4 November 2010, 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 (1).** Let $X_1, \ldots, X_n$ be a sample of independent, identically distributed Uniform$(\theta, 1)$ random variables, with density

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

   (a) Determine the maximum likelihood estimator (MLE) of θ. [10 Marks]

   (b) Let $\hat{\theta}_n$ be the MLE of θ,

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

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

      iii. Determine whether $\hat{\theta}_n$ is sufficient. [5 Marks]

   (c) Why doesn't the Cramer-Rao lower bound apply to unbiased estimates of θ for this distribution? [5 Marks]

2. **Point estimation (2).** Let $X_1, \ldots, X_n$ be a sample of independent, identically distributed Poisson(θ) random variables, with density

$$f_\theta(k) = e^{-\theta}\frac{\theta^k}{k!}, \quad k = 0, 1, 2, \ldots.$$

   (a) Determine the Cramer-Rao lower bound for the variance of an unbiased estimate of θ. [5 Marks]

   (b) Determine the MLE $\hat{\theta}$ of θ. [5 Marks]

   (c) Check that $\hat{\theta}$ is unbiased an whether it attains the Cramer-Rao lower bound. [10 Marks]

   (d) In fact, we are told that $n = 4$ and that the data are given as

$$x_1 = 2, \quad x_2 = 5, \quad x_3 = 4, \quad x_4 = 5.$$

      i. Determine an approximate 95% confidence interval based on the asymptotic normality of $\hat{\theta}$. [5 Marks]

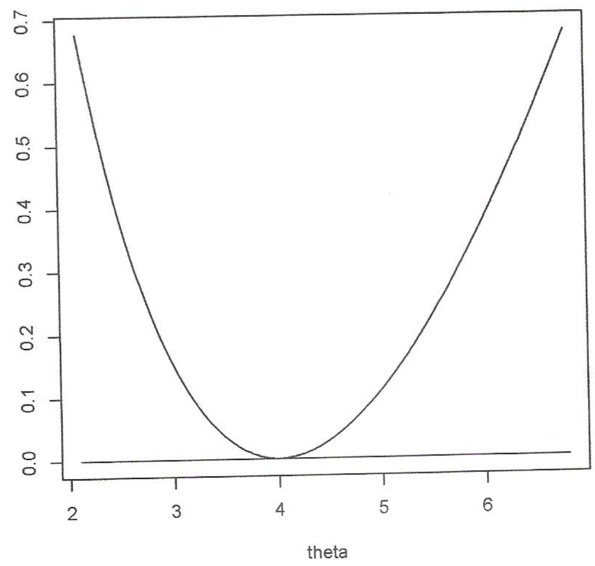

Figure 1: This is the function $g(\theta) = \theta - 4\log(\theta) + 4\log 4 - 4$.

ii. Determine an approximate 95% confidence interval based on the asymptotic distribution of the deviance $D(\theta)$. (The solution does not have to be exact and you can use Figure 1, but pay attention to the definition of be g!) [5 Marks]

3. **Hypothesis testing.** Let $X = (X_1, \ldots, X_n)$ be a sample of independent, identically distributed random variables with density f_θ. The Neyman-Pearson Lemma describes the optimal test-statistic and associated rejection region for testing two simple hypotheses:

$$H_0 : \theta = \theta_0$$
$$H_1 : \theta = \theta_1$$

Let $L_i(X) = \prod_{i=1}^n f_{\theta_i}(X_i)$ be the likelihood associated with θ_i. Then

$$C = \left\{X \mid \frac{L_0(X)}{L_1(X)} \leq k\right\},$$

such that

$$P(C|H_0) = \alpha,$$

is the most powerful critical region of size α.

2

(a) The following questions ask you to prove this statement, i.e., for any other critical region D of size α,

$$P(C|H_1) \geq P(D|H_1).$$

i. Let D be a critical region of size α. Show that

$$\int_{C\cap D^c} L_0(x)dx = \int_{D\cap C^c} L_0(x)dx.$$

[5 Marks]

ii. Use (a) and the definition of C to show that

$$\int_{C\cap D^c} L_1(x)dx \geq \int_{D\cap C^c} L_1(x)dx.$$

[5 Marks]

iii. Use (b) to shown that

$$P(C|H_1) \geq P(D|H_1).$$

[5 Marks]

(b) On the packaging of a particular light bulb it states that it will burn on average 2 time units (measured in 10,000 hours). A review in magazine however reports that its burning average is only 1 time unit. We do an experiment to find out. Let $X_1, \ldots, X_5$ be a sample of independent, identically distributed Exp(θ) random variables, with density

$$f_\theta(x) = \begin{cases} \frac{1}{\theta}e^{-x/\theta} & x \geq 0 \\ 0 & \text{elsewhere} \end{cases}$$

We want test,

$$H_0 : \theta = 2$$
$$H_1 : \theta = 1$$

i. Determine the optimal rejection region of size 0.05. (Note: (i) the rejection region is a subspace of $\mathbb{R}^5$; (ii) the sum of n Exp(2) distributed variables is χ^2_{2n} distributed (iii) a table of the chi-squared quantiles is given at the end of the exam.) [10 Marks]

ii. If we observe the following five burning times, $x_1 = 0.9, x_2 = 0.8, x_3 = 0.8, x_4 = 0.6, x_5 = 0.9$, then use the optimal rejection region to decide whether H_0 can be rejected at the 0.05 level in favour of H_1. [5 Marks]

Next page contains 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 $\chi^2_{\alpha,\nu}$ as found in the book: 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.

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 as found in the book. This means that values in the table correspond to probabilities $P(0 < Z \leq z)$, where Z is a standard normal distributed variable.

4