Mathematical Statistics
Centre for Mathematical Sciences
Lund University

Exam: 2019–06–03 kl 08^{00}–13^{00}
FMS F65 / MAS C05
Design of Experiments

Correct, adequately motivated, and explained answers to problems 1 to 3 give 10 points each. Problems 4 to 6 give 20 points each. Total maximal score is 90 points. The threshold for passing the exam is 40 points. Use the departments paper both for drafting and final answer. The answer to each new problem shall start at the top of a new page. Red pen or pencil is not allowed. Write full name or id on all paper sheets. Approved aids: statistical tables including Mathematical Statistics support sheets, and calculator.

---

1. Describe in detail, e.g. by giving a small example, what is meant by *Latin Square Design* and *Graeco-Latin Square Design*. (10p)

2. a) When checking that a certain assumption is fulfilled in ANOVA-models we may use a qq-plot. In this context, which model assumption is investigated by using the qq-plot?

   b) Assume that we want to investigate how well a data set $x_1, x_2, \ldots, x_n$ fits a theoretical distribution (with distribution function F). To this end, we may plot the following points: $(x, y) = (x_{(i)}, F^{-1}(i/n))$ for $i = 1, \ldots, n$. Here, $x_{(i)}$, $i = 1, \ldots, n$, is the sorted data. Explain in detail the logic behind such a plot by drawing the empirical distribution function and the theoretical distribution function in the same picture. If the fit is good, what would the plot look like?

   (10p)

3. A car company is planning to test two types of so called "coolant" for electrical car engines. Engine coolant is a water based liquid that is circulated around the battery packs to lower the temperature. In more detail, the company wants to test two types of coolant (A - the old one, and B - an improved version) on ten different engines. To this end, the company randomized the allocation of coolant-type to each of the ten engines and got the result (for example, B (y_i) in the table below means that coolant-type B was applied to engine i, and the observed temperature was y_i):

B (y_1)	A (y_2)	B (y_3)	A (y_4)	B (y_5)
A (y_6)	A (y_7)	A (y_8)	B (y_9)	B (y_{10})

   a) Given the result of the experiment above, describe in detail how the company could apply a *distribution free test* (or a so called randomization test) to see if there is a significant difference in engine-temperature between the two coolant types A and B.

   b) Describe an obvious practical downside of the *distribution free test* in a).

   (10p)

4. The following are the lifetimes of a certain sub-component in two different smartphones (Type I and Type II). The hardware-team at the company is interested in both the means and variances of the lifetime.

Type I		**Type II**	
65	82	64	56
81	67	71	69
57	59	83	74
66	75	59	82
82	70	65	79

   a) Test the hypotheses that the two variances (for Type I and Type II, respectively) are equal. Use significance level $\alpha = 0.05$.

   b) Using the results of a), test the hypotheses that the mean lifetimes are equal. Use significance level $\alpha = 0.05$

   c) Assume that $X \in N(0,1)$, and $Y \in \chi^2(f)$. Then, $T = \frac{X}{\sqrt{Y/f}} \in t(f)$. Given this information, derive the distribution of T^2.

(20p)

5. In the following table the design and the results of a two times replicated experiment are presented. The sample mean and the sample variance are given for each experiment. The 16 runs were done in random order. From process knowledge we know that the third and higher order interactions between factors A, B, C and D can be neglected.

A	B	C	D	y_1	y_2	$\bar{y}$	s^2
$-$	$-$	$-$	$-$	8.34	7.47	7.91	0.3785
$+$	$-$	$-$	$+$	5.73	5.62	5.68	0.0061
$-$	$+$	$-$	$+$	-0.92	1.04	0.06	1.9208
$+$	$+$	$-$	$-$	13.4	13.6	13.5	0.0200
$-$	$-$	$+$	$+$	-1.75	-2.46	-2.11	0.2521
$+$	$-$	$+$	$-$	10.43	9.91	10.17	0.1352
$-$	$+$	$+$	$-$	4.71	3.56	4.14	0.6613
$+$	$+$	$+$	$+$	3.01	3.65	3.33	0.2048

   a) What are the confounding relations and what is the resolution of the design?

   b) Is it possible to estimate the main effects alone?

   c) Which factors are active (i.e., significant) at a 3-sigma-level?

(20p)

6. Three diet methods (α, β and γ) for weight reduction of overweight men are compared. Fifteen men (all equally motivated to loose weight) were included in the study. The trials were carried out by first dividing the subjects into "matched" groups. That is, men in any group were chosen to be as alike as possible. The loss of weight (in pounds) after 3 months for the three diets was as follows:

	Diets		
Matched groups	α	β	γ
1	15	10	8
2	24	15	17
3	31	28	34
4	37	36	34
5	33	37	39

a) What type of experimental design is this?

b) Fill in the missing pieces (the question marks) of the ANOVA-table (it is enough to say whether the p-value is less or greater than 0.05):

	SS	df	MS	F-ratio	p-value
Diet	19.73	?	?	?	?
Group	1507.73	?	?	?	?
Residuals	92.27	?	?		
Total	1619.73	?			

c) Interpret all the results in the ANOVA-table. State also all assumptions made when interpreting the results.

d) Suppose that you are told that the average weight (in pounds) at the beginning of the trial for members in each group is as follows:

Group	1	2	3	4	5
Weight	250	309	327	356	379

Would you analyze the results differently given this information? In such a case, how?

(20p)

THE END!