Lecture 15: UMP tests and unbiased tests

To complete the proof of Theorem 6.2, we need the following lemma.

Lemma 6.3

Suppose that the distribution of X is in a parametric family $\mathscr{P}$ indexed by a real-valued θ and that $\mathscr{P}$ has monotone likelihood ratio in $Y(X)$. If ψ is a nondecreasing function of Y, then $g(\theta) = E[\psi(Y)]$ is a nondecreasing function of θ.

Proof.

For $\theta_1 < \theta_2$, define $A = \{x : f_{\theta_1}(x) > f_{\theta_2}(x)\}$
$a = \sup_{x \in A} \psi(Y(x))$
$B = \{x : f_{\theta_1}(x) < f_{\theta_2}(x)\}$
$b = \inf_{x \in B} \psi(Y(x))$.
Since $\mathscr{P}$ has monotone likelihood ratio in $Y(X)$ and ψ is nondecreasing in Y, $b \geq a$.
Then the result follows from

$$
\begin{aligned}
g(\theta_2) - g(\theta_1) &= \int \psi(Y(x))(f_{\theta_2} - f_{\theta_1})(x) d\nu \\
&\geq a \int_A (f_{\theta_2} - f_{\theta_1})(x) d\nu + b \int_B (f_{\theta_2} - f_{\theta_1})(x) d\nu \\
&= (b-a) \int_B (f_{\theta_2} - f_{\theta_1})(x) d\nu \\
&\geq 0
\end{aligned}
$$

An important consequence

If $\psi(y) = I_{(t,\infty)}(y)$, then $g(\theta) = P(Y > t) = 1 - F_Y(t)$ is nondecreasing in θ.

Example 6.6

Let $X_1, ..., X_n$ be i.i.d. from the $N(\mu, \sigma^2)$ distribution with an unknown $\mu \in \mathscr{R}$ and a known σ^2.
Consider $H_0 : \mu \leq \mu_0$ versus $H_1 : \mu > \mu_0$, where μ_0 is a fixed constant.
The p.d.f. of $X = (X_1, ..., X_n)$ is from a one-parameter exponential family with $Y(X) = \bar{X}$ and $\eta(\mu) = n\mu/\sigma^2$.
By Corollary 6.1 and the fact that $\bar{X}$ is $N(\mu, \sigma^2/n)$, the UMP test is $T_*(X) = I_{(c_\alpha,\infty)}(\bar{X})$, where $c_\alpha = \sigma z_{1-\alpha}/\sqrt{n} + \mu_0$ and $z_a = \Phi^{-1}(a)$.

Discussion

To derive a UMP test for testing $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$ when X has a p.d.f. in a one-parameter exponential family, it is essential to know the distribution of $Y(X)$.

Typically, a nonrandomized test can be obtained if the distribution of Y is continuous; otherwise UMP tests are randomized.

Example 6.8

Let $X_1, ..., X_n$ be i.i.d. random variables from the Poisson distribution $P(\theta)$ with an unknown $\theta > 0$.

The p.d.f. of $X = (X_1, ..., X_n)$ is from a one-parameter exponential family with $Y(X) = \sum_{i=1}^n X_i$ and $\eta(\theta) = \log \theta$.

Note that Y has the Poisson distribution $P(n\theta)$.

By Corollary 6.1, a UMP test for $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$ is given by Theorem 6.2 with c and γ satisfying

$$\alpha = \sum_{j=c+1}^{\infty} \frac{e^{n\theta_0}(n\theta_0)^j}{j!} + \gamma \frac{e^{n\theta_0}(n\theta_0)^c}{c!}.$$

Example 6.9

Let $X_1, ..., X_n$ be i.i.d. random variables from the uniform distribution $U(0, \theta)$, $\theta > 0$.
Consider the hypotheses $H_0 : \theta \leq \theta_0$ and $H_1 : \theta > \theta_0$.
The p.d.f. of $X = (X_1, ..., X_n)$ is in a family with monotone likelihood ratio in $Y(X) = X_{(n)}$ (Example 6.4).
By Theorem 6.2, a UMP test is T_*.
Since $X_{(n)}$ has the Lebesgue p.d.f. $n\theta^{-n}x^{n-1}I_{(0,\theta)}(x)$, the UMP test T_* is nonrandomized and

$$\alpha = \beta_{T_*}(\theta_0) = \frac{n}{\theta_0^n}\int_c^{\theta_0} x^{n-1}dx = 1 - \frac{c^n}{\theta_0^n}.$$

Hence $c = \theta_0(1-\alpha)^{1/n}$.
The power function of T_* when $\theta > \theta_0$ is

$$\beta_{T_*}(\theta) = \frac{n}{\theta^n}\int_c^{\theta} x^{n-1}dx = 1 - \frac{\theta_0^n(1-\alpha)}{\theta^n}.$$

In this problem, however, UMP tests are not unique.
Note that the condition $P_\theta(f_\theta(X) = cf_{\theta_0}(X)) = 0$ in Theorem 6.2(iv) is not satisfied.

It can be shown that the following test is also UMP with size α:

$$T(X) = \begin{cases} 1 & X_{(n)} > \theta_0 \\ \alpha & X_{(n)} \leq \theta_0. \end{cases}$$

The following result is useful for finding optimal tests for two sided hypotheses.

Proposition 6.1 (Generalized Neyman-Pearson lemma)

Let $f_1, ..., f_{m+1}$ be Borel functions on $\mathscr{R}^p$ integrable w.r.t. a σ-finite ν. For given constants $t_1, ..., t_m$, let $\mathscr{T}$ be the class of Borel functions ϕ (from $\mathscr{R}^p$ to $[0,1]$) satisfying

$$\int \phi f_i d\nu \leq t_i, \quad i = 1, ..., m, \tag{1}$$

and $\mathscr{T}_0$ be the set of ϕ's in $\mathscr{T}$ satisfying (1) with all inequalities replaced by equalities. If there are constants $c_1, ..., c_m$ such that

$$\phi_*(x) = \begin{cases} 1 & f_{m+1}(x) > c_1 f_1(x) + \cdots + c_m f_m(x) \\ 0 & f_{m+1}(x) < c_1 f_1(x) + \cdots + c_m f_m(x) \end{cases}$$

is a member of $\mathscr{T}_0$, then ϕ_* maximizes $\int \phi f_{m+1} d\nu$ over $\phi \in \mathscr{T}_0$. If $c_i \geq 0$ for all i, then ϕ_* maximizes $\int \phi f_{m+1} d\nu$ over $\phi \in \mathscr{T}$.

The proof is left as an exercise.
The existence of constants c_i's in ϕ_* is considered in the following lemma whose proof can be found in Lehmann (1986, pp. 97-99).

Lemma 6.2

Let $f_1, ..., f_m$ and ν be given by Proposition 6.1.
Then the set $M = \left\{ (\int \phi f_1 d\nu, ..., \int \phi f_m d\nu) : \ \phi \text{ is from } \mathscr{R}^p \text{ to } [0,1] \right\}$ is convex and closed.
If $(t_1, ..., t_m)$ is an interior point of M, then there exist constants $c_1, ..., c_m$ such that the function ϕ_* defined in Proposition 6.1 is in $\mathscr{T}_0$.

Two-sided hypotheses

The following hypotheses are called two-sided hypotheses:

$$H_0: \ \theta \leq \theta_1 \text{ or } \theta \geq \theta_2 \quad \text{versus} \quad H_1: \ \theta_1 < \theta < \theta_2, \tag{2}$$

$$H_0: \ \theta_1 \leq \theta \leq \theta_2 \quad \text{versus} \quad H_1: \ \theta < \theta_1 \text{ or } \theta > \theta_2, \tag{3}$$

$$H_0: \ \theta = \theta_0 \quad \text{versus} \quad H_1: \ \theta \neq \theta_0, \tag{4}$$

where θ_0, θ_1, and θ_2 are given constants and $\theta_1 < \theta_2$.

Theorem 6.3 (UMP tests for two-sided hypotheses)

Suppose that X has a p.d.f. in a one-parameter exponential family, i.e., the p.d.f. is

$$f_\theta(x) = \exp\{\eta(\theta)Y(x) - \xi(\theta)\}h(x)$$

w.r.t. a σ-finite measure, where η is a strictly increasing function of θ.
(i) For testing hypotheses (2), a UMP test of size α is

$$T_*(X) = \begin{cases} 1 & c_1 < Y(X) < c_2 \\ \gamma_i & Y(X) = c_i,\ i = 1,2 \\ 0 & Y(X) < c_1 \text{ or } Y(X) > c_2, \end{cases} \tag{5}$$

where c_i's and γ_i's are determined by

$$\beta_{T_*}(\theta_1) = \beta_{T_*}(\theta_2) = \alpha. \tag{6}$$

(ii) T_* minimizes $\beta_T(\theta)$ over all $\theta < \theta_1$, $\theta > \theta_2$, and T satisfying (6).
(iii) If T_* and T_{**} are two tests satisfying (5) and $\beta_{T_*}(\theta_1) = \beta_{T_{**}}(\theta_1)$ and if the region $\{T_{**} = 1\}$ is to the right of $\{T_* = 1\}$, then $\beta_{T_*}(\theta) < \beta_{T_{**}}(\theta)$ for $\theta > \theta_1$ and $\beta_{T_*}(\theta) > \beta_{T_{**}}(\theta)$ for $\theta < \theta_1$.
If both T_* and T_{**} satisfy (5) and (6), then $T_* = T_{**}$ a.s. $\mathscr{P}$.

Proof

(i) Since Y is sufficient for θ, we only need to consider tests of the form $T(Y)$.
By Theorem 2.1, the distribution of Y has a p.d.f.

$$g_\theta(y) = \exp\{\eta(\theta)y - \xi(\theta)\} \tag{7}$$

Let $\theta_1 < \theta_3 < \theta_2$.
Consider the problem of testing $\theta = \theta_1$ or $\theta = \theta_2$ versus $\theta = \theta_3$.
(α, α) is an interior point of the set of all points $(\beta_T(\theta_1), \beta_T(\theta_2))$ as T ranges over all tests of the form $T(Y)$.
By (7) and Lemma 6.2, there are constants $\tilde{c}_1$ and $\tilde{c}_2$ such that

$$T_*(Y) = \begin{cases} 1 & a_1 e^{b_1 Y} + a_2 e^{b_2 Y} < 1 \\ 0 & a_1 e^{b_1 Y} + a_2 e^{b_2 Y} > 1 \end{cases}$$

satisfies (6), where $a_i = \tilde{c}_i e^{\xi(\theta_3) - \xi(\theta_i)}$ and $b_i = \eta(\theta_i) - \eta(\theta_3)$, $i = 1, 2$.
Clearly a_i's cannot both be ≤ 0.

If one of the a_i's is ≤ 0 and the other is > 0, then $a_1 e^{b_1 Y} + a_2 e^{b_2 Y}$ is strictly monotone (since $b_1 < 0 < b_2$) and

$$T_*(\text{ or } 1 - T_*) = \begin{cases} 1 & Y(X) > c \\ \gamma & Y(X) = c \\ 0 & Y(X) < c, \end{cases}$$

which has a strictly monotone power function (Theorem 6.2) and, therefore, cannot satisfy $\beta_{T_*}(\theta_1) = \beta_{T_*}(\theta_2) = \alpha$.

Thus, both a_i's are positive.

The function $a_1 e^{b_1 Y} + a_2 e^{b_2 Y}$ is convex (since $b_1 < 0 < b_2$).

$a_1 e^{b_1 Y} + a_2 e^{b_2 Y} < 1$ is equivalent to $c_1 < Y < c_2$ for some c_1 and c_2.

Then, T_* is of the form (5) and it follows from Proposition 6.1 that T_* is UMP for testing $\theta = \theta_1$ or $\theta = \theta_2$ versus $\theta = \theta_3$.

Since T_* does not depend on θ_3, it follows from Lemma 6.1 that T_* is UMP for testing $\theta = \theta_1$ or $\theta = \theta_2$ versus H_1.

To show that T_* is a UMP test of size α for testing H_0 versus H_1, it remains to show that $\beta_{T_*}(\theta) \leq \alpha$ for $\theta \leq \theta_1$ or $\theta \geq \theta_2$, which follows from part (ii) of the theorem by comparing T_* with the test $T(Y) \equiv \alpha$.

(ii) The proof is similar to that in (i) and is left as an exercise.

(iii) The first claim follows from Lemma 6.4, since $T_{**} - T_*$ has a single change of sign; the second claim follows from the first claim.

Part (iii) of Theorem 6.3 shows that the c_i's and γ_i's are uniquely determined by (5) and (6), and indicates how to select the c_i's and γ_i's.

One can start with some trial values $c_1^{(0)}$ and $\gamma_1^{(0)}$, find $c_2^{(0)}$ and $\gamma_2^{(0)}$ such that $\beta_{T_*}(\theta_1) = \alpha$, and compute $\beta_{T_*}(\theta_2)$.

If $\beta_{T_*}(\theta_2) < \alpha$, by Theorem 6.3(iii), the correct rejection region $\{T_* = 1\}$ is to the right of the one chosen so that one should try $c_1^{(1)} > c_1^{(0)}$ or $c_1^{(1)} = c_1^{(0)}$ and $\gamma_1^{(1)} < \gamma_1^{(0)}$; the converse holds if $\beta_{T_*}(\theta_2) > \alpha$.

Example 6.10

Let $X_1, ..., X_n$ be i.i.d. from $N(\theta, 1)$.
By Theorem 6.3, a UMP test for testing (2) is $T_*(X) = I_{(c_1,c_2)}(\bar{X})$, where c_i's are determined by

$$\Phi\big(\sqrt{n}(c_2 - \theta_1)\big) - \Phi\big(\sqrt{n}(c_1 - \theta_1)\big) = \alpha$$

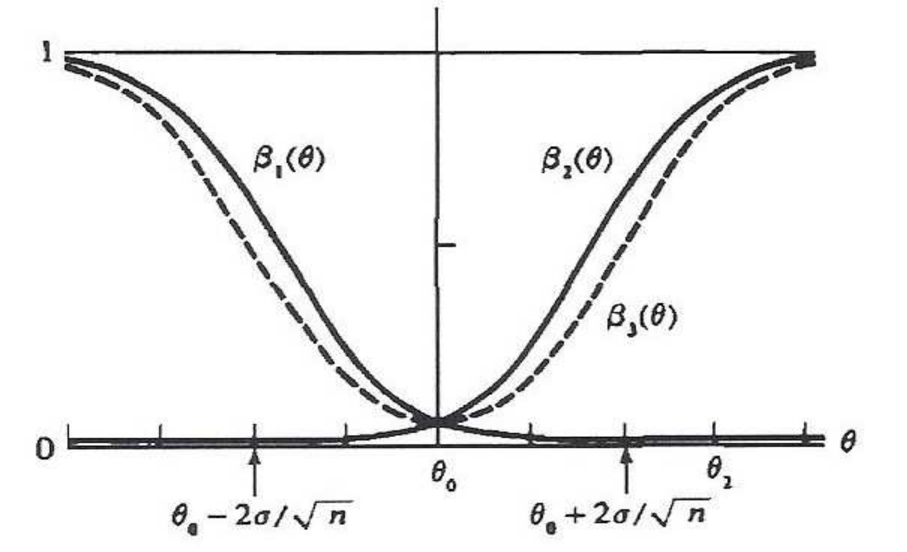
and

$$\Phi\big(\sqrt{n}(c_2 - \theta_2)\big) - \Phi\big(\sqrt{n}(c_1 - \theta_2)\big) = \alpha.$$

Nonexistence of UMP tests

- When the distribution of X is not from a one-parameter exponential family, UMP tests for hypotheses (2) exist in some cases (see Exercises 17 and 26).
- Unfortunately, a UMP test does not exist in general for testing hypotheses (3) or (4) (Exercises 28 and 29).
- A key reason for this phenomenon is that UMP tests for testing one-sided hypotheses do not have level α for testing (2); but they are of level α for testing (3) or (4) and there does not exist a single test more powerful than all tests that are UMP for testing one-sided hypotheses.
- Although UMP tests for testing one-sided hypotheses are of level α for testing (3) or (4) and have very good performances when θ is on one side of θ_0, they have very bad performances for θ on the other side of θ_0.
- If we eliminate these type of tests, then we may be able to find an optimal test.

Figure. Power functions of three tests for two sided hypotheses



Unbiased tests

When a UMP test does not exist, we may use the same approach used in estimation problems, i.e., imposing a reasonable restriction on the tests to be considered and finding optimal tests within the class of tests under the restriction.

Two such types of restrictions in estimation problems are unbiasedness and invariance.

A UMP test T of size α has the property that

$$\beta_T(P) \le \alpha, \quad P \in \mathscr{P}_0 \qquad \text{and} \qquad \beta_T(P) \ge \alpha, \quad P \in \mathscr{P}_1,$$

since T is at least as good as the silly test $T \equiv \alpha$.

Since $\beta_T(P)$, $P \in \mathscr{P}_1$ is the probability of correctly rejecting H_0, it is desired to have $\beta_T(P) \ge \alpha$, i.e., T is better than the silly test $\equiv \alpha$.

We want to consider tests that are at least better than the silly test $\equiv \alpha$.

This leads to the following definition of “unbiased tests”.

Definition 6.3

Let α be a given level of significance.
A test T for $H_0 : P \in \mathscr{P}_0$ versus $H_1 : P \in \mathscr{P}_1$ is said to be unbiased of level α if and only if

$$\beta_T(P) \le \alpha, \quad P \in \mathscr{P}_0 \qquad \text{and} \qquad \beta_T(P) \ge \alpha, \quad P \in \mathscr{P}_1,$$

A test of size α is called a *uniformly most powerful unbiased* (UMPU) test if and only if it is UMP within the class of unbiased tests of level α.

Discussion

Since a UMP test is UMPU, the discussion of unbiasedness of tests is useful only when a UMP test does not exist.

In a large class of problems for which a UMP test does not exist, there do exist UMPU tests.

Suppose that U is a sufficient statistic for $P \in \mathscr{P}$.

Then, similar to the search for a UMP test, we need to consider functions of U only in order to find a UMPU test, since, for any unbiased test $T(X)$, $E(T|U)$ is unbiased and has the same power function as T.

Consider the following hypotheses:

$$H_0 : \theta \in \Theta_0 \qquad \text{versus} \qquad H_1 : \theta \in \Theta_1,$$

where $\theta = \theta(P)$ is a functional from $\mathscr{P}$ onto Θ and Θ_0 and Θ_1 are two disjoint Borel sets with $\Theta_0 \cup \Theta_1 = \Theta$. ($\mathscr{P}_j = \{P : \theta \in \Theta_j\}$, $j = 0, 1$.)

For instance, $X_1, ..., X_n$ are i.i.d. from F but we are interested in testing $H_0 : \theta \leq 0$ versus $H_1 : \theta > 0$, where $\theta = EX_1$.

Definition 6.4 (Similarity)

Consider the hypotheses $H_0 : \theta \in \Theta_0$ vs $H_1 : \theta \in \Theta_1$.
Let α be a given level of significance and let $\bar{\Theta}_{01}$ be the common boundary of Θ_0 and Θ_1, i.e., the set of points θ that are points or limit points of both Θ_0 and Θ_1.
A test T is *similar* on $\bar{\Theta}_{01}$ if and only if $\beta_T(P) = \alpha$ for all $\theta \in \bar{\Theta}_{01}$.

Remark

It is more convenient to work with similarity than to work with unbiasedness for testing $H_0 : \theta \in \Theta_0$ vs $H_1 : \theta \in \Theta_1$.

Continuity of the power function

For a given test T, the power function $\beta_T(P)$ is said to be continuous in θ if and only if for any $\{\theta_j : j = 0, 1, 2, ...\} \subset \Theta$, $\theta_j \to \theta_0$ implies $\beta_T(P_j) \to \beta_T(P_0)$, where $P_j \in \mathscr{P}$ satisfying $\theta(P_j) = \theta_j$, $j = 0, 1, ...$.
If β_T is a function of θ, then this continuity property is simply the continuity of $\beta_T(\theta)$.

Lemma 6.5

Consider hypotheses $H_0 : \theta \in \Theta_0$ vs $H_1 : \theta \in \Theta_1$.
Suppose that, for every T, $\beta_T(P)$ is continuous in θ.
If T_* is uniformly most powerful among all similar tests and has size α, then T_* is a UMPU test.

Proof

Under the continuity assumption on β_T, the class of similar tests contains the class of unbiased tests.
Since T_* is uniformly at least as powerful as the test $T \equiv \alpha$, T_* is unbiased.
Hence, T_* is a UMPU test.