Chapter 03 Likelihood-Based Tests and Confidence Regions
Part 02
3.1 Simple Null Hypothesis

$H_0: \theta = \theta_0$ $\quad H_a: \theta \neq \theta_0$ $\quad$ where $\theta_{b\times 1}$ and
$Y_1, \dots, Y_n$ are iid from $f(y;\theta)$

The identity matrix is denoted by $\mathbb{I}$ in the following.

The asymptotic distribution results for the three statistics stated below hold under suitable regularity conditions.

- The Wald statistic is

$$T_W = \underset{1\times b}{(\hat{\theta}_{MLE} - \theta_0)^T} \underset{b\times b}{\{I_T(\hat{\theta}_{MLE})\}} \underset{b\times 1}{(\hat{\theta}_{MLE} - \theta_0)}$$

$\hat{\theta}_{MLE}$ is $AN(\theta, \{I_T(\theta)\}^{-1})$ or equivalently,
$\{I_T(\theta)\}^{-\frac{1}{2}}(\hat{\theta}_{MLE} - \theta) \xrightarrow{d} N(0, \mathbb{I}_b)$ where $\mathbb{I}_b$ is the b-dimensional identity matrix. It follows that $T_W \xrightarrow{d} \chi^2_b$ as $n \to \infty$ under H_0 provided $I_T(\hat{\theta}_{MLE})\{I_T(\theta_0)\}^{-1}$ converges in probability to $\mathbb{I}_b$ as $n\to\infty$ under H_0.

In fact, $I_T(\hat{\theta}_{MLE})$ can be replaced by $I_T(\theta_0)$ or by the sample information matrix $I_T(Y, \hat{\theta}_{MLE})$ in the definition of T_W and the asymptotic distribution of the resulting statistic is unchanged.

· The score test statistic is

$$T_S = S(\theta_0)^T \{I_T(\theta_0)\}^{-1} S(\theta_0)$$

Under H_0, $S(\theta_0)$ has mean 0, variance $I_T(\theta_0)$, and by The Central Limit Theorem is $AN(0, I_T(\theta_0))$. Thus $\{I_T(\theta_0)\}^{-\frac{1}{2}} S(\theta_0)$ converges in distribution to $N(0, \mathbb{I}_b)$ under H_0. It follows that $T_S \xrightarrow{d} \chi_b^2$ as $n \to \infty$ under H_0.

· The likelihood ratio statistic is

$$T_{LR} = -2\log\left\{\frac{\sup_{\theta \in H_0} L(\theta|Y)}{\sup_{\theta \in \Theta} L(\theta|Y)}\right\} = -2\{l(\theta_0) - l(\hat{\theta}_{MLE})\}$$

where Θ is the parameter space and sup = least upper bound plays the role of "maximum" in cases where the maximum is not attained.

Example 3.1 (Normal model with known variance).

Suppose that $Y_1, \cdots, Y_n$ are iid $N(\mu, 1)$ and $H_0: \mu = \mu_0$. Then

$$f(y) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{1}{2\sigma^2}(y-\mu)^2\right)$$

$$L(\mu) = (2\pi)^{-\frac{n}{2}} \sigma^{-n} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i-\mu)^2\right)$$

$$= (2\pi)^{-\frac{n}{2}} \exp\left(-\sum_{i=1}^{n}(Y_i-\mu)^2\right)$$

$$l(\mu) = \log L(\mu|Y) = -\frac{n}{2}\log(2\pi) - \sum_{i=1}^{n}(Y_i-\mu)^2$$

$$S(\mu) = -\frac{1}{2}\sum_{i=1}^{n}(Y_i-\mu)(-2) = \sum_{i=1}^{n}(Y_i-\mu) \Rightarrow \hat{\mu}_{MLE} = \bar{Y}$$

and $I_T(Y,\mu) = -\frac{\partial}{\partial\mu} S(\mu) = n$,

so that $\hat{\mu}_{MLE} = \bar{Y}$ and $I_T(\mu) = E\{I_T(Y,\mu)\} = n$.

It follows that

$T_W = (\bar{Y}-\mu_0)(n)(\bar{Y}-\mu_0) = n(\bar{Y}-\mu_0)^2$ and

$$T_S = \sum_{i=1}^{n}(Y_i-\mu_0)(n^{-1})\sum_{i=1}^{n}(Y_i-\mu_0) = \frac{1}{n}\sum_{i=1}^{n}(Y_i-\mu_0)\cdot\frac{1}{n}\sum_{i=1}^{n}(Y_i-\mu_0)\,n$$

$= n(\bar{Y}-\mu_0)^2$, are identical for this model.

Example 3.2 (Bernoulli data).

Suppose that $Y_1, \cdots, Y_n$ are iid from a Bernoulli distribution with parameter p and $H_0 : p = p_0$. With $X = \sum_{i=1}^{n} Y_i$,

$$f(y) = p^{y}(1-p)^{1-y}$$

$$\mathcal{L}(p) = \prod_{i=1}^{n} f(y_i) = p^{\Sigma y_i}(1-p)^{n-\Sigma y_i} = p^{x}(1-p)^{n-x}$$

$$l(p) = x \log p + (n-x)\log(1-p)$$

$$S(p) = \frac{\partial l(p)}{\partial p} = \frac{x}{p} - \frac{n-x}{1-p} = \frac{x(1-p)}{p(1-p)} - \frac{p(n-x)}{p(1-p)} = \frac{x - \cancel{xp} - pn + \cancel{xp}}{p(1-p)}$$

$$= \frac{x - pn}{p(1-p)} \quad \Rightarrow \hat{p}_{MLE} = \hat{p} = \frac{x}{n}$$

and $I_T(Y,p) = -\frac{\partial}{\partial p} S(p) = -\left(-\frac{x}{p^2} - \frac{n-x}{(1-p)^2}\right) = \frac{x}{p^2} + \frac{n-x}{(1-p)^2}$

Thus $p_{MLE} = \hat{p} = \frac{x}{n}$ and $I_T(p) = E\{I_T(Y,p)\} = \frac{np}{p^2} + \frac{n-np}{(1-p)^2}$

$$= \frac{n}{p} + \frac{n\cancel{(1-p)}}{(1-p)^{\cancel{2}}} = \frac{n}{p} + \frac{n}{(1-p)}$$

$$= n\,\frac{1-p+p}{p(1-p)} = \frac{n}{p(1-p)}$$

The test statistics are

$$T_W = (\hat{p}-p_0)\left\{\frac{n}{\hat{p}(1-\hat{p})}\right\}(\hat{p}-p_0) = \frac{n(\hat{p}-p_0)}{\hat{p}(1-\hat{p})}$$

$$T_S = \left\{\frac{X-np_0}{p_0(1-p_0)}\right\}\left\{\frac{n}{p_0(1-p_0)}\right\}^{-1}\left\{\frac{X-np_0}{p_0(1-p_0)}\right\}$$

$$= \frac{(X-np_0)}{\cancel{p_0(1-p_0)}}\,\frac{\cancel{p_0(1-p_0)}}{n}\,\frac{(X-np_0)}{p_0(1-p_0)}\cdot\frac{1}{n}\,n$$

$$= \left(\frac{X}{n}-p_0\right)\left(\frac{X}{n}-p_0\right)\frac{n}{p_0(1-p_0)} = \frac{n(\hat{p}-p_0)}{p_0(1-p_0)}$$ $-2\{l(\theta_0)-l(\hat{\theta}_{MLE})\}$

$$T_{LR} = -2\{l(p_0)-l(\hat{p}_{MLE})\}$$

$$= -2\{X\log p_0 + (n-X)\log(1-p_0) - X\log\hat{p}_{MLE} - (n-X)\log(1-\hat{p}_{MLE})\}$$

$$= -2\left\{X\log\frac{p_0}{\hat{p}_{MLE}} + (n-X)\log\frac{1-p_0}{1-\hat{p}_{MLE}}\right\}$$