CSUEB – STAT 6305 – Winter 2017 – Prof. Yan Yan Zhou

Homework 3 - Henry Lankin, Gui Larangeira

February 01, 2017

HW 3: 9.7a,b,c,d,e, 9.12a,b,c, 15.6a,b, 15.10a,b,c,d

9.7

9.7 Refer to Example 8.4. The researchers were interested in determining if the mean oxygen content was lower for samples taken near the mouth of the Mississippi River in comparison to the samples further away from the mouth. Write a contrast to answer each of the following questions and test if the contrast is different from 0 using $\alpha = .05$.

a) **Question: Is the mean oxygen content at 20 km different than the average of the mean oxygen content at 1 km, 5 km, and 10 km?**

We must test the null hypothesis:

$$H_o: l_1 = (3, -1, -1, -1)$$

From first row of Table II (AOV Linear Contrast Analysis), we find the p-value is small enough to reject the H_o and we conclude that the mean oxygen content at 20km is different from the average of the others.

b) **Question: Is the mean oxygen content at 10 km different than the average of the mean oxygen content at 1 km and 5 km?**

Likewise, we test the H_o and conclude by examining the second row of Table II that the oxygen content at 10km is different than the average of those at 1 and 5km.

$$H_o : l_2 = (0, 2, -1, -1)$$

c) **Question: Is the mean oxygen content at 5 km different than the average of the mean oxygen content at 1 km?**

Again, Table II row 3 supports the falsehood of the null and we accept the alternative hypothesis that the O_2 content at 5 and 1km are different at the proposed significance level.

$$H_o: l_3 = (0, 0, 1, -1)$$

d) Are the three contrasts defined above mutually orthogonal?

We have,

$$l_1 \cdot l_2 = (3)(0) + (-1)(2) + (-1)(-1) + (-1)(-1) = 0$$
$$l_2 \cdot l_3 = (0)(0) + (2)(0) + (1)(-1) + (-1)(-1) = 0$$
$$l_1 \cdot l_3 = (3)(0) + (-1)(0) + (-1)(1) + (-1)(-1) = 0$$

Since the dot product of each pair is 0, the three contrast statements are mutually orthogonal.

e) Do the three contrasts sum of squares total to SS_{TRT}?

From the SAS output, we have the following AOV table:

	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	5927.60	1975.87	122.22	<.0001
Error	36	582.00	16.17		
Corrected Total	39	6509.60			

This shows that $SS_{treatment} = 5927.6$.

Also from the SAS output, for the contrast statements in parts (a)-(c), we have the following results,

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
20 vs. 1, 5, 10	1	2218.80	2218.80	137.25	<.0001
10 vs. 1, 5	1	117.60	117.60	7.27	0.0106
5 vs. 1	1	3591.20	3591.20	222.14	<.0001

We can verify that indeed

$$SS_{C1} + SS_{C2} + SS_{C3} = 2218.8 + 117.6 + 3591.2 = 5927.6 = SS_{Treatment}$$

Thus, the treatment sum of squares, which has degrees of freedom 3, is broken up into 3 orthogonal sum of squares each with 1 degree of freedom.

9.12

9.12 Refer to Exercise 7.20. The wildlife biologist was interested in determining if the mean weights of deer raised in a zoo would be lower than those from a more uncontrolled environment, for example, either from the wild or raised on a ranch.

a) Use a multiple comparison procedure to determine if the mean weight of the deer raised in the wild or on a ranch is significantly higher than the mean weight of deer raised in a zoo.

We performed both a Tukey and an LSD test in SAS, and the output is shown below:

Alpha	0.05
Error Degrees of Freedom	21
Error Mean Square	784.02
Critical Value of Studentized Range	3.56
Minimum Significant Difference	35.29

Tukey Pairwise Comparison

Means with the same letter are not significantly different.			
Tukey Grouping	Mean	N	treatment
A	122.86	8	wild
A			
A	118.39	8	ranch
A			
A	102.88	8	zoo

Alpha	0.05
Error Degrees of Freedom	21
Error Mean Square	784.02
Critical Value of t	2.08
Least Significant Difference	29.12

LSD Pairwise Comparison

Means with the same letter are not significantly different.			
t Grouping	Mean	N	treatment
A	122.86	8	wild
A			
A	118.39	8	ranch
A			
A	102.88	8	zoo

We see from both multiple comparisons tests that, at the $\alpha = 0.05$ level, there is not a significant difference between the mean weight of deer raised in the zoo and those raised in the wild or on a ranch.

b) Write a linear contrast to compare the average weight of deer raised in a zoo or on a ranch to the mean weight of deer raised in the wild.

$$H_o: l = (-1, -1, 2)$$

c) Test at the $\alpha = 0.05$ level if your contrast in (b) is significantly different from zero. What conclusions can you make from this test?

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
wild, ranch vs. zoo	1	1680.22	1680.22	2.14	0.1580

At a significance level of .05, given the p-value above, there is not enough evidence to support the alternative hypothesis that the mean weight of deer raised in a zoo and those raised in the wild or on a ranch are significantly different.

15.6

15.6 An industrial psychologist working for a large corporation designs a study to evaluate the effect of background music on the typing efficiency of secretaries. The psychologist selects a random sample of seven secretaries from the secretarial pool. Each subject is exposed to three types of background music: no music, classical music, and hard rock music. The subject is given a standard typing test that combines an assessment of speed with a penalty for typing errors. The particular order of the three experiments is randomized for each of the seven subjects. The results are given here with a high score indicating a superior performance. This is a special type of randomized complete block design in which a single experimental unit serves as a block and receives all treatments.

a) Write a statistical model for this experiment and estimate the parameters in your model.

Using a CRBD design with:

y_{ij} – the score from each test-music type combination: 21 observations

τ_i – the effect due to the type of music: 3 treatment levels

β_j – the effect due to the secretary: 7 block levels

ε_{ij} – random error associated with each test-music type combination: 21 residual errors

μ – overall score mean

CRBD design model: $y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$

The following table from our SAS output shows the parameter estimates for the overall mean score $\hat{\mu}$, the music treatments $\hat{\tau}_i$, and the subject blocks $\hat{\beta}_j$, with NoMusic and subject 7 set to be the base parameters:

Parameter	Estimate		Standard Error	t Value	Pr > |t|
Intercept	17.1904761	B	1.00677974	17.07	<.0001
subject 1	3.66666667	B	1.25567495	2.92	0.0128
subject 2	0.66666667	B	1.25567495	0.53	0.6052
subject 3	7.00000000	B	1.25567495	5.57	0.0001
subject 4	2.33333333	B	1.25567495	1.86	0.0878
subject 5	4.66666667	B	1.25567495	3.72	0.0029
subject 6	7.33333333	B	1.25567495	5.84	<.0001
subject 7	0.00000000	B	.	.	.
music Classical	2.14285714	B	0.82203221	2.61	0.0229
music HardRock	-0.71428571	B	0.82203221	-0.87	0.4019
music NoMusic	0.00000000	B	.	.	.

b) Are there differences in the mean typing efficiency for the three types of music? Use $\alpha = 0.05$.

Source	DF	Type I SS	Mean Square	F Value	Pr > F
subject	6	149.3333333	24.8888889	10.52	0.0003
music	2	30.9523810	15.4761905	6.54	0.0120

The table above shows a p-value of $0.0120 < 0.05$, implying that the null hypothesis should be rejected. We conclude that there is a significant difference between the mean scores of the 3 music types.

15.10

15.10 A petroleum company was interested in comparing the miles per gallon achieved by four different gasoline blends (A, B, C, and D). Because there can be considerable variability due to differences in driving characteristics and car models, these two extraneous sources of variability were included as "blocking" variables in the study. The researcher selected four different brands of cars and four different drivers. The drivers and brands of cars were assigned to blends in the manner displayed in the following table. The mileage (in mpg) obtained over each test run was recorded as follows.

a) **Write a model for this experimental setting.**

Using a **Latin Square Design** with:

y_{ij} – the mpg from each driver-model-blend combination: 16 observations

τ_k – the effect due to the type of blend, where k is determined by i, j: 4 treatment levels applied through 16 block combinations

β_i – the effect due to the driver: 4 block levels

γ_j – the effect due to the model: 4 block levels

ε_{ij} – random error associated with each driver-model-blend combination: 16 residual errors

μ – overall mpg mean

Latin squares design model: $\boldsymbol{y_{ij} = \mu + +\beta_i + \gamma_j \ + \tau_k + \varepsilon_{ij}}$

b) **Estimate the parameters in the model.**

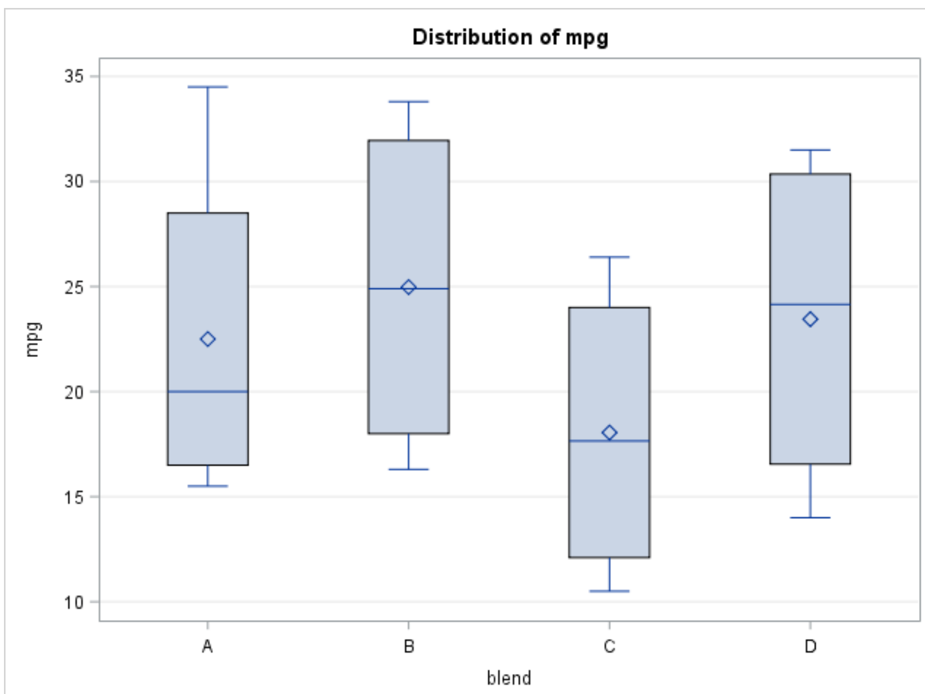
The following table shows the parameter estimates for the overall mpg mean $\hat{\mu}$, the blend treatments $\hat{\tau}_{ij}$, the driver blocks $\hat{\beta}_i$, and the model blocks $\hat{\gamma}_j$.

Parameter	Estimate
Intercept	27.2625
driver 1	0.6000
driver 2	-1.37500
driver 3	-0.05000
driver 4	0.000000
model 1	-11.77500

Parameter	Estimate
model 2	5.7000000
model 3	-8.350000
model 4	0.0000000
blend A	-0.950000
blend B	1.5250000
blend C	-5.40000000
blend D	0.0000000

c) Conduct an analysis of variance. Use $\alpha = 0.05$.

We start by including the boxplots of the data for sake of a quick visual examination:



The following tables shows the results of the ANOVA test on the Latin squares design.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	9	869.975625	96.6639583	22.42	0.0006
Error	6	25.8637500	4.3106250		
Corrected Total	15	895.839375			

R-Square	Coeff Var	Root MSE	mpg Mean
0.971129	9.333878	2.076204	22.24375

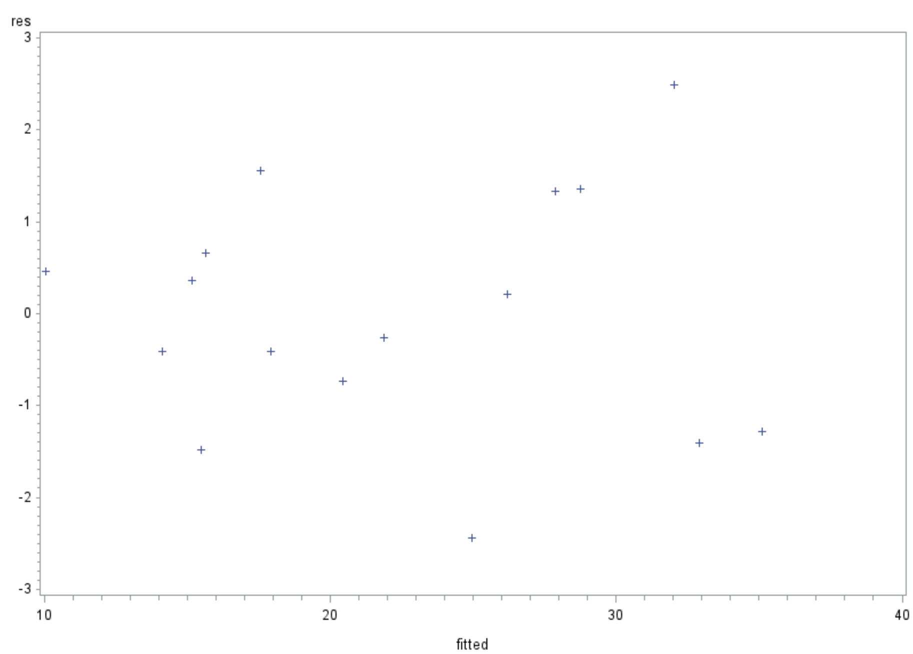
Source	DF	Type I SS	Mean Square	F Value	Pr > F
driver	3	8.3318750	2.7772917	0.64	0.6143
model	3	755.371875	251.7906250	58.41	<.0001
blend	3	106.271875	35.4239583	8.22	0.0151

From the AOV tables above, the p-value of 0.0006 shows that the model does explain the variability in the data with significance. Further, we see from the p-values of $0.6143, 0.0001,$ and 0.0151 for the driver block, model block and blend treatment, respectively, that the driver block does not have a significant effect on mean mpg while the model block and blend treatment does have a significant effect on the mean mpg.

At a significance level of .05, we are only able to determine a worst performing blend, C, while all the others, A, B and D are not significantly different.

Next, we check the model assumptions.

- For equal variance:

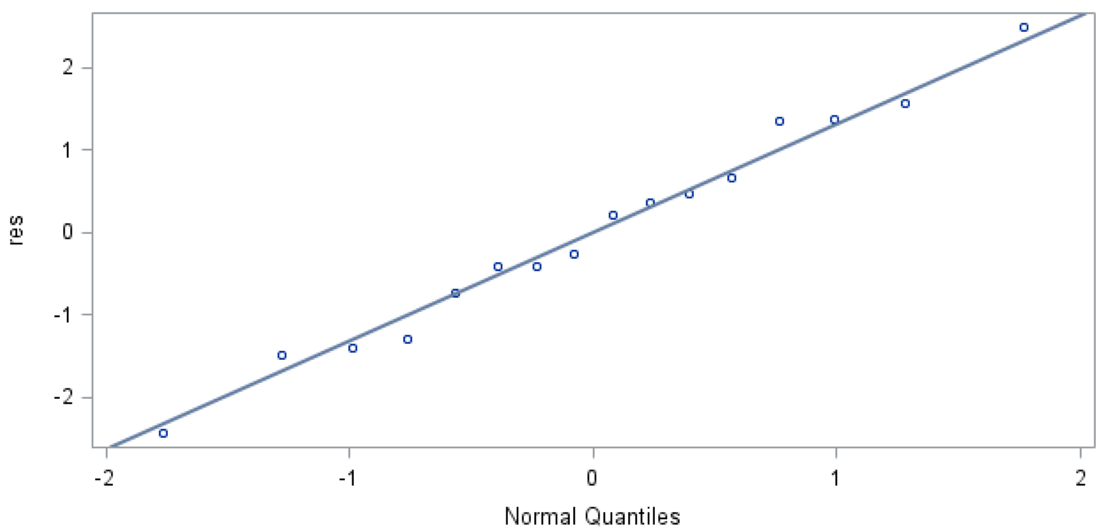


The above graph shows residual values vs. fitted values. There is no distinct pattern signifying equal variances. To examine further, we run a Levene variance test on a one-way ANOVA with blend as the treatment we get the results:

Levene's Test for Homogeneity of mpg Variance ANOVA of Squared Deviations from Group Means					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
blend	3	532.3	177.4	0.11	0.9538
Error	12	19712.0	1642.7		

The table above shows a p-value of $0.9538 > 0.05$, implying that we fail to reject the null hypothesis of equal variances. Thus, there is significant evidence to conclude that the variances are equal.

- Normality of residuals:



Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	W	0.985159	Pr < W	0.9915

The residuals QQ-plot shows little deviation of the residuals from the regression line. Further, the Shapiro-Wilk test results in a p-value of 0.9915, further corroborating our initial impression of no significant departure from normality.

- Independence:

Both the driver and models can be reasonably assumed to be selected independently of each other.

d) What conclusions can you draw concerning the best gasoline blend?

The AOV results show a p-value of $0.0151 < 0.05$, implying that there is significant evidence that blend has an effect on mpg. From the parameters in part (b), blend B has the highest positive estimation of the mpg difference from blend D. This suggests that blend B may be the best gasoline blend.
Running Fisher's LSD and Tukey's HSD multiple comparisons test for further investigation, we have

Fisher LSD test	
Alpha	0.05
Error Degrees of Freedom	6
Error Mean Square	4.310625
Critical Value of t	2.44691
Least Significant Difference	3.5923

Means with the same letter are not significantly different.			
t Grouping	Mean	N	blend
A	24.975	4	B
A			
A	23.450	4	D
A			
A	22.500	4	A
B	18.050	4	C

Tukey HSD test	
Alpha	0.05
Error Degrees of Freedom	6
Error Mean Square	4.310625
Critical Value of Studentized Range	4.89559
Minimum Significant Difference	5.0821

Means with the same letter are not significantly different.				
Tukey Grouping		Mean	N	blend
	A	24.975	4	B
	A			
	A	23.450	4	D
	A			
B	A	22.500	4	A
B				
B		18.050	4	C

Both tests show a significant difference in mean mpg between blend C and blends B and D, while not showing a significant difference between blends B, D and A. While blend B does have the highest mean mpg, it is not possible to conclude with significance that blend B is the best gasoline blend.

SAS code:

```
* ex 19.7;
* input data;
data oxycontent;
        input sample dist1 dist5 dist10 dist20;
        cards;
1 1 4 20 37
2 5 8 26 30
3 2 2 24 26
4 1 3 11 24
5 2 8 28 41
6 2 5 20 25
7 4 6 19 36
8 3 4 29 31
9 0 3 21 31
10 2 3 24 33
;
run;

* print oxycontent data;
proc print data=oxycontent;
run;

* turn oxycontent data into a stacked table;
data oxycontent_flat; set oxycontent;
length treatment $10;
treatment = "dist1"; content = dist1; output;
treatment = "dist5"; content = dist5; output;
treatment = "dist10"; content = dist10; output;
treatment = "dist20"; content = dist20; output;
keep content treatment;
run;

proc print data=oxycontent_flat;
run;
```

```
* run glm procedure on stacked table;
* run contrast statement test for the 3 mutually orthogonal contrast statements;
proc glm data=oxycontent_flat;
        class treatment;
        model content = treatment;
        means treatment /deponly;
        contrast '20 vs. 1, 5, 10' treatment 3 -1 -1 -1;
        contrast '10 vs. 1, 5' treatment 0 2 -1 -1;
        contrast '5 vs. 1' treatment 0 0 1 -1;
run;
quit;

* ex9.12;
* input data unstacked;
data deersize;
        input wild ranch zoo;
        cards;
114.7 120.4 103.1
128.9 91.0 90.7
111.5 119.6 129.5
116.4 119.4 75.8
134.5 150.0 182.5
126.7 169.7 76.8
120.6 100.9 87.3
129.59 76.1 77.3
;
run;

proc print data=deersize;
run;

* convert deersize to stacked table;
data deersizeflat; set deersize;
        treatment = "wild"; size = wild; output;
        treatment = "ranch"; size = ranch; output;
        treatment = "zoo"; size = zoo; output;
        keep size treatment;
run;

proc print data=deersizeflat;
run;

* run anova test for H0: treatment means are equal;
* run multiple comparisons test -- LSD and Tukey;
* test contrast statement (-1,-1,2);
proc glm data=deersizeflat;
        class treatment;
        model size = treatment;
        means treatment / lsd;
        means treatment / tukey;
        means treatment / deponly;
        contrast 'wild,ranch vs. zoo' treatment -1 -1 2;
run;
quit;

* ex 15.6
* input data;
data typing;
        input subject music :$10. score;
        cards;
1       NoMusic         20
2       NoMusic         17
3       NoMusic         24
4       NoMusic         20
5       NoMusic         22
6       NoMusic         25
7       NoMusic         18
1       HardRock        20
2       HardRock        18
3       HardRock        23
4       HardRock        18
5       HardRock        21
6       HardRock        22
7       HardRock        19
1       Classical       24
```

```
2       Classical       20
3       Classical       27
4       Classical       22
5       Classical       24
6       Classical       28
7       Classical       16
;
run;

proc print data=typing;
run;

* calculate means by music (treatment);
proc means data = typing;
        class music;
        var score;
        output out = music_means mean = ;
run;

* calculate means by subject (block);
proc means data = typing;
        class subject;
        var score;
        output out = subject_means mean = ;
run;

* run anova test;
* estimate parameters of the model;
proc glm data = typing;
        class subject music;
        model score = subject music / solution;
        output out=residuals r=res;
run;
quit;

proc print data=residuals;
run;

* ex 15.10;
* input data;
data mpgblend;
        input driver$ model$ blend$     mpg;
        cards;
1       1       A       15.5
2       1       B       16.3
3       1       C       10.5
4       1       D       14
1       2       B       33.8
2       2       C       26.4
3       2       D       31.5
4       2       A       34.5
1       3       C       13.7
2       3       D       19.1
3       3       A       17.5
4       3       B       19.7
1       4       D       29.2
2       4       A       22.5
3       4       B       30.1
4       4       C       21.6
;
run;

proc print data = mpgblend;
run;

* run glm procedure for anova test;
* estimate parameters of the model;
* output residuals res and fitted values p;
proc glm data = mpgblend;
        class driver model blend;
        model mpg = driver model blend / solution;
        means blend / lsd;
        means blend / tukey;
        output out=residuals r=res p=fitted;
run;
quit;
```

```
* plot residuals vs. fitted values;
proc gplot data=residuals;
plot res*fitted;
run;
quit;

* anova test is "faked" into one-way to run levene variance test;
/**
proc glm data = mpgblend;
        class blend;
        model mpg = blend;
        means blend / hovtest=levene;
run;
quit;
*/
```