

Name Jerry

Student Number _____

STA 441s 2024 Quiz 10

1. (5 points) In the Tooth Growth data, guinea pigs were randomly assigned to a combination of Supplement Type (Vitamin C or orange juice), and Dosage Level (Low, Medium, High). The response variable was tooth length after a certain length of time.

- (a) Write $E(y|x)$ for a regression model with *effect coding*. That's the model with an intercept, and zeros, ones and minus ones. You don't have to say how the dummy variables are defined. That will become clear in the next part. In spite of the notation, your dummy variables do not have to be called x .

-5 if no interaction

$$E(y|x) = \beta_0 + \beta_1 s + \beta_2 d_1 + \beta_3 d_2 + \beta_4 sd_1 + \beta_5 sd_2$$

- (b) Make a table with six rows, one for each treatment combination. Make columns showing how your dummy variables are defined. **You are not being asked for the expected values.** They are too messy.

3 if not effect coding

SUPP	Dosage	s	d_1	d_2
VitC	Low	1	1	0
VitC	Med	1	0	1
VitC	High	1	-1	-1
OJ	Low	-1	1	0
OJ	Med	-1	0	1
OJ	High	-1	-1	-1

- (c) Suppose you want to test whether, averaging over Dosage Level, Supplement Type has an effect on average tooth growth. Give the null hypothesis in terms of β values from your regression equation.

If not effect coding, the tests will not be this

$$H_0: \beta_1 = 0$$

- (d) Suppose you want to test whether the effect of Dosage Level depends on Supplement Type. Give the null hypothesis in terms of β values from your regression equation.

$$H_0: \beta_4 = \beta_5 = 0$$

- (e) Suppose you want to test whether, averaging over Supplement Type, Dosage Level has an effect on average tooth growth. Give the null hypothesis in terms of β values from your regression equation.

$$H_0: \beta_2 = \beta_3 = 0$$

If 0-1 dummy coding with product term test (F) for interaction will be correct.

2. For the bunny study, you tested the effect of drug on force required to extract the tooth, conditional on each healing time. You followed up this test of conditional independence with Bonferroni-corrected tests for drug, separately for each healing time.

- (a) (2 points) We want to know if the drug had an effect at 12 days of healing time. Fill in the table below. On your printout, circle the *uncorrected* p -value and write "Question 2" beside it.

-1 for
 $p=0.0011$

Bonferroni-corrected p -value (a number)	Reject Null Hypothesis? (Yes or No)	Statistically Significant? (Yes or No)
0.0044	Yes	Yes

- (b) (3 points) The reason for doing the bunny study was to find out if the drug helped dental implants become more firmly attached to the bone. Based on the results of your statistical analysis (and only on the results of your statistical analysis), would you recommend proceeding to clinical trials with humans? *Answer Yes or No and briefly explain.* You can get full marks for a single sentence with no numbers and no statistical terminology. You have a lot more room than you need.

No! More force was required to extract the teeth without the drug.

or
No! The drug caused the teeth to be less firmly attached.

(No part marks on 2b if the answer was Yes, or if a non-directional conclusion

Attach your log file and your results file to the quiz paper. Make sure your name and student number are written clearly on both printouts.

STA441 S24 Assignment 10 Check

Q4b: Sample sizes

The FREQ Procedure

drugtime	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	5	12.50	5	12.50
2	5	12.50	10	25.00
3	5	12.50	15	37.50
4	5	12.50	20	50.00
11	5	12.50	25	62.50
12	5	12.50	30	75.00
13	5	12.50	35	87.50
14	5	12.50	40	100.00

Frequency	Table of drug by time					
	time(Healing time in days)					Total
	drug	3	6	9	12	
	Saline	5	5	5	5	20
	HEBP	5	5	5	5	20
	Total	10	10	10	10	40

STA441 S24 Assignment 10 Check

Q4c: Cell means and marginal means

The MEANS Procedure

Analysis Variable : force						
day	N Obs	N	Mean	Std Dev	Minimum	Maximum
3	10	10	46.4200000	21.5490036	18.2000000	84.5000000
6	10	10	61.6500000	37.7397006	21.2000000	133.2000000
9	10	10	131.2200000	108.3577080	25.2000000	335.0000000
12	10	10	182.8100000	154.7708733	28.5000000	544.0000000

STA441 S24 Assignment 10 Check

Q4c: Cell means and marginal means

The MEANS Procedure

Analysis Variable : force						
drug	N Obs	N	Mean	Std Dev	Minimum	Maximum
Saline	20	20	148.4850000	136.3112509	21.2000000	544.0000000
HEBP	20	20	62.5650000	40.9851741	18.2000000	188.5000000

STA441 S24 Assignment 10 Check

Q4c: Cell means and marginal means

The MEANS Procedure

Analysis Variable : force							
day	drug	N Obs	N	Mean	Std Dev	Minimum	Maximum
3	Saline	5	5	56.2200000	26.0708074	27.0000000	84.5000000
	HEBP	5	5	36.6200000	11.1813237	18.2000000	47.6000000
6	Saline	5	5	70.6800000	49.8099588	21.2000000	133.2000000
	HEBP	5	5	52.6200000	22.7981798	29.0000000	80.5000000
9	Saline	5	5	191.1800000	125.5112027	37.0000000	335.0000000
	HEBP	5	5	71.2600000	40.9520207	25.2000000	123.6000000
12	Saline	5	5	275.8600000	169.0493656	78.6000000	544.0000000
	HEBP	5	5	89.7600000	60.6061713	28.5000000	188.5000000

STA441 S24 Assignment 10 Check

Q4d: Two-way ANOVA

The GLM Procedure

Class Level Information		
Class	Levels	Values
day	4	3 6 9 12
drug	2	HEBP Saline

Number of Observations Read	40
Number of Observations Used	40

STA441 S24 Assignment 10 Check

Q4d: Two-way ANOVA

The GLM Procedure

Dependent Variable: force

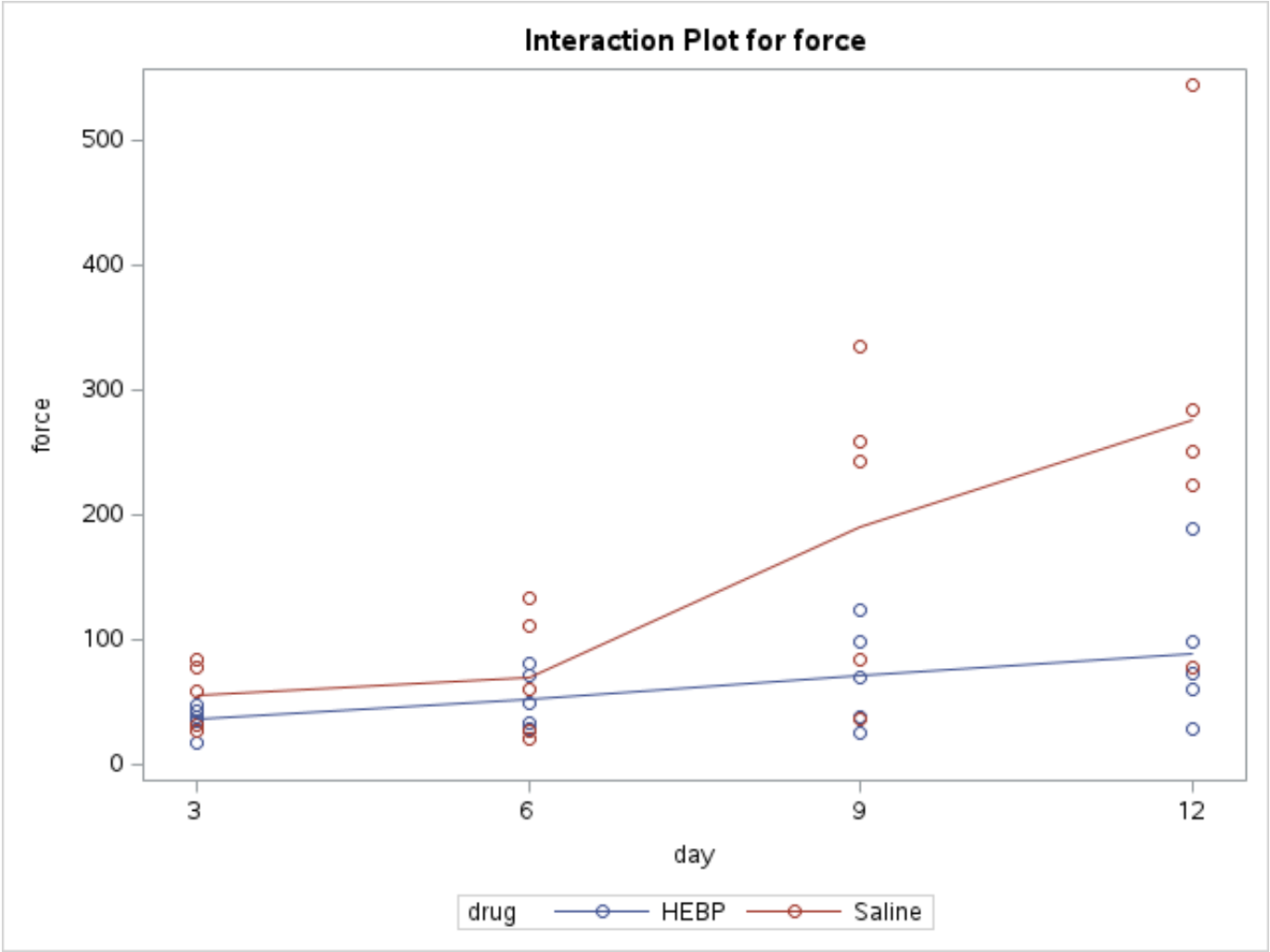
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	244827.0590	34975.2941	5.23	0.0005
Error	32	213945.6960	6685.8030		
Corrected Total	39	458772.7550			

R-Square	Coeff Var	Root MSE	force Mean
0.533656	77.48568	81.76676	105.5250

Source	DF	Type I SS	Mean Square	F Value	Pr > F
--------	----	-----------	-------------	---------	--------

day	3	120516.2090	40172.0697	6.01	0.0023
drug	1	73822.4640	73822.4640	11.04	0.0022
day*drug	3	50488.3860	16829.4620	2.52	0.0757

Source	DF	Type III SS	Mean Square	F Value	Pr > F
day	3	120516.2090	40172.0697	6.01	0.0023
drug	1	73822.4640	73822.4640	11.04	0.0022
day*drug	3	50488.3860	16829.4620	2.52	0.0757



STA441 S24 Assignment 10 Check
Look at the ODS output tables

Obs	Dependent	Source	DF	SS	MS	FValue	ProbF
1	force	Model	7	244827.0590	34975.2941	5.23	0.0005
2	force	Error	32	213945.6960	6685.8030	—	—
3	force	Corrected Total	39	458772.7550	—	—	—

STA441 S24 Assignment 10 Check

Look at the ODS output tables

Obs	Dependent	HypothesisType	Source	DF	SS	MS	FValue	ProbF
1	force	1	day	3	120516.2090	40172.0697	6.01	0.0023
2	force	1	drug	1	73822.4640	73822.4640	11.04	0.0022
3	force	1	day*drug	3	50488.3860	16829.4620	2.52	0.0757
4	force	3	day	3	120516.2090	40172.0697	6.01	0.0023
5	force	3	drug	1	73822.4640	73822.4640	11.04	0.0022
6	force	3	day*drug	3	50488.3860	16829.4620	2.52	0.0757

STA441 S24 Assignment 10 Check

Q4d(ii): Proportions of remaining variation

Take a look at the matrices

anova1				
7	244827.06	34975.294	5.231278	0.0004724
32	213945.7	6685.803	—	—
39	458772.76	—	—	—

anova2					
1	3	120516.21	40172.07	6.0085632	0.0022872
1	1	73822.464	73822.464	11.041675	0.002239
1	3	50488.386	16829.462	2.5171938	0.0757373
3	3	120516.21	40172.07	6.0085632	0.0022872
3	1	73822.464	73822.464	11.041675	0.002239
3	3	50488.386	16829.462	2.5171938	0.0757373

Proportion of remaining variation

Day	Drug	Day_by_Drug
0.3603287	0.2565345	0.1909299

STA441 S24 Assignment 10 Check

Q4n: Does drug have an effect at any time period?

Custom tests with cell means coding

The REG Procedure

Model: MODEL1

Dependent Variable: force

Number of Observations Read	40
Number of Observations Used	40

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

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	8	690248	86281	12.91	<.0001
Error	32	213946	6685.80300		
Uncorrected Total	40	904194			

Root MSE	81.76676	R-Square	0.7634
Dependent Mean	105.52500	Adj R-Sq	0.7042
Coeff Var	77.48568		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
mu11	1	56.22000	36.56721	1.54	0.1340
mu12	1	70.68000	36.56721	1.93	0.0621
mu13	1	191.18000	36.56721	5.23	<.0001
mu14	1	275.86000	36.56721	7.54	<.0001
mu21	1	36.62000	36.56721	1.00	0.3241
mu22	1	52.62000	36.56721	1.44	0.1599
mu23	1	71.26000	36.56721	1.95	0.0601
mu24	1	89.76000	36.56721	2.45	0.0197

STA441 S24 Assignment 10 Check

Q4n: Does drug have an effect at any time period?

Custom tests with cell means coding

The REG Procedure
Model: MODEL1

Test ConditionalIndependence Results for Dependent Variable force				
Source	DF	Mean Square	F Value	Pr > F
Numerator	4	31078	4.65	0.0045
Denominator	32	6685.80300		

STA441 S24 Assignment 10 Check

Q4n: Does drug have an effect at any time period?

Custom tests with cell means coding

The REG Procedure
Model: MODEL1

Test AtDay3 Results for Dependent Variable force				
		Mean		

Source	DF	Square	F Value	Pr > F
Numerator	1	960.40000	0.14	0.7072
Denominator	32	6685.80300		

STA441 S24 Assignment 10 Check
Q4n: Does drug have an effect at any time period?
Custom tests with cell means coding

The REG Procedure
Model: MODEL1

Test AtDay6 Results for Dependent Variable force				
Source	DF	Mean Square	F Value	Pr > F
Numerator	1	815.40900	0.12	0.7292
Denominator	32	6685.80300		

STA441 S24 Assignment 10 Check
Q4n: Does drug have an effect at any time period?
Custom tests with cell means coding

The REG Procedure
Model: MODEL1

Test AtDay9 Results for Dependent Variable force				
Source	DF	Mean Square	F Value	Pr > F
Numerator	1	35952	5.38	0.0269
Denominator	32	6685.80300		

STA441 S24 Assignment 10 Check
Q4n: Does drug have an effect at any time period?
Custom tests with cell means coding

The REG Procedure
Model: MODEL1

Test AtDay12 Results for Dependent Variable force				
Source	DF	Mean Square	F Value	Pr > F
Numerator	1	86583	12.95	0.0011
Denominator	32	6685.80300		

Question 2

STA441 S24 Assignment 10 Check
Q4o: Follow up with randomization

The Multtest Procedure

Model Information	
Test for continuous variables	Mean t-test
Degrees of Freedom Method	Pooled
Tails for continuous tests	Two-tailed
Strata weights	None
P-value adjustment	Permutation
Center continuous variables	No
Number of resamples	20000
Seed	99999

Contrast Coefficients									
Contrast		drugtime							
		1	2	3	4	11	12	13	14
AtDay3	Centered	1	0	0	0	-1	0	0	0
AtDay6	Centered	0	1	0	0	0	-1	0	0
AtDay9	Centered	0	0	1	0	0	0	-1	0
AtDay12	Centered	0	0	0	1	0	0	0	-1

Continuous Variable Tabulations				
Variable	drugtime	NumObs	Mean	Standard Deviation
force	1	5	56.2200	26.0708
force	2	5	70.6800	49.8100
force	3	5	191.1800	125.5112
force	4	5	275.8600	169.0494
force	11	5	36.6200	11.1813
force	12	5	52.6200	22.7982
force	13	5	71.2600	40.9520
force	14	5	89.7600	60.6062

p-Values			
Variable	Contrast	Raw	Permutation
force	AtDay3	0.7072	0.9941
force	AtDay6	0.7292	0.9957
force	AtDay9	0.0269	0.1149
force	AtDay12	0.0011	0.0062

```

1      OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;
NOTE: ODS statements in the SAS Studio environment may disable some output features.
73
74      /* 441s24HW10bunnies.sas */
75      title 'STA441 S24 Assignment 10 Check';
76
77      proc format;
78          value dfmt 0 = 'Saline' 1 = 'HEBP';
NOTE: Format DFMT is already on the library WORK.FORMATS.
NOTE: Format DFMT has been output.
79          value hfmt 1 = '3' 2 = '6' 3 = '9' 4 = '12';
NOTE: Format HFMT is already on the library WORK.FORMATS.
NOTE: Format HFMT has been output.
80

```

```

NOTE: PROCEDURE FORMAT used (Total process time):
      real time           0.00 seconds
      user cpu time       0.00 seconds
      system cpu time     0.00 seconds
      memory              246.59k
      OS Memory           31396.00k
      Timestamp           03/22/2024 03:50:19 PM
      Step Count          82  Switch Count  0
      Page Faults         0
      Page Reclaims       24
      Page Swaps          0
      Voluntary Context Switches 0
      Involuntary Context Switches 0
      Block Input Operations 0
      Block Output Operations 32

```

```

81      data implant;
82          infile '/home/u1407221/441s24/data/bunnies.data.txt';
83          input id $ day drug stiff force preload;
84          time = day/3;
85          drugtime = 10*drug + time;
86          /* Dummy vars for cell means coding */
87          if drugtime= 1 then mu11=1 ; else mu11=0;
88          if drugtime= 2 then mu12=1 ; else mu12=0;
89          if drugtime= 3 then mu13=1 ; else mu13=0;
90          if drugtime= 4 then mu14=1 ; else mu14=0;
91          if drugtime=11 then mu21=1 ; else mu21=0;
92          if drugtime=12 then mu22=1 ; else mu22=0;
93          if drugtime=13 then mu23=1 ; else mu23=0;
94          if drugtime=14 then mu24=1 ; else mu24=0;
95          format drug dfmt.;
96          format time hfmt.;
97          label time = 'Healing time in days';
98

```

```

NOTE: The infile '/home/u1407221/441s24/data/bunnies.data.txt' is:
      Filename=/home/u1407221/441s24/data/bunnies.data.txt,
      Owner Name=u1407221,Group Name=oda,
      Access Permission=-rw-r--r--,
      Last Modified=18Mar2024:13:29:55,
      File Size (bytes)=1027

```

```

NOTE: 40 records were read from the infile '/home/u1407221/441s24/data/bunnies.data.txt'.
      The minimum record length was 21.
      The maximum record length was 26.

```

```

NOTE: The data set WORK.IMPLANT has 40 observations and 16 variables.

```

```

NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      user cpu time       0.01 seconds
      system cpu time     0.00 seconds
      memory              782.00k
      OS Memory           32168.00k
      Timestamp           03/22/2024 03:50:19 PM
      Step Count          83  Switch Count  2
      Page Faults         0
      Page Reclaims       145
      Page Swaps          0
      Voluntary Context Switches 20
      Involuntary Context Switches 0
      Block Input Operations 0
      Block Output Operations 264

```

```

99      proc freq data=implant;
100          title2 'Q4b: Sample sizes';
101          tables drugtime;
102          tables drug*time / norow nocol nopercnt;
103

```

NOTE: There were 40 observations read from the data set WORK.IMPLANT.

NOTE: PROCEDURE FREQ used (Total process time):

real time	0.04 seconds
user cpu time	0.04 seconds
system cpu time	0.01 seconds
memory	3715.96k
OS Memory	33456.00k
Timestamp	03/22/2024 03:50:19 PM
Step Count	84 Switch Count 5
Page Faults	0
Page Reclaims	373
Page Swaps	0
Voluntary Context Switches	28
Involuntary Context Switches	0
Block Input Operations	0
Block Output Operations	536

```

104      proc means data=implant;
105          title2 'Q4c: Cell means and marginal means';
106          class day;
107          var force;

```

NOTE: There were 40 observations read from the data set WORK.IMPLANT.

NOTE: PROCEDURE MEANS used (Total process time):

real time	0.02 seconds
user cpu time	0.03 seconds
system cpu time	0.01 seconds
memory	9102.34k
OS Memory	41916.00k
Timestamp	03/22/2024 03:50:19 PM
Step Count	85 Switch Count 2
Page Faults	0
Page Reclaims	2148
Page Swaps	0
Voluntary Context Switches	15
Involuntary Context Switches	0
Block Input Operations	0
Block Output Operations	8

```

108      proc means data=implant;
109          class drug;
110          var force;

```

NOTE: There were 40 observations read from the data set WORK.IMPLANT.

NOTE: PROCEDURE MEANS used (Total process time):

real time	0.02 seconds
user cpu time	0.01 seconds
system cpu time	0.00 seconds
memory	8983.59k
OS Memory	41916.00k
Timestamp	03/22/2024 03:50:19 PM
Step Count	86 Switch Count 1
Page Faults	0
Page Reclaims	1820
Page Swaps	0
Voluntary Context Switches	12
Involuntary Context Switches	0
Block Input Operations	0
Block Output Operations	24

```

111      proc means data=implant;
112          class day drug;
113          var force;
114

```

NOTE: There were 40 observations read from the data set WORK.IMPLANT.

NOTE: PROCEDURE MEANS used (Total process time):

real time	0.03 seconds
user cpu time	0.03 seconds
system cpu time	0.01 seconds

```

memory          8735.53k
OS Memory       42684.00k
Timestamp       03/22/2024 03:50:19 PM
Step Count      87  Switch Count  1
Page Faults     0
Page Reclaims   1828
Page Swaps      0
Voluntary Context Switches  24
Involuntary Context Switches 0
Block Input Operations  0
Block Output Operations  0

```

```

115      proc glm data=implant;
116          title2 'Q4d: Two-way ANOVA';
117          class day drug;
118          model force = day|drug;
119          ods output OverallANOVA = Overall
120                  ModelANOVA    = TwoWay;
121      run;

```

NOTE: The data set WORK.TWOWAY has 6 observations and 8 variables.

NOTE: The data set WORK.OVERALL has 3 observations and 7 variables.

122

NOTE: PROCEDURE GLM used (Total process time):

```

real time      0.27 seconds
user cpu time   0.13 seconds
system cpu time 0.02 seconds
memory         23271.39k
OS Memory       53440.00k
Timestamp       03/22/2024 03:50:20 PM
Step Count     88  Switch Count  8
Page Faults    0
Page Reclaims  5836
Page Swaps     0
Voluntary Context Switches  712
Involuntary Context Switches 1
Block Input Operations  0
Block Output Operations  1904

```

```

123      proc print data=Overall;
124          title2 'Look at the ODS output tables';

```

NOTE: There were 3 observations read from the data set WORK.OVERALL.

NOTE: PROCEDURE PRINT used (Total process time):

```

real time      0.01 seconds
user cpu time   0.01 seconds
system cpu time 0.00 seconds
memory         674.65k
OS Memory       51880.00k
Timestamp       03/22/2024 03:50:20 PM
Step Count     89  Switch Count  1
Page Faults    0
Page Reclaims  71
Page Swaps     0
Voluntary Context Switches  7
Involuntary Context Switches 0
Block Input Operations  0
Block Output Operations  8

```

```

125      proc print data=TwoWay;
126

```

NOTE: There were 6 observations read from the data set WORK.TWOWAY.

NOTE: PROCEDURE PRINT used (Total process time):

```

real time      0.01 seconds
user cpu time   0.02 seconds
system cpu time 0.00 seconds
memory         679.34k
OS Memory       51880.00k
Timestamp       03/22/2024 03:50:20 PM
Step Count     90  Switch Count  0
Page Faults    0
Page Reclaims  69
Page Swaps     0
Voluntary Context Switches  0
Involuntary Context Switches 0

```

```
Block Input Operations      0
Block Output Operations     8
```

```
127      proc iml;
NOTE: IML Ready
128          title2 'Q4d(ii): Proportions of remaining variation';
129          use overall;
129      !      read all into anova1;
129      !      /* "all" means all cases */
130          use TwoWay;
130      !      read all into anova2;
131          print "Take a look at the matrices";
132          print anova1;
132      !      print anova2;
133          print "Proportion of remaining variation";
134          NminusP = anova1[2,1];
135          s1 = anova2[4,2];
135      !      F1 = anova2[4,5];
135      !      /* Main effect for Day */
136          s2 = anova2[5,2];
136      !      F2 = anova2[5,5];
136      !      /* Main effect for Drug */
137          s3 = anova2[6,2];
137      !      F3 = anova2[6,5];
137      !      /* Interaction */
138          Day = s1*F1/(s1*F1 + NminusP);
139          Drug = s2*F2/(s2*F2 + NminusP);
140          Day_by_Drug = s3*F3/(s3*F3 + NminusP);
141          print Day Drug Day_by_Drug;
142
NOTE: Exiting IML.
NOTE: PROCEDURE IML used (Total process time):
real time          0.03 seconds
user cpu time      0.03 seconds
system cpu time    0.00 seconds
memory            1142.71k
OS Memory          52140.00k
Timestamp          03/22/2024 03:50:20 PM
Step Count         91  Switch Count  1
Page Faults        0
Page Reclaims      131
Page Swaps         0
Voluntary Context Switches  8
Involuntary Context Switches 0
Block Input Operations  0
Block Output Operations 16

143      proc reg data=implant plots=none;
144          title2 'Q4n: Does drug have an effect at any time period?';
145          title3 'Custom tests with cell means coding';
146          model force = mu11--mu24 / noint;
147          ConditionalIndependence:
148              test mu11=mu21, mu12=mu22, mu13=mu23, mu14=mu24;
149          AtDay3: test mu11=mu21;
150          AtDay6: test mu12=mu22;
151          AtDay9: test mu13=mu23;
152          AtDay12: test mu14=mu24;
153
154      /* Cross-checked in 442s12HW8 with contrasts. 2018 had randomization. */
155
```

```
NOTE: PROCEDURE REG used (Total process time):
real time          0.08 seconds
user cpu time      0.08 seconds
system cpu time    0.00 seconds
memory            3184.40k
OS Memory          53952.00k
Timestamp          03/22/2024 03:50:20 PM
Step Count         92  Switch Count  3
Page Faults        0
Page Reclaims      338
Page Swaps         0
Voluntary Context Switches 27
Involuntary Context Switches 0
Block Input Operations  0
Block Output Operations 96
```

```
156      proc multtest permutation nsample=20000 seed=99999 data=implant;
157          title2 'Q4o: Follow up with randomization';
158          class drugtime; /* 01 02 03 04 11 12 13 14 */
159          contrast 'AtDay3'      1 0 0 0 -1 0 0 0;
160          contrast 'AtDay6'      0 1 0 0 0 -1 0 0;
161          contrast 'AtDay9'      0 0 1 0 0 0 -1 0;
162          contrast 'AtDay12'     0 0 0 1 0 0 0 -1;
163          test mean(force);
164
165      run;
```

NOTE: The multiple testing procedure for this run is not closed. In cases with badly heteroskedastic data, tests for individual null hypotheses can have inflated familywise Type I error rates.

NOTE: There were 40 observations read from the data set WORK.IMPLANT.

NOTE: PROCEDURE MULTTEST used (Total process time):

real time	0.07 seconds
user cpu time	0.08 seconds
system cpu time	0.00 seconds
memory	776.90k
OS Memory	52392.00k
Timestamp	03/22/2024 03:50:20 PM
Step Count	93 Switch Count 1
Page Faults	0
Page Reclaims	62
Page Swaps	0
Voluntary Context Switches	9
Involuntary Context Switches	0
Block Input Operations	0
Block Output Operations	80

```
166      quit;
167
168
169
170      OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;
182
```