# Estatística e Probabilidade

## Probabilidade condicionada e Teorema de Bayes

**Desenvolvimento do material**

Gregório Dalle Vedove Nosaki

**1ª Edição**



# Sumário

## Probabilidade condicionada e Teorema de Bayes

## Para início de conversa...

No último capítulo deste livro, vamos abordar as probabilidades condicionadas ou probabilidades condicionais. Trata-se de variações da definição de probabilidade clássica apresentada no capítulo anterior quando temos mais informações sobre as possibilidades de ocorrência de um evento. Iremos considerar que um dado evento B ocorre e, a partir daí, calcular a probabilidade de outro evento A. Veremos como essa condição inicial pode alterar o valor da probabilidade do evento A. Traremos diversos exemplos utilizando experimentos probabilísticos clássicos para ilustrar as definições e as propriedades que serão abordadas neste capítulo. Finalizaremos com o Teorema de Bayes, também chamado de Regra de Bayes, que pode auxiliar no cálculo e na compreensão das probabilidades condicionais e suas relações. Com as probabilidades condicionais começamos a explorar um pouco mais a fundo essa área da Matemática, que possui diferentes aplicações e tornou-se um ramo próprio de estudo, tendo como ponto de partida questões bastante simples, mas que deram origem a uma teoria muito rica e complexa.

## Objetivos

- Aplicar as propriedades operatórias da Teoria das Probabilidades em casos práticos.
- Calcular probabilidades aplicando as regras da adição e da multiplicação.

# 1. Probabilidade Condicionada

As probabilidades condicionadas ou probabilidades condicionais também partem de um experimento probabilístico com um determinado espaço amostral, mas quando calculamos a chance de um determinado evento ocorrer, podemos ter informações adicionais que podem alterar as chances de que os resultados do nosso evento inicial sejam obtidos. Nesse cenário, utilizamos a probabilidade condicional. ROSS (2010) destaca alguns aspectos das aplicações e características das probabilidades condicionais:

> “Em primeiro lugar, estamos frequentemente interessados em calcular probabilidades quando temos alguma informação parcial a respeito do resultado de um experimento; em tal situação, as probabilidades desejadas são condicionais. Em segundo lugar, mesmo quando não temos nenhuma informação parcial sobre o resultado de um experimento, as probabilidades condicionais podem ser frequentemente utilizadas para computar mais facilmente as probabilidades desejadas. (ROSS, 2010, p. 81).”

Apresentaremos a definição formal de probabilidade condicional e, em seguida, um exemplo simples da aplicação desta definição.

**Definição:** sejam $A, B \subset S$ eventos do espaço amostral $S$ tal que $P(B) \neq 0$. A **probabilidade condicional** de $A$ dado $B$ é denotada por $P(A|B)$ e vale a seguinte fórmula

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

A probabilidade condicional $P(A|B)$ pode ser interpretada como a probabilidade de o evento $A$ ocorrer caso já tenhamos de antemão que o evento $B$ ocorreu.

Vamos considerar o experimento do dado de seis faces. O espaço amostral neste caso é o conjunto

$$S = \{1, 2, 3, 4, 5, 6\}.$$

Se considerarmos o evento $A$ sendo obter a face 2 voltada para cima e o evento $B$ como sendo obter uma face par, então temos que

$$A = \{2\} \text{ e } B = \{2, 4, 6\}.$$

Claro que

$$A \cap B = \{2\}$$

e, pela fórmula do cálculo de probabilidade em casos onde o espaço amostral é finito e todos os resultados têm a mesma chance de serem obtidos, temos que

$$P(B) = \frac{3}{6} = \frac{1}{2} \text{ e } P(A \cup B) = \frac{1}{6}.$$

Dessa forma, a probabilidade de obtermos a face 2 caso já saibamos de antemão que uma face par foi obtida é dada pela probabilidade condicional $P(A|B)$ cujo valor é

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

### Importante

Em algumas bibliografias, você poderá encontrar a notação $AB$ para denotar a interseção $A \cap B$. Isso ocorre pois a interseção $A \cap B$ significa, em probabilidade, que estamos considerando os resultados que pertencem tanto ao evento $A$ quanto ao evento $B$ simultaneamente. Essa operação lógica é do tipo multiplicativa, ou seja, estamos considerando os resultados que pertencem a $A$ **e** a $B$. Utilizaremos a notação $A \cap B$ para evitar qualquer confusão e nos aproximarmos mais da Teoria de Conjuntos, mas em algumas referências a notação multiplicativa dos eventos pode ser encontrada.

**Exemplo 1:** considere o lançamento de dois dados honestos de seis faces. Denotando por $S$ o espaço amostral deste experimento, considere os seguintes eventos $A, B, C \subset S$ descritos como sendo:

- $A$ é evento onde um dos dois dados tem a face 2 voltada para cima;
- $B$ é o evento cuja soma das faces voltadas para cima é maior ou igual a 8;
- $C$ é o evento tal que o produto das duas faces é um número ímpar.

A partir das definições dos eventos acima, podemos descrever os conjuntos que representam os eventos $A$, $B$ e $C$ como sendo

$$A = \{(x, y) \in S : x = 2 \text{ ou } y = 2\}$$

$$B = \{(x, y) \in S : x + y \geq 8\}$$

$$C = \{(x, y) \in S : x \cdot y \text{ é ímpar }\}.$$

Outra forma de descrever os eventos apresentados anteriormente é listar todos os resultados possíveis que pertencem a cada um dos eventos. Para isso, podemos utilizar uma tabela que apresenta todo o espaço amostral e assim determinarmos quais resultados pertencem a cada evento. Por exemplo, para o evento $A = \{(x, y) \in S : x = 2 \text{ ou } y = 2\}$ temos que os pares ordenados destacados abaixo pertencem ao conjunto $A$:

| | | | | | |
|---|---|---|---|---|---|
| (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) |

Já para o evento $B = \{(x,y) \in S : x + y \geq 8\}$, os pares que pertencem a esse evento são destacados na tabela abaixo:

| | | | | | |
|---|---|---|---|---|---|
| (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)** |

Fazendo o mesmo processo para o evento $C = \{(x,y) \in S : x \cdot y \text{ é ímpar } \}$, obtemos:

| | | | | | |
|---|---|---|---|---|---|
| **(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) |

Vamos agora calcular algumas probabilidades condicionais baseadas nestes eventos. A primeira probabilidade será a probabilidade do evento $A$ dado a ocorrência do evento $B$. O evento $A \cap B$ contém só dois resultados possíveis,

$$A \cap B = \{(2,6);(6,2)\}$$

pois trata-se do evento no qual pelo menos um dos dois dados lançados tem a face 2 voltada para cima e a soma das duas faces deve ser maior ou igual a 8. Todas as outras combinações com uma das faces dos dados iguais a 2 têm a soma menor que 8 e, portanto, os dois únicos resultados possíveis para o evento $A \cap B$ são os apresentados acima.

A partir dessa análise, vamos calcular a probabilidade condicional de $A$ dado $B$. Analisando a cardinalidade do conjunto $B$ e do conjunto $A \cap B$, além de todo o espaço amostral $S$, obtemos que

$$P(A \cap B) = \frac{2}{36} \text{ e } P(B) = \frac{15}{36}$$

portanto

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{2}{36}}{\frac{15}{36}} = \frac{2}{15}.$$

A probabilidade condicional de $B$ dado $A$, no entanto, é igual a

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{2}{36}}{\frac{11}{36}} = \frac{2}{11}.$$

Para calcularmos as probabilidades condicionais $P(B|C)$ e $P(C|B)$, devemos determinar os resultados possíveis do evento $B \cap C$. Analisando as tabelas apresentadas anteriormente, obtemos que

$$B \cap C = \{(3,3);(3,5);(5,3);(5,5)\}.$$

Dessa forma, temos que

$$P(B|C) = \frac{P(B \cap C)}{P(C)} = \frac{\frac{4}{36}}{\frac{9}{36}} = \frac{4}{9}$$

e

$$P(C|B) = \frac{P(B \cap C)}{P(B)} = \frac{\frac{4}{36}}{\frac{15}{36}} = \frac{4}{15}.$$

## Importante

Quando temos dois eventos independentes, então a probabilidade condicional entre eles será sempre igual a zero. Tome, por exemplo, os eventos $A$ e $C$ do Exemplo 1. Como $A \cap C = \varnothing$, vale que $P(A \cap C) = P(\varnothing) = 0$. Portanto, temos que $P(A|C) = 0 = P(C|A)$.

**Proposição:** considere $A, B \subset S$ eventos de um espaço amostral $S$. Valem as seguintes regras

1. $P(A \cup B) + P(A \cap B) = P(A) + P(B)$;
2. $P(A \cap B) = P(B) \cdot P(A|B)$ desde que $P(B) \neq 0$.

Os itens 1 e 2 são popularmente conhecidos pelos nomes de **regra da adição** e **regra da multiplicação**, respectivamente.

Vejamos a aplicação dessas regras em um exemplo utilizando um baralho de jogos de cartas.

**Exemplo 2:** considere um baralho comum com 52 cartas divididas em 4 naipes com 13 cartas cada. Dentro de cada naipe, as cartas são numeradas de 1 a 10, sendo que a carta na posição do número 1 é representada por um A e é chamada de 'ás'. Além disso, cada naipe possui três cartas com figuras: o valete (J), a rainha (Q) e o rei (K).

Consideremos ainda que todas as cartas têm a mesma chance de serem sorteadas em um sorteio aleatório. Para este exemplo, considere o evento $A$ como sendo sortear uma carta do naipe de copas representado pelo símbolo ♥ e o evento $B$ como sendo sortear uma carta que seja representada por uma letra, ou seja, um 'ás', um valete, uma rainha ou um rei, independentemente do naipe.

As probabilidades dos eventos $A$ e $B$ podem ser descritas como

$$P(A) = \frac{13}{52} = \frac{1}{4} \text{ e } P(B) = \frac{16}{52} = \frac{4}{13}.$$

O evento $A \cap B$ é descrito pelo conjunto formado pelos resultados como indicado a seguir

$$A \cap B = \{ A♥, J♥, Q♥, K♥\},$$

logo, a probabilidade do evento $A \cap B$ pode ser calculada como sendo

$$P(A \cap B) = \frac{4}{52} = \frac{1}{13}.$$

Pela definição de probabilidade condicional, temos que

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{13}}{\frac{4}{13}} = \frac{1}{4}.$$

A regra da multiplicação pode ser vista como uma variação da própria fórmula, que determina o valor da probabilidade condicional $P(A|B)$, pois

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \Rightarrow P(A \cap B) = P(B) \cdot P(A|B).$$

Para determinar o valor de $P(A \cup B)$ podemos utilizar a regra da adição. Dessa forma, temos que

$$P(A \cup B) + P(A \cap B) = P(A) + P(B) \Rightarrow P(A \cup B) + \frac{1}{13} = \frac{1}{4} + \frac{4}{13}$$

e, portanto,

$$P(A \cup B) = \frac{25}{52}.$$

A regra da adição também pode ser vista como consequência de uma das propriedades apresentadas no capítulo anterior.

A regra da multiplicação citada anteriormente pode ser aplicada invertendo a ordem dos eventos, desde que não haja inconsistência nas definições. Se $P(A) \neq 0$ e $P(B) \neq 0$, então valem as igualdades

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

e

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

Além disso, podemos generalizar a regra da multiplicação para um número finito qualquer de interseções de conjuntos. Sejam $A_1, A_2, ..., A_n$

eventos de um mesmo espaço amostral tais que $P(A_i) \neq 0$ para todo $i$. Vale a seguinte igualdade

$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdot \ldots \cdot P(A_n|A_1 \cap \ldots \cap A_{n-1})$.

Confira ROSS (2010) para ver exemplos comentados da aplicação da regra da multiplicação generalizada para um número finito de eventos.

## 2. Teorema de Bayes

Nesta seção, veremos mais uma aplicação das probabilidades condicionais conhecida como Teorema de Bayes ou Regra de Bayes.

**Teorema de Bayes:** sejam $A, B \subset S$ eventos de um espaço amostral $S$. Se $P(B) \neq 0$, vale que

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

### Curiosidade

Thomas Bayes foi um reverendo presbiteriano de origem inglesa que viveu no início do século XVIII. Ele possui apenas uma publicação voltada para os estudos matemáticos cujo nome é "A doutrina dos fluxions" ("The doctrine of fluxions"). No entanto, sua obra mais relevante no contexto acadêmico foi encontrada entre seus pertences após sua morte. Em um dos achados, havia o rascunho de um artigo que mais tarde foi publicado e, posteriormente, trazido ao conhecimento do mundo todo pelo matemático Laplace. O ramo da inferência bayesiana recebe esse nome em homenagem a esse matemático e sua contribuição para a Matemática, sendo esse um ramo com diferentes aplicações, que variam de inteligência artificial à mineração de dados. (PENA, 2009).

Vamos analisar uma aplicação do Teorema de Bayes em um exemplo.

**Exemplo 3:** considere uma caixa contendo 100 bolinhas numeradas de 1 a 100 e que todas elas têm a mesma chance de serem sorteadas em uma tiragem aleatória. Além disso, considere $A$ como sendo o evento de sortear uma bolinha cujo valor seja maior que 60 e $B$ o evento de sortear uma bolinha onde o algarismo 7 aparece, independentemente da posição que ele ocupa. Dessa forma, temos que

$$A = \{61, 61, 63, 64, ..., 99, 100\}$$

e

$$B = \{7, 17, 27, 37, 47, 57, 67, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 87, 97\}.$$

Calculando as probabilidades $P(A)$ e $P(B)$ obtemos que

$$P(A) = \frac{40}{100} \text{ e } P(B) = \frac{19}{100}.$$

Além disso, o evento $A \cap B$ é

$$A \cap B = \{67, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 87, 97\}.$$

Vamos determinar a probabilidade condicional $P(B|A)$, isto é, a probabilidade da ocorrência do evento $B$ dado a ocorrência do evento $A$. Pela definição de probabilidade condicional, temos que

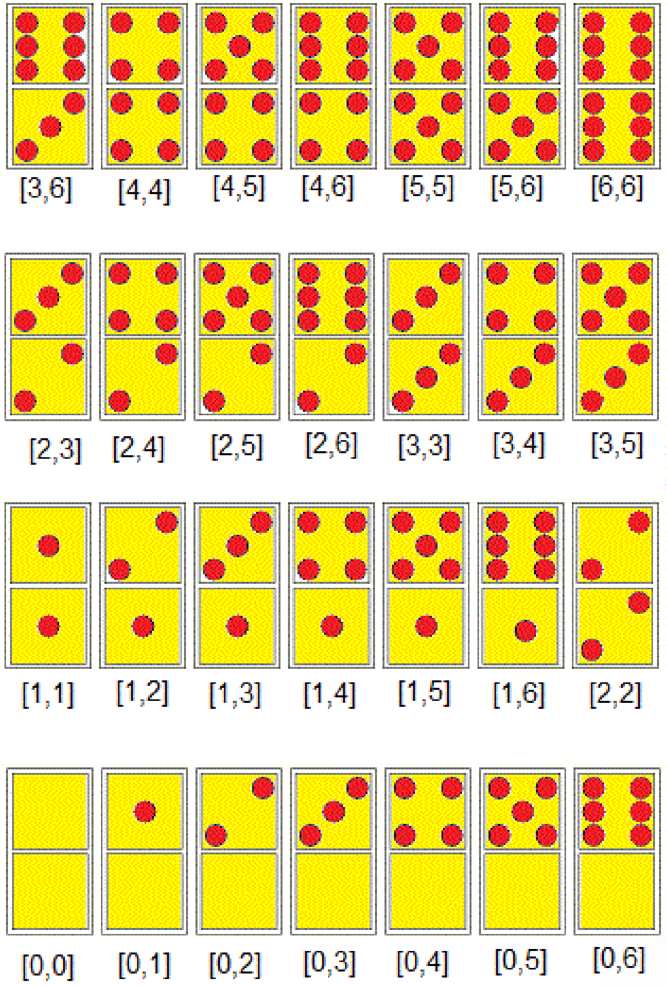
$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{13}{100}}{\frac{19}{100}} = \frac{13}{19}.$$

Utilizando o Teorema de Bayes, obtemos

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} = \frac{\frac{13}{19} \cdot \frac{19}{100}}{\frac{40}{100}} = \frac{13}{40}.$$

No exemplo anterior, usamos tanto a definição de probabilidade condicional quanto o Teorema de Bayes para determinar os valores de probabilidades condicionais. Além disso, é importante compreender o que cada probabilidade condicional significa dentro do experimento probabilístico. A probabilidade $P(A|B)$ significa a probabilidade de obtermos uma bolinha cuja numeração seja maior que 60 dado que foi sorteada uma bolinha que apresenta o algarismo 7 em alguma posição. De modo análogo, a probabilidade $P(B|A)$ determina a razão das chances de sortear uma bolinha onde o algarismo 7 apareça sabendo que sorteamos uma bolinha cujo valor é maior que 60.

**Exemplo 4:** considere um jogo de dominó com 28 peças, em que cada peça possui duas faces ou dois lados. Em cada lado, há um número que varia de 0 a 6, como ilustrado na figura abaixo.




Vamos utilizar a notação $[n, m]$ para a peça que apresenta os lados com os valores $n$ e $m$. Além disso, é importante notar que a peça $[n, m]$ pode ser representada também como $[m, n]$, pois não se tratam de pares ordenados. A notação $[n, m]$ indica que em um dos lados da peça há o valor $n$ e em outro o valor $m$ independentemente da posição que esses valores ocupam.

Seja $A$ o evento tal que a soma dos números inscritos dos dois lados de uma mesma peça seja maior ou igual a 7 e $B$ o evento de que um dos lados da peça apresente o valor 5. Os eventos $A$ e $B$ são descritos pelos conjuntos:

$A = \{[1,6]; [2,5]; [2,6]; [3,4]; [3,5]; [3,6]; [4,4]; [4,5]; [4,6]; [5,5]; [5,6]; [6,6]\}$

$B = \{[0,5]; [1,5]; [2,5]; [3,5]; [4,5]; [5,5]; [6,5]\}$.

Assim, vale que

$$P(A) = \frac{12}{28} = \frac{3}{7} \text{ e } P(B) = \frac{7}{28} = \frac{1}{4}.$$

A probabilidade condicional $P(A|B)$ pode ser calculada como sendo

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{5}{28}}{\frac{1}{4}} = \frac{5}{7}.$$

Portanto, pelo Teorema de Bayes, podemos afirmar que

$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)} = \frac{\frac{5}{7} \cdot \frac{1}{4}}{\frac{3}{7}} = \frac{5}{12}.$$

De fato, pela definição de probabilidade condicional podemos confirmar que

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{5}{28}}{\frac{3}{7}} = \frac{5}{12}.$$

Neste capítulo, trabalhamos com a definição de probabilidade condicional e vimos como a ocorrência de um evento pode influenciar a probabilidade de ocorrência de outro evento. Trata-se de uma definição bastante intuitiva que possui diversas aplicações, mas requer um certo cuidado ao manipular tais conceitos e propriedades. Vimos como trabalhar com a definição clássica de probabilidade condicional e também a Regra de Bayes que pode auxiliar na interpretação e até mesmo no cálculo das probabilidades condicionais. Apresentamos diversos exemplos com diferentes experimentos probabilísticos para ilustrar as definições e estruturas que estão envolvidas ao trabalharmos com as probabilidades condicionais.

# Referências


BONAFINI, F. C. (org.). **Matemática**. São Paulo: Pearson Prentice Hall, 2012.

IEZZI, G.; MURAKAMI, C. **Fundamentos de Matemática Elementar**, v. 1: Conjuntos e Funções. São Paulo: Atual, 1977.

LOESCH, C. **Probabilidade e estatística**. Rio de Janeiro: Grupo Gen, 2012.

NAVIDI, W. **Probabilidade e estatística para ciências exatas**. Porto Alegre: Grupo A - AMGH, 2012.

PENA, S. D. Thomas Bayes: o cara! **Ciência Hoje**, 2009. Disponível em: https://cienciahoje.org.br/artigo/thomas-bayes-o-cara/. Acesso em: 13 set. 2021.

ROSS, S. **Probabilidade**: um curso moderno com aplicações. Porto Alegre: Bookman, 2010.

SPIEGEL, M; SCHILLIER, J; SRINIVASAN, A. **Probabilidade e estatística**. 3. ed. Porto Alegre: Bookman, 2013.