Advanced Biostatistics **Quiz 3** **Name** _____
March 16, 2005 **9 or 10 Total Points**

Directions: Thoroughly, clearly and neatly answer the following two problems in the space given, showing all relevant calculations. Unless otherwise noted, use $\alpha = 5\%$ throughout.

1. (1 + 1 + 1.5 + 2 = 5.5 points) Medical researchers, wishing to test the efficacy of six (6) drugs, labeled 1-6, used healthy volunteers in a study. In the study, the efficacy endpoint was labeled Y, where higher levels of Y are associated with higher efficacy. Twelve volunteers were chosen and grouped into three "blocks", where blocks correspond to groups of people that are similar in age, diet, and lifestyle. Since only four volunteers were available in each block, an incomplete block design was used in which the following drugs are randomized to patients in the respective blocks. The data are given and analyzed on pp.3-4 of the *Appendix*.

Block 1	1	4	2	5
Block 2	2	5	3	6
Block 3	3	6	1	4

(a) Give the calculated test statistic and p-value for these data used to test whether the response is the same for the six drugs.

(b) Why can this output ***not*** be used to test for the significance of blocks?

(c) Which drug(s) is (are) best? Be as clear and as detailed as possible.

(d) Is this design balanced? Support your claim(s) in a clear and detailed manner.

2. (1 + 1 + 1.5 + 1= 3.5 or 4.5 points) A crossover design was used to compare drugs for the control of hypertension. Two drugs, A and B, were used alone and in combination. The combination of the two drugs was labeled as drug C in the experiment. Subjects were randomly assigned to one of the six sequences of the drug treatments listed on p.5 of the *Appendix*. Each treatment period lasted four weeks with a one-week washout period between treatments. The systolic blood pressure (the response variable, labeled 'sbp') of the subjects was measured at the end of each period, and this data is analyzed on pp.5-6 of the *Appendix*. Based on this analysis, answer the following.

   (a) Do you feel that the one-week washout period was long enough? Why or why not?

   (b) Is there evidence that the average sbp's differed for the six sequences? Give the relevant calculated test statistics and p-value along with your conclusion.

   (c) Noting that drug corresponds to the factor 'trt' (short for treatment) in the output, summarize the treatment results here ***using the underline method***. Which drug(s) [or drug combination(s)] is (are) best?

   (d) [Mandatory for G students; EC for UG students] Do these drugs exhibit interaction? Support your claim(s).

First Exercise – Data, Program and Output

Block 1	65 (1)	99 (4)	74 (2)	91 (5)
Block 2	106 (2)	104 (5)	125 (3)	96 (6)
Block 3	143 (3)	128 (6)	95 (1)	140 (4)

```
proc glm;
  class block drug;
  model y=drug block;
  means drug/snk;
  lsmeans drug/pdiff;
run;
```

```
                                The GLM Procedure
Dependent Variable: y

                                          Sum of
      Source                     DF         Squares     Mean Square    F Value    Pr > F
      Model                       7     6353.500000      907.642857      16.69    0.0082
      Error                       4      217.500000       54.375000
      Corrected Total            11     6571.000000

                      R-Square     Coeff Var      Root MSE        y Mean
                      0.966900      6.989517      7.373941      105.5000

      Source                     DF       Type I SS     Mean Square    F Value    Pr > F
      drug                        5     4010.000000      802.000000      14.75    0.0110
      block                       2     2343.500000     1171.750000      21.55    0.0072

      Source                     DF     Type III SS     Mean Square    F Value    Pr > F
      drug                        5     2407.000000      481.400000       8.85    0.0276
      block                       2     2343.500000     1171.750000      21.55    0.0072
```

Student-Newman-Keuls Test for y

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	4
Error Mean Square	54.375

Number of Means	2	3	4	5	6
Critical Range	20.473352	26.28064	30.018238	32.781571	34.968458

Means with the same letter are not significantly different.

SNK Grouping	Mean	N	drug
A	134.000	2	3
A			
B A	119.500	2	4
B A			
B A	112.000	2	6
B			
B C	97.500	2	5
B C			
B C	90.000	2	2
C			
C	80.000	2	1

The GLM Procedure
Least Squares Means

drug	y LSMEAN	LSMEAN Number
1	79.583333	1
2	100.083333	2
3	124.333333	3
4	119.083333	4
5	107.583333	5
6	102.333333	6

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

Dependent Variable: y

i/j	1	2	3	4	5	6
1		0.0617	0.0049	0.0059	0.0245	0.0461
2	0.0617		0.0382	0.0755	0.3666	0.7916
3	0.0049	0.0382		0.5458	0.1033	0.0406
4	0.0059	0.0755	0.5458		0.2223	0.1033
5	0.0245	0.3666	0.1033	0.2223		0.5458
6	0.0461	0.7916	0.0406	0.1033	0.5458	

NOTE: To ensure overall protection level, only probabilities associated with pre-planned comparisons should be used.

Second Exercise

Sequence	Patients
ABC	1-4
ACB	5-8
BAC	9-12
BCA	13-16
CAB	17-20
CBA	21-24

```
proc glm;
  class seq subject per trt co;
  model sbp=seq subject(seq) per trt co;
run;
```

```
Dependent Variable: sbp

                                        Sum of
      Source                    DF       Squares     Mean Square    F Value    Pr > F
      Model                     29     40456.56944      1395.05412       5.03    <.0001
      Error                     42     11648.08333       277.33532
      Corrected Total           71     52104.65278

                      R-Square     Coeff Var      Root MSE      sbp Mean
                      0.776448      9.352187      16.65339      178.0694


      Source                    DF       Type I SS     Mean Square    F Value    Pr > F
      seq                        5      3888.90278       777.78056       2.80    0.0284
      subject(seq)              18     27066.41667      1503.68981       5.42    <.0001
      per                        2      1189.19444       594.59722       2.14    0.1298
      trt                        2      8219.69444      4109.84722      14.82    <.0001
      co                         2        92.36111        46.18056       0.17    0.8472


      Source                    DF     Type III SS     Mean Square    F Value    Pr > F
      seq                        5      3974.57500       794.91500       2.87    0.0258
      subject(seq)              18     27066.41667      1503.68981       5.42    <.0001
      per                        1       609.18750       609.18750       2.20    0.1458
      trt                        2      6010.61667      3005.30833      10.84    0.0002
      co                         2        92.36111        46.18056       0.17    0.8472
```

```
proc glm data=one;
  class seq subject per trt;
  model sbp=seq subject(seq) per trt carry carry2;
  test h=seq e=subject(seq)/htype=1 etype=1;
  lsmeans trt/pdiff;
run;
```

```
Dependent Variable: sbp

                                        Sum of
      Source                    DF       Squares     Mean Square    F Value    Pr > F
      Model                     29     40456.56944      1395.05412       5.03    <.0001
      Error                     42     11648.08333       277.33532
      Corrected Total           71     52104.65278
```

R-Square	Coeff Var	Root MSE	sbp Mean
0.776448	9.352187	16.65339	178.0694

Source	DF	Type I SS	Mean Square	F Value	Pr > F
seq	5	3888.90278	777.78056	2.80	0.0284
subject(seq)	18	27066.41667	1503.68981	5.42	<.0001
per	2	1189.19444	594.59722	2.14	0.1298
trt	2	8219.69444	4109.84722	14.82	<.0001
carry	1	0.52083	0.52083	0.00	0.9656
carry2	1	91.84028	91.84028	0.33	0.5681

Source	DF	Type III SS	Mean Square	F Value	Pr > F
seq	5	3974.57500	794.91500	2.87	0.0258
subject(seq)	18	27066.41667	1503.68981	5.42	<.0001
per	2	781.79310	390.89655	1.41	0.2556
trt	2	6010.61667	3005.30833	10.84	0.0002
carry	1	88.02083	88.02083	0.32	0.5762
carry2	1	91.84028	91.84028	0.33	0.5681

Tests of Hypotheses Using the Type I MS for subject(seq) as an Error Term

Source	DF	Type I SS	Mean Square	F Value	Pr > F
seq	5	3888.902778	777.780556	0.52	0.7599

The GLM Procedure
Least Squares Means

trt	sbp LSMEAN	LSMEAN Number
1	190.652778	1
2	165.631944	2
3	177.923611	3

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

Dependent Variable: sbp

i/j	1	2	3
1		<.0001	0.0225
2	<.0001		0.0273
3	0.0225	0.0273	

NOTE: To ensure overall protection level, only probabilities associated with pre-planned comparisons should be used.