Week8: nested f-test

Nested F-test

- A way to test if multiple predictors taken together are adding to the usefulness of the model
- The key is to compare a full (larger) model with a reduced model that eliminates the group of predictors we are interested in testing.
  - A good way to check if you need higher order terms.
  - To test if a categorical variable with more than two levels is needed

Dataset: cigarette content

- The Federal Trade Commission ranks cigarettes according to tar, nicotine, and carbon monoxide contents – three substances thought to be hazardous to health. Past studies have show that increases in tar and nicotine are accompanied by an increase in carbon monoxide. They also recorded weight for 25 brands of cigarettes.

$CarbonMonoxide = \beta_0 + \beta_1 X_{tar} + \beta_2 X_{weight} + \varepsilon$

- Lets compare this to a complete second order model

$CarbonMonoxide = \beta_0 + \beta_1 X_{tar} + \beta_2 X_{weight} + \beta_3 X_{tar3} X_{weight} + \beta_4 X^2_{tar} + \beta_5 X^2_{weight} + \varepsilon$

Nested F-test: nested model

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.958[a]	.917	.909	1.4277

a. Predictors: (Constant), weight, tar
b. Dependent Variable: co

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	494.306	2	247.153	121.251	.000[b]
	Residual	44.844	22	2.038		
	Total	539.150	24			

a. Dependent Variable: co
b. Predictors: (Constant), weight, tar

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	3.114	3.416		.912	.372		
	tar	.804	.059	.961	13.622	.000	.759	1.317
	weight	-.423	3.813	-.008	-.111	.913	.759	1.317

a. Dependent Variable: co

Nested F-test: full model

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.974[a]	.949	.936	1.2025

a. Predictors: (Constant), weightsq, tarsq, tar, weight, tarweight
b. Dependent Variable: co

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
	Regression	511.676	5	102.335	70.769	.000[b]
1	Residual	27.475	19	1.446		
	Total	539.150	24			

a. Dependent Variable: co
b. Predictors: (Constant), weightsq, tarsq, tar, weight, tarweight

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-6.462	31.789		-.203	.841
	tar	1.444	.930	1.726	1.552	.137
	weight	10.819	70.742	.200	.153	.880
	tarweight	-.386	1.267	-.517	-.304	.764
	tarsq	-.009	.014	-.325	-.655	.520
	weightsq	-3.021	40.311	-.111	-.075	.941

a. Dependent Variable: co

Compare overall model statistics

- Coefficient of determination

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.958[a]	.917	.909	1.4277

a. Predictors: (Constant), weight, tar
b. Dependent Variable: co

Nested model

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.974[a]	.949	.936	1.2025

a. Predictors: (Constant), weightsq, tarsq, tar, weight, tarweight
b. Dependent Variable: co

Full model

Nested F-test

H_0: $\beta_i = 0$ for all predictors in the subset
H_a: $\beta_i \neq 0$ for at least one predictor in the subset

- Test statistic:

$$F = \frac{(SSModel_{full} - SSModel_{nested}) / \text{\# predictors tested}}{SSE_{full}/(n-k-1)}$$

k=number of predictors in Full model
# of predictors tested = number
Dropped from full model to make nested

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	494.306	2	247.153	121.251	.000[b]
	Residual	44.844	22	2.038		
	Total	539.150	24			

a. Dependent Variable: co
b. Predictors: (Constant), weight, tar

Nested model

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	511.676	5	102.335	70.769	.000[b]
	Residual	27.475	19	1.446		
	Total	539.150	24			

a. Dependent Variable: co
b. Predictors: (Constant), weightsq, tarsq, tar, weight, tarweight

Full model

Nested F-test

k=number of predictors in Full model
# of predictors tested = number dropped from full model to make subset

$$F = \frac{(SSModel_{full} - SSModel_{nested}) / \text{\# predictors tested}}{SSE_{full}/(n-k-1)}$$

$$F = \frac{(511.7 - 494.3)/3}{27.5/(25-5-1)} \qquad F = \frac{5.8}{1.4} = 4.1$$

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	494.306	2	247.153	121.251	.000[b]
	Residual	44.844	22	2.038		
	Total	539.150	24			

a. Dependent Variable: co
b. Predictors: (Constant), weight, tar

Nested model
$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	511.676	5	102.335	70.769	.000[b]
	Residual	27.475	19	1.446		
	Total	539.150	24			

a. Dependent Variable: co
b. Predictors: (Constant), weightsq, tarsq, tar, weight, tarweight

Full model
$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \beta_4 x_1^2 + \beta_5 x_2^2 + \varepsilon$

- Look up pvalue using an F-distribution
  - With numerator degrees of freedom equal to the number of predictors being tested,
  - denominator degrees of freedom equal to the error degrees of freedom for the full model
- Nice table: http://www.socr.ucla.edu/applets.dir/f_table.html
- numerator=3, denominator=19 -> critical value =3.1

Nested F-test

$F = \frac{5.8}{1.4} = 4.1$ > 3.1, so p<.05

Data suggests one of the tested betas is useful, or some combination of them is, in predicting average CO content, after accounting for tar and weight.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	494.306	2	247.153	121.251	.000[b]
	Residual	44.844	22	2.038		
	Total	539.150	24			

a. Dependent Variable: co
b. Predictors: (Constant), weight, tar

Nested model
$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	511.676	5	102.335	70.769	.000[b]
	Residual	27.475	19	1.446		
	Total	539.150	24			

a. Dependent Variable: co
b. Predictors: (Constant), weightsq, tarsq, tar, weight, tarweight

Full model
$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \beta_4 x_1^2 + \beta_5 x_2^2 + \varepsilon$

Nested F-test: full model

- Something useful in higher order terms even though none of the individual betas are significant.
  - Can try taking them out one by one
- So why not just start that way?
  - Too many tests increase risk of type I error.

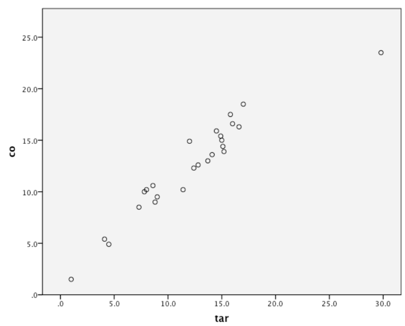
Coefficients[a]

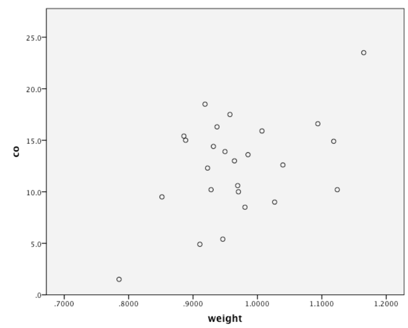
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-6.462	31.789		-.203	.841
	tar	1.444	.930	1.726	1.552	.137
	weight	10.819	70.742	.200	.153	.880
	tarweight	-.386	1.267	-.517	-.304	.764
	tarsq	-.009	.014	-.325	-.655	.520
	weightsq	-3.021	40.311	-.111	-.075	.941

a. Dependent Variable: co

Final model?

- Can usually get good idea if we need a quadratic term from graphs
- Start with?
  - Interaction term





Final model

- Could still test square terms, but probably would stop

$Y = -10.3 + 1.8x_{tar} + 13.2x_{weigt} -1.1\ \beta_3 x_{tar} x_{weight}$

- What is the direction of the relationship between tar and CO?
  - Depends on the value of weight, due to the higher order

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-10.320	4.707		-2.192	.040
	tar	1.884	.310	2.252	6.085	.000
	weight	13.220	4.949	.245	2.671	.014
	tarweight	-1.075	.305	-1.442	-3.530	.002

a. Dependent Variable: co

Summary

- Nested F is a way to test the significance of several predictors at once
- Can help us avoid type I errors that might occur when you run a bunch of individual tests

Exam!

- I want to kill Canvas!!!!

ANOVA

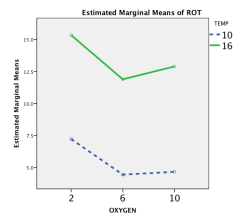
- Farmers were interested in the amount of rot on potatoes grown in different environments. They randomly selected several potato plants and assigned them to several environments that varied in the level of bacteria present (BACT: low, medium, high); temperature (TEMP: 10 degrees C or 16 degrees C); and percent saturation of oxygen level in the soil (OXYGEN: 2 %, 6%, 10%). First, they investigated whether the amount of rot varied for different levels of oxygen and temperature. The output is below.

Tests of Between-Subjects Effects

Dependent Variable: ROT

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	947.481[a]	5	189.496	5.169	.001
Intercept	4778.963	1	4778.963	130.368	.000
TEMP	848.074	1	848.074	23.135	.000
OXYGEN	97.815	2	48.907	1.334	.273
TEMP * OXYGEN	1.593	2	.796	.022	.979
Error	1759.556	48	36.657		
Total	7486.000	54			
Corrected Total	2707.037	53			

a. R Squared = .350 (Adjusted R Squared = .282)



Tests of Between-Subjects Effects

Dependent Variable: ROT

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	947.481[a]	5	189.496	5.169	.001
Intercept	4778.963	1	4778.963	130.368	.000
TEMP	848.074	1	848.074	23.135	.000
OXYGEN	97.815	2	48.907	1.334	.273
TEMP * OXYGEN	1.593	2	.796	.022	.979
Error	1759.556	48	36.657		
Total	7486.000	54			
Corrected Total	2707.037	53			

a. R Squared = .350 (Adjusted R Squared = .282)

Are the main effects significant?

(11) H_0: $\mu_{10c} = \mu_{16c}$

H_a: $\mu_{10c} \neq \mu_{16c}$

For the main effect of TEMP, the pvalue is small ($p<.001$), and the F is large (23.135) so we can reject the null. This suggests that the mean amount of rot on potatoes differs depending on what temperature they were grown in. From the graph, we can see that mean rot levels are greater for higher temperatures.

H_0: $\mu_2 = \mu_6 = \mu_{10}$

H_a: the means are not all equal

For the main effect of OXYGEN, the p-value is large ($p=.273$) and F is small (1.33), we fail to reject the null hypothesis. This suggests that the mean amount of rot is similar for different saturation levels of OXYGEN, or it differs by too small an amount for us to detect with the current study.

Tests of Between-Subjects Effects

Dependent Variable: ROT

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	947.481[a]	5	189.496	5.169	.001
Intercept	4778.963	1	4778.963	130.368	.000
TEMP	848.074	1	848.074	23.135	.000
OXYGEN	97.815	2	48.907	1.334	.273
TEMP * OXYGEN	1.593	2	.796	.022	.979
Error	1759.556	48	36.657		
Total	7486.000	54			
Corrected Total	2707.037	53			

a. R Squared = .350 (Adjusted R Squared = .282)

– Is there a significant interaction? What are the implications?

(6) H_0: the main effect of each factor is the same for each level of the other factor

H_a: the two factors interact

Since the p-value is large (p=.979) and the F small (.022), we fail to reject the null hypothesis. This suggests that the main effect of temperature is the same for the different levels of oxygen saturation. This means that we can interpret the main effects using the current model

– Would you run posthoc tests on either of the main effects in the model, why or why not?

- (4) I would not run any posthoc tests. TEMP is the only significant main effect, and there are only two levels so a posthoc test is unnecessary. I would not run it on OXYGEN because it is nonsignificant.
- Posthoc – need hypotheses, p-values, and to think about what you want to know.

Regression/general

- P is never 0
- Need hypotheses to go with pvalues.
- Multiple regression: 'after accounting for other variables!' and 'average!'
  - H_0: $\beta_3=0$
  - H_a: $\beta_3 \neq 0$
  - Since the p-value is small (p<.001) and t is large (t=4.303), we reject the null hypothesis. This suggests that retail sales are useful in predicting the average sale price of a station after accounting for when it was built and the number of homes with TVs.

Regression/general

- Assumptions: not linearity!
- Units!!
- Dummy variables for categorical predictors