EAH5002 – Planejamento de Experimentos e Análise Estatística de Dados

Revisão Teórica: 02 – Probabilidade Condicional

Marcelo de Souza Lauretto
Sistemas de Informação – EACH
www.each.usp.br/lauretto

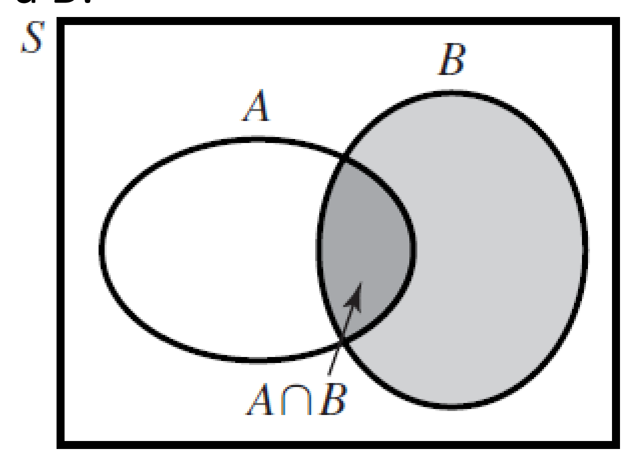
Referência:
Morris DeGroot, Mark Schervish.
Probability and Statistics. 4th Ed. – 2º capítulo

2.1 Definição de probabilidade condicional

- Exemplo: Bilhete da Mega-Sena
  - Considere o jogo da Mega-Sena, em que seis números são sorteados sem reposição dentre os números 1 – 60. O objetivo é acertar o conjunto de seis números sorteados, independentemente da ordem de sorteio. Suponha que você tem um bilhete com os números 3,15,20,35,41,58.
  - Você liga a televisão para assistir ao sorteio mas você só vê o sorteio do 1º número, 15, quando a energia cai subitamente.
  - Agora que você sabe que o número 15 aparece no sorteio vencedor, a probabilidade de seu bilhete ser um ganhador deve ser maior do que era antes de você ter visto o sorteio.
  - Como atualizar a probabilidade de seu bilhete ser ganhador?

2.1 Definição de probabilidade condicional

- Esse exemplo é típico da seguinte situação:
  - Considere que um experimento deve ser realizado; seu espaço amostral é conhecido, assim como as probabilidades de ocorrência para todos os eventos de interesse.
  - Nosso interesse é calcular a probabilidade de um evento A.
  - Durante a realização do experimento, tomamos conhecimento de que outro evento B ocorreu, e queremos saber como a probabilidade de A se altera com essa nova informação.
  - Se sabemos que o evento B ocorreu, então sabemos que o resultado do experimento é um daqueles pertencentes a B.
  - Logo, para calcular a probabilidade de que A ocorrerá, precisamos considerar o conjunto dos resultados que também implicam na ocorrência de B, ou seja, nos resultados pertencentes a $A \cap B$.



2.1 Definição de probabilidade condicional

- **Definição 2.1.1: Probabilidade condicional**
  - Suponha que aprendemos que um evento *B* ocorreu e queremos calcular a probabilidade de outro evento *A* levando em conta que sabemos da ocorrência de *B*.
  - A nova probabilidade de *A* é denominada a **probabilidade condicional de que o evento *A* ocorreu dado que o evento *B* tenha ocorrido**. Notação: $\Pr(A|B)$.
  - Se $\Pr(B) > 0$, calculamos essa probabilidade como

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

  - A probabilidade condicional $\Pr(A|B)$ não é definida quando $\Pr(B) = 0$.

2.1 Definição de probabilidade condicional

- Voltando ao exemplo do bilhete da Mega-Sena
  - O evento de que tomamos conhecimento foi:
    B = {um dos números sorteados é 15}
  - O evento cuja probabilidade queremos calcular é:
    A = {os números 3,15,20,35,41,58 foram sorteados}.
  - O espaço amostral *S* consiste de todos os sorteios de 6 dentre 60 números (independentemente da ordem de sorteio):

$$|S| = \binom{60}{6} = \frac{60!}{6!\,54!} = 50.063.860$$

  - Evento B consiste das combinações que incluem 15, e portanto

$$\Pr(B) = \frac{|B|}{|S|} = \frac{\binom{59}{5}}{\binom{60}{6}} = \frac{59!\,6!\,54!}{60!\,5!\,54!} = \frac{6}{60} \simeq 0{,}1$$

2.1 Definição de probabilidade condicional

- Voltando ao exemplo do bilhete da Mega-Sena
  - O evento A de que seu bilhete é vencedor consiste de um único resultado que também está em B, de forma que $A \cap B = A$, e
$$\Pr(A \cap B) = \Pr(A) = \frac{1}{\binom{60}{6}} = \frac{1}{50.063.860} \simeq 2 \times 10^{-8}$$
  - Segue que a probabilidade condicional de *A* dado *B* é:
$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} \simeq \frac{2 \times 10^{-8}}{10^{-1}} \simeq 2 \times 10^{-7}$$
  - Note que, embora $\Pr(A|B)$ seja ainda muito baixa, é uma ordem de grandeza superior a $\Pr(A)$ antes de você saber que *B* havia ocorrido.

2.1 Definição de probabilidade condicional

- A probabilidade condicional possui tanto uma interpretação subjetiva como também frequentista:
  - Interpretação subjetiva: $\Pr(A|B)$ é a probabilidade atribuída ao evento *A* após o conhecimento de que o evento *B* ocorreu.
  - Interpretação frequentista: Se um processo experimental é repetido um número grande de vezes, então a proporção de repetições nos quais o evento B ocorrerá será aproximadamente $\Pr(B)$, e a proporção de repetições nas quais ambos os eventos A e B ocorrerão será aproximadamente $\Pr(A \cap B)$.
    Portanto, entre as repetições nas quais o evento B ocorre, a proporção de repetições em que A também ocorrerá é aproximadamente igual a

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

2.1 Definição de probabilidade condicional

- Exemplo 2.1.4: Ensaio clínico
  - Prien et al. (1984) estudaram 3 tratamentos contra reincidência de depressão: imipramina, carbonato de lítio e uma combinação.
  - Para tal, conduziram um ensaio clínico com quatro grupos, dos quais três receberam um dos tratamentos e o 4º grupo recebeu um placebo.
  - Neste exemplo, consideramos 150 pacientes (que entraram após um episódio de depressão), divididos nos quatro grupos.

Table 2.1 Results of the clinical depression study in Example 2.1.4

	Treatment group				
Response	Imipramine	Lithium	Combination	Placebo	*Total*
Relapse	18	13	22	24	77
No relapse	22	25	16	10	73
Total	40	38	38	34	150

2.1 Definição de probabilidade condicional

- Exemplo 2.1.4: Ensaio clínico
  - Seja *P* o evento de que o paciente recebeu o placebo, e seja *R* o evento de que o paciente teve uma reincidência. Podemos *estimar* $\Pr(P) = 34/150$ e $\Pr(R \cap P) = 24/150$ diretamente da tabela. Então $\Pr(R|P) = 24/34 = 0.706$.
  - Por outro lado, se o paciente tiver recebido lítio (chamemos evento *L*), então temos as estimativas $\Pr(L) = 38/150$, $\Pr(R \cap L) = 13/150$ e $\Pr(R|L) = 13/38 = 0.342$
  - Ou seja, conhecer o tratamento que o paciente recebeu parece fazer diferença na probabilidade de reincidência.
  - Futuramente neste curso estudaremos *testes de hipóteses*, que servem para avaliar se essa diferença observada é estatisticamente significante ou se é meramente fruto do acaso na amostra particular que foi coletada.

2.1 Definição de probabilidade condicional

- Exemplo: A caixa de Bertrand:
  - Você passou a vida toda procurando as Pérolas Gêmeas (um tesouro arqueológico muito valioso) e finalmente as encontrou, porém com uma armadilha:
  - Há 3 porta-joias fechados, cada qual com duas gavetas, e você sabe que as Pérolas Gêmeas estão em um deles.
  - Quando você força a abertura de uma das gavetas, encontra o que parece ser uma das Pérolas Gêmeas, junto com o seguinte bilhete:

Fonte: Jeremy Stangroom. O Enigma de Einstein. Ed. Marco Zero, 2010.

2.1 Definição de probabilidade condicional

- Exemplo: A caixa de Bertrand (cont)

Caro Grande Aventureiro,

À sua frente há três porta-joias. Um deles contém as Pérolas Gêmeas, uma em cada gaveta. Outro contém uma pérola comum numa gaveta e um pedaço de carvão na outra. O último contém dois pedaços de carvão, um em cada gaveta. Infelizmente, não há como distinguir as três pérolas: você só vai saber que está com as genuínas Pérolas Gêmeas pelo fato de estarem no mesmo porta-joias. E mais uma última coisa: você só poderá abrir mais uma gaveta. Se errar, os três porta-joias se autodestruirão.

2.1 Definição de probabilidade condicional

- Exemplo: A caixa de Bertrand (cont):
  - Qual a probabilidade de você encontrar outra pérola na 2ª gaveta do mesmo porta-joias?
    - Refraseando: Qual a probabilidade de que a 1ª gaveta que você abriu era do porta-joias correto?
  - Chamemos os três porta-joias de PP, PC e CC, de acordo com seu conteúdo (2 pérolas, 1 pérola e 1 carvão, 2 carvões).
  - O que queremos saber é a probabilidade de nosso porta-joias ser o PP.
  - Resposta: via probabilidade condicional:
    - Evento PP: o porta-joias aberto é o PP
    - Evento P: a gaveta aberta contém uma pérola.
    - Note que $PP \cap P = PP$, e portanto $\Pr(PP \cap P) = \Pr(PP) = 1/3$

$$\Pr(PP|P) = \frac{\Pr(PP \cap P)}{Pr(P)} = \frac{1/3}{1/2} = 2/3$$

  - Exercício: Calcular as probabilidades $\Pr(PC|P)$ e $\Pr(CC|P)$

2.1 Probabilidade condicional – Regra da multiplicação

- Em alguns experimentos, certas probabilidades condicionais são relativamente simples de atribuir diretamente. Nesses casos, é possível calcular a probabilidade de que os dois eventos ocorram, por inversão direta da equação que define as probabilidades condicionais.

- **Teorema 2.1.1: Regra da multiplicação para probabilidades condicionais (dois eventos)**
  Sejam A e B eventos.
  Se $\Pr(B) > 0$, então
  $$\Pr(A \cap B) = \Pr(B)\Pr(A|B).$$
  Se $\Pr(A) > 0$, então
  $$\Pr(A \cap B) = \Pr(A)\Pr(B|A).$$

2.1 Probabilidade condicional – Regra da multiplicação

- Exemplo: Seleção de duas bolas
  - Suponha que duas bolas devem ser selecionadas aleatoriamente, sem reposição, de uma caixa contendo *r* bolas vermelhas e *b* bolas azuis.
  - Devemos determinar a probabilidade *p* de que a primeira bola será vermelha e a segunda bola será azul.
  - Resp. Seja A o evento de que a primeira bola é vermelha, e B o evento de que a segunda bola é azul. Claramente, $\Pr(A) = r/(r+b)$. Além disso, se o evento A ocorreu, então uma bola vermelha foi removida da caixa. Logo, a probabilidade de obter uma bola azul no segundo sorteio será:

$$\Pr(B|A) = \frac{b}{r+b-1}.$$

Segue então que:

$$\Pr(A \cap B) = \frac{r}{r+b} \cdot \frac{b}{r+b-1}.$$

- Esse princípio pode ser estendido para qualquer número de eventos, conforme descrito no seguinte teorema.

2.1 Probabilidade condicional – Regra da multiplicação

- **Teorema 2.1.2: Regra da multiplicação para probabilidades condicionais (múltiplos eventos)**
  Suponha que $A_1, A_2, \dots, A_n$ sejam eventos tais que $\Pr(A_1 \cap A_2 \cap \cdots \cap A_n) > 0$. Então

$$\Pr(A_1 \cap A_2 \cap \cdots \cap A_n)$$
$$= \Pr(A_1)\Pr(A_2|A_1)\Pr(A_3|A_1 \cap A_2)\cdots\Pr(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1}).$$

- Demonstração algébrica direta:
  - O produto das probabilidades do lado direito da Equação acima é:

$$\Pr(A_1) \cdot \frac{\Pr(A_1 \cap A_2)}{\Pr(A_1)} \cdot \frac{\Pr(A_1 \cap A_2 \cap A_3)}{\Pr(A_1 \cap A_2)} \cdots \frac{\Pr(A_1 \cap A_2 \cap \cdots \cap A_n)}{\Pr(A_1 \cap A_2 \cdots \cap A_{n-1})}.$$

  - Todos os numeradores e denominadores se cancelam, exceto o último numerador, que é justamente o lado esquerdo da equação.

2.1 Probabilidade condicional – Regra da multiplicação

- **Teorema 2.1.2: Regra da multiplicação para probabilidades condicionais (múltiplos eventos)**
  Demonstração por indução (fraca) em *n*:
  - Caso base (n=2):
    $\Pr(A_1 \cap A_2) = \Pr(A_1).\Pr(A_2|A_1)$ (segue diretamente do Teo 2.1.1)
  - Hipótese de indução: Suponha que, para n>2,
    $\Pr(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) =$
    $\Pr(A_1).\Pr(A_2|A_1).\Pr(A_3|A_1 \cap A_2) \ldots \Pr(A_{n-1}|A_1 \cap A_2 \cap \cdots \cap A_{n-2})$
  - Passo da Indução: Pelo Teo 2.1.1,
    $\Pr(A_1 \cap A_2 \cap \cdots \cap A_{n-1} \cap A_n)$
    $= \Pr(A_1 \cap A_2 \cap \cdots \cap A_{n-1}).\Pr(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})$
    (basta lembrar que $A_1 \cap A_2 \cap \cdots \cap A_{n-1}$ é um evento).
    Aplicando a hipótese de indução sobre $\Pr(A_1 \cap A_2 \cap \cdots \cap A_{n-1})$ na equação acima, temos
    $\Pr(A_1 \cap A_2 \cap \cdots \cap A_{n-1} \cap A_n)$
    $= \Pr(A_1).\Pr(A_2|A_1).\Pr(A_3|A_1 \cap A_2) \ldots$
    $\Pr(A_{n-1}|A_1 \cap A_2 \cap \cdots \cap A_{n-2}).\Pr(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})$,
    c.q.d.

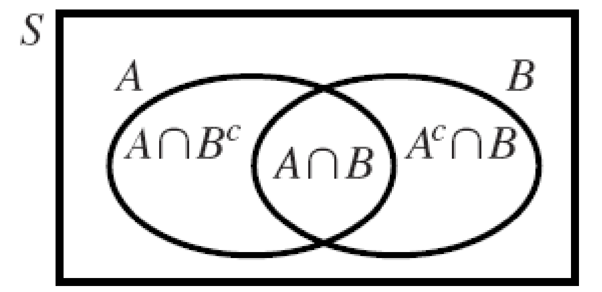
2.1 Probabilidade condicional – Regra da multiplicação

- Exemplo: Seleção de quatro bolas
  - Suponha que quatro bolas devem ser selecionadas aleatoriamente, sem reposição, de uma caixa contendo *r* bolas vermelhas e *b* bolas azuis, $r \geq 2, b \geq 2$.
  - Devemos determinar a probabilidade de que a sequência de bolas sorteadas seja vermelha, azul, vermelha, azul.
  - Resp. Se denotarmos R_j o evento de que uma bola vermelha foi obtida no *j*-ésimo sorteio, e B_j o evento de que uma bola azul foi obtida no *j*-ésimo sorteio (*j*=1,...,4), então:

$$\begin{aligned}\Pr(R_1 \cap B_2 \cap R_3 \cap B_4) &= \Pr(R_1)\Pr(B_2|R_1)\Pr(R_3|R_1 \cap B_2)\Pr(B_4|R_1 \cap B_2 \cap R_3) \\ &= \frac{r}{r+b} \cdot \frac{b}{r+b-1} \cdot \frac{r-1}{r+b-2} \cdot \frac{b-1}{r+b-3}.\end{aligned}$$

2.1 Probabilidade condicional e partições

- Resultados anteriores importantes (revisar!):
  - **Teorema 1.4.11:** Partição de um conjunto em dois:
    Para todos conjuntos A e B, $A \cap B$ e $A \cap B^c$ são disjuntos e $A = (A \cap B) \cup (A \cap B^c)$.



  - **Teorema 2.1.1:** Regra da multiplicação para probabilidades condicionais:
    Sejam A e B eventos. Se Pr(B)>0, então $\Pr(A \cap B) = \Pr(A)\Pr(A|B)$.

- A generalização do Teorema 1.4.11 para partições maiores, combinada com o Teorema 2.1.1, fornece uma ferramenta poderosa para cálculo de probabilidades, como veremos a seguir.

2.1 Probabilidade condicional e partições

- **Definição 2.1.2: Partição**
  Seja *S* o espaço amostral de um experimento, e considere *k* eventos <u>disjuntos</u> $B_1, B_2, \dots, B_k$ em *S* tais que $\bigcup_{i=1}^{k} B_i = S$. Dizemos que esses eventos formam uma **partição** de S.
  – Relembrando a definição geral de partição em teoria dos conjuntos: Uma *partição* de um conjunto não vazio *A* é uma coleção de subconjuntos $A_1, A_2, \dots$ tais que $A_i \neq \emptyset$, $A_i \cap A_j = \emptyset$, $\bigcup_i A_i = A$

- **Teorema 2.1.4: Lei da probabilidade total**
  Suponha que os eventos $B_1, B_2, \dots, B_k$ formam uma partição do espaço *S* e que $\Pr(B_j) > 0$ para *j*=1,...,*k*. Então, para qualquer evento A em S,

$$\Pr(A) = \sum_{j=1}^{k} \Pr(B_j) \Pr(A|B_j).$$

2.1 Probabilidade condicional e partições

- **Teorema 2.1.4: Lei da probabilidade total**
  - Dem: Os eventos $B_1 \cap A,\ B_2 \cap A, \dots, B_k \cap A$ formam uma partição de A, como ilustrado na figura abaixo. Assim, podemos escrever

$$A = (B_1 \cap A) \cup (B_2 \cap A) \cup \cdots \cup (B_k \cap A).$$

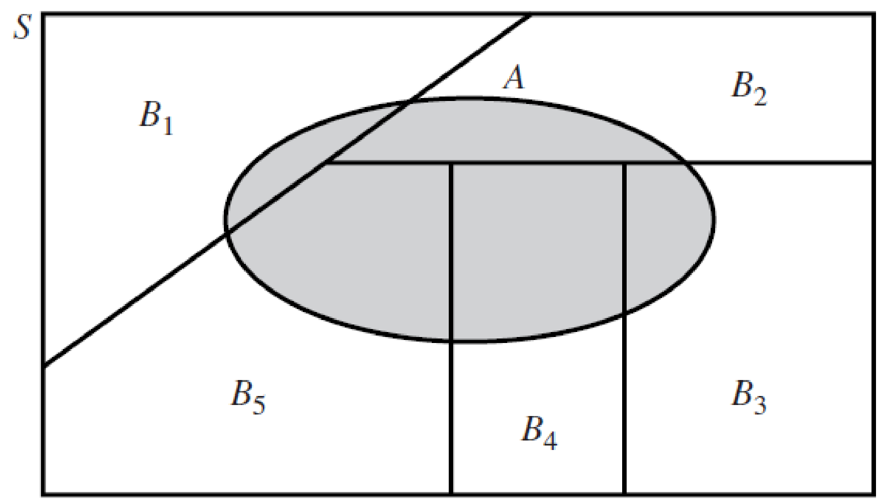
Além disso, como os *k* eventos do lado direito da equação são disjuntos, temos:

$$\Pr(A) = \sum_{j=1}^{k} \Pr(B_j \cap A).$$

Como $\Pr(B_j) > 0$ para todo j, vale

$$\Pr(B_j \cap A) = \Pr(B_j)\Pr(A|B_j)$$

e o resultado é imediato.



2.1 Probabilidade condicional e partições

- Exemplo 2.1.11: Seleção de parafusos
  - Duas caixas contêm parafusos longos e parafusos curtos:
    - Cx 1: 60 longos + 40 curtos; Cx 2: 10 longos + 20 curtos
  - Uma das caixas é selecionada ao acaso e um parafuso é então selecionado aleatoriamente daquela caixa.
  - Queremos determinar a probabilidade desse parafuso ser longo.
  - Resposta:
    - Seja B1 o evento indicando que a caixa 1 foi selecionada e B2 o evento de que a caixa 2 foi selecionada. Seja A o evento indicando que um parafuso longo foi selecionado.
    - Pelo Teo 2.1.4, $\Pr(A) = \Pr(B_1)\Pr(A|B_1) + \Pr(B_2)\Pr(A|B_2).$
    - Sabemos que:
      - $Pr(B_1) = \Pr(B_2) = 1/2;$
      - $Pr(A|B_1) = 60/100 = 3/5;$ $Pr(A|B_2) = 10/30 = 1/3$
    - Logo: $Pr(A) = \frac{1}{2} \cdot \frac{60}{100} + \frac{1}{2} \cdot \frac{10}{30} = \frac{7}{15}$

2.2 Eventos independentes

- Se a descoberta de que *B* ocorreu não altera a probabilidade de *A*, dizemos que *A* e *B* são *independentes*.

- **Definição 2.2.1. Eventos independentes.**
  Dois eventos *A* e *B* são independentes se
  $$\Pr(A \cap B) = \Pr(A)\Pr(B).$$

- Suponha que $\Pr(A) > 0$ e $\Pr(B) > 0$. Então segue diretamente das definições de independência e probabilidade condicional que *A* e *B* são independentes se e somente se $\Pr(A|B) = \Pr(A)$ e $\Pr(B|A) = \Pr(B)$.

2.2 Eventos independentes

- Exemplo 2.2.2:

Machine Operation. Suppose that two machines 1 and 2 in a factory are operated independently of each other. Let A be the event that machine 1 will become inoperative during a given 8-hour period, let B be the event that machine 2 will become inoperative during the same period, and suppose that $\Pr(A) = 1/3$ and $\Pr(B) = 1/4$. We shall determine the probability that at least one of the machines will become inoperative during the given period.

The probability $\Pr(A \cap B)$ that both machines will become inoperative during the period is

$$\Pr(A \cap B) = \Pr(A)\Pr(B) = \left(\frac{1}{3}\right)\left(\frac{1}{4}\right) = \frac{1}{12}.$$

Therefore, the probability $\Pr(A \cup B)$ that at least one of the machines will become inoperative during the period is

$$\begin{aligned} \Pr(A \cup B) &= \Pr(A) + \Pr(B) - \Pr(A \cap B) \\ &= \frac{1}{3} + \frac{1}{4} - \frac{1}{12} = \frac{1}{2}. \end{aligned}$$ ◀

2.2 Eventos independentes

- Exemplo 2.2.3:

Rolling a Die. Suppose that a balanced die is rolled. Let A be the event that an even number is obtained, and let B be the event that one of the numbers 1, 2, 3, or 4 is obtained. We shall show that the events A and B are independent.

In this example, $\Pr(A) = 1/2$ and $\Pr(B) = 2/3$. Furthermore, since $A \cap B$ is the event that either the number 2 or the number 4 is obtained, $\Pr(A \cap B) = 1/3$. Hence, $\Pr(A \cap B) = \Pr(A)\Pr(B)$. It follows that the events A and B are independent events, even though the occurrence of each event depends on the same roll of a die. ◀

  – Interpretação desse resultado: no próximo slide

2.2 Eventos independentes

- Exemplo 2.2.3 (cont):
  - A independência dos eventos A e B pode ser interpretada como segue:
  - Suponha que uma pessoa quer apostar se o número obtido no lançamento do dado é par ou ímpar. Como há 3 resultados de cada tipo (par ou ímpar), a pessoa não terá nenhuma preferência inicial sobre em qual dos resultados (par ou ímpar) apostar.
  - Antes do apostador fazer sua aposta, o dado é lançado mas o resultado não é mostrado. O apostador só é informado que o resultado real foi um dos 4 números 1,2,3,4, ou seja, que o evento B ocorreu.
  - Como 2 desses resultados são pares e 2 são ímpares, o apostador continuará a não ter nenhuma preferência por apostar no número par ou no ímpar.
  - Em outras palavras, a informação de que o evento B ocorreu não ajuda em nada o apostador adivinhar se o evento A ocorreu ou não, ou seja, $\Pr(A|B) = \Pr(A)$.

2.2 Eventos independentes

- Exemplo 2.2.3 (cont):
  - Observação: Este exemplo não serve para quaisquer pares de eventos sobre um dado.
  - P.ex. se considerarmos
    A={um número par é obtido} e *B*={1,2 ou 3 é obtido},
    então os eventos *A* e *B* <u>não são</u> independentes:
    $$\Pr(A \cap B) = \Pr(B)\,.\Pr(A|B) = \frac{1}{2}.\frac{1}{3} = \frac{1}{6} \neq \Pr(A)\Pr(B) = \frac{1}{2}.\frac{1}{2} = \frac{1}{4}$$
  - Neste caso, considerando a situação hipotética de uma aposta, apresentada no slide anterior, teremos
    $$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{2}{6} = \frac{1}{3}$$
    Logo, o apostador deve apostar que um <u>número ímpar</u> foi obtido (ou seja, deve apostar no evento A^C)

2.2 Eventos independentes

- Independência de complementos:
  - Se A e B são independentes, a ocorrência de A não deve estar relacionada com a ocorrência de B.
  - O mesmo vale para a não-ocorrência de A em relação à ocorrência (ou não-ocorrência) de B.
  - Logo, se A e B são independentes, então:
    - A e B^c são independentes;
    - A^c e B são independentes;
    - A^c e B^c são independentes.

- Teorema a seguir mostra um desses resultados.

2.2 Eventos independentes

- **Teorema 2.2.1:**
  Se dois eventos A e B são independentes, então os eventos *A* e B^c são independentes.
  - Dem: O Teorema 1.5.6 estabelece que

    $$\Pr(A \cap B^c) = \Pr(A) - \Pr(A \cap B).$$

    (facilmente verificável pelo diagrama de Venn)

    Adicionalmente, como A e B são independentes, $\Pr(A \cap B) = \Pr(A)\Pr(B)$. Segue então que:

    $$\begin{aligned}\Pr(A \cap B^c) &= \Pr(A) - \Pr(A)\Pr(B) = \Pr(A)[1 - \Pr(B)] \\ &= \Pr(A)\Pr(B^c).\end{aligned}$$

    e portanto *A* e B^c são independentes.

2.2 Eventos independentes – múltiplos eventos

- A definição de eventos independentes pode ser estendida para quaisquer número de eventos, $A_1, A_2, \dots, A_k$.
  - Intuitivamente: se a descoberta de que alguns desses eventos ocorre (ou não) não altera nossas probabilidades para quaisquer eventos que dependam apenas dos eventos restantes, então dizemos que todos os *k* eventos são independentes.

- **Definição 2.2.2: Eventos mutuamente independentes.** Os *k* eventos $A_1, A_2, \dots, A_k$ são *independentes* (ou *mutuamente independentes*) se, para todo subconjunto $A_{i_1}, \dots, A_{i_j}$ desses eventos (*j* = 2,3,...,*k*),

$$\Pr(A_{i_1} \cap \cdots \cap A_{i_j}) = \Pr(A_{i_1}) \cdots \Pr(A_{i_j}).$$

2.2 Eventos independentes – múltiplos eventos

- Por exemplo, para que três eventos A, B e C sejam mutuamente independentes, as quatro relações abaixo precisam ser satisfeitas:

$$\begin{aligned} \Pr(A \cap B) &= \Pr(A)\Pr(B), \\ \Pr(A \cap C) &= \Pr(A)\Pr(C), \\ \Pr(B \cap C) &= \Pr(B)\Pr(C), \end{aligned} \tag{2.2.1}$$

e

$$\Pr(A \cap B \cap C) = \Pr(A)\Pr(B)\Pr(C). \tag{2.2.2}$$

2.2 Eventos independentes – múltiplos eventos

- Exemplo 2.2.5:

Inspecting Items. Suppose that a machine produces a defective item with probability p $(0 < p < 1)$ and produces a nondefective item with probability $1 - p$. Suppose further that six items produced by the machine are selected at random and inspected, and that the results (defective or nondefective) for these six items are independent. We shall determine the probability that exactly two of the six items are defective.

It can be assumed that the sample space S contains all possible arrangements of six items, each one of which might be either defective or nondefective. For $j = 1, \dots, 6$, we shall let D_j denote the event that the jth item in the sample is defective so that D_j^c is the event that this item is nondefective. Since the outcomes for the six different items are independent, the probability of obtaining any particular sequence of defective and nondefective items will simply be the product of the individual probabilities for the items. For example,

$$\begin{aligned}\Pr(D_1^c \cap D_2 \cap D_3^c \cap D_4^c \cap D_5 \cap D_6^c) &= \Pr(D_1^c)\Pr(D_2)\Pr(D_3^c)\Pr(D_4^c)\Pr(D_5)\Pr(D_6^c)\\ &= (1-p)p(1-p)(1-p)p(1-p) = p^2(1-p)^4.\end{aligned}$$

2.2 Eventos independentes – múltiplos eventos

- Exemplo 2.2.5 (cont):

It can be seen that the probability of any other particular sequence in S containing two defective items and four nondefective items will also be $p^2(1-p)^4$. Hence, the probability that there will be exactly two defectives in the sample of six items can be found by multiplying the probability $p^2(1-p)^4$ of any particular sequence containing two defectives by the possible number of such sequences. Since there are $\binom{6}{2}$ distinct arrangements of two defective items and four nondefective items, the probability of obtaining exactly two defectives is $\binom{6}{2}p^2(1-p)^4$. ◄

- Exemplo 2.2.6 :

Obtaining a Defective Item. For the conditions of Example 2.2.5, we shall now determine the probability that at least one of the six items in the sample will be defective.

Since the outcomes for the different items are independent, the probability that all six items will be nondefective is $(1-p)^6$. Therefore, the probability that at least one item will be defective is $1-(1-p)^6$. ◄

2.2 Eventos independentes – múltiplos eventos

- Exemplo 2.2.7:

Tossing a Coin Until a Head Appears. Suppose that a fair coin is tossed until a head appears for the first time, and assume that the outcomes of the tosses are independent. We shall determine the probability p_n that exactly n tosses will be required.

The desired probability is equal to the probability of obtaining $n-1$ tails in succession and then obtaining a head on the next toss. Since the outcomes of the tosses are independent, the probability of this particular sequence of n outcomes is $p_n = (1/2)^n$.

The probability that a head will be obtained sooner or later (or, equivalently, that tails will not be obtained forever) is

$$\sum_{n=1}^{\infty} p_n = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = 1.$$

Since the sum of the probabilities p_n is 1, it follows that the probability of obtaining an infinite sequence of tails without ever obtaining a head must be 0. ◀

2.2 Eventos independentes – múltiplos eventos

- Exemplo 2.2.8:

Inspecting Items One at a Time. Consider again a machine that produces a defective item with probability p and produces a nondefective item with probability $1 - p$. Suppose that items produced by the machine are selected at random and inspected one at a time until exactly five defective items have been obtained. We shall determine the probability p_n that exactly n items ($n \geq 5$) must be selected to obtain the five defectives.

The fifth defective item will be the nth item that is inspected if and only if there are exactly four defectives among the first $n - 1$ items and then the nth item is defective. By reasoning similar to that given in Example 2.2.5, it can be shown that the probability of obtaining exactly four defectives and $n - 5$ nondefectives among the first $n - 1$ items is $\binom{n-1}{4} p^4 (1 - p)^{n-5}$. The probability that the nth item will be defective is p. Since the first event refers to outcomes for only the first $n - 1$ items and the second event refers to the outcome for only the nth item, these two events are independent. Therefore, the probability that both events will occur is equal to the product of their probabilities. It follows that

$$p_n = \binom{n-1}{4} p^5 (1 - p)^{n-5}. \qquad \blacktriangleleft$$

2.3 Teorema de Bayes

LII. *An Eſſay towards ſolving a Problem in the Doctrine of Chances. By the late Rev. Mr.* Bayes, *F. R. S. communicated by Mr.* Price, *in a Letter to* John Canton, *A. M. F. R. S.*

Mr. Bayes and Mr. Price

Phil. Trans. 1763 **53**, 370-418, published 1 January 1763

PROBLEM.

Given the number of times in which an unknown event has happened and failed: *Required* the chance that the probability of its happening in a ſingle trial lies ſomewhere between any two degrees of probability that can be named.

2.3 Teorema de Bayes

- Interpretação frequentista de probabilidade:
  - Se um grande número de repetições forem realizados para observação de um evento, a frequência relativa da ocorrência do evento deve estar situada em um intervalo próximo da (verdadeira) probabilidade de ocorrência do evento.
- Interpretação de Bayes para probabilidade: como inferir causas a partir dos efeitos
  - Como podemos aprender a probabilidade de um evento futuro ocorrer se apenas sabemos quantas vezes ele ocorreu ou não ocorreu no passado?
  - Solução de Bayes:
    - Criar uma estimativa inicial e melhorar sua estimativa à medida em que se coleta mais informação.
    - Crença inicial + Novos dados → Crença melhorada

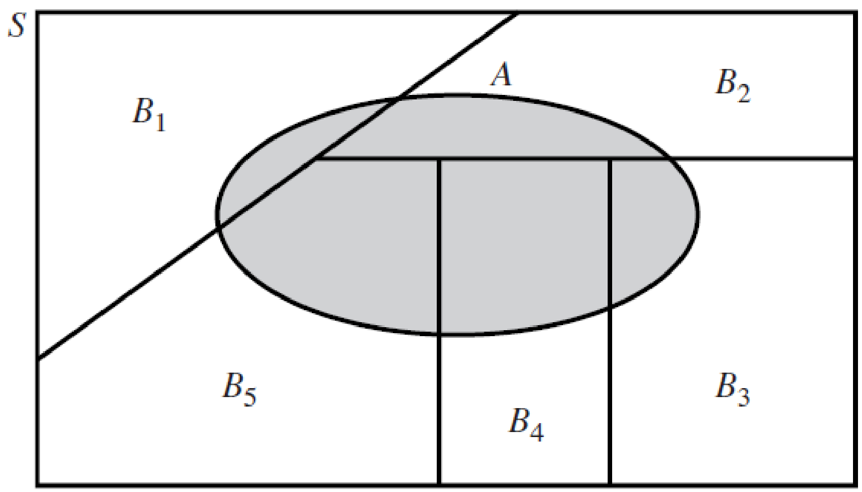
2.3 Teorema de Bayes

- Exemplo 2.3.1: Teste de uma doença
  - Você está caminhando na rua e nota que o posto de saúde está fornecendo um teste gratuito para uma certa doença.
  - O teste tem a seguinte confiabilidade:
    - *Sensibilidade:* Se uma pessoa tem a doença, o teste tem 90% de probabilidade de dar um resultado <u>positivo</u>
    - *Especificidade:* Se uma pessoa não tem a doença, o teste tem 90% de probabilidade de dar um resultado <u>negativo</u>. (Portanto, só 10% de probabilidade de dar resultado positivo)
  - Dados epidemiológicos indicam que prevalência da doença é de apenas 1 em 10.000, mas como o teste é gratuito e não invasivo, você decide fazer.
  - Alguns dias depois você recebe uma carta informando que seu teste deu positivo.
  - Agora, qual é a chance de que você tem a doença?

    →Resposta adiante

2.3 Teorema de Bayes

- Intuição (e demonstração) do teorema de Bayes:
  - Suponha que você esteja interessado em obter as probabilidades de k eventos $B_1, B_2, \ldots, B_k$ após a realização de um experimento
    - Esses eventos formam um partição de S → um deles deve ocorrer
  - Você possui estimativas iniciais de $\Pr(B_i) > 0, i = 1 \ldots k$.
  - Seu interesse é calcular as probabilidades condicionais $\Pr(B_i|A)$, $i = 1, \ldots, k$, dado que evento A ocorreu no experimento.



2.3 Teorema de Bayes

- Intuição (e demonstração) do teorema de Bayes:
  - Seu interesse é calcular as probabilidades condicionais $\Pr(B_i|A)$, $i = 1, .., k$, dado que evento A ocorreu no experimento.
  - **Resposta:**
  - Da definição de probabilidade condicional + Teorema 2.1.1:

$$\Pr(B_i|A) = \frac{\Pr(B_i \cap A)}{\Pr(A)} = \frac{\Pr(B_i)\Pr(A|B_i)}{\Pr(A)}, \qquad i = 1, \dots, k$$

  - Agora, pelo Teorema 2.4.1 (ver diagrama no slide anterior):

$$\Pr(A) = \sum_{j=1}^{k} \Pr(B_j \cap A) = \sum_{j=1}^{k} \Pr(B_j)\Pr(A|B_j)$$

  - Substituindo $\Pr(A)$ no denominador da 1ª equação, temos:

$$\Pr(B_i|A) = \frac{\Pr(B_i)\Pr(A|B_i)}{\sum_{j=1}^{k}\Pr(B_j)\Pr(A|B_j)}.$$

2.3 Teorema de Bayes

- **Teorema 2.3.1: Teorema de Bayes**
  Sejam $B_1, B_2, \dots, B_k$ eventos que formam uma partição do espaço S tais que $\Pr(B_i) > 0, i = 1 \dots k$, e seja A um evento tal que $\Pr(A) > 0$. Então, para $i = 1 \dots k$,
  $$\Pr(B_i|A) = \frac{\Pr(B_i)\Pr(A|B_i)}{\sum_{i=1}^{k}\Pr(B_i)\Pr(A|B_i)}.$$

- $\Pr(B_i|A)$ é calculado exclusivamente a partir das probabilidades a priori $\Pr(B_i)$ e da verossimilhança $\Pr(A|B_i)$ de A ocorrer, condicionado a cada um dos eventos alternativos $B_1, B_2, \dots, B_k$

2.3 Teorema de Bayes

- Retornando ao Exemplo 2.3.1: Teste de uma doença
  - O teste para uma doença tem a seguinte confiabilidade: Pr(positivo|doença) = 90%; Pr(positivo|não doença) = 10%
  - Prevalência da doença na população: Pr(doença) = 1/10000
  - Seu teste deu positivo.
  - Agora, qual é a chance de que você tem a doença?
  - Resp:
    - B_1: você possui a doença; B_2: você não possui a doença
    - A: Seu teste deu positivo

$$\Pr(B_1|A) = \frac{\Pr(B_1)\Pr(A|B_1)}{\Pr(B_1)\Pr(A|B_1) + \Pr(B_2)\Pr(A|B_2)}$$

$$= \frac{0.0001 \times 0.9}{0.0001 \times 0.9 + 0.9999 \times 0.1} = 0.00090.$$

  - A probabilidade condicional de você ter a doença dado o resultado do teste aumentou 9x, mas continua baixa: aprox. 1 em 1000

2.3 Teorema de Bayes

- Retornando ao Exemplo 2.3.1: Teste de uma doença
  - A probabilidade condicional de você ter a doença dado o resultado do teste aumentou 9x, mas continua baixa: aprox. 1 em 1000
  - Explicação:
    - Doença é relativamente rara (1 em 10.000)
    - O teste tem uma alta taxa de resultados positivos (1 em 10), ou seja, o número de positivos pelo teste é 1000 vezes o número de pessoas que realmente têm a doença.
    - Em outras palavras, de cada 1.000 pessoas para as quais o teste dá um resultado positivo, apenas uma pessoa tem a doença.
  - Questões:
    1. E se o teste fosse mais robusto?
       - P.ex. com menor número de falsos positivos?
    2. E se o teste tivesse dado negativo?

    Exercícios no próximo slide

2.3 Teorema de Bayes

- Retornando ao Exemplo 2.3.1: Teste de uma doença
  - Neste exemplo, a evidência de você ter a doença aumentaria significativamente se a taxa de falsos positivos do teste fosse menor
  - Exercício 1:
    - Clique aqui para baixar a planilha desse exemplo
    - Teste diferentes valores de Pr(A|B1) e Pr(A|B2):
      - Pr(A|B1) = 0.9, 0.99, 0.999, etc
      - Pr(A|B2) = 0.1, 0.01, 0.001, 0.0001, etc.
    - Observe que a Pr(B1|A) é mais sensível à taxa de falsos positivos – Pr(A|B2) – do que à taxa de verdadeiros positivos – Pr(A|B1)
  - Exercício 2:
    - Suponha que seu teste deu negativo (chamemos de evento N)
    - Calcule a probabilidade de você não ter a doença, dado esse resultado:
      - Pr(B1|N)

2.3 Teorema de Bayes

- Voltando ao Exemplo: A caixa de Bertrand:
  - Você passou a vida toda procurando as Pérolas Gêmeas (um dos tesouros arqueológico muito valioso) e finalmente as encontrou.
  - Há 3 porta-joias fechados, cada qual com duas gavetas, e você sabe que as Pérolas Gêmeas estão em um deles.
  - Quando você força a abertura de uma das gavetas, encontra uma pérola (mas não sabe distinguir se é uma pérola comum ou uma das pérolas gêmeas)
  - Problema:
    - Qual a probabilidade de você encontrar outra pérola na 2ª gaveta do mesmo porta-joias?
    - Ou: Qual a probabilidade de que a 1ª gaveta que você abriu era do porta-joias correto?
  - Decisão a tomar: escolher outro gaveteiro ou não

2.3 Teorema de Bayes

- Exemplo: A caixa de Bertrand (cont):
  - Chamemos os três porta-joias de PP, PC e CC, de acordo com seu conteúdo (2 pérolas, 1 pérola e 1 carvão, 2 carvões).
  - O que queremos saber é a probabilidade de nosso porta-joias ser o PP.
  - Anteriormente: resposta via probabilidade condicional:
    - Evento PP: o porta-joias aberto é o PP
    - Evento A: a gaveta aberta contém uma pérola.

$$\Pr(PP|A) = \frac{\Pr(PP \cap A)}{PP(A)} = \frac{1/3}{1/2} = 2/3$$

2.3 Teorema de Bayes

- Exemplo: A caixa de Bertrand (cont):
  - Outra forma de responder: Teorema de Bayes:
    - $\Pr(PP) = \Pr(PC) = \Pr(CC) = 1/3$
    - Probabilidades do evento A sob cada uma das três hipóteses:
      - $\Pr(A|PP) = 1;\ \Pr(A|PC) = \frac{1}{2};\ \Pr(A|CC) = 0$

$$\Pr(PP|A)$$
$$= \frac{\Pr(PP)\Pr(A|PP)}{\Pr(PP)\Pr(A|PP) + \Pr(PC)\Pr(A|PC) + \Pr(CC)\Pr(A|CC)}$$
$$= \frac{\frac{1}{3}\times 1}{\frac{1}{3}\times 1 + \frac{1}{3}\times\frac{1}{2} + \frac{1}{3}\times 0} = 2/3$$

2.3 Teorema de Bayes

- Exemplo 2.3.4:

Identifying the Source of a Defective Item. Three different machines M_1, M_2, and M_3 were used for producing a large batch of similar manufactured items. Suppose that 20 percent of the items were produced by machine M_1, 30 percent by machine M_2, and 50 percent by machine M_3. Suppose further that 1 percent of the items produced by machine M_1 are defective, that 2 percent of the items produced by machine M_2 are defective, and that 3 percent of the items produced by machine M_3 are defective. Finally, suppose that one item is selected at random from the entire batch and it is found to be defective. We shall determine the probability that this item was produced by machine M_2.

2.3 Teorema de Bayes

- Exemplo 2.3.4 - solução:

Let B_i be the event that the selected item was produced by machine M_i ($i = 1, 2, 3$), and let A be the event that the selected item is defective. We must evaluate the conditional probability $\Pr(B_2|A)$.

The probability $\Pr(B_i)$ that an item selected at random from the entire batch was produced by machine M_i is as follows, for $i = 1, 2, 3$:

$$\Pr(B_1) = 0.2, \quad \Pr(B_2) = 0.3, \quad \Pr(B_3) = 0.5.$$

Furthermore, the probability $\Pr(A|B_i)$ that an item produced by machine M_i will be defective is

$$\Pr(A|B_1) = 0.01, \quad \Pr(A|B_2) = 0.02, \quad \Pr(A|B_3) = 0.03.$$

It now follows from Bayes' theorem that

$$\Pr(B_2|A) = \frac{\Pr(B_2)\Pr(A|B_2)}{\sum_{j=1}^{3}\Pr(B_j)\Pr(A|B_j)}$$

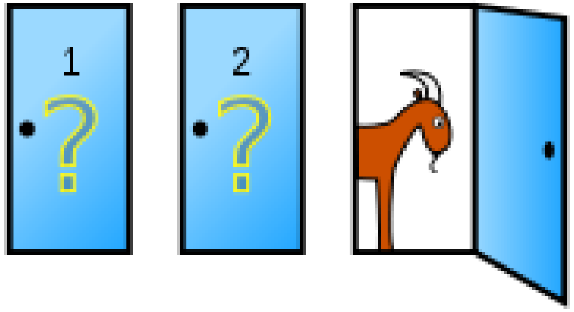
$$= \frac{(0.3)(0.02)}{(0.2)(0.01) + (0.3)(0.02) + (0.5)(0.03)} = 0.26.$$

2.3 Teorema de Bayes – Probabilidades *a priori* e *a posteriori*

- No exemplo 2.3.4, uma probabilidade como $\Pr(B_2)$ é comumente denominada a probabilidade *a priori* de que um item selecionado ao acaso tenha sido produzido pela máquina M_2
  - $\Pr(B_2)$ é a probabilidade do evento B_2 antes do item ser selecionado e antes de sabermos se o item selecionado é defeituoso ou não
- Uma probabilidade como $\Pr(B_2|A)$ é usualmente denominada probabilidade *a posteriori* de que o item selecionado tenha sido produzido pela máquina M_2
  - $\Pr(B_2|A)$ é a probabilidade do evento B_2 depois de sabermos que o item selecionado é defeituoso

2.3 Teorema de Bayes

- Exemplo: O problema de Monty Hall
  - Você ficou muito feliz ao ser escolhido para participar do quadro "Ferrari ou bode" de um programa de televisão, em que poderá sair dirigindo uma Ferrari zero quilômetro, ou ganhar um bode de estimação.
  - Há três portas. Atrás de uma está o carro, atrás das outras estão os bodes (todos colocados de forma aleatória).
  - O jogo:
    - Você escolhe uma porta inicial.
    - Monty Hall (o apresentador), que sabe onde o carro está, abrirá uma porta com o bode, e perguntará se você quer trocar.



  - Existe diferença entre trocar de porta ou não?
    - Para responder a essa pergunta, deve-se inferir qual a probabilidade de ganho em cada estratégia.

Fonte: Jeremy Stangroom. O Enigma de Einstein. Ed. Marco Zero, 2010.

2.3 Teorema de Bayes

- Exemplo: O problema de Monty Hall (cont)
- Solução direta:
  - Estratégia: Não trocar de porta:
    - Você só ganha o carro se sua 1ª escolha de porta for a correta
    - Probabilidade de ganho com esta estratégia: 1/3
  - Estratégia: Trocar de porta:
    - Se sua 1ª escolha de porta for a correta, você não ganha o carro (já que vai trocar por uma das duas outras contendo bodes).
    - Mas, se a 1ª porta que você tiver escolhido tiver um bode, então Monty Hall será obrigado a abrir a porta com o outro bode, deixando para você a porta com a Ferrari.
      - Não interessa sua 1ª porta era aquela com o bode 1 ou com o bode 2, sua 2ª porta será sempre aquela com a Ferrari.
    - Ou seja: você só <u>perde</u> o carro o carro se sua 1ª escolha for a correta.

      → Probabilidade de ganho: 2/3

2.3 Teorema de Bayes

- Exemplo: O problema de Monty Hall (cont)
- Solução via Teorema de Bayes:
  - Vamos considerar os eventos C1, C2, C3 indicando que a Ferrari está atrás da porta 1,2 e 3, respectivamente.
  - Vamos supor, sem perda de generalidade (s.p.g), que você tenha escolhido a porta 1.
  - Também vamos supor, s.p.g, que o apresentador tenha aberto a porta 3 (evento denotado por A)
  - Inicialmente, a probabilidade do carro estar em qualquer uma das 3 portas é a mesma, ou seja, $\Pr(C1) = \Pr(C2) = \Pr(C3) = 1/3$
  - As probabilidades do apresentador abrir a porta 3 em cada cenário são:
    - $\Pr(A|C1) = \frac{1}{2}$ (se o carro estiver na porta que você escolheu, o apresentador poderia ter escolhido as portas 1 ou 2)
    - $\Pr(A|C2) = 1;\ \ \Pr(A|C3) = 0$ (por quê?)

2.3 Teorema de Bayes

- Exemplo: O problema de Monty Hall (cont)
- Solução via Teorema de Bayes (cont):
  - $\Pr(C1) = \Pr(C2) = \Pr(C3) = 1/3$
  - As probabilidades do apresentador abrir a porta 3 em cada cenário são: $\Pr(A|C1) = \frac{1}{2}$; $\Pr(A|C2) = 1$; $\Pr(A|C3) = 0$
  - Logo, se você inicialmente selecionar a porta 1 e o apresentador abrir a porta 3, então a troca implica necessariamente em escolher a porta 2.
  - A probabilidade condicional de vitória na estratégia de troca é:

$$\Pr(C2|A) = \frac{\Pr(C2)\Pr(A|C2)}{\Pr(C1)\Pr(A|C1) + \Pr(C2)\Pr(A|C2) + \Pr(C3)\Pr(A|C3)}$$

$$= \frac{\frac{1}{3}\times 1}{\frac{1}{3}\times\frac{1}{2} + \frac{1}{3}\times 1 + \frac{1}{3}\times 0} = 2/3,$$

o que implica que a probabilidade condicional de vitória na estratégia de não trocar é de apenas 1/3!

2.3 Teorema de Bayes

- Breve histórico do problema de Monty Hall:
  - Desafio inspirado em um programa de televisão americana, “Let’s Make a Deal”. Foi apresentado por Steve Selvin para o periódico American Statistician em 1975
  - Ficou famoso quando foi apresentado por um leitor para a coluna “Ask Marilyn”, na revista Parade em 1990
  - O argumento comum (defendido por vários estatísticos) era de que, uma vez que o apresentador (Monty Hall) abriu uma das portas, a probabilidade das outras duas portas conterem o prêmio aumentava para ½, o que significaria que manter a porta original ou trocar teriam as mesmas chances de sucesso
  - Marilyn vos Savant apresentou várias formas de mostrar que trocar de porta era vantajoso, mas recebeu várias cartas ridicularizando seus argumentos
  - As posições dos leitores só mudaram após vos Savant recomendar que as pessoas simulassem o jogo (repetindo várias vezes o jogo sob cada uma das estratégias e verificando a taxa de sucesso em cada uma)
- Referências e links:

Leonard Mlodinow. O Andar do Bêbado: Como o Acaso Determina Nossas Vidas. Ed. Jorge Zahar, 2009.

https://en.wikipedia.org/w/index.php?title=Monty_Hall_problem&oldid=771023549

https://www.jstor.org/stable/2683689?seq=1#page_scan_tab_contents

http://marilynvossavant.com/game-show-problem/