Introdução à Bioestatística

Marcelo Goulart Correia

Instituto Nacional de Cardiologia

February 22, 2016

- Existem dois tipos de variáveis aleatórias
  - Variáveis aleatórias discretas
  - Variáveis aleatórias contínuas

- Variável aleatória discreta
  - Assume um número finito ou infinito enumerável de valores

$$\boxed{X : E \longrightarrow \mathbb{N}} \tag{1}$$

- Variável aleatória contínua
  - Assume um número infinito não-enumerável de valores

$$\boxed{X : E \longrightarrow \mathbb{R}} \tag{2}$$

- Variável aleatória discreta
  - Função de probabilidade (f)
  - $f(x_i)$ a probabilidade de que X assuma um determinado valor

$$\boxed{\begin{aligned} f : \mathbb{N} &\longrightarrow [0,1] \\ x_i &\longmapsto f(x_i) = P[X = x_i] \end{aligned}} \tag{3}$$

  - Se x_i não é um valor que X pode assumir –> $f(x_i) = 0$

- Exemplo
  - Probabilidade de tirar 3 caras em 3 lançamentos de moedas $f(3) = P[X = 3] = P[\{KKK\}] = 1/2 * 1/2 * 1/2 = 1/8$
  - Probabilidade de tirar 2 caras em 3 lançamentos de moedas $f(2) = P[X = 2] = P[\{KKC, KCK, CKK\}] = 1/8 + 1/8 + 1/8 = 3/8$
  - Probabilidade de tirar o número 6 em 3 lançamentos de moedas $f(N6) = P[X = N6] = P[\{N6\}] = 0$

- Variável aleatória discreta
  - Função de distribuição acumulada (F)
  - $F(x_i)$ a probabilidade de que X assuma um valor inferior ou igual a x_i

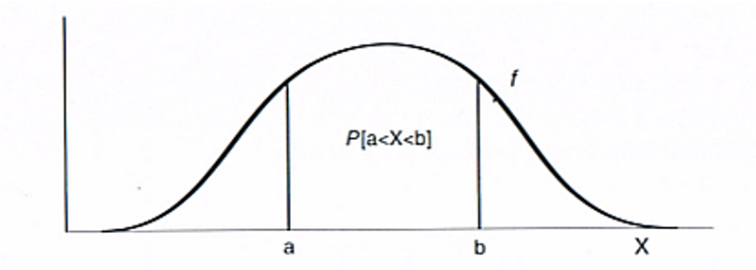
- Exemplo
  - Probabilidade de tirar 1 ou nenhuma cara em 3 lançamentos de moedas $F(1) = P[X \leq 1] = f(0) + f(1) = 1/8 + 3/8 = 4/8$
  - Probabilidade de tirar 2 ou menos caras em 3 lançamentos de moedas
    $F(2) = P[X \leq 2] = f(0)+f(1)+f(2) = 1/8+3/8+3/8 = 7/8$
  - Probabilidade de tirar um número menor que 5 em 3 lançamentos de moedas $F(N6) = P[X \leq N5] = 0$

- Variável aleatória contínua
  - Função de probabilidade (f)
  - $f(x)$ a faixa de probabilidade de ocorrer algum evento

$$\boxed{\begin{aligned} f(x) \geq 0 \\ \int_{-\infty}^{+\infty} f(x)dx = 1 \end{aligned}} \tag{4}$$

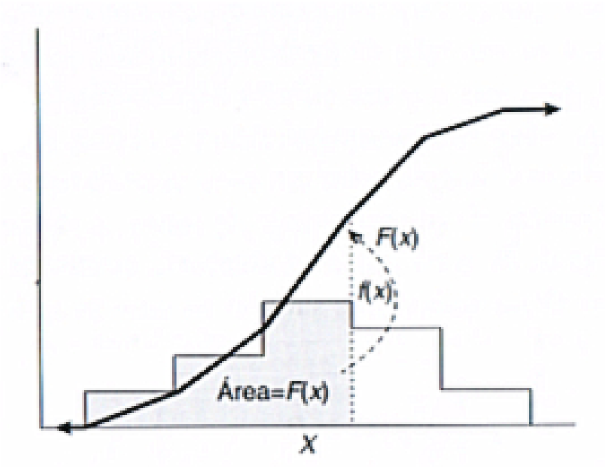
  - É possível também calcular uma área mais restrita de valores utilizando:

$$\boxed{P[a \leq X \leq b] = \int_{a}^{b} f(x)dx} \tag{5}$$


f
P[a<X<b]
a
b
X

- Variável aleatória contínua
  - Função de distribuição acumulada (F)
  - $F(x)$ a probabilidade de que X assuma um valor inferior ou igual a x

$$\boxed{\begin{aligned} F : \mathbb{R} &\longrightarrow [0,1] \\ x \longmapsto F(x) = P[X \leq x] &= \int_{-\infty}^{x} f(t)dt \end{aligned}} \tag{6}$$


F(x)
f(x)
Área=F(x)
X

- Medidas de tendência central e dispersão para variáveis aleatórias
    - Esperança (ou valor esperado)

$$\boxed{E[X] = \sum_{i \in \mathbb{I}} x_i f(x_i)} \tag{7}$$

$$\boxed{E[X] = \int_{-\infty}^{+\infty} x * f(x) dx} \tag{8}$$

- Medidas de tendência central e dispersão para variáveis aleatórias
  - Variância

$$\boxed{\mathrm{Var}[X] = E[(X - E[X])^2]} \tag{9}$$

$$\boxed{\mathrm{Var}[X] = \sum_{i \in \mathbb{I}} (x_i - E[X])^2 * f(x_i)} \tag{10}$$

$$\boxed{\mathrm{Var}[X] = \int_{-\infty}^{+\infty} (x - E[X])^2 * f(x) dx} \tag{11}$$

- Exercício

- Existem diversas distribuições de probabilidades:
    - Discretas
        - Bernoulli, **Binomial**, Geométrica, Binomial negativa, Hipergeométrica e Poisson
    - Contínuas
        - Uniforme, Exponencial, **Normal**, χ^2, t de Student, F de Snedcor

- Distribuição Binomial
  - Variáveis aleatórias binárias

$$\boxed{f(k) = P[X = k] = \binom{n}{k} p^k q^{(n-k)} \; \forall \; k = 0, 1, ..., n} \tag{12}$$

$$\boxed{E[X] = np} \tag{13}$$

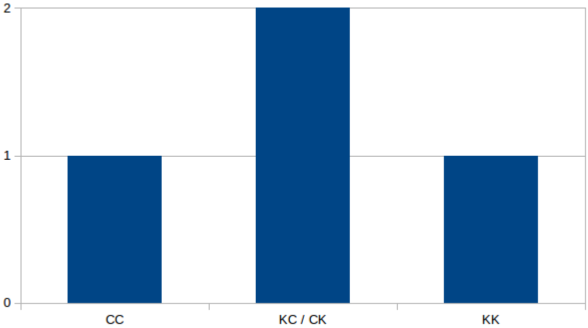
$$\boxed{\mathrm{Var}[X] = npq} \tag{14}$$

- n –> Total de amostras
- k –> Total de eventos ocorridos
- p –> Probabilidade de um evento ocorrer
- q –> Probabilidade de um evento não ocorrer

- Entendendo melhor...
  - Vamos lançar uma moeda "honesta" duas vezes e ver os possíveis.
  - "Cara" = C e "Coroa" = K

Segundo \Primeiro	C	K
C	CC	CK
K	KC	KK

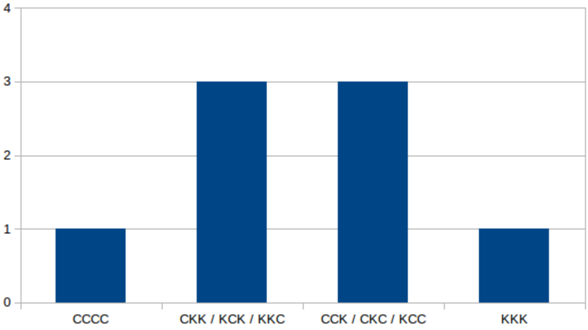
- Entendendo melhor...



- Entendendo melhor...
  - Repetindo o experimento agora lançando 3 moedas
  - "Cara" = C e "Coroa" = K

Terceiro \Anterior	CC	KC	CK	KK
C	CCC	KCC	CKC	KKC
K	CCK	KCK	CKK	KKK

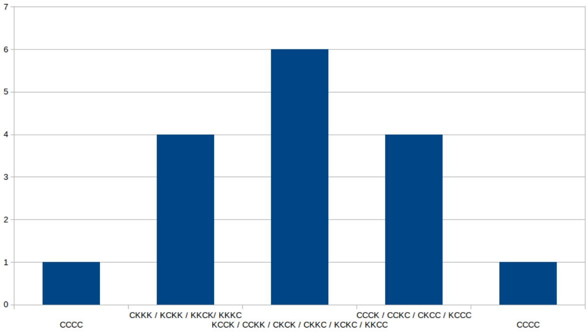
- Entendendo melhor...



- Entendendo melhor...
  - Repetindo o experimento agora lançando 4 moedas
  - "Cara" = C e "Coroa" = K

Quarto \Anterior	CCC	KCC	CKC	CCK	KKC	KCK	CKK	KKK
C	CCCC	KCCC	CKCC	CCKC	KKCC	KCKC	CKKC	KKKC
K	CCCK	KCCK	CKCK	CCKK	KKCK	KCKK	CKKK	KKKK

- Entendendo melhor...



- Entendendo melhor...
  - Resumindo:
    - Os valores de X são sempre inteiros
    - Os eventos são independentes
    - As probabilidades pressupõesm reposição
    - A ordem dos fatores não altera o produto :P

- Exemplo
  - A probabilidade de uma pessoa apresentar IC num teste de esforço é de 10%. Após analisar 10 pacientes, qual a probabilidade de que 5 deles apresentarem IC? Qual a esperança e a variância?

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

$$\begin{aligned} f(5) = P[X = 5] &= \binom{10}{5} 0,1^5 0,9^{(10-5)} \\ &= \frac{10!}{5!(10-5)!} 0,1^5 0,9^{(10-5)} \\ &= 0,0015 = 0,15\% \end{aligned}$$

$$E[X] = 10 * 0,1 = 1$$
$$\mathrm{Var}[X] = 10 * 0,1 * 0,9 = 0,9$$

- Exercício

- Distribuição Normal
  - Distribuição essencial para diversos testes e hipóteses trabalhadas na Estatística
  - Grande parte dos fenômenos estudados possuem esse comportamento

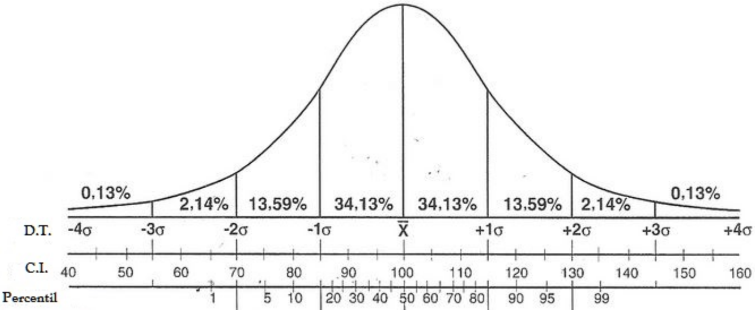
$$\boxed{f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}*(\frac{x-\mu}{\sigma})^2}, \forall x \in \mathbb{R}} \tag{15}$$

- σ –> Desvio-padrão dos dados
- μ –> Média dos dados
- x –> Faixa de valores que podem ocorrer

$$\boxed{E[X] = \mu} \tag{16}$$

$$\boxed{Var[X] = \sigma^2} \tag{17}$$

- Entendendo melhor...

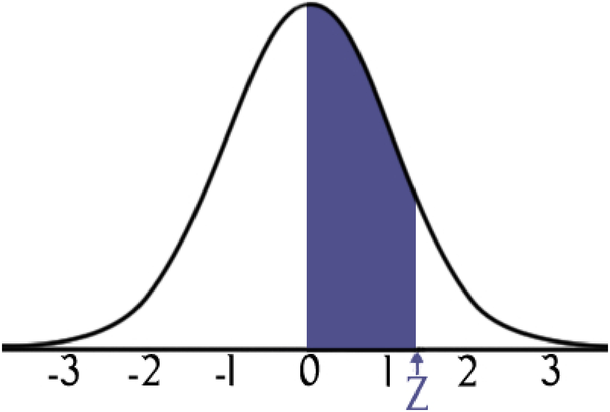


- Distribuição Normal
  - Assimetria e Curtose igual à zero
  - Forma de sino (Depende de μ e σ)
  - Outras distribuições podem ser aproximadas à Normal (Binomial, Poisson, ...)

- Estatística Z - Curva Normal Padronizada
  - Artifício utilizado para cálculo da área (probabilidades) da curva Normal

$$\boxed{Z = \frac{X - \mu}{\sigma}} \tag{18}$$

- Entendendo melhor...



- Exemplo
  - Probabilidade de uma pessoa apresentar entre 200-225mg por 100ml de plasma. $\mu = 200$ e $\sigma = 20$

$$z = \frac{225-200}{20} = 1,25$$
$$P(0 \leq z \leq 1,25) = 0,3944$$

- Exercício