Lecture 28: Wald vs. LRT

Equivalence of the Wald and LRT statistics

$H_0: \theta = \theta_0$ vs. $H_A: \theta \neq \theta_0$ $\quad \theta \in \mathbb{R}^q$

($X_1, \ldots, X_n$ from distribution with parameter θ)

Suppose we want to look performance for a given value of θ

Fixed alternative: true $\theta = \theta_0 + d$ $\quad d \in \mathbb{R}^q$

difference between null & alternative hypotheses

Local alternative: true $\theta = \theta_0 + \frac{d}{\sqrt{n}}$

Intuition: under a fixed alternative $\theta = \theta_0 + d$,
expect Power(θ) $\to 1$ as $n \to \infty$
under a local alternative, maybe Power(θ) $\to ? \in (0,1)$

Key points

- Under H_0, Wald & LRT are asymptotically equivalent as $n \to \infty$

- For a fixed alternative $\theta = \theta_0 + d$, Wald and LRT are not equivalent (if H_A is true)

- For a local alternative $\theta = \theta_0 + \frac{d}{\sqrt{n}}$, Wald and LRT are asymptotically equivalent as $n \to \infty$ (if H_A is true)

For a local alternative $\theta = \theta_0 + \frac{d}{\sqrt{n}}$, or under $H_0: \theta = \theta_0$, Wald & LRT statistics are asymptotically equivalent

Why? Consider $\theta \in \mathbb{R}$

From previous class: if $\hat{\theta} \approx \theta_0$ (either H_0 is true, or $\theta = \theta_0 + \frac{d}{\sqrt{n}}$)

then $$2\ell(\hat{\theta}) - 2\ell(\theta_0) \approx \underbrace{-\frac{1}{n}\ell''(\hat{\theta})}_{I_1(\theta)}(\sqrt{n}(\hat{\theta}-\theta_0))^2$$

$$I_1(\theta)\, n(\hat{\theta}-\theta_0)^2$$

$$\approx I_1(\theta_0)\, n(\hat{\theta}-\theta_0)^2$$

$$= \left(\frac{\sqrt{n}(\hat{\theta}-\theta_0)}{I_1^{-1}(\theta_0)}\right)^2$$

= Wald test statistic!

more generally,

$$2\ell(\hat{\theta}) - 2\ell(\theta_0)$$

$$\approx (\hat{\theta}-\theta_0)^T n I_1(\theta_0)(\hat{\theta}-\theta_0)$$

(Wald test statistic)

Asymptotic normality of MLE:

$$\hat{\theta} \approx N(\theta, \mathcal{I}^{-1}(\theta)) \qquad \mathcal{I}(\theta) = n\mathcal{I}_1(\theta)$$

Test $H_0: \theta = \theta_0$ vs. $H_A: \theta \neq \theta_0$

$$W = (\hat{\theta} - \theta_0)^T n\mathcal{I}_1(\theta_0)(\hat{\theta} - \theta_0)$$

Under H_0, $W \approx \chi^2_q$ $\qquad q =$ dimension of θ, $\theta \in \mathbb{R}^q$

what happens if H_A is true?

$\hat{\theta} \approx N(\theta, \mathcal{I}^{-1}(\theta)) \Rightarrow \hat{\theta} - \theta_0 \approx N(\theta - \theta_0, \mathcal{I}^{-1}(\theta))$

<u>Def</u>: If $Z \sim N(\mu, I)$, $Z^T Z \sim \chi^2_q(\lambda)$, $\lambda = \mu^T \mu$

$\Rightarrow \mathcal{I}^{\frac{1}{2}}(\theta)(\hat{\theta} - \theta_0) \approx N(\mathcal{I}^{\frac{1}{2}}(\theta)(\theta - \theta_0), I)$

$\Rightarrow (\hat{\theta} - \theta_0)^T n\mathcal{I}_1(\hat{\theta})(\hat{\theta} - \theta_0) \approx$

$W \approx \chi^2_q(\lambda)$

Under $\theta = \theta_0 + \delta$: $\lambda = n\delta^T \mathcal{I}_1(\theta)\delta \to \infty$ as $n \to \infty$

Under $\theta = \theta_0 + \frac{\delta}{\sqrt{n}}$: $\lambda = \delta^T \mathcal{I}_1(\theta_0)\delta \in \mathbb{R}$

$$I^{\frac{1}{2}}(\theta)(\hat{\theta}-\theta_0) \approx N\left(I^{\frac{1}{2}}(\theta)(\theta-\theta_0), I\right)$$

$$Z \sim N(\mu, I)$$

$$\implies Z^T Z \sim \chi^2_q(\lambda)$$

$$(\hat{\theta}-\theta_0)^T n I_1(\hat{\theta})(\hat{\theta}-\theta_0)$$

$$\approx \chi^2_q(\lambda)$$

$$\lambda = (\theta-\theta_0)^T n I_1(\theta)(\theta-\theta_0)$$

<u>if</u> $\theta - \theta_0 = d$ (fixed alternative):

$$\lambda = d^T n I_1(\theta)\, d \to \infty$$

<u>if</u> $\theta - \theta_0 = \frac{d}{\sqrt{n}}$

$$\lambda = n\left(\frac{d}{\sqrt{n}}\right)^T \underbrace{I_1\left(\theta_0 + \frac{d}{\sqrt{n}}\right)}_{\approx I_1(\theta_0)} \left(\frac{d}{\sqrt{n}}\right)$$

$$= d^T I_1(\theta_0)\, d \quad \text{as } n \to \infty$$

$$\not\to \infty$$