# Stat 359 Assignment 2

1.  Use R to take 10, 100, 1000 samples of size 10, 20, 50, 100 from 3 distributions.

    i) Uniform(a=0, b=1)

    ii) Poisson($\lambda = 5$)

    iii) Bernoulli(p=0.20) (or Binomial(n=1, p=0.20))

    (a) Investigate how the Central Limit Theorem works when sampling from all of the above distributions, and for each sample size. Plot a histogram illustrating the distributions of the sample mean for each sample size, number of samples, and distribution.

    (b) What do you notice as sample sizes increase?

    (c) What do you notice as the number of samples increases?

    (d) How does the distribution affect the outcome?

2.  A mixture of salt and sucrose was tasted to investigate how saltiness was judged depending on sucrose concentration and the data are contained in the file: *salt.txt*

    (a) Examine several graphical summaries to determine whether the data come from a symmetric distribution.

    (b) Estimate the skew ($\gamma_1$) and kurtosis ($\gamma_2$) of this distribution using the data.

    (c) Using the bootstrap construct a 95% confidence interval for the population skewness? Does it seem that the population distribution generating the data may be skewed?

    (d) Using the bootstrap construct a 95% confidence interval for the population kurtosis. Does it seem that the population distribution generating the data may have non-zero kurtosis?

    (e) Based on your analysis above, what do you conclude about the distribution?


3.  Data on the per diem fecundity (*fecundity.txt*) (number of eggs laid per female per day for the first 14 days of life) for 25 females on 2 genetic lines of the fruit fly Drosophila melanogaster are provided. Resistant (RS) to DDT were selectively bred and non-selected (NS) was the control. Do RS and NS lines differ in population variance? Do RS and NS lines differ in population mean fecundity?

4.    Fusible interlinings are being used with increasing frequency to support outer fabrics and improve the shape and drape of various pieces of clothing. The data on extensibility (100%) at 100 gm/cm for both high quality fabric (H) and poor-quality fabric (P) specimens is given in *fabric.txt*.

(a) Construct normal qq plots to verify the plausibility of both samples having been selected from normal population distributions.

(b) Construct a comparative boxplot. Does it suggest that there is a difference between true average extensibility for high-quality fabric specimens and that for poor-quality specimens?

(c) Decide whether true average extensibility differs for the two types of fabric.