ECE531 Screencast 9.3: Bayesian Detection with Finite Observations

D. Richard Brown III

Worcester Polytechnic Institute
The Bayesian Approach

We assume a prior state distribution $\pi \in \mathcal{P}_N$ such that

$$\text{Prob(state is } x_j \text{)} = \pi_j$$

Like Bayesian estimation, this prior reflects our belief of the state probabilities prior to the observation.
The Bayesian Approach

We assume a prior state distribution \( \pi \in \mathcal{P}_N \) such that

\[
\text{Prob(state is } x_j \text{)} = \pi_j
\]

Like Bayesian estimation, this prior reflects our belief of the state probabilities prior to the observation.

We denote the (scalar) Bayes Risk of the decision rule \( \rho \) as

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

This is simply the weighted overall risk, or average risk, given our prior belief of the state probabilities. A decision rule that minimizes this risk is called a Bayes decision rule for the prior \( \pi \).
Geometric Intuition

The prior $\pi$ weights the conditional risks and establishes a family of “level sets”. Given a constant $c \in \mathbb{R}$, the level set of value $c$ is defined as

$$L_c^\pi := \{ x \in \mathbb{R}^N : \pi^\top x = c \}$$
Given a decision matrix \( D \), we can write the Bayes Risk as

\[
 r(D, \pi) = \sum_{j=0}^{N-1} \pi_j R_j(D) = \sum_{j=0}^{N-1} \pi_j c_j^\top D p_j
\]

\[
 = \sum_{j=0}^{N-1} \pi_j \sum_{i=0}^{M-1} c_{ij} \sum_{\ell=0}^{L-1} D_{i\ell} p_{\ell j}
\]

\[
 = \sum_{\ell=0}^{L-1} \left( \sum_{i=0}^{M-1} D_{i\ell} \left[ \sum_{j=0}^{N-1} \pi_j c_{ij} p_{\ell j} \right] \right)
\]

\[
 = \sum_{\ell=0}^{L-1} \left( \sum_{i=0}^{M-1} D_{i\ell} G_{i\ell} \right) = \sum_{\ell=0}^{L-1} d_{\ell}^\top g_{\ell}
\]
Given a decision matrix \( D \), we can write the Bayes Risk as

\[
r(D, \pi) = \sum_{j=0}^{N-1} \pi_j R_j(D) = \sum_{j=0}^{N-1} \pi_j c_j^\top D p_j
\]

\[
= \sum_{j=0}^{N-1} \pi_j \sum_{i=0}^{M-1} \sum_{\ell=0}^{L-1} C_{ij} D_{i\ell} P_{\ell j}
\]

\[
= \sum_{\ell=0}^{L-1} \left( \sum_{i=0}^{M-1} D_{i\ell} \left[ \sum_{j=0}^{N-1} \pi_j C_{ij} P_{\ell j} \right] \right)
\]

\[
= \sum_{\ell=0}^{L-1} \left( \sum_{i=0}^{M-1} D_{i\ell} G_{i\ell} \right) = \sum_{\ell=0}^{L-1} d_{\ell}^\top g_{\ell}
\]

Note that we can minimize each term in this sum separately.
Example Part 1

Let $G_{i\ell} := \sum_{j=0}^{N-1} \pi_j C_{ij} P_{\ell j}$ and $G \in \mathbb{R}^{M \times L}$ be the matrix composed of elements $G_{i\ell}$. Suppose

$$G = \begin{bmatrix}
0.3 & 0.5 & 0.2 & 0.8 \\
0.4 & 0.2 & 0.1 & 0.5 \\
0.5 & 0.1 & 0.7 & 0.6
\end{bmatrix}$$

What decision rule minimizes the Bayes Risk?
Example Part 1

Let \( G_{i\ell} := \sum_{j=0}^{N-1} \pi_j C_{ij} P_{\ell j} \) and \( G \in \mathbb{R}^{M \times L} \) be the matrix composed of elements \( G_{i\ell} \). Suppose

\[
G = \begin{bmatrix}
0.3 & 0.5 & 0.2 & 0.8 \\
0.4 & 0.2 & 0.1 & 0.5 \\
0.5 & 0.1 & 0.7 & 0.6
\end{bmatrix}
\]

What decision rule minimizes the Bayes Risk?

\[
D = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 \\
0 & 1 & 0 & 0
\end{bmatrix}
\]

This is a deterministic decision rule and it is unique. Note that the Bayes Risk of this decision rule is then simply

\[
r(\pi, D) = 0.3 + 0.1 + 0.1 + 0.5 = 1.0.
\]
Example Part 2

What happens if

\[
G = \begin{bmatrix}
0.3 & 0.5 & 0.2 & 0.8 \\
0.4 & 0.2 & 0.2 & 0.5 \\
0.5 & 0.1 & 0.7 & 0.6
\end{bmatrix}
\]

In this case, both

\[
D = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 \\
0 & 1 & 0 & 0
\end{bmatrix}
\]

or

\[
D = \begin{bmatrix}
1 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0
\end{bmatrix}
\]

achieve the same Bayes Risk \( r(\pi, D) = 1.1 \).
Example Part 2

What happens if

\[ G = \begin{bmatrix}
0.3 & 0.5 & 0.2 & 0.8 \\
0.4 & 0.2 & 0.2 & 0.5 \\
0.5 & 0.1 & 0.7 & 0.6
\end{bmatrix} \]

In this case, both

\[ D = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 \\
0 & 1 & 0 & 0
\end{bmatrix} \quad \text{or} \quad D = \begin{bmatrix}
1 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0
\end{bmatrix} \]

achieve the same Bayes Risk \( r(\pi, D) = 1.1 \).

In fact, any decision matrix of the form

\[ D = \begin{bmatrix}
1 & 0 & \alpha & 0 \\
0 & 0 & 1 - \alpha & 1 \\
0 & 1 & 0 & 0
\end{bmatrix} \]

for \( \alpha \in [0, 1] \) will also achieve \( r(\pi, D) = 1.1 \).
Summary: Finite Observations

To minimize the Bayes Risk for finite \( \mathcal{Y} \), we just find the index

\[
m_\ell = \arg\min_{i \in \{0, \ldots, M-1\}} \sum_{j=0}^{N-1} \pi_j C_{ij} P_{\ell j} \left( G_{i \ell} \right)
\]

for each \( \ell = 0, \ldots, L - 1 \) and set \( D_{m_\ell, \ell}^{B\pi} = 1 \) and \( D_{i, \ell}^{B\pi} = 0 \) for all \( i \neq m_\ell \).