TABELA N° 12

| N.º de animais tratados | $\chi^2$ 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 |

Muito mais haveria a dizer sobre testes de significância, mas, embora o assunto seja de grande interesse para agrónomos, não pode ser desenvolvido nesta cadeira. E, se nos alongámos um pouco a respeito do teste do $\chi^2$, foi sobretudo para dar uma ideia de como o Cálculo das Probabilidades pode intervir nas aplicações de carácter agronómico.

## B – Probabilidades

### 1. Lógica indutiva

Suponhamos que, em $n$ realizações, duma mesma prova $\mathcal{P}$, um dado acontecimento $\alpha$ se realizou $n$ vezes: então, a frequência relativa do acontecimento $\alpha$, na referida série de provas, será $f = n/n = 1$. Como já tivemos ocasião de observar, este facto não habilita a concluir que $\alpha$ seja um acontecimento certo. No entanto, se o número $n$ de provas realizadas for *muito grande*, é-se levado a admitir que o acontecimento é *praticamente certo*, embora não se possa concluir que seja *certo* (*em absoluto*). Nisto consiste, esquematicamente, a *indução* ou *raciocínio indutivo*, que se encontra na base de toda a ciência experimental. É bem sabido que as *leis físicas* ou, melhor, as *leis da Natureza*, têm carácter *contingente*: não se pode garantir que sejam absolutamente infalíveis.

Seja, por exemplo, a seguinte experiência: aproximar uma chama dum frasco, com a boca para baixo, no qual se tenha introduzido uma mistura de hidrogénio e oxigénio, em proporções convenientes. É, então, praticamente certo que se verifica o fenómeno "explosão".

Trata-se, aqui, duma lei química qualitativa, de cuja contingência nos apercebemos, sobretudo, se não tiverem sido definidas, com certa precisão, as condições em que a experiência deve ser realizada. Muitos outros exemplos poderíamos citar, de leis qualitativas, com o mesmo carácter contingente.

As leis quantitativas correspondem a um grau mais elevado de conhecimento, a uma fase mais avançada da ciência: mas nelas subsiste o carácter de contingência. Primeiro que tudo, devemos lembrar-nos de que as medidas das grandezas físicas *nunca* podem ser exactas (não faz, mesmo, sentido falar de medida exacta duma grandeza empírica – seja comprimento, seja velocidade, seja carga eléctrica, seja um pH, seja qualquer outra). Já daqui resulta uma certa margem de incerteza, isto é, de contingência.

Consideremos, por exemplo, esta experiência: aquecer um pedaço de chumbo a uma temperatura de cerca de 335 graus centígrados, em condições normais. É praticamente certo que se verifica o fenómeno *fusão*. Mas no próprio carácter aproximativo da temperatura indicada, bem como no significado da expressão "condições normais", há uma origem de incerteza, que pode ser reduzida (indicando, por exemplo, uma vizinhança de 335 na qual se prevê o acontecimento), mas nunca eliminada.

Recordemos, ainda, leis físicas, de nível mais elevado: a lei da gravitação, as leis da termodinâmica, as leis do electro-magnetismo, etc. Como é sabido, nenhuma destas leis se adapta exactamente à realidade: assentam todas em hipóteses simplificadoras, que só aproximadamente se podem realizar. Essas leis são, pois, apenas esquematizações duma realidade que é sempre demasiado complexa, demasiado mutável, para se deixar traduzir fielmente na simplicidade dos nossos símbolos.

Assim, a lei dos gases perfeitos deve o seu nome ao facto de se ter convencionado chamar "gases perfeitos" aos gases que a seguem rigorosamente, o que imprime a essa lei um certo carácter de definição ou de postulado. A verdade, porém, é que não há gases perfeitos: há, apenas, gases que se aproximam, mais ou menos, dessa forma ideal. Uma lei que substitui aquela, dando uma melhor aproximação da realidade, é a de Van der Waals; mas nem mesmo essa poderá ser exacta. Porque é lei metafísica das leis físicas, o serem contingentes.

O método experimental, e com ele o método matemático, tem-se estendido progressivamente, do âmbito restrito das ciências físico-químicas, ao das ciências biológicas, ao das ciências sociais, etc. Nestes novos domínios, mais acentuado se torna o carácter contingente das leis naturais. No entanto, algumas se apresentam com visos de certeza inabalável, como aquela que tem servido de premissa a um exemplo clássico de silogismo:

"Todos os homens são mortais".

Há aqui, de certo modo, uma lei biológica qualitativa. Mas se tentarmos precisá-la quantitativamente, afirmando, por exemplo:

"Todos os homens morrem antes dos 300 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 futuro?

O que pode dizer-se é que, nas condições actuais, *é extremamente improvável, praticamente impossível* que um ser humano atinja a idade de 300 anos. Um outro exemplo análogo é o que se refere a alturas: é praticamente impossível que um ser humano cresça até atingir a altura de 4 metros.

Se formos baixando estes limites, o grau de incerteza aumentará – e entraremos, abertamente, no campo das probabilidades. É de salientar que o Cálculo das Probabilidades e a Estatística 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 moderna tem carácter probabilista.

NOTA. Uma fonte primária de contingência está na impossibilidade que há em definir exactamente os conceitos relativos ao mundo empírico, – em delimitar com precisão os atributos dos entes naturais. Esta impossibilidade tem sido origem constante de especulações filosóficas através dos séculos. Platão resolvia a dificuldade, distinguindo duas formas de existência: a *realidade sensível* ou *mundo dos fenómenos*, que conhecemos por meio dos sentidos, e a *realidade inteligível* ou *mundo das Ideias,* que conhecemos por meio da razão e da qual a primeira é, apenas, imitação grosseira. Recorde-se, ainda, a querela que, na Idade Média, dividiu os filósofos em *realistas* e *nominalistas*, os primeiros proclamando a *realidade* dos universais (isto é, das classes, dos entes abstractos) e a sua supremacia sobre o contingente e o transitório; os segundos afirmando que os conceitos abstractos são, apenas, *nomes*, ficções cómodas, e que só os indivíduos existem.

No fundo, trata-se de duas tendências complementares, inerentes ao nosso espírito, com predomínio duma ou doutra conforme as pessoas e as situações. Assim, o matemático assume geralmente a atitude platónica (realista), reportando ao mundo dos seres ideais, de maneira mais ou menos consciente, os conceitos abstractos que são objecto do seu pensamento. Por sua vez, o experimentador tende para a atitude nominalista (ou empirista). Do equilíbrio das duas tendências, segundo o bom-senso, é que pode resultar o êxito da actividade intelectual: toda a teoria deve ser controlada pela experiência, assim como a prática deve sempre ser guiada pela teoria.

## 2. Lógica dedutiva

Ao raciocínio indutivo das ciências experimentais contrapõe a Matemática o raciocínio dedutivo. À contingência *física* das leis físicas opõe-se a certeza *matemática* das deduções matemáticas. Considerem-se, por exemplo, os seguintes teoremas:

"O quadrado dum número ímpar é sempre um número ímpar".

"Toda a equação de coeficientes complexos admite pelo menos uma solução no campo complexo".

"Toda a série de potências de $x$ é derivável, termo a termo (em ordem a $x$), no seu intervalo de convergência".

Encontramos aqui um carácter de certeza e precisão que contrasta vivamente com a natureza dúbia e aproximativa das leis naturais. Contudo, já se nota diferença, quando se passa da matemática pura para as matemáticas aplicadas. A própria Geometria é, sob certo aspecto, um ramo da Física. Consideremos, por exemplo, o bem conhecido teorema da geometria euclideana:

"A soma dos ângulos dum triângulo (plano) é sempre igual a um ângulo raso".

Dois métodos diferentes se podem seguir para estabelecer a veracidade desta proposição:

1.º – *Método indutivo ou experimental* – Medem-se os ângulos internos de vários triângulos e verifica-se que a soma dos ângulos de cada triângulo se *aproxima bastante* de 180°, tanto mais quanto mais perfeito for o traçado dos triângulos e mais precisa for a medição. É este o método seguido na primeira fase do ensino da Matemática no liceu.

2.º – *Método dedutivo ou racional* – Demonstra-se a proposição, *deduzindo-a logicamente* de proposições mais simples, cuja veracidade se considera *evidente*. Estas proposições evidentes são os chamados *axiomas* ou *postulados* da geometria euclideana. No teorema em questão, o postulado que intervém de maneira essencial é precisamente o postulado de Euclides:

"Dados um ponto e uma recta, existe sempre uma recta, e uma só, que passa pelo ponto dado e é paralela à recta dada".

Mas o declarar *evidentes* os postulados é uma atitude cómoda, que exige reflexão. Têm eles, de facto, carácter de certeza absoluta?

Um dos grandes progressos da história do pensamento, realizado no século passado, foi precisamente o de reconhecer que os postulados da geometria euclideana têm a natureza contingente e aproximativa das leis naturais. Nem leis se podem dizer, propriamente, mas sim hipóteses, cuja legitimidade se avalia indirectamente pelas suas consequências lógicas. Insinuam-se no nosso espírito, por um longo processo indutivo, dificilmente controlável, que se desenvolve na experiência quotidiana, no exercício constante da nossa actividade neuro-muscular. É esta uma forma de indução que se confunde com aquele processo directo de conhecimento a que se dá o nome de *intuição*.

Mas a nossa experiência quotidiana refere-se a uma região limitada do espaço. A teoria da relatividade veio mostrar, precisamente, que, nos espaços astronómicos, não é a geometria euclideana a que mais se aproxima da realidade: aí, a soma dos ângulos internos dum triângulo pode tornar-se sensivelmente superior a 180°.

O que se disse a respeito do teorema citado aplica-se a qualquer outro. A contingência dos postulados transmite-se a todos os teoremas. O que é matematicamente certo não é o teorema, mas, sim, o facto de o teorema ser consequência lógica dos postulados.

Entretanto registem-se as vantagens do método racional. Enquanto o método empírico exige um grande número de verificações fastidiosas, o método racional, com elegante simplicidade e perfeito rigor, reduz a veracidade do teorema à de proposições já reconhecidas como verdadeiras. Estas vantagens patenteiam-se, em particular, no Cálculo das Probabilidades.

## 3. Conceito natural de probabilidade

As considerações precedentes mostram que os termos "verdadeiro" e "falso", aplicados a proposições, se tornam insuficientes, quando, da realidade inteligível da Matemática, se passa à realidade sensível dos fenómenos. Os conceitos de "verdadeiro" e "falso" cedem, então, o lugar ao conceito de "probabilidade", correlativo do de "incerteza" ou "contingência": um facto dir-se-á tanto mais *provável* quanto menos contingente for.

O conceito de probabilidade, como todos os conceitos relativos ao mundo empírico, não é susceptível de definição lógica: gera-se no nosso espírito por um processo indutivo. Mais até, faz parte intrínseca do próprio mecanismo da indução. É o que vamos ver, tentando esclarecer este conceito.

Seja $\alpha$ um dos acontecimentos a prever numa certa experiência $\mathcal{P}$ e suponhamos que esta experiência foi efectuada em várias séries de provas, todas em *grande número*:

Série 1 – $\mathcal{P}_{1,1}, \mathcal{P}_{1,2}, \ldots\ldots\ldots, \mathcal{P}_{1,n_1}$ ($n_1$ provas)
Série 2 – $\mathcal{P}_{2,1}, \mathcal{P}_{2,2}, \ldots\ldots\ldots, \mathcal{P}_{2,n_2}$ ($n_2$ provas)
..............................................................................................
Série $m$ – $\mathcal{P}_{m,1}, \mathcal{P}_{m,2}, \ldots\ldots\ldots, \mathcal{P}_{m,n_m}$ ($n_m$ provas).

Se forem, respectivamente, $\nu_1, \nu_2, \ldots, \nu_m$ as frequências absolutas do acontecimento $\alpha$ nestas séries de provas, serão

$$\frac{\nu_1}{n_1}, \frac{\nu_2}{n_2}, \ldots, \frac{\nu_m}{n_m},$$

as correspondentes frequências relativas. Suponhamos, ainda, que se verificou uma sensível concordância entre estas frequências, isto é, que todas se localizaram num pequeno intervalo $]f_0-\varepsilon, f_0+\varepsilon[$:

$$f_0 - \varepsilon < \frac{\nu_i}{n_i} < f_0+\varepsilon, \quad \text{para } i = 1, 2, \ldots, m.$$

*Aplicando o raciocínio indutivo*, exactamente como se faz ao estabelecer as leis naturais, seremos levados a admitir, como praticamente certo, que:

*Em toda a série formada por muitas realizações de $\mathcal{P}$ (em número não inferior ao maior dos números $n_1, n_2, \ldots, n_m$) a frequência relativa de $\alpha$ ficará situada entre $f_0 - \varepsilon$ e $f_0 + \varepsilon$.*

É, então, natural dizer que $f_0$ é um *valor aproximado da probabilidade de $\alpha$, com erro inferior a $\varepsilon$ (ou a menos de $\varepsilon$).*

Por exemplo, se em tais séries de provas se registou sempre uma frequência relativa de $\alpha$ entre 7% e 11% (ou seja, entre 0,09 − 0,02 e 0,09 + 0,02), dir-se-á que 0,09 é um valor aproximado, a menos de 0,02, da probabilidade de $\alpha$.

Suponhamos, agora, que se efectuaram novas séries de realizações de $\mathcal{P}$, em *números bastante maiores* que os anteriores, e que passou a registar-se uma frequência relativa de $\alpha$, entre

$$f_1 - \frac{\varepsilon}{10} \text{ e } f_1 + \frac{\varepsilon}{10};$$

dir-se-á, então, que $f_1$ é valor aproximado da probabilidade de $\alpha$, a menos de $\varepsilon/10$. Analogamente, poderíamos imaginar valores $f_2, f_3, \ldots$, aproximados da probabilidade de $\alpha$ a menos de $\varepsilon/10^2$, de $\varepsilon/10^3$, etc.

A probabilidade aparece-nos, assim, como *grandeza que se mede*, tal como se fosse uma grandeza física (uma velocidade, um ponto de fusão, uma densidade, etc.), com o carácter contingente e aproximativo que revestem todas as medições empíricas. Já as expressões "grandes números", "números bastante maiores", atrás usadas, envolvem a aplicação de critérios subjectivos, humanos, a que falta rigor matemático, mas que são inevitáveis.

Tal como sucede com as grandezas físicas, é-se levado a idealizar, para cada acontecimento $\alpha$ (em certos tipos de experiências), um valor *exacto*, $p$, da probabilidade de $\alpha$, que seria, por assim dizer, o *limite* da sucessão dos valores aproximados $f_0, f_1, f_2, \ldots$ Mas é claro que um tal valor exacto *não existe nos casos concretos* – assim como não existe o valor exacto duma grandeza física: trata-se apenas de convenções cómodas, que o nosso espírito transporta para o mundo platónico dos entes ideais, para sobre as mesmas poder construir uma teoria racional – *uma teoria matemática*[1].

Note-se, ainda, que a unidade adoptada para medir a probabilidade do acontecimento $\alpha$ é a probabilidade do acontecimento certo – isto é, a *certeza*[2]. Por sua vez, ao acontecimento impossível é atribuída a probabilidade 0. Deste modo, a probabilidade $p$ de $\alpha$ deverá ser um número tal que

$$0 \leq p \leq 1 .$$

Se $p$ for aproximadamente igual a 1, dir-se-á que $\alpha$ é *praticamente certo*. Se $p$ for aproximadamente igual a 0, dir-se-á que $\alpha$ é *praticamente impossível*. Mas o grau de aproximação a partir do qual se dirá uma ou outra coisa é questão manifestamente subjectiva. Registe-se, entretanto, que, para um acontecimento ser considerado praticamente certo, não é necessário que se realize em *todas* as provas que tenham sido efectuadas: admite-se a possibilidade de excepções, contanto que estas sejam extremamente raras.

E, analogamente, no caso oposto.

---

(1) – Esta concepção da probabilidade como constante física dos acontecimentos é devida ao matemático francês MAURICE FRÉCHET.

(2) – Recorde-se que, na medida dos arcos, se pode tomar unidade a circunferência. Uma das maneiras sugestivas de indicar graficamente as probabilidades (ou as frequências relativas) de vários acontecimentos incompatíveis, consiste em representá-los por meio de sectores dum círculo, de amplitudes proporcionais a essas probabilidades (ou frequências).

O termo "probabilidade", tal como acabamos de o precisar, é usado na prática com maior ou menor propriedade. Recordemos o exemplo das tábuas de mortalidade, sobre as quais os actuários das companhias de seguros baseiam o cálculo dos prémios a pagar em seguros de vida. Essas tábuas são, no fundo, estatísticas, isto é, tabelas de frequência, relativas a populações numerosas. Assim, por exemplo, a tábua de mortalidade alemã para homens, correspondente ao decénio 1901-1910, indica que, entre 10.000 crianças com menos de 1 ano, 3.608 chegaram à idade de 65 anos. Deste modo, a probabilidade de que um recém-nascido venha a atingir os 65 anos será, aproximadamente:

$$\frac{3.608}{10.000} \approx 36\%.$$

Por sua vez, a mesma tábua mostra que, entre 10.000 crianças com menos de 1 ano, 7.065 chegam aos 20 anos. Daqui se deduz que a probabilidade de um rapaz de 20 anos chegar aos 65 anos é

$$\frac{3.608}{7.065} \approx 51\%,$$

maior que a precedente.

A mesma tábua dá para crianças de menos de 1 ano, a probabilidade 0,68% de chegar aos 90 anos, a probabilidade 0,0038% de atingir os 100 anos, etc.

Mas note-se que as tábuas de mortalidade têm de ser revistas de tempos a tempos e não devem ser as mesmas para todas as regiões, para todas as raças, para os dois sexos, etc.

A probabilidade dum acontecimento pode, assim, estar sujeita a uma evolução imprevisível no tempo e no espaço. É este um facto que, pelo menos aparentemente, está em contradição com o conceito de probabilidade, tal como foi atrás explanado. O mesmo é dizer que as leis físicas evoluem de maneira imprevisível no tempo e no espaço – *deixando, portanto, de ser leis*. A explicação do facto está, ainda, no carácter contingente do raciocínio indutivo, que impõe sempre a seguinte norma prudencial ao experimentador: *as induções* *só merecem confiança em regiões limitadas do espaço e do tempo; não podem ser extrapoladas muito para além do domínio das experiências efectuadas.*

Portanto, só quando a evolução for lenta será lícito falar de probabilidades.

Recordemos, ainda, o exemplo dos desastres de avião: é hoje menor do que há vinte anos a frequência relativa dos desastres de aviação em percursos iguais. Neste caso, embora seja mais rápida a evolução, ainda, de certo modo, é lícito falar de probabilidade no sentido atrás considerado(1).

Casos há, porém, em que o emprego desta palavra escapa, de todo, a critérios quantitativos, que lhe confiram interesse científico.

## 4. Axiomatização do conceito de probabilidade

Viu-se no número anterior que, para fundar uma teoria matemática das probabilidades, se idealiza para cada acontecimento, em certos tipos de experiências, um valor exacto da sua probabilidade. Contudo, seja esse valor uma convenção ou uma realidade, o que interessa para o desenvolvimento lógico da teoria não é, propriamente, o conceito em si, mas, antes, o conjunto das suas propriedades formais que intervêm nos cálculos e nos raciocínios. O que há a fazer, então, é escolher, entre essas propriedades, algumas mais simples, das quais se possam deduzir logicamente todas as outras. As propriedades escolhidas dir-se-ão *propriedades fundamentais* ou *axiomas* e o seu conjunto terá o nome de *axiomática.* Já em Matemáticas Gerais foi estudada uma axiomática dos grupos contínuos de grandezas. Trata-se, agora, de estabelecer uma axiomática do conceito de probabilidade, no caso dos corpos finitos de acontecimentos, considerando aquele conceito como *primitivo*, isto é, não definível logicamente à custa de outros.

Seja, pois, $\mathcal{R}$ um corpo finito de eventualidades a considerar numa dada prova $\mathcal{P}$, e suponhamos que a cada eventualidade $\alpha \in \mathcal{R}$, corresponde uma determinada probabilidade, que designaremos pelo

(1) – Nos seguros relativos a acidentes de aviação, as companhias americanas fazem hoje o cálculo dos prémios como se houvesse morte de um passageiro em 20.000 viagens de ida e volta.

símbolo $Pr(\alpha)$. Tomaremos como fundamentais as propriedades seguintes[1]:

AXIOMA 1. *A probabilidade de* $\alpha$ *é sempre um número real não negativo, isto é*:

$$Pr(\alpha) \geq 0, \textit{ qualquer que seja } \alpha \in \mathcal{R}.$$

AXIOMA 2. *Se* $\alpha$ *e* $\beta$ *são acontecimentos incompatíveis de* $\mathcal{R}$, *tem-se*:

$$Pr(\alpha + \beta) = Pr(\alpha) + Pr(\beta).$$

(Convém aqui rever, cuidadosamente, as noções do n.° 2 e do n.° 3 relativas às álgebras de atributos e de acontecimentos).

AXIOMA 3. *A probabilidade do acontecimento certo é* 1; *isto é, tem-se*:

$$Pr(\alpha) = 1, \textit{ se } \alpha \textit{ é certo}.$$

Estas propriedades concordam com o conceito empírico de probabilidade, tal como foi atrás introduzido. Sendo as probabilidades, por assim dizer, limites de frequências relativas, é natural que se conservem as propriedades das frequências relativas que não mudam na passagem ao limite.

Os axiomas 1 e 2 dizem-nos, simplesmente, que a função $Pr(\alpha)$ é uma distribuição definida no corpo $\mathcal{R}$. O axioma 3 acrescenta que essa distribuição é relativa. Diremos, então, que se trata duma *distribuição de probabilidade* definida em $\mathcal{R}$. Mas só nas aplicações concretas será possível distinguir uma distribuição de probabilidade duma distribuição de frequência, pois que a anterior axiomática se limita a dar os caracteres duma distribuição abstracta.

---

(1) – Vários autores consideram, ainda, como axioma a propriedade relativa à probabilidade do produto. Mas essa propriedade pode considerar-se como definição ou como consequência de definição de probabilidade condicional, como veremos.

Da anterior axiomática deduzem-se, desde logo, as seguintes consequências (que se estendem a qualquer distribuição):

TEOREMA 1. *Se* $\alpha$ *implica* $\beta$, *então* $\Pr(\alpha) \leq \Pr(\beta)$.

Com efeito, se $\alpha \subset \beta$, tem-se $\beta = \alpha + \tilde{\alpha}\beta$ e, como $\alpha$ e $\tilde{\alpha}\beta$ são incompatíveis, vem, pelo axioma 2,

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

donde se conclui, pelo axioma 1, que $\Pr(\beta) \geq \Pr(\alpha)$.

COROLÁRIO. *Qualquer que seja* $\alpha \in \mathcal{R}$, *tem-se:*

$$\Pr(\alpha) \leq 1.$$

Com efeito, todo o acontecimento $\alpha$ de $\mathcal{R}$ implica o acontecimento certo: isto é, em fórmula:

$$\alpha \subset I,$$

donde, pelo teorema 1 e pelo axioma 3:

$$\Pr(\alpha) \leq 1.$$

TEOREMA 2[1]. *Se* $\alpha_1, \alpha_2, \ldots, \alpha_n$ *são acontecimentos (de* $\mathcal{R}$*) incompatíveis dois a dois, a probabilidade de que se realize um dos acontecimentos* $\alpha_1, \alpha_2, \ldots, \alpha_n$ *é a soma das probabilidades destes acontecimentos, isto é, tem-se*:

$$\Pr(\sum \alpha_i) = \sum \Pr(\alpha_i), \textit{ se } \alpha_i\alpha_k = \varnothing \text{ para } i \neq k.$$

Basta aplicar, repetidamente, o axioma 2, observando que é $\alpha_1 + \alpha_2 + \alpha_3 = (\alpha_1 + \alpha_2) + \alpha_3$, $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = (\alpha_1 + \alpha_2 + \alpha_3) + \alpha_4$, etc.,

(1) – Também conhecido por *princípio das probabilidades totais.*

e que, se $\alpha_1, \alpha_2, \ldots, \alpha_n$ são incompatíveis dois a dois, também $\alpha_3$ é incompatível com $\alpha_1 + \alpha_2$, $\alpha_4$ com $\alpha_1 + \alpha_2 + \alpha_3$, etc.

TEOREMA 3. *Quaisquer que sejam* $\alpha, \beta \in \mathcal{R}$, *tem-se*

$$\Pr(\alpha + \beta) = \Pr(\alpha) + \Pr(\beta) - \Pr(\alpha\beta).$$

Basta notar que $\alpha + \beta = \alpha\tilde{\beta} + \tilde{\alpha}\beta + \alpha\beta$ e aplicar o teorema 2, notando que $\Pr(\alpha) = \Pr(\alpha\tilde{\beta}) + \Pr(\alpha\beta)$, $\Pr(\beta) = \Pr(\tilde{\alpha}\beta) + \Pr(\alpha\beta)$. O teorema pode generalizar-se ao caso de vários acontecimentos, obtendo-se a fórmula de DANIEL DA SILVA (cf. A-10).

TEOREMA 4. *A probabilidade do acontecimento contrário de* $\alpha$ *é a diferença para* 1 *da probabilidade de* $\alpha$; *isto é*:

$$\Pr(\tilde{\alpha}) = 1 - \Pr(\alpha).$$

Consequência imediata dos axiomas 2, 3 e da definição de contrário ($\alpha + \tilde{\alpha} = I$, $\alpha\tilde{\alpha} = \varnothing$), conforme já se viu para as distribuições em geral.

COROLÁRIO. *Se* $\alpha$ *é acontecimento impossível, tem-se*

$$\Pr(\alpha) = 0.$$

Basta lembrar que o acontecimento impossível é contrário do acontecimento certo e aplicar o teorema 4 com o axioma 3.

É claro que, em virtude do teorema 2, uma distribuição de probabilidade num corpo finito $\mathcal{R}$ fica determinada, logo que se conheçam as probabilidades das células de $\mathcal{R}$: a probabilidade dum acontecimento $\alpha \in \mathcal{R}$ será a soma das probabilidades dos acontecimentos celulares de que $\alpha$ é a soma. Tudo está, pois, em determinar as probabilidades das células.

## 5. Alguns exemplos de cálculo de probabilidades *a priori*

Por vezes, a distribuição de probabilidade num corpo finito pode ser determinada *a priori* (isto é, antes de qualquer experiência deliberada) por simples considerações ditadas pelo *senso-comum.*

Um dos exemplos clássicos, já atrás invocados, é o que se refere à extracção casual de bolas ou de cartões duma caixa. Consideremos o caso duma urna com $N$ bolas, que designaremos por $x_1, x_2, \ldots, x_N$. Seja $U$ o conjunto destas bolas e seja $\mathscr{P}$ a experiência que consiste em tirar *ao acaso* uma bola da urna.

Os acontecimentos elementares que se podem verificar nesta experiência são: "sair a bola $x_1$", "sair a bola $x_2$", ..., "sair a bola $x_N$"; ou, em notação proposicional:

$$x = x_1, x = x_2, \ldots, x = x_N.$$

(Como se disse atrás, $x$ é neste caso uma *variável casual*, que toma, em cada prova, um e um só dos valores $x_1, x_2, \ldots, x_N$).

Estes acontecimentos elementares geram um corpo de que são as células. Os acontecimentos que formam o corpo serão, além das células e do acontecimento impossível, todos aqueles que se exprimem como somas lógicas de células (1), por exemplo:

$$(x = x_1) + (x = x_2) \text{ (sair } x_1 \text{ ou } x_2)$$

$$(x = x_1) + (x = x_3) + (x = x_5) \text{ (sair } x_1 \text{ ou } x_3 \text{ ou } x_5\text{), etc.}$$

Entre estes está incluído o acontecimento certo:

$$\sum_{i=1}^{p} (x = x_i) \text{ (sair uma das bolas } x_1, x_2, \ldots, x_N).$$

É claro, agora, que se conhecermos as probabilidades dos acontecimentos elementares, estamos aptos a conhecer a probabilidade de qualquer outro acontecimento do corpo, por simples aplicação do teorema 2. Designemos por $p_1, p_2, \ldots, p_N$, respectivamente, as probabilidades de "sair $x_1$", de "sair $x_2$", ..., de "sair $x_N$", isto é, ponhamos

$$p_i = \Pr(x = x_i), \quad \text{para } i = 1, 2, \ldots, N.$$

---

(1) – É evidente que os acontecimentos considerados são incompatíveis dois a dois, uma vez estabelecido que se tira *uma só* bola de cada vez.

Então, a probabilidade de "sair $x_1$ ou $x_2$" será

$$\Pr[(x = x_1) + (x = x_2)] = p_1 + p_2,$$

a probabilidade de "sair $x_1$ ou $x_3$ ou $x_5$" será $p_1 + p_3 + p_5$.

Como a probabilidade do acontecimento certo é igual à soma das probabilidades de todas as células, deverá ter-se

$$p_1 + p_2 + \cdots + p_N = 1.$$

Suponhamos que as bolas $x_1, x_2, \ldots x_\nu$ são vermelhas e as restantes brancas. Designando por $V$ o conjunto das bolas vermelhas e por $B$ o conjunto das bolas brancas, a probabilidade de sair bola vermelha será

$$\Pr(x \in V) = p_1 + p_2 + \cdots + p_\nu$$

e a de sair bola branca

$$\Pr(x \in B) = p_{\nu+1} + p_{\nu+2} + \cdots + p_N = 1 - \Pr(x \in V)$$

Mas como determinar as probabilidades elementares $p_1, p_2, \ldots, p_N$, se é que existem?

Para isso, há que introduzir hipóteses suplementares. *Suponhamos, por exemplo, que as bolas são sensivelmente iguais, em substância, forma e dimensões*. Então, para dar a todas as bolas igual probabilidade de serem extraídas (*casualização*), ocorre a ideia de fechar a urna e agitá-la várias vezes antes de tirar a bola.

Esta ideia, ou melhor, esta intuição, é confirmada pela experiência: efectuando *muitas vezes* a prova nas condições indicadas (com reposição da bola retirada), serão sensivelmente iguais as frequências relativas com que aparecem as diferentes bolas[1].

---

(1) – Donde nos vem esta intuição? É, na verdade, anterior a qualquer esperiência? Não parece que o seja. Será, antes, o produto de inúmeras experiências que fazemos, sem dar por isso, no decurso da nossa existência. Qualquer coisa de semelhante à intuição que nos leva a admitir os postulados da geometria euclideana.

Uma vez admitido que todas as bolas têm a mesma probabilidade de saída, isto é, que:

$$p_1 = p_2 = \cdots = p_N,$$

como se tem $p_1 + p_2 + \cdots + p_N = 1$, será

$$p_i = \frac{1}{N}$$

a probabilidade de saída de qualquer bola. Então, a probabilidade de sair bola vermelha, será

$$\Pr(x \in V) = \frac{1}{N} + \frac{1}{N} + \cdots + \frac{1}{N} \quad (v \text{ vezes}) = \frac{v}{N}.$$

Este último resultado corresponde directamente à definição clássica de probabilidade:

"A probabilidade dum acontecimento é o quociente do número de casos favoráveis ao acontecimento (1) pelo número total de casós possíveis, supondo que estes são todos igualmente prováveis".

Segundo o ponto de vista adoptado nesta definição, considera-se como primitivo, não o conceito de probabilidade, mas, sim, o de "igualmente provável".

É claro que a hipótese de as bolas serem iguais (em substância, forma e dimensões) nunca se realiza exactamente na prática: podem ser, apenas, *aproximadamente iguais* e, então, as respectivas probabilidades, $p_1, p_2, \ldots, p_N$, serão aproximadamente iguais, com um grau de aproximação correspondente. Se as bolas diferem entre si de maneira sensível, as probabilidades elementares $p_1$, $p_2, \ldots$, $p_N$ serão, também, sensivelmente diferentes e só poderão ser avaliadas *a posteriori,* efectuando numerosas extracções com reposição e registando as frequências relativas dos acontecimentos $x = x_1$, $x = x_2$, ..., $x = x_N$. Por comodidade, em tudo o que segue, ao tratar de problemas de bolas numa urna, supomos verificada a hipótese da igualdade.

---

(1) – Aqui a palavra "caso" é empregada na acepção de acontecimento celular. Chamam-se "casos favoráveis ao acontecimento" os casos que *implicam* o acontecimento.

Recordemos que a extracção casual de bolas com números constitui a base de certos jogos, como o loto ou a lotaria. O cálculo das probabilidades nasceu precisamente das reflexões de certos matemáticos, nomeadamente PASCAL, sobre questões relativas a *jogos de azar*[1].

Um jogo de azar conhecido desde a antiguidade é o dos dados. Em princípio, um dado deve ser um cubo formado de substância homogénea (*dado perfeito*), mas é claro que, na prática, estas condições só aproximadamente se realizam. As faces do cubo estão numeradas de 1 a 6. Admite-se que, lançando o dado, depois de o agitar dentro dum copo (*casualização*), todas as faces têm igual probabilidade de se apresentar em cima e será, pois, 1/6 a probabilidade correspondente a cada face. Esta hipótese é, geralmente, confirmada pela experiência: quando não houver confirmação, a experiência vem, apenas, revelar imperfeições que tenham passado despercebidas.

Recordemos, ainda, os jogos de cartas, alguns dos quais são puramente de azar. Nestes jogos, a casualização consiste em *baralhar* as cartas.

É, ainda, de citar o jogo da roleta – círculo dividido em sectores numerados, que roda em torno dum eixo que passa pelo centro. A casualização consiste em imprimir à roleta um impulso que a faça dar várias voltas: se os sectores foram iguais, a probabilidade de parar num dado ponto é a mesma para todos. Se os sectores forem diferentes, as probabilidades serão proporcionais às respectivas amplitudes, de modo que , se foram $\Theta_1, \Theta_2, \ldots, \Theta_N$, essas amplitudes em radianos, serão

$$p_1 = \frac{\Theta_1}{2\pi}, p_2 = \frac{\Theta_2}{2\pi}, \ldots, p_N = \frac{\Theta_N}{2\pi},$$

as respectivas probabilidades, pois que se deve ter

$$p_1 + p_2 + \cdots + p_N = 1.$$

(Este exemplo conduz, naturalmente, à consideração de probabilidade no contínuo).

---

(1) – Aqui ''azar'' não é usado como sinónimo de ''pouca sorte'', mas, simplesmente, como sinónimo de ''acaso'' (tal como o francês ''hasard'').

Recordemos, por último, o mais simples de todos os jogos de azar: o lançamento duma moeda ao ar. Se a moeda for bem *balançada*, as duas eventualidades "sair coroa" e "sair face" terão igual probabilidade, ou seja, 1/2. Se não, haverá uma diferença de probabilidade que a experiência acusará.

Para ilustrar as considerações precedentes, convém apresentar aqui alguns exemplos numéricos, que, no fundo, se reduzem geralmente a problemas de Cálculo Combinatório.

Exemplo 1 – *Determinar a probabilidade de que, no lançamento dum dado perfeito, se obtenha um múltiplo de* 3.

São 2 os casos favoráveis ao acontecimento: $x = 3$, $x = 6$. Visto haver ao todo 6 casos possíveis (igualmente prováveis), a probabilidade pedida será

$$\frac{2}{6} = \frac{1}{3}.$$

Exemplo 2 – *Determinar a probabilidade de que, em dois lançamentos sucessivos dum dado perfeito, se obtenham* 2 *números pares.*

As células, neste caso, podem ser representadas pelos pares ordenados $(x, y)$, em que $x$ designa o número saído no 1.º lançamento e $y$ o número saído no 2.º lançamento. Estes pares são em número de $6^2 = 36$ e todos igualmente prováveis, como é evidente. Os casos que implicam o acontecimento "saída de dois números pares" são: (2,2), (2,4), (2,6), (4,2), (4,4), (4,6), (6,2), (6,4), (6,6), em número de 9. A probabilidade pedida será, pois,

$$\frac{9}{36} = \frac{1}{4}.$$

NOTA. Como se disse em Matemáticas Gerais, chama-se produto cartesiano dum conjunto $A$ por um conjunto $B$, e designa-se por $A \times B$, o conjunto de todos os pares ordenados $(a, b)$ que se obtém tomando um elemento $a$ em $A$ e um elemento $b$ em $B$; o número de elementos de $A \times B$ será, então, o produto dos números de elementos de $A$ e de $B$. Analogamente, chama-se produto cartesiano de 3 conjuntos $A$, $B$, $C$, na ordem por que estão escritos, e designa-se por $A \times B \times C$, o conjunto de todos os ternos ordenados $(a, b, c)$ com $a \in A$, $b \in B$, $c \in C$; o número destes ternos é igual ao produto dos números de elementos de $A$, $B$, e $C$. E assim por diante.

Também convencionámos escrever $A^2$ como abreviatura de $A \times A$, $A^3$ como abreviatura de $A \times A \times A$, etc.

Deste modo, os casos possíveis, no exemplo 2, serão os elementos do produto cartesiano

$$\{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\} = \{1, 2, 3, 4, 5, 6\}^2$$

e os casos que implicam o acontecimento considerado serão os elementos do produto cartesiano $\{2, 4, 6\}^2$.

Note-se que os pares ordenados, os ternos ordenados, etc., também se chamam *arranjos com repetição* (*dois a dois, três a três*, etc.).

Exemplo 3 – *Determinar a probabilidade de que, em dois lançamentos sucessivos dum dado perfeito, saia pelo menos uma vez número ímpar.*

O acontecimento "sair pelo menos uma vez número ímpar" é o contrário do acontecimento "sair duas vezes número par" cuja probabilidade, há pouco calculada, é 1/4. A probabilidade pedida será, pois,

$$1 - \frac{1}{4} = \frac{3}{4}.$$

Exemplo 4 – *Determinar a probabilidade de que, em dois lançamentos sucessivos dum dado perfeito, saia primeiro um número par e depois um múltiplo de* 3.

Os casos possíveis (células) são os mesmos que no exemplo 2 e no exemplo 3. Os casos favoráveis ao acontecimento considerado obtêm-se desenvolvendo o produto cartesiano

$$\{2, 4, 6\} \times \{3, 6\},$$

cujos elementos são em número de $3 \times 2 = 6$. A probabilidade pedida será, pois, $6/36 = 1/6$.

Exemplo 5 – *Determinemos a probabilidade de que, em dois lançamentos sucessivos dum dado perfeito, saia uma vez um número par e outra vez um múltiplo de* 3.

O acontecimento considerado é a soma lógica dos dois seguintes: "sair primeiro número par e depois um múltiplo de 3" e "sair primeiro um múltiplo de 3 e depois um número par". Estes dois acontecimentos, que designaremos por $\alpha_1$ e $\alpha_2$, têm ambos a probabilidade 1/6 (exemplo 4), *mas não são incompatíveis*. O acontecimento $\alpha_1\alpha_2$ equivale a "sair 6 duas vezes"; como se trata de um só caso entre 36 possíveis, a sua probabilidade é de 1/36. A probabilidade pedida será, pois (teorema 3):

$$\Pr(\alpha_1) + \Pr(\alpha_2) - \Pr(\alpha_1\alpha_2) = 2 \cdot \frac{1}{6} - \frac{1}{36} = \frac{11}{36}.$$

NOTA. Nos exemplos anteriores, excepto o primeiro, as considerações ficam essencialmente as mesmas, se substituirmos a expressão "em dois lançamentos sucessivos dum dado perfeito" pela expressão "num lançamento simultâneo de dois dados perfeitos".

Exemplo 6 – *Duma urna que contem* 12 *bolas, das quais* 4 *são brancas e* 8 *pretas, tiram-se duas bolas à sorte, uma após a outra, sem reposição. Calcular a probabilidade de que*: a) *sejam ambas brancas*; b) *sejam ambas pretas*; c) *sejam da mesma cor*; d) *sejam de cores diferentes*; e) *sejam a 1.ª branca e a 2.ª preta*; f) *sejam a 2.ª preta e a 1.ª branca.*

Suponhamos as bolas numeradas de 1 a 12, sendo as 4 primeiras brancas; designemos por $x$ o primeiro número saído e por $y$ o segundo. Neste caso, os acontecimentos celulares são representados pelos pares ordenados, $(x, y)$, com $x \neq y$; trata-se, portanto, de arranjos (sem repetição!) de 12 elementos 2 a 2; o número desses arranjos é $12 \times 11 = 132$.

O acontecimento considerado em a) corresponde aos pares $(x, y)$, em que $x$ e $y$ variam de 1 a 4, mas sendo $x \neq y$. Trata-se, portanto, de arranjos de 4 elementos 2 a 2, cujo número é $4 \times 3 = 12$. A probabilidade pedida será, pois, 12/132.

Analogamente, se reconhece que a probabilidade pedida em b) é

$$\frac{8 \times 7}{132} = \frac{56}{132} .$$

O acontecimento considerado em c) é a soma lógica dos acontecimentos anteriores, que são incompatíveis. Portanto, a probabilidade pedida em c) será

$$\frac{12}{132} + \frac{56}{132} = \frac{68}{132} .$$

O acontecimento considerado em d) é o contrário do anterior. Logo, a sua probabilidade será

$$1 - \frac{68}{132} = \frac{64}{132} .$$

É fácil reconhecer, finalmente, que as probabilidades pedidas em e) e f) são ambas iguais a

$$\frac{8 \times 4}{132} = \frac{32}{132} ;$$

a sua soma dá a anterior, como era de esperar.

Exemplo 7 – *Duma urna que contem N bolas, sendo a brancas e b pretas, extrai-se, ao acaso, uma amostra de n bolas (tiradas, por exemplo, uma após outra, sem reposição). Achar a probabilidade de que, na amostra obtida, haja* $\nu$ *bolas brancas e* $n-\nu$ bolas pretas.

Comecemos por notar que a ordem pela qual são tiradas as bolas não interessa à questão. Os casos possíveis (todos igualmente prováveis) serão, pois, as combinações das $N$ bolas tomadas $n$ a $n$, cujo número é $\binom{N}{n}$. Por sua vez, as amostras com $\nu$ bolas brancas e $n-\nu$ bolas pretas podem obter-se, sem omissão nem repetição, do seguinte modo:

1) – Formando, por um lado, as combinações das $a$ bolas brancas tomadas $\nu$ a $\nu_1$ cujo número é $\binom{a}{\nu}$.

2) – Formando, por outro lado, as combinações das $b$ bolas pretas tomadas $n-\nu$ a $n-\nu$, cujo número é $\binom{b}{n-\nu}$.

3) – Arranjando todos os possíveis pares $A_\nu\ B_{n-\nu}$ constituídos por uma combinação $A_\nu$ obtida em 1) e uma combinação $B_{n-\nu}$ obtida em 2). Como o número destes pares é $\binom{a}{\nu}\binom{b}{n-\nu}$, a probabilidade pedida será

$$\frac{\binom{a}{\nu}\binom{b}{n-\nu}}{\binom{N}{n}}.$$

Esta distribuição de variável casual $\nu$ é chamada *distribuição hipergeométrica.*

Exemplo 8 – *Duma urna que contem N bolas, sendo a brancas e b pretas, tiram-se, ao acaso, sucessivamente, n bolas, repondo a bola retirada após cada extracção. Achar a probabilidade de que saia* $\nu$ *vezes bola branca* e $n-\nu$ *vezes bola preta.*

Discorrendo como há pouco, seríamos levados a considerar, agora, os casos possíveis sob a forma de *combinações com repetição*. Porém, esses casos *já não são igualmente prováveis, como é fácil ver*. Suponhamos, por exemplo, $n=3$; neste caso, a combinação $x_1x_2x_3$ será mais provável que a combinação $x_1x_1x_3$, pois que a primeira se pode apresentar com as seguintes ordens de saída de bolas:

$$x_1x_2x_3,\ x_1x_3x_2,\ x_2x_1x_3,\ x_2x_3x_1,\ x_3x_1x_2,\ x_3x_2x_1,$$

ao passo que a segunda se pode apresentar só dos seguintes modos:

$$x_1x_1x_3,\ x_1x_3x_1,\ x_3x_1x_1.$$

Este exemplo mostra, bem, o cuidado que é necessário ter na escolha dos acontecimentos celulares, para que sejam igualmente prováveis. Neste caso, há que tomar para células os arranjos com repetição das $N$ bolas $n$ a $n$ (isto é, sistema de $n$ bolas).

Veremos, mais adiante, como se resolve este problema, considerado sob um aspecto mais geral (distribuição de BERNOULLI).

Exemplo 9 – *Nas condições do exemplo* 7, *achar a probabilidade de que, numa amostra, o número* $\nu$ de bolas *brancas:* a) *não seja inferior a um dado número* $L_1$ *nem superior a um dado número* $L_2$; b) *seja inferior a* $L_1$ *ou superior a* $L_2$.

Em a) trata-se de achar a probabilidade do acontecimento

$$L_1 \leq \nu \leq L_2$$

o qual é a soma lógica dos acontecimentos

$$\nu = L_1,\ \nu = L_1 + 1,\ \ldots,\ \nu = L_2 - 1,\ \nu = L_2,$$

incompatíveis dois a dois. Ter-se-á, portanto,

$$\Pr(L_1 \leq \nu \leq L_2) = \Pr(\nu = L_1) + \Pr(\nu = L_1 + 1) + \cdots + \Pr(\nu = L_2),$$

ou seja, atendendo ao resultado do exemplo 7:

$$\Pr(L_1 \leq \nu \leq L_2) = \sum_{\nu=L_1}^{L_2} \frac{\binom{a}{\nu}\binom{b}{n-\nu}}{\binom{N}{n}}.$$

Em b) pede-se a probabilidade do acontecimento

$$(\nu < L_1) + (L_2 < \nu) \quad (\nu \text{ inferior a } L_1 \text{ ou } L_2 \text{ inferior a } \nu)$$

que é o contrário do acontecimento $L_1 \leq \nu \leq L_2$. Será, pois,

$$\Pr[(\nu < L_1) + (\nu > L_2)] = 1 - \Pr(L_1 \leq \nu \leq L_2).$$

Eis como se poderiam calcular exactamente as probabilidades que fomos levados a considerar em A-16, a propósito da experiência hipotética sobre animais. Porém, estas fórmulas, embora simples na aparência, exigem cálculos muito laboriosos. É possível substituí-las por outras que, com muito menos trabalho, fornecem boas aproximações quando o número $n$ excede um certo limite.

## NOTA IMPORTANTE RELATIVA ÀS NOTAÇÕES E TERMINOLOGIA

Convencionámos no n.º 4 designar por $\Pr(\alpha)$ a probabilidade do acontecimento $\alpha$. Nesta ordem de ideias, representámos, atrás, por $\Pr(x=x_1)$ a probabilidade de "sair a bola $x_1$", por $\Pr(x \in V)$ a probabilidade de "sair bola vermelha", etc. Mas poderíamos, igualmente, para abreviar, designar por $\Pr(x_1)$, $\Pr(V)$, etc., aquelas probabilidades, dizendo "a probabilidade do elemento $x_1$", a "probabilidade do conjunto $V$", etc. Assim, a distribuição de probabilidade passa a conceber-se como distribuição definida, não num corpo de acontecimentos, mas, sim, num corpo de conjuntos: todos os possíveis conjuntos de bolas da urna (cf. A-4). Poderíamos, também, conceber esta distribuição como função $\Pr(x)$ da variável casual $x$, tendo-se, na hipótese de as bolas serem iguais,

$$\Pr(x) = \frac{1}{N}\,, \text{ para todo o valor de } x.$$

Estas considerações estendem-se, *mutatis mutandis,* a qualquer distribuição de probabilidade.

## 6. Independência e associação de acontecimentos

Sendo as probabilidades valores ideais de frequências relativas, as definições de independência e associação que demos para o caso das frequências traduzem-se, imediatamente, em termos de probabilidade. Seja $\mathcal{R}$ um corpo de eventualidades a considerar numa dada prova $\mathcal{P}$, com determinadas probabilidades. Dados dois acontecimentos $\alpha$, $\beta$, *chama-se probabilidade condicional de* $\beta$ *em relação a* $\alpha$ e *designa-se por* $\Pr(\beta|\alpha)$ ao *número dado pela fórmula*

$$(6.1) \qquad \Pr(\beta|\alpha) = \frac{\Pr(\alpha\beta)}{\Pr(\alpha)}.$$

Para avaliar $\Pr(\beta|\alpha)$ empiricamente, o processo a seguir consistiria em efectuar a prova $\mathcal{P}$ um grande número de vezes e registar: 1) a frequência absoluta ($\alpha$) de $\alpha$; 2) a frequência absoluta ($\alpha\beta$) de $\alpha\beta$. O quociente de ($\alpha\beta$) por ($\alpha$) dar-nos-ia, então, um valor aproximado de $\Pr(\beta|\alpha)$.

É claro que, da definição (6.1), resulta, logo, a *fórmula do produto*:

$$(6.2) \qquad \Pr(\alpha\beta) = \Pr(\alpha) \,.\, \Pr(\beta|\alpha) = \Pr(\beta) \,.\, \Pr(\alpha|\beta),$$

a qual nos diz que: *a probabilidade de se verificarem ao mesmo tempo os acontecimentos* $\alpha$ e $\beta$ *é o produto da probabilidade de* $\alpha$ *pela probabilidade condicional de* $\beta$ *na hipótese de se verificar* $\alpha$ (*ou vice-versa*).

Os acontecimentos $\alpha$, $\beta$ *dizem-se estocasticamente independentes* ou, apenas, *independentes*, quando $\Pr(\beta|\alpha) = \Pr(\beta)$ ou, o que é equivalente, quando $\Pr(\alpha|\beta) = \Pr(\alpha)$; no caso contrário, dizem-se *associados*. Desta definição e da fórmula do produto deduz-se, logo, o

TEOREMA DO PRODUTO. *Se os acontecimentos* $\alpha$, $\beta$ *são independentes* (*e só neste caso*), *a probabilidade de realização simultânea de* $\alpha$ *e* $\beta$ *é igual ao produto das probabilidades de* $\alpha$ *e de* $\beta$.

Suponhamos, por exemplo, que uma urna contém 8 bolas brancas e 24 pretas, estando 6 bolas brancas marcadas com o sinal +, 18 bolas pretas com este mesmo sinal, e todas as restantes com o sinal –. Então, como existem na urna 32 bolas, ao todo, sendo 24 marcadas com o sinal +, a probabilidade (incondicional) de aparecer o sinal + é 24/32 = 3/4. Por sua vez, a probabilidade de aparecer o sinal + em bola branca é 6/8 = 3/4, igual à primeira. Os acontecimentos "sair bola branca" e "sair sinal +" são, pois, independentes; deste modo, a probabilidade de "sair bola branca e sinal +" é:

$$\frac{8}{32} \cdot \frac{3}{4} = \frac{1}{4} \cdot \frac{3}{4} = \frac{3}{16}$$

(probabilidade de sair bola branca vezes probabilidade de sair sinal +).

Um exemplo mais sugestivo, embora menos rigoroso, será o seguinte. Consideremos a prova que consiste em *semear uma certa quantidade de trigo num certo campo*. Seja $p$ a probabilidade (incondicional) *de colher então uma quantidade de sementes superior a um dado limite L e* seja $p'$ a probabilidade do mesmo acontecimento, *na hipótese de a altura pluviométrica, nos meses de outono e inverno, ser inferior a um dado limite* $\lambda$. Será, então, $p'$ uma probabilidade condicional do primeiro acontecimento em relação ao segundo; e é natural que seja $p' \neq p$, isto é, que a eventualidade "colher trigo em quantidade superior a $L$" depende da eventualidade "chover em quantidade inferior a $\lambda$".

Mas, geralmente, dados vários acontecimentos $\alpha$, $\beta$, $\gamma$, ... do corpo $\mathcal{R}$ tem-se, como é fácil ver,

$$\Pr(\alpha\beta\gamma \ldots) = \Pr(\alpha)\Pr(\beta|\alpha)\Pr(\gamma|\alpha\beta)\ldots . \tag{6.3}$$

*Os acontecimentos* $\alpha$, $\beta$, $\gamma$, ... *são independentes, se for verdadeira, não só a igualdade*

$$\Pr(\alpha\beta\gamma \ldots) = \Pr(\alpha)\Pr(\beta)\Pr(\gamma) \ldots$$

*como todas as que resultam desta substituindo um ou mais dos acontecimentos* $\alpha$, $\beta$, $\gamma$, ... *pelos seus contrários* (cf. A-15)[1].

## 7. Sistema de duas experiências

Os conceitos de associação e independência de acontecimentos têm interesse, principalmente, quando se trata de várias experiências combinadas. Comecemos por considerar o caso das duas experiências $\mathcal{P}$, $\mathcal{P}$', iguais ou diferentes, e sejam $\mathcal{R}$ e $\mathcal{R}$' dois corpos de eventualidades a considerar, respectivamente, nas provas $\mathcal{P}$ e $\mathcal{P}$'. A realização destas duas provas, ao mesmo tempo ou uma após a outra, pode

(1) – Pode considerar-se esta proposição como definição de independência, no caso de vários acontecimentos.

conceber-se como *experiência* composta, única, que designaremos por ($\mathcal{P}$, $\mathcal{P}'$). Sejam $\alpha_1, \alpha_2, \ldots \alpha_m$ as células de $\mathcal{R}$ e $\beta_1, \beta_2, \ldots, \beta_n$ as células de $\mathcal{R}'$. Designaremos, então, por

$$(\alpha_i, \beta_k)$$

o acontecimento que consiste em realizar-se $\alpha_i$ na prova $\mathcal{P}$ e $\beta_k$ na prova $\mathcal{P}'$ (*acontecimento composto de* $\alpha_i$ *e* $\beta_k$). Todos estes acontecimentos $(\alpha_i, \beta_k)$ são resultados possíveis da prova ($\mathcal{P}$, $\mathcal{P}'$), são as células dum novo corpo, que designaremos por $\mathcal{R} \times \mathcal{R}'$. Note-se que cada acontecimento $\alpha_i$ pode ser considerado como elemento do corpo $\mathcal{R} \times \mathcal{R}'$, pondo

$$\alpha_i = (\alpha_i, \beta_1) + (\alpha_i, \beta_2) + \cdots + (\alpha_i, \beta_n),$$

visto que, $\beta_1 + \beta_2 + \cdots + \beta_n = I$. Analogamente:

$$\beta_k = (\alpha_1, \beta_k) + (\alpha_2, \beta_k) + \cdots + (\alpha_m, \beta_k).$$

Nesta ordem de ideias, podemos identificar o acontecimento composto $(\alpha_i, \beta_k)$ com o produto lógico $\alpha_i \beta_k$, desde que não haja perigo de confusão[(1)].

Suponhamos, agora, definido no corpo $\mathcal{R} \times \mathcal{R}'$ uma distribuição de probabilidade. Esta poderá indicar-se numa tábua do seguinte tipo (cf. tabela-tipo de A-14):

| | $\beta_1$ | $\beta_2$ | ...... | $\beta_n$ | Total |
|---|---|---|---|---|---|
| $\alpha_1$ | $Pr(\alpha_1 \beta_1)$ | $Pr(\alpha_1 \beta_2)$ | ...... | $Pr(\alpha_1 \beta_n)$ | $Pr(\alpha_1)$ |
| $\alpha_2$ | $Pr(\alpha_2 \beta_1)$ | $Pr(\alpha_2 \beta_2)$ | ...... | $Pr(\alpha_2 \beta_n)$ | $Pr(\alpha_2)$ |
| ...... | ...... | ...... | ...... | ...... | ...... |
| $\alpha_m$ | $Pr(\alpha_m \beta_1)$ | $Pr(\alpha_m \beta_2)$ | ...... | $Pr(\alpha_m \beta_n)$ | $Pr(\alpha_m)$ |
| Total | $Pr(\beta_1)$ | $Pr(\beta_2)$ | ...... | $Pr(\beta_n)$ | 1 |

(1) – A confusão pode surgir quando $\mathcal{P}'$ é apenas a repetição de $\mathcal{P}$. Neste caso, um acontecimento $\alpha$ deverá ser designado por dois símbolos diferentes, conforme se realize em $\mathcal{P}$ ou em $\mathcal{P}'$. Um outro modo de evitar a confusão é chamar a $(\alpha_i, \beta_k)$ o *produto cartesiano de* $\alpha_i$ *por* $\beta_k$. Este produto, *que não é comutativo*, será designado por $\alpha_i \times \beta_k$ para o distinguir do produto lógico vulgar $\alpha_i \beta_k$.

Note-se que, para $i = 1, \ldots, m$, $k = 1, 2, \ldots, n$, se tem:

$$\Pr(\alpha_i) = \Pr(\alpha_i\beta_1) + \Pr(\alpha_i\beta_2) + \cdots + \Pr(\alpha_i\beta_n),$$

$$\Pr(\beta_k) = \Pr(\alpha_1\beta_2) + \Pr(\alpha_2\beta_k) + \cdots + \Pr(\alpha_m\beta_k).$$

Como estas probabilidades estão registadas à margem (sob a indicação "total"), dá-se-lhes, também, o nome de *probabilidades marginais* da distribuição considerada. Pode acontecer, em particular, que se tenha

$$\Pr(\alpha_i\beta_k) = \Pr(\alpha_i)\Pr(\beta_k),$$

quaisquer que sejam $i$, $k$. Então, é fácil ver que *a probabilidade de qualquer acontecimento* $\alpha\beta$, *composto dum acontecimento* $\alpha$ *de* $\mathcal{R}$ *e dum acontecimento* $\beta$ *de* $\mathcal{R}'$ *será igual ao produto das probabilidades de* $\alpha$ *e de* $\beta$[1]. Os acontecimentos de $\mathcal{R}$ dir-se-ão, neste caso, *independentes* dos acontecimentos de $\mathcal{R}'$. De contrário, terá de usar-se a fórmula (6.2).

Estas considerações tornam-se mais intuitivas quando se fala em termos de variáveis casuais. Consideremos um par $(x, y)$ de variáveis casuais com uma determinada distribuição de probabilidade, $\Pr(x, y)$. Chama-se *probabilidade condicional dum valor* $x_i$ *de* $x$ *a respeito dum valor* $y_k$ de $y$, e representa-se por $\Pr(x_i|y_k)$, o número dado por

$$\Pr(x_i|y_k) = \frac{\Pr(x_i, y_k)}{\Pr(y_k)}.$$

As variáveis $x$, $y$ dizem-se *independentes* (*estocasticamente*), se for $\Pr(x|y) = \Pr(x)$ para todos os valores de $x$ e $y$. Tem-se, pois, nesta hipótese

$$\Pr(x, y) = \Pr(x)\Pr(y).$$

(1) – Esta regra, ou, mais geralmente, a fórmula (6.3), é conhecida tradicionalmente, como *princípio das probabilidades compostas.*

NOTA. Seria mais correcto designar por símbolos diferentes as distribuições de probabilidade de $x$ e $y$: por exemplo, a primeira por $\Pr_1(x)$ e a segunda por $\Pr_2(y)$. Mas, para não sobrecarregar as notações, e não havendo perigo de confusão, usamos aqui, para ambas as distribuições, o mesmo símbolo.

Exemplos – Consideremos uma urna que contenha $M$ bolas $x_1, x_2, \ldots, x_M$, sendo as $v$ primeiras vermelhas e as $b$ últimas brancas. E consideremos uma outra urna, que contenha $N$ bolas $y_1, y_2, \ldots, y_N$, sendo as $v'$ primeiras vermelhas e as $b'$ últimas brancas. Sejam $\mathcal{P}$, $\mathcal{P}'$, respectivamente, as experiências que consistem em tirar à sorte uma bola da primeira urna e tirar à sorte uma bola da segunda urna. Neste caso, os resultados elementares da experiência composta $(\mathcal{P}, \mathcal{P}')$ serão expressos pelos diferentes valores $(x_i, y_k)$ da variável casual $(x, y)$. Como estes valores são em número de $MN$ e *todos igualmente prováveis*, a probabilidade de cada par $(x_i, y_k)$ será

$$\frac{1}{MN} = \frac{1}{M} \cdot \frac{1}{N} = \Pr(x_i)\Pr(y_k),$$

o que significa que *as duas variáveis casuais $x$, $y$ são independentes* (*estocasticamente*). Os resultados da prova $\mathcal{P}'$ são pois independentes dos resultados da prova $\mathcal{P}$. Por exemplo, a probabilidade do acontecimento *sair bola branca em $\mathcal{P}$ e bola vermelha em $\mathcal{P}'$*, será

$$\frac{b}{M} \cdot \frac{v'}{N} = \frac{bv'}{MN}.$$

Casos análogos ao anterior serão, ainda, todos aqueles em que $\mathcal{P}'$ é simplesmente a repetição de $\mathcal{P}$. Suponhamos, por exemplo, que se trata de duas extracções sucessivas duma bola da 1.ª urna, *com reposição*. A probabilidade de sair, em primeiro lugar, bola vermelha e, depois, bola branca será

$$\frac{v}{M} \cdot \frac{b}{M} = \frac{bv}{M^2}.$$

igual à probabilidade de sair, primeiro, branca e, depois, vermelha (será, portanto, $2bv/M^2$ a probabilidade do acontecimento "sairem cores diferentes").

Mas suponhamos, agora, que $\mathcal{P}$ consiste em tirar uma bola da 1.ª urna, e $\mathcal{P}'$ em tirar uma segunda bola da mesma urna *sem repor a primeira bola tirada*. Os resultados elementares da prova composta ($\mathcal{P}$, $\mathcal{P}'$) serão, ainda, os valores da variável casual $(x, y)$, *mas com a condição suplementar* $x \neq y$. Isto basta para ver que, neste caso, as variáveis casuais $x$, $y$ não são independentes. Por exemplo, a probabilidade (condicional) de sair bola vermelha na 2.ª extracção, *tendo saído branca na* 1.ª, é

$$\frac{v}{M-1},$$

ao passo que a probabilidade (condicional) de sair vermelha na 2.ª, *tendo saído vermelha* na 1.ª é

$$\frac{v-1}{M-1}.$$

Assim, a probabilidade de *sair vermelha duas vezes* é

$$\frac{v}{M} \cdot \frac{v-1}{M-1} = \frac{v(v-1)}{M(M-1)}$$

e a de *sair primeiro branca e depois vermelha*:

$$\frac{b}{M} \cdot \frac{v}{M-1} = \frac{bv}{M(M-1)}.$$

Estes resultados podem recolher-se na seguinte tabela:

TABELA N.º 13

| 1.ª \ 2.ª | Vermelha | Branca | Total |
|---|---|---|---|
| Vermelha | $\frac{v(v-1)}{M(M-1)}$ | $\frac{vb}{M(M-1)}$ | $\frac{v}{M}$ |
| Branca | $\frac{bv}{M(M-1)}$ | $\frac{b(b-1)}{M(M-1)}$ | $\frac{b}{M}$ |
| Total | $\frac{v}{M}$ | $\frac{b}{M}$ | 1 |

Na margem direita estão indicadas as probabilidades de sair bola vermelha e a de sair bola branca na 2.ª extracção, respectivamente, iguais às probabilidades de sair vermelha e de sair branca na 1.ª extracção (registadas na margem inferior). A tabela patenteia, assim, a associação dos acontecimentos considerados.

## 8. Sistema de várias experiências

Podemos, agora, conceber, mais geralmente, uma experiência $(\mathcal{P}, \mathcal{P}', \mathcal{P}'', \ldots)$, composta de várias experiências $\mathcal{P}, \mathcal{P}', \mathcal{P}'', \ldots$, (em número finito). Sendo $\alpha, \beta, \gamma, \ldots$ eventualidades a considerar em cada uma destas experiências, designaremos por $(\alpha, \beta, \gamma, \ldots)$ a eventualidade que consiste em acontecer $\alpha$ em $\mathcal{P}$, $\beta$ em $\mathcal{P}'$, $\gamma$ em $\mathcal{P}''$, etc., e diremos que $(\alpha, \beta, \gamma, \ldots)$ é o *acontecimento composto* de $\alpha, \beta, \gamma, \ldots$. Porém, segundo o ponto de vista explanado no número anterior, podemos considerar este novo acontecimento como o produto lógico dos acontecimentos $\alpha, \beta, \gamma, \ldots$, desde que se tomem as devidas precauções[1]. Suponhamos que, a todo o acontecimento a considerar na prova $(\mathcal{P}, \mathcal{P}', \mathcal{P}'', \ldots)$, corresponde uma determinada probabilidade. Então, o que se disse em B-6 é aqui aplicável: os acontecimentos $\alpha, \beta, \gamma, \ldots$ dizem-se *independentes* (*estocasticamente*), se for verdadeira não só a igualdade

$$(8.1) \qquad \Pr(\alpha\beta\gamma \ldots) = \Pr(\alpha) \Pr(\beta) \Pr(\gamma) \ldots,$$

como todas as que se deduzem desta substituindo um ou mais dos acontecimentos $\alpha, \beta, \gamma, \ldots$ pelos seus contrários. Não sendo assim, terá de usar-se a fórmula geral do produto, com as probabilidades condicionais.

Poderíamos, de novo, dar aqui exemplos de tiragens de bolas de uma ou várias urnas, com ou sem reposição, mas as considerações do número seguinte bastam para esclarecer as precedentes.

---

(1) – Isto é, se $\mathcal{P}', \mathcal{P}'', \ldots$ consistem apenas na repetição de $\mathcal{P}$, um mesmo acontecimento deverá ser designado de modos diversos, conforme as provas em que se realiza. É o que faremos mais adiante.

## 9. Distribuição binomial ou de BERNOULLI

Seja $\alpha$ um acontecimento a esperar com determinada probabilidade numa certa prova $\mathcal{P}$ e consideremos $n$ realizações $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_n$, da prova genérica $\mathcal{P}$ (1). Estas provas particulares constituem uma experiência composta $(\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_n)$. Tendo em vista o conceito empírico de probabilidade, é fácil reconhecer que a probabilidade do acontecimento $\alpha$ deverá ser a mesma em cada uma das provas, quaisquer que sejam os resultados das restantes provas. Por outras palavras:

*Os acontecimentos a esperar na repetição duma dada prova são independentes dos resultados das provas já efectuadas.*

É claro que este facto não se deduz da anterior axiomática das probabilidades. Poderia, sim, assumir-se como novo axioma, porém com carácter bem diverso do dos primeiros.

Um exemplo típico é o das sucessivas extracções casuais de bolas de uma urna, com reposição da bola após cada extracção.

Pode, pois, aplicar-se nestes casos a fórmula (8.1).

Proponhamo-nos, então, resolver o seguinte problema:

*Determinar a probabilidade de que, em n realizações da prova $\mathcal{P}$, um acontecimento $\alpha$ de probabilidade p se realize x vezes.*

(Para concretizar, podemos supor que $\alpha$ consiste em tirar à sorte uma bola duma urna com bolas brancas e pretas, sendo $\alpha$ o acontecimento "sair bola branca". Porém, as considerações que se seguem são de todo gerais).

Representemos por $q$ a probabilidade de $\tilde{\alpha}$, isto é, ponhamos $q=1-p$. Para evitar confusões, designaremos por $\alpha_i$ o acontecimento particular que consiste em realizar-se $\alpha$ na prova $\mathcal{P}_i$ $(i = 1, 2, \ldots, n)$; é claro que será sempre

$$\Pr(\alpha_i) = p, \Pr(\tilde{\alpha}_i) = q.$$

Suponhamos, por exemplo, $n=3$ e $x=2$. Três são os modos de $\alpha$ se realizar 2 vezes numa série de 3 provas: na 1.ª e na 2.ª, na 1.ª e na 3.ª ou na 2.ª e na 3.ª; isto é, em símbolos:

$$\alpha_1\alpha_2\tilde{\alpha}_3, \qquad \alpha_1\tilde{\alpha}_2\alpha_3, \qquad \tilde{\alpha}_1\alpha_2\alpha_3.$$

(1) – Dizendo que $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_n$ são realizações da mesma prova $\mathcal{P}$, fica implícito que $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_n$ se realizam em condições idênticas.

O acontecimento que consiste em $\alpha$ *se realizar* 2 *vezes nas* 3 *provas*, será, pois, a soma lógica

$$\alpha_1\alpha_2\tilde{\alpha}_3 + \alpha_1\tilde{\alpha}_2\alpha_3 + \tilde{\alpha}_1\alpha_2\alpha_3.$$

Ora, as três modalidades consideradas são incompatíveis duas a duas e têm todas as mesmas probabilidades; por exemplo:

$$\Pr(\alpha_1\tilde{\alpha}_2\alpha_3) = \Pr(\alpha_1)\Pr(\tilde{\alpha}_2)\Pr(\alpha_3) = pqp = p^2q.$$

A probabilidade pedida será, pois, neste caso, $3p^2q$.

Passemos, agora, ao caso geral. Um dos modos de $\alpha$ se realizar $x$ vezes nas $n$ provas será

$$\alpha_1\alpha_2 \ldots \alpha_x\tilde{\alpha}_{x+1} \ldots \tilde{\alpha}_n.$$

Qualquer outra modalidade se obtém escolhendo, entre as $n$ provas, as $x$ provas em que se realize $\alpha$. As modalidades possíveis correspondem, pois, às combinações das $n$ provas $x$ a $x$. Como o número total destas combinações é $\binom{n}{x}$, o acontecimento que consiste em $\alpha$ *se realizar x vezes nas n provas* é a soma lógica de $\binom{n}{x}$ eventualidades, incompatíveis duas a duas e todas com a mesma probabilidade, que é

$$\Pr(\alpha_1)\Pr(\alpha_2) \ldots \Pr(\alpha_x)\Pr(\tilde{\alpha}_{x+1}) \ldots \Pr(\tilde{\alpha}_n) = p^x q^{n-x}.$$

A probabilidade pedida será, pois,

$$\Pr(x) = \binom{n}{x} p^x q^{n-x}$$

fórmula esta muito importante. Note-se que a variável casual *x representa, aqui, a frequência absoluta de* $\alpha$ *numa série de n provas.* A distribuição de probabilidade desta variável, dada pela fórmula anterior, tem o nome de *distribuição de BERNOULLI.* Também se lhe chama distribuição *binominal*, atendendo a que os diferentes valores de $\Pr(x)$ são os termos do desenvolvimento da potência $n$ do binómio $p + q$:

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

É claro que $(p + q)^n = 1$. Procuremos, agora, a probabilidade de que a frequência absoluta de $\alpha$ seja inferior ou igual a um dado limite $L$ (com $L \leq n$). Como o acontecimento $x \leq L$ é a soma lógica dos acontecimentos incompatíveis $x = 0, x = 1, \ldots, x = L$, virá:

$$\Pr(x \leq \mathrm{L}) = \Pr(x=0) + \Pr(x=1) + \cdots + \Pr(x=L) = \sum_{x=0}^{L} \binom{n}{x} p^x q^{n-x}.$$

Analogamente, se reconhece que

$$\Pr(L_1 \leq x \leq L_2) = \sum_{x=L_1}^{L_2} \binom{n}{x} p^x q^{n-x}.$$

(Confrontar com os exemplos 7, 8 e 9 do n.° 5).

Se pusermos $\xi = x/n$ (frequência relativa de $\alpha$ nas $n$ provas), a distribuição da variável $\xi$ será, manifestamente:

$$\Pr(\xi) = \binom{n}{n\xi} p^{n\xi} q^{n(1-\xi)}.$$

A tabela n.° 14 representa a distribuição binominal para $p = 1/2$, $n = 10$. Por exemplo, a probabilidade de que em 10 lançamentos sucessivos duma moeda "correcta" se apresente 3 vezes coroa, é

$$\Pr(3) = \binom{10}{3} \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^7 = \left(\frac{10.9.8}{3.2.1}\right) \cdot \left(\frac{1}{2^{10}}\right) = \frac{15}{128},$$

valor que concorda com o correspondente da tabela (a menos de 0,001).

TABELA N.º 14

| $x$ | $\Pr(x)$ | $x$ | $\Pr(x)$ |
|---|---|---|---|
| 0 | 0,001 | 6 | 0,205 |
| 1 | 0,010 | 7 | 0,117 |
| 2 | 0,044 | 8 | 0,044 |
| 3 | 0,117 | 9 | 0,010 |
| 4 | 0,205 | 10 | 0,001 |
| 5 | 0,246 | | |

Na Fig. 3 é dado o histograma desta distribuição.

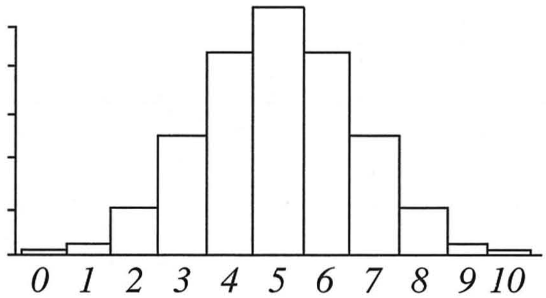


Fig. 3

Note-se que o histograma é *simétrico* a respeito da recta $x = 5$ e que a função $\Pr(x)$ atinge um máximo no ponto 5.

Podemos ver um exemplo curioso desta distribuição no campo da genética. Sabe-se que, no cruzamento dum touro avermelhado do Shorthorn com uma vaca malhada da mesma raça, há a probabilidade de 1/2 de se obter um vitelo avermelhado sem malhas(1). Então, a probabilidade de que, em 10 destes cruzamentos, se obtenham $x$ vitelos avermelhados (sem malhas), é dada pela tabela n.º 14.

(1) – Traduz-se por ''touro avermelhado'' a expressão inglesa ''red bull'' e por ''vaca malhada'' a expressão ''roan cow''. Neste caso, ''malhado'' significa ''avermelhado, com malhas brancas ou cinzentas dispersas''.

Por sua vez, sabe-se que a probabilidade de que, no cruzamento dum touro malhado Shorthorn com uma vaca malhada da mesma raça se obtenha um vitelo avermelhado (sem malhas) é 1/4. Então, a probabilidade de que, em 10 destes cruzamentos, se obtenham $x$ vitelos avermelhados (sem malhas) será:

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

Na Fig. 4 é dado o histograma desta distribuição, que já não apresenta simetria.

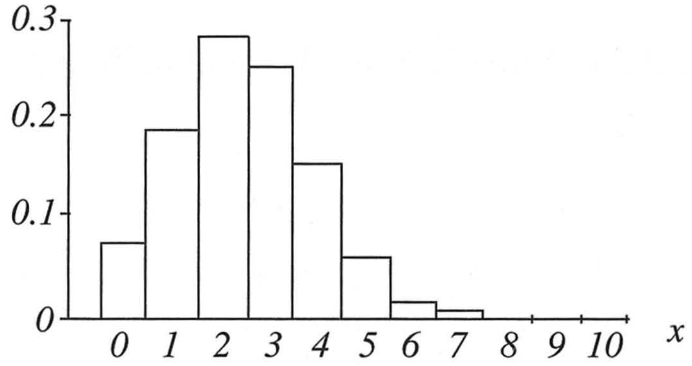


Fig. 4

Os dois casos extremos na distribuição binominal são os seguintes:

1) $x = n$ (o acontecimento $\alpha$ realiza-se nas $n$ provas)

2) $x = 0$ (o acontecimento $\alpha$ não se realiza em nenhuma das provas).

A probabilidade do primeiro caso é (como se poderia reconhecer mesmo directamente):

$$\Pr(n) = p^n.$$

A probabilidade do segundo caso é:

$$\Pr(0) = q^n.$$

Note-se, porém, que a negação do acontecimento $x = 0$ não é o acontecimentos $x = n$, mas sim o acontecimento que consiste em $\alpha$ *se realizar pelo menos uma vez em n provas.*

A probabilidade deste acontecimento será, pois,

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

Exemplo – *Calcular a probabilidade de que, jogando* 1.000 *vezes num número duma lotaria com* 30.000 *números, se tenha pelo menos uma vez a sorte grande.*

A probabilidade de, em cada extracção, se ter a sorte grande, é, manifestamente, 1/30.000. Então, a probabilidade pedida será

$$1 - \left(1 - \frac{1}{30.000}\right)^{1.000} = \frac{1.000}{30.000} - \binom{1.000}{2} \frac{1}{(30.000)^2} + \cdots \approx 0{,}0328$$

probabilidade esta ainda muito fraca, apesar do grande número de tentativas.

## 10. Conceito de moda. Caso da distribuição normal

Chama-se *moda* duma distribuição de probabilidade duma variável $x$ (com um número finito de valores) todo o valor de $x$, cuja probabilidade seja superior ou igual à de qualquer outro valor de $x$. Há distribuições com uma só moda, distribuições com mais de uma moda e distribuições sem moda.

Vamos ver que a distribuição binominal

$$\Pr(x) = \frac{n!}{x!\,(n-x)!} p^x q^{n-x} \quad (\text{com } q = 1 - p)$$

apresenta uma ou, quando muito, duas modas.

Considerando três valores consecutivos, $X-1$, $X$, $X+1$ da variável $x$, tem-se

$$\Pr(X-1) = \frac{n!}{(X-1)!\,(n-X+1)!} p^{X-1} q^{n-X+1}$$

$$\Pr(X) = \frac{n!}{X!\,(n-X)!} p^X q^{n-X}$$

$$\Pr(X+1) = \frac{n!}{(X+1)!\,(n-X-1)!} p^{X+1} q^{n-X-1}.$$

Então, virá, como é fácil verificar,

$$\frac{\Pr(X)}{\Pr(X-1)} = \frac{(n-X+1)p}{Xq},$$

$$\frac{\Pr(X)}{\Pr(X+1)} = \frac{(X+1)q}{(n-X)p}.$$

Por conseguinte, para que $X$ seja uma moda, deve ser, simultaneamente,

$$(n-X+1)p \geq Xq,$$
$$(X+1)q \geq (n-X)p.$$

Ora, esta dupla condição equivale à seguinte

$$np - q \leq X \leq np + p.$$

Como a diferença entre $np+p$ e $np-q$ é igual a $p+q=1$, a anterior condição será verificada por um só valor inteiro de $X$ (a moda), a não ser que $np+p$ e $np-q$ sejam números inteiros, que serão, nesse caso, as duas modas existentes.

Assim, a moda ou as modas da distribuição binominal são sempre números inteiros que diferem de $np$ menos de uma unidade.

No primeiro exemplo atrás considerado ($p=1/2$, $n=10$), a moda é precisamente $np=10.1/2=5$. No segundo exemplo ($p=1/4$, $n=10$), a moda $X$ deve verificar a condição

$$1{,}75 = 10 \cdot \frac{1}{4} - \frac{3}{4} \leq X \leq 10 \cdot \frac{1}{4} + \frac{1}{4} = 2{,}75;$$

só poderá, então, ser $X=2$. Mas, se em vez de $n=10$, tivéssemos tomado $n=15$, com $p=1/4$, já teríamos duas modas: $X_1=3$, $X_2=4$.

## 11. Distribuição polinomial. Amostras casuais

Consideremos agora, mais geralmente, uma partição formada por $r$ acontecimentos $\alpha_1, \alpha_2, \dots, \alpha_r$, a prever numa certa prova $\mathcal{P}$, com probabilidades $p_1, p_2, \dots, p_r$ respectivamente[1]. Pergunta-se:

*Qual a probabilidade de que, em n realizações da prova $\mathcal{P}$, as frequências absolutas dos acontecimentos $\alpha_1, \alpha_2, \dots, \alpha_r$, sejam, respectivamente, $x_1, x_2, \dots, x_r$?*

É claro que, sendo as eventualidades $\alpha_1, \alpha_2, \dots, \alpha_r$, por hipótese, incompatíveis duas a duas, e sendo a sua soma lógica o acontecimento certo, deverá ter-se

$$p_1 + p_2 + \cdots + p_r = 1, \qquad x_1 + x_2 + \cdots + x_r = n.$$

Suponhamos que os acontecimentos $\alpha_1, \alpha_2, \dots, \alpha_r$ se verificam numa determinada ordem, por exemplo:

$\alpha_1$ nas provas $\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_{x_1}$

$\alpha_2$ nas provas $\mathcal{P}_{x_1+1}, \mathcal{P}_{x_1+2}, \dots, \mathcal{P}_{x_1+x_2}$

..........................................

$\alpha_r$ nas provas $\mathcal{P}_{n-x_r}; \mathcal{P}_{n-x_r+1}, \dots, \mathcal{P}_n$ .

Obtém-se, então, um acontecimento composto, cuja probabilidade é $p_1^{x_1} p_2^{x_2} \dots p_r^{x_r}$. É claro que, qualquer que seja a ordem de realização, a probabilidade será esta; a probabilidade pedida será, pois, o produto de $p_1^{x_1} p_2^{x_2} \dots p_r^{x_r}$ pelo número total de ordens possíveis. Mas este número não é $n!$ (permutações das $n$ provas), como poderia parecer, visto que, permutando entre si as $x_1$ provas em que se verifica $\alpha_1$, as $x_2$ provas em que se verifica $\alpha_2$, ..., as $x_r$ provas em que se verifica $\alpha_r$, se obtém sempre o mesmo acontecimento. Como se vê facilmente, o número das ordens possíveis será $n!$ dividido por $x_1!\, x_2! \dots x_r!$, e a probabilidade pedida será, então,

$$\Pr(x_1, x_2, \dots x_r) = \frac{n!}{x_1!\, x_2! \dots x_r!}\, p_1^{x_1} p_2^{x_2} \dots p_r^{x_r}.$$

(1) – As eventualidades $\alpha_1, \alpha_2, \dots, \alpha_r$ serão pois incompatíveis duas a duas e a sua soma lógica será o acontecimento certo.

Temos aqui, pois, um exemplo duma distribuição de $r$ variáveis casuais $x_1, x_2, \ldots, x_r$. É claro que, por ser $x_1 + x_2 + \ldots + x_r = n$, estas variáveis não são independentes, nem sequer algebricamente.

A referida distribuição diz-se *polinominal*, atendendo a que os valores de $\Pr(x_1, x_2, \ldots, x_r)$ são os termos do desenvolvimento da potência $n$ do polinómio $p_1 + p_2 + \ldots + p_r$. Tem-se, com efeito, segundo a fórmula de LEIBNIZ:

$$(p_1 + p_2 + \cdots + p_r)^n = \sum_{x_1 + \ldots + x_r = n} \frac{n!}{x_1!\, x_2! \ldots x_r!} p_1^{x_1} p_2^{x_2} \ldots p_r^{x_r}.$$

Note-se que a distribuição binominal é um caso particular desta, correspondente a $r = 2$ (partição dicotómica). Tem-se, então, $x_2 = n - x_1$, donde

$$\frac{n!}{x_1!\, x_2!} = \frac{n!}{x_1!\,(n - x_1)!} = \binom{n}{x_1}.$$

Este resultado aplica-se em questões de *amostragem casual.* Chama-se *amostra casual* a todo o sistema $(u_1, u_2, \ldots, u_n)$ de $n$ indivíduos tirados *ao acaso* duma dada população $U$.

Como já tivemos ocasião de verificar em exemplos, a amostragem casual pode fazer-se de dois modos: *com reposição ou sem reposição.* No primeiro caso, cada um dos indivíduos $u_1, u_2, \ldots, u_n$ é tirado e, em seguida, reposto na população, antes de se tirar o seguinte: *pode assim acontecer que um mesmo indivíduo apareça repetido na amostra.* No segundo caso, os indivíduos são tirados sucessivamente, sem regressarem à população após as tiragens.

Suponhamos dada no universo $U$ uma partição em $r$ atributos $\alpha_1, \alpha_2, \ldots, \alpha_r$ e sejam $a_1, a_2, \ldots, a_r$, respectivamente, as frequências absolutas de $\alpha_1, \alpha_2, \ldots, \alpha_r$ em $U$. Suponhamos, ainda, que todos os indivíduos têm igual probabilidade de ser tirados. Então, a probabilidade de aparecer um indivíduo com o atributo $\alpha_i$ será, manifestamente,

$$p_i = \frac{a_i}{N}, \; i = 1, 2, \ldots, r,$$

em que $N = a_1 + a_2 + \ldots + a_r$ (número de elementos de $U$).

*Qual é, neste caso, a probabilidade de que, numa amostragem casual de n elementos, com reposição, as frequências absolutas de* $\alpha_1, \alpha_2, \ldots, \alpha_r$ *sejam, respectivamente,* $x_1, x_2, \ldots, x_r$?

A resposta é, segundo o que vimos,

$$\Pr(x_1, x_2, \ldots, x_r) = \frac{n!}{x_1!\, x_2! \ldots x_r!} \left(\frac{a_1}{N}\right)^{x_1} \left(\frac{a_2}{N}\right)^{x_2} \cdots \left(\frac{a_r}{N}\right)^{x_r}.$$

Mas seja, agora, este outro problema:

*Qual é a probabilidade de que, numa amostragem casual de n elementos, sem reposição, as frequências absolutas de* $\alpha_1, \alpha_2, \ldots, \alpha_r$, *sejam, respectivamente,* $x_1, x_2, \ldots, x_r$?

Raciocinando como no exemplo 7 do n.° 21, chega-se, agora, ao resultado

$$\Pr(x_1, x_2, \ldots, x_r) = \frac{\binom{a_1}{x_1}\binom{a_2}{x_2} \cdots \binom{a_r}{x_r}}{\binom{N}{n}}.$$

Eis aqui um novo exemplo de distribuição de $r$ variáveis casuais, que, no caso particular $r = 2$, tem o nome de *distribuição hipergeométrica.*

É fácil ver que, se $n$ é bastante pequeno em relação a $N$ (isto é, se a razão $n/N$ é bastante pequena), esta distribuição coincide, sensivelmente, com a anterior.

Note-se, ainda, que, na prática, as fórmulas obtidas exigem cálculos muito laboriosos quando $n$ é grande. Têm de ser então substituídas por outras que, embora não sejam exactas, se adaptam muito melhor ao cálculo numérico, dando uma boa aproximação, para valores de $n$ elevados.