Probability basics

Part 1

9/30/24

Housekeeping

Probability basics

We spend a whole semester on this in MATH/STAT 310!

Key terms

- **Random process**: a situation in which a particular result, called an **outcome**, is random/not known ahead of time
  - Examples: flipping a coin, rolling six-sided die, sports game, if a treatment is effective
- A **sample space** S is the set of all possible outcomes of the random process
  - What are possible sample spaces for the above examples?
- An **event** is a set of outcomes from a random process

Random variable

- A **random variable** is a variable whose value is unknown and depends on random events
  - Often denoted with a capital letter like X or Y
- There are two types: discrete and continuous (just like in numeric variables)
  - **Discrete**: represents random process where sample space is countable (i.e. finite, or distinct counts)
  - **Continuous**: sample space is uncountable (i.e. can take on any value within a specified interval with infinite number of possible values)
- **NOTE**: we will focus on *discrete* random variables for now

Probability

- For us, the **probability** of an outcome is the proportion of times the outcome would occur if we observed the random process an infinite number of times
  - Probability is used to express the likelihood that some outcome or event will or will not occur
  - Think of as a proportion
- Let A denote some outcome or event. We denote the probability of A occurring as $\mathrm{P}(A)$ or $\Pr(A)$.
- When the sample space S is discrete with a finite size, then $\Pr(A) = \frac{\text{number of outcomes favorable to } A}{\text{number of total outcomes possible}}$

Example

Let the random process rolling a fair, six-sided die. Let X a random variable representing the value of the die.

- For each of the following, determine the outcome(s) and event under consideration, along with the value of the probability itself:
  - $\Pr(X = 1)$
  - $\Pr(X = 1 \text{ and } 2)$
  - $\Pr(X \text{ is even})$

Operations with events

Let A and B be two possible events.

- The **intersection** of A and B is denoted as $A \cap B$, and is the set of outcomes that belong to *both* events A and B
- The **union** of A and B is denoted as $A \cup B$, and is the set of outcomes that belong to A and/or B

When we have only two or three events, Venn diagrams can be very useful for visualizing probabilities!

Disjoint events

Two events are **disjoint** or **mutually exclusive** if they cannot simultaneously happen.

- That is, if A and B are disjoint, then $\Pr(A \cap B) = ?$
- If our random process is rolling a six-sided die one time, what are some examples of disjoint events?

Rules of probability

Kolmogorov axioms

1. The probability of any event is non-negative real number
2. The probability of the entire sample space 1
3. If A and B are disjoint, then $\Pr(A \cup B) = \Pr(A) + \Pr(B)$

These axioms imply that all probabilities are between 0 and 1 inclusive, and lead to some important rules!

Addition rule

Let A and B be two possible events. Then the **addition rule** states that the probability that at least one will occur is:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- Venn diagram
- Example: in a standard deck of 52 cards, we have four suits (diamond, heart, club, spade) with 13 cards within each suit (1-10, Jack, Queen, King).
  - Suppose we randomly draw one card from the shuffled deck.
  - Let A be the event that the card is a spade.
  - Let B be the event that the card is a face card (Jack, Queen or King).
  - Find $P(A \cup B)$.

Complement

- The **complement** of an event A is the set of all outcomes in S that are not in A
  - Denoted as A^c
- Continuing the dice example, if A is the event that a 1 or 2 is rolled, what is A^c?
- **Complement rule**: $\Pr(A^c) = 1 - \Pr(A)$
- Let our random process be the sum of two dice. What is the probability that…
  - the sum of the dice is not 6?
  - the sum is *at least* 4?

DeMorgan's Laws

Let's use Venn diagrams to try and determine formulas for the following:

1. Complement of union: $(A \cup B)^c \; = \; ?$
2. Complement of intersection: $(A \cap B)^c \; = \; ?$

Independence

- Qualitatively, two processes are **independent** if knowing the outcome of one does not provide any information about the outcome of the other process
  - Examples and non-examples? How to formalize this?
- If A and B are independent events from two different and independent processes, then $\Pr(A \cap B) = \Pr(A) \times \Pr(B)$
- More generally, if $\Pr(A \cap B) = \Pr(A) \times \Pr(B)$ and A and B are events from the same process, then A and B are independent.
  - This is known as the **multiplication rule for independent events**

Probability distributions

When a random variable is discrete, it can be useful to discuss its **probability distribution**, which is a table of all (disjoint) outcomes and their associated probabilities.

- Let X be the sum of two fair, six-sided dice. What is the sample space S?
- Fill out the table below to display the probability distribution of X:

X	2	3	4	5	6	7
Probability						
X	8	9	10	11	12	
Probability						

Probability distributions (cont.)

The probability distribution of a discrete random variable must satisfy the following three rules:

1. The outcomes listed must be disjoint
2. Each probability must be between 0 and 1 (inclusive)
3. The probabilities must sum to 1

Let's confirm that the distribution we found on the previous slide satisfies these rules!

Practice

A Pew Research survey asked 2,373 randomly sampled registered voters their political affiliation (Republican, Democrat, or Independent) and whether or not they identify as swing voters. 35% of respondents identified as Independent, 23% identified as swing voters, and 11% identified as both.

a. What percent of voters are Independent but not swing voters?

b. What percent of voters are Independent or swing voters?

c. What percent of voters are neither Independent nor swing voters?

d. Is the event that someone is a swing voter independent of the event that someone is a political Independent?