Linear Regression: Test Scores

Dr. Patrick Toche

Textbook:

James H. Stock and Mark W. Watson, *Introduction to Econometrics*, 4th Edition, Pearson.

Other references:

Joshua D. Angrist and Jörn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, 1st Edition, Princeton University Press.
Jeffrey M. Wooldridge, *Introductory Econometrics: A Modern Approach*, 7th Edition, Cengage Learning.

The textbook comes with online resources and study guides. Other references will be given from time to time.

Test Scores

A researcher, using data on class size (CS) and average test scores from 100 third-grade classes, estimates the OLS regression:

$$\widehat{TestScore} = 520.4 - 5.82 \times CS, \quad R^2 = 0.08, \quad SER = 11.5$$

a. A classroom has 22 students. What is the regression's prediction for that classroom's average test score?

b. Last year a classroom had 19 students, and this year it has 23 students. What is the regression's prediction for the change in the classroom average test score?

c. The sample average class size across the 100 classrooms is 21.4. What is the sample average of the test scores across the 100 classrooms?

d. What is the sample standard deviation of test scores across the 100 classrooms?

Test Scores

A researcher, using data on class size (CS) and average test scores from 100 third-grade classes, estimates the OLS regression:

$$\widehat{TestScore} = 520.4 - 5.82 \times CS, \quad R^2 = 0.08, \quad SER = 11.5$$

a. A classroom has 22 students. What is the regression's prediction for that classroom's average test score?
b. Last year a classroom had 19 students, and this year it has 23 students. What is the regression's prediction for the change in the classroom average test score?
c. The sample average class size across the 100 classrooms is 21.4. What is the sample average of the test scores across the 100 classrooms?
d. What is the sample standard deviation of test scores across the 100 classrooms?

Test Scores

A researcher, using data on class size (CS) and average test scores from 100 third-grade classes, estimates the OLS regression:

$$\widehat{TestScore} = 520.4 - 5.82 \times CS, \quad R^2 = 0.08, \quad SER = 11.5$$

a. A classroom has 22 students. What is the regression's prediction for that classroom's average test score?
b. Last year a classroom had 19 students, and this year it has 23 students. What is the regression's prediction for the change in the classroom average test score?
c. The sample average class size across the 100 classrooms is 21.4. What is the sample average of the test scores across the 100 classrooms?
d. What is the sample standard deviation of test scores across the 100 classrooms?

Test Scores

A researcher, using data on class size (CS) and average test scores from 100 third-grade classes, estimates the OLS regression:

$$\widehat{TestScore} = 520.4 - 5.82 \times CS, \quad R^2 = 0.08, \quad SER = 11.5$$

a. A classroom has 22 students. What is the regression's prediction for that classroom's average test score?
b. Last year a classroom had 19 students, and this year it has 23 students. What is the regression's prediction for the change in the classroom average test score?
c. The sample average class size across the 100 classrooms is 21.4. What is the sample average of the test scores across the 100 classrooms?
d. What is the sample standard deviation of test scores across the 100 classrooms?

Test Scores

$$\widehat{TestScore} = 520.4 - 5.82 \times CS, \quad R^2 = 0.08, \quad SER = 11.5$$

a. A classroom has 22 students. What is the regression's prediction for that classroom's average test score?

   The regression's prediction is:

$$\begin{aligned} \widehat{TestScore}_{|CS=22} &= 520.4 - 5.82 \times 22 \\ &= 392.36 \end{aligned}$$

Test Scores

$$\widehat{TestScore} = 520.4 - 5.82 \times CS, \quad R^2 = 0.08, \quad SER = 11.5$$

a. A classroom has 22 students. What is the regression's prediction for that classroom's average test score?

The regression's prediction is:

$$\begin{aligned}\widehat{TestScore}_{|CS=22} &= 520.4 - 5.82 \times 22 \\ &= 392.36\end{aligned}$$

Test Scores

$$\widehat{TestScore} = 520.4 - 5.82 \times CS, \quad R^2 = 0.08, \quad SER = 11.5$$

b. Last year a classroom had 19 students, and this year it has 23 students. What is the regression's prediction for the change in the classroom average test score?

The regression's prediction is:

$$\begin{aligned} \Delta\widehat{TestScore}_{|CS=19\to 23} &= (520.4 - 5.82 \times 23) - (520.4 - 5.82 \times 19) \\ &= -5.82 \times (23 - 19) \\ &= -23.28 \end{aligned}$$

The classroom size was increased by 4 extra students, so we expect the average test score to fall by about 23 points.

Test Scores

$$\widehat{TestScore} = 520.4 - 5.82 \times CS, \quad R^2 = 0.08, \quad SER = 11.5$$

b. Last year a classroom had 19 students, and this year it has 23 students. What is the regression's prediction for the change in the classroom average test score?

The regression's prediction is:

$$\begin{aligned} \Delta\widehat{TestScore}_{|CS=19\to 23} &= (520.4 - 5.82 \times 23) - (520.4 - 5.82 \times 19) \\ &= -5.82 \times (23 - 19) \\ &= -23.28 \end{aligned}$$

The classroom size was increased by 4 extra students, so we expect the average test score to fall by about 23 points.

Test Scores

$$\widehat{TestScore} = 520.4 - 5.82 \times CS, \quad R^2 = 0.08, \quad SER = 11.5$$

c. The sample average class size across the 100 classrooms is 21.4. What is the sample average of the test scores across the 100 classrooms?

The sample average across the 100 classrooms is:

$$\begin{aligned}\overline{TestScore} &= \hat{\beta}_0 + \hat{\beta}_1 \times \overline{CS} \\ &= 520.4 - 5.82 \times 21.4 \\ &= 395.852\end{aligned}$$

Test Scores

$$\widehat{TestScore} = 520.4 - 5.82 \times CS, \quad R^2 = 0.08, \quad SER = 11.5$$

c. The sample average class size across the 100 classrooms is 21.4. What is the sample average of the test scores across the 100 classrooms?

The sample average across the 100 classrooms is:

$$\begin{aligned} \overline{TestScore} &= \hat{\beta}_0 + \hat{\beta}_1 \times \overline{CS} \\ &= 520.4 - 5.82 \times 21.4 \\ &= 395.852 \end{aligned}$$

Test Scores

$$\widehat{TestScore} = 520.4 - 5.82 \times CS, \quad R^2 = 0.08, \quad SER = 11.5$$

d. What is the sample standard deviation of test scores across the 100 classrooms?

The sample variance may be computed as the ratio of the total sum of squares to the residual sum of squares, TSS/RSS. The residual sum of squares, RSS, may be computed from the standard error of the regression, SER:

$$SER = \sqrt{\frac{\sum_{i=1}^{n} \hat{u}_i^2}{n-2}} = \sqrt{\frac{RSS}{n-2}} \implies RSS = (n-2)SER^2 = 98 \cdot 11.5^2 = 12,961$$

The total sum of squares, TSS, may be computed from the formula for R^2:

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} \implies TSS = \frac{RSS}{1-R^2} = \frac{12,961}{1-0.08^2} = 13,044$$

And finally the variance of test scores:

$$s_{YY} = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2 = \frac{TSS}{n-1} = \frac{13,044}{99} = 131.8 \implies s_Y \approx 11.5$$

Test Scores

$$\widehat{TestScore} = 520.4 - 5.82 \times CS, \quad R^2 = 0.08, \quad SER = 11.5$$

d. What is the sample standard deviation of test scores across the 100 classrooms?

The sample variance may be computed as the ratio of the total sum of squares to the residual sum of squares, TSS/RSS. The residual sum of squares, RSS, may be computed from the standard error of the regression, SER:

$$SER = \sqrt{\frac{\sum_{i=1}^{n} \hat{u}_i^2}{n-2}} = \sqrt{\frac{RSS}{n-2}} \implies RSS = (n-2)SER^2 = 98 \cdot 11.5^2 = 12,961$$

The total sum of squares, TSS, may be computed from the formula for R^2:

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} \implies TSS = \frac{RSS}{1-R^2} = \frac{12,961}{1-0.08^2} = 13,044$$

And finally the variance of test scores:

$$s_{YY} = \tfrac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2 = \frac{TSS}{n-1} = \frac{13,044}{99} = 131.8 \implies s_Y \approx 11.5$$