ECE531 Screencast 8.1: Introduction to Detection

D. Richard Brown III

Worcester Polytechnic Institute

Detection = Hypothesis Testing

Detection is the process of deciding among two or more possible underlying statistical situations ("hypotheses") from noisy observations.

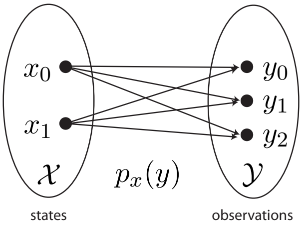
Examples of hypotheses:

- The coin is fair ($\mathcal{H}_0$) or not fair ($\mathcal{H}_1$).
- The approaching airplane is friendly ($\mathcal{H}_0$) or unfriendly ($\mathcal{H}_1$).
- This email is spam ($\mathcal{H}_1$) or not spam ($\mathcal{H}_0$).
- The medical treatment is effective ($\mathcal{H}_1$) or ineffective ($\mathcal{H}_0$).
- Communication receiver: Given a codebook with M codewords, which codeword was sent ($\{\mathcal{H}_0, \dots, \mathcal{H}_{M-1}\}$)?

More generally, we want to specify a decision rule that maps observations to decisions optimally in some sense.

States and Observations

- Let $x \in \mathcal{X} = \{x_0, \dots, x_{N-1}\}$ denote the **state**, a hidden variable about which we wish to make an inference.
- The available **observation** is modeled as a random variable Y taking on values in the set $\mathcal{Y}$.
- For each state $x \in \mathcal{X}$, we assume we have a **probabilistic description** of the random variable Y when the state is x. The notation $p_x(y) = p_{Y|X}(y|X = x)$ (random X) or $p_x(y) = p_Y(y; x)$ (nonrandom X) is the density of the random variable Y when the state is x.



Hypotheses and Decisions

- **Hypotheses** can be represented as a partition of $\mathcal{X}$, denoted by $\mathcal{H} = \{\mathcal{H}_0, \mathcal{H}_1, \dots, \mathcal{H}_{M-1}\}$ where

$$
\begin{aligned}
\mathcal{H}_i &\subseteq \mathcal{X} \\
\mathcal{H}_i &\neq \emptyset \\
\mathcal{H}_i \bigcap \mathcal{H}_j &= \emptyset \text{ for } i \neq j \text{ and} \\
\bigcup_i \mathcal{H}_i &= \mathcal{X}
\end{aligned}
$$

- The set of possible **decisions** is then $\mathcal{Z} = \{0, 1, \dots, M-1\}$ where decision i indicates the selection of hypothesis $\mathcal{H}_i$. In other words, decision i is the decision that $x \in \mathcal{H}_i$.
- If $\mathcal{X}$ is finite, then we must have $M \leq N$.

Types of Hypothesis Testing Problems

Recall $N = |\mathcal{X}|$ is the number of states (assume $\mathcal{X}$ is finite for now) and $M = |\mathcal{H}|$ is the number of hypotheses.

- If $M = 2$, then we have a **binary** hypothesis testing problem.
- If $M = N$, then we seek to decide the actual state. In this case we can take $\mathcal{H}_i = \{x_i\}$ and we have a **simple** hypothesis testing problem.
- If $M < N$ or $\mathcal{X}$ is infinite, then we have a **composite** hypothesis testing problem. At least one hypothesis contains more than one state.

Unlike a simple hypothesis with underlying distribution $p_x(y)$, a composite hypothesis does not completely specify the underlying distribution.

Our focus will be on simple hypothesis testing problems for now, but we will return to composite hypothesis testing in a few weeks.

Examples

We have a coin with $\mathrm{Prob}(\mathsf{H}) = q$ unknown.

1. Suppose q can only take on two values: q_0 or q_1. What kind of hypothesis testing problem is this? Binary, simple.
2. Suppose q can take on any value in the set $\{q_0, q_1, \dots, q_{M-1}\}$ and we wish to determine which value it is. What kind of hypothesis testing problem is this? M-ary, simple.
3. Suppose q can take on any value in the set $\{q_0, q_1, \dots, q_{N-1}\}$ but only wish to know if it is q_0 or not (e.g. $q_0 = 0.5$ "is the coin fair?"). What kind of hypothesis testing problem is this? Binary, composite $M = 2 < N$.
4. Suppose q can be any value in $[0, 1]$ and we want to determine this value. What kind of problem is this? Estimation.

Model Summary