



ME414 - Estatística para Experimentalistas

Parte 5

Probabilidade

Introdução

Diariamente, tomamos decisões com relação a eventos incertos:

Devo investir na bolsa?

Vale a pena fazer um plano odontológico?

Devo contratar um seguro para o meu carro?

Devo levar um guarda-chuva?

Devo me matricular numa disciplina eletiva com baixa taxa de aprovação?

Experimentos Aleatórios

Experimento: qualquer processo que produza uma observação ou resultado

Experimento Determinístico: é aquele que, dada uma ação controlada, sabemos exatamente qual será o resultado obtido



Exemplo: lançamento de um dado com todas as faces iguais a 6
Único resultado possível? 6

Experimento Aleatório: é aquele em que não se tem certeza sobre seus resultados, a priori. Múltiplos resultados podem ser obtidos a partir de uma única ação. Cada vez que se repete o experimento, o resultado pode ser diferente.



Exemplo: lançamento de um dado de seis faces
Resultados possíveis: {1, 2, 3, 4, 5, 6}

Probabilidade

Probabilidade: medida de incerteza sobre certos eventos ou características de interesse.

Tais eventos, em geral, estão associados a **experimentos aleatórios**.

Aleatorização:

- Jogar um dado.
- Jogar uma moeda.
- Girar uma roleta.

Ex: para aleatorizar dois tratamentos entre pacientes, pode-se lançar uma moeda. Se sair "cara" o paciente recebe a droga A, se sair "coroa", recebe a droga B.

Exemplo: Lançamento de dado



Você está jogando Ludo: um dado é usado para movimentar as peças.

Em certo ponto do jogo, durante a sua vez, o 6 sai 3 vezes seguidas e você vence o jogo!

Dentre os de 100 lançamentos do dado durante a sua vez, seu oponente no jogo comenta que o 6 saiu 23 vezes.

Seu oponente então reclama que o dado estava te favorecendo com tantos 6, portanto o dado não era "justo".

Exemplo: Lançamento de dado

Se o dado é "justo", quantos 6 você espera que ocorram em 100 lançamentos?

Se um dado "justo" é lançado diversas vezes, esperamos que o 6 ocorra $1/6$ das vezes.

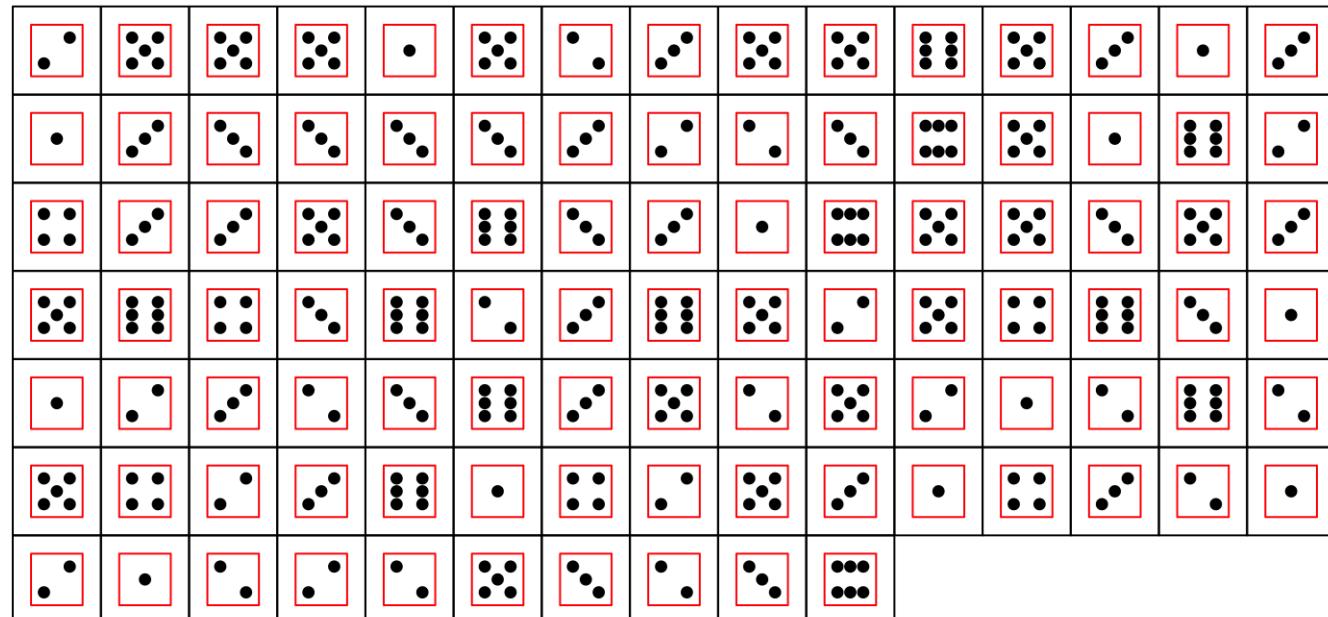
100 lançamentos: $100/6 \approx 17$ vezes.

É muito improvável que o 6 saia 23 vezes em 100 lançamentos? Como verificar?

- Lance o dado 100 vezes.
- Conte o número de 6 que aparecem.
- Repita várias vezes esse processo.

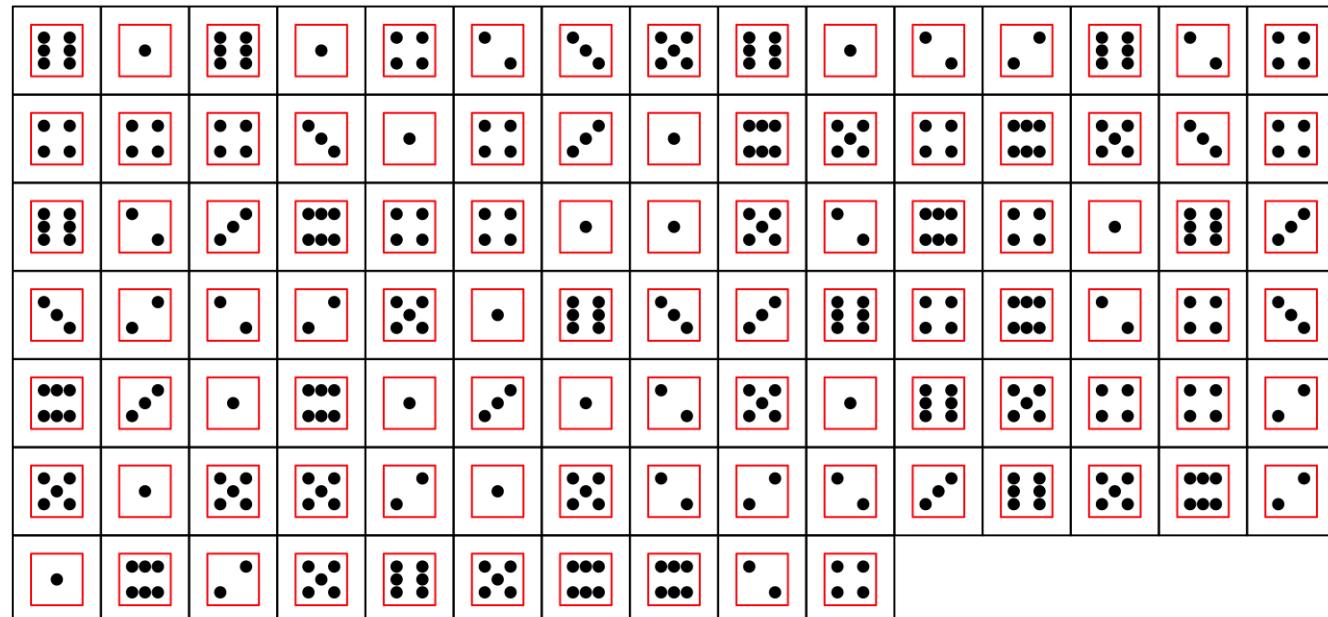
Você obtém assim a distribuição de frequências do 6 em 100 lançamentos do dado.

Simulação 1: lançamento de um dado 100 vezes



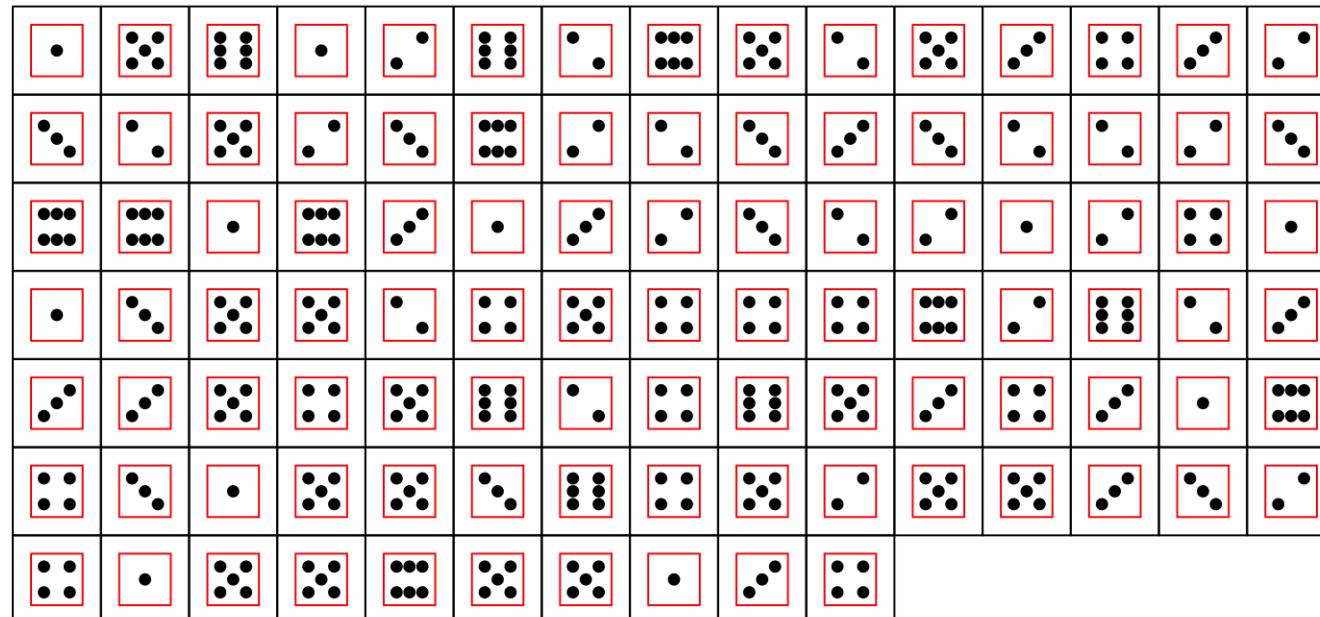
	1	2	3	4	5	6
Freq	12	21	28	6	20	13

Simulação 2: lançamento de um dado 100 vezes



	1	2	3	4	5	6
Freq	16	19	13	16	14	22

Simulação 3: lançamento de um dado 100 vezes



	1	2	3	4	5	6
Freq	11	21	22	13	19	14

Simulação 3: lançamento de um dado 100 vezes

Show 10 entries

Search:

Lançamento	Face 6 ocorre?	Soma Acumulada de 6	Proporção Acumulada de 6
1	não	0	0
2	não	0	0
3	sim	1	0.33
4	não	1	0.25
5	não	1	0.2
6	sim	2	0.33
7	não	2	0.29
8	sim	3	0.38
9	não	3	0.33
10	não	3	0.3

Showing 1 to 10 of 100 entries

Previous

1

2

3

4

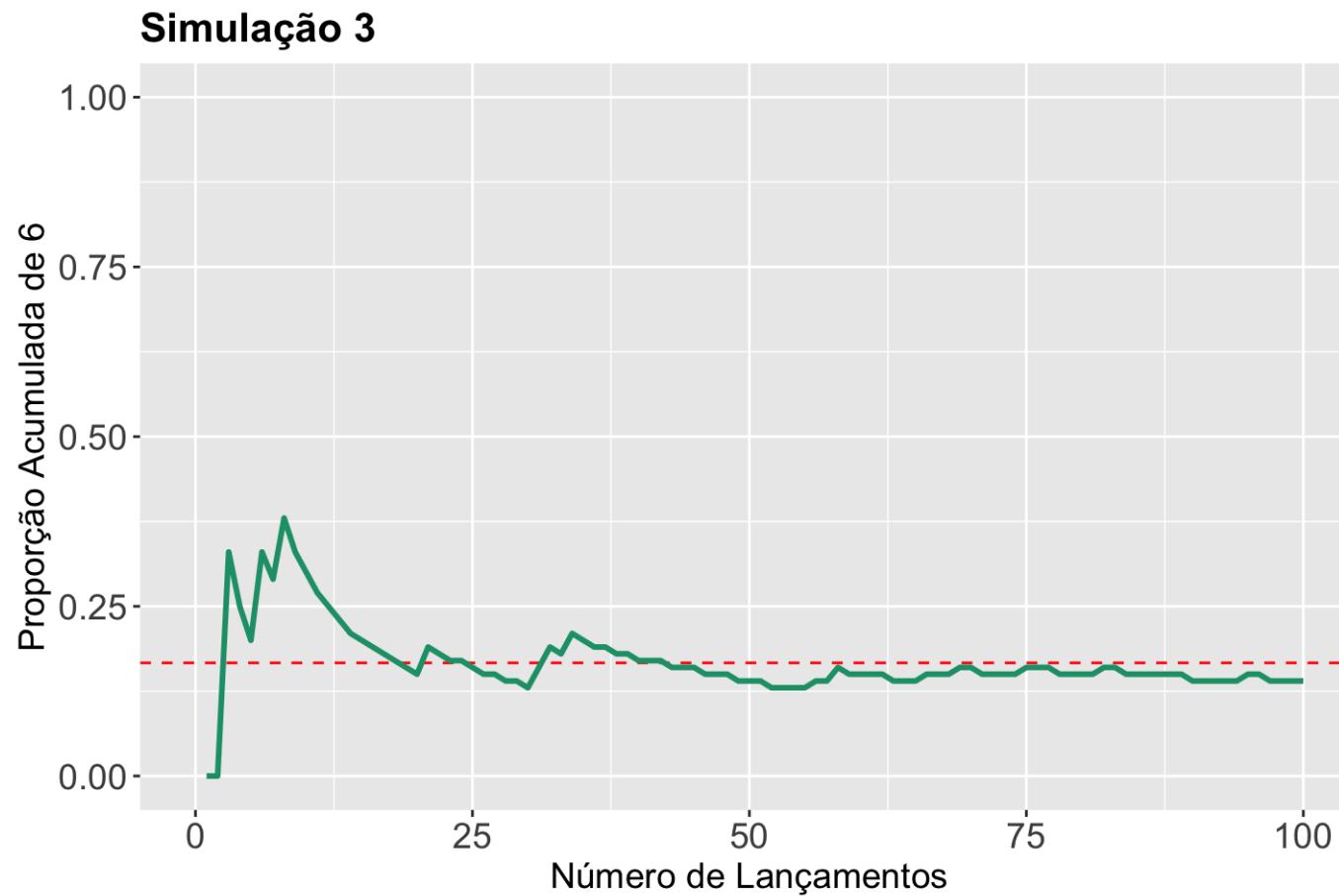
5

...

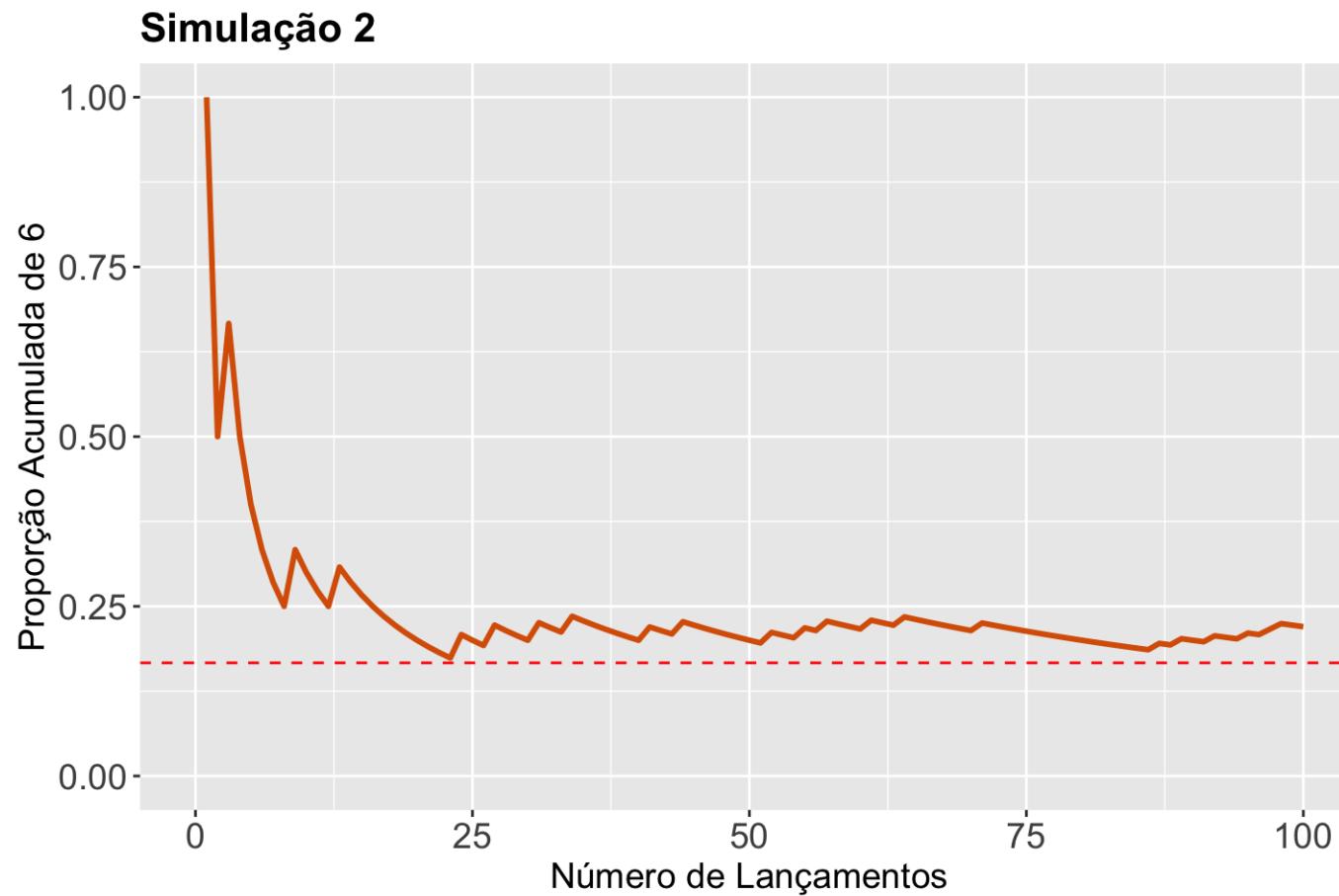
10

Next

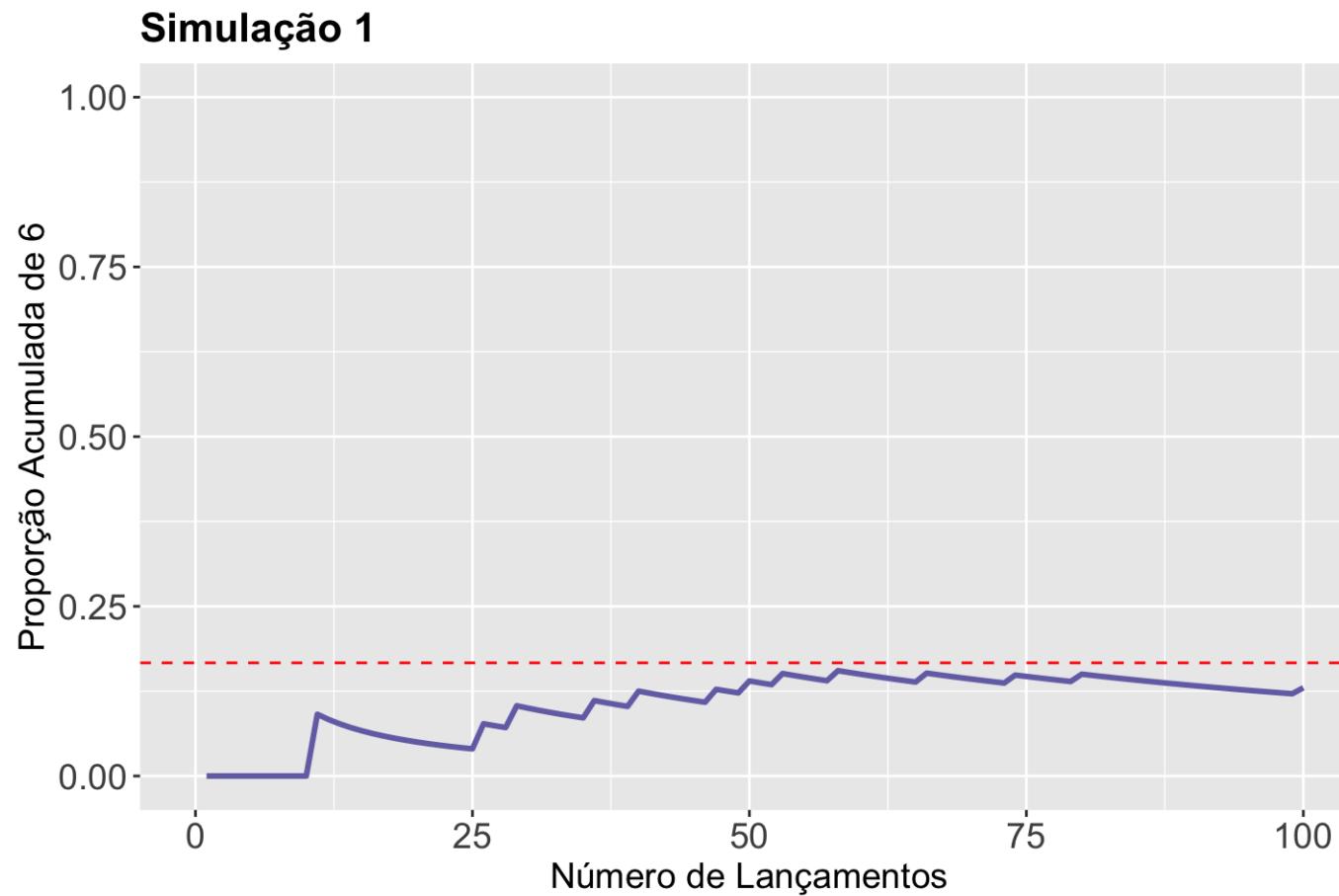
Simulação 3: lançamento de um dado 100 vezes



Simulação 2: lançamento de um dado 100 vezes



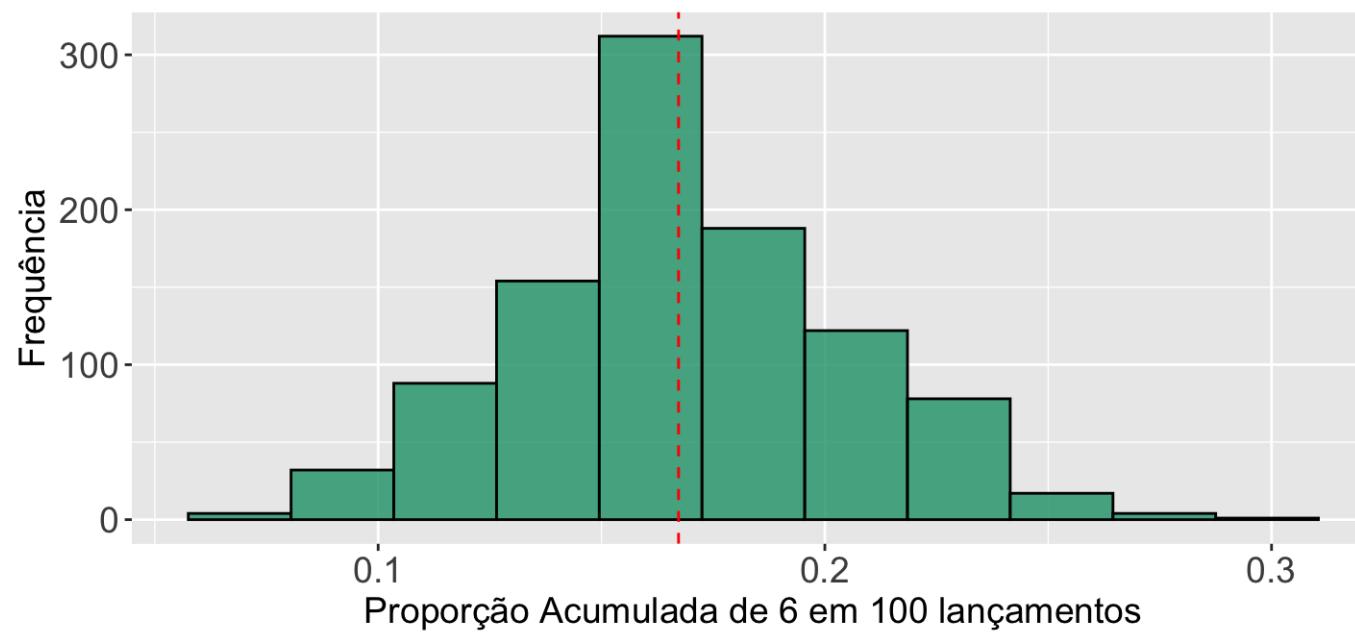
Simulação 1: lançamento de um dado 100 vezes



Simulação: lançamento de um dado 100 vezes

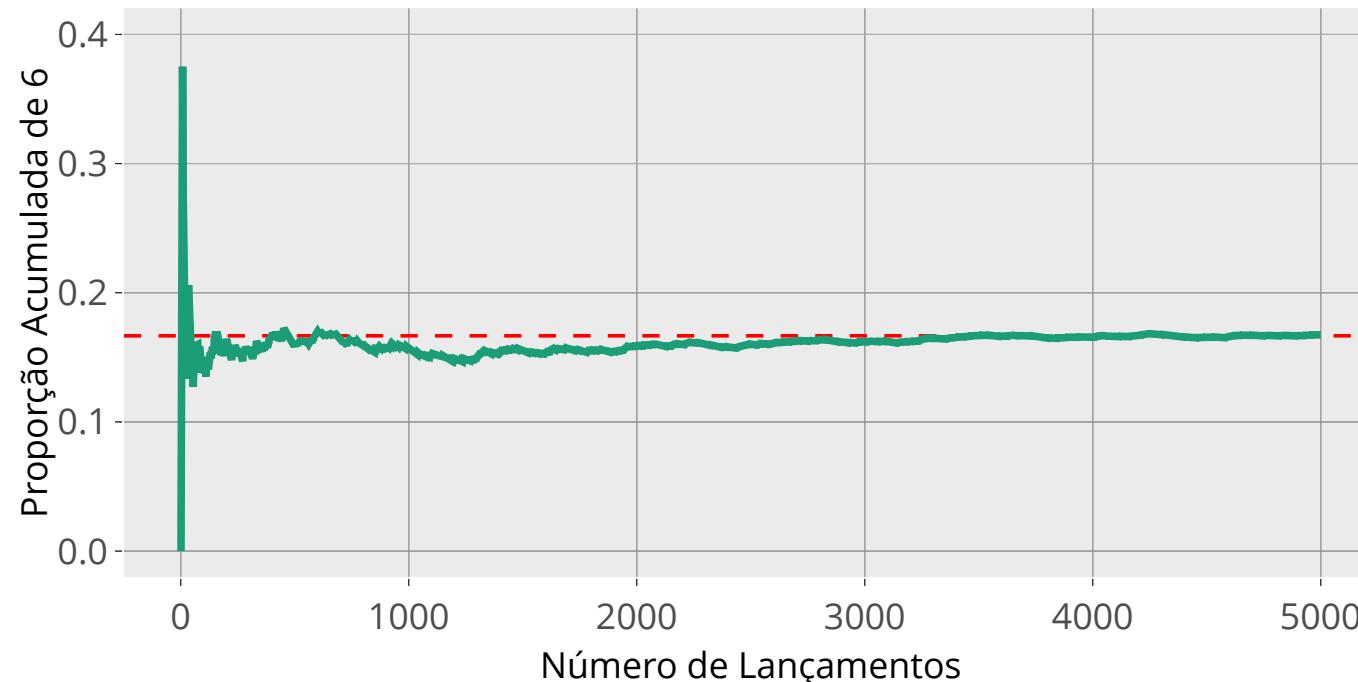
A cada simulação (100 lançamentos e anotando o total de 6) obtivemos um resultado diferente: 13, 22 e 14.

Se repetirmos a simulação 1000 vezes, temos uma idéia da distribuição de frequências da proporção de 6 em 100 lançamentos.



Média: 0.167. Mediana: 0.17.

Simulação 4: lançamento de um dado 5000 vezes



Com poucos lançamentos, a proporção de 6 pode flutuar bastante, mas com o aumento do número de lançamentos, a proporção acumulada de 6 estabiliza em $1/6$.

Lei dos Grandes Números

O resultado da simulação é um caso particular da [Lei dos Grandes Números](#), resultado provado em 1689 pelo matemático suíço Jacob Bernoulli.



Se um evento de probabilidade p é observado repetidamente em ocasiões independentes, a proporção da frequência observada deste evento em relação ao total número de repetições converge em direção a p à medida que o número de repetições se torna arbitrariamente grande.

Probabilidade

Em um fenômeno aleatório, a **probabilidade** de um resultado acontecer é a proporção de vezes que o resultado ocorreu quando consideramos muitas observações do fenômeno em questão.

- Quando dizemos que a probabilidade do **6** sair no dado é $1/6$, estamos dizendo que a proporção esperada de **6** em **vários lançamentos** (observações) do dado é $1/6$.
- Quando a previsão do tempo diz que a chance de chuva para hoje é 70%, quer dizer que para **vários dias** observados no passado com condições atmosféricas equivalentes ao dia de hoje a proporção observada de dias de chuva foi 0.7.

Como calcular probabilidades?

Em um fenômeno aleatório, a probabilidade de um resultado acontecer é a proporção de vezes que o resultado ocorreu quando consideramos muitas observações do fenômeno em questão.

- Esta definição nem sempre é útil.
- Quando a NASA lançou o primeiro ônibus espacial, como os cientistas sabiam a probabilidade de sucesso? Não havia nenhum dado sobre lançamentos no passado para que se pudesse calcular a probabilidade de sucesso.



Ônibus espacial Columbia

Probabilidade

Algumas vezes, é possível fazer alguma suposição sobre o fenômeno aleatório considerado.

- Ao lançar um dado, podemos assumir que cada valor de 1 a 6 tenha a mesma chance de ocorrer: $1/6$.
- Ao lançar uma moeda, podemos assumir que ela pode cair de um lado ou de outro com a mesma chance: $1/2$.

Outras vezes, podemos utilizar a distribuição de frequências observadas como uma estimativa das probabilidades.

Exemplo: dado

Estudar as probabilidades de ocorrência das faces de um dado.

Procedimento Empírico: lançar o dado um certo número n de vezes e contar o número de vezes, n_i , que a face $i = 1, 2, 3, 4, 5, 6$ ocorre.

Distribuição empírica das probabilidades:

$$f_i = \frac{n_i}{n}.$$

Para diferentes vezes que esse experimento for realizado, a distribuição de frequência terá resultados diferentes (exemplo anterior, lançamento de 100 dados, várias vezes).

No entanto, espera-se que esses resultados, apesar de distintos, sejam semelhantes.

Distribuição de Probabilidade

Procedimento Teórico: construir a distribuição de frequências populacionais (probabilidades) através de suposições teóricas.

Suposições:

- só podem ocorrer 6 faces: $\{1, 2, 3, 4, 5, 6\}$;
- o dado é perfeitamente equilibrado;
- então, cada face deve ocorrer o mesmo número de vezes, ou seja $f_i = \frac{1}{6}$.

Face	1	2	3	4	5	6	Total
Freq. Teórica	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

Espaço Amostral

Para quantificar incerteza em fenômenos aleatórios usando probabilidades, precisamos primeiro especificar o conjunto de todos os possíveis resultados do fenômeno em questão.

- **Espaço Amostral:** todos os resultados possíveis do experimento (aleatório), denotado por $\Omega = \{\omega_1, \omega_2, \dots\}$.
- **Probabilidade:** $P(\omega)$, para cada "ponto amostral" ω .

Exemplos de Espaço amostral

1. Se o fenômeno considerado é observar o sexo de uma criança ao nascer:



$$\Omega = \{F, M\}$$

1. Se o experimento consiste em observar os resultados ao lançar uma moeda duas vezes:

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$$

$$\omega_1 = (C, C); \omega_2 = (C, X); \omega_3 = (X, C); \omega_4 = (X, X)$$

$$C = \text{cara} \quad \text{e} \quad X = \text{coroa}$$

Exemplo: dois dados

Experimento é lançar dois dados (1 vermelho e 1 verde) e anotar os valores:

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

•	•	•	•	•	•	•
•	•	•	•	•	•	•
•	•	•	•	•	•	•
•	•	•	•	•	•	•
•	•	•	•	•	•	•
•	•	•	•	•	•	•
•	•	•	•	•	•	•

Exemplo: testinho surpresa

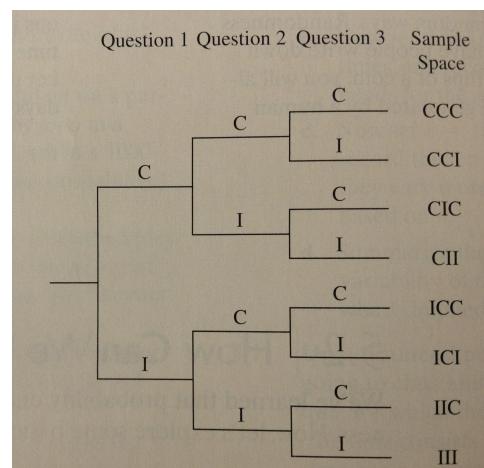
Testinho surpresa com três questões de múltipla escolha.

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

Experimento: anotar o resultado do aluno no testinho.

Ex: *CCI* significa que o aluno acertou as duas primeiras questões e errou a última.

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



Evento

Evento é um subconjunto do espaço amostral. Denotamos eventos pelas letras A, B, C , etc...

Dizemos que o evento A ocorreu sempre que o resultado observado pertencer ao subconjunto de elementos do evento A .

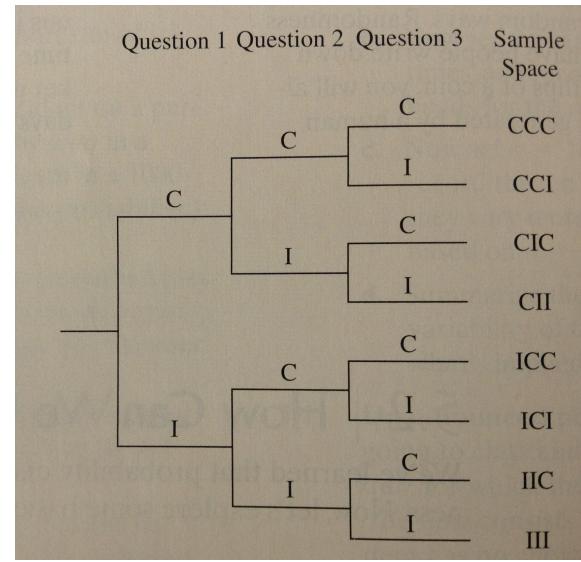
- Experimento é lançar dois dados (1 vermelho e 1 verde) e anotar os valores:

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

Evento: soma dos valores é igual a 3.

$$A = \{(1, 2), (2, 1)\}$$

Exemplo: testinho surpresa



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

Evento: o aluno acertou pelo menos duas questões e foi aprovado.

$$A = \{CCC, CCI, CIC, ICC\}$$

Probabilidade de um evento

Cada elemento do espaço amostral tem uma probabilidade de ocorrer.

Portanto, cada evento (subconjunto do espaço amostral) também tem uma probabilidade.

Duas regras:

- A probabilidade de cada elemento do espaço amostral deve estar entre 0 e 1.
- A soma das probabilidades de cada elemento do espaço amostral deve ser igual a 1.

Exemplo: lançar uma moeda duas vezes

C = cara

X = coroa

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\} = \{(C, C), (C, X), (X, C), (X, X)\}$$

Então $\omega_1 = (C, C)$; $\omega_2 = (C, X)$; $\omega_3 = (X, C)$ e $\omega_4 = (X, X)$.

- Considerando que a moeda é honesta:

$$P(\omega_i) = \frac{1}{4}, \quad \forall i = 1, 2, 3, 4.$$

- Seja o evento $A = \{\omega_1, \omega_4\}$ = obter duas faces iguais:

$$P(A) = P(\{\omega_1, \omega_4\}) = P(\omega_1) + P(\omega_4) = \frac{1}{4} + \frac{1}{4} = \frac{2}{4} = \frac{1}{2}$$

Equiprobabilidade

$\Omega = \{\omega_1, \dots, \omega_n\}$ finito.

Equiprobabilidade: Todos os elementos do espaço amostral tem a mesma probabilidade de acontecer, ou seja,

$$P(\omega_i) = \frac{1}{n}, \quad \forall i = 1, 2, \dots, n$$

Seja $A = \{\omega_{A_1}, \dots, \omega_{A_m}\}$ um evento em Ω com $m \leq n$ pontos amostrais, então

$$P(A) = \frac{m}{n}$$

Probabilidade de um evento

A probabilidade de um evento A , denotada por $P(A)$, é obtida somando as probabilidades de cada elemento do espaço amostral que pertence ao evento A .

Quando cada elemento do espaço amostral tem a mesma probabilidade de ocorrer:

$$P(A) = \frac{\text{número de elementos no evento } A}{\text{número de elementos do espaço amostral}}$$

Exemplos

Exemplo 1: moeda honesta é lançada uma vez $\Omega = \{C, X\}$

$$P(C) = P(X) = \frac{1}{2}$$

$$A = \{C\} \quad \rightarrow \quad P(A) = \frac{1}{2}$$

Exemplo 2: moeda honesta é lançada duas vezes

$$\Omega = \{(C, C), (X, C), (C, X), (X, X)\}$$

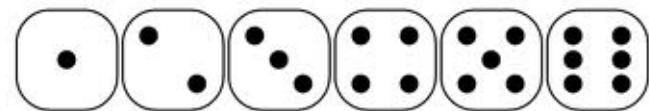
$$P(C, C) = P(X, C) = P(C, X) = P(X, X) = \frac{1}{4}$$

$$A = \{(X, X), (C, C)\} \quad \rightarrow \quad P(A) = \frac{2}{4} = \frac{1}{2}$$

Exemplo: dado honesto é lançado uma vez

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$$

em que ω_i = face i , $\forall i = 1, 2, 3, 4, 5, 6$.



Como o dado é honesto, $P(\omega_i) = \frac{1}{6}$.

Seja o evento $A = \{\text{a face é um número par}\} = \{\omega_2, \omega_4, \omega_6\} = \{2, 4, 6\}$

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

Propriedades: Modelo Teórico

(Ω, P) :

- Ω é o espaço amostral.
- P é probabilidade em Ω .
- Seja A um evento em Ω .
- \emptyset é um conjunto vazio ou evento impossível.

Propriedades:

- $0 \leq P(A) \leq 1, \forall$ evento A em Ω .
- $P(\Omega) = 1$.
- $P(\emptyset) = 0$.

Exemplo: Alunos de um Instituto

Curso	Masculino (M)	Feminino (F)	Total
Matemática Pura (MP)	70	40	110
Matemática Aplicada (A)	15	15	30
Estatística (E)	10	20	30
Computação (C)	20	10	30
Total	115	85	200

Escolhendo um aluno ao acaso:

$$P(MP) = \frac{110}{200} = 0.550$$

$$P(E) = \frac{30}{200} = 0.150$$

$$P(M) = \frac{115}{200} = 0.575$$

$$P(F) = \frac{85}{200} = 0.425$$

Exemplo: Alunos de um Instituto

Curso	Masculino (M)	Feminino (F)	Total
Matemática Pura (MP)	70	40	110
Matemática Aplicada (A)	15	15	30
Estatística (E)	10	20	30
Computação (C)	20	10	30
Total	115	85	200

Seja o evento I : escolher ao acaso um aluno e ele ser estudante de estatística e do sexo masculino.

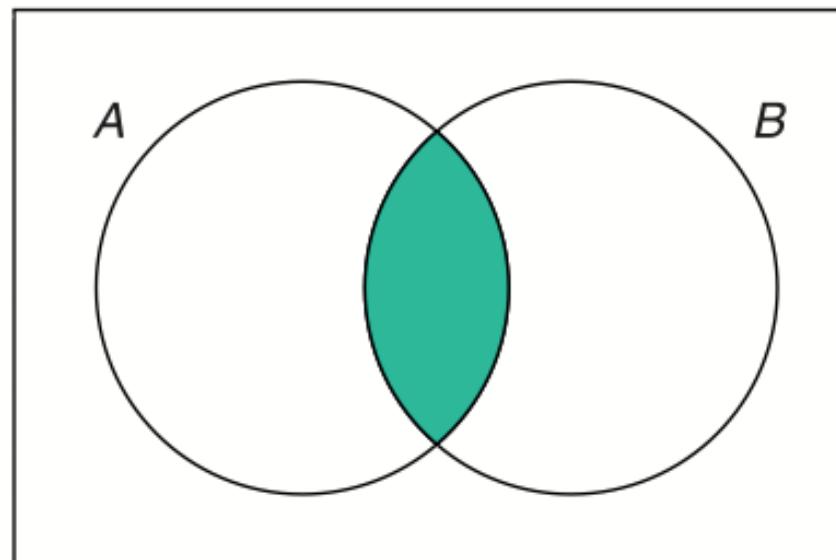
$I = E \cap M$, o evento I é a interseção dos eventos E e M .

$$P(E \cap M) = \frac{10}{200} = 0.05$$

Interseção de Eventos

A interseção de A e B consiste de elementos do espaço amostral que pertencem tanto ao evento A quanto ao evento B .

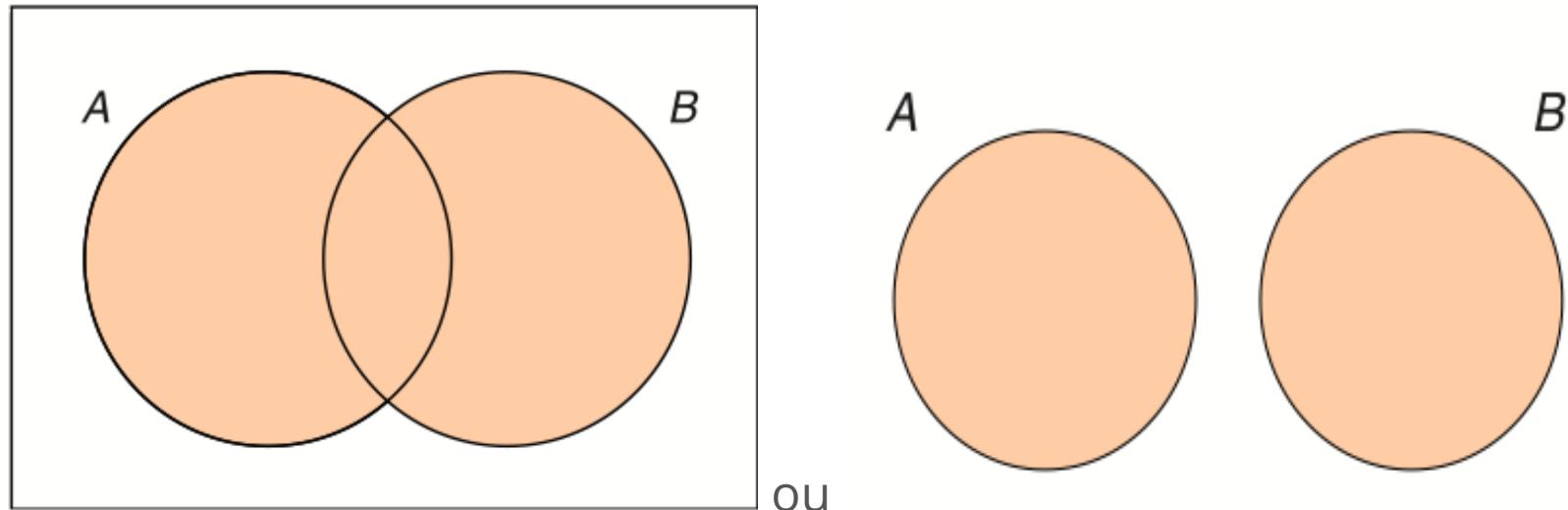
Denotamos $A \cap B$:



União de Eventos

A união de A e B consiste de elementos do espaço amostral que pertencem ao evento A ou ao evento B .

Denotamos $A \cup B$:



Na figura da direita, A e B são denominados **disjuntos**, pois $A \cap B = \emptyset$.

Probabilidade de União de Eventos

$A \cup B$ contém elementos dos eventos A ou B .

Para calcular a probabilidade de $A \cup B$, podemos então somar a probabilidade de A ocorrer e a probabilidade de B ocorrer.

Problema: ao fazer isso, estamos somando a probabilidade de $A \cap B$ duas vezes.

Forma correta: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= \text{brown circle} + \text{red circle} - \text{overlap} \\ &= \text{non-overlapping red circle} + \text{non-overlapping blue circle} \end{aligned}$$

Exemplo: Alunos de um Instituto

Curso	Masculino (M)	Feminino (F)	Total
Matemática Pura (MP)	70	40	110
Matemática Aplicada (A)	15	15	30
Estatística (E)	10	20	30
Computação (C)	20	10	30
Total	115	85	200

Seja U o evento: escolher um aluno ao acaso e ele ser estudante de estatística ou do sexo masculino.

$U = E \cup M$, o evento U é uma união dos eventos E e M .

$$P(E \cup M) = P(E) + P(M) - P(E \cap M)$$

$$P(E) = \frac{30}{200} = 0.150 \quad P(M) = \frac{115}{200} = 0.575 \quad P(E \cap M) = \frac{10}{200} = 0.050$$

$$P(E \cup M) = 0.150 + 0.575 - 0.050 = 0.675$$

Exemplo: Alunos de um Instituto

Curso	Masculino (M)	Feminino (F)	Total
Matemática Pura (MP)	70	40	110
Matemática Aplicada (A)	15	15	30
Estatística (E)	10	20	30
Computação (C)	20	10	30
Total	115	85	200

No caso de eventos mutuamente exclusivos ou disjuntos, a interseção é vazia (\emptyset).

Probabilidade de escolher um aluno ao acaso e ele ser estudante da matemática pura **e** da computação:

$$P(MP \cap C) = P(\emptyset) = 0$$

Probabilidade de escolher um aluno ao acaso e ele ser estudante da matemática pura **ou** da computação:

$$P(MP \cup C) = P(MP) + P(C) - P(MP \cap C) = P(MP) + P(C) = \frac{140}{200} = 0.700$$

Exemplo: Qual a chance de cair na malha fina?

80.2 milhões de declarações (2002).

(para simplificar, uma frequência de 90 representa 90.000).

Renda	Caiu na malha fina	Não caiu na malha fina	Total
D - abaixo de 25000	90	14010	14100
C - 25000 a 49999	71	30629	30700
B - 50000 a 99999	69	24631	24700
A - acima de 100000	80	10620	10700
Total	310	79890	80200

Qual o espaço amostral?

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

É um fenômeno equiprovável?

Exemplo: Qual a chance de cair na malha fina?

Renda	Caiu na malha fina	Não caiu na malha fina	Total
D - abaixo de 25000	90	14010	14100
C - 25000 a 49999	71	30629	30700
B - 50000 a 99999	69	24631	24700
A - acima de 100000	80	10620	10700
Total	310	79890	80200

Se escolhermos uma declaração de 2002 aleatoriamente, qual a probabilidade dela ter caído na malha fina (evento Z)?

$$Z = \{(A, \text{sim}), (B, \text{sim}), (C, \text{sim}), (D, \text{sim})\} \implies P(Z) = \frac{310}{80200} = 0.004$$

Qual a probabilidade dela ter renda acima de 100.000 (evento Y)?

$$Y = \{(A, \text{sim}), (A, \text{não})\} \implies P(Y) = \frac{10700}{80200} = 0.133$$

Exemplo: Qual a chance de cair na malha fina?

Renda	Caiu na malha fina	Não caiu na malha fina	Total
D - abaixo de 25000	90	14010	14100
C - 25000 a 49999	71	30629	30700
B - 50000 a 99999	69	24631	24700
A - acima de 100000	80	10620	10700
Total	310	79890	80200

Se escolhermos uma declaração de 2002 aleatoriamente, qual a probabilidade dela ter renda acima de 100.000 e ter caído na malha fina (evento W)?

$$W = Z \cap Y = \{(A, \text{sim})\}$$

$$P(W) = P(Z \cap Y) = \frac{80}{80200} = 0.001$$

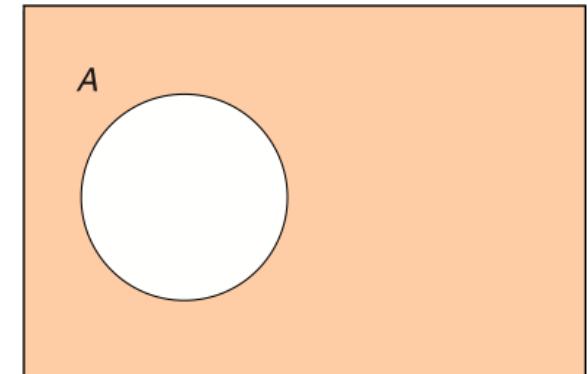
Evento complementar

No caso geral, sejam A e B subconjuntos de Ω :

- $A \cap B$ = evento em que A e B ocorrem simultaneamente.
- $A \cup B$ = evento em que A ou B ocorrem.
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
- se $\{A \cap B\} = \emptyset$, então $P(A \cup B) = P(A) + P(B)$.

Dois eventos A e B são **complementares** se $A \cap B = \emptyset$ e $A \cup B = \Omega$.

- como $P(A) + P(B) = 1$, então $P(B) = 1 - P(A)$
- B é denotado por $B = A^C$ (indicado na cor laranja).



Exemplo: Cartões de Crédito

Um estabelecimento aceita Visa ou Mastercard. Dentre os clientes, 22% possuem Mastercard, 58% possuem Visa e 14% possuem os dois.

Qual a probabilidade de que um cliente tenha pelo menos um destes cartões?



Espaço amostral: $\Omega = \{V, M, VM, N\}$, onde V="tem só Visa", M="tem só Mastercard", VM="tem Visa e Mastercard", N="não tem Visa nem Mastercard".

Evento A: cliente possui Mastercard. $A = \{M, VM\}$

Evento B: cliente possui Visa. $B = \{V, VM\}$

$$P(A) = 0.22$$

$$P(B) = 0.58$$

$$P(A \cap B) = 0.14$$

Exemplo: Cartões de Crédito

$A \cup B$: cliente possui pelo menos um dos cartões. $A \cup B = \{V, M, VM\}$.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.22 + 0.58 - 0.14 = 0.66.$$

Evento C: cliente não possui nenhum dos cartões. $C = \{N\}$

C é complementar de $A \cup B$, pois:

$$C \cup (A \cup B) = \Omega \text{ e } C \cap (A \cup B) = \emptyset.$$

Então $P(C) = 1 - P(A \cup B) = 0.34$.



Exemplo: Cinto de segurança e acidentes

Usava cinto	Sobreviveu (S)	Não sobreviveu (\bar{S})	Total
Sim (C)	414368	510	414878
Não (\bar{C})	162527	1601	164128
Total	576895	2111	579006

$$\Omega = \{(C, S), (C, \bar{S}), (\bar{C}, S), (\bar{C}, \bar{S})\}$$

Se selecionarmos um registro ao acaso, qual a probabilidade dele conter uma morte registrada?

$$P(\bar{S}) = \frac{2111}{579006} = 0.0034$$

Exemplo: Cinto de segurança e acidentes

Usava cinto	Sobreviveu (S)	Não sobreviveu (\bar{S})	Total
Sim (C)	414368	510	414878
Não (\bar{C})	162527	1601	164128
Total	576895	2111	579006

$$\Omega = \{(C, S), (C, \bar{S}), (\bar{C}, S), (\bar{C}, \bar{S})\}$$

Se selecionarmos um registro ao acaso, qual a probabilidade de constar que o cinto não foi usado?

$$P(\bar{C}) = \frac{164128}{579006} = 0.283$$

Leituras

- [OpenIntro](#): seção 2.1.
- [Ross](#): seções 4.1, 4.2, 4.3, 4.4.

Slides produzidos pelos professores:

- Samara Kiihl
- Tatiana Benaglia
- Benilton Carvalho