ST5224: Advanced Statistical Theory II

Chen Zehua

Department of Statistics & Applied Probability

Thursday, February 22, 2012
Lemma 6.7

Suppose that $X$ has the following p.d.f. w.r.t. a $\sigma$-finite measure:

$$f_{\theta,\varphi}(x) = \exp \left\{ \theta Y(x) + \varphi^T U(x) - \zeta(\theta, \varphi) \right\},$$

where $\theta$ is real-valued, $\varphi$ is vector-valued, and $Y$ and $U$ are statistics. Suppose also that $V(Y, U)$ is a statistic independent of $U$ when $\theta = \theta_j$, where $\theta_j$'s are known values given in the hypotheses in (i)-(iv) of Theorem 6.4. (i) If $V(y, u)$ is increasing in $y$ for each $u$, then the UMPU tests in (i)-(iii) of Theorem 6.4 are equivalent to those given by Theorem 6.4 with $Y$ and $(Y, U)$ replaced by $V$ and with $c_i(U)$ and $\gamma_i(U)$ replaced by constants $c_i$ and $\gamma_i$, respectively. (ii) If there are Borel functions $a(u) > 0$ and $b(u)$ such that $V(y, u) = a(u)y + b(u)$, then the UMPU test in Theorem 6.4(iv) is equivalent to that given by Theorem 6.4(iv) with $Y$ and $(Y, U)$ replaced by $V$ and with $c_i(U)$ and $\gamma_i(U)$ replaced by constants $c_i$ and $\gamma_i$, respectively.
Proof

(i) Since $V$ is increasing in $y$, $Y > c_i(u)$ is equivalent to $V > d_i(u)$ for some $d_i$.

The result follows from the fact that $V$ is independent of $U$ so that $d_i$’s and $\gamma_i$’s do not depend on $u$ when $Y$ is replaced by $V$.

(ii) Since $V = a(U)Y + b(U)$, the UMPU test in Theorem 6.4(iv) is the same as

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

subject to $E_{\theta_0}[T_*(V, U)|U = u] = \alpha$ and

$$E_{\theta_0} \left[ T_*(V, U) \frac{V - b(U)}{a(U)} \bigg| U \right] = \alpha E_{\theta_0} \left[ \frac{V - b(U)}{a(U)} \bigg| U \right]. \quad (1)$$

Under $E_{\theta_0}[T_*(V, U)|U = u] = \alpha$, (1) is the same as

$$E_{\theta_0}[T_*(V, U)V|U] = \alpha E_{\theta_0}(V|U).$$

Since $V$ and $U$ are independent when $\theta = \theta_0$, $c_i(u)$’s and $\gamma_i(u)$’s do not depend on $u$ and, therefore, $T_*$ does not depend on $U$. 
Remarks

- If the conditions of Lemma 6.7 are satisfied, then UMPU tests can be derived by working with the distribution of $V$ instead of $P_{Y|U=u}$.
- In exponential families, a $V(Y, U)$ independent of $U$ can often be found by applying Basu’s theorem (Theorem 2.4).
- When we consider normal families, $\gamma_i$’s can be chosen to be 0 since the c.d.f. of $Y$ given $U = u$ or the c.d.f. of $V$ is continuous.

One-sample problems

Let $X_1, ..., X_n$ be i.i.d. from $N(\mu, \sigma^2)$ with unknown $\mu \in \mathcal{R}$ and $\sigma^2 > 0$, where $n \geq 2$. The joint p.d.f. of $X = (X_1, ..., X_n)$ is

$$
\frac{1}{(2\pi\sigma^2)^{n/2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} x_i^2 + \frac{\mu}{\sigma^2} \sum_{i=1}^{n} x_i - \frac{n\mu^2}{2\sigma^2} \right\}.
$$
One-sample problems

Consider first hypotheses concerning $\sigma^2$. The p.d.f. of $X$ is in a multiparameter exponential family with $\theta = -(2\sigma^2)^{-1}$, $\varphi = n\mu/\sigma^2$, $Y = \sum_{i=1}^n X_i^2$, and $U = \bar{X}$. By Basu’s theorem, $V = (n - 1)S^2$ is independent of $U = \bar{X}$ (Example 2.18), where $S^2$ is the sample variance. Also,

$$\sum_{i=1}^n X_i^2 = (n - 1)S^2 + n\bar{X}^2, \text{ i.e., } V = Y - nU^2.$$

Hence the conditions of Lemma 6.7 are satisfied. Since $V/\sigma^2$ has the chi-square distribution $\chi^2_{n-1}$ (Example 2.18), values of $c_i$’s for hypotheses in (i)-(iii) of Theorem 6.4 are related to quantiles of $\chi^2_{n-1}$. For testing $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$ (which is equivalent to testing $H_0 : \sigma^2 = \sigma_0^2$ vs $H_1 : \sigma^2 \neq \sigma_0^2$), $d_i = c_i/\sigma_0^2$, $i = 1, 2$, are determined by

$$\int_{d_1}^{d_2} f_{n-1}(v)dv = 1 - \alpha \quad \text{ and } \quad \int_{d_1}^{d_2} v f_{n-1}(v)dv = (n-1)(1-\alpha),$$
One-sample problems

where \( f_m \) is the Lebesgue p.d.f. of the chi-square distribution \( \chi^2_m \).

Since \( v f_{n-1}(v) = (n - 1) f_{n+1}(v) \), \( d_1 \) and \( d_2 \) are determined by

\[
\int_{d_1}^{d_2} f_{n-1}(v) dv = \int_{d_1}^{d_2} f_{n+1}(v) dv = 1 - \alpha.
\]

If \( n - 1 \approx n + 1 \), then \( d_1 \) and \( d_2 \) are nearly the \((\alpha/2)\)th and \((1 - \alpha/2)\)th quantiles of \( \chi^2_{n-1} \), respectively, in which case the UMPU test in Theorem 6.4(iv) is the same as the “equal-tailed” chi-square test for \( H_0 \) in elementary textbooks.

Consider next hypotheses concerning \( \mu \).

The p.d.f. of \( X \) is in a multiparameter exponential family with

\[
Y = \bar{X}, \quad U = \sum_{i=1}^{n} (X_i - \mu_0)^2, \quad \theta = n(\mu - \mu_0)/\sigma^2, \quad \varphi = -(2\sigma^2)^{-1}.
\]

For testing hypotheses \( H_0 : \mu \leq \mu_0 \) versus \( H_1 : \mu > \mu_0 \), we take \( V \) to be

\[
t(X) = \sqrt{n}(\bar{X} - \mu_0)/S.
\]

By Basu’s theorem, \( t(X) \) is independent of \( U \) when \( \mu = \mu_0 \).

Hence it satisfies the conditions in Lemma 6.7(i).
One-sample problems

From Examples 1.16 and 2.18, \( t(X) \) has the t-distribution \( t_{n-1} \) when \( \mu = \mu_0 \).

Thus, \( c(U) \) in Theorem 6.4(i) is the \((1 - \alpha)\)th quantile of \( t_{n-1} \).

For the two-sided hypotheses \( H_0 : \mu = \mu_0 \) versus \( H_1 : \mu \neq \mu_0 \), the statistic \( V = (\bar{X} - \mu_0) / \sqrt{U} \) satisfies the conditions in Lemma 6.7(ii) and has a distribution symmetric about 0 when \( \mu = \mu_0 \).

Then the UMPU test in Theorem 6.4(iv) rejects \( H_0 \) when \(|V| > d|\), where \( d \) satisfies \( P(|V| > d) = \alpha \) when \( \mu = \mu_0 \).

Since

\[
t(X) = \frac{\sqrt{(n-1)nV(X)}}{\sqrt{1 - n[V(X)]^2}},
\]

the UMPU test rejects \( H_0 \) if and only if \(|t(X)| > t_{n-1,\alpha/2} \), where \( t_{n-1,\alpha} \) is the \((1 - \alpha)\)th quantile of the t-distribution \( t_{n-1} \).

The UMPU tests derived here are the so-called one-sample t-tests in elementary textbooks.

The power function of a one-sample t-test is related to the noncentral t-distribution introduced in §1.3.1 (see Exercise 36).
Two-sample problems

The problem of comparing the parameters of two normal distributions arises in the comparison of two treatments, products, and so on.

Suppose that we have two independent samples, $X_{i1}, \ldots, X_{in_i}$, $i = 1, 2$, i.i.d. from $N(\mu_i, \sigma_i^2)$, $i = 1, 2$, respectively, where $n_i \geq 2$.

The joint p.d.f. of $X_{ij}$’s is

$$C(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2) \exp \left\{ - \sum_{i=1}^{2} \frac{1}{2\sigma_i^2} \sum_{j=1}^{n_i} x_{ij}^2 + \sum_{i=1}^{2} \frac{n_i\mu_i}{\sigma_i^2} \bar{x}_i \right\},$$

where $\bar{x}_i$ is the sample mean based on $x_{i1}, \ldots, x_{in_i}$ and $C(\cdot)$ is a known function.
Two-sample problems

Consider first the hypothesis \( H_0 : \sigma_2^2 / \sigma_1^2 \leq \Delta_0 \) or \( H_0 : \sigma_2^2 / \sigma_1^2 = \Delta_0 \). The p.d.f. of \( X_{ij} \)'s is in a multiparameter exponential family with

\[
\theta = \frac{1}{2 \Delta_0 \sigma_1^2} - \frac{1}{2 \sigma_2^2}, \quad \varphi = \left( -\frac{1}{2 \sigma_1^2}, \frac{n_1 \mu_1}{\sigma_1^2}, \frac{n_2 \mu_2}{\sigma_2^2} \right),
\]

\[
Y = \sum_{j=1}^{n_2} X_{2j}^2, \quad U = \left( \sum_{j=1}^{n_1} X_{1j}^2 + \frac{1}{\Delta_0} \sum_{j=1}^{n_2} X_{2j}^2, \bar{X}_1, \bar{X}_2 \right).
\]

To apply Lemma 6.7, consider

\[
V = \frac{(n_2 - 1)S_2^2 / \Delta_0}{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2 / \Delta_0} = \frac{(Y - n_2 U_3^2) / \Delta_0}{U_1 - n_1 U_2^2 - n_2 U_3^2 / \Delta_0},
\]

where \( S_i^2 \) is the sample variance based on \( X_{i1}, \ldots, X_{ini} \) and \( U_j \) is the \( j \)th component of \( U \).

By Basu’s theorem, \( V \) and \( U \) are independent when \( \theta = 0 \) \( (\sigma_2^2 = \Delta_0 \sigma_1^2) \).
Two-sample problems

Since $V$ is increasing and linear in $Y$, the conditions of Lemma 6.7 are satisfied.

Thus, a UMPU test rejects $H_0 : \theta \leq 0$ (which is equivalent to $H_0 : \sigma_2^2/\sigma_1^2 \leq \Delta_0$) when $V > c_0$, where $c_0$ satisfies $P(V > c_0) = \alpha$ when $\theta = 0$; and a UMPU test rejects $H_0 : \theta = 0$ (which is equivalent to $H_0 : \sigma_2^2/\sigma_1^2 = \Delta_0$) when $V < c_1$ or $V > c_2$, where $c_i$’s satisfy $P(c_1 < V < c_2) = 1 - \alpha$ and $E[VT_*(V)] = \alpha E(V)$ when $\theta = 0$. Note that

$$V = \frac{(n_2 - 1)F}{n_1 - 1 + (n_2 - 1)F} \quad \text{with} \quad F = \frac{S_2^2/\Delta_0}{S_1^2}.$$

It follows from Example 1.16 that $F$ has the F-distribution $F_{n_2-1,n_1-1}$ when $\theta = 0$.

Since $V$ is a strictly increasing function of $F$, a UMPU test rejects $H_0 : \theta \leq 0$ when $F > F_{n_2-1,n_1-1,\alpha}$, where $F_{a,b,\alpha}$ is the $(1 - \alpha)$th quantile of the F-distribution $F_{a,b}$.

This is the F-test in elementary textbooks.
Two-sample problems

When $\theta = 0$, $V$ has the beta distribution $B((n_2 - 1)/2, (n_1 - 1)/2)$ and $E(V) = (n_2 - 1)/(n_1 + n_2 - 2)$ (Table 1.2). Then, $E[VT_\star(V)] = \alpha E(V)$ when $\theta = 0$ is the same as

$$
(1 - \alpha)(n_2 - 1) \over n_1 + n_2 - 2 = \int_{c_1}^{c_2} v f_{(n_2 - 1)/2, (n_1 - 1)/2}(v) dv,
$$

where $f_{a,b}$ is the p.d.f. of the beta distribution $B(a, b)$. Using the fact that

$$
v f_{(n_2 - 1)/2, (n_1 - 1)/2}(v) = (n_1 + n_2 - 2)^{-1} (n_2 - 1) f_{(n_2 + 1)/2, (n_1 - 1)/2}(v),
$$

we conclude that a UMPU test rejects $H_0 : \theta = 0$ when $V < c_1$ or $V > c_2$, where $c_1$ and $c_2$ are determined by

$$
1 - \alpha = \int_{c_1}^{c_2} f_{(n_2 - 1)/2, (n_1 - 1)/2}(v) dv = \int_{c_1}^{c_2} f_{(n_2 + 1)/2, (n_1 - 1)/2}(v) dv.
$$

If $n_2 - 1 \approx n_2 + 1$ (i.e., $n_2$ is large), then this UMPU test can be approximated by the F-test that rejects $H_0 : \theta = 0$ if and only if

$$
F < F_{n_2 - 1, n_1 - 1, 1 - \alpha/2} \text{ or } F > F_{n_2 - 1, n_1 - 1, \alpha/2}.
$$
Two-sample problems

Consider next the hypothesis $H_0 : \mu_1 \geq \mu_2$ or $H_0 : \mu_1 = \mu_2$. If $\sigma_1^2 \neq \sigma_2^2$, the problem is the so-called Behrens-Fisher problem and is not accessible by the method introduced in this section. We now assume that $\sigma_1^2 = \sigma_2^2 = \sigma^2$ but $\sigma^2$ is unknown. The p.d.f. of $X_{ij}$’s is then

$$C(\mu_1, \mu_2, \sigma^2) \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{2} \sum_{j=1}^{n_i} x_{ij}^2 + \frac{n_1\mu_1}{\sigma^2} \bar{x}_1 + \frac{n_2\mu_2}{\sigma^2} \bar{x}_2 \right\},$$

which is in a multiparameter exponential family with

$$\theta = \frac{\mu_2 - \mu_1}{(n_1^{-1} + n_2^{-1})\sigma^2}, \quad \varphi = \left( \frac{n_1\mu_1 + n_2\mu_2}{(n_1 + n_2)\sigma^2}, -\frac{1}{2\sigma^2} \right),$$

$$Y = \bar{X}_2 - \bar{X}_1, \quad U = \left( n_1 \bar{X}_1 + n_2 \bar{X}_2, \sum_{i=1}^{n_i} \sum_{j=1}^{2} x_{ij}^2 \right).$$
Two-sample problems

For testing $H_0 : \theta \leq 0$ (i.e., $\mu_1 \geq \mu_2$) versus $H_1 : \theta > 0$, we consider $V$ in Lemma 6.7 to be

$$t(X) = \frac{(\bar{X}_2 - \bar{X}_1)/\sqrt{n_1^{-1} + n_2^{-1}}}{\sqrt{[(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2]}/(n_1 + n_2 - 2)}.$$

When $\theta = 0$, $t(X)$ is independent of $U$ (Basu's theorem) and satisfies the conditions in Lemma 6.7(i); the numerator and the denominator of $t(X)$ (after division by $\sigma$) are independently distributed as $N(0, 1)$ and the chi-square distribution $\chi_{n_1+n_2-2}^2$, respectively. Hence $t(X)$ has the t-distribution $t_{n_1+n_2-2}$ and a UMPU test rejects $H_0$ when $t(X) > t_{n_1+n_2-2, \alpha}$ (the $(1 - \alpha)$th quantile of $t_{n_1+n_2-2}$). This is the so-called (one-sided) two-sample t-test. For testing $H_0 : \theta = 0$ (i.e., $\mu_1 = \mu_2$) versus $H_1 : \theta \neq 0$, it follows from a similar argument used in the derivation of the (two-sided) one-sample t-test that a UMPU test rejects $H_0$ when $|t(X)| > t_{n_1+n_2-2, \alpha/2}$ (exercise). This is the (two-sided) two-sample t-test. The power function of a two-sample t-test is related to a noncentral t-distribution.
Likelihood ratio
When both $H_0$ and $H_1$ are simple (i.e., $\Theta_0 = \{\theta_0\}$ and $\Theta_1 = \{\theta_1\}$), Theorem 6.1 applies and a UMP test rejects $H_0$ when
\[
\frac{f_{\theta_1}(X)}{f_{\theta_0}(X)} > c_0
\]
for some $c_0 > 0$.
The following definition is a natural extension of this idea.

Definition 6.2
Let $\ell(\theta) = f_\theta(X)$ be the likelihood function. For testing $H_0 : \theta \in \Theta_0$ versus $H_1 : \theta \in \Theta_1$, a likelihood ratio (LR) test is any test that rejects $H_0$ if and only if $\lambda(X) < c$, where $c \in [0, 1]$ and $\lambda(X)$ is the likelihood ratio defined by
\[
\lambda(X) = \sup_{\theta \in \Theta_0} \ell(\theta) \bigg/ \sup_{\theta \in \Theta} \ell(\theta).
\]
Discussions

If $\lambda(X)$ is well defined, then $\lambda(X) \leq 1$. The rationale behind LR tests is that when $H_0$ is true, $\lambda(X)$ tends to be close to 1, whereas when $H_1$ is true, $\lambda(X)$ tends to be away from 1. If there is a sufficient statistic, then $\lambda(X)$ depends only on the sufficient statistic. LR tests are as widely applicable as MLE’s in §4.4 and, in fact, they are closely related to MLE’s.

If $\hat{\theta}$ is an MLE of $\theta$ and $\hat{\theta}_0$ is an MLE of $\theta$ subject to $\theta \in \Theta_0$ (i.e., $\Theta_0$ is treated as the parameter space), then

$$\lambda(X) = \ell(\hat{\theta}_0)/\ell(\hat{\theta}).$$

For a given $\alpha \in (0, 1)$, if there exists a $c_\alpha \in [0, 1]$ such that

$$\sup_{\theta \in \Theta_0} P_\theta(\lambda(X) < c_\alpha) = \alpha,$$

then an LR test of size $\alpha$ can be obtained. Even when the c.d.f. of $\lambda(X)$ is continuous or randomized LR tests are introduced, it is still possible that such a $c_\alpha$ does not exist.
Optimality

When a UMP or UMPU test exists, an LR test is often the same as this optimal test.

Proposition 6.5

Suppose that $X$ has a p.d.f. in a one-parameter exponential family:

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

w.r.t. a $\sigma$-finite measure $\nu$, where $\eta$ is a strictly increasing and differentiable function of $\theta$.

(i) For testing $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$, there is an LR test whose rejection region is the same as that of the UMP test $T_*$ given in Theorem 6.2.

(ii) For testing $H_0 : \theta \leq \theta_1$ or $\theta \geq \theta_2$ versus $H_1 : \theta_1 < \theta < \theta_2$, there is an LR test whose rejection region is the same as that of the UMP test $T_*$ given in Theorem 6.3.

(iii) For testing the other two-sided hypotheses, there is an LR test whose rejection region is equivalent to $Y(X) < c_1$ or $Y(X) > c_2$ for some constants $c_1$ and $c_2$. 

Proof
We prove (i) only.
Let $\hat{\theta}$ be the MLE of $\theta$.
Note that $\ell(\theta)$ is increasing when $\theta \leq \hat{\theta}$ and decreasing when $\theta > \hat{\theta}$.
Thus,
\[
\lambda(X) = \begin{cases} 
1 & \hat{\theta} \leq \theta_0 \\
\frac{\ell(\theta_0)}{\ell(\hat{\theta})} & \hat{\theta} > \theta_0.
\end{cases}
\]

Then $\lambda(X) < c$ is the same as $\hat{\theta} > \theta_0$ and $\ell(\theta_0)/\ell(\hat{\theta}) < c$.
From the property of exponential families, $\hat{\theta}$ is a solution of the likelihood equation
\[
\frac{\partial \log \ell(\theta)}{\partial \theta} = \eta'(\theta) Y(X) - \xi'(\theta) = 0
\]
and $\psi(\theta) = \xi'(\theta)/\eta'(\theta)$ has a positive derivative $\psi'(\theta)$.
Since $\eta'(\hat{\theta}) Y - \xi'(\hat{\theta}) = 0$, $\hat{\theta}$ is an increasing function of $Y$ and $\frac{d\hat{\theta}}{dY} > 0$. 
Proof (continued)

Consequently, for any $\theta_0 \in \Theta$,

$$
\frac{d}{dY} \left[ \log \ell(\hat{\theta}) - \log \ell(\theta_0) \right] = \frac{d}{dY} \left[ \eta(\hat{\theta}) Y - \xi(\hat{\theta}) - \eta(\theta_0) Y + \xi(\theta_0) \right]
$$

$$
= \frac{d\hat{\theta}}{dY} \eta'(\hat{\theta}) Y + \eta(\hat{\theta}) - \frac{d\hat{\theta}}{dY} \xi'(\hat{\theta}) - \eta(\theta_0)
$$

$$
= \frac{d\hat{\theta}}{dY} [\eta'(\hat{\theta}) Y - \xi'(\hat{\theta})] + \eta(\hat{\theta}) - \eta(\theta_0)
$$

$$
= \eta(\hat{\theta}) - \eta(\theta_0),
$$

which is positive (or negative) if $\hat{\theta} > \theta_0$ (or $\hat{\theta} < \theta_0$), i.e., $\log \ell(\hat{\theta}) - \log \ell(\theta_0)$ is strictly increasing in $Y$ when $\hat{\theta} > \theta_0$ and strictly decreasing in $Y$ when $\hat{\theta} < \theta_0$.

Hence, for any $d \in \mathcal{R}$, $\hat{\theta} > \theta_0$ and $\ell(\theta_0)/\ell(\hat{\theta}) < c$ is equivalent to $Y > d$ for some $c \in (0, 1)$. 
Example 6.20

Consider the testing problem $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$ based on i.i.d. $X_1, \ldots, X_n$ from the uniform distribution $U(0, \theta)$. We now show that the UMP test with rejection region $X_{(n)} > \theta_0$ or $X_{(n)} \leq \theta_0 \alpha^{1/n}$ given in Exercise 19(c) is an LR test. Note that $\ell(\theta) = \theta^{-n} I_{(X_{(n)}, \infty)}(\theta)$. Hence

$$\lambda(X) = \begin{cases} 
(X_{(n)}/\theta_0)^n & X_{(n)} \leq \theta_0 \\
0 & X_{(n)} > \theta_0 
\end{cases}$$

and $\lambda(X) < c$ is equivalent to $X_{(n)} > \theta_0$ or $X_{(n)}/\theta_0 < c^{1/n}$. Taking $c = \alpha$ ensures that the LR test has size $\alpha$.

Example 6.21

Consider normal linear model $X = N_n(Z\beta, \sigma^2 I_n)$ and the hypotheses $H_0 : L\beta = 0$ versus $H_1 : L\beta \neq 0$, where $L$ is an $s \times p$ matrix of rank $s \leq r$ and all rows of $L$ are in $\mathcal{R}(Z)$. 
The likelihood function in this problem is
\[ \ell(\theta) = \left( \frac{1}{2\pi\sigma^2} \right)^{n/2} \exp \left\{ -\frac{1}{2\sigma^2} \|X - Z\beta\|^2 \right\}, \quad \theta = (\beta, \sigma^2). \]
Since \( \|X - Z\beta\|^2 \geq \|X - Z\hat{\beta}\|^2 \) for any \( \beta \) and the LSE \( \hat{\beta} \),
\[ \ell(\theta) \leq \left( \frac{1}{2\pi\sigma^2} \right)^{n/2} \exp \left\{ -\frac{1}{2\sigma^2} \|X - Z\hat{\beta}\|^2 \right\}. \]
Treating the right-hand side of this expression as a function of \( \sigma^2 \), it is easy to show that it has a maximum at \( \sigma^2 = \hat{\sigma}^2 = \|X - Z\hat{\beta}\|^2 / n \) and
\[ \sup_{\theta \in \Theta} \ell(\theta) = (2\pi\hat{\sigma}^2)^{-n/2} e^{-n/2}. \]
Similarly, let \( \hat{\beta}_{H_0} \) be the LSE under \( H_0 \) and \( \hat{\sigma}^2_{H_0} = \|X - Z\hat{\beta}_{H_0}\|^2 / n \). Then
\[ \sup_{\theta \in \Theta_0} \ell(\theta) = (2\pi\hat{\sigma}^2_{H_0})^{-n/2} e^{-n/2}. \]
Thus,
\[ \lambda(X) = \left( \hat{\sigma}^2 / \hat{\sigma}^2_{H_0} \right)^{n/2} = \left( \frac{\|X - Z\hat{\beta}\|^2}{\|X - Z\hat{\beta}_{H_0}\|^2} \right)^{n/2}. \]
For a two-sample problem, we let $n = n_1 + n_2$, $\beta = (\mu_1, \mu_2)$, and 

$$Z = \begin{pmatrix} J_{n_1} & 0 \\ 0 & J_{n_2} \end{pmatrix}.$$ 

Testing $H_0 : \mu_1 = \mu_2$ versus $H_1 : \mu_1 \neq \mu_2$ is the same as testing $H_0 : L\beta = 0$ versus $H_1 : L\beta \neq 0$ with $L = \begin{pmatrix} 1 & -1 \end{pmatrix}$.

Since $\hat{\beta}_{H_0} = \bar{X}$ and $\hat{\beta} = (\bar{X}_1, \bar{X}_2)$, where $\bar{X}_1$ and $\bar{X}_2$ are the sample means based on $X_1, ..., X_{n_1}$ and $X_{n_1+1}, ..., X_n$, respectively, we have 

$$n\hat{\sigma}^2 = \sum_{i=1}^{n_1} (X_i - \bar{X}_1)^2 + \sum_{i=n_1+1}^{n} (X_i - \bar{X}_2)^2 = (n_1 - 1)S^2_1 + (n_2 - 1)S^2_2$$

and 

$$n\hat{\sigma}^2_{H_0} = (n - 1)S^2 = n^{-1}n_1n_2(\bar{X}_1 - \bar{X}_2)^2 + (n_1 - 1)S^2_1 + (n_2 - 1)S^2_2.$$
Therefore, \( \lambda(X) < c \) is equivalent to \(|t(X)| > c_0\), where

\[
t(X) = \frac{(\bar{X}_2 - \bar{X}_1)/\sqrt{n_1^{-1} + n_2^{-1}}}{\sqrt{[(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2]/(n_1 + n_2 - 2)}},
\]

and LR tests are the same as the two-sample two-sided t-tests in §6.2.3.