EEE 598C: Statistical Pattern Recognition
Lecture Note 2: Neyman / Pearson Decision Theory

Darryl Morrell

September 3, 1996

1 Introduction

We consider the two-category classification problem, in which there are two states of nature ω_1 and ω_2, and two possible actions α_1 and α_2; α_j is the action of choosing ω_j. For the Neyman / Pearson approach to decision theory, we assume that the conditional densities $f(\mathbf{x}|\omega_1)$ and $f(\mathbf{x}|\omega_2)$ are known, but we do not use any prior probabilities nor do we use a loss function.

We define two types of errors:

False Alarm: A false alarm is when we choose action α_2 when ω_1 is the true state of nature. We denote P_{FA} as the probability of a false alarm.

Missed Detection: A missed detection is when we choose action α_1 when ω_2 is the true state of nature. We denote P_{MD} as the probability of a missed detection, and P_D as the probability of a correct detection.

The terminology for these two types of errors comes from detection problems, in which ω_1 is the state of nature in which an object of interest (ie. target, signal, etc.) is absent, and ω_2 is the state of nature in which the object is present.

As in the Bayesian Decision theory case, our decision rule $\alpha(\mathbf{x})$ divides the space of possible observations into two regions:

$$\mathcal{R}_1 = \{\mathbf{x} : \alpha(\mathbf{x}) = \alpha_1\}$$

$$\mathcal{R}_2 = \{\mathbf{x} : \alpha(\mathbf{x}) = \alpha_2\}$$

The probability of false alarm can be expressed as

$$P_{FA} = \int_{\mathcal{R}_2} f(\mathbf{x}|\omega_1) d\mathbf{x} \tag{1}$$

The probability of correct detection can be expressed as

$$P_D = \int_{\mathcal{R}_2} f(\mathbf{x}|\omega_2) d\mathbf{x}$$

A *Receiver Operating Curve* (ROC) is a plot of P_D as a function of P_{FA}.

A Neyman/Pearson test is guaranteed to maximize P_D subject to the constraint that $P_{FA} \leq a$. It takes the form of a likelihood ratio test:

Choose α_1 if

$$\frac{f(\mathbf{x}|\omega_1)}{f(\mathbf{x}|\omega_2)} \geq T$$

where T is a threshold chosen to mee the P_{FA} constraint.

Note that this result need not be interpreted in terms of detection; a Neyman Pearson test minimizes the probability of one type of missclassification error subject to the constraint that the other type of missclassification errror is no larger than a.

The threshold T is chosen to obtain the desired P_{FA} as follows: given T and the likelihood ratio test, $\mathcal{R}_2$ can be written as

$$\mathcal{R}_2 = \{\mathbf{x} : f(\mathbf{x}|\omega_1) < T f(\mathbf{x}|\omega_2)\}$$

P_{FA} is given in terms of $\mathcal{R}_2$ by (1). We can use this relationship to determine the value of T.

Exercise 1 *Suppose that under ω_1, $\mathbf{X}$ is Gaussian with mean one and variance one, and that under ω_2, $\mathbf{X}$ is Gaussian with mean negative one and variance one:*

$$f(\mathbf{x}|\omega_1) = \frac{1}{\sqrt{2\pi}} e^{-(x-1)^2/2}$$

$$f(\mathbf{x}|\omega_2) = \frac{1}{\sqrt{2\pi}} e^{-(x+1)^2/2}$$

1. *Find the likelihood ratio test.*
2. *How would you choose T for a given P_{FA}?*
3. *How would you compute P_D?*

2 Proof of Neyman-Pearson Lemma

For this section, we represent the decision rule using a function $\phi(\mathbf{x})$ defined as

$$\phi(\mathbf{x}) = \begin{cases} 1, & \frac{f(\mathbf{X}|\omega_1)}{f(\mathbf{X}|\omega_2)} \geq T \\ 0, & \frac{f(\mathbf{X}|\omega_1)}{f(\mathbf{X}|\omega_2)} < T \end{cases} \tag{2}$$

where ϕ takes a value of one when we choose ω_2 and a value of zero when we choose ω_1. This representation of a decision rule has the advantage that the probabilities of false alarm and detection can be represented in terms of expected values of $\phi(\mathbf{X})$:

$$P_{FA} = \int_{\mathcal{R}_2} f(\mathbf{x}|\omega_1)d\mathbf{x} = \int \phi(\mathbf{x})f(\mathbf{x}|\omega_1)d\mathbf{x} = E\left[\phi(\mathbf{X})|\omega_1\right]$$

$$P_D = \int_{\mathcal{R}_2} f(\mathbf{x}|\omega_2)d\mathbf{x} = \int \phi(\mathbf{x})f(\mathbf{x}|\omega_2)d\mathbf{x} = E\left[\phi(\mathbf{X})|\omega_2\right]$$

where we have used the fact that $\phi(\mathbf{x})$ is one when $\mathbf{x} \in \mathcal{R}_2$ and zero when $\mathbf{x} \notin \mathcal{R}_2$.

In the proof of the Neyman Pearson lemmma, we will consider the decision rule ϕ given in (2) and another decision rule ϕ'; we will denote probabilities of detection and false alarm for these two decision rules as $P_{FA}(\phi)$, $P_{FA}(\phi')$, $P_D(\phi)$, and $P_D(\phi')$, to denote the fact that these probabilities depend on the decision rule used.

We now state the Neyman-Pearson lemma[1]:

Theorem 1 *Consider the decision rule of (2) with T chosen to give $P_{FA}(\phi) = a$. There is no decision rule ϕ' such that $P_{FA}(\phi') \leq a$ and $P_D(\phi') > P_D(\phi)$.*

Proof: Let ϕ' be a decision rule with

$$P_{FA}(\phi') = E\left[\phi'(\mathbf{X})|\omega_1\right] \leq a$$

The following identity holds for any decision rule ϕ':

$$\int \left[\phi(\mathbf{x}) - \phi'(\mathbf{x})\right]\left[f(\mathbf{x}|\omega_2) - Tf(\mathbf{x}|\omega_1)\right]d\mathbf{x} \geq 0 \qquad (3)$$

This can be seen by considering all possible values of $\mathbf{x}$: for those values for which $\phi(\mathbf{x}) = 1$, $\phi(\mathbf{x}) - \phi'(\mathbf{x}) \geq 0$, since $0 \leq \phi'(\mathbf{x}) \leq 1$, and $f(\mathbf{x}|\omega_2) - Tf(\mathbf{x}|\omega_1) \geq 0$; similarly, for those values for which $\phi(\mathbf{x}) = 0$, $\phi(\mathbf{x}) - \phi'(\mathbf{x}) \leq 0$ and $f(\mathbf{x}|\omega_2) - Tf(\mathbf{x}|\omega_1) < 0$. Multiplying out (3) and writing the result in terms of probabilities of detection and false alarms, we get

$$\begin{aligned}
&\int \left[\phi(\mathbf{x}) - \phi'(\mathbf{x})\right]\left[f(\mathbf{x}|\omega_2) - Tf(\mathbf{x}|\omega_1)\right]d\mathbf{x} \\
&= \int \phi(\mathbf{x})f(\mathbf{x}|\omega_2)d\mathbf{x} - \int \phi'(\mathbf{x})f(\mathbf{x}|\omega_2)d\mathbf{x} - T\int \phi(\mathbf{x})f(\mathbf{x}|\omega_1)d\mathbf{x} + \int \phi'(\mathbf{x})f(\mathbf{x}|\omega_1)d\mathbf{x} \\
&= \left[P_D(\phi) - \P_D(\phi')\right] - T\left[P_{FA}(\phi) - P_{FA}(\phi')\right] \\
&\geq P_D(\phi) - \P_D(\phi') \\
&\geq 0
\end{aligned}$$

where the first inequality follows from the assumption that $P_{FA}(\phi') \leq a = P_{FA}(\phi)$, and the second inequality follows from (3). Thus, $P_D(\phi) \geq P_D(\phi')$, and $\phi(\mathbf{x})$ is a decision rule that maximizes the probability of detection for a given probability of false alarm. □.

So likelihood ratio tests are both Bayes optimal and Neyman-Pearson optimal.

---

[1] Note that this is actually a somewhat simplified version of the lemma that only applies when the set $\{\mathbf{x} : f(\mathbf{x}|\omega_1) = Tf(\mathbf{x}|\omega_2)\}$ has measure zero for all $T > 0$; when this condition does not hold, then the Neyman-Pearson optimal decision rule may be a *random decision rule*. In a random decision rule, there are values of $\mathbf{x}$ for which one would choose ω_1 with a given probability p and ω_2 with probability $1 - p$.