Regressiong & General Factorial Design

1. Regression

1.1. Multiple Linear Regression Models

- Model $y = b_0 + b_1x_1 + b_2x_2 + \cdots + b_kx_k + e$
- Given a sample of n observations with k predictors

$$\{(x_{11}, x_{21}, \ldots, x_{k1}, y_1), \ldots, (x_{1n}, x_{2n}, \ldots, x_{kn}, y_n)\}$$

$$\begin{aligned}
y_1 &= b_0 - b_1x_{11} - b_2x_{21} - \cdots - b_kx_{k1} + e_1 \\
y_2 &= b_0 - b_1x_{12} - b_2x_{22} - \cdots - b_kx_{k2} + e_2 \\
&\vdots \\
y_n &= -b_0 - b_1x_{1n} - b_2x_{2n} - \cdots - b_kx_{kn} + e_n
\end{aligned}
\quad\Longrightarrow\quad
\begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix}
=
\begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{k1} \\ 1 & x_{12} & x_{22} & \cdots & x_{k2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{1n} & x_{2n} & \cdots & x_{kn} \end{bmatrix}
\begin{bmatrix} b_0 \\ b_1 \\ \cdot \\ \cdot \\ \cdot \\ b_k \end{bmatrix}
+
\begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ \cdot \\ e_n \end{bmatrix}$$

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$$

$$b = (X^TX)^{-1}X^Ty$$
$$SSE = y^Ty - b^Tx^Ty$$

1.2. Anova for Multiple Linear Regression Models

$$\text{SST} = \text{SSY} - \text{SS0} = \text{SSR} + \text{SSE}$$

- Degrees of freedom = Number of independent values required to compute

$$\begin{array}{ccccccccc}
\text{SST} & = & \text{SSY} & - & \text{SS0} & = & \text{SSR} & + & \text{SSE} \\
n-1 & = & n & - & 1 & = & k & + & (n-k-1)
\end{array}$$

$$\mathrm{MSR} = \frac{\mathrm{SSR}}{k} \qquad \text{and} \qquad \mathrm{MSE} = \frac{\mathrm{SSE}}{n-k-1}$$

Then using **F-test**

2. General Factorial Design