AP Statistics   Name_____

Unit 08 – HW #3   Period_____

Regression Inference

1. *The European School Study Project on Alcohol and other Drugs,* published in 1995, investigated the use of marijuana and other drugs. Data from 11 countries are summarized in the following scatterplot and regression analysis. They show the association between the percentage of a country's ninth graders who report having smoked marijuana and who have used other drugs such as LSD, amphetamines, and cocaine.
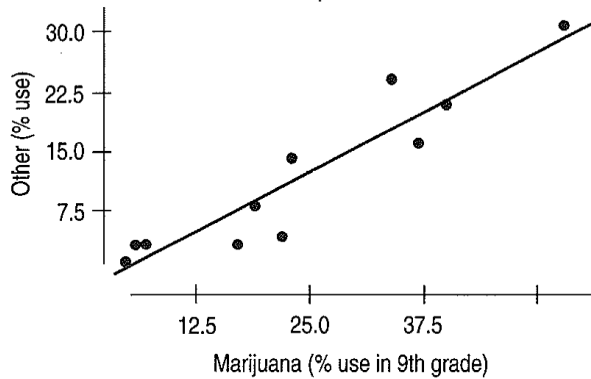
Dependent variable is: Other
R-squared = 87.3%
s = 3.853 with 11 − 2 = 9 degrees of freedom

Variable	Coefficient	SE(Coeff)	t-ratio	P-value
Intercept	−3.06780	2.204.	−1.39	0.1974
Marijuana	0.615003	0.0784	7.85	<0.0001

a. Explain in context what the regression says.

b. State the hypothesis about the slope (both numerically and in words) that describes how use of marijuana is associated with other drugs.

c. Assuming that the assumptions for inference are satisfied, perform a hypothesis test and state your conclusion in context.

d. Explain what R-squared means in context.

e. Do these results indicate that marijuana use leads to the use of harder drugs? Explain.

2. Does a person's cholesterol level tend to change with age? Data collected from 1406 adults aged 45-62 produced the regression analysis shown. Assuming that the data satisfy the conditions for inference, examine the association between age and cholesterol level.
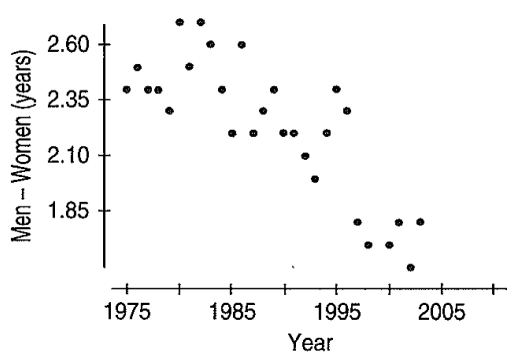
Dependent variable is: Chol

$s = 46.16$

Variable	Coefficient	SE(Coeff)	t-ratio	P-value
Intercept	194.232	13.55	14.3	≤0.0001
Age	0.771639	0.2574	3.00	0.0056

a. State the appropriate hypothesis for the slope.

b. Test your hypothesis and state your conclusion in the proper context.

3. The scatterplot suggests a decrease in the difference in ages at first marriage for men and women since 1975. We want to examine the regression to see if this decrease is significant.

Dependent variable is: Men − Women
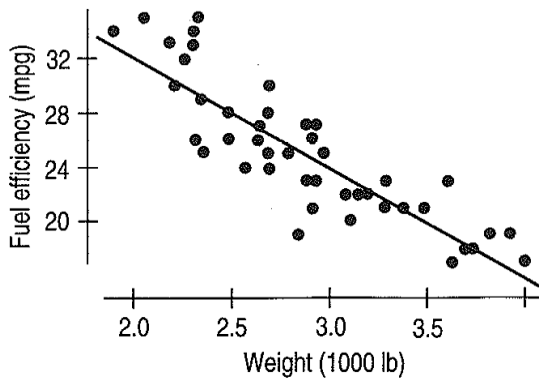R squared = 65.6%
s = 0.1869 with 28 − 2 = 26 degrees of freedom

Variable	Coefficient	SE(Coeff)	t-ratio	P-value
Intercept	61.8067	8.468	7.30	≤0.0001
Year	−0.02996	0.0043	−7.04	≤0.0001

a. Write appropriate hypotheses.

b. Test the hypothesis and state your conclusion about the trend in age at first marriage. Assume conditions have been met.

c. Based on the analysis of marriage ages since 1975, give a 95% confidence interval for the rate at which the age gap is closing. Explain what your confidence interval means.

4. A consumer organization has reported test data for 50 car models. We will examine the association between the weight of the car (in thousands of pounds) and the fuel efficiency (in miles per gallon).

Variable	Count	Mean	StdDev
MPG	50	25.0200	4.83394
wt/1000	50	2.88780	0.511656

Dependent variable is: MPG
R-squared = 75.6%
$s = 2.413$ with $50 - 2 = 48$ df

Variable	Coefficient	SE(Coeff)	t-ratio	P-value
Intercept	48.7393	1.976	24.7	≤ 0.0001
Weight	−8.21362	0.6738	−12.2	≤ 0.0001

a. Is there strong evidence of an association between the weight of a car and its gas mileage? Write the appropriate hypothesis.

b. Test your hypothesis and state your conclusion. Assume conditions have been met.

c. Create a 95% confidence interval for the slope of the regression line.

d. Explain in this context what your confidence interval means.