Norwegian University of Science and Technology
Department of Mathematical Sciences

English

Contact during examination: Øyvind Bakke
Telephone: 73 59 81 26, 990 41 673

ST0201 Statistics with Applications

Friday 15 May 2009
9:00–13:00
Permitted aids: Any written and printed material. One calculator.
Grades to be announced: 5 June 2009

In the assessment each of the ten points counts equally.

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

A laboratory has developed a new procedure to perform a chemical reaction. The old procedure had a success rate of 0.357 – that is, the probability that the reaction would succeed was 0.357. They do 50 trials using the new reaction, and of these 24 succeed.

a) Find a 95% confidence interval for the new success rate.

b) How many trials would have to be performed for the confidence interval to have a guaranteed length of less than 0.1, regardless of the value of the new success rate?

c) Perform a test with significance level 0.05 to investigate whether the new procedure has a higher success rate than 0.357. The null hypothesis is that the new procedure has a success rate that is not higher than 0.357.

d) We do 50 new trials. What is the probability of rejecting the null hypothesis if the success rate is 0.5?

Problem 2

Twelve animals from a rodent population are equipped with GPS transmitters in order to examine the distance they cover during the month of June.

The results for the twelve animals (in km) was:

43.1 43.0 35.2 54.7 62.9 37.1 51.4 64.7 43.6 30.6 37.4 48.3

Assume that the observations are independent and come from the same normal distribution. It is given that the sample mean is 46.0 and the sample variance is 116.0.

a) Find a 95% confidence interval for the expected (mean) distance a randomly selected animal from the population covered during June.

It is suspected that the animals of a population of the same species in another part of the country cover a greater distance during June, so 14 animals of this population were equipped with GPS transmitters, too. They covered the following distances:

45.9 41.3 52.5 48.3 58.0 50.5 51.3 64.6 58.9 60.8 39.9 59.4 67.7 39.9

It is given that the sample mean is 52.8 and the sample variance is 82.6.

b) Perform a hypothesis test to examine whether the latter population has a greater expected (mean) covered distance than the former (the null hypothesis is that this is not the case). Use significance level 0.05.

c) Perform a non-parametric test with significance level 0.05 to examine whether the median of the covered distance is larger in the latter than in the former population. It is given that the rank sum of the sample of 12 animals (when we rank from the lowest to the highest observation) is 131.5. Compare with the answer to (b), and comment.

Problem 3

It is assumed that earthquakes in California over a certain magnitude (4 on the Richter scale) have a magnitude X (on the Richter scale) that follows a probability distribution with density $f(x) = \frac{1}{\theta}e^{-(x-4)/\theta},\ \ x > 4$, where θ is a positive parameter.

a) 100 earthquakes were examined. Of those, 64 had a magnitude between 4 and 4.2, 21 had a magnitude between 4.2 and 4.4, 10 had a magnitude between 4.4 and 4.6, while 5 had a magnitude larger than 4.6.

Perform a test with significance level 0.05 to examine whether the magnitudes follow the above mentioned distribution with $\theta = 0.21$ (the alternative hypothesis being that they do not).

b) Find the maximum likelihood estimator $\hat{\theta}$ for θ based on n earthquakes. What is the estimate based on 100 earthquakes of mean magnitude 4.21? Find the expected value and variance of $\hat{\theta}$.

c) Find a 95% confidence interval for θ based on 100 earthquakes of mean magnitude 4.21. It is given that $P(Y < 162.7) = 0.025$ and $P(Y > 241.1) = 0.025$ when Y has the chi-square distribution with 200 degrees of freedom.