Lecture 16: Wald tests

Ciaran Evans

Wald test for one parameter

Wald test for one parameter: examples

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

- Population proportion: $H_0 : p = p_0$

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

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?

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?

Wald tests for multiple parameters

For the dengue example:

$$\widehat{\beta} = \begin{pmatrix} \widehat{\beta}_0 \\ \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{pmatrix}$$

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

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

```
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
```

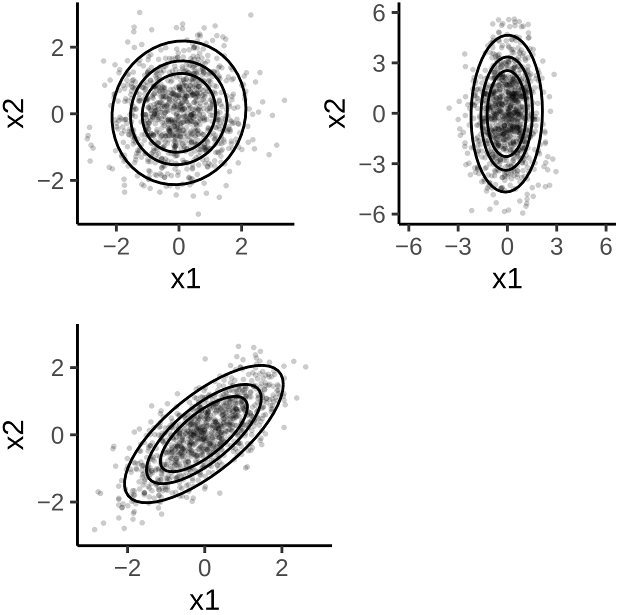
Multivariate normal distribution

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.

- $\mu =$

- $\Sigma =$

Multivariate normal distribution



Wald tests for multiple parameters

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

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

Wald tests for multiple parameters

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

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
```

Wald tests for multiple parameters

For the dengue example:

```r
C
```

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

```r
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
```

```r
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

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

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

$$\mathbf{A}X \sim$$

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

Example

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

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

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

```
## [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
```