Teoria da Probabilidade

Rodrigo R. Pescim

Universidade Estadual de Londrina

07 de maio 2020

- *"A razão do número de todos os casos favoráveis à um acontecimento, para o de todos os casos possíveis é a probabilidade buscada, a qual é portanto uma fração* $[\cdots]$*"*

- *"A teoria da probabilidade nada mais é do que o senso comum reduzido à cálculo."*

Pierre Simon Laplace
Ensaio Filosófico Sobre as Probabilidades (1812)

Teoria da Probabilidade

- É a base sobre a qual a estatística é desenvolvida.

Teoria da Probabilidade

- É a base sobre a qual a estatística é desenvolvida.

- Fornece um meio para modelar populações, experimentos ou **algum fenômeno aleatório** qualquer.

Teoria da Probabilidade

- É a base sobre a qual a estatística é desenvolvida.
- Fornece um meio para modelar populações, experimentos ou **algum fenômeno aleatório** qualquer.
- **objetivo:** Descrever ideias básicas desta teoria, que são fundamentais para o estudo da estatística.

Experimento aleatório ($\mathcal{E}$)

- Um dos maiores objetivos de um estatístico é chegar a conclusões sobre certa população de objetos por meio da **realização de um experimento**.

Experimento aleatório ($\mathcal{E}$)

- Um dos maiores objetivos de um estatístico é chegar a conclusões sobre certa população de objetos por meio da **realização de um experimento**.

- Um experimento é qualquer processo de observação.

Experimento aleatório ($\mathcal{E}$)

- Um dos maiores objetivos de um estatístico é chegar a conclusões sobre certa população de objetos por meio da **realização de um experimento**.

- Um experimento é qualquer processo de observação.

- Em muitos experimentos, **observa-se um elemento de incerteza** e essencialmente a **impossibilidade** de predizer seu comportamento em futuras realizações.

Experimento aleatório ($\mathcal{E}$)

- Um dos maiores objetivos de um estatístico é chegar a conclusões sobre certa população de objetos por meio da **realização de um experimento**.
- Um experimento é qualquer processo de observação.
- Em muitos experimentos, **observa-se um elemento de incerteza** e essencialmente a **impossibilidade** de predizer seu comportamento em futuras realizações.
- **objetivo:** Descrever ideias básicas desta teoria, que são fundamentais para o estudo da estatística.

Razões para a nossa falta de habilidade em predizer:

- Pode-se **não ter conhecimento** de todas as causas envolvidas;

Razões para a nossa falta de habilidade em predizer:

- Pode-se **não ter conhecimento** de todas as causas envolvidas;
- Pode-se não ter **dados suficientes** sobre as condições iniciais do experimento;

Razões para a nossa falta de habilidade em predizer:

- Pode-se **não ter conhecimento** de todas as causas envolvidas;

- Pode-se não ter **dados suficientes** sobre as condições iniciais do experimento;

- As **causas podem ser tão complexas** que o cálculo do seu efeito combinado não é possível;

Razões para a nossa falta de habilidade em predizer:

- Pode-se **não ter conhecimento** de todas as causas envolvidas;
- Pode-se não ter **dados suficientes** sobre as condições iniciais do experimento;
- As **causas podem ser tão complexas** que o cálculo do seu efeito combinado não é possível;
- **Existe alguma aleatoriedade** fundamental no experimento

Estaremos interessados em uma classe particular de experimentos, chamados **experimentos aleatórios**.

Um experimento aleatório é caracterizado por:

Um experimento aleatório é caracterizado por:

1. Repete-se, **sob as mesmas condições**, o experimento e seus resultados em diferentes realizações **podem ser diferentes**.

Um experimento aleatório é caracterizado por:

1. Repete-se, **sob as mesmas condições**, o experimento e seus resultados em diferentes realizações **podem ser diferentes**.
   - **Exemplo 1**: Jogar uma moeda diversas vezes com bastante cuidado para que cada jogada seja realizada da mesma maneira.

Um experimento aleatório é caracterizado por:

1. Repete-se, **sob as mesmas condições**, o experimento e seus resultados em diferentes realizações **podem ser diferentes**.
    - **Exemplo 1**: Jogar uma moeda diversas vezes com bastante cuidado para que cada jogada seja realizada da mesma maneira.

2. Apesar de não ser capaz de afirmar que resultado particular ocorre, pode-se descrever o **conjunto de todos os possíveis resultados** do experimento.

Um experimento aleatório é caracterizado por:

1. Repete-se, **sob as mesmas condições**, o experimento e seus resultados em diferentes realizações **podem ser diferentes**.
    - **Exemplo 1**: Jogar uma moeda diversas vezes com bastante cuidado para que cada jogada seja realizada da mesma maneira.

2. Apesar de não ser capaz de afirmar que resultado particular ocorre, pode-se descrever o **conjunto de todos os possíveis resultados** do experimento.

3. Ao repetir o experimento um grande número de vezes, pode-se observar uma certa **regularidade**. É esta regularidade que torna possível construir um **modelo probabilístico**.

Um experimento aleatório é caracterizado por:

1. Repete-se, **sob as mesmas condições**, o experimento e seus resultados em diferentes realizações **podem ser diferentes**.
    - **Exemplo 1**: Jogar uma moeda diversas vezes com bastante cuidado para que cada jogada seja realizada da mesma maneira.
2. Apesar de não ser capaz de afirmar que resultado particular ocorre, pode-se descrever o **conjunto de todos os possíveis resultados** do experimento.
3. Ao repetir o experimento um grande número de vezes, pode-se observar uma certa **regularidade**. É esta regularidade que torna possível construir um **modelo probabilístico**.
    - **Retornando ao exemplo 1**: É fato empírico conhecido que, depois de um grande número de jogadas, **a proporção de caras e de coroas serão aproximadamente iguais** (assumindo que a moeda é simétrica).

Exemplos de experimentos aleatórios

$\mathcal{E}_1$: Lançamento de uma moeda honesta;

$\mathcal{E}_2$: Lançamento de um dado honesto;

$\mathcal{E}_3$: Lançamento de duas moedas honestas;

$\mathcal{E}_4$: Selecionar um morador da cidade de Londrina e medir sua altura;

$\mathcal{E}_5$: Observar o tempo de falha de um componente mecânico;

$\mathcal{E}_6$: Número peças defeituosas num processo de fabricação

Espaço Amostral (Ω)

- O espaço amostral é o conjunto de todos os resultados possíveis associados a um experimento aleatório (ε), repesentado por Ω.
- Dos exemplos anteriores temos:

Espaço Amostral (Ω)

- O espaço amostral é o conjunto de todos os resultados possíveis associados a um experimento aleatório (ε), repesentado por Ω.

- Dos exemplos anteriores temos:

  ε_1: $\Omega_1 = \{C, K\}$, em que C = cara e K = coroa

Espaço Amostral (Ω)

- O espaço amostral é o conjunto de todos os resultados possíveis associados a um experimento aleatório (ε), repesentado por Ω.
- Dos exemplos anteriores temos:

ε_1: $\Omega_1 = \{C, K\}$, em que $C =$ cara e $K =$ coroa

ε_2: $\Omega_2 = \{1, 2, 3, 4, 5, 6\}$, em que $i = 1, \ldots, 6$ são as faces

Espaço Amostral (Ω)

- O espaço amostral é o conjunto de todos os resultados possíveis associados a um experimento aleatório (ε), repesentado por Ω.

- Dos exemplos anteriores temos:

  ε_1: $\Omega_1 = \{C, K\}$, em que $C =$ cara e $K =$ coroa

  ε_2: $\Omega_2 = \{1, 2, 3, 4, 5, 6\}$, em que $i = 1, \ldots, 6$ são as faces

  ε_3: $\Omega_3 = \{(C, C), (C, K), (K, C), (K, K)\}$;

Espaço Amostral (Ω)

- O espaço amostral é o conjunto de todos os resultados possíveis associados a um experimento aleatório (ε), repesentado por Ω.
- Dos exemplos anteriores temos:

ε_1: $\Omega_1 = \{C, K\}$, em que $C =$ cara e $K =$ coroa

ε_2: $\Omega_2 = \{1, 2, 3, 4, 5, 6\}$, em que $i = 1, \ldots, 6$ são as faces

ε_3: $\Omega_3 = \{(C, C), (C, K), (K, C), (K, K)\}$;

ε_4: $\Omega_4 = \{h \in \mathbb{R}, \ h > 0\}$

Espaço Amostral (Ω)

- O espaço amostral é o conjunto de todos os resultados possíveis associados a um experimento aleatório (ε), repesentado por Ω.

- Dos exemplos anteriores temos:

  ε_1: $\Omega_1 = \{C, K\}$, em que $C =$ cara e $K =$ coroa

  ε_2: $\Omega_2 = \{1, 2, 3, 4, 5, 6\}$, em que $i = 1, \ldots, 6$ são as faces

  ε_3: $\Omega_3 = \{(C, C), (C, K), (K, C), (K, K)\}$;

  ε_4: $\Omega_4 = \{h \in \mathbb{R},\ h > 0\}$

  ε_5: $\Omega_5 = \{t \in \mathbb{R},\ t > 0\}$

Espaço Amostral (Ω)

- O espaço amostral é o conjunto de todos os resultados possíveis associados a um experimento aleatório (ε), repesentado por Ω.

- Dos exemplos anteriores temos:

  ε_1: $\Omega_1 = \{C, K\}$, em que $C =$ cara e $K =$ coroa

  ε_2: $\Omega_2 = \{1, 2, 3, 4, 5, 6\}$, em que $i = 1, \ldots, 6$ são as faces

  ε_3: $\Omega_3 = \{(C, C), (C, K), (K, C), (K, K)\}$;

  ε_4: $\Omega_4 = \{h \in \mathbb{R},\ h > 0\}$

  ε_5: $\Omega_5 = \{t \in \mathbb{R},\ t > 0\}$

  ε_6: $\Omega_6 = \{0, 1, 2, 3, 4, \ldots\}$

Tipos de Espaço Amostral

- Podemos classificar espaços amostrais em dois tipos de acordo com os elementos que eles contém.

- Espaços amostrais podem ser **enumeráveis** ou **não enumeráveis**.

- Se os elementos do espaço amostral podem ser colocados em **uma correspondência 1-1** com um subconjunto dos inteiros, o espaço amostral é **enumerável**.

Para refletir ...

Em um nível filosófico, pode-se argumentar que **só existem espaços amostrais enumeráveis**, visto que **medidas não podem ser realizadas com infinita precisão**. Enquanto na prática isto é verdadeiro, métodos estatísticos e probabilísticos associados com espaços amostrais não enumeráveis são, em geral, menos complicados que aqueles para espaços amostrais enumeráveis, e proporcionam uma boa aproximação para a situação (enumerável) real.

Eventos e Coleção de Eventos

- Um **evento** é um subconjunto do espaço amostral, ou seja, é um conjunto de resultados possíveis do experimento aleatório.
- Se ao realizarmos um experimento aleatório, o resultado pertence a um dado evento A, dizemos que A **ocorreu**.
- Estaremos interessados no estudo da ocorrência de combinações de eventos. Para isso, pode-se utilizar as **operações Booleanas de conjuntos** (complementar, união, intersecção, diferença) para expressar eventos combinados de interesse.

Definição 1. Os eventos A e B são **disjuntos** ou **mutuamente exclusivos** se a interseção entre eles é vazia, isto é, $A \cap B = \emptyset$.

Exemplo 1: Sejam A, B, e C eventos em um mesmo espaço amostral Ω. Expresse os seguintes eventos em função de A, B, e C e operações Booleanas de conjuntos.

(a) Pelo menos um deles ocorre.

(b) Exatamente um deles ocorre.

(c) Apenas A ocorre.

(d) Pelo menos dois ocorrem.

(e) No máximo dois deles ocorrem.

(f) Nenhum deles ocorre.

(g) Ambos A e B ocorrem, mas C não ocorre.

(a) $(A \cup B \cup C)$

(a) $(A \cup B \cup C)$

(b) $(A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A^c \cap B^c \cap C)$

(a) $(A \cup B \cup C)$

(b) $(A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A^c \cap B^c \cap C)$

(c) $(A \cap B^c \cap C^c)$

(a) $(A \cup B \cup C)$

(b) $(A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A^c \cap B^c \cap C)$

(c) $(A \cap B^c \cap C^c)$

(d) $(A \cap B \cap C^c) \cup (A \cap B^c \cap C) \cup (A^c \cap B \cap C) \cup (A \cap B \cap C)$

(a) $(A \cup B \cup C)$

(b) $(A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A^c \cap B^c \cap C)$

(c) $(A \cap B^c \cap C^c)$

(d) $(A \cap B \cap C^c) \cup (A \cap B^c \cap C) \cup (A^c \cap B \cap C) \cup (A \cap B \cap C)$

(e) $(A \cap B \cap C)^c$

(a) $(A \cup B \cup C)$

(b) $(A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A^c \cap B^c \cap C)$

(c) $(A \cap B^c \cap C^c)$

(d) $(A \cap B \cap C^c) \cup (A \cap B^c \cap C) \cup (A^c \cap B \cap C) \cup (A \cap B \cap C)$

(e) $(A \cap B \cap C)^c$

(f) $(A^c \cap B^c \cap C^c)$

(a) $(A \cup B \cup C)$

(b) $(A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A^c \cap B^c \cap C)$

(c) $(A \cap B^c \cap C^c)$

(d) $(A \cap B \cap C^c) \cup (A \cap B^c \cap C) \cup (A^c \cap B \cap C) \cup (A \cap B \cap C)$

(e) $(A \cap B \cap C)^c$

(f) $(A^c \cap B^c \cap C^c)$

(g) $(A \cap B \cap C^c)$

Leis de Morgan

(i) $$\left(\bigcup_{i=1}^{\infty} A_i\right)^c = \bigcap_{i=1}^{\infty} A_i^c$$

(ii) $$\left(\bigcap_{i=1}^{\infty} A_i\right)^c = \bigcup_{i=1}^{\infty} A_i^c$$

Observa-se que (i) e (ii) são chamadas de Leis de Morgan.

(iii) $$A \cap \left(\bigcup_{i=1}^{\infty} B_i\right) = \bigcup_{i=1}^{\infty} (A \cap B_i)$$

(iv) $$A \cup \left(\bigcap_{i=1}^{\infty} B_i\right) = \bigcap_{k=1}^{\infty} (A \cup B_i)$$

Partição

Definição 2. Dado um espaço amostral Ω, **uma partição** $\Pi = \{A_i, i \in \mathbb{N}\}$ de Ω é uma **coleção de eventos** (subconjuntos de Ω) e satisfaz:

P1. $A_i \cap A_j = \emptyset$, $\forall\ i \neq j$.

P2. $\displaystyle\bigcup_{i \in \mathbb{N}} A_i = \Omega$.

Partição

Definição 2. Dado um espaço amostral Ω, **uma partição** $\Pi = \{A_i, i \in \mathbb{N}\}$ de Ω é uma **coleção de eventos** (subconjuntos de Ω) e satisfaz:

P1. $A_i \cap A_j = \emptyset, \ \forall \ i \neq j$.

P2. $\displaystyle\bigcup_{i \in \mathbb{N}} A_i = \Omega$.

Observação: Deste modo os eventos de uma partição são **mutuamente exclusivos** e **cobrem todo o espaço amostral**. Portanto, cada elemento $\omega \in \Omega$ pertence a um, e somente um, dos eventos A_i de uma partição.

Partição

Definição 2. Dado um espaço amostral Ω, **uma partição** $\Pi = \{A_i, i \in \mathbb{N}\}$ de Ω é uma **coleção de eventos** (subconjuntos de Ω) e satisfaz:

P1. $A_i \cap A_j = \emptyset, \ \forall \ i \neq j$.

P2. $\displaystyle\bigcup_{i \in \mathbb{N}} A_i = \Omega$.

Observação: Deste modo os eventos de uma partição são **mutuamente exclusivos** e **cobrem todo o espaço amostral**. Portanto, cada elemento $\omega \in \Omega$ pertence a um, e somente um, dos eventos A_i de uma partição.

Exemplo 2: Seja $\Omega = \{1, 2, 3, 4\}$ um espaço amostral e considere $A_1 = \{1, 2, 3\}$ e $A_2 = \{4\}$ eventos. A_1 e A_2 formam uma partição de Ω?

Partição

Definição 2. Dado um espaço amostral Ω, **uma partição** $\Pi = \{A_i, i \in \mathbb{N}\}$ de Ω é uma **coleção de eventos** (subconjuntos de Ω) e satisfaz:

P1. $A_i \cap A_j = \emptyset, \forall\ i \neq j$.

P2. $\bigcup\limits_{i \in \mathbb{N}} A_i = \Omega$.

Observação: Deste modo os eventos de uma partição são **mutuamente exclusivos** e **cobrem todo o espaço amostral**. Portanto, cada elemento $\omega \in \Omega$ pertence a um, e somente um, dos eventos A_i de uma partição.

Exemplo 2: Seja $\Omega = \{1, 2, 3, 4\}$ um espaço amostral e considere $A_1 = \{1, 2, 3\}$ e $A_2 = \{4\}$ eventos. A_1 e A_2 formam uma partição de Ω?

Exemplo 3: A coleção de intervalos $\{(n, n+1] : n \in Z\}$ é uma partição dos números reais $\mathbb{R}$?

Sigma-álgebra (ou σ-álgebra) de Eventos

Embora possa-se pensar que, para certo um Ω, necessariamente é de interesse analisar todos os seus subconjuntos (e isto eventualmente é verdadeiro), temos três razões para considerar alguns subconjuntos de Ω.

Sigma-álgebra (ou σ-álgebra) de Eventos

Embora possa-se pensar que, para certo um Ω, necessariamente é de interesse analisar todos os seus subconjuntos (e isto eventualmente é verdadeiro), temos três razões para considerar alguns subconjuntos de Ω.

1. O espaço amostral pode conter um grau de detalhamento superior ao que estamos interessados.

Sigma-álgebra (ou σ-álgebra) de Eventos

Embora possa-se pensar que, para certo um Ω, necessariamente é de interesse analisar todos os seus subconjuntos (e isto eventualmente é verdadeiro), temos três razões para considerar alguns subconjuntos de Ω.

1. O espaço amostral pode conter um grau de detalhamento superior ao que estamos interessados.

2. Pode-se associar cada evento A com uma probabilidade $P(A)$. Como essas probabilidades estão baseadas em algum conhecimento sobre a ocorrência do evento A, nosso conhecimento sobre P **pode não estender para todos os subconjuntos de** Ω.

Sigma-álgebra (ou σ-álgebra) de Eventos

Embora possa-se pensar que, para certo um Ω, necessariamente é de interesse analisar todos os seus subconjuntos (e isto eventualmente é verdadeiro), temos três razões para considerar alguns subconjuntos de Ω.

1. O espaço amostral pode conter um grau de detalhamento superior ao que estamos interessados.

2. Pode-se associar cada evento A com uma probabilidade $P(A)$. Como essas probabilidades estão baseadas em algum conhecimento sobre a ocorrência do evento A, nosso conhecimento sobre P **pode não estender para todos os subconjuntos de** Ω.

3. Existem condições em P pelos **axiomas de Kolmogorov**, que estudaremos adiante, que podem não permitir que P **seja definida em todos os subconjuntos de** Ω, em particular isto pode ocorrer quando Ω for não enumerável. (Resultado da **Teoria da Medida**).

Sigma-álgebra (ou σ-álgebra) de Eventos

Estamos interessados em uma **coleção $\mathcal{F}$ de subconjuntos de Ω** (note que $\mathcal{F}$ é um conjunto cujos elementos também são conjuntos!) que são eventos de interesse no que se refere ao experimento aleatório $\mathcal{E}$ e os quais **temos conhecimento sobre a sua probabilidade de ocorrência**. $\mathcal{F}$ é chamado de uma σ-**álgebra de eventos**.

Sigma-álgebra (ou σ-álgebra) de Eventos

Estamos interessados em uma **coleção $\mathcal{F}$ de subconjuntos de** Ω (note que $\mathcal{F}$ é um conjunto cujos elementos também são conjuntos!) que são eventos de interesse no que se refere ao experimento aleatório $\mathcal{E}$ e os quais **temos conhecimento sobre a sua probabilidade de ocorrência**. $\mathcal{F}$ é chamado de uma σ-**álgebra de eventos**.

Definição 3. Uma σ-álgebra de eventos $\mathcal{F}$ é uma coleção de subconjuntos de Ω que satisfaz:

1. $\Omega \in \mathcal{F}$;
2. Se $A \in \mathcal{F}$ então $A^c \in \mathcal{F}$;
3. Se $A_i \in \mathcal{F}$, com $i \in \mathbb{N}$, então $\displaystyle\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Exercícios

1. $\mathcal{F} = \{\emptyset, \Omega\}$ é σ-álgebra?

2. $\mathcal{F} = \{\emptyset, A, A^c, \Omega\}$ é σ-álgebra?

3. Considere $\Omega = \{1, 2, 3\}$ e as seguintes coleções:

   $\mathcal{F}_1 = \{\emptyset, \Omega, \{1\}, \{2, 3\}\}$ e $\mathcal{F}_2 = \{\emptyset, \Omega, \{1\}, \{2\}, \{1, 3\}, \{2, 3\}\}$.

   $\mathcal{F}_1$ e $\mathcal{F}_2$ são σ-álgebras? Justifique.

4. Sejam $\Omega = [0, 1] \cap \mathbb{R}$ um espaço amostral e $\mathcal{F}_1$ e $\mathcal{F}_2$ duas sigma-álgebras em Ω. Mostre que $\mathcal{F} = \{A : A \in \mathcal{F}_1 \cap \mathcal{F}_2\}$ é uma sigma-álgebra, mas $\mathcal{G} = \{A : A \in \mathcal{F}_1 \cup \mathcal{F}_2\}$ não é.

Exemplo: σ-álgebra de Borel

Seja $\Omega = \mathbb{R}$. A σ-álgebra de Borel, denotada por $\mathcal{B}(\mathbb{R})$, e se define como a **menor σ-álgebra** de subconjuntos de $\mathbb{R}$ que contém todos os intervalos da seguinte forma: $(-\infty, x]$, $x \in \mathbb{R}$.

Exemplo: σ-álgebra de Borel

Seja $\Omega = \mathbb{R}$. A σ-álgebra de Borel, denotada por $\mathcal{B}(\mathbb{R})$, e se define como a **menor σ-álgebra** de subconjuntos de $\mathbb{R}$ que contém todos os intervalos da seguinte forma: $(-\infty, x]$, $x \in \mathbb{R}$.

Observações:

- Pode-se mostrar que nem todos os subconjuntos de $\mathbb{R}$ estão em $\mathcal{B}(\mathbb{R})$.
- De maneira análoga, pode-se definir $\mathcal{B}(\mathbb{R}^2), \mathcal{B}(\mathbb{R}^3), \ldots, \mathcal{B}(\mathbb{R}^n)$
- Em geral, todas as questões de interesse na teoria da probabilidade se referem a σ-álgebra de Borel.

Pode-se classificar os fundamentos da probabilidade nas seguintes situações:

- O **grau de precisão**, isto é o que podemos esperar da probabilidade como representação dos fenômenos aleatórios.

- A **interpretação** que proporciona a base com a qual probabilidade deve ser determinada.

- A **estrutura matemática** formal de probabilidade dada por um conjunto de axiomas.

Observação:
A compreensão dos **fundamentos de probabilidade** é importante, pois esses fundamentos influeciam na escolha dos **métodos estatísticos** a serem utilizados (Clássicos/Frequentistas, Logicista, Bayesianos, etc).

Interpretação Clássica de Probabilidade

i) Cardano (1663), De Moivre (1718), Laplace (1812)

ii) A probabilidade é definida com base em dados do experimento aleatório. Além disso, baseia-se no conceito de **resultados equiprováveis**;

iii) Observa-se que todo evento de Ω tem uma probabilidade.

Interpretação Clássica de Probabilidade

i) Cardano (1663), De Moivre (1718), Laplace (1812)

ii) A probabilidade é definida com base em dados do experimento aleatório. Além disso, baseia-se no conceito de **resultados equiprováveis**;

iii) Observa-se que todo evento de Ω tem uma probabilidade.

Dado um espaço amostral Ω, a probabilidade de um evento A, tal que $A \in \mathcal{F} = \mathcal{P}(\Omega)$, que associa um valor númerico ao evento A é definida por

$$P(A) = \frac{\#A}{\#\Omega} = \frac{\text{n}^o \text{ de resultados favoráveis do evento A}}{\text{n}^o \text{ de resultados possíveis}}$$

Interpretação Frequentista de Probabilidade

i) John Venn (1866), Von Mises (1928)

ii) Pode-se **repetir** o experimento aleatório n vezes;

iii) E assim contar quantas vezes o evento A ocorre, $n(A)$ (número de ocorrências de A).

Dessa forma a **frequência relativa de** A nas n repetições é dada por

$$f_n(A) = \frac{\mathrm{n(A)}}{\mathrm{n}}.$$

Para $n \to \infty$ repetições sucessivas e independentes, a frequência relativa de A tende para uma constante p

$$\lim_{n\to\infty} f_n(A) = \lim_{n\to\infty} \frac{\mathrm{n(A)}}{\mathrm{n}} = P(A) = p$$

Exemplo: Se um dado fosse lançado **10** vezes, e contássemos quantas vezes saiu a face 4, qual seria a probabilidade desse evento?

```
## Tamanho da amostra
n <- 10

## Objeto para armazenar os resultados
x <- numeric(n)

## Estrutura de repetição
# Repetir n vezes

for(i in 1:n){
# Amostra aleatória de tamanho 1 dos números 1 a 6
x[i] <- sample(1:6, size = 1)
}

## Total de valores igual a 4 => n(A)
sum(x == 4)
[1] 3

## Proporção de valores igual a 4 => n(A)/n
sum(x == 4)/length(x)
[1] 0.3
```

Exemplo: Se um dado fosse lançado **100** vezes, e contássemos quantas vezes saiu a face 4, qual seria a probabilidade desse evento?

```
## Tamanho da amostra
n <- 100

## Objeto para armazenar os resultados
x <- numeric(n)

## Estrutura de repetição
# Repetir n vezes

for(i in 1:n){
# Amostra aleatória de tamanho 1 dos números 1 a 6
x[i] <- sample(1:6, size = 1)
}

## Total de valores igual a 4 => n(A)
sum(x == 4)
[1] 13

## Proporção de valores igual a 4 => n(A)/n
sum(x == 4)/length(x)
[1] 0.13
```

Exemplo: Se um dado fosse lançado **1000** vezes, e contássemos quantas vezes saiu a face 4, qual seria a probabilidade desse evento?

```
## Tamanho da amostra
n <- 1000

## Objeto para armazenar os resultados
x <- numeric(n)

## Estrutura de repetição
# Repetir n vezes

for(i in 1:n){
# Amostra aleatória de tamanho 1 dos números 1 a 6
x[i] <- sample(1:6, size = 1)
}

## Total de valores igual a 4 => n(A)
sum(x == 4)
[1] 146

## Proporção de valores igual a 4 => n(A)/n
sum(x == 4)/length(x)
[1] 0.146
```

Exemplo: Se um dado fosse lançado **10000** vezes, e contássemos quantas vezes saiu a face 4, qual seria a probabilidade desse evento?

```
## Tamanho da amostra
n <- 10000

## Objeto para armazenar os resultados
x <- numeric(n)

## Estrutura de repetição
# Repetir n vezes

for(i in 1:n){
# Amostra aleatória de tamanho 1 dos números 1 a 6
x[i] <- sample(1:6, size = 1)
}

## Total de valores igual a 4 => n(A)
sum(x == 4)
[1] 1586

## Proporção de valores igual a 4 => n(A)/n
sum(x == 4)/length(x)
[1] 0.1586
```

Exemplo: Se um dado fosse lançado **100000** vezes, e contássemos quantas vezes saiu a face 4, qual seria a probabilidade desse evento?

```
## Tamanho da amostra
n <- 100000

## Objeto para armazenar os resultados
x <- numeric(n)

## Estrutura de repetição
# Repetir n vezes

for(i in 1:n){
# Amostra aleatória de tamanho 1 dos números 1 a 6
x[i] <- sample(1:6, size = 1)
}

## Total de valores igual a 4 => n(A)
sum(x == 4)
[1] 16616

## Proporção de valores igual a 4 => n(A)/n
sum(x == 4)/length(x)
[1] 0.16616
```

Exemplo: Se um dado fosse lançado **1000000** vezes, e contássemos quantas vezes saiu a face 4, qual seria a probabilidade desse evento?

```
## Tamanho da amostra
n <- 1000000

## Objeto para armazenar os resultados
x <- numeric(n)

## Estrutura de repetição
# Repetir n vezes

for(i in 1:n){
# Amostra aleatória de tamanho 1 dos números 1 a 6
x[i] <- sample(1:6, size = 1)
}

## Total de valores igual a 4 => n(A)
sum(x == 4)
[1] 166911

## Proporção de valores igual a 4 => n(A)/n
sum(x == 4)/length(x)
[1] 0.16691
```

Assim,

$$\lim_{n\to\infty} f_n(A) = \lim_{n\to\infty} \frac{\mathrm{n(A)}}{\mathrm{n}} = P(A) \approx 0,1667$$

As probabilidades calculadas a partir de **frequências relativas**, são **estimativas** da verdadeira probabilidad.e

Observação

Nós prosseguiremos como se existisse alguma **base empírica** (ou metafísica) que garanta que $f_n(A) \to P(A)$, embora que o sentido de convergência quando n cresce só será explicado pela **Lei dos Grandes Números**. Esta tendência da frequência relativa de estabilizar em um certo valor é conhecida como **regularidade estatística**.

Interpretação de Probabilidade

1. **Clássica:** Baseada em uma enumeração de casos *igualmente prováveis*.

Interpretação de Probabilidade

i) **Clássica:** Baseada em uma enumeração de casos *igualmente prováveis*.

ii) **Frequentista:** Se refere ao limite da frequência relativa de ocorrência do evento A *em repetidas realizações independentes* do experimento aleatório ε.

Interpretação de Probabilidade

i) **Clássica:** Baseada em uma enumeração de casos *igualmente prováveis*.

ii) **Frequentista:** Se refere ao limite da frequência relativa de ocorrência do evento A *em repetidas realizações independentes* do experimento aleatório ε.

iii) **Subjetiva:** Se refere ao grau de crença pessoal na ocorrência do evento A e é medida por meio da interpretação comportamental de disposição a apostar ou agir.

Interpretação de Probabilidade

i) **Clássica:** Baseada em uma enumeração de casos *igualmente prováveis*.

ii) **Frequentista:** Se refere ao limite da frequência relativa de ocorrência do evento A *em repetidas realizações independentes* do experimento aleatório ε.

iii) **Subjetiva:** Se refere ao grau de crença pessoal na ocorrência do evento A e é medida por meio da interpretação comportamental de disposição a apostar ou agir.

iv) **Lógica:** É o grau de confirmação da hipótese de uma proposição que "A ocorre" dada uma evidência por meio da proposição que "B ocorreu".

Interpretação de Probabilidade

1. **Clássica:** Baseada em uma enumeração de casos *igualmente prováveis*.

2. **Frequentista:** Se refere ao limite da frequência relativa de ocorrência do evento A *em repetidas realizações independentes* do experimento aleatório ε.

3. **Subjetiva:** Se refere ao grau de crença pessoal na ocorrência do evento A e é medida por meio da interpretação comportamental de disposição a apostar ou agir.

4. **Lógica:** É o grau de confirmação da hipótese de uma proposição que "A ocorre" dada uma evidência por meio da proposição que "B ocorreu".

5. **Geométrica:** É uma interpretação relacionada com quantidades geométricas, isto é, dois eventos A e B possuem a mesma probabilidade se, e somente se, eles têm a mesma área.

- Os axiomas que descreveremos a seguir não descrevem um único modelo probabilístico, eles apenas determinam uma família de **modelos probabilísticos**.

- Os axiomas que descreveremos a seguir não descrevem um único modelo probabilístico, eles apenas determinam uma família de **modelos probabilísticos**.

- A escolha de um **modelo específico satisfazendo os axiomas** é realizado por meio do **fenômeno aleatório** que está sendo estudado.

Axiomas de Kolmogorov

- Os axiomas que descreveremos a seguir não descrevem um único modelo probabilístico, eles apenas determinam uma família de **modelos probabilísticos**.

- A escolha de um **modelo específico satisfazendo os axiomas** é realizado por meio do **fenômeno aleatório** que está sendo estudado.

- Motivado pelas propriedades de frequência relativa, **Andrei Kolmogorov** propõe, em 1933, **a axiomatização do conceito de probabilidade** por meio de três axiomas.

Definição 4. Uma função P, definida na σ-álgebra $\mathcal{F}$ de subconjuntos de Ω e com valores em $[0,1]$, $P : \mathcal{F} \rightarrow [0,1]$ é uma **medida de probabilidade** se satisfaz os axiomas de Kolmogorov:

1. **Normalização Unitária**. $P(\Omega) = 1$;
2. **Não-negatividade**. $\forall\ A \in \mathcal{F}$, $P(A) \geq 0$;
3. σ-**aditividade**. $\forall\ A_1, A_2, A_3, \ldots \in \mathcal{F}$ então $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$, sendo disjuntos dois a dois.

Observações:

1. Motivado pelas propriedades de frequência relativa (**interpretação frequentista de probabilidade**), Andrei Kolmogorov por volta de 1933, propôs os axiomas acima.

2. Os axiomas de Kolmogorov não descrevem um único modelo probabilístico, eles apenas **determinam uma família de modelos probabilísticos**, com os quais pode-se utilizar métodos matemáticos para estudar propriedades **que serão verdadeiras em qualquer modelo probabilístico**.

3. Assim, a escolha de um modelo específico **satisfazendo os axiomas de Kolmogorov** é feito pelo probabilista/estatístico familiar com o fenômeno aleatório que está sendo modelado.

Espaço de Probabilidade

Definição 5. Seja ε um **experimento aleatório** associado ao **espaço amostral** Ω. Suponhamos que $\mathcal{F}$ é uma σ-**álgebra de subconjuntos de** Ω onde se pode definir uma medida de probabilidade P, isto é, $P : \mathcal{F} \to [0, 1]$. Então, a trinca $(\Omega, \mathcal{F}, P)$ chama-se **espaço de probabilidade** e representa um **modelo probabilístico**.

Definição 5. Seja ε um **experimento aleatório** associado ao **espaço amostral** Ω. Suponhamos que $\mathcal{F}$ é uma σ-**álgebra de subconjuntos de** Ω onde se pode definir uma medida de probabilidade P, isto é, $P : \mathcal{F} \rightarrow [0, 1]$. Então, a trinca $(\Omega, \mathcal{F}, P)$ chama-se **espaço de probabilidade** e representa um **modelo probabilístico**.

Observações:

1. Intuitivamente quando se propõe um modelo para um problema por meio de probabilidade, basicamente, o que se faz é especificar cada uma das componentes da trinca acima.

2. Eventos são os elementos de $\mathcal{F}$, aos quais se pode atribuir probabilidade. Assim, probabilidade é uma função cujo argumento é um conjunto.

3. Portanto, não somente conjuntos, como também as operações sobre eles, têm uma importância fundamental em teoria da probabilidade.

Exemplo 1. Se Ω for um conjunto finito, então temos que a probabilidade clássica que assume que todos os resultados são igualmente prováveis, é um exemplo de uma medida de probabilidade. Neste caso, temos que

$$P(A) = \frac{||A||}{||\Omega||},$$

definido para todo A subconjunto e Ω. O fato de $0 \leq ||A|| \leq ||\Omega||$ e que

$$||A \cup B|| = ||A|| + ||B|| - ||A \cap B||,$$

permitem que P satisfaça os axiomas de Kolmogorov.

Exemplo 2. Seja $\Omega = \{\omega_1, \omega_1, \ldots, \omega_n\}$ um conjunto finito e considere $P(\{\omega_i\}) = p_i$ em que $i \geq 1$. Além disso, vamos admitir que $\sum_{i=1}^{n} p_i = 1$. A probabilidade de um evento A é dada por

$$P(A) = \sum_{\omega_i \in A} P(\{\omega_i\}).$$

Pode-se notar que P é uma medida de probabilidade, verificando os axiomas de Kolmogorov.

Exemplo 3. Seja $\Omega = [0, 1]$ e $\mathcal{B}_0$ a σ-álgebra de Borel restrita a eventos contidos em $[0, 1]$. Pode-se mostrar que existe uma medida de probabilidade μ em $(\Omega, \mathcal{B}_0)$ tal que para todo intervalo I em $[0, 1]$, $\mu(I)$ é igual ao comprimento I. Esta medida de probabilidade μ é conhecida como **medida de Lebesgue**.

Propriedades de uma Medida de Probabilidade

Teorema 1: Seja A um evento de Ω qualquer e P uma medida de probabilidade. Então,

P1. $P(A^c) = 1 - P(A)$;

P2. $P(\emptyset) = 0$;

P3. $P(A) \leq 1$.

Teorema 2: Monotonicidade. Se $A \subseteq B$ então $P(A) \leq P(B)$.

Teorema 3: Uma expressão exata para a probabilidade de uma união não-disjunta é dada por

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

Teorema 4: Probabilidade de Partições. Seja $\{A_i\}$ uma partição enumerável de Ω, de conjuntos em $\mathcal{F}$. Então, para todo $B \in \mathcal{F}$

$$P(B) = \sum_i P(B \cap A_i).$$

Teorema 5: Desigualdade de Boole. Para n eventos arbitrários $\{A_1, \ldots, A_n\}$, a desigualdade de Boole é representada por

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i).$$

Exercício. Mostre que se $P(A_i) = 1$ então $P\left(\bigcap_{i=1}^{n} A_i\right) = 1$, para $i = 1, 2, 3 \ldots$.

Probabilidade Condicional

- Como vimos anteriormente, existem muitas interpretações do conceito de probabilidade.

- Pode-se interpretar **probabilidade de um evento A como um limite das frequências relativas** de ocorrência do evento A em realizações independentes de um experimento.

- Em ambos os casos, probabilidade é baseada em informação e conhecimento. Em particular, **o conhecimento sobre um determinado evento que ocorreu pode influenciar na probabilidade** dos demais eventos.

Interpretação frequentista: Qual a probabilidade de um evento A, sabendo-se que um evento B ocorreu?

- Vamos realizar um experimento n vezes das quais o evento A ocorre N_A vezes.
- Observa-se que $r_A = \dfrac{N_A}{n}$ é a frequência relativa do evento A nestas n realizações do experimento.
- Assim, a probabilidade condicional do evento A dado que B ocorreu é igual ao limite das frequências relativas condicionais do evento A dado B, isto é,

$$P(A \text{ sabendo-se que } B \text{ ocorreu}) = \lim_{n\to\infty} \frac{N_{A\cap B}}{N_B} = \lim_{n\to\infty} \frac{r_{A\cap B}}{r_B} = \frac{P(A\cap B)}{P(B)}.$$

Probabilidade Condicional

Definição 6. Seja $(\Omega, \mathcal{F}, P)$ um espaço de probabilidade. Considere $A, B \in \mathcal{F}$ e $P(B) > 0$, a **probabilidade condicional de A dado B** é definida por

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

A probabilidade condicional também satisfaz as seguintes propriedades:

- $P(B|B) = 1$;
- $P(A|B) = P(A \cap B|B)$;
- Se $B \subseteq A$ então $P(A|B) = 1$;
- $P(A \cap B|C) = P(A|B \cap C) \cdot P(B|C)$.

Utilizando indução matemática, pode-se mostrar que

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1) \cdot P(A_2|A_1) \ldots P(A_n|A_1 \cap \ldots \cap A_{n-1}).$$