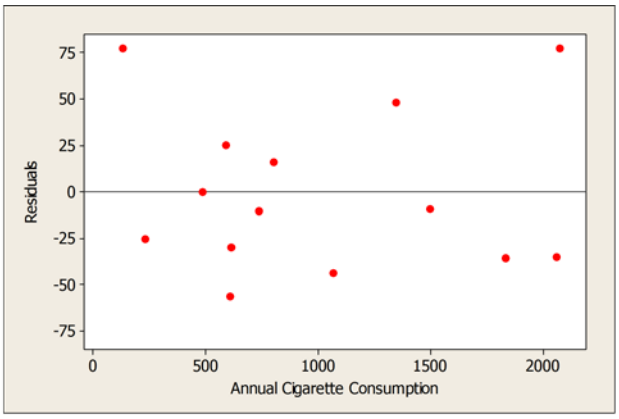
Quiz 12.1A AP Statistics Name:

Is there a link between the amount of cigarette smoking in countries and death rates from coronary heart disease (CHD)? Below is computer output from a regression analysis of this relationship for 14 randomly-selected countries from around the world, along with a residual plot. The explanatory variable is annual consumption of cigarettes per person and the response variable is annual deaths from coronary heart disease per 100,000 people.

```
Predictor      Coef  SE Coef     T      P
Constant      65.44    22.82  2.87  0.014
Cigarettes  0.02268  0.01926  1.18  0.262

S = 45.2672  R-Sq = 10.4%  R-Sq(adj) = 2.9%
```



1. What is the equation of the least-squares regression line based on these data? Define any variables used.

2. Interpret the slope of the regression line.

3. If we are trying to determine the relationship between these two variables throughout the world, is the slope you provided in part (a) a statistic or a parameter? Explain.

4. Based on the information given, discuss whether the conditions have been met to use *t*-procedures to make inferences about the slope of the regression line. If you do not have enough information to determine if a condition is met, indicate what other information or analysis is required.

5. Assuming all conditions have been met, construct and interpret a 90% confidence interval for the slope of the least squares regression of annual CHD deaths on annual cigarette consumption.

6. If you were to perform a test of the hypotheses $H_0 : \beta = 0$ *versus* $H_a : \beta \neq 0$ at the $\alpha = 0.10$ level, what would you conclude? Justify your answer using your result in Question 5.

Chapter 12 Solutions

Quiz 12.1A

1. $\hat{y} = 65.44 + 0.02268x$, where $\hat{y} =$ predicted deaths per year per 100,000 people from CHD and x = cigarettes consumed per person per year. **2.** A one-cigarette increase in the annual number of cigarettes consumed per peson in a country is associated with a predicted increase of 0.02268, on average, in annual deaths per 100,000 people from CHD. **3.** This is a statistic: it is an estimate of the population regression line's slope based on this particular random sample of 14 countries. **4.** **L**inear: The residual plot shows a random scattering of points around the residual = 0 line, suggesting that the liner model is a good fit for the data. **I**ndependent: observations for one country should be independent of observations for any other country. The population is finite, but there are more than $10 \times 14 = 140$ countries in the world. **N**ormal: to determine whether values of y for a given x are approximately Normally distributed, we need to determine if the residuals are roughly Normally distributed, using either a histogram or Normal probability plot. This information has not been provided. **E**qual standard deviation: the residual plot shows approximately the same spread around residual = 0 for all values of x. **R**andom: the data came from a random sample of 14 countries. **5.** State: We want to estimate β, the true slope of the population regression line relating annual cigarette consumption per person to annual deaths from CHD per 100,000 people, with 90% confidence. Plan: We are told to assume all conditions for inference have been met, so we will use a t-interval for the slope to estimate β. Do: We use the t distribution with 14 – 2 = 12 degrees of freedom to find the critical value. For a 90% confidence level, the critical value is $t^* = 1.782$, so the 90% confidence interval for β is $0.02268 \pm 1.782(0.01926) \approx 0.0227 \pm 0.0343$ or $(-0.0116, 0.0570)$. Conclude: We are 90% confident that the interval from –0.0116 to 0.0570 captures the actual slope of the population regression line relating annual deaths by CHD per 100,000 people to annual cigarette consumption per person in all countries. **6.** Since the 90% confidence interval contains the null value of 0, we cannot reject H_0 at the $\alpha = 0.10$ level. We do not have convincing evidence that the slope of the population regression line relating annual cigarette consumption per person to annual deaths from CHD per 100,000 people is different from 0.