

TENTAMEN: Experimental design (TMS031/MSA250)

Tisdagen den 15 mars, 2011, kl 8:30-13:30

Lärare och jour: Aila Särkkä, tel. 772 35 42

Hjälpmedel: A valfri miniräknare och tabellbok (minicalculator and a book of tables).

- 1) In which situations (and why) should one use
 - a) paired T test? Give also assumptions needed.
 - b) split-plot design?
 - c) randomization? (8p)

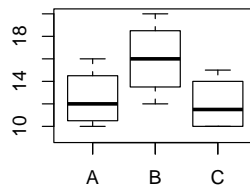
- 2) In a 2^{4-1} factorial design the influence of four catalysts A, B, C and D on the temperature of the process was investigated. The following result was obtained (notation -1 means that the catalyst in question is not present):

A	B	C	D	Temperature increase
-1	+1	+1	+1	12
-1	-1	+1	-1	16
-1	-1	-1	+1	11
+1	-1	-1	-1	18
+1	-1	+1	+1	10
+1	+1	+1	-1	20
-1	+1	-1	-1	15
+1	+1	-1	+1	11

- a) Write down the defining relation of the experiment.
- b) Present the confounding pattern.
- c) What is the resolution of the design?
- d) Estimate the main effect of the catalyst A. What assumptions are needed in order to guarantee that the estimate of the main effect is unbiased? (8p)

- 3) An experiment was conducted to determine the effects of three methods of soil preparation on the first-year growth of pine seedlings. Four locations were selected, and each location was divided into three plots. The methods of soil preparation were A (no preparation), B (light fertilization), and C (burning). Each soil preparation was randomly applied to a plot within each location. On each plot, the same number of seedlings were planted and the average first-year growth of the seedlings was recorded on each plot. The data and a boxplot of the methods are given below.

Method	Location			
	1	2	3	4
A	11	13	16	10
B	15	17	20	12
C	10	15	13	10



- a) An incomplete ANOVA table:

Source	df	SS	MS	F value	<i>p</i> value
Soil preparation		38			
Residual		73			
Total					

Fill in the ANOVA table (it is enough to say whether the *p* value is less or larger than 0.05). Explain all the numbers in it (both those you add and those that are already there).

- b) Based on the ANOVA table, do the data provide evidence to indicate a difference in the mean growths for the three soil preparations? Give the assumptions you have made.
- c) How would you improve the analysis? Explain.
- d) Use a 95% confidence interval to estimate the difference in mean growth for methods A and B. Interpret the result. (14p)

- 4) It is believed that radiation from cell phones possibly increases the cancer rate. Assume that the cancer frequency (for some kind of cancer) is 1.0% (with variance 0.25) for people not using cell phones. Suppose that the cancer rate is increased to 1.2% for regular cell phone users. We want that the increased cancer frequency would be detected with probability 99% on the 5% level of significance. How many people would be needed in the study? What assumptions have you made? (8p)
- 5) A chemical engineer is studying the yield of a process. There are two variables, temperature A (low= 80°C, high= 100°C) and pressure B (low= 75psi, high= 120psi). The experimental layout and yield data are:

Point	A	B	y
1	-1	-1	63.4
2	+1	-1	60.3
3	-1	+1	64.2
4	+1	+1	63.1
5	0	0	68.1
5	0	0	68.7
5	0	0	69.0
5	0	0	68.6
5	0	0	69.1

- a) How do you have to scale the variables to obtain the values given in the design matrix?
- b) Give and fit a plane (1st order) model to the data.
- c) Does the plane model fit well? Why/why not?
- d) What would be the next step in order to find the maximum yield value?
- e) Give the assumptions you have made? (12p)

Good luck!