ECE531 Lecture 3a: A Mathematical Model for Hypothesis Testing (Infinite Number of Possible Observations)

D. Richard Brown III

Worcester Polytechnic Institute

26-January-2011

Hypothesis Testing with Infinite Observation Spaces

Last week, we covered the case of observation sets with a **finite** number of possibilities. Lots of real-world problems have observation sets with an infinite number of possibilities. For example:

1. Communications: We transmit a binary symbol $s \in \{s_0, s_1\}$ and the signal is received in additive white Gaussian noise

$$y = s + w$$

   with $w \sim \mathcal{N}(0, \sigma^2)$. The observation $y \in \mathbb{R} = \mathcal{Y}$. The decision rule is a mapping from $\mathbb{R}$ to $\mathcal{Z} = \{0, 1\}$.

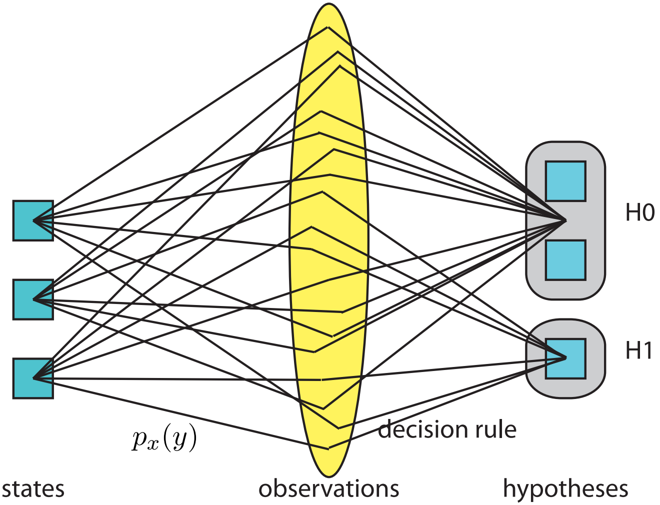
2. Drug testing: A test provides values for the level of testosterone, red blood cell count, and creatinine. The observation $y \in \mathbb{R}^3 = \mathcal{Y}$. The decision rule is a mapping from $\mathbb{R}^3$ to $\mathcal{Z} = \{0, 1\}$.

In the case of finite observation spaces, we developed the concept of **conditional risk vectors** $R(D) = [R_0(D), \dots, R_{N-1}(D)]^\top$ with

$$R_j(D) = c_j^\top D p_j \text{ (finite observation spaces)}$$

We would like to extend this notion to **infinite observation spaces**.

Model Summary



Infinite Observation Sets: Part 1

We can generalize our insight from the finite observation space as follows:

1. We can no longer use a decision matrix. Our randomized decision rule is denoted as $\rho = [\rho_0, \dots, \rho_{M-1}] : \mathcal{Y} \mapsto \mathcal{P}_M$ where $\mathcal{P}_M$ is the set of pmfs on $\mathcal{Z}$. We still use $\mathcal{D}$ to denote the set of decision rules $\rho \in \mathcal{D}$.
2. We denote $\rho_i(y)$ as the probability of deciding $\mathcal{H}_i$ when the observation is y.
3. The cost of deciding $\mathcal{H}_i$ when the state is x_j is still denoted as C_{ij}. Hence, when we start in state x_j and receive the observation y, the expected cost of using decision rule ρ is

$$C_j(\rho) = \sum_{i=0}^{M-1} \rho_i(y) C_{ij}$$

Infinite Observation Sets: Part 2

The conditional risk for state x_j is then

$$R_j(\rho) = \int_{y \in \mathcal{Y}} C_j(\rho) p_j(y)\, dy = \int_{y \in \mathcal{Y}} \left[\sum_{i=0}^{M-1} \rho_i(y) C_{ij} \right] p_j(y)\, dy$$

where $p_j(y)$ is the known conditional density that probabilistically describes the relationship between state x_j and the observations.

As before, we can group these individual conditional risks into a conditional risk vector $R(\rho) \in \mathbb{R}^N$.

Theorem

The function $R : \rho \mapsto R(\rho)$ is linear.

Proof.

Same idea as the case with finite $\mathcal{Y}$. □

Infinite Observation Sets: Part 3

If we let the decision rule ρ range over all of $\mathcal{D}$, $R(\rho)$ traces out the set $\mathcal{Q}$ of achievable conditional risk vectors in $\mathbb{R}^N$.

Theorem

$\mathcal{Q}$ is a closed and bounded convex subset of $\mathbb{R}^N$.

The proof of this is omitted here since it is a bit messy and requires some understanding of topology, pointwise convergence, and the dominated convergence theorem.

The main point: The concepts of **Pareto optimal decision rules** and the **optimal tradeoff surface** of $\mathcal{Q}$ also apply to the case of infinite $\mathcal{Y}$.

Note that $\mathcal{Q}$ is probably not a polytope anymore.

Summary of Main Results

Conditional risks as a way of quantifying the performance/consequences of a decision rule when the state is x_j:

$$R_j(D) = c_j^\top D p_j \text{ (finite observation spaces)}$$

$$R_j(\rho) = \int_{y \in \mathcal{Y}} \left[\sum_{i=0}^{M-1} \rho_i(y) C_{ij} \right] p_j(y)\, dy \text{ (infinite observation spaces)}$$

Remarks:

1. We can only use decision matrices in the case when $\mathcal{Y}$ is finite.
2. The conditional risks for finite and infinite $\mathcal{Y}$ are conceptually similar:
   - Both are an inner product of the cost-weighted decision rule and the conditional observation probabilities
   - Both yield a set of achievable CRVs that is closed, bounded, and convex
   - Convexity implies that minimizing all conditional risks simultaneously is impossible. The conditional risks must be traded off against each other on the optimal tradeoff surface.