Probability

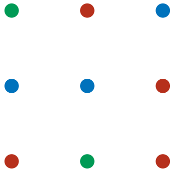
Adam Klepáč

January 5, 2024

Probabilistic Intuition

WHAT IS CHANCE?

Imagine you have 9 balls of different colours.



- If you pick a ball **at random**, what colour is it most likely to be?
- How many times more likely is picking a red ball than picking a green ball?

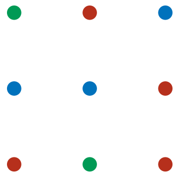
Quantifying Probability

Probability

A **probability** is a number between 0 and 1 measuring how **likely** is something to happen.

Quantifying Probability – Example

In our example of 9 balls



what is the probability of picking a ball of a specific colour?

- For red, it's $4/9$.
- For blue, it's $3/9$.
- For green, it's $2/9$.

The probabilities above **sum up to 1** because I am certain to pick *some* ball.

Quantifying Probability – Example

We'll all the outcome of a random choice, a **random variable** and typically write it as X.

A random variable always lies in the set of all possible outcomes.

In this case, the variable X must lie in the set of possible colours, $\{\text{red}, \text{blue}, \text{green}\}$.

We'll write the probability that X is equal to one of the elements in the set as $P(X = \text{colour})$.

So, for the 9-ball example from before, we would have

$$P(X = \text{red}) = \frac{4}{9}, \quad P(X = \text{blue}) = \frac{3}{9}, \quad P(X = \text{green}) = \frac{2}{9}.$$

Calculating Probability

In the case the set of outcomes is **finite**, the probability of X being one of the possible outcomes is always

$$P(X \in S) = \frac{|S|}{|O|},$$

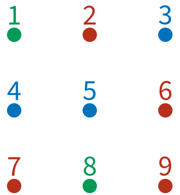
where S is a certain subset of O – all the possible outcomes.

Calculating Probability – Example

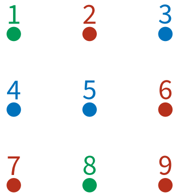
We'll describe our 9-ball example more formally.

We'll assign the balls number from 1 to 9. The set of all possible outcomes of picking a random ball is then

$$O = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}.$$



Calculating Probability – Example



We'll form three subsets of O:

$$R = \{2, 6, 7, 9\},$$

$$B = \{3, 4, 5\},$$

$$G = \{1, 8\}.$$

We can use the formula from before to calculate the probability that X will be a green ball:

$$P(X \in G) = \frac{|G|}{|O|} = \frac{2}{9}.$$

Probability Equations

Sums of Probabilities

What if I asked about the probability that the ball I pick is red or blue?

We can literally use the same formula as before. Now, the set of outcomes we're interested in is $R \cup B$ and so

$$P(X \in R \cup B) = \frac{|R \cup B|}{|O|} = \frac{|R| + |B|}{|O|} = \frac{4+3}{9} = \frac{7}{9}.$$

However, this example cannot be easily generalized. We'll see why.

Sums of Probabilities – Counterexample

Suppose we're instead choosing from a set of numbers between 1 and 20.

We want to calculate the probability that a randomly picked number is **even or divisible by** 5.

So, we have

$$O = \{1, 2, \ldots, 20\},$$

$$E = \{2, 4, 6, \ldots, 20\},$$

$$F = \{5, 10, 15, 20\}.$$

and we want to figure out the probability $P(X \in E \cup F)$.

Sums of Probabilities – Counterexample

Let's try to use the same formula as before:

$$P(X \in E \cup F) = \frac{|E \cup F|}{|O|} \overset{??}{=} \frac{|E| + |F|}{|O|} = \frac{10 + 4}{20} = \frac{14}{20}.$$

This doesn't quite add up.

If we count such numbers by hand, we get the set

$$\{2, 4, 5, 6, 8, 10, 12, 14, 15, 16, 18, 20\}.$$

There's **only 12 of them.**

The problem is that **we counted the numbers 10 and 20 twice!**

So, to get the size of $E \cup F$, we cannot just add the size of E to the size of F but we also have to subtract the elements that appear twice – the size of $E \cap F$.

Sums of Probabilities – Formula

The previous example applies in general. If A, B are two subsets of the set of outcomes, O, then

$$P(X \in A \cup B) = \frac{|A \cup B|}{|O|} = \frac{|A| + |B| - |A \cap B|}{|O|}.$$

Sums of Probabilities – Formula

We have a formula for two sets but how about three sets? Four sets? Million sets?

We need a **general formula** to calculate the size

$$|A_1 \cup A_2 \cup \ldots \cup A_n|$$

where $A_1, A_2, \ldots, A_n$ are any sets.

Such a formula is widely known as the **principle of inclusion and exclusion**.

Principle of Inclusion and Exclusion – Example

Let's consider the following setup: There are three language groups – English, French and German.

- 40 people speak English, 23 speak German and 11 speak French.
- 10 people speak both English and German, 5 speak both English and French and only 3 speak both German and French.
- Finally, just one person speaks all three languages.

How many people speak at least one language?

Principle of Inclusion and Exclusion – Example

Let's tackle this formally.

Label the three language groups E, F and G. The setup from the previous slide can be summarized as

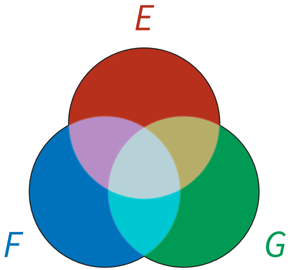
$\lvert E\rvert$	$\lvert F\rvert$	$\lvert G\rvert$	$\lvert E \cap F\rvert$	$\lvert E \cap G\rvert$	$\lvert F \cap G\rvert$	$\lvert E \cap F \cap G\rvert$
40	11	23	5	10	3	1

We're trying to calculate $\lvert E \cup F \cup G\rvert$.

Principle of Inclusion and Exclusion – Example

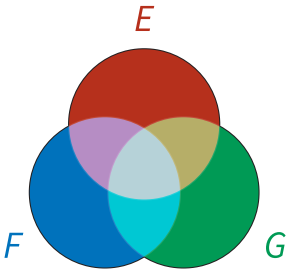
Let's picture the problem first.

When working with sets, Venn diagrams are often a great choice.



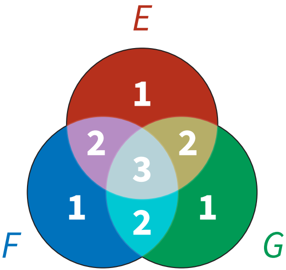
There are 7 regions in total (differentiated by colour) in this picture, corresponding to the 7 sets in the previous slide.

Principle of Inclusion and Exclusion – Example



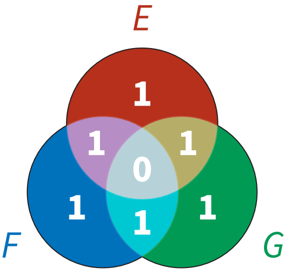
What we need to count is the total number of elements inside this entire shape. Let's start by counting the number of elements in each of the regions separately and assign numbers to regions corresponding to **how many times we've counted all the elements in that region.**

Principle of Inclusion and Exclusion – Example



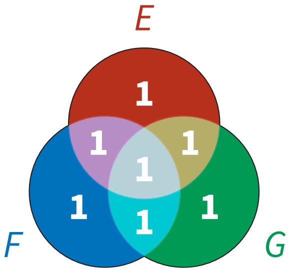
$$|E \cup F \cup G| = |E| + |F| + |G| \ldots$$

Principle of Inclusion and Exclusion – Example



$$|E \cup F \cup G| = |E| + |F| + |G| - |E \cap F| - |E \cap G| - |F \cap G|$$

Principle of Inclusion and Exclusion – Example



$$|E \cup F \cup G| = |E| + |F| + |G| - |E \cap F| - |E \cap G| - |F \cap G| + |E \cap F \cap G|.$$

Principle of Inclusion and Exclusion – Example

$$|E \cup F \cup G| = |E| + |F| + |G| - |E \cap F| - |E \cap G| - |F \cap G| + |E \cap F \cap G|.$$

Apply this formula to our example with language groups gives

$$|E \cup F \cup G| = 40 + 11 + 23 - 5 - 10 - 3 + 1 = 57.$$

So, 57 people speak at least one language.

Principle of Inclusion and Exclusion – Formula

The previous example can be generalized to any number of sets.
The basic idea is

1. Add the sizes of all the sets.
2. Subtract the size of all two-set intersections.
3. Add the sizes of all three-set intersections.
4. Subtract the sizes of all four-set intersections.
5. ...

Principle of Inclusion and Exclusion – Formula

If $A_1, A_2, \ldots, A_n$ are sets with $n \in \mathbb{N}$, then

Principle of Inclusion and Exclusion

$$
\begin{aligned}
|A_1 \cup A_2 \cup \ldots A_n| = {} & |A_1| + |A_2| + |A_3| + \ldots + |A_n| \\
& - |A_1 \cap A_2| - \ldots - |A_1 \cap A_n| - |A_2 \cap A_3| - \ldots - |A_{n-1} \cap A_n| \\
& + |A_1 \cap A_2 \cap A_3| + \ldots + |A_1 \cap A_2 \cap A_n| + \ldots |A_{n-2} \cap A_{n-1} \cap A_n| \\
& \quad \vdots \\
& + (-1)^n |A_1 \cap A_2 \cap \ldots \cap A_n|.
\end{aligned}
$$

The $(-1)^n$ only means that if n is odd, then I subtract the last term, and I add it if n is even.

Principle of Inclusion and Exclusion – Problems

Probabilistic problems requiring the **principle of exclusion and inclusion** are those with multiple desirable outcomes.
Let's start with something familiar:

Out of the numbers 1 to 100, what is the probability that a randomly picked number is a multiple of 2, 3 or 7?

Let's define the sets

$$E = \{\text{multiples of } 2\}, \quad T = \{\text{multiples of } 3\}, \quad S = \{\text{multiples of } 7\}$$

and

$$O = \{1, 2, \ldots, 100\}.$$

Principle of Inclusion and Exclusion – Problems

We're figuring out the probability

$$P(X \in E \cup T \cup S) = \frac{|E \cup T \cup S|}{|O|}.$$

Using the **inclusion-exclusion principle**, we count

$$\begin{aligned}
|E \cup T \cup S| &= |E| + |T| + |S| - \underbrace{|E \cap T|}_{\text{multiples of 6}} - \underbrace{|E \cap S|}_{\text{multiples of 14}} - \underbrace{|T \cap S|}_{\text{multiples of 21}} + \underbrace{|E \cap T \cap S|}_{\text{multiples of 42}} \\
&= 50 + 33 + 14 - 16 - 7 - 4 + 2 = 72.
\end{aligned}$$

So,

$$P(X \in E \cup T \cup S) = \frac{72}{100}.$$

Principle of Inclusion and Exclusion – Problems

Given two circles and a triangle in the plane, what's the maximum number of points that can belong to at least two of these shapes?
Let's label the circles C_1, C_2 and the triangle T. We're interested in points that lie in **at least one** of the sets $C_1 \cap C_2$, $C_1 \cap T$ and $C_2 \cap T$.
In other words, we want to determine the maximum size of

$$|(C_1 \cap C_2) \cup (C_1 \cap T) \cup (C_2 \cap T)|.$$

The maximum number of points

- two circles can share is 2. So, let's set $|C_1 \cap C_2| = 2$.
- circle and a triangle can share is 6. So $|C_1 \cap T| = |C_2 \cap T| = 6$.
- all three objects share is zero if the number of intersections is maximized. So $|C_1 \cap C_2 \cap T| = 0$.

Principle of Inclusion and Exclusion – Problems

Let's apply the inclusion-exclusion principle. We get

$$\begin{aligned}
&|(C_1 \cap C_2) \cup (C_1 \cap T) \cup (C_2 \cup T)| \\
&= |C_1 \cap C_2| + |C_1 \cap T| + |C_2 \cap T| \\
&- |(C_1 \cap C_2) \cap (C_1 \cap T)| - |(C_1 \cap C_2) \cap (C_2 \cap T)| - |(C_1 \cap T) \cap (C_2 \cap T)| \\
&+ |(C_1 \cap C_2) \cap (C_1 \cap T) \cap (C_2 \cap T)|.
\end{aligned}$$

Principle of Inclusion and Exclusion – Problems

This is less scary than it looks. Actually, most of the intersections there are one and the same. Really,

$$
\begin{aligned}
(C_1 \cap C_2) \cap (C_1 \cap T) &= C_1 \cap C_2 \cap T \\
(C_1 \cap C_2) \cap (C_2 \cap T) &= C_1 \cap C_2 \cap T \\
(C_1 \cap T) \cap (C_2 \cap T) &= C_1 \cap C_2 \cap T \\
(C_1 \cap C_2) \cap (C_1 \cap T) \cap (C_2 \cap T) &= C_1 \cap C_2 \cap T.
\end{aligned}
$$

So, the previous expression just ends up being

$$|C_1 \cap C_2| + |C_1 \cap T| + |C_2 \cap T| - 2 \cdot |C_1 \cap C \cap T| = 2 + 6 + 6 - 2 \cdot 0 = 14.$$

Events

What Is An Event?

Formally, an **event** is just an element which **has some probability**.

However, we typically think of events as **things that have some chance of happening**.

For example

- the fact that a random variable X lies in some set is an event.
- the fact that a randomly chosen ball has a specific colour is an event.
- the fact that the universe ends today at midnight is an event.

Operations On Events

Ultimately, an event is **logical sentence**, meaning it's a sentence which is either **true** or **false**.

Logical sentences can be negated (written as $\neg$) and joined together using logical conjunctions

- **and** (written as $\wedge$),
- **or** (written as $\vee$).

We want to understand how to calculate $P(\neg A), P(A \wedge B), P(A \vee B)$ for two events A, B whose probabilities we know.

Negating Events

Suppose we have an event $A =$ 'It's going to rain in 5 minutes.' with probability 0.2.
What's the probability of $\neg A$?
Quite naturally, it's 0.8 because it either is going to rain or it's not. The total of the probabilities of those two events has to be 1.

Negation Formula

If A is an event with probability $P(A) = p$, then

$$P(\neg A) = 1 - p.$$

Independent and Incompatible Events

Two events are called **independent** if the result of one event doesn't at all influence the result of the other.

For example, two tosses of a coin are independent.

Two events are called **incompatible** if they cannot *both* happen.

For example, the events 'I'm going to spend vacation in Florida.' and 'I'm going to spend vacation in Spain.' are incompatible.

An event is always incompatible with its own negation.

Conjunction of Events

Conjunction Formula

If A, B are two **independent** events, then

$$P(A \wedge B) = P(A) \cdot P(B).$$

If the two events are **dependent**, then calculating the probability of their conjunction is much more difficult. We'll need *conditional probability* for that.

Disjunction of Events

The disjunction formula for events is really just the inclusion-exclusion principle.

Disjunction Formula

If A, B are any events, then

$$P(A \vee B) = P(A) + P(B) - P(A \wedge B).$$

If A, B are **incompatible**, then $P(A \wedge B) = 0$ and the formula above becomes $P(A \vee B) = P(A) + P(B)$.

Conditional Probability

Conditional is the probability of an event happening **given another event has already happened**.

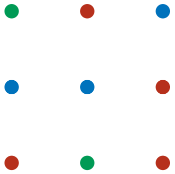
Conditional Probability

If A, B are events, then

$$P(A \mid B)$$

is the probability that A happens supposing B has already happened.

Conditional Probability – Example

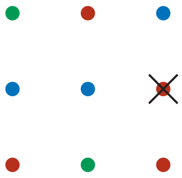


In our balls example, suppose

- A is the event that the **second** randomly chosen ball is red.
- B is the event that the **first** randomly chosen ball is red.

What is the probability $P(A \mid B)$?

Conditional Probability – Example



If B has happened, then there are only 3 red balls left in the set of 8 balls.
Therefore $P(A \mid B) = 3/8$.

Conditional Probability & Conjunction

We can use conditional probability to compute $P(A \wedge B)$ for **any** events A and B, not necessarily independent.

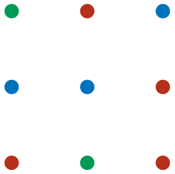
If, for example, A is dependent on B, then A can **only happen if B happened as well**. In other words, the probability that $A \wedge B$ happens is the probability that B happens times the probability that A happens **supposing B happened**.

Event Conjunction Formula

If A, B are **any** events, then

$$P(A \wedge B) = P(B) \cdot P(A \mid B) = P(A) \cdot P(B \mid A).$$

Conditional Probability & Conjunction



The event $A \wedge B$ in the ball example means that the first two randomly chosen balls are red. We know that $P(B) = 4/9$ and $P(A \mid B) = 3/8$. Therefore,

$$P(A \wedge B) = P(B) \cdot P(A \mid B) = \frac{4}{9} \cdot \frac{3}{8} = \frac{1}{6}.$$

Conditional Probability – Examples

Suppose the probability that a woman will live to at least 70 years is 0.7 and that she will live to at least 80 years is 0.55. What is the probability that she will live to 80 supposing she has already turned 70?

Let's set the problem up formally:

- S is the event that she will live to at least 70 and E is the event that she will live to at least 80.
- We know that $P(S) = 0.7$ and $P(E) = 0.55$.
- We want to know $P(E \mid S)$.

Conditional Probability – Examples

Suppose the probability that a woman will live to at least 70 years is 0.7 and that she will live to at least 80 years is 0.55. What is the probability that she will live to 80 supposing she has already turned 70?

Using the formula $P(E \wedge S) = P(S) \cdot P(E \mid S)$, we can calculate

$$P(E \mid S) = \frac{P(E \wedge S)}{P(S)}.$$

Quite clearly, $P(E \wedge S) = P(E)$, so the above becomes $P(E)/P(S)$. This means that

$$P(E \mid S) = \frac{P(E)}{P(S)} = \frac{0.55}{0.7} = 0.786.$$

Conditional Probability – Examples

From a deck of 32 cards (8 ranks and 4 suits) two cards are drawn. What's the probability that the first is of diamonds and the second is of a different suit?

Let's formalize, again. Denote by D the event that the first card is of diamonds and by S the event that the second card is of a different suit. We want to know $P(D \wedge S)$.

We'll use the formula $P(D \wedge S) = P(D) \cdot P(S \mid D)$.

It's easy to see that $P(D) = 1/4$. If D has already happened, there are only 31 cards left in the deck, 24 of them being of a different suit than diamonds. This means that $P(S \mid D) = 24/31$.

Multiplying these two values, we get our result:

$$P(D \wedge S) = P(D) \cdot P(S \mid D) = \frac{1}{4} \cdot \frac{24}{31} = \frac{6}{31}.$$

Bayes' Theorem

Consider the following problem: *In a clinic, 10 % of patients are prescribed pain killers. Overall, 5 % of the patients are addicted to narcotics (whether pain killers or illegal substances). Out of all the people with prescribed pain killers, 8 % are addicts.* ***If a patient is an addict, what's the probability that he will be prescribed pain killers?***

Let's formalize this:

- Let's denote by A the event that a patient is prescribed pain killers.
- By E, we denote the event that a patient is an addict.

We know that $P(A) = 0.1$ and $P(B) = 0.05$. We also know that **if** a patient is prescribed pain killers, then he is addicted with probability 0.08. In symbols, $P(B \mid A) = 0.08$. However, we're asking for the probability of being prescribed pain killers to an addicted patient. That is, we want to know $P(A \mid B)$.

Bayes' Theorem

Fortunately, there's a way to compute $P(A \mid B)$ having determined $P(B \mid A)$.
Remember the formula for calculating $P(A \wedge B)$. We have

$$P(A \wedge B) = P(B) \cdot P(A \mid B) = P(A) \cdot P(B \mid A).$$

The second equality is important. If we divide by $P(B)$, we get that

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

This (either of the two equalities) is called the **Bayes' Theorem**.

Bayes' Theorem

Bayes' Theorem

If A, B are events, then

$$P(B) \cdot P(A \mid B) = P(A) \cdot P(B \mid A).$$

Bayes' Theorem

Using Bayes' Theorem, we can solve the problem. Recall that

$$P(A) = 0.1, \quad P(B) = 0.05 \quad \text{and} \quad P(B \mid A) = 0.08.$$

This means that

$$P(A \mid B) = \frac{P(A) \cdot P(B \mid A)}{P(B)} = \frac{0.1 * 0.08}{0.05} = 0.16.$$

Conditional Probability Problems

Problem #1

A nuclear power plant hosts three reactors. One is quite old and the probability of failure is about 0.001. The other two are newer and their share an estimated probability of failure of 0.0002. *What's the probability that a reactor fails?*

Problem #2

A box contains three coins: two regular coins and one fake two-headed coin.

- You pick a coin at random and toss it. What is the probability that it lands heads up?
- You pick a coin at random and toss it, and get heads. What is the probability that it is the two-headed coin?

Problem # 3

A spam filter is designed by looking at commonly occurring phrases in spam. Suppose that 80 % of email is spam. In 10 % of the spam emails, the phrase "free money" is used, whereas this phrase is only used in 1 % of non-spam emails. A new email has just arrived, which does mention "free money". What is the probability that it is spam?