Ap Stats Ch15.1 Interpreting Linear Regression Models

2) Below is the computer output for the appraised value (in thousands of dollars) and number of rooms for houses in East Meadow, New York. Answer the following questions.

Predictor	Coef	Stdev	t-ratio
Constant	74.80	19.04	3.93
rooms	19.718	2.631	7.49

s = 29.05 R-sq = 43.8% R-sq(adj) = 43.0%

Analysis of Variance

Source	DF	SS	MS
Regression	1	47398	47398
Error	72	60775	844
Total	73	108173	

LSRL: $\widehat{value} = 74.8 + 19.718\ rooms$

$r = .661$ $r^2 = .438$

$s = 29.05$ $s_b = 2.631$

$t = 7.49$ $df = 72$

p-value = 0.000

Interpret the slope.

Interpret the slope.

There is an approximate increase of $19,718 in appraised value for an increase in one room.

Interpret the coefficient of determination.

Approximately 43.8% of the variation in appraised value is explained by the LSRL.

Predict the appraised value of a house with 13 rooms. $330,900

3) The manager of a chain of package delivery stores would like to predict the weekly sales (in $1000) for individual stores based on the number of customers. Find or calculate the related values for a two-sided test about the slope β.

Predictor	Coef	Stdev	t-ratio
Constant	2.4230	0.4810	
Customer	0.0087293	0.0006397	

s = 0.5015 R-sq = 91.2% R-sq(adj) = 90.7%

Analysis of Variance

Source	DF	SS	MS	F	p
Regression	1	46.834	46.834	186.22	0.000
Error	18	4.527	0.251		
Total	19	51.360			

LSRL: $\widehat{sales} = 2.423 + .0087\ rooms$

$r = .9549$ $r^2 = .912$

$s = 0.5015$ $s_b = .0006397$

$t = 13.64$ $df = 18$

p-value = 0.000

90% confidence interval = (.00759, .00981)

4) A real estate agent would like to predict the selling price of a single-family house by predicting the price (in \$1000) based on the square footage (in 100 ft^2). Find or calculate the following answers.

LSRL: $\widehat{price} = 18.35 + 3.87\ footage$

$r = .805$ $r^2 = .648$

$s = 12.9965$ $s_b = .7936$

$t = 4.887$ $df = 13$

p-value = .0003

Predictor	Coef	Stdev	t-ratio
Constant	18.3538	14.8077	
Customer	3.8785	0.7936	

s = 12.9965 R-sq = 64.8% R-sq(adj) = 62.0%

Analysis of Variance

Source	DF	SS	MS	F
Regression	1	4034.4144	4034.4144	23.885
Error	13	2195.8215	168.9093	
Total	14	6230.2360		

Interpret the slope & correlation coefficient.

S: For an increase of 100 square feet, there is an approximate increase of \$3870 in the price of a house.
R: There is a strong, positive, linear association between square footage & price of a house.

. A regression model was run to analyze how the "Metabolic Rate" (y) is related to "Lean Body Mass" (x) using a random sample. Below is the Minitab regression output.

Regression Analysis: MetabolicRate versus LeanBodyMass

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	892500	892500	50.40	0.000
Error	17	301051	17709		
Total	18	1193551			

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	113	180	0.63	0.537	
LeanBodyMass	26.88	3.79	7.10	0.000	1.00

a. Write down the equation of the least squares regression line. (10)
b. How many observations were included in this analysis? (5)
c. What is the correlation coefficient "r" between the response and predictor variables? (10)
d. Find the standard error of the estimate (regression standard error) s. (5)
e. Provide the details of a t-test to assess whether "Lean Body Mass" is a significant predictor of "Metabolic Rate". State the hypotheses. Provide the t-statistic, its degrees of freedom, and the p-value (all directly from the output). State the conclusion in context at α=5%. (15)
f. Provide a 95% confidence interval for the population slope $\beta 1$. Clearly specify the degrees of freedom used and the t* used. (15)

2011

5. Windmills generate electricity by transferring energy from wind to a turbine. A study was conducted to examine the relationship between wind velocity in miles per hour (mph) and electricity production in amperes for one particular windmill. For the windmill, measurements were taken on twenty-five randomly selected days, and the computer output for the regression analysis for predicting electricity production based on wind velocity is given below. The regression model assumptions were checked and determined to be reasonable over the interval of wind speeds represented in the data, which were from 10 miles per hour to 40 miles per hour.

Predictor	Coef	SE Coef	T	P
Constant	0.137	0.126	1.09	0.289
Wind velocity	0.240	0.019	12.63	0.000

S = 0.237 R-Sq = 0.873 R-Sq (adj) = 0.868

(a) Use the computer output above to determine the equation of the least squares regression line. Identify all variables used in the equation.

(b) How much more electricity would the windmill be expected to produce on a day when the wind velocity is 25 mph than on a day when the wind velocity is 15 mph? Show how you arrived at your answer.

(c) What proportion of the variation in electricity production is explained by its linear relationship with wind velocity?

(d) Is there statistically convincing evidence that electricity production by the windmill is related to wind velocity? Explain.