Probabilidades e Aplicações

LCC

2021/2022

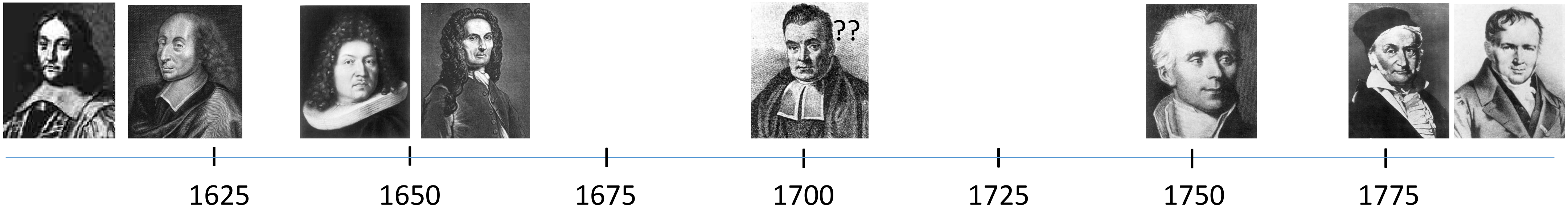
T

Emilia Athayde
mefqa@math.uminho.pt

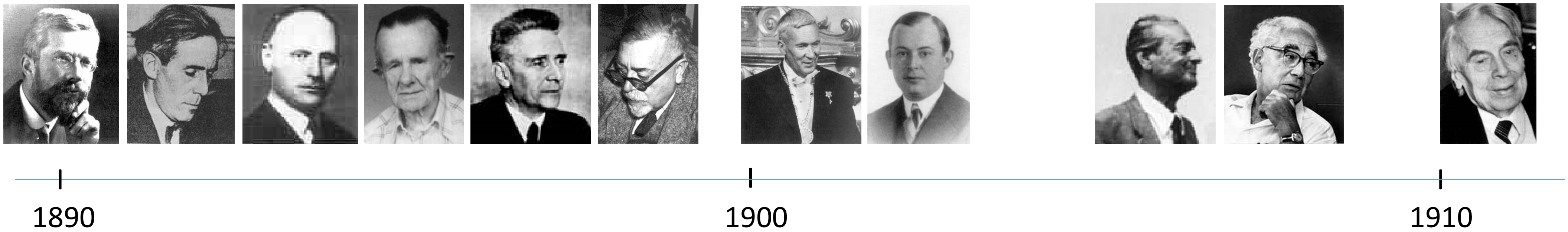
Programa*

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.

- Conforme Plano de Estudos (15 semanas = 75 horas de contacto, incluindo avaliação).
  O calendário escolar 2021/2022 da ECUM abrange apenas 65 horas de contacto (13.3% de redução).


??
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
1655 - 1705

Ars Conjectandi, 1713

Abraham de Moivre
1667-1754

The Doctrine of Chances , 1718

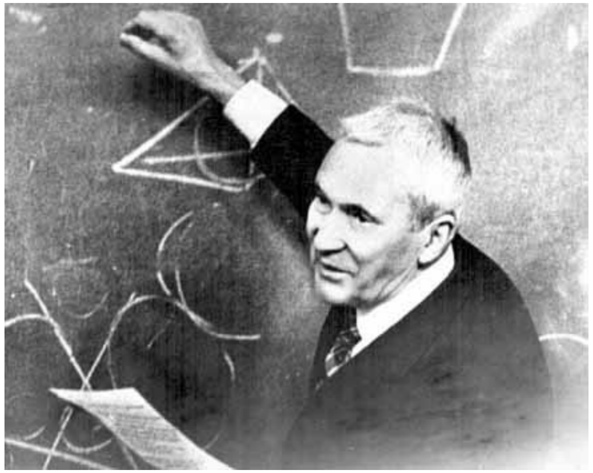
teoria clássica

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

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

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)

The history of the CLT as a universal law begins with Laplace; all relevant studies in the 18th century, starting with [de Moivre 1733], essentially contained only approximations of the binominal distribution and their scope of application remained narrow. Laplace's finding of 1810, according to which the additive coaction of a large number of independent random variables generally leads to probabilities that can, at least approximately, be calculated according to the normal distribution, substantially expanded the numerical possibilities of probability theory, especially in the discussion of mass phenomena.

H. Fisher, *A History of the CLT*, 2011.

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?

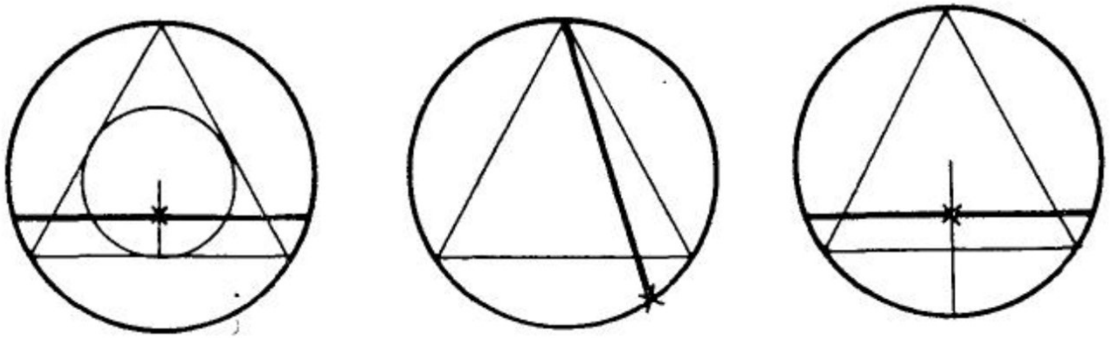
… the history of mathematics has revealed many interesting paradoxes some of which have served as starting-points for great changes. The mathematics of randomness is especially rich in paradoxes. According to Charles Sanders Peirce no branch of mathematics is as easy to slip up in as probability theory.

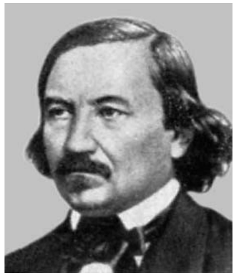
Gábor J. Székely, 1986
Paradoxes in Probability Theory and Mathematical Statistics

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?





J. L. Bertrand
1822 - 1900

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

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

Fixo 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

$$p = \frac{1}{3}$$

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

$$p = \frac{1}{2}$$

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?

Experiências aleatórias – exemplos

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 certo 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 fechada para a complementação e para a união numerável (note-se que as σ-álgebras têm propriedades análogas às do conjunto das partes 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 fechada para a intersecção numerável, pois $\bigcap_{i\geq 1} B_i = \overline{\bigcup_{i\geq 1} \bar{B}_i}$

Motivação

Só falta "definir" probabilidade. A motivação vem da interpretação frequencista (repetindo muitas vezes a experiência, a frequência relativa do acontecimento A é aproximadamente igual à sua probabilidade)

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

Versão teórica: se a moeda é equilibrada, a probabilidade de A é 0.25 (pois os 4 resultados elementares são equiprováveis), e escrevemos

$$P(A) = 0.25$$

Versão empírica: em 200 lançamentos, "CC" saiu 47 vezes. A probabilidade de A é *estimada* pela frequência relativa 47/200 = 0.235 e escrevemos

$$\hat{P}(A) = 0.235$$

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 usuais e inspiraram 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

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

II. $P(\Omega) = 1$

III. Se $A_1, A_2, \ldots \in \mathcal{A}$, e $\underset{i \neq j}{\forall}\ A_i \cap A_j = \varnothing$, então $P\left(\bigcup_i A_i\right) = \sum_i P(A_i)$ (σ-aditividade)

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

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: teoria clássica (de Laplace) ou regra de Laplace

Laplace

Sendo Ω finito (n elementos)
e $\mathcal{A} = \mathcal{P}(\Omega)$, define-se

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

Regra de Laplace

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

Nota: Dado um conjunto Ω qualquer, chama-se "<u>conjunto das partes</u> de Ω" ao conjunto de todos os subconjuntos de Ω e representa-se por $\mathcal{P}(\Omega)$.
Exemplo: se Ω ={0,1}, então $\mathcal{P}(\Omega)$ = { $\emptyset$, {0}, {1}, {0,1} }
Recorde que se $\#\Omega = n$ então $\#\mathcal{P}(\Omega) = 2^n$.

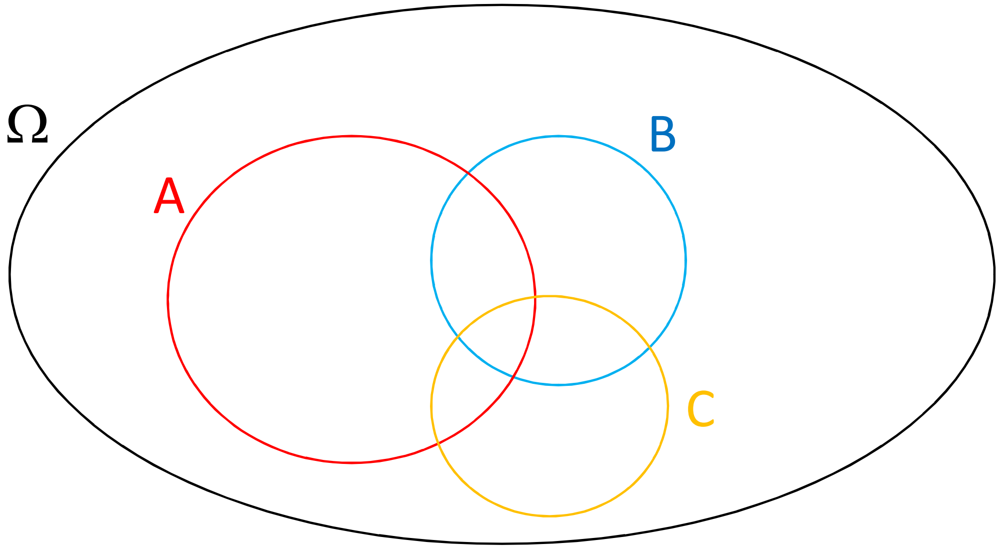
Exemplo: lançamento de um dado equilibrado (faces 1 a 6): $\Omega = \{1, 2, 3, 4, 5, 6\}$

$$P(\{i\}) = 1/6, \; i = 1, 2, ..., 6$$

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

Diagramas de Venn

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).



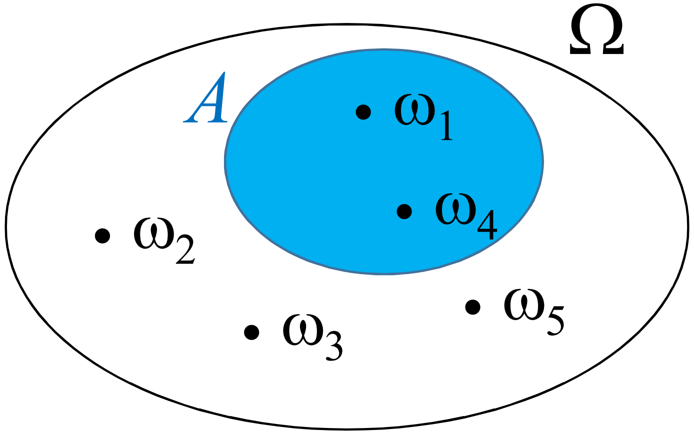
John Venn
1834-1923

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 $P(A) = \sum_{i:\omega_i \in A} p_i$



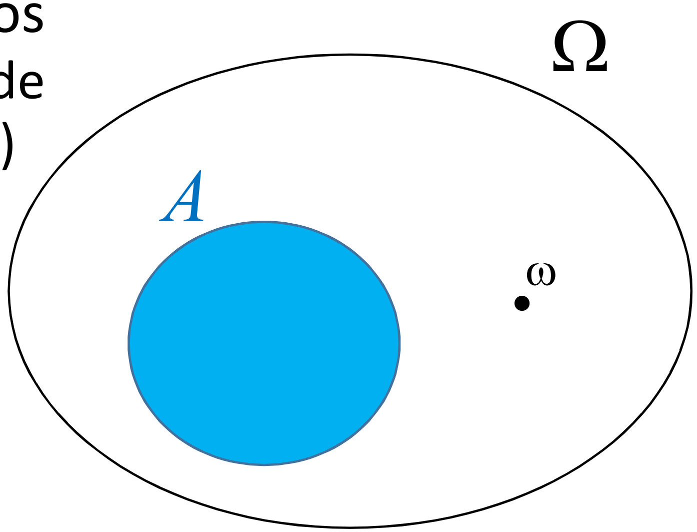
- Ω qualquer, $A \subset \Omega$ (com $A \neq \emptyset$, $A \neq \Omega$) , $\underbrace{\mathcal{A} = \{\emptyset, A, \bar{A}, \Omega\}}_{\sigma\text{-álgebra gerada por } A}$, $0 \leq p \leq 1$

  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), $\mathcal{A}$ é a σ-álgebra dos borelianos (formada pelos subconjuntos *mensuráveis* de Ω, i.e., que têm área definida; é gerada pelos rectângulos)



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

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). 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). Portanto, 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. $P(A) \leq 1$

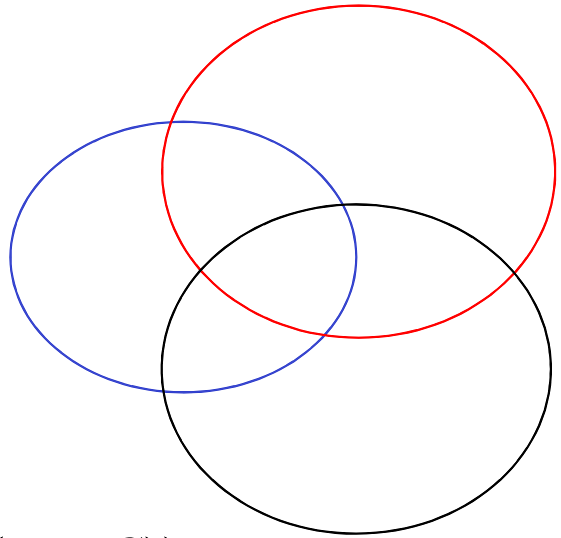
VII. $P(B - A) = P(B) - P(A \cap B)$

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:

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

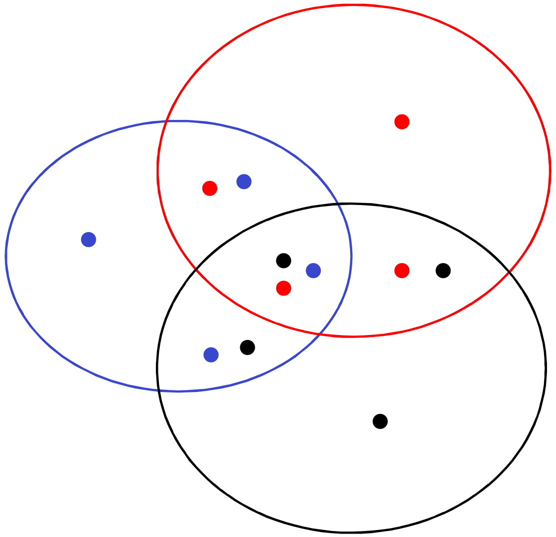
$$=P(A)+\overbrace{P(B)+P(C)-P(B\cap C)}^{*}-\overbrace{P(A\cap B)-P(A\cap C)+P(\underbrace{A\cap B\cap A\cap C}_{A\cap B\cap C})}^{*}$$

Logo

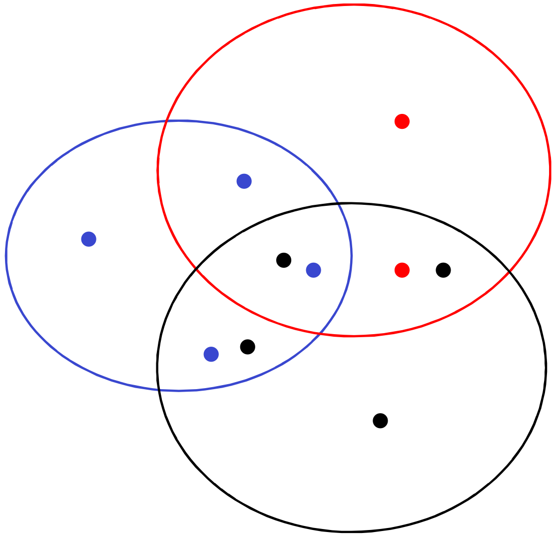
$$\boxed{P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(B\cap C)-P(A\cap B)-P(A\cap C)+P(A\cap B\cap C)}$$

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

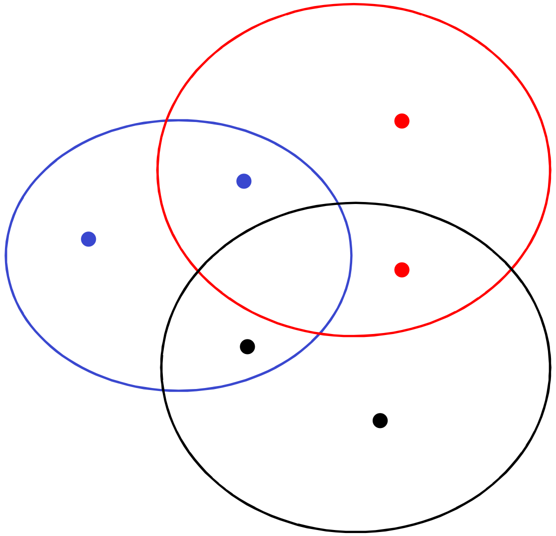
$P(A) + P(B) + P(C)$



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



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



agora falta acrescentar $P(A \cap B \cap C)$ 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 provar 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:

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

n termos $\qquad$ $\binom{n}{2}$ termos $\qquad$ $\binom{n}{3}$ termos

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 p_n 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 $p_n = P(A_1 \cup A_2 \cup \ldots \cup A_n)$. Pela regra de Laplace, temos

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

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

Aplicando a regra de inclusão-exclusão e a série exponencial, $e^x = \sum_{i \geq 0} \frac{x^i}{i!} \ , \forall_{x \in \mathbb{R}}$ temos

Resolução (cont.): (iii)

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

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

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

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

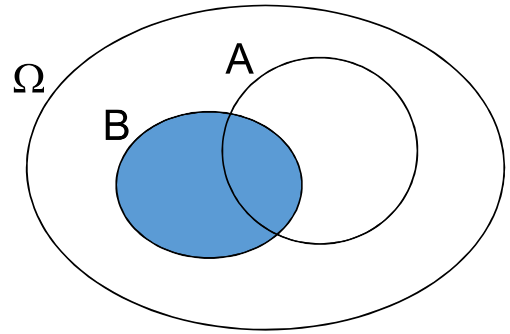
```
# a convergência é rápida:
n <- 12
cumsum( 1/factorial(1:n)*(-1)^(2:(n+1)) )
[1] 1.0000000 0.5000000 0.6666667 0.6250000 0.6333333 0.6319444
[7] 0.6321429 0.6321181 0.6321208 0.6321205 0.6321206 0.6321206
options(digits=10)
1-exp(-1)
[1] 0.6321205588
```

$\underbrace{0.6321206}_{p_{11}}$ $\underbrace{0.6321206}_{p_{12}}$

Probabilidade condicional

Se, numa dada experiência aleatória, o acontecimento B ocorreu, define-se, para qualquer $A \in \mathcal{A}$, a “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 $P(B \mid B) = 1$.

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

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

Atenção: Não confundir com acontecimentos disjuntos...

Se A e B forem acontecimentos independentes, e se ambos tiverem probabilidade positiva, então $P(A \cap B) = P(A)\, P(B) > 0$, donde $A \cap B \neq \varnothing$, donde A e B não são disjuntos (têm que se intersectar).

---

*de facto, supondo $P(A \mid B) = P(A)$, então temos $P(B \mid A) = \dfrac{P(B \cap A)}{P(A)} = \dfrac{P(A \mid B) P(B)}{P(A)} = \dfrac{P(A) P(B)}{P(A)} = P(B)$

ou seja, $P(B \mid A) = P(B)$, i.e., B é independente de A.

Exemplo: Duas extrações, s/ reposição, de um baralho de cartas (52 cartas):

A – "sair ás♠ na 1ª";

B – "sair ás♠ na 2ª"

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

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

(e se fosse com reposição? Ver adiante)

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

Espaço de probabilidades (condicional)

Num espaço de probabilidades (Ω, $\mathcal{A}$, P), dado $B \in \mathcal{A}$ tal que $P(B) > 0$, prova-se que a probabilidade condicional $P_B(A) = P(A \mid B)$, definida para todo $A \in \mathcal{A}$, é uma medida de probabilidade. Ou seja, (Ω, $\mathcal{A}$, P_B) é um espaço de probabilidades (e portanto satisfaz às fórmulas IV a IX).

Demonstração no slide 50

Em particular, $P_B(\overline{A}) = 1 - P_B(A)$

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

etc.

Independência mútua

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

$$\boxed{\begin{gathered} P(A_i \cap A_j) = P(A_i)P(A_j) \\ P(A_i \cap A_j \cap A_k) = P(A_i)P(A_j)P(A_k) \\ \vdots \\ P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)P(A_2)\ldots P(A_n) \end{gathered}}$$

2 a 2

3 a 3

n a n

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.

Resposta: são independentes 2 a 2; não são mutuamente independentes (resolvido na aula T)

(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:

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

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

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

A e B são independentes

sem reposição:

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

$$P(A) = \frac{1 \times 51}{52 \times 51} = \frac{1}{52}$$

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

A e B são disjuntos, <u>não são independentes</u>

(com reposição e sem reposição: $P(A)$ = $P(A)$, $P(B)$ = $P(B)$)

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

Conclusão

- 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 supondo que a amostra tem uma dimensão pequena em comparação com N, prova-se que 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:

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

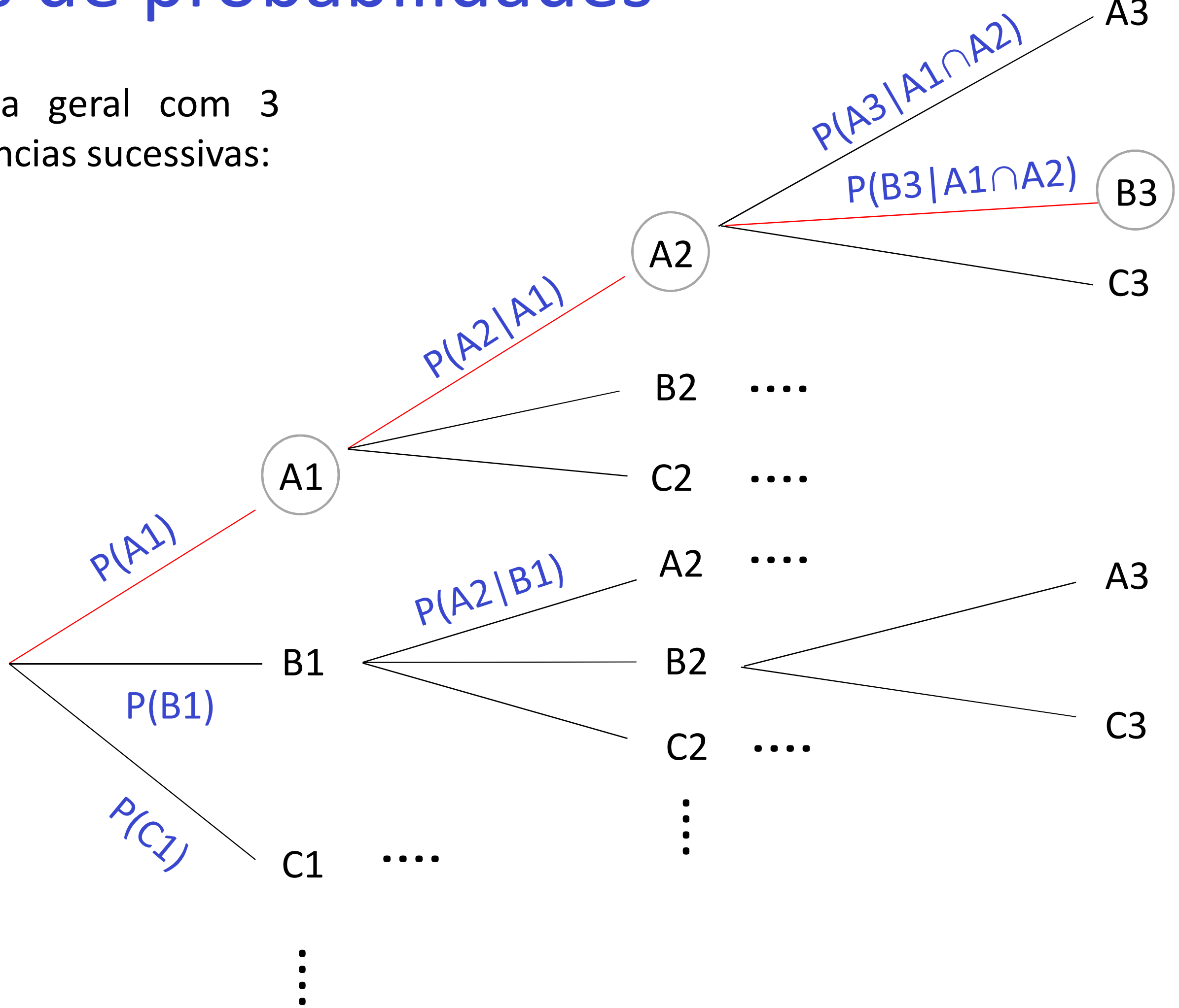
demonstrar…
(muito fácil: desenvolver 2º membro e simplificar)

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 – i.e., disjuntos dois a dois e com probabilidade total 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:



$P(A1 \cap A2 \cap B3)$ é o produto das probabilidades que estão no caminho vermelho (pela regra da multiplicação)

Exemplo: Calcular a probabilidade p de “não sair a bola nº1 nas 7 extrações sucessivas de um totoloto” (bolas 1:49), considerando A_i – “não sair bola nº1 na extracção i”.

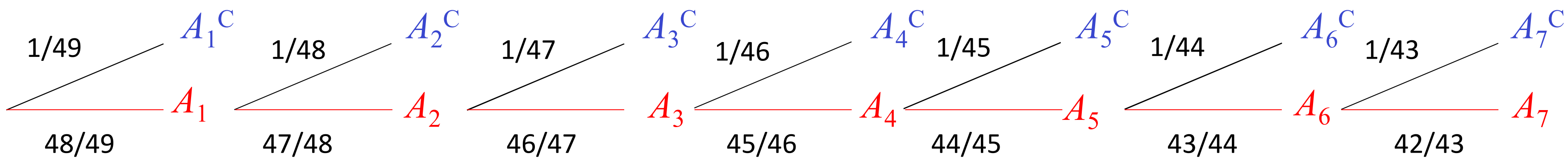
(i) pela regra de Laplace: nº casos favoráveis: $48 \times 47 \times \ldots \times (48\text{-}7+1)$
nº casos possíveis: $49 \times 48 \times \ldots \times (49\text{-}7+1)$ $\Rightarrow \quad p = \dfrac{42}{49}$

(ii) pela regra da multiplicação:

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

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

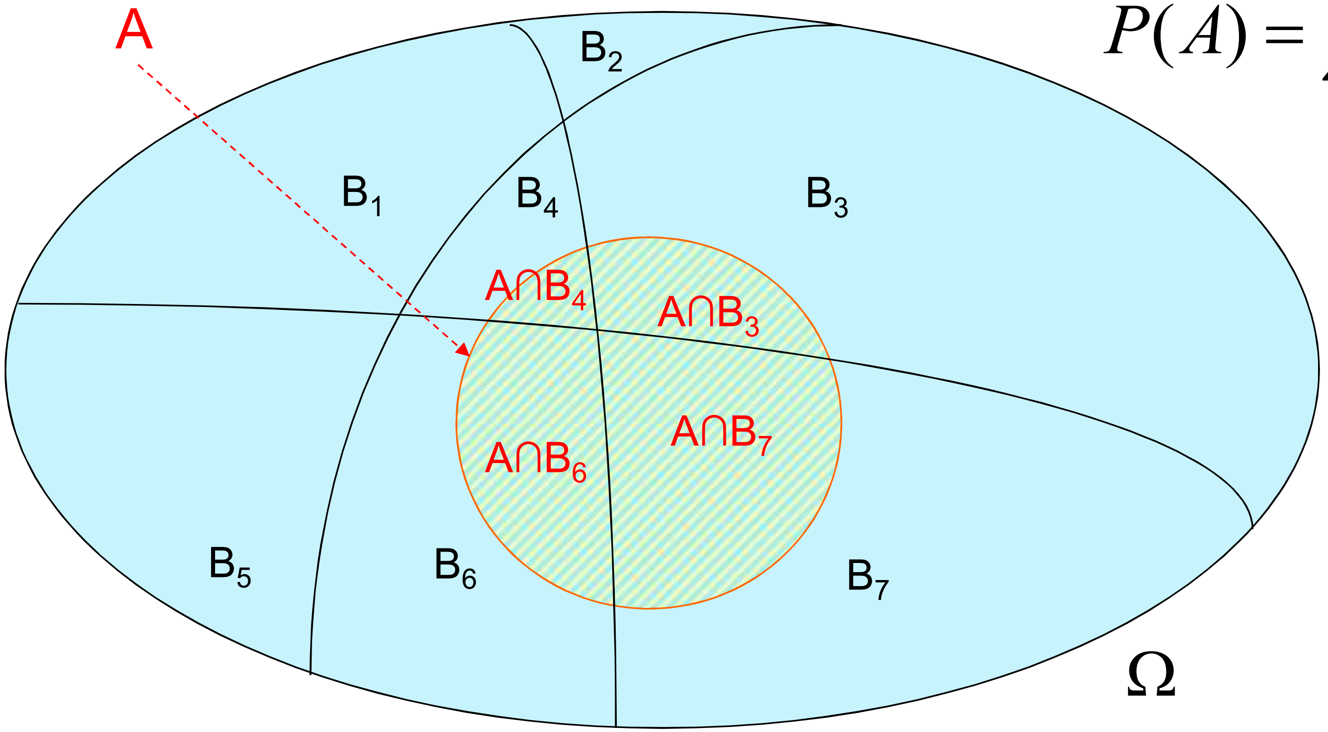
(iii) com uma árvore de probabilidades:



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

Teorema da Probabilidade Total (TPT)

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 $B_i \cap B_j = \emptyset$ ($\forall\, i \neq j$) e $B_1 \cup B_2 \cup B_3 \cup \ldots = \Omega$

Demonstração:

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

Porque os B_j $(j \geq 1)$
são uma partição de Ω

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

Exemplo 1: Numa dada população, 0.5% dos indivíduos têm determinada doença. Um certo teste clínico é tal que se o indivíduo tem a doença, o resultado é positivo com prob 0.95; se o indivíduo não tem a doença, o resultado é negativo com prob 0.99.

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

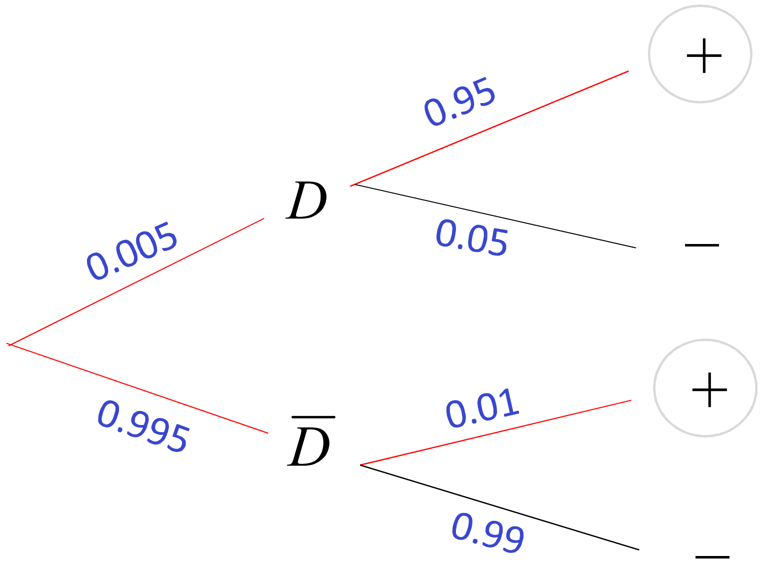
Nota: Diz-se que a prevalência da doença é 0.5%, a sensibilidade do teste é 95% e a especificidade do teste é 99%.

Resolução: O indivíduo (escolhido ao acaso) ou está doente (acontecimento D) ou não está; $P(D) = 0.005$; o teste ou dá positivo (+) ou negativo (–), e temos $P(+\mid D) = 0.95, \quad P(+\mid \overline{D}) = 0.01$. Pretendemos calcular $P(+)$.

Aplicando o TPT (com partição $\{D, \overline{D}\}$ do espaço de resultados; note que há 4 resultados: $D\cap +$, $D\cap -$, $\overline{D}\cap +$, $\overline{D}\cap -$, e que D é união dos dois 1ºs). Então

$$P(+) = P(+\mid D)P(D) + P(+\mid \overline{D})P(\overline{D}) = 0.95\times 0.005 + 0.01\times 0.995 = 0.0147$$

Resolução: (com uma árvore de probabilidades)



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

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

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

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

Total: 1

probabilidade total 1, que se decompõe em 4 parcelas (que são as probabilidades dos 4 resultados da experiência)

Então

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

Teorema de Bayes ou Teorema da Probabilidade Inversa

(Bayes, 1763)
(Laplace, 1781)

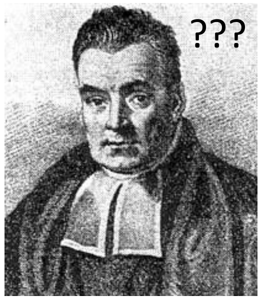
Nas mesmas condições do TPT, tem-se

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

"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.

Thomas Bayes*

An Essay towards the Solution of a Problem in the Doctrine of Chances,
Phil. Trans. Royal Soc., 1763.



Thomas Bayes
1701/2(?) – 1761

O Teorema de Bayes é irrefutável. Mas o postulado de Bayes (*o princípio da razão insuficiente*: não havendo qualquer conhecimento prévio, considerem-se todas as probabilidades *a priori* como iguais) sobre atribuição das probabilidades *a priori* a parâmetros (desconhecidos), utilizado na aplicação do teorema à inferência estatística, é que foi muito controverso. Mais tarde (no séc. XX) Jeffreys introduziria os *Jeffreys' priors* para contornar esta questão.

> "… considero o Bayes como o primeiro homem que pôs claramente o problema fundamental da Estatística: de que maneira, a partir das observações, é possível saber alguma coisa relativamente a um certo universo."
>
> J. Tiago de Oliveira, *A Herança de Bernoulli-Bayes-Laplace: Um Esboço da História da Estatística* (SPM, 1988)

---

* Reverendo presbiteriano, estudou em Edinburgo, FRS em 1742. Deixou manuscritos sobre vários temas de matemática e física. Publicou apenas duas obras em vida: 1. *Divine Benevolence* (1731).
2. *An Introduction to the Doctrine of Fluxions, and a Defence of the Mathematicians Against the Objections of the Author of The Analyst* (1736). Nesta obra, Bayes defende e esclarece a teoria das fluxões (cálculo infinitesimal) de Newton, em resposta aos ataques do bispo George Berkeley a respeito dos fundamentos dessa teoria.

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

Aplicando o teorema de Bayes, temos

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

(dos indivíduos com resultado positivo, só $\cong$ 1/3 estão de facto doentes)

Exercício: Mantendo $p = P(+\,|\,D) = 0.95$ e $q = P(+\,|\,\overline{D}) = 0.01$, refaça o cálculo para
(i) doença extremamente rara, com $P(D) = 10^{-5}$; o resultado é $P(D\,|\,+) = 0.00095$. ☹
(ii) doença comum, p. ex. com $P(D) = 0.4$; neste caso o resultado é $P(D\,|\,+) = 0.98$. ☺

```
d.se.posit <- function(d,p,q) p*d/(p*d+q*(1-d))
d.se.posit(0.005,0.95,0.01)
[1] 0.3231293
d.se.posit(10^-5,0.95,0.01)
[1] 0.0009491078
d.se.posit(0.4,0.95,0.01)
[1] 0.984456
```

O teste (fixando $p = 0.95$ e $q = 0.01$) é bom (útil) ou não?
Depende da prevalência da doença...

Se a % de saudáveis é $\cong$1, haverá imensos falsos positivos; para que o teste fosse útil, precisava ser $q << 0.01$

Exemplo 2: (exº 21) Numa população, certa doença está presente sob forma grave em 3% dos indivíduos, sob forma moderada em 10% e ausente nos restantes 87%. Um exame clínico assinala a presença da doença (resultado positivo) em 90% dos casos graves, em 70% dos casos moderados e em 10% dos saudáveis.

(a) Um indivíduo (escolhido ao acaso na população) submete-se ao exame clínico. Qual a probabilidade de que o resultado seja positivo?

Resolução: O indivíduo ou está gravemente doente (acontecimento G) ou está moderadamente doente (M) ou não está doente (A); $P(G) = 0.03$; $P(M) = 0.10$; $P(A) = 0.87$; $P(D) = 0.13$; o teste ou dá positivo (+) ou negativo (–), e

$$P(+\,|\,G) = 0.90,\ P(+\,|\,M) = 0.70,\ P(+\,|\,A) = 0.10$$

Então, aplicando o TPT (partição $\{G, M, A\}$ do espaço de resultados; cf. Ex. 1),

$$\begin{aligned} P(+) &= P(+\,|\,G)P(G) + P(+\,|\,M)P(M) + P(+\,|\,A)P(A) \\ &= 0.90\times 0.03\ +\ 0.70\times 0.10\ +\ 0.10\times 0.87 = 0.184 \end{aligned}$$

(b) E sabendo que o resultado do exame desse indivíduo foi positivo, qual a probabilidade de que o indivíduo tenha de facto a doença?

Resolução: Passando ao complementar (vd. slide 30) e usando a alínea anterior, temos

$$P(D \mid +) = 1 - P(A \mid +) = 1 - \frac{P(+ \mid A)P(A)}{P(+)} = 1 - \frac{0.10 \times 0.87}{0.184} = 0.5271739$$

Em alternativa, resolve-se (sem passar ao complementar) usando $D = G \cup M$ (união disjunta, donde a sua probabilidade – condicional – é aditiva), fazendo

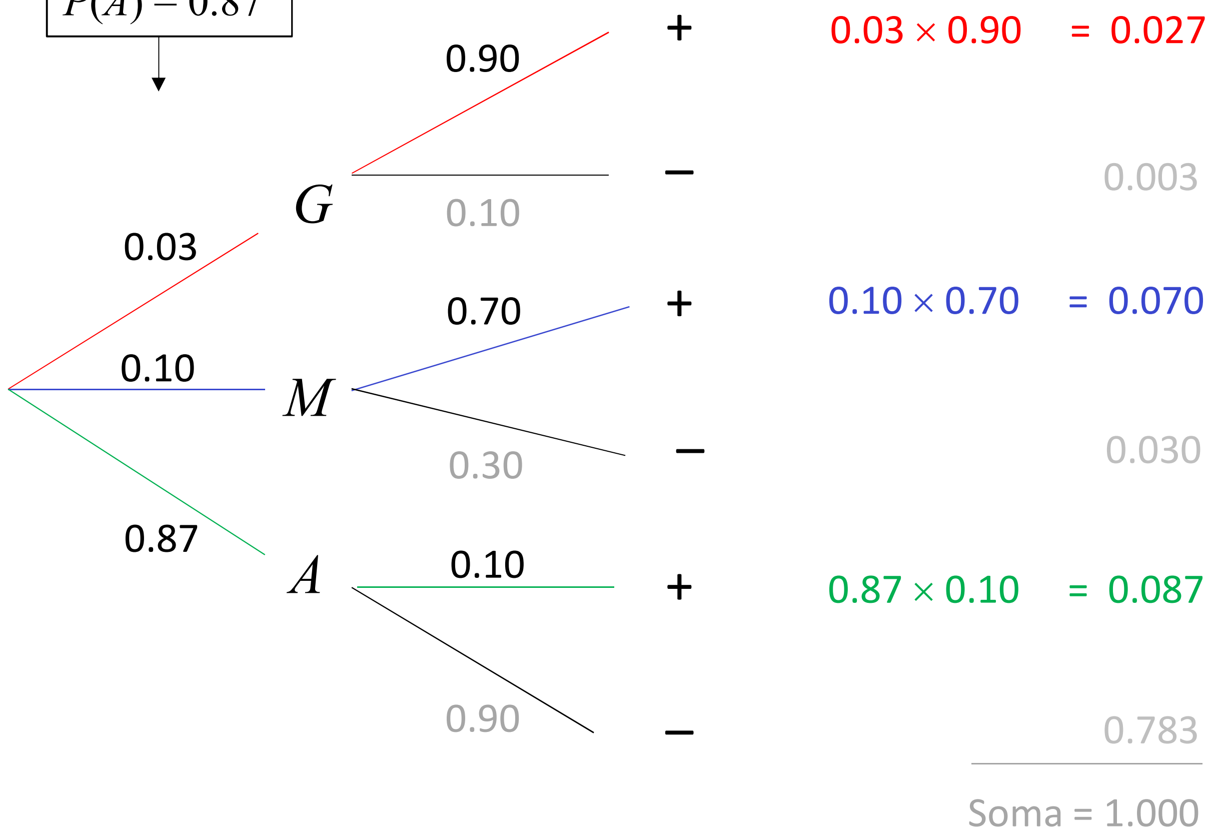
$$P(D \mid +) = P(G \cup M \mid +) = P(G \mid +) + P(M \mid +) = \frac{P(+ \mid G)P(G) + P(+ \mid M)P(M)}{P(+)} =$$

$$= \frac{0.90 \times 0.03 + 0.70 \times 0.10}{0.184} = 0.5271739$$

Exemplo 2 – Resolução usando uma árvore de probabilidades:

$$P(G)=0.03 \quad P(M)=0.10 \quad P(A)=0.87$$

$$P(+\mid G)=0.90,\ P(+\mid M)=0.70,\ P(+\mid A)=0.10$$



(a) Equivalente ao TPT:

$$P(+) = 0.027 + 0.070 + 0.087$$
$$= 0.184$$

(b) Equivalente ao T. Bayes; usando a alínea anterior:

$$P(D\mid +)=\frac{P(D\cap +)}{P(+)}=$$
$$=\frac{P(G\cap +)+P(M\cap +)}{0.184}=$$
$$=\frac{P(+\mid G)P(G)+P(+\mid M)P(M)}{0.184}=$$
$$=0.5271739$$

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

Exemplo 3: Numa investigação criminal (numa grande cidade), 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"; $P(C) = 0.6$ (probabilidade *a priori*)
L – "o suspeito tem a característica (louro) do criminoso"

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

Exercícios resolvidos

14) Prove que se dois acontecimentos são independentes, então
(i) qualquer um deles e o complementar do outro também o são

(ii) os seus complementares também o são.

(i) Sejam A e B independentes, i.e., $P(A\cap B)=P(A)P(B)$. Então

$$P(A\cap\overline{B})=P(A-(A\cap B))=P(A)-P(A\cap B)=P(A)-P(A)P(B)=P(A)(1-P(B))=P(A)P(\overline{B})$$

donde A e $\overline{B}$ são independentes. Analogamente (trocando A com B) temos que $\overline{A}$ e B são independentes.

(ii) Aplicando (i) duas vezes, temos A e B indep $\Rightarrow$ $\overline{A}$ e B indep $\Rightarrow$ $\overline{A}$ e $\overline{B}$ indep

o complementar de um deles (A) e o outro (B) são independentes

o complementar de um deles (B) e o outro (A^C) são independentes

16) 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.

Prova-se que a função P_B definida em $(\Omega, \mathcal{A})$ satisfaz aos axiomas I, II, III:

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

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

III. Se $A_1, A_2, \ldots \in\mathcal{A}$ forem disj. 2 a 2 (i.e., $A_i \cap A_j = \varnothing$, para $i \neq j$) então também o são os acontecimentos $A_i \cap B$, donde

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

Logo $(\Omega, \mathcal{A}, P_B)$ é um espaço de probabilidades.

17) Em três lançamentos de um dado equilibrado, calcule, recorrendo ao TPT, a probabilidade de "a soma dos resultados dos dois primeiros ser igual ao resultado do terceiro".

Seja R_i o resultado no i-ésimo lançamento; calcula-se $P(R_1 + R_2 = R_3)$ condicionando no 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

20) Pretende-se ensinar um rato a virar à direita num labirinto. Coloca-se o rato num compartimento com duas saídas, uma à esquerda e outra à direita. O rato será recompensado se sair pela direita e castigado se sair pela esquerda. Supõe-se que na 1ª tentativa o rato sai ao acaso por qualquer das saídas, e que a seguir a uma recompensa [castigo] sai pela direita com probabilidade 0.6 [0.8]. Seja D_n o acontecimento "o rato sai pela direita na n-ésima tentativa" e $p_n = P(D_n)$. Calcule

(i) p_2 e p_3

(ii) p_{n+1} em função de p_n

(iii) $\lim_{n\to\infty} p_n$

Temos

$$p = P(D_1) = 0.5, \quad P(D_{n+1} \mid D_n) = 0.6, \quad P(D_{n+1} \mid \overline{D_n}) = 0.8$$

Então calcula-se

(i) p_2 e p_3 :

Aplicando o TPT com partição $\{D_1, \overline{D_1}\}$, temos

$$P(D_2) = P(D_2 \mid A_1)P(D_1) + P(D_2 \mid \overline{D_1})P(\overline{D_1}) = 0.6 \times 0.5 + 0.8 \times 0.5 = 0.7$$

Analogamente $P(D_3) = \ldots = 0.66$

(ii) p_{n+1} em função de p_n :

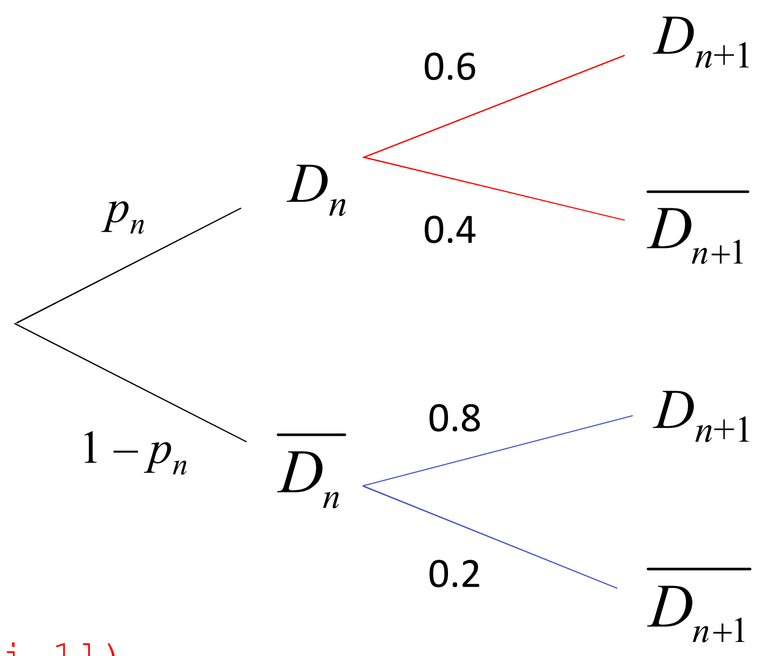
Aplica-se o TPT com a partição $\{D_n, \overline{D_n}\}$

$$p_{n+1} = P(D_{n+1}) = P(D_{n+1} \mid D_n)P(D_n) + P(D_{n+1} \mid \overline{D_n})P(\overline{D_n}) = 0.6 \times p_n + 0.8 \times (1 - p_n) = 0.8 - 0.2 \times p_n$$

(iii) $\lim_{n \to \infty} p_n$

$$\lim_{n \to \infty} p_{n+1} = 0.8 - 0.2 \times \lim_{n \to \infty} p_n \quad \Rightarrow \quad p = 0.8 - 0.2 \times p \quad \Rightarrow \quad p = \frac{8}{12} = \frac{2}{3}$$

$$p_1 = 0.5$$
$$p_2 = 0.6 \times 0.5 + 0.8 \times (1 - 0.5) = 0.7$$
$$p_3 = 0.6 \times 0.7 + 0.8 \times (1 - 0.7) = 0.66$$
$$\vdots$$




```
pnext <- function(p) 0.6*p + 0.8*(1-p)
p <- 0.5; for (i in 2:100) p[i] <- pnext(p[i-1])
p[1:10]
 [1] 0.5000000 0.7000000 0.6600000 0.6680000 0.6664000
 [6] 0.6667200 0.6666560 0.6666688 0.6666662 0.6666668
options(digits=12)
p[100]
 [1] 0.666666666667
```

22) "Duelo triangular" – A, B e C jogam por ordem alfabética, ciclicamente. Na sua vez de jogar, cada um escolhe um dos restantes para alvo e dispara (os tiros são "mortais" caso acertem no alvo). A, B e C acertam no alvo com prob 0.3, 1, 0.5 (resp.). O 1º a jogar, A, tem 3 estratégias à escolha: Escolher B para alvo, escolher C, ou passar a vez a B. Qual a melhor estratégia para sobreviver ao duelo?

Calcula-se $P(A \text{ sobreviver})$ de acordo com cada estratégia.

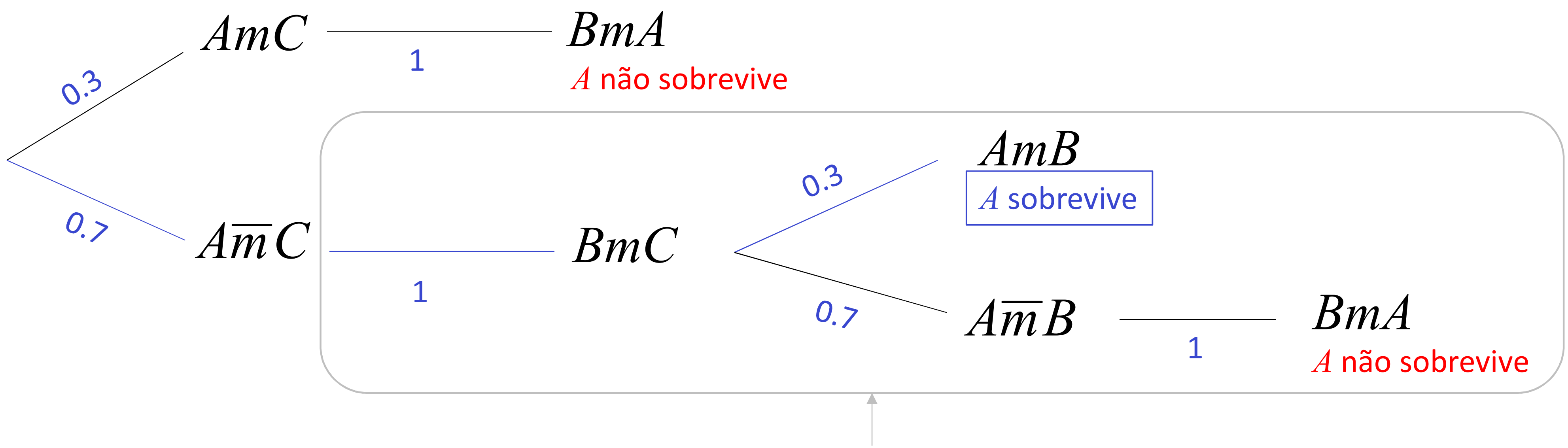
Estratégia 1 – A escolhe B para alvo

Estratégia 2 – A escolhe C para alvo

Estratégia 3 – A passa a vez a B

Escolhe-se a estratégia que maximiza $P(A \text{ sobreviver})$

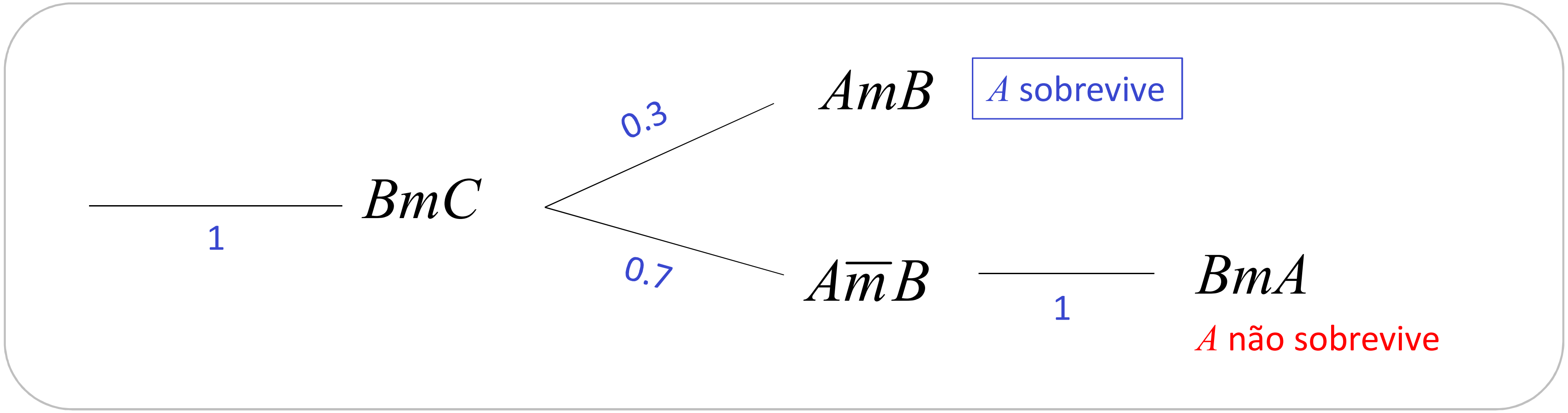
Estratégia 2 – A escolhe C; então AmC (A mata C) ou não; a seguir joga B; se AmC então BmA (A não sobrevive); se A~mC então BmC; depois joga A, que sobrevive sse AmB



$$P(A\ \ sobreviver) = 0.7 \times 1 \times 0.3 = 0.21$$

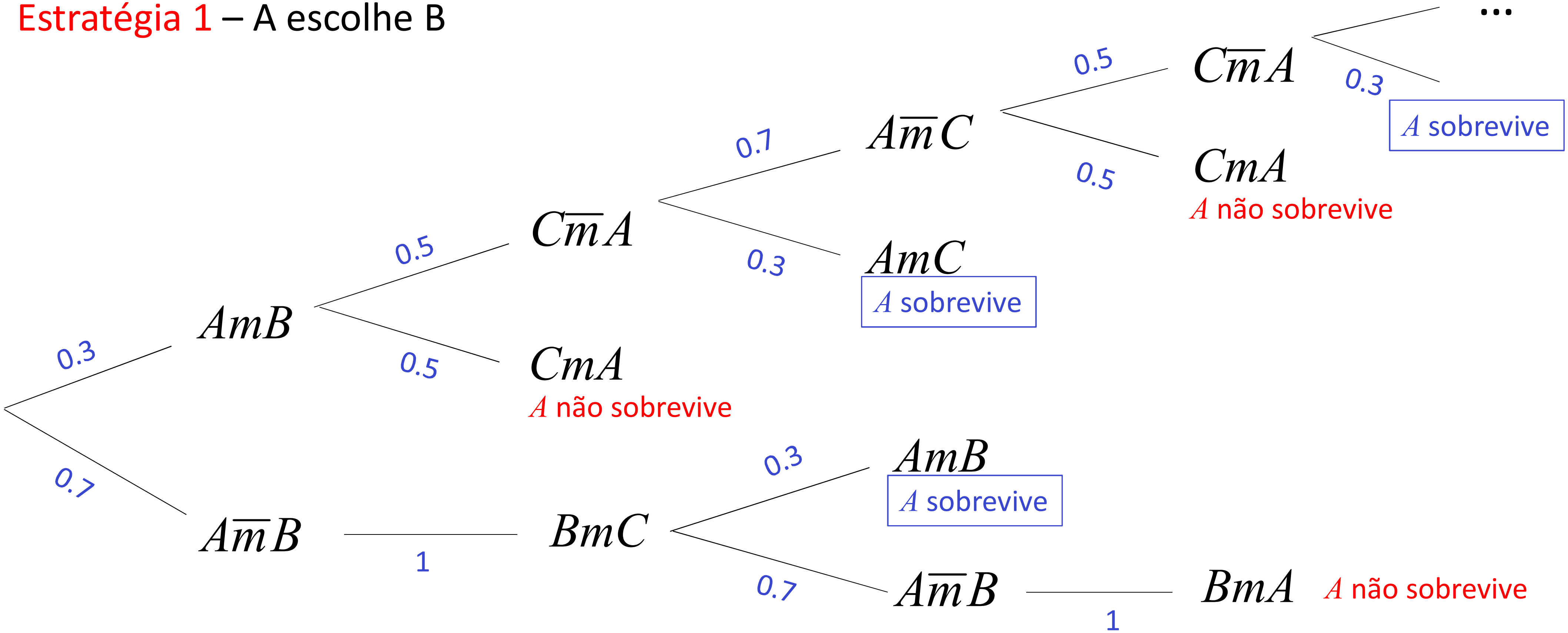
Mais valia passar a vez a B!

Estratégia 3 – A passa a vez; a seguir joga B; B escolhe C; BmC (B mata C); a seguir joga A; ou AmB (e então A sobrevive) ou não (A não sobrevive).



$$P(A\ \ sobreviver) = 1 \times 0.3 = 0.3$$

Estratégia 1 – A escolhe B



$$P(A\ sobreviver) = 0.7\times1\times0.3\ +\ 0.3\times0.5\times0.3\ +\ 0.3\times0.5\times(0.7\times0.5)\times0.3+\ldots=$$

$$=0.21+(0.3\times0.5\times0.3)\sum_{j\geq0}(0.7\times0.5)^j\ =0.21+0.045\tfrac{1}{1-0.35}=0.2792308$$

Fórmulas úteis:

Binómio de Newton:

$$(a+b)^n = \sum_{j=0}^{n} \binom{n}{j} a^j \, b^{n-j}$$

$$\Downarrow$$

caso $a = p$ e $b = 1 - p$:

$$1 = \sum_{j=0}^{n} \binom{n}{j} p^j \, (1-p)^{n-j}$$

caso $a = 1$ e $b = 1$:

$$2^n = \sum_{j=0}^{n} \binom{n}{j}$$

Se $\# \, \Omega = n$, então $\# \, \mathcal{P}(\Omega) = 2^n$

$$1+2+3+\ldots+n = \frac{n(n+1)}{2}$$

$$1^2+2^2+3^2+\ldots+n^2 = \frac{n(n+1)(2n+1)}{6}$$

Soma dos termos de uma progressão geométrica de razão x :

$$1+x+x^2+x^3+\ldots+x^n = \frac{1-x^{n+1}}{1-x}$$

$$\sum_{j=0}^{\infty} x^j = \frac{1}{1-x}, \quad |x|<1$$

Desenvolvimento em série de e^x :

$$e^x = 1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\ldots = \sum_{j\geq 0} \frac{x^j}{j!} \; , \; x \in \mathbb{R}$$

$$\lim_{x \to +\infty} \left(1 + \frac{a}{x} \right)^x = e^a$$

Desenvolvimento em série de $\log(1+x)$:

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots = \sum_{j\geq 1} (-1)^{j-1} \frac{x^j}{j} \; , \; |x|<1$$