TENTAMEN: Experimental design (TMS031/MSA250)
Tisdagen den 6 mars, 2012, kl 8:30-12:30
Lärare och jour: Aila Särkkä, tel. 772 35 42
Hjälpmedel: A valfri miniräknare (med tömt minne) och tabellbok (minicalculator (with emptied memory) and a book of tables).

1) In a soda factory a new orange drink is going to be designed. There is a choice of oranges from two different countries, two levels of sugar, two different water qualities, two levels of added lemon flavour, and two different pH values. The response variable is the concentration of ascorbic acid. Eight runs will be carried out.

   a) Which runs would you suggest? Explain.

   b) What is the resolution of your experimental plan? Explain.

   c) Give the confounding pattern for the main effects. (10p)

2) a) Give an example of a latin square design. When is it useful?

   b) Give an example of a split-plot design? When should it be used?

   c) Give an example of a randomization (permutation) test when two groups are compared. Explain how you would perform the randomization and how you would compute the p-value.

   (8p)

3) Suppose you wish to estimate the difference between two mean acidity (pH) for rainfalls at two different locations, one in a relatively unpolluted area along the ocean and the other in an area subject to heavy air pollution. If you wish your estimate to be correct to the nearest 0.1 pH with probability 90%, approximately how many rainfalls (pH values) would have to be included in each sample? Assume that the variance of the pH measurements is 0.16 at both locations and that the samples are of equal size. (8p)

4) A local school board is interested in comparing test scores on a standardized reading test for fourth grade students in their district. They selected a random sample of five male and five female fourth grade students at each of four different elementary schools in the district and recorded the test scores. The results are shown in the table below

Gender	School 1	School 2	School 3	School 4
Male	631	642	651	350
	566	710	611	565
	620	649	755	543
	542	596	693	509
	560	660	620	494
Female	669	722	709	505
	644	769	545	498
	600	723	657	474
	610	649	722	470
	559	766	711	463

a) What type of experimental design is this? What are the experimental units? What are the factors and levels of interest to the school board?

b) An incomplete ANOVA table:

Source	df	SS	MS	F value
Gender	?	6200	?	?
School	?	246726	?	?
Gender:School	?	10575	?	?
Residuals	?	94826	?	
Total	?	?		

Fill in the ?'s in the ANOVA table and explain every item.

c) Do the data indicate that the effect of gender on the average test score is different depending on the student's school? Test the appropriate hypothesis using $\alpha = 0.05$.

d) Which of the effects are significant (if any)?

e) What are the assumptions you have to make in order to draw the conclusions in c) and d)?

(12p)

5) We would like to investigate how the response variable (percentage of yield) Y depends on the temperature X. The data are as follows:

		Value	of X	
30	40	50	60	70
13.7	15.5	18.5	17.7	15.0
14.0	16.0	20.0	18.1	15.6
14.6	17.9	21.1	18.5	16.5

We have fitted a first order (plane) model to the data by using the least squares method. We have then done a lack-of-fit analysis and obtained the following ANOVA table:

Source	df	SS	MS	F-ratio
Model	?	4235		
Residual	?	59	?	
Lack of fit	?	52	?	?
Pure error	?	7	?	
Total	?	?		

a) Give the model that has been fitted.

b) Fill in the ?'s in the ANOVA Table. Explain all the items in the table, especially what we mean by Lack of fit and Pure error.

c) Is there some evidence of lack of fit? What would be the next step in your analysis? Why?

d) Give an estimate for the error variance.

e) Give the assumptions you need for the estimation of the parameters and for the lack of fit analysis.

(12p)

Good luck!