Math/Stat 251 Fall 2015
Summary of Lecture from September 23, 2015

The purpose of today's class is to derive two of the most useful formulas in probability, namely the *Law of Total Probability* and *Bayes' Rule*. With a bit of guidance, it is not so difficult to derive these results.

Suppose that A and B are events with $\mathbf{P}(A) > 0$ and $\mathbf{P}(B) > 0$. From a Venn diagram, we see that A can be decomposed into two pieces, namely (i) that part of A which is also in B, and (ii) that part of A which is not in B. In other words, $A = (A \cap B) \cup (A \cap B^c)$ expresses A as a disjoint union. Using the property that says the probability of a disjoint union is the sum of the probabilities, we conclude

$$\mathbf{P}(A) = \mathbf{P}((A \cap B) \cup (A \cap B^c)) = \mathbf{P}(A \cap B) + \mathbf{P}(A \cap B^c). \tag{$*$}$$

Recall that the definition of the conditional probability of A given B is

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

Solving for $\mathbf{P}(A \cap B)$ gives

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

Similarly, we find

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

Substituting the previous two expressions into $(*)$ gives the **Law of Total Probability**.

$$\boxed{\mathbf{P}(A) = \mathbf{P}(A \mid B)\mathbf{P}(B) + \mathbf{P}(A \mid B^c)\mathbf{P}(B^c)}$$

However, we can also consider the conditional probability of B given A which is

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

so that solving for $\mathbf{P}(B \cap A)$ gives

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

But, of course, the events that $A \cap B$ and $B \cap A$ are the same! This means that $(\dagger)$ and $(\ddagger)$ give two different, but equal, expressions for $\mathbf{P}(A \cap B)$. Equating them gives

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

Finally, we can solve this for $\mathbf{P}(B \mid A)$. This leads to the first version of **Bayes' Rule**.

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

Finally, if we substitute the expression for $\mathbf{P}(A)$ from the Law of Total Probability into the first version of Bayes' Rule, then we arrive at another version of **Bayes' Rule**.

$$\boxed{\mathbf{P}(B\,|\,A) = \frac{\mathbf{P}(A\,|\,B)\,\mathbf{P}(B)}{\mathbf{P}(A\,|\,B)\,\mathbf{P}(B) + \mathbf{P}(A\,|\,B^c)\,\mathbf{P}(B^c)}.}$$

The importance of Bayes' Rule is that it tells us that if we know the conditional probability of A given B, then we can determine the conditional probability of B given A.

Example. Suppose that patients are randomly assigned to one of two treatment regimes. After completing a course of treatment, the investigator examines each patient and assesses whether or not the patient has shown marked improvement. Assume that 60% of patients are assigned to treatment 1 and the remaining 40% are assigned to treatment 2. Assume further that 70% of those receiving treatment 1 show marked improvement and 65% of those receiving treatment 2 show marked improvement. Suppose that one patient is selected at random.

(a) Determine the probability the patient shows marked improvement.

(b) If the patient shows marked improvement, determine the probability that the patient received treatment 1.

Solution. Let T_1 and T_2 denote the events that the patient received treatments 1 and 2, respectively. Let I be the event that the patient shows marked improvement. In terms of probabilities, the information given in the problem is that

$$\mathbf{P}(T_1) = 0.6, \quad \mathbf{P}(T_2) = 0.4,$$

$$\mathbf{P}(I\,|\,T_1) = 0.70, \quad \mathbf{P}(I^c\,|\,T_1) = 0.30, \quad \mathbf{P}(I\,|\,T_2) = 0.65, \quad \mathbf{P}(I^c\,|\,T_2) = 0.35.$$

(a) The Law of Total Probability implies that

$$\begin{aligned}
\mathbf{P}(I) = \mathbf{P}(I \cap T_1) + \mathbf{P}(I \cap T_2) &= \mathbf{P}(I\,|\,T_1)\,\mathbf{P}(T_1) + \mathbf{P}(I\,|\,T_2)\,\mathbf{P}(T_2) \\
&= (0.70)(0.60) + (0.65)(0.40) \\
&= 0.68.
\end{aligned}$$

(b) The first version of Bayes' Rule implies that

$$\mathbf{P}(T_1\,|\,I) = \frac{\mathbf{P}(I\,|\,T_1)\,\mathbf{P}(T_1)}{\mathbf{P}(I)} = \frac{(0.70)(0.60)}{0.68} = \frac{42}{68} = \frac{21}{34}.$$