ECE531 Screencast 9.4: Bayesian Detection with Infinite Possible Observations

D. Richard Brown III

Worcester Polytechnic Institute
Infinite Observation Spaces: Bayes Risk

Given the prior $\pi$ and the decision rule $\rho$, we have the Bayes Risk

$$r(\rho, \pi) := \sum_{j=0}^{N-1} \pi_j R_j(\rho)$$

where $\pi_j \geq 0$ and $\sum_j \pi_j = 1$.

Recall that the randomized decision rule $\rho_i(y) \in [0, 1]$ specifies the probability of deciding $\mathcal{H}_i$ when the observation is $y$.

When the observations are specified by conditional pdfs, our conditional risk function becomes

$$R_j(\rho) = \int_{y \in \mathcal{Y}} \left[ \sum_{i=0}^{M-1} \rho_i(y) C_{ij} \right] p_j(y) \, dy$$
Infinite Observation Spaces: Minimizing the Bayes Risk

The problem is then to solve

$$\rho^* = \arg \min_{\rho \in \mathcal{D}} r(\rho, \pi) = \arg \min_{\rho \in \mathcal{D}} \sum_{j=0}^{N-1} \pi_j \int_{y \in \mathcal{Y}} \left[ \sum_{i=0}^{M-1} \rho_i(y) C_{ij} \right] p_j(y) \, dy$$

where $\mathcal{D}$ is the set of all valid decision rules satisfying $\rho_i(y) \geq 0$ for all $i = 0, \ldots, M - 1$ and $y \in \mathcal{Y}$, as well as $\sum_{i=0}^{M-1} \rho_i(y) = 1$ for all $y \in \mathcal{Y}$. 
The problem is then to solve

$$\rho^* = \arg\min_{\rho \in \mathcal{D}} r(\rho, \pi) = \arg\min_{\rho \in \mathcal{D}} \sum_{j=0}^{N-1} \pi_j \int_{y \in \mathcal{Y}} \left[ \sum_{i=0}^{M-1} \rho_i(y) C_{ij} \right] p_j(y) \, dy$$

where $\mathcal{D}$ is the set of all valid decision rules satisfying $\rho_i(y) \geq 0$ for all $i = 0, \ldots, M - 1$ and $y \in \mathcal{Y}$, as well as $\sum_{i=0}^{M-1} \rho_i(y) = 1$ for all $y \in \mathcal{Y}$.

We can rearrange terms to write

$$\rho^* = \arg\min_{\rho \in \mathcal{D}} \sum_{j=0}^{N-1} \pi_j \int_{y \in \mathcal{Y}} \left[ \sum_{i=0}^{M-1} \rho_i(y) C_{ij} \right] p_j(y) \, dy$$

$$= \arg\min_{\rho \in \mathcal{D}} \int_{y \in \mathcal{Y}} \left( \sum_{i=0}^{M-1} \rho_i(y) \left[ \sum_{j=0}^{N-1} \pi_j C_{ij} p_j(y) \right] \right) \, dy$$
Infinite Observation Spaces: Minimizing the Bayes Risk

Recall for observation spaces with a finite number of possibilities, we had

$$D^* = \arg \min_{D \in \mathcal{D}} \sum_{\ell=0}^{L-1} \left( \sum_{i=0}^{M-1} D_{i\ell} \left[ \sum_{j=0}^{N-1} \pi_j C_{ij} p_j(y) \right] \right)$$

When we have an infinite number of possible observations, the only difference is that the sum over $\ell$ is replaced by an integral and $P_{\ell j} \to p_j(y)$. The minimization inside the sum/integral is the same.

Hence, when we have an infinite number of possible observations, we just do what we did before. Find the index of the minimum

$$m(y) = \arg \min_{i \in \{0, \ldots, M-1\}} \sum_{j=0}^{N-1} \pi_j C_{ij} p_j(y)$$

for each $y \in \mathcal{Y}$ and set $\rho_m(y) = 1$ and all other $\rho_m(y) = 0$. 
Bayesian Detection: Putting it All Together

Given a prior $\pi$, we can find the decision rule that minimizes the Bayes risk by computing

$$G_{i\ell} = \sum_{j=0}^{N-1} \pi_j C_{ij} P_{\ell j} \quad \text{(finite observation space)}$$

$$g_i(y) = \sum_{j=0}^{N-1} \pi_j C_{ij} p_j(y) \quad \text{(infinite observation space)}$$

and then finding the index of the minimum cost

$$m_\ell = \arg \min_{i \in \{0, \ldots, M-1\}} G_{i\ell} \quad \text{(finite observation space)}$$

$$m(y) = \arg \min_{i \in \{0, \ldots, M-1\}} g_i(y) \quad \text{(infinite observation space)}$$

for each $\ell = 0, \ldots, L-1$ (finite) or $y \in \mathcal{Y}$ (infinite) and setting

$$D_{m_\ell} = \rho_{m_\ell}(y) = 1 \quad \text{and} \quad D_{n,\ell} = \rho_n(y) = 0 \quad \text{for all} \ n \neq m_\ell \ \text{or} \ n \neq m(y).$$
Remarks

1. There is always at least one deterministic decision rule that minimizes the Bayes Risk.

2. The decision rule that minimizes the Bayes Risk may not be unique.

3. If there is more than one deterministic decision rule that minimizes the Bayes Risk then there will also be a corresponding convex hull of randomized decision rules that also achieve the minimum Bayes Risk.

4. Note that one of way of specifying deterministic decision rules is $\delta : \mathcal{Y} \mapsto \mathcal{Z}$. Hence,

$$\delta^{B\pi}(y_\ell) = \arg \min_{i \in \{0, \ldots, M-1\}} G_{i\ell} \quad \text{(finite observation space)}$$

$$\delta^{B\pi}(y) = \arg \min_{i \in \{0, \ldots, M-1\}} g_i(y) \quad \text{(infinite observation space)}$$

is a common notation for Bayesian decision rules.