Math 490 HW #4

Maxwell Levin

January 31, 2018

Question 1.

Let random variable U be uniformly distributed over the interval [a, b]. Find $E(U^2)$ and $Var(U^2)$.

Since U is a continuous random uniform distribution, we calculate $E(U^2)$ using the equation

$$E(U^2) = \int_a^b U^2 f(U) dU.$$

Since U is uniformly distributed, we know that f(U) is a constant, namely $f(U) = \frac{1}{b-a}$. Thus we have

$$E(U^2) = \frac{1}{b-a} \int_a^b U^2 dU,$$

$$E(U^2) = \frac{a^2 + ab + b^2}{3}.$$

We now calculate the variance of U^2 :

$$Var(U^2) = E((U^2 - \mu)^2),$$

$$Var(U^2) = E(U^4) - 2\mu E(U^2) + \mu^2,$$

$$Var(U^2) = \frac{b^5 - a^5}{5(b-a)} - \frac{b^3 - a^3}{3} + \left(\frac{b-a}{2}\right)^2.$$

Question 2.

Consider sampling n = 4 random numbers U_1, U_2, U_3, U_4 from the Uniform[0,1] distribution, and let

$$\overline{V_4} = \frac{U_1^2 + U_2^2 + U_3^2 + U_4^2}{4}.$$

a. What are the mean and standard deviation of $\overline{V_4}$?

We calculate the mean:

$$E(\overline{V_4}) = E\left(\frac{U_1^2 + U_2^2 + U_3^2 + U_4^2}{4}\right),$$

$$E(\overline{V_4}) = E\left(\frac{1}{4}E(U_1^2) + \frac{1}{4}E(U_2^2) + \frac{1}{4}E(U_3^2) + \frac{1}{4}E(U_4^2)\right),$$

$$E(\overline{V_4}) = E(U^2).$$

Thus we can use our result from question 1 with $a = 0, b = 1$ to get

$$E(\overline{V_4}) = \frac{1}{3}.$$

To calculate variance we use

$$Var(\overline{V_4}) = Var\left(\frac{U_1^2 + U_2^2 + U_3^2 + U_4^2}{4}\right),$$

$$Var(\overline{V_4}) = \left(\frac{1}{4}\right)^2 Var(U_1^2) + \left(\frac{1}{4}\right)^2 Var(U_2^2) + \left(\frac{1}{4}\right)^2 Var(U_3^2) + \left(\frac{1}{4}\right)^2 Var(U_4^2),$$

$$Var(\overline{V_4}) = \frac{1}{4}Var(U^2).$$

Thus we can again apply our results from question 1 to achieve

$$Var(\overline{V_4}) = \frac{7}{240} \approx 0.0292,$$

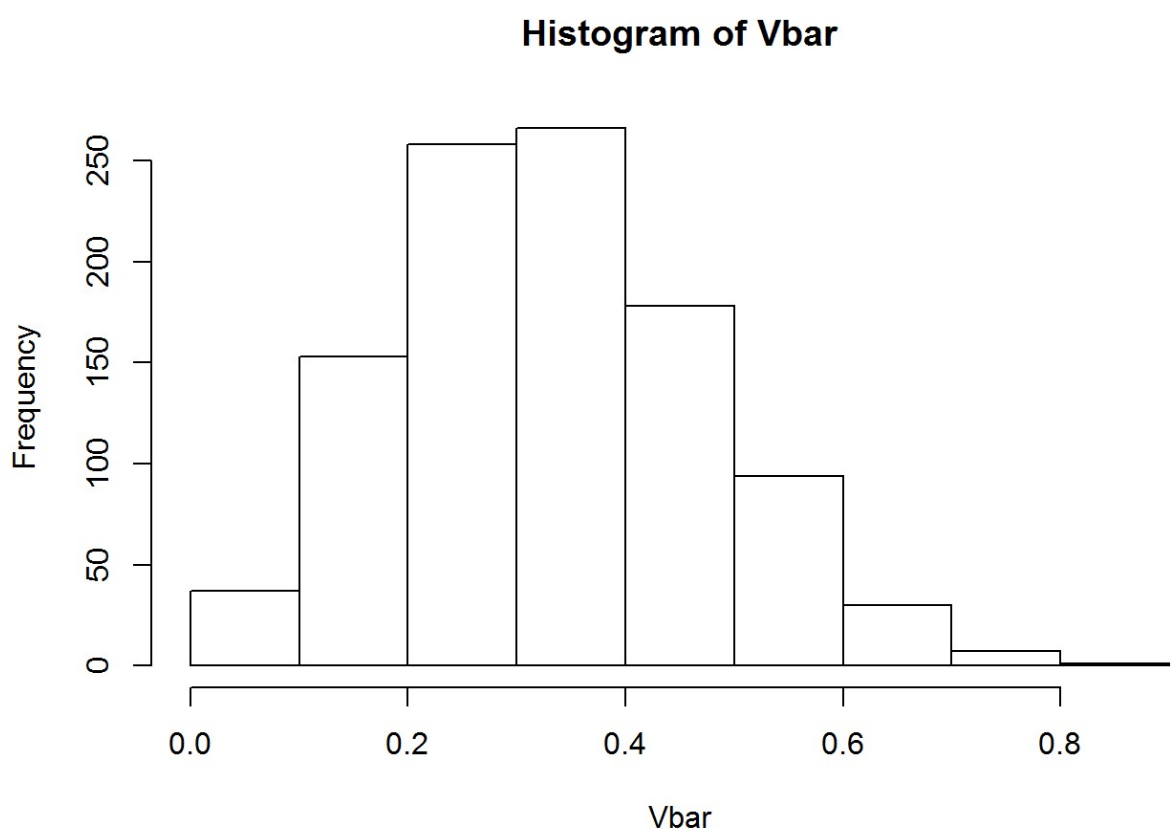
$$SD(\overline{V_4}) = \sqrt{Var(\overline{V_4})} \approx 0.1708$$

b. Use R to simulate 1024 sample means $\overline{V_4}$?, make a histogram, and report the mean (center) and standard deviation (spread) of the simulated sample means.

```
u1 = runif(1024, 0, 1)
u2 = runif(1024, 0, 1)
u3 = runif(1024, 0, 1)
u4 = runif(1024, 0, 1)

Vbar = (u1^2 + u2^2 + u3^2 + u4^2) / 4

hist(Vbar)
```



The Mean is

```
mean(Vbar)
```

```
## [1] 0.3322512
```

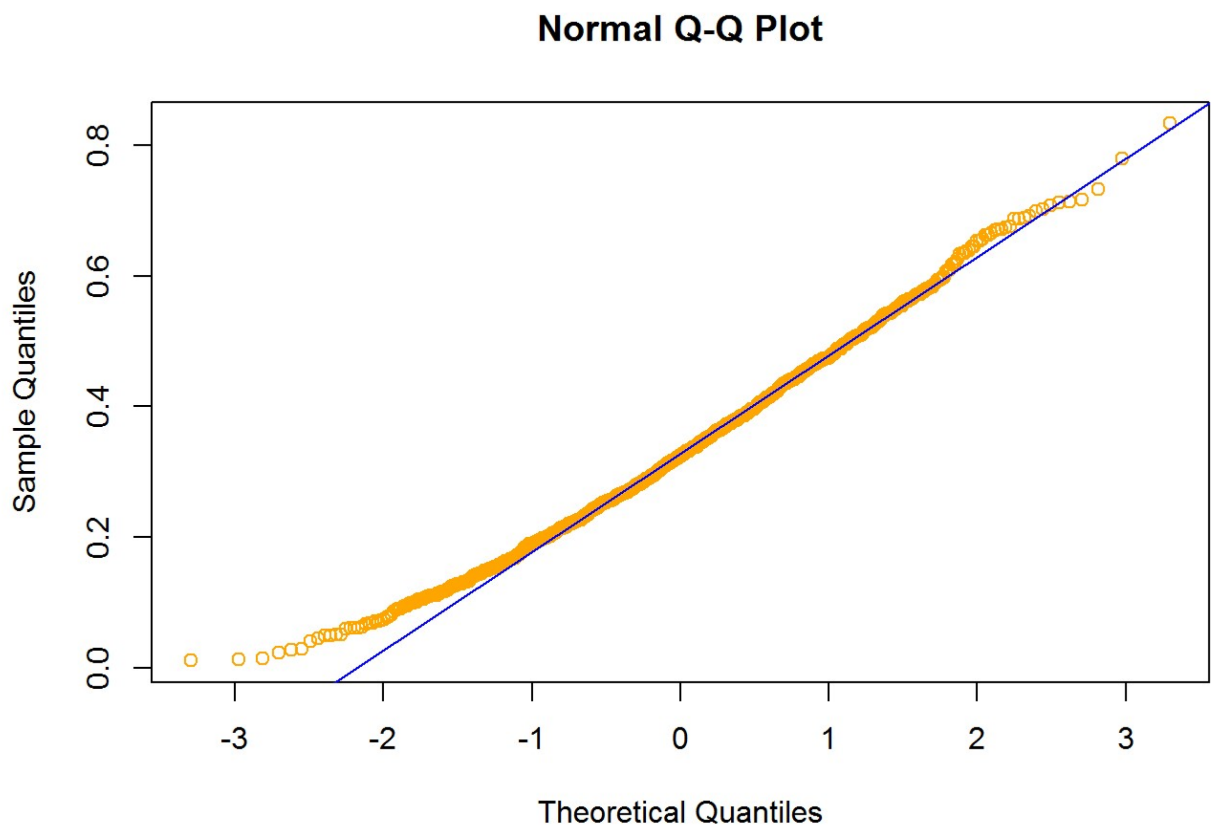
The Standard Deviation is

```
SD = sqrt(sum((Vbar - mean(Vbar))^2)/1024)
SD
```

```
## [1] 0.1425198
```

c. Use a normal probability plot to access the normality of the 1024 sample means simulated in part b.

```
qqnorm(Vbar, col="orange")
qqline(Vbar, col="blue")
```



Our normal probability plot indicates that our sample means are pretty close to a normal distribution.

d. Use the Shapiro-Wilk test to assess the normality of the 1024 sample means simulated in part b.

```
shapiro.test(Vbar)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  Vbar
## W = 0.9911, p-value = 7.448e-06
```

Our Shapiro-Wilk test tells us that there is very strong evidence for normality of our sample means.

Question 3.

Consider sampling n = 25 random numbers $U_1, U_2, \ldots, U_25$ from the Uniform [0,1] distribution, and let

$$\overline{V_{25}} = \frac{U_1^2 + U_2^2 + \cdots + U_{25}^2}{25}.$$

a. What are the mean and standard deviation of $\overline{V_4}$?

We calculate the mean:

$$E(\overline{V_{25}}) = E\left(\frac{U_1^2 + U_2^2 + \cdots + U_{25}^2}{25}\right),$$

$$E(\overline{V_{25}}) = E\left(\frac{1}{25}E(U_1^2) + \frac{1}{25}E(U_2^2) + \cdots + \frac{1}{25}E(U_{25}^2)\right),$$

$$E(\overline{V_{25}}) = E(U^2).$$

Thus for an interval from 0 to 1 our mean is

$$E(\overline{V_{25}}) = \frac{1}{3}.$$

We calculate the variance:

$$Var(\overline{V_{25}}) = Var\left(\frac{U_1^2 + U_2^2 + \cdots + U_{25}^2}{25}\right)$$

$$Var(\overline{V_{25}}) = \left(\frac{1}{25}\right)^2 Var(U_1^2) + \left(\frac{1}{25}\right)^2 Var(U_2^2) + \cdots + \left(\frac{1}{25}\right)^2 Var(U_{25}^2)$$

$$Var(\overline{V_{25}}) = \frac{1}{25}Var(U^2).$$

We can apply our results from question 2 to get

$$Var(\overline{V_{25}}) = \frac{7}{1500} \approx 0.0047$$

That is, the standard deviation of our sample means when n = 25 is

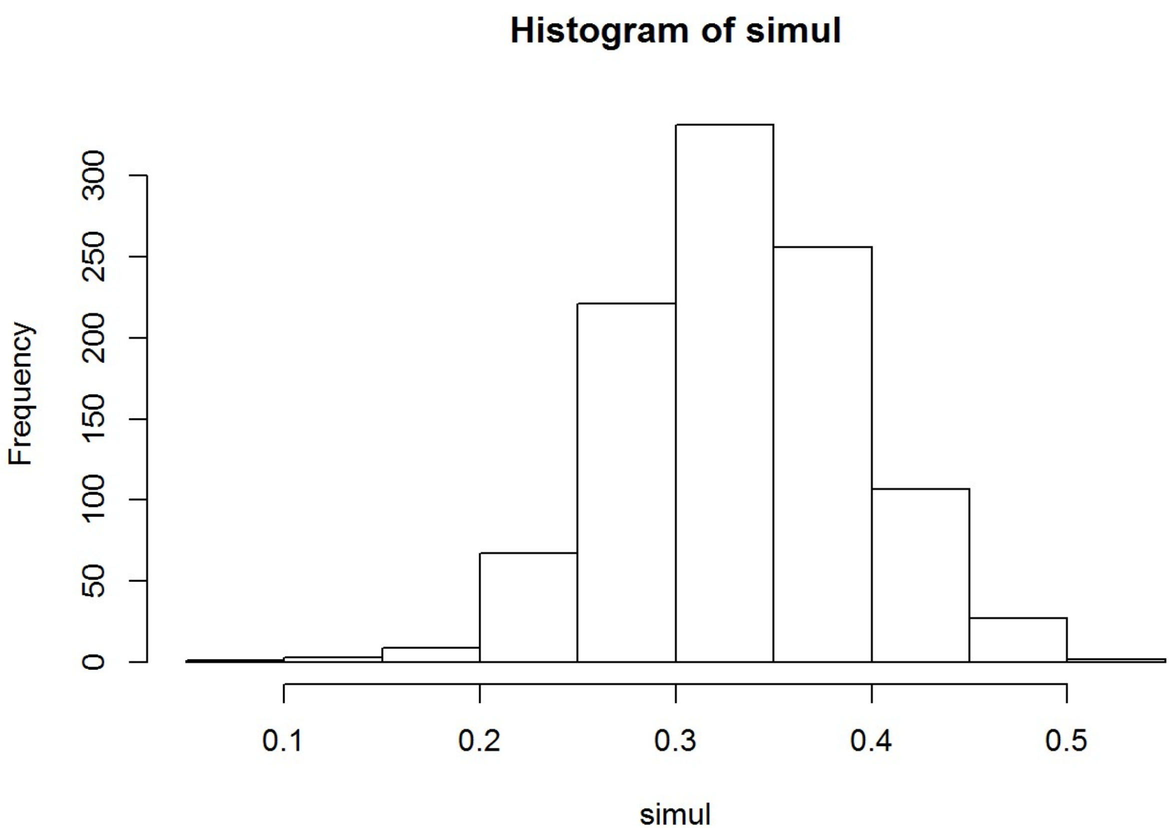
$$SD(\overline{V_{25}}) = \sqrt{Var(\overline{V_{25}})} \approx 0.0683.$$

b. Use R to simulate 1024 sample means $\bar{V}_{25}$?, make a histogram, and report the mean (center) and standard deviation (spread) of the simulated sample means.

```
clt = function(sam, rep) {
  obs = NULL
  for (i in 1:rep) {
    v = runif(sam, 0, 1)^2
    Vbar = mean(v)
    obs = c(obs, Vbar)
  }
  obs;
}

simul = clt(25, 1024)

hist(simul)
```



The mean is

```
mean(simul)
```

```
## [1] 0.3319743
```

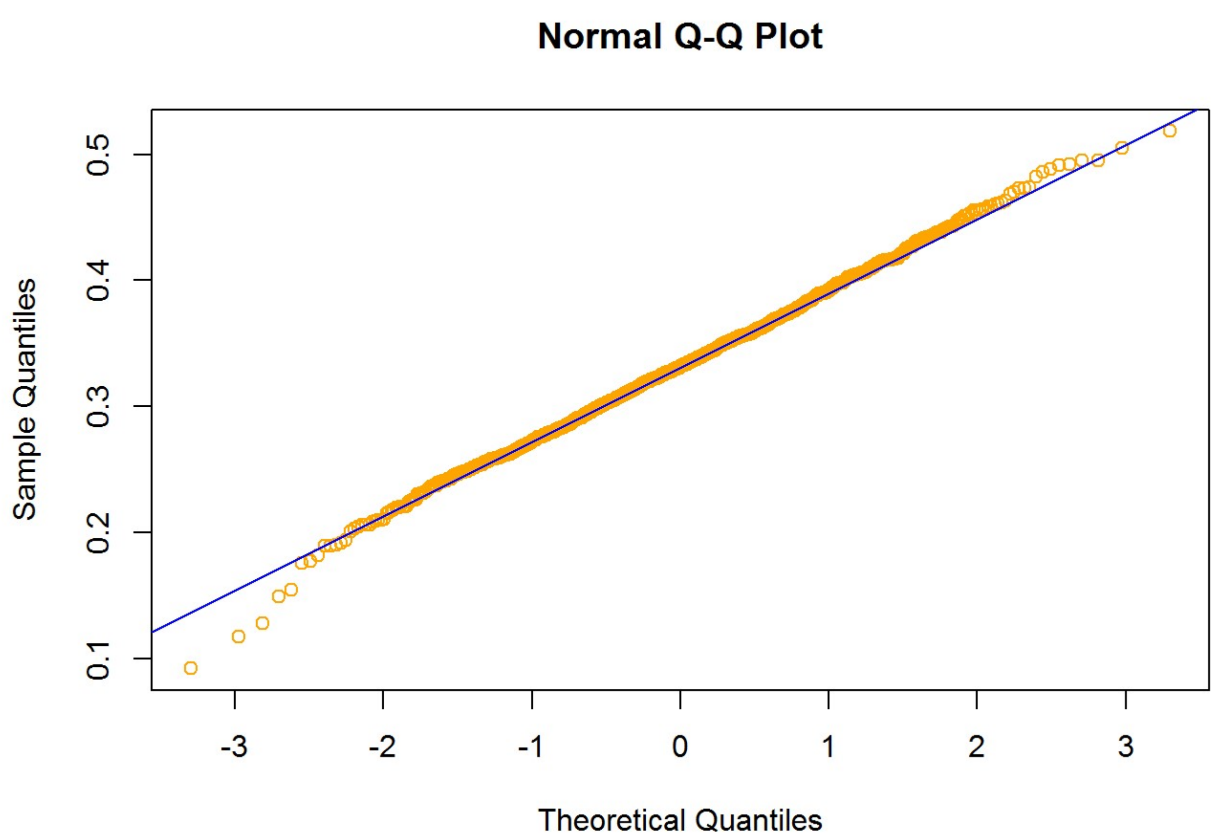
The standard deviation is

```
SD = sqrt(sum((simul - mean(simul))^2)/1024)
SD
```

```
## [1] 0.06025003
```

c. Use a normal probability plot to access the normality of the 1024 sample means simulated in part b.

```
qqnorm(simul, col="orange")
qqline(simul, col="blue")
```



Our normal probability plot suggests that our simulation of sample means is pretty close to a normal distribution.

d. Use the Shapiro-Wilk test to assess the normality of the 1024 sample means simulated in part b.

```
shapiro.test(simul)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  simul
## W = 0.99801, p-value = 0.2693
```

Our Shapiro-Wilk normality test gives us no evidence for normality since the P-value is larger than 0.1.