STAT 3119

Week 11: 11/7/2019 @GWU

Outline

- ANOVA Inferences when Treatment Means Are of Unequal Importance (ch 23.5)
- Recitation and Quiz

Two-factor studies with normal observations

We can use the standard two-way ANOVA factor effects model:

$$Y_{ijk} = \mu_{..} + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

- There are two factors: factor A with a levels and factor B with b levels
- There are ab treatment levels, with treatment mean μ_{ij}, and n_{ij}= sample size for the treatment consisting of the ith level of factor A and the jth level of factor B.

Unweights vs. weighted mean

- When treatment means μ_{ij} are equally important, one can make inferences on the unweighted factor level means $\mu_{i.} = \sum_j \mu_{ij}/b$, and $\mu_{.j} = \sum_i \mu_{ij}/a$.
- Sometimes, the treatment means μ_{ij} are not of equal importance, one may want to make inferences on weighted factor level or treatment level means with certain given weights. We can still use the formulas for linear combination L in Table 23.5 for estimation and testing.

Examples of making inferences on weighted means

Example 1: In a breakfast cereal study

- Factor A was type of sweetener (i = 1: corn syrup, i = 2: low-calorie sweetener) and factor B was consumer category (j = 1: child, j = 2: male adult, j = 3: female adult).
- The company wishes to determine if a change to a low-calorie sweetener will change the mean rating (Y) of its product in the population of consumers (i.e. compare the response between two levels for factor A).
- It is known: 60 percent of the consumers of this product were children, 20 percent male adults, and 20 percent female adults.
- Therefore, the treatment means μ_{ij} (i ~ for sweetener; j ~ for consumers) in the model for the study data have "unequal importance" (i.e. need to reweight the sample treatment means to represent the distribution of the custumers in the population) and the company therefore wishes to compare the two weighted means for factor A (sweetener):

$$L_{A1} = .6\mu_{11} + .2\mu_{12} + .2\mu_{13} \text{ (Corn syrup), vs}$$

$$L_{A2} = .6\mu_{21} + .2\mu_{22} + .2\mu_{23} \text{ (Low-calorie sweetener)}$$

That is, they can estimate the difference (or the contrast)

$$L = L_{A1} - L_{A2} = (.6\mu_{11} + .2\mu_{12} + .2\mu_{13}) - (.6\mu_{21} + .2\mu_{22} + .2\mu_{23})$$

or testing the hypothesis of

$$H_0 : L = 0$$

to determine if the mean ratings for the two sweeteners are the same. The weights of 0.6, 0.2 and 0.2 are decided based on the population of consumers from their market research (not from the study design).

Examples of using weighted means (2)

Example 2: Mathematics learning example in Table 19.11.

TABLE 19.11 Results—Mathematics Learning Example.

(a) Mean Learning Scores ($n = 21$)

Teaching Method i	Quantitative Ability (j): Excellent	Good	Moderate
Abstract	92 ($\bar{Y}_{11\cdot}$)	81 ($\bar{Y}_{12\cdot}$)	73 ($\bar{Y}_{13\cdot}$)
Standard	90 ($\bar{Y}_{21\cdot}$)	86 ($\bar{Y}_{22\cdot}$)	82 ($\bar{Y}_{23\cdot}$)

(b) ANOVA Table

Source of Variation	*SS*	*df*	*MS*
Factor *A* (teaching methods)	504	1	504
Factor *B* (quantitative ability)	3,843	2	1,921.5
AB interactions	651	2	325.5
Error	3,360	120	28
Total	8,358	125	

A junior college system studied the effects of teaching method (factor A) and student's quantitative ability (factor B) on learning of college mathematics.

- Teaching (factor A): standard and abstract , 2 levels
- Ability (factor B) determined from a standard test: excellent, good, or moderate, 3 levels
- $n = 21$ at each combination of Teaching and Ability combination.

Now:

1. A school administrator requested information about which teaching method leads to better learning of college mathematics when 20 percent of the students in the class have excellent quantitative ability, 50 percent have good ability, and 30 percent have moderate ability.
2. The treatment means μ_{ij} (i ~ for teaching method; j ~ for students) would have unequal importance, i.e., we need to reweight the μ_{ij} to represent the certain distribution of students with different levels of ability.
3. The mean learning scores for such a class mix with the two teaching methods are the following linear combinations:

$$L_{A1} = .2\mu_{11} + .5\mu_{12} + .3\mu_{13} \text{ (standard teaching), vs}$$

$$L_{A2} = .2\mu_{21} + .5\mu_{22} + .3\mu_{23} \text{ (abstract teaching)}$$

The difference between the means for the two teaching methods is the contrast:

$$L = L_{A1} - L_{A2}$$

With the experiment data table, we can calculate the estimate and CI for L. We can use R as a calculator.

```
# mij
meanA1 = c(92, 81, 73 )
meanA2 = c(90, 86, 82 )

# get the estimate of L
weight= c(.2, .5, .3)

(L.A1 = sum(weight*meanA1))
```

```
## [1] 80.8
```

```
(L.A2 = sum(weight*meanA2))
```

```
## [1] 85.6
```

```
(L= L.A1- L.A2 )
```

```
## [1] -4.8
```

```
# get the estimated variance, MSE use to estimate sigma^2
MSE= 28
(s.square = MSE/21*( sum(weight^2+ weight^2 )))
```

```
## [1] 1.013333
```

```
(sd= sqrt(s.square))
```

```
## [1] 1.006645
```

```
# t-value, df= 21*6-6 =120
tvalue= qt(.975, 120)

# confidence interval
(LCI= L- tvalue*sd )
```

```
## [1] -6.793086
```

```r
(UCI= L+ tvalue*sd )
```

```
## [1] -2.806914
```

Conclusion: With 95 percent confidence we conclude that the standard teaching method is better for the specified class mix, leading to a mean learning score that is at least 2.81 points greater than that for the abstract teaching method and may be as much as 6.79 points greater.

Tests weighted means by Equivalent Regression Models

Example 3: In the growth hormone example in Table 23.1

TABLE 23.1 **Sample Data and Notation—Growth Hormone Example (growth rate difference in centimeters per month).**

	Bone Development (factor B) j		
Gender (factor A) i	**Severely Depressed (B_1)**	**Moderately Depressed (B_2)**	**Mildly Depressed (B_3)**
Male (A_1)	1.4 (Y_{111}) 2.4 (Y_{112}) 2.2 (Y_{113})	2.1 (Y_{121}) 1.7 (Y_{122})	.7 (Y_{131}) 1.1 (Y_{132})
Mean	2.0 ($\bar{Y}_{11\cdot}$)	1.9 ($\bar{Y}_{12\cdot}$)	.9 ($\bar{Y}_{13\cdot}$)
Female (A_2)	2.4 (Y_{211})	2.5 (Y_{221}) 1.8 (Y_{222}) 2.0 (Y_{223})	.5 (Y_{231}) .9 (Y_{232}) 1.3 (Y_{233})
Mean	2.4 ($\bar{Y}_{21\cdot}$)	2.1 ($\bar{Y}_{22\cdot}$)	.9 ($\bar{Y}_{23\cdot}$)

- The investigator was interested in the effects of a child's gender (factor A) and bone development (factor B: severely depressed, moderately depressed, mildly depressed) on the rate of growth induced by hormone administration.

Now:

1. It is known that twice as many male as female children undergo growth hormone treatment therapy, for each given bone development group.
2. The treatment means μ_{ij} (i ~ for gender; j ~ for bone development state) would have unequal importance, i.e., we want to reweight the μ_{ij} based on the gender distribution for the target population with the given problem.
3. Desired Inference: they wish to test whether or not the state of bone development affects the change in growth rate in the target population. Then for each bone development (factor B) level, the estimated mean for target population would be

$$L_{B_j} = (2\mu_{1j} + \mu_{2j})/3$$

Then the hypotheses to be tested are as follows: (To compare L_{B_1} vs. L_{B_2} vs. L_{B_3})

$$H_0: \frac{2\mu_{11} + \mu_{21}}{3} = \frac{2\mu_{12} + \mu_{22}}{3} = \frac{2\mu_{13} + \mu_{23}}{3} \tag{23.37}$$

H_a: not all equalities hold

We restate the alternative H_0 in the following equivalent fashion:

$$H_0: \begin{cases} \dfrac{2\mu_{11} + \mu_{21}}{3} - \dfrac{2\mu_{12} + \mu_{22}}{3} = 0 \\ \dfrac{2\mu_{11} + \mu_{21}}{3} - \dfrac{2\mu_{13} + \mu_{23}}{3} = 0 \end{cases} \tag{23.37a}$$

Example 3 (continued)

To test these hypothesis, we can compare a full regression model for a two-factor study and a reduced model (assuming H_0 is true), then use anova() to compare the SS difference with a F-test.

Step 1: We fit the full equivalent regression model corresponding to the cell mean models for two-way ANOVA.

In such regression model, we use six indicator functions to indicate the six treatment levels (two genders, 3 bone groups). The regression model (23.38) is used without intercept

$$\begin{aligned} Y_{ijk} = \mu_{11}X_{ijk1} + \mu_{12}X_{ijk2} + \mu_{13}X_{ijk3} + \mu_{21}X_{ijk4} \\ + \mu_{22}X_{ijk5} + \mu_{23}X_{ijk6} + \varepsilon_{ijk} \qquad \text{Full model} \end{aligned} \tag{23.38}$$

where:

$$X_1 = \begin{cases} 1 & \text{if case from level 1 of factor } A \text{ and level 1 of factor } B \\ 0 & \text{otherwise} \end{cases}$$

$$X_2 = \begin{cases} 1 & \text{if case from level 1 of factor } A \text{ and level 2 of factor } B \\ 0 & \text{otherwise} \end{cases}$$

$$\vdots$$

$$X_6 = \begin{cases} 1 & \text{if case from level 2 of factor } A \text{ and level 3 of factor } B \\ 0 & \text{otherwise} \end{cases}$$

A) We follow the same steps/code to read the data Ex23 from Lecture 11A.

```
Ex23  =read.table(
        url("https://raw.githubusercontent.com/npmldabook/Stat3119/master/Week-11/CH23TA01.txt"))

names(Ex23) =  c("response", "Gender", "Bone", "units")

# get the  treatment variable to use for estimate contrasts
Ex23$Treament = Ex23$Gender*(Ex23$Bone+2)
```

```r
# make  categorical variables for factor A and B
Ex23$Gender =  as.factor(Ex23$Gender)
Ex23$Bone =  as.factor(Ex23$Bone)
Ex23$Treament =  as.factor(Ex23$Treament)

levels(Ex23$Gender) = c("M", "F")
levels(Ex23$Bone) = c("Severe", "Moderate", "Mild")
levels( Ex23$Treament) <- c("A1B1","A1B2","A1B3", "A2B1", "A2B2", "A2B3")

# check frequencies
xtabs(~ Gender+ Bone, data=Ex23 )
```

```
##       Bone
## Gender Severe Moderate Mild
##      M      3        2    2
##      F      1        3    3
```

```r
table( Ex23$Treament)
```

```
##
## A1B1 A1B2 A1B3 A2B1 A2B2 A2B3
##    3    2    2    1    3    3
```

B) Then we can define the indicator functions.

```r
Ind1 =  (Ex23$Gender =="M" &  Ex23$Bone=="Severe")*1     #A1 B1
Ind2 =  (Ex23$Gender =="M" &  Ex23$Bone=="Moderate")*1   #A1 B2
Ind3 =  (Ex23$Gender =="M" &  Ex23$Bone=="Mild")*1       #A1 B3

Ind4 =  (Ex23$Gender =="F" &  Ex23$Bone=="Severe")*1     #A2 B1
Ind5 =  (Ex23$Gender =="F" &  Ex23$Bone=="Moderate")*1   #A2 B2
Ind6 =  (Ex23$Gender =="F" &  Ex23$Bone=="Mild")*1       #A2 B3

# Full models without intercept
LM.full =  lm( response~ 0+ Ind1 + Ind2 + Ind3+Ind4 + Ind5 + Ind6, data=Ex23)

summary(LM.full)
```

```
##
## Call:
## lm(formula = response ~ 0 + Ind1 + Ind2 + Ind3 + Ind4 + Ind5 +
##     Ind6, data = Ex23)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##   -0.6   -0.2    0.0    0.2    0.4
##
## Coefficients:
##      Estimate Std. Error t value Pr(>|t|)
## Ind1   2.0000     0.2327   8.593 2.60e-05 ***
## Ind2   1.9000     0.2850   6.666 0.000158 ***
```

```
## Ind3   0.9000     0.2850   3.157 0.013447 *
## Ind4   2.4000     0.4031   5.954 0.000341 ***
## Ind5   2.1000     0.2327   9.023 1.82e-05 ***
## Ind6   0.9000     0.2327   3.867 0.004761 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4031 on 8 degrees of freedom
## Multiple R-squared:  0.9702, Adjusted R-squared:  0.9478
## F-statistic: 43.34 on 6 and 8 DF,  p-value: 1.134e-05
```

Step 2: We fit a reduced model assuming H_0 is true.

Under H_0, we can derive the relationship of treatments from these equations, plug in the full models, collect the terms and obtain the reduced regression model. We rewrite the equations as follows

$$\begin{aligned} \mu_{21} &= 2\mu_{12} + \mu_{22} - 2\mu_{11} \\ \mu_{23} &= 2\mu_{12} - 2\mu_{13} + \mu_{22} \end{aligned} \qquad \textbf{(23.39)}$$

Replacing μ_{21} and μ_{23} in full model (23.38) by the expressions in (23.39), we obtain the reduced model:

$$\begin{aligned} Y_{ijk} = {} & \mu_{11}X_{ijk1} + \mu_{12}X_{ijk2} + \mu_{13}X_{ijk3} + (2\mu_{12} + \mu_{22} - 2\mu_{11})X_{ijk4} \\ & + \mu_{22}X_{ijk5} + (2\mu_{12} - 2\mu_{13} + \mu_{22})X_{ijk6} + \varepsilon_{ijk} \end{aligned}$$

This model can be simplified algebraically, as follows:

$$Y_{ijk} = \mu_{11}Z_{ijk1} + \mu_{12}Z_{ijk2} + \mu_{13}Z_{ijk3} + \mu_{22}Z_{ijk4} + \varepsilon_{ijk} \qquad \text{Reduced model} \qquad \textbf{(23.40)}$$

where:

$$\begin{aligned} Z_{ijk1} &= X_{ijk1} - 2X_{ijk4} \\ Z_{ijk2} &= X_{ijk2} + 2X_{ijk4} + 2X_{ijk6} \\ Z_{ijk3} &= X_{ijk3} - 2X_{ijk6} \\ Z_{ijk4} &= X_{ijk4} + X_{ijk5} + X_{ijk6} \end{aligned}$$

Now we have four new covariates in the regression model, we can fit the reduced model as follows.

```
Ind.new1= Ind1 - 2*Ind4
Ind.new2= Ind2 + 2*Ind4 +2*Ind6
Ind.new3= Ind3 - 2*Ind6
Ind.new4= Ind4 + Ind5 + Ind6

LM.reduced =  lm( response~ 0+ Ind.new1 + Ind.new2 + Ind.new3+ Ind.new4, data=Ex23)

summary(LM.reduced)
```

```
##
```

```
## Call:
## lm(formula = response ~ 0 + Ind.new1 + Ind.new2 + Ind.new3 +
##     Ind.new4, data = Ex23)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1857 -0.2964  0.0500  0.5714  0.8000
##
## Coefficients:
##          Estimate Std. Error t value Pr(>|t|)
## Ind.new1   1.6000     0.3132   5.108 0.000458 ***
## Ind.new2   1.5143     0.3192   4.744 0.000787 ***
## Ind.new3   1.8857     0.3192   5.908 0.000149 ***
## Ind.new4   1.9714     0.3787   5.206 0.000398 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6895 on 10 degrees of freedom
## Multiple R-squared:  0.8909, Adjusted R-squared:  0.8472
## F-statistic: 20.41 on 4 and 10 DF,  p-value: 8.447e-05
```

Then run model comparisons to test our H_0,

```r
anova( LM.reduced, LM.full)
```

```
## Analysis of Variance Table
##
## Model 1: response ~ 0 + Ind.new1 + Ind.new2 + Ind.new3 + Ind.new4
## Model 2: response ~ 0 + Ind1 + Ind2 + Ind3 + Ind4 + Ind5 + Ind6
##   Res.Df    RSS Df Sum of Sq      F  Pr(>F)
## 1     10 4.7543
## 2      8 1.3000  2    3.4543 10.629 0.00559 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Results: To compare the RSS from the two models, F-test statistic= 10.629, following a $F(2, 8)$ distribution, with a p-value=0.00559, thus we reject H_0 and concluded that weighted mean change in the growth rate is not the same for the three bone development groups.

Example 3 (continued): use multiple comparison procedures

Instead of the above testing hypothesis of $H_0 : L_{B_1} - L_{B_2} = 0$ and $L_{B_1} - L_{B_3} = 0$, we can also use estimation approaches. We then need to estimate the two contrast

$$L_1 = L_{B_1} - L_{B_2} = (2\mu_{11} + \mu_{21})/3 - (2\mu_{12} + \mu_{22})/3$$

and

$$L_2 = L_{B_1} - L_{B_3} = (2\mu_{11} + \mu_{21})/3 - (2\mu_{13} + \mu_{23})/3$$

with a multiple comparison procedure (e.g., the Bonferroni procedure) and noting whether or not both confidence intervals include zero or testing $L_1 = 0$ and $L_2 = 0$ adjusted for multiple testing.

In our previous example for Chapter 17 on contrasts (Lecture 3A), we tried to use **contrast()** function in **emmeans** package to estimate and test the linear combinations of treatment means. We can also apply here for the weighted means (a special case of the linear combination). *Note*: The **multcomp** package with **glht()** functions can run similar comparisons.

```
library(emmeans)
fit = lm( response ~ Treament , data= Ex23)
Est.mean<- emmeans(fit, ~ Treament)

#Set the c1-c6 corresponding to treatment levels
L = list( L1= c( 2/3, -2/3,  0,  1/3,  -1/3, 0),
          L2 = c(2/3,  0 , -2/3, 1/3,  0, -1/3))

contrast(Est.mean, L, adjust='bonferroni')
```

```
##  contrast estimate   SE df t.ratio p.value
##  L1          0.167 0.29  8 0.574   1.0000
##  L2          1.233 0.29  8 4.249   0.0056
##
## P value adjustment: bonferroni method for 2 tests
```

```
confint(contrast(Est.mean, L, adjust='bonferroni'))
```

```
##  contrast estimate   SE df lower.CL upper.CL
##  L1          0.167 0.29  8   -0.632    0.965
##  L2          1.233 0.29  8    0.435    2.032
##
## Confidence level used: 0.95
## Conf-level adjustment: bonferroni method for 2 estimates
```

Results: Based the multiple comparisons for the two contrasts (testing or CI), we can conclude that the weight mean for B1 and B2 are not significantly different but the weight mean for B1 and B3 are significantly different.

Note: Tests for Factor Main Effects by Use of Matrix Formulation (p 975-976) is an alternative way to perform the test for a number of contrasts. It provides the same results as fitting and comparing two regression models. (*We skip this section and the matrix formulation to test contrasts will not be tested in the final exam.*)

Tests for Factor Effects: Weights are proportional to n_{ij}

Sometimes, the weights of the treatment means may be based on the sample size n_{ij} for the treatment levels. In some situations, for the given research questions, it may be useful to use the weights proportional to the n_{ij}, then a treatment with a large sample size would provide more weights.

In the previous examples, factor A has two levels and factor B has 3 levels. To test whether the two mean effects of factor A are equal or not , we may use weights proportional to treatment sample sizes and conduct the following hypothesis testing: (the weights are proportional to n_{ij} and the denominators are just the sample size at factor levels to normalize the weights to add to one and make this a weighted average)

$$H_0: \left(\frac{n_{11}}{n_{1\cdot}}\right)\mu_{11} + \left(\frac{n_{12}}{n_{1\cdot}}\right)\mu_{12} + \left(\frac{n_{13}}{n_{1\cdot}}\right)\mu_{13} = \left(\frac{n_{21}}{n_{2\cdot}}\right)\mu_{21} + \left(\frac{n_{22}}{n_{2\cdot}}\right)\mu_{22} + \left(\frac{n_{23}}{n_{2\cdot}}\right)\mu_{23}$$

H_a: equality does not hold

For factor B, we may also use weights proportional to treatment sample sizes and conduct the following hypothesis testing to compare the weighted averages across the 3 levels:

$$H_0: \left(\frac{n_{11}}{n_{\cdot 1}}\right)\mu_{11} + \left(\frac{n_{21}}{n_{\cdot 1}}\right)\mu_{21} = \left(\frac{n_{12}}{n_{\cdot 2}}\right)\mu_{12} + \left(\frac{n_{22}}{n_{\cdot 2}}\right)\mu_{22} = \left(\frac{n_{13}}{n_{\cdot 3}}\right)\mu_{13} + \left(\frac{n_{23}}{n_{\cdot 3}}\right)\mu_{23}$$
$$H_a\text{: not all equalities hold}$$

In general,

When sample sizes do constitute appropriate weights, the alternatives for testing for weighted factor A effects can be stated in general as follows:

$$H_0: \sum_j \left(\frac{n_{1j}}{n_{1\cdot}}\right)\mu_{1j} = \cdots = \sum_j \left(\frac{n_{aj}}{n_{a\cdot}}\right)\mu_{aj} \qquad \textbf{(23.50)}$$
$$H_a\text{: not all equalities hold}$$

and the alternatives for testing for weighted factor B effects are:

$$H_0: \sum_i \left(\frac{n_{i1}}{n_{\cdot 1}}\right)\mu_{i1} = \cdots = \sum_i \left(\frac{n_{ib}}{n_{\cdot b}}\right)\mu_{ib} \qquad \textbf{(23.51)}$$
$$H_a\text{: not all equalities hold}$$

To conduct the hypothese testing: Similarly, we can compare a full and reduced model to test the hypothesis, or estimate/test one or multiple contrasts with multiple comparison adjustments.

Or, for these weighting options ($\sim n_{ij}$), we can get some easy formulas to set the test-statistics. It can be shown to test for factor A, we can construct a F-test based on

$$F = [SSA/(a-1)] \div [SSE/(n_T - ab)]$$

which follows a $F(a-1, n_T - ab)$ distribution. Here SSA can be calculated as the ordinary **single-factor** treatment sum of squares with the factor A levels considered to be the treatments. Note that the test for factor B weighted effects using weights ($\sim n_{ij}$) can be carried out in similar fashion .

Example 3: In the growth hormone example of Table 23.1,

suppose that the treatment sample sizes n_{ij} reflect the relative importance of the factor means. We now wish to test whether gender affects the weighted mean change in the growth rate. Then we are testing

$$H_0: \frac{3}{7}\mu_{11} + \frac{2}{7}\mu_{12} + \frac{2}{7}\mu_{13} = \frac{1}{7}\mu_{21} + \frac{3}{7}\mu_{22} + \frac{3}{7}\mu_{23}$$
$$H_a\text{: equality does not hold}$$

Approach 1: using SSA from one-way ANOVA because of this special weighting

```r
# We calculate the SSA from one-factor ANOVA
summary(aov(response ~ Gender, data=Ex23 ))
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## Gender        1  0.003  0.0029   0.006   0.94
## Residuals    12  5.771  0.4810
```

```r
# check SSE from the full model
anova(LM.full)
```

```
## Analysis of Variance Table
##
## Response: response
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Ind1       1  12.00 12.0000 73.8462 2.600e-05 ***
## Ind2       1   7.22  7.2200 44.4308 0.0001582 ***
## Ind3       1   1.62  1.6200  9.9692 0.0134474 *
## Ind4       1   5.76  5.7600 35.4462 0.0003406 ***
## Ind5       1  13.23 13.2300 81.4154 1.819e-05 ***
## Ind6       1   2.43  2.4300 14.9538 0.0047614 **
## Residuals  8   1.30  0.1625
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We find the SSA= 0.003 with df=1, then from the previous full model, SSE= 1.3000 with df=8. We can get the test statistic and p-value as follows:

```r
(Fstat= 0.003/(1.3/8))
```

```
## [1] 0.01846154
```

```r
#p-value
(Pv= 1- pf(Fstat, 1, 8))
```

```
## [1] 0.8952783
```

Approach 2: We estimate the test the contrasts for the two gender differences

We cam get the testing results if $L = 0$ or the confidence interval to compare with 0.

```r
#L1 was the same as last example#
L = list( L= c( 3/7, 2/7,  2/7,  -1/7,  -3/7, -3/7))

contrast(Est.mean, L, adjust='none')
```

```
##  contrast estimate    SE df t.ratio p.value
##  L          0.0286 0.215  8 0.133   0.8978
```

```r
confint(contrast(Est.mean, L, adjust='none'))
```

```
##  contrast estimate    SE df lower.CL upper.CL
##  L          0.0286 0.215  8   -0.468    0.525
##
## Confidence level used: 0.95
```

Approach 3: Regression approach: following the approach with more general weighting options in the previous example.

Summary this week

- Reading: Chapter 23.1-23.5
- Homework Assignment for ch 23: (due 11/14)
  - 23.6,23.20 : cash offer data (CH19PR10.txt) previously studied in chapter 19
  - 23.18: Auditor training data (CH21PR05.txt) previously studied in chapter 21
  - Here each problem assumed some data are missing to make the study 'unbalanced'.
- Project due on 12/5