Lecture 26: Wald vs. LRT

Ciaran Evans

Comparing Wald and LRT statistics

Suppose we observe data $X_1, \ldots, X_n$ from some distribution with parameter $\theta \in \mathbb{R}^q$, and we wish to test

$$H_0 : \theta = \theta_0 \qquad H_A : \theta \neq \theta_0$$

Consider three possible scenarios:

- H_0 is true: $\theta = \theta_0$
- H_A is true, **fixed** alternative: $\theta = \theta_0 + \delta$ $\qquad \delta \neq 0$
- H_A is true, **local** alternative: $\theta = \theta_0 + \frac{\delta}{\sqrt{n}}$ $\qquad \delta \neq 0$

- If H_0 is true, $\text{Power}(\theta_0) \approx \alpha$ regardless of n
- If $\theta = \theta_0 + \delta$, expect $\text{Power}(\theta) \to 1$ as $n \to \infty$ for both tests (even if test stats aren't the same)
- If $\theta = \theta_0 + \frac{\delta}{\sqrt{n}}$, $\text{Power}(\theta) \to ? \in (0,1)$ as $n \to \infty$

Comparing Wald and LRT statistics

Under H_0, or for a local alternative $\theta = \theta_0 + \frac{d}{\sqrt{n}}$, Wald and LRT are asymptotically equivalent as $n \to \infty$ (under certain regularity conditions).

Consider $\theta \in \mathbb{R}$. Previously: if $\hat{\theta} \approx \theta_0$ (either H_0 is true, or $\theta = \theta_0 + \frac{d}{\sqrt{n}}$) then

$$\underbrace{2\ell(\hat{\theta}) - 2\ell(\theta_0)}_{\text{LRT stat}} \approx \underbrace{-\frac{1}{n}\ell''(\hat{\theta})}_{\mathcal{I}_1(\theta)} \left(\sqrt{n}(\hat{\theta} - \theta_0)\right)^2$$

$$\approx \mathcal{I}_1(\theta_0)\, n\, (\hat{\theta} - \theta_0)^2$$

$$= \left(\frac{\sqrt{n}(\hat{\theta} - \theta_0)}{\mathcal{I}_1^{-\frac{1}{2}}(\theta_0)} \right)^2$$

Wald test statistic

Power under a local alternative

$\mathcal{I}(\theta) = n\mathcal{I}_1(\theta)$

Recall asymptotic normality of the MLE: $\widehat{\theta} \approx N(\theta, \mathcal{I}^{-1}(\theta))$

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

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

Under H_0, $W \approx \chi^2_q$. what happens under H_A?

<u>Def</u>: If $Z \sim N(\mu, I)$, $Z \in \mathbb{R}^q$, then

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

(non-central χ^2 distribution with noncentrality parameter λ)

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

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

$$\Rightarrow W = (\hat{\theta} - \theta_0)^T \mathcal{I}(\hat{\theta}) (\hat{\theta} - \theta_0) \approx \chi^2_q(\lambda)$$

$$\lambda = (\theta - \theta_0)^T \mathcal{I}(\theta) (\theta - \theta_0)$$

$$W = (\hat{\theta} - \theta_0)^T \mathcal{I}(\hat{\theta})(\hat{\theta} - \theta_0) \approx \chi^2_q(\lambda)$$

$$\lambda = (\theta - \theta_0)^T \mathcal{I}(\theta)(\theta - \theta_0)$$

If H_0 is true: $(\theta = \theta_0)$ $\Rightarrow$ $\theta - \theta_0 = 0 \Rightarrow \lambda = 0$

$$\chi^2_q(\lambda) = \chi^2_q \quad \checkmark$$

If $\theta = \theta_0 + d$ (fixed alternative)

$$\lambda = d^T \mathcal{I}(\theta) d \qquad \mathcal{I}(\theta) = n\mathcal{I}_1(\theta)$$

$$= \underset{\to\infty}{n}\underbrace{(d^T \mathcal{I}_1(\theta) d)}_{\text{fixed values}}$$

(i.e. $\mathbb{E}[W] \to \infty$ as $n \to \infty$ if fixed alternative)

$$\to \infty \text{ as } n \to \infty$$

If $\theta = \theta_0 + \frac{d}{\sqrt{n}}$ (local alternative):

$$\lambda = \left(\frac{d}{\sqrt{n}}\right)^T n\mathcal{I}_1(\theta)\left(\frac{d}{\sqrt{n}}\right)$$

$$= d^T \mathcal{I}_1(\theta) d = d^T \mathcal{I}_1\left(\theta_0 + \frac{d}{\sqrt{n}}\right) d$$

$$\to d^T \mathcal{I}_1(\theta_0) d$$

(not 0, but not ∞) converges to some stable value

Class activity

- Simulate data under a local alternative
- Verify that the Wald statistic follows a non-central χ^2 distribution