

# Textbook Exercises: Chapters 14 and 16

Due on the 24th of April, 2025.

The problems below are not your textbook original, but are my adaptations. Your textbook focuses more on showing you the relevance of statistics to the real world, and on using software to solve problems. In contrast, I focus more on hand calculation because that is what you will need to do in traditional examinations. And it helps to demystify statistical software: You are not just entering data into some black box and believing it. By knowing how it works, you can have more confidence in the results that it returns.

As always, show your calculations. Keep two or more decimal places for your answers.

articles	salary	evaluation	articles	salary	evaluation
0	41.5	5.19	22	88.4	7.35
12	82.1	6.77	2	58.4	5.81
9	60.0	6.34	16	83.0	7.18
6	57.1	5.76	18	77.3	7.38
3	54.2	6.46	4	50.9	5.74
12	61.5	6.23	4	57.0	6.61
16	63.2	6.38	4	45.5	7.40
0	63.2	6.45	1	51.1	5.83
2	49.2	7.20	8	69.8	6.85
31	75.7	8.16	9	60.3	6.63
3	65.6	6.49	5	55.3	5.23
9	63.9	6.28	14	61.8	5.40
3	57.0	5.39	15	50.5	6.52
8	63.2	8.25	34	80.1	7.39
21	82.9	6.53	19	71.1	7.77
42	103.3	8.46	5	59.6	5.06
29	83.7	9.48	1	59.2	4.31
4	59.5	7.01	0	52.0	5.83
11	65.7	6.80	23	87.2	6.48
7	52.3	6.20	14	65.3	5.50
20	70.4	6.01	2	49.8	5.13
5	58.1	6.24	9	58.0	6.14
22	75.0	7.99	14	74.2	6.88
1	49.9	5.58	9	65.7	5.02
28	85.0	6.93	4	60.0	6.05

1. A student organisation surveyed 50 economics professors. They recorded the professors' recent teaching evaluation (on a 10-point scale), their normal salary (in thousand US dollars), and the total number of their published research articles.
  - (a) Split the sample into two subsamples by whether or not an economics professor has authored more than 10 published research articles. Answer the following questions for each subsample.
    - i. Calculate the mean of Salary. (2 marks)
    - ii. Calculate the mean of Evaluation. (2 marks)
    - iii. Calculate the standard deviation of Salary. (4 marks)
    - iv. Calculate the standard deviation of Evaluation. (4 marks)
    - v. Calculate Pearson's coefficient of correlation between the two. (8 marks)
    - vi. Assume a linear model. Regress Evaluation on Salary using the ordinary least-square method. What is the fitted equation? (8 marks)
    - vii. Draw a residual plot to see whether or not there is any obvious sign of residual heteroscedasticity or dependence on Salary. Briefly explain how you answer this by looking at the plot. (8 marks)
    - viii. Calculate the standard error of the regression. (6 marks)
    - ix. Assume that linear extrapolation is appropriate in the given case. Estimate with 95% confidence the teaching evaluation of such an economics professor who does not have any normal salaries. (8 marks)
    - x. Repeat Part ix., but assume the professor's normal salary to equal the mean of the subsample. (8 marks)
    - xi. Draw a scatter plot of the recorded data. Focus your plot on the following range: Salary between 40 and 110, and Evaluation between 4 and 10. Include the fitted line. Also, visualise the 95% confidence interval of an economics professor's teaching evaluation in the plot. (8 marks)
  - (b) Compare your findings in Part (a):
    - i. Which subsample has a higher proportion of unexplained variation? (1 mark)
    - ii. Which subsample has a greater standard error of the regression? (1 mark)
  - (c) Thus far, you have regressed Evaluation on Salary; you should find a positive relationship (be the relationship statistically significant or not). Suppose that a student argues that the data suggest the possibility that raising a professor's normal salary may help to improve the professor's teaching evaluation by motivating him to put more effort into delivering knowledge to students. Try coming up with two alternative explanations of the observed relationship. (7 marks)

2. Review Problem 2 in Homework 3. Suppose that a different manager attempted to replicate the operations manager's findings, so he conducted a balanced  $5 \times 3$  factorial experiment where the procedure was as before. However, with 95% confidence, he found an interaction between the method being used and the worker's skill level. The following table summarises the results. The first number in each of the 15 pairs is the average number of minutes that the task (i.e., assembling the specialised desk) took each worker, and the second number is the standard deviation of the time that each worker spent.

method	skilled	average	unskilled
A	(25.19, 2.59)	(24.38, 2.45)	(19.67, 2.61)
B	(22.31, 2.43)	(24.10, 2.13)	(21.02, 2.31)
C	(25.10, 1.85)	(22.76, 2.16)	(21.87, 2.29)
D	(28.02, 2.41)	(24.31, 2.61)	(20.69, 2.83)
E	(25.06, 2.33)	(25.49, 2.44)	(21.99, 2.16)

The number of participants in the experiment is not disclosed, for some reason. Find the range of the proportion of unexplained variation. (25 marks)