Norwegian University
of Science and Technology
Department of Mathematical Sciences

English

Contact during the exam:

Arvid Næss	99 53 83 50
Jarle Tufto	99 70 55 19
Ola Diserud	93 21 88 23

EXAM IN TMA4245 STATISTICS

August, xx. 2010

Time: 09:00–13:00

Permitted aids: *Tabeller og formler i statistikk*, Tapir Forlag
K. Rottmann: *Matematisk formelsamling*
Calculator HP30S
Yellow, stamped A5-sheet with handwritten notes.

Problem 1

Let X be a stochastic variable with probability density (pdf)

$$\begin{aligned} f_X(x) &= \frac{1}{\sqrt{2\pi}\tau x} e^{-\frac{1}{2\tau^2}(\ln x - \nu)^2}, \quad x > 0\,, \\ &= 0\,, \qquad x \le 0\,, \end{aligned} \tag{1}$$

where $\tau > 0$ and ν are real numbers, and $\ln x$ denotes the natural logarithm of x.

Consider n independent soil samples taken from a given area, each having a weight of one kilogram, and let $x_1, \ldots, x_n$ denote the measured content of nickel (in milligrams) in each sample. Suppose that these measurement can be regarded as realisations of independent and identically distributed random variables $X_1, \ldots, X_n$ with probability density given by (1).

a) Show that if X is a random variable with pdf given by equation (1), then $Y = \ln X$ is normally distributed with expectation $\nu = \mathrm{E}[Y]$ and variance $\tau^2 = \mathrm{Var}[Y]$.

b) Show that $F_X(x) = \text{Prob}(X \leq x) = \Phi\left(\frac{\ln x - \nu}{\tau}\right)$, where Φ denotes the the cumulative probability density of a random variable $Z \sim N(0,1)$.

Thus, if we introduce $Y_j = \ln X_j$, $j = 1, \ldots, n$, then $Y_1, \ldots, Y_n$ are indepenent and identically distributed stochastic variables with $Y_j \sim N(\nu, \tau^2)$, $j = 1, \ldots, n$.

c) Assume that $\nu = 1.0$ and $\tau = 0.8$. Find $\text{Prob}(X_1 \leq 1.0)$ and $\text{Prob}(X_1 \cdot X_2 \leq 1.0)$.

d) Assume that $\nu = 1.0$, $\tau = 0.8$ and $n = 5$. Find the probability that the measured content of nickel in at least 4 out of 5 samples is less than 2.72 mg.

e) Show that

$$\mu = \text{E}[X] = e^{\nu + \tau^2/2} \quad , \quad \sigma^2 = \text{Var}[X] = e^{2\nu}\left(e^{2\tau^2} - e^{\tau^2}\right)$$

Hint: Use the substitution $t = (\ln x - \nu)/\tau$.

f) Suppose that we have measured the content of nickel $x_1, \ldots, x_n$ in n soil samples and that we want an estimate of μ based on these measurements. A possible estimator is

$$\hat{\mu} = \frac{1}{n}\sum_{j=1}^{n} X_j = \overline{X}.$$

Briefly discuss the properties of this estimator and compute the corresponding estimate of μ using the data given below.

g) An alternative estimator of μ can be constructed by first considering the maximum likelihood estimators (MLE) $\hat{\nu}$ and $\hat{\tau}^2$ of ν og τ^2, and then, in turn, making use of the result in e). Find the estimators $\hat{\nu}$ and $\hat{\tau}^2$. What are the properties of these estimators? Find a corresponding estimator μ^* of μ based on $\hat{\nu}$ and $\hat{\tau}^2$. Using μ^*, compute an estimate of μ based on the data given below.

h) Suppose that τ^2 is known and equal to τ_0^2. Based on the random $x_1, \ldots, x_n$ we wish to test

$$H_0 \ : \mu \leq \mu_0$$

againts

$$H_1 \ : \mu > \mu_0$$

where μ_0 is a given quantity.

Use the result in e) to express H_0 and H_1 in terms of ν, and show that the above hypothesis test is equivalent to testing

$$H_0' : \nu \leq \nu_0$$

against

$$H_1' : \nu > \nu_0$$

where ν_0 is a known quantity.

Use this to construct a reasonable test of H_0 against H_1 using α as the level of significance.

i) Derive an expression for the power of the test in h) under the alternative hypothesis $\mu = \gamma\mu_0$ ($\gamma > 1$) when $\alpha = 0.05$ and $\tau_0^2 = 0.36$.

What is the minimum sample size n if the probability of rejecting H_0 is 0.9 when $\mu = 1.5\mu_0$?

j) Suppose that τ^2 is known and equal to τ_0^2. Derive a $100(1-\alpha)\%$ confidence interval for ν, og use the result to find a corresponding confidence interval for μ.

Compute the confidence interval for μ when $n = 10$, $\alpha = 0.05$, $\tau_0^2 = 0.36$ using the data below.

x_j	57	38	150	29	65	44	36	24	51	131
$y_j = \ln x_j$	4.04	3.64	5.01	3.37	4.17	3.78	3.58	3.18	3.93	4.48

Table 1: Observations

$$\sum_{j=1}^{10} y_j = 39.58 \qquad \sum_{j=1}^{10} (y_j - \overline{y})^2 = 3.2360$$