Norwegian University
of Science and Technology
Department of Mathematical Sciences

English

Contact during exam:

Arvid Næss	73 59 70 53/ 99 53 83 50
Jarle Tufto	99 70 55 19
Ola Diserud	93 21 88 23

EXAM IN TMA4245 STATISTICS

3. June 2010
Time: 09:00–13:00
Grades will announced on June 24

Aids: *Tabeller og formler i statistikk*, Tapir Forlag
K. Rottmann: *Matematisk formelsamling*
Calculator HP30S
One stamped, yellow paper (A5) with your personal handwritten notes.

Problem 1 A rare gene variant a predisposes for a given illness. In a population, the relative frequency of individuals that carries two copies of this gene variant (individuals of genotype aa) is 0.0001, the frequency of individuals that carries one copy (genotype Aa) is 0.0198, and the frequency of individuals that do not carry this rare gene variant (genotype AA) is 0.9801. Assume further that the probabilities for the illness to be expressed among persons with genotypes aa, aA and AA are 0.6, 0.02 and 0.01 respectively.

a) Find the probability for the illness to be expressed in a randomly chosen individual in the population.

b) What are the probabilities that an individual is of genotype aa, Aa and AA respectively, given that the illness has been expressed?

Problem 2 A factory produces a special type of machine components. The time from a component starts to be used until it malfunctions for the first time is called the components lifetime. Experience has shown that the lifetime T, measured in weeks, can be modeled as a continuous random variable with probability distribution

$$\begin{aligned} f_T(t) &= 2\lambda t e^{-\lambda t^2}\,, \quad t \geq 0\,, \\ &= 0\,, \qquad \text{otherwise}\,, \end{aligned} \tag{1}$$

where $\lambda > 0$ is an unknown parameter.

a) Find the cumulative distribution function $F_T(t)$ for T, and calculate $P(20 < T \leq 30)$ when $\lambda = 1.5 \cdot 10^{-3}$.

b) The parameter λ shall be estimated from the lifetimes $T_1, \ldots, T_n$ for $n > 2$ randomly chosen components. $T_1, \ldots, T_n$ are assumed to be independently, identically distributed with probability distribution $f_T(t)$. Show that the maximum likelihood estimator (MLE) for λ becomes

$$\Lambda^* = \frac{n}{\sum_{i=1}^{n} T_i^2}\,. \tag{2}$$

c) Let X be χ^2-distributed with $2n$ degrees of freedom ($n > 2$). Show that

$$\mathrm{E}(X^{-1}) = \frac{1}{2(n-1)} \quad \text{and} \quad \mathrm{E}(X^{-2}) = \frac{1}{4(n-1)(n-2)}\,. \tag{3}$$

Show also that $Y = 2\lambda T^2$ is χ^2-distributed with 2 degrees of freedom (remember that $T \geq 0$). Then use this result together with equation (3) to check if Λ^* is unbiased. If necessary, correct the estimator so that it becomes unbiased. Find this estimator's variance. (Tip: Some helpful results can be found in 'Tabeller og formler i statistikk'.)

d) Derive a $100(1-\alpha)\%$ confidence interval for λ. Find the numerical solution for the interval when $\alpha = 0.05$, $n = 5$ and the observed values are

23.63 35.97 18.65 18.18 11.59

e) One day an error is detected in the production process. The components are tested by choosing 5 components randomly from this day's production, and the lifetimes for these components are found by accelerated lifetime-testing. These observations are used to test

$$H_0\ : \lambda \leq \lambda_0 = 1.5 \cdot 10^{-3}$$

vs.

$$H_1\ : \lambda > \lambda_0 = 1.5 \cdot 10^{-3}$$

Show that $2\lambda_0 \sum_{i=1}^{n} T_i^2$ can be used as a test statistic, and find the critical area for a test with significance level α. Conclude for $\alpha = 0.05$ when the observations are

$$12.06 \quad 18.02 \quad 19.86 \quad 16.60 \quad 9.36$$

The components are packed in boxes with 5 components in each box. The factory guarantees that all components in a box will have a lifetime of at least a weeks. The factory will get a complaint on a box if one or more of the 5 components in a box has lifetime less than a weeks.

f) If $\lambda = 1.5 \cdot 10^{-3}$, how large can a at most be for the probability of a complaint to be no more than 0.05? You can assume independent lifetimes.

g) Let a and λ be as in point f), and assume that 1000 boxes are sold in a given period. Let U be the number of boxes with complaints. Which assumptions must be accepted for U to be binomially distributed? Assume that these assumptions are fulfilled and find $P(U \leq 60)$ by the normal approximation.