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

English

Contact during exam:
Nikolai Ushakov 45128897

ST1201/ST6201 STATISTICAL METHODS

Saturday 8 June 2013
Time: 09:00–13:00

Permitted material:
Statistiske tabeller og formler, Tapir forlag.
K.Rottman. Matematisk formelsamling.
Yellow A4-sheet of paper with handwritten notes.
Calculator.

Results: 29 June 2013

Oppgave 1

77 goals were scored on the 2012 European Football Championship. The table shows the number of games where 0,1,2 etc. goals were scored.

The number of goals in a game	The number of games
0	2
1	6
2	10
3	6
4	3
5	3
6	1
7+	0

a) Test the hypothesis that the number of goals, scored in a game, has the Poisson distribution with parameter $\lambda = 1$. The significance level is 0.05.

b) Test the hypothesis that the number of goals, scored in a game, has a Poisson distribution (with unknown parameter λ). The significance level is 0.05.

Oppgave 2

Let $X_1, ..., X_n$ be a random sample (independent identically distributed random variables) from an exponential distribution with unknown parameter λ, i.e. from the distribution with the probability density

$$f(x) = \lambda e^{-\lambda x}, \;\; x \geq 0, \;\; \lambda > 0.$$

The null hypothesis $H_0 : \lambda = 1$ is tested versus the alternative $H_1 : \lambda > 1$. There are two tests. Both are based on the test statistic

$$T = \min\{X_1, ..., X_n\},$$

but the first test rejects H_0 for large values of T while the second one rejects H_0 for small values of T. In other words

Test 1: if $T \geq c_1$, then H_0 is rejected.

Test 2: if $T \leq c_2$, then H_0 is rejected.

The significance level is α.

a) Find c_1 and c_2.

b) Find the power function of each test.

c) Which test is better? Why?

d) Even the best of these two tests is bad. Explain why.

Oppgave 3

Let Y be measured percentage of fat in a certain type of sausages. A laboratory has measured the fat percentage in 15 sausages and results $y_1, y_2, ..., y_{15}$ are supposed to be realisations of independent continuous random variables with a symmetric (about the unknown expectation) distribution. The results are

$$19.2, 27.6, 25.6, 32.2, 17.7, 20.5, 23.9, 20.2, 24.2, 26.1, 32.0, 24.8, 28.9, 16.2, 18.7.$$

Fat percentage 20.0 is considered as normal. We wish to test the hypothesis that the fat percentage is normal (i.e. equals 20.0) versus the alternative that it is greater than 20.0. The significance level is 0.05.

a) Test the hypothesis using the large-sample sign test.

b) Test the hypothesis using the large-sample Wilcoxon signed rank test.

c) Suppose that in addition to the conditions above it is known that the distribution of the fat percentage is approximately normal. Test the hypothesis using the $t-$test (for simple calculations you can use that $\sum_{i=1}^{15} y_i = 357.8$ and $\sum_{i=1}^{15} y_i^2 = 8888.62$).

Oppgave 4

The following table is an ANOVA table in which some entries are lost (stars).

Source	df	SS	MS	F
Treatment	3	$\star$	$\star$	1.6
Error	$\star$	$\star$	5.1	
Total	43	$\star$		

a) Find lost values and fill in the ANOVA table. Show how you calculate values where there are $\star$ in the table.

b) Test the hypothesis that

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4.$$

The significance level $\alpha = 0.05$.