**Probability**

There are three forms of basic probability: theoretical, empirical, and subjective. The *theoretical*, or *a priori* probability is based on a model in which all outcomes are equally likely. We can count precisely the number of possibilities used to calculate the probabilities. For example, we know that a standard die (singular of dice) has six sides and the outcome of any side landing up is equal to all the other sides. Since there is only one 2 on a die, the probability of a die landing on a 2 is one out of six, or simply 1/6 (0.1667). Given a coin, the probability of a head landing up is 1/2 (0.5). Given three doors on a game show, the probability of selecting any door is 1/3 (0.333).

In an *empirical* probability we base the probability on the results of observations or experiments. We cannot count all the possible outcomes because we are basing our probabilities on what we have observed thus far. For example, if it rains an average of 100 days a year and someone asks what is the probability of rain on any one day, the answer is what we observed: 100/365 (0.2739). If history has shown that a big hurricane hits Tampa, Florida once every 15 years and someone asks what is the probability that Tampa will have a big hurricane this year, then the answer is 1/15 (0.0667).

The third form of probability is subjective. With this probability one uses personal judgment or intuition. For example, if you go to college today, you will be more successful in the future. We will not examine subjective probability in this course.

Before we get into the study of probabilities in any more depth, let’s discuss the topic of *possible outcomes*. You need to understand the possible outcomes of an event before you can calculate probabilities. As we have already seen, if a die has six sides, then there are six possible outcomes in tossing a die. If we toss two dice, then the total number of possible outcomes is a multiplication of the individual outcomes, or in this case, 36 (6 x 6). If a restaurant offers two choices of an appetizer, five choices for a main entrée, and three choices for a dessert, then there are 30 possible three-course meals (2 appetizers \* 5 entrées \* 3 desserts). If a college offers 12 natural science courses, 15 social science courses, 10 English courses, and 8 fine arts courses, how many possible outcomes of all the courses are there? (12 \* 15 \* 10 \* 8 = 14400)

Let’s consider another example – license plates. What if a license plate can have six symbols (positions on the plate) and the symbols can only be the digits 0 – 9. How many possible outcomes of license plate numbers are there? Since each position has 10 possible digits and there six positions, the answer is 10 \* 10 \* 10 \* 10 \* 10 \* 10, or 1,000,000. (Or 106) What if the symbols can be either the digits 0 – 9 or letters of the alphabet A – Z? Now each position can have 36 different symbols, so our answer is 366. How many possible postal codes are there in the U.S.? 10 \* 10 \* 10 \* 10 \* 10. What about Canada? How many different telephone area codes are possible (knowing that an area code cannot start with a 0 or a 1)? 8 \* 10 \* 10.

Now let’s go back to basic probability. When we write P(A) we are saying “the probability of A occurring”, or (number of ways A can occur) / (total number of possible outcomes). As we have already seen, the probability of a head landing in a coin toss is 1/2. The probability of rolling a 6 on a single die is 1/6. What is the probability of rolling a 7 using two dice? First figure out how many ways A can occur. In other words, how many ways can 7 occur when tossing two dice? You can toss a 1 and a 6, a 6 and a 1, a 2 and 5, a 5 and a 2, a 3 and a 4, and a 4 and a 3. That is six different ways. Now figure out the total number of possible outcomes. As we have already seen, with two dice there are 36 possible outcomes. Thus our answer is 6/36, or 1/6, or 0.1667.

What is the probability that a family with three children will have two boys and one girl? For this one, let’s list all the possible outcomes first. We can have BBB, BBG, BGB, BGG, GBB, GBG, GGB, and GGG. That gives us eight possible outcomes. Of those eight, how many have one girl and two boys? BBB, **BBG**, **BGB**, BGG, **GBB,** GBG, GGB, and GGG. That is 3/8.

It is often interesting to compare theoretical probabilities to empirical probabilities (when we can calculate the theoretical probabilities). For example, if we toss two coins together, there are four possible outcomes: HH, HT, TH, and TT. (2 \* 2) Thus the probability of tossing a Heads and a Tails (HT) is 1/4. This is the theoretical probability. But what if we actually take two coins and toss them, let’s say 10 times. We will record the results:

HH

HT

HT

TH

HH

HT

TT

HT

HH

TT

TH

Let’s count the number of occurrences we actually received (the empirical). There were 3 HH, 4 HT, 3 TH, and 2 TT. So out of 10 tosses, we had 3/10 HH, 4/10 HT, 3/10 TH, and 2/10 TT. How do those results compare to the theoretical? Well, we just stated that the theoretical for any of the four possible outcomes is 1/4. We had 4/10 HT, which is more the 1/4 theoretical. We only had 2/10 TT, which is less than 1/4.

If we know the probability (P(A)) of some event occurring, it is very simple to calculate the probability of some event not happening: 1 – P(A). For example, if the probability of rolling a 7 with two dice is 6/36, then the probability of not rolling a 7 with two dice is 30/36.

How do we calculate the probability of two independent events happening? (We can say the events are independent if the outcome of one event does not affect the outcome of the next event.) The probability of A and B occurring together is simply the P(A) \* P(B). Or in other words, to calculate the probability of this happening AND this happening, you multiply the two probabilities. For example, suppose you toss one coin three times in a row. What is the probability of getting three tails in a row? Each toss is independent of the others, so the answer is 1/2 \* 1/2 \* 1/2, or 1/8. Notice that is 1/2 for the first coin toss, 1/2 for the second toss, and 1/2 for the third toss.

What is the probability that a particular city will be hit by a 100-year flood in two consecutive years? Well, a 100-year flood means that this kind of flood happens once every 100 years, or 1/100. For two consecutive years, you would have 1/100 \* 1/100, or 1/10000, or 0.0001.

You should note that when events are independent of each other then one event really has nothing to do with another. For example, you are playing craps in Vegas. You have had a string of bad luck and figure it is time to have some good luck because “things have to balance out”. This is a bad assumption and is called Gambler’s Fallacy. The next throw of the dice is independent from the previous and has the same probabilities as the first throw.

What if you ask what is the probability of either this happening OR that happening (not this happening AND that happening, as above)? If the events are independent of each other, then the calculation of P(A OR B) = P(A) + P(B). For example, you roll a single die. What is the probability of rolling either a 2 or a 3? P(2 OR 3) = P(2) + P(3) = 1/6 + 1/6 = 2/6 = 1/3 = 0.333.

Calculating probabilities so far has been pretty easy. To calculate the probability of this AND that, you multiply. To calculate the probability of this OR that, you add. But what is the probability of something happening at least once given a number of trials. The formula is:

P(at least one event A in n trials) = 1 – P(not A)n

where n is the number of trials. For example, if we ask what is the probability of a 100-year flood happening this year, it is simply 1/100, as we noted earlier. But what is the probability of at least one 100-year flood in, say, the next 50 years? (The key phrase here is “at least”.)

P(A) = 1/100

P(not A) = 99/100

1 – P(not A)n = 1 – (99/100)50 = 1 – 0.605 = 0.395

What is the probability of at least one 100-year flood in the next 100 years? You would think it would be 100%, but it is not:

1 – P(not A)n = 1 – (99/100)100 = 1 – 0.366 = 0.634

It is only 63.4%.

Let’s try another one. Your favorite hamburger joint has a new promotion. If you buy a large drink, your cup has a scratch-off label on it. You are told that 1 in 20 cups will win a free quarter pounder hamburger. If you purchase five large drinks, what is the probability that you will win a quarter pounder?

1 – P(not A)n = 1 – (19/20)5 = 1 - 0.7738 = 0.2262

Finally, let’s examine a concept called *expected value*. The expected value is what you expect to gain or lose after considering a number of possible outcomes. Simply stated, the expected value can be calculated as event1’s probability \* event1’s value + event2’s probability \* event2’s value + …

For example, suppose lottery tickets cost $1 each and have the following probabilities and values: 1 in 5 tickets win a free $1 ticket; 1 in 100 win $5; 1 in 100,000 win $1000; 1 in 10 million win $1,000,000. What is the expected value of the lottery ticket to the person that purchases the ticket?

Expected value = 100% \* -$1 (you have to buy a ticket!) + 1/5 \* $1 + 1/100 \* $5 + 1/100000 \* $1000 + 1/10,000,000 \* $1,000,000

Note that since you bought the ticket, that is $1 out of your pocket, or -$1.

So expected value = -$1 + $0.20 + $0.05 + $0.01 + $0.10 = -$0.64. You should expect to lose 64 cents for every dollar you spend on this lottery.

Let’s try one more example. Suppose an insurance company sells policies for $500 each. The company knows that 10% of all customers will submit a claim that year and that the claims average $1500 each. How much money can the company expect to make on each customer?

Expected value = 100% \* $500 (the money the company makes when a customer buys a policy) + 10% \* -$1500 (the company pays out this amount, so it is negative to the company) = $500 + -$150 = $350. If the company sells 10,000 policies, they will make 10,000 \* $350, or $3,500,000.

It is important to note the perspective of the expected value. In the insurance company example, we were considering the expected value to the company. Since they make money selling policies and lose money settling claims, the $500 policy cost was positive while the $1500 claim payout was negative. If you were asked to calculate the expected value to the customer buying the policy, the $500 to buy the policy would be negative while the $1500 claim would be positive. (Go back and see how this works with respect to the lottery example.)