HW10 - Solutions

1) Sincich, 11.27

The graph in **b** would have the smallest s^2 because the width of the data points is the smallest.

2) Sincich, 11.32

2nd edition:

a. From the printout, SSE = 20,608.39, s^2 = MSE = 420.58, and s = 20.51.

b. s = 20.51. We would expect approximately 95% of the observed values of y (2007 SAT Score) to fall within $2s$ or 2(20.51) = 41 units of their least squares predicted values.

3rd edition:

a. SSE = 6952.757, s2 = MSE = 141.893, and s = 11.912
b. s = 11.912. We would expect approximately 95% of the observed values of y (2011 Math SAT Score) to fall within 2s or 2(11.912) = 24 units of their least squares predicted values.

3) Ex. 11.58

a. From Exercise 10.24, $SS_{xy} = -787.51087$, $SS_{xx} = 6{,}906.6087$, $\sum y = 60.1$, $\sum y^2 = 262.271$, and $\hat{\beta}_1 = -0.114022801$.

$$SS_{yy} = \sum y^2 - \frac{\left(\sum y\right)^2}{n} = 262.271 - \frac{(60.1)^2}{23}$$
$$= 262.271 - 157.043913 = 105.227087$$

$$SSE = SS_{yy} - \hat{\beta}_1 SS_{xy} = 105.227087 - (-0.114022801)(-787.51087) = 15.43289179$$

$$s^2 = MSE = \frac{SSE}{n-2} = \frac{15.43289179}{23-2} = 0.734899609 \text{ and } s = \sqrt{0.734899609} = 0.8573$$

$$s_{\hat{\beta}_1} = \frac{\sqrt{MSE}}{\sqrt{SS_{xx}}} = \frac{\sqrt{0.734899609}}{\sqrt{6,906.6087}} = 0.010315$$

To determine if the mass of the spill tends to diminish linearly as time increases, we test:

H_0: $\beta_1 = 0$
H_a: $\beta_1 < 0$

The test statistic is $t = \dfrac{\hat{\beta}_1 - 0}{s_{\hat{\beta}_1}} = \dfrac{-0.114022801}{0.010315} = -11.05$

The rejection region requires $\alpha = .05$ in the lower tail of the t-distribution with df $= n - 2 = 23 - 2 = 21$. From Table V, Appendix B, $t_{.05} = 1.721$. The rejection region is $t < -1.721$.

Since the observed value of the test statistic falls in the rejection region ($t = -11.05 < -1.721$), H_0 is rejected. There is sufficient evidence to indicate the mass of the spill tends to diminish linearly as time increases at $\alpha = .05$.

b. For confidence coefficient .95, $\alpha = .05$ and $\alpha/2 = .05/2 = .025$. From Table V, Appendix B, with df $= n - 2 = 23 - 2 = 21$, $t_{.025} = 2.080$. The 95% confidence interval is:

$$\hat{\beta}_1 \pm t_{.025} s_{\hat{\beta}_1} \Rightarrow -0.1140 \pm 2.080(0.010315) \Rightarrow -0.1140 \pm 0.02146$$
$$\Rightarrow (-0.13546,\ -0.09254)$$

We are 95% confident that for each additional minute of elapsed time, the decrease in spill mass is between 0.13546 and 0.09254.

4) a. Yes; the points look like they lie roughy on a line.

b. beta0 = 4.567 has no interpretation (0 is outside the range of the data)

beta1 = 4.9446; For every 1% increase in the interest rate, we expect the mean number of defaults per 1000 loans to go up by 4.9946

c. s = 2.31404. Roughly 95% of the Y values are within 2s = 4.6 defaults per thousand loans of the mean value 4.57 + 4.95 X.

sY = sqrt(SST/(n-1)) = sqrt(170.00/8) = 4.61. Roughly 95% of the Y values are within 2 sY = 9.2 defaults per thousand loans of the mean (40.33 as computed by Stat -> Basic Statistics -> Display Descriptive Statistics).

We can see that there is about half as much spread around the regression line than there is around the mean value of Y (s/sY = 0.5).

d. H0 : beta1 = 0; Ha : beta1 != 0; test stat T = 4.9446 / 0.9940 = 4.97.

Reject H0 if |T| > t.005, with n - 2 = 7 degrees of freedom. t.005 = 3.499.

Since T > t.005, we reject H0. There is a statistically significant linear relationship between Default rate and Interest rate.

e. R2 = 78.0%

5) a. Minitab plots

b. The histogram does not look bell shaped but the normal probability plot is roughly a straight line. The latter is more reliable since the sample size is small. There does not seem to be a departure from normality.

c. No clear pattern. The x axes are “yhat” and “i”.

d. No large residuals (all are between -2 and 2).