Vetores aleatórios discretos (parte 3)

Prof. Caio Azevedo

Vetores aleatórios

- Seja $(\Omega, \mathcal{A}, \mathcal{P})$ um espaço de probabilidade.
- Consideraremos agora (algumas) extensões de conceitos e resultados vistos aqui para o caso de vead (vetores aleatórios discretos) trivariados.
- Seja $\boldsymbol{W} = (X, Y, Z)'$ um ved (vetor aleatório discreto) trivariado.

Definições e propriedades

- Uma função de probabilidade (ou função densidade de probabilidade) (fdp) trivariada, associada à $\boldsymbol{W}$ é uma função trivariada ($p_{\boldsymbol{Z}} : \mathcal{R}^3 \rightarrow [0, 1]$), tal que:
  - $0 \leq p_{\boldsymbol{W}}(x, y, z) \leq 1, \forall (x, y, z) \in \mathcal{R}^3$.
  - $\sum\limits_{x,y,z} p_{\boldsymbol{W}}(x, y, z) = 1$,

    em que (a menos que mencionado algo diferente) $\sum_{x,y,z}$ representa o somatório triplo em qualquer ordem.

Definições e propriedades

- Notações mais comuns (a menos da função indicadora) (em que $\boldsymbol{w} = (x, y, z)'$):

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

Definições e propriedades

- Obs: a definição de fdp trivariada (fdpt) é geral, ou seja, não precisa estar associada a um vea.
- Pela definição, notamos que a fdpt fornece probabilidades conjuntas relativas à um vead.
- Invariavelmente, à uma fdpt, associaremos seu suporte, ou seja um conjunto $A \subset \mathcal{R}^3$ na qual aquela terá valor no intervalo $(0, 1)$.

Definições e propriedades

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

$$
\begin{aligned}
f_X(x) &= P(X = x) = p_X = \sum_{y,z} p_{\boldsymbol{W}}(X = x, Y = y, Z = z). \\
f_Y(y) &= P(Y = y) = p_Y = \sum_{x,z} p_{\boldsymbol{W}}(X = x, Y = y, Z = z) \\
f_Z(z) &= P(Z = z) = p_Z = \sum_{x,y} p_{\boldsymbol{W}}(X = x, Y = y, Z = z)
\end{aligned}
$$

em que os somatórios acima também representam (a menos que seja mencionado o contrário) ordem invariantes.

Definições e propriedades

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

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

Definições e propriedades

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

Definições e propriedades

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

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

Definições e propriedades

- (Cont.)

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

Definições e propriedades

- (Cont.)

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

Definições e propriedades

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

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

Definições e propriedades

- (Cont.)

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

Definições e propriedades

- (Cont.)

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

- 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:

$$p_{\mathbf{W}}(x, y, z) = P(X = x, Y = y, Z = z)\mathbb{1}_A(x, y, z).$$

Definições e propriedades

- Sejam X, Y e Z três va's definidas no mesmo espaço de probabilidade. Então

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

- $X \perp Y \perp Z$ significa: X, Y e Z são mutuamente independentes.
- A definição acima implica que (independência dois a dois)

$$X \perp Y, X \perp Z, Y \perp Z.$$

Definições e propriedades

- Contudo, independência dois a dois não implica, necessariamente, em independência conjunta (trivariada), veja aqui.
- Independência trivariada implica que todas as distribuições condicionais, são iguais as distribuições marginais, ou seja:

$$
\begin{aligned}
P(X = x|Z = z) &= P(X = x); P(Z = z|X = x) = P(Z = z) \\
P(X = x|Y = Y) &= P(X = x); P(Y = y|X = x) = P(Y = y) \\
P(Y = y|Z = z) &= P(Y = y); P(Z = z|Y = y) = P(Z = z)
\end{aligned}
$$

Definições e propriedades

- (Cont.)

$$
\begin{aligned}
P(X = x, Y = y|Z = z) &= P(X = x, Y = y)\\
P(X = x, Z = z|Y = y) &= P(Y = y, Z = z)\\
P(Y = y, Z = z|X = x) &= P(Y = y, Z = z)
\end{aligned}
$$

Definições e propriedades

- Sejam X, Y e Z três va's definidas no mesmo espaço de probabilidade. Então $(X \perp Y | Z)$ significa que X e Y são condicionalmente independentes dado Z. Com efeito:

$$
\begin{aligned}
(X \perp Y)|Z &\leftrightarrow P(X = x, Y = y | Z = z) \\
&= P(X = x | Z = z) P(Y = y | Z = z), \\
(X \perp Z)|Y &\leftrightarrow P(X = x, Z = z | Y = y) \\
&= P(X = x | Y = y) P(Z = z | Y = y), \\
(Y \perp Z)|X &\leftrightarrow P(Y = y, Z = z | X = x) \\
&= P(Y = y | X = x) P(Z = z | X = x).
\end{aligned}
$$

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:

$$\mathcal{E}\left[g(X, Y, Z)\right] = \sum_{x,y,z} g(x, y, z) P(X = x, Y = y, Z = z) \mathbb{1}_A(x, y, z)$$

- Se $X \perp Y \perp Z$ então (exercício)

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

- Está além do escopo desta disciplina estudar medidas de dependência para três (ou mais) variáveis simultaneamente.

Definições e propriedades

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

$$
\begin{aligned}
\mathcal{E}(X) &= \mathcal{E}(\mathcal{E}(X|Y,Z)). \\
\mathcal{V}(X) &= \mathcal{E}(\mathcal{V}(X|Y,Z)) + \mathcal{V}(\mathcal{E}(X|Y,Z)).
\end{aligned}
$$

- Dem: Exercício.

Definições e propriedades

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

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

- Temos que $X \perp Y \perp Z \leftrightarrow F_{\boldsymbol{W}}(x, y, z) = F_X(x) F_Y(y) F_Z(z), \forall x, y, z$.
- Pesquisar sobre as propriedades equivalentes ao caso bivariado hrefaqui.

Exemplo

- Considere, novamente, 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.
  - Z: a cor da primeira bola sorteada em que 1 se a bola foi B e 0, se foi V.
- Defina $\boldsymbol{W} = (X, Y, Z)'$. 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 $\boldsymbol{W}$, conseguimos verificar que:

$$A(:\Omega \rightarrow \mathcal{R}^3) = \{\{2,0,1\},\{1,1,1\},\{1,0,0\},\{0,1,0\}\}.$$

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

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

(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:

	Z				
	0		1		
	X		X		
Y	0	1	0	1	2
0	0	3/10	0	0	1/10
1	3/10	0	0	3/10	0

(Cont.) Exemplo

- Notemos que:

$$
\begin{aligned}
P(X = 0) &= \sum_{y,z} P(X = 0, Y = y, Z = z) = \frac{3}{10} \\
P(X = 1) &= \sum_{y,z} P(X = 1, Y = y, Z = z) = \frac{6}{10} \\
P(X = 2) &= \sum_{y,z} P(X = 2, Y = y, Z = z) = \frac{1}{10}
\end{aligned}
$$

(Cont.) Exemplo

- Notemos que:

$$
\begin{aligned}
P(Y=0) &= \sum_{x,z} P(X=x, Y=1, Z=z) = \frac{4}{10} = \frac{2}{5} \\
P(Y=1) &= \sum_{x,z} P(X=x, Y=1, Z=z) = \frac{6}{10} = \frac{3}{5} \\
P(Z=0) &= \sum_{y,z} P(X=x, Y=y, Z=0) = \frac{6}{10} = \frac{3}{5} \\
P(Z=1) &= \sum_{y,z} P(X=x, Y=y, Z=1) = \frac{4}{10} = \frac{2}{5}
\end{aligned}
$$

(Cont.) Exemplo

- Notemos que:

	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

- Além disso:

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

(Cont.) Exemplo

- Além disso:

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