J. SEBASTIÃO E SILVA

COMPÊNDIO DE MATEMÁTICA

1.º volume

2.º tomo

Curso Complementar do Ensino Secundário

Edição GEP

LISBOA

CAPÍTULO VII

INTRODUÇÃO À ESTATÍSTICA E AO CÁLCULO DAS PROBABILIDADES

1. **Lógica de atributos e lógica de conjuntos.** Já vimos que toda a propriedade (ou atributo) num dado universo U define um conjunto, que é o *conjunto de todos os indivíduos que têm essa propriedade*. Por exemplo, seja U o conjunto dos portugueses existentes numa dada época e sejam *e, c, s, r, a, m, p, b*, respectivamente, os atributos *estudante, casado, solteiro, ruivo, algarvio, menor de* 25 *anos, poeta, com bigode*. Cada um destes atributos define, no universo U, um conjunto; designemos, respectivamente, por E, C, S, R, A, M, P, B os conjuntos assim definidos. Deste modo, E é o *conjunto dos estudantes portugueses,* C o *conjunto dos portugueses casados,* etc. na referida época ([1]).

Também vimos como as operações lógicas de *conjunção, disjunção* e *negação* sobre atributos se traduzem, respectivamente, nas

---

([1]) Alguns destes conjuntos, como por exemplo R, não estão na realidade definidos. Tratando-se dum inquérito estatístico, relativo a indivíduos ruivos, haverá casos de dúvida, em que o autor do inquérito pode incluir um dado indivíduo no conjunto R ou no seu complementar, de modo *mais ou menos* arbitrário.

operações lógicas de *intersecção, reunião e complementação* sobre conjuntos. Trata-se, apenas, de duas linguagens ou pontos de vista equivalentes: o *ponto de vista da compreensão* (relativo a atributos) e o *ponto de vista da extensão* (relativo a conjuntos).

Assim, tornando ao exemplo anterior, sabemos que os referidos atributos se podem associar entre si de diferentes modos por meio das operações lógicas fundamentais, dando lugar a novos atributos. Por exemplo os atributos

$$e \wedge c \;,\; e \vee m \;,\; \sim s, \; e \wedge \sim s \;,\; e \wedge a \wedge m \;,\; p \vee b \;,\; \sim p \wedge \sim b,$$

que, em linguagem comum, se traduzem pelas expressões 'estudante casado', 'estudante ou menor de 25 anos', 'não solteiro', 'estudante não solteiro', 'estudante algarvio, menor de 25 anos', 'poeta ou com bigode', 'não poeta e sem bigode', definem respectivamente os conjuntos

$$E \cap C \;,\; E \cup M \;,\; \complement S \;,\; E \cap \complement S \;,\; E \cap A \cap M \;,\; P \cup B \;,\; \complement P \cap \complement B$$

que o aluno facilmente pode identificar (convém recorrer a diagramas de Venn).

Também vimos que a implicação entre atributos se traduz na inclusão entre conjuntos. Por exemplo, no universo considerado, o atributo c *implica* o atributo $\sim s$ (não lhe sendo, contudo, equivalente); por isso, o conjunto C *está contido* no conjunto $\complement S$. Assim, a implicação $c \Rightarrow \sim s$ traduz-se na inclusão $C \subset \complement S$.

Por sua vez, a equivalência entre atributos traduz-se na *identidade* entre conjuntos. Por exemplo:

$$\text{'}\sim(p \vee b) \Leftrightarrow \sim p \wedge \sim b\text{'} \quad \text{traduz-se por} \quad \text{'}\complement(P \cup B) = \complement P \cap \complement B\text{'}$$

Finalmente, vimos que, deste modo, a um *atributo universal* corresponde o *universo* e a um *atributo impossível* corresponde o *conjunto vazio;* e que dois *atributos incompatíveis* definem dois *conjuntos disjuntos*. Por exemplo, o atributo $s \vee \sim s$ é universal (portanto $S \cup \complement S = U$), o atributo $c \wedge s$ é impossível (ou seja $C \cap S = \emptyset$), os atributos c, s são incompatíveis (ou seja, os conjuntos C e S são disjuntos), etc.

2. **Terminologia e notações.** Daqui por diante será cómodo chamar *'produto lógico'* quer à conjunção (de atributos) quer à intersecção (de conjuntos), *'soma lógica'* quer à disjunção (de atributos) quer à reunião (de conjuntos) e *'contrário'* quer à negação (de atributos) quer ao complementar (de conjuntos). Ao mesmo tempo, dados dois atributos α, β ou dois conjuntos A, B, designaremos por

$$\alpha\beta, \quad AB, \quad \alpha + \beta, \quad A + B, \quad \tilde{\alpha}, \quad \tilde{A}\,(^1)$$

respectivamente, os produtos lógicos de α por β e de A por B, as somas lógicas de α com β e de A com B, e os contrários de α e de A (o símbolo $\tilde{\alpha}$ pode ler-se 'não α' e a expressão $\alpha + \beta$ pode continuar a ler-se 'α ou β').

Tornando ao exemplo do número anterior será fácil, agora, reconhecer os significados das expressões

$$ec, \quad e + m, \quad \tilde{s}, \quad e\tilde{s}, \quad e\,a\,m, \quad p + b, \quad \tilde{p}\tilde{b}$$
$$EC, \quad E + M, \quad \tilde{S}, \quad E\tilde{S}, \quad E\,A\,M, \quad P + B, \quad \tilde{P}\tilde{B}$$

Se convencionarmos considerar como *idênticos* dois atributos quando são equivalentes, é manifesto que o conjunto dos atributos

(1) O autor representa, aqui, $\complement A$ por $\tilde{A}$ — (N. do E.).

definidos num dado universo U é uma álgebra de Boole isomorfa à álgebra dos subconjuntos de U. Designaremos, neste caso, por **1** o atributo universal e por (0 o atributo impossível.

3. **Frequência absoluta de um atributo numa população.** Em vez de 'universo' usa-se, muitas vezes, em estatística o termo 'população', com significado idêntico. Assim, um conjunto de pessoas, um conjunto de árvores, um conjunto de livros, um conjunto de lâmpadas eléctricas, etc., etc. podem ser tomados como *populações* (ou *universos*) em diversas situações.

DEFINIÇÃO. *Chama-se frequência absoluta dum atributo* α *numa população finita* U *o número de indivíduos que possuem o atributo* α *(ou seja o número de elementos do conjunto* A *definido por* α *em* U).

Assim, *frequência absoluta do atributo* é o mesmo que *cardinal do conjunto* A *correspondente* — número que convencionámos representar pela notação $\#$ A. Representaremos pela notação $\Phi(\alpha)$ a frequência absoluta de α. É claro que se tem:

$$0 \leqslant \Phi(\alpha) \leqslant \# \mathrm{U}$$

Tornando ao exemplo do n.º 1 vemos que $\Phi(e)$ é a frequência absoluta do atributo *estudante* na população dos portugueses da referida época, isto é, *o número total de estudantes portugueses existentes nessa época.*

Atendendo às considerações dos dois números anteriores, bem como ao que foi estabelecido no Cap. III, n.º 2, (pág. 137, 1.º tomo), vemos que

Se dois atributos α, β são incompatíveis, a frequência absoluta do atributo $\alpha+\beta$ é igual à soma das frequências absolutas dos atributos α, β na população considerada; isto é, em fórmula

$$\boxed{\Phi(\alpha + \beta) = \Phi(\alpha) + \Phi(\beta), \quad \textit{se} \quad \alpha, \beta \textit{ são incompatíveis}}$$

Com efeito, se α, β, são incompatíveis, os conjuntos A, B correspondentes são disjuntos e, então, a frequência absoluta de $\alpha + \beta$ (ou seja o número de elementos de $A + B$) é a soma das frequências absolutas de α e de β (ou seja a soma dos cardinais de A e B).

É claro que a propriedade anterior se estende a um número qualquer (finito) de atributos, incompatíveis entre si dois a dois.

Se os atributos α, β, não são necessariamente incompatíveis, tem-se a seguinte propriedade que resulta do estabelecido no Cap. III, n.º 12 (pág. 157, 1.º tomo):

$$\boxed{\Phi(\alpha + \beta) = \Phi(\alpha) + \Phi(\beta) - \Phi(\alpha\,\beta)}$$

Assim, tornando ao exemplo do número I, vê-se que

$$\Phi(e + s) = \Phi(e) + \Phi(s) - \Phi(e\,s)$$

isto é: a frequência absoluta do atributo *estudante ou solteiro,* no universo dos portugueses, é a soma das frequências absolutas do atributo *estudante* e do atributo *solteiro,* menos a frequência do atributo *estudante solteiro,* no referido universo.

No caso geral, de *n* atributos quaisquer $a_1, \ldots, a_n$, num universo U, tem-se um novo aspecto da FÓRMULA DE DANIEL DA SILVA (pág. 159, 1.º tomo), com atributos no lugar de conjuntos.

EXERCÍCIO. Sendo α um atributo qualquer numa população finita U, indique condições equivalentes às seguintes: 1) $\Phi(\alpha)=0$, 2) $\Phi(\alpha) = \# U$.

4. **Frequência relativa.** Chama-se *frequência relativa* dum atributo α numa população finita U ao quociente da frequência absoluta de α pelo número de elementos de U. Designaremos por $fr(\alpha)$ a frequência relativa de α. Então, se for $\nu = \Phi(\alpha)$ e $n = \# U$, será, por definição:

$$fr(\alpha) = \frac{\nu}{n}$$

Por exemplo, sabe-se que, numa cidade com 23 528 habitantes, 9 253 desses pertencem ao sexo masculino e os restantes 14 275 ao sexo feminino. Os dois últimos números são, pois, as frequências absolutas do sexo masculino e do sexo feminino na referida população. As frequências relativas dos mesmos atributos serão, *com aproximação até às centésimas:*

$$\frac{9\,253}{23\,528} \approx 0{,}39 \quad , \quad \frac{14\,275}{23\,528} \approx 0{,}61$$

Também podemos dizer, neste caso, que a frequência relativa do primeiro atributo é de 39 % e a do segundo é de 61%. Dum modo geral, se for ν a frequência absoluta dum atributo num universo finito U e n o cardinal de U, a *frequência relativa de α em percentagem,* neste universo, será:

$$fr(\alpha) = \frac{100\,\nu}{n}\ \%$$

Vejamos um outro exemplo. Num dado país, com 8 438 250 habitantes, 82 % dos habitantes têm cabelos castanhos ou pretos, 14 %

têm os cabelos louros e 4 % têm os cabelos ruivos. Serão, pois, 0,82, 0,14 e 0,04 as frequências relativas dos referidos atributos, *até às centésimas.* As respectivas frequências absolutas serão:

$$0{,}82 \times 8\,432\,250 \approx 6\,900\,000 \quad \text{(castanhos ou pretos)}$$

$$0{,}14 \times 8\,432\,250 \approx 1\,200\,000 \quad \text{(louros)}$$

$$0{,}04 \times 8\,432\,250 \approx 300\,000 \quad \text{(ruivos)}$$

Assim, da definição resulta que é sempre:

$$0 \leqslant \mathrm{fr}(\alpha) \leqslant 1,$$

tendo-se $\mathrm{fr}(\alpha) = 0$ *sse* α é impossível e $\mathrm{fr}(\alpha) = 1$ *sse* α é universal.

Do estabelecido no número anterior resulta o seguinte:

Se α e β são atributos incompatíveis, a frequência relativa de $\alpha+\beta$ é igual à soma das frequências relativas de α e β; isto é, em fórmula:

$$\mathrm{fr}(\alpha + \beta) = \mathrm{fr}(\alpha) + \mathrm{fr}(\beta), \text{ se } \alpha, \beta \text{ são incompatíveis.}$$

Com efeito, pondo $\# \, U = n$, vem (justifique):

$$\mathrm{fr}(\alpha+\beta) = \frac{\Phi(\alpha+\beta)}{n} = \frac{\Phi(\alpha)+\Phi(\beta)}{n} = \frac{\Phi(\alpha)}{n} + \frac{\Phi(\beta)}{n} = \mathrm{fr}(\alpha) + \mathrm{fr}(\beta)$$

Se α, β não são necessariamente incompatíveis, tem-se a fórmula mais geral (justifique):

$$\boxed{\mathrm{fr}(\alpha + \beta) = \mathrm{fr}(\alpha) + \mathrm{fr}(\beta) - \mathrm{fr}(\alpha\beta)}$$

No caso geral de n atributos quaisquer num universo finito, tem-se um novo aspecto da FÓRMULA DE DANIEL DA SILVA. Da propriedade anterior deduz-se o seguinte corolário:

A frequência relativa do atributo contrário de α é $1 - \mathrm{fr}(\alpha)$.

Em fórmula:

$$\boxed{\mathrm{fr}(\tilde{\alpha}) = 1 - \mathrm{fr}(\alpha)}$$

Vamos ilustrar as considerações anteriores com um exemplo usual. Suponhamos que num dado exame se apresentaram 189 alunos e que os resultados foram os que constam da seguinte tabela:

Tabela 1

Classificações	N.º de alunos	Classificações	N.º de alunos
0	0	11	45
1	0	12	28
2	0	13	17
3	0	14	9
4	2	15	12
5	1	16	7
6	5	17	5
7	11	18	2
8	7	19	1
9	0	20	0
10	37		

Assim, nesta prova, a frequência absoluta da classificação 3 foi 0; a frequência absoluta da classificação 10 foi 37, etc. As frequências relativas das classificações, em percentagens, são as que constam da seguinte tabela:

Tabela 2

Classificações	Percentagem	Classificações	Percentagem
0	0,0	10	19,6
1	0,0	11	23,8
2	0,0	12	14,8
3	0,0	13	9,0
4	1,1	14	4,8
5	0,5	15	6,3
6	2,6	16	3,8
7	5,8	17	2,6
8	3,7	18	1,1
9	0,0	19	0,5
		20	0,0

Designemos por m, M, S, b, B, A, R, respectivamente, os atributos *medíocre, mau, suficiente, bom, muito bom, aprovado, reprovado*. Como habitualmente, chama-se 'suficiente' à soma lógica dos atributos correspondentes às classificações 10, 11, 12, 13, 'bom' à soma lógica dos atributos correspondentes a 14, 15, 16, 17, etc Por outro lado:

$$R = M + m, \quad A = S + b + B$$

Visto que as diferentes classificações são incompatíveis entre si, tem-se:

$$fr(S) = 0{,}196 + 0{,}238 + 0{,}148 + 0{,}090 = 0{,}672$$

$$fr(b) = 0{,}048 + 0{,}063 + 0{,}038 + 0{,}026 = 0{,}175$$

$$fr(B) = 0{,}011 + 0{,}005 = 0{,}016$$

Por sua vez, como os atributos S, b, B também são incompatíveis entre si, vem:

$$fr(A) = 0{,}672 + 0{,}175 + 0{,}016 = 0{,}863$$

Finalmente, como R é o *atributo contrário de* A, vem:

$$fr(R) = 1 - 0{,}863 = 0{,}137$$

5. **Frequência relativa do produto lógico. Primeiro exemplo de probabilidade.** Quando se pretende averiguar em que medida dois atributos ou dois fenómenos estão *ligados entre si*, numa dada população ou num dado tipo de experiências, o que há a fazer é um inquérito estatístico, de que vamos indicar os passos preliminares.

Suponhamos, por exemplo, que se trata de saber se, entre pessoas, a miopia está de qualquer modo associada com o facto de *ter olhos azuis*. O que desde logo ocorre é indagar, numa população tão numerosa e variada quanto possível, quais as percentagens de míopes:

1.º *entre os indivíduos com olhos azuis;*

2.º *na população total.*

Se estas duas percentagens são sensivelmente iguais, há razões para pensar que os atributos 'míope' e 'com olhos azuis' são *independentes*; se as duas percentagens se afastam consideravelmente uma da outra, seremos inclinados a admitir que os dois atributos estão *associados* ou *correlacionados*: em sentido *positivo*, se a primeira percentagem é maior que a segunda, em sentido *negativo* no caso oposto.

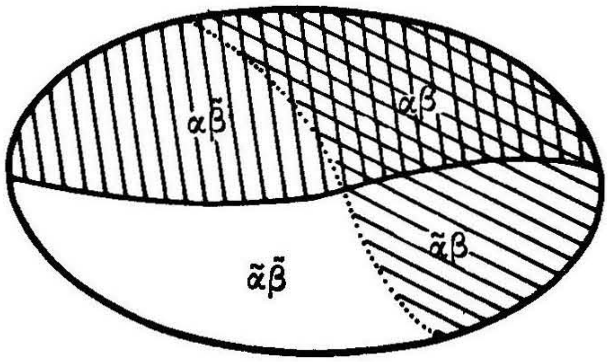
Analogamente se procederia para averiguar, por exemplo, se o *fumar* está ou não ligado com o *ter cancro dos pulmões*, se um insecticida é ou não eficaz no tratamento de certas plantas, etc. Nestes exemplos, começa a desenhar-se a hipótese duma relação *causa-efeito* entre os fenómenos considerados.

As considerações precedentes podem ser teorizadas do seguinte modo: sejam α, β, dois atributos num universo finito U constituído por n indivíduos. Para indicar as frequências dos atributos $\alpha\beta$, $\alpha\tilde{\beta}$, etc. podemos fazer uso duma tabela de duas entradas do seguinte tipo (chamada *tabela de contingência*):

	β	β	Total
α	$\Phi(\alpha\beta)$	$\Phi(\alpha\tilde{\beta})$	$\Phi(\alpha)$
$\tilde{\alpha}$	$\Phi(\tilde{\alpha}\beta)$	$\Phi(\tilde{\alpha}\tilde{\beta})$	$\Phi(\tilde{\alpha})$
Total	$\Phi(\beta)$	$\Phi(\tilde{\beta})$	n

Por exemplo, se α e β são, respectivamente, os atributos 'míope' e 'com olhos azuis', então $\alpha\beta$ é o atributo 'míope com olhos azuis', $\Phi(\alpha\tilde{\beta})$ é *o número de indivíduos míopes com olhos não azuis (no universo* U), etc. Neste caso, a frequência relativa do atributo 'míope' entre indivíduos com olhos azuis será:

$$\frac{\Phi(\alpha\beta)}{\Phi(\beta)}, \quad \text{supondo} \quad \Phi(\beta) \neq 0$$



Chamar-lhe-emos *frequência relativa de α se β* (isto é, *frequência relativa de α na hipótese de β se verificar)* e representá-la-emos por $\mathrm{fr}(\alpha|\beta)$, quaisquer que sejam os atributos α e β. Será, pois, por definição:

$$\mathrm{fr}(\alpha|\beta) = \frac{\Phi(\alpha\beta)}{\Phi(\beta)} \tag{1}$$

Por sua vez, a frequência relativa do atributo α (no universo considerado) será:

$$\mathrm{fr}\,(\alpha) = \frac{\Phi(\alpha)}{n}$$

Ora, segundo as considerações precedentes, os atributos α e β serão chamados *independentes* (em U), sse

$$\mathrm{fr}(\alpha|\beta) = \mathrm{fr}(\alpha) \tag{2}$$

ou, o que é equivalente, sse

$$\frac{\Phi(\alpha\beta)}{\Phi(\beta)} = \frac{\Phi(\alpha)}{n}$$

o que pode escrever-se de maneira mais simétrica:

$$\Phi(\alpha\beta) = \frac{\Phi(\alpha)\cdot\Phi(\beta)}{n} \tag{3}$$

Assim, qualquer das fórmulas (2) e (3) exprime que *os atributos α e β são independentes.*

Dividindo ambos os membros de (3) por *n* vem:

$$\frac{\Phi(\alpha\beta)}{n} = \frac{\Phi(\alpha)}{n} \cdot \frac{\Phi(\beta)}{n}$$

ou seja:

$$fr(\alpha\,\beta) = fr(\alpha)\,.\,fr(\beta),$$

o que é uma nova forma de exprimir a independência dos atributos α e β. Assim, teremos:

$$(4) \qquad \boxed{fr(\alpha\,\beta) = fr(\alpha)\cdot fr(\beta), \text{ sse } \alpha \text{ e } \beta \text{ são independentes}}$$

isto é: *a frequência relativa do produto lógico $\alpha\beta$ é o produto das frequências relativas de α e de β, sse estes atributos são independentes* (pela própria definição de 'atributos independentes').

Notemos, agora, que a fórmula (1) se pode escrever

$$\Phi(\alpha\beta) = \Phi(\beta)\cdot fr(\alpha|\beta)$$

ou seja, dividindo ambos os membros por n:

$$fr(\alpha\beta) = fr(\beta)\cdot fr(\alpha|\beta)$$

É claro que podemos trocar os papéis de α e de β:

$$(5) \qquad \boxed{fr(\alpha\beta) = fr(\alpha)\cdot fr(\beta|\alpha)}$$

isto é: quaisquer que sejam os atributos α e β, *a frequência relativa de α β é o produto de frequência relativa de α pela frequência relativa de β se α* (por definição de 'frequência relativa de β se α').

A fórmula (4) é, evidentemente, um caso particular de (5), pois que $fr(\beta|\alpha) = fr(\beta)$, sse α e β são independentes.

É claro que, na prática, só por coincidência se tem *exactamente* $fr(\beta|\alpha) = fr(\beta)$ ou, o que é equivalente:

$$\Phi(\alpha\beta) = \frac{\Phi(\alpha) \cdot \Phi(\beta)}{n}$$

O segundo membro desta igualdade representa-se por $\Phi_0(a\beta)$:

$$\Phi_0(\alpha\beta) = \frac{\Phi(\alpha) \cdot \Phi(\beta)}{n}$$

e chama-se *valor de independência* ou *valor esperado de* $\Phi(\alpha\beta)$ (isto é, *valor esperado na hipótese de* α *e* β *serem independentes*). Quase sempre o valor esperado é diferente do *valor observado* (ou *valor real*), $\Phi(\alpha\beta)$, e a diferença

$$\delta = \Phi(\alpha\beta) - \Phi_0(\alpha\beta)$$

é chamada *desvio* ou *discrepância* (entre o valor observado e o valor esperado).

Se o desvio δ é *relativamente pequeno* (ou *insignificante*), podemos atribuí-lo ao acaso e dizer que os atributos α e β são *independentes*, do ponto de vista da *estatística*. Se o desvio é *relativamente grande* (ou *significante*), já não é atribuível ao acaso e diremos que os atributos α e β estão *associados* (na população U).

Quanto ao significado das expressões 'relativamente pequeno', 'relativamente grande', 'acaso', etc., só mais tarde e progressivamente poderão ir sendo esclarecidos.

Note-se que, em particular, se pode ter

$$fr(\alpha|\beta) = 1 \quad \text{ou seja} \quad \Phi(\alpha\beta) = \Phi(\beta)$$

Por exemplo, se α e β são os atributos 'míope' e 'com olhos azuis' isto quereria dizer que o número de indivíduos *míopes com olhos azuis* é igual ao número de indivíduos *com olhos azuis* e que, portanto, *todos os indivíduos com olhos azuis são míopes* (isto é, que o atributo β *implica* o atributo α). Dum modo geral tem-se, quaisquer que sejam os atributos α e β:

$$\boxed{\mathrm{fr}(\alpha|\beta) = 1 \quad \text{sse} \quad \beta \Rightarrow \alpha}$$

(Demonstre, considerando conjuntos em vez de atributos e atendendo à propriedade característica dos conjuntos finitos considerada no Cap. III).

Diz-se que α e β estão *completamente associados* sse $\alpha \Rightarrow \beta$ ou $\beta \Rightarrow \alpha$, isto é sse $\mathrm{fr}(\beta|\alpha) = 1 \vee \mathrm{fr}(\alpha|\beta) = 1$.

• Um outro caso particular é aquele em que

$$\mathrm{fr}(\alpha|\beta) = 0 \text{ ou seja } \Phi(\alpha\beta) = 0 \; [\text{supomos } \Phi(\beta) \neq 0]$$

Isto significa manifestamente, que α e β são incompatíveis ou, o que é equivalente, que $\alpha \Rightarrow \tilde{\beta}$ ou ainda que $\beta \Rightarrow \tilde{\alpha}$ (no exemplo concreto considerado: 'nenhum indivíduo míope tem olhos azuis' ou 'nenhum indivíduo com olhos azuis é míope').

Diz-se que α e β estão *completamente dissociados* sse são incompatíveis os atributos α e β ou os atributos $\tilde{\alpha}$ e $\tilde{\beta}$, isto é, sse $\Phi(\alpha\beta) = 0 \vee \Phi(\tilde{\alpha}\tilde{\beta}) = 0$.

Facilmente se reconhece que:

Os atributos α e $\tilde{\beta}$ estão completamente dissociados, sse os atributos α e β estão completamente associados.

É claro que neste enunciado os papéis de α e β podem ser trocados.

EXEMPLO. Imaginemos uma experiência destinada a avaliar em que medida a inoculação de certa vacina imuniza contra a cólera. Suponhamos que os resultados obtidos são os que constam da tabela 3.

Tabela 3

	Não atacados	Atacados	Total
Inoculados	276	3	279
Não inoculados	473	66	539
Total	749	69	818

Sejam α e β, respectivamente, os atributos 'inoculado' e 'não atacado'. Então:

$$(\mathrm{fr})\ (\beta|a) = \frac{\Phi(a\beta)}{\Phi(a)} = \frac{276}{279} \approx 0{,}99$$

$$\mathrm{fr}\ (\beta) = \frac{\Phi(\beta)}{n} = \frac{749}{818} \approx 0{,}88$$

Neste caso, há uma diferença apreciável entre as duas frequências calculadas, indicativa de uma associação positiva entre os referidos atributos; e, como $\mathrm{fr}(\beta|\alpha)$ se aproxima bastante de 1, ao contrário de $\mathrm{fr}(\alpha|\beta)$ ($\approx$ 0,34), vemos que a referida associação se dá no sentido $\alpha \Rightarrow \beta$.

São, também, elucidativas as seguintes indicações:

1) Percentagem de inoculados que foram atacados: 1 %

2) Percentagem de não inoculados que foram atacados: 12 %

Parece, pois, concluir-se daqui que, embora a referida vacina não imunize *em absoluto,* confere no entanto uma *imunidade relativa* bastante apreciável, contra a cólera.

Se várias outras experiências, em condições análogas, conduzirem a um resultado sensivelmente igual a este quanto a $\mathrm{fr}(\alpha|\beta)$, seremos levados a dizer, por indução, que a vacina imuniza em 99 % dos casos. Exprimiremos também este facto dizendo que a *probabilidade de ficar imunizado* é 99 % (ou 0,99) e que, portanto, a *probabilidade de não ficar imunizado é* 1 % (ou 0,01).

NOTA IMPORTANTE. A implicação $\alpha \Rightarrow \beta$ deduzida de vários inquéritos estatísticos, com uma certa probabilidade, nem sempre pode ser interpretada como relação de *causa-efeito*. Suponhamos, por exemplo, que a frequência relativa de indivíduos calvos nas duas primeiras filas dos teatros se aproxima de 1 mais do que nas restantes filas. É óbvio que tal observação nunca poderia ser interpretada deste modo:

O facto de uma pessoa se sentar numa das duas primeiras filas de um teatro pode produzir-lhe a calvície.

Um dos problemas da estatística será, pois, o de estabelecer critérios que permitam distinguir entre implicações de causalidade e outras que o não são.

5. Coeficiente de associação*. Sejam ainda α e β dois atributos definidos num universo finito U. Já vimos que se chama *desvio* (ou *discrepância*) à diferença

$$\delta = \Phi(\alpha\beta) - \frac{\Phi(\alpha)\cdot\Phi(\beta)}{n}$$

entre o *valor observado* e o *valor de independência* de $\Phi(\alpha\beta)$. Mais tarde estudaremos um teste estatístico que nos habilita em certos casos a decidir, com maior ou menor *segurança*, se tal desvio não é devido ao *acaso* ([1]).

([1]) Aqui, 'segurança' significa 'probabilidade de não errar'. Por enquanto, temos de nos contentar com o significado intuitivo de tais expressões, que irá sendo progressivamente esclarecido.

Vamos, agora, substituir o desvio δ por um *desvio relativo,* que nos dá melhor ideia do grau de associação dos atributos α, β. Tal é, por exemplo, o número Q dado pela fórmula:

$$Q = \frac{n\delta}{\Phi(\alpha\,\beta)\,\Phi(\tilde{\alpha}\,\tilde{\beta}) + \Phi(\alpha\,\tilde{\beta})\,\Phi(\tilde{\alpha}\,\beta)} \tag{1}$$

ou seja:

$$Q = \frac{(n\Phi\alpha\,\beta) - \Phi(\alpha)\,\Phi(\beta)}{\Phi(\alpha\,\beta)\,\Phi(\tilde{\alpha}\,\tilde{\beta}) + \Phi(\alpha\,\tilde{\beta})\,\Phi(\tilde{\alpha}\,\beta)} \tag{2}$$

Este é chamado, precisamente, o *coeficiente de associação de* α *e* β. Notemos que se tem (justifique):

$$(\alpha + \tilde{\alpha})\,(\beta + \tilde{\beta}) = \alpha\,\beta + \tilde{\alpha}\,\beta + \alpha\,\tilde{\beta} + \tilde{\alpha}\,\tilde{\beta} = \mathbf{1}\ \text{(1)}$$

Visto que as quatro parcelas são atributos incompatíveis entre si, daqui se deduz:

$$\Phi(\alpha\,\beta) + \Phi(\tilde{\alpha}\,\beta) + \Phi(\alpha\,\tilde{\beta}) + \Phi(\tilde{\alpha}\,\tilde{\beta}) = n \tag{3}$$

Por outro lado:

$$\alpha = \alpha\,\beta + \alpha\,\tilde{\beta} \quad , \quad \beta = \alpha\,\beta + \tilde{\alpha}\,\beta \qquad \textit{(Porquê?)}$$

Donde:

$$\Phi(\alpha) = \Phi(\alpha\,\beta) + \Phi(\alpha\,\tilde{\beta}), \quad \Phi(\beta) = \Phi(\alpha\,\beta) + \Phi(\tilde{\alpha}\,\beta) \tag{4}$$

---

(1) Lembremos que se designa por **1** o atributo universal e por **0** o atributo impossível (ver n.º 2).

Finalmente, substituindo em (2) as expressões de n, $\Phi(\alpha)$ e $\Phi(\beta)$ dadas por (3) e (4), e simplificando, vem:

$$(5) \qquad Q = \frac{\Phi(\alpha\,\beta)\,\Phi(\tilde{\alpha}\,\tilde{\beta}) - \Phi(\alpha\,\tilde{\beta})\,\Phi\,(\tilde{\alpha}\,\beta)}{\Phi(\alpha\,\beta)\,\Phi(\tilde{\alpha}\,\tilde{\beta}) + \Phi(\alpha\,\tilde{\beta})\,\Phi\,(\tilde{\alpha}\,\beta)}$$

É fácil ver que se tem sempre:

$$-1 \leqslant Q \leqslant 1$$

Posto isto, há a distinguir três casos notáveis:

1.° *Tem-se* $Q = 0$, *sse* α *e* β *são independentes.* Para o reconhecer, basta atender a (1) e lembrar que α e β são independentes, por definição, sse $\delta = 0$.

2.° *Tem-se* $Q = -1$, *sse* α *e* β *estão completamente dissociados.*

Com efeito, se α e β estão completamente dissociados, tem-se por definição

$$\Phi(\alpha\,\beta) = 0 \quad \vee \quad \Phi\,(\tilde{\alpha}\,\tilde{\beta}) = 0$$

o que implica $Q = -1$. Reciprocamente, é fácil ver que, se $Q = -1$, será necessariamente $\Phi(\alpha\,\beta)\ \Phi\ (\tilde{\alpha}\,\tilde{\beta}) = 0$ e, portanto:

$\Phi(\alpha\,\beta) = 0 \vee \Phi(\tilde{\alpha}\,\tilde{\beta}) = 0$. Basta lembrar que em $\mathbb{R}$:

$$\frac{x-y}{x+y} = -1 \Leftrightarrow x = 0, \text{ se } x + y \neq 0$$

3.° *Tem-se* $Q = 1$, *sse* α *e* β *estão completamente associados.*

Este caso reduz-se ao anterior, lembrando que α e β estão completamente associados, sse α e $\tilde{\beta}$ estão completamente dissociados.

Finalmente demonstraremos o seguinte

TEOREMA: *Se* (α, β) *é um par de atributos independentes, também* $(\tilde{\alpha}, \beta)$ $(\alpha, \tilde{\beta})$ *e* $(\tilde{\alpha}\tilde{\beta})$ *são pares de atributos independentes.*

Com efeito, como vimos, se α e β são independentes, tem-se $Q = 0$. Então, por exemplo, o coeficiente de associação de $\tilde{\alpha}$ e β será:

$$Q' = \frac{\Phi(\tilde{\alpha}\,\beta)\,\Phi(\alpha\,\tilde{\beta}) - \Phi(\tilde{\alpha}\tilde{\beta})\,\Phi(\alpha\,\beta)}{\Phi(\tilde{\alpha}\,\beta)\,\Phi(\alpha\,\tilde{\beta}) + \Phi(\tilde{\alpha}\tilde{\beta})\,\Phi(\alpha\,\beta)} = -\,Q = 0$$

e, portanto, $\tilde{\alpha}$ e β são independentes. Analogamente se procede nos restantes casos.

6. **Extensão dos conceitos do n.º 4 a mais de dois atributos.** Consideremos três atributos α, β, γ num universo finito U. Por exemplo, α, β, γ podem ser, respectivamente, os atributos 'fumador', 'alcoólico', 'com úlcera gástrica ou duodenal', num conjunto de pessoas. Visto que

$$\alpha\beta\gamma = (\alpha\beta)\gamma,$$

tem-se, aplicando a fórmula (5) do n.º 4:

$$\mathrm{fr}(\alpha\beta\gamma) = \mathrm{fr}(\alpha\beta)\cdot\mathrm{fr}(\gamma|\alpha\beta)$$

e, como $\mathrm{fr}(\alpha\beta) = \mathrm{fr}(\alpha)\cdot\mathrm{fr}(\beta\,|\,\alpha)$, vem, finalmente:

$$\boxed{\mathrm{fr}(\alpha\beta\gamma) = \mathrm{fr}(\alpha)\cdot\mathrm{fr}(\beta|\alpha)\cdot\mathrm{fr}(\gamma|\alpha\beta)}$$

É claro que nesta fórmula as letras α, β, γ podem ser permutadas de todos os modos possíveis, visto que α, β, γ são atributos *quaisquer*.

Em particular, *se for*

$$(1) \qquad \mathrm{fr}(\beta|\alpha) = \mathrm{fr}(\beta) \ \text{ e } \ \mathrm{fr}(\gamma|\alpha\,\beta) = \mathrm{fr}(\gamma),$$

virá:

$$(2) \qquad \boxed{\mathrm{fr}(\alpha\,\beta\,\gamma) = \mathrm{fr}(\alpha)\cdot\mathrm{fr}(\beta)\cdot\mathrm{fr}(\gamma)}$$

DEFINIÇÃO. *Diz-se que três atributos α, β, γ são independentes (no universo U), sse verificam não só as condições (1), mas ainda as que se deduzem dessas, substituindo um ou mais dos atributos α, β, γ pelos seus contrários, como por exemplo* [1]*:*

$$\mathrm{fr}(\gamma|\tilde{\alpha}\,\beta) = \mathrm{fr}(\gamma), \quad \mathrm{fr}(\gamma|\tilde{\alpha}\,\tilde{\beta}) = \mathrm{fr}(\gamma), \ \text{etc.}$$

Desde logo se vê que esta definição equivale à seguinte proposição:

Os atributos α, β, γ, são independentes, sse, além da condição (2), *verificam as condições tais como*

$$\mathrm{fr}(\alpha\tilde{\beta}\gamma) = \mathrm{fr}(\alpha)\cdot\mathrm{fr}(\tilde{\beta})\cdot\mathrm{fr}(\gamma),$$
$$\mathrm{fr}(\tilde{\alpha}\beta\tilde{\gamma}) = \mathrm{fr}(\tilde{\alpha})\cdot\mathrm{fr}(\beta)\cdot\mathrm{fr}(\tilde{\gamma}), \ \text{etc.}$$

que se deduzem de (2) *substituindo um ou mais dos atributos α, β, γ pelos respectivos contrários.*

Isto mostra imediatamente que *a relação ternária assim definida entre atributos é simétrica.*

---

(1) Parte das condições assim obtidas são consequência das restantes. Por exemplo, a equação $\mathrm{fr}(\beta|\alpha) = \mathrm{fr}(\beta)$ equivale a $\mathrm{fr}(\tilde{\beta}|\alpha) = \mathrm{fr}(\tilde{\beta})$, etc. (Ver teorema do número anterior).

Mas, notem-se os dois seguintes pontos importantes:

1) Os atributos α, β, γ podem ser *independentes dois a dois* (isto é, α independente de β, α independente de γ, e β independente de γ), sem serem os três independentes.

2) Ao contrário do que sucede no caso de dois atributos (ver teorema do número anterior), a fórmula (2) pode ser verificada sem que os atributos α, β, γ sejam independentes.

É preciso também não confundir 'atributos independentes' com 'atributos incompatíveis'.

É claro que estas considerações se podem generalizar imediatamente ao caso de um número finito *n* qualquer de atributos.

7. **A lógica em termos de acontecimentos.** Recordemos que, em alguns dos exemplos anteriores, os atributos considerados também se podem interpretar como *acontecimentos*. Assim, no exemplo final do número 4, os atributos 'inoculado' e atacado' correspondem, respectivamente, aos acontecimentos 'ser inoculado (com vacina)' e 'ser atacado (de cólera)'.

Não pretendemos aqui definir 'acontecimento', assim como não tentámos definir 'atributo' nem 'conjunto': trata-se de conceitos psicologicamente primitivos, gerados por indução ou intuição no nosso espírito. O que será possível e conveniente é esclarecer progressivamente a terminologia que lhes diz respeito.

Comecemos por notar que, em vez de 'acontecimento', se usam como significado semelhante os termos 'facto', 'fenómeno', 'eventualidade', etc.

Imaginemos uma prova ou experiência *que se possa repetir várias vezes em condições idênticas,* conduzindo, de cada vez, a um ou mais *resultados,* entre vários que são de prever. É a cada um desses resultados da prova que, em cálculo das probabilidades, se costuma dar o nome de *'acontecimento'*.

São inúmeros os exemplos que neste sentido se podem apresentar: desastre ou ausência de desastre numa viagem aérea, resultados dum exame ou duma prova desportiva, etc.

Mas convém registar o que há pouco foi observado: que muitas vezes os acontecimentos se traduzem por atributos (e vice-versa). Tal é, por exemplo, o caso dos resultados dum exame ou o caso análogo duma competição desportiva; assim, o acontecimento 'ficar reprovado' traduz-se pelo atributo 'reprovado'; o acontecimento 'vitória' traduz-se pelo atributo 'vencedor', relativamente a um dado clube; etc., etc.

Ainda aqui há portanto, de certo modo, uma questão de ponto de vista psicológico, semelhante à que se põe na distinção entre atributos e conjuntos. É, assim, de prever que a lógica de atributos se traduza numa lógica de acontecimentos. Com efeito, sejam α e β dois acontecimentos relativos a uma dada prova $\mathcal{P}$:

1) Chama-se *conjunção* (ou *produto lógico*) de α e β, e representa-se por $\alpha\beta$, o acontecimento que consiste na realização simultânea de α e de β.

2) Chama-se *disjunção* (ou *soma lógica*) de α e β e representa-se por $\alpha + \beta$, o acontecimento que consiste em se realizar *um, pelo menos,* dos acontecimentos α e β.

3) Chama-se *contrário de* α, e representa-se por $\tilde{\alpha}$, o acontecimento que consiste em não se realizar α.

4) Diz-se que α *implica* β, e escreve-se $\alpha \Rightarrow \beta$, sse β se realiza todas as vezes que se realiza α.

5) Dois acontecimentos α, β dizem-se *equivalentes* e escreve-se $\alpha \Rightarrow \beta$, sse $\alpha \Rightarrow \beta$ e $\beta \Rightarrow \alpha$.

6) Um acontecimento diz-se *certo,* sse sabemos, com certeza absoluta, que se realizará na prova $\mathcal{P}$, todas as vezes que esta for efectuada. Um acontecimento diz-se *impossível,* sse o seu contrário é certo.

7) Dois acontecimentos dizem-se *incompatíveis*, sse a sua conjunção é acontecimento impossível.

Por exemplo, no exame dum aluno é certo o acontecimento *aprovação* ou *reprovação* ou *desistência*, é impossível o acontecimento *aprovado com* 21 *valores* (em escolas portuguesas) e são incompatíveis os dois acontecimentos *aprovação* e *desistência*.

8. **Expressões proposicionais de acontecimentos; conceito de variável casual; passagem a conjuntos.** Tal como sucede com os atributos, os acontecimentos habitualmente considerados em estatística e em cálculo das probabilidades podem ser indicados por meio de expressões proposicionais com variáveis. Vejamos dois exemplos:

a) Imaginemos um saco que contenha várias bolas, umas brancas e outras pretas. Designando por U o universo das bolas contidas no saco, por B o conjunto das bolas brancas e por P o conjunto das bolas pretas, teremos duas expressões proposicionais:

$$x \in B \quad , \quad x \in P,$$

definidas em U. Os valores possíveis da variável x serão, pois, as bolas do conjunto U. Ora, o valor de x pode ser determinado extraindo *ao acaso* uma bola de U (no final discutiremos o significado da palavra 'acaso'). Nesta prova — extracção de uma bola — realiza-se então necessariamente um dos seguintes acontecimentos contrários: $x \in B$ (isto é: *sair bola branca*), $x \in P$ (isto é: *sair bola preta*). Vê-se, pois, como as referidas expressões proposicionais passam a indicar acontecimentos.

b) Seja x a classificação dum aluno numa prova $\mathcal{P}$ a realizar. Os resultados 'mau', 'medíocre', 'suficiente', 'bom', e 'muito bom' são agora indicados, respectivamente, pelas expressões proposicionais seguintes:

$$0 \leqslant x < 6, \quad 6 \leqslant x < 10, \quad 10 \leqslant x < 14, \quad 14 \leqslant x < 18, \quad 18 \leqslant x < 20$$

O resultado 'aprovado', soma lógica de 'suficiente', 'bom' e 'muito bom', é indicado pela expressão $x \geqslant 10$.

Assim, em casos como estes, cada um dos acontecimentos a que pode dar lugar a prova $\mathcal{P}$ considerada aparece sob a forma de uma expressão proposicional com uma variável x, num determinado universo U. O valor da variável x (elemento de U) é, então, determinado em cada realização da prova $\mathcal{P}$ por um processo em que intervém mais ou menos o *acaso* e que, por isso, não permite prever qual será exactamente esse valor. Exprime-se este facto dizendo que x é uma *variável casual* (ou uma *variável aleatória*).

É claro que a conjunção de dois acontecimentos α e β será indicada pela conjunção das expressões proposicionais que indicam α e β, e analogamente para a disjunção, para a negação, etc.

Deste modo, a cada acontecimento α fica a corresponder um determinado conjunto A em U: o conjunto dos elementos que verificam a expressão indicativa de α. Aos dois acontecimentos α e β corresponde então um mesmo conjunto, sse $\alpha \Leftrightarrow \beta$.

Reciprocamente, a cada conjunto A em U corresponde o acontecimento indicado pela expressão $x \in A$.

Vemos assim que, tal como sucede com os atributos:

A lógica de acontecimentos é traduzida, segundo a referida correspondência, na lógica de conjuntos.

Notemos ainda que, tal como para atributos:

Se convencionarmos considerar como *idênticos* dois acontecimentos α e β quando são *equivalentes*, escrevendo então $\alpha = \beta$ em vez de $\alpha \Leftrightarrow \beta$, o conjunto dos acontecimentos relativos a uma dada prova $\mathcal{P}$ constitui uma álgebra de Boole isomorfa a uma álgebra de conjuntos. Designaremos, então, por $\mathbb{1}$ o *acontecimento certo* e por $\mathbb{O}$ o *acontecimento impossível*.

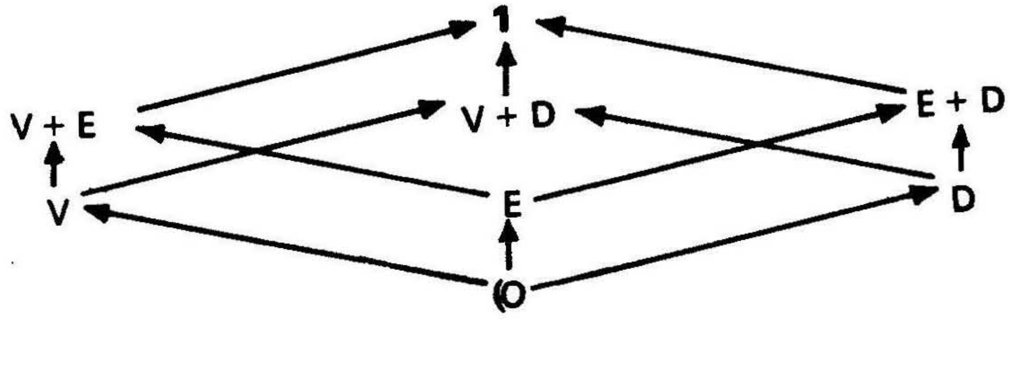
Por exemplo, consideremos as seguintes eventualidades relativas a um desafio de futebol entre duas equipas A e B:

$$V = \text{vitória de A}, \quad E = \text{empate}, \quad D = \text{derrota de A}$$

Então $V \cdot E = V \cdot D = E \cdot D = \mathbb{O}$. Por outro lado, tratando-se dum *desafio normal*, tem-se:

$$V + E + D = \mathbb{1}$$

Além destas, há ainda a considerar as seguintes eventualidades: V + E, V + D, E + D. Temos, assim, ao todo 8 hipóteses (que se podem considerar, por exemplo, em cada uma das provas indicadas num bilhete de TOTOBOLA). É evidente que o conjunto dessas oito hipóteses, com as operações de soma lógica e produto lógico, é uma álgebra de Boole, representada no seguinte esquema, em que usamos setas simples como símbolos de implicação:



É também fácil reconhecer que esta *álgebra de acontecimentos é isomorfa à álgebra dos subconjuntos do conjunto* $\{V, E, D\}$.

NOTA SOBRE O CONCEITO DE 'ACASO'. Numa primeira aproximação, poderíamos dizer que o termo 'acaso' significa 'ausência de causa' ou, pelo menos, 'ausência de causa conhecida', o que torna impossível a previsão de certos acontecimentos. Por exemplo, quando lançamos *ao acaso* uma moeda de um escudo ao ar ou quando tiramos *à sorte* uma bola de loto de um saco, somos incapazes de prever se sairá escudo ou face, ou qual o número da bola que vai aparecer. Diz-se, então, que se trata de acontecimentos *casuais (fortuitos, aleatórios* ou *eventuais),* isto é, de acontecimentos que não estão determinados *a priori,* e que, portanto, se podem verificar umas vezes e outras não.

Para certos autores, o *acaso* consistiria na acumulação de um grande número de pequenas causas desconhecidas, que actuam em diversos sentidos, tornando praticamente impossível a previsão do efeito global. Mas esta interpretação filia-se ainda no ponto de vista do *determinismo mecanicista,* que se admitiu no século passado e que é sintetizado pelas seguintes palavras de LAPLACE no seu *Essai philosophique sur les probabilités:*

'Um intelecto de tal modo vasto que conhecesse o estado e as posições relativas de todos os entes da natureza num dado instante, assim como todas as forças que os solicitam; intelecto que fosse, além disso, bastante poderoso para submeter todos esses dados à análise matemática — uma tal inteligência poderia abranger numa só fórmula o movimento dos maiores corpos e das mais ínfimas partículas do universo: então, nada ficaria incerto e tanto o passado como o futuro se tornariam presentes aos seus olhos' ([1]).

Ora, a ciência do século XX veio mostrar que esta posição é ilusória sob vários aspectos. É-se levado hoje a admitir que, na evolução do mundo físico, subsiste sempre algo de essencialmente imprevisível, isto é, um certo grau de incerteza radical, que não resulta apenas da nossa ignorância.

([1]) O germe do determinismo mecanicista encontra-se já em DESCARTES, que dizia: *'Dai-me o espaço e o movimento, eu vos darei o mundo'.*

9. **Frequência dum acontecimento numa sequência de provas.** Comecemos por dois exemplos:

a) Numa série de 30 000 viagens efectuadas por uma dada companhia de aviação, houve desastres apenas em duas viagens. Diremos, então, que a *frequência absoluta* do acontecimento *desastre* nesta sequência de viagens foi 2 e que a *frequência relativa* do mesmo acontecimento nessa mesma sequência foi:

$$\frac{2}{30\,000} \approx 0{,}000067 \approx 7/100\,000$$

b) Numa série de competições, um dado clube desportivo teve 9 vitórias, 4 derrotas e 1 empate. Nesta sequência de provas as frequências relativas dos acontecimentos *vitória, derrota* e *empate* foram, respectivamente:

$$\frac{9}{14} \approx 0{,}64 = 64\,\%, \quad \frac{4}{14} \approx 0{,}29 = 29\,\%, \quad \frac{1}{14} \approx 0{,}07 = 7\,\%$$

Dum modo geral:

DEFINIÇÃO 1. *Chama-se frequência absoluta dum acontecimento* α *numa sequência de* n *provas* $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_n$, *consideradas como realizações de uma mesma prova-tipo* $\mathcal{P}$, *o número* ν *de vezes que* α *se verifica nessas provas. Chama-se frequência relativa de* α *na mesma sequência de provas o quociente* ν/n *da frequência absoluta pelo número total de provas da sequência*. Designaremos por $\Phi(\alpha)$ a frequência absoluta de α e por $fr(\alpha)$ a frequência relativa de α na referida sequência.

Será, pois, por definição:

$$(1) \qquad fr(\alpha) = \frac{\Phi(\alpha)}{n}$$

Desta definição resultam imediatamente as propriedades:

I. *Tem-se sempre* $0 \leqslant \mathrm{fr}(\alpha) \leqslant 1$

II. *Se* α *é certo, então* $\mathrm{fr}(\alpha) = 1$ (*Porquê?*)

III. *Se* α *é impossível, então* $\mathrm{fr}(\alpha) = 0$ (*Porquê?*)

Como no caso dos atributos, também se demonstra:

IV. *Se* α *e* β *são acontecimentos relativos à prova* $\mathcal{P}$, *então*

$$\mathrm{fr}(\alpha+\beta) = \mathrm{fr}(\alpha) + \mathrm{fr}(\beta) - \mathrm{fr}(\alpha\beta)$$

Por sua vez de IV e III deduz-se (justifique):

V. *Se* α *e* β *são incompatíveis, então*

$$\mathrm{fr}(\alpha+\beta) = \mathrm{fr}(\alpha) + \mathrm{fr}(\beta)$$

Assim, no exemplo *b*) anterior, os acontecimentos *vitória* e *derrota* são incompatíveis e, portanto:

$$\mathrm{fr}(V + D) = \mathrm{fr}(V) + \mathrm{fr}(D) = \frac{9}{14} + \frac{4}{14} = \frac{13}{14} \approx 93\,\%$$

Finalmente de II e V deduz-se:

VI. $$\mathrm{fr}(\tilde{\alpha}) = 1 - \mathrm{fr}(\alpha)$$

Com efeito tem-se, por definição, $\alpha + \tilde{\alpha} = \mathbb{1}$ e $\alpha\tilde{\alpha} = \mathbb{O}$, donde, aplicando II, $\mathrm{fr}(\alpha + \tilde{\alpha}) = 1$ e, aplicando V,

$$\mathrm{fr}(\alpha) + \mathrm{fr}(\tilde{\alpha}) = 1 \quad \text{e portanto} \quad \mathrm{fr}(\tilde{\alpha}) = 1 - \mathrm{fr}(\alpha)$$

Tornando ao exemplo 2) anterior, temos $V + D = \bar{E}$ e, portanto:

$$\mathrm{fr}(V + D) = \mathrm{fr}(\bar{E}) = 1 - \mathrm{fr}(E) = 1 - \frac{1}{14} = \frac{13}{14} \approx 93\,\%$$

Passemos, agora, ao caso do produto lógico:

DEFINIÇÃO 2. *Se* α *e* β *são acontecimentos relativos a uma prova* $\mathcal{P}$, *chama-se 'frequência relativa de* α *se* β', *numa sequência de realizações de* $\mathcal{P}$, *e representa-se por* $\mathrm{fr}(\alpha\,\beta)$, *o quociente* μ/ν *do número* μ *de vezes que se realiza* $\alpha\beta$ *pelo número* ν *de vezes que se realiza* β Será, pois:

$$\mathrm{fr}(\alpha|\beta) = \frac{\Phi(\alpha\beta)}{\Phi(\beta)}$$

Dividindo ambos os termos desta fracção pelo número *n* de realizações de $\mathcal{P}$ e atendendo a (1), obtém-se:

$$\mathrm{fr}(\alpha|\beta) = \frac{\mathrm{fr}(\alpha\,\beta)}{\mathrm{fr}(\beta)}$$

ou seja: $\mathrm{fr}(\alpha\,\beta) = \mathrm{fr}(\beta)\cdot\mathrm{fr}(\alpha|\beta)$, ou ainda, invertendo os papéis de α e β:

$$\boxed{\mathrm{fr}(\alpha\,\beta) = \mathrm{fr}(\alpha)\;\;\mathrm{fr}(\beta|\alpha)}$$

Os acontecimentos α, β dizem-se *independentes da referida sequência de provas,* sse $\mathrm{fr}(\alpha|\beta) = \mathrm{fr}(\alpha)$, ou, o que é equivalente, $\mathrm{fr}(\beta|\alpha) = \mathrm{fr}(\beta)$, ou ainda

$$\Phi(\alpha\,\beta) = \frac{\Phi(\alpha)\;\Phi\;(\beta)}{n}$$

Suponhamos, por exemplo, que a prova $\mathcal{P}$ é *desafio de futebol dum clube* A *com um outro clube qualquer,* sendo α o acontecimento vitória de A e β o acontecimento *presença do jogador* X na equipa (concretize com casos do seu conhecimento). Neste exemplo, $\mathrm{fr}(\alpha|\beta)$ é a frequência relativa de vitórias do clube A nos desafios em que entra X. Se $fr(\alpha|\beta) = fr(\alpha)$, diremos que, na sequência de jogos considerados, o acontecimento *vitória de* A foi independente do acontecimento *presença do jogador* X.

Um outro exemplo: suponhamos que $\mathcal{P}$ designa *viagem de automóvel,* sendo α o acontecimento *desastre* e β o acontecimento *chuva.* É então fácil interpretar o significado do símbolo $fr(\alpha|\beta)$.

Estes exemplos começam a sugerir a ideia de causalidade. Tal ideia está intimamente ligada ao processo de indução, a que já nos referimos no Cap. I, n.º 17, págs. 49-50, 1.º tomo, e ao conceito de probabilidade, de que trataremos seguidamente.

10. **Lógica indutiva; certeza absoluta e certeza prática.** Como vimos no número anterior, *se um acontecimento* α *é certo numa prova* $\mathcal{P}$*, então* $fr(\alpha) = 1$ *em qualquer sequência de realizações* de $\mathcal{P}$. Mas será a recíproca verdadeira, isto é: *se* $fr(\alpha) = 1$ numa sequência de provas, podemos daí concluir que α é acontecimento certo? Claro que não.

Consideremos um exemplo. Imaginemos um saco que contém 50 bolas, *todas brancas,* numeradas de 1 a 50 e seja $\mathcal{P}$ a seguinte prova: *tirar ao acaso uma bola do saco, tornando a colocá-la depois no saco.* Então o acontecimento *saída de bola branca* é certo (ao contrário dos acontecimentos *saída do n.º* 5, *saída de número par,* etc.) e, por isso, numa sequência qualquer de realizações de $\mathcal{P}$, a frequência relativa de *saída de bola branca* é necessariamente 1. Mas imaginemos, agora, a situação inversa: o saco contém bolas cuja cor ignoramos e alguém efectua a prova $\mathcal{P}$ um grande número de vezes. Suponhamos que a frequência relativa do acontecimento *saída de bola branca* é então 1 (isto é, que sai bola branca em todas as extracções, com reposição). Podemos nós daí concluir que tal acontecimento é certo, isto é, que se verificará *sempre,* em qualquer prova futura? Claro que não: só podemos ter a *certeza absoluta,* tirando todas as bolas do saco e examinando-as uma a uma.

Porém, se a frequência relativa desse acontecimento for 1 num grande número de provas, por exemplo 1000, começamos a *convencer-nos* de que o acontecimento é certo, e essa convicção aumentará com o número de provas, se a frequência relativa continuar a ser 1. Em vez de uma *certeza autêntica* (ou *certeza absoluta*) passamos a ter uma *certeza prática* (ou *certeza relativa*).

A maior parte das certezas em que nos baseamos no decorrer da nossa vida são *certezas práticas* e não *certezas absolutas*. Por exemplo, temos a certeza dos seguintes factos: '*O Sol nasce amanhã*', '*Se largarmos um copo sem apoio, o copo cai*', '*A água quando gela aumenta de volume*', '*O gelo flutua na água líquida*', etc. Mas trata-se aqui apenas de *certezas práticas*, resultantes de um enorme número de provas, em que a frequência relativa de tais acontecimentos tem sido sempre 1. Aliás, tais factos só são certos *em determinadas condições:* por exemplo, o Sol *não nasce amanhã* em certos pontos da Terra, um copo *não cai* se for largado no interior de uma nave espacial a grande altitude, etc.

Em contraste com as *certezas práticas das ciências experimentais*, encontramos *certezas absolutas na matemática*, tais como:

> '*O quadrado dum número ímpar é sempre um número ímpar*'
> '*Não existe nenhum número racional cujo quadrado seja* 3'
> '*Uma equação algébrica de grau* n *não pode ter mais de* n *raízes num corpo*'

e muitas outras mais.

Dum modo geral, sempre que a frequência relativa dum acontecimento é 1 numa sequência de provas *muito numerosa*, é-se levado a admitir que o acontecimento é *praticamente certo*, isto é, que se realizará em qualquer outra prova futura, efectuada em condições idênticas. Nisto consiste, essencialmente, a *indução* ou *raciocínio indutivo*, que se encontra na base de toda a ciência experimental. Porém, à luz da lógica dedutiva, não se trata propriamente dum raciocínio (como os das demonstrações matemáticas), mas antes de um *paralogismo*, visto que se está a concluir ilicitamente do particular para o geral. Na verdade, as leis das ciências experimentais, em particular as *leis físicas*, têm todas carácter *contingente*: não se pode garantir que sejam infalíveis (ver Cap. I, n.º 17).

Vejamos mais alguns exemplos:

Suponhamos que, na Lotaria da Santa Casa da Misericórdia de Lisboa, nunca saiu a sorte grande no número 21212 (não vamos averiguar se isto é verdade ou não; interessa-nos apenas a hipótese, *que não é inverosímil*). Podemos nós daqui concluir que o acontecimento

α = *não sair a sorte grande no n.º 21212*

é certo, ou (o que é equivalente) que o acontecimento

$\tilde{\alpha}$ = *sair a sorte grande no n.º 21212*

é impossível?

Claro que não. Se o sistema de lotaria é correcto (e não deve haver dúvidas a esse respeito), há *tanta razão* para sair a sorte grande nesse número, como em qualquer dos outros em que tem saído. Por outras palavras: se o sistema é correcto, todos os números têm a *mesma probabilidade* de sair.

Portanto, o acontecimento $\tilde{\alpha}$, em rigor, é possível, embora possamos dizer que é *praticamente impossível* numa extracção isolada; no mesmo sentido em que podemos dizer:

'É praticamente impossível que, no próximo ano, o Tamisa suba até ao ponto de inundar a Abadia de Westminster'.

Recordemos, a propósito, a máxima popular:

A sorte grande sai sempre aos outros.

O significado disto é que *a sorte grande sai sempre num número diferente do nosso:* trata-se afinal de uma lei baseada na indução, exactamente como sucede com as leis da física. No entanto, há pessoas felizes que têm razão para não acreditar nesta lei...

Note-se que o grau de certeza prática pode, por vezes, equiparar-se ao da certeza matemática. Tal é, por exemplo, o que se observa com a proposição que tem servido de premissa a exemplos clássicos de silogismo:

'Todos os homens são mortais'

Trata-se aqui de uma *lei biológica qualitativa.* Mas se tentarmos precisá-la quantitativamente, afirmando por exemplo:

'Todos os homens morrem antes dos 1000 anos de idade'

já não sentiremos o mesmo grau de segurança. Conhecemos nós suficientemente o passado da espécie humana? E que sabemos nós sobre o seu futuro, quanto às possibilidades que vêm abrir os progressos da ciência, por exemplo as viagens espaciais?

O que pode dizer-se é que, nas condições actuais, é *extremamente improvável, praticamente impossível,* que um ser humano atinja a idade de 1000 anos.

Um outro exemplo análogo é o que se refere a alturas de pessoas: é praticamente impossível que um ser humano cresça até atingir a altura de 5 metros.

Porém, se formos baixando estes limites sucessivamente para 200 anos, 150 anos, etc. (no primeiro caso) ou para 3m, 2,50 m, etc. (no segundo caso), o grau de incerteza irá aumentando — e entraremos abertamente no campo das probabilidades (ver no fim do volume a *Tabela de Mortalidade*). É de salientar que o Cálculo das Probabilidades e a Estatística Matemática se têm desenvolvido principalmente no sector das ciências biológicas e das ciências sociais. Mas certo é também que, por um movimento de retrocesso, acabaram por invadir o campo das ciências físicas, principalmente no que se refere ao estudo do átomo. *A física deixou de ser determinista para se tornar probabilista.*

11. **Conceito quantitativo de probabilidade.** As considerações precedentes mostram como os atributos 'verdadeiro' e 'falso' aplicados a proposições se tornam insuficientes ([1]) quando, do rigor abstracto das teorias matemáticas, se passa à imprecisão inevitável dos *conhecimentos empíricos,* relativos ao mundo em que vivemos, e que são, por isso mesmo, *indispensáveis na vida prática.* Os conceitos de 'verdade' e 'falsidade' cedem, então, lugar ao conceito de 'probabilidade', complementar do de 'incerteza' (ou 'contingência'); um facto dir-se-á tanto mais *provável* quanto menos *incerto* (ou *contingente*) for.

O conceito de probabilidade, como todas as noções primitivas de fonte empírica, não é susceptível de definição lógica: gera-se no nosso espírito por um processo indutivo. Mais até: é inseparável do próprio raciocínio indutivo, como veremos.

Em muitos casos, a *probabilidade dum acontecimento* é algo que se pode *medir,* algo que se pode *exprimir por um número,* como se fosse uma grandeza mensurável – comprimento, velocidade, energia eléctrica ou qualquer outra. São esses, é claro, os casos que interessam no Cálculo das Probabilidades.

Vejamos um exemplo. Suponhamos que os alunos de uma turma foram encarregados da seguinte experiência: cada aluno lança 100 vezes uma moeda de um escudo ao ar e verifica quantas vezes sai *escudo* (em vez de *cara*). Suponhamos que todos acharam um número compreendido entre 35 e 65. Assim, se designarmos por α o acontecimento *sair escudo,* temos, em todas as referidas sequências:

(1) $$0{,}35 < \mathrm{fr}(\alpha) < 0{,}65$$

Suponhamos que se chega à mesma conclusão em várias outras experiências efectuadas em idênticas condições. Então, pelo *raciocínio*

([1]) Releia o n.º 9 do Cap. I (págs. 20-22, 1.º tomo).

indutivo, somos levados a prever que, em qualquer sequência futura de 100 lançamentos, a frequência relativa de α verifica a condição (1). Pois bem, exprimiremos este facto dizendo que a *probabilidade de* α é um número compreendido entre 0,35 e 0,65, ou ainda que a *probabilidade de* α *é aproximadamente* 0,5 *com erro inferior a* 0,15 (ou 50 % *com erro inferior a* 15 %).

Suponhamos agora que, em vez de 100 lançamentos se fazem *sequências de* 1000 *lançamentos* e que se obtém *sempre* o resultado

$$0{,}45 < \mathrm{fr}(\alpha) < 0{,}55$$

Então, com significado análogo ao anterior, diremos que a *probabilidade* de α é um número compreendido entre 0,45 e 0,55, e escreveremos:

$$0{,}45 < \mathrm{P}(\alpha) < 0{,}55$$

em que 'P(α)' é uma abreviatura de 'probabilidade de α'. Também diremos, neste caso, que a *probabilidade de* α *é aproximadamente* 0,5 *com erro inferior a* 0,05 (ou 50 % *a menos de* 5 %).

Passando a sequências de 10 000 lançamentos, é de esperar que se obtenha o resultado

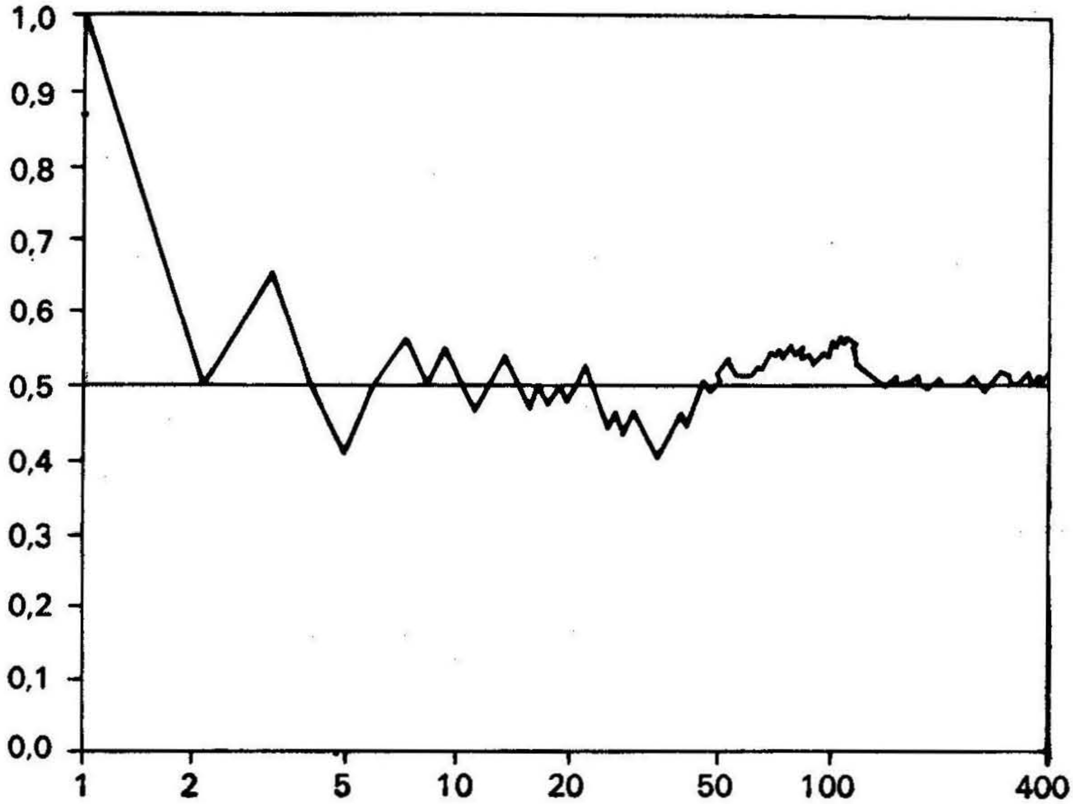
$$0{,}485 < \mathrm{fr}(\alpha) < 0{,}515$$

Então, escreveremos:

$$0{,}485 < \mathrm{P}(\alpha) < 0{,}515$$

e diremos que a *probabilidade de* α *é aproximadamente* 0,5 *a menos de* 0,015 (ou 50 % *a menos de* 1,5 %).

Assim, à medida que o número de lançamentos aumenta, a frequência relativa parece aproximar-se cada vez mais do número 0,5 que assumiremos como *valor aproximado* da probabilidade de α. Mas não tem sentido falar de *valor exacto* da probabilidade de α, do mesmo modo que não tem sentido falar de *medida exacta* duma grandeza física. Por exemplo, quando se diz que o comprimento de uma mesa é 1,65 m está-se a indicar um valor aproximado do comprimento da mesa, a menos de 1 cm; pode-se levar a aproximação até ao meio centímetro, até ao milímetro, etc., mas, abaixo de certo limite, já *não tem interesse* ou mesmo *não tem sentido* o grau de aproximação considerado.



Frequência relativa do acontecimento 'escudo' ao longo duma sequência de 400 lançamentos duma moeda ao ar. As frequências vão indicadas em escala logarítmica (isto é, as abcissas são logaritmos das frequências). (Exemplo dado por CRAMER em *Mathematical Methods of Statistics*.)

Em vez do lançamento de uma moeda ao ar, podíamos considerar outro tipo $\mathcal{P}$ de provas, por exemplo:

1) Lançar uma *punaise* (ou *attache*) ao ar e tomar nota da fre-

quência relativa do acontecimento *'cair de bico'*, que é contrário do acontecimento *'cair de cabeça'*. Ambas estas eventualidades vêm indicadas na figura seguinte:

2) Lançar um dado *ao acaso,* como é costume fazer-se, e tomar nota da frequência relativa de acontecimentos tais como 'sair o número 6', 'sair número par', 'sair número primo', etc.

Qualquer destas provas pode ser efectuada um grande número de vezes por uma equipa de alunos, utilizando *punaises* ou dados sensivelmente *iguais* em forma, dimensões e substância.

Em qualquer dos casos se deve verificar a chamada REGULARIDADE ESTATÍSTICA, isto é:

À medida que o número n *aumenta, a frequência relativa de cada acontecimento* α, *numa sequência de* n *provas, tende a ficar situada em intervalos*

$$[f_1, f'_1],\ [f_2, f'_2],\ \ldots,\ [f_n, f'_n],\ \ldots$$

cada vez mais pequenos, cada um deles contendo um seguinte. Nestas condições, os pontos médios

$$p_1 = \frac{f_1 + f'_1}{2}\ ,\ p_2 = \frac{f_2 + f'_2}{2}, \ldots,\ p_n = \frac{f_n + f'_n}{2}, \ldots$$

podem ser chamados 'valores aproximados da probabilidade do acontecimento α*' (com erro inferior a metade do comprimento do intervalo considerado).*

No exemplo da moeda, a probabilidade de sair escudo é aproximadamente 1/2, com erro bastante pequeno.

No exemplo da *attache,* a probabilidade do acontecimento

$$\alpha = \textit{cair de bico}$$

será *sensivelmente diferente* da probabilidade de

$$\tilde{\alpha} = \textit{cair de cabeça}$$

Essas probabilidades poderão ser determinadas com *suficiente aproximação,* por uma equipa de alunos, mas desde já é evidente que deverá ser:

$$P(\alpha) = 1 - P(\tilde{\alpha}) \text{ e, portanto, } P(\alpha) \neq \frac{1}{2} \qquad (\textit{Porquê?})$$

Quanto ao exemplo do dado, se designarmos por $\alpha_1, \alpha_2, \ldots, \alpha_6$ os acontecimentos *'sair o n.º 1', 'sair o n.º 2', ..., 'sair o n.º 6',* é de esperar que se tenha *aproximadamente*:

$$(1) \qquad P(\alpha_1) = P(\alpha_2) = \ldots = P(\alpha_6)$$

Mas, mesmo que o dado seja muito bem construído, aproximando-se bastante de um *cubo homogéneo,* nunca faz sentido afirmar *em absoluto* que as igualdades (1) são verdadeiras, do mesmo modo que não faz sentido afirmar *em absoluto* que duas réguas têm o mesmo comprimento.

Convenciona-se chamar *dado perfeito* a um dado que verifique a condição (1). Mas desde já vemos que se trata de uma noção abstracta: não existem *dados perfeitos,* do mesmo modo que não existem *gases perfeitos, água pura, pontos, rectas, cubos, esferas,* etc. Tais conceitos são apenas *esquemas,* isto é, modelos simplificados de entes concretos, idealizações que o nosso espírito elabora, tentando guiar-nos com maior ou menor êxito no mundo em que vivemos. Um desses esquemas é precisamente o conceito de probabilidade. *O que importa é saber aplicá-lo, de acordo com um conjunto* *adequado de regras (axiomas), que nos permita raciocinar logicamente sobre tal conceito.* (No final do capítulo trataremos do *conceito qualitativo* de probabilidade.)

12. **Axiomatização do conceito de probabilidade.** Das considerações precedentes podemos tirar, como súmula, a seguinte norma prática:

Dizer que a *probabilidade de um acontecimento* α, relativo a uma prova $\mathcal{P}$, é um determinado número p, significa que, dado um número positivo ε, tão pequeno quanto se queira, existe sempre um número n bastante grande tal que, numa sequência de n ou mais realizações de $\mathcal{P}$, é *praticamente certo* que a frequência relativa de α estará compreendida entre $p - \varepsilon$ e $p + \varepsilon$.

Não se trata aqui propriamente duma definição. É preciso notar que, no conceito de 'certeza prática' já está implícito o de 'probabilidade': um acontecimento diz-se *praticamente certo,* quando a sua probabilidade é *aproximadamente* 1, com erro desprezível ([1]). No entanto, a regra anterior elucida bastante sobre o uso prático do termo 'probabilidade'. Uma vez que as probabilidades são *frequências relativas previstas,* é natural atribuir-lhes as mesmas propriedades formais que se aplicam a frequências relativas (ver n.º 9). Somos, assim, levados a admitir, como *axiomas,* as seguintes propriedades, sendo α e β dois acontecimentos relativos a uma prova $\mathcal{P}$:

AXIOMA 1. *A probabilidade de* α *— que se designa abreviadamente por* $P(\alpha)$ *— é sempre um número real não negativo, isto é:* $P(\alpha) \geqslant 0$.

([1]) Analogamente, um acontecimento diz-se *praticamente impossível,* quando a sua probabilidade é aproximadamente 0, com erro desprezível. Por exemplo, é praticamente impossível que um macaco escreva à máquina a *Eneida* (exemplo do *macaco dactilógrafo,* de Emílio Borel).

AXIOMA 2. *A probabilidade do acontecimento certo é* 1, *isto é:* $P(\alpha) = 1$ *se* $\alpha = \mathbb{1}$.

AXIOMA 3. *Se* α *e* β *são acontecimentos incompatíveis, tem-se:*

$$P(\alpha + \beta) = P(\alpha) + P(\beta)$$

Destes axiomas (como premissas) deduzem-se, logicamente, vários *teoremas* (como conclusões), entre os quais os seguintes:

TEOREMA 1. *A probabilidade do acontecimento contrário de* α *é* $1 - P(\alpha)$, *isto é:*

$$P(\tilde{\alpha}) = 1 - P(\alpha)$$

Com efeito, como $\alpha\,\tilde{\alpha} = \mathbb{O}$ (*porquê?*) tem-se:

$$P(\alpha + \tilde{\alpha}) = P(\alpha) + (P\tilde{\alpha}) \qquad (\textit{Porquê?})$$

Por outro lado, como $\alpha + \tilde{\alpha} = \mathbb{1}$ (*porquê?*) tem-se:

$$P(\alpha) + P(\tilde{\alpha}) = 1 \quad (\textit{porquê?})$$

e, portanto, $P(\tilde{\alpha}) = 1 - P(\alpha)$.

COROLÁRIO 1. *Se* α *é acontecimento impossível, tem-se:*

$$P(\alpha) = 0 \quad (\textit{demonstre})$$

COROLÁRIO 2. *Qualquer que seja* α, *tem-se:*

$$P(\alpha) \leqslant 1 \quad (\textit{demonstre})$$

Portanto: $P(\alpha)$ é sempre um número real do intervalo $[0.1]$.

TEOREMA 2. *Se $\alpha_1, \alpha_2, \ldots, \alpha_n$ são* n *acontecimentos incompatíveis dois a dois (relativos à mesma prova $\mathcal{P}$) a probabilidade de que se realize um, pelo menos, dos acontecimentos $\alpha_1, \alpha_2, \ldots, \alpha_n$ é a soma das probabilidades destes acontecimentos, isto é:*

$$P\left(\sum_{i=1}^{n} \alpha_i\right) = \sum_{i=1}^{n} P(\alpha_i), \; \textit{se } \alpha_i \alpha_k = (0 \; \textit{para } i \neq k$$

Para demonstrar este teorema, basta aplicar, repetidamente, o axioma 2, observando que

$$\alpha_1 + \alpha_2 + \alpha_3 = (\alpha_1 + \alpha_2) + \alpha_3,$$

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = (\alpha_1 + \alpha_2 + \alpha_3) + \alpha_4, \text{ etc.}$$

e que, se $\alpha_1, \alpha_2, \ldots, \alpha_n$ são incompatíveis dois a dois, também α_3 é incompatível com $\alpha_1 + \alpha_2$ (*prove*), α_4 com $\alpha_1 + \alpha_2 + \alpha_3$, etc.

TEOREMA 3. *Quaisquer que sejam os acontecimentos α, β relativos à prova $\mathcal{P}$, tem-se:*

$$P(\alpha + \beta) = P(\alpha) + P(\beta) - P(\alpha\beta)$$

Com efeito, tem-se (*prove*):

(1) $$\alpha = \alpha\tilde{\beta} + \alpha\beta \quad , \quad \beta = \tilde{\alpha}\beta + \alpha\beta$$

(2) $$\alpha + \beta = \alpha\tilde{\beta} + \tilde{\alpha}\beta + \alpha\beta$$

De (1) vem:

(3) $$P(\alpha) = P(\alpha\tilde{\beta}) + P(\alpha\beta) \quad , \quad P(\beta) = P(\tilde{\alpha}\beta) + P(\alpha\beta) \qquad (\textit{Porquê?})$$

Por sua vez de (2) vem:

$$P(\alpha + \beta) = P(\alpha\tilde{\beta}) + P(\tilde{\alpha}\beta) + P(\alpha\beta) \qquad (\textit{Porquê?})$$

Donde, atendendo a (3):

$$P(\alpha+\beta) = P(\alpha) + P(\beta) - P(\alpha\,\beta)$$

Este teorema pode generalizar-se ao caso de *n* acontecimentos, obtendo-se a FÓRMULA DE DANIEL DA SILVA em termos de probabilidade.

13. **Exemplos de aplicação.** Vamos ver como as regras anteriores (axiomas e teoremas) podem ser aplicadas na prática. Os exemplos que vão seguir-se referem-se quase todos a *jogos de sorte,* também chamados *jogos de azar* (atribuindo aqui à palavra 'azar' o significado de 'acaso'). Na verdade, foram as reflexões de alguns matemáticos sobre jogos de azar que deram origem ao cálculo das probabilidades. Mais adiante trataremos de exemplos mais importantes.

EXEMPLO 1. Consideremos o caso do lançamento de um dado *ao acaso* e sejam α_1, α_2, ..., α_6 os acontecimentos *'sair o n.º* 1' *'sair o n.º* 2', ..., *'sair o número* 6'. Se o dado é *imperfeito,* isto é, se não se aproxima bastante de um cubo ou se é sensivelmente não homogéneo (por exemplo, mais denso nuns pontos que noutros), as probabilidades

$$P(\alpha_1),\ P(\alpha_2),\ ...,\ P(\alpha_6)$$

não são sensivelmente iguais, mas poderão ser determinadas *empiricamente* com maior ou menor aproximação. (*Como?*)

Seja porém como for, o acontecimento *'sair número primo menor que* 5' é, neste caso, $\alpha_2 + \alpha_3$ e, portanto, a sua probabilidade será:

$$P(\alpha_2+\alpha_3) = P(\alpha_2) + P(\alpha_3) \qquad (\textit{Porquê?})$$

Por sua vez, o acontecimento 'sair número par' é $\alpha_2+\alpha_4+\alpha_6$ e portanto, a sua probabilidade será:

$$P(\alpha_2 + \alpha_4 + \alpha_6) = P(\alpha_2) + P(\alpha_4) + P(\alpha_6) \quad \textit{(porquê?)}$$

e analogamente em outros casos.

Suponhamos, agora, que o dado é perfeito. Então será por defi nição:

$$P(\alpha_1) = P(\alpha_2) = P(\alpha_3) = P(\alpha_4) = P(\alpha_5) = P(\alpha_6)$$

isto é, os casos possíveis $\alpha_1,\alpha_2,\ldots,\alpha_6$ são todos *igualmente prováveis.* Designemos por p o valor comum destas probabilidades. Visto que os acontecimentos $\alpha_1,\alpha_2,\ldots,\alpha_6$ são incompatíveis dois a dois e a sua soma lógica é $\mathbb{1}$ *(porquê?)*, virá:

$$P(\alpha_1) + P(\alpha_2) + P(\alpha_3) + P(\alpha_4) + P(\alpha_5) + P(\alpha_6) = 6\ p = 1$$

e, portanto:

$$p = \frac{1}{6}$$

isto é, *a probabilidade de sair um determinado número, qualquer que ele seja, é* 1/6.

Posto isto, a probabilidade de '*sair número primo menor que* 5' será:

$$P(\alpha_2+\alpha_3) = P(\alpha_2) + P(\alpha_3) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3};$$

a probabilidade de *sair número par será:*

$$P(\alpha_2+\alpha_4+\alpha_6) = P(\alpha_2) + P(\alpha_4) + P(\alpha_6) = \frac{3}{6} = \frac{1}{2}, \quad \text{etc.}$$

Estes exemplos conduzem-nos à seguinte regra prática, que durante muito tempo foi tomada como definição da probabilidade[1]:

REGRA. *Quando os casos possíveis numa dada prova $\mathcal{P}$ são todos igualmente prováveis, a probabilidade de um acontecimento α relativo a $\mathcal{P}$ é igual ao quociente* ν/n *do número* ν *de casos favoráveis a* α *pelo número* n *de casos possíveis.*

Chama-se aqui 'casos possíveis' àqueles acontecimentos possíveis, de que todos os outros são somas lógicas, e 'casos favoráveis a um acontecimento α', precisamente, aos casos de que α é soma lógica. Pressupõe-se, além disso, que os *casos possíveis são em número finito.*

EXEMPLO 2. Imaginemos uma esfera de lotaria que contenha 20 bolas, numeradas de 1 a 20, das quais 8 são brancas, 7 vermelhas e 5 amarelas[2]. Consideremos a prova $\mathcal{P}$ que consiste em *extrair uma bola da esfera ao acaso* e designemos por $\alpha_1, \alpha_2, \ldots, \alpha_{20}$ os acontecimentos 'sair o n.º 1', 'sair o n.º 2', ..., 'sair o n.º 20'. Serão estes, portanto, os *casos possíveis da prova* $\mathcal{P}$. Serão eles *equiprováveis*? Se as bolas são sensivelmente iguais em forma, dimensões e substância, e se além disso a *casualização* é bem feita (isto é, se as bolas são bem agitadas na esfera e se a extracção é feita automaticamente, sem escolha humana deliberada), é de admitir que sim, isto é, mostra a experiência que

$$P(\alpha_1) = P(\alpha_2) = \ldots = P(\alpha_{20}) \text{ (com grande aproximação).}$$

---

(1) Tal definição era inaceitável por conter um *círculo vicioso:* não se pode definir um conceito, utilizando esse mesmo conceito.

(2) Em vez de uma esfera da lotaria, podemos também considerar uma caixa com pequena abertura ou um saco, como os que se usam no jogo do loto. Dum modo geral, chama-se *urna*, em cálculo das probabilidades, um recipiente destinado a conter bolas, ou outros objectos, para fazer sorteios.

Nestas condições, podemos aplicar a regra prática anterior ao cálculo de probabilidades de vários acontecimentos:

a) *Probabilidade de que saia o n.º 10.* Número de casos possíveis: 20. Número de casos favoráveis: 1. Probabilidade pedida: 1/20.

b) *Probabilidade de que saia bola branca.* Como agora o número de casos favoráveis é 8, a probabilidade pedida será:

$$\frac{8}{20} = 0{,}4 = 40\,\%$$

c) *Probabilidade de que saia bola vermelha.* Obtém-se de modo análogo: $\frac{7}{20} = 0{,}35 = 35\,\%$.

d) *Probabilidade de sair bola amarela:* $\frac{5}{20} = 0{,}25$.

e) *Probabilidade de sair bola amarela ou vermelha.* Este acontecimento é soma lógica dos acontecimentos 'sair bola amarela' e 'sair bola vermelha', que são *incompatíveis entre si.* Logo, a probabilidade pedida é a soma das duas anteriores:

$$\frac{7}{20} + \frac{5}{20} = \frac{12}{20} = 0{,}6 = 60\,\%$$

O mesmo resultado se podia obter, notando que o acontecimento considerado é o contrário de 'sair bola branca'. Assim, a probabilidade pedida é:

$$1 - \frac{8}{20} = \frac{12}{20} = 0{,}6$$

f) *Probabilidade de sair bola preta:* 0 (*Porquê?*)

g) *Probabilidade de sair número inferior a cem:* 1 (*Porquê?*)

Se designarmos por B, V e A, respectivamente, os acontecimentos *'sair bola branca'*, *'sair bola vermelha'* e *'sair bola amarela'*, relativos à prova $\mathcal{P}$, teremos em resumo:

$$P(B) = 0{,}4, \quad P(V) = 0{,}35, \quad P(A) = 0{,}25$$

$$P(A + V) = 0{,}6, \quad P(B + V) = 0{,}75, \quad P(A + B + V) = 1$$

Muitas vezes, usa-se em cálculo das probabilidades o mesmo símbolo para designar um acontecimento e o conjunto correspondente. Por exemplo, o símbolo B pode designar indistintamente o acontecimento *'sair bola branca'* ou o *conjunto das bolas brancas* (existentes na esfera).

EXEMPLO 3. Consideremos, agora, duas esferas de lotaria, E_1 e E_2. Suponhamos que E_1 contém 20 bolas, nas condições do exemplo anterior, e que E_2 contém 10 bolas, numeradas de 1 a 10, sendo 7 brancas e 3 vermelhas (sensivelmente iguais em forma, volume e substância). Seja agora $\mathcal{P}$ a prova que consiste em extrair *ao acaso* uma bola de cada uma das esferas. É claro que os casos possíveis podem ser esquematicamente indicados pelos elementos do *produto cartesiano*,

$$\{1,2,...,20\} \times \{1,2,...,10\}$$

Por exemplo, o par ordenado (14,5) indica o caso que consiste em sair a bola 14 da esfera E_1 e a bola 5 da esfera E_2. Ora, o número de tais pares ordenados é $20 \times 10 = 200$. *Será, pois,* 200 *o número de casos possíveis*. Serão estes casos equiprováveis? É de admitir que sim, uma vez que as bolas sejam sensivelmente iguais, como se disse, e que a casualização seja bem feita; em particular as esferas E_1 e E_2 devem ser *independentes*, isto é, *a bola que sair de uma das esferas não depender de modo algum da bola que sair da outra esfera.*

Aceites estas premissas, podemos efectuar os seguintes cálculos:

a) *Probabilidade de que saia bola branca das duas esferas.* Já sabemos que o número de casos possíveis é 20×10. O número de casos favoráveis será, manifestamente, o cardinal do produto cartesiano do conjunto das bolas brancas de E_1 (em número de 8) pelo conjunto das bolas brancas de E_2 (em número de 7). Será, pois, 8×7 o número de casos favoráveis e, assim, a probabilidade pedida é:

$$\frac{8 \times 7}{20 \times 10} = \frac{28}{100} = 0{,}28$$

b) *Probabilidade de que saia bola branca de* E_1 *e bola vermelha de* E_2. Obtém-se de modo análogo:

$$\frac{8 \times 3}{20 \times 10} = \frac{12}{100} = 0{,}20$$

c) *Probabilidade de sair bola amarela das duas esferas.*

d) *Probabilidade de sair de* E_1 *número duplo do que sair de* E_2. Resposta: 5 %.

e) *Probabilidade de não sair bola branca de nenhuma das esferas.*

f) *Probabilidade de sair bola branca de uma, pelo menos, das esferas.* (Sugestão: acontecimento contrário do anterior.)

g) *Probabilidade de sair bola branca de uma e uma só esfera.* (Sugestão: este acontecimento é soma lógica dos acontecimentos incompatíveis 'sair bola branca de E_1 e não sair bola branca de E_2', 'sair bola branca de E_2 e não sair bola branca de E_1'.)

EXEMPLO 4. Consideremos, novamente, uma esfera com a composição do exemplo 2 (20 bolas equiprováveis, sendo 8 brancas, 7 vermelhas e 5 amarelas). Designemos por $\mathcal{P}'$ a prova que consiste em duas extracções sucessivas, *repondo na esfera a bola que sai na*

1ª extracção; e por $\mathcal{P}''$ a prova que consiste em duas extracções sucessivas, *sem reposição*. Calculemos, então:

a) *Probabilidade de sair duas vezes bola branca na prova* $\mathcal{P}'$. Número de casos possíveis (equiprováveis): 20 × 20. Número de casos favoráveis: 8 × 8. Probabilidade pedida:

$$\frac{8 \times 8}{20 \times 20} = 0{,}4 \times 0{,}4 = 0{,}16$$

b) *Probabilidade de sair duas vezes bola branca na prova* $\mathcal{P}''$. Número de casos possíveis (equiprováveis): 20 × 19. Número de casos favoráveis: 8 × 7. Probabilidade pedida:

$$\frac{8 \times 7}{20 \times 19} = \frac{28}{190} \approx 0{,}19$$

c) *Probabilidade de sair primeiro bola branca e depois bola vermelha na prova* $\mathcal{P}'$

d) *Idem na prova* $\mathcal{P}''$

e) *Probabilidade de sair uma vez bola branca e outra vez bola vermelha, independentemente de ordem (prova* $\mathcal{P}'$ *e prova* $\mathcal{P}''$*)* (Sugestão soma lógica dos acontecimentos 'sair primeiro bola branca e depois bola evermelha sair primeiro bola vermelha e depois bola branca'.)

f) *Probabilidade de não sair vez nenhuma bola branca (prova* $\mathcal{P}'$ *e prova* $\mathcal{P}''$*)* (Sugestão acontecimento equivalente a 'sair duas vezes bola vermelha ou amarela')

g) *Probabilidade de sair alguma vez bola branca (prova* $\mathcal{P}'$ *e prova* $\mathcal{P}''$*)* (Acontecimento contrário do anterior.)

h) *Probabilidade de sairem duas bolas da mesma cor (prova* $\mathcal{P}'$ *e prova* $\mathcal{P}''$*)*.

i) *Probabilidade de sairem duas bolas de cor diferente (prova* $\mathcal{P}'$ *e prova* $\mathcal{P}''$*)*.

EXEMPLO 5. Tornemos ao caso de um dado perfeito. É agora fácil calcular as seguintes probabilidades:

a) *Probabilidade de que, em dois lances sucessivos, se obtenha uma vez um número par e outra vez um múltiplo de 3.*

b) *Probabilidade de que, em* 3 *lances sucessivos, não saiam os números 1 e 6.*

c) *Probabilidade de que, em 2 lances sucessivos, a soma dos números saídos seja menor que 5.*

EXEMPLO 6. Um cofre tem um segredo de 3 discos, com 24 letras cada um (ver pág. 163, 1.º tomo). *Calcular a probabilidade que um ladrão teria de descobrir o segredo, fazendo no máximo 100 tentativas.* Resposta: aproximadamente 0,007 ou seja 7 ‰.

EXEMPLO 7. *Calcular a probabilidade de, jogando uma vez no totobola com o equivalente a 9 600 apostas simples* [1]*, se acertar nos 13 resultados, supondo que as apostas são feitas inteiramente ao acaso.* Resposta: aproximadamente 0,006, probabilidade ainda extremamente pequena (compare com o resultado anterior).

14. **Probabilidade do produto lógico.** Da axiomática do conceito de probabilidade (n.º 12), nada se pode deduzir quanto à probabilidade do produto lógico. Mas, por analogia com o que se fez para o conceito de frequência relativa, é natural introduzir a seguinte definição:

DEFINIÇÃO 1. *Sendo α e β dois acontecimentos relativos a uma dada prova, chama-se probabilidade de α se β, e representa-se por* $P(\alpha|\beta)$, *o quociente de* $P(\alpha\beta)$ *por* $P(\beta)$, *isto é:*

(1) $$P(\alpha|\beta) = \frac{P(\alpha\beta)}{P(\beta)}$$

---

(1) O que importa em 14 400$00 — (Valor calculado ao preço de cada aposta, ao tempo — N. do E.).

Por exemplo, seja α o acontecimento 'haver acidente numa viagem de automóvel em 100 km de estrada' e β o acontecimento 'ter chovido'. Então, supondo que podemos atribuir probabilidades aos acontecimentos α, β, $\alpha\beta$, o símbolo $P(\alpha|\beta)$ designa a probabilidade de *haver um tal acidente, após ter chovido.*

DEFINIÇÃO 2. *Diz-se que dois acontecimentos* α, β *são independentes, sse* $P(\alpha|\beta) = P(\alpha)$. *Caso contrário,* α *e* β *dizem-se dependentes ou associados.*

Notemos, agora, que (1) se pode escrever:

$$(P\alpha\beta) = P(\beta) \cdot P(\alpha|\beta)$$

ou ainda, trocando os papéis de α e de β:

$$\boxed{P(\alpha\,\beta) = P(\alpha) \cdot P(\beta|\alpha)} \tag{2}$$

É evidente que esta fórmula apenas exprime a definição 1, sob uma forma diferente. Por outro lado, tem-se, atendendo à definição 2:

$$\boxed{P(\alpha\beta) = P(\alpha)\,P(\beta),\ \text{sse } \alpha \text{ e } \beta \text{ são independentes}} \tag{3}$$

NOTA IMPORTANTE. O conceito de 'acontecimentos independentes' desempenha, em relação ao produto lógico, um papel análogo ao de 'acontecimentos incompatíveis', em relação à soma lógica. *Não confunda estes dois conceitos!*

Tratando-se de três acontecimentos α, β, γ relativos a uma prova $\mathcal{P}$ é fácil reconhecer qual o significado de símbolos tais como $\mathcal{P}(\alpha|\beta\gamma)$, $P(\gamma|\alpha\,\beta)$, etc. e ver que (cf. n.º 6, pág. 216):

$$P(\alpha\,\beta\,\gamma) = P(\alpha) \cdot P(\beta|\alpha) \cdot P(\gamma|\alpha\,\beta) \tag{4}$$

Em particular, pode ter-se:

(5) $$P(\alpha\beta\gamma) = P(\alpha)\cdot P(\beta)\cdot P(\gamma)$$

DEFINIÇÃO 3. *Diz-se que três acontecimentos* α, β, γ *são independentes, sse verificam as condições* $P(\beta) = P(\beta|\alpha)$, $P(\gamma) = P(\gamma|\alpha\beta)$ *e todas as que se deduzem destas, substituindo um ou mais dos acontecimentos* α, β, γ *pelos respectivos contrários.*

Em termos intuitivos poderíamos dizer:

'Três acontecimentos são *independentes,* sse a probabilidade de cada um deles é a mesma, quer se verifique ou não um só ou os dois outros acontecimentos'.

É evidente que a anterior definição equivale à seguinte proposição:

Três acontecimentos α, β, γ, *são independentes sse, além da condição* (5), *verificam todas as que desta se deduzem, substituindo um ou mais dos acontecimentos* α, β, γ, *pelos seus contrários.*

Isto mostra que a relação ternária definida é simétrica.

É porém de notar que, tal como no caso das frequências relativas:

1) Três acontecimentos α, β, γ podem ser independentes entre si dois a dois sem serem *os três* independentes (segundo a definição anterior).

2) Pode verificar-se a fórmula (5), sem que os acontecimentos α, β, γ sejam independentes.

15. **Probabilidade do produto cartesiano. Sistemas de lotaria.** Para melhor compreensão do que vai seguir-se, convém recordar algumas situações que surgem, em geometria aplicada, a propósito de produtos cartesianos. Viu-se que, por exemplo, no

universo $\mathbb{R}$, a condição $x \geqslant 2$, com *uma só* variável, pode ser identificada à condição *com duas variáveis*

$$x \geqslant 2 \wedge y \in \mathbb{R}$$

que é verificada por todos os pares ordenados (x, y) de números reais tais que: $x \geqslant 2$ e y é *qualquer*. Deste modo, a condição $x \geqslant 2$, que na recta representa geometricamente uma *semi-recta*, no plano passa a representar um *semiplano*. Analogamente, a condição $y \geqslant 1$, que na recta (ou mais precisamente no eixo dos y) representa uma semi-recta, pode ser identificada com a condição

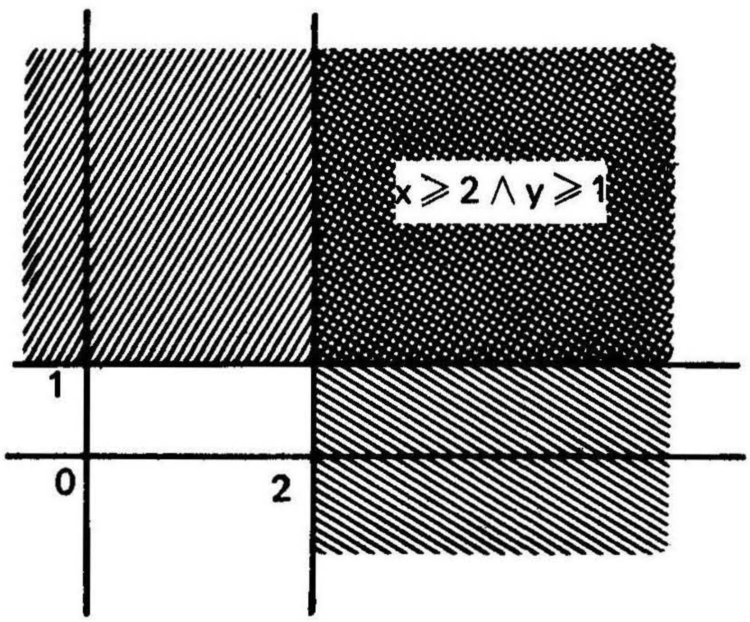
$$x \in \mathbb{R} \wedge y \geqslant 1,$$

e assim, no plano, passa a representar um semiplano.

Por sua vez, a condição

$$x \geqslant 2 \wedge y \geqslant 1$$

representa no plano um ângulo recto, que é a *intersecção* (ou *produto lógico*) dos semiplanos $x \geqslant 2$ e $y \geqslant 1$. Este produto lógico resulta



assim identificado ao *produto cartesiano* das semi-rectas representadas pelas condições $x \geqslant 2$ e $y \geqslant 1$, respectivamente no eixo dos x e no eixo dos y.

Consideremos, agora, duas esferas de lotaria, E_1 e E_2, nas condições indicadas no exemplo 3 do n.º 13. Designemos por $\mathcal{P}_1$ a prova que consiste em extrair ao acaso uma bola de E_1, por $\mathcal{P}_2$ a prova análoga para E_2 e por $\mathcal{P}'$ a prova que consiste em efectuar $\mathcal{P}_1$ e $\mathcal{P}_2$. Diremos, então, que $\mathcal{P}'$ é o produto cartesiano de $\mathcal{P}_1$ e $\mathcal{P}_2$ e escreveremos:

$$\mathcal{P}' = \mathcal{P}_1 \times \mathcal{P}_2$$

Sejam B_1, V_1 e A_1, respectivamente, os acontecimentos *sair bola branca, sair bola vermelha* e *sair bola amarela,* na prova $\mathcal{P}_1$. Sejam ainda B_2 e V_2 os acontecimentos sair *bola branca* e *sair bola vermelha,* na prova $\mathcal{P}_2$. Então, designaremos por $B_1 \times V_2$ (produto cartesiano de B_1 por V_2) o acontecimento sair bola branca de E_1 e bola vermelha de E_2. Análogos significados terão as expressões:

$$V_1 \times B_2,\ A_1 \times V_2,\ B_1 \times B_2,\ V_1 \times V_2,\ A_1 \times B_2$$

Notemos, agora, que o acontecimento B_1 relativo a $\mathcal{P}_1$ se pode identificar a um acontecimento B_1^* relativo a $\mathcal{P}_1 \times \mathcal{P}_2$: o acontecimento

Sair bola branca de E_1 *e bola qualquer de* E_2

De modo análogo, V_1, A_1, B_2, V_2 podem ser identificados a acontecimentos, respectivamente, V_1^*, A_1^*, B_1^* e V_2^*, relativos a $\mathcal{P}_1 \times \mathcal{P}_2$. Assim, podemos reduzir os *produtos cartesianos* a *produtos lógicos:*

$$B_1 \times V_2^* = B_1^* V_2^* \quad , \quad V_1 \times B_2 = V_1^* B_2^* \quad , \quad \text{etc.}$$

Posto isto, recordemos que, segundo as condições indicadas no exemplo 3 do n.º 13, os acontecimentos B_1 e V_2^* são independentes (o mesmo podendo dizer-se agora dos acontecimentos B_1 e V_2).

Será assim:

$$P(B_1 \times V_2) = P(B_1^*) \cdot P(V_2^*) = P(B_1) \cdot P(B_2)$$

E, como $P(B_1) = 4/10$, $P(V_2) = 3/10$, virá:

$$P(B_1) \cdot P(V_2) = \frac{4}{10} \cdot \frac{3}{10} = \frac{12}{100}$$

tal como tínhamos achado. Analogamente:

$$P(B_1 \times B_2) = P(B_1) \cdot P(B_2) = 0{,}4 \times 0{,}7 = 0{,}28$$

$$P(V_1 \times B_2) = 0{,}7 \times 0{,}7 = 0{,}49, \text{ etc.}$$

Dum modo geral, se α e β são acontecimentos relativos a duas provas distintas $\mathcal{P}_1$ e $\mathcal{P}_2$, representa-se por $\mathcal{P}_1 \times \mathcal{P}_2$ a prova que consiste em efectuar $\mathcal{P}_1$ e $\mathcal{P}_2$, e por $\alpha \times \beta$ (*produto cartesiano de* α *por* β) o acontecimento que consiste em *realizar-se* α *em* $\mathcal{P}_1$ *e* β *em* $\mathcal{P}_2$. Os acontecimentos α e β podem ser identificados a acontecimentos, respectivamente α^* e β^*, relativos a $\mathcal{P}_1 \times \mathcal{P}_2$, tal como foi indicado no exemplo anterior, permitindo reduzir o produto cartesiano a produto lógico

$$\alpha \times \beta = \alpha^* \beta^*$$

Os acontecimentos α e β dizem-se *independentes,* sse α^* e β^* o forem. Assim, virá:

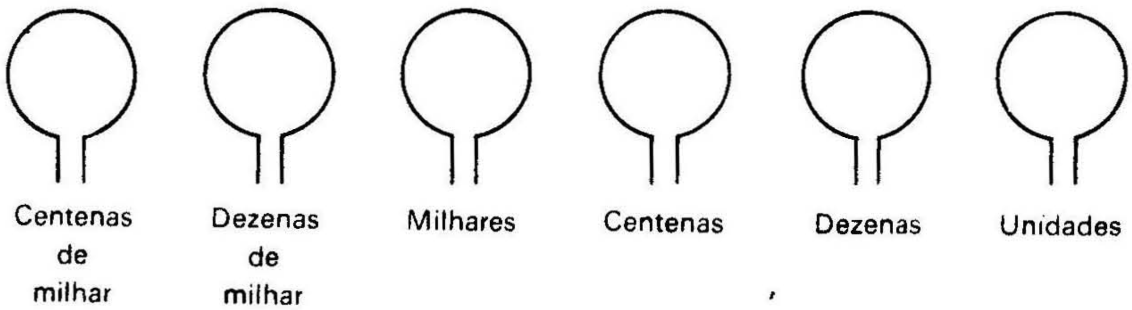
$$(1) \qquad \boxed{P(\alpha \times \beta) = P(\alpha) \cdot P(\beta), \text{ sse } \alpha \text{ e } \beta \text{ são independentes}}$$

As provas $\mathcal{P}_1$ e $\mathcal{P}_2$ dizem-se *independentes,* sse todo o acontecimento relativo a $\mathcal{P}_1$ é independente de todo o acontecimento relativo a $\mathcal{P}_2$.

A regra (1) permite resolver mais simplesmente vários dos problemas que foram resolvidos no n.º 13 por contagem de casos possíveis e casos favoráveis (veja quais são esses problemas e resolva-os pelo novo processo).

Estas considerações podem estender-se de maneira óbvia ao caso de três provas, quatro provas, etc., tendo em conta o que foi estabelecido no número anterior.

Por exemplo, um dos modernos sistemas de lotaria mais em uso consiste em utilizar, por exemplo, 6 esferas em que se introduzem bolas cuja numeração pode ir de 0 a 9; uma das esferas é destinada a dar o algarismo das unidades, outra a dar o número das centenas, etc.



Em princípio, as extracções das diferentes esferas devem ser provas independentes, isto é, a probabilidade de saída de uma dada sequência de algarismos deve ser exactamente igual ao produto das probabilidades de saída de um algarismo de cada esfera. Além disso, os algarismos de cada esfera devem ter todos a mesma probabilidade de saída. Assim, por exemplo, num jogo com 300 000 bilhetes, a probabilidade de ser premiado um determinado número na primeira extracção deverá ser *exactamente* igual a

$$\frac{1}{3}\cdot\frac{1}{10}\cdot\frac{1}{10}\cdot\frac{1}{10}\cdot\frac{1}{10}\cdot\frac{1}{10}=\frac{1}{300\,000}$$

Já sabemos que não é possível, nem sequer faz sentido, falar de *exactidão absoluta*. O que se requer na prática é que seja impossível detectar estatisticamente alguns números com maior probabilidade de saída, visto que esse facto poderia ser explorado ilicitamente por uma casa de jogo.

Assim, quando uma casa de jogo anuncia comercialmente que é contemplada com maior número de prémios, isso apenas deve significar que a referida casa vende um maior número de bilhetes, o que, evidentemente, aumenta a probabilidade de prémio. Qualquer outra interpretação equivaleria a admitir que essa casa está a usar ilicitamente o conhecimento de algum vício do sistema.

16. **Problema das provas repetidas; distribuição binomial.** Consideremos, por exemplo, o seguinte problema:

Determinar a probabilidade de que, em três lançamentos sucessivos dum dado perfeito, saia duas vezes (e só duas) o número 6.

Os três lançamentos sucessivos são três provas $\mathcal{P}_1$, $\mathcal{P}_2$, $\mathcal{P}_3$, que constituem realizações distintas de uma mesma prova-tipo $\mathcal{P}$ (lançamento do dado) A sequência dessas três provas constitui por sua vez, a prova

$$\mathcal{P}_1 \times \mathcal{P}_2 \times \mathcal{P}_3$$

Designemos por s_1, s_2, s_3, respectivamente, os acontecimentos *'sair o n.º 6 em* $\mathcal{P}_1$*', 'sair o n.º 6 em* $\mathcal{P}_2$*', 'sair o n.º 5 em* $\mathcal{P}_3$*'* Então o acontecimento *'sair o n.º 6 duas vezes (e duas só) em* $\mathcal{P}_1 \times \mathcal{P}_2 \times \mathcal{P}_3$*'* pode realizar-se de três maneiras distintas e incompatíveis:

(1) $$s_1 \times s_2 \times \tilde{s}_3, \quad s_1 \times \tilde{s}_2 \times s_3, \quad \tilde{s}_1 \times s_2 \times s_3$$

Ora, $P(s_1) = P(s_2) = P(s_3) = 1/6$ e, portanto,

$$P(\tilde{s}_1) = P(\tilde{s}_2) = P(\tilde{s}_3) = 1 - 1/6.$$

Por outro lado, é fácil reconhecer que os três acontecimentos s_1, s_2, s_3 são independentes (discutiremos este ponto mais adiante, no caso geral). Logo, os acontecimentos (1) têm todos probabilidade igual a

$$P(s_1 \times s_2 \times \tilde{s}_3) = P(s_1)\,P(s_2)\,P(\tilde{s}_3) = \left(\frac{1}{6}\right)^2 \frac{5}{6} = \frac{5}{216}$$

Mas o acontecimento cuja probabilidade se pede é a soma lógica dos referidos acontecimentos incompatíveis:

$$s_1 \times s_2 \times \tilde{s}_3 + s_1 \times \tilde{s}_3 \times s_3 + \tilde{s}_1 \times s_2 \times s_3$$

A probabilidade pedida será, pois, a soma de 3 números iguais a 5/216, ou seja:

$$3 \cdot \frac{5}{216} = \frac{5}{72} \approx 7\,\%$$

Consideremos, ainda, o problema análogo:

De uma urna que contém 10 bolas, das quais 3 são brancas e 7 pretas, tiram-se 4 bolas à sorte, sucessivamente, com reposição. Determinar a probabilidade de que seja 2 o número de bolas brancas saídas nas 4 extracções.

Designando por B_1, B_2, B_3, B_4 o acontecimento *saída de bola branca* nas sucessivas extracções, o acontecimento cuja probabilidade se pede será a soma lógica das seguintes hipóteses incompatíveis duas a duas:

$$B_1 \times B_2 \times \tilde{B}_3 \times \tilde{B}_4,\quad B_1 \times \tilde{B}_2 \times B_3 \times \tilde{B}_4,\quad B_1 \times \tilde{B}_2 \times \tilde{B}_3 \times B_4$$

$$\tilde{B}_1 \times B_2 \times B_3 \times \tilde{B}_4,\quad \tilde{B}_1 \times B_2 \times \tilde{B}_3 \times B_4,\quad \tilde{B}_1 \times \tilde{B}_2 \times B_3 \times B_4$$

Como interpretar o número destas hipóteses em cálculo combinatório? Pense por si uns momentos, antes de ver a resposta que vem a seguir.

É fácil reconhecer que estas hipóteses estão em correspondência biunívoca com as *combinações*

$$B_1B_2,\ B_1B_3,\ B_1B_4,\ B_2B_3,\ B_2B_4,\ B_3B_4,$$

dos acontecimentos B_1, B_2, B_3, B_4 tomados 2 a 2. O seu número é, pois:

$$\binom{4}{2} = \frac{4 \times 3}{1 \times 2} = 6$$

Por outro lado, os acontecimentos B_1, B_2, B_3, B_4 têm todos a probabilidade 3/10 e, portanto, os seus contrários têm a probabilidade 7/10. E, como se trata de acontecimentos independentes ([1]), segue-se que todas as referidas hipóteses têm a mesma probabilidade:

$$P(B_1 \times B_2 \times \tilde{B}_3 \times \tilde{B}_4) = \left(\frac{3}{01}\right)^2 \left(\frac{7}{10}\right)^2$$

Por conseguinte, a probabilidade pedida será a soma de 6 parcelas iguais a este número, ou seja:

$$\binom{4}{2} \cdot \left(\frac{3}{10}\right)^2 \left(\frac{7}{10}\right)^2 = 6 \times 21^2 \times 10^{-4} = 0{,}2646$$

Estes dois exemplos ajudam-nos a resolver o seguinte problema geral, muito importante:

PROBLEMA DAS PROVAS REPETIDAS. *Sendo* α *um acontecimento de probabilidade* p, *relativo a uma prova* $\mathcal{D}$, *determinar a probabilidade de que, em* n *realizações de* $\mathcal{D}$, *seja* x *o número de vezes que* α *se verifica.*

Como se vê, o que se pede afinal é a *probabilidade de que em* n *realizações de* $\mathcal{D}$, *a preferência absoluta de* α *seja um dado número inteiro* x $(0 \leqslant x \leqslant n)$.

([1]) Este ponto será esclarecido a seguir, no caso geral.

Designemos por $\mathcal{P}_1, \mathcal{P}_2, \ \ , \mathcal{P}_n$ as n realizações de $\mathcal{P}$ e, dum modo geral, por α_k o acontecimento que consiste em *verificar-se* α *na prova* $\mathcal{P}_k$ (k=1, 2, ..., n) Então, o acontecimento

$$x\alpha = \textit{verificação de } \alpha \textit{ em } x \textit{ das } n \textit{ provas consideradas}$$

é a soma lógica de todos os acontecimentos que resultam de

$$\alpha_1 \times \alpha_2 \times \cdots \times \alpha_n,$$

deixando ficar x factores como estão e substituindo os n-x restantes pelos acontecimentos contrários. Um deles será o acontecimento

$$\alpha_1 \times a_2 \times \cdots \times \alpha_x \times \tilde{\alpha}_{x+1} \times \cdots \times \tilde{\alpha}_n,$$

que consiste em verificar-se α precisamente nas primeiras x provas da sequência.

Qual é o número de hipóteses assim obtidas? Pensando um momento, vê-se que é igual ao número de combinações dos acontecimentos $\alpha_1, \alpha_2, \ldots, \alpha_n$ tomados x a x, ou seja $\binom{n}{x}$. Ora

$$P(\alpha_k) = p, \quad P(\tilde{\alpha}_k) = 1 - p, \quad \text{para } k=1, 2, \ldots, n$$

Além disso, os acontecimentos $\alpha_1, \alpha_2, \ldots, \alpha_n$ são *independentes*. Com efeito, do próprio conceito de probabilidade (n.os 11 e 12) decorre naturalmente que a probabilidade de α se verificar em cada uma das provas $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_n$ (n realizações sucessivas de $\mathcal{P}$, em condições idênticas) é sempre a mesma, qualquer que seja o resultado das restantes provas (cf. n.os 14 e 15). Logo, as hipóteses consideradas têm todas a probabilidade

$$P(\alpha_1 \times \cdots \times \alpha_x \times \tilde{\alpha}_{x+1} \times \cdots \times \tilde{\alpha}_n) = p^x(1-p)^{n-x}$$

Ora, estas hipóteses são incompatíveis entre si duas a duas e o seu número é, como vimos $\binom{n}{x}$. Logo, a probabilidade pedida será:

$$P(x\alpha) = \binom{n}{x} p^x(1-p)^{n-x}$$

Assim o problema está resolvido.

Costuma simplificar-se a fórmula anterior escrevendo P(x) em vez de P(xα) e pondo $1-p=q$:

$$(1) \qquad \boxed{P(x) = \binom{n}{x} p^x q^{n-x}}$$

Como se vê, x é uma variável casual (ver n.º 8), cujos valores possíveis são 0,1,...,n, e, a cada um destes valores, a fórmula (1) faz corresponder uma determinada probabilidade P(x). Exprime-se este facto dizendo que a fórmula (1) define *a distribuição de probabilidade da variável casual* x *considerada* (frequência absoluta de α em *n* realizações de P). Como, além disso, o segundo membro de (1) é precisamente o termo em p^x do desenvolvimento de $(p+q)^x$ segundo a *fórmula do binómio,* diz-se que a distribuição de probabilidade definida por (1) é a DISTRIBUIÇÃO BINOMIAL (também chamada *'distribuição de BERNOULLI'*).

Apresentam-se, na prática, muitas outras distribuições de probabilidade [1]. Por exemplo, o número que sai no lançamento dum dado é uma variável casual x, cujos valores possíveis são 1, 2, 3, 4, 5, 6 e tal que

$$P(x) = \frac{1}{6}, \text{ sse o dado é perfeito.}$$

Se o dado é imperfeito a distribuição será outra.

---

[1] A par do conceito de 'distribuição de probabilidade' apresenta-se naturalmente o conceito de *'distribuição de frequência'*.

NOTA. A fórmula (1) pode ser demonstrada de maneira mais breve e mais rigorosa pelo *método de indução matemática*, que estudaremos mais tarde.

17. Aplicações de distribuição binomial; exemplo da genética. Diz-se que uma moeda é *equilibrada*, sse a probabilidade de aparecer escudo num lançamento da moeda ao ar é 1/2. Determinemos a probabilidade de que, em 10 lançamentos de uma moeda equilibrada, apareça x vezes escudo. Neste caso, basta aplicar a distribuição binomial com $p = 1/2$ e $q = 1 - 1/2 = 1/2$, o que dá:

$$P(x) = \binom{10}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{10-x}$$

Por exemplo, se $x = 3$, vem:

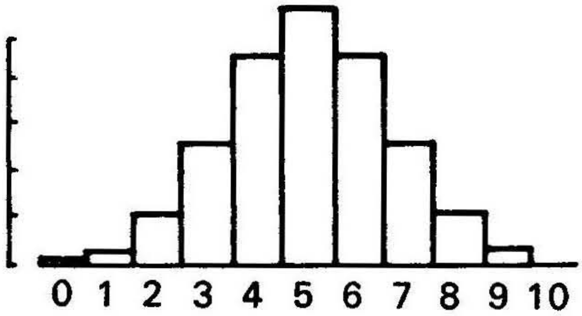
$$P(3) = \binom{10}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^7 = \frac{10\cdot 9\cdot 8}{3\cdot 2\cdot 1}\cdot\left(\frac{1}{2}\right)^{10} = \frac{15}{128} \approx 0{,}117$$

Os valores de P(x) para x = 0, 1,...,10, são dados pela tabela seguinte, com erro inferior a 0,001

Tabela

x	P(x)	x	P(x)
0	0,001	6	0,205
1	0,010	7	0,177
2	0,044	8	0,044
3	0,177	9	0,010
4	0,205	10	0,001
5	0,246		

Para ter uma imagem desta distribuição convém recorrer a um *gráfico de colunas* (ou *histograma*), que consiste neste caso em construir, para cada um dos valores de x considerados, um rectângulo de altura P(x) e de base 1, assente no eixo das abcissas, tendo esse valor de x por ponto médio. Convém, neste caso, adoptar uma unidade de comprimento bastante maior para as ordenadas do que para as abcissas.



Note-se que este histograma é simétrico (a distribuição *é simétrica*) e que P(x) atinge um valor máximo (= 0,246) para x = 5. É, portanto, 5 o *valor mais provável* da variável casual considerada, o que aliás era de esperar, visto que, sendo 1/2 a probabilidade de aparecer escudo, a frequência relativa mais provável deverá ser 50 % e, como o número de lançamentos é 10, a frequência absoluta mais provável é 0,5 × 10 = 5. Mas, note-se o seguinte facto curioso, aparentemente paradoxal:

> *É mais provável não se verificar o valor mais provável do que verificar-se este valor.*

Com efeito, a probabilidade do valor mais provável é 0,246, enquanto a do acontecimento contrário é 1 − 0,246 = 0,754. A explicação do *aparente paradoxo* é simples: o acontecimento contrário é a soma lógica dos acontecimentos que correspondem a *todos* os outros valores possíveis; embora menos prováveis, a soma das probabilidades destes é 0,754 ≈ 75 %.

Chama-se *moda* de uma distribuição de probabilidade de uma variável x (com um número finito de valores) todo o valor de x com probabilidade máxima. A distribuição anterior tem uma única moda: o valor 5 (distribuição *unimodal*). Mas, se em vez de $n = 10$ tivéssemos tomado $n = 11$ (com $p = 1/2$), já encontraríamos duas modas: os valores 5 e 6 (distribuição *bimodal*). Há ainda distribuições *trimodais*, etc. A distribuição binomial só pode ser unimodal ou bimodal.

Tornando ao caso anterior ($p = 1/2$, $n = 10$), calculemos a probabilidade de que x satisfaça, por exemplo, à condição $3 \leqslant x \leqslant 7$, equivalente à soma lógica das condições $x = 3$, $x = 4$, $x = 5$, $x = 6$, $x = 7$. Será, pois:

$$P(3 \leqslant x \leqslant 7) = \sum_{k=3}^{7} \binom{10}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^k$$

$$\approx 2 \times 0{,}117 + 2 \times 0{,}205 + 0{,}246 \approx 0{,}89$$

Podemos também dizer que *é 89 % a probabilidade duma frequência relativa* f *do acontecimento 'aparecer escudo' tal que* $0{,}3 \leqslant f \leqslant 0{,}7$, *numa sequência de* 10 *provas.*

Passando a uma sequência de 100 provas, as frequências absolutas correspondentes às frequências relativas 0,3 e 0,7 são $0{,}3 \times 100 = 30$ e $0{,}7 \times 100 = 70$. Será, então:

$$P(30 \leqslant x \leqslant 70) = \sum_{k=30}^{70} \binom{100}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^k$$

Porém, os cálculos exigidos por esta fórmula são demasiado laboriosos, mesmo que se recorra ao *triângulo aritmético*. Por métodos que se estudam em matemática superior, é possível achar o valor aproximado 0,9994 para esta probabilidade, efectuando cálculos muito simples e utilizando uma tabela numérica (tabela da *distribuição normal* ou *distribuição de Gauss*) [1]. Podemos, portanto, garantir com

[1] Em vez de 'distribuição binomial', 'distribuição normal', etc. também se diz 'lei binomial', 'lei normal', etc.

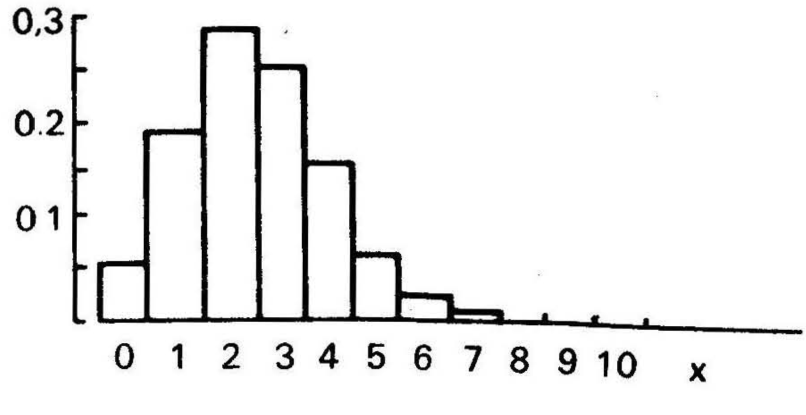
grande segurança (isto é, com pequeníssima probabilidade de errar) que, numa sequência de 100 provas, a frequência relativa do acontecimento considerado estará no intervalo fechado de extremos 0,3 e 0,7 (cf. considerações do n.º 9, pág. 224).

Assim, a distribuição binomial vem lançar nova luz sobre as relações entre o conceito de probabilidade e o conceito de frequência relativa.

Suponhamos, agora, $p = 1/4$ e $n = 10$. Neste caso, a distribuição binomial será, concretamente:

$$(1) \qquad P(x) = \binom{10}{x}\left(\frac{1}{4}\right)^{x}\left(\frac{3}{4}\right)^{10-x}$$

Na figura que a seguir se apresenta é dado o histograma desta distribuição, que já *não é simétrica*. Temos ainda aqui uma distribuição unimodal, sendo a *moda* o valor $x = 2$.



A distribuição binomial encontra importantes aplicações nos mais diversos domínios: na biologia, na psicologia, etc. Vamos apresentar um exemplo relativo à *genética,* ramo da biologia que estuda cientificamente os fenómenos da hereditariedade.

Segundo as experiências de MENDEL (1), relativas ao cruzamento de certas plantas *de flor genuinamente vermelha com outras*

---

(1) Monge austríaco (Silésia), fundador da Genética (1822-1884).

de flor genuinamente branca, verifica-se que, na primeira geração, todas as plantas dão flor vermelha *(carácter dominante),* ao passo que, na segunda geração (resultante dos cruzamentos entre híbridos da primeira geração), há a probabilidade 3/4 de aparecer uma planta de flor vermelha e a probabilidade 1/4 de aparecer uma planta de flor branca ('carácter recessivo') ([1]).

Assim, a probabilidade de, em 10 indivíduos da 2.ª geração, aparecerem x plantas de cor branca é dada pela fórmula (1).

Se, em vez de 10, consideramos 500 indivíduos, o *valor esperado* do número de indivíduos de flor branca (na 2.ª geração), será:

$$\frac{1}{4} \times 500 = 125$$

Mas o *valor observado* será geralmente diferente do valor esperado, sendo a diferença entre o primeiro e o segundo *(desvio* ou *discrepância)* devido ao *acaso.*

Por exemplo, a probabilidade de que, no conjunto dos referidos 500 indivíduos, o número de plantas de flor branca esteja compreendido entre 100 e 150, é dada pela expressão:

$$P\,(100 < x < 150) = \sum_{x=101}^{149} \binom{500}{x} \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^x$$

Todavia, tal como num dos exemplos anteriores, os cálculos exigidos por esta fórmula são impraticáveis. Utilizando o processo a que fizemos alusão (com o emprego duma *tabela da distribuição normal,* que não podemos aqui estudar), consegue-se facilmente obter o valor aproximado 0,003 para aquela probabilidade. É, portanto, quase certo que, no referido conjunto de 500 indivíduos, a frequência

---

([1]) Recordemos que, em biologia, se usa o termo 'carácter' (plural 'caracteres'), como sinónimo de 'atributo'.

relativa do atributo considerado esteja no intervalo aberto de extremos 0,2 e 0,3 ou mesmo numa vizinhança mais pequena de 0,25([1]).

Para interpretar estes e outros resultados análogos de Mendel, o biologista americano T. H. MORGAN (1866-1945, Prémio Nobel 1933), introduziu a hipótese da existência de partículas materiais, chamadas '*genes*', como factores de hereditariedade localizados nos cromossomas, conseguindo mesmo identificar a posição de alguns deles. Cada gene representa um carácter elementar: p. ex. 'flor vermelha' ou 'flor branca'. Todas as células de um indivíduo biológico (em particular, de cada um de nós, seres humanos) têm a mesma composição genética, sendo os genes de cada célula comparáveis às bolas de uma esfera de lotaria. Na verdade, como bolas de lotaria se comportam, em certa medida, essas minúsculas porções de matéria, no cruzamento de dois indivíduos, dando por vezes resultados imprevistos, ao entrarem na composição genética do novo ser: por exemplo, de pais com olhos castanhos pode nascer uma criança com olhos azuis, que recebe de um antepassado mais ou menos próximo o gene desse carácter, *segregado* na resultante genética dos progenitores.

18. **Casos extremos da distribuição binomial.** Como vimos, a distribuição binomial dá-nos a probabilidade de que, em n realizações duma prova $\mathcal{P}$, um acontecimento α de probabilidade p se verifique x vezes, sendo x um dado número inteiro tal que $0 \leqslant x \leqslant n$. Os dois casos extremos são, pois:

1) $x = n$ (o acontecimento α verifica-se nas n provas)

2) $x = 0$ (o acontecimento $\tilde{\alpha}$ verifica-se nas n provas)

([1]) Chama-se *vizinhança* dum número real a qualquer intervalo com centro em *a*.

A probabilidade do primeiro caso é, evidentemente:

$$P(n) = p^n$$

A probabilidade do segundo caso é:

$$P(0) = q^n \quad (\text{com } q = 1 - p)$$

Note-se, porém, que o contrário do acontecimento $x = 0$ não é o acontecimento $x = n$, mas sim o acontecimento que consiste em *α se verificar, pelo menos, uma vez nas* n *provas.*

A probabilidade desse acontecimento será, pois:

$$P(x \neq 0) = 1 - q^n = 1 - (1 - p)^n = \sum_{x=1}^{n} (-1)^{x-1} \binom{n}{x} p^x .$$

EXEMPLO 1. *Calcular a probabilidade de que, em* 6 *lances dum dado perfeito saia, pelo menos, uma vez o número* 1.

Neste caso $p = 1/6$, $n = 6$ e, portanto:

$$P(x \neq 0) = 1 - \left(\frac{5}{6}\right)^6 = \frac{31031}{46656} \approx 66\,\%$$

EXEMPLO 2. *Calcular a probabilidade de que, comprando ao todo mil bilhetes em mil lotarias sucessivas de* 100 000 *números, se tenha ao menos uma vez a sorte grande.*

Agora é $p = 10^{-5}$, $n = 1000$ e, portanto:

$$P(x \neq 0) = 1 - \left(1 - \frac{1}{10^5}\right)^{1000} = \frac{1000}{100\,000} - \binom{1000}{2} \frac{1}{100\,000^2} + \ldots$$

$$= 0{,}01 - 0{,}000\,0495 + \ldots \approx 0{,}00995$$

A probabilidade pedida é, pois, praticamente igual a 1 % (os termos omitidos no desenvolvimento são desprezíveis).

19. Valor médio, esperança matemática. Jogos equitativos. Tornemos ao exemplo da tabela n.º 1 da pág. 204. Suponhamos que se trata de achar a *média* das classificações dadas no exame em questão. Como se procede? É bem conhecido o processo:

Multiplica-se cada classificação pelo número de alunos que a obtiveram e divide-se o resultado pelo número total de alunos.

Dum modo geral, consideremos uma variável casual x que toma *m* valores reais

$$x_1, x_2, \ldots, x_m$$

respectivamente, com as *frequências absolutas*

$$\nu_1, \nu_2, \ldots, \nu_m$$

e seja $n = \nu_1 + \nu_2 + \ldots + \nu_m$. Chama-se *valor médio da variável* x (com esta distribuição de frequência) o número

$$\frac{\nu_1 x_1 + \nu_2 x_2 + \ldots + \nu_m x_m}{n}$$

que se representa por $M\{x\}$. É claro que, se pusermos

$$f_1 = \frac{\nu_1}{n}, \quad f_2 = \frac{\nu_2}{n}, \quad \ldots, f_m = \frac{\nu_m}{n}$$

(frequências relativas de $x_1, x_2, \ldots, x_m$), virá:

$$M\{x\} = f_1 x_1 + f_2 x_2 + \ldots + f_m x_m = \sum_{k=1}^{m} f_k x_k \quad (\textit{Porquê?})$$

Em rigor, não deveríamos dizer '*valor médio da variável casual* x', mas sim '*valor médio da sua distribuição de frequência (absoluta ou relativa)*'. Todavia, na prática, quando se fala de uma variável casual x, subentende-se geralmente que é dada uma distribuição de frequência ou de probabilidade dessa variável.

Como vimos, as probabilidades são *frequências relativas esperadas no futuro*. Haverá, portanto, um conceito correspondente ao de valor médio, quando passarmos de frequências a probabilidades.

DEFINIÇÃO. *Seja* x *uma variável casual que só pode tomar um número finito* r *de valores reais* $x_1,\dots,x_r$, *com probabilidades respectivamente* $p_1,\dots,p_r$. Chama-se *esperança matemática* (ou *valor esperado*) da variável x, e representa-se por $E\{x\}$, o número:

$$E\{x\} = p_1x_1+p_2x_2+\dots+p_rx_r = \sum_{k=1}^{r} p_kx_k \text{ (}^1\text{)}$$

Como exemplo, calculemos a esperança matemática duma variável x com distribuição binomial (ou, mais precisamente, a *esperança matemática da distribuição binomial)*. Por definição, é:

$$E\{x\} = \sum_{x=0}^{n} P(x)\cdot x,$$

com $P(x) = \binom{n}{x} p^xq^{n-x}$ e $q = 1 - p$. Será, pois:

$$E\{x\} = \sum_{x=0}^{n} x \cdot \frac{n!}{x!\ (n-x)!} p^xq^{n-x}$$

(1) Também se define 'esperança matemática' de uma variável casual com uma infinidade de valores possíveis, mas isso ultrapassa o carácter elementar desta introdução.

Visto o primeiro termo ser nulo (x=0), pode escrever-se:

$$E\{x\} = pn \cdot \sum_{x=0}^{n} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1}q^{n-x}$$
$$= pn(p+q)^{n-1}$$

e, portanto:

$$E\{x\} = np$$

Por exemplo, sendo $p = 1/2$ e $n = 10$, a esperança matemática (ou valor esperado) da distribuição binomial é 5: exactamente igual à moda da distribuição (ver n.º 17). Mas, sendo $p=1/2$ e $n=11$, a distribuição binomial tem, como vimos, duas modas (5 e 6) e a sua esperança matemática é agora 5,5.

Muitas vezes, em vez de 'esperança matemática' ou 'valor esperado', também se diz 'valor médio', mesmo quando se trata de distribuições de probabilidade.

Posto isto, chama-se *esperança matemática dum jogador* à soma $p_1x_1+...+p_nx_n$ dos produtos p_kx_k das quantias x_k que ele pode ganhar num dado jogo, pelas respectivas probabilidades p_k de as ganhar. Um jogo diz-se *equitativo,* quando a entrada de cada jogador (isto é, a quantia com que entra no jogo) é igual à sua esperança matemática.

São habitualmente equitativos os jogos *puramente de sorte* (como, por exemplo, o loto) entre pessoas que se reúnem para jogar a dinheiro. Não é equitativo, em geral, por exemplo, o bridge a dinheiro, precisamente porque não é um jogo de sorte, mas antes de *perícia.*

Mas há, também, jogos de sorte que não podem ser equitativos: por exemplo a roleta, a lotaria, etc. No caso da roleta ou da lotaria, a entrada do jogador tem de ser superior à sua esperança matemática: o excesso destina-se a contribuir para as despesas e receitas da banca ou instituição emissora. *Simplesmente a entrada deve ser proporcional* *à esperança matemática, isto é, deve existir um número* $k > 1$, *igual para todos os jogadores (constante de proporcionalidade), tal que, sendo* E' *a entrada de cada jogador e* E *a sua esperança matemática, se tenha sempre*:

$$E' = kE.$$

Podemos exprimir este facto, dizendo que o jogo é *equitativo entre o público,* o que equivale afinal a dizer, no caso da lotaria, que a *probabilidade de saída dos diferentes prémios é a mesma para todos os números da emissão.*

Por exemplo, na Santa Casa da Misericórdia de Lisboa, 60 % dos preços dos bilhetes de lotaria são destinados a prémios e os restantes 40 % a despesas e receitas da Instituição. Tem-se pois, neste caso:

$$0{,}6 \times E' = E \quad \text{ou seja} \quad E' = \frac{5}{3} E$$

Surge agora, naturalmente, a pergunta:

É o totobola um jogo equitativo para o público?

Claro que não.

O totobola não é um jogo puramente de sorte, mas sim um jogo de sorte e de perícia.

Na verdade, o facto de um jogador de totobola estar *bem informado* acerca das equipas que se defrontam aumenta a sua probabilidade de acertar, probabilidade que será muito maior, ainda, se o concorrente juntar a essa condição as seguintes: 1) ter uma apreciável *intuição* para prever os resultados dos desafios, intuição essa baseada na experiência; 2) dispor de *bastante dinheiro,* para poder, com uma aposta múltipla, cobrir uma zona de grande probabilidade (1).

(1) Esta última condição *por si só* é demasiado fraca e pode constituir uma tentação perigosa para pessoas ingénuas que desejam, à viva força, ser milionárias (ver pág. 246, ex. 7).

Os exemplos de pessoas que são únicas a acertar nos 13 resultados, com uma aposta dupla *feita ao acaso,* nada provam contra o anterior argumento; não só porque, em estatística, os casos individuais não contam, mas ainda por uma razão, aparentemente paradoxal, semelhante à que já foi indicada a propósito da distribuição binomial (cf. n.º 17):

Como o número dos peritos em totobola é extremamente reduzido em comparação com a massa total dos concorrentes, pode acabar por ser mais provável que acerte um concorrente não perito, principalmente se tivermos em conta a possibilidade de resultados que estejam fora de todas as previsões admissíveis.

Também não se pode inferir daqui que, quando os resultados estão ou parecem estar inteiramente fora das previsões admissíveis, só um *não perito* pode ser totalista. Entre os indivíduos a que podemos chamar 'peritos em totobola', há naturalmente uma grande diversidade de *graus de perícia* e serão, portanto, os *peritos-mores* (chamemos-lhes assim) os concorrentes com mais probabilidade de se *tornarem milionários do totobola.*

Como argumento final decisivo, a experiência parece confirmar as considerações de carácter apriorístico que acabamos de expor.

20. **Aplicação da teoria das probabilidades nos seguros.** Como vimos, foram os jogos de azar que deram origem ao cálculo das probabilidades, mas este acabou por ter numerosas aplicações importantes nos mais diversos domínios. Uma das mais antigas aplicações importantes do cálculo das probabilidades encontra-se na teoria matemática dos seguros. Esta é denominada *'cálculo actuarial'* e os especialistas neste ramo são chamados *'actuários'.*

O sistema dos seguros assemelha-se, em vários aspectos, ao das lotarias. Comecemos por um exemplo simples:

Um chefe de família vai fazer uma viagem de avião e, prevendo a possibilidade de um acidente fatal, procura proteger a família, fazendo um seguro de 2000 contos. Para 24 horas de viagem esta importa em cerca de 160 escudos *(prémio do seguro)*. É claro que neste caso, o acidente (ou *sinistro*) desempenha o papel da *sorte* na lotaria e que, portanto, o prémio do seguro deve ser superior à respectiva *esperança matemática,* isto é, ao produto do capital segurado pela probabilidade p do referido acidente, a fim de contribuir para as despesas e lucros da companhia seguradora. No exemplo concreto considerado deverá ter-se, pois:

$$p \times 2\,000\,000\$00 < 160\$00$$

Donde:

$$p < \frac{160}{2\,000\,000} = \frac{1}{12\,500}$$

Aliás, a probabilidade p, cujo valor só grosseiramente poderá ser avaliado, deve ser muito inferior a este limite, talvez inferior a 1/100 000, o que deve tranquilizar as pessoas que viajam de avião...

Assim, este ramo de seguros funciona como um *jogo equitativo entre os segurados:* tudo se passa como se, em caso de acidente, *todos* os segurados contribuissem para o pagamento do capital devido à família do sinistrado, sendo essa contribuição *equitativa,* isto é, proporcional à importância que cada segurado estipula para emergência análoga.

Um exemplo semelhante, embora mais complexo, é o dos *seguros de vida*. O cálculo destes seguros é baseado em *tabelas de mortalidade,* construídas por métodos de estatística matemática a partir de dados empíricos respeitantes a populações muito numerosas, distribuídas por vastas regiões e por longos períodos. Essas tabelas fornecem *frequências relativas ajustadas* (isto é, submetidas a certas correcções teóricas), que podem ser assumidas como *probabilidades,* num futuro não muito distante da época em que foram elaboradas tabelas.

As companhias portuguesas de seguros adoptam ainda, por lei, as tábuas francesas AF (abreviatura de 'assurés français') e RF (abreviatura de 'rentiers français'), que foram elaboradas no século passado.

No fim do volume apresenta-se um extracto da tábua AF. A tábua refere-se à evolução de uma população humana hipotética, que começa com um milhão de indivíduos, sem distinção de sexos. Na primeira coluna, indica-se o *número* x *de anos de idade* e, na segunda coluna, em correspondência por linhas, *o número de vivos com a idade* x, número este que vamos designar por v_x (os actuários usam, neste caso, o símbolo ℓ_x, abreviatura do inglês 'number living at age x'). Por exemplo, consultando a tábua AF para x = 20, x = 50 e x = 102, encontra-se:

$$v_{20} = 824159 \quad , \quad v_{50} = 628727 \quad , \quad v_{102} = 3$$

Quer isto dizer, segundo a referida tábua, que, entre um milhão de indivíduos que nascem, 824159 atingem os 20 anos, 628727 atingem os 50, e só 3 atingem os 102 anos; por outras palavras, as probabilidades de atingir essas idades são, respectivamente:

$$\frac{v_{20}}{v_0} = 0{,}824159, \quad \frac{v_{50}}{v_0} = 0{,}628727, \quad \frac{v_{102}}{v_0} = 0{,}000003$$

A última é, como se vê, extremamente pequena, *em todo o caso superior à probabilidade de acertar nos* 13 *resultados do totobola, com uma aposta dupla puramente casual.*

Segundo a mesma tábua, nenhum dos indivíduos da população inicial de um milhão atinge os 104 anos. Não quer isto dizer que seja *impossível atingir* essa idade ou ainda outras superiores (citam-se casos, naturalmente *raríssimos,* de macróbios com 120 anos e mais).

A única conclusão a tirar daí é esta: *a probabilidade que tem um recém-nascido de ultrapassar os* 103 *anos é inferior a um milionésimo* (segundo a referida tábua).

É fácil reconhecer, agora, o seguinte:

A probabilidade que um indivíduo de idade x *tem de chegar a uma idade* $y \geqslant x$ *é* v_y/v_x.

Por exemplo, a probabilidade que um indivíduo de 20 anos tem de sobreviver até aos 50 (pelo menos) é:

$$\frac{v_{50}}{v_{20}} = \frac{628727}{824159} \approx 0,76$$

portanto, superior à probabilidade que tem um recém-nascido de chegar aos 50 ($\approx 0,63$).

Como exercício, calcule a probabilidade de que duas pessoas com idades x e x' atinjam ambas as idades y e y' respectivamente (com $y \geqslant x$ e $y' \geqslant x'$).

Um conceito intimamente relacionado com este assunto é o de *vida média* duma pessoa (após uma idade x). A questão pode pôr-se intuitivamente nestes termos:

Calcular o número médio de anos de vida que têm ainda à sua frente as pessoas de idade x.

Para obter este valor médio *em primeira aproximação,* admite-se uma hipótese simplificadora que, evidentemente, não é verificada na prática: *todas as pessoas vivem um número inteiro de anos, falecendo imediatamente após o último aniversário.* Nesta hipótese, o número de pessoas de idade x que sobrevivem só 1 ano é $v_{x+1} - v_{x+2}$; o número de pessoas de idade x que sobrevivem só 2 anos é $v_{x+2} - v_{x+3}$, e assim por diante. Nestas condições, a vida média após a idade x obtém-se dividindo por v_x (número total de indivíduos que atingiram a idade x) a soma dos produtos dos anos de sobrevivência pelos respectivos números de indivíduos. Assim, se designarmos por ω o número máximo de anos de vida (103 na tabela AF) e por e_x a vida média após a idade x, teremos:

$$(1) \qquad e_x = \frac{(v_{x+1} - v_{x+2}) + 2(v_{x+2} - v_{x+3}) + \ldots + (\omega - x)v_\omega}{v_x}$$

$$= \frac{v_{x+1} + v_{x+2} + \ldots + v_\omega}{v_x} = \frac{1}{v_x} \sum_{n=1}^{\omega - x} v_{x+n}$$

Para obter aproximações cada vez melhores da vida média após a idade x, deveríamos dispor de tábuas de mortalidade sucessivamente mais minuciosas, por exemplo por trimestres, por meses, por semanas, etc., que nos permitissem substituir a hipótese inicial por hipóteses cada vez mais próximas da realidade. Todavia, na prática, tem-se revelado suficiente uma segunda aproximação, que consiste em adicionar meio ano ao valor médio anterior. Esta segunda aproximação costuma ser designada pelo símbolo $\mathring{e}_x$. Tem-se, pois, por definição:

$$\mathring{e}_x = \frac{1}{2} + e_x$$

Em vez de 'vida média após a idade x', também se diz '*esperança de vida na idade* x' (donde a escolha da letra *e* nas anteriores notações). Porém, a primeira expressão é mais de origem *estatística*, referindo-se ao *passado*, e só enquanto se aplica ao futuro pode, com propriedade, ser substituída pela segunda, de carácter *probabilista*. Sob este aspecto, trata-se, na realidade, de uma esperança matemática, como vamos ver, relativamente a e_x.

Com efeito, de (1) resulta imediatamente que

$$e_x = 1 \cdot \frac{v_{x+1} - v_{x+2}}{v_x} + 2 \cdot \frac{v_{x+2} - v_{x+3}}{v_x} + \ldots + (\omega - x) \frac{v_\omega}{v_x}$$

Ora $(v_{x+1}-v_{x+2})/v_x$ é a *probabilidade* que tem um indivíduo de idade x de viver só mais 1 *ano*, $(v_{x+2}-v_{x+3})/v_x$ a *probabilidade* que tem um indivíduo de idade x de viver só mais 2 *anos*, e assim sucessivamente. Logo, a expressão anterior dá, efectivamente, segundo a definição do n.º 19, a esperança matemática da variável casual

$$n = \textit{número de anos de vida após a idade } x,$$

adoptando a hipótese simplificadora inicial.

Posto isto, suponhamos que uma pessoa de idade x pretende fazer um seguro de vida na importância de C escudos, que deverá ser entregue à pessoa ou às pessoas designadas pela primeira, após a sua morte. Suponhamos que esta decide pagar para esse fim um prémio anual de *c* escudos, enquanto viva. Se o sistema fosse equitativo e se os prémios não rendessem juro, o produto de *c* pela parte inteira da vida média dessa pessoa após a idade x deveria ser igual a C, isto é:

$$mc = C,$$

sendo *m* a parte inteira (ou característica) de $\overset{\circ}{e}_x$.

Mas, já sabemos que o sistema só pode ser equitativo em relação ao público entre si, visto que o prémio deve contribuir não só para a constituição do capital C, mas também para as despesas, riscos e receitas da companhia seguradora.

Por outro lado, os prémios devem ficar a render juros compostos, a favor do segurado, durante a vida deste, com determinada taxa r anual (geralmente 4 %).

Estas e outras circunstâncias vêm complicar o cálculo dos seguros de vida inteira (bem como de outras modalidades afins), com pormenores técnicos que ultrapassam o âmbito desta iniciação elementar.

21. **As variações da probabilidade no cálculo de seguros.** Como vimos nos n.os 11 e 12, a atribuição de uma probabilidade p a um acontecimento α assenta no método de indução e está, portanto, sujeita a todas as limitações próprias deste método, nomeadamente no espaço e no tempo: *não podemos extrapolar os resultados da experiência para além de certos limites.* Isto verifica-se com a própria física e até com a geometria. Por exemplo, o teorema segundo o qual a soma dos ângulos dum triângulo é igual a 180° é válida até às aproximações exigidas na prática, numa região do espaço que não exceda muito as dimensões da Terra; porém, a TEORIA DA RELATIVIDADE veio mostrar que, em domínios astronómicos mais extensos, a soma dos ângulos dum triângulo pode ser sensivelmente superior a 180°(1). Isto revela que os postulados da geometria euclidiana, e em especial o próprio POSTULADO DE EUCLIDES (2), deixam de dar resultados com aproximação razoável em tais domínios.

É, portanto, natural que a probabilidade dum acontecimento (ou melhor, de um tipo de acontecimentos) esteja sujeita a variações mais ou menos imprevisíveis. Por vezes, essas variações são tão rápidas, que nem sequer faz sentido falar de probabilidade, por *ausência de regularidade estatística,* mesmo em zonas ou períodos bastante limitados. Outras vezes as variações de probabilidade são bastante lentas, para que os números atribuídos como probabilidades aos acontecimentos possam ser aceites com relativa confiança; é o que sucede, por exemplo, com as probabilidades indicadas por tábuas de mortalidade não muito antigas, abstraindo, é claro, de períodos excepcionais de guerras, grandes epidemias, etc. e de regiões do Globo a que essas tábuas não se apliquem.

---

(1) Neste caso chamam-se *segmentos de recta* as trajectórias dos fotões no vazio.

(2) 'Dados um ponto e uma recta, existe sempre uma recta e uma só que passa pelo ponto dado e é paralela à recta dada'. Na geometria que melhor se adapta a domínios astronómicos, não existem rectas paralelas.

Já atrás foi mencionado que as companhias portuguesas de seguros ainda usam, por lei, as tábuas francesas AF e RF, elaboradas no século passado. *Essas tábuas fornecem probabilidades diferentes.* Como se explica esta discordância em tábuas do mesmo país e da mesma época? As razões são as seguintes:

A tábua AF ('assurés français') destina-se essencialmente a *seguros de vida* e a tábua RF ('rentiers français') a *rendas vitalícias.* É natural que tenham maior tendência para a segunda modalidade pessoas saudáveis que esperam ter uma longa vida e que se decidem a confiar capitais a companhias de seguros ou ao Estado, a fim de serem convertidos, com os respectivos juros, em *rendas vitalícias,* de carácter periódico (mensais, trimestrais, etc.) conforme o estipulado. Pelo contrário, é natural que tenham maior tendência para seguros de vida pessoas com preocupações de saúde. Ora, as tábuas foram organizadas não só utilizando os resultados de censos populacionais, mas também com base na experiência das próprias companhias *— e essa experiência parece confirmar as previsões anteriores, apesar de as companhias tomarem a precaução de sujeitar a inspecção médica as pessoas que pretendem fazer seguros de vida.* Assim se explica, portanto, que as tábuas RF dêem probabilidades de vida um pouco superiores às correspondentes da tábua AF.

Deve acrescentar-se, no entanto, que estão actualmente em estudo entre nós tábuas de mortalidade recentes, também de origem francesa, uma para cada sexo: tábua PM ('population masculine') e tábua PF ('population féminine'). Verifica-se que as pessoas do sexo feminino têm, para cada idade x, vida média geralmente superior às do sexo masculino para a mesma idade, o que está de acordo com certos factos que chamam a atenção: por exemplo, parece ser mais frequente a viuvez no sexo feminino do que no masculino.

No final do volume é apresentado um extracto da tábua PM. Comparando-a com a tábua AF observa-se um considerável acréscimo das vidas médias (ou esperanças de vida) para cada idade: assim, para a idade 0, a tábua AF dá como vida média 52,0187 anos enquanto a tábua PM dá 66,9551, *cerca de* 15 *anos mais;* para a idade 50, a tábua AF dá como vida média 19,6593 anos, enquanto a tábua PM dá 22,5495, *cerca de* 3 *anos mais,* etc. Estes acréscimos, que podem ser explicados pelos progressos da medicina e pelas condições mais higiénicas da vida moderna, são confirmados, ao que parece, pela própria experiência das companhias que continuam a adoptar as tábuas AF e RF: o ramo dos seguros de vida tem-se tornado mais lucrativo (aumentando a vida média, aumenta o número de prémios pagos pelos segurados na modalidade *vida inteira*); ao passo que as rendas vitalícias se tornam menos vantajosas (vivendo mais tempo, as pessoas recebem maior número de rendas).

Observa-se, no entanto, um facto estranho ao comparar as duas tábuas: no primeiro ano de vida a tábua PM dá um número de mortes (68620) quase duplo do que é dado pela tábua AF (36015). Este facto contradiz a afirmação corrente de que os progressos da medicina e da higiene, associados ao progresso social, têm diminuído enormemente a mortalidade infantil. Deve notar-se, porém, que a tábua AF não merece grande confiança relativamente às primeiras idades, quer pela insuficiência de dados estatísticos, quer pelos métodos de ajustamento utilizados. Convém assinalar entretanto que, logo na idade 1, a tábua PM apresenta um número de mortes (6706), que é cerca de 30 % do valor correspondente dado pela tábua AF e esta forte diminuição da mortalidade mantém-se em todos os anos de infância e da adolescência.

Finalmente deve dizer-se que foi elaborada, há poucos anos, uma tábua de mortalidade portuguesa.

22. **Interpretação estatística duma tábua de contingência. Teste do qui-quadrado com um grau de liberdade.** Imaginemos uma experiência destinada a averiguar da eficácia de um certo tratamento em determinada doença. Suponhamos que o expe-

rimentador dispõe de 300 animais atacados dessa doença e que decide submeter 100 destes ao referido tratamento, reservando os outros 200 para *contrôle* (ou *testemunho*). Suponhamos, além disso, que se obtiveram os resultados indicados na seguinte tabela [1]:

TABELA DOS VALORES OBSERVADOS

	Sobreviventes	Mortos	Total
Não tratados	152	48	200
Tratados	88	12	100
TOTAL	240	60	300

Para averiguar *em que medida* estes resultados autorizam a admitir uma real influência do tratamento na doença (no sentido da cura ou no sentido do agravamento), adopta-se uma atitude mental semelhante ao método de demonstração por redução ao absurdo usado em matemática pura. Consiste essa atitude em admitir uma hipótese, chamada *hipótese nula* (ou *hipótese de independência*), que é exactamente a negação daquilo que se pretende estabelecer.

HIPÓTESE NULA. *O tratamento adoptado não influi na mortalidade da doença, nem num sentido nem no outro.*

Admitida esta hipótese, seria de esperar que os resultados obti-

[1] Exemplo dado por FINNEY na sua obra 'An introduction to statistical science on agriculture'.

dos fossem aqueles a que, no n.º 5, chamámos *valores de independência* (ou *valores esperados*) e que são os indicados na seguinte tabela:

TABELA DOS VALORES ESPERADOS (¹)

	Sobreviventes	Mortos	Total
Não tratados	160	40	200
Tratados	80	20	100
TOTAL	240	60	300

Estes valores foram obtidos segundo a fórmula definidora de $\Phi^o(\alpha\,\beta)$ (ver pág. 210), que se aplica igualmente a $\Phi_o(\tilde{\alpha}\,\beta)$, $\Phi_o(\tilde{\alpha}\,\tilde{\beta})$, etc. Assim, tem-se:

$$160 = \frac{240 \times 200}{300}, \qquad 80 = \frac{240 \times 100}{300} = 240 - 160,$$

$$40 = \frac{60 \times 200}{300} = 200 - 160, \qquad 20 = \frac{60 \times 100}{300} = 100 - 80.$$

Quer dizer: segundo a hipótese nula, o número de sobreviventes não tratados deveria ser 160, a dos sobreviventes tratados 40, etc. Com efeito, só deste modo a percentagem dos sobreviventes entre os tratados seria igual à dos sobreviventes entre os não tratados, ou seja:

$$\frac{80}{100} = \frac{160}{200} = 80\,\%,$$

(¹) Aqui 'valores esperados' é abreviatura de 'valores esperados *na hipótese da independência*'.

e, analogamente, a percentagem dos mortos entre os tratados seria igual à dos mortos entre os não tratados, ou seja:

$$\frac{20}{100} = \frac{40}{200} = 20\,\%$$

Porém, os valores observados não coincidem com os correspondentes valores esperados. As diferenças entre os primeiros e os segundos, chamados *desvios* (ou *discrepâncias*) são indicados na seguinte tabela:

TABELA DOS DESVIOS

	Sobreviventes	Mortos	Total
Não tratados	− 8	8	0
Tratados	8	− 8	0
TOTAL	0	0	0

É claro que, em tabelas de contingência deste tipo, basta conhecer um dos desvios para que os restantes fiquem determinados, uma vez que *as somas dos desvios, tanto por linhas como por colunas,* têm de ser zero. Exprime-se este facto dizendo que as tabelas deste tipo têm *um só grau de liberdade.*

Surge, agora, a pergunta:

Estão estes desvios em contradição com a hipótese nula?

Em rigor, isto é, em matemática pura baseada na lógica bivalente do 'sim ou não', a resposta só pode ser 'sim'. Não esqueçamos, porém, que não se trata agora de proposições matemáticas, mas de *factos empíricos,* necessariamente *contingentes.*

Por outras palavras:

O facto de o tratamento não influir na doença não obriga, de modo nenhum, a uma rígida confirmação dos valores de independência, pois pode haver desvios não nulos *devidos ao acaso*; só se os desvios excederem, em valor absoluto, um *certo limite* é que haverá razão para *rejeitar a hipótese nula*. Tudo está, portanto, em avaliar aproximadamente esse limite, para além do qual deixará de ser admissível a hipótese nula.

* * *

A situação anterior é análoga à que se levanta a propósito de lançamentos sucessivos duma moeda ao ar. Supondo que a moeda é equilibrada, é de *esperar* que, por exemplo, numa sequência de 100 lançamentos, se apresente *escudo* 50 vezes, isto é, que a frequência relativa deste acontecimento seja 1/2. Mas, é possível (e até, como vimos, *muito mais provável*), que a frequência absoluta desse acontecimento em 100 provas seja diferente de 50; simplesmente, a frequência absoluta será tanto menos provável quanto mais se afastar do *valor esperado* (ou *esperança matemática*), que é, como vimos, 50. Aliás, a distribuição binomial (n.^{os} 17 e 18) habilita-nos a calcular, neste caso, a probabilidade de que o desvio (diferença entre o valor esperado e o valor observado) tenha módulo inferior a um certo limite. Todavia, como vimos, os cálculos exigidos pela distribuição binomial para 100 provas são demasiado laboriosos, o que o obriga a um cálculo aproximado (tanto mais aproximado quanto maior é o número de provas), recorrendo à *distribuição normal*, que ainda não estamos em condições de estudar.

Tornando à experiência do exemplo inicial, podemos imaginar um tipo de provas semelhantes ao que consiste em atirar uma moeda ao ar, para ver em que medida os desvios verificados são atribuíveis unicamente ao acaso.

Tomem-se 300 bocados de cartão, sensivelmente iguais em forma, dimensões e substância, e distingam-se 100 desses cartões com a cor vermelha (representativos dos 100 animais tratados) e os 200 restantes com a cor branca (representativos dos animais não tratados).

Executa-se, depois, um *grande número de vezes seguidas* a prova $\mathcal{P}$ que consiste nas seguintes operações:

1.º *Fechar todos os cartões numa caixa bastante espaçosa.*

2.º *Agitar fortemente a caixa várias vezes.*

3.º *Tirar da caixa, ao acaso,* 60 *cartões (representativos dos animais mortos).*

4.º *Registar o número de cartões vermelhos existentes nessa amostra de* 60 *cartões.*

É claro que a prova $\mathcal{P}$ implica a reposição dos cartões retirados na prova anterior.

O número x de cartões vermelhos em cada amostra indica o *número de animais tratados mortos,* numa experiência hipotética, em que o desvio $x-20$ entre x e o valor esperado 20 seja atribuível exclusivamente ao acaso (segundo a HIPÓTESE NULA). Uma vez conhecido este desvio, imediatamente se determinam os restantes, visto que as somas por linhas e por colunas são nulas, como vimos.

Ora, no caso em estudo, o módulo dos desvios verificados é 8. Portanto, o que está indicado é *determinar a frequência relativa com que, nas referidas provas casuais, se apresentam desvios cujo módulo é igual ou superior a* 8.

Quem tiver a paciência de efectuar um número muito grande de tais provas (pelo menos 1000) verificará que *só em cerca de* 2% *das provas se apresentam desvios com módulo igual a* 8. Somos, assim, levados a concluir, por indução, que:

A probabilidade de desvios puramente casuais, com módulo igual ou superior a 8, *é aproximadamente* 2%;

ou ainda, o que é equivalente:

A probabilidade de desvios puramente casuais, com módulo inferior a 8, *é cerca de* 98%.

Como esta probabilidade é muito próxima de 1, podemos tomar como *quase certo* que os desvios puramente casuais têm módulo inferior a 8. *Mas isto é contrário à hipótese nula.* Somos, portanto, levados a *rejeitar a hipótese nula,* isto é, a concluir, com uma *boa margem de segurança,* que:

O tratamento influi realmente na evolução da doença, sendo o efeito benéfico, visto ser positivo o desvio correspondente a 'animais tratados sobreviventes'.

* * *

Releia, com muita atenção, todas as considerações anteriores acerca da experiência em estudo. *Analise-as com espírito crítico.* A argumentação usada é de tipo muito diferente do das demonstrações matemáticas, apesar da analogia com as demonstrações por redução ao absurdo. Para quem se habituou ao chamado 'rigor lógico da matemática', deixam uma certa insatisfação no espírito, não é verdade? Mas é tempo de se ir habituando também à ideia de que o rigor matemático se refere a um ideal platónico de perfeição absoluta, que jamais se encontra realizado neste mundo em que vivemos — *muito embora nos seja indispensável para interpretar, com êxito cada vez maior, esta mesma realidade, sempre mutável e imperfeita.* Já vimos que todas as leis experimentais sofrem do mesmo carácter contingente, que resulta do método indutivo. Mas são essas leis que marcam o progresso real do homem no conhecimento do mundo empírico.

Um dos inconvenientes do método anterior é o de obrigar a um enorme número de provas fastidiosas. Todavia esse inconveniente pode ser evitado, recorrendo precisamente ao *método dedutivo da matemática,* segundo os princípios do cálculo das probabilidades. Com efeito, não é difícil reconhecer que a probabilidade de, numa amostra casual de 60 cartões, aparecerem x cartões vermelhos é [1]:

$$P(x) = \frac{\binom{100}{x}\binom{200}{60-x}}{\binom{300}{60}} \quad \textit{(distribuição hipergeométrica)}$$

---

(1) A dedução desta fórmula pode ficar como exercício para os alunos mais interessados. É fácil determinar o número de casos possíveis e o número de casos favoráveis.

Esta fórmula permite, pelo menos teoricamente, calcular a probabilidade de um desvio x−20 com módulo igual ou superior a 8. Essa probabilidade, até à ordem das centésimas, é precisamente:

$$P(|x-20| \geqslant 8) = 0{,}0210 \approx 2\,\%$$

Simplesmente, tal como no caso da distribuição binomial, os cálculos exigidos pela utilização da fórmula (1) são impraticáveis. Mas, também como no caso da distribuição binomial, podemos rodear a dificuldade recorrendo a um cálculo aproximado que se baseia numa outra distribuição. Com efeito:

A probabilidade de que os *desvios casuais* não excedam, em valor absoluto, um certo limite pode ser calculada aproximadamente, recorrendo a uma lei de probabilidade chamada DISTRIBUIÇÃO DE PEARSON. Embora não estejamos ainda em condições de fazer o estudo teórico desta lei de probabilidade, *podemos, desde já, indicar como pode ser usada na prática, recorrendo a tabelas numéricas dessa distribuição*. Para isso, começa-se por calcular um *índice estatístico* da tabela de contingência, que se designa pelo símbolo χ^2 (lê-se 'qui-quadrado', visto 'χ' ser a letra grega chamada 'qui') e que é assim definida:

O χ^2 duma tabela de contingência é a soma dos números que se obtêm dividindo o quadrado de cada desvio pelo valor esperado correspondente. Assim, no exemplo anterior, o valor de χ^2 é:

$$\chi^2 = \frac{(-8)^2}{160} + \frac{8^2}{80} + \frac{8^2}{40} + \frac{(-8)^2}{20} = 6$$

Ora bem, demonstra-se em Estatística Matemática a seguinte proposição, cujo significado irá sendo compreendido posteriormente com exemplos práticos(1):

---

(1) Como estamos a seguir uma exposição de tipo prático, renunciamos a precisar o sentido exacto desta proposição.

O χ^2 duma tabela de contingência inteiramente casual segue aproximadamente a lei de Pearson, sendo a aproximação tanto maior quanto maior for o máximo valor esperado.

Por este facto, a lei de Pearson também é chamada *'distribuição do qui-quadrado'*. No final do volume encontra-se uma tábua desta distribuição, apresentada do seguinte modo:

A tábua tem duas entradas, uma para o *número de graus de liberdade* (n) *e a outra para a probabilidade* (P). Suponhamos, por exemplo, n = 1 (o caso que nos interessa por agora) e P = 0,02 ou P = 0,01. No cruzamento da linha 1 com as colunas 0,02 e 0,01 encontramos, respectivamente, os valores 5,41 e 6,64 do χ^2. Quer isto dizer o seguinte:

A probabilidade de um χ^2 igual ou superior a 5,41 *é* 0,02; *a probabilidade de um X^2 igual ou superior a* 6,64 *é* 0,01 *(supondo que os desvios se devem exclusivamente ao acaso)*. Simbolicamente:

$$P(\chi^2 \geqslant 5{,}41) = 2\,\%, \quad P(\chi^2 \geqslant 6{,}64) = 1\,\%$$

Analogamente se procede em qualquer outro caso. No exemplo anterior, o valor obtido do χ^2 é 6, portanto *compreendido entre* 5,41 e 6,64. Nestas condições, a probabilidade de um χ^2 *igual ou superior a* 6 *estará entre* 1 % *e* 2 %.

Mas é preciso não esquecer que o uso da distribuição de Pearson em casos como este envolve aproximações que só são aceitáveis quando os valores esperados são relativamente grandes. *Uma regra que costuma ser recomendada na prática é a de não utilizar a distribuição de Pearson quando algum dos valores esperados é inferior a* 5. Contudo, a aproximação pode ser melhorada consideravelmente, em qualquer caso, por um simples artifício chamado *'correcção de YATES'*, que consiste em diminuir 0,5 ao módulo de cada desvio. Assim, no caso em estudo, temos o *valor corrigido* (1):

$$\chi^2 = \frac{(-7{,}5)^2}{160} + \frac{7{,}5^2}{40} + \frac{7{,}5^2}{80} + \frac{(-7{,}5)^2}{20} = 5{,}27$$

---

(1) Por comodidade continuamos a designar por χ^2 o valor corrigido.

Ora, este valor aproxima-se muito mais do valor 5,41 correspondente a P = 0,02, que do valor 3,84 correspondente a P = 0,05. Assim se confirma o que já tinha sido indicado atrás como resultado de métodos muito mais laboriosos:

A probabilidade de, na referida experiência, haver desvios puramente casuais com módulo igual ou superior a 8, *é aproximadamente* 0,02.

* * *

Esta conclusão levou-nos a rejeitar a hipótese nula, mas aí está um outro motivo de insatisfação que nos deixa o método estatístico usado. Nós concluímos que é *quase impossível* haver desvios casuais com módulo igual ou superior a 8 (ou, o que é equivalente, que é *quase certo* os desvios casuais terem módulo inferior a 8). Mas o que se entende por 'quase certo' e por 'quase impossível'? Em matemática pura todo o conceito não primitivo é definido com rigor absoluto e de modo inteiramente *objectivo,* isto é, com um critério independente das pessoas e das circunstâncias. Agora, pelo contrário, é necessário usar critérios subjectivos, que dependem da intuição e do bom senso do experimentador, assim como da natureza da própria experiência e das circunstâncias que a rodeiam.

Em certas investigações biológicas é usual considerar já como *muito pequena* a probabilidade de 1 por 20 (0,05), embora muitas vezes se prefira considerar como *muito pequena* a de 1 por 50 (0,02) ou a de 1 por 100. Cada uma destas probabilidades é chamada *'nível de significância'*, no método estatístico descrito, o qual por sua vez é denominado *'teste de significância do* χ^2,' ou simplesmente *'teste do* χ^2'.

Convenciona-se chamar *'significante'* o nível de 5 % e *'altamente significante'* o nível de 1 %. Mas convém salientar que se trata aqui apenas de convenções cómodas de linguagem, adoptadas por estatísticos e experimentadores.

Aliás, não são estes os únicos níveis de significância adoptados na prática. Casos há em que o nível de 1 por 10 é já considerado como significante (especialmente em investigaçãos médicas em que o tratamento consiste numa ligeira alteração de regime, pouco dispendiosa e sem risco para o doente); e casos há em que, pelo contrário, se requer um nível de 1 por 1000 (quando se trate de alterações profundas ou muito dispendiosas).

Além disso, o teste do χ^2 não é senão um exemplo, entre muitos, dos *testes de significância* oferecidos pela estatística matemática. Nas suas linhas gerais, um teste de significância compreende os seguintes passos:

1) *Formular a hipótese nula, que nega o facto experimental a estabelecer.*

2) *Escolher o nível de significância que pareça mais adequado à natureza da experiência e ao tipo de decisão a tomar.*

3) *Calcular a probabilidade de que os valores absolutos dos desvios puramente casuais sejam iguais ou superiores àqueles observados.*

4) *Se a probabilidade obtida em* 3) *é inferior ao nível de significância adoptado em* 2), *rejeitar a hipótese nula; caso contrário, deixar a conclusão em suspenso.*

* * *

Note-se que o papel da estatística matemática não se limita, de modo nenhum, à *interpretação de resultados experimentais:* os métodos estatísticos devem ser usados, primeiro que tudo, no *planeamento das experiências*. Não podemos aqui estudar este assunto de alto interesse. Limitar-nos-emos, por isso, a breves indicações sobre a natureza do problema, no caso particular do tratamento hipotético atrás considerado.

É claro que, quanto maior for o número total de casos individuais estudados, melhor informação fornecem os resultados obtidos; mas, esse número é forçosamente limitado, especialmente por razões de ordem económica, *e, por isso, mais necessário se torna tirar o maior rendimento possível do material de que se dispõe.*

Suponhamos, por exemplo, que se tinham escolhido apenas 2 ou 3 animais para serem tratados entre os 300, reservando os restantes para *contrôle* (ou vice-versa); é manifesto mesmo *a priori* que os resultados não autorizariam um juízo seguro. Deve, portanto, haver uma proporção óptima para revelar a eficácia do tratamento, caso este tenha de facto efeito. Porém, essa proporção não é, como poderia parecer à primeira vista, a de 50 % para os dois grupos. Na tabela seguinte estão indicados os valores do χ^2 para diferentes proporções de animais tratados, entre os 300, na hipótese de a percentagem de mortos entre animais tratados ser a mesma que se observou na experiência imaginada (12 %).

N.º de animais tratados	para 12 % de mortalidade
25	1,2
50	2,8
100	5,3
125	6,1
150	6,5
175	6,6
200	6,3
250	4,0
275	2,0

Vê-se que o máximo valor do χ^2 é atingido na proporção de 175 tratados para 125 não tratados. Deve ser, portanto, essa a proporção ideal para a experiência.

Há, ainda, outros pormenores relativos à realização da experiência que precisam de ser observados, para que o teste de significância não resulte ilusório. É óbvio que, se os animais tratados tiverem uma proveniência diferente da dos animais não tratados, as diferenças registadas podem não ser devidas ao tratamento nem ao acaso. Para atenuar o mais possível os factores estranhos ao tratamento, o que há a fazer é destacar os 300 animais duma população tão homogénea quanto possível, e entre esses escolher, *completamente ao acaso*, os 100 animais a serem tratados, dando a todos os 300 igual probabilidade de serem escolhidos. Esta operação, chamada *amostragem casual*, tem de ser efectuada por um processo objectivo, em que não intervenha qualquer factor pessoal de escolha. Um tal processo pode consistir no seguinte: *numerar os animais de* 1 *a* 300, *deitar numa caixa* 300 *cartões iguais, numerados de* 1 *a* 300, *e tirar* 100 *à sorte; os números saídos serão os dos animais a tratar.*

Nunca é de mais encarecer a importância da *amostragem casual*, neste e noutros tipos de experiências, como base duma investigação estatística bem orientada. Não podemos, contudo, indicar aqui os pormenores de técnica, a que tem por vezes de obedecer uma tal operação, que é hoje objecto de estudos desenvolvidos em estatística matemática.

NOTA SOBRE A TERMINOLOGIA. Entre as palavras 'teste', 'significância' e 'significante', atrás usadas, as duas primeiras são anglicismos. Por isso, alguns autores portugueses preferem a essas, respectivamente, as palavras 'prova', 'significação' e 'significativo'. A conveniência em adoptar as primeiras está em que, quando se trata de termos técnicos (como é aqui o caso), importa evitar qualquer confusão com acepções da linguagem comum.

23. **Teste do qui-quadrado com mais de um grau de liberdade.** Limitar-nos-emos a um exemplo ([1]). Trata-se de colheitas de cevada em 260 campos, feitas em 1942. No Inverno e na Primavera antecedentes, tinha-se avaliado em cada campo a frequência

---

([1]) Extraído da obra de FINNEY, já citada.

da larva dum coleóptero designado em inglês por 'wireworm' (variedade de insecto chamado em português 'alfinete'). Segundo determinado critério, classificou-se a infestação dos diferentes campos em 'baixa', 'moderada', 'alta', e 'muito alta'. Por sua vez, a qualidade das colheitas foi classificada em 'satisfatória' e 'não satisfatória'. Os resultados são os que constam da seguinte tabela:

Resultados das colheitas	Infestação				Total
	Baixa	Moderada	Alta	Muito alta	
Satisfatórios	94	62	31	15	202
Não satisfatórios	15	15	17	11	58
Total	109	77	48	26	260

Daqui se deduziu uma tabela dos valores esperados e uma outra dos desvios (tabelas que não reproduzimos). O valor do χ^2 será ainda, neste caso, *a soma dos números que se obtêm, dividindo o quadrado de cada desvio, pelo respectivo valor esperado*. Acha-se, então:

$$\chi^2 = \frac{9{,}3^2}{84{,}7} + \frac{2{,}2^2}{59{,}9} + \frac{(-6{,}3)^2}{37{,}3} + \frac{(-5{,}2)^2}{20{,}2} + \frac{(-9{,}3)^2}{24{,}3} + $$
$$+ \frac{(-2{,}2)^2}{17{,}2} + \frac{6{,}3^2}{10{,}9} + \frac{5{,}2^2}{5{,}8} = 15{,}7$$

A hipótese nula formula-se, agora, do seguinte modo:

HIPÓTESE NULA. *A infestação não influi em nada na qualidade da colheita.*

Adoptemos o nível 1 % (altamente significante). Para utilizar o teste do χ^2 há que notar, agora, os seguintes factos:

1) A distribuição de Pearson não depende do número de casos

considerados, mas depende do *número de graus de liberdade* da tabela de contingência. Neste caso, trata-se de uma tabela 2 × 4 (isto é, com 2 linhas e 4 colunas). Como, na correspondente tabela de desvios, são nulas as somas dos desvios quer por linhas quer por colunas, é necessário e suficiente conhecer os desvios numa linha e em três colunas (portanto, ao todo, 3), para que os restantes desvios fiquem determinados. Exprime-se este facto dizendo que a tabela tem 3 *graus de liberdade*. Dum modo geral, uma tabela com *m* linhas e *n* colunas tem (m−1) (n−1) graus de liberdade.

2) O teste do χ^2 não deve aplicar-se quando algum dos valores esperados for *demasiado pequeno*.

3) A correcção de Yates é somente aplicável nas tábuas 2 × 2 (com 1 grau de liberdade).

Portanto, como nenhum dos valores esperados é demasiado pequeno e a tabela tem 3 graus de liberdade, podemos utilizar o anterior valor do χ^2 atrás obtido (15,7). Procurando na tábua da distribuição de Pearson o valor do χ^2 correspondente ao nível de probabilidade 0,01 e a 3 graus de liberdade, achamos 11,34. Ora, o χ^2 obtido é muito superior a este e aproxima-se até do que corresponde ao nível 0,001 (que é 16,27).

Podemos, pois, rejeitar a hipótese nula com nível altamente significante.

24. **Conceitos qualitativo e quantitativo de probabilidade.** Consideremos frases tais como:

'É provável que, já antes de Cristóvão Colombo, navegadores portugueses tenham chegado à América'.

'É provável que existam seres vivos em Marte', etc.

O primeiro juízo refere-se ao passado e o segundo ao presente; nenhum deles se baseia em frequências relativas observadas em sequências estatísticas. Trata-se, em ambos os casos, de um *conceito* *qualitativo de probabilidade* — muito diverso do conceito quantitativo atrás considerado.

O conceito quantitativo de probabilidade é *objectivo,* isto é, depende só do *objecto* (facto considerado) e não do *sujeito* (pessoa que pensa e emite o juízo), visto que se baseia em dados estatísticos.

Pelo contrário, o conceito qualitativo de probabilidade é *subjectivo,* isto é, depende do sujeito que o aplica: constitui uma *opinião,* que pode variar de pessoa para pessoa e de época para época.

Em vez da expressão 'é provável' também se usam, neste caso, com significado idêntico expressões tais como *'é verosímil', 'é de crer', 'é crível', 'é admissível', 'é natural'* e ainda *'é possível'* (agora com significado diferente do que temos atribuído a esta última palavra).

Por vezes diz-se mesmo, falando na primeira pessoa gramatical: *'não me repugna admitir', 'não excluo a hipótese',* etc., o que põe mais em evidência o carácter subjectivo do conceito em questão.

Podemos chamar *credibilidade* à probabilidade no sentido qualitativo. Embora se trate de um conceito que parece estar fora da alçada da ciência, há cientistas que pensam diversamente. Mas trata-se, ainda aqui, de uma opinião pessoal discutível...

Porém, a delimitação entre os conceitos qualitativo e quantitativo de probabilidade não é tão nítida como possa parecer à primeira vista.

Consideremos a seguinte afirmação:

'Com os progressos da técnica, a probabilidade de acidente numa viagem aérea tem diminuído nos últimos 30 anos, relativamente a percursos iguais. É, portanto, *provável* que essa probabilidade continue a diminuir'.

Neste caso, a palavra 'provável' sublinhada poderia substituir-se por 'admissível'. Refere-se a um conceito de probabilidade situado entre o qualitativo e o quantitativo, porém mais próximo do segundo, visto que a conclusão se baseia em dados estatísticos *objectivos,* de acordo com o *método de indução.* Mediante uma análise estatística aprofundada, essa probabilidade talvez pudesse mesmo vir a ser definida quantitativamente.

Muitas vezes, a avaliação de uma probabilidade começa por uma *suspeita* ou *intuição*, de carácter mais ou menos subjectivo. Por exemplo, alguns cientistas começaram a ter a suspeita de que o fumar contribui para se ter cancro nos pulmões. Depois disso, numerosas investigações têm sido efectuadas sobre o assunto, com a aplicação de testes tais como o do χ^2 (que implicam já, eles próprios, o conceito de probabilidade). Há pouco tempo, uma revista americana de divulgação indicava o seguinte resultado:

'*A probabilidade que tem um grande fumador de contrair cancro nas vias respiratórias é cerca de* 1/8'.

Considera-se aqui como 'grande fumador', ao que parece, um indivíduo que fume pelo menos 40 cigarros por dia. Estas investigações têm sido fortemente dificultadas pelo facto de haver enormes interesses económicos em causa. No entanto, a ser válida a conclusão anterior, o comércio de tabacos deveria ser equiparado ao negócio de estupefacientes, para efeitos legais...

Atitude análoga se deveria, aliás, adoptar em relação a muitas das *drogas* que inundam os mercados e que só têm servido para fazer grandes fortunas à custa da bolsa e da saúde, física ou mental, das pessoas incautas. A propaganda comercial não conhece limites morais para fazer vingar os seus fins e, assim, todos os sofismas lhe podem servir.

Sabe-se, por exemplo, que Puccini, fumador excessivo, morreu com um cancro na garganta. Pois não faltará quem justifique este caso, alegando que, se Puccini não tivesse fumado tanto, talvez não tivesse produzido a *Bohème*, a *Madame Butterfly*, etc. — além de que já tinha 66 anos quando faleceu...

Quanto àqueles que argumentam com o facto de haver muitos casos de não fumadores que contraem o cancro, esses apenas revelam ausência de espírito científico, semelhante ao das pessoas ingénuas que concluem, a partir de exemplos, que tem maior probabilidade de ganhar no totobola quem não conhecer nada de futebol.

Um facto é certo: os métodos estatísticos têm sido uma arma poderosa na luta contra o cancro e outras doenças que afligem a humanidade, assim como no combate à superstição e ao charlatanismo.

Estes e outros exemplos não vêm senão confirmar um princípio que deve estar presente no espírito de todo o cientista:

Embora partindo de intuições de carácter mais ou menos subjectivo, a investigação científica, para conduzir a resultados seguros, tem de ser sempre submetida a critérios de lógica dedutiva ou indutiva, tanto quanto possível objectivos, desinteressados, alheios a toda a paixão humana.

FIM DO 1.º VOLUME

TÁBUAS DE MORTALIDADE

TÁBUA AF

Idade x	N.º de vivos v_x	Esperança de vida: $\mathring{e}_x$	Idade x	N.º de vivos v_x	Esperança de vida: $\mathring{e}_x$
0	1 000 000	52,0187	26	791 780	37,4315
1	963 985	52,9434	27	786 713	36,6693
2	937 488	53,4257	28	781 578	35,9070
3	917 939	53,5528	29	776 368	35,1446
4	903 486	53,4015	30	771 075	34,3824
5	892 765	53,0368	31	765 690	33,6207
6	884 754	52,5125	32	760 203	32,8597
7	878 676	51,8723	33	754 606	32,0997
8	873 932	51,1511	34	748 887	31,3411
9	870 056	50,3768	35	746 036	30,5839
10	866 684	49,8708	36	737 038	29,8287
11	863 529	48,7501	37	730 884	29,0757
12	860 371	47,9272	38	724 556	28,3253
13	857 043	47,1114	39	718 042	26,5777
14	853 426	46,3090	40	711 324	26,8334
15	849 446	45,5236	41	704 386	26,0928
16	845 069	44,7568	42	697 210	25,3562
17	840 298	44,0081	43	689 777	24,6241
18	835 173	43,2750	44	682 067	23,8967
19	829 762	42,5540	45	674 058	23,1748
20	824 159	41,8399	46	665 729	22,4584
21	818 171	41,1272	47	657 056	21,7483
22	812 809	40,4102	48	648 015	21,0477
23	807 271	39,6840	49	638 581	20,3483
24	801 926	38,9451	50	628 727	19,6593
25	796 786	38,1932	51	618 429	18,9784

TÁBUAS DE MORTALIDADE

TÁBUA AF

(Continuação)

Idade x	N.º de vivos v_x	Esperança de vida: $\mathring{e}_x$	Idade x	N.º de vivos v_x	Esperança de vida: $\mathring{e}_x$
52	607 659	18,3059	78	143 530	5,1134
53	596 389	17,6424	79	124 896	4,8016
54	584 594	16,9882	80	107 354	4,5046
55	572 246	16,3440	81	91 047	4,2218
56	559 322	15,7101	82	76 094	3,9531
57	545 797	15,0870	83	62 588	3 6983
58	531 649	14,4752	84	50 588	3,4570
59	516 861	13,8751	85	40 118	3,2287
60	501 417	13,2870	86	31 159	3,0132
61	485 307	12,7115	87	23 658	2,81012
62	468 525	12,1489	88	17 523	2,61893
63	451 075	11,5995	89	12 632	2,43936
64	432 964	11,0638	90	8 841	2,27095
65	414 214	10,5420	91	5 992	2,11316
66	394 851	10,0345	92	3 920	1,96566
67	374 918	9,5414	93	2 468	1,82804
68	354 468	9,0630	94	1 490	1,69984
69	333 567	8,5995	95	859	1,58062
70	312 299	8,1511	96	471	1,47000
71	290 735	7,7179	97	245	1,36754
72	269 062	7,3000	98	120	1,27286
73	247 333	6,8974	99	55	1,18555
74	225 714	6,5101	100	23	1,10410
75	204 359	6,1382	101	9	1,03130
76	183 430	5,7815	102	3	0,96307
77	163 096	5,4399	103	1	0,89864

TÁBUAS DE MORTALIDADE

TÁBUA PM

Idade x	N.º de vivos v_x	Esperança de vida: $\mathring{e}_x$	Idade x	N.º de vivos v_x	Esperança de vida: $\mathring{e}_x$
0	1 000 000	66,9551	30	879 672	39,4565
1	931 380	65,4829	31	877 007	38,5748
2	924 674	64,9541	32	874 244	37,6952
3	922 002	64,1409	33	871 377	36,8176
4	920 268	63,2608	34	868 388	35,9426
5	918 851	62,3576	35	865 262	35,0706
6	917 638	61,4394	36	862 000	34,2014
7	916 574	60,5101	37	858 578	33,3357
8	915 630	59,5720	38	854 980	32,4739
9	914 787	58,6264	39	851 201	31,6159
10	914 010	57,6758	40	847 218	30,7622
11	913 260	56,7228	41	843 007	29,9133
12	912 502	55,7695	42	838 556	29,0695
13	911 699	54,8182	43	833 843	28,2309
14	910 815	53,8709	44	828 840	27,3983
15	909 813	52,9297	45	823 527	26,5719
16	908 667	51,9958	46	817 870	25,7522
17	907 349	51,0706	47	811 842	24,9397
18	905 861	50,1537	48	805 412	24,1348
19	904 194	49,2452	49	798 550	23,3379
20	902 359	48,3443	50	791 219	22,5495
21	900 383	47,4493	51	783 386	21,7700
22	898 294	46,5585	52	775 012	20,9998
23	896 165	45,6680	53	766 053	20,2395
24	893 996	44,7775	54	756 462	19,4898
25	891 770	43,8881	55	746 197	18,7510
26	889 487	42,9994	56	735 213	18,0237
27	887 139	42,1119	57	723 464	17,3083
28	884 726	41,2254	58	710 905	16,6052
29	882 240	40,3402	59	697 483	15,8152

TÁBUAS DE MORTALIDADE

TÁBUA PM

(Continuação)

Idade x	N.º de vivos v_x	Esperança de vida: $\mathring{e}_x$	Idade x	N.º de vivos v_x	Esperança de vida: $\mathring{e}_x$
60	683 156	15,2384	85	89 057	3,5776
61	667 881	14,5755	86	71 598	3,3280
62	651 611	13,9270	87	56 353	3,0931
63	634 311	13,2932	88	43 331	2,8723
64	615 948	12,6746	89	32 476	2,6653
65	596 502	12,0715	90	23 665	2,4715
66	575 953	11,4843	91	16 720	2,2904
67	554 297	10,9135	92	11 420	2,1213
68	531 549	10,3591	93	7 514	1,9641
69	507 730	9,8216	94	4 746	1,8180
70	482 882	9,3013	95	2 866	1,6825
71	457 076	8,7982	96	1 648	1,5564
72	430 397	8,3126	97	897	1,4409
73	402 963	7,8445	98	460	1,3348
74	374 909	7,3941	99	221	1,2376
75	346 409	6,9613	100	99	1,1465
76	317 657	6,5461	101	41	1,0610
77	288 874	6,1485	102	16	0,9375
78	260 310	5,7683	103	5	0,9000
79	232 225	5,4055	104	2	0,5000
80	204 897	5,0597	105	0	
81	178 609	4,7308			
82	153 637	4,4185			
83	130 241	4,1224			
84	108 651	3,8422			

TÁBUA DE DISTRIBUIÇÃO DO χ^2 DE PEARSON

	PROBABILIDADE (P)										
n	0,90	0,80	0,70	0,50	0,30	0,20	0,10	0,05	0,02	0,01	0,001
1	0,016	0,064	0,15	0,46	1,07	1,64	2,71	3,84	5,41	6,64	10,83
2	0,21	0,45	0,71	1,39	2,41	3,22	4,61	5,99	7,82	9,21	13,82
3	0,58	1,01	1,42	2,37	3,67	4,64	6,25	7,82	9,84	11,34	16,27
4	1,06	1,65	2,20	3,36	4,88	5,99	7,78	9,49	11,77	13,28	18,47
5	1,61	2,34	3,00	4,35	6,06	7,29	9,24	11,07	13,39	15,09	20,52
6	2,20	3,07	3,83	5,35	7,23	8,56	10,65	12,59	15,03	16,81	22,46
7	2,83	3,82	4,67	6,35	8,38	9,80	12,02	14,07	16,62	18,48	24,32
8	3,49	4,59	5.53	7,34	9,52	11,03	13,36	15,51	18,17	20,09	26,13
9	4,17	5,38	6,39	8,34	10,66	12,24	14,68	16,92	19,68	21,67	27,88
10	4,87	6,18	7,27	9,34	11,78	13,44	15,99	18,31	21,16	23,21	29,59
12	6,30	7,81	9,03	11,34	14,01	15,81	18,55	21,03	24,05	26,22	32,91
14	7,79	9,47	10,82	13,34	16,22	18,15	21,06	23,69	26,87	29,14	36,12
16	9,31	11,15	12,62	15,34	18,42	20,47	23,54	26,30	29,63	32,00	39,25
18	10,87	12,86	14,44	17,34	20,60	22,76	25,99	28,87	32,35	34,81	42,31
20	12,44	14,58	16,27	19,34	22,78	25,04	28,41	31,41	35,02	37,57	45,32
22	14,04	16,31	18,10	21,34	24,94	27,30	30,81	33,92	37,66	40,29	48,27
24	15,66	18.06	19,94	23,34	27,10	29,55	33,20	36,42	40,27	42,98	51,18
26	17,29	19,82	21,79	25,34	29,25	31,80	35,56	38,89	42,86	45,64	54,05
28	18,94	21,59	23,65	27,34	31,39	34,03	37,92	41,34	45,42	48,28	56,89
30	20,60	23,36	25,51	29,34	33,53	36,25	40,26	43,77	47,96	50,89	59,70

Índice

Composto e impresso na
Tipografia Guerra — Viseu
e concluiu-se
em Março de 1975

GABINETE DE ESTUDOS E PLANEAMENTO
DO MINISTÉRIO DA EDUCAÇÃO E CULTURA