# Chapter 2

# Basic concepts of binary hypothesis testing

Most signal detection problems can be cast in the framework of $M$-ary hypothesis testing, where from some observations (data) we wish to decide among $M$ possible situations. For example, in a communication system, the receiver observes an electric waveform that consists of one of the $M$ possible signals corrupted by channel or receiver noise, and we wish to decide which of the $M$ possible signals is present during the observation. Obviously, for any given decision problem, there are a number of possible decision strategies or rules that can be applied, however we would like to choose a decision rule that is optimal in some sense. There are several classical useful criteria of optimality for such problems. The main object of this chapter is to give all the necessary basic definitions to define these criteria and their practical signification. Before going to the general case of $M$-ary hypothesis testing problem, let us start by a particular problem of binary (M=2) hypothesis testing which allows us to introduce the main basis more easily.

## 2.1 Binary hypothesis testing

The primary problem that we consider as an introduction is the simple hypothesis testing problem in which we assume that the observed data belong only on two possible processes with well known probability distributions $P_0$ and $P_1$:

$$\begin{cases} H_0 & : X \sim P_0 \\ H_1 & : X \sim P_1 \end{cases} \tag{2.1}$$

where "$X \sim P$" denotes "$X$ has distribution $P$" or "Data come from a stochastic process whose distribution is $P$". The hypotheses $H_0$ and $H_1$ are respectively referred to as *null* and *alternative* hypotheses. A decision rule $\delta$ for $H_0$ versus $H_1$ is any partition of the observation space $\Gamma$ into $\Gamma_1$ and $\Gamma_0 = \Gamma_1^c$ such that we choose $H_1$ when $\boldsymbol{x} \in \Gamma_1$ and $H_0$ when $\boldsymbol{x} \in \Gamma_0$. The sets $\Gamma_1$ and $\Gamma_0$ are respectively called the *rejection* and *acceptance* regions. So, we can think of the decision rule $\delta$ as a function on $\Gamma$ such that

$$\delta(\boldsymbol{x}) = \begin{cases} \delta_1 = 1 & \text{if } \boldsymbol{x} \in \Gamma_1 \\ \delta_0 = 0 & \text{if } \boldsymbol{x} \in \Gamma_0 = \Gamma_1^c \end{cases} \tag{2.2}$$

so that the value of $\delta$ for a given $\boldsymbol{x}$ is the index of the hypothesis accepted by the decision rule $\delta$.

We can also think of the decision rule $\delta(\boldsymbol{x})$ as a probability distribution $\{\delta_0, \delta_1\}$ on the space $\mathcal{D}$ of all the possible decisions where $\delta_j$ is the probability of deciding $H_j$ in the light of the data $\boldsymbol{x}$. In both cases $\delta_0 + \delta_1 = 1$.

We would like to choose $H_0$ or $H_1$ in some optimal way and, with this in mind, we may assign costs to our decisions. In particular we may assign costs $c_{ij} \geq 0$ to pay if we make the decision $H_i$ while the true decision to make was $H_j$. With the partition $\Gamma = \{\Gamma_0, \Gamma_1\}$ of the observation set, we can then define the conditional probabilities

$$P_{ij} = \Pr\{\boldsymbol{X} = \boldsymbol{x} \in \Gamma_i | H = H_j\} = P_j(\Gamma_i) = \int_{\Gamma_i} p_j(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \tag{2.3}$$

and then the average or expected *conditional risks* $R_j(\delta)$ for each hypothesis as

$$R_j(\delta) = \sum_{i=0}^{1} c_{ij} P_{ij} = c_{1j} P_j(\Gamma_1) + c_{0j} P_j(\Gamma_0), \quad j = 0, 1 \tag{2.4}$$

## 2.2 Bayesian binary hypothesis testing

Now assume that we can assign *prior* probabilities $\pi_0$ and $\pi_1 = 1 - \pi_0$ to the hypotheses $H_0$ and $H_1$, either to translate our preferences or to translate our prior knowledge about these hypotheses. Note that $\pi_j$ is the probability that $H_j$ is true unconditional (or independent) of the observation data $\boldsymbol{x}$ of $\boldsymbol{X}$. This is why they are called *prior* or *a priori* probabilities. For given priors $\{\pi_0, \pi_1\}$ we can define the *posterior* or *a posteriori* probabilities

$$\pi_j(\boldsymbol{x}) = \Pr\{H = H_j | \boldsymbol{X} = \boldsymbol{x}\} = \frac{p_j(\boldsymbol{x})\,\pi_j}{m(\boldsymbol{x})} \tag{2.5}$$

where

$$m(\boldsymbol{x}) = \sum_j P_j(\boldsymbol{x})\,\pi_j \tag{2.6}$$

is the overall density of $\boldsymbol{X}$.

We can also define an average or Bayes risk $r(\delta)$ as the overall average cost incurred by the decision rule $\delta$:

$$r(\delta) = \sum_j \pi_j\, R_j(\delta) \tag{2.7}$$

We may now use this quantity to define an optimum decision rule as the one that minimizes, over all decision rules, the Bayes risk. Such a decision rule is known as a Bayes decision rule.

To go a little further in details, let combine (2.4) and (2.7) to give

$$\begin{aligned} r(\delta) &= \sum_j \pi_j\, R_j(\delta) = \sum_j \pi_j \sum_i c_{ij} P_j(\Gamma_i) \\ &= \sum_j \pi_j c_{0j} P_j(\Gamma_0) + \pi_j c_{1j} P_j(\Gamma_1) \\ &= \sum_j \pi_j c_{0j} (1 - P_j(\Gamma_1)) + \pi_j c_{1j} P_j(\Gamma_1) \end{aligned}$$

$$\begin{aligned} &= \sum_j \pi_j c_{0j} + \sum_j \pi_j (c_{1j} - c_{0j}) P_j(\Gamma_1) \\ &= \sum_j \pi_j c_{0j} + \int_{\Gamma_1} \sum_j \pi_j \left(c_{1j} - c_{0j}\right) p_j(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \end{aligned} \tag{2.8}$$

Thus, we see that $r(\delta)$ is a minimum over all $\Gamma_1$ if we choose

$$\Gamma_1 = \left\{ \boldsymbol{x} \in \Gamma | \sum_j (c_{1j} - c_{0j}) \pi_j p_j(\boldsymbol{x}) \leq 0 \right\} \tag{2.9}$$

$$= \{ \boldsymbol{x} \in \Gamma | (c_{11} - c_{01}) \pi_1 p_1(\boldsymbol{x}) \leq (c_{00} - c_{10}) \pi_0 p_0(\boldsymbol{x}) \} \tag{2.10}$$

In general, the costs $c_{jj} < c_{ij}$ which means that the cost of correctly deciding $H_i$ is less than the cost of incorrectly deciding it. Then, (2.10) can be written

$$\Gamma_1 = \left\{ \boldsymbol{x} \in \Gamma | \frac{\pi_1 p_1(\boldsymbol{x})}{\pi_0 p_0(\boldsymbol{x})} \geq \tau_1 = \frac{c_{10} - c_{00}}{c_{01} - c_{11}} \right\} \tag{2.11}$$

$$= \left\{ \boldsymbol{x} \in \Gamma | \frac{p_1(\boldsymbol{x})}{p_0(\boldsymbol{x})} \geq \tau_2 = \frac{\pi_0}{\pi_1} \frac{c_{10} - c_{00}}{c_{01} - c_{11}} \right\} \tag{2.12}$$

$$= \{ \boldsymbol{x} \in \Gamma | c_{10} \pi_0(\boldsymbol{x}) + c_{11} \pi_1(\boldsymbol{x}) \leq c_{00} \pi_0(\boldsymbol{x}) + c_{01} \pi_1(\boldsymbol{x}) \} \tag{2.13}$$

This decision rule is known as a *likelihood-ratio* test or *posterior probability ratio* test due to the fact that $L(\boldsymbol{x}) = \frac{\pi_1(\boldsymbol{x})}{\pi_0(\boldsymbol{x})}$ is the ratio of the likelihoods and $\frac{\pi_1 p_1(\boldsymbol{x})}{\pi_0 p_0(\boldsymbol{x})}$ is the ratio of the posterior probabilities.

Note also that the quantity $c_{i0}\pi_0(\boldsymbol{x}) + c_{i1}\pi_1(\boldsymbol{x})$ is the average cost incurred by choosing the hypothesis $H_i$ given that $\boldsymbol{X} = \boldsymbol{x}$. This quantity is called the *posterior cost* of choosing $H_i$ given the observation $\boldsymbol{X} = \boldsymbol{x}$. Thus, the Bayes rule makes its decision by choosing the hypothesis that yields the minimum posterior cost.

This test plays a central role in the theory of hypothesis testing. It computes the likelihood ratio $L(\boldsymbol{x})$ for a given observed value $\boldsymbol{X} = \boldsymbol{x}$ and then makes its decision by comparing this ratio to a threshold $\tau_1$, *i.e.*;

$$\delta(\boldsymbol{x}) = \begin{cases} 1 & \text{if } L(\boldsymbol{x}) \geq \tau_1 \\ 0 & \text{if } L(\boldsymbol{x}) < \tau_1 \end{cases} \tag{2.14}$$

A commonly cost assignment is the *uniform costs* given by

$$c_{ij} = \begin{cases} 0 & \text{if } i = j \\ 1 & \text{if } i \neq j \end{cases} \tag{2.15}$$

The Bayes risk $\delta$ then becomes

$$r(\delta) = \pi_0 P_0(\Gamma_1) + \pi_1 P_1(\Gamma_0) = \pi_0 P_{01} + \pi_1 P_{10} \tag{2.16}$$

Noting that $P_i(\Gamma_j)$ is the probability of choosing $H_j$ when $H_i$ is true, $r(\delta)$ in this case becomes the average *probability of error* incurred by the rule $\delta$. This decision rule is then a minimum probability of error decision scheme.

Note also that with the uniform cost coefficients (2.15) the decision rule can be rewritten as

$$\delta(\boldsymbol{x}) = \begin{cases} 1 & \text{if } \pi_1(\boldsymbol{x}) \geq \pi_0(\boldsymbol{x}) \\ 0 & \text{if } \pi_1(\boldsymbol{x}) < \pi_0(\boldsymbol{x}) \end{cases} \tag{2.17}$$

This test is called the *maximum a posteriori (MAP)* decision scheme.

**Example : Detection of a constant signal in a Gaussian noise**
Let consider

$$\begin{cases} H_0 & X = \epsilon \\ H_1 & X = \mu + \epsilon \end{cases} \tag{2.18}$$

where $\epsilon$ is a Gaussian random variable with zero mean and a known variance $\sigma^2$ and where $\mu > 0$ is a known constant. In terms of distributions we can rewrite these hypotheses as

$$\begin{cases} H_0 & X \sim \mathcal{N}\left(0, \sigma^2\right) \\ H_1 & X = \mathcal{N}\left(\mu, \sigma^2\right) \end{cases} \tag{2.19}$$

where $\mathcal{N}\left(\mu, \sigma^2\right)$ means

$$\mathcal{N}\left(\mu, \sigma^2\right) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right] \tag{2.20}$$

It is then easy to calculate the likelihood ratio $L(x)$

$$L(x) = \frac{p_1(x)}{p_0(x)} = \exp\left[\frac{\mu}{\sigma^2}(x - \mu/2)\right] \tag{2.21}$$

Thus, the Bayes test for these hypotheses becomes

$$\delta(x) = \begin{cases} 1 & \text{if } L(x) \geq \tau \\ 0 & \text{if } L(x) < \tau \end{cases} \tag{2.22}$$

where $\tau$ is an appropriate threshold. We can remark that $L(x)$ is a striclty increasing function of $x$, so comparing $L(x)$ to a threshold $\tau$ is equivalent to comparing $x$ itself to another threshold $\tau' = L^{-1}(\tau) = \frac{\sigma^2}{\mu}\log(\tau) + \mu/2$:

$$\delta(x) = \begin{cases} 1 & \text{if } x \geq \tau' \\ 0 & \text{if } x < \tau' \end{cases} \tag{2.23}$$

where $L^{-1}$ is the inverse function of $L$.

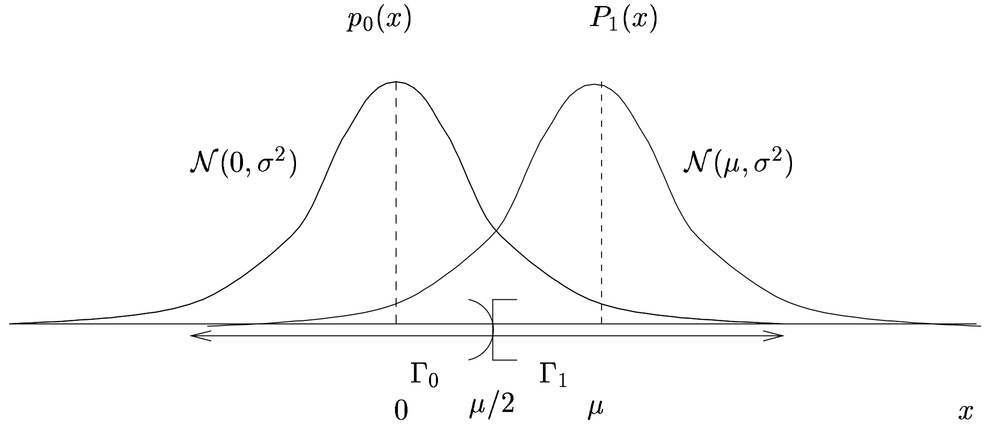


Figure 2.1: Location testing with Gaussian errors, uniform costs and equal priors.

In the special case of uniform costs and equal priors, we have $\tau = 1$ and so $\tau' = \mu/2$. Then, it is not difficult to show that the conditional probabilities are

$$P_{i1} = \Pr\{X = x \in \Gamma_1 | H = H_j\} = P_j(\Gamma_1) = \int_{\tau'}^{\infty} p_j(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \begin{cases} 1 - \Phi\left(\frac{\mu}{2\sigma}\right) & \text{for } j = 1 \\ 1 - \Phi\left(-\frac{\mu}{2\sigma}\right) & \text{for } j = 0 \end{cases} \tag{2.24}$$

and the minimum Bayes risk $r(\delta)$ is

$$r(\delta) = 1 - \Phi\left(\frac{\mu}{2\sigma}\right). \tag{2.25}$$

This is a decreasing function of $\frac{\mu}{\sigma}$, a quantity which is a simple version of the signal to noise ratio.

| Summary of notations for binary hypothesis testing | | |
|---|---|---|
| Hypotheses $H$ | $H_0$ | $H_1$ |
| Processes | $P_0$ | $P_1$ |
| Conditional densities or likelihood functions | $p_0(\boldsymbol{x})$ | $p_1(\boldsymbol{x})$ |
| Observation space $\Gamma$ partition | $\Gamma_0$ | $\Gamma_1$ |
| Decisions $\delta(\boldsymbol{x})$ | $\delta_0(\boldsymbol{x})$ | $\delta_1(\boldsymbol{x})$ |
| Conditional probabilities $P_{ij} = \int_{\Gamma_i} p_j(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ | $P_{00}, P_{01}$ | $P_{10}, P_{11}$ |
| Costs $c_{ij}$ | $c_{00}, c_{01}$ | $c_{10}, c_{11}$ |
| Conditional risks $R_j = \sum_i c_{ij} P_{ij}$ | $R_0 = c_{00}P_{00} + c_{10}P_{10}$ | $R_1 = c_{01}P_{01} + c_{11}P_{11}$ |
| Prior probabilities $\pi_j$ | $\pi_0$ | $\pi_1$ |
| Posterior probabilities $\pi_j(\boldsymbol{x}) = \frac{p_j(\boldsymbol{x})\pi_j}{m(\boldsymbol{x})}$ | $\pi_0(\boldsymbol{x})$ | $\pi_1(\boldsymbol{x})$ |
| Joint probabilities $Q_{ij} = \pi_j P_{ij}$ | $Q_{00}, Q_{01}$ | $Q_{10}, Q_{11}$ |
| Posterior costs $\bar{c}_i(\boldsymbol{x}) = \sum_j c_{ij}\pi_j(\boldsymbol{x})$ | $\bar{c}_0(\boldsymbol{x}) = c_{00}\pi_0(\boldsymbol{x}) + c_{01}\pi_1(\boldsymbol{x})$ | $\bar{c}_1(\boldsymbol{x}) = c_{10}\pi_0(\boldsymbol{x}) + c_{11}\pi_1(\boldsymbol{x})$ |
| Bayes risk $r(\delta)$ | $r(\delta) = \sum_j \pi_j R_j(\boldsymbol{x})$ | |
| Likelihoods ratio $L(\boldsymbol{x})$ | $L(\boldsymbol{x}) = \frac{p_1(\boldsymbol{x})}{p_0(\boldsymbol{x})}$ | |
| Posteriors ratio | $\frac{\pi_1(\boldsymbol{x})}{\pi_0(\boldsymbol{x})} = \frac{\pi_1 p_1(\boldsymbol{x})}{\pi_0 p_0(\boldsymbol{x})}$ | |
| Posterior costs ratio | $\frac{\bar{c}_1(\boldsymbol{x})}{\bar{c}_0(\boldsymbol{x})} = \frac{c_{10}\pi_0(\boldsymbol{x})+c_{11}\pi_1(\boldsymbol{x})}{c_{00}\pi_0(\boldsymbol{x})+c_{01}\pi_1(\boldsymbol{x})}$ | |

| The following equivalent tests minimize the Bayes risk $r(\delta)$ | |
|---|---|
| Likelihoods ratio test | $L(\boldsymbol{x}) = \frac{p_1(\boldsymbol{x})}{p_0(\boldsymbol{x})} > \tau_1 = \frac{\pi_0}{\pi_1}\frac{c_{10}-c_{00}}{c_{01}-c_{11}}$ |
| Posteriors ratio test | $\frac{\pi_1(\boldsymbol{x})}{\pi_0(\boldsymbol{x})} > \tau_2 = \frac{c_{10}-c_{00}}{c_{01}-c_{11}}$ |
| Posterior costs ratio test | $\frac{\bar{c}_1(\boldsymbol{x})}{\bar{c}_0(\boldsymbol{x})} > 1$ |

| Special case of uniform costs binary hypothesis testing | | |
|---|---|---|
| Costs $c_{ij} = \begin{cases} 0 & \text{if } i = j \\ 1 & \text{if } i \neq j \end{cases}$ | $c_{00} = 0, \quad c_{01} = 1$ | $c_{10} = 1, \quad c_{11} = 0$ |
| Conditional risks $R_j = \sum_i c_{ij} P_{ij}$ | $R_0 = P_{00}$ | $R_1 = P_{11}$ |
| Prior probabilities $\pi_j$ | $\pi_0$ | $\pi_1$ |
| Posterior probabilities $\pi_j(\boldsymbol{x}) = \frac{p_j(\boldsymbol{x})\pi_j}{m(\boldsymbol{x})}$ | $\pi_0(\boldsymbol{x})$ | $\pi_1(\boldsymbol{x})$ |
| Joint probabilities $Q_{ij} = \pi_j P_{ij}$ | $Q_{00}, \quad Q_{01}$ | $Q_{10}, \quad Q_{11}$ |
| Posterior costs $\bar{c}_j(\boldsymbol{x}) = \sum_j c_{ij} \pi_j(\boldsymbol{x})$ | $\bar{c}_0(\boldsymbol{x}) = \pi_1(\boldsymbol{x})$ | $\bar{c}_1(\boldsymbol{x}) = \pi_0(\boldsymbol{x})$ |
| Bayes risk $r(\delta)$ | $r(\delta) = \sum_j \pi_j R_j(\boldsymbol{x}) = \sum_j \pi_j P_{jj}$ | |
| Likelihoods ratio $L(\boldsymbol{x})$ | $L(\boldsymbol{x}) = \frac{p_1(\boldsymbol{x})}{p_0(\boldsymbol{x})}$ | |
| Posteriors ratio | $\frac{\pi_1(\boldsymbol{x})}{\pi_0(\boldsymbol{x})} = \frac{\pi_1 p_1(\boldsymbol{x})}{\pi_0 p_0(\boldsymbol{x})}$ | |
| Posterior costs ratio | $\frac{\bar{c}_1(\boldsymbol{x})}{\bar{c}_0(\boldsymbol{x})} = \frac{\pi_1(\boldsymbol{x})}{\pi_0(\boldsymbol{x})}$ | |

| The following equivalent tests minimize the Bayes risk $r(\delta)$ | |
|---|---|
| Likelihoods ratio test | $L(\boldsymbol{x}) = \frac{p_1(\boldsymbol{x})}{p_0(\boldsymbol{x})} > \tau_1 = \frac{\pi_0}{\pi_1}$ |
| Posteriors ratio test | $\frac{\pi_1(\boldsymbol{x})}{\pi_0(\boldsymbol{x})} > \tau_2 = 1$ |
| Posterior costs ratio test | $\frac{\bar{c}_1(\boldsymbol{x})}{\bar{c}_0(\boldsymbol{x})} = \frac{\pi_1(\boldsymbol{x})}{\pi_0(\boldsymbol{x})} > 1$ |

## 2.3 Minimax binary hypothesis testing

In the previous section, we saw how the Bayesian hypothesis testing gives us a complete procedure to the hypothesis testing problems. However, in some applications, we may not be able to assign the prior probabilities $\{\pi_0, \pi_1\}$. Then, one approache is to choose arbitrarily $\pi_0 = \pi_1 = 1/2$ and continue all the Bayesian procedure as in the last section. An alternative approach is to choose another design criterion than the expected penalty $r(\delta)$. For example, we may use the conditional risks $R_0(\delta)$ and $R_1(\delta)$ and design a decision rule that minimizes, over all $\delta$, the following criterion

$$\max\{R_0(\delta), R_1(\delta)\} \tag{2.26}$$

The decision rule based on this criterion is known as the *minimax rule.*

To design this decision rule, it is useful to consider the function $r(\pi_0, \delta)$, defined for a given prior $\pi_0 \in [0, 1]$ and a given decision rule $\delta$ as the average risk

$$r(\pi_0, \delta) = \pi_0 R_0(\delta) + (1 - \pi_0) R_1(\delta) \tag{2.27}$$

Noting that $r(\pi_0, \delta)$ is a linear function of $\pi_0$, then for fixed $\delta$, its maximum occurs at either $\pi_0 = 0$ or $\pi_0 = 1$ with the maximum value respectively either $R_1(\delta)$ or $R_0(\delta)$. So, the optimization problem of minimizing the criterion (2.26) over $\delta$ is equivalent to minimizing the quantity

$$\max_{\pi_0 \in [0,1]} r(\pi_0, \delta) \tag{2.28}$$

over $\delta$.

Now, for each prior $\pi_0$, let $\delta_{\pi_0}$ denote a Bayes rule corresponding to that prior and let $V(\pi_0) = r(\pi_0, \delta_{\pi_0})$; that is $V(\pi_0)$ is the minimum Bayes risk for the prior $\pi_0$. Then, it is not difficult to show that $V(\pi_0)$ is a concave function of $\pi_0$ with $V(0) = c_{11}$ and $V(1) = c_{00}$.

Now consider the function $r(\pi_0, \delta_{\pi_0'})$ which is a straight line tangent to $V(\pi_0)$ at $\pi_0 = \pi_0'$ and parallel to $r(\pi_0, \delta)$ (see figure 2.3).

From this figure, we can see that only Bayes rules can possibly be minimax rules. Indeed, we see that the minimax rule, in this case, is a Bayes rule for the prior value $\pi_0 = \pi_L$ that maximizes $V$, and for this prior $r(\pi_0, \delta_{\pi_L})$ is constant over $\pi_0$ and so $R_0(\delta_{\pi_L})) = R_1(\delta_{\pi_L}))$. This decision rule (with equal conditional risks) is called an *equalizer rule.* Because $\pi_L$ maximizes the minimum Bayes risk, it is called the *least-favorable prior.* Thus, in this case, a minimax decision rule is the Bayes rule for the least-favorable prior $\pi_L$.

Even if we arrived at this conclusion through an example, it can be prouved that this fact is true in all practical situations. This result is stated as the following proposition:

Suppose that $\pi_L$ is a solution to $V(\pi_L) = max_{\pi_0 \in [0,1]} V(\pi_0)$. Assume further that either $R_0(\delta_{\pi_L}) = R_1(\delta_{\pi_L})$ or $\pi_L = \{0, 1\}$. Then $\delta_{\pi_L}$ is a minimax rule. (see V. Poor for the proof). We will be back more in details on the minimax rule in chapter x.

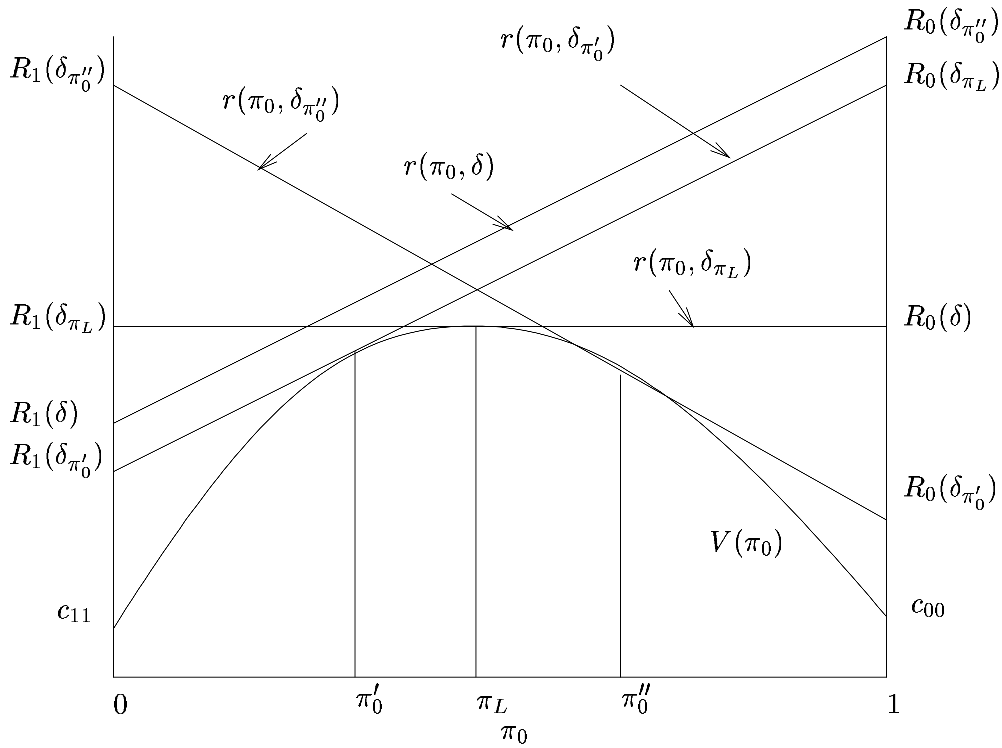


Figure 2.2: Illustration of minimax rule.

## 2.4 Neyman-Pearson hypothesis testing

In previous sections, we examined first the the Bayes hypothesis testing where the optimality was defined through the overall expected cost $r(\delta)$. Then, we considered the case where the prior probabilities $\{\pi_0, \pi_1\}$ are not available and described the minimax decision rule in terms of the maximum value of the conditional risks $R_0(\delta)$ and $R_1(\delta)$.

In both cases, we need to define the costs $c_{ij}$. In some applications, imposing a special cost structure on the decisions may not be available or not desirable. In such cases, an alternative criterion, known as the Neyman-Pearson criterion, is designed which is based on the probability of making a false decision. The main idea of this procedure is to choose one of the hypotheses as to be the main hypothesis and test other hypotheses against it. For example, in testing $H_0$ against $H_1$, two kinds of errors can be made:

- Falsely rejecting $H_0$ (or in this case falsely detecting $H_1$). This error is called either a *Type I error* or a *false alarm* or still a *false detection.*

- Falsely rejecting $H_1$ (or in this case falsely detecting $H_0$). This error is called either a *Type II error* or a *miss.*

The terms "false alarm" and "miss" come from radar applications in which $H_0$ and $H_1$ usually represent the absence and presence of a target.

For a decision rule $\delta$, the probability of a Type I error is known as *false alarm probability* and denoted by $P_F(\delta)$. Similarly, the probability of a Type II error is called the *miss* *probability* and denoted by $P_M(\delta)$. The quantity $P_D(\delta) = 1 - P_M(\delta)$ is called as the *detection probability* or still the *power* of $\delta$.

The Neyman-Pearson decision rule criterion is based on these quantities. It tries to place a bound on the *false alarm probability* and minimizes the *miss probability* within this constraint, *i.e.*;

$$\max P_D(\delta) \quad \text{subject to} \quad P_F(\delta) = 1 - P_D(\delta) \leq \alpha \tag{2.29}$$

where $\alpha$ is known as the *significance level* of the test. Thus the Neyman-Pearson decision rule criterion is to find the *most powerful* $\alpha$*-level* test of $H_0$ against $H_1$.

Note that, in the Neyman-Pearson test, as opposed to the Bayesian and minimax tests, the two hypotheses are not considered symetrically.

The general form of the Neyman-Pearson decision rule takes the forme

$$\delta(\boldsymbol{x}) = \begin{cases} 1 & \text{if } L(\boldsymbol{x}) > \tau \\ \gamma(\boldsymbol{x}) & \text{if } L(\boldsymbol{x}) = \tau \\ 0 & \text{if } L(\boldsymbol{x}) < \tau \end{cases} \tag{2.30}$$

where $\tau$ is a threshold.

The false alarm probability and the detection probability of a decision rule $\delta$ can be calculated respectively by

$$P_F(\delta) = \mathrm{E}_0\{\delta(\boldsymbol{x})\} = \int \delta(\boldsymbol{x}) p_0(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \tag{2.31}$$

$$P_D(\delta) = \mathrm{E}_1\{\delta(\boldsymbol{x})\} = \int \delta(\boldsymbol{x}) p_1(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \tag{2.32}$$

A parametric plot of $P_D(\delta)$ as a function of $P_F(\delta)$ is called the *receiver operation characterization* (ROCs).