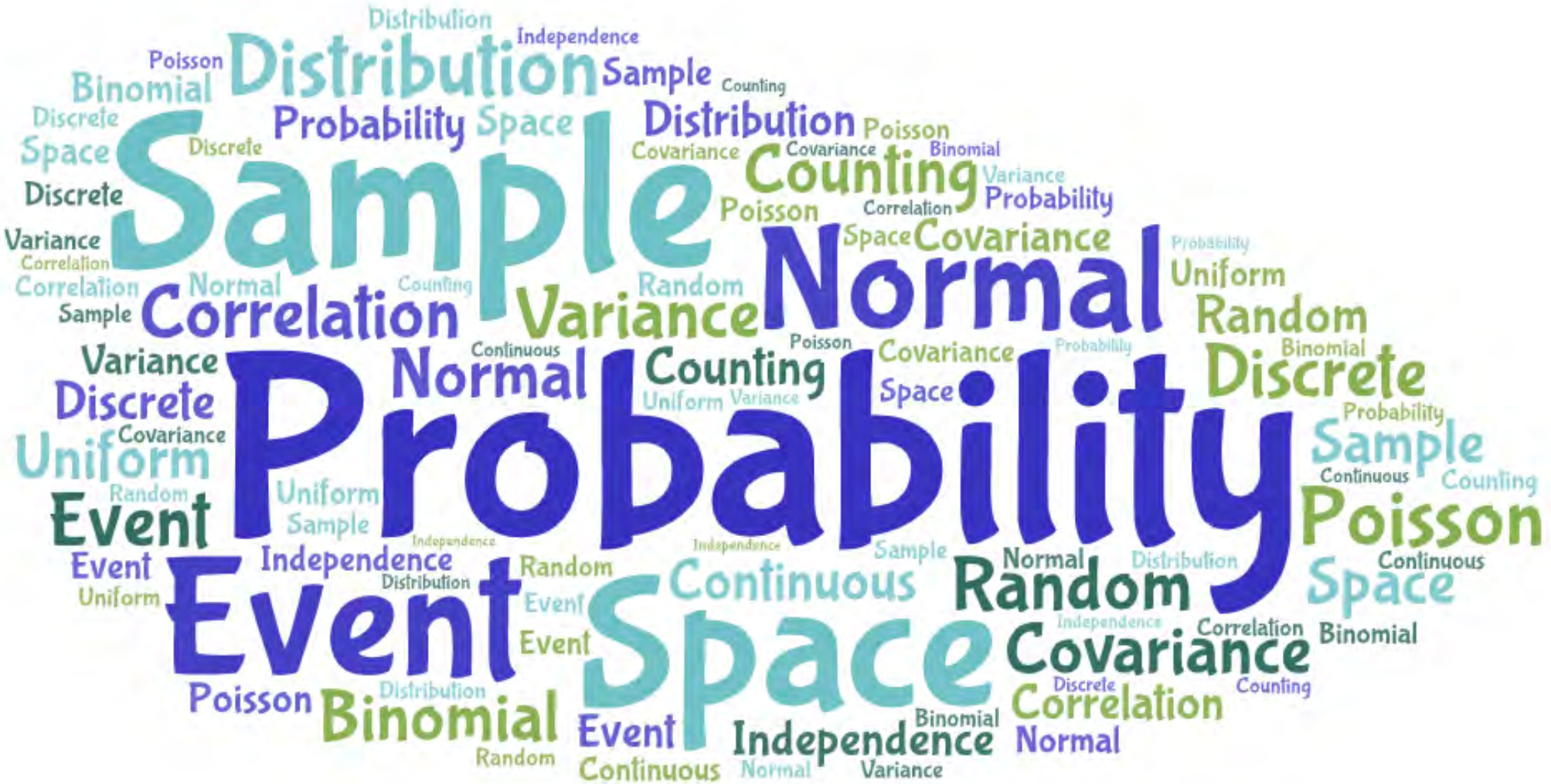
Discrete Mathematics and Probability

Week 7



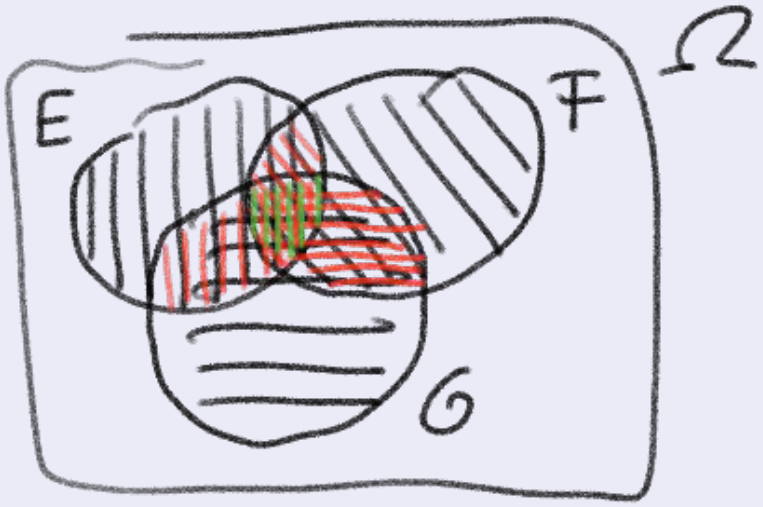
MARKULF Chris Heunen KOHLWEISS

Inclusion-exclusion (3 events)

Proposition

For any events:

$$\begin{aligned}\mathbf{P}(E \cup F \cup G) = \mathbf{P}(E) + \mathbf{P}(F) + \mathbf{P}(G) \\ - \mathbf{P}(E \cap F) - \mathbf{P}(E \cap G) - \mathbf{P}(F \cap G) \\ + \mathbf{P}(E \cap F \cap G).\end{aligned}$$



E, F mutually exclusive

$P(E \cup F) = P(E) + F)$

$E \subset F \quad P(F - E) = P(F) - P(E)$

Example

wooclap

Example

In a sports club, with N members

T	36 members play tennis,	22 play tennis and squash,	$T \cap S$
S	28 play squash,	12 play tennis and badminton,	$T \cap B$
B	18 play badminton,	9 play squash and badminton,	$S \cap B$
	4 play tennis, squash and badminton.		$T \cap S \cap B$

What is probability that a random member ~~How many~~ plays at least one of these games?

$$P(T \cup S \cup B) = P(T) + P(S) + P(B) - P(T \cap S) - P(T \cap B) - P(S \cap B) + P(T \cap S \cap B)$$

$$= \frac{36}{N} + \frac{28}{N} + \frac{18}{N} - \frac{22}{N} - \frac{12}{N} - \frac{9}{N} + \frac{4}{N} = \frac{43}{N}$$

Example

Example

In a sports club, with N members

36 members play tennis,	22 play tennis and squash,
28 play squash,	12 play tennis and badminton,
18 play badminton,	9 play squash and badminton,

4 play tennis, squash and badminton.

What is probability that a random member
~~How many~~ plays at least one of these games?

$$P(T \cup S \cup B) = P(T) + P(S) + P(B) - P(T \cap S) - P(T \cap B) - P(S \cap B) + P(T \cap S \cap B)$$

$$= \frac{36}{N} + \frac{28}{N} + \frac{18}{N} - \frac{22}{N} - \frac{12}{N} - \frac{9}{N} + \frac{4}{N} = \frac{43}{N}$$

Topics

- Recap: examples with equally likely outcomes
- Conditional probability: how knowledge influences probability
- Bayes' theorem: link probabilities of related events

Equally likely outcomes

The return of counting

Finite sample space, $|\Omega| = N < \infty$, has special important case where each experiment outcome has *equal probability*:

$$\mathbf{P}(\omega) = \frac{1}{N} \qquad \text{for all } \omega \in \Omega$$

Definition

Outcomes $\omega \in \Omega$ are also called *elementary events*.

Example

Example

Rolling two dice, what is the probability that the sum of the numbers shown is 7?

What's wrong with this solution? "The number 7 is one out of the possible values $2, 3, \ldots, 12$ for the sum, and the answer is $\frac{1}{11}$."

$\Omega = \{(i,j) \mid i,j \in \{1,\ldots 6\}\}$ $|\Omega| = 36$

what is wrong with the answer?

$E = \text{sum is } 7 = \{(i,j) \mid i+j = 7\} = \{(1,6),(2,5)(3,4)(4,3),(5,2),(1,6)\}$

$P(E) = \frac{6}{30} = \frac{1}{6}$

Example

Example

Rolling two dice, what is the probability that the sum of the numbers shown is 7?

What's wrong with this solution? "The number 7 is one out of the possible values $2, 3, \ldots, 12$ for the sum, and the answer is $\frac{1}{11}$."

sums are not equally likely:
e.g. 12 is only 1 out of 36.

what is wrong with the answer?

Example

Example

Rolling two dice, what is the probability that the sum of the numbers shown is 7?

What's wrong with this solution? "The number 7 is one out of the possible values $2, 3, \ldots, 12$ for the sum, and the answer is $\frac{1}{11}$."

sums are not equally likely: e.g. 12 is only 1 out of 36.

what is wrong with the answer?

$$\Omega = \{(i,j) \mid i,j \in \{1,2,\ldots,6\}\}$$

$$E = \text{sum is } 7 = \{(i,j) \mid i+j=7\} = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$$

$$P(E) = \frac{6}{36} = \frac{1}{6}$$

Conditional probability

Conditional probability

Often you have *partial information* about the outcome of an experiment. This alters the likelihoods for various outcomes.

Example

Roll two dice. What is the probability that the sum of the numbers is 8? What if we know that the first die shows a 5?

$\Omega = \{1,\dots,6\}^2 = \{1,\dots,6\} \times \{1,\dots 6\}$

$E = \text{"sum 8"} = \{(2,6),(3,5),(4,4),(5,3),(6,2)\}$

$P(E) = \frac{|E|}{|\Omega|} = \frac{5}{36}$

what if the first die is 5

$\Omega' = \{(5,1),\dots (5,6)\}$ $\quad |\Omega'| = 6$

$E' = \text{"sum 8"} = \{5,3\}$ $\quad P_{\omega}(E') = \frac{|E'|}{|\Omega'|} = \frac{1}{6}$

Conditional probability

Often you have *partial information* about the outcome of an experiment. This alters the likelihoods for various outcomes.

Example

Roll two dice. What is the probability that the sum of the numbers is 8? What if we know that the first die shows a 5?

$\Omega = \{1, \dots 6\}^2$

$E = "sum = 8" = \{(2,6), (3,5), (4,4), (5,3), (6,2)\}$

Conditional probability

Often you have *partial information* about the outcome of an experiment. This alters the likelihoods for various outcomes.

Example

Roll two dice. What is the probability that the sum of the numbers is 8? What if we know that the first die shows a 5?

$\Omega = \{1, \dots 6\}^2$

$E = \text{"sum} = 8\text{"} = \{(2,6), (3,5), (4,4), (5,3), (6,2)\}$

$P(E) = \frac{|E|}{|\Omega|} = \frac{5}{36}$

what if first die = 5

Conditional probability

Often you have *partial information* about the outcome of an experiment. This alters the likelihoods for various outcomes.

Example

Roll two dice. What is the probability that the sum of the numbers is 8? What if we know that the first die shows a 5?

$\Omega = \{1, \dots 6\}^2$

$E = \text{"sum} = 8\text{"} = \{(2,6), (3,5), (4,4), (5,3), (6,2)\}$

$P(E) = \frac{|E|}{|\Omega|} = \frac{5}{36}$

what if first die = 5

$\Omega' = \{(5,1), \dots (5,6)\}$

$E' = \text{"sum} = 8\text{"} = \{(5,3)\}$

$P(E') = \frac{|E'|}{|\Omega'|} = \frac{1}{6}$

Partial information about experiment changes probability

Reduced sample space

We reduced our world to the event we were given:
$F = \{\text{first die shows } 5\} = \{(5, 1), (5, 2), \ldots, (5, 6)\}$

Definition

The event that is given to us is called a *reduced sample space*. We can simply work in this set to figure out the conditional probabilities given this event.

The event F has 6 equally likely outcomes. Only one of them, $(5, 3)$, provides a sum of 8. Hence the conditional probability is $\frac{1}{6}$.

Definition of conditional probability

The question can be reformulated.

$$E = \{\text{the sum is } 8\} = \{(2, 6), (3, 5), \ldots, (6, 2)\}$$

"In what proportion of cases in F will E also occur?"
"How does probability of 'E and F' compare to probability of F?"

Definition

Let F be an event with $\mathbf{P}(F) > 0$.
The *conditional probability of* E *given* F is:

$$\mathbf{P}(E \mid F) = \frac{\mathbf{P}(E \cap F)}{\mathbf{P}(F)}$$

"probability of E given F"

$E \cap F = \{(5,3)\}$ $\quad P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{\frac{1}{36}}{\frac{1}{6}} = \frac{6}{36} = \frac{1}{6}$

Definition of conditional probability

The question can be reformulated.

$$E = \{\text{the sum is } 8\} = \{(2, 6), (3, 5), \ldots, (6, 2)\}$$

"In what proportion of cases in F will E also occur?"
"How does probability of 'E and F' compare to probability of F?"

Definition

Let F be an event with $\mathbf{P}(F) > 0$.
The *conditional probability of* E *given* F is:

$$\mathbf{P}(E \mid F) = \frac{\mathbf{P}(E \cap F)}{\mathbf{P}(F)}$$

"probability of E given F"

$E \cap F = \{(5,3)\}$ $\quad P(E|F) = \dfrac{P(E \cap F)}{P(F)} = \dfrac{1/36}{1/6} = \dfrac{1}{6}$

Axioms

Proposition

Conditional probability $\mathbf{P}(\cdot \mid F)$ *satisfies the axioms of probability:*

1. *conditional probability is non-negative:* $\mathbf{P}(E \mid F) \geq 0$;
2. *conditional probability of sample space is one:* $\mathbf{P}(\Omega \mid F) = 1$;
3. *for countably many mutually exclusive events* $E_1, E_2, \ldots$:

$$\mathbf{P}\left(\bigcup_i E_i \,\middle|\, F\right) = \sum_i \mathbf{P}(E_i \mid F)$$

How to compute conditional probabilities

Corollary

- $\mathbf{P}(E^c \mid F) = 1 - \mathbf{P}(E \mid F)$
- $\mathbf{P}(\emptyset \mid F) = 0$
- $\mathbf{P}(E \mid F\} = 1 - \mathbf{P}(E^c \mid F) \leq 1$
- $\mathbf{P}(E \cup G \mid F) = \mathbf{P}(E \mid F) + \mathbf{P}(G \mid F) - \mathbf{P}(E \cap G \mid F)$
- *If* $E \subseteq G$, *then* $\mathbf{P}(G - E \mid F) = \mathbf{P}(G \mid F) - \mathbf{P}(E \mid F)$
- *If* $E \subseteq G$, *then* $\mathbf{P}(E \mid F) \leq \mathbf{P}(G \mid F)$

BUT: Don't change the condition!

$\mathbf{P}\{E \mid F\}$ and $\mathbf{P}\{E \mid F^c\}$ have nothing to do with each other.

Really important that the same given F.

Multiplication rule

Proposition (*Multiplication rule*)

$$\mathbf{P}(E_1 \cap \cdots \cap E_n) = \mathbf{P}(E_1) \cdot \mathbf{P}(E_2 \mid E_1) \cdot \mathbf{P}(E_3 \mid E_1 \cap E_2) \cdots \mathbf{P}(E_n \mid E_1 \cap \cdots \cap E_{n-1})$$

$$P(E|F) = \frac{P(E \cap F)}{P(F)} \qquad P(E \cap F) = P(F) \cdot P(E|F)$$

Multiplication rule

Proposition (*Multiplication rule*)

$$\mathbf{P}(E_1 \cap \cdots \cap E_n) = \mathbf{P}(E_1) \cdot \mathbf{P}(E_2 \mid E_1) \cdot \mathbf{P}(E_3 \mid E_1 \cap E_2) \cdots \mathbf{P}(E_n \mid E_1 \cap \cdots \cap E_{n-1})$$

$$RH = P(E_1) \cdot \frac{P(E_1 \cap E_2)}{P(E_1)} \cdot \frac{E(E_3 \cap (E_1 \cap E_2))}{P(E_1 \cap E_2)} \cdots \frac{P(E_1 \cap \cdots \cap E_n)}{P(E_1 \cap \cdots \cap E_{n-1})}$$

Example again

Example

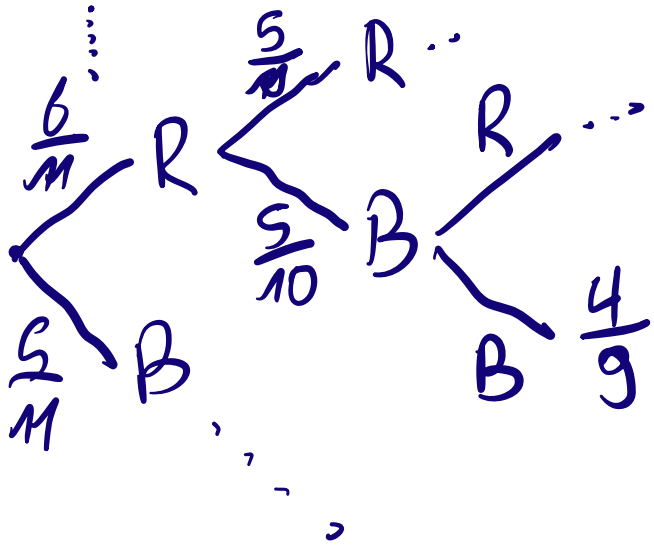
An urn contains 6 red and 5 blue balls. We draw three balls at random, at once (that is, without replacement). What is the chance of drawing one red and two blue balls?

$$P(1\text{ red} + 2\text{ blue}) = P(R_1 \wedge B_2 \wedge B_3) + P(B_1 \wedge R_2 \wedge B_3) + P(B_1 \wedge B_2 \wedge R_3)$$

$$= P(R_1) \cdot P(B_2 \mid R_1) \cdot P(B_3 \mid R_1 \wedge B_2) = \frac{6}{11} \cdot \frac{5}{10} \cdot \frac{4}{9}$$

$$+ \frac{5 \cdot 6 \cdot 4}{11 \cdot 10 \cdot 9}$$

$$+ \frac{5 \cdot 4 \cdot 6}{11 \cdot 10 \cdot 9} =$$



Example again

Example

An urn contains 6 red and 5 blue balls. We draw three balls at random, at once (that is, without replacement). What is the chance of drawing one red and two blue balls?

$$P(R_1 \cap B_2 \cap B_3) + P(B_1 \cap R_2 \cap B_3) + P(B_1 \cap B_2 \cap R_3)$$
$$= P(R_1) \cdot P(B_2 | R_1) \cdot P(B_3 | R_1 \cap B_2)$$
$$+ P(B_1) \cdot P(R_2 | B_1) \cdot P(B_3 | B_1 \cap R_2)$$
$$+ P(B_1) \cdot P(B_2 | B_1) \cdot P(R_3 | B_1 \cap B_2)$$
$$= \frac{6 \cdot 5 \cdot 4}{11 \cdot 10 \cdot 9} + \frac{5 \cdot 6 \cdot 4}{11 \cdot 10 \cdot 9} + \frac{5 \cdot 4 \cdot 6}{11 \cdot 10 \cdot 9} =$$

Example again

Example

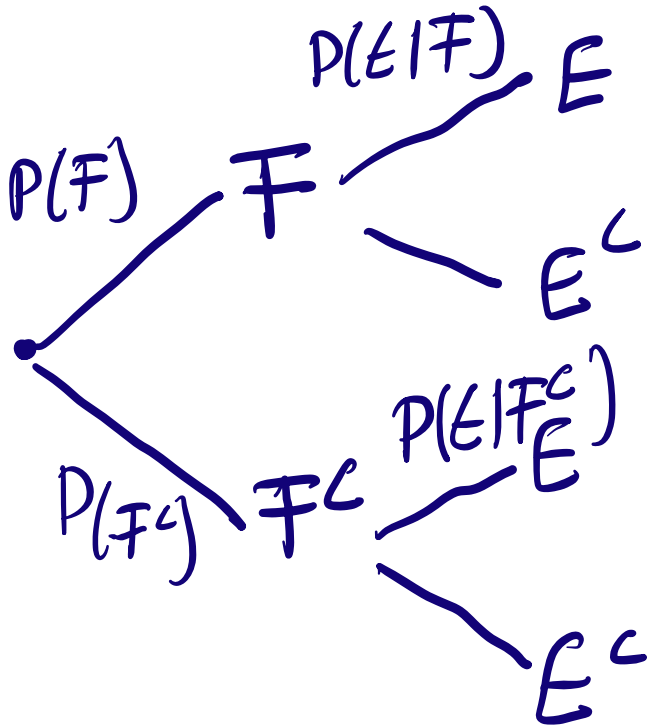
An urn contains 6 red and 5 blue balls. We draw three balls at random, at once (that is, without replacement). What is the chance of drawing one red and two blue balls?

$$P(R_1 \cap B_2 \cap B_3) + P(B_1 \cap R_2 \cap B_3) + P(B_1 \cap B_2 \cap R_3)$$
$$= P(R_1) \cdot P(B_2 | R_1) \cdot P(B_3 | R_1 \cap B_2)$$
$$+ P(B_1) \cdot P(R_2 | B_1) \cdot P(B_3 | B_1 \cap R_2)$$
$$+ P(B_1) \cdot P(B_2 | B_1) \cdot P(R_3 | B_1 \cap B_2)$$
$$= \frac{6}{11} \cdot \frac{5}{10} \cdot \frac{4}{9} + \frac{5}{11} \cdot \frac{6}{10} \cdot \frac{4}{9} + \frac{5}{11} \cdot \frac{4}{10} \cdot \frac{6}{9} = \frac{3 \cdot 8 \cdot 4 \cdot 6}{11 \cdot 10 \cdot 9} = \frac{4}{11}$$

Bayes' theorem

Bayes' Theorem

The aim is to say something about $\mathbf{P}(F \mid E)$, once we know $\mathbf{P}(E \mid F)$ (and other things...). This will be very useful, and serve as a fundamental tool in probability and statistics.

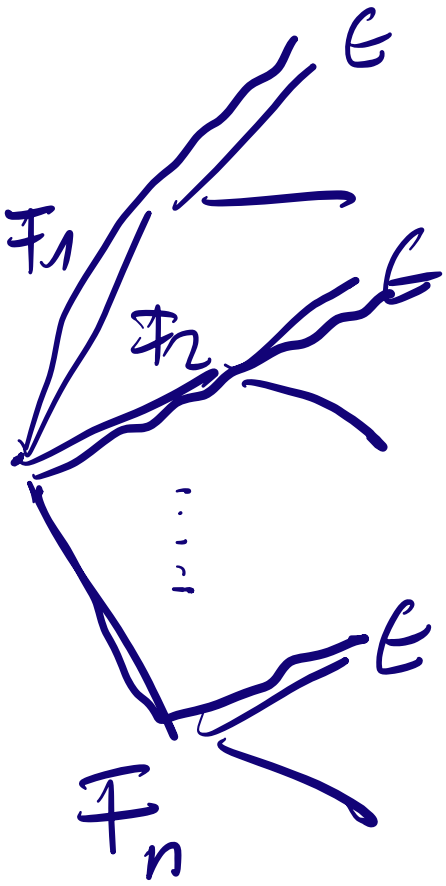


$$P(F \mid E) = \frac{P(F \cap E)}{P(E)}$$

The Law of Total Probability

Theorem (*Partition Theorem*)

$$\mathbf{P}(E) = \mathbf{P}(E \mid F) \cdot \mathbf{P}(F) + \mathbf{P}(E \mid F^c) \cdot \mathbf{P}(F^c)$$



The Law of Total Probability

Theorem (*Partition Theorem*)

$$\mathbf{P}(E) = \mathbf{P}(E \mid F) \cdot \mathbf{P}(F) + \mathbf{P}(E \mid F^c) \cdot \mathbf{P}(F^c)$$

Definition

Countably many events $F_1, F_2, \ldots$ form a *partition* of Ω if $F_i \cap F_j = \emptyset$ and $\bigcup_i F_i = \Omega$.

Theorem (*Partition Theorem*)

For any event E and any partition $F_1, F_2, \ldots$:

$$\mathbf{P}(E) = \sum_i \mathbf{P}(E \mid F_i) \cdot \mathbf{P}(F_i)$$

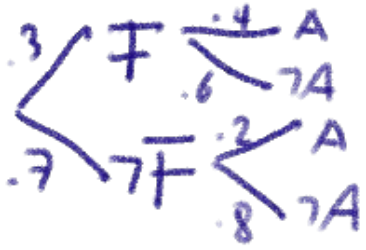
Example

Example

According to an insurance company:

- 30% of population are *accident-prone*:
  they will have an accident in any given year with 0.4 chance.
- 70% of population are *careful*:
  they have an accident in any given year with 0.2 chance.

How likely is a new customer to have an accident in 2023?



$$P(A) = P(A|F) \cdot P(F) + P(A|F^c) \cdot P(F^c)$$
$$= 0.4 \cdot 0.3 + 0.2 \cdot 0.7 = 0.26$$

Bayes' Theorem

$P(E \cap F)$

Theorem (Bayes' Theorem)

$$\mathbf{P}\{F \mid E\} = \frac{\mathbf{P}\{E \mid F\} \cdot \mathbf{P}\{F\}}{\mathbf{P}\{E \mid F\} \cdot \mathbf{P}\{F\} + \mathbf{P}\{E \mid F^c\} \cdot \mathbf{P}\{F^c\}}$$

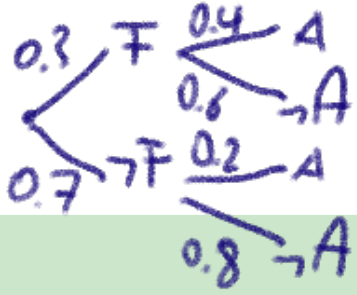
$P(E)$

If $\{F_i\}_i$ *partitions* Ω, *then:*

$$\mathbf{P}\{F_i \mid E\} = \frac{\mathbf{P}\{E \mid F_i\} \cdot \mathbf{P}\{F_i\}}{\sum_j \mathbf{P}\{E \mid F_j\} \cdot \mathbf{P}\{F_j\}}$$

Proof: Def of conditional prob + Law of total probability

Belief update



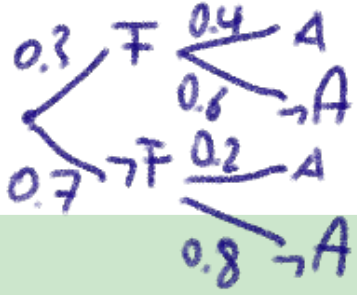
Example

Consider the insurance company again. Imagine it's now 2024. We learn that the new customer did have an accident in 2023. Now what is the chance that they are accident-prone?

$$P(F|A) = \frac{0.3 \cdot 0.4}{0.3 \cdot 0.4 + 0.7 \cdot 0.2} = 0.46$$

$$P(F) = 0,3$$

Belief update



Example

Consider the insurance company again. Imagine it's now 2024. We learn that the new customer did have an accident in 2023. Now what is the chance that they are accident-prone?

$$P(F|A) = \frac{P(A|F) \cdot P(F)}{P(A|F) \cdot P(F) + P(A|F^c) \cdot P(F^c)}$$

$$= \frac{0.4 \cdot 0.3}{0.4 \cdot 0.3 + 0.2 \cdot 0.7} = 46\%$$

instead of

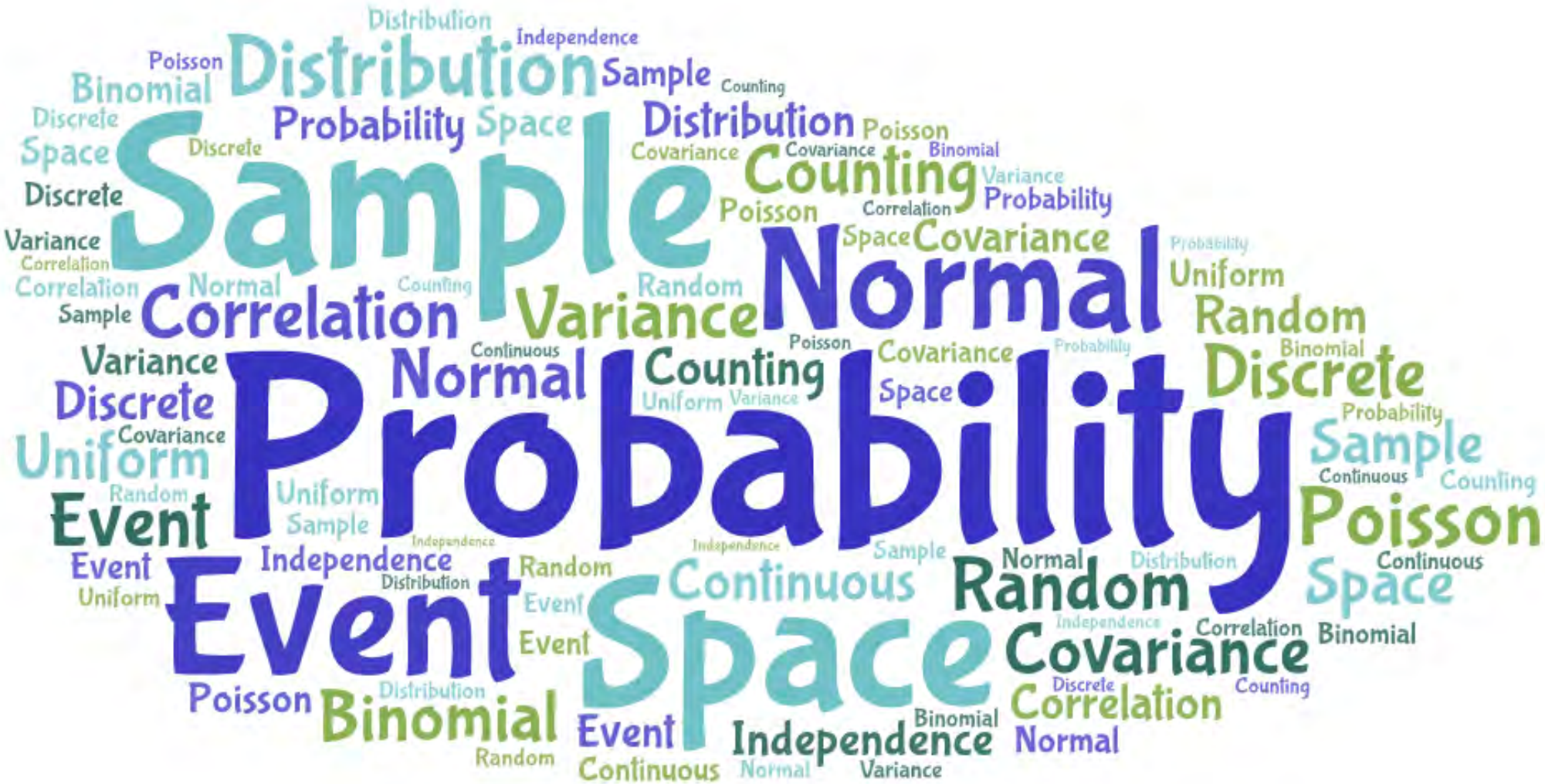
$P(F) = 0.3$ 30%

Summary

- Probability: multiple ways to compute
- Conditional probability: reduced sample space, multiplication rule
- Bayes' theorem: partition theorem, belief update

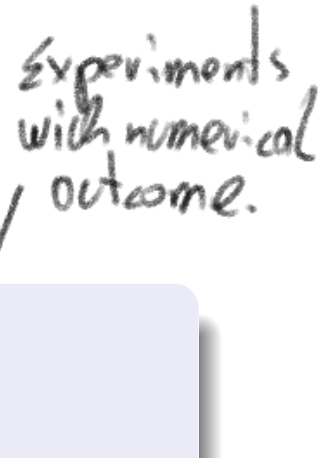
Discrete Mathematics and Probability

Week 8

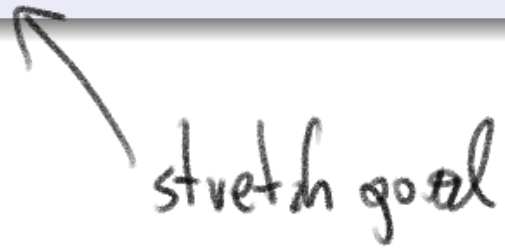


Mavkulf Chris Heunen Kohlweiss

Topics



- Independence: what information changes probability
- Random variables: when variables depend on chance
- Expectation: most likely outcomes of experiment
- Variance: how much the experiment can deviate



Independence

Independence

Sometimes partial information on an experiment
does not change the likelihood of an event.

Definition

Events E and F are *independent* if $\mathbf{P}(E \mid F) = \mathbf{P}(E)$.
Equivalently: $\mathbf{P}(E \cap F) = \mathbf{P}(E) \cdot \mathbf{P}(F)$.
Equivalently: $\mathbf{P}(F \mid E) = \mathbf{P}(F)$.

Proposition

If E and F are independent events,
then E and F^c are also independent.

Independence

Sometimes partial information on an experiment does not change the likelihood of an event.

Definition

Events E and F are *independent* if $\mathbf{P}(E \mid F) = \mathbf{P}(E)$.
Equivalently: $\mathbf{P}(E \cap F) = \mathbf{P}(E) \cdot \mathbf{P}(F)$.
Equivalently: $\mathbf{P}(F \mid E) = \mathbf{P}(F)$.

Proposition

If E and F are independent events, then E and F^c are also independent.

Always give you some example propositions.

independent events $\neq$ mutually exclusive events

independence usually either trivial or tricky

Examples

Ω = two dice

E = {sum is 6}

F = {first is 3}

$\frac{1}{36} = P(E \cap F) \neq P(E) \cdot P(F) = \frac{5}{36} \cdot \frac{1}{6}$

not independent

Examples

Ω = two dice

E = {sum is 6}

F = {first is 3}

$\frac{1}{36} = P(E \cap F) \neq P(E) \cdot P(F) = \frac{5}{36} \cdot \frac{1}{6}$

not independent

E' = {sum is 7}

F = {first is 3}

$\frac{1}{36} = P(E' \cap F) = P(E') \cdot P(F) = \frac{6}{36} \cdot \frac{1}{6}$

independent!

Examples

Ω = two dice

E = {sum is 6}

F = {first is 3}

$\frac{1}{36} = P(E \cap F) \neq P(E) \cdot P(F) = \frac{5}{36} \cdot \frac{1}{6}$

not independent

E' = {sum is 7}

F = {first is 3}

G = {second is 4}

$\frac{1}{36} = P(E' \cap F) = P(E') \cdot P(F) = \frac{6}{36} \cdot \frac{1}{6}$

independent!

pairwise independent

$\frac{1}{36} = P(E' \cap G) = P(E') \cdot P(G) = \frac{6}{36} \frac{1}{6}$

$\frac{1}{36} = P(F \cap G) = P(F) \cdot P(G) = \frac{1}{6} \cdot \frac{1}{6}$

Examples

Ω = two dice

E = {sum is 6}

F = {first is 3}

$\frac{1}{36} = P(E \cap F) \neq P(E) \cdot P(F) = \frac{5}{36} \cdot \frac{1}{6}$

not independent

E' = {sum is 7}

F = {first is 3}

G = {second is 4}

$\frac{1}{36} = P(E' \cap F) = P(E') \cdot P(F) = \frac{6}{36} \cdot \frac{1}{6}$

independent!

pairwise independent

$\frac{1}{36} = P(E' \cap G) = P(E') \cdot P(G) = \frac{6}{36} \frac{1}{6}$

$\frac{1}{36} = P(F \cap G) = P(F) \cdot P(G) = \frac{1}{6} \cdot \frac{1}{6}$

But: $\frac{1}{36} = P(E' \cap F \cap G) \stackrel{?}{=} P(E') \cdot P(F) \cdot P(G) = \frac{6}{36} \cdot \frac{1}{6} \cdot \frac{1}{6}$

Kind of makes sense: degrees of freedom

Independence

Definition

Three events E, F, G are *(mutually) independent* if:

$$\begin{aligned}
\mathbf{P}\{E \cap F\} &= \mathbf{P}\{E\} \cdot \mathbf{P}\{F\}, \\
\mathbf{P}\{E \cap G\} &= \mathbf{P}\{E\} \cdot \mathbf{P}\{G\}, \\
\mathbf{P}\{F \cap G\} &= \mathbf{P}\{F\} \cdot \mathbf{P}\{G\}, \\
\mathbf{P}\{E \cap F \cap G\} &= \mathbf{P}\{E\} \cdot \mathbf{P}\{F\} \cdot \mathbf{P}\{G\}.
\end{aligned}$$

For more events the definition is that any (finite) subset of them have this factorisation property.

have to go all the way to n

Examples

Examples

Fix parameter $0 < p < 1$

n independent experiments, each succeeds with probability p

. chance that every one succeeds?

Examples

Fix parameter $0 < p < 1$

n independent experiments, each succeeds with probability p

- chance that every one succeeds?

$$p^n \xrightarrow[n \to \infty]{} 0$$

- chance that at least one succeeds?

$$1 - (1-p)^n \xrightarrow[n \to \infty]{} 1$$

Murphy's Law

Examples

Fix parameter $0 < p < 1$

n independent experiments, each succeeds with probability p

- chance that every one succeeds?

$$p^n \xrightarrow[n \to \infty]{} 0$$

- chance that at least one succeeds?

$$1 - P(\text{each one fails}) = 1 - (1-p)^n \xrightarrow[n \to \infty]{} 1$$

Murphy's Law

- chance that exactly k succeed?

$$\binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

Random variables

Random variables

"Random variables $\approx$ random numbers". But *random* means that there must be some kind of experiment behind these numbers.

Definition

A *random variable* is a function from the sample space Ω to the real numbers $\mathbb{R}$.

flip 3 coins $\Omega = \{H,T\}^3$ X number of Hs

$X(TTT) = 0$

$X(HTT) = X(THT) = X(TTH) = 1$

$X(HHT) = X(HTH) = X(THH) = 2$

$X(HHH) = 3$

$P(X=1) = P(X^{-1}(1)) = P(\{HTT, THT, TTH\}) = \frac{3}{8}$

Random variables

"Random variables $\approx$ random numbers". But *random* means that there must be some kind of experiment behind these numbers.

Definition

A *random variable* is a function from the sample space Ω to the real numbers $\mathbb{R}$.

not between 0 or 1

flip 3 coins X = random variable that counts heads

$X : \Omega \to \mathbb{R}$

$X(TTT) = 0$

$X(HTT) = X(THT) = X(TTH) = 1$

$X(HHT) = X(HTH) = X(THH) = 2$

$X(HHH) = 3$

Random variables

"Random variables $\approx$ random numbers". But *random* means that there must be some kind of experiment behind these numbers.

Definition

A *random variable* is a function from the sample space Ω to the real numbers $\mathbb{R}$.

not between 0 or 1

flip 3 coins X = random variable that counts heads

$X : \Omega \to \mathbb{R}$

$X(TTT) = 0$

$X(HTT) = X(THT) = X(TTH) = 1$

$X(HHT) = X(HTH) = X(THH) = 2$

$X(HHH) = 3$

$P(X=1) = P(X^{-1}(1)) = P(\{HTT, THT, TTH\}) = \frac{3}{8}$

Discrete random variables

Definition

A random variable X that can take on countably many possible values is called *discrete*.

Discrete random variables

Definition

A random variable X that can take on countably many possible values is called *discrete*.

e.g. – nr. heads in 3 coin flips

– nr. of flips needed to first see a head

discrete things I can always count.
don't need to integrate

Probability mass function



Definition

The *probability mass function (pmf)*, or *distribution* of a discrete random variable X gives the probabilities of its possible values:

$$\mathfrak{p}_X(x_i) = \mathbf{P}(X = x_i),$$

Proposition

$$\mathfrak{p}(x_i) \geq 0 \qquad \textit{and} \qquad \sum_i \mathfrak{p}(x_i) = 1$$

Vice versa: any function $f: \mathbb{N} \to \mathbb{R}$ s.t. $f(x) \geq 0$, $\Sigma f(x) = 1$ is always a pmf

Probability mass function

Definition

The *probability mass function (pmf)*, or *distribution* of a discrete random variable X gives the probabilities of its possible values:

$$\mathfrak{p}_X(x_i) = \mathbf{P}(X = x_i),$$

Proposition

$$\mathfrak{p}(x_i) \geq 0 \qquad \textit{and} \qquad \sum_i \mathfrak{p}(x_i) = 1$$

Examples

Examples

Fix parameter $\lambda > 0$

Define $p(i) = c \cdot \frac{\lambda^i}{i!}$

which c makes this a pmf.?

$p(i) \geq 0 \iff c \geq 0$

Examples

Fix parameter $\lambda > 0$

Define $p(i) = c \cdot \frac{\lambda^i}{i!}$

which c makes this a pmf.?

$p(i) \geq 0 \iff c \geq 0$

Maclaurin infinite series expansion of e^λ

$$\sum_{i=0}^{\infty} p(i) = \sum_{i=0}^{\infty} c \cdot \frac{\lambda^i}{i!} = c \cdot \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = c \cdot e^\lambda \iff c = e^{-\lambda}$$

so e.g. $P(x=0) = p(0) = e^{-\lambda} \cdot \frac{\lambda^0}{0!} = e^{-\lambda}$

$$\begin{aligned} P(x>2) &= 1 - P(x \leq 2) \\ &= 1 - P(x=0) - P(x=1) - P(x=2) \\ &= 1 - e^{-\lambda} - e^{-\lambda}\lambda - e^{-\lambda}\frac{\lambda^2}{2} \end{aligned}$$

Cumulative distribution function

Definition

The *cumulative distribution function (cdf)* of a random variable X:

$$F : \mathbb{R} \to [0, 1], \qquad x \mapsto F(x) = \mathbf{P}(X \leq x).$$

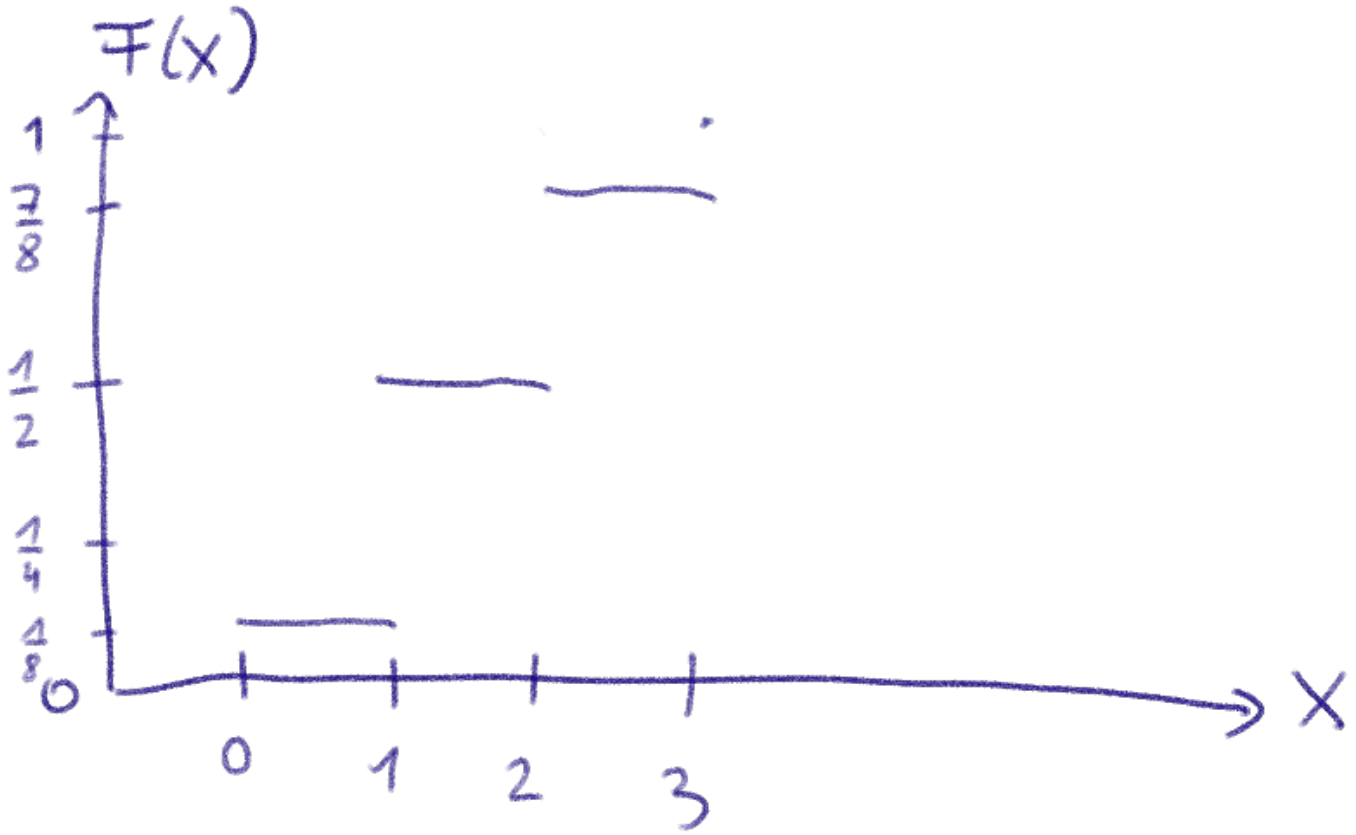
Cumulative distribution function



Definition

The *cumulative distribution function (cdf)* of a random variable X:

$$F : \mathbb{R} \to [0, 1], \qquad x \mapsto F(x) = \mathbf{P}(X \leq x).$$



Cumulative distribution function

Will be very important when working with continuous random variables.

Proposition

A cumulative distribution function F*:*

- *is non-decreasing: if* $x \leq y$ *then* $F(x) \leq F(y)$
- *has limit* $\lim_{x \to -\infty} F(x) = 0$ *on the left*
- *has limit* $\lim_{x \to \infty} F(x) = 1$ *on the right*

Expectation

Expectation

Once we have a random variable, we want to quantify its *typical* behaviour in some sense. Two of the most often used quantities for this are the *expectation* and the *variance*.

Definition

The *expectation* of a discrete random variable X is:

$$\mathbf{E}X = \sum_i x_i \cdot \mathfrak{p}(x_i)$$

provided the sum exists. Also called *mean*, or *expected value*.

weighted average / center of mass

Examples

Examples

X = nr on fair die after roll

$$EX = \sum_{i=1}^{6} x_i \cdot p(X=i) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} \dots + 6 \frac{1}{6} = \frac{7}{2}$$

expectation need not be a possible value

Properties of expectation

Proposition (expectation of a function of a random variable)

If X is a discrete random variable, and $g : \mathbb{R} \to \mathbb{R}$ a function, then:

$$\mathbf{E}g(X) = \sum_i g(x_i) \cdot \mathfrak{p}(x_i) \qquad \textit{(if it exists)}$$

Corollary (expectation is linear)

If X is a discrete random variable, and a, b fixed real numbers:

$$\mathbf{E}(aX + b) = a \cdot \mathbf{E}X + b.$$

Moments

Definition (moments)

Let $n \in \mathbb{N}$. The n^{th} *moment* of a random variable X is:

$$\mathbf{E}X^n$$

The n^{th} *absolute moment* of X is:

$$\mathbf{E}|X|^n$$

Variance

3. Variance

Definition (variance, standard deviation)

The *variance* and the *standard deviation* of a random variable are:

- $\mathbf{Var}\, X = \mathbf{E}(X - \mathbf{E}X)^2$.
- $\mathrm{SD}\, X = \sqrt{\mathbf{Var}\, X}$.

Properties of the variance

Proposition (equivalent form of the variance)

$\mathbf{Var}\, X = \mathbf{E}X^2 - (\mathbf{E}X)^2$ *for any random variable* X.

Corollary

$\mathbf{E}X^2 \geq (\mathbf{E}X)^2$ *for any random variable* X,
with equality only if X *is constant.*

Properties of the variance

Proposition (variance is not linear)

Let X be a random variable, a and b fixed real numbers. Then:

$$\mathbf{Var}(aX + b) = a^2 \cdot \mathbf{Var}\, X$$

Summary

- Independence: what information changes probability
- Random variables: when variables depend on chance
- Expectation: most likely outcomes of experiment
- Variance: how much the experiment can deviate