Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

English

Contact during exam:

Steinar Engen	73 59 17 47/ 90 63 50 53
Arvid Næss	73 59 70 53/ 99 53 83 50
Turid Follestad	73 59 35 37

EXAM IN COURSE SIF5062/SIF5506 STATISTICS

Saturday May 24th 2003
Time: 09:00–14:00

Permitted aids: *Tabeller og formler i statistikk*, Tapir Forlag
K. Rottmann: *Matematisk formelsamling*
Calculator HP30S

The examination results are due June 28th 2003

Problem 1

Let X be the height of a randomly chosen 6 year old girl. We assume that the height is normally distributed with expectation $\mathrm{E}(X) = 115$ cm and standard deviation $\mathrm{SD}(X) = 5$ cm.

a) Compute the probabilities

$$P(X \leq 120) \text{ and } P(120 < X \leq 125).$$

Let the following two events be defined:

$A: \quad X \leq 120$
$B: \quad X > 125$

Are A and B disjoint? Are A and B independent? Justify the answers.

We imagine going to the year 2010, and are interested in the names among girls in the first year of school, born in the year 2004. The name statistics shows that 2% of the girls born this year were given the name Maud.

b) Let Z be the number of girls named Maud in a randomly selected class with n girls. Explain why it is reasonable to assume that Z is binomially distributed with parameters n and p, where $p = 0.02$.

In the rest of this problem we assume that Z is binomially distributed with $n = 15$ and $p = 0.02$. Let the events C and D be defined by

C : at least one of the girls is named Maud
D : exactly two of the girls are named Maud

Compute the probabilities $P(C)$ and $P(D \mid C)$.

Assume that there are totally 25 pupils in the class. What is the probability that a randomly chosen pupil among all pupils in the class is named Maud?

Problem 2

The producer of a given car model claims that this model can be expected to drive at least 16 km per liter gasoline on a freeway. The consumers organization CO tests this postulate by driving a random sample of cars of this type a suitable distance on a representative freeway and measuring the amount of gasoline used.

Based on experience from earlier experiments of the same type, the CO assumes that the consumption of gasoline for a randomly selected car of this type can be modelled approximately by a normally distributed variable X with mean μ and variance σ^2, that is, $X \sim N(\mu, \sigma^2)$. The mean μ as well as the standard deviation σ are initially unknown quantities.

For practical reasons CO restricted the sample size to $n = 20$ cars. When the experiment was finished, all measurements were analyzed and resulted in a sample mean $\bar{x} = 15.56$ and a sample standard deviation $s = 0.94$.

a) Formulate a hypothesis test for this experiment. Choose the producer's postulate as the null hypothesis. Which test statistic will you choose to check the hypothesis? Give a short justification for your choice of statistic. In relation to a chosen significance level $\alpha = 0.05$, will you accept the producers postulate?

b) What is the P-value (significance probability) for the hypothesis test in a) for the given observations?

Which approximation can you do to make the test statistic normally distributed? Which P-value do you find using this approximation?

c) Determine the power of the test for the alternative hypothesis $\mathrm{H}_1' : \mu = 15.5$ for significance level $\alpha = 0.05$ by using the same normal approximation as in b). Give a suggestion for how the power of the test can be increased.

Problem 3

An apparatus contains k equal components and functions only if all these are functioning. The components' lifetimes $T_1, T_2, \ldots T_k$ are independent and exponentially distributed random variables with parameter β (> 0), that is, the probability density is

$$f(t) = \begin{cases} \dfrac{1}{\beta} e^{-t/\beta} & \text{for} \quad t \geq 0, \\ 0 & \text{for} \quad t < 0. \end{cases}$$

a) Find the cumulative distribution function of the lifetime of one component. Compute $P(T_1 < 3)$ and $P(2 < T_1 < 4)$ when $\beta = 5$.

b) Let X denote the lifetime of the apparatus. Show that X is exponentially distributed with parameter β/k. What is the expected lifetime of the apparatus when $k = 4$ and $\beta = 5$?

The company has made several versions of the apparatus with different number of components. The apparatus functions better with many components, but will at the same time have smaller expected lifetime. Let $X_1, X_2, \ldots, X_n$ be the lifetimes of n apparatuses with $k_1, k_2, \ldots, k_n$ components, respectively. Two estimators for β based on the lifetimes of the apparatuses are proposed,

$$\hat{\beta} = \frac{1}{n} \sum_{i=1}^{n} X_i k_i$$

and

$$\tilde{\beta} = \frac{\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} k_i^{-1}}.$$

c) Find the mean and variance of both estimators.

d) Show that one of the estimators in (b) is the maximum likelihood estimator (MLE) and show that this always has a variance that is smaller than or equal to the variance of the other estimator.

Hint: Write $r_i = 1/k_i$ and apply the result $\frac{1}{n} \sum r_i^2 - (\frac{1}{n} \sum r_i)^2 \geq 0$.

e) Show that $2k_iX_i/\beta$ is χ^2-distributed with 2 degrees of freedom and that $2n\hat{\beta}/\beta$ is χ^2-distributed. Use this to derive a $(1-\alpha)\cdot 100\%$ confidence interval for β. Find this interval when $\alpha = 0.05$, $n = 8$ and $\hat{\beta} = 8.3$.