Problem 1.

(a) Let $X \sim \text{Exp}(\lambda)$ be an exponential random variable with parameter $\lambda > 0$. Give a formula for its mean and its median, and compute both when $\lambda = 10$.

(b) Now suppose that $n$ samples are generated from a $\text{Exp}(\lambda)$ distribution. Give the log likelihood function, and using this, compute the maximum likelihood estimator of $\lambda$.

(c) Use `rexp` to generate 100,000 samples of an exponential random variable with parameter $\lambda = 10$. Plot the histogram with 1000 buckets, and compute the mean and the median of the sample.

(d) Fix $n = 10$, and create a sample of $n$ i.i.d. exponential random variables with parameter $\lambda = 10$. Using your sample, write code to compute $B$ bootstrapped estimates of the mean and median, and plot resulting histograms for each.

(e) Use the resulting bootstrap distribution to compute a 95% confidence interval for both the mean and the median; does it contain the true theoretical value?

(f) Now write a function `my-bootstrap` that takes as input $n$, $B$, and $\lambda$, and automates the preceding process: generate bootstrap samples, plot the histogram of the resulting bootstrap distribution, and return a 95% confidence interval for mean and median.

(g) Using the preceding function, plot the histograms, as well as the evolution of upper and lower bounds of the confidence intervals for the mean and median, for $n = 10, 12, 15, 20, 25, 40, 70, 100, 150, 300, 500, 1000, 2000, 5000, 10000, 20000$, and $B = 30, 100, 200, 400$ (we keep $\lambda = 10$ fixed). Display the middle point of the each interval, together with the true theoretical value.

(h) Comment on your observations from the preceding plots. Do the mean and median behave similarly? Based on what you see, can you use each of them to construct consistent estimators of $\lambda$? Based on what you see, can you suggest whether one of these is a more efficient estimator than the other?

Some tips:

- **You are welcome to use the** `boot` **library** ([http://www.statmethods.net/advstats/bootstrapping.html](http://www.statmethods.net/advstats/bootstrapping.html)) **to compute bootstrap estimates.** Start by typing

  ```r
  library(boot)
  ```
We use the function \texttt{boot(data=statistic=R)} to create \( R \) bootstrapped datasets with replacement out of the dataset passed to \texttt{boot}. Each new dataset has the same size as the original one. To use \texttt{boot}, you must pass in a function (this is the argument \texttt{statistic}) that returns the desired statistic; see the examples in Lecture 13.

- If you store the output of \texttt{boot} in, say, \texttt{boot.out}, you can pass the resulting object to \texttt{boot.ci} to get confidence intervals; type \texttt{?boot.ci} to get more information on how to get various types of confidence intervals (normal, percentile, etc.).

**Problem 2.** In this problem we explore another approach to parameter estimation, called the *method of moments*. The basic idea of the method of moments is explained in Section 9.2 of *All of Statistics*.

In this problem we apply the method of moments to estimation of the parameters of a beta distribution. The *beta distribution* \( \text{Beta}(\alpha, \beta) \) with left parameter \( \alpha > 0 \) and right parameter \( \beta > 0 \) has pdf:

\[
g(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}, \quad 0 < x < 1
\]

where \( B(\alpha, \beta) \) is a normalizing term. Specifically:

\[
B(\alpha, \beta) = \frac{(\alpha + \beta - 1)!}{(\alpha - 1)! (\beta - 1)!}.
\]

It can be shown that if \( X \sim \text{Beta}(\alpha, \beta) \), then:

\[
\mathbb{E}[X] = \frac{\alpha}{\alpha + \beta} = \mu_1(\alpha, \beta); \quad \mathbb{E}[X^2] = \frac{\alpha(\alpha + 1)}{(\alpha + \beta)(\alpha + \beta + 1)} = \mu_2(\alpha, \beta).
\]

Now suppose \( Y = (Y_1, Y_2, ..., Y_n) \) is a draw of \( n \) i.i.d. samples from the beta distribution with left and right parameters \( \alpha \) and \( \beta \) respectively. Let \( M_1^{(n)} = \frac{1}{n} \sum_{i=1}^{n} Y_i \) and \( M_2^{(n)} = \frac{1}{n} \sum_{i=1}^{n} Y_i^2 \). That is \( M_1 \) is the *sample mean* and \( M_2 \) is the *sample second moment*.

(a) Use the law of large numbers to give the limits as \( n \to \infty \) of \( M_1^{(n)} \) and \( M_2^{(n)} \).

(b) In the method of moments, we solve the following two equations to estimate \( \alpha \) and \( \beta \) from our sample:

\[
M_1^{(n)} = \mu_1(\alpha, \beta);
\]

\[
M_2^{(n)} = \mu_2(\alpha, \beta).
\]

Check that the following pair solve the preceding equations:

\[
\hat{\alpha} = \frac{M_1(M_1 - M_2)}{M_2 - M_1^2};
\]

\[
\hat{\beta} = \frac{(1 - M_1)(M_1 - M_2)}{M_2 - M_1^2}.
\]

These are the *method of moment estimators* for \( \alpha \) and \( \beta \).
(c) Now we use simulations to see the estimator in action. Use the function `rbeta(n, a, b)` to generate \( n = 100 \) data points using a beta distribution with both left and right parameters \( \alpha = \beta = 1 \). Find the estimates of the first and second moments using your random observations. Now using the above results and your estimates of the moments, find the method of moments estimators of the parameters of the beta distribution. How far are you away from the true parameters? Repeat the entire experiment for \( n = 1000 \) observations and \( n = 10000 \) observations. What do you observe?

**Notes:**

- The method of moments estimator is sometimes useful when other approaches to estimation are likely to be intractable (e.g., maximum likelihood estimation).

- There are fairly general conditions under which method of moments estimation is consistent.

- At the same time, note that in this case, the method of moments estimators are unique and we can obtain them in closed form; in general they might not even exist, they might not be unique, and there may be no closed form!

- In addition, note that in general method of moments estimation is not going to be asymptotically efficient. If you want to get more practice with the bootstrap, you can check this point by computing bootstrap estimates for the standard error, for progressively increasing sample sizes \( n \); and comparing with the same estimates for the standard error of the MLE. You will see that the MLE has smaller standard errors as \( n \) grows.