McCOMBS SCHOOL OF BUSINESS
Salem Center for Policy

Probability

David Puelz

September 14, 2021

Outline

The basics and conditional probability

Independence

Paradoxes, mixtures, and the rule of total probability

Random variables, distributions, and simulation

What is probability?

- A measure of uncertainty
- Answering the question: "How likely is a given event?"
- As with any mathematical concept, there are a set of axioms setting the "ground rules"
- Separately, there are different ways to interpret probability ...

  (i) **frequentist**: limit of relative frequency after repeating an experiment an infinite number of times (coin flip!)

  (ii) **Bayesian**: subjective belief about the likelihood of an event occurence

Probability basics

If A denotes some event, then $P(A)$ is the probability that this event occurs:

- $P(\text{coin lands heads}) = 0.5$
- $P(\text{rainy day in Ireland}) = 0.85$
- $P(\text{cold day in Hell}) = 0.0000001$

And so on...

Probability basics

If A denotes some event, then $P(A)$ is the probability that this event occurs:

- $P(\text{coin lands heads}) = 0.5$
- $P(\text{rainy day in Ireland}) = 0.85$
- $P(\text{cold day in Hell}) = 0.0000001$

And so on...

Probability basics

Some probabilities are estimated from direct experience over the long run:

- $P(\text{newborn baby is a boy}) = \frac{106}{206}$
- $P(\text{death due to car accident}) = \frac{11}{100{,}000}$
- $P(\text{death due to any cause}) = 1$

Probability basics

Some probabilities are estimated from direct experience over the long run:

- $P(\text{newborn baby is a boy}) = \frac{106}{206}$
- $P(\text{death due to car accident}) = \frac{11}{100{,}000}$
- $P(\text{death due to any cause}) = 1$

Others are synthesized from our best judgments about unique events:

- $P(\text{Apple stock goes up after next earnings call}) = 0.54$
- $P(\text{Djokovic wins next US Open}) = 0.4$ (6 to 4 odds)
- etc.

Probability basics: conditioning

A conditional probability is the chance that one thing happens, given that some other thing has already happened.

A great example is a weather forecast: if you look outside this morning and see gathering clouds, you might assume that rain is likely and carry an umbrella.

We express this judgment as a conditional probability: e.g. "the conditional probability of rain this afternoon, given clouds this morning, is 60%."

Probability basics: conditioning

In statistics, we write this a bit more compactly:

- $P(\text{rain this afternoon} \mid \text{clouds this morning}) = 0.6$
- That vertical bar means "given" or "conditional upon."
- The thing on the left of the bar is the event we're interested in.
- The thing on the right of the bar is our knowledge, also called the "conditioning event" or "conditioning variable": what we believe or assume to be true.

$P(A \mid B)$: "the probability of A, given that B occurs."

Probability basics: conditioning

Conditional probabilities are how we express judgments in a way that reflects our partial knowledge.

- You just gave *Manifest* a high rating. What's the conditional probability that you will like *The Bourne Identity* or *Saving Private Ryan*?

Probability basics: conditioning

Conditional probabilities are how we express judgments in a way that reflects our partial knowledge.

- You just gave *Manifest* a high rating. What's the conditional probability that you will like *The Bourne Identity* or *Saving Private Ryan*?
- You just bought organic dog food on Amazon. What's the conditional probability that you will also buy a GPS-enabled dog collar?

Probability basics: conditioning

Conditional probabilities are how we express judgments in a way that reflects our partial knowledge.

- You just gave *Manifest* a high rating. What's the conditional probability that you will like *The Bourne Identity* or *Saving Private Ryan*?
- You just bought organic dog food on Amazon. What's the conditional probability that you will also buy a GPS-enabled dog collar?
- You follow Lionel Messi (@leomessi) on Instagram. What's the conditional probability that you will respond to a suggestion to follow Cristiano Ronaldo (@cristiano)?

Probability basics: conditioning

Conditional probabilities are how we express judgments in a way that reflects our partial knowledge.

- You just gave *Manifest* a high rating. What's the conditional probability that you will like *The Bourne Identity* or *Saving Private Ryan*?
- You just bought organic dog food on Amazon. What's the conditional probability that you will also buy a GPS-enabled dog collar?
- You follow Lionel Messi (@leomessi) on Instagram. What's the conditional probability that you will respond to a suggestion to follow Cristiano Ronaldo (@cristiano)?

A really important fact is that conditional probabilities are **not symmetric**:

$$P(A \mid B) \neq P(B \mid A)$$

As a quick counter-example, let the events A and B be as follows:

- A: "you can dribble a basketball"
- B: "you play in the NBA"

Probability basics: conditioning

- A: "you can dribble a basketball"
- B: "you play in the NBA"

Clearly $P(A \mid B) = 1$: every NBA player can dribble a basketball!

Probability basics: conditioning

- A: "you can dribble a basketball"
- B: "you play in the NBA"

But $P(B \mid A)$ is nearly zero!

Uncertain outcomes and probability models

An uncertain outcome (more formally called a “random process”) has two key properties:

1. The set of possible outcomes, called the sample space, *is known* beforehand.

2. The particular outcome that occurs is *not known* beforehand.

We denote the **sample space** as Ω, and some particular element of the sample space as $\omega \in \Omega$

Uncertain outcomes and probability models

Examples:

1. NBA finals, Golden State vs. Toronto:

$$\Omega = \{\text{4-0, 4-1, 4-2, 4-3, 3-4, 2-4, 1-4, 0-4}\}$$

Uncertain outcomes and probability models

Examples:

1. NBA finals, Golden State vs. Toronto:

$$\Omega = \{\text{4-0, 4-1, 4-2, 4-3, 3-4, 2-4, 1-4, 0-4}\}$$

2. Temperature in degrees F in Austin on a random day:

$$\Omega = [10, 115]$$

Uncertain outcomes and probability models

Examples:

1. NBA finals, Golden State vs. Toronto:

$$\Omega = \{\text{4-0, 4-1, 4-2, 4-3, 3-4, 2-4, 1-4, 0-4}\}$$

2. Temperature in degrees F in Austin on a random day:

$$\Omega = [10, 115]$$

3. Number of no-shows on an AA flight from Austin to DFW:

$$\Omega = \{0, 1, 2, \ldots, N_{\text{seats}}\}$$

Uncertain outcomes and probability models

Examples:

1. NBA finals, Golden State vs. Toronto:

$$\Omega = \{\text{4-0, 4-1, 4-2, 4-3, 3-4, 2-4, 1-4, 0-4}\}$$

2. Temperature in degrees F in Austin on a random day:

$$\Omega = [10, 115]$$

3. Number of no-shows on an AA flight from Austin to DFW:

$$\Omega = \{0, 1, 2, \ldots, N_{\text{seats}}\}$$

4. Poker hand

$$\Omega = \text{all possible five-card deals from a 52-card deck}$$

An **event** is a *subset of the sample space*, i.e. $A \subset \Omega$. For example:

1. NBA finals, Golden State vs. Toronto. Let A be the event "Toronto wins". Then

$$A = \{\text{3-4, 2-4, 1-4, 0-4}\} \subset \Omega$$

2. Austin weather. Let A be the event "cooler than 90 degrees". Then

$$A = [10, 90) \subset [10, 115]$$

3. Flight no-shows. Let A be "more than 5 no shows":

$$A = \{6, 7, 8, \ldots, N_{\text{seats}}\}$$

Some set theory concepts

We need some basic set-theory concepts to make sense of probability, since the sample space Ω is a set, and since "events" are subsets of Ω.

Union: $A \cup B = \{\omega : \omega \in A \text{ or } \omega \in B\}$

Intersection: $A \cap B = \{\omega : \omega \in A \text{ and } \omega \in B\}$

Complement: $A^C = \tilde{A} = \{\omega : \omega \notin A\}$

Difference: $A \setminus B = \{\omega : \omega \in A, \omega \notin B\}$

Disjointness: A and B are disjoint if $A \cap B = \emptyset$ (the empty set).

Axioms of probability (Kolmogorov)

These are the ground rules!

Consider an uncertain outcome with sample space Ω. "Probability" $P(\cdot)$ is a set function that maps Ω to the real numbers, such that:

1. Non-negativity: For any event $A \subset \Omega$, $P(A) \geq 0$.
2. Normalization: $P(\Omega) = 1$ and $P(\emptyset) = 0$.
3. Finite additivity: If A and B are disjoint, then $P(A \cup B) = P(A) + P(B)$.

3a. Finite additivity (general): For any sets A and B, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
(bonus: prove this with set theory!)

Not that intuitive! Notice no mention of frequencies...

Summary of terms

- Uncertain outcome/"random process": we know the possibilities ahead of time, just not the specific one that occurs
- Sample space: the set of possible outcomes
- Event: a subset of the sample space
- Probability: a function that maps events to real numbers and that obeys Kolmogorov's axioms

OK, so how do we actually *calculate* probabilities?

Calculation

Now that we have an understanding of the axioms, notation, and interpretation, how do we calculate probabilities?

Counting!

(review 6.1.3-6.1.4 in the QSS book … ways to count objects in structured sets are discussed)

Suppose our sample space Ω is a finite set consisting of N elements $\omega_1, \ldots, \omega_N$.

Suppose further that $P(\omega_i) = 1/N$: each outcome is equally likely, i.e. we have a discrete uniform distribution over possible outcomes.

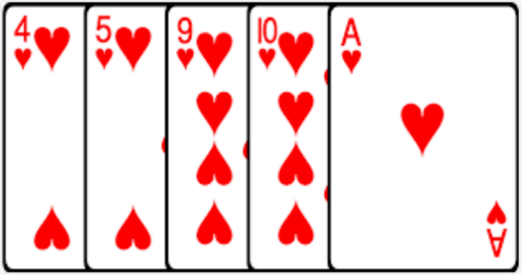
Then for each set $A \subset \Omega$,

$$P(A) = \frac{|A|}{N} = \frac{\text{Number of elements in } A}{\text{Number of elements in } \Omega}$$

That is, to compute $P(A)$, **we just need to count how many elements are in** A.

Counting example

Someone deals you a five-card poker hand from a 52-card deck. What is the probability of a flush (all five cards the same suit)?



Note: this is a very historically accurate illustration of probability, given its origins among bored French aristocrats!

Counting example

- Our sample space has $N = \binom{52}{5} = 2{,}598{,}960$ possible poker hands, each one equally likely.

- How many possible flushes are there? Let's start with hearts:
  $\rightarrow$ There are 13 hearts

Counting example

- Our sample space has $N = \binom{52}{5} = 2{,}598{,}960$ possible poker hands, each one equally likely.
- How many possible flushes are there? Let's start with hearts:
  $\rightarrow$ There are 13 hearts
  $\rightarrow$ To make a flush with hearts, you need any 5 of these 13 cards.

Counting example

– Our sample space has $N = \binom{52}{5} = 2{,}598{,}960$ possible poker hands, each one equally likely.

– How many possible flushes are there? Let's start with hearts:
$\rightarrow$ There are 13 hearts
$\rightarrow$ To make a flush with hearts, you need any 5 of these 13 cards.
$\rightarrow$ Thus there are $\binom{13}{5} = 1287$ possible flushes with hearts.

Counting example

- Our sample space has $N = \binom{52}{5} = 2{,}598{,}960$ possible poker hands, each one equally likely.

- How many possible flushes are there? Let's start with hearts:
  $\rightarrow$ There are 13 hearts
  $\rightarrow$ To make a flush with hearts, you need any 5 of these 13 cards.
  $\rightarrow$ Thus there are $\binom{13}{5} = 1287$ possible flushes with hearts.
  $\rightarrow$ The same argument works for all four suits, so there are $4 \times 1287 = 5{,}148$ flushes. Thus:

Counting example

- Our sample space has $N = \binom{52}{5} = 2{,}598{,}960$ possible poker hands, each one equally likely.

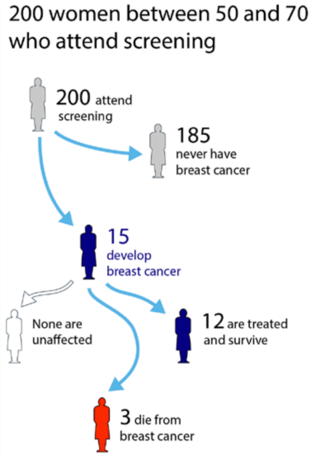
- How many possible flushes are there? Let's start with hearts:
  $\rightarrow$ There are 13 hearts
  $\rightarrow$ To make a flush with hearts, you need any 5 of these 13 cards.
  $\rightarrow$ Thus there are $\binom{13}{5} = 1287$ possible flushes with hearts.
  $\rightarrow$ The same argument works for all four suits, so there are 4 $\times$ 1287 = 5,148 flushes. Thus:

$$P(\text{flush}) = \frac{|A|}{|\Omega|} = \frac{5148}{2598960} = 0.00198079$$

So we know how to count, but what about conditioning?

Probability trees are very useful for this task! This involves counting at different levels of the tree.

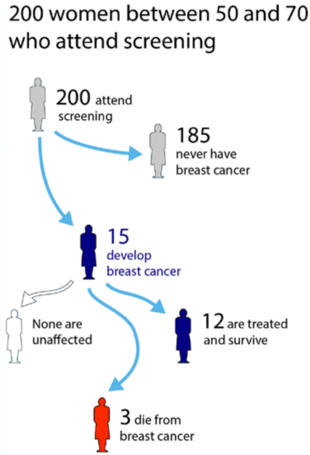
Conditioning example: mammograms



- $P(\text{cancer}) = \frac{15}{200}$
- $P(\text{die,cancer}) = \frac{3}{200}$
- $P(\text{die} \mid \text{cancer}) = \frac{3}{15}$

– In general, we can estimate the conditional probability as:

Conditioning example: mammograms



- $P(\text{cancer}) = \frac{15}{200}$
- $P(\text{die,cancer}) = \frac{3}{200}$
- $P(\text{die} \mid \text{cancer}) = \frac{3}{15}$
  - In general, we can estimate the conditional probability as:

$$P(A \mid B) = \frac{\text{Frequency of } A \text{ and } B \text{ both happening}}{\text{Frequency of } B \text{ happening}}$$

This is actually a new axiom

The multiplication rule – it is an axiom since it can't be derived from the original axioms.

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

Alternate version

We can also use this alternative version if we want to go in reverse, from a conditional probability to a joint probability.

It says the same thing with the terms rearranged.

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