School of Economics

Doctoral Programme in Economics

Probability and statistics

Written examination

September 1st, 2022

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.	
1.			•	•	
2.					
3.			•	•	
4.			•	•	
Total					

1. (25) Suppose we have a population with N units. The values of the statistical variable are $x_1, x_2, \ldots, x_N$. Denote by μ the population mean and by σ^2 the population variance.

a. (5) Suppose you chose a simple random sample of size n. Denote

$$\gamma = \frac{1}{N} \sum_{k=1}^{N} x_k^2 .$$

Suggest an unbiased estimate for γ. Explain why it is unbiased.

Solution: an unbiased estimate of γ is the sample average of the squares of sample values. We also have

$$\sigma^2 = \frac{1}{N} \sum_{k=1}^{N} x_k^2 - \mu^2 = \gamma - \mu^2 .$$

b. (5) Suppose you chose a simple random sample of size n. Suggest an unbiased estimate for μ^2.

Hint: Note that $\sigma^2 = \gamma - \mu^2$.

Solution: we know that

$$\hat{\sigma}^2 = \frac{N-1}{N(n-1)} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

is an unbiased estimate of σ^2. We have denoted the sample values by $X_1, \ldots, X_n$. Since in the above equation in the hint we have unbiased estimates for two of the three quantities and the relationship is linear, we can estimate the third, i.e. μ^2, in an unbiased way.

c. (5) Assume now that the population is divided into K equally sized groups of size M so that $N = KM$. A sample is chosen in such a way that k groups are chosen from all the K groups by simple random sampling. Then all the units from the chosen groups are included into the sample. For the estimator we chose the average of all the kM sample values. Denote by μ_k the population average for the k-the group and by σ_k^2 the population variance for the k-th group. Find the standard error of the suggested estimator using the quantity

$$\tau^2 = \frac{1}{K} \sum_{k=1}^{K} (\mu_k - \mu)^2 .$$

Solution: since all the groups are of equal size we have $\mu = \frac{1}{K} \sum_{k=1}^{K} \mu_k$. We can think that we are choosing a simple random sample from a population of groups. The estimator is therefore unbiased and its variance is given by

$$\mathrm{var}(\bar{X}) = \frac{\tau^2}{k} \cdot \frac{K-k}{K-1} ,$$

where

$$\tau^2 = \frac{1}{K}\sum_{r=1}^{K}(\mu_r - \mu)^2 \,.$$

d. (10) Assume that the sample is as in c. We would like to estimate the population variance σ^2 on the basis of the sample. Suggest and unbiased estimate. Explain why it is unbiased.

Hint: look at a.

Solution: we think of groups as our primary sampling units. From a. we know that μ^2 can be estimated in an unbiased way. Returning to our sampling procedure we see that we have an unbiased estimator of

$$\frac{1}{N}\sum_{k=1}^{N} x_k^2 \,.$$

Since

$$\sigma^2 = \frac{1}{N}\sum_{k=1}^{N} x_k^2 - \mu^2$$

and we know how to estimate both quantities on the right we can estimate σ^2 in an unbiased way.

To express the estimator explicitly denote by X_{ij} the value of the variable for the jth unit in the ith group selected and let A_i be the average in this group, and by $\bar{A}$ the average of all the group averages which is our estimator. We have

$$\begin{aligned}\hat{\sigma}^2 &= \frac{1}{kM}\sum_{i=1}^{k}\sum_{j=1}^{M} X_{ij}^2 - \frac{1}{k}\sum_{i=1}^{k} A_i^2 + \frac{K-1}{K(k-1)}\sum_{i=1}^{k}(A_i - \bar{A})^2 \\ &= \frac{1}{kM}\sum_{i=1}^{k}\sum_{j=1}^{M}(X_{ij} - A_i)^2 + \frac{K-1}{K(k-1)}\sum_{i=1}^{k}(A_i - \bar{A})^2 \,.\end{aligned}$$

2. (25) Assume the sample values $x_1, x_2, \ldots, x_n$ are in independent identically distributed sample from the gamma distribution with parameters $a = 2$ and λ. The density of the distribution is

$$f(x) = \lambda^2 x e^{-\lambda x}$$

for $x > 0$. Note that the density of the $\Gamma(a, \lambda)$ distribution is given by

$$f(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x}$$

for $x > 0$ and $a, \lambda > 0$, and the expectation is a/λ.

a. (5) Find explicitly the maximum likelihood estimator for the parameter λ.

Solution: the log-likelihood function is

$$\ell\left(\lambda|\mathbf{x}\right) = 2n \log \lambda + \sum_{k=1}^{n} \log x_k - \lambda \sum_{k=1}^{n} x_k\,.$$

Taking derivatives and equation to 0 *we have*

$$\hat{\lambda} = \frac{2n}{\sum_{k=1}^{n} x_k}\,.$$

b. (10) Fix the maximum likelihood estimator so that it will be unbiased.

Hint: if U and V are independent with $U \sim \Gamma(a, \lambda)$ and $V \sim \Gamma(b, \lambda)$ then $U + V \sim \Gamma(a + b, \lambda)$.

Solution: following the hint we have $\sum_{k=1}^{n} X_k \sim \Gamma(2n, \lambda)$. We compute

$$\begin{aligned}
E(\hat{\lambda}) &= E\left(\frac{2n}{\sum_{k=1}^{n} X_k}\right) \\
&= 2n\,\frac{\lambda^{2n}}{\Gamma(2n)} \int_0^{\infty} \frac{1}{x} \cdot x^{2n-1} e^{-\lambda x}\, \mathrm{d}x \\
&= 2n\,\frac{\lambda^{2n}}{\Gamma(2n)} \cdot \frac{\Gamma(2n-1)}{\lambda^{2n-1}} \\
&= \frac{2n\lambda}{2n-1} \\
&= \frac{2n}{2n-1}\,\lambda\,.
\end{aligned}$$

The unbiased estimator is

$$\tilde{\lambda} = \frac{2n-1}{2n}\,\hat{\lambda} = \frac{2n-1}{\sum_{k=1}^{n} X_k}\,.$$

c. (5) Using Fisher information find the approximate standard error for the maximum likelihood estimator.

Solution: we compute for $n = 1$.

$$\ell'' = -\frac{2}{\lambda^2}\,.$$

The approximate standard error is

$$\mathrm{se}(\hat{\lambda}) = \frac{\lambda}{\sqrt{2n}}\,.$$

d. (5) Find the exact variance for the maximum likelihood estimator.

Solution: we need $E\big(\tilde{\lambda}^2\big)$. We compute

$$\begin{aligned} E\big(\tilde{\lambda}^2\big) &= E\left[\left(\frac{2n-1}{\sum_{k=1}^n X_k}\right)^2\right] \\ &= (2n-1)^2 \cdot \frac{\lambda^{2n}}{\Gamma(2n)} \int_0^\infty \frac{1}{x^2}\, x^{2n-1}\, e^{-\lambda x}\, \mathrm{d}x \\ &= (2n-1)^2 \cdot \frac{\lambda^{2n}}{\Gamma(2n)} \cdot \frac{\Gamma(2n-2)}{\lambda^{2n-2}} \\ &= \frac{(2n-1)^2\lambda^2}{(2n-1)(2n-2)} \\ &= \frac{2n-1}{2(n-1)}\,\lambda^2\,. \end{aligned}$$

It follows

$$\mathrm{var}\big(\tilde{\lambda}\big) = \lambda^2\left(\frac{2n-1}{2(n-1)} - 1\right) = \frac{\lambda^2}{2(n-1)}\,.$$

Further, we have

$$\mathrm{var}\big(\hat{\lambda}\big) = \frac{2n^2}{(2n-1)^2(n-1)}\,\lambda^2\,.$$

3. (25) Suppose the observed values are pairs $(x_1, y_1), \ldots, (x_n, y_n)$. Assume the pairs are an i.i.d. sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ from the density

$$f(x, y) = e^{-x} \cdot \frac{1}{\sigma\sqrt{2\pi x}}\, e^{-\frac{(y-\theta x)^2}{2\sigma^2 x}}$$

for $x > 0$ and $-\infty < y < \infty$ and $\sigma^2 > 0$. The testing problem is

$$H_0\colon \theta = 0 \quad \text{versus} \quad H_1\colon \theta \neq 0\,.$$

a. (10) Find the Wilks's test statistic λ.

Solution: the log-likelihood function is

$$\ell\,(\theta, \sigma|\mathbf{x}, \mathbf{y}) = -\frac{n}{2}\log 2\pi - n\log\sigma - \frac{1}{2}\sum_{k=1}^{n}\Big[-\log x_k - \frac{(y_k-\theta x_k)^2}{\sigma^2 x_k}\Big]\,.$$

Computing partial derivatives we get

$$\begin{aligned}\frac{\partial \ell}{\partial \theta} &= \sum_{k=1}^{n}\frac{(y_k-\theta x_k)}{\sigma^2}\\ \frac{\partial \ell}{\partial \sigma} &= -\frac{n}{\sigma} + \sum_{k=1}^{n}\frac{(y_k-\theta x_k)^2}{\sigma^3 x_k}\end{aligned}$$

Equating with 0 *we get*

$$\hat{\theta} = \frac{\sum_{k=1}^{n} y_k}{\sum_{k=1}^{n} x_k}$$

and the second equation gives

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{k=1}^{n}\frac{(y_k-\hat{\theta}x_k)^2}{x_k}\,.$$

When we maximize only over σ^2 *taking derivatives gives*

$$\frac{\partial \ell}{\partial \sigma} = -\frac{n}{\sigma} + \sum_{k=1}^{n}\frac{y_k^2}{\sigma^3 x_k}\,.$$

It fololows

$$\tilde{\sigma}^2 = \frac{1}{n}\sum_{k=1}^{n}\frac{y_k^2}{x_k}\,.$$

After some calculations we get

$$\lambda = -2n\log\hat{\sigma} + 2n\log\tilde{\sigma}\,.$$

b. (15) Assume that H_0 is rejected when $\lambda > \lambda_\alpha$ where λ_α is chosen in such a way that the size of the test is $\alpha \in (0, 1)$. Give an approximate value for λ_α using an appropriate $\chi^2(r)$ distribution?

Solution: Wilks's theorem gives the rejection region as $\{\lambda > \lambda_\alpha\}$ *where* λ_α *is the* $(1 - \alpha)$*th percentile of the* $\chi^2(1)$ *distribution.*

4. (25) Assume the regression model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}\,,$$

where $E(\boldsymbol{\epsilon}) = 0$ and $\mathrm{var}(\boldsymbol{\epsilon}) = \sigma^2(\mathbf{I} + a\mathbf{1}\mathbf{1}^T)$ for some $a > 0$. Assume a is known and $\mathbf{X}$ is a $n \times m$ matrix with rank m.

a. (15) Show that

$$\hat{\boldsymbol{\beta}} = \left[\mathbf{X}^T\left(\mathbf{I} + c\mathbf{1}\mathbf{1}^T\right)\mathbf{X}\right]^{-1}\mathbf{X}^T\left(\mathbf{I} + c\mathbf{1}\mathbf{1}^T\right)\mathbf{Y}$$

for

$$c = -\frac{a}{1 + an}$$

is the best unbiased linear estimator of $\boldsymbol{\beta}$.

Hint: check that

$$\left(\mathbf{I} + a\mathbf{1}\mathbf{1}^T\right)\left(\mathbf{I} + c\mathbf{1}\mathbf{1}^T\right) = \mathbf{I}\,.$$

Solution: let $\tilde{\boldsymbol{\beta}}$ be an unbiased linear estimator of $\boldsymbol{\beta}$. We can write

$$\tilde{\boldsymbol{\beta}} = \mathbf{L}\mathbf{Y}$$

for a matrix $\mathbf{L}$ satisfying

$$\mathbf{L}\mathbf{X}\boldsymbol{\beta} = \boldsymbol{\beta}\,.$$

We compute

$$\begin{aligned}
\mathrm{var}\big(\tilde{\boldsymbol{\beta}}\big) &= \mathrm{var}\big(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}\big) \\
&= \mathrm{var}\big(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}\big) + \mathrm{var}\big(\hat{\boldsymbol{\beta}}\big) + 2\,\mathrm{cov}\big(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}},\ \hat{\boldsymbol{\beta}}\big)\,.
\end{aligned}$$

Denote

$$\mathbf{A} = \left(\mathbf{I} + a\mathbf{1}\mathbf{1}^T\right) \quad \textit{and} \quad \mathbf{C} = \mathbf{I} + c\mathbf{1}\mathbf{1}^T\,.$$

Compute

$$\mathbf{A}\mathbf{C} = \mathbf{I} + a\mathbf{1}\mathbf{1}^T + c\mathbf{1}\mathbf{1}^T + ac\mathbf{1}\mathbf{1}^T\mathbf{1}\mathbf{1}^T = \mathbf{I} + (a + c + nac)\mathbf{1}\mathbf{1}^T = \mathbf{I}\,.$$

Taking into account that $\mathrm{cov}\,(\mathbf{Y}, \mathbf{Y}) = \sigma^2\mathbf{A}$ we get

$$\begin{aligned}
\mathrm{cov}\left(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\beta}}\right) &= \mathrm{cov}\left(\left(\mathbf{L} - \left(\mathbf{X}^T\mathbf{C}\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{C}\right)\mathbf{Y},\ \left(\mathbf{X}^T\mathbf{C}\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{C}\mathbf{Y}\right) \\
&= \sigma^2\left(\mathbf{L} - \left(\mathbf{X}^T\mathbf{C}\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{C}\right)\mathbf{A}\mathbf{C}\mathbf{X}\left(\mathbf{X}^T\mathbf{C}\mathbf{X}\right)^{-1} \\
&= \sigma^2\left(\mathbf{L}\mathbf{X} - \mathbf{I}\right)\left(\mathbf{X}^T\mathbf{C}\mathbf{X}\right)^{-1} \\
&= 0\,.
\end{aligned}$$

The conclusion follows the same way as in the proof of the standard Gauss/Markov theorem.

b. (10) Suggest an unbiased estimator for the parameter σ^2. Explain why it is unbiased.

Solution: one possibility is to use residuals. Denote

$$\hat{\boldsymbol{\epsilon}} = \begin{bmatrix} \hat{\epsilon}_1 \\ \hat{\epsilon}_2 \\ \vdots \\ \hat{\epsilon}_n \end{bmatrix} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}\,.$$

We have

$$\hat{\boldsymbol{\epsilon}} = \Big(\mathbf{I} - \mathbf{X}\big(\mathbf{X}^T\mathbf{C}\mathbf{X}\big)^{-1}\mathbf{X}^T\mathbf{C}\Big)\boldsymbol{\epsilon}$$

and

$$\begin{aligned}
\sum_{k=1}^n \hat{\epsilon}_k^2 &= \big(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}\big)^T\big(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}\big) \\
&= \boldsymbol{\epsilon}^T\Big(\mathbf{I} - \mathbf{C}\mathbf{X}\big(\mathbf{X}^T\mathbf{C}\mathbf{X}\big)^{-1}\mathbf{X}^T\Big)\Big(\mathbf{I} - \mathbf{X}\big(\mathbf{X}^T\mathbf{C}\mathbf{X}\big)^{-1}\mathbf{X}^T\mathbf{C}\Big)\boldsymbol{\epsilon} \\
&= \mathrm{Sl}\Big[\Big(\mathbf{I} - \mathbf{C}\mathbf{X}\big(\mathbf{X}^T\mathbf{C}\mathbf{X}\big)^{-1}\mathbf{X}^T\Big)\Big(\mathbf{I} - \mathbf{X}\big(\mathbf{X}^T\mathbf{C}\mathbf{X}\big)^{-1}\mathbf{X}^T\mathbf{C}\Big)\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T\Big]\,.
\end{aligned}$$

Since $\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T = \sigma^2\mathbf{A}$ *we get*

$$\begin{aligned}
E\left(\sum_{k=1}^n \hat{\epsilon}_k^2\right) &= \sigma^2\,\mathrm{Sl}\Big[\Big(\mathbf{I} - \mathbf{C}\mathbf{X}\big(\mathbf{X}^T\mathbf{C}\mathbf{X}\big)^{-1}\mathbf{X}^T\Big)\Big(\mathbf{I} - \mathbf{X}\big(\mathbf{X}^T\mathbf{C}\mathbf{X}\big)^{-1}\mathbf{X}^T\mathbf{C}\Big)\mathbf{A}\Big] \\
&= \sigma^2\,\mathrm{Sl}\big(\mathbf{A} - \mathbf{X}(\mathbf{X}^T\mathbf{C}\mathbf{X})^{-1}\mathbf{X}^T\big)\,.
\end{aligned}$$

It follows that

$$\hat{\sigma}^2 = \frac{1}{\mathrm{Sl}\big(\mathbf{A} - \mathbf{X}(\mathbf{X}^T\mathbf{C}\mathbf{X})^{-1}\mathbf{X}^T\big)}\sum_{k=1}^n \hat{\epsilon}_k^2$$

is an unbiased estimator of σ^2.