

SPSS

Correlation and Simple Linear Regression

A Evans

1. Correlation:

SPSS: Analyze: Correlate: Bivariate: select *Pearson* and *Spearman* and *two-tailed*; enter *variables*; **Options:** *means and standard deviations*; *exclude pairwise*.

2. Test Assumptions (Eyeball test):

SPSS: Graphs: Scatter: Simple: Define: enter two variables; double click on the graph (may need to do it twice) until it shows “SPSS Chart Editor” at the top of the screen; go to menu: **Chart: Options:** Show Sunflowers; Fit Line: Total; Fit Options: Linear Regression; Regression Prediction Line: Individual; Include Constant; Display R squared.

- Is there a straight line relationship or some curved line relationship? How closely do the points cluster around the best straight line?
- Are the assumptions of the Pearson correlation satisfied? Are there lots of ties?

3. Linear Regression:

SPSS: Analyze: Regression: Linear: select outcome variable (Dependent) and predictor variable (Independent); **Statistics:** select *Estimates*, *Confidence intervals*, and *Descriptives*.

- How would you describe the strength and direction of the association?
- What is the P value? Does this indicate a strong association?
- What does the correlation coefficient mean in words?
- For every one standard deviation change in the *independent variable* (X), what's the expected change in the *dependent variable* (Y)?
- Can you say with confidence that patients with dramatically different values for X have different values for Y?
- What's your 95% confidence band for an individual patient's value of Y for some value of X?

4. Linear Regression Assumptions:

Are the assumptions of linear regression satisfied?

1. **Linearity:** Is a straight line the best summary of the data; or is some other shaped line better?

Eyeball test: **SPSS: Graphs: Scatter: Simple: Define:** enter Y and X; double click on graph; once in Chart Editor: **Chart: Options:** Show sunflowers; Fit line: Total.

2. **Independence:** Are some individuals represented by more than one dot in the scatter plot (ie, did some individuals get the posttest more than once)?

3. **Normality:** For each value of X are the values of Y roughly “normal?”

Eyeball test of the scatter plot: most points clustered near the line, nearly equally distributed on either side, and without any outliers.

4. **Equality of variance:** For all values of X is there a similar amount of variation in Y?

Same eyeball test as above.

However, a better eyeball test for (3) and (4):

SPSS: Analyze: Regression: Linear: Plots: select *ZPRED* for X axis and *SDRESID* for Y axis; select *Histogram*.

- Does the histogram look somewhat normal?
- Is the scatter plot an oblong cloud that’s symmetric around a horizontal line running through zero on the Y axis?

5. Interactions:

SPSS: Graphs: Scatter: Simple: select appropriate X and Y variables; Set markers by: *variable that defines the subgroups of interest*. Double click on graph until in Chart Editor; **Chart: Options:** Show subgroups; Fit line: Subgroups.

- Is the relationship between X and Y the same for all values of the *subgrouping variable*?

If YES: then there is **no** modification of the effect of X on Y by subgroup; ie, **no interaction**.

If NO: then there **is** a modification of the effect of X on Y by subgroup; ie, **interaction is present**.