**Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag**

English

Contact during exam:
John Tyssedal 73 59 35 34
Arvid Næss 99 53 83 50

EXAM IN COURSE SIF5060 STATISTICS

Tuesday November 28th 2002
Time: 09:00–14:00

Hjelpemiddel: Godkjent lommekalkulator med tomt minne.
Statistiske tabeller og formler, Tapir forlag.
K. Rottman: Matematisk formelsamling.

The examination results are due January 12th 2003.

Problem 1 Earthquake damages

A town is located in an area where earthquakes occur. The next earthquake that will hit the town, can be classified as strong (K), moderate (M) or weak (S). An arbitrary building of a certain type will give way if accumulated damage D becomes greater than a critical value $d_0 = 1.0$. Assume that for each of the three earthquake classes D can be modeled as a random variable with probability density function

$$f_D(d) = \begin{cases} \lambda e^{-\lambda d} & , \quad d \geq 0 \\ 0 & , \quad d < 0. \end{cases}$$

The parameter λ has a value determined by the intensity of the earthquake. The associated parameter values for the three classes K, M and S are $\lambda_K = 1.61$, $\lambda_M = 3.00$ and $\lambda_S = 4.60$.

a) Determine for each of the three earthquake classes the probability that an arbitrary building of the given type will give way.

It is estimated that the next earthquake that hits the area, will be strong, moderate or weak with probabilities $\mathrm{P}(K) = 0.02$, $\mathrm{P}(M) = 0.20$ and $\mathrm{P}(S) = 0.78$, respectively.

b) Determine the probability that a building of the given type gives way in the next earthquake.

What is the probability that the eartquake was moderate given that the building gives way?

Assume that A and B represent two buildings of the given type. It is estimated that if buiding A gives way, then the probability that building B gives way is 0.50 for K, 0.15 for S and 0.02 for S.

c) Determine the probability that both buildings give way in the next earthquake.

What is the probability that the earthquake was not strong given that building A gave way after an eartquake, while building B did not?

Problem 2 Norwegian earthquakes

The absolute strength of an earthquake is measured on the Richter-scale, which is logarithmic. The strongest Norwegian earthquake that is registered, is measured to 5.8 on this scale. In **2a**) og **2b**) we shall in particular consider earthquakes registered offshore outside Mid-Norway of such a strength that it can be registered by human beings. Based on historical data up to 1980, it is reasonable to assume that the absolute strength of such earthquakes is normally distributed with expectation $\mu = 4.2$ and standard deviation $\sigma = 0.4$.

a) In order for substantial damages to occur, it is assumed that the absolute strength measured on the Richter scale must exceed 5.4. What is the probasbility that the absolute strength of an arbitrary earthquake registered offshore outside Mid-Norway shall exceed 5.4?

Determine a value k such that the probability that the absolute strength of such earthquakes shall exceed k is only 5%.

Since 1980 it is registered 9 such earthquakes of the given type with the following absolute strength measured on the Richter scale: 4.8, 3.5, 3.8, 5.3, 3.8, 3.5, 3.5, 3.5, 4.0. These values shall be considered as realizations of 9 independent identical normally distributed variables $Y_1, Y_2, \ldots, Y_9$. For these numbers

$$\bar{y} = 3.967 \quad \text{and} \quad \sum_{i=1}^{9} (y_i - \bar{y})^2 = 3.40.$$

The variance is now assumed to be unknown.

b) It is of interest to find out if expected absolute strength has increased as a consequence of the Norwegian oil activity. State this as a hypothesis test.

Perform the test. What is the conclusion at a 5% level of significance.

It is also of interest to find out if the frequency of earthquakes has changed as a consequence of the Norwegian oil activity. We shall now consider all types of earthquakes registered all over the country and assume the number of earthquakes, X in a certain time interval $[0, t]$ measured in years is Poisson distributed with parameter, λt, i.e.

$$\mathrm{P}(X = x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}.$$

c) Assume in **c**) that $\lambda = 5$. What is the probability that it will be registered exactly 8 earthquakes in a year?

What is the probability that it will be registered 5 earthquakes or more in half a year?

What is the probability that the largest number of earthquakes in a year for a time period of 10 years shall exceed 10?

Let T be the period of time from registration is started until the first earthquake occur.

d) Show that T is exponentially distributed with parameter $1/\lambda$.

Also show that $2\lambda T$ is chisquare distributed with 2 degrees of freedom.

e) The time period between two following earthquakes has the same distribution as T. Let $T_1, T_2, \ldots, T_n$ be times between two following earthquakes and let be the corresponding registered times. Find the maximum likelihood estimator for λ, $\widehat{\lambda}$, based upon the random sample.

Show how it is possible to construct a 95% confidence interval for λ based upon $2\lambda n/\widehat{\lambda}$.

What numerical values do you get for the confidence interval if the average of 10 time periods between succeeding earthquakes is 3 months. Is there any reason to conclude that $\lambda \neq 5$? Explain your answer.

Problem 3 Automatized laboratory

In a laboratory it is of interest to evaluate the relationship between two variables Y and x. The apparatus is put up in such a way that a value for x can be set and thereafter Y can be measured. It is decided to use the following model for the relationship between the two variables

$$Y = \alpha + \beta x + \varepsilon$$

where α and β are two independent unknown coefficients and ε is a normal distributed random variable with expectation equal to 0 and unknown variance σ^2. Let $x_1, x_2, \ldots, x_n$ be n values of the variable x and $y_1, y_2, \ldots, y_n$ the corresponding values measured for Y. These can be considered as realizations of independent variables $Y_1, Y_2, \ldots, Y_n$.. The least squares estimators A and B, for the coefficients α and β are then given by

$$A = \frac{1}{n}\sum_{i=1}^{n} Y_i - B\bar{x}$$

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

where

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

a) Show that the estimators A and B have expectations equal to α and β, respectively.

Use that the covariance between $\bar{Y} = (1/n)\sum_{i=1}^{n} Y_i$ and B is 0 and deduce the variance of the estimators A and B.

What are the probability distributions of the estimators A and B? Explain your answer.

The laboratory experiments are very work consuming, but the apparatus has been made automatic such that the experiments can be performed for equidistant distances of x. Consider two series of measurements with $n = 10$

Series 1:	$x_1 = 1, x_2 = 2, \ldots, x_{10} = 10$
Series 2:	$x_1 = 2, x_2 = 4, \ldots, x_{20} = 20$

The goal of the experiment was to predict Y_0 for $x_0 = 5.5$. The following predictor is used: $\widehat{Y}_0 = A + Bx_0$.

b) Deduce the variance of $Y_0 - \widehat{Y}_0$.

What series of measurements should be used to predict Y_0 as good as possible for $x_0 = 5.5$? Explain your answer and make comments on it.