

Homework 1

Linear Regression

Spring 2019 Prof. Chase Meyer

Instructions and Information: Answer each question in a new document, and show your work where applicable. Type and staple your homework (hand calculations can be written out). Present exercises in numerical order, and label your answers clearly. Embed any figures or tables within the homework. Each question is worth one point, unless otherwise noted. There are 20 points possible.

1. You have a sample of 100 people. The mean of x in your sample is 150.00. The standard deviation of x in your sample is 50.00. What is the 95% confidence interval for the mean value of X in the population?

- (a) The confidence interval is from 100.00 to 200.00.
- (b) The confidence interval is from 149.02 to 150.98.
- (c) The confidence interval is from 132.75 to 167.25.
- (d) The confidence interval is from 140.10 to 159.90.

2. Which is the best informal interpretation of the confidence interval you constructed, where lb is your lower bound and ub is your upper bound?

- (a) We can be 95% confident that our sample mean, \bar{x} , is between lb and ub .
- (b) We can be 95% confident that our sample standard deviation, s , is equal to 50.
- (c) We can be 95% confident that the population standard deviation, σ , is equal to 5.
- (d) We can be 95% confident that the mean in the population, μ , is between lb and ub .

3. (2 points) I am conducting a test of a null hypothesis that the population mean of a variable, X , is zero. In my sample, the mean is 3.0, the standard deviation is 9.0, and the number of observations is 21. I want to conduct a two-sided hypothesis test at the 5% significance level. Can I reject the null hypothesis? Why or why not?

4. (3 points) The below table gives information on gender and drink of choice for 30 individuals. You want to know if these variables are associated. Calculate the chi-squared statistic and lambda by hand. Next, use software or a chi-squared table to get a p -value for the chi-squared statistic. If you use a chi-squared table, the degrees of freedom is $(\text{columns} - 1) \times (\text{rows} - 1)$. What do the chi-squared statistic and lambda tell you about the association between gender and drink of choice? (The associated data set is called “homework1_1” and is available on the website and Canvas in both Stata and Excel formats.)

		Gender		Total
		Female	Male	
Drink Choice	Wine	11	5	16
	Liquor	1	3	4
	Beer	2	8	10
Total		14	16	30

For questions 5-6, consider the following regression equation:

$$\hat{y}_i = 1.0 + 1.5x_i$$

5. How much is y expected to change when x increases by one unit?

- (a) 0.5 unit
- (b) 1.0 unit
- (c) 1.5 units
- (d) 2.5 units

6. What is the expected value of y when x equals zero?

- (a) 0.5
- (b) 1.0
- (c) 1.5
- (d) 2.5

7. The R^2 value in a bivariate regression equation is 0.358. What percentage of the variation in the dependent variable is not explained by the independent variable?

- (a) 35.8%
- (b) 23.4%
- (c) 64.2%
- (d) 40.6%

For questions 8-9, consider the following regression output, in which the dependent variable is monthly income in dollars, and *education* captures an individual's education level using a continuous scale. The number of observations is 1,567.

	<i>Coefficient</i>	<i>Standard Error</i>	<i>t-stat</i>	<i>p-value</i>
Education	193.00	92.60	2.08	0.037
Intercept	3277.00	2390.90	1.40	0.162

8. Which bracket below represents the 95% confidence interval for the additional expected income associated with each additional unit of education?

- (a) between \$11.50 and \$374.50
- (b) between \$-455.20 and \$841.20
- (c) between \$179.28 and \$206.72
- (d) between \$169.28 and \$216.72

9. Based on the p -value associated with the coefficient on the *education* variable, which of the following is the correct interpretation?

- (a) If the true association between a unit change in education and expected monthly income was an increase of \$193.00, the likelihood of observing an effect as large in size as we did would be 3.7%.
- (b) If the true association between a unit change in education and expected monthly income was an increase of \$0.00, the likelihood of observing an effect as large in size as we did would be 3.7%.
- (c) There is a 3.7% chance that expected income is associated with education, assuming the null effect of a unit change in education is \$193.00.
- (d) There is a 3.7% chance that expected income is associated with education, assuming the null effect of a unit change in education is \$0.00.

10. (8 points) The below table gives information on the crime rate (y) and the poverty rate (x) across ten precincts in Detroit. Using these data (a copy of this dataset can be found in .dta and excel form titled “homework1_2”):

- Without using statistical software (you may use a calculator), calculate the bivariate correlation between crime rate and poverty rate.

- Without using statistical software (you may use a calculator), calculate the slope and intercept coefficients for a bivariate regression of the crime rate (y) on the poverty rate (x).
- Given the regression results you calculated, estimate the expected crime rate in a new precinct that has a poverty rate of 30.
- The standard error of b (the slope coefficient) is 0.257. Using this information, conduct a two-sided hypothesis test of the null hypotheses that $\beta = 0$ at the 1% significance level.
- The following information has been obtained: The variance of the crime rate variable is 72.22. The variance of the residuals is 12.99. Using this information, compute the R^2 value for the regression.
- Check all of your by-hand calculations using statistical software. (You'll need to input the data yourself.) Paste the regression table produced by the software into your homework document.
- Finally, pretend like you are explaining your findings to a friend who knows nothing about statistics. In plain words, what can you conclude about the degree to which one can use poverty rates to predict crime rates in Detroit?

Precinct	Crime Rate (y)	Poverty Rate (x)
1	39	24
2	12	9
3	23	13
4	32	18
5	24	16
6	14	8
7	20	11
8	25	15
9	34	16
10	27	20