NTNU
Norwegian University of
Science and Technology

Department of Mathematical Sciences

Examination paper for **TMA4255 Applied statistics**

Academic contact during examination: Nikolai Ushakov
Phone: 45128897

Examination date: 02 June 2018
Examination time (from–to): 09:00 – 13:00
Permitted examination support material: C:
– Tabeller og formler i statistikk, Tapir forlag,
– Stamped yellow A4 sheet with your own handwritten notes,
– Specified calculator.

Other information:
– In outputs from MINITAB comma is used as decimal separator.
– Significance level 5% should be used unless a different level is specified.
– All answers need to be justified.

Language: English
Number of pages: 3
Number of pages enclosed: 0

Checked by:

Informasjon om trykking av eksamensoppgave	
Originalen er:	
1-sidig ☐	**2-sidig** ☒
sort/hvit ☒	**farger** ☐
skal ha flervalgskjema ☐	

Date Signature

Problem 1

Two independent samples of sizes $n = 200$ and $m = 240$ are taken from normal distributions with unknown expectations μ_X, μ_Y and known variances $\sigma_X^2 = 1$ and $\sigma_Y^2 = 1.2$, respectively. $H_0 : \mu_X = \mu_Y$ is being tested against $H_1 : \mu_X \neq \mu_Y$.

a) Find the $P-$value if the observed sample means are $\bar{x} = 2.1$ and $\bar{y} = 2.0$.

Problem 2

At a laboratory the connection between reaction velocity Y (in micromoles per hour) and consentration x (in micromoles per dm^3) of a catalyst is investigated. Ten measurements of reaction velocity Y_i and concentration x_i are made, $1 \leq i \leq 10$. Assume that the measurements are independent, and that Y_i has a normal distribution with expected value $\beta_0 + \beta_0 x_i$ and standard deviation σ, where β_0, β_1 and σ are unknown parameters.

a) By the method of least squares the estimate of β_1 is 1.12. The estimate of σ^2 is 2.3, and $\sum_{i=1}^{10}(x_i - \bar{x})^2 = 4.1$. Perform a hypothesis test to investigate whether there is a connection between x og Y. Use significance level 0.05.

Problem 3 A multiple linear regression model is considered. It is assumed that

$$Y_{ijk} = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2j} + \epsilon_{ijk},$$
$$i = 1, 2, 3;\ j = 1, 2, 3, 4;\ k = 1, 2,$$

where ϵ_{ijk} are independent and have the same normal distribution. There are some data, and the output from a regression analysis using MINITAB (a part) is given below. Use this when solving the Problem.

```
The regression equation is
Y = 15,1 - 10,0 x1 + 0,808 x2

Predictor        Coef     SE Coef
Constant       15,140       1,848
x1           -10,0250      0,9466
x2            0,80842     0,01653
```

a) Test the hypotheses $H_0 : \beta_1 = 0$ vs. $H_1 : \beta_1 \neq 0$ and $H_0 : \beta_2 = 0$ vs. $H_1 : \beta_2 \neq 0$. Write down the relevant test statistics and perform the two tests with significance level 0.01.

Problem 4

A research institute has five different types of instruments for measuring the amount of infrared radiation. An experiment is done to check whether the different instruments give similar measurements. For each of 6 objects the amount of infrared radiation is measured with each of the five instruments. The six objects used all differ in terms of type of material, temperature and size. Assume that all measurements are independent and normally distributed with the same variance. A analysis of variance model (ANOVA) was fitted to the data, and the MINITAB output (a part) is given below.

Source	DF	SS	MS	F	P
Instrument	?	8,00	?	?	0,025
Object	?	?	1,54	?	0,050
Error	?	?	?		
Total	?	?			

a) What design of experiment is used in the situation described above? Eleven of the entries of this MINITAB printout are replaced by a question mark (?). Find numerical values for these eleven missing entries. Show how you calculate these values.

b) Two $P-$values are given in the ANOVA table. Specify the null hypotheses, H_0, corresponding to each of these two $P-$values.

Which of the two $P-$values is of interest for the research institute? What is the conclusion of the corresponding test if the significance level is 0.05?

Problem 5

Let Y be measured percentage of fat in a certain type of sausages. A laboratory has measured the fat percentage in 15 sausages and the results $y_1, y_2, ..., y_{15}$ are supposed to be realisations of independent continuous random variables with a symmetric (about the unknown expectation) distribution. The results are

$$19.2, 27.6, 25.6, 32.2, 17.7, 20.5, 23.9, 20.2, 24.2, 26.1, 32.0, 24.8, 28.9, 16.2, 18.7.$$

Fat percentage 20.0 is considered as normal. We wish to test the hypothesis that the fat percentage is normal (i.e. equals 20.0) versus the alternative that it is greater than 20.0. The significance level is 0.05.

a) Test the hypothesis using the large-sample sign test.

b) Test the hypothesis using the large-sample Wilcoxon signed rank test.

c) Suppose that in addition to the conditions above it is known that the distribution of the fat percentage is approximately normal. Test the hypothesis using the $t-$test (for simple calculations you can use that $\sum_{i=1}^{15} y_i = 357.8$ and $\sum_{i=1}^{15} y_i^2 = 8888.62$).