29 January 2016, 9:00–12:00

Rijksuniversiteit Groningen
Statistiek

Hertentamen

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 chi-squared table.
- Your exam mark : 10 + 90× your score/75 .

1. **Asymptotic distribution Likelihood Ratio Test statistic** $\boxed{\textbf{20 Marks}}$**.**
   Let $X = (X_1, \ldots, X_n)$ be the observed data, such that

   $$X_i \overset{\text{i.i.d.}}{\sim} f_{\theta_0}.$$

   Let f_θ be twice continuously differentiable with support not depending on θ.
   Let $\hat{\theta} = \hat{\theta}(X)$ be the maximum likelihood estimate of θ.

   (a) Derive the second order Taylor expansion of the log-likelihood $\ell(\theta) = \log f_\theta(X_1, \ldots, X_n)$ at the true value $\ell(\theta_0)$ around the MLE $\hat{\theta}$. [5 Marks]

   (b) Show that $\frac{1}{n}\frac{d^2\ell}{d\theta^2}(\theta_0) \to -I(\theta_0) = E\frac{d^2}{d\theta^2}\log f_\theta(X_1)|_{\theta=\theta_0}$ as $n \to \infty$. [5 Marks]

   (c) Use the second order Taylor expansion to show that

   $$-2\log LRT \approx -(\hat{\theta} - \theta_0)^2 \frac{d^2\ell}{d\theta^2}(\hat{\theta}).$$

   where LRT is the likelihood ratio test statistic.[5 Marks]

   (d) Taking the approximation in (c) as an equality, use (b) and (c) together with the asymptotic efficiency of the MLE $\hat{\theta}$ to show that

   $$-2\log LRT \to \chi^2_1$$

   in distribution as $n \to \infty$. [5 Marks]

2. **Point estimation** $\boxed{\textbf{10 Marks}}$**.**
   Let $X_1, \ldots, X_n$ be a sample of independent, identically distributed random variables, with Unif$(0, \theta)$ density.

   (a) Determine the Method of Moments estimator $\hat{\theta}$ of θ. [5 Marks]

   (b) Determine whether $\hat{\theta}$ is consistent. [5 Marks]

3. **Survival regression** $\boxed{\textbf{30 Marks}}$.
Let $(Y_1, x_1), ..., (Y_n, x_n)$ be the data, where $\{Y_i\}_{i=1}^n$ are independently and exponentially distributed random variables in the following way:

$$Y_i \sim Ex(\lambda x_i), \quad i = 1, 2, \ldots, n$$

i.e.

$$f_{Y_i}(y) = \lambda x_i e^{-\lambda x_i y} 1_{y \geq 0}.$$

The known constants $\{x_i\}$ are strictly positive.

(a) Derive the maximum likelihood estimator for λ. [Hint: Don't forget to show that this is really a maximum.] [10 Marks]

(b) Determine whether the MLE is a sufficient statistic. [5 Marks]

(c) Exponential distributions are often used to model survival times. A researcher wants to find the relationship between the price of a light bulb and its life span. She samples 40 different light bulbs and hypothesizes the above model, whereby for $i = 1, \ldots, 40$

- $x_i = \frac{1}{\text{(euro) price of light bulb } i}$
- $Y_i =$ life span of light bulb i (in years).

The plot of the relationship between price and lifespan is given in the plot below. We are given that $\sum_{i=1}^{40} x_i y_i = 80$.

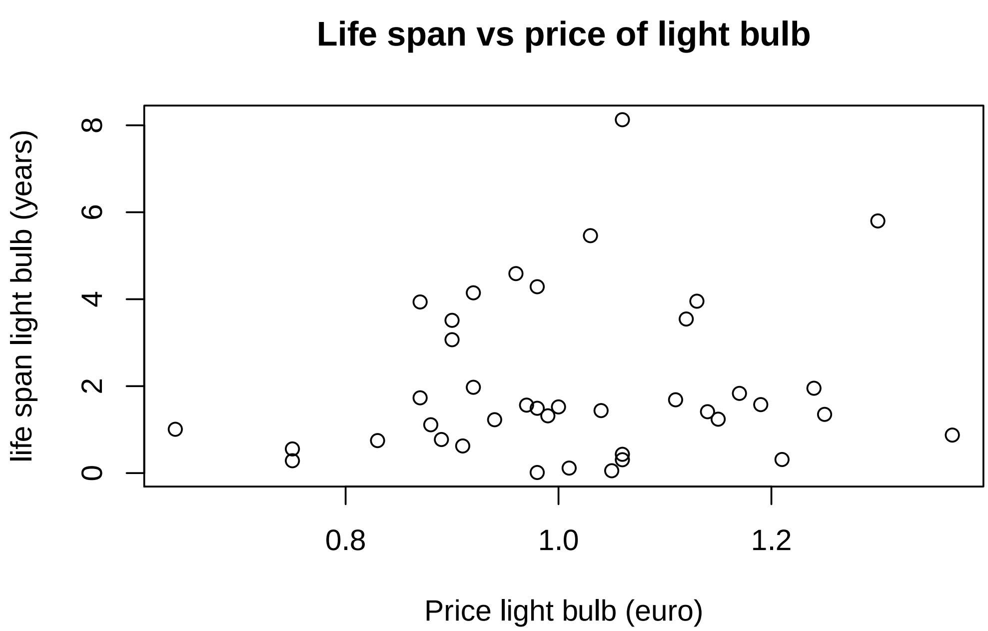

Figure 1: 40 measurements of light bulb life spans versus their price.

i. A consumer organization wants to know whether, based on these data, it can reject $\lambda = 1$. Based on the asymptotic distribution of the MLE as test-statistic, set up a hypothesis test with null and alternative hypothesis to see if there is sufficient evidence for a relationship at a significance level of $\alpha = 0.05$ (hint: the standard normal 0.975 quantile is 1.96). [10 Marks]

ii. Calculate the (numeric!) 95% confidence interval for λ based on the asymptotic distribution of the likelihood ratio statistic. Draw this confidence interval in a plot of the likelihood ratio statistic versus λ. [Hint: you are allowed to "read off" the numeric CI from this plot, but provide the inequality that you would like to solve.] [5 Marks]

4. **Optimal testing** $\boxed{\textbf{15 Marks}}$**.** Consider a single observation X from a exponential distribution with mean μ, i.e. with density

$$f_X(x) = \frac{1}{\mu} e^{-x/\mu}, \qquad x > 0$$

We want to test the following hypotheses:

$$\begin{aligned} H_0 : &\quad \mu = 1 \\ H_1 : &\quad \mu = 2 \end{aligned}$$

(a) We want to perform an optimal test with a significance level of at most 5% of the null hypothesis against the alternative. Determine the critical region. [10 Marks]

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

Below 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.