Capítulo I: Probabilidades

Elementos de Probabilidades e Teoria de Números

Licenciatura em Engenharia Informática
e
Mestrado Integrado em Engenharia Informática

Universidade do Minho
Ano Letivo 2021/2022

• -> está no formulário

2. Definição de Probabilidade

O que sabemos sobre uma frequência relativa:

A frequência relativa de um acontecimento A em n repetições de uma experiência aleatória é dada pelo seguinte quociente:

$$f_A = \frac{\text{número de vezes que } A \text{ ocorreu nas } n \text{ repetições}}{n}.$$

Temos que, qualquer que seja o acontecimento A,

$$f_A \geq 0 \tag{1}$$

e que, em particular,

$$f_\Omega = 1. \tag{2}$$

Por outro lado, se A e B são dois acontecimentos disjuntos (i.e., $A \cap B = \varnothing$), temos

$$f_{A \cup B} = f_A + f_B. \tag{3}$$

Estas três propriedades da frequência relativa inspiraram a definição axiomática de probabilidade, apresentada de seguida.

Definição

Uma probabilidade sobre um espaço amostral Ω é uma função que a cada acontecimento A associa um número real, $P(A)$, designado de *probabilidade de* A, e que satisfaz os seguintes três axiomas:

i) $P(A) \geq 0$, para qualquer acontecimento A (não negatividade);

ii) $P(\Omega) = 1$;

iii) $P(A_1 \cup A_2 \cup A_3 \cup \ldots) = P(A_1) + P(A_2) + P(A_3) + \ldots$, para quaisquer acontecimentos $A_1, A_2, A_3, \ldots$ disjuntos 2 a 2 (i.e., $A_i \cap A_j = \emptyset$, para $i \neq j$).

Observação: Quando Ω é um conjunto finito, o axioma iii) equivale a

$$P(A_1 \cup A_2) = P(A_1) + P(A_2)$$

para quaisquer dois acontecimentos A_1 e A_2 disjuntos.

2. Definição de Probabilidade

Uma definição particular de probabilidade, que se utiliza quando Ω é finito e os acontecimentos elementares são equiprováveis, é

$$P(A) = \frac{\text{número de elementos do subconjunto } A}{\text{número de elementos do conjunto } \Omega} = \frac{\sharp(A)}{\sharp(\Omega)}. \tag{4}$$

É conhecida como a *probabilidade de Laplace*.

Na experiência do lançamento de um dado equilibrado, é precisamente esta a probabilidade que utilizamos. Por exemplo, a probabilidade do acontecimento "saiu uma face ímpar" é $\frac{1}{2}$ $\left(= \frac{3}{6}\right)$ porque o acontecimento em causa corresponde ao subconjunto $\{1, 3, 5\}$, que tem 3 elementos, e Ω tem cardinal igual a 6.

Apesar de muito popular, a definição (4) tem limitações óbvias:

- Ω tem que ser finito (o que nem sempre acontece - ver exemplos de experiências aleatórias atrás referidas);
- nem sempre temos acontecimentos elementares equiprováveis (por exemplo, lançamentos de dados ou moedas viciados).

2. Definição de Probabilidade

Algumas propriedades de uma probabilidade:

Da definição axiomática de probabilidade, deduzem-se muito facilmente várias propriedades, nomeadamente:

i) $P(\overline{A}) = 1 - P(A)$, onde $\overline{A} = \Omega \setminus A = \{x \in \Omega : x \notin A\}$.

ii) Se $A \subseteq B$ tem-se que

$$P(A) \leq P(B)$$

e

$$P(B \setminus A) = P(B) - P(A),$$

onde $B \setminus A = B \cap \overline{A}$.

iii) $P(\emptyset) = 0$ e, para qualquer acontecimento A, $0 \leq P(A) \leq 1$.

iv) Para quaisquer dois acontecimentos A e B, tem-se

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

v) Para quaisquer dois acontecimentos A e B, tem-se

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

Da última propriedade deduz-se ainda que, para quaisquer três acontecimentos A, B e C, tem-se

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

Mais, deduz-se também a conhecida *Fórmula de Poincaré*: para quaisquer n acontecimentos, $A_1, A_2, \ldots, A_n$, tem-se

$$\begin{aligned} P(A_1 \cup A_2 \cup \ldots \cup A_n) &= \sum_{i=1}^{n} P(A_i) - \sum_{i=1}^{n} \sum_{j=i+1}^{n} P(A_i \cap A_j) \\ &+ \sum_{i=1}^{n} \sum_{j=i+1}^{n} \sum_{k=j+1}^{n} P(A_i \cap A_j \cap A_k) \\ &- \sum_{i=1}^{n} \sum_{j=i+1}^{n} \sum_{k=j+1}^{n} \sum_{l=k+1}^{n} P(A_i \cap A_j \cap A_k \cap A_l) \\ &+ \ldots + (-1)^{n+1} P(A_1 \cap A_2 \cap \ldots \cap A_n) \end{aligned}$$

É natural pensar que, sabendo que B ocorreu, a nova probabilidade de um acontecimento A vai depender do que existir em comum entre A e B, i.e., vai depender de $A \cap B$. Mais, a nova probabilidade de B deverá ser agora igual a 1. Surge assim a seguinte definição.

Definição

Seja Ω o espaço amostral de uma experiência aleatória e $B \subseteq \Omega$ um acontecimento tal que $P(B) > 0$. Para um acontecimento $A \subseteq \Omega$, a *probabilidade de A condicionada por B* (ou *probabilidade de A dado B* ou ainda *probabilidade de A sabendo que B ocorreu*) é denotada por $P(A \mid B)$ e é dada por

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

Nota: Desde que $P(B) > 0$, podemos agora calcular $P(A \cap B)$ do seguinte modo:

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

E desde que $P(A) > 0$, também temos: $P(A \cap B) = P(B \mid A)\, P(A)$.

Teorema da Probabilidade Total e Fórmula de Bayes:

Sejam $A_1, A_2, \ldots, A_n, \ldots$ acontecimentos disjuntos 2 a 2 decorrentes de uma certa experiência aleatória, com espaço amostral Ω, e tais que

$$\Omega = A_1 \cup A_2 \cup \ldots \cup A_n \cup \ldots = \bigcup_{i \in \mathbb{N}} A_i$$

(ou seja, $A_1, A_2, \ldots, A_n, \ldots$ formam uma *partição de* Ω). Se $P(A_i) > 0$, então:

- Para um qualquer acontecimento B, tem-se

$$P(B) = P(B \mid A_1)P(A_1) + P(B \mid A_2)P(A_2) + \ldots + P(B \mid A_n)P(A_n) + \ldots; \quad (5)$$

- Se B é um acontecimento tal que $P(B) > 0$, tem-se

$$P(A_i \mid B) = \frac{P(B \mid A_i)\, P(A_i)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \ldots + P(B|A_n)P(A_n) + \ldots}. \quad (6)$$

Os resultados (5) e (6) são conhecidos por **Teorema da Probabilidade Total** e **Fórmula de Bayes**, respetivamente. A sua dedução (TPC) é um exercício simples, que faz uso de propriedades de operações entre conjuntos e de propriedades de uma probabilidade.

Independência de acontecimentos:

Intuitivamente, dois acontecimentos, A e B, serão independentes se a ocorrência de um deles não alterar a probabilidade de ocorrência do outro (i.e., $P(A \mid B) = P(A)$ e $P(B \mid A) = P(B)$, quando $P(A)P(B) > 0$).
A definição de independência é, no entanto, a seguinte:

Definição

Dois acontecimentos A e B dizem-se independentes se

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

Observações:

1. Não confundir acontecimentos independentes com disjuntos.
2. Se A e B são acontecimentos independentes e tais que $P(A) > 0$ e $P(B) > 0$, tem-se que

$$P(A \mid B) = P(A) \;\text{ e }\; P(B \mid A) = P(B),$$

o que condiz com a ideia intuitiva de independência acima indicada.

De uma forma geral, dizemos que n acontecimentos são independentes se, para quaisquer r desses acontecimentos, com $2 \leq r \leq n$, a probabilidade da intersecção dos r acontecimentos é igual ao produto das respectivas probabilidades.

Por exemplo, três acontecimentos A, B e C, dizem-se independentes se satisfazem as seguintes condições:

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

ii) $P(A \cap C) = P(A)\,P(C)$

iii) $P(B \cap C) = P(B)\,P(C)$

iv) $P(A \cap B \cap C) = P(A)\,P(B)\,P(C)$

Exemplo/Exercício: Na experiência que consiste em efetuar dois lançamentos consecutivos de uma moeda equilibrada, os seguintes acontecimentos A, B e C são independentes?

A: “saiu cara primeiro lançamento”,

B: “saiu coroa no segundo lançamento”,

C: “sairam duas faces distintas”.

Solução: Não! (mas são independentes 2 a 2...)

4. Árvores de probabilidade

Se B for um acontecimento com $P(B) > 0$, a partir da probabilidade condicionada, $P(A|B)$, chega-se à fórmula que permite obter a probabilidade da interseção

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

que pode ser facilmente generalizada à intersecção de n acontecimentos sendo, nesse caso, conhecida como regra da multiplicação.

Regra da Multiplicação

Sejam $A_1, A_2, \ldots, A_n$ acontecimentos. Se $P\left(\bigcap_{i=1}^{n-1} A_i\right) > 0$ então

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2)\ldots P(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}).$$

(TPC) Demonstre o resultado, tendo o cuidado que verificar que todas as probabilidades condicionadas estão bem definidas.

4. Árvores de probabilidade

Exemplo/Exercício$^{(*)}$: No concurso do totoloto (7 extracções, sucessivas e sem reposição, de uma urna com 49 bolas, numeradas de 1 a 49) qual é a probabilidade de não sair a bola numerada com 1?

Resolução: Se considerarmos os acontecimentos

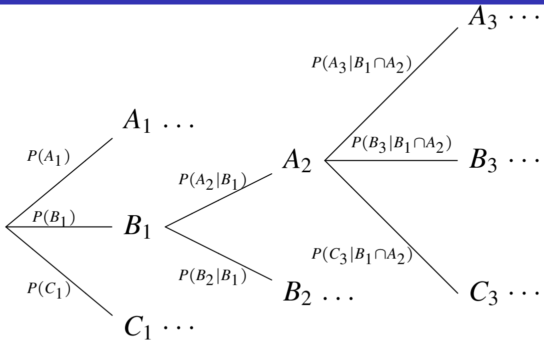
A_i: "na i-ésima extracção não saiu a bola numerada com 1", $\quad i = 1, \ldots, 7,$

a probabilidade pedida é

$$P\left(\bigcap_{i=1}^{7} A_i\right).$$

Pela regra de multiplicação, tem-se:

$$\begin{aligned}
P\left(\bigcap_{i=1}^{7} A_i\right) &= P(A_1) \times P(A_2|A_1) \times P(A_3|A_1 \cap A_2) \times P(A_4|A_1 \cap A_2 \cap A_3) \\
&\times P(A_5|A_1 \cap A_2 \cap A_3 \cap A_4) \times P(A_6|A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5) \\
&\times P(A_7|A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5 \cap A_6) \\
&= \frac{48}{49} \times \frac{47}{48} \times \frac{46}{47} \times \frac{45}{46} \times \frac{44}{45} \times \frac{43}{44} \times \frac{42}{43} \\
&= \frac{42}{49}
\end{aligned}$$



Em cada ramo da árvore tem-se a probabilidade do acontecimento indicado no nó à direita desse ramo, condicional à intersecção de todos os acontecimentos que surgem no caminho da raíz até esse nó (para os ramos com origem na raíz, convenciona-se que esta intersecção é o espaço amostral Ω_1). Observe que, em cada ramificação, a soma das probabilidades é naturalmente igual a 1.

Exemplo/Exercício: Desenhe a árvore de probabilidade adequada ao problema do totoloto $(*)$ e identifique o caminho nela percorrido que lhe permite calcular a probabilidade pedida.

Exemplos: Considere as seguintes experiências aleatórias:

a) Lançamento de uma moeda equilibrada duas vezes consecutivas;
b) Lançamento de uma moeda seguido do lançamento de um dado, ambos equilibrados.

Observe que os espaços amostrais destas experiências são, na verdade, produtos cartesianos de conjuntos muito simples:

a) $\Omega_a = M \times M$, em que $M = \{Ca, Co\}$. De facto:

$$M \times M = \{(Ca, Ca), (Ca, Co), (Co, Ca), (Co, Co)\}.$$

E facilmente concluímos que $\#(\Omega_a) = \#(M) \times \#(M) = 2 \times 2 = 4$.

b) $\Omega_b = M \times D$, em que $M = \{Ca, Co\}$ e $D = \{1, 2, 3, 4, 5, 6\}$. De facto:

$$\begin{aligned} M \times D &= \{(Ca, 1), (Ca, 2), (Ca, 3), (Ca, 4), (Ca, 5), (Ca, 6), \\ &\quad (Co, 1), (Co, 2), (Co, 3), (Co, 4), (Co, 5), (Co, 6)\}. \end{aligned}$$

E facilmente concluímos que $\#(\Omega_b) = \#(M) \times \#(D) = 2 \times 6 = 12$.

Exemplo/Exercício:

Quantas matrículas automóveis (distintas) se podem formar com o atual sistema português: 2 letras, 2 algarismos e novamente 2 letras (as letras são de a a z e incluem k, w e y)? E qual é a probabilidade de uma matrícula ser formada apenas por algarismos pares? E qual a probabilidade de uma matrícula ter pelo menos um algarismo ímpar?

Solução: Observe que o número de matrículas corresponde ao cardinal do seguinte produto cartesiano de conjuntos:

$$L \times L \times A \times A \times L \times L, \tag{7}$$

em que $L = \{a, b, \ldots, k, \ldots, w, x, y, z\}$ e $A = \{0, 1, \ldots, 9\}$. Supondo que todas as configurações de letras e de algarismos podem ser utilizadas, uma matrícula é simplesmente uma sequência formada por 6 elementos, $(a_1, a_2, a_3, a_4, a_5, a_6)$, em que a_1, a_2, a_5 e a_6 podem ser quaisquer elementos do conjunto L e a_3 e a_4 podem ser quaisquer elementos do conjunto A. Temos assim que o número total de matrículas é de

$$26 \times 26 \times 10 \times 10 \times 26 \times 26 = 45697600$$

.

Exemplo/Exercício(cont.):
Usando a definição de Laplace, a probabilidade de uma matrícula ser formada apenas por algarismos pares é:

$$\frac{\#(L \times L \times P \times P \times L \times L)}{\#(\Omega)},$$

em que P é o conjunto formado pelos algarismos pares, i.e., $P = \{0, 2, 4, 6, 8\}$, e Ω é o conjunto (7) atrás indicado. A probabilidade pedida é assim de

$$\frac{26 \times 26 \times 5 \times 5 \times 26 \times 26}{26 \times 26 \times 10 \times 10 \times 26 \times 26} = \frac{1}{2 \times 2} = \frac{1}{4}.$$

E a probabilidade de uma matrícula ter pelo menos um algarismo ímpar é então igual a $\frac{3}{4}$ (note-se que os acontecimentos “matrícula é formada apenas por algarismos pares” e “matrícula tem pelo menos um algarismo ímpar” são complementares).

[TPC] E a probabilidade de uma matrícula ter as letras todas iguais?

Quando efetuamos o produto cartesiano de r conjuntos todos iguais, isto é, $A_1 = A_2 = \ldots = A_r = S$, com S um conjunto que tem n elementos (S diz-se uma população), obtemos n^r sequências ordenadas com r elementos. Tais sequências também são designadas por **amostras ordenadas com reposição** (ou *com repetição*) **de dimensão** r.

Se pretendermos o número de **amostras ordenadas sem reposição de dimensão** r, com $r \leq n$, podemos aplicar novamente o princípio básico da análise combinatória, em que A_1 é o conjunto S (que tem n elementos), A_2 é o conjunto formado por todos os elementos de S excepto o que foi escolhido para o 1^{o} elemento da amostra (e este conjunto tem $n-1$ elementos), ..., e finalmente A_r é o conjunto formado por todos os elementos de S excepto os $r-1$ anteriormente escolhidos para a amostra (e este conjunto tem $n-(r-1)$ elementos). Concluímos assim que o número total de amostras agora pretendidas é dado por

$$n \times (n-1) \times (n-2) \times \ldots \times (n-(r-1)) = \frac{n!}{(n-r)!}, \tag{8}$$

em que $n!$ (lê-se “n factorial”) é igual a $n \times (n-1) \times (n-2) \times \ldots \times 1$.

Resumindo/Recordando

Seja S um conjunto (ou população) com n elementos.

1) O número de amostras ordenadas, formadas por r elementos retirados do conjunto S, com reposição, é igual a n^r. Tais amostras são conhecidas na literatura por *arranjos com repetição de* n, r *a* r.
2) O número de amostras ordenadas, formadas por r elementos retirados de S do conjunto S, sem reposição, é igual a

$$\frac{n!}{(n-r)!}, \quad 1 \leq r \leq n.$$

Tais amostras são conhecidas na literatura por *arranjos sem repetição de* n, r *a* r. No caso particular em que $r = n$, temos as chamadas *permutações de* n, que são em número iguais a $n!$.

Exemplo/Exercício: Usando letras da palavra $ALUNO$, quantas palavras (com ou sem sentido), de tamanho

i. 5, podemos formar se usarmos todas as 5 letras?
ii. 3, podemos formar se usarmos apenas letras distintas (e, portanto,não há repetição de letras na palavra formada)?
iii. 3, podemos formar se permitirmos a reposição de letras (e, portanto, pode haver letras repetidas na palavra formada)?

Solução:

i. $5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$ (permutações de 5)
ii. $5 \times 4 \times 3 = \frac{5!}{(5-3)!} = \frac{5!}{2!} = 60$ (arranjos sem repetição de 5, 3 a 3)
iii. $5^3 = 125$ (arranjos com repetição de 5, 3 a 3)

Concluímos assim que o **número de subconjuntos** pretendido (em que a ordem que os elementos ocupam não é relevante) é igual ao **número de arranjos sem repetição de** 4**,** 3 **a** 3, **dividido por** $3!$, isto é,

$$\frac{4!/(4-3)!}{3!} = \frac{4!}{(4-3)!3!} = \frac{4 \times 3 \times 2 \times 1}{1!(3 \times 2 \times 1)} = 4.$$

De uma forma geral, temos que a partir de **um único subconjunto formado por** r **elementos** (distintos), retirados do conjunto S, **é possível obter** $r!$ **amostras ordenadas** formadas pelos elementos desse mesmo subconjunto (e portanto, nessas amostras ordenadas não há elementos repetidos). Concluímos assim que o número de tais subconjuntos é igual ao **número de arranjos sem repepetição de** n**,** r **a** r, **dividido por** $r!$, ou seja, igual a

$$\frac{n!}{(n-r)!r!}.$$

Este número é usualmente designado de *combinações de* n*,* r *a* r, e é denotado por $\binom{n}{r}$.

Resumindo/Recordando

Seja S um conjunto (ou população) com n elementos. O número de subconjuntos formados por r elementos (distintos) escolhidos em S, com $0 \leq r \leq n$, é igual a

$$\binom{n}{r} \equiv \frac{n!}{(n-r)!r!},$$

e é conhecido na literatura como *combinações de* n, r *a* r.
Tais subconjuntos também são conhecidos na literatura como *amostras não ordenadas sem reposição de* n, r *a* r.

Notas:

1. Estas combinações surgem no desenvolvimento do Binómio de Newton: para quaisquer a e b números reais e para qualquer $n \in \mathbb{N}$,

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

2. É fácil ver que $\binom{n}{r} = \binom{n}{n-r}$. [TPC - Interpretar o resultado.]