Probabilidade e Estatística
Monitoria 2 - Resolução da Lista 2

Guilherme Vianna

3 de fevereiro de 2026

Questão 1(a) — Enunciado

Enunciado

Seja $(\Omega, \Sigma, \mathbb{P})$ um espaço de probabilidade e $(Z_n)_{n\in\mathbb{N}}$ uma sequência de variáveis aleatórias não negativas ($Z_n \geqslant 0$). Mostre que

$$\mathbb{E}\left[\sum_{n=1}^{\infty} Z_n\right] = \sum_{n=1}^{\infty} \mathbb{E}[Z_n].$$

Questão 1(a) — Resolução

Defina as somas parciais

$$S_N := \sum_{n=1}^{N} Z_n \qquad \text{e} \qquad S := \sum_{n=1}^{\infty} Z_n \quad \text{(valor em } [0, \infty]\text{)}.$$

Como $Z_n \geqslant 0$, temos $S_N(\omega) \uparrow S(\omega)$ ponto a ponto em Ω.

Questão 1(a) — Resolução

Defina as somas parciais

$$S_N := \sum_{n=1}^{N} Z_n \qquad \text{e} \qquad S := \sum_{n=1}^{\infty} Z_n \quad \text{(valor em } [0, \infty]\text{)}.$$

Como $Z_n \geqslant 0$, temos $S_N(\omega) \uparrow S(\omega)$ ponto a ponto em Ω.
Pelo **Teorema da Convergência Monótona**:

$$\mathbb{E}[S] = \mathbb{E}\left[\lim_{N\to\infty} S_N\right] = \lim_{N\to\infty} \mathbb{E}[S_N].$$

Questão 1(a) — Resolução

Defina as somas parciais

$$S_N := \sum_{n=1}^{N} Z_n \qquad \text{e} \qquad S := \sum_{n=1}^{\infty} Z_n \quad (\text{valor em } [0,\infty]).$$

Como $Z_n \geqslant 0$, temos $S_N(\omega) \uparrow S(\omega)$ ponto a ponto em Ω.
Pelo **Teorema da Convergência Monótona**:

$$\mathbb{E}[S] \; = \mathbb{E}\left[\lim_{N\to\infty} S_N\right] = \lim_{N\to\infty} \mathbb{E}[S_N].$$

Resta identificar $\mathbb{E}[S_N]$ em termos de $(\mathbb{E}[Z_n])_{n\geqslant 1}$.

Questão 1(a) — Resolução: aditividade

Relembrando:

Se $X, Y \geqslant 0$ são mensuráveis, então

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y] \qquad \text{(em } [0,\infty]\text{)}.$$

Questão 1(a) — Resolução: aditividade

Relembrando:

Se $X, Y \geqslant 0$ são mensuráveis, então

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y] \qquad \text{(em } [0,\infty]).$$

Prova (caso simples). Se $X = \sum_{i=1}^{k} a_i \, \mathbf{1}_{A_i}$ e $Y = \sum_{j=1}^{m} b_j \, \mathbf{1}_{B_j}$ são simples não negativas, refine para a partição disjunta $(A_i \cap B_j)_{i,j}$ e use a definição de integral de Lebesgue para funções simples:

$$\mathbb{E}[X+Y] = \sum_{i,j} (a_i + b_j) \, \mathbb{P}(A_i \cap B_j) = \sum_i a_i \, \mathbb{P}(A_i) + \sum_j b_j \, \mathbb{P}(B_j) = \mathbb{E}[X] + \mathbb{E}[Y].$$

Questão 1(a) — Resolução: aditividade (continuação)

Prova (caso geral). Sejam $(X_k)_{k\geqslant 1}$ e $(Y_k)_{k\geqslant 1}$ sequências simples[1] tais que $X_k \uparrow X$ e $Y_k \uparrow Y$. Então $X_k + Y_k \uparrow X + Y$ e, pelo Teorema da Convergência Monótona,

$$\mathbb{E}[X+Y] = \lim_{k\to\infty} \mathbb{E}[X_k + Y_k] = \lim_{k\to\infty} \big(\mathbb{E}[X_k] + \mathbb{E}[Y_k]\big) = \mathbb{E}[X] + \mathbb{E}[Y].$$

[1] Construídas com as funções escada s_k dos slides da aula, i.e $X_k = s_k \circ X$ por exemplo

Questão 1(a) — Resolução: aditividade (continuação)

Prova (caso geral). Sejam $(X_k)_{k\geqslant 1}$ e $(Y_k)_{k\geqslant 1}$ sequências simples[1] tais que $X_k \uparrow X$ e $Y_k \uparrow Y$. Então $X_k + Y_k \uparrow X + Y$ e, pelo Teorema da Convergência Monótona,

$$\mathbb{E}[X+Y] = \lim_{k\to\infty} \mathbb{E}[X_k + Y_k] = \lim_{k\to\infty} \big(\mathbb{E}[X_k] + \mathbb{E}[Y_k]\big) = \mathbb{E}[X] + \mathbb{E}[Y].$$

Por indução, a aditividade vale para qualquer soma finita de v.as não-negativas:

$$\mathbb{E}\left[\sum_{n=1}^{N} Z_n\right] = \sum_{n=1}^{N} \mathbb{E}[Z_n], \qquad \forall\, N \in \mathbb{N}.$$

[1] Construídas com as funções escada s_k dos slides da aula, i.e $X_k = s_k \circ X$ por exemplo

Questão 1(a) — Resolução: conclusão

Lembrando que $S_N = \sum_{n=1}^{N} Z_n$ e $S = \sum_{n=1}^{\infty} Z_n$,

$$\mathbb{E}\left[\sum_{n=1}^{\infty} Z_n\right] = \mathbb{E}[S] = \mathbb{E}\left[\lim_{N\to\infty} S_N\right] = \lim_{N\to\infty} \mathbb{E}[S_N] =$$

$$= \lim_{N\to\infty} \mathbb{E}\left[\sum_{n=1}^{N} Z_n\right] = \lim_{N\to\infty} \sum_{n=1}^{N} \mathbb{E}[Z_n] = \sum_{n=1}^{\infty} \mathbb{E}[Z_n].$$

Questão 1(a) — Resolução: conclusão

Lembrando que $S_N = \sum_{n=1}^{N} Z_n$ e $S = \sum_{n=1}^{\infty} Z_n$,

$$\mathbb{E}\left[\sum_{n=1}^{\infty} Z_n\right] = \mathbb{E}[S] = \mathbb{E}\left[\lim_{N\to\infty} S_N\right] = \lim_{N\to\infty} \mathbb{E}[S_N] =$$

$$= \lim_{N\to\infty} \mathbb{E}\left[\sum_{n=1}^{N} Z_n\right] = \lim_{N\to\infty} \sum_{n=1}^{N} \mathbb{E}[Z_n] = \sum_{n=1}^{\infty} \mathbb{E}[Z_n].$$

Ou seja,

$$\mathbb{E}\left[\sum_{n=1}^{\infty} Z_n\right] = \sum_{n=1}^{\infty} \mathbb{E}[Z_n].$$

Questão 1(a) — Resolução: conclusão

Lembrando que $S_N = \sum_{n=1}^{N} Z_n$ e $S = \sum_{n=1}^{\infty} Z_n$,

$$\mathbb{E}\left[\sum_{n=1}^{\infty} Z_n\right] = \mathbb{E}[S] = \mathbb{E}\left[\lim_{N\to\infty} S_N\right] = \lim_{N\to\infty} \mathbb{E}[S_N] =$$

$$= \lim_{N\to\infty} \mathbb{E}\left[\sum_{n=1}^{N} Z_n\right] = \lim_{N\to\infty} \sum_{n=1}^{N} \mathbb{E}[Z_n] = \sum_{n=1}^{\infty} \mathbb{E}[Z_n].$$

Ou seja,

$$\mathbb{E}\left[\sum_{n=1}^{\infty} Z_n\right] = \sum_{n=1}^{\infty} \mathbb{E}[Z_n].$$

Observação

A igualdade é válida sem supor integrabilidade individual dos Z_n; ambos os lados são entendidos como valores em $[0, \infty]$.

Questão 1(b) — Enunciado

Enunciado

Seja $(\Omega, \Sigma, \mathbb{P})$ um espaço de probabilidade e $(Z_n)_{n\in\mathbb{N}}$ uma sequência de variáveis aleatórias não negativas ($Z_n \geqslant 0$). Mostre que

$$\mathbb{E}\left[\sum_{n=1}^{\infty} Z_n\right] < \infty \implies \mathbb{P}\left(\left\{\omega : \sum_{n=1}^{\infty} Z_n(\omega) < \infty\right\}\right) = 1.$$

Questão 1(b) — Resolução

Defina as somas parciais

$$S_N := \sum_{n=1}^{N} Z_n \qquad \text{e} \qquad S := \sum_{n=1}^{\infty} Z_n = \lim_{N\to\infty} S_N.$$

Como $Z_n \geqslant 0$, temos $S_N \uparrow S$ ponto a ponto. Em particular, cada S_N é mensurável, logo S é mensurável e $S \geqslant 0$.

Questão 1(b) — Resolução

Defina as somas parciais

$$S_N := \sum_{n=1}^{N} Z_n \qquad \text{e} \qquad S := \sum_{n=1}^{\infty} Z_n = \lim_{N\to\infty} S_N.$$

Como $Z_n \geqslant 0$, temos $S_N \uparrow S$ ponto a ponto. Em particular, cada S_N é mensurável, logo S é mensurável e $S \geqslant 0$.
Assuma $\mathbb{E}[S] < \infty$.
Pelo lema padrão de integração de Lebesgue para funções não negativas:

$$\boxed{\text{se } f \geqslant 0 \text{ é mensurável e } \int f\,d\mu < \infty, \text{ então } \mu(\{f = \infty\}) = 0},$$

tomando $\mu = \mathbb{P}$ e $f = S$, obtemos

$$\mathbb{P}(S = \infty) = 0.$$

Questão 1(b) — Resolução

Defina as somas parciais

$$S_N := \sum_{n=1}^{N} Z_n \qquad \text{e} \qquad S := \sum_{n=1}^{\infty} Z_n = \lim_{N\to\infty} S_N.$$

Como $Z_n \geqslant 0$, temos $S_N \uparrow S$ ponto a ponto. Em particular, cada S_N é mensurável, logo S é mensurável e $S \geqslant 0$.

Assuma $\mathbb{E}[S] < \infty$.

Pelo lema padrão de integração de Lebesgue para funções não negativas:

$$\boxed{\text{se } f \geqslant 0 \text{ é mensurável e } \int f\, d\mu < \infty, \text{ então } \mu(\{f = \infty\}) = 0},$$

tomando $\mu = \mathbb{P}$ e $f = S$, obtemos

$$\mathbb{P}(S = \infty) = 0.$$

Logo,

$$\mathbb{P}\left(\left\{\omega : \sum_{n=1}^{\infty} Z_n(\omega) < \infty\right\}\right) = \mathbb{P}(S < \infty) = 1,$$

isto é, $\sum_{n=1}^{\infty} Z_n$ é finita $\mathbb{P}$-quase certamente.

Questão 1(c) — Enunciado

Enunciado

Mostre que

$$\mathbb{P}\left(\left\{\omega : \sum_{n=1}^{\infty} Z_n(\omega) < \infty\right\}\right) = 1 \implies \mathbb{P}(\{\omega : Z_n(\omega) \nrightarrow 0\}) = 0.$$

Questão 1(c) — Resolução

Seja

$$E := \left\{ \omega : \sum_{n=1}^{\infty} Z_n(\omega) < \infty \right\}, \qquad \mathbb{P}(E) = 1.$$

Questão 1(c) — Resolução

Seja

$$E := \left\{ \omega : \sum_{n=1}^{\infty} Z_n(\omega) < \infty \right\}, \qquad \mathbb{P}(E) = 1.$$

Para $\omega \in E$, defina as somas parciais

$$S_N(\omega) := \sum_{n=1}^{N} Z_n(\omega), \qquad S_0(\omega) := 0,$$

e a soma total (finita em E)

$$S(\omega) := \sum_{n=1}^{\infty} Z_n(\omega) < \infty.$$

Como $Z_n(\omega) \geqslant 0$, temos $S_N(\omega) \uparrow S(\omega)$ para todo $\omega \in E$.

Questão 1(c) — Resolução

Seja

$$E := \left\{ \omega : \sum_{n=1}^{\infty} Z_n(\omega) < \infty \right\}, \qquad \mathbb{P}(E) = 1.$$

Para $\omega \in E$, defina as somas parciais

$$S_N(\omega) := \sum_{n=1}^{N} Z_n(\omega), \qquad S_0(\omega) := 0,$$

e a soma total (finita em E)

$$S(\omega) := \sum_{n=1}^{\infty} Z_n(\omega) < \infty.$$

Como $Z_n(\omega) \geqslant 0$, temos $S_N(\omega) \uparrow S(\omega)$ para todo $\omega \in E$.

Questão 1(c) — Resolução

Considere as caudas

$$T_n(\omega) := \sum_{k=n}^{\infty} Z_k(\omega) = S(\omega) - S_{n-1}(\omega), \qquad n \geqslant 1.$$

Então, para $\omega \in E$, como $S_{n-1}(\omega) \uparrow S(\omega)$, segue que

$$T_n(\omega) \downarrow 0 \quad \text{quando } n \to \infty.$$

Questão 1(c) — Resolução (conclusão)

Para todo $n \geqslant 1$ e $\omega \in E$, vale

$$0 \leqslant Z_n(\omega) \leqslant T_n(\omega),$$

pois $T_n(\omega)$ é a soma de $Z_n(\omega)$ com termos não negativos.

Questão 1(c) — Resolução (conclusão)

Para todo $n \geqslant 1$ e $\omega \in E$, vale

$$0 \leqslant Z_n(\omega) \leqslant T_n(\omega),$$

pois $T_n(\omega)$ é a soma de $Z_n(\omega)$ com termos não negativos.
Como $T_n(\omega) \to 0$ em E, obtemos

$$Z_n(\omega) \to 0 \quad \text{para todo } \omega \in E.$$

Logo,

$$E \subseteq \{\omega : Z_n(\omega) \to 0\} \implies \mathbb{P}(E) \leqslant \mathbb{P}(\{\omega : Z_n(\omega) \to 0\}).$$

Questão 1(c) — Resolução (conclusão)

Para todo $n \geqslant 1$ e $\omega \in E$, vale

$$0 \leqslant Z_n(\omega) \leqslant T_n(\omega),$$

pois $T_n(\omega)$ é a soma de $Z_n(\omega)$ com termos não negativos.
Como $T_n(\omega) \to 0$ em E, obtemos

$$Z_n(\omega) \to 0 \quad \text{para todo } \omega \in E.$$

Logo,

$$E \subseteq \{\omega : Z_n(\omega) \to 0\} \implies \mathbb{P}(E) \leqslant \mathbb{P}(\{\omega : Z_n(\omega) \to 0\}).$$

Portanto, (lembrando que $\mathbb{P}(E) = 1$),

$$1 - \mathbb{P}(E^c) \leqslant 1 - \mathbb{P}(\{\omega : Z_n(\omega) \nrightarrow 0\}) \implies \mathbb{P}(\{\omega : Z_n(\omega) \nrightarrow 0\}) \leqslant 0,$$

isto é,

$$\mathbb{P}\{\omega : Z_n(\omega) \nrightarrow 0\} = 0.$$

Questão 1(c) — Resolução (conclusão)

Para todo $n \geqslant 1$ e $\omega \in E$, vale

$$0 \leqslant Z_n(\omega) \leqslant T_n(\omega),$$

pois $T_n(\omega)$ é a soma de $Z_n(\omega)$ com termos não negativos.
Como $T_n(\omega) \to 0$ em E, obtemos

$$Z_n(\omega) \to 0 \quad \text{para todo } \omega \in E.$$

Logo,

$$E \subseteq \{\omega : Z_n(\omega) \to 0\} \implies \mathbb{P}(E) \leqslant \mathbb{P}(\{\omega : Z_n(\omega) \to 0\}).$$

Portanto, (lembrando que $\mathbb{P}(E) = 1$),

$$1 - \mathbb{P}(E^c) \leqslant 1 - \mathbb{P}(\{\omega : Z_n(\omega) \nrightarrow 0\}) \implies \mathbb{P}(\{\omega : Z_n(\omega) \nrightarrow 0\}) \leqslant 0,$$

isto é,

$$\mathbb{P}\{\omega : Z_n(\omega) \nrightarrow 0\} = 0.$$

Comentário: o item (c) é exatamente o análogo probabilístico do "teste do termo geral" para séries: para cada ω, $\sum_{n \geqslant 1} Z_n(\omega)$ é uma série de números reais, se converge, então $Z_n(\omega) \to 0$.

Exercício 2 — Enunciado

Enunciado

Seja $(\Omega, \Sigma, \mathbb{P})$ um espaço de probabilidade, e X uma variável aleatória. Mostre que, para toda função mensurável $f : \mathbb{R} \to \mathbb{R}$ tal que $\mathbb{E}[\,|f(X)|\,] < \infty$,

$$\mathbb{E}[f(X)] = \int_{\mathbb{R}} f(x)\, \mathbb{P}_X(dx),$$

onde $\mathbb{P}_X$ é a medida induzida por X em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, isto é, $\mathbb{P}_X(B) = \mathbb{P}(X \in B)$ para todo $B \in \mathcal{B}(\mathbb{R})$.

Dica: prove para f simples; depois aproxime $f \geqslant 0$ por funções simples via TCM; por fim, trate o caso geral via partes positiva/negativa.

Exercício 2 — Solução (função simples)

Caso 1: f simples. Suponha que

$$f = \sum_{i=1}^{k} a_i \, \mathbf{1}_{A_i}, \qquad A_i \in \mathcal{B}(\mathbb{R}) \text{ disjuntos.}$$

Então

$$f(X) = \sum_{i=1}^{k} a_i \, \mathbf{1}_{A_i}(X) = \sum_{i=1}^{k} a_i \, \mathbf{1}_{\{X \in A_i\}}.$$

Exercício 2 — Solução (função simples)

Caso 1: f simples. Suponha que

$$f = \sum_{i=1}^{k} a_i \, \mathbf{1}_{A_i}, \qquad A_i \in \mathcal{B}(\mathbb{R}) \text{ disjuntos.}$$

Então

$$f(X) = \sum_{i=1}^{k} a_i \, \mathbf{1}_{A_i}(X) = \sum_{i=1}^{k} a_i \, \mathbf{1}_{\{X \in A_i\}}.$$

Logo,

$$\mathbb{E}[f(X)] = \sum_{i=1}^{k} a_i \, \mathbb{E}[\mathbf{1}_{\{X \in A_i\}}] = \sum_{i=1}^{k} a_i \, \mathbb{P}(X \in A_i) = \sum_{i=1}^{k} a_i \, \mathbb{P}_X(A_i).$$

Exercício 2 — Solução (função simples)

Caso 1: f simples. Suponha que

$$f = \sum_{i=1}^{k} a_i \, \mathbf{1}_{A_i}, \qquad A_i \in \mathcal{B}(\mathbb{R}) \text{ disjuntos.}$$

Então

$$f(X) = \sum_{i=1}^{k} a_i \, \mathbf{1}_{A_i}(X) = \sum_{i=1}^{k} a_i \, \mathbf{1}_{\{X \in A_i\}}.$$

Logo,

$$\mathbb{E}[f(X)] = \sum_{i=1}^{k} a_i \, \mathbb{E}[\mathbf{1}_{\{X \in A_i\}}] = \sum_{i=1}^{k} a_i \, \mathbb{P}(X \in A_i) = \sum_{i=1}^{k} a_i \, \mathbb{P}_X(A_i).$$

Pela definição da integral de Lebesgue para funções simples,

$$\sum_{i=1}^{k} a_i \, \mathbb{P}_X(A_i) = \int_{\mathbb{R}} f(x) \, \mathbb{P}_X(\mathrm{d}x),$$

concluindo o caso simples.

Exercício 2 — Solução ($f \geqslant 0$)

Caso 2: $f \geqslant 0$. Tome $(f_n)_{n \geqslant 1}$ funções simples não negativas tais que

$$f_n(x) \uparrow f(x) \quad \text{para todo } x \in \mathbb{R}.^2$$

Então $f_n(X) \uparrow f(X)$ em Ω.

[2] por exemplo, aproximações pelas funções-escada $f_k = s_k \circ f$

Exercício 2 — Solução ($f \geqslant 0$)

Caso 2: $f \geqslant 0$. Tome $(f_n)_{n \geqslant 1}$ funções simples não negativas tais que

$$f_n(x) \uparrow f(x) \quad \text{para todo } x \in \mathbb{R}.^2$$

Então $f_n(X) \uparrow f(X)$ em Ω. Pelo Teorema da Convergência Monótona (em $(\Omega, \Sigma, \mathbb{P})$),

$$\mathbb{E}[f(X)] = \lim_{n \to \infty} \mathbb{E}[f_n(X)].$$

[2] por exemplo, aproximações pelas funções-escada $f_k = s_k \circ f$

Exercício 2 — Solução ($f \geqslant 0$)

Caso 2: $f \geqslant 0$. Tome $(f_n)_{n \geqslant 1}$ funções simples não negativas tais que

$$f_n(x) \uparrow f(x) \quad \text{para todo } x \in \mathbb{R}.^2$$

Então $f_n(X) \uparrow f(X)$ em Ω. Pelo Teorema da Convergência Monótona (em $(\Omega, \Sigma, \mathbb{P})$),

$$\mathbb{E}[f(X)] = \lim_{n \to \infty} \mathbb{E}[f_n(X)].$$

Pelo caso simples, para cada n,

$$\mathbb{E}[f_n(X)] = \int_{\mathbb{R}} f_n(x)\, \mathbb{P}_X(\mathrm{d}x).$$

[2] por exemplo, aproximações pelas funções-escada $f_k = s_k \circ f$

Exercício 2 — Solução ($f \geqslant 0$)

Caso 2: $f \geqslant 0$. Tome $(f_n)_{n\geqslant 1}$ funções simples não negativas tais que

$$f_n(x) \uparrow f(x) \quad \text{para todo } x \in \mathbb{R}.$$[2]

Então $f_n(X) \uparrow f(X)$ em Ω. Pelo Teorema da Convergência Monótona (em $(\Omega, \Sigma, \mathbb{P})$),

$$\mathbb{E}[f(X)] = \lim_{n\to\infty} \mathbb{E}[f_n(X)].$$

Pelo caso simples, para cada n,

$$\mathbb{E}[f_n(X)] = \int_{\mathbb{R}} f_n(x)\, \mathbb{P}_X(\mathrm{d}x).$$

Aplicando novamente o TCM agora no espaço $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_X)$,

$$\mathbb{E}[f(X)] = \lim_{n\to\infty} \mathbb{E}[f_n(X)] = \lim_{n\to\infty} \int f_n(x)\, \mathbb{P}_X(\mathrm{d}x) = \int f(x)\, \mathbb{P}_X(\mathrm{d}x).$$

[2] por exemplo, aproximações pelas funções-escada $f_k = s_k \circ f$

Exercício 2 — Solução (caso geral)

Caso 3: f geral com $\mathbb{E}[|f(X)|] < \infty$. Escreva

$$f = f^+ - f^-, \qquad f^+ = \max\{f, 0\}, \quad f^- = \max\{-f, 0\},$$

com $f^+, f^- \geqslant 0$ e $|f| = f^+ + f^-$.

Exercício 2 — Solução (caso geral)

Caso 3: f geral com $\mathbb{E}[|f(X)|] < \infty$. Escreva

$$f = f^+ - f^-, \qquad f^+ = \max\{f, 0\}, \;\; f^- = \max\{-f, 0\},$$

com $f^+, f^- \geqslant 0$ e $|f| = f^+ + f^-$.
A hipótese $\mathbb{E}[|f(X)|] < \infty$ implica

$$\mathbb{E}[f^+(X)] < \infty \quad \text{e} \quad \mathbb{E}[f^-(X)] < \infty.$$

Exercício 2 — Solução (caso geral)

Caso 3: f geral com $\mathbb{E}[|f(X)|] < \infty$. Escreva

$$f = f^+ - f^-, \qquad f^+ = \max\{f, 0\}, \quad f^- = \max\{-f, 0\},$$

com $f^+, f^- \geqslant 0$ e $|f| = f^+ + f^-$.
A hipótese $\mathbb{E}[|f(X)|] < \infty$ implica

$$\mathbb{E}[f^+(X)] < \infty \quad \text{e} \quad \mathbb{E}[f^-(X)] < \infty.$$

Aplicando o caso $f \geqslant 0$ a f^+ e f^-,

$$\mathbb{E}[f^+(X)] = \int_{\mathbb{R}} f^+(x)\,\mathbb{P}_X(\mathrm{d}x), \qquad \mathbb{E}[f^-(X)] = \int_{\mathbb{R}} f^-(x)\,\mathbb{P}_X(\mathrm{d}x).$$

Exercício 2 — Solução (caso geral)

Caso 3: f geral com $\mathbb{E}[|f(X)|] < \infty$. Escreva

$$f = f^+ - f^-, \qquad f^+ = \max\{f, 0\}, \;\; f^- = \max\{-f, 0\},$$

com $f^+, f^- \geqslant 0$ e $|f| = f^+ + f^-$.
A hipótese $\mathbb{E}[|f(X)|] < \infty$ implica

$$\mathbb{E}[f^+(X)] < \infty \quad \text{e} \quad \mathbb{E}[f^-(X)] < \infty.$$

Aplicando o caso $f \geqslant 0$ a f^+ e f^-,

$$\mathbb{E}[f^+(X)] = \int_{\mathbb{R}} f^+(x)\,\mathbb{P}_X(\mathrm{d}x), \qquad \mathbb{E}[f^-(X)] = \int_{\mathbb{R}} f^-(x)\,\mathbb{P}_X(\mathrm{d}x).$$

Subtraindo, (pela linearidade da esperança e da integral),

$$\mathbb{E}[f(X)] = \mathbb{E}[f^+(X)] - \mathbb{E}[f^-(X)] = \int_{\mathbb{R}} f^+(x)\,\mathbb{P}_X(\mathrm{d}x) - \int_{\mathbb{R}} f^-(x)\,\mathbb{P}_X(\mathrm{d}x)$$

$$= \int_{\mathbb{R}} f(x)\,\mathbb{P}_X(\mathrm{d}x).$$

Exercício 2 — Comentário

Este resultado mostra que, para calcular $\mathbb{E}[f(X)]$, basta conhecer a *distribuição de* X (isto é, $\mathbb{P}_X$), sem precisar determinar explicitamente a lei de $f(X)$. Pela definição,

$$\mathbb{E}[f(X)] = \int_\Omega f(X(\omega))\,\mathbb{P}(\mathrm{d}\omega) = \int_\mathbb{R} y\,\mathbb{P}_{f(X)}(\mathrm{d}y),$$

e o teorema garante a forma equivalente

$$\mathbb{E}[f(X)] = \int_\mathbb{R} f(x)\,\mathbb{P}_X(\mathrm{d}x).$$

Veja que **nada sabemos (pelo exposto em aula) sobre** $\mathbb{P}_{f(X)}$, esse teorema mostra que isso é uma complicação desnecessária e pode ser facilmente ignorada (tanto que na prática muitas pessoas esquecem que isso é um resultado que precisa ser provado e tomam isso como parte da definição), por isso coloquialmente é chamado de **Teorema do Estatístico Inconsciente**.

Exercício 3 — Jensen (estrita): Enunciado

Enunciado

Demonstre a desigualdade de Jensen *estrita*: seja X uma v.a. integrável em $(\Omega, \Sigma, \mathbb{P})$ e $f : A \to \mathbb{R}$ **estritamente** convexa, com $A \subseteq \mathbb{R}$ convexo e aberto. Suponha $\mathbb{P}(X \in A) = 1$, $\mathbb{E}[|X|] < \infty$ e $\mathbb{E}[|f(X)|] < \infty$. Se $\mathbb{P}(X \neq \mathbb{E}[X]) > 0$, então

$$f(\mathbb{E}[X]) \; < \; \mathbb{E}[f(X)].$$

Exercício 3 — Resolução

Denote

$$m := \mathbb{E}[X].$$

Como A é um intervalo aberto e $X \in A$ q.c., segue que $m \in A$.

Exercício 3 — Resolução

Denote

$$m := \mathbb{E}[X].$$

Como A é um intervalo aberto e $X \in A$ q.c., segue que $m \in A$.
Assuma $\mathbb{P}(X \neq m) > 0$ e considere a partição

$$B := \{X \leqslant m\}, \qquad C := \{X > m\}.$$

Exercício 3 — Resolução

Denote

$$m := \mathbb{E}[X].$$

Como A é um intervalo aberto e $X \in A$ q.c., segue que $m \in A$.
Assuma $\mathbb{P}(X \neq m) > 0$ e considere a partição

$$B := \{X \leqslant m\}, \qquad C := \{X > m\}.$$

Afirmamos que $\mathbb{P}(B), \mathbb{P}(C) \in (0,1)$. De fato, se $\mathbb{P}(C) = 0$, então $X \leqslant m$ q.c.; como $\mathbb{P}(X \neq m) > 0$ isso implica $\mathbb{P}(X < m) > 0$, e portanto $\mathbb{E}[X] < m$, contradição. Analogamente, $\mathbb{P}(B) \neq 0$ e $\mathbb{P}(B) \neq 1$.

Exercício 3 — Resolução (médias condicionais)

Defina as médias condicionais nos dois lados de m:

$$m_1 := \mathbb{E}[X \mid B] = \frac{\mathbb{E}[X\mathbf{1}_B]}{\mathbb{P}(B)}, \qquad m_2 := \mathbb{E}[X \mid C] = \frac{\mathbb{E}[X\mathbf{1}_C]}{\mathbb{P}(C)}.$$

(As quantidades são finitas pois $\mathbb{E}[|X|] < \infty$.)

Exercício 3 — Resolução (médias condicionais)

Defina as médias condicionais nos dois lados de m:

$$m_1 := \mathbb{E}[X \mid B] = \frac{\mathbb{E}[X\mathbf{1}_B]}{\mathbb{P}(B)}, \qquad m_2 := \mathbb{E}[X \mid C] = \frac{\mathbb{E}[X\mathbf{1}_C]}{\mathbb{P}(C)}.$$

(As quantidades são finitas pois $\mathbb{E}[|X|] < \infty$.)
Como $X \leqslant m$ em B e $\mathbb{P}(X < m\,;\, B) > 0$, segue que $m_1 < m$.
Analogamente, como $X > m$ em C, segue que $m_2 > m$. Em particular, $m_1 \neq m_2$.

Exercício 3 — Resolução (médias condicionais)

Defina as médias condicionais nos dois lados de m:

$$m_1 := \mathbb{E}[X \mid B] = \frac{\mathbb{E}[X\mathbf{1}_B]}{\mathbb{P}(B)}, \qquad m_2 := \mathbb{E}[X \mid C] = \frac{\mathbb{E}[X\mathbf{1}_C]}{\mathbb{P}(C)}.$$

(As quantidades são finitas pois $\mathbb{E}[|X|] < \infty$.)
Como $X \leqslant m$ em B e $\mathbb{P}(X < m\,;\,B) > 0$, segue que $m_1 < m$.
Analogamente, como $X > m$ em C, segue que $m_2 > m$. Em particular, $m_1 \neq m_2$.
Além disso, pela decomposição da esperança,

$$m = \mathbb{E}[X] = \mathbb{E}[X\mathbf{1}_B] + \mathbb{E}[X\mathbf{1}_C] = \mathbb{P}(B)\,m_1 + \mathbb{P}(C)\,m_2.$$

Exercício 3 — Resolução (estrita convexidade)

Como $\mathbb{P}(B) \in (0,1)$ e $m_1 \neq m_2$, temos que m é combinação convexa *estrita* de m_1 e m_2:

$$m = \mathbb{P}(B)\, m_1 + \mathbb{P}(C)\, m_2, \qquad \mathbb{P}(C) = 1 - \mathbb{P}(B).$$

Exercício 3 — Resolução (estrita convexidade)

Como $\mathbb{P}(B) \in (0, 1)$ e $m_1 \neq m_2$, temos que m é combinação convexa *estrita* de m_1 e m_2:

$$m = \mathbb{P}(B)\, m_1 + \mathbb{P}(C)\, m_2, \qquad \mathbb{P}(C) = 1 - \mathbb{P}(B).$$

Pela **estrita convexidade** de f em A,

$$f(m) = f(\mathbb{P}(B)\, m_1 + \mathbb{P}(C)\, m_2) \;<\; \mathbb{P}(B)\, f(m_1) + \mathbb{P}(C)\, f(m_2).$$

Exercício 3 — Resolução (Jensen e conclusão)

Aplique Jensen (não-estrita) sob as probabilidades condicionais:

$$\mathbb{E}[f(X) \mid B] \;\geqslant\; f(\mathbb{E}[X \mid B]) = f(m_1).$$

$$\mathbb{E}[f(X) \mid C] \;\geqslant\; f(\mathbb{E}[X \mid C]) = f(m_2).$$

Exercício 3 — Resolução (Jensen e conclusão)

Aplique Jensen (não-estrita) sob as probabilidades condicionais:

$$\mathbb{E}[f(X) \mid B] \;\geqslant\; f(\mathbb{E}[X \mid B]) = f(m_1).$$

$$\mathbb{E}[f(X) \mid C] \;\geqslant\; f(\mathbb{E}[X \mid C]) = f(m_2).$$

Multiplicando pelos pesos e somando (lei da esperança total),

$$\mathbb{E}[f(X)] = \mathbb{P}(B)\,\mathbb{E}[f(X) \mid B] + \mathbb{P}(C)\,\mathbb{E}[f(X) \mid C] \;\geqslant\; \mathbb{P}(B)\,f(m_1) + \mathbb{P}(C)\,f(m_2).$$

Exercício 3 — Resolução (Jensen e conclusão)

Aplique Jensen (não-estrita) sob as probabilidades condicionais:

$$\mathbb{E}[f(X) \mid B] \;\geqslant\; f(\mathbb{E}[X \mid B]) = f(m_1).$$

$$\mathbb{E}[f(X) \mid C] \;\geqslant\; f(\mathbb{E}[X \mid C]) = f(m_2).$$

Multiplicando pelos pesos e somando (lei da esperança total),

$$\mathbb{E}[f(X)] = \mathbb{P}(B)\,\mathbb{E}[f(X) \mid B] + \mathbb{P}(C)\,\mathbb{E}[f(X) \mid C] \;\geqslant\; \mathbb{P}(B)\,f(m_1) + \mathbb{P}(C)\,f(m_2).$$

Combinando com a desigualdade estrita do slide anterior,

$$f(\mathbb{E}[X]) = f(m) \;<\; \mathbb{P}(B)\,f(m_1) + \mathbb{P}(C)\,f(m_2) \;\leqslant\; \mathbb{E}[f(X)].$$

Logo, $f(\mathbb{E}[X]) < \mathbb{E}[f(X)]$, como desejado.

Exercício 4 — Enunciado

Enunciado

Seja $(\Omega, \Sigma, \mathbb{P})$ um espaço de probabilidade. Para $A, B \in \mathcal{L}_2(\Omega, \Sigma, \mathbb{P})$, a covariância é definida por

$$\mathbb{C}(A, B) \coloneqq \mathbb{E}[AB] - \mathbb{E}[A]\mathbb{E}[B].$$

Para duas variáveis aleatórias X, Y em $(\Omega, \Sigma, \mathbb{P})$, mostre que X é independente de Y se, e somente se, para quaisquer funções $h, g : \mathbb{R} \to \mathbb{R}$ limitadas e mensuráveis,

$$\mathbb{C}\big(h(X), g(Y)\big) = 0.$$

Dica: para "$\Leftarrow$", use indicadoras $h = \mathbf{1}_A$, $g = \mathbf{1}_B$. Para "$\Rightarrow$", prove para funções simples e aproxime funções limitadas por funções-escada, usando convergência limitada.

Exercício 4 — Resolução (redução ao produto das esperanças)

Como h e g são limitadas, $h(X), g(Y) \in \mathcal{L}_2$ e as esperanças envolvidas são finitas.

Além disso,

$$\mathbb{C}\big(h(X), g(Y)\big) = 0 \quad \iff \quad \mathbb{E}\big[h(X)g(Y)\big] = \mathbb{E}[h(X)]\,\mathbb{E}[g(Y)].$$

Assim, basta provar a equivalência entre **independência** e a **fatoração** acima para toda h, g limitadas e mensuráveis.

Exercício 4 — Resolução ($\Leftarrow$)

Suponha que, para quaisquer h, g limitadas e mensuráveis,

$$\mathbb{E}\big[h(X)g(Y)\big] = \mathbb{E}[h(X)]\,\mathbb{E}[g(Y)].$$

Exercício 4 — Resolução ($\Leftarrow$)

Suponha que, para quaisquer h, g limitadas e mensuráveis,

$$\mathbb{E}\big[h(X)g(Y)\big] = \mathbb{E}[h(X)]\,\mathbb{E}[g(Y)].$$

Tome $A, B \in \mathcal{B}(\mathbb{R})$ e escolha $h = \mathbf{1}_A$ e $g = \mathbf{1}_B$. Então

$$\mathbb{E}[\mathbf{1}_A(X)\mathbf{1}_B(Y)] = \mathbb{E}[\mathbf{1}_A(X)]\;\mathbb{E}[\mathbf{1}_B(Y)]\,.$$

Exercício 4 — Resolução ($\Leftarrow$)

Suponha que, para quaisquer h, g limitadas e mensuráveis,

$$\mathbb{E}[h(X)g(Y)] = \mathbb{E}[h(X)]\,\mathbb{E}[g(Y)].$$

Tome $A, B \in \mathcal{B}(\mathbb{R})$ e escolha $h = \mathbf{1}_A$ e $g = \mathbf{1}_B$. Então

$$\mathbb{E}[\mathbf{1}_A(X)\mathbf{1}_B(Y)] = \mathbb{E}[\mathbf{1}_A(X)]\;\mathbb{E}[\mathbf{1}_B(Y)]\,.$$

Usando $\mathbf{1}_A(X) = \mathbf{1}_{\{X \in A\}}$ e $\mathbb{E}[\mathbf{1}_E] = \mathbb{P}(E)$,

$$\mathbb{P}(X \in A,\ Y \in B) = \mathbb{P}(X \in A)\,\mathbb{P}(Y \in B).$$

Como A, B são arbitrários, segue que X e Y são independentes.

Exercício 4 — Resolução ($\Rightarrow$): indicadoras

Agora suponha que X e Y são independentes, isto é, para todo $A, B \in \mathcal{B}(\mathbb{R})$,

$$\mathbb{P}(X \in A,\ Y \in B) = \mathbb{P}(X \in A)\,\mathbb{P}(Y \in B).$$

Exercício 4 — Resolução ($\Rightarrow$): indicadoras

Agora suponha que X e Y são independentes, isto é, para todo $A, B \in \mathcal{B}(\mathbb{R})$,

$$\mathbb{P}(X \in A,\ Y \in B) = \mathbb{P}(X \in A)\,\mathbb{P}(Y \in B).$$

Para indicadoras, obtemos imediatamente:

$$\mathbb{E}[\mathbf{1}_A(X)\mathbf{1}_B(Y)] = \mathbb{P}(X \in A,\ Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B) = \mathbb{E}[\mathbf{1}_A(X)]\,\mathbb{E}[\mathbf{1}_B(Y)$$

Exercício 4 — Resolução ($\Rightarrow$): funções simples

Sejam

$$h = \sum_{i=1}^{k} a_i \, \mathbf{1}_{A_i}, \qquad g = \sum_{j=1}^{m} b_j \, \mathbf{1}_{B_j},$$

funções simples, com $A_i, B_j \in \mathcal{B}(\mathbb{R})$ (podemos supor as famílias disjuntas).

Exercício 4 — Resolução ($\Rightarrow$): funções simples

Sejam

$$h = \sum_{i=1}^{k} a_i \, \mathbf{1}_{A_i}, \qquad g = \sum_{j=1}^{m} b_j \, \mathbf{1}_{B_j},$$

funções simples, com $A_i, B_j \in \mathcal{B}(\mathbb{R})$ (podemos supor as famílias disjuntas).
Então

$$h(X)g(Y) = \sum_{i=1}^{k} \sum_{j=1}^{m} a_i b_j \, \mathbf{1}_{A_i}(X) \mathbf{1}_{B_j}(Y),$$

e, pela linearidade da esperança,

$$\mathbb{E}[h(X)g(Y)] = \sum_{i=1}^{k} \sum_{j=1}^{m} a_i b_j \, \mathbb{E}\Big[\mathbf{1}_{A_i}(X) \mathbf{1}_{B_j}(Y)\Big] .$$

Exercício 4 — Resolução ($\Rightarrow$): funções simples

Sejam

$$h = \sum_{i=1}^{k} a_i \, \mathbf{1}_{A_i}, \qquad g = \sum_{j=1}^{m} b_j \, \mathbf{1}_{B_j},$$

funções simples, com $A_i, B_j \in \mathcal{B}(\mathbb{R})$ (podemos supor as famílias disjuntas).
Então

$$h(X)g(Y) = \sum_{i=1}^{k} \sum_{j=1}^{m} a_i b_j \, \mathbf{1}_{A_i}(X) \mathbf{1}_{B_j}(Y),$$

e, pela linearidade da esperança,

$$\mathbb{E}[h(X)g(Y)] = \sum_{i=1}^{k} \sum_{j=1}^{m} a_i b_j \, \mathbb{E}\Big[\mathbf{1}_{A_i}(X) \mathbf{1}_{B_j}(Y)\Big].$$

Pelo caso de indicadoras (independência),

$$\mathbb{E}\Big[\mathbf{1}_{A_i}(X) \mathbf{1}_{B_j}(Y)\Big] = \mathbb{E}[\mathbf{1}_{A_i}(X)] \, \mathbb{E}[\mathbf{1}_{B_j}(Y)].$$

Exercício 4 — Resolução (⇒): funções simples

Logo,

$$\begin{aligned}
\mathbb{E}[h(X)g(Y)] &= \sum_{i=1}^{k}\sum_{j=1}^{m} a_i b_j \,\mathbb{E}[\mathbf{1}_{A_i}(X)]\,\mathbb{E}[\mathbf{1}_{B_j}(Y)] \\
&= \left(\sum_{i=1}^{k} a_i \,\mathbb{E}[\mathbf{1}_{A_i}(X)]\right)\left(\sum_{j=1}^{m} b_j \,\mathbb{E}[\mathbf{1}_{B_j}(Y)]\right) \\
&= \mathbb{E}[h(X)]\,\mathbb{E}[g(Y)].
\end{aligned}$$

Exercício 4 — Resolução ($\Rightarrow$): caso geral

Sejam agora $h, g : \mathbb{R} \to \mathbb{R}$ limitadas e mensuráveis. Defina $M_h := \|h\|_\infty$ e $u := h + M_h$, de modo que $0 \leqslant u \leqslant 2M_h$. Escolha funções simples u_n tais que

$$u_n \uparrow u \quad \text{e} \quad 0 \leqslant u_n \leqslant u,$$

e defina $h_n := u_n - M_h$. Então h_n é simples, $|h_n| \leqslant M_h$ e $h_n \to h$ ponto a ponto. Fazendo o mesmo para g, obtemos simples g_n com $|g_n| \leqslant M_g := \|g\|_\infty$ e $g_n \to g$ ponto a ponto.

Exercício 4 — Resolução ($\Rightarrow$): caso geral

Sejam agora $h, g : \mathbb{R} \to \mathbb{R}$ limitadas e mensuráveis. Defina $M_h := \|h\|_\infty$ e $u := h + M_h$, de modo que $0 \leqslant u \leqslant 2M_h$. Escolha funções simples u_n tais que

$$u_n \uparrow u \quad \text{e} \quad 0 \leqslant u_n \leqslant u,$$

e defina $h_n := u_n - M_h$. Então h_n é simples, $|h_n| \leqslant M_h$ e $h_n \to h$ ponto a ponto. Fazendo o mesmo para g, obtemos simples g_n com $|g_n| \leqslant M_g := \|g\|_\infty$ e $g_n \to g$ ponto a ponto.
Pelo passo anterior, para todo n,

$$\mathbb{E}[h_n(X)g_n(Y)] = \mathbb{E}[h_n(X)]\,\mathbb{E}[g_n(Y)].$$

Além disso,

$$h_n(X) \to h(X), \quad g_n(Y) \to g(Y), \quad h_n(X)g_n(Y) \to h(X)g(Y),$$

e $|h_n(X)g_n(Y)| \leqslant M_h M_g$.

Exercício 4 — Resolução (⇒): caso geral

Pelo Teorema da Convergência Limitada (ou Dominada),

Exercício 4 — Resolução ($\Rightarrow$): caso geral

Pelo Teorema da Convergência Limitada (ou Dominada),

$$\mathbb{E}[h_n(X)] \to \mathbb{E}[h(X)], \quad \mathbb{E}[g_n(Y)] \to \mathbb{E}[g(Y)],$$

$$\mathbb{E}[h_n(X)g_n(Y)] \to \mathbb{E}[h(X)g(Y)].$$

Passando ao limite na identidade para cada n, obtemos

$$\mathbb{E}[h(X)g(Y)] = \mathbb{E}[h(X)]\,\mathbb{E}[g(Y)].$$

Logo,

$$\mathbb{C}(h(X), g(Y)) = 0.$$

Combinando com a implicação anterior, concluímos a equivalência.

Exercício 5 — Enunciado

Contexto

Seja $(\Omega, \Sigma, \mathbb{P})$ um espaço de probabilidade, $Y \in L^2(\Omega, \Sigma, \mathbb{P})$ e $\mathcal{G}$ uma sub-σ-álgebra de Σ.

Objetivos

Mostrar:

(a) $\mathbb{E}[Y \mid \mathcal{G}] \in L^2(\Omega, \Sigma, \mathbb{P})$;

(b) $\mathbb{E}[Y \mid \mathcal{G}]$ é o único minimizador (q.c.) de $\min_{S \in L^2(\Omega, \mathcal{G}, \mathbb{P})} \mathbb{E}[(Y - S)^2]$;

(c) Casos extremos: $\mathcal{G} = \{\varnothing, \Omega\}$ e $\mathcal{G} = \Sigma$ e a interpretação via (b);

(d) Caso $\mathcal{G} = \sigma(E) = \{\varnothing, E, E^c, \Omega\}$ e interpretação via (b).

Exercício 5(a) — Enunciado

Enunciado

Se $Y \in L^2(\Omega, \Sigma, \mathbb{P})$ e $\mathcal{G} \subseteq \Sigma$ é uma sub-σ-álgebra, mostre que

$$\mathbb{E}[Y \mid \mathcal{G}] \in L^2(\Omega, \Sigma, \mathbb{P}).$$

Exercício 5(a) — Resolução

Como $Y \in L^2(\Omega, \Sigma, \mathbb{P})$, em particular $Y \in L^1$ (pois $\|Y\|_1 \leqslant \|Y\|_2$ em medida de probabilidade). Logo $\mathbb{E}[Y \mid \mathcal{G}]$ está bem definida.

Exercício 5(a) — Resolução

Como $Y \in L^2(\Omega, \Sigma, \mathbb{P})$, em particular $Y \in L^1$ (pois $\|Y\|_1 \leqslant \|Y\|_2$ em medida de probabilidade). Logo $\mathbb{E}[Y \mid \mathcal{G}]$ está bem definida.
Aplique a Jensen condicional à função convexa $\varphi(t) = t^2$:

$$\varphi(\mathbb{E}[Y \mid \mathcal{G}]) \leqslant \mathbb{E}[\varphi(Y) \mid \mathcal{G}] \quad \Longrightarrow \quad \big(\mathbb{E}[Y \mid \mathcal{G}]\big)^2 \leqslant \mathbb{E}[Y^2 \mid \mathcal{G}] \qquad \text{q.c.}$$

Exercício 5(a) — Resolução

Como $Y \in L^2(\Omega, \Sigma, \mathbb{P})$, em particular $Y \in L^1$ (pois $\|Y\|_1 \leqslant \|Y\|_2$ em medida de probabilidade). Logo $\mathbb{E}[Y \mid \mathcal{G}]$ está bem definida.
Aplique a Jensen condicional à função convexa $\varphi(t) = t^2$:

$$\varphi(\mathbb{E}[Y \mid \mathcal{G}]) \leqslant \mathbb{E}[\varphi(Y) \mid \mathcal{G}] \quad \Longrightarrow \quad \big(\mathbb{E}[Y \mid \mathcal{G}]\big)^2 \leqslant \mathbb{E}[Y^2 \mid \mathcal{G}] \qquad \text{q.c.}$$

Tomando esperança e usando a propriedade de preservação da esperança ("torre"),

$$\mathbb{E}\Big[\big(\mathbb{E}[Y \mid \mathcal{G}]\big)^2\Big] \leqslant \mathbb{E}\Big[\mathbb{E}[Y^2 \mid \mathcal{G}]\Big] = \mathbb{E}[Y^2] < \infty.$$

Portanto, $\mathbb{E}[Y \mid \mathcal{G}] \in L^2(\Omega, \Sigma, \mathbb{P})$.

Exercício 5(b) — Enunciado

Enunciado

Mostre que $\mathbb{E}[Y \mid \mathcal{G}]$ é a única (a menos de um evento de probabilidade zero) solução do problema:

$$\min_{S \in L^2(\Omega, \mathcal{G}, \mathbb{P})} \mathbb{E}\big[(Y - S)^2\big].$$

Exercício 5(b) — Resolução (decomposição)

Fixe $S \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ e defina

$$M := \mathbb{E}[Y \mid \mathcal{G}].$$

Pelo item (a), $M \in L^2$ e M é $\mathcal{G}$-mensurável, logo $S - M \in L^2(\Omega, \mathcal{G}, \mathbb{P})$.

Exercício 5(b) — Resolução (decomposição)

Fixe $S \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ e defina

$$M := \mathbb{E}[Y \mid \mathcal{G}].$$

Pelo item (a), $M \in L^2$ e M é $\mathcal{G}$-mensurável, logo $S - M \in L^2(\Omega, \mathcal{G}, \mathbb{P})$. Expanda:

$$(Y-S)^2 = (Y-M+M-S)^2 = (Y-M)^2+(S-M)^2+2(Y-M)(M-S).$$

Tomando esperança,

$$\mathbb{E}[(Y-S)^2] = \mathbb{E}[(Y-M)^2] + \mathbb{E}[(S-M)^2] + 2\,\mathbb{E}[(Y-M)(M-S)].$$

Exercício 5(b) — Resolução (decomposição)

Fixe $S \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ e defina

$$M := \mathbb{E}[Y \mid \mathcal{G}].$$

Pelo item (a), $M \in L^2$ e M é $\mathcal{G}$-mensurável, logo $S - M \in L^2(\Omega, \mathcal{G}, \mathbb{P})$. Expanda:

$$(Y-S)^2 = (Y-M+M-S)^2 = (Y-M)^2+(S-M)^2+2(Y-M)(M-S).$$

Tomando esperança,

$$\mathbb{E}[(Y-S)^2] = \mathbb{E}[(Y-M)^2] + \mathbb{E}[(S-M)^2] + 2\,\mathbb{E}[(Y-M)(M-S)].$$

Mostremos que o termo cruzado é nulo. Primeiro,

$$\mathbb{E}[Y - M \mid \mathcal{G}] = \mathbb{E}[Y \mid \mathcal{G}] - \mathbb{E}[M \mid \mathcal{G}] = M - M = 0 \quad \text{q.c.}$$

Como $M - S$ é $\mathcal{G}$-mensurável,

$$\mathbb{E}[(Y-M)(M-S) \mid \mathcal{G}] = (M-S)\,\mathbb{E}[Y - M \mid \mathcal{G}] = 0 \quad \text{q.c.}$$

Logo, pela torre,

$$\mathbb{E}[(Y-M)(M-S)] = 0.$$

Exercício 5(b) — Resolução (minimalidade e unicidade)

Com o termo cruzado nulo,

$$\mathbb{E}[(Y-S)^2] = \mathbb{E}[(Y-M)^2] + \mathbb{E}[(S-M)^2] \geqslant \mathbb{E}[(Y-M)^2],$$

com igualdade se, e somente se, $\mathbb{E}[(S-M)^2] = 0$.

Exercício 5(b) — Resolução (minimalidade e unicidade)

Com o termo cruzado nulo,

$$\mathbb{E}[(Y-S)^2] = \mathbb{E}[(Y-M)^2] + \mathbb{E}[(S-M)^2] \;\geqslant\; \mathbb{E}[(Y-M)^2],$$

com igualdade se, e somente se, $\mathbb{E}[(S-M)^2] = 0$.
Assim, $M = \mathbb{E}[Y \mid \mathcal{G}]$ minimiza $\mathbb{E}[(Y-S)^2]$ em $L^2(\Omega, \mathcal{G}, \mathbb{P})$.

Exercício 5(b) — Resolução (minimalidade e unicidade)

Com o termo cruzado nulo,

$$\mathbb{E}[(Y-S)^2] = \mathbb{E}[(Y-M)^2] + \mathbb{E}[(S-M)^2] \;\geqslant\; \mathbb{E}[(Y-M)^2],$$

com igualdade se, e somente se, $\mathbb{E}[(S-M)^2] = 0$.
Assim, $M = \mathbb{E}[Y \mid \mathcal{G}]$ minimiza $\mathbb{E}[(Y-S)^2]$ em $L^2(\Omega, \mathcal{G}, \mathbb{P})$.
Unicidade (q.c.). Se $S^\star \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ também atinge o mínimo, então

$$\mathbb{E}[(Y-S^\star)^2] = \mathbb{E}[(Y-M)^2] \quad \Longrightarrow \quad \mathbb{E}[(S^\star - M)^2] = 0.$$

Como $(S^\star - M)^2 \geqslant 0$, segue que $(S^\star - M)^2 = 0$ quase certamente, isto é,

$$S^\star = M \quad \text{q.c.}$$

Portanto, $\mathbb{E}[Y \mid \mathcal{G}]$ é o único minimizador q.c.

Exercício 5(c) — Enunciado

Enunciado

Mostre que:

- se $\mathcal{G} = \{\varnothing, \Omega\}$, então $\mathbb{E}[Y \mid \mathcal{G}] = \mathbb{E}[Y]$;
- se $\mathcal{G} = \Sigma$, então $\mathbb{E}[Y \mid \mathcal{G}] = Y$.

À luz do item (b), qual a interpretação desses resultados?

Exercício 5(c) — Resolução e interpretação

Caso $\mathcal{G} = \{\varnothing, \Omega\}$. Toda v.a. $\mathcal{G}$-mensurável é constante quase certamente. Logo existe $c \in \mathbb{R}$ tal que $\mathbb{E}[Y \mid \mathcal{G}] = c$ q.c. Pela propriedade de preservação da esperança,

$$c = \mathbb{E}[c] = \mathbb{E}[\mathbb{E}[Y \mid \mathcal{G}]] = \mathbb{E}[Y].$$

Portanto, $\mathbb{E}[Y \mid \mathcal{G}] = \mathbb{E}[Y]$ q.c.

Exercício 5(c) — Resolução e interpretação

Caso $\mathcal{G} = \{\varnothing, \Omega\}$. Toda v.a. $\mathcal{G}$-mensurável é constante quase certamente. Logo existe $c \in \mathbb{R}$ tal que $\mathbb{E}[Y \mid \mathcal{G}] = c$ q.c. Pela propriedade de preservação da esperança,

$$c = \mathbb{E}[c] = \mathbb{E}[\mathbb{E}[Y \mid \mathcal{G}]] = \mathbb{E}[Y].$$

Portanto, $\mathbb{E}[Y \mid \mathcal{G}] = \mathbb{E}[Y]$ q.c.

Caso $\mathcal{G} = \Sigma$. Aqui Y é Σ-mensurável e Y satisfaz a caracterização de esperança condicional: para todo $A \in \Sigma$, $\mathbb{E}[Y\mathbf{1}_A] = \mathbb{E}[Y\mathbf{1}_A]$. Logo,

$$\mathbb{E}[Y \mid \Sigma] = Y \quad \text{q.c.}$$

Exercício 5(c) — Resolução e interpretação

Caso $\mathcal{G} = \{\varnothing, \Omega\}$. Toda v.a. $\mathcal{G}$-mensurável é constante quase certamente. Logo existe $c \in \mathbb{R}$ tal que $\mathbb{E}[Y \mid \mathcal{G}] = c$ q.c. Pela propriedade de preservação da esperança,

$$c = \mathbb{E}[c] = \mathbb{E}[\mathbb{E}[Y \mid \mathcal{G}]] = \mathbb{E}[Y].$$

Portanto, $\mathbb{E}[Y \mid \mathcal{G}] = \mathbb{E}[Y]$ q.c.

Caso $\mathcal{G} = \Sigma$. Aqui Y é Σ-mensurável e Y satisfaz a caracterização de esperança condicional: para todo $A \in \Sigma$, $\mathbb{E}[Y\mathbf{1}_A] = \mathbb{E}[Y\mathbf{1}_A]$. Logo,

$$\mathbb{E}[Y \mid \Sigma] = Y \quad \text{q.c.}$$

Interpretação (item (b)). O item (b) diz que $\mathbb{E}[Y \mid \mathcal{G}]$ é a *melhor aproximação quadrática* de Y dentre as variáveis $\mathcal{G}$-mensuráveis. Quando $\mathcal{G}$ é trivial, só podemos escolher constantes, e a melhor é $\mathbb{E}[Y]$. Quando $\mathcal{G} = \Sigma$, podemos escolher $S = Y$, e o erro mínimo é 0.

Exercício 5(d) — Enunciado

Enunciado

Seja $E \in \Sigma$ com $1 > \mathbb{P}(E) > 0$ e tome

$$\mathcal{G} = \{\varnothing, E, E^c, \Omega\}.$$

Mostre que a função

$$g(\omega) = \begin{cases} \dfrac{\mathbb{E}[Y\mathbf{1}_E]}{\mathbb{P}(E)}, & \omega \in E, \\ \dfrac{\mathbb{E}[Y\mathbf{1}_{E^c}]}{\mathbb{P}(E^c)}, & \omega \in E^c, \end{cases}$$

é uma versão de $\mathbb{E}[Y \mid \mathcal{G}]$. Interprete o resultado à luz do item (b).

Exercício 5(d) — Resolução (verificação)

Defina $Z : \Omega \to \mathbb{R}$ por

$$Z(\omega) = \frac{\mathbb{E}[Y\mathbf{1}_E]}{\mathbb{P}(E)}\,\mathbf{1}_E(\omega) + \frac{\mathbb{E}[Y\mathbf{1}_{E^c}]}{\mathbb{P}(E^c)}\,\mathbf{1}_{E^c}(\omega).$$

Então Z é constante em E e em E^c, logo Z é $\mathcal{G}$-mensurável.

Exercício 5(d) — Resolução (verificação)

Defina $Z : \Omega \to \mathbb{R}$ por

$$Z(\omega) = \frac{\mathbb{E}[Y\mathbf{1}_E]}{\mathbb{P}(E)}\,\mathbf{1}_E(\omega) + \frac{\mathbb{E}[Y\mathbf{1}_{E^c}]}{\mathbb{P}(E^c)}\,\mathbf{1}_{E^c}(\omega).$$

Então Z é constante em E e em E^c, logo Z é $\mathcal{G}$-mensurável.
Além disso, $Z \in L^2$. De fato, por Cauchy–Schwarz,

$$\big(\mathbb{E}[Y\mathbf{1}_E]\big)^2 \leqslant \mathbb{E}[Y^2\mathbf{1}_E]\,\mathbb{P}(E) \leqslant \mathbb{E}[Y^2]\,\mathbb{P}(E),$$

e analogamente para E^c. Assim,

$$\mathbb{E}[Z^2] = \frac{(\mathbb{E}[Y\mathbf{1}_E])^2}{\mathbb{P}(E)} + \frac{(\mathbb{E}[Y\mathbf{1}_{E^c}])^2}{\mathbb{P}(E^c)} \leqslant \mathbb{E}[Y^2] + \mathbb{E}[Y^2] = 2\mathbb{E}[Y^2] < \infty.$$

Exercício 5(d) — Resolução (verificação)

Para verificar que Z é uma versão de $\mathbb{E}[Y \mid \mathcal{G}]$, basta checar a identidade

$$\mathbb{E}[Z\mathbf{1}_A] = \mathbb{E}[Y\mathbf{1}_A] \quad \text{para todo } A \in \mathcal{G}.$$

Exercício 5(d) — Resolução (verificação)

Para verificar que Z é uma versão de $\mathbb{E}[Y \mid \mathcal{G}]$, basta checar a identidade

$$\mathbb{E}[Z\mathbf{1}_A] = \mathbb{E}[Y\mathbf{1}_A] \quad \text{para todo } A \in \mathcal{G}.$$

Como $\mathcal{G}$ é gerada por E e E^c, é suficiente verificar para $A = E$ e $A = E^c$:

$$\mathbb{E}[Z\mathbf{1}_E] = \mathbb{E}\left[\frac{\mathbb{E}[Y\mathbf{1}_E]}{\mathbb{P}(E)}\,\mathbf{1}_E\right] = \frac{\mathbb{E}[Y\mathbf{1}_E]}{\mathbb{P}(E)}\,\mathbb{P}(E) = \mathbb{E}[Y\mathbf{1}_E],$$

e analogamente, $\mathbb{E}[Z\mathbf{1}_{E^c}] = \mathbb{E}[Y\mathbf{1}_{E^c}]$. Logo $Z = g$ é uma versão de $\mathbb{E}[Y \mid \mathcal{G}]$.

Exercício 5(d) — Interpretação (via item (b))

Pelo item (b), $\mathbb{E}[Y \mid \mathcal{G}]$ é a melhor aproximação de Y em L^2 dentre as variáveis $\mathcal{G}$-mensuráveis.

Exercício 5(d) — Interpretação (via item (b))

Pelo item (b), $\mathbb{E}[Y \mid \mathcal{G}]$ é a melhor aproximação de Y em L^2 dentre as variáveis $\mathcal{G}$-mensuráveis.
Quando $\mathcal{G} = \{\varnothing, E, E^c, \Omega\}$, a restrição "$S$ é $\mathcal{G}$-mensurável" significa:

$$S \text{ é constante em } E \text{ e constante em } E^c.$$

Ou seja, o previsor só pode usar a *informação binária* "ocorreu E ou ocorreu E^c".

Exercício 5(d) — Interpretação (via item (b))

Pelo item (b), $\mathbb{E}[Y \mid \mathcal{G}]$ é a melhor aproximação de Y em L^2 dentre as variáveis $\mathcal{G}$-mensuráveis.
Quando $\mathcal{G} = \{\varnothing, E, E^c, \Omega\}$, a restrição "$S$ é $\mathcal{G}$-mensurável" significa:

$$S \text{ é constante em } E \text{ e constante em } E^c.$$

Ou seja, o previsor só pode usar a *informação binária* "ocorreu E ou ocorreu E^c".
O minimizador escolhe, em cada átomo da partição $\{E, E^c\}$, o melhor valor constante:

$$Z(\omega) = \underbrace{\mathbb{E}[Y \mid E]}_{= \mathbb{E}[Y\mathbf{1}_E]/\mathbb{P}(E)} \quad \text{em } E, \qquad Z(\omega) = \underbrace{\mathbb{E}[Y \mid E^c]}_{= \mathbb{E}[Y\mathbf{1}_{E^c}]/\mathbb{P}(E^c)} \quad \text{em } E^c.$$

Assim, g é exatamente a previsão quadrática ótima de Y quando só se observa se E ocorreu.

Exercício 6 — Enunciado (contexto)

Contexto

Seja $(\Omega, \Sigma, \mathbb{P})$ e X, Y duas variáveis aleatórias reais. Considere a medida de probabilidade induzida por (X, Y) em $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$,

$$\mathbb{P}_{X,Y}(B) := \mathbb{P}\big((X, Y) \in B\big), \qquad B \in \mathcal{B}(\mathbb{R}^2).$$

Dizemos que $\mathbb{P}_{X,Y}$ admite densidade f com respeito à medida de Lebesgue λ^2 (seja $\lambda^2 = \lambda \otimes \lambda$ em $\mathbb{R}^2$)[a] se, para todo $B \in \mathcal{B}(\mathbb{R}^2)$,

$$\mathbb{P}_{X,Y}(B) = \int_{\mathbb{R}^2} f(\omega_1, \omega_2)\, \mathbf{1}_B(\omega_1, \omega_2)\, \lambda^2(d\omega_1\, d\omega_2).$$

[a] A medida de Lebesgue em $\mathbb{R}^2$ é a medida produto $\lambda \otimes \lambda$ sobre $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^2)$.

Exercício 6(a) — Enunciado

Enunciado

Mostre que

$$f_1(\omega) := \int f(\omega, y)\, \lambda(dy)$$

define uma densidade para a probabilidade $\mathbb{P}_X$ induzida por X em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Dica: tome $A \in \mathcal{B}(\mathbb{R})$ e use Fubini para mostrar que

$$\mathbb{P}_X(A) = \mathbb{P}_{X,Y}(A \times \mathbb{R}) = \int \mathbf{1}_{A\times\mathbb{R}}\, f\, d\lambda^2 = \int \mathbf{1}_A\, f_1\, d\lambda.$$

Exercício 6(a) — Resolução

Como $\mathbb{P}_{X,Y}$ tem densidade f em relação a λ^2, temos $f \geqslant 0$ mensurável e

$$\int_{\mathbb{R}^2} f(x,y)\,\lambda^2(\mathrm{d}(x,y)) = \mathbb{P}_{X,Y}(\mathbb{R}^2) = 1,$$

logo $f \in L^1(\mathbb{R}^2, \lambda^2)$.

Exercício 6(a) — Resolução

Como $\mathbb{P}_{X,Y}$ tem densidade f em relação a λ^2, temos $f \geqslant 0$ mensurável e

$$\int_{\mathbb{R}^2} f(x,y)\,\lambda^2(\mathrm{d}(x,y)) = \mathbb{P}_{X,Y}(\mathbb{R}^2) = 1,$$

logo $f \in L^1(\mathbb{R}^2, \lambda^2)$.

Pelo Teorema de Fubini (ou Tonelli),

$$f_1(x) := \int_{\mathbb{R}} f(x,y)\,\lambda(\mathrm{d}y)$$

está bem definida para λ-q.t.p. x, é mensurável e não negativa.

Exercício 6(a) — Resolução

Como $\mathbb{P}_{X,Y}$ tem densidade f em relação a λ^2, temos $f \geqslant 0$ mensurável e

$$\int_{\mathbb{R}^2} f(x,y)\,\lambda^2(\mathrm{d}(x,y)) = \mathbb{P}_{X,Y}(\mathbb{R}^2) = 1,$$

logo $f \in L^1(\mathbb{R}^2, \lambda^2)$.
Pelo Teorema de Fubini (ou Tonelli),

$$f_1(x) := \int_{\mathbb{R}} f(x,y)\,\lambda(\mathrm{d}y)$$

está bem definida para λ-q.t.p. x, é mensurável e não negativa.
Para $A \in \mathcal{B}(\mathbb{R})$,

$$\begin{aligned}
\mathbb{P}_X(A) &= \mathbb{P}(X \in A) = \mathbb{P}\big((X,Y) \in A \times \mathbb{R}\big) = \mathbb{P}_{X,Y}(A \times \mathbb{R}) \\
&= \int_{\mathbb{R}^2} \mathbf{1}_{A\times\mathbb{R}}(x,y)\, f(x,y)\,\lambda^2(\mathrm{d}(x,y)) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathbf{1}_A(x)\, f(x,y)\,\lambda(\mathrm{d}y) \right) \lambda(\mathrm{d}x) \\
&= \int_{\mathbb{R}} \mathbf{1}_A(x)\, f_1(x)\,\lambda(\mathrm{d}x) = \int_A f_1(x)\,\lambda(\mathrm{d}x).
\end{aligned}$$

Exercício 6(a) — Resolução

Como $\mathbb{P}_{X,Y}$ tem densidade f em relação a λ^2, temos $f \geqslant 0$ mensurável e

$$\int_{\mathbb{R}^2} f(x,y)\,\lambda^2(\mathrm{d}(x,y)) = \mathbb{P}_{X,Y}(\mathbb{R}^2) = 1,$$

logo $f \in L^1(\mathbb{R}^2, \lambda^2)$.

Pelo Teorema de Fubini (ou Tonelli),

$$f_1(x) := \int_{\mathbb{R}} f(x,y)\,\lambda(\mathrm{d}y)$$

está bem definida para λ-q.t.p. x, é mensurável e não negativa.

Para $A \in \mathcal{B}(\mathbb{R})$,

$$\begin{aligned}\mathbb{P}_X(A) &= \mathbb{P}(X \in A) = \mathbb{P}\big((X,Y) \in A \times \mathbb{R}\big) = \mathbb{P}_{X,Y}(A \times \mathbb{R}) \\ &= \int_{\mathbb{R}^2} \mathbf{1}_{A\times\mathbb{R}}(x,y)\,f(x,y)\,\lambda^2(\mathrm{d}(x,y)) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathbf{1}_A(x)\,f(x,y)\,\lambda(\mathrm{d}y)\right)\lambda(\mathrm{d}x) \\ &= \int_{\mathbb{R}} \mathbf{1}_A(x)\,f_1(x)\,\lambda(\mathrm{d}x) = \int_A f_1(x)\,\lambda(\mathrm{d}x).\end{aligned}$$

Em particular, tomando $A = \mathbb{R}$ obtemos $\int_{\mathbb{R}} f_1\,\lambda(\mathrm{d}x) = 1$. Portanto, f_1 é uma densidade de $\mathbb{P}_X$ com respeito a λ.

Exercício 6(b) — Enunciado

Enunciado

Suponha que $\mathbb{E}[|Y|] < \infty$. Mostre que a função $g : \mathbb{R} \to \mathbb{R}$ definida por

$$g(x) = \begin{cases} \dfrac{1}{f_1(x)} \displaystyle\int y\, f(y,x)\, \lambda(dy), & \text{se } f_1(x) > 0, \\ 0, & \text{se } f_1(x) = 0, \end{cases}$$

define uma função de esperança condicional de Y em X, isto é, $g(X)$ é uma versão de $\mathbb{E}[Y \mid X]$.

A escolha do valor 0 quando $f_1(x) = 0$ faz alguma diferença? Por quê?

Exercício 6(b) — Resolução (preparação)

Como $\mathbb{E}[|Y|] < \infty$ e $\mathbb{P}_{X,Y}$ tem densidade f, vale (pelo **Teorema do Estatístico Inconsciente** em $\mathbb{R}^2$)

$$\mathbb{E}[|Y|] = \int_{\mathbb{R}^2} |y| \, f(x,y) \, \lambda^2(\mathrm{d}(x,y)) < \infty.$$

Logo, a função $(x,y) \mapsto y f(x,y)$ é integrável em valor absoluto.

Exercício 6(b) — Resolução (preparação)

Como $\mathbb{E}[|Y|] < \infty$ e $\mathbb{P}_{X,Y}$ tem densidade f, vale (pelo **Teorema do Estatístico Inconsciente** em $\mathbb{R}^2$)

$$\mathbb{E}[|Y|] = \int_{\mathbb{R}^2} |y|\, f(x,y)\, \lambda^2(\mathrm{d}(x,y)) < \infty.$$

Logo, a função $(x,y) \mapsto yf(x,y)$ é integrável em valor absoluto.
Pelo Teorema de Fubini, a função

$$m(x) \coloneqq \int_{\mathbb{R}} y\, f(x,y)\, \lambda(\mathrm{d}y)$$

é (finita para λ-q.t.p. x e) mensurável.

Exercício 6(b) — Resolução (preparação)

Como $\mathbb{E}[|Y|] < \infty$ e $\mathbb{P}_{X,Y}$ tem densidade f, vale (pelo **Teorema do Estatístico Inconsciente** em $\mathbb{R}^2$)

$$\mathbb{E}[|Y|] = \int_{\mathbb{R}^2} |y| \, f(x,y) \, \lambda^2(\mathrm{d}(x,y)) < \infty.$$

Logo, a função $(x,y) \mapsto y f(x,y)$ é integrável em valor absoluto.
Pelo Teorema de Fubini, a função

$$m(x) \coloneqq \int_{\mathbb{R}} y \, f(x,y) \, \lambda(\mathrm{d}y)$$

é (finita para λ-q.t.p. x e) mensurável.
Defina então

$$g(x) \coloneqq \begin{cases} \dfrac{m(x)}{f_1(x)}, & f_1(x) > 0, \\ 0, & f_1(x) = 0. \end{cases}$$

Então g é mensurável e $g(X)$ é $\sigma(X)$-mensurável.

Exercício 6(b) — Resolução (preparação)

Além disso, $g(X)$ é integrável:

$$\begin{aligned}\mathbb{E}[|g(X)|] &= \int_{\mathbb{R}} |g(x)|\, f_1(x)\, \lambda(\mathrm{d}x) = \int_{\{f_1>0\}} \left|\frac{m(x)}{f_1(x)}\right| f_1(x)\, \lambda(\mathrm{d}x) \\ &= \int_{\{f_1>0\}} |m(x)|\, \lambda(\mathrm{d}x) \leqslant \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |y|\, f(x,y)\, \lambda(\mathrm{d}y)\right) \lambda(\mathrm{d}x) \\ &= \int_{\mathbb{R}^2} |y|\, f(x,y)\, \lambda^2(\mathrm{d}(x,y)) = \mathbb{E}[|Y|] < \infty.\end{aligned}$$

Exercício 6(b) — Resolução (caracterização)

Para mostrar que $g(X)$ é uma versão de $\mathbb{E}[Y \mid X]$, basta verificar que, para todo $A \in \mathcal{B}(\mathbb{R})$,

$$\mathbb{E}\big[g(X)\mathbf{1}_{\{X\in A\}}\big] = \mathbb{E}\big[Y\mathbf{1}_{\{X\in A\}}\big].$$

Exercício 6(b) — Resolução (caracterização)

Para mostrar que $g(X)$ é uma versão de $\mathbb{E}[Y \mid X]$, basta verificar que, para todo $A \in \mathcal{B}(\mathbb{R})$,

$$\mathbb{E}\big[g(X)\mathbf{1}_{\{X \in A\}}\big] = \mathbb{E}\big[Y\mathbf{1}_{\{X \in A\}}\big].$$

Usando que X tem densidade f_1 (item (a)),

$$\mathbb{E}\big[g(X)\mathbf{1}_{\{X \in A\}}\big] = \int_{\mathbb{R}} g(x)\mathbf{1}_A(x)\, f_1(x)\,\lambda(\mathrm{d}x) = \int_A g(x) f_1(x)\,\lambda(\mathrm{d}x).$$

Exercício 6(b) — Resolução (caracterização)

Para mostrar que $g(X)$ é uma versão de $\mathbb{E}[Y \mid X]$, basta verificar que, para todo $A \in \mathcal{B}(\mathbb{R})$,

$$\mathbb{E}\big[g(X)\mathbf{1}_{\{X\in A\}}\big] = \mathbb{E}\big[Y\mathbf{1}_{\{X\in A\}}\big].$$

Usando que X tem densidade f_1 (item (a)),

$$\mathbb{E}\big[g(X)\mathbf{1}_{\{X\in A\}}\big] = \int_{\mathbb{R}} g(x)\mathbf{1}_A(x)\, f_1(x)\, \lambda(\mathrm{d}x) = \int_A g(x) f_1(x)\, \lambda(\mathrm{d}x).$$

Observação: se $f_1(x) = 0$, então $\int f(x,y)\,\lambda(\mathrm{d}y) = 0$ e, como $f(x,\cdot) \geqslant 0$, temos $f(x,y) = 0$ para λ-q.t.p. y, logo $m(x) = 0$. Portanto, para todo x (fora de um conjunto λ-nulo),

$$g(x) f_1(x) = m(x).$$

Assim,

$$\begin{aligned}\int_A g(x) f_1(x)\,\lambda(\mathrm{d}x) &= \int_A m(x)\,\lambda(\mathrm{d}x) = \int_A \left(\int_{\mathbb{R}} y\, f(x,y)\,\lambda(\mathrm{d}y)\right)\lambda(\mathrm{d}x) \\ &= \int_{\mathbb{R}^2} y\,\mathbf{1}_A(x)\, f(x,y)\,\lambda^2(\mathrm{d}(x,y)).\end{aligned}$$

Exercício 6(b) — Resolução (caracterização)

Assim,

$$\int_A g(x) f_1(x)\,\lambda(\mathrm{d}x) = \int_A m(x)\,\lambda(\mathrm{d}x) = \int_A \left(\int_{\mathbb{R}} y\, f(x,y)\,\lambda(\mathrm{d}y) \right) \lambda(\mathrm{d}x)$$
$$= \int_{\mathbb{R}^2} y\, \mathbf{1}_A(x)\, f(x,y)\,\lambda^2(\mathrm{d}(x,y)).$$

De novo sendo um"estatístico inconsciente" em $\mathbb{R}^2$),

$$\int_{\mathbb{R}^2} y\, \mathbf{1}_A(x)\, f(x,y)\,\lambda^2(\mathrm{d}(x,y)) = \mathbb{E}\big[Y\, \mathbf{1}_{\{X \in A\}}\big].$$

Logo, $g(X)$ satisfaz a caracterização e é uma versão de $\mathbb{E}[Y \mid \sigma(X)]$, isto é, de $\mathbb{E}[Y \mid X]$.

Exercício 6(b) — Comentário: o que ocorre quando $f_1(x) = 0$?

A escolha de $g(x) = 0$ quando $f_1(x) = 0$ **não altera** a esperança condicional, pois o conjunto

$$N \coloneqq \{x \in \mathbb{R} : f_1(x) = 0\}$$

tem $\mathbb{P}_X$-medida nula:

$$\mathbb{P}\big(f_1(X) = 0\big) = \mathbb{P}_X(N) = \int_N f_1(x)\,\lambda(\mathrm{d}x) = 0.$$

Exercício 6(b) — Comentário: o que ocorre quando $f_1(x) = 0$?

A escolha de $g(x) = 0$ quando $f_1(x) = 0$ **não altera** a esperança condicional, pois o conjunto

$$N \coloneqq \{x \in \mathbb{R} : f_1(x) = 0\}$$

tem $\mathbb{P}_X$-medida nula:

$$\mathbb{P}\big(f_1(X) = 0\big) = \mathbb{P}_X(N) = \int_N f_1(x)\,\lambda(\mathrm{d}x) = 0.$$

Assim, qualquer outra definição de g em N produz uma função que coincide com $g(X)$ quase certamente. Como versões de esperança condicional são únicas a menos de eventos de probabilidade zero, a escolha do valor em $\{f_1 = 0\}$ é irrelevante para $\mathbb{E}[Y \mid X]$.

Exercício 7 — Enunciado

Enunciado

Seja $(\Omega, 2^\Omega, \mathbb{P})$ um espaço de probabilidade em que Ω é finito. Defina a *medida de contagem* c sobre $(\Omega, 2^\Omega)$ como, para todo $B \subseteq \Omega$,

$$c(B) = |B| = \sum_{\omega \in \Omega} \mathbf{1}_B(\omega).$$

Mostre que $\mathbb{P}$ admite densidade com respeito a c, e que essa densidade é dada por

$$g(\omega) = \mathbb{P}[\{\omega\}], \qquad \omega \in \Omega.$$

Exercício 7 — Resolução

Como Ω é finito, temos $\Sigma = 2^{\Omega}$ e toda função $g : \Omega \to \mathbb{R}$ é mensurável.

Exercício 7 — Resolução

Como Ω é finito, temos $\Sigma = 2^{\Omega}$ e toda função $g : \Omega \to \mathbb{R}$ é mensurável. Para todo $B \subseteq \Omega$, podemos escrever a decomposição disjunta

$$B = \bigcup_{\omega \in B} \{\omega\}.$$

Como a união é finita, pela aditividade (finita) de $\mathbb{P}$ em uniões disjuntas,

$$\mathbb{P}(B) = \sum_{\omega \in B} \mathbb{P}(\{\omega\}).$$

Exercício 7 — Resolução

Como Ω é finito, temos $\Sigma = 2^{\Omega}$ e toda função $g : \Omega \to \mathbb{R}$ é mensurável. Para todo $B \subseteq \Omega$, podemos escrever a decomposição disjunta

$$B = \bigcup_{\omega \in B} \{\omega\}.$$

Como a união é finita, pela aditividade (finita) de $\mathbb{P}$ em uniões disjuntas,

$$\mathbb{P}(B) = \sum_{\omega \in B} \mathbb{P}(\{\omega\}).$$

Defina

$$g(\omega) := \mathbb{P}(\{\omega\}), \qquad \omega \in \Omega.$$

Exercício 7 — Resolução (densidade em relação a c)

A integral em relação à medida de contagem coincide com soma: para toda $h : \Omega \to \mathbb{R}_+$,

$$\int_\Omega h(\omega)\, c(\mathrm{d}\omega) = \sum_{\omega \in \Omega} h(\omega).$$

Exercício 7 — Resolução (densidade em relação a c)

A integral em relação à medida de contagem coincide com soma: para toda $h : \Omega \to \mathbb{R}_+$,

$$\int_\Omega h(\omega)\, c(\mathrm{d}\omega) = \sum_{\omega \in \Omega} h(\omega).$$

Logo, para todo $B \subseteq \Omega$,

$$\begin{aligned} \int_\Omega \mathbf{1}_B(\omega)\, g(\omega)\, c(\mathrm{d}\omega) &= \sum_{\omega \in \Omega} \mathbf{1}_B(\omega)\, g(\omega) = \sum_{\omega \in B} g(\omega) \\ &= \sum_{\omega \in B} \mathbb{P}(\{\omega\}) = \mathbb{P}(B). \end{aligned}$$

Portanto, $\mathbb{P}$ admite densidade com respeito a c, e uma densidade é exatamente $g(\omega) = \mathbb{P}(\{\omega\})$.

Exercício 7 — Resolução (densidade em relação a c)

A integral em relação à medida de contagem coincide com soma: para toda $h : \Omega \to \mathbb{R}_+$,

$$\int_\Omega h(\omega)\, c(\mathrm{d}\omega) = \sum_{\omega\in\Omega} h(\omega).$$

Logo, para todo $B \subseteq \Omega$,

$$\begin{aligned}\int_\Omega \mathbf{1}_B(\omega)\, g(\omega)\, c(\mathrm{d}\omega) &= \sum_{\omega\in\Omega} \mathbf{1}_B(\omega)\, g(\omega) = \sum_{\omega\in B} g(\omega)\\ &= \sum_{\omega\in B} \mathbb{P}(\{\omega\}) = \mathbb{P}(B).\end{aligned}$$

Portanto, $\mathbb{P}$ admite densidade com respeito a c, e uma densidade é exatamente $g(\omega) = \mathbb{P}(\{\omega\})$.

Como verificação adicional,

$$\int_\Omega g(\omega)\, c(\mathrm{d}\omega) = \sum_{\omega\in\Omega} \mathbb{P}(\{\omega\}) = \mathbb{P}(\Omega) = 1,$$

como deve ocorrer para uma densidade de uma medida de probabilidade.

Exercício 8 — Enunciado

Enunciado

Se $(\Omega, \Sigma, \mathbb{P})$ é um espaço de probabilidade e X é uma variável aleatória real cuja imagem está contida num intervalo $[a, b] \subset \mathbb{R}$. Suponha que a função de distribuição de X, F_X, é contínua em $[a, b]$ e diferenciável em (a, b). Mostre que a função $g : \mathbb{R} \to \mathbb{R}$, dada por

$$g(x) = \begin{cases} F_X'(x), & \text{se } x \in (a, b), \\ 0, & \text{se } x \notin (a, b), \end{cases}$$

define uma densidade de $\mathbb{P}_X$ com respeito à medida de Lebesgue em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Dica: use o Teorema Fundamental do Cálculo e o lema do π-sistema.

Exercício 8 — Resolução (definição de μ e propriedades de g)

Defina uma medida μ em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ por

$$\mu(B) \coloneqq \int_{\mathbb{R}} \mathbf{1}_B(x)\, g(x)\, \lambda(\mathrm{d}x), \qquad B \in \mathcal{B}(\mathbb{R}),$$

onde

$$g(x) = \begin{cases} F_X'(x), & x \in (a,b), \\ 0, & x \notin (a,b). \end{cases}$$

Exercício 8 — Resolução (definição de μ e propriedades de g)

Defina uma medida μ em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ por

$$\mu(B) \coloneqq \int_{\mathbb{R}} \mathbf{1}_B(x)\, g(x)\, \lambda(\mathrm{d}x), \qquad B \in \mathcal{B}(\mathbb{R}),$$

onde

$$g(x) = \begin{cases} F_X'(x), & x \in (a,b), \\ 0, & x \notin (a,b). \end{cases}$$

Necessito:
Provar que μ coincide com $\mathbb{P}_X$!

Exercício 8 — Resolução (definição de μ e propriedades de g)

(i) Mensurabilidade. Como $X(\Omega) \subseteq [a, b]$, temos $F_X(t) = 0$ se $t < a$ e $F_X(t) = 1$ se $t \geqslant b$; com a continuidade em a, b, F_X é contínua em $\mathbb{R}$. Para cada $n \geqslant 1$, a função

$$x \longmapsto n\Big(F_X\Big(x + \tfrac{1}{n}\Big) - F_X(x)\Big)$$

é contínua (logo Borel-mensurável). Em (a, b),

$$F_X'(x) = \lim_{n\to\infty} n\Big(F_X\Big(x + \tfrac{1}{n}\Big) - F_X(x)\Big),$$

portanto F_X' é mensurável em (a, b) como limite pontual, e estendendo por 0 fora de (a, b), g é mensurável em $\mathbb{R}$.

Exercício 8 — Resolução (definição de μ e propriedades de g)

(i) Mensurabilidade. Como $X(\Omega) \subseteq [a,b]$, temos $F_X(t) = 0$ se $t < a$ e $F_X(t) = 1$ se $t \geqslant b$; com a continuidade em a, b, F_X é contínua em $\mathbb{R}$. Para cada $n \geqslant 1$, a função

$$x \longmapsto n\Big(F_X\Big(x + \tfrac{1}{n}\Big) - F_X(x)\Big)$$

é contínua (logo Borel-mensurável). Em (a,b),

$$F_X'(x) = \lim_{n\to\infty} n\Big(F_X\Big(x + \tfrac{1}{n}\Big) - F_X(x)\Big),$$

portanto F_X' é mensurável em (a,b) como limite pontual, e estendendo por 0 fora de (a,b), g é mensurável em $\mathbb{R}$.

(ii) Não-negatividade. Como F_X é não-decrescente, para $h > 0$,

$$\frac{F_X(x+h) - F_X(x)}{h} \geqslant 0 \quad \Rightarrow \quad F_X'(x) \geqslant 0 \text{ em } (a,b),$$

logo $g \geqslant 0$.

Exercício 8 — Resolução (μ é medida de probabilidade)

Como $X(\Omega) \subseteq [a, b]$, vale

$$F_X(t) = 0 \text{ se } t < a, \qquad F_X(t) = 1 \text{ se } t \geqslant b.$$

Pela continuidade de F_X em $[a, b]$, em particular em a e b,

$$F_X(a) = 0, \qquad F_X(b) = 1.$$

Exercício 8 — Resolução (μ é medida de probabilidade)

Como $X(\Omega) \subseteq [a,b]$, vale

$$F_X(t) = 0 \text{ se } t < a, \qquad F_X(t) = 1 \text{ se } t \geqslant b.$$

Pela continuidade de F_X em $[a,b]$, em particular em a e b,

$$F_X(a) = 0, \qquad F_X(b) = 1.$$

Pelo Teorema Fundamental do Cálculo (em $[a,b]$),

$$\int_a^b F_X'(x)\,\lambda(\mathrm{d}x) = F_X(b) - F_X(a) = 1.$$

Como $g = F_X'$ em (a,b) e $g = 0$ fora de (a,b), obtemos

$$\int_{\mathbb{R}} g(x)\,\lambda(\mathrm{d}x) = 1,$$

isto é, $\mu(\mathbb{R}) = 1$. Logo, μ é uma medida de probabilidade.

Exercício 8 — Resolução (coincidência em um π-sistema)

Considere o π-sistema

$$\mathcal{P} := \{(-\infty, t] : t \in \mathbb{R}\}.$$

Temos $\mathcal{P}$ é um π-sistema pois

$$(-\infty, t_1] \cap (-\infty, t_2] = (-\infty, \min\{t_1, t_2\}],$$

e $\sigma(\mathcal{P}) = \mathcal{B}(\mathbb{R})$.

Exercício 8 — Resolução (coincidência em um π-sistema)

Considere o π-sistema

$$\mathcal{P} \coloneqq \{(-\infty, t] : t \in \mathbb{R}\}.$$

Temos $\mathcal{P}$ é um π-sistema pois

$$(-\infty, t_1] \cap (-\infty, t_2] = (-\infty, \min\{t_1, t_2\}],$$

e $\sigma(\mathcal{P}) = \mathcal{B}(\mathbb{R})$.
Mostremos que μ e $\mathbb{P}_X$ coincidem em $\mathcal{P}$. Fixe $t \in \mathbb{R}$. Então

$$\mu((-\infty, t]) = \int_{(-\infty, t]} g(x)\, \lambda(\mathrm{d}x).$$

Exercício 8 — Resolução (coincidência em um π-sistema)

Analisando casos:

Exercício 8 — Resolução (coincidência em um π-sistema)

Analisando casos:

1) Se $t \leqslant a$**:** $g = 0$ em $(-\infty, t]$, logo $\mu((-\infty, t]) = 0$. Além disso, $F_X(t) = \mathbb{P}(X \leqslant t) = 0$, então

$$\mu((-\infty, t]) = 0 = F_X(t) = \mathbb{P}_X((-\infty, t]).$$

Exercício 8 — Resolução (coincidência em um π-sistema)

Analisando casos:

1) Se $t \leqslant a$: $g = 0$ em $(-\infty, t]$, logo $\mu((-\infty, t]) = 0$. Além disso, $F_X(t) = \mathbb{P}(X \leqslant t) = 0$, então

$$\mu((-\infty, t]) = 0 = F_X(t) = \mathbb{P}_X((-\infty, t]).$$

2) Se $a < t < b$:

$$\mu((-\infty, t]) = \int_a^t F'_X(x)\,\lambda(\mathrm{d}x) = F_X(t) - F_X(a) = F_X(t) = \mathbb{P}_X((-\infty, t]).$$

Exercício 8 — Resolução (coincidência em um π-sistema)

Analisando casos:

1) Se $t \leqslant a$: $g = 0$ em $(-\infty, t]$, logo $\mu((-\infty, t]) = 0$. Além disso, $F_X(t) = \mathbb{P}(X \leqslant t) = 0$, então

$$\mu((-\infty, t]) = 0 = F_X(t) = \mathbb{P}_X((-\infty, t]).$$

2) Se $a < t < b$:

$$\mu((-\infty, t]) = \int_a^t F_X'(x)\,\lambda(\mathrm{d}x) = F_X(t) - F_X(a) = F_X(t) = \mathbb{P}_X((-\infty, t]).$$

3) Se $t \geqslant b$:

$$\mu((-\infty, t]) = \int_a^b F_X'(x)\,\lambda(\mathrm{d}x) = 1 = F_X(t) = \mathbb{P}_X((-\infty, t]).$$

Exercício 8 — Resolução (coincidência em um π-sistema)

Analisando casos:

1) Se $t \leqslant a$**:** $g = 0$ em $(-\infty, t]$, logo $\mu((-\infty, t]) = 0$. Além disso, $F_X(t) = \mathbb{P}(X \leqslant t) = 0$, então

$$\mu((-\infty, t]) = 0 = F_X(t) = \mathbb{P}_X((-\infty, t]).$$

2) Se $a < t < b$**:**

$$\mu((-\infty, t]) = \int_a^t F'_X(x)\,\lambda(\mathrm{d}x) = F_X(t) - F_X(a) = F_X(t) = \mathbb{P}_X((-\infty, t]).$$

3) Se $t \geqslant b$**:**

$$\mu((-\infty, t]) = \int_a^b F'_X(x)\,\lambda(\mathrm{d}x) = 1 = F_X(t) = \mathbb{P}_X((-\infty, t]).$$

Logo, μ e $\mathbb{P}_X$ coincidem em $\mathcal{P}$.

Exercício 8 — Conclusão (lema do π-sistema)

Como μ e $\mathbb{P}_X$ são medidas de probabilidade em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ e coincidem no π-sistema $\mathcal{P}$ que gera $\mathcal{B}(\mathbb{R})$, pelo lema do π-sistema, segue que

$$\mu(B) = \mathbb{P}_X(B) \qquad \forall\, B \in \mathcal{B}(\mathbb{R}).$$

Exercício 8 — Conclusão (lema do π-sistema)

Como μ e $\mathbb{P}_X$ são medidas de probabilidade em $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ e coincidem no π-sistema $\mathcal{P}$ que gera $\mathcal{B}(\mathbb{R})$, pelo lema do π-sistema, segue que

$$\mu(B) = \mathbb{P}_X(B) \qquad \forall\, B \in \mathcal{B}(\mathbb{R}).$$

Pela definição de μ,

$$\mathbb{P}_X(B) = \int_{\mathbb{R}} \mathbf{1}_B(x)\, g(x)\, \lambda(\mathrm{d}x), \qquad \forall\, B \in \mathcal{B}(\mathbb{R}).$$

Portanto, $\mathbb{P}_X$ admite densidade em relação à medida de Lebesgue, e essa densidade é g.