Stat 710: Mathematical Statistics
Lecture 21

Jun Shao

Department of Statistics
University of Wisconsin
Madison, WI 53706, USA

Chapter 6. Hypothesis Tests
Lecture 21: UMP tests and Neyman-Pearson lemma

Theory of testing hypotheses

X: a sample from a population P in $\mathscr{P}$, a family of populations.
Based on the observed X, we test a given hypothesis

$$H_0 : P \in \mathscr{P}_0 \quad \text{vs} \quad H_1 : P \in \mathscr{P}_1$$

where $\mathscr{P}_0$ and $\mathscr{P}_1$ are two disjoint subsets of $\mathscr{P}$ and $\mathscr{P}_0 \cup \mathscr{P}_1 = \mathscr{P}$.
A test for a hypothesis is a statistic $T(X)$ taking values in $[0,1]$.
When $X = x$ is observed, we reject (or accept) H_0 with probability $T(x)$ (or $1 - T(x)$).
If $T(X) = 1$ or 0 a.s. $\mathscr{P}$, then $T(X)$ is a nonrandomized test; otherwise $T(X)$ is randomized.
For a given test $T(X)$, the *power function* of $T(X)$ is defined to be

$$\beta_T(P) = E[T(X)], \qquad P \in \mathscr{P},$$

which is the type I error probability of $T(X)$ when $P \in \mathscr{P}_0$ and one minus the type II error probability of $T(X)$ when $P \in \mathscr{P}_1$.

Significance tests

With a sample of a fixed size, we are not able to minimize two error probabilities simultaneously.
Our approach involves maximizing the power $\beta_T(P)$ over all $P \in \mathscr{P}_1$ (i.e., minimizing the type II error probability) and over all tests T satisfying

$$\sup_{P \in \mathscr{P}_0} \beta_T(P) \leq \alpha,$$

where $\alpha \in [0,1]$ is a given level of significance.
The left-hand side of the last expression is defined to be the size of T.

Definition 6.1

A test T_* of size α is a *uniformly most powerful* (UMP) test if and only if $\beta_{T_*}(P) \geq \beta_T(P)$ for all $P \in \mathscr{P}_1$ and T of level α.

Using sufficient statistics

If $U(X)$ is a sufficient statistic for $P \in \mathscr{P}$, then for any test $T(X)$, $E(T|U)$ has the same power function as T and, therefore, to find a UMP test we may consider tests that are functions of U only.

Significance tests

With a sample of a fixed size, we are not able to minimize two error probabilities simultaneously.
Our approach involves maximizing the power $\beta_T(P)$ over all $P \in \mathscr{P}_1$ (i.e., minimizing the type II error probability) and over all tests T satisfying

$$\sup_{P \in \mathscr{P}_0} \beta_T(P) \leq \alpha,$$

where $\alpha \in [0,1]$ is a given level of significance.
The left-hand side of the last expression is defined to be the size of T.

Definition 6.1

A test T_* of size α is a *uniformly most powerful* (UMP) test if and only if $\beta_{T_*}(P) \geq \beta_T(P)$ for all $P \in \mathscr{P}_1$ and T of level α.

Using sufficient statistics

If $U(X)$ is a sufficient statistic for $P \in \mathscr{P}$, then for any test $T(X)$, $E(T|U)$ has the same power function as T and, therefore, to find a UMP test we may consider tests that are functions of U only.

Significance tests

With a sample of a fixed size, we are not able to minimize two error probabilities simultaneously.

Our approach involves maximizing the power $\beta_T(P)$ over all $P \in \mathscr{P}_1$ (i.e., minimizing the type II error probability) and over all tests T satisfying

$$\sup_{P \in \mathscr{P}_0} \beta_T(P) \leq \alpha,$$

where $\alpha \in [0,1]$ is a given level of significance.

The left-hand side of the last expression is defined to be the size of T.

Definition 6.1

A test T_* of size α is a *uniformly most powerful* (UMP) test if and only if $\beta_{T_*}(P) \geq \beta_T(P)$ for all $P \in \mathscr{P}_1$ and T of level α.

Using sufficient statistics

If $U(X)$ is a sufficient statistic for $P \in \mathscr{P}$, then for any test $T(X)$, $E(T|U)$ has the same power function as T and, therefore, to find a UMP test we may consider tests that are functions of U only.

Theorem 6.1 (Neyman-Pearson lemma)

Suppose that $\mathscr{P}_0 = \{P_0\}$ and $\mathscr{P}_1 = \{P_1\}$.
Let f_j be the p.d.f. of P_j w.r.t. a σ-finite measure ν (e.g., $\nu = P_0 + P_1$), $j = 0, 1$.
(i) Existence of a UMP test.
For every α, there exists a UMP test of size α, which is

$$T_*(X) = \begin{cases} 1 & f_1(X) > cf_0(X) \\ \gamma & f_1(X) = cf_0(X) \\ 0 & f_1(X) < cf_0(X), \end{cases}$$

where $\gamma \in [0,1]$ and $c \geq 0$ are some constants chosen so that $E[T_*(X)] = \alpha$ when $P = P_0$ ($c = \infty$ is allowed).
(ii) Uniqueness.
If T_{**} is a UMP test of size α, then

$$T_{**}(X) = \begin{cases} 1 & f_1(X) > cf_0(X) \\ 0 & f_1(X) < cf_0(X) \end{cases} \qquad \text{a.s. } \mathscr{P}.$$

Remarks

- Theorem 6.1 shows that when both H_0 and H_1 are simple (a hypothesis is simple iff the corresponding set of populations contains exactly one element), there exists a UMP test that can be determined by Theorem 6.1 uniquely (a.s. $\mathscr{P}$) except on the set $B = \{x : f_1(x) = cf_0(x)\}$.
- If $\nu(B) = 0$, then we have a unique nonrandomized UMP test; otherwise UMP tests are randomized on the set B and the randomization is necessary for UMP tests to have the given size α
- We can always choose a UMP test that is constant on B.

Proof of Theorem 6.1

The proof for the case of $\alpha = 0$ or 1 is left as an exercise.

Assume now that $0 < \alpha < 1$.

(i) We first show that there exist γ and c such that $E_0[T_*(X)] = \alpha$, where E_j is the expectation w.r.t. P_j.

Let $\gamma(t) = P_0(f_1(X) > tf_0(X))$.

Then $\gamma(t)$ is nonincreasing, $\gamma(0) = 1$, and $\gamma(\infty) = 0$ (why?).

Remarks

- Theorem 6.1 shows that when both H_0 and H_1 are simple (a hypothesis is simple iff the corresponding set of populations contains exactly one element), there exists a UMP test that can be determined by Theorem 6.1 uniquely (a.s. $\mathscr{P}$) except on the set $B = \{x : f_1(x) = cf_0(x)\}$.
- If $\nu(B) = 0$, then we have a unique nonrandomized UMP test; otherwise UMP tests are randomized on the set B and the randomization is necessary for UMP tests to have the given size α
- We can always choose a UMP test that is constant on B.

Proof of Theorem 6.1

The proof for the case of $\alpha = 0$ or 1 is left as an exercise.
Assume now that $0 < \alpha < 1$.
(i) We first show that there exist γ and c such that $E_0[T_*(X)] = \alpha$, where E_j is the expectation w.r.t. P_j.
Let $\gamma(t) = P_0(f_1(X) > tf_0(X))$.
Then $\gamma(t)$ is nonincreasing, $\gamma(0) = 1$, and $\gamma(\infty) = 0$ (why?).

Proof (continued)

Thus, there exists a $c \in (0,\infty)$ such that $\gamma(c) \le \alpha \le \gamma(c-)$.
Set

$$\gamma = \begin{cases} \frac{\alpha-\gamma(c)}{\gamma(c-)-\gamma(c)} & \gamma(c-) \ne \gamma(c) \\ 0 & \gamma(c-) = \gamma(c). \end{cases}$$

Note that $\gamma(c-)-\gamma(c) = P(f_1(X) = cf_0(X))$.
Hence

$$E_0[T_*(X)] = P_0\big(f_1(X) > cf_0(X)\big) + \gamma P_0\big(f_1(X) = cf_0(X)\big) = \alpha.$$

Next, we show that T_* is a UMP test.
Suppose that $T(X)$ is a test satisfying $E_0[T(X)] \le \alpha$.
If $T_*(x) - T(x) > 0$, then $T_*(x) > 0$ and $f_1(x) \ge cf_0(x)$.
If $T_*(x) - T(x) < 0$, then $T_*(x) < 1$ and $f_1(x) \le cf_0(x)$.
In any case,

$$[T_*(x) - T(x)][f_1(x) - cf_0(x)] \ge 0$$

and, therefore,

Proof (continued)

$$\int [T_*(x) - T(x)][f_1(x) - cf_0(x)]d\nu \geq 0,$$

i.e.,

$$\int [T_*(x) - T(x)]f_1(x)d\nu \geq c \int [T_*(x) - T(x)]f_0(x)d\nu.$$

The left-hand side is $E_1[T_*(X)] - E_1[T(X)]$ and the right-hand side is

$$c\{E_0[T_*(X)] - E_0[T(X)]\} = c\{\alpha - E_0[T(X)]\} \geq 0.$$

This proves the result in (i).
(ii) Let $T_{**}(X)$ be a UMP test of size α.
Define

$$A = \{x : T_*(x) \neq T_{**}(x),\ \ f_1(x) \neq cf_0(x)\}.$$

Then $[T_*(x) - T_{**}(x)][f_1(x) - cf_0(x)] > 0$ when $x \in A$ and $= 0$ when $x \in A^c$, and

$$\int [T_*(x) - T_{**}(x)][f_1(x) - cf_0(x)]d\nu = 0,$$

since both T_* and T_{**} are UMP tests of size α.
By Proposition 1.6(ii), $\nu(A) = 0$.
This proves the result.

Example 6.1

Suppose that X is a sample of size 1, $\mathscr{P}_0 = \{P_0\}$, and $\mathscr{P}_1 = \{P_1\}$, where P_0 is $N(0,1)$ and P_1 is the double exponential distribution $DE(0,2)$ with the p.d.f. $4^{-1}e^{-|x|/2}$.

Since $P(f_1(X) = cf_0(X)) = 0$, there is a unique nonrandomized UMP test.

By Theorem 6.1, the UMP test $T_*(x) = 1$ if and only if $\frac{\pi}{8}e^{x^2-|x|} > c^2$ for some $c > 0$, which is equivalent to $|x| > t$ or $|x| < 1-t$ for some $t > \frac{1}{2}$.

Suppose that $\alpha < \frac{1}{3}$. To determine t, we use

$$\alpha = E_0[T_*(X)] = P_0(|X| > t) + P_0(|X| < 1-t).$$

If $t \le 1$, then $P_0(|X| > t) \ge P_0(|X| > 1) = 0.3374 > \alpha$.

Hence t should be larger than 1 and

$$\alpha = P_0(|X| > t) = \Phi(-t) + 1 - \Phi(t).$$

Thus, $t = \Phi^{-1}(1-\alpha/2)$ and $T_*(X) = I_{(t,\infty)}(|X|)$.

Note that it is not necessary to find out what c is.

Example 6.1 (continued)

Intuitively, the reason why the UMP test in this example rejects H_0 when $|X|$ is large is that the probability of getting a large $|X|$ is much higher under H_1 (i.e., P is the double exponential distribution $DE(0,2)$).

The power of T_* when $P \in \mathscr{P}_1$ is

$$E_1[T_*(X)] = P_1(|X| > t) = 1 - \frac{1}{4}\int_{-t}^{t} e^{-|x|/2}dx = e^{-t/2}.$$

Example 6.2

Let $X_1,...,X_n$ be i.i.d. binary random variables with $p = P(X_1 = 1)$. Suppose that $H_0 : p = p_0$ and $H_1 : p = p_1$, where $0 < p_0 < p_1 < 1$. By Theorem 6.1, a UMP test of size α is

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

where $Y = \sum_{i=1}^{n} X_i$ and

$$\lambda(Y) = \left(\frac{p_1}{p_0}\right)^{Y}\left(\frac{1-p_1}{1-p_0}\right)^{n-Y}.$$

Example 6.1 (continued)

Intuitively, the reason why the UMP test in this example rejects H_0 when $|X|$ is large is that the probability of getting a large $|X|$ is much higher under H_1 (i.e., P is the double exponential distribution $DE(0,2)$). The power of T_* when $P \in \mathscr{P}_1$ is

$$E_1[T_*(X)] = P_1(|X| > t) = 1 - \frac{1}{4}\int_{-t}^{t} e^{-|x|/2} dx = e^{-t/2}.$$

Example 6.2

Let $X_1, ..., X_n$ be i.i.d. binary random variables with $p = P(X_1 = 1)$. Suppose that $H_0 : p = p_0$ and $H_1 : p = p_1$, where $0 < p_0 < p_1 < 1$. By Theorem 6.1, a UMP test of size α is

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

where $Y = \sum_{i=1}^{n} X_i$ and

$$\lambda(Y) = \left(\frac{p_1}{p_0}\right)^{Y} \left(\frac{1-p_1}{1-p_0}\right)^{n-Y}.$$

Example 6.2 (continued)

Since $\lambda(Y)$ is increasing in Y, there is an integer $m > 0$ such that

$$T_*(Y) = \begin{cases} 1 & Y > m \\ \gamma & Y = m \\ 0 & Y < m, \end{cases}$$

where m and γ satisfy $\alpha = E_0[T_*(Y)] = P_0(Y > m) + \gamma P_0(Y = m)$.
Since Y has the binomial distribution $Bi(p,n)$, we can determine m and γ from

$$\alpha = \sum_{j=m+1}^{n} \binom{n}{j} p_0^j (1-p_0)^{n-j} + \gamma \binom{n}{m} p_0^m (1-p_0)^{n-m}.$$

Unless

$$\alpha = \sum_{j=m+1}^{n} \binom{n}{j} p_0^j (1-p_0)^{n-j}$$

for some integer m, in which case we can choose $\gamma = 0$, the UMP test T_* is a randomized test.

Remark

An interesting phenomenon in Example 6.2 is that the UMP test T_* does not depend on p_1.
In such a case, T_* is in fact a UMP test for testing $H_0 : p = p_0$ versus $H_1 : p > p_0$.

Lemma 6.1

Suppose that there is a test T_* of size α such that for every $P_1 \in \mathscr{P}_1$, T_* is UMP for testing H_0 versus the hypothesis $P = P_1$.
Then T_* is UMP for testing H_0 versus H_1.

Proof

T_* is a test since it does not depend on P_1.
For any test T of level α, T is also of level α for testing H_0 versus the hypothesis $P = P_1$ with any $P_1 \in \mathscr{P}_1$.
Hence $\beta_{T_*}(P_1) \geq \beta_T(P_1)$.
Since P_1 is arbitrary, this proves that T_* is UMP for testing H_0 versus H_1.

Remark

An interesting phenomenon in Example 6.2 is that the UMP test T_* does not depend on p_1.
In such a case, T_* is in fact a UMP test for testing $H_0 : p = p_0$ versus $H_1 : p > p_0$.

Lemma 6.1

Suppose that there is a test T_* of size α such that for every $P_1 \in \mathscr{P}_1$, T_* is UMP for testing H_0 versus the hypothesis $P = P_1$.
Then T_* is UMP for testing H_0 versus H_1.

Proof

T_* is a test since it does not depend on P_1.
For any test T of level α, T is also of level α for testing H_0 versus the hypothesis $P = P_1$ with any $P_1 \in \mathscr{P}_1$.
Hence $\beta_{T_*}(P_1) \geq \beta_T(P_1)$.
Since P_1 is arbitrary, this proves that T_* is UMP for testing H_0 versus H_1.

Remark

An interesting phenomenon in Example 6.2 is that the UMP test T_* does not depend on p_1.
In such a case, T_* is in fact a UMP test for testing $H_0 : p = p_0$ versus $H_1 : p > p_0$.

Lemma 6.1

Suppose that there is a test T_* of size α such that for every $P_1 \in \mathscr{P}_1$, T_* is UMP for testing H_0 versus the hypothesis $P = P_1$.
Then T_* is UMP for testing H_0 versus H_1.

Proof

T_* is a test since it does not depend on P_1.
For any test T of level α, T is also of level α for testing H_0 versus the hypothesis $P = P_1$ with any $P_1 \in \mathscr{P}_1$.
Hence $\beta_{T_*}(P_1) \geq \beta_T(P_1)$.
Since P_1 is arbitrary, this proves that T_* is UMP for testing H_0 versus H_1.