



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: 27 May 2017

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: 4

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

A manufacturer of fertilizers studies the difference between the effect of three fertilizers A, B and C on the growth of plants. An experiment was conducted where plants were grown under identical conditions and one randomly assigned fertilizer: A to 7 plants, B to 7 plants, C to 7 plants. After a certain time the height of the plants were measured in cm. Data from the experiment (rounded to the nearest integer) are presented in the table below.

A	46	39	53	44	46	50	54	sum is 332	sum of squares is 15914
B	58	40	51	49	53	49	52	sum is 352	sum of squares is 17880
C	61	57	55	57	62	67	59	sum is 418	sum of squares is 25058

- a) Assume that all measurements are independent and normally distributed with the same variance. A one-way analysis of variance model (ANOVA) was fitted to the data, and the MINITAB output is given below.

One-way ANOVA: A;B;C

Source	DF	SS	MS	F	P
Factor	?	?	?	?	0,001
Error	?	?	24.7		
Total	?	1023,2			

S=4,970 R-Sq=56,55% R-Sq(adj)=51,72%

Seven of the entries of this MINITAB printout are replaced by a question mark (?) Find numerical values for these seven missing entries. Show how you calculate these values.

Perform a hypothesis test that there is a difference between the effect of the three fertilizers. Write down the null hypothesis and the alternative. Base the test on what you have in the MINITAB printout above. Use significance level $\alpha = 0.05$.

- b) Perform a multiple comparison, using the Tukey test. Again $\alpha = 0.05$. Use that $q(0.05, 3, 18) = 3.61$, where $q(\alpha, m, n)$ is such a value that

$$P(Q_{m,n} \geq q(\alpha, m, n)) = \alpha,$$

where the random variable $Q_{m,n}$ has the studentized range distribution with m and n degrees of freedom.

- c) Test the hypothesis that effects of fertilizers A and B are equal, using the two-sample t -test. Compare the result with the result obtained in b).
- d) Now we do not assume that the observations are normally distributed. Solve the hypothesis testing problem in a) using a suitable nonparametric test.

Problem 2

77 goals were scored on the 2012 European Football Championship. The table shows the number of games where 0,1,2 etc. goals were scored.

The number of goals in a game	The number of games
0	2
1	6
2	10
3	6
4	3
5	3
6	1
7+	0

- a) Test the hypothesis that the number of goals, scored in a game, has the Poisson distribution with parameter $\lambda = 1$. The significance level is 0.05.
- b) Now we test the hypothesis that the number of goals, scored in a game, has a Poisson distribution, where the parameter is not specified. This is done using MINITAB. The MINITAB output is given below (except P-Value). Explain how the numbers in the column "Probability" have been obtained. Why the value of "DF" is 4? How is the value of "Chi-Sq" obtained from columns "Observed" and "Expected"? What is the conclusion now? Is the missing P-Value in the MINITAB output greater or less than 0.05?

Poisson mean for C1 = 2,48387

C1	Poisson			Contribution to Chi-Sq
	Observed	Probability	Expected	
0	2	0,083420	2,58601	0,132795
1	6	0,207204	6,42332	0,027898
2	10	0,257334	7,97734	0,512845

3	6	0,213061	6,60490	0,055398
4	3	0,132304	4,10143	0,295786
>=5	4	0,106678	3,30700	0,145220

N	N*	DF	Chi-Sq	P-Value
31	0	4	1,16994	

3 cell(s) (50,00%) with expected value(s) less than 5.

Problem 3

The amounts (in grams) of a chemical compound Y that dissolved in 100 grams of water at various temperatures x , were recorded as follows:

$x(^{\circ}\text{C})$	16	20	24	28	32	36	40	44	48	52	56	60	64	68	72
Y	13	18	24	26	27	30	31	35	42	43	44	47	49	53	58

For these data

$$\sum_{i=1}^{15} x_i = 660, \quad \sum_{i=1}^{15} Y_i = 540, \quad \sum_{i=1}^{15} x_i^2 = 33520, \quad \sum_{i=1}^{15} x_i Y_i = 27048.$$

We would like to fit a simple linear regression model

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, 2, \dots, 15,$$

where the ϵ_i are independent and normally distributed random variable with zero expectation and (unknown) variance σ^2 .

- Find the equation of the regression line. Estimate the amount of chemical that will dissolve in 100 grams of water at 70°C .
- A part of MINITAB output for these data is given below.

S = 1,72856 R-Sq = 98,4% R-Sq(adj) = 98,3%

Analysis of variance

Source	DF	SS	MS	F
--------	----	----	----	---

Regression	1	2413,2	2413,2	807,64
Residual Error	13	38,8	3,0	
Total	14	2452,0		

Using this output find the estimated standard deviation of the estimator of the slope. Find a 95% confidence interval for the slope.

How is the R-Sq value obtained from the numbers given in the Analysis of variance table?

The observed value 807.64 of some test statistic F is given in the Analysis of variance table. Which hypothesis is tested? What is the conclusion if the significance level is $\alpha = 0.01$?