MÉTODOS ESTATÍSTICOS

Distribuições Teóricas

Licenciatura em Engenharia Informática

Departamento de Matemática
Escola Superior de Tecnologia de Setúbal
Instituto Politécnico de Setúbal
2023-2024

Experiência Aleatória

Designa-se por **experiência** como sendo qualquer processo capaz de produzir resultados observáveis.

Quando uma experiência está sujeita à influência de fatores casuais e conduz a resultados incertos diz-se uma **experiência aleatória**.

As experiências aleatórias caracterizam-se por:

- poder repetir-se um grande número de vezes nas mesmas condições;
- não existir um conhecimento suficiente para prever o resultado;
- existência de regularidade quando se repete a experiência um grande número de vezes.

Espaço de Resultados

É o conjunto formado por todos os resultados possíveis de uma experiência aleatória.
O espaço de resultados representa-se por Ω.

Acontecimento

Os subconjuntos de Ω designam-se por acontecimentos.
Em geral, os acontecimentos representam-se com letras maiúsculas: A, B, ...

Definição Clássica de Probabilidades

$$P(A) = \frac{\text{número de casos favoráveis}}{\text{número de casos possíveis}}$$

Só é aplicável quando o espaço de resultados é finito e os elementos do espaço de resultados possuem igual probabilidade de ocorrerem.

Definição Axiomática de Probabilidades

Seja Ω o espaço de resultados e $\mathcal{P}(\Omega)$ o conjunto formado por todos os subconjuntos de Ω, então **Probabilidade** é uma função $P : \mathcal{P}(\Omega) \to [0, 1]$ que verifica as seguintes propriedades:

- $P(A) \geq 0, \forall A \subseteq \Omega$;
- $P(\Omega) = 1$;
- Se $A \cap B = \emptyset$, então $P(A \cup B) = P(A) + P(B)$.

Propriedades das Probabilidades de Acontecimentos

- $0 \leq P(A) \leq 1$
- $P(\Omega) = 1$
- $P(\varnothing) = 0$
- $P\left(\overline{A}\right) = 1 - P(A)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - $
  $-P(B \cap C) + P(A \cap B \cap C)$
- $P(A - B) = P(A) - P(A \cap B)$
- $P(A/B) = \frac{P(A \cap B)}{P(B)}, \qquad P(B) > 0$
- Se A e B acontecimentos independentes, $P(A \cap B) = P(A) \times P(B)$

Exemplo 1

Experiência aleatória: lançamento de um dado não viciado.

Espaço de resultados:

$$\Omega = \{\text{face } 1, \text{face } 2, \text{face } 3, \text{face } 4, \text{face } 5, \text{face } 6\}$$

Neste caso, como o espaço de resultados é finito e o dado não é viciado, é possível calcular as probabilidade de acontecimentos recorrendo à definição clássica de probabilidades.

1. acontecimento $A =$ num lançamento do dado sair face 4 $= \{\text{face } 4\}$
   A diz-se acontecimento elementar (só 1 elemento)
   $P(A) = \frac{1}{6}$

2. acontecimento $B =$ num lançamento do dado sair face superior a 2 $=$
   $= \{\text{face } 3, \text{face } 4, \text{face } 5, \text{face } 6\}$
   $P(B) = \frac{4}{6} = \frac{2}{3}$

Exemplo 1

3. acontecimento $C =$ num lançamento do dado sair face inferior a 8 $= \Omega$

   C diz-se acontecimento certo (todos os elementos do espaço de resultados)

   $P(C) = P(\Omega) = \frac{6}{6} = 1$

4. acontecimento $D =$ num lançamento do dado sair face igual a 20$= \{\}$

   D diz-se acontecimento impossível (nenhum elemento do espaço de resultados)

   $P(D) = \frac{0}{6} = 0$

5. acontecimento $A \cap B =$ num lançamento do dado sair face 4 **e** num lançamento do dado sair face superior a 2 $= \{$face 4$\}$

   $P(A \cap B) = \frac{1}{6}$

Exemplo 1

6. acontecimento $A \cup B$ = num lançamento do dado sair face 4 **ou** num lançamento do dado sair face superior a 2 $= \{\text{face } 3, \text{face } 4, \text{face } 5, \text{face } 6\}$

   $P(A \cup B) = \frac{4}{6} = \frac{2}{3}$

   ou

   $P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{6} + \frac{2}{3} - \frac{1}{6} = \frac{2}{3}$

7. acontecimento $\overline{B}$ = num lançamento do dado **não** sair face superior a 2 =

   = num lançamento do dado sair face inferior ou igual a 2 =

   $= \{\text{face } 1, \text{face } 2\}$

   $\overline{B}$ diz-se o acontecimento contrário de B

   $P(\overline{B}) = \frac{2}{6} = \frac{1}{3}$

   ou

   $P(\overline{B}) = 1 - P(B) = 1 - \frac{2}{3} = \frac{1}{3}$

Exemplo 1

8. $P(A|B) \longmapsto$ probabilidade de num lançamento do dado sair face 4 **sabendo** que no lançamento do dado saiu face superior a 2

   $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{6}}{\frac{2}{3}} = \frac{1}{4}$

9. e podiam ser definidos muitos outros acontecimentos...

Variáveis Aleatórias

Definição

Chama-se variável aleatória (v.a.) e representa-se por X, a uma função de domínio Ω e conjunto de chegada $\mathbb{R}$, cujo valor é determinado pelo resultado de uma experiência aleatória, isto é

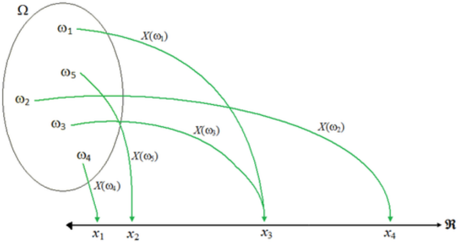
$$\begin{array}{rccc} X: & \Omega & \rightarrow & \mathbb{R} \\ & w & \rightarrow & X(w) = x \end{array}$$

Observação

Uma variável aleatória é uma função e não uma variável no sentido em que é habitualmente empregue em Matemática.

Variáveis Aleatórias

Muitas vezes o resultado de uma experiência aleatória não é numérico ou sendo-o não interessa lidar com os resultados possíveis de Ω, mas pretende-se associar-lhe uma quantidade numérica.



Ou seja, uma variável aleatória é uma função que "associa um código" numérico aos acontecimentos, permitindo desta forma deixar de trabalhar em Ω e passar a trabalhar em $\mathbb{R}$.

Variáveis Aleatórias

No entanto convém ter atenção que o "código"definido para construir a função variável aleatória tem de ser tal que permita andar ao contrário e conseguir saber qual o acontecimento pretendido. Ou seja, de uma forma mais formal:

Associemos a cada elemento $\omega \in \Omega$ um número real $X(\omega) \in E \subset \mathbb{R}$.
Sendo A um acontecimento, tem-se

$$X(A) = \{X(\omega) : \omega \in A\}$$

por outro lado, tem-se para cada subconjunto $E \subset \mathbb{R}$

$$X^{-1}(E) = \{\omega : X(\omega) \in E\}$$

Exemplo 1

Experiência aleatória: lançamento de um dado não viciado.

Espaço de resultados:

$$\Omega = \{\text{face } 1, \text{face } 2, \text{face } 3, \text{face } 4, \text{face } 5, \text{face } 6\}$$

Neste caso é fácil pensar numa variável aleatória que represente de forma numérica o espaço de resultados:

Variável Aleatória: $X =$ número da face virada para cima no lançamento de um dado não viciado

Neste caso estamos a considerar face $1 = 1$, face $2 = 2$, face $3 = 3$, face $4 = 4$, face $5 = 5$ e face $6 = 6$.

Agora os acontecimentos podem ser definidos à custa da variável aleatória:

Exemplo 1

Variável Aleatória: $X =$ número da face virada para cima no lançamento de um dado não viciado

1. acontecimento $A =$ num lançamento do dado sair face 4 $= \{\text{face } 4\}$
   $P(A) = P(X = 4) = \frac{1}{6}$

2. acontecimento $B =$ num lançamento do dado sair face superior a 2 $=$
   $= \{\text{face } 3, \text{face } 4, \text{face } 5, \text{face } 6\}$
   $P(B) = P(X > 2) = \frac{4}{6} = \frac{2}{3}$

3. acontecimento $C =$ num lançamento do dado sair face inferior a 8 $= \Omega$
   $P(C) = P(X < 8) = 1$

4. acontecimento $D =$ num lançamento do dado sair face igual a 20$= \{\}$
   $P(D) = P(X = 20) = 0$

Exemplo 1

5. acontecimento $A \cap B =$ num lançamento do dado sair face 4 **e** num lançamento do dado sair face superior a 2 $= \{$face 4$\}$
   $P(A \cap B) = P(X = 4 \wedge X > 2) = P(X = 4) = \frac{1}{6}$

6. acontecimento $A \cup B =$ num lançamento do dado sair face 4 **ou** num lançamento do dado sair face superior a 2 $= \{$face 3, face 4, face 5, face 6$\}$
   $P(A \cup B) = P(X = 4 \vee X > 2) = P(X > 2) = \frac{2}{3}$

7. acontecimento $\overline{B} =$ num lançamento do dado **não** sair face superior a 2 $=$
   $=$ num lançamento do dado sair face inferior ou igual a 2 $=$
   $= \{$face 1, face 2$\}$
   $P(\overline{B}) = P(X \leq 2) = \frac{2}{6} = \frac{1}{3}$
   ou
   $P(\overline{B}) = P(X \leq 2) = 1 - P(X > 2) = 1 - \frac{2}{3} = \frac{1}{3}$

8. $P(A|B) = P(X = 4|X > 2) = \frac{P(X=4 \wedge X>2)}{P(X>2)} = \frac{P(X=4)}{P(X>2)} = \frac{\frac{1}{6}}{\frac{2}{3}} = \frac{1}{4}$

Exemplo 2

Experiência aleatória: lançamento de duas moedas não viciadas.

Espaço de resultados:

$$\Omega = \{(\text{Cara}, \text{Cara}), (\text{Cara}, \text{Coroa}), (\text{Coroa}, \text{Cara}), (\text{Coroa}, \text{Coroa})\}$$

Variável Aleatória: X = número da caras que saem no lançamento de duas moedas não viciadas

Ou seja, equivale a considerar Cara $= 1$ e Coroa $= 0$.

Agora podemos definir as probabilidades pretendidas mas à custa da variável aleatória definida, no entanto a forma de pensar para calcular as probabilidades pode ser feita da mesma forma que eram feitas para os acontecimentos (não esquecer que a função variável aleatória tem de ser tal que permita determinar qual o acontecimento que está associado a cada condição).

Também aqui é possível calcular as probabilidade recorrendo à definição clássica de probabilidades.

Exemplo 2

Experiência aleatória: lançamento de duas moedas não viciadas.

Espaço de resultados:

$$\Omega = \{(\text{Cara}, \text{Cara}), (\text{Cara}, \text{Coroa}), (\text{Coroa}, \text{Cara}), (\text{Coroa}, \text{Coroa})\}$$

Variável Aleatória: X = número da Caras que saem no lançamento de duas moedas não viciadas

1. $P(X = 0) = \frac{1}{4}$
2. $P(X > 1) = P(X = 2) = \frac{1}{4}$
3. $P(X < 8) = 1$
4. $P(X = 20) = 0$

Tipos de Variáveis Aleatórias

Variáveis Aleatórias Discretas

Uma variável aleatória diz-se **Discreta** se pode assumir um número finito ou infinito numerável de valores $\rightarrow$ associada a contagens.

Exemplos:

- número da face virada para cima no lançamento de um dado não viciado;
- número de caras que saem no lançamento de duas moedas não viciadas;
- número de pessoas numa fila na caixa de um supermercado durante uma hora.

Variáveis Aleatórias Contínuas

Uma variável aleatória diz-se **Contínua** se pode assumir um número infinito não numerável de valores $\rightarrow$ associada a medidas.

Exemplos:

- peso dos habitantes de Setúbal;
- altura dos alunos da ESTSetúbal;
- tempo, em minutos, que os alunos da ESTSetúbal demoram de casa até à escola.

Variáveis Aleatórias Discretas

Uma variável aleatória diz-se **Discreta** se pode assumir um número finito ou infinito numerável de valores.

Uma variável aleatória discreta fica perfeitamente identificada através da:

- **função de probabilidade**

  ou

- **função de distribuição**

e através dos seus parâmetros (apenas vamos considerar 3):

- **valor esperado** ou média ou esperança matemática
- **variância**
- **desvio padrão**

Variáveis Aleatórias Discretas

Função de Probabilidade (f.p.)

Se X é uma variável aleatória discreta, que assume valores distintos $x_1, x_2, \ldots, x_n$, então a função de probabilidade (ou função massa de probabilidade) é representada por $f(x)$ e é definida por

$$f(x) = P(X = x) = \begin{cases} P(X = x) & , x = x_j \\ 0 & , x \neq x_j \end{cases} \qquad j = 1, \ldots, n$$

ou esquematicamente por

x	x_1	x_2	$\cdots$	x_j	$\cdots$	x_n
$f(x)$	$P(X = x_1)$	$P(X = x_2)$	$\cdots$	$P(X = x_j)$	$\cdots$	$P(X = x_n)$

e satisfaz as seguintes propriedades

1. $f(x) \geq 0, \quad \forall x;$
2. $\sum\limits_{x} f(x) = 1.$

Função de probabilidade = f(x)

- Nas variáveis aleatórias **discretas** tem-se a **função de probabilidade** que permite calcular as probabilidades pontuais:

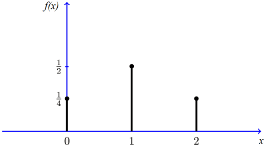
$$f(x) = P(X = x)$$

(**observação:** corresponde à coluna das frequências relativas nas tabelas de frequências)

- Como a variável aleatória discreta tem um domínio numerável (finito ou infinito), a função de probabilidade indica o comportamento probabilístico ponto a ponto do domínio.
- A forma mais usual de representar uma função de probabilidade é recorrendo a uma tabela, na primeira linha são colocados os pontos do domínio (os valores que não estão representados na tabela significa que têm probabilidade zero de acontecer) e na segunda linha coloca-se a respetiva probabilidade pontual.
- Graficamente as funções de probabilidade são representadas por gráficos de barras.
- Podemos ver alguns exemplos de representações da função de probabilidade de variáveis aleatórias discretas:

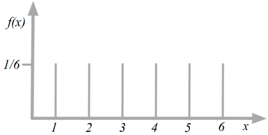
Exemplo 3 X variável aleatória discreta com domínio $D_X = \{0, 1, 2\}$ e função de probabilidade:

x	0	1	2
$f(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$



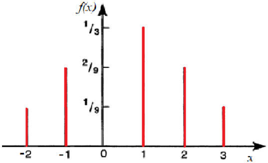
Exemplo 4 X variável aleatória discreta com domínio $D_X = \{1, 2, 3, 4, 5, 6\}$ e função de probabilidade:

$$f(x) = \begin{cases} \frac{1}{6} & , x \in D_X \\ 0 & , x \notin D_X \end{cases}$$



Exemplo 5 X variável aleatória discreta com domínio $D_X = \{-2, -1, 1, 2, 3\}$ e função de probabilidade:

x	-2	-1	1	2	3
$f(x)$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{3}$	$\frac{2}{9}$	$\frac{1}{9}$



Função de probabilidade = f(x)

No entanto é necessário ter em atenção que uma função $f(x)$ só é uma **função de probabilidade** se verificar as seguintes propriedades:

1. $f(x) \geq 0, \quad \forall x;$
2. $\sum_{x} f(x) = 1.$

Ou seja, a função de probabilidade não pode ser negativa e a soma de todas as probabilidades pontuais é 1.

Função de probabilidade = f(x)

Exemplo 6

Seja X uma variável aleatória discreta que representa o número de pessoas que visitam, por dia, um determinado site e seja f a sua função de probabilidade representada por:

x	0	1	2	3	4	5
$f(x)$	0.04	0.5	0.24	0.12	0.05	0.05

1. Mostre que f é de facto uma função de probabilidade.

O domínio da variável aleatória X é $D_X = \{0, 1, 2, 3, 4, 5\}$.

$f(x)$ é função de probabilidade sse

1. $f(x) \geq 0, \quad \forall x;$
2. $\sum_x f(x) = 1.$

propriedade 1:

$$f(0) = 0.04 \geq 0 \qquad f(3) = 0.12 \geq 0$$
$$f(1) = 0.5 \geq 0 \qquad f(4) = 0.05 \geq 0 \qquad f(x) = 0 \geq 0 \text{ se } x \notin D_X$$
$$f(2) = 0.24 \geq 0 \qquad f(5) = 0.05 \geq 0$$

propriedade 2:

$$\sum_x f(x) = f(0) + f(1) + f(2) + f(3) + f(4) + f(5) =$$

$$= 0.04 + 0.5 + 0.24 + 0.12 + 0.05 + 0.05 = 1$$

Como verifica as duas propriedades, então é função de probabilidade.

Função de probabilidade = f(x)

Exemplo 6

Seja X uma variável aleatória discreta que representa o número de pessoas que visitam, por dia, um determinado site e seja f a sua função de probabilidade representada por:

x	0	1	2	3	4	5
$f(x)$	0.04	0.5	0.24	0.12	0.05	0.05

2. Calcule

a) $P(X = 3)$
b) $P(X \leq 3)$
c) $P(X < 3)$
d) $P(X > 3)$
e) $P(X \geq 3)$
f) $P(3 \leq X < 5)$
g) $P(3 < X \leq 5)$
h) $P(X > 3 | X < 5)$

Vamos calcular as probabilidades recorrendo apenas à função de probabilidade:

a) $P\left(X=3\right)=f(3)=0.12$

b) $P\left(X\leq 3\right)=f(0)+f(1)+f(2)+f(3)=0.04+0.5+0.24+0.12=0.9$

c) $P\left(X<3\right)=P\left(X\leq 2\right)=f(0)+f(1)+f(2)=0.04+0.5+0.24=0.78$

d) $P\left(X>3\right)=1-P\left(X\leq 3\right)=1-0.9=0.1$

d) $P\left(X>3\right)=f(4)+f(5)=0.05+0.05=0.1$

e) $P\left(X\geq 3\right)=1-P\left(X<3\right)=1-0.78=0.22$

e) $P\left(X\geq 3\right)=f(3)+f(4)+f(5)=0.12+0.05+0.05=0.22$

f) $P\left(3\leq X<5\right)=f(3)+f(4)=0.12+0.05=0.17$

g) $P\left(3<X\leq 5\right)=f(4)+f(5)=0.05+0.05=0.1$

h) $P\left(X>3|X<5\right)=\frac{P(X>3\wedge X<5)}{P(X<5)}=\frac{P(3<X<5)}{1-P(X=5)}=\frac{f(4)}{1-f(5)}=\frac{0.05}{1-0.05}=0.0526$

Variáveis Aleatórias Discretas

Função de Distribuição (f.d.)

Se X é uma variável aleatória discreta, que assume valores distintos $x_1, x_2, \ldots, x_n$, então a função de distribuição (ou função de distribuição acumulada) é representada por $F(x)$ e é definida por

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$$

isto é,

$$F(x) = \begin{cases} 0 & , \ x < x_1 \\ P(X = x_1) & , \ x_1 \leq x < x_2 \\ P(X = x_1) + P(X = x_2) & , \ x_2 \leq x < x_3 \\ \vdots & , \ \vdots \\ P(X = x_1) + \cdots + P(X = x_{n-1}) & , \ x_{n-1} \leq x < x_n \\ 1 & , \ x \geq x_n \end{cases}$$

Variáveis Aleatórias Discretas

Função de Distribuição (f.d.)

e satisfaz as seguintes propriedades:

1. $0 \leq F(x) \leq 1;$
2. $F(x)$ é uma função não decrescente;
3. $F(x)$ é contínua à direita;
4. $\lim_{x\to-\infty} F(x) = 0$ e $\lim_{x\to+\infty} F(x) = 1;$
5. $P(X = a) = F(a) - P(X < a);$
6. $$P(a < X \leq b) = F(b) - F(a);$$
   $$P(a < X < b) = F(b) - F(a) - P(X = b);$$
   $$P(a \leq X \leq b) = F(b) - F(a) + P(X = a);$$
   $$P(a \leq X < b) = F(b) - F(a) - P(X = b) + P(X = a).$$

Função de distribuição = F(x)

- A função de distribuição de uma variável aleatória discreta permite calcular as probabilidades acumuladas:

$$F(x) = P(X \leq x)$$

  (**observação:** corresponde à coluna das frequências relativas acumuladas nas tabelas de frequências)

- A ideia é ir somando cada uma das probabilidades pontuais. Ou seja, começamos com uma "caixa vazia" e depois colocamos cada uma das probabilidades pontuais uma a uma e a função distribuição vai indicando qual a probabilidade que se vai obtendo à medida que se vai enchendo a "caixa".

- O domínio da função distribuição é sempre $\mathbb{R}$.

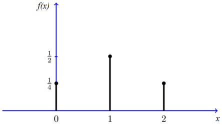
- O menor valor da função de distribuição é sempre zero ("caixa vazia") e o maior valor da função de distribuição é sempre 1 ("caixa cheia").

Função de distribuição = F(x)

- A função distribuição de uma variável aleatória discreta é uma função não decrescente e contínua à direita.
- Graficamente a função distribuição de uma variável aleatória discreta é um gráfico em escada, em que a altura de cada degrau corresponde ao valor da função de probabilidade.
- Podemos ver alguns exemplos de representações de funções de distribuição de variáveis aleatórias discretas:

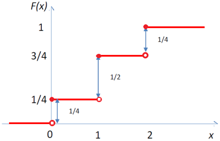
Exemplo 7 X variável aleatória discreta com domínio $D_X = \{0, 1, 2\}$ e com **função de probabilidade**:

x	0	1	2
$f(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$



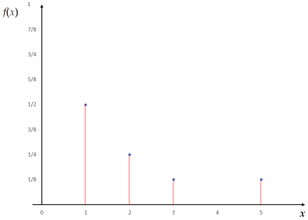
A **função de distribuição** é:

$$F(x) = \begin{cases} 0 & , x < 0 \\ 1/4 & , 0 \leq x < 1 \\ 3/4 & , 1 \leq x < 2 \\ 1 & , x \geq 2 \end{cases}$$



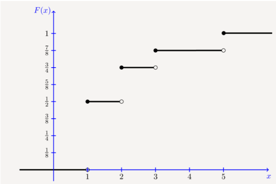
Exemplo 8 X variável aleatória discreta com domínio $D_X = \{1, 2, 3, 5\}$ e com **função de probabilidade**:

x	1	2	3	5
$f(x)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$



A **função de distribuição** é:

$$F(x) = \begin{cases} 0 & , \; x < 1 \\ 1/2 & , \; 1 \leq x < 2 \\ 3/4 & , \; 2 \leq x < 3 \\ 7/8 & , \; 3 \leq x < 5 \\ 1 & , \; x \geq 5 \end{cases}$$



Função de distribuição = F(x)

Exemplo 6

Seja X uma variável aleatória discreta que representa o número de pessoas que visitam, por dia, um determinado site e seja f a sua função de probabilidade representada por:

x	0	1	2	3	4	5
$f(x)$	0.04	0.5	0.24	0.12	0.05	0.05

3. Calcule a função de distribuição.

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i) = \begin{cases} 0 & , \ x < 0 \\ 0.04 & , \ 0 \leq x < 1 \quad (1) \\ 0.54 & , \ 1 \leq x < 2 \quad (2) \\ 0.78 & , \ 2 \leq x < 3 \quad (3) \\ 0.90 & , \ 3 \leq x < 4 \quad (4) \\ 0.95 & , \ 4 \leq x < 5 \quad (5) \\ 1 & , x \geq 5 \quad (6) \end{cases}$$

(1) $F(x) = \sum_{x_i \leq x} f(x_i) = f(0) = 0.04$

(2) $F(x) = \sum_{x_i \leq x} f(x_i) = f(0) + f(1) = F(0) + f(1) = 0.04 + 0.5 = 0.54$

(3) $F(x) = \sum_{x_i \leq x} f(x_i) = f(0) + f(1) + f(2) = F(1) + f(2) = 0.54 + 0.24 = 0.78$

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i) = \begin{cases} 0 & , \ x < 0 \\ 0.04 & , \ 0 \leq x < 1 & (1) \\ 0.54 & , \ 1 \leq x < 2 & (2) \\ 0.78 & , \ 2 \leq x < 3 & (3) \\ 0.90 & , \ 3 \leq x < 4 & (4) \\ 0.95 & , \ 4 \leq x < 5 & (5) \\ 1 & , x \geq 5 & (6) \end{cases}$$

(4) $F(x) = \sum_{x_i \leq x} f(x_i) = f(0) + f(1) + f(2) + f(3) =$

$= F(2) + f(3) = 0.78 + 0.12 = 0.90$

(5) $F(x) = \sum_{x_i \leq x} f(x_i) = f(0) + f(1) + f(2) + f(3) + f(4) =$

$= F(3) + f(4) = 0.90 + 0.05 = 0.95$

(6) $F(x) = \sum_{x_i \leq x} f(x_i) = f(0) + f(1) + f(2) + f(3) + f(4) + f(5) =$

$= F(4) + f(5) = 0.95 + 0.05 = 1$

Exemplo 6

Seja X uma variável aleatória discreta que representa o número de pessoas que visitam, por dia, um determinado site e seja f a sua função de probabilidade representada por:

x	0	1	2	3	4	5
$f(x)$	0.04	0.5	0.24	0.12	0.05	0.05

4. Calcule (sempre que possível use apenas a função de distribuição)

a) $P(X \leq 3)$
b) $P(X < 3)$
c) $P(X = 3)$
d) $P(X > 3)$
e) $P(X \geq 3)$
f) $P(3 < X \leq 5)$
g) $P(3 \leq X \leq 5)$
h) $P(3 \leq X < 5)$
i) $P(3 < X < 5)$

Como a função de distribuição é:

$$F(x) = P(X \leq x) = \begin{cases} 0 & , \ x < 0 \\ 0.04 & , \ 0 \leq x < 1 \\ 0.54 & , \ 1 \leq x < 2 \\ 0.78 & , \ 2 \leq x < 3 \\ 0.90 & , \ 3 \leq x < 4 \\ 0.95 & , \ 4 \leq x < 5 \\ 1 & , x \geq 5 \end{cases}$$

a) $P(X \leq 3) = F(3) = 0.90$

b) $P(X < 3) = P(X \leq 2) = F(2) = 0.78$

c) $P(X = 3) = F(3) - F(2) = 0.90 - 0.78 = 0.12$
c) $P(X = 3) = f(3) = 0.12$

d) $P(X > 3) = 1 - P(X \leq 3) = 1 - F(3) = 1 - 0.9 = 0.1$

e) $P(X \geq 3) = 1 - P(X < 3) = 1 - P(X \leq 2) = 1 - F(2) = 1 - 0.78 = 0.22$

f) $P(3 < X \leq 5) = F(5) - F(3) = 1 - 0.90 = 0.10$

Como a função de distribuição é:

$$F(x) = P(X \leq x) = \begin{cases} 0 & , \ x < 0 \\ 0.04 & , \ 0 \leq x < 1 \\ 0.54 & , \ 1 \leq x < 2 \\ 0.78 & , \ 2 \leq x < 3 \\ 0.90 & , \ 3 \leq x < 4 \\ 0.95 & , \ 4 \leq x < 5 \\ 1 & , x \geq 5 \end{cases}$$

g) $P(3 \leq X \leq 5) = P(2 < X \leq 5) = F(5) - F(2) = 1 - 0.78 = 0.22$

g) $P(3 \leq X \leq 5) = F(5) - F(3) + f(3) = 1 - 0.90 + 0.12 = 0.22$

h) $P(3 \leq X < 5) = P(2 < X \leq 4) = F(4) - F(2) = 0.95 - 0.78 = 0.17$

h) $P(3 \leq X < 5) = F(5) - F(3) - f(5) + f(3) = 1 - 0.90 - 0.05 + 0.12 = 0.17$

i) $P(3 < X < 5) = P(3 < X \leq 4) = F(4) - F(3) = 0.95 - 0.90 = 0.05$

i) $P(3 < X < 5) = F(5) - F(3) - f(5) = 1 - 0.90 - 0.05 = 0.05$

Exemplo 6

Seja X uma variável aleatória discreta que representa o número de pessoas que visitam, por dia, um determinado site e seja f a sua função de probabilidade representada por:

x	0	1	2	3	4	5
$f(x)$	0.04	0.5	0.24	0.12	0.05	0.05

5. Sabendo que em 90% dos dias no máximo k pessoas visitam o site, determine k.

Exemplo 6

Seja X uma variável aleatória discreta que representa o número de pessoas que visitam, por dia, um determinado site e seja f a sua função de probabilidade representada por:

x	0	1	2	3	4	5
$f(x)$	0.04	0.5	0.24	0.12	0.05	0.05

5. Sabendo que em 90% dos dias no máximo k pessoas visitam o site, determine k.

Pretende-se determinar k tal que

$$P(X \leq k) = 0.90 \Leftrightarrow F(k) = 0.90 \Leftrightarrow 3 \leq k < 4$$

Ou seja, $k = 3$.

Exemplo 9

Considere uma variável aleatória, X, com a seguinte função de distribuição:

$$F(x) = \begin{cases} 0 & , \quad x < -2 \\ 0.42 & , \quad -2 \leq x < 1 \\ 0.95 & , \quad 1 \leq x < 3 \\ 1 & , \quad x \geq 3 \end{cases}$$

1. Calcule $P(X > 0.5)$.

Exemplo 9

Considere uma variável aleatória, X, com a seguinte função de distribuição:

$$F(x) = \begin{cases} 0 & , \quad x < -2 \\ 0.42 & , \quad -2 \leq x < 1 \\ 0.95 & , \quad 1 \leq x < 3 \\ 1 & , \quad x \geq 3 \end{cases}$$

1. Calcule $P(X > 0.5)$.

$$P(X > 0.5) = 1 - P(X \leq 0.5) = 1 - F(0.5) = 1 - 0.42 = 0.58$$

Exemplo 9

Considere uma variável aleatória, X, com a seguinte função de distribuição:

$$F(x) = \begin{cases} 0 & , \quad x < -2 \\ 0.42 & , \quad -2 \leq x < 1 \\ 0.95 & , \quad 1 \leq x < 3 \\ 1 & , \quad x \geq 3 \end{cases}$$

2. Determine a função de probabilidade de X.

Exemplo 9

Considere uma variável aleatória, X, com a seguinte função de distribuição:

$$F(x) = \begin{cases} 0 & , \quad x < -2 \\ 0.42 & , \quad -2 \leq x < 1 \\ 0.95 & , \quad 1 \leq x < 3 \\ 1 & , \quad x \geq 3 \end{cases}$$

2. Determine a função de probabilidade de X.

A função de probabilidade é,

x	-2	1	3
$f(x)$	0.42	0.53	0.05

$$f(-2) = P(X = -2) = F(-2) - P(X < -2) = 0.42 - 0 = 0.42$$

$$f(1) = P(X = 1) = F(1) - F(-2) = 0.95 - 0.42 = 0.53$$

$$f(3) = P(X = 3) = F(3) - F(1) = 1 - 0.95 = 0.05$$

Variáveis Aleatórias Discretas

Valor Esperado ou Média ou Esperança Matemática

O valor esperado ou média ou esperança matemática de uma variável aleatória discreta X representa-se por

$$\mu = \mu_X = E[X]$$

e calcula-se

$$\mu = E[X] = \sum_x x f(x)$$

Observação

O valor esperado é um **parâmetro de localização**, que pretende localizar o centro da distribuição de probabilidade, ou seja, pretende identificar o "centro de gravidade" da variável aleatória.

Valor Esperado = E[ ]

- O cálculo do valor esperado de uma variável aleatória discreta X,

$$\mu = E\left[X\right] = \sum_{x} x f\left(x\right)$$

corresponde a uma média ponderada cujos pesos são as probabilidades pontuais.

Exemplo 6

Seja X uma variável aleatória discreta que representa o número de pessoas que visitam, por dia, um determinado site e seja f a sua função de probabilidade representada por:

x	0	1	2	3	4	5
$f(x)$	0.04	0.5	0.24	0.12	0.05	0.05

6. Em média quantas pessoas visitam o site diariamente?

Exemplo 6

Seja X uma variável aleatória discreta que representa o número de pessoas que visitam, por dia, um determinado site e seja f a sua função de probabilidade representada por:

x	0	1	2	3	4	5
$f(x)$	0.04	0.5	0.24	0.12	0.05	0.05

6. Em média quantas pessoas visitam o site diariamente?

Pretende-se o número médio de visitas diárias ao site, $\mu = E[X]$:

$$
\begin{aligned}
E[X] &= \sum_x x f(x) = \\
&= 0 \times f(0) + 1 \times f(1) + 2 \times f(2) + 3 \times f(3) + 4 \times f(4) + 5 \times f(5) = \\
&= 0 \times 0.04 + 1 \times 0.5 + 2 \times 0.24 + 3 \times 0.12 + 4 \times 0.05 + 5 \times 0.05 = \\
&= 1.79 \text{ visitas diárias}
\end{aligned}
$$

Valor Esperado = E[ ]

- Atenção $E[\,]$ representa uma função que calcula uma média, ou seja, calcula a média do que se colocar "dentro" da função.

- No exemplo anterior pretendia-se a média da variável aleatória discreta X, logo a função valor esperado ficou

$$E[X] = \sum_x x f(x)$$

- Se pretender calcular a média de uma transformação da variável aleatória discreta X, por exemplo $g(X)$, então o seu valor médio é

$$E[g(X)] = \sum_x g(x) f(x)$$

.

- Algumas das propriedades da função valor esperado são:

Variáveis Aleatórias Discretas

Propriedades

Sejam X uma variável aleatória discreta e a e b constantes reais.

1. Se $X = a$, então $E[X] = E[a] = a$;
2. $E[aX + b] = aE[X] + b$;
3. Sejam $g(X)$ e $h(X)$ funções de X

$$E[g(X) + h(X)] = E[g(X)] + E[h(X)].$$

Valor Esperado = E[ ]

De forma simplista, as propriedades da função valor esperado são:

1. O valor esperado de uma constante é a própria constante.

2. Se efetuar uma transformação na variável aleatória X que apenas depende de constantes (multiplicar por uma constante e/ou somar uma constante), então para calcular o valor esperado dessa transformação basta efetuar exatamente a mesma transformação no valor esperado da variável aleatória X. Pois apenas efetuou uma mudança de escala na variável aleatória X, logo basta efetuar a mesma mudança de escala no valor esperado da variável aleatória X.

3. O valor esperado da soma de transformações na variável aleatória X é igual a somar os valores esperados de cada uma das transformações.

Valor Esperado = E[ ]

Exemplo 6

Seja X uma variável aleatória discreta que representa o número de pessoas que visitam, por dia, um determinado site e seja f a sua função de probabilidade representada por:

x	0	1	2	3	4	5
$f(x)$	0.04	0.5	0.24	0.12	0.05	0.05

7. Calcule $E[3X-2]$.

Recorrendo à propriedade $E\left[aX+b\right]=aE\left[X\right]+b$ tem-se

$$E\left[3X-2\right]=3E[X]-2=3\times 1.79-2=3.37$$

Claro que também era possível calcular $E\left[3X-2\right]$ sem usar a propriedade:

$$\begin{aligned} E\left[3X-2\right] &= \sum_x (3x-2)\times f\left(x\right) = \\ &= (3\times 0-2)\times f(0)+(3\times 1-2)\times f(1)+(3\times 2-2)\times f(2)+ \\ &+ (3\times 3-2)\times f(3)+(3\times 4-2)\times f(4)+(3\times 5-2)\times f(5) = \\ &= (3\times 0-2)\times 0.04+(3\times 1-2)\times 0.5+(3\times 2-2)\times 0.24+ \\ &+ (3\times 3-2)\times 0.12+(3\times 4-2)\times 0.05+(3\times 5-2)\times 0.05 = \\ &= 3.37 \end{aligned}$$

Variáveis Aleatórias Discretas

Variância

A variância de uma variável aleatória discreta X representa-se por

$$\sigma^2 = \sigma_X^2 = Var\left[X\right] = V\left[X\right] = E\left[(X-\mu)^2\right]$$

e calcula-se

$$\sigma^2 = V\left[X\right] = \sum_x (x-\mu)^2 f(x)$$

Observação

A variância é um **parâmetro de dispersão**. Mede a dispersão (ao quadrado) da variável aleatória em torno do seu valor esperado.

Variância = V[ ]

- A variância corresponde a um valor médio, neste caso ao valor médio das diferenças ao quadrado em relação à média da variável aleatória:

$$\sigma^2 = V\left[X\right] = E\left[\left(X-\mu\right)^2\right]$$

com $\mu = E[X]$.

- Portanto o seu cálculo é simples se recorrermos à observação feita em relação à função valor esperado:

$$E\left[g\left(X\right)\right] = \sum_x g\left(x\right) f\left(x\right)$$

portanto tem-se

$$\sigma^2 = V\left[X\right] = E\left[\left(X-\mu\right)^2\right] = \sum_x \left(x-\mu\right)^2 f\left(x\right)$$

com $\mu = E[X]$.

- A variância goza das seguintes propriedades:

Variáveis Aleatórias Discretas

Propriedades

Sejam X uma variável aleatória discreta e a e b constantes reais.

1. $V\left[X\right] = E\left[X^2\right] - E^2\left[X\right]$;
2. $V\left[X\right] \geq 0$;
3. Se $X = a$, então $V\left[X\right] = V\left[a\right] = 0$;
4. $V\left[aX + b\right] = a^2 V\left[X\right]$.

Variância = V[ ]

De forma simplista, as propriedades da variância são:

1. $V\left[X\right] = E\left[X^2\right] - E^2\left[X\right]$ forma alternativa para calcular a variância.

2. Não há variâncias negativas.

3. A variância de uma constante é nula pois não há qualquer variablidade.

4. Se efetuar uma transformação na variável aleatória X que apenas depende de constantes (multiplicar por uma constante e/ou somar uma constante), então para calcular a variância dessa transformação só interessa a constante que foi multiplicada e basta multiplicar essa constante ao quadrado à variância da variável aleatória X. Pois a dispersão só é alterada quando se multiplica uma constante, a soma de uma constante apenas muda a escala dos dados mas não altera a dispersão dos dados.

Variáveis Aleatórias Discretas

Desvio padrão

O desvio padrão de uma variável aleatória discreta X representa-se por

$$\sigma = \sigma_X$$

e calcula-se

$$\sigma = \sigma_X = \sqrt{V\left[X\right]}$$

Observação

O desvio padrão é um **parâmetro de dispersão**, é a raiz quadrada da variância. Mede a dispersão da variável aleatória em torno do seu valor esperado na mesma unidade de medida em que a variável aleatória vem expressa.

Variância e Desvio Padrão

Exemplo 6

Seja X uma variável aleatória discreta que representa o número de pessoas que visitam, por dia, um determinado site e seja f a sua função de probabilidade representada por:

x	0	1	2	3	4	5
$f(x)$	0.04	0.5	0.24	0.12	0.05	0.05

8. Calcule a variância da variável aleatória X.

Pretende-se a variância de X, $\sigma^2 = V[X]$:

- Usando a **definição** tem-se:

$$V[X] = E\left[(X-\mu)^2\right] = \sum_x (x-\mu)^2 f(x) = \qquad \text{(tem-se } \mu = E[X] = 1.79\text{)}$$

$$= E\left[(X-1.79)^2\right] = \sum_x (x-1.79)^2 f(x) =$$

$$= (0-1.79)^2 \times f(0) + (1-1.79)^2 \times f(1) + (2-1.79)^2 \times f(2) +$$

$$+ (3-1.79)^2 \times f(3) + (4-1.79)^2 \times f(4) + (5-1.79)^2 \times f(5) =$$

$$= (0-1.79)^2 \times 0.04 + (1-1.79)^2 \times 0.5 + (2-1.79)^2 \times 0.24 +$$

$$+ (3-1.79)^2 \times 0.12 + (4-1.79)^2 \times 0.05 + (5-1.79)^2 \times 0.05 =$$

$$= 1.3859 \text{ visitas diárias}^2$$

Pretende-se a variância de X, $\sigma^2 = V[X]$:

- Usando a **propriedade** $V[X] = E\left[X^2\right] - E^2\left[X\right]$ tem-se:

$$E[X] = 1.79$$

$$\begin{aligned} E[X^2] &= \sum_x x^2 f\left(x\right) = \\ &= 0^2 \times f(0) + 1^2 \times f(1) + 2^2 \times f(2) + 3^2 \times f(3) + 4^2 \times f(4) + 5^2 \times f(5) = \\ &= 0^2 \times 0.04 + 1^2 \times 0.5 + 2^2 \times 0.24 + 3^2 \times 0.12 + 4^2 \times 0.05 + 5^2 \times 0.05 = \\ &= 4.59 \end{aligned}$$

$V[X] = E\left[X^2\right] - E^2\left[X\right] = 4.59 - 1.79^2 = 1.3859$ visitas diárias^2

Variância e Desvio Padrão

Exemplo 6

Seja X uma variável aleatória discreta que representa o número de pessoas que visitam, por dia, um determinado site e seja f a sua função de probabilidade representada por:

x	0	1	2	3	4	5
$f(x)$	0.04	0.5	0.24	0.12	0.05	0.05

9. Calcule o desvio padrão da variável aleatória X.

Variância e Desvio Padrão

Exemplo 6

Seja X uma variável aleatória discreta que representa o número de pessoas que visitam, por dia, um determinado site e seja f a sua função de probabilidade representada por:

x	0	1	2	3	4	5
$f(x)$	0.04	0.5	0.24	0.12	0.05	0.05

9. Calcule o desvio padrão da variável aleatória X.

Pretende-se o desvio padrão de X:

$$\sigma = \sqrt{V[X]} = \sqrt{1.3859} = 1.1772 \text{ visitas diárias}$$

Exemplo 6

Seja X uma variável aleatória discreta que representa o número de pessoas que visitam, por dia, um determinado site e seja f a sua função de probabilidade representada por:

x	0	1	2	3	4	5
$f(x)$	0.04	0.5	0.24	0.12	0.05	0.05

10. Calcule $V[-3X-2]$.

Exemplo 6

Seja X uma variável aleatória discreta que representa o número de pessoas que visitam, por dia, um determinado site e seja f a sua função de probabilidade representada por:

x	0	1	2	3	4	5
$f(x)$	0.04	0.5	0.24	0.12	0.05	0.05

10. Calcule $V[-3X-2]$.

Recorrendo à propriedade $V[aX+b] = a^2 V[X]$ tem-se

$$V[-3X-2] = (-3)^2 V[X] = 9 \times 1.3859 = 12.4731$$

Exemplo 9

Considere uma variável aleatória, X, com a seguinte função de distribuição:

$$F(x) = \begin{cases} 0 & , \quad x < -2 \\ 0.42 & , \quad -2 \leq x < 1 \\ 0.95 & , \quad 1 \leq x < 3 \\ 1 & , \quad x \geq 3 \end{cases}$$

3. Calcule $E[1 - 2X]$.

Exemplo 9

Considere uma variável aleatória, X, com a seguinte função de distribuição:

$$F(x) = \begin{cases} 0 & , \quad x < -2 \\ 0.42 & , \quad -2 \leq x < 1 \\ 0.95 & , \quad 1 \leq x < 3 \\ 1 & , \quad x \geq 3 \end{cases}$$

3. Calcule $E[1-2X]$.

Para calcular $E[X]$ tem de conhecer-se a função de probabilidade logo,

x	-2	1	3
$f(x)$	0.42	0.53	0.05

$$E[X] = \sum_x x f(x) = -2 \times 0.42 + 1 \times 0.53 + 3 \times 0.05 = -0.16$$

Tem-se,

$$E[1-2X] = 1 - 2E[X] = 1 - 2 \times (-0.16) = 1.32.$$

- Agora que já sabem caracterizar todo o modelo probabilístico associado às variáveis aleatórias discretas, vamos "dar nomes" a alguns desses modelos probabilísticos.

- Vamos analisar pormenorizadamente três modelos probabilísticos ou, como é mais usual dizer, três Distribuições Teóricas Discretas:
    - Distribuição Uniforme Discreta;
    - Distribuição Binomial;
    - Distribuição de Poisson.

Distribuições Teóricas Discretas:

- **Distribuição Uniforme Discreta**;
- Distribuição Binomial;
- Distribuição de Poisson.

Distribuição Uniforme Discreta

Definição

Diz-se que uma variável aleatória discreta X definida em $D_X = \{x_1, x_2, \ldots, x_n\}$ tem **distribuição Uniforme Discreta** e representa-se por

$$X \sim U_{(n)}$$

se **assume em todos os n pontos do seu domínio a mesma probabilidade**, ou seja, se a sua função de probabilidade é dada por

x	x_1	x_2	$\ldots$	x_n
$f(x)$	$\frac{1}{n}$	$\frac{1}{n}$	$\ldots$	$\frac{1}{n}$

ou, de forma análoga,

$$f(x) = P(X = x) = \begin{cases} \frac{1}{n} & , \; x \in D_X \\ 0 & , \; \text{caso contrário} \end{cases}$$

Distribuição Uniforme Discreta

Definição (continuação)

e a sua função de distribuição é dada por

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i) = \sum_{x_i \leq x} \frac{1}{n} = \begin{cases} 0 & , \ x < x_1 \\ \frac{1}{n} & , \ x_1 \leq x < x_2 \\ \frac{2}{n} & , \ x_2 \leq x < x_3 \\ \frac{3}{n} & , \ x_3 \leq x < x_4 \\ \vdots & \vdots \\ \frac{n-1}{n} & \ x_{n-1} \leq x < x_n \\ 1 & , \ x \geq x_n \end{cases}$$

Distribuição Uniforme Discreta

Teorema

Se a variável aleatória X tem distribuição Uniforme Discreta, $X \sim U_{(n)}$, então

$$\mu = E\left[X\right] = \sum_{i=1}^{n} x_i f\left(x_i\right) = \sum_{i=1}^{n} x_i \times \frac{1}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

e

$$\sigma^2 = V\left[X\right] = \sum_{i=1}^{n} \left(x_i - \mu\right)^2 f\left(x_i\right) = \sum_{i=1}^{n} \left(x_i - \mu\right)^2 \frac{1}{n} = \frac{1}{n}\sum_{i=1}^{n} \left(x_i - \mu\right)^2 =$$
$$= \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2$$

Distribuição Uniforme Discreta

Teorema

Se a variável aleatória X tem distribuição Uniforme Discreta, $X \sim U_{(n)}$, então

$$E[X] = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \text{e} \qquad V[X] = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2$$

Caso Particular

Se a variável aleatória X, com distribuição Uniforme Discreta, está definida num conjunto de **inteiros consecutivos**, $D_X = \{a, a+1, a+2, \ldots, b\}$, então

$$E[X] = \frac{a+b}{2} \qquad \text{e} \qquad V[X] = \frac{(b-a+1)^2 - 1}{12}.$$

Distribuição Uniforme Discreta

- No ensino secundário fizeram, provavelmente, imensos exercícios do tipo:
  (i) "Qual a probabilidade de retirar o às de ouros de um baralho (não viciado) de 52 cartas?"
  (ii) "Qual a probabilidade de lançar um dado (não viciado) e sair a face 4?"

- para calcular as probabilidades pedidas limitavam-se a fazer contagens:
  (i) número de casos possíveis = 52, número de casos favoráveis = 1, então a probabilidade pedida é $\frac{1}{52}$.
  (ii) número de casos possíveis = 6, número de casos favoráveis = 1, então a probabilidade pedida é $\frac{1}{6}$

- Este método (definição clássica de probabilidade) só é possível por estarem a considerar (sem definir) que têm uma variável aleatória X que assume em todos os n pontos do seu domínio a mesma probabilidade, ou seja, uma variável aleatória X com distribuição Uniforme discreta:

$$X \sim U_{(n)}$$

  portanto todos esses exercícios foram feitos usando a distribuição uniforme discreta (só não foi usado este nome por estarem a trabalhar com acontecimentos em vez de variáveis aleatórias).

Exemplo 10

Um sistema de comunicações de uma empresa possui 4 linhas externas. Admita que cada uma das possibilidades, em termos do número de linhas externas em utilização, tem a mesma probabilidade de ocorrer.

1. Indique e represente graficamente a função de probabilidade da variável aleatória número de linhas externas em utilização.

Exemplo 10

Um sistema de comunicações de uma empresa possui 4 linhas externas. Admita que cada uma das possibilidades, em termos do número de linhas externas em utilização, tem a mesma probabilidade de ocorrer.

1. Indique e represente graficamente a função de probabilidade da variável aleatória número de linhas externas em utilização.

Seja X a variável aleatória discreta que representa o número de linhas externas em utilização, definida em $D_X = \{0, 1, 2, 3, 4\}$.

O enunciado diz "Admita que cada uma das possibilidades, em termos do número de linhas externas em utilização, tem a mesma probabilidade de ocorrer", então X segue uma distribuição Uniforme Discreta com 5 elementos, ou seja,

$$X \sim U_{(5)}$$

Agora que identificámos o modelo, é fácil responder à questão:

$X =$ número de linhas externas em utilização, com $D_X = \{0, 1, 2, 3, 4\}$

$$X \sim U_{(5)}$$

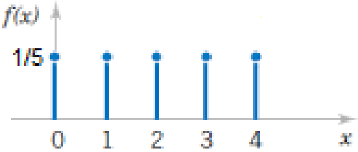
função de probabilidade:

$$f(x) = P(X = x)$$

x	0	1	2	3	4
$f(x)$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

$\Leftrightarrow$

$$f(x) = \begin{cases} \frac{1}{5} & , x \in D_X \\ 0 & , \text{caso contrário} \end{cases}$$



Observação: O gráfico da função de probabilidade de uma distribuição Uniforme Discreta tem sempre as barras todas com a mesma altura.

Exemplo 10

Um sistema de comunicações de uma empresa possui 4 linhas externas. Admita que cada uma das possibilidades, em termos do número de linhas externas em utilização, tem a mesma probabilidade de ocorrer.

2. Indique e represente graficamente a função de distribuição da variável aleatória número de linhas externas em utilização.

Exemplo 10

Um sistema de comunicações de uma empresa possui 4 linhas externas. Admita que cada uma das possibilidades, em termos do número de linhas externas em utilização, tem a mesma probabilidade de ocorrer.

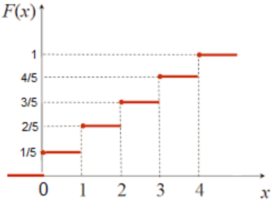
2. Indique e represente graficamente a função de distribuição da variável aleatória número de linhas externas em utilização.

$X =$ número de linhas externas em utilização, com $D_X = \{0, 1, 2, 3, 4\}$

$$X \sim U_{(5)}$$

Função de distribuição: $F(x) = P(X \leq x)$

$$F(x) = \begin{cases} 0 & , \ x < 0 \\ \frac{1}{5} & , \ 0 \leq x < 1 \\ \frac{2}{5} & , \ 1 \leq x < 2 \\ \frac{3}{5} & , \ 2 \leq x < 3 \\ \frac{4}{5} & \ 3 \leq x < 4 \\ 1 & , \ x \geq 4 \end{cases}$$



Observação: Todos os "degraus" do gráfico da função de distribuição de uma distribuição Uniforme Discreta têm sempre a mesma altura. Neste caso $\frac{1}{5}$.

Exemplo 10

Um sistema de comunicações de uma empresa possui 4 linhas externas. Admita que cada uma das possibilidades, em termos do número de linhas externas em utilização, tem a mesma probabilidade de ocorrer.

3. Qual a probabilidade do número de linhas externas em utilização ser inferior a 3 sabendo que existem linhas externas em utilização?

Exemplo 10

Um sistema de comunicações de uma empresa possui 4 linhas externas. Admita que cada uma das possibilidades, em termos do número de linhas externas em utilização, tem a mesma probabilidade de ocorrer.

3. Qual a probabilidade do número de linhas externas em utilização ser inferior a 3 sabendo que existem linhas externas em utilização?

$$P(X < 3 | X > 0) = \frac{P(X < 3 \wedge X > 0)}{P(X > 0)} = \frac{P(0 < X < 3)}{1 - P(X \leq 0)} = \frac{f(1) + f(2)}{1 - F(0)} =$$

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

Exemplo 10

Um sistema de comunicações de uma empresa possui 4 linhas externas. Admita que cada uma das possibilidades, em termos do número de linhas externas em utilização, tem a mesma probabilidade de ocorrer.

4. Qual o número médio de linhas externas em utilização?

Exemplo 10

Um sistema de comunicações de uma empresa possui 4 linhas externas. Admita que cada uma das possibilidades, em termos do número de linhas externas em utilização, tem a mesma probabilidade de ocorrer.

4. Qual o número médio de linhas externas em utilização?

$X =$ número de linhas externas em utilização, com $D_X = \{0, 1, 2, 3, 4\}$

$$X \sim U_{(5)}$$

Como $D_X = \{0, 1, 2, 3, 4\}$ são inteiros consecutivos, então podemos recorrer ao caso particular com $a = 0$ e $b = 4$:

Se a variável aleatória X, com distribuição Uniforme Discreta, está definida num conjunto de **inteiros consecutivos**, $D_X = \{a, a+1, a+2, \ldots, b\}$, então

$$E[X] = \frac{a+b}{2} \quad \text{e} \quad V[X] = \frac{(b-a+1)^2 - 1}{12}.$$

$$E[X] = \frac{0+4}{2} = \frac{4}{2} = 2 \text{ linhas externas}$$

Exemplo 10

Um sistema de comunicações de uma empresa possui 4 linhas externas. Admita que cada uma das possibilidades, em termos do número de linhas externas em utilização, tem a mesma probabilidade de ocorrer.

5. Qual o desvio padrão do número de linhas externas em utilização?

Exemplo 10

Um sistema de comunicações de uma empresa possui 4 linhas externas. Admita que cada uma das possibilidades, em termos do número de linhas externas em utilização, tem a mesma probabilidade de ocorrer.

5. Qual o desvio padrão do número de linhas externas em utilização?

$X =$ número de linhas externas em utilização, com $D_X = \{0, 1, 2, 3, 4\}$

$$X \sim U_{(5)}$$

Como $D_X = \{0, 1, 2, 3, 4\}$ são inteiros consecutivos, então podemos recorrer ao caso particular com $a = 0$ e $b = 4$:

Se a variável aleatória X, com distribuição Uniforme Discreta, está definida num conjunto de **inteiros consecutivos**, $D_X = \{a, a+1, a+2, \ldots, b\}$, então

$$E[X] = \frac{a+b}{2} \qquad \text{e} \qquad V[X] = \frac{(b-a+1)^2 - 1}{12}.$$

variância: $V[X] = \frac{(4-0+1)^2-1}{12} = \frac{24}{12} = 2$ linhas externas^2

desvio padrão: $\sqrt{V[X]} = \sqrt{2} = 1.414$ linhas externas

Claro que as alíneas anteriores podiam ter sido calculadas com recurso às definições e propriedades que vimos para as Variáveis Aleatórias Discretas.

$$X \sim U_{(5)}, \text{ com } D_X = \{0, 1, 2, 3, 4\}$$

função de probabilidade:

$$f(x) = \begin{cases} \frac{1}{5} & , \ x \in D_X \\ 0 & , \text{ caso contrário} \end{cases}$$

$$E[X] = \sum_x x f(x) = 0 \times f(0) + 1 \times f(1) + 2 \times f(2) + 3 \times f(3) + 4 \times f(4) =$$
$$= 0 \times \frac{1}{5} + 1 \times \frac{1}{5} + 2 \times \frac{1}{5} + 3 \times \frac{1}{5} + 4 \times \frac{1}{5} = \frac{1}{5} \times (0 + 1 + 2 + 3 + 4) = 2$$

$$E[X^2] = \sum_x x^2 f(x) = 0^2 \times f(0) + 1^2 \times f(1) + 2^2 \times f(2) + 3^2 \times f(3) + 4^2 \times f(4) =$$
$$= 0^2 \times \frac{1}{5} + 1^2 \times \frac{1}{5} + 2^2 \times \frac{1}{5} + 3^2 \times \frac{1}{5} + 4^2 \times \frac{1}{5} = \frac{1}{5} \times \left(0^2 + 1^2 + 2^2 + 3^2 + 4^2\right) = 6$$

$$V[X] = E[X^2] - E^2[X] = 6 - 2^2 = 2$$

$$\sigma = \sqrt{V[X]} = \sqrt{2} = 1.414$$

Exemplo 11

Seja X uma variável aleatória com distribuição Uniforme Discreta, definida no conjunto $D_X = \{3, 6, 9\}$. Calcule o valor esperado e a variância da variável aleatória X.

Exemplo 11

Seja X uma variável aleatória com distribuição Uniforme Discreta, definida no conjunto $D_X = \{3, 6, 9\}$. Calcule o valor esperado e a variância da variável aleatória X.

$$X \sim U_{(3)}, \text{ com } D_X = \{3, 6, 9\}$$

Neste caso não é possível utilizar o "caso particular" pois $D_X = \{3, 6, 9\}$, embora seja formado por inteiros, **não são inteiros consecutivos**.
No entanto como é possível escrever:

- $3 = 3 \times 1$
- $6 = 3 \times 2$
- $9 = 3 \times 3$

então é possível considerar que $X = 3Y$ e a variável aleatória Y tem um comportamento probabilístico igual à variável X, ou seja

$$Y \sim U_{(3)}$$

mas $D_Y = \{1, 2, 3\}$.

$$Y \sim U_{(3)}, \text{ com } D_Y = \{1, 2, 3\}$$

Como $D_Y = \{1, 2, 3\}$ são inteiros consecutivos, então podemos recorrer ao caso particular com $a = 1$ e $b = 3$:

Se a variável aleatória X, com distribuição Uniforme Discreta, está definida num conjunto de **inteiros consecutivos**, $D_X = \{a, a+1, a+2, \ldots, b\}$, então

$$E[X] = \frac{a+b}{2} \qquad \text{e} \qquad V[X] = \frac{(b-a+1)^2 - 1}{12}.$$

$$E[Y] = \frac{1+3}{2} = 2 \qquad \text{e} \qquad V[Y] = \frac{(3-1+1)^2 - 1}{12} = \frac{2}{3}$$

Agora basta usar as propriedades de valor esperado e da variância:

$$E[X] = E[3Y] = 3E[Y] = 3 \times 2 = 6$$

$$V[X] = V[3Y] = 3^2 V[Y] = 9 \times \frac{2}{3} = 6$$

Claro que podia ter sido calculado por definição e com as propriedades que vimos nas Variáveis Aleatórias Discretas.

$$X \sim U_{(3)}, \text{ com } D_X = \{3, 6, 9\}$$

função de probabilidade:

$$f(x) = \begin{cases} \frac{1}{3} & , \ x \in D_X \\ 0 & , \ \text{caso contrário} \end{cases}$$

$$E[X] = \sum_x x f(x) = 3 \times f(3) + 6 \times f(6) + 9 \times f(9) = 3 \times \frac{1}{3} + 6 \times \frac{1}{3} + 9 \times \frac{1}{3} = \frac{1}{3} \times (3 + 6 + 9) = \\ = 6$$

$$E[X^2] = \sum_x x^2 f(x) = 3^2 \times f(3) + 6^2 \times f(6) + 9^2 \times f(9) = 3^2 \times \frac{1}{3} + 6^2 \times \frac{1}{3} + 9^2 \times \frac{1}{3} = \\ = \frac{1}{3} \times \left(3^2 + 6^2 + 9^2\right) = 42$$

$$V[X] = E[X^2] - E^2[X] = 42 - 6^2 = 6$$

Exemplo 12

Considere uma variável aleatória X que assume a mesma probabilidade em todos os pontos do conjunto $\{2, 5, 7, 11\}$. Calcule o valor esperado e a variância da variável aleatória X.

Exemplo 12

Considere uma variável aleatória X que assume a mesma probabilidade em todos os pontos do conjunto $\{2, 5, 7, 11\}$. Calcule o valor esperado e a variância da variável aleatória X.

O enunciado diz

"...que assume a mesma probabilidade em todos os pontos..."

então X segue uma distribuição Uniforme Discreta com 4 elementos, ou seja,

$$X \sim U_{(4)}, \text{ com } D_X = \{2, 5, 7, 11\}$$

Neste caso não é possível utilizar o "caso particular" pois $D_X = \{2, 5, 7, 11\}$, embora seja formado por inteiros, **não são inteiros consecutivos** nem é fácil de estabelecer uma relação que permita construir uma nova variável cujo domínio sejam inteiros consecutivos.

Então a única possibilidade é recorrer às definições e propriedades que vimos nas Variáveis Aleatórias Discretas.

$$X \sim U_{(4)}, \text{ com } D_X = \{2, 5, 7, 11\}$$

função de probabilidade:

$$f(x) = \begin{cases} \frac{1}{4} & , \ x \in D_X \\ 0 & , \ \text{caso contrário} \end{cases}$$

$$E[X] = \sum_x x f(x) = 2 \times f(2) + 5 \times f(5) + 7 \times f(7) + 11 \times f(11) =$$
$$= 2 \times \frac{1}{4} + 5 \times \frac{1}{4} + 7 \times \frac{1}{4} + 11 \times \frac{1}{4} = \frac{1}{4} \times (2 + 5 + 7 + 11) = 6.25$$

$$E[X^2] = \sum_x x^2 f(x) = 2^2 \times f(2) + 5^2 \times f(5) + 7^2 \times f(7) + 11^2 \times f(11) =$$
$$= 2^2 \times \frac{1}{4} + 5^2 \times \frac{1}{4} + 7^2 \times \frac{1}{4} + 11^2 \times \frac{1}{4} = \frac{1}{4} \times \left(2^2 + 5^2 + 7^2 + 11^2\right) = 49.75$$

$$V[X] = E[X^2] - E^2[X] = 49.75 - 6.25^2 = 10.69$$

Distribuições Teóricas Discretas:

- Distribuição Uniforme Discreta;
- **Distribuição Binomial**;
- Distribuição de Poisson.

Distribuição Binomial

Provas de Bernoulli

É uma experiência aleatória que tem apenas dois resultados possíveis: sucesso ou insucesso.

O sucesso ocorre com probabilidade p (fixo),

$$P\left(sucesso\right) = p$$

e o insucesso com probabilidade $q = 1 - p$,

$$P\left(insucesso\right) = q = 1 - p.$$

Distribuição Binomial

Exemplo 13

Prova de Bernoulli:

- Experiência aleatória: retirar uma carta de um baralho (não viciado) com 52 cartas

- sucesso: sair carta de copas
- insucesso: não sair carta de copas

- O sucesso ocorre com probabilidade $p = \frac{13}{52} = 0.25$

- O insucesso ocorre com probabilidade $q = 1 - p = 1 - \frac{13}{52} = 0.75$,

Distribuição Binomial

Experiência Binomial

É uma sucessão de provas de Bernoulli e caracteriza-se por:

- a experiência ser constituída por n provas de Bernoulli, em que uma prova é uma repetição em condições idênticas;
- as provas são independentes;
- em cada prova pode-se realizar um dos dois acontecimentos possíveis:

  sucesso ou insucesso;

  onde

$$P\,(sucesso) = p \qquad \text{e} \qquad P\,(insucesso) = q = 1 - p$$

Distribuição Binomial

Exemplo 14

Experiência Binomial: repetir $n = 4$ vezes a seguinte Prova de Bernoulli:

- Experiência aleatória: retirar uma carta de um baralho (não viciado) com 52 cartas
- sucesso: sair carta de copas
- insucesso: não sair carta de copas
- O sucesso ocorre com probabilidade $p = \frac{13}{52} = 0.25$
- O insucesso ocorre com probabilidade $q = 1 - p = 1 - \frac{13}{52} = 0.75$,

Distribuição Binomial

Definição

Diz-se que uma variável aleatória discreta X, que representa o número de sucessos em n provas de Bernoulli, tem **distribuição Binomial** com os parâmetros n e p (fixos)

$$X \sim B\left(n, p\right)$$

se a sua função de probabilidade é dada por

$$f\left(x\right) = P\left(X = x\right) = \begin{cases} {}^{n}C_{x} p^{x} \left(1-p\right)^{n-x} & , \ x = 0, 1, \ldots, n \\ 0 & , \ \text{caso contrário} \end{cases}$$

e a sua função de distribuição é dada por

$$F\left(x\right) = P\left(X \leq x\right) = \sum_{x_i \leq x} \left({}^{n}C_{x_i} p^{x_i} \left(1-p\right)^{n-x_i} \right)$$

onde $0 < p < 1$ representa a probabilidade de sucesso numa prova de Bernoulli.

Distribuição de Binomial

Exemplo 14

Considere a variável a aleatória discreta:

$X =$ número de vezes que sai carta de copas quando se tira 4 cartas do baralho

ou seja,

$X =$ número de sucessos em n provas de Bernoulli

Portanto a variável aleatória X tem $D_X = \{0, 1, 2, 3, 4\}$ e tem distribuição Binomial com os parâmetros $n = 4$ e $p = 0.25$, ou seja

$$X \sim B\left(4, 0.25\right)$$

Observação: Na distribuição Binomial está subjacente que houve reposição, por isso é que a probabilidade de sucesso é fixa.

Distribuição Binomial

Exemplo 14

$X =$ número de vezes que sai carta de copas quando se tira 4 cartas do baralho

$$X \sim B\,(4, 0.25)$$

função de probabilidade $f(x) = P(X = x)$:

x	0	1	2	3	4
$f\,(x)$	0.3164	0.4219	0.2109	0.0469	0.0039

função de distribuição $F(x) = P(X \leq x)$:

$$F\,(x) = \begin{cases} 0 & , \ x < 0 \\ 0.3164 & , \ 0 \leq x < 1 \\ 0.7383 & , \ 1 \leq x < 2 \\ 0.9492 & , \ 2 \leq x < 3 \\ 0.9961 & , \ 3 \leq x < 4 \\ 1 & , \ x \geq 4 \end{cases}$$

Para calcular os valores da função de probabilidade e da função de distribuição vamos ver duas possibilidades:

1. **cálculos diretos**

Se $X \sim B(n, p)$, a função de probabilidade é

$$f(x) = P(X = x) = \begin{cases} {}^nC_x p^x (1-p)^{n-x} & , \ x = 0, 1, \ldots, n \\ 0 & , \ \text{caso contrário} \end{cases}$$

No exemplo 14 tem-se $X \sim B(4, 0.25)$, então

$$f(0) = P(X = 0) = {}^4C_0 \times 0.25^0 \times (1 - 0.25)^{4-0} = 0.3164$$
$$f(1) = P(X = 1) = {}^4C_1 \times 0.25^1 \times (1 - 0.25)^{4-1} = 0.4219$$
$$f(2) = P(X = 2) = {}^4C_2 \times 0.25^2 \times (1 - 0.25)^{4-2} = 0.2109$$
$$f(3) = P(X = 3) = {}^4C_3 \times 0.25^3 \times (1 - 0.25)^{4-3} = 0.0469$$
$$f(4) = P(X = 4) = {}^4C_4 \times 0.25^4 \times (1 - 0.25)^{4-4} = 0.0039$$

Agora basta construir a função distribuição com base na função de probabilidade.

2. **tabelas com recurso ao R**:
   - função de probabilidade: **dbinom**$(x, size = n, prob = p)$
   - função de distribuição: **pbinom**$(x, size = n, prob = p)$
   - inversa da função de distribuição: **qbinom**$(p, size = n, prob = p)$

No exemplo 14 tem-se $X \sim B\,(4, 0.25)$, então

$$
\begin{aligned}
f(0) &= P(X = 0) = dbinom(0, 4, 0.25) = 0.3164 \\
f(1) &= P(X = 1) = dbinom(1, 4, 0.25) = 0.4219 \\
f(2) &= P(X = 2) = dbinom(2, 4, 0.25) = 0.2109 \\
f(3) &= P(X = 3) = dbinom(3, 4, 0.25) = 0.0469 \\
f(4) &= P(X = 4) = dbinom(4, 4, 0.25) = 0.0039
\end{aligned}
$$

$$
\begin{aligned}
F(0) &= P(X \leq 0) = pbinom(0, 4, 0.25) = 0.3164 \\
F(1) &= P(X \leq 1) = pbinom(1, 4, 0.25) = 0.7383 \\
F(2) &= P(X \leq 2) = pbinom(2, 4, 0.25) = 0.9492 \\
F(3) &= P(X \leq 3) = pbinom(3, 4, 0.25) = 0.9961 \\
F(4) &= P(X \leq 4) = pbinom(4, 4, 0.25) = 1
\end{aligned}
$$

Distribuição Binomial

Teorema

Se a variável aleatória X tem distribuição Binomial com parâmetros n e p,

$$X \sim B(n, p)$$

então

$$E[X] = np \qquad \text{e} \qquad V[X] = npq = np(1-p).$$

Distribuição Binomial

Propriedade: Aditividade da Binomial

Sejam $X_1, X_2, \ldots, X_k$ variáveis aleatórias independentes com distribuição Binomial e com a mesma probabilidade de sucesso, isto é

$$X_i \sim B\left(n_i, p\right) \quad i = 1, \ldots, k,$$

então

$$Y = \sum_{i=1}^{k} X_i$$

também é uma variável aleatória com distribuição Binomial, isto é

$$Y = \sum_{i=1}^{k} X_i \sim B\left(\sum_{i=1}^{k} n_i, p\right).$$

Exemplo 15

Sabe-se que 10% dos testes feitos para detetar uma determinada doença dá falsos negativos quando são efetuados no laboratório A. Considere cinco testes efetuados no laboratório A de forma independente.

1. Calcule a probabilidade de 3 desses testes serem falsos negativos.

Exemplo 15

Sabe-se que 10% dos testes feitos para detetar uma determinada doença dá falsos negativos quando são efetuados no laboratório A. Considere cinco testes efetuados no laboratório A de forma independente.

1. Calcule a probabilidade de 3 desses testes serem falsos negativos.

$X =$ número de testes que são falsos negativos, num grupo de cinco testes efetuados no laboratório A

A variável aleatória discreta X tem $D_X = \{0, 1, 2, 3, 4, 5\}$ e tem distribuição Binomial com os parâmetros $n = 5$ e $p = 0.10$, ou seja

$$X \sim B\,(5, 0.10)$$

$$P(X = 3) = f(3) = 0.0081$$

1. **cálculo direto**

$$f(3) = {}^5C_3 \times 0.10^3 \times (1 - 0.10)^{5-3} = 0.0081$$

2. **tabelas com recurso ao R**

$$f(3) = dbinom(3, 5, 0.10) = 0.0081$$

Exemplo 15

Sabe-se que 10% dos testes feitos para detetar uma determinada doença dá falsos negativos quando são efetuados no laboratório A. Considere cinco testes efetuados no laboratório A de forma independente.

2. Calcule a probabilidade de mais de 2 testes serem falsos negativos.

Exemplo 15

Sabe-se que 10% dos testes feitos para detetar uma determinada doença dá falsos negativos quando são efetuados no laboratório A. Considere cinco testes efetuados no laboratório A de forma independente.

2. Calcule a probabilidade de mais de 2 testes serem falsos negativos.

$X =$ número de testes que são falsos negativos, num grupo de cinco testes efetuados no laboratório A

$$X \sim B\,(5, 0.10)$$

$$P(X > 2) = 1 - P(X \leq 2) = 1 - F(2) = 1 - 0.9914 = 0.0086$$

1. **cálculo direto**

$$F(2) = f(0) + f(1) + f(2) =$$
$$= {}^5C_0 \times 0.10^0 \times (1-0.10)^{5-0} + {}^5C_1 \times 0.10^1 \times (1-0.10)^{5-1} + {}^5C_2 \times 0.10^2 \times (1-0.10)^{5-2} =$$
$$= 0.9914$$

1. **tabelas com recurso ao R**

$F(2) = pbinom(2, 5, 0.10) = 0.9914$

Exemplo 15

Sabe-se que 10% dos testes feitos para detetar uma determinada doença dá falsos negativos quando são efetuados no laboratório A. Considere cinco testes efetuados no laboratório A de forma independente.

3. Determine k sabendo que a probabilidade de no máximo ter k testes falsos negativos é 0.75?

Exemplo 15

Sabe-se que 10% dos testes feitos para detetar uma determinada doença dá falsos negativos quando são efetuados no laboratório A. Considere cinco testes efetuados no laboratório A de forma independente.

3. Determine k sabendo que a probabilidade de no máximo ter k testes falsos negativos é 0.75?

$X =$ número de testes que são falsos negativos, num grupo de cinco testes efetuados no laboratório A

$$X \sim B\left(5, 0.10\right)$$

Pretende-se determinar k tal que

$$P(X \leq k) = 0.75 \Leftrightarrow F(k) = 0.75 \Leftrightarrow k = F^{-1}(0.75) = 1 \text{ teste}$$

pois $F^{-1}(0.75) = qbinom(0.75, 5, 0.10) = 1$

Exemplo 15

Sabe-se que 10% dos testes feitos para detetar uma determinada doença dá falsos negativos quando são efetuados no laboratório A. Considere cinco testes efetuados no laboratório A de forma independente.

4. Calcule, em média, quantos dos 5 testes espera que sejam falsos negativos.

Exemplo 15

Sabe-se que 10% dos testes feitos para detetar uma determinada doença dá falsos negativos quando são efetuados no laboratório A. Considere cinco testes efetuados no laboratório A de forma independente.

4. Calcule, em média, quantos dos 5 testes espera que sejam falsos negativos.

$X =$ número de testes que são falsos negativos, num grupo de cinco testes efetuados no laboratório A

$$X \sim B\,(5, 0.10)$$

$$E[X] = np = 5 \times 0.10 = 0.5 \text{ testes}$$

Exemplo 15

Sabe-se que 10% dos testes feitos para detetar uma determinada doença dá falsos negativos quando são efetuados no laboratório A. Considere cinco testes efetuados no laboratório A de forma independente.

5. Se os testes forem efetuados num laboratório C, apenas 10% dos testes são falsos negativos. Considere 4 testes efetuados de forma independente no laboratório C. Calcule a probabilidade de, entre os 9 testes (5 do laboratório A e 4 do laboratório C), serem recebidos entre 4 e 6 testes (inclusive) falsos negativos.

Exemplo 15

Sabe-se que 10% dos testes feitos para detetar uma determinada doença dá falsos negativos quando são efetuados no laboratório A. Considere cinco testes efetuados no laboratório A de forma independente.

5. Se os testes forem efetuados num laboratório C, apenas 10% dos testes são falsos negativos. Considere 4 testes efetuados de forma independente no laboratório C. Calcule a probabilidade de, entre os 9 testes (5 do laboratório A e 4 do laboratório C), serem recebidos entre 4 e 6 testes (inclusive) falsos negativos.

$X =$ número de testes que são falsos negativos, num grupo de cinco testes efetuados no laboratório A , $X \sim B\,(5, 0.10)$

$V =$ número de testes que são falsos negativos, num grupo de quatro testes efetuados no laboratório C, $V \sim B\,(4, 0.10)$

Pretende-se $T = X + V$, como X e V podem ser consideradas variáveis aleatórias independentes e têm a mesma probabilidade de sucesso, então pela aditividade da distribuição Binomial tem-se

$$T \sim B\,(9, 0.10)$$

$$P(4 \leq T \leq 6) = 0.0083$$

Resolvido recorrendo à função de probabilidade

1. **cálculo direto**:

$$P(4 \leq T \leq 6) = f(4) + f(5) + f(6) =$$
$$= {}^9C_4 \times 0.10^4 \times (1-0.10)^{9-4} + {}^9C_5 \times 0.10^5 \times (1-0.10)^{9-5} + {}^9C_6 \times 0.10^6 \times (1-0.10)^{9-6} =$$
$$= 0.0083$$

3. **tabelas com recurso ao R**

$$P(4 \leq T \leq 6) = f(4) + f(5) + f(6) =$$
$$= dbinom(4, 9, 0.10) + dbinom(5, 9, 0.10) + dbinom(6, 9, 0.10) = 0.0083$$

Resolvido recorrendo à função de distribuição

1. **cálculo direto**: dá muito trabalho

2. **tabelas com recurso ao R**

$$P(4 \leq T \leq 6) = P(3 < T \leq 6) = F(6) - F(3) =$$
$$= pbinom(6, 9, 0.10) - pbinom(3, 9, 0.10) = 0.0083$$

Distribuições Teóricas Discretas:

- Distribuição Uniforme Discreta;
- Distribuição Binomial;
- **Distribuição de Poisson**.

Distribuição de Poisson

Processo de Poisson

Suponha que se procede à contagem do número de ocorrências de um acontecimento num determinado intervalo. Tem-se um processo de Poisson com parâmetro $\lambda > 0$ quando se verificam as seguintes condições:

- a probabilidade de uma ocorrência do acontecimento é a mesma para quaisquer dois intervalos de igual amplitude (apenas depende da amplitude do intervalo e não da posição em que se situa o intervalo);
- a ocorrência ou não ocorrência do acontecimento num determinado intervalo é independente da ocorrência ou não ocorrência do acontecimento num outro qualquer intervalo (não sobreposto);

Distribuição de Poisson

Processo de Poisson (continuação)

- a probabilidade de ocorrer um acontecimento em qualquer intervalo de amplitude Δt (arbitrariamente pequeno) é proporcional à dimensão do intervalo: $\lambda\Delta t$;

- a probabilidade de ocorrerem dois ou mais acontecimentos em qualquer intervalo de amplitude muito pequena é negligenciável, quando comparada com a probabilidade de se verificar apenas uma ocorrência.

Distribuição de Poisson

Definição

Diz-se que uma variável aleatória discreta X, que representa o número de ocorrências por unidade de medida, tem **distribuição de Poisson** com o parâmetro λ (fixo)

$$X \sim P(\lambda)$$

se a sua função de probabilidade é dada por

$$f(x) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots, \quad \lambda > 0$$

e a sua função de distribuição é dada por

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!}$$

onde λ representa a intensidade da distribuição.

Exemplo 16

Suponha que a variável aleatória discreta X, que representa o número de defeitos (por metro quadrado) na superfície de painéis de plástico usados no interior de uma máquina, segue uma distribuição de Poisson de parâmetro igual a 0.2.

Exemplo 16

Suponha que a variável aleatória discreta X, que representa o número de defeitos (por metro quadrado) na superfície de painéis de plástico usados no interior de uma máquina, segue uma distribuição de Poisson de parâmetro igual a 0.2.

Considere a variável a aleatória discreta:

$$X = \text{ número de defeitos, por metro quadrado, na superfície de painéis de plástico}$$

ou seja,

$$X = \text{ número de ocorrências por unidade de medida}$$

Portanto a variável aleatória X tem $D_X = \{0, 1, 2, \ldots\}$ e tem distribuição Poisson com o parâmetro $\lambda = 0.2$, ou seja

$$X \sim P\,(0.2)$$

Observação: Na distribuição Binomial também é efetuada uma contagem de ocorrências (a que se chama de sucessos) mas o domínio é finito, enquanto na distribuição Poisson é infinito (embora numerável).

Exemplo 16

$X =$ número de defeitos, por metro quadrado, na superfície de painéis de plástico

$$X \sim P\,(0.2)$$

função de probabilidade $f(x) = P(X = x)$:

x	0	1	2	3	$\cdots$
$f\,(x)$	0.8187	0.1637	0.0164	0.0011	$\cdots$

função de distribuição $F(x) = P(X \leq x)$:

$$F\,(x) = \begin{cases} 0 & , \; x < 0 \\ 0.8187 & , \; 0 \leq x < 1 \\ 0.9825 & , \; 1 \leq x < 2 \\ 0.9989 & , \; 2 \leq x < 3 \\ 0.9999 & , \; 3 \leq x < 4 \\ \cdots & , \; \cdots \end{cases}$$

Para calcular os valores da função de probabilidade e da função de distribuição vamos ver duas possibilidades:

1. **cálculos diretos**

Se $X \sim P(\lambda)$, a função de probabilidade é

$$f(x) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots, \quad \lambda > 0$$

No exemplo 16 tem-se $X \sim P(0.2)$, então

$$f(0) = P(X = 0) = \frac{e^{-0.2} \times 0.2^0}{0!} = 0.8187$$

$$f(1) = P(X = 1) = \frac{e^{-0.2} \times 0.2^1}{1!} = 0.1637$$

$$\ldots$$

Agora basta construir a função distribuição com base na função de probabilidade.

2. **tabelas com recurso ao R**:
    - função de probabilidade: **dpois**$(x, lambda = \lambda)$
    - função de distribuição: **ppois**$(x, lambda = \lambda)$
    - inversa da função de distribuição: **qpois**$(p, lambda = \lambda)$

No exemplo 16 tem-se $X \sim P\,(0.2)$, então

$$f(0) = P(X = 0) = dpois(0, 0.2) = 0.8187$$
$$f(1) = P(X = 1) = dpois(1, 0.2) = 0.1637$$
$$f(2) = P(X = 2) = dpois(2, 0.2) = 0.0164$$
$$\ldots$$

$$F(0) = P(X \leq 0) = ppois(0, 0.2) = 0.8187$$
$$F(1) = P(X \leq 1) = ppois(1, 0.2) = 0.9825$$
$$F(2) = P(X \leq 2) = ppois(2, 0.2) = 0.9989$$
$$\ldots$$

Distribuição de Poisson

Teorema

Se a variável aleatória X tem distribuição de Poisson com parâmetro λ,

$$X \sim P(\lambda)$$

então

$$E[X] = \lambda \qquad \text{e} \qquad V[X] = \lambda.$$

Distribuição de Poisson

Propriedade: Aditividade da Poisson

Sejam $X_1, X_2, \ldots, X_k$ variáveis aleatórias independentes com distribuição de Poisson, isto é

$$X_i \sim P(\lambda_i) \quad i = 1, \ldots, k,$$

então

$$Y = \sum_{i=1}^{k} X_i$$

também é uma variável aleatória com distribuição de Poisson, isto é

$$Y = \sum_{i=1}^{k} X_i \sim P\left(\sum_{i=1}^{k} \lambda_i\right).$$

Exemplo 17

A Central que gere os pedidos de ambulância de uma determinada região reencaminha esses pedidos para dois postos de socorro: A e B. Os pedidos serão reencaminhados para o posto de socorro A se os pedidos se dirigirem à zona norte da região, caso contrário serão reencaminhados para o posto de socorro B. Sabe-se que o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro A segue uma distribuição de Poisson de parâmetro 2 e o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro B segue uma distribuição de Poisson com média 3.

1. Em relação ao posto de socorro A, calcule a probabilidade de que num dia, haja 3 pedidos de ambulância.

Exemplo 17

A Central que gere os pedidos de ambulância de uma determinada região reencaminha esses pedidos para dois postos de socorro: A e B. Os pedidos serão reencaminhados para o posto de socorro A se os pedidos se dirigirem à zona norte da região, caso contrário serão reencaminhados para o posto de socorro B. Sabe-se que o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro A segue uma distribuição de Poisson de parâmetro 2 e o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro B segue uma distribuição de Poisson com média 3.

1. Em relação ao posto de socorro A, calcule a probabilidade de que num dia, haja 3 pedidos de ambulância.

$X =$ número de pedidos de ambulância, por dia, reencaminhados para o posto de socorro A

A variável aleatória discreta X tem $D_X = \{0, 1, 2, 3, \ldots\}$ e tem distribuição Poisson com o parâmetro $\lambda = 2$, ou seja

$$X \sim P(2)$$

$$P(X = 3) = f(3) = 0.1804$$

1. **cálculo direto**

$$f(3) = \frac{e^{-2} \times 2^3}{3!} = 0.1804$$

2. **tabelas com recurso ao R**

$$f(3) = dpois(3, 2) = 0.1804$$

Exemplo 17

A Central que gere os pedidos de ambulância de uma determinada região reencaminha esses pedidos para dois postos de socorro: A e B. Os pedidos serão reencaminhados para o posto de socorro A se os pedidos se dirigirem à zona norte da região, caso contrário serão reencaminhados para o posto de socorro B. Sabe-se que o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro A segue uma distribuição de Poisson de parâmetro 2 e o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro B segue uma distribuição de Poisson com média 3.

2. Em relação ao posto de socorro A, calcule a probabilidade de que num dia, haja pelo menos 4 pedidos de ambulância.

Exemplo 17

A Central que gere os pedidos de ambulância de uma determinada região reencaminha esses pedidos para dois postos de socorro: A e B. Os pedidos serão reencaminhados para o posto de socorro A se os pedidos se dirigirem à zona norte da região, caso contrário serão reencaminhados para o posto de socorro B. Sabe-se que o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro A segue uma distribuição de Poisson de parâmetro 2 e o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro B segue uma distribuição de Poisson com média 3.

2. Em relação ao posto de socorro A, calcule a probabilidade de que num dia, haja pelo menos 4 pedidos de ambulância.

$X =$ número de pedidos de ambulância, por dia, reencaminhados para o posto de socorro A

$$X \sim P(2)$$

$$P(X \geq 4) = 1 - P(X < 4) = 1 - P(X \leq 3) = 1 - F(3) = 1 - 0.8571 = 0.1429$$

1. **cálculo direto**

$$F(3) = f(0) + f(1) + f(2) + f(3) = \frac{e^{-2} \times 2^0}{0!} + \frac{e^{-2} \times 2^1}{1!} + \frac{e^{-2} \times 2^2}{2!} + \frac{e^{-2} \times 2^3}{3!} =$$
$$= 0.8571$$

2. **tabelas com recurso ao R**

$$F(3) = ppois(3, 2) = 0.8571$$

Exemplo 17

A Central que gere os pedidos de ambulância de uma determinada região reencaminha esses pedidos para dois postos de socorro: A e B. Os pedidos serão reencaminhados para o posto de socorro A se os pedidos se dirigirem à zona norte da região, caso contrário serão reencaminhados para o posto de socorro B. Sabe-se que o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro A segue uma distribuição de Poisson de parâmetro 2 e o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro B segue uma distribuição de Poisson com média 3.

3. Em relação ao posto de socorro A, calcule a probabilidade de que num dia, haja 3 pedidos de ambulância e no dia seguinte também se verifiquem 3 pedidos de ambulância.

Exemplo 17

A Central que gere os pedidos de ambulância de uma determinada região reencaminha esses pedidos para dois postos de socorro: A e B. Os pedidos serão reencaminhados para o posto de socorro A se os pedidos se dirigirem à zona norte da região, caso contrário serão reencaminhados para o posto de socorro B. Sabe-se que o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro A segue uma distribuição de Poisson de parâmetro 2 e o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro B segue uma distribuição de Poisson com média 3.

3. Em relação ao posto de socorro A, calcule a probabilidade de que num dia, haja 3 pedidos de ambulância e no dia seguinte também se verifiquem 3 pedidos de ambulância.

Sabemos

X = número de pedidos de ambulância, **num dia**, reencaminhados para o posto de socorro A, $\qquad X \sim P(2)$

Agora pretende-se:

Y = número de pedidos de ambulância, **no dia seguinte**, reencaminhados para o posto de socorro A

X = número de pedidos de ambulância, **num dia**, reencaminhados para o posto de socorro A, $\quad X \sim P(2)$ (*)

Y = número de pedidos de ambulância, **no dia seguinte**, reencaminhados para o posto de socorro A, $\quad Y \sim P(2)$ (*)

$$P(X=3 \wedge Y=3) \underset{(*)}{=} P(X=3) \times P(Y=3) = f(3) \times f(3) = 0.1804^2 = 0.0325$$

Recorrendo às condições de um Processo de Poisson, sabe-se:

(*) A probabilidade de uma ocorrência do acontecimento é a mesma para quaisquer dois intervalos de igual amplitude (apenas depende da amplitude do intervalo e não da posição em que se situa o intervalo).

(*) A ocorrência ou não ocorrência do acontecimento num determinado intervalo é **independente** da ocorrência ou não ocorrência do acontecimento num outro qualquer intervalo (não sobreposto).

Exemplo 17

A Central que gere os pedidos de ambulância de uma determinada região reencaminha esses pedidos para dois postos de socorro: A e B. Os pedidos serão reencaminhados para o posto de socorro A se os pedidos se dirigirem à zona norte da região, caso contrário serão reencaminhados para o posto de socorro B. Sabe-se que o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro A segue uma distribuição de Poisson de parâmetro 2 e o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro B segue uma distribuição de Poisson com média 3.

4. Em relação ao posto de socorro A, calcule a probabilidade de que em 2 dias, sejam pedidas 6 ambulâncias.

Exemplo 17

A Central que gere os pedidos de ambulância de uma determinada região reencaminha esses pedidos para dois postos de socorro: A e B. Os pedidos serão reencaminhados para o posto de socorro A se os pedidos se dirigirem à zona norte da região, caso contrário serão reencaminhados para o posto de socorro B. Sabe-se que o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro A segue uma distribuição de Poisson de parâmetro 2 e o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro B segue uma distribuição de Poisson com média 3.

4. Em relação ao posto de socorro A, calcule a probabilidade de que em 2 dias, sejam pedidas 6 ambulâncias.

Sabemos

$X =$ número de pedidos de ambulância, **num dia**, reencaminhados para o posto de socorro A, $\qquad X \sim P(2)$

Agora pretende-se:

$W =$ número de pedidos de ambulância, **em 2 dias**, reencaminhados para o posto de socorro A

X = número de pedidos de ambulância, **num dia**, reencaminhados para o posto de socorro A, $\quad X \sim P(2)$

W = número de pedidos de ambulância, **em 2 dias**, reencaminhados para o posto de socorro A

$$W \sim P(4) \quad (*)$$

$$P(W=6) = f_W(6) = 0.1042$$

Recorrendo às condições de um Processo de Poisson, sabe-se:

(*) A probabilidade de ocorrer um acontecimento em qualquer intervalo de amplitude Δt (arbitrariamente pequeno) é proporcional à dimensão do intervalo: $\lambda \Delta t$.

Esta condição indica que basta recorrer à regra de três simples para atualizar o parâmetro:

unidade de medida		parâmetro da Poisson
1 dia	$\mapsto$	$\lambda = 2$
2 dias	$\mapsto$	?

logo $\lambda_W = 4$

Outra possibilidade de resolução é:

W = número de pedidos de ambulância, **em 2 dias**, reencaminhados para o posto de socorro A

Já tínhamos definido:

X = número de pedidos de ambulância, **num dia**, reencaminhados para o posto de socorro A, $\quad X \sim P(2)$

Y = número de pedidos de ambulância, **no dia seguinte**, reencaminhados para o posto de socorro A, $\quad Y \sim P(2)$

Portanto

$$W = X + Y$$

como X e Y são variáveis aleatórias independentes, então pela **aditividade da distribuição de Poisson** tem-se

$$W \sim P(4)$$

pois $\lambda_W = \lambda_X + \lambda_Y = 2 + 2 = 4$

$$P(W = 6) = f_W(6) = 0.1042$$

1. **cálculo direto**

$$f_W(6) = \frac{e^{-4} \times 4^6}{6!} = 0.1042$$

2. **tabelas com recurso ao R**

$$f_W(6) = dpois(6, 4) = 0.1042$$

Exemplo 17

A Central que gere os pedidos de ambulância de uma determinada região reencaminha esses pedidos para dois postos de socorro: A e B. Os pedidos serão reencaminhados para o posto de socorro A se os pedidos se dirigirem à zona norte da região, caso contrário serão reencaminhados para o posto de socorro B. Sabe-se que o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro A segue uma distribuição de Poisson de parâmetro 2 e o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro B segue uma distribuição de Poisson com média 3.

5. Em relação ao posto de socorro A, calcule a probabilidade de que em 12 horas (metade de um dia), sejam pedidas mais de 2 ambulâncias.

Exemplo 17

A Central que gere os pedidos de ambulância de uma determinada região reencaminha esses pedidos para dois postos de socorro: A e B. Os pedidos serão reencaminhados para o posto de socorro A se os pedidos se dirigirem à zona norte da região, caso contrário serão reencaminhados para o posto de socorro B. Sabe-se que o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro A segue uma distribuição de Poisson de parâmetro 2 e o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro B segue uma distribuição de Poisson com média 3.

5. Em relação ao posto de socorro A, calcule a probabilidade de que em 12 horas (metade de um dia), sejam pedidas mais de 2 ambulâncias.

Sabemos

X = número de pedidos de ambulância, **num dia**, reencaminhados para o posto de socorro A, $\qquad X \sim P(2)$

Agora pretende-se:

V = número de pedidos de ambulância, **em 12 horas**, reencaminhados para o posto de socorro A

X = número de pedidos de ambulância, **num dia**, reencaminhados para o posto de socorro A, $\quad X \sim P(2)$

V = número de pedidos de ambulância, **em 12 horas**, reencaminhados para o posto de socorro A

$$V \sim P(1) \quad (*)$$

$$P(V > 2) = 1 - P(V \leq 2) = 1 - F_V(2) = 1 - 0.9197 = 0.0803$$

Recorrendo às condições de um Processo de Poisson, sabe-se:

(*) A probabilidade de ocorrer um acontecimento em qualquer intervalo de amplitude Δt (arbitrariamente pequeno) é proporcional à dimensão do intervalo: $\lambda \Delta t$.

Esta condição indica que basta recorrer à regra de três simples para atualizar o parâmetro:

unidade de medida		parâmetro da Poisson
1 dia = 24 horas	$\mapsto$	$\lambda = 2$
12 horas	$\mapsto$	?

logo $\lambda_V = 1$

1. **cálculo direto**

$$F_V(2) = f_V(0) + f_V(1) + f_V(2) = \frac{e^{-1} \times 1^0}{0!} + \frac{e^{-1} \times 1^1}{1!} + \frac{e^{-1} \times 1^2}{2!} = 0.9197$$

2. **tabelas com recurso ao R**

$$F_V(2) = ppois(2, 1) = 0.9197$$

Exemplo 17

A Central que gere os pedidos de ambulância de uma determinada região reencaminha esses pedidos para dois postos de socorro: A e B. Os pedidos serão reencaminhados para o posto de socorro A se os pedidos se dirigirem à zona norte da região, caso contrário serão reencaminhados para o posto de socorro B. Sabe-se que o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro A segue uma distribuição de Poisson de parâmetro 2 e o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro B segue uma distribuição de Poisson com média 3.

6. Em 75% dos dias qual o número máximo de pedidos de ambulância no posto de socorro A?

Exemplo 17

A Central que gere os pedidos de ambulância de uma determinada região reencaminha esses pedidos para dois postos de socorro: A e B. Os pedidos serão reencaminhados para o posto de socorro A se os pedidos se dirigirem à zona norte da região, caso contrário serão reencaminhados para o posto de socorro B. Sabe-se que o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro A segue uma distribuição de Poisson de parâmetro 2 e o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro B segue uma distribuição de Poisson com média 3.

6. Em 75% dos dias qual o número máximo de pedidos de ambulância no posto de socorro A?

$X =$ número de pedidos de ambulância, por dia, reencaminhados para o posto de socorro A

$$X \sim P\left(2\right)$$

Pretende-se determinar m tal que

$$P(X \leq m) = 0.75 \Leftrightarrow F(m) = 0.75 \Leftrightarrow m = F^{-1}(0.75) = 3 \text{ ambulâncias}$$

pois $F^{-1}(0.75) = qpois(0.75, 2) = 3$

Exemplo 17

A Central que gere os pedidos de ambulância de uma determinada região reencaminha esses pedidos para dois postos de socorro: A e B. Os pedidos serão reencaminhados para o posto de socorro A se os pedidos se dirigirem à zona norte da região, caso contrário serão reencaminhados para o posto de socorro B. Sabe-se que o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro A segue uma distribuição de Poisson de parâmetro 2 e o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro B segue uma distribuição de Poisson com média 3.

7. Em relação à Central, calcule a probabilidade de que, num dia, haja no mínimo 7 pedidos de ambulância na Central.

Exemplo 17

A Central que gere os pedidos de ambulância de uma determinada região reencaminha esses pedidos para dois postos de socorro: A e B. Os pedidos serão reencaminhados para o posto de socorro A se os pedidos se dirigirem à zona norte da região, caso contrário serão reencaminhados para o posto de socorro B. Sabe-se que o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro A segue uma distribuição de Poisson de parâmetro 2 e o número de pedidos de ambulância reencaminhados, por dia, para o posto de socorro B segue uma distribuição de Poisson com média 3.

7. Em relação à Central, calcule a probabilidade de que, num dia, haja no mínimo 7 pedidos de ambulância na Central.

$X =$ número de pedidos de ambulância, num dia, reencaminhados para o posto de socorro A, $\quad X \sim P(2)$

$S =$ número de pedidos de ambulância, num dia, reencaminhados para o posto de socorro B, $\quad S \sim P(3) \quad$ pois $E[S] = \lambda = 3$

$T = X + S =$ número de pedidos de ambulância, num dia, na Central

$X =$ número de pedidos de ambulância, num dia, reencaminhados para o posto de socorro A

$$X \sim P\left(2\right)$$

$S =$ número de pedidos de ambulância, num dia, reencaminhados para o posto de socorro B

$$S \sim P\left(3\right)$$

$T = X + S =$ número de pedidos de ambulância, num dia, na Central

como X e S podem ser consideradas variáveis aleatórias independentes, então pela **aditividade da distribuição de Poisson** tem-se

$$T \sim P\left(5\right)$$

pois $\lambda_T = \lambda_X + \lambda_S = 2 + 3 = 5$

$$P(T \geq 7) = 1 - P(T < 7) = 1 - P(T \leq 6) = 1 - F_T(6) = 1 - 0.7622 = 0.2378$$

1. **cálculo direto**: dá muito trabalho

2. **tabelas com recurso ao R**

$$F_T(6) = ppois(6, 5) = 0.7622$$

Distribuição de Poisson

Teorema

A distribuição Binomial, $B(n,p)$, converge para a distribuição de Poisson, $P(\lambda)$, quando $n \to +\infty$ (o número de provas é muito grande), $p \to 0$ (a probabilidade de sucesso é muito pequena) e o produto (np) mantém-se aproximadamente constante, $np = \lambda > 0$ (o número médio de sucessos mantém-se aproximadamente constante ao longo das provas).

$$X \sim B(n,p) \underset{n \to +\infty \text{ e } p \to 0}{\longrightarrow} X \,\dot{\sim}\, P\left(\underbrace{np}_{=\lambda}\right)$$

Observação

Na prática a distribuição de Poisson é uma boa aproximação da distribuição Binomial se $n \geq 30$ e $np \leq 5$ (ou $nq \leq 5$).

Variáveis Aleatórias Contínuas

Uma variável aleatória diz-se **Contínua** se pode assumir um número infinito não numerável de valores.

Uma variável aleatória contínua fica perfeitamente identificada através da:

- **função densidade de probabilidade**

  ou

- **função de distribuição**

e através dos seus parâmetros (apenas vamos considerar 3):

- **valor esperado** ou média ou esperança matemática
- **variância**
- **desvio padrão**

Variáveis Aleatórias Contínuas

Função Densidade de Probabilidade (f.d.p.)

Se X é uma variável aleatória contínua, então existe uma função $f\left(x\right)$ tal que

$$F\left(x\right)=\int_{-\infty}^{x} f\left(t\right) dt.$$

À função $f\left(x\right)$ dá-se o nome de função densidade de probabilidade e pode ser representada por

$$f\left(x\right)=\begin{cases} F'\left(x\right) & \text{, caso exista} \\ 0 & \text{, outros casos} \end{cases}$$

e satisfaz as seguintes propriedades:

1. $f\left(x\right)\geq 0, \quad \forall x;$
2. $\int_{-\infty}^{+\infty} f\left(x\right) dx=1.$

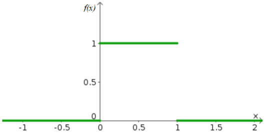
Função densidade de probabilidade = f(x)

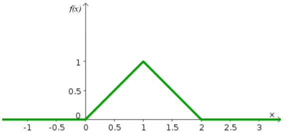
- Nas variáveis aleatórias **discretas** tem-se a **função de probabilidade** que permite calcular as probabilidades pontuais: $f(x) = P(X = x)$.

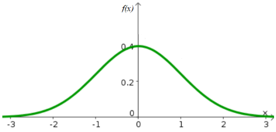
  Aqui o interesse não são as probabilidades pontuais mas a probabilidade de estar dentro de um intervalo: $P(a \leq X \leq b)$.

- Nas variáveis aleatórias **contínuas** tem-se a **função densidade de probabilidade**, esta função não tem como objetivo calcular probabilidades pontuais (pois aqui não têm interesse) mas descrever a probabilidade relativa de uma variável aleatória, ou seja, o seu comportamento: onde cresce, onde diminui ou onde se mantém constante.

  Alguns gráficos de funções densidade de probabilidade:





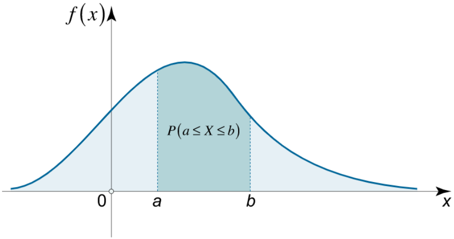


Função densidade de probabilidade = f(x)

- Nas variáveis aleatórias contínuas o interesse está na probabilidade de uma variável cair dentro de um intervalo, então essa probabilidade é dada pelo integral da função densidade nesse intervalo.

- Ou seja, a probabilidade no caso das variáveis contínuas é dada pela área compreendida entre a função densidade de probabilidade, o eixo dos xx e os limites do intervalo pretendido:

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



Função densidade de probabilidade = f(x)

No entanto é necessário ter em atenção que uma função $f(x)$ só é uma **função densidade de probabilidade** se verificar as seguintes propriedades:

1. $f(x) \geq 0, \quad \forall x;$
2. $\int_{-\infty}^{+\infty} f(x)\, dx = 1.$

Ou seja, a função densidade de probabilidade não pode ser negativa e a área total entre a função densidade de probabilidade e o eixo dos xx é 1.

Função densidade de probabilidade = f(x)

Exemplo 18

De acordo com determinadas especificações técnicas de uma marca de aparelhos de radiologia, sabe-se que a sua duração (em anos) é uma variável aleatória X com a seguinte função densidade de probabilidade:

$$f(x) = \begin{cases} \frac{x}{4} & , \ 0 < x \leq 2 \\ 1 - \frac{x}{4} & , \ 2 < x \leq 4 \\ 0 & , \ \text{caso contrário} \end{cases}$$

1. Mostre que f é de facto uma função densidade de probabilidade.

$f(x)$ é função densidade de probabilidade sse

1. $f(x) \geq 0, \quad \forall x;$
2. $\int_{-\infty}^{+\infty} f(x)\, dx = 1.$

propriedade 1:

se $0 < x \leq 2$, $\quad f(x) \geq 0 \Leftrightarrow \dfrac{x}{4} \geq 0 \Leftrightarrow x \geq 0$ o que é verdade pois $0 < x \leq 2$

se $2 < x \leq 4$, $\quad f(x) \geq 0 \Leftrightarrow 1 - \dfrac{x}{4} \geq 0 \Leftrightarrow x \leq 4$ o que é verdade pois $2 < x \leq 4$

se $x \leq 0$ ou $x > 4$, $\quad f(x) \geq 0 \Leftrightarrow 0 \geq 0$ sempre verdade, qualquer que seja x

propriedade 2:

$$\int_{-\infty}^{+\infty} f(x)\, dx = \int_{-\infty}^{0} 0 dx + \int_{0}^{2} \frac{x}{4} dx + \int_{2}^{4} \left(1 - \frac{x}{4}\right) dx + \int_{4}^{+\infty} 0 dx =$$

$$= 0 + \left[\frac{x^2}{8}\right]_0^2 + \left[x - \frac{x^2}{8}\right]_2^4 + 0 = \left(\frac{2^2}{8} - 0\right) + \left(4 - \frac{4^2}{8} - \left(2 - \frac{2^2}{8}\right)\right) = 1$$

Como verifica as duas propriedades, então é função densidade de probabilidade.

Função densidade de probabilidade = f(x)

Exemplo 18

De acordo com determinadas especificações técnicas de uma marca de aparelhos de medição da tensão arterial, sabe-se que a sua duração (em anos) é uma variável aleatória X com a seguinte função densidade de probabilidade:

$$f(x) = \begin{cases} \frac{x}{4} & , \ 0 < x \leq 2 \\ 1 - \frac{x}{4} & , \ 2 < x \leq 4 \\ 0 & , \ \text{caso contrário} \end{cases}$$

2. Calcule $P(1 \leq X \leq 3)$.

Função densidade de probabilidade = f(x)

Exemplo 18

De acordo com determinadas especificações técnicas de uma marca de aparelhos de medição da tensão arterial, sabe-se que a sua duração (em anos) é uma variável aleatória X com a seguinte função densidade de probabilidade:

$$f(x) = \begin{cases} \frac{x}{4} & , \ 0 < x \leq 2 \\ 1 - \frac{x}{4} & , \ 2 < x \leq 4 \\ 0 & , \ \text{caso contrário} \end{cases}$$

2. Calcule $P(1 \leq X \leq 3)$.

$$P(1 \leq X \leq 3) = \int_1^3 f(x)\,dx = \int_1^2 \frac{x}{4}dx + \int_2^3 \left(1 - \frac{x}{4}\right)dx =$$

$$= \left[\frac{x^2}{8}\right]_1^2 + \left[x - \frac{x^2}{8}\right]_2^3 = \left(\frac{2^2}{8} - \frac{1^2}{8}\right) + \left(3 - \frac{3^2}{8} - \left(2 - \frac{2^2}{8}\right)\right) = \frac{3}{4}$$

Variáveis Aleatórias Contínuas

Função de Distribuição (f.d.)

Se X é uma variável aleatória contínua, então a função de distribuição é representada por $F(x)$ e é definida por

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt$$

e satisfaz as seguintes propriedades:

1. $0 \leq F(x) \leq 1$;
2. $F(x)$ é uma função não decrescente;
3. $F(x)$ é contínua em $\mathbb{R}$;
4. $\lim\limits_{x \to -\infty} F(x) = 0$ e $\lim\limits_{x \to +\infty} F(x) = 1$;
5. $P(X = a) = 0$;
6. $P(a < X \leq b) = P(a \leq X < b) = P(a \leq X \leq b) =$
   $= P(a < X < b) = F(b) - F(a)$.

Função de distribuição = F(x)

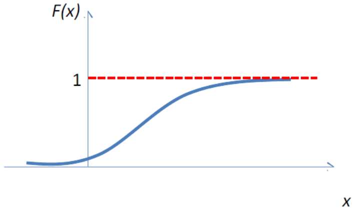
- Tal como nas variáveis aleatórias discretas, a função de distribuição de uma variável aleatória contínua permite calcular as probabilidades acumuladas:

$$F(x) = P(X \leq x)$$

- A diferença é que, em vez de somar probabilidades pontuais, vamos somar áreas:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$$

- Em relação às propriedades, agora $F(x)$ é contínua em $\mathbb{R}$, o seu gráfico tem o seguinte aspeto:



Função de distribuição = F(x)

Exemplo 18

De acordo com determinadas especificações técnicas de uma marca de aparelhos de radiologia, sabe-se que a sua duração (em anos) é uma variável aleatória X com a seguinte função densidade de probabilidade:

$$f(x) = \begin{cases} \frac{x}{4} & , \ 0 < x \leq 2 \\ 1 - \frac{x}{4} & , \ 2 < x \leq 4 \\ 0 & , \ \text{caso contrário} \end{cases}$$

3. Calcule a função de distribuição.

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt = \begin{cases} 0 & , \ x \leq 0 & \color{red}{(1)} \\ \frac{x^2}{8} & , \ 0 < x \leq 2 & \color{red}{(2)} \\ -\frac{x^2}{8} + x - 1 & , \ 2 < x \leq 4 & \color{red}{(3)} \\ 1 & , x > 4 & \color{red}{(4)} \end{cases}$$

(1) $F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{x} 0dt = 0$

(2) $F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{0} f(t)dt + \int_{0}^{x} \frac{t}{4}dt = 0 + \left[\frac{t^2}{8}\right]_0^x = \frac{x^2}{8}$

(3) $F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{2} f(t)dt + \int_{2}^{x} \left(1 - \frac{t}{4}\right)dt = \frac{2^2}{8} + \left[t - \frac{t^2}{8}\right]_2^x = -\frac{x^2}{8} + x - 1$

(4) $F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{4} f(t)dt + \int_{4}^{x} 0dt = -\frac{4^2}{8} + 4 - 1 + 0 = 1$

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt = \begin{cases} 0 & , \ x \leq 0 & (1) \\ \frac{x^2}{8} & , \ 0 < x \leq 2 & (2) \\ -\frac{x^2}{8} + x - 1 & , \ 2 < x \leq 4 & (3) \\ 1 & , x > 4 & (4) \end{cases}$$

(1) $F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{x} 0dt = 0$

$F(0)$

(2) $F(x) = \int_{-\infty}^{x} f(t)\,dt = \underbrace{\int_{-\infty}^{0} f(t)dt}_{F(0)} + \int_{0}^{x} \frac{t}{4}dt = 0 + \left[\frac{t^2}{8}\right]_0^x = \frac{x^2}{8}$

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt = \begin{cases} 0 & , x \leq 0 & (1) \\ \frac{x^2}{8} & , 0 < x \leq 2 & (2) \\ -\frac{x^2}{8} + x - 1 & , 2 < x \leq 4 & (3) \\ 1 & , x > 4 & (4) \end{cases}$$

$$(2)\quad F(x) = \int_{-\infty}^{x} f(t)\,dt = \underbrace{\int_{-\infty}^{0} f(t)dt}_{F(0)} + \int_{0}^{x} \tfrac{t}{4}dt = 0 + \left[\frac{t^2}{8}\right]_0^x = \frac{x^2}{8}$$

F(2)

$$(3)\quad F(x) = \int_{-\infty}^{x} f(t)\,dt = \underbrace{\int_{-\infty}^{2} f(t)dt}_{F(2)} + \int_{2}^{x} \left(1 - \tfrac{t}{4}\right)dt = \frac{2^2}{8} + \left[t - \frac{t^2}{8}\right]_2^x = -\frac{x^2}{8} + x - 1$$

F(4)

$$(4)\quad F(x) = \int_{-\infty}^{x} f(t)\,dt = \underbrace{\int_{-\infty}^{4} f(t)dt}_{F(4)} + \int_{4}^{x} 0dt = -\frac{4^2}{8} + 4 - 1 + 0 = 1$$

Função de distribuição = F(x)

- O cálculo da função distribuição pode dar "muito trabalho" mas compensa, pois o cálculo das probabilidades torna-se muito mais simples no caso das variáveis aleatórias contínuas.

- Como $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt$, então

$$\boxed{\int_a^b f(x)\, dx = F(b) - F(a)}$$

- Também permite verificar que as probabilidades pontuais não têm interesse no caso das variáveis aleatórias contínuas, pois

$$P(X = a) = \int_a^a f(x)\, dx = F(a) - F(a) = 0$$

Função de distribuição = F(x)

- Como no caso das variáveis aleatórias contínuas tem-se $P(X = a) = 0$, logo quando X é uma variável aleatória contínua tem-se:

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

e em todos os caso basta calcular:

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

Exemplo 18

De acordo com determinadas especificações técnicas de uma marca de aparelhos de radiologia, sabe-se que a sua duração (em anos) é uma variável aleatória X com a seguinte função densidade de probabilidade:

$$f(x) = \begin{cases} \frac{x}{4} & , \ 0 < x \leq 2 \\ 1 - \frac{x}{4} & , \ 2 < x \leq 4 \\ 0 & , \ \text{caso contrário} \end{cases}$$

4. Calcule

a) $P(X \leq 1)$
b) $P(X = 1)$
c) $P(X < 1)$
d) $P(X > 1)$
e) $P(1 \leq X < 3)$
f) $P(1 \leq X \leq 3)$
g) $P(1 < X \leq 3)$
h) $P(1 < X < 3)$

Como a função de distribuição é:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt = \begin{cases} 0 & , x \leq 0 \\ \frac{x^2}{8} & , 0 < x \leq 2 \\ -\frac{x^2}{8} + x - 1 & , 2 < x \leq 4 \\ 1 & , x > 4 \end{cases}$$

a) $P(X \leq 1) = F(1) = \frac{1^2}{8} = \frac{1}{8}$
b) $P(X = 1) = 0$
c) $P(X < 1) = F(1) = \frac{1^2}{8} = \frac{1}{8}$
d) $P(X > 1) = 1 - P(X \leq 1) = 1 - F(1) = 1 - \frac{1^2}{8} = \frac{7}{8}$
e) $P(1 \leq X < 3) = F(3) - F(1) = -\frac{3^2}{8} + 3 - 1 - \frac{1^2}{8} = \frac{3}{4}$
f) $P(1 \leq X \leq 3) = F(3) - F(1) = -\frac{3^2}{8} + 3 - 1 - \frac{1^2}{8} = \frac{3}{4}$
g) $P(1 < X \leq 3) = F(3) - F(1) = -\frac{3^2}{8} + 3 - 1 - \frac{1^2}{8} = \frac{3}{4}$
h) $P(1 < X < 3) = F(3) - F(1) = -\frac{3^2}{8} + 3 - 1 - \frac{1^2}{8} = \frac{3}{4}$

Exemplo 18

De acordo com determinadas especificações técnicas de uma marca de aparelhos de radiologia, sabe-se que a sua duração (em anos) é uma variável aleatória X com a seguinte função densidade de probabilidade:

$$f(x) = \begin{cases} \frac{x}{4} & , \ 0 < x \leq 2 \\ 1 - \frac{x}{4} & , \ 2 < x \leq 4 \\ 0 & , \ \text{caso contrário} \end{cases}$$

5. Sabendo que 90% dos aparelhos de radiologia duram no máximo k anos, determine k.

Pretende-se determinar k tal que

$$P(X \leq k) = 0.90$$

Ou seja

$$P(X \leq k) = 0.90 \Leftrightarrow F(k) = 0.90$$

Sabemos que k existe, pois F é uma função contínua em $\mathbb{R}$.

Como a função de distribuição é:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt = \begin{cases} 0 & , \; x \leq 0 \\ \frac{x^2}{8} & , \; 0 < x \leq 2 \\ -\frac{x^2}{8} + x - 1 & , \; 2 < x \leq 4 \\ 1 & , x > 4 \end{cases}$$

tem-se:

- Se $k \leq 0, \quad F(k) = 0 \neq 0.90$

- Se $0 < k \leq 2, \quad F(k) = \frac{k^2}{8}$
  então $F(k) = 0.90 \Leftrightarrow \frac{k^2}{8} = 0.90 \Leftrightarrow k = \sqrt{7.2} = 2.68 \vee k = -\sqrt{7.2} = -2.68$
  $k = 2.68$ impossível pois $0 < k \leq 2$
  $k = -2.68$ impossível pois $0 < k \leq 2$

- Se $2 < k \leq 4, \quad F(k) = -\frac{k^2}{8} + k - 1$
  então $F(k) = 0.90 \Leftrightarrow -\frac{k^2}{8} + k - 1 = 0.90 \Leftrightarrow -\frac{k^2}{8} + k - 1,9 = 0 \Leftrightarrow$
  $\Leftrightarrow k = 3.11 \vee k = 4.89$
  $k = 4.89$ impossível pois $2 < k \leq 4$
  $k = 3.11$ possível

- Se $k > 4, \quad F(k) = 1 \neq 0.90$

$k = 3.11$ anos.

Função de distribuição = F(x)

Exemplo 19

A quantidade de álcool (em dezenas de litros) que um produtor-engarrafador vende por dia é uma variável aleatória X com função de distribuição definida por:

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{x^2}{50} & , \quad 0 \leq x < 5 \\ \frac{2}{5}x - \frac{x^2}{50} - 1 & , \quad 5 \leq x < 10 \\ 1 & , \quad x \geq 10 \end{cases} .$$

1. Determine a probabilidade do produtor-engarrafador vender entre 40 a 80 litros de álcool, por dia.

Função de distribuição = F(x)

Exemplo 19

A quantidade de álcool (em dezenas de litros) que um produtor-engarrafador vende por dia é uma variável aleatória X com função de distribuição definida por:

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{x^2}{50} & , \quad 0 \leq x < 5 \\ \frac{2}{5}x - \frac{x^2}{50} - 1 & , \quad 5 \leq x < 10 \\ 1 & , \quad x \geq 10 \end{cases} .$$

1. Determine a probabilidade do produtor-engarrafador vender entre 40 a 80 litros de álcool, por dia.

$$P(4 \leq X \leq 8) = F(8) - F(4) = \left(\frac{2}{5} \times 8 - \frac{8^2}{50} - 1\right) - \frac{4^2}{50} = \frac{3}{5}$$

Exemplo 19

A quantidade de álcool (em dezenas de litros) que um produtor-engarrafador vende por dia é uma variável aleatória X com função de distribuição definida por:

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{x^2}{50} & , \quad 0 \leq x < 5 \\ \frac{2}{5}x - \frac{x^2}{50} - 1 & , \quad 5 \leq x < 10 \\ 1 & , \quad x \geq 10 \end{cases}.$$

2. Calcule a função densidade de probabilidade da variável aleatória X.

Exemplo 19

A quantidade de álcool (em dezenas de litros) que um produtor-engarrafador vende por dia é uma variável aleatória X com função de distribuição definida por:

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{x^2}{50} & , \quad 0 \leq x < 5 \\ \frac{2}{5}x - \frac{x^2}{50} - 1 & , \quad 5 \leq x < 10 \\ 1 & , \quad x \geq 10 \end{cases} .$$

2. Calcule a função densidade de probabilidade da variável aleatória X.

A função densidade de probabilidade obtém-se a partir da derivada da função de distribuição de X,

$$f(x) = F'(x)$$

isto é, $f(x)$ é dada por

$$f(x) = F'(x) = \begin{cases} (0)' & , \quad x < 0 \\ \left(\frac{x^2}{50}\right)' & , \quad 0 \leq x < 5 \\ \left(\frac{2}{5}x - \frac{x^2}{50} - 1\right)' & , \quad 5 \leq x < 10 \\ (1)' & , \quad x \geq 10 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow f(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{x}{25} & , \quad 0 \leq x < 5 \\ \frac{2}{5} - \frac{x}{25} & , \quad 5 \leq x < 10 \\ 0 & , \quad x \geq 10 \end{cases} \Leftrightarrow f(x) = \begin{cases} \frac{x}{25} & , \quad 0 \leq x < 5 \\ \frac{2}{5} - \frac{x}{25} & , \quad 5 \leq x < 10 \\ 0 & , \quad \text{caso contrário} \end{cases}$$

Variáveis Aleatórias Contínuas

Valor Esperado ou Média ou Esperança Matemática

O valor esperado ou média ou esperança matemática de uma variável aleatória contínua X representa-se por

$$\mu = \mu_X = E[X]$$

e calcula-se

$$\mu = E[X] = \int_{-\infty}^{+\infty} x f(x)\, dx.$$

Observação

O valor esperado é um **parâmetro de localização**, que pretende localizar o centro da distribuição de probabilidade, ou seja, pretende identificar o "centro de gravidade" da variável aleatória.

Valor Esperado = E[ ]

- O cálculo do valor esperado nas variáveis aleatórias contínuas faz-se de forma idêntica à realizada nas variáveis aleatórias discretas, mas em vez do "símbolo de somatório aparece o integral".

- Ou seja, nas variáveis aleatórias discretas o valor esperado é uma média ponderada cujos pesos são as probabilidades pontuais e nas variáveis aleatórias contínuas o valor esperado refere-se à área média.

- Todas as observações e propriedades referidas para o valor esperado no caso das variáveis aleatórias discretas continuam a ser válidas quando a variável é contínua:

Variáveis Aleatórias Contínuas

Observação

Seja $g(X)$ uma função da variável aleatória X. Se X uma variável aleatória contínua, então

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x) f(x) dx.$$

Propriedades

Sejam X uma variável aleatória e a e b constantes reais.

1. Se $X = a$, então $E[X] = E[a] = a$;
2. $E[aX + b] = aE[X] + b$;
3. Sejam $g(X)$ e $h(X)$ funções de X

$$E[g(X) + h(X)] = E[g(X)] + E[h(X)].$$

Valor Esperado = E[ ]

Exemplo 18

De acordo com determinadas especificações técnicas de uma marca de aparelhos de radiologia, sabe-se que a sua duração (em anos) é uma variável aleatória X com a seguinte função densidade de probabilidade:

$$f(x) = \begin{cases} \frac{x}{4} & , \; 0 < x \leq 2 \\ 1 - \frac{x}{4} & , \; 2 < x \leq 4 \\ 0 & , \; \text{caso contrário} \end{cases}$$

6. Qual a duração média dos aparelhos de radiologia?

Pretende-se a duração média dos aparelhos de radiologia, $\mu = E[X]$:

$$E[X] = \int_{-\infty}^{+\infty} x f(x)\, dx =$$

$$= \int_{-\infty}^{0} x \times 0 dx + \int_{0}^{2} x \times \frac{x}{4} dx + \int_{2}^{4} x \times \left(1 - \frac{x}{4}\right) dx + \int_{4}^{+\infty} x \times 0 dx =$$

$$= \int_{-\infty}^{0} 0 dx + \int_{0}^{2} \frac{x^2}{4} dx + \int_{2}^{4} \left(x - \frac{x^2}{4}\right) dx + \int_{4}^{+\infty} 0 dx =$$

$$= 0 + \left[\frac{x^3}{12}\right]_0^2 + \left[\frac{x^2}{2} - \frac{x^3}{12}\right]_2^4 + 0 =$$

$$= \left(\frac{2^3}{12} - 0\right) + \left[\frac{4^2}{2} - \frac{4^3}{12} - \left(\frac{2^2}{2} - \frac{2^3}{12}\right)\right] = 2 \text{ anos}$$

Valor Esperado = E[ ]

Exemplo 18

De acordo com determinadas especificações técnicas de uma marca de aparelhos de radiologia, sabe-se que a sua duração (em anos) é uma variável aleatória X com a seguinte função densidade de probabilidade:

$$f(x) = \begin{cases} \frac{x}{4} & , \ 0 < x \leq 2 \\ 1 - \frac{x}{4} & , \ 2 < x \leq 4 \\ 0 & , \ \text{caso contrário} \end{cases}$$

7. Calcule $E[3X - 2]$.

Recorrendo à propriedade $E\left[aX+b\right]=aE\left[X\right]+b$ tem-se

$$E\left[3X-2\right]=3E[X]-2=3\times 2-2=4$$

Claro que também era possível calcular $E\left[3X-2\right]$ sem usar as propriedades:

$$\begin{aligned}
E\left[3X-2\right] &= \int_{-\infty}^{+\infty}(3x-2)\,f(x)\,dx= \\
&= \int_{-\infty}^{0}(3x-2)\times 0dx+\int_{0}^{2}(3x-2)\times\frac{x}{4}dx+ \\
&\qquad +\int_{2}^{4}(3x-2)\times\left(1-\frac{x}{4}\right)dx+\int_{4}^{+\infty}(3x-2)\times 0dx= \\
&= \int_{-\infty}^{0}0dx+\int_{0}^{2}\left(\frac{3x^2}{4}-\frac{2x}{4}\right)dx+ \\
&\qquad +\int_{2}^{4}\left(-\frac{3x^2}{4}+-\frac{7x}{2}-2\right)dx+\int_{4}^{+\infty}0dx= \\
&= \dots=4
\end{aligned}$$

Variáveis Aleatórias Contínuas

Variância

A variância de uma variável aleatória contínua X representa-se por

$$\sigma^2 = \sigma_X^2 = Var\left[X\right] = V\left[X\right] = E\left[\left(X-\mu\right)^2\right].$$

e calcula-se

$$\sigma^2 = V\left[X\right] = \int_{-\infty}^{+\infty} \left(x-\mu\right)^2 f\left(x\right) dx.$$

Observação

A variância é um **parâmetro de dispersão**. Mede a dispersão (ao quadrado) da variável aleatória em torno do seu valor esperado.

Variância = V[ ]

- Tal como foi referido para o valor médio, o cálculo da variância nas variáveis aleatórias contínuas faz-se de forma idêntica à realizada nas variáveis aleatórias discretas, mas em vez do "símbolo de somatório aparece o integral".

- Todas as observações e propriedades referidas para a variância na caso das variáveis aleatórias discretas continuam a ser válidas quando a variável é contínua:

Variáveis Aleatórias Contínuas

Propriedades:

Sejam X uma variável aleatória e a e b constantes reais.

1. $V[X] = E[X^2] - E^2[X]$;
2. $V[X] \geq 0$;
3. Se $X = a$, então $V[X] = V[a] = 0$;
4. $V[aX + b] = a^2 V[X]$.

Desvio Padrão

- Como já vimos, o desvio padrão é a raíz quadrada da variância:

Variáveis Aleatórias Contínuas

Desvio Padrão

O desvio padrão de uma variável aleatória X representa-se por

$$\sigma = \sigma_X.$$

e calcula-se

$$\sigma = \sigma_X = \sqrt{V[X]}.$$

Observação

O desvio padrão é um **parâmetro de dispersão**, é a raiz quadrada da variância. Mede a dispersão da variável aleatória em torno do seu valor esperado na mesma unidade de medida em que a variável aleatória vem expressa.

Variância e Desvio Padrão

Exemplo 18

De acordo com determinadas especificações técnicas de uma marca de aparelhos de radiologia, sabe-se que a sua duração (em anos) é uma variável aleatória X com a seguinte função densidade de probabilidade:

$$f(x) = \begin{cases} \frac{x}{4} & , \ 0 < x \leq 2 \\ 1 - \frac{x}{4} & , \ 2 < x \leq 4 \\ 0 & , \ \text{caso contrário} \end{cases}$$

8. Calcule a variância da variável aleatória X.

Pretende-se a variância de X, $\sigma^2 = V[X]$:

- Usando a **definição** tem-se:

$$V[X] = E\left[(X-\mu)^2\right] = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)\,dx = \qquad (\text{tem-se } \mu = E[X] = 2)$$

$$= E\left[(X-2)^2\right] = \int_{-\infty}^{+\infty} (x-2)^2 f(x)\,dx =$$

$$= \int_{-\infty}^{0} (x-2)^2 \times 0dx + \int_{0}^{2} (x-2)^2 \times \frac{x}{4}dx +$$

$$+ \int_{2}^{4} (x-2)^2 \times \left(1-\frac{x}{4}\right) dx + \int_{4}^{+\infty} (x-2)^2 \times 0dx =$$

$$= \int_{-\infty}^{0} 0dx + \int_{0}^{2} \left(\frac{x^3}{4} - x^2 + x\right) dx +$$

$$+ \int_{2}^{4} \left(-\frac{x^3}{4} + 2x^2 - 5x + 4\right) dx + \int_{4}^{+\infty} 0dx =$$

$$= 0 + \left[\frac{x^4}{16} - \frac{x^3}{3} + \frac{x^2}{2}\right]_0^2 + \left[-\frac{x^4}{16} + \frac{2x^3}{3} - \frac{5x^2}{2} + 4x\right]_2^4 + 0 = 0.67 \text{ anos}^2$$

Pretende-se a variância de X, $\sigma^2 = V[X]$:

- Usando a **propriedade** $V[X] = E\left[X^2\right] - E^2\left[X\right]$ tem-se:

$E[X] = 2$

$$E[X^2] = \int_{-\infty}^{+\infty} x^2 f\left(x\right) dx =$$

$$= \int_{-\infty}^{0} x^2 \times 0dx + \int_{0}^{2} x^2 \times \frac{x}{4} dx + \int_{2}^{4} x^2 \times \left(1 - \frac{x}{4}\right) dx + \int_{4}^{+\infty} x^2 \times 0dx =$$

$$= \int_{-\infty}^{0} 0dx + \int_{0}^{2} \frac{x^3}{4} dx + \int_{2}^{4} \left(x^2 - \frac{x^3}{4}\right) dx + \int_{4}^{+\infty} 0dx =$$

$$= 0 + \left[\frac{x^4}{16}\right]_0^2 + \left[\frac{x^3}{3} - \frac{x^4}{16}\right]_2^4 + 0 =$$

$$= \left(\frac{2^4}{16} - 0\right) + \left[\frac{4^3}{3} - \frac{4^4}{16} - \left(\frac{2^3}{3} - \frac{2^4}{16}\right)\right] = \frac{14}{3}$$

$$V[X] = E\left[X^2\right] - E^2\left[X\right] = \frac{14}{3} - 2^2 = \frac{2}{3} = 0.67 \text{ anos}^2$$

Exemplo 18

De acordo com determinadas especificações técnicas de uma marca de aparelhos de radiologia, sabe-se que a sua duração (em anos) é uma variável aleatória X com a seguinte função densidade de probabilidade:

$$f(x) = \begin{cases} \frac{x}{4} & , \ 0 < x \leq 2 \\ 1 - \frac{x}{4} & , \ 2 < x \leq 4 \\ 0 & , \ \text{caso contrário} \end{cases}$$

9. Calcule o desvio padrão da variável aleatória X.

Exemplo 18

De acordo com determinadas especificações técnicas de uma marca de aparelhos de radiologia, sabe-se que a sua duração (em anos) é uma variável aleatória X com a seguinte função densidade de probabilidade:

$$f(x) = \begin{cases} \frac{x}{4} & , \ 0 < x \leq 2 \\ 1 - \frac{x}{4} & , \ 2 < x \leq 4 \\ 0 & , \ \text{caso contrário} \end{cases}$$

9. Calcule o desvio padrão da variável aleatória X.

Pretende-se o desvio padrão de X:

$$\sigma = \sqrt{V[X]} = \sqrt{\frac{2}{3}} = 0.816 \text{ anos}$$

Exemplo 18

De acordo com determinadas especificações técnicas de uma marca de aparelhos de radiologia, sabe-se que a sua duração (em anos) é uma variável aleatória X com a seguinte função densidade de probabilidade:

$$f(x) = \begin{cases} \frac{x}{4} & , \ 0 < x \leq 2 \\ 1 - \frac{x}{4} & , \ 2 < x \leq 4 \\ 0 & , \ \text{caso contrário} \end{cases}$$

10. Calcule $V[-3X - 2]$.

Exemplo 18

De acordo com determinadas especificações técnicas de uma marca de aparelhos de radiologia, sabe-se que a sua duração (em anos) é uma variável aleatória X com a seguinte função densidade de probabilidade:

$$f(x) = \begin{cases} \frac{x}{4} & , \ 0 < x \leq 2 \\ 1 - \frac{x}{4} & , \ 2 < x \leq 4 \\ 0 & , \ \text{caso contrário} \end{cases}$$

10. Calcule $V[-3X-2]$.

Recorrendo à propriedade $V[aX+b] = a^2 V[X]$ tem-se

$$V[-3X-2] = (-3)^2 V[X] = 9 \times \frac{2}{3} = 6$$

Exemplo 19

A quantidade de álcool (em dezenas de litros) que um produtor-engarrafador vende por dia é uma variável aleatória X com função de distribuição definida por:

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{x^2}{50} & , \quad 0 \leq x < 5 \\ \frac{2}{5}x - \frac{x^2}{50} - 1 & , \quad 5 \leq x < 10 \\ 1 & , \quad x \geq 10 \end{cases}.$$

Qual a quantidade média de álcool vendida por dia?

Exemplo 19

A quantidade de álcool (em dezenas de litros) que um produtor-engarrafador vende por dia é uma variável aleatória X com função de distribuição definida por:

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ \frac{x^2}{50} & , \quad 0 \leq x < 5 \\ \frac{2}{5}x - \frac{x^2}{50} - 1 & , \quad 5 \leq x < 10 \\ 1 & , \quad x \geq 10 \end{cases} .$$

Qual a quantidade média de álcool vendida por dia?

Para o cálculo do valor esperado é necessário conhecer-se a função densidade de probabilidade, anteriormente calculada e dada por

$$f(x) = \begin{cases} \frac{x}{25} & , \quad 0 \le x < 5 \\ \frac{2}{5} - \frac{x}{25} & , \quad 5 \le x < 10 \\ 0 & , \quad \text{caso contrário} \end{cases}$$

Vindo

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} x f(x)\, dx = \\ &= \int_{-\infty}^{0} x \times 0 dx + \int_{0}^{5} x \left(\tfrac{x}{25}\right) dx + \int_{5}^{10} x \left(\tfrac{2}{5} - \tfrac{x}{25}\right) dx + \int_{10}^{+\infty} x \times 0 dx = \\ &= 0 + \left[\frac{x^3}{75}\right]_0^5 + \left[\frac{x^2}{5} - \frac{x^3}{75}\right]_5^{10} + 0 = \\ &= \frac{5^3}{75} - 0 + \frac{10^2}{5} - \frac{10^3}{75} - \left(\frac{5^2}{5} - \frac{5^3}{75}\right) = \\ &= 5 \end{aligned}$$

Logo, a quantidade média de álcool vendida por dia é de 50 litros.

- Agora que já sabem caracterizar todo o modelo probabilístico associado às variáveis aleatórias contínuas, vamos "dar nomes" a alguns desses modelos probabilísticos.

- Vamos analisar pormenorizadamente três modelos probabilísticos ou, como é mais usual dizer, três Distribuições Teóricas Contínuas:
  - Distribuição Exponencial;
  - Distribuição Uniforme Contínua;
  - Distribuição Normal.

Distribuições Teóricas Contínuas:

- **Distribuição Exponencial**;
- Distribuição Uniforme Contínua;
- Distribuição Normal.

Distribuição Exponencial

Definição

Diz-se que uma variável aleatória contínua X tem **distribuição Exponencial** com o parâmetro θ (fixo),

$$X \sim Exp\left(\theta\right),$$

se a sua função densidade de probabilidade é dada por

$$f\left(x\right) = \begin{cases} 0 & , \ x < 0 \\ \frac{1}{\theta}e^{-\frac{x}{\theta}} & , \ x \geq 0 \end{cases} , \quad \theta > 0$$

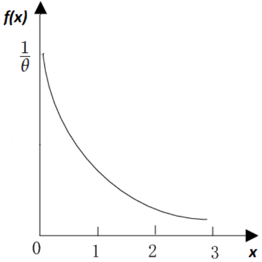
e a sua função de distribuição é dada por

$$F\left(x\right) = P\left(X \leq x\right) = \begin{cases} 0 & , \ x < 0 \\ 1 - e^{-\frac{x}{\theta}} & , \ x \geq 0 \end{cases} , \quad \theta > 0.$$

Distribuição Exponencial

Habitualmente diz-se apenas que a variável aleatória X tem Distribuição Exponencial de parâmetro θ, $X \sim Exp(\theta)$, mas é necessário ter em atenção que é **Distribuição Exponencial Negativa**, pois a sua função densidade de probabilidade é escrita à custa da função exponencial mas com expoente negativo:

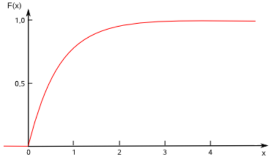
$$f(x) = \begin{cases} 0 & , \ x < 0 \\ \frac{1}{\theta} e^{-\frac{x}{\theta}} & , \ x \geq 0 \end{cases} \quad , \quad \theta > 0$$



Distribuição Exponencial

A função de distribuição de uma variável aleatória contínua $X \sim Exp\,(\theta)$, não tem o aspeto usual: o último ramo ser igual a 1.

$$F\,(x) = \begin{cases} 0 & , \; x < 0 \\ 1 - e^{-\frac{x}{\theta}} & , \; x \geq 0 \end{cases} \quad , \qquad \theta > 0$$



Isso deve-se o facto do domínio (intervalo onde existe probabilidade diferente de zero) ser $D_X = [0, +\infty[$, se calcularmos

$$\lim_{x \to +\infty} F(x) = \lim_{x \to +\infty} (1 - e^{-\frac{x}{\theta}}) = 1 - 0 = 1$$

podemos ver que no "valor máximo" do domínio a função de distribuição é 1.

Distribuição Exponencial

Teorema

Se a variável aleatória X tem distribuição Exponencial com parâmetro θ,

$$X \sim Exp\,(\theta)$$

então

$$E\,[X] = \theta \qquad \text{e} \qquad V\,[X] = \theta^2.$$

Propriedade: "Falta de Memória"

Seja $X \sim Exp\,(\theta)$, então

$$P\,(X \geq a + b \mid X \geq a) = P\,(X \geq b)\,, \qquad a, b > 0.$$

Exemplo 20

O tempo de vida, em horas, de um certo tipo de componente eletrónico tem distribuição exponencial de parâmetro 100.

1. Qual o tempo de vida médio de um destes componentes?

Exemplo 20

O tempo de vida, em horas, de um certo tipo de componente eletrónico tem distribuição exponencial de parâmetro 100.

1. Qual o tempo de vida médio de um destes componentes?

Considere a variável aleatória contínua:

$$X = \text{ tempo de vida, em horas, de certo componente eletrónico}$$

com

$$X \sim Exp\,(100)$$

pois $\theta = 100$.

Portanto

$$E[X] = \theta = 100 \text{ horas.}$$

Exemplo 20

O tempo de vida, em horas, de um certo tipo de componente eletrónico tem distribuição exponencial de parâmetro 100.

2. Qual o desvio padrão do tempo de vida desses componentes?

Exemplo 20

O tempo de vida, em horas, de um certo tipo de componente eletrónico tem distribuição exponencial de parâmetro 100.

2. Qual o desvio padrão do tempo de vida desses componentes?

Considere a variável aleatória contínua:

$$X = \text{ tempo de vida, em horas, de certo componente eletrónico}$$

$$X \sim Exp\,(100)$$

Portanto

$$\sigma^2 = V[X] = \theta^2 = 100^2 = 10000 \text{ horas}^2$$

$$\sigma = \sqrt{V[X]} = \theta = 100 \text{ horas.}$$

Exemplo 20

O tempo de vida, em horas, de um certo tipo de componente eletrónico tem distribuição exponencial de parâmetro 100.

3. Qual a probabilidade de um destes componentes funcionar, sem falhas, pelo menos 200 horas?

Exemplo 20

O tempo de vida, em horas, de um certo tipo de componente eletrónico tem distribuição exponencial de parâmetro 100.

3. Qual a probabilidade de um destes componentes funcionar, sem falhas, pelo menos 200 horas?

Considere a variável aleatória contínua:

$$X = \text{ tempo de vida, em horas, de certo componente eletrónico}$$

$$X \sim Exp\,(100)$$

Portanto

$$F\,(x) = P\,(X \leq x) = \begin{cases} 0 & , \; x < 0 \\ 1 - e^{-\frac{x}{100}} & , \; x \geq 0 \end{cases}$$

Pretende-se

$$P(X \geq 200) = 1 - P(X < 200) \underset{\text{v.a. contínua}}{=} 1 - F(200) = 1 - 0.8647 = 0.1353$$

Para calcular os valores da função de distribuição vamos ver duas possibilidades:

1. **cálculos diretos**

$$F(200) = 1 - e^{-\frac{200}{100}} = 0.8647$$

2. **com recurso ao R**:
   - função densidade de probabilidade: **dexp**$\left(x, rate = \frac{1}{\theta}\right)$
   - função de distribuição: **pexp**$\left(x, rate = \frac{1}{\theta}\right)$
   - inversa da função de distribuição: **qexp**$\left(p, rate = \frac{1}{\theta}\right)$

$$F(200) = pexp\,(200, 1/100) = 0.8647$$

Exemplo 20

O tempo de vida, em horas, de um certo tipo de componente eletrónico tem distribuição exponencial de parâmetro 100.

4. Qual o tempo máximo de vida de 75% destes componentes?

Exemplo 20

O tempo de vida, em horas, de um certo tipo de componente eletrónico tem distribuição exponencial de parâmetro 100.

4. Qual o tempo máximo de vida de 75% destes componentes?

Considere a variável aleatória contínua:

$$X = \text{ tempo de vida, em horas, de certo componente eletrónico}$$

$$X \sim Exp\,(100)$$

Portanto

$$F\,(x) = P\,(X \leq x) = \begin{cases} 0 & , \; x < 0 \\ 1 - e^{-\frac{x}{100}} & , \; x \geq 0 \end{cases}$$

Pretende-se determinar m tal que

$$P(X \leq m) = 0.75 \Leftrightarrow F(m) = 0.75 \Leftrightarrow m = F^{-1}(0.75) = 138.6294 \text{ horas}$$

Temos duas possibilidades de resolução:

1. **cálculos diretos**

$$F(m) = 0.75 \Leftrightarrow 1 - e^{-\frac{m}{100}} = 0.75 \Leftrightarrow m = -100 \times ln(-(0.75 - 1)) = 138.6294$$

2. **com recurso ao R**:

$$m = F^{-1}(0.75) = qexp(0.75, 1/100) = 138.6294$$

Observação: no R a função "ln()" é "log()".

Exemplo 20

O tempo de vida, em horas, de um certo tipo de componente eletrónico tem distribuição exponencial de parâmetro 100.

5. Qual a probabilidade de um destes componentes funcionar, sem falhas, pelo menos 500 horas sabendo que já está a funcionar há pelo menos 300 horas?

Exemplo 20

O tempo de vida, em horas, de um certo tipo de componente eletrónico tem distribuição exponencial de parâmetro 100.

5. Qual a probabilidade de um destes componentes funcionar, sem falhas, pelo menos 500 horas sabendo que já está a funcionar há pelo menos 300 horas?

Considere a variável aleatória contínua:

$$X = \text{ tempo de vida, em horas, de certo componente eletrónico}$$

$$X \sim Exp\,(100)$$

Portanto

$$F\,(x) = P\,(X \leq x) = \begin{cases} 0 & , \; x < 0 \\ 1 - e^{-\frac{x}{100}} & , \; x \geq 0 \end{cases}$$

logo

$$P(X \geq 500 | X \geq 300) \underset{(*)}{=} P(X \geq 500 - 300) = P(X \geq 200) = 0.1353$$

$(*)$ Propriedade "Falta de memória" da distribuição Exponencial.

Claro que era possível calcular a probabilidade recorrendo à definição de probabilidade condicional:

$$P(X \geq 500 | X \geq 300) = \frac{P(X \geq 500 \wedge X \geq 300)}{P(X \geq 300)} = \frac{P(X \geq 500)}{P(X \geq 300)} =$$

$$= \frac{1 - P(X < 500)}{1 - P(X < 300)} \underset{\text{v.a. contínua}}{=} \frac{1 - F(500)}{1 - F(300)} =$$

$$= \frac{1 - \left(1 - e^{-\frac{500}{100}}\right)}{1 - \left(1 - e^{-\frac{300}{100}}\right)} = 0.1353$$

Recorrendo ao R:

$$F(300) = pexp\left(300, \frac{1}{100}\right) = 0.9502$$

$$F(500) = pexp\left(500, \frac{1}{100}\right) = 0.9933$$

Relação entre a Distribuição Exponencial e a Distribuição de Poisson

Se

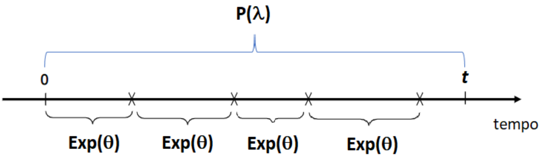
$$X \sim P(\lambda)$$
$$X - \text{ número de ocorrências num intervalo de tempo } t$$
$$\lambda = \text{ número médio de ocorrências num intervalo de tempo } t$$

e

$$Y \sim Exp(\theta)$$
$$Y - \text{ tempo de espera entre ocorrências sucessivas}$$
$$\theta = \text{ tempo de espera médio entre ocorrências sucessivas,}$$

então

$$\theta = \frac{t}{\lambda} \Leftrightarrow \lambda = \frac{t}{\theta}.$$



Exemplo 21

Uma máquina que funciona em contínuo tem, em média, 2 avarias por cada turno de 8 horas e o número de avarias segue uma distribuição de Poisson. Determine a probabilidade de que o tempo entre avarias consecutivas na máquina seja superior a 5 horas.

Exemplo 21

Uma máquina que funciona em contínuo tem, em média, 2 avarias por cada turno de 8 horas e o número de avarias segue uma distribuição de Poisson. Determine a probabilidade de que o tempo entre avarias consecutivas na máquina seja superior a 5 horas.

$$X = \text{ número de avarias por cada turno de 8 horas, com } X \sim P\,(2)$$

pois $E[X] = \lambda = 2$ avarias/turno.

$Y =$ tempo, em horas, entre avarias consecutivas na máquina, com $Y \sim Exp\,(4)$

pois, recorrendo à relação entre as distribuições Poisson e Exponencial, tem-se $\theta = \frac{t}{\lambda} = \frac{8}{2} = 4$ horas/avaria. Como

$$F_Y(y) = P\,(Y \leq y) = \begin{cases} 0 & ,\ y < 0 \\ 1 - e^{-\frac{y}{4}} & ,\ y \geq 0 \end{cases}$$

tem-se

$$P(Y > 5) = 1 - P(Y \leq 5) = 1 - F_Y(5) = 1 - \left(1 - e^{-\frac{5}{4}}\right) = 0.2865$$

Exemplo 22

O tempo, em minutos, entre a chegada consecutiva de utentes a um centro de saúde é uma variável aleatória Exponencial com média 90 segundos. Qual a probabilidade de chegarem pelo menos 3 utentes em 6 minutos?

Exemplo 22

O tempo, em minutos, entre a chegada consecutiva de utentes a um centro de saúde é uma variável aleatória Exponencial com média 90 segundos. Qual a probabilidade de chegarem pelo menos 3 utentes em 6 minutos?

$$Y = \text{ tempo, em minutos, entre a chegada de utentes, com } Y \sim Exp\,(1.5)$$

pois $E[Y] = \theta = \frac{90}{60} = 1.5$ minutos/chegada.

$$X = \text{ número de utentes que chegam em 6 minutos, com } X \sim P\,(4)$$

pois, recorrendo à relação entre as distribuições Poisson e Exponencial, tem-se $\lambda = \frac{t}{\theta} = \frac{6}{1.5} = 4$ chegadas/períodos de 6 minutos. Tem-se

$$P(X \geq 3) = 1 - P(X < 3) = 1 - P(X \leq 2) = 1 - F_X(2) = 1 - 0.2381 = 0.7619$$

pois

$$F_X(2) = ppois(2, 4) = 0.2381$$

Distribuições Teóricas Contínuas:

- Distribuição Exponencial;
- **Distribuição Uniforme Contínua**;
- Distribuição Normal.

Distribuição Uniforme Contínua

Definição

Diz-se que uma variável aleatória contínua X, definida no intervalo real $[a, b]$, tem **distribuição Uniforme Contínua**,

$$X \sim U_{(a,b)},$$

se a sua função densidade de probabilidade é dada por

$$f(x) = \begin{cases} \frac{1}{b-a} & , \ a \leq x \leq b \\ 0 & , \ \text{caso contrário} \end{cases}$$

e a sua função de distribuição é dada por

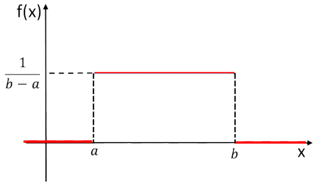
$$F(x) = P(X \leq x) = \begin{cases} 0 & , \ x < a \\ \frac{x-a}{b-a} & , \ a \leq x \leq b \\ 1 & , \ x > b \end{cases}.$$

A distribuição Uniforme Contínua é idêntica à distribuição Uniforme Discreta, a única diferença é a variável aleatória ser contínua, logo em vez de pontos isolados tem-se um intervalo, $D_X = [a, b]$. Portanto, diz-se que uma variável aleatória contínua tem distribuição Uniforme Contínua,

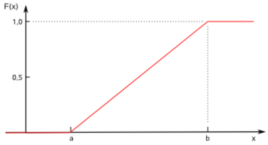
$$X \sim U_{(a,b)}$$

se assume em todo o intervalo, que define o seu domínio, a mesma probabilidade. Graficamente tem-se:

$$f(x) = \begin{cases} \frac{1}{b-a} & , a \leq x \leq b \\ 0 & , \text{caso contrário} \end{cases}$$



$$F(x) = P(X \leq x) = \begin{cases} 0 & , x < a \\ \frac{x-a}{b-a} & , a \leq x \leq b \\ 1 & , x > b \end{cases}$$



Distribuição Uniforme Contínua

Teorema

Se a variável aleatória X tem distribuição Uniforme Contínua, $X \sim U_{(a,b)}$, então

$$E[X] = \frac{a+b}{2} \qquad \text{e} \qquad V[X] = \frac{(b-a)^2}{12}.$$

Exemplo 23

Suponha que o tempo, em minutos, necessário para que um técnico afine cada componente de uma máquina numa linha de produção é bem descrito por uma variável aleatória X com distribuição uniforme no intervalo $[5, 10]$.

1. Qual a probabilidade de que uma componente escolhida ao acaso tenha necessitado de mais de 7 minutos para ser afinada?

Exemplo 23

Suponha que o tempo, em minutos, necessário para que um técnico afine cada componente de uma máquina numa linha de produção é bem descrito por uma variável aleatória X com distribuição uniforme no intervalo $[5, 10]$.

1. Qual a probabilidade de que uma componente escolhida ao acaso tenha necessitado de mais de 7 minutos para ser afinada?

Considere a variável aleatória contínua:

$$X = \text{ tempo, em minutos, para afinar uma componente, com } X \sim U_{(5,10)}$$

logo

$$F(x) = P(X \leq x) = \begin{cases} 0 & , \ x < 5 \\ \frac{x-5}{10-5} & , \ 5 \leq x \leq 10 \\ 1 & , \ x > 10 \end{cases} \Leftrightarrow F(x) = \begin{cases} 0 & , \ x < 5 \\ \frac{x-5}{5} & , \ 5 \leq x \leq 10 \\ 1 & , \ x > 10 \end{cases}$$

Portanto

$$P(X > 7) = 1 - P(X \leq 7) = 1 - F(7) = 1 - 0.4 = 0.6$$

Para calcular os valores da função de distribuição vamos ver duas possibilidades:

1. **cálculos diretos**

$$F(7) = \frac{7-5}{5} = \frac{2}{5} = 0.4$$

2. **com recurso ao R**:
   - função densidade de probabilidade: **dunif**$(x, min = a, max = b)$
   - função de distribuição: **punif**$(x, min = a, max = b)$
   - inversa da função de distribuição: **qunif**$(p, min = a, max = b)$

$$F(7) = punif\,(7, 5, 10) = 0.4$$

Exemplo 23

Suponha que o tempo, em minutos, necessário para que um técnico afine cada componente de uma máquina numa linha de produção é bem descrito por uma variável aleatória X com distribuição uniforme no intervalo $[5, 10]$.

2. Qual a probabilidade de que uma componente escolhida ao acaso tenha exigido ao técnico um tempo de afinação inferior a 9 minutos sabendo que aquele tempo foi superior a 7 minutos?

Exemplo 23

Suponha que o tempo, em minutos, necessário para que um técnico afine cada componente de uma máquina numa linha de produção é bem descrito por uma variável aleatória X com distribuição uniforme no intervalo $[5, 10]$.

2. Qual a probabilidade de que uma componente escolhida ao acaso tenha exigido ao técnico um tempo de afinação inferior a 9 minutos sabendo que aquele tempo foi superior a 7 minutos?

$X =$ tempo, em minutos, para afinar uma componente, com $X \sim U_{(5,10)}$

$$F(x) = P(X \leq x) = \begin{cases} 0 & , \ x < 5 \\ \frac{x-5}{5} & , \ 5 \leq x \leq 10 \\ 1 & , \ x > 10 \end{cases}$$

Portanto

$$P(X < 9 | X > 7) = \frac{P(X < 9 \wedge X > 7)}{P(X > 7)} = \frac{P(7 < X < 9)}{1 - P(X \leq 7)} \underset{\text{v.a. contínua}}{=}$$

$$= \frac{F(9) - F(7)}{1 - F(7)} = \frac{2}{3} = 0.6667$$

Para calcular os valores da função de distribuição vamos ver duas possibilidades:

1. **cálculos diretos**
   - $F(7) = \frac{7-5}{5} = 0.4$
   - $F(9) = \frac{9-5}{5} = 0.8$
2. **com recurso ao R**:
   - $F(7) = punif(7, 5, 10) = 0.4$
   - $F(9) = punif(9, 5, 10) = 0.8$

Exemplo 23

Suponha que o tempo, em minutos, necessário para que um técnico afine cada componente de uma máquina numa linha de produção é bem descrito por uma variável aleatória X com distribuição uniforme no intervalo $[5, 10]$.

3. Determine k sabendo que a probabilidade de que uma componente escolhida ao acaso tenha necessitado no máximo de k minutos para ser afinada é 0.80.

Exemplo 23

Suponha que o tempo, em minutos, necessário para que um técnico afine cada componente de uma máquina numa linha de produção é bem descrito por uma variável aleatória X com distribuição uniforme no intervalo $[5, 10]$.

3. Determine k sabendo que a probabilidade de que uma componente escolhida ao acaso tenha necessitado no máximo de k minutos para ser afinada é 0.80.

Considere a variável aleatória contínua:

$$X = \text{ tempo, em minutos, para afinar uma componente, com } X \sim U_{(5,10)}$$

logo

$$F(x) = P(X \leq x) = \begin{cases} 0 & , \ x < 5 \\ \frac{x-5}{10-5} & , \ 5 \leq x \leq 10 \\ 1 & , \ x > 10 \end{cases} \Leftrightarrow F(x) = \begin{cases} 0 & , \ x < 5 \\ \frac{x-5}{5} & , \ 5 \leq x \leq 10 \\ 1 & , \ x > 10 \end{cases}$$

Pretende-se determinar k tal que

$$P(X \leq k) = 0.80 \Leftrightarrow F(k) = 0.80 \Leftrightarrow k = F^{-1}(0.80) = 9 \text{ minutos}$$

Vamos ver duas possibilidades para calcular o valor pretendido:

1. **cálculos diretos**

$$F(k) = 0.80 \Leftrightarrow \frac{k-5}{5} = 0.80 \Leftrightarrow k = 5 + 5 \times 0.80 = 9$$

2. **com recurso ao R**:

$$F^{-1}(0.80) = qunif(0.80, 5, 10) = 9$$

Exemplo 23

Suponha que o tempo, em minutos, necessário para que um técnico afine cada componente de uma máquina numa linha de produção é bem descrito por uma variável aleatória X com distribuição uniforme no intervalo $[5, 10]$.

4. Seja $Y = 20 + 3X$ o custo, em euros, de afinação de cada componente. Qual o custo esperado de afinação de uma componente escolhida ao acaso? E qual o desvio padrão?

Exemplo 23

Suponha que o tempo, em minutos, necessário para que um técnico afine cada componente de uma máquina numa linha de produção é bem descrito por uma variável aleatória X com distribuição uniforme no intervalo $[5, 10]$.

4. Seja $Y = 20 + 3X$ o custo, em euros, de afinação de cada componente. Qual o custo esperado de afinação de uma componente escolhida ao acaso? E qual o desvio padrão?

$$X = \text{ tempo, em minutos, para afinar uma componente, com } X \sim U_{(5,10)}$$

$$Y = \text{ custo, em euros, de afinação de cada componente, com } Y = 20 + 3X$$

$$\mu_Y = E[Y] = E[20 + 3X] = 20 + 3E[X] = 20 + 3 \times \frac{5 + 10}{2} = 42.5 \text{ euros}$$

$$\sigma_Y^2 = V[Y] = V[20 + 3X] = 3^2 V[X] = 3^2 \times \frac{(10 - 5)^2}{12} = 18.75 \text{ euros}^2$$

$$\sigma_Y = \sqrt{V[Y]} = \sqrt{18.75} = 4.3301 \text{ euros}$$

Distribuições Teóricas Contínuas:

- Distribuição Exponencial;
- Distribuição Uniforme Contínua;
- **Distribuição Normal**.

Distribuição Normal

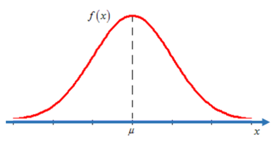
Definição

Diz-se que uma variável aleatória contínua X tem **distribuição Normal (ou Gaussiana)** com parâmetros μ e σ (fixos), e representa-se por

$$X \sim N(\mu, \sigma)$$

se a sua função densidade de probabilidade é dada por

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \mu \in \mathbb{R}, \quad \sigma > 0, \quad x \in \mathbb{R}$$



e a sua função de distribuição é dada por

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt.$$

Distribuição Normal

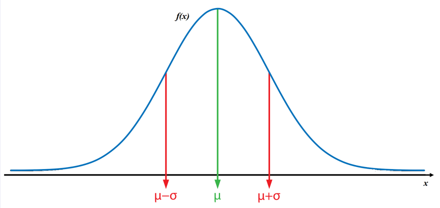
Teorema

Se a variável aleatória X tem distribuição Normal com parâmetros μ e σ, $X \sim N(\mu, \sigma)$, então

$$E[X] = \mu \qquad \text{e} \qquad V[X] = \sigma^2.$$

Propriedades: associadas à representação gráfica da função densidade de probabilidade, $f(x)$

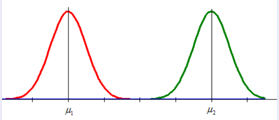
- É simétrica relativamente a μ.
- Atinge um máximo absoluto em $x = \mu$.
- $\mu \pm \sigma$ são os pontos de inflexão da curva.
- O eixo dos xx é uma assimptota horizontal.



Distribuição Normal

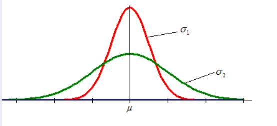
Propriedades: associadas à representação gráfica da função densidade de probabilidade, $f(x)$

- A curva é simétrica em relação ao valor esperado μ:



o mesmo desvio padrão (σ), valores esperados diferentes $(\mu_1 < \mu_2)$

- A curva é tanto mais achatada quanto maior o valor do desvio padrão σ:



o mesmo valor esperado (μ), desvios padrão diferentes $(\sigma_1 < \sigma_2)$

Distribuição Normal Reduzida

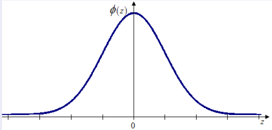
Definição

Diz-se que uma variável aleatória contínua Z tem **distribuição Normal Reduzida (standard ou padrão)** se a variável aleatória Z tem distribuição Normal com os parâmetros $\mu = 0$ e $\sigma = 1$

$$Z \sim N(0, 1).$$

A sua função densidade de probabilidade é dada por

$$\phi(z) = f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad z \in \mathbb{R}$$



e a sua função de distribuição é dada por

$$\Phi(z) = F(z) = P(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt.$$

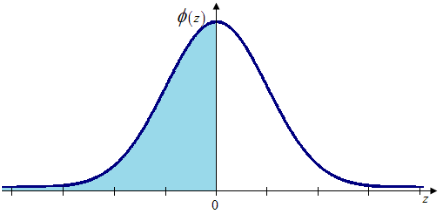
Distribuição Normal Reduzida: $\mathbf{Z} \sim \mathbf{N}\left(\mathbf{0}, \mathbf{1}\right)$

- A variável aleatória representa-se habitualmente pela letra Z.
- A função densidade de probabilidade representa-se habitualmente pela letra ϕ em vez de f: $\phi(z) = f\left(z\right)$
- a função distribuição representa-se habitualmente pela letra Φ em vez de F:

$$\Phi(z) = P\left(Z \leq z\right)$$

- O eixo de simetria é o eixo dos yy:

$$P\left(Z \leq 0\right) = P\left(Z \geq 0\right) = 0.5$$

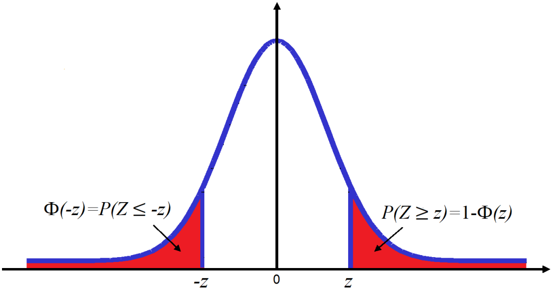


área total = 1 e área sombreada = 0.5

Distribuição Normal Reduzida: $\mathbf{Z} \sim \mathbf{N}\left(\mathbf{0}, \mathbf{1}\right)$

Propriedade

Se $Z \sim N\left(0, 1\right)$, então $\Phi\left(-z\right) = 1 - \Phi\left(z\right)$.



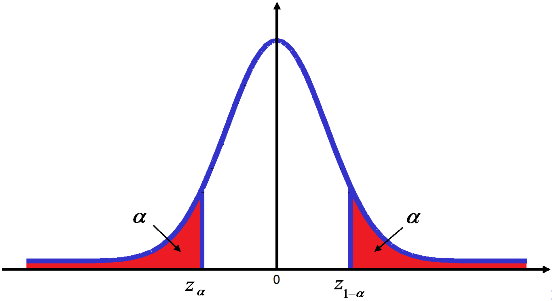
Distribuição Normal Reduzida: $\mathbf{Z} \sim \mathbf{N}(\mathbf{0}, \mathbf{1})$

- Seja z_α tal que

$$P(Z \leq z_\alpha) = \alpha \Leftrightarrow \Phi(z_\alpha) = \alpha \Leftrightarrow z_\alpha = \Phi^{-1}(\alpha)$$

então a z_α chama-se **quantil de probabilidade** α da distribuição Normal reduzida.

- Os quantis de probabilidade são simétricos: $z_\alpha = -z_{1-\alpha}$

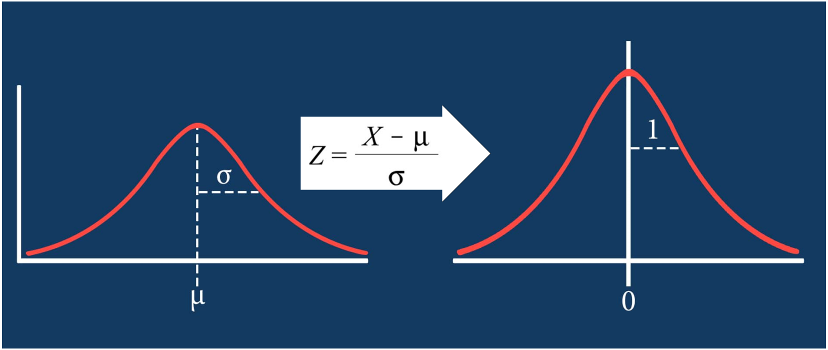


Distribuição Normal

Propriedade

Se $X \sim N(\mu, \sigma)$, então $Z = \frac{X-\mu}{\sigma} \sim N(0, 1)$.

Todas as Distribuições Normais podem ser transformadas na Distribuição Normal Reduzida:



1. **com recurso ao R**:
   - $X \sim N(\mu, \sigma)$
     - ⋆ função densidade de probabilidade: **dnorm**$(x, mean = \mu, sd = \sigma)$
     - ⋆ função de distribuição: **pnorm**$(x, mean = \mu, sd = \sigma)$
     - ⋆ inversa da função de distribuição: **qnorm**$(p, mean = \mu, sd = \sigma)$
   - $Z \sim N(0, 1)$
     - ⋆ função densidade de probabilidade: **dnorm**(x)
     - ⋆ função de distribuição: **pnorm**(x)
     - ⋆ inversa da função de distribuição ou quantil de probabilidade da normal reduzida: **qnorm**(p)

Exemplo 24 (estava Exemplo 29)

Suponha que o QI de uma pessoa adulta segue uma distribuição normal com valor médio 100 e desvio padrão 15. Um indivíduo que tenha um QI superior a 115 considera-se que possui um QI elevado.

1. Qual é a probabilidade de um certo indivíduo ter um QI elevado?

Exemplo 24 (estava Exemplo 29)

Suponha que o QI de uma pessoa adulta segue uma distribuição normal com valor médio 100 e desvio padrão 15. Um indivíduo que tenha um QI superior a 115 considera-se que possui um QI elevado.

1. Qual é a probabilidade de um certo indivíduo ter um QI elevado?

Considere a variável aleatória contínua:

$$X = \text{ QI de uma pessoa adulta, com } X \sim N(100, 15)$$

pois $\mu = E[X] = 100$ e $\sigma = \sqrt{V[X]} = 15$.

$$P(X > 115) = 1 - P(X \leq 115) = 1 - F(115) = 1 - 0.8413 = 0.1587$$

pois

$$F(115) = pnorm(115, 100, 15) = 0.8413$$

Exemplo 24 (estava Exemplo 29)

Suponha que o QI de uma pessoa adulta segue uma distribuição normal com valor médio 100 e desvio padrão 15. Um indivíduo que tenha um QI superior a 115 considera-se que possui um QI elevado.

2. Qual é a probabilidade de um certo indivíduo ter um QI entre 80 e 110?

Exemplo 24 (estava Exemplo 29)

Suponha que o QI de uma pessoa adulta segue uma distribuição normal com valor médio 100 e desvio padrão 15. Um indivíduo que tenha um QI superior a 115 considera-se que possui um QI elevado.

2. Qual é a probabilidade de um certo indivíduo ter um QI entre 80 e 110?

$$X = \text{ QI de uma pessoa adulta, com } X \sim N(100, 15)$$

$$P(80 < X < 110) \underset{\text{v.a. contínua}}{=} F(110) - F(80) = 0.7475 - 0.0912 = 0.6563$$

pois

- $F(110) = pnorm(110, 100, 15) = 0.7475$
- $F(80) = pnorm(80, 100, 15) = 0.0912$

Exemplo 24

Suponha que o QI de uma pessoa adulta segue uma distribuição normal com valor médio 100 e desvio padrão 15. Um indivíduo que tenha um QI superior a 115 considera-se que possui um QI elevado.

3. Qual o valor máximo do QI que possui 80% dos adultos?

Exemplo 24

Suponha que o QI de uma pessoa adulta segue uma distribuição normal com valor médio 100 e desvio padrão 15. Um indivíduo que tenha um QI superior a 115 considera-se que possui um QI elevado.

3. Qual o valor máximo do QI que possui 80% dos adultos?

$$X = \text{ QI de uma pessoa adulta, com } X \sim N(100, 15)$$

Pretende-se determinar m tal que

$$P(X \leq m) = 0.80 \Leftrightarrow F(m) = 0.80 \Leftrightarrow m = F^{-1}(0.80) = 112.6243$$

pois

- $F^{-1}(0.80) = qnorm(0.80, 100, 15) = 112.6243$

Exemplo 25

A altura (em metros) a que crescem os pinheiros é uma variável aleatória X normalmente distribuída com desvio padrão igual a 1.1 metros. Supondo que 90% dos pinheiros atingem uma altura de pelo menos 16 metros, qual a altura média dos pinheiros?

Exemplo 25

A altura (em metros) a que crescem os pinheiros é uma variável aleatória X normalmente distribuída com desvio padrão igual a 1.1 metros. Supondo que 90% dos pinheiros atingem uma altura de pelo menos 16 metros, qual a altura média dos pinheiros?

Considera a variável aleatória contínua:

$$X = \text{ altura, em metros, dos pinheiros, } \text{com } X \sim N(\mu, 1.1)$$

pois $\sigma = \sqrt{V[X]} = 1.1$ metros. Como pretende-se determinar $\mu = E[X]$ é necessário recorrer à distribuição Normal Reduzida:

$$Z = \frac{X - \mu}{1.1} \sim N(0,1)$$

Sabendo que

$$P(X \geq 16) = 0.90 \Leftrightarrow 1 - P(X < 16) = 0.90 \underset{\text{v.a. contínua}}{\Leftrightarrow} 1 - F(16) = 0.90 \Leftrightarrow F(16) = 0.10 \Leftrightarrow$$

$$\Leftrightarrow \Phi\left(\frac{16-\mu}{1.1}\right) = 0.10 \Leftrightarrow \frac{16-\mu}{1.1} = z_{0.10} \Leftrightarrow \frac{16-\mu}{1.1} = -1.282 \Leftrightarrow \mu = 17.41 \text{metros}$$

pois $z_{0.10} = qnorm(0.10) = -1.282$

Distribuição Normal

Propriedade 1: Aditividade da Normal

Sejam $X_1, X_2, \ldots, X_k$ variáveis aleatórias independentes com distribuição Normal, isto é

$$X_i \sim N\left(\mu_i, \sigma_i\right) \quad i = 1, \ldots, k,$$

então

$$Y = \sum_{i=1}^{k} X_i$$

também é uma variável aleatória com distribuição Normal, isto é

$$Y = \sum_{i=1}^{k} X_i \sim N\left(\sum_{i=1}^{k} \mu_i, \sqrt{\sum_{i=1}^{k} \sigma_i^2}\right).$$

Distribuição Normal

Propriedade 2: Combinação Linear da Normal

Qualquer combinação linear de variáveis aleatórias independentes com distribuição Normal, ainda tem distribuição Normal, isto é

$$X_i \sim N\left(\mu_i, \sigma_i\right) \quad i = 1, \ldots, k, \quad \text{independentes}$$

então

$$Y = \sum_{i=1}^{k} a_i X_i \sim N\left(\sum_{i=1}^{k} a_i \mu_i, \sqrt{\sum_{i=1}^{k} a_i^2 \sigma_i^2}\right).$$

Exemplo 26

Uma empresa, com várias lojas a nível nacional, realiza anualmente uma espécie de concurso nacional em que os participantes são os funcionários com mais vendas em cada loja. Uma das provas desse concurso é classificada, de forma independente, por dois jurados, J1 e J2, numa escala de 0 a 100. Seja X_1 a variável aleatória que representa a pontuação atribuída pelo jurado J1 e seja X_2 a variável aleatória que representa a pontuação atribuída pelo jurado J2. Sabe-se que X_1 tem distribuição normal com média 84 e variância 25 e que X_2 também é normalmente distribuída com média 85 e desvio padrão 8.

1. Calcule a probabilidade da soma das pontuações atribuídas pelos dois jurados exceder 176.

Exemplo 26

Uma empresa, com várias lojas a nível nacional, realiza anualmente uma espécie de concurso nacional em que os participantes são os funcionários com mais vendas em cada loja. Uma das provas desse concurso é classificada, de forma independente, por dois jurados, J1 e J2, numa escala de 0 a 100. Seja X_1 a variável aleatória que representa a pontuação atribuída pelo jurado J1 e seja X_2 a variável aleatória que representa a pontuação atribuída pelo jurado J2. Sabe-se que X_1 tem distribuição normal com média 84 e variância 25 e que X_2 também é normalmente distribuída com média 85 e desvio padrão 8.

1. Calcule a probabilidade da soma das pontuações atribuídas pelos dois jurados exceder 176.

$$X_1 = \text{ pontuação atribuída pelo jurado J1}, \text{ com } X_1 \sim N(84, 5)$$

pois $\mu_1 = E[X_1] = 84$ e $\sigma_1 = \sqrt{V[X_1]} = \sqrt{25} = 5$

$$X_2 = \text{ pontuação atribuída pelo jurado J2}, \text{ com } X_2 \sim N(85, 8)$$

pois $\mu_2 = E[X_2] = 85$ e $\sigma_2 = \sqrt{V[X_2]} = 8$

Seja

$$Y = \text{ soma das pontuações atribuídas pelos dois jurados, com } Y = X_1 + X_2$$

Como "...provas desse concurso é classificada, de forma independente, por dois jurados...", então X_1 e X_2 podem ser consideradas variáveis aleatórias independentes. Recorrendo à propriedade Aditividade da Distribuição Normal tem-se

$$Y = X_1 + X_2 \sim N\left(169, \sqrt{89}\right)$$

pois

$$\mu_Y = E\left[Y\right] = E[X_1 + X_2] = \sum_{i=1}^{2}\mu_i = \mu_1 + \mu_2 = 84 + 85 = 169$$

$$\sigma_Y^2 = V\left[Y\right] = V[X_1 + X_2] = \sum_{i=1}^{2}\sigma_i^2 = \sigma_1^2 + \sigma_2^2 = 5^2 + 8^2 = 89$$

$$\sigma_Y = \sqrt{V\left[Y\right]} = \sqrt{89}$$

logo

$$P(Y > 176) = 1 - P(Y \leq 176) = 1 - F_Y(176) = 1 - 0.771 = 0.229$$

pois $F_Y(176) = pnorm(176, 169, sqrt(89)) = 0.771$

Exemplo 26

Uma empresa, com várias lojas a nível nacional, realiza anualmente uma espécie de concurso nacional em que os participantes são os funcionários com mais vendas em cada loja. Uma das provas desse concurso é classificada, de forma independente, por dois jurados, J1 e J2, numa escala de 0 a 100. Seja X_1 a variável aleatória que representa a pontuação atribuída pelo jurado J1 e seja X_2 a variável aleatória que representa a pontuação atribuída pelo jurado J2. Sabe-se que X_1 tem distribuição normal com média 84 e desvio padrão 5 e que X_2 também é normalmente distribuída com média 85 e desvio padrão 8.

2. O jurado J2 tem fama de ser MUITO mais generoso do que o jurado J1. Calculando uma probabilidade que lhe pareça adequada, estude se essa reputação é exagerada.

Exemplo 26

Uma empresa, com várias lojas a nível nacional, realiza anualmente uma espécie de concurso nacional em que os participantes são os funcionários com mais vendas em cada loja. Uma das provas desse concurso é classificada, de forma independente, por dois jurados, J1 e J2, numa escala de 0 a 100. Seja X_1 a variável aleatória que representa a pontuação atribuída pelo jurado J1 e seja X_2 a variável aleatória que representa a pontuação atribuída pelo jurado J2. Sabe-se que X_1 tem distribuição normal com média 84 e desvio padrão 5 e que X_2 também é normalmente distribuída com média 85 e desvio padrão 8.

2. O jurado J2 tem fama de ser MUITO mais generoso do que o jurado J1. Calculando uma probabilidade que lhe pareça adequada, estude se essa reputação é exagerada.

$$X_1 = \text{ pontuação atribuída pelo jurado J1, com } X_1 \sim N(84, 5)$$

$$X_2 = \text{ pontuação atribuída pelo jurado J2, com } X_2 \sim N(85, 8)$$

Vamos calcular

$$P(X_1 < X_2) = P(X_1 - X_2 < 0)$$

se este valor for muito elevado, então a reputação não deverá ser exagerada.

Seja

$$W = X_1 - X_2$$

Como "...provas desse concurso é classificada, de forma independente, por dois jurados...", então X_1 e X_2 podem ser consideradas variáveis aleatórias independentes. Recorrendo à propriedade 2 Combinação Linear da Distribuição Normal tem-se

$$W = X_1 - X_2 = X_1 + (-1)X_2 \sim N\left(-1, \sqrt{89}\right)$$

pois

$$\begin{aligned}
\mu_W &= E[W] = E[X_1 + (-1)X_2] = \mu_1 + (-1)\mu_2 = 84 - 85 = -1 \\
\sigma_W^2 &= V[W] = V[X_1 + (-1)X_2] = \sigma_1^2 + (-1)^2\sigma_2^2 = 5^2 + (-1)^2 \times 8^2 = 89 \\
\sigma_W &= \sqrt{V[W]} = \sqrt{89}
\end{aligned}$$

logo

$$P(X_1 < X_2) = P(X_1 - X_2 < 0) = P(W < 0) \underset{\text{v.a. contínua}}{=} F_W(0) = 0.5422$$

com $F_W(0) = pnorm(0, -1, sqrt(89)) = 0.5422$

A reputação parece ser exagerada pois apenas em cerca de 54% da vezes é que a pontuação do J2 é superior à do J1.

Exemplo 27

Uma fábrica produz e comercializa rolos de cabos elétricos cuja a capacidade, em metros, é uma variável aleatória normal com valor médio 100 metros e variância $121\ m^2$. Sabendo que o fornecimento é feito em caixas de 300 rolos, calcule a probabilidade de uma caixa conter mais de $30.5\ km$ de cabo elétrico.

Exemplo 27

Uma fábrica produz e comercializa rolos de cabos elétricos cuja a capacidade, em metros, é uma variável aleatória normal com valor médio 100 metros e variância $121\ m^2$. Sabendo que o fornecimento é feito em caixas de 300 rolos, calcule a probabilidade de uma caixa conter mais de $30.5\ km$ de cabo elétrico.

$$X_i = \text{ capacidade, em metros, do rolo } i, \text{ com } X_i \sim N\left(100, 11\right)$$

pois $\mu_i = E[X_i] = 100$ e $\sigma_i = \sqrt{V[X_i]} = \sqrt{121} = 11, \quad i = 1, 2, \ldots, 300$

$$T = \text{ capacidade, em metros, de uma caixa }, \text{ com } T = \sum_{i=1}^{300} X_i$$

Como uma caixa é constituída por 300 rolos diferentes, então $X_1, X_2, \ldots, X_{300}$ podem ser consideradas variáveis aleatórias independentes. Recorrendo à propriedade Aditividade da Distribuição Normal tem-se

$$T = \sum_{i=1}^{300} X_i \sim N\left(30000, \sqrt{36300}\right)$$

pois

$$\mu_T = E[T] = E\left[\sum_{i=1}^{300} X_i\right] = \sum_{i=1}^{300} \mu_i = \sum_{i=1}^{300} 100 = 300 \times 100 = 30000$$

$$\sigma_T^2 = V[T] = V\left[\sum_{i=1}^{300} X_i\right] = \sum_{i=1}^{300} \sigma_i^2 = \sum_{i=1}^{300} 11^2 = 300 \times 11^2 = 36300$$

$$\sigma_T = \sqrt{V[T]} = \sqrt{36300}$$

logo

$$P(T > 30500) = 1 - P(T \leq 30500) = 1 - F_T(30500) = 1 - 0.9956 = 0.0044$$

pois $F_T(30500) = pnorm(30500, 30000, sqrt(36300)) = 0.9956$

Distribuição Normal

Teorema do Limite Central

Se $X_1, X_2, \ldots, X_n$ são variáveis aleatórias independentes e identicamente distribuídas, com média μ e variância σ^2 (finitas), então a distribuição da soma

$$S_n = \sum_{i=1}^{n} X_i = X_1 + X_2 + \cdots + X_n$$

tende a aproximar-se da distribuição Normal quando $n \to +\infty$, isto é

$$S_n = \sum_{i=1}^{n} X_i \dot\sim N\left(n\mu, \sqrt{n\sigma^2}\right) \qquad \Leftrightarrow \qquad S_n = \sum_{i=1}^{n} X_i \dot\sim N\left(n\mu, \sigma\sqrt{n}\right).$$

Observação

Na prática considera-se uma boa aproximação se $n \geq 30$.

Exemplo 28

O conteúdo de certo tipo de garrafas é uma variável aleatória cuja a média é 1 litro e o desvio padrão 0.0201 litros. Se 500 garrafas forem despejadas para um recipiente, qual a probabilidade do recipiente ficar com um conteúdo superior a 500.1 litros?

Exemplo 28

O conteúdo de certo tipo de garrafas é uma variável aleatória cuja a média é 1 litro e o desvio padrão 0.0201 litros. Se 500 garrafas forem despejadas para um recipiente, qual a probabilidade do recipiente ficar com um conteúdo superior a 500.1 litros?

$$X_i = \text{ conteúdo, em litros, da garrafa } i, \quad i = 1, 2, \ldots, 500$$

com $\mu_i = E[X_i] = 1$ e $\sigma_i = \sqrt{V[X_i]} = 0.0201$

$$T = \text{ capacidade, em litros, de um recipiente }, \text{ com } T = \sum_{i=1}^{500} X_i$$

Como são despejadas 500 garrafas diferentes, então $X_1, X_2, \ldots, X_{500}$ podem ser consideradas variáveis aleatórias independentes.

Embora este exemplo seja parecido com o Exemplo 27, não é possível recorrer à propriedade Aditividade da Distribuição Normal pois a distribuição das variáveis aleatórias X_i não é conhecida (para recorrer à aditividade da distribuição Normal era obrigatório ter distribuição Normal).

No entanto sabe-se que as variáveis aleatórias X_i têm o mesmo comportamento probabilístico (a média e a variância são iguais e como todas representam o mesmo têm a mesma distribuição de probabilidade (embora desconhecida), logo diz-se que são identicamente distribuídas).

Portanto tem-se $X_1, X_2, \ldots, X_{500}$ variáveis aleatórias independentes e identicamente distribuídas e $n = 500 \geq 30$, então pelo Teorema do Limite Central tem-se

$$T = \sum_{i=1}^{500} X_i \;\dot{\sim}\; N\left(500, 0.0201\sqrt{500}\right)$$

pois

$$\mu_T = E\left[T\right] = E\left[\sum_{i=1}^{500} X_i\right] = \sum_{i=1}^{500} \mu_i = \sum_{i=1}^{500} 1 = 500 \times 1 = 500$$

$$\sigma_T^2 = V\left[T\right] = V\left[\sum_{i=1}^{500} X_i\right] = \sum_{i=1}^{500} \sigma_i^2 = \sum_{i=1}^{500} 0.0201^2 = 500 \times 0.0201^2$$

$$\sigma_T = \sqrt{V\left[T\right]} = \sqrt{500 \times 0.0201^2} = 0.0201\sqrt{500}$$

logo

$$P(T > 500.1) = 1 - P(T \leq 500.1) = 1 - F_T(500.1) = 1 - 0.588 = 0.412$$

pois $F_T(500.1) = pnorm(500.1, 500, 0.0201 \times sqrt(500)) = 0.588$

Aplicações do Teorema do Limite Central

O Teorema do Limite Central garante que, sob determinadas condições, a distribuição Normal é uma boa aproximação para o cálculo de probabilidades de variáveis aleatórias que representam somas.

A distribuição Binomial e a distribuição de Poisson podem ser entendidas como variáveis aleatórias que representam somas, logo uma das aplicações do Teorema do Limite Central é na utilização da distribuição Normal como aproximação destas distribuições.

Este resultado é muito útil quando se utilizam tabelas em papel. As tabelas em papel da distribuição Binomial apresentam limites em relação aos valores de n e da distribuição de Poisson apresentam limites em relação aos valores de λ.

Teorema: Aproximação da Binomial pela Normal

A distribuição Binomial converge para a distribuição Normal quando $n \rightarrow +\infty$ (o número de provas é muito grande), isto é

$$X \sim B(n, p) \underset{n \rightarrow +\infty}{\longrightarrow} X \dot{\sim} N(np, \sqrt{npq})$$

Teorema: Aproximação da Poisson pela Normal

A distribuição de Poisson converge para a distribuição Normal quando $\lambda \rightarrow +\infty$ (o número médio de ocorrências é muito grande), isto é

$$X \sim P(\lambda) \underset{\lambda \rightarrow +\infty}{\longrightarrow} X \dot{\sim} N\left(\lambda, \sqrt{\lambda}\right)$$

Observações

1. Na prática a distribuição Normal é uma boa aproximação da distribuição Binomial se $n \geq 30$, $np > 5$ e $nq > 5$.

   (Se $n \geq 30$ e $np \leq 5$ (ou $nq \leq 5$), a distribuição de Binomial é aproximada à distribuição de Poisson.)

2. Na prática a distribuição Normal é uma boa aproximação da distribuição de Poisson se $\lambda > 20$.

3. Com esta aproximação estamos a transformar uma variável aleatória discreta numa variável aleatória contínua (onde as probabilidades pontuais são nulas), torna-se necessário proceder à correção por continuidade:

$$P_{v.a.discreta}\,(X = x) \approx P_{Normal}\,(x - 0,5 \leq X < x + 0,5)$$

$$P_{v.a.discreta}\,(X \leq x) \approx P_{Normal}\,(X < x + 0,5).$$

Antes de entrar no próximo capítulo serão referidas mais algumas **distribuições teóricas contínuas** que irão surgir recorrentemente quando se pretende efetuar inferência estatística:

Distribuições Teóricas Contínuas:

- Distribuição Qui-Quadrado;
- Distribuição t de Student;
- Distribuição F de Snedecor.

Distribuições Teóricas Contínuas:

- **Distribuição Qui-Quadrado**;
- Distribuição t de Student;
- Distribuição F de Snedecor.

Distribuição Qui-Quadrado

Definição

Diz-se que uma variável aleatória contínua X tem **distribuição Qui-Quadrado** com n graus de liberdade,

$$X \sim \chi^2_{(n)}$$

se a sua função densidade de probabilidade é dada por

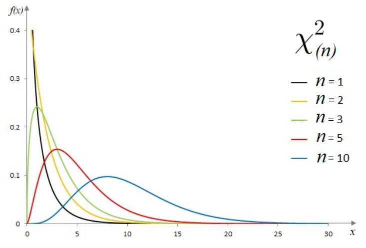
$$f(x) = \frac{e^{-\frac{x}{2}} x^{\frac{n}{2}-1}}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} = \frac{e^{-\frac{x}{2}} x^{\frac{n}{2}-1}}{2^{\frac{n}{2}} \int_0^{+\infty} e^{-x} x^{\frac{n}{2}-1} dx}, \quad n > 0, \quad x > 0$$

e a sua função de distribuição é dada por

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} \frac{e^{-\frac{t}{2}} t^{\frac{n}{2}-1}}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} dt.$$

Distribuição Qui-Quadrado: $X \sim \chi^2_{(n)}$

Graficamente a sua função densidade de probabilidade, $f(x)$, é dada por:



- Domínio: $\mathbb{R}^+$
- Distribuição assimétrica com cauda longa do lado direito.
- quanto maior n (graus de liberdade) menos assimétrica é a distribuição

Propriedade

Sejam $X_1, X_2, \ldots, X_n$ variáveis aleatórias independentes e identicamente distribuídas com distribuição Normal, isto é, $X_i \sim N(\mu, \sigma) \quad i = 1, \ldots, n$, então

$$Y = \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2 \sim \chi^2_{(n)}.$$

Distribuição Qui-Quadrado: $X \sim \chi^2_{(n)}$

Observação

Seja $X \sim \chi^2_{(n)}$ e seja $x^2_{\alpha;n}$ tal que

$$P\left(X \leq x^2_{\alpha;n}\right) = \alpha \Leftrightarrow F\left(x^2_{\alpha;n}\right) = \alpha \Leftrightarrow x^2_{\alpha;n} = F^{-1}\left(\alpha\right)$$

então a $x^2_{\alpha;n}$ chama-se **quantil de probabilidade** α de uma distribuição Qui-Quadrado com n graus de liberdade.

Tabelas

A distribuição Qui-Quadrado encontra-se tabelada para diversos valores de n:

- **com recurso ao R**:
  - função densidade de probabilidade: **dchisq**(x, n)
  - função de distribuição: **pchisq**(x, n)
  - inversa da função de distribuição ou quantil de probabilidade: **qchisq**$(\text{probabilidade}, n)$

Exemplo 29

Seja $X \sim \chi^2_{(6)}$.

- Calcule $P\left(X \leq 16.8\right)$.

Exemplo 29

Seja $X \sim \chi^2_{(6)}$.

- Calcule $P\left(X \leq 16.8\right)$.

$$P\left(X \leq 16.8\right) = F(16.8) = 0.99$$

1. **tabela com recurso ao R**

   $F(16.8) = pchisq(16.8, 6) = 0.99$

Exemplo 29

Seja $X \sim \chi^2_{(6)}$.

- Calcule $P\left(X \geq 7.84\right)$.

Exemplo 29

Seja $X \sim \chi^2_{(6)}$.

- Calcule $P\left(X \geq 7.84\right)$.

$$P\left(X \geq 7.84\right) = 1 - P(X < 7.84) \underset{\text{v.a. contínua}}{=} 1 - F(7.84) = 1 - 0.75 = 0.25$$

1. **tabela com recurso ao R**

   $F(7.84) = pchisq(7.84, 6) = 0.75$

Exemplo 29

Seja $X \sim \chi^2_{(6)}$.

- Calcule o quantil de probabilidade $x^2_{0.95;6}$.

Exemplo 29

Seja $X \sim \chi^2_{(6)}$.

- Calcule o quantil de probabilidade $x^2_{0.95;6}$.

$$x^2_{0.95;6} = F^{-1}(0.95) = 12.6$$

1. **tabela com recurso ao R**

   $x^2_{0.95;6} = qchisq(0.95, 6) = 12.6$

Exemplo 29

Seja $X \sim \chi^2_{(6)}$.

- Calcule o quantil de probabilidade $x^2_{0.01;6}$.

Exemplo 29

Seja $X \sim \chi^2_{(6)}$.

- Calcule o quantil de probabilidade $x^2_{0.01;6}$.

$$x^2_{0.01;6} = F^{-1}(0.01) = 0.872$$

1. **tabela com recurso ao R**

   $x^2_{0.01;6} = qchisq(0.01, 6) = 0.872$

Distribuições Teóricas Contínuas:

- Distribuição Qui-Quadrado;
- **Distribuição t de Student**;
- Distribuição F de Snedecor.

Distribuição t de Student

Definição

Diz-se que uma variável aleatória contínua X tem **distribuição t de Student** com n graus de liberdade,

$$X \sim t_{(n)}$$

se a sua função densidade de probabilidade é dada por

$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}} = \frac{\int_0^{+\infty} e^{-x}x^{\frac{n+1}{2}-1}dx}{\sqrt{n\pi}\int_0^{+\infty} e^{-x}x^{\frac{n}{2}-1}dx}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}},$$

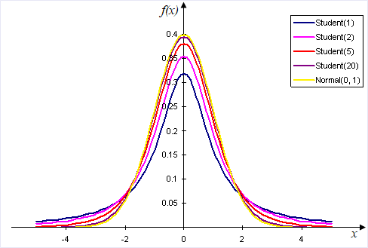
com $n > 0, \quad x \in \mathbb{R}$

e a sua função de distribuição é dada por

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}} dt.$$

Distribuição t de Student: $X \sim t_{(n)}$

Graficamente a sua função densidade de probabilidade, $f(x)$, é dada por:



- Domínio: $\mathbb{R}$
- A distribuição é simétrica em relação ao eixo dos yy.
- Devido à simetria: $F(-x) = 1 - F(x)$
- $X \sim t_{(n)} \underset{n \to +\infty}{\longrightarrow} X \dot{\sim} N(0,1)$

Propriedade

Se X e Y são variáveis aleatórias independentes, tais que $X \sim N(\mu, \sigma)$ e $Y \sim \chi^2_{(n)}$, então

$$T = \frac{X}{\sqrt{\frac{Y}{n}}} \sim t_{(n)}.$$

Distribuição t de Student: $X \sim t_{(n)}$

Observação

1. Seja $X \sim t_{(n)}$ e seja $t_{\alpha;n}$ tal que

$$P\left(X \leq t_{\alpha;n}\right) = \alpha \Leftrightarrow F\left(t_{\alpha;n}\right) = \alpha \Leftrightarrow t_{\alpha;n} = F^{-1}\left(\alpha\right)$$

então a $t_{\alpha;n}$ chama-se **quantil de probabilidade** α de uma distribuição t de Student com n graus de liberdade.

2. Os quantis de probabilidade são simétricos: $t_{\alpha;n} = -t_{1-\alpha;n}$

Tabelas

A distribuição t de Student encontra-se tabelada para diversos valores de n:

- **com recurso ao R**:
  - função densidade de probabilidade: **dt**(x, n)
  - função de distribuição: **pt**(x, n)
  - inversa da função de distribuição ou quantil de probabilidade: **qt**$(\text{probabilidade}, n)$

Exemplo 30

Seja $X \sim t_{(17)}$.

- Calcule $P\left(X \leq 0.689\right)$.

Exemplo 30

Seja $X \sim t_{(17)}$.

- Calcule $P\left(X \leq 0.689\right)$.

$$P\left(X \leq 0.689\right) = F(0.689) = 0.75$$

1. **tabela com recurso ao R**

   $F(0.689) = pt(0.689, 17) = 0.75$

Exemplo 30

Seja $X \sim t_{(17)}$.

- Calcule $P\left(X \geq -2.57\right)$.

Exemplo 30

Seja $X \sim t_{(17)}$.

- Calcule $P\left(X \geq -2.57\right)$.

$$P\left(X \geq -2.57\right) = 1 - P\left(X < -2.57\right) \underset{\text{v.a. contínua}}{=} 1 - F(-2.57) = 1 - 0.01 = 0.99$$

1. **tabela com recurso ao R**

   $F(-2.57) = pt(-2.57, 17) = 0.01$

Exemplo 30

Seja $X \sim t_{(17)}$.

- Calcule o quantil de probabilidade $t_{0.95;17}$.

Exemplo 30

Seja $X \sim t_{(17)}$.

- Calcule o quantil de probabilidade $t_{0.95;17}$.

$$t_{0.95;17} = F^{-1}(0.95) = 1.74$$

1. **tabela com recurso ao R**

   $t_{0.95;17} = qt(0.95, 17) = 1.74$

Exemplo 30

Seja $X \sim t_{(17)}$.

- Calcule o quantil de probabilidade $t_{0.01;17}$.

Exemplo 30

Seja $X \sim t_{(17)}$.

- Calcule o quantil de probabilidade $t_{0.01;17}$.

$$t_{0.01;17} = F^{-1}(0.01) = -2.57$$

1. **tabela com recurso ao R**

   $t_{0.01;17} = qt(0.01, 17) = -2.57$

Distribuições Teóricas Contínuas:

- Distribuição Qui-Quadrado;
- Distribuição t de Student;
- **Distribuição F de Snedecor**.

Distribuição F de Snedecor

Definição

Diz-se que uma variável aleatória contínua X tem **distribuição F de Snedecor** com m e n graus de liberdade,

$$X \sim F_{(m,n)}$$

se a sua função densidade de probabilidade é dada por

$$f(x) = \frac{m\left(\frac{m}{n}x\right)^{\frac{m}{2}-1}}{nB\left(\frac{m}{2},\frac{n}{2}\right)\left(1+\frac{m}{n}x\right)^{\frac{m+n}{2}}} = \frac{m\left(\frac{m}{n}x\right)^{\frac{m}{2}-1}}{n\int_0^{+\infty}\frac{x^{\frac{m}{2}-1}}{(1+x)^{\frac{m}{2}+\frac{n}{2}}}dx\left(1+\frac{m}{n}x\right)^{\frac{m+n}{2}}} \quad ,$$

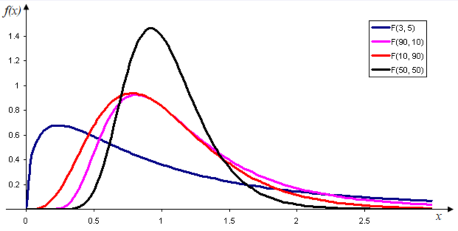
com $m > 0, \quad n > 0, \quad x > 0$

e a sua função de distribuição é dada por

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} \frac{m\left(\frac{m}{n}t\right)^{\frac{m}{2}-1}}{nB\left(\frac{m}{2},\frac{n}{2}\right)\left(1+\frac{m}{n}t\right)^{\frac{m+n}{2}}}dt.$$

Distribuição F de Snedecor: $X \sim F_{(m,n)}$

Graficamente a sua função densidade de probabilidade, $f(x)$ é dada por:



- Domínio: $\mathbb{R}^+$
- Distribuição assimétrica com cauda longa do lado direito.
- quanto maiores m e n (graus de liberdade) menos assimétrica é a distribuição.

Propriedade

Se X e Y são variáveis aleatórias independentes, tais que $X \sim \chi^2_{(m)}$ e $Y \sim \chi^2_{(n)}$, então

$$\mathrm{F} = \frac{\frac{X}{m}}{\frac{Y}{n}} \sim F_{(m,n)}.$$

Distribuição F de Snedecor: $X \sim F_{(m,n)}$

Observação

1. Seja $X \sim F_{(m,n)}$ e seja $f_{\alpha;m,n}$ tal que

$$P\left(X \leq f_{\alpha;m,n}\right) = \alpha \Leftrightarrow F\left(f_{\alpha;m,n}\right) = \alpha \Leftrightarrow f_{\alpha;m,n} = F^{-1}\left(\alpha\right)$$

então a $f_{\alpha;m,n}$ chama-se **quantil de probabilidade** α da distribuição F de Snedecor com m e n graus de liberdade.

Tabelas

A distribuição F de Snedecor encontra-se tabelada para diversos valores de m e n:

- **com recurso ao R**:
  - função densidade de probabilidade: **df**(x, m, n)
  - função de distribuição: **pf**(x, m, n)
  - inversa da função de distribuição ou quantil de probabilidade: **qf**(probabilidade, m, n)

Exemplo 31

Seja $X \sim F_{(7,4)}$.

- Calcule $P\left(X \leq 15\right)$.

Exemplo 31

Seja $X \sim F_{(7,4)}$.

- Calcule $P\left(X \leq 15\right)$.

$$P\left(X \leq 15\right) = F(15) = 0.99$$

1. **tabela com recurso ao R**

   $F(15) = pf(15, 7, 4) = 0.99$

Exemplo 31

Seja $X \sim F_{(7,4)}$.

- Calcule $P(X \geq 9.07)$.

Exemplo 31

Seja $X \sim F_{(7,4)}$.

- Calcule $P\left(X \geq 9.07\right)$.

$$P\left(X \geq 9.07\right) = 1 - P\left(X < 9.07\right) \underset{\text{v.a. contínua}}{=} 1 - F(9.07) = 1 - 0.975 = 0.025$$

1. **tabela com recurso ao R**

   $F(9.07) = pf(9.07, 7, 4) = 0.975$

Exemplo 31

Seja $X \sim F_{(7,4)}$.

- Calcule o quantil de probabilidade $f_{0.95;7,4}$.

Exemplo 31

Seja $X \sim F_{(7,4)}$.

- Calcule o quantil de probabilidade $f_{0.95;7,4}$.

$$f_{0.95;7,4} = F^{-1}(0.95) = 6.09$$

1. **tabela com recurso ao R**

   $f_{0.95;7,4} = qf(0.95, 7, 4) = 6.09$

Exemplo 31

Seja $X \sim F_{(7,4)}$.

- Calcule o quantil de probabilidade $f_{0.01;7,4}$.

Exemplo 31

Seja $X \sim F_{(7,4)}$.

- Calcule o quantil de probabilidade $f_{0.01;7,4}$.

$$f_{0.01;7,4} = F^{-1}(0.01) = 0.13$$

1. **tabela com recurso ao R**

   $f_{7,4;0.01} = qf(0.01, 7, 4) = 0.13$