More Multiple Regression

Approaches to Regression Analysis, Types of Correlations and Advanced Regression

Types of Regression Analysis

Standard Regression

- Standard or Simultaneous Regression
  - Put all of the predictors in at one time and the coefficients are calculated for all of them controlling for all others
  - Method equals enter in SPSS Sequential

Forward Sequential

- What does a predictor add to the prediction equation, over and above the variables already in the equation?
- You think the X1 is a more important predictor and your interest in X2 is what does it add to the X1 -> Y prediction
- Real Forward Sequential in SPSS is setting it to Enter and using the blocks function (user specified)

Statistical Forward Sequential

- starts with Y'=a, all potential predictors are assessed and compared to an Entry Criterion; the variable with the lowest F probability ($p<.05$) enters it into the equation
- Remaining predictors are re-evaluated given the new equation ($Y'=a + X_{\text{first entered}}$) and the next variable with the lowest probability enters, etc…
- This continues until either all of the variables are entered or no other variables meet the entry criterion.
- Once variables enter the equation they remain.
- Method equals Forward in SPSS

Backward Sequential

- Can predictors be removed from an equation without hurting the prediction of Y? In other words, can a prediction equation be simplified?
- You know there are a set of predictors of a certain variable and you want to know if any of them can be removed without weakening the prediction
- In SPSS put all predictors in block one method equals enter, in block 2 any variables you want removed method equals removed, etc…

Statistical backward sequential

- All variables entered in and then each are tested against an Exit Criteria; F probability is above a set criteria (p>.10).
- The variable with the worst probability is then removed.
- Re-evaluation of remaining variables given the new equation and the next variable with the worst probability is then removed.
- This continues until all variables meet the criteria or all variables removed.
- In SPSS this is setting method equals backward.

Stepwise (Purely Statistical Regression)

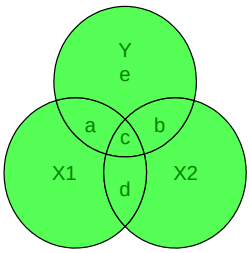
- at each step of the analysis variables are tested for both entry and exit criteria.
- Starts with intercept only then tests all of the variables to see if any match entry criteria.
- Any matches enter the equation
- The next step tests un-entered variables for both entry and entered variables for exit criteria, and so on…

Stepwise

- This cycles through adding and removing variables until none meet the entry or exit criteria
- Variables can be added or removed over and over given the new state of the equation each time.
- Considered a very post-hoc type of analysis and is not recommended

Correlations and Effect size

Ballantine



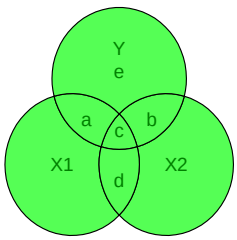
- Regular Correlation (Zero – Order, Pearson)

$$r_{y1}^2 = a + c$$

$$r_{y2}^2 = b + c$$

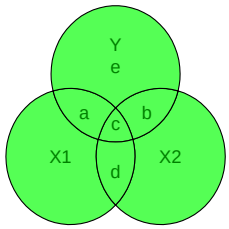
$$r_{12}^2 = c + d$$

Standard Regression



- Partial Correlation
  - correlation between Y and X1 with the influence of X2 removed from both
  - Yres, X1res
  - area a/(a + e) for x1 and b/(b + e) for x2 in the ballantine

Semipartial or Part Correlation

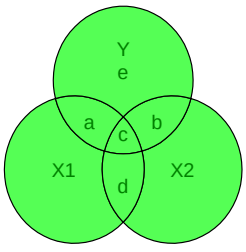


- correlation between Y and X1 with the influence of X2 removed from X1 only
- Y, X1res
- area a/(a + b + c + e) for x1 and b/(a + b + c +e) for x2

Semipartials and Bs

- Bs and semipartials are very similar
- B is the amount of change in Y for every unit change in X, while controlling for other Xs on Xi.
- Semipartials are measures of the relationship between Y and Xi controlling for other Xs on Xi.

Sequential



- Assuming x1 enters first
- The partial correlations would be
  (a + c)/(a + c + e) for x1 and unchanged for x2
- The part correlation would be
  (a + c)/(a + b + c + e) for x1 and x2 is unchanged.

Advanced Regression

Moderation, Mediation and Curve Estimation

Centering the data

- If you want to include powers, Moderation (interactions) or mediation you should first center the data
- Subtract the mean from every score
- You don't need to standardize by dividing by the SD
- This helps form creating multicollinearity in the data

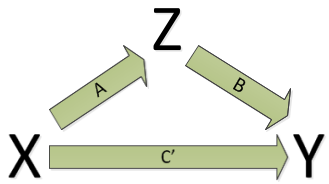
Moderation (interaction)

- Testing for moderation can be accomplished by simply cross multiplying the variables and adding the new variable in as another predictor
- If A and B are predictors of Y
  - First Center A and B separately (if they don't already have a meaningful zero)
  - Multiply the Centered A and B variables to create AB
  - Use A, B and AB as predictors of Y
  - If the slope predicting Y from AB is significant than A moderates B and vice versa (i.e., there is an interaction)

Mediation

- Regression can be used to test if a mediating effect is present in the data
- Defined - a given variable functions as a mediator to the extent that it accounts for the relation between a predictor and an outcome variable
- Often though of as an indirect effect of one variable on another.
  - X predicts Y through Z

Mediation



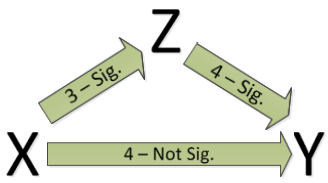
- C is the total effect of X on Y
- A*B is the indirect effect
- C' is the direct effect

Mediation

- 4 steps to establishing mediation (Baron and Kenny/ Regression Method)
  1. Establish x predicts y significantly
  2. Establish z predicts y significantly
  3. Establish x predicts z significantly
  4. Establish that x no longer predicts y when both x and z are in the prediction (C' is zero or at least non-significant)
- Partial Mediation – steps 1-3 are the same but in step 4 C' is less than C but still significant

Mediation

- Baron and Kenny



Mediation

- Sobel Method – Indirect Effect

$$Z_{Sobel} = \frac{a*b}{\sqrt{a^2 s_b^2 + b^2 s_a^2}}$$

- Where a and b are the unstandardized regression coefficients for paths a and b
- And s_a and s_b are the standard errors for paths a and b

Powers

- Even though we're talking about linear regression the equations can be altered to account for curves and interactions between variables
- Adding squares, cubes, etc. to account for curves in the relationship
- If you think X can predict an curved Y simply square X and add X^2 as an additional predictor