NTNU – Trondheim
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: 11 December 2014

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:

Language: English

Number of pages: 4

Number pages enclosed: 0

Checked by:

Date Signature

Problem 1

An agent goes to regular shooting practice. Experience tells him that the probability of hit is $p = 0.6$. During a practice session he has 20 trials. Assume that each shot is either a hit or a miss, and that the trials are independent. The boss decides that the agent should have a new gun. They hope that this new one results in a better hitting probability. They want to check if this may hold, and the agent does a usual practice session consisting of 20 trials with the new gun.

a) Formulate the problem as a hypothesis test.

Use the common normal approximation to perform the test at significance level $\alpha = 0.05$ when the observed number of hits is 18.

b) What is the P-value of the test when he hits on 18 shots?

Problem 2

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 positive and nonstochastic while $Y_1, Y_2, ..., Y_n$ are assumed to be independent random variables with

$$Y_i \sim \mathrm{N}(\beta x_i, \sigma_0^2 x_i^2).$$

Thus, the variance of Y_i is assumed to be proportional to x_i^2. In this exercise we will assume the value of σ_0^2 to be known while the parameter β is estimated from the available data.

a) Work out the maximum likelihood estimator (MLE) for β and show that it can be written on the form

$$\hat{\beta} = \frac{1}{n} \sum_{i=1}^{n} \frac{Y_i}{x_i}.$$

b) Show that $\hat{\beta}$ is unbiased and find the variance of $\hat{\beta}$.

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

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

Problem 3

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.

A (partialy finished) analysis of variance (ANOVA) table for the measurements reads as follows.

Source	df	SS	MS	F	P-value
Instrument	$\star$	8	$\star$	$\star$	0.025
Object	$\star$	$\star$	1.54	$\star$	0.05
Error	$\star$	$\star$	$\star$		
Total	$\star$	$\star$			

a) What design of experiment is used in the situation described above? Write down the complete ANOVA table. In particular, show how you obtain the values for the positions with $\star$ in the above table.

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 4

Darwin (1876) studied the growth of pairs of corn plants, where one plant was produced by cross-fertilization and the other produced by self-fertilization. His goal was to demonstrate the greater fitness (e.g. survival and growth) of cross-fertilized plants compared to self-fertilized plants.

Fifteen pairs of plants were grown together under identical conditions in every pair (but possibly under different conditions in different pairs). The data recorded are the height (in inches) of the plants in each pair.

For pair i let X_{1i} denote the height of the plant from the seedling produced by cross-fertilization and X_{2i} denote the height of the plant from seedling produced by self-fertilization, $i = 1, ..., 15$. Further, let $D_i = X_{1i} - X_{2i}$. Data from the experiment are presented below.

i	1	2	3	4	5	6	7	8	9	10
x_{1i}	188	96	168	176	153	172	177	163	146	173
x_{2i}	130	163	160	160	147	149	149	122	132	144

i	11	12	13	14	15
x_{1i}	186	168	177	184	96
x_{2i}	130	144	102	124	144

Descritive measures for this dataset are

$$\bar{d} = \frac{1}{15}\sum_{i=1}^{15} d_i = 21.53,$$

$$s_d = \sqrt{\frac{1}{14}\sum_{i=1}^{15}(d_i - \bar{d})^2} = 38.29.$$

a) Assume that X_{1i} and X_{2i} are normally distributed, $X_{1i} \sim N(\mu_1 + \beta_i, \sigma^2)$ and $X_{2i} \sim N(\mu_2 + \beta_i, \sigma^2)$, $i = 1, ..., 15$.

Based on this experiment, could Darwin conclude that cross-fertilazed plants are taller than self-fertilized plants? Write down the null hypothesis and the alternative hypothesis, choose a test statistics and perform a hypothesis test. Use significance level $\alpha = 0.05$.

b) Assume that X_{1i} og X_{2i} are not normally distributed (but have simmetrical around expectations distributions). Perform a sign test to test if cross-fertilized plants are taller than self-fertilized plants.

Problem 5

A study of the independence of the temperament of husbands and wives was conducted. 111 married couples were randomly selected and a relative of the couple crossclassified the husband and wife into either having good ar bad temperament.

a) Is it any reason to believe that the temperament (good/bad) of the husband is dependent on the temperament (good/bad) of the wife? Write down the null hypothesis and the alternative hypothesis and perform a hypothesis test on the basis of the table. Use the significance level 0.05.

	Good wife	Bad wife
Good husband	24	27
bad husband	34	26