Examination paper for TMA4295 Statistical inference.
Date: 13.08.26
Time: 0900 - 1300

Course contact: Håkon Tjelmeland.
Present at the exam location: Yes, at approximately 11.00.

Permitted examination support material: C: Tabeller og formler i statistikk, Akademika; A specific basic calculator.

OTHER INFORMATION

Read the questions carefully and make your own assumptions. In your answers, explain clearly what assumptions you have made and how you have understood or limited the assignment.

If there are direct errors or omissions in the assignment set and you cannot make your own assumptions, please refer to the information about complaints regarding formal errors on the NTNU website “Explanation of grades and appeals”.

SPECIFIC INFORMATION FOR YOUR COURSE

All answers should be justified. In the evaluation it will be emphasised that it from the solution is easy to understand the reasoning leading to the results. In particular, it will be emphasised that the solution includes all natural intermediate calculations.

Paper drawings: The answers of all problems should be given on handwritten sheets.
At the bottom of the question, you will find a seven-digit code. Fill in this code in the top left corner of the sheets you wish to submit. We recommend that you do this during the exam. If you require access to the codes after the examination time ends, click “Show submission”.

Please note that only blue or black ballpoint pens are allowed

You are responsible for filling in the correct codes on the handwritten sheets. Therefore, read the cover sheet carefully. The Examination Office cannot guarantee that incorrectly completed sheets will be added to your assignment.

Weighting: The problem set consists of eight problems (1, 2, 3a, 3b, 4a, 4b, 4c and 4d). In the evaluation, equal weight will be given to each of these eight problems.

Withdrawing from the exam:
If you wish to submit a blank test/withdraw from the exam for another reason, go to the “hamburger menu” in the top right-hand corner and click “Submit blank”. This cannot be undone, even if the test is still open.

Access to your answers:
After the exam, you can find your answers under previous tests in Inspera. Be aware that it may take a working day until any paper drawings material is available in “previous tests”.

1 **Introduction**: Let $X_1, \ldots, X_n$ be independent and identically gamma distributed random variables, i.e. the density function for each of these random variables is

$$f(x; \alpha, \beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} \exp\left\{-\frac{x}{\beta}\right\}; x \geq 0,$$

where $\alpha > 0$ and $\beta > 0$ are assumed to be unknown parameters. As usual we let

$$\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i \quad \text{and} \quad S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2.$$

Problem: Find the moment estimators for α and β based on $X_1, \ldots, X_n$.

In particular, find formulas for these moment estimators as functions of $\bar{X}$ and S^2.

Maximum marks: 10

2 **Introduction**: Let $X_1, \ldots, X_n$ be independent and identically normal distributed random variables with mean zero and some variance θ.

To estimate θ based on $X_1, \ldots, X_n$ we consider the estimator

$$\theta^2 = k_n \sum_{i=1}^n X_i^2,$$

where k_n is some constant (that may be a function of n).

Problem: Using mean square error as a criterion, find how to choose k_n to make $\hat{\theta}$ the best estimator for θ.

Maximum marks: 10

3(a) **Introduction**: Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables with density function

$$f(x|\theta) = \begin{cases} \frac{2x}{\theta}\exp\left\{-\frac{x^2}{\theta}\right\}, & \text{for } x > 0, \\ 0, & \text{otherwise,} \end{cases}$$

where $\theta > 0$ is un unknown parameter. Considering the problem of estimating the value of θ in the Bayesian framework, we let the prior for θ be given by

$$\pi(\theta) = \begin{cases} \frac{b^a}{\Gamma(a)}\frac{\exp\{-\frac{b}{\theta}\}}{\theta^{a+1}}, & \text{for } \theta > 0, \\ 0, & \text{otherwise,} \end{cases}$$

where $a > 0$ and $b > 0$ are assumed known constants. This prior distribution is an inverse gamma distribution and in this problem we denote it by $\mathrm{IG}(a, b)$.

In this problem you can use (without proof) that the mean and the variance of an $\mathrm{IG}(a, b)$ distribution are given by

$$E[\theta] = \frac{b}{a-1} \text{ when } a > 1,$$

and

$$\mathrm{Var}[\theta] = \frac{b^2}{(a-1)^2(a-2)} \text{ when } a > 2,$$

respectively.

Problem: Show that the posterior distribution of θ given $x_1, x_2, \ldots, x_n$ is the $\mathrm{IG}\left(a + n, b + \sum_{i=1}^n x_i^2\right)$ distribution.

Maximum marks: 10

3(b) **Introduction**: Consider the following two loss functions,

$$L_1(\theta, \hat{\theta}) = (\hat{\theta} - \theta)^2,$$
$$L_2(\theta, \hat{\theta}) = \theta \cdot (\hat{\theta} - \theta)^2.$$

Problem: Find the Bayes estimator associated to each of these two loss functions.

Maximum marks: 10

4(a) **Introduction**: A large company has a certain number of machines of a particular type. Whenever an error occurs in one of these machines, it is classified as either *critical* or *non-critical*. The number of critical failures in one specific year we denote by U_1 and the number of non-critical failures in the same year we denote by V_1. We assume U_1 and V_1 to be independent and Poisson distributed with

$$E[U_1] = \mu, \quad E[V_1] = \varphi.$$

The following year the maintenance routines the company use are changed. We let U_2 and V_2 denote the number of critical and non-critical failures in that year, respectively. Moreover, we assume U_2 and V_2 to be independent of each other and independent of U_1 and V_1, and Poisson distributed with

$$E[U_2] = a\mu, \quad E[V_2] = a\varphi,$$

where $a > 0$ is a parameter representing the change in the tendency of the machine to error. Note in particular that we assume this factor to be the same for critical and non-critical failures.

So in the situation we observe the vector $\mathbf{X} = (U_1, V_1, U_2, V_2)$ and the vector of unknown parameters is $\theta = (\mu, \varphi, a)$, where $\mu, \varphi, a > 0$.

Problem: Show that the probability mass function of $\mathbf{X}$ can be expressed as an exponential family with $k = 3$, $c(\theta) = \exp\{-(\mu+\varphi)(a+1)\}$, $w_1(\theta) = \ln\mu$, $w_2(\theta) = \ln\varphi$, $w_3(\theta) = \ln a$, and $t_1(\mathbf{x}) = u_1 + u_2$, $t_2(\mathbf{x}) = v_1 + v_2$, $t_3(\mathbf{x}) = u_2 + v_2$.

Identify an expression for the corresponding function $h(\mathbf{x})$ in the exponential family formulation.

Identify also a sufficient statistic for θ based on $\mathbf{X}$.

Maximum marks: 10

4(b) **Problem**: Derive the maximum likelihood estimator for θ, $\hat{\theta} = (\hat{\mu}, \hat{\varphi}, \hat{a})$, and show that it can be expressed as

$$\hat{\mu} = \frac{(U_1+U_2)(U_1+V_1)}{U_1+V_1+U_2+V_2}, \quad \hat{\varphi} = \frac{(V_1+V_2)(U_1+V_1)}{U_1+V_1+U_2+V_2}, \quad \hat{a} = \frac{U_2+V_2}{U_1+V_1}.$$

Maximum marks: 10

4(c) **Introduction**: Assume we want to use the observed value of $\mathbf{X} = (U_1, V_2, U_2, V_2)$ to test

$$H_0 : a = 1 \quad \text{versus} \quad H_1 : a \neq 1.$$

Problem: Derive a likelihood ratio test statistic for this situation and specify the associated rejection criterion. **Note**: You do not need to decide the critical value associated to a given significance level α.

Maximum marks: 10

4(d) **Introduction**: Consider now the problem of estimating the parameter

$$\tau(\theta) = \mu + \varphi.$$

Problem: Indentify the uniform minimum variance unbiased estimator (UMVUE) for $\tau(\theta)$ based on $\mathbf{X} = (U_1, V_1, U_2, V_2)$. **Note**: You both need to find a formula for the estimator, and give arguments for why it is UMVUE.

Maximum marks: 10

Formulas from Casella and Berger

Leibnitz rule:

$$\frac{d}{d\theta}\int_{a(\theta)}^{b(\theta)} f(x,\theta)dx = f(b(\theta),\theta)\cdot\frac{db}{d\theta}(\theta) - f(a(\theta),\theta)\cdot\frac{da}{d\theta}(\theta) + \int_{a(\theta)}^{b(\theta)}\frac{\partial f}{\partial\theta}(x,\theta)dx$$

Double expectation and double variance:

$$E[X] = E[E[X|Y]], \quad \mathbf{Var}[X] = E[\mathbf{Var}[X|Y]] + \mathbf{Var}[E[X|Y]]$$

Chebychev's inequality:

$$P(g(X)\geq r) \leq \frac{E[g(X)]}{r}$$

Cauchy-Schwarz inequality:

$$|E[XY]| \leq E[|XY|] \leq \sqrt{E\left[|X|^2\right]E\left[|Y|^2\right]}$$

Jensen's inequality:

If $g(x)$ is a convex function we have

$$E[g(X)] \geq g(E[X]).$$

Equality holds if and only if, for every line $a+bx$ that is a tangent to $g(x)$ at $x=E[X]$,
$P(g(X)=a+bX)=1.$

Delta method:

Let Y_n be a sequence of random variables that satisfy $\sqrt{n}(Y_n-\theta)\longrightarrow N(0,\sigma^2)$ in distribution. For a given function $g(\theta)$ and a specific value of θ, suppose $g'(\theta)$ exists and is not 0. Then

$$\sqrt{n}(g(Y_n)-g(\theta))\longrightarrow N(0,\sigma^2(g'(\theta))^2)$$

in distribution.

Second-order delta method:

Let Y_n be a sequence of random variables that satisfy $\sqrt{n}(Y_n-\theta)\longrightarrow N(0,\sigma^2)$ in distribution. For a given function $g(\theta)$ and a specific value of θ, suppose $g'(\theta)=0$ and $g''(\theta)$ exists and is not 0. Then

$$\sqrt{n}(g(Y_n)-g(\theta))\longrightarrow \frac{\sigma^2 g''(\theta)}{2}\chi_1^2$$

in distribution.

Exponential family:

$$f(x|\theta) = h(x)c(\theta)\exp\left\{\sum_{i=1}^{k} w_i(\theta)t_i(x)\right\}$$

Expectation and variance formulas for an exponential family:

When X is a random variable from an exponential family, we have

$$E\left[\sum_{i=1}^{k} \frac{\partial w_i(\theta)}{\partial \theta_j} t_i(X)\right] = -\frac{\partial}{\partial \theta_j} \ln c(\theta),$$
$$\mathrm{Var}\left[\sum_{i=1}^{k} \frac{\partial w_i(\theta)}{\partial \theta_j} t_i(X)\right] = -\frac{\partial^2}{\partial \theta_j^2} \ln c(\theta) - E\left[\sum_{i=1}^{k} \frac{\partial^2 w_i(\theta)}{\partial \theta_j^2} t_i(X)\right]$$

Complete statistic:

Let $f(t|\theta)$ be a family of pdfs/pmfs for a statistic $T(\mathbf{X})$. The family of probability distributions is called complete if $E_\theta[g(T)] = 0$ for all θ implies $P_\theta(g(T) = 0) = 1$ for all θ. Equivalently, $T(\mathbf{X})$ is called a complete statistic.

Basu's theorem:

If $T(\mathbf{X})$ is a complete and minimal sufficient statistic, then $T(\mathbf{X})$ is independent of every ancillary statistic.

Mean squared error:

The mean squared error (MSE) of an estimator W of a parameter θ is the function of θ defined by $E[(W-\theta)^2]$.

Cramer-Rao inequality:

$$\mathrm{Var}_\theta[W(\mathbf{X})] \geq \frac{\left(\frac{d}{d\theta} E_\theta[W(\mathbf{X})]\right)^2}{E_\theta\left[\left(\frac{\partial}{\partial \theta} \ln f(\mathbf{X}|\theta)\right)^2\right]}$$

Rao-Blackwell:

Let W be any unbiased estimator of $\tau(\theta)$, and let T be a sufficient statistic for θ. Define

$$\phi(T) = E[W|T].$$

Then $E_\theta[\phi(T)] = \tau(\theta)$ and $\mathrm{Var}_\theta[\phi(T)] \leq \mathrm{Var}_\theta[W]$ for all θ; that is, $\phi(T)$ is a uniformly better unbiased estimator of $\tau(\theta)$.

Risk function:

$$R(\theta, \delta) = E_\theta[L(\theta, \delta(\mathbf{X}))]$$

Bayes risk:

$$r(\delta) = E[R(\theta, \delta)]$$

Power function:

$$\beta(\theta) = P_\theta(\mathbf{X} \in R)$$

Likelihood ratio test:

$$\lambda(\mathbf{x}) = \frac{\sup_{\theta \in \Theta_0} L(\theta|\mathbf{x})}{\sup_{\theta \in \Theta} L(\theta|\mathbf{x})}$$

Size α test:

$$\sup_{\theta \in \Theta_0} \beta(\theta) = \alpha$$

Level α test:

$$\sup_{\theta \in \Theta_0} \beta(\theta) \leq \alpha$$

Neyman-Pearson lemma:

Consider testing $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1,$ where the pdf or pmf corresponding to θ_i is $f(\mathbf{x}|\theta_i), i = 0, 1,$ using a test with rejection region R that satisfy

$\mathbf{x} \in R$ if $f(\mathbf{x}|\theta_1) > k\, f(\mathbf{x}|\theta_0)$

and

$\mathbf{x} \in R^C$ if $f(\mathbf{x}|\theta_1) < k\, f(\mathbf{x}|\theta_0)$

for some $k \geq 0,$ and

$$\alpha = P_{\theta_0}(\mathbf{X} \in R).$$

Then any test that satisfy these three conditions is an UMP level α test.

P-value:

A p-value $p(\mathbf{X})$ is a test statistic satisfying $0 \leq p(\mathbf{x}) \leq 1$ for every sample point $\mathbf{x}.$ Small values of $p(\mathbf{X})$ give evidence that H_1 is true. A p-value is valid if, for every $\theta \in \Theta_o$ and every $0 \leq \alpha \leq 1,$

$$P_\theta(p(\mathbf{X}) \leq \alpha) \leq \alpha.$$

Valid p-value:

Let $W(\mathbf{X})$ be a test statistic such that large values of W give evidence that H_1 is true. For each sample point $\mathbf{x}$, define

$$p(\mathbf{x}) = \sup_{\theta \in \Theta_0} P_\theta(W(\mathbf{X} \geq W(\mathbf{x})).$$

Then $p(\mathbf{X})$ is a valid p-value.

Information number:

$$\mathbf{I}(\theta) = E_\theta\left[\left(\frac{\partial}{\partial\theta}\ln f(\mathbf{X}|\theta)\right)^2\right] = -E_\theta\left[\frac{\partial^2}{\partial\theta^2}\ln f(\mathbf{X}|\theta)\right]$$

Score test:

$$S(\theta) = \frac{\partial}{\partial\theta}\ln f(\mathbf{X}|\theta)$$