

Multiple Linear Regression II



Lecture 8

Survey Research & Design in Psychology
James Neill, 2015
[Creative Commons Attribution 4.0](#)

Overview



1. Summary of MLR I
2. Semi-partial correlations
3. Residual analysis
4. Interactions
5. Analysis of change
6. Writing up an MLR
7. Summary of MLR II
8. MLR II quiz – Practice questions

2

Readings

1. As per previous lecture, plus
2. **Howitt & Cramer (2011/2014)**
 - Ch 38/39: Moderator variables on relationships between two variables

3

Summary of MLR I

Purpose of MLR

- To examine linear relations between:
 - two or more predictors (IVs; X) and
 - a single outcome variable (DV; Y)

Model

- Establish a model
 - Path diagram or
 - Venn diagram
- Express as one hypothesis per IV

4

Summary of MLR I

Check assumptions

- Level of measurement
- Sample size
- Normality
- Linearity
- Homoscedasticity
- Collinearity
- Multivariate outliers
- Residuals should be normally distributed

5

Summary of MLR I

Choose type

- Standard
- Hierarchical
- Stepwise, Forward, Backward

6

Summary of MLR I

Interpret output

- Overall:
 - R , R^2 , Adjusted R^2
 - Changes in R^2 (if hierarchical)
 - Significance (F , p)
- For each IV:
 - Standardised coefficient
 - size, direction and significance
 - Unstandardised coefficient
 - report equation (if useful)
 - Partial correlations (s^2)

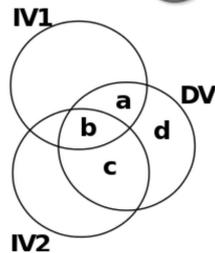
7

Semi-partial correlation



If this Venn Diagram depicts %s of variance and shared variance,

- $a + b + c + d =$ variance in DV
- $a + b + c = R^2$
- $a + c =$ uniquely explained variance
- $b =$ non-uniquely explained variance
- what is a ?
- what is c ?



a and c are semi-partial correlations squared (s^2):

$a = s^2$ between $IV1$ and DV after controlling for (partialling out) the influence of $IV2$
 $c = s^2$ between $IV2$ and DV after controlling for (partialling out) the influence of $IV1$

8

Semi-partial correlations in MLR

When interpreting MLR coefficients:

- Draw a Venn diagram
- Compare 0-order (r) and semi-partial correlations (s^2) for each IV to help understand the nature of relations amongst the IVs and the DV:
 - A semi-partial correlation (s^2) will be less than or equal to the correlation (r)
 - If a s^2 equals a r , then the IV independently predicts the DV
 - To the extent that a s^2 is less than the r , the IV's explanation of the DV is shared with other IVs
 - An IV may have a significant r with the DV, but a non-significant s^2 . This indicates that the unique variance explained by the IV may be 0.
- Compare the relative importance of the predictors using betas and/or s^2 s (latter preferred)

9

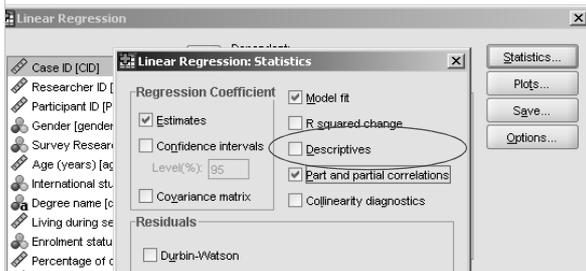
Semi-partial correlations (s_r) in MLR

- SPSS provides semi-partial correlations
- The s_r^2 indicates the % of variance in the DV which is uniquely explained by an IV.
- In SPSS, the s_r s are labelled “part”. Square these to get s_r^2 .

10

Semi-partial correlations for MLR in SPSS

In Linear Regression - Statistics dialog box, check “Part and partial correlations”



Semi-partial correlations for MLR in SPSS - Example

Coefficients^a

Model		Correlations		
		Zero-order	Partial	Part
1	Zscore: GA: General Ability Total	.285	.107	.090
	Zscore: PSA: Perceptual Speed Ability Total	.419	.319	.280
	Zscore: PMA: Psychomotor Ability Total	.275	.034	.028
	QBK: Total Knowledge of Text Editing Keys (% correct)	.386	.368	.329

a. Dependent Variable: Typing Test: Speed (Words Per Minute)

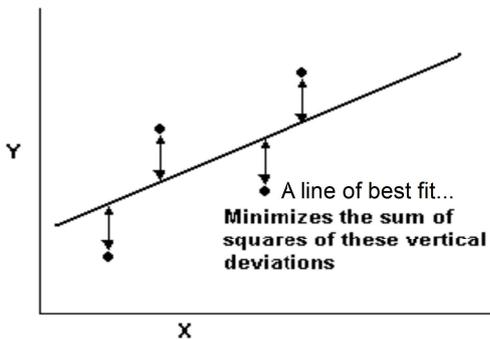
Semi-partial correlations (s_r) indicate the relative importance of each of the predictors; s_r^2 indicate the % of variance uniquely explained by each predictor.

Summary:

Semi-partial correlation (sr)

1. In MLR, sr is labelled "part" in the regression coefficients table SPSS output
2. sr^2 is the unique % of the DV variance explained by each IV

Residual analysis



Residual analysis

Three key assumptions can be tested using plots of residuals:

1. **Linearity**: IVs are linearly related to DV
2. **Equal variances** (Homoscedasticity))
3. **Normality of residuals**

Residual analysis

Assumptions about residuals:

- Random noise (error)
- Sometimes positive, sometimes negative but, on average, 0
- Normally distributed about 0

16

Residual analysis

Linear Regression: Plots [X]

DEPENDENT
*ZPRED
*ZRESID
*DRESID
*ADJPRED
*SRESID
*SDRESID

Scatter 1 of 1
Previous Next

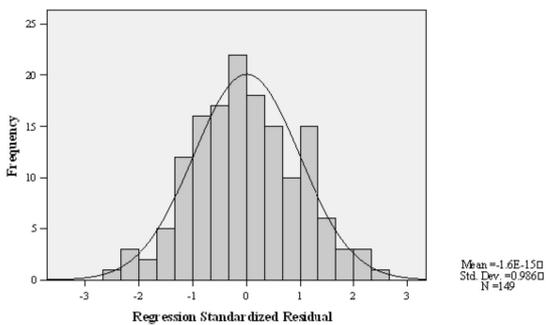
Y: *ZRESID
X: *ZPRED

Standardized Residual Plots
 Histogram
 Normal probability plot
 Produce all partial plots

Continue
Cancel
Help

Residual analysis

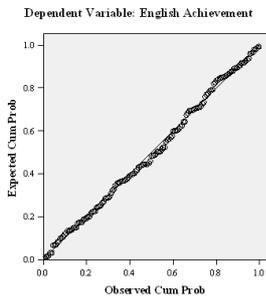
Dependent Variable: English Achievement



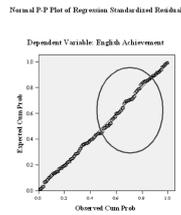
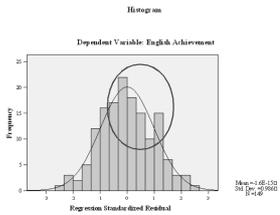
“The **Normal P-P (Probability) Plot of Regression Standardized Residuals** can be used to assess the assumption of normally distributed residuals. **If the points cluster reasonably tightly along the diagonal line (as they do here), the residuals are normally distributed.** Substantial deviations from the diagonal may cause for concern.”
 Allen & Bennett, 2008, p. 183

Residual analysis

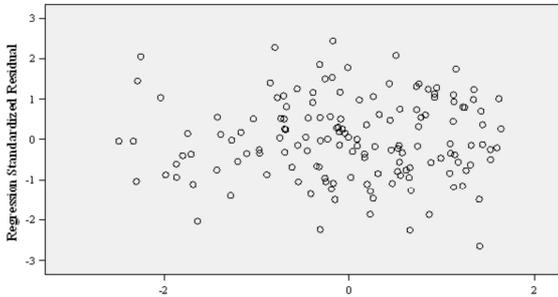
Normal P-P Plot of Regression Standardized Residual



Residual analysis



“The Scatterplot of standardised residuals against standardised predicted values can be used to assess the assumptions of normality, linearity and homoscedasticity of residuals. The absence of any clear patterns in the spread of points indicates that these assumptions are met.”
 Allen & Bennett, 2008, p. 183



Why the big fuss about residuals?

↑ assumption violation

→

↑ Type I error rate
(i.e., more false positives)

22

Why the big fuss about residuals?

- Standard error formulae (which are used for confidence intervals and sig. tests) work when residuals are well-behaved.
- If the residuals don't meet assumptions these formulae tend to underestimate coefficient standard errors giving overly optimistic p -values and too narrow CIs.

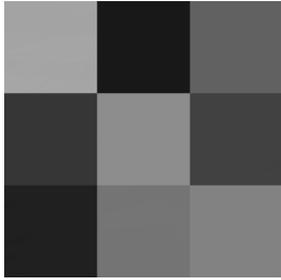
23

Summary: Residual analysis

1. Residuals are the difference between predicted and observed Y values
2. MLR assumption is that residuals are normally distributed.
3. Examining residuals also helps to assess:
 1. Linearity
 2. Homoscedasticity

24

Interactions



Interactions

- Additivity - when IVs act independently on a DV, i.e., they do not interact.
- However, there may also be interaction effects - when the magnitude of the effect of one IV on a DV varies as a function of a second IV.
- Also known as an interaction or moderation effect.

26

Interactions: Example 1

Some drugs interact with each other to reduce or enhance other's effects e.g.,

Pseudoephedrine → ↑ Arousal

Caffeine → ↑ Arousal

Pseudoeph. X Caffeine → ↑↑↑ Arousal



X



27

Interactions: Example 2

Physical exercise (IV1)
and
Natural environments (IV2)
may provide
multiplicative benefits
in reducing stress e.g.,

Natural environment → ↓ Stress

Physical exercise → ↓ Stress

Natural env. X Phys. ex. → ↓↓↓ Stress

28

Interactions: Example 3

University student satisfaction (IV1)
and
Level of coping (IV2)
may predict level of stress e.g.,

Satisfaction → ↓ Stress

Coping → ↓ Stress

Satisfaction X Coping → ↓↓↓ Stress

Dissatisfaction X Not coping → ↑↑↑ Stress

29

Interactions

Test interactions in MLR by computing
a cross-product term e.g.,:

- Pseudoephedrine (IV1)
- Caffeine (IV2)
- Pseudoephedrine x Caffeine (IV3)

30

Interactions

$$Y = b_1x_1 + b_2x_2 + b_{12}x_{12} + a + e$$

- b_{12} is the product of the first two slopes ($b_1 \times b_2$)
- b_{12} can be interpreted as the amount of change in the slope of the regression of Y on b_1 when b_2 changes by one unit.

31

Interactions

- Conduct Hierarchical MLR
- Step 1:
 - Pseudoephedrine
 - Caffeine
- Step 2:
 - Pseudo x Caffeine (cross-product)
- Examine ΔR^2 , to see whether the interaction term explains additional variance above and beyond the direct effects of Pseudo and Caffeine.

32

Interactions

Possible effects of Pseudo and Caffeine on Arousal:

- None
- Pseudo only (incr./decr.)
- Caffeine only (incr./decr.)
- Pseudo + Caffeine (additive inc./dec.)
- Pseudo x Caffeine (synergistic inc./dec.)
- Pseudo x Caffeine (antagonistic inc./dec.)

33

Interactions – SPSS example

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change
	1	.543 ^a	.295	.288	17.34399
2	.547 ^b	.299	.289	17.33241	.004

a. Predictors: (Constant), Caffeine, Pseudoephedrine

b. Predictors: (Constant), Caffeine, Pseudoephedrine, PseudoCaffeine interaction

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	57.705	5.109		11.296	.000
	Pseudoephedrine	11.511	1.510	.464	7.625	.000
	Caffeine	4.735	1.780	.162	2.660	.008
2	(Constant)	42.730	14.137		3.023	.003
	Pseudoephedrine	18.137	6.025	.732	3.010	.003
	Caffeine	10.126	5.069	.346	1.998	.047
	PseudoCaffeine interaction	-2.313	2.036	-.380	-1.136	.257

Interactions

- Cross-product interaction terms may be highly correlated (multicollinear) with the corresponding simple IVs, creating problems with assessing the relative importance of main effects and interaction effects.
- An alternative approach is to conduct separate regressions for each level of the interacting variable.

Summary: Interactions

1. In MLR, IVs may interact to:
 1. Increase effect on DV
 2. Decrease effect on DV
2. Model interactions with hierarchical MLR:
 1. Step 1: Enter IVs
 2. Step 2: Enter cross-product of IVs
 3. Examine change in R^2

Analysis of Change

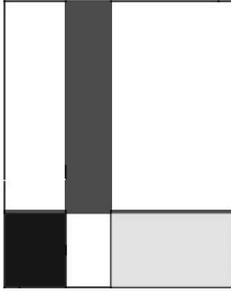


Image source: <http://commons.wikimedia.org/wiki/File:PMinspiret.jpg>, CC-by-SA 3.0

Analysis of change

There are two ways MLR can be used to analyse the variance of changes in outcome measures over time:

1. Compute difference scores (Time 2 – Time 1) and use these as a DV in a standard MLR
2. Hierarchical MLR:
 1. "Partial out" baseline measures from post-scores: Enter post-score as the DV and the pre-score as the IV in Step 1
 2. Explain remaining variance in the post-measure using IV(s): Enter the other IVs of interest in Step 2 and examine change in R^2

38

Analysis of change

Example research question:

Does the quality of social support (IV2) explain changes in mental health between the beginning (IV1) and end (DV) of an intervention ?

39

Analysis of change

Step 1: partial out baseline mental health from post mental health

Step 1:
Baseline
Mental Health
(IV1)

Post
Mental
Health
(DV)

Step 2:
Social
Support
(IV2)

Step 2: analyse the effect of Social Support on the change (remaining variance) in Mental Health

40

Analysis of change

Hierarchical MLR

- DV = Mental health after the intervention
- Step 1
 - IV1 = Mental health before the intervention
- Step 2
 - IV2 = Support from group members

41

Analysis of change

Results of interest

- Change in R^2 – amount of additional variance in the DV (MH2) explained by the predictor (IV2) added in Step 2 (after having baseline IV1 variance has been accounted for in Step 1)
- Regression coefficients for the predictor (IV2) in Step 2 – indicates variance explained in the DV (MH2) by the IV (SS) after controlling for baseline MH (MH1) in Step 1

42

Summary: Analysis of change

1. Analysis of changes over time can be assessed by

1. Standard regression

1. Calculate difference scores (Time 2 minus Time 1) and use as DV

2. Hierarchical MLR

1. Step 1: "Partial out" baseline scores

2. Step 2: Enter other IVs to help predict variance in changes over time.

43

Writing up an MLR

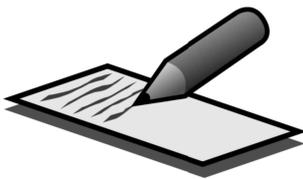


Image source: <http://commons.wikimedia.org/wiki/File:Writing.svg>, Public domain

Writing up an MLR (Results)

- Describe purpose
- Present a model, research question, and/or hypotheses
- Table of univariate descriptive statistics (M , SD , Skewness, Kurtosis)
- Table of correlations
- Describe type of MLR and testing of assumptions
- Table of regression coefficients (unstandardised and standardised), t , p , and sr^2 values

45

Table of regression coefficients

Table 2

Standard multiple regression of non-productive coping strategies on Psychological Distress reported by adolescents during outdoor education programs

Variables	PD ^a (DV)	NC	WT	W	SC	B	β	r^2 ^b
						(unique)		
Not Coping (NC)	.43					.41*	.29	.07
Wishful Thinking (WT)	.38	.36				.15*	.15	.02
Worry (W)	.36	.25	.40			.13*	.13	.01
Somatic Coping (SC)	.42	.22	.30	.38		.26*	.26	.06
Intercept						.60*		

Note. ^a $p < .05$; ^bPD = Psychological Distress; ^c r^2 = the squared semipartial correlations indicate the unique variance predicted by the independent variable.

Summary: Writing up an MLR

1. Introduction:

1. Purpose
2. Describe model, research question, and/or hypotheses

2. Results:

1. Univariate descriptive statistics
2. Correlations
3. Type of MLR and assumptions
4. Regression coefficients
5. Equation (if appropriate)

47

Questions

?

48

**MLR II Quiz –
Practice question 1**

In an MLR, if the r between the two IVs is 1, R will equal the r between one of the IVs and the DV.

[Hint: Draw a Venn Diagram.]

- True
- False

49

**MLR II Quiz –
Practice question 2**

In MLR, if two IVs are somewhat correlated with the DV and with one another, the srs between the IVs and the DV will be _____ in magnitude than the rs :

[Hint: Draw a Venn Diagram.]

- Equal
- Smaller
- Larger
- Impossible to tell

50

**MLR II Quiz –
Practice question 3**

In MLR, the unique variance in the DV explained by a particular IV is estimated by its:

- Zero-order correlation squared (r^2)
- Multiple correlation coefficient squared (R^2)
- Semi-partial correlation squared (s^2)

51

MLR II Quiz – Practice question 4

Interaction effects can be tested in MLR by using IVs that represent:

- Cross-products between the IVs and DV
- Cross-products of IVs
- Semi-partial correlations squared (s^2)

52

MLR II Quiz – Practice question 5

To assess the extent to which social support from group members during an intervention program explain changes in mental health (MH) between the beginning and end of an intervention program, what MLR design could be used?

- Hierarchical with pre-MH in Step 1
- Hierarchical with cross-products of IVs in Step 2

53

References

- Allen, P. & Bennett, K. (2008). *SPSS for the health and behavioural sciences*. South Melbourne, Victoria, Australia: Thomson.
- Francis, G. (2007). *Introduction to SPSS for Windows: v. 15.0 and 14.0 with Notes for Studentware* (5th ed.). Sydney: Pearson Education.
- Howell, D. C. (2010). *Statistical methods for psychology* (7th ed.). Belmont, CA: Wadsworth.
- Howitt, D. & Cramer, D. (2011). *Introduction to statistics in psychology* (5th ed.). Harlow, UK: Pearson.

54

Open Office Impress

- This presentation was made using Open Office Impress.
- Free and open source software.
- <http://www.openoffice.org/product/impress.html>