Delft University of Technology
Faculty EEMCS, Department DIAM

van Mourik Broekmanweg 6
2628 XE Delft

Practice Midterm AM2080

This written midterm exam contains 5 questions, each question counts for 20% of the final grade of the written test. You are only allowed to use the two sheets with information on probability distributions and R-commands. You are not allowed to use any books or notes.

1. For $-\infty < s < t < \infty$, the distribution function of a uniform distribution on the interval $[s, t]$, is given by

$$F(x) = \begin{cases} 0 & , x < s \\ \dfrac{x-s}{t-s} & , s \leq x \leq t \\ 1 & , x > t \end{cases}$$

   (a) Does the uniform distribution with parameters s and t belong to a location-scale family $\{F_{a,b} : a \in \mathbb{R}, b > 0\}$ associated with the standard uniform distribution on $[0, 1]$? If yes, then specify the location parameter a and scale parameter b. If no, then explain why.

   (b) Derive the expression of the quantile function corresponding to the uniform distribution with parameters s and t.

2. Let $X_1, \ldots X_n$ be independent random variables with an exponential distribution with parameter $\lambda > 0$.

   (a) Show that $X_{(1)} = \min\{X_1, \ldots, X_n\}$ has an exponential distribution with parameter $n\lambda$.

   (b) Consider the following estimators for $1/\lambda$:

$$T_1 = nX_{(1)} \quad \text{and} \quad T_2 = \overline{X}$$

   Which of these estimators has the smallest mean square error?
   You may use that $X_1 + \cdots + X_n$ has a gamma distribution with parameters n and λ.

3. Let $X_1, \ldots, X_n$ be a sample from a distribution with probability density given by

$$p_\theta(x) = \frac{1}{2}\theta e^{-\theta|x|} \quad \text{for } x \in (-\infty, \infty)$$

   where $\theta > 0$ is an unknown parameter.

   (a) Determine the maximum likelihood estimator for θ.

(b) As prior distribution we choose a gamma distribution with parameters $\alpha > 0$ and $\lambda > 0$. Determine the Bayes estimator for θ with respect to this prior.

4. Let $X_1, \ldots, X_n$ be independent identically distributed random variables with a $N(\mu, 1)$ distribution. We test $H_0 : \mu \leq \mu_0$ against $\mu > \mu_0$ at significance level α_0 with test statistic $T = \sqrt{n}(\overline{X} - \mu_0)$. You may use that large values of T are in favor of $H_1 : \mu > \mu_0$ and that the critical region is given by $[\xi_{1-\alpha_0}, \infty)$, where $\xi_{1-\alpha_0}$ is the $(1 - \alpha_0)$-quantile of the standard normal distribution.

   (a) For an observed value $\overline{x}$, give the definition of the corresponding p-value and show that it can be written as $1 - \Phi(\sqrt{n}(\overline{x} - \mu_0))$, where Φ denotes the standard normal distribution function.

   (b) For a given alternative $\mu_1 > \mu_0$ we want the power function to be at least $1 - \beta$, for some $0 < \beta < 1 - \alpha_0$. Determine the minimal sample size for which this holds.

5. Let $X_1, \ldots, X_n$ be independent identically distributed with a $N(0, 1)$ distribution and let

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

Let O be an orthonormal $n \times n$-matrix, with first row

$$f_1 = \left(\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}}\right)$$

and define the column vector Y by $Y = OX$, where X is the column vector consisting of $X_1, \ldots, X_n$.

   (a) Show that

$$Y_1 = \sqrt{n}\,\overline{X} \quad \text{and} \quad \sum_{i=2}^{n} Y_i^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2.$$

   (b) Prove that $\overline{X}$ and S_X^2 are independent and that $(n-1)S_X^2$ has a χ^2_{n-1} distribution.

   You may use that for any orthonormal matrix O, the random vector $Y = OX$ has components $Y_1, \ldots, Y_n$ that are independent identically distributed with a $N(0, 1)$ distribution.

Solutions:

1a The distribution function of the $U[s,t]$ distribution satisfies

$$F(x) = F_{0,1}\left(\frac{x-s}{t-s}\right)$$

where $F_{0,1}$ is the distribution function of the standard uniform. This is of the form

$$F_{0,1}\left(\frac{x-a}{b}\right)$$

with $a = s$ and $b = t - s$. So yes, the $U[s,t]$ distribution is a member of a location family associated with the standard uniform with location parameter $a = s$ and scale parameter $b = t - s$.

1b For $0 < \alpha < 1$ solve

$$\alpha = \frac{x-s}{t-s} \Leftrightarrow x = s + \alpha(t-s).$$

Hence the quantile function is given by

$$F^{-1}(\alpha) = s + \alpha(t-s)$$

2a We compute the distribution function

$$\begin{aligned} P(X_{(1)} \le t) &= 1 - P(X_{(1)} > t) \\ &= 1 - P(X_1 > t, \ldots, X_n > t). \end{aligned}$$

Because of independence of $X_1, \ldots, X_n$, we have

$$P(X_{(1)} \le t) = 1 - P(X_1 > t) \cdots P(X_n > t).$$

Since each X_i has an exponential distribution

$$P(X_1 > t) = \int_t^\infty \lambda e^{-\lambda x}\, dx = e^{-\lambda t}$$

It follows that

$$P(X_{(1)} \le t) = 1 - \left(e^{-\lambda t}\right)^n = 1 - e^{-n\lambda t}$$

which is the distribution function of the exponential distribution with parameter $n\lambda$.

2b From part (a) we know that $X_{(1)}$ has an exponential distribution with parameter $n\lambda$. This means that T_1 is unbiased for $1/\lambda$:

$$\mathrm{E}T_1 = n\mathrm{E}X_{(1)} = n \times \frac{1}{n\lambda} = \frac{1}{\lambda}.$$

Therefore, the MSE of T_1 is

$$\text{MSE}(T_1) = \text{var}T_1 = n^2\text{var}X_{(1)} = n^2\frac{1}{(n\lambda)^2} = \frac{1}{\lambda^2}.$$

Write $S_n = \sum_{i=1}^n X_i$, then we may use that S_n has a gamma distribution with parameters n and λ. This means that

$$\text{E}S_n = \frac{n}{\lambda} \quad \text{and} \quad \text{var}S_n = \frac{n}{\lambda^2}.$$

This means that $T_2 = S_n/n$ is unbiased for $1/\lambda$:

$$\text{E}T_2 = \frac{1}{n}\text{E}S_n = \frac{1}{n} \times \frac{n}{\lambda} = \frac{1}{\lambda}.$$

It follows that the MSE of $T_2 = (1/n)S_n$ is

$$\text{MSE}(T_2) = \text{var}T_2 = \frac{1}{n^2}\text{var}S_n = \frac{1}{n^2} \times \frac{n}{\lambda^2} = \frac{1}{n\lambda^2}.$$

We conclude that T_2 has the smallest MSE.

3a The loglikelihood is given by

$$\log L(\theta; X_1, \ldots, X_n) = \log \prod_{i=1}^n \frac{1}{2}\theta \mathrm{e}^{-\theta|X_i|} = -n\log 2 + n\log\theta - \theta\sum_{i=1}^n |X_i|$$

Putting the derivative equal to zero gives

$$\frac{n}{\theta} - \sum_{i=1}^n |X_i| = 0 \Leftrightarrow \theta = \frac{n}{\sum_{i=1}^n |X_i|}$$

Since $\partial^2 L/\partial\theta^2 = -n/\theta^2 < 0$, this is maximum. Therefore, the maximum likelihood estimator is

$$\widehat{\theta} = \frac{n}{\sum_{i=1}^n |X_i|}.$$

3b The posterior density is proportional to

$$\begin{aligned}\prod_{i=1}^n \frac{1}{2}\theta\mathrm{e}^{-\theta|X_i|} \times \theta^{\alpha-1}e^{-\lambda\theta} &\propto \theta^n e^{-\theta\sum_{i=1}^n|X_i|} \times \theta^{\alpha-1}e^{-\lambda\theta} \\ &= \theta^{n+\alpha-1}e^{-(\lambda+\sum_{i=1}^n|X_i|)\theta}\end{aligned}$$

The right hand side is proportional to the density of a gamma distribution with parameters $\alpha' = n + \alpha$ and $\lambda' = \lambda + \sum_{i=1}^n |X_i|$. Hence the posterior distribution is a gamma distribution with these parameters. The Bayes estimator for θ is the expectation of this distribution, which is

$$\frac{\alpha'}{\lambda'} = \frac{n+\alpha}{\lambda + \sum_{i=1}^n |X_i|}.$$

4a For the observed value $\overline{x}$, the value of the test statistic is $t = \sqrt{n}(\overline{x} - \mu_0)$. Since the null hypothesis is rejected for large values of T, by definition the p-value is

$$\sup_{\mu \leq \mu_0} \mathrm{P}_\mu(T \geq t)$$

We can write this as

$$\begin{aligned}
\sup_{\mu \leq \mu_0} \mathrm{P}_\mu(T \geq t) &= \sup_{\mu \leq \mu_0} \mathrm{P}_\mu\left(\sqrt{n}(\overline{X} - \mu_0) \geq t\right) \\
&= \sup_{\mu \leq \mu_0} \mathrm{P}_\mu\left(\sqrt{n}(\overline{X} - \mu) \geq t + \sqrt{n}(\mu_0 - \mu)\right) \\
&= \sup_{\mu \leq \mu_0} \Big(1 - \Phi(t + \sqrt{n}(\mu_0 - \mu))\Big)
\end{aligned}$$

where we use that, under P_μ, the random variable $\overline{X}$ has a $N(\mu, 1/n)$ distribution, so that $\sqrt{n}(\overline{X} - \mu) \sim N(0, 1)$. Since $1 - \Phi(t + \sqrt{n}(\mu_0 - \mu))$ is increasing in μ, its supremum over $\mu \leq \mu_0$ is attained at $\mu = \mu_0$. Hence

$$\sup_{\mu \leq \mu_0} \mathrm{P}_\mu(T \geq t) = 1 - \Phi(t) = 1 - \Phi(\sqrt{n}(\overline{x} - \mu_0)).$$

4b The power function is given by

$$\begin{aligned}
\mu \mapsto \mathrm{P}_\mu(T \geq \xi_{1-\alpha_0}) &= \mathrm{P}_\mu\left(\sqrt{n}(\overline{X} - \mu_0) \geq \xi_{1-\alpha_0}\right) \\
&= \mathrm{P}_\mu\left(\sqrt{n}(\overline{X} - \mu) \geq \xi_{1-\alpha_0} + \sqrt{n}(\mu_0 - \mu)\right) \\
&= 1 - \Phi(\xi_{1-\alpha_0} + \sqrt{n}(\mu_0 - \mu))
\end{aligned}$$

where the last equality again comes from the fact that, under P_μ, the random variable $\sqrt{n}(\overline{X} - \mu)$ has a $N(0, 1)$ distribution.
We want the power function at $\mu = \mu_1$ to be at least $1 - \beta$, so

$$1 - \Phi(\xi_{1-\alpha_0} + \sqrt{n}(\mu_0 - \mu_1)) \geq 1 - \beta$$

or equivalently

$$\Phi(\xi_{1-\alpha_0} + \sqrt{n}(\mu_0 - \mu_1)) \leq \beta.$$

It follows

$$\xi_{1-\alpha_0} + \sqrt{n}(\mu_0 - \mu_1) \leq \xi_\beta.$$

This means that

$$\sqrt{n} \geq \frac{\xi_{1-\alpha_0} - \xi_\beta}{\mu_1 - \mu_0}$$

which is positive because $\beta < 1 - \alpha_0$ and $\mu_1 > \mu_0$. We find

$$n \geq \left(\frac{\xi_{1-\alpha_0} - \xi_\beta}{\mu_1 - \mu_0}\right)^2$$

5a We have that

$$Y_1 = f_1 X = \begin{pmatrix} \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} \end{pmatrix} \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i = \sqrt{n}\overline{X}$$

Furthermore

$$\sum_{i=2}^{n} Y_i^2 = \sum_{i=2}^{n} Y_i^2 - Y_1^2 = \|Y\|^2 - n\overline{X}^2 = \|OX\|^2 - n\overline{X}^2 = \|X\|^2 - n\overline{X}^2$$

because O is orthonormal. The right hand side is equal to

$$\sum_{i=1}^{n} X_i^2 - n\overline{X}^2 = \sum_{i=1} (X_i - \overline{X})^2.$$

The latter can be seen as follows

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

5b We may use that $Y = OX$ has independent standard normal distributed components $Y_1, \ldots, Y_n$. In particular Y_1 is independent of $Y_2, \ldots, Y_n$, and hence from part (a) we conclude that $\overline{X} = Y_1/\sqrt{n}$ and $\sum_{i=1}^{n}(X_i - \overline{X})^2 = \sum_{i=2}^{n} Y_i^2$ are independent.

Finally,

$$\sum_{i=1}^{n} (X_i - \overline{X})^2 = \sum_{i=2}^{n} Y_i^2$$

which is the sum of $n-1$ independent squared standard normal random variables. By definition this has a chi-square distribution with $n-1$ degrees of freedom.