Stats & Research

How do we acquire knowledge?

- Tenacity, authority, experience, reason and logic

Qualitative and quantitative

- *Quantitative research question*
  - Effectiveness of intervention?
  - Association between variables?
- *Qualitative research question*
  - Meaning of phenomenon or experience?
- *Mixed Research Methods Question*
  - Generalise findings to population and understand meaning?
  - Qualitative and quantitative
- Assumptions
  - Positivism → quantitative
  - Constructivist → qualitative

Positivist assumptions

- Causes determine effects
  - Experiments examine cause and effect (probability)
- Reductionist
  - Hypotheses (directional or non-directional)
- World works via orderly laws
  - Scientific method verifies and tests theory
- Premium on objectivity
  - Validity and reliability of methods

Constructivist assumptions

- People construct meaning as they interact with the world
  - Open-ended questions allow expression of views
- View world through historical and cultural lenses
  - Understand context and interpret it → acknowledge researcher's experience
- Meaning-making is social
  - Data collection/analysis inductive → based on data

Qualitative → e.g. case studies
Quantitative → e.g. ANOVA

Elements of scientific method

- Objectivity
- Confirmation
  - Replication
  - Failure to replicate
- Self-correction
- Control

Confirmation bias

- Search for info consistent with view, ignore inconsistent
- Impact on researcher
  - Literature reviews → show studies that confirm but not the ones that failed

Goals of science

- Describe behaviour
- Predict "
- Determine causes "
- Understand and explain "
- Apply knowledge to solve problems

Research process

- Find problem and review literature

- Role of theory
    - Theories used to make hypothesis
    - Test hypothesis to develop theory
- Design and conduct experiment
- Analyse and interpret data
- Share results → reports, presentations

Ethics
- Scandals in psych regarding research ethics
- Data irregularities
    - Brian Wansink
    - Michael LaCour
- External data sleuth: Uri Simonsohn
    - Search for patterns in data that are suspicious

Stapel debacle
- Diederik Stapel
    - 2011: suspended for fabricating data and other scientific misconduct
    - 55 publications affected

Variables and descriptive stats

Stats
- **Organise**
    - Stats allow us to organise data to make it easier to understand
- **Describe**
    - Stats allow us to summarise data in sensible ways
- **Analyse**
    - Allow us to make inferences about patterns of behaviour e.g. compare stats before and after teaching

Variables
- **What are the variables?**
    - Conceptual (e.g. self-esteem) vs operational (make it measurable e.g. self-esteem scale) distinction
- **What type of variables?**
    - Categorical vs continuous
    - Manipulated vs measured

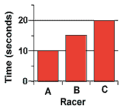
Independent variable (IV) – what you manipulate
Dependent variable (DV) – what you measure

Nominal scale
- Each category has own unique identity e.g. categorical data
- Can only make same-different comparisons

Ordinal scale
- Levels of scale can be ordered according to whether they are higher/lower/same
- Can make same-different and greater than-less than

Interval scale
- Units equally spaced over scale but no absolute zero (i.e. no absence of the quantity)
- Can make same-different, greater than-less than, equal intervals

Ratio scale
- Units equally spaced over scale and there is absolute zero
- Same-diff, greater-less, equal intervals, ratios



Scales of measurement

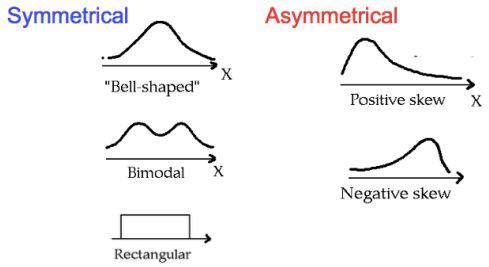
Scale	Features
Nominal	• Same vs. Different
Ordinal	• Same vs. Different • Greater than / Less than
Interval	• Same vs. Different • Greater than / Less than • Equal intervals
Ratio	• Same vs. Different • Greater than / Less than • Equal intervals • Absolute zero (allows ratios to be formed)

Qualitative
Categorical
Discrete

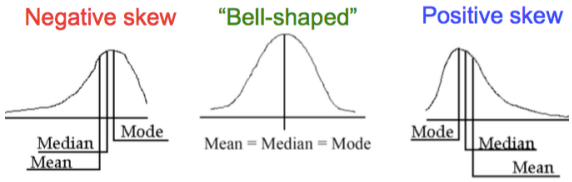
Quantitativ
Continuous

Describing data

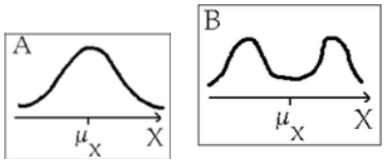
- **Frequency distribution –** order scores and tabulate frequencies
- **Distributions –** diff shapes



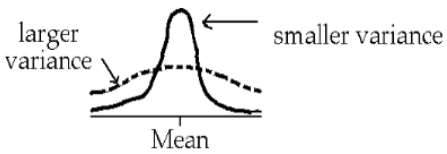
- **Histogram –** graph of frequency distribution
- **Measures of central tendency**
  - **Mode** – score with highest frequency
    - May not be unique (e.g. bimodal), may not exist (e.g. rectangular)
  - **Median** – score divides distribution into two roughly equal parts
    - For ordered scores, (N+1)/2
    - Not affected by outliers but not sensitive to all scores
  - **Mean** – balance point of distribution
    - Sensitive to all scores



- **Deviation score –** how far away a score is from mean (X- μ)
  - The sum of all deviation scores in set of observations is 0
- But very different distributions can have same mean
  - A – less variance
  - B – more variance



- **Variance σ^2_x –** measure of dispersion of scores about the mean
  - All other things being equal, if one distribution has greater variance, it will be fatter and flatter



  - Because of individual differences, there is variance due to factors other than what we are interested in → additional factors are called 'error' variance – we try reduce this
- **Standard deviation σ_x –** average distance each score deviates from mean → square root of variance

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

  - Sd is always positive as it is a measure of 'distance'
  - Mean and sd help us compare both within and across distributions

Linear transformations and z-scores

Z-scores – standard score that indicates relative standing in a distribution

deviation score for the i^{th} person

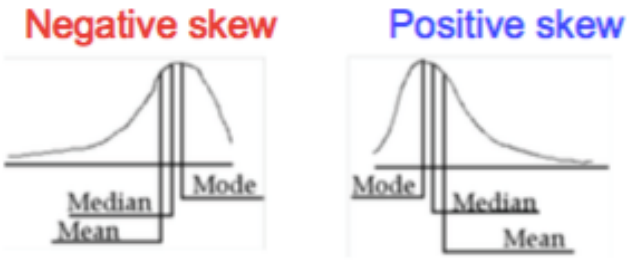
z-score for the i^{th} person $z_i = \frac{X_i - \mu_X}{\sigma_X}$

standard deviation

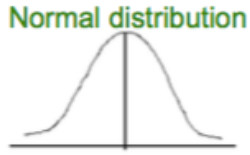
- Z-score is how many standard deviations from mean a given observation falls

If $X < \mu$, $z < 0$ If $X = \mu$, $z = 0$ If $X > \mu$, $z > 0$

- Comparisons of relative standing across distributions are more meaningful if distributions have same shape
- In positively skewed distribution, e.g. having score on mean does not indicate same sort of relative standing as being on the mean in a negatively skewed distribution



- In psych, we assume most variables we measure are *normally distributed* i.e. bell-shaped



Effects of transformation

- *Adding* a constant to each score does not affect variance BUT *multiplying* by a constant (other than +/- 1) does
- These are all *linear transformations*

Linear transformations

- Prevent acquiesce bias – change direction of scale in answer e.g. 1 is good and then 5 is good in next question
- Linear transformations are used for e.g. mm to cm, or temp C to F
- Linear transformation does not affect shape of distribution (normal → normal, skewed → skewed)
  - Mean is transformed same way X scores are

    $\mu_y = a + b\mu_x$
  - Sd only affected by b

    $\sigma_y = |b|\sigma_x$
  - Variance only affected by b

    $\sigma^2_y = b^2\sigma^2_x$
  - Z-score does not change

    $z_y = z_x$

The Normal Distribution

Z transformation