# Homework Assignment 4

DS4043, Spring 2023

**Due on April 28, 2023 at 11:59 pm**

1. Consider random variables $X_1, \ldots, X_n$ are i.i.d. $\mathrm{N}\left(\mu = 30, \sigma^2 = 100\right)$, given $n = 50$ and $\alpha = 0.05$.

   a) Obtain the **Monte Carlo estimate of the confidence level** for the 95% confidence interval includes the true value of $\mu$. Let the number of replicate as $m = 1000$. (Hint: you need to construct a 95% confidence interval of $\mu$; the statistic is the sample mean.)

   **Solution**:

   ```
   m = 1000; n = 50; mu <- 30
   set.seed(15)
   ```

   b) For the hypotheses, $H_0 : \mu = 30$ vs $H_1 : \mu \neq 30$, use Monte Carlo method to compute an empirical probability of type-I error, and compare it with the true value. Let the number of replicate as $m = 10000$.

   **Solution**:

   ```
   m = 10000; n = 50; mu <- 30
   set.seed(13)
   ```

2. Consider the random variables $X_1, \ldots, X_n$ are i.i.d. with a mixture normal density, i.e.

$$(1-p)N\left(\mu = 0, \sigma^2 = 1\right) + pN\left(\mu = 1, \sigma^2 = 9\right)$$

   We have $\alpha = 0.05, p = 0.4$ and $n = 50$. Let $\beta_1$ denote the skewness of random variable $X$ and its sample estimate is denoted by $b_1$. The hypotheses are $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$. Use the Monte Carlo method to estimate **empirical power** of the hypotheses. For finite samples one should use

$$\mathrm{Var}\left(b_1\right) = \frac{6(n-2)}{(n+1)(n+3)}.$$

   Let the number of replicate as $m = 10000$. To generate number from mixture density. Suppose $X_1 \sim N(0,1)$ and $X_2 \sim N(3,1)$ are independent. We can define a 50% normal mixture $X$, denoted $F_X(x) = 0.5F_{X_1}(x) + 0.5F_{X_2}(x)$. Unlike the convolution, the distribution of the mixture $X$ is distinctly non-normal; it is bimodal. To simulate the mixture:

   1. Generate an integer $k \in \{1, 2\}$, where $P(1) = P(2) = 0.5$.
   2. If $k = 1$ deliver random $x$ from $\mathrm{N}(0,1)$; if $k = 2$ deliver random $x$ from $\mathrm{N}(3,1)$.

   **Solution**:

   ```
   set.seed(19)
   ## function to calculate b1 skewness
   ```

3. Compute a jackknife estimate of the bias and the standard error of the correlation statistic in the *law* data example. Compare the result with the bootstrap method.

   **Solution**:

```r
library(bootstrap)
n <- nrow(law)
```

4. Refer to the air-conditioning data set *aircondit* provided in the *boot* package. The 12 observations are the times in hours between failures of airconditioning equipment:

$$3, 5, 7, 18, 43, 85, 91, 98, 100, 130, 230, 487.$$

Assume that the times between failures follow an exponential model $\mathrm{Exp}(\lambda)$. Obtain the MLE of the hazard rate $\lambda$ and use bootstrap to estimate the bias and standard error of the estimate. Let the number of replicates as $m = 200$.

**Solution**:

```r
library(boot)
set.seed(20)
m=200
```