University of Ljubljana, School of Economics

Quantitative finance and actuarial science

Probability and statistics

Written examination

January 30$^{\text{th}}$, 2025

Name and surname: ______________________ ID: ☐☐☐☐☐☐☐☐

## Instructions

Read the problems carefull before staring your work. There are four problems. You can have a sheet with formulae and a mathematical handbook. Write your solutions on the paper provided. You have two hours.

| Problem | a. | b. | c. | d. | Total |
|---|---|---|---|---|---|
| 1. | | | • | • | |
| 2. | | | • | • | |
| 3. | | | | | |
| 4. | | | | • | |
| Total | | | | | |

**1.** (25) A population of size $N$ is divided into $K$ groups of equal size $M = N/K$. A sample is selected in such a way that $k$ groups are selected by simple random sampling, and then all the units in the selected groups are selected.

a. (10) Show that the sample average $\bar{Y}$ is an unbiased estimate of the population mean.

   *Solution: let $\mu_i$ be the population mean in the $i$-th group. In the sampling procedure described we are choosing a simple random sample of groups and we observe $\mu_i$ for this group. The estimator $\bar{Y}$ is just a sample average of the $\mu_i$ selected. The expectation is therefore the average of all $\mu_i$s which is $\mu$.*

b. (15) Let $\mu_i$ be the population mean in group $i$ for $i = 1, 2, \ldots, K$ and let $\mu$ be the population mean. Define

$$\sigma_b^2 = \frac{1}{K} \sum_{i=1}^{K} (\mu_i - \mu)^2 .$$

   Show that

$$\operatorname{se}\left(\bar{Y}\right) = \frac{\sigma_b}{\sqrt{k}} \cdot \sqrt{\frac{K-k}{K-1}} .$$

   *Solution: think of groups as units selected and to each group assign the value $\mu_i$. The formula is then the formula for the standard error of such a sample average. But $\bar{Y}$ is equal to this sample average.*

**2.** (25) Let $x_1, x_2, \ldots, x_n$ be an i.i.d. sample from the distribution with density

$$f(x) = \frac{\lambda^2}{12} x\, e^{-\sqrt{\lambda x}}$$

for $x > 0$ and $\lambda > 0$.

a. (15) Find the Fisher information. Assume as known that

$$\int_0^\infty x^{3/2} e^{-\sqrt{\lambda x}}\, \mathrm{d}x = \frac{48}{\lambda^{5/2}}\,.$$

*Solution: the log-likelihood function is*

$$\ell(\lambda|x) = 2\log\lambda - \log 12 + \log x - \sqrt{\lambda x}\,.$$

*Taking the second derivative we get*

$$\ell'' = -\frac{2}{\lambda^2} + \frac{\sqrt{x}}{4\lambda^{3/2}}\,.$$

*It follows*

$$\begin{aligned} I(\lambda) &= \frac{2}{\lambda^2} - \frac{1}{4\lambda^{3/2}}\, E(\sqrt{X}) \\ &= \frac{2}{\lambda^2} - \frac{1}{4\lambda^{3/2}} \cdot \frac{\lambda^2}{12} \cdot \int_0^\infty x^{3/2} e^{-\sqrt{\lambda x}}\, \mathrm{d}x \\ &= \frac{2}{\lambda^2} - \frac{1}{4\lambda^{3/2}} \cdot \frac{\lambda^2}{12} \cdot \frac{48}{\lambda^{5/2}} \\ &= \frac{1}{\lambda^2}\,. \end{aligned}$$

b. (10) Write explicitely the 99%-confidence interval for $\lambda$ on the basis of the data $x_1, x_2, \ldots, x_n$.

*Solution: the log-likelihood function is*

$$\ell(\lambda|x_1,\ldots,x_n) = 2n\log\lambda - n\log 12 + \sum_{k=1}^n \log x_k - \sqrt{\lambda}\sum_{k=1}^n \sqrt{x_k}\,.$$

*Taking derivatives we get the equation*

$$\frac{2n}{\lambda} - \frac{1}{2\sqrt{\lambda}}\sum_{k=1}^n \sqrt{x_k} = 0$$

*with the solution*

$$\hat{\lambda} = \left( \frac{4n}{\sum_{k=1}^{n} \sqrt{x_k}} \right)^2 .$$

*The 99%-confidence interval is*

$$\hat{\lambda} \pm 2.56 \cdot \frac{\hat{\lambda}}{\sqrt{n}} .$$

**3.** (20) The $\chi^2$ statistic can be used to test whether a roulette wheel is unbiased. If $O_i$ is the number of observed occurences of $i$ and $E_i$ is the number of expected occurences we define

$$\chi^2 = \sum_{i=0}^{36} \frac{(O_i - E_i)^2}{E_i}\,.$$

Large values of the $\chi^2$ statistic indicate that the roulette wheel is biased. We are assuming that individual spins are independent and that the probabilities are constant throughout the observation period.

Suppose the gambling house tests all the weels at the end of every month on the basis of data collected in that month. The rule is that a wheel is examined more closely if the $p$-value is below 0.01.

a. (5) Suppose that for a roulette wheel we got the $p$-value $p = 0.005$. Can this happen with an unbiased wheel? With what probability?

   *Solution: yes, it can happen with probability 0.005.*

b. (5) Suppose that for a roulette wheel the $p$-value was $p = 0,23$. Is this conclusive evidence that the wheel is unbiased? Explain.

   *Solution: no, it is not conclusive evidence.*

c. (5) Suppose a gambling house has 100 roulette wheels which are tested every month on the basis of data collected. Suppose all the wheels were unbiased. How many wheels per month would be examined on average over a long period of time. Explain.

   *Solution: the probability of examining an unbiased wheel is 0.01. So on average one wheel would be examined.*

d. (5) Suppose one of the wheels is biased. Is the probability that it will be examined more or less than 0.01? Explain.

   *Solution: any sensible test would have to have power exceeding its size.*

**4.** (25) Assume the regression model

$$Y_k = \beta x_k + \epsilon_k$$

for $k = 1, 2, \ldots, n$ where $\epsilon_1, \ldots, \epsilon_n$ are uncorrelated, $E(\epsilon_k) = 0$ and $\mathrm{var}(\epsilon_k) = \sigma^2$ for $k = 1, 2, \ldots, n$. Assume that $x_k > 0$ for all $k = 1, 2, \ldots, n$. Consider the following linear estimators of $\beta$:

$$\begin{aligned} \hat{\beta}_1 &= \frac{\sum_{k=1}^n x_k Y_k}{\sum_{k=1}^n x_k^2} \\ \hat{\beta}_2 &= \frac{1}{n} \sum_{k=1}^n \frac{Y_k}{x_k} \\ \hat{\beta}_3 &= \frac{\sum_{k=1}^n Y_k}{\sum_{k=1}^n x_k} \end{aligned}$$

a. (5) Are all estimators unbiased?

*Solution: since $E(Y_k) = \beta x_k$ for all $k = 1, 2, \ldots, n$ all the estimators are unbiased.*

b. (10) Which of the estimators has the smallest standard error? Justify your answer.

*Solution: all the estimators are unbiased. Gauss-Markov tells us that the best estimator is the one given by least squares and that is $\hat{\beta}_1$.*

c. (10) Write down the standard errors for all three estimators.

*Solution: we first compute the theoretical variances. Since $Y_1, \ldots, Y_n$ are uncorrelated we have*

$$\begin{aligned} \mathrm{var}(\hat{\beta}_1) &= \frac{\sigma^2}{\sum_{k=1}^n x_k^2} \\ \mathrm{var}(\hat{\beta}_2) &= \frac{\sigma^2 \sum_{k=1}^n x_k^{-2}}{n^2} \\ \mathrm{var}(\hat{\beta}_3) &= \frac{n\sigma^2}{\left(\sum_{k=1}^n x_k\right)^2} . \end{aligned}$$

*We need an unbiased estimate of $\sigma^2$. Theoretically, we have that*

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{k=1}^n (Y_k - \hat{\beta} x_k)^2$$

*is an unbiased estimator $\sigma^2$. This gives us an unbiased estimator of $\sigma^2$.*