EE 520: Random Processes **Fall 2021**

Lecture 2
Conditional Probability and Combinatorics

Instructor Name: John Lipor

Recommended Reading: Pishro-Nik: 1.4 - 2.1; Gubner: 1.5 - 1.7

1 Conditional Probability

We're often interested in statements of the form, "If B occurs, then the probability of A is p," where A and B are events. This is known as *conditional probability* and is a foundational part of probability.

Definition 1. If $P(B) > 0$ for some event B, then the **conditional probability** that A occurs given B occurs is

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

This is referred to as the probability of A given B.

Example 1. Roll two fair dice. Given that the first shows a 3, what is the probability the total exceeds 6?

$$A = \{\text{total of dice exceeds } 6\} = \{(a, b) : a + b > 6\}$$
$$B = \{\text{roll a 3 on the first die}\} = \{(3, b) : 1 \le b \le 6\}$$

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{3/36}{6/36} = \frac{1}{2},$$

where the above probabilities come from counting the number of outcomes in each set (check for yourself).

Now assume $P(A) > 0$ and suppose we know $P(A \mid B)$. How can we find $P(B \mid A)$? Rearrange the definition of conditional probability to get

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

Now apply the definition of $P(B \mid A)$ to see that

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

This is useful if we know $P(A)$. When we cannot find $P(A)$ directly, we instead use a (very important) tool called *the law of total probability.*

Definition 2. A **partition** $\{B_1, B_2, \ldots, B_n\}$ of a set Ω is a collection of sets such that

- $B_i \cap B_j = \emptyset$, $\forall i \neq j$ (B_i's are pairwise disjoint)
- $\bigcup_{i=1}^{n} B_i = \Omega$ (B_i's cover the entire set Ω).

Definition 3. Let $\{B_1, B_2, \ldots, B_n\}$ be a partition of the set Ω in the probability space $(\Omega, \mathcal{F}, P)$. The **law of total probability** states that for any $\mathcal{F}$-measurable event A, we can write

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i)P(B_i).$$

Proof. Consider $n = 2$. We have

$$A = (A \cap B) \cup (A \cap B^{\mathsf{c}}).$$

Since B and B^{c} are disjoint, the axioms of probability tell us that

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \cap B^{\mathsf{c}}) \\ &= P(A \mid B)P(B) + P(A \mid B^{\mathsf{c}})P(B^{\mathsf{c}}). \end{aligned}$$

The proof of the general case can be found in Section 1.4.2 of Pishro-Nik. □

Now that we can compute $P(A)$, we can go back to our original problem of finding $P(B \mid A)$. This process is so common that it gets its own definition.

Definition 4. Let $\{B_1, B_2, \dots, B_n\}$ form a partition of Ω. Then **Bayes' rule** states that

$$P(B_i \mid A) = \frac{P(A \mid B_i)P(B_i)}{\sum_{i=1}^{n} P(A \mid B_i)P(B_i)}.$$

For $n = 2$, we get

$$P(B \mid A) = \frac{P(A \mid B)P(B)}{P(A \mid B)P(B) + P(A \mid B^{\mathsf{c}})P(B^{\mathsf{c}})}.$$

In words, suppose we observe A but want to know about B_i. If we know the *prior probabilities* $\{P(B_i)\}_{i=1}^{n}$ and the conditional probabilities $\{P(A \mid B_i)\}_{i=1}^{n}$, we can compute the *posterior probability* $P(B_i \mid A)$ for each $i \in \{1, 2, \dots, n\}$.

Example 2. Assume you are a zoggle magnate and own two factories. Based on your experience, you know the defect rate at each factory

$$P(\text{defective zoggle} \mid \text{factory } 1) = 0.2$$
$$P(\text{defective zoggle} \mid \text{factory } 2) = 0.05.$$

Further, factory 1 produces twice as many zoggles as factory 2. What is the probability that a random zoggle is satisfactory? Let A be the event that a zoggle is satisfactory and B be the event that a zoggle is from factory 1 (what is B^{c}?). Then

$$P(A) = P(A \mid B)P(B) + P(A \mid B^{\mathsf{c}})P(B^{\mathsf{c}}) = 4/5 \times 2/3 + 19/20 \times 1/3 = 51/60.$$

Given the above, what is the *posterior* probability that a zoggle is from factory 1 given that it is defective?

$$P(B \mid A^{\mathsf{c}}) = \frac{P(A^{\mathsf{c}} \mid B)P(B)}{P(A^{\mathsf{c}})} = \frac{1/5 \times 2/3}{9/60} = \frac{8}{9}.$$

2 Independence

Conditional probability implies that the occurrence of some event B impacts the probability of another event A. When this is not the case, we say that these events are *independent*.

Definition 5. Events A and B are called **independent** if

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

More generally, the collection $\{A_i\}_{i \in I}$ is (mutually) independent if

$$P\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} P(A_i)$$

for any finite subset $J \subset I$.

Caution: There is a strong tendency to conflate disjoint sets with independent events. Recall that A and B are disjoint if they do not intersect, i.e., if $A \cap B = \emptyset$. This has nothing to do with the probability measure used, whereas independence is a function of the two sets and the probability measure.

Example 3. Roll a fair die. Let

$$\begin{aligned} A = \{1, 2\} \quad &\Longrightarrow \quad P(A) = 1/3 \\ B = \{2, 4, 6\} \quad &\Longrightarrow \quad P(B) = 1/2. \end{aligned}$$

Note that

$$P(A \cap B) = P(\{2\}) = 1/6 = P(A)P(B),$$

so we say that A and B are independent events. This example illustrates that independence does not necessarily follow our intuition—we sometimes need to follow the math.

Example 4. Choose a playing card at random. The suit is independent of the rank (which may match our intuition). For example

$$P(\text{king of spades}) = 1/52 = 1/4 \times 1/13 = P(\text{spade})P(\text{king}).$$

3 Combinatorics and Probability

Many problems of interest in probability are called *counting problems*, where we compute a probability by counting the cardinality of an event of interest and dividing it by the cardinality of the sample space. Formally, let E be some event of interest on a sample space Ω. Then

$$P(E) = \frac{|E|}{|\Omega|}.$$

The major difficulty in these problems is determining the cardinality of the event of interest. There are four types of counting problems, and we focus on three of them in this course.

Ordered sampling with replacement. In these experiments, the order of outcomes matters. Suppose you perform k experiments with $n_1, n_2, \ldots, n_k$ possible outcomes for each. Then there are $\prod_{i=1}^{k} n_i$ possible outcomes.

Example 5. Draw k cards with replacement from a deck of n cards. The number of possible sequences (note that sequences have an order, unlike sets) is n^k.

Ordered sampling without replacement. This is the case where order again matters, but once an object is removed, it cannot be drawn again. Given n objects, the number of ordered permutations of these objects is

$$n \times (n-1) \times (n-2) \times \cdots \times 3 \times 2 \times 1 = n!,$$

since for the first draw, we have n options, for the second we have $n-1$, and so on. If we draw k objects without replacement, we have

$$n \times (n-1) \times (n-2) \times \cdots \times (n-(k-1)) = \frac{n!}{(n-k)!}$$

possible outcomes.

Example 6. Draw 5 cards without replacement. The number of possible sequences is $52 \times 51 \times 50 \times 49 \times 48$, which is really large. Note that

$$\frac{52!}{(52-5)!} = \frac{52 \times 51 \times 50 \times 49 \times 48 \times 47 \times 46 \cdots \times 1}{47 \times 46 \times \cdots \times 1} = 52 \times 51 \times 50 \times 49 \times 48.$$

Unordered sampling without replacement. Given n objects, the number of different groups of k objects that could be formed (called *combinations*) is

$$\frac{n!}{(n-k)!\,k!} = \binom{n}{k},$$

which is called the *binomial coefficient* and is read as "n choose k." To see why this is, note that there are $n!/(n-k)!$ possible ordered sets of k. But each set has $k!$ possible permutations, so it is counted $k!$ times. Dividing by $k!$ normalizes for this double counting.

Example 7. Deal a hand of 5 cards. The number of possible hands is $\binom{52}{5}$.

From counting to probability. We can apply the above principles to determine the probabilities of events of interest. This may seem straightforward, but counting problems are often the most dreaded in this course. Skill with these is best obtained by working as many examples as possible.

Example 8. An urn contains 11 green balls and 9 red balls. If 12 balls are chosen at random, what is the probability of choosing exactly 5 green balls and 7 red balls?

First note that we are interested in sets of balls, so order does not matter, i.e., we are in the setting of unordered sampling without replacement. In this case, the sample space has size $\binom{20}{12}$. To determine the cardinality of E, we first note that the number of ways to choose 5 green balls is $\binom{11}{5}$, and similarly the number of ways to choose 7 red balls is $\binom{9}{7}$. To get the cardinality of E, we multiply these to see that

$$P(E) = \frac{|E|}{|\Omega|} = \frac{\binom{11}{5}\binom{9}{7}}{\binom{20}{12}}.$$