Norwegian University
of Science and Technology
Department of Mathematical Sciences

English

Contact during examination: Jarle Tufto
Telephone: 73 59 18 88, 997 05 519

ST0201 Statistics with Applications

Tuesday 20 May 2008
9:00–13:00
Permitted aids: Any written and printed material. One calculator.
Grades to be announced: 10 June 2008

The final examination consists of two parts:

1. The problems on the next two pages.
2. Appendix with a multiple choice questionnaire.

The Appendix is to be submitted with the form filled in together with the answer to part (1). Part (1) and (2) count equally in the evaluation of the final examination.

In addition to the final examination the mid-term examination counts 20% if it is advantageous to the candidate.

In the evaluation of part (1) (the next two pages) each of the six points counts equally.

In part (1) you should demonstrate how you arrive at your answers (e.g. by including intermediate answers or referral to theory). Answers based on calculator or table look-up only will not be accepted.

Problem 1

The literature states that the mean length in Norway of the tail of a mammalian species is 30 cm. A biologist think that the mean (that is, the expected value of the tail length of a randomly chosen individual) is greater, and she does an experiment to investigate this. She has the tail lengths y_i measured for a random sample of 10 individuals, and obtains the following results (in cm):

y_i	32.8	36.8	30.9	34.0	38.2	33.4	21.0	33.7	34.6	26.2

It is given that $\bar{y} = 32.16$ and $\sum(y_i - \bar{y})^2 = 233.324$.

a) Perform a test to investigate whether the expected tail length is greater than 30 cm (the null hypothesis is that it is less than or equal to 30 cm). Assume that tail length is normally distributed. Use significance level $\alpha = 0.05$.

b) Find et 95 % confidence interval for the expected tail length.

c) Also the latitude x_i where the animal stayed when the tail was measured was recorded. The table shows latitude and tail length for each animal:

x_i	63.9	61.5	64.8	65.5	59.0	58.5	68.5	66.0	66.0	66.8
y_i	32.8	36.8	30.9	34.0	38.2	33.4	21.0	33.7	34.6	26.2

Assume a linear regression model, where latitude is explanatory variable and tail length response variable. The biologist has a suspicion that the tail length decreases with increasing latitude.

It is given that $\bar{x} = 64.05$, $\sum(x_i - \bar{x})^2 = 100.465$, $\sum(x_i - \bar{x})(y_i - \bar{y}) = -105.88$ and $SS_E = 121.737$. Estimate the regression line (find $\hat{\alpha}$ and $\hat{\beta}$). Perform a test to investigate the biologist's suspicion. Use significance level 0.05.

d) Find an approximate 95 % confidence interval for the correlation coefficient between latitude and tail length.

Problem 2

The time between two events in a series of chemical reactions is exponentially distributed with mean (expected value) μ, that is, with probability density $\frac{1}{\mu}e^{-x/\mu}$. In a variant of the reaction series this time is exponentially distributed with mean $\mu/2$, that is, with probability density $\frac{2}{\mu}e^{-2y/\mu}$.

The time with expected value (mean) μ is measured 8 times, and independent observations $X_1, X_2, \ldots, X_8$ are obtained, and the time with expected value $\mu/2$ is measured 4 times, and independent observations Y_1, Y_2, Y_3, Y_4 are obtained. Two estimators for μ, $\hat{\mu} = \frac{2}{3}(\overline{X} + \overline{Y})$ and $\mu^* = \frac{1}{2}\overline{X} + \overline{Y}$, are suggested.

a) Find the mean (expected value) and variance of the two estimators. Which would you prefer?

b) Show that $\hat{\mu}$ is the maximum likelihood estimator.