# ECE531 Homework Assignment Number 4

Solution

Make sure your reasoning and work are clear to receive full credit for each problem.

1. 4 points. Consider our standard coin flipping problem where you have an unknown coin, either fair (HT) or double headed (HH), and you observe the outcome of $n$ flips of this coin. Assume a uniform cost assignment. For notational consistency, let the state and hypothesis $x_0$ and $\mathcal{H}_0$ be the case when the coin is HT and $x_1$ and $\mathcal{H}_1$ be the case when the coin is HH. When $n = 2$, find the Neyman-Pearson decision rule and corresponding power $\beta$ for a false alarm probability $0 < \alpha < 1$. Repeat this for $n = 3$ and comment on any changes.

   **Solution:** When $n = 2$, we can form the conditional probability matrix as

$$P = \begin{bmatrix} 0.25 & 0 \\ 0.5 & 0 \\ 0.25 & 1 \end{bmatrix}$$

   and the likelihood ratio vector as

$$L = \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix}.$$

   - Case 1: $0 < \alpha < 0.25 \Rightarrow v = 4$, accordingly we have

$$\rho^{NP}(y) = \begin{cases} \gamma & \text{if } L(y) = 4 \\ 0 & \text{if } L(y) < 4 \end{cases}.$$

     where,

$$\gamma = \frac{\alpha - \sum_{\ell: L_\ell > v} P_{\ell,0}}{\sum_{\ell: L_\ell = v} P_{\ell,0}} = 4\alpha.$$

     The power of the test is then $\beta = P_D = 4\alpha$.

   - Case 2: $\alpha = 0.25 \Rightarrow v = 0$ and

$$\rho^{NP}(y) = \begin{cases} 1 & \text{if } L(y) > 0 \\ 0 & \text{if } L(y) \leq 0 \end{cases}.$$

     The power of the test is then $\beta = P_D = 1$.

   - Case 3: $0.25 < \alpha < 1 \Rightarrow v = 0$ Notice that, in this case, if the state is $x_1$, it is impossible for $y = 0$ and $y = 1$ to be observed. Thus increasing $\alpha$ can not increase the probability of detection. Hence the decision rule in Case 2 also applies here in Case 3.

When $n = 3$, the conditional probability matrix is

$$P = \begin{bmatrix} 0.125 & 0 \\ 0.375 & 0 \\ 0.375 & 0 \\ 0.125 & 1 \end{bmatrix}$$

and the likelihood ratio vector is

$$L = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 8 \end{bmatrix}.$$

- Case 1: $0 < \alpha < 0.125 \Rightarrow v = 8$, hence

  $$\rho^{NP}(y) = \begin{cases} \gamma & \text{if } L(y) = 8 \\ 0 & \text{if } L(y) < 8 \end{cases}.$$

  where

  $$\gamma = \frac{\alpha - \sum_{\ell : L_\ell > v} P_{\ell,0}}{\sum_{\ell : L_\ell = v} P_{\ell,0}} = 8\alpha.$$

  and the power of the test $\beta = P_D = 8\alpha$. Note that the additional observation effectively doubles the power of the test with respect to the $n = 2$ case when $0 < \alpha < 0.125$.

- Case 2: $\alpha = 0.125 \Rightarrow v = 0$

  $$\rho^{NP}(y) = \begin{cases} 1 & \text{if } L(y) > 0 \\ 0 & \text{if } L(y) \leq 0 \end{cases}.$$

  and the power of the test $\beta = P_D = 1$.

- Case 3: $0.125 < \alpha < 1 \Rightarrow v = 0$ This case is the same as when we had $n = 2$ and $0.25 < \alpha < 1 \Rightarrow v = 0$.

2. 4 points. Poor textbook Chapter II, Problem 2 (c).

   **Solution:** From the N-P lemma, the optimum decision rule must be of the form

   $$\rho^{NP}(y) = \begin{cases} 1 & \text{if } \frac{3}{2(y+1)} > v \\ \gamma & \text{if } \frac{3}{2(y+1)} = v \\ 0 & \text{if } \frac{3}{2(y+1)} < v \end{cases},$$

   Since $L(y)$ is monotone decreasing in $y$, the above test is equivalent to

   $$\rho^{NP}(y) = \begin{cases} 1 & \text{if } y < \tau \\ \gamma & \text{if } y = \tau \\ 0 & \text{if } y > \tau \end{cases},$$

   where $\tau = \frac{3}{2v} - 1$. Thus, the false-positive probability is:

   $$P_{\mathbf{fp}}(\rho^{NP}) = P_0(Y < \tau) = \int_0^\tau \frac{2}{3}(y+1)dy = \begin{cases} 0 & \text{if } \tau \leq 0 \\ \frac{2\tau}{3}\left(\frac{\tau}{2} + 1\right) & \text{if } 0 < \tau < 1 \\ 1 & \text{if } \tau \geq 1 \end{cases}.$$

The threshold for $P_{\mathbf{fp}}(\rho^{NP}) = \alpha$ is the solution to

$$\frac{2\tau}{3}\left(\frac{\tau}{2}+1\right) = \alpha,$$

which is $\tau = \sqrt{1+3\alpha}-1$. Hence, the $\alpha$-level Neyman-Pearson test is

$$\rho^{NP}(y) = \begin{cases} 1 & \text{if } y \leq \sqrt{1+3\alpha}-1 \\ 0 & \text{if } y > \sqrt{1+3\alpha}-1 \end{cases}$$

where randomization is not necessary because of the continuous observations space. The power of the test is

$$\beta = P_D(\rho^{NP}) = \int_0^\tau dy = \tau = \sqrt{1+3\alpha}-1, \quad 0 < \alpha < 1.$$

3. 4 points. Poor textbook Chapter II, Problem 6 (c).

**Solution:** Here we have $p_0(y) = p_N(y+s)$ and $p_1(y) = p_N(y-s)$, which gives

$$L(y) = \frac{1+(y+s)^2}{1+(y-s)^2}.$$

Figure 1 shows $L(y)$ versus $y$. Note that there is no monotonic relationship.

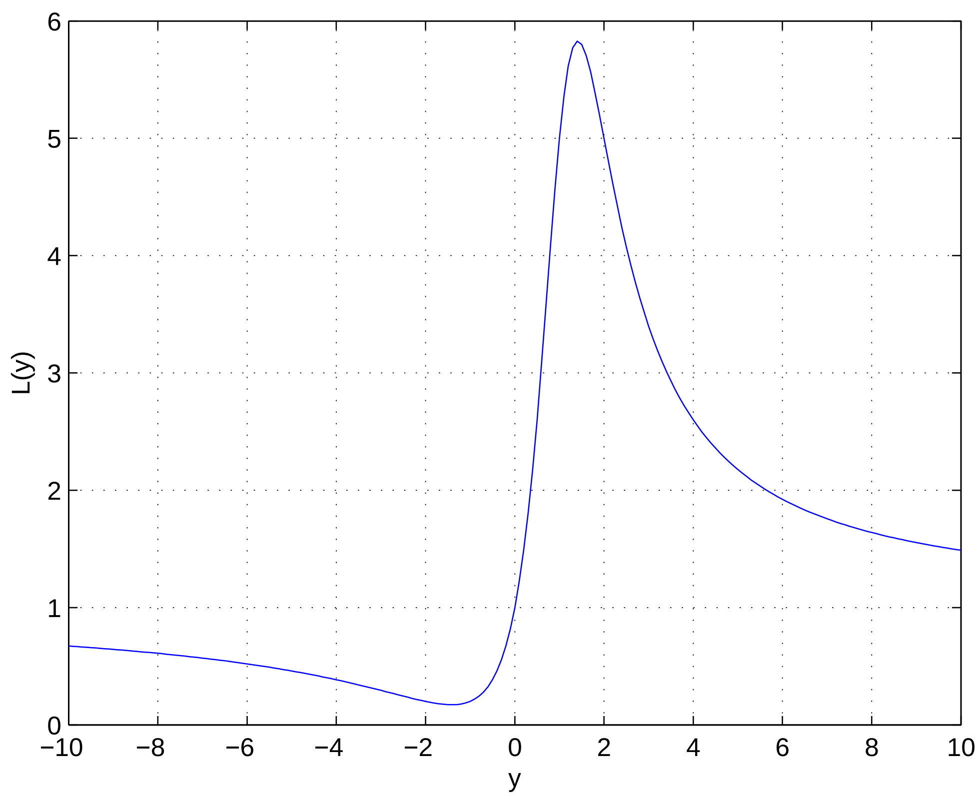


Figure 1: $L(y)$ versus $y$.

The critical region where we decide $\mathcal{H}_1$ is

$$\Gamma_1 = \{y \,|\, L(y) > v.\}\,.$$

It should be clear from Figure 1 that as we decrease $v$ from $\infty$ to 1, the region $\Gamma_1$ goes from $\emptyset$ to a single interval eventually filling $[0,\infty)$. As we decrease $v$ from 1 to 0, $\Gamma_1$ becomes two disjoint intervals $(-\infty,\gamma_1]$ and $[\gamma_2,\infty)$, where $\gamma_1 < \gamma_2$ are the solutions to $L(y) = v$ for $0 < v < 1$. We can formalize this with a precise description the following three cases:

- Case 1: $v > 1$,

$$\Gamma_1 = \left\{ y \,|\, \frac{s(v+1) - \sqrt{4s^2v - (v-1)^2}}{v-1} < y < \frac{s(v+1) + \sqrt{4s^2v - (v-1)^2}}{v-1} \right\}$$

- Case 2: $v = 1$,

$$\Gamma_1 = \{y \,|\, 0 < y < \infty\}$$

- Case 3: $0 < v < 1$,

$$\Gamma_1 = \left\{ y \mid y > \frac{s(v+1) - \sqrt{4s^2v - (v-1)^2}}{v-1} \text{ or } y < \frac{s(v+1) + \sqrt{4s^2v - (v-1)^2}}{v-1} \right\}$$

Given $\alpha$ and $s$, it is easy to determine which case you are in by computing

$$p = \int_0^\infty p_0(y)dy = \frac{\arctan(s)}{\pi} + \frac{1}{2}.$$

If $p > \alpha$, then $v > 1$. If $p = \alpha$, then $v = 1$. Otherwise, $0 < v < 1$. If you are in Case 1 or Case 3, you will need to numerically determine $v$ by integrating $p_0(y)$ over the appropriate critical region $\Gamma_1$ such that

$$\int_{y\in\Gamma_1} p_0(y)dy = \alpha.$$

With $v$, the detection probability can be found numerically by

$$\beta = \int_{y\in\Gamma_1} p_1(y)dy.$$

4. 4 points. Poor textbook Chapter II, Problem 19.

   **Solution:**

   (a) The likelihood ratio is given by

$$\begin{aligned} L(y) &= \frac{\prod_{k=1}^n \frac{1}{\sqrt{2\pi}\sigma_1} e^{-(y_k-\mu_1)^2/2\sigma_1^2}}{\prod_{k=1}^n \frac{1}{\sqrt{2\pi}\sigma_0} e^{-(y_k-\mu_0)^2/2\sigma_0^2}} \\ &= \left(\frac{\sigma_0}{\sigma_1}\right)^n \exp\left\{\frac{n}{2}\left(\frac{\mu_0^2}{\sigma_0^2} - \frac{\mu_1^2}{\sigma_1^2}\right)\right\} \exp\left\{\left(\frac{1}{2\sigma_0^2} - \frac{1}{2\sigma_1^2}\right)\sum_{k=1}^n y_k^2\right\} \exp\left\{\left(\frac{\mu_1}{\sigma_1^2} - \frac{\mu_0}{\sigma_0^2}\right)\sum_{k=1}^n y_k\right\} \end{aligned}$$

   which has the desired form.

(b) If $\mu_1 = \mu_0 \equiv \mu$ and $\sigma_1^2 > \sigma_0^2$, then we can simplify the comparison $L(y) > v$. Skipping the algebraic details, the Neyman-Pearson test in this case can be written as

$$\rho^{NP}(y) = \begin{cases} 1 & \sum_{k=1}^{n} (y_k - \mu)^2 > \eta \\ 0/1 & \sum_{k=1}^{n} (y_k - \mu)^2 = \eta \\ 0 & \sum_{k=1}^{n} (y_k - \mu)^2 < \eta \end{cases}$$

where $\eta$ is the appropriate threshold selected to satisfy the false positive probability constraint.

Alternatively, if $\mu_1 > \mu_0$ and $\sigma_1^2 = \sigma_0^2$, then the NP test can be written in the form

$$\rho^{NP}(y) = \begin{cases} 1 & \sum_{k=1}^{n} y_k > \eta' \\ 0/1 & \sum_{k=1}^{n} y_k = \eta' \\ 0 & \sum_{k=1}^{n} y_k < \eta' \end{cases}$$

where $\eta'$ is the appropriate threshold selected to satisfy the false positive probability constraint. Note that, in the first case, the test statistic is *quadratic* in the observations, and in the second case it is *linear*.

(c) For $n = 1$, $\mu_1 = \mu_0 \equiv \mu$, and $\sigma_1^2 > \sigma_0^2$, the NP test is of the form

$$\rho^{NP}(y) = \begin{cases} 1 & (y_1 - \mu)^2 > \eta \\ 0/1 & (y_1 - \mu)^2 = \eta \\ 0 & (y_1 - \mu)^2 < \eta \end{cases}$$

where $\eta > 0$ is an appropriate threshold. We have

$$\begin{aligned} P_{\mathbf{fp}}(\rho^{NP}) &= \text{Prob}[(Y_1 - \mu)^2 > \eta \,|\, x_0] \\ &= 2Q\left(\frac{\sqrt{\eta}}{\sigma_0}\right) \end{aligned}$$

where $Q(x)$ is the usual $Q$-function. Thus, for a test with significance-level $\alpha$ we have to solve $2Q\left(\frac{\sqrt{\eta}}{\sigma_0}\right) = \alpha$ which can be done in Matlab numerically or, even better, with the `erfcinv` function. The detection probability is

$$\begin{aligned} P_D(\rho^{NP}) &= \text{Prob}[(Y_1 - \mu)^2 > \eta \,|\, x_1] \\ &= 2Q\left(\frac{\sqrt{\eta}}{\sigma_1}\right) \end{aligned}$$

where $\eta$ is the solution to $2Q\left(\frac{\sqrt{\eta}}{\sigma_0}\right) = \alpha$.

5. 4 points. Poor textbook Chapter III, Problem 3. Also, try part (a) for the case when the noise is distributed as $\mathcal{N}(0, \Sigma)$ where $\Sigma \in \mathbb{R}^{n \times n}$ is the covariance matrix of the noise.

**Solution:**

(a) Since this is an $M$-ary hypothesis testing problem with equiprobable priors, and we wish to minimize error probability, we will have the critical regions:

$$\Gamma_k = \left\{ \underline{y} \in \mathbb{R}^n \,|\, p_k(\underline{y}) = \max_{m \in \{0,\ldots,M-1\}} p_m(\underline{y}) \right\}$$

Since $p_m(\underline{y})$ has the density $\mathcal{N}(\underline{s}_m, \sigma^2 \mathbf{I})$, this critical region can be reduced to

$$\begin{aligned}\Gamma_k &= \left\{ \underline{y} \in \mathbb{R}^n \,|\, \|\underline{y} - \underline{s}_k\|^2 = \min_{m \in \{0,\ldots,M-1\}} \|y - s_m\|^2 \right\} \\ &= \left\{ \underline{y} \in \mathbb{R}^n \,|\, \underline{s}_k^T \underline{y} = \max_{m \in \{0,\ldots,M-1\}} \underline{s}_m^T \underline{y} \right\}.\end{aligned}$$

Intuitively, the detector here is just computing the deterministic correlation between each signal vector $\underline{s}_m$ and the observation vector $\underline{y}$ and selecting the one that is largest. The minimum error probability decision rule is simply

$$\rho = \arg \max_{m \in \{0,\ldots,M-1\}} \underline{s}_m^T \underline{y}$$

If we have noise that is distributed as $\mathcal{N}(0, \Sigma)$, we can use the decorrelation trick discussed in Lecture 4. We factor $\Sigma = S^\top S$ and do a coordinate transformation on the observation and the signal vectors so that the noise in the transformed coordinate space is white. Then all of the above analysis applies directly.

(b) Under the assumption of $M$ equiprobable signal vectors, the error probability can be written as

$$P_e = \frac{1}{M} \sum_{k=0}^{M} \text{Prob}[\underline{Y} \in \Gamma_k^c \,|\, x_k]$$

where $\Gamma_k^c = \mathcal{Y} \backslash \Gamma_k$ and $x_k$ means that signal vector $\underline{s}_k$ was sent. We can write

$$\text{Prob}[\underline{Y} \in \Gamma_k^c \,|\, x_k] = 1 - \text{Prob}[\underline{Y} \in \Gamma_k \,|\, x_k] = 1 - \text{Prob}\left[\arg \max_{m \in \{0,\ldots,M-1\}} \underline{s}_m^T \underline{Y} = k \,|\, x_k\right].$$

Under the assumed orthogonality of the signal vectors $\underline{s}_1, \cdots, \underline{s}_n$, it is not difficult to show that, conditioned on $x_k$, the deterministic correlations $\underline{s}_1^T \underline{Y}, \underline{s}_2^T \underline{Y}, \cdots, \underline{s}_n^T \underline{Y}$ are independent Gaussian random variables with variances $\sigma^2 \|s_1\|^2$, and with means $\mu_m = 0$ for $m \neq k$ and mean $\mu_m = \|s_1\|^2$ for $m = k$. Thus

$$\text{Prob}\left[\arg \max_{m \in \{0,\ldots,M-1\}} \underline{s}_m^T \underline{Y} = k \,|\, x_k\right] =$$

$$\frac{1}{\sqrt{2\pi}\sigma \|s_1\|} \int_{-\infty}^{\infty} \text{Prob}\left[\max_{m \in \{0,\ldots,k-1,k+1,\ldots,M\}} \underline{s}_m^T \underline{Y} < z \,|\, x_k\right] e^{-(z - \|s_1\|^2)/2\sigma^2 \|s_1\|^2} \, dz$$

Now

$$\begin{aligned}\text{Prob}\left[\max_{m \in \{0,\ldots,k-1,k+1,\ldots,M\}} \underline{s}_m^T \underline{Y} < z \,|\, x_k\right] &= \prod_{m \in \{0,\ldots,k-1,k+1,\ldots,M\}} \text{Prob}\left[\underline{s}_m^T \underline{Y} < z \,|\, x_k\right] \\ &= \left[\Phi\left(\frac{z}{\sigma \|s_1\|}\right)\right]^{M-1}\end{aligned}$$

where the first inequality is a consequence of the independence of the individual deterministic correlations and $\Phi(x)$ is the usual CDF of a zero-mean, unit variance Gaussian random variable.

Combining the above and setting $x = z/\sigma\|s_1\|$ yields

$$1 - \text{Prob}\left[\underline{Y} \in \Gamma_k \,|\, x_k\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} [\Phi(x)]^{M-1} \, e^{-(x-d)^2/2} dx$$

for $k = 0, \ldots, M-1$. Since these are all the same, the desired expression for $P_e$ immediately follows.