Análise Estatística de dados

Inteligência Artificial

PECE Programa de Educação Continuada
Escola Politécnica da USP

USP

AULA 07 – PROBABILIDADE

Larissa Driemeier

PROGRAMA DO CURSO

Aula	Data	Conteúdo da Aula
01	18/02	Aula Inaugural
02	25/02	Introdução ao Curso. Noções de Álgebra Linear , Geometria Analítica Parte I
03	11/03	Noções de Álgebra Linear , Geometria Analítica Parte II
04	18/03	Decomposição de valor singular (SVD)
05	25/03	Otimização: derivadas, derivadas parciais (operadores gradiente, Jacobiano, Hessiano e Laplaciano), algoritmos de gradiente, noções de algoritmos genéticos
06	01/04	Variáveis independentes e não independentes. Estatística Descritiva e Indutiva. Definições de medidas de dispersão e tendência central.
07	08/04	Probabilidade e Teorema de Bayes
08	15/04	Modelos de probabilidade discretos/contínuos
09	22/04	Intervalo de confiança e teste de hipóteses
10	29/04	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.

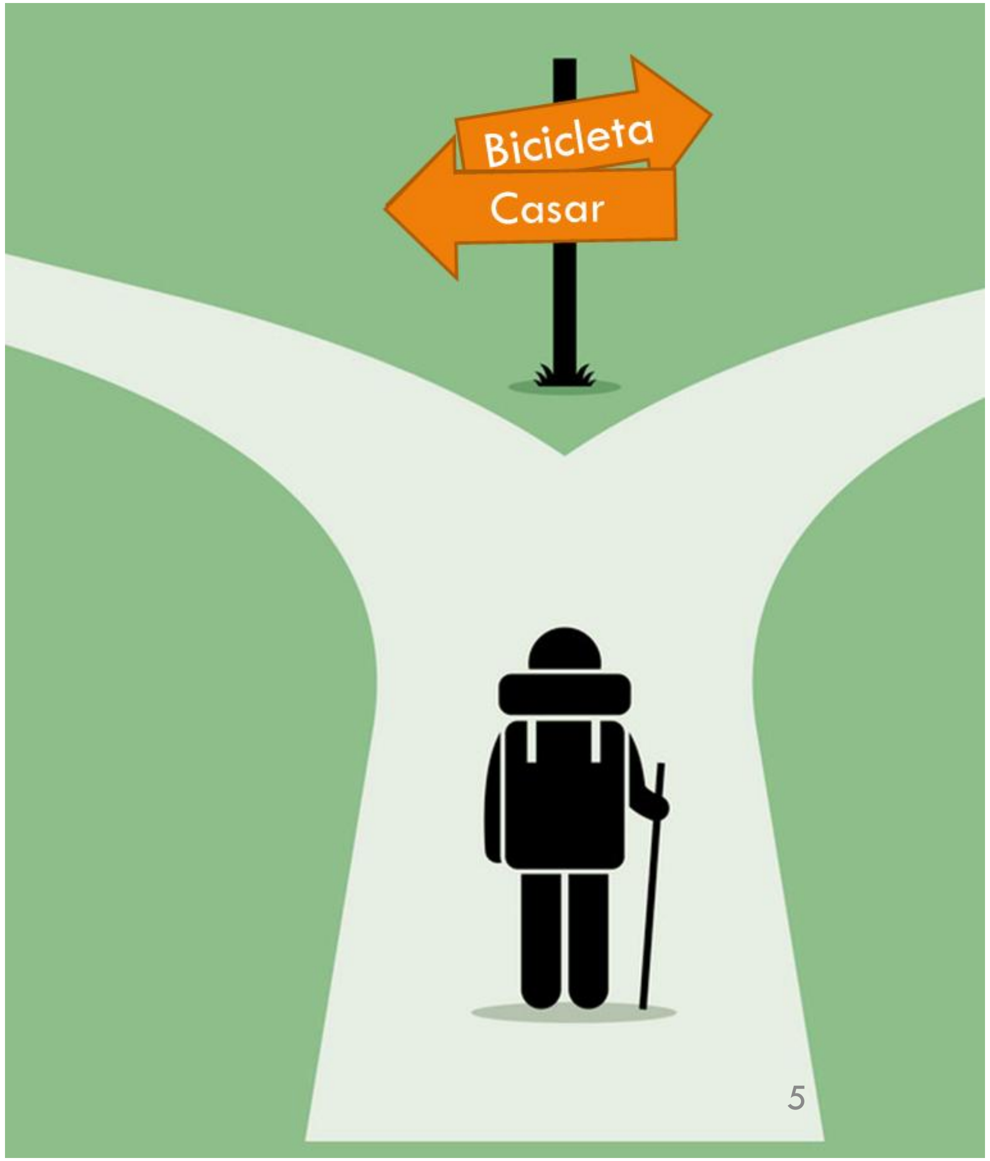
QUAL É A VANTAGEM DA PROBABILIDADE E ESTATÍSTICA?

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

O que é variável aleatória?
Como se define a probabilidade?
O que é um modelo probabilístico?
O 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):

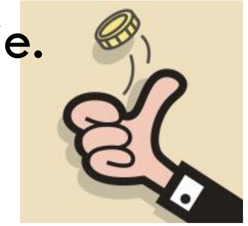
- **Face obtida** em cada jogada - variável qualitativa - **não é uma variável aleatória**.
- **Número de caras** - **variável aleatória** associada à variável qualitativa estudada.

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:

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

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*, $P(\omega_i)$

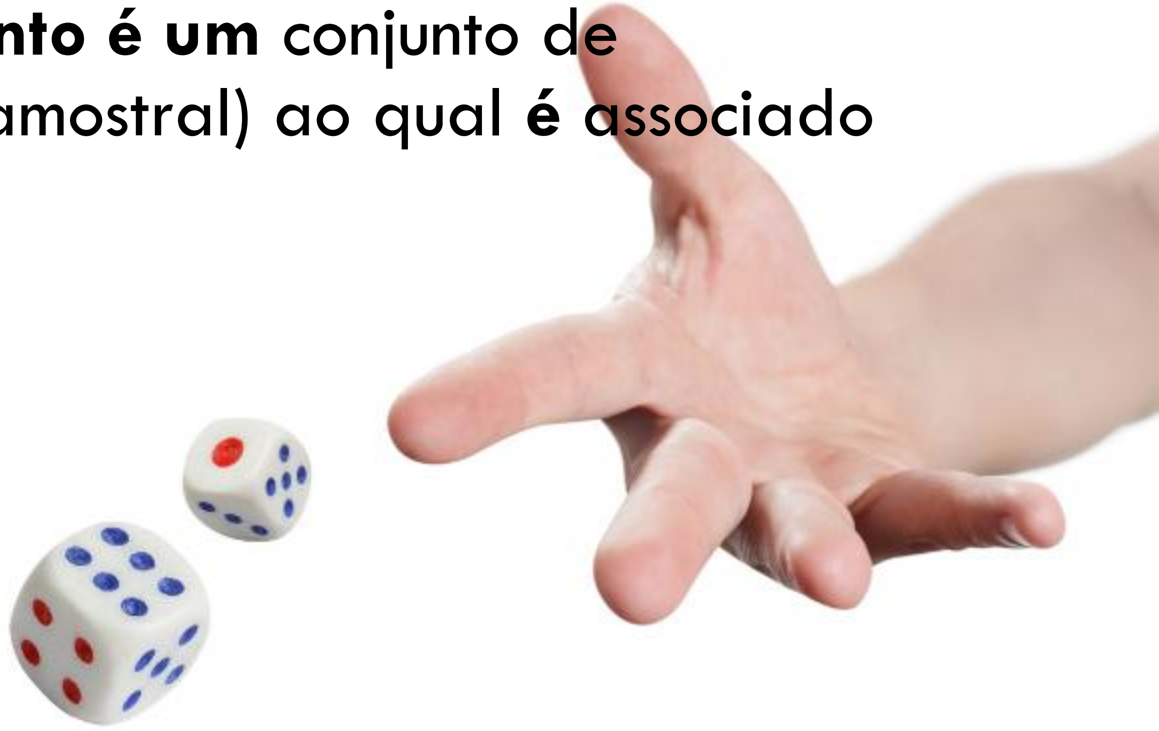
Para cada *ponto amostral*, podemos considerar a probabilidade como uma função que recebe um elemento do espaço amostral e mapeia o resultado para um número real não negativo e menor ou igual a 1.

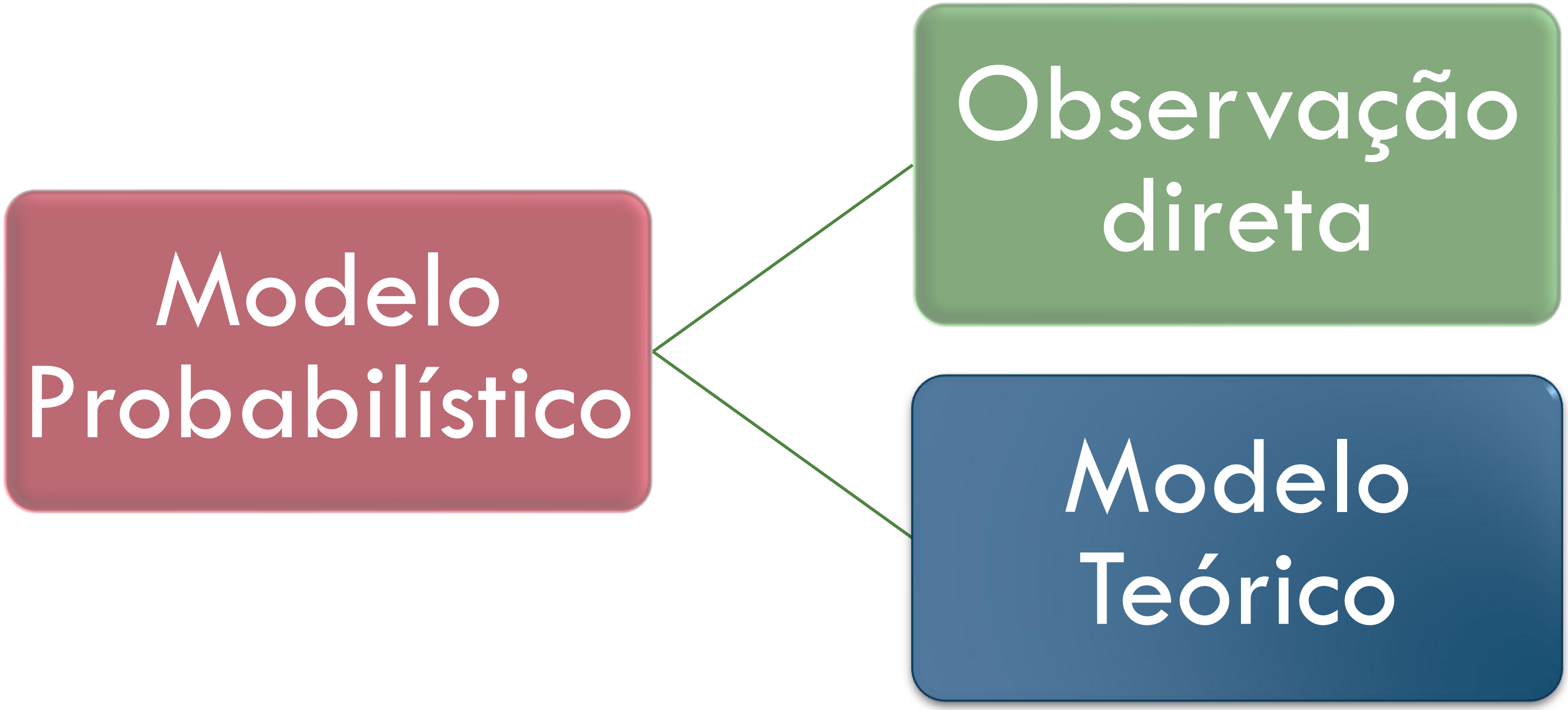
EVENTO

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.




Modelo Probabilístico
Observação direta
Modelo Teórico

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 freqüências 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.

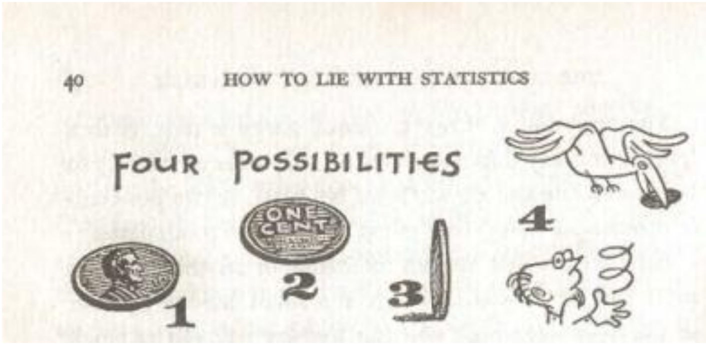
Por exemplo, no lançamento de um dado, as suposições teóricas são:

- 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:

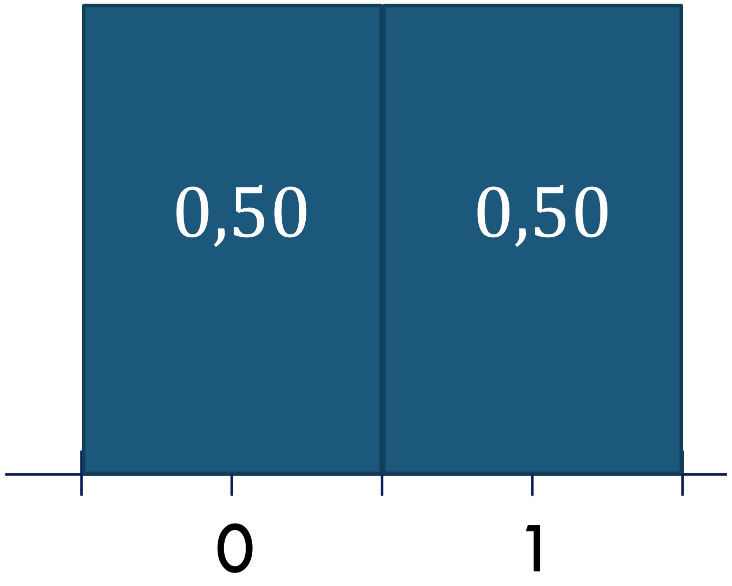
Face	1	2	3	4	5	6
Frequência teórica	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

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** $P(X = x)$
$Coroa$	0	0,5
$Cara$	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 ½ de sair **Cara** e ½ de sair **Coroa**, se a moeda é perfeitamente simétrica e homogênea.

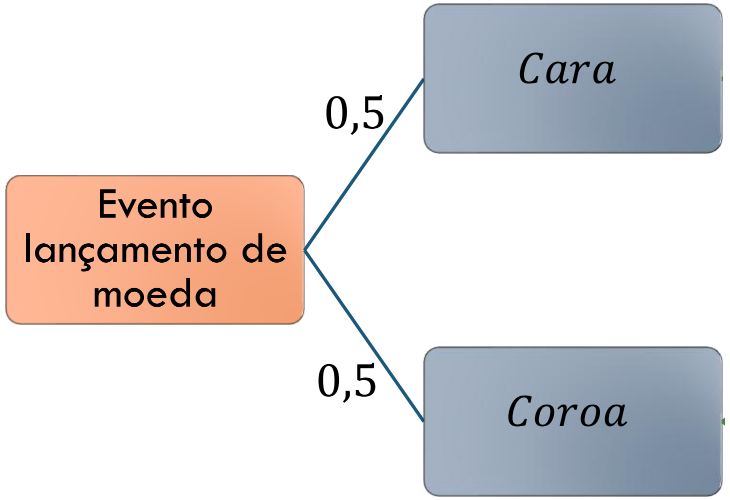
EXERCÍCIO

Resultados Possíveis	Variáveis aleatórias (x)	Probabilidade $P(X = x)$
$Coroa - Coroa$	0	$0{,}5x0{,}5 = 0{,}25$
$Cara - Coroa$	1	$0{,}5x0{,}5 = 0{,}25$
$Coroa - Cara$	1	$0{,}5x0{,}5 = 0{,}25$
$Cara - Cara$	2	$0{,}5x0{,}5 = 0{,}25$
	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

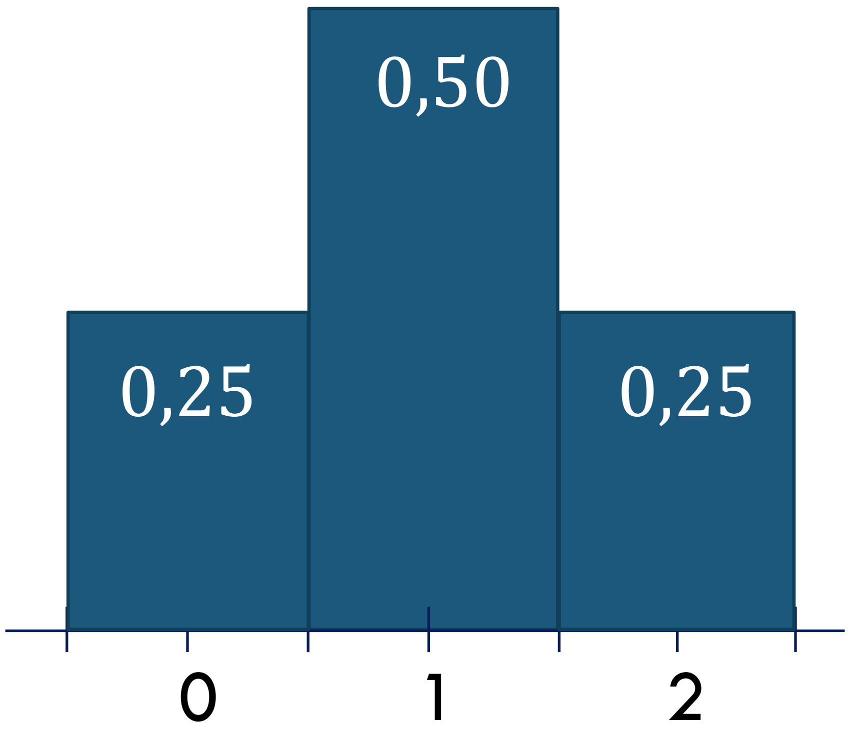


Se A é o evento "número de caras", então:

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

PECE Programa de Educação Continuada
Escola Politécnica da USP

Se A é um evento, então:

$P(A) = \Sigma\, P(\omega_j)$, 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

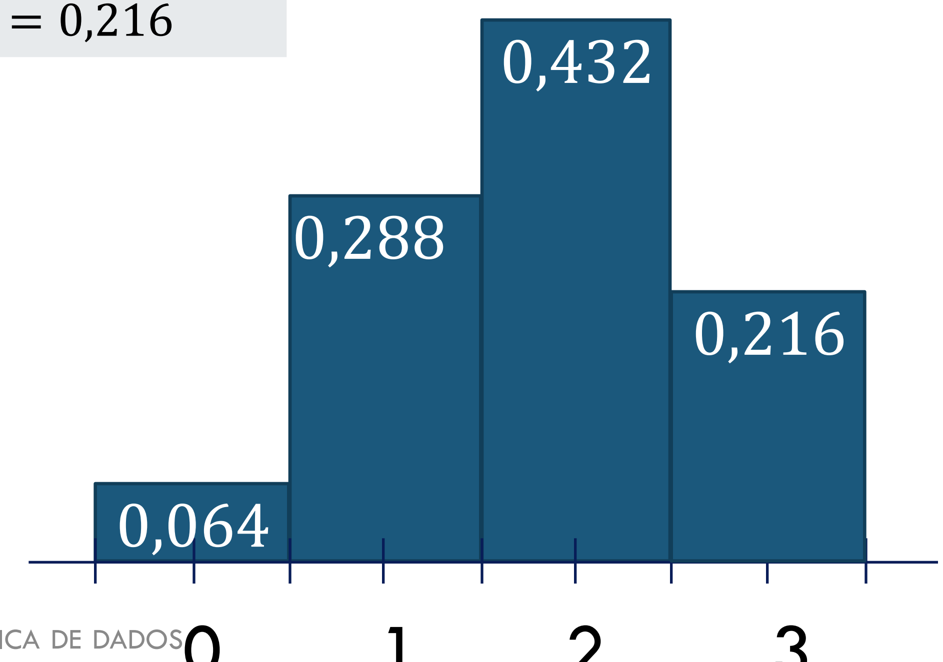
Um grande lote de peças possui 60% dos itens com algum tipo de defeito. Construir a distribuição de probabilidades para a variável aleatória **número de itens com defeito dentre 2 sorteados aleatoriamente.**

EXERCÍCIO 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
$B\ B\ B$	0	$0,4 \times 0,4 \times 0,4 = 0,064$
$B\ B\ D$	1	$0,4 \times 0,4 \times 0,6 = 0,096$
$B\ D\ B$	1	$0,4 \times 0,6 \times 0,4 = 0,096$
$D\ B\ B$	1	$0,6 \times 0,4 \times 0,4 = 0,096$
$B\ D\ D$	2	$0,4 \times 0,6 \times 0,6 = 0,144$
$D\ B\ D$	2	$0,6 \times 0,4 \times 0,6 = 0,144$
$D\ D\ B$	2	$0,6 \times 0,6 \times 0,4 = 0,144$
$D\ D\ D$	3	$0,6 \times 0,6 \times 0,6 = 0,216$

x	P(X=x)
0	0,064
1	$0,096x3 = 0,288$
2	$0,144x3 = 0,432$
3	0,216

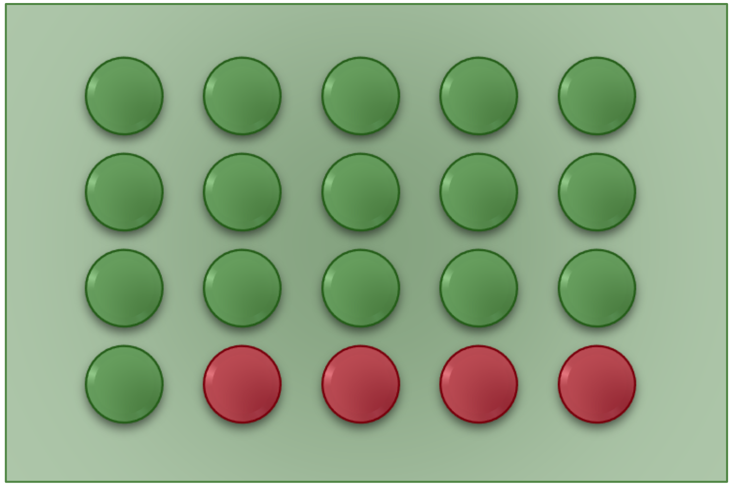


EXEMPLO FINAL...

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?

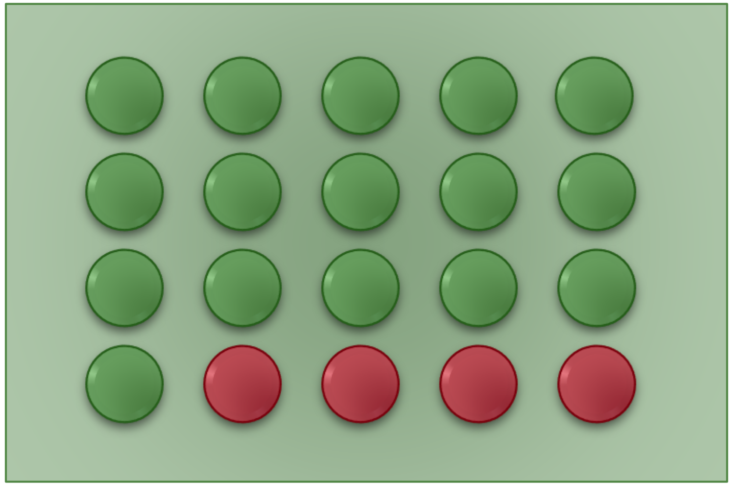
SITUAÇÃO INICIAL,

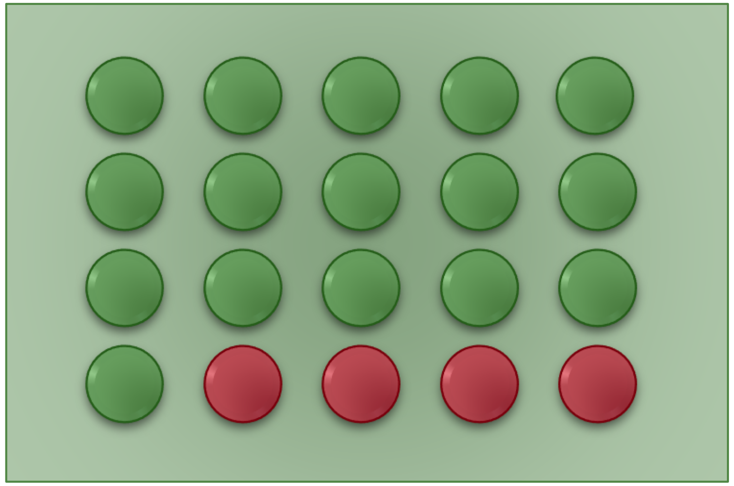


B - Peça Boa

D - Peça Defeituosa

PRIMEIRA PEÇA





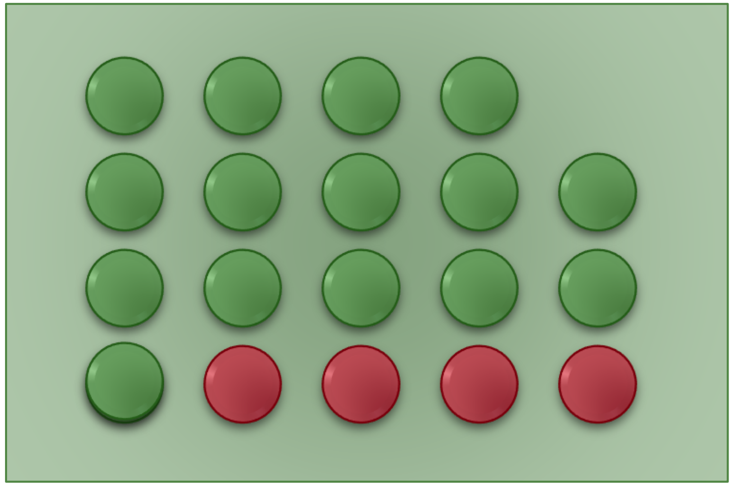
$$P(B) = \frac{16}{20}$$

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

RETIRADA DA SEGUNDA PEÇA,

Se a primeira peça for

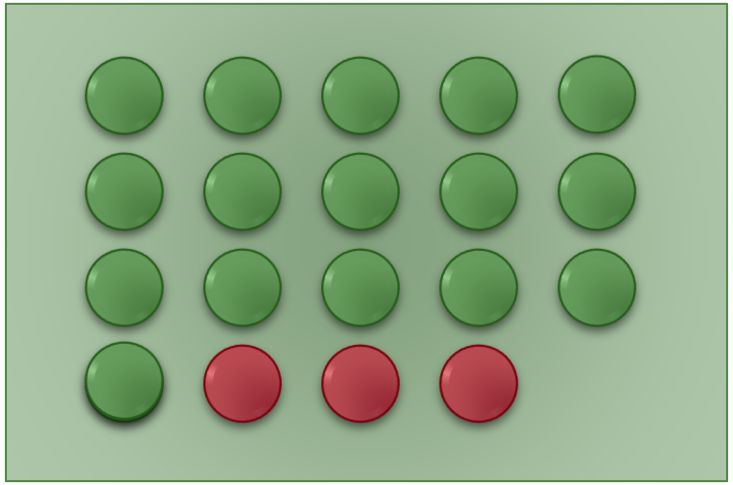
Boa



P(2ª peça ser boa) = 15/19

P(2ª peça ser defeituosa) = 4/19

Defeituosa



P(2ª peça ser boa) = 16/19

P(2ª peça ser defeituosa) = 3/19

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,

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

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

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 $P(X = x)$
Sim Não	1	$0{,}4 \times 0{,}5 = \mathbf{0,2}$
Sim Sim	2	$0{,}4 \times 0{,}5 = \mathbf{0,2}$
Não Sim	1	$0{,}6 \times 0{,}5 = \mathbf{0,3}$
Não Não	0	$0{,}6 \times 0{,}5 = \mathbf{0,3}$

$\mathbf{0,7}$

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.

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

PROPRIEDADES DA PROBABILIDADE

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

Como toda freqüência 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

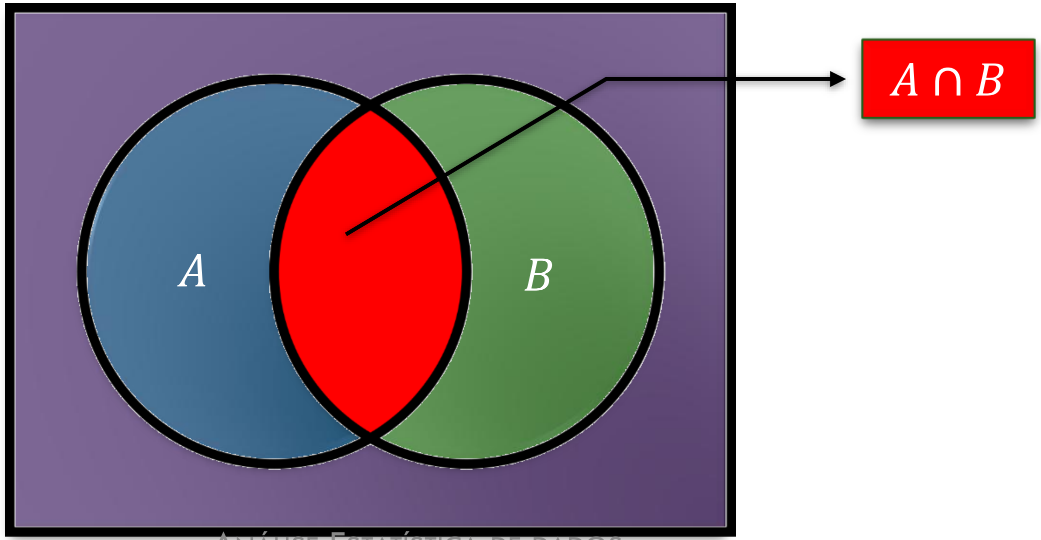
$$P(S) = 1$$

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

$$P(\emptyset) = 0$$

EVENTO INTERSECÇÃO

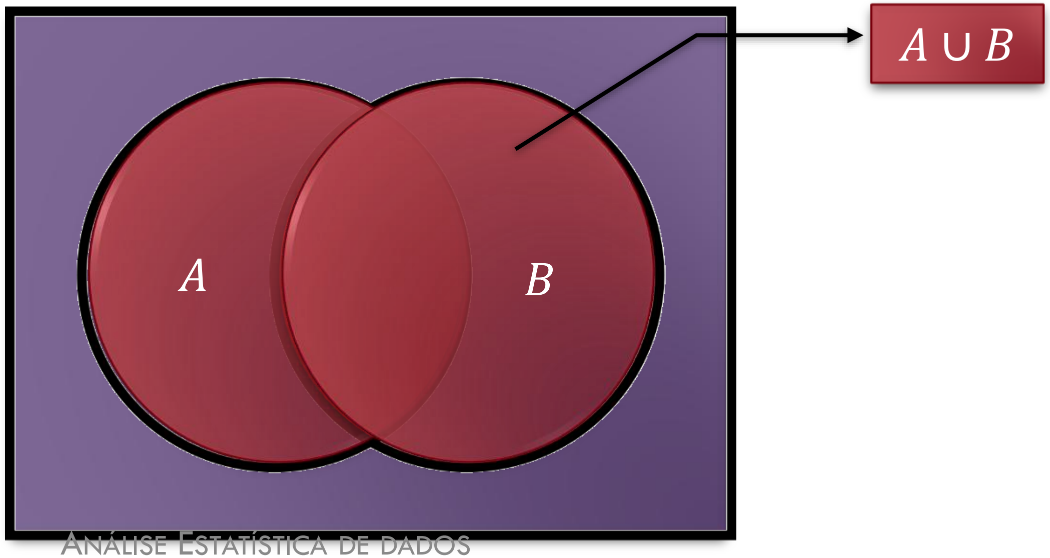
O evento intersecção $(A \cap B)$ significa que A e B ocorrem simultaneamente. Esta probabilidade é calculada com o emprego do *Teorema da Probabilidade Condicionada*, a ser visto na seqüência desta aula.



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...

PECE Programa de Educação Continuada
Escola Politécnica da USP

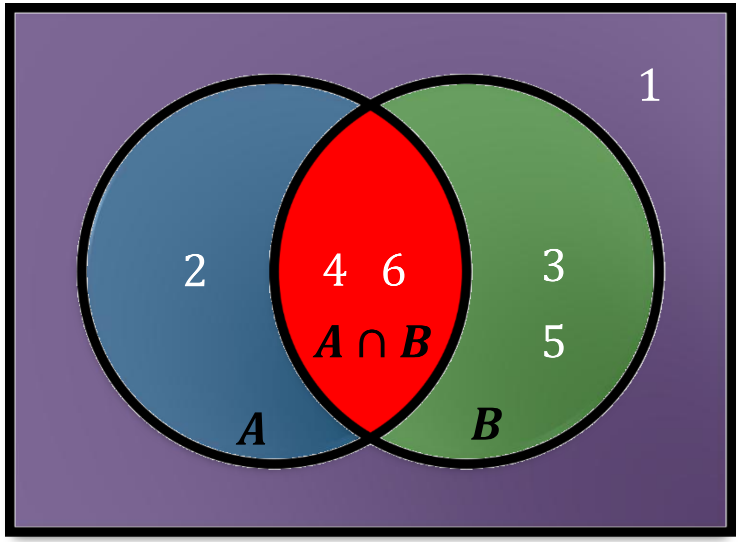
Suponha que você joga um dado, se A é o evento números pares e B é o evento números maiores ou iguais a 3, qual a probabilidade de ocorrer A **e** B ao mesmo tempo?

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

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

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

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



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

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

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?

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

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

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

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

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

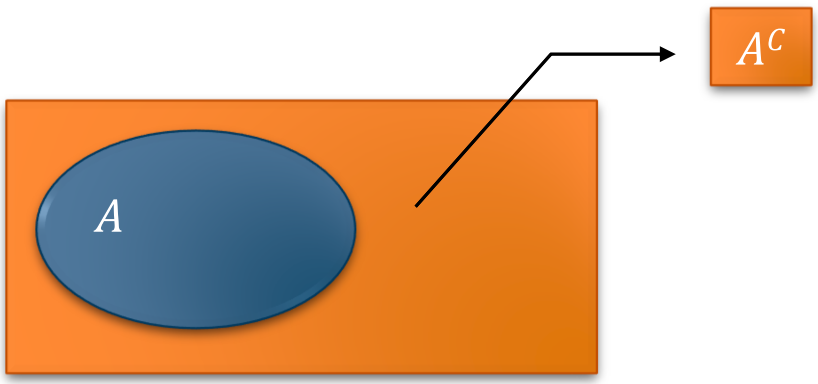
Resolva o problema do aluno respondendo às questões da prova usando a equação: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

EVENTO COMPLEMENTAR

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

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

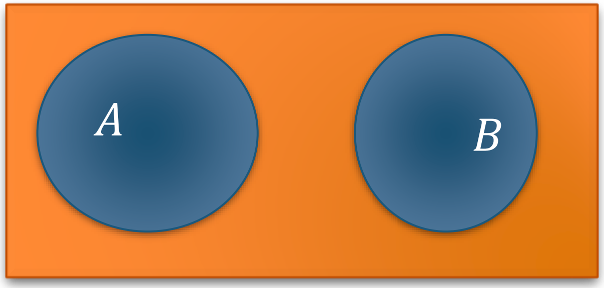
$$P(A) + P(A^C) = 1$$



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:

$$\begin{bmatrix} 1,1 & 1,2 & 1,3 & 1,4 & 1,5 & 1,6 \\ 2,1 & 2,2 & 2,3 & 2,4 & 2,5 & 2,6 \\ 3,1 & 3,2 & 3,3 & 3,4 & 3,5 & 3,6 \\ 4,1 & 4,2 & 4,3 & 4,4 & 4,5 & 4,6 \\ 5,1 & 5,2 & 5,3 & 5,4 & 5,5 & 5,6 \\ 6,1 & 6,2 & 6,3 & 6,4 & 6,5 & 6,6 \end{bmatrix}$$

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

Evento A:

$P(total < 4) = 3/36$

$$\begin{bmatrix} 1,1 & 1,2 & 1,3 & 1,4 & 1,5 & 1,6 \\ 2,1 & 2,2 & 2,3 & 2,4 & 2,5 & 2,6 \\ 3,1 & 3,2 & 3,3 & 3,4 & 3,5 & 3,6 \\ 4,1 & 4,2 & 4,3 & 4,4 & 4,5 & 4,6 \\ 5,1 & 5,2 & 5,3 & 5,4 & 5,5 & 5,6 \\ 6,1 & 6,2 & 6,3 & 6,4 & 6,5 & 6,6 \end{bmatrix}$$

Evento B:

$P(\,sair\ o\ número\ 4) = 11/36$

$$\therefore P(A \cup B) = \frac{3}{36} + \frac{11}{36}$$

$$= 14/36 = \mathbf{7/18}$$

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:

$$\begin{bmatrix} 1,1 & 1,2 & 1,3 & 1,4 & 1,5 & 1,6 \\ 2,1 & 2,2 & 2,3 & 2,4 & 2,5 & 2,6 \\ 3,1 & 3,2 & 3,3 & 3,4 & 3,5 & 3,6 \\ 4,1 & 4,2 & 4,3 & 4,4 & 4,5 & 4,6 \\ 5,1 & 5,2 & 5,3 & 5,4 & 5,5 & 5,6 \\ 6,1 & 6,2 & 6,3 & 6,4 & 6,5 & 6,6 \end{bmatrix}$$

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

Evento A:

$P(sair\ o\ número\ 5)\ =\ 11/36$

$$\begin{bmatrix} 1,1 & 1,2 & 1,3 & 1,4 & 1,5 & 1,6 \\ 2,1 & 2,2 & 2,3 & 2,4 & 2,5 & 2,6 \\ 3,1 & 3,2 & 3,3 & 3,4 & 3,5 & 3,6 \\ 4,1 & 4,2 & 4,3 & 4,4 & 4,5 & 4,6 \\ 5,1 & 5,2 & 5,3 & 5,4 & 5,5 & 5,6 \\ 6,1 & 6,2 & 6,3 & 6,4 & 6,5 & 6,6 \end{bmatrix}$$

Evento B:

$P(\ sair\ o\ número\ 6)$
$=\ 11/36$

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

$\therefore P(\ sair\ 5\ ou\ 6)\ =$
$= 11/36\ +\ 11/36\ -\ 2/36$
$=\ 20/36\ =\ 5/9$

E SE A PERGUNTA FOSSE...

PECE Programa de Educação Continuada
Escola Politécnica da USP

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.

$$\begin{bmatrix} & & 1,3 & & & \\ & & 2,3 & & & \\ 3,1 & 3,2 & 3,3 & 3,4 & 3,5 & 3,6 \\ & & 4,3 & & & \\ & & 5,3 & & & \\ & & 6,3 & & & \end{bmatrix}$$

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

Probabilidade conjunta (*joint*)

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

Probabilidade condicional

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.

Idade		Posição na carreira				
		R_1	R_2	R_3	R_4	Total
	< 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. $P(50|R_3) = \frac{P(50,R_3)}{P(R_3)} = \frac{36/1164}{320/1164} = 0{,}113$

RESPONDA

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

PARA VOCÊ FAZER EM CASA

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

Queremos saber,

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

	$Peru$	$Brasil$	$Bolívia$	Total
$Compra = 1$	20	50	10	80
$Não\ compra = 0$	300	500	200	1000
Total	320	550	210	1080

	Peru	*Brasil*	*Bolívia*	Total
$Compra = 1$	20	50	10	80
$Não\ compra = 0$	300	500	200	1000
Total	320	550	210	1080

PECE Programa de Educação Continuada
Escola Politécnica da USP

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

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

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

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

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

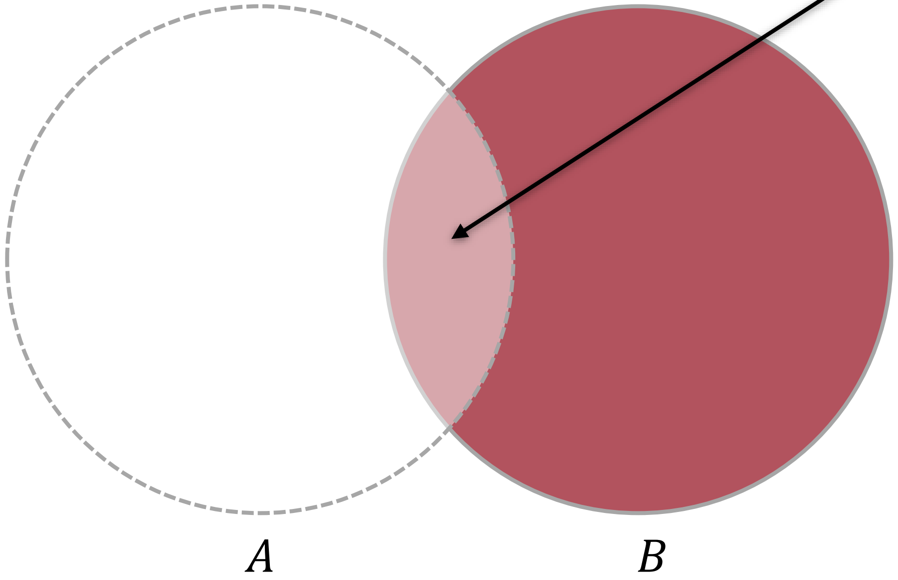
$$P(1|Bolívia) = \frac{P(1, Bolívia)}{P(Bolívia)} = \frac{0{,}009}{0{,}19} = \mathbf{0,05}$$

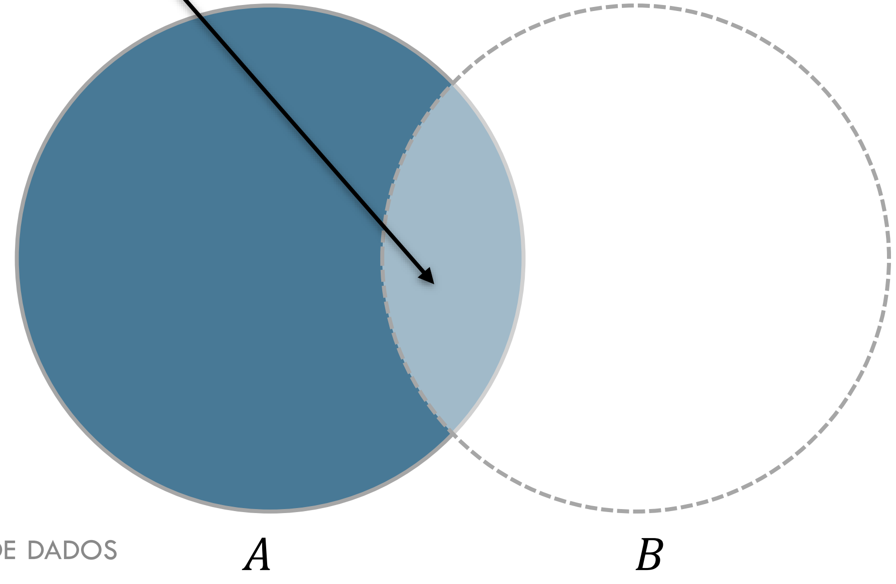
$$P(0|Bolívia) = \frac{P(0, Bolívia)}{P(Bolívia)} = \frac{0{,}185}{0{,}19} = \mathbf{0,97}$$

UMA DEFINIÇÃO EXTRA...

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

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



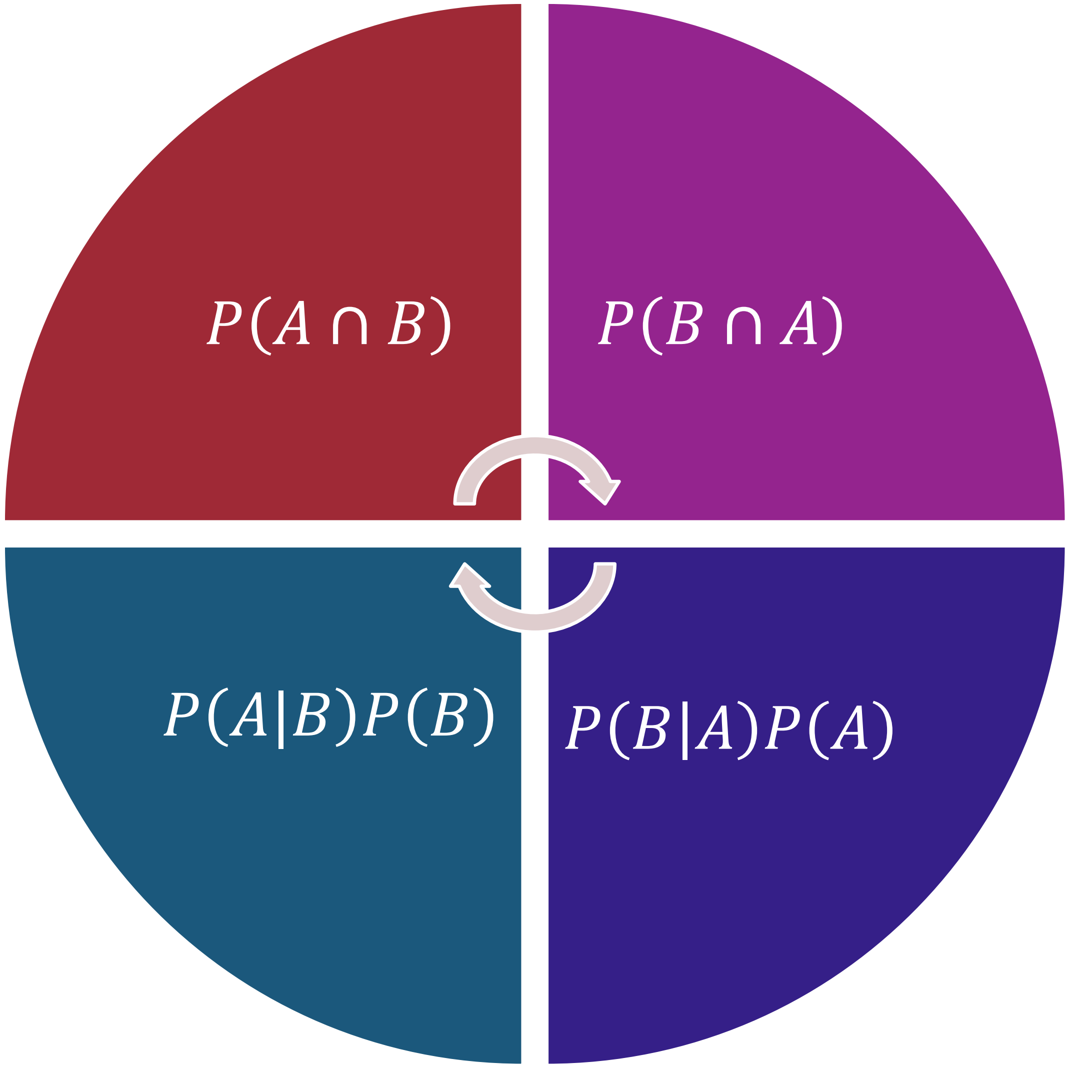


INTERSECÇÃO

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

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

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



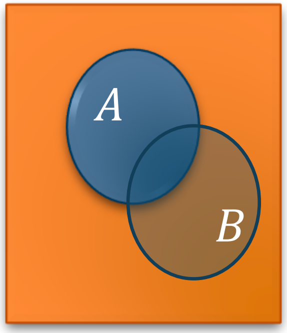
EVENTOS INDEPENDENTES

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

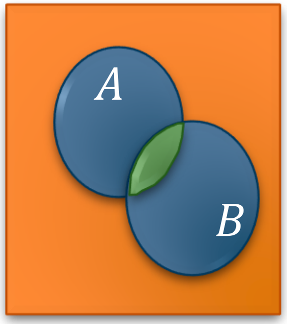
$P(B|A) = P(B)$ ou $P(A|B) = P(A)$



$A \cup B$

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

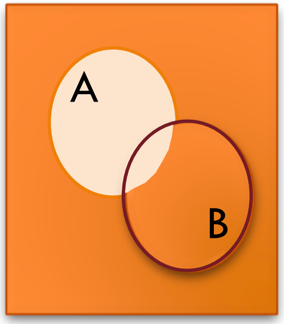
$$P(A \cup B) = P(A) + P(B)$$ eventos mutuamente exclusivos



$A \cap B$

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

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



A^c

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

TEOREMA DE BAYES

Uma das equações famosas do mundo da estatística e probabilidade.

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

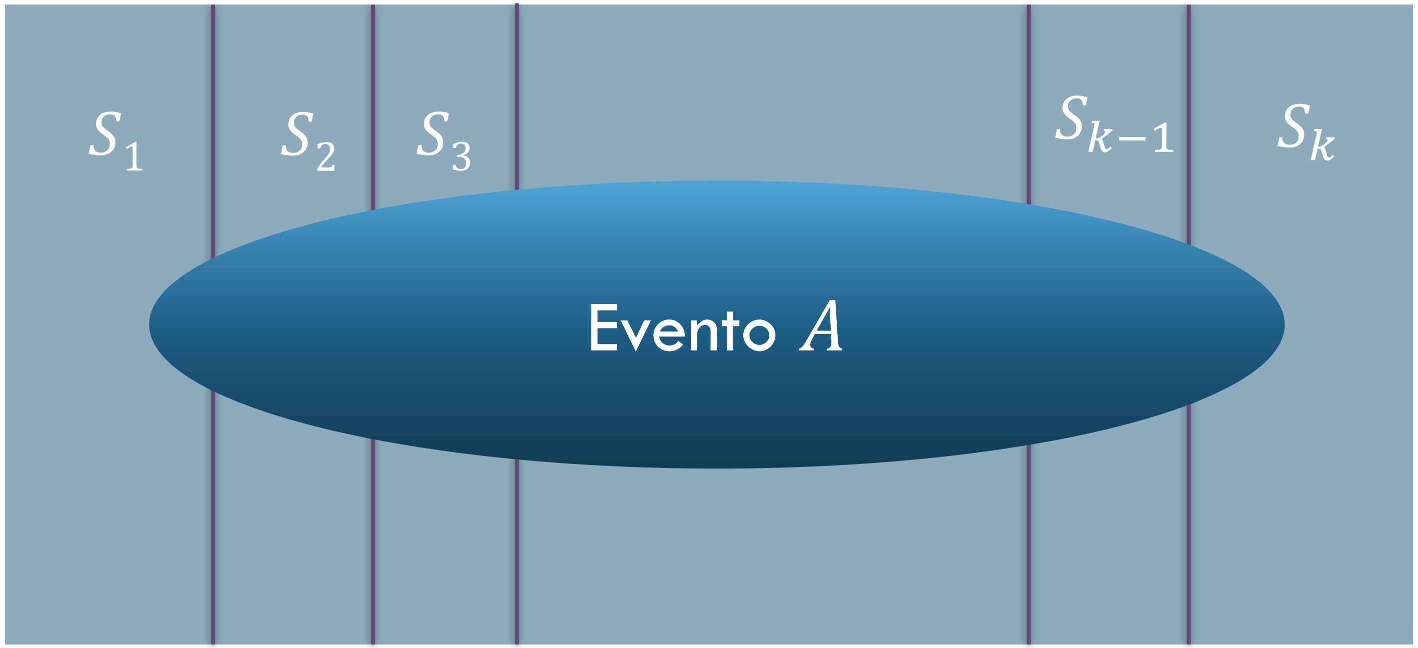
Regra de Bayes é apenas uma formalização da lógica que seguimos até agora...

Thomas BAYES
(1702 - 1761)

TEOREMA DA PROBABILIDADE TOTAL

Vamos dividir um espaço amostral em subespaços, tal que:

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



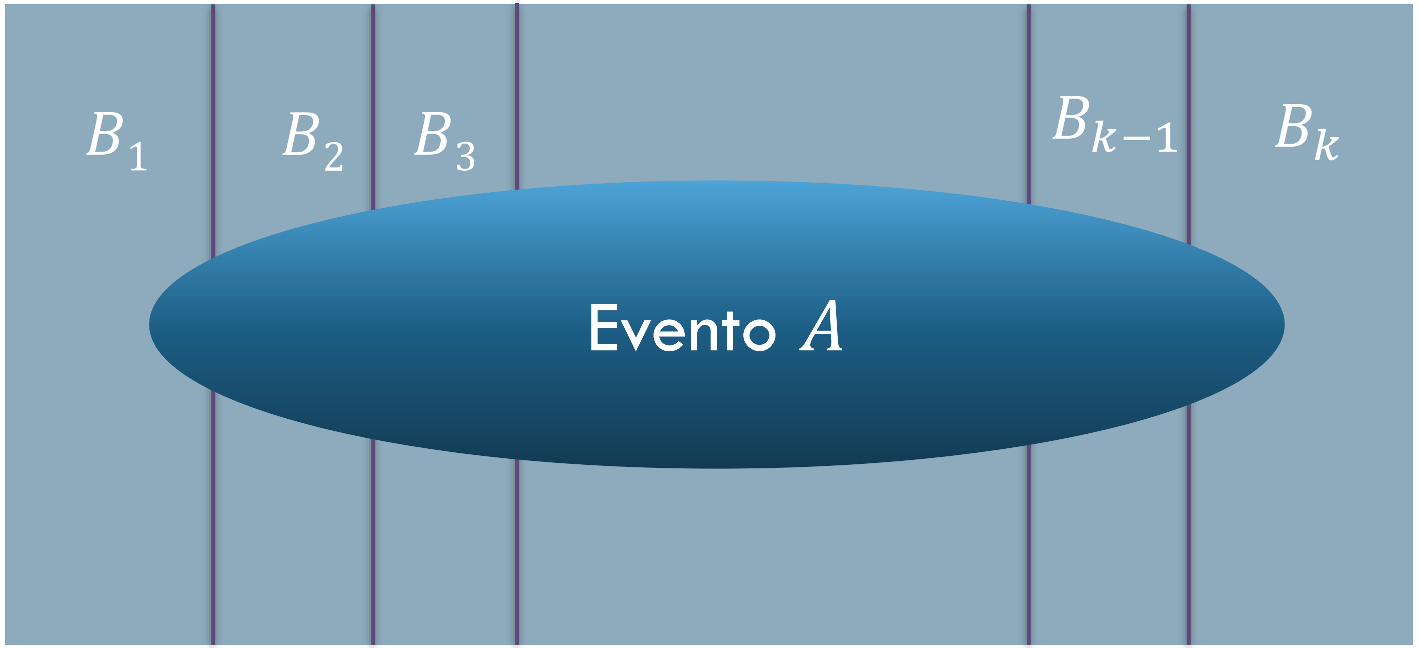
Eventos mutuamente exclusivos e exaustivos.

CONT...

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

$$P(A) = P(A/B_1)P(B_1) + P(A/B_2)P(B_2)+\ldots+P(A/B_k)P(B_k)$$

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

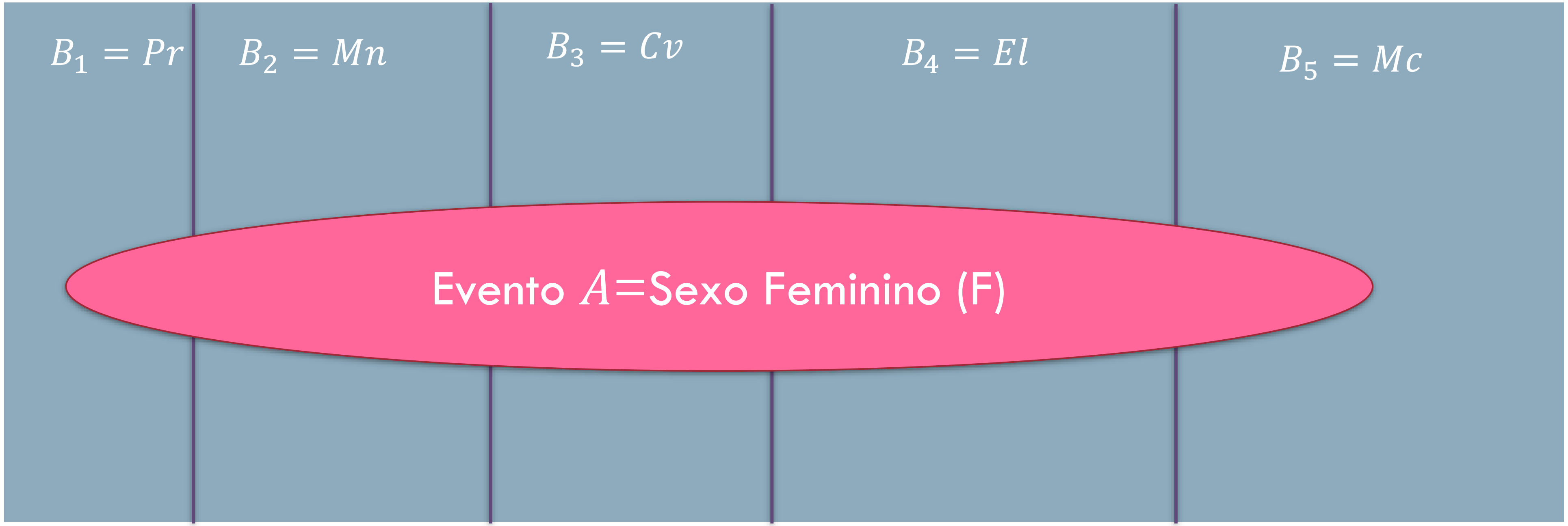


EXEMPLO

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?

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



PECE Programa de Educação Continuada
Escola Politécnica da USP

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???? $P(S_1|\,F) = P(S_1)\frac{P(F/S_1)}{P(F)}$

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.

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

sendo $i = 1, \cdots, k$ 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*

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

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

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?

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 uma 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

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

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

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

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

DETECÇÃO DE SPAM

O Mário 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.

PROBLEMA CLÁSSICO...FILTRO DE SPAM COM NAIVE BAYES

Queremos que uma IA identifique automaticamente se um e-mail é **spam** ou **não-spam** com base no seu conteúdo.

O algoritmo calcula a **probabilidade de um e-mail ser spam dado o conjunto de palavras que ele contém**. Isso é exatamente o que o **Teorema de Bayes** faz:

$$P(Spam|palavras) = P(spam)\frac{P(palavras|spam)}{P(palavras)}$$

$P(spam)$ probabilidade prévia de um e-mail ser spam (com base em e-mails anteriores).

$P(palavras|spam)$ probabilidade de ver essas palavras em e-mails classificados como spam.

$P(palavras)$ probabilidade de ver essas palavras em qualquer e-mail.

$P(Spam|palavras)$ probabilidade do e-mail ser spam dado o que ele contém.

$P(Spam) = 50\%$

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

O algoritmo aprende com uma base de dados de e-mails rotulados e, com isso, passa a classificar novos e-mails automaticamente.

BAYES E TESTE DE DIAGNÓSTICO MÉDICO

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

VPP (valor preditivo positivo) é a proporção de pessoas com teste positivo que realmente têm a doença,

$$VPP = P(Doente|Exame\ positivo)$$

VPN (valor preditivo negativo) é a proporção daqueles com resultado negativo que não têm a doença,

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

A **sensibilidade** de um teste refere-se à sua capacidade de identificar corretamente os casos positivos, medindo a proporção de indivíduos que realmente possuem a condição alvo e que são corretamente identificados como positivos pelo teste.

A **especificidade** é a taxa de verdadeiros negativos. É a capacidade de identificar corretamente as pessoas que não têm uma determinada doença. É uma medida que avalia a qualidade e confiabilidade dos resultados de exames.

	DOENTE	SAUDÁVEL	
TESTE SAUDÁVEL	Verdadeiro Positivo (VP)	Falso Positivo (FP)	$VPP = \dfrac{VP}{VP + FP}$
TESTE DOENTE	Falso Negativo (FN)	Verdadeiro Negativo (VN)	$VPN = \dfrac{VN}{VN + FN}$
	sensibilidade $\dfrac{VP}{VP + FN}$	especificidade $\dfrac{VN}{VN + FP}$	

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 2%, ou seja, apenas 2% 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.

$P(+|nDo) = 0{,}02$

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

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 bem menor que isso

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

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

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?

SOLUÇÃO USANDO BAYES

Supomos que $A(B, C)$ é o evento que diz que o prêmio está atrás da porta $1(2, 3)$.

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

Suponha – sem perda de generalização – que o jogador escolhe a **porta 1** e que o Monty abre a **porta 2** e todos veem que tem uma cabra!

$$P(abrir\ 2|A) = \frac{1}{2} \quad P(abrir\ 2|B) = 0 \quad P(abrir\ 2|C) = 1$$

$$P(abrir\ 2) = P(abrir\ 2|A)\ P(A) + P(abrir\ 2|B)\ P(B) + P(abrir\ 2|C)\ P(C)$$
$$= \frac{1}{2} \times \frac{1}{3} + 0 \times \frac{1}{3} + 1 \times \frac{1}{3} = \frac{1}{2}$$

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

Portanto,

$$P(C|abrir\ 2) = \frac{2}{3}$$

MAS ONDE APLICAREMOS NOSSO MODELO?

PECE Programa de Educação Continuada
Escola Politécnica da USP

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

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

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.

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.

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

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

$P(Y =$  $|Médio, Não\ ágil) = P(G)\frac{P(Médio\ e\ não\ ágil|G)}{P(Médio\ e\ não\ ágil)}$

$P(Y =$  $|Médio, Não\ ágil) = P(C)\frac{P(Médio\ e\ não\ ágil|C)}{P(Médio\ e\ não\ ágil)}$

$P(Y =$  $|Médio, Não\ ágil) = P(H)\frac{P(Médio\ e\ não\ ágil|H)}{P(Médio\ e\ não\ ágil)}$

ACABOU...

Reveja a aula antes
de fazer o teste.