

STAT362 Homework Assignment 11

Sharon O'Boyle

Problem 1, Problem 7.2, p.231

SAS Program

```
* Sharon O'Boyle;
* Stat 362;
* Homework Assignment 11;

* Problem 1, 7.2 p, 231 ;
* Topic: ANOVA;

* Create dataset;

DATA study;
  DO GROUP = 'STATIN A','STATIN B', 'PLACEBO ';
    DO SUBJ = 1 TO 10;
      INPUT CHOLESTEROL @;
      OUTPUT;
    END;
  END;
DATALINES;
220 190 180 185 210 170 178 200 177 189
160 168 178 200 172 155 159 167 185 199
240 220 246 244 198 238 277 255 190 188
;
Run;

proc print data=study;
title 'study dataset';
run;

*** Perform ANOVA ;

PROC ANOVA DATA=study;
  CLASS GROUP ;
  MODEL CHOLESTEROL = GROUP;
  MEANS GROUP / SNK;
  TITLE "Analysis of Cholesterol Treatment Data";
RUN;
```

SAS Log

```
14 DATA study;
15     DO GROUP = 'STATIN A','STATIN B', 'PLACEBO ';
16         DO SUBJ = 1 TO 10;
17             INPUT CHOLESTEROL @;
18             OUTPUT;
19         END;
20     END;
21 DATALINES;
```

NOTE: SAS went to a new line when INPUT statement reached past the end of a line.

NOTE: The data set WORK.STUDY has 30 observations and 3 variables.

NOTE: DATA statement used (Total process time):

real time	0.03 seconds
cpu time	0.03 seconds

```
25 ;
26 Run;
```

```
27 proc print data=study;
NOTE: Writing HTML Body file: sashtml.htm
28 title 'study dataset';
29 run;
```

NOTE: There were 30 observations read from the data set WORK.STUDY.

NOTE: PROCEDURE PRINT used (Total process time):

real time	15.83 seconds
cpu time	0.26 seconds

```
30
31 *** Perform ANOVA ;
32
33 PROC ANOVA DATA=study;
34     CLASS GROUP ;
35     MODEL CHOLESTEROL = GROUP;
36     MEANS GROUP / SNK;
37     TITLE "Analysis of Cholesterol Treatment Data";
38 RUN;
```

SAS Output

Analysis of Cholesterol Treatment Data

The ANOVA Procedure

Dependent Variable: CHOLESTEROL

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	16258.46667	8129.23333	17.58	<.0001
Error	27	12483.40000	462.34815		
Corrected Total	29	28741.86667			

The F Statistic is 17.58 and the p-value <0.0001. So we would reject the null hypothesis and we would say that there is a difference in the treatments.

Student-Newman-Keuls Test for CHOLESTEROL

Means with the same letter are not significantly different.			
SNK Grouping	Mean	N	GROUP
A	229.600	10	PLACEBO
B	189.900	10	STATIN A
B			
B	174.300	10	STATIN B

The Student-Newman-Keuls tells us that there is a significant difference between the Placebo and the other 2 Statin treatments. It also tells us that there is no significant difference between the Statin A and Statin B treatments.

Problem 2, Problem 7.4

SAS Program

```
* Problem 2, 7.4 p, 232 ;
```

```
* Topic: ANOVA;
```

```
* Create dataset;
```

```
DATA EXAM;
```

```
INPUT PROGRAM $ SCORE;
```

```
DATALINES;
```

```
A 560
```

```
A 520
```

```
A 530
```

```
A 525
```

```
A 575
```

```
A 527
```

```
A 580
```

```
A 620
```

```
B 565
```

```
B 522
```

```
B 520
```

```
B 530
```

```
B 510
```

```
B 522
```

```
B 600
```

```
B 590
```

```
C 512
```

```
C 518
```

```
C 555
```

```
C 502
```

```
C 510
```

```
C 520
```

```
C 516
```

```
D 505
```

```
D 508
```

```
D 512
```

```
D 520
```

```
D 543
```

```
D 523
```

```
D 517
```

```
;
```

```
run;
```

```
proc print data=EXAM;
```

```
title 'EXAM dataset';
```

```
run;
```

```
*** USE PROC GLM BECAUSE IT IS UNBALANCED AND TO GET CONTRASTS;
```

```
PROC GLM DATA=EXAM;
```

```
  TITLE "Analysis of College Entrance Exam - Planned Comparisons";
```

```
  CLASS PROGRAM;
```

```
  MODEL SCORE = PROGRAM;
```

```
  MEANS PROGRAM / SNK ALPHA=0.05;
```

```
  CONTRAST 'A AND B VS. C AND D' PROGRAM  -1 -1 1 1;
```

```
  CONTRAST 'A B AND C VS. D' PROGRAM  -1 -1 -1 3;
```

```
RUN;
```

SAS Log

```
40 DATA EXAM;
41 INPUT PROGRAM $ SCORE;
42 DATALINES;
```

NOTE: The data set WORK.EXAM has 30 observations and 2 variables.

NOTE: DATA statement used (Total process time):

real time	0.00 seconds
cpu time	0.00 seconds

```
73 ;
74 run;
```

```
75
76 proc print data=EXAM;
77 title 'EXAM dataset';
78 run;
```

NOTE: There were 30 observations read from the data set WORK.EXAM.

NOTE: PROCEDURE PRINT used (Total process time):

real time	0.01 seconds
cpu time	0.01 seconds

```
79
80 *** USE PROC GLM BECAUSE IT IS UNBALANCED AND TO GET CONTRASTS;
81 PROC GLM DATA=EXAM;
82     TITLE "Analysis of College Entrance Exam - Planned
Comparisons";
83     CLASS PROGRAM;
84     MODEL SCORE = PROGRAM;
85     MEANS PROGRAM / SNK ALPHA=0.05;
86     CONTRAST 'A AND B VS. C AND D' PROGRAM    -1 -1 1 1;
87     CONTRAST 'A B AND C VS. D' PROGRAM    -1 -1 -1 3;
88 RUN;
```

SAS Output

Analysis of College Entrance Exam - Planned Comparisons

The GLM Procedure

Dependent Variable: SCORE

Source	DF	Type I SS	Mean Square	F Value	Pr > F
PROGRAM	3	7607.188095	2535.729365	3.28	0.0366

The F statistic is 3.28 and p-value = 0.0366. So we would reject the null hypothesis at the .05 level and we would say that there is a difference in the programs.

Means with the same letter are not significantly different.			
SNK Grouping	Mean	N	PROGRAM
A	554.63	8	A
A			
A	544.88	8	B
A			
A	519.00	7	C
A			
A	518.29	7	D

The Student-Newman-Keuls tells us that there is no significant difference between any of the Programs. This seems to contradict the results of the F test.

The GLM Procedure

Dependent Variable: SCORE

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
A AND B VS. C AND D	1	7225.152381	7225.152381	9.36	0.0051
A B AND C VS. D	1	2413.012158	2413.012158	3.13	0.0888

The Contrasts show that A and B are significantly different from C and D. However, A, B and C are not significantly different from D.

Problem 3, Problem 7.6

SAS Program

```
* Problem 3: 7.6 p, 233;
* Topic: 2-way ANOVA;

* Create dataset;

DATA dep_study;
    LENGTH GROUP $9 CONDITION $20 ;
    INPUT GROUP $ DRUG $ DEPRESSION;
    CONDITION = TRIM(GROUP) || '-' || DRUG;
DATALINES;
DEFICIENT DRUG 9
DEFICIENT PLACEBO 9
DEFICIENT DRUG 11
DEFICIENT PLACEBO 6
DEFICIENT DRUG 10
DEFICIENT PLACEBO 6
DEFICIENT DRUG 10
DEFICIENT PLACEBO 7
NORMAL DRUG 5
NORMAL PLACEBO 12
NORMAL DRUG 4
NORMAL PLACEBO 11
NORMAL DRUG 7
NORMAL PLACEBO 10
NORMAL DRUG 7
NORMAL PLACEBO 11
;
RUN;

proc print data=dep_study;
title 'depression dataset';
run;

*** Perform ANOVA ;

PROC ANOVA DATA=dep_study;
    TITLE "Analysis of depression Data";
    CLASS GROUP DRUG;
    MODEL DEPRESSION = GROUP | DRUG;
    MEANS GROUP | DRUG / SNK;
RUN;

*** Do follow-up tests ***;

*** 1-way ANOVA using condition ;

PROC ANOVA DATA=dep_study;
    TITLE "One-way ANOVA of depression Data";
    CLASS CONDITION;
    MODEL DEPRESSION = CONDITION;
    MEANS CONDITION / SNK;
RUN;
```

SAS Log

```
145
146 DATA dep_study;
147     LENGTH GROUP $9 CONDITION $20 ;
148     INPUT GROUP $ DRUG $ DEPRESSION;
149     CONDITION = TRIM(GROUP) || '-' || DRUG;
150 DATALINES;
```

NOTE: The data set WORK.DEP_STUDY has 16 observations and 4 variables.

NOTE: DATA statement used (Total process time):

real time	0.00 seconds
cpu time	0.00 seconds

```
167 ;
168 RUN;
```

```
169
170 proc print data=dep_study;
171 title 'depression dataset';
172 run;
```

NOTE: There were 16 observations read from the data set WORK.DEP_STUDY.

NOTE: PROCEDURE PRINT used (Total process time):

real time	0.01 seconds
cpu time	0.01 seconds

```
173
174 *** Perform ANOVA ;
175
176 PROC ANOVA DATA=dep_study;
177     TITLE "Analysis of depression Data";
178     CLASS GROUP DRUG; /* GROUP | AGE is the same as GROUP AGE GROUP*AGE;
*/
179     MODEL DEPRESSION = GROUP | DRUG;
180     MEANS GROUP | DRUG / SNK;
181 RUN;
```

```
182
183 *** 1-way ANOVA using condition ;
184
```

NOTE: PROCEDURE ANOVA used (Total process time):

real time	2:33.60
cpu time	6.98 seconds

```
185 PROC ANOVA DATA=dep_study;
186     TITLE "One-way ANOVA of depression Data";
187     CLASS CONDITION;
188     MODEL DEPRESSION = CONDITION;
189     MEANS CONDITION / SNK;
190 RUN;
```


SAS Output

The ANOVA Procedure

Dependent Variable: DEPRESSION

Source	DF	Anova SS	Mean Square	F Value	Pr > F
GROUP	1	0.06250000	0.06250000	0.04	0.8360
DRUG	1	5.06250000	5.06250000	3.63	0.0811
GROUP*DRUG	1	68.06250000	68.06250000	48.76	<.0001

We need to look at the interaction first. Since there is a significant interaction (p-value <.001) we should not look further at this test result. We need to do another type of test. I chose the 1-way ANOVA using CONDITION. See results below:

The ANOVA Procedure Student-Newman-Keuls Test for DEPRESSION

Means with the same letter are not significantly different.			
SNK Grouping	Mean	N	CONDITION
A	11.0000	4	NORMAL-PLACEBO
A			
A	10.0000	4	DEFICIENT-DRUG
B	7.0000	4	DEFICIENT-PLACEBO
B			
B	5.7500	4	NORMAL-DRUG

This shows that there **is not** a significant difference between the NORMAL-PLACEBO and DEFICIENT-DRUG GROUPS. This also shows that there **is not** a significant difference between the DEFICIENT-PLACEBO and NORMAL-DRUG GROUPS. However, there **is** a significant difference between the NORMAL-PLACEBO GROUP and both DEFICIENT-PLACEBO and NORMAL-DRUG GROUPS. And there **is** a significant difference between the DEFICIENT-DRUG GROUP and both DEFICIENT-PLACEBO and NORMAL-DRUG GROUPS.