Vetores aleatórios discretos (parte 1)

Prof. Caio Azevedo

Vetores aleatórios

- Seja $(\Omega, \mathcal{A}, \mathcal{P})$ um espaço de probabilidade. Um vetor aleatório n-variado é um vetor de funções que associa cada valor do espaço amostral Ω a um vetor no $\mathcal{R}^n = \mathcal{R} \times \mathcal{R} \times ... \times \mathcal{R}$.
- Em outras palavras um vetor aleatório (ve) $\boldsymbol{X} = (X_1, X_2, ...X_n)'$ é um vetor cujas componentes são variáveis aleatórias. O valor observado de um ve é denotado por $\boldsymbol{x} = (x_1, x_2, ...x_n)'$.
- Notação: Letras maiúsculas para cada componente do vetor aleatório (por exemplo $(X, Y, ..., Z, W)'$) e valor assumido pelo vetor aleatório (por exemplo $(x, y, ..., z, w)$'). Vetores aleatórios serão denotados por letras em negrito.

Vetores aleatórios

- Nosso **foco** serão os vetores aleatórios bi e tridimensionais e discretos.
- Para vetores aleatórios contínuos veja, por exemplo, o curso de Probabilidade II (ME310).
- Observação: Cada componente de uma vetor aleatório bivariado é uma variável aleatória univariada, todos definidos no mesmo espaço de probabilidade.
- Um vetor aleatório discreto, digamos $\boldsymbol{X} = (X_1, .., X_n)'$ é um vetor aleatório em que cada componente é uma vad.
- Apresentaremos várias definições e propriedades para o caso bivariado, ficando as devidas extensões para o caso trivariado, como exerício.

Definições e propriedades

- Seja $\boldsymbol{Z} = (X, Y)'$ um ved (vetor aleatório discreto) bivariado.
- Uma função de probabilidade (ou função densidade de probabilidade) (fdp) bivariada, associada à $\boldsymbol{Z}$ é uma função bivariada ($p_{\boldsymbol{Z}} : \mathcal{R}^2 \to [0, 1]$), tal que:
  - $0 \leq p_{\boldsymbol{Z}}(x, y) \leq 1, \forall (x, y) \in \mathcal{R}^2$.
  - $\sum_x \sum_y p_{\boldsymbol{Z}}(x, y) = \sum_y \sum_x p_{\boldsymbol{Z}}(x, y) = 1$.
- Notações mais comuns (a menos da função indicadora) (em que $\boldsymbol{z} = (x, y'))$:

$$
\begin{aligned}
f_{(X,Y)}(x, y) &= p_{(X,Y)}(x, y) = P(X = x, Y = y) = f_{\boldsymbol{Z}}(x, y) \\
&= p_{\boldsymbol{Z}}(x, y) = f_{\boldsymbol{Z}}(x, y) = p_{\boldsymbol{Z}}(\boldsymbol{z}) = P(\boldsymbol{Z} = \boldsymbol{z}).
\end{aligned}
$$

Definições e propriedades

- Obs: a definição de fdp bivariada (fdpb) é geral, ou seja, não precisa estar associada a um ve.
- Pela definição, notamos que a fdpb fornece probabilidades conjuntas relativas à um ved.
- Invariavelmente, à uma fdpb, associaremos seu suporte, ou seja um conjunto $A \subset \mathcal{R}$ na qual aquela terá valor no intervalo $(0, 1)$.
- Discutiremos, posteriormente, as definições e conceitos através de exemplos.

Definições e propriedades

- Dado um vetor aleatório $\boldsymbol{Z} = (X, Y)'$ e sua fdpb $p_{\boldsymbol{Z}}$, as distribuições marginais de X e Y são dadas, respectivamente, por:

$$
\begin{aligned}
f_X(x) &= P(X = x) = p_X = \sum_y p_Z(X = x, Y = y). \\
f_Y(y) &= P(Y = y) = p_Y = \sum_x p_Z(X = x, Y = y).
\end{aligned}
$$

- Para as distribuições (fdp) marginais, podemos usar as mesmas notações vistas aqui.
- As definições acima são compatíveis com o teorema da probabilidade total.

Definições e propriedades

- Dado um vetor aleatório $Z = (X, Y)$ e sua fdpb p_Z, as distribuições condicionais de $X|Y = y$ e $Y|X = x$ são dadas, respectivamente, por:

$$
\begin{aligned}
P(X = x|Y = y) &= \begin{cases} \dfrac{P(X = x, Y = y)}{P(Y = y)}, \text{se } P(Y = y) \neq 0 \\ 0, \text{se } P(Y = y) = 0. \end{cases} \\
P(Y = y|X = x) &= \begin{cases} \dfrac{P(X = x, Y = y)}{P(X = x)}, \text{se } P(X = x) \neq 0 \\ 0, \text{se } P(X = x) = 0. \end{cases}
\end{aligned}
$$

- As distribuições condicionais tem interpretações semelhantes às probabilidades condicionais de eventos (aqui).

Definições e propriedades

- Um conceito importante é o de dependência estocástica (probabilística) já comentada anteriormente (veja aqui e aqui).
- Essencialmente, tal dependência é totalmente definida pela distribuição conjunta, ou seja:

$$p_Z(x,y) = P(X = x, Y = y)\mathbb{1}_A(x,y).$$

Definições e propriedades

- Sejam X e Y duas va's definidas no mesmo espaço de probabilidade.

  Então

$$X \perp Y \leftrightarrow P(X = x, Y = y) = P(X = x)P(Y = y) \forall x, y$$

- $X \perp Y$ significa: X e Y são independentes.
- A definição acima implica que

$$\begin{aligned} X \perp Y \quad & \leftrightarrow P(X = x | Y = y) = P(X = x) \\ \text{e } & \leftrightarrow P(Y = y | X = x) = P(Y = y) \forall x, y. \end{aligned}$$

Definições e propriedades

- Se $Z = (X, Y)'$ é um ve com fdpc (fdp conjunta) p_Z com suporte A e $g(X, Y), g : A \to \mathcal{R}$ (uma va) então:

$$
\begin{aligned}
\mathcal{E}\left[g(X,Y)\right] &= \sum_x \sum_y g(x,y) P(X = x, Y = y) \mathbb{1}_A(x,y) \\
&= \sum_y \sum_x g(x,y) P(X = x, Y = y) \mathbb{1}_A(x,y)
\end{aligned}
$$

- Se $X \perp Y$ então (exercício)

$$\mathcal{E}(XY) = \mathcal{E}(X)\mathcal{E}(Y).$$

Definições e propriedades

- Vamos agora estudar como "quantificar" o grau de dependência entre duas variáveis aleatórias.
- Inicialmente, vamos considerar o conceito de covariância entre duas variáveis X e Y.
- A covariância entre duas variáveis aleatórias $X(\mathcal{E}(X) = \mu_X)$ e $Y(\mathcal{E}(Y) = \mu_Y)$ é dada (se existir) por:

$$\sigma_{XY} = Cov(X, Y) = \mathcal{E}\left[(X - \mu_X)(Y - \mu_Y)\right].$$

Definições e propriedades

- Uma fórmula alternativa (e mais útil para seu cálculo) da covariância, é dada por (Exercício):
- Note que: $\sigma_{XY} = Cov(X, Y) = \mathcal{E}(XY) - \mu_X \mu_Y$.
  - Se quanto maior/menor for o valor de X (ou de Y) maior/menor for o valor de Y (ou de X), maior (mais positiva) será σ_{XY}.
  - Se quanto maior/menor for o valor de X (ou de Y) menor/maior for o valor de Y (ou de X), maior (mais positiva) será, maior (mais negativa) será σ_{XY}.
  - Se quanto maior/menor for o valor de X (ou de Y) o valor de Y (ou de X), variar indistintamente mais próximo de zero será σ_{XY}.

Definições e propriedades

- Com efeito, assuma que $Y = a + bX$, $a, b \in \mathcal{R}$. Logo

$$
\begin{aligned}
\sigma_{XY} &= \mathcal{E}(X(a+bX)) - \mathcal{E}(a+bX)\mu_x = a\mu_X + b\mathcal{E}(X^2) - a\mu_X - b\mu_X^2 \\
&= b\left(\mathcal{E}(X^2) - \mu_X^2\right) = b\mathcal{V}(X).
\end{aligned}
$$

- Assim, se :
  - Se $b \to \infty \Rightarrow \sigma_{XY} \to \infty$.
  - Se $b \to -\infty \Rightarrow \sigma_{XY} \to -\infty$.
  - Se $b \to 0 \Rightarrow \sigma_{XY} \to 0$.

Definições e propriedades

- Logo, σ_{XY} mede um tipo de dependência estocástica linear entre X e Y. Contudo, a interpretação da magnitude não é simples.
- Dessa forma, faz-se mister a definição de uma quantidade mais interpretável.
- A correlação (de Pearson) entre duas variáveis aleatórias $X(\mathcal{E}(X) = \mu_X, \mathcal{DP}(X) = \sigma_X)$ e $Y(\mathcal{E}(Y) = \mu_Y, \mathcal{DP}(Y) = \sigma_Y)$ é dada (se existir) por:

$$\rho_{XY} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}.$$

Definições e propriedades

- Teorema: Se ρ_{XY} existir, então $-1 \leq \rho_{XY} \leq 1$.
- Dem. : (Usaremos a Desigualdade de Cauchy-Schwarz (CS)). Temos que:

$$
\begin{aligned}
|\sigma_{XY}| &= |\mathcal{E}\left[(X-\mu_X)(Y-\mu_Y)\right]| \underbrace{\leq}_{\text{CS}} \sqrt{\mathcal{E}\left[(X-\mu_X)^2\right]\mathcal{E}\left[(Y-\mu_Y)^2\right]} \\
&= \sqrt{\mathcal{V}(X)}\sqrt{\mathcal{V}(Y)} \rightarrow \frac{|\sigma_{XY}|}{\mathcal{DP}(X)\mathcal{DP}(Y)} \leq 1 \rightarrow |\rho_{XY}| \leq 1.
\end{aligned}
$$

■

Definições e propriedades

- Suponha que $Y = a + bX + \Delta$, $a, b \in \mathcal{R}$, $\Delta \perp X$, $\mathcal{E}(\Delta) = 0, \mathcal{V}(\Delta) = 1$.

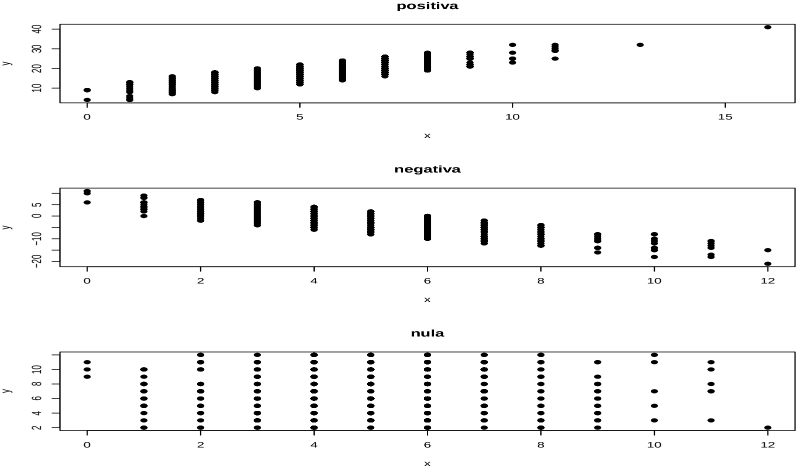
  Temos que:

$$
\begin{aligned}
\rho_{XY} &= \frac{\mathcal{E}(X(a + bX + \Delta)) - \mu_X(a + b\mu_X)}{\mathcal{DP}(X)\sqrt{b^2\mathcal{V}(X) + 1}} \\
&= \frac{a\mu_X + b\mathcal{E}(X^2) - \mu_X(a + b\mu_X)}{\mathcal{DP}(X)\sqrt{b^2\mathcal{V}(X) + 1}} = \frac{b\mathcal{V}(X)}{\mathcal{DP}(X)\sqrt{b^2\mathcal{V}(X) + 1}}.
\end{aligned}
$$

Definições e propriedades

- (Cont.)
  - Se $b \to \infty \Rightarrow \rho_{XY} \to 1$.
  - Se $b \to -\infty \Rightarrow \rho_{XY} - 1$.
  - Se $b \to 0 \Rightarrow \rho_{XY} \to 0$.
- No próximo slide apresentaremos valores simulados de duas va's com padrões de correlações distintos.

Definições e propriedades



Definições e propriedades

- Sejam X e Y duas variáveis aleatórias, $\mathcal{E}(|X^k|) < \infty$ e $\mathcal{E}(|Y^k|) < \infty, k = 1, 2$. Então
  1. $\mathcal{E}(X \pm Y) = \mathcal{E}(X) \pm \mathcal{E}(Y)$.
  2. $\mathcal{V}(X \pm Y) = \mathcal{V}(X) + \mathcal{V}(Y) \pm 2Cov(X, Y)$.
- Dem. (item 1, Exercício). Item 2) (considerando só a soma, a diferença fica como exercício)

$$
\begin{aligned}
\mathcal{V}(X+Y) &= \mathcal{E}\left[(X+Y)^2\right] - \mathcal{E}^2\left[(X+Y)\right] \\
&= \mathcal{E}(X^2) - \mathcal{E}^2(X) + \mathcal{E}(Y^2) - \mathcal{E}^2(Y) \\
&+ 2\mathcal{E}(XY) - 2\mathcal{E}(X)\mathcal{E}(Y) \\
&= \mathcal{V}(X) + \mathcal{V}(Y) + 2Cov(X, Y)
\end{aligned}
$$

Definições e propriedades

- Momentos condicionais. Sejam X e Y duas vad's definidas no mesmo espaço de probabilidade, então

$$
\begin{aligned}
\mathcal{E}(g(X)|Y=y) &= \sum_{x} g(x)P(X=x|Y=y)\mathbb{1}_A(x) \\
&= \sum_{x \in A} g(x)P(X=x|Y=y),
\end{aligned}
$$

em que A é o suporte da distribuição de $X|Y=y$.

Definições e propriedades

- Em particular:

$$\begin{aligned}
\mathcal{E}(X|Y=y) &= \sum_x xP(X=x|Y=y)\mathbb{1}_A(x) \\
\mathcal{V}(X|Y=y) &= \mathcal{E}(X^2|Y=y) - \mathcal{E}^2(X|Y=y) \\
&= \sum_x x^2P(X=x|Y=y)\mathbb{1}_A(x) \\
&- \left(\sum_x xP(X=x|Y=y)\mathbb{1}_A(x)\right)^2
\end{aligned}$$

- Observação: Enquanto que $\mathcal{E}(g(X)|Y=y)$ é um número, $\mathcal{E}(g(X)|Y)$ é uma va (nesse caso, discreta).

Definições e propriedades

- Teorema: Sejam X e Y duas vad's definidas no mesmo espaço de probabilidade, então

$$\mathcal{E}(X) = \mathcal{E}(\mathcal{E}(X|Y)). \quad (1)$$

$$\mathcal{V}(X) = \mathcal{E}(\mathcal{V}(X|Y)) + \mathcal{V}(\mathcal{E}(X|Y)). \quad (2)$$

- Dem: A Equação (2) fica com exercício. Com relação à Equação (1), temos que (próximo slide):

Definições e propriedades

- Cont.

$$
\begin{aligned}
\mathcal{E}(X) &= \sum_{x \in A} xP(X = x) = \sum_{x} x \sum_{y} P(X = x, Y = y) \\
&= \sum_{y} P(Y = y) \underbrace{\sum_{x} xP(X = x | Y = y)}_{\mathcal{E}(X|Y=y)} \\
&= \sum_{y} P(Y = y)\mathcal{E}(X|Y = y) = \mathcal{E}(\mathcal{E}(X|Y))
\end{aligned}
$$

■

Definições e propriedades

- Função distribuição acumulada bivariada de $\boldsymbol{Z} = (X, Y)'$:

$$\begin{aligned} F_{\boldsymbol{Z}}(x,y) &= P(X \leq x, Y \leq y) = \\ &= \sum_{z \leq x} \sum_{w \leq y} P(X = z, Y = w) \end{aligned}$$

- Temos que $X \perp Y \leftrightarrow F_{\boldsymbol{Z}}(x,y) = F_X(x) F_Y(y), \forall x, y$.

Definições e propriedades

- Seja $F_Z(.)$ uma fdab qualquer (discreta ou contínua), então:

  1. $F_Z(-\infty, x) = \lim_{x \to -\infty} F_Z(x,y) = 0$, $F_Z(x, -\infty) = \lim_{y \to -\infty} F_Z(x,y) = 0$ e $F_Z(\infty, \infty) = \lim_{x \to \infty,\ y \to \infty} F_Z(x,y) = 1$.
  2. Se $x_1 < x_2$ e $y_1 < y_2$, então

$$
\begin{aligned}
P(a_1 \leq X \leq a_2, b_1 \leq Y \leq b_2) &= F_Z(a_2,b_2) - F_Z(a_2,b_1) - F_Z(a_1,b_2) \\
&+ F_Z(a_1,b_1) \geq 0.
\end{aligned}
$$

  3. $F_Z(.)$ é contínua à direita em cada componente, ou seja, $\lim_{0<h \to 0} F_Z(x+h,y) = F_Z(x,y)$ e $\lim_{0<h \to 0} F_Z(x,y+h) = F_Z(x,y)$.

- Prova: exercício.

Definições e propriedades

- De uma forma geral, $\forall C_1, C_2 \subset \mathcal{R}$, temos que:

$$P(X \in C_1, Y \in C_2) = \sum_{x \in C_1} \sum_{y \in C_2} P(X = x, Y = y) \geq 0.$$

- Exemplos:

$$\begin{aligned}
P(X \geq x, Y < y) &= \sum_{z \geq x} \sum_{w < y} P(X = z, Y = w). \\
P(X > x, Y = y) &= \sum_{z > x} P(X = z, Y = y). \\
P(X > w, Y \geq X) &= \sum_{x > w} \sum_{y \geq x} P(X = x, Y = y).
\end{aligned}$$

Definições e propriedades

- Se algumas "áreas" (e/ou intervalos) definidas pelas regiões C_1 e/ou C_2 estiveram fora do suporte de $\boldsymbol{Z} = (X, Y)'$, então esses valores terão probabilidade zero.
- Assim como no caso univariado, considerar ($<$ ou $\leq$) e/ou ($>$ ou $\geq$), pode levar a probabilidades diferentes.

Definições e propriedades

- Função geradora de momentos bivariada ($\boldsymbol{t} = (t_1, t_2)'$) de $\boldsymbol{Z} = (X, Y)'$:

$$\phi_{\boldsymbol{Z}}(\boldsymbol{t}) = \mathcal{E}(e^{\boldsymbol{t}'\boldsymbol{z}}) = \mathcal{E}(e^{t_1 x + t_2 y}) = \sum_x \sum_y e^{t_1 x + t_2 y} P(X = x, Y = y).$$

- Temos que $\phi_X(t_1) = \phi_{\boldsymbol{Z}}(t_1, 0)$ e $\phi_Y(t_2) = \phi_{\boldsymbol{Z}}(0, t_2)$.
- É possível provar que ($\boldsymbol{0} = (0, 0)'$):

$$\left.\frac{\partial^{r+s}}{\partial t_1^r \partial t_2^s} \phi_{\boldsymbol{Z}}(\boldsymbol{t})\right|_{\boldsymbol{t}=\boldsymbol{0}} = \mathcal{E}(X^r Y^s).$$

- $X \perp Y \leftrightarrow \phi_{\boldsymbol{Z}}(\boldsymbol{t}) = \phi_X(t_1)\phi_Y(t_2)$.

Definições e propriedades

- Função geradora de probabilidades bivariada ($\boldsymbol{t} = (t_1, t_2)'$) de $\boldsymbol{Z} = (X, Y)'$:

$$G_{\boldsymbol{Z}}(\boldsymbol{t}) = \mathcal{E}(t_1^X t_2^Y) = \sum_x \sum_y t_1^X t_2^Y P(X = x, Y = y).$$

- Temos que $\phi_X(t_1) = \phi_{\boldsymbol{Z}}(t_1, 1)$ e $\phi_Y(t_2) = \phi_{\boldsymbol{Z}}(1, t_2)$.
- É possível provar que ($\boldsymbol{0} = (0, 0)'$):

$$\left.\frac{\partial^{r+s}}{\partial t_1^r \partial t_2^2} G_{\boldsymbol{Z}}(\boldsymbol{t})\right|_{\boldsymbol{t}=\boldsymbol{0}} = \frac{1}{r!s!} P(X = r, Y = s).$$

- $X \perp Y \leftrightarrow G_{\boldsymbol{Z}}(\boldsymbol{t}) = G_X(t_1) G_Y(t_2)$.

Exemplo

- Considere o experimento aleatório que consiste em sortear duas bolas, sem reposição de uma urna com que contem 3 bolas vermelhas (V) e 2 brancas (B).
- Defina as seguintes variáveis aleatórias:
  - X: o número de bolas brancas observadas.
  - Y: a cor da segunda bola sorteada em que 1 se a bola foi V e 0, se foi B.
- Defina $\boldsymbol{Z} = (X, Y)'$. Vamos obter diversos resultados probabilísticos.

Exemplo

- Primeiramente, seguindo a notação usual, note que:

$$\Omega = \{(1B, 2B), (1B, 2V), (1V, 2B), (1V, 2V)\}.$$

- Podemos provar que :

$$\begin{aligned}
P[(1B, 2B)] &= P(1B)P(2B|1B) = \frac{2}{5} \times \frac{1}{4} = \frac{1}{10} \\
P[(1B, 2V)] &= P(1B)P(2V|1B) = \frac{2}{5} \times \frac{3}{4} = \frac{3}{10} \\
P[(1V, 2B)] &= P(1V)P(2B|1V) = \frac{3}{5} \times \frac{2}{4} = \frac{3}{10} \\
P[(1V, 2V)] &= P(1V)P(2V|11) = \frac{3}{5} \times \frac{2}{4} = \frac{3}{10}
\end{aligned}$$

(Cont.) Exemplo

- Sendo A o suporte de $\boldsymbol{Z}$, conseguimos verificar que:

$$A(:\Omega \to \mathcal{R}^2) = \{\{2,0\},\{1,1\},\{1,0\},\{0,1\}\}.$$

- Logo, podemos expressar a fdp de $\boldsymbol{Z}$ através de

$$P(X = x, Y = y) = \frac{1}{10}\mathbb{1}_{\{2,0\}}(\boldsymbol{z}) + \frac{3}{10}\mathbb{1}_{\{1,1\},\{1,0\},\{0,1\}}(\boldsymbol{z}).$$

(Cont.) Exemplo

- Em casos como esse, em que o suporte é formado por um número muito pequeno de valores, podemos representar a fdpc através de uma tabela de contingência, ou seja:

	X			
Y	0	1	2	
0	0	$\frac{3}{10}$	$\frac{1}{10}$	$\frac{2}{5}$
1	$\frac{3}{10}$	$\frac{3}{10}$	0	$\frac{3}{5}$
	$\frac{3}{10}$	$\frac{6}{10}$	$\frac{1}{10}$	1

(Cont.) Exemplo

- Podemos notar que as distribuições marginais (que ficam nas margens da tabela anterior) de X e Y são dadas respectivamente por:

$$
\begin{aligned}
P(X = x) &= \frac{3}{10}\mathbb{1}_{\{0\}}(x) + \frac{6}{10}\mathbb{1}_{\{1\}}(x) + \frac{1}{10}\mathbb{1}_{\{2\}}(x) \\
&= \frac{\binom{2}{x}\binom{3}{2-x}}{\binom{5}{2}}\mathbb{1}_{\{0,1,2\}}(x), \\
P(Y = y) &= \left(\frac{3}{5}\right)^{y}\left(\frac{2}{5}\right)^{1-y}\mathbb{1}_{\{0,1\}}(y).
\end{aligned}
$$

- Assim, temos que $X \sim HG(5,2,2)$ e $Y \sim$ Bernoulli$(3/5)$ (como esperado).

(Cont.) Exemplo

- Exercício, deduzir as distribuições marginais de X e Y, ou seja, $P(X = x)$ e $P(Y = y)$ através das características do experimento aleatório e das respectivas definições da Bernoulli e da hipergeométrica.
- Por outro lado, note que (por exemplo)

$$P(X = 2, Y = 1) = 0 \neq P(X = 2)P(Y = 1) = \frac{1}{10} \times \frac{3}{5}.$$

Logo, as variáveis são (estocastimente) dependentes.

- Com relação aos momentos, temos que $\mathcal{E}(X) = \frac{4}{5}$ e $\mathcal{V}(X) = \frac{9}{25}$, $\mathcal{E}(Y) = \frac{3}{5}$ e $\mathcal{V}(Y) = \frac{6}{25}$ (veja aqui).

(Cont.) Exemplo

- Para calcular $Cov(X,Y)$, temos que calcular $\mathcal{E}(XY)$, ou seja

$$\mathcal{E}(XY) = \sum_{x=0}^{2}\sum_{y=0}^{1} xyP(X=x, Y=y) = 1 \times 1 \times \frac{3}{10} + 2 \times 1 \times 0 = \frac{3}{10}.$$

- Logo

$$\sigma_{XY} = \frac{3}{10} - \frac{4}{5} \times \frac{3}{5} = -\frac{9}{50} \approx -0,180.$$

(Cont.) Exemplo

- Finalmente, para a correlação, temos que:

$$\rho_{XY} = \frac{-\frac{9}{50}}{\sqrt{\frac{9}{25} \times \frac{6}{25}}} \approx -0,612.$$

- Portanto, X e Y são negativa, e altamente correlacionadas.

(Cont.) Exemplo

- Distribuições condicionais.
- Note que temos $X|Y=0$ e $X|Y=1$. Analogamente, temos $Y|X=0$, $Y|X=1$ e $Y|X=2$.
- Com efeito, obtemos que

$$
\begin{aligned}
P(X=0|Y=0) &= \frac{P(X=0,Y=0)}{P(Y=0)} = \frac{0}{2/5}, \\
P(X=1|Y=0) &= \frac{P(X=1,Y=0)}{P(Y=0)} = \frac{3/10}{2/5} = \frac{3}{4}, \\
P(X=2|Y=0) &= \frac{P(X=2,Y=0)}{P(Y=0)} = \frac{1/10}{2/5} = \frac{1}{4}.
\end{aligned}
$$

(Cont.) Exemplo

- Analogamente, temos que:

$$\begin{aligned} P(X=0|Y=1) &= \frac{P(X=0,Y=1)}{P(Y=1)} = \frac{3/10}{3/5} = \frac{1}{2}, \\ P(X=1|Y=1) &= \frac{P(X=1,Y=1)}{P(Y=1)} = \frac{3/10}{3/5} = \frac{1}{2}, \\ P(X=2|Y=1) &= \frac{P(X=2,Y=1)}{P(Y=1)} = 0. \end{aligned}$$

- Assim

$$\begin{aligned} P(X=x|Y=0) &= \frac{3}{4}\mathbb{1}_{\{1\}}(x) + \frac{1}{4}\mathbb{1}_{\{2\}}(x) \\ X|Y=1 &\sim \text{Bernoulli}(1/2) \equiv U_{\{0,1\}}. \end{aligned}$$

(Cont.) Exemplo

- Note ainda que $X|Y = 0 \equiv Z + 1$, em que $Z \sim \text{Bernoulli}(1/4)$.
- Por outro lado, temos que:

$$
\begin{aligned}
P(Y=0|X=0) &= \frac{P(Y=0, X=0)}{P(X=0)} = 0, \\
P(Y=1|X=0) &= \frac{P(Y=1, X=0)}{P(X=0)} = \frac{3/10}{3/10} = 1.
\end{aligned}
$$

- Analogamente, temos que:

$$
\begin{aligned}
P(Y=0|X=1) &= \frac{P(Y=0, X=1)}{P(X=1)} = \frac{3/10}{6/10} = \frac{1}{2}, \\
P(Y=1|X=1) &= \frac{P(Y=1, X=1)}{P(X=1)} = \frac{3/10}{6/10} = \frac{1}{2}.
\end{aligned}
$$

(Cont.) Exemplo

- Finalmente, temos que:

$$
\begin{aligned}
P(Y=0|X=2) &= \frac{P(Y=0,X=2)}{P(X=2)} = \frac{1/10}{1/10} = 1, \\
P(Y=1|X=2) &= \frac{P(Y=1,X=2)}{P(X=1)} = 0.
\end{aligned}
$$

- Assim,

$$
\begin{aligned}
P(Y=1|X=0) &\equiv 1, \\
Y|X=1 &\sim \text{Bernoulli}(1/2) \equiv U_{\{0,1\}}, \\
P(Y=0|X=2) &\equiv 1.
\end{aligned}
$$

(Cont.) Exemplo

- Logo, $Y|X = 0$ e $Y|X = 2$ são exemplos de va's degeneradas, nos valores 1 e 0, respectivamente.
- Vamos agora ilustrar a validade das Equações (1) e (2).
- Inicialmente, notemos que:

$$\begin{aligned}
\mathcal{E}(X|Y=0) &= \frac{5}{4}, \mathcal{E}(X|Y=1) = \frac{1}{2}, \\
\mathcal{V}(X|Y=0) &= \frac{3}{16}, \mathcal{V}(X|Y=1) = \frac{1}{4}.
\end{aligned}$$

(Cont.) Exemplo

- Assim, temos que:

$$
\begin{aligned}
\mathcal{E}(\mathcal{E}(X|Y)) &= \sum_{y=0}^{1} \mathcal{E}(X|Y=y)P(Y=y) \\
&= \mathcal{E}(X|Y=0)P(Y=0) + \mathcal{E}(X|Y=1)P(Y=1) \\
&= \frac{5}{4}\frac{2}{5} + \frac{1}{2}\frac{3}{5} = \frac{8}{10} = \frac{4}{5}.
\end{aligned}
$$

(Cont.) Exemplo

- Por outro lado, temos que:

$$\begin{aligned}
\mathcal{E}(\mathcal{V}(X|Y)) &= \sum_{y=0}^{1} \mathcal{V}(X|Y=y)P(Y=y) \\
&= \mathcal{V}(X|Y=0)P(Y=0) + \mathcal{V}(X|Y=1)P(Y=1) \\
&= \frac{3}{16}\frac{2}{5} + \frac{1}{4}\frac{3}{5} = \frac{9}{40}.
\end{aligned}$$

(Cont.) Exemplo

- Adicionalmente, tem-se que:

$$
\begin{aligned}
\mathcal{V}(\mathcal{E}(X|Y)) &= \mathcal{E}\left[(\mathcal{E}(X|Y))^2\right] - \mathcal{E}^2\left[(\mathcal{E}(X|Y))\right] \\
&= \mathcal{E}\left[(\mathcal{E}(X|Y))^2\right] - \mathcal{E}^2(X) \\
&= \mathcal{E}\left[(\mathcal{E}(X|Y))^2\right] - \frac{16}{25}
\end{aligned}
$$

(Cont.) Exemplo

- Agora, note que:

$$
\begin{aligned}
\mathcal{E}\left[(\mathcal{E}(X|Y))^2\right] &= \sum_{y=0}^{1} (\mathcal{E}(X|Y=y))^2 P(Y=y) \\
&= (\mathcal{E}(X|Y=0))^2 P(Y=0) \\
&+ (\mathcal{E}(X|Y=1))^2 P(Y=1) = \frac{25}{16}\frac{2}{5} + \frac{1}{4}\frac{3}{5} \\
&= \frac{31}{40}.
\end{aligned}
$$

(Cont.) Exemplo

- Portanto, temos que:

$$\mathcal{V}(\mathcal{E}(X|Y)) = \frac{31}{40} - \frac{16}{25} = \frac{155 - 128}{200} = \frac{27}{200}$$

- Finalmente, vem que:

$$\mathcal{E}(\mathcal{V}(X|Y)) + \mathcal{V}(\mathcal{E}(X|Y)) = \frac{9}{40} + \frac{27}{200} = \frac{72}{200} = \frac{9}{25} = \mathcal{V}(X).$$

- Exercício: repetir a verificação para $\mathcal{E}(Y)$ e $\mathcal{V}(Y)$.

Mais resultados

- Sejam $X_i, i = 1, 2, .., n$ e $Y_j, j = 1, 2, ..., m$ va's definidas no mesmo espaço de probabilidade e $a_i, i = 1, 2, .., n, b_j, j = 1, 2, ..., m(a_i, b_i \in \mathcal{R})$. Então:
  - $Cov\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{m} b_j Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j Cov(X_i, Y_j)$.
  - $\mathcal{V}\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 \mathcal{V}(X_i) + 2\sum_{i<j} a_i b_j Cov(X_i, X_j)$.
- Dem. Exercício.