1 Linear regression models

Simple linear regression model

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i \qquad i = 1,..,n \qquad \epsilon_i \sim N(0, \sigma^2)$$

Regression through the origin model

$$Y_i = \beta x_i + \epsilon_i \qquad i = 1,..,n \qquad \epsilon_i \sim N(0, \sigma^2)$$

Multiple linear regression model

$$Y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_{p-1} x_{i,p-1} + \epsilon_i \qquad i = 1,..,n \qquad \epsilon_i \sim N(0, \sigma^2)$$

Or

$$bY = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \qquad \boldsymbol{\epsilon} \sim MN(\mathbf{0}, \sigma^2 \mathbf{I})$$

---

2 Sums of squares

- Simple linear regression

$$S_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i y_i - \frac{1}{n}\Big(\sum_{i=1}^{n} x_i\Big)\Big(\sum_{i=1}^{n} y_i\Big)$$

$$S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - \frac{1}{n}\Big(\sum_{i=1}^{n} x_i\Big)^2$$

$$SST = S_{yy} = \sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - \frac{1}{n}\Big(\sum_{i=1}^{n} y_i\Big)^2$$

$$SSR = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 = \hat{\beta}_1^2 S_{xx}$$

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

- Multiple linear regression

$$SST = S_{yy} = \sum_{i=1}^{n}(y_i - \bar{y})^2 = \mathbf{Y}^T\big(\mathbf{I} - \frac{1}{n}\mathbf{J}\big)\mathbf{Y}$$

$$SSR = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 = \mathbf{Y}^T\big(\mathbf{H} - \frac{1}{n}\mathbf{J}\big)\mathbf{Y}$$

$$SSE = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \mathbf{Y}^T\big(\mathbf{I} - \mathbf{H}\big)\mathbf{Y}$$

Additive identity:	$SST = SSR + SSE$
Mean squares:	$MSR = SSR/(p-1)$ $\quad\quad \hat{\sigma}^2 = MSE = SSE/(n-p)$

3 Estimation and hat-values

Model	Coefficients	Variance
$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$	$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} \quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$	$\hat{\sigma}_{\hat{\beta}_1} = \frac{MSE}{S_{xx}}$
$Y_i = \beta x_i + \epsilon_i$	$\hat{\beta} = \frac{\sum y_i x_i}{\sum x_i^2}$	$\hat{\sigma}_{\hat{\beta}} = \frac{MSE}{\sum x_i^2}$
$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$	$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$	$\hat{\mathbf{\Sigma}}_{\hat{\boldsymbol{\beta}}} = MSE(\mathbf{X}^T\mathbf{X})^{-1}$

Model	Hat-values/matrix	Properties
$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$	$h_{ij} = \frac{1}{n} + \frac{1}{S_{xx}}(x_i - \bar{x})(x_j - \bar{x})$ $i, j = 1, 2, \ldots, n$	$h_{ij} = h_{ji}$ $\sum_{j=1}^{n} h_{ij} = 1$ $\sum_{j=1}^{n} h_{ij} x_j = x_i$ $\sum_{j=1}^{n} {h_{ij}}^2 = h_{ii}$
$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$	$\mathbf{H} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$	$\mathbf{H}^T = \mathbf{H}$ $\mathbf{H}\mathbf{X} = \mathbf{X}$ $\mathbf{H}^2 = \mathbf{H}$

4 Expectation and covariance properties for random vectors

- Let $E[\mathbf{Y}] = \mu$, $Var[\mathbf{Y}] = \mathbf{\Sigma}$ and $\mathbf{A}$ be a matrix of scalars. Then

$$E[\mathbf{W}] = E[\mathbf{AY}] = \mathbf{A}E[\mathbf{Y}] = \mathbf{A}\mu,$$

$$Var[\mathbf{W}] = Var[\mathbf{AY}] = \mathbf{A}Var[\mathbf{Y}]\mathbf{A}^T = \mathbf{A\Sigma A}^T.$$

- Let A be symmetric, then the *quadratic form* $\mathbf{Y}^T\mathbf{AY}$ has expectation

$$E[\mathbf{Y}^T\mathbf{AY}] = tr(\mathbf{A\Sigma}) + \mu^T\mathbf{A}\mu,$$

---

5 Inferential procedures for coefficients

<u>T-test (simple linear regression)</u>

Test statistic	Confidence interval
$t_{calc} = \frac{\hat{\beta}_1 - \beta_{10}}{\sqrt{\frac{MSE}{S_{xx}}}}$	$\hat{\beta}_1 \pm t_{\alpha/2,n-2}\sqrt{\frac{MSE}{S_{xx}}}$

<u>ANOVA table (simple and multiple linear regression)</u>

Source	Degrees of freedom	Sums of squares	Mean square	F-ratio
Regression	$p-1$	SSR	MSR	$f_{calc} = \frac{MSR}{MSE}$
Error	$n-p$	SSE	MSE	
Total	$n-1$	SST		

<u>General linear f-statistic (simple and multiple linear regression)</u>

$$f_{calc} = \left(\frac{SSE_R - SSE_F}{df_R - df_F}\right) \div \left(\frac{SSE_F}{df_F}\right)$$

<u>Inferential procedure for linear combination of $\boldsymbol{\beta}$</u>

- For known vector $\mathbf{c} \in \mathbb{R}^p$, define the linear parameter ψ and its estimator $\hat{\psi}$ by

$$\psi = \mathbf{c}^T\boldsymbol{\beta} \quad \text{and} \quad \hat{\psi} = \mathbf{c}^T\hat{\boldsymbol{\beta}}$$

- T-statistic for testing $H_0 : \psi = \psi_0$

$$t_{calc} = \frac{\hat{\psi} - \psi_0}{\hat{\sigma}_{\hat{\psi}}} = \frac{\mathbf{c}^T \hat{\boldsymbol{\beta}} - \psi_0}{\sqrt{\mathbf{c}^T (\widehat{\text{var}}[\hat{\boldsymbol{\beta}}])\mathbf{c}}} \qquad T_{calc} \sim t(df = n - p)$$

---

6 Correlation

Sample correlation coefficient	Coefficient of determination
$r = \frac{S_{xy}}{\sqrt{(S_{xx})(S_{yy})}}$	$r^2 = 1 - \frac{SSE}{SST}$ (R^2 is the same for multiple regression)

Inferential procedure for linear correlation

$$t_{calc} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

---

7 Maximum likelihood and likelihood-ratio test

- Likelihood function:

$$\mathcal{L}(\theta; y_1, y_2, \ldots, y_n) = f(y_1, y_2, \ldots, y_n|\theta) = f(y_1|\theta) \times f(y_2|\theta) \times \cdots \times f(y_n|\theta)$$

- Maximum likelihood estimator:

$$\hat{\theta} = \underset{\theta}{\text{argmax}} \mathcal{L}(\theta; y_1, y_2, \ldots, y_n) = \underset{\theta}{\text{argmax}} \mathcal{L}(\theta)$$

- Consider the null alternative pair $H_0 : \theta \in \Theta_0$ versus $H_A : \theta \in \Theta_0^C$. The likelihood ratio test statistic is

$$\lambda(y_1, y_2, \ldots, y_n) = \frac{\max_{\theta \in \Theta_0} \mathcal{L}(\theta; y_1, y_2, \ldots, y_n)}{\max_{\theta \in \Theta} \mathcal{L}(\theta; y_1, y_2, \ldots, y_n)}$$

Reject when $\lambda(y_1, y_2, \ldots, y_n) \leq c$, for some $0 \leq c \leq 1$.

8 Prediction

Simple: $\hat{y}_h = \hat{\beta}_0 + \hat{\beta}_1 x_h$ Multiple: $\hat{y}_h = \mathbf{X}_h^T \widehat{\boldsymbol{\beta}}$

	Confidence interval for EY_h	Prediction interval for future value $Y_{h(new)}$
Simple	$\hat{y}_h \pm t_{\alpha/2,n-2}\sqrt{MSE\left(\frac{1}{n} + \frac{(x_h-\bar{x})^2}{S_{xx}}\right)}$	$\hat{y}_h \pm t_{\alpha/2,n-2}\sqrt{MSE\left(1 + \frac{1}{n} + \frac{(x_h-\bar{x})^2}{S_{xx}}\right)}$
Multiple	$\hat{y}_h \pm t_{\alpha/2,n-p}\sqrt{MSE(\mathbf{X}_h^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}_h)}$	$\hat{y}_h \pm t_{\alpha/2,n-p}\sqrt{MSE(1 + \mathbf{X}_h^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}_h)}$

---

9 Other results

Normal density

$$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(y-\mu)^2 \right), \quad -\infty < y < \infty$$

Inverse of 2×2 matrix

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad \mathbf{A}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$