LECTURE - 30

<u>Agenda:</u>

1. Likelihood Ratio Tests
2. Examples

As we noticed in the last two lectures, the Neyman-Pearson lemma is useful only when there is ONLY ONE UNKNOWN PARAMETER IN THE POPULATION. But more often than not, there are several unknown parameters in the population. Let us take a very simple example where we have a random sample $Y_1, Y_2, \ldots, Y_n$ from a Normal(μ, σ^2) population, where μ and σ^2 are both unknown. Suppose we want to test

$$H_0: \mu = \mu_0 \quad \text{v.s.} \quad H_A: \mu > \mu_0.$$

Note that neither the null hypothesis nor the alternative hypothesis are simple hypotheses in this case, because they do not specify the value of σ^2. Hence, σ^2 is an unknown parameter which is not involved in the two hypotheses. Such parameters are known as NUISANCE PARAMETERS.

IS THERE A STANDARD TESTING PROCEDURE FOR TESTING HYPOTHESES IN THE PRESENCE OF A NUISANCE PARAMETER?

<u>GENERAL PROCEDURE:</u> Suppose there are k unknown parameters $\theta_1, \theta_2, ---, \theta_k$ in the population. Let

$\Theta = (\theta_1, \theta_2, ---, \theta_k)$, i.e., Θ is the vector of unknown parameters. Suppose we want to test

$$H_0: \Theta \in \Omega_0 \quad \text{v.s.} \quad H_A: \Theta \in \Omega_A.$$

Recall that $L(\Theta)$ denotes the likelihood of the data at the parameter value Θ.

LIKELIHOOD RATIO BASED TESTING PROCEDURE FOR OBTAINING LEVEL-α TEST

TEST STATISTIC: $$\lambda \triangleq \frac{\max_{\Theta \in \Omega_0} L(\Theta)}{\max_{\Theta \in \Omega_0 \cup \Omega_A} L(\Theta)}.$$

REJECTION REGION: $\{\lambda \leq k\}$, where k is chosen so that the level of the test is α.

__Example:__ Suppose $Y_1, Y_2, ---, Y_n$ is a random sample from a population which is Normal (μ, σ^2), with μ and σ^2 unknown. Suppose we want to test

$$H_0 : \mu = 0 \quad \text{v.s.} \quad H_1 : \mu > 0 .$$

Provide the likelihood ratio testing procedure for this testing problem.

Note that for this testing problem, $\Theta = (\mu, \sigma^2)$.

$$\Omega_0 = \{ (\mu, \sigma^2) : \mu = 0, \sigma^2 > 0 \}, \quad \Omega_A = \{ (\mu, \sigma^2) : \mu > 0, \sigma^2 > 0 \}$$

__STEP I:__ Find $\max\limits_{\Theta \in \Omega_0} L(\Theta)$.

Note that
$$L(\Theta) = \prod_{i=1}^{n} f_{Y_i}(y_i \mid \Theta)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_i-\mu)^2}{2\sigma^2}}$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^n e^{-\frac{\Sigma (y_i-\mu)^2}{2\sigma^2}}$$

Hence,

$$\max_{\Theta \in \Omega_0} L(\Theta) = \max_{\mu=0, \sigma^2>0} \frac{1}{(\sqrt{2\pi\sigma^2})^n} e^{-\frac{\Sigma(y_i-\mu)^2}{2\sigma^2}}$$

$$= \max_{\sigma^2>0} \frac{1}{(\sqrt{2\pi\sigma^2})^n} e^{-\frac{\Sigma y_i^2}{2\sigma^2}}.$$

We have seen in Exercises during MLE that the value of σ^2 that maximizes the above function is

$$\hat{\sigma}_0^2 = \frac{1}{n}\sum_{i=1}^{n} y_i^2.$$

Hence,

$$\max_{\Theta \in \Omega_0} L(\Theta) = \frac{1}{(\sqrt{2\pi\hat{\sigma}_0^2})^n} e^{-\frac{\Sigma y_i^2}{2\hat{\sigma}_0^2}}$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n \cdot \frac{1}{\hat{\sigma}_0^n} e^{-\frac{n}{2}}.$$

STEP II: Find $\max_{\Theta \in \Omega_0 \cup \Omega_A} L(\Theta)$

Note that,

$$\Omega_0 \uplus \Omega_A = \{(\mu, \sigma^2): \mu \geq 0, \sigma^2 > 0\}$$

From the equations involving $\frac{\partial}{\partial \sigma^2} \log L(\Theta)$ it follows

that $$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{\mu})^2 .$$

Hence, $$\max_{\Theta \in \Omega_0 \cup \Omega_a} L(\Theta) = L(\hat{\mu}, \hat{\sigma}^2)$$

$$= \frac{1}{(\sqrt{2\pi})^n} \left(\frac{1}{\hat{\sigma}^2} \right)^{n/2} e^{-n/2} .$$

Hence, the rejection region of the likelihood ratio test is given by

$$\lambda = \frac{\frac{1}{(\sqrt{2\pi})^n} \frac{1}{\hat{\sigma}_0^n} e^{-n/2}}{\frac{1}{(\sqrt{2\pi})^n} \frac{1}{\hat{\sigma}^n} e^{-n/2}} \leq k$$

$$\Leftrightarrow \quad \left(\frac{\hat{\sigma}}{\hat{\sigma}_0} \right)^n \leq k .$$

<u>STEP III:</u> Find the value of k using the level-α condition.

We will skip this computation, but just note that after simplification, k can be found as a function of an appropriate t-quantile.

Hence, we need to maximize $L(\Theta)$ or equivalently

$$\log L(\Theta) = -\frac{n}{2}\log\sigma^2 - \frac{n}{2}\log 2\pi - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2$$

over the space $\{\mu \geq 0, \sigma^2 > 0\}$.

Note that,

$$\frac{\partial}{\partial\mu}\log L(\Theta) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i-\mu) = \frac{n}{\sigma^2}(\bar{y}-\mu)$$

$$\frac{\partial}{\partial\sigma^2}\log L(\Theta) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(y_i-\mu)^2$$

Note that,

$$\frac{\partial}{\partial\mu}\log L(\Theta) \begin{cases} < 0 & \text{if } \mu > \bar{y} \\ = 0 & \text{if } \mu = \bar{y} \\ > 0 & \text{if } \mu < \bar{y} \end{cases}$$

Hence, $\log L(\mu, \sigma^2)$ is maximized at

$$\hat{\mu} = \begin{cases} \bar{y} & \text{if } \bar{y} > 0, \\ 0 & \text{if } \bar{y} \leq 0. \end{cases}$$