Lecture 16: Wald tests

Ciaran Evans

Wald test for one parameter

Let $\theta \in \mathbb{R}$ be a parameter of interest, and let $\hat{\theta}_n$ be an estimator s.t. $\frac{\hat{\theta}_n - \theta}{s_n} \xrightarrow{d} N(0,1)$

for some sequence s_n $\quad (s_n^2 \approx \mathrm{Var}(\hat{\theta}_n))$

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

- let $Z_n = \dfrac{\hat{\theta}_n - \theta_0}{s_n}$
- reject when $|Z_n| > z_{\frac{\alpha}{2}}$ ← upper $\frac{\alpha}{2}$ quantile of $N(0,1)$
- p-value: $2\,\Phi(-|Z_n|)$

↑

$N(0,1)$ cdf

Wald test for one parameter: examples

- Population mean: $H_0 : \mu = \mu_0$

$$\hat{\theta}_n = \bar{x}_n \ , \quad \theta_0 = \mu_0 \ , \quad s_n = \frac{\sigma}{\sqrt{n}} \quad \text{or} \quad \frac{s}{\sqrt{n}}$$

$$s = \sqrt{\frac{1}{n-1} \sum_i (x_i - \bar{x}_n)^2}$$

- Population proportion: $H_0 : p = p_0$

$$\hat{\theta}_n = \hat{p}_n \ , \quad \theta_0 = p_0 \ , \quad s_n = \sqrt{\frac{p_0(1-p_0)}{n}} \quad \text{or} \quad \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

- Regression coefficient: $H_0 : \beta_j = 0$

$$\hat{\theta}_n = \hat{\beta}_j \ , \quad \theta_0 = 0 \ , \quad s_n = \sqrt{\underbrace{\left[\underset{\sim}{I}^{-1}(\beta)\right]_{jj}}_{j^{th} \text{ diagonal entry}}}$$

Testing multiple parameters

Logistic regression model for the dengue data:

$$Y_i \sim Bernoulli(p_i)$$

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

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

Question: What hypotheses should they test?

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

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

Testing multiple parameters

```
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)     2.642      0.121  21.772        0
## WBC            -0.289      0.013 -21.533        0
## PLT            -0.007      0.001 -11.061        0
```

$H_0 : \beta_1 = \beta_2 = 0$

Can the researchers test their hypotheses using this output?

No: each p-value only tests $H_0: \beta_j = 0$

Wald tests for multiple parameters

$\mathcal{I}(\beta) = X^T W X$

(Fisher info for the full data)

For the dengue example:

$$\widehat{\beta} = \begin{pmatrix} \widehat{\beta}_0 \\ \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{pmatrix} \approx N(\beta, \mathcal{I}^{-1}(\beta))$$

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

$\hat{\beta}$

```
##  (Intercept)          WBC          PLT
##  2.641506279 -0.289290446 -0.006561464
```

```
vcov(m1)
```

$\hat{\mathcal{I}}^{-1}(\beta)$ $\quad X^T \hat{W} X$

```
##               (Intercept)           WBC           PLT
## (Intercept)  1.471934e-02 -4.937020e-04 -5.125888e-05
## WBC         -4.937020e-04  1.804972e-04 -3.221337e-06
## PLT         -5.125888e-05 -3.221337e-06  3.518938e-07
```

Multivariate normal distribution

Definition: Let $X = (X_1, ..., X_k)^T$. We say that $X \sim N(\mu, \Sigma)$ if for any $\mathbf{a} \in \mathbb{R}^k$, $\mathbf{a}^T X$ follows a (univariate) normal distribution.

- $\mu = \mathbb{E}[X] = \begin{bmatrix} \mathbb{E}[X_1] \\ \mathbb{E}[X_2] \\ \vdots \\ \mathbb{E}[X_k] \end{bmatrix} \in \mathbb{R}^k$

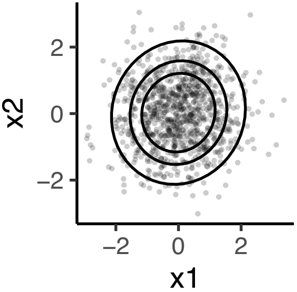
- $\Sigma = \mathrm{Var}(X) = \begin{bmatrix} \mathrm{Var}(X_1) & \mathrm{Cov}(X_1, X_2) & \cdots & \mathrm{Cov}(X_1, X_k) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Var}(X_2) & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathrm{Cov}(X_k, X_1) & \cdots & & \mathrm{Var}(X_k) \end{bmatrix} \in \mathbb{R}^{k \times k}$

  (variance-covariance matrix)

Multivariate normal distribution

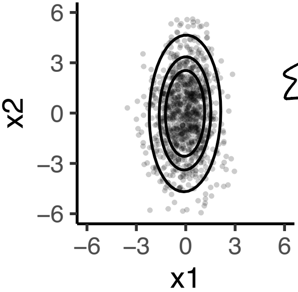
$N(0, I)$

$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$



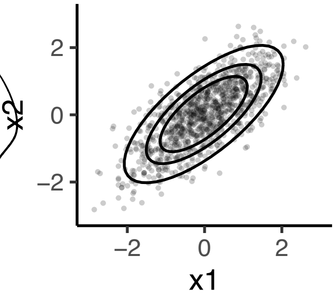
$N(0, \Sigma)$

$\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}$



$N(0, \Sigma)$

$\Sigma = \begin{pmatrix} 1 & 0.75 \\ 0.75 & 1 \end{pmatrix}$



Wald tests for multiple parameters

For the dengue example: $\widehat{\beta} = \begin{pmatrix} \widehat{\beta}_0 \\ \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{pmatrix} \quad \approx$ Normal

We want to test: $\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad H_0: C\beta = 0$

rewrite:

$$\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}$$

$$= C\beta$$

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = C\hat{\beta}$$

vector under H_0

In general: hypotheses of form $H_0: C\beta = \gamma_0$

Wald tests for multiple parameters

want:
$var(C\hat{\beta})$

$$H_0 : \mathbf{C}\beta = \gamma_0$$

For the dengue example:

```
C <- matrix(c(0, 1, 0,
              0, 0, 1), nrow=2, byrow=T)


C
```

```
##      [,1] [,2] [,3]
## [1,]    0    1    0
## [2,]    0    0    1
```

```
C %*% coef(m1)
```

```
##              [,1]
## [1,] -0.289290446
## [2,] -0.006561464
```

$\hat{\beta}_1$

$\hat{\beta}_2$

Wald tests for multiple parameters

For the dengue example:

```
C
```

```
##       [,1] [,2] [,3]
## [1,]    0    1    0
## [2,]    0    0    1
```

```
vcov(m1)
```

```
##               (Intercept)           WBC           PLT
## (Intercept)  1.471934e-02 -4.937020e-04 -5.125888e-05
## WBC         -4.937020e-04  1.804972e-04 -3.221337e-06
## PLT         -5.125888e-05 -3.221337e-06  3.518938e-07
```

```
C %*% vcov(m1) %*% t(C)
```

```
##               [,1]          [,2]
## [1,]  1.804972e-04 -3.221337e-06
## [2,] -3.221337e-06  3.518938e-07
```

Wald tests for multiple parameters

univariate intuition:

$X \sim N(\mu, \sigma^2)$

$aX \sim N(a\mu, a^2\sigma^2)$

- $H_0 : \mathbf{C}\beta = \gamma_0$
- Look at $\mathbf{C}\widehat{\beta}$

Fact: Suppose $X \sim N(\mu, \Sigma)$ (multivariate normal), and $\mathbf{A}$ is a matrix (not random). Then:

$$\mathbf{A}X \sim N(A\mu, A\Sigma A^T)$$

(pf: HW?)

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

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

Under H_0: $C\hat{\beta} \approx N(\gamma_0, C\mathcal{I}^{-1}(\beta)C^T)$

Test statistic and p-value

- $H_0 : \mathbf{C}\beta = \gamma_0$
- $\mathbf{C}\widehat{\beta} \approx N(\mathbf{C}\beta,\ \mathbf{C}\mathcal{I}^{-1}(\beta)\mathbf{C}^T)$
- Want to turn $\mathbf{C}\widehat{\beta}$ into a scalar test statistic

Facts:
- if $Z \sim N(0,1)$ (scalar), then $Z^2 \sim \chi^2_1$
- if $Z_1, \dots, Z_k \overset{iid}{\sim} N(0,1)$, then $\sum_{i=1}^k Z_i^2 \sim \chi^2_k$

$$\Rightarrow \quad \sum_i Z_i^2 = [Z_1 \ \cdots \ Z_k] \begin{bmatrix} Z_1 \\ \vdots \\ Z_k \end{bmatrix}$$

$\Rightarrow$ if $Z \sim N(0, I)$ (multivariate normal),

then $Z^T Z \sim \chi^2_k$

Test statistic and p-value

$$Z \sim N(0, I)$$
$$\Rightarrow Z^T Z \sim \chi^2_k$$

- $H_0 : \mathbf{C}\beta = \gamma_0$
- $\mathbf{C}\widehat{\beta} \approx N(\mathbf{C}\beta,\ \mathbf{C}\mathcal{I}^{-1}(\beta)\mathbf{C}^T)$
- Want to turn $\mathbf{C}\widehat{\beta}$ into a scalar test statistic

under H_0: $\quad C\hat{\beta} \approx N(\gamma_0,\ C\mathcal{I}^{-1}(\beta)C^T)$

$$C\hat{\beta} - \gamma_0 \approx N(0,\ C\mathcal{I}^{-1}(\beta)C^T)$$

$$(C\mathcal{I}^{-1}(\beta)C^T)^{-\frac{1}{2}}(C\hat{\beta} - \gamma_0) \approx N(0, I)$$

$$\Rightarrow (C\hat{\beta} - \gamma_0)^T (C\mathcal{I}^{-1}(\beta)C^T)^{-\frac{1}{2}} (C\mathcal{I}^{-1}(\beta)C^T)^{-\frac{1}{2}} (C\hat{\beta} - \gamma_0) \approx \chi^2_q$$

$$\Rightarrow (C\hat{\beta} - \gamma_0)^T (C\mathcal{I}^{-1}(\beta)C^T)^{-1} (C\hat{\beta} - \gamma_0) \approx \chi^2_q$$

$q = \dim(C\beta)$
$= \#$ parameters tested

Example

$H_0 : \mathbf{C}\beta = \gamma_0$

Dengue example: $H_0 : C\beta = 0$

$$(\mathbf{C}\widehat{\beta} - \gamma_0)^T (\mathbf{C}\mathcal{I}^{-1}(\beta)\mathbf{C}^T)^{-1} (\mathbf{C}\widehat{\beta} - \gamma_0) \approx \chi^2_q \qquad \text{under } H_0$$

```
C %*% coef(m1)
```

$\hat{\beta}_1$ $\hat{\beta}_2$

```
## [1] -0.289290446 -0.006561464
```

```
test_stat <- t(C %*% coef(m1)) %*% solve(C %*% vcov(m1) %*%
  (C %*% coef(m1))
test_stat
```

```
##          [,1]
## [1,] 930.5777
```

```
pchisq(test_stat, 2, lower.tail=F)
```

```
##               [,1]
## [1,] 8.464858e-203
```