NTNU
Norwegian University of
Science and Technology

Department of Mathematical Sciences

Examination paper for
TMA4295 Statistical inference — solution sketch

Academic contact during examination: Håkon Tjelmeland

Phone: 48 22 18 96

Academic contact present at the exam location: Yes, at approximately 11.00

Examination date: August 2026

Examination time (from–to): 09.00-13.00

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

Other information:
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.
Please note that only blue or black ballpoint pens are allowed.

Language: English

Number of pages: 8

Number of pages enclosed: 0

Informasjon om trykking av eksamensoppgave	
Originalen er:	
1-sidig ☐	**2-sidig** ☒
sort/hvit ☒	**farger** ☐
skal ha flervalgskjema ☐	

Students will find the examination results in Studentweb. Please contact the department if you have questions about your results. The Examinations Office will not be able to answer this.

Problem 1

For the gamma distribution we have that

$$E[X] = \alpha\beta \qquad \text{and} \qquad \text{Var}[X] = \alpha\beta^2.$$

This gives

$$E[X^2] = \text{Var}[X] + E[X]^2 = \alpha\beta^2 + (\alpha\beta)^2 = \alpha\beta^2(\alpha+1).$$

The moment estimators $\widehat{\alpha}$ and $\widehat{\beta}$ are thereby given by the expressions

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \widehat{\alpha}\widehat{\beta}, \tag{1}$$

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 = \widehat{\alpha}\widehat{\beta}^2(\widehat{\alpha}+1). \tag{2}$$

To solve this system of equations one may start by using (1) to get

$$\widehat{\alpha} = \frac{\overline{X}}{\widehat{\beta}}. \tag{3}$$

Inserting this expression for $\widehat{\alpha}$ into (2) we get

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 = \frac{\overline{X}}{\widehat{\beta}} \cdot \widehat{\beta}^2\left(\frac{\overline{X}}{\widehat{\beta}} + 1\right) = \widehat{\beta}\overline{X} \qquad \Rightarrow \qquad \widehat{\beta} = \frac{\sum_{i=1}^{n} X_i^2}{n\overline{X}} - \overline{X}.$$

Inserting this expression for $\widehat{\beta}$ into (3) we get

$$\widehat{\alpha} = \overline{X} \cdot \frac{n\overline{X}}{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2} = \frac{n\overline{X}^2}{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2}.$$

The moment estimators are thereby

$$\underline{\underline{\widehat{\alpha} = \frac{n\overline{X}^2}{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2}, \qquad \widehat{\beta} = \frac{\sum_{i=1}^{n} X_i^2}{n\overline{X}} - \overline{X}.}}$$

To derive expression of $\widehat{\alpha}$ and $\widehat{\beta}$ as a function of $\overline{X}$ and S^2, we may start by noting that

$$\begin{aligned}\sum_{i=1}^{n}(X_i - \overline{X})^2 &= \sum_{i=1}^{n} X_i^2 - 2\overline{X}\sum_{i=1}^{n} X_i + n\overline{X}^2 \\ &= \sum_{i=1}^{n} X_i^2 - 2\overline{X} \cdot n\overline{X} + n\overline{X}^2 \\ &= \sum_{i=1}^{n} X_i^2 - n\overline{X}^2.\end{aligned}$$

Thereby we have

$$\begin{aligned}\sum_{i=1}^{n} X_i^2 &= \sum_{i=1}^{n}(X_i - \overline{X})^2 + n\overline{X}^2 \\ &= (n-1)S^2 + n\overline{X}^2.\end{aligned}$$

Inserting this into our expressions for $\widehat{\alpha}$ and $\widehat{\beta}$ we get

$$\widehat{\alpha} = \frac{n\overline{X}^2}{(n-1)S^2 + n\overline{X}^2 - n\overline{X}^2} = \frac{n}{n-1} \cdot \frac{\overline{X}^2}{S^2}$$

and

$$\widehat{\beta} = \frac{(n-1)S^2 + n\overline{X}^2}{n\overline{X}} - \overline{X} = \frac{(n-1)S^2 + n\overline{X} - n\overline{X}^2}{n\overline{X}} = \frac{n-1}{n} \cdot \frac{S^2}{\overline{X}}.$$

So we have shown that

$$\underline{\underline{\widehat{\alpha} = \frac{n}{n-1} \cdot \frac{\overline{X}^2}{S^2}, \qquad \widehat{\beta} = \frac{n-1}{n} \cdot \frac{S^2}{\overline{X}}.}}$$

Problem 2

The MSE for $\widehat{\theta}$ becomes

$$\mathrm{MSE}[\widehat{\theta}] = \mathrm{E}\left[\left(\widehat{\theta} - \theta\right)^2\right] = \mathrm{Var}[\widehat{\theta}] + \left(E[\widehat{\theta}] - \theta\right)^2.$$

To find the expected value and the variance of $\widehat{\theta}$ one can use that we know that

$$\frac{\widehat{\theta}}{k_n\theta} = \sum_{i=1}^{n}\left(\frac{X_i}{\sigma}\right)^2 \sim \chi^2_n.$$

So we get

$$\frac{E[\widehat{\theta}]}{k_n\theta} = E\left[\frac{\widehat{\theta}}{k_n\theta}\right] = n \qquad \Rightarrow \qquad E[\widehat{\theta}] = nk_n\theta$$

and

$$\frac{\mathrm{Var}[\widehat{\theta}]}{k_n^2\theta^2} = \mathrm{Var}\left[\frac{\widehat{\theta}}{k_n\theta}\right] = 2n \qquad \Rightarrow \qquad \mathrm{Var}[\widehat{\theta}] = 2nk_n^2\theta^2.$$

Inserting these expression in the expression we have for the MSE, we get

$$\begin{aligned}\text{MSE}[\widehat{\theta}] &= 2nk_n^2\theta^2 + (nk_n\theta - \theta)^2 \\ &= \theta^2\left(2nk_n^2 + n^2k_n^2 - 2nk_n + 1\right)\end{aligned}$$

To find for what value of k_n this expression in minimised we take the partial derivative with respect to k_n,

$$\frac{\partial \text{MSE}[\widehat{\theta}]}{\partial k_n} = \theta^2\left(4nk_n + 2n^2k_n - 2n\right).$$

Requiring this partial derivative to be equal to zero, we get

$$4nk_n + 2n^2k_n = 2n \qquad \Rightarrow \qquad \underline{\underline{k_n = \frac{1}{n+2}}}.$$

Problem 3

a) For the posterior distribution we get

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \pi(\theta)\cdot\prod_{i=1}^{n} f(x_i|\theta) \\ &\propto \frac{\exp\left\{-\frac{b}{\theta}\right\}}{\theta^{a+1}}\cdot\prod_{i=1}^{n}\left[\frac{1}{\theta}\exp\left\{-\frac{x_i^2}{\theta}\right\}\right] \\ &= \frac{\exp\left\{-\frac{b+\sum_{i=1}^{n}x_i^2}{\theta}\right\}}{\theta^{a+n+1}},\end{aligned}$$

where the proportionality is as a function of θ. We observe that this expression is again proportional to the density of an inverse gamma distribution with parameters

$$\widetilde{a} = a + n, \qquad \widetilde{b} = b + \sum_{i=1}^{n} x_i^2.$$

Thereby we must have

$$\underline{\underline{\theta|\mathbf{x} \sim \text{IG}\left(a+n, b+\sum_{i=1}^{n}x_i^2\right)}}.$$

b) The $L_1(\theta, \widehat{\theta})$ is the quadratic loss function, for which we know the associated Bayes estimator is the posterior mean, i.e.

$$\widehat{\theta} = E[\theta|\mathbf{X}] = \frac{\widetilde{b}}{\widetilde{a}-1} = \underline{\underline{\frac{b+\sum_{i=1}^n X_i^2}{a+n-1}}}.$$

For the $L_2(\theta, \widehat{\theta})$ loss function we start by finding the expected loss,

$$\begin{aligned} E[L(\theta,t)|\mathbf{x}] &= E[\theta\cdot(t-\theta)^2|\mathbf{x}] = \int_0^\infty \theta(t-\theta)^2\pi(\theta|\mathbf{x})d\theta \\ &= \int_0^\infty (t^2\theta - 2t\theta^2 + \theta^3)\pi(\theta|\mathbf{x})d\theta \\ &= t^2E[\theta|\mathbf{x}] - 2tE[\theta^2|\mathbf{x}] + E[\theta^3|\mathbf{x}]. \end{aligned}$$

To find the Bayes estimator we should minimise this expected loss with respect to t, so we start by deriving the partial derivative

$$\frac{\partial E[L(\theta,t)]}{\partial t} = 2tE[\theta|\mathbf{x}] - 2E[\theta^2|\mathbf{x}].$$

Equating this derivative to zero we get

$$\frac{\partial E[L(\theta,t)]}{\partial t} = 0 \qquad \Rightarrow \qquad t = \frac{E[\theta^2|\mathbf{x}]}{E[\theta|\mathbf{x}]}.$$

From the given formulas we have

$$E[\theta|\mathbf{x}] = \frac{\widetilde{b}}{\widetilde{a}-1}$$

and

$$\begin{aligned} E[\theta^2|\mathbf{x}] &= \mathrm{Var}[\theta|\mathbf{x}] + E[\theta|\mathbf{x}]^2 \\ &= \frac{\widetilde{b}^2}{(\widetilde{a}-1)^2(\widetilde{a}-2)} + \frac{\widetilde{b}^2}{(\widetilde{a}-1)^2} \\ &= \frac{\widetilde{b}^2(1+\widetilde{a}-2)}{(\widetilde{a}-1)^2(\widetilde{a}-2)} \\ &= \frac{\widetilde{b}^2}{(\widetilde{a}-1)(\widetilde{a}-2)}. \end{aligned}$$

Thereby the expected loss is minimised for

$$\begin{aligned} t = \frac{E[\theta^2|\mathbf{x}]}{E[\theta|\mathbf{x}]} &= \frac{\widetilde{b}}{\widetilde{a}-2} \\ &= \frac{b+\sum_{i=1}^n x_i^2}{a+n-2}. \end{aligned}$$

So the Bayes estimator for loss function $L_2(\theta, \widehat{\theta})$ is

$$\underline{\underline{\theta^\star = \frac{b + \sum_{i=1}^n X_i^2}{a + n - 2}}}.$$

Problem 4

a) We have

$$\begin{aligned}
f(\mathbf{x}|\theta) &= f(u_1|\mu) f(v_1|\varphi) f(u_2|a\mu) f(v_2|a\varphi) \\
&= \frac{\mu^{u_1}}{u_1!} e^{-\mu} \mathbf{I}(u_1 \in \{0, 1, \ldots\}) \times \frac{\varphi^{v_1}}{v_1!} e^{-\varphi} \mathbf{I}(v_1 \in \{0, 1, \ldots\}) \\
&\times \frac{(a\mu)^{u_2}}{u_2!} e^{-a\mu} \mathbf{I}(u_2 \in \{0, 1, \ldots\}) \times \frac{(a\varphi)^{v_2}}{v_2!} e^{-a\varphi} \mathbf{I}(v_2 \in \{0, 1, \ldots\}) \\
&= \left[\frac{\mathbf{I}\{u_1, v_1, u_2, v_2 \in \{0, 1, \ldots\}}{u_1! v_1! u_2! v_2!} \right] \times \left[\exp\left\{ -(\mu + \varphi + a\mu + a\varphi) \right\} \right] \\
&\times \exp\left\{ u_1 \ln \mu + v_1 \ln \varphi + u_2 (\ln a + \ln \mu) + v_2 (\ln a + \ln \varphi) \right\} \\
&= \left[\frac{\mathbf{I}\{u_1, v_1, u_2, v_2 \in \{0, 1, \ldots\}}{u_1! v_1! u_2! v_2!} \right] \times \left[\exp\left\{ -(\mu + \varphi)(a + 1) \right\} \right] \\
&\times \exp\left\{ (u_1 + u_2) \ln \mu + (v_1 + v_2) \ln \varphi + (u_2 + v_2) \ln a \right\} \\
&= h(\mathbf{x}) c(\theta) \exp\left\{ w_1(\theta) t_1(\mathbf{x}) + w_2(\theta) t_2(\mathbf{x}) + w_3(\theta) t_3(\mathbf{x}) \right\},
\end{aligned}$$

where

$$\begin{aligned}
h(\mathbf{x}) &= \frac{\mathbf{I}\{u_1, v_1, u_2, v_2 \in \{0, 1, \ldots\}}{u_1! v_1! u_2! v_2!}, \\
c(\theta) &= \exp\left\{ -(\mu + \varphi)(a + 1) \right\}, \\
w_1(\theta) &= \ln \mu, \\
w_2(\theta) &= \ln \varphi, \\
w_3(\theta) &= \ln a, \\
t_1(\mathbf{x}) &= u_1 + u_2, \\
t_2(\mathbf{x}) &= v_1 + v_2, \\
t_3(\mathbf{x}) &= u_2 + v_2.
\end{aligned}$$

So thereby it can be expressed as an exponential family with $k = 3$ and the specified functions. In particular we have that

$$\underline{\underline{h(\mathbf{x}) = \frac{\mathbf{I}\{u_1, v_1, u_2, v_2 \in \{0, 1, \ldots\}}{u_1! v_1! u_2! v_2!}}},$$

and a sufficient statistic is $t(\mathbf{X}) = (t_1(\mathbf{X}), t_2(\mathbf{X}), t_3(\mathbf{X})) = \underline{\underline{(U_1 + U_2, V_1 + V_2, U_2 + V_2)}}$.

b) Using the exponential family expression for $f(\mathbf{x}|\theta)$ we derived in **a**), the log-likelihood function becomes

$$\ell(\theta) = \ln h(\mathbf{x}) - (\mu+\varphi)(a+1) + (u_1+u_2)\ln\mu + (v_1+v_2)\ln\varphi + (u_2+v_2)\ln a.$$

The three partial derivatives become

$$\begin{aligned}
\frac{\partial\ell}{\partial\mu} &= -(a+1) + \frac{u_1+u_2}{\mu}, \\
\frac{\partial\ell}{\partial\varphi} &= -(a+1) + \frac{v_1+v_2}{\varphi}, \\
\frac{\partial\ell}{\partial a} &= -(\mu+\varphi) + \frac{u_2+v_2}{a}.
\end{aligned}$$

Setting $\frac{\partial\ell}{\partial\mu} = 0$ and solving for μ, we get

$$\mu = \frac{u_1+u_2}{a+1}, \tag{4}$$

and setting $\frac{\partial\ell}{\partial\varphi} = 0$ and solving for φ, we correspondingly get

$$\varphi = \frac{v_1+v_2}{a+1}. \tag{5}$$

Setting $\frac{\partial\ell}{\partial a} = 0$ and solving for a, and thereafter inserting the above expressions for μ and φ, we get

$$\begin{aligned}
a = \frac{u_2+v_2}{\mu+\varphi} &= \frac{u_2+v_2}{\frac{u_1+u_2}{a+1} + \frac{v_1+v_2}{a+1}} = \frac{u_2+v_2}{u_1+v_1+u_2+v_2}\cdot(a+1). \\
\Rightarrow a\left(1 - \frac{u_2+v_2}{u_1+v_1+u_2+v_2}\right) &= \frac{u_2+v_2}{u_1+v_1+u_2+v_2} \\
\Rightarrow a\cdot\frac{u_1+v_1}{u_1+v_1+u_2+v_2} &= \frac{u_2+v_2}{u_1+v_1+u_2+v_2} \\
\Rightarrow a &= \frac{u_2+v_2}{u_1+v_1}.
\end{aligned}$$

This gives

$$a+1 = \frac{u_2+v_2}{u_1+v_1} + 1 = \frac{u_1+v_1+u_2+v_2}{u_1+v_1}, \tag{6}$$

and inserting this expression in the expression we found for μ and φ above, we get

$$\mu = \frac{(u_1+u_2)(u_1+v_1)}{u_1+v_1+u_2+v_2}$$

and

$$\varphi = \frac{(v_1+v_2)(u_1+v_1)}{u_1+v_1+u_2+v_2}.$$

The maximum likelihood estimators are thereby

$$\underline{\underline{\widehat{\mu} = \frac{(U_1+U_2)(U_1+V_1)}{U_1+V_1+U_2+V_2}, \qquad \widehat{\varphi} = \frac{(V_1+V_2)(U_1+V_1)}{U_1+V_1+U_2+V_2}, \qquad \widehat{a} = \frac{U_2+V_2}{U_1+V_1}.}}$$

c) The likelihood ratio test statistic is given from

$$\lambda(\mathbf{x}) = \frac{\sup_{\mu,\varphi>0} f(\mathbf{x}|(\mu,\varphi,1))}{\sup_{\mu,\varphi,a>0} f(\mathbf{x}|\theta)} = \frac{\exp\{\ell((\widetilde{\mu},\widetilde{\varphi},1))\}}{\exp\{\ell((\widehat{\mu},\widehat{\varphi},\widehat{a}))\}},$$

where $\widehat{\mu},\widehat{\varphi},\widehat{a}$ are the maximum likelihood estimates found above, and $\widetilde{\mu}$ and $\widetilde{\varphi}$ are the parameter values that maximise $\ell(\mu,\varphi,1)$. From (4) and (5) we see that

$$\widetilde{\mu} = \frac{u_1+u_2}{2} \qquad \text{and} \qquad \widetilde{\varphi} = \frac{v_1+v_2}{2}.$$

Introducing the notation $s = u_1+v_1+u_2+v_2$, we can observe that

$$\widehat{\mu}+\widehat{\varphi} = u_1+v_1 \qquad \text{and} \qquad \widetilde{\mu}+\widetilde{\varphi} = \frac{s}{2}.$$

Using this and (6) we get

$$\begin{aligned}\ell(\widehat{\mu},\widehat{\varphi},\widehat{a}) &= \ln h(\mathbf{x}) - (u_1+v_1)\cdot\frac{s}{u_1+v_1} + (u_1+u_2)\left(\ln(u_1+u_2)+\ln(u_1+v_1)-\ln s\right)\\ &\quad + (v_1+v_2)\left(\ln(v_1+v_2)+\ln(u_1+v_1)-\ln s\right) + (u_2+v_2)\left(\ln(u_2+v_2)-\ln(u_1+v_1)\right)\\ &= \ln h(\mathbf{x}) - s + (u_1+u_2)\ln(u_1+u_2) + (v_1+v_2)\ln(v_1+v_2)\\ &\quad + (u_1+v_1)\ln(u_1+v_1) + (u_2+v_2)\ln(u_2+v_2) - s\ln s\end{aligned}$$

and

$$\begin{aligned}\ell(\widetilde{\mu},\widetilde{\varphi},1) &= \ln h(\mathbf{x}) - \frac{s}{2}\cdot 2 + (u_1+u_2)\left(\ln(u_1+u_2)-\ln 2\right) + (v_1+v_2)\left(\ln(v_1+v_2)-\ln 2\right)\\ &= \ln h(\mathbf{x}) - s + (u_1+u_2)\ln(u_1+u_2) + (v_1+v_2)\ln(v_1+v_2) - s\ln 2.\end{aligned}$$

Thereby we get

$$\begin{aligned}\ln\lambda(\mathbf{x}) &= \ell(\widetilde{\mu},\widetilde{\varphi},1) - \ell(\widehat{\mu},\widehat{\varphi},\widehat{a})\\ &= -s\ln 2 - (u_1+v_1)\ln(u_1+v_1) - (u_2+v_2)\ln(u_2+v_2) + s\ln s\\ &= s\ln\left(\frac{s}{2}\right) - (u_1+v_1)\ln(u_1+v_1) - (u_2+v_2)\ln(u_2+v_2).\end{aligned}$$

So a likelihood ratio test statistic is given by

$$\begin{aligned}\ln \lambda(\mathbf{X}) = (U_1 + V_1 + U_2 + V_2) \ln \left(\frac{U_1 + V_1 + U_2 + V_2}{2} \right) \\ - (U_1 + V_1) \ln(U_1 + V_1) - (U_2 + V_2) \ln(U_2 + V_2),\end{aligned}$$

where we should reject H_0 if $\ln \lambda(\mathbf{X}) \leq c$ for some critical value c.

d) A naïve estimator for $\tau(\theta)$ is

$$\tau(\widehat{\theta}) = \widehat{\mu} + \widehat{\varphi} = U_1 + V_1.$$

One should note that

$$E[\tau(\widehat{\theta})] = E[U_1] + E[V_1] = \mu + \varphi = \tau(\theta),$$

so $\tau(\widehat{\theta})$ is an unbiased estimator for $\tau(\theta)$. Moreover, using the sufficient statisics found in a) we have

$$\tau(\widehat{\theta}) = U_1 + V_1 = (U_1 + U_2) + (V_1 + V_2) - (U_2 + V_2) = t_1(\mathbf{X}) + t_2(\mathbf{X}) - t_3(\mathbf{X}),$$

so $\tau(\widehat{\theta})$ is a function of the sufficient statistic $t(\mathbf{X})$. Since $\{(w_1(\theta), w_2(\theta), w_3(\theta) | \theta \in \mathbf{R}^3_+\} = \{(\ln \mu, \ln \varphi, \ln a) | \mu, \varphi, a > 0\} = \mathbf{R}^3$ contains an open set in $\mathbf{R}^3$, general theory for exponential families gives that the sufficient statistic $t(\mathbf{X})$ is complete.

Moreover, general theory (Theorem 7.3.23 in our textbook) also gives that any unbiased estimator $\widehat{\tau}$ of τ that is only a function of a sufficient and complete statistic is UMVUE for τ. Thereby $\tau(\widehat{\theta}) = U_1 + V_1$ is UMVUE for $\tau(\theta)$.