

## The Language of Statistical Inference

1. What are you testing? What does it mean to say something is statistically significant?
  - a. Testing a relationship between a dependent variable and one or more independent variables.
  - b. Seeing if there is covariation between the dependent variable and the independent variables.
  - c. Seeing if categorizing by an independent variable helps us better to explain the world (the 100%).
  - d. *“The probability that our observed data would show the relationship that we observe if our hypothesis is true” (or alternatively, is the null hypothesis of no effect true) (reading, pg. 339) (refer back to Z-score on previous handout)*
  
2. For a Difference of Means (sometimes called “t-test”)
  - a. Used when you have a continuous or interval dependent variable and a categorical, nominal, or “dummy” independent variable.
    - i. (Note: can use a 0,1 dependent variable, such as vote/non-vote, where the mean is a proportion)
    - ii. Test statistic: Z or T-score
    - iii. Result: *The difference between men and women on ratings of Gore and Bush is statistically significant at the .05 level.*
  
3. For a Crosstab
  - a. Used when you have two or more categorical variables
    - i. Test statistic: Chi-square
    - ii. Result: *My hypothesis that variable X and variable Y are related fails to be rejected at the .05 level.*
  
4. For a Regression
  - a. Used when you are testing a multivariate relationship with a continuous dependent variable
    - i. Test statistic: overall, R-squared tells you what proportion of variance in the dependent variable is explained by your model.
    - ii. Test statistic: for each variable, a t-statistic is provided for each “b” or slope, tests whether  $b=0$  or  $b\neq 0$ .
    - iii. Result: *My model explains X% of the variance in my dependent variable. A change of one unit in variable 1 is associated with  $b \cdot X$  change in Y, and this effect is significant at the .05 level. Etc.*