Module 3

Probability: The Language of Uncertainty

James Baglin

Last updated: 13 July, 2020

Overview

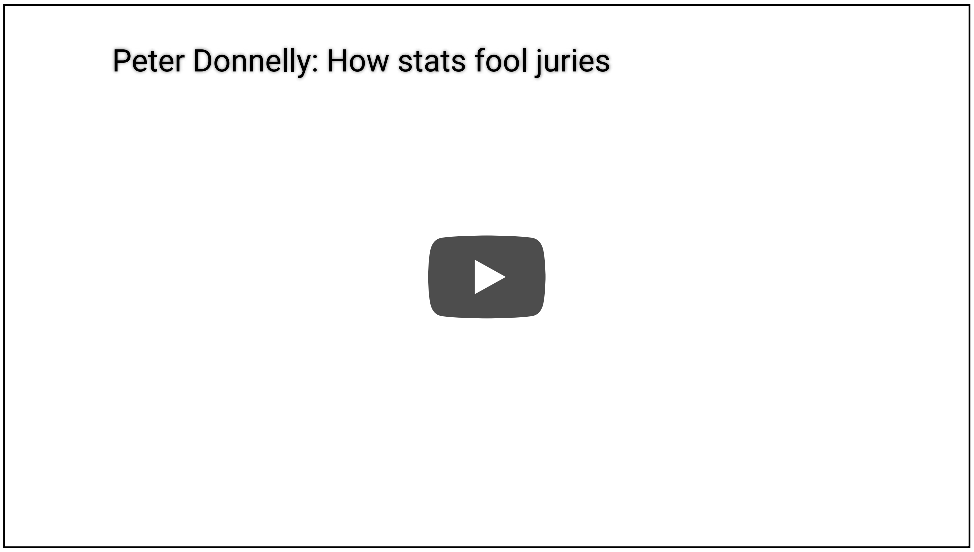
##Summary

Learning Objectives

The learning objectives associated with this module are:

- List the basic principles of probability.
- Express uncertainty using probability.
- Define common probability terms.
- Solve basic probability problems.
- Solve problems involving permutations and combinations.

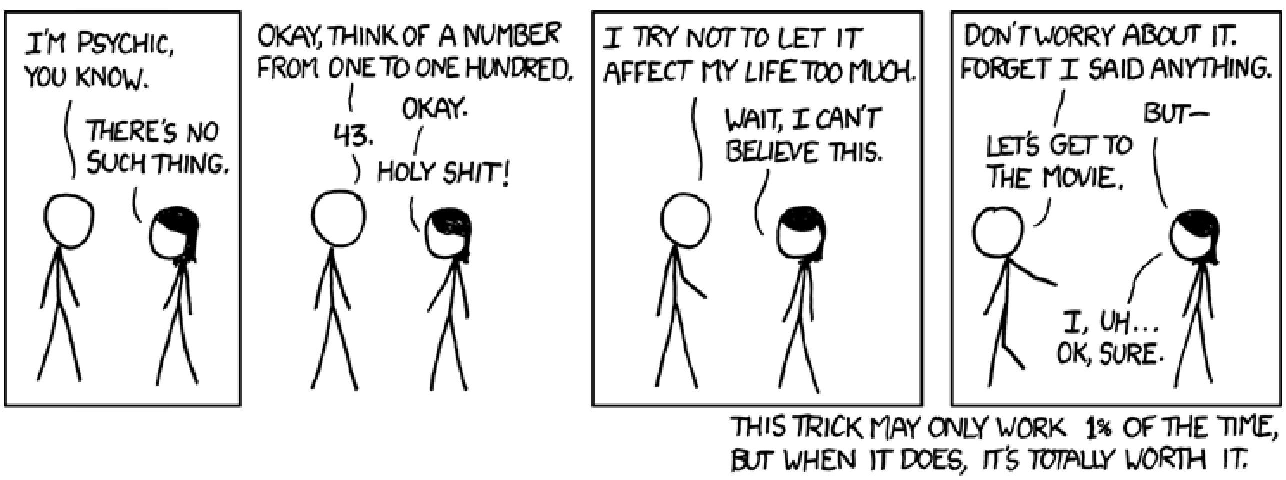
Module Video



Probability

Ouch!

Probability is defined as the proportion of times a random **event** occurs in a very large number of **trials**. Probability must always be a value between 0 and 1. What we define as an “event” and a “trial” depends on the context. In statistics, we estimate the probability of an event using a sample and note its relative frequency, f/n, where f is the frequency or number of times an event occurs and n is the total sample size. As the sample size n increases, the sample will begin to approximate the true population probability.



The **sample space** is the set of all possible outcomes of an **experiment**. We use the term “experiment” very loosely here to refer to any activity where an outcome can vary. The most common “experiment” in statistics is taking a sample from the population. Let’s consider an example estimating the probability of randomly selecting people from the Australian population who have different levels of daily fruit consumption.

If we sampled random Australians we could categorise them into “< 1 serve”, “1 serve”, “2 serves” or “3 serves or more”. These outcomes or categories make up the sample space and the process of the sampling from the population is our “experiment”. An event is an outcome of the experiment, e.g. that someone we sample consumes < 1 serve of fruits per day. The probability of an event measures “how often” we expect the event to occur in the long run and is estimated using f/n. A probability estimated using a sample is an example of a statistic.

There is a entire abstract language used to express the laws of probability. We will look at the conventions here and relate these to real world examples. Once we have the basics down, we will concentrate on using these principles for practical purposes. Consider the following table adapted from Table 10 of the

Australian Bureau of Statistics 2011 - 2012 National Health Survey (http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4364.0.55.0012011-12?OpenDocument). The table shows the estimated persons (reported in hundreds of thousands, ’000) that occupy the different levels of fruit intake. These estimates were taken from self-reports of a representative sample of approximately 18,400 households. The table splits these estimates across age and gender. We will use this table to explain the language of probability in a meaningful way and learn to solve problems that you will face in this course and later modules.

You can download this table here (https://docs.google.com/a/rmit.edu.au/spreadsheets/d/1uVEsRAdn4RlXhAd44L331qds49BUGD3cCzulYZezeLk/edit?usp=sharing).

Rules

We denote probability using $Pr(A)$, where A refers to an event of interest, e.g. $Pr(2 \text{ serves})$. Let’s run an experiment by randomly selecting an Australian adult from the population. We will use the ABS data table as the probabilities of different events occurring. We use the relative frequency formula, f/n, to estimate this basic probability. First, we find the trial size, which is the estimated population size of Australians over the age of 18 as of 2012, 17,042, which equates to 17,042,000 people (the table is reported in ’000). Next, we find the frequency of people who eat two serves of fruit a day, 4984. Using the f/n formula, we write…

$$Pr(2 \text{ serves}) = 4984/17042 = 0.292$$

Therefore, the probability of randomly selecting an Australian adult who consumes two serves of fruits a day is $.292$. The probability of any event, $Pr(A)$, must be between 0 and 1. If event A can never occur, $Pr(A) = 0$. If event A always occurs when the experiment is performed, then $Pr(A) = 1$. The sum of all probabilities of all possible events must equal 1:

$$Pr(<1 \text{ serve}) + Pr(1 serve) + Pr(2 \text{ serves}) + Pr(3 \text{ serves or more}) = 1$$

$$\frac{3368 + 5445 + 4984 + 3244}{17042} = \frac{17042}{17042} = 1$$

Two events are said to be **mutually exclusive** if, when one event occurs, the other cannot and vice versa. Mutually exclusive sets have no **intersection**: $Pr(A \cap B) = 0$. We use $\cap$ to denote an intersection. The levels of fruit consumption are mutually exclusive. A person cannot occupy more than one category at a particular time.

$$Pr(1 \text{ serve} \cap 2 \text{ serves}) = 0$$

If two events can occur simultaneously, the events are not mutually exclusive and an intersection is possible, e.g. $Pr(<1 \text{ serve} \cap \text{Male})$. Note that the probability of an intersection can still be 0 even though an intersection is possible. For example, what if all Australian males had their favourite fruit prepared and delivered to them daily? Then, $Pr(<1 \text{ serve} \cap \text{Male})$ would probably be close to 0.

When we talk about intersections, the term "and" is often used. For example, what is the probability that a randomly sampled Australian adult is in the "< 1 serve" interval AND male? You need to become comfortable with the language of probability. However, don't worry, its a lot easier than you think when we apply the rules to real world examples. Using the Fruit Intake table, we can find this probability to be:

$$Pr(<1 \text{ serve} \cap \text{Male}) = \frac{2036}{17042} = .119$$

The **union** of two events A or B, is an event when either A or B, or A and B occur. We write a union using the $\cup$ symbol. The word "or" is used to refer to a union. For example, consider $Pr(1 \text{ serve} \cup <1 \text{ serve})$, or in words, "What is the probability that a randomly sampled Australian will consume 1 serve of fruit OR less?"

$$Pr(1 \text{ serve} \cup <1 \text{ serve}) = \frac{3368 + 5445}{17042} = .517$$

The **complement** of an event consists of all outcomes of the experiment that do not result in an event. For example, $Pr(\overline{<1 \text{ serve}}) = Pr(1 \text{ serve}) + Pr(2 \text{ serves}) + Pr(3 \text{ serves or more})$. The complement of an event usually has a bar over the top of the event which is read as "not". For example:

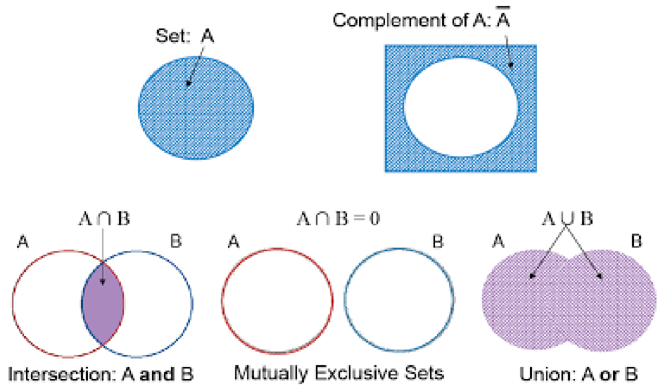
$$Pr(\overline{<1 \text{ serve}}) = Pr(1 \text{ serve}) + Pr(2 \text{ serves}) + Pr(3 \text{ serves or more}) =$$

$$\frac{5445 + 4984 + 3244}{17042} = .802$$

or

$$Pr(\overline{<1 \text{ serve}}) = 1 - Pr(<1 \text{ serve}) = 1 - \frac{3368}{17042} = .802$$

The figure below provides a notational overview of the basic probability concepts.



##Multiplication Law

If two events are **independent** (i.e. the probability of the first event does not impact the probability of the second event), then the intersection is equal to the probability of the first event multiplied by the second event, $Pr(A \cap B) = Pr(A) \times Pr(B)$. However, if independence does not hold, the two events are said to be **dependent**.

For example, if the consumption of fruit was independent of gender, then $Pr(< 1 \text{ serve} \cap \text{Male}) = Pr(< 1 \text{ serve}) \times Pr(\text{Male})$. However, is the assumption of independence safe for fruit consumption and gender? It's often believed that adult males tend to consume less fruits than females. Therefore, the assumption of independence is not safe and the multiplication rule will not hold. Let's check this assumption using the multiplication rule.

Recall…

$$Pr(< 1 \text{ serve} \cap \text{Male}) = \frac{2036}{17042} = .119$$

Now, according to the multiplication rule for independent events:

$$Pr(< 1 \text{ serve} \cap \text{Male}) = Pr(< 1 \text{ serve}) \times \text{Pr}(\text{Male}) = \frac{3368}{17042} \times \frac{8406}{17042} = .097$$

The two probabilities are not the same, therefore, as suspected, fruit consumption and gender are not independent. Males are more likely to be in the "< 1 serve" category. The take home message is that the multiplication rule does not apply when events are dependent.

Addition Laws

Finding the probability of a union depends on whether or not the events are mutually exclusive. If the events are **mutually exclusive**, we simply add the events. Recall…

$$Pr(1 \text{ serve} \cup < 1 \text{ serve}) = \frac{3368 + 5445}{17042} = .517$$

If the events are **not mutually exclusive**, we need to subtract the intersection from the addition law. For example:

$$Pr(< 1 \text{ serve} \cup \text{Male}) = Pr(< 1 \text{ serve}) + Pr(\text{Male}) - Pr(< 1 \text{ serve} \cap \text{Male}) =$$

$$\frac{3368}{17042} + \frac{8406}{17042} - \frac{2036}{17042} = .571$$

Conditional Probability

The probability that an event, B, will occur given that another event, A, has already occurred is called the **conditional probability** of B given A. The "|" symbol, read as "given" is used to denote a condition. Conditional probability can be written as follows:

$$Pr(B|A) = \frac{Pr(A \cap B)}{Pr(A)}$$

Using an example...

$$Pr(< 1 \text{ serve}|\text{Male}) = \frac{Pr(\text{Male} \cap < 1 \text{ serve})}{Pr(Male)} = \frac{2036/17042}{8406/17042} = .242$$

We can also use conditional probability to check independence. The two events A and B are independent if and only if $Pr(A|B) = Pr(A)$ or $Pr(B|A) = Pr(B)$. Otherwise, the events are dependent. Let's use this rule to reconfirm that gender and fruit consumption are dependent.

$$Pr(< 1 \text{ serve}|\text{Male}) = .242$$

$$Pr(< 1 \text{ serve}) = \frac{3368}{17042} = .198$$

We find $Pr(B|A) \neq Pr(A)$. The probabilities are not equal, therefore, dependency is present. Given that you're male, you are more likely to be consuming less than 1 serve of fruit per day than females. Males need to eat more fruit!

#Permutations and Combinations

##Permutations

Let's assume you are voting in a local council election. There are six candidates. You need to vote for the top three. How many possible ways can you assign your votes, 1st, 2nd and 3rd preference? This is an example of a permutation problem. **Permutations** refers to all the possible ways of selecting something where order matters. Here are three possible permutations for the voting example:

Veronica Paskett	Milagros Depaolo	Loraine Muntz	Thuy Silverberg	Myriam Hakes	Maude Dimery
1st	2nd	3rd	-	-	-
-	1st	2nd	3rd	-	-
-	2nd	1st	-	3rd	-
...	...	...	...	...	...

As you can see, there are many more possible ways to assign your votes. To quickly calculate all the possible permutations, we can use the following formula:

$$P(n, k) = \frac{n!}{(n-k)!}$$

The $!$ symbol refers to the factorial of a number. For example, $6! = 6 \times 5 \times 4 \times 4 \times 3 \times 2 \times 1 = 720$. In R, we can use `factorial()`:

```
factorial(6)
```

```
## [1] 720
```

k is the number to choose, in this example, 3.

Solving the voting problem:

$$P(6,3) = \frac{6!}{(6-3)!} = \frac{6!}{3!} = \frac{720}{6}$$

Using R:

```
factorial(6)/factorial(6-3)
```

```
## [1] 120
```

Combinations

Now let's change the problem. You have four spare tickets for a sport grand final. You have ten friends that you know would like to go. You need to weigh up the social impact of inviting different combinations of friends. First you need to know how many possible combinations of selecting four out of ten friends need to be considered. **Combinations** refer to all the possible ways of selecting a certain number of things from a larger group. Here are three possible combinations:

Leah	Rosalie	Marlena	Tarra	Graham	Gilberto	Marcos	Gladis	Otha	Jeremiah
Ticket	-	Ticket	-	Ticket	-	-	-	-	Ticket
-	-	-	-	-	-	Ticket	Ticket	Ticket	Ticket
-	Ticket	-	Ticket	-	Ticket	-	-	-	Ticket

Notice with combinations that order does not matter. Leah, Marlena, Graham and Jeremiah is the same combination as Marlena, Leah, Jeremiah and Graham. All four friends will go to the game and will be your best friends forever, regardless of the order by which they're selected.

The formula for combinations is as follows:

$$C(n,k) = \frac{n!}{(n-k)!k!}$$

This is known as the "choose"" formula or the binomial coefficient (we will revisit this in Module 4). Solving, we find:

$$C(n,k) = \frac{n!}{(n-k)!k!} = \frac{10!}{(10-4)!4!} = \frac{10!}{(6)!4!} = \frac{3628800}{17280} = 210$$

Using R:

```
choose(10,4)
```

```
## [1] 210
```

That's incredible. There are 210 different combinations of friends that you could end up taking to the final. Considering the social ramifications of each combination will take weeks of deliberation, leave it to chance. Use a raffle instead. That way no one can claim favouritism… Who said statistics wasn't useful!