}{\sqrt{k\log N}}\right) - \frac{\xi^2}{2k\log N}\right| \leqslant \frac{C\xi^4}{2k^2(\log N)^2}.$$

*Dai obtemos que*

$$\left|\log \Phi_{Z_N}(\xi) - \sum_{k=1}^{N} \frac{\xi^2}{2k\log N}\right| \leqslant C \sum_{k=1}^{N} \frac{C\xi^4}{2k^2(\log N)^2}.$$

*O membro de direita e menor que $C'\xi^4(\log N)^{-2}$ (pois $k^{-2}$ é sumável). Daí podemos concluir que*

$$\lim_{N\to\infty} \log \Phi_{Z_N}(\xi) = \lim_{N\to\infty} \sum_{k=1}^{N} \frac{\xi^2}{2k\log N} = \frac{\xi^2}{2}$$

*de que concluemos que $Z_N$ converge em distribuição para uma Gaussiana padrão.*