Biostats VHM 802 Course
Winter 2008, Atlantic Veterinary College, PEI
Henrik Stryhn

Final examination, 23 April 2008

All aids are allowed, except personal computer and personal assistance. The exam consists of 3 questions, which have equal weight (*10 points each*) and should all be answered; further detail about the points is given for specific parts of each question. The duration of the exam is 3 hours.

Generally, all statistical models used should be specified, and to such detail that it is clear which terms are present and in which form. Your answers should generally (unless specified otherwise) be based on the information provided: the study descriptions, data listings and computer listing. Nevertheless, if at some point you think it is necessary to carry out additional analysis in statistical software, explain carefully the purpose of your proposed analysis and how you would implement it in Minitab or Stata.

Question 1.

Most universities teach many sections of introductory calculus, and faculty are constantly looking for a method to evaluate students consistently across sections. An exam service claims that it can supply different exams that consistently evaluate students. Some faculty are sceptical of this claim, in part because they believe that the text(book) used may affect how specific exams work. Three math departments at three different universities (A, B, C) propose the following experiment. Three random final exams are obtained from the service (E1, E2, E3). At each of the universities, the three exams will be used in random order in the fall, winter, and spring quarters. The three universities all use the same two texts (T1, T2). Sections of the courses within each quarter and university will be divided at random into two groups, with half of the sections using text T1 and the other half using text T2. At the end of the year, the mean test scores were tallied with the following results.

Mean test score	Exam E1		Exam E2		Exam E3	
University	Text T1	Text T2	Text T1	Text T2	Text T1	Text T2
A	81	87	79	85	70	78
B	84	82	81	81	83	84
C	87	98	82	93	86	90

Use the description of the study and the information contained in the subsequent Minitab and Stata listings (for Question 1) to answer to the following questions.

A) (*5 points*) Describe the experimental design in statistical terms, and formulate a statistical model that properly reflects the experimental design. Explain the meaning of each term in your model. (*Hint:* It may be helpful to display the experimental design or the data structure by a sketch/diagram.)

B) (*3 points*) Use the attached output from statistical software to assess, by means of estimates or statistical tests, the different effects of your model. Draw conclusions about the questions of interest that motivated the study (as described above).

C) (*2 points*) In continuation of the discussion in B), use the information provided to draw specific conclusions about how the specific exams, texts, and universities affected the exam scores.

Minitab listing for Question 1:
note: the Stata listing contains additional information

```
MTB > GLM 'score' = exam|text univ;
SUBC>   Brief 2 ;
SUBC>   EMS.
General Linear Model: score versus exam, text, univ

Factor  Type   Levels  Values
exam    fixed       3  E1, E2, E3
text    fixed       2  T1, T2
univ    fixed       3  A, B, C

Analysis of Variance for score, using Adjusted SS for Tests

Source     DF  Seq SS  Adj SS  Adj MS     F      P
exam        2   67.11   67.11   33.56  2.00  0.186
text        1  112.50  112.50  112.50  6.70  0.027
exam*text   2    1.33    1.33    0.67  0.04  0.961
univ        2  280.11  280.11  140.06  8.34  0.007
Error      10  167.89  167.89   16.79
Total      17  628.94

S = 4.09742   R-Sq = 73.31%   R-Sq(adj) = 54.62%

Variance Components, using Adjusted SS

        Estimated
Source      Value
Error       16.79

MTB > GLM 'score' = exam|text univ|exam;
SUBC>   Random 'univ';
SUBC>   Brief 2 ;
SUBC>   EMS.
General Linear Model: score versus exam, text, univ

Factor  Type    Levels  Values
exam    fixed        3  E1, E2, E3
text    fixed        2  T1, T2
univ    random       3  A, B, C

Analysis of Variance for score, using Adjusted SS for Tests

Source     DF  Seq SS  Adj SS  Adj MS     F      P
exam        2   67.11   67.11   33.56  1.63  0.303
text        1  112.50  112.50  112.50  7.88  0.031
exam*text   2    1.33    1.33    0.67  0.05  0.955
univ        2  280.11  280.11  140.06  6.81  0.051
exam*univ   4   82.22   82.22   20.56  1.44  0.328
Error       6   85.67   85.67   14.28
Total      17  628.94

S = 3.77859   R-Sq = 86.38%   R-Sq(adj) = 61.41%
```

```
Variance Components, using Adjusted SS

           Estimated
Source         Value
univ          19.917
exam*univ      3.139
Error         14.278

MTB > GLM 'score' = exam|text univ|text;
SUBC>   Random 'univ';
SUBC>   Brief 2 ;
SUBC>   EMS.

General Linear Model: score versus exam, text, univ

Factor  Type    Levels  Values
exam    fixed        3  E1, E2, E3
text    fixed        2  T1, T2
univ    random       3  A, B, C

Analysis of Variance for score, using Adjusted SS for Tests

Source     DF  Seq SS  Adj SS  Adj MS     F      P
exam        2   67.11   67.11   33.56  2.66  0.130
text        1  112.50  112.50  112.50  3.36  0.208
exam*text   2    1.33    1.33    0.67  0.05  0.949
univ        2  280.11  280.11  140.06  4.18  0.193
text*univ   2   67.00   67.00   33.50  2.66  0.130
Error       8  100.89  100.89   12.61
Total      17  628.94

S = 3.55121   R-Sq = 83.96%   R-Sq(adj) = 65.91%

Variance Components, using Adjusted SS

           Estimated
Source         Value
univ          17.759
text*univ      6.963
Error         12.611
```

Stata listing for Question 1:
note: the Minitab listing contains additional information

```
. table text exam, contents(mean score ) row col
----------------------------------------------
          |                exam
     text |      E1       E2       E3    Total
----------+-----------------------------------
       T1 |      84  80.6667  79.6667  81.4444
       T2 |      89  86.3333       84  86.4444

    Total |    86.5     83.5  81.8333  83.9444
----------------------------------------------

. table univ exam, contents(mean score ) col
----------------------------------------------
          |                exam
     univ |      E1       E2       E3    Total
----------+-----------------------------------
        A |      84       82       74       80
        B |      83       81     83.5     82.5
        C |    92.5     87.5       88  89.3333
----------------------------------------------

. table text univ, contents(mean score )
------------------------------------
          |           univ
     text |       A        B        C
----------+-------------------------
       T1 | 76.6667  82.6667       85
       T2 | 83.3333  82.3333  93.6667
------------------------------------

. encode univ, gen(Univ)
. encode exam, gen(Exam)
. encode text, gen(Text)

. anova score Exam Text Exam*Text Univ

                          Number of obs =      18     R-squared     =  0.7331
                          Root MSE      = 4.09742     Adj R-squared =  0.5462

                  Source |  Partial SS    df       MS           F     Prob > F
              -----------+----------------------------------------------------
                   Model |  461.055556     7  65.8650794       3.92     0.0255
                         |
                    Exam |  67.1111111     2  33.5555556       2.00     0.1861
                    Text |       112.5     1       112.5       6.70     0.0270
                Exam*Text |  1.33333333     2  .666666667       0.04     0.9612
                    Univ |  280.111111     2  140.055556       8.34     0.0074
                         |
                Residual |  167.888889    10  16.7888889
              -----------+----------------------------------------------------
                   Total |  628.944444    17   36.996732
```

```
. anova score Exam Univ / Exam*Univ Text Exam*Text

                           Number of obs =      18     R-squared     =  0.8638
                           Root MSE      = 3.77859     Adj R-squared =  0.6141

                  Source |  Partial SS    df       MS           F     Prob > F
              -----------+----------------------------------------------------
                   Model |  543.277778    11  49.3888889       3.46     0.0700
                         |
                    Exam |  67.1111111     2  33.5555556       1.63     0.3032
                    Univ |  280.111111     2  140.055556       6.81     0.0515
                Exam*Univ |  82.2222222     4  20.5555556
              -----------+----------------------------------------------------
                    Text |       112.5     1       112.5       7.88     0.0309
                Exam*Text |  1.33333333     2  .666666667       0.05     0.9547
                         |
                Residual |  85.6666667     6  14.2777778
              -----------+----------------------------------------------------
                   Total |  628.944444    17   36.996732


. anova score Text Univ / Text*Univ Exam Exam*Text

                           Number of obs =      18     R-squared     =  0.8396
                           Root MSE      = 3.55121     Adj R-squared =  0.6591

                  Source |  Partial SS    df       MS           F     Prob > F
              -----------+----------------------------------------------------
                   Model |  528.055556     9  58.6728395       4.65     0.0207
                         |
                    Text |       112.5     1       112.5       3.36     0.2083
                    Univ |  280.111111     2  140.055556       4.18     0.1930
                Text*Univ |          67     2        33.5
              -----------+----------------------------------------------------
                    Exam |  67.1111111     2  33.5555556       2.66     0.1301
                Exam*Text |  1.33333333     2  .666666667       0.05     0.9488
                         |
                Residual |  100.888889     8  12.6111111
              -----------+----------------------------------------------------
                   Total |  628.944444    17   36.996732
```

Questions 2-3.

Not shown.