PSY2202 RESEARCH METHODS IN PSYCHOLOGY 2

EXAM

8th June, 8.30am

3 hours + 5 minutes reading

Closed book, calculator allowed

Multiple choice (90Q's) + short answers (5Q's)

50% of semester grade

Week 1, lecture 1 (p. 1 – 8) NO TUTORIAL

Introduction: Getting research started – Review of hypothesis testing, t-tests

- Independent variable (IV): A variable whose variation does not depend on that of another
- Dependent variable (DV): A variable whose value depends on another
- Behaviour can be varied due to different variables
- Theory leads to test which leads to predictions
- Taking what has been done and conducting our own research
- Conflicting findings – what may go against your theory – research that shows why different inconsistencies exist between findings.
- Evidence shapes theories
- Overlooked variables i.e. gender
- Always know how you intend to analyse data (i.e. what stats you're using)
- Unfolding preferences analysis – outlines preference probability
- Only use journals as books reflect the opinion of the author
- Journal articles reflect what has been done, the IV and DV, how it was conducted, theory, evidence, outcome, how it's measured
- What makes a good research question?

  1. Answerable (empirical question involves measurement)
- Be specific when talking about the IV and DV i.e. in class scenarios

  Scenario 1: Different health promotion campaigns (IV)

  Attitudes to condoms (DV)

  Scenario 2: Age in year/months (IV)

  IQ number measured on a specific scale (DV)
- Theory: An organised way of thinking about two or more variables and how they relate.

Week 2, lecture 2 (p. 9 – 11)

Review of hypothesis testing, t-tests

- **Parametric statistics** are the most commonly used. It assumes that the sample comes from a population that follows a probability distribution based on a fixed set of parameters. If the assumptions are violated, we can still read our data as long as it's clear.
- **Scales of measurement**

  1. Nominal scales – simple categorization

  2. Ordinal scales – some quantitative information

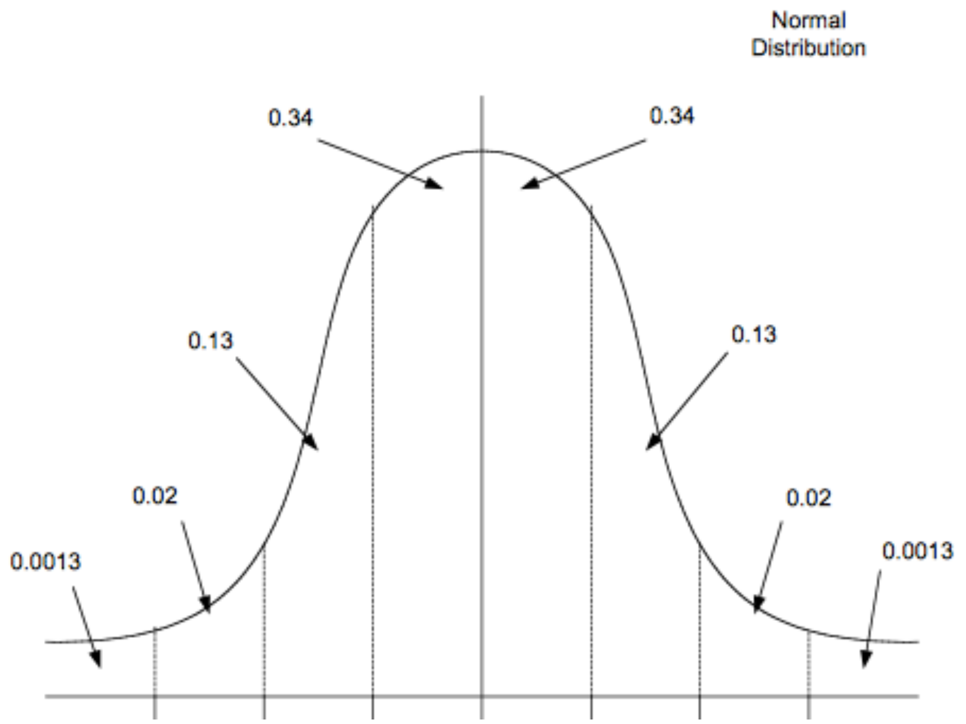
  3. Interval/ratio
- Descriptive statistics include standard deviation of the mean, mean, sum of all score etc
- Effect size – Is the variance of the groups significant?
- Inferential statistics is how we analyse the results
- **Finding the SD of the mean:**

Standard deviation of group 1 + standard deviation of group 2 \ 2

- **Finding the D score:**

D = mean of group 1 – mean of group 2 / the standard deviation

- Normal distribution characteristics: Mean of any set of score = 0, 68% of people fall with -1 and +1 of the standard deviation of the mean, the distribution is symmetrical, all of the scores fall under the distribution, but the distribution never ends.



- In sampling distribution, the standard deviation is equal to the standard error of the mean

Formula:

574 APPENDIX B STATISTICAL TABLES

(A) z	(B) Proportion in Body	(C) Proportion in Tail	(D) Proportion Between Mean and z	(A) z	(B) Proportion in Body	(C) Proportion in Tail	(D) Proportion Between Mean and z
2.50	.9938	.0062	.4938	2.95	.9984	.0016	.4984
2.51	.9940	.0060	.4940	2.96	.9985	.0015	.4985
2.52	.9941	.0059	.4941	2.97	.9985	.0015	.4985
2.53	.9943	.0057	.4943	2.98	.9986	.0014	.4986
2.54	.9945	.0055	.4945	2.99	.9986	.0014	.4986
2.55	.9946	.0054	.4946	3.00	.9987	.0013	.4987
2.56	.9948	.0052	.4948	3.01	.9987	.0013	.4987
2.57	.9949	.0051	.4949	3.02	.9987	.0013	.4987
2.58	.9951	.0049	[illegible]	3.03	.9988	.0012	.4988
2.59	.9952	.0048	[illegible]	3.04	.9988	.0012	.4988
2.60	.9953	.0047	.4953	3.05	.9989	.0011	.4989
2.61	.9955	.0045	.4955	3.06	.9989	.0011	.4989
2.62	.9956	.0044	.4956	3.07	.9989	.0011	.4989
2.63	.9957	.0043	.4957	3.08	.9990	.0010	.4990
2.64	.9959	.0041	.4959	3.09	.9990	.0010	.4990
2.65	.9960	.0040	.4960	3.10	.9990	.0010	.4990
[illegible]	[illegible]	[illegible]	[illegible]	3.11	.9991	.0009	.4991
2.67	.9962	.0038	.4962	3.12	.9991	.0009	.4991
[illegible]	[illegible]	[illegible]	[illegible]	3.13	.9991	.0009	.4991
2.69	.9964	.0036	.4964	3.14	.9992	.0008	.4992
2.70	.9965	.0035	.4965	3.15	.9992	.0008	.4992
2.71	.9966	.0034	.4966	3.16	.9992	.0008	.4992
2.72	.9967	.0033	.4967	3.17	.9992	.0008	.4992
2.73	.9968	.0032	.4968	3.18	.9993	.0007	.4993
2.74	.9969	.0031	.4969	3.19	.9993	.0007	.4993
2.75	.9970	.0030	.4970	3.20	.9993	.0007	.4993
2.76	.9971	.0029	.4971	3.21	.9993	.0007	.4993
2.77	.9972	.0028	.4972	3.22	.9994	.0006	.4994
2.78	.9973	.0027	.4973	3.23	.9994	.0006	.4994
2.79	.9974	.0026	.4974	3.24	.9994	.0006	.4994

- The population in your study is what YOU define it to be. i.e. everyone in the room wearing glasses, everyone in the room who lives NOR etc.
- **T-TEST:** Statistical examination of two population means. A t-test examines whether two samples are different and is commonly used when the variance of the two normal distributions are unknown and when an experiment uses a small sample size.
- **Three different kinds of T-tests:**

  1. Single samples – comparing a group score vs. a single number

  2. Related/dependent/correlated samples – E.g. one person tested at 2 different times (pre-score vs. post score) or two conditions, same sample.

  3. Unrelated/independent/uncorrelated samples
- Comparing a group mean vs. another group mean or one score at one time vs. another score at time 2 by the same person.

- **Z-TEST:** Statistical test used in inference which determines if the difference between a sample mean and the population mean is large enough to be statistically significant.
- **Z Scores**
  Zobs(erved): Z score that we calculate
  Zcrit(ical): Z score that is related to alpha (reject null hypothesis)
- **Z score formula**

Z = the mean of the sample – the mean of the population/the standard error of the mean
i.e. in class example, the probability of musicians having a mean IQ of 104 is less than 5% meaning 5% is the chance that I'm wrong. The independent variable is the musicians.

- **Cohen's (1988) Effect Size Guidelines:**

Effect size	D score	Percentage of overlap
Small	0.2	85%
Medium	0.5	67%
Large	0.8	53%

Difference between the mean of the two groups.

- **Sampling error** is whenever we take a sample from a population, we are bound to make an error in estimating the population mean (unless our sample is the population) i.e. the difference between a sample and a population is sampling error. – When you take two samples from one population, the statistics will differ from the statistics of the whole population and you have to make judgement whether this is representative of the population, how well you measure the population is sampling error.
- **Alpha value:** Arbitrary probability level which delineates the uncommon from the unlikely
- **Alpha** (typically taken to be) = 0.05
- If we obtain a sample mean, which has a probability of occurrence (e.g. less than alpha < 0.05), then we would claim that it is not consistent with sampling error. The result is then due to the manipulation of the independent variable.
- **Inference** is a type of reasoning, drawing a specific inference from specific instances to form a general conclusion.
- **Inductive inference** is when we take a sample and assume it infers something about the population. (Null hypothesis: The IV has no effect on DV).
- If we are wrong, we reject the null hypothesis and accept the alternative hypothesis.
- *Inductive reasoning isn't error free:*
  1. Our reasoning is inductive; it is not error-free
  2. Assume the null hypothesis is really true; An extreme sample from the population could provide a statistic falling in the region where we would (normally) reject the null hypothesis.

- **Type I error** is when we say there isn't a difference but there really is.
- **One tailed test** – Directional
- **Two tailed test** – Non-directional

Week 3, lecture 3 (p. 12)

Controlling extraneous variance (single factor between and within groups design) ANOVA: Between-groups design, multiple comparison procedures (Planned and Post Hoc comparisons)

- Between: The researcher distributes the effects of extraneous participant variables evenly across the experimental and control conditions
- Within: Effects of participant variables are balanced across conditions
- Confounding variables e.g. Motivation, IQ, temperature, learning spaces
- Maximising internal validity: Keep some things constant e.g. interventions, common characteristics
- Include a control group: Receives no intervention
- Placebo effect: when one pill doesn't have the drug you want to measure but the other group does
- Reactivity effects (Hawthorne effect): Participants change their behaviour because they know they're in a study
- Demand characteristics – when participants guess the hypothesis of the study and act accordingly
- Randomly assign people to groups
- Assess equivalence before the study using one or more pretests
- Within subject designs are more powerful – bigger effect with fewer participants
- Advantages: fewer subjects, error variance reduced
- Disadvantage within subject – order effects
- Sensitisation effects: Any type of reactivity in experimental situation – change in criterion, attitude etc
- One way to control order effects = counterbalancing: different groups of participants run under the experimental variable in a different order
- Latin squares: Counterbalance all treatment conditions so that all subject perform in each condition and each condition appears in each test order across subjects

Week 4, lecture 4 (p. 12)

ANOVA: Between groups design

- 3 way ANOVA = 3 Independent variables, 4 way ANOVA = 4 IV etc
- Saline means control condition

- Between-group variance is the variability between the groups (All possible combinations of variability within the groups)
- Within group variability is the variability within the groups.
- By doing many t-tests, we have to do three t-tests opposed to one comparison (F-test).
- Increases chances of type one error (increasing t-tests).
- Non pair comparisons: AB vs. CD – ABC vs. D – C vs. ABD etc
- Double sigma sign = summed over everything
- We rely on inferential statistics for more complicated designs (ANOVA)
- Analyses of variance – different variables and difference between those variables
- The larger the difference between the two means the more likely the groups are not equal
- Sampling distribution of F – all possible sample size from the same population, estimate population variance from each of the samples, calculate F for all possible combinations
- F critical = e.g. 2.87, if calculate value less than we accept null hypothesis
- Unsystematic effects: Random – motivation, ability, biorhythm
- Systematic effects: Factors we manipulate
- Scores differ within groups because of unsystematic effects
- Scores differ between groups because of both
- F statistic = between group variance/within group variance
- F statistic = unsystematic effects + systematic effects/unsystematic effects
- We can get F less than 1
- ANOVA summary source table *
- Effects we do not control are randomly distributed (unsystematic effects)
- Form of the distribution of random effects is a NORMAL curve
- Size of F is not a good indicator of size of effect
- Much bigger effect when change in denominator
- Total sums of squared – individual scores – group mean
- Between sums of squares + within sums of squares
- Eta square of .5 mean = 50% of variance in performance can be explained by effect of the drug
- Normality: Observations for each group are sampled from normally distributed populations
- Homogeneity of variance: The variances in the population from which the subjects are sampled are equal
- Group independence: Observations within each group are independent from those of any other
- Parametric tests – interval and ratio

- Non-parametric tests – ordinal and nominal

Week 5, lecture 5/6 (p. 12)

ANOVA: Between-groups design, multiple comparison procedures (Planned and Post Hoc comparisons)

- Alternative ways of examining pair wise comparisons
  - Independent comparisons – do not depend on the outcome of any other comparison
  - 3 groups, we want to make comparisons between all possible pairs of means, then there are 3 possible pairwise comparisons that can be made but only 2 are independent comparisons
- Comparison wise error rate = level of significance for each comparison
- Experiment wise error rate = probability of making at least one type I error for the set of all possible comparisons
- Post hoc – occurring or done after the event
- Planned comparisons
  - Do not require a significant overall F value
- Post hoc comparisons are performed in such a way to control for experiment wise error rates
- Within group variance is also known as error variance
- Divide SS by DF to get MS
- F is MS between / MS within
- Scheffes – paired comparisons and complex comparisons between groups of means
- Most versatile post hoc test, tends to be conservative
- Multiply each weight with mean
- J refers to number of groups
- Apriori = relating or denoting reasoning or knowledge that proceeds from theoretical deduction rather than observation or experience
- Comparisons are orthogonal when it equals 0
- Orthogonal: The results are independent to each other
- "The weighted average of group 1 and 2 equals …"
- SSc1 (sums of squares for comparison 1) = (16.48)(1) + (15.80)(-1)2 / 1(weight squared)/15 + -1(weight squared)/15 = 3.38
- If comparison is orthogonal, comparisons should equal between sums of squares
- Df for comparisons should also equal the between df
- MS/between group MS = F
- "Trend" – how a set of numbers look i.e. linear? Non-linear?
- Linear – variables that fall on a straight line

- Yerkes-dodson law/inverted u shape: optimal level of arousal (anxiety, stress etc) but there is some level of optimal arousal, any less results in poorer performance and any higher poorer performance
- Cubic goes up and down again
- Quadratic inverted u
- To do trend analysis properly – IV must be equally spaced (interval/ratio level)
- Relationship between IV and DV not always clear
- Relationships should be overdetermined

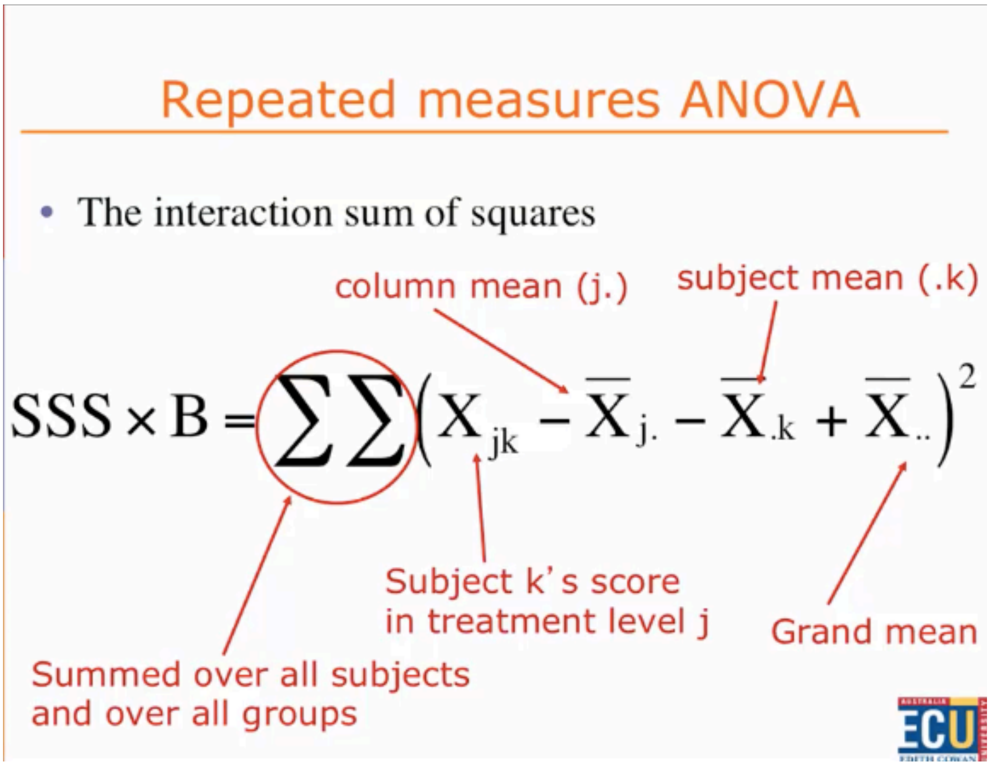
Week 7, lecture 7/8 (p .13)

ANOVA: Within-groups design/ Repeated Measures ANOVA

- Problems with between subject designs: lots of subjects, error variances
- Within groups variance: Measure of error variance
- Reduce error variance by:
  - Reducing number of subjects
  - Comparing subjects with themselves
- Within-subjects design: the same set of subjects are examined in all conditions within an experiment
  - Fewer subjects necessary
  - Error variance reduced
- Cell scores = Sums of squares (SS) x condition interaction
- People perform differently in different circumstances
- 3 sources of variance
  - Condition
  - Subjects
  - Condition x subjects
- No such thing as a within source of variability
- Subjects is a random factor
- Calculation F
  - F = MS condition/MS condition x subject
  - Compare F with MS between/MS within
- Three variance estimates (MSs)
  - The treatment (column) means
  - The subject (row) means
  - The treatment x subject interaction
- Sum of squares based on the subject means:

- N = subjects, first X = column mean, second X = grand mean

- Interaction sum of squares



- Interpretation: i.e. 32.6% of the variance in the experiment can be attributed to the treatment effect (the different conditions)
- Mean squared residual used as the error term
- Tukeys procedure
  - Figure out critical q at alpha = .05
  - There are 4 ordered means
  - Take into consideration the degrees of freedom
- Between subject factor = experimental x control group
- Pairwise comparions
  - Experimental before – exp after
  - Exp before – control before
  - Exp before – control after
- The more tests you do the higher the risk of type 1 error
- Advantage of nested design: increase in generality
  - Disadvantage: without an interaction term, too much info is sacrificed
- Friedman rank test = sum of ranks lowest score is most preferred

Friedman's rank test

$$\chi_F^2 = \frac{12}{Nk(k+1)} \sum_{i=1}^{k} R_i^2 - 3N(k+1)$$

χ^2_F represents the amount by which the current rankings differ from the null rankings.

R_i = the sum of the ranks for the ith condition
N = the number of subjects
k = the number of conditions (songs)

Week 9, lecture 9 (p. 13)

ANOVA: Factorial designs

- Designs which include more than one independent variables
- All of the levels of each IV are completely crossed with each other
- Examples of factorial designs:
  - Two-way factorial between-subject design (2 IV)
  - 4-way (4IV etc)
  - Between-subjects: different cells comprised of different individuals
- Factorial refers to the factorial combination of the levels of the 2 IVs
- E.g. Gender: Male & Female, Handedness: right & left (different levels of IV)
- Factorial design provides more efficiency in terms of having more than one IV being analysed at one time
- Allows us to understand interactions – ways in which different levels of each IV interacts with each other
- When talking about the main effect – is there a difference between the means?
- Benefit of a two-way ANOVA? Able to determine within one experiment the effect of two main effects
  - Much more efficient
  - Collecting data on wide range of people all at once
- Interactions: two IVs have interacted to create an effect over and above their individual effects
- Interaction = when the effect of one IV is different at different levels of another IV
- If the lines are not parallel then an interactions may exist
- If an interaction exists, main effects should be interpreted with caution
- Interpret interactions before main effects
- When there are no main effects of the two factors, they may still interact

- Total variance = within-group variance + between-subject variance
- Two way ANOVA
  - SS within + SS associated with factor a + SS associated with factor b + SS interactions of axb
  - Null hypothesis – no difference
  - Alternative – there is a difference
- Main effect = Mean squared divided by error term and get an F value
- "Solving one type of problem takes longer than solving another problem"
- i.e. significant main effect of cognitive style: "visualisers take longer to solve problems (overall) than verbalisers"
- Two way – F = MS effect / MS within
- Fixed IV's = variables where the levels chosen mean something for the issue being examined
- Interaction will not intrude on marginal means
  - Gender, marital status, problem type
- Levels can be collapsed into smaller number of groups (i.e. age 4 – 10, 11 – 20, 21 – 40 etc)
  - Socioeconomic status – lower, middle upper
- Systematic selection of levels (i.e. set size, 2, 4, 6, 8)
- Random IV's – levels of the IV are randomly selected (interaction will intrude on marginal means)
  - Schools in an education study
  - Set of words in a memory study
  - Order of questions in a survey
- Three ANOVA models
  - Fixed – both IV's are fixed
  - Random – both IV's are random
  - Mixed – one IV is fixed, one IV is random
- There are differences between IV's that have an effect on the size of the F ratio

Week 10, lecture 10 (p. 14)

Simple main effects

- Main point of interest: main effects – is there a difference between ... and …
- Ordinal interaction: lines that are parallel – start out the same then at different levels they go in different directions
- Disordinal: lines are crossed X
- Cell scores = Sums of squares (SS) x condition interaction
- Mean square = SS calculation/degrees of freedom

- A significant simple effect means for that factor it differs at different levels
- F (simple main effect) = MS i.e. reinforcement type at simple / MS error
- SS(reinforcement) + SS(reinforcement x problem)
- SS (simple) + SS (complex) = 30.40 + 68.13 = 98.53
- Df(reinforcement type) = 3 – 1 = 2
- Type 1 error rates: keeps error rates low
- Family-wise error rate
- Simple comparisons
- Error terms

Week 11, lecture 11

Correlation and multiple regressions

- Whenever you talk about a correlation you must describe the relationship in terms of strength and direction
  - “Weak negative correlation” “strong positive correlation”
  - Strong = two variables increase at the same time
  - Negative = one value goes up, one value goes down
- When you work out Pearson’s R you are working out the relationship between two variables
- Correlation provides a description of the relationship between the two variables i.e. early experience and reading test score
- No causality can be inferred
- A correlation is not a causal link
- Multivariate influences: variables that have influences on others

Correlation and regression

$$r_{xy} = \frac{\sum XY - \frac{\sum X \sum Y}{n}}{\sqrt{\left[\sum X^2 - \frac{\left(\sum X\right)^2}{n}\right]\left[\sum Y^2 - \frac{\left(\sum Y\right)^2}{n}\right]}}$$

- Sample size effects correlation coefficients if the sample size is very small this can affect the interpretation of r
- Outliers – extreme scores can have a large effect on the overall correlation
- Combining several samples – or different types of data can effect the correlation
- Restriction of range – i.e. if you only take only high or low results it wont be representative
- Skewness
- Criterion variable = variable predicted
- Predictor variable = variable used as basis for prediction

- B = how many units the criterion variables will change
- Regression = line of best fit
- Least squared analysis
- 1 – r squared or coefficient of nondetermination = the proportion of variance in Y due to other unknown factors not considered in the regression equation
- Residual = the difference between the observed value of the criterion variable and the value of the criterion variable predicted by the regression line
- Ordinary least squares regression = method of determining the best fit regression line which is the one that minimises the sum of the squared residuals

Week 12, lecture 12

Multivariate statistics

- Multivariate statistics aren't used a lot in psychology
- Univariate statistics – one dependent variable
- Multivariate statistics – more than one DV
- MANOVA = multivariate analysis of variance
- Alpha level problems = doing more tests increases our chance of finding significant results by chance
- One multivariate tests are preferable to many univariate tests
- Power of multivariate = comes from consideration of correlations between variables
- Problem with this
    - Alpha level (several tests need to be conducted)
    - Information loss (taking into account the relationships between all variables)
- Multiple regression is better – criterion and predictor variables
- Criterion variable = DV
- Predictor variable = IV
- Multiple r = correlation between Y and Y'
- B = regression weights
- Beta weights to standardise scores
    - Use standard scores for each of the predictor variables, instead of raw scores
- Sr2 = semi-partial correlations: the amount that the R2 would be reduced if a particular IV was deleted from the regression equation
- With multiple regression you can pick and choose data in your regression equations which is values – t values, overall F calculation, semi-partial correlations
- Path analysis = examines questions about the minimum number of relationships among variables and their directions

- Represented in diagram form
- Path analysis is a confirmatory method
- Factor analysis = performance on many variables explained by the operation of a small number of factors
- Factor analysis reveals
    - Factors underlying the scores/processes
    - Its up to the researcher to interpret what the factors represent
    - Patterns are examined to extract underlying factors
- Eigenvalues = importance of each of the factors
- Orthogonal rotation: rotation of axes in which you get perpendicular (independent/orthogonal) to each other
- Oblique rotation: Rotation of axes are not always going to be perpendicular to each other
- Exploratory factor analysis: researcher tries to describe and summarise data by grouping together variables that are correlated
    - Conducted in early stages of research
- Confirmatory factor analysis: tries to 'test' a hypothesis
    - Variables are carefully and specifically chosen to reveal underlying processes
- Problems with factors analysis:
    - Lacks objective criteria of interpretation adequacy
    - Infinite rotations available
    - Factor analysis often associated with sloppy research
- Cluster analysis: create groupings of variables where the members of a particular cluster correlate highly with each other
    - Clusters of subjects are produced where members of a particular cluster are more similar to each other i.e. correlate more highly with each other
- MANOVA: creates a new DV
    - A linear combination of measured DVs, combined so as to separate the groups as much as possible
- Factorial MANOVA: same study but including gender as a factor
- MANOVA advantages
    - Improved chance of discovering nature of effects
    - Protects against inflated type I error
    - MANOVA sometimes more powerful than separate ANOVAs
- Disadvantages of MANOVA
    - Complicated
    - Assumptions

- Ambiguity
- Greater power limited

Textbook notes

- Statistics are defined as facts and figures i.e. average income, crime rate, birth rate, baseball batting averages etc
- They condense large amounts of information into simple figures
- Statistics are used to organise and summarise the information so that the researcher can see what happened in the research study and can communicate the results to others
- Statistics help the researcher to answer the questions that initiated the research by determining exactly what general conclusions are justified based on the specific results that were obtained
- The term statistics refers to a set of mathematical procedures for organising, summarising and interpreting information
- Research in the behavioural sciences typically begins with a general question about a specific group of individuals
- A population is the entire set of the individuals of interest for a particular research question
- Populations can vary in size depending on how the researcher defines the population
- The set of individuals selected from a population to represent the population in the research study is called a sample group
- A variable is a characteristic or condition that changes or has different values for different individuals i.e. height, weight, gender or personality
- The measurement obtained for each individual is called a datum, score or raw score
- The complete set of scores is called the data set, or simply, the data
- Data are measurements or observations – a data set is a collection of measurements or observations – a datum is a single measurement or observation
- The average score for a population is a parameter
- A characteristic that describes the sample is call a statistic
- Descriptive statistics are statistical procedures used to summarise, organise and simplify data
- Take raw scores and organise or summarise them in a form that is more manageable
- Often scores are organised in a table or a graph
- Summarises a mean for the scores
- Inferential statistics are methods that use sample data to make general statements about a population
- They consist of techniques that allow us to study samples and then make generalisations about the populations from which they were selected

- Sampling error is the naturally occurring discrepancy or error that exists between a sample statistic and the corresponding population parameter
- Correlational method: two different variables are observed to determine whether there is a relationship between them
    - Limitations: do not provide explanation for relationship
    - Cannot demonstrate cause-effect relationship
- The goal of the experimental method is to demonstrate a cause-effect relationship between two variables
    - Manipulation: Researcher manipulates one variable by changing its value from one level to another. A second variable is observed to determine whether the manipulation causes change to occur
    - Control: The researcher must exercise control over the research situation to ensure that other, extraneous variables do not influence the relationship being examined
- Participant variables: Characteristics such as age, gender, intelligence that vary from one to another