Introdução à Bioestatística
Conceitos de Probabilidade

Enrico A. Colosimo/UFMG

http://www.est.ufmg.br/~enricoc/

Depto. Estatística - ICEx - UFMG

Tipos de Fenômenos

1. Aleatório: Situação ou acontecimentos que não pode ser previsto com certeza.

2. Determinístico: Situação ou acontecimentos que pode ser previsto com certeza. Ex.: Leis físicas de Newton

Teoria de Probabilidade

DEF.: É um conjunto de métodos para lidar com a incerteza existente nos fenômenos aleatórios.

1. É fundamental para entender as técnicas estatísticas.
2. Exemplos: Prevalência , testes diagnósticos (sensibilidade, especificidade, PFP, PFN).

Situação: Corrida de Cavalos

Suponha que você tenha R$10.000,00 para apostar em uma corrida de cavalos.

Abaixo estão as probabilidades dos cavalos A,B e C ganharem a corrida. Em qual cavalo você apostaria?

Cavalo A: 75%

Cavalo B: 20%

Cavalo C: 5%

Conceitos Básicos

01. Espaço Amostral:

    Conjunto de todos os resultados possíveis de um experimento (Ω)

02. Eventos:

    Subconjuntos de Ω (representados por letras maiúsculas).
    Conjunto vazio: ϕ

03. Operações com Eventos/Conjuntos:

    União, intersecção, complementar, etc.

Exemplo (a)

Um dado equilibrado é lançado:

- Experimento: Lançar um dado (ν).
- Característica de interesse: Número da face superior.
- Espaço Amostral: Ω = $\{$1,2,3,4,5,6$\}$.
- Eventos:
    - A = $\{$1,3,5$\}$ (resultados ímpares)
    - B = $\{$2,4,6$\}$ (resultados pares)
    - C = $\{$1,2,3$\}$ (os três menores resultados)

Exemplo (b)

Um certo procedimento cirúrgico é testado para a extração de um tumor:

- Experimento: Realizar a cirurgia para extrair o tumor (ν).
- Característica de interesse: Tempo de sobrevida do paciente.
- Espaço Amostral: Ω = $\{$0,1,2,3,4,5,6,...$\}$ anos
- Eventos:
  - A = $\{0\}$ (morte com menos de um ano)
  - B = $\{1,2,3,4,5\}$ (sobrevida de no máximo 5 anos)
  - C = $\{6, ...\}$ (sobrevida maior que 5 anos)

Operações sobre conjuntos

A **união** de dois eventos A e B representa a ocorrência de, pelo menos, um dos eventos A ou B. A união é denotada por $A \cup B$.

- Exemplo: (experimento do dado)
  - A = 'observa-se um número ímpar'
  - B = 'observa-se um número $\leqslant$ 3'
  - $A \cup B = \{1,2,3,5\}$

Operações sobre conjuntos

A **interseção** entre os (ou dos) eventos A e B é a ocorrência simultânea de A e B. A interseção é denotada por $A \cap B$.

- Exemplo: (experimento do dado)
  - A = 'observa-se um número ímpar'
  - B = 'observa-se um número $\leqslant 3$
  - $A \cap B = \{1{,}3\}$

Operações sobre conjuntos

Dois eventos A e B são **disjuntos** ou **mutuamente exclusivos** quando não têm elementos em comum. Isto é, $A \cap B = \phi$.

- Exemplo: (experimento do dado)
  - A = 'observa-se um número ímpar'
  - B = 'observa-se um número par'
  - $A \cap B = \phi$

Operações sobre conjuntos

Dizemos que A e B são **complementares** se sua união é o espaço amostral e a sua interseção é vazia. O complementar de A é representado por A^c e temos:

$$A \cup A^c = \Omega$$

e

$$A \cap A^c = \phi$$

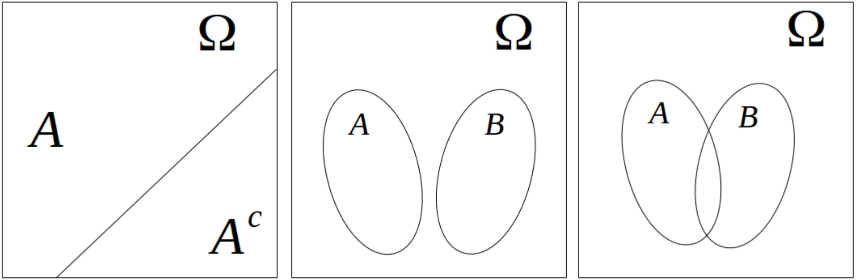
- Exemplo: (experimento do dado)
  - A = 'observa-se um número $\leqslant$ 3'
  - $A^c = \{4,5,6\}$

Axiomas de Probabilidade

- Uma função P(.) é denominada probabilidade se satisfaz as seguintes condições:

1. $0 \leqslant P(A) \leqslant 1$, para qualquer evento A em Ω
2. $P(\Omega) = 1$
3. $P(A \cup B) = P(A) + P(B)$ se A e B são disjuntos

Diagrama de Venn



$A \cup A^c = \Omega$

$A \cap A^c = \Phi$

$A \cap B = \Phi$

$A \cup B$

$A \cap B$

Pelo Diagrama de Venn mostre que:

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

Exemplo (Hipertensão)

Seja A o evento em que uma pessoa tenha pressão diastólica (PD) normotensiva (DBP $<$ 90), e seja B o evento que a pessoa tenha PD na 'borderline' DBP (90 $\leqslant$ DBP $\leqslant$ 95).

Suponha que P(A)=0,7 e P(B)=0,1.

Seja Z o evento que a pessoa tenha DBP $\leqslant$ 95. Assim:

$$P(Z) = P(A \cup B) = P(A) + P(B) = 0{,}8$$

pois os eventos não podem ocorrer simultaneamente ($A \cap B = \phi$)

Exemplo (Hipertensão)

Sejam os eventos A e B definidos como no exemplo anterior:

$$A = \{DBP < 90\} \text{ e } B \ \{90 \leqslant DBP \leqslant 95\}$$

Então, $A \cup B = \{DBP < 95\}$.

Sejam os eventos C e D:

$$C = \{DBP \geq 90\} \text{ e } D \ \{75 \leqslant DBP \leqslant 100\}$$

Os eventos C e D não são mutuamente excludentes. Ambos podem ocorrer quando $90 \leqslant DBP \leqslant 100$ ou seja:

$$C \cap D = \{90 \leqslant DBP \leqslant 100\}$$

Situação Real (livro: Andar do Bêbado)

Uma situação clínica foi exposta para um grupo de residentes: Atribua escore 1 (menos provável), 2 ou 3 (mais provável) para os seguintes sintomas de um paciente com embolia pulmonar:

- A: paralisia parcial
- B: falta de ar
- C: paralisia parcial e falta de ar

91% dos residentes apontaram a opção C como a mais provável.

Você concorda?

Probabilidade Condicional

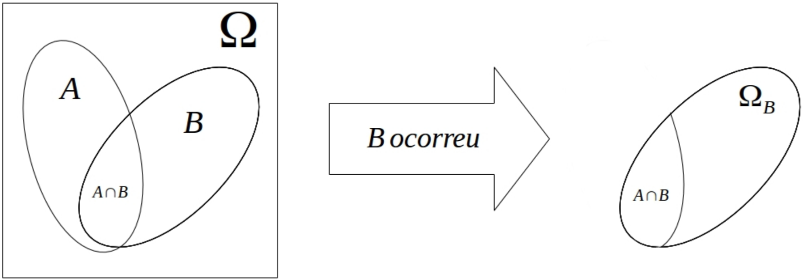
Dados dois eventos A e B, a probabilidade condicional de A dado que ocorreu B, P($A|B$) é dada por:

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

P($A|B$) é a proporção de A dentro de B onde o 'novo' espaço amostral (Ω_B) é o espaço amostral de B tal que:

$$0 \leq P(A|B) \leq 1$$

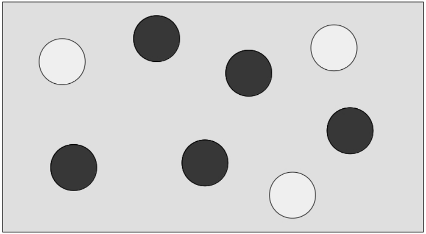
Probabilidade Condicional - Interpretação Gráfica



Exemplo Didático

Uma caixa contém 3 bolas brancas e 5 pretas, todas idênticas em peso e tamanho.

- Qual a probabilidade de tirar ao acaso uma bola branca?
- Qual a probabilidade de tirar ao acaso uma bola preta?
- Qual a probabilidade de tirar uma segunda bola branca se a primeira foi preta (sem reposição)?



Independência de Eventos

Dois eventos A e B são independentes se a informação da ocorrência (ou não) de B não altera a probabilidade da ocorrência do evento A.

- $P(A \mid B) = P(A)$ ou
- $P(A \cap B) = P(A) \cdot P(B)$

Exemplo (Hipertensão)

Suponha que estamos realizando uma avaliação de hipertensão em famílias. Seja o espaço amostral constituído de todos os pares da forma (X,Y) em que representam as medições de DBP da mãe e do pai. Sejam os eventos abaixo:

$$A = \{\text{DBP da mãe} \geq 95\}$$

$$B = \{\text{DBP do pai} \geq 95\}$$

e P(A)=0,1 e P(B)=0,2

Se A e B são independentes, então a probabilidade que a mãe e o pai sejam hipertensos é:

$$P(A \cap B) = P(A) \cdot P(B) = 0{,}1 \cdot 0{,}2 = 0{,}02 = 2\%$$

Exemplo (Hipertensão)

Uma forma de interpretar isso é considerar que o status hipertensivo da mãe não depende do status hipertensivo do pai.
Assim, considerando independência:

- em 10% das casas onde o pai é hipertensivo a mãe é também hipertensiva
- em 10% das casas onde o pai não é hipertensivo a mãe é hipertensiva

Admitimos nesse exemplo que as causas da hipertensão são genéticas. Se fossem, em certo ponto, ambientais poderíamos admitir mais probabilidade de a mãe ser hipertensa quando o pai é hipertenso (B ocorre) do que quando o pai não é hipertenso (B não ocorre). Neste último caso, A e B não mais seriam independentes.

Exemplo

Sabendo que a probabilidade de nascer um menino é 0,5 encontre as seguintes probabilidades:

1. Um casal ter duas meninas?
2. Um casal ter duas meninas dado que o primeiro filho é uma menina?
3. Um casal ter duas meninas dado que um dos partos foi de uma menina?
4. Um casal ter duas meninas dado que já nasceu uma ”Marina”?

Exemplo (DST)

Suponha que dois médicos, A e B, testem os pacientes para verificar se possuem sífilis. Sejam os eventos:

- $A^+ = \{$ médico A dá diagnóstico positivo $\}$
- $B^+ = \{$ médico B dá diagnóstico positivo $\}$

Suponha que:

- O médico A diagnosticou 10% dos pacientes como positivo,
- O médico B diagnosticou 17% dos pacientes como positivo e
- Ambos os médicos diagnosticaram 8% dos pacientes como positivo.

Os eventos A e B são independentes?

$$P(A^+) \cdot P(B^+) = 0{,}10 \cdot 0{,}17 = 0{,}017$$
$$P(A^+ \cap B^+) = 0{,}08$$

Ou seja, os eventos não são independentes. Esse resultado faz sentido?

Exemplo (DST)

Qual a probabilidade de que o médico B apresente um diagnóstico positivo de sífilis, dado que o médico A fez um diagnóstico positivo?

$$P(A^+|B^+) = \frac{P(A^+ \cap B^+)}{P(A^+)} = \frac{0,08}{0,1} = 0,8$$

Assim, o médico B confirmará o diagnóstico positivo dado pelo médico A em 80% das vezes.

Resultados para dois eventos A e B em um espaço amostral Ω

1. Axioma 3:

   $P(A \cup B) = P(A) + P(B)$ se A e B forem disjuntos

2. Independência:

   $P(A \cap B) = P(A) \cdot P(B)$ se A e B forem independentes

3. Resultado Geral:

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

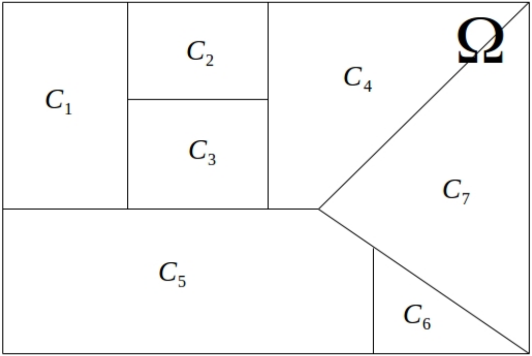
4. Resultado Geral:

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

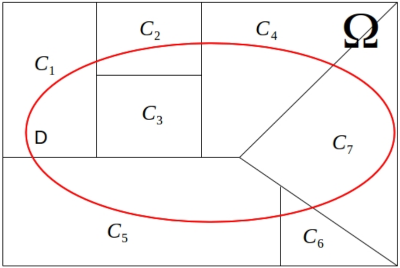
Partição do espaço amostral

Os eventos $C_1, C_2, ..., C_k$ formam uma partição do espaço amostral se eles não tem interseção (são disjuntos) entre si e se sua união é igual ao espaço amostral.

$$C_i \cap C_j = \phi, i \neq j \qquad C_1 \cup C_2 \cup ... \cup C_k = \Omega$$



Partição do espaço amostral



$$D = (C_1 \cap D) \cup (C_2 \cap D) \cup (C_3 \cap D) \cup \ldots \cup (C_7 \cap D)$$

$$\begin{aligned} P(D) &= P(C_1 \cap D) + P(C_2 \cap D) + P(C_3 \cap D) + \ldots + P(C_7 \cap D) \\ &= P(D|C_1)P(C_1) + P(D|C_2)P(C_2) + P(D|C_3)P(C_3) + \ldots + P(D|C_7)P(C_7) \end{aligned}$$

Partição do espaço amostral

As probabilidades condicionais P(B | A) e P(B | A^c) e P(B) estão relacionadas da seguinte forma:

$$P(B) = P(B|A) \cdot P(A) + P(B|A^c) \cdot P(A^c)$$

Exemplo (Câncer)

Suponha que A={mamografia positiva} e B= {câncer de mama} :

- Entre 100.000 mulheres com mamografias negativas 20 terão câncer de mama em dois anos:

$$P(B|A^c) = \frac{20}{100.000} = 0,0002$$

- Enquanto que 1 em 10 com mamograficas positivas terão câncer em dois anos:

$$P(B|A) = \frac{1}{10} = 0,1$$

- E, ainda, que 7% da população geral das mulheres terão mamografia positiva. Ou seja,

$$P(A) = \frac{7}{100} = 0,07$$

Exemplo (Câncer)

Qual a probabilidade de uma mulher desenvolver câncer nos próximos dois anos?

$$
\begin{aligned}
P(B) =& P(B|A)P(A) + P(B|A^c)P(A^c) \\
=& (0,1) \cdot (0,07) + (0,0002) \cdot (0,93) \\
=& 0,00719
\end{aligned}
$$

Exemplo (Oftalmologia)

Estamos planejando um estudo de 5 anos sobre catarata numa população de cinco mil pessoas com 60 anos ou mais.
Sabemos a partir do último Censo que nessa população:

- 45% tem entre 60-64 anos,
- 28% tem entre 65-69 anos,
- 20% tem entre 70-74 anos e
- 7% tem 75 anos ou mais

Também sabemos de um estudo anterior que 2,4%, 4,6%, 8,8% e 15,3% das pessoas nos respectivos grupos de idade desenvolverão catarata nos próximos 5 anos.

Exemplo (Oftalmologia)

Qual porcentagem dessa população desenvolverá catarata em 5 anos e quantas pessoas isso representa?

- A_1 = { idade entre 60-64 anos } e P(A_1) = 0,45
- A_2 = { idade entre 65-69 anos } e P(A_2) = 0,28
- A_3 = { idade entre 70-74 anos } e P(A_3) = 0,20
- A_4 = { idade $\geq$ 75 anos } e P(A_4) = 0,07
- B = { desenvolverá catarata em 5 anos }

Exemplo (Oftalmologia)

- P(B|A_1) = 0,024
- P(B|A_2) = 0,046
- P(B|A_3) = 0,088
- P(B|A_4) = 0,153

De modo que:

$$P(B) = P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2) + P(B|A_3) \cdot P(A_3) + P(B|A_4) \cdot P(A_4)$$

= (0,024) $\cdot (0,45) + (0,046) \cdot (0,28) + (0,088) \cdot (0,20) + (0,153) \cdot (0,07)$
= 0,052

CONCLUSÃO: 5,2% da população, ou seja, 260 (5.000 $\cdot$ 0,052) pessoas desenvolverão catarata nos próximos 5 anos.

Exercício

Em um cidade a população pode ser dividida em três grupos etários: A, B e C representando 25%, 35% e 40% do total, respectivamente.

Desses, 5%, 4% e 2% são alérgico a Penicilina.

Escolhe-se um indivíduo e verifica-se que é alérgico.

Qual a probabilidade dele ter vindo do grupo A, do B e do C?

Exercício - Monty Hall

Você foi convidado pelo Sílvio Santos para participar de um programa de auditório.

No palco existem três portas: em uma delas tem um carro e nas outras duas um livro em árabe.

Você ganha o carro se escolher a porta certa.

Digamos que você escolhe uma das portas ao acaso. Das duas restantes, o apresentador abre uma delas e vê-se que tem um dos livros.

Você mudaria sua aposta?

Resposta: https://www.youtube.com/watch?v=hcFkic2I8zU