Research Methods

Primary Data Collection: Quantitative Data Analysis

Two main types of data: Categorical and Numerical data.

- Categorical data:
  - grouped into a descriptive set (nominal);
  - put in rank order (ordinal).
- Numerical data - measured using numbers.
  - discrete data –discrete units
  - continuous data –any value

Describing data

- Description:
  - • The central tendency, which is the common, middle or average value: mode, median, mean.
  - • The dispersion, which is how the data values are spread, or dispersed, around the central tendency: standard deviation. IQR, range.
- Inference
  - Exploring Relationships: correlation, regression
  - Comparing Groups: t-test, ANOVA, ANCOVA

Presenting your Data

- Tables
- Bar Graph
- Histogram
- Pie Chart
- Trend Line
- Scatter graph

Population/Sample

- Your sample must represent your population.

Exploring Relationships:

- The association between two variables.
- The correlation between two variables.

Correlation

Correlation Coefficient (r)

- Categorical data: Chi-squared coefficient
- Ranked data: Kendall's rank correlation coefficient, Spearman's rank correlation coefficient.
- Numerical data: Pearson’s product moment correlation coefficient

Hypotheses

H_0 - ***null hypothesis*** **-** no relationship between two measured phenomena.
H_1 –**alternative hypothesis**
Rejecting or disproving the **null hypothesis** is a central task in modern scientific practice.
Statistically significant means that the null hypothesis has been rejected. This means we can accept the **alternative hypothesis**.

Significance
How likely the relationship between your variables has occurred by chance.

- Significance (alpha level) p
- Significant at p<.05 or 0.01
- There is less than a 5% (1%) chance that the result is due to chance.
- This means that the finding has a 95% (99%) chance of being true.

NB:
Statistical Significance ≠ Meaningful or Important
Correlation ≠ Causality: Third variable/Confounding variable

Regression

Coefficient of Determination - $R^{2:\,a}$ measure of how much variance is shared, how much of one variable can be explained by another.

Comparing Groups: Difference

t-test (unrelated) – independent samples t-test
Compares 2 unrelated groups (e.g. different towns) on same test. Need one categorical independent variable with only two groups (e.g. M/F) and one continuous dependent variable (e.g. items purchased).
t-test (related) – paired samples / matched t-tests / matched pairs
Compares the same group on 2 occasions (e.g. same product at two times). Need one categorical independent variable (e.g. time 1 & time 2) and one continuous dependent variable (e.g. product).

Effect Size
Effect sizes either measure the sizes of associations between variables or the sizes of differences between group means.
Several measures have been proposed and used:
• coefficient of determination – R
• the standardized mean difference - Cohen's d – and η^2 - eta squared - the correlation ratio.
Benchmarks (Cohen, 1988)

	Small	Medium	Large
R^2	.01	.09	.25
Cohen's d	.20	.50	.80
Eta squared - η^2	.01	.06	.14

Analysis of Variance (ANOVA): If three or more distinct groups:

One-Way Between-Groups ANOVA: Single independent variable.
Two-Way Between-Groups ANOVA: Factorial/Multifactorial ANOVA
- 2 independent variables - two factors
- Post Hoc Tests
- Homogeneous Subsets

Analysis of Covariance (ANCOVA): If three or more distinct groups – need to control for 3rd variable:

Two-Way ANCOVA) – two factors, controlling for 3rd variable

Summary

- We identify the variable of interest (dependent variable).
- We use descriptive statistics to get ideas about potential patterns in the data.
- These patterns then help us to devise experimental and null hypotheses.
- We then use inferential statistics to test the null hypothesis.
- If these inferential statistics return a p-value less than or equal to 0.05, then we have statistical significance and can reject the null hypothesis.
- If the p-value is greater than 0.05, then the null hypothesis cannot be rejected and we are unable to support the claims made by the experimental hypothesis.

References & Reading

Aldrich, J. O. & Rodriguez, H. M. (2013). *Building SPSS graphs to understand data*. London: Sage.

Anderson, D., Sweeney, D. & Williams, T. (2001). *Modern business statistics with Microsoft Excel*. New York: Cengage.

Brown, J. D. & Rogers, T. S. (2002). *Doing second language research*. Oxford: Oxford University Press.

Burns, R. B. & Burns, R. A. (2008). *Business research methods and statistics using SPSS.* London: Sage.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Routledge.

Colman, A. M. & Pulford, B. D. (2008). *A crash course in SPSS for Windows* (4th ed.). Oxford: Blackwell.

Davis, G. & Pecar, B. (2013). *Business statistics using Excel*. Oxford: Oxford University Press.

Dörnyei, Z. (2007*). Research methods in applied linguistics*. Oxford: Oxford University Press

Ellis, P. D. (2010). *The essential guide to effect sizes*. Cambridge: Cambridge University Press.

Hand, D. J. (2008) *Statistics: A very short introduction*. Oxford: Oxford University Press.

Jelen, B. (2013). *Excel 2013 charts and graphs*. Indianapolis: Que.

Lee, N. & Peters, M. (2015). *Business statistics using Excel and SPSS*. London: Sage.

Keller, G. (2000). *Applied statistics with Microsoft Excel*. London: Duxbury.

Kline, R. (2004). *Beyond significance testing: Reforming data analysis methods in behavioral research.* Washington, DC: American Psychological Association.

Knapp, H. (2016). *Introductory statistics using SPSS* (2nd ed.). London: Sage.

Larson-Hall, J. (2012). How to run statistical analyses. In A. Mackey & S. M. Gass (Eds.), *Research methods in second language acquisition* (pp. 245-274). London: Wiley-Blackwell.

Levon, E. (2018). Organising and processing your data: The nuts and bolts of quantitative analysis. In L. Litosseliti (Ed.), *Research methods in linguistics* (2nd. ed., pp. 138-166.). London: Bloomsbury Academic.

Lowie, W. & Seton, B. (2013). *Essential statistics for applied linguistics*. London: Palgrave Macmillan.

Meyers, L. S., Gamst, G. & Guarino, A. J. (2017). *Applied multivariate research: Design and interpretation* (3rd ed.). London: Sage.

Pallant, J. (2016). *SPSS survival manual* (6th ed.). London: McGraw Hill.

Popham, W. J. , & Sirotnik, K. A. (1973). *Educational statistics: Use and interpretation* (2nd ed.). New York: Harper & Row

Robson, C. (2011). *Real world research* (3rd ed.). London: Wiley.

Salkind, N. J. (2011). *Statistics for people who (think they) hate statistics* (4th ed.). London: Sage.

Salkind, N. J. (2015). *Excel statistics: A quick guide* (3rd ed.). London: Sage.

Saunders, M. & Lewis, P. (2012). *Doing research in business and management*. London: FT/Prentice Hall.

Saunders, M., Lewis, P. & Thornhill, A. (2016). *Research methods for business students* (7th ed.). London: FT/Prentice Hall.

Schmuller, J. (2016*). Statistical analysis with Excel for dummies* (4th ed.). London: Wiley.

Shavelson, R. (1988). *Statistical reasoning for the behavioral sciences*. Boston: Allyn and Bacon.

Speigelhalter, D. (2020). *The art of statistics: Learning from data*. London: Pelican.

Tabachnick, B. G. & Fidell, L. S. (2007). *Using multivariate statistics*. (5th ed.). Boston: Allyn & Bacon.