Directions: Answer all three exercises, showing all relevant work. As always, conclusions and justifications are to be given in clear detailed English. For each exercise and part, be sure to clearly write down all needed assumptions and requirements. Please type up your solutions or write ***very*** neatly.

1. [**F**rom Kuehl, p.326] A horticulturalist studied the germination of tomato seed with four different temperatures (25C, 30C, 35C and 40C) such that each "run" of the experiment included only two different temperatures because there were only two growth chambers available for the study. The two experimental temperatures were randomly assigned to the chambers for each run, but the researcher feels certain there may be run-to-run variability. The data that follow are germination rates of the tomato seed.
   (a) It is claimed the design is an IBD. Explain why this is so. Is this IBD balanced? Thoroughly justify your answer and give the IBD "parameters" (see the 6 symbols defined in the box on p.13 of the Chapter 3 Course Notes). Pay particular attention to how you define the λ parameter here.
   (b) Give all the needed assumptions and thoroughly analyze these data. Summarize your findings using the underline method for the treatments (25C, 30C, 35C and 40C), and comment on the model fit.
   (c) What is the commonly used term for "runs" in this experiment?

	25C	**30C**	**35C**	**40C**
Run 1	**24.65**	**----**	**----**	**18.62**
Run 2	**----**	**24.11**	**----**	**17.08**
Run 3	**22.31**	**21.25**	**----**	**----**
Run 4	**----**	**----**	**17.95**	**18.93**
Run 5	**28.90**	**----**	**18.27**	**----**
Run 6	**----**	**25.53**	**20.91**	**----**

2. Associative effects occur in animal diets when feedstuffs are combined and diet utilization or animal performance is different from that predicted from a sum of the individual ingredients. The addition of roughage to the diets of ruminant animals has been shown to influence various diet utilization factors such as ruminal retention time. However, information about the relative associative effects of different roughage was scarce, especially in mixed feedlot diets. An animal scientist hypothesized roughage source could influence utilization of mixed diets of beef steers by altering ruminal digestion of other diet ingredients. The basic mixed diet for the beef steers was a 65% concentrate based on steam flaked milo and 35% roughage. Three roughage treatments were used with

   (A) 35% alfalfa hay as a control treatment
   (B) 17.5% wheat straw and 17.5% alfalfa
   (C) 17.5% cottonseed hulls and 17.5% alfalfa.

   Twelve beef steers were available for the study. Each of the three roughage diets was fed to the steers in one of six possible sequences of the three diets:

Sequence 1	**Sequence 2**	**Sequence 3**	**Sequence 4**	**Sequence 5**	**Sequence 6**
A → B → C	**B → C → A**	**C → A → B**	**A → C → B**	**B → A → C**	**C → B → A**

   Each diet in each sequence was fed to two steers for 30 days. The steers were allowed a period of 21 days to adapt to a diet change before any data was collected. The Neutral Detergent Fiber (NDF) digestion coefficient, which indicates the percent of dietary fiber digested by the steer, was calculated for each steer on each diet. The raw data are given below (and analyzed in the attached), with each row corresponding to a different steer and the variables are sequence, treatment, NDF, treatment, NDF, treatment and NDF for each steer in each of the 3 periods.

Sequence	Treatment	Y	Treatment	Y	Treatment	Y
ABC	A	50	B	61	C	53
ABC	A	55	B	63	C	57
BCA	B	44	C	42	A	57
BCA	B	51	C	46	A	59
CAB	C	35	A	55	B	47
CAB	C	41	A	56	B	50
ACB	A	54	C	48	B	51
ACB	A	58	C	51	B	54
BAC	B	50	A	57	C	51
BAC	B	55	A	57	C	55
CBA	C	41	B	56	A	58
CBA	C	46	B	58	A	61

(a) Using the correct test, test for a "sequence" effect, giving the test statistic, degrees of freedom, and clear conclusion.
(b) Do you feel that there is a carryover effect here? Thoroughly justify your answer.
(c) Do you feel the treatments differ? Summarize your findings using the underline method. Is it correct to use the "Means" here or the "LSMeans"? Why?
(d) Comment on the fit here – do all the needed assumptions appear to be met? Be clear and detailed.

3. Discuss the uses of blocking in biostatistical studies addressing the following:
(a) Why is it (blocking) used and useful?
(b) Why is it (blocking) not used more?
(c) In what sense (if any) is blocking used in CODs and split-plot designs?
(d) How is the analysis of a RCBD with one blocking and one treatment factor different from a two-way ANOVA with two treatment factors?

SAS Program for Exercise One

```
data tomato;
  do run=1 to 6;
  do rep=1 to 2;
    input y temp @@; drop rep;
  output; end; end; datalines;
24.65 25 18.62 40 24.11 30 17.08 40 22.31 25 21.25 30
17.95 35 18.93 40 28.90 25 18.27 35 25.53 30 20.91 35
;
proc glm;
  class run temp;
  model y=run temp;
  lsmeans temp/ pdiff;
run;
```

SAS Output for Exercise One

```
                          The GLM Procedure

Dependent Variable: y
                                    Sum of
Source                  DF         Squares     Mean Square    F Value    Pr > F
Model                    8     139.5838000      17.4479750       7.85    0.0588
Error                    3       6.6678250       2.2226083
Corrected Total         11     146.2516250
```

R-Square	Coeff Var	Root MSE	y Mean
0.954409	6.920466	1.490841	21.54250

Source	DF	Type I SS	Mean Square	F Value	Pr > F
run	5	35.1480750	7.0296150	3.16	0.1861
temp	3	104.4357250	34.8119083	15.66	0.0245

Source	DF	Type III SS	Mean Square	F Value	Pr > F
run	5	32.4003083	6.4800617	2.92	0.2038
temp	3	104.4357250	34.8119083	15.66	0.0245

Least Squares Means

temp	y LSMEAN	LSMEAN Number
25	25.9725000	1
30	24.1900000	2
35	17.4850000	3
40	18.5225000	4

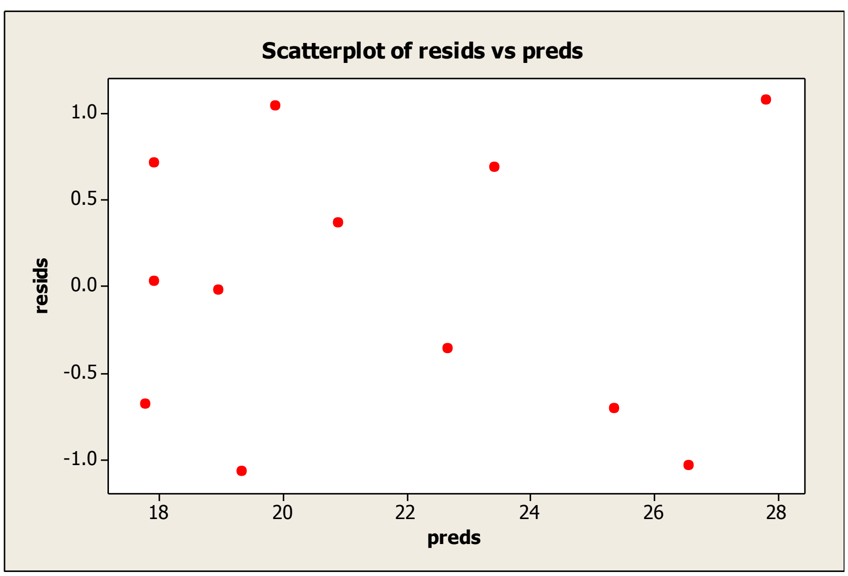
Least Squares Means for effect temp
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: y

i/j	1	2	3	4
1		0.3177	0.0107	0.0154
2	0.3177		0.0205	0.0320
3	0.0107	0.0205		0.5365
4	0.0154	0.0320	0.5365	

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

Residual Plot for Exercise One



First SAS Program for Exercise Two

```
data one;
  do seq='abc','bca','cab','acb','bac','cba';
  do rep=1,2;
  do period=1,2,3;
    input trt $ y carry @@; carry2=carry*carry; output;
  end; end; end; datalines;
a 50 0 b 61 1 c 53 2 a 55 0 b 63 1 c 57 2
b 44 0 c 42 2 a 57 3 b 51 0 c 46 2 a 59 3
c 35 0 a 55 3 b 47 1 c 41 0 a 56 3 b 50 1
a 54 0 c 48 1 b 51 3 a 58 0 c 51 1 b 54 3
b 50 0 a 57 2 c 51 1 b 55 0 a 57 2 c 55 1
c 41 0 b 56 3 a 58 2 c 46 0 b 58 3 a 61 2
;
proc glm;
  class seq rep period trt carry;
  model y=seq rep(seq) period trt carry;
  test h=seq e=rep(seq)/etype=1 htype=1; run;
```

First SAS Output for Exercise Two

```
                              The GLM Procedure
                            Class Level Information
                     Class         Levels    Values
                     seq                6    abc acb bac bca cab cba
                     rep                2    1 2
                     period             3    1 2 3
                     trt                3    a b c
                     carry              4    0 1 2 3
                     Number of Observations Read          36
                     Number of Observations Used          36

Dependent Variable: y
Source                      DF    Sum of Squares     Mean Square    F Value    Pr > F
Model                       17       1267.513889       74.559641       7.89    <.0001
Error                       18        170.125000        9.451389
Corrected Total             35       1437.638889

                 R-Square     Coeff Var      Root MSE        y Mean
                 0.881664      5.877600      3.074311      52.30556

Source                      DF       Type I SS     Mean Square    F Value    Pr > F
seq                          5     318.4722222      63.6944444       6.74    0.0011
rep(seq)                     6     111.8333333      18.6388889       1.97    0.1235
period                       2     284.3888889     142.1944444      15.04    0.0001
trt                          2     532.3888889     266.1944444      28.16    <.0001
carry                        2      20.4305556      10.2152778       1.08    0.3603

Source                      DF     Type III SS     Mean Square    F Value    Pr > F
seq                          5     317.0916667      63.4183333       6.71    0.0011
rep(seq)                     6     111.8333333      18.6388889       1.97    0.1235
period                       1       0.3750000       0.3750000       0.04    0.8443
trt                          2     425.4750000     212.7375000      22.51    <.0001
carry                        2      20.4305556      10.2152778       1.08    0.3603

    Tests of Hypotheses Using the Type I MS for rep(seq) as an Error Term
Source                      DF       Type I SS     Mean Square    F Value    Pr > F
seq                          5     318.4722222      63.6944444       3.42    0.0833
```

Second SAS Program for Exercise Two

```
proc glm;
  class seq rep period trt;
  model y=seq rep(seq) period trt carry carry2;
  test h=seq e=rep(seq)/etype=1 htype=1;
  means trt/tukey;
  lsmeans trt/pdiff;
run;
```

Second SAS Output for Exercise Two

```
                         The GLM Procedure
                       Class Level Information
                Class        Levels    Values
                seq               6    abc acb bac bca cab cba
                rep               2    1 2
                period            3    1 2 3
                trt               3    a b c

                    Number of Observations Read          36
                    Number of Observations Used          36

                                   Sum of
Source                     DF         Squares     Mean Square    F Value    Pr > F
Model                      17     1267.513889       74.559641       7.89    <.0001
Error                      18      170.125000        9.451389
Corrected Total            35     1437.638889

                 R-Square     Coeff Var      Root MSE        y Mean
                 0.881664      5.877600      3.074311      52.30556

Source                     DF       Type I SS     Mean Square    F Value    Pr > F
seq                         5     318.4722222      63.6944444       6.74    0.0011
rep(seq)                    6     111.8333333      18.6388889       1.97    0.1235
period                      2     284.3888889     142.1944444      15.04    0.0001
trt                         2     532.3888889     266.1944444      28.16    <.0001
carry                       1       0.3750000       0.3750000       0.04    0.8443
carry2                      1      20.0555556      20.0555556       2.12    0.1624

Source                     DF     Type III SS     Mean Square    F Value    Pr > F
seq                         5     317.0916667      63.4183333       6.71    0.0011
rep(seq)                    6     111.8333333      18.6388889       1.97    0.1235
period                      2      56.3922414      28.1961207       2.98    0.0760
trt                         2     425.4750000     212.7375000      22.51    <.0001
carry                       1      20.4294218      20.4294218       2.16    0.1588
carry2                      1      20.0555556      20.0555556       2.12    0.1624

    Tests of Hypotheses Using the Type I MS for rep(seq) as an Error Term
Source                     DF       Type I SS     Mean Square    F Value    Pr > F
seq                         5     318.4722222      63.6944444       3.42    0.0833
```

```
                    Tukey's Studentized Range (HSD) Test for y
NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher
                             Type II error rate than REGWQ.

                       Alpha                                 0.05
                       Error Degrees of Freedom                18
```

```
                Error Mean Square                   9.451389
                Critical Value of Studentized Range  3.60930
                Minimum Significant Difference        3.2032
        Means with the same letter are not significantly different.

          Tukey Grouping          Mean      N    trt
                       A        56.417     12    a
                       A
                       A        53.333     12    b

                       B        47.167     12    c
```

```
                              Least Squares Means
                                                  LSMEAN
                          trt         y LSMEAN      Number
                          a         56.7430556           1
                          b         52.8055556           2
                          c         47.3680556           3

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

                              Dependent Variable: y
                  i/j              1             2             3
                    1                       0.0117        <.0001
                    2         0.0117                      0.0011
                    3         <.0001        0.0011

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

Residual Plot for Exercise Two