Notas de Aula - Capítulo 3

Probabilidade

Caio Gomes Alves

19/05/2025

1 Esperança

1.1 Definição

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

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

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

1. Se **I** e **II** são finitos, então X é integrável;
2. Se **I** é finito e $\mathbf{II} = +\infty$, então $E(X) = +\infty$;
3. Se **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. $X \sim U(0,1)$, $Y = \min\left(X, \frac{1}{2}\right)$:

$$P\left(Y = \frac{1}{2}\right) = P\left(X > \frac{1}{2}\right) = 1 - F_X\left(\frac{1}{2}\right) = 1 - \frac{1}{2} = \frac{1}{2} = P_Y\left(Y = \frac{1}{2}\right)$$

$$\begin{aligned} E(Y) = \int_{-\infty}^{\infty} ydF(y) &= \int_{0}^{1/2} y.1dy + \frac{1}{2}P_Y\left(Y = \frac{1}{2}\right) \\ &= \left.\frac{y^2}{2}\right|_{0}^{1/2} + \frac{1}{4} \\ &= \frac{1}{8} + \frac{1}{2} = \frac{3}{8} \end{aligned}$$

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

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

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 $b > 0$:

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

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

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

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 $\int_0^\infty xdF(x) = \lim_{b\to\infty} \int_0^b [F(b) - F(x)]dx \geq \lim_{b\to\infty}\{\lambda[F(b) - 1] + \int_0^\lambda [1 - F(x)]dx\} = \int_0^\lambda [1 - F(x)]dx$. Assim:

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

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

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 $X \sim Exp(\lambda)$, 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:

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

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:

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

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

1.2 Propriedades da esperança

- $\mathbf{E_1}$: Se $X = c$, com c uma constante, $E(X) = c$;
- $\mathbf{E_2}$**(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:

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

E pela proposição 1.1, temos que:

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

$\square$

- $\mathbf{E_3}$**(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 $a > 0, b > 0; F_{aX+b}(x) = P(aX + b \leq x) = P\left(X \leq \frac{x-b}{a}\right) = F_X\left(\frac{x-b}{a}\right)$. Logo:

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

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

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

$\square$

- $\mathbf{E_4}$**(Desigualdade de Jansen)**: Seja φ uma função convexa, definida na reta, com X integrável, então:

$$E(\varphi(X)) \geq \varphi(E(X)) \tag{1}$$

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:

$$\begin{gathered}\varphi(x) \geq L(x) = \varphi(x_0) + \lambda(x - x_0) \\ \Downarrow \mathbf{E_2} \\ E(\varphi(x)) \geq E(L(x)) \overset{\mathbf{E_1},\mathbf{E_3}}{=} \varphi(x_0) + \lambda\left[E(x) - x_0\right]\end{gathered}$$

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

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

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:

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

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 $\varphi(X) = \ln(X), \varphi$ 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 $k = 1, 2, \ldots$, em 2 casos: k par e k ímpar:

k par:

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

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

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

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

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

□

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 $E(X^2)$:

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

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) $E\left([X-b]^k\right)$: $k-$ésimo momento de X em torno de b;

b) $E\left(X^k\right)$: $k-$ésimo momento em torno de 0;

c) Se em **(a)**, $b = E(X)$, o momento é central;

d) $t > 0, E\left(|X|^t\right)$: $t-$ésimo momento absoluto de X.

Definition 1.2 (Variância de uma variável aleatória)**.**

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

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:

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

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

- $\mathbf{E_5}$: Se $X = c$, com c uma constante, $\text{Var}(X) = 0$;
- $\mathbf{E_6}$: $\text{Var}(X + b) = \text{Var}(X), \text{Var}(aX + b) = a^2\text{Var}(X)$, com $a, b \in \mathbb{R}$;

Prova.

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

□

- **$\mathbf{E_7}$(Desigualdade de Tchebychev)**: Seja X uma variável aleatória, com $X \geq 0$. Para todo $\lambda > 0$:

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

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 $\lambda > 0$:

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

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

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

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 $\mathbf{E_7}$ usando $\lambda^2 : P(Y \geq \lambda^2) \leq \frac{E(Y)}{\lambda^2}$. Note que $E(Y) = E([X - E(X)]^2) = \mathrm{Var}(X)$. Logo:

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

b) Seja $Y = |X|^t$, aplicamos $\mathbf{E_7}$ 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 $P(Y \geq \lambda^t) = P(|X|^t \geq \lambda^t) = P(|X| > \lambda)$. Logo:

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

c) $Z = 0$ quase certamente, usamos $\mathbf{E_7}$ na variável Z e em $\lambda = \frac{1}{n}$, então:

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

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

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

□

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:

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

$\square$

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 $m < c$. $\square$

1.5 Esperanças e funções de vetores

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

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

Caso discreto: Seja $\underline{X}$ discreto, tomando valores $\underline{X}_i = (x_{i1}, \ldots, x_{in})$, 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 $f(x_1, \ldots, x_n)$. Então:

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

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 $f_1, \ldots, f_n$, sendo a densidade conjunta dada por $f = \prod_{i=1}^n f_i$, então:

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

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 $\varphi(X, Y) = X.Y$:

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

$\square$

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

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

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 $Cov(X,Y) = 0$ e mesmo assim $X \not\perp Y$.

Notas:

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

Proposition 1.6. *Sejam $X_1, \ldots, X_n$ 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.

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

$\square$

Corollary 1.3. *Sejam $X_1, \ldots, X_n$ 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 $\text{Cov}(X,Y) = E\left\{(X - E(X))(Y - E(Y))\right\}$, 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, $Y \overset{q.c}{=} E(Y) + \frac{\sigma_Y}{\sigma_X}(X - E(X))$, 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:

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

1.6 Convergência

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

Prova. Como $X_n \underset{n\to\infty}{\uparrow} X, X_n \geq 0$, por E_2 temos que $0 \leq E(X_n) \leq E(X)$. Devemos então provar que $\forall\epsilon > 0, \lim_{n\to\infty} E(X_n) \geq E(X) - \epsilon$ (ou seja, $\forall\epsilon > 0\ \exists\ n_0(\epsilon) : E(X_n) \geq E(X) - \epsilon, \forall n : n \geq n_0(\epsilon)$).

Defina $Y = \sum_{n=0}^{\infty} n\epsilon I_{B_n}$, onde $B_n = [x\epsilon < X \leq (n+1)\epsilon], n = 0, 1, \ldots$. Assim:

$$
\begin{aligned}
n = 0 &: B_0 = [0 < X \leq \epsilon] \\
n = 1 &: B_1 = [\epsilon < X \leq 2\epsilon] \\
n = 2 &: B_2 = [2\epsilon < X \leq 3\epsilon] \\
&\vdots
\end{aligned}
$$

$$
Y = \begin{cases} n\epsilon & \text{, se } n\epsilon < X \leq (n+1)\epsilon \\ 0 & \text{, se } X = 0 \end{cases}
$$

Temos então que mostrar que $X - \epsilon \leq Y \leq X$. Como $Y = n\epsilon < X$, o lado direito é dado diretamente. Para o lado esquerdo temos que:

$$
X \leq (n+1)\epsilon = n\epsilon + \epsilon \Leftrightarrow X - \epsilon \leq n\epsilon = Y \Rightarrow E(X) - \epsilon \leq E(Y) \leq E(X)
$$

Note que, se $E(Y) \leq \lim_{n\to\infty} E(X_n)$, então teremos provado o resultado. Seja $A_k = [X_k \geq Y]$, com k grande o suficiente. Se ω é tal que $X_k(\omega) \geq Y(\omega) \Rightarrow X_{k+1}(\omega) \geq Y(\omega) \Rightarrow A_k \subseteq A_{k+1} \subseteq A_{k+2} \subseteq \ldots$.

Formalmente, pelo limite $X_k \uparrow X$ se k é suficientemente grande, $X_k(\omega) \geq Y(\omega)$. Assim, $\bigcup A_k = [Y \leq X] = \Omega$, e:

$$
YI_k = \begin{cases} Y(\omega) & \text{, se } \omega \in A_k \\ 0 & \text{, se } \omega \notin A_k \end{cases} = \begin{cases} n\epsilon & \text{, se } \omega \in B_n \cap A_k, n = 0, 1, \ldots \\ 0 & \text{, se } \omega \notin \bigcup_{n=0}^{\infty} B_n \cap A_k \end{cases}
$$

Logo, $0 \leq YI_k \leq X_k \Rightarrow 0 \leq E(YI_k) \leq E(X_k)$. Assim, $E(YI_k)$ será:

$$
E(YI_k) = \sum_{n=0}^{\infty} n\epsilon P(B_n \cap A_k) \geq \sum_{n=0}^{m} n\epsilon P(B_n \cap A_k)
$$

Logo, $\lim_{k\to\infty} E(X_k) \geq \lim_{k\to\infty} \sum_{n=0}^{m} n\epsilon P(B_n \cap A_k) = \sum_{n=0}^{m} n\epsilon P(B_n), \forall m$. Como m é arbitrário, $\lim_{k\to\infty} E(X_k) \geq \sum_{n=0}^{m} n\epsilon P(B_n) = E(Y)$. $\square$

Theorem 1.4 (Teorema da convergência dominada)**.** *Sejam $Y, X_1, X_2, \ldots, X$ variáveis aleatórias em $(\Omega, \mathcal{A}, P)$, tais que Y é integrável, $|X_n| \leq Y$ $\forall n$ e $X_n \to X$ (ou seja, dado $\omega : X_n(\omega) \xrightarrow[n\to\infty]{} X(\omega)$). Então X e X_n são integráveis e $E(X_n) \xrightarrow[n\to\infty]{} E(X)$.*

Prova. Por hipótese, temos a integrabilidade de X e X_n, tais que:

$$|X| = \lim_{n\to\infty} |X_n| \overset{\text{hip}}{\leq} Y$$

Assim, X e X_n são dominadas, então por E_2, X e X_n são integráveis. Defina $Y_n = \inf_{k\geq n} X_k$. Tomar o ínfimo provoca a sequência a se movimentar pela esquerda, de modo que:

$$X(\omega) \overset{\text{hip}}{=} \lim_{n\to\infty} X_n(\omega) = \lim_{n\to\infty} \inf X_n(\omega) = \lim_{n\to\infty} \left(\inf_{k\geq n} X_k(n) \right) = \lim_{n\to\infty} Y_n$$

E por definição de $Y_n : Y_n \uparrow X \Rightarrow (Y_n + Y) \underset{n\to\infty}{\uparrow} (X + Y)$. Aplicando o teorema da convergência monótona, temos que $|X_n| \leq Y$ $\forall n \Rightarrow -Y \leq X_n$, logo $\inf_{k\geq n} X_k \geq -Y \Rightarrow Y_n \geq -Y \Rightarrow Y_n + Y \geq 0$. Logo:

$$E(Y_n + Y) \underset{n\to\infty}{\uparrow} E(X + Y)$$

Defina $Z_n(\omega) = \sup_{k\geq n} X_k(\omega)$. Note que $Z_n(\omega) \underset{n\to\infty}{\downarrow} X(\omega)$, logo $(Y - Z_n) \underset{n\to\infty}{\uparrow} (Y - X)$. Note que $|X_n| \leq Y, \forall n \Rightarrow X_n \leq Y$, de modo que:

$$\sup_{k\geq n} X_k \leq Y \Rightarrow Z_n \leq Y \Rightarrow 0 \leq Y - Z_n$$

Agora que temos a positividade e a monotonia do crescimento, utilizamos o teorema da convergência monótona, de modo que:

$$E(Y - Z_n) \underset{n\to\infty}{\uparrow} E(Y - X) \Rightarrow E(Z_n) \underset{n\to\infty}{\downarrow} E(X)$$

Juntando as convergências de Y_n e Z_n:

$$Y_n = \inf_{k\geq n} X_k \leq X_n \leq \sup_{k\geq n} X_k = Z_n \Longrightarrow E(Y_n) \leq E(X_n) \leq E(Z_n)$$

De modo que $E(X_n) \to E(X)$. $\square$

1.6.1 Observações sobre o teorema da convergência dominada

1. Há casos nos quais $X_n \to X$, no entanto $E(X_n) \not\to E(X)$;

Example 1.7. $X \sim$ Cauchy $\to f(x) = \frac{1}{\pi(1+x^2)}$. Sabemos que $E(X) = \infty$. Seja $X_n = I_{[-n\leq x\leq n]} = \begin{cases} X & , -n \leq x \leq n \\ 0 & , |x| > n \end{cases}$. Então:

i) $X_n \underset{n\to\infty}{\to} X$;

ii) Cada X_n é limitado pois $|X_n| \leq n$, ou seja, podemos computar $E(X_n)$;

iii) $E(X_n) = 0$ $\forall n$ fixo, pois X_n é simétrica em torno de 0, de modo que $E(X_n) \not\to E(X) = \infty$.

2. Levemos em consideração que se $0 < s < t \Rightarrow |X|^s \leq 1 + |X|^t$. Então, se X é uma variável aleatória tal que $E(|X|^t) < \infty$ para $t > 0 \Rightarrow g(s) = E(|X|^s)$ é contínua para todo $s \in (0, t]$. Basta ver que se $s_n \to s \Rightarrow g(s_n) \to g(s)$. Note que $|X|^{s_n} \leq |X|^t + 1 \Rightarrow E(|X|^{s_n}) \leq E(|X|^t) + 1 < \infty$. Assim, pelo teorema da convergência dominada:

$$E(|X|^{s_n}) \to E(|X|^s)$$

3. **(Teorema de Arzelà)** Sejam $f, f_1, f_2, \ldots$, funções reais (borel-mensuráveis) definidas em $[a, b]$, com $a < b$ e integráveis a Riemann. Se $f_n \to f$ e se $|f_n| \leq M \ \forall n$, então:

$$\int_a^b f_n(x)dx \to \int_a^b f(x)dx$$

Prova. Considere $Omega = [a, b]$ e $\mathcal{A} = \mathcal{B}_{[a,b]}$ (ou seja, a $\sigma-$álgebra dos borelianos em $[a, b]$). Defina $\forall \omega \in \Omega$:

$$X_n(\omega) = (b - a)f_n(\omega)$$
$$X(\omega) = (b - a)f(\omega)$$

Já que $f_n \to f$ por hipótese, então $X_n(\omega) \to X(\omega) \ \forall \omega \in \Omega$. Também $|f_n| \leq M \Rightarrow (a - b)M \geq |X_n|$, logo $\{X_n\}$ é uma sequência de variáveis aleatórias integráveis, e sob o teorema da convergência dominada:

$$E(X_n) \underset{n\to\infty}{\longrightarrow} E(X)$$

□

4. **(Convergência de séries)** Se $a_{mn} \geq 0$ para $m, n = 1, 2, \ldots$ e $a_{mn} \underset{n\to\infty}{\uparrow} a_m \ \forall m$, então:

$$\sum_{m=1}^{\infty} a_{mn} \uparrow \sum_{m=1}^{\infty} a_m$$

Prova. Podemos escrever $a_{mn} = X_n(m)p_m$ onde $X_n(m) = \frac{a_{mn}}{p_m}, p_m = p(m) \ \forall m, m = \{1, 2, \ldots\}$, sendo p_m tal que $\sum_{m=1}^{\infty} p_m = 1$. Analogamente $a_m = X(m)p_m$, e sendo $p_m \geq 0 \ \forall m$, como $a_{mn} \geq 0$ e $p_m \geq 0 \Rightarrow X_n(m) \geq 0$. Já que $a_{mn} \underset{n\to\infty}{\uparrow} a_m \Rightarrow 0 \leq X_n(m) \underset{n\to\infty}{\uparrow} X(m), \forall m$. Logo, pelo teorema da convergência monótona, $E(X_n) \underset{n\to\infty}{\uparrow} E(X)$, e assim:

$$E(X_n) = \sum_{m=1}^{\infty} X_n(m)p_m = \sum_{m=1}^{\infty} a_{mn}$$
$$E(X) = \sum_{m=1}^{\infty} X(m)p_m = \sum_{m=1}^{\infty} a_m$$

Ou seja, $\sum_{m=1}^{\infty} a_{mn} \underset{n\to\infty}{\uparrow} \sum_{m=1}^{\infty} a_m$. □

1.7 Exercícios

TODO