3.1.5. Cálculo de probabilidades usando teoria dos conjuntos

3.1.5.1. Introdução
(i) Porque usar *conjuntos*
(ii) Definições e notação
(iii) Tipos de espaço amostral
(iv) Diagrama de Euler-Venn
3.1.5.2. Relações entre conjuntos: complemento, união, exclusão, interseção,
3.1.5.3. Probabilidades das relações entre conjuntos
(i) Probabilidade da união
(ii) Probabilidade da interseção; probabilidade condicional
(iii) Probabilidade do evento complementar
(iv) Probabilidade da exclusão
3.1.5.4. Correspondência entre Lógica, Álgebra e Probabilidades

---

3.1.5.1. Introdução

(i) Porque usar conjuntos

Até o início do século XX, a Teoria de Probabilidades ainda não era uma área muito respeitável da Matemática. Apesar dos avanços que tinha conseguido nas suas aplicações (por exemplo, na indústria dos seguros), suas bases teóricas ainda estavam longe de serem sólidas. Não havia relação lógica entre as definições propostas para a "probabilidade" (veja as definições *clássica* e *frequencista* na seção **3.1.1.3**) e os métodos sugeridos para resolver os problemas; estes métodos pareciam ser apenas um conjunto meio desorganizado de regras empíricas, algumas das quais poderiam ser úteis para resolver alguns tipos de problemas, mas não outros tipos. Para os matemáticos, isto era deplorável - um conjunto de regras que funcionam bem, mas não são logicamente encadeadas, pode ser muito útil para a vida prática, mas não merece o nome de "Matemática".

No início do século XX, os matemáticos passaram a exigir mais rigor na teoria. Uma das obras fundamentais foi o livro de Russell e Whitehead [[1]], que se propuseram a reorganizar de forma lógica toda a Matemática. Para as Probabilidades, a situação começou a mudar no intervalo entre as duas grandes guerras, quando foram publicados os artigos e livros que definiram os rumos que a área de Probabilidades segue até hoje (Stigler [[2]] argumenta que a *Estatística Matemática* só surgiu em 1933). A base foi dada pelos trabalhos do matemático russo Kolmogorov, que estabeleceram a definição *axiomática* de probabilidade, e uma nova reorganização da matéria usando os conceitos e símbolos da teoria de conjuntos. Este novo formalismo, embora a primeira vista possa parecer que tenha complicado as coisas, tornou na verdade a notação do cálculo de probabilidades mais clara e logicamente coerente.

Introduziremos neste capítulo o tratamento da teoria das Probabilidades que se baseia no conceito de *conjunto*. Existe ainda uma outra abordagem da Probabilidade, baseada na definição *subjetiva*; é a chamada corrente *bayesiana*, que diverge consideravelmente da abordagem "clássica" que temos estudado, e não será vista neste capítulo.

(ii) Definições e notação

Antes de iniciar o cálculo de probabilidades por meio de conjuntos, é preciso definirmos alguns termos:

- Qualquer ação que produza um resultado aleatório observável é chamada de *experimento*; por exemplo, lançar um dado e verificar o número mostrado, ou sortear uma pessoa e medir o seu peso.
- Cada resultado possível de um experimento é um *ponto amostral.*
- O conjunto dos pontos amostrais forma o *espaço amostral.*
- Qualquer subconjunto destes pontos forma um *evento.*

Os espaços amostrais podem ser *finitos* ou *infinitos*; os infinitos podem ser *discretos* ou *contínuos* (também chamados de *enumeráveis* ou *não-enumeráveis*) .

Exemplo 1
Experimento: lançar um dado e anotar o número *X* mostrado.
Espaço amostral: o conjunto *S* de seis pontos amostrais que correspondem às faces do dado,
S = {1,2,3,4,5,6}
Neste espaço definimos por exemplo os *eventos* A e B:
A = {conjunto dos números menores que 3} = $\{X \mid X < 3\} = \{1,2\}$
B = {conjunto dos números pares} = {X | X é par} = {2,4,6}

Exemplo 2
Experimento: retiramos aleatoriamente uma carta de um baralho comum, e anotamos seu naipe.
Espaço amostral: o conjunto *S* de 52 cartas de um baralho
S = {A♣, K♣, Q♣, J♣, 10♣, ..., 2♣, A♥, K♥, Q♥, J♥, 10♥, ..., 2♥, etc.}
Neste espaço definimos por exemplo o *evento* A:
A = {cartas do naipe ♣} = { A♣, K♣, Q♣, J♣, 10♣, ..., 2♣}

Exemplo 3
Experimento: Testamos um isqueiro, acendendo-o repetidamente até que ele pare de funcionar, e anotamos o número *X* de repetições que fizemos.
Espaço amostral: o conjunto dos números inteiros $S = \{X \mid X \in N\} = \{1,2,3,4...\}$.
Neste espaço, definimos por exemplo o *evento A*:
$A = \{ X \mid X > 10\} = \{11,12,13,14,...\}$

Exemplo 4
Experimento: Sorteamos uma pessoa e medimos seu peso.
Espaço amostral: O conjunto dos números reais positivos, $S=\{ X \mid X \in \Re+\}$.
Neste espaço, definimos por exemplo o *evento A*:
$A = \{X \mid 50 < X < 60\}$

(iii) Tipos de espaços amostrais

Os espaços amostrais dos Exemplos 1 e 2 são do tipo *discreto*: conjuntos finitos de valores que podem ser *enumerados,* isto é, podem ser colocados na forma de enumerações, ou listas. Foi a partir de problemas com espaços amostrais deste tipo que foi desenvolvida toda a teoria clássica de Probabilidades, nos séculos XVII e XVIII. O espaço amostral do Exemplo 3 é *infinito discreto*: o número de tentativas feitas pode ainda ser representado em uma lista de números inteiros, embora esta lista seja infinita – não sabemos qual será o número máximo de tentativas.

O espaço amostral do Exemplo 4 é *contínuo*: o peso de uma pessoa é dado por um número real, e este número pode assumir infinitos valores; estes valores não podem ser enumerados, porque entre quaisquer dois valores sempre existem infinitos outros: entre 50kg e 51kg, existe por exemplo 50,5kg; entre 50,5g e 51kg, existe por exemplo 50,6 kg; entre 50,000kg e 50,001kg existe 50,0005kg, etc. Em teoria, sempre podemos subdividir infinitamente qualquer intervalo, e podemos medir o peso com qualquer número de decimais que quisermos.

Esta distinção entre tipos de espaços amostrais será importante quando estudarmos *variáveis aleatórias*, na seção 3.2, pois teremos então que usar definições e modelos diferentes para as variáveis *discretas* e para as *contínuas*. Na prática, às vezes estes tipos de espaços amostrais se confundem; por exemplo, quando pesamos pessoas adultas, em geral arredondamos as medições e somente consideramos os valores inteiros em quilogramas, o que transformaria o espaço em *infinito discreto*; na teoria, contudo, os dois tipos de espaço devem ser tratados de formas diferentes.

(iv) Diagramas de Euler-Venn

Uma vez que eventos e espaços amostrais são conjuntos, podemos representá-los graficamente por meio dos conhecidos *diagramas de Euler-Venn*. Esta representação pode ser muito útil para ajudar a resolver problemas de probabilidade.

Para o Exemplo 1 acima (lançamento de um dado), podemos representar o espaço amostral S e os dois eventos A e B pelo diagrama da Fig. 1; os pontos amostrais estão representados por pequenos círculos.

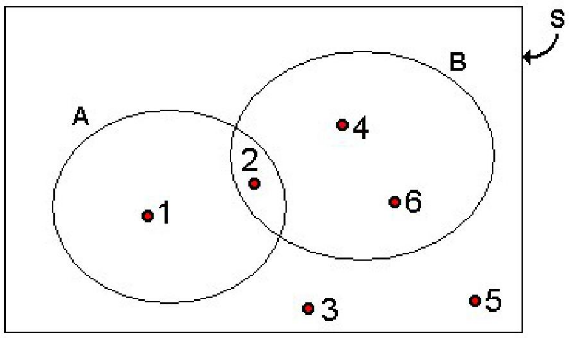


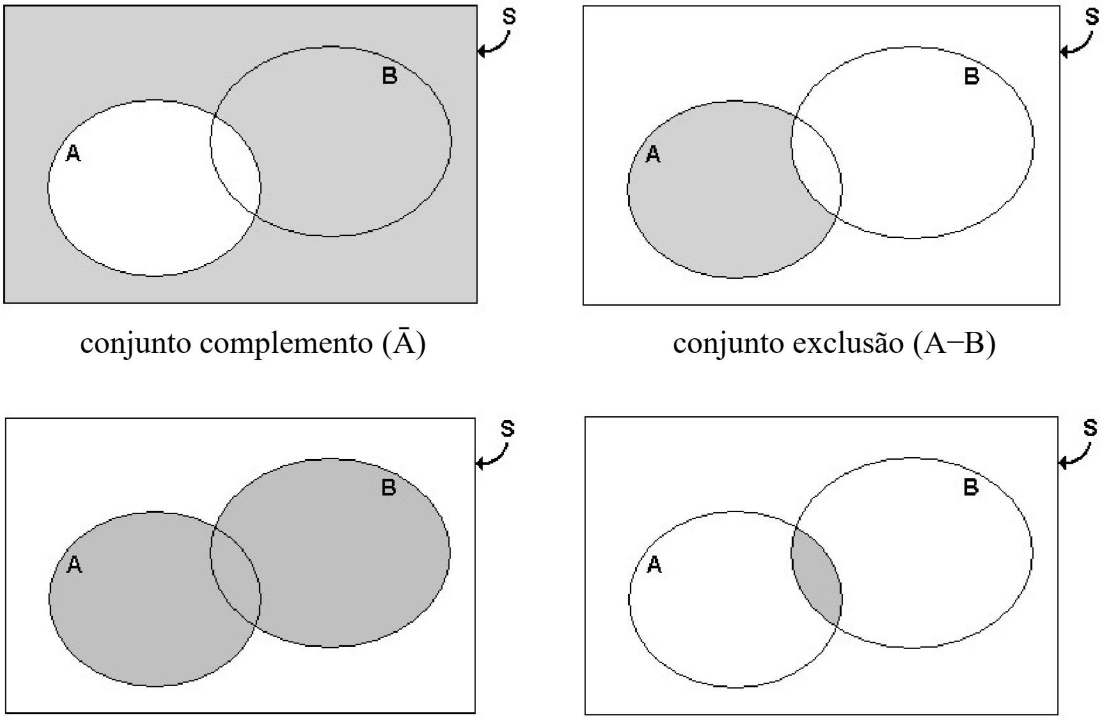
Figura 1. Diagrama de Euler-Venn para o Exemplo 1.

3.1.5.2. Relações entre dois conjuntos: complemento, união, exclusão e interseção

Dados dois conjuntos A e B, eles podem ser relacionados de forma a produzir quatro novos conjuntos:

nome	notação	descrição
complemento	$\bar{A}$	conjunto de pontos que não pertencem a A
união	$A \cup B$	conj. de pontos que pertencem a A, ou a B, ou a ambos
exclusão	$A - B$	conj. de pontos que pertencem a A, mas não a B
interseção	$A \cap B$	conj. de pontos que pertencem a A e a B

Note que o sinal “−” é usado aqui para indicar *exclusão* de dois conjuntos, não a *diferença* entre dois números. Os quatro conjuntos resultantes estão representados na cor cinza nos diagramas de Euler-Venn da Fig. 2.



conjunto complemento ($\bar{A}$) conjunto exclusão (A−B)

conjunto união (A∪B) conjunto interseção (A∩B)

Figura 2. Diagramas de Euler-Venn das relações possíveis entre dois conjuntos

3.1.5.3. Probabilidades das relações entre conjuntos

Veremos agora como esta representação dos eventos aleatórios por meio de *conjuntos* está relacionada com o cálculo de probabilidades. Esta notação já foi usada quando apresentamos a definição *axiomática* de probabilidade; reproduzimos novamente esta definição no quadro abaixo, porque suas propriedades (3) e (4) são necessárias para o cálculo das probabilidades.

Seja um E um experimento, e S o espaço amostral a ele associado. A cada evento A associaremos um número real representado por $P(A)$ e denominado *probabilidade de A*, que satisfaça às seguintes propriedades:

1) $0 \leq P(A) \leq 1$
2) $P(S) = 1$
3) Se A e B forem eventos mutuamente excludentes,
   $P(A \cup B) = P(A) + P(B)$
4) Se $A_1, A_2, A_3, \dots A_n$ forem eventos mutuamente excludente dois a dois s, então
   $P(A_1 \cup A_2 \cup A_3 \cup \dots \cup A_n) = P(A_1) + P(A_2) + P(A_3) + \dots + P(A_n)$

(i) Probabilidade da união de dois conjuntos

(a) Quando os eventos são mutuamente excludentes

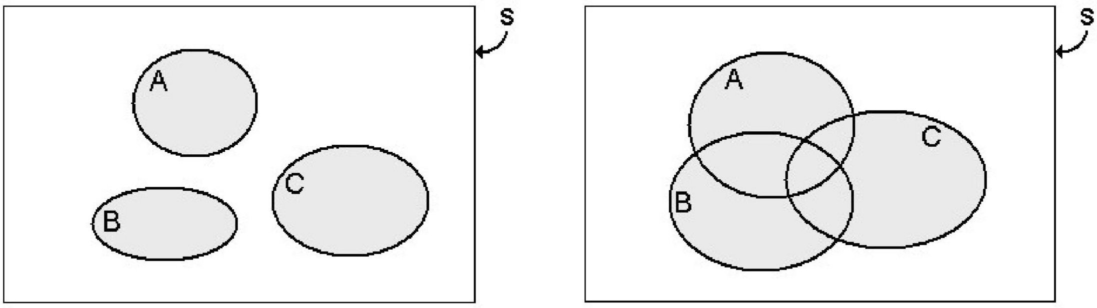
Se dois eventos A e B são representados por conjuntos que não têm interseção entre si, são chamados de eventos "mutuamente excludentes". Se queremos calcular a probabilidade de que ocorra *ou o evento A ou o evento B*, queremos a probabilidade do evento formado pela *união* de A e B; esta probabilidade será calculada pela soma das probabilidades de A e de B:

$$P(A \cup B) = P(A) + P(B) \tag{1}$$

Esta regra pode ser estendida para mais de dois eventos. Se são *n* eventos mutuamente excludentes, representados por $A_1, A_2, \ldots, A_n$, então

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n) \tag{2}$$

A Fig.3 mostra, como exemplo, dois casos possíveis de união de três conjuntos.



A. Três eventos mutuamente excludentes

B. Três eventos **não** mutuamente excludentes

Figura 3 – União de três conjuntos

Exemplo

Experimento: retiramos aleatoriamente uma carta de um baralho comum (veja Fig. 4). Qual são as probabilidades de que esta carta:

(i) seja de um naipe "vermelho" – *ouros* (O) ou *copas* (C)?

(ii) seja uma figura (rei K, dama D ou valete J)?

Os naipes formam eventos mutuamente excludentes (uma carta não pode pertencer a dois naipes ao mesmo tempo). Por isso, as probabilidades pedidas são:

(i) P(*ouros* ou *copas*) $= P(O \cup C) = P(O) + P(C) =$

$= 13/52 + 13/52 = 1/2$

(ii) P(K ou Q ou J) $= P(K \cup Q \cup J) = P(K) + P(Q) + P(J) =$

$= 4/52 + 4/52 + 4/52 = 12/52 = 3/13$

Note que estas duas regras nas eqs. (1) e (2) não são *teoremas* que tenham que ser demonstrados, e sim *axiomas* - isto é, afirmativas que são aceitas como base de uma teoria, e a partir das quais se demonstram os teoremas (veja a definição axiomática de Probabilidade, acima: estas regras são as propriedades de números 3 e 4).

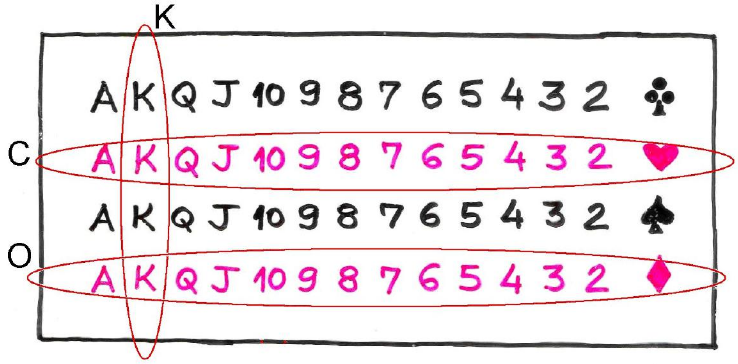


Figura 4 – Baralho completo, com três eventos definidos

(b) Quando os eventos não são mutuamente excludentes

Se os eventos A e B *podem* ocorrer juntos, os conjuntos que os representam têm interseção. Neste caso, a probabilidade da união entre A e B será dada por

$$P(A\cup B) = P(A) + P(B) - P(A\cap B) \qquad (3)$$

Na prática, isto significa: se queremos calcular a probabilidade de que ocorra *ou o evento A, ou o evento B ou ambos juntos*, somamos a probabilidade de A com a de B e subtraírmos a probabilidade da interseção. Note que a regra em (1) é apenas um caso particular da regra geral na eq. (3), quando fazemos $P(A\cap B) = 0$.

Exemplo:
Experimento: Retirarmos uma carta aleatoriamente de um baralho completo. Quais são as probabilidades de que seja um *Rei* (K) de qualquer naipe ou qualquer carta do naipe de *Copas* (C)? (Fig. 4)

Os eventos agora não são mutuamente excludentes, e a probabilidade desejada é:

$$P(K\cup C) = P(K) + P(C) - P(K\cap C)$$
$$= 4/52 + 13/52 - 1/52 = 16/52 = 4/13 = 0{,}3077$$

Ao contrário da regra anterior, esta é um *teorema* que exige demonstração. Não veremos esta demonstração, mas é fácil entender intuitivamente a lógica da regra, consultando o diagrama de Venn na Fig. 4: se calcularmos P(K) contando os reis, e P(C) contando as cartas de copas, acabaremos contando duas vezes a carta “rei de copas”, que está na interseção; por isso, temos que subtrair uma vez sua probabilidade .

Esta regra pode ser estendida para o caso de mais de dois eventos, embora a notação fique um pouco complicada. No caso de três eventos A, B e C (Fig. **3B**), a probabilidade da união dos três será dada por:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A\cap B) - P(B\cap C) + P(A\cap B\cap C) \quad (4)$$

Em resumo:
As regras em (1), (2) e (3) costumam ser chamadas de “regras da adição das probabilidades”. Estas regras dizem que a probabilidade da união de conjuntos é dada pela soma das probabilidades destes conjuntos:

Para eventos mutuamente excludentes:	$P(A\cup B) = P(A) + P(B)$
Para eventos não mutuamente excludentes:	$P(A\cup B) = P(A) + P(B) - P(A\cap B)$

(ii) Probabilidade da interseção de dois eventos

Para calcular a probabilidade de dois eventos ocorrerem juntos (isto é, de ocorrer a *interseção* destes dois eventos), precisamos primeiro verificar se estes eventos são *dependentes probabilisticamente*; se forem, devemos calcular suas *probabilidades condicionais*. Estes conceitos serão discutidos abaixo.

(a) Dependência probabilística

O idéia de *dependência probabilística* pode ser apresentado através de um exemplo.

Exemplo
Experimento: A Tabela **1** mostra os resultados do levantamento do número de casos de daltonismo numa amostra de 1000 pessoas (dados fictícios). Uma pessoa é escolhida aleatoriamente na amostra.
a) Qual é a probabilidade de que seja daltônica?
b) Se esta pessoa for uma mulher, qual é a probabilidade de que seja daltônica?
c) Se esta pessoa for um homem, qual é a probabilidade de que seja daltônico?

Tabela 1. Casos de daltonismo numa amostra

	Sexo		
Daltonismo	Masc.(M)	Fem.(F)	Total
Presente (D+)	36	12	48
Ausente (D-)	364	588	952
Total	400	600	1000

A probabilidade pedida no item (a) é, usando a definição clássica:
P(D+) = 48/1000 = 0,048

A probabilidade pedida no item (b) é uma *probabilidade condicional*, representada como
P(D+|F)

Esta probabilidade é chamada de *probabilidade condicional de a pessoa ser daltônica, dado que ela é do sexo feminino*; ou, mais resumidamente, *probabilidade de D+ dado F* (a barra vertical é lida como "dado que"). Esta probabilidade pode ser calculada pela razão entre o número de mulheres daltônicas e o número de total mulheres na amostra:
P(D+|F) = 12/600 = 0,02

A probabilidade pedida no item (c) é a *probabilidade condicional de a pessoa ser daltônica, dado que ela é do sexo masculino*, representada por:
P(D+|M)

Esta probabilidade pode ser calculada também usando a definição clássica:
P(D+|M) = 36/400 = 0,09

Observe que, na amostra acima, estas duas probabilidades condicionais são diferentes:
P(D+|F) = 0,02
P(D+|M) = 0,09

A probabilidade *não-condicional* de uma ser daltônica nesta amostra foi calculada acima, sem levar em conta o sexo da pessoa:
P(D+) = 48/1000 = 0,048

Note que estas três probabilidades de D+ são diferentes entre si. Neste caso, dizemos então que *Daltonismo* e *Sexo* são variáveis "probabilisticamente dependentes". Isto é: a probabilidade de ocorrer um valor da variável *Daltonismo* (D+ ou D−) depende do que ocorreu com a variável *Sexo* (M ou F).

O Exemplo 2 mostra uma situação diferente.

Exemplo 2:
Experimento: A Tabela 2 mostra os resultados de um levantamento de número de casos de resfriado numa amostra de 600 pessoas. Uma pessoa é escolhida aleatoriamente nesta amostra.
a) Qual é a probabilidade de que ela esteja resfriada?
b) Se esta pessoa for uma mulher, qual é a probabilidade de que ela esteja resfriada?
c) Se esta pessoa for um homem, qual é a probabilidade de que ele esteja resfriado?

Tabela 2. Ocorrência de resfriado, por sexo

	Sexo		
Resfriado	M	F	Totais
Presente (R+)	80	20	100
Ausente (R−)	400	100	500
Totais	480	120	600

As probabilidades pedidas são

(a) $P(R+) = 100/600 = 1/6$
(b) $P(R+ \mid F) = 20/120 = 1/6$
(c) $P(R+ \mid M) = 80/480 = 1/6$

Note que as três probabilidades de R+ são iguais:

$$P(R+ \mid F) = P(R+ \mid M) = P(R+) = 1/6$$

Dizemos então que as variáveis *Sexo* e *Resfriado* são *probabilisticamente independentes*.

Em resumo**:** dois eventos são ditos probabilisticamente *dependentes* quando a ocorrência de um deles afeta a probabilidade de ocorrência do outro (como no Ex. 1, no qual a probabilidade de uma pessoa ser daltônica depende do seu sexo); são ditos *independentes*, no caso contrário (como no Ex. 2, no qual a probabilidade de uma pessoa estar resfriada não depende de seu sexo).

(b) Cálculo da probabilidade da interseção

Formalmente, a probabilidade condicional $P(B \mid A)$ é *definida* em função da probabilidade da interseção, na forma:

$$P(B \mid A) = P(A \cap B) / P(A) \qquad (5)$$

Note que a expressão em (5) é dada como uma *definição*, não como um teorema que deva ser demonstrado. A probabilidade da interseção de dois eventos, portanto, pode ser calculada como:

$$P(A \cap B) = P(A) \,.\, P(B/A) \qquad (6)$$

Se as probabilidades condicionais são iguais entre si, e iguais à probabilidade não-condicional:

$$P(B \mid A) = P(B \mid \bar{A}) = P(B)$$

os dois eventos A e B são *independentes,* e a probabilidade da interseção é dada por:

$$P(A \cap B) = P(A) \,.\, P(B) \qquad (7)$$

Exemplo: (eventos independentes)
Se lançarmos duas moedas, qual é a probabilidade de que ambas mostrem *caras*?
Representando os eventos por:
C_1: *cara* na primeira moeda
C_2: *cara* na segunda moeda
$P(C_1 \cap C_2) = P(C_1) \times P(C_2 | C_1) = P(C_1) \times P(C_2) = ½ \times ½ = ¼$

Exemplo: (eventos dependentes)
Se retirarmos duas cartas de um baralho, sem reposição, qual é a probabilidade de que ambas sejam *ases* ?
Representando os eventos por:
A_1: *ás* na primeira carta
A_2: *ás* na segunda carta
$P(A_1 \cap A_2) = P(A_1) \times P(A_2 | A_1) = 4/52 \times 3/51 = 1/221 = 0{,}0045$

Em resumo:
As regras em (5) e (6) costumam ser chamadas de "regras do produto das probabilidades". Estas regras dizem que:
- se quisermos calcular a probabilidade que os dois eventos *independentes* A e B aconteçam juntos, simplesmente multiplicamos suas probabilidades.

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

- se quisermos calcular a probabilidade que os dois eventos *dependentes* A e B aconteçam juntos, multiplicamos a probabilidade de um deles pela *probabilidade condicional* do outro.

$$P(A \cap B) = P(A) \times P(B|A) \;\text{ ou }\; P(A \cap B) = P(B) \times P(A|B)$$

(iii) Probabilidade do evento complementar

Dado um evento A, a probabilidade de seu evento complementar $\bar{A}$ será dada por:

$$P(\bar{A}) = 1 - P(A) \qquad (8)$$

O evento complementar $\bar{A}$ é o conjunto dos pontos que não pertencem a um conjunto A; corresponde à *negação* em Lógica. Se queremos que um evento *não* ocorra, queremos que seu evento *complementar* ocorra. Por exemplo, na Tabela **1**, se *D* é conjunto das pessoas que são daltônicas, então $\overline{D}$ é o conjunto das pessoas que *não* são daltônicas. A partir da eq. (8), podemos calcular a probabilidade de uma pessoa sorteada *não* ser daltônica desta forma:

Probabilidade de ser daltônica: $P(D) = 48/1000$
Probabilidade de não ser daltônica: $P(\overline{D}) = 1 - P(D) = 1 - 0{,}048 = 0{,}952$

Neste exemplo, é claro, poderíamos ter calculado $P(\overline{D})$ diretamente da tabela, ao invés de usar a eq. (**8**); mas veremos nos exercícios alguns casos nos quais é mais fácil calcular $P(\bar{A})$ e depois usar a eq. (8), do que tentar calcular P(A) diretamente.

Formalmente, os eventos A e $\bar{A}$ são ditos *complementares* se

$$A \cap \bar{A} = \phi \qquad \text{e} \qquad A \cup \bar{A} = S$$

onde ϕ é o conjunto vazio, e S o espaço amostral. Note que a notação do conjunto complementar pode variar de livro para livro; além do símbolo $\bar{A}$ também costumam ser usados os símbolos A' ou A^c.

(iv) Probabilidade da exclusão de conjuntos

Dados dois conjuntos A e B, a probabilidade do conjunto *exclusão* (representado por A−B) será dada por:

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

Este resultado é um teorema, mas pode ser facilmente compreendido, sem demonstração, se olhamos o diagrama de Euler-Venn na Fig. 5.

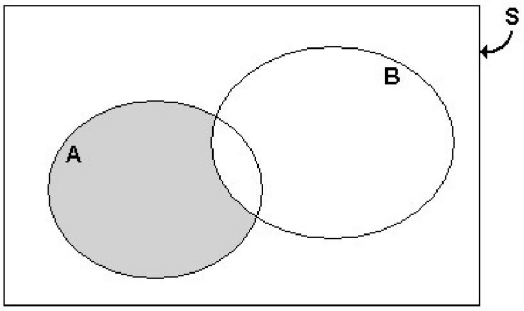


Figura 5. Conjunto exclusão A−B

Tome cuidado com a notação: lembre-se de aqui que o símbolo "−" não significa *subtração* (de números) e sim *exclusão* (de conjuntos). O fato de o mesmo símbolo ser usado para as duas operações pode causar alguma confusão. Note que podemos escrever

$$P(A) + P(B), \text{ ou } P(A) - P(B)$$

mas não $\quad P(A) \cup P(B)$, ou $P(A) \cap P(B)$

porque P(A) e P(B) são números, e não conjuntos; por outro lado, podemos escrever

$$A\cup B \text{ ou } A\cap B \text{ ou } A-B$$

mas não $\quad A+B$

porque A e B são conjuntos, não números, e não definimos a operação "+" para conjuntos.

3.1.5.4. Correspondência entre Lógica, Álgebra e Probabilidades

Como as seções anteriores devem ter sugerido, existem algumas relações entre as operações lógicas, as operações de conjunto, e as probabilidades.

As conjunções *ou* e *e* usadas na lógica correspondem às operações de *união* e *interseção*, que são calculadas por *soma* e *produto*, respectivamente. Se queremos que ocorram os eventos mutuamente excludentes A ***ou*** B, queremos que ocorra o evento A∪B, cuja probabilidade é dada pela soma $P(A\cup B) = P(A) + P(B)$; se queremos que ocorram os eventos independentes A ***e*** B, queremos que ocorra o evento $A \cap B$, cuja probabilidade é dada pelo produto $P(A \cap B) = P(A) \times P(B)$. Outra relação importante existe entre a *negação* lógica e o *complemento* de um conjunto: se quero que A não ocorra, quero que $\bar{A}$ ocorra.

Estas correspondências estão mostradas na Tab. 3; é bom tê-las em mente, pois servem de orientação quanto temos que resolver problemas de probabilidade.

Tabela 3. Correspondência entre as operações

Operações lógicas	Operações com conjuntos	Probabilidades
ou	união	soma
e	interseção	produto
não	complemento	subtração

Referência
[1] Russell, B. e Whitehead, A. N. *Principia Mathematica*. Cambridge: Cambridge Univ. Press, 1910-1912.
[2] Stigler, Stephen M. *Statistics on the Table.* Harvard Univ. Presss, 1999.