Lecture 19

In the last lecture we found the Bayes decision rule that minimizes the Bayes eror

$$\alpha = \sum_{i=1}^{k} \xi(i)\alpha_i = \sum_{i=1}^{k} \xi(i)\mathbb{P}_i(\delta \neq H_i).$$

Let us write down this decision rule in the case of two simple hypothesis H_1, H_2. For simplicity of notations, given the sample $X = (X_1, \ldots, X_n)$ we will denote the joint p.d.f. by

$$f_i(X) = f_i(X_1) \ldots f_i(X_n).$$

Then in the case of two simple hypotheses the Bayes decision rule that minimizes the Bayes error

$$\alpha = \xi(1)\mathbb{P}_1(\delta \neq H_1) + \xi(2)\mathbb{P}_2(\delta \neq H_2)$$

is given by

$$\delta = \begin{cases} H_1 : & \xi(1)f_1(X) > \xi(2)f_2(X) \\ H_2 : & \xi(2)f_2(X) > \xi(1)f_1(X) \\ H_1 \text{ or } H_2 : & \xi(1)f_1(X) = \xi(2)f_2(X) \end{cases}$$

or, equivalently,

$$\delta = \begin{cases} H_1 : & \frac{f_1(X)}{f_2(X)} > \frac{\xi(2)}{\xi(1)} \\ H_2 : & \frac{f_1(X)}{f_2(X)} < \frac{\xi(2)}{\xi(1)} \\ H_1 \text{ or } H_2 : & \frac{f_1(X)}{f_2(X)} = \frac{\xi(2)}{\xi(1)} \end{cases} \tag{19.1}$$

(Here $\frac{1}{0} = +\infty$, $\frac{0}{1} = 0$.) This kind of test if called *likelihood ratio test* since it is expressed in terms of the ratio $f_1(X)/f_2(X)$ of likelihood functions.

Example. Suppose we have only one observation X_1 and two simple hypotheses $H_1 : \mathbb{P} = N(0, 1)$ and $H_2 : \mathbb{P} = N(1, 1)$. Let us take an apriori distribution given by

$$\xi(1) = \frac{1}{2} \text{ and } \xi(2) = \frac{1}{2},$$

i.e. both hypothesis have equal weight, and find a Bayes decision rule δ that minimizes

$$\frac{1}{2}\mathbb{P}_1(\delta \neq H_1) + \frac{1}{2}\mathbb{P}_2(\delta \neq H_2)$$

Bayes decision rule is given by:

$$\delta(X_1) = \begin{cases} H_1 : & \frac{f_1(X)}{f_2(X)} > 1 \\ H_2 : & \frac{f_1(X)}{f_2(X)} < 1 \\ H_1 \ or \ H_2 : & \frac{f_1(X)}{f_2(X)} = 1 \end{cases}$$

This decision rule has a very intuitive interpretation. If we look at the graphs of these p.d.f.s (figure 19.1) the decision rule picks the first hypothesis when the first p.d.f. is larger, to the left of point C, and otherwise picks the second hypothesis to the right of point C.

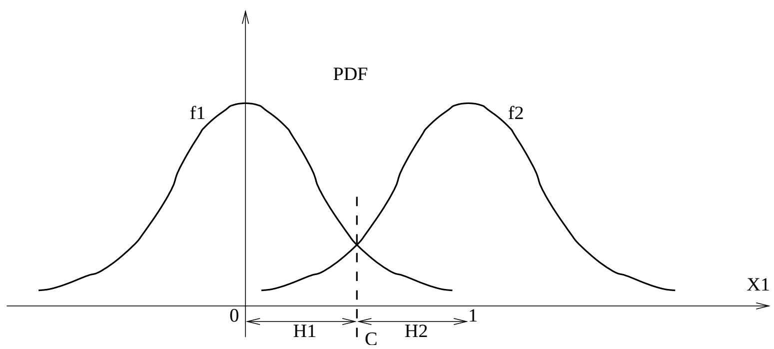

Figure 19.1: Bayes Decision Rule.

Example. Let us now consider a similar but more general example when we have a sample $X = (X_1, \cdots, X_n)$, two simple hypotheses $H_1 : \mathbb{P} = N(0, 1)$ and $H_2 : \mathbb{P} = N(1, 1)$, and arbitrary apriori weights $\xi(1), \xi(2)$. Then Bayes decision rule is given by (19.1). The likelihood ratio can be simplified:

$$\begin{aligned} \frac{f_1(X)}{f_2(X)} &= \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2}\sum X_i^2} \Big/ \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2}\sum (X_i-1)^2} \\ &= e^{\frac{1}{2}\sum_{i=1}^k ((X_i-1)^2 - X_i^2)} = e^{\frac{n}{2} - \sum X_i} \end{aligned}$$

Therefore, the decision rule picks the first hypothesis H_1 when

$$e^{\frac{n}{2} - \sum X_i} > \frac{\xi(2)}{\xi(1)}$$

or, equivalently,

$$\sum X_i < \frac{n}{2} - \log \frac{\xi(2)}{\xi(1)}.$$

Similarly, we pick the second hypothesis H_2 when

$$\sum X_i > \frac{n}{2} - \log \frac{\xi(2)}{\xi(1)}.$$

In case of equality, we pick any one of H_1, H_2.

$\square$

19.1 Most powerful test for two simple hypotheses.

Now that we learned how to construct the decision rule that minimizes the Bayes error we will turn to our next goal which we discussed in the last lecture, namely, how to construct the decision rule with controlled error of type 1 that minimizes error of type 2. Given $\alpha \in [0,1]$ we consider the class of decision rules

$$K_\alpha = \{\delta : \mathbb{P}_1(\delta \neq H_1) \leq \alpha\}$$

and will try to find $\delta \in K_\alpha$ that makes the type 2 error $\alpha_2 = \mathbb{P}_2(\delta \neq H_2)$ as small as possible.

Theorem: *Assume that there exist a constant c, such that*

$$\mathbb{P}_1\Big(\frac{f_1(X)}{f_2(X)} < c\Big) = \alpha.$$

Then the decision rule

$$\delta = \begin{cases} H_1 : & \frac{f_1(X)}{f_2(X)} \geq c \\ H_2 : & \frac{f_1(X)}{f_2(X)} < c \end{cases} \tag{19.2}$$

is the most powerful in class K_α.

Proof. Take $\xi(1)$ and $\xi(2)$ such that

$$\xi(1) + \xi(2) = 1, \frac{\xi(2)}{\xi(1)} = c,$$

i.e.

$$\xi(1) = \frac{1}{1+c} \text{ and } \xi(2) = \frac{c}{1+c}.$$

Then the decision rule δ in (19.2) is the Bayes decision rule corresponding to weights $\xi(1)$ and $\xi(2)$ which can be seen by comparing it with (19.1), only here we break the tie in favor of H_1. Therefore, this decision rule δ minimizes the Bayes error which means that for any other decision rule δ',

$$\xi(1)\mathbb{P}_1(\delta \neq H_1) + \xi(2)\mathbb{P}_2(\delta \neq H_2) \leq \xi(1)\mathbb{P}_1(\delta' \neq H_1) + \xi(2)\mathbb{P}_2(\delta' \neq H_2). \quad (19.3)$$

By assumption in the statement of the Theorem, we have

$$\mathbb{P}_1(\delta \neq H_1) = \mathbb{P}_1\Big(\frac{f_1(X)}{f_2(X)} < c\Big) = \alpha,$$

which means that δ comes from the class K_α. If $\delta' \in K_\alpha$ then

$$\mathbb{P}_1(\delta' \neq H_1) \leq \alpha$$

and equation (19.3) gives us that

$$\xi(1)\alpha + \xi(2)\mathbb{P}_2(\delta \neq H_2) \leq \xi(1)\alpha + \xi(2)\mathbb{P}_2(\delta' \neq H_2)$$

and, therefore,

$$\mathbb{P}_2(\delta \neq H_2) \leq \mathbb{P}_2(\delta' \neq H_2).$$

This exactly means that δ is more powerful than any other decision rule in class K_α.

$\square$

Example. Suppose we have a sample $X = (X_1, \ldots, X_n)$ and two simple hypotheses $H_1 : \mathbb{P} = N(0,1)$ and $H_2 : \mathbb{P} = N(1,1)$. Let us find most powerful δ with the error of type 1

$$\alpha_1 \leq \alpha = 0.05.$$

According to the above Theorem if we can find c such that

$$\mathbb{P}_1\Big(\frac{f_1(X)}{f_2(X)} < c\Big) = \alpha = 0.05$$

then we know how to find δ. Simplifying this equation gives

$$\mathbb{P}_1\Big(\sum X_i > \frac{n}{2} - \log c\Big) = \alpha = 0.05$$

or

$$\mathbb{P}_1\Big(\frac{1}{\sqrt{n}}\sum X_i > c' = \frac{1}{\sqrt{n}}(\frac{n}{2} - \log c)\Big) = \alpha = 0.05.$$

But under the hypothesis H_1 the sample comes from standard normal distribution $\mathbb{P}_1 = N(0,1)$ which implies that the random variable

$$Y = \frac{1}{\sqrt{n}}\sum X_i$$

is standard normal. We can look up in the table that

$$\mathbb{P}(Y > c') = \alpha = 0.05 \Rightarrow c' = 1.64$$

and the most powerful test δ with level of significance $\alpha = 0.05$ will look like this:

$$\delta = \begin{cases} H_1 : & \frac{1}{\sqrt{n}} \sum X_i \leq 1.64 \\ H_2 : & \frac{1}{\sqrt{n}} \sum X_i > 1.64. \end{cases}$$

Now, what will the error of type 2 be for this test?

$$\begin{aligned} \alpha_2 &= \mathbb{P}_2(\delta \neq H_2) = \mathbb{P}_2\Big(\frac{1}{\sqrt{n}} \sum X_i \leq 1.64\Big) \\ &= \mathbb{P}_2\Big(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - 1) \leq 1.64 - \sqrt{n}\Big). \end{aligned}$$

The reason we subtracted 1 from each X_i is because under the second hypothesis X's have distribution $N(1, 1)$ and random variable

$$Y = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - 1)$$

will be standard normal. Therefore, the error of type 2 for this test will be equak to the probability $\mathbb{P}(Y < 1.64 - \sqrt{n})$. For example, when the sample size $n = 10$ this will be

$$\alpha_2 = \mathbb{P}(Y < 1.64 - \sqrt{10}) = 0.087$$

from the table of normal distribution.