WLLN, CLT and Simulation

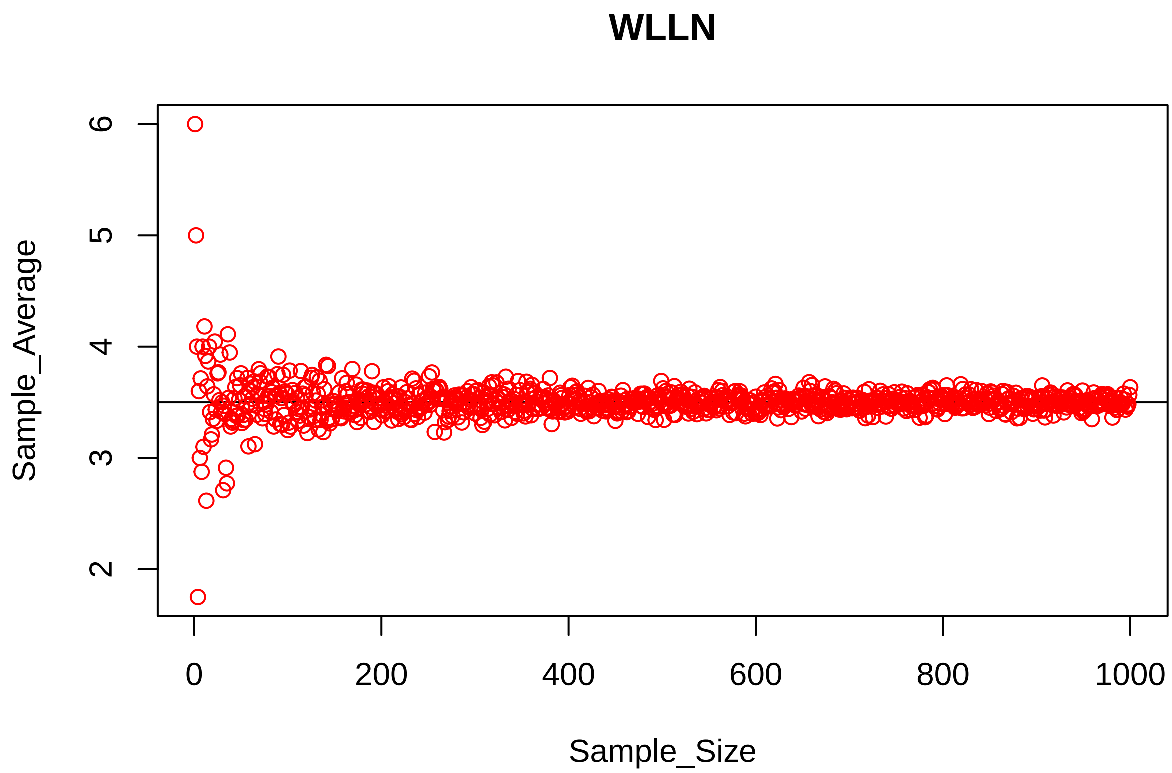
Kingsuk Jana

2023-03-25

Demonstration of Weak Law of Large Number (WLLN)

If we draw sample from a population with finite mean and finite variance, then the sample averages approach towards the population average as sample size grows larger. Here we take a population as discrete uniform distribution where random variable can take six values with probability 1/6 each.

```r
m=1000 # Number of iteration.
Sample_Average <- c() # Empty vector for sample averages.

Sample_Size <- c() # vector to store sample sizes.

for(i in 1:m){
  x = mean(sample(1:6,i,replace = TRUE))
  # Drawing i th sample from discrete uniform distribution.

  Sample_Average[i] = x
  Sample_Size[i] = i }

Sample_Average
Sample_Size
plot(Sample_Size, Sample_Average, main = "WLLN", abline(h = 3.5), col = "red")
```



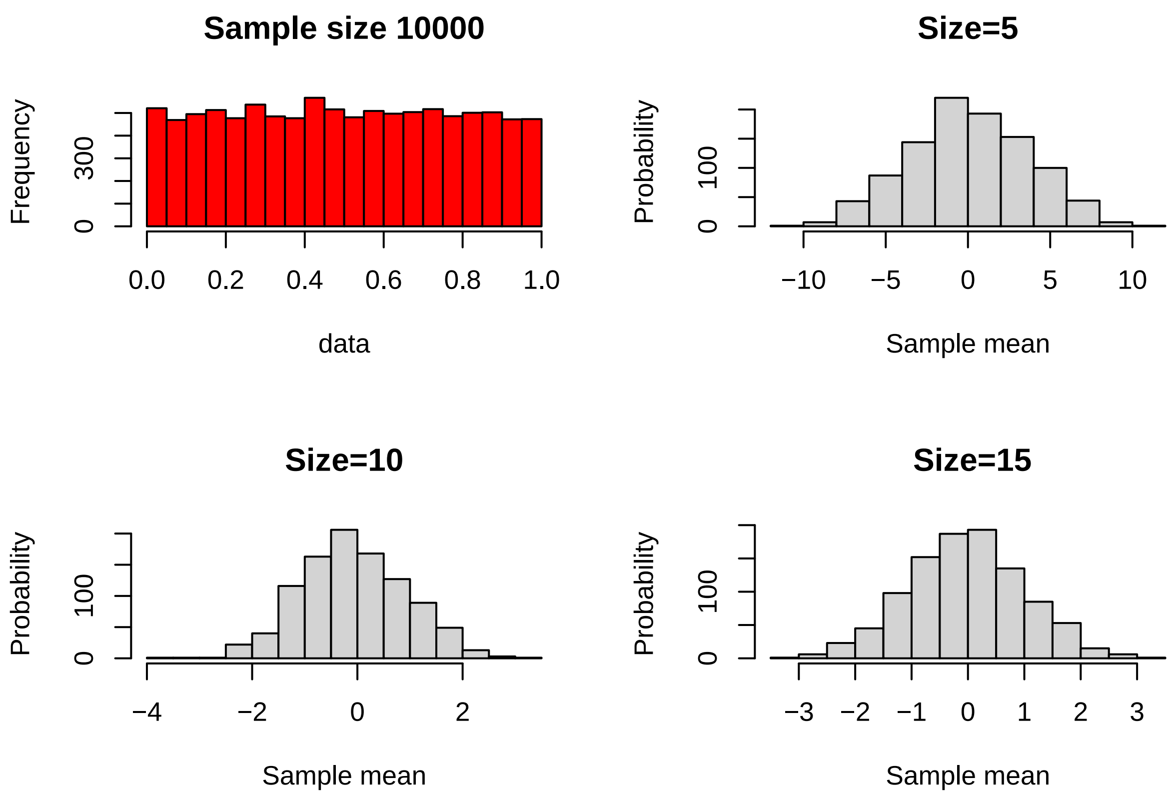
Here we simulate samples with sizes 2,..,1000 from discrete uniform distribution where random variable can take six values with probability 1/6 each. Corresponding to each samples we calculate the sample average. After that we plot the sample averages against respective sample sizes. As we can see from the figure that the sample averages are fluctuating around the population mean values 3.5 corresponding to the small sample sizes. When the sample size increases the sample average approaching towards the population average.

Demostration of Central Limit Theorem :

The central limit theorem states that the distribution of standardized sample mean tends to standard normal distribution as sample size gets larger.

```
set.seed(123)
par(mfrow = c(2,2))

data = runif(n = 10000, min = 0, max = 1)
# Draw sample from uniform distribution.

hist(data, col = "red", main = "Sample size 10000")
# We can visualize the distribution of  sample using the histogram.



# Simulate 1000 many samples each with size 5 from the population.

Sample_Mean_5 <- c() # Vector to store the sample means.
```

```r
n = 1000 # Number of iteration.

for(i in 1:n){

  Sample_Mean_5[i] = mean(sample(data, 5, replace = TRUE))

}

Sample_Mean_5
Standardized_5 = sqrt(5)*(Sample_Mean_5 - .5)/(0.08)
# Standardized sample mean.

hist(Standardized_5, main ="Size=5", xlab = "Sample mean", ylab = "Probability")
# Histogram of standardized sample mean.

# Simulate 1000 many random samples each with size 15 from the population.

Sample_Mean_10<-c() # vector to store the sample means.

n=1000 # Number of iteration.

for(i in 1:n){
  Sample_Mean_10[i] = mean(sample(data, 10, replace = TRUE))
  # Sample mean for i th sample.
}

Sample_Mean_10
Standardized_10 = sqrt(10)*(Sample_Mean_10 - 0.5)/sqrt(0.08)
hist(Standardized_10, xlab = "Sample mean",ylab = "Probability", main = "Size=10")
#quantile(Standardized_10, probs = c(0.25, 0.975, 1))

## Simulate 1000 many random samples each with size 20 from the population

Sample_Mean_15 <- c() # vector to store the sample means.

n=1000 # Number of iteration.

for(i in 1:n){
  Sample_Mean_15[i] = mean(sample(data, 15, replace = TRUE))
}

Sample_Mean_15
Standardized_15 = sqrt(15)*(Sample_Mean_15 - 0.5)/sqrt(0.08)
hist(Standardized_15,xlab = "Sample mean", ylab = "Probability", main = " Size=15")
```




```
#quantile(Standardized_15, probs = c(0.25, 0.975, 1))
```

In the the above, Figure 1 represents population distribution of standard uniform distribution. We draw 1000 random sample of different sizes from Figure 1 population. Figure 2 shows the histogram of standardized samples mean computed from samples of size 5. Similarly, Figure 3 is the histogram of standardized sample mean based on samples of size 10, and Figure 4 is the histogram of standardized sample mean based on the sample of size 15. As sample size increases the histogram of standardized samples mean tends to standard normal distribution.