Resit Exam - Statistics (WBMA009-05) 2020/2021

Date and time: January 29, 2021, 18.45-21.45h
Place: Nestor Online Exam

Rules to follow:

- This is an online open book exam. Consultation of books and notes is permitted.
- You can use a simple (non-programmable) calculator.
- There are three exercises and you can reach 90 points.
- Always include the relevant equation(s) and/or short descriptions.
- **We wish you success with the completion of the exam!**

START OF EXAM

1. **Hypothesis testing, p-value and confidence interval.** [20]
   Consider a random sample of size $n = 16$ from a Gaussian distribution with known variance parameter $\sigma^2 = 4$, symbolically:

$$X_1, \ldots, X_{16} \sim \mathcal{N}(\mu, 4)$$

and the two-sided test problem

$$H_0 : \mu = -2 \ \textbf{vs.} \ H_1 : \mu \neq -2$$

Under H_0 we have that the statistic:

$$T(X) := 4 + 2\bar{X}_{16}$$

has a standard Gaussian $\mathcal{N}(0, 1)$ distribution.
For solving the exercise use <u>and only use</u> the quantiles provided in Table 1.

   (a) Give the rejection region of a two-sided test to the level $\alpha = 0.05$. [5]
   (b) Give the power of the two-sided level $\alpha = 0.05$ test at the true parameter value $\mu = -1$. [5]

Now we assume that we have observed the realization $\bar{X}_{16} = -2.65$.

   (c) Give the p-value of the two-sided test. [5]
   (d) Give an exact two-sided 80% confidence interval for μ. [5]

α	0.5	0.75	0.9	0.95	0.975	0.99	0.99997
q_α	0	0.7	1.3	1.6	2	2.3	4

Table 1: Approximate quantiles q_α of the $\mathcal{N}(0, 1)$ distribution.

2. **Random sample I.** $\boxed{30}$
Consider a random sample of size n from a discrete distribution

$$X_1, \ldots, X_n \sim \mathcal{D}(\theta)$$

that depends on one single parameter

$$0 < \theta < 1$$

and whose density is given by:

$$f_\theta(x) = (1-\theta)^{x-1} \cdot \theta \qquad (x = 1, 2, 3, \ldots)$$

It can be shown that this implies the expectation $E[X_1] = \frac{1}{\theta}$.

(a) Compute the ML estimator of θ. $\boxed{5}$
HINT: Check via the 2nd derivative whether you have a maximum.

(b) Compute the Fisher information $I(\theta)$ for a sample of size $n = 1$. $\boxed{5}$

For the following two exercise parts (c-d) consider the test problem

$$H_0 : \theta = 0.6 \ \textbf{vs.} \ H_1 : \theta = 0.4$$

(c) Show that a statistical test that rejects H_0 if $\sum_{i=1}^{n} X_i > k_0$, where $k_0 > 0$ is a constant, is the UMP test for the test problem. $\boxed{5}$
HINT: Apply the Neyman Person lemma.

(d) Consider the UMP test from (c) with $n = 1$ and $k_0 = 3$.
Compute the test level α. $\boxed{5}$

For the last exercise part (e) assume $n = 5$ and the five realisations:
$X_1 = 1$, $X_2 = 3$, $X_3 = 1$, $X_4 = 3$, and $X_5 = 2$.

(e) Make use of the asymptotic efficiency of the ML estimator to construct an asymptotic two-sided 80% confidence interval for θ. $\boxed{10}$
HINT: See Table 1 from Exercise 1 for a list of quantiles.

3. **Random sample II.** $\boxed{40}$
Consider a distribution $\mathcal{D}(\theta_1, \theta_2)$ that depends on two parameters

$$0 < \theta_1 < 1 \quad \text{and} \quad \theta_2 > 0$$

and whose density is given by:

$$f_{\theta_1,\theta_2}(x) = \frac{1-\theta_1}{\theta_2} \cdot \exp\left\{\frac{x}{\theta_2}\right\} \cdot I_{(-\infty,0)}(x) + \frac{\theta_1}{\theta_2} \cdot \exp\left\{-\frac{x}{\theta_2}\right\} \cdot I_{[0,\infty)}(x)$$

We consider a random sample of size n:

$$X_1, \ldots, X_n \sim \mathcal{D}(\theta_1, \theta_2)$$

(a) Give the likelihood without using indicator functions or case distinctions. $\boxed{5}$
<u>HINT</u>: Use the order statistics and assume that you have the relationship:

$$X_{(1)} \leq \ldots \leq X_{(n-K)} < 0 \leq X_{(n-K+1)} \leq \ldots < X_{(n)}$$

so that K out of n random variables take non-negative values.

(b) Show that the log likelihood is then given by:

$$l(\theta_1, \theta_2) = K \cdot \log(\theta_1) + (n-K) \cdot \log(1-\theta_1) - n \cdot \log(\theta_2) - \frac{S}{\theta_2}$$

where $S := \sum_{i=1}^{n} |X_i|$ $\boxed{5}$

$$\underline{\text{HINT}}: \quad \sum_{i=1}^{n-K} X_{(i)} + \sum_{i=n-K+1}^{n} -X_{(i)} = -S$$

For the following exercise parts (c-g), we assume that θ_2 is a known parameter, and we let K denote the number of non-negative realisations in the sample.

(c) Show that K is a sufficient statistic for θ_1. $\boxed{5}$

(d) Show that $P(X_1 > 0) = \theta_1$. $\boxed{5}$

(e) Compute $E[K]$. $\boxed{5}$.

(f) Compute the ML estimator of θ_1 and show that it is unbiased. $\boxed{5+5}$

(g) Show that the Fisher information for sample size $n = 1$ is $I(\theta_1) = \frac{1}{\theta_1(1-\theta_1)}$. $\boxed{5}$