4 December 2013, 18:30 – 21:30

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 normal table and a chi-squared table.
- Your exam mark : 10 + 90× your score/85 .

1. **Cramer-Rao: best unbiased estimators.**
   Let $X = (X_1, \ldots, X_n)$ be the observed data, such that
   $$X_i \overset{\text{i.i.d.}}{\sim} f_\theta.$$
   Let $\hat{\theta} = \hat{\theta}(X)$ be an unbiased estimate of θ. Let $Y = \frac{d}{d\theta} \log f_{\theta,\text{joint}}(X)$.
   (a) Show that $EY = 0$. [5 Marks]
   (b) Show that $\text{Cov}(\hat{\theta}, Y) = 1$. [5 Marks]
   (c) Use Cauchy-Schwarz to show that $V(\hat{\theta}) \geq 1/E(Y^2)$. [5 Marks]
   (d) Use the above to show that
   $$V(\hat{\theta}) \geq \frac{1}{nE(\frac{d}{d\theta} \log f_\theta(X_1))^2}.$$
   [5 Marks]

2. **Linear regression.** Let $(x_i, Y_i) \in \mathbb{R}^2$ be independent observations on n subjects, such that
   $$Y_i | x_i \sim N(x_i \beta, \sigma^2),$$
   where $(\beta, \sigma^2) \in \mathbb{R}^2$ are unknown coefficients.
   (a) Derive the maximum likelihood estimate $\hat{\beta}$ of β. [5 Marks]
   (b) Determine whether $\hat{\beta}$ is unbiased. [5 Marks]
   (c) Derive the variance of $\hat{\beta}$. [5 Marks]
   (d) Derive the maximum likelihood estimate of σ^2. [5 Marks]

3. **Point estimation.** Let $X_1, ..., X_n$ be independently Poisson distributed random variables with parameter θ, i.e.

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

(a) Find a sufficient statistic $\hat{\theta}(X_1, \ldots, X_n)$ for θ. [5 Marks]

(b) Determine the Cramer-Rao lowerbound for an unbiased estimator of θ. [5 Marks]

(c) Let $\hat{\theta}_n = \bar{X}$ be an estimator of θ. Show that [5 Marks]

$$\forall \epsilon > 0 : \lim_{n\to\infty} P(|\hat{\theta}_n - \theta| \leq \epsilon) = 1.$$

(d) Assume the asymptotic normality, unbiasedness and efficiency of the estimator $\hat{\theta}_{100}$. Based on this statistic, determine the usual (i.e. symmetric or minimum length) 95% confidence interval, if you know that $\sum_{i=1}^{100} x_i = 200$. [10 Marks]

4. **Optimal testing.** Consider a single observation X from an exponential distribution, $X \sim \text{Exp}(\mu)$, i.e. with density

$$f_X(x) = \frac{e^{-x/\mu}}{\mu}, \quad x \geq 0$$

and cumulative distribution function

$$F_X(x) = 1 - e^{-x/\mu}, \quad x \geq 0.$$

We want to test the following hypotheses:

$$\begin{aligned} H_0 : &\quad \mu = 1 \\ H_1 : &\quad \mu = 3 \end{aligned}$$

(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. [15 Marks]

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

Below a statistical table 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.