4 February 2013, 14:00 – 17:00

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.
- Your exam mark : 10 + your score (max 90) .

1. **Maximum likelihood.** Let $X_1, ..., X_n$ be independently Poisson distributed with parameter θ, i.e.

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

   (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. [10 Marks]

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

   (d) 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]

   (e) 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]

2. **Normal data.** Let $X_1, \ldots, X_n$ be a sample of independent, identically distributed Normal(μ, σ^2) random variables.

   (a) Determine the MLE for μ. [5 Marks]

   (b) [10 Marks] Determine a 95% CI for μ keeping in mind that

      - we don't know σ^2,
      - $n = 9$,
      - $\sum_{i=1}^{9} x_i = 27$ and $\sum_{i=1}^{9} x_i^2 = 113$.
      - We know the following quantiles $t_{n,\alpha}$ of the t-distribution:
        - $t_{1,.95} = 6.31$ $\quad t_{1,.975} = 12.71$
        - $t_{8,.95} = 1.86$ $\quad t_{8,.975} = 2.31$
        - $t_{9,.95} = 1.83$ $\quad t_{9,.975} = 2.26$

      whereby $P(T_n \leq t_{n,\alpha}) = \alpha$ for a t-distributed random variable T_n with n degrees of freedom.

3. **Hypothesisl testing.** An Atomic Energy Agency is worried that a particular nuclear plant has leaked radio-active material. They plan to do n independent measurements, which are assumed to Poisson(θ) distributed, with probability mass function

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

(a) The natural background radiation has an average of 2 counts (per unit time) and one employee suggests to use the likelihood ratio test to test the hypotheses,

$$H_0: \quad \theta = 2$$
$$H_1: \quad \theta \neq 2$$

i. Determine the log-likelihood $l_x(\theta)$ [5 Marks].

ii. Determine the logarithm of the likelihood ratio statistic $\log \lambda(x)$. [5 Marks]

iii. The junior employee believes that they will get a lot of data and wants to test the above hypotheses based the asymptotic distribution of the likelihood ratio statistic. Derive an expression for the Critical Region, $CR \subset \mathbb{R}^n$ associated with a 5% significance level. [5 Marks]

(b) However, only 5 independent Geiger counter measurements in the direct neighbourhood of the reactor are taken. Moreover, the agency should only be worried if the radiation rate would be at the level of $\theta = 5$. They therefore decide to test,

$$H_0: \quad \theta = 2$$
$$H_1: \quad \theta = 5$$

Use the Neyman-Pearson testing procedure at the 5% significance level to test whether to reject H_0 based on the following measurements

observation	1	2	3	4	5
count	4	2	3	2	6

Hint: you can use the following information:

- the sum of independent Poissons is Poisson distributed with a mean equal to the sum of the original means.
- the 0.05 quantile of a Poisson(10) is 5, whereas the 0.95 quantile of a Poisson(10) is 15.

[15 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.