Homework 10 Redo

Nooreen Dabish

January 12, 2015

1 Problem 1: US Crime Data

1.1 Plot the scatterplot matrix between the variables.

1.2 Construct a linear model to study the relationship between Crime(Y) and Prob, adjusting for the effect of the 13 char. variables.

1.2.1 Description of the model:

The adjusted effect model is:

$$\mathbf{Y} = \beta_0 + \mathbf{X_1}\beta_1 + \ldots + \mathbf{X_{p-1}}\beta_{p-1} + \mathbf{X_p}\beta_p + \epsilon$$

where:

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} ; \mathbf{X} = \begin{pmatrix} 1 & X_{1,1} & \cdots & X_{1,p-1} \\ 1 & X_{2,1} & \cdots & X_{2,p-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n,1} & \cdots & X_{n,p-1} \end{pmatrix} ; \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{pmatrix} ; \epsilon = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

There are `15` covariates in the data set, so $p = 15$. There are \$ n = =47=\$ observations in the data set. Our model assumes that:

- $E(\epsilon) = 0$
- $\text{Var}(\epsilon) = \sigma^2\mathbf{I}$

*** E

1.2.2 Estimating the paramaters in the model:

$$\hat{\beta}_{LSE} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{Y}$$

and

$$\hat{\sigma}^2_{LSE} = \frac{\mathbf{Y}^{\mathrm{T}}\mathbf{Y} - \mathbf{Y}^{\mathrm{T}}\mathbf{P}_{\mathbf{X}}\mathbf{Y}}{n-p}$$

and

$$\hat{\mathrm{Var}}(\hat{\beta}_{LSE}) = \hat{\sigma}^2(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}$$

where
The projection matrix $\mathbf{P}_{\mathbf{X}}$ is $\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}$.

1.2.3 Test for the effect of Prob, the probability of imprisonment on Crime, adjusting for the other variables.

Prob is determined as the ratio of the number of commitments to the number of offenses.

- Hypotheses:

– **Null** $H_0 :\ \beta 14 = 0$

  - **Alternate** $H_A :\ \beta 14 \neq 0$

$$\begin{aligned}
\hat{\beta}_{LSE} &\sim N(\beta, \sigma^2(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}) \\
q^{\mathrm{T}}\hat{\beta}_{LSE} &\sim N(q^{\mathrm{T}}\beta, \sigma^2 q^{\mathrm{T}}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}q) \\
\frac{q^{\mathrm{T}}\hat{\beta}_{LSE} - q^{\mathrm{T}}\beta}{\hat{\sigma}\sqrt{q^{\mathrm{T}}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}q}} &\sim T_{n-p}\text{Under } H_0\text{:} \\
\frac{q^{\mathrm{T}}\hat{\beta}_{LSE}}{\hat{\sigma}\sqrt{q^{\mathrm{T}}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}q}} &\sim T_{n-p}
\end{aligned}$$

Rejection criterion

$$\left|\frac{q^{\mathrm{T}}\hat{\beta}_{LSE}}{\hat{\sigma}\sqrt{q^{\mathrm{T}}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}q}}\right| > t^{-1}_{n-p}(1-\alpha/2)$$

15.1	1	9.1	5.8	5.6	0.51	95	33	30.1	0.108	4.1	3940	26.1	0.084602
14.3	0	11.3	10.3	9.5	0.583	101.2	13	10.2	0.096	3.6	5570	19.4	0.029599
14.2	1	8.9	4.5	4.4	0.533	96.9	18	21.9	0.094	3.3	3180	25	0.083401
13.6	0	12.1	14.9	14.1	0.577	99.4	157	8	0.102	3.9	6730	16.7	0.015801
14.1	0	12.1	10.9	10.1	0.591	98.5	18	3	0.091	2	5780	17.4	0.041399
12.1	0	11	11.8	11.5	0.547	96.4	25	4.4	0.084	2.9	6890	12.6	0.034201

The $(1-\alpha)$ Confidence Interval of $q^{\mathrm{T}}\beta$

$$q^{\mathrm{T}}\hat{\beta} \pm t_{n-p}^{-1}(1-\alpha/2)s\sqrt{q^{\mathrm{T}}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}q}$$