Probabilité et Simulation

PolyTech INFO4 (Grenoble) – TD

References

- Fundamentals: [1] B. Jourdain, *Probabilités et statistique pour l'ingénieur*. 2018.

- Further reading: [2] P. Billingsley, *Probability and Measure*. John Wiley & Sons, 2012.

Websites

- CM: https://github.com/jonatha-anselmi/INFO4-PS

- TD: https://github.com/davidelegacci/probasim24

Contents

1. Theory recap - Probability on finite set

1.1. Probability space

- *Random process*: one random outcome out of finitely many

- *Sample space* Ω = finite set of all possible outcomes ω

- *Probability* on Ω = set of weights $\mathbb{P}(\omega) \in \mathbb{R}$ on each $\omega \in \Omega$ such that
  - $\mathbb{P}(\omega) > 0 \forall \omega \in \Omega$
  - $\sum_{\omega \in \Omega} \mathbb{P}(\omega) = 1$

- *Event* $A \subseteq \Omega$ = subset of the sample space

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega) \tag{1}$$

- *Complementary event* $A^c = \Omega/A$ (pronounced "*not A*")
- "*A and B*" $= A \cap B$
- "*A or B*" $= A \cup B$

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) \tag{2}$$

- Probability of complementary event

$$\Omega = A^c \sqcup A \Rightarrow 1 = \mathbb{P}(A^c) + \mathbb{P}(A) \tag{3}$$

- *Indicator function* of event $A \subseteq \Omega$

$$\mathbb{1}_A : \Omega \to \mathbb{R}, \quad \mathbb{1}_A(\omega) = \begin{cases} 1 \text{ if } \omega \in A \\ 0 \text{ if } \omega \notin A \end{cases} \tag{4}$$

- Denoted by $|S|$ the *cardinality* of a set S

- Denoted by S^n the *cartesian product* of S with itself n times

$$S^n = S \times S \times ... \times S = \{(s_1, ..., s_n) : s_i \in S \text{ for all } i = 1, ..., n\} \tag{5}$$

- *Cardinality of cartesian product*

$$|S^n| = |S|^n \tag{6}$$

1.2. Uniform probability

- *Every outcome $\omega \in \Omega$ has the same weight*

$$\mathbb{P}(\omega) = \frac{1}{|\Omega|} \tag{7}$$

- *Uniform probability of the event $A \subseteq \Omega$*

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega) = \sum_{\omega \in A} \frac{1}{|\Omega|} = \frac{|A|}{|\Omega|} \tag{8}$$

1.3. Counting

- Number of *permutations* of k elements:
  - ‣ Number of ways to *order k elements*
  - ‣ **Only order matters**

$$P_k = k!$$ (9)

Permutation of 5 elements

1	5	4	3	2

- Number of *dispositions* of k elements out of n $(k \leq n)$:
  - ‣ Number of ways to *choose and order k elements out of n*
  - ‣ **Order and elements** matter
  - ‣ Number of injections : $\{1, ..., k\} \rightarrow \{1, ..., n\}$

$$D_{n,k} = \underbrace{n(n-1)...}_{k \text{ times}} = n(n-1)...(n-k+1) = \frac{n!}{(n-k)!}$$ (10)

Disposition of 3 elements out of 5

	2	1	3	

- Number of *combinations* of k elements out of n $(k \leq n)$:
  - ‣ Number of ways to *choose k elements out of n*
  - ‣ **Only elements** matter
  - ‣ Number of subsets of cardinality k of a set of cardinality n
  - ‣ Number of dispositions modulo number of permutations

$$C_{n,k} = \frac{D_{n,k}}{P_k} = \frac{n!}{k!(n-k)!} = \binom{n}{k} = \text{choose}(n, k)$$ (11)

Combination of 3 elements out of 5

		X	X	X

- *Binomial theorem*

$$(x + y)^n = \sum_{j=0}^{n} \binom{n}{j} x^{n-j} y^j$$ (12)

Exercises

Exercise 1.1 *(Handshakes and kisses)*

There are f girls and g boys in a room. Boys exchange handshakes, girs exchange kisses, boys and girls exchange kisses. How many kisses in total?

The number of kisses exchanegs among girls is the number of subsets of cardinality 2 of a set of cardinality f, that is $\binom{f}{2} = \frac{f(f-1)}{2}$. Or, think that each girl gives $f-1$ kisses, and one needs a factor of one half to avoid double counting.

For the number of kisses exchanged between boys and girls: the first girl gives g kisses, the second girl gives g kisses, and so on, so we have fg in total.

$$\text{number of kisses} = \frac{f(f-1)}{2} + fg \tag{13}$$

Exercise 1.2 *(Throwing a dice)* Throw a fair dice with f faces n times. What's the prob to never get the same result twice?

General strategy
- Identify sample space Ω (write in set-theoretic notation!) and its cardinality $|\Omega|$
- Identify event $A \subseteq \Omega$ (write in set-theoretic notation!) and its cardinality $|A|$
- Uniform probability? If so, use $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$

Let $\mathcal{N} = \{1, ..., n\}$ and $\mathcal{F} = \{1, ..., f\}$. The sample space is

$$\Omega = \{\omega = (\omega_1, ..., \omega_n) : \omega_i \in \mathcal{F} \text{ for all } i \in \mathcal{N}\} = \mathcal{F}^n \tag{14}$$

with cardinality

$$|\Omega| = |\mathcal{F}^n| = |\mathcal{F}|^n = f^n \tag{15}$$

Endow the sample space with the uniform probability (since every outcome of the experiment is equiprobable).

The event we're looking at is

$$A = \{\omega \in \Omega : \omega_i \neq \omega_j \text{ for all } i \neq j \in \mathcal{N}\} \tag{16}$$

Clearly if $n > f$ then $\mathbb{P}(A) = 0$. Let $n \leq f$. The (uniform) probability of the event A is $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$, with

$$|A| = \# \text{ of ways to choose and order } n \text{ elements out of } f$$

$$= \underbrace{f(f-1)...}_{n} = f(f-1)...(f-n+1) = \frac{f!}{(f-n)!} \tag{17}$$

$$\mathbb{P}(A) = \frac{f!}{f^n(f-n)!} \tag{18}$$

Exercise 1.3 *(Birthday paradox)* What is the probability that at least 2 people out of n have the same birthday? (Assume: uniform birth probability and year with y number of days).

Quick solution

$$\mathbb{P}(A) = 1 - \mathbb{P}\left(\underbrace{\text{no two people have the same birthday}}_{\text{Ex. 2}}\right) \tag{19}$$

$$= 1 - \frac{y!}{y^n (y-n)!}$$

Formal solution Let $\mathcal{N} = \{1, ..., n\}$ and $\mathcal{Y} = \{1, ..., y\}$ with $n \leq y$. The sample space is

$$\Omega = \text{distributions of possible birthdays of } n \text{ people}$$
$$= \{\omega = (\omega_1, ..., \omega_n) : \omega_i \in \mathcal{Y} \text{ for all } i \in \mathcal{N}\} = \mathcal{Y}^n \tag{20}$$

where ω_i is the birthday of the i-th person. The cardinality of the sample space is

$$|\Omega| = |\mathcal{Y}^n| = |\mathcal{Y}|^n = y^n \tag{21}$$

The event we're looking at is

$$A = \{\omega \in \Omega : \exists i \neq j \in \mathcal{N} : \omega_i = \omega_j\} \tag{22}$$

Note that this is the complementary event to the event defined in Equation 16 of Exercise 2. Thus we can compute its probability as

$$\mathbb{P}(A) = 1 - \mathbb{P}(A^c) \tag{23}$$
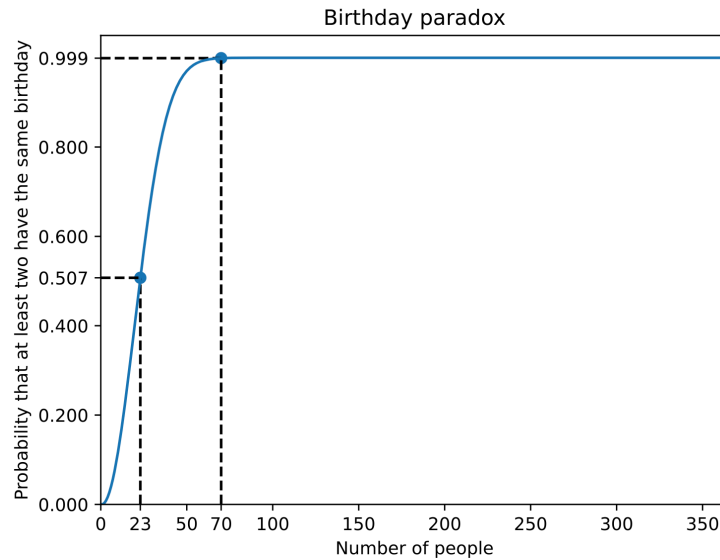
in agreement with Equation 19.

Figure 4: Birthday paradox probability. <u>Code available.</u>

Exercise 1.4 *(Same birthday as the prof)* What is the probability that at least 1 student out of n has the same birthday of the prof? (Assume: uniform birth probability and year with y number of days).

Quick solution

$$\mathbb{P}(A) = 1 - \mathbb{P}\left(\underbrace{\text{nobody has the prescribed birth date}}\right)$$
$$= 1 - \left(\frac{y-1}{y}\right)^n \tag{24}$$

Formal solution 1 As above $\mathcal{N} = \{1, ..., n\}$ and $\mathcal{Y} = \{1, ..., y\}$ with $n \leq y$. The sample space is $\Omega = \mathcal{Y}^{n+1}$, that is the set of possible birthdays of $n + 1$ people, the $(n + 1)$-th being the prof. Its cardinality is $|\Omega| = y^{n+1}$. The event we're looking at is

$$A = \left\{\omega \in \Omega : \exists i \in \mathcal{N} : \omega_i = \omega_{n+1}\right\} \tag{25}$$

with complementary event

$$A^c = \left\{\omega \in \Omega : \omega_i \neq \omega_{n+1} \, \forall i \in \mathcal{N}\right\} \tag{26}$$

As usual $\mathbb{P}(A) = 1 - \mathbb{P}(A^c) = 1 - \frac{|A^c|}{|\Omega|}$, with

$$|A^c| = \underbrace{y}_{\text{prof}} \cdot \underbrace{(y-1)^n}_{\text{students}} \tag{27}$$

So, $\mathbb{P}(A) = 1 - \frac{y(y-1)^n}{y^{n+1}} = 1 - \left(\frac{y-1}{y}\right)^n$, in agreement with Equation 24.

Note that a factor $\frac{y}{y}$, corresponding to the prof's birthday, simplifies in the last step. Alternatively, you can fix the birthday of the prof and exclude it from the analysis from the beginning. In this case $|\Omega| = y^n$ and $|A^c| = (y-1)^n$, leading to the same result.

Formal solution 2 Using the probability of the complementary event is often the smartest way to proceed, but for the sake of completeness let's see how to get the same result directly. Let A_j be the event *"exactly j students out of n have the same birthday as the prof"*. The event we look at then is

$$A = \sqcup_{j \in \mathcal{N}} A_j \tag{28}$$

with probability (cf Equation 2)

$$\mathbb{P}(A) = \sum_{j \in \mathcal{N}} \mathbb{P}(A_j) = \frac{\sum_{j \in \mathcal{N}} |A_j|}{|\Omega|} \tag{29}$$

The cardinality of A_j is

$$|A_j| = \underbrace{1...1}_{j \text{ times}} \cdot \underbrace{(y-1)...(y-1)}_{n-j \text{ times}} \cdot \underbrace{y}_{\text{prof}} \cdot \underbrace{\binom{n}{j}}_{\text{number of ways to choose } j \text{ elements out of } n} \tag{30}$$

$$= y(y-1)^{n-j} \binom{n}{j}$$

By an application of the binomial theorem (Equation 12) and a short manipulation,

$$\sum_{j=1}^{n} |A_j| = y\left(y^n - (y-1)^n\right) \tag{31}$$

which leads back to Equation 24.

2. Theory recap - Conditional probability and independence

2.1. Conditional probability
- Let $\mathbb{P}$ be a probability on Ω, and consider the events $A, B \subseteq \Omega$.
- *Conditional probability*: probability of *A given B*

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \text{ if } \mathbb{P}(B) \neq 0 \tag{32}$$

- not reallty defined if $\mathbb{P}(B) = 0$, cf [2] pag. 427.
- often used as

$$\mathbb{P}(A \cap B) = \mathbb{P}(A \mid B)\mathbb{P}(B) \tag{33}$$

- Conditional probability and complementary event (proof: simple exercise.)

$$\mathbb{P}(A \mid B) + \mathbb{P}(A^c \mid B) = 1 \tag{34}$$

- *Total Probability Theorem*

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c) = \mathbb{P}(A \mid B)\mathbb{P}(B) + \mathbb{P}(A \mid B^c)\mathbb{P}(B^c) \tag{35}$$

- *Bayes theorem*

$$\mathbb{P}(A \cap B) = \mathbb{P}(B \cap A) \Rightarrow \mathbb{P}(A \mid B)\mathbb{P}(B) = \mathbb{P}(B \mid A)\mathbb{P}(A) \tag{36}$$

See this notebook for an example of Bayes theorem in action.

2.2. Independent events

Let Ω be equipped with a probability $\mathbb{P}$.
- two events $A, B \subseteq \Omega$ are said *independent* if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) \tag{37}$$

- equivalently, by definition of conditional probability, A and B are independent if

$$\mathbb{P}(A \mid B) = \mathbb{P}(A) \tag{38}$$

- n events $A_1, ..., A_n$ are said *independent* if

$$\mathbb{P}(\cap_{i \in I} A_i) = \prod_{i \in I} \mathbb{P}(A_i) \text{ for all } I \subseteq \{1, ..., n\} \tag{39}$$

- pairwise independence does not imply independence of all events!

Bibliography

[1] B. Jourdain, *Probabilités et statistique pour l'ingénieur.* 2018.

[2] P. Billingsley, *Probability and Measure.* John Wiley & Sons, 2012.