Simple Linear Regression (Chapter #11)

for part (c) Residuals

pg.395

1)

Temperature, x (x_i)	Converted Sugar, y (y_i)	$e_i = y_i - \hat{y}_i$
1.0	8.1	$8.1 - [6.4136 + 1.8091(1)] = -0.12$
1.1	7.8	-0.60
1.2	8.5	-0.08
1.3	9.8	1.03
1.4	9.5	0.55
1.5	8.9	-0.22
1.6	8.6	-0.71
1.7	10.2	0.71
1.8	9.3	-0.36
1.9	9.2	-0.65
2.0	10.5	0.47

pg.396

(a) $\sum_i x_i = 16.5$, $\sum_i y_i = 100.4$, $\sum_i x_i^2 = 25.85$, $\sum_i x_i y_i = 152.59$

$n = 11$

so,

pg.396

$$b_1 = \frac{n \sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

$$b_1 = \frac{(11)(152.90) - (16.5)(100.4)}{(11)(25.85) - (16.5)^2} = 1.8091$$

pg.396

$$b_0 = \frac{\sum_{i=1}^{n} y_i - b_1 \sum_{i=1}^{n} x_i}{n}$$

$$b_0 = \frac{100.4 - (1.8091)(16.5)}{11} = 6.4136$$

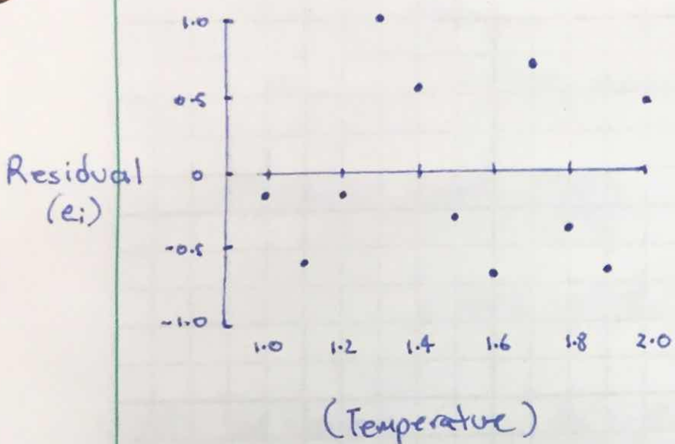
so, linear regression line : $\hat{y} = 6.4136 + 1.8091(x)$

(b) For $x = 1.75$,

$\hat{y} = 6.4136 + 1.8091(1.75)$

$\hat{y} = 9.580$

(c) Plot: Residuals .vs. Temperature



→ Please see table on first sheet for residual values.

→ No pattern observed in the residuals.

(d) $S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2 = 25.85 - \frac{165^2}{11} = 1.1$

pg. 402

$S_{yy} = \sum_{i=1}^{n} (y_i - \bar{y})^2 = 923.58 - \frac{100.4^2}{11} = 7.2018$

$S_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) = 152.59 - (165)\frac{(100.4)}{11} = 1.99$

pg. 402 $s^2 = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n-2} = \frac{S_{yy} - b_1 S_{xy}}{n-2} = \frac{7.2018 - (1.8091)(1.99)}{9}$

$s^2 = 0.40$

(e) 95% C.I. for β_0.

pg. 403 $b_0 - t_{\frac{\alpha}{2}} \cdot s \cdot \sqrt{\frac{\sum x_i^2}{n \cdot S_{xx}}} < \beta_1 < b_0 + t_{\frac{\alpha}{2}} \cdot s \cdot \sqrt{\frac{\sum x_i^2}{n \cdot S_{xx}}}$

so, $\alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025$, $(n-2)$ DOF

$t_{0.025, 11-2} = t_{0.025, 9} \longrightarrow$ from t-table: 2.262

so,

$6.4136 - 2.262\sqrt{0.4} \cdot \sqrt{\frac{(25.85)}{11(1.1)}} < \beta_0 < 6.4136 + 2.262 \cdot \sqrt{0.4} \cdot \sqrt{\frac{(25.85)}{11(1.1)}}$

$4.324 < \beta_0 < 8.503$

(f) 95% C.I. for β_1

pg. 403 $$b_1 - t_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{S_{xx}}} < \beta_1 < b_1 + t_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{S_{xx}}}$$

$t_{0.025, 9} \longrightarrow 2.262$ (from t-table)

$$1.8091 - 2.262 \cdot \frac{\sqrt{0.4}}{\sqrt{1.1}} < \beta_1 < 1.8091 + 2.262 \cdot \frac{\sqrt{0.4}}{\sqrt{1.1}}$$

$$0.446 < \beta_1 < 3.172$$

(g) $\alpha = 0.05$

$H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$

pg. 415

Analysis-of-variance Approach. (F-test)

$\Rightarrow f_{\alpha,(1,n-2)} = f_{0.05,(1,9)} \Rightarrow 5.12$ (From F-table)

pg. 414 $SSR = b_1 \cdot S_{xy} = (1.8091)(1.99) = 3.60$

pg. 414 $SSE = S_{yy} - SSR = 7.20 - 3.60 = 3.60$

Construct table:

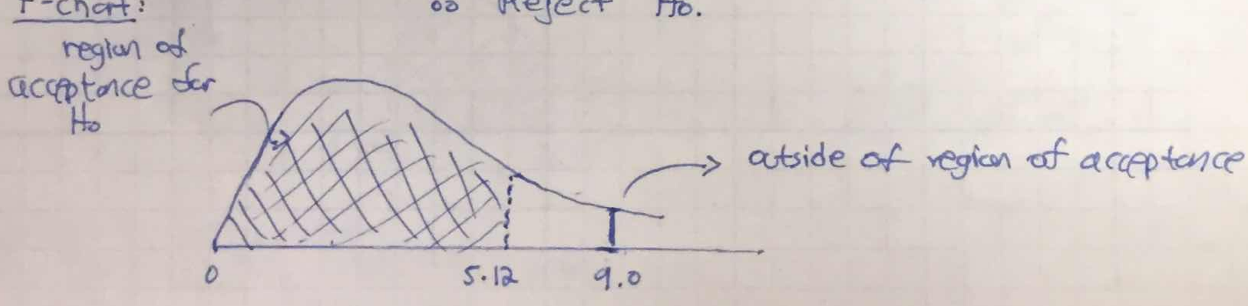
	Source of Variation	Sum of Squares	DoF	Mean Square	Computed f
pg. 415	Regression	SSR = 3.60	1	3.60	$f_c = SSR/s^2 = 9.0$
	Error	SSE = 3.60	n-2 = 9	$s^2 = 0.40$	
	Total	7.20	10		

so, compare the computed f with the significance f value.

$$f_c = 9.0 > f_{\alpha,(1,n-2)} = 5.12$$

$\therefore$ Reject H_0.

F-chart:



2)

(x_i) Math Grade	70	92	80	74	65	83
(y_i) English Grade	74	84	63	87	78	90

(a) Correlation coeff.

$$r = b_1 \sqrt{\frac{S_{xx}}{S_{yy}}} = \frac{S_{xy}}{\sqrt{S_{xx} \cdot S_{yy}}} \quad \text{pg. 432}$$

pg. 402 $S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2$

$$= \sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n} = 36{,}354 - 35{,}882.667 = 471.333$$

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

$$= \sum y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n} = 38{,}254 - 37{,}762.667 = 491.333$$

$$S_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) = 36{,}926 - 36{,}810.667 = 115.333$$

so, $r = \dfrac{115.333}{\sqrt{(471.333)(491.333)}} = 0.240.$

pg. 432

∴ The correlation coeff. between the math grade and english grade is only about 24%.

→ There doesn't seem to be a strong linear relationship.

(b) $\alpha = 0.05$

(2-tail test)

pg. 434

$H_0: \rho = 0$

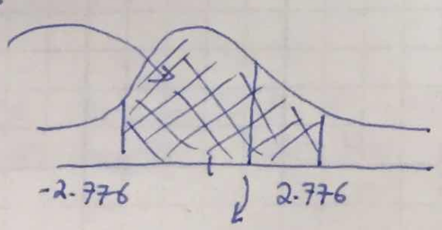
$H_1: \rho \neq 0$

critical significance region

$t_{\frac{\alpha}{2}, n-2} = t_{0.025, 4}$

→ $t = 2.776$ (from t-table)

acceptance region for H_0.



$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \quad \text{pg. 434}$$

$$t = \frac{0.240\sqrt{4}}{\sqrt{1-0.240^2}} = 0.49$$

∴ since the computed t value is in the acceptance region we don't reject H_0.

3)

(x_i) Weight (kg)	(y_i) Chest size (cm)
2.75	29.5
2.15	26.3
4.41	32.2
5.52	36.5
3.21	27.2
4.32	27.7
2.31	28.3
4.30	30.3
3.71	28.7

(a) $r = ?$ $\quad r = \dfrac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}$ pg. 432

pg. 402 $\quad S_{xx} = 128.6602 - \dfrac{32.68^2}{9} = 9.9955$

$$S_{yy} = 7980.83 - \frac{266.7^2}{9} = 77.62$$

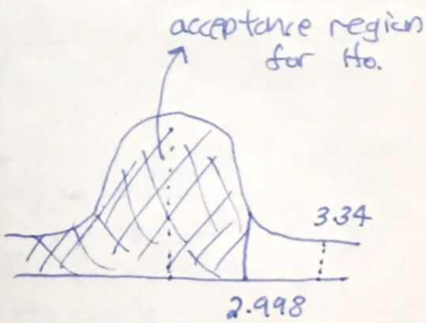
$$S_{xy} = 990.268 - \frac{(32.68)(266.7)}{9} = 21.8507$$

$$r = \frac{21.8507}{\sqrt{(9.9955)(77.62)}} = 0.784$$

(b) $\alpha = 0.01$ $\quad H_0: \rho = 0$ pg. 434 one-tail test.

$H_1: \rho > 0$

$t_{\alpha, n-2} = t_{0.01, 7} = 2.998$ (from t table)



$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$ pg. 434

$$t = \frac{(0.784)\sqrt{7}}{\sqrt{1-0.784^2}} = 3.34$$

∴ since the calculated t-value is outside of the acceptance region for Ho. we reject Ho.

(c) Variation in chest sizes explained by difference in weight:

$r^2 \cdot (100\%) = (0.784)^2 \cdot (100\%) = 61.5\%$

pg. 433