Written Exam
(2 hours)

General remarks:

- 10 hand- or typewritten A4 pages (= 5 sheets written on both sides). Pocket calculator. No communication devices (mobile phones, WLAN, etc.).
- 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.
- 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. (10 Points) An experiment was conducted to compare the quality of four brands of jeans with a completely randomized design. Equal number of jeans were randomly selected from each brand and all selected jeans were given a fabric quality score (higher score indicates better quality).

We fit a one-way ANOVA model:

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij} \text{ for } i = 1, \ldots, 4, \text{ and } j = 1, \ldots, J$$

where α_i is the effect of the i-th brand on fabric score.

With `R` we obtain the following (shortened) summary output:

```
            Df Sum Sq Mean Sq F value Pr(>F)
brand        3   47.9   15.96    6.56 0.0014
Residuals   32   77.8    2.43
```

a) **(1 Pt)** How many jeans were selected from each brand?

b) **(2 Pts)** What null hypothesis should we test to assess whether the brands have an effect on fabric quality? Based on the given `R`-output, can we conclude that the four brands have different fabric quality? Justify your answer by explaining why you accept or reject the null hypothesis at a 5% significance level.

c) **(1.5 Pts)** Brands 1 and 4 are expensive brands and the other two brands are relatively inexpensive. Propose some contrasts to test the following null hypotheses. Write down each contrast in the form of a vector, e.g. $(1, -1, 0, 0)$.

- $L1$: Both expensive brands have the same fabric quality.
- $L2$: Brands 2 and 3 have the same fabric quality.
- $L3$: The fabric quality of expensive brands are no different than the other two brands.

d) **(1 Pt)** Are the previous 3 contrasts orthogonal to each other? Justify your answer.

e) **(1 Pt)** Is there any other contrast which is orthogonal to the previous 3 contrasts? Justify your answer.

f) **(2 Pts)** Compute the values of the appropriate F-ratio for testing the hypotheses $L1$, $L2$ and $L3$ given in part **c)** based on the `R`-output given above and the following information:

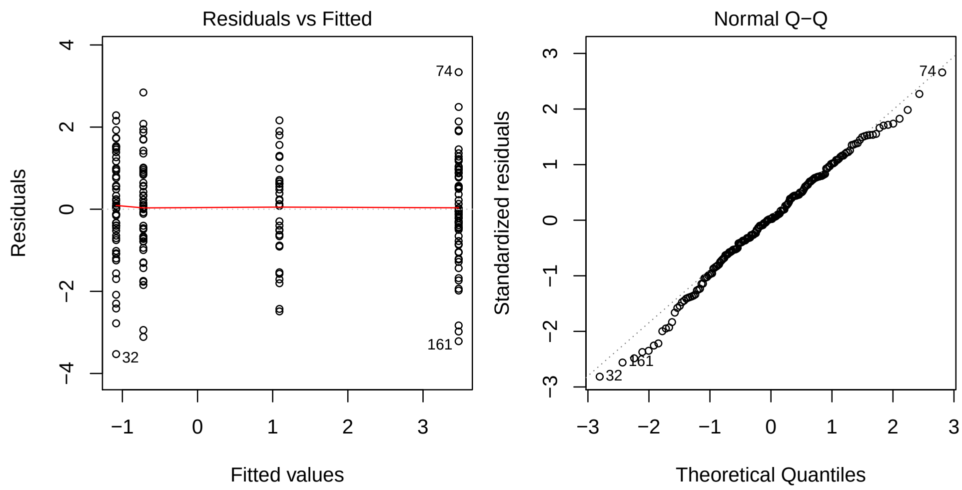
$$SS_{L1} = 5.65 \text{ and } SS_{L2} = 20.68,$$

where SS_c denotes the sum of squares associated with a contrast c.

g) **(1.5 Pts)** Which of the hypotheses $L1$, $L2$ and $L3$ can we reject based on the values of the test statistics computed in part **f)**? Justify your answer based on the information that the 95%-quantile of the $F_{1,32}$ distribution is 4.15. Note that we do not adjust for multiple testing in this case.

2. (13 Points) The Swiss Federal Statistical Office wants to study the influence of age and education on voting patterns and gives the task to Gian, one of its interns. Gian decides to call a random sample of 200 persons and collects their age (variable `age` with levels "young" or "old") and education level (variable `educ` with levels "low" or "high"). The response variable `y` is on a numerical scale from -10 to 10, with negative values meaning that the participant leans towards party A and positive values that he or she leans more towards party B.

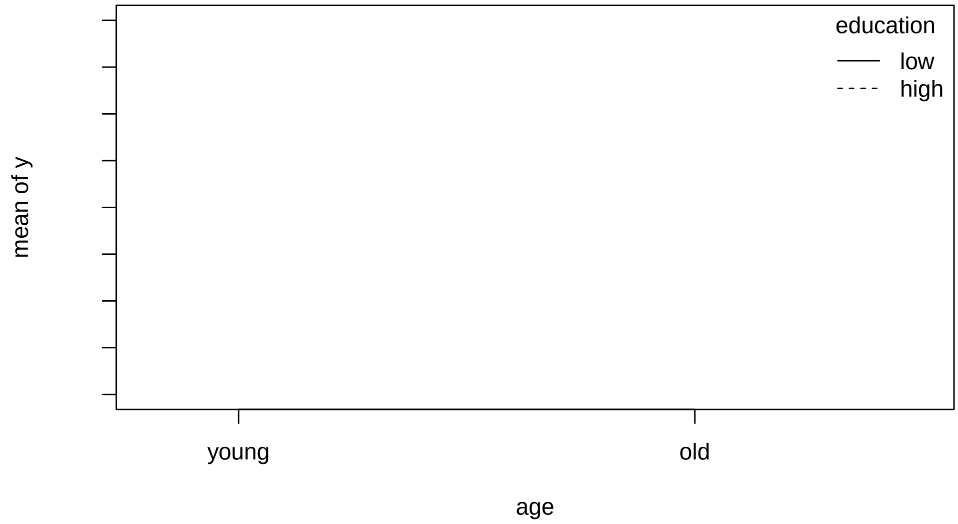
a) **(3 Pts)** Gian fits a two-way ANOVA model of y versus age and education with an interaction and obtains the following diagnostic plots. What assumptions can you check with each of these plots? Are these assumptions fulfilled in the present case?



b) **(3 Pts)** Gian looks at the estimated coefficients and wants to understand his model. Draw the interaction plot corresponding to the estimated (dummy-) coefficients shown below. Focus on the qualitative aspects. **Remark: Use the (empty) plot on the next page as template but reproduce the plot on the answer sheet!**

```
Full coefficients are

(Intercept):          -0.72
age:                  young     old
                        0.0     1.8
educ:                   low    high
                       0.00   -0.36
age:educ:         young:low old:low young:high old:high
                        0.0     0.0        0.0      2.7
```



c) (2 Pts) Have a look at the following output. Complete the missing parts.
Remark: write your answer on the answer sheet!

```
                Df  Sum Sq  Mean Sq   F-value      Pr(>F)
age              1     240      240        20    1.31e-05
educ             1       2        2      0.16       0.689
age:educ       ...     ...      ...         7       0.009
Residuals      196     ...       12
```

d) (1 Pt) Given the ANOVA table above, what is the estimated standard deviation of the errors?

e) (2 Pts) Given the ANOVA table above, what terms would you keep in the final model? Motivate your answer.

f) (2 Pts) The previous ANOVA table was produced by the function call:

```
fit1 <- aov(y ~ age * educ, data = survey)
summary(fit1)
```

To check that he did not make any mistake, Gian asks a colleague to repeat the analysis. The latter uses the function call:

```
fit2 <- aov(y ~ educ * age, data = survey)
summary(fit2)
```

To his surprise he obtains different p-values for both education and age. What is the reason that their outputs are different and does the p-value of the interaction term differ too? Motivate your answer.

3. (11 Points) A japanese tea vendor is interested in optimizing the tea he offers. To find the most popular tea, he performs a double-blind experiment with 10 randomly selected subjects and 4 different types of tea. Each tea is steeped for 90 seconds at 75 °C. Then the tea is cooled down to 40 °C. He offers each tea to each of the subjects in random order; in between, the subjects drink water to neutralize their tastebuds. The subjects report their ratings on a numerical scale. The vendor decides to use the model

$$\text{Rating}_{ij} = \mu + \text{Type}_i + \text{Subject}_j + \epsilon_{ij},$$

where `Type` is a fixed effect, `Subject` a random (block) factor and ϵ_{ij} is the error term.

a) (2 Pts) Specify the model assumptions on the random effect of `Subject` and on the error term.

b) (1 Pt) For fixed j, are the Rating_{ij} uncorrelated? Justify your answer.

c) (1.5 Pts) Specify the hypothesis that is tested in the following output and state the p-value.

```
> fit <- lmer(Rating ~ Type + (1 | Subject), data = tea)
> anova(fit)

Analysis of Variance Table of type III  with  Satterthwaite
approximation for degrees of freedom
     Sum Sq Mean Sq NumDF DenDF F.value    Pr(>F)
Type 49.524  16.508     3    27      30 9.544e-09 ***
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Now the tea vendor decides to perform an in-depth analysis with the R-function `summary(fit)` and `confint(fit, oldNames = FALSE)`. Below you see part of the R-output he obtains. He uses the encoding scheme `contr.treatment`, i.e. Type1 as reference level.

```
Random effects:
 Groups   Name        Variance Std.Dev.
 Subject  (Intercept) 0.8151   0.9029
 Residual             0.5503   0.7418
Number of obs: 40, groups:  Subject, 10

Fixed effects:
            Estimate Std. Error      df t value Pr(>|t|)
(Intercept)   0.8191     0.3695 17.3980   2.217   0.0402
Type2         1.7216     0.3317 27.0000   5.190 1.83e-05
Type3         2.9916     0.3317 27.0000   9.018 1.24e-09
Type4         2.3400     0.3317 27.0000   7.054 1.39e-07

> confint(fit, oldNames = FALSE)

                             2.5 %    97.5 %
sd_(Intercept)|Subject 0.51636160 1.5059552
sigma                  0.55725513 0.9276822
(Intercept)            0.09331001 1.5448264
Type2                  1.08446392 2.3586975
Type3                  2.35452624 3.6287598
Type4                  1.70286489 2.9770985
```

d) (1 Pt) Calculate the point estimate for the expected rating of the second tea type.

e) (1.5 Pts) Give point estimates for the variance of the random effect of `Subject` and for the variance of `Rating`.

f) (1 Pt) The vendor wants to make sure that the estimate for the expected rating of the second tea type is correct. A friend of his uses a fixed effect model and performs the analysis shown below. As above, he uses the encoding scheme `contr.treatment`. Based on the R-output below, he calculates $0.048+1.72 = 2.2$. What does this number refer to?

```
> fit2 <- aov(Rating ~  Type + Subject, data = tea)
> coef(fit2)
```

```
(Intercept)       Type2       Type3       Type4    Subject2
 0.04753849  1.72158070  2.99164302  2.33998167  0.29019204
   Subject3    Subject4    Subject5    Subject6    Subject7
 1.44925105 -0.55735453  1.01772272  0.39756117  1.24372812
   Subject8    Subject9   Subject10
 2.31235864 -0.38621618  1.94805373
```

g) (1.5 Pts) Can we reject the hypothesis that the true variance of the random effect of `Subject` is equal to 4 on a significance level of 0.05? Give reasons for your answer.

h) (1.5 Pts) The vendor wants to check the model assumptions. Which plots should he look at?

4. (7 Points) A company producing cheese would like to compare the effects of 4 different production techniques on the quality of cheese. 5 batches of milk are available for the experiment and each batch is divided into 4 equal parts.
The company initially considered to use a completely randomized design (on the 20 experimental units), but later (before doing the randomization) decided to use a randomized complete block design, where they blocked on the batches of milk.

D

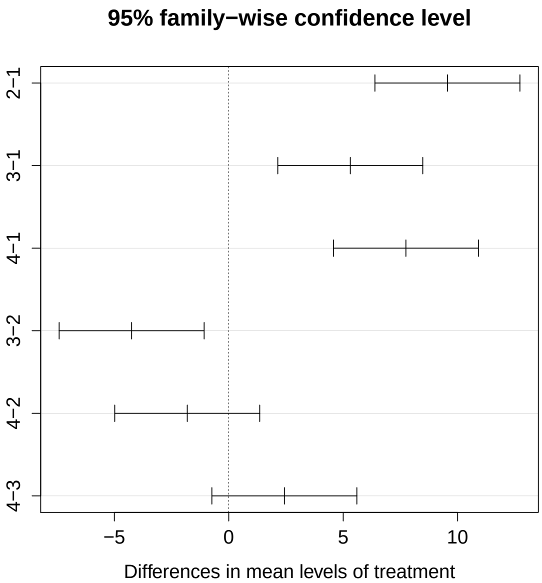
1) The company used blocking instead of a completely randomized design ...
   a) ...to test differences between the blocks.
   b) ...to ensure that the variance of the error is constant.
   c) ...to decrease the degrees of freedom of the error.
   d) ...to reduce the variance of the error.

~~D~~ b

2) Choose which one of the following statements about the randomized complete block design is true.
   a) A randomized complete block design tries to ensure heterogeneity *within* blocks.
   b) The error corresponding to a randomized complete block design with r blocks has $r-1$ less degrees of freedom than the error corresponding to a completely randomized design on the same number of experimental units.
   c) The number of experimental units in a block of a randomized complete block design can can be less than the number of treatments.
   d) A randomized complete block design can be constructed from a completely randomized design by imposing blocking *after* randomization.

3) We fitted the randomized complete block experiment proposed above in R. The fitted model is stored in `fit`. Look at the following R code and the resulting plot.

```
plot(TukeyHSD(x = fit, "treatment", conf.level = 0.95))
```



(The question continues on the next page!)

e

Which treatment differences are significant at the (family-wise) 5% level?

a) $2-1$, $3-1$, $4-1$
b) $3-2$
c) $4-2$, $4-3$
d) $2-1$
e) $2-1$, $3-1$, $4-1$, $3-2$
f) It cannot be determined from the given information.

e

4) Consider that we changed the (family-wise) significance level from 5% to 10% in part 3) above. Which statement would then be true?

a) All the confidence intervals would be longer.
b) There would be no change in the confidence intervals.
c) Some treatment differences that were significant may become insignificant (i.e. not significant).
d) All the confidence intervals would be shifted to the right.
e) Some treatment differences that were not significant may become significant.

d

5) In addition to blocking on the batches, the company would also like to block on "temperature", which has 5 different levels. They again have 5 batches of milk, each batch divided into 4 parts. Which design can they use for this experiment?

a) They can still use a randomized complete block design
b) Latin square
c) Graeco-Latin square
d) None of the above designs can be used

b

6) (Independent of above story) We designed a two-factor factorial with factors A and B in a randomized complete block design with a blocking factor C. Which of the following models should we use if we assume an additive effect of the block factor? Remark: α_i are the coefficients of A, β_j the coefficients of B and γ_k the coefficients of C.

a) $Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + \gamma_k + \epsilon_{ijk}$
b) $Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_k + \epsilon_{ijk}$
c) $Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \epsilon_{ijk}$
d) $Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_k + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijk}$

~~a~~ b

7) A company wants to study the effect of irrigation (with three levels A, B, C) on the quality of crops. Which of the following designs should they prefer if we are interested in treatment differences? Assume that the variance of the units is the same for both designs.

D1: Rent 2 pieces of land with three sub-fields each and form two complete blocks (by blocking on the pieces of land).

D2: Rent 3 pieces of land with two sub-fields each and form three blocks (by blocking on the pieces of land): Block 1: A,B, Block 2: A,C, Block 3: B,C.

a) They should prefer design D1.
b) They should prefer design D2.

c) There is no important difference between the designs.
d) It cannot be decided which design is preferable based on the given information.

5. (6 points)

d

1) Consider the following experimental design with one factor with levels A, B, C and D.

Block 1	Block 2	Block 3	Block 4	Block 5	Block 6
A	B	B	A	C	D
C	D	A	D	B	C

This is a

A

B

a) Two-series factorial design
b) Complete block design
c) Latin square
d) Balanced incomplete block design
e) Partially balanced incomplete block design

d

2) A chocolate company wants to study the effect of 4 different recipes on the quality of chocolate. They have 4 production workers and 4 machines available. Each worker will try each recipe exactly once, and each recipe will be tried on each machine exactly once. This is a

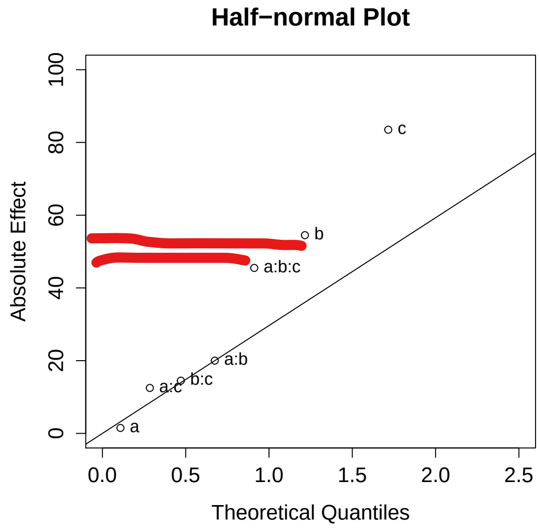
a) Split plot design
b) Complete block design
c) Balanced incomplete block design
d) Latin square
e) Two-series factorial design

c

3) The chocolate company further wants to study the effect of cooking temperature and amount of milk on the smoothness of the chocolate. There are 9 days available for the experiment. On each day, 4 batches of chocolate liquor are produced. The machine used for cooking the chocolate liquor operates at one of 3 different temperatures, and this setting applies for the whole day. Each level of cooking temperature is randomly assigned to 3 days. On each day, 4 different amounts of milk (1, 2, 2.5, 3 liters of milk) are randomly assigned to the 4 batches. This is a

a) Complete block design
b) Balanced incomplete block design
c) Split plot design
d) Latin square
e) Two-series factorial design

4) We fitted a two-series factorial experiment in R. Based on this fit, we plotted the following half-normal plot.



Which of the (incomplete) R outputs below corresponds to the given half-normal plot?

a

a)
```
   a      b      c    a:b  a:b:c
 1.5   54.5  -83.5  -20.0  -45.5
```
b)
```
   a      b      c    a:b  a:b:c
 1.5   45.5   83.5  -20.0  -45.5
```
c)
```
   a      b      c    a:b  a:b:c
 1.5   54.5   83.5   -1.5   45.5
```
d)
```
   a      b      c    a:b  a:b:c
 1.5   54.5   53.5  -20.0  -45.5
```

5) We fitted a two-series factorial model ~~without replicates~~ in R, including all the interaction effects. Using this fit we cannot test hypotheses about the effects or construct confidence intervals for the effects. This is because

c

a) the errors are not independent
b) the errors are not identically distributed
c) there are not enough observations compared to the number of parameters in the model
d) the residuals only have 1 degree of freedom
e) some main effect is confounded with another main effect

6) We consider a split-plot experimental design with factor A applied on the whole plots and factor C applied on the split plots. There is an additional factor B which represents complete blocks on the whole-plot level. To fit a model to this experiment in R, we should use the command:

a) `lmer(Y ~ B * A + A * C)`
b) `lmer(Y ~ B + A * B + (1 | B))`
c) `lmer(Y ~ B + A * C + (1 | B:A))`
d) `lmer(Y ~ B + A * B + (1 | B:C))`