MA 3140: Statistical Inference

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

Example 3

Let $X_1, \ldots, X_n$ be a random sample from an exponential distribution with density $\frac{1}{\sigma} e^{-x/\sigma}$, $x > 0$, $\sigma > 0$. Find the MP test for testing

$$H_0 : \sigma = \sigma_0 \quad vs. \quad H_1 : \sigma = \sigma_1$$

Case 1: $\sigma_1 > \sigma_0$

Solution: The joint density of $X_1, \ldots, X_n$ under H_0 and H_1 are:

$$f_0(\boldsymbol{x}) = \frac{1}{(\sigma_0)^n} e^{-\sum x_i/\sigma_0}$$

$$f_1(\boldsymbol{x}) = \frac{1}{(\sigma_1)^n} e^{-\sum x_i/\sigma_1}$$

Example 3 contd.

NP Lemma gives the form of the MP test

$$\text{Reject } H_0 \text{ if } \frac{f_1(\boldsymbol{x})}{f_0(\boldsymbol{x})} \geq k$$

where k is determined by the size condition.

This is equivalent to

$$\begin{aligned} \frac{f_1(\boldsymbol{x})}{f_0(\boldsymbol{x})} = \left(\frac{\sigma_0}{\sigma_1}\right)^n e^{-\frac{\sum x_i}{\sigma_1} + \frac{\sum x_i}{\sigma_0}} &\geq k \\ \Leftrightarrow \sum x_i \left(\frac{1}{\sigma_0} - \frac{1}{\sigma_1}\right) &\geq k_1 \\ \Leftrightarrow \sum x_i &\geq k_2 \end{aligned}$$

Example 3 contd.

In order to determine k_2, we employ the size condition, i.e.,

$$\alpha = P(\text{Type I error}) = P(\text{Rejecting } H_0 \text{ when it is true})$$
$$= P_{\sigma_0}\left(\sum X_i \geq k_2\right)$$

Note that $\frac{X_i}{\sigma_0} \sim e^{-x}$,

$Y = \frac{\sum X_i}{\sigma_0} \sim Gamma(n, 1)$ and $W = 2Y = \frac{2\sum X_i}{\sigma_0} \sim \chi^2_{2n}$.

Thus the MP test is

$$\text{Reject } H_0 \text{ if } \frac{2\sum X_i}{\sigma_0} \geq \chi^2_{2n,\alpha}$$

Example 3 contd.

Case 2: $\sigma_1 < \sigma_0$

Solution: The test procedure will get modified as follows:

$$\text{Reject } H_0 \text{ if } \sum X_i \leq k_2^*.$$

In order to determine k_2^*, we employ the size condition, i.e.,

$$\alpha = P(\text{Reject } H_0 \text{ when it is true}) = P_{\sigma_0}\left(\frac{2\sum X_i}{\sigma_0} \leq k_3^*\right)$$

$$\implies k_3^* = \chi^2_{2n,1-\alpha}$$

Thus the MP test is

$$\text{Reject } H_0 \text{ if } \frac{\sum X_i^2}{\sigma_0^2} \leq \chi^2_{2n,1-\alpha}.$$

Example 4

Let X be an observation from a density $f(x)$. Find the MP test for testing

$$H_0 : f(x) = f_0(x) \quad vs. \quad H_1 : f(x) = f_1(x)$$

where

$$f_0(x) = \begin{cases} x, & 0 < x \leq 1 \\ 2 - x, & 1 < x \leq 2 \\ 0, & \text{otherwise} \end{cases}; \qquad f_1(x) = \begin{cases} \frac{1}{2}, & 0 < x \leq 2 \\ 0, & \text{otherwise} \end{cases}$$

Solution: NP Lemma gives the form of the MP test

$$\frac{f_1(x)}{f_0(x)} = \begin{cases} \frac{1}{2x}, & 0 < x \leq 1 \\ \frac{1}{2(2-x)}, & 1 < x < 2 \end{cases}$$

Example 4 contd.

Reject H_0

- if $\frac{1}{2x} > k \implies x < \frac{1}{2k}$ where k is determined by

$$P_0\left(0 < X < \frac{1}{2k}\right) = \int_0^{1/2k} x \, dx = \frac{1}{8k^2} \qquad (\because 0 < x \leq 1)$$

- if $\frac{1}{2(2-x)} > k \implies 2 - x < \frac{1}{2k}$ or $x > 2 - \frac{1}{2k}$ where k is determined by

$$P_0\left(X > 2 - \frac{1}{2k}\right) = \int_{2-\frac{1}{2k}}^{2} (2-x)dx = \frac{1}{8k^2} \quad (\because 1 < x < 2)$$

Example 4 contd.

The size condition gives

$$P_0\left(0 < X < \frac{1}{2k}, 0 < X \leq 1\right) + P_0\left(X > 2 - \frac{1}{2k}, 1 < X < 2\right) = \alpha$$

$$\implies \frac{1}{8k^2} + \frac{1}{8k^2} = \alpha \implies \frac{1}{4k^2} = \alpha \implies \frac{1}{2k} = \sqrt{\alpha}$$

Thus, the MP test of size α for testing H_0 against H_1 is

$$\text{Reject } H_0 \text{ if } X < \sqrt{\alpha} \text{ or } X > 2 - \sqrt{\alpha}$$

Example: Suppose $\alpha = 0.01$, $\sqrt{\alpha} = 0.1$.

The test will be Reject H_0 if $X < 0.1$ or $X > 1.9$ else it will accept H_0.

Example 5

Let $X_1, \ldots, X_n \sim Bin(1, p)$. Find the MP test for testing

$$H_0 : p = p_0 \quad vs. \quad H_1 : p = p_1, \quad p_1 > p_0.$$

Solution: The MP size α test of H_0 against H_1 is of the form

$$\phi(\boldsymbol{x}) = \begin{cases} 1, & \lambda(\boldsymbol{x}) = \frac{p_1^{\sum x_i}(1-p_1)^{n-\sum x_i}}{p_0^{\sum x_i}(1-p_0)^{n-\sum x_i}} > k, \\ \gamma, & \lambda(\boldsymbol{x}) = k, \\ 0, & \lambda(\boldsymbol{x}) < k, \end{cases}$$

where k and γ are determined from

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

Example 5 contd.

Note that for $p_1 > p_0$.

$$\lambda(\boldsymbol{x}) = \frac{p_1^{\sum x_i}(1-p_1)^{n-\sum x_i}}{p_0^{\sum x_i}(1-p_0)^{n-\sum x_i}}$$

is an increasing function of $\sum x_i$.

It follows that $\lambda(\boldsymbol{x}) > k$ iff $\sum x_i > k_1$, where k_1 is a constant.

Thus the MP size α test is

$$\phi(\boldsymbol{x}) = \begin{cases} 1, & \sum x_i > k_1, \\ \gamma, & \sum x_i = k_1, \\ 0, & \sum x_i < k_1, \end{cases}$$

Example 5 contd.

where k_1 and γ are determined from

$$\alpha = E_{p_0}\phi(\boldsymbol{X}) = P_{p_0}\left\{\sum X_i > k_1\right\} + \gamma P_{p_0}\left\{\sum X_i = k_1\right\}$$
$$= \sum_{r=k_1+1}^{n} \binom{n}{r} p_0^r (1-p_0)^{n-r} + \gamma \binom{n}{k_1} p_0^{k_1}(1-p_0)^{n-k_1}.$$

- Suppose $n = 5$, $p_0 = 1/2$, $p_1 = 3/4$ and $\alpha = 0.5$, then k and γ are determined from

$$0.05 = \alpha = \sum_{k+1}^{5} \binom{5}{r} \left(\frac{1}{2}\right)^5 + \gamma \binom{5}{k} \left(\frac{1}{2}\right)^5.$$

It follows that $k = 4$ and $\gamma = 0.122$.

Example 5 contd.

Thus, the MP test is

$$\phi(\boldsymbol{x}) = \begin{cases} 1, & \sum x_i > 4, \\ 0.122, & \sum x_i = 4, \\ 0, & \sum x_i < 4. \end{cases}$$

i.e., reject $p = 1/2$ in favor of $p = 3/4$ if $\sum X_i = 5$ and reject $p = 1/2$ with probability 0.122 if $\sum X_i = 4$.

- In case of testing $H_0 : p = p_0 \quad vs. \quad H_1 : p = p_1, \quad p_1 < p_0$, the test form is

$$\phi(\boldsymbol{x}) = \begin{cases} 1, & \sum x_i < k, \\ \gamma, & \sum x_i = k, \\ 0, & \sum x_i > k. \end{cases}$$

Neyman Pearson Lemma & Sufficiency

Theorem: Consider the hypothesis problem posed in previous theorem. Suppose $T(\boldsymbol{X})$ is a sufficient statistic for θ and $g(t|\theta_i)$ is the pdf or pmf of T corresponding to θ_i, $i = 1, 2$.

Then any test based on T with rejection region S is a MP level α test if it satisfies

$$\begin{cases} t \in S, & \text{when } g(t|\theta_1) > kg(t|\theta_0) \\ t \in S^c, & \text{when } g(t|\theta_1) < kg(t|\theta_0) \end{cases} \tag{1}$$

for some $k \geq 0$, where

$$P_{\theta_0}(T \in S) = \alpha. \tag{2}$$

Proof

In terms of the original sample $\boldsymbol{X}$, the test based on T has the rejection region $R = \{\boldsymbol{x} : T(\boldsymbol{x}) \in S\}$.
By Factorization Theorem, $f(\boldsymbol{x}|\theta_i) = g(T(\boldsymbol{x}|\theta_i))h(\boldsymbol{x})$, $i = 1, 2$, for some non-negative function $h(\boldsymbol{x})$.
Multiplying the inequalities in (1) by $h(\boldsymbol{x})$, we see that R satisfies

$$\boldsymbol{x} \in R, \text{ if } f(\boldsymbol{x}|\theta_1) = g(T(\boldsymbol{x}|\theta_1))h(\boldsymbol{x}) > kg(T(\boldsymbol{x}|\theta_0))h(\boldsymbol{x}) = kf(\boldsymbol{x}|\theta_0)$$

and

$$\boldsymbol{x} \in R^c, \text{ if } f(\boldsymbol{x}|\theta_1) = g(T(\boldsymbol{x}|\theta_1))h(\boldsymbol{x}) < kg(T(\boldsymbol{x}|\theta_0))h(\boldsymbol{x}) = kf(\boldsymbol{x}|\theta_0)$$

Also by (2),

$$P_{\theta_0}(\boldsymbol{X} \in R) = P_{\theta_0}(T(\boldsymbol{X}) \in S) = \alpha.$$

Thus, the test based on T is the MP test.

UMP tests

Situations where NP Lemma fails

Let $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$, both parameters are unknown.

$$H_0 : \mu = 0 \quad vs. \quad H_1 : \mu \neq 0$$

Note that both hypotheses are composite. Hence NP Lemma does not give us a MP Test.

In general, we may have a family of distributions $f(x, \theta)$, where we are interested to test

$$H_0 : \theta \leq \theta_0 \quad vs. \quad H_1 : \theta > \theta_0$$

or

$$H_0 : \theta \geq \theta_0 \quad vs. \quad H_1 : \theta < \theta_0$$

Families with Monotone Likelihood Ratio

Let $f(x, \theta)$ be a pmf (pdf) of a random variable X.

Define

$$r(x) = \frac{f(x, \theta_1)}{f(x, \theta_2)}, \quad \theta_1 > \theta_2.$$

If $r(x)$ is an increasing function of $T(x)$, we say that the family of densities $\{f(x, \theta) : \theta \in \Omega\}$ has monotone likelihood ratio (MLR) in $(\theta, T(x))$.

Example 1

Let $X \sim N(\theta, 1)$ with pdf

$$f(x, \theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\theta)^2}.$$

Check whether the family has MLR in (θ, x).

Solution: Consider

$$\begin{aligned} r(x) = \frac{f(x, \theta_1)}{f(x, \theta_2)} &= \exp\left[-\frac{1}{2}(x-\theta_1)^2 + \frac{1}{2}(x-\theta_2)^2 \right] \\ &= \exp\left[\frac{1}{2}(\theta_2^2 - \theta_1^2) + (\theta_1 - \theta_2)x \right] \end{aligned}$$

Since $r(x)$ is an increasing function of x (if $\theta_1 > \theta_2$), we say that $\{N(\theta, 1) : \theta \in \mathbb{R}\}$ has monotone likelihood ratio (MLR) in (θ, x).

Example 2

Let $X_1, \ldots, X_n \sim N(0, \sigma^2)$ with joint density

$$f(\boldsymbol{x}, \sigma^2) = \frac{1}{(\sigma\sqrt{2\pi})^n} e^{-\frac{\sum x_i^2}{2\sigma^2}}, x_i \in \mathbb{R}, \sigma^2 > 0.$$

Check whether the family has MLR.

Solution: Consider

$$r(\boldsymbol{x}) = \frac{f(\boldsymbol{x}, \sigma_1^2)}{f(\boldsymbol{x}, \sigma_2^2)} = \left(\frac{\sigma_2}{\sigma_1}\right)^n \exp\left[\frac{\sum x_i^2}{2}\left(\frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2}\right)\right]$$

This is an increasing function in $T(\boldsymbol{x}) = \sum x_i^2$.

Hence, the family has monotone likelihood ratio (MLR) in $(\sigma^2, \sum x_i^2)$.

Example 3: One-parameter Exponential Family

Consider one parameter exponential family

$$f(x,\theta) = c(\theta)e^{Q(\theta)T(x)}h(x)$$

where $Q(\theta)$ is strictly monotonic. Check whether the family has MLR.

Solution: For $\theta_1 > \theta_2$, consider

$$r(x) = \frac{f(x,\theta_1)}{f(x,\theta_2)} = \frac{c(\theta_1)}{c(\theta_2)}e^{(Q(\theta_1)-Q(\theta_2))T(x)}$$

- If $Q(\theta)$ is monotonically increasing then $r(x)$ is increasing in $T(x)$.
  So, $\{f(x,\theta) : \theta \in \Omega\}$ has MLR in $(\theta, T(x))$.
- If $Q(\theta)$ is monotonically decreasing then $r(x)$ is decreasing in $T(x)$.
  So, $\{f(x,\theta) : \theta \in \Omega\}$ has MLR in $(\theta, -T(x))$.

Example 4: Uniform Distribution

Let $X_1, \ldots, X_n \sim U[0, \theta]$, $\theta > 0$, with joint density

$$f(\boldsymbol{x}, \theta) = \frac{1}{\theta^n}, \ 0 \le x_{(n)} \le \theta.$$

Check whether the family has MLR.

Solution: For $\theta_2 > \theta_1$, consider

$$r(\boldsymbol{x}) = \frac{f(\boldsymbol{x}, \theta_2)}{f(\boldsymbol{x}, \theta_1)} = \frac{\theta_1^n}{\theta_2^n} \frac{I_{[x_{(n)} \le \ \theta_2]}}{I_{[x_{(n)} \le \ \theta_1]}} = \begin{cases} 1, & x_{(n)} \in [0, \theta_1], \\ \infty, & x_{(n)} \in [\theta_1, \theta_2]. \end{cases}$$

This is a nondecreasing function of $X_{(n)}$.

Hence, the family has monotone likelihood ratio (MLR) in $(\theta, X_{(n)})$. Recall that it does not belong to exponential family of distributions.

Theorem 1 on UMP Tests (One-Tailed Hypothesis)

Let the random variable X has pmf (pdf) $f(x,\theta)$ with MLR in $(\theta, T(x))$, $\theta \in \Theta \subseteq \mathbb{R}$.

(i) For testing $H_0 : \theta \leq \theta_0$ against $H_1 : \theta > \theta_0$, there exists a Uniformly Most Powerful (UMP) test, given by

$$\phi(x) = \begin{cases} 1, & \text{if } T(x) > c \\ \gamma, & \text{if } T(x) = c \\ 0, & \text{if } T(x) < c \end{cases} \tag{3}$$

where c and γ are determined by

$$E_{\theta_0}\phi(X) = \alpha \tag{4}$$

Theorem 1 on UMP Tests (One-Tailed Hypothesis)

(ii) The power function $\beta^*(\theta) = E_\theta \phi(X)$ is strictly increasing for all points θ; for which $0 < \beta^*(\theta) < 1$.

Remark: If we consider the dual problem $H_0 : \theta \geq \theta_0$ against $H_1 : \theta < \theta_0$, the inequalities in (3) gets reversed.

UMP test for exponential family

Let X have a prob. density in 1-parameter exponential family

$$f(x, \theta) = c(\theta) e^{Q(\theta) T(x)} h(x),$$

where Q is a monotonic function, then there exists a UMP test for

$$H_0 : \theta \leq \theta_0 \quad vs. \quad H_1 : \theta > \theta_0.$$

If Q is increasing, the test is of the form

$$\phi(x) = \begin{cases} 1, & \text{if } T(x) > c \\ \gamma, & \text{if } T(x) = c \\ 0, & \text{if } T(x) < c \end{cases} \tag{5}$$

If Q is decreasing, the inequalities will get reversed.
Here, c and γ are determined by $E_{\theta_0} \phi(x) = \alpha$.

Example 1

Let $X_1, \dots, X_n$ be a random sample from double exponential distribution

$$f(x, \theta) = \frac{1}{2\theta} \exp\left[-\frac{|x|}{\theta}\right], \ x \in \mathbb{R}, \ \theta > 0.$$

Find the UMP test for testing

$$H_0 : \theta \leq \theta_0 \quad vs. \quad H_1 : \theta > \theta_0.$$

Solution: It belongs to one-parameter exponential family with $Q(\theta) = -\frac{1}{\theta}$ is increasing in θ.
The joint pdf of $X_1, \dots, X_n$

$$f(\boldsymbol{x}, \theta) = \frac{1}{(2\theta)^n} \exp\left[-\frac{\sum |x_i|}{\theta}\right]$$

So, MLR in $(\theta, \sum |X_i|)$.

Example 1 contd.

UMP test is given by

$$\text{Reject } H_0 \text{ if } \sum |X_i| \geq c$$

where c is to be determined from the size condition

$$E_{\theta_0}\phi(\boldsymbol{X}) = \alpha$$

Note that $Y_i = |X_i| \sim \frac{1}{\theta} exp[-\frac{y_i}{\theta}]$, $y_i > 0$, $\theta > 0$.
Further, $\frac{\sum Y_i}{\theta} = \frac{\sum |X_i|}{\theta} \sim Gamma(n, 1)$ and
$\frac{2\sum |X_i|}{\theta_0} \sim \chi^2_{2n}$ under $\theta = \theta_0$.

Example 1 contd.

$$P_{\theta=\theta_0}\left(\frac{2\sum|X_i|}{\theta_0} \geq \frac{2c}{\theta_0}\right) = \alpha \Longrightarrow \frac{2c}{\theta_0} = \chi^2_{2n,\alpha}.$$

So the UMP test is

$$\text{Reject } H_0 \text{ if } \frac{2\sum|X_i|}{\theta_0} \geq \chi^2_{2n,\alpha}$$

$$\text{Accept } H_0 \text{ if } \frac{2\sum|X_i|}{\theta_0} < \chi^2_{2n,\alpha}$$

Thanks for your patience!