Wald vs. likelihood ratio tests

Class activity

https://sta711-s23.github.io/class_activities/ca_lecture_27.html

Take-aways:

- Under H_0, Wald and LRT are asymptotically equivalent as $n \to \infty$
- For a fixed alternative, Wald & LRT are not asymptotically equivalent
- For a local alternative, Wald & LRT are asymptotically equivalent

Asymptotic distribution of the Wald statistic

Asymptotic normality of the MLE:

$$\sqrt{n}\,(\hat{\theta} - \theta) \xrightarrow{d} N(0, \mathcal{I}_1^{-1}(\theta)) \qquad \theta \in \mathbb{R}^q$$

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

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

identity matrix

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

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

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

For a fixed alternative $\theta = \theta_0 + \delta$, $\lambda = n\delta^T \mathcal{I}_1(\theta)\,\delta \rightarrow \infty$

Local alternative: $\theta = \theta_0 + \frac{\delta}{\sqrt{n}} \Rightarrow \lambda = \delta^T \mathcal{I}_1(\theta_0)\,\delta$

For a local alternative, or when $\theta = \theta_0$, Wald and LRT are asymptotically equivalent

Equivalence of the Wald and LRT statistics

For a local alternative, or when $\theta = \theta_0$, Wald and LRT are asymptotically equivalent

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

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

$$
\begin{aligned}
2\ell(\hat{\theta}) - 2\ell(\theta_0) &\approx -\frac{1}{n}\ell''(\hat{\theta})\left(\sqrt{n}(\hat{\theta}-\theta_0)\right)^2 \\
&\approx n\mathcal{I}_1(\theta_0)(\hat{\theta}-\theta_0)^2 \\
&= \underbrace{\mathcal{I}(\theta_0)(\hat{\theta}-\theta_0)^2}_{\text{Wald statistic when } \theta \in \mathbb{R}\,!}
\end{aligned}
$$

More generally,

$$2\ell(\hat{\theta}) - 2\ell(\theta_0) \approx (\hat{\theta}-\theta_0)^T \mathcal{I}(\theta_0)(\hat{\theta}-\theta_0)$$