# Probabilidade

## Síntese

- **Conjuntos**
    - **Relação de Pertinência $\in$**
    - **Classificação de Função**
    - **Cardinalidade**
    - **Enumerando os Racionais**
    - **Conjunto Vazio e Conjunto Universo**
    - **Relação de Inclusão**
        - $\subset$ **Contido**
        - $\not\subset$ **Não contido**
        - $\supset$ **Contém**
        - $\not\supset$ **Não Contém**
        - **Propriedades da Inclusão**
    - **Operações com Conjuntos**
        - $A^c$ **Complementação**
        - $(A^c)^c$ **Idempotência:**
        - $\cup$ **União**
        - $\cap$ **Intersecção**
        - **Comutatividade (Simetria)**
        - $-$ **Diferença**
    - **Exemplo e Propriedades das Operações**
    - **Uniões e Intersecções de Vários Conjuntos**
    - **Produto Cartesiano**
    - **Conjunto das Partes**
        - **Teorema de Cantor**
    - **Partição**
    - **Função Indicadora**
- **Técnicas de Contagem**

- Introdução à Probabilidade
  - Experimento Aleatório
    - Espaço Amostral
    - Evento
      - Álgebra de Eventos
      - Sigma Álgebra
    - Probabilidade
      - Paradoxo De Ellsbergue
      - Aversão a Ambiguidade
      - Frequência Relativa
      - Propriedades de Frequência Relativa
      - Axiomas de Kolmogorov
      - Espaço de Probabilidade
    - Linguagem Probabilística
    - Propriedades de uma Medida de Probabilidade
      - Propriedade de Monotonicidade
      - Propriedade da União de Conjuntos Não-Disjuntos
      - Propriedade da Probabilidade de Partições
      - Propriedade da Desigualdade de Boole
      - Propriedade da Inclusão-Exclusão


---

## Conjuntos

▼

- Um conjunto é uma coleção de elementos distintos onde os elementos não são ordenados.
- Formalmente na teoria dos conjuntos os elementos de um conjunto devem ser distintos.
- Um conjunto pode ser especificado, listando seus elementos dentro de chaves $\{\}$. Por exemplo:

$$A = \{0, 1, 2, 3, 5, 8, 13\},\ B = \{0, 1, 2, \ldots, 1000\}.$$

- Como em um conjunto a ordem dos elementos não importa, tem-se que:

$$\{1, 2, 3\} = \{2, 3, 1\}$$

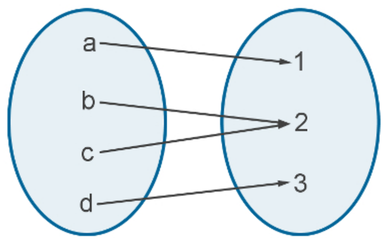
▼ **Relação de Pertinência** $\in$

- Se um dado elemento faz parte de um conjunto, diz-se que ele **pertence** ao conjunto e denota-se isso com símbolo $\in$.
- Por outro lado, se um dado elemento não faz parte de um conjunto, diz-se que ele **não pertence** ao conjunto e denota-se isso com o símbolo $\notin$.
- **Obs:** É preciso ter cuidado ao distinguir entre um elemento, como $2$ e o conjunto contendo somente este elemento $\{2\}$ .

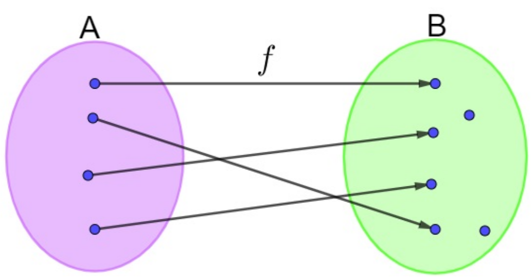
  **Exemplo:** $2 \in F = \{2, 3, 5\}$, nesse exemplo o $2$ pertence ao conjunto $F$, enquanto no exemplo a seguir o conjunto $\{2\}$ não pertence ao conjunto $G$, $\quad \{2\} \notin G = \{2, 3, 5\}$.

## ▼ Classificação de Função

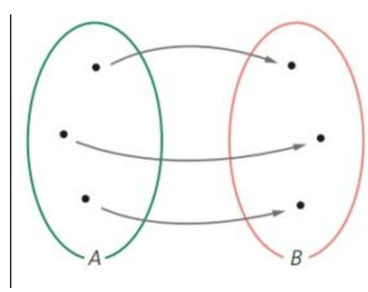
- **Sobrejetora:** É quando a imagem da função é igual ao contradomínio.




- **Injetora:** Se $a \neq b$ então $f(a) \neq f(b)$. Ou seja o domínio tem uma imagem exclusiva para ele.




- **Bijetora:** A função tem que ser **sobrejetora** e **injetora** ao mesmo tempo.




## ▼ Cardinalidade

- O tamanho de um conjunto $A$, $||A||$, é a quantidade de elementos que ele possui, a qual é chamada de sua **cardinalidade**.
- A cardinalidades pode ser:
  - **Finita:** Um conjunto é finito quando existe uma função **bijetiva** cujo domínio é igual a este conjunto e a imagem é o conjunto dos inteiros não-negativos menores que um número finito.
  - **Infinita Enumerável:** Um conjunto infinito enumerável tem exatamente a mesma quantidade de elementos que os naturais, ou seja, existe uma função **bijetiva** cujo domínio é igual a este conjunto e a imagem é igual ao conjunto dos naturais.
  - **Obs:** Um conjunto é **enumerável**, se ele for finito ou infinito enumerável.
  - **Infinita não-Enumerável:** Um conjunto é não-enumerável se ele não for enumerável.

### Enumerando os Racionais

- Um número $x$ é racional se pode ser escrito sob a forma $\frac{q}{p}$ , onde $p$ e $q$ são inteiros e $q \neq 0$. A seguinte matriz contém todos os racionais não-negativos:

$$
\begin{array}{cccccc}
0/1 & & 0/2 & & 0/3 & \cdots \\
& \swarrow & & \swarrow & & \\
1/1 & & 1/2 & & 1/3 & \cdots \\
& \swarrow & & \swarrow & & \\
2/1 & & 2/2 & & 2/3 & \cdots \\
& \swarrow & & \swarrow & & \\
3/1 & & 3/2 & & 3/3 & \cdots \\
& \swarrow & & \swarrow & & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}
$$

  **Obs:** Essa matriz é feita com o método **Diagonalização de Cantor**.

- Utilizando o método da diagonalização, os elementos da matriz são ordenados, sem repetição, da seguinte forma:

$$0/1, 1/1, 1/2, 2/1, 1/3, 3/1, ...$$

- Logo, para enumerar todos os racionais, basta alternar os positivos e os negativos na seguinte ordem:

$$0/1, +1/1, -1/1, +1/2, -1/2, +2/1, -2/1, +1/3, -1/3, +3/1, -3/1, \ldots$$

### Conjunto Vazio e Conjunto Universo

- O conjunto que contém todos os elementos que queremos considerar é chamado de **Conjunto Universo** e é denotado por $\Omega$.

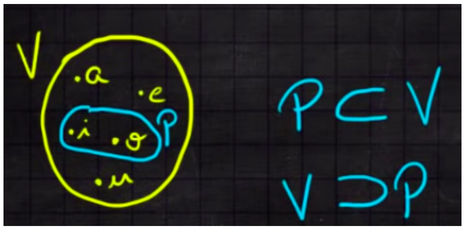
- Por outro lado, o conjunto especial que não possui elementos é chamado de **Conjunto Vazio** e é denotado por $\emptyset$. Este conjunto tem cardinalidade $0$ e portanto é finito. Por exemplo:

$$\emptyset = \{\} = \{x : x \in \mathbb{R} \text{ e } x < x\} \text{ ou } \emptyset = (a, a).$$

- **Obs:** Todo conjunto possuir (contem) um subconjunto vazio: $A \subset \emptyset = True$, mais nem todo conjunto possuir o elemento vazio: $A \in \emptyset$.

▼ **Relação de Inclusão**

- $\subset$ **Contido:** É quando todos os elementos de um conjunto $A$ estiver dentro de outro conjunto $B$. Ou seja $A \subset B$. E lida como $A$ esta contido em $B$ ou $A$ é um subconjunto de $B$.
- $\not\subset$ **Não Contido:** Um conjunto $A$ não esta contido em um outro conjunto $B$ se os elementos de $A$ não estiverem em $B$.
- $\supset$ **Contém:** É quando um conjunto grande $B$ contém (tem) um conjunto menor $A$ dentro dele. Ou seja $B \supset A$.



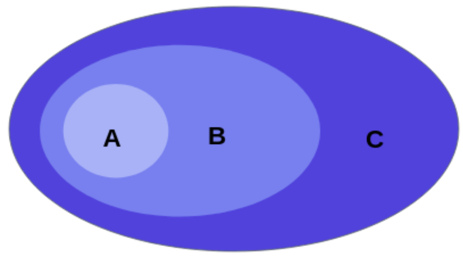

- $\not\supset$ **Não Contém:** É quando um conjunto $B$ não tem um outro conjunto $A$ dentro dele.
- **Obs:** Podemos prova que dois conjuntos são iguais se cada conjunto estiver contido no outro conjunto: $A \subset B$ e $B \subset A$.

**Exemplo:** $A = \{1, 2, 3\}$ $B = \{\{1, 2\}1, 2, 3\}$ $C = \{1, 2\}$

- $1 \in A$ (V)
- $7 \in A$ (F)
- $4 \notin A$ (V)
- $\{1, 2\} \in B$ (V)
- $\{1, 2\} \in A$ (F) O conjunto $\{1, 2\}$ não é elemento de $A$.
- $A \subset B$ (V)
- $B \subset A$ (F)
- $C \subset B$ (F)
- $C \not\subset B$ (V)
- $A \subset A$ (V)
- $\#(A \cup C) = \#A + \#B$ (F) O $\#(A \cup C) = \{1, 2, 3\} = 3$ e $\#A + \#B = \{1, 2, 3, 1, 2\} = 5$
- $\#(A \cup C) = \#A + \#B - \#(A \cap B)$ (V)

- **Propriedades da Inclusão**
  - **Reflexividade:** Um conjunto esta contido nele mesmo: $A \subset A$.

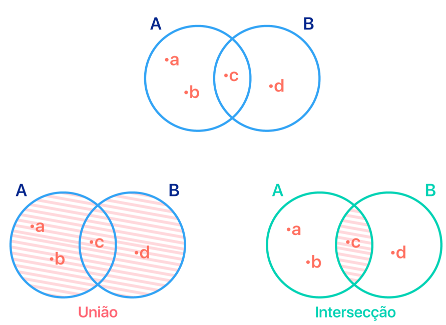
- **Transitividade:** Se um conjunto $A$ esta contido em $B$, e $B$ esta contido em $C$, isso implica que $A$ também esta contido em $C$. $A \subset B, B \subset C \rightarrow A \subset C$.




- **Anti-simetria:** Se $A$ esta contido em $B$, e $B$ esta contido em A, então $A$ é igual a $B$. Ou seja: $A \subset B, B \subset A, A = B$

**Operações com Conjuntos**

- $A^c$ **Complementação:** O complementa de um conjunto $A$ é todo o $\Omega$ exceto os elementos de $A$: $A^c = \{w \in \Omega : w \notin A\}$. Observe que de acordo com esta definição, para todo $w \in \Omega$ e todo conjunto $A$ , não existe outra opção além de $w \in A$ ou $w \in A^c$ , além disso não pode ser verdade que $w \in A$ e $w \in A^c$simultaneamente.
- **Obs:** Se houver dois conjuntos $A$ e $B$, o $A^c$ pode ser representado por $B - A = A^c$, pois a diferença de $B$ e $A$, é todos os elementos de $B$ que não estão em $A$.
- $(A^c)^c$ **Idempotência:** $(A^c)^c = A$, o complementar do complementar de um conjunto é igual ao próprio conjunto.
- $\cup$ **União:** $A \cup B = \{w : w \in A \textbf{ ou } w \in B\}$, ou seja a união é basicamente criar um novo conjunto com todos os elementos de $A$ e de $B$, sem repetir os elementos que estão tando no A quanto no $B$.
- $\cap$ **Intersecção:** $A \cap B = \{\omega : \omega \in A \textbf{ e } \omega \in B\}$, ou seja a união é um novo conjunto composto só pelos os elementos que estão tando no $A$ quanto no $B$.




- **Comutatividade (Simetria):** $A \cup B = B \cup A$ e $A \cap B = B \cap A$.

**Prova:** Suponha que $\omega \in A \cup B$. Então, $\omega \in A$, ou $\omega \in B$, o que por sua vez implica que $\omega \in B \cup A$, ou seja, $A \cup B \subseteq B \cup A$. Agora suponha que $\omega \in B \cup A$. Então, $\omega \in B$, ou $\omega \in A$, o que por sua vez implica que $\omega \in A \cup B$, ou seja, $B \cup A \subseteq A \cup B$. Portanto, $A \cup B = B \cup A$.

- $-$ **Diferença:** $A - B = A \cap B^c = \{w : w \in A \text{ e } w \notin B\}$, ou seja a diferença de $A$ e $B$ é todos os elementos de $A$ que não estão em $B$.

▼ **Exemplo 1:**

Considere os conjuntos

$$A = \{w \in \Omega : \text{a soma dos pontos sobre as duas faces é um número par}\},$$

$$B = \{w \in \Omega : 1 \text{ aparece em pelo menos um dos dados}\}.$$

Descreva os conjuntos:

a. $A - B$:

$$A - B = \{(x, y) \in \Omega : x + y \text{ é par}, x \neq 1 \text{ e } y \neq 1\}$$

b. $B - A$:

$$B - A = \{(x, y) \in \Omega : (x = 1 \text{ ou } y = 1) \text{E } x + y \text{ é ímpar}\}$$

▼ **Exemplo 2:**

Sejam $A$ , $B$ e $C$ eventos arbitrários. Mostre que:

a. $A \subset B \Leftrightarrow A \cap B = A$

Assuma que $A \subset B$. Por definição da interseção, se $w \in A \cap B$, então $w \in A$. Seja $w \in A$. Como $A \subset B$, temos $w \in B$. Logo, $w \in A \cap B$. Portanto, $A \cap B = A$.

Assuma que $A \cap B = A$. Seja $w \in A$. Como $A = A \cap B$, temos que $w \in A \cap B$. Logo $w \in B$. Portanto, $A \subset B$.

Existe outra maneira de resolver esse problema, que é pela função indicadora:

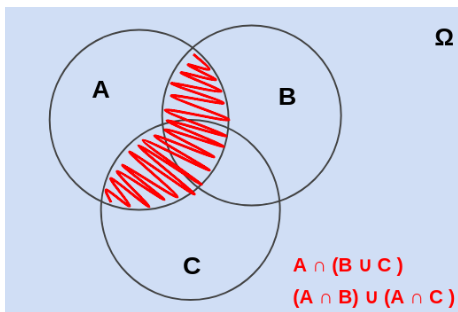
$$A \subset B \Leftrightarrow I_A \leq I_B \Leftrightarrow min(I_A, I_B) = I_A \Leftrightarrow I_{A \cap B} = I_A$$
$$\Leftrightarrow A \cap B = A$$

- **Obs:** Se $A \cap B = \emptyset$ , então $A$ e $B$ não tem nenhum elemento em comum, e nós dizemos que $A$ e $B$ são disjuntos.

| Operações entre | | | |
|---|---|---|---|
| **Conjuntos** | | **Função Indicadora** | |
| $A^c$ | V | $I_A \cdot I_B$ | V |
| $A \cap B$ | V | $I_A + I_B$ | V |
| $A \cup B$ | V | $I_A - I_B$ | V |
| A-B | V | $(I_{A^c})$ | V |
| $A \cdot B$ | F | $(I_A)^c$ | F |
| A+B | F | $I_A \cup I_B$ | F |
| | | $I_A \cap I_B$ | F |
| | | $I_{A \cup B}$ | V |

▼ **Exemplo e Propriedades das Operações**

- **Associatividade:** $A \cup (B \cup C) = (A \cup B) \cup C$ e $A \cap (B \cap C) = (A \cap B) \cap C$.
- **Distributividade:** $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ e $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$




- **Leis de De Morgan:** $(A \cup B)^c = A^c \cap B^c$ e $(A \cap B)^c = A^c \cup B^c$.

  **Prova:** Suponha que $\omega \in (A \cup B)^c$. Então, $\omega \notin (A \cup B)$, o que por sua vez implica que $\omega \notin A$ e $\omega \notin B$. Logo, $\omega \in A^c$ e $\omega \in B^c$, ou seja, $\omega \in (A^c \cap B^c)$. Então, $(A \cup B)^c \subseteq (A^c \cap B^c)$. Agora suponha que $\omega \in (A^c \cap B^c)$. Então, $\omega \in A^c$ e $\omega \in B^c$, o que por sua vez implica que $\omega \notin A$ e $\omega \notin B$. Logo, $\omega \notin (A \cup B)$, ou seja, $\omega \in (A \cup B)^c$. Então, $(A^c \cap B^c) \subseteq (A \cup b)^c$. Portanto, $(A^c \cap B^c) = (A \cup b)^c$.

- **Exemplo:** Sejam $A$, $B$, $C$ e $D$ subconjuntos do conjunto universo $\Omega$ tal que $A \cup B = \Omega$ , $C \cap D = \emptyset$ , $A \subset C$ e $B \subset D$ . Prove que $A = C$ e $B = D$.

**Solução:** Basta provar que $C \subseteq A$ e $D \subseteq B$. Seja $\omega \in C$, então como $C \cap D = \emptyset$, temos que $\omega \notin D$. Logo, como $B \subseteq D$, segue que $\omega \notin B$. Mas como $A \cup B = \Omega$, temos que $\omega \in A$. Portanto, $C \subseteq A$.
Para provar que $D \subseteq B$, seja $\omega \in D$, então como $C \cap D = \emptyset$, temos que $\omega \notin C$. Logo, como $A \subseteq C$, segue que $\omega \notin A$. Mas como $A \cup B = \Omega$, temos que $\omega \in B$. Portanto, $D \subseteq B$.

### Uniões e Intersecções de Vários Conjuntos

- Se temos uma coleção $\{A_\alpha : \alpha \in I\}$ de subconjuntos de $\Omega$ indexados pelo conjunto de índices $I$ , então:

$$\cup_{\alpha \in \mathcal{I}} A_\alpha = \{\omega : (\exists \alpha \in \mathcal{I}, \omega \in A_\alpha)\} \text{ e } \cap_{\alpha \in \mathcal{I}} A_\alpha = \{\omega : (\forall \alpha \in \mathcal{I}, \omega \in A_\alpha)\}$$

- Por exemplo, se $\Omega = \{0, 1, 2, ..., I\}$ é o conjunto de inteiros positivos divisíveis por $3$ e $A_\alpha = N_\alpha = \{0, 1, 2, ..., \alpha{-}1\}$, então

$$\cup_{\alpha \in \mathcal{I}} N_\alpha = \Omega \text{ e } \cap_{\alpha \in \mathcal{I}} N_\alpha = N_3.$$

- **Exemplo 1:** Dado os conjuntos $A_1, A_2, A_3, A_4$, determine a união desses conjuntos.

  A formula para representa a união desses quatros conjuntos é:

$$\cup_{\alpha = I} A_\alpha \qquad I = \{1, 2, 3, 4\}$$

  Essa formula acima é o mesmo que:

$$\cup_{i=1}^{4} A_i$$

  Também pode ser escrita como:

$$A_1 \cup A_2 \cup A_3 \cup A_4$$

  Todas essas formulas é definida como

$$\{w \in \Omega : \exists \alpha \in I, w \in A_\alpha\}$$

  Essa formula acima pode ser rescrita de uma maneira mais simples:

$$\{w \in \Omega : w \in A_1 \textbf{ ou } w \in A_2 \textbf{ ou } w \in A_3 \textbf{ ou } w \in A_4 \}$$

- **Exemplo 2:** Dado o conjunto dos números reais: $A_n = [0, 1 + \frac{1}{n}], n = \{1, 2, 3, \cdots\}$, defina quem é a $\cup_{n=1}^{\infty} A_n$ e quem é $\cap_{n=1}^{\infty} A_n$.

$$\cup_{n=1}^{\infty} A_n = [0,2]$$
$$\cap_{n=1}^{\infty} A_n = [0,1]$$

### ▼ Produto Cartesiano

- O produto Cartesiano $A \times B$ de dois conjuntos dados $A$ e $B$ é o conjunto de todos os pares ordenados de elementos, onde o primeiro pertence à $A$ e o segundo pertence à $B$:

$$A \times B = \{(a,b) : a \in A, b \in B\}$$

- Por exemplo, se $A = \{1,2,3\}$ e $B = \{c,d\}$ , então:

$$A \times B = \{(1,c),(1,d),(2,c),(2,d),(3,c),(3,d)\}$$
$$B \times A = \{(c,1),(c,2),(c,3),(d,1),(d,2),(d,3)\}$$

- A noção de produto cartesiano pode ser estendida da seguinte maneira:

  Se $A_1, ..., A_n$ forem conjuntos, então,

$$A_1 \times A_2 \times ... \times A_n = \{(a_1, a_2, ..., a_n) : a_i \in A_i\}$$

  ou seja, o conjunto de todas as ênuplas ordenadas.

- Exemplo com 3 conjuntos: se $A = \{1,2,3\}$, $B = \{c,d\}$ e $C = \{-,*\}$, então

$$A \times B \times C = \{(1,c,-),(1,c,*),(1,d,-),(1,d,*),(2,c,-),\cdots,(3,d,*)\}$$

### ▼ Conjunto das Partes

- Dado um conjunto qualquer $A$ , pode-se definir um outro conjunto, conhecido como **conjuntos das partes** de $A$ , e denotado por $2^A$ , cujos elementos são todos os subconjuntos de $A$.
- Por exemplo: Seja $A = \{1,2,3\}$ , então temos que:

$$2^A = \{\emptyset, A, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}\}$$

- **Obs:** Todo **conjunto das partes** tem por definição o conjunto vazio $\varnothing$ e o próprio conjunto $A$.
- **Exemplo 1:** $A = \{1, \{2,3\}\}$

$$2^A = \{\varnothing, A, \{1\}, \{\{2,3\}\}\}$$

- **Exemplo 2:** $B = \{\varnothing, \{1\}, \{1,3\}\}$

$$2^B = \left\{\varnothing, B, \{\varnothing\}, \left\{\{1\}\right\}, \left\{\{1,3\}\right\}, \left\{\varnothing, \{1\}\right\}, \left\{\varnothing, \{1,3\}\right\}, \left\{\ \{1\}, \{1,3\}\right\}\right\}$$

a. $\{1\} \subset 2^B$ **False** (O $\{1\}$ precisa ter mais uma $\{\}$ para ser subconjunto de $2^B$)

b. $\{1\} \in 2^B$ **False** (O $\{1\}$ é elemento do $B$, e não um subconjunto do $2^B$)

c. $\{\{1\}\} \subset B$ **True**

d. $\{\{1\}\} \in 2^B$ **True**

e. $\{\{\{1\}\}\} \subset 2^B$ **True**

f. $\varnothing \in B$ **True**

g. $\{\varnothing\} \in B$ **True**

h. $\{\varnothing\} \subset B$ **True**

### Teorema de Cantor

- O próximo Teorema conhecido como **Teorema de Cantor**, prova que a cardinalidade do conjunto das partes de qualquer conjunto dado $A$ é maior que
a cardinalidade de $A$.

- **Teorema:** Se $A$ é um conjunto, e $2^A$ é o **conjunto das partes** de $A$, não existe uma função $f : A \to 2^A$ que seja sobrejetiva. Isso quer dizer que a imagem sempre vai ser maior que o domínio, logo nunca vai ter uma função sobrejetiva de $f : A \to 2^A$.

  **Prova:** Recorde que uma função $g : D \to I$ é sobrejetiva se para todo $y \in I$, existe $x \in D$ tal que $g(x) = y$. Suponha por contradição, que existe uma função sobrejetiva $f : A \to 2^A$. Defina o conjunto, $B = \{x \in A : x \notin f(x)\}$. Como $f$ por suposição é sobrejetiva e $B \in 2^A$, temos que existe $b \in A$ tal que $f(b) = B$. Temos que considerar dois casos: $b \in B$ ou $b \in B^c$. Se $b \in B$, então $b \notin f(b)$. Mas como $B = f(b)$, temos que $b \notin B$, absurdo. Se $b \in B^c$, então $b \in f(b)$. Mas como $B = f(b)$, temos que $b \in B$, absurdo. ∎

## Partição

- Dado um conjunto universo $\Omega$ , uma partição $\Pi = \{A_\alpha, \alpha \in I\}$ de $\Omega$ é uma coleção de subconjuntos de $\Omega$ (neste caso, indexados por $\alpha$ que toma valores no conjunto de índices $I$) e satisfaz:

  P1. Para todo $\alpha \neq \beta$, $A_\alpha \cap A_\beta = \emptyset$;

  P2. $\cup_{\alpha \in \mathcal{I}} A_\alpha = \Omega$.

  Ou seja, a interseção de varias partições diferentes de um conjunto universo é igual ao vazio: $\Pi_1 \cap \Pi_2 \cap \Pi_3 = \varnothing$. E a união de todas as partições é igual ao conjunto universo: $\Pi_1 \cap \Pi_2 \cap \cdots \cap \Pi_\alpha = \Omega$.

- Deste modo os conjuntos de uma partição são disjuntos par a par e cobrem todo o conjunto universo. Portanto, cada elemento $w \in \Omega$ pertence a um, e somente um, dos conjuntos $A_\alpha$

de uma partição.

Exemplo : $\Omega = \{1, 2, 3\}$

$2^{\Omega} = \{\emptyset, \Omega, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}\}$

$\Pi_1 = \{\{1\}, \{2,3\}\}$ ; $\Pi_2 = \{\Omega\}$ ; $\Pi_3 = \{\emptyset, \{2\}, \{1,3\}\}$

**Exemplo**

Se $\Omega = \{1, 2, 3, 4\}$, então $\{A_1, A_2\}$, onde $A_1 = \{1, 2, 3\}$ e $A_2 = \{4\}$, é uma partição de $\Omega$. ∎

### Função Indicadora

- Função indicadora é uma maneira de indica um conjunto por meio de uma função.
- É sempre conveniente representar um conjunto $A$ por uma função $I_A$ tendo domínio (conjunto dos argumentos da função) $\Omega$ e contra-domínio (conjunto dos possíveis valores da função) binário $\{0, 1\}$.
- Função Indicadora. A função indicadora $I_A : \Omega \to \{0, 1\}$ de um conjunto $A$ é dada por

$$I_A(\omega) = \begin{cases} 1 & \text{se } \omega \in A, \\ 0 & \text{se } \omega \notin A. \end{cases}$$

  Ou seja, o domínio é todo o conjunto universo $\Omega$, e o contra domínio só tem dois valores $0$ ou $1$. Onde a função indicadora de um conjunto ela vai sempre assumir o valor $1$ ou $0$.

- Note que podemos determinar $A$ a partir de sua função indicadora: $A = \{w : I_A(w) = 1\}$.
  - Se $I_A(\omega)$ for sempre igual a $1$, ou seja, $I_A(\omega) = 1$ , $\forall w \in \Omega$ , então $A$ é o conjunto universo $\Omega$. Ou seja todo elemento pertence ao conjunto $A$.
  - Se $I_A(\omega)$ for sempre igual a $0$, então $A$ é o conjunto vazio $\emptyset$. Ou seja todo elemento não pertencem ao conjunto $A$.
  - Se $I_A(\omega)$ for igual a $1$ somente quando $w = w_0$ , então $A$ é o conjunto $\{w_0\}$ que contém somente o elemento $w_0$.
- Note que existe uma correspondência $1-1$ entre conjuntos e suas funções indicadoras:

$$A = B \Leftrightarrow (\forall \omega \in \Omega) I_A(\omega) = I_B(\omega).$$

- O fato que conjuntos são iguais se, e somente se, suas funções indicadoras forem idênticas nos permitem explorar a aritmética de funções indicadoras:

1. A função indicadora do complementar de um conjunto vai ser igual a função indicadora do conjunto menos $1$: $I_{A^c} = 1 - I_A$.
    a. Se houver um conjunto $A \subset B$, a função indicadora desse conjunto vai sempre $I_A \leq I_B$.
2. A função indicadora da interseção é $I_{A \cap B} = min(I_A, I_B) = I_A \cdot I_B$.
3. A função indicadora da união é $I_{A \cup B} = max(I_A, I_B) = I_A + I_B - I_{A \cap B}$.
4. A função indicadora da indiferença é $I_{A-B} = max(I_A - I_B, 0) = I_A \cdot I_{B^c}$.

- **Exemplo 1:** Prove que $A \subset B \Leftrightarrow B^c \subset A^c$.

$$A \subset B \Leftrightarrow I_A \leq I_B \Leftrightarrow 1 - I_A \geq 1 - I_B \Leftrightarrow I_{A^c} \geq I_{B^c} \Leftrightarrow B^c \subset A^c$$

- **Exemplo 2:** Se $I_A \cdot I_B$ for identicamente igual a zero, o que sabemos a respeito da relação entre $A$ e $B$ ?

  Vimos que a multiplicação entre as indicadoras vai ser sempre igual a $0$, logo se $I_A \cdot I_B = 0$ a indicadora da interseção também vai ser igual a $0$: $I_{A \cap B} = 0$. Isso significa que qualquer elemento do universo não pertence a interseção $A \cap B = \varnothing$, logo essa interseção é igual ao vazio, ou seja esses conjuntos são disjuntos.

$$I_A \cdot I_B = 0 \rightarrow I_A(w) \cdot I_B(w) = 0, \forall w \in \Omega \rightarrow$$
$$I_{A \cap B} = 0 \rightarrow A \cap B = \varnothing$$

- **Exemplo 3:** Se $A \cap B^c = B \cap A^c$ , o que sabemos a respeito da relação entre $A$ e $B$ ?

$$I_{A \cap B^c} = I_{A^c \cap B} \rightarrow I_A \cdot I_{B^c} = I_{A^c} \cdot I_B \rightarrow$$
$$I_A \cdot (1 - I_B) = (1 - I_A) \cdot I_B \rightarrow I_A - I_A \cdot I_B = I_B - I_A \cdot I_B \rightarrow I_A = I_B \rightarrow A = B$$

- **Exemplo 4:** Mostre que $B - A = B \cap A^c$.

$$I_{B-A} = max(I_B - I_A, 0) = I_B \cdot (1 - I_A) = I_B \cdot I_{A^C} = I_{B \cap A^C}$$
$$\text{Ou seja}$$
$$I_{B-A} = I_{B \cap A^C} \text{ é o mesmo que } A - B = B \cap A^C$$

- **Exemplo 5:** Se $I_A - I_B$ for identicamente igual a $1$, o que podemos concluir sobre $A$ e $B$?

  Para $I_A - I_B = 1$ o $I_A = 1$ e o $I_B = 0$. Isso é o mesmo que $A = \Omega$ e o $B = \varnothing$

- **Exemplo 6:** Se $A \cap B = B \cup A$ , o que podemos concluir sobre $A$ e $B$?

$$A \cap B = B \cup A \rightarrow I_{A \cap B} = I_{B \cup A} \rightarrow \min(I_A, I_B) = \max(I_A, I_B)$$
$$\text{Se o } \mathbf{min} \text{ e o } \mathbf{max} \text{ são iguais, o}$$
$$I_A = I_B \rightarrow A = B$$
$$\text{também vai ser igual.}$$

- **Exemplo 7:** Mostre que $A \subset B \Leftrightarrow A \cap B = A$:

$$A \subset B \Leftrightarrow I_A \leq I_B \Leftrightarrow min(I_A, I_B) = I_A \Leftrightarrow I_{A \cap B} = I_A$$
$$\Leftrightarrow A \cap B = A$$

| Operações entre | | | |
|---|---|---|---|
| **Conjuntos** | | **Função Indicadora** | |
| $A^c$ | V | $I_A \cdot I_B$ | V |
| $A \cap B$ | V | $I_A + I_B$ | V |
| $A \cup B$ | V | $I_A - I_B$ | V |
| A-B | V | $(I_{A^c})$ | V |
| $A \cdot B$ | F | $(I_A)^c$ | F |
| A+B | F | $I_A \cup I_B$ | F |
| | | $I_A \cap I_B$ | F |
| | | $I_{A \cup B}$ | V |

## Técnicas de Contagem

▼

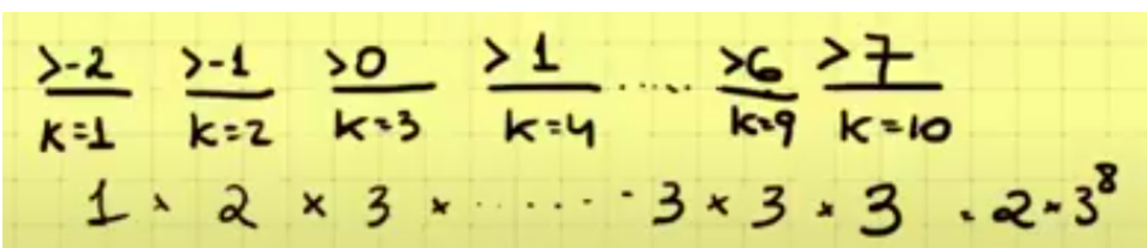
▼ **Exemplo 1:** Quantas são as permutações simples dos números $\{1, 2, ..., 10\}$ nas quais o elemento que ocupa o lugar de ordem $k$, da esquerda para a direita, é sempre maior que $k-3$?

Na primeira posição $k = 1$, os valores que poderão fica nessa posição terão quer ser maior que $> -2$, pois $1 - 3 = -2$. Na segunda posição os valores terão que ser $> -1$, e assim por diante, até a última posição $k = 10$, onde os valores terão que ser $> 7$.

Vamos contar o número de valores possíveis para cada posição, vamos começa da direita para a esquerda, pois sempre devemos começar a contar da posição que é mais restrita até a posição que é menos restrita.

A posição $k = 10$, tem $3$ valores possíveis para essa posição. A posição $k = 9$, tem $4$ valores possíveis, mais como já foi usado um valor na posição $k = 10$, resta para a posição $k = 9$, $3$ valores. A posição $k = 8$, tem $3$ posições possíveis, pois já foi usado dois valores. Isso se repete até a posição $k = 3$. Na posição $k = 2$ resta $2$ valores, e na posição $k = 1$ resta um valor.

Logo o número de permutações vai ser igual a $2 \cdot 3^8$.




**A mesma solução para esse problema, mais com palavras diferentes:**

**Solução:** Comecemos escolhendo os números da direita para esquerda. Observe que o número no lugar de ordem 10, tem que ser maior que 7, portanto existem 3 opções. O número no lugar de ordem 9, tem que ser maior que 6, existem, portanto, 3 opções visto que um dos números maiores que 6 já foi utilizado na última posição. De maneira similar pode-se ver que existem 3 opções para os números que ocupam do terceiro ao oitavo lugar. O número no lugar de ordem 2, tem somente 2 opções, pois oito números já foram escolhidos anteriormente. Finalmente, resta apenas um número para o lugar de ordem 1. Portanto, existem $2 \times 3^8$ permutações deste tipo.

## Introdução à Probabilidade

- Probabilidade é uma função que tem como domínio uma sigma álgebra de eventos e um certo espaço amostral, de forma que cada evento dessa sigma álgebra vai está associado a um número real. E tem que satisfazer algumas propriedades derivada lá da frequência relativa, como a probabilidade nunca é negativa; a probabilidade do espaço amostra é sempre igual a $1$; Se $A$ e $B$ são disjuntos, então $P(A \cup B) = P(A) + P(B)$.

### Experimento Aleatório

- Um experimento é qualquer processo de observação. Experimentos aleatórios são aqueles em que não somos capazes de predizer seu comportamento em futuras realizações.
- Motivos pela qual, nós não conseguimos predizer o resultado de experimentos aleatórios com precisão:
  - Nós podemos não saber de todas as causas envolvidas;
  - Nós podemos 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;
  - Ou na verdade existe alguma aleatoriedade fundamental no experimento.
- **Obs:** É importante ressaltar que frequentemente são encontradas situações práticas onde não se consegue descrever todos os possíveis resultados de um experimento. Uma maneira de contornar este problema é assumir que um resultado possível do experimento é a não ocorrência de qualquer um dos resultados descritos, contudo, em problemas práticos, tal suposição pode acarretar em dificuldades quando se tenta elicitar ou deduzir probabilidades.
- Os resultados de um experimento aleatório são caracterizados pelos seguintes componentes:
  1. O conjunto de resultados possíveis **Espaço Amostral** $\Omega$;
  2. A coleção de conjuntos de resultados de interesse **Evento** $A$;
  3. Um valor numérico $P$ da verossimilhança ou probabilidade de ocorrência de cada um dos conjuntos de resultados de interesse.

#### Espaço Amostral

- O conjunto de possíveis resultados de um experimento aleatório é chamado de **Espaço Amostral** $\Omega$.
- Podemos fazer uma associação com a teoria dos conjuntos, onde o **Espaço Amostral** é como se fosse o **Conjunto Universo**.
- Em um dado experimento aleatório a especificação do espaço amostral deve ser tal que este:
    1. Liste todos os possíveis resultados do experimento sem duplicação.
    2. Faça em um nível de detalhamento suficiente para os interesses desejados, omitindo resultados que, embora logicamente ou fisicamente possíveis, não tenham qualquer implicação prática na sua análise.
- **Por exemplo**, uma única jogada de uma moeda pode ter o espaço amostral tradicional $\Omega = \{cara, coroa\}$, ou podemos considerar que a moeda pode fisicamente ficar equilibrada na borda $\Omega = \{cara, coroa, borda\}$.
- De acordo com o número de elemento de um conjunto, podemos classificar os espaços amostrais em dois tipos: **Enumeráveis** ou **Não Enumeráveis**.
- Em um nível filosófico, pode-se argumentar que só existem espaços amostrais enumeráveis, visto que medidas não podem ser feitas 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.

▼ **Eventot**

- Um **evento** $A$ 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 tanto, utilizaremos as operações Booleanas de conjuntos (**complementar**, **união**, **intersecção**, **diferença**) para expressar eventos combinados de interesse.
- Os eventos $A$ e $B$ são **Disjuntos** (também pode ser chamado de: **Mutuamente Excludentes** ou **Mutuamente Exclusivos**) se não puderem ocorrer juntos, ou, em linguagem de conjuntos, $A \cap B = \emptyset$ .

▼ **Exemplo:** Sejam $A$, $B$, e $C$ eventos em um mesmo espaço amostral $\Omega$. Expresse os seguintes eventos em função de $A$ , $B$ , e $C$ e operações Booleanas de conjuntos.

a. Pelo menos um deles ocorre: $A \cup B \cup C$

b. Apenas $A$ ocorre: $A \cap B^c \cap C^c$

c. Exatamente um deles ocorre: $(A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A^c \cap B^c \cap C)$

d. Pelo menos dois ocorrem: $(A \cap B) \cup (A \cap C) \cup (B \cap C)$ outra alternativa seria $(A \cap B \cap C^c) \cup (A \cap B^c \cap C) \cup (A^c \cap B \cap C) \cup (A \cap B \cap C)$

e. No máximo dois deles ocorrem: $(A \cap B \cap C)^c$, pela lei De Morgan isso também poderia ser: $A^c \cup B^c \cup C^c$

f. Nenhum deles ocorre: $(A \cup B \cup C)^c$, pela lei De Morgan isso também poderia ser: $A^c \cap B^c \cap C^c$

g. Ambos $A$ e $B$ ocorrem, mas $C$ não ocorre: $(A \cap B \cap C^c)$.

- Embora possa-se pensar que, dado um espaço amostral, necessariamente é de interesse analisar todos os seus subconjuntos (e isto eventualmente é verdadeiro), temos três razões para esperar que estejamos apenas interessados em alguns subconjuntos do espaço amostral:

  1. O espaço amostral pode conter um grau de detalhamento superior ao que estamos interessados no momento. Por exemplo, ele pode representar uma única jogada de um dado com 6 elementos, mas nós apenas estamos interessados em saber se o resultado é par ou ímpar.
  2. Nós vamos querer associar cada evento $A$ com uma probabilidade numérica $P(A)$. Como essas probabilidades estão baseadas em algum conhecimento sobre a tendência de ocorrer do evento ou no grau de nossa crença que determinado evento ocorrerá, nosso conhecimento sobre $P$ pode não estender para todos os subconjuntos de $\Omega$.
  3. Condições impostas em $P$ pelos axiomas de Kolmogorov, que estudaremos adiante, podem não permitir que $P$ seja definida em todos os subconjuntos de $\Omega$, em particular isto pode ocorrer quando $\Omega$ for não enumerável, mas não iremos demonstrar este fato que está fora do escopo deste curso.

▼ **Álgebra de Eventos**

- Uma álgebra é um subconjunto do conjunto das partes.
- Uma álgebra de eventos $\mathcal{F}$ é uma coleção de subconjuntos do espaço amostral $\Omega$ que satisfaça três condições:
  1. Não pode ser vazia;
  2. Fechada com respeito a complementos (se $A \in \mathcal{F}$, então $A^c \in \mathcal{F}$);
  3. Fechada com respeito a uniões nitas (se $A, B \in \mathcal{F}$, então $A \cup B \in \mathcal{F}$);
- Uma álgebra de eventos sempre vai ter os elementos $\Omega$ e $\varnothing$.

▼ **Propriedades:**

- Se $A$ é uma álgebra, então $\Omega \in A$.

  **Prova:** Como $A$ é não vazia, seja $A$ um elemento qualquer seu. Pela segunda propriedade de álgebras, temos que $A^c \in A$, e pela terceira propriedade temos que $\Omega = A \cup A^c \in A$.

- Sejam $A_1$ e $A_2$ álgebras de subconjuntos de $\Omega$ e seja $A = A_1 \cap A_2$ a coleção de subconjuntos comuns as duas álgebras. Então, $A$ é uma álgebra. Ou seja a interseção de duas álgebra, pertencente ao mesmo $\Omega$, vai gera uma nova álgebra.

**Prova:** Como $\mathcal{A}_1$ e $\mathcal{A}_2$ são álgebras, ambos contém $\Omega$. Então, $\Omega \in \mathcal{A}$. Se $A \in \mathcal{A}$, então $A$ está em ambos $\mathcal{A}_1$ e $\mathcal{A}_2$. Logo, $A^c$ está em ambos $\mathcal{A}_1$ e $\mathcal{A}_2$, e portanto na sua intersecção $\mathcal{A}$. Se $A, B \in \mathcal{A}$, então eles estão em ambos $\mathcal{A}_1$ e $\mathcal{A}_2$. Consequentemente, $A \cup B$ está em ambos $\mathcal{A}_1$ e $\mathcal{A}_2$ e, portanto, em $\mathcal{A}$. Como $\mathcal{A}$ satisfaz as três condições da definição de álgebra de eventos, $\mathcal{A}$ é uma álgebra de eventos. ∎

- Existe uma menor (no sentido de inclusão) álgebra contendo qualquer família dada de subconjuntos de $\Omega$.

**Prova:** Seja $\mathcal{C}$ uma coleção qualquer de subconjuntos de $\Omega$, defina $\mathcal{A}(\mathcal{C})$ como sendo o conjunto que é igual a intercessão de todas as álgebras de eventos que contém $\mathcal{C}$, isto é:

$$\mathcal{A}(\mathcal{C}) = \bigcap_{\mathcal{A} \supseteq \mathcal{C} : \mathcal{A} \text{ é uma álgebra de eventos}} \mathcal{A}.$$

Pelo Teorema 3.8, $\mathcal{A}(\mathcal{C})$ é uma álgebra de eventos, e consequentemente é a menor álgebra de eventos contendo $\mathcal{C}$. $\mathcal{A}(\mathcal{C})$ é conhecida como a *álgebra de eventos gerada por* $\mathcal{C}$. ∎

- A diferença de **Álgebra** para **Sigma Álgebra**, é que a **Sigma Álgebra** tem uma propriedade a mais. Sendo assim toda **Sigma Álgebra** é uma **Álgebra**.

▼ **Construindo uma álgebra:**

- Uma maneira de construir uma álgebra de eventos, é primeiro particionar $\Omega$ em um número finito de subconjuntos e depois considerar álgebra que consiste dos eventos que são uniões finitas dos subconjuntos da partição.
- Se quisemos construir uma Álgebra, pegamos todas as possíveis uniões finitas.
- Se quisemos construir uma Sigma Álgebra, pegamos todas as possíveis uniões enumeráveis.

**Exemplo**

Por exemplo, $\Omega = \{a, b, c, d\}$. Considere a partição, $\{\{a, c\}, \{b, d\}\}$, então considere a coleção de eventos que consiste de uniões finitas dos eventos desta partição: $\mathcal{A} = \{\emptyset, \Omega, \{a, c\}, \{b, d\}\}$. É fácil ver que $\mathcal{A}$ é uma álgebra de eventos.

▼ **Sigma Álgebra**

- Uma sigma álgebra é representada por $\sigma - álgebra$.
- Uma $\sigma - álgebra$ é uma coleção de eventos que satisfaz as três propriedades da Álgebra de Eventos e satisfaz mais uma propriedade:
  - Se houver uma quantidade enumerável de eventos na $\sigma - álgebra$: $A_i \in \mathcal{A}$, a união de todos eles tem que está também na $\sigma - álgebra$: $\cup_{i \in I} A_i \in \mathcal{A}$. Ou seja a

união de dois elementos, ou mais por exemplo, tem que resultar em um conjunto que esteja dentro da sigma álgebra também.

- Uma definição mais precisa seria:
  - Uma $\sigma - \acute{a}lgebra$ de eventos $\mathcal{A}$ é uma **álgebra de eventos** de um espaço amostral $\Omega$ que também é fechada com relação a uma união enumerável de eventos, ou seja, se $\mathcal{I}$ for um conjunto enumerável:

$$A_i \in \mathcal{A}, \forall_i \in \mathcal{I} \Rightarrow \cup_{i \in I} A_i \in \mathcal{A}$$

- Estamos usando o símbolo $\mathcal{A}$ para representar uma $\sigma - \acute{a}lgebra$ de eventos.

**Exemplo**

1. A menor $\sigma$-álgebra de eventos é $\mathcal{A} = \{\emptyset, \Omega\}$;
2. A maior $\sigma$-álgebra de eventos é o conjunto das partes de $\Omega$;
3. Um exemplo intermediário, temos:

$$\Omega = \{1, 2, 3\},\ \mathcal{A} = \{\Omega, \emptyset, \{1\}, \{2, 3\}\}.$$

4. A álgebra de eventos finitos e co-finitos. Seja $\Omega = \boldsymbol{R}$ e

$$\mathcal{A} = \{A \subseteq \boldsymbol{R} : A \text{ é finito}\} \cup \{A \subseteq \boldsymbol{R} : A^c \text{ é finito}\},$$

ou seja, $\mathcal{A}$ consiste dos subconjuntos de $\boldsymbol{R}$ que ou são finitos ou têm complementos finitos. $\mathcal{A}$ é uma álgebra de eventos, mas não é uma $\sigma$-álgebra.

- O exemplo 3 acima, $\mathcal{A} = \{\Omega, \emptyset, \{1\}, \{2, 3\}\}$, é uma $\sigma - \acute{a}lgebra$ pois satisfaz as propriedades, ela é uma coleção não vazia; cada elemento dela está acompanhado de seu complementar, por exemplo, o elemento $\Omega$ está acompanhado de seu complementa $\varnothing$; A união de qualquer elemento da $\sigma - \acute{a}lgebra$ vai gerar um conjunto que esta dentro da $\sigma - \acute{a}lgebra$, por exemplo, a $\{1\} \cup \{2, 3\}$ vai resultar no $\Omega$ que também está na $\sigma - \acute{a}lgebra$.
- **Objs:** A $\sigma - \acute{a}lgebra$ não precisar ser igual ao conjunto das partes, ou seja não precisar ter todos os subconjuntos de $\Omega$.

▼ **Proposição 1:**

Seja $\mathcal{F}$ uma $\sigma - \acute{a}lgebra$ de subconjuntos de $\Omega$. Então valem as seguintes propriedades:

- $\varnothing \in \mathcal{F}$
- $\forall n, \forall A_1, \cdots, A_n \in \mathcal{F}, \text{ então } \cap_{n=1}^{\infty} A_n \in \mathcal{F}$.

▼ **Proposição 2:**

- A interseção infinita de $\sigma - \acute{a}lgebra$ é também uma $\sigma - \acute{a}lgebra$.

▼ **Gerando uma $\sigma - álgebra$**

- Dada uma coleção finita eventos $C = \{A_1, A_2, ..., A_n\}$ , define-se um **átomo** de $C$ como sendo qualquer evento $B$ da seguinte forma: $B = B_1 \cap B_2 \cap ... \cap B_n$ , onde $B_i = A_i$ ou $B_i = A_i^c$ para $i = 1, 2, ..., n$ .
- Note que existem no máximo $2^{||C||}$ **átomos** diferentes e que eles formam uma partição de $C$ (verifique!).
- Quando $C$ for uma coleção finita de eventos, um evento pertencerá a $\mathcal{A}(C)$ se e somente se, for igual a uma união finita de **átomos** de $C$.
- Note que $\mathcal{A}(C)$ terá no máximo $2^{2^{||C||}}$ elementos (verifique!).

▼ **Exemplo 1:** Se $\Omega = \{a, b, c, d, e, f\}$, encontre a $\sigma - álgebra$ gerada por $C = \{\{a, b, d\}, \{b, d, f\}\}$.

Vamos considerar o $A_1 = \{a, b, d\}$ e o $A_2 = \{b, d, f\}$.

Os átomos (as possíveis interseções entre $A_1$ e $A_2$) desse exemplo são 4:

1. $A_1 \cap A_2 = \{b, d\}$
2. $A_1 \cap A_2^c = \{a\}$
3. $A_1^c \cap A_2 = \{f\}$
4. $A_1^c \cap A_2^c = \{c, e\}$

Como temos $4$ átomos distintos não vazios, a $\sigma - álgebra$ vai ter $2^4$ elementos.

$$\mathcal{A} = \Big\{\varnothing, \Omega\, \{b, d\}, \{a\}, \{f\}, \{c, e\}, \{a, b, d\}, \{b, d, f\}, \\ \{b, d, c, e\}, \{a, f\}, \{a, c, e\}, \{c, e, f\}, \{a, b, d, f\}, \\ \{a, b, d, c, e\}, \{b, d, f, c, e\}, \{a, f, c, e\}\Big\}$$

▼ **Exemplo 2:** Se $\Omega = \{1, 2, 3, 4, 5\}$, encontre a menor $\sigma - álgebra$ quem contém $C = \{\{1, 3, 5\}, \{1, 3\}, \{2, 4\}\}$.

Os átomos são:

1. $A_1 \cap A_2 \cap A_3 = \varnothing$
2. $A_1 \cap A_2 \cap A_3^c = \{1, 3\}$
3. $A_1 \cap A_2^c \cap A_3 = \varnothing$
4. $A_1 \cap A_2^c \cap A_3^c = \{5\}$
5. $A_1^c \cap A_2 \cap A_3 = \varnothing$
6. $A_1^c \cap A_2 \cap A_3^c = \varnothing$
7. $A_1^c \cap A_2^c \cap A_3 = \{2, 4\}$
8. $A_1^c \cap A_2^c \cap A_3^c = \varnothing$

O número de elementos vai ser $2$ elevado ao número de átomos não vazio $2^3$.

$$\mathcal{A} = \Big\{\varnothing, \Omega, \{1,3\}, \{5\}, \{2,4\}, \{1,3,5\}, \{1,2,3,4\}, \{2,4,5\}\Big\}$$

▼ **Probabilidade**

- Nós podemos classificar as formas de raciocínio probabilístico nas seguintes dimensões:
    - **Grau de precisão — o conceito estrutural.**

      O conceito estrutural determina a precisão com que podemos esperar que probabilidade represente fenômenos aleatórios.
    - **O significado, ou interpretação a ser dada a probabilidade.**

      A interpretação proporciona a base com a qual probabilidade deve ser determinada e indica o que podemos esperar aprender com ela, ou seja, o que uma armação probabilística significa.
    - **Estrutura matemática formal de probabilidade dada por um conjunto de axiomas.**

      O conceito estrutural e a interpretação guiam a escolha dos axiomas. O conjunto de axiomas, contudo, pode somente capturar uma parte do que entendemos da interpretação.
- Compreensão de fundamentos de probabilidade é importante, pois aplicações de teoria da probabilidade dependem fortemente de seus fundamentos. Por exemplo, os fundamentos influem na escolha dos métodos estatísticos a serem utilizados (Frequentistas, Bayesianos, . . . ) e na interpretação dos resultados obtidos.
- **Paradoxo De Ellsbergue:**

  > **Paradoxo de Ellsbergue.** Suponha que existam duas urnas cada uma com 60 bolas. A urna 1 contém 30 bolas azuis e 30 bolas verdes. Tudo que se sabe sobre a urna 2 é que ela contém bolas azuis e verdes, mas não sabe-se a distribuição das bolas. Considere que existem duas loteria com prêmios baseados no sorteio de bolas dessas urnas. Loteria $L_1$ paga R\$1.000,00 se uma bola azul for sorteada na urna 1, e R\$0,00 caso contrário. Loteria $L_2$ paga R\$1.000,00 se uma bola azul for sorteada na urna 2, e R\$0,00 caso contrário. A maioria das pessoas quando questionada se prefere um bilhete da Loteria $L_1$ ou $L_2$ prefere um bilhete da loteria $L_1$.

- **Aversão a Ambiguidade:**

Suponha agora que temos duas outras loterias $L_3$ e $L_4$, onde a primeira paga R$1.000,00 somente se uma bola verde for sorteada da urna 1, e a segunda para R$1.000,00 somente se uma bola verde for sorteada da urna 2. Também, é verificado que a maioria das pessoas que preferiram a loteria $L_1$ a loteria $L_2$ preferem a loteria $L_3$ a loteria $L_4$. Com estas preferências, não é possível que o decisor possua uma única distribuição de probabilidade subjetiva sobre as cores das bolas na urna 2, pois a primeira preferência ($L_1$ sobre $L_2$) indica que o decisor considera que existam mais bolas verdes que azuis na urna 2, e a segunda ($L_3$ sobre $L_4$) indica que o decisor considera que existam mais bolas azuis que verdes na urna 2. Esse fenômeno é conhecido na literatura como *aversão a ambiguidade*, e pode-se modelar a incerteza do decisor por um conjunto de medidas de probabilidade ao invés de uma única medida de probabilidade.

- Nós discutiremos uma variedade de conceitos estruturais e interpretações de probabilidade. Depois nós focaremos na probabilidade numérica tradicional que satisfaz os famosos axiomas de Kolmogorov e em uma interpretação baseada em frequências de ocorrência.
- Os seguintes são exemplos de uma variedade de conceitos estruturais de probabilidade:

**Possivelmente:**

"Possivelmente $A$" é o conceito mais rudimentar e menos preciso, e o usado pelos antigos Gregos para distinguir entre o que era necessário e o que era contingente. Existe um número de conceitos de possibilidade que incluem os seguintes:

- **Possibilidade Lógica:** No sentido que não se contradiz logicamente. Ou seja uma preposição matemática não pode ser verdadeira e falsa ou mesmo tempo.
- **Possibilidade Epistêmica:** Segundo a qual ocorrência de $A$ não contradiz nosso conhecimento, que inclui, mas estende mais que mera lógica. Por exemplo, logicamente, é possível que neve em Fortaleza, mais nos sabemos que esse fenômeno é impossível.
- **Possibilidade Física:** A ocorrência de $A$ é compatível com leis físicas, contudo ela pode ser extremamente improvável, por exemplo, jogar uma moeda, e ela cair equilibrada na borda em uma superfície rígida é algo possível.
- **Possibilidade Prática:** A noção do dia-a-dia segundo a qual $A$ praticamente possível se ele tem pelo menos uma verossimilhança não tão pequena de ocorrer.

**Provavelmente:**

Provavelmente $A$ é um fortalecimento da noção de possibilidade que significa mais que provável que não. Enquanto ela pode corresponder ao caso que a probabilidade numérica de $A$ seja maior que 1/2, este conceito não requer nenhum comprometimento com probabilidade numérica nem com o preciso estado de conhecimento que probabilidade numérica requer.

▼ **Probabilidade Comparativa:**

"$A$ é pelo menos tão provável quanto $B$". A probabilidade comparativa inclui "provavelmente $A$" através de "$A$ é pelo menos tão provável quanto $A^c$ ". Pode ser relacionada com probabilidade numérica através de $P(A) \geq P(B)$; embora como nos dois exemplos anteriores, probabilidade comparativa não requer nenhum comprometimento com probabilidade numérica.

▼ **Probabilidade Intervalar:**

"$A$ tem probabilidade intervalar, ou probabilidade inferior e superior $(\underline{P}(A), \overline{P}(A))$". Isto permite um grau de indeterminação variável sem nenhum comprometimento com que exista um verdadeiro valor no intervalo.

▼ **Probabilidade Numérica:**

"A probabilidade de $A$ é o número real $P(A)$". Este é o conceito usual como qual nos ocuparemos neste curso. Enquanto este conceito absorveu quase toda atenção de pessoas envolvidas com fenômenos de chance e incerteza e provou ser frutífero na prática científica, este não é o único conceito utilizado em linguagem ordinária e no raciocínio probabilístico do dia-a-dia. É duvidoso que probabilidade numérica seja adequada a todas as aplicações que ela é utilizada, e é provável que ela tenha inibido o desenvolvimento de teorias matemáticas apropriadas para outros fenômenos aleatórios.

- De agora em diante focaremos no conceito estrutural **mais utilizado** e preciso que é a **probabilidade numérica**.
- Parece não ser possível reduzir probabilidade a outros conceitos; ela é uma noção em si mesma. O melhor que podemos fazer é relacionar probabilidade a outros conceitos através de uma interpretação. Os cinco mais comuns grupos de interpretação são os seguintes: **Lógica**, **Subjetiva**, **Frequentista**, **Propensidade**, **Clássica**.

▼ **Lógica:**

Grau de confirmação da hipótese de uma proposição que "$A$ ocorre" dada uma evidência através da proposição que "$B$ ocorreu". Esta interpretação está ligada a um sistema lógico formal e não, digamos, ao mundo físico. Ela é usada para tornar o raciocínio indutivo quantitativo. Quando as evidências ou premissas são insuficientes para deduzir logicamente a hipótese ou conclusão, podemos ainda medir quantitativamente o grau de suporte que uma evidência da a uma hipótese através de probabilidade lógica.

▼ **Subjetiva:**

Se refere ao grau de crença pessoal na ocorrência do evento $A$ e é medida através da interpretação comportamental de disposição a apostar ou agir.

▼ **Frequentista:**

Se refere ao limite da frequência relativa de ocorrência do evento $A$ em repetidas realizações não relacionadas do experimento aleatório $\mathcal{E}$. Note que limites de frequência relativas são uma idealização, pois não se pode realizar infinitas realizações de um experimento.

▼ **Propensidade:**

Tendência, propensidade, ou disposição para um evento $A$ ocorrer. Por exemplo, considerações de simetria, podem levar a conclusão que um dado tem a mesma propensão ou tendência a cair em qualquer uma de suas faces.

▼ **Clássica:**

Baseada em uma enumeração de casos igualmente prováveis.

- Resta-nos discutir o terceiro elemento para modelagem do raciocínio probabilístico, a associação de uma medida numérica a eventos que representam a probabilidade com que eles ocorrem.
- As propriedades desta associação são motivadas em grande parte pelas propriedades de **frequência relativas**.

▼ **Frequência Relativas:**

Considere uma coleção de experimentos aleatórios $\mathcal{E}_1$ que possuem a mesma $\sigma - álgebra$ de eventos $\mathcal{A}$ e tem resultados individuais não necessariamente numéricos $\{w_i\}$. Fixando uma dada sequência de resultados $\{w_i\}$, se estamos interessados na ocorrência de um dado evento $A$, a frequência relativa de $A$ nada mas é que uma média aritmética da função indicadora de $A$ calculada em cada um dos termos da sequência $\{w_i\}$, ou seja

**Definição**

A frequência relativa de um evento $A$, determinada pelos resultados $\{\omega_1, \ldots, \omega_n\}$ de $n$ experimentos aleatórios, é

$$r_n(A) = \frac{1}{n}\sum_{i=1}^{n} I_A(\omega_i) = \frac{N_n(A)}{n}.$$

**Exemplo:** Um dado foi jogado 12 vezes. E foi anotado o resultado de cada jogada: $\{1, 2, 4, 2, 2, 6, 5, 5, 1, 3, 4, 4\}$, qual é a frequência relativa dos eventos:

a. O número ser par, $A = \{2, 4, 6\}$.

$$r_{12}(A) = \frac{6}{12}$$

b. O número ter sido o 4, $B = \{4\}$.

$$r_{12}(B) = \frac{3}{12}$$

▼ **Propriedades de Frequência Relativa**

1. Ela é uma função que para cada evento ela associa uma valor real.

$$\text{FR0.} \quad r_n : \mathcal{A} \to \boldsymbol{R}.$$

2. Frequência relativa nunca é negativa.

$$\text{FR1.} \quad r_n(A) \geq 0.$$

3. A frequência do espaço amostral sempre vai ser igual a 1.

$$\text{FR2.} \quad r_n(\Omega) = 1.$$

4. FR3. Se $A$ e $B$ são disjuntos, então $r_n(A \cup B) = r_n(A) + r_n(B)$.

5. FR4. Se $A_1, A_2, \cdots A_n, \cdots$ é uma sequência de eventos disjuntos dois a dois, então $r_n(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} r_n(A_i)$.

- Nós prosseguiremos como se existisse algo que garanta que $r_n(A) \to P(A)$ , embora que o sentido de convergência quando $n$ cresce só será explicado pela Lei dos Grandes Números, que não será discutida neste curso. Esta tendência da frequência relativa de estabilizar em um certo valor é conhecida como regularidade estatística. Deste modo, $P$ herdará propriedades da frequência relativa $r_n$.
- Questões de probabilidade em situações práticas basicamente constituem-se, como seria o esperado, em como calcular probabilidades. Aí é onde a situação se complica. Portanto a construção axiomática da teoria da probabilidade, abstrai o cálculo de probabilidades de casos particulares e nos provê de um método formal para resolver problemas probabilísticos.
- Os axiomas que descreveremos a seguir não descrevem um único modelo probabilístico, eles apenas determinam uma família de modelos probabilísticos, com os quais poderemos utilizar métodos matemáticos para descobrir propriedades que serão verdadeiras em qualquer modelo probabilístico. A escolha de um modelo específico satisfazendo os axiomas é feito pelo analista/estatístico familiar com o fenômeno aleatório sendo modelado.

**Axiomas de Kolmogorov:**

- Motivados pelas propriedades de frequência relativa, impõe-se os primeiros quatro axiomas de Kolmogorov:
- **k1 Inicial:** O experimento aleatório é descrito pelo espaço de probabilidade $(\Omega, \mathcal{A}, P)$, que consiste do espaço amostral $\Omega$; de uma $\sigma - álgebra\ \mathcal{A}$ , e de uma função de valores reais $P : \mathcal{A} \to R$.
- **k2 Não-negatividade:** $\forall A \in A, P(A) \geq 0.$
- **K3 Normalização Unitária:** $P(\Omega) = 1$
- **k4 Aditividade Finita:** Se $A$, $B$ são disjuntos, então $P(A \cup B) = P(A) + P(B)$.
- É fácil provar (tente!) utilizando indução matemática que $K4$ é válida para qualquer coleção finita de eventos disjuntos par a par, ou seja, se $A_i, i = 1, 2, \cdots, n$, são

eventos disjuntos par a par, então $P(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$

- Esses quatro axiomas acima são suficiente para eventos **finitos**.
- Quando temos um espaço amostral **infinito**, necessitamos de mais um axioma.
- Esse quinto axioma foi proposto por Kolmogorov para garantir um certo grau de continuidade da medida de probabilidade.
- **k5 Continuidade Monotônica:** Se para todo $i > 0, A_{i+1} \subseteq A_i$ e $\cap_i A_i = \varnothing$, então

$$\lim_{i \to \infty} P(A_i) = 0$$

- Um forma equivalente de **K5** é a seguinte:
- **k5**$'$ $\sigma - \text{aditividade}$: Se $\{A_i\}$ é uma coleção enumerável de eventos disjuntos dois a dois, então

$$P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$$

- Vamos prova que qualquer função que satisfaça os 4 primeiro axiomas e satisfaça também o k5, ela por definição vai satisfazer o k5$'$. O mesmo vale para o inverso, se a função satisfaça os 4 primeiros axiomas e o k5$'$, ela por definição vai satisfazer o k5.
- Primeiro, vamos provar que K1 até K5 implica o axioma da $\sigma - \text{aditividade}$ K5$'$:
  - Seja $\{A_i\}$ qualquer sequência enumerável de eventos disjuntos par a par, e defina para todo $n$

$$B_n = \cup_{i>n} A_i,$$
$$\cup_{i=1}^{\infty} A_i = B_n \cup (\cup_{i=1}^{n} A_i).$$

Claramente, para todo $i \leq n$, temos que $A_i$ e $B_n$ são disjuntos. Por K4, temos

$$P(\cup_{i=1}^{\infty} A_i) = P(B_n) + \sum_{i=1}^{n} P(A_i).$$

Por definição de série numérica,

$$\lim_n \sum_{i=1}^{n} P(A_i) = \sum_{i=1}^{\infty} P(A_i).$$

K5$'$ segue se conseguirmos mostrar que $\lim_n P(B_n) = 0$. Note que $B_{n+1} \subseteq B_n$, e que $\cap_{n=1}^{\infty} B_n = \emptyset$. Então por K5, temos que o limite acima é zero e K5$'$ é verdadeiro.

- Agora, vamos provar que K1 até K4, k5$'$ implica o axioma da continuidade monotônica K5:

Agora, vamos provar que K1—K4, K5′ implicam o axioma da continuidade monotônica K5. Seja $\{B_n\}$ qualquer coleção enumerável de eventos satisfazendo as hipóteses do axioma K5: $B_{n+1} \subseteq B_n$ e $\cap_{n=1}^{\infty} B_n = \emptyset$. Defina, $A_n = B_n - B_{n+1}$ e observe que $\{A_n\}$ é uma coleção enumerável de eventos disjuntos par a par. Note que

$$B_n = \cup_{j \geq n} A_j.$$

Então, por K5′ temos que

$$P(B_n) = P(\cup_{j \geq n} A_j) = \sum_{j \geq n} P(A_j).$$

Como por K5′,

$$\sum_{j=1}^{\infty} P(A_j) = P(\cup_{j=1}^{\infty} A_j) \leq 1,$$

temos que

$$\lim_n P(B_n) = \lim_n \sum_{j \geq n} P(A_j) = 0,$$

logo K5 é verdadeiro. ▮

- **Espaço de Probabilidade:**
  - A terna $(\Omega, \mathcal{A}, P)$ é chamada de **espaço de probabilidade**.
  - Intuitivamente quando se modela um problema através de probabilidade, basicamente, o que se faz é especificar cada uma das componentes da terna acima.
  - Eventos são os elementos de $\mathcal{A}$, aos quais se pode atribuir probabilidade.
  - Probabilidade é uma função cujo argumento é um conjunto.

- **Linguagem Probabilística**
  - Podemos fazer um paralelo entre **teoria dos conjuntos** e a **teoria da probabilidade**.
  - Entretanto, é preciso que a linguagem de conjuntos seja traduzida para a linguagem de probabilidade.
  - A tabela, a seguir, exibe algumas dessas traduções.
  - A ideia subjacente é que um experimento aleatório foi realizado e aconteceu algum evento.

| | | |
|---|---|---|
| $\Omega$ | conjunto universo | espaço amostral, evento certo |
| $\omega$ | elemento | resultado do experimento |
| $A$ | conjunto $A$ | evento $A$ |
| $\emptyset$ | conjunto vazio | evento impossível |
| $A^c$ ou $\overline{A}$ | complemento de $A$ | não ocorreu o evento $A$ |
| $A \cap B$ | $A$ intersecção $B$ | os eventos $A$ e $B$ ocorreram |
| $A \cup B$ | $A$ união $B$ | os eventos $A$ ou $B$ ocorreram |
| $\cap_n A_n$ | intersecção dos conjuntos $A_n$ | todos os eventos $A_n$ ocorreram |
| $\cup_n A_n$ | união dos conjuntos $A_n$ | ao menos um dos eventos $A_n$ ocorreu |

A segunda coluna é referente a teoria dos conjuntos e a terceira coluna é referente a teoria da probabilidade.

**Propriedades de uma Medida de Probabilidade**

- Se $P$ é uma medida de probabilidade, então ela deve satisfazer as seguintes propriedades:

1. $P(A^c) = 1 - P(A)$.
2. $P(\emptyset) = 0$.
3. $P(A) \leq 1$.

**Prova:** Parte 1, segue do fato que $\Omega = A \cup A^c$, K3, e K4, pois

$$1 = P(\Omega) = P(A) + P(A^c).$$

Parte 2, segue da Parte 1, do fato que $\Omega^c = \emptyset$, e K3, pois

$$P(\emptyset) = 1 - P(\Omega) = 0.$$

Parte 3, segue do fato que $1 = P(\Omega) = P(A) + P(A^c) \geq P(A)$, já que $P(A^c) \geq 0$ por K2. ∎

**Propriedade de Monotonicidade:**

$$\textit{Se } A \subseteq B, \textit{ então } P(A) \leq P(B).$$

**Prova:** Note que podemos separar o $B$ em duas partes: $B = A \cup (B - A)$. Como o $A$ e o $B - A$ são disjuntos, usaremos o axioma k4. Então a probabilidade de $B$ vai ser igual a $P(B) = P(A) + P(B - A)$, como toda probabilidade não é negativa, essa soma vai ser maior ou igual que a probabilidade do $A$: $P(A) + P(B - A) \geq P(A)$.

Usando essa propriedade podemos prova que:

$$P(A \cup B) \geq \max(P(A), P(B)) \geq \min(P(A), P(B)) \geq P(A \cap B).$$

Ou seja, usando a monotonicidade, sabemos que qualquer evento de $A$ está contido na $A \cup B$, e que qualquer evento de $B$ esta contido em $A \cup B$, logo a $P(A \cup B)$ é maior ou igual que a maior das duas $P(A), P(B)$, ou seja $P(A \cup B) \geq \max\Big(P(A), P(B)\Big)$, logo a maior das duas probabilidades é maior ou igual a menor das duas probabilidades. Usando a monotonicidade novamente, temos que a $P(A \cap B)$ é menor ou igual que o menor evento entre as duas probabilidades, ou seja $\min\Big(P(A), P(B)\Big) \geq P(A \cap B)$.

- Usamos o Propriedade de Monotonicidade quando os conjuntos são disjuntos. Quando não forem disjuntos, usamos a propriedade a seguir.

▼ **Propriedade da União de Conjuntos Não-Disjuntos:**

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

---

**Prova:** Como $A \cup B = A \cup (B - A)$, e $A$ e $B - A$ são disjuntos, K4 implica que $P(A \cup B) = P(A) + P(B - A)$. E como $B = (A \cap B) \cup (B - A)$, $A \cap B$ e $B - A$ são disjuntos, K4 implica que $P(B) = P(A \cap B) + P(B - A)$. Logo,

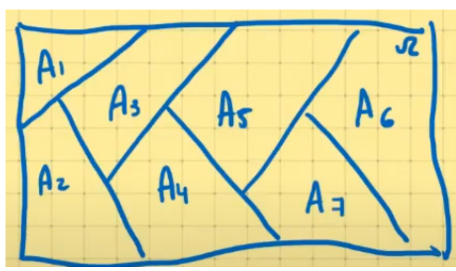
$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

▼ **Propriedade da Probabilidade de Partições**

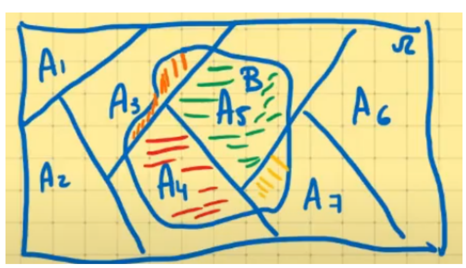
- Se $\{A_i\}$ é uma partição enumerável de $\Omega$ feita de conjuntos em $\mathcal{A}$, então para todo $B \in \mathcal{A}$

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

- Ou seja se o $\Omega$ está particionado em vários $\{A_i\}$




E $B \in \mathcal{A}$




Logo a probabilidade de $B$ vai ser a soma da probabilidade de cada interseção entre $B$ e $\{A_i\}$.

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

- **Prova:** Como $\{A_i\}$ é uma partição, segue que o $B = B \cap \Omega$, como o $\Omega$ é igual a união de todos os eventos da partição $\{A_i\}$, logo o $B = B \cap (\cup_i A_i)$ isso é o mesmo que $B = \cup_i (B \cap A_i)$. Como cada interseção do $B$ com os $\{A_i\}$ vai gerar valores disjuntos e enumeráveis, então usando o axioma k5$'$, teremos a soma das probabilidades das interseções entre $B$ e $\{A_i\}$: $P(B) = \sum_i P(B \cap A_i)$
- A mesma prova mais sendo mais direta:

**Prova:** Como $\{A_i\}$ é uma partição, segue que

$$B = B \cap \Omega = B \cap (\cup_i A_i) = \cup_i (B \cap A_i).$$

O resultado segue então por K5$'$. ∎

▼ **Propriedade da Desigualdade de Boole**

- Para $n$ eventos arbitrários $\{A_1, ..., A_n\}$ disjuntos ou não-disjuntos, a desigualdade de Boole é que a probabilidade da união dos $\{A_i\}$ é menor ou igual a soma das probabilidades de todos os $\{A_i\}$

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

- **Prova:**

---

**Prova:** Provaremos por indução matemática em $n$. A desigualdade é trivialmente verdadeira para $n = 1$ e verdadeira para $n = 2$, pois é uma consequência imediata do Teorema 4.13. Assuma que a desigualdade é válida para $n = k$ e vamos provar que ela é válida para $n = k + 1$. Para ver isto, escrevemos $\cup_{i=1}^{k+1} A_i = A_{k+1} \cup \cup_{i=1}^{k} A_i$.
Pela desigualdade para $n = 2$,

$$P(\cup_{i=1}^{k+1} A_i) \leq P(A_{k+1}) + P(\cup_{i=1}^{k} A_i).$$

Pela hipótese do passo indutivo, para $n = k$,

$$P(\cup_{i=1}^{k} A_i) \leq \sum_{i=1}^{k} P(A_i),$$

portanto, a desigualdade de Boole é verdadeira. ∎

**Corolário**

*Para $n$ eventos arbitrários $\{A_1, \ldots, A_n\}$,*

$$P(\cap A_i) \geq \sum_{i=1}^{n} P(A_i) - (n-1).$$

**Prova:** Utilizando a Lei de De Morgan e a desigualdade de Boole para os eventos $\{A_1^c, \ldots, A_n^c\}$, temos

$$P(\cup_{i=1}^{n} A_i^c) = 1 - P(\cap A_i) \leq \sum_{i=1}^{n} P(A_i^c) = \sum_{i=1}^{n} (1 - P(A_i)).$$

Logo,

$$P(\cap A_i) \geq \sum_{i=1}^{n} P(A_i) - (n-1).$$

$\blacksquare$

### Propriedade da Inclusão-Exclusão

- Esse teorema permite que possamos calcular de maneira exata a probabilidade $P(\cup_{i=1}^{n} A_i)$ para $n$ eventos arbitrários.

**Princípio da Inclusão-Exclusão.** *Seja $I$ um conjunto genérico de índices que é um subconjunto não-vazio qualquer de $\{1, 2, \ldots, n\}$. Para eventos arbitrários $\{A_1, \ldots, A_n\}$,*

$$P(\cup_{i=1}^{n} A_i) = \sum_{\emptyset \neq I \subseteq \{1,\ldots,n\}} (-1)^{||I||+1} P(\cap_{i \in I} A_i),$$

*onde o somatório é sobre todos os $2^n - 1$ conjuntos de índices excluindo apenas o conjunto vazio.*

No caso particular de $n = 3$, o princípio de inclusão-exclusão afirma que

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) &= P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) \\ &- P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3). \end{aligned}$$

**Prova:** A prova é por indução matemática em $n$ e será omitida. $\blacksquare$

## Exemplos

### Exemplo 1:

Se $\Omega$ 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

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

**Provando k1:** Temos que prova que todo evento esta associado a um número real.

$$(\Omega, 2^\Omega, P), \ P : 2^\Omega \to \mathbb{R}, \ \text{tal que}$$
$$P(A) = \frac{||A||}{||\Omega||}, \ \forall A \subset \Omega$$

**Provando k2:** Temos que prova que a probabilidade de A vai ser sempre maior que zero.

$$P(A) = \frac{||A||}{||\Omega||} \geq 0$$
$$\text{Pois o número de elementos nunca vai ser negativo.}$$

**Provando k3:** Temos que prova que a probabilidade do  espaço amostral vai ser sempre igual a 1.

$$P(\Omega) = \frac{||\Omega||}{||\Omega||} = 1$$

**Provando k4:**

$$\text{Se } A \cap B = \varnothing, \ \text{ então } ||A \cup B|| = ||A|| + ||B|| \text{ Logo}$$
$$P(A \cup B) = \frac{||A \cup B||}{||\Omega||} = \frac{||A|| + ||B||}{||\Omega||} = P(A) + P(B)$$

Como esse $\Omega$ é finito, só basta prova os 4 primeiros axiomas de Kolmogorov, para prova que é uma medida de probabilidade.

▼ **Exemplo 2:**

- Se $\Omega = \{a, b, c\}$ , e a álgebra $\mathcal{A}$ é o conjunto das partes de $\Omega$, e a medida de probabilidade $P$ é parcialmente definida por

$$P(\{a, b\}) = 0.5, \ P(\{b, c\}) = 0.8, \ P(\{a, c\}) = 0.7,$$

então complete a especificação de $P$ para todos os eventos em $\mathcal{A}$.

- Para resolver esse problema vamos usar as propriedades da Medida de Probabilidade.
- Primeiro temos que encontra os outros eventos do $2^{A}$(conjunto das partes).

- O enunciado já deu 3 eventos do $2^A$: $\{a, b\}, \{b, c\}, \{a, c\}$.
- Como sabemos quantos eventos tem o $2^A$, basta pegar o 2 e elevar ao número de elementos de $\Omega$, $2^3 = 8$.
- Para saber quais são os eventos que faltam, usaremos as propriedades de $2^A$.
- Os elementos que faltam são o $\Omega, \varnothing$ pois todo $2^A$ tem esses dois eventos. E os eventos unitários de cada elemento de $\Omega$: $\{a\}, \{b\}, \{c\}$.
- Agora vamos calcular as probabilidades desses eventos que encontramos.
- Por definição o $P(\Omega) = 1$ e a $P(\varnothing) = 0$.
- Para calcular a probabilidade dos outros eventos vamos usar a propriedade do complementar, que diz que a $P(A^c) = 1 - P(A)$.
- Então teremos que

  $P(\{a\}) = P(\{b, c\}^c) = 1 - 0.8 = 0.2$

  $P(\{b\}) = P(\{a, c\}^c) = 1 - 0.7 = 0.3$

  $P(\{c\}) = P(\{a, b\}^c) = 1 - 0.5 = 0.5$

**Exemplo 3:**

- Se $\{A_i\}$ for uma partição enumerável de $\Omega$ e $P(A_i) = ab^i$ , $i \geq 1$, então quais as condições que $a$ e $b$ devem satisfazer para que $P$ seja uma medida de probabilidade?
- A primeira propriedade é que o $ab^i$ nunca pode ser negativa. Logo o $a$ e $b$ não podem ser negativos
- Outra propriedade é a seguinte.

$$1 = P(\Omega) = P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i) = \sum_{i=1}^{\infty} a \cdot b^i$$

$$= a \sum_{i=1}^{\infty} b^i = a \cdot \frac{b}{1-b} \Rightarrow a = \frac{1-b}{b}$$

$$|b| < 1$$
$$b \neq 0$$
$$a \neq 0$$

$$\therefore \boxed{a > 0,\ b < 1,\ b > 0,\ a = \frac{1-b}{b}}$$