Variáveis Aleatórias Multidimensionais

Prof. José Roberto Silva dos Santos

Fortaleza, 09 de novembro de 2022

Variáveis aleatórias multidimensionais

- Em muitas situações práticas há o interesse na descrição probabilística de mais de uma característica numéricas, associada a um experimento aleatório.

- Por exemplo, na distribuição de alturas e pesos de indivíduos de uma determinada população.

- Portanto, é natural a extensão do conceito de variáveis aleatórias para mais de uma dimensão.

- Suponha que estamos interessados em estudar a composição de famílias com três crianças, quanto ao sexo. Definamos:

X = número de meninos,

$$Y = \begin{cases} 1, \text{ se o primeiro filho for homem} \\ 0, \text{ se o primeiro filho for mulher,} \end{cases}$$

Z = número de vezes em que houve variação do sexo entre um nascimento e outro, dentro da mesma família.

- A seguir temos os possíveis valores para as três variáveis aleatórias.

Figura 1: Composição de famílias com três crianças, quanto ao sexo.

Eventos	Probabilidade	X	Y	Z
HHH	1/8	3	1	0
HHM	1/8	2	1	1
HMH	1/8	2	1	2
MHH	1/8	2	0	1
HMM	1/8	1	1	1
MHM	1/8	1	0	2
MMH	1/8	1	0	1
MMM	1/8	0	0	0

Figura 2: Distribuições de probabilidade unidimensionais.

(a)

x	0	1	2	3
$p(x)$	1/8	3/8	3/8	1/8

(b)

y	0	1
$p(y)$	1/2	1/2

(c)

z	0	1	2
$p(z)$	1/4	1/2	1/4

- A seguir temos as probabilidades associadas aos pares de valores nas variáveis X e Y, denominada *distribuição conjunta* de X e Y.

Figura 3: Distribuição conjunta da v.a (X, Y).

(x, y)	$p(x, y)$
(0, 0)	1/8
(1, 0)	2/8
(1, 1)	1/8
(2, 0)	1/8
(2, 1)	2/8
(3, 1)	1/8

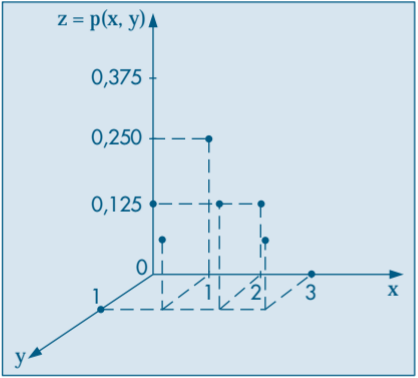
- Nesse caso, também é possível utilizar uma tabela de dupla entrada da seguinte forma:

Figura 4: Distribuição conjunta da v.a (X, Y).

$Y \backslash X$	0	1	2	3	$p(y)$
0	1/8	2/8	1/8	0	1/2
1	0	1/8	2/8	1/8	1/2
$p(x)$	1/8	3/8	3/8	1/8	1

- Podemos ver, então, as probabilidades conjuntas $p(x, y)$ e as probabilidades marginais $p(x)$ e $p(y)$.

Figura 5: Representação gráfica da distribuição da variável aleatória (X, Y).



- Os conjuntos de pares $(x, p(x))$ e $(y, p(y))$ são denominadas *distribuições marginais* das v.a's X e Y, respectivamente.
- Qual seria a distribuição do número de meninos, sabendo-se que o primeiro filho é do sexo masculino? Da definição de probabilidade condicional obtemos

$$\mathbb{P}(X = x|Y = 1) = \frac{\mathbb{P}(X = x; Y = 1)}{\mathbb{P}(Y = 1)} = p(x|Y = 1)$$

para todo $x = 0, 1, 2, 3$.

- Tal distribuição é denominada *distribuição condicional*.

Figura 6: Distribuição condicional de X, dado que $Y = 1$.

x	1	2	3
$p(x\|Y=1)$	1/4	1/2	1/4

Figura 7: Distribuição condicional de Y, dado que $X = 2$.

y	0	1
$p(y\|X=2)$	1/3	2/3

Definição:

Seja x um valor da v.a X, tal que $p(x) = \mathbb{P}(X = x) > 0$. A probabilidade

$$p(y|x) = \mathbb{P}(Y = y|X = x) = \frac{\mathbb{P}(Y = y; X = x)}{\mathbb{P}(X = x)} \text{ para todo } x$$

é denominada probabilidade de $Y = y$ dado que $X = x$ e, conjunto de pares $(y, p(y|x))$ é denominado *distribuição condicional* de Y dado que $X = x$.

- Considere a distribuição condicional de X, dado que $Y = 1$. Podemos determinar a esperança dessa distribuição, ou seja,

$$\mathbb{E}(X|Y=1) = 1 \times \frac{1}{4} + 2 \times \frac{1}{2} + 3 \times \frac{1}{4} = 2$$

- Observe que $\mathbb{E}(X) = 1,5$, ao passo que $\mathbb{E}(X|Y=1) = 2$.
- De modo geral, no caso discreto, a esperança condicional de X dado que $Y = y$ é definida por

$$\mathbb{E}(X|Y=y) = \sum_x x\mathbb{P}(X=x|Y=y) \text{ para qualquer } y$$

- Considere agora a distribuição conjuntas das v.a's Y e Z

Figura 8: Distribuição conjunta de Y e Z.

Y \ Z	0	1	2	$p(y)$
0	1/8	2/8	1/8	1/2
1	1/8	2/8	1/8	1/2
$p(z)$	1/4	2/4	1/4	1

- Nesse caso, observamos que

$$\mathbb{P}(Z = z|Y = y) = \frac{\mathbb{P}(Z = z; Y = y)}{\mathbb{P}(Y = y)} = \mathbb{P}(Z = z)$$

para quaisquer z e y. O que significa dizer que

$$\mathbb{P}(Z = z; Y = y) = \mathbb{P}(Z = z)\mathbb{P}(Y = y)$$

- Também é verdade que

$$\mathbb{P}(Y = y|Z = z) = \mathbb{P}(Y = y)$$

qualquer que seja y e z. Neste caso, dizemos Y e Z são independentes.

Definição:

As variáveis aleatórias X e Y discretas são independentes se, e somente se, para quaisquer pares (x, y) tivermos

$$\mathbb{P}(X = x; Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y).$$

Variáveis aleatórias multidimensionais

Definição 1

- Sejam X e Y variáveis aleatórias. Denominaremos o vetor (X, Y) uma *variável aleatória bidimensional*.
- Se $X_1, X_2, \ldots, X_n$ forem n variáveis aleatórias, denominaremos $(X_1, \ldots, X_n)$ uma *variável aleatória n-dimensional* (ou vetor aleatório).

Definição 2

(a) Seja (X, Y) uma variável aleatória discreta bidimensional. A cada resultado possível (x, y) associaremos um número $p(x, y)$ representando $\mathbb{P}(X = x, Y = y)$ e satisfazendo às seguintes condições:

1. $p(x, y) \geq 0$ para todo (x, y),
2. $\sum_{x=1}^{\infty} \sum_{y=1}^{\infty} p(x, y) = 1$.

(b) Seja (X, Y) uma variável aleatória contínua tomando todos os valores em alguma região B do plano euclidiano. A *função densidade de probabilidade conjunta* f é uma função que satisfaz às seguintes condições:

1. $f(x, y) \geq 0$ para todo $(x, y) \in B$
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$.

Definição 3

(a) Seja $(X_1, X_2, \ldots, X_n)$ uma variável aleatória discreta n-dimensional. A cada resultado possível associaremos um número $p(x_1, x_2, \ldots, x_n)$ representando $\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n)$ e satisfazendo às seguintes condições:

1. $p(x_1, x_2, \ldots, x_n) \geq 0$ para todo $(x_1, \ldots, x_n)$,
2. $\sum_{x_1=1}^{\infty} \cdots \sum_{x_n=1}^{\infty} p(x_1, x_2, \ldots, x_n) = 1$.

(b) Seja $(X_1, \ldots, X_n)$ uma variável aleatória contínua tomando todos os valores em alguma região $B \subset \mathbb{R}^n$. A *função densidade de probabilidade conjunta* f é uma função que satisfaz às seguintes condições:

1. $f(x_1, \ldots, x_n) \geq 0$ para todo $(x_1, \ldots, x_n) \in B$
2. $\int_B \cdots \int f(x_1, \ldots, x_n) dx_1 \cdots dx_n = 1$.

- Suponha que um fabricante de lâmpadas esteja interessado no número de lâmpadas encomendadas a ele durante os meses de janeiro e fevereiro. Sejam X e Y, respectivamente, o número de lampâdas encomendadas durante esses dois meses. Admitiremos que (X, Y) seja uma variável aleatória contínua bidimensional, com a seguinte f.d.p conjunta

$$f(x, y) = c \underset{[5.000,10.000]}{I}(x) \underset{[4.000,9.000]}{I}(y)$$

(a) Determine o valor de c.
(b) Qual a probabilidade de que as vendas de fevereiro sejam no máximo igual as vendas de janeiro?

- Seja (X, Y) uma variável aleatória bidimensional com f.d.p conjunta definida por

$$f(x,y) = \left(x^2 + \frac{xy}{3}\right) \underset{[0,1]}{I}(x) \underset{[0,2]}{I}(y)$$

  (a) Demostre que $f(x,y)$ é de fato uma f.d.p conjunta.
  (b) Determine a probabilidade do evento $B = \{X + Y \geq 1\}$.

Distribuições Marginais

Definição:

- Seja (X, Y) um vetor aleatório bidimensional com f.d.p conjunta $f(x, y)$ as funções densidade de probabilidade marginais de X e Y são dados por

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy; \quad f_Y(y) = \int_{-\infty}^{\infty} f(x, y)dx$$

- No caso multidimensional para determinar a distribuição marginal da k-ésima componente, é necessário integrar $f(x_1, x_2, \ldots, x_n)$ para todas as variáveis, exceto x_i. Ou seja

$$f_{X_k}(x) = \int_B \cdots \int f(x_1, \ldots, x_n)dx_1 \cdots dx_n \text{ para todo } i \neq k.$$

- Seja (X, Y, Z) um vetor aleatório com densidade conjunta definida por

$$f(x,y,z) = kxy^2z \underset{(0,1)}{I}(x) \underset{(0,1)}{I}(y) \underset{(0,\sqrt{2})}{I}(z)$$

(a) Determine o valor de k.
(b) Determine as distribuições marginais de X, Y e Z.