

Module 11: Statistical inference for simple linear regression

(The attached PDF file has better formatting.)

Homework assignment: Estimating regression parameters

Some final exam problems give a set of points and ask to compute ordinary least squares estimators, sums of squares, t values, confidence intervals, and other regression statistics. The final exam may give one set of points and ask about several statistics and estimates, or separate points for each statistic.

This homework assignment continues the exercise from Module 8.

An actuary fits a two-variable regression model ($Y_i = \alpha + \beta \times X_i + \varepsilon_i$) to the relation between the incurred loss ratio (x) and the retrospective ratio (y):

Policy Year	(x)	(y)	($x - \bar{x}$)	($x - \bar{x}$) ²	($y - \bar{y}$)	($y - \bar{y}$) ²	($x - \bar{x}$)($y - \bar{y}$)
20X1	66.00%	22.50%	0.00%	0.00%	0.60%	0.0036%	0.0000%
20X2	67.00%	19.50%	1.00%	0.01%	-2.40%	0.0576%	-0.024%
20X3	68.00%	21.00%	2.00%	0.04%	-0.90%	0.0081%	-0.018%
20X4	65.00%	22.50%	-1.00%	0.01%	0.60%	0.0036%	-0.006%
20X5	64.00%	24.00%	-2.00%	0.04%	2.10%	0.0441%	-0.042%
Average	66.00%	21.90%	0.00%	0.02%	0.00%	0.0234%	-0.0180%

- ~ The column captions in the table use lower case x and y for the variables; the deviations are shown explicitly as $(x - \bar{x})$ and $(y - \bar{y})$.
- ~ The last line has averages, not totals. Some formulas in the textbook use totals.

We regress retrospective premium ratios on reported loss ratios to estimate premium assets for retrospectively rated business. Actuarial issues of retrospectively rated business are not important; the homework deals with the regression analysis only.

The Module 8 homework assignment solves for the least squares estimators of α and β .

- A. What is the *variance* of the ordinary least squares estimator of β ?
- B. What is the t statistic for testing the null hypothesis that $\beta = 0$?
- C. What is the 95% confidence interval for the true value of β ?
- D. What is the p value for testing the null hypothesis that $\beta = 0$?
- E. What is the *variance* of the ordinary least squares estimator of α ?
- F. What is the t statistic for testing the null hypothesis that $\alpha = 21.90\%$?
- G. What is the 95% confidence interval for the true value of α ?

H. What is the p value for testing the null hypothesis that $\alpha = 21.90\%$?

Show the computations for the homework assignment, not just the solution. You can check your solutions with Excel or other statistical software.

The null hypotheses for α and β depend on the scenario.

- A null hypothesis of $\beta = 0$ means that retrospective rating has no effect on the premium.
- A null hypothesis of $\beta = 1$ means that a dollar of loss causes a dollar of premium.

The null hypothesis for the value of α depends on the situation.

- A null hypothesis of $\alpha = 0$ means that retrospective rating is the standard premium.
 - This hypothesis is not realistic, since large insureds get premium discounts.
- A null hypothesis of $\alpha = 21.90\%$ means retrospective rating has no effect on premium.
 - The average observed discount is the expected discount regardless of losses.

These retrospective rating issues are not part of the regression analysis course.

The confidence intervals require a t distribution for the appropriate degrees of freedom. Use Excel to find the t values for a 95% confidence interval.