MATH 308 Assignment 19: Exercises 9.7

Nakul Joshi

April 22, 2014

7

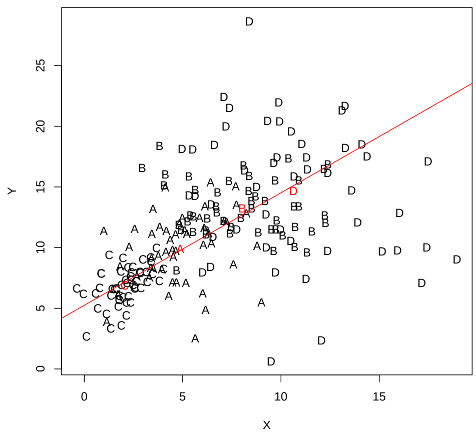
a)

$\rho \approx 0.50$

b)

	A	B	C	D
$\overline{X}$	4.88	8.03	2.06	10.64
$\overline{Y}$	9.91	13.24	6.89	14.68

c)



$\rho = 0.99$, which is much higher than the result from (a).

9

$$
\begin{aligned}
g(a,b) &= \Sigma(y_i - \hat{y}_i)^2 \\
\implies \frac{\partial g}{\partial a} &= \Sigma\left(2(y_i - \hat{y}_i)\frac{\partial(y_i - (a + bx_i))}{\partial a}\right) \\
\implies 0 &= -2\Sigma(y_i - \hat{y}_i) \\
\implies \Sigma(y_i - \hat{y}_i) &= 0 \quad \square
\end{aligned}
$$

10

a)

$$
\begin{aligned}
r &= \frac{ss_{xy}}{\sqrt{ss_x ss_y}}, b = \frac{ss_{xy}}{ss_x} \\
\Sigma(\hat{y}_i - \overline{y})^2 &= \Sigma(a + bx_i - (a + b\overline{x}))^2 \\
&= b^2\Sigma(x_i - \overline{x})^2 = b^2 ss_x = \frac{ss_{xy}^2}{ss_x} \\
&= r^2 ss_y \\
\implies s_{\hat{y}} &= \sqrt{\frac{\Sigma(\hat{y}_i - \overline{y})^2}{n-1}} = r\sqrt{\frac{ss_y}{n-1}} \\
&= r\sqrt{\frac{\Sigma(y_i - \overline{y})^2}{n-1}} = rs_y \quad \square
\end{aligned}
$$

b)

$$
\begin{aligned}
\overline{e} &= \Sigma(y_i - \hat{y}_i)/n = 0 \\
\implies s_e &= \sqrt{\frac{\Sigma(y_i - \hat{y}_i - 0)^2}{n-1}}
\end{aligned}
$$

$$
\begin{aligned}
\implies & (n-1)s_e^2 \\
&= \Sigma(y_i - \hat{y}_i)^2 \\
&= \Sigma(y_i - \overline{y} + \overline{y} - \hat{y}_i)^2 \\
&= \Sigma(y_i - \overline{y})^2 + \Sigma(\overline{y} - \hat{y}_i)^2 \\
&- 2\Sigma(y_i - \overline{y})(\hat{y}_i - \overline{y}) \\
&= ss_y + r^2 ss_y - 2\Sigma(y_i\hat{y}_i - \cancel{y_i}\overline{y} - \hat{y}_i\overline{y} + \overline{y}^{\cancel{2}}) \\
&= ss_y(1 + r^2) - 2\Sigma(\hat{y}_i\cancel{(y_i - \overline{y})}) \\
\implies & s_e = \sqrt{\frac{(1+r^2)ss_y}{n-1}} = \sqrt{1+r^2}s_y \quad \square
\end{aligned}
$$

11

a)

$$
\begin{aligned}
b = r\frac{s_w}{s_h} &\approx 1.6 a = \overline{w} - b\overline{h} \approx 20.1 \\
\implies \hat{w} &= a + bh \approx 20.1 + 1.6h
\end{aligned}
$$

b)

$\hat{w}(5 \text{ ft}) = \hat{w}(60 \text{ in}) = 116.5$ pounds

c)

$r^2 \approx 0.56$, which means that the model explains 56% of the variability of the weights of the individuals.

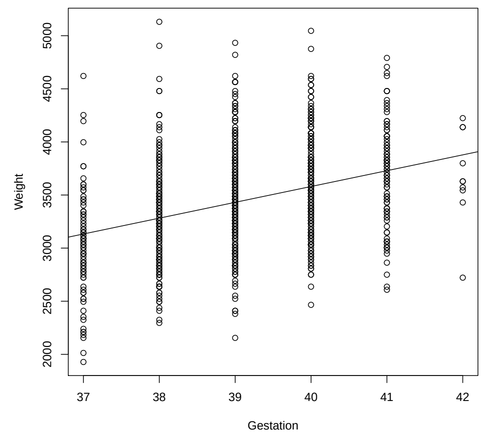
17

a)

Correlation is ≈ 0.35

b)

$y = -2380 + 149x$

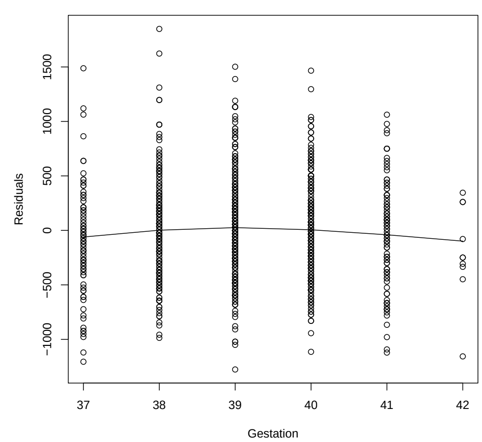


c)

The slope is interpreted as 'a week increase in gestation period is correlated with a 149g increase in weight'.

The $r^2 \approx 0.12$ is interpreted as 'the linear model explains 12% of the variation in weights'.

d)



The residual plot shows a curve peaking at week 39. The linear model is therefore inappropriate.

18

a)

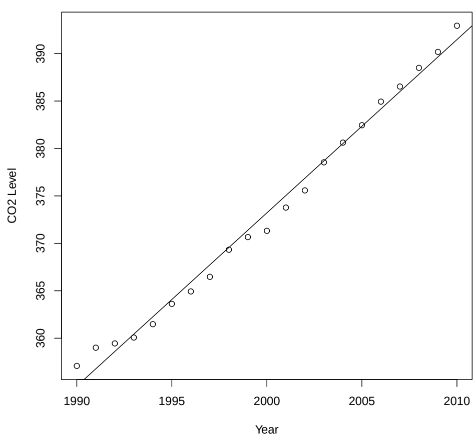
457.4 g

b)

$1-\alpha = 0.95 \implies 1-\alpha/2 = 0.975$
$\implies t \equiv t_{n-2,1-\alpha/2} = t_{1007,0.975} \approx 1.96$
$\therefore$ 95% CI: $\hat{\beta} \pm \frac{tS}{\sqrt{ss_x}} = 149 \pm 24.8$
$= (124.2, 173.8)$

21

a)

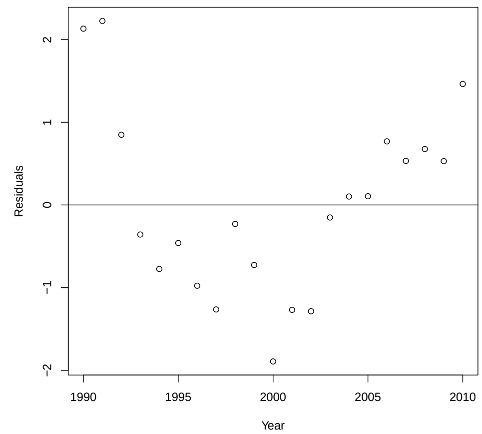


The data is strongly linear, with a correlation of over 0.99.

b)

$$y = -3279.593 + 1.826x$$

c)



At first, it appears that there is a pattern to the residuals, decreasing until year 2000 and then increasing again. However, looking at the scale of the y-axis we see that the largest residual is only 2, but the CO_2 Levels are of order 350.

Thus, the apparent pattern can be put down to random error and ignored, making a linear model appropriate.