Economics 520, Fall 2009

Lecture Note 2: Conditional Probability and Independence (CB 1.3)
REVISED 08/30/09

A fundamental topic of probability theory is how to update or modify probabilities to reflect the arrival of new information. This plays a major role in certain parts of statistics, and also in many economic models involving choice under uncertainty. We will start with a very simple example to try to develop some intuition.

Example 1: A randomly chosen child is either male or female, and either right-handed or left-handed. The probabilities of each possible combination are:

note: these probabilities are made up, to keep things simple.

	girl	boy
RH	.45	.4275
LH	.05	.0725

Note that the 4 probabilities sum to 1. Let E_1 denote the event of drawing a girl. Thus $E_1 = \{(\text{girl,RH}), (\text{girl,LH})\}$, and since $\{(\text{girl,RH})\}$ and $\{(\text{girl,LH})\}$ are disjoint, $P(E_1) = .45 + .05 = .5$ We call this a *marginal* or *unconditional* probability.

Any $\{\cdot\}$ is a subset of Ω. What is Ω here?

Next, we want to know what is the probability of left-handedness *among girls*? To do this, we need to define "probabilities given certain events."

Definition 1 *The conditional probability of an event E_2 given an event E_1 is*

$$P(E_2|E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)},$$

provided $P(E_1) > 0$.

Example 1 continued: Let E_2 be the event $\{(\text{girl,LH})\}$. Note that $P(E_2) = .05$. Also, since $E_2 \subset E_1$,

$$P(E_1 \cap E_2) = P(E_2) = .05.$$

So the conditional probability of left-handedness given the child is a girl, is:

$$P(E_2|E_1) = \frac{.05}{.45 + .05} = .10.$$

This says that 10% of girls are left-handed. By similar calculation, we can see that 13.5% of boys are left-handed. So in a certain sense, the gender of the child helps to predict whether the child is left-handed.

This is important: for a fixed E_1, $P(\cdot|E_1)$ acts just like an ordinary probability function.

Notice that by dividing by $P(E_1)$, we are essentially renormalizing by the marginal probability. In essence, we are restricting attention to the left column of probabilities in the Table, and renormalizing them so that they sum to 1.

Example 2: You draw a two cards out of a deck of 52. What is the probability that exactly one of them is an ace?

$$\begin{aligned} P(\text{exactly one ace}) &= \frac{\text{\# ways to draw exactly one ace}}{\text{\# ways to draw 2 cards out of 52}} \\ &= \frac{\text{\# ways to draw 1 of 48 non-ace cards} \times \text{\# ways to draw 1 of 4 ace cards}}{\text{\# ways to draw 2 cards out of 52}} \\ &= \frac{\binom{48}{1}\binom{4}{1}}{\binom{52}{2}} \\ &= \frac{32}{221}. \end{aligned}$$

Again, we call this a marginal or unconditional probability, to emphasize that it does not reflect any additional information.

In R, you can use "choose(n,k)" to calculate $\binom{n}{k}$

Now, someone tells you that at least one of the two cards is an ace. We would like to be able to say, given that at least one of the two cards is an ace, what is the probability that exactly one is an ace?

Let $E_1 = \{\text{At least one ace}\}$ and $E_2 = \{\text{Exactly one ace}\}$. Since $E_2 \subset E_1$, the numerator probability is equal to $P(E_2) = 32/221$, as we calculated previously. The probability of the conditioning event E_1 (at least one ace) is the sum of the probabilities of one ace and two aces. The latter is 1/221, so the probability of the conditioning event is 33/221. (Alternatively the probability of at least one ace is 1 minus the probability of no aces which is $1-(48/52)\cdot(47/51)$). Then the conditional probability is the ratio (32/221)/(33/221)=32/33. □

$E_1 \cap E_2 = E_2$

$P(E) = 1 - P(E^c)$

Example 3: A simple example is that of two coin tosses. What is the probability of two heads given that you have at least one head in the two tosses. $E_1 = \{HH, TH, HT\}$, with probability 3/4, $E_2 = \{HH\}$, so $E_1 \cap E_2 = \{HH\}$ with probability 1/4, so

$$P(E_2|E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)} = \frac{1/4}{3/4} = \frac{1}{3},$$

not 1/2 as many people think at first.

Example 4: It is easy to get mixed up when thinking about conditional probabilities. A famous example is the Monty Hall problem. You are a contestant in a game show and have to choose one of three doors. Behind one of the doors is a prize; the other doors are empty. After you choose a door, the game show host (Monty Hall) opens one of the other two doors and shows you there is no prize behind that door. The host then offers you the opportunity to switch from the door you chose to the third door. Should you switch?

To solve this problem we first have to remove some of the ambiguities in the description above. We assume that the door behind which the prize is located is choosen randomly, with probability 1/3 for each door. More importantly, we assume that the host will always open one of the doors not chosen by the contestant and not containing the prize. If there is a choice for the host, for example if you choose door A and the prize is in fact behind door A, the host will choose one of the eligible doors (B or C) with equal probability. Let L be the location of the prize, and let H be the door opened by Monty Hall. Suppose that you pick door A. Then

this notation is a bit loose. $L = A$ really means the *event* that prize is in A, etc.

$$P(L = A) = P(L = B) = P(L = C) = \frac{1}{3};$$

$$P(H = B|L = A) = \frac{1}{2} = P(H = C|L = A);$$

$$P(H = B|L = B) = 0; \quad P(H = C|L = C) = 0;$$

$$P(H = B|L = C) = 1; \quad P(H = C|L = B) = 1.$$

Given your initial choice of door A, you have two possible strategies:

1. Stay with door A.
2. Change your door. For example, if Monty Hall opens door B, you would pick door C.

Clearly, under strategy 1, your probability of winning is 1/3.

To calculate the probability of winning under strategy 2, we need to calculate $P(L = C|H = B)$. A useful fact here is that:

$$P(Y|X) = \frac{P(X \cap Y)}{P(X)} \quad \Rightarrow \quad P(X \cap Y) = P(X)P(Y|X).$$

Using this fact, we can write:

$$P(L = C|H = B) = \frac{P(L = C \cap H = B)}{P(H = B)} = \frac{P(L = C)P(H = B|L = C)}{P(H = B)}.$$

Expand the denominator:

In the first line, we are partitioning the event $H = B$ into 3 disjoint subevents

$$\begin{aligned} P(H = B) &= P(H = B \cap L = A) + P(H = B \cap L = B) + P(H = B \cap L = C) \\ &= P(L = A)P(H = B|L = A) + P(L = B)P(H = B|L = B) + P(L = C)P(H = B|L = C) \\ &= \frac{1}{2} \cdot \frac{1}{3} + 0 + 1 \cdot \frac{1}{3} \\ &= \frac{1}{6} + \frac{2}{6} = \frac{3}{6} \end{aligned}$$

So

$$P(L = C|H = B) = \frac{2/6}{3/6} = \frac{2}{3}.$$

By a symmetric argument,

$$P(L = B | H = C) = \frac{2}{3}.$$

So your probability of winning under strategy 2 is $\frac{2}{3}$. You should follow strategy 2.

For a nice graphical explanation, see

`http://math.ucr.edu/~jdp/Monty_Hall/Monty_Hall.html`

Example 5: Another typical example is that of a diagnostic test for diseases. Suppose 1 in 10,000 people in the population have a particular disease. A test exists with the following properties. If you have the disease and get tested the test will come out positive 99% of the time and negative (a "false negative") 1% of the time. If you do not have the disease, the test will come out positive (a "false positive") 5% of the time, and negative 95% of the time. What is the probability that a randomly choosen person from the population who tests positive actually has the disease?

We are interested in the probability

$$P(Disease|Positive) = \frac{P(Disease, Positive)}{P(Positive)}$$

$$= \frac{P(Positive|Disease)P(Disease)}{P(Positive, Disease) + P(Positive, Disease^c)}$$

$$= \frac{P(Positive|Disease)P(Disease)}{P(Positive|Disease)P(Disease) + P(Positive|Disease^c)P(Disease^c)}$$

$$= \frac{(99/100)(1/10000)}{(99/100)(1/10000) + (5/100)(9999/10000)}$$

$$= \frac{99}{99 + 49995} \approx .002.$$

Even though the test seems very good, giving the correct result at least 95% of the time, the probability of actually having the disease once you test positive is still very small since it is so small to begin with (0.0001). In fact, it has gone up by a factor of 20, because the test gives the wrong answer in only one in twenty cases for healthy people. □

In the last example we were given conditional probabilities in one direction (test results given health status) but were interested in conditional probabilities in the other direction (health status given test results). A general result concerning this type of calculation, is referred to as Bayes' theorem:

Result 1 *If $E_1, E_2, \ldots, E_k$ form a partition of Ω, then*

Try proving this for $k = 2$

$$P(E_j|E) = \frac{P(E_j) \cdot P(E|E_j)}{\sum_{i=1}^{k} P(E_i) \cdot P(E|E_i)}.$$

Calculations involving conditional probabilities can be greatly simplified if a particular relation holds:

Definition 2 *Two events E_1 and E_2 are independent if*

$$P(E_1 \cap E_2) = P(E_1) \cdot P(E_2).$$

If two events E_1 and E_2 are independent and $P(E_1) > 0$, then $P(E_2|E_1) = P(E_2)$. The condition that $P(E_1) > 0$ is important, however. Note that the empty set $\emptyset$ is independent of any event because both left-hand side and right-hand side probabilities are equal to zero.

Definition 3 *Three events E_1, E_2 and E_3 are jointly independent if :*

1. (a) E_1 and E_2 are independent,
   (b) E_1 and E_3 are independent,
   (c) E_2 and E_3 are independent.
2. $P(E_1 \cap E_2 \cap E_3) = P(E_1) \cdot P(E_2) \cdot P(E_3)$.

Similarly, joint independence of four events requires that all combinations of three events are jointly independent as well as that the probability of the intersection is equal to the product of the probabilities.

To see why both conditions 1 and 2 are necessary in the three event definition of independence consider the case where $E_1 = E_2$ and $E_3 = \emptyset$. In that case condition 2 would be satisfied but no one would want to call the three events independent. To see that condition 2 is necessary consider a roulette wheel with 8 numbers. Event E_1 is an odd number, E_2 is a number (strictly) less than 5, and E_3 is a number in the set $\{1, 3, 6, 8\}$. In that case E_1 is independent of E_2, E_1 is independent of E_3, and E_2 is independent of E_3, but the probability of the intersection is not equal to the product of the three marginal probabilities:

$$P(E_1) \cdot P(E_2) \cdot P(E_3) = 1/8 \neq 1/4 = P(E_1 \cap E_2 \cap E_3),$$

and again this does not correspond to the intuitive notion of independence.