

## 11.4: Biometrics Lab #4

Name: \_\_\_\_\_

### Experiment 1

The following data were collected on Old Faithful geyser in Yellowstone Park. The **x**-variable is time between eruptions and the **y**-variable is length of eruptions.

X	Y
12.17	1.88
11.63	1.77
12.03	1.83
12.15	1.83
11.30	1.70
11.70	1.82
12.27	1.93
11.60	1.77
11.72	1.83
12.10	1.89
11.70	1.80
11.40	1.72
11.22	1.75
11.42	1.73
11.53	1.74
11.50	1.77
11.90	1.87
11.86	1.84

a) Determine if a relationship exists between the 2 variables using a scatterplot and the linear correlation coefficient. Select **Graph> Scatterplot**. Select the **Simple** plot and click OK. Enter the response variable (length of eruptions) in the **Y variables** box, and the predictor variable (time between eruptions) in the **X variables** box. Click OK. Describe the relationship that you see.

---



---



---



---

b) Calculate the linear correlation coefficient. **Statistics> Basic Stats> Correlation**. Enter the 2 variables in the **Variables** box and click OK.

$r =$  \_\_\_\_\_

What two pieces of information about the relationship between these two variables does the linear correlation coefficient tell you?

---

c) Find a least squares regression line treating “time between eruptions” as the predictor variable (x) and “length of eruptions” as the response variable (y). **Stat>Regression> General Regression.** Enter “length of eruptions” in the **Response** box. Enter “time between eruptions” in the **Model** box. Click on **Options** and make sure that 95% is selected for all confidence intervals. Click on **Graphs** and select the **Residual plot** “Residual versus fits.” Click **Results** and make sure the Regression equation, Coefficient table, Display confidence intervals, Summary of model, Analysis of Variance table, and prediction tables are checked. Click OK.

Write the regression equation \_\_\_\_\_

What is the value of R<sup>2</sup>? \_\_\_\_\_

What does this mean?

Examine the residual model. Do you see any problems?

What is the value of the regression standard error? \_\_\_\_\_

Write the confidence intervals for the y-intercept \_\_\_\_\_  
and slope \_\_\_\_\_

Use the output to test if the slope is significantly different from zero. Write the null and alternative hypotheses for this test.

H<sub>0</sub>: \_\_\_\_\_

H<sub>1</sub>: \_\_\_\_\_

Using the test statistic and p-value from the Minitab output to test this claim.

Test statistic \_\_\_\_\_ p-value \_\_\_\_\_

Conclusion:

d) Using the regression equation, what would be the length of the eruption if the time between eruptions is 11.42 min.?

## Experiment 2

The index of biotic integrity (IBI) is a measure of water quality in streams. The sample data given in the table below comes from the Piedmont forest region. The table gives the data for IBI and forested area in square kilometers. Let Forest Area be the predictor variable (x) and IBI be the response variable (y).



Create a scatterplot and describe the relationship between these variables. Compute the linear correlation coefficient.

r = \_\_\_\_\_

Create a regression model for this data set following the steps from the first example. Write the regression model.

Is there significant evidence to support the claim that IBI increases with Forest Area? Write the test statistic/p-value used for this slope test along with your answer.

The researcher wants to estimate the population mean IBI for streams that have an average forested area of 48 sq. km. Click **STAT>REGRESSION> GENERAL REGRESSION**. Making sure that IBI is in the Response box and Forest Area is in the Model box, click on **Prediction** and enter 48 in the **New observation for continuous predictors** box and check Confidence limits. Click OK. Write the 95% confidence interval for mean IBI for streams in an average forested area of 48 sq. km.

You are working with a stream in an area with 19 sq. km. of forested area. Your management plan includes an afforestation project that will increase the forested area to 23 sq. km. You need to predict what the specific IBI would be for this stream when the forested area is increased. Create a prediction interval to estimate this IBI if the forested area increased to 23 sq. km.

Click **STAT>REGRESSION>GENERAL REGRESSION**. Making sure that IBI is in the Response box and Forest Area is in the Model box, click on **Prediction** and enter 23 in the **New observation for continuous predictors** box and check **Prediction limits**. Click OK. Write the 95% prediction interval for the IBI for this stream when the forested area is increased to 23 sq. km.

Explain the difference between the confidence and prediction intervals you just computed.

This page titled [11.4: Biometrics Lab #4](#) is shared under a [CC BY-NC-SA 3.0](#) license and was authored, remixed, and/or curated by [Diane Kiernan \(OpenSUNY\)](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.