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

**ENS 495 Fall 2017 Homework 10**

**Question 1:** Which **TWO** statements in the list below about **correlation analysis** are **TRUE** and can be used to complete the following sentence:

**Correlation analysis investigates/determine if . . .** (circle 2)

1. There is a significant relationship between 2 variables
2. The relationship between variables is positive or negative
3. The relationship between 3 or more variables is significant
4. The relationship between a binomial variable and a numeric variable is significant.
5. One variable is physically or biological causing change in another

**Question 2:** Regression analysis is **different** than correlation analysis for which of the following reasons (that is, what does regression analysis assume or do that correlation analysis does not or can not?)

1. It can determine if there is a positive or negative relationship between y and x.
2. It assumes no causal relationship between y and x.
3. It cannot be used for prediction
4. It assumes a directional relationship between variables, such as y is caused by x.

**Question 3:** What mathematical method is used to fit a regression line to data?

1. Maximum Entropy
2. Least Squares
3. Correlation analysis
4. Mean square error (MSE)
5. All of the above
6. None of the above

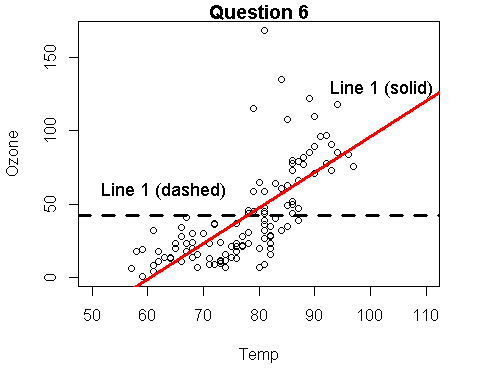
**Question 4:** When you fit a regression line to data, the **TWO** parameters that describe form of line through the scatterplot are: (circle 2)

1. Slope
2. p-value
3. F-statistic
4. Standard error (SE)
5. R2
6. Intercept

**Question 5:** When you run a regression, **uncertainty** about the true value of the slope of the line can be characterized using

1. p-value for the regression model
2. Standard error (SE) and/or confidence interval (CI)
3. F-statistics for the model
4. R2 value

**Question 8)** **TRUE / FALSE** (circle one) Regression, ANOVA and t-tests are all fundamentally **different** methods that require different statistical approaches. (1 point)



**Question 6** For the following questions consider the graph to the left with 2 regression lines.

**6a:** Which line in the graph represents the **NULL Hypothesis** or model (Ho)? (1 point)

1. Line 1 (Flat dashed line)
2. Line 2 (Angled solid line)

**6b:** Which line represents the **Alternative Hypothesis** or model (Ha)? (1 point)

1. Line 1 (Flat dashed line)
2. Line 2 (Angled solid line)

**6c)** Just looking at the graph above, which hypothesis do you think is most likely to be true? (1 point)

1. Null hypothesis (Ho)
2. Alternative hypothesis (Ha)

Why?

**6d)** What statement best characterizes Line 1

1. Positive slope (+)
2. Negative slope (-)
3. Slope of zero (0)
4. The term "slope" is not relevant

Why?

**6e)** What is the **"word equation"** that would be appropriate for a regression model for these data in the plot?

1. Ozone ~ Intercept + Slope\*Temp
2. Temp ~ Intercept + Ozone\*Temp
3. Ozone ~ Temp
4. y ~ mx + b

**In the geometry equation y = mx + b (circle)**

m = slope / intercept / p-value / F-statistics / R^2

b = slope / intercept / p-value / F-statistics / R^2