

1. (Nominal scale variables in boldface).
nested ANOVA

$$\underline{Y} = \beta_0 + \beta_1 \mathbf{X1} + \beta_2 \mathbf{X2}(\mathbf{X1}) + \epsilon$$

t-test

$$\underline{Y} = \beta_0 + \beta_1 \mathbf{X1} + \epsilon$$

Regression

$$\underline{Y} = \beta_0 + \beta_1 \underline{X1} + \epsilon$$

Analysis of Covariance

$$\underline{Y} = \beta_0 + \beta_1 \mathbf{X1} + \beta_2 \underline{X2} + \beta_{1*2} \mathbf{X1} \underline{X2} + \epsilon$$

2. C5 shows residuals
residuals invariant? no
No, because residuals show pattern of zero values at low t, high values at intermediate t, and low values at high t.
3. ST is correct. It has dimensions of $M^0 L^0 T^0$
ST* is not correct, it has dimensions of $M^{-1} L^3 T^0$
4. For a thesis, all three are appropriate.
For newspaper, use verbal models, diagrams only sparingly.
5. histogram, nscores, fit to normal distribution (rootogram).
1. Use randomization to obtain better estimate of p-value. 2. Use transformation.
New model sometimes uninterpretable (e.g. arcsin transformation).
6. $L = \underline{11}$ R after $L = \underline{5}$ $E(R \text{ after } L) = (11)(0.5) = \underline{5.5}$
Bootstrap by
-sampling the 21 observations with replacement
-computing observed statistic, R.L = number of R after L
-repeat at least 500 times to get good estimate of distribution R.L
-identify the 5% extreme values in the distribution (2.5% in each tail).
-report these values.
7. $D = 11$ grams

(nominal scale variables in bold)

Analysis A $\underline{W} = \beta_0 + \beta_{ST} \mathbf{ST} + \epsilon$
 Analysis B $\underline{W} = \beta_0 + \beta_{ST} \mathbf{ST} + \beta_{YR} \mathbf{YR} + \epsilon$
 Analysis C $\underline{W} = \beta_0 + \beta_{ST} \mathbf{ST} + \beta_{YR} \mathbf{YR} + \epsilon$

Analysis B is better than A because Type II error will be lower. Removing the year effect reduces the MSerror, increasing the power of the test to detect a difference.

Analysis C is no improvement over B because there is no further reduction in MSerror