Tutorial 12

Week of April 8, 2019

Question 12.2.16, Page 507

An article gave a scatterplot, along with the least squares line, where x represents rainfall volume (m^3) and y represents runoff volume (m^3) for a particular location. The accompanying values were read from the plot.

x	5	12	14	17	23	30	40	47	55	67
	72	81	96	112	127					
y	4	10	13	15	15	25	27	46	38	46
	53	70	82	99	100					

$$\sum x = 798 \quad \sum y = 643 \quad \sum x_i^2 = 63040 \quad \sum y_i^2 = 41999 \quad \sum x_i y_i = 51232$$

(a) Calculate point estimates of the slope and intercept of the population regression line.

(b) Calculate a point estimate of the true average runoff volume when rainfall volume is 50.

(c) Calculate a point estimate of the standard deviation σ.

(d) What proportion of the observed variation in runoff volume can be attributed to the simple linear regression relationship between runoff and rainfall?

(e) In `R`: Does a scatterplot of the data support the use of the simple linear regression model? Repeat parts (a)–(d) in `R`.

Height/Weight of Women

Consider the data set `women` in `R`, but pretend that each observation represents a single person instead of an average. Using `R`:

(a) Compute the regression parameters using `weight` as the predictor and `height` as the response.

(b) Give a 95% confidence interval for the true slope of the line.

(c) Conduct a hypothesis test at the 5% significance level of whether the true slope differs from zero.

(d) What does the coefficient of determination measure? What is its value for this model?

(e) State the assumptions of linear regression. Create diagnostic plots to check these assumptions. Are the assumptions satisfied?