PSYC2012: STATISTICS AND RESEARCH METHODS FOR PSYCHOLOGY

WEEK 1

ETHICS

There have been many recent scandals in psychology regarding research ethics. The most common ethical issue is what is known as 'data irregularities'.

ELEMENTS OF THE SCIENTIFIC METHOD

- Objectivity
- Confirmation
    - Replication / failure to replicate
- Self-correction
- Control (experimental control)

VARIABLES AND DESCRIPTIVE STATISTICS

VARIABLES

- Conceptual vs. operational (theory vs. practice)
- Categorical vs. continuous
- Manipulated vs. measured
    - Independent variable – what you manipulate
    - Dependent variable – what you measure

DESCRIPTIVE STATISTICS

Scales of measurement

Nominal/discrete/categorical

- Each category has its own identity
- Can only make same/different comparisons
- e.g. colour scale

Ordinal scale

- The levels of the scale can be ordered as higher, lower or equal
- e.g. race places - 1st, 2nd, 3rd

Interval scale

- Units are equally spaced but there is not inherent zero
- Cannot have an absence of the quantity
- e.g. temperature

Ratio scale

- Equally spaced units and an absolute zero
- e.g. time

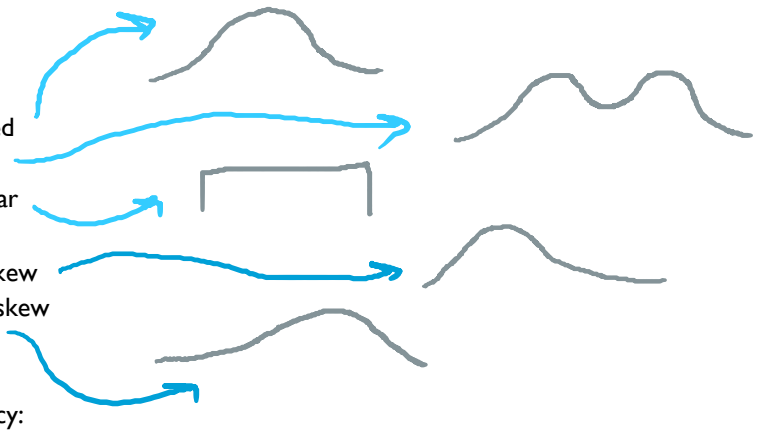
Describing data visually

Frequency distribution

Histogram – a graph of a frequency distribution

Shape of distribution

- Symmetrical
  - Bell-shaped
  - Bimodal
  - Rectangular
- Asymmetrical
  - Positive skew
  - Negative skew



Describing data numerically

Measures of central tendency:

- Mean – the average score
  i.e. $\mu = \frac{\sum X}{N}$
- Median – the middle score of ordered data
  i.e. $\frac{N+1}{2}$
- Mode – the most common score

VALIDITY

Validity is dependent on the scale of measurement.

MORE ON MEAN

The mean has the smallest deviation score from the mean i.e. the value of X-μ is the smallest. The sum of all deviation scores in a set of observations is equal to 0 at the mean.
i.e. $\Sigma(X - \mu) = 0$

Summation notation

$$\mu_X = \sum_{i=1}^{N} X_i = \frac{(X_1 - \mu_X) + (X_2 - \mu_X) + \cdots + (X_N - \mu_X)}{N}$$

VARIANCE (σ_X^2)

The variance is a measure of the dispersion of scores about the mean. Also known as 'the mean square deviation'.

$$\sigma_X^2 = \frac{\sum_{i=1}^{N}(X_i - \mu_X)^2}{N}$$

STANDARD DEVIATION (σ_X)

Two distributions can have the same variance, and variance is in squared units…so we use standard deviation as it is in the same unit as the original scale e.g. temperature. The SD is the average distance each score deviates from the mean. It is always positive, and it helps us compare within and across distributions.

$$\sigma_X = \sqrt{\sigma_X^2}$$

$$\sigma_X = \sqrt{\frac{\sum(X - \mu_X)^2}{N}}$$

WEEK 2

LINEAR TRANSFORMATIONS AND Z-SCORES

Z-SCORES

A z-score is a standard score which indicates relative position in a distribution. It translates to how many SDs from the mean an observation falls. It can be used to compare scores across different distributions. A z-score is most meaningful when distributions have the same shape (but most are normally distributed in psychology so this is ok!).

$$Z_i = \frac{X_i - \mu_X}{\sigma_X}$$

LINEAR TRANSFORMATIONS

Data can be transformed with predictable effects. Adding a constant does not affect variance or SD but multiplying data does. i.e. the score is transformed as follows: $Y_i = a + bX_i$ where X is the original score. Z-scores do not change for either (unless b is negative), it is a standardised score.

Descriptive	Transformation of data
SD	$\sigma_Y = \lvert b \rvert \sigma_X$
Variance	$\sigma_Y^2 = b^2\sigma_X^2$
Mean	$\mu_Y = a + b\mu_X$
Z-scores	$Z_Y = Z_X$

Example: adding 10 (a=10, b=1)

Original data:

X_i	μ_X	σ_X^2	σ_X	$X_i - \mu_X$	Z_i
19	15	16	4	4	1

Transformed data:

Y_i	μ_Y	σ_Y^2	σ_Y	$Y_i - \mu_Y$	Z_i
$a + X_i$	$a + b\mu_X$	$b^2\sigma_X^2$	$\lvert b \rvert \sigma_X$	-	-
29	25	16	4	4	1

THE NORMAL DISTRIBUTION, STANDARDISATION, AND Z-SCORES

Z-TRANSFORMATION

A z-score transformation is a linear transformation – it does not change the original distribution shape. A z distribution has the following properties:

- $\mu_Z = 0$
- $\sigma_Z^2 = 1$
- $\sigma_Z = 1$

STANDARD NORMAL CURVE

A standard normal curve has a mean of 0, a variance of 1, and a SD of 1. The curve area is 1 and the curve is symmetrical about the mean. It can be used to determine probability, percentage, proportion and percentiles.
50% of scores are above the mean i.e. $p(z > \mu) = 0.5$
50% of scores are below the mean i.e. $p(z < \mu) = 0.5$

QUALITATIVE RESEARCH

WHY USE QUALITATIVE RESEARCH?

The goal of qualitative research is to **understand the concept holistically and representationally**. It is the study of **process** not just the outcome, the examination of meaning not just causes, and looks at **subjective experience** in a **naturalistic setting**.

Two main viewpoints

1) **Positivist: realist** – content analysis, data display model
2) **Interpretivist: constructivist** – analysing meaning, cultural context, reflexivity, interpreting discourse

QUALITATIVE METHODS

- Observation
- Field studies
- In-depth interviews
- Multimethod research

QUALITATIVE THEORIES

1) **Grounded theory**
   - Aims to **construct theory** to explain social processes
   - Following the path of discovery
   - Constant comparative method – data matched for commonalities and emerging themes
   - Inductive and deductive
   - Theory generation
   - Explaining how/why, not what

2) **Ethnography**
   - Studies **patterns** between people and cultures
   - Bridges psychology, sociology and anthropology
   - Involves immersion into the setting
   - Inductive

3) **Phenomenology**
   - Reflection and study of **experiences, perceptions and emotions**
     - Describing and interpreting experience meaning
   - Thematic structural analysis
   - Explaining what, not how/why

WEEK 3

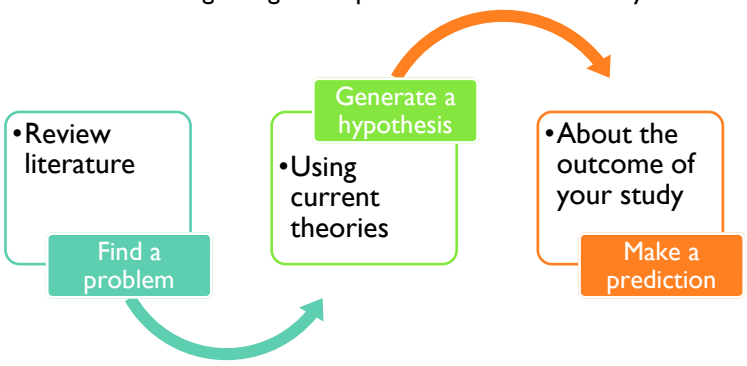
THE RESEARCH PROCESS

GOALS OF SCIENCE

Science aims to describe, predict, understand and determine causes of behaviour.

THE RESEARCH PROCESS

Theory – an overarching framework which organises and explains phenomena
Hypothesis – a tentative statement about a relationship that may or may not be true
Prediction – a specific statement regarding the expected outcome of a study



What makes a good research question

A good research question should be:

- Directional – the hypothesis should have a direction
- Testable
- Falsifiable

RESEARCH DESIGNS

Experimental

- Independent variables
  - At least 2 levels of IVs are needed
- Classification variables – pre-existing non-manipulated groups e.g. age, gender
  - **Quasi-experimental** design
  - Issue of internal validity

Non-experimental

- **Archival data**
  - Analysis of existing information

- **Observational data**
  - **Case studies**
    - Generate hypotheses
    - Study rare phenomena
    - Challenging accepted theories
  - **Naturalistic observation**
    - When subjects are observed in their day-to-day life without notice
    - High external validity, low control
    - Reduces reactivity (a bias caused when people know they are being observed)
  - **Participant-observer studies**
    - The researcher becomes part of the group being studied
    - A type of field study

- **Correlational designs**
  - Measuring the relation between two variables
  - Can conclude the **magnitude** (strength) and **direction** of a relationship
    - Directionality issue – does A cause B or B cause A?
    - Third variable issue
  - **Correlation ≠ causation**
  - There are three types of correlational designs…
    - **Single-strata**
      - But, issue of cohort effect
    - **Cross-sectional**
      - Better but, issue of non-random selection
    - **Longitudinal**
      - Better again but, issue of cohort effect and history of participants

VARIABLES

Dependent variables

When selecting the dependent variable, it is important to avoid **floor** (too difficult) or **ceiling** (too easy) effects. The DV can be measured by number, rate, frequency, degree, duration etc.

Extraneous variables

Extraneous variables are **any variable (other than the IV) which could influence the DV**, i.e. unintentional influences on outcome.

Nuisance variables (noise)

- Make it more difficult to find a difference between groups
- Increase variability within groups

Confounding variables

- Make groups look more similar or more different than they actually are
- INVALIDATES the experiment :(

CONTROL

Increasing control in an experiment is achieved through:

- Randomisation
  - Random selection
  - Random assignment
- Having a sample size n > 1 in each condition
- Constant levels / systematic varying of the IV
- Having a control group
- Doing a randomised control trail (RCT) – often used in clinical research

Random assignment

A 'between-subjects' design, where participants are only exposed to one level of the IV.

Non-random assignment

A 'within-subjects' design, where participants are exposed to every level of the IV. Pairs are often matched to ensure equal 'types' of people in each condition group. This is sometimes achieved by using **natural sets** e.g. twins (matching in all characteristics). This design is **quasi-experimental**.

INFLUENCING THE DATA

Participants