Lecture 18: Wald Tests

Hypothesis test for a population proportion

Let $Y_1, Y_2, \ldots \stackrel{\text{iid}}{\sim}$ Bernoulli(p). We want to test

$H_0 : p = p_0 \quad H_A : p \neq p_0$

$$Z_n = \frac{\hat{p} - p_0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}$$

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} Y_i = \bar{Y}_n$$

$$\text{Var}(\hat{p}) = \text{Var}(\bar{Y}_n) = \frac{p(1-p)}{n}$$

CLT: $\dfrac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \xrightarrow{d} N(0,1)$

Under H_0: CLT + Slutsky's $\Rightarrow \dfrac{\hat{p} - p_0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} \xrightarrow{d} N(0,1)$

Alternatively: $\dfrac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$ under H_0, both $\xrightarrow{d} N(0,1)$

Wald test for one parameter (scalar)

Let $\theta \in \mathbb{R}$ be a parameter of interest, and let $\hat{\theta}_n$ be an estimator such that $\dfrac{\hat{\theta}_n - \theta}{S_n} \xrightarrow{d} N(0,1)$

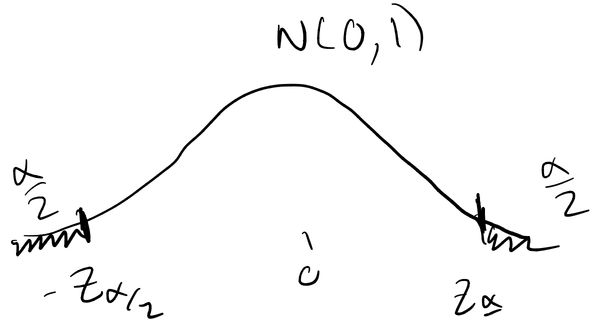
for some sequence S_n $(S_n^2 \approx \mathrm{var}(\hat{\theta}_n))$

To test $H_0: \theta = \theta_0$ vs. $\theta \neq \theta_0$:

- let $Z_n = \dfrac{\hat{\theta}_n - \theta_0}{S_n}$

- the <u>Wald test</u> rejects H_0 when $|Z_n| > z_{\alpha/2}$

  where $z_{\alpha/2}$ is the upper $\frac{\alpha}{2}$ quantile of $N(0,1)$



- p-value for observed test statistic z, is $2\Phi(-|z|)$ ← N(0,1) cdf

$$Z_n = \frac{\hat{\theta}_n - \theta_0}{S_n}$$

MLE: (under some assumptions, e.g. iid data)

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

$$\Rightarrow \mathcal{I}^{-1}(\theta) = \frac{1}{n}\mathcal{I}_1^{-1}(\theta)$$

Examples:

① $H_0: \mu = \mu_0$

$\hat{\theta}_n = \bar{X}_n \qquad \theta_0 = \mu_0$

$S_n = \frac{\sigma}{\sqrt{n}}$ or $S_n = \frac{s}{\sqrt{n}}$

② $H_0: p = p_0$

$\hat{\theta}_n = \hat{p}_n \qquad \theta_0 = p_0$

$S_n = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ or $S_n = \sqrt{\frac{p_0(1-p_0)}{n}}$

③ Logistic regression: $H_0: \beta_j = 0$ $\qquad \hat{\theta}_n = \hat{\beta}_j \qquad \theta_0 = 0$

$S_n = \sqrt{[\hat{\mathcal{I}}^{-1}(\beta)]_{jj}}$

Comment: any asymptotically normal statistic can be used to construct a Wald test

Testing multiple parameters

(later: return to LRT / "drop-in-deviance" test for multiple parameters)

Logistic regression model for the dengue data:

$$Y_i \sim \text{Bernoulli}(p_i)$$

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 WBC_i + \beta_2 PLT_i$$

Researchers want to know if there is any relationship between white blood cell count or platelet count, and the probability a patient has dengue. What hypotheses should they test?

$H_0: \quad \beta_1 = \beta_2 = 0$

$H_A: \quad$ at least one of $\beta_1, \beta_2 \neq 0$

Testing multiple parameters

```
              Estimate    Std. Error    z value       Pr(>|z|)
(Intercept)  2.641506279 0.1213233066  21.77246 4.233346e-105
WBC         -0.289290446 0.0134349261 -21.53272 7.689284e-103
PLT         -0.006561464 0.0005932064 -11.06101  1.938945e-28
```

testing $\beta_1 = 0$ (w/ PLT model)

testing $\beta_2 = 0$ (w/ WBC in model)

Can the researchers test their hypotheses using this output?

No – these are hypothesis tests for single coefficients

Wald tests for multiple parameters

Consider dengue example: (asymptotic normality of MLE)

$$\underbrace{\hat{\beta}}_{\text{MLE for } \beta} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} \approx N(\beta, \mathcal{I}^{-1}(\beta))$$

we want to test $H_0: \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Facts: Suppose $X \sim N(\mu, \Sigma)$

matrix A

$$AX \sim N(A\mu, A\Sigma A^T)$$

(cf: HW)

$$\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = C\beta = \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{C} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$

$$\mathcal{I}^{-1}(\beta) = \begin{bmatrix} Var(\hat{\beta}_0) & \cdots & \\ \cdots & Var(\hat{\beta}_1) & \cdots \\ \cdots & \cdots & Var(\hat{\beta}_2) \end{bmatrix}$$

$C\mathcal{I}^{-1}(\beta)C^T$

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = C\hat{\beta} \approx N(C\beta, C\mathcal{I}^{-1}(\beta)C^T)$$

$$\begin{bmatrix}\hat\beta_1\\ \hat\beta_2\end{bmatrix} = C\hat\beta \approx N(C\beta, C\mathcal{I}^{-1}(\beta)C^T)$$

want to turn $C\hat\beta$ into a real-valued test statistic

Motivation: $Z \sim N(0,1) \in \mathbb{R}$ $\qquad Z^2 \sim \chi^2_1$

$Z \sim N(0, I) \in \mathbb{R}^q$ $\qquad Z^T Z \sim \chi^2_q$

$$C\hat\beta \approx N(C\beta, C\mathcal{I}^{-1}(\beta)C^T)$$

$$(C\mathcal{I}^{-1}(\beta)C^T)^{-\frac{1}{2}}(C\hat\beta - C\beta) \approx N(0, I)$$

$$\Rightarrow \quad (C\hat\beta - C\beta)^T (C\mathcal{I}^{-1}(\beta)C^T)^{-1}(C\hat\beta - C\beta) \approx \chi^2_q$$

$q =$ length of $C\beta$

H_0: $C\beta = \gamma_0$ $\qquad$ (e.g., $\gamma_0 = \begin{bmatrix}0\\0\end{bmatrix}$)

test statistic: $W = (C\hat\beta - \gamma_0)^T (C\mathcal{I}^{-1}(\beta)C^T)^{-1}(C\hat\beta - \gamma_0)$

Under H_0: $W \approx \chi^2_q$

Class activity

https://sta711-s24.github.io/class_activities/ca_lecture_18.html