LCE0216
Introdução à Bioestatística Florestal

4. Variáveis Aleatórias

Profa. Dra. Clarice Garcia Borges Demétrio

Monitores: Eduardo E. R. Junior & Giovana Fumes

Escola Superior de Agricultura "Luiz de Queiroz"
Universidade de São Paulo

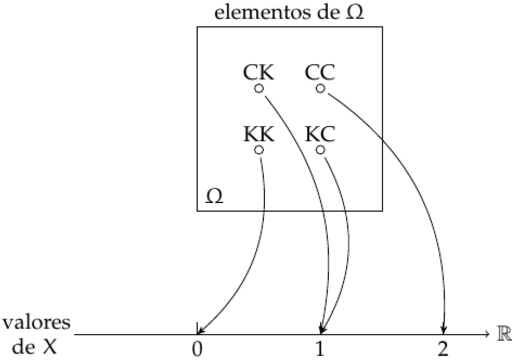
Piracicaba, 05 de abril de 2018

Variáveis aleatórias

Definição

Uma variável aleatória (v.a) é uma função que associa, a cada ponto pertencente a um espaço amostral (Ω), um único número real.

$$\text{v.a } X: \quad \Omega \rightarrow \mathbb{R}$$



Exemplo

Consideremos o experimento lançamento de duas moedas não viciadas e a observação das faces voltadas para cima.
O espaço amostral associado a esse experimento é dado por:

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

Seja, por exemplo, X o número de caras e Y o número de coroas

Possíveis resultados	x	y
(Cara, Cara)		
(Cara, Coroa)		
(Coroa, Cara)		
(Coroa, Coroa)		

Exemplo

Consideremos o experimento lançamento de duas moedas não viciadas e a observação das faces voltadas para cima.
O espaço amostral associado a esse experimento é dado por:

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

Seja, por exemplo, X o número de caras e Y o número de coroas

Possíveis resultados	x	y
(Cara, Cara)	2	0
(Cara, Coroa)	1	1
(Coroa, Cara)	1	1
(Coroa, Coroa)	0	2

Definições

$$\text{Variáveis aleatórias} \begin{cases} \text{discretas} \\ \text{contínuas} \end{cases}$$

Variável aleatória discreta

Uma quantidade X, associada a cada possível resultado do espaço amostral, é denominada **variável aleatória discreta**, se assume valores em um conjunto enumerável, com certa probabilidade.

Variável aleatória contínua

Uma quantidade X, associada a cada possível resultado do espaço amostral, é denominada **variável aleatória contínua**, se seu conjunto de valores é qualquer intervalo dos números reais, o que seria um conjunto não enumerável.

Um problema

Em 24 de março de 2013 foi divulgado o resultado de uma experiência realizada por um grupo de meninas da nona série da Hjallerup School, as quais concluíram que a proximidade dos roteadores WiFi prejudica o desenvolvimento de plantas (leia a matéria)

Suposição

Desejamos verificar a validade do estudo de tais meninas e, para tanto, iremos realizar um experimento com plantas de feijão.

- São necessários para um certo ensaio, 20 copos com ao menos uma muda;
- Restrição: 40 copos disponíveis e apenas 120 sementes;
- Suposição: porcentagem de germinação da semente do feijão, em condições iguais às do ensaio, é de 30%.

Ideia: formar os copos com ao menos uma muda para verificar se a proximidade do roteador prejudica o desenvolvimento da planta.

Quantos feijões por copo devemos plantar para a obtenção dos 20 copos com ao menos uma muda?

A) Se forem utilizados 3 feijões por copo...

- Qual é a porcentagem esperada de copos com pelo menos um feijão germinado? Com três feijões germinados? Com nenhum feijão germinado?
- Qual é o número médio de feijões germinados por copo?
- Dê uma ideia da variação esperada do número de feijões germinados.
- Qual é o número médio de copos com ao menos um feijão germinado?

B) Será que não seria melhor utilizar quatro feijões por copo e apenas 30 copos? Nesse caso,

- Qual é a porcentagem esperada de copos com pelo menos um feijão germinado? Com três feijões germinados? Com nenhum feijão germinado?
- Qual é o número médio esperado de feijões germinados por copo?
- Dê uma ideia de variação esperada do número de feijões germinados.
- Qual é o número médio de copos com ao menos um feijão germinado?

Análise da situação A

Seja G o evento germinar e $\bar{G}$ o evento não germinar.

(a) Construir o espaço amostral associado a esse experimento.

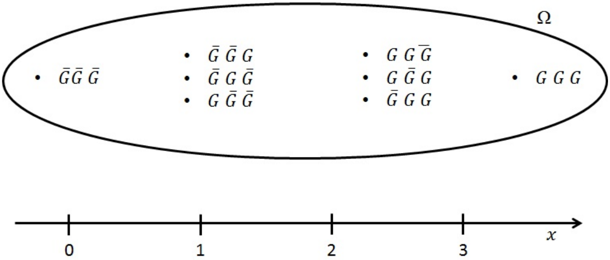
(b) Calcular as probabilidades associadas a cada um dos elementos do espaço amostral.

(c) Considerar X a variável número de feijões germinados e associar um valor x a cada um dos elementos do espaço amostral.

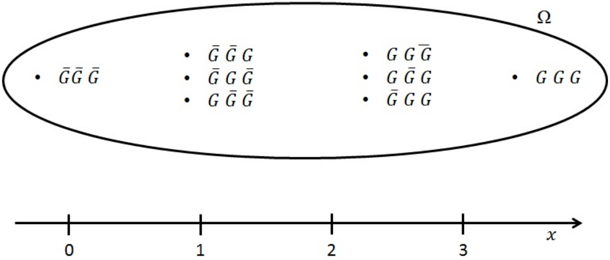
(d) Considerar Y a variável que associa o valor 0 ao resultado em que não há nenhum feijão germinado e o valor 1 aos resultados em que há pelo menos um feijão germinado. Associar um valor y a cada um do elementos do espaço amostral.

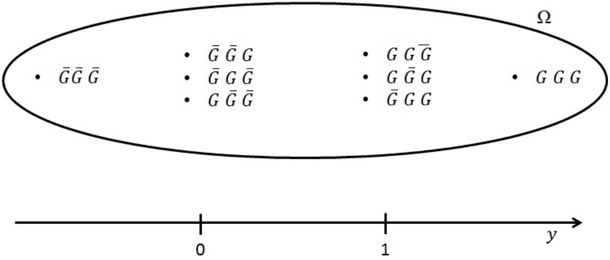
Possíveis Resultados	Probabilidades	x	y
$\bar{G}\bar{G}\bar{G}$	0,343	0	0
$\bar{G}\bar{G}G$	0,147	1	1
$\bar{G}G\bar{G}$	0,147	1	1
$G\bar{G}\bar{G}$	0,147	1	1
$\bar{G}GG$	0,063	2	1
$G\bar{G}G$	0,063	2	1
$GG\bar{G}$	0,063	2	1
GGG	0,027	3	1
Total	1,000		

Variáveis aleatórias discretas



Variáveis aleatórias discretas





Variáveis aleatórias discretas

Distribuição de uma variável aleatória discreta

Damos o nome de **distribuição de probabilidade** (modelo probabilístico) da variável aleatória discreta X, ao conjunto de pontos $(x_i, P(x_i))$, em que x_i representa os diferentes valores da variável aleatória e $P(x_i)$ a probabilidade de ocorrência de x_i, satisfazendo:

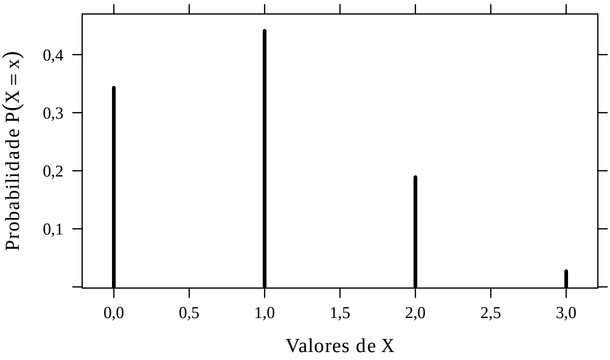
$$P(x_i) \geq 0 \quad \text{e} \quad \sum_i P(x_i) = 1.$$

- Costuma-se adotar, também, a notação $P(X = x_i)$ para designar a probabilidade da variável aleatória X assumir o valor x_i.

Variáveis aleatórias discretas

No exemplo com feijões...

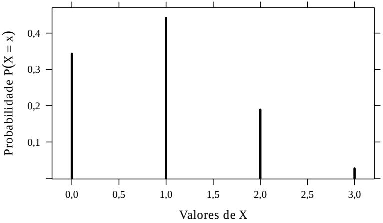
x	$P(X = x) = P(x)$
0	0,343
1	0,441
2	0,189
3	0,027
Total	1,000



Variáveis aleatórias discretas

No exemplo com feijões ...

x	$P(X = x) = P(x)$
0	0,343
1	0,441
2	0,189
3	0,027
Total	1,000

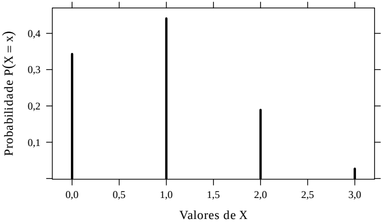


Qual é a porcentagem esperada de copos

- com três feijões germinados?
- com nenhum feijão germinado?
- com pelo menos um feijão germinado?

No exemplo com feijões ...

x	$P(X = x) = P(x)$
0	0,343
1	0,441
2	0,189
3	0,027
Total	1,000



Qual é a porcentagem esperada de copos

- com três feijões germinados?

  2,7%

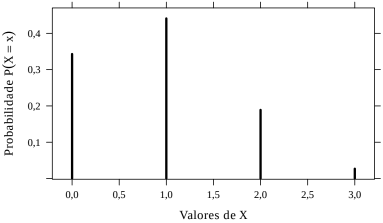
- com nenhum feijão germinado?

- com pelo menos um feijão germinado?

Variáveis aleatórias discretas

No exemplo com feijões ...

x	$P(X = x) = P(x)$
0	0,343
1	0,441
2	0,189
3	0,027
Total	1,000



Qual é a porcentagem esperada de copos

- com três feijões germinados?

  2,7%

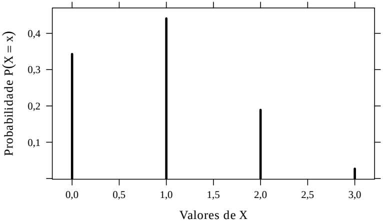
- com nenhum feijão germinado?

  34,3%

- com pelo menos um feijão germinado?

No exemplo com feijões ...

x	$P(X = x) = P(x)$
0	0,343
1	0,441
2	0,189
3	0,027
Total	1,000



Qual é a porcentagem esperada de copos

- com três feijões germinados?

  2,7%

- com nenhum feijão germinado?

  34,3%

- com pelo menos um feijão germinado?

  65,7%

Exercício: Obter a distribuição da variável aleatória Y e um gráfico que a represente.

Variáveis aleatórias discretas

Função de probabilidades

A função que fornece as probabilidades de ocorrências dos valores que a variável aleatória pode assumir é chamada **função de probabilidades** (f.p.)

Variáveis aleatórias discretas

Função de probabilidades

A função que fornece as probabilidades de ocorrências dos valores que a variável aleatória pode assumir é chamada **função de probabilidades** (f.p.)

Exemplo: A função de probabilidades da variável X = número de feijões germinados, é dada por:

$$P(x) = \binom{3}{x} 0,3^x 0,7^{(3-x)}, \quad \text{para} \quad x = 0, 1, 2, 3.$$

Função de probabilidades

A função que fornece as probabilidades de ocorrências dos valores que a variável aleatória pode assumir é chamada **função de probabilidades** (f.p.)

Exemplo: A função de probabilidades da variável X = número de feijões germinados, é dada por:

$$P(x) = \binom{3}{x} 0,3^x 0,7^{(3-x)}, \quad \text{para} \quad x = 0, 1, 2, 3.$$

Calcular $P(0), P(1), P(2)$ e $P(3)$ por meio da função de probabilidades.

Mas ...

Qual é o número médio esperado de feijões germinados por copo?

Valor médio ou esperança matemática de X

Dada a variável aleatória X, assumindo os valores $x_1, x_2, \ldots$ com as respectivas probabilidades $P(x_1), P(x_2), \ldots$, chamamos **valor médio** ou **esperança matemática de** X ao valor:

$$\mu_X = \mathrm{E}(X) = \sum_i x_i P(x_i).$$

Variáveis aleatórias discretas

No exemplo com feijões ...
Calcular o valor médio ou esperança da variável aleatória X.

x	$P(X = x) = P(x)$	$xP(x)$
0	0,343	
1	0,441	
2	0,189	
3	0,027	
Total	1,000	

Interpretação: Espera-se, na observação de um número grande de copos, obter um número médio de   feijões germinados por copo.

No exemplo com feijões ...
Calcular o valor médio ou esperança da variável aleatória X.

x	$P(X = x) = P(x)$	$xP(x)$
0	0,343	0
1	0,441	0,441
2	0,189	0,378
3	0,027	0,081
Total	1,000	0,9

Interpretação: Espera-se, na observação de um número grande de copos, obter um número médio de 0,9 feijões germinados por copo.

Exercício: Calcular o valor médio ou esperança da variável aleatória Y.

Valor médio ou esperança matemática de uma função de X

Dada uma variável aleatória X, assumindo os valores $x_1, x_2, \ldots,$, com as respectivas probabilidades $P(x_1), P(x_2), \ldots,$ chamamos **valor médio** ou **esperança matemática de uma função** $h(X)$ ao valor:

$$\mathrm{E}[h(X)] = \sum_i h(x_i) P(x_i).$$

Variáveis aleatórias discretas

Exercício: Calcular:

1. o valor médio ou esperança da função $3X$
2. o valor médio ou esperança da função X^2
3. o valor médio ou esperança da função $(X - 0,5)^2$
4. o valor médio ou esperança da função $(X - \mu_X)^2$
5. $\mathrm{E}[|X - \mu_X|]$

Observação: Sejam a e b duas constantes quaisquer e $h(X) = a + bX$, então

$$
\begin{aligned}
\mathrm{E}(a + bX) &= \sum_i (a + bx_i)P(x_i) \\
&= \sum_i \left[aP(x_i) + bx_iP(x_i)\right] \\
&= \sum_i aP(x_i) + \sum_i bx_iP(x_i) \\
&= a\sum_i P(x_i) + b\sum_i x_iP(x_i) \\
&= a + b\mathrm{E}(X)
\end{aligned}
$$

Exercício: Calcular $\mathrm{E}(30X)$, $\mathrm{E}(10 + X)$, $\mathrm{E}(1 - 2X)$ e $\mathrm{E}(X - \mu_X)$

Variância de X

Dada a variável aleatória X, chamamos de **variância** de X ao valor médio ou esperança da função $(X - \mu_X)^2$,

$$\sigma_X^2 = \text{Var}(X) = \text{E}\left[(X - \mu_X)^2\right].$$

Automaticamente ficam definidos o desvio padrão e o coeficiente de variação da variável aleatória X, dados respectivamente por:

$$\sigma_X = \sqrt{\sigma_X^2} \qquad \text{e} \qquad CV_X = 100 \times \frac{\sigma_X}{\mu_x}.$$

Exemplo: Calcular para as variáveis aleatórias X e Y a variância, o desvio padrão e o coeficiente de variação.

Observação:

$$
\begin{aligned}
\sigma_X^2 &= \mathrm{Var}(X) = \mathrm{E}\left[(X-\mu_x)^2\right] \\
&= \mathrm{E}(X^2 - 2X\mu_X + \mu_X^2) \\
&= \mathrm{E}(X^2) - 2\mathrm{E}(X)\mu_X + \mu_X^2 \\
&= \mathrm{E}(X^2) - 2\mu_X^2 + \mu_X^2 \\
&= \mathrm{E}(X^2) - \mu_X^2 \\
&= \mathrm{E}(X^2) - \Big[\mathrm{E}(X)\Big]^2
\end{aligned}
$$

Variância de X

Fórmula mais prática para o cálculo da variância de X:

$$\sigma_X^2 = \mathrm{Var}(X) = \mathrm{E}(X^2) - \Big[\mathrm{E}(X)\Big]^2.$$

Exemplo: Recalcular a variância de X utilizando a fórmula mais prática.

Exercício: Refazer todos os cálculos considerando 4 feijões por vaso e responder a sequência **B)** de questões iniciais.

Exemplo: Seja $F(X) = P(X \leq x)$.

x	$P(X = x) = P(x)$	$F(x) = P(X \leq x)$
0	0,343	
1	0,441	
2	0,189	
3	0,027	
Total	1,000	

Exemplo: Seja $F(X) = P(X \leq x)$.

x	$P(X = x) = P(x)$	$F(x) = P(X \leq x)$
0	0,343	0,343
1	0,441	0,784
2	0,189	0,973
3	0,027	1,000
Total	1,000	

Variáveis aleatórias discretas

Função de distribuição acumulada

Dada a variável aleatória X, chamaremos de **função de distribuição acumulada** ou simplesmente **função de distribuição** a função $F : \mathbb{R} \to [0,1]$ definida por:

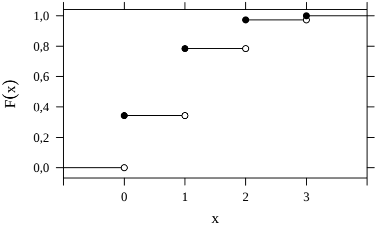
$$F(x) = P(X \leq x).$$

Exercício: Calcular para a variável aleatória X = número de feijões germinados,

- $F(-1) = P(X \leq -1)$
- $F(0) = P(X \leq 0)$
- $F(0,5) = P(X \leq 0,5)$
- $F(1) = P(X \leq 1)$
- $F(3) = P(X \leq 3)$
- $F(4) = P(X \leq 4)$

A função de distribuição acumulada da variável aleatória X = número de feijões germinados é dada a seguir, bem como o gráfico que a representa.

$$F(x) = \begin{cases} 0,000 & \text{para} \quad x < 0 \\ 0,343 & \text{para} \quad 0 \leq x < 1 \\ 0,784 & \text{para} \quad 1 \leq x < 2 \\ 0,973 & \text{para} \quad 2 \leq x < 3 \\ 1,000 & \text{para} \quad x \geq 3 \end{cases}$$



Exemplo: A distribuição de frequências da velocidade máxima diária do vento em 2014, é apresentada a seguir:

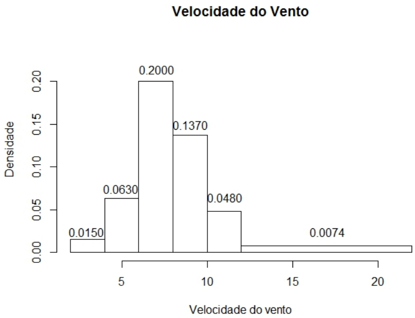


Tabela: Distribuição de frequências da velocidade máxima do vento

X_i	x_i^*	f_i	f_i'
2,00 ⊢ 4,00	3,00	11	0,0301
4,00 ⊢ 6,00	5,00	46	0,1260
6,00 ⊢ 8,00	7,00	146	0,4000
8,00 ⊢ 10,00	9,00	100	0,2740
10,00 ⊢ 12,00	11,00	35	0,0959
12,00 ⊢ 22,00	17,00	27	0,0740
Total		365	1,000

Exemplo: A distribuição de frequências da velocidade máxima diária do vento em 2014, é apresentada a seguir:

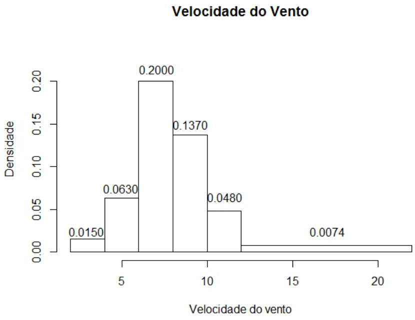
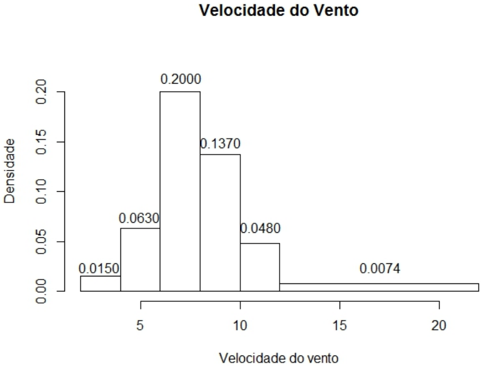


Tabela: Distribuição de frequências da velocidade máxima do vento

X_i	x_i^*	f_i	f_i'
$2,00 \vdash 4,00$	3,00	11	0,0301
$4,00 \vdash 6,00$	5,00	46	0,1260
$6,00 \vdash 8,00$	7,00	146	0,4000
$8,00 \vdash 10,00$	9,00	100	0,2740
$10,00 \vdash 12,00$	11,00	35	0,0959
$12,00 \vdash 22,00$	17,00	27	0,0740
Total		365	1,000

$$\text{Densidade} = \frac{\text{freq. rel.}}{\text{amplitude}}$$

Variáveis aleatórias contínuas



Dado o histograma acima, obter aproximadamente, a porcentagem de dias com velocidade máxima do vento avaliada

- entre 4 e 8 (m/s)
- entre 6 e 10 (m/s)
- entre 2 e 22 (m/s)

Variáveis aleatórias contínuas

Função densidade de probabilidade

Condições para que uma função seja uma função densidade de probabilidade:

(i) $f(x) \geq 0, \ \forall x \in D_f$

(ii) A área entre o gráfico da função f e o eixo x é igual a 1, ou seja

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

Função densidade de probabilidade

Condições para que uma função seja uma função densidade de probabilidade:

(i) $f(x) \geq 0, \; \forall x \in D_f$

(ii) A área entre o gráfico da função f e o eixo x é igual a 1, ou seja

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

Consequências ...

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

Se $a = b = c$, então $P(X = c) = 0$

Exemplo: Seja uma função $f(x)$ dada por:

$$f(x) = \begin{cases} 0 & \text{para} \quad x \leq 0 \\ ax^3 & \text{para} \quad 0 < x \leq 2 \\ 0 & \text{para} \quad x > 2 \end{cases}$$

em que a é uma constante.
Obter a de modo que $f(x)$ seja uma função densidade de probabilidade de uma variável aleatória contínua X.

Variáveis aleatórias contínuas

Exemplo: Seja uma função $f(x)$ dada por:

$$f(x) = \begin{cases} 0 & \text{para} \quad x \leq 0 \\ 0,2 - 0,02x & \text{para} \quad 0 < x \leq 10 \\ 0 & \text{para} \quad x > 10 \end{cases}$$

(a) Construir o gráfico dessa função;

(b) Verifique que $f(x)$ é uma função densidade de probabilidade;

(c) Calcular as porcentagens esperadas para
- X entre 5 e 10 unidades;
- X entre 3 e 5 unidades;
- X entre 0 e 2 unidades;
- X entre 0 e 10 unidades;
- X maior do que 10 unidades;

Variáveis aleatórias contínuas

Valor médio ou esperança matemática de X

$$\mu_X = \mathrm{E}(X) = \int_{-\infty}^{+\infty} x f(x) d(x)$$

Moda

$$Mo_X = \max f(x),\ x \in Df$$

Variáveis aleatórias contínuas

Valor médio ou esperança matemática de uma função $h(X)$

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

Variância de X

$$\begin{aligned}
\sigma_X^2 = \mathrm{Var}(X) &= \mathrm{E}\left[(X-\mu_X)^2\right] \\
&= \int_{-\infty}^{+\infty} (x-\mu_X)^2 f(x)dx \\
&= \ldots \\
&= \mathrm{E}(X^2) - \left[\mathrm{E}(X)\right]^2
\end{aligned}$$

Variáveis aleatórias contínuas

Função de distribuição acumulada

Dada a variável aleatória X, com função densidade de probabilidade $f(x)$, temos que a função de distribuição acumulada é dada por:

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

- Percentil:

$$P_{100p} \text{ é o valor de } t \text{ tal que } F(t) = p$$

- Caso particular: Mediana

$$Md_X = P_{50} \text{ é o valor de } t \text{ tal que } F(t) = 0,5.$$

Exemplo: Calcular, supondo o modelo teórico,

$$f(x) = \begin{cases} 0 & \text{para} \quad x \leq 0 \\ ax^3 & \text{para} \quad 0 < x \leq 2 \\ 0 & \text{para} \quad x > 2 \end{cases}$$

1. o valor médio de X (μ_X)
2. $\mathrm{E}(X^2)$
3. a variância e o desvio padrão de X.

Variáveis aleatórias contínuas

Exercício: Para a função $f(x)$, dada por:

$$f(x) = \begin{cases} 0 & \text{para} \quad x \leq 0 \\ 0,2 - 0,02x & \text{para} \quad 0 < x \leq 10 \\ 0 & \text{para} \quad x > 10 \end{cases}$$

Pede-se:

(a) Calcular μ_X

(b) Calcular σ_X^2.