MA204: Mathematical Statistics

Tutorial 9

T9.1 Uniformly Most Powerful Test (UMPT)

A test φ with critical region $\mathbb{C}$ is a UMPT of size α for testing H_0: $\theta \in \Theta_0$ against H_1: $\theta \in \Theta_1 = \Theta - \Theta_0$, if

(i) $\sup\limits_{\theta \in \Theta_0} p_\varphi(\theta) = \sup\limits_{\theta \in \Theta_0} \Pr(\boldsymbol{X} \in \mathbb{C} \mid \theta) = \alpha$, where $\boldsymbol{X} = (X_1, \dots, X_n)^\top$.

(ii) For any other test ψ with critical region $\mathbb{A}$ satisfying

$$\sup_{\theta \in \Theta_0} p_\psi(\theta) = \sup_{\theta \in \Theta_0} \Pr(\boldsymbol{X} \in \mathbb{A} \mid \theta) \leqslant \alpha,$$

we have

$$p_\varphi(\theta) = \Pr(\boldsymbol{X} \in \mathbb{C} \mid \theta) \geqslant p_\psi(\theta) = \Pr(\boldsymbol{X} \in \mathbb{A} \mid \theta), \quad \forall\, \theta \in \Theta_1.$$

T9.2 Likelihood Ratio Test (LRT)

— Suppose that we wish to test

$$H_0\colon\ \theta \in \Theta_0 \quad \text{against} \quad H_1\colon\ \theta \in \Theta_1,$$

where Θ_0 and Θ_1 are disjoint, i.e., $\Theta_0 \cap \Theta_1 = \emptyset$. Let $\Theta \triangleq \Theta_0 \cup \Theta_1$, then $\Theta \subseteq \Theta^*$, where Θ^* denotes the parameter space.

— Let $L(\theta)$ denote the likelihood function. The *likelihood ratio statistic* is defined as

$$\lambda(\boldsymbol{X}) = \frac{\sup_{\theta \in \Theta_0} L(\theta)}{\sup_{\theta \in \Theta} L(\theta)} = \frac{L(\hat{\theta}^{\mathrm{R}})}{L(\hat{\theta})},$$

where $\hat{\theta}^{\mathrm{R}}$ denotes the restricted MLE of θ in Θ_0 and $\hat{\theta}$ denotes the (restricted) MLE of θ in Θ. Note that $0 < \lambda(\boldsymbol{x}) \leqslant 1$, where $\boldsymbol{x} = (x_1, \dots, x_n)^\top$.

— The LRT of size α is a test with critical region

$$\mathbb{C} = \{\boldsymbol{x}\colon\ \lambda(\boldsymbol{x}) \leqslant \lambda_\alpha\}, \quad 0 < \lambda_\alpha < 1,$$

and λ_α is determined by

$$\sup_{\theta \in \Theta_0} \Pr\{\lambda(\boldsymbol{X}) \leqslant \lambda_\alpha \mid \theta\} = \alpha.$$

Example T9.1 (Exponential distribution). Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \text{Exponential}(\lambda)$ with density $\lambda \exp(-\lambda x)$ for $x \geqslant 0$ and $\lambda > 0$. Find a UMPT of size α for testing $H_0\colon\ \lambda \leqslant \lambda_0$ versus $H_1\colon\ \lambda > \lambda_0$.

Solution: First, we consider a test of size α for testing $H_0'\colon\ \lambda = \lambda_0$ versus $H_1'\colon\ \lambda = \lambda_1 > \lambda_0$. The likelihood function is

$$L(\lambda) = \prod_{i=1}^n \lambda \exp(-\lambda x_i) = \lambda^n \exp\left(-\lambda \sum_{i=1}^n x_i\right).$$

Then

$$\frac{L(\lambda_0)}{L(\lambda_1)} = \frac{\lambda_0^n \exp(-\lambda_0 \sum_{i=1}^n x_i)}{\lambda_1^n \exp(-\lambda_1 \sum_{i=1}^n x_i)} = \left(\frac{\lambda_0}{\lambda_1}\right)^n \exp\left\{(\lambda_1 - \lambda_0)\sum_{i=1}^n x_i\right\} \leqslant k$$

is equivalent to

$$\bar{x} \leqslant \frac{\log(k)}{n(\lambda_1 - \lambda_0)} + \frac{\log(\lambda_1/\lambda_0)}{\lambda_1 - \lambda_0} \quad \triangleq \quad c,$$

when $\lambda_1 > \lambda_0$. To determine c, we noted that

$$\begin{aligned}
X_i &\overset{\text{iid}}{\sim} \text{Exponential}(\lambda) \\
\Rightarrow n\bar{X} = \sum_{i=1}^n X_i &\sim \text{Gamma}(n, \lambda) \\
\Rightarrow \qquad 2\lambda n\bar{X} &\sim \text{Gamma}\left(n, \frac{1}{2}\right) = \chi^2(2n),
\end{aligned}$$

and

$$\begin{aligned}
\alpha &= \Pr(\bar{X} \le c \mid \lambda = \lambda_0) = \Pr(2\lambda n\bar{X} \le 2\lambda nc \mid \lambda = \lambda_0) \\
&= \Pr(\chi^2(2n) \le 2\lambda_0 nc) = 1 - \Pr(\chi^2(2n) \ge 2\lambda_0 nc) \\
\Rightarrow\ 1-\alpha &= \Pr(\chi^2(2n) \ge 2\lambda_0 nc) \\
\Rightarrow\ 2\lambda_0 nc &= \chi^2(1-\alpha, 2n) \\
\Rightarrow\ c &= \frac{\chi^2(1-\alpha, 2n)}{2\lambda_0 n}.
\end{aligned}$$

By the Neyman–Pearson Lemma, a test φ with critical region

$$\mathbb{C} = \left\{ \boldsymbol{x}\colon\ \bar{x} \leqslant \frac{\chi^2(1-\alpha, 2n)}{2\lambda_0 n} \right\}$$

is the most powerful test of size α for testing H_0': $\lambda = \lambda_0$ versus H_1': $\lambda = \lambda_1 > \lambda_0$. Since the critical region $\mathbb{C}$ depends only on n, λ_0, α and the fact $\lambda_1 > \lambda_0$, but not on the value of λ_1, the test φ is also a UMPT of size α for testing H_0': $\lambda = \lambda_0$ versus H_1: $\lambda > \lambda_0$.

Then, consider φ as a test for testing H_0: $\lambda \leqslant \lambda_0$ versus H_1: $\lambda > \lambda_0$. The size of φ becomes

$$\begin{aligned}
\sup_{\lambda \leqslant \lambda_0} p_\varphi(\lambda) &= \sup_{\lambda \leqslant \lambda_0} \Pr(\bar{X} \leqslant c \mid \lambda) = \sup_{\lambda \leqslant \lambda_0} \Pr(2\lambda n\bar{X} \leqslant 2\lambda nc) \\
&= \sup_{\lambda \leqslant \lambda_0} \Pr(\chi^2(2n) \leqslant 2\lambda nc) = \Pr(\chi^2(2n) \leqslant 2\lambda_0 nc) \\
&= \Pr(\chi^2(2n) \leqslant \chi^2(1-\alpha, 2n)) = 1 - \Pr(\chi^2(2n) > \chi^2(1-\alpha, 2n)) \\
&= 1 - (1-\alpha) = \alpha = p_\varphi(\lambda_0),
\end{aligned}$$

where the fact that $\Pr(\chi^2(2n) \leqslant 2\lambda nc)$ is an increasing function of λ is utilized. Then, the test φ is also a UMPT of size α for testing H_0: $\lambda \leqslant \lambda_0$ versus H_1: $\lambda > \lambda_0$. ‖

Example T9.2 (A normal distribution). Consider two independent samples $X_1, \ldots, X_n \stackrel{\text{iid}}{\sim} N(\theta_1, \theta_3)$ and $Y_1, \ldots, Y_m \stackrel{\text{iid}}{\sim} N(\theta_2, \theta_3)$, where θ_1, θ_2 and θ_3 are unknown parameters. Find the LRT of size α for testing H_0: $\theta_1 = \theta_2$ against H_1: $\theta_1 \ne \theta_2$.

Solution: Let $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)^\top$, $\boldsymbol{\Theta}_0 = \{\boldsymbol{\theta}\colon\ \theta_1 = \theta_2,\ -\infty < \theta_1, \theta_2 < \infty,\ \theta_3 > 0\}$ and $\boldsymbol{\Theta} = \{\boldsymbol{\theta}\colon\ -\infty < \theta_1, \theta_2 < \infty,\ \theta_3 > 0\} = \boldsymbol{\Theta}^*$. let $\hat{\boldsymbol{\theta}}$ be the MLEs of $\boldsymbol{\theta}$ in $\boldsymbol{\Theta}$ and $\hat{\boldsymbol{\theta}}^{\mathrm{R}} = (\hat{\theta}_1^{\mathrm{R}}, \hat{\theta}_1^{\mathrm{R}}, \hat{\theta}_3^{\mathrm{R}})^\top$ be the restricted MLEs of $\boldsymbol{\theta}$ in $\boldsymbol{\Theta}_0$.

Note that $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ are $n + m\,(> 2)$ mutually independent random variables. Under $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, the likelihood function is

$$L(\boldsymbol{\theta}) = \left(\frac{1}{2\pi\theta_3}\right)^{\frac{n+m}{2}} \exp\left[-\frac{\sum_{i=1}^n (x_i - \theta_1)^2 + \sum_{j=1}^m (y_j - \theta_2)^2}{2\theta_3}\right].$$

By partially differentiating the log-likelihood function with respect to θ_1, θ_2 and θ_3 and letting them equal zeros, we have

$$\begin{aligned}
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_1} &= \frac{1}{\theta_3}\sum_{i=1}^n (x_i - \theta_1) = 0,\\
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_2} &= \frac{1}{\theta_3}\sum_{j=1}^m (y_j - \theta_2) = 0,\\
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_3} &= \frac{1}{2\theta_3}\left\{-(n+m) + \frac{1}{\theta_3}\left[\sum_{i=1}^n (x_i - \theta_1)^2 + \sum_{j=1}^m (y_j - \theta_2)^2\right]\right\} = 0.
\end{aligned}$$

The solutions for θ_1, θ_2 and θ_3 are

$$\begin{aligned}
\theta_1 &= \frac{\sum_{i=1}^n x_i}{n} = \bar{x}, \quad \theta_2 = \frac{\sum_{j=1}^m y_j}{m} = \bar{y} \quad \text{and}\\
\theta_3 &= \frac{\sum_{i=1}^n (x_i - \theta_1)^2 + \sum_{j=1}^m (y_j - \theta_2)^2}{n+m},
\end{aligned}$$

respectively. Therefore, the MLEs of $\boldsymbol{\theta}$ in $\boldsymbol{\Theta}$ are

$$\begin{aligned}
\hat{\theta}_1 &= \frac{\sum_{i=1}^n X_i}{n} = \bar{X}, \quad \hat{\theta}_2 = \frac{\sum_{j=1}^m Y_j}{m} = \bar{Y} \quad \text{and}\\
\hat{\theta}_3 &= \frac{\sum_{i=1}^n (X_i - \hat{\theta}_1)^2 + \sum_{j=1}^m (Y_j - \hat{\theta}_2)^2}{n+m},
\end{aligned}$$

so that

$$L(\hat{\boldsymbol{\theta}}) = \left(2\pi \mathrm{e}\hat{\theta}_3\right)^{-\frac{n+m}{2}}.$$

Under H_0, the likelihood function for θ_1 and θ_3 is

$$L(\theta_1, \theta_3) = \left(\frac{1}{2\pi\theta_3}\right)^{\frac{n+m}{2}} \exp\left[-\frac{\sum_{i=1}^n (x_i - \theta_1)^2 + \sum_{j=1}^m (y_j - \theta_1)^2}{2\theta_3}\right].$$

By partially differentiating the log-likelihood function with respect to θ_1 and θ_3 and letting them equal zeros, we have

$$\begin{aligned}
\frac{\partial \ell(\theta_1, \theta_3)}{\partial \theta_1} &= \frac{1}{\theta_3}\left[\sum_{i=1}^n (x_i - \theta_1) + \sum_{j=1}^m (y_j - \theta_1)\right] = 0, \\
\frac{\partial \ell(\theta_1, \theta_3)}{\partial \theta_3} &= \frac{1}{2\theta_3}\left\{-(n+m) + \frac{1}{\theta_3}\left[\sum_{i=1}^n (x_i - \theta_1)^2 + \sum_{j=1}^m (y_j - \theta_1)^2\right]\right\} = 0.
\end{aligned}$$

The solutions for θ_1 and θ_3 are

$$\begin{aligned}
\theta_1 &= \frac{\sum_{i=1}^n x_i + \sum_{j=1}^m y_j}{n+m} = \frac{n\bar{x} + m\bar{y}}{n+m} \quad \text{and} \\
\theta_3 &= \frac{\sum_{i=1}^n (x_i - \theta_1)^2 + \sum_{j=1}^m (y_j - \theta_1)^2}{n+m},
\end{aligned}$$

respectively. Therefore, the restricted MLEs of $\boldsymbol{\theta}$ in $\boldsymbol{\Theta}_0$ are

$$\begin{aligned}
\hat{\theta}_1^{\mathrm{R}} &= \frac{\sum_{i=1}^n X_i + \sum_{j=1}^m Y_j}{n+m} = \frac{n\bar{X} + m\bar{Y}}{n+m} \quad \text{and} \\
\hat{\theta}_3^{\mathrm{R}} &= \frac{\sum_{i=1}^n (X_i - \hat{\theta}_1^{\mathrm{R}})^2 + \sum_{j=1}^m (Y_j - \hat{\theta}_1^{\mathrm{R}})^2}{n+m},
\end{aligned}$$

so that

$$L(\hat{\boldsymbol{\theta}}^{\mathrm{R}}) = \left(2\pi \mathrm{e} \hat{\theta}_3^{\mathrm{R}}\right)^{-\frac{n+m}{2}}.$$

The likelihood ratio statistic is defined as

$$\lambda(\boldsymbol{X}, \boldsymbol{Y}) = \frac{L(\hat{\boldsymbol{\theta}}^{\mathrm{R}})}{L(\hat{\boldsymbol{\theta}})} = \left(\frac{\hat{\theta}_3}{\hat{\theta}_3^{\mathrm{R}}}\right)^{\frac{n+m}{2}} \triangleq \lambda,$$

where $\boldsymbol{X} = (X_1, \ldots, X_n)^\top$ and $\boldsymbol{Y} = (Y_1, \ldots, Y_m)^\top$. Note that

$$\begin{aligned}
\sum_{i=1}^{n} \left(X_i - \frac{n\bar{X} + m\bar{Y}}{n+m} \right)^2 &= \sum_{i=1}^{n} \left[(X_i - \bar{X}) + \left(\bar{X} - \frac{n\bar{X} + m\bar{Y}}{n+m} \right) \right]^2 \\
&= \sum_{i=1}^{n} (X_i - \bar{X})^2 + n \left(\bar{X} - \frac{n\bar{X} + m\bar{Y}}{n+m} \right)^2 \\
&= \sum_{i=1}^{n} (X_i - \bar{X})^2 + \frac{nm^2}{(n+m)^2} (\bar{X} - \bar{Y})^2
\end{aligned}$$

and

$$\begin{aligned}
\sum_{j=1}^{m} \left(Y_j - \frac{n\bar{X} + m\bar{Y}}{n+m} \right)^2 &= \sum_{j=1}^{m} \left[(Y_j - \bar{Y}) + \left(\bar{Y} - \frac{n\bar{X} + m\bar{Y}}{n+m} \right) \right]^2 \\
&= \sum_{j=1}^{m} (Y_j - \bar{Y})^2 + m \left(\bar{Y} - \frac{n\bar{X} + m\bar{Y}}{n+m} \right)^2 \\
&= \sum_{j=1}^{m} (Y_j - \bar{Y})^2 + \frac{n^2 m}{(n+m)^2} (\bar{X} - \bar{Y})^2.
\end{aligned}$$

Then

$$\begin{aligned}
\lambda &= \left[\frac{\sum_{i=1}^{n} (X_i - \bar{X})^2 + \sum_{j=1}^{m} (Y_j - \bar{Y})^2}{\sum_{i=1}^{n} \left(X_i - \frac{n\bar{X}+m\bar{Y}}{n+m} \right)^2 + \sum_{j=1}^{m} \left(Y_j - \frac{n\bar{X}+m\bar{Y}}{n+m} \right)^2} \right]^{\frac{n+m}{2}} \\
&= \left[\frac{\sum_{i=1}^{n} (X_i - \bar{X})^2 + \sum_{j=1}^{m} (Y_j - \bar{Y})^2}{\sum_{i=1}^{n} (X_i - \bar{X})^2 + \sum_{j=1}^{m} (Y_j - \bar{Y})^2 + \frac{nm}{n+m} (\bar{X} - \bar{Y})^2} \right]^{\frac{n+m}{2}} \\
&= \left[\frac{1}{1 + \frac{\frac{nm}{n+m} (\bar{X} - \bar{Y})^2}{\sum_{i=1}^{n} (X_i - \bar{X})^2 + \sum_{j=1}^{m} (Y_j - \bar{Y})^2}} \right]^{\frac{n+m}{2}} \\
&= \left(\frac{n+m-2}{n+m-2+T^2} \right)^{\frac{n+m}{2}} \leqslant \lambda_\alpha
\end{aligned}$$

is equivalent to

$$|T| \geqslant \left[(n+m-2)(\lambda_\alpha^{-\frac{2}{n+m}} - 1)\right]^{\frac{1}{2}} \;\triangleq\; c,$$

where

$$T = \frac{\sqrt{\frac{nm}{n+m}}(\bar{X} - \bar{Y})}{\sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2 + \sum_{j=1}^{m}(Y_j - \bar{Y})^2}{n+m-2}}} \sim t(n+m-2) \quad \text{under } H_0.$$

Since the size of the test equals α,

$$\alpha = \Pr(|T| \geqslant c \mid H_0) \quad \Rightarrow \quad c = t\left(\frac{\alpha}{2}, n+m-2\right).$$

Therefore, the LRT of size α for testing H_0: $\theta_1 = \theta_2$ against H_1: $\theta_1 \neq \theta_2$ is a test with critical region

$$\mathbb{C} = \left\{ \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} : \; |t_{\rm obs}| \geqslant t\left(\frac{\alpha}{2}, n+m-2\right) \right\},$$

where $\boldsymbol{x} = (x_1, \ldots, x_n)^\top$, $\boldsymbol{y} = (y_1, \ldots, y_m)^\top$, $t_{\rm obs}$ is the observed value of T. $\|$