

Statistical Inference
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture No. # 24
UMP Tests

So far the testing procedure that we have discussed was based on the Neyman Pearson fundamental lemma. The main assumption that we made in deriving the test procedure was that the null hypothesis and the alternative hypothesis both were considered to be simple; and in this case, when we fix the probability of type one error, then we were able to derive the test which is having the minimum probability of type two error or the maximum power, and we called it the most powerful test. However, in most of the real life situations, we do not come across the simple hypothesis versus simple hypothesis problems. In most of the complex situations, we have composite hypothesis.

(Refer Slide Time: 01:06)

Lecture 24

Families with Monotone Likelihood Ratio

$N(\mu, \sigma^2)$ both parameters are unknown

$H_0: \mu = 0$
 $H_1: \mu \neq 0$ } composite

$f(\theta)$ $f(x, \theta)$

$H_0: \theta \leq \theta_0$ $H_0: \theta \geq \theta_0$
 $H_1: \theta > \theta_0$ $H_1: \theta < \theta_0$

Let $f(x, \theta)$ be a prob. m.f (d.f.) of a r. v. X .

$r(x) = \frac{f(x, \theta_1)}{f(x, \theta_2)}$, $\theta_1 > \theta_2$.

If $r(x)$ is an increasing fn. of $T(x)$, we say that the family of

As a very simple case, we may have the family of distributions as normal mu sigma square distributions, and we may like to even now we may like to test something like, whether mu naught mu is equal to zero or mu is not equal to zero. Note here that, now

this H_0 is not a simple hypothesis; this is composite, because σ^2 is unknown. Here we have assumed both parameters to be unknown; both parameters are unknown. Therefore these are now composite both hypothesis are composite hypothesis. And therefore, the Neyman Pearson lemma does not help us to give a solution in this particular problem; that means, does not give a most powerful test.

The simplest composite hypothesis are of this nature, that we may have a one parameter family, say family of distribution with one parameter θ say effects θ . And we may like to test about, say $H_0: \theta \leq \theta_0$ against say $H_1: \theta > \theta_0$, or alternative we may have say $H_0: \theta \geq \theta_0$ against $H_1: \theta < \theta_0$.

Now, let us remember our cases that some examples we consider for the Neyman Pearson lemma, where we had considered $\theta = \theta_0$ against $\theta = \theta_1$. I had consider two cases; one was $\theta < \theta_1$, and another was $\theta > \theta_1$. When $\theta < \theta_1$, we got a one sided testing region; that is the rejection region, that is for larger values of \bar{x} we were rejecting H_0 .

Now, in that problem, in place of θ_1 suppose we replaced by another value θ_2 , suppose we replace by another value θ_3 , the testing procedure remains the same as long as this second value in the alternative hypothesis remains larger than θ_0 . In a similar way if we are considering the reverse case $\theta > \theta_1$, then the rejection region was for smaller values of \bar{x} , and once again if we replace this alternative hypothesis θ_1 in the same direction; that means, value which is greater than θ_0 or less than θ_0 , then the rejection region does not get effected. What does it mean? It means, that for those values we are getting the most powerful tests; that means, this normal distribution with one parameter, the second parameter σ was considered to be known has certain property.

Now, in these situations for the changing values we get the maximum power at each of the values, this is called uniformly most powerful test. Now, this family distributions which will satisfy this property; that means, where we will get such test; it is having some particular name, it is call the family's with monotone likelihood ratio property. In particular, for the one sided testing of hypothesis problems like $\theta \leq \theta_0$ are equal

to theta naught against theta greater than theta naught, or theta greater than or equal to theta naught against theta less than theta naught etcetera, For such cases we are actually getting the uniformly most powerful test. The result that are proved, their actually you can say they are extension of the Neyman Pearson fundamental lemma.

So Firstly, let me define this family's, so let $f(x, \theta)$ be a probability mass function or density function of a random variable say x , let us write down the ratio $f(x, \theta_1)$ divided by $f(x, \theta_2)$, let us call this name; this ratio let me call it $r(x)$. And let us takes a θ_1 greater than θ_2 , if $r(x)$ is an increasing function of some variables say $T(x)$, then we say that the family of densities.

(Refer Slide Time: 05:58)

densities $\{f(x, \theta) : \theta \in \Omega\}$ has monotone likelihood ratio (MLR) in $(\theta, T(x))$.

Examples: 1. $X \sim N(\theta, 1)$
 $f(x, \theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\theta)^2}$

$r(x) = \frac{f(x, \theta_1)}{f(x, \theta_2)} = e^{-\frac{1}{2}(x-\theta_1)^2 + \frac{1}{2}(x-\theta_2)^2} = e^{\frac{1}{2}(\theta_2^2 - \theta_1^2) + (\theta_1 - \theta_2)x}$
 is an increasing fn. of x (if $\theta_1 > \theta_2$).

So $\{N(\theta, 1) : \theta \in \mathbb{R}\}$ has MLR in (θ, x) . $\bar{x} = \frac{1}{n} \sum x_i$

$X_1, \dots, X_n \sim N(\theta, 1)$
 The joint density $f_{\underline{X}}(x_1, \dots, x_n)$ is
 $f(\underline{x}, \theta) = \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2} \sum (x_i - \theta)^2} = \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2} \sum x_i^2 - \frac{n\theta^2}{2} + n\bar{x}\theta}$

The word densities means, it includes the probability mass functions. So, that is $f(x, \theta)$, θ belonging to the parameter is phase has monotone likelihood ratio, that we call MLR in θ $T(x)$. Let me given example here, let us consider say x following a normal distribution with mean θ and known variants 1, let us write down the distribution $f(x, \theta)$ is $\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\theta)^2}$. Let us consider this ratio $r(x)$, that is $f(x, \theta_1)$ divided by $f(x, \theta_2)$, now when we write this ratio this gets cancelled out, and you have e to the power minus half x minus θ_1 square plus half x minus θ_2 square, that is equal to e to the power half θ_2 square minus θ_1 square, and then you will have plus θ_1 minus θ_2 x .

So, you can look at this, this is an increasing function if I am taking increasing function of x , if θ_1 is greater than θ_2 because this is constant, and if θ_1 is greater than θ_2 e to the power this becomes an increasing function of x . So, this family of distributions normal θ_1 ; where θ belongs to real line, this has monotone likelihood ratio in θ and x .

Now, I have return here the distribution of one observation, suppose in place of x I have $x_1 \times 2 \times n$. **suppose I have $x_1 \times 2 \times n$** In this case, $f(x|\theta)$ we have to write the joint distribution of $x_1 \times 2 \times n$, so the joint density of $x_1 \times 2 \times n$. So, let me give the notation $f(x)$, where x is standing for the values $x_1 \times 2 \times n$ of $X_1 \times 2 \times n$. So, this becomes $\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x - \theta)^2}$.

Let us simplify this, we can write it as $\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x - \theta)^2}$, now you have the cross product term twice $x - \theta$ with minus sign, and minus minus will become plus and 2 will cancel out. So, you get twice $n\bar{x}\theta$, where \bar{x} is the $\frac{1}{n} \sum x_i$.

(Refer Slide Time: 10:03)

Handwritten notes on a blue background:

$$\frac{f(x|\theta_1)}{f(x|\theta_2)} = e^{\frac{n}{2}(\theta_2^2 - \theta_1^2) + n\bar{x}(\theta_1 - \theta_2)}$$

This is an increasing fn. of $T(x) = \bar{x}$ when $\theta_1 > \theta_2$.
So f has MLR in (θ, \bar{X}) .

2. $X \sim N(0, \sigma^2)$, $\sigma^2 > 0$
 $f(x, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$, $x \in \mathbb{R}$

$$\frac{f(x, \sigma_1^2)}{f(x, \sigma_2^2)} = \frac{\sigma_2}{\sigma_1} e^{\frac{x^2}{2} \left(\frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2} \right)}$$

\uparrow fn. of x^2 . $\sigma_1^2 > \sigma_2^2 \Rightarrow \frac{1}{\sigma_1^2} < \frac{1}{\sigma_2^2}$

$\{N(0, \sigma^2) : \sigma^2 > 0\}$ has MLR in (σ^2, x^2)

So, now you write down this ratio $f(x|\theta_1)$ divided by $f(x|\theta_2)$, that is turning out to be, now, when you write the ratio, this constant determinate get cancelled out e to the power minus half sigma x i square will get cancelled out, we will be left with e to the power n by 2 θ_2^2 minus θ_1^2 plus $n\bar{x}\theta_1$ minus θ_2 . Now,

this is constant, for θ_1 greater than θ_2 this becomes an increasing function of \bar{x} , so this ratio and increasing function of $T\bar{x}$ is equal to \bar{x} when θ_1 is greater than θ_2 . So, this family of distribution normal θ_1 , when we are having n observations, so we have MLR in θ_1 and \bar{x} we can say. Now, the similar thing we can observe for various distributions; let me give a couple of more examples, here I have considered the normal distribution when the variances assume to be known. Now, there can be a other case where mean may be known and the variance may be unknown, let us state that case, let me again consider say 1 observation and then I will consider n observations, generally we are dealing with the sample.

So, let me taking this case, here σ^2 is positive parameter, if we consider the density function here; it is $\frac{1}{\sigma\sqrt{2\pi}}$ into the power minus x^2 by $2\sigma^2$ is square, where x is any real number. Therefore, if I consider the ratio $f(x, \sigma_1^2)$ divided by $f(x, \sigma_2^2)$, now this when will give me σ_2^2 by σ_1^2 , this $\frac{1}{\sigma_1\sqrt{2\pi}}$ will get cancelled out to the power x^2 by $2\sigma_1^2$ minus $\frac{1}{\sigma_2\sqrt{2\pi}}$ by σ_2^2 minus x^2 by $2\sigma_2^2$. Now, let us takes a σ_1^2 greater than σ_2^2 ; that means, $\frac{1}{\sigma_1^2}$ is less than $\frac{1}{\sigma_2^2}$. So, this term because positive and therefore, this is increasing function of x^2 , so this family of normal zero σ^2 distributions, this has monotone likelihood ratio in $\sigma^2 x^2$.

(Refer Slide Time: 13:16)

Let $X_1, \dots, X_n \sim N(0, \sigma^2)$
 The joint density of X_1, \dots, X_n is $\frac{\sigma^{-2n}}{(2\pi)^{n/2}}$
 $f(x, \sigma^2) = \frac{1}{(\sigma\sqrt{2\pi})^n} e^{-\frac{\sum x_i^2}{2\sigma^2}}$, $x_i \in \mathbb{R}$, $\sigma^2 > 0$
 $\frac{f(x, \sigma_1^2)}{f(x, \sigma_2^2)} = \left(\frac{\sigma_2}{\sigma_1}\right)^n e^{-\frac{\sum x_i^2}{2} \left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right)}$, $\sigma_1^2 > \sigma_2^2$
 \uparrow in $\sum x_i^2 = T(x)$
 So MLR in $(\sigma^2, \sum X_i^2)$.

Now, not here, that suppose I take a sample here in place of x , let us take sample $x_1 \times 2 \times n$ following normal zero σ^2 , and let us write down the same thing of once again. The joint distribution; the joint density of $x_1 \times 2 \times n$, that will become 1 by σ root 2π to the power then e to the power minus $\sum x_i^2$ by $2 \times n$ square, where σ^2 is positive and each x_i 's on the real line. So, when we write down the ratio, now this term get cancel out will get σ^2 by σ^2 to the power n e to the power minus $\sum x_i^2$ by $2 \times n$ square minus 1 by σ^2 square, so this i will put plus here. Once again, you not here this is positive if σ_1^2 is less than σ_2^2 **Sorry**. Now, if σ_1^2 is greater than σ_2^2 is square, then this term becomes positive, so this is increasing in $\sum x_i^2$, that we will call T_x . So, this family has monotone likelihood ratio in σ^2 and $\sum x_i^2$. Now, this T_x has a special role, then we will derive the uniformly most powerful test; you will see that the test will depend upon this itself. So, I will discuss a few more applications little later; let us look at the main result of this section of now, that is as an application of the monotone likelihood ratio property. How the uniformly most powerful test exist?

(Refer Slide Time: 15:06)

Theorem (Lehmann & Romano, 2005, Rohatgi's Selah.)

Let the r.v. X have pmf (pdf) $f(x, \theta)$ with MLR in $(\theta, T(x))$.
 $\theta \in \Theta \subseteq \mathbb{R}$.

(i) For testing $H_0: \theta \leq \theta_0$ against $H_1: \theta > \theta_0$, there exists a Uniformly most powerful (UMP) test, given by

$$\phi(x) = \begin{cases} 1 & \text{if } T(x) > c \\ \gamma & \text{if } T(x) = c \\ 0 & \text{if } T(x) < c \end{cases} \quad \dots (1)$$

where c & γ are determined by

$$E_{\theta_0} \phi(X) = \alpha \quad \dots (2)$$

(ii) The power function $\beta^*(\theta) = E_{\theta} \phi(X)$ is strictly increasing for all points θ for which $0 < \beta^*(\theta) < 1$.

So, I state the theorem; for a proper statement of this theorem you may look at the books of Lehmann and Romano 2005, or you may look at Rohatgi and Selah, the proofs are also given there. So, I am not discussing the proof here.

So, let us consider, let the random variable x have probability mass function or probability density function, say $f(x; \theta)$ with monotone likelihood ratio in $T(x)$, And of course, here θ is a real parameter; θ belongs to say Θ which is subset of the real line.

So, the result that we are having here is that, for testing $H_0: \theta \leq \theta_0$ against $H_1: \theta > \theta_0$, there exists a uniformly most powerful test; that is, UMP test given by as before we will use $\phi(x)$ notation for the test function. So, you reject if $T(x)$ is greater than c , you reject with probability γ if $T(x)$ is equal to c , and you accept if $T(x) < c$. Where c and this γ determined by expectation of $T(x)$ is equal to α , let me call this conditions one and two, note here that similarity with Neyman Pearson lemma, in the Neyman Pearson lemma we had return f_1 by f_0 greater than k . Now, if f_1 by f_0 is an increasing function of $T(x)$, then that region is transformed to $T(x) > c$. So, it is as I mentioned, it is direct extension of the Neyman Pearson fundamental lemma only, the result is coming from there. The power function; that is, we have use the notation say $\beta(\theta)$, that is equal to expectation $E_\theta[\phi(x)]$ is strictly increasing for all points θ for which it lies between zero and 1.

(Refer Slide Time: 19:04)

(iii) For all θ^* , the test determined by (1) & (2) is UMP for testing $H_0: \theta \leq \theta^*$ against $H_1: \theta > \theta^*$ at level $\alpha = \beta^*(\theta^*)$.

(iv) For any $\theta < \theta_0$ the test minimized $\beta^*(\theta)$ among all tests satisfying (2).

Remark: If we consider the dual problem $H_0: \theta \geq \theta_0$, $H_1: \theta < \theta_0$, the inequalities in (1) get reversed.

Example: $X_1, \dots, X_n \sim P(\lambda)$, $\lambda > 0$.
 $H_0: \lambda \leq \lambda_0$
 $H_1: \lambda > \lambda_0$.

The joint pmf of X_1, \dots, X_n is

$$f(x; \lambda) = \prod_{i=1}^n \frac{e^{-\lambda} \lambda^{x_i}}{x_i!} = \frac{e^{-n\lambda} \cdot \lambda^{\sum x_i}}{n!}$$

For all θ , the test determined by 1 and 2 is uniformly most powerful for testing $H_0: \theta \leq \theta_0$ against $H_1: \theta > \theta_0$.

prime at level say α prime is equal to, let me put star here, this μ hypothesis I am calling H_0 and H_1 star at α star is equal to β star of θ star. And for $n \rightarrow \infty$ $\theta < \theta_0$, the test minimizes β star θ among all tests satisfying the condition 2. I will skip the proof here; one can look at the book of Lehmann for the detailed proof of this is statements.

Now, one may not here, I have considered $\theta < \theta_0$ against $\theta > \theta_0$. As I gave the heuristic argument, that in the Neyman Pearson lemma, and as I also I give the normal distribution example when we were testing for the mean, the rejection region for the larger value of \bar{x} , and here it is for the larger value of $T(x)$. So, if we reverse, like for the null hypothesis region we consider greater, and for the null alternative hypothesis we consider less than or equal to, then the rejection region will also get the reverse. So, what I just give it as a comment here, if we considered the say dual problem $H_0: \theta > \theta_0$ against $H_1: \theta < \theta_0$, the inequalities in 1 get reversed. So, you have the solution in a similar manner way.

Let me take an application here, say we have a random sample say X_1, X_2, \dots, X_n from say poisson λ distribution, and we consider say hypothesis $\lambda < \lambda_0$ against say $\lambda > \lambda_0$. Now, let us look at this family of poisson distributions, whether it as monotone likelihood ratio or not. So, the joint probability mass function of X_1, X_2, \dots, X_n , so we write it as $f(x; \lambda) = \prod_{i=1}^n \frac{e^{-\lambda} \lambda^{x_i}}{x_i!}$, that is equal to $e^{-n\lambda} \lambda^{\sum x_i} / \prod x_i!$.

(Refer Slide Time: 22:59)

For $\lambda_1 > \lambda_2$, $r(x) = \frac{f(x, \lambda_1)}{f(x, \lambda_2)} = \frac{e^{-n\lambda_1} \lambda_1^{\sum x_i}}{e^{-n\lambda_2} \lambda_2^{\sum x_i}}$

$= e^{n(\lambda_2 - \lambda_1)} \left(\frac{\lambda_1}{\lambda_2}\right)^{\sum x_i}$. This is an increasing fn. in $T(x) = \sum x_i$.

So MLR in $(\lambda, \sum X_i)$

So the UMP test is given by

$$\phi(x) = \begin{cases} 1 & \text{if } \sum x_i > c \\ \gamma & \text{if } \sum x_i = c \\ 0 & \text{if } \sum x_i < c \end{cases} \quad (1)$$

c & γ are determined by $E_{\lambda} \phi(X) = \alpha$ (2)

$Y = \sum X_i \sim P(n, \lambda)$

So, if we consider the ratio for lambda 1 greater than lambda 2, let us consider the ratio f x lambda 1 by f x lambda 2. So, it is becoming e to the power minus n lambda 1 lambda 1 to the power sigma x i divided by product x i factorial divided by e to the power minus n lambda 2 lambda 2 to the power sigma x i, and this term will get cancelled out.

So, we can write it in a simplified fashion as, lambda 2 minus lambda 1 lambda 1 by lambda 2 the power sigma x i. This lambda 1 by lambda 2 is greater than 1, because lambda 1 is greater than lambda 2 therefore, this will become an increasing function, **this is an increasing function** in T x is equal to sigma x i. So, we have monotone likelihood ratio in lambda and sigma x i.

So, we can apply the theorem that i gave, if the family has monotone likelihood ratio and theta and T x, then for one sided null hypothesis verses one sided alternative hypothesis the uniformly most powerful test is obtained here. So, let me write it here, so the UMP test is given by phi x; this here x means x 1 x 2 x n, it is rejecting if sigma x i is greater than c, it is rejecting with probability gamma if sigma x i is equal to c, it is zero if sigma x i is less than c.

Now, the sigma x i is actually, let me write say it has equal to y then that will follow poisson distribution n lambda. Now, c and gamma r determined by the condition expectation of the lambda naught phi x is equal to alpha, now this is reducing to, let me write it as a one and this as two, so this condition to let me simplify.

(Refer Slide Time: 25:43)

The condition (2) is simplified as

$$P_{\lambda_0}(Y > c) + \gamma P_{\lambda_0}(Y = c) = \alpha, \quad Y \sim \mathcal{P}(n\lambda_0)$$

$\lambda_0 = 1, n = 5 \quad \mathcal{P}(5), \quad \alpha = 0.1$

Consider one parameter exponential family

$$f(x, \theta) = c(\theta) e^{Q(\theta)T(x)}$$

Here $Q(\theta)$ is strictly monotonic

$$\sigma(x) = \frac{f(x, \theta_1)}{f(x, \theta_2)} = \frac{c(\theta_1)}{c(\theta_2)} e^{\frac{Q(\theta_1) - Q(\theta_2)}{>0} T(x)}$$

If Q is monotonically increasing, then $\theta_1 > \theta_2 \Rightarrow Q(\theta_1) > Q(\theta_2)$
 \uparrow in $T(x)$

So $\{f(x, \theta) : \theta \in \Omega\}$ has MLR in $(\theta, T(x))$

The condition two .So, expectation of lambda naught phi x, since this phi x is completely dependent upon sigma x i that is y, so it is becoming probability of y greater than c plus gamma times probability of y is equal to c, this is equal to alpha when the true parameter value is lambda naught. Now, another point which I would like to explain here, in the case of simple verses simple hypothesis we had the probability of type one error as a single value, but when we have composite hypothesis for the null hypothesis, then the probability of type one error is a function.

However, this is an increasing function which i mentioned in the statement of the theorem also, that the power function is strictly increasing function, so the probability of type one error is increasing. So, when you are getting theta is equal to theta naught, then at that point the maximum value is adoptant. So, effectively this condition is actually the size condition that is expectation of phi x equal to alpha lambda naught, this is the maximum probability of type one error here, that we are fixing to be equal to alpha. So, the size condition now, it is reduced to a condition which is involving the distribution poison n lambda naught therefore, from the tables of the poison distribution one can calculate this. Suppose, I say lambda naught is equal to 1, and n is equal to say phi, then basically we are looking at the tables of poison phi distribution. Suppose, I say alpha is equal to 0.1, then basically what we are seeing here is that what is the point from where... Now, the c could be; need not be an integer actually, we may fix that thing in

such a way, if it is an integer then this value may be positive, if it is naught integer then this may become zero.

Now, you may see from the tables that whether this randomization with probability gamma is required or not, if it is not required then this probability can be taken to be zero, there will be a point where after you will have the probability alpha. In case, that is naught possible, then we suitably choose a value where we lot of probability and then we may give some value of gamma also.

So, now this can be calculated from the tables of the **from the tables of the** poisson distribution. We can also see like a binomial distribution, we suppose we are hyper geometric distribution, suppose we have negative binomial distribution, in all of these distributions we are able to find out the uniformly most powerful tests. I will consider this derivation of the test in the following lecture; let me further develop this theory of the UMP tests here.

So, let us consider one parameter exponential family, so we are considering the form of the probability mass function or the probability density function as $c(\theta) e^{Q(\theta) T(x) - h(x)}$. Here Q function is strictly monotonic function; that means, it could be monotonically increasing or monotonically decreasing. Let us write down the ratio $f(x; \theta_1)$ by $f(x; \theta_2)$, then this is becoming $c(\theta_1) e^{Q(\theta_1) T(x) - h(x)}$ divided by $c(\theta_2) e^{Q(\theta_2) T(x) - h(x)}$, then $T(x)$ and $h(x)$ will get cancelled out by $c(\theta_2)$.

(Refer Slide Time: 31:26)

If $Q(\theta)$ is monotonically decreasing then
 $r(x)$ is decreasing in $T(x)$
 So MLR is $(\theta_1, -T(x))$

Corollary: Let X have a prob. density in one parameter exponential family
 $f(x; \theta) = c(\theta) e^{Q(\theta) T(x) - h(x)}$,
 where Q is monotonic fn., then \exists a UMP test for $H_0: \theta \leq \theta_0$
 vs. $H_1: \theta > \theta_0$.

If Q is \uparrow the test is of the form $\phi_1(x) = \begin{cases} 1 & \text{if } T(x) > c \\ \gamma & \text{if } T(x) = c \\ 0 & \text{if } T(x) < c \end{cases}$

If Q is \downarrow the inequalities will get reversed.
 c & γ are determined by $E_{\theta_0} \phi_1(X) = \alpha$.

If Q is monotonically increasing, then θ_1 greater than θ_2 will imply $Q(\theta_1)$ greater than $Q(\theta_2)$; that means, this term will become positive, and you will have this as increasing function of increasing function of $T(x)$. So, this ratio becomes an increasing function of $T(x)$, so the family $f(x; \theta)$; this will have monotone likelihood ratio in $T(x)$. On the other hand, if I consider say $Q(\theta_2)$ be monotonically decreasing, **if Q theta is monotonically decreasing** then this $r(x)$ term, what will happen here? That $Q(\theta_1)$ will become less than $Q(\theta_2)$, if θ_1 is greater than θ_2 therefore, this term will become decreasing function of $T(x)$, and therefore, monotone likelihood ratio will be in minus $T(x)$; decreasing in $T(x)$, so MLR will be in θ and minus $T(x)$; that means, the test function will get reversed, inequalities like here we are $T(x) > c$ you will it will become $T(x) < c$.

So, as a corollary of the previous theorem, we can write then let x have a probability density in one parameter exponential family, that is $f(x; \theta)$ is equal to $c(\theta) e^{-Q(\theta)T(x)}$ into $h(x)$, where Q is monotonic function then there exist a UMP test for $H_0: \theta \leq \theta_0$ less than or equal to θ_0 against $H_1: \theta > \theta_0$ greater than θ_0 . If Q is increasing, the test is of the form $\phi(x) = 1$ if $T(x)$ is greater than c **if $T(x)$ is equal to c** it is zero if $T(x)$ is less than c , if Q is decreasing the inequalities will get reversed, and here c and γ are determined by expectation of θ_0 $\phi(x) = 1$ if $T(x)$ is equal to c .

(Refer Slide Time: 34:31)

Example: Let X_1, \dots, X_n be a random sample from double exponential distⁿ. with pdf

$$f(x, \theta) = \frac{1}{2\theta} e^{-\frac{|x|}{\theta}}, \quad x \in \mathbb{R}, \theta > 0$$

$H_0: \theta \leq \theta_0$
 $H_1: \theta > \theta_0$

The joint pdf of X_1, \dots, X_n is

$$f(\mathbf{x}, \theta) = \frac{1}{(2\theta)^n} e^{-\frac{\sum |x_i|}{\theta}}$$

So MLR in $(\theta, \sum |x_i|)$

So UMP test is given by $T(x)$

$Q(\theta) = -\frac{1}{\theta} \ln \theta$

NPTL

Let me consider one example, let x_1, x_2, \dots, x_n be a random sample from double exponential distribution, with pdf given by say $f(x; \theta)$ is equal to $\frac{1}{2\theta} e^{-|x|/\theta}$. Here x is a real number, and θ has to be a positive parameter here, let us consider say, $\theta < \theta_0$ or $\theta > \theta_0$. You can easily see that this is a one parameter exponential family, and the monotone likelihood ratio; here you may consider $Q(\theta)$ as equal to $-1/\theta$.

So, naturally this is increasing in θ because $-1/\theta$ is decreasing, so $-1/\theta$ is increasing. So, this is strictly a monotonic function, so this suppose I write down the joint distribution of x_1, x_2, \dots, x_n , $\frac{1}{(2\theta)^n} e^{-\sum |x_i|/\theta}$ that is equal to $\frac{1}{(2\theta)^n} e^{-\sum |x_i|/\theta}$ to the power n $e^{-\sum |x_i|/\theta}$, so monotone likelihood ratio in θ and $\sum |x_i|$, this is $T = \sum |x_i|$. Therefore, by an application of this corollary that I mentioned UMP test for this problem,

(Refer Slide Time: 37:00)

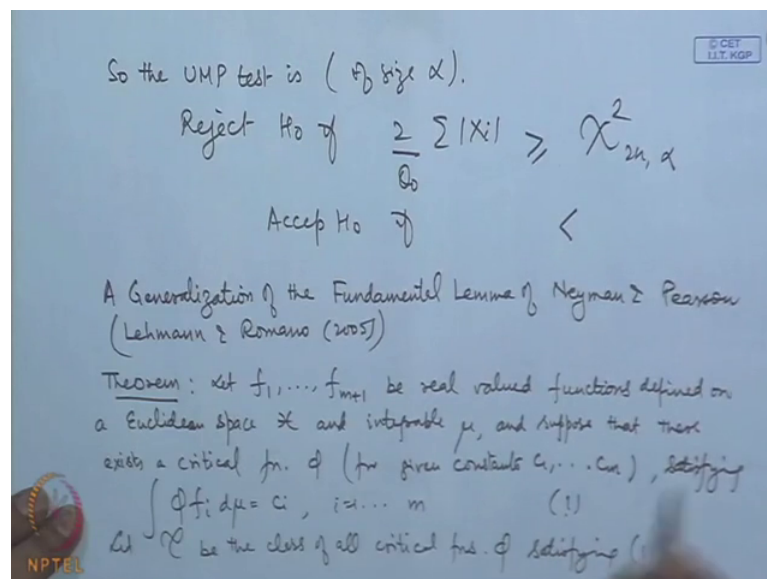
• Reject H_0 if $\sum |X_i| \geq c$
 where c is to be determined from the size condition
 $E_{\theta_0} \phi(X) = \alpha$
 $Y_i = |X_i| \sim \frac{1}{\theta} e^{-\frac{Y_i}{\theta}}, Y_i > 0, \theta > 0$
 $\frac{\sum Y_i}{\theta} = \frac{\sum |X_i|}{\theta} \sim \text{Gamma}(n, 1)$
 $\frac{2 \sum |X_i|}{\theta_0} \sim \chi^2_{2n}$ under $\theta = \theta_0$.
 $P_{\theta=\theta_0} \left(\frac{2 \sum |X_i|}{\theta_0} \geq \frac{2c}{\theta_0} \right) = \alpha \Rightarrow \frac{2c}{\theta_0} = \chi^2_{2n, \alpha}$

Let us say reject H_0 if $\sum |x_i|$ is greater than c . Now, note here that we are dealing with the continuous distributions, so I have returned this part only, this part will be the acceptance region. Now, $\sum |x_i|$ is equal to c , we need not write this portion here because this will have probability zero, so without loss of generality I am including the equality here.

Now, what we need to do is to determine this, where c is to be determined from the size condition, that is expectation of $\phi(x)$ is equal to α . Now, let us consider say y_i is equal to modulus of x_i , if x_i is having double exponential distribution, then modulus of x_i will have simple exponential distribution, that is $\frac{1}{\theta} e^{-y_i/\theta}$. So, this will have distribution $\frac{1}{\theta} e^{-y_i/\theta}$.

So, $\sum y_i$ will have gamma distribution with parameters n and 1 ; that means, $2 \sum |x_i|$ will follow chi square distribution on $2n$ degrees of freedom under $\theta = \theta_0$. So, when we consider this size condition, that is probability of $\sum |x_i| \geq c$ when $\theta = \theta_0$, this is equal to α when c is equal to $\chi^2_{2n, \alpha}$; that means, on the chi square curve with the $2n$ degrees of freedom this probability is equal to α , so this is $\chi^2_{2n, \alpha}$.

(Refer Slide Time: 39:46)



So, the test is written as, so the UMP test is reject H_0 if $\frac{2}{\theta_0} \sum |x_i| \geq \chi^2_{2n, \alpha}$; this is the UMP test of size α , and of course, accept H_0 if this is less. So, you can see this extension of the Neyman Pearson theory to the family's with the monotone likelihood ratio is helpful in providing the uniformly most powerful tests for one sided testing problems. If the

families have monotonic likelihood ratio, we are able to directly use these things here, and we are having exact test here; that means, once we have the observations, and we our testing problem is clearly specified, then at a given level of significance we can provide decision, whether we should accept a null hypothesis are not. On the other hand, if you do not is specify alpha in advance, then we can find out the p value here.

Now, let me proceed further with this theory here. Now, for further extension of this theory of most powerful tests, a generalization of the Neyman Pearson fundamental lemma was done. Let me state this results without any proof here, and these results are used for solving further problem; that means, here we are considering theta less than are equal to theta naught, so it is strictly one sided.

Now, there can be cases were we may have two sided also for example, if I am considering say I binomial proportion, whether it lies in a range or it is outside a range. Now, if I say it is within a range then it is like an interval, but if you i say it is an outside a range for example, I may say it is outside the interval 1 by 4 to 3 by 4; that means, I am saying the hypothesis $p \leq 1/4$ or $p \geq 3/4$, is a two sided thing. Now, in families with the monotone likelihood ratio etcetera, this Neyman Pearson theory applicable to this also, and then there is another point; that is regarding the determination of the constant in the test. In the one sided thing, the maximum was occurring at the cut of point that is theta naught here, when we have two sided then you will have two cut of points, it will increase and then... So, what will happen, that we will consider the maximum value at both the end points, that is at both end points of the intervals.

So, these results are proved using a certain extended features of the Neyman Pearson lemma. So, the result is known as the generalization of the fundamental lemma. let me give it here first of all, a generalization of the fundamental lemma of Neyman and Pearson. This is statement and the proof one can find out in the book of Lehmann and Romano, I will be skipping the details of the proof; I will only give the statement here. Let f_1, f_2, \dots, f_{m+1} be real valued functions defined on a Euclidean space x and integrable μ , and suppose that there exists a critical function ϕ for given constants c_1, c_2, \dots, c_n satisfying $\int \phi f_i d\mu$ is equal to c_i i is equal to 1 to m . Let us say C is the class of all critical functions ϕ satisfying 1.

(Refer Slide Time: 45:26)

(i) Among all members of \mathcal{C} \exists one that maximizes $\int \phi f_{m+1} d\mu$.

(ii) A sufficient condition for a member of \mathcal{C} to maximize $\int \phi f_{m+1} d\mu$ is the existence of constants $k_1, \dots, k_m \geq 0$ such that

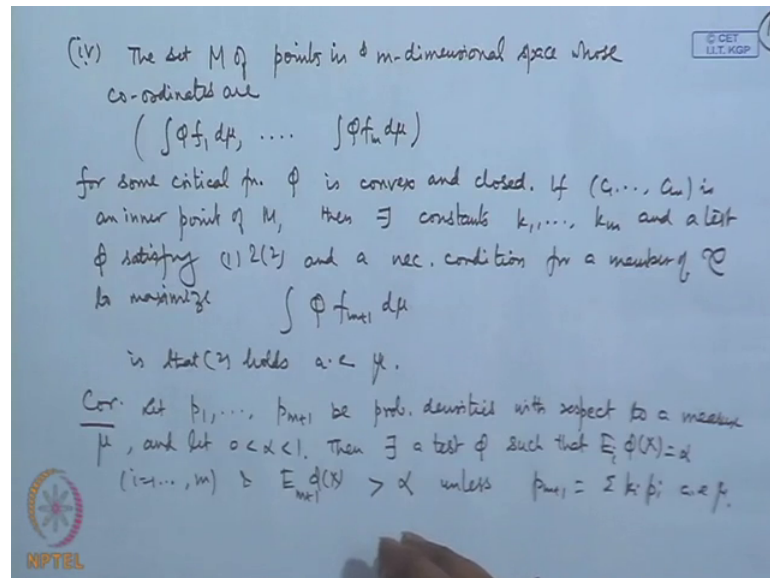
$$\phi(x) = \begin{cases} 1 & \text{when } f_{m+1}(x) > \sum_{i=1}^m k_i f_i(x) \\ 0 & \text{when } < \end{cases} \quad (2)$$

(iii) If a member of \mathcal{C} satisfies (2) with $k_1, \dots, k_m \geq 0$, then it maximizes $\int \phi f_{m+1} d\mu$ among all critical functions satisfying

$$\int \phi f_i d\mu \leq c_i, \quad i=1, \dots, m. \quad \dots (3)$$

Then among all members of \mathcal{C} there exists 1 that maximizes integral $\phi f_{m+1} d\mu$. A sufficient condition for a member of \mathcal{C} to maximize integral $\phi f_{m+1} d\mu$ is the existence of constants k_1, k_2, \dots, k_m , such that $\phi(x)$ is equal to 1 when $f_{m+1}(x)$ is greater than $\sum_{i=1}^m k_i f_i(x)$ and it is equal to zero when it is less. Thirdly, if a member of \mathcal{C} satisfies (2) with $k_1, k_2, \dots, k_m \geq 0$, then it maximizes integral $\phi f_{m+1} d\mu$ among all critical functions satisfying $\int \phi f_i d\mu \leq c_i$ for $i=1$ to m .

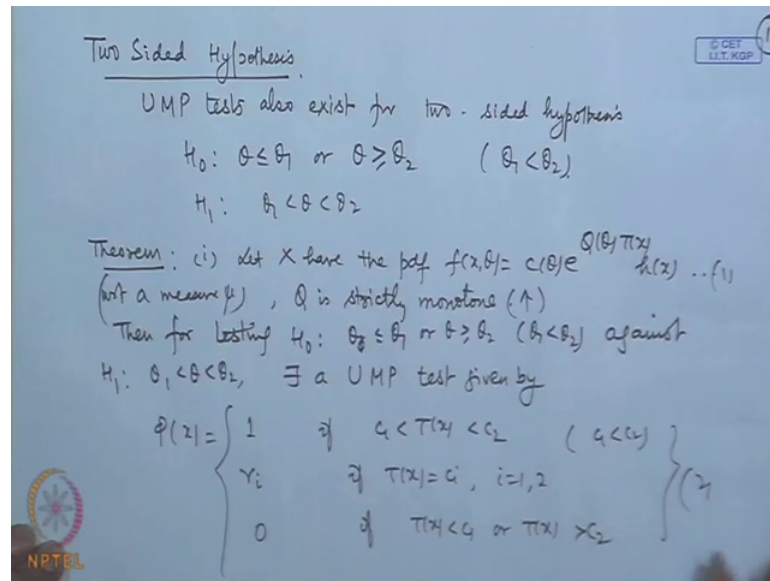
(Refer Slide Time: 47:42)



And then lastly, the set m of points in the m dimensional space whose co-ordinates are say $\int \phi f_1 d\mu$ and so on, $\int \phi f_m d\mu$ for some critical function ϕ this is convex and closed. If (c_1, \dots, c_m) is the inner point of M , then there exist constants k_1, \dots, k_m , and a test ϕ satisfying (1) and (2) and a necessary condition for a member of M to maximize integral $\int \phi f_{m+1} d\mu$ is that (2) holds almost everywhere. As i mentioned i will not be giving the proof of these results, one can see the book of Lehmann.

Now, this extension is helpful for solving more general testing problems, as a corollary I state the following, let p_1, \dots, p_{m+1} be probability densities with respect to a measure μ , and let $0 < \alpha < 1$, then there exists a test ϕ such that expectation of $\phi(X)$ is equal to α for $i = 1, \dots, m$, and expectation of $\phi(X)$ for $m+1$ is greater than α unless of course, $p_{m+1} = \sum k_i p_i$ almost everywhere. So, this will actually give the solution to more general two sided null hypothesis testing problems.

(Refer Slide Time: 51:02)



So, we have the following result than, that is if I am considering two sided hypothesis. So, we can say that UMP tests also exists for certain to sided hypothesis of this nature h naught, say theta less than or equal to theta 1 or theta greater than or equal to theta 2, where theta 1 is less than theta 2. So, we may like to test whether for example, say theta is the error measurements. So, we may like to check, whether the error measurements like within a certain range or there outside the range. It could be like, we are producing certain items and say certain ball bearings are produced, and we are looking at the diameter of the ball bearings.

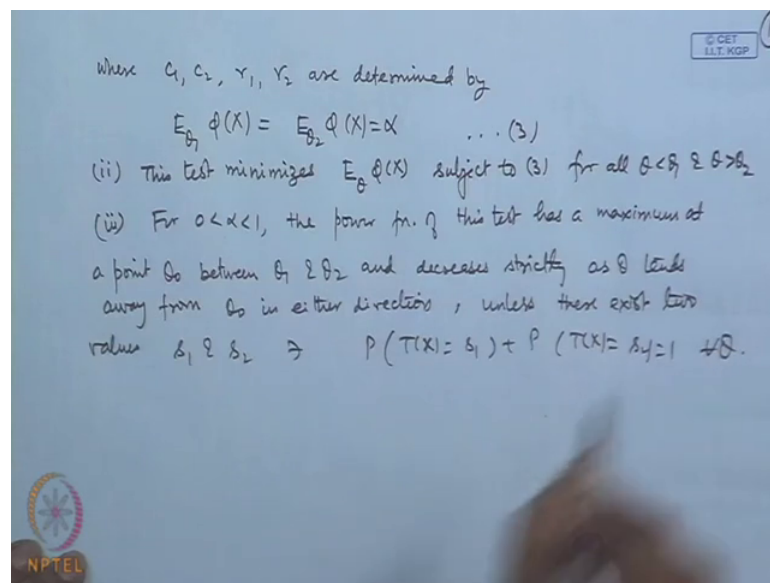
So, whether the ball bearings diameters are within a range a range or it is outside the range. If it is within the range we will be excepting the product as the good item, if it is outside then we will be rejecting that. So therefore, this is a perfect case for the two sided testing hypothesis problems, we may have say H 1 as theta 1 less than theta less than theta 2.

So, the result is that, by the use of the generalization of the Neyman Pearson fundamental lemma, we can actually give the uniformly most powerful test for these situations also. So, we have the following theorem, which I will the state, let x have the probability density function with respect to a measure mu and Q is strictly monotone, then for testing theta 1 theta less than or equal to theta 1 or theta greater than or equal to theta 2, where theta 1 is less than theta 2 against the alternatives theta 1 less than theta less than

theta 2, there exists a uniformly most powerful test. Of course, here against c_1 has to be less than c_2 , it is γ_1 if $T(x)$ is equal to c_1 for i is equal to 1 2. So, there are two points of randomization here, and we accept if $T(x)$ is less than c_1 or $T(x)$ is greater than c_2 . Once again, if we are considering Q to be strictly monotone here, then the familiar distribution has monotone likelihood ratio in $\theta T(x)$ or θ minus $T(x)$ and therefore....

So, here I would take an for example, increasing say because we are writing the **region in the** rejection region in this one, so I am considering monotonically increasing. So, we are rejecting when the value lies between two ranges, and we are accepting for a smaller values of $T(x)$ or larger values of $T(x)$. If it is decreasing, then the inequalities will get reversed, and at the boundary points of the interval we have done the randomization.

(Refer Slide Time: 55:25)



So let me consider this as 1 this as 2 say, where these constants c_1 c_2 γ_1 γ_2 , they are determined by $E_{\theta_1} \phi(x) = E_{\theta_2} \phi(x) = \alpha$. This test minimizes $E_{\theta} \phi(x)$ subject to (3) for all θ less than θ_1 , and θ greater than θ_2 . And for zero less than α less than 1, the power function of this test has a maximum at a point θ_0 between θ_1 and θ_2 , and decreases strictly as θ tends away from θ_0 in either direction, unless of course, there exist 2 values; say s_1 and s_2 , such that $P(T(x) = s_1) + P(T(x) = s_2) = 1$ for all θ .

So, here you can see the probability of type one error will be maximized at the end point; that is, at θ_1 and θ_2 , that is why we are fixing that value equal to α , so this is the size condition in the two sided null hypothesis problem. When we have one sided hypothesis problem, than the maximum value is occurring at the cut point; that cut of point, where the null and alternative hypothesis points are changing. But when we have two sided then will have two points; one is below and another is above, and at both the points we are having the maximum value of the probability of type one error, that value we are fixing as the α value. In the next lecture, I will be considering further amplification of these results; certain applications of this, and we will also consider certain properties of this power function here, which are based on actually the monotone likelihood ratio property. So, basically their properties of the expectations, I will be discussing at in the next lecture.