

## Chapter 2: Linear Regression

Distribution of the RSS Values and Hypothesis Testing for 수정하기

---

JuHyun, Kang

February 16, 2025

The Three Sisters of Newton

School of Mathematics, Statistics and Data Science

Sungshin Women's University

1 Distribution of the RSS Values

2 Hypothesis Testing for  $\hat{\beta}_j \neq 0$

- Hat matrix defined by

$$\hat{y} = X\hat{\beta} = X(X^T X)^{-1} X^T y = Hy$$

$$H := X(X^T X)^{-1} X^T$$

- Some properties

1.  $H^2 = H$

2.  $(I - H)^2 = I - H$

3.  $HX = X$

4.  $H^T = H$

- RSS defined

$$\text{RSS} := \|y - \hat{y}\|^2$$

- Using hat matrix

$$\begin{aligned} y - \hat{y} &= y - Hy = (I - H)y = (I - H)(X\beta + \varepsilon) \\ &= (X - HX)\beta + (I - H)\varepsilon = (I - H)\varepsilon \end{aligned}$$

$$\|y - \hat{y}\|^2 = \{(I - H)\varepsilon\}^T (I - H)\varepsilon = \varepsilon^T (I - H)^2 \varepsilon = \varepsilon^T (I - H)\varepsilon$$

- They are only zeros and ones
- Dimensions of the eigenspaces of  $H$  and  $I - H$  are both  $p + 1$

**Proof** using  $\text{rank}(X) = p + 1$

$$\text{rank}(H) \leq \min\{\text{rank}(X(X^T X)^{-1}), \text{rank}(X)\} \leq \text{rank}(X) = p + 1$$

$$\text{rank}(H) \geq \text{rank}(HX) = \text{rank}(X) = p + 1$$

## Eigenvalues of $H$ and Null space of $(I - H)$

- Proof by contrapositive

$$Hx = x \Rightarrow (I - H)x = 0$$

$$(I - H)x = 0 \Rightarrow Hx = x$$

- Dimensions of the eigenspaces of  $H$  is  $p + 1$

**Proof** using  $\text{rank}(X) = p + 1$

$$\text{rank}(H) \leq \min\{\text{rank}(X(X^T X)^{-1}), \text{rank}(X)\} \leq \text{rank}(X) = p + 1$$

$$\text{rank}(H) \geq \text{rank}(HX) = \text{rank}(X) = p + 1$$

- Dimensions of the null space of  $I - H$  is  $N - (p + 1)$

$$P(I - H)P^T = \text{diag}(\underbrace{1, \dots, 1}_{N-p-1}, \underbrace{0, \dots, 0}_{p+1})$$

- We define  $v = P\varepsilon \in \mathbb{R}^N$ , then from  $\varepsilon = P^T v$

$$\text{RSS} = \varepsilon^T (I - H) \varepsilon = (P^T v)^T (I - H) P^T v = v^T P (I - H) P^T v$$

$$= [v_1, \dots, v_{N-p-1}, v_{N-p}, \dots, v_N] \begin{bmatrix} 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & \ddots & 0 & \dots & \dots & \vdots \\ \vdots & 0 & 1 & \dots & \dots & 0 \\ 0 & 0 & 0 & \dots & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & \dots & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_{N-p-1} \\ v_{N-p} \\ \vdots \\ v_N \end{bmatrix} = \sum_{i=1}^{N-p-1} v_i^2$$

- Let  $w \in \mathbb{R}^{N-p-1}$

- Average

$$E[v] = E[P\varepsilon] = 0$$

$$E[w] = 0$$

- Covariance

$$E[vv^t] = E[P\varepsilon(P\varepsilon)^T] = PE[\varepsilon\varepsilon^t]P = P\sigma^2IP^T = \sigma^2I$$

$$E[ww^T] = \sigma^2I$$

- We have RSS

$$\frac{RSS}{\sigma^2} \sim \chi_{N-p-1}^2$$



1 Distribution of the RSS Values

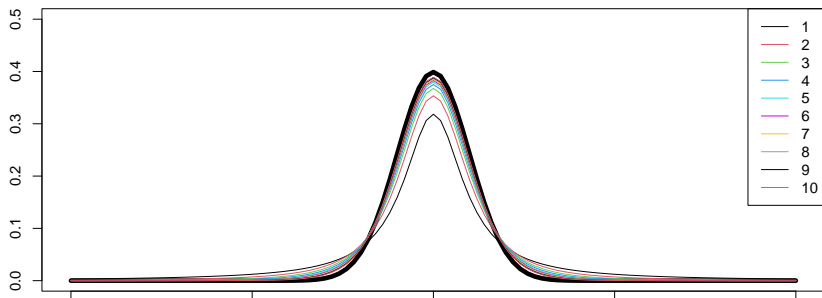
2 Hypothesis Testing for  $\hat{\beta}_j \neq 0$

## Test statistic $T$

- A  $t$  distribution with  $N - P - 1$  degrees of freedom
- We decide that hypothesis  $\beta_j = 0$  should be rejected.
- $U \sim N(0, 1)$ ,  $V \sim \chi_m^2$ ,

$$T \triangleq U / \sqrt{V/m}$$

```
curve(dnorm(x), -10, 10, ann = FALSE, ylim = c(0, 0.5), lwd = 5)
for(i in 1:10)curve(dt(x, df= i), -10, 10, col = i, add = TRUE, ann = FALSE)
legend("topright", legend = 1:10, lty = 1, col = 1:10)
```



- $\alpha = 0.01, 0.05$
- Reject the null hypothesis

# Q & A

**Thank you :)**