Chap 12

12.

a. $S_{xx} = 39{,}095 - \frac{(517)^2}{14} = 20{,}002.929$, $S_{xy} = 25{,}825 - \frac{(517)(346)}{14} = 13047.714$;

$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{13{,}047.714}{20{,}002.929} = .652$; $\hat{\beta}_0 = \frac{\Sigma y - \hat{\beta}_1 \Sigma x}{n} = \frac{346 - (.652)(517)}{14} = .626$, so the equation of the least squares regression line is $y = .626 + .652x$.

b. $\hat{y}_{(35)} = .626 + .652(35) = 23.456$. The residual is $y - \hat{y} = 21 - 23.456 = -2.456$.

c. $S_{yy} = 17{,}454 - \frac{(346)^2}{14} = 8902.857$, so

$SSE = 8902.857 - (.652)(13047.714) = 395.747$. $\hat{\sigma} = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{395.747}{12}} = 5.743$.

d. $SST = S_{yy} = 8902.857$; $r^2 = 1 - \frac{SSE}{SST} = 1 - \frac{395.747}{8902.857} = .956$.

e. Without the two upper extreme observations, the new summary values are $n = 12, \Sigma x = 272, \Sigma x^2 = 8322, \Sigma y = 181, \Sigma y^2 = 3729, \Sigma xy = 5320$. The new $S_{xx} = 2156.667, S_{yy} = 998.917, S_{xy} = 1217.333$. New $\hat{\beta}_1 = .56445$ and $\hat{\beta}_0 = 2.2891$, which yields the new equation $y = 2.2891 + .56445x$. Removing the two values changes the position of the line considerably, and the slope slightly. The new $r^2 = 1 - \frac{311.79}{998.917} = .6879$, which is much worse than that of the original set of observations.

19. $n = 14$, $\Sigma x_i = 3300$, $\Sigma y_i = 5010$, $\Sigma x_i^2 = 913{,}750$, $\Sigma y_i^2 = 2{,}207{,}100$, $\Sigma x_i y_i = 1{,}413{,}500$

a. $\hat{\beta}_1 = \frac{3{,}256{,}000}{1{,}902{,}500} = 1.71143233$, $\hat{\beta}_0 = -45.55190543$, so we use the equation $y = -45.5519 + 1.7114x$.

b. $\hat{\mu}_{Y\cdot 225} = -45.5519 + 1.7114(225) = 339.51$

c. Estimated expected change $= -50\hat{\beta}_1 = -85.57$

d. No, the value 500 is outside the range of x values for which observations were available (the danger of extrapolation).

38.

a. From Exercise 23, which also refers to Exercise 19, SSE = 16.205.45, so $s^2 = 1350.454$, $s = 36.75$, and $s_{\hat{\beta}_1} = \frac{36.75}{368.636} = .0997$. Thus $t = \frac{1.711}{.0997} = 17.2 > 4.318 = t_{.0005,14}$, so p-value $< .001$. Because the p-value $< .01$, $H_o : \beta_1 = 0$ is rejected at level .01 in favor of the conclusion that the model is useful $(\beta_1 \neq 0)$.

b. The C.I. for $\beta_1$ is $1.711 \pm (2.179)(.0997) = 1.711 \pm .217 = (1.494, 1.928)$. Thus the C.I. for $10\beta_1$ is $(14.94, 19.28)$.

Chap 12

52.

a. We wish to test $H_o : \beta_1 = 0$ vs $H_a : \beta_1 \neq 0$. The test statistic $t = \frac{10.6026}{.9985} = 10.62$ leads to a p-value of < .006 ( 2P( t > 4.0 ) from the 7 df row of table A.8), and $H_o$ is rejected since the p-value is smaller than any reasonable $\alpha$. The data suggests that this model does specify a useful relationship between chlorine flow and etch rate.

b. A 95% confidence interval for $\beta_1$: $10.6026 \pm (2.365)(.9985) = (8.24,12.96)$. We can be highly confident that when the flow rate is increased by 1 SCCM, the associated expected change in etch rate will be between 824 and 1296 A/min.

c. A 95% CI for $\mu_{Y\cdot 3.0}$: $38.256 \pm 2.365\left(2.546\sqrt{\frac{1}{9}+\frac{9(3.0-2.667)^2}{58.50}}\right)$

$= 38.256 \pm 2.365(2.546)(.35805) = 38.256 \pm 2.156 = (36.100,40.412)$, or 3610.0 to 4041.2 A/min.

d. The 95% PI is $38.256 \pm 2.365\left(2.546\sqrt{1+\frac{1}{9}+\frac{9(3.0-2.667)^2}{58.50}}\right)$

$= 38.256 \pm 2.365(2.546)(1.06) = 38.256 \pm 6.398 = (31.859,44.655)$, or 3185.9 to 4465.5 A/min.

e. The intervals for x* = 2.5 will be narrower than those above because 2.5 is closer to the mean than is 3.0.

f. No. A value of 6.0 is not in the range of observed x values, therefore predicting at that point is meaningless.

59.

a. $S_{xx} = 251{,}970 - \frac{(1950)^2}{18} = 40{,}720$, $S_{yy} = 130.6074 - \frac{(47.92)^2}{18} = 3.033711$, and

$S_{xy} = 5530.92 - \frac{(1950)(47.92)}{18} = 339.586667$, so $r = \frac{339.586667}{\sqrt{40{,}720}\sqrt{3.033711}} = .9662$.

There is a very strong positive correlation between the two variables.

b. Because the association between the variables is positive, the specimen with the larger shear force will tend to have a larger percent dry fiber weight.

c. Changing the units of measurement on either (or both) variables will have no effect on the calculated value of r, because any change in units will affect both the numerator and denominator of r by exactly the same multiplicative constant.

d. $r^2 = (.966)^2 = .933$

e. $H_0 : \rho = 0$ vs $H_a : \rho > 0$. $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$ ; Reject $H_o$ at level .01 if $t \geq t_{.01,16} = 2.583$.

$t = \frac{.966\sqrt{16}}{\sqrt{1-.966^2}} = 14.94 \geq 2.583$, so $H_o$ should be rejected. The data indicates a positive linear relationship between the two variables.