Chapter 03 Likelihood-Based Tests and Confidence Regions

Part 04

3.2.5 Wald, Score, and Likelihood Ratio Tests – $H_0: h(\theta)=0$

Equivalence of the two types of composite H_0:

1. For $\theta = (\theta_1^T, \theta_2^T)^T$, if $H_0: \theta_1 = \theta_{10}$, then $H_0: h(\theta)=0$ with $h(\theta) = \theta_1 - \theta_{10}$.

2. If $H_0: h(\theta)=0$, then construct a one-to-one transformation $\theta = (\theta_1, \cdots, \theta_p)^T \longrightarrow \beta = (\beta_1, \cdots, \beta_p)^T$ such that $\beta_1 = h(\theta)$. Then $H_0: \beta_1 = 0$ is partitioned null parameter.

Example 3.4 Consider bivariate normal model $N(\mu, \Sigma)$ with $\mu = (\mu_1, \mu_2)^T$ and $\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$.

Let $\theta = (\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)^T$.

a) To test $H_0: \mu_1 = \mu_2$, define transformation

$$\theta \to \beta = (\mu_1 - \mu_2, \mu_2, \sigma_1, \sigma_2, \rho)$$

then $H_0: \beta_1 = 0$.

b) To test $H_0: \mu_1 = \mu_2$ & $\sigma_1 = \sigma_2$, define

$\theta \to \beta = (\mu_1 - \mu_2, \sigma_1 - \sigma_2, \mu_2, \sigma_2, \rho)^T$

then $H_0: (\beta_1, \beta_2)^T = (0, 0)^T$.

We now test $H_0: h(\theta) = 0$ directly without reparameterization. Generally, $h(\theta)$ is an $r \times 1$ vector function with first partial derivatives $H(\theta) = \frac{\partial h(\theta)}{\partial \theta^T}$ an $r \times p$ matrix. Assume $r \leq p$ and $H(\theta)$ is of full rank r.

Example 3.4 (continued)

$p = 5$

a) $h(\theta) = \mu_1 - \mu_2$ $(r=1)$, $H(\theta) = (1, -1, 0, 0, 0)$

b) $r = 2$ $\quad h(\theta) = \begin{pmatrix} h_1(\theta) \\ h_2(\theta) \end{pmatrix} = \begin{pmatrix} \mu_1 - \mu_2 \\ \sigma_1 - \sigma_2 \end{pmatrix}$

$H(\theta) = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \end{pmatrix}$

1. Wald statistic

Since the MLE $\hat{\theta} \sim AN(\theta, I_T^{-1}(\theta))$, Delta Method gives
$h(\hat{\theta}) \sim AN(h(\theta), H(\theta) I_T^{-1}(\theta) H^T(\theta))$.

Thus $T_W = h^T(\hat{\theta}) [H(\theta) I_T^{-1}(\theta) H^T(\theta)]^{-1} h(\hat{\theta})$

Properties:

a. Intuitive and easily computed.

b. Non-invariant to reparameterization and choice of h. E.g. $H_0: \mu_1 = \mu_2$ and $H_0: \mu_1/\mu_2 - 1 = 0$ result in different T_W.

2. Score statistic

$T_S = S^T(\tilde{\theta}) \tilde{I}_T^{-1} S(\tilde{\theta})$

where $\tilde{\theta}$ and $\tilde{I}_T$ are the restricted MLE and the estimated total information under $h_0: h(\theta) = 0$ respectively.

An equivalent form: With Lagrange multipliers, we need to maximize $g(\theta) = l(\theta) - h^T(\theta)\lambda$. Then

$$\frac{\partial g(\theta)}{\partial \theta} = S(\theta) - H^T(\theta)\lambda = 0,$$

$$\frac{\partial g(\theta)}{\partial \lambda} = h(\theta) = 0$$

give solution $\tilde{\theta}$, the restricted MLE, and data-dependent $\tilde{\lambda}$.

Plugging $S(\tilde{\theta}) = H^T(\tilde{\theta})\tilde{\lambda}$ into T_S gives

$T_S = \tilde{\lambda}^T H(\tilde{\theta}) I_T^{-1} H^T(\tilde{\theta}) \tilde{\lambda}$ called Lagrange Multiplier Test in econometrics.

Example 3.4 (continued)

Assume σ_1, σ_2, ρ are known, then $\theta = (\mu_1, \mu_2)^T$. To test $H_0: \mu_1 = \mu_2$, i.e. $h(\theta) = \mu_1 - \mu_2$ with $H(\theta) = (1, -1)$, solve

$$S(\theta) - \begin{pmatrix} 1 \\ -1 \end{pmatrix} \lambda = 0$$

$$\mu_1 - \mu_2 = 0$$

i.e.

$$\frac{\partial l(\theta)}{\partial \mu_1} - \lambda = 0$$

$$\frac{\partial l(\theta)}{\partial \mu_2} + \lambda = 0$$

$$\mu_1 - \mu_2 = 0$$

One can equivalently find $\tilde{\theta}$ directly by substituting $\mu_1 = \mu_2$ into $l(\theta)$.

3. LR statistic

$T_{LR} = -2\,[l(\tilde{\theta}) - l(\hat{\theta})]$