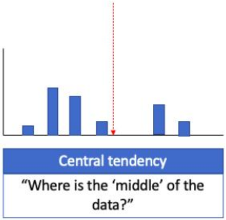
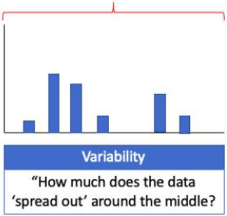
COMP229: Introduction to Data Science
Lecture 4: Probabilities, exclusivity and independence.

Olga Anosova, O.Anosova@liverpool.ac.uk
Autumn 2023, Computer Science department
University of Liverpool, United Kingdom

Reminder: location and spread

Measurements of location and spread:





mean (several types), median and mode,

variance, standard deviation, IQR, outliers.

Talking about probabilities

- *What's the probability to meet a dinosaur walking down the street?*
- *50%*
- *Why?*
- *Either you meet a dinosaur, or you don't.*

Conclusions like this can be avoided (or justified) by applying a rigorous theory. Axioms of probability theory were introduced by A. Kolmogorov in 1933.

Important questions: what are all possible events and how can we define their probabilities?

Probability space $(\Omega, \mathcal{E}, P)$

Definition 4.1. A *probability space* has 3 parts:

1. a **sample space** Ω = set of **elementary events** (all possible outcomes),
2. an **event space** $\mathcal{E}$ = set of all possible combinations of elementary events,
3. a **probability** $P : \mathcal{E} \to [0, 1]$ = function that satisfies **Kolmogorov's axioms**

A1: positivity $P(A) \geqslant 0$ for any event $A \in \mathcal{E}$

A2: something always happens $P(\Omega) = 1$

A3: additivity $P(\bigsqcup_{n=1}^{+\infty} A_n) = \sum_{n=1}^{+\infty} P(A_n)$ for disjoint A_n.

Flipping a coin

Problem 4.2. Flipping a coin that will either land heads (H) or tails (T). Then $\Omega = \{H, T\}$ consists of 2 elementary events. What is an event space $\mathcal{E}$?

Solution 4.2. When Ω is finite, $\mathcal{E}$ is often the set of all subsets of Ω, i.e. $\mathcal{E} = \{\emptyset, \{H\}, \{T\}, \{H, T\}\}$. The empty set symbol $\emptyset$ means an **empty set** with no elements, i.e. impossible events (the coin landing on its edge, or turning into a pink elephant). The set $\{H, T\} = \Omega$ means that the coin lands.

A coin is *fair* if $P(\{H\}) = 0.5 = P(\{T\})$.

Flipping a fair coin twice

Problem 4.3. Flip a fair coin twice. Find the probability to get at least one head.

Solution 4.3.

$P(\text{head at least once}) = P(\text{head on 1st flip}) + P(\text{head on 2nd flip}) = 0.5 + 0.5 = 1.$

Which is wrong as there is a non-zero possibility of getting both tails. Where's the catch?

The A3 axiom of addition is only for disjoint events!

Mutually exclusive events

Definition 4.4. Events A_n are called **mutually exclusive** (disjoint, non-intersecting) if any two of these events A_i, A_j can not happen simultaneously.

For any event A, the **complementary** event $\Omega - A = \overline{A} = A^C = \neg A$ consists of all elementary events that are mutually exclusive with A.

For one flip of a coin, the 'head' event $\{H\}$ and the 'tail' event $\{T\}$ are mutually exclusive and are complementary to each other.

Useful facts

Claim 4.5. For a complementary event $P(\neg A) = 1 - P(A)$.

Proof. $(\Omega - A) \cup A = \Omega$, A and $\Omega - A$ are mutually exclusive by definition.

Then by A3, A2 $P(\Omega - A) + P(A) = P(\Omega) = 1$.

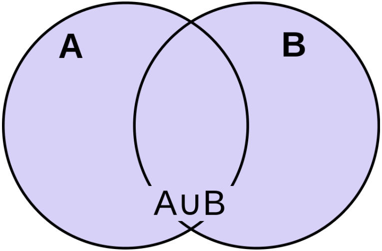
Claim 4.6. $P(\varnothing) = 0$.

Proof. By contradiction: suppose $P(\varnothing) = p > 0$ (by A1 p can not be < 0). $\varnothing$ is mutually exclusive with every other set. Then by A3, A2 $P(\Omega) + P(\varnothing) = 1 + p = 1$, which contradicts the initial assumption of $p > 0$, so $P(\varnothing) = 0$.

The general sum rule

$P(A \cup B) = P(A) + P(B)$ only for mutually exclusive A, B.

In general case, use inclusion-exclusion principle:
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ for any A, B.



The picture on the left illustrates the sum rule: imagine you compute the area of $A \cup B$ from the areas A, B and $A \cap B$.

Correctly flipping a coin twice

Solution 4.3. All 4 elementary events $\{HH\}, \{HT\}, \{TH\}, \{TT\}$ are equally likely with probabilities of 0.25 each.

Using the sum rule: $P(\text{head at least once}) =$
$= P(\text{head on 1st flip} \cup \text{head on 2nd flip}) =$
$= P(HT) + P(TH) - P(HH) = 1 - 0.25 = 0.75$

Alternatively, add the 3 elementary events:
$P(\text{head at least once}) = P(HT) + P(TH) + P(HH) = 0.25 + 0.25 + 0.25 = 0.75$

Independent events

Is there a way to find the intersection (both heads) from the knowledge about a single coin?

Sometimes, yes.

Definition 4.7. Events A, B are called **independent** iff $P(A \cap B) = P(A)P(B)$.

The above formula is confusingly called the **product rule**, because in practice one says: A, B are assumed to be independent, hence $P(A \cap B) = P(A)P(B)$.

Notable cases of independence

Can an event A be independent with itself?

If yes, then $P(A) = P(A \cap A) = (P(A))^2$, hence $P(A) = 0$ or $P(A) = 1$, so $A = \emptyset$ or $A = \Omega$.

Can A be independent with its complementary $\neg A$?

Then $P(A \cap \neg A) = P(\emptyset) = 0 = P(A)P(B)$. It's possible iff $A = \emptyset$ or $A = \Omega$.

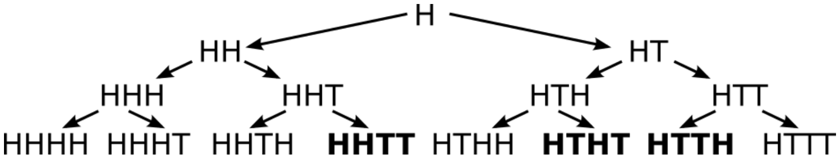
Are mutually exclusive events independent?

By the definition of mutually exclusive events, $P(A \cap B) = P(\emptyset) = 0$. When is it equal to $P(A)P(B)$? Iff at least one of A, B is impossible!

More fair coin flips

Problem 4.8. Find P(2 heads in 4 coin flips).

Solution 4.8. The 2-head event consists of 6 basic events, here's half of the tree starting with H:



Each of $2^4 = 16$ events has probability $\frac{1}{16}$ since all flips are independent. All 16 events are mutually exclusive. The sum rule gives $\frac{6}{16} = \frac{3}{8} = 37.5\%$.

More consequences

Claim 4.9. $0 \leqslant P(A) \leqslant 1$ for any event A.

Proof. From A1, $0 \leqslant P(A)$.

From A1 and the complement rule

$P(\neg A) = 1 - P(A) \geqslant 0$, hence $P(A) \leqslant 1$.

Claim 4.10. Monotonicity: if $A \subset B$, then

$P(A) \leqslant P(B)$.

Proof. Since $B = A \cup (B - A)$, A and $B - A$ are mutually exclusive and $P(B - A) \geqslant 0$,

$P(B) = P(A) + P(B - A) \geqslant P(A)$.

Happy Weekend

Problem 4.11. The probability of rain on Saturday is 25%, on Sunday it is (independently) 25%. Is the probability of rain during the weekend 50% or less?

A popular question in the last year exam: what is a weekend? It is the union of 2 days: Saturday and Sunday. A condition of independence should be explicitly stated or explicitly assumed in a solution.

Rain during the weekend means rain on Saturday or rain on Sunday: rain at least on one of two days.

Two solutions

Solution 4.11. Use the complementary event:
$P(\text{rain during weekend}) =$
$= 1 - P(\text{no rain at the weekend}) = 1 -$
$P(\text{no rain on Saturday})P(\text{no rain on Sunday}) =$
$1 - \frac{3}{4} \cdot \frac{3}{4} = 1 - \frac{9}{16} = \frac{7}{16} < \frac{1}{2} = 50\%.$

Or use the sum rule:

Two solutions

Solution 4.11. Use the complementary event:
$P(\text{rain during weekend}) =$
$= 1 - P(\text{no rain at the weekend}) = 1 -$
$P(\text{no rain on Saturday})P(\text{no rain on Sunday}) =$
$1 - \frac{3}{4} \cdot \frac{3}{4} = 1 - \frac{9}{16} = \frac{7}{16} < \frac{1}{2} = 50\%.$

Or use the sum rule: $P(\text{rain on Sat or Sun}) =$
$P(\text{rain on Sat}) + P(\text{rain on Sun}) -$
$P(\text{rain on Sat and Sun}) = \frac{1}{4} + \frac{1}{4} - \frac{1}{4} \cdot \frac{1}{4} = \frac{7}{16}.$

Mutual independence of 2+ events

Definition 4.12. For multiple events, events A_i are **pairwise independent** if every pair of events is independent, that is $P(A_i \cap A_j) = P(A_i)P(A_j)$ for all distinct i, j.

Definition 4.13. Events $A_1, \ldots, A_n$, are **mutually independent** if $P(\bigcap_{i=1}^{k} B_i) =$ product $\prod_{i=1}^{k} P(B_i)$ for any $2 \leqslant k \leqslant n$ events $B_1, \ldots, B_k \in \{A_1, \ldots, A_n\}$.

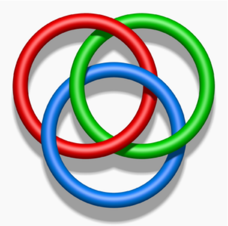
Multiple independence

If $n = 3$, the mutual independence means that the 4 conditions hold:

$$P(A_1 \cap A_2) = P(A_1)P(A_2),$$
$$P(A_2 \cap A_3) = P(A_2)P(A_3),$$
$$P(A_3 \cap A_1) = P(A_3)P(A_1),$$
$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3).$$

Mutual vs pairwise independence

Problem 4.14. (discussed in tutorial 2) A binary number machine gives 4 equally likely outcomes: (000), (011), (101), (110). Let $A_i = \{1 \text{ on } i\text{-th place}\}$. Are the events A_1, A_2, A_3 independent?



Pairwise independence doesn't imply mutual independence, which is illustrated by the 3 Borromean rings: any two are unlinked.

Independent events in finances

The concept of independent events is central in probability theory. In the modern world almost all events are linked to each other in some way.

The financial crisis in the subprime mortgage market (2008) happened, because (amongst other reasons) potential failures of homeowners to repay their mortgages were assumed to be *independent*. Why not?

Usually, people behave independently, but not in a crisis, see more in the 2015 movie "The Big Short".

Independence & exclusivity: question of life & death

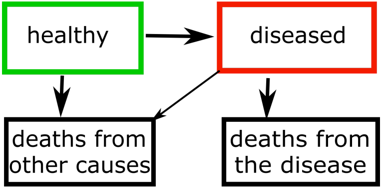
Basic statistical notions can be messed-up in advanced applications: the Global Burden of Disease project relied at some point on DisMod II model to compute disease epidemiology. One of the main assumptions was that *deaths from different diseases happen independently*.

Independence & exclusivity: question of life & death

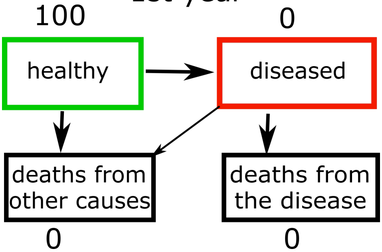
Basic statistical notions can be messed-up in advanced applications: the Global Burden of Disease project relied at some point on DisMod II model to compute disease epidemiology. One of the main assumptions was that *deaths from different diseases happen independently*.

Which, as we know, can be true for mutually exclusive events only if probablity of one of them is zero.

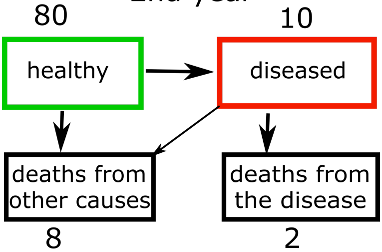
Dynamics of the simple error



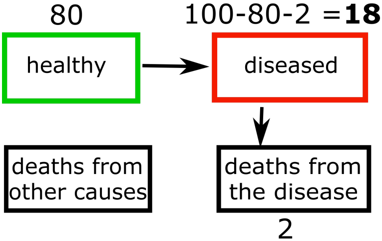
1st year



2nd year



2nd year, estimation



Error legacy is still alive, and kicking!

In a dynamic model this trivial error led to overestimation of disease numbers.

Dismod is discontinued, but models calibrated on it are still very much in use.

Andrew Gregory *Health editor*

@andrewgregory

Tue 25 Jul 2023 00.01 BST

'Catastrophic' forecast shows 9m people in England with major illnesses by 2040

Cases of dementia, diabetes, cancer, depression and kidney disease expected to soar as growing numbers reach old age

The Health Foundation says an extra 2.5 million people will be living with major illnesses by 2040, up 37% compared with 2019. Photograph: Julian Claxton/Alamy

Time to revise and ask questions

To benefit from the lecture, now you could

- ask or submit your questions on CANVAS or after the lecture;
- write down your summary in 2-3 phrases, e.g. list key concepts you have learned;
- talk to your classmates to revise the lecture.

Problem 4.15. Flipping a fair coin, you got 2019 heads. What's the probability to get a head again?

Final solution and summary

Solution 4.15. The probability to get a head by flipping a fair coin is always 0.5, because this event doesn't depend on previous flips.

Final solution and summary

Solution 4.15. The probability to get a head by flipping a fair coin is always 0.5, because this event doesn't depend on previous flips. The probability to get 2020 heads in a row is very small: 0.5^{2020}.

Final solution and summary

Solution 4.15. The probability to get a head by flipping a fair coin is always 0.5, because this event doesn't depend on previous flips. The probability to get 2020 heads in a row is very small: 0.5^{2020}.

- Probabilities can be computed after defining a probability space (Ω, E, P) satisfying 3 axioms.
- The sum rule for any events A, B is $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $P(A \cap B) = P(A)P(B)$ for independent events.