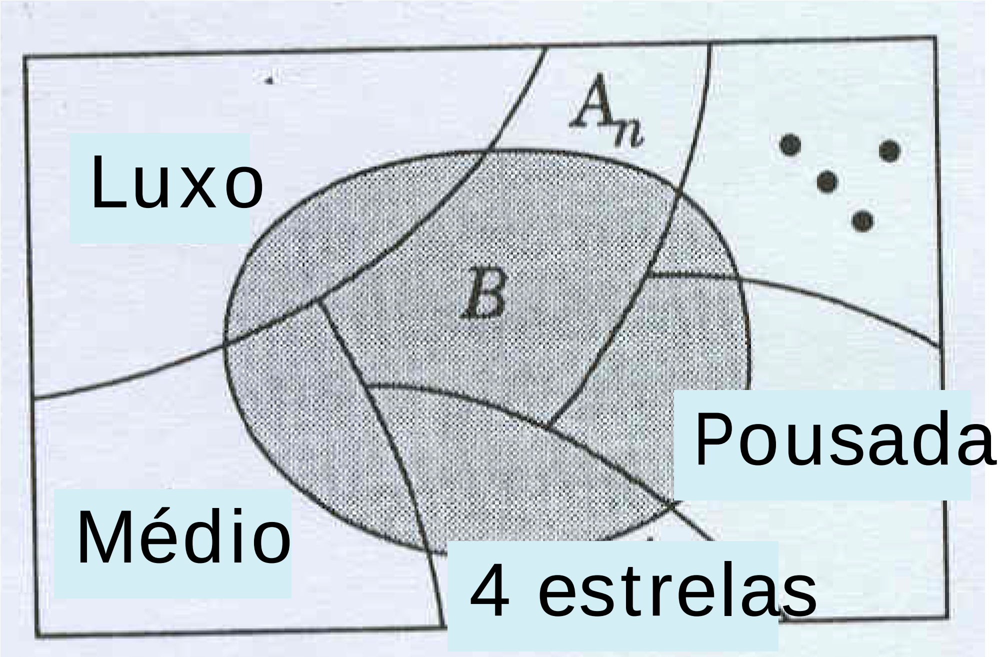
Teorema de Bayes

Anne Magály de Paula Canuto

Cenários mais complexos..

- Qual a probabilidade de um hotel apresentar uma taxa de ocupação num determinado mês entre 40 e 50%?
  - Esta pergunta abre respostas diversas, já que não delimitamos o tipo de hotel-luxo, médio, pousada, rural, praia,...
- Contudo, se perguntássemos: Qual a probabilidade de um hotel padrão luxo apresentar uma taxa de ocupação num determinado mês entre 40 e 50%?
- O espaço amostral está delimitado à análise de hotéis de luxo.

Um exemplo...



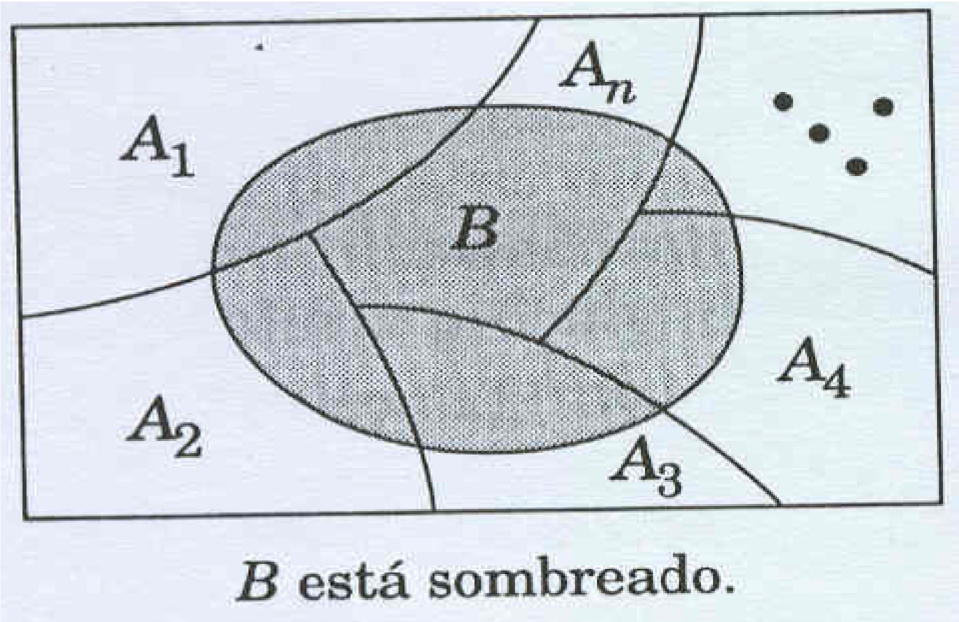
B está sombreado.

- B $\rightarrow$ Amostra de Hotéis com taxa de Ocupação ente 40 a 50%

Um exemplo

- Teste de Independência
  - Se $P(A|B) = P(A)$
    - Ocorrência do evento B não influi na probabilidade do evento A
  - Se $P(A|B) \neq P(A)$
    - Então A e B são eventos dependentes
- Ou seja:
  - $P(A \text{ e } B) = P(A) * P(B) \rightarrow$ independentes
  - $P(A \text{ e } B) \neq P(A) * P(B) \rightarrow$ dependentes

Partição do espaço amostral



(a) $A_i \cap A_j = \varnothing$, para todo $i \neq j$.

(b) $A_1 \cup A_2 \ldots \cup A_n = S$

(c) $P(A_i) > 0$, para todo i

- Neste caso:

$$\boxed{B = S \cap B = (A_1 \cup A_2 \cup \ldots A_n) \cap B = (A_1 \cap B) \cup (A_2 \cap B) \cup \ldots \cup (A_n \cap B)}$$

- Consequentemente:

$$\boxed{P(B) = P(A_1 \cap B) + P(A_2 \cap B) + \ldots + P(A_n \cap B)}$$

Teorema de Probabilidade Total

- Neste caso:

$$B=(\Omega\cap B)=(A_1,A_2,...A_n)\cap B=(B\cap A_1)\cup(B\cap A_2)\cup...(B\cap A_n)$$

- Consequentemente:

$$P(B)=P(\Omega\cap B)=P[(A_1,A_2,...A_n)\cap B]$$

$$=P[(B\cap A_1)\cup(B\cap A_2)\cup...(B\cap A_n)]$$

$$=P(B\cap A_1)+P(B\cap A_2)+...+P(B\cap A_n)$$

- Teorema de Probabilidade Total (TPT)

Uma vez particionado..

- De acordo com a regra de multiplicação e TPT, podemos escrever:

$$P(B) = P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2) + \ldots + P(A_n)P(B \mid A_n)$$

- Essa afirmação somente é possível quando os eventos são mutuamente exclusivos

Uma vez particionado..

- É importante deixar claro que $P(B|A_i)$ é diferente de $P(A_i|B)$
- Intuitivamente, $P(B|A_i)$ é a probabilidade de um hotel ter 40 a 50% de ocupação dado ser de luxo
  - Analisa a probabilidade de um hotel de luxo ter 40 a 50% de ocupação
- $P(A_i|B)$ define a probabilidade de um hotel ser de luxo dado que tenha 40 a 50% de ocupação
  - Analisa a probabilidade de um hotel ser de luxo, entre todos os hotéis que apresentam 40 a 50% de ocupação

Neste caso...

$$P(A_i|B) \neq P(B|A_i)$$

$$\boxed{P(A_1 \mid B) = \frac{P(A \cap B)}{P(B)}} \quad \text{e} \quad \boxed{P(B \mid A_1) = \frac{P(A \cap B)}{P(A_1)}}$$

$$\boxed{P(A \cap B) = P(B).P(A_1 \mid B)} \qquad \boxed{P(A \cap B) = P(A_1).P(B \mid A_1)}$$

$$\boxed{\text{Portanto:} \quad P(B).P(A_1 \mid B) = P(A_1).P(B \mid A_1)}$$

$$\boxed{P(A_1 \mid B) = \frac{P(A_1).P(B \mid A_1)}{P(B)}}$$

- Porém, muitas vezes não dispomos do P(B)
- O que fazer??

Neste caso..

- Mas sabemos que:

$$P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \ldots + P(A_n)P(B|A_n)$$
$$= \sum P(A_i)\,P(B|A_i)$$

- Assim, a equação fica:

$$\boxed{P(A_i|B) = \frac{P(A_i)\,P(B|A_i)}{\sum P(A_i)\,P(B|A_i)}}$$

- Esse é o teorema de Bayes

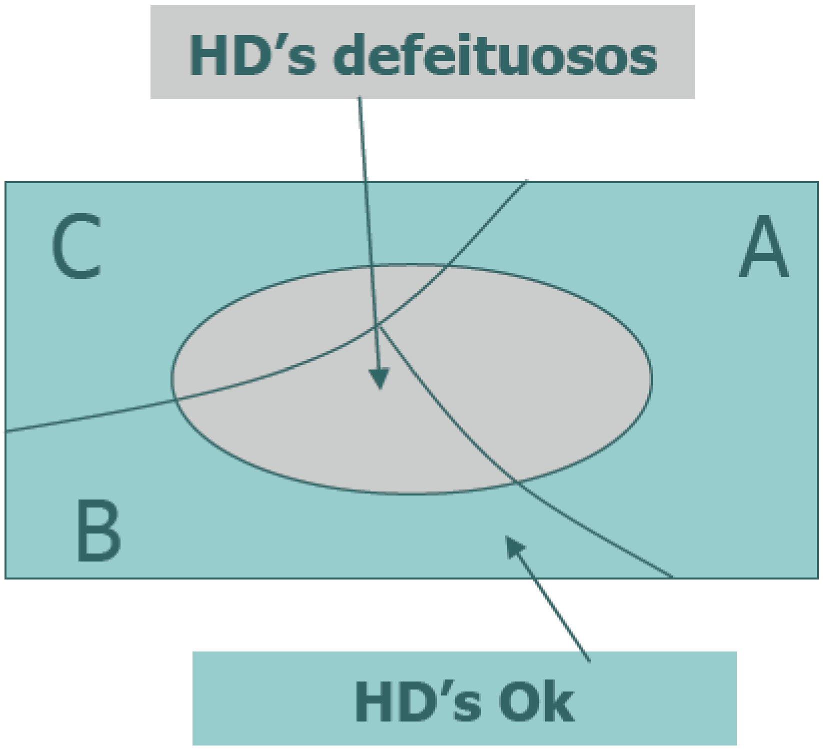
O teorema de Bayes é comumente utilizado para realizar inferência: Existe um certo numero de “causas” que podem resultar em um determinado “efeito”. Observamos o efeito e desejamos inferir a causa

Um exemplo..

- Um fabricante produz HDs em três fábricas (A, B e C), que respondem, respectivamente, por 40%, 35% e 25% de sua produção total.

- Registros históricos indicam que 2% da produção de A é defeituosa, assim como 1% da de B, e 3% da fábrica C.
  - Escolhemos 1 HD aleatoriamente, e ele é defeituoso.

- Qual a probabilidade dele ter sido produzido na fábrica B ?



Um exemplo

- Chamando *B o evento "fabricado em B" e d o evento "HD defeituoso", podemos escrever:*
- Uma peça defeituosa pode provir de qualquer uma das 3 fábricas (e só de uma!). Logo, eventos mutuamente excludentes.
- Portanto:

Um exemplo..

$$P(d) = P(A)P(d|A) + P(B)P(d|B) + P(C)P(d|C)$$

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

- Podemos calcular a P(d)

P(*d*) = (0,40 x 0,02) + (0,35 x 0,01) + (0,25 x 0,03) = 0,019

- Desta forma:

$$\boxed{P(B|d) = \frac{(0{,}35\,x\,0{,}01)}{(0{,}40\,x\,0{,}02) + (0{,}35\,x\,0{,}01) + (0{,}25\,x\,0{,}03)} = 0{,}184 = 18{,}4\%}$$

Um desafio para vocês..

- Desafio: O paradoxo do Monty-hall
  - Você está em um programa de TV que vai te dar um carro
  - Para tal, existem três portas (o carro está atrás de uma delas
  - Apenas uma das portas é premiada
  - O participante escolhe uma porta
  - O apresentador abre uma das portas não selecionadas
  - Você deve mudar de opção?

Um desafio para vocês..

Solução???

Um desafio para vocês..

- Critério de escolha do apresentador:
  1. Abre uma das portas aleatoriamente
  2. Sempre abre uma porta que sabe estar vazia
  3. Só abre a porta quando sabe que o participante escolheu a porta premiada
- Em cada um desses casos:
  1. A probabilidade de o prêmio encontrar-se atrás das outras portas se redistribui igualmente entre elas;
  2. A probabilidade de o prêmio estar atrás da porta escolhida não muda;
  3. A probabilidade de o participante ter escolhido a porta premiada vai a 100%.
- Vamos analisar o caso 2...

Um desafio para vocês..

- A resposta intuitiva ao problema: quando o apresentador revelou uma porta não-premiada, o concorrente teria à frente um novo dilema com apenas duas portas e um prêmio
  - As chances de que o prêmio esteja em qualquer uma das duas portas seriam de 50%.
    - Nossas chances subiram de 1/3 para ½
  - Assim, não faria diferença trocar ou não de porta pois ambas teriam as mesmas chances de possuírem o prêmio.

Um desafio para vocês..

- No entanto, esta resposta está errada, pois a porta que o apresentador abre depende da porta que o concorrente escolher inicialmente.
  - O apresentador sabe desde o começo onde está o prêmio (ele nunca abrirá uma porta premiada).
  - Ao abrir uma porta, ele não está criando um jogo todo novo, mas está dando informações valiosas sobre o jogo original.
- É por isso que a resposta é tão contra-intuitiva: parece-nos que o apresentador abriu uma porta aleatoriamente, mas isso está muito longe da verdade.

Um desafio para vocês..

- Vale a pena mudar.
  - Suas chances de ganhar são duas vezes maiores do que quando você permanece com sua escolha original.
  - Surpreso???
  - Vamos entender...

Um desafio para vocês..

- Você tem uma chance de 1/3 de escolher a porta que esconde o carro.
- Como o fato do apresentador escolher uma outra porta pode fazer alguma diferença?
  - O carro não mudou de posição.
- Se escolher a porta **A**, você terá uma chance de 1/3 de vencer,

Um desafio para vocês..

- A probabilidade de o carro estar atrás da porta **B** é de 1/3 e a probabilidade de o carro estar atrás da porta **C** também é de 1/3
- Neste caso, a probabilidade de o carro estar atrás da porta **B** ou da porta **C** é de 2/3.

- Vamos supor que o apresentador abra a porta **B** e descubra que ela está vazia.

Um desafio para vocês..

- A probabilidade de o carro estar atrás da porta **B** ou da porta **C** ainda é de 2/3,
  - Agora, a probabilidade de o carro estar atrás da porta **B** é igual 0 (o carro não está lá)

- Portanto, a probabilidade de o carro estar atrás da porta **C** é agora igual a 2/3.
- A soma das probabilidades ainda é igual a 1: 1/3 para **A**, 0 para **B**, 2/3 para **C**.

Exercício

- Um fabricante de sorvetes recebe 20% de todo leite de uma fazenda F1, 30% de uma fazenda F2 e 50% de F3. Um órgão de fiscalização inspecionou as fazendas de surpresa e descobriu que 20% do leite da fazenda F1 está adulterado, enquanto que para F2 e F3, essa proporção era de 5% e 2%, respectivamente.
  - Dado um galão de leite adulterado, qual a probabilidade de que este galão seja da fazenda F1?