Test statistic of the test $H_0: \beta_1 = \beta_2, H_1: \beta_1 \neq \beta_2$

For a linear regression

$$Y = X\beta + \varepsilon$$

where the dimensions are $Y(N \times 1)$, $X(N \times k)$, $\beta(k \times 1)$, and $\varepsilon(N \times 1) \sim N(0, \sigma^2 I)$

The Least Square Estimator (LSE) is

$$\hat{\beta} = \hat{\beta}_{LSE} = (X^T X)^{-1} X^T Y$$

The distribution of the LSE s

$$\hat{\beta} \sim N_k(\beta, \sigma^2 (X^T X)^{-1})$$

And the coefficient of the i^{th} variable is

$$\hat{\beta}_i = u^T \hat{\beta}$$

where $u^T = [0,0,\ldots,1,\ldots,0]$, the i^{th} element is 1 and the rest are 0.

So

$$\hat{\beta}_i = u^T \hat{\beta} \sim N(u^T \beta, u^T [\sigma^2 (X^T X)^{-1}] u) = N(\beta_i, \sigma^2 A_{ii})$$

where A_{ij} is the element in the i^{th} row and j^{th} column of the matrix $(X^T X)^{-1}$.

Note that $N(\beta_i, \sigma^2 A_{ii})$ is a univariate distribution, both β_i and A_{ii} scalars.

So we have

$$\hat{\beta}_1 \sim N(\beta_1, \sigma^2 A_{11}) \text{ and } \hat{\beta}_2 \sim N(\beta_2, \sigma^2 A_{22})$$

Note that we CANNOT conclude that

$$\hat{\beta}_1 - \hat{\beta}_2 \sim N(\beta_1 - \beta_2, \sigma^2 A_{11} + \sigma^2 A_{22})$$

Because the variance part holds only when $\hat{\beta}_1$ and $\hat{\beta}_2$ are independent, which is not true here.

Instead, we consider $\hat{\beta}_1 - \hat{\beta}_2$ as a random variable

$$\hat{\beta}_1 - \hat{\beta}_2 = u^T \hat{\beta}$$

Here $u(k \times 1)$ is a matrix: the 1^{st} element is 1, the 2^{nd} element is -1, and the rest are 0.

$$u^T = [1, -1, 0, \ldots, 0]$$

(We can use this method to expression any linear combination of $\hat{\beta}_i$ and $\hat{\beta}_j$. For example, if the null hypothesis is $H_0: 3\beta_1 + 4\beta_3 = 0$, we consider $3\hat{\beta}_1 + 4\hat{\beta}_3 = u^T \hat{\beta}$ where $u^T = [3,0,4,\ldots,0]$.)

So

$$\hat{\beta}_1 - \hat{\beta}_2 = u^T \hat{\beta} \sim N(u^T \beta, u^T [\sigma^2 (X^T X)^{-1}] u)$$

The distribution of this random variable is

- $E(\hat{\beta}_1 - \hat{\beta}_2) = u^T\beta = \beta_1 - \beta_2,$
- $Var(\hat{\beta}_1 - \hat{\beta}_2)$

$$= u^T[\sigma^2(X^TX)^{-1}]u$$

$$= \sigma^2 \times [1,-1,0,\dots,0]\begin{bmatrix} A_{11} & A_{12} & \dots & A_{1k} \\ \vdots & \vdots & \vdots & \vdots \\ A_{k1} & A_{k2} & \dots & A_{kk} \end{bmatrix}\begin{bmatrix} 1 \\ -1 \\ \vdots \\ 0 \end{bmatrix}$$

$$= \sigma^2 \times [A_{11} - A_{21}, \quad A_{12} - A_{22}, \quad \dots, A_{1k} - A_{2k}]\begin{bmatrix} 1 \\ -1 \\ \vdots \\ 0 \end{bmatrix}$$

$$= \sigma^2 \times [A_{11} - A_{21} - (A_{12} - A_{22})]$$

$$= \sigma^2(A_{11} + A_{22} - 2A_{21})$$

So

$$\hat{\beta}_1 - \hat{\beta}_2 \sim N\big(\beta_1 - \beta_2, \sigma^2(A_{11} + A_{22} - 2A_{21})\big)$$

where A_{ij} is the element in the i^{th} row and j^{th} row of the matrix $(X^TX)^{-1}$.

And the test statistic is

$$\text{test statistic} = ts = \frac{(\hat{\beta}_1 - \hat{\beta}_2) - 0}{se(\hat{\beta}_1 - \hat{\beta}_2)} = \frac{\hat{\beta}_1 - \hat{\beta}_2}{\sqrt{\widehat{\sigma^2}(A_{11} + A_{22} - 2A_{21})}} \sim t_{(n-p)}$$