Estatística Básica

Variáveis Aleatórias Contínuas

Renato Dourado Maia

Instituto de Ciências Agrárias

Universidade Federal de Minas Gerais

Lembrando...

- Uma quantidade X, associada a cada possível resultado do espaço amostral, é denominada de variável aleatória (VA) discreta, se assume valores num conjunto enumerável, com certa probabilidade. Por outro lado, será denominada variável aleatória contínua, se seu conjunto de valores é qualquer intervalo dos números reais, o que seria um conjunto não enumerável.

Exemplo 1

- Estudos anteriores revelam a existência de um grande lençol de água no subsolo de uma região. No entanto, sua profundidade ainda não foi determinada, sabendo-se apenas que o lençol pode estar situado em qualquer ponto entre 20 e 100 metros.

- Vamos supor que escolhemos, ao acaso, um ponto nessa região, e dispomos de uma sonda que, ao fazer a perfuração, detecta com precisão a profundidade do reservatório de água. X será a VA que representará a profundidade.

Exemplo 1

- Notemos que, apesar de X poder ser qualquer número entre 20 e 100, o instrumento com o qual trabalhamos pode não ser tão preciso quanto gostaríamos. Por exemplo, uma profundidade de 32,571 metros poderia ser medida por 32,6 metros. Vamos assumir, entretanto, que temos um instrumento ideal que não faz aproximações. Nessas condições, podemos supor a sonda acoplada a um instrumento indicador da profundidade e um dispositivo que, quando a sonda encontrar água, interrompa imediatamente a perfuração.

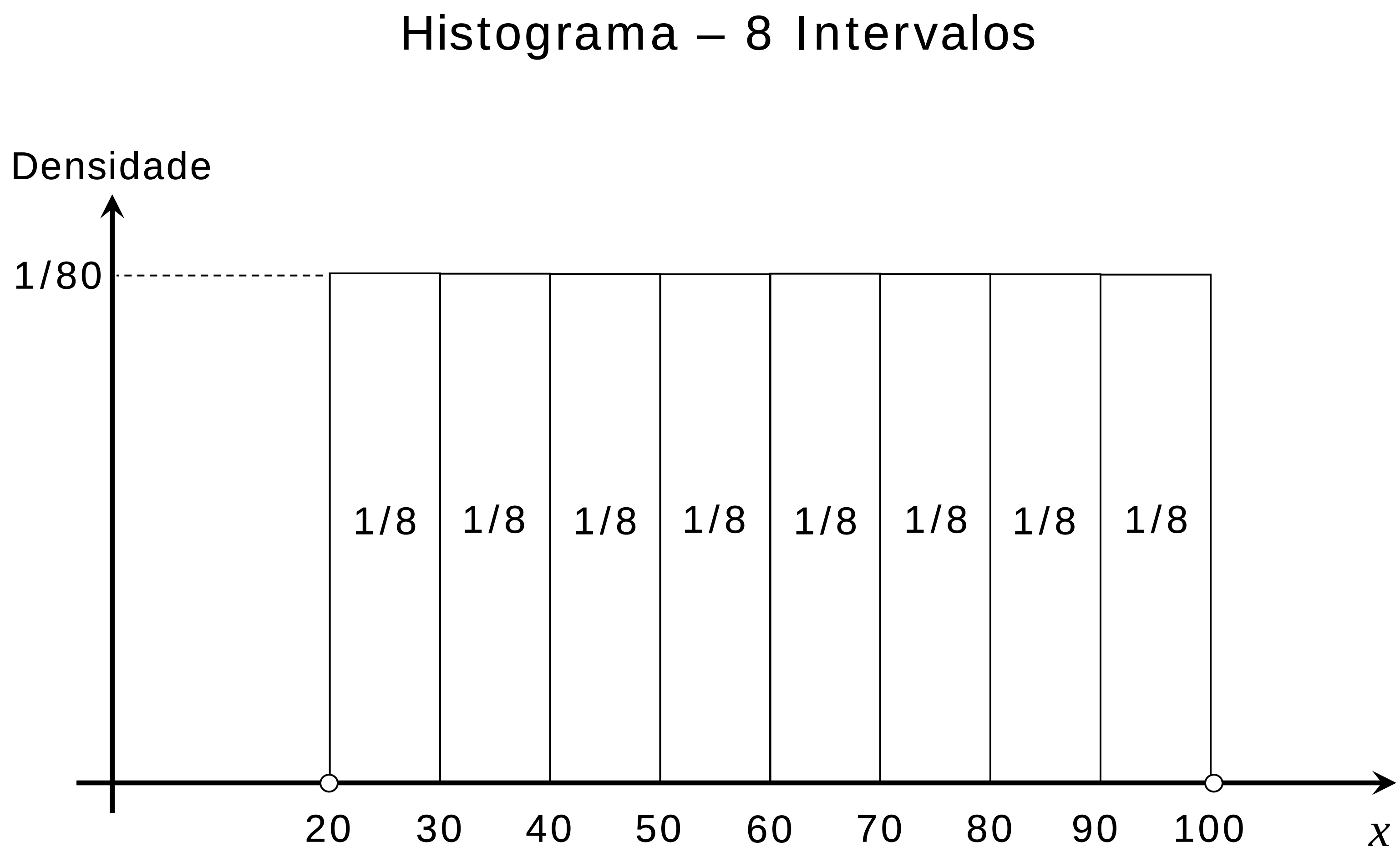
Exemplo 1

- Uma vez que não temos informações adicionais a respeito da profundidade do lençol, é razoável assumirmos que a sonda pode parar em qualquer ponto entre 20 e 100 metros, sem que tenhamos motivos para privilegiar essa ou aquela profundidade. Assim, consideraremos todos os pontos como igualmente prováveis. Se utilizarmos a mesma ideia de atribuir a cada possível ponto uma probabilidade, teremos uma dificuldade extra... Vocês percebem o porquê disso?

Exemplo 1

- A solução é considerar intervalos de valores na atribuição de probabilidades. Neste exemplo, sabemos que o espaço amostral corresponde ao intervalo [20 100] e que as profundidades são igualmente prováveis. Suponha, por um momento, que dividimos o espaço amostral em 8 intervalos de comprimento 10 ([20, 30), [30, 40), ..., [90, 100)). Logo, é razoável atribuir aos intervalos a probabilidade 1/8.

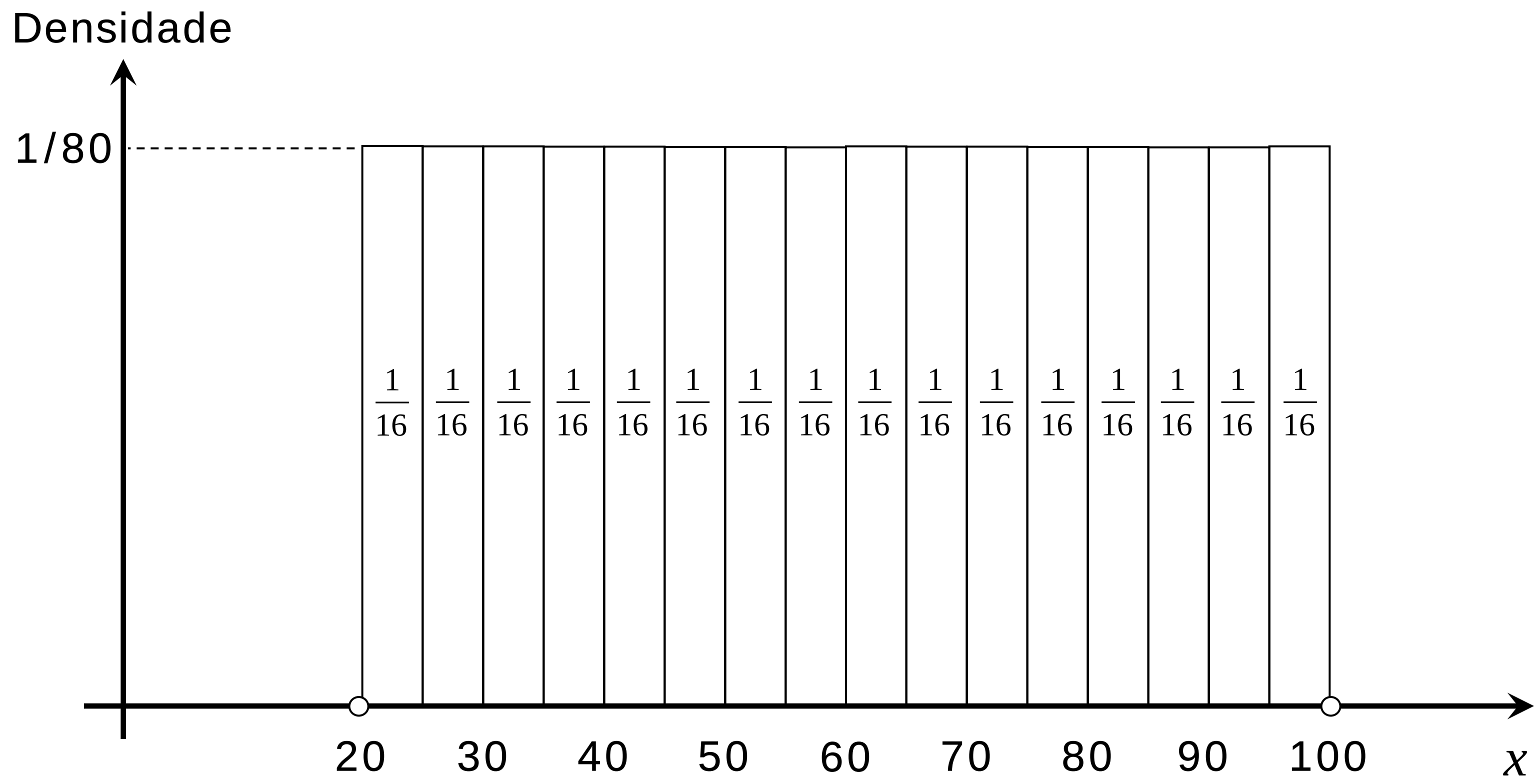
Exemplo 1



Histograma – 8 Intervalos
Densidade
1/80
1/8
1/8
1/8
1/8
1/8
1/8
1/8
1/8
20
30
40
50
60
70
80
90
100
x

Exemplo 1

Histograma – 16 Intervalos

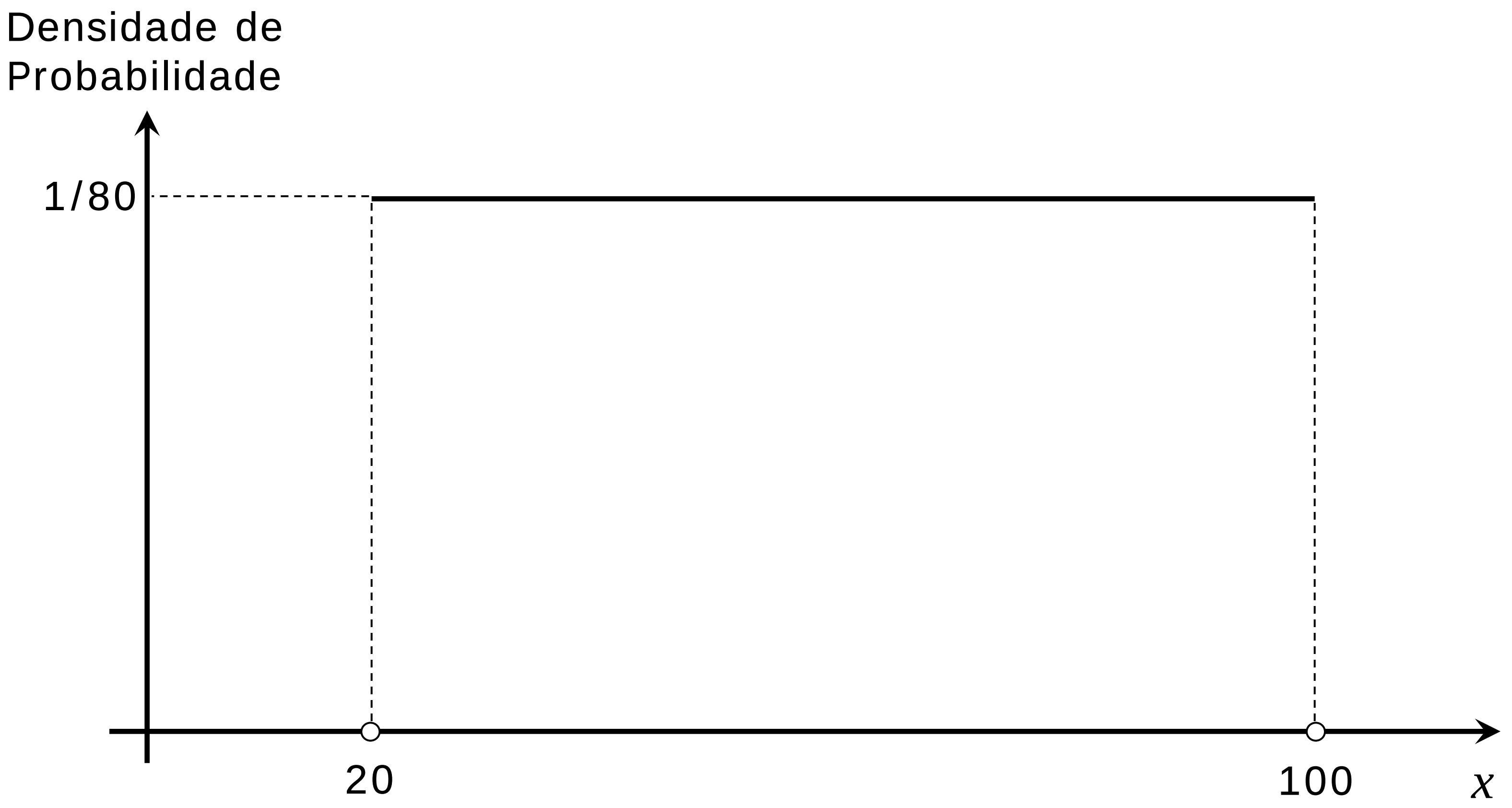


Exemplo 1

- Os histogramas dos *slides* 7 e 8, apesar de possuírem uma quantidade diferente de intervalos, apresentam o mesmo valor de densidade, 1/80. O procedimento de aumento de intervalos pode continuar de tal forma que, numa situação teórica, com infinitos intervalos, seria obtido um histograma com qual aspecto?

Exemplo 1

Histograma – "Infinitos" Intervalos



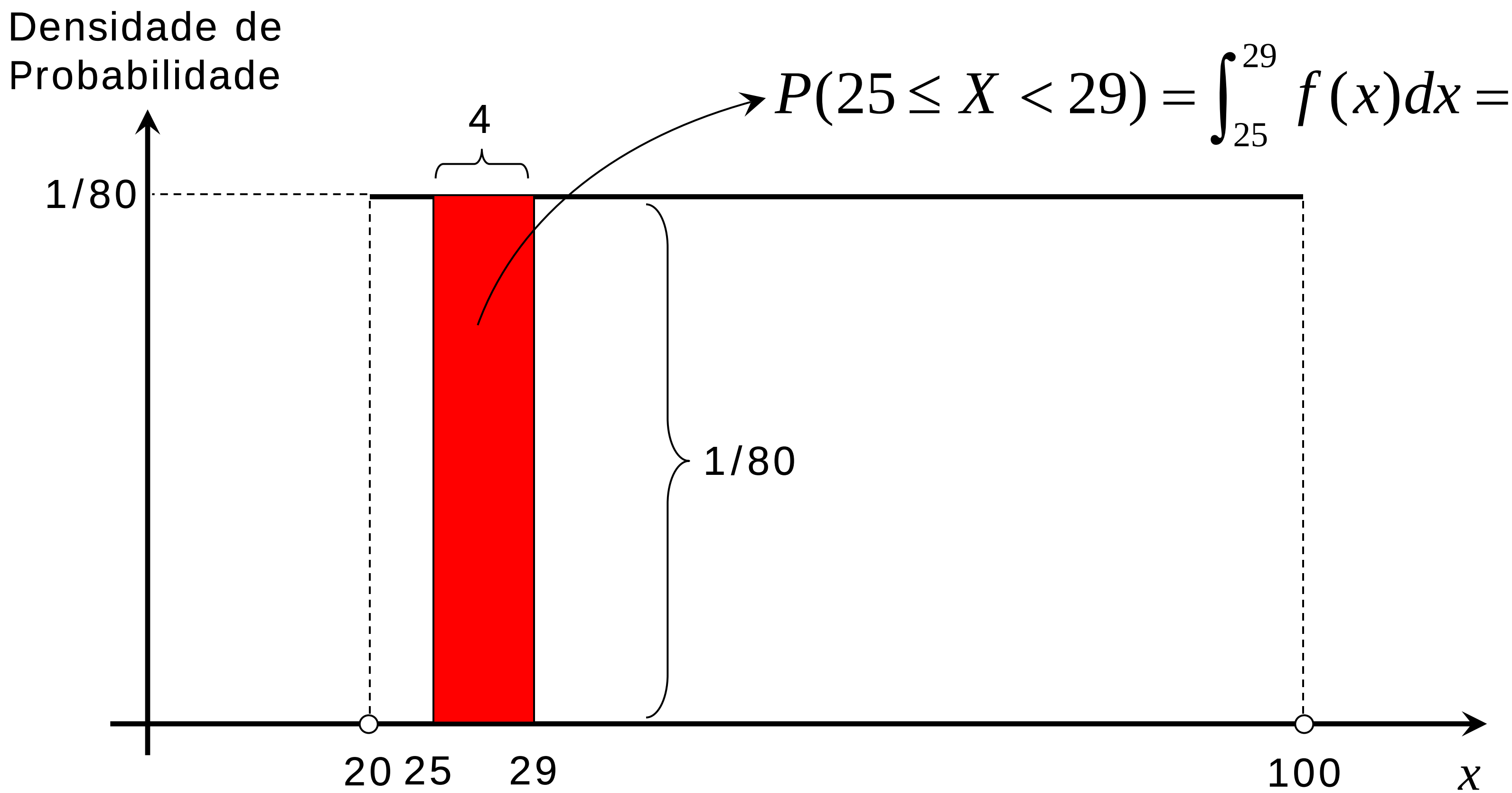
Exemplo 1

- Estamos agora em condições de caracterizar, completamente, a atribuição de probabilidades para o caso contínuo: ela será definida pela área abaixo de uma função, denominada densidade de probabilidade.

  - Observe que a densidade em si não é uma probabilidade, mas uma função matemática que nos auxilia na atribuição de probabilidades. Para a variável aleatória X, que representa a profundidade do lençol de água:

$$f(x) = \begin{cases} \dfrac{1}{80}, & para\ 20 \le x \le 100; \\ 0, & para\ x < 20\ ou\ x > 100. \end{cases}$$

Exemplo 1



$$P(25 \le X < 29) = \int_{25}^{29} f(x)dx = 4/80$$

Função Densidade de Probabilidade

- Dizemos que $f(x)$ é uma função contínua de probabilidade, ou função densidade de probabilidade, para uma variável aleatória contínua X, se satisfaz duas condições:

$$i)\ f(x) \geq 0,\ para\ todo\ x\ \in\ (-\infty,\ \infty).$$

$$ii)\ \ A\ área\ definida\ por\ f(x)\ é\ igual\ a\ 1,\ \text{ou seja,}\ \int_{-\infty}^{\infty} f(x)dx = 1.$$

Atribuição de Probabilidades:

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

$$P(X = k) = 0$$

Exemplo 2

- Arqueólogos estudaram certa região e estabeleceram um modelo teórico para a variável C, comprimento de fósseis da região (em cm). Suponha que C é uma variável aleatória contínua com a seguinte função densidade de probabilidade:

$$f(c) = \begin{cases} \dfrac{1}{40}\left(\dfrac{c}{10}+1\right), & \textit{para } 0 \le c \le 20; \\ 0, & \textit{caso contrário.} \end{cases}$$

a) Comprove que a função é efetivamente uma densidade.

b) Qual é a probabilidade de um fóssil, escolhido ao acaso nessa região, apresentar comprimento inferior a 8 cm?

Exemplo 2

$$\text{a) } \int_{-\infty}^{\infty} f(c)dc = \int_{0}^{20} \frac{1}{40}\left(\frac{c}{10}+1\right)dc = \left(\frac{c^2}{800}+\frac{c}{40}\right)\Bigg|_{0}^{20}$$

$$= \frac{400}{800}+\frac{20}{40} = \frac{1}{2}+\frac{1}{2} = 1$$

$$\text{b) } P(C<8) = \int_{-\infty}^{8} f(c)dc = \int_{0}^{8} \frac{1}{40}\left(\frac{c}{10}+1\right)dc = \left(\frac{c^2}{800}+\frac{c}{40}\right)\Bigg|_{0}^{8}$$

$$= \frac{64}{800}+\frac{8}{40} = \frac{7}{25}$$

Exemplo 3

- Num teste educacional com crianças, o tempo para a realização de uma bateria de questões de raciocínio verbal e lógico é medido e anotado para ser comparado com um modelo teórico. Esse teste é utilizado para identificar o desenvolvimento das crianças e auxiliar na aplicação de medidas corretivas. O modelo teórico considera T, tempo de teste em minutos, como uma VA contínua com a seguinte função densidade de probabilidade:

Exemplo 3

$$f(t) = \begin{cases} \dfrac{1}{40}(t-4), & se\ 8 \le t < 10; \\ \dfrac{3}{20}, & se\ 10 \le t \le 15; \\ 0, & caso\ contrário. \end{cases}$$

a) Comprove que a função é efetivamente uma densidade.
b) Calcule $P(9 < T \le 12)$.

Exemplo 3

a) $$\int_{-\infty}^{\infty} f(t)dt = \int_{8}^{10} \frac{1}{40}(t-4)\,dt + \int_{10}^{15} \frac{3}{20}\,dt = \left(\frac{t^2}{80} - \frac{4t}{40}\right)\Bigg|_{8}^{10} + \frac{3t}{20}\Bigg|_{10}^{15}$$

$$= \frac{10}{40} + \frac{15}{20} = 1$$

b) $$P(9 < T \le 12) = \int_{9}^{12} f(t)dt = \int_{9}^{10} \frac{1}{40}(t-4)\,dt + \int_{10}^{12} \frac{3}{20}\,dt$$

$$= \left(\frac{t^2}{80} - \frac{4t}{40}\right)\Bigg|_{9}^{10} + \frac{3t}{20}\Bigg|_{10}^{12} = \frac{7}{16}$$

Medidas de Posição para VA Contínuas

- Seja X uma VA contínua, e $f(x)$ a sua função densidade de probabilidade.
  - A média, valor esperado ou esperança matemática de X, denotada por $E(X)$ ou , é dada pela expressão:

$$E(X) = \mu = \int_{-\infty}^{\infty} xf(x)dx.$$

  - A mediana é o valor Md que satisfaz às seguintes condições:

$$P(X \geq M_d) \geq 0{,}5 \; e \; P(X \leq M_d) \geq 0{,}5.$$

  - Para a moda (Mo):

$$f(M_o) = \max_x f(x)$$

Variância para VA Contínuas

- Para uma variável aleatória X, com função densidade de probabilidade $f(x)$, a variância é dada pela seguinte expressão:

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx.$$

- Assim com no caso discreto, a variância é a medida de dispersão mais utilizada, e também pode ser utilizada a expressão alternativa:

$$\sigma^2 = E(X^2) - \mu^2, \; com \; E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx$$

Variância para VA Contínuas

- O desvio padrão é a raiz quadrada positiva da variância e, como já mencionado anteriormente, tem a mesma unidade de medida da variável original, o que facilita a interpretação dos seus valores.

- As propriedades do valor esperado e da variância apresentadas para VA discretas permanecem válidas, e a verificação pode ser feita utilizando-se as propriedades da integral.

Exemplo 4

- Calcular o valor esperado, a variância, a moda e a mediana para a variável aleatória C do exemplo 2, cuja função densidade de probabilidade é dada pela seguinte expressão:

$$f(c) = \begin{cases} \dfrac{1}{40}\left(\dfrac{c}{10}+1\right), & para\ \ 0 \le c \le 20; \\ 0, & caso\ contrário. \end{cases}$$

Exemplo 4

$$E(C) = \mu = \int_{-\infty}^{\infty} cf(c)dc = \int_{0}^{20} \frac{1}{40}\left(\frac{c^2}{10} + c\right)dc = \left(\frac{c^3}{1200} + \frac{c^2}{80}\right)\Bigg|_0^{20}$$

$$= \frac{8000}{1200} + \frac{400}{80} = \frac{20}{3} + 5 = \frac{35}{3}$$

$$E(C^2) = \int_{-\infty}^{\infty} c^2 f(c)dc = \int_{0}^{20} \frac{1}{40}\left(\frac{c^3}{10} + c^2\right)dc = \left(\frac{c^4}{1600} + \frac{c^3}{120}\right)\Bigg|_0^{20} = \frac{500}{3}$$

$$\sigma_C^2 = E(C^2) - \mu^2 = \frac{500}{3} - \left(\frac{35}{3}\right)^2 = \frac{275}{9} = 30,56\ cm^2$$

$$\sigma_C = \sqrt{30,56} = 5,53\ cm$$

Exemplo 4

$$Mo = 20$$

$$P(C \geq Md) = 0,5$$

$$\int_{Md}^{\infty} f(c)dc = \int_{Md}^{20} \frac{1}{40}\left(\frac{c}{10}+1\right)dc = \left(\frac{c^2}{800}+\frac{c}{40}\right)\Bigg|_{Md}^{20} = 0,5$$

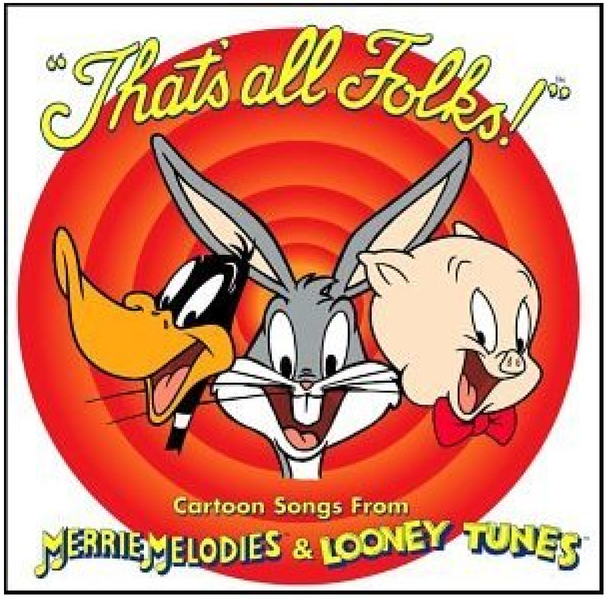
$$Md^2 + 20Md - 400 = 0$$

$$Md = 12,36$$

Ainda há Tempo?

Exercícios da seção 6.1!

Por hoje é só!



"That's all Folks!"
Cartoon Songs From
MERRIE MELODIES & LOONEY TUNES