P_M6_1

December 18, 2021

1 Module 6 Peer Review Assignment

2 Problem 1

Suppose X and Y are independent normal random variables with the same mean μ and the same variance σ^2. Do the random variables $W = X + Y$ and $U = 2X$ have the same distribution? Explain.

No because not every single element in X equals the respective element in Y so X + Y is not equal 2X

3 Problem 2: Central Limit Theorem and Simulation

a) For this problem, we will be sampling from the Uniform distribution with bounds [0, 100]. Before we simulate anything, let's make sure we understand what values to expect. If $X \sim U(0, 100)$, what is $E[X]$ and $Var(X)$?

[ ]:
```
E(X) = (a+b)/2 = 50
Var(X) = 1/12 * (b-a)**2 = 2500/3
```

b) In real life, if we want to estimate the mean of a population, we have to draw a sample from that population and compute the sample mean. The important questions we have to ask are things like:

- Is the sample mean a good approximation of the population mean?
- How large does my sample need to be in order for the sample mean to well-approximate the population mean?

Complete the following function to sample n rows from the $U(0, 100)$ distribution and return the sample mean. Start with a sample size of 10 and draw a sample mean from your function. Is the estimated mean a good approximation for the population mean we computed above? What if you increase the sample size?

[62]:
```
uniform.sample.mean = function(n){

    a = runif(n, 0, 100)
    sample.mean = mean(a)
```

```
    return(sample.mean)
}

uniform.sample.mean(10)
uniform.sample.mean(100)
uniform.sample.mean(999)
```

42.5144718633965

49.3557154964656

49.360384404388

The mean of 10 samples might be as close to our E(X) since the sample is quite small. If we increase n, the mean value will come closer to E(X).

c) Notice, for a sample size of n, our function is returning an estimator of the form

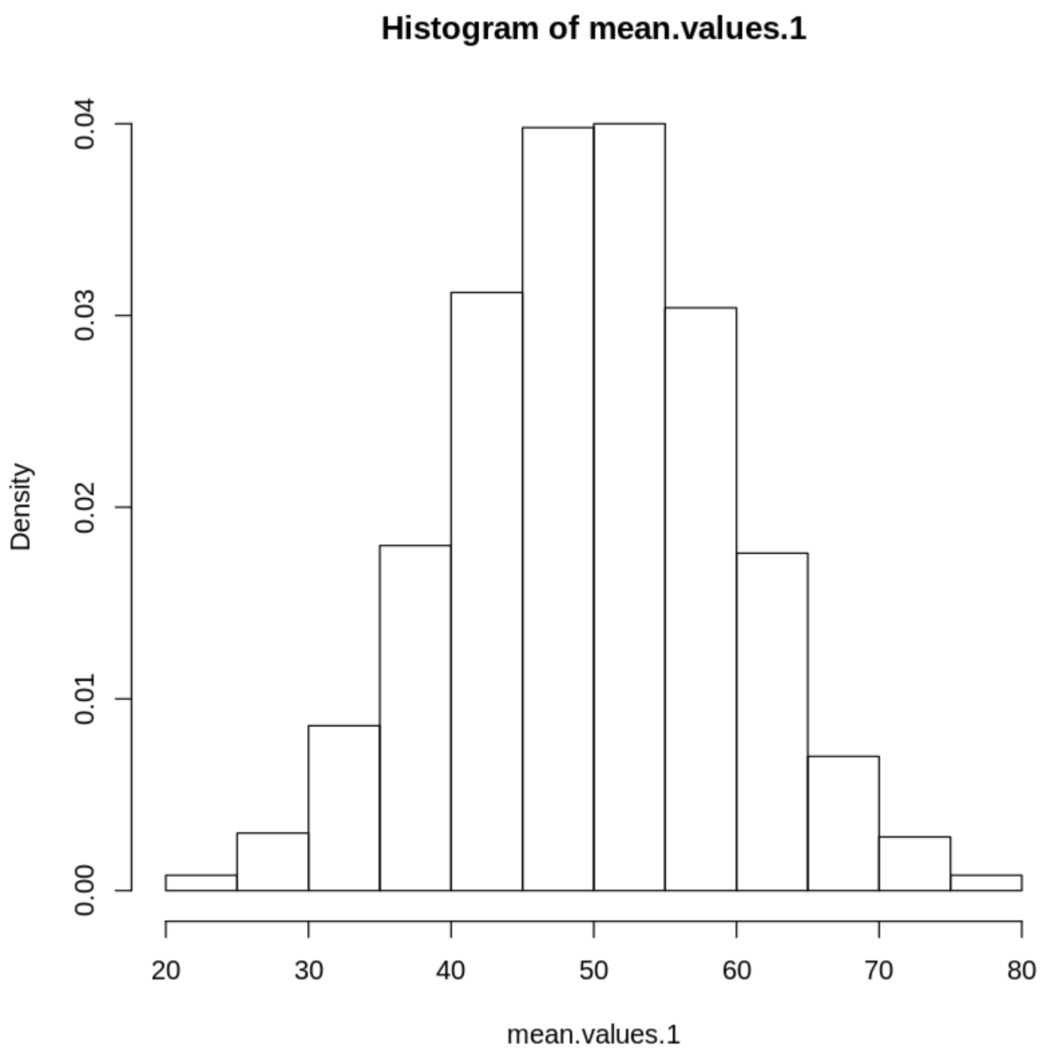
$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

That means, if each X_i is a random variable, then our sample mean is also a random variable with its own distribution. We call this disribution the sample distribution. Let's take a look at what this distribution looks like.

Using the `uniform.sample.mean` function, simulate $m = 1000$ sample means, each from a sample of size $n = 10$. Create a histogram of these sample means. Then increase the value of n and plot the histogram of those sample means. What do you notice about the distribution of $\bar{X}$? What is the mean μ and variance σ^2 of the sample distribution?

[16]:
```
mean.values.1 = matrix(, nrow = 1, ncol = 1000)
var.1 = matrix(, nrow = 1, ncol = 1000)
for (i in 1:1000){
    a = runif(10, 0, 100)
    mean.values.1[1, i] = sum(a)/10
    var.1[1, i] = var(a)
}
```

[22]:
```
hist(mean.values.1, freq = FALSE)
```



[7]:
```
mean.values.2
```

A matrix: 1×1000 of type dbl 50.96975 50.32648 50.4059 50.15236 49.69397 51.27966 49.27623 49.30

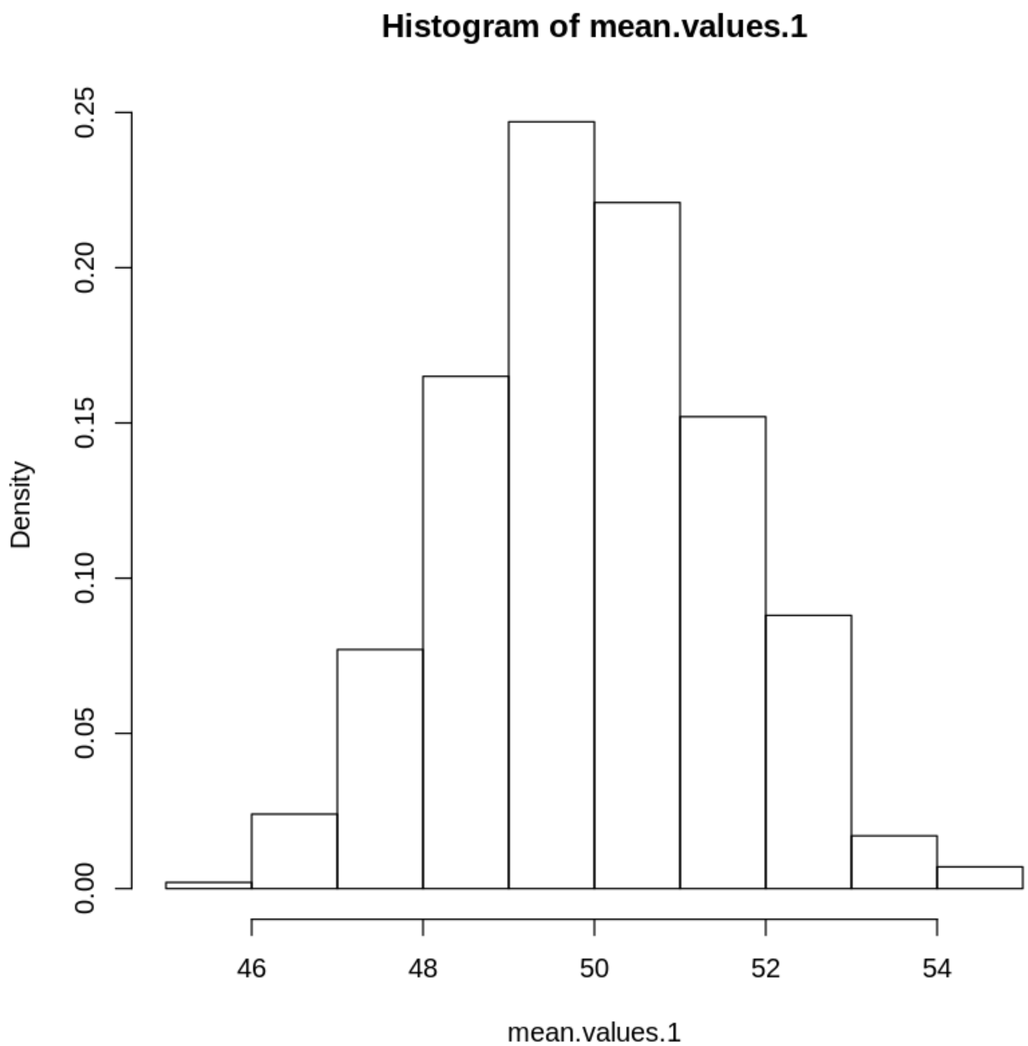
X.bar has a normal distribution with mu is approximately 50. And, its variance decreases as n increases.

d) Recall that our underlying population distribution is $U(0, 100)$. Try changing the underlying distribution (For example a binomial(10, 0.5)) and check the sample distribution. Be sure to explain what you notice.

[63]:
```
mean.values.1 = matrix(, nrow = 1, ncol = 1000)
var.1 = matrix(, nrow = 1, ncol = 1000)
for (i in 1:1000){
```

```
    a = rbinom(10, 100, 0.5)
    mean.values.1[1, i] = mean(a)
    var.1[1, i] = var(a)
}
```

[64]:
```
hist(mean.values.1, freq = FALSE)
```



It's still pretty much the same. It is normally distributed with mu is approximately 50 and its variance decreases as n in creases.

4 Problem 3

Let X be a random variable for the face value of a fair d-sided die after a single roll. X follows a discrete uniform distribution of the form unif$\{1, d\}$. Below is the mean and variance of unif$\{1, d\}$.

$$E[X] = \frac{1+d}{2} \qquad Var(X) = \frac{(d-1+1)^2 - 1}{12}$$

a) Let $\bar{X}_n$ be the random variable for the mean of n die rolls. Based on the Central Limit Theorem, what distribution does $\bar{X}_n$ follow when $d = 6$.

X.bar ~ N(3.5, 35/(12n))

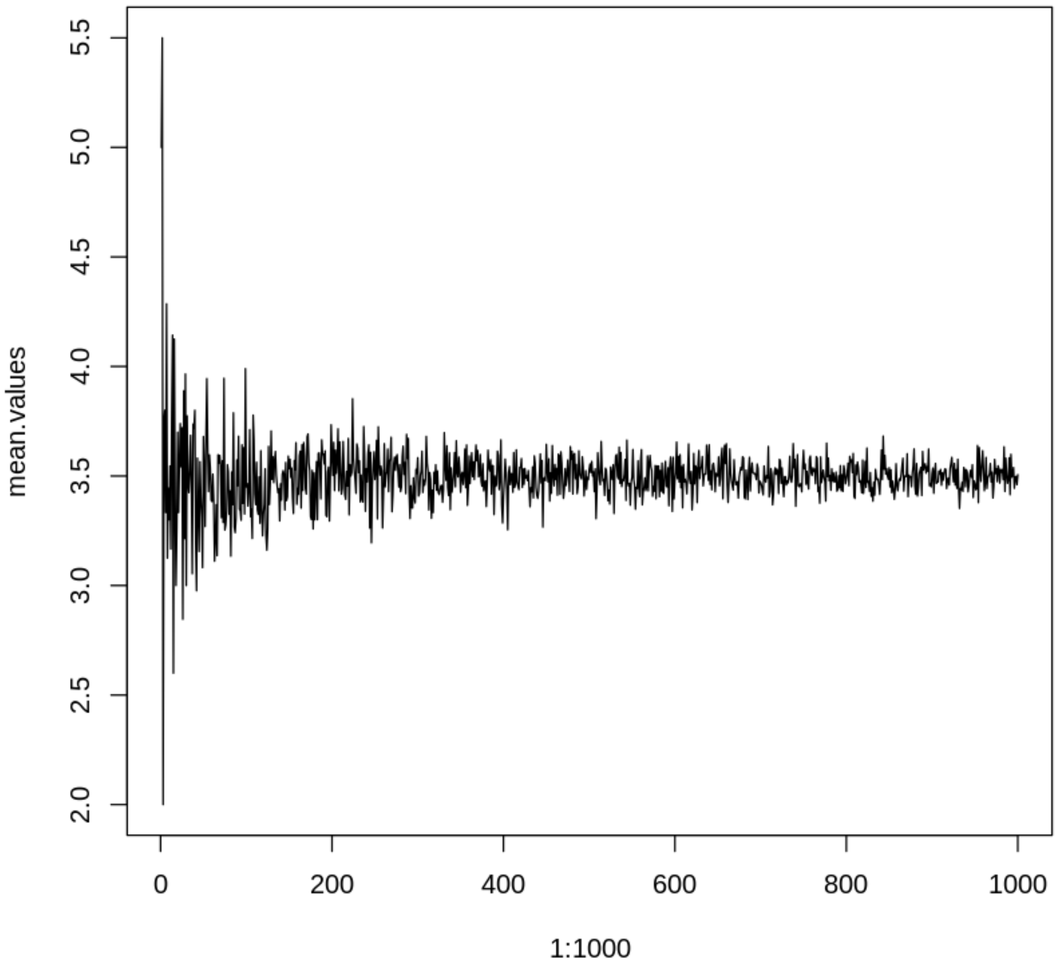
b) Generate $n = 1000$ die values, with $d = 6$. Calculate the running average of your die rolls. In other words, create an array r such that:

$$r[j] = \sum_{i=1}^{j} \frac{X_i}{j}$$

Finally, plot your running average per the number of iterations. What do you notice?

[65]:
```
mean.values = matrix(, nrow = 1, ncol = 1000)
for (i in 1:1000){
    mean.values[1, i] = sum(sample(1:6, i, replace = TRUE))/i
}
mean.values
plot(x = 1:1000, y = mean.values, type="l")
```

A matrix: 1×1000 of type dbl 5 5.5 2 3.75 3.8 3.333333 4.285714 3.125 3.444444 3.3 3.41473



The average value gets closer to 3.5 as i approaches 1000.