ST 501 R Project

Jimmy Hickey, Shaleni Kovach, Meredith Saunders, Stephanie Stewart

July 25, 2019

Contents

Part I - Convergence in Probability . . . **2**
1. . . . 2
    a. . . . 2
    b. . . . 3
    c. . . . 3
    d. & e. . . . 4

Part II - Convergence in Distribution . . . **6**
2. . . . 6
3. . . . 7
4. . . . 9
5. & 6. . . . 11

Part I - Convergence in Probability

1.

Consider the "double exponential" or Laplace Distribution. A RV $Y \sim Laplace(\mu, b)$ has the PDF given by

$$f_Y(y) = \frac{1}{2b} e^{-\left(\frac{|y-\mu|}{b}\right)}$$

for $-\infty < y < \infty$, $-\infty < \mu < \infty$, and $b > 0$.

We will consider having a random sample of Laplace RVs with $\mu = 0$ and $b = 5$. We'll look at the limiting behavior of $L = \frac{1}{n}\sum_{i=1}^{n} Y_i^2$ using simulation.

a.

Give a derivation of what L converges to in probability. You should show any moment calculations and state the theorem(s) you use.

By the Weak Law of Large Numbers, we know that,

$$L = \frac{1}{n}\sum_{i=1}^{n} Y_i^2 \xrightarrow{\text{p}} E(Y^2).$$

We can calculate $E(Y^2)$ using the definition of an expected value.

$$
\begin{aligned}
E(Y^2) &= \int_{-\infty}^{\infty} y^2 \cdot \frac{1}{2b} \cdot e^{-\left(\frac{|y-\mu|}{b}\right)} dy \\
&= \int_{-\infty}^{\infty} (x+\mu)^2 \cdot \frac{1}{2b} \cdot e^{-\frac{|x|}{b}} dx && \text{taking } x = y - \mu \\
&= \frac{1}{2b} \int_{-\infty}^{\infty} (x^2 + 2\mu x + \mu^2) \cdot e^{-\frac{|x|}{b}} dx \\
&= \frac{1}{2b} \Big[\int_{-\infty}^{\infty} x^2 \cdot e^{-\frac{|x|}{b}} dx + \int_{-\infty}^{\infty} 2\mu x \cdot e^{-\frac{|x|}{b}} dx + \int_{-\infty}^{\infty} \mu^2 e^{-\frac{|x|}{b}} dx \Big] \\
&= \frac{1}{2b} [4b^3 + 0 + 2b\mu^2] \\
&= \frac{4b^3}{2b} + \frac{2b\mu^2}{2b} \\
&= 2b^2 + \mu
\end{aligned}
$$

We can confirm this by checking $E(Y)^2 = Var(Y) + E(Y)^2$. From Wikipedia, we can see that $E(Y) = \mu$ and $Var(Y) = 2b^2$.

$$Var(Y) + E(Y)^2 = 2b^2 + (\mu)^2 = 2b^2 + \mu^2 = E(Y)^2$$

In the case of $\mu = 0,\ b = 5$, we get that $L \xrightarrow{\text{p}} 2 \cdot 5^2 + 0 = 50$.

b.

Explain what $K = \sqrt{L}$ converges to and why.

By the Continuity Theorem, we can see that $K = \sqrt{L} \xrightarrow{\text{p}} \sqrt{2b^2 + \mu^2}$. In the case of $\mu = 0,\ b = 5$, we get that $K \xrightarrow{\text{p}} \sqrt{50}$.

c.

Derive the CDF of Y . Note you'll have two cases and you should show your work.

Our CDF looks like

$$F_Y(y) = \int_{-\infty}^{y} \frac{1}{2b} e^{-\left(\frac{|x-\mu|}{b}\right)} dx$$

Using the absolute value, we can split the density function into two pieces, $y < \mu$ and $y \geq \mu$. Let us examine the first case.

$$\begin{aligned} F_Y(y) &= \int_{\infty}^{y} \frac{1}{2b} e^{-\frac{\mu-x}{b}} dx && \text{for } y < \mu \\ &= \int_{\infty}^{y} \frac{1}{2b} e^{\frac{x-\mu}{b}} dx \\ &= \frac{1}{2} e^{\frac{y-\mu}{b}} \end{aligned}$$

Next we can examine the $y \geq \mu$ case.

$$\begin{aligned} F_Y(y) &= \int_{\infty}^{y} \frac{1}{2b} e^{-\frac{x-\mu}{b}} dx && \text{for } y \geq \mu \\ &= \int_{\infty}^{\mu} \frac{1}{2b} e^{\frac{\mu-x}{b}} dx + \int_{\mu}^{y} \frac{1}{2b} e^{\frac{\mu-x}{b}} dx \\ &= \frac{1}{2} + \left(\frac{1}{2} - \frac{1}{2} e^{\frac{\mu-y}{b}}\right) \\ &= 1 - \frac{1}{2} e^{\frac{\mu-y}{b}} \end{aligned}$$

Putting the pieces together gives,

$$F_Y(y) = \begin{cases} \frac{1}{2} e^{\frac{y-\mu}{b}} & \text{for } y < \mu \\ 1 - \frac{1}{2} e^{\frac{\mu-y}{b}} & \text{for } y \geq \mu. \end{cases}$$

d. & e.

The code for parts d. and e. can be found in the `Problem_1.R` file. Here are the resulting graphs.

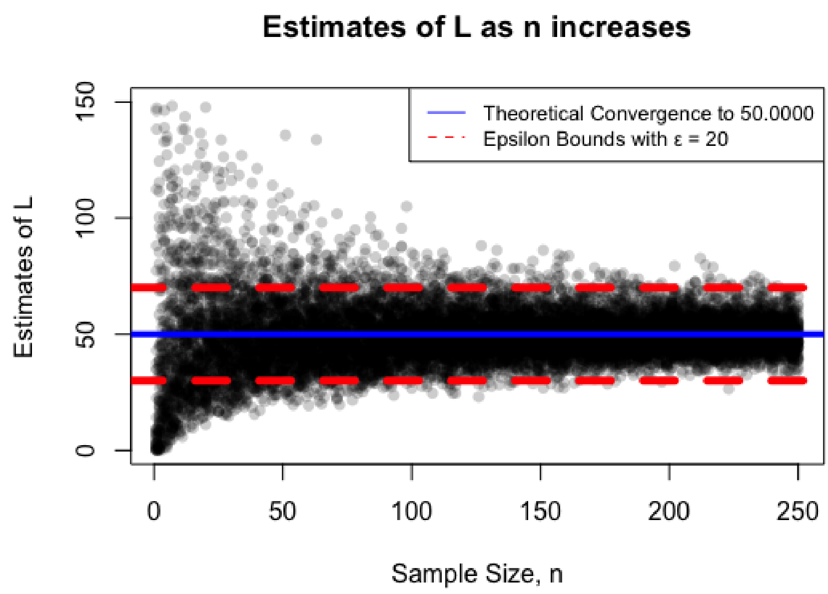

Figure 1: L as n increases

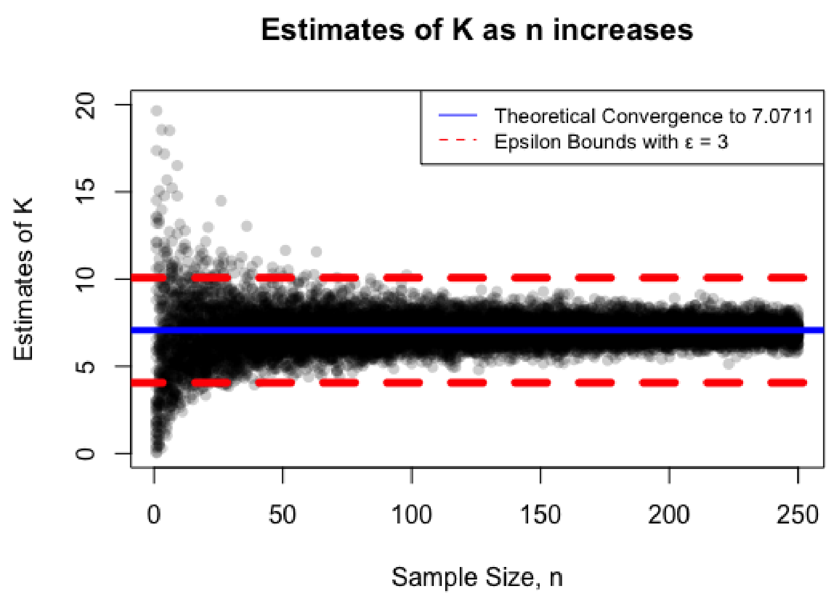

Figure 2: K as n increases

These graphs both demonstrate that our RVs L and K are converging in probability At each sample size n, the same number of samples (50) were taken. Notice the trend as the number of observations (n) in each sample increases. It is clear that there is far less spread. As n increases, the RVs are converging to the blue line. The observed values of L are geting closer and closer to 50. And the observed values of K are approaching $\sqrt{50}$. As n increases we could continue to shrink our epsilon bubble around the expected value and we will continue to see this convergence.

Part II - Convergence in Distribution

2.

This time we'll look at the limiting distribution of our statistics.

We can see from the plots that as n increases, the estimates of L do appear to converge to a normal distribution. Similarly, as the n increases, the estimates of K also appear to converge to a normal distribution, just with a smaller variance.

This behavior is most noticeable by plotting the graphs side by side and by putting them on the same axes.

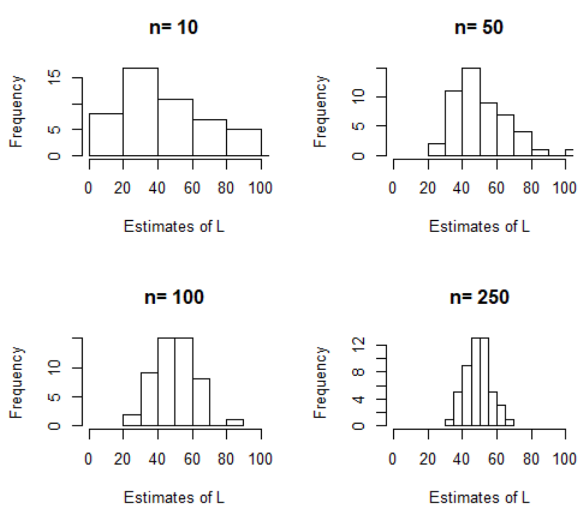

Figure 3: Estimates of L with the same axes

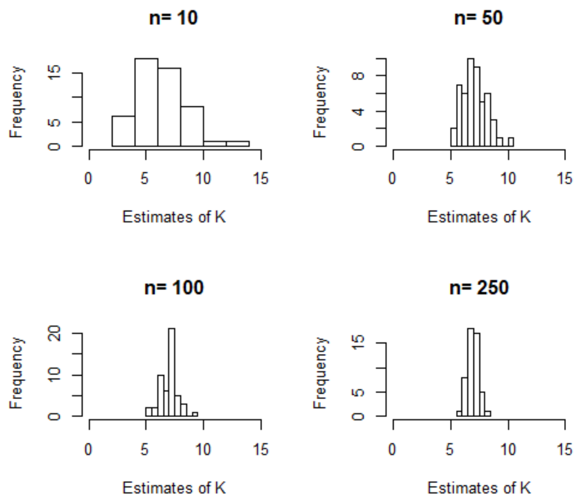

Figure 4: Estimates of K with the same axes

3.

Theoretically, since L is an average of *iid* RVs (Y_i^2 are each RVs) with finite variance we know that, properly standardized, L should have a standard normal limiting distribution by the CLT.

Derive the appropriate standardization that will converge to a standard normal distribution for $\mu = 0$ and an arbitrary b. Show your work. Note: the kurtosis of the Laplace distribution is 6 and we have $\mu = 0$. Find the formula for kurtosis (book or wikipedia) and the calculation of the fourth moment won't be too bad!

Recall that $L = \frac{1}{n}\sum_{i=1}^{n} Y_i^2$. Let's start by finding $E(L)$.

$$
\begin{aligned}
E(L) &= E\Big(\frac{1}{n}\sum_{i=1}^{n} Y_i^2\Big) \\
&= \frac{1}{n}\sum_{i=1}^{n} E(Y_i^2) && \text{since } E \text{ is a linear operator} \\
&= \frac{1}{n}\cdot nE(Y_i^2) && \text{by identically distributed RVS} \\
&= E(Y_i^2) \\
&= 2b^2 + \mu
\end{aligned}
$$

With this, we can calculate the variance.

$$
\begin{aligned}
Var(L) &= Var\Big(\frac{1}{n}\sum_{i=1}^{n} Y_i^2\Big) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} Var(Y_i^2) && \text{by properties of Variance} \\
&= \frac{1}{n^2}\cdot n \cdot Var(Y_i^2) && \text{by identically distributed RVs} \\
&= \frac{1}{n}Var(Y_i^2) \\
&= \frac{1}{n}\Big[E((Y_i^2)^2) - E(Y_i^2)^2\Big] && \text{definition of variance} \\
&= \frac{1}{n}\Big[E(Y_i^4) - E(Y_i^2)^2\Big]
\end{aligned}
$$

Note that the kurtosis is 6, so we can find $E(Y_i^4)$.

$$
\begin{aligned}
6 &= \frac{E(Y_i^4)}{E(Y_i^2)^2} && \text{defintion of kurtosis} \\
E(Y_i^4) &= 6 \cdot E(Y_i^2)^2
\end{aligned}
$$

$$\begin{aligned} Var(L) &= \frac{1}{n}\Big[E(Y_i^4) - E(Y_i^2)^2\Big] \\ &= \frac{1}{n}\Big[6 \cdot E(Y_i^2)^2 - E(Y_i^2)^2\Big] \\ &= \frac{1}{n}\Big[5 \cdot E(Y_i^2)^2\Big] \\ &= \frac{1}{n}\Big(5 \cdot (4b^4 - 4b^2 \cdot \mu + \mu^2)\Big) \end{aligned}$$

Given that $\mu = 0$, we get

$$Var(L) = \frac{1}{n}20b^4.$$

With those parameters,

$$\begin{aligned} &\frac{L}{n} \dot\sim N(2b^2 + 0, \frac{1}{n}20b^4) \\ &L \dot\sim N(n \cdot 2b^2, n \cdot 20b^4) \end{aligned}$$

4.

Redo your above 4 plots using the standardization.

The graphs were generated using the standardization of L, along with the value of b that was provided in part I. Each graph shows to have a similar shape that is approaching a normal distribution. Further increasing the size of n, would continue to show the distribution converging to a normal distribution.

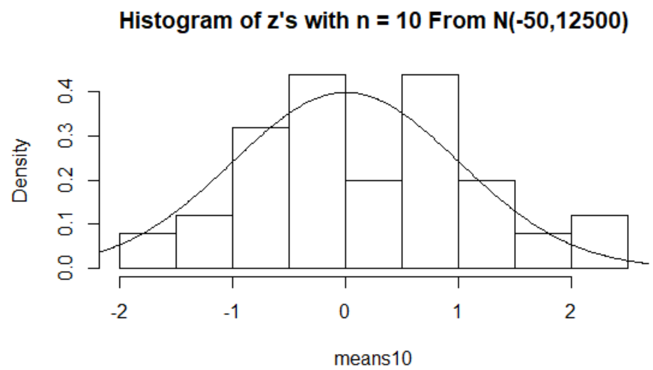

Figure 5: 50 samples at $n = 10$

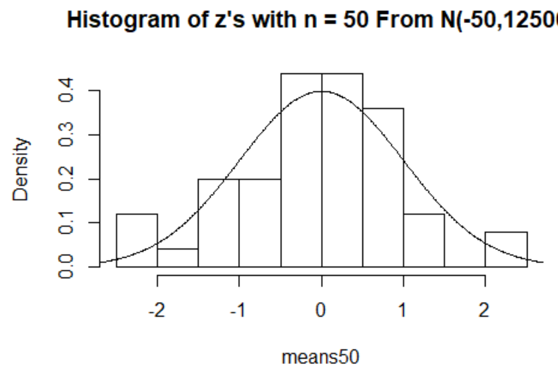

Figure 6: 50 samples at $n = 50$

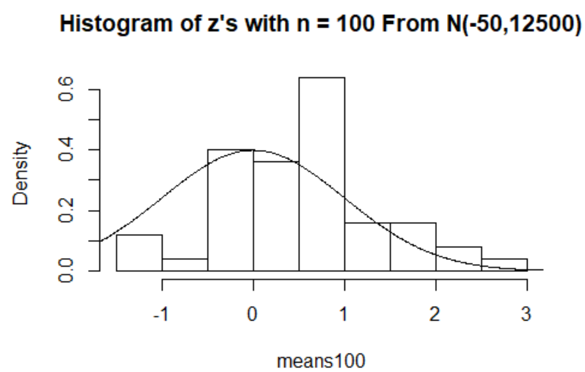

Figure 7: 50 samples at $n = 100$

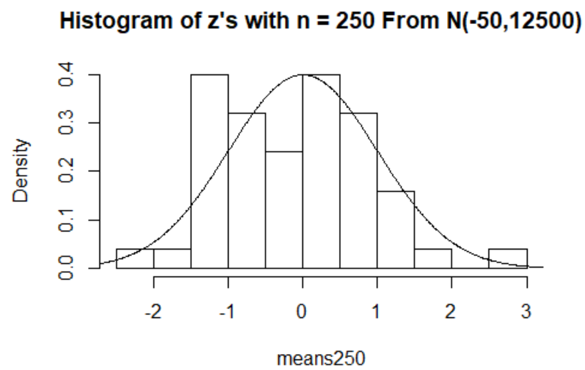

Figure 8: 50 samples at $n = 250$

5. & 6.

Now let's see if convergence is occurring for larger values of n. Generate data using the same method as above except do so for $n = 1000$ and $n = 10,000$. Use $N = 10,000$ data sets as well. Create similar plots for these two n values. In a comment discuss how the CLT is manifesting for this problem. Does n > 30 work?

Using $N = 50$, while the samples seem to be converging to normal, they still do not appear quite normal even for sample size $n = 10,000$. When we change to $N = 10,000$ both $n = 1,000$ and $n = 10,000$ appear normal. This shows it is not enough for $n > 30$ the number of sample repetitions also plays a role in CLT..

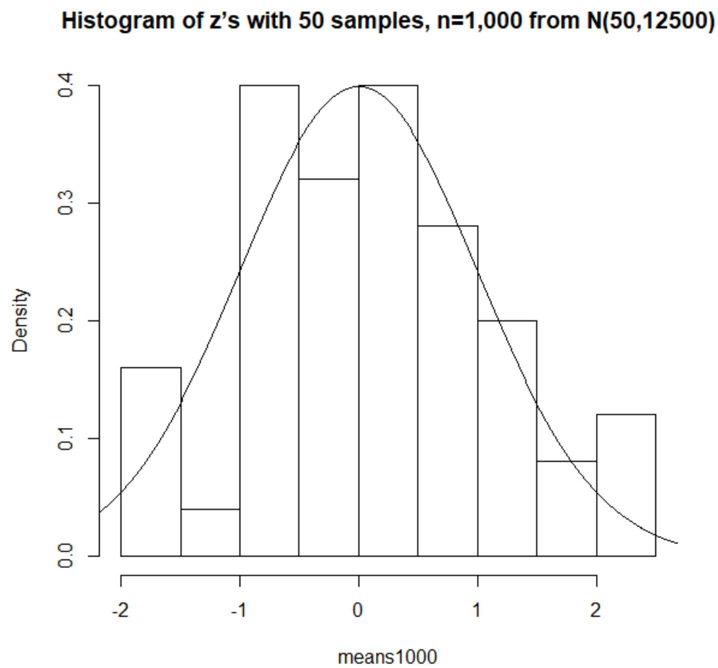

Figure 9: 50 samples at $n = 1,000$

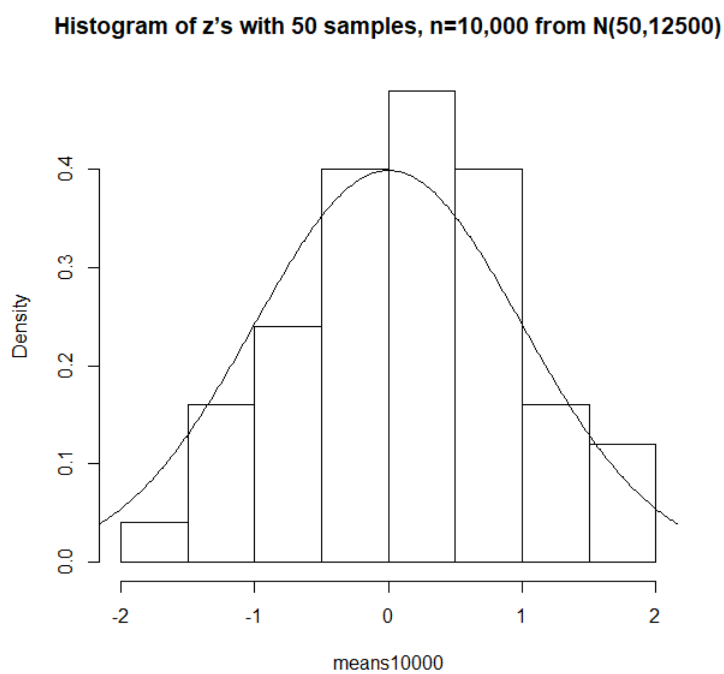

Figure 10: 50 samples at $n = 10,000$

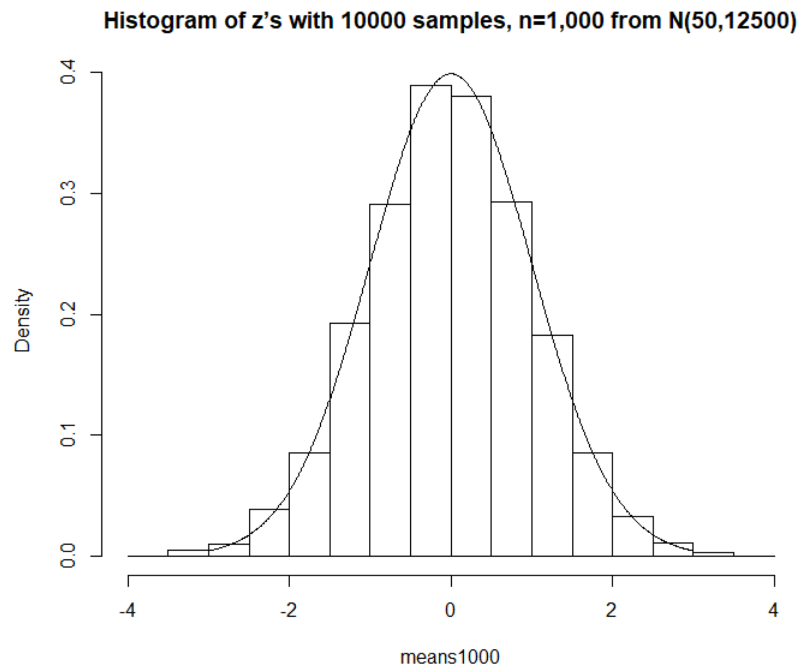

Figure 11: 10,000 samples at $n = 1,000$

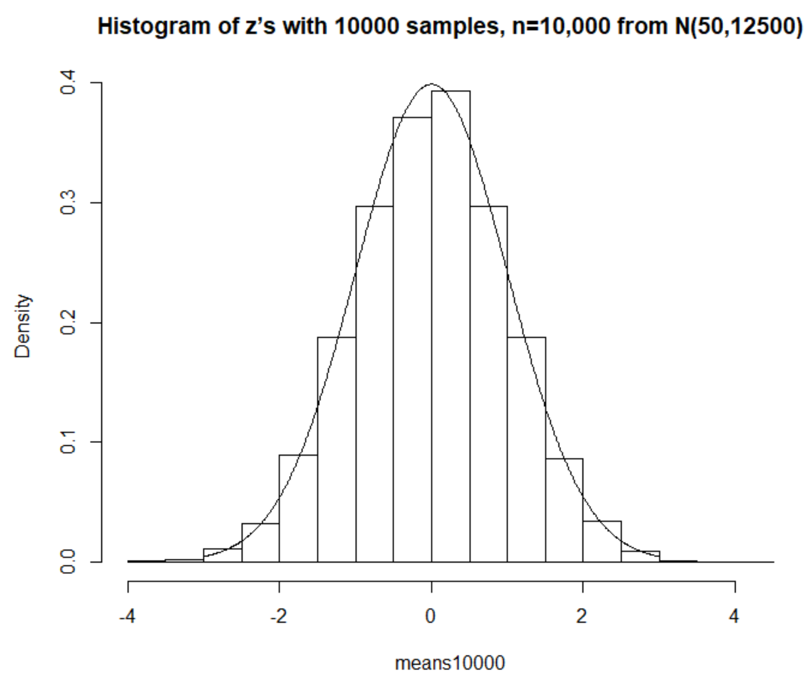

Figure 12: 10,000 samples at $n = 10,000$