**Using Regression Models to Examine The Effect of Carbon Dioxide (CO2) on Fluid pH**

**Objectives:**

* Explain why carbon dioxide levels in the extracellular fluid affect the pH of blood.
* Measure pH changes as air is exhaled into water.
* Compare the rate of pH changes in someone at rest and the same person after exercising.
* Determine and graph the regression equations for linear and logarithmic models for the data.
* Interpret the slope and y-intercept of the linear equation.
* Use a regression model to make predictions, calculate residuals, and interpret residuals.

**Introduction:**

The cells of all living organisms need energy to live and do the work they are meant to do. For example, in the human body, a nerve cell needs energy to carry impulses, the cilia on cells lining the trachea need energy to sweep mucus and debris from your lungs, and heart muscle cells need energy to contract and pump blood. These are just a few examples.

The most common form of cellular energy is adenosine triphosphate, abbreviated as ATP. Large amounts of energy are held in the chemical bonds between the three phosphate groups in ATP, and this energy is released when those bonds are broken. In human cells, ATP is made constantly by several processes. The method used by most cells under normal conditions relies on organelles called mitochondria. Mitochondria make ATP using oxygen and sugar in a chemical process called cellular respiration (**Equation 1**). A waste product of mitochondrial ATP production is the gas carbon dioxide (CO2). Because all cells require a constant supply of ATP to function, they are also constantly producing carbon dioxide and releasing it into fluid outside the cells where it is picked up by blood.

C6H12O6 + 6O2 6CO2 + 6H2O

**Equation 1.** Chemical equation for Cellular Respiration

In the blood, carbon dioxide causes an acid to be formed, which lowers the pH of the blood. If excess carbon dioxide is produced by cells or if it is not removed from the body efficiently, too much CO2 builds up in the blood it can lower the blood pH excessively and result in a serious medical condition called acidosis.

In healthy individuals the blood levels of carbon dioxide are tightly regulated by buffers in the bloodstream, removal, or excretion by the urinary system in urine, and changes in breathing rates (**Equation 2**). Breathing is very important for the removal of CO2 and maintenance of healthy cells and tissues. During breathing, CO2 leaves the blood in the lungs and is exhaled, thus reducing its levels in the blood, and raising blood pH to become more alkaline (basic).

CO2 + H2O ⇌ H2CO3 ⇌ H+ + HCO3-

**Equation 2.** Chemical equation for the bicarbonate buffer system present in human blood. The respiratory system helps maintain blood pH by utilizing this buffer system.

In this experiment we will observe and measure the effect of CO2 levels on fluid pH by exhaling into water and measuring the pH of water under different conditions such as repeated breaths and exercise.

**Assessing Specific Knowledge (Biology) (Pre-lab)**

1. All muscles need \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ to contract and move the body.
2. carbon dioxide
3. Acid
4. Energy
5. Base
6. Carbon dioxide is a byproduct of \_\_\_\_\_\_\_\_\_\_\_\_\_\_ production.
7. Carbon dioxide
8. Acid
9. Energy
10. Base
11. Energy is produced in organelles called \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.
12. Mitochondria
13. Golgi apparatus
14. Nuclei
15. Rough Endoplasmic Reticulum
16. The body removes carbon dioxide from the blood by carrying it to the lungs where it is \_\_\_\_\_\_\_\_\_\_\_.
17. Inhaled
18. Converted to water
19. Kept in small spaces
20. Exhaled
21. The term used for the condition that occurs when blood pH is too low is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.
22. Hyperventilation
23. Hypoventilation
24. Acidemia
25. Alkalemia
26. The term used to describe a breathing rate that is too fast is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.
27. Hyperventilation
28. Hypoventilation
29. Acidemia
30. Alkalemia

**Statistical Analysis Background**

The purpose of an experiment is to investigate the relationship between two variables and to determine if one variable (**independent** **variable** or **explanatory variable**) causes a change in the other variable (**dependent variable** or **response variable**). In an experiment, the researcher manipulates values of the independent variable and measures the resulting changes in the dependent variable. Additional variables that can cloud a study are called **lurking variables**. These are variables that are not considered in an experiment but may also influence the dependent variable.

To find a mathematical model to describe the relationship between the independent and dependent variables, we use **regression**. The goal of regression is to use the independent variable, , to make predictions for the dependent variable, . Outputs from the regression model represent predicted values of the dependent variable, denoted ŷ. The simplest regression model is a linear model, meaning that the pattern of the data follows a straight line. If the data has a **linear model**, then the regression model will be of the form ŷ = mx + b. Where m is the slope of the line, b is the y-intercept of the line, and ŷ denotes the predicted value of y.

The **y-intercept** of a graph is where it crosses the y-axis, which occurs when x = 0. Therefore, the y-intercept can be interpreted as the value of the dependent variable, y, whenever the independent variable, x, is 0. The **slope** of a line determines how steep the line is and the direction of the line. The slope of a line can be calculated by the equation m = Δy/Δx [[1]](#footnote-1) where the line moves uphill from left to right if the slope is positive, and the line moves downhill from left to right if the slope is negative. So, the slope of a line can be interpreted as the expected increase/decrease in the dependent variable, y, if the independent variable, x, increases by one unit. For example, say that a regression model using the age of a car (x) to predict the price of the car (y) is given by . The y-intercept of 30,000 means that a brand-new car (0 years old) would cost about $30,000. The slope of -2,250 means that each year, the price of the car is expected to decrease by $2,250.

While a linear model is the simplest regression method, it is not always the best approach if the variables have a nonlinear relationship. If the two variables have a logarithmic relationship, then a **logarithmic curve** would be the best model for the data. Logarithmic regression examines the linear relationship between a transformation of the independent variable x, ln(x), and the dependent variable y. In this case, a regression model could be of the form ŷ = a lnx +b. Where “ln” is an abbreviation for the natural logarithm (or the logarithm with base e), and ŷ denotes the predicted value of y.

To understand logarithms, it may be helpful to think about the relationship between logarithms and exponentials. These are inverse functions (see Figure 1), so ax = y means that logay = x. For example, we can see that log5 125 = 3 because 53 = 125. In other words, the output of a logarithmic function is an exponent.  For the natural logarithm (abbreviated “ln”), the base is implied to be e ≈ 2.71828. So, if we write ln1 = x, we interpret this as ex = 1.

Using this definition, we can also see that the input of a logarithm is restricted. This means there are certain numbers for which we cannot take the logarithm. For example, consider (**Figure 1**). Start by rewriting this as , where we want to solve for . Converting this into exponential form implies that . We can now see that there is no number as an exponent that will result in 0 (recall that . We can also see graphically in Figure 1 that a logarithmic graph cannot have an x-value of 0. Therefore, is undefined, meaning that it cannot be computed. This will also hold for any base, meaning that is undefined for any base .

A picture containing text, line, parallel, plot

Description automatically generated

**Figure 1:** Example of exponential and logarithmic graphs with base 8.

To measure the strength of a regression model, we can use the values of R2 (coefficient of determination) and the residuals. The value of R2 (range 0 to 1) measures the variability in the dependent variable that can be explained by the independent variable, thus measuring the strength of a model. A value of R2 close to 1 (or 100%) is evidence of a good-fit regression model. For example, say that a regression model using the age of a car to predict the price of the car has R2 = 0.97. This means that 97% of the variability in prices can be explained by the age of the car, where the remaining 3% of variability is due to other unexplained factors.

The residual, also called the “error”, is the vertical distance from a data point to the regression line (**Figure 2**). The residual is calculated using the equation *residual* = y – ŷ. That is, the residual for a given value of x is the difference between the observed value of y (from the collected data) and the predicted value of y (from the model). If the observed data point lies above the line, the residual is positive, and the line underestimates the actual data value for y. If the observed data point lies below the line, the residual is negative, and the line overestimates the actual data value for y.

A picture containing line, diagram, plot, screenshot

Description automatically generated 

**Figure 2**. Example of a linear regression with a visualization of residuals (inset **A**).

Before making a prediction using a regression model, it is important to consider the data that was used to create the model. While the model may hold for a specific range of values, the trend may change once you move outside the scope of the model. So, for example, say that regression was used to create a model that uses the age of a car (x) to predict the price of the car (y). If the data used to construct the model had ages spanning from 0 to 30 years, then it may not be appropriate to use the model to predict the price of a car that is 100 years old. An age of 100 years old would fall outside the scope of the model, and the model may not hold anymore.

**Assessing Specific Knowledge (Statistics)**

Suppose that a sports analyst is interested in determining the relationship between the speed that a baseball bat is swung and the distance the baseball travels. She conducts an experiment where a baseball bat is swung at various speeds between 60mph and 80mph. She then measures the distance the ball travels in feet at each speed.

* 1. Determine the independent, dependent, and possible lurking variables in her experiment.
  2. Suppose that a linear regression model is given to be . Determine the slope and the y-intercept of this equation.
  3. Would it be appropriate to use this regression model to make a prediction for the distance a baseball travels when the bat is swung at 40mph? Why or why not?
  4. Suppose that the collected data shows that a bat swung at 68 mph resulted in the ball traveling 350 ft. Find the residual for this point on the linear model.

**Procedure 1: Measuring pH Changes Caused by Ventilation at Rest**

****

**Figure 2.** Materials needed to conduct experiment. (**A**) deionized or distilled water (**B**) Digital pH meter with probe (**C**) 200 mL beaker (**D**) drinking straw

1. Place 50 mL of deionized or distilled water in a 100 mL beaker. Alternatively, 100 mL of deionized or distilled water can be placed in a 200 mL beaker. The objective is to have your beakers only half full to reduce splash.
2. Measure the pH of the water using a pH meter and record this value in **Table 2** (page 12).

A picture containing text, indoor

Description automatically generated

**Figure 3.** Example of a digital pH meter reading for deionized or distilled water.

1. Measure your partner’s ventilation rate in breaths per minute while your partner is sitting down and at rest. You can do this by observing the rise and fall of your partner’s chest for 10 seconds and then multiplying that number times 6. Record this value above **Table 2**. It is important that your partner is resting and calm. This will be your baseline measurement for ventilation rate.
2. Place a drinking straw in the water and have your partner gently exhale for 1 second. You should only see one or two bubbles.

A picture containing indoor

Description automatically generated

**Figure 4.** Example of rinsing pH meter between measurements. **IT IS VERY IMPORTANT TO MAKE SURE TO RINSE THE METER WITH DISTILLED WATER AFTER EACH pH MEASUREMENT IS TAKEN!**

1. Measure the pH of the water using a pH meter and record the value in **Table 2**.
2. Have your partner gently exhale for one more second. Repeat step 5.
3. Repeat steps 4 and 5 eight more times until you have measured any pH changes that occurred in 10 one-second breaths and recorded each in **Table 2**.
4. Now, have your partner gently exhale into the water for 10 seconds straight. Measure the pH of the water using the pH meter and record this value in **Table 2**.

**Procedure 2: Measuring pH Changes Caused by Ventilation after Exercise**

1. Place 50 mL of fresh deionized or distilled water in a 100 ml beaker. It is important that you do NOT use the water from Procedure 1. Alternatively, 100 mL of fresh deionized or distilled water can be placed in a 200 ml beaker. The objective is to have your beakers only half full to reduce splash.
2. Measure the pH of the water using a pH meter and record this value in **Table 3** (on pg. 12).
3. Have the same person from Procedure 1 exercise for 5 minutes. Your lab partner should exercise enough to increase their heart rate and ventilation rate. Have your partner sit down and then quickly measure your partner’s ventilation rate in breaths per minute. You can do this by observing the rise and fall of your partner’s chest for 10 seconds and then multiplying that number times 6. Record the initial respiratory rate above **Table 3**.
4. Place a drinking straw in the water and have your partner gently exhale for 1 second. You should only see one or two bubbles.
5. Measure the pH of the water using a pH meter and record the value in **Table 3**.
6. Have your partner gently exhale for one more second. Repeat step 5.
7. Repeat steps 4 and 5 eight more times until you have measured any pH changes that occurred in 10 one-second breaths and record each in **Table 3**.
8. Now, have your partner gently exhale into the water for 10 seconds straight. Measure the pH of the water using the pH meter and record this value in **Table 3**.
9. Measure your partner’s respiratory rate again and note the final respiratory rate below **Table 3**.

**Critical Thinking Question:** In Procedure 2, what do you predict would happen to the respiratory rate and the pH change of the water if your partner rested 10 minutes after exercising before blowing into the water?

**Student notes and observations:**

As you are conducting the experiment, write down anything out of the ordinary that you encounter while performing the experiment in the box below. Also note anything that is done differently than as directed in the instructions or any other observations that you feel may affect your results.

Here are a few examples of possible observations:

“The person blowing through the straw told me that they have asthma.”

“I coughed while exhaling into the water during exhale 8.”

“My lab partner could only blow for 7 seconds during the last reading.”

**Observations**

|  |
| --- |
|  |

**Results**

**Table 2.** pH measurements at rest.

**\*Initial Respiratory (breathing) rate:** \_\_\_\_\_

|  |  |
| --- | --- |
| **Measurement** | **pH** |
| Distilled water |  |
| Exhale 1 |  |
| Exhale 2 |  |
| Exhale 3 |  |
| Exhale 4 |  |
| Exhale 5 |  |
| Exhale 6 |  |
| Exhale 7 |  |
| Exhale 8 |  |
| Exhale 9 |  |
| Exhale 10 |  |
| 10 second exhalation |  |

**Table 3.** pH measurements after exercise.

**\*Initial Respiratory (breathing) rate:** \_\_\_\_\_

|  |  |
| --- | --- |
| **Measurement** | **pH** |
| Distilled water |  |
| Exhale 1 |  |
| Exhale 2 |  |
| Exhale 3 |  |
| Exhale 4 |  |
| Exhale 5 |  |
| Exhale 6 |  |
| Exhale 7 |  |
| Exhale 8 |  |
| Exhale 9 |  |
| Exhale 10 |  |
| 10 second exhalation |  |

**\*Final Respiratory (breathing) rate:** \_\_\_\_\_\_

**Statistical Analysis Assessment[[2]](#footnote-2)**

1. What are the dependent and independent variables for Procedures 1 and 2? Were there any lurking variables in your experiment?
2. Find the equation for the linear model for your data set and graph the line. Copy and paste the graph here or print and attach to the laboratory when you submit it.
3. What does the y-intercept represent in the context of the problem? What does the slope tell you about how pH level changes over time?
4. What does the change in pH tell you about the CO2 content of the solution?
5. Try to create a logarithmic model to fit the data. Excel will give you an error when constructing this model. Use the exponential definition of a logarithm to explain why the data point (0,b), where b is the initial reading of distilled water, is invalid.
6. To correct the above issue, change the point (0,b) in your data to (0.001,b). Now create an equation for the logarithmic model for your data set and construct a graph. Copy and paste the graph here or print and attach to the laboratory when you submit it.
7. Of the two models you have constructed, which is a better fit for your data? Why?
8. Would it be appropriate to use these two models to predict the pH after blowing into the solution for 2 seconds? Why?
9. Using the linear model, predict what the pH will be after blowing into the solution for 2 seconds.
10. Find the residual on the linear model when time is 2 seconds. Does the model over- or underestimate the actual value?
11. Using the logarithmic model, predict what the pH will be after blowing into the solution for 2 seconds.
12. Find the residual on the logarithmic model when time is 2 seconds. Does the model over- or underestimate the actual value?
13. Compare and discuss the residuals for both the linear and logarithmic models when time is 2 seconds.
14. Would it be appropriate to make a prediction for the pH after blowing into the solution for 400 seconds? Why or why not?

**Critical Understanding Assessment**

1. Describe the pH changes that occurred with each exhalation into the water in Procedure 1 (at rest). Did pH increase, decrease, or stay the same?
2. Describe the pH changes that occurred with each exhalation into the water in Procedure 2 (after exercise). Did the pH increase, decrease, or stay the same?
3. Analyzing both tables, what do you infer about changes in CO2 levels in the water that occur with each breath you exhale?
4. Look at your tables for Procedure 1 and Procedure 2. Do you see any difference between your data in the two tables? If so, what measurements are different?
5. Looking at your tables, what happens to the respiratory rate after exercise?
6. Compare both tables. Is there any difference between your pH readings between the person at rest and after they exercise? If so, summarize the differences that you see.
7. Using what you know about how and why carbon dioxide is produced, explain why the breathing rate of someone who has exercised may be different than that of the same person at rest and how this change in breathing rate relates to carbon dioxide levels and blood pH.
8. During an anxiety attack, sometimes called a “panic attack,” a person can begin breathing faster than normal. This is called hyperventilation. When someone is hyperventilating, they exhale more carbon dioxide than normal because they are blowing out more carbon dioxide per minute than someone with a normal breathing rate.
9. Using what you know about the relationship between carbon dioxide levels and blood pH, do you think that blood pH in someone who is hyperventilating goes up (becomes more basic), goes down (becomes more acidic) or stays the same? Explain your answer.
10. If someone followed Procedure 1 and collected data from someone who had just finished hyperventilating and is now breathing normally, what changes do you predict in the first pH values that are recorded if you compare those values to the data in Table 1a? Would the pH values of water blown into by someone who had just finished hyperventilating be higher (more basic), lower (more basic) or stay the same when compared to the data you collected today? Explain your answer.
11. If you observed and recorded anything out of the ordinary or done differently under “Student Notes and Observations,” discuss how these changes may have affected your data.

**Appendix A.** Directions for constructing graphs using the Excel spreadsheet.[[3]](#footnote-3)

**Linear and Logarithmic Regression: Excel**

**Adding a Linear Trendline**

1. To create a graph of your linear model, highlight the time and pH columns.
2. Go to the Insert tab and select the scatterplot chart.
3. In the chart design tab, select Add Chart Element and select the linear trendline.
4. To display your linear model on the scatterplot, select Add Chart Element, select trendline, select more trendline options, and select Display Equation on chart.

**Adding Logarithmic Trendline**

1. Highlight both the “Time” column and the “pH of solution” column. Do this by clicking and dragging across both columns at the same time (see Appendix B for example).
2. Select “Insert” from the top ribbon in Excel.  In the “Charts” category, select the first scatterplot labeled “Scatter.”
3. Label the graph by changing the chart title to “pH vs. Time.”
4. Right-click on any data point in the graph.  Select “Add Trendline.”
5. From the Trendline Options menu, select “logarithmic.”  Scroll down and checkmark the last two boxes “Display Equation on chart” and “Display R-squared value on chart.”
6. Select and drag the equation so that it is visible on the graph.

**Appendix B.** An example data set.  The data was collected, and then converted into seconds based on the cumulative amount of time blowing into the solution.

|  |  |
| --- | --- |
| **Trial** | **pH** |
| Distilled water (Baseline) | 6.9 |
| Exhale 1 | 5.2 |
| Exhale 2 | 5 |
| Exhale 3 | 5 |
| Exhale 4 | 4.9 |
| Exhale 5 | 4.9 |
| Exhale 6 | 4.9 |
| Exhale 7 | 4.8 |
| Exhale 8 | 4.8 |
| Exhale 9 | 4.8 |
| Exhale 10 | 4.8 |
| 10 second exhalation | 4.6 |

|  |  |
| --- | --- |
| **Time (seconds)** | **pH** |
| 0 | 6.9 |
| 1 | 5.2 |
| 2 | 5 |
| 3 | 5 |
| 4 | 4.9 |
| 5 | 4.9 |
| 6 | 4.9 |
| 7 | 4.8 |
| 8 | 4.8 |
| 9 | 4.8 |
| 10 | 4.8 |
| 20 | 4.6 |

1. Δ is the Greek letter for Delta, which represents the change in a variable. [↑](#footnote-ref-1)
2. Many of the questions in this section will require you to refer to the Directions for constructing graphs in Excel included in Appendix A at the end of this laboratory exercise. [↑](#footnote-ref-2)
3. See the spreadsheet provided by your instructor. [↑](#footnote-ref-3)