Chapter 1: Experiments with random outcomes

Jess Kunke

MATH/STAT 394: Probability I (Summer 2022 A-term)

Outline

Outline

Motivation

Probability theory: building mathematical models of experiments that have random/uncertain/noisy outcomes

- Recall: probability is a tool to measure uncertainty
- How? Probabilities measure sets
- To measure finite sets, we will count the number of elements in them
  - Clever rules
  - Three basic types of problems
- We will also learn useful problem-solving strategies based on the rules of probability

We will introduce terms, concepts, and notation from set theory and counting rules **in context along the way**

- See Appendices B-C for a concise presentation with practice exercises
- Refer to App. A and D if needed for calculus, series, and more

Probability model

coin toss: sample space includes H, T

A **probability model**, **probability space**, or **random experiment** consists of three parts:

omega

1. The **sample space** Ω is the set of all possible **outcomes** of the experiment
2. The **event space** $\mathcal{F}$ is the set of all possible **events** (subsets of Ω)
   - This essentially corresponds to all the questions you can ask about the experiment's result
3. The third part is a **probability measure**; more on this soon

Let's look at some examples...

- We're going to introduce a bunch of terminology and notation in context
- Don't get too hung up on remembering it all now! We will practice it together in class and you will practice it in your homework
- Do take notes and practice the correct terminology and notation
- Refer to the textbook, annotated slides, lecture recordings as needed

Probability model

Example

Consider the result of a **single coin toss**.

- What is the sample space Ω? What is the event space $\mathcal{F}$?

$\{ \{ \} \; \{$

$\}$

$$\Omega = \{H, T\} = \{T, H\}$$

set = unordered collection of items

$$\mathcal{F} = \{\text{all subsets of } \Omega\} = \{\{H\}, \{T\}, \{H, T\}, \emptyset\}$$

elements of Ω = cardinality of Ω $= |\Omega| = 2$

$$|\mathcal{F}| = 4$$

Probability model

$\underline{H,T}$ $\underline{H,T}$ $\underline{H,T}$

3-tuples

(H, H, T)

(T, H, H)

$\cancel{\{H, H, T\} = \{T, H, H\}}$

Example

Consider a series of **three coin flips**.

- What is the sample space Ω? What is the event space $\mathcal{F}$?

$$\Omega = \{(H,H,H),\ (H,H,T),\ (H,T,H),\ (T,H,H),\ (H,T,T),\ (T,H,T),\ (T,T,H),\ (T,T,T)\}$$

$$= \{H,T\} \times \{H,T\} \times \{H,T\}$$

$$= \{H,T\}^3$$

$$= \{(x_1, x_2, x_3) : x_i \in \{H,T\}\ \forall i = 1,2,3\}$$

is an element of

for all

the set of all 3-tuples ... such that

$$|\Omega| = 2^3 = 8$$

ordered pair

$(2, 2)$

2-tuple

$R \times R$

R^2

$\mathcal{F} = \{\underline{\text{power set of } \Omega}\}$

$= \{\emptyset,$ all one-element subsets of Ω such as $\boxed{\{(H, H, H)\}}$,

all two-element subsets of Ω such as $\{(H, H, H), (H, H, T)\}$,

$\ldots, \Omega\}$

$|\mathcal{F}| = 2^{|\Omega|} = 2^8 = 256$

$\{(H, H, T), (H, T, H), (T, H, H),$

$(H, H, H)\}$ $\longleftarrow$ <u>at least 2 heads</u>

Key terms/concepts so far

- Set (unordered collection of items without repetition)
- Tuple (ordered list of items, can repeat)
- Different notation for a set $\{1, 2, 3\}$ versus a tuple $(1, 2, 3)$
- Writing a set in words, in formulas, or as a list of elements
- Subsets, elements of a set
  - (Sometimes useful) definition of a subset: We say $A \subseteq B$ if is a subset of

$$\text{if } \boxed{\omega \in A} \overset{\text{implies}}{\underset{\text{then}}{\Longrightarrow}} \boxed{\omega \in B} \quad \text{(see next slide too)}$$

- Special subsets of a set A: singletons, $\emptyset$, A itself

  $\{(H, H, H)\}$
- Cartesian products

Note:

- Authors made unfortunate choice to use lowercase and uppercase omega for elements and sample spaces (makes it awkward to read the math aloud)
- Nothing special about this choice of letters, but it is typical to use lowercase (uppercase) letters for elements (sets)

A quick but handy note on subsets

Given two sets A and B, how do you prove that A is a subset of B?

if $\underline{w \in A} \Rightarrow \underline{w \in B}$, then $\underline{A \subseteq B}$

if every element of A is also in B, then $A \subseteq B$.

$A = \{1, 2, 3\}$ $B = \{1, 2, 4\}$ $3 \in A$ but $3 \notin B$.

$\therefore A \not\subseteq B$.

How do you prove that A is **not** a subset of B? $A \not\subseteq B$

$\cancel{w \in B \not\Rightarrow w \in A}$

Show $\exists\, w \in A$ such that $w \notin B$.

(there exists)

that shows that $w \in A \not\Rightarrow w \in B$

We'll see an example of this later, on slide 34

Probability model: some more detail

A **probability model**, **probability space**, or **random experiment** consists of three parts:

1. The **sample space** Ω is the set of all possible **outcomes** of the experiment
2. The **event space** $\mathcal{F}$ is the set of all possible **events** (subsets of Ω)
    - This essentially corresponds to all the questions you can ask about the experiment's result
    - The set of all subsets of Ω is called the **power set** of Ω
    - Sometimes we either have to or want to choose something other than the power set; more on this later
3. The **probability measure** P is a function from $\mathcal{F}$ into the real numbers $\mathbb{R}$
    - Defines how we are measuring probability/uncertainty
    - Also called probability, or probability distribution

What parts you need to define depends on the question:

- "How many ways...?" $\rightarrow$ only need sample space, event space
- "What is the probability that...?" $\rightarrow$ need full probability model

Set operations

To define probability measures, we will need some fundamental set operations:

Definition

intersection

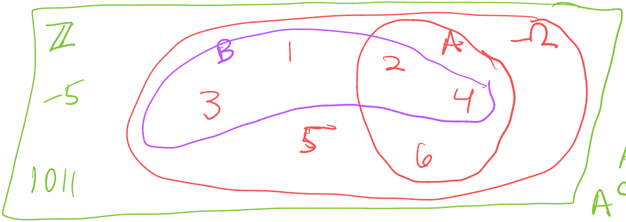
Let A, B be two subsets of a set Ω.

- Their **intersection** is $A \cap B = \{\omega \in \Omega : \omega \in A \text{ and } \omega \in B\}$
  - Two sets A, B are called **disjoint** if $A \cap B = \emptyset$
- Their **union** is $A \cup B = \{\omega \in \Omega : \omega \in A \text{ or } \omega \in B\}$ Union
  - In logic/sets/probability, "or" means "and/or"
- Their **difference** is $A \setminus B := \{\omega \in \Omega : \omega \in A \text{ and } \omega \notin B\}$
- The **complement** of A in Ω is $A^c = \{\omega \in \Omega : \omega \notin A\}$
  - Notice that to talk about the complement of A, we have to talk about A as a subset of some larger set Ω

$A^c = \mathbb{Z} \setminus \{2,4,6\}$

Let's visualize this using a Venn diagram in an example: words

$$\Omega = \{1,2,3,4,5,6\}, \quad A = \{x \in \Omega : x \text{ is even}\}, \quad B = \{1,2,3,4\}$$



$A \cap B = \{2,4\}$

$A \cup B = \{1,2,3,4,6\}$
$= \Omega \setminus \{5\}$

$A \setminus B = \{6\}$

$A^c = \{1,3,5\}$

A **probability model**, **probability space**, or **random experiment** consists of three parts:

1. The **sample space** Ω is the set of all possible **outcomes** of the experiment.
2. The **event space** $\mathcal{F}$ is the set of all possible **events**, or subsets of Ω.
3. The **probability measure** P is a function from $\mathcal{F}$ into the real numbers $\mathbb{R}$. Each event A has a probability $P(A)$, and P satisfies the following three axioms (Kolmogorov's axioms):

$\cancel{-3, 10}$

i $0 \leq P(A) \leq 1$ for each event A

with probability 1

ii $P(\Omega) = 1$ and $P(\emptyset) = 0$

$A_1 = \{1 \text{ head}\}$

$A_2 = \{2 \text{ heads}\}$

iii If $A_1, A_2, A_3, \ldots$ is a sequence of **pairwise disjoint** events (i.e.[1] mutually exclusive, $A_i \cap A_j = \emptyset$ for $i \neq j$), then

countable additivity

$A_1 \cap A_2 = \emptyset$

$P(A_1 \cup A_2) = P(\{1 \text{ or } 2 \text{ heads}\}) = P(A_1) + P(A_2)$

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

Prove that the sum of the probabilities of all possible outcomes must be 1.

$A_3, A_4, \ldots = \emptyset$

$P\left(\bigcup_{w \in \Omega} \{w\}\right) \overset{iii}{=} P(\Omega) \overset{ii}{=} 1$

[1]Note: i.e. means "in other words" while e.g. means "for example"

Fact

If $A_1, A_2, \ldots, A_n$ are **pairwise disjoint** *events, then*

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n).$$

Example

Roll a fair six-sided die once. Fair means that the outcomes are equally likely.

- ✓ What is the sample space Ω? $\Omega = \{1, 2, 3, 4, 5, 6\}$
- ✓ What is the probability of each possible outcome?
- ✓ What is the probability of rolling an even number?

$P(1) = \ldots = P(6)$, $\sum_{i=1}^{6} P(i) = P(\Omega) = 1.$

$\Rightarrow P(i) = 1/6 \quad \forall i \in \Omega.$

Let $A = \{2, 4, 6\}$. $P(A) = P(2) + P(4) + P(6) = \frac{1}{2}.$

A useful consequence of Axiom iii

Fact

If $A_1, A_2, ..., A_n$ are **pairwise disjoint** *events, then*

$$P(A_1 \cup A_2 \cup ... \cup A_n) = P(A_1) + P(A_2) + ... + P(A_n).$$

Example

1 2 2 4 5 6

Now suppose we change the 3 on the die to another 2. Define the new probability measure.

$$\ell = \underbrace{\tilde{P}(1) = \tilde{P}(4) = \tilde{P}(5) = \tilde{P}(6)} = \underbrace{\tfrac{1}{2}\tilde{P}(2)}, \quad \underbrace{\tilde{P}(3) = 0}$$

$P(1) = ?$ $\quad P(2) = ?$

$$\boxed{1 = \sum_{i=1}^{6} P(i)} = \underline{P(1) + P(2) + P(3) + P(4) + P(5) + P(6)}$$

$$= \ell + 2\ell + 0 + \ell + \ell + \ell$$

$$= 6\ell \qquad \ell = 1/6$$

Outline

Probability models: an introduction

Random sampling

Infinitely many outcomes

Useful consequences of Kolmogorov's axioms

Random variables: an introduction

Sampling

Sampling: choosing objects from a given set according to a probability model

- Many problems can be framed as sampling problems
- Sampling methods are the mechanisms we use to model our random experiments
- Outcomes may be equally likely or not depending on the problem

Note: Choosing "at random" (imprecise) typically implies choosing **uniformly** at random (precise), meaning that all outcomes are equally likely.

We will start by considering experiments with equally likely outcomes.

Experiments with equally likely outcomes

Probabilities must sum to 1, and if the outcomes are equally likely then the probabilities must all be equal. In this case we have the following handy result:

Fact

If the sample space Ω has finitely many elements and equally likely outcomes, then we have

$$P(\omega) = \frac{1}{|\Omega|} \textit{ for any outcome } \omega \in \Omega \textit{ and } P(A) = \frac{|A|}{|\Omega|} \textit{ for any event } A,$$

where $|A|$ is the number of elements of A, also called the **cardinality** *of A.*

Example

Let's return to the example of flipping a fair coin three times.

- What is the sample space Ω? $\Omega = \{H, T\}^3$ $|\Omega| = 2^3 = 8$
- What is the probability of getting at least two heads in a row?

$$\left| \underbrace{\{(H,H,T), (T,H,H), (H,H,H)\}}_{\equiv A} \right| = 3$$

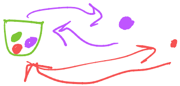
$$P(A) = \frac{3}{8}$$

Types of sampling mechanisms

Purple red

Three basic types of sampling with equally likely outcomes:

1. Ordered sampling with replacement
   - e.g. the last example (three coin flips)
   - Repetitions of a simple experiment
   - Cartesian products, n-tuples



2. Ordered sampling without replacement
   - Permutations, arrangements

3. Unordered sampling without replacement
   - Combinations/sets

Comments:

- What about unordered sampling with replacement? (cliffhanger! later)
- Keep straight the sample space Ω versus the underlying sets of objects
  - e.g. $C = \{H, T\}$ vs. $\Omega = \{(x_1, x_2, x_3) : x_i \in C \; \forall \, i\}$
- Remember the different notation for *unordered sets* $\{a_1, a_2, \ldots\}$ versus *ordered n-tuples* $(a_1, a_2, \ldots)$
- In general, problems can involve a combination of mechanisms
- We will start by examining one at a time

1. Ordered sampling with replacement

Let's see a slightly more general example than the three coin flips:

Example

Suppose license plates are generated randomly by choosing three letters (say from the English alphabet) followed by three digits (0 through 9), and repeats are allowed. What is the probability that you get a license plate whose first two numbers are in $\{1, 2, 3, 4, 5\}$, possibly with repeats?

______	______	______	______	______	______
letter	letter	letter	number	number	number

$L = \{A, B, \ldots, Z\}$ $\quad |L| = 26$ $\quad |\Omega| = 26^3 10^3$

$D = \{0, \ldots, 9\}$ $\quad |D| = 10$

$\Omega = L \times L \times L \times D \times D \times D = L^3 D^3$

$A = \{$ license plates with... $\}$

$= L^3 \times \{1, \ldots, 5\}^2 \times D \Rightarrow |A| = 26^3 \cdot 5^2 \cdot 10$

$P(A) = 26^3 \cdot 5^2 \cdot 10 / 26^3 10^3 = 5^2 / 10^2$.