

Multiple Regression: The Basics

Multiple regression is not a technique to use on small samples where the distribution of scores can be skewed. MR makes a number of assumptions about the data you have collected and input and is not all that forgiving if they are violated.

Assumptions:

1. Your DV must be measured on a continuous scale. Either interval or ratio data. If your DV is categorical then you would need to use a different type of regression analysis (e.g., ordinal regression)
2. You must have at least two IV's. Your IV's can be continuous or categorical.
3. Normality, linearity, homoscedasticity, independence of residuals
 - a. Normality: the residuals should be normally distributed about the predicted DV scores
 - b. Linearity: the residuals should have a straight-line relationship with predicted DV scores
 - c. Homoscedasticity: the variance along the line of best fit should remain similar as you move along the line
4. Multicollinearity
 - a. This refers to the relationship among the IV's. Multicollinearity exists when the IV's are highly correlated ($r = \text{above } 0.9$).
5. Outliers
 - a. MR is very sensitive to outliers. You should always check for extreme scores when screening your data initially, and this should be done for the IV's and the DV. Outliers can either be deleted from the data set or, alternatively, given a score for that variable that is high but not too different from the remaining cluster of scores.

What the MR does:

Multiple regression will tell you how much of the variance in your DV can be explained by your IV(s). It also gives you an indication of the relative contribution of each IV. The tests run within the MR allow you to determine the statistical significance of the results, in terms of both the model itself and the individual IV's.

Procedure:

- 1 •Analyze → Regression → Linear
- 2 •Move continuous DV into **Dependent** box
- 3 •Move IVs to **Independent** box
- 4 •Method → Make sure **ENTER** is selected
- 5 •Statistics → select **Estimates, Confidence Intervals, Model Fit, Descriptives, Part and partial correlations, and Collinearity diagnostics.**
•Under residuals, select **Casewise diagnostics and outliers outside 3 SD's**
•Click **Continue**
- 6 •Options → in the **Missing Values** section, select **Exclude cases pairwise.**
•Click **Continue**
- 7 •Plots → Move ***ZRESID** to the **Y** box; move ***ZPRED** to the **X** box.
• In the **Standardised Residual Plots** section, tick the **Normal probability plot** option
•Click **Continue**
- 8 •Click on **OK** and the output will be generated.