Psychology Research Methods (Introductory) – Week 1

Part 1

Population vs Sample

- Population: "the entire collection of events in which you are interested" (Howell, p. 2)
- All individuals that you are interested in studying
- Sample: A subset of individuals selected from the entire population

If you're interested in doing a study comparing males and females on anxiety levels, the entire population would be all males and all females.

But because we can't access the entire population for a study, we recruit a sample from that population.

Research questions are generally phrased at the population level, but the studies are normally just conducted on a sample drawn from that population.

Example

- A psychologist wants to determine whether a motivational program is effective in improving performance in Australian office workers.
- Population? Every Australian office worker
- Problem – not feasible, and not sure if it works yet
- Solution – select a **sample** from the population
- Sample – e.g., 100 workers from each state

The view then would be to infer the effects of the motivational program on the sample to the population; if it works in this particular sample, then we have evidence that it might be effective at the population level.

Parameter vs Statistic

- **Parameter:** characteristic of the **p**opulation
- Example: average income of all Australian office workers
- **Statistic:** characteristic of the **s**ample
- Example: measure the age of the 100 depressed patients

A parameter is any value that we obtain that is characteristic of the population; any score that we get from a population. Census data allows us to access the whole population.

If we were interested in studying all university students but we could only get the 200 people that come to the first seminar to do a survey, and we got the average score from that sample then the average would be a statistic.

Two main types of statistics

1. **Descriptive statistics:** simply used to *describe* the data
    - Summarise the data
    - Averages and ranges the scores etc.
    - Makes the data more 'manageable'

2. **Inferential statistics:** used when we want to answer research questions
    - Allow us to make generalisations from the sample to the population
    - *Infer* the behaviour of all office workers based on the data we collect from the sample of 100 office workers.

Descriptive statistics are simply used to describe the data; summarise and put the data into a simpler form that we can more easily analyse.

Assuming we're studying differences between males and females on levels of anxiety in the population, because we cannot get the entire population of males and females into our study, we recruit a sample of 100 males and 100 females, and we give them an anxiety questionnaire. Because we've 200 people in the study, we'll get 200 anxiety questionnaire scores. Because this a lot of data of data to deal with, we would need a way of simplifying that data into a form that it is easier understand and work with; it makes the data more manageable. So we could get the average anxiety score for males and the average anxiety score for females, and so instead of having 200 scores, we've now got just 2.

The mean or average is a descriptive statistic; it describes the data that we have. We could also get the range of scores for descriptive statistics; the average anxiety score for males was 50 and they ranged from say 20 to 80.

We cannot infer anything from descriptive statistics; we cannot use them for anything other than simplifying and summarising the data. Inferential statistics on the other hand are used when we want to answer our research questions. If you had a research question that was looking at differences between males and females on anxiety, and you have a prediction that based on reading of research or clinical knowledge that males are going to have higher anxiety levels than females, you would need an inferential statistic to enable you to provide some support for that hypothesis or that prediction. The descriptive statistic simply gives you data; the males had an average score of 50 and the females a 40. So the males would be higher, but the question is could that difference just be due to chance? How much of a difference would you need to see before you can say that is a meaningful difference between males and females and that your prediction has been supported? Because it could be possible that for whatever reason you happened to sample males who were more anxious and so it is not reflected in the population. So inferential statistics allow us to answer those research questions; they give us a way of determining whether the differences we've seen in our data are actually significant.

Another example would be a case where you're studying a treatment program and you're looking at treating at depression. You administer this program and you find that the depression scores of the people who were in the program reduced by 20 points, from before to after the program; the treatment program has worked in some way. But how much of a reduction would we need to see before we can say that this is meaningful difference/significant change in depression? Could the change just have been a chance finding? So we need a way to rule out chance as an explanation for our data, and that's what inferential statistics allow us to do. The decrease in 20 points is just a descriptive statistic; it's just describing what happened, but the inferential statistics allow us to come to a verdict on whether it is a significant reduction in depression. In other words, they allow us to make generalisations from the sample to the population.

Inferential statistics use descriptive statistics to enable us to answer our research questions.

Sampling Error

- Statistics are used as estimates of the actual population parameters
- They are not perfect values – there is some error in the estimation
- Sampling error: the difference between the sample statistic and the corresponding population parameter

When doing studies on samples, one would hope they would be representative of the population in some way. For example, hoping a study sampling 100 university students is reflective of all university students.

But the fact about statistics is that they are estimates of the actual population parameters; the data you obtain from the sample is representative of the actual population parameters only to a certain extent.

The problem with statistics is that it only accounts for a sample of a population and not the entire population, and so they are often prone to error; there's going to be some error in the estimation of the population parameter based on the sample statistic. The extent of this difference is known as the sampling error.

Research methods account for sampling error whenever samples are studied, and therefore accept that they are not likely to truly reflective of what is going on at the population level.

Part 2

Variables

- Before we can use statistics we need to have data
- Data is obtained by measuring specific **variables**
- A variable is anything that can take on different values
  - Age
  - Self-esteem
  - Anxiety
  - Mood
  - IQ
  - Gender

Before we can get a descriptive statistic, we need to have data; data is obtained by measuring specific variables. A variable is anything that can take on different values.

Types of Variables

- Discrete vs continuous variables
- Discrete variables have only a limited number of values
  - e.g., Gender (dichotomous variable), categories
- Continuous variables can take on many different values
  - e.g., Age, IQ score, height, reaction time, etc.

A discrete variable is the most basic form of a variable; can only take on a limited number of values. Gender, for example is a discrete variable; either male or female. Also known as a categorical variable.

Continuous variables can take on many different values; for example, age is a continuous variable ranging from 0 to say 80. Anything that is measured along a continuum is called a continuous variable.

Discrete and Continuous Variables

- Often depends on the way you set up the measurement
- Anxiety levels of HPS771 students
  - Discrete or continuous?
- Generally continuous – score on an anxiety questionnaire
- But, often people will be categorised into groups:
  - High anxiety, Average anxiety, Low anxiety
  - This is a discrete variable

Sometimes it can be difficult to know if a variable is discrete or continuous; often depends on the way the measurement is set up in the study. How does the research choose to use that variable in the data analysis?

The anxiety levels of HPS771 students could be either discrete or continuous? But generally, we could think of anxiety as a continuous variable; more often than not we measure anxiety by a score on an anxiety questionnaire. But it could be discrete if values like 'low', 'average' and 'high' are used.

Measurement Data vs Categorical Data

- When we are dealing with continuous variables we have measurement data
  - Scores have been *measured* along a continuum
- When we are dealing with discrete variables we have **categorical data**
  - People are in *categories* (e.g., male/female; high anxiety/low anxiety)
- The type of data we collect influences the type of statistical approach we use
- Measurement data are usually summarised using means (averages)
  - Mean age of HPS201 students is 19.8 (made up)
  - Mean height of AFL footballers is 188cm
- Categorical data are usually summarised using percentages
  - 25% of the sample were female
  - 12% of students have black hair, 50% have brown hair, 30% have blonde hair, 8% have red hair
  - 23% of students reported low anxiety, 50% reported moderate anxiety, 27% reported high anxiety

Whether or not we have discrete or continuous variables has implications for the type of data that we can obtain from those variables. When we are dealing with continuous variables, we have what is called measurement data; scores have been measured along a continuum. But when we are dealing with discrete variables, we have what is called categorical data; our variable is in categories.

The type of data we collect, whether we get measurement data or categorical data has direct ramifications for the types of statistics we can obtain; it influences the type of statistical approach we can use.

Normally, we can do sophisticated data analyses' with measurement data. The most common type of analysis we can do on are averages; we can get the average score if we are obtaining people's IQ scores.

With categorical data, what we normally look at is percentages or proportions; we can only look at the number of people in each category and so we cannot get any averages.

Discrete Variables → Categorical Data → Percentages, frequencies

Continuous Variables → Measurement Data → Means, *Variance*, *Standard deviation*

Independent and dependent variables

- Independent variable (IV):
  - The IV is controlled by the researcher
  - For example, group membership – the researcher assigns participants to either high anxiety or low anxiety groups
  - Researcher randomly allocates people to either a treatment group or a control group
  - The IV 'causes' differences/changes in the dependent variable
- Dependent variable (DV):
  - The actual measured data
  - The variable that is observed for differences/changes

- For example, levels of depression in control vs treatment groups

The independent variable is controlled by the researcher; something that is manipulated by the researcher. Group membership is an example of an independent variable; based on some score, the researcher might assign people to either high anxiety or low anxiety groups, so they are influencing that variable. Another example is in looking at a treatment study and the researcher might allocate people to either a treatment group or a control group, and that is being controlled by the researcher.

The independent variable is the variable that is 'causing' the differences or changes in some other variable, the other variable being the dependent variable.

The dependent variable is actual measured data i.e. the data that we are interested in; the data that we are looking at for changes or differences. The dependent variable depends on the independent variable.

E.g., Differences between males and females on levels of anxiety; the independent variable would be the gender because in proposing a research question that males would have higher anxiety levels than females, there is the implication that gender is influencing anxiety in some way. Whether you are male, or female is going to lead to a difference in levels of anxiety, which is the dependent variable.

E.g., Control vs treatment groups for depression treatments; implying the research question predicts the treatment should reduce depression levels, group membership would be the independent variable purely because whether the proposed treatment is received (as opposed to a standard treatment) directly affects the reduction in depression levels, which is dependent variable.

Part 3

Measurement Scales

- We measure variables using a variety of measurement scales
- The type of measurement scale depends on the data we are dealing with
- There are 4 main types of measurement scales
- Nominal, Ordinal, Interval, Ratio

Nominal Scales

- Nominal scales are simply categories with different names
- There is no underlying scale and **no ordering**
  - Christian, Muslim, Jewish, Other. (1, 2, 3, 4)
  - Can't say that one category is larger/higher/more than another
  - The values are just labels for the different categories
  - E.g. Hair Colour:
    1. Brown
    2. Black
    3. Blonde
    4. Red
    5. Other
  - The numbers are for convenience only – no statistical operations can be performed on these numbers

The nominal scale has no order associated with it; the scale is simply arbitrary. Religion is a good example of a variable that is measured on a nominal scale. The values associated with each religion are arbitrary; 1 for Christian, 2 for Muslim and so on do not amount to one being greater than the other; no underlying scale. The values are for convenience.

Ordinal Scales

- Ordinal scales are categories with different names **AND** are organised into an ordered sequence
- Example: Degree of illness – None, Mild, Moderate, Severe
- This allows us to determine the direction of the difference
  - We can say that severe is greater than moderate; and moderate is greater than mild
  - But distances between the categories is unknown

Ordinal scales are ordered categories; they are still categories, but the categories are in an ordered sequence. An example of that is degree of illness; you have none, mild, moderate or severe, which the research might categorise people into. We can see that as we go up the scale from none to severe, there is progressively more illness.

This allows us to determine the direction of the difference across the categories. We can now say that being in the category of severe is greater than being in the category of moderate, and moderate greater than being mild.

A characteristic of the ordinal scale though, is that differences between the categories are unknown; we do not know if the differences between the categories of none, mild, moderate and severe are the same. Is the difference from moderate to severe illness the same as the difference from mild to moderate? Unknown.

Interval Scales

- Equal distances between points on the scale
- Many more points than on an ordinal scale (continuous data)
- Temperature (F, C) is measured on an interval scale
- Distance between 10 – 20 degrees is exactly the same as the distance between 50 – 60 degrees
- But, no true zero point
  - 0 degrees does not mean an absence of temperature

Generally, the interval scale has many more points on it than an ordinal scale; ordinal scales are generally in terms of categories, but an interval scale is often a continuously scored variable.

The fact that distances between points can be assessed characterises an interval scale. Another characteristic of the interval scale is that it as no true zero point. 0 degrees in terms of temperature does not equate to an absence of temperature; merely another point on the scale.

Ratio Scales

- Equal distances between points on the scale *and* a true zero point
- Time, length, age etc.
- Such variables have true zero points
- Deciding on the measurement scale is not always easy
- Questionnaire measure of anxiety (0 – 100)?
  - Nominal? Ordinal? Interval? Ratio?

On a ratio scale however, 0 means nothing; none of the construct being measured. Variables such as time, length and age have a true starting point. When we have a true zero point, we can talk in terms of ratios.

Is a questionnaire measure of anxiety from 0 – 100 either nominal, ordinal, interval or ratio?

If it was nominal, then it would be simply categories with no ordered sequence, and we know that on this questionnaire from 0 – 100, as we go up the scale we are getting more anxious. So it cannot be nominal; not an unordered category like gender, hair colour or religion.

It could be ordinal; as we are going up the scale, we are getting more of the variable we are measuring i.e. anxiety.

It could also be interval because it is measured continuously, which is one of the characteristics of an interval scale. But the question there would be, do we have equal distances on the scale? Is the difference between 80 – 100 on this questionnaire, the same as the distance between 20 – 40? It is a difference of 20 points, but does it mean the same for the variable we are measuring, i.e. anxiety. We might think that we would need a lot more anxiety to get from 80 – 100 than we would from 20 – 40. But that is probably unknown on these types of questionnaires, so we would be thinking more of an ordinal scale.

Is it ratio? Probably not. Does the 0 mean no anxiety for the person that is being measured? It is unlikely that the 0 means absolutely no anxiety.

Summary of Measurement Scales (Heffner, 2004)

Type of Scale	Scale Qualities	Example
Nominal	Categories None	Gender Hair Colour Religion
Ordinal	Ordered Categories Magnitude	Small, Medium, Large None, Moderate, Severe
Interval	Magnitude Equal Intervals	Temperature
Ratio	Magnitude Equal Intervals True Zero Point	Age Time Weight & Height

Data Description and Exploration

- Organisation of data
  - Summarise and simplify information

1. Frequency Distributions
2. Histograms
3. Stem-and-leaf displays

In taking an initial look at the data received from the study, because quite often we are dealing with particularly large sample sizes (e.g. online questionnaires), it is important to summarise and simplify that information. Instead of looking at say 2000 numbers, we can take a snapshot of those values and organise the data in a way that makes some sense to us and from which we can make some conclusions from, as an initial step in our data analysis.

We can do this through frequency distributions, how to convert that frequency distribution into a histogram and finally a less-common but still relevant approach seen in steam-and-leaf displays. These three techniques all provide us with exactly the same information; shows us a snapshot of our data in the first instance.

Frequency Distributions

- Frequency = how often each score appears in the set of data
- Example – Self-esteem scores of 16 people:

5 1 3 4 5 6 2 5 5 4 7 4 4 8 6 3

Frequency distribution of the self-esteem data

Self-esteem Score	Frequency
0	0
1	1
2	1
3	2
4	4
5	4
6	2
7	1
8	1
9	0
10	0

Each of the aforementioned techniques deal in terms of frequencies; frequency is simply how often each score appears in the set of data. For example, if we are measuring anxiety via a questionnaire in which a person is scored from 0 – 10, and 200 people are told to complete it which means 200 scores, a way we could summarise that information is to construct what is called a frequency distribution. And that would tell us how often each score appeared in the set of data; how often did people score a score of 1 to how often did they score a score of 10? A frequency distribution that can thereby indicate whether most people were scoring high, medium or low.

Looking at the frequency distribution of the self-esteem data, we can see that most people were scoring on the midpart of the scale and lesser at the extremes.

Histograms

- Graphical illustrations of frequency distributions

In the case of a histogram, it is exactly the same information as a frequency distribution but now it is being graphically represented; plotted onto a histogram.

Stem-and-leaf Displays

- Can summarise complex data in a simple way

Raw data is compiled into a stem-and-leaf display. Each of the stems represent a respective sets of numbers, whether that be in tens or units (e.g. 0 would represent 0-9, 1 would represent 10-19 etc.) and the leaves represent every value within said respective set (e.g. 19 19 would be represented in the 1 stem as 9 9).

Stems are also known as leading digits and the leaves as ending digits.

Types of Distributions

- Normal distribution

  - Most people score in the middle of the scale; scores become less frequent in the extremes
- Bimodal distribution
  - At least two peaks (bimodal or multimodal)

A normal distribution is where we see most people scoring in the middle of the scale and scores less frequent in the extremes. Distributions can vary around what is called normality to a variety of different degrees but generally, when most scores are in the middle and fewer in the extremes, that is called a normal distribution.

The mode is the most commonly occurring score in a set of data. When you have a bimodal or multimodal distribution, you have more than one peak in your distribution at different points on the scale. When you have two peaks, then you have an odd distribution.

Types of Distributions – Skewness

- Negatively skewed distributions
  - More high scores (scores bunched up at the high end of the scale)
- Positively skewed distributions
  - More low scores (scores bunched up at the low end of the scale)

Skewed distributions occur when scores are bunched up at one end of the scale. So instead of being what is called a normal distribution where most of the scores are in the middle of the scale, skewed distributions are when scores are bunched up at one end of the scale.

A negatively skewed distribution is when scores are bunched up at the high end of the scale; there is fewer in the middle and even fewer at the low end. A way to remember this is by thinking how the curve starts from the high end on the right i.e. the positive end and ends in the low end on the left i.e. the negative end.

On the other hand, a positively skewed distribution is when scores are bunched up at the low end of the scale; there is fewer in the middle and even fewer at the high end. Inversely, a way to remember this is by thinking how the curve starts from the low end on the left i.e. the negative end and ends in the high end on the right i.e. the positive end.

A lot of clinical measures end up positively skewed; if you are measuring alcohol-related problems for example, and you are giving that scale to a community sample, most people tend to score lower on that. This is because most clinically designed measures are designed to assess people who have problems with drinking; more appropriate for clinical setting than would be for community setting. Therefore, it is important to keep in mind who the study is sampling when looking at such distributions.

Types of Distributions – Kurtosis

- Kurtosis refers to how 'flat' or how 'peaked' the distribution appears
- Leptokurtic – the distribution is characterised by a high peak in the centre of the scale
- Platykurtic – the distribution is flatter – less scores in the centre

Kurtosis is another way we can characterise distributions. Kurtosis refers to how flat or peaked the distribution appears.

A leptokurtic distribution is when there is a high peak in the centre. There is markedly high peak in the centre of the distribution with no scores toward either extremes, indicating an aggregation of scores in the middle.

A platykurtic distribution on the other hand, is much flatter; all scores are being equally represented on the scale and there is no real distinction happening, i.e. 'plateau'.