UNIVERSITY OF AMSTERDAM

Faculty of Science

Exam

Asymptotic Statistics

Mastermath stochastics

Re-sit exam
Date: Wednesday 17 January 2024
Time: 9.45–12.45h

Number of pages: 4 (including front page)
Number of questions: 3
Maximum number of points: 75
For each question is indicated how many points it is worth.

BEFORE YOU START

- Check if your version of the exam is complete.
- Write down **your name, student ID number**, and if applicable the **version number** on **each sheet** that you hand in. Also **number the pages**.
- Your **mobile phone** has to be switched off and be put in your coat or bag. Your **coat and bag** should be on the ground.
- Upper-right of first page **Write: Name + University + page-nr. on EVERY page**

PRACTICAL MATTERS

- The first 30 minutes you are not allowed to leave the room, not even to visit the toilet.
- 15 minutes before the end, you will be warned that the time to hand in is approaching.
- If applicable, fill out the evaluation form at the end of the exam.
- You are obliged to identify yourself at the request of the examiner (or his representative) with a proof of your registration and a valid ID.
- During the examination it is not permitted to visit the toilet, unless the invigilator gives permission to do so.

Good luck!

Problem 1 *(Two estimators for the Poisson parameter)*

Let $X_1, X_2, \ldots$ be an *i.i.d.* sample from a Poisson distribution with parameter $\lambda > 0$. That means that,

$$P(X_i = k) = \frac{e^{-\lambda}\lambda^k}{k!},$$

for all $i \geq 1, k \geq 0$. Define two sequences of statistics (T_n) and (S_n) as follows:

$$S_n = \frac{1}{n}\sum_{i=1}^{n} 1\{X_i = 0\}, \quad T_n = \frac{1}{n}\sum_{i=1}^{n} 1\{X_i = 1\},$$

a. *(5 points)*
Show that $\log(1/S_n)$ and T_n/S_n are consistent estimators for the parameter λ.

b. *(5 points)*
Show that $\hat{\lambda}_1 = \log(1/S_n)$ is asymptotically normal and give the limit distribution.

c. *(5 points)*
Show that $\hat{\lambda}_2 = T_n/S_n$ is asymptotically normal and give the limit distribution.

d. *(5 points)*
Describe what the relative efficiency of two estimator sequences signifies (e.g. through sample sizes). Based on your answers at parts *b.* and *c.*, calculate the relative efficiency of $\hat{\lambda}_1$ and $\hat{\lambda}_2$.

Problem 2 *(A transformed exponential distribution)*

Let $X_1, X_2, \ldots$ be *i.i.d.* non-negative real-valued random variables with single-observation distribution P_{μ_0} and Lebesgue density $p_{\mu_0} : \mathbb{R} \to [0, \infty)$ for some $\mu_0 > 0$, with $p_\mu(x) = 0$ for $x < 0$, and

$$p_\mu(x) = 2\mu\, x\, e^{-\mu\, x^2},$$

for $x \geq 0$ and $\mu > 0$. A change of variables $Z = X^2$ leads to $Z \sim \text{Exp}(\mu)$.

Hint: you may use the following integrals,

$$\int_0^\infty x^2 e^{-x^2}\, dx = \frac{\sqrt{\pi}}{4}, \int x^3 e^{-x^2}\, dx = -\frac{e^{-x^2}(x^2+1)}{2} + C, \int x^5 e^{-x^2}\, dx = -\frac{e^{-x^2}(x^4+2x^2+2)}{2} + C,$$

a. *(5 points)*
Find the maximum-likelihood estimator $\hat{\mu}_n$ for μ_0 based on the first n sample points $X_1, \ldots, X_n$.

b. *(5 points)*
Calculate the expectation $E_\mu X_1^2$ of the second moment for a single observation and show that $\hat{\mu}_n$ is a consistent estimator sequence for μ_0.

c. *(5 points)*
Calculate the Fisher information I_μ for a single observation X_1, show that $\mu \mapsto I_\mu$ is continuous and that $I_\mu > 0$ for all $\mu > 0$.

d. *(5 points)*
Show that, for any $x > 0$, the map $\mu \mapsto \log p_\mu(x)$ is Lipschitz in a neighbourhood of μ_0. In other words, prove that for some $\epsilon > 0$ and any $\mu_1, \mu_2 > 0$ such that $|\mu_1 - \mu_0| < \epsilon$ and $|\mu_2 - \mu_0| < \epsilon$,

$$\big|\log p_{\mu_1}(x) - \log p_{\mu_2}(x)\big| \leq \dot{\ell}(x)|\mu_1 - \mu_2|,$$

for some measurable function $\dot{\ell} : \mathbb{R} \to \mathbb{R}$ such that $E_{\mu_0}\dot{\ell}^2 < \infty$.
Hint: for any $\mu_1, \mu_2 \geq \mu > 0$, we have $|\log \mu_1 - \log \mu_2| \leq |\mu_1 - \mu_2|/\mu$.

e. *(5 points)*
State a theorem from the lecture notes and use parts a.–d. to prove that $\sqrt{n}(\mu_n - \mu_0)$ is asymptotically normal under P_{μ_0}. Give the variance of the limit distribution.

The moment estimator for μ_0 is,

$$\tilde{\mu}_n = \frac{\pi}{4}\Big(\frac{1}{\overline{X}_n}\Big)^2,$$

where $\overline{X}_n$ denotes the sample average.

f. *(5 points)*
Use the delta rule to find the limit distribution for $\sqrt{n}(\tilde{\mu}_n - \mu_0)$.

Problem 3 *(Domain boundary estimation)*

Let $Y_1, Y_2, \ldots$ be an *i.i.d.* sample from the uniform distribution P_θ on $[0, \theta]$, for some $\theta > 0$. The distribution function for $X \sim U[0,1]$ is given by $P(X \leq 0) = 0$, $P(X \leq x) = x$, $(0 < x \leq 1)$, $P(X \leq 1) = 1$. Denote the maximum of the first n observations by $Y_{(n)} = \max\{Y_1, \ldots, Y_n\}$. *(Hint: in the problem, you may use that* $(1 + a/n)^n \to e^a$ *as* $n \to \infty$*, for any* $a \in \mathbb{R}$*.)*

a. *(5 points)*
Given $n \geq 1$, find the maximum-likelihood estimator $\hat{\theta}_n$ for θ, based on the first n observations.

b. *(5 points)*
Show that, for given $n \geq 1$ and θ, the distribution function of satisfies,

$$P(Y_{(n)} \leq x) = \left(\frac{x}{\theta}\right)^n,$$

for all $0 < x \leq \theta$.

c. *(5 points)*
Show that $\hat{\theta}_n$ is consistent for estimation of θ.

d. *(5 points)*
Show that $n(\theta - \hat{\theta}_n)$ converges weakly and give the limit distribution.

Given any estimators $\tilde{\theta}_n$ for the parameter θ, define the *bias* Δ_n of $\tilde{\theta}_n$ by the (θ-dependent) expectation $\Delta_n = P_\theta^n(\tilde{\theta}_n - \theta)$.

e. *(5 points)*
For every $n \geq 1$, give the bias Δ_n of $\hat{\theta}_n$. Find a real-valued sequence (a_n) such that the bias of the estimators $a_n\hat{\theta}_n$ is exactly zero.