**#All code and outputs to be sent to** [**oliver.shipley@stonybrook.edu**](mailto:oliver.shipley@stonybrook.edu) **by end of class.**

**MAR504 –Workshop 2 – Probability Theory and Distributions**

**Part A**

1. How many different ways can the first three places be decided in a race with four runners? *Hint# Factorial*
2. How many different arrangements of the letters MISSISSIPPI? *Hint# Factorial*
3. Find the probability for a ***binomial*** experiment. There are 20 trials, the probability of a success is 0.01, and you want to find the probability of 1 success (P(x=1)).
4. For the same trial, find the probability for ***at most*** 6 successes.

**Part B**

1. Randomly sample 100 times from a normal distribution with a mean of 50 and a standard deviation of 10. Save the new data (this is called a vector) as an object.
2. Make a histogram of the resampled data and specify the use of 8 breaks on the x-axis.
3. Generate associated descriptive statistics. Does the resampled data match the original mean and SD specified in part 1?
4. Plot the probability density function of the vector and save this file.
5. Plot the cumulative distribution of the vector and save this file.
6. Describe the difference between the probability density function and cumulative distribution function.

**Part C**

1. Read in the penguin data file ‘penguin\_adapted.csv’.
2. Assess whether each of the four associated variables conform to a normal distribution. Report these results for each variable along with a histogram.
3. For variables that conform to a normal distribution, generated a probability density function and a cumulative distribution function. Save each respective plot.
4. Calculate densities at three specific points for two of these variables and report these values.