Distribuição de Probabilidade Conjunta

1. DISTRIBUIÇÃO DE PROBABILIDADE CONJUNTA

O nosso estudo de variável aleatória e de suas funções de probabilidade até agora se restringiram a espaços amostrais unidimensionais nos quais os valores observados eram assumidos por uma única v. a.

Entretanto, existem situações em que se deseja observar resultados simultâneos de várias variáveis aleatórias. Por exemplo, podemos medir o total precipitado, a umidade e a temperatura, resultado em um espaço amostral tri-dimensional que consiste nos resultados (p, u, t).

1.1. VARIAÇÕES ALEATÓRIAS DISCRETAS

Se x e y são duas variáveis aleatórias, a distribuição de probabilidades de sua ocorrência simultânea pode ser representada pela função com valores f(x, y) para qualquer par de valores (x, y).Costuma-se referir a esta função como Distribuição de Probabilidade Conjunta de x e y.

Para o caso discreto:

$F(x, y) – P (X = x, Y = y) \quad = \quad f.m.p.$

ou seja, os valores f(x, y) dão a probabilidade dos resultados x e y ocorrerem ao mesmo tempo.

A função f(x, y) é a <u>distribuição de probabilidade conjunta</u> ou f.m.p. das variáveis aleatórias x e y se:

1. $f(x, y) \geq 0$ p/ todos (x, y) (claro! É probabilidade!).
2. $\sum_x \sum_y f(x, y) = 1$
3. P (X = x, Y = y) = f(x,y)

Para qualquer região A no plano xy

$$P[(x, y) \in A] = \sum_A \sum f(x, y)$$

Exercício 7.1

Dois refils selecionados aleatoriamente

3B
2R
3G

$\left.\begin{array}{l}\text{X – número de “blue” refils}\\ \text{Y – número de “red” refils}\end{array}\right\}$ selecionados

a) Ache a probabilidade conjunta f(x, y), ou seja a probabilidade de x e y ocorrerem simultaneamente.

S = { (G, G) (G, B) (G, R) (R, R) (R, G) (R, B) (B, B) (B, G) (B, R) }

- Quais os valores assumidos pelas variáveis aleatórias X e Y nestes pontos do espaço amostral?

(x, y)

x = 0	x = 1	x = 0	x = 0	x = 0	x = 1	x = 2	x = 1	x = 1
y = 0	y = 0	y = 1	y = 2	y = 1	y = 1	y = 0	y = 0	y = 1

Assim, os possíveis pares de valores simultâneos de (x, y) são:

(0, 0) (1, 0) (0,1) (0, 2) (2, 0) (1, 1)

- Calculando as probabilidades:

f (0,0) = 2 “GREENS” serem selecionados.

$$P(x=0, y=0) = \frac{\binom{3}{2}}{\binom{8}{2}} = \frac{\frac{3!}{2!\,1!}}{\frac{8!}{2!\;6!}} = \frac{\frac{3\ \text{x}\ 2!}{2!\ .\ 1!}}{\frac{8\ \text{x}\ 7\ \text{x}\ 6!}{2!\ .\ 6!}} = \frac{3}{\frac{56}{2}} = \frac{3}{28}$$

f (1,0) = 1 BLUE ser selecionado

1 GREEN ser selecionado

$$= \frac{\binom{3}{1}\binom{3}{1}}{\binom{8}{2}} = \frac{\frac{3!}{1!\ \ 2!} \cdot \frac{3!}{1!\ \ 2!}}{28} = \frac{\frac{3\ \text{x}\cancel{2!}}{1!\ \cancel{2!}} \cdot \frac{3\ \text{x}\cancel{2!}}{1!\ \cancel{2!}}}{28} = \frac{9}{28}$$

f (0, 1) = 1 RED ser selecionado

1 GREEN ser selecionado

$$= \frac{\binom{2}{1}\binom{3}{1}}{\binom{8}{2}} = \frac{\frac{2!}{1!\ 1!} \cdot \frac{3!\ \text{x}\cancel{2!}}{2!\ \cancel{1!}}}{28} = \frac{6}{28}$$

f (0, 2) = 2 RED serem selecionados

$$= \frac{\binom{2}{2}}{\binom{8}{2}} = \frac{\frac{\cancel{2!}}{\cancel{2!}\ 1!}}{28} = \frac{1}{28}$$

f (2, 0) = 2 “BLUES” serem selecionados

$$= \frac{\binom{3}{2}}{\binom{8}{2}} = \frac{\frac{3!}{2!\ 1!}}{28} = \frac{\frac{3!\ x\ \cancel{2!}}{\cancel{2!}\ 1!}}{28} = \frac{3}{28}$$

f (1,1) = 1 BLUE ser selecionado

1 RED ser selecionado

$$= \frac{\binom{3}{1}\binom{2}{1}}{\binom{8}{2}} = \frac{\frac{3!}{1!\ \cancel{2}} \cdot \frac{\cancel{2!}}{1!\ 1!}}{28} = \frac{6}{28}$$

Assim, a f.m.p. do acontecimento simultâneo de x e y é dado por:

Tabela 7.1

f (x, y)		x			
		0	1	2	
y	0	3/28	9/28	3/28	15/21
	1	6/28	6/28	--	12/28
	2	1/28	--	--	1/28
		10/28	15/28	3/28	1

Σ = 1

Σ = 1

b)

$$P\left[(x, y) \in A \ / \ x + y \leq 1\right] = P(x + y \leq 1)$$

$$= f(0, 0) + f(1, 0) + f(0, 1)$$

$$= \frac{3}{28} + \frac{9}{28} + \frac{6}{28} = \frac{18}{28}$$

1.2. VARIÁVEL ALEATÓRIA CONTÍNUA

A função $f(x, y)$ é uma função densidade de probabilidade conjunta das variáveis aleatórias x e y se:

1. $f(x, y) \geq 0$ p/ todos (x, y)
2. $\int_{-\infty}^{+\infty} \int_{-\infty}^{\infty} f(x, y) \, dx \, dy = 1$
3. $P[(X, Y) \in A] = \int \int f(x, y) \, dx \, dy$ para qualquer região A no plano xy.

Exercício 7.2

Uma fábrica de doces distribuiu caixas de chocolates com mistura de creme, toffees e amêndoas, envolta em chocolate branco e marrom. Para uma caixa selecionada ao acaso, seja x e y, respectivamente, a proporção de chocolate branco e marrom existente no creme e suponha que f.d.p. conjunta é:

$$f(x, y) = \begin{cases} \frac{2}{5}(2x + 3y) & 0 \leq x \leq 1, \ 0 \leq y \leq 1 \\ 0, & \text{outros valores} \end{cases}$$

a) verifique a propriedade 2 (é f.d.p. ?)

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\ dx\ dy = \int_0^1\int_0^1 \frac{2}{5}(2x+3y)\,dx\,dy$$

⇒ integra em relação a "x"; depois em relação a "y".

$$= \frac{2}{5}\int_0^1 \left(\int_0^1 2x\,dx + \int_0^1 3y\,dx \right) dy$$

$$= \frac{2}{5}\int_0^1 \left(2\frac{x^2}{2}\Big|_0^1 + 3yx\Big|_0^1 \right) dy$$

$$= \frac{2}{5}\int_0^1 (1+3y)\,dy$$

$$= \frac{2}{5}\left[\int_0^1 dy + \int_0^1 3y\,dy\right]$$

$$= \frac{2}{5}\left[y\Big|_0^1 + \frac{3y^2}{2}\Big|_0^1 \right]$$

$$= \frac{2}{5}\left[1+\frac{3}{2}\right] = \frac{2}{5}\left[\frac{2+3}{2}\right] = 1 \quad \text{c.q.d!}$$

b) $P[(x,y)\in A]$ onde A é região definida por $\left\{(x,y) \,\middle|\, 0<x<\frac{1}{2}, \frac{1}{4}<y<\frac{1}{2}\right\}$

2. DISTRIBUIÇÃO DE PROBABILIDADE MARGINAL

Dada uma função de probabilidade conjunta f(x, y) das variáveis aleatórias discretas X e Y, a distribuição de probabilidade de X isolado g(x) é obtida pela soma dos valores de f(x, y) ao longo de Y. Do mesmo modo, a distribuição de probabilidade de Y isolado h(y) é dada pela soma dos valores de f(x, y) ao longo de x.

Definimos g(x) e h(y) como sendo as distribuições de probabilidades marginais de x e y, respectivamente.

Assim,

A distribuição de probabilidade de X isolado e Y isolados são:

$g(x) = \sum_{y} f(x, y)$ e $h(y) = \sum_{x} f(x, y)$ variável discreta

$g(x) = \int_{-\infty}^{+\infty} f(x, y)\, dy$ e $h(y) = \sum_{x} f(x, y)$ variável contínua

Exercício 7.3

Dada a f.m.p. conjunta de X e Y dada no exercício 2:

Tabela 7.2

		X			
f(x, y)		0	1	2	
Y	0	3/28	9/28	3/28	15/28
	1	6/28	6/28	-	12/28
	2	1/28	-	-	1/28
		10/28	15/28	3/28	1

$\rightarrow$ Mostre que o somatório de cada coluna dá a distribuição de probabilidade marginal de x.

$$P(X = 0) = g(0) = \sum_{y=0}^{2} f(0, y) = f(0, 0) + f(0, 1) + f(0, 2) = 10/28$$

$$P(X = 1) = g(1) = \sum_{y=0}^{2} f(1, y) = f(1, 0) + f(1, 1) + f(1, 2) = 15/28$$

$$P(X = 2) = g(2) = \sum_{y=0}^{2} f(2, y) = f(2, 0) = 3/28$$

ou seja:

Tabela 7.3

x	0	1	2
g(x)	$\frac{10}{28}$	$\frac{15}{28}$	$\frac{3}{28}$

= distribuição de probabilidade marginal de X.

Exercício 7.4

Ache g(x) para a distribuição de probabilidade conjunta:

$$f(x,\ y) = \begin{cases} \frac{2}{5}(2x+3y) & 0 \le x \le 1,\ \ 0 \le y \le 1 \\ 0 & \text{outro valor} \end{cases}$$

Exercício 7.5

Ache g (x) para a distribuição de probabilidade conjunta:

$$f(x,\ y) = \ x, y) = \begin{cases} \frac{2}{5}(2x+3y) & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0 & \text{outro valor} \end{cases}$$

Solução:

Por definição a distribuição de probabilidade marginal de x é dada por:

$$g(x) \ = \int_{-\infty}^{+\infty} f(x,\ y)\, dy$$

$$= \int_0^1 \frac{2}{5}(2x+3y)\,dy = \frac{2}{5}\left[2xy\Big|_9^1 + \frac{3y^2}{2}\Big|_0^1\right]$$

$$= \frac{2}{5}\left[2x+\frac{3}{2}\right] = \frac{4x}{5}+\frac{3}{5}$$

$$\therefore\ g(x) = \frac{4x+3}{5}$$

ou seja:

$$\therefore g(x)\begin{cases} \dfrac{4x+3}{5} & 0 \le x \le 1 \\ 0 & \end{cases}$$

Exercício 7.6

Ache h(y) para o exemplo anterior (distribuição de marginal de y) por definição:

$$h(y) = \int_{-\infty}^{+\infty} f(x,y)dx = \int_0^1 (2x+3y)dx$$

$$= \frac{2}{5}\left[\frac{2x^2}{2}\Big|_0^1 + 3yx\Big|_0^1\right] = \frac{2}{5}[1+3y]$$

$$= \frac{2+6y}{5}$$

ou seja:

$$h(y)\begin{cases} \dfrac{2+6y}{5} & p/\ 0 \le y \le 1 \\ 0 & \end{cases}$$

3. DISTRIBUIÇÃO DE PROBABILIDADES CONDICIONAL

Sabemos que:

$$P(B/A) = \frac{P(A \cap B)}{P(A)} \quad , \ P(A) > 0$$

Definindo:

A → o evento onde X = x

B → o evento onde Y = y

Temos que:

$$P(Y = y \,/\, X = x) = \frac{P(X = x, Y = y)}{P(X = x)} = \frac{f(x, y)}{9(x)}, \quad 9(x) > 0$$

ou ainda

$$P(X = x \,/\, Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{f(x, y)}{h(y)}, \quad h(y) > 0$$

Exercício 7.7

Referindo-se ao exemplo 2 (refis), ache a distribuição de probabilidade condicional de X dado que Y = 1 e use isto para calcular $P(X = 0/Y = 1)$.

Tabela 7.2 do exemplo ② é:

		X			
f(x, y)		0	1	2	
Y	0	3/28	9/28	3/28	12/28
	1	6/24	6/24	-	
	2	1/28	-	-	

a) Quero:

Tabela 7.3

x	0	1	2
f(x/1)	$\frac{1}{2}$	$\frac{1}{2}$	0

Saberia fazer pelos nossos conhecimentos anteriores de probabilidade condicional:

$$P(x = 0/y = 1) = \frac{6/28}{12/28} = \frac{1}{2}$$

$$P(x = 1/y = 1) = \frac{6/28}{12/28} = \frac{1}{2}$$

$$P(x = 2 / y = 1) = \frac{0}{12/28} = 0$$

Mas usando a definição:

$$P(X = x / y = 1)\frac{f(x, y)}{h(y)}$$

por definição: $h(y) = \sum_{x} f(x, y)$

$$\therefore h(1) = \sum_{x} f(x,1) = \frac{6}{28} + \frac{6}{28} + 0 = \frac{12}{28} = \frac{3}{7}$$

Assim,

$$P(x = 0 / y = 1)\frac{f(0,1)}{h(1)} = \frac{7}{3} f(0,0) = \frac{7}{3} - \frac{6}{28} = \frac{1}{2}$$

$$P(x = 1 / y = 1)\frac{f(1,1)}{h(1)} = \frac{7}{3} f(1,1) = \frac{7}{3} - \frac{6}{28} = \frac{1}{2}$$

$$P(x = 2 / y = 1)\frac{f(2,1)}{h(1)} = \frac{7}{3} f(2,1) = 0$$

Exercício 7.8

Ache A distribuição de probabilidade das variáveis aleatórias X e Y onde:

X - mudança unitária da temperatura

y - mudança unitária da profundidade de um lago

a) Ache as distribuições de probabilidades marginais de X e Y e a distribuição de probabilidade condicional f(y, x:)

1) $$9(x) = \int_{-\infty}^{+\infty} f(x, y)dy = \int_{X}^{1} 10xy^2 dy$$

$$= \frac{10xy^3}{3}\Big|_x^1 = \frac{10x}{3}\left(1 - x^3\right) = \frac{10x - 10x^4}{3}$$

$$\boxed{9(x)=\frac{10x}{3}\Big|_x^1=\frac{10x}{3}\left(1-x^3\right), \qquad 0<x<1}$$

2) $h(y)=\int_{-\infty}^{+\infty} f(x,y)dx=\int_{x=0}^{x=y}10xy^2dx=\frac{10y^2.x^2}{2}\Big|_{x=0}^{x=y}$

$=5y^2\left(y^2-0\right)=5y^4$

$$\boxed{h(y)=5y^4 \quad , \qquad 0<y<1}$$

3) $f(y/x)$

por definição $f(y/x)=\frac{f(x,y)}{9(x)}=\frac{10xy^2}{\frac{10}{3}x\left(1-x^3\right)}$

$$\boxed{=\frac{3.10xy^2}{10x\left(1-x^3\right)}=\frac{3y^2}{\left(1-x^3\right)} \quad , \qquad 0<x<y<1}$$

b) $P\left(y>\frac{1}{2}\Big|x=0,25\right)$

→ acabamos de achar a f.d.p. da distribuição de probabilidade condicional.

Por definição:

$$\boxed{P\left(a<y<b/X=x\right)=\int_a^b f\left(x/y\right).dy}$$

Assim,

$P\left(a<y<b/X=x\right)=\int_a^b f\left(y/x\right).dy$

$$=\int_{1/2}^1\frac{3y^2}{\left(1-x^3\right)}dy=\frac{3}{\left(1-0,25^3\right)}\int_{1/2}^1 y^2dy=\frac{3}{\left(1-0,25^3\right)}\frac{y^3}{3}\Big|_{1/2}^1$$

$$=\frac{1}{\left(1-0,25^3\right)}\left(1-\frac{1}{8}\right)=\left(\frac{1-1/8}{1-1/64}\right)=\frac{\frac{8-1}{8}}{\frac{64-1}{64}}$$

$$= \frac{\frac{7}{8}}{\frac{63}{64}} = \frac{7}{8} \cdot \frac{\overset{8}{64}}{63} = \frac{53}{63} = \frac{8}{9}$$

Exercício 7.9

Dada a f.d.p. conjunta $f(x,y) = \begin{cases} \frac{x(1+3y^2)}{4} & 0 < x < 2,\ 0 < y < 1 \\ 0 & \text{outro valor} \end{cases}$

Ache g(x), h(y), f(x/y) e $P\left(\frac{1}{4} < X < \frac{1}{2} \middle|\ y = \frac{1}{3}\right)$.

Solução:

a) g(x)

Por definição:

$$g(x) = \int_{-\infty}^{+\infty} f(x,y).dy$$

$$g(x) = \int_{y=0}^{y=1} \left(\frac{x}{4} + \frac{3xy^2}{4}\right) dy = \int_{y=0}^{y=1} \frac{x}{4} dy + \int_{y=0}^{y=1} \frac{3}{4} xy^2 dy$$

$$g(x) = \frac{x}{4} y \Big|_0^1 + \frac{3}{4} x \frac{y^3}{3} \Big|_0^1 = \frac{x}{4} + \frac{x}{4} = \frac{x}{2}$$

$$g(x) = \begin{cases} \frac{x}{2} & ,\ 0 < x < 2 \\ 0 & \text{outro valor} \end{cases}$$

b) h(y)

Por definição:

$$h(y) = \int_{-\infty}^{+\infty} f(x, y).dx = \int_{x=0}^{x=2} \frac{x}{4}\,dx + \int_{0}^{2} \frac{3}{4}\,xy^2 dy$$

$$h(y) = \left.\frac{x^2}{8}\right|_0^2 + \frac{3}{4}\,y^2 \left.\frac{x^2}{2}\right|_0^2$$

$$h(y) = \frac{4}{8} + \frac{3}{8}\,y^2\,4 = \frac{1}{\frac{8}{2}}\left(\frac{1}{4} + \frac{3}{12}\,y^2\right) =$$

$$h(y) = \begin{cases} \dfrac{1+3y^2 2}{2} & , \ 0 < y < 1 \\ 0 & \text{outro valor} \end{cases}$$

c) f (x, y)

Por definição:

$$f\,(x, y) = \frac{f(x, y)}{h.(y)}$$

$$\frac{\dfrac{x\cancel{(1+3y^2)}}{4^2}}{\dfrac{\cancel{(1+3y^2)}}{2}} = \frac{x}{2}$$

Assim,

$$f(x, y) = x, y) = \begin{cases} \dfrac{x}{2} & , \ 0 < x < 2 \\ 0 & \text{outro valor} \end{cases}$$

d) $P\left(\frac{1}{4} < x < \frac{1}{2} \middle| Y = y\right) = \int_a^b f(x|y)dx$

$$\int_{1/4}^{1/2} \frac{x}{2}dx = \frac{x^2}{4}\Big|_{1/4}^{1/2} = \frac{1}{4}\left[\frac{1}{4} - \frac{1}{16}\right] = \frac{1}{4}\left(\frac{4-1}{16}\right)$$

$$= \frac{3}{64}$$

4. INDEPENDÊNCIA ESTATÍSTICA

(Dedução com analogia Teoria das Probabilidades)

Sabemos que:

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \Rightarrow f(x/y) = \frac{f(x,y)}{h(y)}$$

Mas se A e B forem independentes:

$$P(A/B) = P(A) \Rightarrow f(x/y) = g(x)$$

Assim,

$$P(A \cap B) = P(A) \,.\, P(B) \Rightarrow f(x,y) = g(x).h(y)$$

Sejam x e y duas variáveis aleatórias (discretas ou contínuas) com distribuição de probabilidade conjunta f (x, y) e distribuição de probabilidade marginais g(x) e h(y), respectivamente. As variáveis aleatórias x e y são consideradas Estatisticamente Independentes se e somente se:

$$f(x, y) = g(x) \,.\, h(y)$$

para todos (x, y) dependendo do intervalo).

Exercício 7.10

Mostre que as variáveis aleatórias do exemplo ① não são estatisticamente independentes.

$$f(x, y) \overset{?}{=} g(x) \cdot h(x)$$

Tabela 7.4

		X			
f(x, y)		0	1	2	
Y	0	3/28	9/28	3/28	15/28
	1	6/24	6/24	-	12/28
	2	1/28	-	-	1/28
		10/28	15/28	3/28	

Vamos verificar em um par (x, y)

Suponha (0, 0)

f (x, y) = f (0, 0) = 3/28

9 (0) = 10/28

h (0) = 15/28

<u>Vemos que:</u>

$$\frac{3}{28} \neq \frac{10}{28} \cdot \frac{15}{28}$$ **não são estatisticamente independentes.**

Generalização

Seja $X_1, X_2, \ldots X_n$, n variáveis aleatórias (discretas ou contínuas) com distribuição de probabilidade conjunta $f(x_1, x_2 \ldots x_n)$ e distribuição de probabilidade marginais $f_1(x_1)$, $f_2(x_2) \ldots f_3(x_n)$, respectivamente. As variáveis aleatórias, $X_1, X_2, \ldots X_n$, são ditas <u>Estatisticamente Independentes</u> se e somente se:

$$f(x_1, x_2 \ldots x_n) = f_1(x_1) \cdot f_2(x_2) \ldots f_3(x_n)$$