Variáveis Aleatórias Bidimensionais

Revisão - Capítulo 5 - Magalhães e Lima

Exemplo 5.5 - Pag 142

Região dividida em 10 sub-regiões.

X - número de poços artesianos

Y - Número de rios presentes na sub-região

Sub-região	1	2	3	4	5	6	7	8	9	10
X	0	0	0	0	1	2	1	2	2	0
Y	1	2	1	0	1	0	0	1	2	2

Selecionando-se uma sub-região ao acaso, a distribuição de probabilidades conjunta de (X, Y) é

(X, Y)	Probabilidade $p(x, y)$
$(0,0)$	$1/10$
$(0,1)$	$2/10$
$(0,2)$	$2/10$
$(1,0)$	$1/10$
$(1,1)$	$1/10$
$(2,0)$	$1/10$
$(2,1)$	$1/10$
$(2,2)$	$1/10$
	$1,0$

$p(x,y) = P(X = x, Y = y)$

$x = 0, 1, 2$

$y = 0, 1, 2$

Distribuição de Probabilidades Conjunta

$X \backslash Y$	0	1	2	$P(X = x)$
0	$1/10$	$2/10$	$2/10$	$5/10$
1	$1/10$	$1/10$	0	$2/10$
2	$1/10$	$1/10$	$1/10$	$3/10$
$P(\cancel{X} = y)$	$3/10$	$4/10$	$3/10$	

distribuição marginal de X

distribuição marginal de Y

$$P(X = 0) = P(X = 0, Y = 0) + P(X = 0, Y = 1) + P(X = 0, Y = 2)$$

$$= \frac{1}{10} + \frac{2}{10} + \frac{2}{10} = \frac{5}{10}$$

Distribuição Marginal de X

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

$p_i = P(X = x_i)$

$i = 1, 2, 3$

$$E(X) = \sum_{i=1}^{3} x_i p_i = 0.\frac{5}{10} + 1.\frac{2}{10} + 2.\frac{3}{10} = \frac{8}{10} = \frac{4}{5} = \mu$$

$$Var(X) = \sum_{i=1}^{3} (x_i - \mu)^2 p_i = \left(0 - \frac{4}{5}\right)^2 \cdot \frac{5}{10} + \left(1 - \frac{4}{5}\right)^2 \cdot \frac{2}{10}$$

$$+ \left(2 - \frac{4}{5}\right)^2 \cdot \frac{3}{10} = \frac{16}{25} \cdot \frac{5}{10} + \frac{1}{25} \cdot \frac{2}{10} + \frac{36}{25} \cdot \frac{3}{10} = \frac{190}{250} = \frac{19}{25}$$

Fórmula alternativa para Var(X)

$$Var(X) = E(X^2) - E^2(X)$$

Analogamente $E(Y) = 1$ $\quad Var(Y) = \frac{60}{100}$

Definição

As variáveis aleatórias X e Y são independentes

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

No exemplo, X e Y não são independentes pois, por ex.

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

No exemplo, a partir da distribuição de (X, Y), obter a distribuição da soma e do produto de X e Y

(X,Y)	$p(x,y)$	$X+Y$	XY
(0,0)	1/10	0	0
(0,1)	2/10	1	0
(0,2)	2/10	2	0
(1,0)	1/10	1	0
(1,1)	1/10	2	1
(2,0)	1/10	2	0
(2,1)	1/10	3	2
(2,2)	1/10	4	4

Dist. de Prob. de X+Y

$X+Y$	prob
0	1/10
1	3/10
2	4/10
3	1/10
4	1/10

Dist. de Prob de XY

XY	prob
0	7/10
1	1/10
2	1/10
4	1/10

Das distribuições de probabilidades de X+Y e de XY,

$$E(X+Y) = 0.\frac{1}{10} + 1.\frac{3}{10} + 2.\frac{4}{10} + 3.\frac{1}{10} + 4.\frac{1}{10} = \frac{18}{10}$$

$$E(XY) = 0.\frac{7}{10} + 1.\frac{1}{10} + 2.\frac{1}{10} + 4.\frac{1}{10} = \frac{7}{10}$$

Observa-se que

$$E(X+Y) = \frac{18}{10} = \frac{4}{5} + 1 = E(X) + E(Y)$$

vale sempre

$$E(XY) = \frac{7}{10} \neq E(X)E(Y)$$

se X e Y são v.a. independentes vale a igualdade

<u>Resultado 1</u>

Sejam X e Y variáveis aleatórias quaisquer, então

$$E(X+Y) = E(X) + E(Y).$$

Resultado 2

Se X e Y são independentes então $E(XY) = E(X)E(Y)$.

No entanto

$$E(XY) = E(X)E(Y) \not\Rightarrow X \text{ e } Y \text{ são independentes}$$

Definição

Dada a v.a. bidimensional (X, Y) define-se a covariância entre X e Y como

$$Cov(X, Y) = E\left[(X - \mu_X)(Y - \mu_Y)\right].$$

↳ Valor esperado do produto dos desvios de cada variável em relação à sua média.

Prova-se que $Cov(X, Y) = E(XY) - E(X)E(Y)$.

Se grandes valores de X tendem a ocorrer com grandes valores de Y e pequenos valores de X com pequenos valores de Y então o produto $(X - \mu_X)(Y - \mu_Y)$ tende a ser positivo em média e $Cov(X, Y) > 0$. Caso contrário, $Cov(X, Y) < 0$.

Resultado 3

Se X e Y são variáveis aleatórias independentes então $Cov(X,Y) = 0$.

$Cov(X,Y) = 0 \not\Rightarrow$ X e Y independentes

Resultado 4

$$Var(X+Y) = Var(X) + Var(Y) + 2\,Cov(X,Y)$$

No exemplo

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$

$$= \frac{7}{10} - \frac{4}{5} \times 1 = -\frac{1}{10}$$

$Var(X+Y) = \frac{116}{100}$ calcule a partir da distribuição de probabilidades de X+Y.

$$Var(X+Y) = \frac{116}{100} = \underbrace{\frac{76}{100}}_{Var(X)} + \underbrace{\frac{60}{100}}_{Var(Y)} + 2\cdot\underbrace{\left(-\frac{1}{10}\right)}_{Cov(X,Y)}$$

Definição

O coeficiente de correlação entre as variáveis aleatórias X e Y é dado por

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

σ_X - desvio padrão de X $= \sqrt{Var(X)}$

σ_Y - desvio padrão de Y $= \sqrt{Var(Y)}$.

Propriedades:

$$-1 \leq \rho(X,Y) \leq 1$$

O coeficiente de correlação é uma medida do grau da relação linear entre X e Y.

$\rho(X,Y) = \pm 1 \iff$ relação linear perfeita entre X e Y da forma $Y = aX + b$.

$\rho(X,Y) = 1 \Rightarrow a > 0$ $\qquad$ $\rho(X,Y) = -1 \Rightarrow a < 0$

$\rho(X,Y) \simeq 1$ forte mas não perfeita relação linear crescente

$\rho(X,Y) \simeq -1$ forte mas não perfeita relação linear decrescente

Resultado 5

Se X e Y são variáveis aleatórias independentes, $Cov(X, Y) = 0$ e portanto

$$Var(X+Y) = Var(X) + Var(Y).$$

De modo geral, se $X_1, X_2, \ldots X_n$ são n variáveis aleatórias

a) $E(X_1 + X_2 + \ldots + X_n) = E(X_1) + E(X_2) + \ldots + E(X_n)$.

b) Se adicionalmente, $X_1, X_2, \ldots X_n$ são independentes

$$Var(X_1 + X_2 + \ldots + X_n) = Var(X_1) + Var(X_2) + \ldots Var(X_n).$$

Obs: X em cm

$E(X) = \mu_X$ cm

$\sigma^2_X = Var(X)$ cm^2

$\sigma_X = DP(X)$ cm

Y em Kg

$Cov(X\,Y)$ cmKg

$\rho(X, Y)$: $\frac{cmKg}{cmKg}$ sem unidade

Coeficiente de Variação

$\frac{\sigma_X}{\mu_X}$: $\frac{cm}{cm}$ $\rightarrow$ sem unidade

Variáveis Aleatórias Contínuas
Revisão - Capítulo 6 - Magalhães e Lima

Definição de Variável Aleatória

$X: \Omega \to \mathbb{R}$
↳ espaço amostral

Se o conjunto de possíveis valores de X é enumerável → X é v.a. discreta

Se o conjunto de possíveis valores de X é um intervalo de números reais → X é v.a. contínua

Ex: Peso, Altura, Renda, Tempo de realização de uma prova, comprimento de uma peça.

Exemplo 6.1 - Pag 177

X - profundidade de um lençol de água em uma região

Sabe-se que pode ser qualquer ponto entre 20 e 100 m.

Intervalo de possíveis valores da variável: $[20, 100]$.

Não se dispõe de informações adicionais sobre a profundidade, de modo que qualquer ponto no intervalo [20, 100] é um possível valor da profundidade e todos os intervalos de mesma amplitude são equiprováveis.

Devido a imprecisões do instrumento, um valor de 32,571 m pode ser arredondado para 32,6. Isso ocorre com altura, peso, etc.

O intervalo [20, 100] possui infinitos pontos. Não é possível atribuir probabilidades a cada um deles, pois se cada um tiver probabilidade $\neq 0$, a soma será > 1 (∞).

Por esse motivo, para variáveis aleatórias contínuas, atribui-se probabilidades a intervalos.

Definição 1

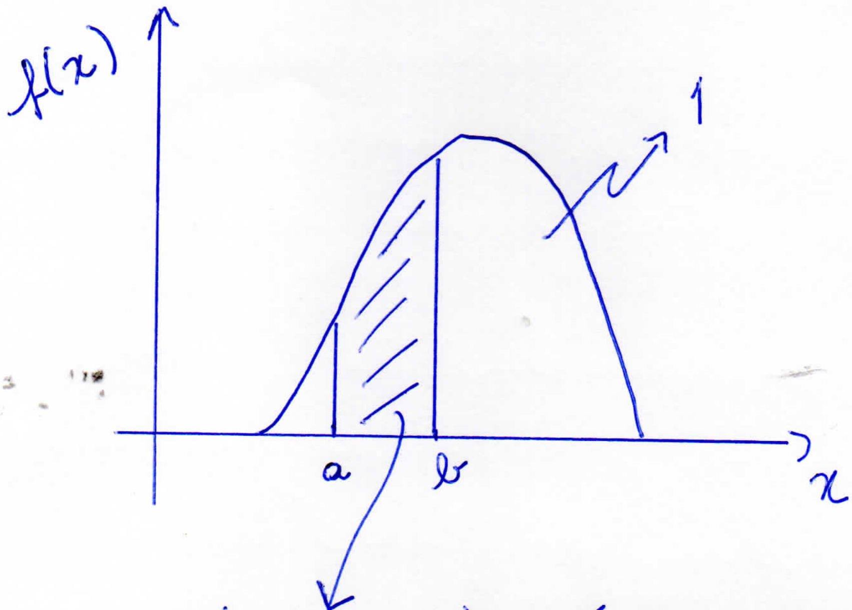
A função $f(x)$ é denominada função densidade de probabilidades da variável aleatória contínua X se

i) $f(x) \geq 0, \quad x \in \mathbb{R}$

ii) $\int_{-\infty}^{\infty} f(x)\,dx = 1$ [a área entre o eixo das abcissas e o gráfico de $f(x)$ é igual a 1.]

Para calcular probabilidades, dados $a, b \in \mathbb{R}$, $a \leq b$,

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx.$$



$P(a \leq X \leq b)$ - área entre o gráfico de $f(x)$ e o eixo das abcissas no intervalo $[a, b]$.

Consequências

$$P(X=K) = P(K \leq X \leq K) = \int_K^K f(x)\,dx = 0$$

$$P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) =$$
$$= P(a < X < b)$$

Exercícios

1) No exemplo

a) Obter $f(x)$.

b) Calcular $P(25 \leq X \leq 29)$.

2) Exemplo 6.3 pag 183

Teste educacional realizado para avaliar o desenvolvimento de crianças.

T - tempo em minutos de execução do teste

Um modelo teórico propõe que a v.a. contínua T tem função densidade de probabilidades

$$f(t) = \begin{cases} \frac{1}{40}(t-4) & 8 \leq t < 10 \\ \frac{3}{20}, & 10 \leq t \leq 15 \\ 0 & \text{caso contrário} \end{cases}$$

a) Construa o gráfico de $f(t)$.

b) Selecionando-se uma criança ao acaso, qual é a probabilidade do seu tempo de execução ser maior que 9 e menor ou igual a 12?