One-Way & Two-Way ANOVA Analysis

Table of Contents

Problem 8.2)

Program Code:

```
28 /*************************************************************************************************************************************************
29 Filename: C:\Users\Eric\Desktop\STAT 482\hw8_2.sas
30 Written by: Eric Brown
31 Date: November 24, 2019
32
33 This program uses a program shown and displayed in the problem to compare good
34 cholesterol (HDL) among the three drugs that each of the patients take during the study.
35 The program also transpose the tall statin data set into a fat data set called
36 fatstatin that contains one observation for each subject, and four variables. From this
37 dataset the program compares the hdl values among the three drugs using the
38 ANOVA procedure's REPEATED statement.
39
40 Input: statin.sas7bdat - displayed in the problem
41 Output:One factor Comparsion Analysis, fatstatin.sas7bdat, & Three possible two-way comparison
42 from statin.sas7bdat produced by PROC ANOVA
43 *************************************************************************************************************************************************/
44 Options ls=89 ps=94 nodate nonumber;
45 PROC ANOVA data=statin; * comparing hdl among the three drugs;
46   TITLE 'ONE-WAY REPEATED MEASURES ANOVA';
47   CLASS SUBJ DRUG; *subject and drug are each main effects & no interaction term between them;
48   MODEL HDL= SUBJ DRUG;
49   MEANS DRUG / SNK; *running Student-Newman_Keuls test;
50 RUN;
51 TITLE;
52
53 DATA fatstatin (drop=LDL HDL TOTAL GENDER DIET DRUG); *not including variables other then ones described in prob.;
54   SET statin; *reading in data from the slim tall dataset;
55   BY SUBJ; *sorting by subject to run conditional statements;
56     if DRUG='A' then hdl1= HDL;
57     else if DRUG='B' then hdl2= HDL;
58     else if DRUG='C' then hdl3= HDL;
59   if last.SUBJ then output; *only outputting until each subject's data is read in;
60   retain hdl1 hdl2 hdl3;
61 RUN;
62
63 PROC PRINT data= fatstatin NOOBS;
64   TITLE 'The fat statin tranposed data set';
67
68 *comparing the hdl values among the three drugs using the ANOVA procedure's REPEATED statement;
69 PROC ANOVA data=fatstatin;
70   Title 'One-way ANOVA Using the Repeated Statement';
71   Model hdl1-hdl3= / NOUNI; *not conducting separate analysis for each of the three HDL varaibles;
72   *no class statement means nothing to put on the right side of the equals sign;
73   Repeated DRUG3 contrast (1) / NOM SUMMARY;
74   Repeated DRUG3 contrast (2) / NOM SUMMARY;
75   Repeated DRUG3 contrast (3) / NOM SUMMARY;
76   *three levels of the DRUG and calling the repeated factor DRUG as well;
77 RUN;
78 TITLE;
79
```

Problem 8.2) SAS Output Window:

```
27
28   /*************************************************************************************
28 ! ************************************
29   Filename: C:\Users\Eric\Desktop\STAT 482\hw8_2.sas
30   Written by: Eric Brown
31   Date: November 24, 2019
32
33   This program uses a program shown and displayed in the problem to compare good
34   cholesterol (HDL) among the three drugs that each of the patients take during the
34 ! study.
35   The program also transpose the tall statin data set into a fat data set called
36   fatstatin that contains one observation for each subject, and four variables. From
36 ! this
37   dataset the program compares the hdl values among the three drugs using the
38   ANOVA procedure's REPEATED statement.
39
40   Input: statin.sas7bdat - displayed in the problem
41   Output:One factor Comparsion Analysis, fatstatin.sas7bdat, & Three possible two-way
41 ! comparison
42   from statin.sas7bdat produced by PROC ANOVA
43   *************************************************************************************
43 ! ***********************************/
44   Options ls=89 ps=94 nodate nonumber;
45   PROC ANOVA data=statin; * comparing hdl among the three drugs;
46     TITLE 'ONE-WAY REPEATED MEASURES ANOVA';
47     CLASS SUBJ DRUG; *subject and drug are each main effects & no interaction term
47 ! between them;
48     MODEL HDL= SUBJ DRUG;
49     MEANS DRUG / SNK; *running Student-Newman_Keuls test;
50   RUN;

51   TITLE;
52

NOTE: PROCEDURE ANOVA used (Total process time):
      real time           0.03 seconds
      cpu time            0.03 seconds


53   DATA fatstatin (drop=LDL HDL TOTAL GENDER DIET DRUG); *not including variables other
53 !  then ones described in prob.;
54     SET statin; *reading in data from the slim tall dataset;
55     BY SUBJ; *sorting by subject to run conditional statements;
56       if DRUG='A' then hdl1= HDL;
57       else if DRUG='B' then hdl2= HDL;
58       else if DRUG='C' then hdl3= HDL;
59     if last.SUBJ then output; *only outputting until each subject's data is read in;
60     retain hdl1 hdl2 hdl3;
61   RUN;

NOTE: There were 60 observations read from the data set WORK.STATIN.
NOTE: The data set WORK.FATSTATIN has 20 observations and 4 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.00 seconds


62
63   PROC PRINT data= fatstatin NOOBS;
64     TITLE 'The fat statin tranposed data set';
65   RUN;

NOTE: There were 20 observations read from the data set WORK.FATSTATIN.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds


66   TITLE;
67
68   *comparing the hdl values among the three drugs using the ANOVA procedure's REPEATED
68 !  statement;
69   PROC ANOVA data=fatstatin;
70     Title 'One-way ANOVA Using the Repeated Statement';
71     Model hdl1-hdl3= / NOUNI; *not conducting separate analysis for each of the three
71 ! HDL varaibles;
72     *no class statement means nothing to put on the right side of the equals sign;
73     Repeated DRUG3 contrast (1) / NOM SUMMARY;
74     Repeated DRUG3 contrast (2) / NOM SUMMARY;
75     Repeated DRUG3 contrast (3) / NOM SUMMARY;
76     *three levels of the DRUG and calling the repeated factor DRUG as well;
77   RUN;

78   TITLE;
```

Problem 8.2) Program Output:

ONE-WAY REPEATED MEASURES ANOVA

The ANOVA Procedure

Class Level Information

Class	Levels	Values
SUBJ	20	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
DRUG	3	A B C

Number of Observations Read	60
Number of Observations Used	60

ONE-WAY REPEATED MEASURES ANOVA

The ANOVA Procedure

Dependent Variable: HDL

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	21	4042.750000	192.511905	1.74	0.0666
Error	38	4193.433333	110.353509		
Corrected Total	59	8236.183333			

R-Square	Coeff Var	Root MSE	HDL Mean
0.490852	22.64807	10.50493	46.38333

Source	DF	Anova SS	Mean Square	F Value	Pr > F
SUBJ	19	2539.516667	133.658772	1.21	0.2991
DRUG	2	1503.233333	751.616667	6.81	0.0030

ONE-WAY REPEATED MEASURES ANOVA

The ANOVA Procedure

Student-Newman-Keuls Test for HDL

NOTE: This test controls the Type I experimentwise error rate under the complete null hypothesis but not under partial null hypotheses.

Alpha	0.05
Error Degrees of Freedom	38
Error Mean Square	110.3535

Number of Means	2	3
Critical Range	6.7247888	8.1016249

Means with the same letter are not significantly different.

SNK Grouping	Mean	N	DRUG
A	53.150	20	B
B	44.800	20	C
B			
B	41.200	20	A

Problem 8.2) Program Output – Continued:

The fat statin tranposed data set

SUBJ	hdl1	hdl2	hdl3
1	48	45	43
2	40	66	38
3	21	25	48
4	22	47	31
5	46	55	64
6	54	57	36
7	39	55	41
8	37	57	51
9	45	64	38
10	30	58	61
11	50	61	42
12	54	58	51
13	49	56	30
14	39	30	40
15	23	64	34
16	51	69	37
17	52	54	57
18	43	42	53
19	43	64	43
20	38	36	58

One-way ANOVA Using the Repeated Statement

The ANOVA Procedure

Number of Observations Read	20
Number of Observations Used	20

One-way ANOVA Using the Repeated Statement

The ANOVA Procedure
Repeated Measures Analysis of Variance

Repeated Measures Level Information

Dependent Variable	hdl1	hdl2	hdl3
Level of DRUG3	1	2	3

One-way ANOVA Using the Repeated Statement

The ANOVA Procedure
Repeated Measures Analysis of Variance
Univariate Tests of Hypotheses for Within Subject Effects

Source	DF	Anova SS	Mean Square	F Value	Pr > F	Adj Pr > F G - G	Adj Pr > F H - F
DRUG3	2	1503.233333	751.616667	6.81	0.0030	0.0047	0.0035
Error(DRUG3)	38	4193.433333	110.353509				

Greenhouse-Geisser Epsilon	0.8687
Huynh-Feldt Epsilon	0.9485

Problem 8.2) Program Output – Continued:

One-way ANOVA Using the Repeated Statement

The ANOVA Procedure
Repeated Measures Analysis of Variance
Analysis of Variance of Contrast Variables

DRUG3_N represents the contrast between the nth level of DRUG3 and the 1st

Contrast Variable: DRUG3_2

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Mean	1	2856.050000	2856.050000	17.60	0.0005
Error	19	3082.950000	162.260526		

Contrast Variable: DRUG3_3

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Mean	1	259.200000	259.200000	1.33	0.2639
Error	19	3714.800000	195.515789		

One-way ANOVA Using the Repeated Statement

The ANOVA Procedure
Repeated Measures Analysis of Variance

Repeated Measures Level Information

Dependent Variable	hdl1	hdl2	hdl3
Level of DRUG3	1	2	3

One-way ANOVA Using the Repeated Statement

The ANOVA Procedure
Repeated Measures Analysis of Variance
Univariate Tests of Hypotheses for Within Subject Effects

Source	DF	Anova SS	Mean Square	F Value	Pr > F	Adj Pr > F G - G	Adj Pr > F H - F
DRUG3	2	1503.233333	751.616667	6.81	0.0030	0.0047	0.0035
Error(DRUG3)	38	4193.433333	110.353509				

Greenhouse-Geisser Epsilon 0.8687
Huynh-Feldt Epsilon 0.9485

One-way ANOVA Using the Repeated Statement

The ANOVA Procedure
Repeated Measures Analysis of Variance
Analysis of Variance of Contrast Variables

DRUG3_N represents the contrast between the nth level of DRUG3 and the 2nd

Contrast Variable: DRUG3_1

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Mean	1	2856.050000	2856.050000	17.60	0.0005
Error	19	3082.950000	162.260526		

Contrast Variable: DRUG3_3

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Mean	1	1394.450000	1394.450000	4.58	0.0455
Error	19	5782.550000	304.344737		

Problem 8.2) Program Output – Continued:

One-way ANOVA Using the Repeated Statement

The ANOVA Procedure
Repeated Measures Analysis of Variance

Repeated Measures Level Information

Dependent Variable	hdl1	hdl2	hdl3
Level of DRUG3	1	2	3

One-way ANOVA Using the Repeated Statement

The ANOVA Procedure
Repeated Measures Analysis of Variance
Univariate Tests of Hypotheses for Within Subject Effects

Source	DF	Anova SS	Mean Square	F Value	Pr > F	Adj Pr > F G - G	Adj Pr > F H - F
DRUG3	2	1503.233333	751.616667	6.81	0.0030	0.0047	0.0035
Error(DRUG3)	38	4193.433333	110.353509				

Greenhouse-Geisser Epsilon 0.8687
Huynh-Feldt Epsilon 0.9485

One-way ANOVA Using the Repeated Statement

The ANOVA Procedure
Repeated Measures Analysis of Variance
Analysis of Variance of Contrast Variables

DRUG3_N represents the contrast between the nth level of DRUG3 and the 3rd

Contrast Variable: DRUG3_1

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Mean	1	259.200000	259.200000	1.33	0.2639
Error	19	3714.800000	195.515789		

Contrast Variable: DRUG3_2

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Mean	1	1394.450000	1394.450000	4.58	0.0455
Error	19	5782.550000	304.344737		

Response: When looking at the analysis of variance table, we find an f value of 6.81 with an associated probability of 0.0030 for the source of the drug variable. We can therefore reject the null hypothesis, 0.0030 < 0.05. Meaning the four drugs are not equally effective for handling a person's good cholesterol (HDL). Now looking at the results of the Student-Newman-Keuls (SNK) test, it shows the two drug groupings.

Assuming that a higher mean indicates a higher raise in patient's good cholesterol, we can say that drug B was the most effective at raising a patient's HDL compared to drug C and drug A. We cannot, at the 0.05 level claim any differences between drug A and drug C.

Problem 8.4)

Program Code:

```
/*************************************************************************************************************************************************************
Filename: C:\Users\Eric\Desktop\STAT 482\hw8_4.sas
Written by: Eric Brown
Date: November 24, 2019

This programuses the statin data set from problem 8.2 and addes the variable GENDER to
the model. As well as test for GENDER and DRUG effects and GENDER by DRUG interaction.
It is understood that the variable DRUG is a repeated measure factor. And the program
uses PROC GLM to start the analysis since the design is unbalanced.

Input: statin.sas7bdat - displayed in previous problem
Output:Two-way ANOVA, Interaction Plot & Analysis
from statin.sas7bdat produced by PROC GLM, PROC MEAN, & PROC GPLOT
*************************************************************************************************************************************************************/
Options ls=89 ps=94 nodate nonumber;
PROC GLM DATA=statin; *using glm instead of anova because design is unbalanced;
  Title 'Two-way ANOVA - Unbalanced Design';
  Class GENDER DRUG; *testing for these effects;
  Model HDL= GENDER | DRUG / SS3; *tell SAS to generate only the Type III sums of squares;
  Lsmeans GENDER | DRUG / PDIFF ADJUST=TUKEY;
RUN;
Title;

PROC MEANS DATA=statin NOPRINT NWAY; *used to get mean hdl value to plot the interaction graph;
  CLASS GENDER DRUG;
  VAR HDL;
  OUTPUT OUT=INTER

SYMBOL1 VALUE=CIRCLE COLOR=BLUE INTERPOL=JOIN;
SYMBOL2 VALUE=SQUARE COLOR=BLACK INTERPOL=JOIN;
PROC GPLOT DATA=INTER;
  TITLE 'Interaction Plot';
  PLOT meanHdl*DRUG=GENDER;
*mean hdl on the y-axis, drug on the x-axis, and gender as the plotting symbol;
RUN;
TITLE;
```

Problem 8.2) SAS Log Window:

```
27   /*************************************************************************************
27 ! ************************************
28   Filename: C:\Users\Eric\Desktop\STAT 482\hw8_4.sas
29   Written by: Eric Brown
30   Date: November 24, 2019
31
32   This programuses the statin data set from problem 8.2 and addes the variable GENDER
32 ! to
33   the model. As well as test for GENDER and DRUG effects and GENDER by DRUG
33 ! interaction.
34   It is understood that the variable DRUG is a repeated measure factor. And the program
35   uses PROC GLM to start the analysis since the design is unbalanced.
36
37   Input: statin.sas7bdat - displayed in previous problem
38   Output:Two-way ANOVA, Interaction Plot & Analysis
39   from statin.sas7bdat produced by PROC GLM, PROC MEAN, & PROC GPLOT
40   *************************************************************************************
40 ! ************************************/
41   Options ls=89 ps=94 nodate nonumber;
42   PROC GLM DATA=statin; *using glm instead of anova because design is unbalanced;
43     Title 'Two-way ANOVA - Unbalanced Design';
44     Class GENDER DRUG; *testing for these effects;
45     Model HDL= GENDER | DRUG / SS3; *tell SAS to generate only the Type III sums of
45 ! squares;
46     Lsmeans GENDER | DRUG / PDIFF ADJUST=TUKEY;
47   RUN;

48   Title;
49

NOTE: PROCEDURE GLM used (Total process time):
      real time           0.07 seconds
      cpu time            0.06 seconds

50   PROC MEANS DATA=statin NOPRINT NWAY; *used to get mean hdl value to plot the
50 ! interaction graph;
51     CLASS GENDER DRUG;
52     VAR HDL;
53     OUTPUT OUT=INTER
54            MEAN=meanHdl;
55   RUN;

NOTE: There were 60 observations read from the data set WORK.STATIN.
NOTE: The data set WORK.INTER has 6 observations and 5 variables.
NOTE: PROCEDURE MEANS used (Total process time):
      real time           0.02 seconds
      cpu time            0.00 seconds


56
57   SYMBOL1 VALUE=CIRCLE COLOR=BLUE INTERPOL=JOIN;
58   SYMBOL2 VALUE=SQUARE COLOR=BLACK INTERPOL=JOIN;
59   PROC GPLOT DATA=INTER;
60     TITLE 'Interaction Plot';
61     PLOT meanHdl*DRUG=GENDER;
62   *mean hdl on the y-axis, drug on the x-axis, and gender as the plotting symbol;
63   RUN;
```

Problem 8.4) Program Output:

Two-way ANOVA - Unbalanced Design

The GLM Procedure

Class Level Information

Class	Levels	Values
GENDER	2	FEMALE MALE
DRUG	3	A B C

Number of Observations Read	60
Number of Observations Used	60

Two-way ANOVA - Unbalanced Design

The GLM Procedure

Dependent Variable: HDL

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	2450.688384	490.137677	4.57	0.0015
Error	54	5785.494949	107.138795		
Corrected Total	59	8236.183333			

R-Square	Coeff Var	Root MSE	HDL Mean
0.297551	22.31575	10.35079	46.38333

Source	DF	Type III SS	Mean Square	F Value	Pr > F
GENDER	1	261.786027	261.786027	2.44	0.1239
DRUG	2	1395.469024	697.734512	6.51	0.0029
GENDER*DRUG	2	685.669024	342.834512	3.20	0.0486

Two-way ANOVA - Unbalanced Design

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

GENDER	HDL LSMEAN	H0:LSMean1=LSMean2 Pr > |t|
FEMALE	44.0740741	0.1239
MALE	48.2727273	

Problem 8.4) Program Output – Continued:

Two-way ANOVA - Unbalanced Design

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey

DRUG	HDL LSMEAN	LSMEAN Number
A	40.8787879	1
B	52.5909091	2
C	45.0505051	3

Least Squares Means for effect DRUG
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: HDL

i/j	1	2	3
1		0.0022	0.4192
2	0.0022		0.0655
3	0.4192	0.0655	

Two-way ANOVA - Unbalanced Design

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

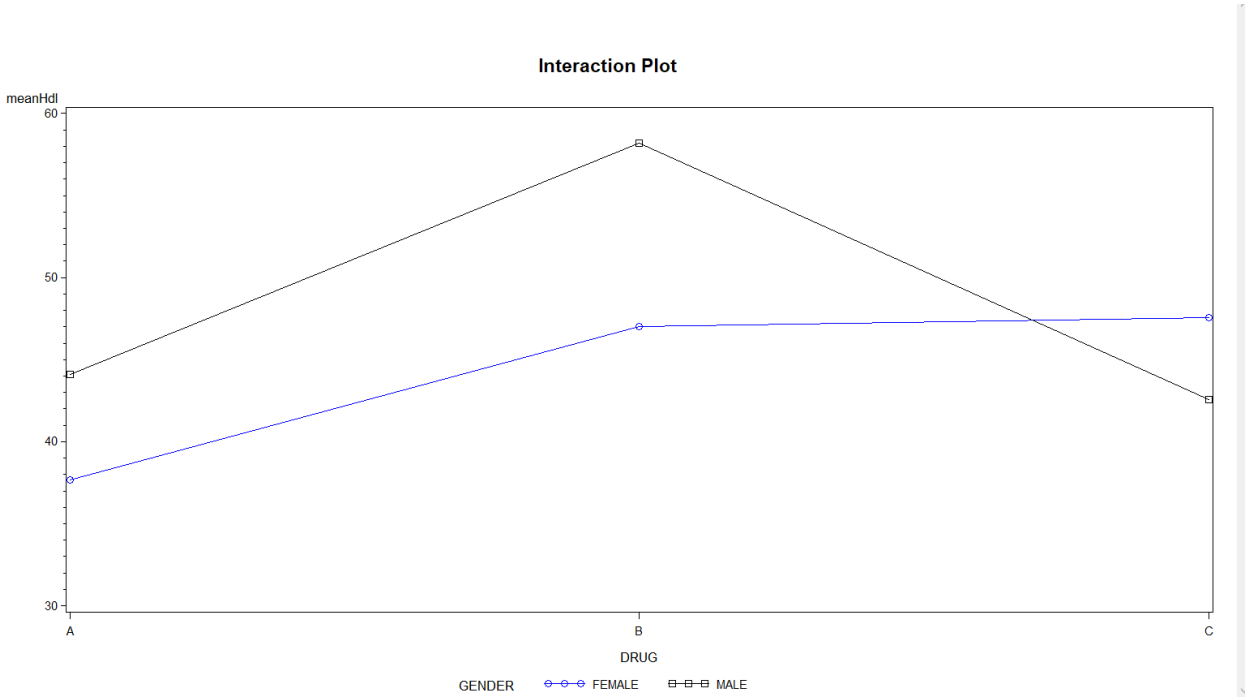
GENDER	DRUG	HDL LSMEAN	LSMEAN Number
FEMALE	A	37.6666667	1
FEMALE	B	47.0000000	2
FEMALE	C	47.5555556	3
MALE	A	44.0909091	4
MALE	B	58.1818182	5
MALE	C	42.5454545	6

Least Squares Means for effect GENDER*DRUG
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: HDL

i/j	1	2	3	4	5	6
1		0.4057	0.3412	0.7381	0.0007	0.8990
2	0.4057		1.0000	0.9887	0.1733	0.9292
3	0.3412	1.0000		0.9753	0.2184	0.8884
4	0.7381	0.9887	0.9753		0.0270	0.9993
5	0.0007	0.1733	0.2184	0.0270		0.0102
6	0.8990	0.9292	0.8884	0.9993	0.0102	

Problem 8.4) Program Output – Continued:



Response: The interaction term is indeed significant since, F-Value for Gender and Drug interaction was 3.20 and the probability was 0.0486, which is less than 0.05.