Inferência Estatística: Conceitos Básicos

Prof. Renzo Flores-Ortiz

- **<u>Experimento aleatório</u>**: é um processo ou procedimento que pode ser repetido nas mesmas condições e cujo **resultado não pode ser previsto com certeza** antes de sua realização.

- Como o resultado de um experimento aleatório não pode ser previsto (isto é, o **resultado é aleatório**), dizemos que um experimento aleatório envolve **incerteza.**

- Apesar da incerteza no resultado individual, o conjunto de todos os resultados possíveis é conhecido e pode ser descrito formalmente (isso é o **espaço amostral**).

- **Exemplos**
  - <u>Lançar uma moeda</u>: os resultados possíveis são “cara” ou “coroa”, mas não se sabe qual face sairá.
  - <u>Lançar um dado</u>: os possíveis resultados são os números de 1 a 6, mas o número que sairá é imprevisível.
  - <u>Medir o tempo de vida de uma lâmpada</u>: o tempo de vida é incerto, mas será um número real positivo.

- Uma **variável aleatória** X é uma função:

$$X: \Omega \longrightarrow \mathbb{R}$$

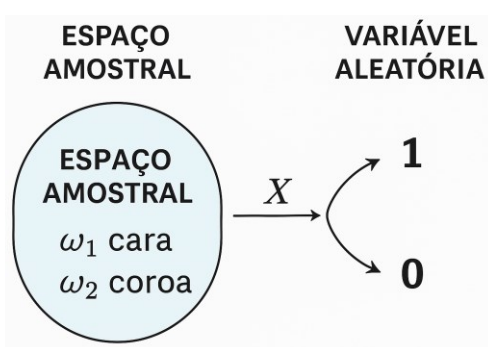
- Ω é o espaço amostral, ou seja, o conjunto de todos os resultados possíveis do experimento aleatório.
- $\mathbb{R}$ é o conjunto dos número reais.
- $X(\omega)$ representa o valor numérico assumido por X quando ocorre o resultado $\omega \in \Omega$.

- Em palavras: uma variável aleatória é uma função que associa a cada resultado possível de um experimento aleatório um número real.

- Ou seja, ela traduz os possíveis resultados de um experimento aleatório em valores numéricos, que podem então ser analisados matematicamente.

Exemplo

- **Experimento aleatório**: lançar uma moeda e observar a face superior (visível).
- **Espaço amostral**: $\Omega = \{cara, coroa\}$
- **Variável aleatória**: $X(\omega) = \begin{cases} 1, se\ \omega = cara \\ 0, se\ \omega = coroa \end{cases}$



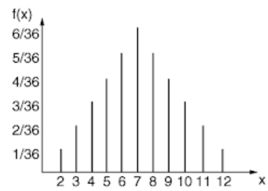
Classificação de variáveis aleatórias quanto à sua natureza ou tipo de valores que podem assumir

- **Variável aleatória discreta** – assume valores inteiros.
  - Exemplo: número de caras ao lançar uma moeda 3 vezes (valores possíveis: 0, 1, 2, 3).
- **Variável aleatória contínua** – assume qualquer valor em um intervalo contínuo.
  - Exemplo: tempo até a falha de um equipamento (pode ser 3,2 horas, 3,21, 3,215 horas, ...).
- **Variável aleatória categórica** (ou qualitativa) – assume valores em forma de categorias ou rótulos, que precisam ser **codificados como números** para permitir o uso em análises e modelos que exigem entradas numéricas.
  - Exemplo: nível de escolaridade (fundamental = 0, médio = 1, superior = 2).

- Embora não seja possível prever qual valor será assumido por uma variável aleatória, é possível atribuir probabilidades aos possíveis valores. Isto é, uma variável aleatória tem uma **distribuição de probabilidade**.
- A **probabilidade** mede a chance de um evento ocorrer, variando entre **0** (evento impossível) e **1** (evento certo).

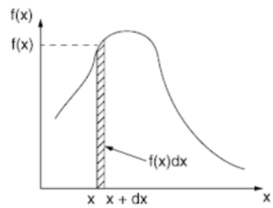
- Uma **distribuição de probabilidade** descreve como as probabilidades são atribuídas aos diferentes valores que uma variável aleatória pode assumir.

- No caso de **variáveis aleatórias discretas**, a distribuição especifica diretamente a **probabilidade de ocorrência de cada valor possível**.



Probability mass function

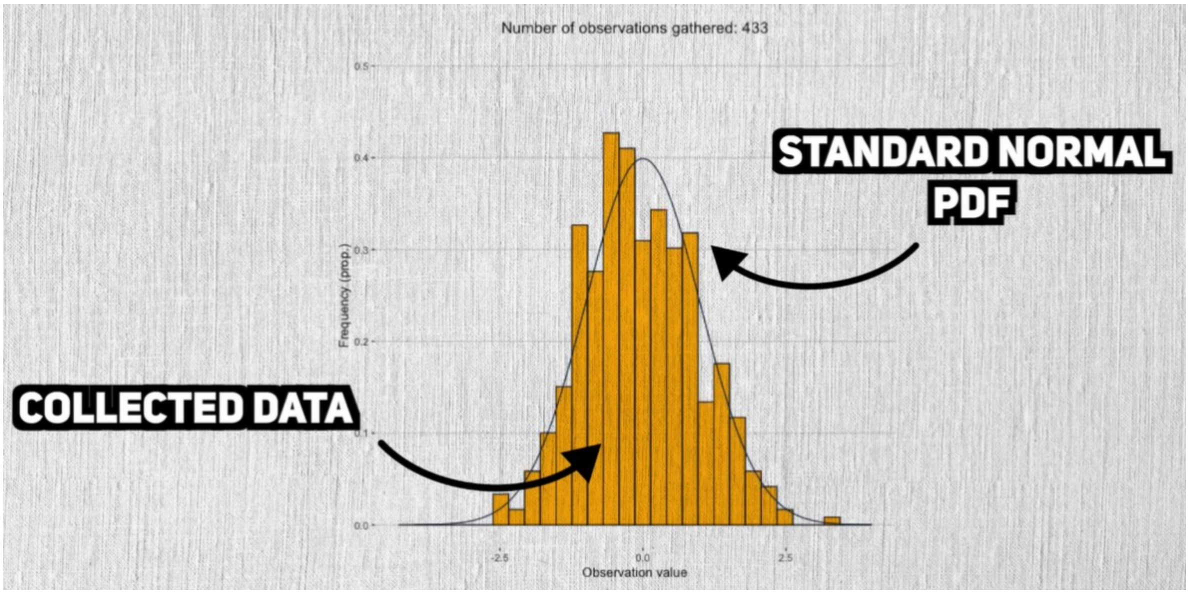
- Já para **variáveis aleatórias contínuas**, a distribuição é representada por uma **função densidade de probabilidade** (PDF), e as probabilidades são calculadas como **áreas sob a curva** dessa função, em intervalos de valores.



Continuous probability density function

- Onde podemos encontrar variáveis aleatórias?

X1	X2	X3
0.876	0	2
0.126	1	5
1.67	1	1



Probability Distributions

Ace Tutors

Continuous

Uniform

$$\mu = \frac{a+b}{2} \qquad \sigma = \sqrt{\frac{(b-a)^2}{12}}$$

$$P(X < x) = \frac{x-a}{b-a}$$

Exponential

$$\mu = \frac{1}{\gamma} \qquad \sigma = \frac{1}{\gamma}$$

$$P(X < x) = 1 - e^{-\gamma x}$$

Normal

$$z = \frac{x-\mu}{\sigma}$$

$P(X < x) \Rightarrow$ Use Z-Chart

Key

- γ = *rate parameter*
- z = *z-score*
- p = *probability of success*
- n = *# of trials*
- N = *population size*
- K = *# of success states*

Discrete

Binomial

$$\mu = n \cdot p \qquad \sigma = \sqrt{n \cdot p \cdot (1-p)}$$

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

Geometric

$$\mu = \frac{1}{p} \qquad \sigma = \frac{\sqrt{1-p}}{p}$$

$$P(X = x) = (1-p)^{x-1} p$$

Hypergeometric

$$\mu = n\frac{K}{N} \qquad \sigma = \sqrt{n\frac{K(N-K)(N-n)}{N^2(N-1)}}$$

$$P(X = x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$