Lecture 2: Introduction to Probability

- Interpretations of Probability
- Set Theory
- The Definition of Probability

Learning Objectives

By the end of class, you would

- Be familiar with the common definitions and interpretations of probability,
- Be able to define events based on simple experiments,
- Know how to verify axioms of set theory, and
- Be able to calculate probabilities of events based on simple experiments.

Interpretations of Probability

– Probability quantifies the uncertainty about the occurrence of an event.

– **If I toss a fair coin once, how likely am I to see one head? What if I toss it twice? What if I used a biased coin?** We will revisit these questions once we have a better understanding of how to assign probabilities.

– In the simplest terms, probability lets us assign numbers (**which need to satisfy some properties we haven't talked about**) to events such that we can make statements about likely the events are.

– How should we develop a formal definition of probability? First, lets review three frequently used interpretations of probability:

- The frequency interpretation
- The classical interpretation
- The subjective interpretation

The Frequency Interpretation of Probability

– Suppose you are interested in an event A. Then the probability of A can be thought of as the proportion of times A occurs in an infinite sequence (or very long run) of separate, independent and identical tries. That is,

$$P(A) = \lim_{n\to\infty} \frac{\text{\# times A happens}}{n}$$

– Thus for a fair coin for example, if A is the probability of heads/tails and we toss the coin n times with n large enough (say, $\geq$ 10,000), we should expect

$$P(A) = \lim_{n\to\infty} \frac{\text{\# heads}}{n} \approx 0.5$$

– **How feasible is this?**

The Classical Interpretation of Probability

– The classical interpretation on the other hand is based on the concept of equally likely outcomes. If the outcomes of an experiment must be one of k different outcomes and all k outcomes are equally likely to occur, then the probability of each outcome is simply $1/k$.

– The textbook discusses the two basic difficulties of developing a formal definition of probability from this interpretation:

- First, the concept of equally likely outcomes is essentially based on the concept of probability, which is exactly what we are trying to define. If two outcomes are equally likely, they must have the same probability.
- Second, there is no provision for defining probability for outcomes that are not equally likely.

The Subjective Interpretation of Probability

– The subjective interpretation relies on the notion that the probability of any event should reflect a person's belief about the likelihood of occurrence.

– Let X denote the temperature at noon tomorrow outside the Social Sciences building at Duke. Pick a number x_1 such that you consider these two outcomes to be equally likely: A = $\{X \leq x_1\}$ and B = $\{X > x_1\}$. Now pick numbers x_2 and x_3 such that the following outcome are equally likely: A = $\{X \leq x_2\}$, B = $\{x_2 < X < x_3\}$ and C = $\{X \geq x_3\}$. **Discuss your value of x_1 with the person seated closest to you. Can both of you agree on the same value? (1 min)**

– Note: we will not spend any more time discussing the pros and cons of each interpretation, but rather we will see as we move on, how each one can be useful depending on the experiment of interest. Interestingly, the theory of probability doesn't actually depend on any particular interpretation.

Experiments and Events

– An **experiment** is any process such that the collection of every possible outcome of the process can be described and "perhaps" even enumerated.

– An **event** is any subset of the **sample space** – the collection of all possible outcomes of an experiment.

– For example, toss a coin twice (the experiment), then the sample space $S = \{HH, HT, TH, TT\}$ and $A = \{HT, TH\}$ is an event: observing only one head in the double coin toss.

Set Relations

– Now for some relations:

- S denotes the sample space.
- $x \in A$ denotes that an element x is a **member** of A,
- $A \subset B$ denotes that a set A is a **subset** of set B, so that A and B are **equal** if and only if $A \subset B$ and $B \subset A$.
- If $A \subset B$ and $B \subset C$, then $A \subset C$.
- $\emptyset$ denotes a set that contains no element; the **empty set**. Thus, $\emptyset \subset A$ for any A.

Set Operations

Let $A, B \subset S$

- **Complement**: $x \in A \Rightarrow x \notin A^c$.
- **Union**: $x \in A \cup B \Rightarrow x \in A$ or $x \in B$ or both
- **Intersection**: $x \in A \cap B \Rightarrow x \in A$ and $x \in B$. Thus, $(A \cap B) \subset (A \cup B)$. A and B are **disjoint** if $(A \cap B) = \emptyset$.
- **Set Difference**: $x \in A \backslash B \Rightarrow x \in A$ and $x \notin B$. That is, $A \backslash B = A \cap B^c$.
- **Symmetric difference**: $x \in A \triangle B \Rightarrow x \in (A \backslash B) \cup (B \backslash A)$.
- **De Morgan's Laws**: $(A \cup B)^c = A^c \cap B^c$ and $(A \cap B)^c = A^c \cup B^c$.
- **Cardinality**: $|A|$ is the number of elements in A.

Examples

Example 1: Toss a fair coin twice and let A be the event that you observe only one head and C be the event that you observe at least one head. Then,

- The sample space $S = \{HH, HT, TH, TT\}$
- $A = \{HT, TH\}$
- $|A| = 2$
- $A^c = \{HH, TT\}$
- $\{HT\}$ and $\{TH\}$ are subsets of A.
- $A \cup A^c = \{HH, HT, TH, TT\} = S$
- $C = \{HH, HT, TH\}$. Then $A \subset C$, $A \cup C = C$, $A \cap C = A$, $A \backslash C = \emptyset$, $C \backslash A = \{HH\} = A \triangle C$.

Disjoint and Non-disjoint Outcomes

Disjoint (mutually exclusive) outcomes: Cannot happen at the same time.

- The outcome of a single coin toss cannot be a head and a tail.
- A student both cannot fail and pass a class.
- A single card drawn from a deck cannot be an ace and a queen.

Non-disjoint outcomes: Can happen at the same time.

- A student can get an A in Stats and A in Econ in the same semester.

Definition of Probability

Let $\mathcal{F}$ be a well-defined collection of subsets of S. A probability measure (or simply a probability) P is a function $P : \mathcal{F} \to [0, 1]$ such that

- For every event $A \in \mathcal{F}$, $P(A) \geq 0$.
- $P(S) = 1$
- For any sequence of disjoint events $A_1, A_2, \ldots \in \mathcal{F}$,

$$P\left(\bigcup_i A_i\right) = \sum_i P(A_i)$$

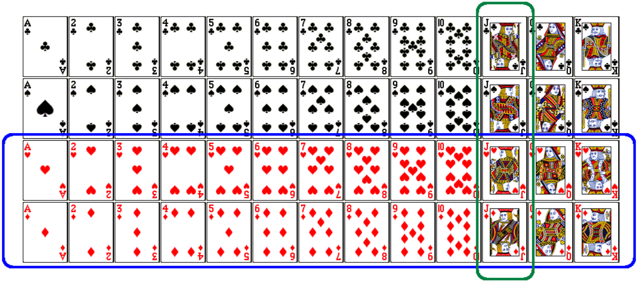
Some Properties of Probability

- $P(\emptyset) = 0$
- $P(A) = 1 - P(A^c)$
- If $A \subset B$, then $P(A) \leq P(B)$
- For any event A, $0 \leq P(A) \leq 1$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- For any sequence of events,

$$P\left(\bigcup_i A_i\right) \leq \sum_i P(A_i) \quad \text{and} \quad P\left(\bigcap_i A_i\right) \geq 1 - \sum_i P(A_i^c)$$

Union of Non-disjoint Events

What is the probability of drawing a jack or a red card from a well shuffled full deck?



$$P(jack\ or\ red) = P(jack) + P(red) - P(jack\ and\ red)$$
$$= \frac{4}{52} + \frac{26}{52} - \frac{2}{52} = \frac{28}{52}$$

Figure from `http://www.milefoot.com/math/discrete/counting/cardfreq.htm`.

Examples

Example 1 again: Since this is a fair coin, all outcomes are equally likely.

- $P(S) = P\{HH\} + P\{HT\} + P\{TH\} + P\{TT\} = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 1$
- $P(A) = P\{HT\} + P\{TH\} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$
- $P(A^c) = 1 - P(A) = \frac{1}{2} = P\{HH\} + P\{TT\}$
- $P(C) = \frac{3}{4}$
- $P(A \cup C) = \frac{3}{4}$
- $P(A \cap C) = \frac{1}{2}$
- $P(A \backslash C) = 0$
- $P(C \backslash A) = P\{HH\} = \frac{1}{4} = P(A \triangle C)$

Independence

Two processes are **independent** if knowing the outcome of one provides no useful information about the outcome of the other.

- Knowing that the coin landed on a head on the first toss does not provide any useful information for determining what the coin will land on in the second toss. $\rightarrow$ Outcomes of two tosses of a coin are independent.
- Knowing that the first card drawn from a deck is an ace does provide useful information for determining the probability of drawing an ace in the second draw. $\rightarrow$ Outcomes of two draws from a deck of cards (without replacement) are dependent.

Product Rule for Independent Events

Product rule for independent events:

$$P(A \text{ and } B) = P(A) \times P(B)$$

Or more generally, $P(A_1 \text{ and } \cdots \text{ and } A_k) = P(A_1) \times \cdots \times P(A_k)$
You toss a coin twice, what is the probability of getting two tails in a row?

$$P(\text{T on the first toss}) \times P(\text{T on the second toss}) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

Examples

Example 1 yet again: Suppose the coin is biased, with $P(H) = \frac{3}{5}$, can you recalculate the same probabilities?

- $P\{HH\} = \frac{9}{25}$ $P\{HT\} = \frac{6}{25}$ $P\{TH\} = \frac{6}{25}$ $P\{TT\} = \frac{4}{25}$
- $P(A) = P\{HT\} + P\{TH\} = \frac{12}{25}$
- $P(A^c) = 1 - P(A) = \frac{13}{25}$
- $P(C) = P\{HH\} + P\{HT\} + P\{TH\} = \frac{21}{25}$
- $P(A \cup C) = P(C) = \frac{21}{25}$, $P(A \cap C) = P(A) = \frac{12}{25}$
- $P(A \backslash C) = P(\emptyset) = 0$, $P(C \backslash A) = P\{HH\} = \frac{9}{25}$
- $P(A \triangle C) = P\{HH\} = \frac{9}{25}$

In-class Exercise

Suppose you roll a fair die twice.

- Denote the sample space of roll a fair die twice as S, what is $|S|$?

$$\begin{aligned} S = \{&(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),\\ &(2,1),(2,2),(2,3),(2,4),(2,5),(2,6),\\ &(3,1),(3,2),(3,3),(3,4),(3,5),(3,6),\\ &(4,1),(4,2),(4,3),(4,4),(4,5),(4,6),\\ &(5,1),(5,2),(5,3),(5,4),(5,5),(5,6),\\ &(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\}, \end{aligned}$$

so $|S| = 36$

- What is the probability of observing an even number for each roll? In each roll, we observe one number from $\{1,2,3,4,5,6\}$, so the probability of seeing an even number for each roll is $1/2$.
- What is the probability of observing an odd number for each roll?
  In each roll, if we do not observe an even number, then we must observe an odd number, so the probability of seeing an odd number for each roll is 1-0.5=1/2.

In-class Exercise (cont'd)

- What is the probability of observing the same number for both rolls?
  No matter which number of the six we saw in the first roll, the probability of getting exactly the same number is $1/6$.
- What is the probability of observing different numbers for both rolls? No matter which number of the six we saw in the first roll, the probability of getting a different number is $5/6$.
- What is the probability of observing numbers whose sum is at most 7? Denote the event of interest as A, which contains all the blue colored outcomes below

$$\begin{aligned} S = \{&(\mathbf{1},\mathbf{1}), (\mathbf{1},\mathbf{2}), (\mathbf{1},\mathbf{3}), (\mathbf{1},\mathbf{4}), (\mathbf{1},\mathbf{5}), (\mathbf{1},\mathbf{6}),\\ &(\mathbf{2},\mathbf{1}), (\mathbf{2},\mathbf{2}), (\mathbf{2},\mathbf{3}), (\mathbf{2},\mathbf{4}), (\mathbf{2},\mathbf{5}), (2,6),\\ &(\mathbf{3},\mathbf{1}), (\mathbf{3},\mathbf{2}), (\mathbf{3},\mathbf{3}), (\mathbf{3},\mathbf{4}), (3,5), (3,6),\\ &(\mathbf{4},\mathbf{1}), (\mathbf{4},\mathbf{2}), (\mathbf{4},\mathbf{3}), (4,4), (4,5), (4,6),\\ &(\mathbf{5},\mathbf{1}), (\mathbf{5},\mathbf{2}), (5,3), (5,4), (5,5), (5,6),\\ &(\mathbf{6},\mathbf{1}), (6,2), (6,3), (6,4), (6,5), (6,6)\}, \end{aligned}$$

  Each of the outcome is equally likely to occur with probability $1/36$, so the probability of the event A is $P(A) = \frac{21}{36} = \frac{7}{12}$.

In-class Exercise (cont'd)

- What is the probability of observing numbers whose sum is at least 4? Denote the event "observing numbers whose sum is at least 4" as B, and its complement "observing numbers whose sum is less than 4" as B^c. B^c contains all the blue colored outcomes below

$$
\begin{aligned}
S = \{&(\mathbf{1},\mathbf{1}), (\mathbf{1},\mathbf{2}), (1,3), (1,4), (1,5), (1,6) \\
&(\mathbf{2},\mathbf{1}), (2,2), (2,3), (2,4), (2,5), (2,6), \\
&(3,1), (3,2), (3,3), (3,4), (3,5), (3,6), \\
&(4,1), (4,2), (4,3), (4,4), (4,5), (4,6), \\
&(5,1), (5,2), (5,3), (5,4), (5,5), (5,6), \\
&(6,1), (6,2), (6,3), (6,4), (6,5), (6,6)\},
\end{aligned}
$$

Each of the outcome is equally likely to occur with probability $1/36$, so the probability of the event B^c is $P(B^c) = \frac{3}{36} = \frac{1}{12}$, and $P(B) = 1 - P(B^c) = \frac{11}{12}$.

Some Questions to Think About

- Probability zero does not mean impossible
- Probability one event has to happen?

Disjoint v.s. Complementary

- Q: Do the sum of probabilities of two disjoint events always add up to 1?
- A: Not necessarily, there may be more than 2 events in the sample space, e.g. party affiliation.
- Q: Do the sum of probabilities of two complementary events always add up to 1?
- A: Yes, that's the definition of complementary, e.g. heads and tails.

Disjoint v.s. Independence

- Disjoint (mutually exclusive) outcomes cannot happen at the same time
- If A and B are independent events, having information on A doesn't tell us anything about B (and vice versa)

Recap

Today, we talked about

- Common interpretations of probability
- Set theory and basic set operations
- How to calculate probability in simple experiments

Suggested reading:

- D.S. Sec. 1.2, 1.3, 1.4, and 1.5
- OpenIntro3: Sec. 2.1