ECE531 Screencast 9.5: Bayesian Detection Interpretation and Special Cases

D. Richard Brown III

Worcester Polytechnic Institute
Prior and Posterior Probabilities

The conditional probability that we are in state $x_j$ given the observation $y$:

$$
\pi_j(y) := \text{Prob}(X = x_j \mid Y = y) = \frac{p_j(y)\pi_j}{p(y)}
$$

where

$$
p(y) = \sum_{j=0}^{N-1} \pi_j p_j(y).
$$

- Recall that $\pi_j$ is the prior probability of state $x_j$, before we have any observations.
- The quantity $\pi_j(y)$ is the posterior probability of state $x_j$, conditioned on the observation $y$. 
A Bayes Decision Rule Minimizes The Posterior Cost

Deterministic Bayes decision rule (infinite observation space):

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

But $\pi_j p_j(y) = \pi_j(y) p(y)$, hence

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

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

since $p(y)$ does not affect the minimizer.
A Bayes Decision Rule Minimizes The Posterior Cost

Deterministic Bayes decision rule (infinite observation space):

\[ \delta^{B\pi}(y) = \arg \min_{i \in \{0, \ldots, M-1\}} \left( \sum_{j=0}^{N-1} \pi_j C_{ij} p_j(y) \right) \]

But \( \pi_j p_j(y) = \pi_j(y) p(y) \), hence

\[ \delta^{B\pi}(y) = \arg \min_{i \in \{0, \ldots, M-1\}} \left( \sum_{j=0}^{N-1} C_{ij} \pi_j(y) p(y) \right) \]

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

since \( p(y) \) does not affect the minimizer. Interpretation: \( \sum_{j=0}^{N-1} C_{ij} \pi_j(y) \) the average cost of choosing hypothesis \( \mathcal{H}_i \) given \( Y = y \), i.e. the posterior cost of choosing \( \mathcal{H}_i \). The Bayes decision rule chooses the hypothesis that yields the minimum expected posterior cost.
Bayesian Hypothesis Testing with UCA: Part 1

The uniform cost assignment (UCA):

\[ C_{ij} = \begin{cases} 
0 & \text{if } x_j \in \mathcal{H}_i \\
1 & \text{otherwise} 
\end{cases} \]

The conditional risk \( R_j(\rho) \) under the UCA is simply the probability of not choosing the hypothesis that contains \( x_j \), i.e. the probability of error when the state is \( x_j \). The Bayes risk in this case is

\[
r(\rho, \pi) = \sum_{j=0}^{N-1} \pi_j R_j(\rho) = \text{Prob(error)}.
\]
Under the UCA, a Bayes decision rule can be written in terms of the posterior probabilities as

\[
\delta^{B\pi}(y) = \arg \min_{i \in \{0, \ldots, M-1\}} \sum_{j=0}^{N-1} C_{ij} \pi_j(y) \\
= \arg \min_{i \in \{0, \ldots, M-1\}} \sum_{x_j \notin \mathcal{H}_i} \pi_j(y) \\
= \arg \min_{i \in \{0, \ldots, M-1\}} \left[ 1 - \sum_{x_j \in \mathcal{H}_i} \pi_j(y) \right] \\
= \arg \max_{i \in \{0, \ldots, M-1\}} \sum_{x_j \in \mathcal{H}_i} \pi_j(y)
\]
Under the UCA, a Bayes decision rule can be written in terms of the posterior probabilities as

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

$$= \arg \min_{i \in \{0, \ldots, M-1\}} \sum_{x_j \notin \mathcal{H}_i} \pi_j(y)$$

$$= \arg \min_{i \in \{0, \ldots, M-1\}} \left[ 1 - \sum_{x_j \in \mathcal{H}_i} \pi_j(y) \right]$$

$$= \arg \max_{i \in \{0, \ldots, M-1\}} \sum_{x_j \in \mathcal{H}_i} \pi_j(y)$$

Hence, for hypothesis tests under the UCA, the Bayes decision rule is the MAP (maximum a posteriori) decision rule. When the hypothesis test is simple, $$\delta^{B\pi}(y) = \arg \max_i \pi_i(y)$$.
Bayesian Hypothesis Testing with UCA and Uniform Prior

Under the UCA and a uniform prior, i.e. $\pi_j = 1/N$ for all $j = 0, \ldots, N - 1$

$$
\delta^{B\pi}(y) = \arg \max_{i \in \{0, \ldots, M-1\}} \sum_{x_j \in \mathcal{H}_i} \pi_j(y)
$$

$$
= \arg \max_{i \in \{0, \ldots, M-1\}} \sum_{x_j \in \mathcal{H}_i} \pi_j p_j(y)
$$

$$
= \arg \max_{i \in \{0, \ldots, M-1\}} \sum_{x_j \in \mathcal{H}_i} p_j(y)
$$

since $\pi_j = 1/N$ does not affect the minimizer.

- In this case, $\delta^{B\pi}(y)$ selects the most likely hypothesis, i.e. the hypothesis which best explains the observation $y$.
- This is called the maximum likelihood (ML) decision rule.
Simple Binary Bayesian Hypothesis Testing: Part 1

We have two states $x_0$ and $x_1$ and two hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$. For each $y \in \mathcal{Y}$, our problem is to compute

$$m(y) = \arg \min_{i \in \{0,1\}} g_i(y, \pi)$$

for each $y \in \mathcal{Y}$ where

$$g_0(y, \pi) = \pi_0 C_{00} p_0(y) + \pi_1 C_{01} p_1(y)$$
$$g_1(y, \pi) = \pi_0 C_{10} p_0(y) + \pi_1 C_{11} p_1(y)$$

We only have two things to compare. We can simplify this comparison:

$$g_0(y, \pi) \geq g_1(y, \pi) \iff \pi_0 C_{00} p_0(y) + \pi_1 C_{01} p_1(y) \geq \pi_0 C_{10} p_0(y) + \pi_1 C_{11} p_1(y)$$
$$\iff p_1(y) \pi_1 (C_{01} - C_{11}) \geq p_0(y) \pi_0 (C_{10} - C_{00})$$
$$\iff \frac{p_1(y)}{p_0(y)} \geq \frac{\pi_0 (C_{10} - C_{00})}{\pi_1 (C_{01} - C_{11})}$$

where we have assumed that $C_{01} > C_{11}$ to get the final result. The expression $L(y) := \frac{p_1(y)}{p_0(y)}$ is known as the likelihood ratio.
Simple Binary Bayesian Hypothesis Testing: Part 2

Given \( L(y) := \frac{p_1(y)}{p_0(y)} \) and

\[
\tau := \frac{\pi_0(C_{10} - C_{00})}{\pi_1(C_{01} - C_{11})},
\]

the Bayes decision rule for simple binary hypothesis testing is then simply

\[
\delta^B\pi(y) = \begin{cases} 
1 & \text{if } L(y) > \tau \\
0/1 & \text{if } L(y) = \tau \\
0 & \text{if } L(y) < \tau.
\end{cases}
\]

Remark:

- For any \( y \in \mathcal{Y} \) that result in \( L(y) = \tau \), the Bayes risk is the same whether we decide \( \mathcal{H}_0 \) or \( \mathcal{H}_1 \). You can deal with this by always deciding \( \mathcal{H}_0 \) in this case, or always deciding \( \mathcal{H}_1 \), or flipping a coin, etc.
Simple Binary Bayesian Hypothesis Testing with UCA

Uniform cost assignment:

$$C_{00} = C_{11} = 0$$
$$C_{01} = C_{10} = 1$$

In this case, the discriminant functions are simply

$$g_0(y, \pi) = \pi_1 p_1(y) = \pi_1(y) p(y)$$
$$g_1(y, \pi) = \pi_0 p_0(y) = \pi_0(y) p(y)$$

and a Bayes decision rule can be written in terms of the posterior probabilities as

$$\delta^{B\pi}(y) = \begin{cases} 1 & \text{if } \pi_1(y) > \pi_0(y) \\ 0/1 & \text{if } \pi_1(y) = \pi_0(y) \\ 0 & \text{if } \pi_1(y) < \pi_0(y). \end{cases}$$

In this case, it should be clear that the Bayes decision rule is the MAP (maximum a posteriori) decision rule.