Estatística

4 - Variáveis Aleatórias Unidimensionais

"Variável Aleatória X" é uma Função que associa um número real a cada Evento

$$X : S \longrightarrow R$$

Notação: X (MAIÚSCULA) .... Varíável Aleatória X

x (minúscula) .... valor da Variável Aleatória X

Exemplo: Pr(X=x) ... Prob. da V.A. X assumir o valor x

1ª. Vantagem: Facilitar o tratamento matemático

EXEMPLOS :

1) Experimento : jogar 1 dado

Variável Aleatória X : "dobro do número obtido menos 1"

X : {1, 2, 3, 4, 5, 6} ⟶ {1, 3, 5, 7, 9, 11}

2) Experimento : jogar 4 moedas (C: Cara e K: Coroa)

Variável Aleatória Y : "números de caras obtidas"

Y : {CCCC, CKCC, ..., KKKK} ⟶ {0, 1, 2, 3, 4}

2ª. Vantagem: Cada variável aleatória tem associada uma distribuição de probabilidades que caracteriza seu comportamento.

Ex. 1: Dado

x	Pr(x)
1	1/6
3	1/6
5	1/6
7	1/6
9	1/6
11	1/6

Ex.2: Moeda

y	Pr(y)
0	1/16
1	4/16
2	6/16
3	4/16
4	1/16

$X : S \longrightarrow \text{Real}$

$Pr(X=x) : \text{Real} \longrightarrow [0, 1]$

Variáveis Aleatórias DISCRETAS

✓ Possíveis resultados do experimento são representados por um conjunto enumerável, isto é, são representados por n números

✓ função distribuição de probabilidade:

associa probabilidades aos possíveis valores da variável aleatória.

Propriedades :

(a) $\Pr(X = x_k) \geq 0, \forall k \quad S = \{x_1, x_2, \ldots, x_n\}$

(b) $$\Pr(x_a \leq X \leq x_b) = \sum_{k=a}^{b} \Pr(X = k), \; b \geq a$$

(c) $$\sum_{k=1}^{n} \Pr(X = x_k) = 1$$

- Possíveis resultados do experimento são representados por valores pertencentes a um intervalo contínuo ou pertencentes a um conjunto de intervalos

- função densidade de probabilidade (fdp): f(x)

  caracteriza como as probabilidades se distribuem num intervalo contínuo

Propriedades :

$$\text{(a)} \quad f(x) \geq 0$$

$$\text{(b)} \quad \Pr(a \leq X < b) = \int_{b}^{a} f(x) \cdot dx \quad , b > a$$

$$\text{(c)} \quad \int_{-\infty}^{+\infty} f(x) \cdot dx = 1$$

Exemplos de Variáveis Aleatórias contínuas

- EXPERIMENTO 1: ponteiro, com origem no centro de um disco fixo, gira até parar

  Variável Aleatória: ângulo de parada do ponteiro com relação a uma marca de referência existente no disco; valores de 0° até 360°.

- EXPERIMENTO 2: identificar o diâmetro de um eixo

  Variável Aleatória: valores entre zero e um limitante superior

IMPORTANTE:

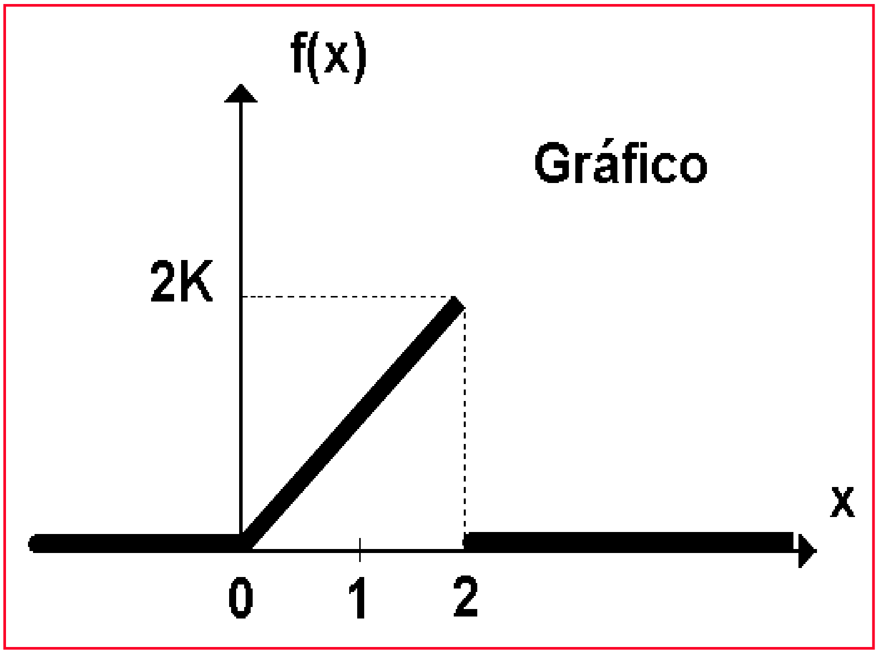
Probabilidade associada a cada valor individual de uma "Variável Aleatória Contínua" é NULA

$\Pr(X = x_a) = 0$

$$\Pr(X = x_a) = \int_{x_a}^{x_a} f(x) \cdot dx = 0$$

Exemplo de função densidade de probabilidade

$$f(x) = \begin{cases} 0 & , \quad x < 0 \\ k \cdot x & , \quad 0 \leq x \leq 2 \\ 0 & , \quad x > 2 \end{cases} \qquad\qquad k = ?$$



$$\text{Área}_{0\to 2} = \frac{2 \times 2k}{2} = 2K = 1$$

$$\Rightarrow k = 1/2$$

$$\Rightarrow f(x) = x/2 \ , \qquad 0 \leq x \leq 2$$

$$\Pr(X \leq 1) = \Pr(0 \leq X \leq 1) = 1/4$$

Função de Distribuição Acumulada

Forma alternativa para caracterizar a distribuição de probabilidades de uma variável aleatória (discreta ou contínua) .

$$F(x) = Pr(X \le x) = \sum_{x_i \le x} Pr(x_i) \quad \text{(DISCRETA)}$$

$$F(x) = Pr(X \le x) = \int_{-\infty}^{x} f(x)\,dx \quad \text{(CONTÍNUA)}$$

PROPRIEDADES :

$(a)\; 0 \le F(x) \le 1 \quad (b)\; F(-\infty) = 0 \quad (c)\; F(+\infty) = 1$

$(d)\; F(x)$ é não decrescente

$(e)\; P(a < X \le b) = F(b) - F(a)\ ,\ b > a$

(f) F(x) é contínua à direita em qualquer ponto

(g) F(x) é descontínua à esquerda nos pontos de probabilidade não nula

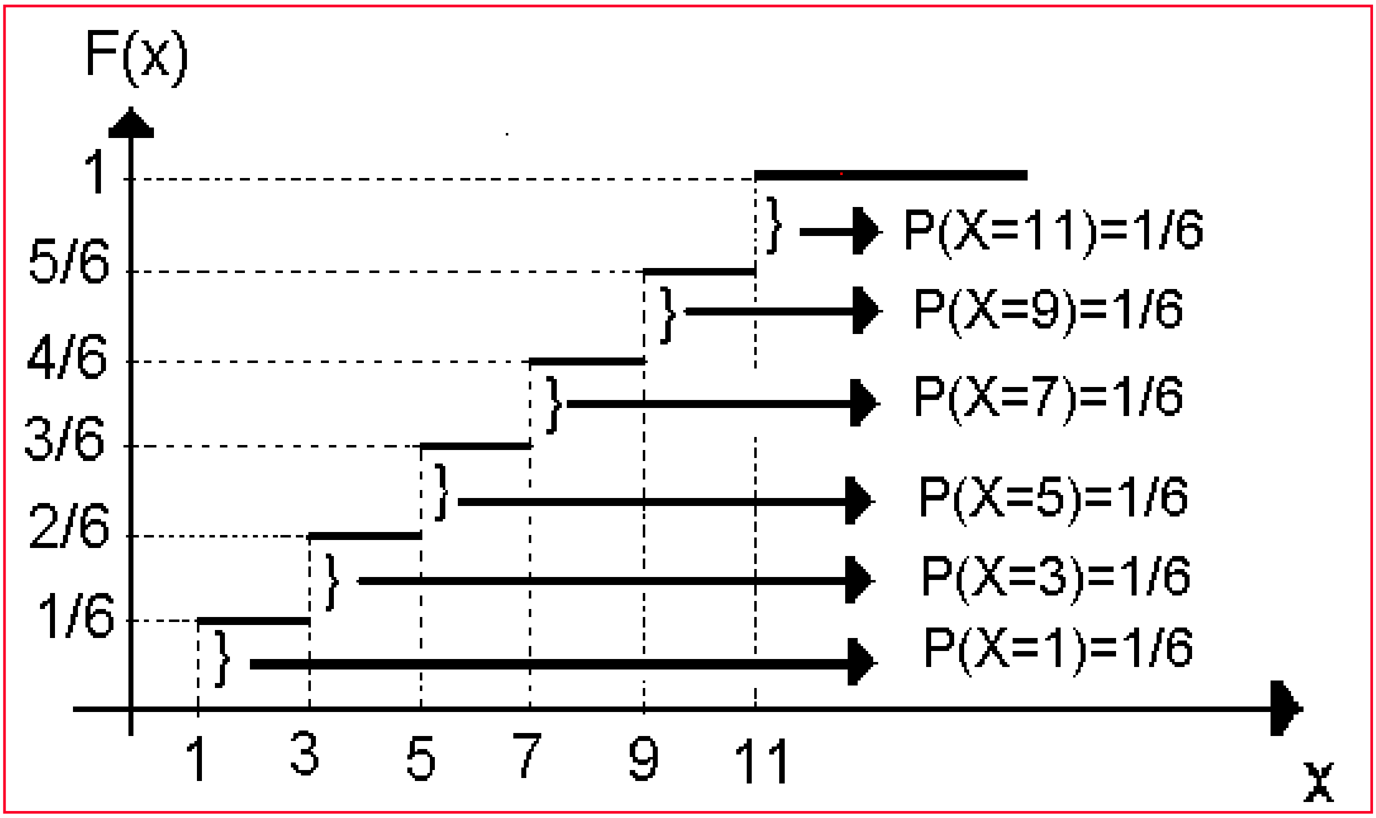
Exemplo de Distribuição Acumulada - Caso Discreto

Experimento : jogar um dado e observar face superior

v. a. X = dobro do número que sair menos 1

x	Pr(x)
1	1/6
3	1/6
5	1/6
7	1/6
9	1/6
11	1/6

F(x)	intervalos
0	$x < 1$
1/6	$1 \leqslant x < 3$
2/6	$3 \leqslant x < 5$
3/6	$5 \leqslant x < 7$
4/6	$7 \leqslant x < 9$
5/6	$9 \leqslant x < 11$
1	$x \geq 11$

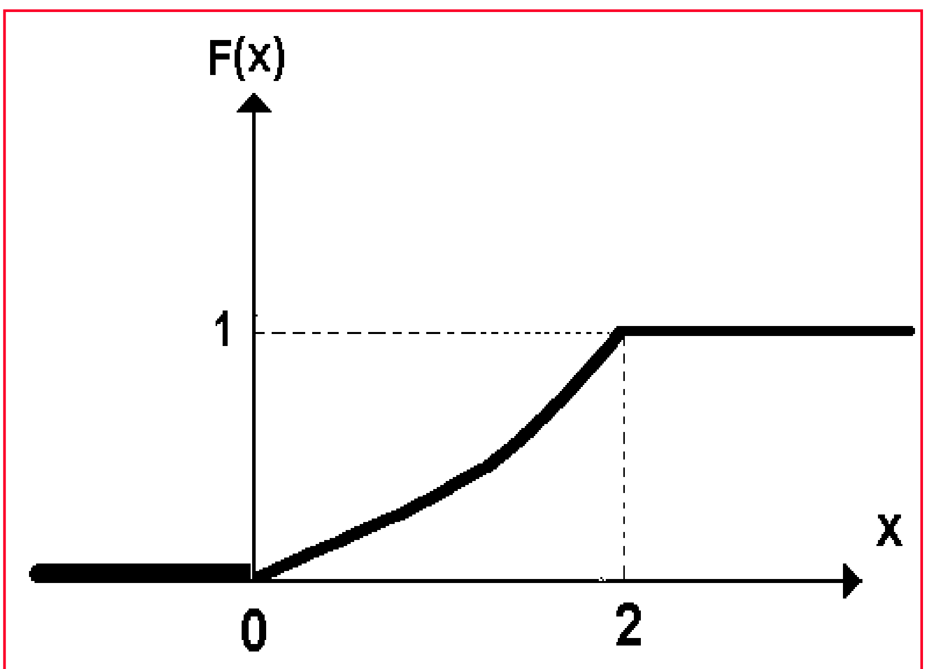


Exemplo de Distribuição Acumulada - Caso Contínuo

$$f(x) = \begin{cases} 0 & , \quad x < 0 \\ \dfrac{x}{2} & , \; 0 \le x \le 2 \\ 0 & , \quad x > 2 \end{cases}$$

$F(x) = ?$

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ \dfrac{1}{4}x^2 & , \; 0 \le x \le 2 \\ 1 & , \quad x > 2 \end{cases}$$



Média (Valor Esperado, Esperança, Expectância)

$$E(X) = \sum_{i} x_i \cdot P(x_i) = \mu_X \qquad \text{(V.A. discreta)}$$

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f(x)\,dx = \mu_X \qquad \text{(V.A. contínua)}$$

Média caracteriza o centróide da distribuição

Propriedades:

(a) $E(k) = k$, k = constante

(b) $E(kX) = kE(X)$

(c) $E(X \pm Y) = E(X) \pm E(Y)$

(d) $E(X \pm k) = E(X) \pm k$

(e) $E(XY) = E(X) . E(Y)$ ⟵ caso X , Y independentes

Mediana (md): divide a distribuição probabilidade em 2 partes equiprováveis

$$\boxed{Pr(X \leq md) = Pr(X \geq md) = 0{,}5 = 50\%}$$

Mediana: menor valor para o qual F(x) > 0,5

Generalizando:

Distribuição pode ser dividida em várias partes equiprováveis.

QUARTIS → 4 partes

DECIS → 10 partes

PERCENTIS → 100 partes

Moda: ponto(s) de
maior probabilidade (V.A.discreta) ou
maior densidade de probabilidade (V.A.contínua)

➔ Indica a região mais provável da distribuição

Caracterizam a variabilidade das Variáveis Aleatórias

Variância : $\boxed{Var(X)}$

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

V.A.Discreta : $Var(X) = \sum_i \left[(x_i - E(X))^2 \cdot \Pr(x_i) \right]$

V.A. Contínua : $Var(X) = \int_{-\infty}^{+\infty} (x - E(X))^2 f(x) dx$

Forma alternativa para cálculo de σ^2 :

$$Var(X) = E(X^2) - (\mathrm{E}(\mathrm{X}))^2$$

V.A. Discreta : $\mathrm{E}(X^2) = \sum_i \left[x^2 \cdot \Pr(x_i) \right]$

V.A.Contínua : $\mathrm{E}(X^2) = \int_{-\infty}^{+\infty} x^2 f(x) dx$

$$a)\ Var(k) = 0 \quad \text{onde k : constante}$$

$$b)\ Var(k \cdot X) = \mathrm{k}^2 \cdot Var(X)$$

$$c)\ Var(X \pm Y) = \mathrm{Var}(X) + VAr(Y) \Leftrightarrow X, Y\ \ V.A.\ Independentes$$

$$d)\ Var(X \pm k) = Var(X)$$

DESVIO-PADRÃO:

$$DP(X) = \sqrt{Var(X)} = \sigma_X = \sigma(X)$$

COEFICIENTE DE VARIAÇÃO

$$CV(X) = \frac{DP(X)}{E(X)}$$

Passa a idéia da dispersão relativa.

AMPLITUDE R(X):

Diferença entre o maior e o menor valor possível da Variável Aleatória X

Exercício 4.1:

Uma máquina automática enche garrafas, saindo a produção com peso bruto médio de 850g e desvio-padrão de 4,5g. As garrafas utilizadas têm peso médio de 220g e desvio-padrão de 2,7g. Determine o peso líquido médio e o seu desvio-padrão se:

a) a máquina pesa o líquido dentro da garrafa;

b) a máquina pesa o líquido e depois o coloca na garrafa.

X: peso bruto (garrafa cheia)

Y: peso da garrafa vazia

W: peso do líquido

a) W = X – Y logo:

E(W) = E(X) - E(Y) = 850 – 220 = 630

Var(W) = Var(X) + Var(Z) = $(4,5)^2 + (2,7)^2 = 27,54 = (5,2)^2$

DP(W)= 5,2 (considerar o peso bruto independente do peso da garrafa vazia)

b) X = Y + W logo:

E(X) = E(Y) + E(W) $\Rightarrow$ 850 = 220 + E(W) $\Rightarrow$ E(W) = 630

Var(X) = Var(Y) + Var(W) $\Rightarrow$ $(4,5)^2 = (2,7)^2$ + Var(W)

$\Rightarrow$ Var(W) = 20,25 – 7,29 = 12,96 $\Rightarrow$ DP(W) = 3,6

(considerar o peso da garrafa vazia independente do peso do líquido)

Exercício 4.2:

Uma revista mensal pretende lançar uma campanha publicitária para conseguir novos assinantes, a qual consistirá no envio de revistas grátis para 10.000 pessoas. Experiências anteriores permitem avaliar o número de prováveis novos assinantes (X), assim:

x	1500	2000	2500	3000
Pr(X=x)	0,10	0,30	0,50	0,10

Sabe-se:
- custo da campanha: R$12,00 / pessoa;
- custo unitário da revista: R$8,00;
- valor mensal de cada assinatura: R$10,00

Qual o prazo de retorno do investimento na campanha ?

X: número de novos assinantes, dentre as 10.000 pessoas

F=10•k•X : faturamento em k meses, ref. aos novos ass.

C=12•10000+8•k•X : custo da campanha + custo da revista

R = F – C : resultado financeiro do período de k meses

E(X) = 1500•0,10 + 2000•0,30 + 2500•0,50 + 3000•0,10 = 2300

E(F) = 10•k•μ(X) = 10•k•2300 =23000•k

E(C) = 120000 + 8•K•μ(X) = 120000 + 18400•K

E(R) = E(F) – E(C) = 23000•k – 120000 – 18400•k

E(R) = 4600•k – 120000

k* : número de meses tal que E(R) = 0 $\Rightarrow$ k*=120000 / 4600 = 26

Logo, a partir do 26º mês tem-se resultado esperado lucrativo