

Multiple Regression

1. $\mathbb{E}(\mathbf{Y}|\mathbf{X}) = \mathbf{X}\beta$, $\text{Var}(\mathbf{Y}|\mathbf{X}) = \sigma^2\mathbf{I}_n$

Specially, in simple linear regression, $y_i = \beta_0 + \beta_1 x_i$,

$$\mathbf{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}, \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}, \mathbb{E}(\mathbf{Y}|\mathbf{X}) = \mathbf{1}_n\beta_0 + \mathbf{x}\beta_1, \text{Var}(\mathbf{Y}|\mathbf{X}) = \sigma^2\mathbf{I}_n$$

2. Least squares estimation: $Q = \sum_{i=1}^n \{y_i - \mathbb{E}(Y_i|\mathbf{X})\}^2 = \mathbf{y}^T\mathbf{y} - 2\mathbf{y}^T\mathbf{X}\beta + \beta^T\mathbf{X}^T\mathbf{X}\beta$

Least squares *unbiased* estimator: $\hat{\beta} = (\mathbf{X}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$

$$\mathbb{E}(\hat{\beta}|\mathbf{X}) = \beta, \text{Var}(\hat{\beta}|\mathbf{X}) = \sigma^2(\mathbf{X}^T\mathbf{X})^{-1}, \text{Var}(\mathbf{c}^T\hat{\beta}|\mathbf{X}) = \sigma^2\mathbf{c}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{c}$$

3. Vector of residuals: $\mathbf{e} = \mathbf{y} - \mathbf{X}\hat{\beta} = (\mathbf{I}_n - \mathbf{P}_\mathbf{X})\mathbf{y}$,

where $\mathbf{P}_\mathbf{X} = \mathbf{X}(\mathbf{X}\mathbf{X})^{-1}\mathbf{X}^T$ is $n \times n$, symmetric, idempotent, rank p ,

$$(\mathbf{I}_n - \mathbf{P}_\mathbf{X})\mathbf{X} = \mathbf{0}, (\mathbf{I}_n - \mathbf{P}_\mathbf{X})\mathbf{P}_\mathbf{X} = \mathbf{0}$$

$$\mathbb{E}(\mathbf{e}|\mathbf{X}) = (\mathbf{I}_n - \mathbf{P}_\mathbf{X})\mathbb{E}(\mathbf{Y}|\mathbf{X}) = (\mathbf{I}_n - \mathbf{P}_\mathbf{X})\mathbf{X}\beta = \mathbf{0}$$

$$\text{Var}(\mathbf{e}|\mathbf{X}) = (\mathbf{I}_n - \mathbf{P}_\mathbf{X})\sigma^2\mathbf{I}_n(\mathbf{I}_n - \mathbf{P}_\mathbf{X}) = \sigma^2(\mathbf{I}_n - \mathbf{P}_\mathbf{X})$$

$$\text{RSS} = \mathbf{e}^T\mathbf{e} = \mathbf{y}^T\mathbf{y} - \hat{\beta}^T\mathbf{X}^T\mathbf{y}, \hat{\sigma}^2 = \frac{\text{RSS}}{n-p} = \frac{\mathbf{y}^T\mathbf{y} - \hat{\beta}^T\mathbf{X}^T\mathbf{y}}{n-p}$$

4. Alternative formulation (for models with an intercept)

$$\mathbb{E}(Y_i|\mathbf{X}) = \gamma + \beta_1(x_{i1} - \bar{x}_1) + \beta_2(x_{i2} - \bar{x}_2) + \cdots + \beta_q(x_{iq} - \bar{x}_q)$$

$$\mathbb{E}(\mathbf{Y}|\mathbf{X}) = \gamma\mathbf{1}_n + \dot{\mathbf{X}}\dot{\beta}, \text{ where } \dot{\mathbf{X}}_{ij} = x_{ij} - \bar{x}_j, \dot{\beta} = (\beta_1 \cdots \beta_q)^T, \gamma = \beta_0 + \beta_1\bar{x}_1 + \cdots + \beta_q\bar{x}_q$$

$$\text{Least squares unbiased estimators: } \hat{\gamma} = \bar{y}, \hat{\beta} = (\dot{\mathbf{X}}_T\dot{\mathbf{X}})^{-1}\dot{\mathbf{X}}^T\mathbf{y}$$

$$\text{Var}(\hat{\beta}|\mathbf{X}) = \sigma^2(\dot{\mathbf{X}}^T\dot{\mathbf{X}})^{-1}, \text{Var}(\hat{\gamma}|\mathbf{X}) = n^{-1}\sigma^2, \text{cov}(\hat{\beta}, \hat{\gamma}|\mathbf{X}) = \mathbf{0}$$

5. Distributional results:

- $\hat{\beta} \sim N(\beta, \sigma^2(\mathbf{X}^T\mathbf{X})^{-1})$
- regression (model) SS $\mathbf{Y}^T\mathbf{P}_\mathbf{X}\mathbf{Y} \sim \sigma^2\chi^2(q, \sigma^{-2}\dot{\beta}^T\dot{\mathbf{X}}^T\dot{\mathbf{X}}\dot{\beta})$
- RSS $\mathbf{Y}^T(\mathbf{I}_n - \mathbf{P}_\mathbf{X})\mathbf{Y} \sim \sigma^2\chi^2(n - q - 1, 0)$
- RSS and regression SS are independent
- $\frac{\mathbf{c}^T\hat{\beta} - \mathbf{c}^T\beta}{\sigma\sqrt{\mathbf{c}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{c}}} \sim N(0, 1)$ and $\frac{\mathbf{c}^T\hat{\beta} - \mathbf{c}^T\beta}{\hat{\sigma}\sqrt{\mathbf{c}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{c}}} \sim t(n-p)$ (test hypotheses about linear funcs of the parameters)

6. 95% Confidence interval: $\mathbf{c}^T\hat{\beta} \pm t_{0.025}\hat{\sigma}\sqrt{\mathbf{c}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{c}}$

$$\text{CI for future response: } \mathbf{x}_*^T\hat{\beta} \pm t_{0.025}\hat{\sigma}\sqrt{\mathbf{x}_*^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_*}$$

$$\text{Prediction interval: } \mathbf{x}_*^T\hat{\beta} \pm t_{0.025}\hat{\sigma}\sqrt{1 + \mathbf{x}_*^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_*}$$

7. To test $\beta = \mathbf{0}$,

$$F = \frac{\text{regression SS}}{\text{RSS}} \sim F(p, n-p) \text{ (chk simple linear regression, but with different SS)}$$

8. To test a more general linear hypothesis about the coefficients of the model, $\mathbf{C}\beta = \mathbf{d}$,

$$\text{Extra SS} = (\mathbf{C}\hat{\beta} - \mathbf{d})^T (\mathbf{C}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{C}^T)^{-1} (\mathbf{C}\hat{\beta} - \mathbf{d})$$

$$F = \frac{(\text{ESS for } H_0)/c}{\text{RMS}} = \frac{(\text{RSS under } H_0 - \text{RSS under full model})/c}{\text{RMS}} \sim F(k, n-p)$$