ECE 541
Stochastic Signals and Systems
Problem Set 8 Solutions

Problem Solutions : Yates and Goodman, 8.1.2 8.1.6 8.2.8 8.2.10 8.3.1 8.3.4 8.3.9 and 8.4.3

Problem 8.1.2 Solution

(a) We wish to develop a hypothesis test of the form

$$P[|K - E[K]| > c] = 0.05 \tag{1}$$

to determine if the coin we've been flipping is indeed a fair one. We would like to find the value of c, which will determine the upper and lower limits on how many heads we can get away from the expected number out of 100 flips and still accept our hypothesis. Under our fair coin hypothesis, the expected number of heads, and the standard deviation of the process are

$$E[K] = 50, \qquad \sigma_K = \sqrt{100 \cdot 1/2 \cdot 1/2} = 5. \tag{2}$$

Now in order to find c we make use of the central limit theorem and divide the above inequality through by σ_K to arrive at

$$P\left[\frac{|K - E[K]|}{\sigma_K} > \frac{c}{\sigma_K}\right] = 0.05 \tag{3}$$

Taking the complement, we get

$$P\left[-\frac{c}{\sigma_K} \le \frac{K - E[K]}{\sigma_K} \le \frac{c}{\sigma_K}\right] = 0.95 \tag{4}$$

Using the Central Limit Theorem we can write

$$\Phi\left(\frac{c}{\sigma_K}\right) - \Phi\left(\frac{-c}{\sigma_K}\right) = 2\Phi\left(\frac{c}{\sigma_K}\right) - 1 = 0.95 \tag{5}$$

This implies $\Phi(c/\sigma_K) = 0.975$ or $c/5 = 1.96$. That is, $c = 9.8$ flips. So we see that if we observe more then $50 + 10 = 60$ or less then $50 - 10 = 40$ heads, then with significance level $\alpha \approx 0.05$ we should reject the hypothesis that the coin is fair.

(b) Now we wish to develop a test of the form

$$P[K > c] = 0.01 \tag{6}$$

Thus we need to find the value of c that makes the above probability true. This value will tell us that if we observe more than c heads, then with significance level $\alpha = 0.01$, we should reject the hypothesis that the coin is fair. To find this value of c we look to evaluate the CDF

$$F_K(k) = \sum_{i=0}^{k} \binom{100}{i} (1/2)^{100}. \tag{7}$$

Computation reveals that $c \approx 62$ flips. So if we observe 62 or greater heads, then with a significance level of 0.01 we should reject the fair coin hypothesis. Another way to obtain this result is to use a Central Limit Theorem approximation. First, we express our rejection region in terms of a zero mean, unit variance random variable.

$$P[K > c] = 1 - P[K \leq c] = 1 - P\left[\frac{K - E[K]}{\sigma_K} \leq \frac{c - E[K]}{\sigma_K}\right] = 0.01 \tag{8}$$

Since $E[K] = 50$ and $\sigma_K = 5$, the CLT approximation is

$$P[K > c] \approx 1 - \Phi\left(\frac{c - 50}{5}\right) = 0.01 \tag{9}$$

From Table 3.1, we have $(c - 50)/5 = 2.35$ or $c = 61.75$. Once again, we see that we reject the hypothesis if we observe 62 or more heads.

Problem 8.1.6 Solution

Since the null hypothesis H_0 asserts that the two exams have the same mean and variance, we reject H_0 if the difference in sample means is large. That is, $R = \{|D| \geq d_0\}$.

Under H_0, the two sample means satisfy

$$E[M_A] = E[M_B] = \mu, \qquad \text{Var}[M_A] = \text{Var}[M_B] = \frac{\sigma^2}{n} = \frac{100}{n} \tag{1}$$

Since n is large, it is reasonable to use the Central Limit Theorem to approximate M_A and M_B as Gaussian random variables. Since M_A and M_B are independent, D is also Gaussian with

$$E[D] = E[M_A] - E[M_B] = 0 \qquad \text{Var}[D] = \text{Var}[M_A] + \text{Var}[M_B] = \frac{200}{n}. \tag{2}$$

Under the Gaussian assumption, we can calculate the significance level of the test as

$$\alpha = P[|D| \geq d_0] = 2\left(1 - \Phi(d_0/\sigma_D)\right). \tag{3}$$

For $\alpha = 0.05$, $\Phi(d_0/\sigma_D) = 0.975$, or $d_0 = 1.96\sigma_D = 1.96\sqrt{200/n}$. If $n = 100$ students take each exam, then $d_0 = 2.77$ and we reject the null hypothesis that the exams are the same if the sample means differ by more than 2.77 points.

Problem 8.2.8 Solution

Given hypothesis H_0 that $X = 0$, $Y = W$ is an exponential ($\lambda = 1$) random variable. Given hypothesis H_1 that $X = 1$, $Y = V + W$ is an Erlang ($n = 2, \lambda = 1$) random variable. That is,

$$f_{Y|H_0}(y) = \begin{cases} e^{-y} & y \geq 0, \\ 0 & \text{otherwise}, \end{cases} \quad f_{Y|H_1}(y) = \begin{cases} ye^{-y} & y \geq 0, \\ 0 & \text{otherwise}. \end{cases} \tag{1}$$

The probability of a decoding error is minimized by the MAP rule. Since $P[H_0] = P[H_1] = 1/2$, the MAP rule is

$$y \in A_0 \text{ if } \frac{f_{Y|H_0}(y)}{f_{Y|H_1}(y)} = \frac{e^{-y}}{ye^{-y}} \geq \frac{P[H_1]}{P[H_0]} = 1; \qquad y \in A_1 \text{ otherwise.} \tag{2}$$

Thus the MAP rule simplifies to

$$y \in A_0 \text{ if } y \leq 1; \qquad y \in A_1 \text{ otherwise.} \tag{3}$$

The probability of error is

$$P_{\text{ERR}} = P[Y > 1|H_0] P[H_0] + P[Y \leq 1|H_1] P[H_1] \tag{4}$$

$$= \frac{1}{2}\int_1^\infty e^{-y}\, dy + \frac{1}{2}\int_0^1 ye^{-y}\, dy \tag{5}$$

$$= \frac{e^{-1}}{2} + \frac{1-2e^{-1}}{2} = \frac{1-e^{-1}}{2}. \tag{6}$$

Problem 8.2.10 Solution

The key to this problem is to observe that

$$P[A_0|H_0] = 1 - P[A_1|H_0], \qquad P[A_1|H_1] = 1 - P[A_0|H_1]. \tag{1}$$

The total expected cost can be written as

$$E[C'] = P[A_1|H_0] P[H_0] C'_{10} + (1 - P[A_1|H_0]) P[H_0] C'_{00} \tag{2}$$

$$+ P[A_0|H_1] P[H_1] C'_{01} + (1 - P[A_0|H_1]) P[H_1] C'_{11}. \tag{3}$$

Rearranging terms, we have

$$E[C'] = P[A_1|H_0] P[H_0] (C'_{10} - C'_{00}) + P[A_0|H_1] P[H_1] (C'_{01} - C'_{11}) + P[H_0] C'_{00} + P[H_1] C'_{11}. \tag{4}$$

Since $P[H_0]C'_{00} + P[H_1]C'_{11}$ does not depend on the acceptance sets A_0 and A_1, the decision rule that minimizes $E[C']$ is the same decision rule that minimizes

$$E[C''] = P[A_1|H_0] P[H_0] (C'_{10} - C'_{00}) + P[A_0|H_1] P[H_1] (C'_{01} - C'_{11}). \tag{5}$$

The decision rule that minimizes $E[C'']$ is the same as the minimum cost test in Theorem 8.3 with the costs C_{01} and C_{10} replaced by the differential costs $C'_{01} - C'_{11}$ and $C'_{10} - C'_{00}$.

Problem 8.3.1 Solution

Since the three hypotheses H_0, H_1, and H_2 are equally likely, the MAP and ML hypothesis tests are the same. From Theorem 8.8, the MAP rule is

$$x \in A_m \text{ if } f_{X|H_m}(x) \geq f_{X|H_j}(x) \text{ for all } j. \tag{1}$$

Since N is Gaussian with zero mean and variance σ_N^2, the conditional PDF of X given H_i is

$$f_{X|H_i}(x) = \frac{1}{\sqrt{2\pi\sigma_N^2}} e^{-(x-a(i-1))^2/2\sigma_N^2}. \tag{2}$$

Thus, the MAP rule is

$$x \in A_m \text{ if } (x-a(m-1))^2 \leq (x-a(j-1))^2 \text{ for all } j. \tag{3}$$

This implies that the rule for membership in A_0 is

$$x \in A_0 \text{ if } (x+a)^2 \leq x^2 \text{ and } (x+a)^2 \leq (x-a)^2. \tag{4}$$

This rule simplifies to

$$x \in A_0 \text{ if } x \leq -a/2. \tag{5}$$

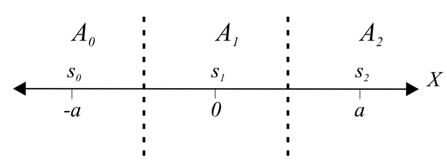
Similar rules can be developed for A_1 and A_2. These are:

$$x \in A_1 \text{ if } -a/2 \leq x \leq a/2 \tag{6}$$
$$x \in A_2 \text{ if } x \geq a/2 \tag{7}$$

To summarize, the three acceptance regions are

$$A_0 = \{x|x \leq -a/2\} \qquad A_1 = \{x| -a/2 < x \leq a/2\} \qquad A_2 = \{x|x > a/2\} \tag{8}$$

Graphically, the signal space is one dimensional and the acceptance regions are



Just as in the QPSK system of Example 8.13, the additive Gaussian noise dictates that the acceptance region A_i is the set of observations x that are closer to $s_i = (i-1)a$ than any other s_j.

Problem 8.3.4 Solution

Let H_i denote the hypothesis that symbol a_i was transmitted. Since the four hypotheses are equally likely, the ML tests will maximize the probability of a correct decision. Given H_i, N_1 and N_2 are independent and thus X_1 and X_2 are independent. This implies

$$f_{X_1,X_2|H_i}(x_1,x_2) = f_{X_1|H_i}(x_1)\, f_{X_2|H_i}(x_2) \tag{1}$$
$$= \frac{1}{2\pi\sigma^2} e^{-(x_1-s_{i1})^2/2\sigma^2} e^{-(x_2-s_{i2})^2/2\sigma^2} \tag{2}$$
$$= \frac{1}{2\pi\sigma^2} e^{-[(x_1-s_{i1})^2+(x_2-s_{i2})^2]/2\sigma^2} \tag{3}$$

From Definition 8.2 the acceptance regions A_i for the ML multiple hypothesis test must satisfy

$$(x_1, x_2) \in A_i \text{ if } f_{X_1,X_2|H_i}(x_1, x_2) \geq f_{X_1,X_2|H_j}(x_1, x_2) \text{ for all } j. \tag{4}$$

Equivalently, the ML acceptance regions are

$$(x_1, x_2) \in A_i \text{ if } (x_1 - s_{i1})^2 + (x_2 - s_{i2})^2 \leq (x_1 - s_{j1})^2 + (x_2 - s_{j2})^2 \text{ for all } j \tag{5}$$

In terms of the vectors $\mathbf{x}$ and $\mathbf{s}_i$, the acceptance regions are defined by the rule

$$\mathbf{x} \in A_i \text{ if } \|\mathbf{x} - \mathbf{s}_i\|^2 \leq \|\mathbf{x} - \mathbf{s}_j\|^2 \tag{6}$$

Just as in the case of QPSK, the acceptance region A_i is the set of vectors $\mathbf{x}$ that are closest to $\mathbf{s}_i$.

Problem 8.3.9 Solution

(a) First we note that

$$\mathbf{P}^{1/2}\mathbf{X} = \begin{bmatrix} \sqrt{p_1} & & \\ & \ddots & \\ & & \sqrt{p_k} \end{bmatrix} \begin{bmatrix} X_1 \\ \vdots \\ X_k \end{bmatrix} = \begin{bmatrix} \sqrt{p_1}X_1 \\ \vdots \\ \sqrt{p_k}X_k \end{bmatrix}. \tag{1}$$

Since each $\mathbf{S}_i$ is a column vector,

$$\mathbf{SP}^{1/2}\mathbf{X} = \begin{bmatrix} \mathbf{S}_1 & \cdots & \mathbf{S}_k \end{bmatrix} \begin{bmatrix} \sqrt{p_1}X_1 \\ \vdots \\ \sqrt{p_k}X_k \end{bmatrix} = \sqrt{p_1}X_1\mathbf{S}_1 + \cdots + \sqrt{p_k}X_k\mathbf{S}_k. \tag{2}$$

Thus $\mathbf{Y} = \mathbf{SP}^{1/2}\mathbf{X} + \mathbf{N} = \sum_{i=1}^{k} \sqrt{p_i}X_i\mathbf{S}_i + \mathbf{N}$.

(b) Given the observation $\mathbf{Y} = \mathbf{y}$, a detector must decide which vector $\mathbf{X} = \begin{bmatrix} X_1 & \cdots & X_k \end{bmatrix}'$ was (collectively) sent by the k transmitters. A hypothesis H_j must specify whether $X_i = 1$ or $X_i = -1$ for each i. That is, a hypothesis H_j corresponds to a vector $\mathbf{x}_j \in B_k$ which has ± 1 components. Since there are 2^k such vectors, there are 2^k hypotheses which we can enumerate as $H_1, \ldots, H_{2^k}$. Since each X_i is independently and equally likely to be ± 1, each hypothesis has probability 2^{-k}. In this case, the MAP and and ML rules are the same and achieve minimum probability of error. The MAP/ML rule is

$$\mathbf{y} \in A_m \text{ if } f_{\mathbf{Y}|H_m}(\mathbf{y}) \geq f_{\mathbf{Y}|H_j}(\mathbf{y}) \text{ for all } j. \tag{3}$$

Under hypothesis H_j, $\mathbf{Y} = \mathbf{SP}^{1/2}\mathbf{x}_j + \mathbf{N}$ is a Gaussian $(\mathbf{SP}^{1/2}\mathbf{x}_j, \sigma^2\mathbf{I})$ random vector. The conditional PDF of $\mathbf{Y}$ is

$$f_{\mathbf{Y}|H_j}(\mathbf{y}) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2}(\mathbf{y}-\mathbf{SP}^{1/2}\mathbf{x}_j)'(\sigma^2\mathbf{I})^{-1}(\mathbf{y}-\mathbf{SP}^{1/2}\mathbf{x}_j)} = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\|\mathbf{y}-\mathbf{SP}^{1/2}\mathbf{x}_j\|^2/2\sigma^2}. \tag{4}$$

The MAP rule is

$$\mathbf{y} \in A_m \text{ if } e^{-\left\|\mathbf{y}-\mathbf{SP}^{1/2}\mathbf{x}_m\right\|^2/2\sigma^2} \geq e^{-\left\|\mathbf{y}-\mathbf{SP}^{1/2}\mathbf{x}_j\right\|^2/2\sigma^2} \text{ for all } j, \tag{5}$$

or equivalently,

$$\mathbf{y} \in A_m \text{ if } \left\|\mathbf{y} - \mathbf{SP}^{1/2}\mathbf{x}_m\right\| \leq \left\|\mathbf{y} - \mathbf{SP}^{1/2}\mathbf{x}_j\right\| \text{ for all } j. \tag{6}$$

That is, we choose the vector $\mathbf{x}^* = \mathbf{x}_m$ that minimizes the distance $\left\|\mathbf{y} - \mathbf{SP}^{1/2}\mathbf{x}_j\right\|$ among all vectors $\mathbf{x}_j \in B_k$. Since this vector $\mathbf{x}^*$ is a function of the observation $\mathbf{y}$, this is described by the math notation

$$\mathbf{x}^*(\mathbf{y}) = \arg\min_{\mathbf{x}\in B_k} \left\|\mathbf{y} - \mathbf{SP}^{1/2}\mathbf{x}\right\|, \tag{7}$$

where $\arg\min_{\mathbf{x}} g(\mathbf{x})$ returns the argument $\mathbf{x}$ that minimizes $g(\mathbf{x})$.

(c) To implement this detector, we must evaluate $\left\|\mathbf{y} - \mathbf{SP}^{1/2}\mathbf{x}\right\|$ for each $\mathbf{x} \in B_k$. Since there 2^k vectors in B_k, we have to evaluate 2^k hypotheses. Because the number of hypotheses grows exponentially with the number of users k, the maximum likelihood detector is considered to be computationally intractable for a large number of users k.

Problem 8.4.3 Solution

With $v = 1.5$ and $d = 0.5$, it appeared in Example 8.14 that $T = 0.5$ was best among the values tested. However, it also seemed likely the error probability P_e would decrease for larger values of T. To test this possibility we use sqdistor with 100,000 transmitted bits by trying the following:

```
>> T=[0.4:0.1:1.0];Pe=sqdistor(1.5,0.5,100000,T);
>> [Pmin,Imin]=min(Pe);T(Imin)
ans =
   0.80000000000000
```

Thus among $\{0.4, 0.5, \cdots, 1.0\}$, it appears that $T = 0.8$ is best. Now we test values of T in the neighborhood of 0.8:

```
>> T=[0.70:0.02:0.9];Pe=sqdistor(1.5,0.5,100000,T);
>>[Pmin,Imin]=min(Pe);T(Imin)
ans =
   0.78000000000000
```

This suggests that $T = 0.78$ is best among these values. However, inspection of the vector `Pe` shows that all values are quite close. If we repeat this experiment a few times, we obtain:

```
>> T=[0.70:0.02:0.9];Pe=sqdistor(1.5,0.5,100000,T);
>> [Pmin,Imin]=min(Pe);T(Imin)
ans =
   0.78000000000000
>> T=[0.70:0.02:0.9];Pe=sqdistor(1.5,0.5,100000,T);
>> [Pmin,Imin]=min(Pe);T(Imin)
ans =
   0.80000000000000
>> T=[0.70:0.02:0.9];Pe=sqdistor(1.5,0.5,100000,T);
>> [Pmin,Imin]=min(Pe);T(Imin)
ans =
   0.76000000000000
>> T=[0.70:0.02:0.9];Pe=sqdistor(1.5,0.5,100000,T);
>> [Pmin,Imin]=min(Pe);T(Imin)
ans =
   0.78000000000000
```

This suggests that the best value of T is in the neighborhood of 0.78. If someone were paying you to find the best T, you would probably want to do more testing. The only useful lesson here is that when you try to optimize parameters using simulation results, you should repeat your experiments to get a sense of the variance of your results.