Hypothesis Testing: Uniformly Most Powerful Tests

In this document, we will consider a simple versus composite hypothesis of the form:

$$
\begin{aligned}
H_0 &: \theta = \theta_0 \\
H_a &: \theta \in \Omega \setminus \{\theta_0\}.
\end{aligned}
$$

(H_a will usually be one of: $\theta \neq \theta_0$, $\theta < \theta_0$, or $\theta . \theta_0$.)

A Note about "Uniformly":

In mathematics, the word "uniformly" can be loosely translated to "for all ____". (In this case, we're going to be doing something "for all" θ.)

For example, you could have a convergence of functions $f_n(x) \overset{n\to\infty}{\to} f(x)$ for one particular x, or you could say that $f_n \overset{n\to\infty}{\to} f$ uniformly in x which means that the convergence holds for all arguments x.

Definition:

For simple H_0 and composite H_a, the critical region C is a uniformly most powerful critical region of size α if C is the most powerful (best) critical region for testing H_0 against each simple hypotheses in H_a.

The corresponding test is a uniformly most powerful (UMP) test with level of significance α for testing the simple H_0 versus the composite (or simple) H_a.

Comments:

1. A UMP test may not exist. (There is usually trouble with the two-sided alternative hypothesis $H_a : \theta \neq \theta_0$.)
2. H_a could be simple. Then the most powerful or best test described in the last handout (for simple versus simple) is, by default, uniformly most powerful.
3. A UMP test may be easily defined for the composite versus composite case:

   $$H_0 : \theta \in \Omega_0 \qquad \text{versus} \qquad H_a : \theta \in \Omega \setminus \Omega_0$$

   C^* is a UMP critical region of size α if

   $$\alpha = \max_{\theta \in \Omega_0} K_{C^*}(\theta)$$

   and

   $$K_{C^*}(\theta) \geq K_C(\theta) \qquad \forall \theta \in \Omega \setminus \Omega_0$$

   and for all critical regions C of size α.

Question: How does one find a UMP test?

Answer: One possible technique is to derive the Neyman-Pearson test for a particular alternative value and then to show that the test does not depend on the specific alternative value.

Example 1: Let $X_1, X_2, \ldots, X_n$ be a random sample form the exponential distribution with rate θ.

Suppose that we want to test,

$$\begin{aligned} H_0 &: \theta = \theta_0 \\ H_a &: \theta > \theta_0. \end{aligned}$$

We could consider the simple versus simple hypothesis:

$$\begin{aligned} H_0 &: \theta = \theta_0 \\ H_a &: \theta = \theta_a, \end{aligned}$$

for some $\theta_a > \theta_0$.

This is exactly what we tested in Example 1 of the previous handout ("Hypothesis Testing: Most Powerful (Best) Tests").

A most powerful or best test for this simple versus simple hypothesis is to reject H_0 if

$$\sum X_i \leq \chi^2_\alpha(2n)/(2\theta_0).$$

The intermediate steps may have depended on θ_a, but the ultimate decision rule does not. Recall though that it was necessary that $\theta_a > \theta_0$ or else the inequality in the decision rule would have flipped.

So, basically, we get this best test for any $\theta_a > \theta_0$. So, by definition of a UMP test, this decision rule gives us the UMP test of $H_0 : \theta = \theta_0$ versus $H_a : \theta > \theta_0$.

Example 2: Let $X_1, X_2, \ldots, X_n$ be a random sample from the distribution with pdf

$$f(x;\theta) = \frac{3x^2}{\theta}\, e^{-x^3/\theta}\, I_{(0,\infty)}(x).$$

Let's find a UMP test for

$$\begin{aligned} H_0 &: \theta = \theta_0 \\ H_a &: \theta > \theta_0 \end{aligned}$$

We begin by reducing the problem to the simple versus simple hypotheses:

$$\begin{aligned} H_0 &: \theta = \theta_0 \\ H_a &: \theta = \theta_a \end{aligned}$$

for some fixed $\theta_a > \theta_0$.

We then use the Neyman-Pearson Lemma for the simple versus simple hypotheses:

$$\lambda(\vec{x};\theta_a,\theta_a) = \frac{\prod_{i=1}^n \left[\frac{3x_i^2}{\theta_0}\, e^{-x_i^3/\theta_0}\right]}{\prod_{i=1}^n \left[\frac{3x_i^2}{\theta_a}\, e^{-x_i^3/\theta_a}\right]} = \left(\frac{\theta_a}{\theta_0}\right)^n e^{-\left(\frac{\theta_a-\theta_0}{\theta_0\theta_a}\right)\sum x_i^3}$$

We will reject H_0 if $\lambda(\vec{x};\theta_0,\theta_a) \leq k$ where k is such that $P(\lambda(\vec{X};\theta_0,\theta_a) \leq k; H_0) = \alpha$.

$$\begin{aligned} P(\lambda(\vec{X};\theta_0,\theta_a) \leq k; H_0) &= P\left(\left(\tfrac{\theta_a}{\theta_0}\right)^n e^{-\left(\frac{\theta_a-\theta_0}{\theta_0\theta_a}\right)\sum X_i^3} \leq k; H_0\right) \\ &= P\left(\sum X_i^3 \geq -\tfrac{\theta_0\theta_a}{\theta_a-\theta_0} \ln\left[\left(\tfrac{\theta_a}{\theta_0}\right)^n k\right]; H_0\right) \end{aligned}$$

(The inequality flipped along the way since $\theta_a > \theta_0$ had us dividing by a negative when we moved the $-\frac{\theta_a-\theta_0}{\theta_0\theta_a}$ to the right side of the inequality.)

So, the critical region has the form

$$C = \left\{(x_1, x_2, \ldots, x_n) : \sum x_i^3 \geq k_1\right\}.$$

To find the actual critical region, we need to find the k_1 such that

$$P(\sum X_i^3 \geq k_1; H_0) = \alpha.$$

What is the distribution of $\sum X_i^3$ if H_0 is true?

Let's first consider just one of the X's:

Let $Y = X^3$. Then $y = g(x) = x^3 \Rightarrow x = g^{-1}(y)(y) = y^{1/3}$.

Then

$$\begin{aligned} f_Y(y) &= f_X(g^{-1}(y)) \cdot \left|\tfrac{\partial}{\partial y} g^{-1}(y)\right| \\ &= \tfrac{3(y^{1/3})^2}{\theta_0}\, e^{-(y^{1/3})^3/\theta_0} \cdot I_{(0,\infty)}(y^{1/3}) \cdot \left|\tfrac{1}{3} y^{-2/3}\right| \\ &= \tfrac{1}{\theta_0} e^{-y/\theta_0} \cdot I_{(0,\infty)(y)}. \end{aligned}$$

(Note: The absolute value went away because $y^{-2/3} = (y^{-1/3})^2 > 0$. Also, the indicator kept us in a region where $y > 0$.)

So, when H_0 is true (which means that $\theta = \theta_0$) $Y = X^3 \sim exp(rate = 1/\theta_0)$.

Therefore,

$$\sum_{i=1}^{n} X_i^3 \sim \Gamma(n, \theta_0)$$

So,

$$\begin{aligned} \alpha &= P(\sum X_i^3 \geq k_1; H_0) \\ &= P(G \geq k_1) \end{aligned}$$

where $G \sim \Gamma(n, \theta_0)$.

As before, you can integrate the gamma pdf from k_1 to ∞, set that equal to α, and solve for k_1. This is not easy but you might consider a numerical approximation. However, it will depend on α, θ_0, and n and you won't be able to see the algebraic dependence but must settle for a numerical approximation for fixed values of α, θ_0, and n.

OR, as before, we can write k_1 in terms of a chi-squared critical value!

See the last handout ("Hypothesis Testing: Most Powerful (Best) Tests") where we showed that

$$G \sim \Gamma(n, \theta_0) \qquad \Rightarrow \qquad 2G/\theta_0 \sim \chi^2(2n).$$

So,

$$\begin{aligned} \alpha &= P(G \geq k_1) \\ &= P(2G/\theta_0 \geq 2k_1/\theta_0) \\ &= P(W \geq 2k_1/\theta_0) \end{aligned}$$

where $W \sim \chi^2(2n)$.

Therefore

$$2k_1/\theta_0 = \chi^2_\alpha(2n) \qquad \Rightarrow \qquad k_1 = \theta_0 \chi^2_\alpha(2n)/2$$

Let us review where we are so far in this example:

We have determined that the best or most powerful test of $H_0 : \theta = \theta_0$ versus $H_a : \theta = \theta_a$ where $\theta_a > \theta_0$ is to

$$\text{reject } \; H_0 \; \text{ if} \sum X_i^3 \geq \theta_0 \chi^2_\alpha(2n)/2.$$

This test does not depend on the specific value of θ_a, although it did depend on the fact that $\theta_a > \theta_0$ (otherwise an inequality would have flipped). So, it would work for any $\theta_a > \theta_0$.

Hence, the test:

$$\text{reject } \; H_0 \; \text{ if} \sum X_i^3 \geq \theta_0 \chi^2_\alpha(2n)/2,$$

is a UMP test of size α for $H_0 : \theta = \theta_0$ versus $H_a : \theta > \theta_0$.

Example 3:

Finally, let us consider the last example again with the hypotheses: $H_0 : \theta = \theta_0$ versus $H_a : \theta \neq \theta_0$.

- First we consider $H_0 : \theta = \theta_0$ versus $H_a : \theta = \theta_a$ for any $\theta_a > \theta_0$. Since we just did this in the previous example, we know that this leads us to a UMP critical region for $H_0 : \theta = \theta_0$ versus $H_a : \theta > \theta_0$ of the form
$$C = \left\{ (x_1, x_2, \ldots, x_n) : \sum x_i^3 \geq c \right\}$$
where c is the solution of $\alpha = P(G \geq c)$ for $G \sim \Gamma(n, \theta_0)$.
- If we now consider $H_0 : \theta = \theta_0$ versus $H_a : \theta = \theta_a$ for any $\theta_a < \theta_0$, and go through the steps of the previous example, we will end up showing that a UMP critical region for $H_0 : \theta = \theta_0$ versus $H_a : \theta < \theta_0$ of the form
$$C = \left\{ (x_1, x_2, \ldots, x_n) : \sum x_i^3 \leq c \right\}$$
where c is the solution of $\alpha = P(G \leq c)$ for $G \sim \Gamma(n, \theta_0)$.
- Since a best test for $H_0 : \theta = \theta_0$ versus $H_a : \theta = \theta_a$ for $\theta_0 < \theta_a$ is different than a best test for $H_0 : \theta = \theta_0$ versus $H_a : \theta = \theta_a$ for $\theta_0 > \theta_a$, there can't be a uniformly best test (most powerful) for $H_0 : \theta = \theta_0$ versus $H_a : \theta ne \theta_0$.

Therefore, there is no UMP test for this example.

Exponential Families and UMP Tests

Theorem:

Suppose that $X_1, X_2, \ldots, X_n$ have a joint pdf in the one-parameter exponential family:

$$f(\vec{x}; \theta) = a(\theta) b(\vec{x}) \exp\left[c(\theta) d(\vec{x})\right]$$

where $c(\theta)$ is an increasing function of θ. Then

1. A UMP test of size α for $H_0 : \theta = \theta_0$ versus $H_a : \theta > \theta_0$ is to reject H_0 if $d(\vec{x}) \geq k$ where k is chosen so that $P(d(\vec{X}) \geq k; \theta_0) = \alpha$.
2. A UMP test of size α for $H_0 : \theta = \theta_0$ versus $H_a : \theta < \theta_0$ is to reject H_0 if $d(\vec{x}) \leq k$ where k is chosen so that $P(d(\vec{X}) \leq k; \theta_0) = \alpha$.

Proof: (of #1)

This is not too difficult to show. For example, to show #1, we consider the simple versus simple hypotheses

$$H_0 : \theta = \theta_0 \qquad \text{versus} \qquad H_a : \theta = \theta_a$$

for $\theta_a > \theta_0$. We apply the N-P Lemma and get that

$$\lambda(\vec{x};\theta_0,\theta_a) = \frac{f(\vec{x};\theta_0)}{f(\vec{x},\theta_a)} = \frac{a(\theta_0)}{a(\theta_a)} \cdot \exp\left[\left(c(\theta_0) - c(\theta_1)\right) d(\vec{x})\right].$$

Setting $\lambda(\vec{x};\theta_0,\theta_a) \leq k$ and "solving down to the x-stuff on the left hand side" will give

$$d(\vec{x}) \geq k_1,$$

and we solve for k_1 by setting $P(d(\vec{X}) \geq k_1) = \alpha$.

Note that the inequality flips at the point when we divide through by $c(\theta_0) - c(\theta_1)$ which is negative since $\theta_a > \theta_0$ and c is, by assumption, increasing in θ.

Finding k_1 will involve the distribution of $d(\vec{X})$ under the assumption that the null hypothesis ($H_0 : theta = \theta_0$) and will not involve θ_a. Hence, this most powerful (best) test provided by the N-P Lemma for $H_0 : \theta = \theta_0$ versus $H_a : \theta = \theta_a$ will be UMP for $H_0 : \theta = \theta_0$ versus $H_a : \theta > \theta_0$.

- Note: If you read and understood the "sketch of proof", it is very easy to see that if $c(\theta)$ is decreasing in θ, all the inequalities directly following $d(\vec{x})$ and $d(\vec{X})$ in the statement of the theorem would flip.
- Note: If we have $H_0 : \theta \leq \theta_0$ in #1 or $H_0 : \theta \geq \theta_0$ in #2, nothing would change because we could show (and we will on the homework!), under the assumptions of this theorem, that for #1

  $$\alpha = \max_{\theta \leq \theta_0} P(d(\vec{X}) \geq k; \theta) = P(d(\vec{X}) \geq k; \theta_0)$$

  and, for #2

  $$\alpha = \max_{\theta \geq \theta_0} P(d(\vec{X}) \geq k; \theta) = P(d(\vec{X}) \geq k; \theta_0).$$

Example 4:

Let $X_1, X_2, \ldots, X_n$ be a random sample from the $N(0, \sigma^2)$ distribution.

Find a UMP test of size α for $H_0 : \sigma = \sigma_0$ versus $H_a : \sigma > \sigma_0$.

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

$$\Rightarrow f(\vec{x};\sigma) = (2\pi)^{-n/2}(\sigma^2)^{-n/2} \cdot \exp\left[-\frac{1}{2\sigma^2}\sum x_i^2\right]$$

$$\Rightarrow c(\sigma) = -\frac{1}{2\sigma^2} \qquad \text{and} \qquad d(\vec{x}) = \sum x_i^2.$$

Note that $c(\sigma)$ is an increasing function of σ.

So, a UMP test of size α is to reject H_0 if $\sum X_i^2 \geq k$ where k is chosen such that $P(\sum X_i^2 \geq k; H_0) = \alpha$.

Now let's find k.

When H_0 is true, $\sigma = \sigma_0$. Under this assumption, we need the distribution of $\sum X_i^2$.

A <u>standard</u> normal squared is a chi-squared, so let's standardize it:

$$\begin{aligned}
\alpha &= P(\sum X_i^2 \geq k; H_0) \\
&= P(\sum X_i^2 \geq k; \sigma_0) \\
&= P\left(\frac{\sum X_i^2}{\sigma_0^2} \geq \frac{k}{\sigma_0^2}; \sigma_0\right) \\
&= P\left(\sum \left(\frac{X_i}{\sigma_0}\right)^2 \geq \frac{k}{\sigma_0^2}; \sigma_0\right) \\
&= P(W \geq k_1)
\end{aligned}$$

where $W \sim \chi^2(n)$ and $k_1 = k/\sigma_0^2$.

We know that $k_1 = \chi^2_\alpha(n)$.

So, the UMP test of size α is to

$$\text{reject } H_0 \text{ if } \frac{\sum X_i^2}{\sigma_0^2} \geq \chi^2_\alpha(n)$$

OR

$$\text{reject } H_0 \text{ if } \sum X_i^2 \geq \sigma_0^2 \chi^2_\alpha(n).$$

"Sub-Example":

Suppose we observe a random sample of size 10 from the $N(0, \sigma^2)$ distribution. Further suppose that we wish to test $H_0 : \sigma = 1$ versus $H_a : \sigma > 1$ at 0.05 level of significance.

We will reject H_0 if

$$\frac{\sum_{i=1}^{10} X_i^2}{\sigma_0^2} = \sum_{i=1}^{10} X_i^2 \geq \chi^2_{0.05}(10) = 18.3.$$