Introduction to probability

Lecture 1a (S24400 F24)

Statistics, uncertainty of outcomes, probability

- Statistics is the science that concerns the collection, organization, analysis, and interpretation of data.
- Often, data are produced from experiments or processes which produce more than one possible outcomes.
- The occurrence of each possible outcome is uncertain.
- We need to describe and quantify the uncertainty of possible outcomes.

$\Longrightarrow$ Probability is the best language to formulate uncertainty.

(Other motivations in using uncertainty/randomness, e.g., optimization algorithm, lottery games.)

Warmup puzzle

Puzzle

A fair coin is flipped repeatedly until
the first time we see the sequence HH or TH.

- Player A wins if HH comes up first.
- Player B wins if TH comes up first.

Question

Is it a fair game?
What are the odds of winning for each player?

Warmup puzzle

Answer This is not a fair game:

- Player A wins if the first two coins are HH (25% chance).
- Otherwise, player B wins (75% chance).

Remarks

Computing probabilities may not be as straightforward as it appears.

(Questions: What are all possible outcomes? What are the probabilities of the outcomes?)

Sample spaces & events

When we perform an experiment or observe the result of a random process, the **sample space** Ω is the set of all possible outcomes.

1. How many inches of rain today?
   $\Omega = [0, \infty)$
2. Measure how many times a die is rolled until you get a 6,
   $\Omega = \{1, 2, 3, \dots\}$
3. There may be multiple sensible choices for how to record outcomes, depending on the probability of event we are interested in, i.e., there can be **multiple sample spaces** for one random experiment.

   Survey 5 people and ask whether they have ever had Covid.

   - $\Omega = \{0, 1, \dots, 5\}$ (total number that had Covid); or
   - $\Omega = \{\texttt{YYYYY}, \texttt{YNYNY}, \texttt{NNYNN}, \dots, \texttt{NNNNN}\}$ (sequence of answers in order)

Sample spaces and events

We may be interested in whether some particular type of outcome has occurred.

An **event** is any subset A of the sample space Ω that we are interested in.

In the aforementioned experiments, we may be interested in:

1. Did it rain today?
   The event of interest is $A = (0, \infty)$ (subset of $\Omega = [0, \infty)$)
2. Did we roll a 6 on our first try?
   The event of interest is $A = \{1\}$ (subset of $\Omega = \{1, 2, 3, \dots\}$)
3. Did fewer than 3 people in our Covid survey of 5 people answer yes?

   Depending on the representation of the outcome in the sample space Ω, the expression of the event of interest is

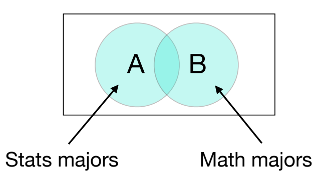
   - $A = \{0, 1, 2\}$ subset of $\Omega = \{0, 1, 2, 3, 4, 5\}$)
   - Or, $A = \{\texttt{NNNNN}, \texttt{YNNNN}, \texttt{YNYNN}, \text{etc}\}$
     (subset of $\Omega = \{\texttt{YYYYY}, \texttt{YNYNY}, \texttt{NNYNN}, \dots, \texttt{NNNNN}\}$)

Set notation (and Venn diagram)

Running example

- Experiment: choose one student from the class, at random
- Ω = set of all students in the class
- A = all stats majors, B = all math majors

Illustration by a Venn diagram:



Set operations

Running example

- Experiment: choose one student from the class
- Ω = set of all students in the class
- A = all stats majors, B = all math majors

	Definition	Example/interpretation
$\varnothing$	the empty set	$B = \varnothing \Leftrightarrow$ there are no math majors in the class
$A \cap B$	intersection of A & B (both A and B occur)	the set of all students who are double majors in stats&math
$A \cup B$	union of A & B (A or B occur)	the set of students who are majoring in stats or math (including both)
B^c ($\Omega \setminus B$)	complement of B (B does not occur)	the set of all students that are not math majors

(Here "or" includes "and", i.e., both.)

Set relations

Running example

- Experiment: choose one student from the class
- Ω = set of all students in the class
- A = all stats majors, B = all math majors

	Definition	Example/interpretation
$A \subset B$ or $A \subseteq B$	A is a subset of B (and possibly $A = B$)	all stats majors in the class are also math majors
$A \subsetneq B$	A is a strict subset of B ($A = B$ not allowed)	all stats majors are also math majors, but not all math majors are stats majors
A and B are disjoint	$A \cap B = \varnothing$	there are no stats/math double majors in the class

Laws of set theory

De Morgan's laws:

- $(A \cup B)^c = A^c \cap B^c$
- $(A \cap B)^c = A^c \cup B^c$

Distributive laws:

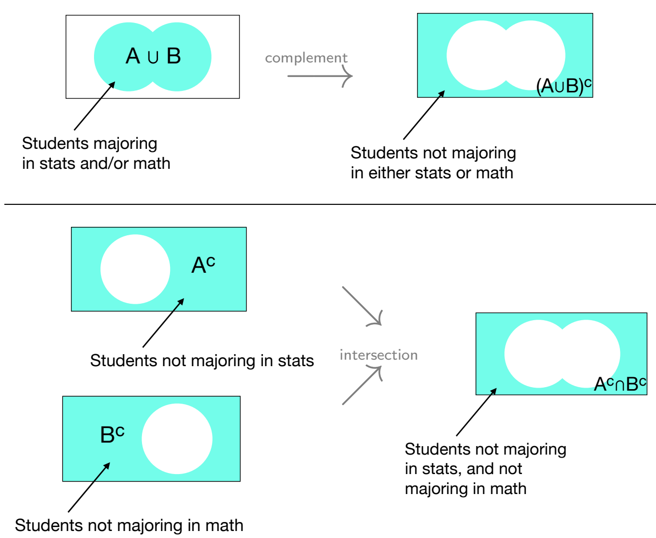
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Commutative laws:

- $A \cup B = B \cup A$
- $A \cap B = B \cap A$

Laws of set theory illustration

- Illustration for De Morgan's law: $(A \cup B)^c = A^c \cap B^c$



Probability measures

A **probability measure** is a function $\mathbb{P}(\cdot)$ that specifies a probability for each subset of Ω that we might consider.

Axioms:

1. $\mathbb{P}(\Omega) = 1$
2. $\mathbb{P}(A) \geq 0$ for any event A
3. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ for any disjoint events A and B

   And more generally,
   $\mathbb{P}(A \cup B \cup C \cup D \dots) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) + \mathbb{P}(D) + \dots$
   for any mutually disjoint events $A, B, C, D, \dots$
   (holds for finite or countably infinite unions)

Probability measure properties

Properties:

(i) $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$

(ii) $\mathbb{P}(\varnothing) = 0$

(iii) If $A \subset B$ then $\mathbb{P}(A) \leq \mathbb{P}(B)$

(iv) Addition law: $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

These properties can be proved from the axioms.

Example — Proof of the first property (i):

$$\underset{\substack{\uparrow \\ \text{by axiom 1}}}{1 =} \mathbb{P}(\Omega) = \mathbb{P}(A \cup A^c) \underset{\substack{\uparrow \\ \text{by axiom 3}}}{=} \mathbb{P}(A) + \mathbb{P}(A^c)$$

Events and probability measures

Example: Measure how many times a die is rolled until you get a 6.
$\Omega = \{1, 2, 3, \dots\}$

- How likely do we roll a 6 on our first try?
  The event of interest is $A = \{1\}$ (subset of $\Omega = \{1, 2, 3, \dots\}$)
- Would it take more than 5 rolls, to roll a 6?
  $B = \{6, 7, 8, \dots\}$ (subset of $\Omega = \{1, 2, 3, \dots\}$)
- Did the first 6 occur on an even count roll?
  $C = \{2, 4, 6, \dots\}$ (subset of $\Omega = \{1, 2, 3, \dots\}$)

A probability measure on Ω would specify the probability of each outcome, consequently the probability of each of these events and any other event.

For example, for a fair die, $\mathbb{P}(A) = \dfrac{1}{6}$

Finite uniform distribution

In many examples, the probability measure will be a uniform distribution over a finite sample space Ω, where each outcome is equally likely to occur.

- Roll a fair die then record the outcome. $\Omega = \{1, 2, 3, 4, 5, 6\}$
  Each number is equally likely to face up (since the die is fair).
- Draw two cards from a shuffled deck. $\Omega = \{A\clubsuit A\diamondsuit, A\clubsuit A\heartsuit, \dots, Q\spadesuit K\spadesuit\}$
  Each hand is equally likely (when the deck is fully shuffled).

Finite uniform distribution (remarks)

Caution:

Probability measures are not always uniform — So, don't assume.

For example,

- Survey 5 people and ask whether they had Covid.
  $\Omega = \{0, 1, \dots, 5\}$
- Roll two fair dice, then record the sum.
  $\Omega = \{2, 3, \dots, 12\}$

Counting, permutations, & combinations (ordered)

Ordered samples

Number of ways to choose an ordered sample of size r from a set of size n:

- n^r, if sampling with replacement
- $n(n-1)(n-2)\dots(n-r+1)$, if sampling without replacement

Choose two from {APPLE ORANGE PEAR}

Sampling with replacement ($3 \times 3 = 3^2$ outcomes):

$$\begin{matrix}\texttt{APPLE}\\\texttt{APPLE}\end{matrix}, \begin{matrix}\texttt{APPLE}\\\texttt{ORANGE}\end{matrix}, \begin{matrix}\texttt{APPLE}\\\texttt{PEAR}\end{matrix}, \begin{matrix}\texttt{ORANGE}\\\texttt{APPLE}\end{matrix}, \begin{matrix}\texttt{ORANGE}\\\texttt{ORANGE}\end{matrix}, \begin{matrix}\texttt{ORANGE}\\\texttt{PEAR}\end{matrix}, \begin{matrix}\texttt{PEAR}\\\texttt{APPLE}\end{matrix}, \begin{matrix}\texttt{PEAR}\\\texttt{ORANGE}\end{matrix}, \begin{matrix}\texttt{PEAR}\\\texttt{PEAR}\end{matrix}$$

Sampling without replacement (3×2 outcomes):

$$\begin{matrix}\texttt{APPLE}\\\texttt{ORANGE}\end{matrix}, \begin{matrix}\texttt{APPLE}\\\texttt{PEAR}\end{matrix}, \begin{matrix}\texttt{ORANGE}\\\texttt{APPLE}\end{matrix}, \begin{matrix}\texttt{ORANGE}\\\texttt{PEAR}\end{matrix}, \begin{matrix}\texttt{PEAR}\\\texttt{APPLE}\end{matrix}, \begin{matrix}\texttt{PEAR}\\\texttt{ORANGE}\end{matrix}$$

Counting, permutations, & combinations (unordered)

Unordered samples

Number of ways to choose an unordered sample of size r from a set of size n (without replacement):

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

Why?

- # of ways of ordered lists of r items: $n(n-1)\dots(n-r+1) = \frac{n!}{(n-r)!}$
- Each possible list has $r!$ many orderings (all equivalent in unordered case), then a factor of $r!$ need to be divided to reduce the overcounting.

Choose two from {APPLE ORANGE PEAR}

$$\begin{matrix}\texttt{APPLE}\\\texttt{ORANGE}\end{matrix}, \begin{matrix}\texttt{APPLE}\\\texttt{PEAR}\end{matrix}, \cancel{\begin{matrix}\texttt{ORANGE}\\\texttt{APPLE}\end{matrix}}, \begin{matrix}\texttt{ORANGE}\\\texttt{PEAR}\end{matrix}, \cancel{\begin{matrix}\texttt{PEAR}\\\texttt{APPLE}\end{matrix}}, \cancel{\begin{matrix}\texttt{PEAR}\\\texttt{ORANGE}\end{matrix}}$$

Counting, permutations, & combinations (sorting into groups)

Sorting into groups:

There are $\binom{n}{n_1, n_2, \dots, n_k} = \frac{n!}{n_1! \cdot n_2! \cdot \dots \cdot n_k!}$ ways

to group n objects into k labeled groups of size $n_1 + n_2 + \dots + n_k = n$.

Special case: $k = 2$ (unordered samples)

Note that $\binom{n}{r} = \binom{n}{n-r} = \binom{n}{r, n-r}$,

i.e. choose a group of size r $\iff$ split into two groups of size r and $n-r$

Counting examples (Same subgroup size)

Be careful about sorting into groups that are the same size:

For example: Do we intend to count these splits as same or different?

$$\begin{matrix}\texttt{APPLE}\\\texttt{ORANGE}\end{matrix} \& \begin{matrix}\texttt{BANANA}\\\texttt{PEAR}\end{matrix} \quad \text{vs.} \quad \begin{matrix}\texttt{BANANA}\\\texttt{PEAR}\end{matrix} \& \begin{matrix}\texttt{APPLE}\\\texttt{ORANGE}\end{matrix}$$

Example Count how many ways each is possible:

1. Split 10 students into Section 1 and Section 2, each with 5 students:

$$\binom{10}{5} = \frac{10!}{5!\,5!} = 252$$

2. Split 10 students into two teams, each with 5 students:

$$\binom{10}{5} \cdot \frac{1}{2} = 126$$

Counting examples (Different subgroup sizes)

Count how many ways each is possible:

- Split 10 students into two teams, with 4 students / 6 students:

$$\binom{10}{4} = \frac{10!}{4!\,6!} = 210$$

- Split 10 students into two teams of any size (≥ 1 student in each):

$$\frac{2^{10}-2}{2} = 511$$

or

$$\binom{10}{1} + \binom{10}{2} + \binom{10}{3} + \binom{10}{4} + \binom{10}{5} \cdot \frac{1}{2} = 511$$

Probability & counting examples (draw 20 cards)

Example

20 cards are drawn randomly from a standard 52 card deck.
What is the probability that exactly half of the cards drawn are red?

Answer:

hands of 10 red cards & 10 black cards

$$\frac{\binom{26}{10} \cdot \binom{26}{10}}{\binom{52}{20}} = 0.224$$

total # hands of 20 cards

(Assumption: Uniform probability measure over $\Omega = \{$all possible hands of 20 cards$\}$)

Probability & counting examples (arrange 10 coins)

Example

If you arrange five pennies and five dimes into a random order,
what is the probability that they alternate?

Answer:

- Total # arrangements is $\binom{10}{5} = 252$
- Total # alternating arrangements $= 2$
- $\rightsquigarrow$ probability $= \frac{2}{252}$

Remarks:

- Assumption: Uniform probability measure over $\Omega = \{$all possible arrangements$\}$
- Counting: Consider having 10 slots in a row, then putting 5 pennies into 5 slots.
- Optional exercise: Consider an alternative sample space treating the coins as distinct.

Warmup puzzle revisit

- **Puzzle:** A fair coin is flipped repeatedly
  until the first time we see the sequence HH or TH.
  - Player A wins if HH comes up first.
  - Player B wins if TH comes up first.

  What are the probabilities of winning for each player?

What is Ω?

```
HH      TH
HTH     TTH
HTTH    TTTH
HTTTH   TTTTH
HTTTTH  TTTTTH
...     ...
```

Questions: What is the probability of each outcome? (Review infinite series)

What are the winning probabilities and the corresponding sample space if B wins when HT comes up first?