Introdução à Probabilidade

Silvia Shimakura

silvia.shimakura@ufpr.br

Probabilidade

- **O que é probabilidade?**

  Medida que quantifica a incerteza de um acontecimento futuro.

- **Como quantificar incerteza?**

  Definição clássica: relaciona eventos favoráveis com eventos possíveis.
  Definição frequentista: baseada em repetições de um experimento, sob condições semelhantes, um grande número de vezes.

Problema trivial 1

- Experimento 1: Lançamento de um dado balanceado

- Espaço amostral: conjunto de todos os resultados possíveis

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

- Evento A: face impar e menor que 5

$$A=\{1, 3\}$$

- Evento B: face par

$$B=\{2, 4, 6\}$$

Cálculo de probabilidades

- Se os eventos simples de E forem equiprováveis a probabilidade de um evento A ocorrer:

$$P(A) = \frac{\text{número de elementos em } A}{\text{número de elementos em } E}$$

- Experimento 1: E={1,2,3,4,5,6} → 6 elementos

  - Evento A: A={1, 3} → P(A)=2/6
  - Evento B: B={2, 4, 6} → P(B)=3/6

Problema trivial 2

- Experimento 2: Lançamento de 2 dados balanceados

- Espaço amostral:

$E=\{(1,1), (1,2), \ldots, (1,6), (2,1), (2,2), \ldots, (2,6), \ldots, (6,6)\}$

$\rightarrow$ 36 elementos $\rightarrow$ equiprováveis

- Evento F: a soma dos dois valores é 10

$$F=\{(4,6), (5,5), (6,4)\}\rightarrow P(F)=3/36$$

- Evento G: os dois valores são iguais

$$G=\{(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)\}\rightarrow P(G)=6/36$$

Problema menos trivial

- Experimento 3: Lançamento de uma moeda
- Espaço amostral: E={Cara, Coroa}
- Evento C: Cara

  C={Cara}
- Eventos simples são equiprováveis?
- P(C)=?

Visão frequentista de probabilidade

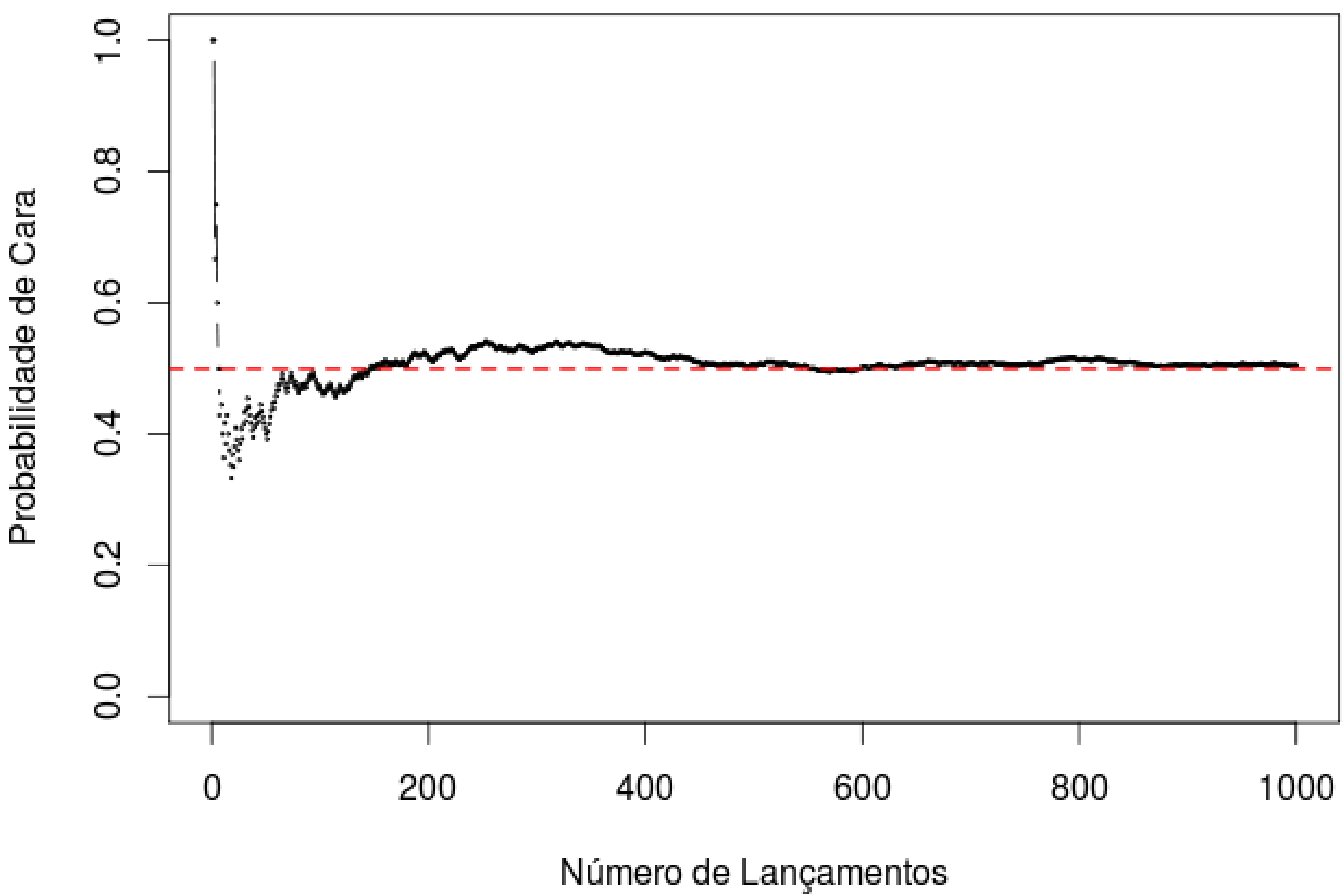
- E se os eventos simples não forem equiprováveis?

- Probabilidade: frequência relativa de ocorrência do evento para um grande número de sorteios

Frequência relativa

- C: Cara   O: Coroa

Resultado	C	C	C	O	C	O	O	O	O	O	O	C
Frequência acumulada de Caras	1	2	3	3	4	4	4	4	4	4	4	5
Número de lançamentos	1	2	3	4	5	6	7	8	9	10	11	12
Freq. relativa de caras	1/1	2/2	3/3	3/4	4/5	4/6	4/7	4/8	4/9	4/10	4/11	5/12
%	100	100	100	75	80	67	57	50	44	40	36	42



Probabilidade de Cara
0.0
0.2
0.4
0.6
0.8
1.0
0
200
400
600
800
1000
Número de Lançamentos

Tipos especiais de eventos

- Evento complementar de A: elementos do espaço amostral E que não estão em A

E={1, 2, 3, 4, 5, 6}   A={1, 3}

$$\bar{A} = \{2, 4, 5, 6\}$$

- Evento interseção: elementos estão em A e em B

A={1, 3}   B={2, 4, 6} $A \cap B = \emptyset$

- Evento união: elementos que estão em A ou B

$A \cup B = \{1, 2, 3, 4, 6\}$

Tipos especiais de eventos

- Eventos mutuamente exclusivos: não existem elementos comuns em A e B

$$A=\{1, 3\} \quad B=\{2, 4, 6\} \quad A \cap B = \emptyset$$

Propriedades de probabilidade

- $0 \leq P(A) \leq 1$, para qualquer evento A

- $P(E)=1$, em que E é o espaço amostral

- $$\boxed{P(\bar{A}) = 1 - P(A)}$$

- Para dois eventos A e B quaisquer, a probabilidade de que A ou B ocorra:

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

- Se A e B são mutuamente exclusivos, a probabilidade de que A ou B ocorra é a soma das probabilidades.

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

Probabilidade condicional

- É a probabilidade de B dado que A ocorreu. Notação: $P(B|A)$

- Para <u>A e B quaisquer</u>

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

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

- Para <u>A e B independentes</u>

$$P(B|A) = P(B)$$

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

Exemplo: Lançamento de dois dados não viciados

- Espaço amostral: 36 elementos com prob 1/36 cada
- Evento B: soma dos dados é 8
- B=?
- P(B)=???

Exemplo: Lançamento de dois dados não viciados

- Espaço amostral: 36 elementos com prob 1/36 cada
- Evento B: soma dos dados é 8
- B={(2, 6), (6, 2), (3, 5), (5, 3), (4, 4)}
- P(B)=?

Exemplo: Lançamento de dois dados não viciados

- Espaço amostral: 36 elementos com prob 1/36 cada
- Evento B: soma dos dados é 8
- B={(2, 6), (6, 2), (3, 5), (5, 3), (4, 4)}
- P(B)=5/36

Exemplo: Lançamento de dois dados não viciados

- Espaço amostral: 36 elementos com prob 1/36 cada
- Evento B: soma dos dados é 8
- B={(2, 6), (6, 2), (3, 5), (5, 3), (4, 4)}
- P(B)=5/36

Se soubermos que o resultado no primeiro dado é 3, qual será a probabilidade da soma dos dois dados ser 8?

Exemplo (cont.)

- Evento B:

  B={(2, 6), (6, 2), (3, 5), (5, 3), (4, 4)} → P(B)=5/36

- Evento A: primeiro dado é 3

  A={(3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6)} → P(A)=6/36

- Evento interseção: **$A \cap B = \{(3,5)\} \rightarrow P(A \cap B) = 1/36$**

- $P(B|A)=?$
- A e B são independentes?
- A e B são mutuamente exclusivos?

Exemplo (cont.)

- Evento B:

  B={(2, 6), (6, 2), (3, 5), (5, 3), (4, 4)} → P(B)=5/36

- Evento A: primeiro dado é 3

  A={(3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6)} → P(A)=6/36

- Evento interseção: **A∩B={(3,5)} → P(A∩B)=1/36**

- P(B|A)= **P(A∩B)/P(A)=(1/36)/(6/36)=1/6**
- A e B são independentes?
- A e B são mutuamente exclusivos?

Exemplo: Distribuição de peso e pressão arterial

Pressão arterial	Peso			
	Excesso	Normal	Deficiente	Total
Elevada	0,10	0,08	0,02	0,2
Normal	0,15	0,45	0,20	0,8
Total	0,25	0,53	0,22	1

- Probabilidade de uma pessoa escolhida ao acaso ter pressão elevada?

- Probabilidade de uma pessoa escolhida ao acaso ter pressão elevada e excesso de peso?

- Sabendo que a pessoa tem excesso de peso, qual a probabilidade de uma pessoa escolhida ao acaso ter pressão elevada?

Exemplo: Distribuição de peso e pressão arterial (cont.)

- Peso em excesso e pressão arterial normal são eventos mutuamente exclusivos?
- Pressão arterial e peso são independentes?

Pressão arterial	Peso			
	Excesso	Normal	Deficiente	Total
Elevada	0,10	0,08	0,02	0,2
Normal	0,15	0,45	0,20	0,8
Total	0,25	0,53	0,22	1

Exemplo: Qualidade de teste diagnóstico

- Suponha que existam dois estados de saúde mutuamente exclusivos e exaustivos: D+ doente e D- não doente

- Seja T+ teste positivo e T- teste negativo

- Num estudo sobre o teste ergométrico, Wriner et al. (1979) compararam os resultados obtidos entre indivíduos com e sem doença coronariana.

- **T+**: mais de 1mm de depressão ou elevação do segmento ST, por pelo menos 0,08s, em comparação com paciente em repouso.
- **D+** e **D-**: angiografia (teste padrão ouro).

Exemplo: Qualidade de teste diagnóstico

Doença coronariana	Teste Ergométrico					
	T+		T-		Total	
D+	815	a	208	b	1023	a+b
D-	115	c	327	d	442	c+d
Total	930	a+c	535	b+d	1465	n

Temos interesse em responder as perguntas:

- Qual a probabilidade do teste ser positivo dado que o paciente é doente?
- Qual a probabilidade do teste ser negativo dado que o paciente não é doente?

Exemplo: Qualidade de teste diagnóstico

Doença coronariana	Teste Ergométrico					
	T+		T-		Total	
D+	815	a	208	b	1023	a+b
D-	115	c	327	d	442	c+d
Total	930	a+c	535	b+d	1465	n

Temos interesse em responder as perguntas:

- Qual a probabilidade do teste ser positivo dado que o paciente é doente? $P(T+|D+)=s$
- Qual a probabilidade do teste ser negativo dado que o paciente não é doente? $P(T-|D-)=e$

Exemplo: Qualidade de teste diagnóstico

Doença coronariana	Teste Ergométrico					
	T+		T-		Total	
D+	815	a	208	b	1023	a+b
D-	115	c	327	d	442	c+d
Total	930	a+c	535	b+d	1465	n

Temos interesse em responder as perguntas:

- Qual a probabilidade do teste ser positivo dado que o paciente é doente? $P(T+|D+)=s=815/1023=0,80$
- Qual a probabilidade do teste ser negativo dado que o paciente não é doente? $P(T-|D-)=e=327/442=0,74$

Exemplo: Qualidade de teste diagnóstico (cont.)

Doença coronariana	Teste Ergométrico					
	T+		T-		Total	
D+	815	a	208	b	1023	a+b
D-	115	c	327	d	442	c+d
Total	930	a+c	535	b+d	1465	n

- Qual a probabilidade de que uma pessoa com resultado de teste positivo realmente tenha a doença?

Exemplo: Qualidade de teste diagnóstico (cont.)

Doença coronariana	Teste Ergométrico					
	T+		T-		Total	
D+	815	a	208	b	1023	a+b
D-	115	c	327	d	442	c+d
Total	930	a+c	535	b+d	1465	n

- Qual a probabilidade de que uma pessoa com resultado de teste positivo realmente tenha a doença? $P(D+|T+)=VPP$

Exemplo: Qualidade de teste diagnóstico (cont.)

Doença coronariana	Teste Ergométrico					
	T+		T-		Total	
D+	815	a	208	b	1023	a+b
D-	115	c	327	d	442	c+d
Total	930	a+c	535	b+d	1465	n

- Qual a probabilidade de que uma pessoa com resultado de teste positivo realmente tenha a doença? $P(D+|T+)=VPP=815/930=0,88$

Exemplo: Qualidade de teste diagnóstico (cont.)

Doença coronariana	Teste Ergométrico					
	T+		T-		Total	
D+	815	a	208	b	1023	a+b
D-	115	c	327	d	442	c+d
Total	930	a+c	535	b+d	1465	n

- Qual a probabilidade de que uma pessoa com resultado de teste positivo realmente tenha a doença? P(D+|T+)=VPP=815/930=0,88

Teorema de Bayes

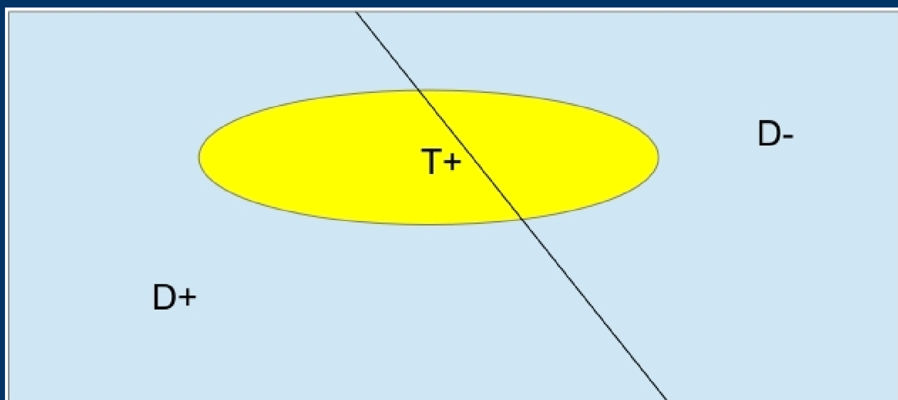
- Se A1, A2, … An são n eventos mutuamente exclusivos e exaustivos, tais que:

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n) = 1$$

- Pelo Teorema de Bayes:

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(A_1)P(B|A_1) + \ldots + P(A_n)P(B|A_n)}$$

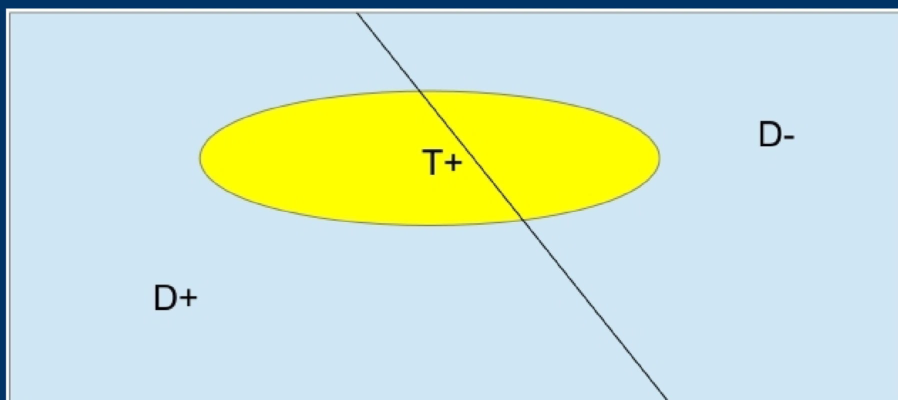
Exemplo: Qualidade de teste diagnóstico



- D+ ∩ D- = Ø
- D+ ∪ D- = E
- Conhecidos:
  - P(D+) e P(D-)
  - P(T+|D+) e P(T-|D-)

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

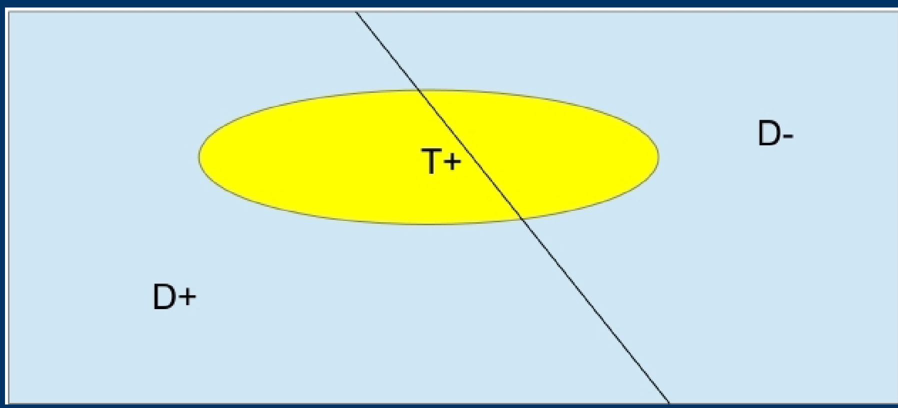
Exemplo: Qualidade de teste diagnóstico



- D+ ∩ D- = Ø
- D+ ∪ D- = E
- Conhecidos:
  - P(D+) e P(D-)
  - P(T+|D+) e P(T-|D-)

$$P(D+|T+)=\frac{P(D+\cap T+)}{P(T+)}=\frac{P(D+\cap T+)}{P(D+\cap T+)+P(D-\cap T+)}=\frac{\overset{p}{P(D+)}\,\overset{s}{P(T+|D+)}}{P(D+)P(T+|D+)+\underset{(1-p)}{P(D-)}\,\underset{(1-e)}{P(T+|D-)}}$$

Exemplo: Qualidade de teste diagnóstico



- D+ ∩ D- = Ø
- D+ ∪ D- = E
- Conhecidos:
  - P(D+) e P(D-)
  - P(T+|D+) e P(T-|D-)

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

$$VPP=P(D+|T+)=\frac{p\,s}{p\,s+(1-p)(1-e)}$$

Exemplo: Qualidade de teste diagnóstico (cont.)

Doença coronariana	Teste Ergométrico					
	T+		T-		Total	
D+	815	a	208	b	1023	a+b
D-	115	c	327	d	442	c+d
Total	930	a+c	535	b+d	1465	n

- Qual a probabilidade de que uma pessoa com resultado de teste positivo realmente tenha a doença? Suponha que a prevalência é 0,10.

$$VPP = \frac{p\,s}{p\,s + (1-p)(1-e)} = \frac{0{,}1 \times 0{,}80}{0{,}1 \times 0{,}8 + (1-0{,}1) \times (1-0{,}74)} = 0{,}25$$