Preditiva.ai

Probabilidades
Introdução

Probabilidades

Introdução

Em nossas vidas temos basicamente dois tipos de fenômenos: **Determinísticos** e **Aleatórios**.

Fenômenos Determinísticos

Quando temos **apenas um resultado** possível como consequência de uma ação.

Exemplos:

- Lucro de uma empresa
- Fatura de cartão de crédito
- Água aquecida a 100º C no nível do mar entra em ebulição
- Entre outros...

Fenômenos Aleatórios

Quando temos **mais de um resultado** possível como consequência de uma ação.

Exemplos:

- Previsão do tempo
- Resultado de um campeonato
- Volume de vendas no mês que vem
- Pessoas que pedirão demissão
- Ações que serão julgadas improcedentes
- Efetividade de um tratamento médico
- Entre outros...

Probabilidades
Introdução

Preditiva.ai

Fenômenos **aleatórios** são muito recorrentes na natureza. Logo, estudar técnicas que "quantificam" sua **probabilidade de ocorrência** é fundamental.

A teoria matemática utilizada para se **estudar a incerteza oriunda de fenômenos de caráter aleatório** é denominada **Probabilidade.**

Existem várias abordagens para estudar probabilidades, sendo as 3 principais:

- Teoria Clássica
- Teoria Frequentista
- Teoria Axiomática

Probabilidades
Teoria Clássica

Preditiva.ai

Probabilidades
Teoria Clássica

Em 2019 o IBGE divulgou um estudo no qual apresentou a quantidade de **homens** e **mulheres** **com** e **sem restrição** de acesso à Educação[1]. Um resumo desses dados é apresentado abaixo (em milhares).

Gênero	Com Restrição (C)	Sem Restrição (S)	Total
Mulheres (M)	28.708	78.813	**107.521**
Homens (H)	28.695	71.638	**100.333**
Total	**57.368**	**150.486**	**207.854**

Sorteando uma pessoa **aleatoriamente**, quais são os **possíveis resultados** que podemos obter considerando **Gênero** e **Restrição à Educação**?

Fonte: (1) https://biblioteca.ibge.gov.br/visualizacao/livros/liv101678.pdf

Probabilidades

Teoria Clássica – Espaço Amostral

Preditiva.ai

Em 2019 o IBGE divulgou um estudo no qual apresentou a quantidade de **homens** e **mulheres** **com** e **sem** **restrição** de acesso à Educação[1]. Um resumo desses dados é apresentado abaixo (em milhares).

Gênero	Com Restrição (C)	Sem Restrição (S)	Total
Mulheres (M)	28.708	78.813	**107.521**
Homens (H)	28.695	71.638	**100.333**
Total	**57.368**	**150.486**	**207.854**

Mulher e Com Restrição (M,C)

Mulher e Sem Restrição (M,S)

Homem e Com Restrição (H,C)

Homem e Sem Restrição (H,S)

O conjunto de **todos os possíveis resultados** de um fenômeno de caráter aleatório é chamado de **Espaço Amostral**.

O **Espaço Amostral** deste estudo é definido como: {**(M,C),(M,S),(H,C),(H,S)**}

Probabilidades

Teoria Clássica – Evento

Podemos estar interessados em subconjuntos do Espaço Amostral:

- **Evento A** - Mulher: **{(M,C),(M,S)}**
- **Evento B** - Homem e Com Restrição: **{(H,C)}**

Espaço Amostral

{(M,C),(M,S),(H,C),(H,S)}

Gênero	Com Restrição (C)	Sem Restrição (S)	Total
Mulheres (M)	28.708	78.813	**107.521**
Homens (H)	28.695	71.638	**100.333**
Total	**57.368**	**150.486**	**207.854**

Mulher (M)

Homem e Com Restrição (H,C)

Qualquer **subconjunto** do Espaço Amostral de um fenômeno de caráter aleatório é chamado de Evento.

Probabilidades

Teoria Clássica

Em qualquer teoria de probabilidades, deseja-se **atribuir um valor à um evento do espaço amostral**.

Na **teoria clássica**, consideramos que **todos os eventos têm a mesma chance** de ocorrência.

$$P(evento\ de\ interesse) = \frac{Tamanho\ do\ evento\ de\ interesse}{Tamanho\ do\ espaço\ amostral}$$

(Leia-se “Probabilidade do evento de interesse”)

Probabilidades
Teoria Clássica – Evento

Sorteando uma pessoa **aleatoriamente**, qual a **probabilidade de ocorrência do Evento A**? E do **Evento B**?

Gênero	Com Restrição (C)	Sem Restrição (S)	Total
Mulheres (M)	28.708	78.813	**107.521**
Homens (H)	28.695	71.638	**100.333**
Total	**57.368**	**150.486**	**207.854**

Evento A - Mulher: **{(M,C),(M,S)}**

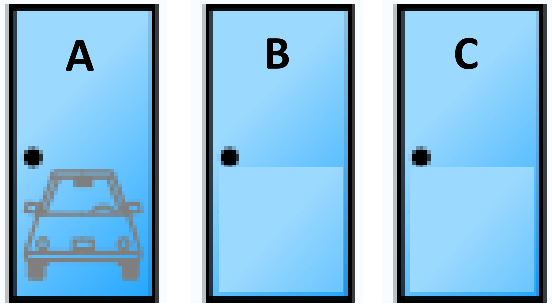
$$P(\text{Mulher}) = \frac{\boxed{28.708} + \boxed{78.813}}{\boxed{207.854}} = 51{,}7\%$$

Evento B - Homem e Com Restrição: **{(H,C)}**

$$P(\text{Homem e Com Restrição}) = \frac{\boxed{28.695}}{\boxed{207.854}} = 13{,}8\%$$

Problema de **Monty Hall** ou **"Porta da Esperança"**

Neste famoso jogo, o jogador escolhe uma porta buscando o prêmio por trás dela.
O apresentador abre uma das outras duas portas e pergunta ao jogador: "Você **quer ficar** com a mesma porta ou **quer trocar**?"

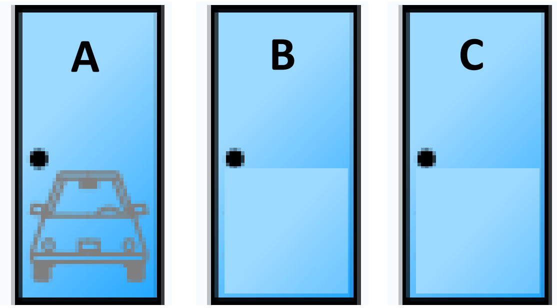


Probabilidades

Teoria Clássica

Preditiva.ai

1ª Possibilidade

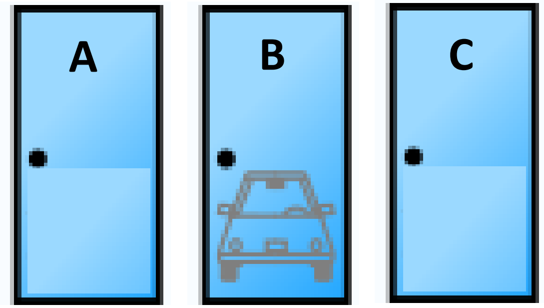


1. O jogador escolhe: **A**
2. O apresentador abre: **B** ou **C**
3. Se jogador **trocar**?: **Perde**

1. O jogador escolhe: **B**
2. O apresentador abre: **C**
3. Se jogador **trocar**?: **Ganha**

1. O jogador escolhe: **C**
2. O apresentador abre: **B**
3. Se jogador **trocar**?: **Ganha**

2ª Possibilidade

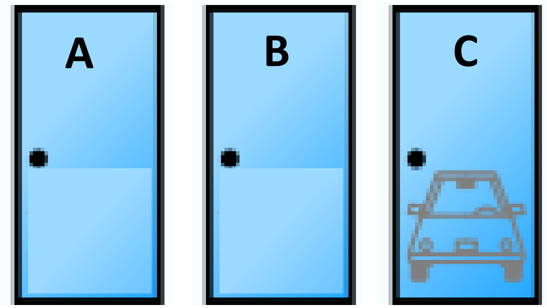


1. O jogador escolhe: **A**
2. O apresentador abre: **C**
3. Se jogador **trocar**?: **Ganha**

1. O jogador escolhe: **B**
2. O apresentador abre: **A** ou **C**
3. Se jogador **trocar**?: **Perde**

1. O jogador escolhe: **C**
2. O apresentador abre: **A**
3. Se jogador **trocar**?: **Ganha**

3ª Possibilidade



1. O jogador escolhe: **A**
2. O apresentador abre: **B**
3. Se jogador **trocar**?: **Ganha**

1. O jogador escolhe: **B**
2. O apresentador abre: **A**
3. Se jogador **trocar**?: **Ganha**

1. O jogador escolhe: **C**
2. O apresentador abre: **A** ou **B**
3. Se jogador **trocar**?: **Perde**

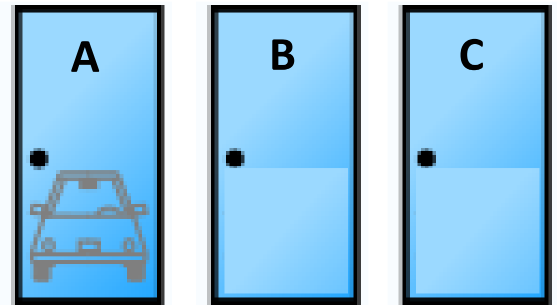
P(Ganhar) = 2/3 = **66%**

Probabilidades

Teoria Clássica

Preditiva.ai

1ª Possibilidade

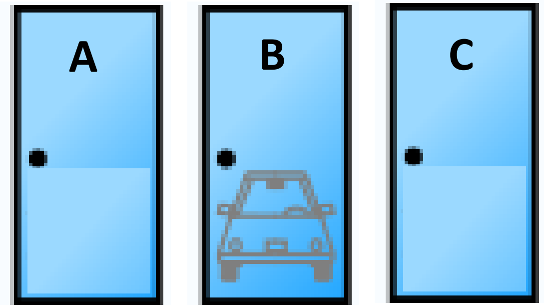


1. O jogador escolhe: **A**
2. O apresentador abre: **B** ou **C**
3. Se jogador **ficar**?: **Ganha**

1. O jogador escolhe: **B**
2. O apresentador abre: **C**
3. Se jogador **ficar**?: **Perde**

1. O jogador escolhe: **C**
2. O apresentador abre: **B**
3. Se jogador **ficar**?: **Perde**

2ª Possibilidade

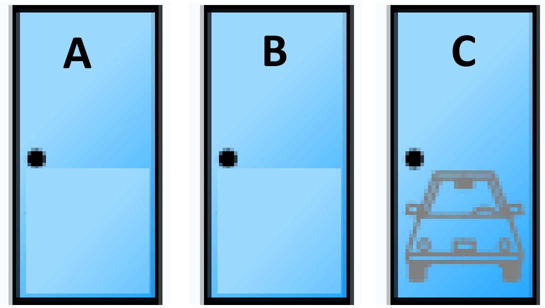


1. O jogador escolhe: **A**
2. O apresentador abre: **C**
3. Se jogador **ficar**?: **Perde**

1. O jogador escolhe: **B**
2. O apresentador abre: **A** ou **C**
3. Se jogador **ficar**?: **Ganha**

1. O jogador escolhe: **C**
2. O apresentador abre: **A**
3. Se jogador **ficar**?: **Perde**

3ª Possibilidade



1. O jogador escolhe: **A**
2. O apresentador abre: **B**
3. Se jogador **ficar**?: **Perde**

1. O jogador escolhe: **B**
2. O apresentador abre: **A**
3. Se jogador **ficar**?: **Perde**

1. O jogador escolhe: **C**
2. O apresentador abre: **A** ou **B**
3. Se jogador **ficar**?: **Ganha**

P(Ganhar) = 1/3 = **33%**

Revisão

Vimos que a **Teoria Clássica de Probabilidades** tem como características:

- Todos os eventos têm a mesma probabilidade de ocorrência
- A probabilidade do evento de interesse é igual ao tamanho do evento de interesse dividido pelo tamanho do espaço amostral.

Como um exemplo, vimos que é bastante trivial resolver o **Problema de Monty Hall** usando a **Teoria Clássica**.

Probabilidades
Teoria Frequentista

Probabilidades
Teoria Frequentista

Estudar Probabilidades sob a **teoria frequentista**, significa basear-se em um número suficiente de experimentos e **anotar a frequência de ocorrência** do evento de interesse.

Imagine que você gostaria de saber qual é a "**taxa de compradores**" de uma loja.

Qual experimento poderia ser realizado para verificar?

Probabilidades

Teoria Frequentista

Preditiva.ai

Experimento: Observar as pessoas que entram na loja e anotar quantos deles realizam compras.

Conclusão: A probabilidade empírica, ou seja, a frequência relativa de Compradores **é de cerca de 40%.**

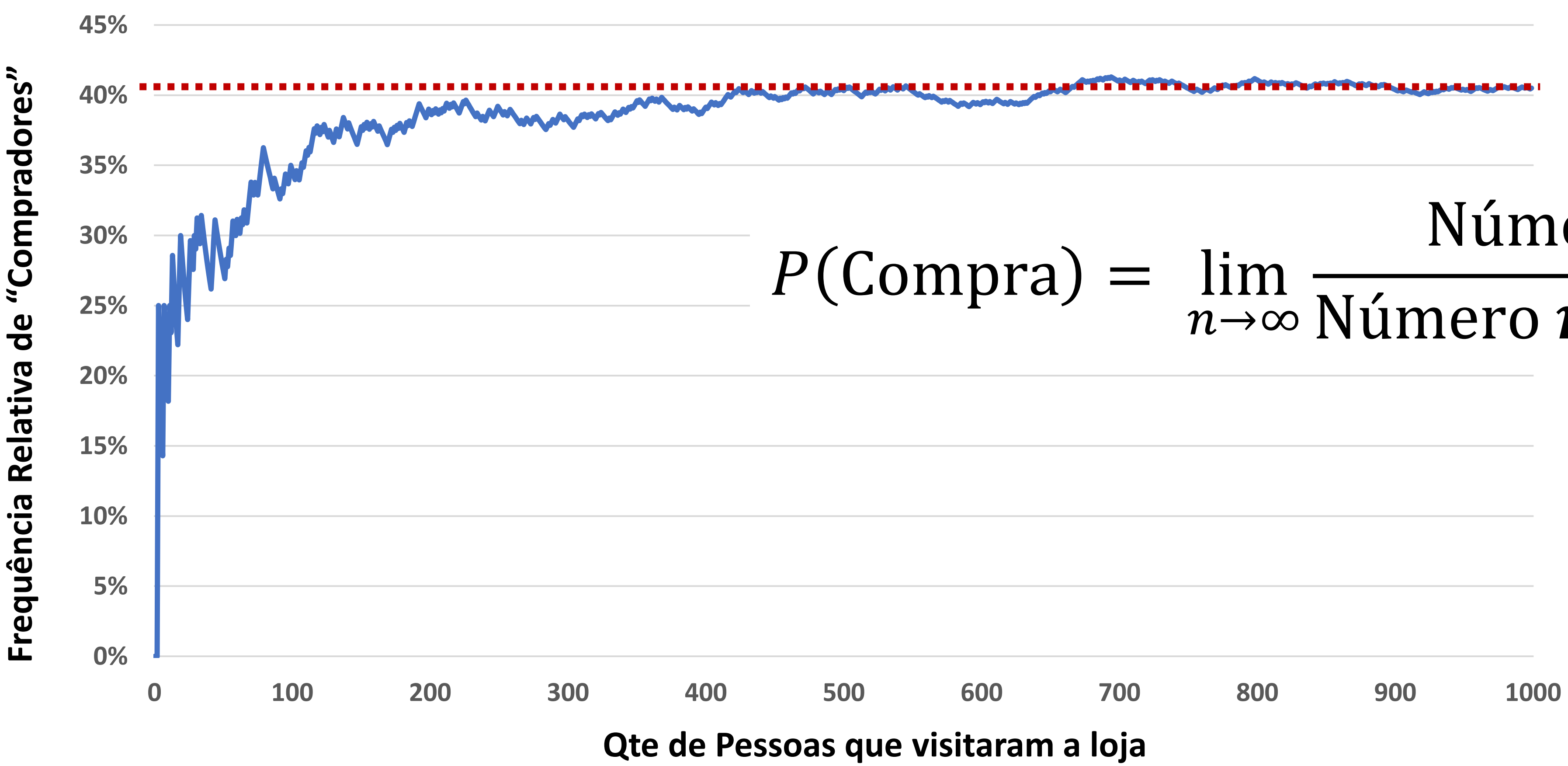
Portanto, podemos dizer que quando uma pessoa entra na loja, a probabilidade dela realizar uma compra é de cerca de **40%**.

Qte de Pessoas que visitaram a loja	Comprou	Número de Compradores	Frequência Relativa de Compradores
1	Não	0	0,0%
2	Não	0	0,0%
3	Não	0	0,0%
4	Sim	1	25,0%
5	Não	1	20,0%
...	...	...	...
30	Sim	9	30,0%
...	...	...	...
50	Não	14	28,0%
...	...	...	...
100	Sim	35	35,0%
...	...	...	...
500	Não	202	40,4%
...	...	...	...
998	Não	404	40,5%
999	Não	404	40,4%
1000	Sim	405	40,5%

Probabilidades

Teoria Frequentista

O ponto central da **Teoria Frequentista** é que a **Probabilidade** do evento de interesse é **confiável** ao se obter um **número suficiente de experimentos**.



Ou seja, em termos formais,

$$P(\text{Compra}) = \lim_{n \to \infty} \frac{\text{Número de Compras}}{\text{Número } \boldsymbol{n} \text{ de pessoas na loja}}$$

Probabilidades

Teoria Frequentista - Exemplos

Preditiva.ai

Suponha que você estime a probabilidade de 0,000001% de defeito em um processo qualquer.

No entanto, a probabilidade real de defeito é 0,01%. Você acha que realizou uma boa estimativa?

Nem sempre, pois com essa **diferença** teremos como **resultado:**

- 12 recém nascidos serão entregues para pais errados por dia
- Correio entregaria 30.000 encomendas erroneamente por dia
- 2,5 milhões de livros seriam entregues com as capas erradas ao ano
- Dois voos por dia pousariam mal só no aeroporto de Chicago
- 210 palavras do dicionário Aurélio estariam soletradas incorretamente
- 20.000 prescrições médicas estariam erradas por ano
- 880.000 cartões de crédito estariam com defeito no mundo
- 291 operações de marca-passo seriam feitas de forma errada por ano

Probabilidades

Teoria Fr... Exemplos

Suponha q

No entant

Nem semp

- 12 re
- Corr
- 2,5 r
- Dois
- 210
- 20.0
- 880.
- 291

Muitas vezes, cometemos o **erro de assumir como verdade uma probabilidade** sem antes verificar o experimento realizado para chegar à essa estimativa.

Como vimos, só podemos ter mais certeza do número quando temos uma quantidade N de experimentos “suficiente”. Mas qual N é suficiente? Estudaremos uma técnica adequada para estimar esse valor mais adiante.

Por enquanto, sempre tenha em mente o seguinte: **Qualquer probabilidade só é confiável quando temos um número SUFICIENTE de experimentos.**

Revisão

Vimos que a **Teoria Frequentista de Probabilidades** se baseia no fato de que uma **estimativa de probabilidade** só é robusta quando temos um **número suficiente** de experimentos.

Pequenas desvios na estimação da probabilidade pode ter efeitos relevantes, considerando sua **frequência de realização**.

Preditiva.ai

Probabilidades

Teoria Axiomática

O que você verá nessa aula?

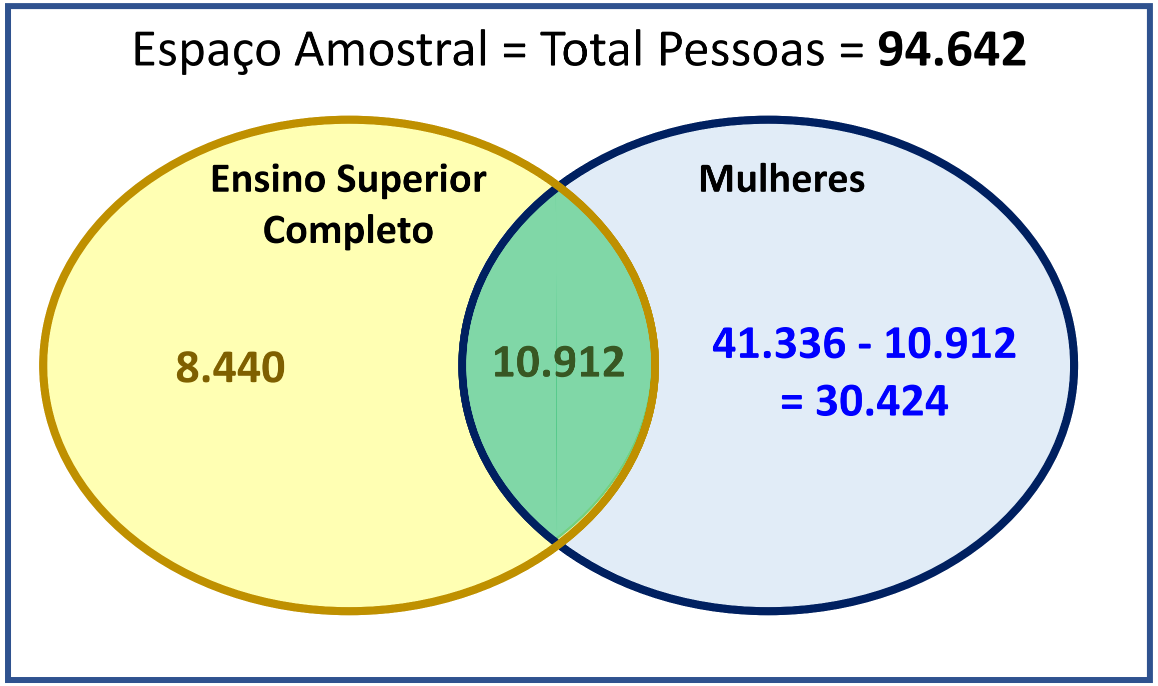
- ❑ Introdução a Probabilidade
- ❑ Teoria Clássica
- ❑ Teoria Frequentista
- ❑ Teoria Axiomática
- ❑ Probabilidades: uma arma poderosa

Probabilidades

Teoria Axiomática

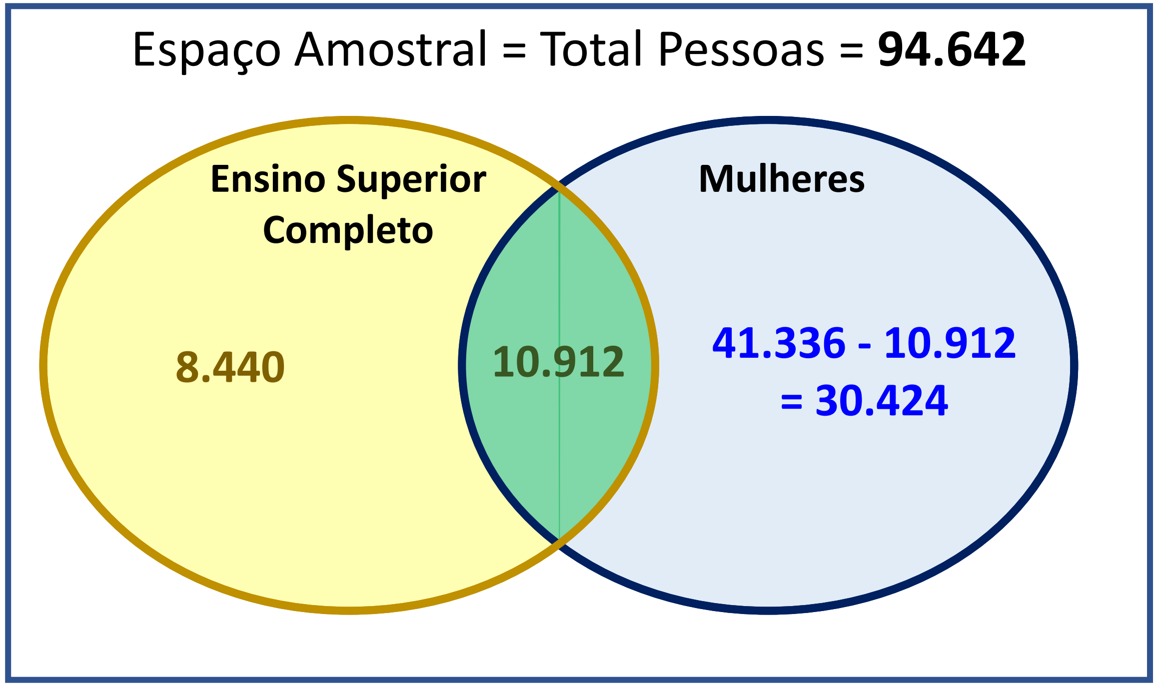
O IBGE divulgou as estatísticas de 2019 sobre nível de instrução para pessoas de 14 anos ou mais, separadas por gênero:

Categorias	Nº Pessoas (mil)
Mulheres	**41.336**
Ensino Superior Completo	**10.912**
Homens	53.306
Ensino Superior Completo	**8.440**
Total	**94.642**



Probabilidades
Teoria Axiomática

Considerando o **Diagrama de Venn** construído, vamos então calcular as probabilidades dos seguintes eventos:



Evento A: Ensino Superior Completo

$$P(\text{Ensino Superior Completo}) = \frac{8.440 + 10.912}{94.642} = 20{,}4\%$$

Evento B: Mulheres

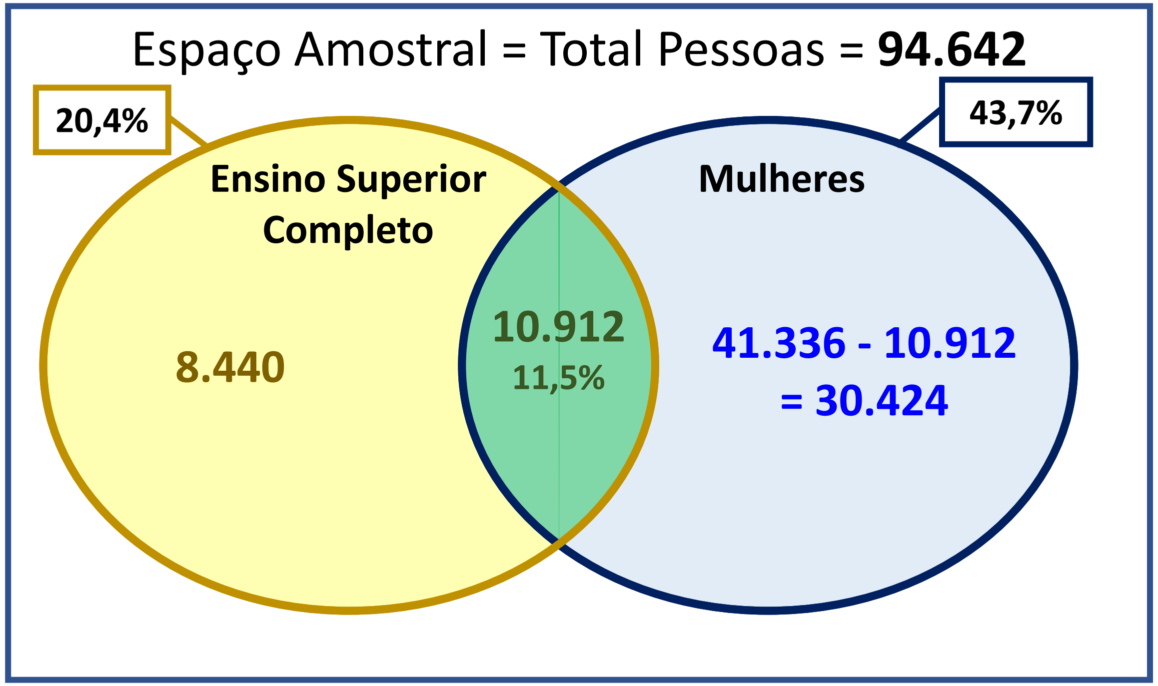
$$P(\text{Mulheres}) = \frac{10.912 + 30.424}{94.642} = 43{,}7\%$$

Evento C: Ensino Superior Completo e Mulheres

$$P(\text{Ens. Superior} \cap \text{Mulheres}) = \frac{10.912}{94.642} = 11{,}5\%$$

Probabilidades
Teoria Axiomática

Considerando o diagrama de Venn construído, vamos então calcular as probabilidades dos seguintes eventos:



Evento D: Superior Completo **ou** Mulheres

$$P(\text{Superior} \cup \text{Mulheres}) =$$

$$P(\text{Superior}) + P(\text{Mulheres}) - P(\text{Superior} \cap \text{Mulheres}) =$$

$$20{,}4\% + 43{,}7\% - 11{,}5\% = 52{,}6\%$$

Evento E: Mulheres **ou** Homens

$$P(\text{Mulheres} \cup \text{Homens}) =$$

$$P(\text{Mulheres}) + P(\text{Homens}) =$$

$$43{,}7\% + \frac{53.306}{94.642} = 100\%$$

Probabilidades

Teoria Axiomática

Preditiva.ai

Embora a teoria Frequentista já resolva grande parte dos problemas para se estimar a probabilidade de um evento, pode-se **argumentar que não é possível garantir que essa probabilidade seja a mesma** em uma segunda ou terceira repetição do experimento.

Para eliminar esta dúvida, outra forma de estudar probabilidades é utilizando a **Teoria Axiomática**. Essa teoria baseia-se nos 4 axiomas (premissas incontestáveis) listados a seguir:

1) $\mathbf{0 \leq P(\text{Evento}) \leq 1}$	A probabilidade de **um evento** é um número entre 0 e 1.
2) $\boldsymbol{P}\ (\textbf{Espaço Amostral}) = \mathbf{1}$	A probabilidade de ocorrência de **pelo menos um dos eventos** do espaço amostral é igual a 1.
3) $\boldsymbol{P\ (A + B) = P(A) + P(B)}$	A probabilidade da **união dos eventos A e B** é igual a **soma das probabilidades de A e B, se e somente se A e B não tiverem chance de ocorrer simultaneamente**. Ex.: Não é possível sair cara e coroa simultaneamente.
4) $\boldsymbol{P\ (A + B) = P(A) + P(B) - \mathrm{P}(\mathrm{A} \cap B)}$	A probabilidade da **união dos eventos A e B** é igual a **soma das probabilidades de A e B menos a probabilidade de ocorrência de A e B simultaneamente**.

Probabilidades

Teoria Axiomática

Preditiva.ai

Outra regra muito relevante é a **regra de multiplicação de probabilidades**:

$$P\ (\text{Evento } A \cap B) = P(\ A\ ) * P(\ B\ )$$

Importante: Esta regra só é verdadeira **quando A e B são independentes**. Ou seja, quando a ocorrência de A não depende da ocorrência de B.

Exemplo: Qual a probabilidade de lançar **duas vezes** uma moeda e sair *cara* nas duas vezes ?

$$P\ (\text{Cara} \cap \text{Cara}) = P(\ \text{Cara}\ ) * P(\ \text{Cara}\ ) = \tfrac{1}{2} * \tfrac{1}{2} = \tfrac{1}{4} = 25\%$$

Revisão

Vimos que a **Teoria Axiomática de Probabilidades** se utiliza de 4 premissas incontestáveis para o cálculo das probabilidade.

A **multiplicação de probabilidades** só pode ser realizada em eventos que sejam **independentes**.

Probabilidades
Uma arma poderosa

“ A probabilidade **nunca** erra. Quem erra são as pessoas que **usam** a probabilidade...”

Autor desconhecido

Probabilidades

Uma arma poderosa

Preditiva.ai

Quando **mal usada**, a probabilidade pode causar **grandes problemas**. Vejamos alguns erros comuns:

Erro 1) Não entender a diferença dos conceitos a seguir:

Evento possível		**Evento provável**
❑ Arrumar um bom emprego $$$ sem formação.	✖	❑ As melhores oportunidades ficam com os bem formados.
❑ Ganhar na loteria!!!		❑ Só ganhar dívidas e frustração.
❑ A TV pifar durante o período de garantia estendida.		❑ Funcionar durante anos e anos...
❑ Um avião cair no mar e todos morrerem.		❑ Morrer no trânsito caótico de SP.

Probabilidades

Uma arma poderosa

Preditiva.ai

Você tem **quase 2.000 vezes mais chance de morrer em um acidente de carro** do que um de aviação*.

Causa da morte - EUA 2017	Probabilidade
Doença cardíaca	16,6667%
Câncer	14,2857%
Doença Respiratória Inferior Crônica	3,7037%
Suicídio	1,1364%
Overdose de drogas	1,0417%
Acidente de carro	**0,9709%**
Cair	0,8772%
Bala perdida	0,3509%
Incidente de pedestres	0,1799%
Acidente de moto	0,1166%
Afogamento	0,0895%
Fogo ou fumaça	0,0678%
Intoxicação	0,0371%
Acidente de bicicleta	0,0247%
Acidente com armas	0,0117%
Insolação	0,0112%
Eletrocussão, Radiação, Temperaturas Extremas e Pressão	0,0064%
Objetos pontiagudos	0,0036%
Tempestade Cataclísmica	0,0032%
Superfícies e substâncias quentes	0,0022%
Picadas de vespas e abelhas	0,0021%
Ataque de cachorro	0,0009%
Passageiros em um avião	**0,0005%**
Relâmpago	0,0005%
Acidente de trem	0,0004%

Estima-se que mais de 2.170 americanos morreram de acidentes de carro após 11/9 por terem deixado de voar de avião.**

* https://injuryfacts.nsc.org/all-injuries/preventable-death-overview/odds-of-dying/

** http://blalock.dyson.cornell.edu/wp/fatalities_120505.pdf

Preditiva.ai

Esse cálculo está correto?

A probabilidade de **um dos motores** de um avião falhar é de **1 em 100 mil**.
Portanto, a probabilidade dos **2 falharem ao mesmo tempo** é de **1 em 10 Bilhões**, ou seja:

$$\frac{1}{100 \text{ mil}} \cdot \frac{1}{100 \text{ mil}} = \frac{1}{10 \text{ bilhões}}$$

Resposta: Não!

Motivo:
Erro 2) Pressupor que eventos **sejam** independentes **quando não são.**

Preditiva.ai

O que acha desse caso?

A **morte súbita de bebês é estimada: 1 em 8.500**. Quando mais de um caso como este acontecia na mesma família, um pediatra famoso chamado Roy Meadow era chamado para depor. Segundo Meadow, a **chance de 2 bebês morrerem na mesma família era de 1 em 73 milhões**, ou seja:

$$\frac{1}{8.500} \cdot \frac{1}{8.500} = \frac{1}{73 \text{ milhões}}$$

Com base nesse argumento, os júris colocaram muitos pais na cadeia no Reino Unido*.

Mais um caso de pressupor que eventos **sejam** independentes **quando não são.**

Após amplo debate, o governo britânico **anunciou que reveria 258 julgamentos** nos quais os pais foram condenados por assassinar seus bebês.

* https://www.economist.com/leaders/2004/01/22/the-probability-of-injustice

Probabilidades

Uma arma poderosa

Preditiva.ai

Agora imagine a seguinte situação:

Um jogador de basquete acerta uma cesta, depois mais outra. A probabilidade de ele acertar a terceira cesta aumenta. **Verdadeiro ou falso?**

Uma **pesquisa* mostrou que 91% dos fãs** de basquete **acreditam na "mão quente"** e acham a afirmação acima verdadeira.

Infelizmente, essa sequência de acertos **não tem comprovação estatística.**

Motivo:

Erro 3) Pressupor que eventos **não sejam** independentes **quando de fato são.**

Esse é o porquê da "banca sempre ganhar" em um cassino.

* https://www.economist.com/leaders/2004/01/22/the-probability-of-injustice

Preditiva.ai