Math/Stat 251 Fall 2015
Summary of Lectures from September 9, 2015, through September 28, 2015

The goal of the first part of the course was to introduce the basic concepts of probability theory. Our approach was to consider probability as a way of modelling a chance experiment. The collection S of all possible outcomes is known as the *sample space*. An *event* is a subset of the sample space. Probabilities are assigned to events. Thus, a *probability* is a function $\mathbf{P} : \{\text{events}\} \to [0, 1]$ with the properties that

(a) $\mathbf{P}(\emptyset) = 0$, $\mathbf{P}(S) = 1$, and

(b) $\mathbf{P}(A \text{ or } B) = \mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$ whenever A and B are disjoint (i.e., mutually exclusive).

We then spent time learning various techniques to compute probabilities for certain events. Our results included the following.

Addition Rule

If A and B are any events, then

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B).$$

In fact, this is a special case of something called the *Inclusion-Exclusion Formula*. (See Problem #5 on Assignment #2.) That is, if A, B, and C are any events, then

$$\begin{aligned}&\mathbf{P}(A \cup B \cup C)\\ &\quad = \mathbf{P}(A) + \mathbf{P}(B) + \mathbf{P}(C) - \mathbf{P}(A \cap B) - \mathbf{P}(A \cap C) - \mathbf{P}(B \cap C) + \mathbf{P}(A \cap B \cap C).\end{aligned}$$

To verify this formula, think about drawing a Venn diagram with three overlapping circles. Note that to compute the probability of A or B or C you add together their probabilities. This accounts for the $\mathbf{P}(A) + \mathbf{P}(B) + \mathbf{P}(C)$ piece. However, you have double-counted all of their common pairwise intersections so you need to subtract them. This accounts for the $-\mathbf{P}(A \cap B) - \mathbf{P}(A \cap C) - \mathbf{P}(B \cap C)$ piece. But by doing so, you have removed their common intersection one too many times so you need to put $\mathbf{P}(A \cap B \cap C)$ back. In fact, this idea can be extended to more than 3 events. If $A_1, A_2, \ldots, A_n$ are any events, then

$$\begin{aligned}&\mathbf{P}(A_1 \cup \cdots \cup A_n)\\ &= \sum_{j=1}^{n} \mathbf{P}(A_j) - \sum_{i<j} \mathbf{P}(A_i \cap A_j) + \sum_{i<j<k} \mathbf{P}(A_i \cap A_j \cap A_k) + \cdots + (-1)^{n+1}\mathbf{P}(A_1 \cap \cdots \cap A_n).\end{aligned}$$

Definition of Conditional Probability

If A and B are any events with $\mathbf{P}(B) > 0$, then the *conditional probability* of A given B is defined as

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

Multiplication Rule

Note that the definition of conditional probability implies that we always have

$$\mathbf{P}(A \cap B) = \mathbf{P}(A \,|\, B)\,\mathbf{P}(B).$$

Sometimes this is called the *Multiplication Rule*.

Law of Total Probability

Suppose that $B_1, B_2, \ldots, B_k$ partition S. That is, $B_1 \cup B_2 \cup \cdots \cup B_k = S$ with $B_i \cap B_j = \emptyset$ for every $i \neq j$. If A is any event, then

$$\mathbf{P}(A) = \mathbf{P}(A \,|\, B_1)\,\mathbf{P}(B_1) + \cdots + \mathbf{P}(A \,|\, B_k)\,\mathbf{P}(B_k).$$

Bayes' Rule

By combining the definition of conditional independence and the Law of Total Probability, we arrive at Bayes' Rule which states that

$$\mathbf{P}(B_1 \,|A) = \frac{\mathbf{P}(A \,|\, B_1)\,\mathbf{P}(B_1)}{\mathbf{P}(A)} = \frac{\mathbf{P}(A \,|\, B_1)\,\mathbf{P}(B_1)}{\mathbf{P}(A \,|\, B_1)\,\mathbf{P}(B_1) + \cdots + \mathbf{P}(A \,|\, B_k)\,\mathbf{P}(B_k)}.$$

(See Problem #3 on Assignment #3 and Problems #1, #2, #3, #4 on Assignment #4.)

Definition of Independence

We say that the events A and B are independent if and only if $\mathbf{P}(A \cap B) = \mathbf{P}(A)\,\mathbf{P}(B)$. This is sometimes called the *Multiplication Rule for Independent Events*. Note that from the definition of conditional independence we conclude that A and B are independent if and only if $\mathbf{P}(A \,|\, B) = \mathbf{P}(A)$.

Repeated Sampling and Repeated Trials

Finally, we have computed probabilities for events that can best be described as repeated trials. These computations involved using both the addition rule for disjoint events and the multiplication rule. (See, for example, Problems #1, #2, #4 ,#5, #6 on Assignment #3 and Problem #5 on Assignment #4 for variations on this theme.) In fact, by multiplying together probabilities on branches of a tree diagram, we've derived the *General Multiplication Rule*. If $B_1, B_2, \ldots, B_k$ are any events, then

$$\mathbf{P}(B_1 \cap B_2 \cap \cdots \cap B_k) = \mathbf{P}(B_1)\,\mathbf{P}(B_2 \,|\, B_1)\,\mathbf{P}(B_3 \,|\, B_1 \cap B_2) \cdots \mathbf{P}(B_k \,|\, B_1 \cap \cdots \cap B_{k-1}).$$