

Teorema de Bayes

Profs. Alejandra, Cristian, Paula e Rafael

Introdução à Probabilidade e Estatística



- 1 Introdução
- 2 Teorema da probabilidade total
- 3 Teorema de Bayes

Estes slides tem por objetivo apenas complementar e esclarecer os textos principais indicados para a disciplina, e não devem ser usados como único texto de referência.

Eles tratam apenas dos principais aspectos da teoria abordada.

Junto à essas notas o aluno deve ler os textos da bibliografia.

Comecemos com um exemplo...

Em uma linha de produção de certa fábrica, determinada peça é produzida em duas máquinas. A máquina 1, mais antiga, é responsável por 35% da produção, e os 65% restantes vêm da máquina 2. A partir dos dados passados e das informações do fabricante das máquinas, estima-se em 5% a proporção de peças defeituosas produzidas pela máquina 1 e em 2,5% a proporção de peças defeituosas produzidas pela máquina 2. As peças produzidas pelas duas máquinas seguem para o departamento de armazenamento e embalagem para venda posterior, sem distinção de qual máquina a produziu.

Já no departamento de armazenamento, sorteamos uma peça aleatoriamente.

- 1 Qual é a proporção de peças defeituosas colocadas no mercado por essa fábrica?
- 2 Se um cliente identificar uma peça defeituosa, qual será a probabilidade de ela ter sido produzida pela máquina 1?

Vamos organizar os dados do problema.

Nosso experimento aleatório é o sorteio de uma peça produzida por essa fábrica, e com isso nosso espaço amostral é o conjunto de todas as peças produzidas.

Podemos ver que o espaço amostral está dividido em 2 eventos mutuamente exclusivos: $M_1 = \{\text{peças produzidas pela máquina 1}\}$, e $M_2 = \{\text{peças produzidas pela máquina 2}\}$, isto é,

$$S = M_1 \cup M_2 \text{ e } M_1 \cap M_2 = \emptyset.$$

Outro evento de interesse é $D = \{\text{a peça é defeituosa}\}$. Podemos ver que esse evento tem interseção com os eventos M_1 e M_2 ; ou seja, há peças defeituosas produzidas na máquina 1 e na máquina 2.

Por fim, considerando que cada peça tem a mesma probabilidade de ser sorteada, encontramos que a probabilidade de um evento qualquer é dado pela proporção das peças que estão naquele evento.

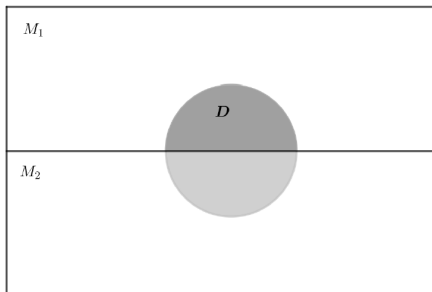


Figura: Espaço amostral e eventos M_1 , M_2 e D

O problema pergunta:

- ① Qual é a probabilidade de que uma peça aleatoriamente selecionada seja defeituosa? Ou seja, quanto é $P(D)$?
- ② Qual é a probabilidade, dado que selecionamos uma peça defeituosa, que ela tenha sido produzida pela maquina 1? Ou seja, quanto vale $P(M_1|D)$?

Já identificamos que queremos $P(D)$ e $P(M_1|D)$, mas o que o problema nos dá?

Vejamos

- 35% das peças vem da máquina 1 e 65% da máquina 2. Ou seja,

$$P(M_1) = 0,35 \quad \text{e} \quad P(M_2) = 0,65$$

- 5% das peças produzidas na máquina 1 são defeituosas. Isso pode ser escrito por

$$P(D|M_1) = 0,05$$

- 2,5% das peças produzidas na máquina 2 são defeituosas. Isso pode ser escrito por

$$P(D|M_2) = 0,025$$

Notem que, como $M_1 \cap M_2 = \emptyset$, podemos escrever o evento D como $D = (D \cup M_1) \cap (D \cup M_2)$. Logo,

$$P(D) = P(D \cap M_1) + P(D \cap M_2) = P(M_1)P(D|M_1) + P(M_2)P(D|M_1)$$

$$P(D) = 0,35 \times 0,05 + 0,65 \times 0,025 = 0,03375.$$

e

$$P(M_1|D) = \frac{P(D \cap M_1)}{P(D)} = \frac{P(M_1)P(D|M_1)}{P(M_1)P(D|M_1) + P(M_2)P(D|M_1)}$$

$$P(M_1|D) = \frac{0,35 \times 0,05}{0,03375} = 0,5185.$$

O problema que acabamos de estudar estava estruturado da seguinte forma.

- 1 O espaço estava completamente descrito a partir eventos mutuamente exclusivos. Neste caso, os eventos M_1 e M_2 , dos quais conhecíamos as probabilidades;
- 2 um novo evento é introduzido no problema. Mais exatamente o evento D ;
- 3 conhecemos as probabilidade de D apenas quando condicionado aos eventos da primeira parte. Ou seja, conhecemos $P(D|M_1)$ e $P(D|M_2)$;
- 4 o problema queria saber a probabilidade deste novo evento, e como esta informação alterava as probabilidades dos eventos usados para descrever o espaço. Ou seja, queremos $P(D)$ e $P(M_1|D)$.

Esta é uma estrutura bastante particular e muito presente em problemas estatísticos. Aparece, por exemplo, quando queremos medir as taxas de falso positivo e falso negativo em exames médicos.

Duas propriedades essenciais dos eventos M_1 e M_2 , sem as quais a solução teria sido muito mais complicado (e talvez impossível), eram:

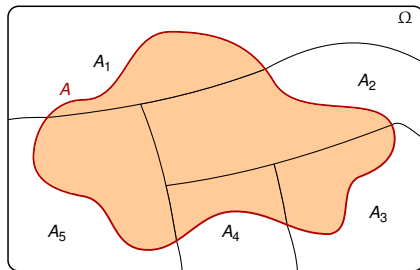
- $M_1 \cap M_2 = \emptyset$ (mutuamente exclusivos)
- $M_1 \cup M_2 = S$

Eventos (podem ser mais de 2) com essas propriedades são chamados de partição, e são centrais nessa parte da teoria.

Começemos estudando-os...

A sequência de eventos $\{A_i\}_{i \in \mathbb{N}}$ é uma partição de S se

$$A_i \cap A_j = \emptyset \text{ se } i \neq j \text{ (são mutuamente exclusivos)} \quad \text{e} \quad \bigcup_{i=1}^{\infty} A_i = S.$$



Note que se $\{A_i\}_{i \in \mathbb{N}}$ é uma partição de S , então para qualquer evento $A \subset S$, vale que

$$A = \bigcup_{i=1}^{\infty} \{A \cap A_i\} = \{A \cap A_1\} \cup \{A \cap A_2\} \cup \dots$$

Teorema

Seja $\{A_i\}_{i \in \mathbb{N}}$ uma partição de S e A um evento qualquer, então

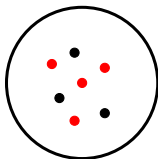
$$P(A) = \sum_{i=1}^{\infty} P(A|A_i)P(A_i) = P(A|A_1)P(A_1) + P(A|A_2)P(A_2) + \dots$$

Prova: Como $\{A_i\}_{i \in \mathbb{N}}$ é uma partição de S , então $A = \bigcup_{i=1}^{\infty} \{A \cap A_i\}$ e a união é sobre eventos mutuamente excludentes. Logo,

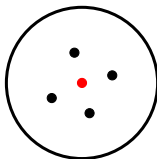
$$P(A) = P\left(\bigcup_{i=1}^{\infty} \{A \cap A_i\}\right) = \sum_{i=1}^{\infty} P(A \cap A_i) = \sum_{i=1}^{\infty} P(A|A_i)P(A_i).$$

Exemplo

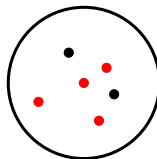
Considere as seguintes urnas contendo bolas vermelhas e pretas:



Urna 1



Urna 2



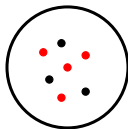
Urna 3

Escolha ao acaso uma urna e, da urna escolhida, retire ao acaso uma bola observando sua cor. Calcule a probabilidade de que seja retirada uma bola vermelha.

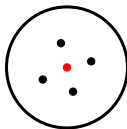
Considere os eventos

$U_i = \{ \text{a urna } i \text{ é escolhida} \}, \quad i \in \{1,2,3\}$

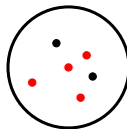
$V = \{ \text{a bola retirada é vermelha} \}.$



Urna 1



Urna 2



Urna 3

Como U_1, U_2, U_3 são uma partição de S temos que:

$$P(V) = \underbrace{P(V|U_1)}_{=4/7} \overbrace{P(U_1)}^{=1/3} + \underbrace{P(V|U_2)}_{=1/5} \overbrace{P(U_2)}^{=1/3} + \underbrace{P(V|U_3)}_{=2/3} \overbrace{P(U_3)}^{=1/3} \approx 0,48.$$

Teorema

Se A_i , $i \in \mathbb{N}$ é uma sequência de eventos que forma uma partição de S , tal que $P(A_i) > 0$ para todo i , então para todo evento A vale que

$$P(A_i|A) = \frac{P(A|A_i)P(A_i)}{\sum_{i=1}^{\infty} P(A|A_i)P(A_i)},$$

para todo $i \in \mathbb{N}$.

Ideia da prova: Note que:

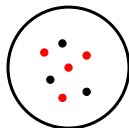
$$P(A_i|A) = \frac{P(A \cap A_i)}{P(A)}$$

regra da multiplicação!

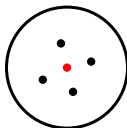
T. da Probabilidade Total!

Voltando ao exemplo anterior...

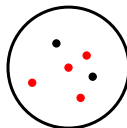
Considere as seguintes urnas contendo bolas vermelhas e pretas, escolha ao acaso uma urna e, da urna escolhida, retire ao acaso uma bola observando sua cor.



Urna 1



Urna 2



Urna 3

Dado que uma bola vermelha é retirada, calcule a probabilidade de que a Urna 2 tenha sido a escolhida.

Solução: Queremos calcular $P(U_2|V)$. Lembrando que já calculamos que $P(V) = 0,48$. *Agora, usando o T. de Bayes:*

$$P(U_2|V) = \frac{P(V|U_2)P(U_2)}{P(V)} = \frac{(1/5) \cdot (1/3)}{0,48} \approx 0,1388.$$

Uma gravidez ectópica (ou extra-uterina) é duas vezes mais provável em mulheres fumantes do que em mulheres não fumantes. Se 32% das mulheres em idade reprodutiva são fumantes, que percentual de mulheres com gravidez ectópica são fumantes?

Vamos considerar o experimento aleatório de sortear uma mulher ao acaso na população. Definamos então os eventos:

$$E = \{\text{A mulher teve gravidez ectópica}\}$$

$$F = \{\text{A mulher é fumante}\}$$

O que queremos determinar corresponde a probabilidade $P(F|E)$.

Do problema, temos que $P(E|F) = 2P(E|F^c)$ e que $P(F) = 0,32$.

Vejam também que os conjuntos F e F^c formam uma partição do espaço amostral (por que?).

Aplicando o teorema de Bayes,

$$\begin{aligned}P(F|E) &= \frac{P(F \cap E)}{P(E)} \\&= \frac{P(F)P(E|F)}{P(F)P(E|F) + P(F^c)P(E|F^c)} \\&= \frac{P(F)P(E|F)}{P(F)P(E|F) + 0,5P(F^c)P(E|F)}, \quad \text{pois } P(E|F) = 2P(E|F^c) \\&= \frac{P(F)}{P(F) + 0,5P(F^c)} \\&= \frac{0,32}{0,32 + (0,5 \times 0,68)} \\&= 0,48.\end{aligned}$$

Suponha que um teste clínico (exame de sangue, ultrassom, tomografia computadorizada, etc.) seja utilizado para o diagnóstico de determinada doença. Vamos analisar as diferentes probabilidades envolvidas neste teste diagnóstico, apresentando alguns conceitos comuns em estudos epidemiológicos.

Em um exame deste tipo temos dois fatores a serem considerados: o estado clínico do paciente e o resultado do teste. O paciente pode ter a doença ($D+$) ou não ($D-$); já o teste pode dar positivo ($T+$), isto é, indicar que o paciente tem a doença, ou negativo ($T-$) indicando que o paciente não tem a doença.

Na análise da qualidade de um teste diagnóstico, há duas características desejáveis: o teste detectar a doença quando o paciente de fato tem a doença e o teste não detectar a doença quando o paciente não tem a doença. As probabilidades associadas recebem nomes especiais.

- Sensibilidade: $P(T+|D+)$.
- Especificidade: $P(T-|D-)$.

As situações inversas, que indicam erros no exame, onde este indica a presença da doença em alguém que não a possui, ou que o teste dá negativo para pessoas com a doença, são conhecidos respectivamente como **falso positivo** e o **falso negativo**, e são definidos por:

- Probabilidade de um falso negativo: $P(T-|D+) = 1 - P(T+|D+)$.
- Probabilidade de um falso positivo: $P(T+|D-) = 1 - P(T-|D-)$.

Sensibilidade, especificidade, falso positivo e falso negativo são medidas de qualidade do teste, uma vez que medem se o teste está indicando corretamente ou não o estado de saúde da pessoa.

Mas essas não são as medidas que um médico precisa ter em mente quando lê um exame. No dia a dia, existem outras duas probabilidades que orientariam melhor as decisões médicas, que envolvem saber se o paciente tem ou não a doença, uma vez que sabemos o resultado do teste. São elas:

- Valor preditivo positivo: $P(D+|T+)$.
- Valor preditivo negativo: $P(D-|T-)$.

O teorema de Bayes é o responsável por conectar estes conceitos: conhecendo-se a probabilidade de prevalência da doença, isto é, $P(D+)$, podemos aplicar o teorema de Bayes para obter os valores preditivos a partir da sensibilidade e da especificidade do teste.

Suponha que um laboratório informe que a sensibilidade de um teste para detectar determinada doença é 0,90, enquanto sua especificidade é 0,95. Um paciente é retirado aleatoriamente de uma população em que a taxa de prevalência da doença é de 1 em 1000 pessoas.

- ❶ Qual é a probabilidade de o teste dar o resultado correto?

Solução: O problema começa nos informando a sensibilidade e a especificidade do teste, como colocado abaixo.

- sensibilidade: $P(T+|D+) = 0,9$,
- especificidade: $P(T-|D-) = 0,95$.

Além disso, o problema nos dá também a prevalência da doença. Ou seja, sabemos que $P(D+) = 0,001$.

Trata-se então de uma doença que aflige apenas 0,1% da população.

O problema pede agora para calcular a probabilidade do teste dar o resultado correto. Ou seja a probabilidade do teste dar positivo e o paciente estar doente, ou do teste dar negativo e o paciente não possuir a doença. Para calcular isso, definamos o evento:

$$C = (D+ \cap T+) \cup (D- \cap T-).$$

Notem que os eventos $(D+ \cap T+)$ e $(D- \cap T-)$ são mutuamente exclusivos, e portanto temos que

$$\begin{aligned} P(C) &= P(D+ \cap T+) + P(D- \cap T-) \\ &= P(D+)P(T+|D+) + P(D-)P(T-|D-) \end{aligned}$$

A quantidade em vermelho ainda não é conhecida! Mas, sabemos que $P(D-) = 1 - P(D+) = 0,999$ (porque?). Deste modo, nossa probabilidade fica

$$P(C) = 0,001 \times 0,9 + 0,999 \times 0,95 = 0,9499.$$

2 Calcule o valor preditivo positivo.

Este valor preditivo positivo é $P(D+|T+)$. Pelo teorema de Bayes, temos que

$$P(D+|T+) = \frac{P(D+)P(T+|D+)}{P(D+)P(T+|D+) + P(D-)P(T+|D-)}$$

Ainda não conhecemos a quantidade em vermelho, e portanto devemos calculá-la. Para isso note que

$$P(T+|D-) = 1 - P(T-|D-) = 1 - 0,95 = 0,05.$$

Assim,

$$\begin{aligned} P(D+|T+) &= \frac{0,001 \times 0,9}{(0,001 \times 0,9) + (0,999 \times 0,05)} \\ &= 0,0177 \end{aligned}$$

Este resultado de apenas 1,7% ocorreu por que a prevalência da doença é muito baixa, e a sensibilidade do teste, apesar de boa, não era alta o suficiente para compensar isso.

Exercício

Jogamos um dado honesto e em seguida lançamos tantas moedas honestas como os pontos indicados no dado.

- 1 Qual é a probabilidade de obter cinco caras?
- 2 Dado que foram obtidas cinco caras, qual é a probabilidade de que o dado tenha mostrado um seis?