1 Definitions

1.0.1 Model Assumptions

The simple linear regression model with Normal errors assumes

- $y = X\beta + \epsilon$ The response is a linear function of the predictors.
- $\mathrm{E}(\epsilon) = 0$ The expectation of the error term is 0.
- $\mathrm{Cov}(\epsilon) = \sigma^2 I_n$ The error terms have constant variance σ^2 and are uncorrelated. This is equivalent to $\mathrm{Cov}(y) = \sigma^2 I_n$.

The normal error model strengthens the second two conditions:

- $\epsilon \sim \mathrm{Normal}(0, \sigma^2 I_n)$ The error terms are normally distributed.

1.0.2 Least Squares

Least squares estimates β by minimizing SSE, the error sum of squares.

$$Q(\hat{\beta}) = e'e$$

1.0.3 Gauss Markov Theorem

The least squares estimates are unbiased and have minimum variance of all linear unbiased estimators.

1.0.4 Design matrix

y can be decomposed into orthogonal vectors

$$y = \hat{y} + e.$$

Estimated $\hat{\beta}$.

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

H is the **hat matrix** because it puts the hat on y. Adorable. It projects y onto $\langle X \rangle$ the column space of the design matrix X.

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

1.0.5 Error vector

is y projected onto the space orthogonal to X.

$$e = y - \hat{y} = y - Hy = (I - H)y$$

1.1 Sums of Squares

1.1.1 SXX

- Variation in X.

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = (x - \bar{x}1)'(x - \bar{x}1) = x_c' x_c$$

1.1.2 SSE

- Error Sum of Squares. The degrees of freedom is the length of the β vector, which equals the number of regressors plus one for the intercept.

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = e'e$$

1.1.3 SSR

- Regression Sum of Squares

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = (\hat{y} - \bar{y}1)'(\hat{y} - \bar{y}1) = \hat{y}_c' \hat{y}_c$$

1.1.4 SSTO

- Total Sum of Squares, Define J_n to be an n by n matrix of all 1's.

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 = y'(I - \frac{1}{n}J_n)y = y_c' y_c$$

1.1.5 MSE

- Error Mean Squared. SSE with degrees of freedom correction. In the case of single regression with an intercept $df(SSE) = n - 2$. This estimates the variance of the errors.

$$MSE = \frac{SSE}{df(SSE)} \quad \mathrm{E}(MSE) = \sigma^2$$

1.2 Confidence Intervals

The following results hold for the normal error model.

1.2.1 Studentized pivotal quantity

In the normal error model with one predictor,

$$\frac{\hat{\beta}_1 - \beta_1}{SE(\hat{\beta}_1}) \sim t_{n-2}$$

Which allows the construction of a $1-\alpha$ confidence interval along with the corresponding t tests.

$$\hat{\beta}_1 + t(1 - \frac{\alpha}{2}; n-2) \times SE(\hat{\beta}_1)$$

1.2.2 Mean Response

Estimates the average y_h given x_h.

$$SE^2(\hat{y_h}) = \sigma^2(\frac{1}{n} + \frac{(x_h - \bar{x})^2}{x_c^t x_c})$$

A large range of x or large n makes for a tighter interval.

Does this hold for multiple regression?

1.2.3 Prediction

Predicts $y_{h(new)}$ given x_h.

$$SE^2(y_{h(new)}) = \sigma^2 + SE^2(\hat{y_h}) = \sigma^2(1 + \frac{1}{n} + \frac{(x_h - \bar{x})^2}{x_c^t x_c})$$

1.2.4 Estimating Whole line

The **Working-Hotelling multiplier** gives a confidence band for the entire regression line.

$$W = \sqrt{2F(1-\alpha; 2, n-2)}$$

The $1-\alpha$ confidence band is:

$$\hat{y_h} + -W \times SE(\hat{y_h}) \quad \forall x \in \mathbb{R}$$