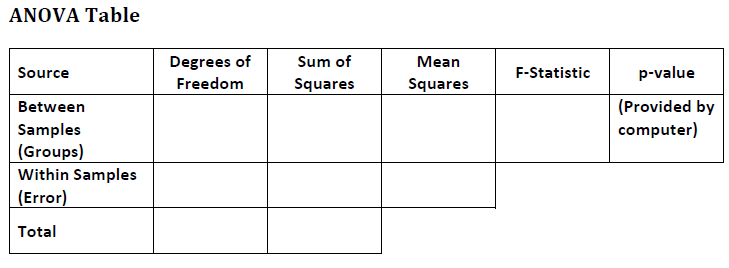
Lecture 8 Inference for More than Two Groups (ANOVA & Chi Square)

8.1 ANOVA

* Sum of Squares: 
  + Measures how far each observation from grand mean
* Sum of Squared Errors: 
  + Measures how far each observation from its group mean
  + 
* Sum of Squared Groups: 
  + How far each sample mean from grand mean
  + SStreatment or SSmodel
* Dv = n – 1 (pg4)
* Mean Squares
  + Variability between groups: 
  + Variability within groups: 



SSG + SSE

n - t

SSE/n -t

SSE

n - t

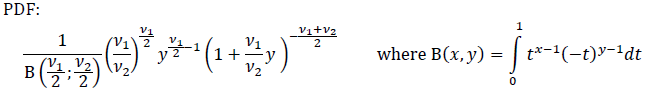
MSG/MSE

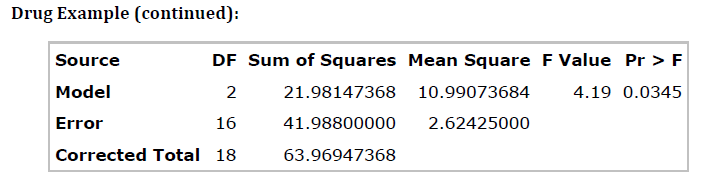
SSG/t-1

SSG

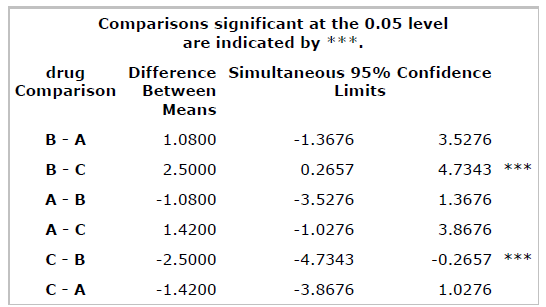
t-1

8.2 Hypothesis Testing for ANOVA (pg5)

* Use: quantitative r.v. and categorical expl variable w/>2 categories
  + Or comparing the means of more than two pop
* Type of test: ANOVA F-Test
* Conditions: random, independent, pop normal or CLT, pop variances are =
* Test Stat: 
  + Test stat represents the ratio of variability that can be explained by groups to the amount that cannot
* Null Dist: F – dist
  + Numerator df = t-1
  + Denominator df = n – t
  + No need to concern about 2 sided vs 1 side
* F Distribution
  + Models ratio of two r.v. that follow Chi Square dist
  + Used often for inference on more than 2 pop means
  + Generally right skewed (values of F Stat cannot be negative)
* 
* V1 =numerator dv = t – 1
* V2 = denominator dv = n – t
* Mean of F Dist: 
* Vari of F Dist: 
* Conditions: random, independent, normal (reasonably symmetric), equal population variacnes
  + Normal Populations: populations look symmetric on box plot
  + Equal Pop Variances: check the range on boxplots to be similar



8.3 ANOVA Multiple Comparisons (pg11)



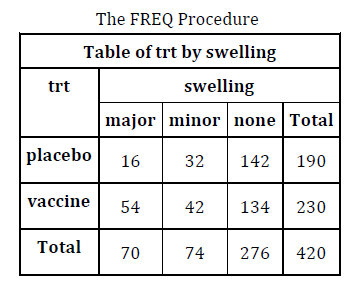
* Only need half the table since the other half is redundant
* If CI does not contain 0, it means the samples are significantly different from each other
  + Can see B and C are significantly different

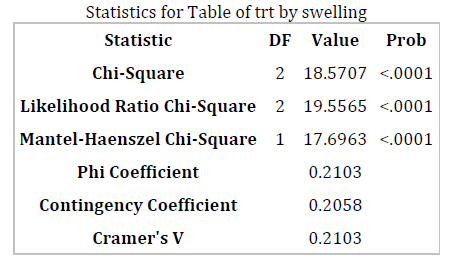
8.4 Intro to Chi Square

* Goodness of Fit: if a model fits a variable in a population or distribution
  + 1 Variable, 1 Pop
  + Extension of 1 sample z test for pop prop
* Test of Homogeneity: if distribution is the same between two or more pop
  + 1 Variable, 2 pop
  + Extension of 2 sample t test, for proportions
* Test of Independence: to check for a relationship between variable in a single pop
  + 2 variables, 1 pop
* All based on Chi Square Statistic
* Chi Square Dist
  + Generally skewed right
  + Mean: v (degrees of freedom)
  + Variance: 2v (degrees of freedom)

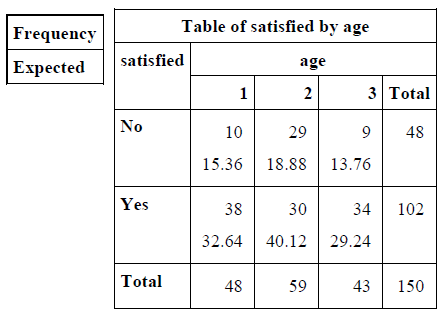
8.5 3 Types of Chi Square Tests

* Goodness of Fit
  + Test Stat: 
  + Null Dist: Chi Square dist df = k – 1
* Homogeneity
  + Test Stat: 
  + Eij = (Row Total X Column Total) / n
  + Null Dist: Chi Square dist df = (R – 1)(C – 1)
  + Contingency table:





* Chi – Square Value = 18.5707
  + Relative distance between observed and expected count = 18.5707
  + On average should be 2
  + .0001
    - Area under the Chi Square dist more extreme
    - DF = 2 = (R-1)(C-1)
* Independence
  + 2 variables, 1 pop
  + Conditions: random, 80% of expected count >5 and none <1
  + Test Stat: 
  + Null Dist: Chi Sq w/ df = (R-1)(C-1)



8.6 Multiple Comparisons (pg24)

meh