Wald tests

Recap

want to test $H_0: \beta_1 = \beta_2 = 0$ H_A: at least one of $\beta_1, \beta_2 \neq 0$

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

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

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

identity matrix

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

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

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

q = length of $C\beta$

logistic regression: $V_n = \mathcal{I}^{-1}(\beta)$ (using MLE $\hat{\beta}$)

General Wald test

Let $\Theta \in \mathbb{R}^p$ and let $\hat{\Theta}_n$ be an estimator such that

$$V_n^{-\frac{1}{2}}(\hat{\Theta}_n - \Theta) \xrightarrow{d} N(0, I)$$

$$(\Rightarrow \quad \hat{\Theta}_n \mathrel{\dot\approx} N(\Theta, V_n))$$

Let $C \in \mathbb{R}^{q \times p}$. To test

$$H_0: C\Theta = \gamma_0 \qquad H_A: C\Theta \neq \gamma_0$$

$Z_n = (CV_nC^T)^{-\frac{1}{2}}(C\hat{\Theta}_n - \gamma_0)$ $\quad Z_n \mathrel{\dot\approx} N(0, I)$ under H_0

Let $W_n = (C\hat{\Theta}_n - \gamma_0)^T (CV_nC^T)^{-1}(C\hat{\Theta}_n - \gamma_0) = Z_n^T Z_n$

Under H_0, $W_n \mathrel{\dot\approx} \chi^2_q$

Wald test: specifies $\alpha \in [0,1]$

rejects H_0 when $W_n > \chi^2_{q,\alpha}$

$\chi^2_{q,\alpha}$ = upper α quantile of χ^2_q

$\chi^2_{q,\alpha}$

Class activity, Part I

$$H_0: \; C\beta = \underbrace{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}_{\gamma_0}$$

$$\hat{\beta} \Rightarrow C\hat{\beta}$$

https://sta711-s23.github.io/class_activities/ca_lecture_18.html

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} \qquad \begin{bmatrix} \beta_3 \\ \beta_4 \end{bmatrix} = C\beta$$

$$V_n = \mathcal{I}^{-1}(\beta) = \begin{bmatrix} Var(\hat{\beta}_0) & & & & \\ & Var(\hat{\beta}_1) & & & \\ & & Var(\hat{\beta}_2) & & \\ & & & Var(\hat{\beta}_3) & \\ & & & & Var(\hat{\beta}_4) \end{bmatrix}$$

(call the covariances)

(call the covariances)

$$C V_n C^T = \begin{bmatrix} Var(\hat{\beta}_3) & Cov(\hat{\beta}_3, \hat{\beta}_4) \\ Cov(\hat{\beta}_3, \hat{\beta}_4) & Var(\hat{\beta}_4) \end{bmatrix}$$

Test statistic (W_n):

$$\begin{bmatrix} \hat{\beta}_3 & \hat{\beta}_4 \end{bmatrix} \begin{bmatrix} Var(\hat{\beta}_3) & Cov(\hat{\beta}_3, \hat{\beta}_4) \\ Cov(\hat{\beta}_3, \hat{\beta}_4) & Var(\hat{\beta}_4) \end{bmatrix} \begin{bmatrix} \hat{\beta}_3 \\ \hat{\beta}_4 \end{bmatrix}$$

$\mathcal{I}^{-1}(\beta) = \text{vcov(model)}$

Class activity

$\hat{\beta}$

```
betahat <- m1$coefficients[4:5]
V <- vcov(m1)[4:5, 4:5]
test_stat <- t(betahat) %*% solve(V) %*% betahat
test_stat
```

```
##              [,1]
## [1,] 85.60437
```

$\leftarrow W_n$

Under H_0, $W_n \approx \chi^2_2 \leftarrow$ # parameters tested

Reject when $W_n > \chi^2_{2,\alpha}$

Class activity

```
betahat <- m1$coefficients[4:5]
V <- vcov(m1)[4:5, 4:5]
test_stat <- t(betahat) %*% solve(V) %*% betahat
test_stat
```

```
##          [,1]
## [1,] 85.60437
```

```
# rejection region for alpha = 0.05
qchisq(0.05, df=2, lower.tail=F)
```

$\chi^2_{2,0.05}$

α

upper quantile

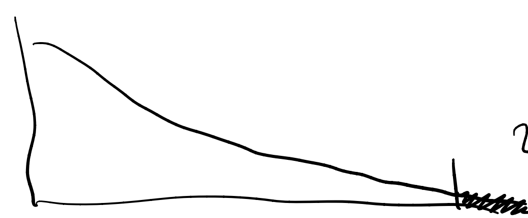
```
## [1] 5.991465
```

$W_n = 85.6 > \chi^2_{2,0.05}$

$\Rightarrow$ reject H_0!

```
# p-value
pchisq(test_stat, df=2, lower.tail=F)
```

```
##              [,1]
## [1,] 2.577787e-19
```



85.6

2.6×10^{-19}

≈ 0

A different question

We have the model

$$Y_i \sim Bernoulli(p_i)$$

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 Sex_i + \beta_2 Age_i + \beta_3 SecondClass_i + \beta_4 ThirdClass_i$$

We want to test whether there is a difference in the chance of survival for second and third class passengers, holding age and sex fixed.

> What hypotheses should we test?

$H_0: \beta_3 = \beta_4$ $\quad$ $H_A: \beta_3 \neq \beta_4$

$\beta_4 - \beta_3 = 0$ $\quad$ $\beta_4 - \beta_3 \neq 0$

Idea: choose C st $C\beta = \beta_4 - \beta_3$

Contrasts

Class activity, Part II

https://sta711-s23.github.io/class_activities/ca_lecture_18.html

Class activity

```r
a <- c(0, 0, 0, -1, 1)
test_stat <- (t(a) %*% coef(m1))/sqrt(t(a) %*% vcov(m1) %*% a)
test_stat
```

```
##           [,1]
## [1,] -5.207289
```

Class activity

```r
a <- c(0, 0, 0, -1, 1)
test_stat <- (t(a) %*% coef(m1))/sqrt(t(a) %*% vcov(m1) %*% a)
test_stat
```

```
##           [,1]
## [1,] -5.207289
```

```r
# rejection region for alpha = 0.05
qnorm(0.025, lower.tail=F)
```

```
## [1] 1.959964
```

```r
# p-value
2*pnorm(abs(test_stat), lower.tail=F)
```

```
##              [,1]
## [1,] 1.916191e-07
```

A two-sample test for a difference in means