

**Final examination, 22. April 2003**

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.

**Questions 1+3.**

Not shown.

**Question 2.**

Many professions have board certification exams. Part of the certification process for *bank examiners* involves a set of tasks that the examinee must complete in a satisfactory fashion in a fixed time period. New task-sets must be developed for each round of examinations, and an experiment was carried out to evaluate two such new task-sets (A and B). The experiment included also 3 old (and approved) task-sets (C, D and E), and all 5 task-sets were completed by 10 paid examinees (so this was not a real exam!). The paid examinees are labeled here 1–10; number 1–6 were certified bank examiners, 7–9 were noncertified bank examiners nearing the end of their training period, and 10 was a public accountant with no bank examining experience or training. The experiment also involved 5 graders (labeled 1–5), who each graded 10 exams. The primary interest is in the exams, which were randomized so that each examinee took each exam, and each grader graded two of each exam. The data are given in the table below; scores are on a scale from 0 to 100, and letters indicate exams (as explained above).

Score and Exam	Grader				
Examinee	1	2	3	4	5
1	68 D	65 A	76 E	74 C	76 B
2	68 A	77 E	84 B	65 D	75 C
3	73 C	85 B	72 D	68 E	62 A
4	74 E	76 C	57 A	79 B	64 D
5	80 B	71 D	76 C	59 A	68 E
6	69 D	75 E	81 B	68 A	68 C
7	60 C	62 D	62 E	66 B	40 A
8	70 B	55 A	62 C	57 E	40 D
9	61 E	67 C	53 A	63 D	69 B
10	37 A	53 B	31 D	48 C	33 E

Again, the data has been analyzed in Minitab, and some listings are given below as an aid for the following questions.

- A) (*5 points*) Describe the statistical design of the experiment, and motivate (as well as the problem description allows) the choice of design for the experiment. Give a statistical model for analysis of the data, and explain the meaning/importance/relevance of the different terms in the model.

B) (5 points) Carry out the statistical analysis, as far as possible from the Minitab listings (you may suggest additional analyses to be carried out), and explain and interpret the significance of the different factors of the experiment. In particular, answer to the primary question: were either or both of the new exams equivalent to the old exams?

```
MTB > GLM 'score' = exam examinee grader;
SUBC> Brief 2 ;
SUBC> Means exam examinee grader.
```

General Linear Model: score versus exam, examinee, grader

```
Factor      Type Levels Values
exam        fixed      5 A B C D E
examinee    fixed     10 1 2 3 4 5 6 7 8 9 10
grader      fixed      5 1 2 3 4 5
```

Analysis of Variance for score, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
exam	4	1889.92	1889.92	472.48	33.96	0.000
examinee	9	5050.32	5050.32	561.15	40.34	0.000
grader	4	443.32	443.32	110.83	7.97	0.000
Error	32	445.16	445.16	13.91		
Total	49	7828.72				

Least Squares Means for score

	Mean	SE Mean
exam		
A	56.40	1.179
B	74.30	1.179
C	67.90	1.179
D	60.50	1.179
E	65.10	1.179
examinee		
1	71.80	1.668
2	73.80	1.668
3	72.00	1.668
4	70.00	1.668
5	70.80	1.668
6	72.20	1.668
7	58.00	1.668
8	56.80	1.668
9	62.60	1.668
10	40.40	1.668
grader		
1	66.00	1.179
2	68.60	1.179
3	65.40	1.179
4	64.70	1.179
5	59.50	1.179