Exercises for Session 7: Multifactorial and block designs

Exercise 7.1
Two-way analysis of variance
An experiment with rats feeding on fat (from pigs, i.e. lard) included a study of the difference in feed consumption between fresh and rancid fat. The consumption during 73 days for 12 rats (6 males and 6 females) of an age between 30 and 34 days are listed in the table below, measured in *g*. (Data from Powick, W. C. (1925): Inactivation of vitamin A by rancid fat, *J. Agric. Res.* **31**, 1017-1027.)

Sex	Fat					
	fresh			rancid		
male	709	679	699	592	538	476
female	657	594	677	508	505	539

Describe the type of experiment, formulate a corresponding statistical model, and analyze the data. Draw conclusions and think about how to present the results.

Exercise 7.2
Latin square
The effect of insulin on the blood concentration of glucose was studied on rabbits. Three rabbits each received insulin doses A, B, and C (corresponding respectively to 0, 1, and 2 units) at different days. The experimental design is given below with the glucose measurements (*mg* pr. 100 *ml* blod) taken 50 minutes after injection. (Data from Young & Romans (1948): Assay of insulin with one blood sample per rabbit per day, *Biometrics* **4**, 122-131.)

Day	Rabbit					
	1		2		3	
1	A	50	C	39	B	36
2	C	37	B	51	A	53
3	B	51	A	60	C	37

Give the type of experimental design, formulate a statistical model and analyze the data. Examine whether the effect of insulin can be described as a linear function of the dose. If so, determine a 95% confidence interval for the regression coefficient.

Exercise 7.3
Multi-way ANOVA
An experiment compared three thiosemicarbazon-type preparations against virus:

I: p-amino
II: substituent p-methoxy
III: unsubstituted.

For each preparation six eggs were infected with virus solutions in three concentrations. The entire experiment was repeated once. If $x_1, \dots x_6$ denotes the survival times (hours) of the virus in the six eggs, we let

$$y = \frac{10^4}{6}\Big(\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \frac{1}{x_4} + \frac{1}{x_5} + \frac{1}{x_6}\Big).$$

The table below gives y-values for every preparation in its three concentrations at the two replications of the experiment. (Data from Hamre *et al.* (1951): Studies on the chemotherapy of vaccinia virus II: The activity of some thiosemicarbazones, *J. Immunology* **67**, 305–312.)

y	Replication 1			Replication 2		
	Concentration			Concentration		
Preparation	$10^{-4.0}$	$10^{-4.3}$	$10^{-4.6}$	$10^{-4.0}$	$10^{-4.3}$	$10^{-4.6}$
I	87	79	77	90	80	81
II	82	73	72	71	72	68
III	72	70	62	77	66	61

Describe the experimental design, formulate a statistical model, and analyze the data.

Exercise 7.4
Multi-way ANOVA
Consider time to intoxication by cyanide in *Phoxinus laevis*, a European minnow, using 5 concentrations of CN^- ions, 3 oxygen concentrations, and 3 temperatures, as follows:

temperature (°C):	5, 15, 25,
oxygen concentration (mg O_2/l):	1.5, 3.0, 9.0,
cyanide concentration (mg CN^-/l):	0.16, 0.8, 4.0, 20.0, 100.0.

The dependent variable is a transformation into logarithms of readings in minutes of survival time, coded and averaged for 10 replicate fishes. (From Sokal & Rohlf (1995): *Biometry*; data from Wuhrmann, K. & Woker, H. (1953): Über die Giftwirkungen von Ammoniak und Zyanidlösungen mit verschiedener Sauerstoffspannung und Temperatur auf Fische, *Schweiz. Z. Hydrol.* **15**, 235-260.)

		Cyanide concentration				
Temperature	Oxygen concentration	0.16	0.8	4.0	20.0	100
5	1.5	20.1	15.0	13.1	13.0	9.7
	3.0	24.6	16.4	13.8	13.6	10.2
	9.0	27.1	17.0	14.9	12.7	9.9
15	1.5	12.4	10.4	8.6	8.9	6.0
	3.0	15.8	11.1	9.9	9.1	7.4
	9.0	20.7	11.7	8.1	8.7	7.2
25	1.5	7.9	6.3	5.0	5.1	3.2
	3.0	12.9	5.4	5.1	5.2	4.6
	9.0	14.2	9.3	6.2	5.1	5.2

State the type of experimental design and suggest statistical models for the data. Note, that it may be of interest to model the data in terms of the quantitative factor levels. Analyze the data and draw conclusions.

Exercise 7.5
Planning of experiments
This exercise is based on the Checklist for Planning of Experiments from Dean & Voss: *Design and Analysis of Experiments*.

Part I
Read critically through the discussion of the Battery experiment in Section 2.5.2 of Dean & Voss.

Pay particular attention to the identification of the sources of variation. Think about whether your planning of a similar experiment would have involved the same considerations and the same choices.

Part II
Go through steps (a)–(d) of the checklist for an experiment outlined as follows (by Clifford Pugh, *Applied Statistics*, 1953):

> The widespread use of detergents for domestic dish washing makes it desirable for manufacturers to carry out tests to evaluate the performance of their products. [...] Since foaming is regarded as the main criterion of performance, the measure adopted is the number of plates washed before the foam is reduced to a thin surface layer. The five main factors which may affect the number of plates washed by a given product are (i) the concentration of the detergent, (ii) the temperature of the water, (iii) the hardness of the water, (iv) the type of "soil' on the plates, and (v) the method of washing used by the operator. [...] The difficulty in standardizing the soil is overcome by using the plates from a works canteen (cafeteria) for the test and adopting a block technique in which plates from any one course forms a block. [...] One practical limitation is the number of plates available in any one block. This permits only four [...] tests to be completed (in a block).

Exercise 7.6
Unbalanced (incomplete) two-way ANOVA
The following experiment was carried out by Dr. H. Wolffhechel, KVL (Danish Veterinary and Agricultural University), in 1986. The purpose was to compare 12 sphagnum moss lots with respect to water and air content. Each of these was applied to four pots with small cucumber plants. The pots were placed in one of six watering troughs, each containing eight pots. The experimental design and the volume (water and air content), in percent, for each pot is given in the table. (From Skovgaard (1994): *Statistisk Forsøgsplanlægning*, in Danish).

Volume (percent)	Watering trough					
Sphagnum moss lot	1	2	3	4	5	6
1	37.0		44.6		42.5	47.1
2			49.0	50.5	51.0	44.8
3		34.6	42.7	41.8	37.8	
4	45.3	42.7	47.7	42.8		
5	32.1	38.5		32.0		31.6
6	34.3	33.3		34.0	22.6	
7	32.3		28.1	28.1	32.3	
8	38.9	36.5		39.7		34.8
9	33.9		31.4	32.1		23.0
10		39.7	41.8		43.5	33.8
11		41.1	38.1		31.1	37.9
12	35.9	7.5			36.2	25.5

Describe the experimental design, and explain why it is not a balanced incomplete block design. Analyse the data using Minitab/Stata; hereby, determine the error sum of squares, SSE, and degrees of freedom, DFE, for the following 4 models,

$$\text{(A)} \quad y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij},$$

$$
\begin{aligned}
\text{(B)} \quad y_{ij} &= \mu + \beta_j + \varepsilon_{ij}, \\
\text{(C)} \quad y_{ij} &= \mu + \alpha_i + \varepsilon_{ij}, \\
\text{(D)} \quad y_{ij} &= \mu + \varepsilon_{ij},
\end{aligned}
$$

where indices i and j correspond to sphagnum moss lots and watering troughs, respectively. Test the reductions of model (A). Make sure you understand the difference between partial (adjusted) and sequential sum of squares, and state which of them you would use to assess lot and trough effects.

In model (A), compare estimation of sphagnum moss lot levels by ordinary means and least squares means. Why do they differ — and which one would you use to summarise the treatment effects? Describe the relationship between the sphagnum moss lots, if necessary using statistical comparison techniques.

Exercise 7.7
ANOVA and transformation
A dataset analysed in a seminal statistical paper by Box & Cox consisted of survival times of 48 animals that were given one of 3 different poisons and prior to the poisoning had been subjected to one of 4 (unspecified) treatments. The data are shown in the table below.

Survival in hours	Treatment			
Poison	A	B	C	D
I	3.1	8.2	4.3	4.5
	4.5	11.0	4.5	7.1
	4.6	8.8	6.3	6.6
	4.3	7.2	7.6	6.2
II	3.6	5.8	4.4	5.6
	2.9	6.1	3.5	10.2
	4.0	4.9	3.1	7.1
	2.3	12.4	4.0	3.8
III	2.2	3.0	2.3	3.0
	2.1	3.7	2.5	3.6
	1.8	3.8	2.4	3.1
	2.3	2.9	2.2	3.3

Analyse the data, while paying particular attention to the points below.

1) Discuss whether it seems appropriate to model the survival times on a transformed scale, and in that case which transformation to use. Carry out all subsequent analyses at the chosen scale.

2) Discuss whether there seems to be outlier(s) present in the data, and substantiate your discussion by statistical test(s).

3) Draw conclusions about the effects of treatments and poisons. For the quantitative comparison of poisons, assume poison III to be of one particular — and new — type, and poisons I and II to be similar and of another type, so that it is of interest to compare the two types of poison. Among the 4 treatments, no special relations exist.

4) Based on your final model (i.e., including the terms you consider relevant), estimate with a 95% confidence interval the expected survival time of animals poisoned with poison III. Same question for the survival time of animals given treatment A and poison III.