Probabilidades e Aplicações

CC e MAT

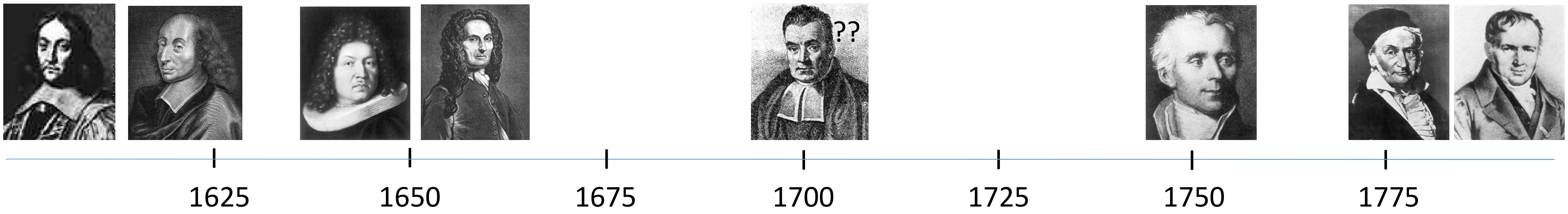
2020/2021

2T + 1TP

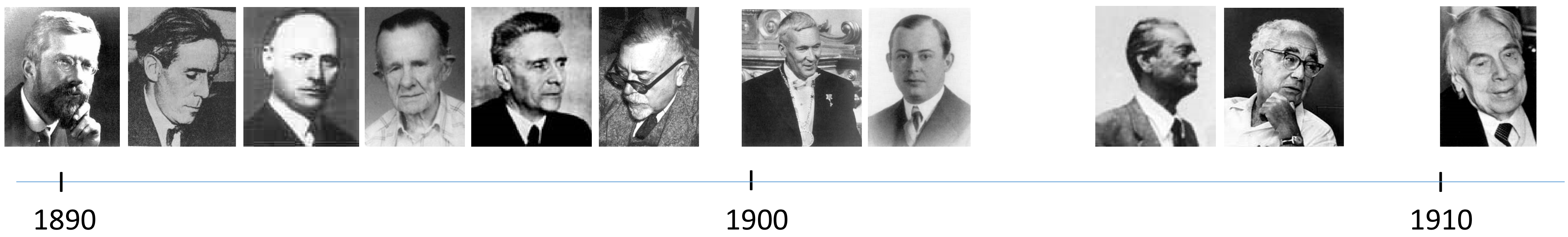
Emilia Athayde
mefqa@math.uminho.pt

Programa

1. Axiomática da probabilidade. Teorema da probabilidade total.
2. Variáveis aleatórias e vetores aleatórios.
3. Medidas de localização, dispersão e forma. Momentos e desigualdades. Transformadas.
4. Distribuições univariadas e multivariadas mais comuns. Condicionamento e independência.
5. Funções de variáveis aleatórias.
6. Processo de Poisson.
7. Distribuição normal e suas propriedades.
8. Convergências estocásticas.
9. Teorema limite central e lei dos grandes números.
10. Simulação.


??
1625
1650
1675
1700
1725
1750
1775


1820
1840
1860
1880


1890
1900
1910

As primeiras obras

Girolamo Cardano (1501-1576): *Liber de ludo aleae*, 1663.

Galileo Galilei (1564-1642): *Sopra le scoperte dei dadi*, 1612.

John Graunt (1620-1674): *Natural and Political Observations Made Upon the Bills of Mortality*, 1663.

Christiaan Huygens (1629-1695): *De ratiociniis in ludo aleae*, 1657.

Jacques Bernoulli (1655-1705): *Ars conjectandi*, 1713.

Pierre Rémond Montmort (1678-1719): *Essay d'analyse sur les jeux de hazard*, 1708.

Abraham de Moivre (1667-1754): *The doctrine of chances*, 1718.

Thomas Bayes (1701?-1761): *Essay towards the solution of a problem in the doctrine of chances*, 1763.

Fermat, Bernoulli, de Moivre, …	Séc. XVII-XVIII
Laplace – teoria clássica, etc. (1812)	Séc. XIX
Kolmogorov – teoria axiomática (1933)	Séc. XX

Pierre de Fermat
1601-1665

Jacob Bernoulli
1654-1705

Ars Conjectandi, 1713

Abraham de Moivre
1667-1754

The Doctrine of Chances , 1718

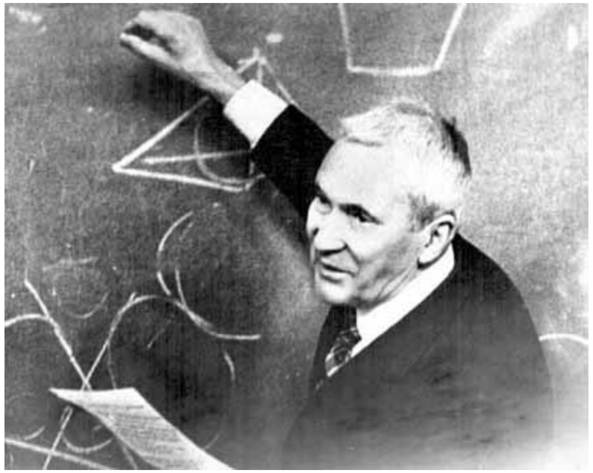
teoria clássica

Théorie Analytique des Probabilités, 1812
Essai Philosophique sur les Probabilités, 1814
Traité de Mécanique Céleste, 1798-1825 (5 volumes)

Pierre Simon de Laplace
("o Newton francês"),1749-1827

distribuição de Bernoulli
Lei dos Grandes Números de Bernoulli
regra de Laplace
distribuição de Laplace
teorema de de Moivre-Laplace
Laplaciano
transformadas de Laplace

teoria axiomática



Andrei N. Kolmogorov, 1903-1987

Grundbegriffe der
Wahrscheinlichkeitsrechnung, 1933
(Fundamentos da Teoria da Probabilidade)

Interpretações de Probabilidade

Frequencista – a probabilidade de um acontecimento A é o limite (quando $n \to \infty$) da frequência relativa de ocorrência de A em n realizações da experiência.

ideia plausível; porém obriga a que a experiência possa ser repetida em iguais condições; e esse limite existe? E caso exista, vai ser independente da sucessão particular de experiências?

Clássica – para uma experiência que tem n resultados possíveis (equiprováveis), a probabilidade de A se realizar é o quociente entre o número de casos favoráveis a A e n.

limita-se ao caso de "resultados possíveis equiprováveis", e de o nº de casos possíveis ser finito; é portanto de aplicação limitada (embora existam algumas extensões…); carece de generalidade

Subjetiva – a probabilidade de um acontecimento A traduz um grau de certeza subjetivo que o observador lhe atribui (que pode ser atualizado em função das observações).

carece de objetividade… é uma espécie de "fezada" (Pestana, 2010); e como atribuir as probabilidades aos acontecimentos com coerência?

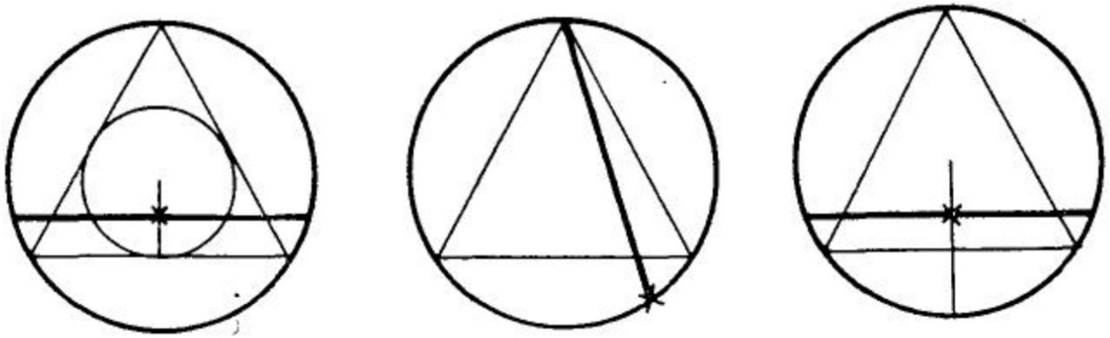
… the history of mathematics has revealed many interesting paradoxes some of which have served as starting-points for great changes. The mathematics of randomness is especially rich in paradoxes. According to Charles Sanders Peirce no branch of mathematics is as easy to slip up in as probability theory.

Gábor J. Székely, 1986
Paradoxes in Probability Theory and Mathematical Statistics

Muitos paradoxos derivam da ambiguidade do significado de "escolher ao acaso"
(como no paradoxo de J. L. Bertrand, 1889, Calcul des Probabilités)

Paradoxo de Bertrand:

Dada uma circunferência (C), escolhe-se ao acaso uma corda. Qual a probabilidade (p) de esta ser maior que o lado do triângulo equilátero (T) inscrito na circunferência?



Escolho um ponto (x) ao acaso no círculo; desenho a corda cujo ponto médio é esse ponto; então T tem lado menor que a corda sse o ponto x está no círculo inscrito em T. Logo a prob é

$$p = \frac{\pi(r/2)^2}{\pi r^2} = \frac{1}{4}$$

Escolho um ponto na C como um dos extremos da corda; escolho outro ponto (x) ao acaso na C para ser o outro extremo; a corda é maior que o lado de T sse o ponto x estiver no terço inferior de C. Logo a prob é

$$p = 1/3$$

Fixo um raio qualquer da C e escolho um ponto (x) ao acaso nesse raio, que será o centro da corda que lhe é perpendicular; desenho o triângulo inscrito T conforme a figura. Logo a prob é

$$p = 1/2$$

Teoria da Probabilidade

J. Bernoulli (1654-1705)

Laplace (1749-1827) – teoria clássica (1812)

Kolmogorov (1903-1987) – teoria axiomática (1933)

A Teoria da Probabilidade estuda modelos para fenómenos incertos ou aleatórios (em que está presente o acaso). Inclui o estudo das "variáveis aleatórias" (quantidades incertas). A probabilidade mede o grau de possibilidade de um acontecimento incerto se realizar.

A probabilidade é uma medida...

Que outras medidas conhecemos? O que têm em comum?

Exemplo de fenómeno aleatório

População dicotómica – cada elemento da população tem ou não tem uma certa característica (e.g., numa população humana – ser fumador ou não, ser portador de um vírus ou não, etc.)

p – proporção de elementos com essa característica, na população

As seguintes experiências (em população dicotómica) têm resultado incerto:

- Extrair um elemento da população, ao acaso
  - é o mesmo que lançar uma moeda-*p* (onde *p* é a probabilidade de sair "cara" em cada lançamento).
- Extrair dois elementos ao acaso (com reposição) da população
  - é o mesmo que lançar duas vezes uma moeda-*p.*

O lançamento de uma moeda (equilibrada ou não) pode portanto ser equiparado ao fenómeno de amostragem (com reposição) numa população dicotómica.

Experiência aleatória

é aquela que tem resultado incerto, i.e.,

- não se sabe qual o resultado que vai ocorrer, antes de ser realizada
- são conhecidos todos os resultados possíveis
- pode ser repetida em condições análogas

Espaço amostral ou espaço de resultados – Ω

é o conjunto de todos os resultados possíveis (i.e., dos resultados elementares – indecomponíveis – da experiência)

Acontecimentos – $A, B, C, \ldots$

são os subconjuntos de Ω "probabilizáveis", i.e., que vão ter uma probabilidade atribuída. O conjunto dos acontecimentos representa-se por $\mathcal{A}$

$$\mathcal{A} = \{A \subseteq \Omega \mid A \text{ é um acontecimento}\}$$

A noção de espaço amostral deve-se a Richard von Mises, 1931 (*cf*. W. Feller, 1968)

Espaço de acontecimentos (ou espaço probabilizável)

é um par $(\Omega\ , \mathcal{A}\ )$ tal que Ω é um espaço amostral e $\mathcal{A}$, formado por subconjuntos (ou partes) de Ω, satisfaz a

- $\emptyset \in \mathcal{A}$
- $B \in \mathcal{A} \Rightarrow \Omega \backslash B \in \mathcal{A}$
- $B_i \in \mathcal{A}\ , i = 1, 2, \ldots \Rightarrow \bigcup_{i \geq 1} B_i \in \mathcal{A}$

($\mathcal{A}$ diz-se então uma σ-álgebra de partes de Ω)

Notações para o complementar de B em Ω :

$\Omega \backslash B$
$\Omega - B$
B^C
$\bar{B}$

Uma σ-álgebra é portanto uma estrutura de subconjuntos de Ω que inclui o conjunto vazio e que é <u>fechada para a complementação e para a união numerável</u> (note-se que as σ-álgebras têm propriedades análogas ao <u>conjunto das partes</u> de Ω).
Como exemplo, dado Ω qualquer e um seu subconjunto próprio não vazio, B, então $\{\emptyset, B, \bar{B}, \Omega\}$ é uma σ-álgebra, chamada "σ-álgebra gerada por B" (é a menor das σ-álgebras que inclui B).

Note-se que uma σ-álgebra é também <u>fechada para a intersecção numerável</u>, pois $\bigcap_{i \geq 1} B_i = \overline{\bigcup_{i \geq 1} \overline{B_i}}$

Só falta "definir" probabilidade. A que propriedades essenciais terá que satisfazer? Às mesmas da versão frequencista…

Exemplo: 2 lançamentos de uma moeda com faces C e E. Ω = {CC, CE, EC, EE}

A = {CE, EC} = "uma só cara"

Versão teórica	Versão empírica
Versão teórica: se a moeda é equilibrada, a probabilidade de A é 0.5, pois os 4 resultados elementares são equiprováveis. $P(A) = 0.5$	Versão empírica: em 200 lançamentos, "CE" saiu 48 vezes e "EC" saiu 49 vezes. A probabilidade de A é ***estimada*** pela frequência relativa 97/200 = 0.485 e escrevemos $\hat{P}(A) = 0.485$

A frequência relativa $f_n(A)$ de A em n repetições da experiência é uma estimativa de $P(A)$, que satisfaz às seguintes propriedades:

$$\boxed{\underbrace{f_n(A) \geq 0}_{\text{não negativa}} \;;\quad f_n(\Omega) = 1 \;;\quad \underbrace{\text{se } A \cap B = \emptyset, \text{ então } f_n(A \cup B) = f_n(A) + f_n(B)}_{\text{aditiva}}}$$

Estas duas propriedades essenciais caracterizam as medidas e motivaram a teoria da medida

Medida de probabilidade

é uma função P que a cada $A \in \mathcal{A}$, num espaço de acontecimentos $(\Omega, \mathcal{A})$, associa um valor real $P(A)$ – a probabilidade de A – tal que

As "medidas" são as funções que satisfazem aos axiomas I e III, tais como: comprimentos, áreas, volumes, pesos ... e probabilidades.

I. $\forall A \in \mathcal{A}, P(A) \geq 0$ (não negatividade)

II. $P(\Omega) = 1$

III. Se $A_i \in \mathcal{A}$, $i = 1, 2, \ldots$, $A_i \cap A_j \underset{i \neq j}{=} \varnothing$, então $P\left(\bigcup_i A_i\right) = \sum_i P(A_i)$ (σ-aditividade)

Ao terno $(\Omega, \mathcal{A}, P)$ chama-se espaço de probabilidades.

Caso Ω seja infinito não numerável, não é possível atribuir probabilidade a todos os subconjuntos de Ω sem violar III (o conjunto de todos os subconjuntos de Ω é demasiado vasto e tem que ser reduzido). No caso $\Omega = \mathbb{R}$, pode usar-se a "σ-álgebra dos borelianos" (gerada pelos intervalos de $\mathbb{R}$). No caso Ω finito, o axioma III ("aditividade numerável") pode ser substituído pelo de "aditividade finita".

Caso particular de medida de probabilidade:

Ω finito (n elementos), $\mathcal{A} = \mathcal{P}(\Omega)$

teoria clássica

Laplace

$$P(A) = \frac{\# A}{\# \Omega}$$

Para um acontecimento elementar, $A=\{a\}$, temos $P(A) = 1/n$ (*resultados equiprováveis*)

Exemplo: lançamento de um dado equilibrado (faces 1 a 6); temos $\Omega = \{1,2,3,4,5,6\}$ e $P(\{i\}) = 1/6,\ i = 1, 2, \ldots, 6$

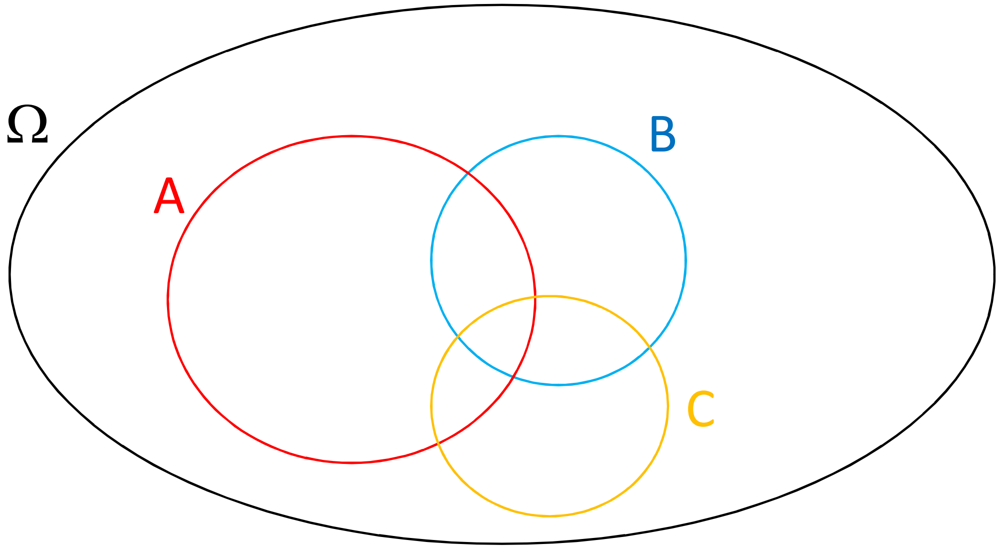
Dado um conjunto Ω qualquer, ao conjunto de todos os subconjuntos de Ω chama-se "conjunto das partes de Ω" e representa-se por $\mathcal{P}(\Omega)$.

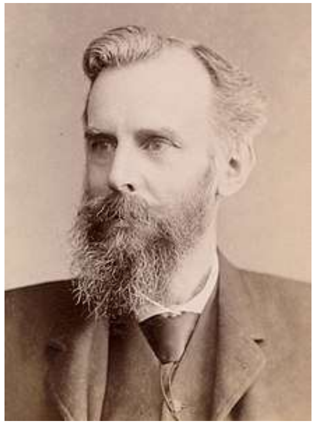
Exemplo: se $\Omega = \{0,1\}$, então $\mathcal{P}(\Omega) = \{ \emptyset, \{0\}, \{1\}, \{0,1\} \}$

Recorde que se $\#\Omega = n$ então $\#\mathcal{P}(\Omega) = 2^n$.

Exercício: (nº 9) Verifique que a "teoria clássica" satisfaz aos 3 axiomas.

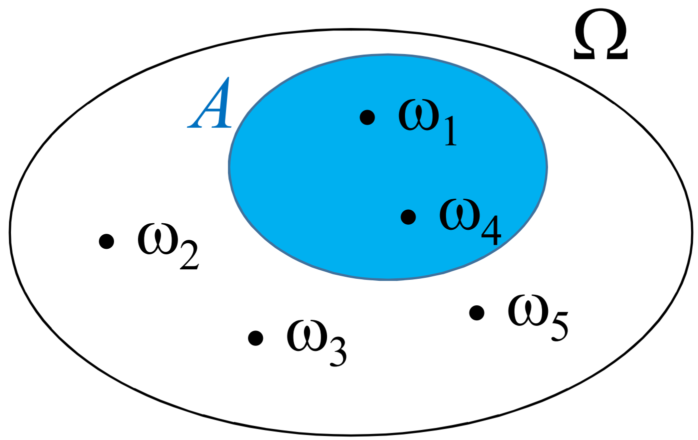
Como a probabilidade é uma medida (tal como a área), é natural usar os "diagramas de Venn" como representação gráfica para os acontecimentos, equiparando a probabilidade a uma área, com a restrição $P(\Omega) = 1$ (ou seja, Ω com área 1).





John Venn
1834-1923

Outros exemplos de espaços de probabilidade:



- $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$, $\mathcal{A} = \mathcal{P}(\Omega)$,

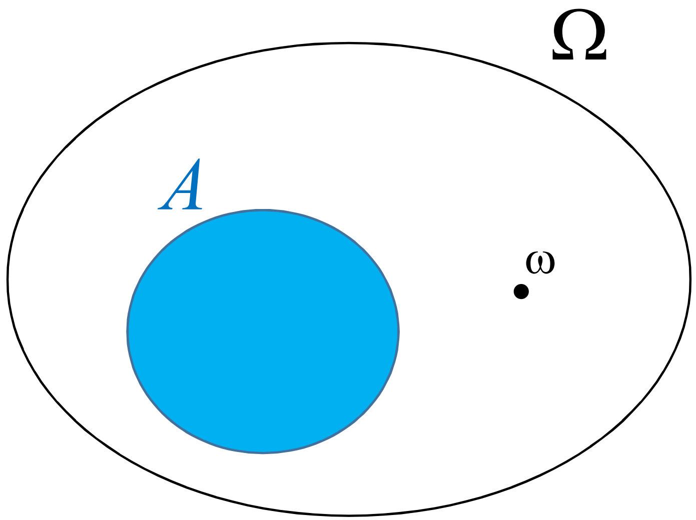
  A cada ω_i associa-se $p_i \geq 0$, com $\sum_{i=1}^{n} p_i = 1$

  Para $A \in \mathcal{A}$, seja $P(A) = \sum_{i:\omega_i \in A} p_i$

- Ω qualquer, $A \subset \Omega$ (com $A \neq \emptyset$, $A \neq \Omega$) , $\mathcal{A} = \{\emptyset, A, \bar{A}, \Omega\}$, $0 \leq p \leq 1$

  Para $B \in \mathcal{A}$, seja $P(B) = \begin{cases} 0 & B = \emptyset \\ p & B = A \\ 1-p & B = \bar{A} \\ 1 & B = \Omega \end{cases}$

Outros exemplos de espaços de probabilidade:



- $\Omega \subset \mathbb{R}^2$ (com área não nula), $\mathcal{A}$ a σ-álgebra dos borelianos (os subconjuntos de Ω que têm área definida, também chamados mensuráveis)

  Para $A \in \mathcal{A}$, seja $$P(A) = \frac{m(A)}{m(\Omega)}$$

  onde $m(A)$ representa a medida (área) de A.

Esta definição generaliza a definição clássica de Laplace a um universo infinito não numerável $\Omega \subset \mathbb{R}^2$. Aqui a probabilidade de A é proporcional à sua área (na teoria clássica é proporcional ao número de elementos de A). Antes da axiomatização, esta formulação gerou alguns paradoxos, tal como este: a probabilidade de um ponto isolado, $B = \{\omega\}$, é zero (pois a área de B é zero), mas isso não implica que esse acontecimento seja impossível de acontecer (veja-se o caso de um tiro ao alvo). Assim, há que distinguir entre um acontecimento impossível e um com probabilidade zero. Ou seja, há acontecimentos com probabilidade zero que não são impossíveis. Mas Ω, que tem probabilidade 1, é uma união de acontecimentos elementares de probabilidade 0…

Fórmulas elementares em espaços de probabilidade
(decorrentes dos axiomas I, II e III):

IV. $P(\overline{A}) = 1 - P(A)$

V. $P(\emptyset) = 0$

VI. $P(A) \leq 1$

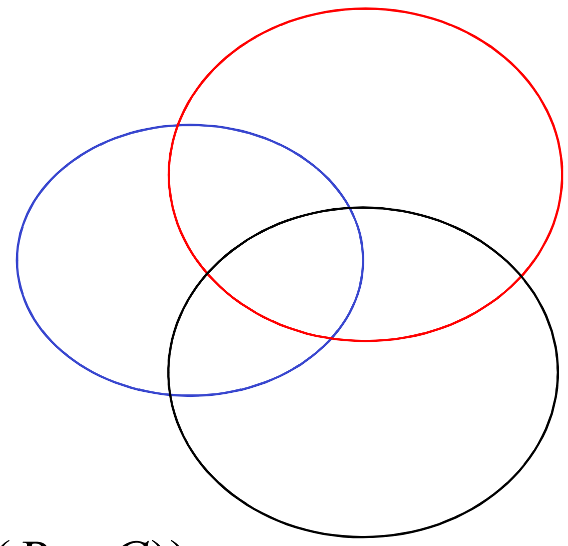
VII. $P(B - A) = P(B) - P(A \cap B)$

VIII. $A \subset B \Rightarrow P(A) \leq P(B)$ monotonia

IX. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ regra da adição

Exercício: (nº 12) Generalize a regra da adição para 3 ou mais acontecimentos ("regra de inclusão-exclusão") e (nº13) aplique ao problema dos chapéus ou presentes (qual a probabilidade de pelo menos uma pessoa receber o presente que trouxe?)

Resolução: (i) Generalização para 3 acontecimentos



Regra da adição:

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

$$P(A \cup B \cup C) = \ ???$$

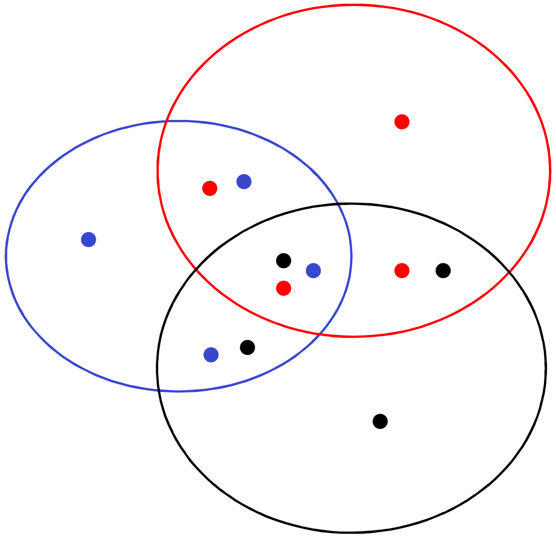
Aplicando 3 vezes a regra da adição:

$$P(A \cup B \cup C) = P(A \cup (B \cup C)) \underset{*}{=} P(A) + \underbrace{P(B \cup C)} - \underbrace{P(A \cap (B \cup C))}_{P((A \cap B) \cup (A \cap C))} =$$

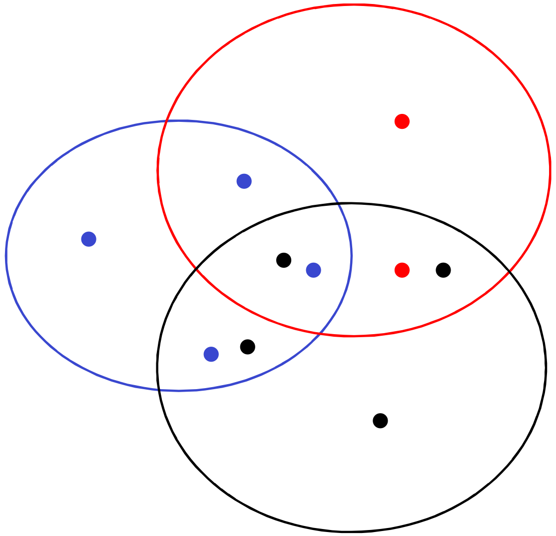
$$= P(A) + \overbrace{P(B) + P(C) - P(B \cap C)}^{*} \overbrace{- P(A \cap B) - P(A \cap C) + P(\underbrace{A \cap B \cap A \cap C}_{A \cap B \cap C})}^{*}$$

Resolução (cont.): (i) O mesmo, com diagramas de Venn:

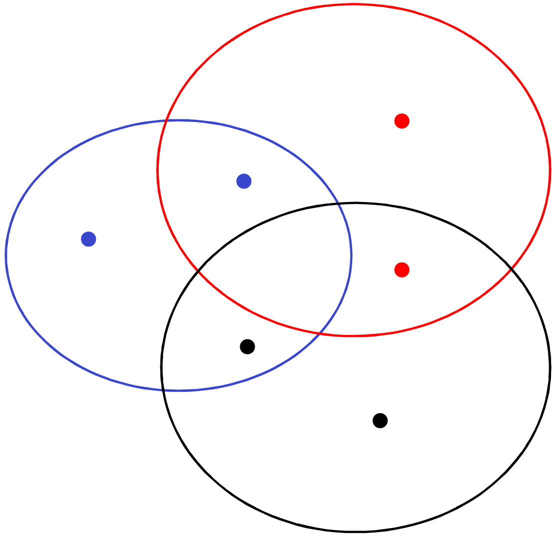
$P(A) + P(B) + P(C)$



$P(A) + P(B) + P(C)$
$- P(A \cap B)$



$P(A) + P(B) + P(C)$
$- P(A \cap B) - P(A \cap C) - P(B \cap C)$



agora falta
$P(A \cap B \cap C)$
para ter
$P(A \cup B \cup C)$

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

Resolução (cont.): (ii) Generalização para n acontecimentos (a demonstrar por indução…)

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

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

$$\vdots$$

Regra de inclusão-exclusão:

$$\boxed{P(A_1 \cup A_2 \cup \ldots \cup A_n) = \underbrace{\sum_{i=1}^{n} P(A_i)}_{n \text{ termos}} - \underbrace{\sum_{i<j} P(A_i \cap A_j)}_{\binom{n}{2} \text{ termos}} + \underbrace{\sum_{i<j<k} P(A_i \cap A_j \cap A_k)}_{\binom{n}{3} \text{ termos}} - \ldots + (-1)^{n+1} P(A_1 \cap A_2 \cap \ldots \cap A_n)}$$

Resolução (cont.): (iii) Aplicação ao problema dos encontros / presentes / chapéus

De uma urna com bolas numeradas de 1 a n fazem-se extrações sucessivas até ficar vazia. Se a bola nº i sair na i-ésima extração diz-se que houve um "encontro". Qual a probabilidade de haver pelo menos um encontro? E qual o limite quando $n \to \infty$?

Seja A_i o acontecimento "encontro na i-ésima extração". Pretende-se calcular $P(A_1 \cup A_2 \cup \ldots \cup A_n)$. Pela regra de Laplace, temos

$$P(A_i) = \frac{(n-1)!}{n!} = \frac{1}{n} \qquad P(A_i \cap A_j) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)} \quad \text{para } i \neq j$$

$$P(A_i \cap A_j \cap A_k) = \frac{(n-3)!}{n!} = \frac{1}{n(n-1)(n-2)} \quad \text{para } i, j, k \text{ distintos,} \quad \text{etc.}$$

Aplicando a regra de inclusão-exclusão e a série exponencial, temos

$$\boxed{e^x = \sum_{i \geq 0} \frac{x^i}{i!} \ , \forall_{x \in \mathbb{R}}}$$

Resolução (cont.): (iii)

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) - \ldots + (-1)^{n+1} P(A_1 \cap A_2 \cap \ldots \cap A_n)$$

$$= n\frac{1}{n} - \binom{n}{2}\frac{1}{n(n-1)} + \binom{n}{3}\frac{1}{n(n-1)(n-2)} - \ldots + (-1)^{n+1}\binom{n}{n}\frac{1}{n!}$$

$$= 1 - \frac{1}{2!} + \frac{1}{3!} - \ldots + (-1)^{n+1}\frac{1}{n!} \xrightarrow[n\to\infty]{} \underbrace{1 - e^{-1}} = 0.6321206$$

A convergência é rápida:

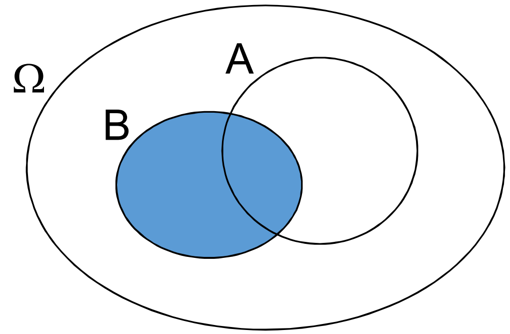
$$\boxed{1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \ldots = -\sum_{i\geq 1}\frac{(-1)^i}{i!} = 1 - \sum_{i\geq 0}\frac{(-1)^i}{i!} = 1 - e^{-1}}$$

```
n <- 12
cumsum( 1/factorial(1:n)*(-1)^(2:(n+1)) )
[1] 1.0000000 0.5000000 0.6666667 0.6250000 0.6333333 0.6319444
[7] 0.6321429 0.6321181 0.6321208 0.6321205 0.6321206 0.6321206
options(digits=10)
1-exp(-1)
[1] 0.6321205588
```

Probabilidade condicional

Se, numa dada experiência aleatória, B ocorreu, define-se "probabilidade de A condicional a B" (ou "probabilidade de A, dado B"), por

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



Esta definição pressupõe $P(B) > 0$. Note que $P(B|B) = 1$.

Na prática, deparamo-nos diretamente com probabilidades condicionais fáceis de calcular; por exemplo, se no totoloto acabaram de sair os nºs 7 e 17 nas duas primeiras extrações, então na terceira extração a probabilidade condicional de sair o 27 é 1/47.

Desta definição resulta (tendo A e B probabilidade > 0) que

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

Independência

Se $P(A|B)$ e $P(A)$ coincidirem, diz-se que A é independente de B ($\Rightarrow$ B é independente de A) e resulta que $P(A\cap B)=P(A)P(B)$. Assim, dois acontecimentos A e B dizem-se independentes se

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

Atenção: Não confundir com acontecimentos disjuntos...
Se A e B são independentes, então $A\cap B \neq \varnothing$ (a menos que algum tenha probabilidade 0)

Exemplo (cartas; 2 extrações s/ reposição): A – "sair ás♠ na 1ª"; B – "sair ás♠ na 2ª"

$$P(B\mid A)=0 \quad , \; P(B)=\frac{51\times 1}{52\times 51}=\frac{1}{52}$$

(e se fosse com reposição? Ver adiante)

Como $P(B\mid A)\neq P(B)$, então A e B não são independentes

Exercício: Extraindo uma carta ao acaso, A –"sair um ás" e B –"sair copas" são independentes?

Independência mútua

Os acontecimentos $A_1, A_2, \ldots, A_n$ dizem-se mutuamente independentes se, para quaisquer r desses acontecimentos ($2 \le r \le n$), "a probabilidade da sua intersecção for igual ao produto das probabilidades de cada um", i.e., se

$$\boxed{\begin{array}{c} P(A_i \cap A_j) = P(A_i)P(A_j) \\ P(A_i \cap A_j \cap A_k) = P(A_i)P(A_j)P(A_k) \\ \vdots \\ P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)P(A_2)\ldots P(A_n) \end{array}}
\qquad \begin{array}{c} 2 \text{ a } 2 \\ 3 \text{ a } 3 \\ \\ n \text{ a } n \end{array}$$

Exercícios: (i) Um tetraedro tem uma face Amarela, uma Branca, uma Cinzenta, e uma com riscas das 3 cores. Analise os acontecimentos "saída de cor A", "saída de cor B", "saída de cor C" quanto à independência.

(ii) Quantas são as equações na definição de independência mútua?

Revisão: amostragem com/sem reposição

Exemplo: Baralho de cartas (52 cartas; 4 naipes: ♠ ♥ ♦ ♣)

Experiência aleatória: 2 extrações ao acaso

	com reposição	sem reposição
Casos possíveis:	52×52	52×51

Acontecimentos:

A – "sair ás♠ na 1ª extração"

B – "sair ás♠ na 2ª extração"

com reposição

$$P(A\cap B)=\frac{1}{52\times52}$$

$$P(A)=\frac{1\times52}{52\times52}=\frac{1}{52}$$

$$P(B)=\frac{52\times1}{52\times52}=\frac{1}{52}$$

A e B são independentes

sem reposição

$$P(A\cap B)=0$$

$$P(A)=\frac{1\times51}{52\times51}=\frac{1}{52}$$

$$P(B)=\frac{51\times1}{52\times51}=\frac{1}{52}$$

A e B são disjuntos, não são independentes

válido para quaisquer outros A e B (referentes à 1ª e 2ª extrações), e para mais extrações… ⟶ Independência mútua

- Em amostragem com reposição,

efetuada ao acaso, quaisquer acontecimentos referentes a extrações distintas são mutuamente independentes (diz-se então que temos **extrações independentes**)

- Em amostragem sem reposição,

efetuada ao acaso, tal já não é verdade (*cf.* totoloto)

Em populações de elevada dimensão, N, e tendo a amostra uma dimensão pequena em comparação com N, é praticamente equivalente efetuar a amostragem *com* reposição ou *sem* reposição. (resultado a demonstrar mais adiante)

O termo **amostragem aleatória** refere-se em rigor a extrações ao acaso e com reposição (logo, **mutuamente independentes**)

Regra da multiplicação

Generaliza a fórmula $P(A \cap B) = P(A)\, P(B|A)$ à intersecção de n acontecimentos:

$$\boxed{P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)\, P(A_2|A_1)\, P(A_3|A_1 \cap A_2) \ldots P(A_n|A_1 \cap A_2 \ldots \cap A_{n-1})}$$

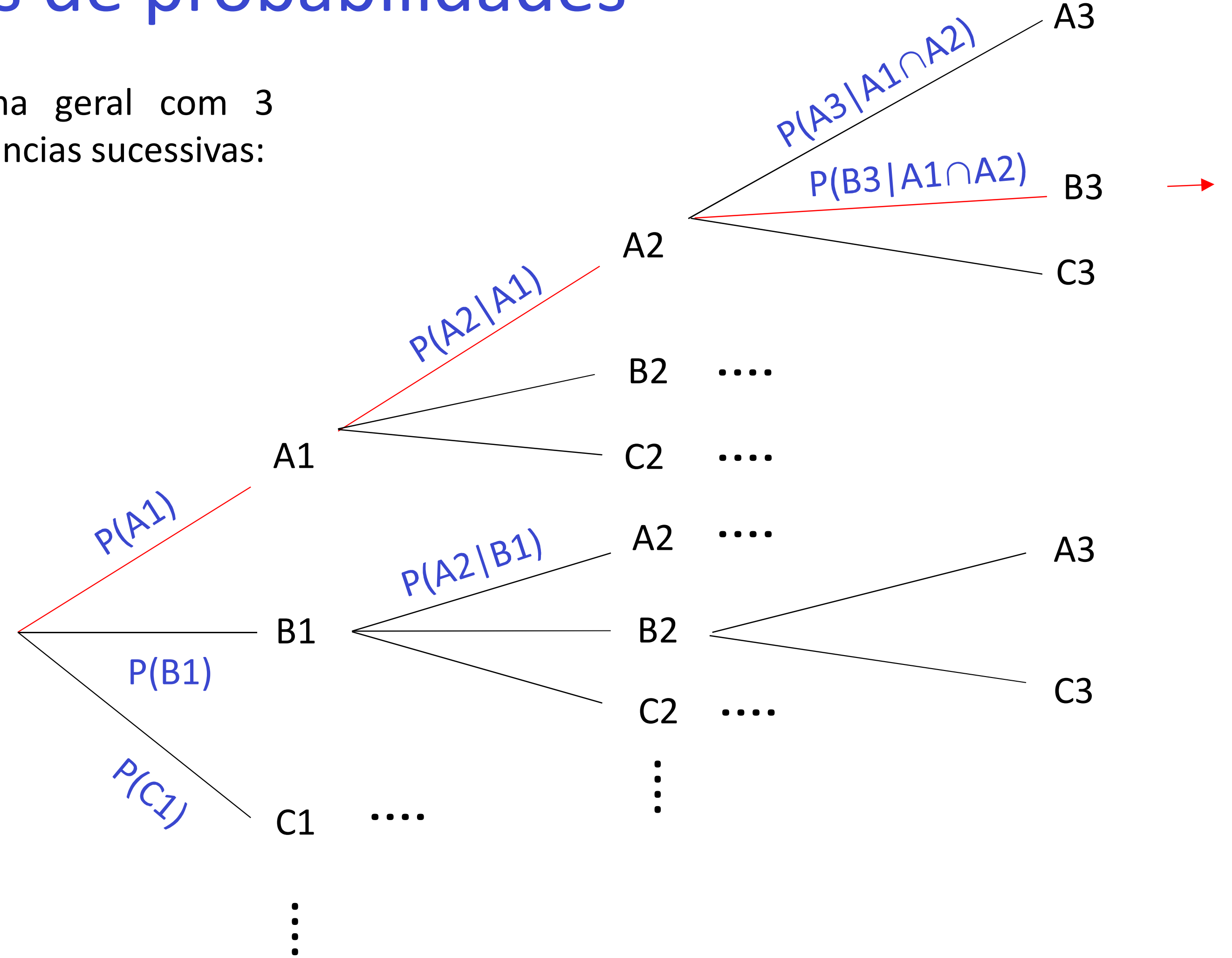
demonstrar…

Nota: no caso de acontecimentos mutuamente independentes esta fórmula reduz-se a $P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)\, P(A_2) \ldots P(A_n)$

As árvores de probabilidades (que recorrem à regra da multiplicação) são muito úteis para esquematizar sucessões de experiências aleatórias.
A 1ª ramificação corresponde à 1ª experiência e leva a acontecimentos A1, B1, C1, … (mutuamente exclusivos e exaustivos – suas probabilidades somam 1) que são os nós da 1ª fase. A partir de cada um destes nós haverá novas ramificações correspondentes à 2ª experiência, levando a acontecimentos A2, B2, C2, …, etc. (ver adiante).
Em cada ramo colocamos as probabilidades condicionais à “intersecção dos acontecimentos que figuram nos nós que conduzem à raiz da árvore”.

Árvores de probabilidades

Esquema geral com 3 experiências sucessivas:



P(A1∩A2∩B3) é o produto das probabilidades que estão no caminho vermelho

Exemplo: Calcule a probabilidade p de não sair a bola nº1 nas 7 extrações sucessivas de um totoloto (com bolas de 1 a 49). Seja A_i – "não sair bola nº1 na extracção i".

(i) pela regra de Laplace: nº casos favoráveis: $48 \times 47 \times \ldots \times (48-7+1)$

nº casos possíveis: $49 \times 48 \times \ldots \times (49-7+1)$

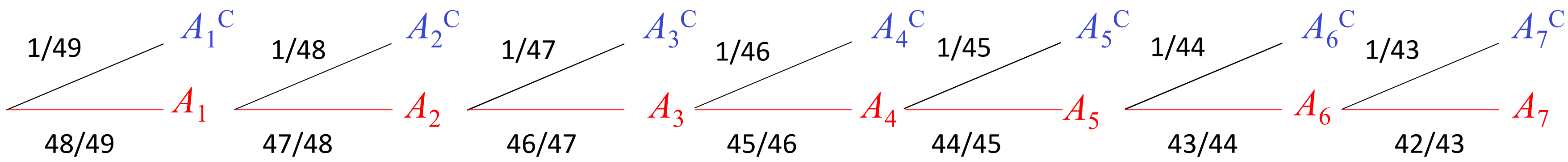
$$\left.\begin{array}{l} 48 \times 47 \times \ldots \times (48-7+1) \\ 49 \times 48 \times \ldots \times (49-7+1) \end{array}\right\} \Rightarrow \quad p = \frac{42}{49}$$

(ii) pela regra da multiplicação:

$$p = P(A_1 \cap A_2 \cap \ldots \cap A_7) = P(A_1)\, P(A_2|A_1)\, P(A_3|A_1 \cap A_2) \ldots P(A_7|A_1 \cap A_2 \ldots \cap A_6) =$$

$$= \frac{48}{49}\,\frac{47}{48}\,\frac{46}{47}\,\frac{45}{46}\,\frac{44}{45}\,\frac{43}{44}\,\frac{42}{43} = \frac{42}{49}$$

(iii) com uma árvore de probabilidades:

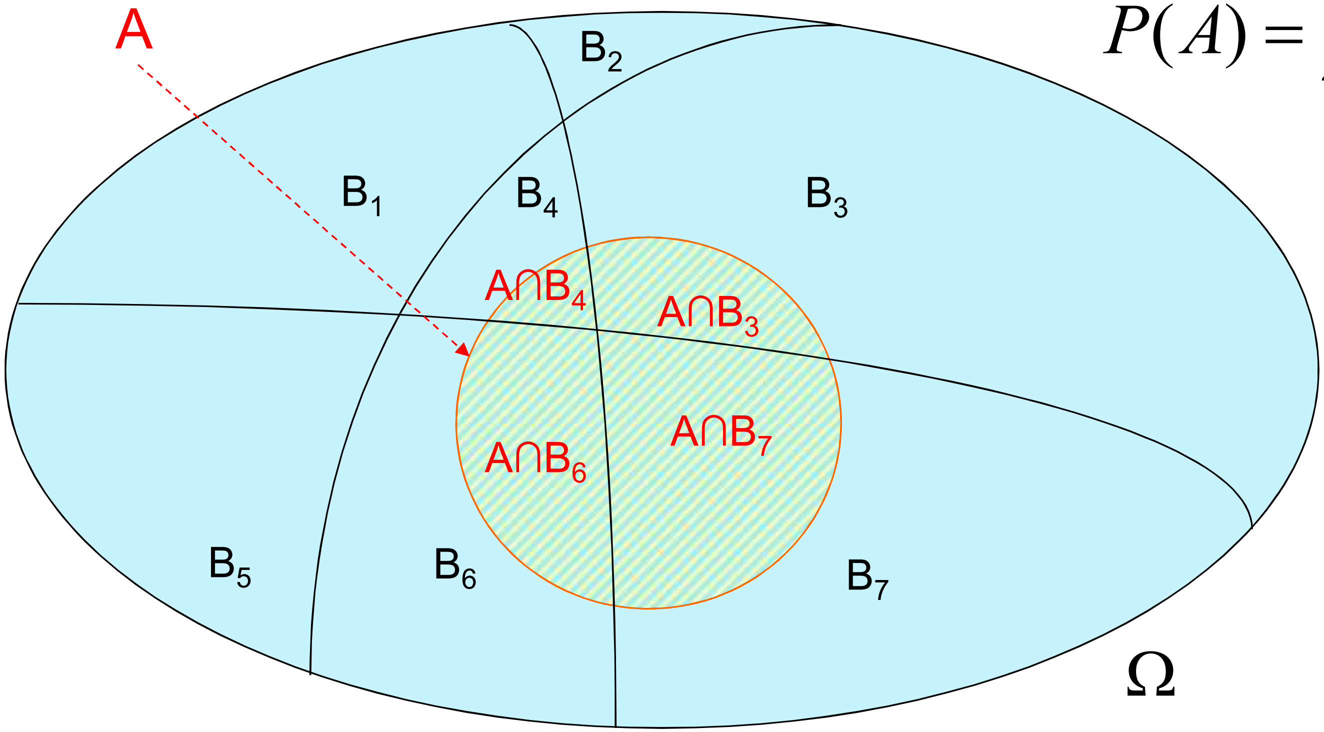


$$p = P(A_1 \cap A_2 \cap \ldots \cap A_7) = \frac{48}{49}\,\frac{47}{48}\,\frac{46}{47}\,\frac{45}{46}\,\frac{44}{45}\,\frac{43}{44}\,\frac{42}{43} = \frac{42}{49}$$ (caminho vermelho)

Teorema da Probabilidade Total

Dados acontecimentos B_j (em número finito ou numerável) formando uma partição* de Ω, então para qualquer $A \in \mathcal{A}$,

$$P(A) = \sum_{j \geq 1} \underbrace{P(A \mid B_j)\, P(B_j)}_{P(A \cap B_j)}$$



*B_1, B_2, ... (em nº finito ou numerável), $B_i \subset \Omega$, formam uma partição de Ω se forem disjuntos dois a dois e a sua união for Ω. Ou seja, se $B_i \cap B_j = \varnothing$ ($\forall\, i \neq j$) e $B_1 \cup B_2 \cup B_3 \cup \ldots = \Omega$

Demonstração:

$$P(A) = P(A \cap \Omega) = P\left(A \cap \left(\bigcup_{j \geq 1} B_j \right) \right) = P\left(\bigcup_{j \geq 1} (A \cap B_j) \right) = \sum_{j \geq 1} P(A \cap B_j) = \sum_{j \geq 1} P(A \mid B_j)\, P(B_j)$$

Pelo axioma III , pois os
acontecimentos $A \cap B_j$
$(j \geq 1)$ são disjuntos 2 a 2

Exemplo 1: Numa população, 0.5% dos indivíduos têm determinada doença (diz-se que a prevalência da doença é de 5/1000 ou 0.5%).

Um teste de sangue tem sensibilidade 95% (i.e., se o indivíduo tem a doença, o teste dá resultado positivo com probabilidade 0.95) e tem especificidade 99% (i.e., se o indivíduo não tem a doença, o teste dá resultado negativo com probabilidade 0.99).

Um indivíduo (escolhido ao acaso) submete-se ao teste. Qual a probabilidade de que o resultado do teste seja positivo?

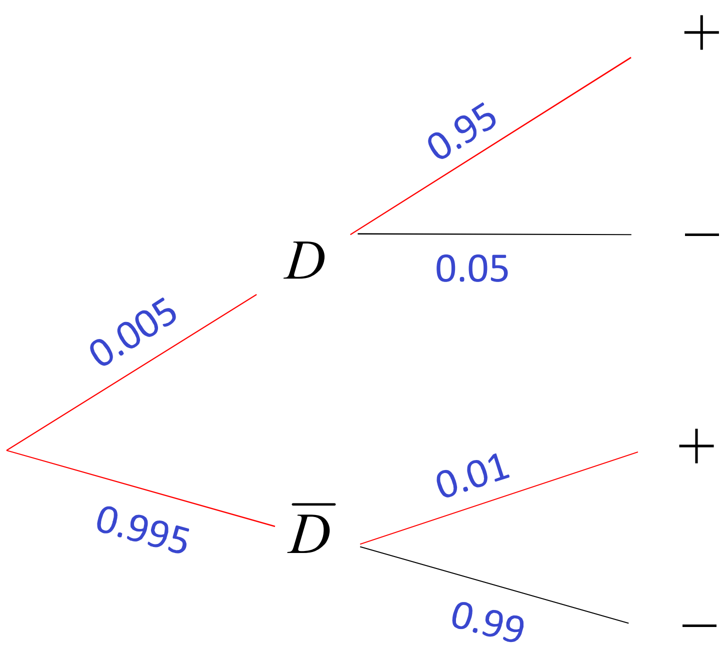
O indivíduo ou está doente (acontecimento D) ou não está; $P(D) = 0.005$; o teste ou dá positivo ($+$) ou negativo ($-$), e

$$P(+\,|\,D) = 0.95, \quad P(+\,|\,\overline{D}) = 0.01$$

Então, aplicando o TPT,

$$\begin{aligned} P(+) &= P(+\,|\,D)P(D) + P(+\,|\,\overline{D})P(\overline{D}) \\ &= 0.95 \times 0.005 + 0.01 \times 0.995 = 0.0147 \end{aligned}$$

Resolução: (com uma árvore de probabilidades)



$$P(D \cap +) = 0.005 \times 0.95 = 0.00475$$

$$P(D \cap -) = 0.005 \times 0.05 = 0.00025$$

$$P(\overline{D} \cap +) = 0.995 \times 0.01 = 0.00995$$

$$P(\overline{D} \cap -) = 0.995 \times 0.99 = 0.98505$$

$$1$$

$$P(+) = 0.00475 + 0.00995 = 0.0147$$

Teorema de Bayes ou Teorema da Probabilidade Inversa

Nas mesmas condições do TPT, tem-se

$$\underbrace{P(B_i \mid A)}_{\text{probabilidade de } B_i \text{ \textit{a posteriori}}} = \frac{P(A \cap B_i)}{P(A)} = \frac{P(A \mid B_i)\overbrace{P(B_i)}^{\text{probabilidade de } B_i \text{ \textit{a priori}}}}{\sum_{j \geq 1} P(A \mid B_j)\, P(B_j)}$$

"Como na axiomática da probabilidade não existe o tempo, o condicionamento pode ser feito num ou noutro sentido. É assim possível procurar a "probabilidade inversa", isto é, inverter o sentido do condicionamento, o que em aplicações é por vezes interpretado como atribuir probabilidade às diversas causas possíveis." – Pestana & Velosa, 2010

Exemplo 1 (cont.): E se o resultado do teste foi positivo, qual a probabilidade de que o indivíduo tenha a doença?

Aplicando o teorema de Bayes, temos

$$P(D \mid +) = \frac{P(+ \mid D)P(D)}{P(+)} = \frac{0.95 \times 0.005}{0.95 \times 0.005 + 0.01 \times 0.995} = 0.3231$$

Exercício: Calcule a probabilidade de que o indivíduo tenha a doença, dado que obteve dois resultados positivos realizados em dois testes independentes.

O teorema de Bayes permite reavaliar uma probabilidade (a priori) depois de ter conhecimento de nova informação.

Exemplo 2: Numa investigação criminal, avaliou-se em 60% a probabilidade de um suspeito ser culpado do crime em causa. Surge entretanto uma nova prova que revela que o culpado tem "cabelo louro" (sabe-se ainda que na população, 20% dos indivíduos são louros). Face a esta nova informação, e dado que o suspeito é louro, qual a probabilidade "a posteriori" de que o suspeito seja o culpado?

C – "o suspeito é culpado"
L – "o suspeito tem a característica (louro) do criminoso"

$$P(C \mid L) = \frac{P(L \mid C)P(C)}{P(L \mid C)P(C) + P(L \mid \overline{C})P(\overline{C})} = \frac{1 \times 0.6}{1 \times 0.6 + 0.2 \times 0.4} = 0.8824$$

Exercícios:

1. Num espaço de probabilidades $(\Omega,\mathcal{A},P)$, dado $B\in\mathcal{A}$ tal que $P(B) > 0$, mostre que a probabilidade condicional $P_B(A) = P(A|B)$, definida para todo $A \in\mathcal{A}$, é uma medida de probabilidade.
2. Numa população com 60% de homens, sabe-se que 10% dos homens são míopes e 15% das mulheres são míopes. Escolhendo ao acaso uma pessoa nesta população, qual a probabilidade de ser míope? E sabendo que tal pessoa é míope, qual a probabilidade de se tratar de uma mulher? Resposta: 0.12; 0.5
3. Em três lançamentos de um dado, calcule (recorrendo ao TPT) a probabilidade de que a soma dos resultados dos dois primeiros lançamentos seja igual ao resultado do terceiro.
4. Um tetraedro tem uma face Amarela, uma Branca, uma Cinzenta, e uma com riscas das 3 cores. Analise os acontecimentos "saída de cor A", "saída de cor B", "saída de cor C" (num lançamento) quanto à independência.

Resposta: são independentes 2 a 2; não são mutuamente independentes

Exercício 1 (nº16) – Resolução

I. $\forall A \in \mathcal{A}, \quad P_B(A) = P(A \mid B) = \dfrac{P(A \cap B)}{P(B)} \geq 0$

II. $P_B(\Omega) = P(\Omega \mid B) = \dfrac{P(\Omega \cap B)}{P(B)} = \dfrac{P(B)}{P(B)} = 1$

III. Se $A_i \in \mathcal{A}$, $i = 1, 2, \ldots$, $A_i \cap A_j = \emptyset$, $i \neq j$, então

$$P_B\left(\bigcup_i A_i\right) = P\left((\bigcup_i A_i) \mid B\right) = \frac{P\left((\bigcup_i A_i) \cap B\right)}{P(B)} = \frac{P\left(\bigcup_i (A_i \cap B)\right)}{P(B)} = \frac{\sum_i P(A_i \cap B)}{P(B)} = \sum_i P(A_i \mid B) = \sum_i P_B(A_i)$$

Logo $(\Omega, \mathcal{A}, P_B)$ é um espaço de probabilidades. Donde P_B satisfaz às fórmulas decorrentes da axiomática (vd. slide 20). Em particular (fórmula IV),

$$P_B(\overline{A}) = 1 - P_B(A) \qquad \text{ou seja} \qquad \boxed{P(\overline{A} \mid B) = 1 - P(A \mid B)}$$

Exercício 3 (nº17) – Resolução

3 lançamentos de um dado; seja R_i o resultado no i-ésimo lançamento; calcular $P(R_1 + R_2 = R_3)$ usando o TPT (partição: resultado do 3º lanç.)

$$P(R_1 + R_2 = R_3) = \sum_{j=1}^{6} P(R_1 + R_2 = R_3 \mid R_3 = j)\ P(R_3 = j) =$$

$$= \sum_{j=1}^{6} P(R_1 + R_2 = j \mid R_3 = j)\ \tfrac{1}{6} =$$

lançamentos independentes

$$= \sum_{j=1}^{6} P(R_1 + R_2 = j)\ \tfrac{1}{6} =$$

$$= (0 + \tfrac{1}{36} + \tfrac{2}{36} + \tfrac{3}{36} + \tfrac{4}{36} + \tfrac{5}{36})\ \tfrac{1}{6} =$$

$$= \tfrac{15}{6^3} = 0.069(4) \cong 0.07$$

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	
3	4	5	6	7		
4	5	6	7			
5	6	7				11
6	7				11	12