T-test for linear regression

Consider a system that can be represented by using a linear-in-parameter function as

$$y = w_0 + w_1 x_1 + \cdots + w_n x_n + e \tag{1}$$

where w_0 and $w_1, \ldots, w_n$ are parameters need to be estimated and $x_1, \ldots, x_n$ are input variables.

Equation (1) can be written into a matrix form as

$$\mathbf{Y} = \mathbf{XW} + \mathbf{e} \tag{2}$$

where

$$\mathbf{Y} = \begin{bmatrix} y(1) \\ \vdots \\ y(N) \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 1 & x_1(1) & \cdots & x_n(1) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_1(N) & \cdots & x_n(N) \end{bmatrix}, \mathbf{W} = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_n \end{bmatrix}, \mathbf{e} = \begin{bmatrix} e(1) \\ \vdots \\ e(N) \end{bmatrix}$$

and $\mathbf{e} \sim \mathrm{N}\left(\mathbf{0}, \sigma^2 \mathbf{I}_N\right)$, $\mathbf{I}_N$ is the $N \times N$ unit matrix

Assuming the evaluated coefficients are normally distributed

$$\hat{w}_j \sim \mathrm{N}\left(w_j, \sigma_j^{\,2}\right),\ j = 1, \ldots, n \tag{3}$$

in which

$$\hat{\mathbf{W}} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1} \mathbf{X}^{\mathrm{T}} (\mathbf{XW} + \mathbf{e}) = \mathbf{W} + \left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1} \mathbf{X}^{\mathrm{T}}\mathbf{e} \tag{4}$$

thus the mean value of $\hat{\mathbf{W}}$ is

$$\mathrm{E}(\hat{\mathbf{W}}) = \mathbf{W} + \left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1} \mathbf{X}^{\mathrm{T}}\mathrm{E}(\mathbf{e}) = \mathbf{W} \tag{5}$$

and the covariance of $\hat{\mathbf{W}}$ is

$$\begin{aligned} \mathrm{Cov}(\hat{\mathbf{W}}) &= \mathrm{E}\{[\hat{\mathbf{W}} - \mathrm{E}(\hat{\mathbf{W}})][\hat{\mathbf{W}} - \mathrm{E}(\hat{\mathbf{W}})]^{\mathrm{T}}\} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1} \mathbf{X}^{\mathrm{T}}\mathrm{E}(\mathbf{e}\mathbf{e}^{\mathrm{T}})\mathbf{X}\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1} \\ &= \left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1} \sigma^2 \end{aligned} \tag{6}$$

Therefore, for a specific parameter $\hat{w}_j$, the variance is

$$\mathrm{Var}(\hat{w}_j) = \left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}_{j+1,\,j+1} \sigma^2 \tag{7}$$

where $\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}_{1,1}$ is corresponding to the constant w_0.