Lecture 2: Set theory and probability spaces

Statistics 251

Yi Sun and Zhongjian Wang

Department of Statistics
The University of Chicago

Where are we?

Sample Spaces

The set S of possible outcomes of an experiment is known as the **sample space** of the experiment.

- If the experiment consists of flipping a single coin, then

$$S = \{H, T\}.$$

- If the experiment consists of flipping two coins, then the sample space consists of the following four points:

$$S = \{(H, H), (H, T), (T, H), (T, T)\}.$$

- If the experiment consists of measuring (in hours) the lifetime of a computer part, then S consists of all nonnegative real numbers; that is,

$$S = \{x : 0 \leq x \leq \infty\}$$

Events

Any subset E of the sample space is known as an **event**. In other words, an event is a subset of possible outcomes of the experiment.

When flipping a coin, $S = \{H, T\}$.

- If $E = \{H\}$, then E is the event that the coin lands heads up.
- If $F = \{T\}$, then F is the event that the coin lands tails up.

Events

When flipping two coins, $S = \{(H, H), (H, T), (T, H), (T, T)\}$:

- $E = (H, H), (H, T)$ is the event that the first coin is heads.

What is the event that there is at least one head?

Where are we?

Sample Spaces and Events

Relations and Operations

Venn diagrams and De Morgan's laws

Axioms of Probability

Union

When flipping two coins

$$S = \{(H, H), (H, T), (T, H), (T, T)\}.$$

Let $E = \{(H, H), (H, T)\}$ be the event that the 1st coin is heads and $F = \{(T, H), (H, H)\}$ be the event that the 2nd coin is heads.

The **union** $E \cup F$ consists of all outcomes in either E or F. What is $E \cup F$ explicitly here?

Intersection

When flipping two coins

$$S = \{(H, H), (H, T), (T, H), (T, T)\}.$$

Let $E = \{(H, H), (H, T)\}$ be the event that the 1st coin is heads and $F = \{(T, H), (H, H)\}$ be the event that the 2nd coin is heads.

The **intersection** $E \cap F$ consists of all outcomes in both E and F. What is $E \cap F$ explicitly here?

Unions and intersections of more than two events

If E_1, E_2,... are events, then the union of these events, denoted by

$$\bigcup_{n=1}^{\infty} E_n$$

is the event consisting of all outcomes in **at least one** of the E_n.

The intersection of the events E_n, denoted by

$$\bigcap_{n=1}^{\infty} E_n,$$

consists of outcomes which are in **all** of the events E_n.

Mutually exclusive events

When flipping two coins, let E be the event the 1st coin is heads, and F the event the first coin is tails. What is $E \cap F$?

Mutually exclusive events

When flipping two coins, let E be the event the 1st coin is heads, and F the event the first coin is tails. What is $E \cap F$?

The event $E \cap F$ does not contain any outcomes and cannot occur. We call this the null event, denoted $\emptyset$.
If $E \cap F = \emptyset$, then E and F are said to be **mutually exclusive.**

Complement

When flipping two coins, let E be the event the 1st coin is heads, and F the event the first coin is tails. What is $E \cup F$?

Complement

When flipping two coins, let E be the event the 1st coin is heads, and F the event the first coin is tails. What is $E \cup F$?

These events satisfy $E \cup F = S$ and $E \cap F = \emptyset$. In this case, we say that F is the **complement** of E, denoted $E = F^c$. Note $S^c = \emptyset$.

Where are we?

Venn diagrams

A Venn diagram is a graphical representation used to illustrate relations between events.

- Union, $E \cup F$
- Intersection, $E \cap F$
- Complement, E^c

De Morgan's Laws

The intersection and union operations satisfy some basic rules:

- Commutativity: $E \cup F = F \cup E$, $E \cap F = F \cap E$.
- Associativity: $(E \cup F) \cup G = E \cup (F \cup G)$, $(E \cap F) \cap G = E \cap (F \cap G)$.
- Distributive principle: $(E \cup F) \cap G = (E \cap G) \cup (F \cap G)$, $(E \cap F) \cup G = (E \cup G) \cap (F \cup G)$.

DeMorgan's Laws

We can relate the complement operation to union and intersection:

$$\left(\bigcup_{i=1}^{n} E_i\right)^c = \bigcap_{i=1}^{n} E_i^c.$$

Another analogue is:

$$\left(\bigcap_{i=1}^{n} E_i\right)^c = \bigcup_{i=1}^{n} E_i^c.$$

Where are we?

What is probability? (frequentist version)

For a sample space S and an event $E \subset S$, define $\#(E, n)$ to be the number of times in the first n **repetitions** of the experiment that the event E occurs. Then the probability $\mathbb{P}(E)$ of the event E is defined as

$$\mathbb{P}(E) = \lim_{n \to \infty} \frac{\#(E, n)}{n}.$$

Note: We assume throughout this course that the probability of any event always exists. In other words, the limit always converges!

What is probability? (Bayesian version)

What if the experiment can only be run once?

For a sample space S and an event $E \subset S$, the probability $\mathbb{P}(E)$ is a number in $[0, 1]$ measuring the **subjective belief** or **reasonable expectation** of the chance that E will occur.

What is probability? (Bayesian version)

What if the experiment can only be run once?

For a sample space S and an event $E \subset S$, the probability $\mathbb{P}(E)$ is a number in $[0, 1]$ measuring the **subjective belief** or **reasonable expectation** of the chance that E will occur.

Sounds good, but what does that mean precisely...?

What is probability? (axiomatic version)

For a sample space S, a probability space on S assigns each event $E \subset S$ a number $\mathbb{P}(E)$ satisfying the following axioms:

1. $0 \leq \mathbb{P}(E) \leq 1$
2. $\mathbb{P}(S) = 1$
3. For any sequence of mutually exclusive events E_1, E_2, ... (that is, events for which $E_i \cap E_j = \emptyset$ when $i \neq j$), we have

$$\mathbb{P}\Big(\bigcup_{i=1}^{\infty} E_i\Big) = \sum_{i=1}^{\infty} \mathbb{P}(E_i).$$

Corollaries of the axioms

What is $\mathbb{P}(\emptyset)$?

Show that $\mathbb{P}(E) \leq 1$.

Corollaries of the axioms

If $E \subset F$, then $\mathbb{P}(E) \leq \mathbb{P}(F)$.

For any E, $\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$.

Example: Tossing a coin

If our experiment consists of tossing a fair coin, then we would have

$$\mathbb{P}(\{H\}) = \mathbb{P}(\{T\}) = \frac{1}{2}.$$

On the other hand, if we believe the coin was manufactured to be biased so that a head is twice as likely as a tail, then we would have

$$\mathbb{P}(\{H\}) = \frac{2}{3}, \mathbb{P}(\{T\}) = \frac{1}{3}.$$