Wald tests and likelihood ratio tests

- HW 2 due tomorrow
- HW 3 released later today, due next Friday
- Challenge 4: Deriving VIFs, released later today

Wald tests for multiple parameters

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

$$\beta = \begin{bmatrix} \beta_{(1)} \\ \beta_{(2)} \end{bmatrix} \qquad \beta_{(1)} = \beta_0 \qquad \beta_{(2)} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

$$H_0: \beta_1 = \beta_2 = 0 \qquad \beta_{(2)} = \beta^0_{(2)} \leftarrow \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

H_A: at least one of $\beta_1, \beta_2 \neq 0$ $\qquad \beta_{(2)} \neq \beta^0_{(2)}$

$$\hat{\beta} \approx N\left(\begin{bmatrix} \beta_{(1)} \\ \beta_{(2)} \end{bmatrix}, \mathcal{I}^{-1}(\beta)\right) \qquad \mathcal{I}^{-1}(\beta) = \begin{bmatrix} \mathcal{I}^{11} & \mathcal{I}^{12} \\ \mathcal{I}^{21} & \mathcal{I}^{22} \end{bmatrix}$$

$$\hat{\beta}_{(2)} \approx N(\beta_{(2)}, \mathcal{I}^{22})$$

If H_0 is true, then $\beta_{(2)} = \beta^0_{(2)}$

$$\Rightarrow \hat{\beta}_{(2)} \approx N(\beta^0_{(2)}, \mathcal{I}^{22}) \qquad \Rightarrow (\mathcal{I}^{22})^{-\frac{1}{2}}(\hat{\beta}_{(2)} - \beta^0_{(2)}) \approx N(0, I)$$

$$(\hat{\Sigma}^{22})^{-\frac{1}{2}} (\hat{\beta}_{(2)} - \beta^0_{(2)}) \approx N(0, I)$$

$$W = (\hat{\beta}_{(2)} - \beta^0_{(2)})^T \underbrace{(\hat{\Sigma}^{22})^{-1}}_{\downarrow} (\hat{\beta}_{(2)} - \beta^0_{(2)}) \qquad W \in \mathbb{R}$$

positive semi-definite matrix

def: A is PSD iff $x^T A x \geq 0 \; \forall x$

$\Rightarrow W \geq 0$

$H_0: \beta_{(2)} = \beta^0_{(2)}$

$H_A: \beta_{(2)} \neq \beta^0_{(2)}$

if H_0 is true, do we expect W to be large or small?

we expect $\hat{\beta}_{(2)} \approx \beta_{(2)} = \beta^0_{(2)}$ if H_0 is true $\Rightarrow$ W should be small

Large values of W are evidence against H_0

Fact: if $Z \sim N(0,1)$ then $Z^2 \sim \chi^2_1$. $\sum_{i=1}^{q} Z_i^2 \sim \chi^2_q$

$$W = (\hat{\beta}_{(2)} - \beta^0_{(2)})^T \, (\mathcal{I}^{22})^{-1} \, (\hat{\beta}_{(2)} - \beta^0_{(2)}) \approx \chi^2_2 \qquad \beta_{(2)} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

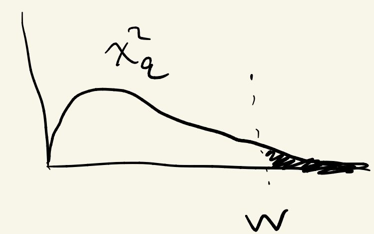
<u>In general</u>: $\beta = \begin{bmatrix} \beta_{(1)} \\ \beta_{(2)} \end{bmatrix}$, $\beta_{(2)} \in \mathbb{R}^q$, $\mathcal{I}^{22} \in \mathbb{R}^{q \times q}$

Then $$W = (\hat{\beta}_{(2)} - \beta^0_{(2)})^T (\mathcal{I}^{22})^{-1} (\hat{\beta}_{(2)} - \beta^0_{(2)})$$

$$\approx \chi^2_q \quad \text{under } H_0$$

↑

\# parameters we test!



$\uparrow W \Rightarrow$ stronger evidence against H_0

<u>p-value</u>: $P(\chi^2_q > W)$

pchisq(w, df = q, lower.tail=F)

Class activity

https://sta712-f22.github.io/class_activities/ca_lecture_11.html

- Wald tests for the dengue data

Likelihood ratio tests

```
m1 <- glm(Dengue ~ WBC + PLT, data = dengue,
          family = binomial)
summary(m1)
```

```
...
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.6415063  0.1213233   21.77   <2e-16 ***
## WBC         -0.2892904  0.0134349  -21.53   <2e-16 ***
## PLT         -0.0065615  0.0005932  -11.06   <2e-16 ***
## ---
##     Null deviance: 6955.8  on 5719  degrees of freedom
## Residual deviance: 5399.7  on 5717  degrees of freedom
## AIC: 5405.7
...
```

What information replaces R^2 and R^2_{adj} in the GLM output?

Deviance

Definition: The *deviance* of a fitted model with parameter estimates $\widehat{\beta}$ is given by

$$2\ell(\text{saturated model}) - 2\ell(\widehat{\beta})$$

Comparing deviances

```
m1 <- glm(Dengue ~ WBC + PLT, data = dengue,
          family = binomial)
summary(m1)
```

```
...
##     Null deviance: 6955.8  on 5719  degrees of freedom
## Residual deviance: 5399.7  on 5717  degrees of freedom
## AIC: 5405.7
...
```

Comparing deviances

Full model: $\log\left(\dfrac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 WBC_i + \beta_2 PLT_i$

Reduced model: $\log\left(\dfrac{p_i}{1-p_i}\right) = \beta_0$

$$G = 2\ell(\widehat{\beta}) - 2\ell(\widehat{\beta}^0)$$

> Why is G always ≥ 0?

Comparing deviances

Full model: $\log\left(\dfrac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 WBC_i + \beta_2 PLT_i$

Reduced model: $\log\left(\dfrac{p_i}{1-p_i}\right) = \beta_0$

$$G = 2\ell(\widehat{\beta}) - 2\ell(\widehat{\beta}^0) = 1556.1$$

> If the reduced model is correct, how unusual is $G = 1556.1$?

Likelihood ratio test