TEL 502E – Detection and Estimation Theory

Spring 2015

Instructors : İlker Bayram
ibayram@itu.edu.tr

Class Meets : Tuesday, 9.30 – 12.30, EEB 1301

Textbook : 'Fundamentals of Statistical Signal Processing' (Vols. I,II), S. M. Kay, Prentice Hall.

Supplementary :
- ◦ 'An Introduction to Signal Detection and Estimation', H. V. Poor, Springer.
- ◦ For a Review of Probability : 'Introduction to Probability', D. P. Bersekas, J. N. Tsitsiklis, Athena Scientific.

Webpage : There's a 'ninova' page, please log in and check.

Grading : Homeworks (10%), Midterm exam (40%), Final Exam (50%).

Tentative Course Outline

(1) Review of probability theory

(2) The estimation problem, minimum variance unbiased estimators

(3) The Cramér-Rao bound, sufficient statistics, Rao-Blackwell Theorem

(4) Best linear unbiased estimators maximum likelihood estimation

(5) Bayesian estimation, minimum mean square estimators, maximum a posteriori estimators

(6) The innovations process, Wiener filtering, recursive least squares, the Kalman filter

(7) Interval Estimation

(8) Simple Hypothesis Testing, the Neyman Pearson Lemma

(9) Bayesian tests, multiple hypothesis testing

(10) The matched filter, detection of stochastic signals

TEL502E – Homework 1

Due 17.02.2015

1. Let X_1, X_2 be iid (independent, identically distributed) zero-mean Gaussian random variables with variance σ^2. Let $Y = (X_1 + X_2)/2$. Also, let the error function be defined as

$$\Phi(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}\, dx.$$

(a) Find the pdf of Y.

(b) Compute the probability of the event $\{Y \leq t\}$, in terms of σ^2 and Φ.

(c) Compute the probability of the event $\{Y = 0\}$, in terms of σ^2 and Φ.

Solution. (a) Recall that the sum of independent random variables distributed as $\mathcal{N}(\theta_1, \sigma_1^2)$, $\mathcal{N}(\theta_2, \sigma_2^2)$ is distributed as $\mathcal{N}(\theta_1 + \theta_2, \sigma_1^2 + \sigma_2^2)$. Thus Y is a $\mathcal{N}(0, \sigma^2/2)$ random variable, with pdf

$$f_Y(t) = \frac{1}{\sqrt{\pi\sigma^2}} e^{-t^2/\sigma^2}.$$

(b) We compute

$$\begin{aligned} P(Y \leq t) &= \int_{-\infty}^{t} f_Y(s)\, ds \\ &= \int_{-\infty}^{t} \frac{1}{\sqrt{\pi\,\sigma^2}} e^{-s^2/\sigma^2}\, ds. \end{aligned}$$

Making the change of variables $s/\sigma = x/\sqrt{2}$, we obtain,

$$P(Y \leq t) = \int_{-\infty}^{\sqrt{2}\,t/\sigma} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx = \Phi\big(\sqrt{2}t/\sigma\big).$$

(c) We have,

$$P(Y = 0) = \int_{0}^{0} \frac{1}{\sqrt{\pi}} e^{-s^2}\, ds = 0.$$

2. Recall that the Fisher information $I(\theta)$ for a given density $f(t, \theta)$ dependent on θ is given by

$$I(\theta) = \mathbb{E}\left(\left[\frac{\partial}{\partial\theta} \ln\Big(f(X, \theta)\Big)\right]^2\right).$$

Note that in general, $I(\theta)$ is a function of θ. In this question, we investigate the behavior of I when f takes special forms.

(a) Suppose f is of the form,

$$f(t, \theta) = g(t - \theta).$$

Show that $I(\theta)$ is a constant (i.e., does not vary with θ).

(b) Suppose f is of the form,

$$f(t, \theta) = \frac{1}{\theta}\, g\left(\frac{t}{\theta}\right),$$

where $\theta > 0$. Let us define the constant $c = I(1)$. Find a simple expression of $I(\theta)$ in terms of c and θ.

Solution. (a) Note that we get a different pdf for each value of θ, but the pdfs are related to each other in a very specific way. Such a collection of pdfs is called a location family. Note that

$$\frac{\partial}{\partial\theta} \ln\Big(f(t, \theta)\Big) = \frac{\partial_\theta f(t, \theta)}{f(t, \theta)} = \frac{-g'(t - \theta)}{g(t - \theta)}$$

Therefore,

$$
\begin{aligned}
I(\theta) &= \mathbb{E}\left(\left[\frac{\partial}{\partial\theta}\ln\Big(f(X,\theta)\Big)\right]^2\right) \\
&= \int_{-\infty}^{\infty}\left(\frac{-g'(t-\theta)}{g(t-\theta)}\right)^2 g(t-\theta)\,dt \\
&= \int_{-\infty}^{\infty}\left(\frac{-g'(s)}{g(s)}\right)^2 g(s)\,ds,
\end{aligned}
$$

where we made the change of variables $s = t - \theta$ to obtain the last line. Since this integral is independent of θ, the claim follows.

(b) As in (a), we get a collection of related pdfs but they are related to each other in a different manner. This collection is called a scale family. Note that (check this!)

$$
\frac{\partial}{\partial\theta}\ln\Big(f(t,\theta)\Big) = \frac{\partial_\theta f(t,\theta)}{f(t,\theta)} = -\frac{1}{\theta} - \frac{t}{\theta^2}\frac{g'(t/\theta)}{g(t/\theta)}.
$$

Therefore,

$$
\begin{aligned}
h(t,\theta) &= \left[\frac{\partial}{\partial\theta}\ln\Big(f(t,\theta)\Big)\right]^2 \\
&= \underbrace{\frac{1}{\theta^2}}_{h_1(t,\theta)} + \underbrace{\frac{2t}{\theta^3}\frac{g'(t/\theta)}{g(t/\theta)}}_{h_2(t,\theta)} + \underbrace{\frac{t^2}{\theta^4}\left[\frac{g'(t/\theta)}{g(t/\theta)}\right]^2}_{h_3(t,\theta)}.
\end{aligned}
$$

Thus we have,

$$
I(\theta) = \mathbb{E}(h(X,\theta)) = \mathbb{E}(h_1(X,\theta)) + \mathbb{E}(h_2(X,\theta)) + \mathbb{E}(h_3(X,\theta)).
$$

Now,

$$
\mathbb{E}(h_1(X,\theta)) = \frac{1}{\theta^2}.
$$

$$
\begin{aligned}
\mathbb{E}(h_2(X,\theta)) &= \int \frac{2t}{\theta^3}\frac{g'(t/\theta)}{g(t/\theta)}\frac{1}{\theta}\,g(t/\theta)\,dt \\
&= \frac{2}{\theta^3}\int t\,\frac{1}{\theta}\,g'(t/\theta)\,dt \\
&= \frac{2}{\theta^3}\left(-\int g(t/\theta)\,dt\right) \\
&= \frac{2}{\theta^3}(-\theta) \\
&= -\frac{2}{\theta^2},
\end{aligned}
$$

where we integrated by parts (assuming that $t\,g(t)$ is absolutely integrable) in the third step.
Finally,

$$
\begin{aligned}
\mathbb{E}(h_3(X,\theta)) &= \int \frac{t^2}{\theta^4}\left[\frac{g'(t/\theta)}{g(t/\theta)}\right]^2\frac{1}{\theta}\,g(t/\theta)\,dt \\
&= \frac{1}{\theta^2}\int s^2\left[\frac{g'(s)}{g(s)}\right]^2 g(s)\,ds \\
&= \frac{1}{\theta^2}\,\mathbb{E}(h_3(X,1)),
\end{aligned}
$$

where we made the change of variables $s = t/\theta$. Thus, we have,

$$
I(\theta) = \mathbb{E}(h(X,\theta)) = \frac{1}{\theta^2}\big[\mathbb{E}(h_3(X,1)) - 1\big].
$$

Plugging in $\theta = 1$, we find that $I(1) = \mathbb{E}(h(X,\theta)) = \mathbb{E}(h_3(X,1)) - 1$. Thus,

$$
I(\theta) = \frac{1}{\theta^2}\,I(1).
$$

3. (a) Compute the Fisher information $I(\theta)$ for the distribution $\mathcal{N}(0,\theta)$.
   (b) Compute the Fisher information $I(\theta)$ for the distribution $\mathcal{N}(\theta,1)$.

   Solution. (a) Note that the family $\mathcal{N}(0,\theta)$ is a scale family, because the pdf associated with a $\mathcal{N}(0,\theta)$ random variable is given by

$$f(t,\theta) = \frac{1}{\sqrt{2\pi\,\theta^2}} \exp\left(-\frac{t^2}{2\theta^2}\right) = \frac{1}{\theta}\, g(t/\theta),$$

for

$$g(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right).$$

Thus, $I(\theta) = I(1)/\theta^2$, by Q2b. Let us now compute $I(1)$.

$$\begin{aligned} I(1) &= \int_{-\infty}^{\infty} \left(\frac{\partial_\theta f(t,\theta)}{f(t,\theta)}\right)^2 f(t,\theta)\, dt \Big|_{\theta=1} \\ &= \int_{-\infty}^{\infty} \left(\frac{g'(t)}{g(t)}\right)^2 g(t)\, dt \\ &= \int_{-\infty}^{\infty} t^2\, g(t)\, dt \\ &= 1. \end{aligned}$$

Thus $I(\theta) = 1/\theta^2$.

   (b) Note now that $\mathcal{N}(\theta,1)$ is a location family, because the pdf associated with a $\mathcal{N}(\theta,0)$ random variable is given by

$$f(t,\theta) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(t-\theta)^2}{2}\right) = g(t-\theta),$$

for

$$g(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right).$$

Therefore, by Q2a, $I(\theta)$ is a constant. To find that constant, we can compute $I(\theta)$ for a special choice of θ. Let us take $\theta = 0$. We have,

$$\begin{aligned} I(0) &= \int_{-\infty}^{\infty} \left(\frac{\partial_\theta f(t,\theta)}{f(t,\theta)}\right)^2 f(t,\theta)\, dt \Big|_{\theta=0} \\ &= \int_{-\infty}^{\infty} \left(\frac{g'(t)}{g(t)}\right)^2 g(t)\, dt \\ &= \int_{-\infty}^{\infty} t^2\, g(t)\, dt \\ &= 1. \end{aligned}$$

4. Suppose X_1 and X_2 are independent random variables distributed as $\mathcal{N}(\theta,1)$ and $\mathcal{N}(2\theta,1)$ respectively, where θ is an unknown parameter. Find the UMVUE for θ in terms of X_1 and X_2.

   Solution. Notice that the joint pdf of X_1 and X_2 is given by

$$f(t_1,t_2,\theta) = \frac{1}{2\pi} \exp\left(-\frac{(t_1-\theta)^2 + (t_2-2\theta)^2}{2}\right).$$

Recall that if CRLB is achieved by an estimator $\hat{\theta}$, then

$$\frac{\partial}{\partial\theta} \ln\Big(f(t_1,t_2,\theta)\Big) = I(\theta)\,(\hat{\theta}(t_1,t_2) - \theta).$$

Let us now compute $I(\theta)$. First,

$$\begin{aligned} \frac{\partial}{\partial\theta} \ln\Big(f(t_1,t_2,\theta)\Big) &= (t_1-\theta) + 2(t_2-2\theta) \\ &= 5\left(\frac{1}{5}(t_1+2t_2) - \theta\right) \end{aligned}$$

Now,

$$\begin{aligned}
I(\theta) &= \mathbb{E}\left(\left[\frac{\partial}{\partial\theta}\ln\Big(f(X_1, X_2, \theta)\Big)\right]^2\right) \\
&= \mathbb{E}\left(\Big[(X_1-\theta)+2(X_2-2\theta)\Big]^2\right) \\
&= \mathbb{E}\Big((X_1-\theta)^2\Big) + \mathbb{E}\Big(4(X_2-2\theta)^2\Big) + \mathbb{E}\Big(2(X_1-\theta)(X_2-2\theta)\Big) \\
&= 5.
\end{aligned}$$

Thus $I(\theta)$ is a constant (in fact this is expected because X_1, X_2 come from a location family). Thus,

$$\frac{\partial}{\partial\theta}\ln\Big(f(t_1, t_2, \theta)\Big) = I(\theta)\ \Big(\hat{\theta}(t_1, t_2) - \theta\Big),$$

for

$$\hat{\theta}(t_1, t_2) = \frac{1}{5}(t_1 + 2t_2).$$

Observe now that

$$\mathbb{E}\left(\hat{\theta}(X_1, X_2)\right) = \frac{1}{5}(\theta + 2\theta) = \theta.$$

Thus $\hat{\theta}(X_1, X_2)$ is an unbiased estimator of θ. Therefore by the CRLB theorem discussed in class, it must be the UMVUE (one also needs to check the two regularity conditions – I leave that to you).

5. Suppose X_1 and X_2 are independent and unit variance random variables with $\mathbb{E}(X_i) = \theta$, where θ is an unknown constant.

   (a) Show that $\hat{\theta} = (X_1 + X_2)/2$ is an unbiased estimator for θ. What is the variance of $\hat{\theta}$?

   (b) Suppose we are interested in the value $\gamma = \theta^2$. Consider $\hat{\gamma} = \hat{\theta}^2$ as an estimator for γ. Is $\hat{\gamma}$ an unbiased estimator of γ?

   Solution. (a) We compute

$$\mathbb{E}\left(\frac{X_1 + X_2}{2}\right) = \frac{1}{2}\Big(\mathbb{E}(X_1) + \mathbb{E}(X_2)\Big) = \theta.$$

   Thus $\hat{\theta}$ is unbiased. Also, since we can add the variances of independent random variables, we have

$$\mathrm{var}(\hat{\theta}) = \mathrm{var}\left(\frac{X_1}{2}\right) + \mathrm{var}\left(\frac{X_2}{2}\right) = \frac{1}{4}\ \mathrm{var}(X_1) + \frac{1}{4}\ \mathrm{var}(X_2) = \frac{1}{2}.$$

   (b) We compute

$$\begin{aligned}
\mathbb{E}\left(\hat{\gamma}\right) &= \frac{1}{4}\Big(\mathbb{E}(X_1^2 + X_2^2 + 2X_1\,X_2)\Big) \\
&= \frac{4\theta^2 + 2}{4} \\
&> \theta.
\end{aligned}$$

   Thus $\hat{\gamma}$ is not an unbiased estimator of $\gamma = \theta^2$.

TEL502E – Homework 2

Due 24.02.2015

1. (a) Suppose Z is a $\mathcal{N}(1,1)$ random variable and we observe $X = \theta\, Z$, where θ is an unknown constant. Find the pdf of X.

   (b) Find the Fisher information $I(\theta)$ for the distribution of X in (a).

   Solution. (a) We know that linear combinations of Gaussian random variables are also Gaussian (why?). Therefore, X is also Gaussian. We compute $\mathbb{E}(X) = \theta\,\mathbb{E}(Z) = \theta$ and $\mathrm{var}(X) = \theta^2\ \mathrm{var}(Z) = \theta^2$. Thus, X is a $\mathcal{N}(\theta, \theta^2)$ random variable.

   (b) The pdf of X is,

$$f(t,\theta) = \frac{1}{\sqrt{2\pi}\,\theta}\ \exp\left(-\frac{(t-\theta)^2}{2\theta^2}\right).$$

   We compute (check this!),

$$\frac{\partial}{\partial\theta}\ \ln\Big(f(t,\theta)\Big) = \frac{1}{\theta^3}\,\big(t^2 - \theta\, t - \theta^2\big).$$

   Thus, we have (please check a table of non-central moments of a Gaussian random variable – it's better to fill such a table from scratch)

$$\begin{aligned}
\mathbb{E}\left(\left[\frac{\partial}{\partial\theta}\ \ln\Big(f(t,\theta)\Big)\right]^2\right) &= \frac{1}{\theta^6}\,\mathbb{E}\big(X^4 + \theta^2\,X^2 + \theta^4 - 2\theta\,X^3 - 2\theta^2\,X^2 + \theta^3\,X\big)\\
&= \frac{1}{\theta^6}\,\big(10\,\theta^4 + 2\,\theta^4 + \theta^4 - 8\,\theta^4 - 4\theta^4 + \theta^4\big)\\
&= \frac{2}{\theta^2}
\end{aligned}$$

   Note that Fisher information increases as $\theta \to 0$ in this scenario.

2. (a) Suppose $X = H\,\theta + W$, where W is a $\mathcal{N}(0, C)$ random vector (here C is the covariance of W), θ is an unknown vector and H is a matrix. Find the pdf of X.

   (b) Find the Fisher information matrix $I(\theta)$ for the pdf of X.

   Solution. (a) Note that X is a linear combination of a Gaussian random vector. Therefore it is also Gaussian and it's sufficient to determine its mean and covariance matrix. Note that $\mathbb{E}(X) = H\,\theta$ and $\mathrm{cov}(X) = \mathrm{cov}(W) = C$. Thus, supposing X is of length n, the pdf of X is of the form

$$f(t,\theta) = \frac{1}{\sqrt{(2\pi)^n |C|}}\ \exp\left(-\frac{1}{2}\,(t - H\,\theta)^T\,C^{-1}(t - H\,\theta)\right).$$

   where $t = \begin{bmatrix} t_1 & t_2 & \dots & t_n\end{bmatrix}^T$.

   (b) Note that

$$v(t) := \nabla_\theta\ \ln\Big(f(t,\theta)\Big) = -H^T\,C^{-1}(H\theta - t)$$

   Thus,

$$I(\theta) = \mathbb{E}\Big(v(X)\,v(X)^T\Big) = \mathbb{E}\Big(H^T\,C^{-1}\,(H\,\theta - X)(H\,\theta - X)^T\,C^{-1}\,H\Big) = H^T\,C^{-1}\,H.$$

   Note that $I(\theta)$ is constant with respect to θ.

3. Suppose $Y = g(X)$ for an invertible function g and the pdfs of X and Y depend on an unknown parameter θ. Suppose also that the estimators of θ based on X and Y, namely $\hat\theta(X)$ and $\tilde\theta(Y)$ are efficient. Show that $\hat\theta(X) = \tilde\theta\big(g(X)\big)$.
   Hint : Recall that an efficient estimator $\hat\theta$ satisfies the equality

$$\frac{\partial}{\partial\theta}\ln f(t,\theta) = I(\theta)\,(\hat\theta - \theta).$$

Solution. Note that g is either increasing or decreasing in order to be invertible. Let us assume g is increasing (a similar analysis can also be carried out for a decreasing g). Note that if $f_X(t,\theta)$ and $f_Y(t,\theta)$ denote the pdfs of X and Y, they satisfy (show this!),

$$f_X(t,\theta) = f_Y(g(t),\theta)\, g'(t).$$

Thus, we have,

$$\frac{\partial_\theta\, f_X(t,\theta)}{f_X(t,\theta)} = \frac{\partial_\theta f_Y(g(t),\theta)}{f_Y(g(t),\theta)}. \tag{1}$$

But by the efficiency of the estimators, we also have (note : $I(\theta)$ is the same for X and Y and this can be shown using (1)),

$$\frac{\partial_\theta\, f_X(t,\theta)}{f_X(t,\theta)} = I(\theta)\,(\hat{\theta}(t) - \theta) \tag{2a}$$

$$\frac{\partial_\theta\, f_Y(t,\theta)}{f_Y(t,\theta)} = I(\theta)\,(\tilde{\theta}(t) - \theta). \tag{2b}$$

Replacing t with $g(t)$ in (2b), we obtain by (1) that

$$I(\theta)\,(\hat{\theta}(t) - \theta) = I(\theta)\,(\tilde{\theta}\big(g(t)\big) - \theta).$$

Cancelling terms, we obtain $\hat{\theta}(t) = \tilde{\theta}\big(g(t)\big)$.

4. Recall that for square integrable functions $g(t)$, $h(t)$, the Cauchy-Schwarz inequality (CSI) is

$$\left(\int g(t)\,h(t)\,dt\right)^2 \leq \left(\int g^2(t)dt\right)\left(\int h^2(t)dt\right).$$

(a) Let X be a random variable with pdf $f(t)$. Use CSI to show that

$$\big[\mathbb{E}\big(g(X)\,h(X)\big)\big]^2 \leq \mathbb{E}\big(g^2(X)\big)\,\mathbb{E}\big(h^2(X)\big).$$

(b) Suppose now that X is a random vector. Also, let $g(X)$, $h(X)$ be random vectors of the form

$$g(X) = \begin{bmatrix} g_1(X) \\ g_2(X) \\ \vdots \\ g_n(X) \end{bmatrix}, \quad h(X) = \begin{bmatrix} h_1(X) \\ h_2(X) \\ \vdots \\ h_n(X) \end{bmatrix},$$

and $\mathbb{E}\big(g(X)\,h^T(X)\big) = I$, where I denotes the $n \times n$ identity matrix. Use part (a) to show that for arbitrary length-n column vectors c, d, we have

$$(c^T\, d)^2 \leq (c^T\, G\, c)\,(d^T\, H\, d),$$

where

$$G = \mathbb{E}\big(g(X)\,g^T(X)\big), \quad H = \mathbb{E}\big(h(X)\,h^T(X)\big).$$

(c) Show that if G and H are symmetric matrices and

$$(c^T\, d)^2 \leq (c^T\, G\, c)\,(d^T\, H\, d),$$

for arbitrary column vectors c, d, then $G - H^{-1}$ is positive semi-definite. (Note that, taken together with part (b), this fills the gap in the proof of the vector valued CRLB discussed in class.)

Solution. (a) Let $f_X(t)$ denote the pdf of X. We have, by CSI,

$$\begin{aligned}\big[\mathbb{E}\big(g(X)\,h(X)\big)\big]^2 &= \left[\int g(t)\,h(t)\,f_X(t)\,dt\right]^2 \\ &\leq \left[\int g^2(t)\,f_X(t)\,dt\right] \cdot \left[\int h^2(t)\,f_X(t)\,dt\right] \\ &= \mathbb{E}\big(g^2(X)\big)\cdot \mathbb{E}\big(h^2(X)\big),\end{aligned}$$

where we used the observation $\left[g(t)\,\sqrt{f_X(t)}\right] \cdot \left[h(t)\,\sqrt{f_X(t)}\right] = g(t)\,h(t)\,f_X(t)$, which is valid since $f_X(t)$ is non-negative.

(b) Note that $c^t\, g(X)$ and $h^T(X)\, d$ can be thought of as scalars. Thus, by part (a),

$$
\begin{aligned}
\mathbb{E}\big(c^T g(X)\, h^T(X)\, d\big)^2 &\le \mathbb{E}\big(\overbrace{c^T g(X)\, g^T(X)\, c}^{(c^T g(X))^2}\big)\, \mathbb{E}\big(d^T h(X)\, h^T(X)\, d\big) \\
&= \Big(c^T\, \mathbb{E}\big[g(X)\, g^T(X)\big]\, c\Big)\, \Big(d^T\, \mathbb{E}\big[h(X)\, h^T(X)\big]\, d\Big) \\
&= \big(c^T G\, c\big)\, \big(d^T H\, d\big).
\end{aligned}
$$

But we also have

$$\mathbb{E}\big(c^T g(X)\, h^T(X)\, d\big) = c^T\, \mathbb{E}\big[g(X)\, g^T(X)\big]\, d = c^T d.$$

Thus follows the inequality.

(c) Take $d = H^{-1} c$. Then we have that

$$(c^T G\, c)\, (c^T H^{-1} c) \ge (c^T H^{-1} c)^2$$

for any c. Cancelling terms, we have

$$(c^T G\, c) \ge (c^T H^{-1} c).$$

Rearranging we obtain

$$c^T (G - H^{-1})\, c \ge 0.$$

for any c. Therefore $G - H^{-1}$ is positive semi definite.

TEL502E – Homework 3

Due 03.03.2015

1. Suppose X and Y random variables.

   (a) Show that

   $$\big[\mathbb{E}\big(X|Y=y\big)\big]^2 \leq \mathbb{E}\big(X^2|Y=y\big),$$

   for any value of y.
   Hint : Note that

   $$\mathbb{E}\big(X|Y=y\big) = \int x\, f_{X|Y}(x|y)\, dx.$$

   Use the Cauchy-Schwarz inequality.

   (b) Show that

   $$\mathbb{E}\Big(\big[\mathbb{E}\big(X|Y\big)\big]^2\Big) \leq \mathbb{E}\Big(\mathbb{E}\big(X^2|Y\big)\Big).$$

   (c) Show that conditioning reduces variance, that is, $\mathrm{var}\big(\mathbb{E}(g(X)|Y)\big) \leq \mathrm{var}(g(X))$ for any function $g(\cdot)$.

Solution. (a) Observe that $\big(x\, f_{X|Y}(x|y)\big) = \big(x\,\sqrt{f_{X|Y}(x|y)}\big) \cdot \sqrt{f_{X|Y}(x|y)}$. Thus, we obtain by CSI that

$$\begin{aligned}
\big[\mathbb{E}\big(X|Y=y\big)\big]^2 &= \left[\int x\, f_{X|Y}(x|y)\, dx\right]^2 \\
&\leq \int x^2\, f_{X|Y}(x|y)\, dx \int f_{X|Y}(x|y)\, dx \\
&=\leq \int x^2\, f_{X|Y}(x|y)\, dx \\
&= \mathbb{E}\big(X^2|Y=y\big).
\end{aligned}$$

(b) Let $h_1(Y) = \mathbb{E}\big(X|Y\big)$ and $h_2(Y) = \mathbb{E}\big(X^2|Y\big)$. Note that by part (a), we know that $h_1^2(t) \leq h_2(t)$ for any t. Also, let $f_Y(t)$ denote the pdf of Y. We have,

$$\begin{aligned}
\mathbb{E}\Big(\big[\mathbb{E}\big(X|Y\big)\big]^2\Big) &= \mathbb{E}\big(h_1^2(Y)\big) \\
&= \int h_1^2(t)\, f_Y(t)\, dt \\
&\leq \int h_2(t)\, f_Y(t)\, dt \\
&= \mathbb{E}\big(h_2(Y)\big) \\
&= \mathbb{E}\Big(\mathbb{E}\big(X^2|Y\big)\Big).
\end{aligned}$$

(c) We have seen in class that $\mathbb{E}\Big(\mathbb{E}\big(X^2|Y\big)\Big) = \mathbb{E}(X)$. Therefore, the inequality in (b) may also be written as,

$$\mathbb{E}\Big(\big[\mathbb{E}\big(X|Y\big)\big]^2\Big) \leq \mathbb{E}\big(X^2\big).$$

Now let us apply this observation. Let $\mu = \mathbb{E}(g(X))$. Observe also that $\mathbb{E}\big(\mathbb{E}(g(X)|Y)\big) = \mu$. Now,

$$\begin{aligned}
\mathrm{var}\Big(\mathbb{E}\big(g(X)|Y\big)\Big) &= \mathbb{E}\Big(\Big[\mathbb{E}\big(g(X)|Y\big) - \mu\Big]^2\Big) \\
&= \mathbb{E}\Big(\Big[\mathbb{E}\big(g(X) - \mu|Y\big)\Big]^2\Big) \\
&\leq \mathbb{E}\Big(\big[g(X) - \mu\big]^2\Big) \\
&= \mathrm{var}\big(g(X)\big).
\end{aligned}$$

2. Suppose X_1 and X_2 are independent random variables distributed as $\mathcal{N}(\theta, 1)$ and $\mathcal{N}(2\theta, 1)$ respectively, where θ is an unknown parameter.

   (a) Find a complete sufficient statistic $T(X_1, X_2)$ for θ.

   (b) Find an unbiased estimator of θ which is a function of T. That is, find $g(T)$ such that $\mathbb{E}\big(g(T)\big) = \theta$.

   Solution. (a) The joint pdf of X_1 and X_2 is,

$$
\begin{aligned}
f_X(x_1, x_2) &= \frac{1}{2\pi} \exp\left(-\frac{(x_1-\theta)^2 + (x_2 - 2\theta)^2}{2}\right) \\
&= \underbrace{\left[\frac{1}{2\pi} \exp\left(-\frac{x_1^2 + x_2^2}{2}\right)\right]}_{h(x_1,x_2)} \cdot \underbrace{\left[\exp\left(\theta\,(x_1 + 2x_2) - \frac{5}{2}\theta^2\right)\right]}_{q(t,\theta)},
\end{aligned}
$$

   where h is independent of θ and q is a function of t and θ only for $t = x_1 + 2x_2$. Thus $T = X_1 + 2\,X_2$ is a sufficient statistic for this problem.

   To see that T is complete, note that since T is a linear combination of Gaussian random variables, it is also Gaussian. In fact it is distributed as $\mathcal{N}(5\theta, 5)$. Suppose now that for a function $g(T)$, we have $\mathbb{E}(g) = 0$ for all θ. Then,

$$
\mathbb{E}\big(g(T)\big) = \frac{1}{\sqrt{10\pi}} \int g(t) \exp\left(-\frac{(t-5\theta)^2}{2}\right) dt = 0, \qquad \text{for all } \theta. \tag{1}
$$

   But for $s = 5\theta$, we can rewrite this condition as

$$
\int g(t) \exp\left(-\frac{(s-t)^2}{2}\right) dt = g(s) * w(s) = 0 \qquad \text{for all } s, \tag{2}
$$

   where $w(s) = \exp\left(-s^2/2\right)$. But convolution with a Gaussian function gives zero if and only if the input function, namely $g(\cdot)$ is zero. Thus T is complete.

   (b) We note that $\mathbb{E}(T) = 5\theta$. Therefore, $g(T) = T/5$ is the UMVUE by the Rao-Blackwell theorem.

3. Suppose X_1 and X_2 are independent random variables distributed as $\mathcal{N}(\theta, 1)$ and $\mathcal{N}(2\theta, 2)$ respectively, where θ is an unknown parameter.

   (a) Find a complete sufficient statistic $T(X_1, X_2)$ for θ.

   (b) Find an unbiased estimator of θ which is a function of T. That is, find $g(T)$ such that $\mathbb{E}\big(g(T)\big) = \theta$.

   Solution. (a) The joint pdf of X_1 and X_2 is,

$$
\begin{aligned}
f_X(x_1, x_2) &= \frac{1}{2\sqrt{2}\pi} \exp\left(-\frac{(x_1-\theta)^2}{2} - \frac{(x_2 - 2\theta)^2}{4}\right) \\
&= \underbrace{\left[\frac{1}{2\sqrt{2}\pi} \exp\left(-\frac{2x_1^2 + x_2^2}{4}\right)\right]}_{h(x_1,x_2)} \cdot \underbrace{\left[\exp\left(\theta\,(x_1 + x_2) - \frac{3}{2}\theta^2\right)\right]}_{q(t,\theta)},
\end{aligned}
$$

   where h is independent of θ and q is a function of t and θ only for $t = x_1 + x_2$. Thus $T = X_1 + X_2$ is a sufficient statistic for this problem.

   As in Q2, T is Gaussian, but this time, it is distributed as $\mathcal{N}(3\theta, 3)$. T can be shown to be complete as in Q2.

   (b) Note that $\mathbb{E}(T) = 3\theta$. Thus $g(T) = T/3$ is the UMVUE by the Rao-Blackwell theorem.

4. (From our supplementary book) For $\theta > 0$, let $A_\theta \subset \mathbb{R}^2$ be the region defined by

$$
A_\theta = \{(x, y) : 0 \le x,\ 0 \le y,\ x + y \le \theta\}.
$$

   Also, suppose $(X_1, Y_1), \ldots, (X_n, Y_n)$ denote iid random variables from the uniform distribution on A_θ, so that their common density is given by

$$
f_{X,Y}(x, y) = \begin{cases} 2/\theta^2, & \text{if } (x, y) \in A_\theta, \\ 0, & \text{if } (x, y) \notin A_\theta. \end{cases}
$$

(a) Find a complete sufficient statistic T for θ.

(b) Find an unbiased estimator of θ which is a function of T. That is, find $g(T)$ such that $\mathbb{E}\big(g(T)\big) = \theta$.

Solution. (a) Let us rewrite the pdf in terms of the step function $u(t)$. Recall that

$$u(t) = \begin{cases} 0, & \text{if } t < 0, \\ 1, & \text{if } 0 \le t. \end{cases}$$

Therefore, we can write

$$f_{X,Y}(x,y) = \frac{2}{\theta^2}\, u(x)\, u(y)\, u\big(\theta - (x+y)\big).$$

Thanks to independence, we can write the joint pdf of (X_1, Y_1), ..., (X_n, Y_n) as

$$\begin{aligned} f(x_1, y_1, \ldots, x_n, y_n) &= \prod_{k=1}^{n} f_{X,Y}(x_k, y_k) \\ &= \left[\prod_{k=1}^{n} u(x_k)\, u(y_k)\right] \left[\left(\frac{2}{\theta}\right)^n \prod_{k=1}^{n} u\big(\theta - (x_k + y_k)\big)\right]. \end{aligned}$$

But now observe that, for an arbitrary collection of numbers s_1, ..., s_n, we have

$$\prod_{k=1}^{n} u\big(\theta - s_k\big) = u\big(\theta - \max_k(s_k)\big),$$

where '$\max_k(s_k)$' denotes the maximum of $s_1, \ldots s_n$. Thus, if we let $t = \max_k\big(x_k + y_k\big)$, then we can write the joint pdf as

$$f(x_1, y_1, \ldots, x_n, y_n) = \left[\prod_{k=1}^{n} u(x_k)\, u(y_k)\right] \left[\left(\frac{2}{\theta}\right)^n u\big(\theta - t\big)\right].$$

From the factorization theorem, we can therefore conclude that

$$T = \max_{1 \le k \le n} (X_k + Y_k)$$

is a sufficient statistic.

Let us now show that T is complete. We will need the pdf of T. We will obtain that in two steps. Note that T can be written as $T = \max_k(S_k)$ where $S_k = X_k + Y_k$. Let us first find the cdf of S_k. Notice that

$$\begin{aligned} F_{S_k}(t) &= P(S_k \le t) \\ &= P(X_k + Y_k \le t) \\ &= \begin{cases} 0, & \text{if } t < 0, \\ t^2/\theta^2, & \text{if } 0 \le t \le \theta, \\ 1, & \text{if } \theta \le t. \end{cases} \end{aligned}$$

Now, since S_k are independent random variables, we have (recall the argument in class)

$$\begin{aligned} F_T(t) &= P(T \le t) \\ &= P\Big((S_1 \le t) \cap (S_2 \le t) \cap \cdots \cap (S_n \le t)\Big) \\ &= P(S_1 \le t) \cdot P(S_2 \le t) \cdot P(S_n \le t) \\ &= F_{S_1}^n(t). \end{aligned}$$

Differentiating, we obtain the pdf of T as

$$\begin{aligned} f_T(t) &= n\, F_{S_1}^{n-1}(t)\, F_{S_1}'(t) \\ &= \begin{cases} 0, & \text{if } t < 0, \\ 2n\, t^{2n-1}/\theta^{2n}, & \text{if } 0 \le t \le \theta, \\ 0, & \text{if } \theta \le t. \end{cases} \end{aligned}$$

Now assume that $\mathbb{E}\big(g(T)\big) = 0$, for all θ. This means that

$$\int_0^\theta g(t)\,\frac{2n}{\theta^{2n}}\,t^{2n-1}\,dt = 0 \text{ for all } \theta.$$

This implies that

$$h(\theta) = \int_0^\theta g(t)\,t^{2n-1}\,dt = 0 \text{ for all } \theta.$$

If we differentiate this with respect to θ, we get

$$h'(\theta) = g(\theta)\,\theta^{2n-1} = 0 \text{ for all } \theta.$$

But this means that $g = 0$. Thus T is complete.

(b) Since we know the pdf of T, let us compute $\mathbb{E}(T)$.

$$\mathbb{E}(T) = \int_0^\theta \frac{2n}{\theta^{2n}}\,t^{2n}\,dt = \frac{2n}{2n+1}\,\theta.$$

Thus, $g(T) = T\,(2n+1)/2n$ is the UMVUE by the Rao-Blackwell theorem.

TEL502E – Homework 4

Due 10.03.2015

1. (From our textbook) Consider the frequency estimation of a sinusoid embedded in white Gaussian noise or,

   $$x(n) = \cos(\omega n) + u(n), \qquad \text{for } n = 0, 1, \dots, N-1,$$

   where $u(n)$ is white Gaussian noise with unit variance. Show that it is not possible to find a sufficient statistic for ω.

2. (From supplementary book) Consider the exponential distribution with failure rate λ, that is,

   $$f(x) = \begin{cases} 0, & \text{if } x < 0, \\ \frac{1}{\lambda}\, e^{-x/\lambda}, & \text{if } 0 \leq x. \end{cases}$$

   Find an invertible function h defining a new parameter $\theta = h(\lambda)$ so that Fisher information $I(\theta)$ is constant.

3. Suppose $Y = X + Z$, where X and Z are independent $\mathcal{N}(0,1)$.

   (a) Compute $\mathbb{E}(Y|X)$.

   (b) Compute $\mathbb{E}(X|Y)$.

4. Suppose X_1, X_2 are iid and distributed as $\mathcal{N}(\theta, 1)$. Also, let $T = X_1 + X_2$. Find an expresssion for $g(T) = \mathbb{E}(X_1|T)$.

5. Suppose $X_1, \dots X_n$ are iid samples from an exponential distribution with failure rate λ, that is,

   $$f_{X_i}(x) = \begin{cases} 0, & \text{if } x < 0, \\ \frac{1}{\lambda}\, e^{-x/\lambda}, & \text{if } 0 \leq x. \end{cases}$$

   Find a complete sufficient statistic for λ.

TEL502E – Homework 5

Due 17.03.2015

1. Suppose X_1, X_2, $\ldots$, X_n are iid random variables, with pdf

$$f(t,\theta) = \begin{cases} 0, & \text{if } t < 0, \\ \lambda\, e^{-\lambda\, t}, & \text{if } t \geq 0. \end{cases}$$

Find the maximum likelihood estimator of θ. Is the MLE unbiased?

Solution. Suppose we are given the realisations x_1, $\ldots$, x_n. Note that the likelihood function is given as

$$L(\lambda) = \lambda^n \exp\left(-\lambda \sum_{k=1}^{n} x_i\right) \prod_{k=1}^{n} u(x_i),$$

where u is the step function. Therefore the derivative of the log-likelihood with respect to λ is,

$$\frac{\partial}{\partial\lambda} \log L(\lambda) = \frac{n}{\lambda} - \sum_{k=1}^{n} x_i.$$

Setting the log-likelihood to zero and solving the resulting equation, we find the ML estimate as

$$\hat{\lambda}_{\text{ML}} = \frac{n}{\sum_{k=1}^{n} x_i}.$$

To see if this estimator is unbiased or not, note that $\mathbb{E}(X_i) = 1/\lambda$ (check this!). Therefore

$$\mathbb{E}\underbrace{\Big(n^{-1}\sum_{i=1}^{n} X_i\Big)}_{T(X)} = 1/\lambda.$$

Now let $g(t) = 1/t$ and observe that for $t > 0$, g is a strictly convex function. Therefore by Jensen's inequality, we have,

$$\mathbb{E}\left(\frac{n}{\sum_{k=1}^{n} X_i}\right) = \mathbb{E}\Big(g(T)\Big) > g\Big(\mathbb{E}(T)\Big) = \lambda.$$

Thus the estimator is biased.

2. Consider a biased coin with $P(\text{Heads}) = p$, where p is an unknown constant of interest. In order to estimate it, we toss the coin n times. Suppose we define the random variables,

$$X_k = \begin{cases} 1, & \text{if the } k^{\text{th}} \text{ toss is a Head,} \\ 0, & \text{if the } k^{\text{th}} \text{ toss is a Tail,} \end{cases}$$

for $k = 1, 2, \ldots, n$. Find the maximum likelihood estimator (MLE) of p in terms of X_1, X_2, $\ldots$, X_n. Is the MLE unbiased?

Solution. Note that we can express the PMF of a single X_i as,

$$P(x) = p^x (1-p)^{1-x}, \text{if } x \in \{0,1\}.$$

Therefore, given the realisations $x_1, \ldots, x_n$, the likelihood function is

$$L(p) = p^{\sum_i x_i} (1-p)^{\sum_i (1-x_i)}.$$

The derivative of the log-likelihood with respect to p is,

$$\frac{\partial}{\partial p} \log L(p) = \frac{1}{p} \sum_i x_i - \frac{1}{1-p} \sum_i (1-x_i).$$

Setting this to zero and solving for $\hat{p}_{\mathrm{ML}}$, we obtain (check this!),

$$\hat{p}_{\mathrm{ML}} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

Observe that

$$\mathbb{E}(\hat{p}_{\mathrm{ML}}) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}(X_i) = p.$$

Thus the MLE is unbiased.

3. Suppose $X_1, X_2, \ldots, X_n$ are iid, $\mathcal{N}(0, \theta)$ random variables.

   (a) Find the MLE for θ. Is the MLE unbiased?

   (b) Let $\gamma = 1/\theta$. Find the MLE for γ. Is the MLE for γ unbiased for γ?

   Solution. (a) Note that the log-likelihood function is given as

   $$L(\theta) = \frac{1}{(2\pi\theta)^{n/2}} \exp\left(-\frac{1}{2\theta} \sum_{i=1}^{n} x_i^2\right).$$

   The derivative of the log-likelihood with respect to θ is,

   $$\frac{\partial}{\partial \theta} \log L(\theta) = -\frac{n}{2\theta} + \frac{1}{2\theta^2} \sum_{i=1}^{n} x_i^2.$$

   Setting this to zero and solving for $\hat{\theta}_{\mathrm{ML}}$, we obtain,

   $$\hat{\theta}_{\mathrm{ML}} = \frac{1}{n} \sum_{i=1}^{n} x_i^2.$$

   We have,

   $$\mathbb{E}(\hat{\theta}_{\mathrm{ML}}) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}(X_i^2) = \theta.$$

   Thus the MLE is unbiased.

(b) Recall that if $\gamma = g(\theta)$, then the ML estimators satisfy $\hat{\gamma}_{\text{ML}} = g(\hat{\theta}_{\text{ML}})$. For this question, the function g is $g(t) = 1/t$. Therefore, the ML estimator for γ is

$$\hat{\gamma}_{\text{ML}} = \frac{n}{\sum_{i=1}^{n} X_i^2}.$$

By Jensen's inequality, it follows that this estimator is biased (see Q1 above).

4. (From textbook) Suppose we have n iid observations of an unknown constant μ of the form

$$X_i = \mu + Z_i$$

where $Z_i \sim \mathcal{N}(0, \sigma^2)$, where σ is unknown. Find the MLE for the signal to noise ratio $\alpha = \mu^2/\sigma^2$.

Solution. Note that the joint pdf of X_i's is,

$$\begin{aligned} f_X(x) &= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 \right) \\ &= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2} \left(\sum_{i=1}^{n} x_i^2 \right) + \frac{\mu}{\sigma^2} \left(\sum_{i=1}^{n} x_i \right) - \frac{n\mu^2}{2\sigma^2} \right). \end{aligned}$$

Therefore, in terms of the unknowns (μ, α) the likelihood function is,

$$L(\mu, \alpha) = \frac{\alpha^{n/2}}{(2\pi\mu^2)^{n/2}} \exp\left(-\frac{\alpha}{2\mu^2} \left(\sum_{i=1}^{n} x_i^2 \right) + \frac{\alpha}{\mu} \left(\sum_{i=1}^{n} x_i \right) - \frac{n}{2}\alpha \right).$$

The partial derivatives with respect to μ and α are given as,

$$\begin{aligned} \frac{\partial}{\partial \mu} L(\mu, \alpha) &= -\frac{n}{\mu} + \frac{\alpha}{\mu^3} s - \frac{\alpha}{\mu^2} t, \\ \frac{\partial}{\partial \alpha} L(\mu, \alpha) &= \frac{n}{2\alpha} - \frac{1}{2\mu^2} s + \frac{1}{\mu} t - \frac{n}{2}, \end{aligned}$$

where $s = \sum_i x_i^2$, and $t = \sum_i x_i$. For $\hat{\mu}_{\text{ML}}$, and $\hat{\alpha}_{\text{ML}}$, these equations evaluate to zero. Therefore we need to solve a nonlinear system of equations given as,

$$\begin{aligned} -\frac{n}{\mu} + \frac{\alpha}{\mu^3} s - \frac{\alpha}{\mu^2} t &= 0 \\ \frac{n}{2\alpha} - \frac{1}{2\mu^2} s + \frac{1}{\mu} t - \frac{n}{2} &= 0. \end{aligned}$$

Multiplying the first equation by μ^3 and the second equation by $2\alpha\,\mu^2$, we obtain an equivalent system as,

$$\begin{aligned} -n\mu^2 + \alpha\, s - \alpha\, \mu\, t &= 0 \\ n\mu^2 - \alpha\, s + 2\alpha\mu\, t - n\alpha\, \mu^2 &= 0. \end{aligned}$$

Summing the first and the second equations we obtain the system

$$\begin{aligned} -n\mu^2 + \alpha\, s - \alpha\, \mu\, t &= 0 \\ \alpha\mu\, t - n\alpha\, \mu^2 &= 0. \end{aligned}$$

From the second equation of this new system, we find $\hat{\mu}_{\text{ML}} = t/n$. Plugging this in the first equation, we find

$$\hat{\alpha}_{\text{ML}} = \frac{n\hat{\mu}_{\text{ML}}^2}{s - \hat{\mu}_{\text{ML}}\, t} = \frac{t^2}{n\, s - t^2} = \frac{\left(\sum_i x_i\right)^2}{n\, \left(\sum_i x_i^2\right) - \left(\sum_i x_i\right)^2}.$$

TEL 502E – Detection and Estimation Theory

Midterm Examination

24.03.2015

(25 pts) 1. Suppose X_1 and X_2 are independent random variables distributed as $\mathcal{N}(2\theta, 1)$ and $\mathcal{N}(3\theta, 1)$ respectively, where θ is an unknown parameter.

(a) Write down the joint pdf of X_1 and X_2.

(b) Compute the Fisher information for θ, that is,

$$I(\theta) = \mathbb{E}\left(\left[\partial_\theta\Big(\ln f(X_1, X_2; \theta)\Big)\right]^2\right),$$

where $f(X_1, X_2; \theta)$ denotes the joint pdf of X_1 and X_2.

(c) Find an unbiased estimator for θ in terms of X_1 and X_2.

(d) Find the UMVUE for θ in terms of X_1 and X_2.

(25 pts) 2. Suppose X_1, X_2 are independent random variables distributed as $\mathcal{N}(0, \theta)$, $\mathcal{N}(0, 2\theta)$, where θ is an unknown positive constant.

(a) Find an unbiased estimator for θ in terms of X_1 and X_2.

(b) Find a sufficient statistic for θ.

(c) Find the UMVUE for θ (please explain briefly why you think the estimator is the UMVUE.)

(25 pts) 3. Suppose $X_1, X_2, \ldots, X_n$ are independent and identically distributed random variables with pdf

$$f_{X_i}(t) = \begin{cases} 0, & \text{if } t < 0, \\ \theta^{-t}/\ln(\theta), & \text{if } t \geq 0, \end{cases}$$

where $\theta > 1$ is an unknown constant.

(a) Find the maximum likelihood estimator for θ in terms of $X_1, X_2, \ldots, X_n$.

(b) Specify whether the estimator you found is biased or not.
(Hint : $\int_0^\infty x\, c^{-x}\, dx = 1/\ln(c)$, if $c > 1$.)

(25 pts) 4. Suppose we observe $X = \theta + Z$, where θ and Z are independent random variables. Suppose also that θ is uniformly distributed over the unit interval and Z is a standard normal random variable (i.e., $\mathcal{N}(0, 1)$). That is, the pdfs of θ and Z are,

$$f_\theta(t) = u(t)\, u(1 - t),$$

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2},$$

where u denotes the step function.

(a) Find the joint pdf of X and θ, that is, $f_{X,\theta}(x, t)$.

(b) Find the maximum a posteriori (MAP) estimator for θ in terms of X.

(c) Evaluate the estimator you found in part (b) if the observation is given as

(c.1) $x = 1/4$,

(c.2) $x = -1$,

(c.3) $x = 2$.

TEL 502E – Detection and Estimation Theory

Final Examination

24.05.2015

Student Name : ____________________

Student Num. : ____________________

4 Questions, 100 Minutes

(25 pts) 1. Suppose X_1, X_2, X_3 are independent random variables and the pdf of X_k is given as,

$$f_k(t) = \begin{cases} \frac{1}{k\theta} \exp\left(-\frac{t}{k\theta}\right), & \text{if } 0 \le t, \\ 0, & \text{if } t \le 0, \end{cases}$$

for $k = 1, 2, 3$, where θ is a positive unknown.

(a) Find a sufficient statistic for θ and compute its expected value.

(b) Find a function of the sufficient statistic which is unbiased as an estimator of θ.

(Note : $\int_0^\infty t\, e^{-t}\, dt = 1$.)

(25 pts) 2. Suppose X is an exponential random variable with probability density function (pdf)

$$f_X(u) = \begin{cases} e^{-u}, & \text{if } 0 \le u, \\ 0, & \text{if } u \le 0, \end{cases}$$

and we observe $Y = X + Z$, where Z is a standard normal random variable (i.e., zero-mean, unit variance Gaussian). Suppose also that X and Z are independent.

(a) Write down the joint pdf Y and X, namely $f_{Y,X}(t, u)$.

(b) Find the maximum a posteriori (MAP) estimator of X given Y.

(c) Evaluate the estimator you found in part (b) for (i) $Y = -2$, (ii) $Y = 0$, (iii) $Y = 2$.

(25 pts) 3. Suppose X_k for $k = 1, 2, 3$ are random variables of the form

$$X_k = k\,\theta + Y_k,$$

where Y_k's are independent standard normal random variables (i.e., zero-mean, unit variance Gaussian).

(a) Find an unbiased estimator for θ.

(b) Find the maximum likelihood estimator (MLE) for θ.

(c) Determine whether the MLE is biased or not.

(25 pts) 4. Suppose X_1, X_2 are independent identically distributed standard normal random variables. We make two observations $Y_k = \theta\, X_k$ where θ is known to be either 1 or 2. We would like to decide which value θ took by studying the realizations of Y_k, namely y_k. We form two hypotheses as

$H_0 : \theta = 1,$

$H_1 : \theta = 2.$

(a) Find the pdf of Y_1 under H_1.

(b) Find the Neyman-Pearson test for the given hypotheses. That is, find a test statistic $g(y_1, y_2)$ such that

$$\begin{cases} \text{if } g(y_1, y_2) > \gamma, & \text{then we decide } H_0, \\ \text{if } g(y_1, y_2) \le \gamma, & \text{then we decide } H_1. \end{cases}$$

(c) For the test in part (b), find the threshold γ so that the probability of a Type-I error is α. For this part, you can assume that $\varphi(t)$ denotes the cdf of a chi-square random variable with two degrees of freedom, and express your answer in terms $\varphi(t)$.
(Recall that we make a Type-I error if we decide H_1 while H_0 is true.)