Ciência dos dados

Teoria da Probabilidade

Insper

Aula de hoje

Ao final desta aula, o aluno deve ser capaz de:

- Traduzir informações descritas em problemas práticos fazendo uso da teoria da probabilidade.
- Aplicar notação de Probabilidade Condicional dentro de contextos práticos.

Conceitos básicos de probabilidade

Denomina-se **fenômeno** (ou experimento) **aleatório** à situação ou acontecimento cujos resultados não podem ser previstos com certeza.

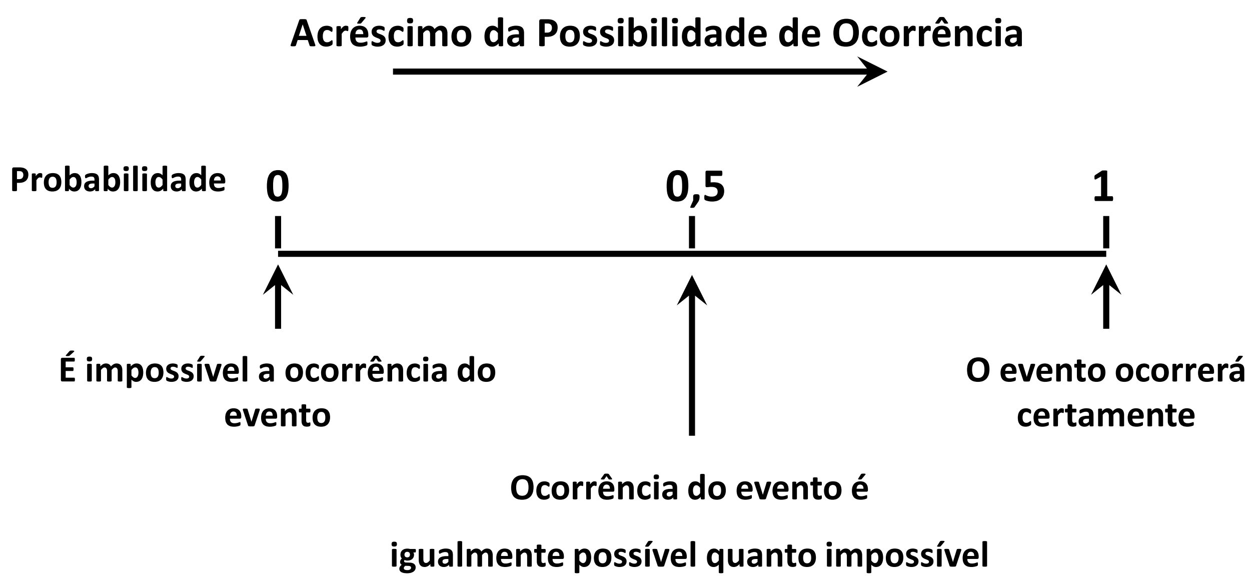
A teoria das probabilidades estabelece modelos matemáticos para os fenômenos aleatórios, tais como:

- ✓ Qual probabilidade do robô virar para direita?
- ✓ Qual probabilidade do projeto terminar no prazo?
- ✓ Qual probabilidade da nova bateria durar mais do 48 horas?

Nestes casos, modelos podem ser estabelecidos para quantificar as incertezas das diversas ocorrências.

Conceitos básicos de probabilidade

Probabilidade: medida da incerteza da ocorrência de um evento



Os principais métodos de estabelecer probabilidades aos eventos analisados são os métodos clássico, frequentista e subjetivo.

Conceitos básicos de probabilidade

- Seja Ω o espaço amostral (evento certo), então $P(\Omega) = 1$

- Seja A um evento qualquer pertencente a Ω, então $0 \leq P(A) \leq 1$

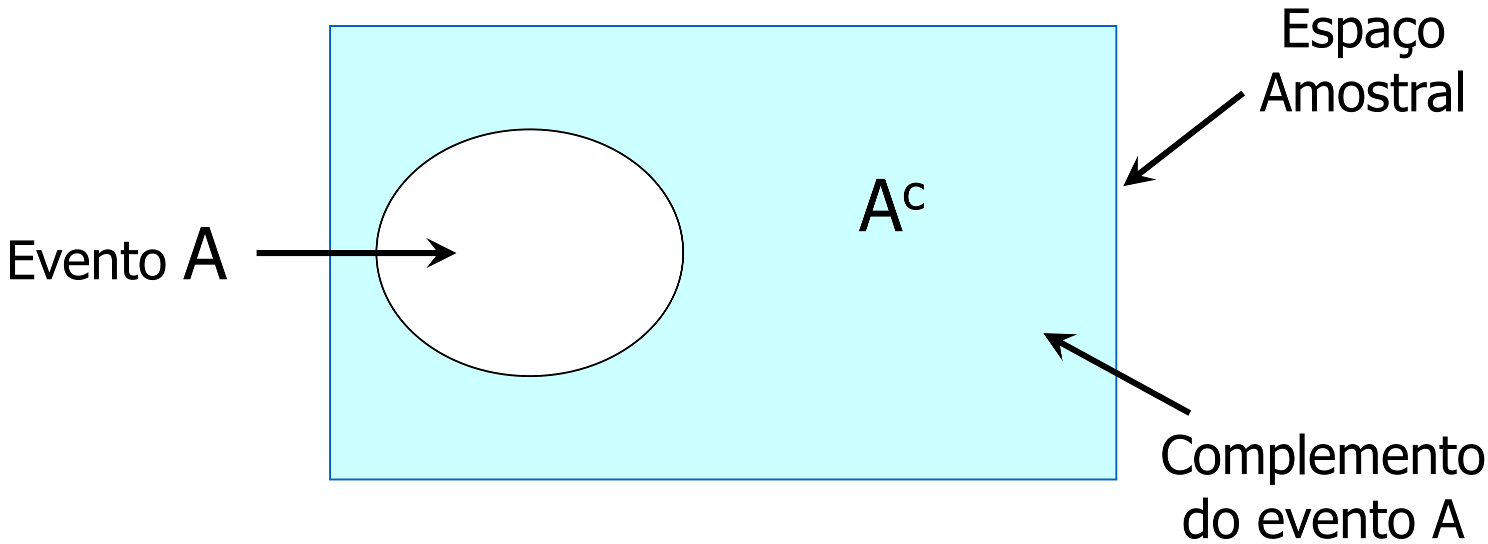
Vamos relembrar probabilidade?

www. socrative.com

LOGIN → STUDENT LOGIN

NOME DA SALA: INSPER

Definição:
Evento Complementar

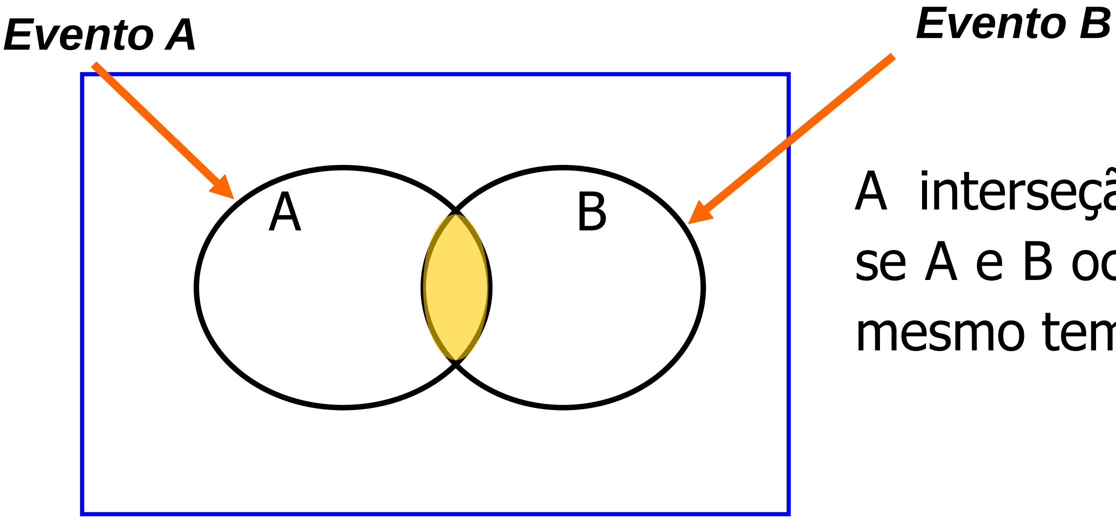


$$P(A^c) = 1 - P(A)$$

Definição:

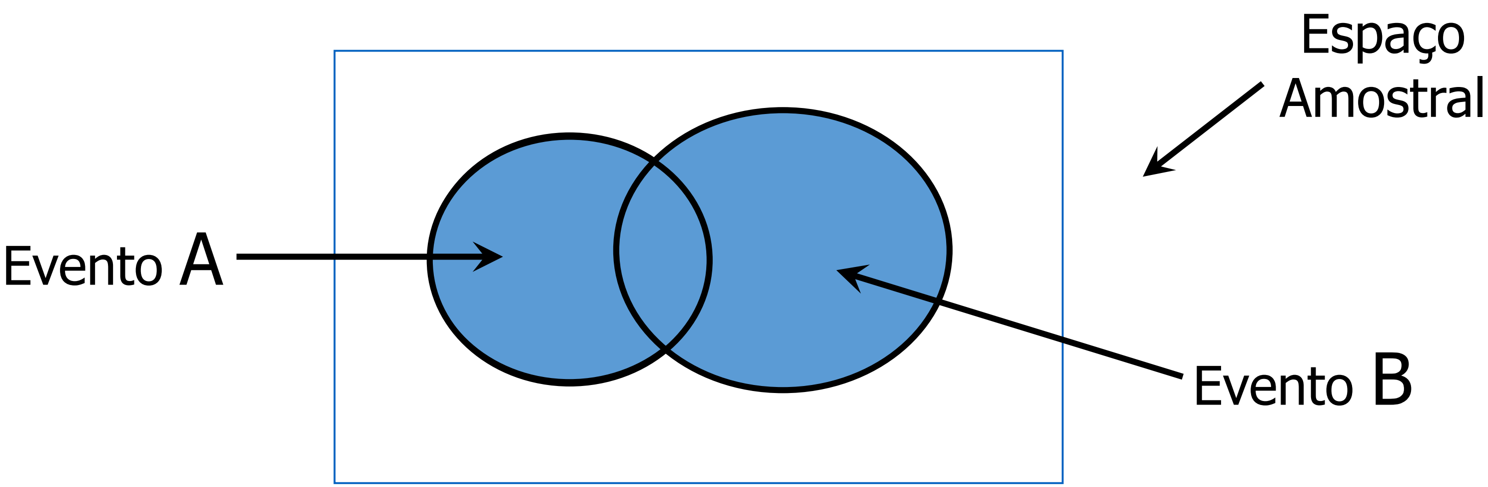
Interseção de 2 eventos

$\{A \cap B\}$: conjunto dos pontos do espaço amostral que pertencem, simultaneamente, aos eventos A e B.



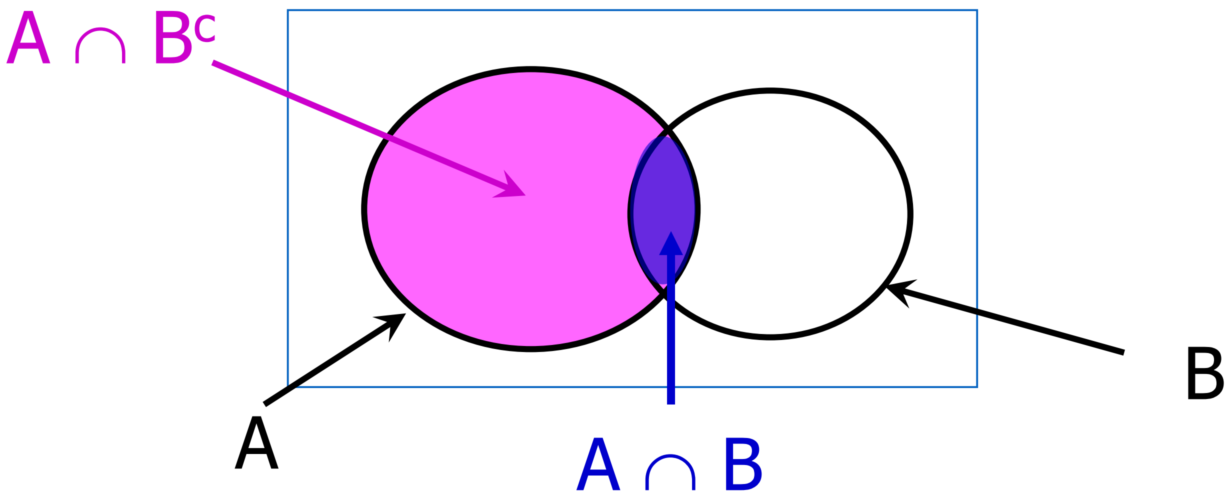
A interseção ocorre se A e B ocorrem ao mesmo tempo.

Definição:
União de 2 eventos



$$P(A \text{ ou } B) =$$
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

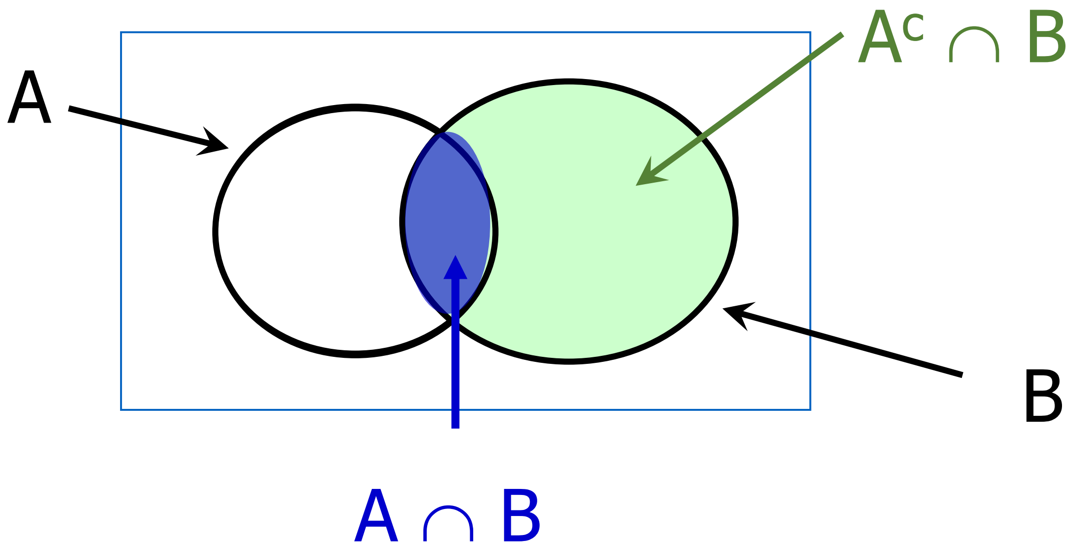
Definição:
Probabilidade marginal



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

Definição:

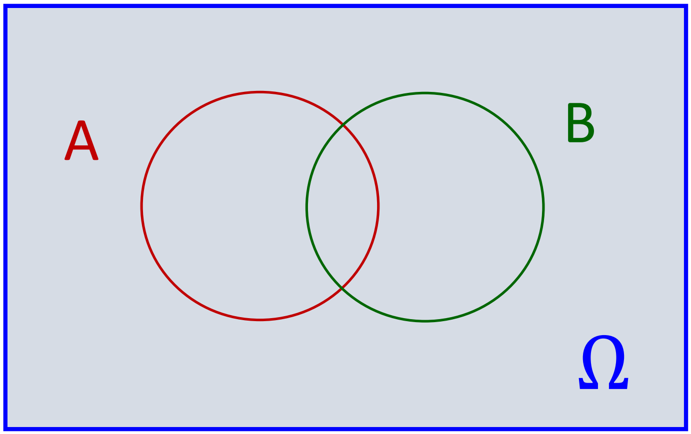
Probabilidade marginal

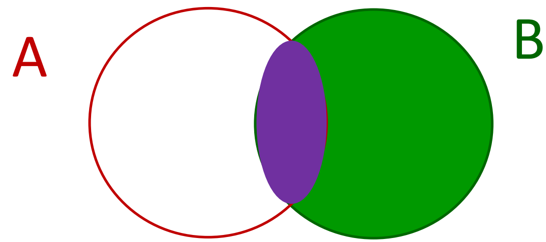


$$P(B) = P(A \cap B) + P(A^c \cap B)$$

Definição:
Probabilidade condicional

$P(A|B) \Rightarrow$ probabilidade de A ocorrer sabendo (dado) que B ocorreu





B faz o papel de um novo espaço amostral → tamanho $P(B)$

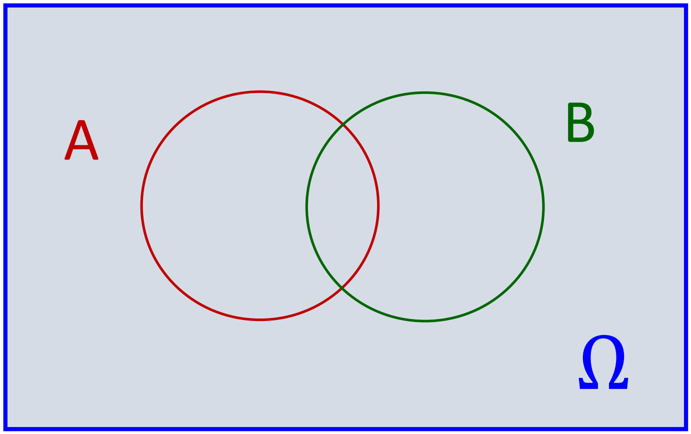
$A \cap B$: é o que sobrou de A no novo espaço amostral → tamanho $P(A \cap B)$

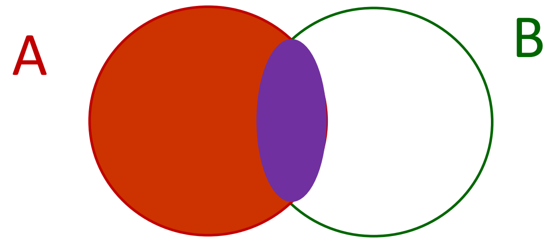
$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Definição:

Probabilidade condicional

$P(B|A) \Rightarrow$ probabilidade de B ocorrer sabendo (dado) que A ocorreu





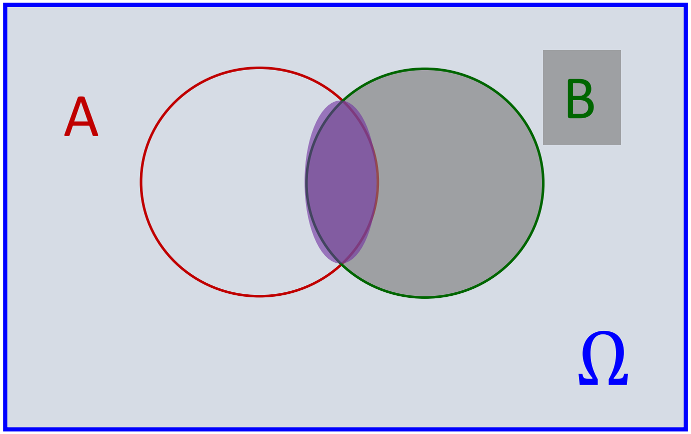
A faz o papel de um novo espaço amostral → tamanho $P(A)$

$A \cap B$: é o que sobrou de A no novo espaço amostral → tamanho $P(A \cap B)$

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

Definição:
Probabilidade para Intersecção

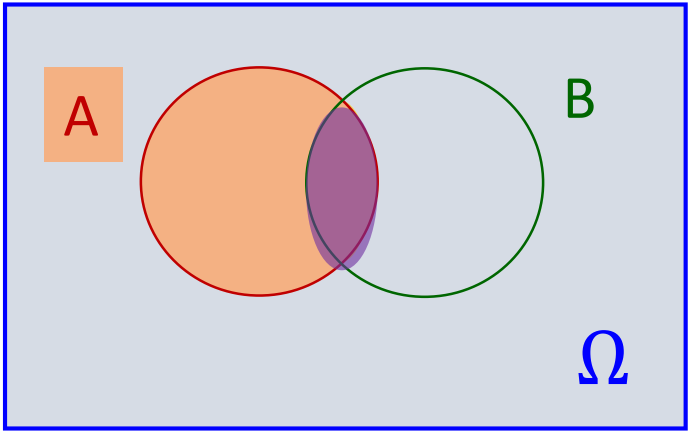
$$P(A|\mathbf{B}) = \frac{P(A \cap B)}{P(B)} \quad \Rightarrow \boldsymbol{P(A \cap B) = P(B)P(A|B)}$$



Definição:

Probabilidade para Intersecção

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \qquad \Rightarrow \boldsymbol{P(A \cap B) = P(A)P(B|A)}$$



Definição:

Independência

Os eventos A e B são independentes quando o fato de ter conhecimento sobre a ocorrência de A **não altera** a expectativa sobre a probabilidade de ocorrência do evento B.

Caso particular:
Eventos independentes

Dois eventos A e B quaisquer contidos ao mesmo espaço amostral são independentes quando :

$$P(A|B) = P(A) \quad \text{ou} \quad P(B|A) = P(B) \qquad \text{(I)}$$

De maneira geral, temos:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \Rightarrow P(A \cap B) = P(A|B)\,P(B) \qquad \text{(II)}$$

Se A e B forem independentes, substituindo (I) em (II) temos:

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

Caso particular:
Eventos independentes

Dois eventos A e B quaisquer contidos ao mesmo espaço amostral são **independentes** quando

$$\boldsymbol{P(A|B) = P(A)}$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \Rightarrow P(A \cap B) = \boldsymbol{P(A|B)}\, P(B)$$

Assim, se A e B forem **independentes** temos

$$P(A \cap B) = \boldsymbol{P(A)}\, P(B)$$

Exemplo da Netflix

A Netflix considera diversas variáveis antes de sugerir uma lista de filmes a um usuário. Entre elas, considere o período que usuário acessa a conta (Diurno e Noturno) e, por consequência, o gênero sugerido ao usuário (Romance e Ação).

A tabela abaixo apresenta as **frequências absolutas (contagens)** que relacionam essas duas variáveis.

Período	Gênero		Total
	Romance	Ação	
Diurno	35	15	50
Noturno	45	105	150
Total	80	120	200

Exemplo da Netflix

Distribuição conjunta: avaliação do comportamento conjunto de 2 variáveis.

Distribuição marginal: avaliação do comportamento individual de cada uma das variáveis.

Período	Gênero		Total
	Romance	Ação	
Diurno	35	15	50
Noturno	45	105	150
Total	80	120	200

Distribuição conjunta das variáveis

Distribuição marginal das variáveis

Exemplo da Netflix

Distribuição conjunta: avaliação do comportamento conjunto de 2 variáveis.

Distribuição marginal: avaliação do comportamento individual de cada variável.

Período	Gênero: Romance	Gênero: Ação	Total
Diurno	0,175	0,075	0,250
Noturno	0,225	0,525	0,750
Total	0,400	0,600	1

Distribuição conjunta entre as variáveis

Distribuição marginal de cada variável

Independentes *vs* Disjuntos

Atenção: não confundir eventos **independentes** com eventos **disjuntos (mutuamente exclusivos)**.

$$\mathbf{P(A \cap B)} = \boldsymbol{P(A)P(B)} \Rightarrow \textbf{Independentes}$$

$$\mathbf{P(A \cap B) = 0} \Rightarrow \textbf{Disjuntos}$$

EXERCÍCIOS

Exercício 1

A probabilidade de que o preço dos combustíveis aumente no mês vindouro é estimada em 0,4.

Se isto ocorrer, a probabilidade de que os preços dos transportes coletivos também aumentem é de 0,5; caso contrário, esta probabilidade é de 0,1.

Se naquele mês o preço das passagens de fato subirem, qual a probabilidade de os preços dos combustíveis não terem sofrido majoração? Resposta: 0,231

Exercício 2

Uma revendedora de veículos trabalha com duas marcas de automóveis: A e B. 70% dos que adquirem carros populares, escolhem a marca A. Dentre os que adquirem carros não populares, 80% compram A. Sabe-se que 60% das vendas são de carros populares.

a. Qual é a probabilidade de um consumidor comprar um carro da marca A?

b. Sabendo que uma pessoa comprou um carro da marca A, qual é a probabilidade de ter sido um carro popular?

Exercício 3 – Montgomery e Runger

Em uma operação de enchimento automático, a probabilidade de um enchimento incorreto será 0,001, quando o processo for operado em baixa velocidade. Quando o processo for operado em alta velocidade, a probabilidade de um enchimento incorreto será de 0,01.

Suponha de 30% dos reservatórios sejam cheios quando o processo for operado em alta velocidade, e o restante seja cheio em baixa velocidade.

a) Qual a probabilidade de um reservatório ser cheio incorretamente?

b) Se um reservatório cheio incorretamente for encontrado, qual é a probabilidade de que ele tenha sido cheio durante uma operação em alta velocidade?

Exercício 4

O gerente de um posto de gasolina sabe da sua experiência que 80% dos clientes usam cartão de crédito quando compram gasolina.

Admita independência entre clientes.

Qual é a probabilidade dos dois próximos clientes comprarem gasolina usando cartão de crédito ?

E de pelo menos 1 usar cartão de crédito?

A = evento de que o 1o. cliente use CC

B = evento de que o 2o. cliente use CC

Resposta: 0,64 e 0,96, respectivamente.

Exercício 5 – Montgomery e Runger

Um sistema de codificação-decodificação consiste em três elementos: codifica, transmite e decodifica.

Uma codificação falha ocorre em 0,5% das mensagens processadas; erro de transmissão ocorrem em 1% das mensagens; e erro de decodificação ocorre em 0,1% das mensagens. Considere os erros como independentes.

a) Qual a probabilidade de se ter uma mensagem completamente livre de defeito?

b) Qual a probabilidade de uma mensagem ter tanto um defeito de codificação como de decodificação?

Exercício 6

Considerando apenas os países que ganharam pelo menos uma medalha nos Jogos Olímpicos de 2012, as seguintes probabilidades foram observadas ao analisar, conjuntamente, o índice de desenvolvimento humano (IDH ≥ 0,75 ou IDH < 0,75) de 2011 e o tipo de medalha olímpica (ouro, prata ou bronze).

- ✓ A probabilidade de um atleta, que pertence a um destes países, ganhar uma medalha de ouro é de 31,5%.
- ✓ Se um país tem IDH igual ou superior a 0,75, então a probabilidade de um atleta ganhar uma medalha de ouro é 35,0% e, na mesma condição, a de um atleta ganhar uma medalha de prata é de 32,9%.
- ✓ Entre países com IDH inferior a 0,75, a probabilidade de um atleta ganhar uma medalha de ouro é 23,3%.
- ✓ Por fim, se um atleta ganhou uma medalha de bronze, a probabilidade de que ele pertença a um país com IDH inferior a 0,75 é de 38,4%.

Considerando apenas os países que ganharam pelo menos uma medalha nos Jogos Olímpicos de 2012, responda:

a) Qual a probabilidade de um atleta ser de um país com IDH igual ou superior a 0,75? 0,70085

b) Qual a probabilidade de uma medalha de prata pertencer a um atleta que representa um país com IDH igual ou superior a 0,75? 0,72106

c) Com base nos resultados mencionados que consideram apenas países que ganharam pelo menos uma medalha nas Olimpíadas de Londres, o IDH de um destes países pode influenciar o tipo de medalha que um atleta pode ganhar? Justifique sua resposta considerando **informações numéricas**.

Exercício 7

O texto a seguir, extraído da internet, fala sobre a medalha de ouro do futebol masculino nas Olimpíadas Rio 2016:


MENU SPORTV RIO 2016

21/08/2016 14h05 - Atualizado em 21/08/2016 14h05

Weverton dedica título da Rio 2016 a Prass: "Ele faz parte da história"

Goleiro do Atlético-PR, que entrou na vaga do jogador do Palmeiras, diz que não esperava participar da Olimpíada do Rio de Janeiro: "Nem imaginava estar aqui agora"

"Um dos destaques na final contra a Alemanha, que terminou com vitória dos donos da casa nos pênaltis, e substituto de Prass na equipe comandada por Rogério Micale, o goleiro Weverton dedicou o ouro nos Jogos do Rio de Janeiro ao titular do Palmeiras.

Em entrevista ao "SporTV", o camisa 1 disse não esperava estar na campanha dos Jogos Olímpicos, apesar de acreditar em uma convocação para a Seleção. Weverton brilhou nas cobranças de pênaltis ao defender a finalização de Petersen, a quinta da Alemanha. O goleiro disse que recebeu um relatório da comissão técnica da CBF sobre como o camisa 18 chutava.

- *Antes do jogo, o treinador de goleiros, junto com o pessoal da análise CBF, passou alguns batedores. Eu tinha o número 11 e o 18, que foi o pênalti que eu peguei. O Petersen tinha quatro batidas no lado direito e quatro no esquerdo. Só que, ao analisar mais friamente, detectamos algumas coisas que foram interessantes para a defesa. Quando ele estava mais relaxado no jogo, ele batia mais no lado direito do goleiro, mas quando era uma partida mais nervosa, ele quase sempre invertia e batia forte. Fiz essa leitura e consegui fazer a defesa.*"

Adaptado de fonte: http://sportv.globo.com/site/programas/rio-2016/noticia/2016/08/weverton-dedica-titulo-da-rio-2016-prass-ele-faz-parte-da-historia.html

Exercício 7

Considere que, de todas as batidas de pênalti feitas pelo jogador Petersen da Alemanha, tenham os seguintes resultados:

- 50% dessas batidas foram no lado direito do goleiro;
- Das batidas feitas no lado direito do goleiro, 85,0% foram em partidas que se sentiu mais relaxado;
- Das batidas feitas em partidas relaxadas, 91,9% foram feitas no lado direito do goleiro;
- Assuma que "partida relaxada" e "partida nervosa" sejam eventos complementares.
- Assuma que "lado direito do goleiro" e "lado esquerdo do goleiro" sejam eventos complementares.

Responda cada item abaixo, **definindo os eventos necessários e considerando, na sua resolução, as notações vistas em aula.** Utilize 3 casas decimais na resolução dos itens.

a) Calcule a probabilidade das batidas de pênalti terem acontecido em partidas relaxadas.

b) Das batidas de pênalti feitas em partidas nervosas, calcule a probabilidade de terem acontecido no lado esquerdo do goleiro. Ainda, por que esse resultado poderia ajudar Weverton a fazer a defesa do pênalti nas Olimpíadas Rio 2016?

Aviso

- **Fazer também os exercícios do Capítulo 2 do Montgomery e Runger (2016).**

Exercício

- Download do notebook pelo Github
- Fazer individual e discutir em grupo
- Tem APS3 no Blackboard