Estatística para Ciência de Dados

Introdução à Teoria das Probabilidades

Francisco A. Rodrigues

Instituto de Ciências Matemáticas e de Computação

Universidade de São Paulo

francisco@icmc.usp.br

Início

- Um erudito jogador: **Girolano Cardano**
- Livro: *Liber de Ludo Aleae* (O livro dos jogos de Azar): **1633**
- Esse foi o primeiro livro a fornecer um tratamento sistemático da matemática da probabilidade. (Há um capítulo sobre como trapacear e se dar bem no jogo…)

Início

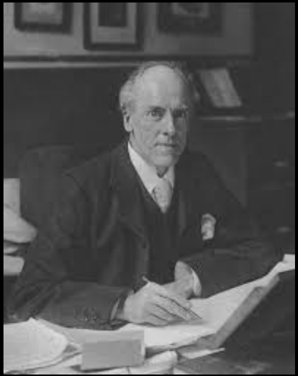
- Em cartas trocadas entre Fermat e Blaise Pascal em 1654, foi criada a teoria das probabilidades.
- Pergunta: Como as apostas devem ser divididas se um jogo for terminado antes do final.

Blaise Pascal
(1623-1662)

Pierre de Fermat
(1601-1665)

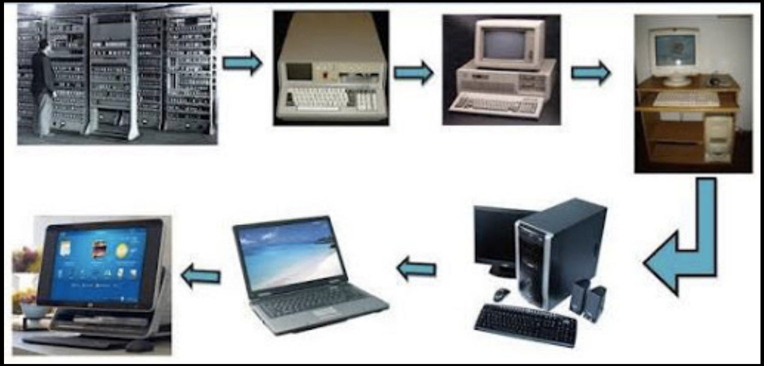
https://www.york.ac.uk/depts/maths/histstat/pascal.pdf

Século XX



Karl Pearson
(1857-1936)

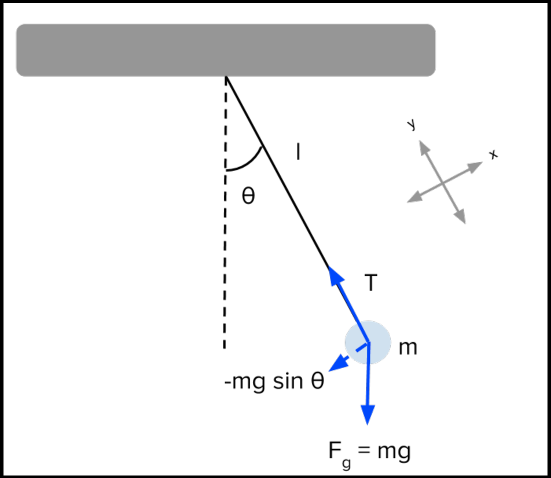
Ronald Fisher
(1890-1962)

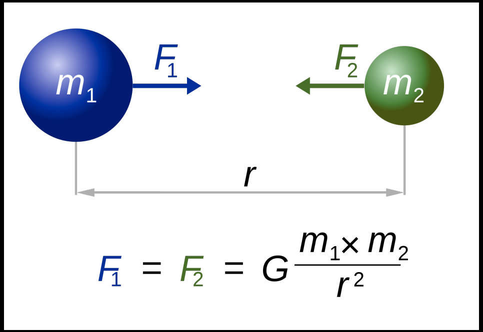


Determinístico X Probabilístico

Modelo determinístico:

Dada as condições iniciais, podemos prever a saída de um experimento com certeza.





$$
\begin{aligned}
\nabla \cdot \mathbf{D} &= \rho \\
\nabla \cdot \mathbf{B} &= 0 \\
\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\
\nabla \times \mathbf{H} &= \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}
\end{aligned}
$$

Determinístico X Probabilístico

Esperimento aleatório:

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




Determinístico X Probabilístico

- Quando lançamos um dado, obtemos uma sequência de valores.

$$X = [1,4,5,6,3,1,2,3,1,6,5,3,2,3,4,3,2,3,3,4,5,..]$$

- Não somos capazes de dizer qual será o próximo valor.
- Mas podemos determinar a probabilidade de cada saída possível.

Determinístico X Probabilístico

- Por exemplo, usando Python:

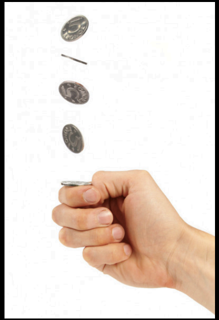
```
import numpy as np
p = 0.6 # probabilidade de sair cara
nsim = 10 # num de experimentos
nhead = 0 # num de caras obtidas
saida = [] # armazena as saidas (cara:1, coroa:0)
for i in range(0, nsim):
    if(np.random.uniform() < p): # se menor que p, cara
        nhead = nhead + 1 # incrementa o contador de caras
        saida.append(1)
    else:
        saida.append(0)
print("Saida:", saida)
print("Frequencia de caras:", nhead/nsim)
```

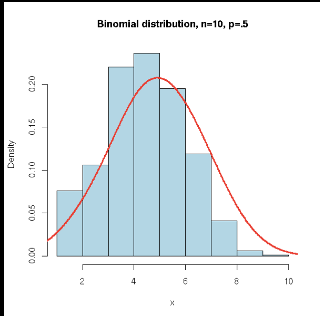
```
Saida: [1, 1, 0, 0, 1, 1, 0, 1, 0, 1]
Frequencia de caras: 0.6
```

Determinístico X Probabilístico

- **Modelos probabilísticos:**
  - As saídas de cada experimento parecem imprevisíveis, mas quando se analisa um grande número de experimentos, um padrão emerge.
  - Não podemos determinar exatamente o valor da saída de um experimento, mas as probabilidades de cada saída possível.
  - Exemplo:

$$P(X=k)=\frac{n!}{(n-k)!k!}p^k(1-p)^{n-k}, \quad k=0,1,\ldots,n$$





Conceitos fundamentais

- **Espaço amostral** (Ω): Conjunto de saídas do experimento.
- **Evento** (A): Um elemento do espaço amostral.
- Relações entre eventos:
  - **Evento impossível**:$\varnothing$
  - **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.
  - $\bar{A}$: é o evento que ocorre se A não ocorre.
  - **Eventos mutualmente exclusivos**: $A \cap B = \varnothing$

Conceitos fundamentais

- **Exemplo**: Sejam A e B dois eventos em um mesmo espaço amostral.

  Escreva na linguagem dos conjuntos:

  - Pelo menos um dos eventos ocorre.

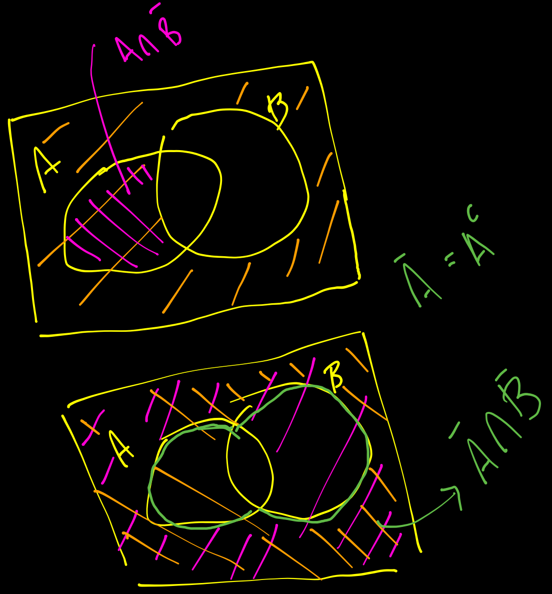
    $$A \cup B$$

  - O evento A ocorre, mas B não ocorre.

    $$A \cap \bar{B}$$

  - Nenhum deles ocorre.

    $$\bar{A} \cap \bar{B}$$



Axiomas da probabilidade

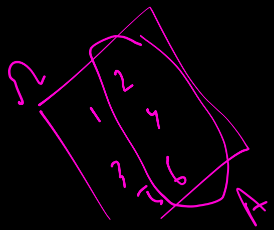
- **Definição: (Kolmogorov, 1933)**
- Uma função P(.) é denominada uma medida de probabilidade se satisfaz:
  - $0 \leq P(A) \leq 1, \quad \forall A \in \Omega$
  - $P(\Omega) = 1$
  - Se $A_1, A_2, \ldots,$ forem eventos mutuamente exclusivos, então:

$$P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i), \quad A_i \cap A_j = \varnothing \quad \forall i \neq j$$

Probabilidade clássica ou a priori

- **Definição:**
- 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:

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



$P(A) = \frac{3}{6} = \frac{1}{2}$

Ex: 6B 4P

$P(A) = \frac{4}{10} = 0,4$

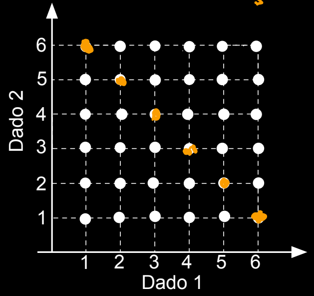
Probabilidades

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

- A soma das faces seja igual a 7:

$P(A) = \frac{6}{36} = \frac{1}{6}$

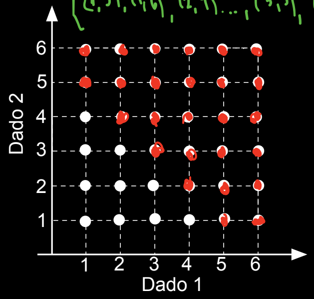
Ω



- Obter uma soma maior do que 5.

$P(A) = \frac{26}{36} = \frac{13}{18}$

$\{(1,5),(1,6),(2,4)\ldots,(3,3),(3,1)\ldots\}$



Probabilidades

- **Solução**

3 - O espaço amostral é dado pelo produto cartesiano entre os valores das faces de cada dado, ou seja, $\Omega = \{(1,1),(1,2),\ldots,(2,1),(2,2),\ldots,(6,4),(6,5),(6,6)\}$. Tais elementos são representados por todos os pontos na figura 1.5. Portanto, temos 36 elementos no espaço amostral.

a) Seja o evento: A : "a soma das faces é igual a 7". Como $A = \{(1,6),(6,1),(4,3),(3,4),(5,2),(2,5)\}$, temos

$$P(A) = \frac{|A|}{|\Omega|} = \frac{6}{36} = \frac{1}{6}.$$

b) Seja o evento: B: "A soma dos dados é maior do que 5". Na figura 1.5 mostramos os elementos do evento B. Portanto, $B = \{(4,2),(4,4),(4,4),(4,5),(4,6),(5,1),(5,2),(5,3),\ldots,(6,6)\}$. Como temos 26 elementos em B,

$$P(B) = \frac{26}{36} = \frac{13}{18}.$$

c) Seja o evento C : "o valor no primeiro dado é maior do que no segundo". Na figura 1.6 mostramos os elementos de C. Assim,

$$P(C) = \frac{15}{36}.$$

Probabilidades

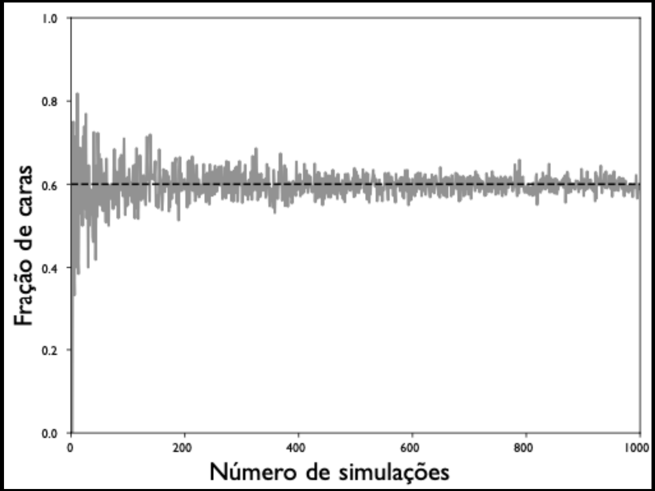
- **Definição**: (frequentista)
- A probabilidade de um evento é igual à sua frequência de ocorrência em muitos experimentos.

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

Probabilidades

- **Definição**: (frequentista): Python

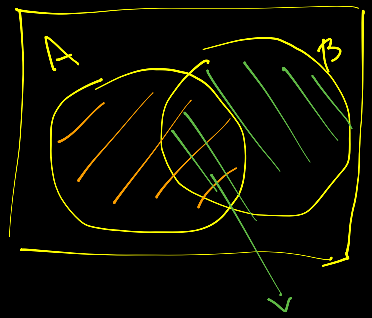
```
import numpy as np
import matplotlib.pyplot as plt
p = 0.6 # chance de sair cara
vp = [] # lista que armazena a frequencia de ocorrencias
vsim = [] # armazena o numero de simulacoes
Nmax = 1000 # numero maximo de simulacoes
for nsim in np.arange(1,Nmax,10): # Simula de 1 ate Nmax
    nhead = 0 # numero de caras
    for i in range(1,nsim): # o numero de simulacoes aumenta
        if(np.random.uniform() < p):
            nhead = nhead + 1
    vp.append(nhead/nsim)
    vsim.append(nsim)
# Mostra os resultados em uma figura
plt.figure(figsize=(8,6))
plt.plot(vsim, vp, linestyle='-', color="blue", linewidth=2,
                                label = 'Valor simulado')
plt.axhline(y=p, color='r', linestyle='--', label = 'Valor
                                teorico')
plt.ylabel("Fracao 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

- Probabilidade da união de eventos:

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



2+

Probabilidades

- **Exemplo**: Em uma escola particular, entre todos os alunos que procuram ajuda, 63% precisam de ajuda em matemática, 34% precisam de ajuda em inglês e 27% precisam de ajuda tanto em matemática quanto em inglês. Qual é a porcentagem de alunos que precisam de ajuda em matemática ou em inglês (ou em ambas disciplinas)?

A B

$P(A) = 0{,}63$

$P(B) = 0{,}34$

$P(A \cap B) = 0{,}27$

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

$= 0{,}63 + 0{,}34 - 0{,}27 = 0{,}97 - 0{,}27 = \boxed{0{,}70}$

Probabilidades

- Solução

Exemplo: Em uma escola particular, entre todos os alunos que procuram ajuda, 63% precisam de ajuda em matemática, 34% precisam de ajuda em inglês e 27% precisam de ajuda tanto em matemática quanto em inglês. Qual é a porcentagem de alunos que precisam de ajuda em matemática ou em inglês (ou em ambas disciplinas)?

Sejam os eventos:
M : "estudantes precisam de ajuda em matemática",
E : "estudantes precisam de ajuda em inglês".

Assim, temos:

$$P(M \cup E) = P(M) + P(E) - P(M \cap E).$$

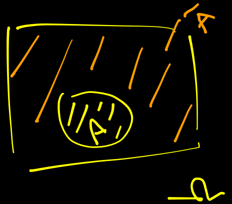
Pelo enunciado do problema, temos $P(M) = 0,63$, $P(E) = 0,34$ e $P(M \cap E) = 0,27$. Assim,

$$P(M \cup E) = P(M) + P(E) - P(M \cap E) = 0,63 + 0,34 - 0,27 = 0,70.$$

Portanto, 70% dos alunos precisam de ajuda tanto em matemática quanto em inglês. ■

Probabilidades

- **Propriedades:** Sejam $\varnothing$ o evento impossível e A e B dois eventos em um mesmo espaço amostral Ω. Seja ainda $A^c = \bar{A}$ o complementar de A. Então:
- $P(\varnothing) = 0$
- $P(A) = 1 - P(A^c)$ $\quad \bar{A}$
- Se $A \subset B$, então $P(A) \leq P(B)$.



$A \cup \bar{A} = \Omega$

$A \cap \bar{A} = \varnothing$

$A \cup \varnothing = A$

$A \cap \varnothing = \varnothing$

$P(A) = P(A \cup \varnothing)$

$P(A) = P(A) + P(\varnothing)$

$\boxed{P(\varnothing) = 0}$

$A \cap B = A$

$B = A \cup (\bar{A} \cap B)$

$P(B) = P(A) + P(\bar{A} \cap B)$

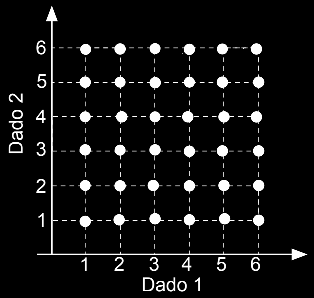
$P(B) \geq P(A)$

$\Omega = A \cup \bar{A}$

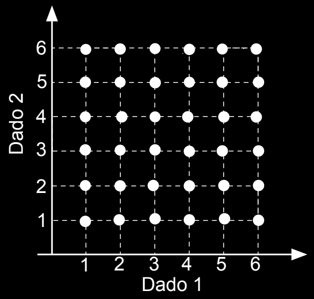
$P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A}) \Rightarrow 1 = P(A) + P(\bar{A}) \rightarrow P(A) = 1 - P(\bar{A})$

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 que 5.

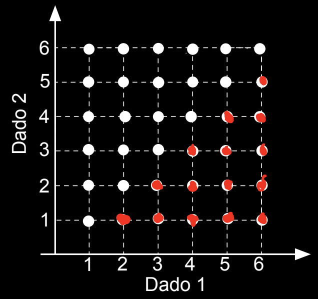


Probabilidades

- **Exemplo:** Lançamos dois dados equilibrados.

  c) Qual é a probabilidade de que a o valor na face do primeiro dado seja maior do que na do segundo?

$P(A) = \frac{15}{36}$



Probabilidades

- **Solução:**

3 - O espaço amostral é dado pelo produto cartesiano entre os valores das faces de cada dado, ou seja, $\Omega = \{(1,1),(1,2),\ldots,(2,1),(2,2),\ldots,(6,4),(6,5),(6,6)\}$. Tais elementos são representados por todos os pontos na figura 1.5. Portanto, temos 36 elementos no espaço amostral.

a) Seja o evento: A : "a soma das faces é igual a 7". Como $A = \{(1,6),(6,1),(4,3),(3,4),(5,2),(2,5)\}$, temos

$$P(A) = \frac{|A|}{|\Omega|} = \frac{6}{36} = \frac{1}{6}.$$

b) Seja o evento: B: "A soma dos dados é maior do que 5". Na figura 1.5 mostramos os elementos do evento B. Portanto, $B = \{(4,2),(4,4),(4,4),(4,5),(4,6),(5,1),(5,2),(5,3),\ldots,(6,6)\}$. Como temos 26 elementos em B,

$$P(B) = \frac{26}{36} = \frac{13}{18}.$$

c) Seja o evento C : "o valor no primeiro dado é maior do que no segundo". Na figura 1.6 mostramos os elementos de C. Assim,

$$P(C) = \frac{15}{36}.$$

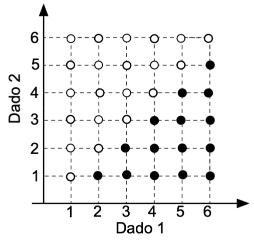


Figura 1.5: Resultados possíveis da saída de dois dados. Os pontos em preto representam o evento B: "A soma dos dados é maior do que 5".

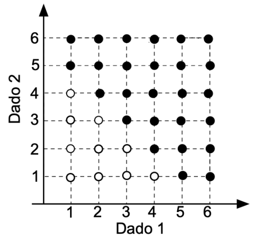


Figura 1.6: Resultados possíveis da saída de dois dados. Os pontos em preto representam o evento C : "o valor no primeiro dado é maior do que no segundo".

Sumário

- Tipos de experimentos
- Definição de probabilidades: clássica, axiomática, frequentista.
- Eventos e espaço amostral

https://www.probabilitycourse.com