1. a. i 5 b. 5 2. a 10 3. a 15
ii 5 b 5 b 5
iii 5 c i 5 d 5
iv 5 ii 5
iii 5

4 November 2011, 9:00 – 12:00

Rijksuniversiteit Groningen
Statistiek

Tentamen

RULES FOR THE EXAM:

- The use of a normal, non-graphical calculator is permitted.
- This is a CLOSED-BOOK exam.
- At the end of the exam you can find a normal table and a chi-squared table.

1. **Point estimation.** Let $X_1, \ldots, X_n$ be a sample of independent, identically distributed random variables, with density

$$f_\theta(x) = \begin{cases} \theta x + \frac{1}{2} & -1 \le x \le 1 \\ 0 & \text{elsewhere} \end{cases}$$

$\int_{-1}^{1} \theta x + \frac{1}{2}$

for $\theta \in (-\frac{1}{2}, \frac{1}{2})$.

(a) Let

$$\hat{\theta}_n = 3\bar{X}/2$$

be an estimator of θ, where $\bar{X}$ is the sample mean.

i. Determine whether $\hat{\theta}_n$ is unbiased. [5 Marks]

ii. Determine whether $\hat{\theta}_n$ is consistent. [5 Marks]

iii. Show that $\hat{\theta}_n$ is not sufficient. [Hint: use the original definition of sufficiency, $p(x|\hat{\theta})$ does not depend on θ, and focus on $n = 2$]. [5 Marks]

iv. Determine whether $\hat{\theta}_n$ is efficient. [5 Marks]

(b) We observe the following 3 data values

observation i	1	2	3
value x_i	-0.5	0	0.75

Determine the value of the maximum likelihood estimator for these data and compare this to the value of $\hat{\theta}_3$ defined above. [5 Marks]

2. **Confidence intervals and Likelihood ratio test.** A study consists of 38 individual experiments, in which counts are measured. Assume $X_1, \dots, X_{38}$ are independent and identically distributed Poisson(μ), i.e.,

$$p_X(k) = e^{-\mu}\frac{\mu^k}{k!}, \quad k = 0, 1, 2, \dots.$$

$E[X] = \mu \quad Var(X) = \mu$

   (a) Let the observed average $\hat{\mu} = \bar{X}$ be the unbiased estimator of μ. Using the normal approximation of $\bar{X}$ (Central Limit Theorem), determine a 95% confidence interval for μ, when we observe

$$\sum_{i=1}^{38} x_i = 95.$$

   [10 Marks.]

   (b) Using the 95% confidence interval for the parameter μ, determine a 95% confidence interval for the probability of observing no counts in one of the individual experiments [5 Marks].

   (c) The experimenter knows that on the basis of standard theory the population mean count $\mu = 2$. He would like to test this hypothesis on the basis of his 38 observations.

      i. Determine the appropriate null and alternative hypotheses. [5 Marks]

      ii. Determine the general expression and the value of log likelihood ratio statistic. [5 Marks]

      iii. Use the likelihood ratio test to determine whether or not we can reject the null hypothesis at the 5% level. [5 points]

3. **Optimal testing.** An Atomic Energy Agency is worried that a particular nuclear plant is not managed carefully. It has decided to monitor the plant for the number of days T between successive "minor problems". It performs a single observation T, assuming $T \sim$ Geometric(p), i.e.

$$p_T(k) = (1-p)^{k-1}p, \quad k = 1, 2, \dots.$$

The management of the plant claims that $p = 0.001$. The law states that the plant should be closed down if $p = 0.003$. The Atomic Energy

$P_T(k) = (1-p)^{k-1}\, p$

Agency decides to test the following hypotheses:

$$H_0: \quad p = 0.001$$
$$H_1: \quad p = 0.003$$

The optimal critical region is given in the form as $\{t \mid L_0(t)/L_1(t) = k\}$.

(a) Determine the cumulative distribution function of T given a general p. [Hint: $\sum_{j=a}^{b} q^j = \frac{q^a - q^{b+1}}{1-q}$ for $|q| < 1$.]

(b) Determine the critical region and the value k. [15 Marks]

(c) What is the power of this test? [5 Marks]

(d) The next minor incident happens after 36 days. Should the Atomic Energy Agency close down this plant? Yes/No answer is not enough. Give a formally correct argument. [5 Marks]

$P\left\{\frac{L_0(t)}{L_1(t)} \le k \mid p = 0.001\right\} = \alpha$

Statistical tables which may be used in the calculations.

$\nu \setminus \alpha$	0.995	0.99	0.975	0.95	0.05	0.025	0.01	0.005
1	0.000	0.000	0.001	0.004	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	11.070	12.833	15.086	16.750
10	2.156	2.558	3.247	3.940	18.307	20.483	23.209	25.188

Table 1: Values of $\chi^2_{\alpha,\nu}$: the entries in the table correspond to values of x, such that $P(\chi^2_\nu > x) = \alpha$, where χ^2_ν correspond to a chi-squared distributed variable with ν degrees of freedom.

$L_0(t) = (1-0.001)^{t-1} \cdot 0.001$

$L_1(t) = (1-0.003)^{t-1} \cdot 0.003$

$0.999^{T-1} \cdot 0.001$

$0.997^{T-1} \cdot 0.001$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.000	0.004	0.008	0.012	0.016	0.020	0.024	0.028	0.032	0.036
0.1	0.040	0.044	0.048	0.052	0.056	0.060	0.064	0.067	0.071	0.075
0.2	0.079	0.083	0.087	0.091	0.095	0.099	0.103	0.106	0.110	0.114
0.3	0.118	0.122	0.126	0.129	0.133	0.137	0.141	0.144	0.148	0.152
0.4	0.155	0.159	0.163	0.166	0.170	0.174	0.177	0.181	0.184	0.188
0.5	0.191	0.195	0.198	0.202	0.205	0.209	0.212	0.216	0.219	0.222
0.6	0.226	0.229	0.232	0.236	0.239	0.242	0.245	0.249	0.252	0.255
0.7	0.258	0.261	0.264	0.267	0.270	0.273	0.276	0.279	0.282	0.285
0.8	0.288	0.291	0.294	0.297	0.300	0.302	0.305	0.308	0.311	0.313
0.9	0.316	0.319	0.321	0.324	0.326	0.329	0.331	0.334	0.336	0.339
1.0	0.341	0.344	0.346	0.348	0.351	0.353	0.355	0.358	0.360	0.362
1.1	0.364	0.367	0.369	0.371	0.373	0.375	0.377	0.379	0.381	0.383
1.2	0.385	0.387	0.389	0.391	0.393	0.394	0.396	0.398	0.400	0.401
1.3	0.403	0.405	0.407	0.408	0.410	0.411	0.413	0.415	0.416	0.418
1.4	0.419	0.421	0.422	0.424	0.425	0.426	0.428	0.429	0.431	0.432
1.5	0.433	0.434	0.436	0.437	0.438	0.439	0.441	0.442	0.443	0.444
1.6	0.445	0.446	0.447	0.448	0.449	0.451	0.452	0.453	0.454	0.454
1.7	0.455	0.456	0.457	0.458	0.459	0.460	0.461	0.462	0.462	0.463
1.8	0.464	0.465	0.466	0.466	0.467	0.468	0.469	0.469	0.470	0.471
1.9	0.471	0.472	0.473	0.473	0.474	0.474	0.475	0.476	0.476	0.477
2.0	0.477	0.478	0.478	0.479	0.479	0.480	0.480	0.481	0.481	0.482
2.1	0.482	0.483	0.483	0.483	0.484	0.484	0.485	0.485	0.485	0.486
2.2	0.486	0.486	0.487	0.487	0.487	0.488	0.488	0.488	0.489	0.489
2.3	0.489	0.490	0.490	0.490	0.490	0.491	0.491	0.491	0.491	0.492
2.4	0.492	0.492	0.492	0.492	0.493	0.493	0.493	0.493	0.493	0.494
2.5	0.494	0.494	0.494	0.494	0.494	0.495	0.495	0.495	0.495	0.495
2.6	0.495	0.495	0.496	0.496	0.496	0.496	0.496	0.496	0.496	0.496
2.7	0.497	0.497	0.497	0.497	0.497	0.497	0.497	0.497	0.497	0.497
2.8	0.497	0.498	0.498	0.498	0.498	0.498	0.498	0.498	0.498	0.498
2.9	0.498	0.498	0.498	0.498	0.498	0.498	0.498	0.499	0.499	0.499
3.0	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499	0.499

Table 2: Standard Normal Distribution. This means that values in the table correspond to probabilities $P(0 < Z \leq z)$, where Z is a standard normal distributed variable.