Lecture 21

21.1 Monotone likelihood ratio.

In the last lecture we gave the definition of the UMP test and mentioned that under certain conditions the UMP test exists. In this section we will describe a property called monotone likelihood ratio which will be used in the next section to find the UMP test for one sided hypotheses.

Suppose the parameter set $\Theta \subseteq \mathbb{R}$ is a subset of a real line and that probability distributions $\mathbb{P}_\theta$ have p.d.f. or p.f. $f(x|\theta)$. Given a sample $X = (X_1, \ldots, X_n)$, the likelihood function (or joint p.d.f.) is given by

$$f(X|\theta) = \prod_{i=1}^{n} f(X_i|\theta).$$

Definition: The set of distributions $\{\mathbb{P}_\theta, \theta \in \Theta\}$ has Monotone Likelihood Ratio (MLR) if we can represent the likelihood ratio as

$$\frac{f(X|\theta_1)}{f(X|\theta_2)} = V(T(X), \theta_1, \theta_2)$$

and for $\theta_1 > \theta_2$ the function $V(T, \theta_1, \theta_2)$ is strictly increasing in T.

Example. Consider a family of Bernoulli distributions $\{B(p) : p \in [0, 1]\}$, in which case the p.f. is qiven by

$$f(x|p) = p^x(1-p)^{1-x}$$

and for $X = (X_1, \ldots, X_n)$ the likelihood function is

$$f(X|p) = p^{\sum X_i}(1-p)^{n-\sum X_i}.$$

We can write the likelihood ratio as folows:

$$\frac{f(X|p_1)}{f(X|p_2)} = \frac{p_1^{\sum X_i}(1-p_1)^{n-\sum X_i}}{p_2^{\sum X_i}(1-p_2)^{n-\sum X_i}} = \Big(\frac{1-p_1}{1-p_2}\Big)^n \Big(\frac{p_1(1-p_2)}{p_2(1-p_1)}\Big)^{\sum X_i}.$$

For $p_1 > p_2$ we have

$$\frac{p_1(1-p_2)}{p_2(1-p_1)} > 1$$

and, therefore, the likelihood ratio is strictly increasing in $T = \sum_{i=1}^{n} X_i$.

Example. Consider a family of normal distributions $\{N(\mu, 1) : \mu \in \mathbb{R}\}$ with variance $\sigma^2 = 1$ and unknown mean μ as a parameter. Then the p.d.f. is

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

and the likelihood

$$f(X|\mu) = \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2}\sum_{i=1}^{n}(X_i-\mu)^2}.$$

Then the likelihood ratio can be written as

$$\frac{f(X|\mu_1)}{f(X|\mu_2)} = e^{-\frac{1}{2}\sum_{i=1}^{n}(X_i-\mu_1)^2 + \frac{1}{2}\sum_{i=1}^{n}(X_i-\mu_2)^2} = e^{(\mu_1-\mu_2)\sum X_i - \frac{n}{2}(\mu_1^2-\mu_2^2)}.$$

For $\mu_1 > \mu_2$ the likelihood ratio is increasing in $T(X) = \sum_{i=1}^{n} X_i$ and MLR holds.

21.2 One sided hypotheses.

Consider $\theta_0 \in \Theta \subseteq \mathbb{R}$ and consider the following hypotheses:

$$H_1 : \theta \leq \theta_0 \text{ and } H_2 : \theta > \theta_0$$

which are called *one sided hypotheses,* because we hypothesize that the unknown parameter θ is on one side or the other side of some threshold θ_0. We will show next that if MLR holds then for these hypotheses there exists a Uniformly Most Powerful test with level of significance α, i.e. in class K_α.

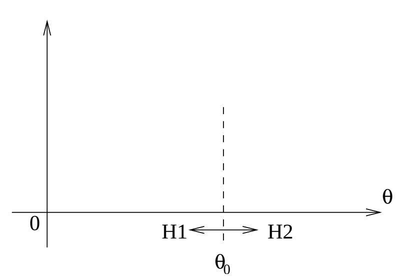

Figure 21.1: One sided hypotheses.

Theorem. *Suppose that we have Monotone Likelihood Ratio with $T = T(X)$ and we consider one-sided hypotheses as above. For any level of significance $\alpha \in [0, 1]$, we can find $c \in \mathbb{R}$ and $p \in [0, 1]$ such that*

$$\mathbb{P}_{\theta_0}(T(X) > c) + (1 - p)\mathbb{P}_{\theta_0}(T(X) = c) = \alpha.$$

Then the following test δ^* will be the Uniformly Most Powerful test with level of significance α:

$$\delta^* = \begin{cases} H_1 : & T < c \\ H_2 : & T > c \\ H_1 \text{ or } H_2 : & T = c \end{cases}$$

where in the last case of $T = c$ we randomly pick H_1 with probability p and H_2 with probability $1 - p$.

Proof. We have to prove two things about this test δ^*:

1. $\delta^* \in K_\alpha$, i.e. δ^* has level of significance α,
2. for any $\delta \in K_\alpha$, $\Pi(\delta^*, \theta) \geq \Pi(\delta, \theta)$ for $\theta > \theta_0$, i.e. δ^* is more powerful on the second hypothesis that any other test from the class K_α.

To simplify our considerations below let us assume that we don't need to randomize in δ^*, i.e. we can take $p = 1$ and we have

$$\mathbb{P}_{\theta_0}(T(X) > c) = \alpha$$

and the test δ^* is given by

$$\delta^* = \begin{cases} H_1 : & T \leq c \\ H_2 : & T > c. \end{cases}$$

Proof of 1. To prove that $\delta^* \in K_\alpha$ we need to show that

$$\Pi(\delta^*, \theta) = \mathbb{P}_\theta(T > c) \leq \alpha \text{ for } \theta \leq \theta_0.$$

Let us for a second forget about composite hypotheses and for $\theta < \theta_0$ consider two simple hypotheses:

$$h_1 : \mathbb{P} = \mathbb{P}_\theta \text{ and } h_2 : \mathbb{P} = \mathbb{P}_{\theta_0}.$$

For these simple hypotheses let us find the most powerful test with error of type 1 equal to

$$\alpha_1 := \mathbb{P}_\theta(T > c).$$

We know that if we can find a threshold b such that

$$\mathbb{P}_\theta\Big(\frac{f(X|\theta)}{f(X|\theta_0)} < b\Big) = \alpha_1$$

then the following test will be the most powerful test with error of type one equal to α_1:

$$\delta_\theta = \begin{cases} h_1 : & \frac{f(X|\theta)}{f(X|\theta_0)} \geq b \\ h_2 : & \frac{f(X|\theta)}{f(X|\theta_0)} < b \end{cases}$$

This corresponds to the situation when we do not have to randomize. But the monotone likelihood ratio implies that

$$\frac{f(X|\theta)}{f(X|\theta_0)} < b \Leftrightarrow \frac{f(X|\theta_0)}{f(X|\theta)} > \frac{1}{b} \Leftrightarrow V(T, \theta_0, \theta) > \frac{1}{b}$$

and, since $\theta_0 > \theta$, this last function $V(T, \theta_0, \theta)$ is strictly increasing in T. Therefore, we can solve this inequality for T (see figure 21.2) and get that $T > c_b$ for some c_b.

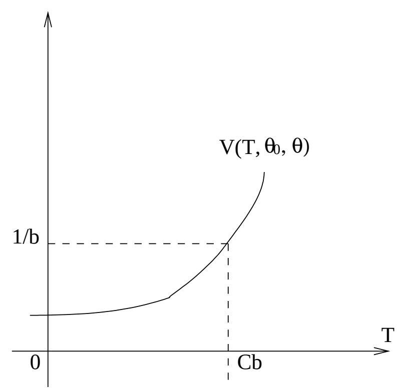

Figure 21.2: Solving for T.

This means that the error of type 1 for the test δ_θ can be written as

$$\alpha_1 = \mathbb{P}_\theta\Big(\frac{f(X|\theta)}{f(X|\theta_0)} < b\Big) = \mathbb{P}_\theta(T > c_b).$$

But we chose this error to be equal to $\alpha_1 = \mathbb{P}_\theta(T > c)$ which means that c_b should be such that

$$\mathbb{P}_\theta(T > c_b) = \mathbb{P}_\theta(T > c) \Rightarrow c = c_b.$$

To summarize, we proved that the test

$$\delta_\theta = \begin{cases} h_1 : & T \leq c \\ h_2 : & T > c \end{cases}$$

is the most powerful test with error of type 1 equal to

$$\alpha_1 = \Pi(\delta^*, \theta) = \mathbb{P}_\theta(T > c).$$

Let us compare this test δ_θ with *completely randomized test*

$$\delta^{\text{rand}} = \begin{cases} h_1: & \text{with probability } 1-\alpha_1 \\ h_2: & \text{with probability } \alpha_1, \end{cases}$$

which picks hypotheses completely randomly regardless of the data. The error of type one for this test will be equal to

$$\mathbb{P}_\theta(\delta^{\text{rand}} = h_2) = \alpha_1,$$

i.e. both tests δ_θ and δ^{rand} have the same error of type one equal to α_1. But since δ_θ is the most powerful test it has larger power than δ^{rand}. But the power of δ_θ is equal to

$$\mathbb{P}_{\theta_0}(T > c) = \alpha$$

and the power of δ^{rand} is equal to

$$\alpha_1 = \mathbb{P}_\theta(T > c).$$

Therefore,

$$\mathbb{P}_\theta(T > c) \le \mathbb{P}_{\theta_0}(T > c) = \alpha$$

and we proved that for any $\theta \le \theta_0$ the power function $\Pi(\delta^*, \theta) \le \alpha$ which this proves 1.