Probability Theory: Trying to Make Sense of Random Processes

- The fundamental tool for statistically describing random phenomena is probability theory.
- Probability theory provides the mathematical foundation for formal statistical inference i.e. making inference about populations from sample information.

Probability Theory: Terminology

- A random process is one for which individual results cannot be predicted exactly, but for which long-term behaviour can be described.
- e.g. A type of random process we have considered so far is that of selecting a some person (or object) at random from a population and recording some measurement(s) about them.
- The key idea is to associate a likelihood, called a probability, to sets of possible outcomes from our random process. i.e. A probability refers to the likelihood of a particular outcome occurring.
- Typically expressed in terms of a fraction or proportion.

In probability theory, each repetition of an experiment is called a trial.

The results we might see from each trial are called outcomes. A subset of possible outcomes of a trial is called an event.

The set of all possible outcomes for an experiment is called the sample space for the experiment. i.e. it is the collection of unique non-overlapping outcomes.

Note that an event might correspond either to one outcome or to a set of possible outcomes.
Probability Theory: Terminology

e.g. Coin Tossing Experiment
- The repeated tossing of the coin is called trials (one toss is one trial).
- Observing either heads or tails is called an outcome.
- The set of all possible outcomes of tossing the coin any number of times is called the sample space.
- Any subset of the sample space is called an event, in this case an event would correspond to either of the two possible outcomes: heads or tails.

Probability Theory: Relative Frequency Interpretation of Probability

- Important concept, we are studying classical, or frequentist statistics in this course.
- The relative frequency probability of any outcome $A$ is the long-term proportion of times that $A$ is expected to occur when we observe a random process.
- The probability of an outcome/event denotes the relative frequency of occurrence of that outcome/event to be expected in the long run.
- The long run relative frequency of occurrence of outcome/event $A$ over trials of the experiment should approach $p(A)$.

Probability Theory: Relative Frequency Interpretation of Probability

- Relative frequency interpretation of probability refers to scenarios where we could imagine observing (counting) the frequencies of outcomes over many repetitions of the same situation or experiment.
- In other words we can observe long-run behaviour.
- Therefore, the probability of a given outcome is the proportion of times it would be expected to be observed in the long-run.
- Aside: by contrast, what we term the personal probability of an outcome is the degree to which a given individual believes an outcome will happen. Sometimes data from similar events in the past and other knowledge (e.g. expert knowledge) are incorporated when determining personal probabilities. Also called subjective probability.

Relative frequency probabilities can be determined by either of these methods:
- Making an assumption about the physical world and using it to define relative frequencies. e.g. coins are fair/balanced.
- Observing relative frequencies of outcomes over many repetitions of the same situation or measuring a representative sample and observing relative frequencies of possible outcomes.
Calculating Probabilities: Some Rules

- Let's denote the probability of event $A$ by $p(A)$.
- Probabilities for the outcomes of a process must all be between 0 and 1 and must add up to 1.
- If the outcomes in the sample space of a random process are all equally likely then, for a single outcome $A$, 
  
  $$p(A) = \frac{1}{\text{number of outcomes}}.$$  

Probability Examples

- Tossing a coin or picking cards from a deck.

Complementary Events

- An event can only either happen or not happen.
- Events that are the opposites of one another are called complementary events.
- Common notation: event $A^c$ is the complement of event $A$, or event $\bar{A}$ is the complement of event $A$.
- Since the sum of all possible events in the sample space must sum to one, we have: $p(A) = 1 - p(A^c)$ or $p(A) + p(A^c) = 1$.
- Notice that two complementary events make up the whole sample space together.

Mutually Exclusive Events

- (In contrast to complementary events,) two non-overlapping events that do not make up the whole sample space are called mutually exclusive events.
Independent and Dependent Events

- Two events are **independent** if knowing whether or not one has occurred has no influence on the probability of the other occurring.

- Two events are **dependent** if knowing whether or not one has occurred does influence the probability of the other occurring.

Conditional Probability

- $p(A|B)$ denotes the **conditional probability** of event $A$ occurring if event $B$ has already occurred.
- This may be conditional in time or might be different attributes of the same person or object.
- Conditional probabilities can be calculated using
  \[ p(A|B) = \frac{p(A \text{ and } B)}{p(B)}. \]
- The above formula can be rearranged to give the probability that both of events $A$ and $B$ occur:
  \[ p(A \text{ and } B) = p(A|B)p(B). \]

Addition Rule for Probabilities

- If events $A$ and $B$ are **independent** then
  \[ p(A|B) = p(A) \quad \text{and} \quad p(B|A) = p(B). \]
- Given the formula:
  \[ p(A \text{ and } B) = p(A|B)p(B). \]
- If events $A$ and $B$ are independent then the probability that both of events $A$ and $B$ occur is
  \[ p(A \text{ and } B) = p(A)p(B). \]
- For more than two independent events:
  \[ p(A_1 \text{ and } A_2 \text{ and } \ldots \text{ and } A_n) = p(A_1)p(A_2) \ldots p(A_n). \]
Common Misconceptions about Probability: The Gambler’s fallacy or Monte Carlo Fallacy

- The false belief that if deviations from expected behaviour are observed in repeated independent trials of some random process, future deviations in the opposite direction are therefore more likely.
- e.g. if a fair coin is tossed repeatedly and tails comes up a larger number of times than is expected, someone may incorrectly believe that this means that heads is more likely in future tosses.
- e.g. Casino gambler playing the big wheel and believing they can come up with a “system” to predict what will fall next - the house always wins!

Common Misconceptions about Probability: A Specific Person vs a Random Person

- e.g. What does a divorce rate of around 50% mean for your marriage?
- Do your grandparents have a 50% chance of getting divorced after all these years?
- This probability refers to the long-run likelihood.
- We might more accurately say, at the start of a randomly chosen union, the probability of it ending in divorce is 0.5.

Common Misconceptions about Probability: Coincidence

- Unlikely coincidences may not be as improbable as we would expect.
- e.g. matching birthdays, random shuffles, psychic readings!

Interpreting Probabilistic Results from Medical Diagnostic Testing

There is a tradeoff for a diagnostic test having a high sensitivity, the probability of correctly diagnosing a person with the disease, and a high specificity, the probability of correctly diagnosing a person without the disease. e.g. For a given HIV test we have

\[ p(\text{test positive} \mid \text{HIV}) = 0.98 \]

and

\[ p(\text{test negative} \mid \text{no HIV}) = 0.926 \]

What is \[ p(\text{HIV} \mid \text{test positive}) \]? Assuming that, in Australia
\[ p(\text{HIV}) = \frac{15000}{19200000} = 0.00078125. \]
Interpreting Probabilistic Results from Medical Diagnostic Testing

We have just used what is known as Baye’s Rule:

\[ p(A|B) = \frac{p(A \text{ and } B)}{p(B|A)p(A) + p(B|A^c)p(A^c)} \]

The Odds of an Event Happening

- Another common way of expressing the probabilities of two mutually exclusive events is in terms of odds. The odds is the ratio of occurrence to non-occurrence.
- Commonly quoted in betting.
- If the probability of an event/outcome is given by \( p \), then the odds in favour of that event are \( p \) to \((1 - p)\) (or \( p : 1 - p \)).
- Conversely, the odds against that event are \((1 - p)\) to \( p \).

The Odds of an Event Happening

- Equivalently, if the probability of an event/outcome is given by \( p \), we might say that the odds of it occurring are \( \frac{p}{1 - p} \).
- If the odds are less than 1 this means the event occurs less than half the time.
- If the odds are equal to 1 this means the event occurs half the time.
- If the odds are more than 1 this means the event occurs more than half the time.

Note, if we know the odds in favour of an event, the we can use that to work out the probability of it occurring via the inverse relationship:

\[ p = \frac{odds}{1 + odds} \]

Odds Ratios

An alternative way of analyzing two groups in terms of how likely some outcome is to occur is through an odds ratio. Odds ratios are more commonly quoted in medical research, as opposed to odds.

The odds ratio is simply the ratio of the odds in two different categories of an explanatory variable.

\[ \text{Odds ratio} = \frac{\text{Odds in category 1}}{\text{Odds in category 2}} \]
Nicotine Inhalers Example: Odds of Reduction in Smoking

The odds for a reduction in the nicotine group are

\[
\frac{0.26}{0.74} = 0.3514 \text{ to } 1,
\]

while in the placebo group the odds are

\[
\frac{0.09}{0.91} = 0.0989 \text{ to } 1,
\]

This gives an odds ratio of

\[
\text{OR} = \frac{0.3514}{0.0989} = 3.55.
\]

That is, the odds of sustaining a reduction in smoking after 4 months are 3.55 times higher if someone is using a nicotine inhaler.