Probabilidade (PPGECD000000001)

Programa de Pós-Graduação em Estatística e Ciência de Dados (PGECD)

Sessão 4

Raydonal Ospina

Departamento de Estatística
Universidade Federal da Bahia
Salvador/BA

Probabilidade Condicional

Motivação

- A probabilidade é uma forma de quantificar a incerteza de um fenômeno. Naturalmente se obtemos mais informações sobre o fenômeno em estudo essa nova informação pode alterar, e por vezes de forma muito significativa a avaliação da probabilidade.
- Probabilidade é baseada em informação e conhecimento.
- Nosso objetivo é saber como atualizar o valor da probabilidade quando esta base de informação ou conhecimento é alterada. Em particular, como alterar a probabilidade de um dado evento A quando sabe-se que um determinado evento B ocorreu?

Interpretação frequentista

- Seja n o número de vezes que repete-se um experimento. Seja N_A (resp., $N_B > 0$ e $N_{A\cap B}$) o número de vezes que o evento A (resp., B e $A \cap B$) ocorre nessas n repetições.
- A probabilidade condicional de A dado que sabe-se que B ocorreu, $P(A|B)$, segundo uma interpretação frequentista, sugere que ela deve ser igual ao limite das frequências relativas condicionais do evento A dado o evento B, isto é, deve ser o limite $N_{A\cap B}/N_B$ quando n tende ao infinito. Ou seja as frequências relativas tendem a se estabilizar ao redor de um valor específico entre 0 e 1.
- Seja $r_A = N_A/n$ a frequência relativa do evento A.
- Note que

$$\frac{N_{A\cap B}}{N_B} = \frac{r_{A\cap B}}{r_B}$$

  e que segundo a interpretação frequentista de probabilidade $r_{A\cap B}/r_B$ é aproximadamente igual a $P(A \cap B)/P(B)$ para valores grandes de n.

Interpretação subjetiva

- Suponha que a incerteza de um agente é descrita por uma probabilidade P em $(\Omega, \mathcal{A})$ e que o agente observa ou fica sabendo que o evento B ocorreu. Como o agente deve atualizar sua probabilidade $P(\cdot|B)$ de modo a incorporar esta nova informação? Claramente, se o agente *acredita* que B é verdadeiro, então parece razoável requerer que

$$P(B^c|B) = 0. \tag{1}$$

- Em relação aos eventos contidos em B, é razoável assumir que sua chance relativa permaneça inalterada se tudo que o agente descobriu foi que o evento B ocorreu, ou seja, se $A_1, A_2 \subseteq B$ com $P(A_2) > 0$, então

$$\frac{P(A_1)}{P(A_2)} = \frac{P(A_1|B)}{P(A_2|B)}. \tag{2}$$

Segue que (1) e (2) determinam completamente $P(\cdot|B)$ se $P(B) > 0$.

Theorem 1

Se $P(B > 0)$ e $P(\cdot|B)$ é uma medida de probabilidade em Ω que satisfaz (1) e (2), então

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

Demonstração.

Como $P(\cdot|B)$ é uma medida de probabilidade e satisfaz $P(B^c|B) = 0$, nós temos que $P(B|B) = 1 - P(B^c|B) = 1$. Considerando $A_1 = A$ e $A_2 = B$ em (2), temos então $P(A|B) = \frac{P(A)}{P(B)}$ para $A \subseteq B$. Se A não é um subconjunto de B, temos que $A = (A \cap B) \cup (A \cap B^c)$. Como $(A \cap B)$ e $(A \cap B^c)$ são eventos disjuntos, temos $P(A|B) = P(A \cap B|B) + P(A \cap B^c|B)$. Como $A \cap B^c \subseteq B^c$ e $P(B^c|B) = 0$, temos que $P(A \cap B^c|B) = 0$. Como $A \cap B \subseteq B$, usando o caso anterior

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

□

Deste modo as interpretações frequentista e subjetivista da probabilidade justificam a seguinte definição.

Definição 1

Seja $(\Omega, \mathcal{A}, P)$ *um espaço de probabilidade. Se* $A, B \in \mathcal{A}$ *e* $P(B) > 0$ *a* probabilidade condicional de A dado B *é definida por*

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

Exemplo 1

Se atiram dois dados honestos ao ar. Qual é a probabilidade condicional de que pelo menos um resultado seja 6 dado que as faces dos dados são diferentes? Para solucionar este problema note que $\Omega = \{(a, b) : a, b \in \{1, 2, \ldots, 6\}\}$ e definamos os eventos

$$\begin{aligned} A &= \text{“Pelo menos um dado cai com a face igual a 6”} \\ B &= \text{“As faces dos dois dados são distintas”} \end{aligned}$$

Neste caso,

$$\begin{aligned} A =& \{(1, 6), (2, 6), (3, 6), (4, 6), (5, 6), (6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)\}, \\ B =& \{(a, b) \in \Omega \,:\, a \neq b\}. \end{aligned}$$

Assim,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{10}{36}}{\frac{30}{36}} = \frac{1}{3}.$$

Teorema 1 (Medida de probabilidade condicional)

Dado que B ocorre, os eventos favoráveis as A são aqueles que pertencem a $A \cap B$*. Assim, para* $(\Omega, \mathcal{F}, P)$ *um espaço de probabilidade em que* $B \in \mathcal{F}$ *com* $P(B) > 0$*. Então:*

A.1 $P(A|B) \geq 0$, *para todo* $A \in \mathcal{F}$,

A.2 $P(\Omega|B) = 1$, *(medida finita).*

A.3 *(σ-aditividade) Se* $A_1, A_2, \ldots$ *é uma sequência de eventos de* $\mathcal{F}$ *mutuamente excludentes, i.e.,* $A_i \cap A_j = \emptyset$ *para todo* $i \neq j$, *então*

$$P\left(\bigcup_{i=1}^{\infty} A_i \middle| B\right) = \sum_{i=1}^{\infty} P(A_i|B)$$

Vamos provar que para um evento fixo B que satisfaz $P(B) > 0$, $P(\cdot|B)$ satisfaz os axiomas de Kolmogorov K1-K4 e realmente é uma medida de probabilidade.

Demonstração.

Para provar $K1$, note que para todo $A \in \mathcal{A}$, como $P(A \cap B) \geq 0$, nós temos

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \geq 0.$$

Para provar $K2$, note que $\Omega \cap B = B$, então

$$P(\Omega|B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1.$$

Finalmente, para provar K4′ (que implica K3), note que se $A_1, A_2, \ldots$ são mutuamente exclusivos $A_1 \cap B, A_2 \cap B, \ldots$ também o são, então

$$P(\cup_i A_i|B) = \frac{P((\cup_i A_i) \cap B)}{P(B)} = \frac{P(\cup_i (A_i \cap B))}{P(B)}$$
$$= \frac{\sum_i P(A_i \cap B)}{P(B)} = \sum_i P(A_i|B).$$

$\square$

Outras propriedades

1. $P(\cdot|A)$ é uma medida de probabilidade sobre Ω, que está "concentrada" em A, isto quer dizer, $P(A|A) = 1$.
2. $P(A|B) = P(A \cap B|B)$;
3. Se $A \cap B = \emptyset$, então $P(B|A) = 0$.
4. se $A \supseteq B$, então $P(A|B) = 1$;
5. $P(A \cap B|C) = P(A|B \cap C)P(B|C)$. Fazendo $C = \Omega$ na propriedade, temos que $P(A \cap B) = P(A|B)P(B)$.

Teorema 2 (Regra da multiplicação)

Consideremos uma sequência finita de eventos aleatórios $A_1, A_2, \ldots, A_n$ tais que os eventos condicionais $A_i|A_1 \cap A_2 \cap \ldots \cap A_{i-1}$ tenham probabilidades positivas. Então temos que a probabilidade de acontecerem todos os eventos é

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2)\ldots P(A_n| \cap_{i=1}^{n-1} A_i).$$

Demonstração.

Use indução. Note também que

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)\frac{P(A_1 \cap A_2)}{P(A_1)}\frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)}\cdots\frac{P(\bigcap_{i=1}^{n} A_i)}{P(\bigcap_{i=1}^{n-1} A_i)}.$$

$\square$

Exemplo 2

Num jogo de cartas, três cartas, são retiradas sem substituição de um baralho (52 cartas). Qual a probabilidade de que nenhuma das cartas retiradas seja um ouro? Note que qualquer carta tem a mesma probabilidade de ser retirada. Agora definamos o evento

$$A_i = \{\text{a } i\text{-ésima carta não é ouro}\},$$

para $i = 1, 2, 3$. Desejamos obter obter $P(A_1 \cap A_2 \cap A_3)$. Lembremos que cada naipe possui 13 cartas. As probabilidades condicionais deste experimento são

$$\begin{aligned} P(A_1) &= \frac{39}{52}, \\ P(A_2|A_1) &= \frac{38}{51}, \\ P(A_3|A_1 \cap A_2) &= \frac{37}{50}. \end{aligned}$$

Logo, pela regra da multiplicação

$$P(A_1 \cap A_2 \cap A_3) = \frac{39}{52} \times \frac{38}{51} \times \frac{37}{50} \approx 0,41.$$

Exercício 1

Um lote contém 15 moldes provenientes de um fornecedor local e 25 de um fornecedor de um estado vizinho. Três moldes são selecionados ao acaso e sem reposição. Seja A_i o evento um que o i-ésimo molde selecionado seja proveniente do fornecedor local. Determine:

(a) $P(A_1)$.

(b) $P(A_2|A_1)$.

(c) $P(A_1 \cap A_2)$.

(d) $P(A_1 \cup A_2)$.

(e) $P(A_1 \cap A_2 \cap A_3)$.

(f) $P(A_1 \cap A_2 \cap A_3^c)$.

Utilizando o seguinte teorema pode-se obter uma probabilidade (incondicional) de uma probabilidade condicional.

Teorema 3 (Teorema da probabilidade total)

Sejam $A_1, A_2, \ldots, A_n$ *eventos dois a dois disjuntos que formam uma partição do espaço amostral, isto é,* $\bigcup_{i=1}^{n} A_i = \Omega$ *e assuma que* $P(A_i) > 0$ *para* $i = 1, 2, \ldots, n$*. Então, para qualquer evento* B*, temos que*

$$\begin{aligned} P(B) &= P(A_1 \cap B) + \cdots + P(A_n \cap B) \\ &= P(A_1)P(B|A_1) + \cdots + P(A_n)P(B|A_n) = \sum_i P(A_i)P(B|A_i). \end{aligned}$$

Uma representação esquemática do teorema anterior pode ser vista no gráfico da Figura 1.

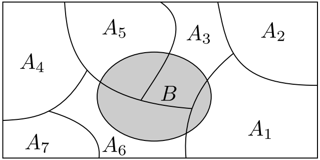

Figura: Representação esquemática do teorema da probabilidade total para $n = 7$.

Demonstração.

Para demonstrarmos o teorema anterior basta observarmos (ajudado pelo gráfico da Figura 1) que a sequência de eventos $A_1, A_2, \ldots$ forma uma partição. Então, para qualquer $B \subset \Omega$, segue que, $B = \bigcup_i (A_i \cap B)$ e como os A_i são disjuntos dois a dois temos que $B \cap A_i$ também são disjuntos e pelo axioma K3 e pelo teorema de probabilidade total concluímos que

$$P(B) = \sum_i P(A_i \cap B) = \sum_i P(A_i) P(B|A_i).$$

$\square$

Interpretação do Teorema da Probabilidade Total

$A_1, A_2, \ldots$ são possíveis causas e o evento A é um efeito particular associado a uma causa, $P(B|A_i)$ especifica a relação estocástica entre a causa A_i e o efeito B.

Exemplo 3

Por exemplo, seja $\{D, D^c\}$ uma partição do espaço amostral, onde D é o evento que um dado indivíduo possui uma certa doença. Seja A o evento que determinado teste para o diagnóstico da doença deu positivo. Então,

- $P(A|D^c)$ - *falso positivo*.
- $P(A^c|D)$ - *falso negativo*.
- Estas probabilidades determinam a qualidade do teste, quanto menores as probabilidades de falso negativo e falso positivo melhor a qualidade do teste.

Caso as probabilidades $P(D)$, $P(A|D)$, $P(A|D^c)$ sejam conhecidas pode-se usando o Teorema da Probabilidade Total obter a probabilidade incondicional de determinado exame dar positivo $P(A)$.

Porém geralmente, o que se busca é saber que dado que o resultado de um exame deu positivo qual a probabilidade de que o indivíduo esteja doente. Pode-se obter esta probabilidade utilizando a famosa *fórmula de Bayes*:

$$P(D|A) = \frac{P(A \cap D)}{P(A \cap D) + P(A \cap D^c)} = \frac{P(A|D)P(D)}{P(A|D)P(D) + P(A|D^c)P(D^c)}.$$

Como corolário do teorema 3 obtemos a muito famosa regra de Bayes, a qual constituí o alicerce da inferência Bayesiana.

Corolário 1 (**Teorema de Bayes ou regra de Bayes)**

Seja $(\Omega, \mathcal{F}, P)$ *um espaço de probabilidade e* $A_1, A_2, \ldots,$ *uma partição finita de* Ω*, então é satisfeita para cada* $B \in \mathcal{F}$ *com* $P(B) > 0$ *a fórmula:*

$$P(A_n|B) = \frac{P(B|A_n)P(A_n)}{\sum_j P(B|A_j)P(A_j)}, \quad \textit{para todo} \quad n.$$

Demonstração.

$$P(A_n|B) = \frac{P(A_n \cap B)}{P(B)} = \frac{P(B|A_n)P(A_n)}{\sum_j P(B|A_j)P(A_j)}.$$

□

Para dar uma interpretação à regra de Bayes, suponhamos que os eventos $A_1, A_2, \ldots$ são todas as possíveis causas, mutuamente excludentes de um evento B. Sob a suposição de que o evento B tenha sido observado, a fórmula de Bayes permite conhecer qual dessas causas é a mais provável de haver produzido o evento B.

Definição 2 (Distribuição a priori e a posteriori)

Seja $A_1, A_2, \ldots,$ uma partição finita ou enumerável de Ω, e seja $B \in \mathcal{F}$ com $P(B) > 0$. Então $P(A_i)$ são chamadas de probabilidades (distribuições)"a priori", isto é, antes de que aconteça B. As probabilidades $P(A_i|B$ são chamadas de probabilidades (distribuições) "a posteriori", isto significa, depois que aconteceu B.

Uma possível e muito clara regra de decisão diante uma situação de interesse, digamos a presença do evento B se considere como ocorrido o evento A_i como aquele que tem a maior probabilidade sob a hipótese de que o evento B acontece, portanto, elege-se entre os possíveis eventos A_i aquele que dando por acontecido B tem a maior probabilidade de ocorrer. Naturalmente, esta decisão não está eximida de erro, mas pode ajudar a indicar a probabilidade de uma decisão falsa.

Exemplo 4

Numa determinada cidade são feitos testes para detectar uma determinada doença. Suponhamos que 1% das pessoas sadias são registradas como doentes, 0,1% da população está realmente doente e 90% dos doentes são de fato registrados como doentes. Qual é a probabilidade de que um cidadão seja registrado como doente? Qual é a probabilidade de que uma pessoa registrada como doente esteja realmente doente?

Solução: Definamos os seguintes eventos S= “sadio”, E= “realmente doente” e T=“registrado como doente”. Então $P(T|\,S) = 0.01$ e $P(T|\,E) = 0,9$. Assim,

$$P(T) = P(T\cap(S\cup E)) = P(T|\,S)P(S)+P(T|\,E)P(E) = 0,01\times 0,999+0.9\times 0,001 \approx 0,01$$

e

$$P(E|\,T) = P(T|\,E)P(E)/P(T) \approx 0,083.$$

Com estes resultados deduz-se para este exemplo, as distribuições a priori e as distribuições a posteriori são respectivamente, (0,001, 0,999) e (0,083, 0,917).

Exemplo 5

Um curso de economia possui três disciplinas de probabilidade obrigatórias (Prob I, II e III) no núcleo básico. Na primeira disciplina há 50% dos estudantes, na segunda há 25% dos estudantes e na terceira há 25% dos estudantes. Todos eles pertencendo ao núcleo básico. As mulheres (M) se encontram distribuídas uniformemente, sendo que elas constituem 60% dos alunos do núcleo básico. No gráfico da Figura 2 estão representadas as probabilidades para este exemplo.

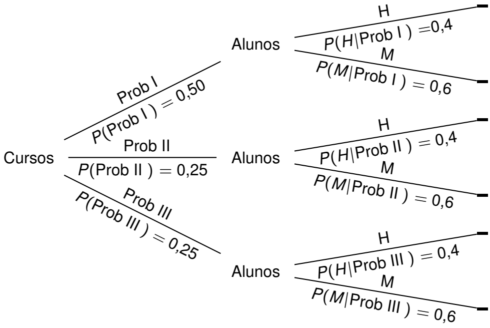

Figura: Diagrama de árvore para o exemplo das disciplinas de probabilidade.

Exemplo 6

Um fabricante produz televisores LED em três fábricas A, B e C, que respondem, respectivamente, por 40%, 35% e 25% de sua produção total. Registros históricos da produção indicam que 2% da produção da fábrica A é defeituosa, assim como 1% da de B, e 3% da fábrica C. Escolhemos um televisor aleatoriamente, e ele é defeituoso. Qual a probabilidade dele ter sido produzido na fábrica B?

Solução: Chamemos por B o evento "fabricado em B" e "*def*" o evento ser defeituoso o qual pode provir de qualquer uma das 3 fábricas (e só de uma!). Logo, os eventos são mutuamente excludentes. Portanto,

$$P(def) = P(A)P(def|A) + P(B)P(def|B) + P(C)P(def|C).$$

Agora,

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

De acordo com os dados fornecidos no problema temos

$$P(def) = (0,40 \times 0,02) + (0,35 \times 0,01) + (0,25 \times 0,03) = 0,019$$

e desta forma,

$$P(B|def) = \frac{0,35 \times 0,01}{(0,40 \times 0,02) + (0,35 \times 0,01) + (0,25 \times 0,03)} = 0,184.$$

A visualização deste problema é simplificada pela visualização do diagramas em árvore apresentado na Figura abaixo

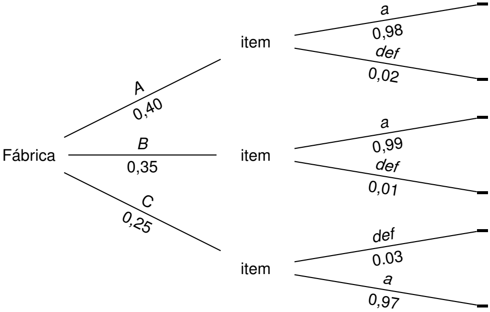

Figura: Diagrama de árvore para o exemplo dos televisores. As siglas *def* indicam defeituosos e *a* não defeituosos.

Exemplo 7

Seja uma imagem com $n \times m$ pixels onde a k-ésima linha contém $d_k(\leq m)$ pixels defeituosos. Escolhe-se uma linha ao acaso e nós não sabemos qual foi a escolha. Em seguida, examina-se um pixel desta linha ao acaso e descobre-se que o pixel é defeituoso (D). Qual a probabilidade de que este pixel defeituoso esteja na linha k?

Solução: Seja $R = k$ se o pixel pertencia a k-ésima linha da imagem. Então, sabe-se que

$$P(R = k) = \frac{1}{n} \quad \text{e} \quad P(D|R = k) = \frac{d_k}{m}.$$

A fórmula de Bayes nos permite determinar

$$P(R = k|D) = \frac{\frac{1}{n}\frac{d_k}{m}}{\sum_{i=1}^{n} \frac{1}{n}\frac{d_i}{m}} = \frac{d_k}{\sum_{i=1}^{n} d_i}.$$

Exercício 2

Jogos do campeonato paulista de futebol ocorrem durante a semana e também nos fins de semana. Suponha que exatamente metade dos jogos ocorram nos fins de semana. Suponha ainda que o São Paulo ganhe 50% dos jogos durante o fim de semana, e perca em 20% de seus jogos no fim de semana. Finalmente, suponha que o São Paulo ganhe todos os jogos que ocorrem durante a semana.

(a) Determine a probabilidade do São Paulo empatar um jogo qualquer.

(b) Dado que o São Paulo ganhou seu último jogo, qual a probabilidade deste jogo ter ocorrido durante a semana?

Exercício 3

Uma urna contém 4 bolas brancas e 6 bolas pretas. Sacam-se, sucessivamente e sem reposição, duas bolas dessa urna. Determine a probabilidade da primeira bola ser branca sabendo que a segunda bola é branca.

Solução: Sejam B_1 e B_2 os eventos a primeira bola é branca e a segunda bola é branca, respectivamente. Queremos calcular $P(B_1|B_2)$. Utilizando a fórmula de Bayes, temos

$$P(B_1|B_2) = \frac{P(B_2|B_1)P(B_1)}{P(B_2|B_1)P(B_1) + P(B_2|B_1^c)P(B_1^c)}.$$

Mas $P(B_2|B_1) = \frac{3}{9}$, $P(B_2|B_1^c) = \frac{4}{9}$, $P(B_1) = \frac{4}{10}$ e $P(B_1^c) = \frac{6}{10}$. Logo,

$$P(B_1|B_2) = \frac{\frac{3}{9} \cdot \frac{4}{10}}{\frac{3}{9} \cdot \frac{4}{10} + \frac{4}{9} \cdot \frac{6}{10}} = \frac{\frac{2}{15}}{\frac{2}{5}} = \frac{1}{3}.$$

Exercício 4

Uma fábrica tem 3 máquinas que produzem o mesmo ítem. As máquinas A e B são responsáveis, cada uma, por 40% da produção. Quanto à qualidade, as máquinas A e B produzem 10% de ítens defeituosos cada uma, enquanto a máquina C apenas 2%. Um ítem é selecionado ao acaso da produção dessa fábrica.

(a) Qual a probabilidade do ítem selecionado ser defeituoso?

(b) Se o ítem selecionado for defeituoso, qual a probabilidade que tenha sido produzido pela máquina A?

Independência

Independência

Intuição

Dois eventos são independentes se eles não têm nada haver um com o outro, eles são totalmente não relacionados; a ocorrência de um não tem nenhuma influência sobre o outro. Por exemplo, resultados de lançamentos sucessivos de uma moeda.

Definição 3

Pode-se usar probabilidades condicionais para formalizar esta intuição da seguinte forma, A é independente de B se $P(A|B) = P(A)$.
Mas usando a definição de probabilidade condicional, chega-se a seguinte conclusão A é independente de B se $P(A \cap B) = P(A)P(B)$.
Como esta última expressão é definida inclusive para o caso de $P(B) = 0$, ela é a expressão adotada como a definição de independência entre eventos.

Definição 4

O evento A é independente do evento B se $P(A \cap B) = P(A)P(B)$.

Observações

Note que esta definição de independência implica que independência é um conceito simétrico em teoria da probabilidade, isto é, A é independente de B se e somente se B é independente de A. Note que esta definição também implica que eventos A e B são independentes se $P(A) = 0$ ou $P(B) = 0$, o que pode gerar algumas conclusões não intuitivas se de fato $P(A) = 0$ ou $P(B) = 0$. Por exemplo, se $P(A) = 0$, então A é independente dele mesmo, porém A certamente não é não relacionado consigo mesmo.

Teorema 4

A é independente dele mesmo se e somente se $P(A) = 0$ ou $P(A) = 1$.

Demonstração.

$$\begin{aligned} & P(A \cap A) = P(A) = P(A)P(A) \\ \Leftrightarrow\ & P(A) = 0 \text{ ou } P(A) = 1. \end{aligned}$$

□

Exemplo 8

Um dado honesto é atirado ao ar duas vezes consecutivas. Definamos os eventos $A =$ "a soma dos resultados obtidos é um número par" e $B =$ " o segundo lançamento resulta ser um número par". Para este caso temos que $P(A) = P(B) = 1/2$ e $P(A \cap B) = 1/4$. Portanto A e B são eventos independentes.

Exemplo 9

Consideremos novamente o lançamento dos dados. Mas agora carregamos o dado de tal maneira que a probabilidade de obter um número par é 2/5. Se consideramos os eventos do exemplo anterior temos que $P(B) = 2/5$, $P(A) = 13/25$ e $P(A \cap B) = 4/25$. Assim, A e B não são independentes.

Propriedades

Intuitivamente, se A é independente de B o fato que B não ocorreu, ou seja que B^c ocorreu, não deve alterar a probabilidade de A. Portanto, é de se esperar que se A e B são independentes, então A e B^c também são. O seguinte teorema prova que esta intuição é verdadeira.

Teorema 5

Se A e B são eventos independentes, A e B^c (resp., A^c e B, A^c e B^c) também o são.

Note que

$$A = A \cap \Omega = A \cap (B \cup B^c) = (A \cap B) \cup (A \cap B^c).$$

Então, como $A \cap B$ e $A \cap B^c$ são mutuamente exclusivos, axioma K3 implica que

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

Como A e B são independentes, nós temos

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

Rearranjando os termos e utilizando o fato que $P(B^c) = 1 - P(B)$, temos $P(A \cap B^c) = P(A)P(B^c)$.

Demonstração.

Note que

$$A = A \cap \Omega = A \cap (B \cup B^c) = (A \cap B) \cup (A \cap B^c).$$

Então, como $A \cap B$ e $A \cap B^c$ são mutuamente exclusivos, axioma K3 implica que

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

Como A e B são independentes, nós temos

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

Rearrajando os termos e utilizando o fato que $P(B^c) = 1 - P(B)$, temos
$P(A \cap B^c) = P(A)P(B^c)$. $\square$

Coleção de Eventos

O conceito de independência também se aplica a uma coleção arbitrária de eventos $\{A_i\}_{i\in\mathcal{I}}$, onde $\mathcal{I}$ é um conjunto de índices. Neste caso, têm-se duas definições.

Definição 5

Uma coleção de eventos $\{A_i\}_{i\in\mathcal{I}}$ *é* independente par a par *se para todo* $i \neq j \in I$, A_i *e* A_j *são eventos independentes.*

Definição 6

Uma sequência finita de eventos
$A_1, A_2, \ldots, A_n$, $n \geq 1$, *é* mutuamente independente *se para todo* $I \subseteq \{1, \ldots, n\}$,

$$P(\cap_{i\in I} A_i) = \prod_{i\in I} P(A_i)$$

E uma coleção qualquer de eventos $\{A_i\}_{i\in\mathcal{I}}$ *é mutuamente independente se para todo* $J \subseteq \mathcal{I}$ *finito,* $\{A_i\}_{i\in J}$ *é mutuamente independente.*

Exemplo 10

Consideremos os seguintes eventos relacionados com o lançamento de um dado corrente duas vezes consecutivas. Definamos os eventos $A =$ " no primero lançamento é obtido um dois", $B =$ " no segundo lançamento é obtido um cinco" e $C =$ " a soma dos resultados em ambos os lançamentos é sete ".
Neste caso, $P(A) = P(B) = P(C) = 1/6$. Por outro lado, $P(A \cap B) = P(A \cap C) = P(B \cap C) = P((2,5)) = 1/36$. Portanto, satisfaz-se que $P(A \cap B) = P(A)P(B)$, $P(A \cap C) = P(A)P(C)$ e $P(B \cap C) = P(B)P(C)$. Contudo

$$P(A \cap B \cap C) = P((2,5)) \neq P(A)P(B)P(C).$$

Exemplo 11

Se $\Omega = \{1,2,3,4\}$ e $P(\{w\}) = 1/4$, então $A = \{1,2\}$, $B = \{1,3\}$, e $C = \{2,3\}$ são eventos independentes par a par. Pode-se verificar isto pelo fato que

$$P(A \cap B) = P(\{1\}) = 1/4 = 1/2 \cdot 1/2 = P(A)P(B).$$

Similarmente, pode-se provar o mesmo resultado para os outros pares. Contudo, a probabilidade

$$P(A \cap B \cap C) = P(\emptyset) = 0 \neq P(A)P(B)P(C) = 1/8.$$

Então, A, B, e C não são mutuamente independentes.

Exemplo 12

Certo experimento consiste em lançar um dado equilibrado duas vezes, independentemente. Dado que os dois números sejam diferentes, qual é a probabilidade condicional de

(a) pelo menos um dos números ser 6,

(b) a soma dos números ser 8?

Solução: Para parte (a), note que existem 30 resultados possíveis para os lançamentos do dado de modo que o mesmo número não se repita, dos quais 10 o número 6 ocorre. Portanto, esta probabilidade é igual a 1/3.

Para parte (b), note que existem 4 resultados possíveis que somam 8 dado que os números são diferentes, logo esta probabilidade é igual a 4/30.

Exemplo 13

Suponha que um determinado experimento é realizado repetidas vezes de forma independente e observa-se a ocorrência de determinado evento A que tem probabilidade p. Qual é a probabilidade que A ocorra n vezes antes de A^c ocorrer m vezes?

Solução: Note que o evento A ocorra n vezes antes de A^c ocorrer m vezes é equivalente ao evento A ocorrer pelo menos n vezes nas primeiras $n+m-1$ repetições do experimento. Como a ordem de ocorrência do evento A nas repetições não é importante e as repetiç es são independentes, temos que o evento A ocorre k vezes em $n+m-1$ repetições do experimento tem probabilidade igual a:

$$P(k \text{ ocorrências de } A \text{ em } n+m-1 \text{ repetições})$$
$$= \binom{n+m-1}{k} p^k (1-p)^{n+m-1-k}.$$

e, então,

$$P(n \text{ ocorrências de } A \text{ antes de } m \text{ ocorrências de } A^c)$$
$$= \sum_{k=n}^{n+m-1} \binom{n+m-1}{k} p^k (1-p)^{n+m-1-k}.$$

Exercício 5

Assuma que $A_1, \ldots, A_n$ são eventos mutuamente independentes e que $P(A_i) = p_i$. Nós calculamos as probabilidades dos seguintes eventos:

- O evento A é o evento que todos estes eventos ocorrem, então

$$P(A) = P(\cap_{i=1}^{n} A_i) = \prod_{i=1}^{n} P(A_i) = \prod_{i=1}^{n} p_i$$

- O evento B é o evento que nenhum desses eventos ocorre, então

$$P(B) = P(\cap_{i=1}^{n} A_i^c) = \prod_{i=1}^{n} P(A_i^c) = \prod_{i=1}^{n} (1 - p_i)$$

- O evento C é o evento que pelo menos um desses eventos ocorre, então $C = B^c$

$$P(C) = P(B^c) = 1 - P(B) = 1 - \prod_{i=1}^{n} (1 - p_i)$$

Exemplo 14

João e José disputam um jogo com uma moeda equilibrada. Cada jogador lança a moeda duas vezes e vence o jogo aquele que primeiro obtiver dois resultados iguais. João começa jogando e se não vencer passa a moeda para José e continuam alternando jogadas. Qual a probabilidade de João vencer o Jogo?

Solução: Seja A_k o evento dois resultados iguais são obtidos na k-ésima tentativa. Note que $P(A_k) = \frac{1}{2}$. Seja B_k o evento João ganha na sua k-ésima jogada. Então,

$$B_1 = A_1;\ B_2 = A_1^c \cap A_2^c \cap A_3;\ B_3 = A_1^c \cap A_2^c \cap A_3^c \cap A_4^c \cap A_5,$$

em geral,

$$B_k = A_1^c \cap A_2^c \cap \cdots \cap A_{2k-2}^c \cap A_{2k-1}.$$

Portanto,

$$P(B_k) = P(A_1^c \cap A_2^c \cap \cdots \cap A_{2k-2}^c \cap A_{2k-1})$$

$$= P(A_1^c)P(A_2^c)\cdots P(A_{2k-2}^c)P(A_{2k-1}) = (\frac{1}{2})^{2k-1},$$

onde a penúltima igualdade se deve ao fato dos lançamentos serem independentes. Logo,

$$P(\text{João vencer}) = P(\cup_{k=1}^{\infty} B_k) = \sum_{k=1}^{\infty} P(B_k)$$

$$= \sum_{k=1}^{\infty} (\frac{1}{2})^{2k-1} = \frac{2}{3}.$$