

Análise Estatística de dados

Inteligência Artificial



AULA 07 — PROBABILIDADE

Arturo Forner-Cordero
Larissa Driemeier

PROGRAMA DO CURSO

Aula	Data	Conteúdo da Aula
01	27/02	Aula Inaugural
02	05/03	Introdução ao Curso. Noções de Álgebra Linear , Geometria Analítica Parte I
03	12/03	Noções de Álgebra Linear , Geometria Analítica Parte II
04	19/03	Decomposição de valor singular (SVD)
05	26/03	Otimização: derivadas, derivadas parciais (operadores gradiente, Jacobiano, Hessiano e Laplaciano), algoritmos de gradiente
06	02/04	Variáveis independentes e não independentes. Estatística Descritiva e Indutiva. Definições de medidas de dispersão e tendência central.
07	09/04	Modelos de probabilidade I
08	16/04	Modelos de probabilidade II
09	23/04	Teoria da Informação
10	30/04	Modelo de Markov. Modelo de Markov oculto.



BENJAMIN FRANKLIN ONCE
SAID THAT TWO THINGS IN LIFE
ARE CERTAIN: DEATH AND
TAXES.

The remaining parts of life are
least predictable.

RELEMBRANDO...

Probabilidade é um conceito filosófico e matemático que permite a quantificação da incerteza. Dessa maneira, ela pode ser aferida, analisada e usada para a realização de previsões ou para a orientação de intervenções.

É a **probabilidade** que torna possível lidar de forma racional com problemas envolvendo o imprevisível.

QUAL É A VANTAGEM DA PROBABILIDADE E ESTATÍSTICA?

Uma de suas aplicações mais importantes é a tomada de decisão sob incerteza.

Quando decidimos sobre uma ação, estamos apostando que concluir a ação nos deixará melhor do que se não a tivéssemos feito.

Mas as apostas são inerentemente incertas, então como você decide se vai continuar com a ação ou não?

Implícita ou explicitamente, você estima uma probabilidade de sucesso - e se a probabilidade for maior que algum limite, você avança.



ONDE QUEREMOS CHEGAR HOJE?

Portanto, ser capaz de estimar com precisão essa probabilidade de sucesso é fundamental para tomar boas decisões.

Embora o acaso sempre tenha um papel importante no resultado, se você puder unir consistentemente as probabilidades a seu favor, deverá se sair muito bem ao longo do tempo.



ALGUMAS DEFINIÇÕES

- que é variável aleatória?
- Como se define a probabilidade?
- que é um modelo probabilístico?
- que é um evento?

VARIÁVEL ALEATÓRIA

Quando os resultados de uma variável são determinados pelo acaso, trata-se de uma **variável aleatória**.

“Uma **variável aleatória** é uma função com valores **numéricos**, determinados por fatores de chance.”

Stevenson, W. (Estatística aplicada à administração)

EXEMPLOS

Selecionando-se uma pessoa de um município através de sorteio, o **peso** é uma **variável aleatória**.

Sorteando-se um setor de uma empresa, o **número de funcionários** é uma **variável aleatória**.

Lança-se uma moeda várias vezes e verifica-se a face obtida (cara ou coroa):

- **Face obtida em cada jogada** - variável qualitativa - **não é uma variável aleatória**.
- **Número de caras** - **variável aleatória** associada à variável qualitativa estudada.

INTRODUÇÃO AO ESTUDO DE PROBABILIDADE

Suponha que existam n possíveis resultados, igualmente prováveis de um experimento. A probabilidade de que um evento A ocorra é igual ao número de maneiras que o evento pode ocorrer, f , dividido pelo número de possíveis resultados, n ,

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

1. Qual a probabilidade de ocorrência da face **2** quando se joga um dado?
2. Qual a probabilidade de ocorrência de **número par** quando se joga um dado?
3. Você e seu amigo tiram a sorte com cara ou coroa para quem vai pagar o lanche hoje. Você escolhe cara. Qual é a probabilidade de você ser escolhido para pagar?



MODELO PROBABILÍSTICO

Modelo de probabilidade ou probabilístico, ou ainda **distribuição de probabilidades** indica, para uma variável aleatória, **quais** são os **resultados** que podem ocorrer e **qual** é a **probabilidade** de cada resultado acontecer.

MODELO PROBABILÍSTICO

Todo fenômeno ou experimento que envolva um elemento aleatório tem seu modelo probabilístico determinado, quando estabelecemos:

Um *espaço amostral*, Ω

No caso de uma variável discreta, o espaço amostral é a enumeração de todos os resultados possíveis do experimento:

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$$

Cada resultado possível ω_i é um *ponto amostral*.

Um exemplo simples é jogar uma moeda, enquanto observa se sai cara ou coroa. Um exemplo onde a variável é real é a observação de uma temperatura.

Uma *probabilidade*, $P(\omega_i)$

Para cada *ponto amostral*, podemos considerar a probabilidade como uma função que recebe um elemento do espaço amostral e mapeia o resultado para um número real não negativo e menor ou igual a 1.

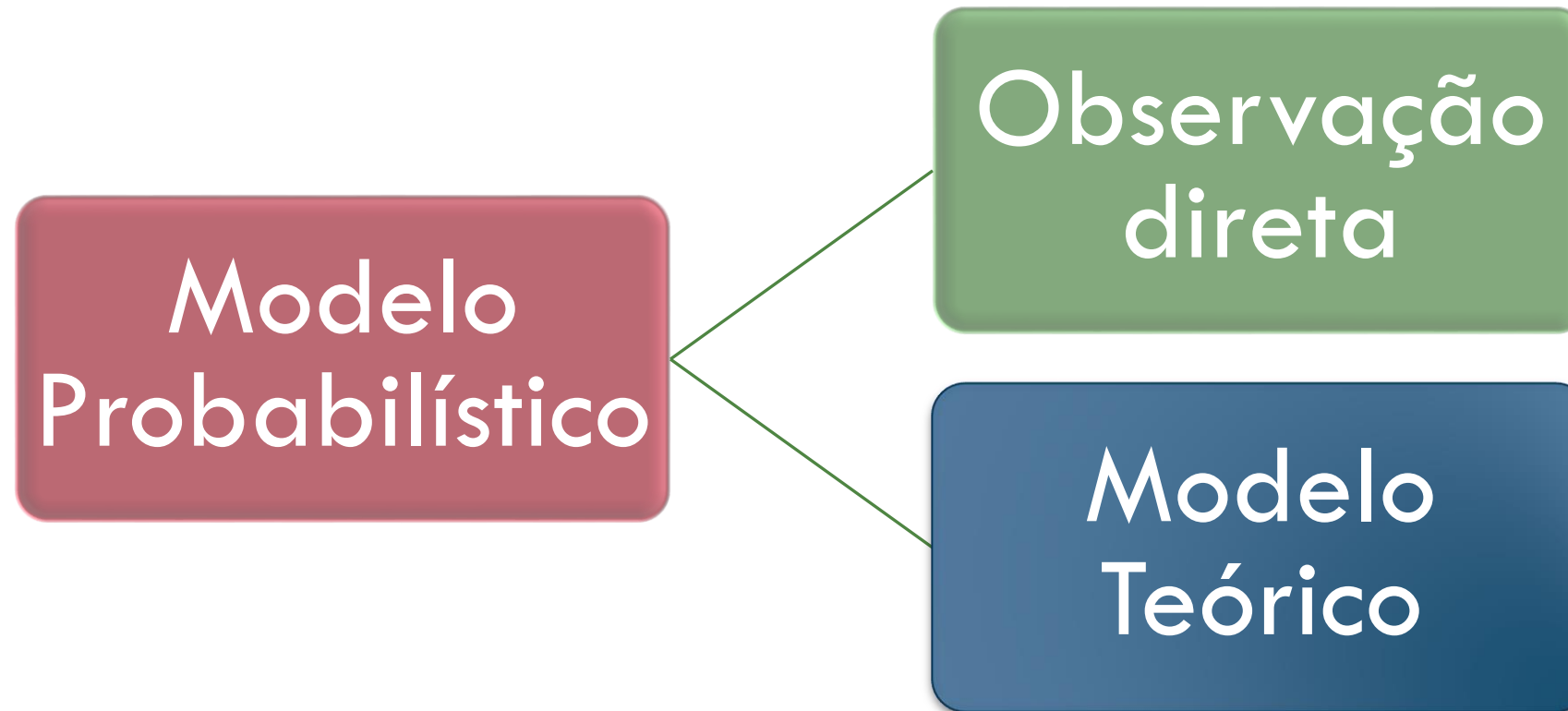
EVENTO

Em teoria das **probabilidades**, um **evento** é um conjunto de resultados (um subconjunto do espaço amostral) ao qual é associado um valor de **probabilidade**.

Por exemplo, quando lançamos dados, o conjunto de todos os números pares é um evento.

Portanto, se lançarmos os dados e sair o número 4, dizemos que o evento ocorreu.





OBSERVAÇÃO DIRETA

- Joga-se o dado n vezes e conta-se o número f_i de vezes em que ocorre cada face i , $i = 1, 2, 3, 4, 5, 6$.
- As proporções f_i/n determinam a distribuição de frequências do fenômeno.

Se jogarmos o dado um número n' vezes teremos outra distribuição de frequências, mas com um padrão que esperamos ser muito próximo ao anterior.

MODELO TEÓRICO

Mesmo sem observar diretamente o fenômeno é possível criar um modelo teórico, que reproduza bem a distribuição das frequências que se verifica quando se observa o próprio fenômeno.

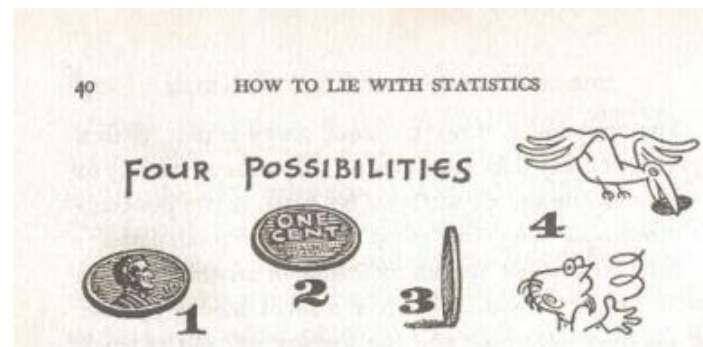
Por exemplo, no lançamento de um dado, as suposições teóricas são:

- só podem ocorrer 6 faces;
- o dado seja perfeitamente equilibrado;
- cada face deve ocorrer o mesmo número de vezes na proporção de $1/6$.

**Modelo
teórico de
frequências:**

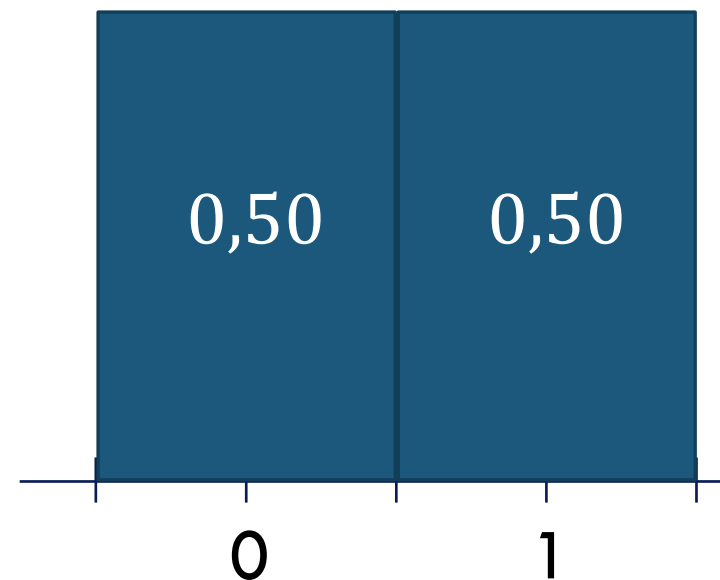
Face	1	2	3	4	5	6
Frequência teórica	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

EXERCÍCIO



Lança-se uma moeda e anota-se a face obtida. Construir a **distribuição de probabilidades** para a variável aleatória **número de caras**.

Resultados Possíveis	Variável aleatória (x)	Probabilidade $P(X = x)$
<i>Coroa</i>	0	0,5
<i>Cara</i>	1	0,5
Total:		1



EXERCÍCIO

Lança-se a moeda 2 vezes, construir a distribuição de probabilidades para a variável aleatória **número de caras**.

Para construir a distribuição de probabilidades, preciso do espaço amostral:

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$$

- $\omega_1 = (Cara, Cara)$
- $\omega_2 = (Cara, Coroa)$
- $\omega_3 = (Coroa, Cara)$
- $\omega_4 = (Coroa, Coroa)$

probabilidade de
ocorrência de cada
ponto amostral ω_i

É razoável supor que, em cada lance, existe a probabilidade de $\frac{1}{2}$ de sair **Cara** e $\frac{1}{2}$ de sair **Coroa**, se a moeda é perfeitamente simétrica e homogênea.

EXERCÍCIO

Resultados Possíveis	Variáveis aleatórias (x)	Probabilidade $P(X = x)$
<i>Coroa – Coroa</i>	0	$0,5 \times 0,5 = 0,25$
<i>Cara – Coroa</i>	1	$0,5 \times 0,5 = 0,25$
<i>Coroa – Cara</i>	1	$0,5 \times 0,5 = 0,25$
<i>Cara – Cara</i>	2	$0,5 \times 0,5 = 0,25$
	Total:	1

Ao repetir experimentos, uma suposição comum é que o resultado de um experimento não tem qualquer influência no resultado dos outros. Em outras palavras, os experimentos são **independentes**.

A probabilidade de que dois eventos **independentes** ocorram é igual à multiplicação das probabilidades individuais.

DIAGRAMA DE ÁRVORE

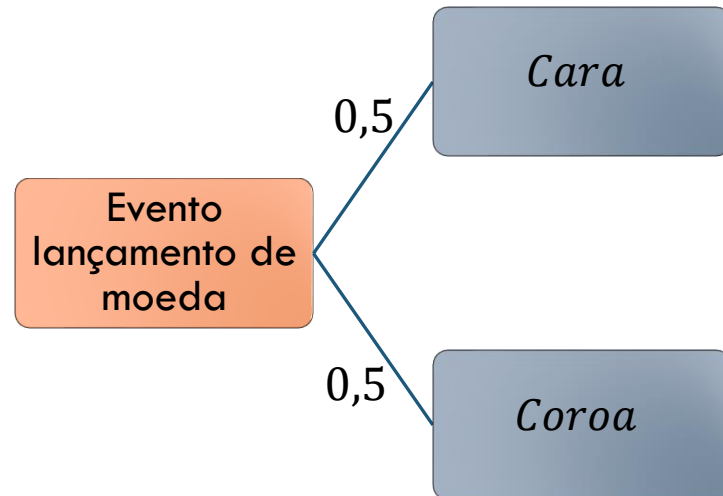


DIAGRAMA DE ÁRVORE

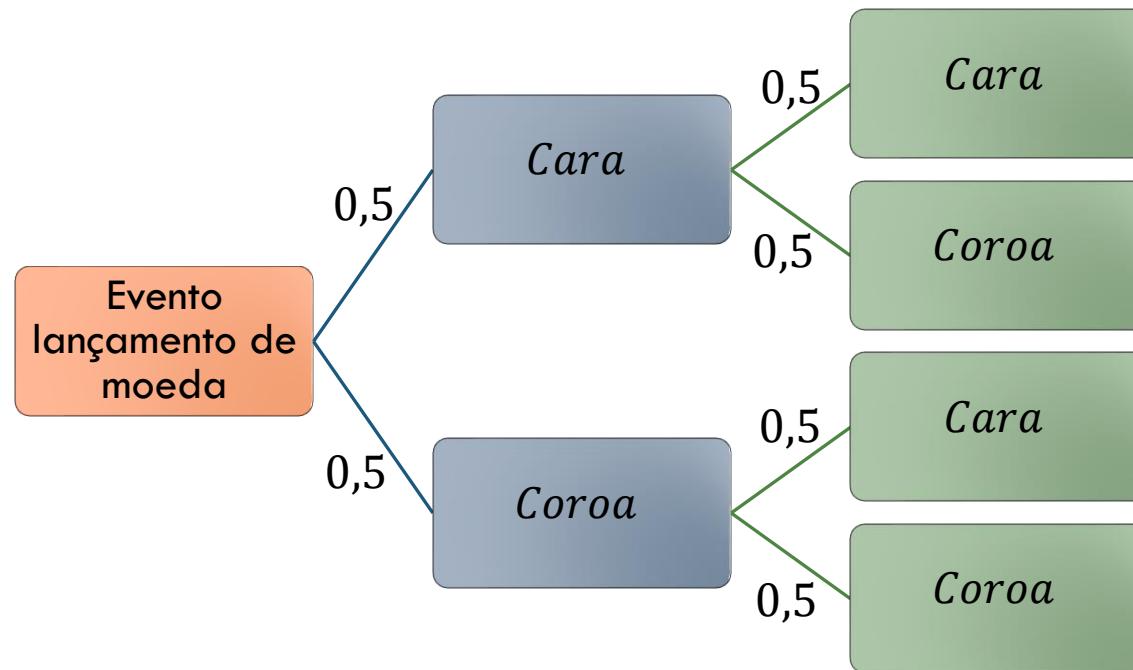
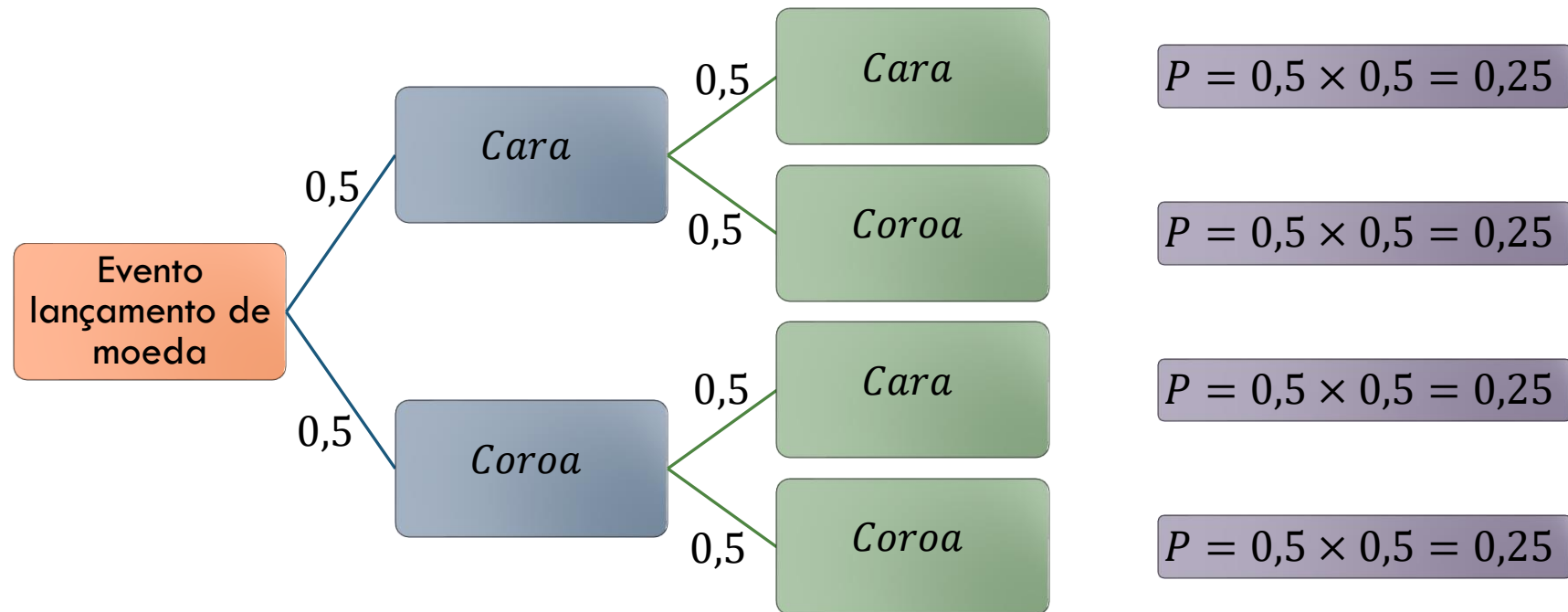


DIAGRAMA DE ÁRVORE

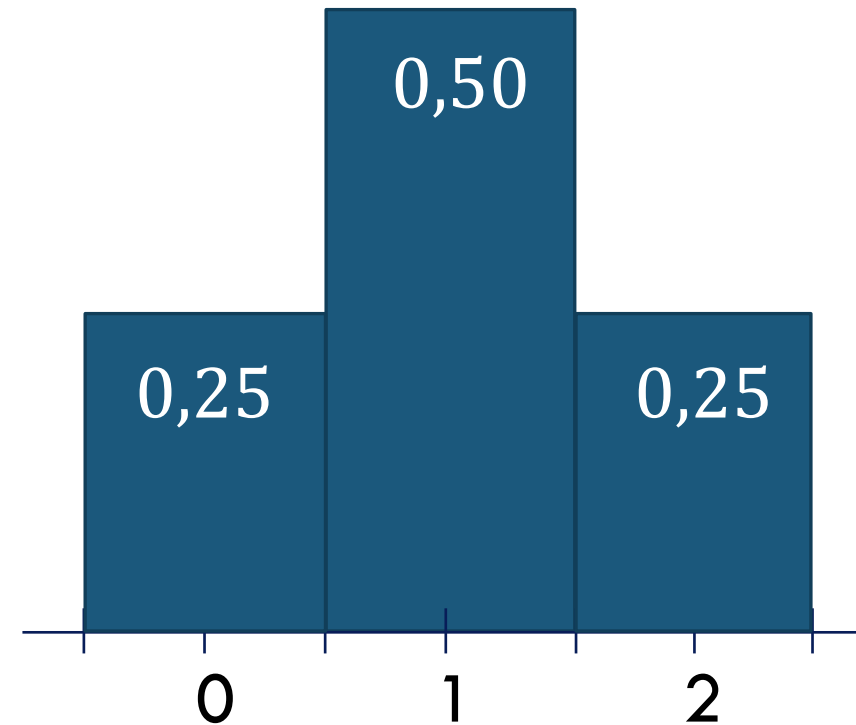


Se A é o evento “número de caras”, então:

x	$P(X = x)$
0	0,25
1	0,50
2	0,25
Total:	1

A probabilidade de que um entre dois eventos **mutuamente excludentes** ocorra (OU) é igual à soma das probabilidades individuais.

Dois eventos são **mutuamente excludentes, ou exclusivos**, se a ocorrência de um impedir a ocorrência do outro.



GENERALIZANDO A PROBABILIDADE DE OCORRÊNCIA DE UM EVENTO

Se A é um evento, então:

$$P(A) = \sum P(\omega_j), \text{ para todos os } \omega_j \in A.$$

Para um mesmo experimento podemos ter vários espaços amostrais, dependendo do nosso interesse.

Exemplo: Uma fábrica produz um determinado artigo. Da linha de produção são retirados 3 artigos, que são classificados como “bom (B)” ou “defeituoso (D)”. Um espaço amostral desse experimento é:

$$\Omega = \{ BBB, BBD, BDB, DBB, DDB, DBD, BDD, DDD \}$$

Se A é o evento “2 artigos defeituosos”, então:

$$A = \{ DDB, DBD, BDD \}$$

Se A é o evento “pelo menos 1 artigo bom”, então:

$$A = \{ BBB, BBD, BDB, DBB, DDB, DBD, BDD \}$$

EXERCÍCIO

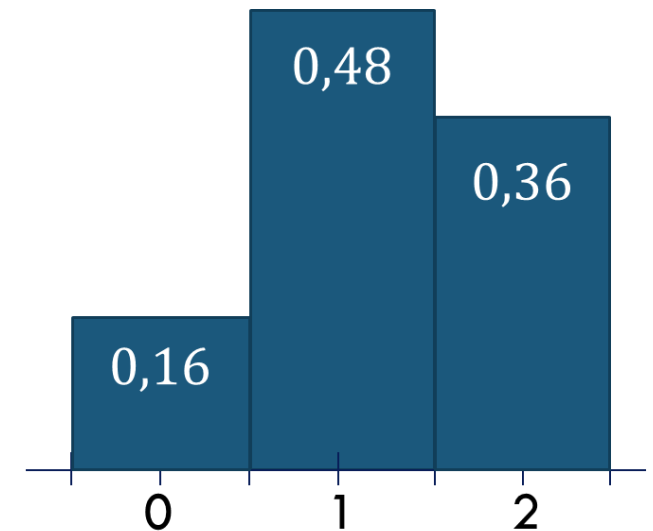
Um grande lote de peças possui 60% dos itens com algum tipo de defeito. Construir a distribuição de probabilidades para a variável aleatória **número de itens com defeito dentre 2 sorteados aleatoriamente.**

EXERCÍCIO

Um grande lote de peças possui 60% dos itens com algum tipo de defeito. Construir a distribuição de probabilidades para a variável aleatória **número de itens com defeito dentre 2 sorteados aleatoriamente.**

Resultados Possíveis	Resultados numéricos (x)	Probabilidade $P(X = x)$
<i>Bom – Bom</i>	0	$0,4 \times 0,4 = 0,16$
<i>Bom – Def</i>	1	$0,4 \times 0,6 = 0,24$
<i>Def – Bom</i>	1	$0,6 \times 0,4 = 0,24$
<i>Def – Def</i>	2	$0,6 \times 0,6 = 0,36$
	Total:	1

x	$P(X=x)$
0	0,16
1	0,48
2	0,36
Total:	1

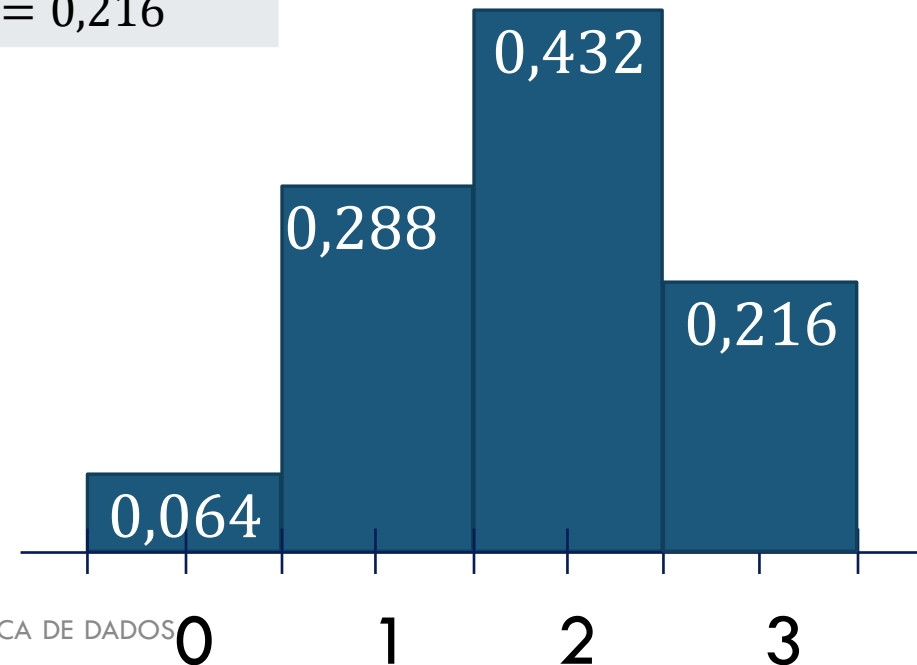


EXERCÍCIO PARA VOCÊ TREINAR EM CASA

Um grande lote de peças possui 60% dos itens com algum tipo de defeito.
Construir a distribuição de probabilidades para a variável aleatória **número de itens com defeito dentre 3 peças sorteadas aleatoriamente**.

Respostas possíveis	Resposta numérica (x)	Probabilidade
<i>B B B</i>	0	$0,4 \times 0,4 \times 0,4 = 0,064$
<i>B B D</i>	1	$0,4 \times 0,4 \times 0,6 = 0,096$
<i>B D B</i>	1	$0,4 \times 0,6 \times 0,4 = 0,096$
<i>D B B</i>	1	$0,6 \times 0,4 \times 0,4 = 0,096$
<i>B D D</i>	2	$0,4 \times 0,6 \times 0,6 = 0,144$
<i>D B D</i>	2	$0,6 \times 0,4 \times 0,6 = 0,144$
<i>D D B</i>	2	$0,6 \times 0,6 \times 0,4 = 0,144$
<i>D D D</i>	3	$0,6 \times 0,6 \times 0,6 = 0,216$

x	P(X=x)
0	0,064
1	$0,096 \times 3 = 0,288$
2	$0,144 \times 3 = 0,432$
3	0,216

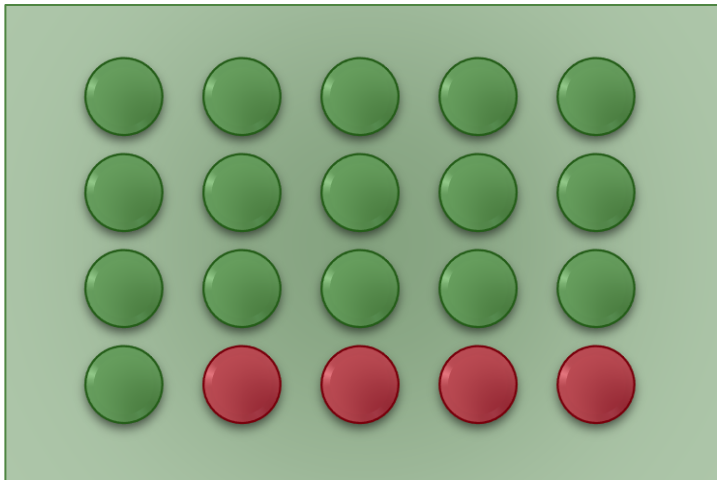


EXEMPLO:

Um lote com 20 peças contém 4 defeituosas. Se forem retiradas duas peças do lote, qual é a probabilidade de serem retiradas:

- a) duas peças boas?
- b) duas peças defeituosas?

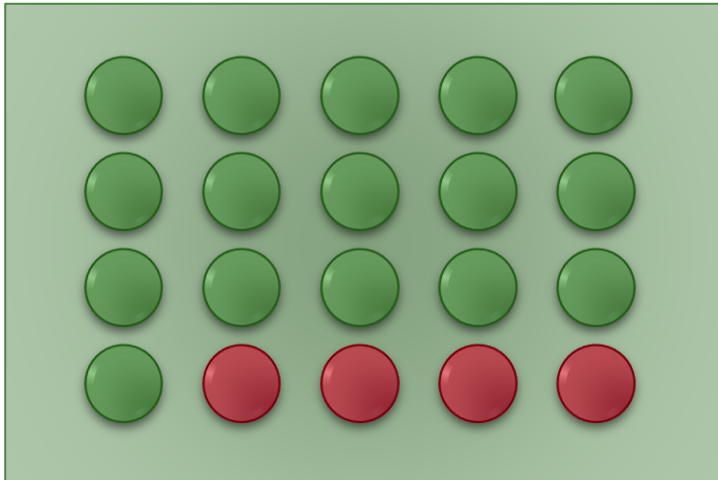
SITUAÇÃO INICIAL,



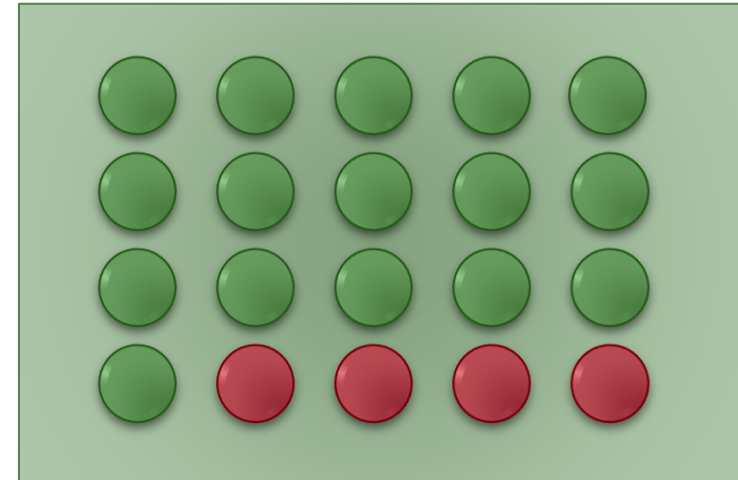
● B - Peça Boa

● D - Peça Defeituosa

PRIMEIRA PEÇA



$$P(B) = \frac{16}{20}$$

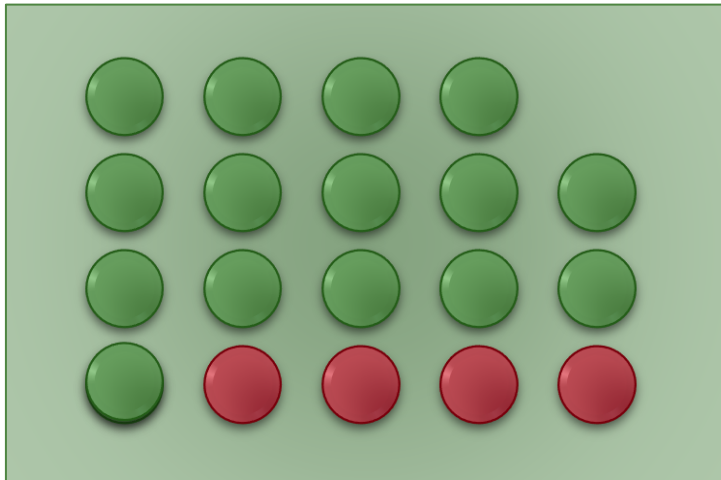


$$P(D) = \frac{4}{20}$$

RETIRADA DA SEGUNDA PEÇA,

Se a primeira peça for

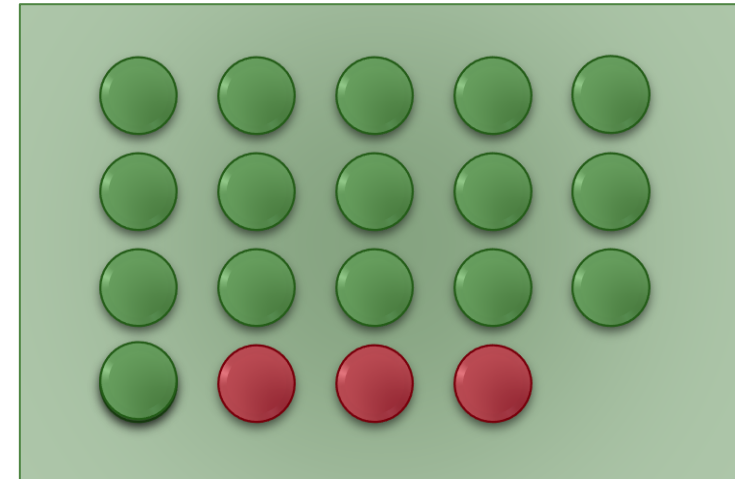
Boa



$$P(2^{\text{a}} \text{ peça ser boa}) = 15/19$$

$$P(2^{\text{a}} \text{ peça ser defeituosa}) = 4/19$$

Defeituosa



$$P(2^{\text{a}} \text{ peça ser boa}) = 16/19$$

$$P(2^{\text{a}} \text{ peça ser defeituosa}) = 3/19$$

ENTÃO,

A. Probabilidade das duas peças serem boas,

B. Probabilidade das duas peças serem defeituosas,

E a probabilidade de ter uma peça defeituosa e uma boa?

E a probabilidade da primeira ser defeituosa e a segunda não nesta ordem?

PORTANTO,

A. Probabilidade das duas peças serem boas,

$$P(BB) = \frac{16}{20} \times \frac{15}{19} = 63,16\%$$

B. Probabilidade das duas peças serem defeituosas,

$$P(DD) = \frac{4}{20} \times \frac{3}{19} = 3,16\%$$

PORTANTO,

E a probabilidade de ter uma peça defeituosa e uma boa?

$$P(BD) + P(DB) = \left(\frac{16}{20} \times \frac{4}{19}\right) + \left(\frac{4}{20} \times \frac{16}{19}\right)$$

E a probabilidade da primeira ser defeituosa e a segunda não nesta ordem?

$$P(DB) = \left(\frac{4}{20} \times \frac{16}{19}\right)$$

MAIS UM...

A Petrobrás perfura um poço quando há probabilidade de, no mínimo, 40 % de encontrar petróleo. Foram encontrados dois poços, com probabilidade de haver petróleo de 40 % e 50 %.

Os dois poços são perfurados.

Qual é a probabilidade de que pelo menos um poço produza petróleo?

Sim: produz
Não: não produz

Resultados Possíveis	Resultados numéricos (x)	Probabilidade $P(X = x)$
<i>Sim Não</i>	1	$0,4 \times 0,5 = 0,2$
<i>Sim Sim</i>	2	$0,4 \times 0,5 = 0,2$
<i>Não Sim</i>	1	$0,6 \times 0,5 = 0,3$
<i>Não Não</i>	0	$0,6 \times 0,5 = 0,3$

0,7

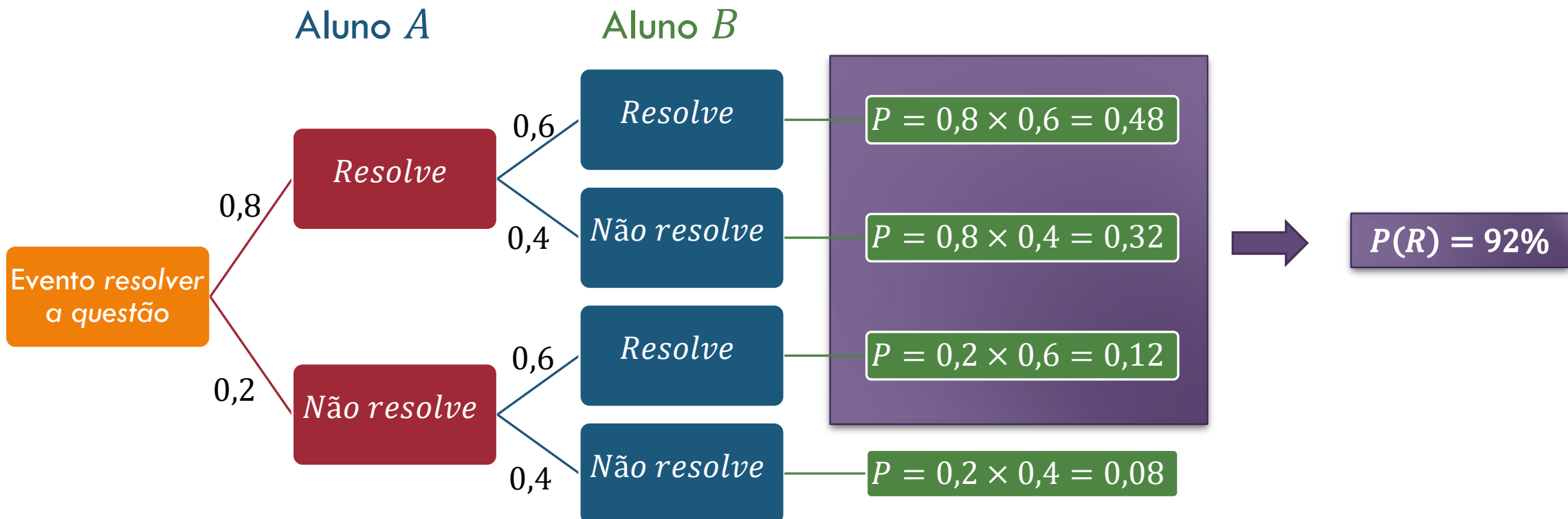
A probabilidade de que pelo menos um poço produza petróleo é de 70%.

RESOLVA...

A probabilidade de um aluno A resolver uma questão de prova é 0,80, enquanto que a do aluno B é 0,60. Qual a probabilidade de que a questão seja resolvida se os dois alunos tentarem resolvê-la independentemente.

RESOLVA...

A probabilidade de um aluno *A* resolver uma questão de prova é 0,80, enquanto que a do aluno *B* é 0,60. Qual a probabilidade de que a questão seja resolvida se os dois alunos tentarem resolvê-la independentemente.





EASY

**Fácil... Está na hora de
complicar um pouco...**

PROPRIEDADES DA PROBABILIDADE

Dado um modelo probabilístico, pode-se verificar as seguintes propriedades ,

Como toda frequência relativa é um número entre 0 e 1, tem-se que:

$$0 \leq P(A) \leq 1$$

para qualquer evento A .

CONT...

Se for considerado todo o espaço amostral (S) como evento, tem-se o denominado *evento certo*, portanto

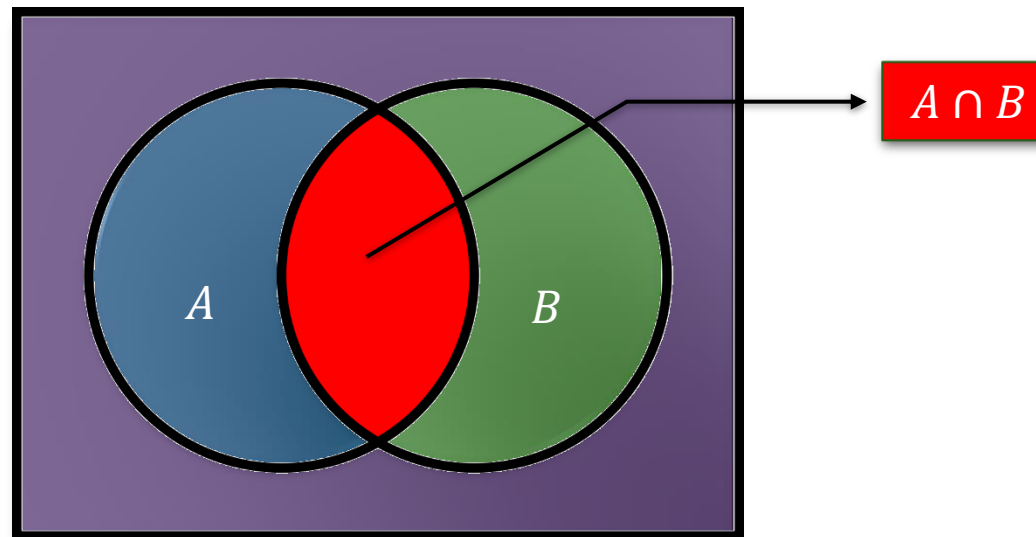
$$P(S) = 1$$

Se for considerado um conjunto vazio como evento (\emptyset), tem-se o denominado *evento impossível*, ou seja

$$P(\emptyset) = 0$$

EVENTO INTERSECÇÃO

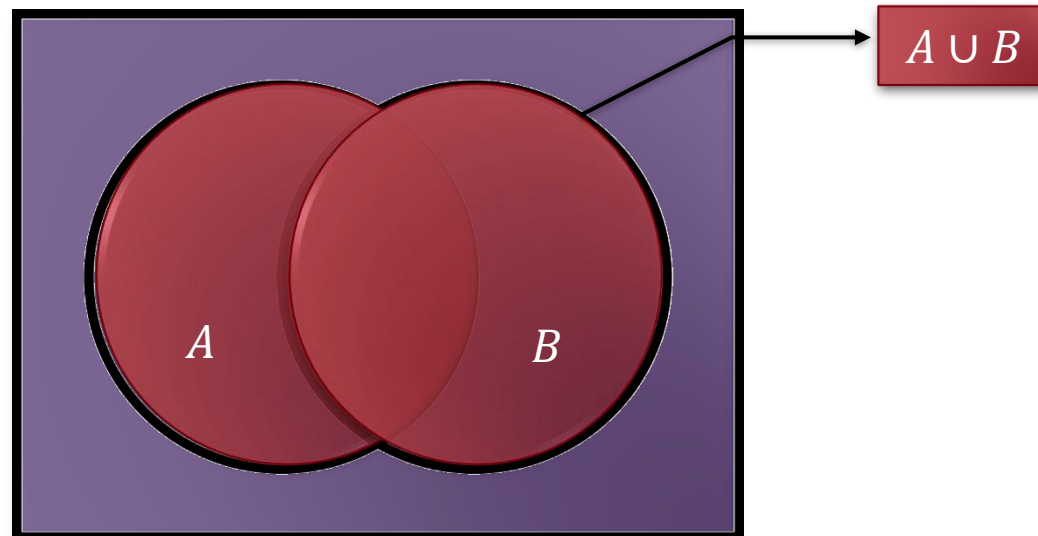
O evento intersecção ($A \cap B$) significa que A e B ocorrem simultaneamente. Esta probabilidade é calculada com o emprego do *Teorema da Probabilidade Condicionada*, a ser visto na seqüência desta aula.



EVENTO UNIÃO

O evento união ($A \cup B$) significa que pelo menos um dos eventos ocorre, sendo calculado pela relação:

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



POR EXEMPLO...

Suponha que você joga um dado, se A é o evento números pares e B é o evento números maiores ou iguais a 3, qual a probabilidade de ocorrer A e B ao mesmo tempo?

$$A = \{2 \quad 4 \quad 6\} \quad P(A) = \frac{3}{6} = \frac{1}{2}$$

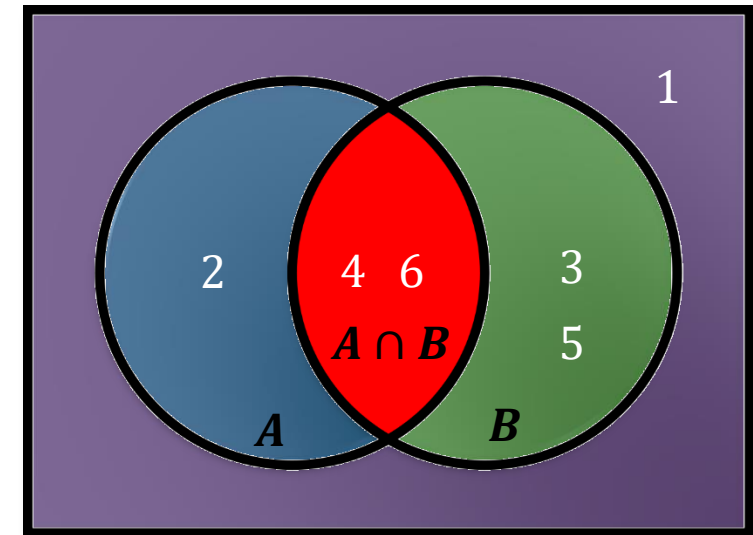
$$B = \{3 \quad 4 \quad 5 \quad 6\} \quad P(B) = \frac{4}{6} = \frac{2}{3}$$

$$A \cup B = \{2 \quad 3 \quad 4 \quad 5 \quad 6\}$$

$$P(A, B) = P(A \cup B) = \frac{5}{6}$$

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

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{3}{6} + \frac{4}{6} - \frac{2}{6} = \frac{5}{6}$$



EXEMPLO

A Petrobrás perfura um poço quando há probabilidade de, no mínimo, 40 % de encontrar petróleo. Foram encontrados dois poços, com probabilidade de haver petróleo de 40 % e 50 %.

Os dois poços são perfurados. Qual é a probabilidade de que pelo menos um poço produza petróleo?

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

$$P(A) = 0,4$$

$$P(B) = 0,5$$

$$P(A \text{ e } B) = P(A \cap B) = 0,4 \times 0,5 = 0,2$$

$$\mathbf{P(A \text{ ou } B) = P(A \cup B) = 0,4 + 0,5 - 0,2 = 0,7}$$

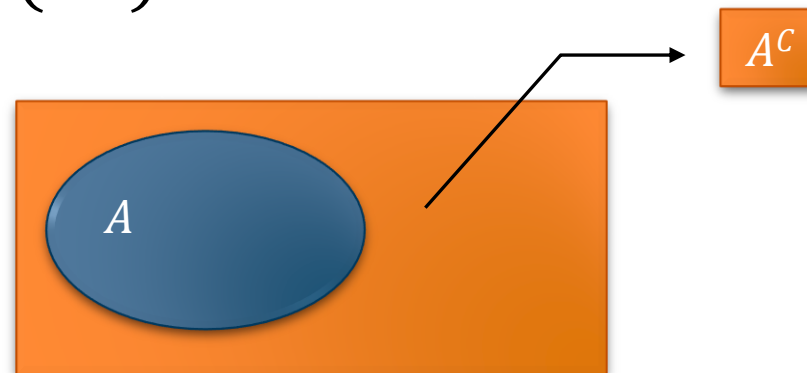
Resolva o problema do aluno respondendo às questões da prova usando a equação: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

EVENTO COMPLEMENTAR

O evento complementar de um evento A , denominado de A^c , tem sua probabilidade calculada pela relação:

$$(A \cup A^c) = S \text{ ou } (A \cap A^c) = \emptyset$$

$$P(A) + P(A^c) = 1$$



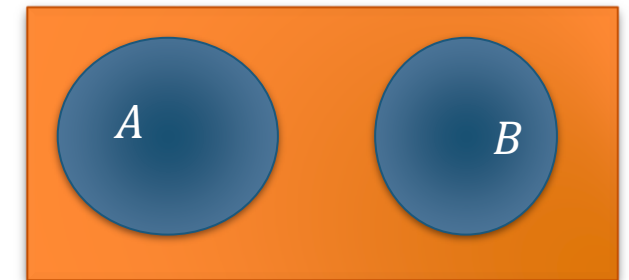
EVENTOS MUTUAMENTE EXCLUSIVOS

Dois eventos A e B são **mutuamente exclusivos** quando a intersecção desses dois eventos for o conjunto vazio:

$$P(A \cap B) = 0$$

Portanto

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



RESOLVA O PROBLEMA DE UNIÃO (ou)

Quando um par de dados é jogado, qual a probabilidade de se obter um resultado que a soma total dos valores seja menor que 4 **OU** que contenha o número 4?

Espaço amostral:

1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

Fonte: Krishnamoorthi, K.S.
“Reliability Methods for Engineers”, 1ª
edição, ASQC Quality Press, 1992.

Evento A:

$$P(\text{total} < 4) = 3/36$$

1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

Evento B:

$$P(\text{sair o número } 4) = 11/36$$

$$\begin{aligned} \therefore P(A \cup B) &= \frac{3}{36} + \frac{11}{36} \\ &= 14/36 = \mathbf{7/18} \end{aligned}$$

AGORA ESSE...

Quando um par de dados é jogado, qual a probabilidade de se obter um resultado que contenha os números 5 **OU** 6?

Espaço amostral:

1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

Fonte: Krishnamoorthi, K.S.
“Reliability Methods for Engineers”, 1ª
edição, ASQC Quality Press, 1992.

Evento A:

$$P(\text{sair o número } 5) = 11/36$$

1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

Evento B:

$$P(\text{sair o número } 6) = 11/36$$

$$P(\text{sair } 5 \text{ e } 6) = 2/36$$

$$\begin{aligned}
 \therefore P(\text{sair } 5 \text{ ou } 6) &= \\
 &= 11/36 + 11/36 - 2/36 \\
 &= 20/36 = 5/9
 \end{aligned}$$

E SE A PERGUNTA FOSSE...

Jogando um par de dados, qual a probabilidade que a soma dos valores obtidos seja menor ou igual a 6 (evento B) **DADO QUE** um dos números obtidos em uma das faces é 3 (evento A).

Já sabemos que o evento A ocorreu, portanto, nosso espaço amostral mudou,

$$P(B|A) = \frac{5}{11}$$

$B|A$ significa: ocorrer B dado que A ocorreu.

			1,3		
			2,3		
3,1	3,2	3,3	3,4	3,5	3,6
			4,3		
			5,3		
			6,3		

PROBABILIDADE CONDICIONADA

Em alguns casos, o fato de ser sabido, à priori, que um dado evento ocorreu, faz com que se modifique a probabilidade de ocorrência de um outro evento.

Foi o que aconteceu no exemplo: o fato de saber que o valor obtido em um dos dados foi 3 diminuiu o tamanho do espaço amostral de 36 para 11!

PRESTE ATENÇÃO

Os dados abaixo referem-se a 200 entrevistados em uma enquete feita pela prefeitura:

	HOMENS	MULHERES	TOTAL
FAVORÁVEIS	60	50	110
DESFAVORÁVEIS	80	10	90
	140	60	200

Qual a probabilidade de uma pessoa, aleatoriamente escolhida,

- A. Ser favorável? $110/200$
- B. Ser favorável, dado ser homem ? $60/140$
- C. Ser homem? $140/200$
- D. Ser homem, dado que é desfavorável? $80/90$
- E. Ser favorável, dado que é mulher? $50/60$

TEOREMA DA PROBABILIDADE CONDICIONADA

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A, B)}{P(B)}, P(B) \neq 0$$

↓
Probabilidade condicional

↑
Probabilidade conjunta (joint)

↓
Probabilidade marginal

DEFINIÇÃO DE PROBABILIDADE CONDICIONADA

De novo: em alguns casos, *o fato de ser sabido, à priori, que um dado evento ocorreu, faz com que se modifique a probabilidade de ocorrência de um outro evento.*

Denomina-se de $P(A|B)$ a probabilidade de ocorrência do evento A , sabendo-se que B ocorreu, ou probabilidade de A condicionada a B .

Tem-se que:

$$P(A|B) = \frac{P(A,B)}{P(B)}, P(B) \neq 0$$

onde $P(A)$ e $P(B)$ são a probabilidade incondicional do evento A e B ocorrerem, respectivamente.

EXEMPLO

A página seguinte apresenta a distribuição em níveis salariais (R_1, R_2, R_3, R_4) da carreira de um engenheiro em determinada empresa, de acordo com sua idade. Suponha que um engenheiro seja selecionado ao acaso.

1. Encontre a probabilidade de que o engenheiro selecionado esteja na casa dos 50 anos.
2. Encontre a possibilidade de que o engenheiro esteja na casa dos 50 anos, sabendo-se que foi selecionado um R_3 .

1. Encontre a probabilidade de que o engenheiro selecionado esteja na casa dos 50 anos.
2. Encontre a possibilidade de que o engenheiro esteja na casa dos 50 anos, sabendo-se que foi selecionado um R_3 .

		Posição na carreira				
		R_1	R_2	R_3	R_4	Total
Idade	< 30	2	3	57	6	68
	30 – 39	52	170	163	17	402
	40 – 49	156	125	61	6	348
	50 – 59	145	68	36	4	253
	≥60	75	15	3	0	93
	Total	430	381	320	33	1164

RESPOSTA

$$1. P(50) = \frac{253}{1164} = 0,217$$

$$2. P(50|R_3) = \frac{P(50,R_3)}{P(R_3)} = \frac{36/1164}{320/1164} = 0,113$$

RESPOSTA

Você encontrou os valores $P(50) = 0,217$ e $P(50|R_3) = 0,113$. O que significam esses valores? Interprete os resultados.

$P(50) = 0,217$ indica que **21,7% dos engenheiros da empresa** estão na casa dos 50.

$P(50|R_3) = 0,113$ indica que **11,3% dos engenheiros nível R_3** estão na casa dos 50.

PARA VOCÊ FAZER EM CASA

$A = \{Compra = 1, Não compra = 0\}, B = \{Brasil, Peru, Bolívia\}$

Queremos saber,

$$P(Opção = Y | País = X) = \frac{P(Y, X)}{P(X)} = ?$$

	<i>Peru</i>	<i>Brasil</i>	<i>Bolívia</i>	Total
<i>Compra = 1</i>	20	50	10	80
<i>Não compra = 0</i>	300	500	200	1000
Total	320	550	210	1080

	<i>Peru</i>	<i>Brasil</i>	<i>Bolívia</i>	Total
<i>Compra = 1</i>	20	50	10	80
<i>Não compra = 0</i>	300	500	200	1000
Total	320	550	210	1080

Probabilidade marginal $P(X)$

$$P(X = Peru) = 320/(210 + 550 + 320) = 0,30$$

$$P(X = Brasil) = 550/(210 + 550 + 320) = 0,51$$

$$P(X = Bolívia) = 210/(210 + 550 + 320) = 0,19$$

Probabilidade conjunta

$$P(1, X = Peru) = 20/(1080) = 0,0185$$

$$P(0, X = Peru) = 300/(1080) = 0,2778$$

$$P(1, X = Brasil) = 50/(1080) = 0,0463$$

$$P(0, X = Brasil) = 500/(1080) = 0,4629$$

$$P(1, X = Bolívia) = 10/(1080) = 0,0093$$

$$P(0, X = Bolívia) = 200/(1080) = 0,1852$$

PROBABILIDADE CONDICIONADA

$$P(1|Peru) = \frac{P(1, Peru)}{P(Peru)} = \frac{0,019}{0,30} = \mathbf{0,06}$$

$$P(0|Peru) = \frac{P(0, Peru)}{P(Peru)} = \frac{0,28}{0,30} = \mathbf{0,93}$$

$$P(1|Brasil) = \frac{P(1, Brasil)}{P(Brasil)} = \frac{0,046}{0,51} = \mathbf{0,09}$$

$$P(0|Brasil) = \frac{P(0, Brasil)}{P(Brasil)} = \frac{0,46}{0,51} = \mathbf{0,91}$$

$$P(1|Bolivia) = \frac{P(1, Bolivia)}{P(Bolivia)} = \frac{0,009}{0,19} = \mathbf{0,05}$$

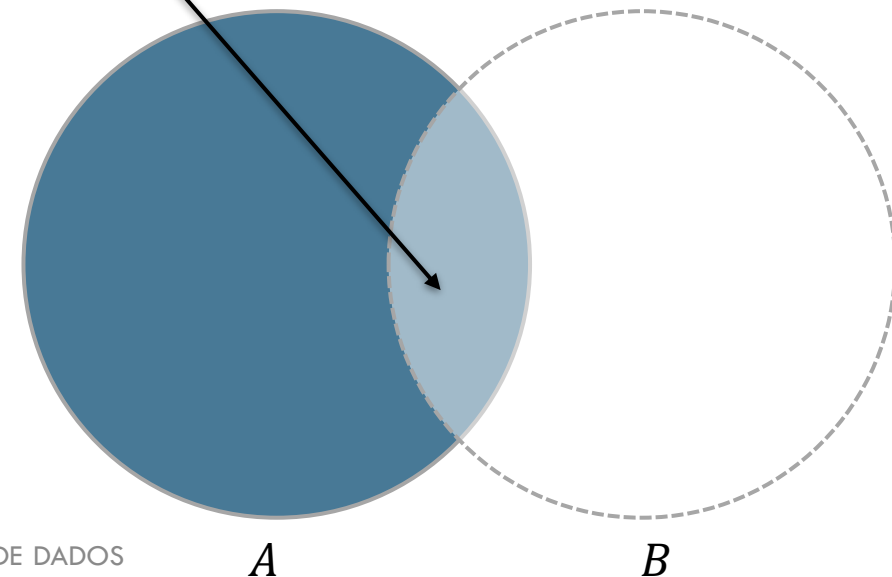
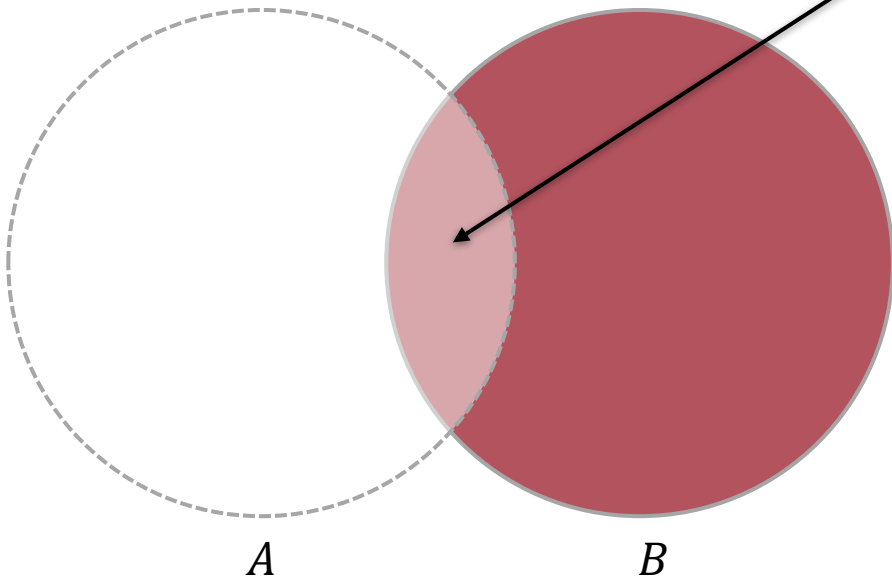
$$P(0|Bolivia) = \frac{P(0, Bolivia)}{P(Bolivia)} = \frac{0,185}{0,19} = \mathbf{0,97}$$

UMA DEFINIÇÃO EXTRA...

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

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

$$P(A, B) = P(A|B)P(B) = P(B|A)P(A)$$



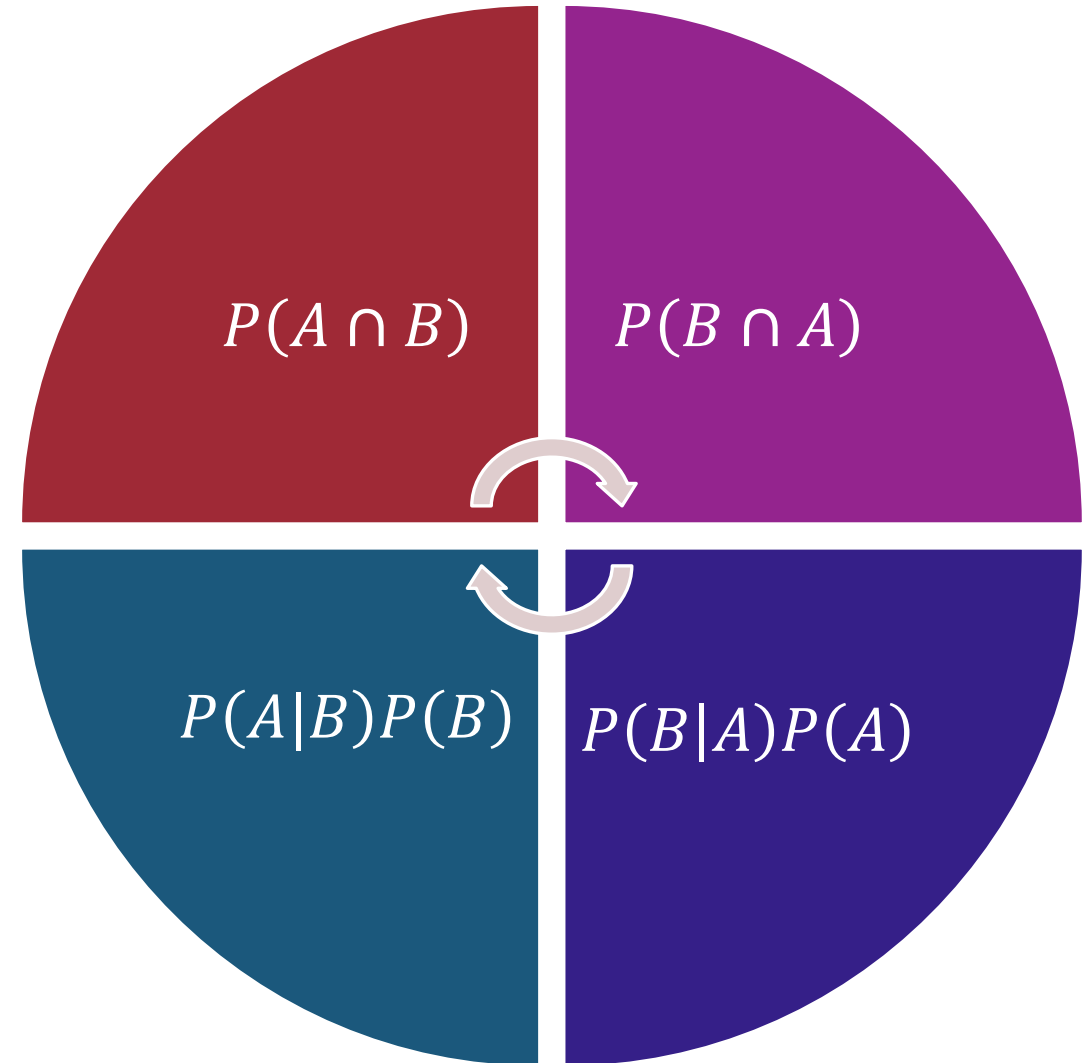
INTERSECÇÃO

Pode-se formular o Teorema do Produto da seguinte forma:

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

$$P(A, B) = P(B, A)$$

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



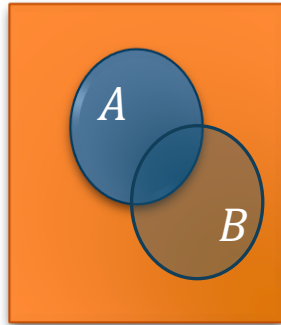
EVENTOS INDEPENDENTES

No caso dos eventos serem independentes, isto é, a ocorrência do evento A não influencia a probabilidade de ocorrência do evento B , a probabilidade de ocorrência de $(A \cap B)$ é calculada pela relação:

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

ou seja

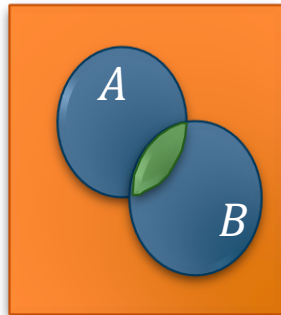
$$P(B|A) = P(B) \text{ ou } P(A|B) = P(A)$$



$A \cup B$

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

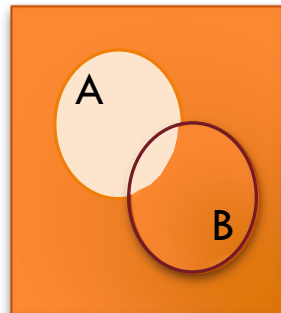
$$P(A \cup B) = P(A) + P(B) \text{ eventos mutuamente exclusivos}$$



$A \cap B$

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

$$P(A \cap B) = P(A)P(B) \text{ eventos independentes}$$



A^c

$$P(A^c) = 1 - P(A)$$

TEOREMA DE BAYES

Uma das equações famosas do mundo da estatística e probabilidade.

$$P(A|B) = \frac{P(A, B)}{P(B)} = P(A) \frac{P(B|A)}{P(B)}, P(B) \neq 0$$

Regra de Bayes é apenas uma formalização da lógica que seguimos até agora...

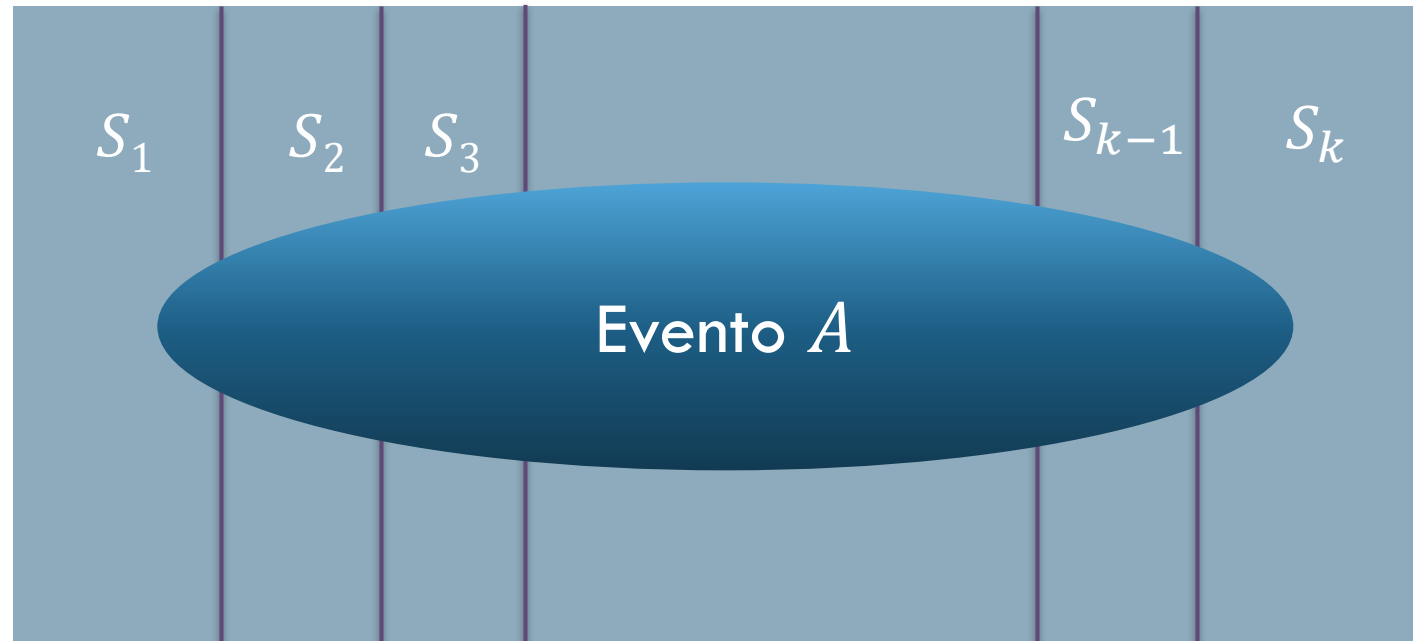


Thomas BAYES
(1702 - 1761)

TEOREMA DA PROBABILIDADE TOTAL

Vamos dividir um espaço amostral em subespaços, tal que:

$$S_i \cap S_j = \emptyset, i \neq j$$



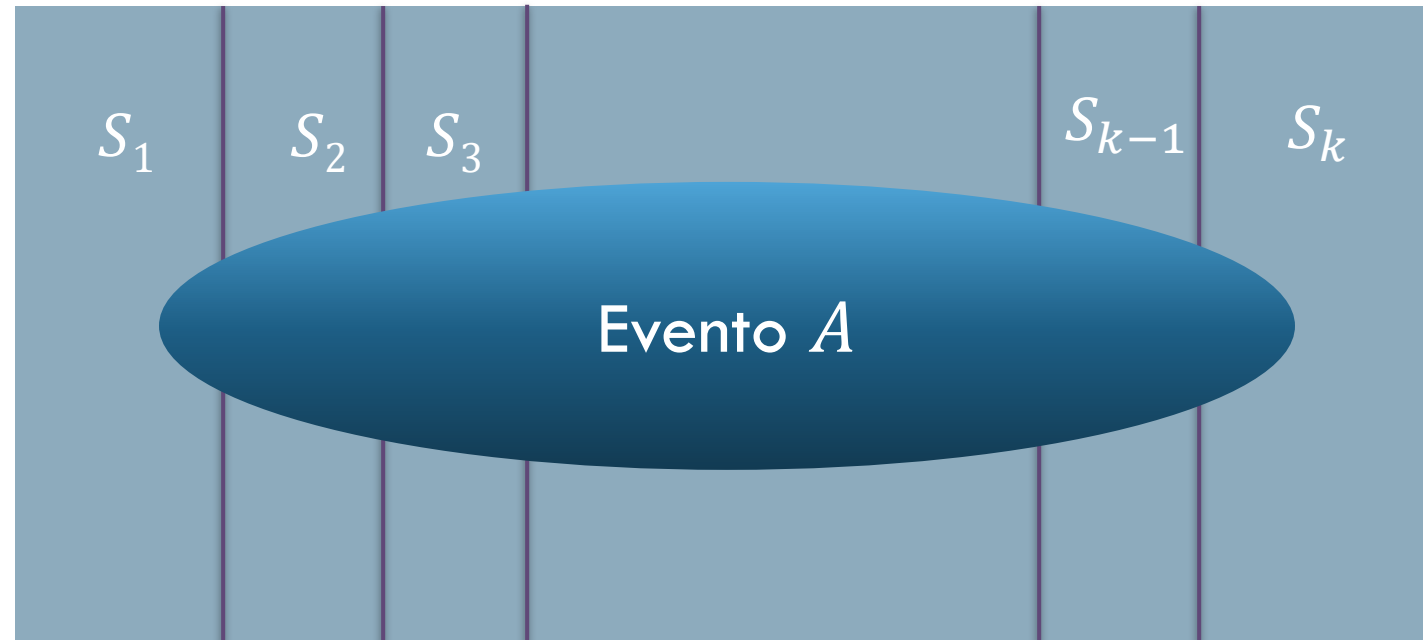
Eventos mutuamente **exclusivos** e **exaustivos**.

CONT...

A probabilidade de ocorrência do evento A é definida pela relação:

$$P(A) = P(A/S_1)P(S_1) + P(A/S_2)P(S_2) + \dots + P(A/S_k)P(S_k)$$

$$P(A) = \sum_{i=1}^k P(A/S_i)P(S_i)$$



EXEMPLO

Na escola de engenharia de uma Universidade, tem-se a seguinte distribuição de alunos por especialidade: **12%** em eng. **Produção**, **19%** em eng. **Manufatura**, **18%** em eng. **Civil**, **26%** em eng. **Elétrica** e **25%** em eng. **Mecânica**. Dentro de cada especialidade, tem-se a seguinte porcentagem de alunos do sexo feminino: **45%** em eng. **Produção**, **4%** em eng. **Manufatura**, **8%** em eng. **Civil**, **5%** em eng. **Elétrica**, **10%** em eng. **Mecânica**.

Caso um aluno seja escolhido ao acaso nesta escola, qual a probabilidade deste ser do sexo feminino?

TENTEM FAZER!!!!

$$S_1 = Pr$$

$$S_2 = Mn$$

$$S_3 = Cv$$

$$S_4 = El$$

$$S_5 = Mc$$

Evento $A = \text{Sexo Feminino (F)}$

A Probabilidade de ser do sexo feminino é dada pela relação:

$$P(F) = P(F|Pr)P(Pr) + P(F|Mn)P(Mn) + P(F|Cv)P(Cv) + P(F|El)P(El) + P(F|Mc)P(Mc)$$

$$P(F) = 0,45 \times 0,12 + 0,04 \times 0,19 + 0,08 \times 0,18 + 0,05 \times 0,26 + 0,10 \times 0,25$$

$$P(F) = 0,114$$

Tarefa: E qual a probabilidade de um estudante escolhido ao acaso ser da civil, dado que é do sexo feminino????

Tarefa: E qual a probabilidade de um estudante escolhido ao acaso ser da civil, dado que é do sexo feminino????

Sabendo a probabilidade de ser do sexo feminino : $P(F) = 0,114$

E aplicando o Teorema da Bayes:

$$P(F) = P(F|Pr)P(Pr) + P(F|Mn)P(Mn) + P(F|Cv)P(Cv) + P(F|El)P(El) + P(F|Mc)P(Mc)$$

$$P(F) = 0,45 \times 0,12 + 0,04 \times 0,19 + 0,08 \times 0,18 + 0,05 \times 0,26 + 0,10 \times 0,25$$

$$P(Cv|F) = P(Cv) \frac{P(F|Cv)}{P(F)} = 0,18 \times \frac{0,08}{0,114}$$

TEOREMA DE BAYES GENERALIZADO

Vamos unir o Teorema da Probabilidade Total em conjunto com o da Probabilidade Condicionada.

Este tem por objetivo definir a probabilidade de ocorrência de um dos eventos **mutuamente exclusivos e independentes** (B_i), dada a ocorrência do evento A .

TEOREMA DE BAYES

$$P(B_j | A) = P(B_j) \frac{P(A/B_j)}{\sum_{i=1}^k P(A/B_i)P(B_i)}$$

sendo $i = 1, \dots, k$ o número de partições independentes do espaço amostral.

Veja que o teorema nada mais é do que a substituição de todas as fórmulas que aprendemos na fórmula da probabilidade condicionada!

APLICAÇÃO

Um aluno tem um despertador que toca na hora pretendida com probabilidade 0,7. Se tocar, a probabilidade do aluno acordar a tempo de ir às aulas é de 0,8, se não tocar, a probabilidade do aluno acordar a tempo de ir às aulas é 0,3.

Qual a probabilidade do aluno chegar no horário à aula?

E ainda, dado que o aluno chegou no horário, qual a probabilidade do despertador ter tocado?

RESPOSTA

$P(T) = 0,70 \rightarrow$ probabilidade do despertador tocar

$P(nT) = 0,30 \rightarrow$ probabilidade do despertador **não** tocar

$P(A|T) = 0,80 \rightarrow$ probabilidade do aluno acordar a tempo de ir às aulas dado que o despertador tocou

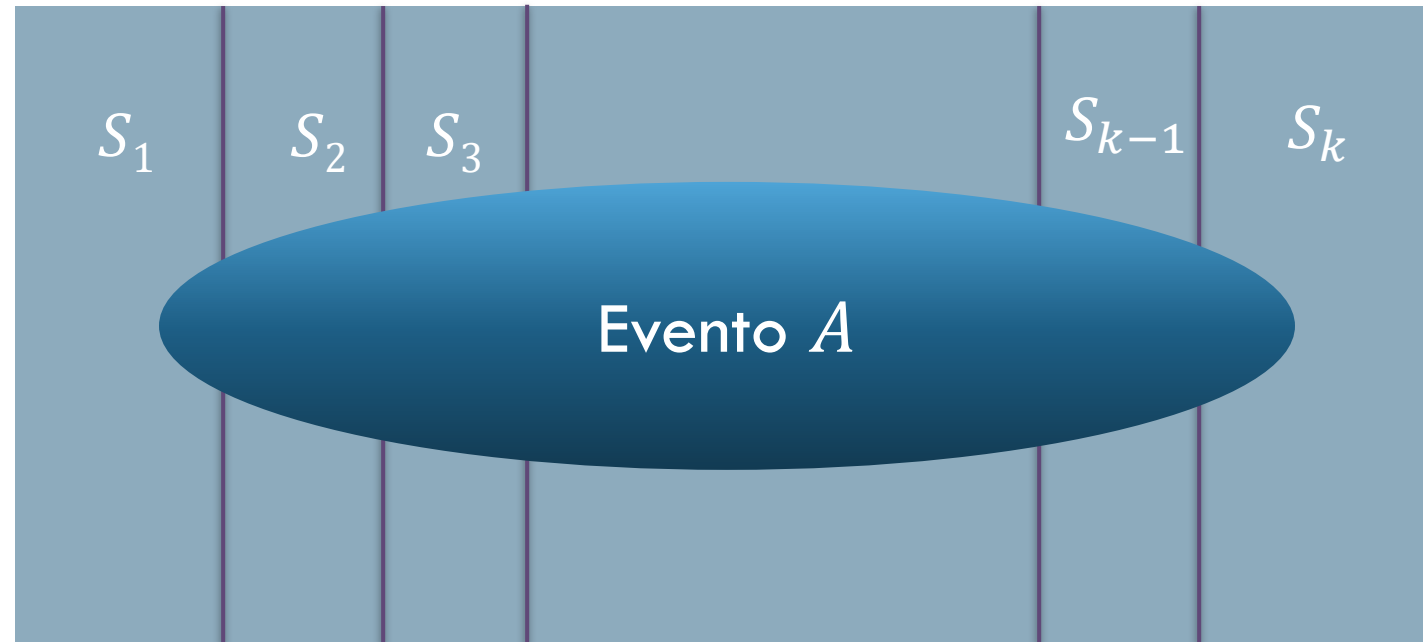
$P(A|nT) = 0,30 \rightarrow$ probabilidade do aluno acordar a tempo de ir às aulas dado que o despertador **não** tocou

CONT...

A probabilidade de ocorrência do evento A é definida pela relação:

$$P(A) = P(A/S_1)P(S_1) + P(A/S_2)P(S_2) + \dots + P(A/S_k)P(S_k)$$

$$P(A) = \sum_{i=1}^k P(A/S_i)P(S_i)$$



RESPOSTA

$P(T) = 0,70 \rightarrow$ probabilidade do despertador tocar

$P(nT) = 0,30 \rightarrow$ probabilidade do despertador **não** tocar

$P(A|T) = 0,80 \rightarrow$ probabilidade do aluno acordar a tempo de ir às aulas dado que o despertador tocou

$P(A|nT) = 0,30 \rightarrow$ probabilidade do aluno acordar a tempo de ir às aulas dado que o despertador **não** tocou

Aplicando o teorema da probabilidade total: $P(A)$ probabilidade de acordar a tempo:

$$P(A) = P(A|T)P(T) + P(A|nT)P(nT) = 0,8 \times 0,7 + 0,3 \times 0,3 = 0,65$$

$$P(T|A) = P(T) \times \frac{P(A|T)}{P(A)} = 0,70 \times \frac{0,80}{0,65}$$

RESOLVA...

Em uma fábrica, os setores A, B e C têm 40, 50 e 10% do total de operários, respectivamente. Dos operários de cada setor, 3, 5 e 2%, respectivamente, estão em férias. Escolhido ao acaso um operário desta fábrica, pede-se:

- a. Qual a probabilidade do operário estar em férias?
- b. Sabendo-se que o operário está em férias, qual a probabilidade dele ser do setor B?

RESOLVA...

Em uma fábrica, os setores A, B e C têm 40, 50 e 10% do total de operários, respectivamente. Dos operários de cada setor, 3, 5 e 2%, respectivamente, estão em férias. Escolhido ao acaso um operário desta fábrica, pede-se:

- a. Qual a probabilidade do operário estar em férias?
- b. Sabendo-se que o operário está em férias, qual a probabilidade dele ser do setor B?

a. R: 3,9%

b. R: 64%

REPAROS EM LINHA DE PRODUÇÃO

Mário é responsável por uma linha de produção automatizada, e recorre a companhias parceiras para fazerem reparos quando ocorrem falhas. A companhia **A** atende 20% das avarias e faz um reparo incompleto 1 vez em 20. A companhia **B** atende 60% das avarias e faz um reparo incompleto 1 vez em 10. A companhia **C** que atende 15% das avarias e faz um reparo incompleto 1 vez em 10. A companhia **D** atende 5% das quebras e faz um reparo incompleto 1 vez em 20. Para o próximo problema com a linha de produção diagnosticada como sendo devido a um reparo inicial incompleto, qual é a probabilidade que esse reparo inicial tenha sido feito pela companhia **A**?

RESPOSTA

$$P(A) = 0,20; \quad P(B) = 0,60; \quad P(C) = 0,15; \quad P(D) = 0,05$$

$$P(I|A) = 1/20; \quad P(I|B) = 1/10; \quad P(I|C) = 1/10; \quad P(I|D) = 1/20$$

$$P(A|I) = P(A) \frac{P(I|A)}{P(I)} = 0,20 \frac{1/20}{1/20 \times 0,20 + 1/10 \times 0,60 + 1/10 \times 0,15 + 1/20 \times 0,05}$$

$$P(A|I) = 11,429\%$$

APLICAÇÃO DE BAYES

Um exemplo excelente e amplamente utilizado dos benefícios do Teorema de Bayes está na análise de um teste de diagnóstico médico.

Testes não são o evento.

Os testes são falhos. Testes detectam coisas que não existem (falso positivo) e perdem coisas que existem (falso-negativo).

O teorema de Bayes converte os resultados do seu teste na probabilidade real do evento.

Leia o *Andar do Bêbado*, págs. 126-130.

DIAGNÓSTICO MÉDICO

A capacidade do teste para detectar uma doença é chamada de sensibilidade, e essa probabilidade é denominada de VPP (valor preditivo positivo).

$$VPP = P(\text{Doente} | \text{Exame positivo})$$

Por outro lado, a capacidade do teste dar negativo, dado que a pessoa não está doente é chamada especificidade e essa probabilidade é denominada de VPN (valor preditivo negativo).

$$VPN = P(\text{Não Doente} | \text{Exame negativo})$$

Às vezes, uma pessoa tem determinada doença, mas o teste não a detecta (falso negativo). Ou a pessoa não tem a doença e o teste a detecta (falso positivo).

DIAGNÓSTICO MÉDICO

A capacidade do teste para detectar uma doença é denominada de **VPP (valor preditivo positivo)**. VPP é a proporção de pessoas com teste positivo que realmente têm a doença,

$$VPP = P(\text{Doente} | \text{Exame positivo})$$

Por outro lado, a capacidade do teste dar negativo, dado que a pessoa não está doente é denominada de **VPN (valor preditivo negativo)**. VPN é a proporção daqueles com resultado negativo que não têm a doença,

$$VPN = P(\text{Não Doente} | \text{Exame negativo})$$

A **sensibilidade** é a taxa de verdadeiros positivos. A **especificidade** é a taxa de verdadeiros negativos.

DOENTE

Verdadeiro
Positivo (VP)

SAUDÁVEL

Falso Positivo
(FP)

$$VPP = \frac{VP}{VP + FP}$$

Falso
Negativo
(FN)

Verdadeiro
Positivo (VP)

$$VPN = \frac{VN}{VN + FN}$$

sensibilidade

$$\frac{VP}{VP + FN}$$

especificidade

$$\frac{VN}{VN + FP}$$

COMO ASSIM?

$$P(Do) = 0,001$$

Mário faz um exame para detectar uma rara e grave doença, que atinge 0,1% da população. O médico informa que o falso positivo do exame é de 1%, ou seja, apenas 1% das pessoas saudáveis recebem erradamente a indicação, pelo exame, de que estão com a doença. Por outro lado, se o paciente tem a doença, o exame dá positivo em 99% dos casos. O resultado do exame de Mário é positivo. Ele se desespera, achando que tem 99% de chance de estar com a doença.

$$P(+|nDo) = 0,01$$

$$P(+|Do) = 0,99$$

Você tem algo a dizer para ele?

Acalme o Mário afirmando que dizer que ele tem 99% de chance de estar com a doença é uma **falácia!!!! Não deve ver a $P(+|Do)$ e sim $P(Do|+)$. São duas coisas diferentes!**

O termo falácia deriva do verbo latino fallere, que significa enganar. Designa-se por falácia um raciocínio errado com aparência de verdadeiro.

PROBLEMA DO MÁRIO...

Diga que está mal em probabilidade, mas talvez não de saúde. Se fizer a análise correta, ele vai descobrir que a chance de ter a doença, com base nestas informações, é de...

$$P(Do) = 0,001 \therefore P(nDo) = 0,999$$

$$P(+|Do) = 0,99 \therefore P(+|nDo) = 0,01$$

$$P(Do|+) = P(Do) \frac{P(+|Do)}{P(+)} = P(Do) \frac{P(+|Do)}{P(+|Do)P(Do) + P(+|nDo)P(nDo)}$$

$$P(Do|+) = 0,001 \frac{0,99}{0,99 \times 0,001 + 0,01 \times 0,999} = 0,09$$

Você, então, pode dizer: Mário...
A situação não é tão ruim assim....

DETECÇÃO DE SPAM

O Mário, que nunca teve a doença do problema anterior, está usando um novo software para detectar spam: **SpamAssassin**. O *SpamAssassin* é um sistema inteligente, que analisa cada mensagem que chega no servidor, e através de regras próprias, determina se o e-mail é um SPAM ou não. O programa é treinado pelos usuários. Ele procura padrões nas palavras dos e-mails marcados como spam pelo usuário. O software é atualizado regularmente, para que novas regras entrem no sistema, e passem a detectar os SPAMs com maior precisão.

Por exemplo, o software pode ter aprendido que a palavra “Rolex” aparece em 10% dos e-mails marcados como spam. Supondo que 0,1% dos e-mails não spam incluam a palavra “Rolex” e 50% de todos os e-mails recebidos pelo usuário são spam, encontre a probabilidade de um e-mail ser spam se a palavra “Rolex” aparecer nele.

$$P(\text{Spam}) = 50\%$$

$$P(\text{Rolex}|\text{Spam}) = 10\% \quad P(\text{Rolex}|n\text{Spam}) = 0,1\%$$

O PROBLEMA DE MONTY HALL

Extraído do livro *Think Bayes*, de Allen B. Downey (livro pode ser baixado grátis pela internet)

O problema de Monty Hall pode ser a questão mais controversa da história das probabilidades. O cenário é simples, mas a resposta correta é tão contraintuitiva que muitas pessoas simplesmente não a aceitam, e muitas pessoas inteligentes se envergonharam não apenas por errar, mas por argumentar agressivamente em público. Monty Hall foi o apresentador original do game show *Let's make a deal*. O problema do Monty Hall é baseado em um dos jogos regulares do programa:

Monty mostra três portas fechadas e diz que há um prêmio atrás de cada porta: um é um carro, os outros dois são menos valiosos, como manteiga de amendoim e unhas postiças. Os prêmios são organizados aleatoriamente. O objetivo do jogo é adivinhar em qual porta está com o carro. Se acertar, fica com o carro.

- Você escolhe uma porta, que chamaremos de porta A. Vamos chamar as outras portas de B e C.
- Antes de abrir a porta que você escolheu, Monty aumenta o suspense abrindo a Porta B ou C. Obviamente, abre a que não estiver com o carro. Se o carro está realmente atrás da porta A, Monty pode abrir aleatoriamente B ou C.
- Então Monty oferece a você a opção de manter sua escolha original ou mudar para a porta que ainda não foi aberta.

A questão é: avaliando a situação estatisticamente, você deve “manter” ou “trocar” de porta, ou não faz diferença?

Solution using Bayes Theorem

Suppose without loss of generality that the player chooses door 1 and that the host opens door 2 to reveal a goat.

Let A (B , C) be the event that the prize is behind door 1, (2, 3).

$$P(A) = P(B) = P(C) = \frac{1}{3}.$$

$$P(\text{opens 2}|A) = \frac{1}{2}. \quad P(\text{opens 2}|B) = 0, \quad P(\text{opens 2}|C) = 1.$$

$$\begin{aligned} P(\text{opens 2}) &= P(\text{opens 2}|A)P(A) + P(\text{opens 2}|B)P(B) + P(\text{opens 2}|C)P(C) \\ &= \frac{1}{2} \times \frac{1}{3} + 0 \times \frac{1}{3} + 1 \times \frac{1}{3} = \frac{1}{2} \end{aligned}$$

$$P(A|\text{opens 2}) = \frac{P(\text{opens 2}|A)P(A)}{P(\text{opens 2})} = \frac{\frac{1}{2} \times \frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}$$

so $P(C|\text{opens 2}) = \frac{2}{3}$ and it is better to switch.

MAS ONDE APLICAREMOS NOSSO MODELO?

A Posteriori: probabilidade da hipótese ser verdadeira, dada a ocorrência da evidência

A Priori: aquilo que acreditamos antes de qualquer nova evidência. Nossa crença.

Redimensionador: probabilidade que a nova evidência ocorra, dada que a hipótese seja verdadeira. Conhecido como verossimilhança.

$$P(\text{Hipótese}|\text{Evidência}) = \frac{P(\text{Hipótese}) \cdot P(\text{Evidência}|\text{Hipótese})}{P(\text{Evidência})}$$

Normalizador (marginal): para ajustar o numerador.

A proporção entre o redimensionador e o normalizador é muito importante.

Quando o redimensionador é maior que o normalizador, aumentamos a probabilidade anterior, caso contrário, a evidência diminui nossas probabilidades.

O TEOREMA DE BAYES EXPLICADO...

$$P(A|B) = P(A) \frac{P(B|A)}{P(B)}, P(B) \neq 0$$

A probabilidade a posteriori é condicionada à probabilidade a priori multiplicada pela verossimilhança normalizada.

A probabilidade a priori é baseada nas informações disponíveis separadamente do experimento. Representa o conhecimento do fenômeno antes dos dados serem observados.

A verossimilhança é deduzida dos dados e expressa todo o conhecimento do fenômeno contido nestes dados.

POR EXEMPLO...

Suponha que você esteja esperando a resposta de várias empresas a entrevistas que você fez. Como você está ficando nervoso, decide calcular a probabilidade de uma empresa específica fazer uma oferta, já que se passaram três dias e ela ainda não ligou para você.

A: receber uma oferta (“O”)

B: nenhuma ligação por 3 dias (“NL”)

20% Procura de
emprego é um
processo longo e
árido...

40% Not bad...

É difícil estimar
direto!

$$P(O|NL) = P(O) \frac{P(NL|O)}{P(NL)}$$

90% existem vários motivos
pelos quais eles podem não ter
chamado

$$P(O|NL) = 0,20 \frac{0,4}{0,9} = 8,9\%$$

PROBLEMAS DE CLASSIFICAÇÃO

Por exemplo, problema de classificação em termos Bayesianos – onde temos um conjunto de crenças anteriores e atualizamos nossas crenças à medida que observamos e coletamos evidências.

O objetivo é prever a que classe uma determinada observação pertence, dadas suas características.



Y = o rótulo da classe que estamos tentando prever. No nosso caso, os rótulos das classes seriam hamster, gato e cachorro.

X_1, X_2, X_3 , etc. = os recursos de nossa observação com os quais tentamos fazer uma previsão. Alguns recursos que podemos usar para diferenciar gatos, cães e hamsters são de tamanho e agilidade.

POR EXEMPLO

$P(Y = \blacksquare | Tamanho = médio, Agilidade = Não) = ????$

$$P(Y = \img alt="cat" data-bbox="251 458 296 545" | Médio, Não ágil) = P(G) \frac{P(\text{Médio e não ágil} | G)}{P(\text{Médio e não ágil})}$$

$$P(Y = \img alt="dog" data-bbox="244 618 301 706" | Médio, Não ágil) = P(C) \frac{P(\text{Médio e não ágil} | C)}{P(\text{Médio e não ágil})}$$

$$P(Y = \img alt="hamster" data-bbox="243 781 300 864" | Médio, Não ágil) = P(H) \frac{P(\text{Médio e não ágil} | H)}{P(\text{Médio e não ágil})}$$



ENOUGH IS
ENOUGH!

ACABOU...

Reveja a aula antes
de resolver os
exercícios.