Lecture 20

20.1 Randomized most powerful test.

In theorem in the last lecture we showed how to find the most powerful test with level of significance α (which means that $\delta \in K_\alpha$), if we can find c such that

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

This condition is not always fulfiled, especially when we deal with discrete distributions as will become clear from the examples below. But if we look carefully at the proof of that Theorem, this condition was only necessary to make sure that the likelihood ratio test has error of type 1 exactly equal to α. In our next theorem we will show that the most powerful test in class K_α can always be found if one randomly breaks the tie between two hypotheses in a way that ensures that the error of type one is equal to α.

Theorem. *Given any $\alpha \in [0,1]$ we can always find $c \in [0,\infty)$ and $p \in [0,1]$ such that*

$$\mathbb{P}_1\Big(\frac{f_1(X)}{f_2(X)} < c\Big) + (1-p)\mathbb{P}_1\Big(\frac{f_1(X)}{f_2(X)} = c\Big) = \alpha. \tag{20.1}$$

In this case, the most powerful test $\delta \in K_\alpha$ is given by

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

where in the last case of equality we break the tie at random by choosing H_1 with probability p and choosing H_2 with probability $1-p$.

This test δ is called a *randomized test* since we break a tie at random if necessary.

Proof. Let us first assume that we can find c and p such that (20.1) holds. Then the error of type 1 for the randomized test δ above can be computed:

$$\alpha_1 = \mathbb{P}_1(\delta \neq H_1) = \mathbb{P}_1\Big(\frac{f_1(X)}{f_2(X)} < c\Big) + (1-p)\mathbb{P}_1\Big(\frac{f_1(X)}{f_2(X)} = c\Big) = \alpha \qquad (20.2)$$

since δ does not pick H_1 exactly when the likelihood ratio is less than c or when it is equal to c in which case H_1 is not picked with probability $1-p$. This means that the randomized test $\delta \in K_\alpha$. The rest of the proof repeats the proof of the last Theorem. We only need to point out that our randomized test will still be Bayes test since in the case of equality

$$\frac{f_1(X)}{f_2(X)} = c$$

the Bayes test allows one to break the tie arbitrarily and we choose to break it randomly in a way that ensures that the error of type one will be equal to α, as in (20.2).

The only question left is why we can always choose c and p such that (20.1) is satisfied. If we look at the function

$$F(t) = \mathbb{P}\Big(\frac{f_1(X)}{f_2(X)} < t\Big)$$

as a function of t, it will increase from 0 to 1 as t increases from 0 to ∞. Let us keep in mind that, in general, $F(t)$ might have jumps. We can have two possibilities: either (a) at some point $t = c$ the function $F(c)$ will be equal to α, i.e.

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

or (b) at some point $t = c$ it will jump over α, i.e.

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

but

$$\mathbb{P}\Big(\frac{f_1(X)}{f_2(X)} \le c\Big) = F(c) + \mathbb{P}\Big(\frac{f_1(X)}{f_2(X)} = c\Big) \ge \alpha.$$

Then (20.1) will hold if in case (a) we take $p = 1$ and in case (b) we take

$$1 - p = (\alpha - F(c)) \Big/ \mathbb{P}\Big(\frac{f_1(X)}{f_2(X)} = c\Big).$$

$\square$

Example. Suppose that we have one observation X with Bernoulli distribution and two simple hypotheses about the probability function $f(X)$ are

$$\begin{aligned} H_1 : f_1(X) &= 0.2^X 0.8^{1-X} \\ H_2 : f_2(X) &= 0.4^X 0.6^{1-X}. \end{aligned}$$

Let us take the level of significance $\alpha = 0.05$ and find the most powerful $\delta \in K_{0.05}$. In figure 20.1 we show the graph of the function

$$F(c) = \mathbb{P}_1\Bigl(\frac{f_1(X)}{f_2(X)} < c\Bigr).$$

Let us explain how this graph is obtained. First of all, the likelihood ratio can take

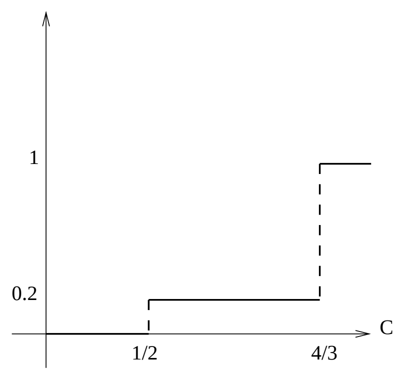

Figure 20.1: Graph of $F(c)$.

only two values:

$$\frac{f_1(X)}{f_2(X)} = \begin{cases} 1/2 \text{ if } & X = 1 \\ 4/3 \text{ if } & X = 0. \end{cases}$$

If $c \le \frac{1}{2}$ then the set

$$\Bigl\{\frac{f_1(X)}{f_2(X)} < c\Bigr\} = \emptyset \text{ is empty and } F(c) = \mathbb{P}_1(\emptyset) = 0,$$

if $\frac{1}{2} < c \le \frac{4}{3}$ then the set

$$\Bigl\{\frac{f_1(X)}{f_2(X)} < c\Bigr\} = \{X = 1\} \text{ and } F(c) = \mathbb{P}_1(X = 1) = 0.2$$

and, finally, if $\frac{4}{3} < c$ then the set

$$\Bigl\{\frac{f_1(X)}{f_2(X)} < c\Bigr\} = \{X = 0 \text{ or } 1\} \text{ and } F(c) = \mathbb{P}_1(X = 0 \text{ or } 1) = 1,$$

as shown in figure 20.1. The function $F(c)$ jumps over the level $\alpha = 0.05$ at the point $c = 1/2$. To determine p we have to make sure that the error of type one is equal to 0.05, i.e.

$$\mathbb{P}_1\Big(\frac{f_1(X)}{f_2(X)} < c\Big) + (1-p)\mathbb{P}_1\Big(\frac{f_1(X)}{f_2(X)} = c\Big) = 0 + (1-p)0.2 = 0.05$$

which gives that $p = \frac{3}{4}$. Therefore, the most powerful test of size $\alpha = 0.05$ is

$$\delta = \begin{cases} H_1: & \frac{f_1(X)}{f_2(X)} > \frac{1}{2} \text{ or } X = 0 \\ H_2: & \frac{f_1(X)}{f_2(X)} < \frac{1}{2} \text{ or never} \\ H_1 \text{ or } H_2: & \frac{f_1(X)}{f_2(X)} = \frac{1}{2} \text{ or } X = 1, \end{cases}$$

where in the last case $X = 1$ we pick H_1 with probability $\frac{3}{4}$ or H_2 with probability $\frac{1}{4}$.

20.2 Composite hypotheses. Uniformly most powerful test.

We now turn to a more difficult situation then the one when we had only two simple hypotheses. We assume that the sample $X_1, \ldots, X_n$ has distribution $\mathbb{P}_{\theta_0}$ that comes from a set of probability distributions $\{\mathbb{P}_\theta, \theta \in \Theta\}$. Given the sample, we would like to decide whether unknown θ_0 comes from the set Θ_1 or Θ_2, in which case our hypotheses will be

$$H_1: \theta \in \Theta_1 \subseteq \Theta$$

$$H_2: \theta \in \Theta_2 \subseteq \Theta.$$

Given some decision rule δ, let us consider a function

$$\Pi(\delta, \theta) = \mathbb{P}_\theta(\delta \neq H_1) \text{ as a function of } \theta,$$

which is called the *power function* of δ. The power function has different meaning depending on whether θ comes from Θ_1 or Θ_2, as can be seen in figure 20.2.

For $\theta \in \Theta_1$ the power function represents the error of type 1, since θ actually comes from the set in the first hypothesis H_1 and δ rejects H_1. If $\theta \in \Theta_2$ then the power function represents the power, or one minus error of type two, since in this case θ belongs to a set from the second hypothesis H_2 and δ accepts H_2. Therefore, ideally, we would like to minimize the power function for all $\theta \in \Theta_1$ and maximize it for all $\theta \in \Theta_2$.

Consider

$$\alpha_1(\delta) = \sup_{\theta\in\Theta_1} \Pi(\delta, \theta) = \sup_{\theta\in\Theta_1} \mathbb{P}_\theta(\delta \neq H_1)$$

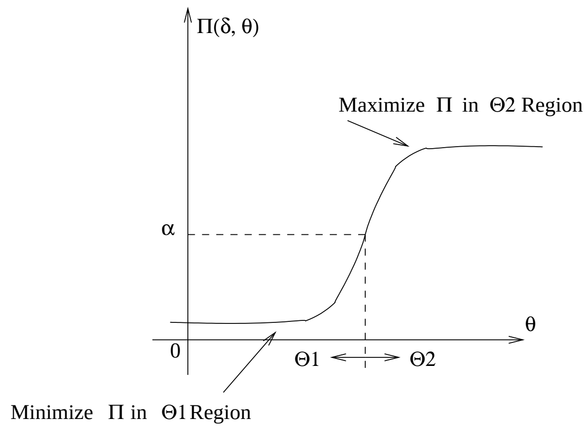

Figure 20.2: Power function.

which is called the *size* of δ and which represents the largest possible error of type 1. As in the case of simple hypotheses it often makes sense to control this largest possible error of type one by some level of significance $\alpha \in [0, 1]$ and to consider decision rules from the class

$$K_\alpha = \{\delta; \alpha_1(\delta) \leq \alpha\}.$$

Then, of course, we would like to find the decision rule in this class that also maximizes the power function on the set Θ_2, i.e. miminizes the errors of type 2. In general, the decision rules $\delta, \delta' \in K_\alpha$ may be incomparable, because in some regions of Θ_2 we might have $\Pi(\delta, \theta) > \Pi(\delta', \theta)$ and in other regions $\Pi(\delta', \theta) > \Pi(\delta, \theta)$. Therefore, in general, it may be impossible to maximize the power function for all $\theta \in \Theta_2$ simultaneously. But, as we will show in the next lecture, under certain conditions it may be possible to find the best test in class K_α that is called the *uniformly most powerful test.*

Definition. If we can find $\delta \in K_\alpha$ such that

$$\Pi(\delta, \theta) \geq \Pi(\delta', \theta) \text{ for all } \theta \in \Theta_2 \text{ and all } \delta' \in K_\alpha$$

then δ is called the Uniformly Most Powerful (UMP) test.