Conditional Probability
TFs: Justin Zhu (justinzhu@college)

Conditional Probability

Reading Right to Left - We read "|" right to left, that is, $A|B$ means B occurred first, then A

Conditional Probability - Suppose we observe event B and are interested in the probability of event A occurring given this information. Then,

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Bayes' Rule - This is arguably one of the most important concepts and tools you will learn in this course.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Bridging Conditional Probability and Sets

An intuitive way to visualize conditional probability is to think about the intersection of sets. In order to find the intersection of two different sets A and B, we establish one of these sets to be our sample space and find the likely occurrence of the other set within this established sample space.

$$P(A \cap B) = P(B|A)P(A) = P(A|B)P(B)$$

$$P(A_1 \cap A_2 \cap A_3 \cap \cdots A_n) = P(A_1|A_2 \cap A_3 \cap \cdots A_n)P(A_2|A_3 \cap \cdots A_n) \cdots P(A_{n-1}|A_n)P(A_n)$$

Law of Total Probability (LOTP)

A common theme in this course is that it is far easier to solve a problem by breaking it down into smaller, simpler components than tackling it head-on. LOTP is one such tool. Suppose you want to find the probability of some event B, and you can partition the sample space into disjoint events $A_1, A_2, \ldots, A_n$. Then,

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i) \tag{1}$$

$$= \sum_{i=1}^{n} P(B \cap A_i) \tag{2}$$

We often use LOTP with Bayes' rule! Specifically, the denominator of Bayes' rule, $P(B)$, is often difficult to calculate outright, so we will instead calculate it in terms of LOTP.

Extra Conditioning

Incorporating extra information C is a simple extension of Bayes' rule and LOTP:

$$P(A|B,C) = \frac{P(A \cap B|C)}{P(B|C)} \tag{3}$$

$$= \frac{P(B|A,C)P(A|C)}{P(B|C)} \tag{4}$$

$$P(B|C) = \sum_{i=1}^{n} P(B|A_i,C)P(A_i|C) \tag{5}$$

$$= \sum_{i=1}^{n} P(B \cap A_i|C) \tag{6}$$

Disjoint vs. Independent

Disjoint, or mutually exclusive, events are events that cannot occur simultaneously. That is, observing event A precludes the possibility of also observing event B. We can state this equivalently as

$$P(A \cap B) = 0$$

Independent events are events such that observing event B yields no information about the possibility of also observing event A. That is, conditioning on observing event B, the probability of observing event A is unchanged.

$$P(A|B) = P(A)$$

We can apply this result to Bayes' rule and quickly demonstrate an alternative definition of independence:

$$P(A \cap B) = P(A)P(B)$$

Another form of independence is conditional independence. Two events A and B are said to be conditionally independent given C if

$$P(A \cap B|C) = P(A|C)P(B|C)$$

However, just as pairwise independence does not imply independence (and vice versa), conditional independence does not imply independence (and vice versa).

Practice Problems and Brainteasers

Example 1. Recruiting.
Suppose the Harvard Consulting, Investment, and Tech Group™currently consists of two freshmen and some number of upperclassmen. A new student joins the group, but she forgot to indicate what year she was in! At the next club meeting, a recruiter from BainBookSachs™comes in and plucks a lucky student to join their ranks. Given that the student is a freshman, what is the probability that the student that just joined was a freshman? Suppose that freshmen and upperclassmen are equally likely to join HCITG.

Example 2. Russian Roulette.
In a game of Russian Roulette, you find a revolver with **six** chambers containing two real bullets **side-by-side** and four empty chambers. You spin the chamber and point the gun at yourself... Click. No bullet. It is your turn again, but do you want to spin the barrel again or just pull the trigger?