Assignment – 4

1. The data in file Asg4Q1.csv is a random sample from an unknown distribution with expectation μ. Use the CLT to obtain a 95% confidence interval for μ.

2. The data in file Asg4Q2.csv is a random sample from an unknown distribution with median *m*. Obtain a 95% Bootstrap Pivotal confidence interval for *m*.

3. Consider the data is file Asg4Q2.csv.

   (a) Obtain a 95% Bootstrap Percentile confidence interval for *Skewness.*
   (b) For a symmetric distribution it is known that Skewness=0. Carry out a test of hypothesis H_0: Skewness = 0 against the alternative H_1: Skewness ≠ 0 at 5% level of significance.
   (c) What can you conclude regarding the possibility that the random sample is from a symmetric distribution?

   (Use the **skewness** command in the **moments** package for this exercise. You can use the following function. You may use the following R function appropriately.

```
sampskew<-function(d,i){
 t2<-d[i]
 return(skewness(t2))
}
)
```

4. The data in file Asg4Q3.csv is a random sample from an unknown distribution. Obtain a 95% confidence interval for $P(X \leq 10)$ where X is a random variable having the unknown distribution.
   (Hint: Let $Z_i = 1$ if $X_i \leq 10$ and = 0 otherwise. What is the distribution of Z_i? Are these i.i.d? Use the CLT to construct the confidence interval.)

5. The data in file Asg4Q4.csv is a random sample from an unknown distribution with standard deviation σ.
   (a) Obtain a 95% Bootstrap Pivotal confidence interval for σ.
   (b) Using the result in (a) test the null hypothesis that $\sigma = 0.2$ against that the alternative that it is not so at 5% level of significance.

6. The data in file Asg4Q5.csv consists of blood pressure readings of 15 randomly chosen individuals. The blood pressure of each individual is measured twice once by an untrained person using an automatic blood pressure measuring machine (*machine*) and another time by an expert (*expert*). The quantity of interest is the difference of the two measurements i.e. $D = |machine - expert|$.
   (a) Obtain a 95% Bootstrap Percentile confidence interval for μ_D, the population mean value of D.
   (b) Using (a) test the hypothesis $H_0: \mu_D = 10$ against the alternative $H_1: \mu_D \neq 10$

7. The data in file Asg4Q6.csv gives information about delays (in minutes) for 20 randomly chosen flights for 4 airlines (denoted as carrier A, B, C and D).
   (a) Obtain 95% Bootstrap Pivotal confidence intervals for the average delays of the four airlines separately.
   (b) Examine the four confidence intervals obtained above and give your comments on the findings.

8. The data in file Asg4Q7.csv gives information about number of dependents in 50 randomly chosen US families. Assume that the number of dependents in a family follow a Poisson distribution with unknown mean (λ).
   (a) Use CLT to obtain a 95% confidence interval for λ.
   (b) An investigator is interested in the estimating *P(No. of Dependents > 5)*. Provide an estimate for the same and then use the confidence interval for λ obtained in (a) to study the variation in the estimate.

9. The data in file Asg4Q8.csv gives information about rent(in $) for 45 randomly chosen similar type apartments in a city.
   (a) Use CLT to construct a 95% confidence interval for the average rent of apartments of the chosen type in the city.
   (b) Obtain 95% Bootstrap pivotal confidence interval for the average rent of apartments of the chosen type in the city.
   (c) Obtain 95% Bootstrap pivotal confidence interval for the standard deviation of rent of apartments of the chosen type in the city.
   (d) Assume that the rent (X) follows a normal distribution with mean μ and standard deviation σ. Obtain an estimate of $P(X \geq 300)$. Use the results obtained in parts (a)-(c) to study the variation in the estimate.