Probability - Lecture notes - Unit One

Unit One notes

Events and outcomes

01 Theory

Events and outcomes – informally

- An **event** is a *description* of something that can happen.
- An **outcome** is a *complete description* of something that can happen.

All outcomes are events. An event is usually a *partial* description. Outcomes are events given with a *complete* description.

Here ‘complete’ and ‘partial’ are within the context of the **probability model**.

- ⚠ It can be misleading to say that an ‘outcome’ is an ‘observation’.
  - ‘Observations’ occur in the *real world*, while ‘outcomes’ occur in the *model*.
  - To the extent the model is a good one, and the observation conveys *complete* information, we can say ‘outcome’ for the observation.

Notice:

- Because outcomes are *complete*, no two distinct outcomes could *actually happen* in a run of the experiment being modeled.

When an event happens, the *fact* that it has happened constitutes **information**.

Events and outcomes – mathematically

- The **sample space** is the *set of possible outcomes*, so it is the set of the complete descriptions of everything that can happen.
- An **event** is a *subset* of the sample space, so it is a *collection of outcomes*.

- For mathematicians: some “wild” subsets are not *valid* events. Problems with infinity and the continuum...

Notation

- Write S for the set of possible outcomes, $s \in S$ for a single outcome in S.
- Write $A, B, C, \dots \subset S$ or $A_1, A_2, A_3, \dots \subset S$ for some events, subsets of S.
- Write $\mathcal{F}$ for the collection of all events. This is frequently a *huge* set!

- Write $|A|$ for the **cardinality** or *size* of a set A, i.e. the *number of elements it contains*.

Using this notation, we can consider an *outcome itself as an event* by considering the “singleton” subset $\{\omega\} \subset S$ which contains that outcome alone.

02 Illustration

Example - Coin flipping

Flip a fair coin two times and record both results.

- *Outcomes:* sequences, like HH or TH.
- *Sample space:* all possible sequences, i.e. the set $S = \{HH, HT, TH, TT\}$.
- *Events:* for example:
  - $A = \{HH, HT\} =$ “first was heads”
  - $B = \{HT, TH\} =$ “exactly one heads”
  - $C = \{HT, TH, HH\} =$ “at least one heads”

With this setup, we may combine events in various ways to generate other events:

- *Complex events:* for example:
  - $A \cap B = \{HT\}$, or in words:

    $$\text{“first was heads” AND “exactly one heads”} = \text{“heads-then-tails”}$$

    Notice that the last one is a *complete description*, namely the *outcome* HT.
  - $A \cup B = \{HH, HT, TH\}$, or in words:

    $$\begin{aligned} &\text{“first was heads” OR “exactly one heads”} \\ &= \text{“starts with heads, else it's tails-then-heads”} \end{aligned}$$

Exercise - Coin flipping: counting subsets

Flip a fair coin five times and record the results.

How many elements are in the sample space? (How big is S?)
How many events are there? (How big is $\mathcal{F}$?)

Solution

There are $2^5 = 32$ possible sequences, so $|S| = 32$.

To count the number of possible subsets, consider that we have 32 distinct items, and a subset is uniquely determined by the binary information – for each item – of whether it is in or out. Thus there are 2^{32} possibilities. So $|\mathcal{F}| = 2^{32}$.

03 Theory

New events from old

Given two events A and B, we can form new events using set operations:

$$A \cup B \quad \longleftrightarrow \quad \text{“event } A \text{ OR event } B\text{”}$$

$$A \cap B \quad \longleftrightarrow \quad \text{“event } A \text{ AND event } B\text{”}$$

$$A^c \quad \longleftrightarrow \quad \textbf{not} \text{ event } A$$

We also use these terms for events A and B:

- They are **mutually exclusive** when $A \cap B = \emptyset$, that is, they have *no elements in common*.
- They are **collectively exhaustive** $A \cup B = S$, that is, when they jointly *cover all possible outcomes*.

- In probability texts, sometimes $A \cap B$ is written “$A \cdot B$” or even (frequently!) “AB”.

Rules for sets

Algebraic rules

- Associativity: $(A \cup B) \cup C = A \cup (B \cup C)$. Analogous to $(A + B) + C = A + (B + C)$.
- Distributivity: $A \cap (B \cup C) = (A \cap B) \ \cup \ (A \cap C)$. Analogous to $A(B + C) = AB + AC$.

De Morgan’s Laws

- $(A \cup B)^c = A^c \cap B^c$
- $(A \cap B)^c = A^c \cup B^c$
  In other words: you can distribute “ ^{c} ” but must simultaneously do a switch $\cap \leftrightarrow \cup$.

Probability models

04 Theory

Axioms of probability

A **probability measure** is a function $P : \mathcal{F} \to \mathbb{R}$ satisfying:

Kolmogorov Axioms:

- **Axiom 1:** $P[A] \geq 0$ for every event A
  (probabilities are not negative!)
- **Axiom 2:** $P[S] = 1$
  (probability of “anything” happening is 1)

- **Axiom 3:** additivity for any *countable collection* of *mutually exclusive* events:

$$P[A_1 \cup A_2 \cup A_3 \cup \cdots] = P[A_1] + P[A_2] + P[A_3] + \cdots$$

$$\text{when:} \quad A_i \cap A_j = \emptyset \quad \text{for all } \; i \neq j$$

- Notation: we write $P[A]$ instead of $P(A)$, even though P is a function, to emphasize the fact that A is a set.

Probability model

A **probability model** or **probability space** consists of a triple $(S, \mathcal{F}, P)$:

- S the sample space
- $\mathcal{F}$ the set of valid events, where every $A \in \mathcal{F}$ satisfies $A \subset S$
- $P : \mathcal{F} \to \mathbb{R}$ a probability measure satisfying the Kolmogorov Axioms

Finitely many exclusive events

It is a consequence of the Kolmogorov Axioms that additivity also works for finite collections of events:

$$P[A \cup B] = P[A] + P[B]$$

$$P[A_1 \cup \cdots \cup A_n] = P[A_1] + \cdots + P[A_n]$$

Inferences from Kolmogorov

A probability measure satisfies these rules.
They can be deduced from the Kolmogorov Axioms.

- **Negation:** Can you find $P[A^c]$ but not $P[A]$? Use negation:

$$P[A] = 1 - P[A^c]$$

- **Monotonicity:** Probabilities grow when outcomes are added:

$$A \subset B \quad \gg\!\gg \quad P[A] \leq P[B]$$

- **Inclusion-Exclusion:** A trick for resolving unions:

$$P[A \cup B] = P[A] + P[B] - P[A \cap B]$$

(even when A and B are *not exclusive!*)

Inclusion-Exclusion

The principle of inclusion-exclusion generalizes to three events:

$$P[A \cup B \cup C] =$$

$$P[A] + P[B] + P[C] - P[A \cap B] - P[A \cap C] - P[B \cap C] + P[A \cap B \cap C]$$

The same pattern works for any number of events!

The pattern goes: “include singles” then “exclude doubles” then “include triples” then ...

Include, exclude, include, exclude, include, ...

05 Illustration

Example - Lucia is Host or Player

Problem: The professor chooses three students at random for a game in a class of 40, one to be Host, one to be Player, one to be Judge. What is the probability that Lucia is either Host or Player?

Solution:

1. Set up the probability model.
    - Label the students 1 to 40. Write L for Lucia's number.
    - *Outcomes:* assignments such as $(H, P, J) = (2, 5, 8)$
      These are ordered triples with *distinct* entries in 1, 2, $\ldots, 40$.
    - *Sample space:* S is the collection of all such distinct triples
    - *Events:* any subset of S
    - *Probability measure*: assume all outcomes are equally likely, so $P[(i, j, k)] = P[(r, l, p)]$ for all i, j, k, r, l, p
    - In total there are $40 \cdot 39 \cdot 38$ triples of distinct numbers.
    - Therefore $P[(i, j, k)] = \frac{1}{40 \cdot 39 \cdot 38}$ for any *specific* outcome (i, j, k).
    - Therefore $P[A] = \frac{|A|}{40 \cdot 39 \cdot 38}$ for any event A. (Recall $|A|$ is the *number* of outcomes in A.)
2. Define the desired event.
    - Want to find $P[$“Lucia is Host or Player”$]$
    - Define $A =$ “Lucia is Host” and $B =$ “Lucia is Player”. Thus:

    $$A = \{(L, j, k) \mid \text{any } j, k\}, \qquad B = \{(i, L, k) \mid \text{any } i, k\}$$

    - So we seek $P[A \cup B]$.
3. Compute the desired probability.
    - Importantly, $A \cap B = \emptyset$ (mutually exclusive).
      There are no outcomes in S in which Lucia is *both* Host and Player.
    - By *additivity*, we infer $P[A \cup B] = P[A] + P[B]$.
    - Now compute $P[A]$.
        - There are $39 \cdot 38$ ways to choose j and k from the students besides Lucia.
        - Therefore $|A| = 39 \cdot 38$.

- Therefore:

$$P[A] \quad \gg\gg \quad \frac{|A|}{40 \cdot 39 \cdot 38} \quad \gg\gg \quad \frac{39 \cdot 38}{40 \cdot 39 \cdot 38} \quad \gg\gg \quad \frac{1}{40}$$

- Now compute $P[B]$. It is similar: $P[B] = \frac{1}{40}$.
- Finally compute that $P[A] + P[B] = \frac{1}{20}$, so the answer is:

$$P[A \cup B] \quad \gg\gg \quad P[A] + P[B] \quad \gg\gg \quad \frac{1}{20}$$

Example - iPhones and iPads

Problem:
At Mr. Jefferson's University, 25% of students have an iPhone, 30% have an iPad, and 60% have neither.

What is the probability that a randomly chosen student has either iProduct? (Q1)

What about both? (Q2)

Solution:

1. Set up the probability model.
    - A student is chosen at random: an *outcome* is the chosen student.
    - *Sample space* S is the set of all students.
    - Write O = "has iPhone" and A = "has iPad" concerning the chosen student.
    - All students are equally likely to be chosen: therefore $P[E] = \frac{|E|}{|S|}$ for any event E.
    - Therefore $P[O] = 0.25$ and $P[A] = 0.30$.
    - Furthermore, $P[O^c A^c] = 0.60$. This means 60% have "not iPhone AND not iPad".
2. Define the desired event.
    - Q1: desired event $= O \cup A$
    - Q2: desired event $= OA$
3. Compute the probabilities.
    - We do not believe O and A are exclusive.
    - Try: apply inclusion-exclusion:

    $$P[O \cup A] = P[O] + P[A] - P[OA]$$

    - We know $P[O] = 0.25$ and $P[A] = 0.30$. So this formula, with given data, RELATES Q1 and Q2.
    - Notice the complements in $O^c A^c$ and try *Negation*.
    - *Negation:*

    $$P[(OA)^c] = 1 - P[OA]$$

    DOESN'T HELP.

- Try again: Negation:*

$$P[(O^c A^c)^c] = 1 - P[O^c A^c]$$

- And De Morgan (or a Venn diagram!):

$$(O^c A^c)^c \quad \gg\gg \quad O \cup A$$

- Therefore:

$$P[O \cup A] \quad \gg\gg \quad P[(O^c A^c)^c]$$

$$\gg\gg \quad 1 - P[O^c A^c] \quad \gg\gg \quad 1 - 0.6 = 0.4$$

- We have found Q1: $P[O \cup A] = 0.40$.
- Applying the RELATION from inclusion-exclusion, we get Q2:

$$P[O \cup A] = P[O] + P[A] - P[OA]$$

$$\gg\gg \quad 0.40 = 0.25 + 0.30 - P[OA]$$

$$\gg\gg \quad P[OA] = 0.15$$