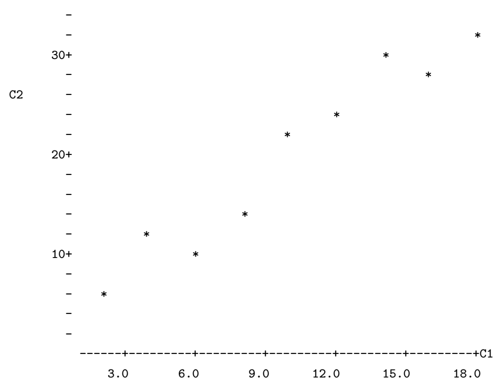
Solutions for Homework 11

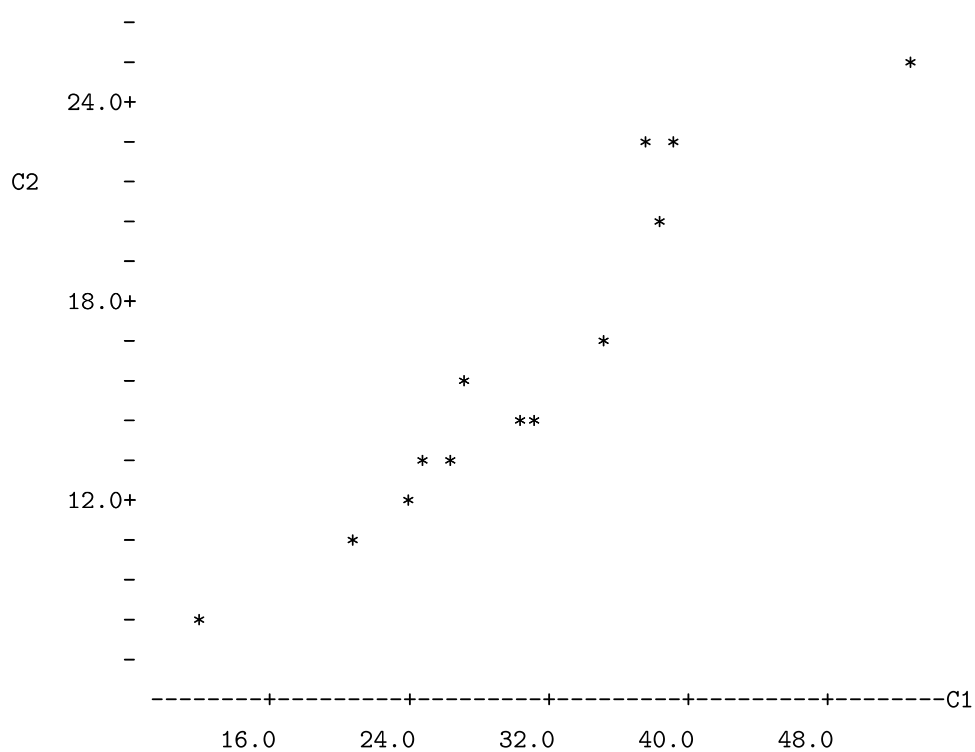
1. $\bar{x} = 10$, $\bar{y} = 19.333$,

   $$\hat{b_1} = \frac{\sum_{i=1}^{9}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{9}(x_i - \bar{x})^2} = 1.7417, \quad \hat{b_0} = \bar{y} - \hat{b_1}\bar{x} = 1.916.$$

```
MTB > plot c2 c1
```



2. (a).



(b). Model: $y_i = b_0 + b_1 x_i + e_i, i = 1, \ldots, 13$.
Here $e_i \sim N(0, \sigma^2)$, and e_i are independent.

(c). The summary statistics are: $\sum_{i=1}^{13} x_i = 401.5$, $\sum_{i=1}^{13} y_i = 211.7$, $\sum_{i=1}^{13} x_i^2 = 13608.11$, $\sum_{i=1}^{13} y_i^2 = 3761.83$, $\sum_{i=1}^{1} x_i y_i = 7119.95$,.

Then least square estimates are:

$$\begin{aligned}
\hat{b_1} &= \frac{\sum_{i=1}^n (x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^n (x_i-\bar{x})^2} = \frac{\sum_{i=1}^n x_i y_i - \frac{\sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n}}{\sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n}} \\
&= 0.482 \\
\hat{b_0} &= \bar{y} - \hat{b_1}\bar{x} \\
&= 1.413
\end{aligned}$$

95% Confidence Interval for b_1

$$\hat{b_1} - T_{n-2,\alpha/2}\frac{\sqrt{MSE}}{\sqrt{\sum(x_i-\bar{x})^2}} \le b_1 \le \hat{b_1} + T_{n-2,\alpha/2}\frac{\sqrt{MSE}}{\sqrt{\sum(x_i-\bar{x})^2}}$$

$$0.482 - 2.2011.76634.755 \le b_1 \le 0.482 + 2.201\frac{1.766}{34.755}$$

$$0.370 \le b_1 \le 0.593$$

95 % Confidence Interval for b_0

$$\hat{b_0} - T_{n-2,\alpha/2}\sqrt{MSE}\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum(x_i-\bar{x})^2}} \le b_0 \le \hat{b_0} + T_{n-2,\alpha/2}\sqrt{MSE}\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum(x_i-\bar{x})^2}}$$

$$1.413 - 2.201\sqrt{3.12}\sqrt{\frac{1}{13} + \frac{30.885^2}{1207.937}} \le b_0 \le 1.413 + 2.201\sqrt{3.12}\sqrt{\frac{1}{13} + \frac{30.885^2}{1207.937}}$$

$$-2.206 \le b_0 \le 5.032$$

(d). 95 % Confidence Interval for the variance of Y about the theoretical straight line

$$\frac{SSE}{\chi^2_{n-2,\alpha/2}} \le \sigma_e^2 \le \frac{SSE}{\chi^2_{n-2,1-\alpha/2}}$$

$$34.2721.92 \le \sigma_e^2 \le \frac{34.27}{3.82}$$

$$1.563 \le \sigma_e^2 \le 8.971.$$

(e). 95 % Confidence Interval for Population Mean Value Corresponding to a Given $x_* = 20$

$$(\hat{b_0} + \hat{b_1}x_*) \pm T_{n-2,\alpha/2}\sqrt{MSE}\sqrt{\frac{1}{n} + \frac{(x_*-\bar{x})^2}{\sum(x_i-\bar{x})^2}}$$

$$11.043 \pm 2.201\sqrt{3.12}\sqrt{\frac{1}{13} + \frac{(118.483)}{1207.937}}$$

$$(9.416, 12.669).$$

(f). 95 % Confidence Interval for the a Predicted Value Corresponding to a Given $x_* = 20$

$$(\hat{b_0} + \hat{b_1}x_*) \pm T_{n-2,\alpha/2}\sqrt{MSE}\sqrt{1 + \frac{1}{n} + \frac{(x_*-\bar{x})^2}{\sum(x_i-\bar{x})^2}}$$

$$11.043 \pm 2.201\sqrt{3.12}\sqrt{1 + \frac{1}{13} + \frac{(118.483)}{1207.937}}$$

$$(6.828, 15.267).$$

This confidence interval is wider than the one we got in part (e).

(g).

```
         -
resid    -                                    *
         -                                      *
         -
      2.0+
         -     *
         -                        *
         -                                     *
         -
      0.0+
         -                     *
         -                    * *          *
         -                *
         -                            *                         *
     -2.0+                           *
         -
         -
           ------+---------+---------+---------+---------+---------+
               8.0      12.0      16.0      20.0      24.0      28.0
                             fitted values
```

From the above residual plot, we can see there is no obvious departure from the assumptions of the linear model.

3. (a) From the equation $\hat{b_1} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = \frac{\sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n}}{\sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n}}$
$= \frac{-4y_1+5+24}{26} = 0.5weget\mathrm{y}_1 = 4$. Then from equation

$$\hat{b}_0 = \bar{y} - \hat{b_1}\bar{x} = 6$$

we get $y_2 = 7$..

(b). The degree of freedom for errors is 2. Only the two values of y can change freely. The missing values are determined by the given two values. So the df = 2.