Examination

Experimental design (MSA250/TMS031) Wednesday, March 14th, 2018, 8:30 - 12:30
Lecturers on call: Aila Särkkä and Torbjörn Lundh, tel 772 5355 and tel 772 3503.
Tools: A Chalmers accepted pocket calculator with emptied memory. At the examination, sheets with statistical distributions and tables will be handed out.
Maximum number of points: 30p
Limits GU: G (15p) and VG (22p)
Limits Chalmers: 3:a (15p), 4:a (20p) and 5:a (25p)
Give explanations to the notation you use and motivation to your conclusions.

1. (2+2+2 p)
   (a) Explain the term "Lack of Fit" and exemplify.
   (b) Describe the term "Design Information Function".
   (c) When do you want to use a higher order model, and how can you find out that you should do that?

2. (5 p) Suppose you would like to repeat Fisher's famous tea experiment, but with the twist of improving the classical design of four cups with milk added to the tea infusion and four cups with tea infusion added to the milk. You consider the following three variants. Analyze them especially with respect to what would be the outcomes where you would reject the null-hypothesis that the tea-taster does not have the ability to tell the difference between the two kinds of tea being served. Which of the test would you suggest in the end?
   (a) You will still serve only eight cups but without the restraint that it should be four of each kind, but instead at probability $\frac{1}{2}$ pick a cup where the milk is added afterwards.
   (b) You serve 12 cups, 6 of each kind (a fact that is revealed to the taster).
   (c) You repeat the classical design ten times.

3. (5 p) You have been given the assignment from a food magazine "Mat är allt" to give a statistical interpretation of a randomized paired design test, at five different malls in Sweden, of which soda that is most "tasty" of Coca Shmola (A) or Preppsi (B). Fourty mall visitors were randomly choosen in each mall and were asked to blindly decide which glass was more tasty. What will be your conclusion and argument?

A	B	B	A	B	A	A	B	B	A
15	25	29	11	21	19	14	26	18	22

4. (5 p) Let us have a completely randomized 2^4 factorial design experiment with four factors A, B, C, and D. The treatment combinations (runs) are divided into four blocks by using ABC and ABD as blocking factors.
   (a) Why is blocking used in factorial designs?
   (b) Give the runs that belong to each of the four blocks?

(c) Which additional effects (if any) are confounded with the block effect?

(d) Is this a good choice of dividing the runs into blocks? Why/why not?

(e) Would it have been better to choose ABCD and one of the three factor interactions or two of the two factor interactions as blocking factors? Why/why not?

5. (4 p) Design an eight-run fractional factorial design for an experimenter with the following five factors: temperature, concentration, pH, agitation rate and catalyst type. She tells you that she is particularly concerned about the two-factor interactions between the temperature and concentration and between catalyst type and temperature. She would like to have a design, if it is possible to construct one, with main effects unconfounded with one another. Make the design for her, tell which desing it is, and explain how you have taken her requests into account.

6. (5 p) Yield of four corn hybrids (A, B, C and D) are compared. 16 corn plants are planted on an area which is divided into 16 squares (4 columns and 4 rows), one plant for each square. The researcher suspects that the area is not homogeneous and would like to account for the variation due to the column and row effects.

   a) Which design would you suggest and why? Draw the design you have chosen.

   b) Give the linear model connected to your design and the hypotheses you want to test.

   c) Fill in the incomplete ANOVA table below and explain all the items in it when it is complete.

Source	SS	df	MS	F ratio
Row	0.030			
Column	0.827			
Treatment	0.427			
Residuals	0.129			
Total	1.414			

   Does the ANOVA table correspond to the design you have suggested in a)?

   d) What are your conclusions based on the ANOVA table and the F test. Give the assumptions that are needed to make these conclusions.

   e) Which design would you have if you ignored the column and row effects? How would you do the experimental set-up in this case?

Good luck!