

LINEAR REGRESSION AND CORRELATION: PREDICTION (MODIFIED R. BLOOM)*

Roberta Bloom

Based on *Linear Regression and Correlation: Prediction*[†] by

Susan Dean

Barbara Illowsky, Ph.D.

This work is produced by OpenStax-CNX and licensed under the
Creative Commons Attribution License 3.0[‡]

Abstract

This module provides an overview of Linear Regression and Correlation: Prediction as a part of R. Bloom's custom Collaborative Statistics collection (col10617). It has been modified from the original module m17092 in the Collaborative Statistics collection (col10522) by Barbara Illowsky and Susan Dean.

Recall the third exam/final exam example.

We examined the scatterplot and showed that the correlation coefficient is significant. We found the equation of the best fit line for the final exam grade as a function of the grade on the third exam. We can now use the least squares regression line for prediction.

Suppose you want to estimate, or predict, the final exam score of statistics students who received 73 on the third exam. The exam scores (*x-values*) range from 65 to 75. **Since 73 is between the *x-values* 65 and 75**, substitute $x = 73$ into the equation. Then:

$$\hat{y} = -173.51 + 4.83(73) = 179.08 \quad (1)$$

We predict that statistic students who earn a grade of 73 on the third exam will earn a grade of 179.08 on the final exam, on average.

Remember: Do not use the regression equation for prediction outside the domain of observed x values in the data.

Example 1

Recall the third exam/final exam example.

Problem 1

What would you predict the final exam score to be for a student who scored a 66 on the third exam?

*Version 1.1: Dec 23, 2009 5:53 pm -0600

[†]<http://cnx.org/content/m17095/1.6/>

[‡]<http://creativecommons.org/licenses/by/3.0/>

Solution

145.27

Problem 2

(Solution on p. 3.)

What would you predict the final exam score to be for a student who scored a 78 on the third exam?

Solutions to Exercises in this Module

Solution to Example 1, Problem 2 (p. 2)

The x values in the data are between 65 and 75. 78 is outside of the domain of the observed x values in the data (independent variables), so you cannot reliably predict the final exam score for this student. (Even though it is possible to enter x into the equation and calculate a y value, you should not do so!)