STA 305s14 Regular Assignment Three[1]

This assignment is preparation for Term Test Two on March 10th, and for the final exam. Your solutions to these homework problems will not be handed in. Use the formula sheet, which is posted on the course home page. As more material is covered, additional problems will be added at the end of the assignment.

Lecture Unit 6: Analysis of variance methods for a one-factor completely randomized design

1. In a study of remedies for lower back pain, volunteer patients at a back clinic were randomly assigned to one of seven treatment conditions:

   - OxyContin: A pain pill in the opiate family.
   - Ibuprofen: A non-steroidal anti-inflammatory drug (Advil, Motrin)
   - Acupuncture: The insertion and manipulation of thin needles into specific points on the body to relieve pain or for therapeutic purposes.
   - Chiropractic: A form of therapy that includes manipulation of the spine, other joints and soft tissue.
   - Stress reduction training based on thinking positive thoughts, a treatment that theoretically should not be effective. This is the non-drug control condition.
   - Placebo: A sugar pill; patients were told that it was a pain killer with few side effects. This is the drug control condition.
   - Waiting list control: Patients were told that the clinic was overcrowded (true), and that they would were on a waiting list. This group received no treatment at all, not even a pretend treatment — until the study was over, at which point they received the most effective treatment based on the results of the study.

   Degree of reported pain was measured by a questionnaire before treatment began, and again after six weeks. The dependent variable was Before-minus-After difference in reported pain, which will be called "improvement," or "effectiveness."

   The idea is that the effectiveness of the drug treatments should be assessed relative to the drug control (placebo), while the effectiveness of the non-drug treatments should be assessed relative to the non-drug control (stress reduction training). Improvement in the control conditions can be measured relative to no treatment at all.

   (a) You will use a regression model with an intercept and indicator (zero-one) dummy variables. Make a table showing how you would set up the dummy variables. There is more than one reasonable way to do this.

   (b) Add another column to the end of your table, showing the expected improvement in back pain in terms of your β parameters.

   (c) For each of the questions below, give the null hypothesis in terms of β parameters. This is a scientific study, and the results will not be ignored if they are the opposite of what's predicted. So even when the question seems to imply a directional alternative, all the tests are non-directional, and the null hypothesis says that something is *equal* to something else.

      i. Does OxyContin work any better than the placebo?

      ii. Does Ibuprofen work any better than the placebo?

[1] Copyright information is at the end of the last page.

iii. Do Chiropractic treatment and Stress reduction training differ in their effectiveness?
iv. Which results in more mean improvement, Acupuncture or Stress reduction training?
v. Is the average improvement from the two drug therapies different from the improvement from the placebo?
vi. Does either drug therapy differ from the placebo in its effectiveness? (This is a single test of two equalities.)
vii. Does either non-drug therapy differ in effectiveness from Stress reduction training?
viii. Is the Placebo better than no treatment at all?
ix. Is Stress reduction training better than no treatment at all?
x. s the average effectiveness of the drug therapies different from the average effectiveness of the non-drug therapies?
xi. Do Stress reduction training and the Placebo differ in their effectiveness?
xii. Does either control condition (Drug or Non-Drug) differ from no treatment at all?
xiii. Is treatment condition (the full independent variable, including the No Treatment condition) related to improvement?

2. For the random sampling model (not the randomization model) explain how the assumption of unit-treatment additivity implies equal variances. Use the example of an experiment with a control condition and just one experimental treatment.

3. Let $Y_1, \ldots, Y_n$ be a random sample (i.i.d.) from a $N(\mu, \sigma^2)$ distribution.

   (a) Write this as a regression model in matrix form.
   i. What is $\mathbf{Y}$? What are its dimensions?
   ii. What is $\mathbf{X}$? What are its dimensions?
   iii. What is $\boldsymbol{\beta}$? What are its dimensions?
   iv. What is $\boldsymbol{\epsilon}$? What is its distribution?
   v. What is $\mathbf{X}'\mathbf{X}$?
   vi. What is $(\mathbf{X}'\mathbf{X})^{-1}$?
   vii. What is $\mathbf{X}'\mathbf{Y}$?
   viii. What is $\widehat{\boldsymbol{\beta}}$?
   ix. What is the $n \times 1$ vector $\widehat{\mathbf{Y}}$?
   x. What is SSE?

   (b) Cite the fact from the formula sheet that tells you $\sum_{i=1}^{n}(Y_i - \overline{Y})^2$ and $\overline{Y}$ are independent.

   (c) Cite the fact from the formula sheet that tells you $\frac{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}{\sigma^2} \sim \chi^2(n-1)$.

4. Consider once again the experiment on scab disease in potatoes, which first appeared in Computer Assignment 2. Remember that there were three levels of sulphur (300, 600 and 1200 pounds per acre) and a control.

   (a) Make a table showing how you would set up the dummy variables for *cell means* dummy variable coding. That's the one with indicators and no intercept. Add another column showing the expected value of the response for each experimental condition, including the control.

   (b) Write this as a regression model in matrix form.
   i. What is $\mathbf{Y}$? What are its dimensions?
   ii. What is $\mathbf{X}$? What are its dimensions?
   iii. What is $\boldsymbol{\beta}$? What are its dimensions?
   iv. What is $\boldsymbol{\epsilon}$? What is its distribution?

v. What is $\mathbf{X}'\mathbf{X}$? This is easier to see if data from the same experimental condition are in adjacent rows.

vi. What is $(\mathbf{X}'\mathbf{X})^{-1}$?

vii. What is $\mathbf{X}'\mathbf{Y}$?

viii. What is $\widehat{\boldsymbol{\beta}}$?

ix. What is the *distribution* of $\widehat{\boldsymbol{\beta}}$? Give the 4×4 covariance matrix explicitly.

x. What is $\widehat{\mathbf{Y}}$?

(c) Now change notation, letting $Y_{ij} = \mu_j + \epsilon_{ij}$, for $j = 1, \ldots, p$ and $i = i, \ldots, n_j$. For the scab disease example, what is p? Now in general,

i. What is the joint distribution of the ϵ_{ij}?

ii. In this new notation, what is $\overline{Y}_j$?

iii. Let $\overline{Y}$ denote the sample mean of all the observations. White a formula for $\overline{Y}$ in terms of the new notation.

iv. In the new notation, what is SST?

v. In the new notation, what is SSE?

vi. In the new notation, what is SSR?

vii. What is the distribution of $\frac{SSE}{\sigma^2}$?

viii. Under $H_0 : \mu_1 = \cdots = \mu_p$, what is the distribution of $\frac{SST}{\sigma^2}$? Why can you just use question 3?

ix. Write $\overline{Y}$ as an explicit function of the $\overline{Y}_j$.

x. How do you know SSR and SSR are independent for this model?

xi. You will notice that the formula sheet now has this, which was proved in STA302: If $W = W_1 + W_2$ with W_1 and W_2 independent, $W \sim \chi^2(\nu_1 + \nu_2)$, $W_2 \sim \chi^2(\nu_2)$ then $W_1 \sim \chi^2(\nu_1)$. Use this fact to find the distribution of $\frac{SSR}{\sigma^2}$. Why does your conclusion depend on H_0 being true?

xii. Based on the newly revised formula sheet, give the formula for a test statistic for testing $H_0 : \mu_1 = \cdots = \mu_p$. What is its distribution when H_0 is true? Why might you expect big values when H_0 is false? It's also the test statistic you'd get if you carried out a general linear test, but that's not obvious. Don't just give the formula for the general linear test.

5. Suppose a completely randomized design is used to compare the expected response for five experimental treatments.

(a) Write a regression model with an intercept for this problem.

(b) Make a table showing how you would set up dummy variables with effect coding. That's the scheme with the minus ones.

(c) Define the "grand mean" by $\mu = \frac{1}{p}\sum_{j=1}^{p} \mu_j$. For this problem (with $p = 5$), what is μ in terms of the β values? Show your work.

(d) In terms of the β values, what is $\tau_3 = \mu_3 - \mu$? Show a little work.

(e) In terms of the β values, what is $\tau_5 = \mu_5 - \mu$? Show a little work.

(f) Write $H_0 : \mu 2 = \mu_3$ in terms of β values.

(g) Write $H_0 : \mu 4 = \mu_5$ in terms of β values. Simplify a bit.

6. In this question, "test a contrast" is a short way to say test the null hypothesis that a contrast of the μ_j values is equal to zero. The "weights" of a contrast are the a_j constants in $c = a_1\mu_1 + \cdots + a_p\mu_p$. Here's the setting. Three hundred university student volunteers who wanted to lose weight were weighed in a clinic under controlled conditions. Then they were randomly assigned to one of six treatment groups:

   1 Free Health Club membership without personal trainer

   2 Free Health Club membership with personal trainer: Emphasis on aerobic conditioning

   3 Free Health Club membership with personal trainer: Emphasis on strength training

   4 Free vegetarian cooking and diet class

   5 Free Exercise video

   6 Waiting list control (They were told "Sorry, we'll call you when there's an opening.")

   After six months they were weighed again. The dependent variable is weight loss in kilograms: Before minus After.

   (a) Is average weight loss in the exercise video condition more than average weight loss for the waiting list control condition?

      i. State the null hypothesis in terms of μ_j values.

      ii. In the table below, give the weights of the contrast or contrasts you would test to answer the question. There should be one row for each contrast.

1	2	3	4	5	6

   (b) Is average weight loss different for the three treatments that include a health club membership?

      i. State the null hypothesis in terms of μ_j values.

      ii. In the table below, give the weights of the contrast or contrasts you would test to answer the question. There should be one row for each contrast.

1	2	3	4	5	6

   (c) Consider a test for differences among the three treatments that include a health club membership, and *at the same time*, for the three treatments that do not include a health club membership

      i. State the null hypothesis in terms of μ_j values.

      ii. In the table below, give the weights of the contrast or contrasts you would test to answer the question. There should be one row for each contrast.

1	2	3	4	5	6

7. Two contrasts $\mathbf{a}_1'\boldsymbol{\mu}$ and $\mathbf{a}_2'\boldsymbol{\mu}$ are said to be *orthoganal* if $\mathbf{a}_1'\mathbf{a}_2 = \mathbf{0}$. Show that if the contrasts c_1 and c_2 are orthoganal and sample sizes are equal, then the estimated contrasts $\widehat{c}_1$ and $\widehat{c}_2$ have zero covariance. Why does this imply that the estimated contrasts are independent if the data are normally distributed?

8. Suppose you have data from an experiment with four treatments and a control. If you reject the null hypothesis that all five treatment means are equal, you would like to test all pairwise comparisons at joint significance level 0.05. But comparisons of treatments with the control are more important to detect than comparisons of the treatments with each other. So you decide to carry out the pairwise tests and compute one-at-a-time p-values as usual, and then reject H_0 with your follow-up tests this way. If $p < 0.01$ for a comparison of treatment to control, reject. If $p < 0.01/6$ for a comparison of treatment to treatment, reject. Will this procedure protect the family of tests against Type I error at the *joint* 0.05 significance level? Answer Yes or No and show your work. See Bonferroni's inequality on the formula sheet.

9. For the general multiple regression model with fixed independent variables and normal error terms, let F_1 denote the test statistic of an initial F-test whose null hypothesis imposes q constraints on $\boldsymbol{\beta}$, and let F_2 denote the test statistic of a follow-up test whose null hypothesis imposes $s < q$ constraints on $\boldsymbol{\beta}$. We have

$$F_1 = \frac{SSR - SSR_1}{q\,MSE} \qquad F_2 = \frac{SSR - SSR_2}{s\,MSE}.$$

   We reject H_0 with a Scheffé test if $F_2 > \frac{q}{s} f_\alpha(q, n-p)$, where $f_\alpha(q, n-p)$ is the critical value of the initial test. Show that if the null hypothesis of the initial test is not rejected, then the Scheffé test's null hypothesis cannot be rejected either.

10. A food company wanted to test four different package designs for a new breakfast cereal. The experimental units were twenty stores with approximately equal sales volume. Each package was used in five randomly chosen stores, but a fire in one of the stores caused it to be dropped from the study. The response variable was total sales of the new cereal, in cases. Here is my SAS code.

```
/********************** kentonreg.sas *************************/
options linesize=79 pagesize=100 noovp formdlim=' ' nodate;
title 'Kenton Oneway Example From Kutner et al.';

proc format;
     value pakfmt 1 = '3Colour Cartoon'   2 = '3Col No Cartoon'
                  3 = '5Colour Cartoon'   4 = '5Col No Cartoon';

data food;
     infile 'kenton.data';
     input package sales;
     label package = 'Package Design'
           sales   = 'Number of Cases Sold';
     format package pakfmt.;
     if package=1 then p1=1; else p1=0;
     if package=2 then p2=1; else p2=0;
     if package=3 then p3=1; else p3=0;
     if package=4 then p4=1; else p4=0;

proc means n mean stddev;
     class package;
     var sales;
```

```
proc reg;
     title2 'Cell means coding';
     model sales = p1 p2 p3 p4 / noint;
     One_vs_2:   test p1=p2;
     One_vs_3:   test p1=p3;
     One_vs_4:   test p1=p4;
     Two_vs_3:   test p2=p3;
     Two_vs_4:   test p2=p4;
     Three_vs_4: test p3=p4;
     Mystery1: test p1+p2=p3+p4;
     Mystery2: test p1+p3=p2+p4;
     Mystery3: test p1-p3=p2-p4;
     Mystery4: test p1=p2, p3=p4;
     Mystery5: test p1=p2=p3=p4;

proc iml;
     title2 'Critical value of initial test';
     numdf = 3;   /* p-1 = Numerator degrees of freedom for initial test */
     dendf = 15;  /* n-p = Denominator degrees of freedom for initial test */
     alpha = 0.05;
     critval = finv(1-alpha,numdf,dendf);  print critval;
```

The output (somewhat edited) appears below.

```
                  Kenton Oneway Example From Kutner et al.                  1

                             The MEANS Procedure

                Analysis Variable : sales Number of Cases Sold

                              N
     Package Design      Obs      N            Mean         Std Dev
     ------------------------------------------------------------
     3Colour Cartoon       5      5      14.6000000       2.3021729

     3Col No Cartoon       5      5      13.4000000       3.6469165

     5Colour Cartoon       4      4      19.5000000       2.6457513

     5Col No Cartoon       5      5      27.2000000       3.9623226
     ------------------------------------------------------------
```

The REG Procedure
Model: MODEL1
Dependent Variable: sales Number of Cases Sold

Number of Observations Read	19
Number of Observations Used	19

NOTE: No intercept in model. R-Square is redefined.

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	7183.80000	1795.95000	170.29	<.0001
Error	15	158.20000	10.54667		
Uncorrected Total	19	7342.00000			

Root MSE	3.24756	R-Square	0.9785
Dependent Mean	18.63158	Adj R-Sq	0.9727
Coeff Var	17.43042		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value
p1		1	14.60000	1.45235	10.05
p2		1	13.40000	1.45235	9.23
p3		1	19.50000	1.62378	12.01
p4		1	27.20000	1.45235	18.73

Parameter Estimates

Variable	Label	DF	Pr > |t|
p1		1	<.0001
p2		1	<.0001
p3		1	<.0001
p4		1	<.0001

The REG Procedure

Test One_vs_2 Results for Dependent Variable sales

Source	DF	Mean Square	F Value	Pr > F
Numerator	1	3.60000	0.34	0.5677
Denominator	15	10.54667		

Test One_vs_3 Results for Dependent Variable sales

Source	DF	Mean Square	F Value	Pr > F
Numerator	1	53.35556	5.06	0.0399
Denominator	15	10.54667		

Test One_vs_4 Results for Dependent Variable sales

Source	DF	Mean Square	F Value	Pr > F
Numerator	1	396.90000	37.63	<.0001
Denominator	15	10.54667		

Test Two_vs_3 Results for Dependent Variable sales

Source	DF	Mean Square	F Value	Pr > F
Numerator	1	82.68889	7.84	0.0135
Denominator	15	10.54667		

Test Two_vs_4 Results for Dependent Variable sales

Source	DF	Mean Square	F Value	Pr > F
Numerator	1	476.10000	45.14	<.0001
Denominator	15	10.54667		

Test Three_vs_4 Results for Dependent Variable sales

Source	DF	Mean Square	F Value	Pr > F
Numerator	1	131.75556	12.49	0.0030
Denominator	15	10.54667		

Test Mystery1 Results for Dependent Variable sales

Source	DF	Mean Square	F Value	Pr > F
Numerator	1	411.40000	39.01	<.0001
Denominator	15	10.54667		

Test Mystery2 Results for Dependent Variable sales

Source	DF	Mean Square	F Value	Pr > F
Numerator	1	49.70588	4.71	0.0464
Denominator	15	10.54667		

Test Mystery3 Results for Dependent Variable sales

Source	DF	Mean Square	F Value	Pr > F
Numerator	1	93.18824	8.84	0.0095
Denominator	15	10.54667		

Test Mystery4 Results for Dependent Variable sales

Source	DF	Mean Square	F Value	Pr > F
Numerator	2	67.67778	6.42	0.0097
Denominator	15	10.54667		

Test Mystery5 Results for Dependent Variable sales

Source	DF	Mean Square	F Value	Pr > F
Numerator	3	196.07368	18.59	<.0001
Denominator	15	10.54667		

Kenton Oneway Example From Kutner et al. 14
Critical value of initial test

critval

3.2873821

Please answer these questions based on the output.

(a) Write the regression equation.

(b) In terms of your β values, what is the null hypothesis of the natural initial test?

(c) Do you reject the null hypothesis of the initial test at $\alpha = 0.05$? Give the value of the test statistic and the p-value.

(d) You decide to follow up with all pairwise comparisons, Bonferroni-corrected. Make a 4×4 table, and put Bonferroni-corrected p-values in the upper triangle. Does this allow you to make a claim that one of the package designs is most effective? If Yes, which one is it?

(e) Now make a similar table for Scheffé tests. Instead of putting Scheffé-corrected p-values, just write the word Yes or No. Naturally you will need a calculator. Are you still able to conclude that one of the designs was most effective?

(f) Now answer each of the following questions in plain, non-statistical language – but base your answers on Scheffé follow-ups to the initial test. Where possible, draw directional conclusions rather than just answering Yes or No.

i. Is average response to the packages with cartoons different from average response to the packages without cartoons?

ii. Is average response to the 3-colour packages different from average response to the 5-colour packages?

iii. Does the effect of colour depend on whether or not the package has a cartoon?

iv. Is there an effect of Cartoon for *either* 3-colour packages or 5-colour packages? This is one test.
