NTNU
Norwegian University of Science and Technology

Department of Mathematical Sciences

Examination paper for **ST1201/ST6201 Statistical methods**

Academic contact during examination: Nikolai Ushakov

Phone: 45128897

Examination date: 20 December 2016

Examination time (from–to): 09:00 – 13:00

Permitted examination support material: C:
– Tabeller og formler i statistikk, Tapir forlag,
– K.Rottman. Matematisk formelsamling,
– Stamped yellow A4 sheet with your own handwritten notes,
– Calculator: HP30S, Citizen SR-270X, Citizen SR-270X College or Casio fx-82ES PLUS.

Other information:
Sensur:

Language: English

Number of pages: 3

Number of pages enclosed: 0

Checked by:

Date Signature

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

Problem 1

At a laboratory, a solvent is used which must have $p\text{H} = 7.45$. n random triales of a portion of the solvent are made, and in each trial pH is measured. The measurements $X_1, X_2, ..., X_n$ are independent normally distributed random variables with expectation μ, which is pH of the solvent, and standard deviation 0.05 (caused by the measurement errors).

A test is worked out with null hypothesis $\mu = 7.45$ versus alternative $\mu > 7.45$. If the null hypothesis is rejected with level of significance 0.05, then the portion is scraped.

a) Suggest how such a test, based on the sample mean $\bar{X}$, can be worked out. What is the conclusion if $n = 20$ and the sample mean is 7.47?

b) What is the probability that the null hypothesis will be rejected if $n = 20$ and pH of the solvent is equal to $\mu = 7.47$?

c) Suppose that pH of the solvent is $\mu = 7.47$. How many trials n must be made for that probability to reject the null hypothesis will be greater than 0.8?

Problem 2

Let $X_1, X_2, ..., X_n$ be a random sample from a normal distribution with (known) mean value $E(X_i) = 1$ and (unknown) variance $\text{Var}(X_i) = \theta$. One wants to use the observed values to test

$$H_0 : \theta = 1 \quad \text{versus} \quad H_1 : \theta \neq 1.$$

a) Using

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}(X_i - 1)^2$$

as the test statistic, work out a decision rule for when to reject H_0. Use significance level α.

Find the power function of the test.

b) Find the generalised likelihood ratio (GLR), λ, for the hypotheses H_0 and H_1 given above.

Explain why the test you worked out in **a**) is not a generalised likelihood ratio test (GLRT).

Problem 3

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 pairs of measurements are independent, and that Y_i has a normal distribution with expected value $\alpha + \beta x_i$ and standard deviation σ, where α, β and σ are unknown parameters.

a) Explain briefly the method of least squares for estimating α and β.

By the method of least squares the estimate of β is 1.12. The estimate of σ^2 is 2.3, and

$$\sum_{i=1}^{10} (x_i - \bar{x})^2 = 4.1$$

(that is, 2.3/4.1 is an estimate of the variance of the estimator of β).

b) Perform a hypothesis test to investigate whether there is a connection between x og Y. Use significance level 0.05.

Problem 4

In this exercise we will consider a regression model that is somewhat modified relative to what is discussed in the textbook. Assume that we have pairs of variables

$$(x_1, Y_1), (x_2, Y_2), ...(x_n, Y_n)$$

where $x_1, x_2, ..., x_n$ are nonstochastic while $Y_1, Y_2, ..., Y_n$ are assumed to be independent random variables with

$$E(Y_i) = \alpha + \beta(x_i - \bar{x}) \text{ and } \operatorname{Var}(Y_i) = \sigma_0^2.$$

Here $\bar{x} = (1/n) \sum_{i=1}^n x_i$, parameters α and β are unknown, while the variance σ_0^2 is known.

a) Work out the maximum likelihood estimators (MLE) for α and β and show, in particular, that the estimator for β can be written on the form

$$\hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x}) Y_i}{\sum_{i=1}^n (x_i - \bar{x})^2}.$$

show that the variance of $\hat{\beta}$ can be written on the form

$$\text{Var}(\hat{\beta}) = \frac{\sigma_0^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}.$$

b) What is the probability distribution of $\hat{\beta}$? Give reason for the answer.

Work out a $100(1-\delta)\%$-confidence interval for β.

Problem 5

The following table is an ANOVA table in which some entries are lost (stars).

Source	df	SS	MS	F
Treatment	$\star$	24.48	8.16	$\star$
Error	40	$\star$	5.1	
Total	$\star$	$\star$		

a) Find lost values and fill in the ANOVA table. Show how you calculate values where there are $\star$ in the table.

Test the hypothesis that

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4.$$

The significance level $\alpha = 0.05$.