4. PROBABILIDADE

Fenômeno aleatório: todo acontecimento ou situação, cujos resultados não podem ser previstos com certeza.

Até agora: <u>descrever</u> as principais características de variáveis associadas a fenômenos aleatórios com o objetivo de entendê-los melhor.

Utilizamos: Gráficos, tabelas e distribuições de frequências; medidas de tendência central, de dispersão, de assimetria e de achatamento.

A partir de agora: Procurar associar à distribuição de frequências dos dados observados do fenômeno um **modelo teórico** que reproduza bem tais dados e possibilite quantificar incertezas.

Probabilidade: associada a jogos de azar; ao cálculo de seguros de cargas (fenícios no século XVII) *etc*.

Características de um jogo de azar: *incerteza* e *regularidade*.

Consigo encontrar uma fórmula para ganhar sempre?

4.1. Definições iniciais

Espaço amostral: conjunto de todos os resultados possíveis de um experimento/fenômeno aleatório. **Notação**: $\boldsymbol{S}$ ou $\boldsymbol{\Omega}$ (Omega).

Exemplos:

$a)$ Altura dos alunos da FZEA $\Rightarrow S = \{x \in \mathbb{R} \mid 1{,}40 \leq x \leq 2{,}10\}$

$b)$ Face superior de um dado não viciado $\Rightarrow S = \{1, 2, 3, 4, 5, 6\}$.

Evento: qualquer subconjunto do espaço amostral, $\boldsymbol{\Omega}$..

Exemplos:

$A =$ "altura dos alunos da turma 2020 de Biossistemas" $\Rightarrow A = \{a \in \mathbb{R} \mid 1{,}40 \leq a \leq 1{,}8\}$

$B =$ "a face superior do dado é um número primo" $\Rightarrow B = \{1, 2, 3, 5\}$

Notação ideal: o evento é nomeado por uma letra maiúscula e os seus elementos, por letras minúsculas. Um evento genérico tem a notação: $\mathrm{X} = \{x_1, x_2, \ldots, x_n\}$.

Exemplo: O experimento aleatório envolvendo o lançamento de um dado tem um espaço amostral $S = \{1, 2, 3, 4, 5, 6\}$.

Alguns eventos:

A: a face superior é um número par $\Rightarrow A = \{2, 4, 6\}$

B: a face superior é um número primo $\Rightarrow B = \{1, 2, 3, 5\}$

C: a face superior é um múltiplo de três $\Rightarrow C = \{3, 6\}$

Vamos conhecer algumas operações com eventos:

- A **união** dos eventos A e B é um novo evento, denotado por $A \cup B$, formado pelos elementos que são de A, de B ou de ambos.

  Exemplo:

  $A \cup B = \{1, 2, 3, 4, 5, 6\} = D$: face par **ou** um número primo"

  $A \cup C = \{2, 3, 4, 6\} = E$: "face par **ou** múltiplo de três"

  $B \cup C = \{1, 2, 3, 5, 6\} = F$: "número primo **ou** múltiplo de três"

- A **intersecção** dos eventos A e B é um novo evento, denotado por $A \cap B$, formado pelos elementos que são de A e de B, simultaneamente. **Exemplo:**

  $A \cap B = \{2\} = G$: "face par **e** número primo"

$A \cap C = \{6\} = H$: “face par **e** múltiplo de três”

$B \cap C = \{3\} = I$: “número primo **e** múltiplo de três”

- O **complementar** do evento A em relação ao espaço amostral é um novo evento denotado por A^c, formado por todos os elementos que não são de A. **Exemplo:**

  $A^c = \{1, 3, 5\} = J$: “face ímpar”

  $B^c = \{4, 6\} = L$: “face não é um número primo”

  $C^c = \{1, 2, 4, 5\} = M$: “face não é múltiplo de três”

- A **diferença** entre os eventos A e B é um novo evento, $A–B$, formado pelos elementos de A que não pertencem a B. **Exemplo:**

  $A - B = \{4, 6\} = N$: “face par, mas não é número primo”

$B - A = \{1, 3, 5\} = O$: "face ímpar, mas não é par" = "face ímpar"

$A - C = \{2, 4\} = O$: "face par, mas não é múltiplo de três"

$B - C = \{1, 2, 5\} = P$: "face é número primo, mas não é múltiplo de três".

- Dois eventos A e B são chamados **mutuamente exclusivos** ou disjuntos se não têm pontos amostrais comuns, ou seja, se $A \cap B = \emptyset$.

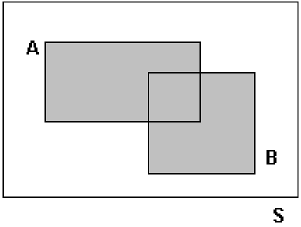
  Exemplo:

  $A \cap B = \{2\} \neq \emptyset$, $A \cap C = \{6\} \neq \emptyset$ e $B \cap C = \{3\} \neq \emptyset \Rightarrow A$ e B, A e C, B e C não são eventos mutuamente exclusivos.

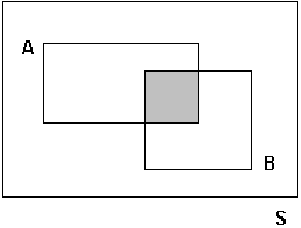
- Dois eventos, A e B, são chamados **complementares** se juntos (unidos) formarem o espaço amostral, ou seja, $A \cup B = S$.

  Exemplo: $A \cup B = \{1, 2, 3, 4, 5, 6\} = S \Rightarrow A$ e B são eventos complementares, ou seja, S é formado por números pares ou primos.

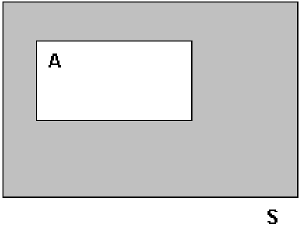
Para visualizar as operações entre eventos usamos os **<u>Diagramas de Venn</u>**



A∪B



A∩B



$\mathbf{A}^{c}$

4.2. PROBABILIDADE

Definição clássica: Suponha que um evento A possa ocorrer de k maneiras diferentes num total de n maneiras possíveis e igualmente prováveis. Então, a $probabilidade$ de ocorrência do evento A é k/n, definida como a **frequência relativa** do evento A.

Definição moderna (axiomática): Uma função $P(.)$ é denominada $probabilidade$ se satisfaz as condições:

$i)$ $0 \leq P(A) \leq 1$, para $\forall$ evento $A \subset \Omega$

$ii)$ $P(\Omega) = 1$

$iii)$ $P\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} P(A_j)$, quando os A_j's disjuntos.

Podemos atribuir probabilidades aos eventos com base em:

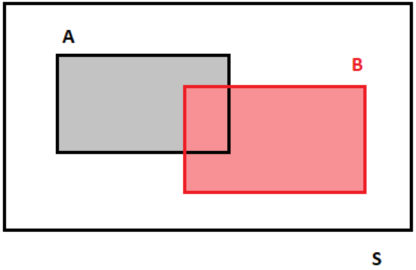
- **Características teóricas** da realização do fenômeno.

  Exemplo: a probabilidade de ocorrer uma das faces no lançamento de um dado é 1/6; no lançamento de uma moeda $P(cara) = P(coroa) = 1/2$

- **Frequência relativa** observada do evento em diversas repetições do fenômeno em que pode ocorrer o evento de interesse.

  Exemplo: acompanhar diversos partos para calcular a probabilidade de o primeiro leitão nascido vivo ser uma fêmea.

Probabilidade da união de eventos: Sejam A e B dois eventos de S.



A probabilidade de ocorrência do evento A **ou** B é

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

Exemplo: Usando os eventos do exemplo dos dados temos $P(A) = 1/2$, $P(B) = 2/3$ e $P(A \cap B) = 1/6$. Então:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{2} + \frac{2}{3} - \frac{1}{6} = \frac{3+4-1}{6} = 1$$

Confirmando a afirmação que A e B são eventos exaustivos.

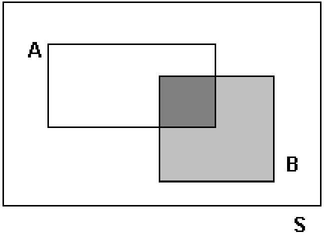
4.3. Probabilidade condicional

Situação: O fenômeno aleatório acontece em etapas e a informação do que ocorreu em uma etapa pode influenciar na probabilidade de ocorrências em outras etapas.

Exemplo: A probabilidade de ganhar na megasena com um jogo simples é $1/50.063.860 \cong 0{,}00000002$. Essa probabilidade pode aumentar se eu souber que a aposta vencedora é do estado de São Paulo, de Pirassununga, da lotérica da avenida...

Definição 4.1. Para dois eventos A e B, com $P(B) > 0$, a probabilidade do evento A ocorrer, dado que o evento B já ocorreu **ou** a probabilidade condicional de A dado B, é definida por:

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



Perceba que:

A ocorrência do evento B diminui o espaço amostral e a probabilidade de A é calculada neste novo espaço amostral.

- Sempre que os eventos A e B não são disjuntos ($A \cap B \neq \Phi$) a ocorrência de B, altera a probabilidade de ocorrência do evento A.

Um evento B é dito **independente** do evento A, se a probabilidade de B ocorrer não é influenciada pelo fato de A já ter ocorrido, isto é,

$$A \text{ é independente de } B \Leftrightarrow P(B) = P(B|A) \text{ ou } P(A) = P(A|B)$$

Usando a fórmula da probabilidade condicional, podemos calcular a probabilidade de ocorrência simultânea de dois eventos:

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

ou

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

Se A e B são dois eventos **independentes** a fórmula simplifica para:

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

Vamos calcular algumas probabilidades com base em um estudo envolvendo 1000 estudantes de Pirassununga, classificados pela área de estudo e a classe socioeconômica da sua família.

Exemplo 4.1. Dados (fictícios) sobre a área de estudo e a classe sócio econômica de 1000 estudantes de Pirassununga.

	Classe socioeconômica			
Área	Alta	Média	Baixa	Total
Exatas	120	156	68	344
Humanas	72	85	112	269
Biológicas	169	145	73	387
Total	361	386	253	1000

Determine a probabilidade de um estudante <u>escolhido ao acaso</u>:

a) Ser da classe econômica mais alta.

b) Estudar na área de exatas.

c) Estudar na área de humanas e ser da classe média.

d) Ser da classe baixa, sabendo-se que estuda na área de biológicas.

e) Ser da classe média ou estudar na área de exatas.

f) Estudar na área de exatas, sabendo-se que é da classe alta.

g) Será que a escolha de um estudante de Pirassununga pela área de Exatas **depende** da classe socioeconômica? E a escolha pela área de Humanas? E pela área de Biológicas?

Solução:

a) $P(\text{Alta}) = 361/1000 = 0{,}361$

b) $P(\text{Exatas}) = 344/1000 = 0{,}344$

c) $P(\text{Humanas} \cap \text{Média}) = 85/1000 = 0{,}085$

d) $P(\text{Baixa} \mid \text{Biológicas}) = 73/387 = 0{,}189$

e) $P(\text{Média} \cup \text{Exatas}) = \frac{386}{1000} + \frac{344}{1000} - \frac{156}{1000} = 0{,}574$

f) $P(\text{Exatas} \mid \text{Alta}) = \frac{120}{361} = 0{,}332$

g) $P(\text{Exatas}|\text{Alta}) = 0{,}332$ $\qquad$ $P(\text{Exatas}|\text{Média}) = \frac{156}{386} = 0{,}404$

$P(\text{Exatas}|\text{Baixa}) = \frac{68}{253} = 0{,}269$ $\qquad$ $P(\text{Exatas}) = 0{,}344$

⇒ concluímos que a escolha pela área de Exatas depende da classe socioeconômica do estudante.

O mesmo acontece com as áreas de Humanas e Biológicas (verifique!)

Exemplo 4.2. Consideremos três baias da granja de suínos com as características:

Baia 1: tem 10 leitões, 4 dos quais já foram vacinados

Baia 2: tem 6 leitões, 1 dos quais já foi vacinado

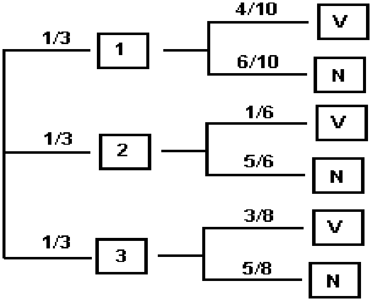
Baia 3: tem 8 leitões, 3 dos quais já foram vacinados

O experimento consiste de duas etapas: sortear uma das três baias e desta baia escolhida sortear um leitão. Pergunta-se:

a) Qual é a probabilidade deste leitão sorteado já estar vacinado?

b) Qual é a probabilidade deste leitão sorteado ser da baia 1, sabendo-se que ele já foi vacinado? E da baia 2? E da baia 3?

$Dica$: Construir um diagrama de árvore.



Evento	Probabilidade
$1 \cap V$	$(1/3)(4/10) = \ \ 48/360 \cong 0{,}1333$
$1 \cap N$	$(1/3)(6/10) = \ \ 72/360 \cong 0{,}2000$
$2 \cap V$	$(1/3)(1/6) \ = \ \ 20/360 \cong 0{,}0556$
$2 \cap N$	$(1/3)(5/6) \ = 100/360 \cong 0{,}2778$
$3 \cap V$	$(1/3)(3/8) \ = \ \ 45/360 \cong 0{,}1250$
$3 \cap N$	$(1/3)(5/8) \ = \ \ 75/360 \cong 0{,}2083$

$a)$ $\mathrm{P}(1 \cap \mathrm{V}) = \mathrm{P}(1)\mathrm{P}(\mathrm{V} \mid 1) = (1/3)(4/10) = 48/360 \cong 0{,}1333.$

$$\mathrm{P}(\mathrm{V}) = \mathrm{P}(1 \cap \mathrm{V}) + \mathrm{P}(2 \cap \mathrm{V}) + \mathrm{P}(3 \cap \mathrm{V})$$

$$\mathrm{P}(\mathrm{V}) = (1/3)(4/10) + (1/3)(1/6) + (1/3)(3/8)$$

$$= 113/360 \Rightarrow \mathrm{P}(\mathrm{V}) \cong 0{,}3139 \qquad \text{e} \ \ \mathrm{P}(\mathrm{N}) = 1 - \mathrm{P}(\mathrm{V}) \cong 0{,}6861$$

- Note que P(V) $\neq \frac{8}{24} = \frac{\#\,vacinados}{\#\,total} = 0{,}3333$, porque o ensaio é realizado em duas fases e as baias têm números diferentes de leitões.

$$b)\ P(1|V) = \frac{P(1 \cap V)}{P(V)} = \frac{48/360}{113/360} = \frac{48}{113} = 0{,}4248$$

$$P(2|V) = \frac{P(2 \cap V)}{P(V)} = \frac{20/360}{113/360} = \frac{20}{113} = 0{,}1770$$

$$P(3|V) = \frac{P(3 \cap V)}{P(V)} = \frac{45/360}{113/360} = \frac{45}{113} = 0{,}3982$$

Esta forma de resolver o item (b) está associada ao Teorema de Bayes, que será conhecido com detalhes a seguir.

O teorema mostra como alterar as probabilidades a $priori$ tendo em vista novas evidências para obter probabilidades a $posteriori$.