Section 6.3: Probability

Suppose you were to roll a six-sided die (one die, two dice – die is the singular of dice). What's the probability of rolling a 1? Well, if the die is fair and all six numbers are equally likely, then the probability of getting any one number is just 1/6.

This is the idea of classical probability – if all outcomes are equally likely, then the probability of event *E* happening is just the number of outcomes in which *E* happens, *n(E)*, divided by the total number of outcomes *n*:

$$P(E) = \frac{n(E)}{n}$$

Example

If you roll two four-sided dice, what's the probability of rolling a total of 5?

Answer: The brute force method involves writing out all possible outcomes.

11	12	13	14
21	22	23	24
31	32	33	34
41	42	43	44

Then, assuming that the dice are fair, there are sixteen equally likely outcomes, and four of them – 41, 32, 23, and 14 – lead to a total of 5. Then P(total of 5) = 4/16 = ¼ or 25%.

Example

What's the probability of winning the Lotto 6/49 if you buy only one ticket?

Answer: You choose six numbers from 49, so there are $_{49}C_6 = 13,983,816$ possible combinations. There's only one way to win, which is if all six numbers drawn match the ones you've chosen. So your probability is

$$P(\text{winning}) = 1/13,983,816$$

Example

The Saanich city council has four members: Alex, Barbara, Charlie, and Dorothy. Two of these members are to be selected to form a subcommittee to study the city's traffic problems.

a) How many different subcommittees are possible? What probability would you assign to each one if there is an equal chance of selecting each council member?

b) What is the probability that Dorothy is a member of the committee?

c) What is the probability that Charlie and Dorothy are both selected?

d) What is the probability either Charlie or Dorothy or both are selected?

Answer:

a) possible outcomes = {AB, AC, AD, BC, BD, CD} for six outcomes (note that AB=BA since order doesn't matter). Or you could just say $_4C_2 = 6$. If they are all equally likely, then each has a probability of 1/6.

b) P(D) = (3 subcommittees with Dorothy)/(6 in total) = ½

c) P(CD) = (1 subcommittee with C & D)/6 = 1/6

d) P(C or D) = (5 subcommittees with C or D)/6 = 5/6

   or P(C or D) = P(C) + P(D) – P(CD) = ½ + ½ – 1/6 = 5/6

   or P(C or D) = 1 – P(AB) = 1 – 1/6 = 5/6

   (note that AB is the only committee with neither C nor D)

Which brings us to the addition rule for probability:

$$P(A \text{ or } B) = \frac{n(A \text{ or } B)}{n}$$
$$= \frac{n(A) + n(B) - n(AB)}{n}$$
$$= P(A) + P(B) - P(AB)$$

However, in real life, you frequently get situations where not all outcomes are equally likely. One tool that we can use in that situation is called a **contingency table**.

Contingency tables

To study contingency tables, it's easiest to look at an example.

To simplify matters, let's assume that students at Interurban are enrolled in either Technology or Business (but not both). Let's also assume that men and women are equally represented in Business, but that only 10% of Technology students are women (which, frankly, is being generous!). Let's also assume that there are the same number of Technology and Business students. Then our entire student population of 100 (to make the numbers easier) would look like this:

	Technology	Business	Total
Male	45	25	70
Female	5	25	30
Total	50	50	100

Let's calculate the probability that if a student were randomly selected from this group, that the student was enrolled in Technology. Then what we wish to calculate is

$$P(T) = \frac{n(T)}{n}$$

where T is technology, $P(T)$ is the probability of being in technology, $n(T)$ is the number of "technology events" or in this case the number of students in technology, and n is the total number of students. Then $P(T) = 50/100 = \frac{1}{2}$ or 50%. (The number 50 came from the total at the bottom of the technology column.)

Let's calculate the probability that the student was a female business student. This would be $P(FB) = 25/100 = \frac{1}{4}$ or 25%. The way we find n(FB) is we look at the cell in the intersection of the "female" row and the "business" column.

Let's calculate the probability that the student was male or in business. P(M or B) can be calculated either by adding up all the students who are male or in business or in both: P(M or B) = (45 + 25 + 25)/100 = 95%. Or you could say that it's going to be the total minus the FB students: (100-5)/100 = 95%. Or you could say that it's P(M)+P(B)–P(MB) = 70% + 50% – 25% = 95%.

Let's calculate the probability that if the student were female, that she was enrolled in Technology. The way we write this in symbols is P(T | F), which we read as P(T "if" F). What we are really asking is that if we only look at the female students (we limit our population), what's the probability of getting a technology student from among those female students? Then P(T | F) = $\frac{n(FT)}{n(F)} = \frac{5}{30} = \frac{1}{6}$.

Let's calculate the probability that that if the student were in Technology, that she were female. This seems like it's the same question in the last paragraph, but it's not. We're now limiting our population to the technology students, and calculating:

$$P(F \mid T) = \frac{n(FT)}{n(T)} = \frac{5}{50} = \frac{1}{10}$$

We can also ask the question: Are the events "student is female" and "student is enrolled in Technology" independent? What this is asking is "are the probabilities of being female the same for the entire population and for the technology population?". The way we tell is to calculate $P(F)$ and $P(F \mid B)$. If these two probabilities are the same, then the probability of being female **does not depend** on whether the student is in technology and we say the events are **independent**. Otherwise, we say that one depends on the other and the events are **dependent**.

So, $P(F) = 30/100 = 30\%$. We already found that $P(F \mid T) = 10\%$. So these probability are not the same, and these events are **dependent**. Notice that we could instead calculate $P(T)$ and $P(T \mid F)$ and compare those probabilities. $P(T) = 50\% = \frac{1}{2}$ and $P(T \mid F) = 1/6$, so we will reach the same conclusion.