O que é e para que serve Inferência Estatística?

- Esta moeda é justa?
- Esta droga "funciona"?
- Quantos casos de Dengue teremos mês que vem?
- Renda básica universal aumenta o PIB?

Todas essas perguntas podem ser abordadas com as ferramentas que a Estatística nos fornece.

Ideia 2

A Estatística é a gramática da Ciência[4]*. O mundo é incerto; medições são imperfeitas. A Estatística é a linguagem que nos permite expressar e quantificar as incertezas associadas às afirmações científicas através da teoria de probabilidades*[5]*.*

[4] Título do livro de Karl Pearson (1857–1936) ("The Grammar of Science"), publicado em 1982.
[5] Chamada por E.T. Jaynes (1922-1998) de lógica da Ciência ("Probability Theory: The Logic of Science").

Modelo estatístico: definição informal

Definição 4

Modelo estatístico *(De Groot, def 7.1.1, pág. 377). Um modelo estatístico consiste na identificação de variáveis aleatórias de interesse (observáveis e potencialmente observáveis), na especificação de uma distribuição conjunta para as variáveis aleatórias observáveis e na identificação dos parâmetros (θ) desta distribuição conjunta. Às vezes é conveniente assumir que os parâmetros são variáveis aleatórias também, mas para isso é preciso especificar uma distribuição conjunta para θ.*

Definição 5

Modelo estatístico *(McCullagh, 2002). Seja $\mathcal{X}$ um espaço amostral qualquer, Θ um conjunto não-vazio arbitrário e $\mathcal{P}(\mathcal{X})$ o conjunto de todas as distribuições de probabilidade em $\mathcal{X}$. Um modelo estatístico paramétrico é uma função $P : \Theta \to \mathcal{P}(\mathcal{X})$, que associa a cada $\theta \in \Theta$ uma distribuição de probabilidade P_θ em $\mathcal{X}$.*

Exemplos:

- Faça $\mathcal{X} = \mathbb{R}$ e $\Theta = (-\infty, \infty) \times (0, \infty)$. Dizemos que P é um modelo[6] estatístico normal se para cada $\theta = \{\mu, \sigma^2\} \in \Theta$,

$$P_\theta(x) \equiv \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \; x \in \mathbb{R}.$$

- Faça $\mathcal{X} = \mathbb{N} \cup \{0\}$ e $\Theta = (0, \infty)$. P é um modelo estatístico Poisson se para $\lambda \in \Theta$,

$$P_\lambda(k) \equiv \frac{e^{-\lambda}\lambda^k}{k!}, \; k = 0, 1, \ldots$$

[6] Note o abuso de notação: estritamente falando, P_θ é uma **medida** de probabilidade e não uma *densidade* como apresentamos aqui.

Exemplo: como sempre, moedas.

Pergunta 3

Suponha que uma moeda tenha sido lançada dez vezes, obtendo o seguinte resultado:

$$KKKCKCCCKC$$

a) Esta moeda é justa?

b) Quanto eu espero ganhar se apostar R\$ 100,00 que é justa?

Podemos formalizar o problema ao, por exemplo, assumir que cada lançamento é uma variável aleatória Bernoulli com probabilidade de cara (K), p. Desta forma $X_i = 1$ se o lançamento deu cara e $X_i = 0$ caso contrário. E queremos saber se $p = 1/2$. Por ora, não temos as ferramentas necessárias para responder a essa pergunta, mas voltaremos a ela no futuro.

Definição 6

***Afirmação probabilística**. Dizemos que uma afirmação é probabilística quando ela utiliza conceitos da teoria de probabilidade para falar de um objeto. Exemplos:*

- $\Pr(\bar{Y}_n \in (0,1)) \leq 2^n$;
- $E[X \mid Y = y] = 2y + 3$;
- $\mathrm{Var}(X) = 4p^2$.
- $\Pr(\mathrm{Var}(X) \leq 4p^2) \leq p^2$

Definição 7

***Inferência Estatística**. Uma inferência estatística é uma afirmação probabilística sobre uma ou mais partes de um modelo estatístico. Considerando o exemplo 3, queremos saber:*

- *Quantos lançamentos até termos 80% de certeza de que a moeda é justa?*
- *Quanto vale $E[\bar{X}_n]$;*
- $\Pr(X_n = 1 \mid X_{n-1} = 1)$.

Definição 8

***Estatística**. Suponha que temos uma coleção de variáveis aleatórias $X_1, X_2, \ldots, X_n \in \boldsymbol{X} \subseteq \mathbb{R}^n$ e uma função $r : \boldsymbol{X} \to \mathbb{R}^m$. Dizemos que a variável aleatória $T = r(X_1, X_2, \ldots, X_n)$ é uma **estatística**.*

São exemplos de estatísticas:

- A média amostral, $\bar{X}_n$;
- A soma, $\sum_{i=1}^{n} X_i$;
- O mínimo, $\min(X_1, X_2, \ldots, X_n)$;
- $r(X_1, X_2, \ldots, X_n) = a, \ \forall X_1, X_2, \ldots, X_n, \ a \in \mathbb{R}$.

Tipos de Inferência Estatística

- **Predição**: prever o valor de uma variável aleatória (ainda) não observada; No exemplo 3, qual será o valor do próximo lançamento, X_{n+1};
- **Decisão Estatística**: Acoplamos o modelo estatístico a uma decisão a ser tomada. Devo emprestar esta moeda ao Duas-Caras? Aqui, temos a *noção* de **risco**.;
- **Desenho experimental**: Quantas vezes é preciso lançar esta moeda para ter 95% de certeza de que ela é (ou não) justa? Quantas pessoas precisam tomar uma droga para sabermos se ela funciona? Onde devemos cavar para procurar ouro/petróleo?;

O que aprendemos?

- Modelo estatístico;
- Inferência Estatística;
- Estatística (amostral);
- Tipos de inferências:
  - Predição;
  - Decisão;
  - Desenho experimental.

Leitura recomendada

- De Groot seção 7.1;
- * McCullagh, 2002.