MA 3140: Statistical Inference

Dr. Sameen Naqvi
Department of Mathematics, IIT Hyderabad
Email id: sameen@math.iith.ac.in

UMP Unbiased tests

Unbiasedness for Hypothesis Testing

- Note that in the class Φ_α of all size α tests, i.e.,

$$\sup_{\theta \in \Theta_0} \beta^*_\phi(\theta) = \sup_{\theta \in \Theta_0} E_\theta \phi(\boldsymbol{X}) = \alpha,$$

  there does not exist UMP test for many hypotheses.

- A size α test ϕ is said to be unbiased if the power function β^*_ϕ satisfies the condition

$$\begin{aligned} \beta^*_\phi &\leq \alpha, \quad \text{for } \theta \in \Theta_0 \\ \text{and } \beta^*_\phi &\geq \alpha, \quad \text{for } \theta \in \Theta_1. \end{aligned} \tag{1}$$

- Thus, an unbiased test rejects a false H_0 more often than a true H_0.

Unbiasedness for Hypothesis Testing

- Let U_α be the class of all unbiased size-α tests of H_0. If there exists a test $\phi \in U_\alpha$ that has maximum power at each $\theta \in \Theta_1$, it is referred to as UMP-unbiased size α test.

- Clearly, $U_\alpha \subset \Phi_\alpha$.

  If a UMP test exists in Φ_α, it is UMP in U_α.

  This follows on comparing the power of the UMP test with that of the trivial test $\phi(x) \equiv \alpha$.

Theorem 1

Let $\boldsymbol{X} = (X_1, \dots, X_n)$ be a random vector with pdf

$$f(\boldsymbol{x}, \theta) = c(\theta) e^{\theta T(\boldsymbol{x})} h(\boldsymbol{x}), \;\; \theta \in \Theta \subset \mathbb{R}.$$

Then for testing

$$H_3 : \theta_1 \leq \theta \leq \theta_2 \quad \text{vs} \quad K_3 : \theta < \theta_1 \text{ or } \theta > \theta_2$$

there exists a UMP unbiased test given by

$$\phi(\boldsymbol{x}) = \begin{cases} 1, & \text{if } T(\boldsymbol{x}) < c_1 \text{ or } T(\boldsymbol{x}) > c_2 \\ \gamma_i, & \text{if } T(\boldsymbol{x}) = c_i, \;\; i = 1, 2 \\ 0, & \text{if } c_1 < T(\boldsymbol{x}) < c_2 \quad (c_1 < c_2). \end{cases}$$

where c_1, c_2, γ_1 γ_2 are determined by

$$E_{\theta_1} \phi(\boldsymbol{X}) = E_{\theta_2} \phi(\boldsymbol{X}) = \alpha.$$

Theorem 2

Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be a random vector with pdf

$$f(\boldsymbol{x}, \theta) = c(\theta) e^{\theta T(\boldsymbol{x})} h(\boldsymbol{x}), \ \theta \in \Theta \subset \mathbb{R}.$$

Then for testing

$$H_4 : \theta = \theta_0 \quad \text{vs} \quad K_4 : \theta \neq \theta_0$$

there exists a UMP unbiased test given by

$$\phi(\boldsymbol{x}) = \begin{cases} 1, & \text{if } T(\boldsymbol{x}) < c_1 \text{ or } T(\boldsymbol{x}) > c_2 \\ \gamma_i, & \text{if } T(\boldsymbol{x}) = c_i, \ i = 1, 2 \\ 0, & \text{if} c_1 < T(\boldsymbol{x}) < c_2 \quad (c_1 < c_2). \end{cases}$$

where c_i's and γ_i's are determined by

$$E_{\theta_0} \phi(\boldsymbol{X}) = \alpha; \quad E_{\theta_0}(T(\boldsymbol{X}) \phi(\boldsymbol{X})) = \alpha E_{\theta_0} T(\boldsymbol{X}).$$

Example 1

Let $X \sim Bin(n, p)$. We want to test

$$H_4 : p = p_0 \quad \text{vs} \quad K_4 : p \neq p_0.$$

Solution: The pmf

$$f(x, p) = \binom{n}{x} p^x (1-p)^{n-x} = \binom{n}{x} (1-p)^n e^{x \log \frac{p}{1-p}}$$

belongs to the exponential family with $T(x) = x$.

The test is

$$\phi(x) = \begin{cases} 1, & \text{if } x < c_1 \text{ or } x > c_2 \\ \gamma_i, & \text{if } x = c_i, \; i = 1, 2 \\ 0, & \text{if } c_1 < x < c_2. \end{cases}$$

Example 1 contd.

The constants c_i's and γ_i's are determined by

$$E_{p_0}\phi(X) = \alpha; \tag{2}$$

$$E_{p_0}\{X\phi(X)\} = \alpha E_{p_0}(X). \tag{3}$$

Condition (2) can be written as

$$P_{p_0}(X < c_1 \text{ or } X > c_2) + \gamma_1 P_{p_0}(X = c_1) + \gamma_2 P_{p_0}(X = c_2) = \alpha$$

$$\iff P(c_1 < X < c_2) + (1 - \gamma_1)P(X = c_1) + (1 - \gamma_2)P(X = c_2) = 1 - \alpha$$

$$\iff \sum_{x=c_1+1}^{c_2-1} \binom{n}{x} p_0^x (1 - p_0)^{n-x} + \sum_{i=1}^{2} (1 - \gamma_i) \binom{n}{c_i} p_0^{c_i} (1 - p_0)^{n-c_i} = 1 - \alpha$$

The LHS can be determined from the tables of Binomial distribution

Example 1 contd.

Condition (3) can be written as

$$E_{p_0}X(1-\phi(X)) = (1-\alpha)E_{p_0}X = (1-\alpha)np_0$$

$$\iff \sum_{x=c_1+1}^{c_2-1} x\binom{n}{x}p_0^x(1-p_0)^{n-x}$$

$$+\sum_{i=1}^{2}(1-\gamma_i)c_i\binom{n}{c_i}p_0^{c_i}(1-p_0)^{n-c_i} = (1-\alpha)np_0$$

$$\iff \sum_{x=c_1+1}^{c_2-1} \binom{n-1}{x-1}p_0^{x-1}(1-p_0)^{n-1-(x-1)}$$

$$+\sum_{i=1}^{2}(1-\gamma_i)\binom{n-1}{c_i-1}p_0^{c_i-1}(1-p_0)^{n-1-(c_i-1)} = 1-\alpha$$

Example 1 contd.

on using

$$x\binom{n}{x}p_0^x(1-p_0)^{n-x} = np_0\binom{n-1}{x-1}p_0^{x-1}(1-p_0)^{n-1-(x-1)}.$$

The LHS can be determined from the tables of Binomial distribution.

- Consider $n = 10, p_0 = 1/2$. Then the first condition is:

$$\sum_{x=c_1+1}^{c_2-1}\binom{10}{x}\left(\frac{1}{2}\right)^{10} + \sum_{i=1}^{2}(1-\gamma_i)\binom{10}{c_i}\left(\frac{1}{2}\right)^{10} = 0.9$$

$$\sum_{x=c_1+1}^{c_2-1}\binom{10}{x} + \sum_{i=1}^{2}(1-\gamma_i)\binom{10}{c_i} = 2^{10}\times 0.9,$$

Example 1 contd.

- and the second condition is:

$$\sum_{x=c_1+1}^{c_2-1} \binom{9}{x-1} \left(\frac{1}{2}\right)^9 + \sum_{i=1}^{2} (1-\gamma_i) \binom{9}{c_i - 1} \left(\frac{1}{2}\right)^9 = 0.9$$

$$\sum_{x=c_1}^{c_2-2} \binom{9}{y} + \sum_{i=1}^{2} (1-\gamma_i) \binom{9}{c_i - 1} = 2^9 \times 0.9$$

These conditions can now be solved using binomial coefficients to get values of c_1, c_2, γ_1 and γ_2.

Example 2

Let $X_1, \ldots, X_n \sim N(0, \sigma^2)$. Find UMP unbiased test for

$$H_4 : \sigma^2 = \sigma_0^2 \quad \text{vs} \quad K_4 : \sigma^2 \neq \sigma_0^2.$$

Solution: The pmf

$$f(\boldsymbol{x}, \sigma) = \frac{1}{(\sigma\sqrt{2\pi})^n} \exp^{-\sum x_i^2/2\sigma^2}$$

belongs to the exponential family with $T(x) = \sum x_i^2$.
The UMP unbiased test is

$$\phi(\boldsymbol{x}) = \begin{cases} 1, & \text{if } \sum x_i^2 < c_1 \text{ or } \sum x_i^2 > c_2 \\ 0, & \text{if } c_1 \leq \sum x_i^2 \leq c_2. \end{cases}$$

Here, γ_i's are 0 as X_i's are constant. Also, recall that $\sum X_i^2/\sigma^2 \sim \chi_n^2$.

Example 2 contd.

Thus, the test becomes

$$\phi(\boldsymbol{x}) = \begin{cases} 1, & \text{if } \frac{\sum x_i^2}{\sigma_0^2} < c_1 \text{ or } \frac{\sum x_i^2}{\sigma_0^2} > c_2 \\ 0, & \text{if } c_1 \leq \frac{\sum x_i^2}{\sigma_0^2} \leq c_2. \end{cases}$$

where c_i's are determined by

$$E_{\sigma_0}\phi(\boldsymbol{X}) = \alpha; \tag{4}$$

$$E_{\sigma_0}\frac{T(\boldsymbol{X})}{\sigma_0^2}\,\phi(\boldsymbol{X}) = \alpha\, E_{\sigma_0}\frac{T(\boldsymbol{X})}{\sigma_0^2}. \tag{5}$$

Here,

$$E_{\sigma_0}(1-\phi(\boldsymbol{X})) = 1-\alpha \Longrightarrow P(c_1 \leq W \leq c_2) = 1-\alpha,\ W \sim \chi_n^2.$$

Example 2 contd.

Also,

$$
\begin{aligned}
& E_{\sigma_0} \frac{T(\boldsymbol{X})}{\sigma_0^2}(1-\phi(\boldsymbol{X})) = (1-\alpha)\ E_{\sigma_0} \frac{T(\boldsymbol{X})}{\sigma_0^2} \\
\Longleftrightarrow \quad & E_{\sigma_0} W(1-\phi(\boldsymbol{X})) = (1-\alpha)\ E_{\sigma_0} W \\
\Longleftrightarrow \quad & \int_{c_1}^{c_2} w g_n(w) dw = (1-\alpha) n \\
\Longleftrightarrow \quad & \int_{c_1}^{c_2} g_{n+2}(w) dw = 1-\alpha,
\end{aligned}
$$

since $w\ g_n = w\ \frac{1}{2^{n/2}\Gamma n/2} \exp^{-w/2} w^{n/2-1} = n\ g_{n+2}$.

Example 2 contd.

So, the 2 conditions are

$$\int_{c_1}^{c_2} g_n(w)dw = 1 - \alpha \text{ and } \int_{c_1}^{c_2} g_{n+2}(w)dw = 1 - \alpha.$$

Integrating by parts in the second condition and use the first condition, we get

$$c_1^{n/2} e^{-c_1/2} = c_2^{n/2} e^{-c_2/2}.$$

These values of c_1 and c_2 are tabulated by Pacheres (1961, AMS).
If n is large and σ_0 is not close to 0 or ∞, we can take an approximate test as

$$c_1 = \chi^2_{n,1-\alpha/2} \qquad c_2 = \chi^2_{n,\alpha/2}.$$

Thanks for your patience!