Quiz 13 Solutions

(a) We have $t = \frac{-0.017}{0.002066} = -8.23$, which has the T-distribution with $n - 2 = 17$ degrees of freedom. Thus $P = 2 \cdot P(T > 8.23) = 2 \cdot \text{tcdf}(8.23, 100, 17) = 2.47 \cdot 10^{-7}$.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.8943
R Square	0.7997
Adjusted R Square	0.7879
Standard Error	0.2149
Observations	19

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	3.13492	3.13492	67.87675	2.44E-07
Residual	17	0.785153	0.046185		
Total	18	3.920074			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	12.449	0.154252	80.70822	2.05E-23	12.124	12.77489
Year since 1900	-0.017	0.002066	-8.23873	2.44E-07	-0.02138	-0.01266

(b) $\text{Winning time (sec)} = 12.449 - 0.017 \cdot \text{Years since 1900}$

(c) Slope is -0.017; this tells us that, on average, the winning time decreases 0.017 seconds per year.

(d) Intercept is 12.449; this tells us that if the winning time had been decreasing at the same rate throughout the 20th century, the winning time in 1900 would have been 12.449 seconds.

(e) $H_0$is that the slope is zero: $\beta_1 = 0$.
H_a is that the slope is not zero: $\beta_1 \neq 0$.

(f) Since the P-value is small, $2.44 \cdot 10^{-7}$, we reject the null hypothesis and conclude the winning time is changing with time. Since the estimated slope, -0.017, is negative (and both ends of the confidence interval are negative), we conclude that the winning time is decreasing with time.

Extra Quiz 6 Solutions

(a) Null hypothesis: There is no significant relationship; $\beta_1 = 0$.
Alternate hypothesis: There is a significant relationship; $\beta_1 \neq 0$.

(b) The *t*-value is

$$t = \frac{-0.5879 - 0}{0.03180} = -18.453.$$

(c) There are 10 data points in the scatter plot, so the *t*-statistic has 8 degrees of freedom.

(d) The *p*-value is approximately 0. This is a 2-sided test, so from a computer, we find

$$p = 2(3.83 \cdot 10^{-8}) = 7.7 \cdot 10^{-8}$$

(e) We reject the null hypothesis; there has been a significant change in farm population over time.

Quiz 2 2010 Answer

A scatter plot of drawn of the weights, in pounds, of a random sample of male college students against their heights, in inches.

(a) The regression line drawn through this scatter plot could have equation: (check one, no reasons needed)

_____ $\text{Weight} = -43.8 - 87.2 \cdot \text{Height}$

_____ $\text{Weight} = 43.8 - 87.2 \cdot \text{Height}$

_____ $\text{Weight} = -43.8 + 87.2 \cdot \text{Height}$

_____ $\text{Weight} = 43.8 + 87.2 \cdot \text{Height}$

_____ $\text{Weight} = -438 - 8.72 \cdot \text{Height}$

_____ $\text{Weight} = 438 - 8.72 \cdot \text{Height}$

__X___ $\text{Weight} = -438 + 8.72 \cdot \text{Height}$

_____ $\text{Weight} = 438 + 8.72 \cdot \text{Height}$

(b) Which is the explanatory variable?
Height

(c) What is the response variable?
Weight

(d) Interpret the slope of the line in terms of heights and weights. *Give units!*

On average, each additional pound corresponds to a 8.72 pound increases in weight.

(e) If $r^2 = 0.417$, what is the value of r?

$$r = \sqrt{0.417} = 0.646$$

(f) If heights are measured in centimeters instead of inches, does the equation of the regression line change? (check one; no reasons needed)

Yes__X___ No _____ Depends on the actual data_____

(g) If heights are measured in centimeters instead of inches, does the correlation coefficient change? (check one; no reasons needed)

Yes_____ No __X____ Depends on the actual data_____

(h) If the roles of explanatory and response variables are interchanged, does the correlation coefficient change? (check one; give a reason)

Yes_____ No ___X___ Depends on the actual data_____

Reason:

Writing s_x for the standard deviation of the x-values and s_y for the standard deviation of the y-values, the formula for the correlation coefficient is

$$r = \frac{1}{n-1}\sum_{i=1}^{i=n} \frac{(x_i - \bar{x})}{s_x} \cdot \frac{(y_i - \bar{y})}{s_y}.$$

Since the formula for r is symmetric in x and y, so if you interchange them, the value of r does not change. Intuitively, the closeness of the data to the line is not affected by which variable is on which axis.