

Notas de Aula - Capítulo 3

Probabilidade

Caio Gomes Alves

06/05/2025

1 Esperança

1.1 Definição

Definition 1.1. Se X é uma variável aleatória com distribuição F , a esperança de X é definida por $E(X = \int_{-\infty}^{\infty} x dF(x))$, sempre que a integral estiver bem definida.

Convenção: Se $E(X) < \infty$, então X é integrável.

Nota: $\int_{-\infty}^{\infty} x dF(x)$ é bem definida se $\int_0^{\infty} x dF(x)$ ou $\int_{-\infty}^0 x dF(x)$ for finita, já que $\int_{-\infty}^{\infty} x dF(x) = \underbrace{\int_{-\infty}^0 x dF(x)}_{\mathbf{I} \leq 0} + \underbrace{\int_0^{\infty} x dF(x)}_{\mathbf{II} \geq 0}$. Assim, podemos separar em quatro casos:

1. Se \mathbf{I} e \mathbf{II} são finitos, então X é integrável;
2. Se \mathbf{I} é finito e $\mathbf{II} = +\infty$, então $E(X) = +\infty$;
3. Se \mathbf{II} é finito e $\mathbf{I} = -\infty$, então $E(X) = -\infty$;
4. Se $\mathbf{I} = -\infty$ e $\mathbf{II} = +\infty$, então $E(X)$ é indefinida.

Propriedade: $E(|X|) = \int |x| dF(x)$. Logo, X é integrável se e somente se $E(|X|) < \infty$.

Example 1.1. $X \sim U(0, 1)$, $Y = \min(X, \frac{1}{2})$:

$$\begin{aligned} P\left(Y = \frac{1}{2}\right) &= P\left(X > \frac{1}{2}\right) = 1 - F_X\left(\frac{1}{2}\right) = 1 - \frac{1}{2} = \frac{1}{2} = P_Y\left(Y = \frac{1}{2}\right) \\ E(Y) &= \int_{-\infty}^{\infty} y dF(y) = \int_0^{1/2} y \cdot 1 dy + \frac{1}{2} P_Y\left(Y = \frac{1}{2}\right) \\ &= \left. \frac{y^2}{2} \right|_0^{1/2} + \frac{1}{4} \\ &= \frac{1}{8} + \frac{1}{4} = \frac{3}{8} \end{aligned}$$

Proposition 1.1. $E(X) = \int_0^{\infty} (1 - F(x)) dx - \int_{-\infty}^0 F(x) dx$. Disso, temos que:

- a) $\int_0^{\infty} x dF(x) = \int_0^{\infty} (1 - F(x)) dx$;
- b) $\int_{-\infty}^0 x dF(x) = - \int_{-\infty}^0 F(x) dx$;

Prova. Vejamos **(a)**: considere que $d(xF(x)) = F(x)dx + xd(F(x)) \Rightarrow xd(F(x)) = d(xF(x)) - F(x)dx$. Seja um $b > 0$:

$$\begin{aligned}\int_0^b x dF(x) &= \int_0^b d(xF(x)) - \int_0^b F(x)dx \\ &= xF(x) \Big|_0^b - \int_0^b F(x)dx \\ &= bF(b) - \int_0^b F(x)dx \\ &= \int_0^b [F(b) - F(x)]dx\end{aligned}$$

Note que $\int_0^b x dF(x) \leq \int_0^\infty [1 - F(x)]dx$, $\forall b > 0$. Basta notar que $F(b) - F(x) \leq 1 - F(x)$ e que $\int_0^b [1 - F(x)]dx \leq \int_0^\infty [1 - F(x)]dx$. Logo:

$$\int_0^\infty x dF(x) = \lim_{b \rightarrow \infty} \int_0^b x dF(x) \leq \int_0^\infty [1 - F(x)]dx \Rightarrow \int_0^\infty x dF(x) \leq \int_0^\infty [1 - F(x)]dx$$

Considere $\lambda > 0$ e $b > 0$, tais que:

$$\begin{aligned}\int_0^b [F(b) - F(x)]dx &\geq \int_0^\lambda [1 - F(x)]dx = \int_0^\lambda [F(b) - 1]dx + \int_0^\lambda [1 - F(x)]dx \\ &= \lambda[F(b) - 1] + \int_0^\lambda [1 - F(x)]dx \\ \int_0^b [F(b) - F(x)]dx &\geq \lambda[F(b) - 1] + \int_0^\lambda [1 - F(x)]dx\end{aligned}$$

Logo, como $\int_0^\infty x dF(x) = \lim_{b \rightarrow \infty} \int_0^b [F(b) - F(x)]dx \geq \lim_{b \rightarrow \infty} \{\lambda[F(b) - 1] + \int_0^\lambda [1 - F(x)]dx\} = \int_0^\lambda [1 - F(x)]dx$. Assim:

$$\int_0^\infty x dF(x) \geq \lim_{\lambda \rightarrow \infty} \int_0^\lambda [1 - F(x)]dx = \int_0^\infty [1 - F(x)]dx$$

E como $\int_0^\infty x dF(x) \leq \int_0^\infty [1 - F(x)]dx$, temos que $\int_0^\infty x dF(x) = \int_0^\infty [1 - F(x)]dx$ □

Corollary 1.1. Se X é tal que $X(\omega) \geq 0 \forall \omega \in \mathbb{R} \Rightarrow E(X) = \int_0^\infty [1 - F(x)]dx = \int_0^\infty P(X \geq x)dx$.

Example 1.2. Seja $X \sim \text{Exp}(\lambda)$, qual a $E(X)$? Como o suporte de X é $(0, \infty)$, aplica-se o corolário anterior, de modo que:

$$\begin{aligned}F_X(x) &= 1 - e^{-\lambda x} \Leftrightarrow P(X > x) = e^{-\lambda x} \\ E(X) &= \int_0^\infty e^{-\lambda x} dx = -\frac{1}{\lambda} e^{-\lambda x} \Big|_0^\infty = \frac{1}{\lambda}\end{aligned}$$

Nota: Suponha X discreta e $X(\omega) \geq 0 \forall \omega$. Então:

$$\begin{aligned}
E(X) &= \sum_{n=0}^{\infty} P(X > n) = \sum_{n=0}^{\infty} P(X \geq n+1) \\
&= \sum_{n=1}^{\infty} P(X \geq n)
\end{aligned}$$

Example 1.3. Considere o lançamento de uma moeda até a 1ª cara. Suponha p = probabilidade de cara e $(1-p)$ = probabilidade de coroa, e X = número de lançamentos até a primeira cara. Tome o evento $[X \geq n]$, logo:

$$E(X) = \sum_{n=1}^{\infty} (1-p)^{n-1} = \sum_{n=0}^{\infty} (1-p)^n = \frac{1}{p}$$

Nota: Sendo X uma variável aleatória, temos pelo corolário 1.1 que:

$$\begin{aligned}
E(|X|) &= \int_0^{\infty} P(|X| > x) dx \\
&= \int_0^{\infty} [P(X > x) + P(X < -x)] dx \\
&= \int_0^{\infty} P(X > x) dx + \int_0^{\infty} P(X < -x) dx \\
&= \int_0^{\infty} (1 - F(x)) dx + \int_0^{\infty} F((-x)^-) dx
\end{aligned}$$

Onde $F((-x)^-) = \lim_{u \uparrow -x} F(u)$, que caso F seja contínua, coincide com $F(-x)$. Logo:

$$E(|X|) = \int_0^{\infty} (1 - F(x)) dx + \int_0^{\infty} F(-x) dx$$

Já que F pode ser descontínua em uma coleção enumerável de pontos. Agora, tomando a transformação de variável $y = -x \Leftrightarrow dy = -dx$:

$$\begin{aligned}
E(|X|) &= \int_0^{\infty} (1 - F(x)) dx + \int_{-\infty}^0 F(y) dy \\
&= \int_0^{\infty} (1 - F(x)) dx + \int_{-\infty}^0 F(x) dx
\end{aligned}$$

Utilizando os resultados **a** e **b** da proposição 1.1, temos que:

$$\begin{aligned}
E(|X|) &= \int_0^{\infty} x dF(x) - \int_{-\infty}^0 x dF(x) \\
&= \int_0^{\infty} |x| dF(x) + \int_{-\infty}^0 |x| dF(x) \\
&= \int_{-\infty}^{\infty} |x| dF(x)
\end{aligned}$$

Onde F é a acumulada de X , ao invés de $|X|$. Assim, a integrabilidade de X depende da finitude de $\int_0^{\infty} x dF(x)$ e $\int_{-\infty}^0 x dF(x)$, logo X é integrável se $E(|X|) < \infty$.

1.2 Propriedades da esperança

- **E₁**: Se $X = c$, com c uma constante, $E(X) = c$;
- **E₂(monotonia)**: Se X e Y são variáveis aleatórias, com $X \leq Y \Rightarrow E(X) \leq E(Y)$, caso ambas as esperanças estejam bem definidas;

Prova. Seja z um valor fixo. Se $Y \leq z \Rightarrow X \leq z$, logo $[Y \leq z] \subseteq [X \leq z]$, assim:

$$\begin{aligned} P(Y \leq z) &\leq P(X \leq z) \\ F_Y(z) &\leq F_X(z) \iff 1 - F_Y(z) \geq 1 - F_X(z) \end{aligned}$$

E pela proposição 1.1, temos que:

$$\begin{aligned} E(Y) &= \int_0^\infty [1 - F_Y(z)] dz - \int_{-\infty}^0 F_Y(z) dz \geq \int_0^\infty [1 - F_X(z)] dz - \int_{-\infty}^0 F_X(z) dz = E(X) \\ E(Y) &\geq E(X) \end{aligned}$$

□

- **E₃(linearidade)**:
 - (i) Se $E(X)$ é bem definida, $a, b \in \mathbb{R}$, então $E(aX + b) = aE(X) + b$;
 - (ii) $E(aX + bY) = aE(X) + bE(Y)$, caso o termo $aE(X) + bE(Y)$ esteja bem definido;
 - Note que se $E(X) = \infty \Rightarrow E(X - X) \neq E(X) - E(X)$.

Prova. Quando $a = 0$; $E(aX + b) = E(b) = b = 0E(X) + b$.

Quando $a > 0, b > 0$; $F_{aX+b}(x) = P(aX + b \leq x) = P(X \leq \frac{x-b}{a}) = F_X\left(\frac{x-b}{a}\right)$. Logo:

$$\begin{aligned} E(aX + b) &= \int_0^\infty [1 - F_{aX+b}(x)] dx - \int_{-\infty}^0 F_{aX+b}(x) dx \\ &= \int_0^\infty \left[1 - F_X\left(\frac{x-b}{a}\right)\right] dx - \int_{-\infty}^0 F_X\left(\frac{x-b}{a}\right) dx \end{aligned}$$

Tome $y = \frac{x-b}{a} \Rightarrow dy = \frac{1}{a} dx$. Então:

$$\begin{aligned} E(aX + b) &= \int_{-b/a}^\infty a[1 - F_X(y)] dy - \int_{-\infty}^{-b/a} aF_X(y) dy \\ &= a \left\{ \int_{-b/a}^\infty [1 - F_X(y)] dy - \int_{-\infty}^{-b/a} F_X(y) dy \right\} \\ &= a \int_0^\infty [1 - F_X(y)] dy - a \int_{-\infty}^0 F_X(y) dy + a \int_{-b/a}^0 [1 - F_X(y)] dy + a \int_{-b/a}^0 F_X(y) dy \\ &= aE(X) + a \int_{-b/a}^0 dy \\ &= aE(X) + a \frac{b}{a} \\ &= aE(X) + b \end{aligned}$$

□

- **E₄(Desigualdade de Jansen):** Seja φ uma função convexa, definida na reta, com X integrável, então:

$$E(\varphi(X)) \geq \varphi(E(X)) \quad (1)$$

Nota: Caso φ seja côncava:

$$E(\varphi(X)) \leq \varphi(E(X))$$

Prova para convexa. Tome x_0 e $\varphi(x_0)$. Então existe uma reta L tal que L passe por $\varphi(x_0)$ e φ fica por cima de L . Logo temos a seguinte equação da reta:

$$L(x) = \varphi(x_0) + \lambda(x - x_0)$$

Onde λ é alguma constante apropriada. Então para todo x temos:

$$\begin{aligned} \varphi(x) &\geq L(x) = \varphi(x_0) + \lambda(x - x_0) \\ &\Downarrow \mathbf{E}_2 \\ E(\varphi(x)) &\geq E(L(x)) \stackrel{\mathbf{E}_1, \mathbf{E}_3}{=} \varphi(x_0) + \lambda[E(x) - x_0] \end{aligned}$$

Que vale para $x_0 = E(x)$, de modo que $E(\varphi(x)) \geq \varphi(E(x)) + \lambda[E(x) - E(x)]$, então:

$$E(\varphi(x)) \geq \varphi(E(x))$$

A prova para funções côncavas segue a mesma metodologia, com a inversão da desigualdade. □

1.2.1 Critério de integrabilidade

Suponha que X é uma variável aleatória dominada por Y (ou seja, $X \leq Y$), sendo Y uma variável aleatória integrável. X é integrável? Temos que:

$$X \leq Y \Rightarrow E(X) \leq E(Y)$$

Se X e Y são tais que $Y \geq 0$ e Y é integrável e $|X| \leq Y \Rightarrow 0 \leq |X| \leq Y$, e como consequência:

$$0 \leq E(X) \leq E(Y) < \infty \implies X \text{ é integrável}$$

De maneira similar, seja X uma variável aleatória qualquer. Então:

$$\sum_{n=1}^{\infty} P(|X| \geq n) \leq E(|X|) \leq 1 + \sum_{n=1}^{\infty} P(|X| \geq n)$$

Assim, X é integrável se e somente se $\sum_{n=1}^{\infty} P(|X| \geq n) < \infty$.

Prova. Seja $x \geq 0$. Tome $[x]$ como a parte inteira de x . Então $[x] = k$ se $k \leq |x| < k + 1$. Então:

$$\begin{aligned} 0 \leq [x] \leq |x| \leq [x] + 1 \\ \Downarrow \mathbf{E_2, E_3} \\ 0 \leq E([x]) \leq E(|x|) \leq E([x]) + 1 \end{aligned}$$

Pelo corolário 1.1, como $[x]$ é discreta e não-negativa, temos que:

$$\begin{aligned} E([x]) &= \sum_{n=1}^{\infty} P([x] \geq n) \\ &= \sum_{n=1}^{\infty} P(|x| \geq n) \leq E(|x|) \leq \sum_{n=1}^{\infty} P(|x| \geq n) + 1 \end{aligned}$$

□

1.2.2 Casos de interesse

a) (Consistência absoluta) $\varphi(X) = |X|$:

$$E(|X|) \geq |E(X)|$$

b) (Consistência quadrática) $\varphi(X) = X^2$:

$$E(X^2) \geq [E(X)]^2$$

c) (Consistência absoluta de ordem p) $\varphi(X) = |X|^p, p \geq 1$:

$$E(|X|^p) \geq |E(X)|^p$$

Nota: φ só precisa ser convexa (ou côncava) em uma região de probabilidade 1. Por exemplo, se X é uma variável aleatória, tal que $P(X > 0) = 1$, ou o suporte da distribuição de X é $(0, \infty)$, $\varphi(X) = \frac{1}{X}$ é convexa em $(0, \infty) \Rightarrow E\left(\frac{1}{X}\right) \geq \frac{1}{E(X)}$. De modo análogo, se $P(X > 0) = 1$ e $\varphi(X) = \ln(X)$, φ é côncava em $(0, \infty)$ logo $E(\ln(X)) \leq \ln(E(X))$.

1.3 Esperança de funções de variáveis aleatórias

Seja X uma variável aleatória, φ uma função mensurável e $Y = \varphi(X)$. Assim, Y é uma variável aleatória, cuja esperança é $E(Y) = \int y dF_{\varphi(X)}(y) = \int_0^{\infty} [1 - F_{\varphi(X)}(y)] dy - \int_{-\infty}^0 F_{\varphi(X)}(y) dy$.

Theorem 1.1. Se X é uma variável aleatória e φ uma função mensurável, com $Y = \varphi(X)$:

$$E(Y) = E(\varphi(X)) = \int \varphi(x) dF_X(x)$$

Prova para caso $\varphi(x) = x^k$. Note que a prova já foi feita para $\varphi(x) = |x|$. Vejamos que a prova é válida para $\varphi(x) = x^k$, com $k = 1, 2, \dots$, em 2 casos: k par e k ímpar:

k par:

$$\begin{aligned} E(X^k) &= \int_0^\infty P(X^k > t) dt \\ &= \int_0^\infty P(X > \sqrt[k]{t}) dt + \int_0^\infty P(X < -\sqrt[k]{t}) dt \\ &= \int_0^\infty [1 - F_X(\sqrt[k]{t})] dt + \int_0^\infty F_X(-\sqrt[k]{t}) dt \end{aligned}$$

Apliquemos as seguintes mudanças de variáveis: $s = t^{\frac{1}{k}}, ds = \frac{1}{k} t^{\frac{1}{k}-1} dt, dt = \frac{(ds)ks^k}{s}, u = -s, du = -ds$:

$$\begin{aligned} E(X^k) &= \int_0^\infty [1 - F_X(s)] ks^{k-1} ds + \int_0^\infty F_X(-s) ks^{k-1} ds \\ &= \int_0^\infty [1 - F_X(s)] ks^{k-1} ds - \int_{-\infty}^0 F_X(u) ku^{k-1} du \\ &= k \left\{ \int_0^\infty [1 - F_X(s)] s^{k-1} ds - \int_{-\infty}^0 F_X(u) u^{k-1} du \right\} \end{aligned}$$

Agora, mostremos que $E(X^k) = \int x^k dF_X(x)$:

$$\begin{aligned} \int_{-\infty}^\infty x^k dF_X(x) &\stackrel{Def}{=} \int_0^\infty [1 - F_X(x)] d(x^k) - \int_{-\infty}^0 F_X(x) d(x^k) \\ &= k \left\{ \int_0^\infty [1 - F_X(x)] x^{k-1} dx - \int_{-\infty}^0 F_X(x) x^{k-1} dx \right\} \\ &= E(X^k) \end{aligned}$$

□

Nota: A propriedade é também válida para polinômios, visto que a esperança opera de maneira linear.

Example 1.4. Seja $X \sim Exp(\lambda)$, vimos que $E(X) = \frac{1}{\lambda}$:

Calcular $E(X^2)$:

$$\begin{aligned} E(X^2) &= 2 \int_0^\infty x e^{-\lambda x} dx = \frac{2}{\lambda} \int_0^\infty x \lambda e^{-\lambda x} dx \\ &= \frac{2}{\lambda} E(X) = \frac{2}{\lambda^2} \end{aligned}$$

Calcular $E(X^3)$:

$$\begin{aligned} E(X^3) &= 3 \int_0^\infty x^2 e^{-\lambda x} dx = \frac{3}{\lambda} \int_0^\infty x^2 \lambda e^{-\lambda x} dx \\ &= \frac{3}{\lambda} E(X^2) = \frac{3}{\lambda^3} \end{aligned}$$

De modo que podemos observar o padrão emergente, e definir $E(X^k) = \frac{k!}{\lambda^k}$.

1.4 Momentos de uma variável aleatória

- a) $E([X - b]^k)$: k -ésimo momento de X em torno de b ;
- b) $E(X^k)$: k -ésimo momento em torno de 0;
- c) Se em (a), $b = E(X)$, o momento é central;
- d) $t > 0, E(|X|^t)$: t -ésimo momento absoluto de X .

Definition 1.2 (Variância de uma variável aleatória).

$$\text{Var}(X) = E\{(X - E(X))^2\} \iff \text{Var}(X) = E(X^2) - (E(X))^2$$

Proposition 1.2. Se X é uma variável aleatória, $f(t) = [E(|X|^t)]^{\frac{1}{t}}$ é não-decrescente em $t, t > 0$.

Prova. Devemos provar que, se $0 < s < t, f(s) \leq f(t)$ (ou $\{E(|X|^s)\}^{\frac{1}{s}} \leq \{E(|X|^t)\}^{\frac{1}{t}}$). Para tanto, consideremos dois casos: **a)** $E(|X|^s) < \infty$, **b)** $E(|X|^s) = \infty$:

a) Defina $\varphi(y) = |y|^{\frac{t}{s}}$ (caso $\frac{t}{s} > 1, \varphi$ será convexa). Pela Desigualdade de Jansen:

$$\begin{aligned} E(\varphi(Y)) &\geq \varphi(E(Y)) \\ E\left(|Y|^{\frac{t}{s}}\right) &\geq |E(Y)|^{\frac{t}{s}} \end{aligned}$$

Tome $Y = |X|^s$. Substituindo temos:

$$\begin{aligned} E\left(|X|^s\right)^{\frac{t}{s}} &\geq |E(|X|^s)|^{\frac{t}{s}} \\ E(|X|^t) &\geq \{E(|X|^s)\}^{\frac{t}{s}} \\ \{E(|X|^t)\}^{\frac{1}{t}} &\geq \{E(|X|^s)\}^{\frac{1}{s}} \end{aligned}$$

b) Como $t > s > 0$, sabemos que $|X|^s \leq 1 + |X|^t$. Como $E(|X|^s) = \infty$, então:

$$\infty = E(|X|^s) \leq 1 + E(|X|^t) = \infty$$

□

Corollary 1.2. Se $E(|X|^t) < \infty \forall t \in (0, \infty) \Rightarrow E(|X|^s) < \infty \forall s$, com $0 < s < t$.

1.4.1 Propriedades

- **E₅**: Se $X = c$, com c uma constante, $\text{Var}(X) = 0$;
- **E₆**: $\text{Var}(X + b) = \text{Var}(X), \text{Var}(aX + b) = a^2\text{Var}(X)$, com $a, b \in \mathbb{R}$;

Prova.

$$\begin{aligned} \text{Var}(aX + b) &= E\left\{[aX + b - E(aX + b)]^2\right\} \\ &= E\left\{[aX + b - aE(X) - b]^2\right\} \\ &= E\left\{a^2[X - E(X)]^2\right\} \\ &= a^2E\left\{[X - E(X)]^2\right\} = a^2\text{Var}(X) \end{aligned}$$

□

- **E₇(Desigualdade de Tchebychev):** Seja X uma variável aleatória, com $X \geq 0$. Para todo $\lambda > 0$:

$$P(X \geq \lambda) \leq \frac{E(X)}{\lambda} \quad (2)$$

Prova. Seja $Y = I_{[X \geq \lambda]} \lambda = \begin{cases} \lambda & , X \geq \lambda \\ 0 & , c.c. \end{cases}$. Por definição, $0 \leq Y \leq X \Rightarrow E(Y) \leq E(X)$ e $E(Y) = \lambda P(X \geq \lambda)$, de modo que:

$$\lambda P(X \geq \lambda) \leq E(X) \Leftrightarrow P(X \geq \lambda) \leq \frac{E(X)}{\lambda}$$

□

1.4.2 Consequências

a) Para todo $\lambda > 0$:

$$P(|X - E(X)| \geq \lambda) \leq \frac{\text{Var}(X)}{\lambda^2}$$

b) **(Desigualdade de Markov)** Seja X uma variável aleatória, para todo t :

$$P(|X| \geq \lambda) \leq \frac{E(|X|^t)}{\lambda^t} \quad (3)$$

c) Se Z é uma variável aleatória, com $Z \geq 0$ e $E(Z) = 0$:

$$P(Z = 0) = 1 \quad (\text{i.e., } Z = 0 \text{ quase certamente})$$

Provas. a) Se $Y = [X - E(X)]^2$, aplicamos **E₇** usando $\lambda^2 : P(Y \geq \lambda^2) \leq \frac{E(Y)}{\lambda^2}$. Note que $E(Y) = E([X - E(X)]^2) = \text{Var}(X)$. Logo:

$$P(|X - E(X)| \geq \lambda) = P(|X - E(X)|^2 \geq \lambda^2) = P(Y \geq \lambda^2) \leq \frac{E(Y)}{\lambda^2} = \frac{\text{Var}(X)}{\lambda^2}$$

b) Seja $Y = |X|^t$, aplicamos **E₇** a Y e $\lambda^t : P(Y \geq \lambda^t) \leq \frac{E(Y)}{\lambda^t}$. Note que $E(Y) = E(|X|^t)$ e que $P(Y \geq \lambda^t) = P(|X|^t \geq \lambda^t) = P(|X| \geq \lambda)$. Logo:

$$P(|X| \geq \lambda) \leq \frac{E(|X|^t)}{\lambda^t}$$

c) $Z = 0$ quase certamente, usamos **E₇** na variável Z e em $\lambda = \frac{1}{n}$, então:

$$P\left(Z \geq \frac{1}{n}\right) \leq E(Z) \cdot n \stackrel{Hip}{=} 0$$

Temos que $[Z > 0] = \bigcup_n [Z \geq \frac{1}{n}]$, de modo que:

$$P(Z > 0) = P\left(\bigcup_n \left[Z \geq \frac{1}{n}\right]\right) = \lim_{n \rightarrow \infty} P\left(Z \geq \frac{1}{n}\right) = 0 \Rightarrow P(Z = 0) = 1 - P(Z > 0) = 1$$

□

Nota: Se X é uma variável tal que $\text{Var}(X) = 0$, temos que $\text{Var}(X) = 0 \Leftrightarrow E([X - E(X)]^2) = 0$, ou seja, se definirmos $Z = [X - E(X)]^2, Z \geq 0$ e $E(Z) = 0$. Logo, por **c)**, $P(Z = 0) = 1$, ou seja, $P([X - E(X)]^2 = 0) = 1 \Leftrightarrow P(X = E(X)) = 1$, ou seja, $X = E(X)$ quase certamente.

Example 1.5. Se X e Y são variáveis aleatórias tais que $E(|X|^t) < \infty$ e $E(|Y|^t) < \infty$, então $E(|X + Y|^t) < \infty$:

- (i) A finitude de $E(|X|^t)$ leva à finitude de $E(|aX|^t)$;
- (ii) Se X e Y forem integráveis (com $t = 1$), então $X + Y$ é integrável. Se X e Y tem variâncias finitas ($t = 2$), então $X + Y$ tem variância finita.

Proposition 1.3. *Seja X uma variável aleatória integrável e $\mu = E(X) \Rightarrow \mu$ minimiza $E([X - c]^2)$, com $c \in \mathbb{R}$.*

Prova. Temos que $(X - c)^2 = (X - \mu + \mu - c)^2 = (X - \mu)^2 + 2(\mu - c)(X - \mu) + (\mu - c)^2$. Logo, pelas propriedades lineares do valor esperado:

$$\begin{aligned} E([X - c]^2) &= E([X - \mu]^2) + 2(\mu - c)E(X - \mu) + (\mu - c)^2 \\ &= \text{Var}(X) + (\mu - c)^2 \end{aligned}$$

□

Proposition 1.4. *Seja X uma variável aleatória e m sua mediana. Assim, m minimiza $E(|X - c|)$, $c \in \mathbb{R}$. Ou seja:*

$$E(|x - m|) = \min_{c \in \mathbb{R}} E(|X - c|)$$

Prova. Considere a definição de mediana: $P(X \leq m) = P(X > m) = \frac{1}{2}$. Suponha que X é integrável, logo $X - c$ também o será para todo c constante real. Vamos ver que, com $m < c$ (o caso em que $m > c$ segue analogamente):

- $X \leq m \Rightarrow |X - c| - |X - m| = \lambda$, onde $\lambda = c - m$;
- $X > m \Rightarrow |X - c| - |X - m| \geq \lambda$.

Seja c tal que $m < c$. Defina $\lambda = c - m > 0$. Então:

- Se $x \leq m \Rightarrow |x - c| = |x - m| + \lambda \Rightarrow |x - c| - |x - m| = \lambda$;
- Se $x > m$ e $x > c$ (os casos intermediários são decorrências), então $x > c \Rightarrow \lambda + |x - c| = |x - m| \Rightarrow |x - c| - |x - m| \geq -\lambda$.

Defina $Y = |X - c| - |X - m| = \begin{cases} \lambda & , \text{se } X \leq m, \\ y & , \text{se } X > m. \end{cases}$. Assim, temos que $y \geq -\lambda$, e que $Y =$

$\begin{cases} \lambda & , P(X \leq m) \geq \frac{1}{2}, \\ y & , P(X > m) \leq \frac{1}{2} \end{cases}$. Logo, $Y \geq \lambda I_{[X \leq m]} - \lambda I_{[X > m]} \Rightarrow E(Y) \geq \lambda E(I_{[X \leq m]}) - \lambda E(I_{[X > m]}) = \lambda P(X \leq m) - \lambda P(X > m) \geq 0$.

Como $E(Y) \geq 0 \Rightarrow E(|X - c|) \geq E(|x - m|) \forall c$, com $m < c$.

□

1.5 Esperanças e funções de vetores

Theorem 1.2. *Seja $\underline{X} = (X_1, \dots, X_n)$ um vetor aleatório e $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ mensurável. Então:*

$$\begin{aligned} E(\varphi(\underline{X})) &= \int y dF_{\varphi(\underline{X})}(y) = \int_{\mathbb{R}^n} \varphi(\underline{x}) dF_{\underline{X}}(\underline{x}) \\ &= \int \cdots \int \varphi(x_1, \dots, x_n) dF_{\underline{X}}(x_1, \dots, x_n) \end{aligned}$$

Caso discreto: Seja \underline{X} discreto, tomando valores $\underline{X}_i = (x_{i1}, \dots, x_{in})$, com probabilidade $P(\underline{X}_i) = \sum_i P(x_i) = 1$. Então:

$$E(\varphi(\underline{X})) = \sum_i \varphi(\underline{x}_i) P(\underline{x}_i)$$

Caso contínuo: Seja \underline{X} contínuo, com densidade $f(x_1, \dots, x_n)$. Então:

$$E(\varphi(\underline{X})) = \int \cdots \int \varphi(x_1, \dots, x_n) f(x_1, \dots, x_n) dx_1 \dots dx_n$$

Example 1.6. Lembrando a propriedade $E_3 : E(X + Y) = E(X) + E(Y)$ desde que existam $E(X)$ e $E(Y)$. Seja $\varphi(x, y) = x + y$ e defina $\varphi_1(x, y) = x$ e $\varphi_2(x, y) = y$. Teremos pelo teorema que:

$$\begin{aligned} E(X + Y) &= E(\varphi(x, y)) = \int \int (x + y) dF_{X,Y}(x, y) = \int \int x dF_{X,Y}(x, y) + \int \int y dF_{X,Y}(x, y) \\ &= E(\varphi_1(x, y)) + E(\varphi_2(x, y)) \\ &= E(X) + E(Y) \end{aligned}$$

Se $\{X_i\}_{i=1}^n$ é conjuntamente independente, com densidades f_1, \dots, f_n , sendo a densidade conjunta dada por $f = \prod_{i=1}^n f_i$, então:

$$\begin{aligned} E(\varphi(\underline{X})) &= \int \cdots \int \varphi(x_1, \dots, x_n) f_1(x_1) \dots f_n(x_n) dx_1 \dots dx_n \\ &= \int \cdots \int \varphi(x_1, \dots, x_n) dF_{X_1}(x_1) \dots dF_{X_n}(x_n) \end{aligned}$$

Proposition 1.5. *Sejam $\{X_i\}_{i=1}^n$ conjuntamente independentes e integráveis. Então:*

$$E\left(\prod_{i=1}^n X_i\right) = \prod_{i=1}^n E(X_i)$$

Prova (para $n = 2$). Seja $\varphi(X, Y) = X.Y$:

$$\begin{aligned} E(X.Y) &= E(\varphi(X, Y)) = \int \int \varphi(x, y) dF_X(x) dF_Y(y) \\ &= \int [y.x dF_X(x)] dF_Y(y) \\ &= \int y E(X) dF_Y(y) \\ &= E(X) \int y dF_Y(y) = E(X)E(Y) \end{aligned}$$

□

Definition 1.3. A covariância entre X e Y será definida por:

$$\text{Cov}(X, Y) = E\left[(X - E(X))(Y - E(Y))\right]$$

Sempre que X e Y sejam integráveis. Assim:

$$\begin{aligned}\text{Cov}(X, Y) &= E\{XY - E(Y)X - E(X)Y + E(X)E(Y)\} \\ &= E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y)\end{aligned}$$

Note que X e Y podem ter $\text{Cov}(X, Y) = 0$ e mesmo assim $X \not\sim Y$.

Notas:

- A existência da covariância entre variáveis integráveis depende da existência de $E(XY)$;
- $\text{Cov}(X, Y) = 0$ é interpretado como “ X e Y são não-correlacionados”;
- Há casos onde $\text{Cov}(X, Y) = 0$ implica independência, como na Normal Bivariada, por exemplo.

Proposition 1.6. *Sejam X_1, \dots, X_n variáveis aleatórias integráveis tais que $\text{Cov}(X_i, X_j) = 0 \forall i \neq j$. Então*

$$\text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i)$$

Prova.

$$\begin{aligned}\text{Var}\left(\sum_{i=1}^n X_i\right) &= E\left\{[X_1 + \dots + X_n] - E[X_1 + \dots + X_n]\right\}^2 \\ &= E\left\{\sum_{i=1}^n [X_i - E(X_i)]^2 + 2 \sum_{i < j} (X_i - E(X_i))(X_j - E(X_j))\right\} \\ &= \sum_{i=1}^n E[(X_i - E(X_i))^2] + 2 \sum_{i < j} E[(X_i - E(X_i))(X_j - E(X_j))] \\ &= \sum_{i=1}^n \text{Var}(X_i) + 2 \sum_{i < j} \text{Cov}(X_i, X_j) \\ &= \sum_{i=1}^n \text{Var}(X_i)\end{aligned}$$

□

Corollary 1.3. *Sejam X_1, \dots, X_n variáveis aleatórias independentes e integráveis. Então:*

$$\text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i)$$

Definition 1.4. Para X e Y variáveis aleatórias, o coeficiente de correlação de Pearson é definido por:

$$\rho_{X,Y} = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}$$

Com $\sigma_X = \sqrt{\text{Var}(X)}$ e $\sigma_Y = \sqrt{\text{Var}(Y)}$, sempre que $\text{Var}(X)$ e $\text{Var}(Y)$ sejam finitas e maiores que 0.

Proposition 1.7. *Sob os supostos da definição 1.4:*

- **a)** $-1 \leq \rho_{X,Y} \leq 1$;
- **b)** $\rho_{X,Y} = 1 \Leftrightarrow P(Y = aX + b) = 1$, para algum $a > 0$ e $b \in \mathbb{R}$;
- **c)** $\rho_{X,Y} = -1 \Leftrightarrow P(Y = aX + b) = 1$, para algum $a < 0$ e $b \in \mathbb{R}$.

Prova. Note primeiramente que $\text{Cov}(X,Y) = E\{(X - E(X))(Y - E(Y))\}$, logo:

$$\rho_{X,Y} = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y} = E\left\{ \frac{(X - E(X))}{\sigma_X} \frac{(Y - E(Y))}{\sigma_Y} \right\}$$

Observe que $E\left(\frac{X - E(X)}{\sigma_X}\right) = 0$ e $\text{Var}\left(\frac{X - E(X)}{\sigma_X}\right) = 1$, e analogamente para Y . Assim:

a)

$$\begin{aligned} 0 &\leq \left(\frac{(X - E(X))}{\sigma_X} - \frac{(Y - E(Y))}{\sigma_Y} \right)^2 \\ 0 &\leq E\left\{ \left(\frac{(X - E(X))}{\sigma_X} - \frac{(Y - E(Y))}{\sigma_Y} \right)^2 \right\} \\ 0 &\leq E\left(\left[\frac{X - E(X)}{\sigma_X} \right]^2 \right) + E\left(\left[\frac{Y - E(Y)}{\sigma_Y} \right]^2 \right) - \frac{2}{\sigma_X \sigma_Y} E((X - E(X))(Y - E(Y))) \\ 0 &\leq \frac{\text{Var}(X)}{\sigma_X^2} + \frac{\text{Var}(Y)}{\sigma_Y^2} - \frac{2\text{Cov}(X,Y)}{\sigma_X \sigma_Y} \\ 0 &\leq 2 - 2\rho_{X,Y} \\ \rho_{X,Y} &\leq 1 \end{aligned}$$

Tomando a diferença ao invés da soma, chegamos que $\rho_{X,Y} \geq -1$.

b)

Suponha $\rho_{X,Y} = 1 \Leftrightarrow E\left\{ \left[\frac{X - E(X)}{\sigma_X} - \frac{Y - E(Y)}{\sigma_Y} \right]^2 \right\} = 0$. Pela propriedade E_7 , temos que:

$$P\left(\frac{X - E(X)}{\sigma_X} = \frac{Y - E(Y)}{\sigma_Y} \right) = 1$$

Ou seja, $Y \stackrel{q.c.}{=} E(Y) + \frac{\sigma_Y}{\sigma_X}(X - E(X))$, então $a = \frac{\sigma_Y}{\sigma_X} > 0$, $b = E(Y) - \frac{\sigma_Y}{\sigma_X}E(X)$. □

Nota: Se $P(Y = aX + b) = 1$, sendo $a \neq 0$, pelo desenvolvimento da prova de **(a)**, temos que:

$$\begin{aligned}
\rho_{X,Y} &= E \left\{ \left(\frac{X - E(X)}{\sigma_X} \right) \left(\frac{aX + b - aE(X) - b}{\sqrt{a^2 \sigma_X^2}} \right) \right\} \\
&= \frac{a}{|a|} E \left\{ \left[\frac{X - E(X)}{\sigma_X} \right]^2 \right\} \\
&= \frac{a}{|a|} = \text{sgn}(a) = \pm 1
\end{aligned}$$

1.6 Convergência

Theorem 1.3 (Teorema da convergência monótona). *Sejam X_1, X_2, \dots e X variáveis aleatórias em (Ω, \mathcal{A}, P) . Se $0 \leq X_n \uparrow_{n \rightarrow \infty} X$ (ou seja, $X_n(\omega) \geq 0, \forall \omega \in \Omega$ e $X_n(\omega) \uparrow_{n \rightarrow \infty} X(\omega), \forall \omega \in \Omega$). Então $E(X_n) \uparrow_{n \rightarrow \infty} E(X)$.*