NTNU – Trondheim
Norwegian University of
Science and Technology

Department of Mathematical Sciences

Examination paper for **ST0103 Statistics with Applications**

Academic contact during examination: Øyvind Bakke

Phone: 73 59 81 26, 990 41 673

Examination date: August 2016

Examination time (from–to): 9:00–13:00

Permitted examination support material: Yellow A4 sheet with your own handwritten notes, specific calculator (Casio fx-82ES Plus, Citizen SR-270X, Citizen SR-270X College or HP 30s), *Tabeller og formler i statistikk* (Tapir forlag or Fagbokforlaget), *Matematisk formelsamling* (K. Rottmann)

Other information:
In the grading, each of the ten points counts equally.

You should demonstrate how you arrive at your answers (e.g. by including intermediate answers or by referring to theory or examples from the reading list).

Language: English

Number of pages: 2

Number of pages enclosed: 0

Checked by:

Date Signature

Problem 1

Alum shale is an acid-forming shale rock. There are numerous problems associated with alum shale, and in connection with construction and digging projects, precautions must be taken when alum shale is handled.

When alum shale is scratched with a knife, the scratch line is always black. But also other shales can give a black scratch line. In an area, 80% of the occurrences of shales are other types than alum shale. The probability that a shale, which is not alum shale, will give a black scratch line is 0.3.

A specimen of shale from this area gives a black scratch line.

a) What is the probability that the specimen consists of alum shale?

A possible problem associated with alum shale, is a high uranium content, which produces radioactivity. Assume that the uranium content measured in mg/kg in a randomly chosen specimen of alum shale from this area is normally distributed with mean (expected value) μ. Ten independent specimens are taken, and the mean uranium content of the 10 specimens is 90.8 and the sample standard deviation 1.4.

b) Test the null hypothesis $\mu \leq 90$ againtst the alternative hypothesis $\mu > 90$. Use significance level 0.05.

Problem 2

Carbon has two stable isotopes that occur naturally, ^{12}C and ^{13}C. On Earth, 98.9% of the carbon atoms are ^{12}C and 1.1% of the carbon atoms ^{13}C. Assume that the number of ^{13}C atoms, X, in a molecule containing n carbon atoms is binomially distributed with parameters n and $p = 0.011$.

The molecule of a biopolymer contains 100 carbon atoms.

a) Which assumptions must be satisfied for X to be binomially distributed? Find the probability that the biopolymer molecule contains 3 or more ^{13}C atoms.

b) Find the probability that a Poisson distributed variable having expected value 1.1 is less than or equal to 2. Use this to find an approximate value of the probability from (a).

A poliovirus has molecular formula $C_{332652}H_{492388}N_{98245}O_{131196}P_{7501}S_{2340}$, and thus contains 332 652 carbon atoms.

c) What is the expected value and the standard deviation of the number of ^{13}C atoms in the virus? Find an approximate probability that this number is 3700 or more.

Problem 3

The duration T measured in days (not necessarily an integer) of the pupal stage of an African butterfly species is exponentially distributed, that is, the probability density is given by $f(t) = \frac{1}{\mu}e^{-t/\mu}$, where $t > 0$, and $\mu > 0$ is a parameter. (In the textbook, the probability density is of form $\lambda e^{-\lambda t}$, where $\lambda = 1/\mu$.)

a) Show, by calculation, that the cumulative distribution function of T is given by $P(T \leq t) = 1 - e^{-t/\mu},\ \ t > 0$.

b) Assume (only here) that $\mu = 20$. Find $P(T \geq 20)$. Find the conditional probability $P(T \geq 30 \mid T \geq 10)$.

Let $T_1, T_2, \ldots, T_n$ be independent observations of the duration of the pupal stage.

c) Show that the maximum likelihood estimator of μ is the mean, $\bar{T}$, of the observations. Find the expected value and the variance of the estimator.

d) Show that $\frac{2}{\mu}T$ is chi-squared distributed with 2 degrees of freedom. (Hint: You can use that $\Gamma(1) = 1$ in the formula, found in the formula booklet, for the probability density of a chi-squared variable.)

More generally, it can be shown that $\frac{2}{\mu}\sum_{i=1}^{n} T_i$ is chi-squared distributed with $2n$ degrees of freedom.

e) Show that

$$\left[\frac{2}{\chi_{\alpha/2}}\sum_{i=1}^{n} T_i, \frac{2}{\chi_{1-\alpha/2}}\sum_{i=1}^{n} T_i\right]$$

is a $100(1-\alpha)\%$ confidence interval for μ, where χ_α is the number such that $P(Y \geq \chi_\alpha) = \alpha$ when Y is chi-squared distributed with $2n$ degrees of freedom. What is the 95% confidence interval if $n = 40$ and $\bar{T} = 28.3$?