The basics of probability – Part III

Conditional probabilities

Conditional probability is defined as the probability of an event given that another event has already occurred. It can be thought of as probability determined in the mode of either "what if" or "given that"

Consider the events A, B and $A \cap B$ in an experiment with sample space U.

Imagine that we make N independent replications of the experiment and let N_A, N_B and $N_{A \cap B}$ denote the number of times that the three events occur, respectively.

We will study the relative frequencies of these events:

1. The event A among all the experiments (i.e., N_A/N), and
2. The event A among the experiments where B occurred (i.e., $N_{A \cap B}/N_B$).

For (1) above, the relative frequency approaches the probability of event A as N tends towards infinity.

For (2), the relative frequency approaches the *conditional probability* of A *given* B, which should be viewed as the probability of A among those outcomes where B occurs.

Mathematically, if we let A and B be events and assume that $P(B) > 0$, the conditional probability of A given B (i.e., that the event A occurs when we already know that event B has occurred) is symbolized by $P(A \mid B)$, and is defined as :

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \qquad (6)$$

or, upon rearrangement, we have

$$P(A \cap B) = P(B) \,.\, P(A \mid B) \qquad (7)$$

Example 4

Several meat samples were analyzed by a chemical diagnostic kit test for the presence of E. coli bacteria O157.

Ideally, the test is positive (+) of the bacteria is present in the meat sample and negative (–) if it is absent. Table 1 below shows the results from such a test on E. coli bacteria O157 investigated on samples known to be with bacteria and another lot of samples without the bacteria:

Table 1: Presence of E. coli: number of positive and negative test results from samples with and without the bacteria

	Positive (+)	Negative (–)
Samples with bacteria	57	5
Sample blanks	4	127

So, we see from Table 1 that the test is not perfect. It is obvious that we sometimes incorrectly identify samples as positive although they do not have the bacteria, and to some cases, we also fail to identify the bacteria in samples where it is present.

We shall use the following two conditional probabilities to describe the test:

$$P\,(\text{positive} \mid \text{bacteria}) = 1 - P\,(\text{negative} \mid \text{bacteria})$$

$$P\,(\text{positive} \mid \text{no bacteria}) = 1 - P\,(\text{negative} \mid \text{no bacteria})$$

Therefore we have two situations:

A **false positive** result : the event {positive ∩ no bacteria} when a sample without bacteria is identified as positive.

A **false negative** result : the event {negative ∩ bacteria} where a sample with bacteria has been tested negative.

From the above example, therefore, we have 4 false positives and 5 false negative.

Now, the *sensitivity* of the test method is defined by the probability P (positive | bacteria). It is the probability that the diagnostic test will show correct result (i.e., that the test is positive) if bacteria are indeed present in the sample.

P (negative | no bacteria) is called *specificity*. It is the probability that the diagnostic test will show the negative result if bacteria are not present in the sample. Of course, we prefer the diagnostic tests are highly sensitive and have high specificity.

So, from Table 1, we estimate the sensitivity to 57/61 = 0.93 and the specificity to 127/132 = 0.96.

Bayer's theorem

We can use the Bayer's theorem to "invert" the conditional probability. If we know the conditional probability $P\,(B \mid A)$ and the two marginal probabilities $P(A)$ and $P(B)$ which are all greater than 0, then we can use the following formula to calculate the inverse conditional probability $P\,(A \mid B)$:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)} \tag{8}$$

Again we can generalize a situation where there are k disjoint sets, $A_1, \ldots, A_k$, such that $A_1 \cup A_2 \cup \ldots \cup A_k$ = sample space U whilst we already have observed event B.

Example 5

Let the Ai's denote different possible disease types and let B be an event corresponding to certain symptoms, such as fever or rash. The conditional probabilities $P\,(B \mid Ai)$ are the probabilities of the symptoms for individual with disease Ai.

We can then calculate P (Ai | B) for the different diseases so we can provide the best possible diagnosis, i.e. telling which disease is most likely given the symptoms we observe.

Let's use a generalized version of Bayer's theorem and the law of total probability as shown below.

If $A_1 \cup A_2 \cup \ldots \cup A_k = U$, and $P\,(Ai) > 0$ for all i, for any event B, we have the law of total probability

$$P\,(B) = P\,(B \mid Ai)\,P(Ai) + \ldots + P\,(B \mid A_k)\,P\,(A_k) \tag{9}$$

and if $P\,(B) > 0$, we have the generalized Bayer's theorem:

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{P(B \mid A_1)P(A_1) + \ldots + P(B \mid A_k)P(A_k)} \tag{10}$$

Independence and dependence

Two events, A & B, are mutually independent if and only if (under the probability distribution P)

$$P(A|B) = P(A) \quad \textbf{\textit{and}} \quad P(B|A) = P(B) \tag{11}$$

and also if

$$P(A \cap B) = P(A) \,.\, P(B) \tag{12}$$

Note:

The property (12) is also commonly referred to as the *multiplication rule*.

When these two events, A & B, are independent, the occurrence of one of the two events does not change the *likelihood* or probability that the other of the two events will occur. On the other hand, if two events A & B are dependent, then occurrence of one of the two events will alter the likelihood and the manner in which the other of the two events may occur.

Therefore, *independence* is an important concept in probability theory and statistics. In laboratory testing, we understand about how experiments and test observations can be independent. Now we can use the mathematical concept of independence to give us a precise meaning to this so-called intuitive understanding.

Example 6

When we throw two regular dice, we would naturally assume that the result of one die has no impact on the result of the other die, i.e., that an observed roll of 6 on the first die would not influence the probability of rolling 6 again on the second die.

In this experiment, the sample space has 36 elements, one for each combination of the two dice. If we assign probability 1/36 to each of the elements, then the outcomes of the two dice are independent. For example, the probability of rolling 6 on the first die and at the same time rolling 6 on the second die is 1/6 x 1/6 or 1/36.

Similarly, rolling at least 5 with the first die and an even number with the second die is 2/6 x 3/6 = 6/36, corresponding to the 6 possible outcomes (5,2), (5,4), (5,6), (6,2), (6,4) and (6,6).

In this example, we have equal probability to choose every possible element in this sample space. We call the probability distribution the *uniform distribution*.