Brian KYANJO
MATH 568
Probability and Statistics Review

1. Let $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$, $E(X) = \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$, $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$

a) Use the fact that $E(aX) = aE(X)$, show that

$$E(AX) = AE(X)$$

Let $Y = AX = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} a_{11}X_1 + a_{12}X_2 \\ a_{21}X_1 + a_{22}X_2 \end{bmatrix}$

$$E(AX) = E(Y) = \begin{bmatrix} E(a_{11}X_1 + a_{12}X_2) \\ E(a_{21}X_1 + a_{22}X_2) \end{bmatrix}$$

$$E(Y) = \begin{bmatrix} a_{11}E(X_1) + a_{12}E(X_2) \\ a_{21}E(X_1) + a_{22}E(X_2) \end{bmatrix}$$

$$E(Y) = \underbrace{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} E(X_1) \\ E(X_2) \end{bmatrix}}_{E(X)}$$

$$E(Y) = AE(X)$$

Hence $\underline{E(AX) = AE(X)}$

b) If $Y = AX$, find $Var(Y_1)$,

$$Y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_{11}X_1 + a_{12}X_2 \\ a_{21}X_1 + a_{22}X_2 \end{bmatrix}$$

<u>Variance of Y_1</u>

$$Y_1 = a_{11}X_1 + a_{12}X_2$$

$$Var(Y_1) = Var(a_{11}X_1 + a_{12}X_2)$$

$$Var(Y_1) = a_{11}^2 Var(X_1) + a_{12}^2 Var(X_2) + 2\,Cov(a_{11}X_1, a_{12}X_2)$$

but $Cov(a_{11}X_1, a_{12}X_2) = a_{11}a_{12}\,Cov(X_1, X_2)$

then $Var(Y_1) = a_{11}^2 Var(X_1) + a_{12}^2 Var(X_2) + 2a_{11}a_{12}\,Cov(X_1, X_2)$

<u>Variance of Y_2</u>

$$Y_2 = a_{21}X_1 + a_{22}X_2$$

$$Var(Y_2) = a_{21}^2 Var(X_1) + a_{22}^2 Var(X_2) + 2\,Cov(a_{21}X_1, a_{22}X_2)$$

$$Var(Y_2) = a_{21}^2 Var(X_1) + a_{22}^2 Var(X_2) + 2a_{21}a_{22}\,Cov(X_1, X_2)$$

Covariance of Y_1 and Y_2

$$Cov(Y_1, Y_2) = E(Y_1 Y_2) - E(Y_1) E(Y_2)$$

but $Y_1 = a_{11} X_1 + a_{12} X_2 \Rightarrow E(Y_1) = a_{11} E(X_1) + a_{12} E(X_2)$

$Y_2 = a_{21} X_1 + a_{22} X_2 \Rightarrow E(Y_2) = a_{21} E(X_1) + a_{22} E(X_2)$

$$Y_1 Y_2 = (a_{11} X_1 + a_{12} X_2)(a_{21} X_1 + a_{22} X_2) = a_{11} a_{21} X_1^2 + a_{11} a_{22} X_1 X_2 + a_{12} a_{21} X_2 X_1 + a_{12} a_{22} X_2^2$$

$$E(Y_1 Y_2) = a_{11} a_{21} E(X_1^2) + a_{11} a_{22} E(X_1 X_2) + a_{12} a_{21} E(X_2 X_1) + a_{12} a_{22} E(X_2^2) \quad \text{— ①}$$

$$E(Y_1) E(Y_2) = a_{11} a_{21} E(X_1) E(X_1) + a_{11} a_{22} E(X_1) E(X_2) + a_{12} a_{21} E(X_2) E(X_1) + a_{12} a_{22} E(X_2) E(X_2) \quad \text{—— ②}$$

$$Cov(Y_1, Y_2) = E(Y_1 Y_2) - E(Y_1) E(Y_2)$$

$$Cov(Y_1, Y_2) = a_{11} a_{21} \left(E(X_1^2) - (E(X_1))^2 \right) + a_{11} a_{22} \left(E(X_1 X_2) - E(X_1) E(X_2) \right) + a_{12} a_{21} \left(E(X_2 X_1) - E(X_2) E(X_1) \right) + a_{12} a_{22} \left(E(X_2^2) - E(X_2) E(X_2) \right)$$

$$Cov(Y_1, Y_2) = a_{11} a_{21} Var(X_1) + a_{12} a_{22} Var(X_2) + a_{11} a_{22} Cov(X_1, X_2) + a_{12} a_{21} Cov(X_2, X_1)$$

$$Cov(Y_1, Y_2) = a_{11} a_{21} Var(X_1) + a_{12} a_{22} Var(X_2) + (a_{11} a_{22} + a_{12} a_{21}) Cov(X_1, X_2)$$

No.2 Do Exercise 8

```
% Given the data
m = [-0.4326 -1.6656 0.1253 0.2877 -1.1465 1.1909 1.1892 -0.0376 0.3273 0.1746];
n = length(m);
```

Find the sample mean

```
mbar = mean(m)
```

```
mbar = 0.0013
```

Find the standard deviation

```
s = std(m)
```

```
s = 0.9034
```

Construct a 95% confidence interval for the mean

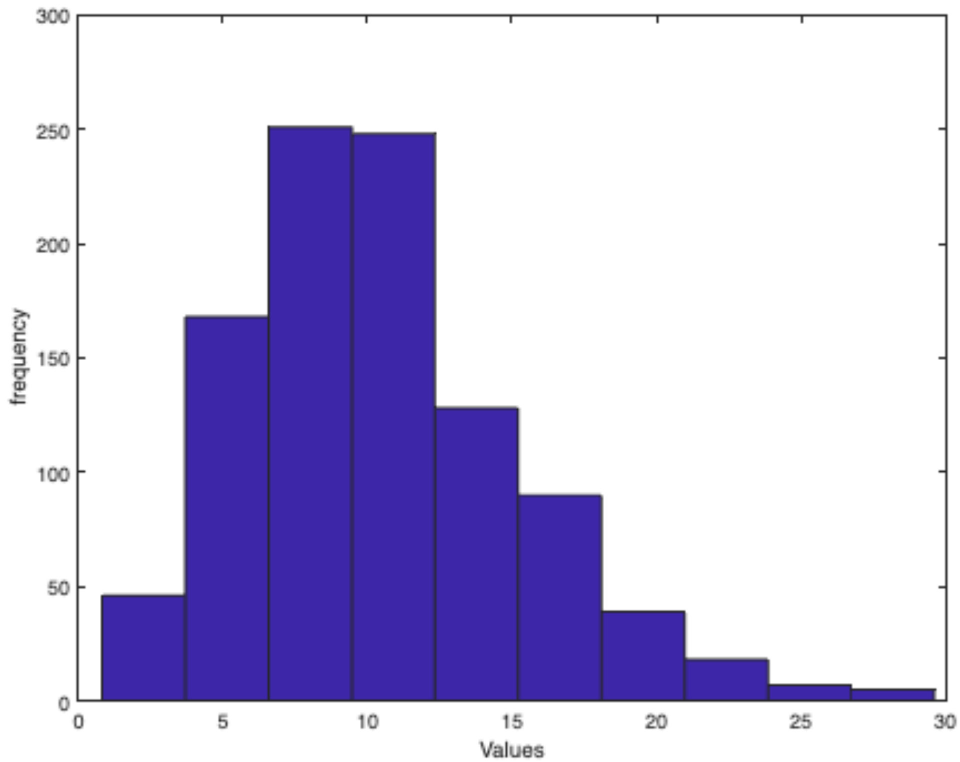
```
ci = 0.95;
top = 1 - (1 - ci)/2;
bottom = (1 - ci)/2;
t_top = tinv(top,n-1);
t_bottom = tinv(bottom,n-1);

% confidence intervals
c1 = mbar + t_bottom*s/sqrt(n);
c2 = mbar + t_top*s/sqrt(n);

cofidence_interval = [c1 c2]
```

```
cofidence_interval = 1×2
   -0.6450    0.6475
```

No.3

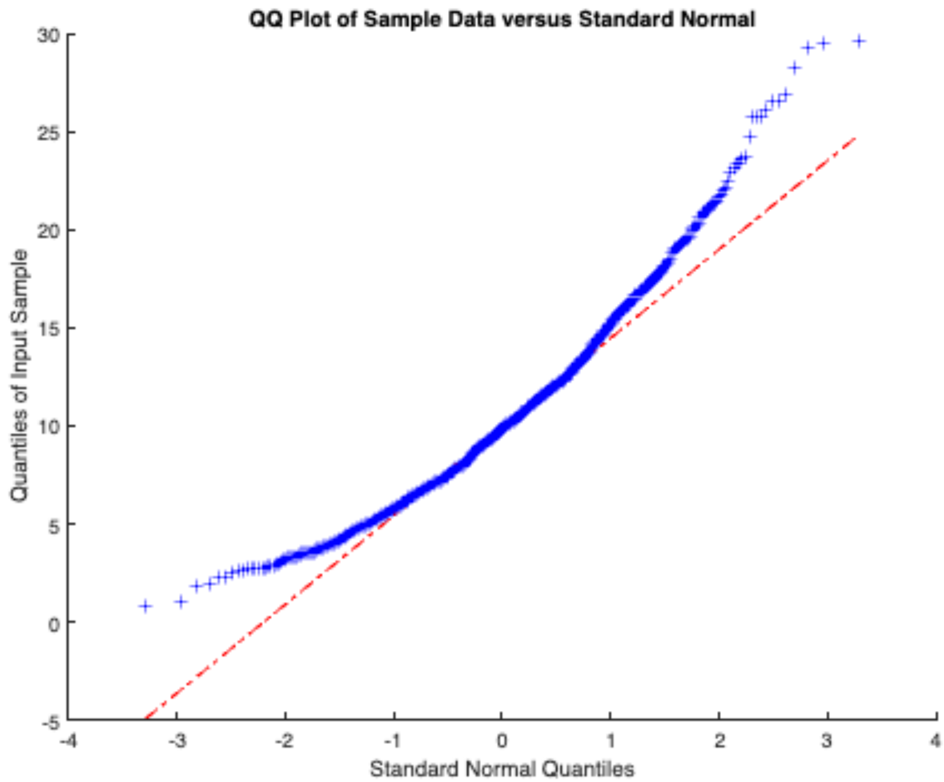
a). Plot a histogram of the 1000 average

```
n3 = 5; % 5 exponentially distributed random numbers

[mbare3] = generate(n3);

% plot a histogram
figure(1)
hist(mbare3)
ylabel('frequency');
xlabel('Values');
```



Make a Q-Q plot of the 1000 averages

```
figure(2)
qqplot(mbare3)
ylabel('Quantiles of Input Sample');
xlabel('Standard Normal Quantiles');
```



Explain why or why not the averages are approximately normally distributed in each case.

In this case the averages are not approximately normally distributed in each case because the histogram shows that the distribution is skewed. The data points enclosed in the area under the curve which can be drawn on the histogram, are very close to the mean at one side, and far from the mean on the other side. And also since the graphical test from the Q-Q plot doesn't depeicts a straight line for the quantile data and standard normal quantiles.

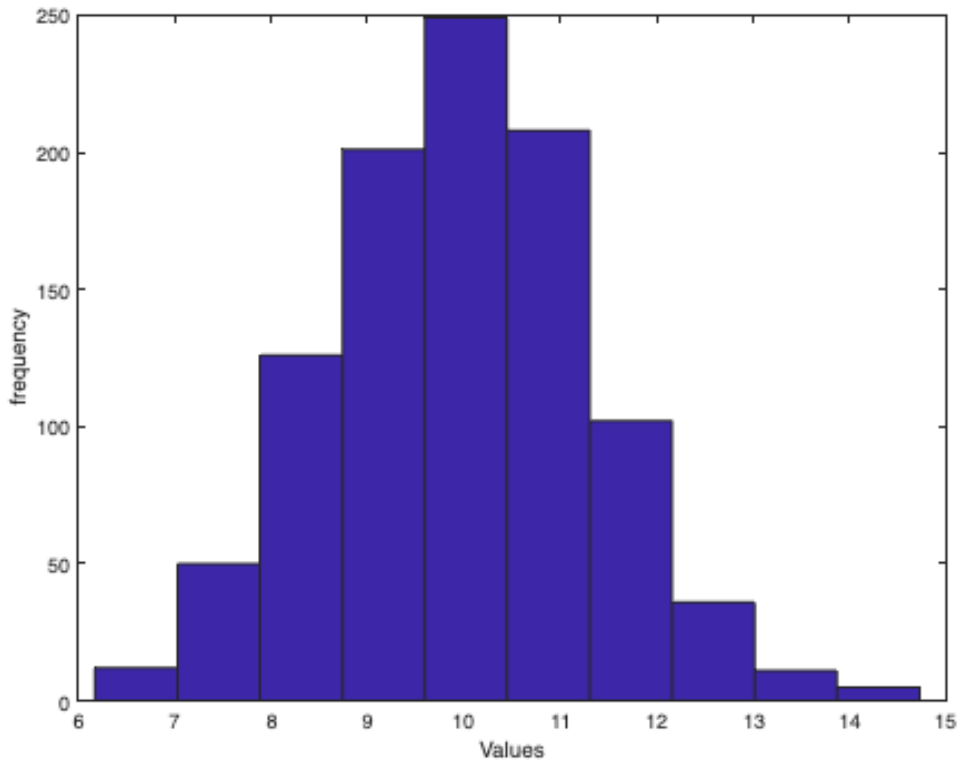
No.4 Repeat the experiment in 3.

Generate 50 exponentially distributed random numbers.

```
nn = 50; % 50 exponentially distributed random numbers

[mbare4] = generate(nn);
```

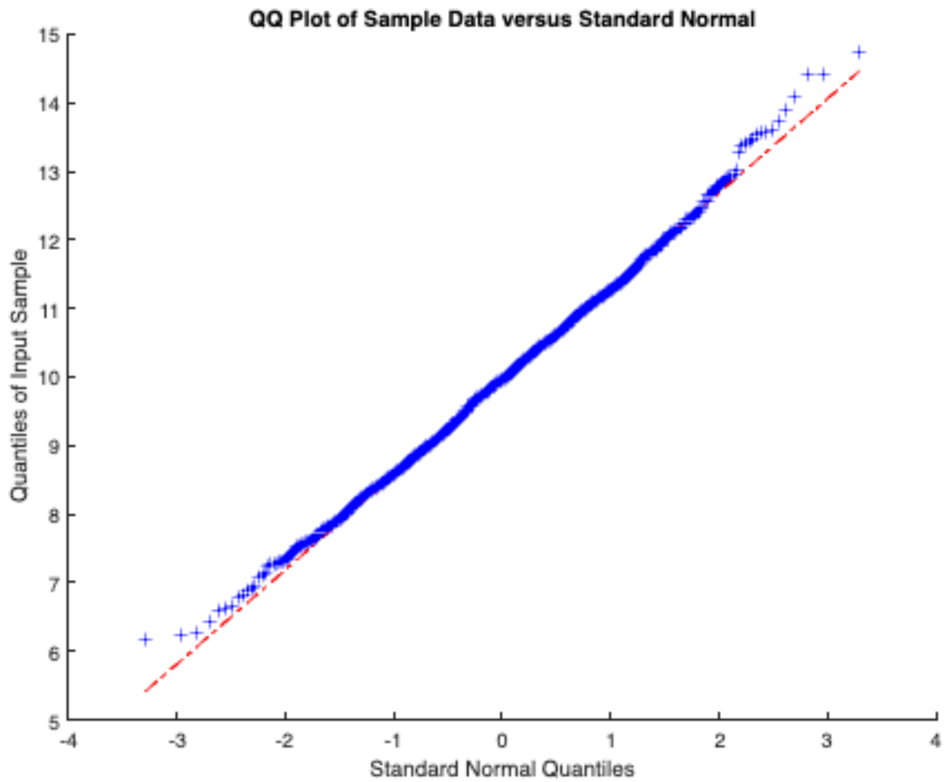
Plot a histogram of the 1000 average

```
% plot a histogram
figure(3)
hist(mbare4)
ylabel('frequency');
xlabel('Values');
```



Make a Q-Q plot of the 1000 averages

```
figure(4)
qqplot(mbare4)
ylabel('Quantiles of Input Sample');
xlabel('Standard Normal Quantiles');
```



Explain why or why not the averages are approximately normally distributed in each case.

In this case the averages are approximately normally distributed because the histogram shows a roughly bell-shaped distribution and its symmetrical around a single peak. The data points enclosed in the area under the curve which can be drawn on the histogram, are very close to the mean, and are at almost equal distances from each either sides of the mean. And also since the graphical test from the Q-Q plot depeicts that the quantile data and standard normal quantiles almost follow a straight line.

```
function [mbare] = generate(nn)
    N = 1000;

    mbare = zeros(N,1);

    for i=1:N
        expd = exprnd(10,nn,1); % nn exponentially distributed random numbers
        mbare(i) = mean(expd); % mean of the five numbers
    end
end
```