Estatística para Ciência de Dados

Aula 2: Probabilidades

Francisco A. Rodrigues
ICMC/USP
francisco@icmc.usp.br

Aula 2: Probabilidades

- Conceitos fundamentais
  Teorema de Bayes
  Variáveis aleatórias

Probabilidades

- **Experimento aleatório**
- Cada experimento pode ser executado um número infinito de vezes sob condições inalteradas.




Probabilidades

- **Espaço amostral (Ω):** Conjunto de saídas do experimento.
- **Evento (A):** Um elemento do espaço amostral.
- **Relações entre eventos:**
    - Evento impossível: $\emptyset$
    - Evento certo: Ω
    - $A \cup B$: é o evento que ocorre se A, ou B (ou ambos) ocorrem.
    - $A \cap B$: é o evento que ocorre se, e somente se, A e B ocorrem.
    - $\overline{A}$: é o evento que ocorre se A não ocorre.
    - Eventos mutualmente exclusivos: $A \cap B = \emptyset$

Probabilidades

- **Probabilidades: (Kolmogorov, 1933)**
- Uma função P(.) é denominada uma medida de probabilidade se satisfaz:
  - $0 \leq P(A) \leq 1, \forall A \in \Omega$
  - $P(\Omega)$=1
  - Se A_1, A_2, A_2, ... forem eventos mutuamente exclusivos, então:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i), A_i \cap A_j = \emptyset, i \neq j$$

Probabilidades

- **Probabilidades: Definição Clássica**
- Se um experimento aleatório tiver $n(\Omega)$ resultados mutuamente exclusivos e igualmente possíveis, e se um evento A tiver $n(A)$ desses resultados, a probabilidade de ocorrer o evento A é:

$$P(A) = \frac{n(A)}{n(\Omega)}$$

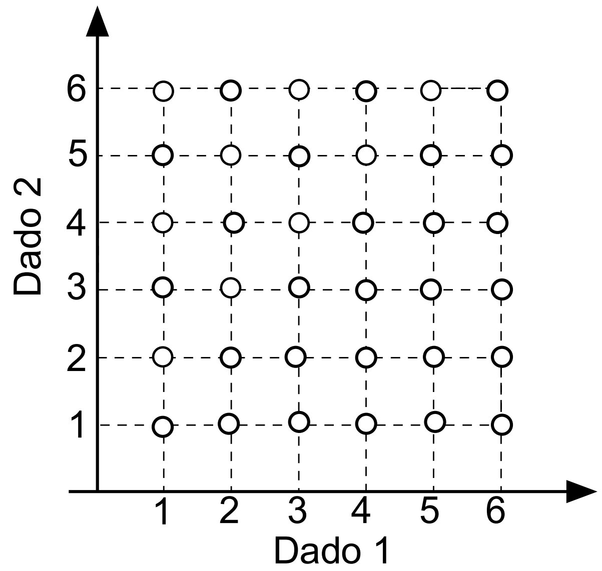
Probabilidades

- **Exemplo:**

Lançando dois dados equilibrados, qual é a probabilidade de que:

A soma das faces seja igual a 7:

Obter uma soma maior do que5.



Probabilidades

- **Definição: (frequentista)**
- A probabilidade de um evento é igual à sua frequência de ocorrência em muitos experimentos.

$$P(A) = \lim_{n \to \infty} \frac{n_A}{n}$$

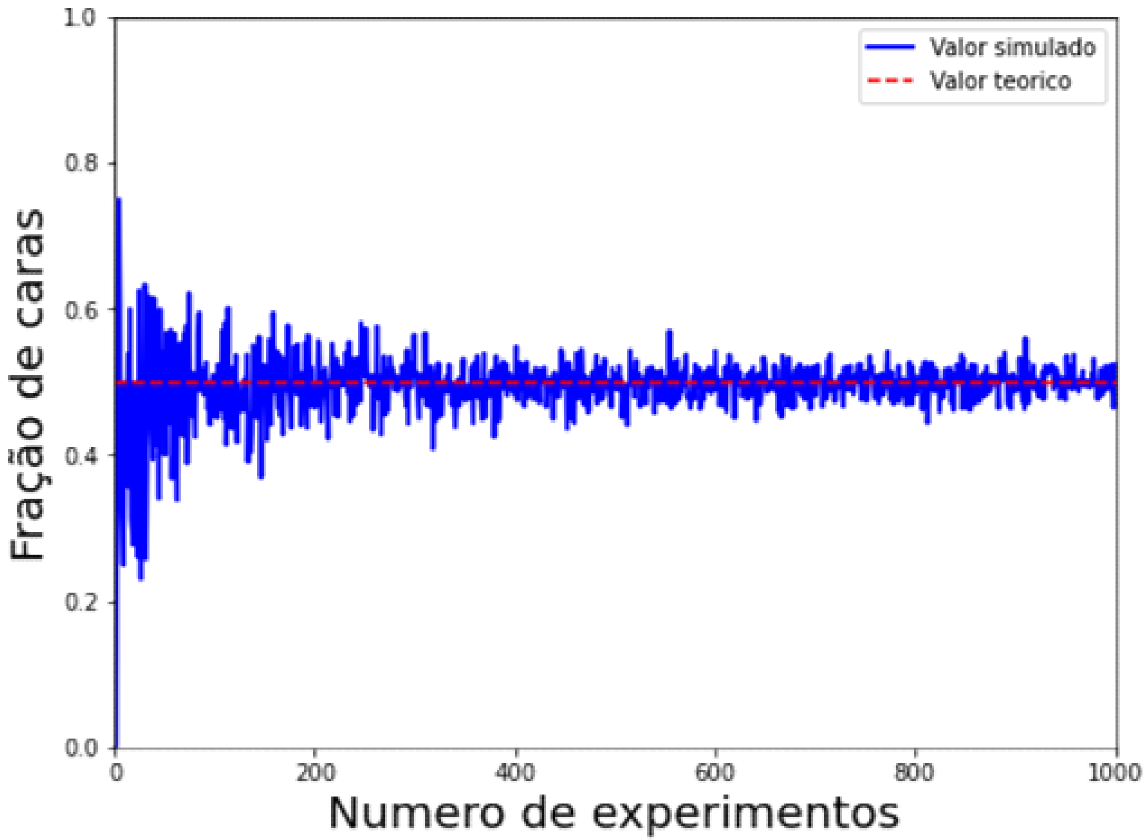
Probabilidades

```python
import random
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline


vp = [] # lista que armazena a fração de ocorrências em função do número de simulações
vsim = [] # armazena o número de simulações
Nmax = 1000 # numero maximo de simulacoes
moeda = ['C','R']
for nsim in np.arange(1,Nmax):
    nhead = 0 # numero de caras
    for i in range(1,nsim):
        face = random.choice(moeda)
        if(face == 'C'):
            nhead = nhead + 1
    vp.append(nhead/nsim)
    vsim.append(nsim)

plt.figure(figsize=(8,6))
plt.plot(vsim, vp, linestyle='-', color="blue", linewidth=2,label = 'Valor simulado')
plt.axhline(y=1/2, color='r', linestyle='--', label = 'Valor teorico')
plt.ylabel("Fração de caras", fontsize=20)
plt.xlabel("Numero de experimentos", fontsize=20)
plt.xlim([0.0, Nmax])
plt.ylim([0.0, 1.0])
plt.legend()
plt.show(True)
```

Probabilidades



Probabilidades

- **Probabilidade da união de eventos**

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- **Exemplo:** Qual é a probabilidade de que em um lançamento de um dado saia um número par ou maior do que três?

Probabilidades

- **Probabilidade Condicional**

dado que ocorreu um evento B a probabilidade de ocorrer o A é a intersecção entre A e B sobre o B

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, P(B) > 0.$$

- **Exemplo:** Calcule P(A|B), onde: $\Omega = \{1,2,3, \dots, 15\}$, A é um número par e B é um número maior do que 5.

Probabilidades

- **Exemplo:**
- Uma caixa usada em um sorteio contém 5 bolas pretas numeradas de 1 a 5, sete bolas brancas numeradas de 1 a 7 e oito bolas vermelhas numeradas de 1 a 8.

  - i) Sorteando-se uma bola dessa caixa, qual é a probabilidade de encontrarmos uma bola preta?

  - ii) Sorteando-se uma bola preta, qual é a probabilidade de tal bola ser par?

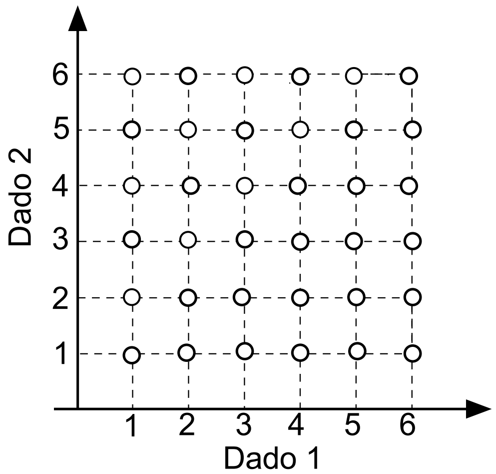
Probabilidades

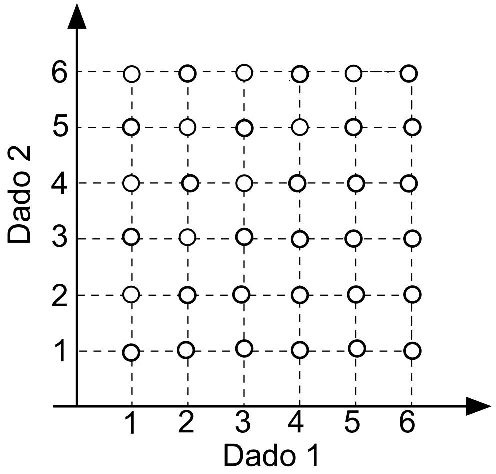
- **Eventos independentes**
- A e B são eventos independentes se, e somente se,

$$P(A \cap B) = P(A)P(B).$$

Probabilidades

- **Exemplo**: Dois dados são lançados. Sejam os eventos:
  - A: “a soma dos dados é igual a 6”
  - B: “saiu o valor 4 no primeiro dado”
- A e B são independentes?





Probabilidades

- **Teorema de Bayes**

Sejam $B\_1,\ldots,B\_k$ uma partição do espaço amostral Ω. Então, qualquer evento A ∈Ω pode ser escrito como:

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^{\infty} P\left(A\middle|B_j\right)P(B_j)}$$

Probabilidades

- **Exemplo:** Um laboratório que faz testes sanguíneos apresenta eficácia de 95% na detecção de uma certa doença quando, de fato, a pessoa está doente (verdadeiro positivo). A taxa de falso positivos, ou seja, quando o teste afirma que o paciente tem a doença, embora seja saudável, é de 2%. Se 0,1% da população realmente tem a doença, qual é a probabilidade de que uma pessoa tenha a doença dado que o teste foi positivo?

Probabilidades

- **Variáveis aleatórias**

- Uma variável aleatória é uma função que associa um valor real a cada elemento do espaço amostral.

Probabilidades

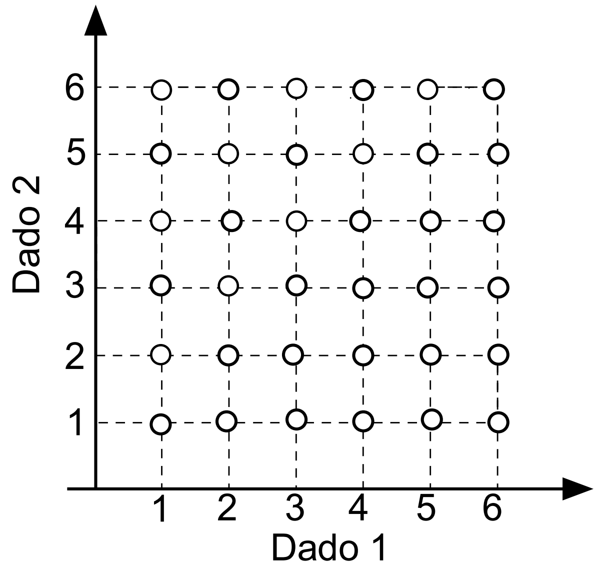
- **Exemplo**
- Lançamos duas moedas. Seja X a variável aleatória que indica o número de caras. Determinar a distribuição de probabilidade de X.

Probabilidades

- Se X é uma variável aleatória discreta:
  - $0 \leq P(X = x) \leq 1$
  - $P(X = x) \geq 0,$
  - $\sum_{i=1}^{\infty} P(X = x_i) = 1$
  - $P(a \leq X \leq b) = \sum_{x=a}^{x=b} P(X = x)$

Probabilidades

- **Exemplo:**
- Lançamos dois dados e observamos a variável aleatória X é igual a 1 se a soma for par ou igual a zero, caso contrário. Determine a distribuição de X.



Probabilidades

- Se X é uma variável aleatória contínua, existe uma função f(x), denominada função densidade de probabilidade, onde:
  - $f(x) \geq 0, \forall x \in \mathbb{R}$
  - $\int_{-\infty}^{\infty} f(x)dx = 1$
  - $P(a < X < b) = \int_{a}^{b} f(x)dx$

Probabilidades

- **Exemplo:**

- A variável aleatória X tem função densidade de probabilidade dada por:

$$\begin{cases} \dfrac{x^2}{3} \; se -1 < x < 2 \\ 0 \;\; caso \; contrário \end{cases}$$

- Calcule $P(0 \leq X \leq 1)$.

Probabilidades

https://www.wolframalpha.com/

WolframAlpha computational intelligence.

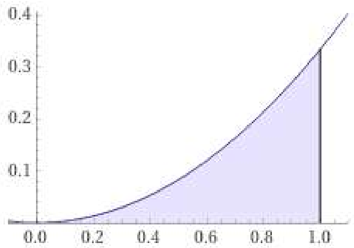
integrate (x^2)/3 from 0 to 1

Extended Keyboard | Upload | Examples | Random

Definite integral: More digits | Step-by-step solution

$$\int_0^1 \frac{x^2}{3}\,dx = \frac{1}{9} \approx 0.11111$$

Visual representation of the integral:



Indefinite integral: Step-by-step solution

$$\int \frac{x^2}{3}\,dx = \frac{x^3}{9} + \text{constant}$$

Sumário

- Probabilidades
  Teorema de Bayes
  Variáveis aleatórias

Leitura Complementar

- Morettin e Bussab, **Estatística Básica**, Saraiva, 2017.

Probabilidades

- **Exemplo:**
- Duas bolas são escolhidas sem reposição de uma caixa com 5 bolas brancas e 8 pretas. Seja X_i igual a um se a *i*-ésima bola selecionada é branca e igual zero, caso contrário. Calcule $P(X_2 = 0|X_1 = 0)$ e $P(X_2 = 1|X_1 = 1)$.

Probabilidades

- **Exemplo**

Suponha que nós temos três cartas idênticas no formato, mas que ambos os lados da primeira carta são vermelhos, ambos os lados da segunda carta são pretos e um lado da terceira carta é vermelho e o outro preto. As três cartas são misturadas em um chapéu e uma carta é removida e colocada sobre uma mesa, de modo que apareça somente um lado. Se o lado mostrado é vermelho, qual é a probabilidade de que o outro lado seja preto?