Eindhoven University of Technology
Faculty of Mathematics and Computer Science

Examination "Mathematical Statistics" (2S990)
Friday 26 November 2004, 9:00-12:00

You are allowed to use the "Statistisch Compendium" and a pocket calculator, but not the textbook. The solutions to the exercises must be clearly formulated, carefully written and explained. All exercises have equal weight.

---

1. Let $X_1,\ldots,X_9$ be an i.i.d. sample drawn from $N(\mu,\sigma^2)$, with μ and σ^2 unknown. The outcome of the sample is:

   220, 205, 192, 198, 201, 207, 195, 201, 204

   (a) Construct a 95% confidence interval for μ.

   (b) Construct a 95% confidence interval for σ.

2. Let Y_1, Y_2 be an i.i.d. sample drawn from UNIF$(-\alpha,\alpha)$, with α unknown

   (a) Is $\max\{|Y_1|,|Y_2|\}$ a sufficient statistic for α?

   (b) Is $|Y_1|$ a sufficient statistic for α?

3. Let $Z_1,\ldots,Z_{20}$ be an i.i.d. sample drawn from EXP(θ), with θ unknown (density $\frac{1}{\theta}e^{-x/\theta}$, $x \geq 0$).

   (a) Is $1/\bar{Z}$ with $\bar{Z} = \frac{1}{20}\sum_{i=1}^{20} Z_i$ an unbiased estimator of $1/\theta$? If so, prove that the bias equals zero. Otherwise, find an expression for the bias.

   (b) Is $-\bar{Z}$ a uniformly minimum variance unbiased estimator of $-\theta$?

4. Twice repeatedly throw a coin that has probability p to produce heads, with p unknown. Let T_1, T_2 be the respective number of throws until the first head comes up. Consider the hypothesis test $H_0 : p = 1/2$ against $H_a : p = 3/4$.

   (a) Show that the critical region

   $$G(c) = \{(t_1,t_2) \in \mathbb{N}^2 : 2 \leq t_1 + t_2 \leq c\}$$

has maximal power among all critical regions with the same significance level $\alpha = \alpha(c)$.

(b) Compute α and $1 - \beta$ for $G(3)$.

(c) Compute α and $1 - \beta$ for the critical region

$$F = \{(1,1),(2,1),(3,1),(1,3)\}.$$

(d) Is $G(3)$ more powerful than F?

5. The Dutch population is split into four categories according to smoking behavior. Nationwide the following percentages apply:

Category	1	2	3	4
Percentage	46	11	27	16

To test whether Dutch students exhibit the same smoking behavior, a sample of 500 students is drawn resulting in the following data:

Category	1	2	3	4
Frequency	206	41	155	98

Carry out a goodness-of-fit for the hypothesis test

H_0 : same behavior
H_a : different behavior

at significance level $\alpha = 0.1$.