Use this page for extra answer or working space

Answer to Q1:

(a) The proposed model is

$$Y_i = \beta_0 + \beta_1 x_{i1} + \epsilon_i$$

so that the fitted value under the model is $\beta_0 + \beta_1 x_{i1}$. Therefore, the vertical mis-fit is $y_i - (\beta_0 + \beta_1 x_{i1})$, which on aggregating the squared values takes

$$S(\beta_0, \beta_1) = \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_{i1}))^2.$$

Least squares estimation proceeds by minimizing this function with respect to the β parameters.

5 Marks

(b) The least squares function is, in vector form,

$$S(\beta) = (\mathbf{y} - \mathbf{X}\beta)^\top (\mathbf{y} - \mathbf{X}\beta)$$

which, on differentiation, yields

$$\frac{\partial S(\beta)}{\partial \beta} = -2\mathbf{X}^\top (\mathbf{y} - \mathbf{X}\beta).$$

Equating to zero and solving, we see that at the solution

$$\mathbf{X}^\top (\mathbf{y} - \mathbf{X}\widehat{\beta}) = \mathbf{0}_2.$$

5 Marks

(c) The solution to the least squares estimation problem can be written

$$\widehat{\beta} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X} \mathbf{y} = \mathbf{A}\mathbf{y}$$

say, which is a linear transform of $\mathbf{y}$. The fitted values are

$$\widehat{\mathbf{y}} = \mathbf{X}\widehat{\beta} = \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X} \mathbf{y} = \mathbf{H}\mathbf{y}$$

which is also a linear combination.

4 Marks

(d) Here are the full results:

```
> summary(lm(y~x))
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.4928     1.7799   0.839   0.4293
x             0.3635     0.1388   2.619   0.0345 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.704 on 7 degrees of freedom
Multiple R-squared:  0.4948,    Adjusted R-squared:  0.4227
F-statistic: 6.857 on 1 and 7 DF,  p-value: 0.03448
```

The answers are as follows:

(i) the sample size n; $n = 9$ (from line 9) 1 Mark

(ii) the `Estimate` omitted from line 4: 1.4928; 1 Mark

(iii) the conclusion of the test of the null hypothesis $\text{H}_0 \ : \ \beta_1 = 0$ omitted from line 5: reject the null hypothesis as the test statistic is significant (note the p-value on line 11 which is identical to the p-value in the test referred to). 1 Mark

(iv) whether `x` is a useful predictor of `y`: it is useful, based on the test result above and the R^2 statistic on line 10. 1 Mark

(v) the three terms in the sums of squares decomposition

$$\text{SS}_\text{T} = \text{SS}_\text{Res} + \text{SS}_\text{R} :$$

We can compute

$$\text{SS}_\text{Res} = (n-2)\widehat{\sigma}^2 = 7 \times 1.704^2 = 20.32531.$$

Then,

$$\text{SS}_\text{T} = \frac{\text{SS}_\text{Res}}{1 - R^2} = \frac{20.32531}{1 - 0.4948} = 40.23221$$

so therefore

$$\text{SS}_\text{R} = \text{SS}_\text{T} - \text{SS}_\text{Res} = 40.23221 - 20.32531 = 19.9069.$$

2 Marks

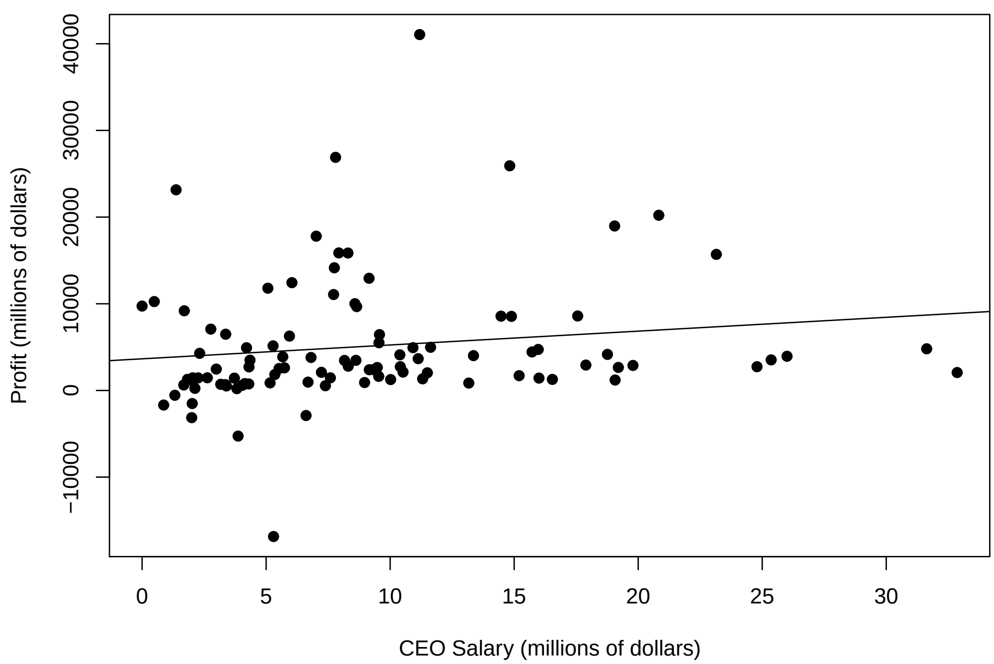
Q2 The following data records Chief Executive Officer (CEO) salary (`CEO.salary`) and annual company profit (`Profit`), each in millions of US dollars, for the most highly paid 100 CEOs in the US for 2012.

The objective of the analysis is to understand whether there is a relationship between annual profit, y, and CEO salary, x. An analysis in `R` is presented below: some values in have been omitted from the output.

```
> fit.CEO<-lm(Profit~CEO.salary)
> summary(fit.CEO)
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   3648.1     1216.1   3.000  0.00343 **
CEO.salary     159.7      105.9   XXXXX  XXXXXXX
---
Signif. codes:  0  ***  0.001  **  0.01  *  0.05  .  0.1     1

Residual standard error: 7301 on 98 degrees of freedom
Multiple R-squared:  0.02268,   Adjusted R-squared:  0.01271
F-statistic:  XXXX  on 1 and 98 DF,  p-value: XXXXX
```

The straight line fit is depicted below.



(a) Is there evidence that the profit of a company increases significantly with CEO salary ? Justify your answer.

Note that the 0.95 and 0.975 quantiles of the Student-t distribution with 98 degrees of freedom are 1.6606 and 1.9845 respectively.

6 Marks

Question continued on the next page.