LECTURE-28

Agenda:
1. Recap: Power of a test
2. Simple and Composite Hypotheses
3. Neyman-Pearson Lemma and the Most Powerful Test

Recall that for testing $H_0$ v.s. $H_A$, which are two hypotheses concerning an unknown parameter $\theta$, we had defined the "power" of a testing procedure at a specific alternative $\theta_A$ as:

$$\text{Power} \left( \theta_A \right) = 1 - \beta(\theta_A)$$

Type II Error probability at $\theta_A$

$$= 1 - P(\text{H}_A \text{ is rejected} | \theta = \theta_A)$$

$$= P(\text{H}_A \text{ is accepted} | \theta = \theta_A)$$

Here, the power at the alternative $\theta_A$ is the probability of correctly accepting $H_A$ when the true value of $\theta$ is $\theta_A$.

So, we would like the power to be as large as possible.
**GENERAL QUESTION:** For a general testing problem, can we find a testing procedure which has the highest power (assuming the level is fixed at, say, \( \alpha \))? 

**ANSWER:** In various situations, we can.

Before we proceed, let us define the concepts of **SIMPLE** and **COMPOSITE** hypothesis.

**DEFINITION:** Suppose we have data from a population with unknown parameter \( \theta \). A hypothesis is said to be **SIMPLE** if it uniquely specifies the distribution of the population from which the sample is taken. [In most cases, this is same as saying, that the hypotheses uniquely identifies the value of \( \theta \).] (But not always)

Any hypothesis that is not a SIMPLE hypothesis is called a **COMPOSITE** hypothesis.

The Neyman-Pearson Lemma provides a method to obtain the **MOST POWERFUL** test among all tests with level-\( \alpha \), if both \( H_0 \) and \( H_a \) are **SIMPLE** hypotheses.
THE NETMAN-PEARSON LEMMA: Suppose that we wish to test the simple null hypothesis $H_0: \theta = \theta_0$ versus the simple alternative hypothesis $H_A: \theta = \theta_1$, based on a random sample $Y_1, Y_2, \ldots, Y_n$ from a distribution with parameter $\theta$. Recall that $L(\theta)$ denotes the likelihood of the sample when the value of the parameter is $\theta$. Then, the test that maximizes the power at $\theta_1$ has a rejection region determined by
\[
\frac{L(\theta_0)}{L(\theta_1)} < k.
\]

The value of $k$ is chosen so that the level is equal to $\alpha$. Such a test is called the MOST POWERFUL LEVEL-$\alpha$ TEST for $H_0$ vs. $H_A$.

Example: Suppose $Y$ is a sample of size $1$ from a population with density
\[
f_Y(y|\theta) = \begin{cases} \theta y^{\theta-1} & \text{if } 0 < y < 1, \\ 0 & \text{otherwise.} \end{cases}
\]

Find the most powerful test with level $\alpha = 0.05$ for testing $H_0: \theta = 2$ vs. $H_A: \theta = 1$.
Note that both $H_0$ and $H_1$ are simply specifying the value of $\theta$; completely specifies the density of the population. Hence, by the Neyman-Pearson lemma, the most powerful test for $H_0$ vs. $H_1$ has a rejection region of the form

$$\frac{L(\theta_0)}{L(\theta_1)} > k,$$

where $k$ is chosen so that level $= 0.05$.

Note that $L(\theta_0) = f(y|\theta = 2) = 2y$, and $L(\theta_1) = f(y|\theta = 1) = 1$, for $0 < y < 1$. Hence, the rejection region is of the form

$$2y > k \quad \text{or} \quad y < \frac{k}{2}.$$

Let us calculate the value of $k$.

level $= 0.05 \Rightarrow P\left( y < \frac{k}{2} \left| \theta = 2 \right. \right) = 0.05$

$$\Rightarrow \int_0^{k/2} 2y \, dy = 0.05$$

$$\Rightarrow \left[ y^2 \right]_0^{k/2} = 0.05$$

$$\Rightarrow \left( \frac{k}{2} \right)^2 = 0.05$$

$$\Rightarrow \frac{k^2}{4} = 0.05$$

$$\Rightarrow k^2 = 0.2$$

$$\Rightarrow k = \sqrt{0.2236}.$$

Hence the rejection region is given by $\{ y < 0.2236 \}$. 