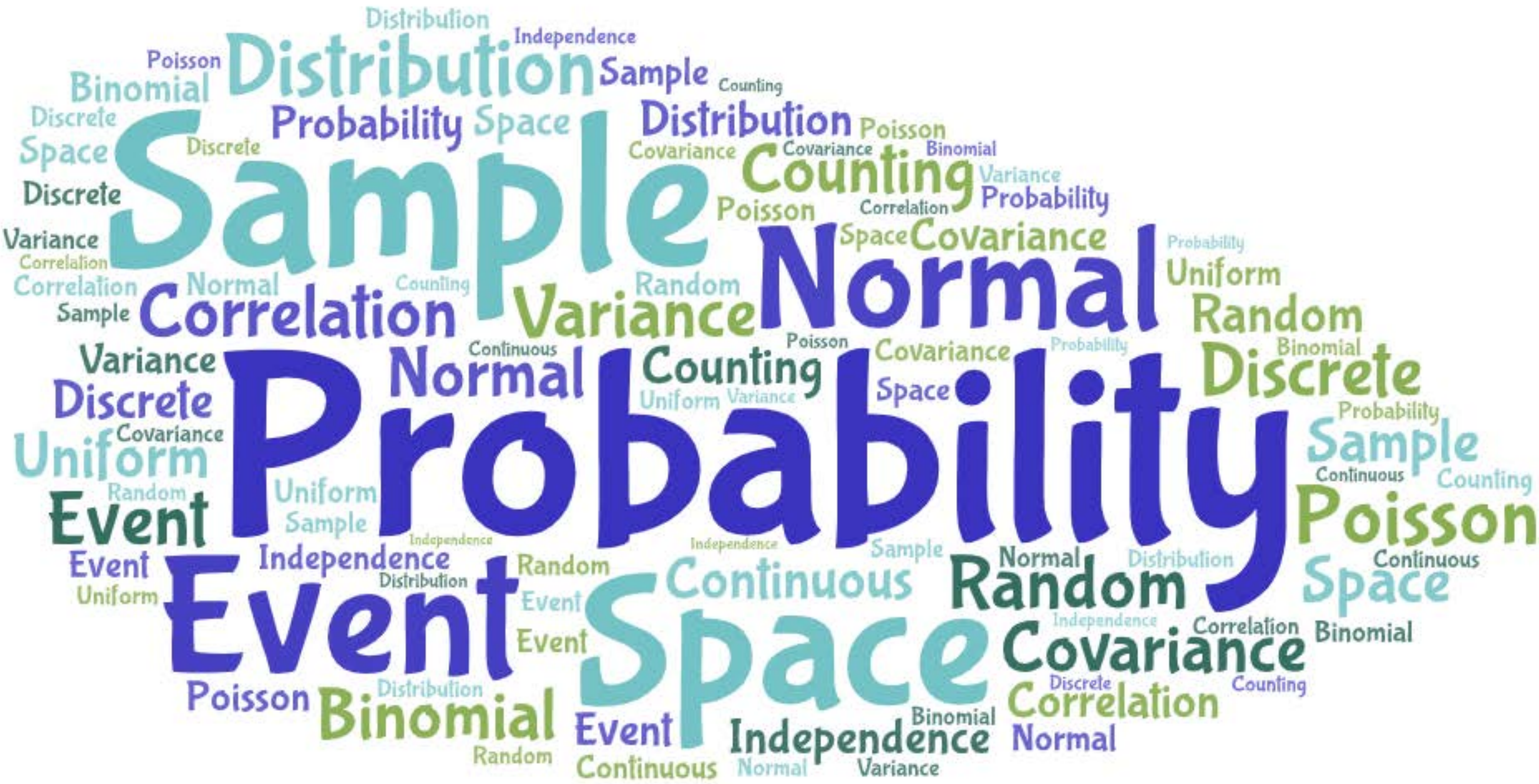
Discrete Mathematics and Probability

Week 7



MARKULF ~~Chris Heunen~~ KOHLWEISS

Events

A mathematical model for experiments:

- *Sample space*: the set Ω of all possible outcomes
- An *event* is a collection[1] of possible outcomes: $E \subseteq \Omega$
- *Union* $E \cup F$ and *intersection* $E \cap F$ of events make sense
- Complement of event $E^c = \Omega - E$

[1]Sometimes Ω is too large, and not all subsets are events. Ignore this now.

De Morgan's law

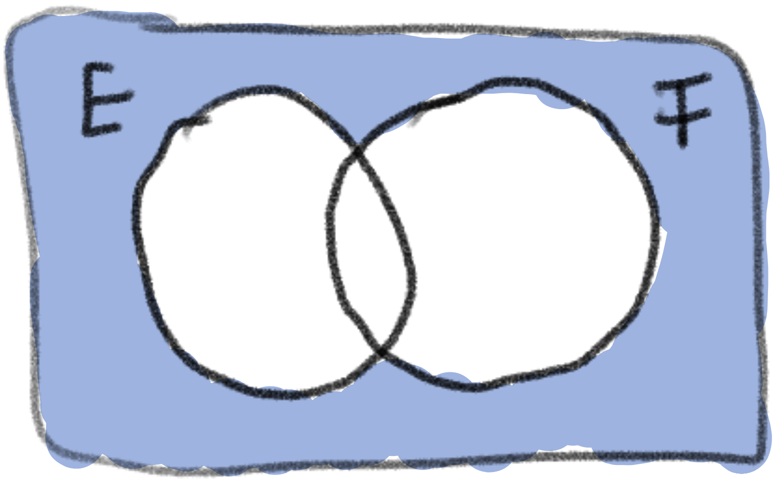
E = you have umbrella
F = I have umbrella

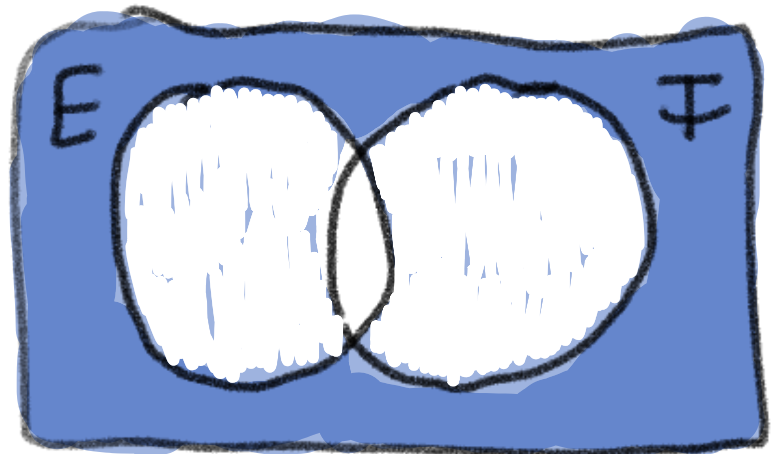
- De Morgan's law: $(E \cup F)^c = E^c \cap F^c$ = not true that one of us has umbrella = both have no umbrella

De Morgan's law

E = you have umbrella
F = I have umbrella

- De Morgan's law: $(E \cup F)^c = E^c \cap F^c$ = not true that one of us has umbrella = both have no umbrella



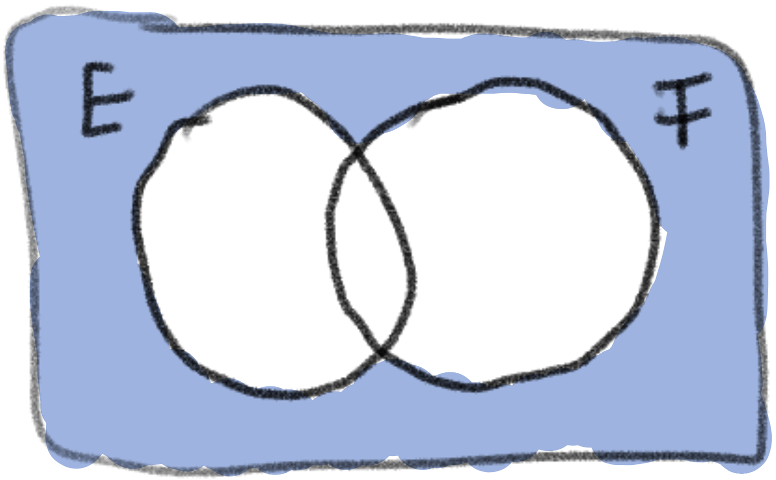


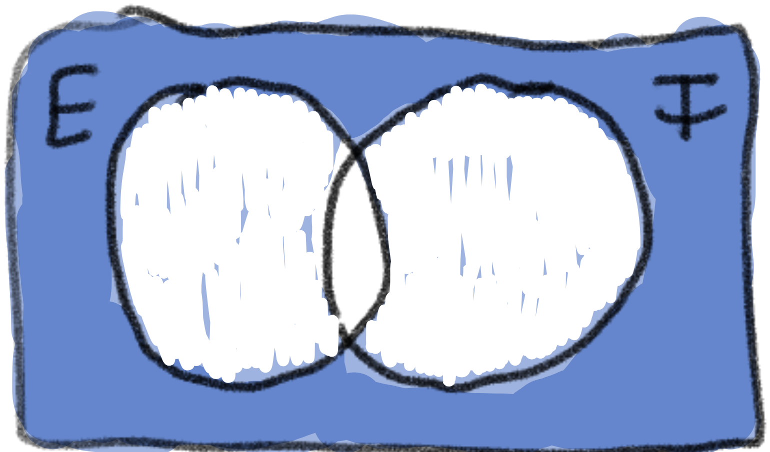
$(E \cap F)^c = E^c \cup F^c$ = not true both have umbrella = one does not have umbrella

De Morgan's law

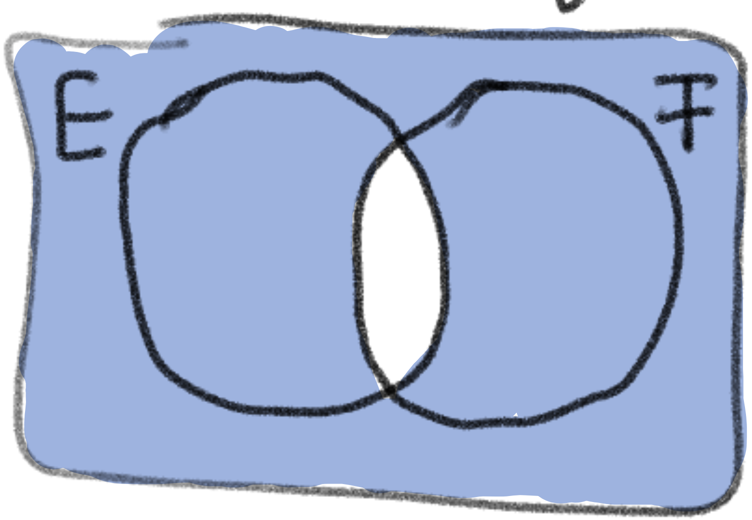
E = you have umbrella
F = I have umbrella

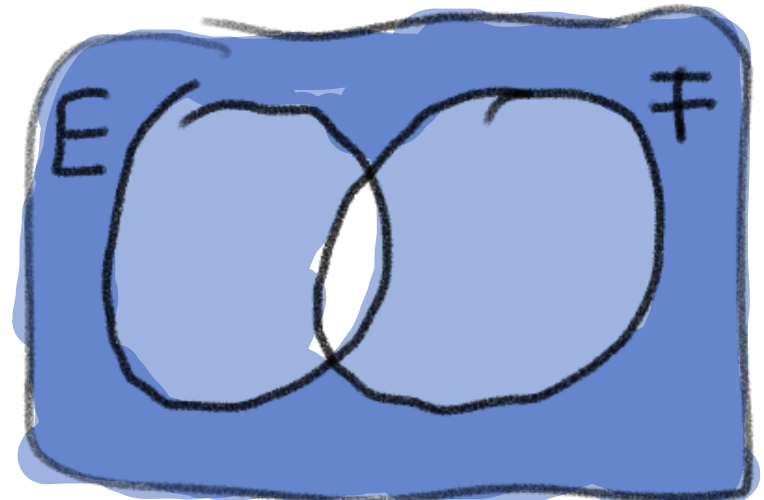
- De Morgan's law: $(E \cup F)^c = E^c \cap F^c$ = not true that one of us has umbrella = both have no umbrella





$(E \cap F)^c = E^c \cup F^c$ = not true both have umbrella = one does not have umbrella





Probability

So far been structural
now we will attach numbers to things
probabilities to events

- Definition by axioms
- How to compute probabilities
- Inclusion-exclusion principle
- Equally likely outcomes

Axioms of probability

Definition

The *probability* $\mathbf{P}$ on a sample space Ω assigns *numbers* to *events* of Ω in such a way that:

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

$$\mathbf{P}\Big(\bigcup_i E_i\Big) = \sum_i \mathbf{P}(E_i)$$

Axioms of probability

Definition

P is a function that satisfies 3 axioms

The *probability* $\mathbf{P}$ on a sample space Ω assigns *numbers* to *events* of Ω in such a way that:

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

$$\mathbf{P}\Big(\bigcup_i E_i\Big) = \sum_i \mathbf{P}(E_i)$$

finite $P(E_1 \cup \cdots \cup E_n) = P(E_1) + \cdots + P(E_n)$

countably infinite $P(E_1 \cup E_2 \cup E_3 \cup \cdots) = P(E_1) + P(E_2) + P(E_3) + \cdots$

How to compute probabilities

Proposition

$P(E) = 1 - P(E^c)$

For any event, $\mathbf{P}(E^c) = 1 - \mathbf{P}(E)$.

1) $E \cap E^c = \emptyset$ thus by axiom 3 $P(E) + P(E^c) = P(E \cup E^c) = P(\Omega)$
$= 1$

2) $P(\Omega) = 1$ by axiom 2

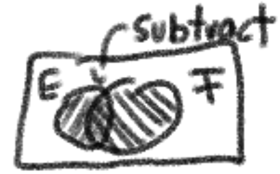
$P(\emptyset) = P(\Omega^c) = 1 - P(\Omega) = 1 - 1 = 0$

Corollary

We have $\mathbf{P}(\emptyset) = \mathbf{P}(\Omega^c) = 1 - \mathbf{P}(\Omega) = 1 - 1 = 0$.

For any event, $\mathbf{P}(E) = 1 - \underbrace{\mathbf{P}(E^c)}_{\geq 0} \leq 1$.

How to compute probabilities



Proposition (Inclusion-Exclusion)

For any two events, $\mathbf{P}(E \cup F) = \mathbf{P}(E) + \mathbf{P}(F) - \mathbf{P}(E \cap F)$.

Proposition (*Boole's inequality*)

For any events $E_1, E_2, \ldots, E_n$:

$$\mathbf{P}\left(\bigcup_{i=1}^{n} E_i\right) \leq \sum_{i=1}^{n} \mathbf{P}(E_i).$$

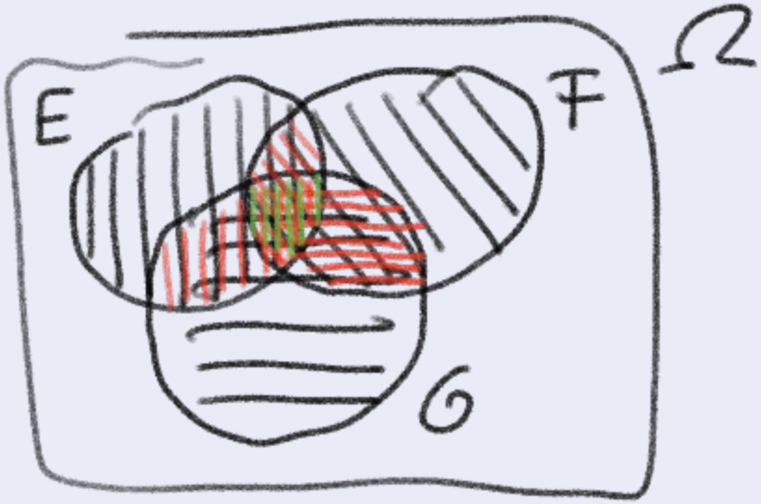
skipping proof by induction initially

Inclusion-exclusion (3 events)

Proposition

For any events:

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



Inclusion-exclusion (n- Events)

Proposition

For any events:

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

$$\begin{aligned}\mathbf{P}(E_1 \cup E_2 \cup \cdots \cup E_n) = &\sum_{1 \le i \le n} \mathbf{P}(E_i) \\ &- \sum_{1 \le i_1 < i_2 \le n} \mathbf{P}(E_{i_1} \cap E_{i_2}) \\ &+ \sum_{1 \le i_1 < i_2 < i_3 \le n} \mathbf{P}(E_{i_1} \cap E_{i_2} \cap E_{i_3}) \\ &- \cdots \\ &+ (-1)^{n+1} \mathbf{P}(E_1 \cap E_2 \cap \cdots \cap E_n).\end{aligned}$$

all the way up to n

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?

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?

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

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

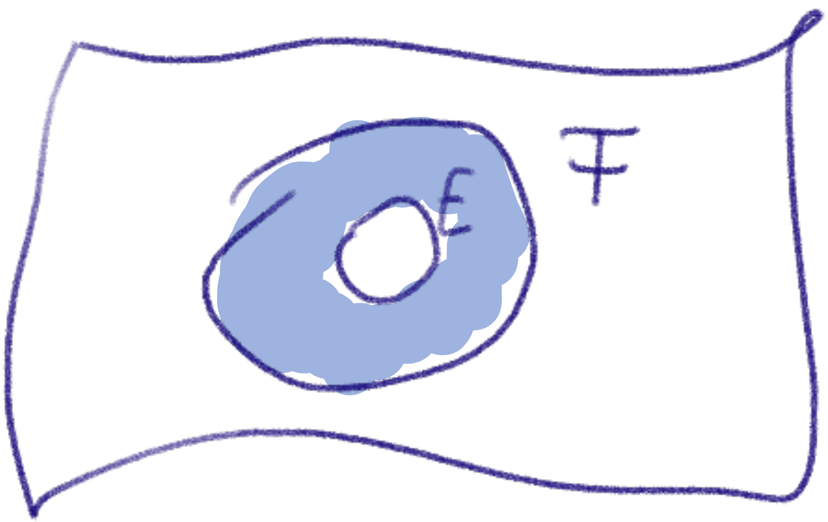
How to compute probabilities

Proposition

If $E \subseteq F$, *then* $\mathbf{P}(F - E) = \mathbf{P}(F) - \mathbf{P}(E)$.

Corollary

If $E \subseteq F$, *then* $\mathbf{P}(E) \leq \mathbf{P}(F)$.



Summary

- Counting: permutations, combinations, repetitions
- Events: sample space, union, intersection, complement
- ~~Experiments:~~ distributivity, De Morgan's law
- Probability: axioms, how to compute, equally likely outcomes

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

The last thing before we all go home
I want to talk about this

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 $\frac{1}{11}$."

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 case 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 case 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?

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"

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"

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

Axioms

Proposition

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

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

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

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*)

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

Multiplication rule

Proposition (*Multiplication rule*)

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

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

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?

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?

$$P(R_1 \cap B_2 \cap B_3) + P(B_1 \cap R_2 \cap B_3) + P(B_1 \cap B_2 \cap R_3)$$

$$= P(R_1) \cdot P(B_2 | R_1) \cdot P(B_3 | R_1 \cap B_2)$$

$$+ P(B_1) \cdot P(R_2 | B_1) \cdot P(B_3 | B_1 \cap R_2)$$

$$+ P(B_1) \cdot P(B_2 | B_1) \cdot P(R_3 | B_1 \cap B_2)$$

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

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.

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 $F_1, F_2, \ldots$:

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

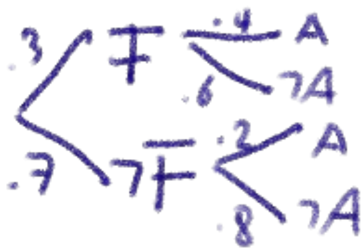
Example

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

Theorem (Bayes' Theorem)

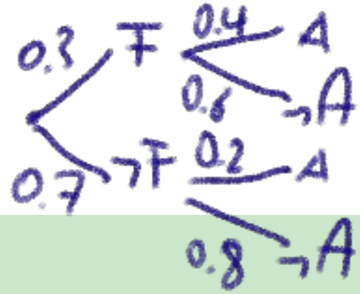
$$\mathbf{P}\{F \mid E\} = \frac{\mathbf{P}\{E \mid F\} \cdot \mathbf{P}\{F\}}{\mathbf{P}\{E \mid F\} \cdot \mathbf{P}\{F\} + \mathbf{P}\{E \mid F^c\} \cdot \mathbf{P}\{F^c\}}$$

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?

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

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

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