TEL 502E – Detection and Estimation Theory

Spring 2014

Instructors : İlker Bayram
ibayram@itu.edu.tr

Class Meets : Tuesday, 9.30 – 12.30, EEB 5307

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.

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

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

Attendance : You need to attend at least 70% of the lectures to sit for the final exam.

Tentative Course Outline

(1) Review of probability theory

(2) Simple Hypothesis Testing, the Neyman Pearson Lemma

(3) Bayesian Tests, Multiple Hypothesis Testing

(4) The detection problem under different scenarios

(5) The estimation problem, minimum variance unbiased estimators

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

(7) Linear Estimators, maximum likelihood estimation

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

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

TEL502E – Homework 1

Due 25.02.2014

1. (a) Suppose that X is a non-negative random variable with a pdf $f_X(t)$ (that is, $f_X(t) = 0$ for $t < 0$). Show that, for any $n > 0$ and $s > 0$,

$$P\big(\{X \geq s\}\big) \leq \frac{\mathbb{E}(X^n)}{s^n}.$$

(b) Using part (a), show that for an arbitrary random variable Y with $\mathbb{E}(Y) = \mu$,

$$P\big(\{\mu - \epsilon \leq Y \leq \mu + \epsilon\}\big) \geq 1 - \frac{\mathrm{var}(Y)}{\epsilon^2}.$$

(c) Suppose that $X_1, X_2, \ldots$ is a sequence of iid random variables with $\mathbb{E}(X_i) = \mu$, $\mathrm{var}(X_i) = \sigma^2$. Also let,

$$Z_n = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

Compute $\mathbb{E}(Z_n)$ and $\mathrm{var}(Z_n)$.

(d) Show that

$$\lim_{n\to\infty} P\big(\{\mu - \epsilon < Z_n < \mu + \epsilon\}\big) = 1,$$

for any $\epsilon > 0$.

Solution. (a) Keeping in mind that $s > 0$, we have,

$$\begin{aligned}
P\big(\{X \geq s\}\big) &= \int_s^\infty f_X(t)\,dt \\
&\leq \int_0^s \frac{t^n}{s} f_X(t)\,dt + \int_s^\infty f_X(t)\,dt \\
&\leq \int_0^s \frac{t^n}{s} f_X(t)\,dt + \int_s^\infty \frac{t^n}{s^n} f_X(t)\,dt \\
&= \frac{\mathbb{E}(X^n)}{s^n}.
\end{aligned}$$

This inequality is known as Markov's inequality.

(b) Using Y suppose we define a new random variable as $Z = |Y - \mu|$. Then, using Markov's inequality with $n = 2$, we have,

$$P\big(\{Z \geq \epsilon\}\big) \leq \frac{\mathbb{E}(Z^2)}{\epsilon^2} = \frac{\mathrm{var}(Y)}{\epsilon^2}.$$

Observe now that

$$P\big(\{Z \geq \epsilon\}\big) + P\big(\{Z < \epsilon\}\big) = 1,$$

since the two events partition the sample space. This implies,

$$P\big(\{Z < \epsilon\}\big) \geq 1 - \frac{\mathrm{var}(Y)}{\epsilon^2}$$

But now observe that

$$\{Z < \epsilon\} = \{|Y - \mu| < \epsilon\} = \{-\epsilon < Y - \mu < \epsilon\} = \{\mu - \epsilon \leq Y \leq \mu + \epsilon\}.$$

Thus the claim follows. This inequality (or an equivalent version) is known as Chebyshev's inequality.

(c) First,

$$\mathbb{Z}_n = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}(X_i) = \mu.$$

Now, note when the random variables are independent, we can add their variances. Thus,

$$\mathrm{var}(Z_n) = \sum_{i=1}^{n} \mathrm{var}\big(X_i/n\big) = \sum_{i=1}^{n} \frac{\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

(d) Since $\mathbb{E}(Z_n) = \mu$, we can use the result of part (b). That gives,

$$P\big(\{\mu - \epsilon < Z_n < \mu + \epsilon\}\big) \geq 1 - \frac{\sigma^2}{\epsilon^2\, n}.$$

Letting $n \to \infty$, the right hand side converges to 1 and the claim follows.

2. (a) Show that if $\mathrm{var}(Y) = 0$, then $P\big(\{Y = \mathbb{E}(Y)\}\big) = 1$.
   (b) Show that if $\mathbb{E}(Y^2) = 0$, then $P\big(\{Y = 0\}\big) = 1$.

Solution. (a) Let A be the event of interest defined as,

$$A = \{Y = \mathbb{E}(Y)\}.$$

Instead of $P(A)$, we will compute the $P(A^c)$. Now observe that,

$$A^c = \{|Y - \mathbb{E}(Y)| > 0\} = \cup_{n=1}^{\infty} \underbrace{\{|Y - \mathbb{E}(Y)| > 1/n\}}_{B_n}.$$

But by part (b) of Question-1, we have that $P(B_n) = 0$. Therefore,

$$P(A^c) \leq \sum_{n=1}^{\infty} P(B_n) = 0.$$

Since $P(A^c) \geq 0$ by definition, it follows that $P(A^c) = 0$. Thus, $P(A) = 1 - P(A^c) = 1$.

(b) Since $\mathrm{var}(Y) = \mathbb{E}(Y^2) - \big(\mathbb{E}(Y)\big)^2 \geq 0$, the condition '$\mathbb{E}(Y^2) = 0$' implies that $\mathbb{E}(Y) = 0$. The desired equality follows therefore follows from part (a).

3. Suppose X is a discrete random variable, taking values on the set of integers $\mathbb{Z}$. Suppose we are testing whether X is distributed according to the probability mass function (PMF) $P_0(t)$ (this is the null hypothesis) or it's distributed according to the PMF $P_1(t)$ (this is the alternative hypothesis). We somehow form the acceptance region $C \subset \mathbb{Z}$ such that if a realization of X, say x falls in C, we accept the null hypothesis, and reject it otherwise. Also, let $p_I(C)$ and $p_{II}(C)$ denote the probabilities of type-I and type-II errors of this test. Below, the parts (a) and (b) are independent of each other.

   (a) Suppose we discover that for some $r \in (\mathbb{Z} \backslash C)$ and $a_1, a_2, \ldots a_n \in C$,
      (i) $P_0(r) = \sum_{i=1}^{n} P_0(a_i)$, and
      (ii) $\big(P_0(r)/P_1(r)\big) > \big(P_0(a_i)/P_1(a_i)\big)$ for $i = 1, 2, \ldots, n$.

   Based on this observation, we decide to update the acceptance region and use $D = C \cup \{r\} \backslash \{a_1, \ldots, a_n\}$ as the acceptance region (i.e., we remove a_i's and include r in the new acceptance region). Let $p_I(D)$ and $p_{II}(D)$ denote the type-I and type-II error probabilities for this updated test. Show that $p_I(D) \leq p_I(C)$, and $p_{II}(D) < p_{II}(C)$.

   (b) Suppose we find that for any $r \in (\mathbb{Z} \cap C^c)$, and $a \in C$, the inequality

$$\frac{P_0(r)}{P_1(r)} < \frac{P_0(a)}{P_1(a)} \tag{1}$$

   is satisfied. Consider now another test than the one described above with an acceptance region given as D, whose type-I and type-II error probabilities are given as $p_I(D)$ and $p_{II}(D)$ respectively. Show that if $p_I(D) \leq p_I(C)$, then $p_I(D) > p_{II}(D)$.

Solution. Notice that, in this setting, for an acceptance region denoted as C, the type-I and type-II error probabilities are given by

$$p_I(C) = \sum_{x \in \mathbb{Z} \cap C^c} P_0(x), \quad p_{II}(C) = \sum_{x \in C} P_1(x).$$

(a) First, observe that, by condition (i), we have,

$$p_I(D) - p_I(C) = \sum_{z \in \mathbb{Z} \cap D^c} P_0(z) - \sum_{z \in \mathbb{Z} \cap C^c} P_0(z) = \sum_{i=1}^{n} P_0(a_i) - P_0(r) = 0.$$

Rewriting (ii) as,

$$\frac{P_1(a_i)}{P_1(r)} > \frac{P_0(a_i)}{P_0(r)}, \text{ for } i = 1, 2, \ldots, n,$$

and summing over i, we obtain,

$$\frac{\sum_{i=1}^n P_1(a_i)}{P_1(r)} > \frac{\sum_{i=1}^n P_0(a_i)}{P_0(r)} = 1,$$

where we made use of (i) again. Now observe that,

$$p_{II}(D) - p_{II}(C) = \sum_{z \in D} P_1(z) - \sum_{z \in C} P_1(z) = P_1(r) - \sum_{i=1}^{n} P_1(a_i) < 0.$$

(b) Suppose $p_I(D) \leq p_I(c)$. This implies,

$$p_I(D) - p_I(C) = \sum_{x \in \mathbb{Z} \cap D^c} P_0(x) - \sum_{x \in \mathbb{Z} \cap C^c} P_0(x) = \sum_{x \in D^c \cap C} P_0(x) - \sum_{x \in C^c \cap D} P_0(x) \leq 0,$$

or

$$\frac{\sum_{x \in C^c \cap D} P_0(x)}{\sum_{x \in D^c \cap C} P_0(x)} \geq 1. \tag{2}$$

Now observe similarly that

$$p_{II}(D) - p_{II}(C) = \sum_{x \in D} P_1(x) - \sum_{x \in C} P_1(x) = \sum_{x \in D \cap C^c} P_1(x) - \sum_{x \in C \cap D^c} P_1(x).$$

Thus, if we can show that

$$\frac{\sum_{x \in D \cap C^c} P_1(x)}{\sum_{x \in C \cap D^c} P_1(x)} > 1, \tag{3}$$

we are done.
For this, we first rewrite (1) in a different form. Note that if $x \in C^c$ and $c \in C$, then

$$P_1(c)\, P_0(x) < P_0(c)\, P_1(x).$$

Fixing $c \in C$, we obtain,

$$P_1(c) \left(\sum_{x \in D \cap C^c} P_0(x) \right) < P_0(c) \left(\sum_{x \in D \cap C^c} P_1(x) \right).$$

Now taking the terms inside the parentheses as fixed, we can write,

$$\left(\sum_{c \in D^c \cap C} P_1(c) \right) \left(\sum_{x \in D \cap C^c} P_0(x) \right) < \left(\sum_{c \in D^c \cap C} P_0(c) \right) \left(\sum_{x \in D \cap C^c} P_1(x) \right).$$

Rewriting and using (2), we obtain (3) :

$$\frac{\sum_{x \in D \cap C^c} P_1(x)}{\sum_{x \in D^c \cap C} P_1(x)} > \frac{\sum_{x \in D \cap C^c} P_0(x)}{\sum_{x \in D^c \cap C} P_0(x)} \geq 1.$$

Notice that throughout, I assumed that $P_i(x)$ is non-zero as long as $x \in \mathbb{Z}$. I leave it to you to consider how to modify the argument if $P_i(x) = 0$ from some x.

TEL502E – Homework 2

Due 11.03.2014

1. Consider a hypothesis testing problem as follows. We are given the realisation of a random vector x of length k. There are two hypothesis concerning x, namely H_0 and H_1, which suggest that x_i's are of the form

$$\begin{aligned}
&H_0 : x_i = n_i,\\
&H_1 : x_i = s_i + n_i,
\end{aligned}$$

where $n_1, \ldots n_k$ is the realisation of a random vector with iid zero-mean Gaussian components with variance σ^2 and $s_1, \ldots, s_k$ is a known constant vector. We have seen in class that the Neyman-Pearson test for this scenario employs the test statistic $T(x) = \sum_{i=1}^{k} s_i\, x_i$ and is of the form

$$\begin{cases}
\text{If } T(x) < \gamma, & \text{then decide } H_0\\
\text{If } T(x) \geq \gamma, & \text{then decide } H_1.
\end{cases}$$

Recall that the $Q(\cdot)$ function is defined in terms of the pdf of a standard Gaussian random variable $f(\cdot)$ as

$$Q(t) = \int_t^\infty f(s)\, ds = \int_t^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}s^2\right)\, ds$$

(a) Determine γ so that the probability of a Type-I error is α. Express this threshold in terms the $Q(\cdot)$ function.

(b) For the threshold computed in part (a), find the probability of a Type-II error.

Solution. (a) Recall that a Type-I error occurs when H_0 is true but we decide H_1. This is equivalent to saying that the event $\{T' < \gamma\}$ where X is assumed to be distributed as an iid Gaussian random vector where each component has variance σ^2. Now we need to determine how $T' = \sum_i s_i\, X_i$ is distributed. Note that since T is a linear combination of zero-mean Gaussian random variables, it is also a zero-mean Gaussian random variable. Also, since $s_i\, X_i$'s are independent, the variances add (why?) and the variance of T' is found to be $\sigma_T^2 = \sigma^2 \sum_i s_i^2 = \sigma^2\, \varepsilon^2$, where ε^2 denotes the energy of the deterministic signal s. Observe now that $\sigma_T^{-1} T'$ is a standard Gaussian random variable. Thus, probability of a type-I error can be computed as,

$$\begin{aligned}
P(\text{Type-I}) &= P\{T' > \gamma\} \quad \text{under } H_0\\
&= P\{\sigma_T^{-1}\, T' > \sigma_T^{-1}\gamma\} \quad \text{under } H_0\\
&= Q(\sigma_T^{-1}\gamma).
\end{aligned}$$

If this probability is desired to be less than or equal to α, the γ value that minimizes the probability of a type-II error is, $\gamma = \sigma_T\, Q^{-1}(\alpha)$.

(b) For the threshold found in part (a), a type-II error occurs if the event $\{T < \gamma\}$ occurs while X is distributed as a random vector with iid entries where X_i is Gaussian with mean s_i and variance σ^2 (this is what H_1 claims). Under this hypothesis, we need to find how $T' = \sum_i s_i\, X_i$ is distributed. Arguing as in (a), T' is a Gaussian r.v. But this time, the mean is $\sum_i s_i^2 = \varepsilon^2$ (why?) and the variance is $\sigma^2 \sum_i s_i^2 = \sigma^2\, \varepsilon^2 = \sigma_T^2$ (as in part (a)). Observe in this case that '$\sigma_T^{-1}(T - \varepsilon^2)$' is a standard Gaussian random variable. Thus,

$$\begin{aligned}
P(\text{Type-II}) &= P\{T' < \gamma\} \quad \text{under } H_1\\
&= P\{\sigma_T^{-1}\,(T' - \varepsilon^2) < \sigma_T^{-1}(\gamma - \varepsilon^2)\} \quad \text{under } H_1\\
&= 1 - Q\big(\sigma_T^{-1}(\gamma - \varepsilon^2)\big)\\
&= 1 - Q\left(Q^{-1}(\alpha) - \frac{\varepsilon}{\sigma}\right).
\end{aligned}$$

2. Let us slightly complicate the problem in Question-1 by introducing another constant signal, namely $z = (z_1, \ldots z_k)$, into the scenario. Suppose that the two hypothesis are now of the form,

$$\begin{aligned}
&H_0 : x_i = z_i + n_i,\\
&H_1 : x_i = s_i + n_i,
\end{aligned}$$

where the rest of the variables are as described in Question-1.

(a) Find a test statistic $T'(x)$ such that the Neyman-Pearson test is of the form

$$\begin{cases} \text{If } T'(x) < \gamma, & \text{then decide } H_0 \\ \text{If } T'(x) \geq \gamma, & \text{then decide } H_1. \end{cases}$$

(b) For the test in part (a), determine the constant γ, in terms of the $Q(\cdot)$ function so that the probability of a Type-I error is equal to α.

(c) Determine the probability of a Type-II error for the threshold computed in part(b).

Solution. Given x, suppose we define the data vector $y = x - z$. Also, let $r = s - z$. Under this change of variables, notice that H_0 and H_1 can be expressed as,

$$H_0 : y_i = n_i,$$
$$H_1 : y_i = r_i + n_i.$$

Thus, the problem is reduced to the problem in Question-1. All we need to do is to translate the results while paying attention to the change of variables.

(a) Note that under this change of variables, one can use

$$\sum_i r_i\, y_i = \sum_i (s_i - z_i)\,(x_i - z_i) = \sum_i (s_i - z_i)\, x_i + \text{const.}$$

as a statistic. Dropping the constant term, we find a test statistic given as $T' = \sum_i (s_i - z_i) X_i$.

(b) Let $\varepsilon^2 = \sum_i (s_i - z_i)^2$. Then, the best threshold is (see Question-1 part(a)), $\gamma = \sigma\,\varepsilon\, Q^{-1}(\alpha)$.

(c) In this case, for $\varepsilon = \sum_i (s_i - z_i)^2$, the probability of a type-II error is (see Question-1, part(b)),

$$P(\text{Type-II}) = 1 - Q\left(Q^{-1}(\alpha) - \frac{\varepsilon}{\sigma}\right).$$

TEL502E – Homework 3

Due 22.04.2014

1. Suppose X_1, X_2 are independent identically distributed random variables with a probability density function (pdf) given as,

$$f(t) = \frac{1}{2}\, e^{-|t|}.$$

We make two observations y_1, y_2, related to realizations of X_i's. Suppose there are two hypotheses concerning the observations

$$H_0 : y_i = x_i, \text{ for } i = 1, 2,$$
$$H_1 : y_i = 2\, x_i, \text{ for } i = 1, 2,$$

where x_i's are realizations of X_i's.

(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. (Recall that we make a Type-I error if we decide H_1 while H_0 is true.)

Solution. (a) To write the pdf let us first find the cdf :

$$F_{Y_1}(t) = P(Y_1 \le t) = P(2\, X_1 \le t) = F_X(t/2).$$

Differentiating, we obtain,

$$f_{Y_1}(t) = \frac{1}{2}\, f_X(t/2) = \frac{1}{4}\, e^{-|t|/2}.$$

(b) The LRT statistic for this problem is,

$$T = \frac{f_0(y_1, y_2)}{f_1(y_1, y_2)} = \frac{1}{2}\, \exp\Big(-\big(|y_1| + |y_2|\big)/2\Big)$$

Note that the test is of the form,

$$\begin{cases} \text{If } T \ge \beta, & \text{decide } H_0 \\ \text{If } T < \beta, & \text{decide } H_1. \end{cases}$$

Taking logarithms, we find an equivalent test as,

$$\begin{cases} \text{If } -\big(|y_1| + |y_2|\big) \ge \gamma, & \text{decide } H_0 \\ \text{If } -\big(|y_1| + |y_2|\big) < \gamma, & \text{decide } H_1. \end{cases}$$

Therefore $g(y_1, y_2) = -\big(|y_1| + |y_2|\big)$ works.

(c) In order to evaluate the probability of error, we need to find the pdf pf the test statistic. Let us first find the pdf of $U_1 = |Y_1|$. Note that the cdf of U_1 is,

$$F_{U_1}(t) = P(U_1 \le t) = \begin{cases} 2\int_0^t f_{Y_1}(s)\, ds, & \text{if } t \ge 0, \\ 0, & \text{if } t < 0. \end{cases}$$

Differentiating with respect to t, we find the pdf of U_1 as,

$$f_{U_1}(t) = \begin{cases} 2\, f_{Y_1}(t) = e^{-|t|}, & \text{if } t \ge 0, \\ 0, & \text{if } t < 0. \end{cases}$$

Now let $S = U_1 + U_2$. Since U_1 and U_2 are independent (note that they are also identically distributed). The pdf of S can be found by convolving the pdfs of U_1 and U_2.

$$f_S(t) = \begin{cases} \int_0^t e^{-(t-s)}\, e^{-s}\, ds = t\, e^{-t}, & \text{if } t > 0, \\ 0, & \text{if } t < 0. \end{cases}$$

Now observe that

$$\begin{aligned} P(g > \gamma) &= P(S < -\gamma) \\ &= \int_0^{-\gamma} t\, e^{-t} dt \\ &= (-t\, e^{-t})\Big|_0^{-\gamma} + \int_0^{-\gamma} e^{-t} dt \\ &= \gamma\, e^{\gamma} + (1 - e^{\gamma}). \end{aligned}$$

Therefore if

$$h(\gamma) = 1 + (\gamma - 1)e^{\gamma} = \alpha,$$

set $\gamma = h^{-1}(\alpha)$ (note that $h(\gamma)$ is invertible for $\gamma < 0$).

2. Suppose we observe a signal x_i for $i = 1, 2, \ldots, k$. There are two hypothesis of the form,

$$\begin{aligned} H_0 &: x_i = z_i + n_i, \\ H_1 &: x_i = s_i + n_i, \end{aligned}$$

where z_i and s_i are deterministic (and known signals) and n_i's are realizations of iid Gaussian random variables with unit variance.

(a) Find a test statistic $T(x)$ such that the Neyman-Pearson test is of the form

$$\begin{cases} \text{If } T(x) < \gamma, & \text{then decide } H_0 \\ \text{If } T(x) \geq \gamma, & \text{then decide } H_1. \end{cases}$$

(b) Recall that the Q function is defined as,

$$Q(t) = \int_t^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-t^2/2}\, dt.$$

For the test in part (a), determine the constant γ, in terms of the $Q(\cdot)$ function so that the probability of a Type-I error is equal to α.

(c) Determine the probability of a Type-II error for the threshold computed in part(b). (Recall that we make a Type-II error if we decide H_0 while H_1 is true.)

Solution. (a) Let f_i, denote pdf of the data under the hypothesis H_i. Then the LRT statistic is given by,

$$\begin{aligned} \frac{f_0}{f_1} &= \exp\left(-\sum_{i=1}^{n} \left[(x_i - z_i)^2 - (x_i - s_i)^2 \right] \right) \\ &= \exp\left(-\sum_{i=1}^{n} \left[2x_i\,(s_i - z_i) + (s_i - z_i)^2 \right] \right) \end{aligned}$$

Taking logarithms and discarding constanat terms, we find a test equivalent to the LRT as,

$$\underbrace{\sum_{i=1}^{n} x_i (s_i - z_i)}_{T(x)} \begin{cases} < \gamma & \Rightarrow \text{ decide } H_0, \\ \geq \gamma & \Rightarrow \text{ decide } H_1. \end{cases}$$

(b) Note that under H_0, $T(X)$ is distributed as a Gaussian with mean $\mu = \sum_i z_i(s_i - z_i)$ and variance $\sigma^2 = \sum_i (z_i - s_i)^2$. Therefore, the probability of a type-I error is,

$$P(T > \gamma) = P\left(\frac{T-\mu}{\sigma} > \frac{\gamma-\mu}{\sigma}\right) = Q\left(\frac{\gamma-\mu}{\sigma}\right).$$

If this is to be equal to α, the threshold γ should be,

$$\gamma = Q^{-1}(\alpha)\sigma + \mu.$$

(c) Under H_1, $T(X)$ is distributed as a Gaussian with mean $\hat{\mu} = \sum_i s_i(s_i - z_i)$ and variance $\sigma^2 = \sum_i (z_i - s_i)^2$. Therefore, the probability of a type-II error is,

$$P(T < \gamma) = P\left(\frac{T-\hat{\mu}}{\sigma} > \frac{\gamma-\hat{\mu}}{\sigma}\right) = 1 - Q\left(\frac{\gamma-\hat{\mu}}{\sigma}\right) = 1 - Q\left(Q^{-1}(\alpha) - \sqrt{\sum_{i=1}^{n}(z_i - s_i)^2}\right).$$

3. Let X_1, X_2 be independent Gaussian random variables with the same mean θ but with different variances σ_1^2, σ_2^2. That is, the probability density functions of X_1 and X_2 are,

$$f_1(t) = \frac{1}{\sqrt{2\pi\,\sigma_1^2}} \exp\left(-\frac{(t-\theta)^2}{2\sigma_1^2}\right),$$
$$f_2(t) = \frac{1}{\sqrt{2\pi\,\sigma_2^2}} \exp\left(-\frac{(t-\theta)^2}{2\sigma_2^2}\right).$$

(a) Find an expression for the Cramér-Rao lower bound for θ (in terms of σ_1^2 and σ_2^2).

(b) Find the uniformly minimum variance unbiased (UMVU) estimator for θ.

Solution. Note that the joint pdf of $X = (X_1, X_2)$ is the product of the two pdfs. We compute

$$\begin{aligned}
\mathbb{E}\left(\left[\frac{\partial}{\partial\theta}\ln f(t_1, t_2; \theta)\right]^2\right) &= \mathbb{E}\left(\left[\frac{1}{\sigma_1^2}(\theta - X_1) + \frac{1}{\sigma_2^2}(\theta - X_2)\right]^2\right) \\
&= \mathbb{E}\left(\frac{1}{\sigma_1^4}(\theta - X_1)^2 + \frac{1}{\sigma_2^4}(\theta - X_2)^2\right) \\
&= \sigma_1^{-2} + \sigma_2^{-2}.
\end{aligned}$$

Therefore the CRLB is $\left(\sigma_1^{-2} + \sigma_2^{-2}\right)^{-1}$.

Note that we can write,

$$\begin{aligned}
\frac{\partial}{\partial\theta}\ln f(X_1, X_2; \theta) &= \underbrace{\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)}_{\sigma^{-2}}\theta - \frac{X_1}{\sigma_1^2} - \frac{X_2}{\sigma_1^2} \\
&= \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)\left[\theta - \underbrace{\left(\frac{\sigma^2}{\sigma_1^2}X_1 + \frac{\sigma^2}{\sigma_2^2}X_2\right)}_{\hat{\theta}(X_1, X_2)}\right].
\end{aligned}$$

Thus, $\hat{\theta}(X_1, X_2)$ must be the UMVUE.

4. Suppose X_1, X_2, X_3 are independent random variables uniformly distributed on $[0, \theta]$. Recall that a random variable uniformly distributed on $[0, \theta]$ has pdf

$$f(t) = \begin{cases} \theta^{-1}, & \text{if } t \in [0, \theta], \\ 0, & \text{if } t \notin [0, \theta]. \end{cases}$$

We are interested in the value of θ.

(a) A student suggests $\hat{\theta} = X_1 + X_2 + X_3$ as an estimator. Find out whether this estimator is biased or not. If it is biased, derive an unbiased estimator $\tilde{\theta}$ which is a function of $\hat{\theta}$ only (i.e., the other random variables do not appear in the new expression).

(b) An unbiased estimator for θ is $\bar{\theta} = X_1 + X_2$. Compare the variances of $\bar{\theta}$ and $\tilde{\theta}$ that you found in part (a). Which would you prefer to use as an estimator of θ – and why?

Solution. (a) Note that $\mathbb{E}(X_i) = \theta/2$. Therefore, $\mathbb{E}(X_1 + X_2 + X_3) = 3\theta/2 \neq \theta$. Therefore $\hat{\theta}$ biased, but $\tilde{\theta} = 2\theta/3$ is unbiased.

(b) Note that $\operatorname{var} X_i = \theta^2/12$. Since X_i's are independent, we can add their variance to find $\operatorname{var}(\bar{\theta}) = \operatorname{var}(X_1) + \operatorname{var}(X_2) = \theta^2/6$. Similarly, we find $\operatorname{var}(\tilde{\theta}) = (4/9) \operatorname{var} X_1 + (4/9) \operatorname{var} X_2 + (4/9) \operatorname{var} X_3 = \theta^2/9$. I'd prefer $\tilde{\theta}$ over $\bar{\theta}$ because it has uniformly lower variance.

TEL 502E – Detection and Estimation Theory

Midterm Examination

25.03.2014

(25 pts) 1. Suppose X_1, X_2 are independent identically distributed random variables with a probability density function (pdf) given as,

$$f(t) = \frac{1}{2}\, e^{-|t|}$$

We make two observations y_1, y_2, related to realizations of X_i's. Suppose there are two hypotheses concerning the observations

$$H_0 : y_i = x_i, \text{ for } i = 1, 2,$$
$$H_1 : y_i = 2\, x_i, \text{ for } i = 1, 2,$$

where x_i's are realizations of X_i's.

(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

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

(c) For the test in part (b), find the threshold γ so that the probability of a Type-I error is α. (Recall that we make a Type-I error if we decide H_1 while H_0 is true.)

(25 pts) 2. Suppose we observe a signal x_i for $i = 1, 2, \ldots, k$. There are two hypothesis of the form,

$$H_0 : x_i = z_i + n_i,$$
$$H_1 : x_i = s_i + n_i,$$

where z_i and s_i are deterministic (and known signals) and n_i's are realizations of iid Gaussian random variables with unit variance.

(a) Find a test statistic $T(x)$ such that the Neyman-Pearson test is of the form

$$\begin{cases} \text{If } T(x) < \gamma, & \text{then decide } H_0 \\ \text{If } T(x) \geq \gamma, & \text{then decide } H_1. \end{cases}$$

(b) Recall that the Q function is defined as,

$$Q(t) = \int_t^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-t^2/2}\, dt.$$

For the test in part (a), determine the constant γ, in terms of the $Q(\cdot)$ function so that the probability of a Type-I error is equal to α.

(c) Determine the probability of a Type-II error for the threshold computed in part(b). (Recall that we make a Type-II error if we decide H_0 while H_1 is true.)

(25 pts) 3. Let X_1, X_2 be independent Gaussian random variables with the same mean θ but with different variances σ_1^2, σ_2^2. That is, the probability density functions of X_1 and X_2 are,

$$f_1(t) = \frac{1}{\sqrt{2\pi\,\sigma_1^2}} \exp\left(-\frac{(t-\theta)^2}{2\sigma_1^2}\right),$$
$$f_2(t) = \frac{1}{\sqrt{2\pi\,\sigma_2^2}} \exp\left(-\frac{(t-\theta)^2}{2\sigma_2^2}\right).$$

(a) Find an expression for the Cramér-Rao lower bound for θ (in terms of σ_1^2 and σ_2^2).

(b) Find the uniformly minimum variance unbiased (UMVU) estimator for θ.

(25 pts) 4. Suppose X_1, X_2, X_3 are independent random variables uniformly distributed on $[0, \theta]$. Recall that a random variable uniformly distributed on $[0, \theta]$ has pdf

$$f(t) = \begin{cases} \theta^{-1}, & \text{if } t \in [0, 1], \\ 0, & \text{if } t \notin [0, 1]. \end{cases}$$

We are interested in the value of θ.

(a) A student suggests $\hat{\theta} = X_1 + X_2 + X_3$ as an estimator. Find out whether this estimator is biased or not. If it is biased, derive an unbiased estimator $\tilde{\theta}$ which is a function of $\hat{\theta}$ only (i.e., the other random variables do not appear in the new expression).

(b) An unbiased estimator for θ is $\bar{\theta} = X_1 + X_2$. Compare the variances of $\bar{\theta}$ and $\tilde{\theta}$ that you found in part (a). Which would you prefer to use as an estimator of θ – and why?

TEL 502E – Detection and Estimation Theory

Final Examination

23.05.2014

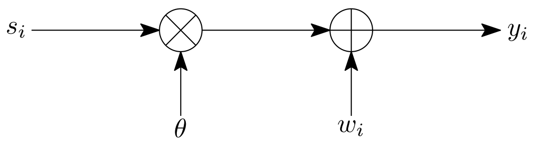
Student Name : ____________________

Student Num. : ____________________

5 Questions, 120 Minutes

Please Show Your Work for Full Credit!

(20 pts) 1. Consider a system as shown below, where a stochastic process s_i is modulated with a constant θ and then contaminated with additive noise w_i giving observations y_i, for $i = 1, 2, \ldots, n$.



Suppose that the samples s_i are independent Gaussian random variables with mean 1 and variance σ^2. Suppose also that w_i's are independent zero-mean Gaussian random variables with variance β^2.

θ can take one of two values $\{0, 1\}$. Let H_0 denote the hypothesis that $\theta = 0$, and H_1 denote the hypothesis that $\theta = 1$.

(a) Find the probability density function (pdf) of the observations y_i under H_1.

(b) Find the Neyman-Pearson test for testing H_0 against H_1.

(20 pts) 2. Consider a discrete random variable X whose probability mass function (pmf) depends on a parameter θ, where $\theta \in \{0, 1, 2\}$. Suppose that X takes values in $\{0, 1, 2, 3\}$ and its pmf for different values of θ, denoted by $P(x|\theta)$, is as given below.

x	$P(x\|\theta = 0)$	$P(x\|\theta = 1)$	$P(x\|\theta = 2)$
0	1/8	1/4	0
1	1/4	1/2	1/3
2	3/8	1/8	1/3
3	1/4	1/8	1/3

Please provide a brief explanation of your answers for full credit.

(a) Suppose we are given a realization of X as $x = 1$. Find the maximum likelihood estimate (MLE) for θ.

(b) Suppose we are given two independent realizations of X as $x_1 = 1$, $x_2 = 2$. Find the MLE for θ.

(20 pts) 3. Consider a disk with an unknown radius r. We are interested in the area of the disk. For this, we measure the radius n times but each measurement contains some error. Specifically, suppose that the measurements are of the form $X_i = r + Z_i$ for $i = 1, 2, \ldots, n$, where Z_i's are independent zero-mean Gaussian random variables with known variance σ^2.

(a) Find a sufficient statistic for r.

(b) A professor suggests that we use

$$\hat{A} = \pi \left(\frac{1}{n} \sum_{i=1}^{n} X_i^2 \right)$$

as an estimator of the area. Determine if $\hat{A}$ is biased or not.

(c) Find the uniformly minimum variance unbiased estimator for the area of the disk.

(20 pts) 4. Let X_1, X_2 be independent Gaussian random variables with mean θ and variance 1. Also, let θ be a random variable uniformly distributed on $[0, 1]$ – that is, the pdf of θ is given by,

$$f_\theta(t) = \begin{cases} 1, & \text{if } t \in [0, 1], \\ 0, & \text{if } t \notin [0, 1]. \end{cases}$$

(a) Find the joint pdf of θ, X_1, X_2. That is, find $f_{\theta, X_1, X_2}(t, x_1, x_2)$.

(b) Find the maximum a posteriori (MAP) estimate of θ.

(c) Evaluate the estimator you found in part (b) if the data is as given below.

(i) $x_1 = 3/4$, $x_2 = 1$.

(ii) $x_1 = 1/2$, $x_2 = 2$.

(20 pts) 5. Suppose a discrete-time system produces data of the form

$$X(n) = A_0 + A_1 n + Z_n,$$

where A_0, A_1 and Z_n are independent random variables. Assume that A_0 and A_1 are identically distributed with zero-mean and variance equal to σ^2. Also, suppose Z_n denotes a sequence of identically distributed zero-mean random variables with variance equal to β^2. Finally let X denote the samples at $n = 0, 1, 2$, which are made available to us, that is,
$X = \begin{bmatrix} X(0) & X(1) & X(2) \end{bmatrix}^T$.

(a) Find the covariance matrix of X, namely C.

(b) Find an expression for the linear minimum mean square estimate (LMMSE) of A_0, given the vector X. (Note : Do not try to invert C, just use C^{-1}.)

(c) Find an expression for the LMMSE of A_1, given the vector X.
