Lecture 3: Minimax & Neyman Pearson Hypothesis Testing

In this lecture we complete the theoretical portion of our discussion of Minimax hypothesis testing rule and begin with the Neyman-Pearson hypothesis testing approach.

1 Minimax Hypothesis Testing (Contd. from Lecture 2)

We have already looked at the Minimax Hypothesis testing approach where the function $V(\pi_0)$ is continuous at the point of maxima. This may not be the case in many problems of interest. We therefore consider the case where V has an interior maximum (i.e. maxima $\pi_L \in (0,1)$) but the derivative does not exist at this point. The situation is shown in Fig. 1. Since the derivative is not unique, there is a concern as to which decision rule δ_{π_L} should be chosen.

We begin by defining two decision rules as limits to the maxima from the right and left,

$$
\begin{aligned}
\delta_{\pi_L}^- &= \lim_{\pi_0 \uparrow \pi_L} \delta_{\pi_0} = \lim_{\pi_0 \uparrow \pi_L} \mathbb{1}_{\{L(y) > \tau(\pi_0)\}}, \\
\delta_{\pi_L}^+ &= \lim_{\pi_0 \downarrow \pi_L} \delta_{\pi_0},
\end{aligned}
\tag{1}
$$

where, δ_{π_0} is the indicator function corresponding to the likelihood function being greater than the threshold. The critical regions Γ_1^-, Γ_1^+ corresponding to the decision rules $\delta_{\pi_L}^-$, and $\delta_{\pi_L}^+$ respectively, are given by,

$$
\begin{aligned}
\Gamma_1^- &= \bigcap_{\pi_0 \uparrow \pi_L} \{L(y) \geq \tau(\pi_0)\} = \{y \in \Gamma | L(y) \geq \tau(\pi_L)\}, \\
\Gamma_1^+ &= \bigcup_{\pi_0 \downarrow \pi_L} \{L(y) > \tau(\pi_0)\} = \{y \in \Gamma | L(y) > \tau(\pi_L)\}.
\end{aligned}
\tag{2}
$$

Let us now consider a randomized decision rule $\tilde{\delta}_{\pi_L}$ that uses Γ_1^- with probability q (where $q \in [0,1]$) and Γ_1^+ with probability $1-q$, at the point $L(y) = \tau(\pi_L)$.

Since we know the optimum decision rule at all other points except at the point of discontinuity, at which we now choose between the limiting decision rules using randomization. This equivalent to throwing a biased coin with probability q and selection a decision rule based on the outcome. We have,

$$\tilde{\delta}_{\pi_L} = \begin{cases} \text{choose } H_1, & \text{if } y \in \Gamma_1^+, \\ \text{choose } H_0, & \text{if } y \in (\Gamma_1^-)^c, \\ \text{choose } H_1 \ w.p. \ q, & \text{if } y \text{ is on boundary of } \Gamma_1^-. \end{cases} \tag{3}$$

Since conditional risk depends on the boundary condition, we write it as,

$$R_j(\tilde{\delta}_{\pi_L}) = qR_j(\delta_{\pi_L}^-) + (1-q)R_j(\delta_{\pi_L}^+), \quad j = [0,1], \tag{4}$$

which is nothing but,

$$\begin{aligned} R_0(\tilde{\delta}_{\pi_L}) &= qR_0(\delta_{\pi_L}^-) + (1-q)R_0(\delta_{\pi_L}^+), \\ R_1(\tilde{\delta}_{\pi_L}) &= qR_1(\delta_{\pi_L}^-) + (1-q)R_1(\delta_{\pi_L}^+). \end{aligned} \tag{5}$$

So for the minimax rule at equality,

$$max\{R_0(\tilde{\delta}_{\pi_L}), R_1(\tilde{\delta}_{\pi_L})\} = R_0(\tilde{\delta}_{\pi_L}) = R_1(\tilde{\delta}_{\pi_L}). \tag{6}$$

Therefore, the conditional risk equations for equality condition are,

$$qR_0(\delta_{\pi_L}^-) + (1-q)R_0(\delta_{\pi_L}^+) = qR_1(\delta_{\pi_L}^-) + (1-q)R_1(\delta_{\pi_L}^+),$$

$$q\{R_0(\delta_{\pi_L}^-) - R_0(\delta_{\pi_L}^+)\} + R_0(\delta_{\pi_L}^+) = q\{R_1(\delta_{\pi_L}^-) - R_1(\delta_{\pi_L}^+)\} + R_1(\delta_{\pi_L}^+),$$

$$R_0(\delta_{\pi_L}^+) - R_1(\delta_{\pi_L}^+) = q\{R_0(\delta_{\pi_L}^+) - R_1(\delta_{\pi_L}^+) + R_1(\delta_{\pi_L}^-) - R_0(\delta_{\pi_L}^-)\}. \tag{7}$$

So,

$$q = \frac{R_0(\delta_{\pi_L}^+) - R_1(\delta_{\pi_L}^+)}{R_0(\delta_{\pi_L}^+) - R_1(\delta_{\pi_L}^+) + R_1(\delta_{\pi_L}^-) - R_0(\delta_{\pi_L}^-)}. \tag{8}$$

This gives us the probability with which we choose the alternate hypothesis at the boundary. We now represent the unconditional risk by V. Since V is concave, it must have left hand and right hand derivative at π_L, which is denoted by $V'(\pi_L^-)$ and $V'(\pi_L^+)$. Now,

$$\begin{aligned} V'(\pi_L^-) &= \frac{d}{d\pi_0} r(\pi_0, \delta_{\pi_L}^-), \\ &= \frac{d}{d\pi_0}\{\pi_0 R_0(\delta_{\pi_L}^-) + (1-\pi_0)R_1(\delta_{\pi_L}^-)\}, \\ &= R_0(\delta_{\pi_L}^-) - R_1(\delta_{\pi_L}^-). \end{aligned} \tag{9}$$

2

Similarly, $V'(\pi_L^+) = R_0(\delta_{\pi_L}^+) - R_1(\delta_{\pi_L}^+)$. Hence eqn. (8) can be written as

$$q = \frac{V'(\pi_L^+)}{V'(\pi_L^+) - V'(\pi_L^-)}. \tag{10}$$

We can analyze the importance of the above equation in the following manner. Figure 1 shows the case in which V is discontinuous at the point of maxima. The decision rule here is $\tilde{\delta}_{\pi_L}$. By varying the probability q from 0 to 1 different slopes of the line $r(\pi_0, \tilde{\delta}_{\pi_L})$ can be obtained. The particular value of q obtained from the equation above gives rise to the horizontal line.
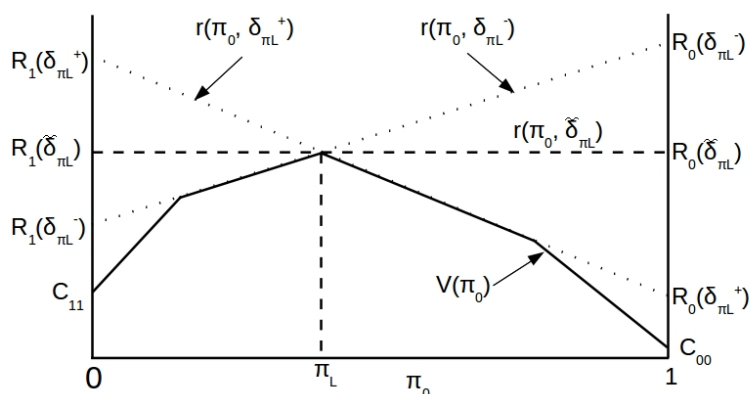

Figure 1: Action of the randomized decision rule

2   Neyman-Pearson Hypothesis Testing

Earlier we saw two different methods, Bayesian and Minimax, for hypothesis testing. It can be noticed that in both cases the predominant assumption is the *cost structure*, i.e. there is a penalty associated with detecting a particular Hypothesis. However, assigning a cost structure may not be possible always. Consider, the hypothesis test corresponding to the detection of enemy aircraft by a RADAR system. In this case, we would rather allow for some false detection (detecting an aircraft when none exists) at the cost of reduced probability of detection. Neyman Pearson criterion formalizes this approach of trading detection probability for false alarm probability. We begin by defining the required terminology.

Definition 2.1. For the hypothesis testing, we define the following terms,

- *Type-I Error or False Alarm:* occurs when the hypothesis H_1 is detected given H_0 is true.

3

- *Type-II Error or Missed Detection:* occurs when the hypothesis H_0 is detected given H_1 is true.

- *Detection:* correct acceptance of H_1.

Definition 2.2. The probability of Type-I Error is termed as the false alarm probability, denoted by $P_F(\delta)$.

Definition 2.3. The probability of Type-II Error is termed as the probability of miss or miss probability, denoted by $P_M(\delta)$. Hence, the probability of detection is $P_D(\delta) = 1 - P_M(\delta)$.

Definition 2.4. The randomized decision rule $\tilde{\delta} : \Gamma \to [0, 1]$ is a function for H_0 versus H_1 with the interpretation that for $y \in \Gamma$, $\tilde{\delta}$ is the conditional probability with which we accept H_1 given that we observe $Y = y$.

For a randomized rule $\tilde{\delta}$ the probability of false alarm is given by,

$$P_F(\tilde{\delta}) = E_0\{\tilde{\delta}(Y)\} = \int_\Gamma \tilde{\delta}(y)p_0(y)\mu(dy), \tag{11}$$

where, E_0 is the expectation under hypothesis H_0. Similarly, detection probability of a randomized detection rule $\tilde{\delta}$ is given by,

$$P_D(\tilde{\delta}) = E_1\{\tilde{\delta}(Y)\} = \int_\Gamma \tilde{\delta}(y)p_1(y)\mu(dy), \tag{12}$$

where, E_1 is the expectation under hypothesis H_1.

2.1 Neyman-Pearson Criterion

The Neyman-Pearson criterion formulates the hypothesis testing problem as a constrained optimization problem in which the probability of detection is maximized subject to an upper-bound on the false alarm probability. Mathematically, the criterion is given by,

$$\max_{\delta} \; P_D(\delta) \; \text{ subject to } P_F(\delta) \le \alpha, \tag{13}$$

where α is the sufficiency level of the test. We will now state and prove the Neyman-Pearson lemma.

Lemma 2.5. *The Neyman-Pearson Lemma*
   Let $\alpha > 0$,

i **Optimality:** *Let $\tilde{\delta}$ be any decision rule satisfying $P_F(\tilde{\delta}) \le \alpha$, and let $\tilde{\delta}'$ be any other decision rule of the form:*

$$\tilde{\delta}'(y) = \begin{cases} 1 & \text{if } p_1(y) > \eta p_0(y) \Rightarrow L(y) > \eta \\ \gamma(y) & \text{if } L(y) = \eta \\ 0 & \text{if } L(y) < \eta, \end{cases} \tag{14}$$

where, $\eta \ge 0$ and $0 \le \gamma(y) \le 1$ such that $P_F(\tilde{\delta}) = \alpha$, then, $P_D(\tilde{\delta}') \ge P_D(\delta')$.

ii **Existence:** *For every $\alpha \in (0,1)$ there is a decision rule $\tilde{\delta}_{NP}$ of the form (14) with $P(y) = \gamma_0$ (a constant), for which $P_F(\tilde{\delta}_{NP}) = \alpha$*

iii **Uniqueness:** *Let δ'', i.e. any α-level Neyman-Pearson decision rule for H_0 versus H_1. Then δ'' must be of the form (14) except possibly on a subset of Γ having zero probability under H_0 and H_1.*

Proof. (i). For any two decision rules of the form in eqn. (14), the following equation holds,

$$[\tilde{\delta}'(y) - \tilde{\delta}(y)][p_1(y) - \eta p_0(y)] \ge 0, \quad \forall y \in \Gamma. \tag{15}$$

We have,

$$\int_\Gamma [\tilde{\delta}'(y) - \tilde{\delta}(y)][p_1(y) - \eta p_0(y)]\mu(dy) \ge 0, \tag{16}$$

$$\int_\Gamma \tilde{\delta}'(y)p_1(y)\mu(dy) - \int_\Gamma \eta\tilde{\delta}'(y)p_0(y)\mu(dy)$$
$$- \int_\Gamma \tilde{\delta}(y)p_1(y)\mu(dy) + \int_\Gamma \eta\tilde{\delta}(y)p_0(y)\mu(dy) \ge 0,$$

$$\int_\Gamma \tilde{\delta}'(y)p_1(y)\mu(dy) - \int_\Gamma \tilde{\delta}(y)p_1(y)\mu(dy)$$
$$\ge \int_\Gamma \eta\tilde{\delta}'(y)p_0(y)\mu(dy) - \int_\Gamma \eta\tilde{\delta}(y)p_0(y)\mu(dy). \tag{17}$$

But we know that,

$$P_D(\tilde{\delta}) = \int_\Gamma \tilde{\delta}(y)p_1(y)\mu(dy) \text{ and } P_F(\tilde{\delta}) = \int_\Gamma \tilde{\delta}(y)p_0(y)\mu(dy). \tag{18}$$

Therefore, we can write the above equation as,

$$\begin{aligned} P_D(\tilde{\delta}') - P_D(\tilde{\delta}) &\ge \eta[P_F(\tilde{\delta}') - P_F(\tilde{\delta})], \\ &\ge \eta(\alpha - P_F(\tilde{\delta})), \quad \because P_F(\tilde{\delta}') = \alpha, \\ &\ge 0, \quad \because P_F(\tilde{\delta}) \le \alpha. \end{aligned} \tag{19}$$

Hence,
$$P_D(\tilde{\delta}') \geq P_D(\tilde{\delta}). \tag{20}$$

(ii). Let η_0 be the smallest number such that
$$\eta_0 = \inf\{\eta \in \mathbb{R}_+ : P_0(\{L(y) > \eta\}) \leq \alpha\}. \tag{21}$$

The term $P_0(\{L(y) > \eta\})$ is indicative of complimentary CDF like form which exhibits right continuity.

Now, if $P_0(\{L(y) > \eta_0\}) < \alpha$, then choose,
$$\begin{aligned}
\gamma_0 &= \frac{\alpha - P_0(\{p_1(y) > \eta_0 p_0(y)\})}{P_0(\{p_1(y) = \eta_0 p_0(y)\})}, \\
&= \frac{\alpha - P_0(\{L(y) > \eta_0\})}{P_0(\{L(y) = \eta_0\})}.
\end{aligned} \tag{22}$$

otherwise, choose γ_0 arbitrarily. Then, on defining $\tilde{\delta}_{NP}$ to be the decision rule of form (14), with $\eta = \eta_0$ and $\gamma(y) = \gamma_0$, we have,
$$\begin{aligned}
P_F(\tilde{\delta}_{NP}) &= E_0\{\tilde{\delta}_{NP}(Y)\}, \\
&= P_0(p_1(Y) > \eta_0 p_0(Y)) + \gamma_0 P_0(p_1(Y) = \eta_0 p_0(Y)), \\
&= \alpha.
\end{aligned} \tag{23}$$

Thus, we have chosen a decision rule of the form (14) with $\gamma(y)$ constant and false-alarm probability α.

(iii). Let $\tilde{\delta}'$ be of the form (14) and $\tilde{\delta}''$ be any other α-level Neyman-Pearson decision rule, then,
$$P_D(\tilde{\delta}'') = P_D(\tilde{\delta}'), \tag{24}$$

and therefore, by equation (19),
$$\begin{aligned}
0 &\geq \alpha - P_F(\tilde{\delta}''), \\
&\geq 0 \quad \text{(from part (i))}.
\end{aligned} \tag{25}$$

Hence, $P_F(\tilde{\delta}'') = \alpha$. Using the relations $P_F(\tilde{\delta}'') = \alpha$, $P_F(\tilde{\delta}') = \alpha$, and $P_D(\tilde{\delta}') = P_D(\tilde{\delta}'')$, we get
$$\{P_D(\tilde{\delta}') - P_D(\tilde{\delta}'')\} - \eta\{P_F(\tilde{\delta}') - P_F(\tilde{\delta}'')\} = 0,$$
$$\int_\Gamma [\tilde{\delta}'(y) - \tilde{\delta}''(y)][p_1(y) - \eta p_0(y)]\mu(dy) = 0. \tag{26}$$

Since the integrand is non-negative, it is zero except possibly on a set of zero probability under H_0 and H_1. Thus $\tilde{\delta}'$ and $\tilde{\delta}''$ differ only on the set $y \in \Gamma | L(y) = \eta$ and hence both have the same form (14) possibly differing only in choice of $\gamma(y)$.

$\square$