

# Research Design and Analysis

## Multiple Regression

Stacey Coleman

---

---

---

---

---

---

---

### Recap so far

- Correlation to linear regression
- Correlation:
  - Parametric test
  - Non-experimental design
- Linear regression:
  - Used to discover the effect of one variable (x) on another variable

2

---

---

---

---

---

---

---

### Recap so far

- Similar to a correlation: However!
  - Correlation only allows us to conclude how strongly 2 variables relate to each other (both in terms of magnitude and direction),
  - Linear Regression also tells us by HOW MUCH one variable will change if the other variable changes.
  - It aims to predict scores on one variable based on scores from another variable.

3

---

---

---

---

---

---

---

## • Terms to remember with regression

### CRITERION VARIABLE

- Same as the DV

### PREDICTOR VARIABLE

- The same as the IV; the variable that is predicting a change in the DV

### PREDICT

- Always looking to predict something with regression, not looking for differences

- **b** = The amount by which y changes, when x changes by 1 unit – this is represented by the slope or tilt of our straight line in a scatterplot.

- **a** = the intercept (also called the Constant) – the value of Y' when X is 0.

4

---

---

---

---

---

---

---

---

## How is multiple regression different?

5

---

---

---

---

---

---

---

---

## • Multiple Regression

- An extension of correlation when using **more than one independent variable**

- Possible research questions?

- Which of a **range** of variables are most strongly related to a specific variable of interest?
- Do these variables do better than chance or simple mean scores at **predicting** scores on a variable of interest?
- How might these variables **combine** to predict a variable?

6

---

---

---

---

---

---

---

---

## • Multiple Regression



- Why not just carry out many simple correlations??
  - We could – this would provide us with several bi-variate correlations
  - But! Still only a non-causal correlation and this cannot indicate our overall success with all the variables used...

7

---

---

---

---

---

---

---

## • Multiple Regression

- Important extension of 'regression' over 'correlation':
  - **Combined** association effect of the predictor variables – multiple regression!
  - Treating the predictors (separate variables) as a combined variable in relation to the dependent variable
  - How 'associated' they are as a whole
  - How 'predictive' they are of the DV as a whole
    - (nb. Also looks at each in isolation...)

8

---

---

---

---

---

---

---

## • Multiple Regression

- Multiple regression allows the researcher to assess the relative importance of each of the variables in question...
- Ideal for investigation of complex real-life research questions
  - Less used in lab-based experimental designs
  - Ideal for longitudinal prospective designs

9

---

---

---

---

---

---

---

## • Multiple Regression

- Often used to test how 'good' a theory is
  - To what extent do the variables identified by a theory explain a given behaviour
  - Can then add additional variables/new element to the theory and test this

10

---

---

---

---

---

---

---

## • Assumptions for data when conducting multiple regression

---

---

---

---

---

---

---

## • Multiple Regression: Assumptions

- MR is one of the fussier statistical techniques and makes a number of assumptions about data
  - And not all that forgiving if they are violated

### 1. Continuous data is best (interval or ratio)

- Nominal data can be used, but some caution around this

### 2. Normal distribution of data (normality)

### 3. Linear relationship between IV's and DV (linearity)

12

---

---

---

---

---

---

---

## • Multiple Regression: Assumptions

### 4. Sample size is important!

- Stevens (1996): approx. 15 participants per predictor
- Tabachnick and Fidell (2013)  $N > 50 + 8m$

### 5. Multicollinearity

- Exists when IV's are highly correlated ( $r = .90+$ )
- You want your IV's to be correlated with your criterion variable, but not necessarily with the other predictors

### 6. Singularity

- When one IV is actually a combination of other IVs (e.g., when subscale and total scores are used)

13

---

---

---

---

---

---

---

---

## • Multiple Regression: Assumptions

### 7. Outliers

- MR is **very** sensitive to outliers
- Always check for extreme scores during cleaning and screening (for IV's and DV)
- Tabachnick and Fidell (2013): outliers are those with standardised residuals above 3.3 (or less than -3.3)

### 4. Homoscedasticity

- Refers to the variance of errors being the same across all IV's
- Check using residuals scatterplot

14

---

---

---

---

---

---

---

---



15

---

---

---

---

---

---

---

---



So, how do you choose which variables to use as your 'predictors' in a regression design??

---

---

---

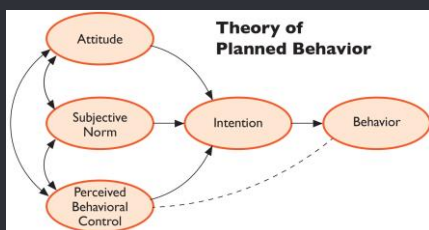
---

---

---

---

### • Specific example of a 'typical' MR



Testing the theory of planned behaviour components to **predict** intentions to engaged in sub-safety behaviour

17

---

---

---

---

---

---

---

### • 'Typical' MR: Research questions

Study example: Do the TPB components predict intentions to engage in sun-safety behaviour?

1. To what extent does the TPB model as a whole predict intentions to engage in sun-safety behaviour?
2. To what extent do each of the **individual predictor variables** predict intentions?
3. Which predictor variable is the **BEST** predictor?

18

---

---

---

---

---

---

---

## Main outputs for Multiple Regression

### 1. MODEL SUMMARY TABLE

What percentage of the variance is explained by the model as a whole?

### 2. ANOVA TABLE

Is the model significantly better at predicting intentions to engage in binge drinking than simply making predictions using the mean score for the DV?

### 3. COEFFICIENTS TABLE

Which of the individual predictor variables predicts intentions to engage in binge drinking? How much change in each predictor variable equates to change in the criterion (DV)

19

## Conduct study, run analysis, CHECK ASSUMPTIONS!

### ◦ Multicollinearity (correlations table)

- Your IV's should show at least some relationship with your DV (preferably **above 0.3**)
- You don't really want your IV's to be too highly correlated with one another – anything **below 0.7** in this case is considered ok
- Collinearity diagnostics can also be requested and performed by SPSS – presented in the coefficients table labelled TOLERANCE and VIF

20

| Model | t      | Sig. | 95.0% Confidence Interval for B |             | Collinearity Statistics |       |
|-------|--------|------|---------------------------------|-------------|-------------------------|-------|
|       |        |      | Lower Bound                     | Upper Bound | Tolerance               | VIF   |
| 1     | .302   | .767 | -4.139                          | 5.513       |                         |       |
| 2     | 6.578  | .000 | .879                            | 1.716       | .766                    | 1.305 |
| 3     | -1.037 | .315 | -.972                           | .333        | .488                    | 2.051 |
| 4     | 1.579  | .134 | -.117                           | .798        | .410                    | 2.439 |

**Tolerance:**  
Values should be above .10 – not violated assumption of collinearity

**VIF:**  
Values should be below 10 – not violated assumption of collinearity

21

### Next: deal with outliers and normality

- Multiple regression is very sensitive to outliers!
- Inspect and interpret this assumption mainly using the plots created by SPSS

22

---

---

---

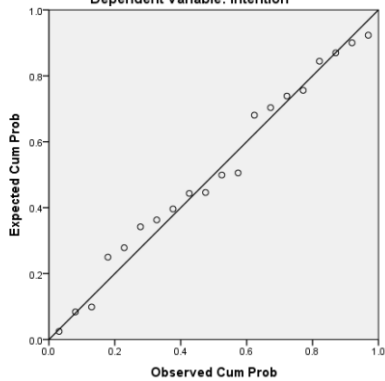
---

---

---

---

Normal P-P Plot of Regression Standardized Residual  
Dependent Variable: Intention



23

---

---

---

---

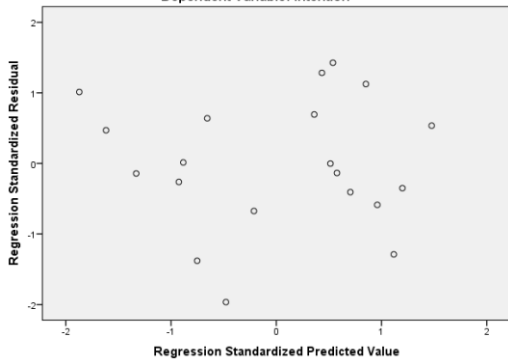
---

---

---

Scatterplot

Dependent Variable: Intention



24

---

---

---

---

---

---

---

## • Regression Statistics

- MR indicates how good a 'fit' all the predictors are in explaining the DV
  - Can assess the degree to which the DV can be predicted from the set of predictors
- R – multiple regression coefficient
  - Between 0 and 1
  - The greater the R, the less residuals
- $R^2$  is of more use ... the % of variance in the DV that the 'model' explains
- Adjusted  $R^2$  - adjusted for sample size
- Beta values represent the amount of contribution (how good a fit) for each of the IV's

25

---

---

---

---

---

---

---

---

## • Reporting MR results

- Today in the workshop!

26

---

---

---

---

---

---

---

---

## • Statistical Power, Effect and Sample Sizes

---

---

---

---

---

---

---

---

## • Statistical power, effect and sample sizes

- Statistical significance versus clinical significance
- The importance of sample size and the factors that need to be considered to calculate an appropriate sample size
- Power
- Effect sizes

28

---

---

---

---

---

---

---

## • Statistical significance

- We are never **certain** of an effect/relation, so we always deal in the area of probability (p-values!)
- We (usually) test for one of two main things:
  - **Differences between scores**
  - Usually sets of scores calculated as grand means e.g., t-test
  - **Associations between scores**
  - Is there a relationship between (x) and (y) e.g., correlation

29

---

---

---

---

---

---

---

## • Reminders!

- 0.05 means there is a 5% likelihood that the results were due to chance/random (null hypothesis)
- 1.0 would mean a 100% certainty that the results were down to chance
- In statistical language, 'significant' simply means that there was a less than 5% chance of getting a result under what would be correct for the null hypothesis to be true...

30

---

---

---

---

---

---

---

## Reminders!

- But! Finding a statistically significant result doesn't necessarily mean that you can now go out and change the world
  - Likewise, not finding a statistically significant result doesn't necessarily mean that there is nothing of interest...Statistical significance is determined through conducting inferential tests (the stuff you've been doing), and is noted as the sig. or sig. value in SPSS

31

---

---

---

---

---

---

---

## Reminders!

- **Type 1 error:** (find a significant p-value in error)
  - Reject the null hypothesis when it is correct
  - Accept the alternative hypothesis in error
  - Displayed like this: Type I error
- **Type 2 error:** (find a non-significant p-value in error)
  - Reject the alternative hypothesis in error
  - Accept the null hypothesis by mistake when there really is an underlying effect
  - Displayed like this: Type II error

32

---

---

---

---

---

---

---

## Clinical significance

- *The practical importance of a study, with a genuine effect on real life!*

- E.g., how effective is a new treatment?
- E.g., does one type of intervention work better than another?

33

---

---

---

---

---

---

---

## • Clinical significance

- If findings from research are considered **clinically significant** then they are likely to influence how actual patients in the real world are treated
- It could amend the way in which health professionals work and what guidance is given from regulatory bodies, and even provide supporting evidence to allow government to pass new or update legislation

34

---

---

---

---

---

---

---

## • Sample Size

- We typically use estimates or % chance/confidence
- It would be great if we could collect data from an entire population – we could remove this chance and estimation element
- But this is impossible to achieve in most cases! So we select a sample of individuals who we believe are representative of the population we are looking to investigate

35

---

---

---

---

---

---

---

## • Sample Size

- Common sense suggests that having the largest sample size possible would be of most benefit in order to make meaningful conclusions about the population
  - If your sample size is too small then you may not detect an important underlying effect!
  - But – in clinical studies in particular, the population that we can draw our sample from may already be small (e.g., patients with a rare disease).

36

---

---

---

---

---

---

---

## • Sample Size

- So, we need to try and get the balance right between getting enough P's to detect an effect without running in to the problem of trying to get an unrealistically large sample.
- We can calculate an optimal sample size – this must be done during **the design stage**

37

---

---

---

---

---

---

---

## • Calculating sample size

- We need to specify the following quantities during the **design stage**:
  - **POWER**
    - The chance of detecting a statistically significant effect (if there is one). Power should be 70%+.
  - **SIGNIFICANCE LEVEL**
    - Usually remains at 0.05 (point where we reject the null)
  - **SMALLEST EFFECT OF INTEREST**
    - Magnitude of the effect that is clinically important – i.e., the difference (this can be expressed as the **Standardised Difference**)

38

---

---

---

---

---

---

---

## • Calculating sample size

- Power and significance are relatively easy to specify
- Obtaining a value for the standardised difference requires two further things:
  1. An estimate of the smallest difference in means needed
  2. An assumed equal standard deviation of the observations (either in groups or across responses)

39

---

---

---

---

---

---

---

## Methods for calculating sample size

- General formulae
- Quick formulae
- Special tables
- Computer packages

40

---

---

---

---

---

---

---

## Methods for calculating sample size

- General formulae
- Quick formulae
- Special tables
- Computer packages

41

---

---

---

---

---

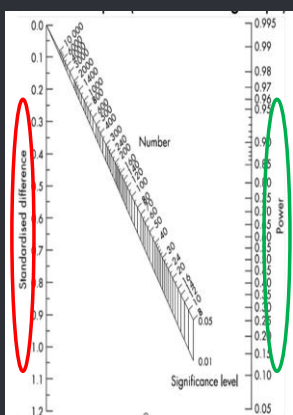
---

---

## 1. Altman's nomogram

Requires the calculation of the **standardised difference** and to join values on the left hand axis of the nomogram to the power we have specified on the right hand side of the axis

NB: also need to specify alpha level - .05 or .01




---

---

---

---

---

---

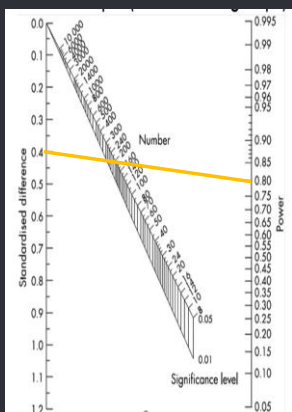
---

## 1. Altman's nomogram

The required sample size is indicated at the point at which the resulting line and sample size axis meet

Example Power = 80%  
Standardised difference = 0.4  
Alpha level = 0.05

200 participants




---

---

---

---

---

---

---

---

## Calculating the standardised difference

- The formulae for calculating the standardised difference will depend upon the test you are looking to conduct:

### Independent t-test:

$$\frac{\delta}{\sigma}$$

Where  $\delta$ : is the smallest difference in means  
 $\sigma$ : assumed equal standard deviation of the observations in each of the 2 groups

---

---

---

---

---

---

---

---

## Calculating the standardised difference

### Paired t-test:

$$\frac{\delta}{\sigma_d}$$

Where  $\delta$ : is the smallest mean difference  
 $\sigma_d$ : the standard deviation of the differences in responses

---

---

---

---

---

---

---

---

## Calculating the standardised difference

### Chi<sup>2</sup> test:

$$p_1 - p_2$$

$$\sqrt{p(1-p)}$$

$p_1 - p_2$  = the smallest difference in the proportion of 'success' in the groups

To calculate  $\bar{p} = \frac{p_1 + p_2}{2}$

46

## 2. Quick formulae: Lehr's formula

For independent t-test and the Chi<sup>2</sup> test, you can use Lehr's formula for calculating the sample size for a power of 80% and a two-tailed significance level of 0.05

$$\frac{16}{(\text{Standardised difference})^2}$$

Note: the numerator of 21 is used (instead of 16) to relate to a power of 90%.

47

## Power Statement

- With applied research in particular, you will more than likely need to include a power statement as part of your study protocol or in the method section of a report to show that you considered the sample size during the designing stages of your study
- A typical statement could be:
  - "84 patients in each group were required to have a 90% chance of detecting a difference in means of 2.5 days (SD = 5 days) at the 5% level of significance using an independent samples t-test."

48

## • Effect size

- The REAL effect in the population Sometimes it's not enough to simply know that an effect is significant!
- We may also want to know the Size of the Effect
  - E.g., we may want to know by how much does a new drug reduce episodes of depression
- Whether the manipulation of an IV has created a sizeable difference between our samples on some measure
  - The further the difference between the 2 means, the larger the effect size

49

---

---

---

---

---

---

---

---

## • Effect size

- The popular measures of effect size to compare groups are Cohen's d and eta squared ( $\eta^2$ )/partial eta squared.
- In Psychology we typically use partial eta squared, which is what SPSS provides us with in the very last column of our inferential statistics box when conducting more complex analyses (such as ANOVA).
  - As long as we have remembered to tick Display 'Estimates of effect size' in the Options Dialogue box!

50

---

---

---

---

---

---

---

---

## • Effect size

- However, SPSS does not provide effect size statistics for t-tests but you can use information from SPSS to calculate the effect size statistic either by hand or by using online effect size calculators:
- <http://www.uccs.edu/~lbecker/>

51

---

---

---

---

---

---

---

---

# Calculating effect size using Eta-Squared

Ranges from 0 to 1, represents the proportion of the variance in the DV that's explained by the IV

Use the values from your SPSS output in the formulas:

• Independent samples t-test =

$$\text{Eta Squared} = \frac{t^2}{t^2 + (N1 + N2 - 2)}$$

Example:

| Group Statistics       |    |         |                |                 |
|------------------------|----|---------|----------------|-----------------|
| Group                  | N  | Mean    | Std. Deviation | Std. Error Mean |
| Anxiety - control      | 10 | 23.0000 | 2.52943        | .80443          |
| anxiety - experimental | 10 | 15.8000 | 3.32666        | 1.05198         |

Difference between a control and experimental group on levels of anxiety:

| Independent Samples Test |                                         |      |                              |        |                 |                 |                       |                                           |          |
|--------------------------|-----------------------------------------|------|------------------------------|--------|-----------------|-----------------|-----------------------|-------------------------------------------|----------|
|                          | Levene's Test for Equality of Variances |      | t-test for Equality of Means |        |                 |                 |                       |                                           |          |
|                          | F                                       | Sig. | t                            | df     | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference |          |
| Anxiety - control        | .800                                    | .449 | 5.214                        | 18     | .000            | 7.20000         | 1.36882               | 4.29900                                   | 10.10100 |
| Anxiety - experimental   |                                         |      | 5.214                        | 17.548 | .000            | 7.20000         | 1.36882               | 4.29381                                   | 10.10619 |

$$\text{Eta Squared} = \frac{t^2}{t^2 + (N1 + N2 - 2)}$$

$$t = 5.21$$

$$N1 = 10$$

$$N2 = 10$$

Example:

$$\text{Eta squared} = \frac{5.21^2}{5.21^2 + (10 + 10 - 2)}$$

$$\frac{27.14}{27.14 + 18} = \frac{27.14}{45.14} = 0.60$$

- **Effect size**

- Guidelines proposed by Cohen (1988) for interpreting the value are:
  - .01 = small
  - .06 = medium
  - .14 = large
- You need to use your stats books to see how effect size is reported!

55

---

---

---

---

---

---

---

- **Any questions?**

---

---

---

---

---

---

---