ME951 - Estatística e Probabilidade I

Parte 7

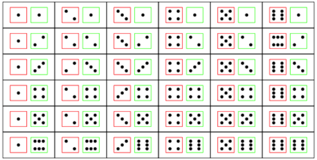
Notas de aula de ME414 produzidas pelos professores **Samara Kiihl**, **Tatiana Benaglia** e **Benilton Carvalho** modificadas e alteradas pela Profa. **Larissa Avila Matos**

Probabilidade Condicional e Independência

Probabilidade condicional

Probabilidade Condicional: encontrar a probabilidade de um evento quando você tem alguma outra informação sobre o evento.

- Considere o lançamento de dois dados. Espaço amostral:



- Considere que cada resultado tenha a mesma chance de ocorrer: 1/36.
- Suponha que você lance primeiro um dos dados e o resultado é 4.
- Qual a probabilidade de que a soma dos resultados dos dois dados seja 10?

Probabilidade condicional

- Como saiu 4 no primeiro dado, há 6 resultados possíveis:

$$\Omega_1 = \{(4,1),(4,2),(4,3),(4,4),(4,5),(4,6)\}$$

- Cada resultado tem a mesma probabilidade de ocorrer: $1/6$.

- Dado que o primeiro dado teve resultado 4, então a probabilidade de cada evento em Ω_1 tem igual chance de ocorrer.

  - $B = \{\text{a soma dos dados é igual a } 10\}$.

  - $A = \{\text{no primeiro dado saiu } 4\}$.

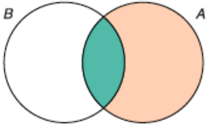
  - **Probabilidade condicional de B dado A:**

$$P(B \mid A)$$

Probabilidade condicional $P(B \mid A)$

- Suponha que o resultado do experimento esteja contido no evento A.
- Para que o resultado esteja também no evento B, ele precisa necessariamente estar tanto em A quanto em B, ou seja, precisa estar em $A \cap B$.
- Mas, como sabíamos desde o início que o resultado estava em A, nosso espaço amostral agora é reduzido para somente os elementos de A.

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



Exemplo: Lançamento de dois dados

Voltando ao exemplo dos dois dados.

- $A =$ no primeiro dado saiu 4.

$$A = \{(4,1),(4,2),(4,3),(4,4),(4,5),(4,6)\}$$

- $B =$ a soma dos dados é igual a 10.

$$B = \{(4,6),(5,5),(6,4)\}$$

- Então $A \cap B = \{(4,6)\}$. Portanto:

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{1/36}{6/36} = \frac{1}{6}$$

Exemplo: Qual a chance de cair na malha fina?

80.2 milhões de declarações.

Table 1: Renda x Caiu na Malha Fina?

	Sim	Não	Total
D - abaixo de 25.000	90	14010	14100
C - 25.000 a 49.999	71	30629	30700
B - 50.000 a 99.999	69	24631	24700
A - acima de 100.000	80	10620	10700
Total	310	79890	80200

Para simplificar, uma frequência de 90 representa 90.000.

Exemplo: Qual a chance de cair na malha fina?

Espaço amostral:

$$\Omega=\{(\text{A, sim}), (\text{A,não}), (\text{B,sim}), (\text{B,não}), (\text{C, sim}), (\text{C,não}), (\text{D,sim}), (\text{D,não})\}$$

Qual a probabilidade de cair na malha fina se a renda for acima de 100.000?

- $\mathcal{A} = \{\text{caiu na malha fina}\} = \{(\text{A,sim}),(\text{B,sim}),(\text{C,sim}),(\text{D,sim})\}$
- $\mathcal{B} = \{\text{renda acima de } 100.000\} = \{(\text{A,sim}),(\text{A,não})\}$

$$\begin{aligned} P(\mathcal{A} \mid \mathcal{B}) &= \frac{P(\mathcal{A} \cap \mathcal{B})}{P(\mathcal{B})} = \frac{P(\{(\text{A,sim})\})}{P(\{(\text{A,sim}),(\text{A,não})\})} \\ &= \frac{80/80200}{10700/80200} = 0.007 \end{aligned}$$

Exemplo: Qual a chance de cair na malha fina?

Probabilidade condicional por faixa de renda em 2002

Renda X Caiu na Malha Fina?	Sim	Não	Total
D - abaixo de 25.000	90/14100	14010/14100	14100/14100
C - 25.000 a 49.999	71/30700	30629/30700	30700/30700
B - 50.000 a 99.999	69/24700	24631/24700	24700/24700
A - acima de 100.000	80/10700	10620/10700	10700/10700

Exemplo: Qual a chance de cair na malha fina?

Probabilidade condicional por faixa de renda em 2002

Renda X Caiu na Malha Fina?	Sim	Não	Total
D - abaixo de 25.000	0.006	0.994	1
C - 25.000 a 49.999	0.002	0.998	1
B - 50.000 a 99.999	0.003	0.997	1
A - acima de 100.000	0.007	0.993	1

Independência

Vimos que:

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

Regra da multiplicação:

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

Quando $P(B \mid A) = P(B)$ (informação sobre A não altera a probabilidade do evento B), dizemos que B e A são **independentes**. Neste caso:

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

Exemplo

Considere o lançamento de dois dados “justos” (36 resultados possíveis têm a mesma probabilidade de ocorrer).

Considere os eventos:

- A: primeiro dado tem resultado 3.
- B: soma dos dados é igual a 8.
- C: soma dos dados é igual a 7.

Exemplo

Eventos A e B são independentes?

$P(A \cap B) = P(\{(3,5)\}) = \frac{1}{36}$

$P(A) = P(\{(3,1),(3,2),(3,3),(3,4),(3,5),(3,6)\}) = \frac{6}{36}$

$P(B) = P(\{(2,6),(3,5),(4,4),(5,3),(6,2)\}) = \frac{5}{36}$

$$P(A \cap B) = \frac{1}{36} \neq P(A) \times P(B) = \frac{6}{36} \times \frac{5}{36}$$

Portanto, A e B não são eventos independentes.

Exemplo

Ainda no mesmo exemplo: os eventos A e C são independentes?

$P(A \cap C) = P(\{(3,4)\}) = \frac{1}{36}$

$P(A) = P(\{(3,1),(3,2),(3,3),(3,4),(3,5),(3,6)\}) = \frac{6}{36}$

$P(C) = P(\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}) = \frac{6}{36}$

$$P(A \cap C) = \frac{1}{36} = P(A) \times P(C) = \frac{6}{36} \times \frac{6}{36}$$

Portanto, A e C são eventos independentes.

Exemplo

Suponha que A e B sejam dois eventos disjuntos.

Suponha que $P(A) > 0$ e $P(B) > 0$.

A e B são independentes?

A e B são disjuntos, então $A \cap B = \varnothing$ e $P(A \cap B) = 0$.
$P(A) > 0$ e $P(B) > 0$, portanto:

$$P(A \cap B) = 0 \neq P(A)P(B)\,.$$

A e B não são independentes.

Além disso: $P(B \mid A) = \frac{P(A \cap B)}{P(A)} = 0$, ou seja, dado que A ocorre, B não ocorre.

Exemplo

Em uma família com duas crianças, considere os eventos:

A={a primeira criança é uma menina} e B={as duas crianças são meninas}.

- Mostre que $P(B \mid A) = 1/2$.

  $\Omega = \{FF, MM, FM, MF\}$

  $A = \{FF, FM\}, \qquad B = \{FF\} \qquad \Longrightarrow \quad B \cap A = B$

  Portanto,

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \frac{P(\{FF\})}{P(\{FF, FM\})} = \frac{1/4}{1/2} = 1/2$$

Exemplo

Em uma família com duas crianças, considere os eventos:

A={a primeira criança é uma menina} e B={as duas crianças são meninas}.

- A e B são eventos independentes?

  $\Omega = \{FF, MM, FM, MF\}$

  $A = \{FF, FM\} \qquad B = \{FF\} \qquad \Longrightarrow \quad B \cap A = B$

  Então, $P(B \cap A) = P(B) = \frac{1}{4}$ e

$$P(A)P(B) = \frac{1}{2} \times \frac{1}{4} = \frac{1}{8} \neq P(B \cap A)$$

Portanto, A e B não são independentes.

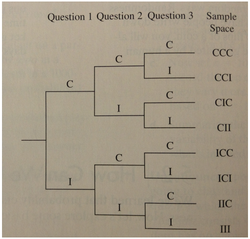
Chutar as respostas e ainda passar na prova

Chutar: escolher as respostas ao acaso.

Prova com três questões de múltipla escolha.

Em cada questão há 5 alternativas, apenas 1 é correta.

Experimento: anotar o resultado do aluno na prova.



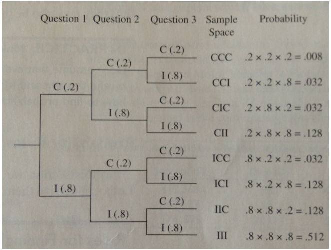
$$\Omega = \{CCC, CCI, CIC, CII, ICC, ICI, IIC, III\}$$

Chutar as respostas e ainda passar na prova

Quais as probabilidades dos eventos do espaço amostral?

Para cada questão: $P(C) = 0.2 \quad \text{e} \quad P(I) = 0.8$

$P(CCC) = P(C) \times P(C) \times P(C) = 0.2^3 = 0.008$



Qual a probabilidade do aluno acertar pelo menos duas questões?

$P(CCC) + P(CCI) + P(CIC) + P(ICC) = 0.008 + 3 \times 0.032 = 0.104$

Cinto de segurança e acidentes

Uso de cinto / Sobreviveu	Sim (S)	Não ($\bar{S}$)	**Total**
Sim (C)	414368	510	**412878**
Não ($\bar{C}$)	162527	1601	**164128**
Total	**574895**	**2111**	**577006**

Qual a probabilidade de que a pessoa morreu no acidente?

$$P(\bar{S}) = \frac{2111}{577006} = 0.004$$

Qual a probabilidade de que a pessoa morreu dado que ela usava o cinto de segurança?

$$P(\bar{S} \mid C) = P(\bar{S} \cap C)/P(C) = \frac{510}{412878} = 0.001$$

Cinto de segurança e acidentes

Uso de cinto / Sobreviveu	Sim (S)	Não ($\bar{S}$)	**Total**
Sim (C)	414368	510	**414878**
Não ($\bar{C}$)	162527	1601	**164128**
Total	**576895**	**2111**	**579006**

Qual a probabilidade de que a pessoa morreu dado que ela não usava o cinto de segurança?

$$P(\bar{S} \mid \bar{C}) = \frac{P(\bar{S} \cap \bar{C})}{P(\bar{C})} = \frac{1601}{164128} = 0.01$$

Morte e uso de cinto são eventos independentes?

$$P(\bar{S} \mid \bar{C}) \neq P(\bar{S}) = \frac{2111}{579006} = 0.004$$
$$P(\bar{S} \mid C) \neq P(\bar{S})$$

Exemplo

Uma sacola contém 10 sementes de flores vermelhas e 5 de flores brancas. Selecionamos duas sementes ao acaso, uma a uma e sem reposição.

Qual é a probabilidade de que :

- a primeira semente seja vermelha?
- a segunda seja branca se a primeira foi vermelha?

Defina os eventos:

A: a primeira semente é vermelha e B: a segunda semente é branca

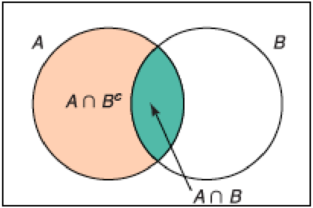
Então:

$$P(A) = \frac{10}{15} = \frac{2}{3} \qquad \text{e} \qquad P(B|A) = \frac{5}{14}$$

Teorema de Bayes

Considere dois eventos quaisquer A e B.

Para que um elemento esteja em A, há duas possibilidades:



- o elemento está em A e em B;
- o elemento está em A, mas não está em B.

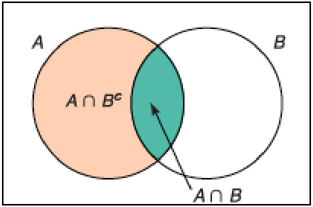
Teorema de Bayes

Portanto, podemos escrever:

$$A = (A \cap B) \cup (A \cap B^c)$$

As duas possibilidades são disjuntas, então:

$$P(A) = P(A \cap B) + P(A \cap B^c)$$



Teorema das Probabilidades Totais

Temos que:

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

$$P(A \cap B^c) = P(A \mid B^c)P(B^c)$$

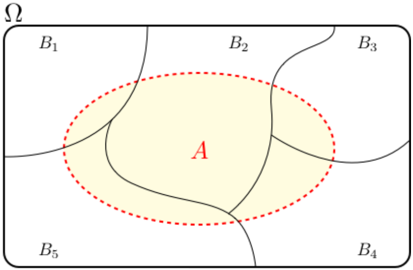
Então reescrevemos:

$$P(A) = P(A \mid B)P(B) + P(A \mid B^c)P(B^c)$$

Interpretação: a probabilidade do evento A é uma média ponderada da probabilidade condicional do evento A dado que B ocorre e da probabilidade condicional do evento A dado que B não ocorre. O peso de cada probabilidade condicional é a probabilidade do evento que está sendo levado em conta ao calcular a probabilidade condicional de A.

Teorema das Probabilidades Totais

Dizemos que os eventos $\{B_1, B_2, \ldots, B_n\}$ formam uma partição do espaço amostral Ω se são mutuamente exclusivos e a união desses eventos é Ω.



Então,

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i) P(B_i)$$

Teorema de Bayes

Se considerarmos a partição B e B^c do espaço amostral Ω e A um evento em Ω. Então:

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A \mid B)P(B)}{P(A \mid B)P(B) + P(A \mid B^c)P(B^c)}$$

No caso geral, seja $\{B_1, \ldots, B_n\}$ uma partição de eventos de Ω e A um evento em Ω:

$$P(B_i \mid A) = \frac{P(A \mid B_i)P(B_i)}{\sum_{i=1}^{n} P(A \mid B_i)P(B_i)}$$

Exemplo: Teste de diagnóstico

Um exame de sangue é 99% efetivo em detectar uma certa doença quando esta está presente. No entanto, 2% são falso-positivos. Suponha que 0,5% da população tem a doença.

Qual a probabilidade condicional de que um indivíduo testado aleatoriamente tenha a doença dado que o teste deu positivo?

Considere os eventos: $D = \{\text{estar doente}\}$ e $TP = \{\text{testar positivo}\}$

$P(TP \mid D) = 0.99 \qquad P(TP \mid D^c) = 0.02 \quad \text{e} \quad P(D) = 0.005$

$$
\begin{aligned}
P(D \mid TP) &= \frac{P(TP \mid D)P(D)}{P(TP \mid D)P(D) + P(TP \mid D^c)P(D^c)} \\
&= \frac{0.99 \times 0.005}{0.99 \times 0.005 + 0.02 \times 0.995} = 0.20
\end{aligned}
$$

Câncer de Mama

Câncer de mama afeta 1% das mulheres.

Mamografia é o teste padrão para detectar câncer de mama. Mas sabe-se que não é um teste perfeito.

Estatísticas mostram que a mamografia é 80% efetiva em detectar o câncer quando este realmente existe. E 9.6% das mamografias resultam em falsos positivos (teste positivo quando o câncer não existe).

Suponha que sua mãe faz uma mamografia e o resultado é positivo.

Qual é a probabilidade dela realmente estar com câncer de mama?

Exemplo: Companhia de Seguros

Uma companhia de seguros acredita que as pessoas podem ser divididas em duas categorias:

1. aquelas que estão mais sujeitas a acidentes.
2. aquelas que não estão mais sujeitas a acidentes.

Os dados indicam que uma pessoa da categoria 1 terá um acidente durante o período de um ano com probabilidade 0.1. A probabilidade para todas as outras pessoas é 0.05.

Suponha que a probabilidade de um novo cliente pertencer à categoria 1 seja 0.2.

Exemplo: Companhia de Seguros

Pergunta: Qual a probabilidade de que o novo cliente tenha um acidente durante o primeiro ano?

Considere os eventos:

A: o novo cliente tem um acidente durante o primeiro ano

B: o novo cliente pertence à categoria 1

B^c: o novo cliente pertence à categoria 2

Pelo Teorema das Probabilidades Totais:

$$\begin{aligned} P(A) &= P(A \mid B)P(B) + P(A \mid B^c)P(B^c) \\ &= 0.1 \times 0.2 + 0.05 \times 0.8 = 0.06 \end{aligned}$$

Exemplo: Companhia de Seguros

Pergunta: Se um novo cliente tem um acidente durante o primeiro ano, qual é a probabilidade de que ele pertença à categoria 1?

A: o novo cliente tem um acidente durante o primeiro ano

B: o novo cliente pertence à categoria 1

Pelo Teorema de Bayes

$$\begin{aligned} P(B \mid A) &= \frac{P(B \cap A)}{P(A)} = \frac{P(A \mid B)P(B)}{P(A)} \\ &= \frac{0.1 \times 0.2}{0.06} = \frac{1}{3} \end{aligned}$$

Exemplo: DNA e crime

Dado que o réu é inocente (I), suponha que a probabilidade do DNA dele ser compatível (C) com o DNA encontrado na cena do crime seja 1 em um milhão.

$$P(C \mid I) = 0.000001$$

Dado que o réu é culpado ($\bar{I}$), suponha que a probabilidade do DNA dele ser compatível com o DNA da cena do crime seja 0.99.

$$P(C \mid \bar{I}) = 0.99$$

O DNA do réu é compatível com o DNA da cena do crime.

Exemplo: DNA e crime

Encontre a probabilidade do réu ser inocente dado que o DNA é compatível, sendo que a probabilidade incondicional dele ser inocente, $P(I)$, é 0.5.

Queremos $P(I \mid C)$, sendo que $P(I) = P(\bar{I}) = 0.5$

Pelo Teorema de Bayes:

$$\begin{aligned}
P(I \mid C) &= \frac{P(I \cap C)}{P(C)} = \frac{P(C \mid I)P(I)}{P(C)} \\
&= \frac{P(C \mid I)P(I)}{P(C \mid I)P(I) + P(C \mid \bar{I})P(\bar{I})} \\
&= \frac{0.000001 \times 0.50}{0.000001 \times 0.5 + 0.99 \times 0.5} = 0.000001
\end{aligned}$$

A chance de ser inocente dado que houve compatibilidade de DNA é 1 em 1 milhão.

Exemplo: DNA e crime

Encontre a probabilidade do réu ser inocente dado que o DNA é compatível, sendo que a probabilidade incondicional dele ser inocente, $P(I)$, é 0.99.

$$
\begin{aligned}
P(I \mid C) &= \frac{P(I \cap C)}{P(C)} = \frac{P(C \mid I)P(I)}{P(C)} \\
&= \frac{P(C \mid I)P(I)}{P(C \mid I)P(I) + P(C \mid \bar{I})P(\bar{I})} \\
&= \frac{0.000001 \times 0.99}{0.000001 \times 0.99 + 0.99 \times 0.01} = 0.00001
\end{aligned}
$$

A chance de ser inocente dado que houve compatibilidade de DNA é 1 em 100 mil.

Leituras

- OpenIntro: seção 2.2.
- Ross: seções 4.5, 4.6
- Magalhães: capítulo 2


HOW COME YOU LIKE THOSE PROBABILITY PEOPLE SO MUCH?
THEY'RE FUN PEOPLE... BESIDES, THEY'RE NEVER NEGATIVE.