Binary hypothesis testing

Decide between two hypotheses: $\mathcal{H}_0$ or $\mathcal{H}_1$.

To do so we find a *decision rule* which maps the observation $\mathbf{x}$ into either $\mathcal{H}_0$ or $\mathcal{H}_1$. Because the observation process is modeled probabilistically, the following errors may be made:

- $P(\mathcal{H}_0; \mathcal{H}_0) = \text{prob(decide } \mathcal{H}_0 \text{ when } \mathcal{H}_0 \text{ is true)} = \text{prob of correct non-detection}$
- $P(\mathcal{H}_0; \mathcal{H}_1) = \text{prob(decide } \mathcal{H}_0 \text{ when } \mathcal{H}_1 \text{ is true)} = \text{prob of missed detection} := P_M$
- $P(\mathcal{H}_1; \mathcal{H}_0) = \text{prob(decide } \mathcal{H}_1 \text{ when } \mathcal{H}_0 \text{ is true)} = \text{prob of false alarm} := P_{FA}$
- $P(\mathcal{H}_1; \mathcal{H}_1) = \text{prob(decide } \mathcal{H}_1 \text{ when } \mathcal{H}_1 \text{ is true)} = \text{prob of detection} := P_D$

More generally $P(\mathcal{H}_i; \mathcal{H}_j)$ is the probability of deciding $\mathcal{H}_i$ when hypothesis $\mathcal{H}_j$ is true.

Binary hypothesis testing

We want to design a "good" detection/decision rule, so need a criterion for "good". Two in this course:

1. Neyman-Pearson (NP): maximize P_D subject to a desired fixed P_{FA}.
2. Generalized Bayesian risk: minimize the Bayesian risk (cost function) for arbitrary costs C_{ij} for deciding $\mathcal{H}_i$ when $\mathcal{H}_j$ is true. Takes into account prior probabilities $P(\mathcal{H}_i)$. Reduces to the following for specific choices of costs C_{ij} and priors $P(\mathcal{H}_i)$:
   - Minimum probability of error ($\min P_E$) or maximum a posteriori (MAP): $C_{ii} = 0, C_{ij} = 1$ for $i \neq j$.
   - Maximum likelihood (ML): $C_{ij} = 0, C_{ij} = 1$ for $i \neq j$ AND all priors are equal, i.e. $P(\mathcal{H}_i) = P(\mathcal{H}_j), \forall i, j$.

Neyman-Pearson hypothesis testing

Neyman-Pearson Theorem 3.1 (pp.65)
To maximize P_D for a given $P_{FA} = \alpha$, decide $\mathcal{H}_1$ if

$$L(x) := \frac{p(\mathbf{x}; \mathcal{H}_1)}{p(\mathbf{x}; \mathcal{H}_0)} > \gamma, \tag{1}$$

where the threshold γ is found from

$$P_{FA} = \int_{\{\mathbf{x}: L(\mathbf{x}) > \gamma\}} p(\mathbf{x}; \mathcal{H}_0) d\mathbf{x} = \alpha \tag{2}$$

$L(\mathbf{x})$ is the *likelihood ratio*, and comparing $L(x)$ to a threshold is termed the *likelihood ratio test.*

Example 1

Consider the two hypotheses:

$$\begin{array}{ll} \mathcal{H}_0 : & x[0] \sim \mathcal{N}(0,1) \quad (\mu = 0) \\ \mathcal{H}_1 : & x[0] \sim \mathcal{N}(1,1) \quad (\mu = 1) \end{array}$$

Based on the single observation $x[0]$, decide which hypothesis it was generated from.

Useful problem 2.1

If $T \sim \mathcal{N}(\mu, \sigma^2)$, then

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

Example 2

We are given N observations $x[n], n = 0, 1, \cdots N-1$, which are i.i.d. and, depending on the hypothesis, are generated as

$$\begin{aligned} \mathcal{H}_0 :& \quad x[n] \sim \mathcal{N}(0, \sigma_0^2) \\ \mathcal{H}_1 :& \quad x[n] \sim \mathcal{N}(0, \sigma_1^2) \quad (\sigma_1^2 > \sigma_0^2) \end{aligned}$$

Determine the Neyman-Pearson hypothesis test.

Example 3

We are given N observations $x[n], n = 0, 1, \cdots N-1$, which, depending on the hypothesis, are generated as

$$\begin{array}{ll} \mathcal{H}_0 : & x[n] = w[n] \qquad\quad w[n] \sim \mathcal{N}(0, \sigma^2) \text{ i.i.d.} \\ \mathcal{H}_1 : & x[n] = A + w[n] \quad w[n] \sim \mathcal{N}(0, \sigma^2) \text{ i.i.d.} \end{array}$$

Determine the Neyman-Pearson hypothesis test.

Deflection coefficient

The deflection coefficient d is defined, for a test statistic T, as

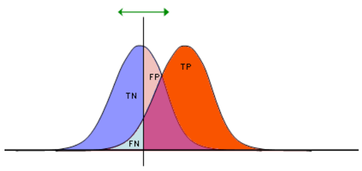
$$d^2 = \frac{(E(T;\mathcal{H}_1) - E(T;\mathcal{H}_0))^2}{\mathrm{var}(T;\mathcal{H}_0)},$$

and is useful in characterizing the performance of a detector. Usually, the larger the deflection coefficient, the easier it is to differentiate between the two signals, and thus the better the detection performance.

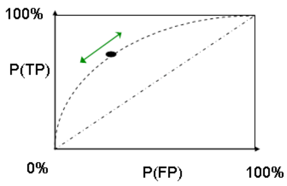
Receiver Operating Characteristics (ROC)

The Receiver Operating Characteristics (ROC) is a graph of P_D (y-axis) versus P_{FA} (x-axis), showing the tradeoff between the two. For $\gamma = +\infty$ you have $P_{FA} = P_D = 0$, while for $\gamma = -\infty$ you have $P_{FA} = P_D = 1$. For intermediate γ you lie on a curve above the $P_{FA} = P_D$ line.

[different terminology - can you map it to ours?]



TP	FP
FN	TN
1	1



Bayesian risk

Associate with each of the four detection possibilities a cost, i.e. C_{ij} is the cost of deciding hypothesis $\mathcal{H}_i$ when hypothesis $\mathcal{H}_j$ is true. In the binary hypothesis testing case, $i, j \in \{0, 1\}$. Let $P(\mathcal{H}_i|\mathcal{H}_j)$ be the probability of deciding $\mathcal{H}_i$ when $\mathcal{H}_j$ is true, and $P(\mathcal{H}_i)$ be the prior probability of hypothesis $\mathcal{H}_i$. Note the Bayesian formulation, assigning priors to the hypotheses is different than in the classical Neyman-Pearson criterion.

$$\text{Bayes risk} := \mathcal{R} = E[C] = \sum_{i=0}^{1} \sum_{j=0}^{1} C_{ij} P(\mathcal{H}_i|\mathcal{H}_j) P(\mathcal{H}_j) \qquad (1)$$

Under the assumption that $C_{10} > C_{00}$, $C_{01} > C_{11}$, the detector that minimizes the Bayes risk is to decide $\mathcal{H}_1$ if

$$\frac{p(\mathbf{x}|\mathcal{H}_1)}{p(\mathbf{x}|\mathcal{H}_0)} > \frac{(C_{10} - C_{00})P(\mathcal{H}_0)}{(C_{01} - C_{11})P(\mathcal{H}_1)} = \gamma$$

Bayesian risk

The Bayesian risk detection framework encompasses:

- Minimum probability of error ($\min P_E$) or maximum a posteriori (MAP) (same): $C_{ii} = 0, C_{ij} = 1$ for $i \neq j$. These detectors decide $\mathcal{H}_1$ if
  - $\min P_E : \frac{p(\mathbf{x}|\mathcal{H}_1)}{p(\mathbf{x}|\mathcal{H}_0)} > \frac{P(\mathcal{H}_0)}{P(\mathcal{H}_1)} = \gamma$
  - MAP: $P(\mathcal{H}_1|\mathbf{x}) > P(\mathcal{H}_0|\mathbf{x})$
- Maximum likelihood (ML): $C_{ij} = 0, C_{ij} = 1$ for $i \neq j$ AND all priors are equal, i.e. $P(\mathcal{H}_i) = P(\mathcal{H}_j), \forall i, j$. This detector decides $\mathcal{H}_1$ if
  - ML: $P(\mathbf{x}|\mathcal{H}_1) > P(\mathbf{x}|\mathcal{H}_0)$

Bayesian risk example

Find the detection rule that minimizes the probability of error for the following binary hypothesis testing problem:

$$\begin{array}{lll} \mathcal{H}_0 : & x[n] = w[n] & w[n] \sim \mathcal{N}(0, \sigma^2) \text{ i.i.d.} \\ \mathcal{H}_1 : & x[n] = A + w[n] & w[n] \sim \mathcal{N}(0, \sigma^2) \text{ i.i.d. } , A > 0 \end{array}$$

Also determine the probability of error achieved by this detector.

Multiple hypothesis testing

In binary hypothesis testing we detected one of 2 hypothesis. We now wish to detect one of $M > 2$ hypotheses. Neyman-Pearson is possible, see reference pp. 81, we only consider Bayes risk minimization, i.e. we wish to minimize

$$\mathcal{R} = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} C_{ij} P(\mathcal{H}_i|\mathcal{H}_j) P(\mathcal{H}_j)$$

The detector that minimizes the Bayes risk choses the hypothesis $\mathcal{H}_i$ for which

$$C_i(x) := \sum_{j=0}^{M-1} C_{ij} P(\mathcal{H}_i|\mathcal{H}_j) P(\mathcal{H}_j)$$

is minimal over all $i = 0, 1, \cdots, M-1$ (picks the one with minimal cost).

Multiple hypothesis testing

Once again, the Bayes risk is a generalization of MAP and ML detectors, which in the multiple hypothesis case reduce to:

- Minimum probability of error ($\min P_E$) or maximum a posteriori (MAP) (same): $C_{ii} = 0, C_{ij} = 1$ for $i \neq j$. These detectors decide $\mathcal{H}_i$ if $P(\mathcal{H}_i|\mathbf{x}) > P(\mathcal{H}_j|\mathbf{x})$, $\forall j$
- Maximum likelihood (ML): $C_{ij} = 0, C_{ij} = 1$ for $i \neq j$ AND all priors are equal, i.e. $P(\mathcal{H}_i) = P(\mathcal{H}_j), \forall i, j$. This detector decides $\mathcal{H}_1$ if $P(\mathbf{x}|\mathcal{H}_i) > P(\mathbf{x}|\mathcal{H}_j)$, $\forall j$

Multiple hypothesis testing example

Assume that we have three hypotheses

$$\begin{array}{lll}\mathcal{H}_0: & x[n] = -A + w[n] & n = 0, 1, \cdots .N-1 \\ \mathcal{H}_1: & x[n] = w[n] & n = 0, 1, \cdots .N-1 \\ \mathcal{H}_2: & x[n] = A + w[n] & n = 0, 1, \cdots .N-1\end{array}$$

where $A > 0$ and $w[n]$ is white Gaussian noise with variance σ^2. Assuming equal priors on the hypotheses, find the detector that minimizes the probability of error and find an expression for this probability of error.