Written Exam
(2 hours)

General remarks:

- Open book exam.
- Switch off your mobile phone!
- Do not stay too long on a part where you experience a lot of difficulties.
- Use only the separate sheet for your answers.
- If not stated otherwise, all tests have to be done at the 5%-level.
- The tables of critical values for the F and t distributions are attached at the end.
- Exercises 4 and 5 are multiple-choice exercises. In each sub-exercise, exactly one answer is correct. A correct answer adds 1 *plus*-point and a wrong answer $\frac{1}{2}$ *minus*-point. You get a minimum of 0 points for each multiple-choice exercise. Tick the correct answer to the multiple choice exercises in the separately added answer sheet.

Good Luck!

1. (12 Points)

A trendy wine bar set up an experiment to evaluate the quality of 3 different wines. Five fine connoisseurs of wine were asked to taste each of the wine and give it a rating between 0 and 10. The order of tasting was randomized and the judges did not know which wine they were drinking.

The following table displays the collected data:

	wine 1	wine 2	wine 3
person 1	1	7	5
person 2	0	4	0
person 3	1	6	4
person 4	1	5	2
person 5	1	8	10

We use the following simple model:

$$Y_{ij} = \mu + W_i + P_j + \epsilon_{ij},$$

with Y_{ij} the ratings, W_i and P_j the fixed effects for the wine sort and person ID, respectively, and the standard assumptions for the errors. We additionally add the constraint that $\sum W_i = \sum P_j = 0$.

a) Compute the estimates of $\hat{\mu}$ and all effects $\hat{W}_i$.

b) Given the data and the model, is there any significant difference of ratings between the different wines? Give the null hypothesis and justify your answer. Hint: the estimated standard error is $\hat{\sigma}_e = 1.79$.

c) Is the difference between the wines 1 (W_1) and 2 (W_2) significant? Hint: use pairwise mean difference without correction.

d) Should we include the effect of the person ID (P_j) in the model or not? Motivate your answer.

e) Compute $\sum_{j=1}^{5} r_{ij}$ for wine 1.

f) What is a potential problem with the design of this experiment? How would you test it?

2. (11 Points)

The last assessment report of the Intergovernmental Panel on Climate Change (IPCC) investigated the global mean temperature change (difference between a future and a reference period) based on 2 different climate models and 3 different emission scenarios. A balanced factorial design was used.

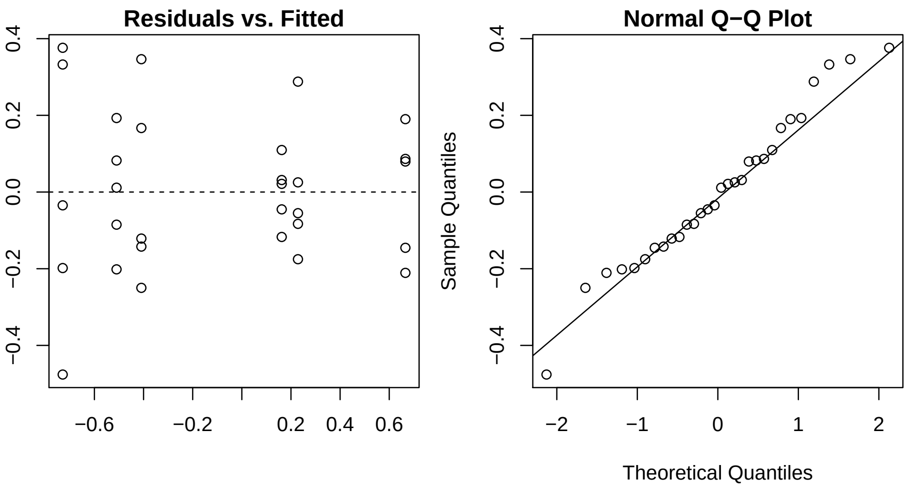
To analyze the data we use the following model:

$$Y_{ijk} = \mu + M_i + E_j + (ME)_{ij} + \epsilon_{ijk},$$

with Y_{ijk} the temperature change, M_i and E_j the effects of the climate model and the emission scenarios, respectively. The term $(ME)_{ij}$ denotes an interaction between M_i and E_j.

a) What are the standard assumptions about the errors?

b) Are the standard assumptions fulfilled? Motivate your answer with the help of the following plots.



c) Complete the following ANOVA table:

	df	SS	MS	F-value
climate model (M)	...	2.03	...	...
emission scenario (E)	...	4.86	...	...
model * scenario (MS)	...	0.22	...	...
Residuals	24	1.13	...	

d) How many replicates of every combination of emission scenarios and climate model were performed?

e) Generally, how does the critical F-value change if tested at a 2.5% instead of 5% significance level? Give a qualitative answer.

We fit the model with R assuming that $\sum M_i = \sum E_j = 0$ and $\sum_i (ME)_{ij} = \sum_j (ME)_{ij} = 0$ and obtain the following estimated effects:

(Intercept)	M1	E1	E2	M1:E1	M1:E2
-0.10	-0.26	-0.47	-0.04	0.10	-0.11

f) Given the fitted model, what is the predicted value for the following cases?

- model 1 and scenario 1 (M_1 and E_1)?

- model 2 and scenario 3 (M_2 and E_3)?

g) Let's say we fit our model and find a p-value of p for the effect of scenario on global warming.
Mark the following statements True or False.

	True	False
The probability that the predicted global temperature is not affected by the emission scenario is equal to p.	◯	◯
If the null hypothesis were true, the probability to observe an effect of size as estimated from the data would be p.	◯	◯
If the null hypothesis were true, the value p would be uniformly distributed between 0 and 1.	◯	◯
If p were bigger than 10% it would mean that the emission scenario is not important.	◯	◯

3. (9 Points)

Benjamin, a PhD student, wants to study the effect of different treatments on the resistance of bones. Three different treatments were administered to horses during their lifetime: T1, T2 or T3. When they died, their 4 femoral bones were exposed to some mechanical force. Benjamin records various properties of the bones and takes the total force needed to break the bone as dependent variable Y.

a) Looking at the data, Benjamin notices that the difference between T1 and T3 is the biggest and would like to test it. Describe how you would do it and explain why.

Benjamin decides to take into account that some bones are posterior and others anterior. He fits a new ANOVA model, with an interaction between the treatment and the bone position (posterior or anterior).

b) Explain in words what a significant interaction would mean.

c) In the model fitted above, can you assume that the errors are i.i.d. ? Motivate your answer.

Benjamin looks again at his data and realizes that the number of horses assigned to each treatment is not the same.

d) Does it have an influence on his ANOVA calculation? Would you recommend the use of the R-function `aov`? Motivate both answers.

Advised by his supervisor, Benjamin includes in the model the age of the horses as a control variable. To his surprise the effect of treatment is not significant at all anymore.

e) What do you think is the reason for this to have happened?

Benjamin wants to get more out of his experiment, so he tries to use other dependent variables that he had measured (number of micro-fractures, difference in pore density before and after, sheer effect, etc.). In total he fits 20 different ANOVA models and finds one which shows a significant effect of treatment and wants to use it for a publication.

f) What is the problem with choosing the dependent variable Y like this? What would you do about it?

4. **(7 Points)**

Part A: Fractional factorial design

A mountain-bike company would like to produce a new type of bike frame. They consider five production factors: A, B, C, D, E. Each of these is a factor with only two levels: "high" $(+)$ or "low" $(-)$. The best bike frame is the one that has the lowest damage score in a crash test. To find the best combination of the factors, the company decides to use a 2^{k-l} fractional factorial design experiment.

1) They initially proposed to use the following defining relations: $I = ABD = EC = ABCDE$. Which of the following statements is true?

    a) Some main effects are confounded with other main effects.

    b) No main effects are confounded with any other main effects, but some main effects may be aliased with two-factor interactions.

    c) No main effects are confounded with any other main effects and two-factor interactions. Some two-factor interactions may be aliased with other two-factor interactions.

    d) No main effects are confounded with any other main effects, two-factor interactions and three-factor interactions. No two-factor interactions are confounded with any other two-factor interactions.

After a consultation with the statistical team, different confounding relations than in 1) were chosen to construct a 2^{k-l} fractional factorial design experiment. The results of the experiment are summarized in the following table.

	Damage score
de	5
a	2
be	3
abd	6
cd	8
ace	4
bc	2
abcde	4

2) What are the confounding relations?

    a) $B = AC, E = AB$

    b) $D = AB, E = AC$

    c) $A = BC, E = ABC$

    d) $D = AB, E = ACD$

3) What is the resolution of the experiment?

    a) 2

    b) 3

    c) 4

    d) 5

4) The following incomplete table gives estimates of the most important effects.

$\hat{A}$	$\hat{B}$	$\hat{C}$	$\hat{D}$	$\hat{E}$	$\hat{BC}$
??	-0.5	0.25	1.5	??	-1

The estimates of the main effects A and E are

a) $\hat{A} = -0.5, \hat{E} = -0.25$
b) $\hat{A} = -0.25, \hat{E} = 1$
c) $\hat{A} = -0.25, \hat{E} = -0.25$
d) $\hat{A} = -0.5, \hat{E} = 1$

5) Further studies show that all other effects which are not given in the table in task 4) are irrelevant. Which of the following combinations then gives the best product? (Hint: You do not need to know $\hat{A}$ and $\hat{E}$ from task 4) to answer this question).

a) $(A, B, C, D, E) = (+, +, +, +, +)$
b) $(A, B, C, D, E) = (+, +, +, -, +)$
c) $(A, B, C, D, E) = (+, +, -, -, +)$
d) $(A, B, C, D, E) = (+, -, -, -, +)$

Part B: Contrasts and `R`-output

In a medical study, there are 6 different types of drugs numbered from 1 to 6. Types 1,2 are natural, types 3,4,5 are synthetic and type 6 is a placebo (a substance that has no therapeutic effect). In each of 8 different hospitals, each type of drug is given to exactly 4 patients.

6) Using contrasts, we would like to test the following:

L1: The difference between natural (types 1,2) and synthetic (types 3,4,5) drugs.
L2: The difference between placebo (type 6) and all other drugs (types 1,2,3,4,5).

Which of the following contrasts can be used?

a) L1: $(1, 1, -1, -1, -1, 0)$, L2: $(1, 1, 1, 1, 1, -5)$
b) L1: $(1, 1, -1, -1, -1, 0)$, L2: $(1, 1, 0, 0, 0, -2)$
c) L1: $(1/2, 1/2, -1/3, -1/3, -1/3, 0)$, L2: $(1, 1, 1, 1, 1, -5)$
d) L1: $(1/2, 1/2, -1/3, -1/3, -1/3, 0)$, L2: $(1, 1, 0, 0, 0, -2)$

7) We test contrast L1 from task 6) in `R` and get the following incomplete ANOVA table.

```
                 Df Sum Sq Mean Sq F value   Pr(>F)
Drug                 41.51
  Drug: L1            7.30
Hospital              6.64
Residuals           154.74
```

What is the F-value for testing contrast L1, rounded to two decimal places?

a) 9.56
b) 7.30
c) 7.74
d) 8.44

5. (7 points)

A coffee company would like to test 7 new types of coffee. They invited 7 people to test them.

1) They decided to construct a BIBD (Balanced incomplete block design), which is summarized in the table below. Each row in the table corresponds to a different person and the types of coffee are numbered from 1 to 7.

Blocks	Treatments		
1	A	2	5
2	B	4	5
3	1	3	6
4	2	3	7
5	2	4	6
6	1	4	7
7	5	6	7

What are the values of A and B?

a) $A = 1, B = 4$
b) $A = 4, B = 1$
c) $A = 3, B = 1$
d) $A = 1, B = 3$

The company extended their offer and they have now 10 types of coffee to test. This time they invited 15 people for testing.

2) Again they constructed a BIBD (Balanced incomplete block design). Each type of coffee will be tested 6 times and each person tests 4 coffees. Which of the following statements is true?

a) Exactly one person will test **both** coffee type 1 and coffee type 2.
b) Exactly two people will test **both** coffee type 1 and coffee type 2.
c) Exactly three people will test **both** coffee type 1 and coffee type 2.
d) Exactly four people will test **both** coffee type 1 and coffee type 2.

3) Order the following designs according to the expected carry-over effect, from the one where you expect the strongest carry-over effect to the design where you expect the weakest carry-over effect.

(i) Each person tests and rates every type of coffee within one hour.
(ii) Each person tests and rates exactly one type of coffee.
(iii) On each of 10 consecutive days, each person tests and rates a different type of coffee.

a) (i), (iii), (ii)
b) (i), (ii), (iii)
c) (ii), (iii), (i)
d) (iii), (ii), (i)
e) (iii), (i), (ii)

Additionally, the company wants to test the effect of different types of cream in combination with different types of coffee. There are 5 types of cream and 10 types of coffee. They have 50 people for testing.

4) Each person gets exactly one type of coffee with one type of cream and is asked to report how satisfied he or she is with the product. This is an example of a:
   a) Split plot design
   b) Latin square design
   c) Randomized complete block design
   d) Cross-over design
   e) Factorial design
   f) None of the previous designs
5) How many cups of coffee would the company need to prepare for a randomized complete block design with people as blocks?
   a) 50
   b) 250
   c) 500
   d) 2500

Look at the following experiment design, where A and B are factors. In each row of the blocks, the factor levels $A_j, j = 1, 2, 3, 4$ are randomly assigned.

Block 1

A_2B_1	A_1B_1	A_3B_1	A_4B_1
A_3B_3	A_4B_3	A_2B_3	A_1B_3
A_1B_2	A_2B_2	A_4B_2	A_3B_2

Block 2

A_1B_1	A_2B_1	A_3B_1	A_4B_1
A_3B_3	A_4B_3	A_2B_3	A_1B_3
A_4B_2	A_2B_2	A_1B_2	A_3B_2

6) This is an example of a:
   a) Factorial design
   b) Split plot design
   c) Latin square design
   d) Randomized complete block design
   e) Cross-over design
   f) None of the previous designs
7) Against which critical value do we test significance of the interaction effect A:B in the analysis corresponding to experiment design from task 6)?
   a) $F^{crit}_{2,8}$
   b) $F^{crit}_{3,9}$
   c) $F^{crit}_{6,9}$
   d) $F^{crit}_{6,36}$
   e) $F^{crit}_{4,8}$

Table of critical values at the 5% level of the F-distributions with ν_1 degrees of freedom in the numerator and ν_2 degrees of freedom in the denominator.

ν_2 \ ν_1 =	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	242.98	243.91	244.69	245.36	245.95
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.40	19.41	19.42	19.42	19.43
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74	8.73	8.71	8.70
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91	5.89	5.87	5.86
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68	4.66	4.64	4.62
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.98	3.96	3.94
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57	3.55	3.53	3.51
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28	3.26	3.24	3.22
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07	3.05	3.03	3.01
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91	2.89	2.86	2.85
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79	2.76	2.74	2.72
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69	2.66	2.64	2.62
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60	2.58	2.55	2.53
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.57	2.53	2.51	2.48	2.46
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48	2.45	2.42	2.40
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.46	2.42	2.40	2.37	2.35
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.41	2.38	2.35	2.33	2.31
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34	2.31	2.29	2.27
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.34	2.31	2.28	2.26	2.23
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28	2.25	2.22	2.20
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.28	2.25	2.22	2.20	2.18
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.26	2.23	2.20	2.17	2.15
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.24	2.20	2.18	2.15	2.13
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.22	2.18	2.15	2.13	2.11
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.20	2.16	2.14	2.11	2.09
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15	2.12	2.09	2.07
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.17	2.13	2.10	2.08	2.06
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.15	2.12	2.09	2.06	2.04
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.14	2.10	2.08	2.05	2.03
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.13	2.09	2.06	2.04	2.01
31	4.16	3.30	2.91	2.68	2.52	2.41	2.32	2.25	2.20	2.15	2.11	2.08	2.05	2.03	2.00
32	4.15	3.29	2.90	2.67	2.51	2.40	2.31	2.24	2.19	2.14	2.10	2.07	2.04	2.01	1.99
33	4.14	3.28	2.89	2.66	2.50	2.39	2.30	2.23	2.18	2.13	2.09	2.06	2.03	2.00	1.98
34	4.13	3.28	2.88	2.65	2.49	2.38	2.29	2.23	2.17	2.12	2.08	2.05	2.02	1.99	1.97
35	4.12	3.27	2.87	2.64	2.49	2.37	2.29	2.22	2.16	2.11	2.07	2.04	2.01	1.99	1.96
36	4.11	3.26	2.87	2.63	2.48	2.36	2.28	2.21	2.15	2.11	2.07	2.03	2.00	1.98	1.95
37	4.11	3.25	2.86	2.63	2.47	2.36	2.27	2.20	2.14	2.10	2.06	2.02	2.00	1.97	1.95
38	4.10	3.24	2.85	2.62	2.46	2.35	2.26	2.19	2.14	2.09	2.05	2.02	1.99	1.96	1.94
39	4.09	3.24	2.85	2.61	2.46	2.34	2.26	2.19	2.13	2.08	2.04	2.01	1.98	1.95	1.93
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.04	2.00	1.97	1.95	1.92

df	$t_{0.60}$	$t_{0.70}$	$t_{0.80}$	$t_{0.90}$	$t_{0.95}$	$t_{0.975}$	$t_{0.99}$	$t_{0.995}$
1	0.325	0.727	1.376	3.078	6.314	12.706	31.821	63.657
2	0.289	0.617	1.061	1.886	2.920	4.303	6.965	9.925
3	0.277	0.584	0.978	1.638	2.353	3.182	4.541	5.841
4	0.271	0.569	0.941	1.533	2.132	2.776	3.747	4.604
5	0.267	0.559	0.920	1.476	2.015	2.571	3.365	4.032
6	0.265	0.553	0.906	1.440	1.943	2.447	3.143	3.707
7	0.263	0.549	0.896	1.415	1.895	2.365	2.998	3.499
8	0.262	0.546	0.889	1.397	1.860	2.306	2.896	3.355
9	0.261	0.543	0.883	1.383	1.833	2.262	2.821	3.250
10	0.260	0.542	0.879	1.372	1.812	2.228	2.764	3.169
11	0.260	0.540	0.876	1.363	1.796	2.201	2.718	3.106
12	0.259	0.539	0.873	1.356	1.782	2.179	2.681	3.055
13	0.259	0.538	0.870	1.350	1.771	2.160	2.650	3.012
14	0.258	0.537	0.868	1.345	1.761	2.145	2.624	2.977
15	0.258	0.536	0.866	1.341	1.753	2.131	2.602	2.947
16	0.258	0.535	0.865	1.337	1.746	2.120	2.583	2.921
17	0.257	0.534	0.863	1.333	1.740	2.110	2.567	2.898
18	0.257	0.534	0.862	1.330	1.734	2.101	2.552	2.878
19	0.257	0.533	0.861	1.328	1.729	2.093	2.539	2.861
20	0.257	0.533	0.860	1.325	1.725	2.086	2.528	2.845
21	0.257	0.532	0.859	1.323	1.721	2.080	2.518	2.831
22	0.256	0.532	0.858	1.321	1.717	2.074	2.508	2.819
23	0.256	0.532	0.858	1.319	1.714	2.069	2.500	2.807
24	0.256	0.531	0.857	1.318	1.711	2.064	2.492	2.797
25	0.256	0.531	0.856	1.316	1.708	2.060	2.485	2.787
26	0.256	0.531	0.856	1.315	1.706	2.056	2.479	2.779
27	0.256	0.531	0.855	1.314	1.703	2.052	2.473	2.771
28	0.256	0.530	0.855	1.313	1.701	2.048	2.467	2.763
29	0.256	0.530	0.854	1.311	1.699	2.045	2.462	2.756
30	0.256	0.530	0.854	1.310	1.697	2.042	2.457	2.750
31	0.255	0.530	0.853	1.309	1.696	2.040	2.452	2.744
32	0.255	0.530	0.853	1.309	1.694	2.037	2.449	2.738
33	0.255	0.530	0.853	1.308	1.693	2.035	2.445	2.733
34	0.255	0.529	0.852	1.307	1.691	2.032	2.441	2.728
35	0.255	0.529	0.852	1.306	1.690	2.030	2.438	2.724
40	0.255	0.529	0.851	1.303	1.684	2.021	2.423	2.704
60	0.254	0.527	0.848	1.296	1.671	2.000	2.390	2.660
90	0.254	0.526	0.846	1.291	1.662	1.987	2.368	2.632
120	0.254	0.526	0.845	1.289	1.658	1.980	2.358	2.617
∞	0.253	0.524	0.842	1.282	1.645	1.960	2.326	2.576