Sage Herrin, ASEN 5044

HW 5, Due 3/5/20

1.) Discrete random variables X & Y can each take on values 1, 3, & 5. Joint probability table of X & Y given as

X	Y = 1	Y = 3	Y = 5
1	1/18	1/18	1/18
3	1/18	1/18	1/6
5	1/18	1/6	1/3

a.) random variables X & Y are not independent because the probabilities of Y change with the value of X, e.g. the second column of the table

b.) The marginal probability $p(Y=3) = \frac{1}{18} + \frac{1}{18} + \frac{1}{6} = \boxed{\frac{5}{18}}$

c.) The conditional probability $p(Y=3 \mid X=3)$

$$\Rightarrow \frac{p(Y=3 \cap X=3)}{p(Y=3)} \Rightarrow \frac{1/18}{5/18} = \boxed{\frac{1}{5}}$$

2.) Generate N = 50 independent random numbers, each uniformly distributed between 0 & 1. Plot histogram of random numbers using 10 bins. What is sample mean & standard deviation of numbers that were generated? What would you expect to see for the mean & standard deviation (i.e., what are theoretical mean & standard deviation? Repeat for N = 500 & N = 5,000 random #s. What changes in histogram do you see as N increases?

Done in Matlab (Histograms shown in appended code), the sample mean & standard deviation for N = 50 was found to be 0.4385 & 0.2684, respectively.

The theoretical mean is calculated by $\int_a^b x f(x) dx$, where $f(x) = \frac{1}{b-a} \in [0,1]$

$$\Rightarrow \text{mean} = \int_0^1 x(1) dx = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2} - 0 = \boxed{\frac{1}{2}}$$

variance is calculated as $\sigma^2 = \int_0^1 (x - 1/2)^2 dx = \int_0^1 x^2 - x + \frac{1}{4} dx = \left[\frac{x^3}{3} - \frac{x^2}{2} + \frac{x}{4}\right]_0^1 = \frac{1}{12} = \sigma^2$

$\Rightarrow$ standard dev. $= \sqrt{\sigma^2} = \sqrt{1/12} \approx 0.289$

for N = 500, mean = 0.4838 & standard deviation = 0.2833

for N = 5,000, mean = 0.4998 & standard deviation = 0.2894

2.15 cont. As N increases, the number of occurrences for ~~the~~ random numbers between 0 & 1 for all 10 bins become closer & closer to being equal, ~~[illegible]~~ so their mean & standard deviation are approaching the theoretical mean & standard deviation.

2.16/b Generate 10,000 of $(x_1+x_2)/2$, where each x_i is a random number uniformly distributed on $[-1/2, +1/2]$ plot 50-bin histogram. Repeat for $(x_1+x_2+x_3+x_4)/4$. Describe difference between histograms, repeat again for $(x_1+x_2+\dots x_{20})/20$ & describe differences between 3 histograms

As the number of x_i points used increases, the distribution on the histogram narrows, & the maximum number of occurrences, or items in each bin, increases, As seen in the 3 histograms in the appended code

2.8/3 Consider 2 zero-mean uncorrelated random variables W & V with standard deviations σ_W & σ_V, respectively. What is the standard deviation of the random variable $X = W+V$?

$$\Rightarrow \text{for } X = W+V,\ \sigma_X^2 = \int_{-\infty}^{\infty} (x-\bar{x})^2 f_X(x)\,dx \text{ , since } X = V+W,\ f_X(x) = f_V(x) + f_W(x)$$

$$\Rightarrow \sigma_X^2 = \int_{-\infty}^{\infty} (x-\bar{x})^2 (f_V(x) + f_W(x))\,dx \Rightarrow \sigma_X^2 = \int_{-\infty}^{\infty} (x-\bar{x})^2 f_V(x)\,dx + \int_{-\infty}^{\infty} (x-\bar{x})^2 f_W(x)\,dx$$

$$\Rightarrow \sigma_X^2 = \sigma_V^2 + \sigma_W^2 \Rightarrow \boxed{\sigma_X = \sqrt{\sigma_V^2 + \sigma_W^2}}$$

2.9/ 4 consider 2 scalar RVs X & Y,

a.) prove that if X & Y are independent, their correlation coefficient $\rho = 0$

$\Rightarrow$ Since RVs X & Y are independent, $P(X \le x, Y \le y) = P(X \le x)P(Y \le y)$ for all x, y. from definition of joint distribution & density functions,

$\Rightarrow F_{XY}(x,y) = F_X(x) \, \& \, F_Y(y)$, $f_{XY}(x,y) = f_X(x) f_Y(y)$

Knowing this & the definition of the correlation coefficient for 2 scalar RVs X & Y as $\rho = \frac{C_{XY}}{\sigma_X \sigma_Y}$. $\rho = 0 \Rightarrow C_{XY} = 0$

C_{XY} defined as $E[(X-\bar{X})(Y-\bar{Y})] \Rightarrow E(XY) - \bar{X}\bar{Y}$

using definition of $\bar{X}$ & E, $\bar{X} = \int_{-\infty}^{\infty} x f_X(x) dx$ & $\bar{Y} = \int_{-\infty}^{\infty} y f_Y(y) dy$

$E(XY) \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f(x,y) dx dy$, since X & Y are independent,

$$E(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f(x,y)\, dx dy \Rightarrow \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_X(x) f_Y(y)\, dx\, dy$$

$$\Rightarrow \int_{-\infty}^{\infty} x f_X(x) dx \int_{-\infty}^{\infty} y f_Y(y) dy$$

$$\Rightarrow C_{XY} = E(XY) - \bar{X}\bar{Y} = \int_{-\infty}^{\infty} x f_X(x) dx \int_{-\infty}^{\infty} y f_Y(y) dy - \int_{-\infty}^{\infty} x f_X(x) dx \int_{-\infty}^{\infty} y f_Y(y)\, dy$$

$= 0 \Rightarrow \boxed{\rho = 0 \text{ for independent RVs}}$

b.) Find example of 2 RVs that are not independent but that have a correlation coefficient of 0

$\Rightarrow \rho = 0 \Rightarrow$ Example would be one random variable x & another RV x^2, both on the range $[-1, 1]$. It is clear that these 2 RVs are not independent. However both variables have a zero mean, & 0 expected value so their correlation coefficient is 0, however both RVs are dependent on x, so they are not independent.

2.9 cont. c) Prove that if Y is a linear function of X, $\rho = \pm 1$

$\rho = \frac{C_{XY}}{\sigma_X \sigma_Y}$, Y linear function of $X \Rightarrow \rho = \pm 1$ $\quad C_{XY} = \pm\sigma_X \sigma_Y$

let the RV $X = X$ & the other RV $Y = \alpha X + \beta$ where α & β are some constants so that Y is a linear function of X

knowing $cor(X,Y) = \frac{cov(X,Y)}{\sqrt{var(X)var(Y)}} \Rightarrow \frac{cov(X, aX+b)}{\sqrt{var(X)var(aX+b)}}$

$cov(X,X) = var(X)$

$\Rightarrow \frac{a\,cov(X,X+b)}{\sqrt{a^2 var(X) var(X+b)}} \Rightarrow \frac{a\,cov(X,X) + cov(X,b)}{\sqrt{a^2 var(X)(var(X) + var(b))}} \Rightarrow \frac{a\,var(X)}{\sqrt{a^2 var(X)^2}}$

$\Rightarrow \boxed{\frac{a}{|a|} = \pm 1}$

2.10 / 5 consider the function

$$f_{XY} = \begin{cases} a e^{-2x} e^{-3y} & x > 0, y > 0 \\ 0 & \text{otherwise} \end{cases}$$

a.) find the value of a so that $f_{XY}(x,y)$ is a valid joint probability density function

given equation set 2.49 in the book, pdfs have the property

$F(x,y) \in [0,1]$

$\Rightarrow 0 \le a e^{-2x} e^{-3y} \le 1 \Rightarrow 0 \le \frac{a}{e^{2x} e^{3y}} \le 1$. Since e^x for $x > 0$,

min value of $e^x = 1 \Rightarrow$ for $0 \le \frac{a}{e^{2x} e^{3y}} \le 1$, $\boxed{0 \le a \le 1 \text{ for this to be a valid pdf to satisfy this equality}}$

b.) calculate $\bar{x}$ & $\bar{y}$

$\Rightarrow$ for $\bar{x}$, need $f_X(x)$. $f_X(x) = \int_0^\infty a e^{-2x} e^{-3y}\, dy \Rightarrow a e^{-2x} \int_0^\infty e^{-3y}\, dy$

let $a = 1$

$\Rightarrow a e^{-2x} \frac{e^{-3y}}{-3} \Big|_0^\infty \Rightarrow \frac{a e^{-2x}}{\frac{e^\infty}{-3}} - \frac{a e^{-2x}}{\frac{e^0}{-3}} = 0 - \frac{e^{2x}}{3} = \frac{e^{-2x}}{3} = f_X(x)$

$\Rightarrow \bar{x} = \int_0^\infty x \frac{e^{-2x}}{3}\, dx$, using IBP, let $u = x$, $dv = e^{-2x}$ $= \frac{1}{3}\int_0^\infty x \frac{e^{-2x}}{2} dx \Rightarrow \frac{e^{-2x}}{-2} x \Big|_0^\infty - \int_0^\infty \frac{e^{-2x}}{-2} dx$

$\Rightarrow v = \frac{e^{-2x}}{-2}$

$= x(0-0) - \frac{1}{3}\left(\frac{e^{-2x}}{4}\right)\Big|_0^\infty = -\frac{1}{3}\left(0 - \frac{1}{4}\right) = \boxed{\frac{1}{12} = \bar{x}}$

2.10 cont. b cont.) Similarly, $\bar{y} \Rightarrow$ need $f_Y(y)$, $f_Y(y) = \int_0^\infty e^{-2x} e^{-3y} dx$

$\Rightarrow e^{-3y} \int_0^\infty e^{-2x} dx = e^{-3y} \frac{e^{-2x}}{-2}\Big|_0^\infty = 0 - -\frac{1}{2} = \frac{1}{2}$

$\Rightarrow f_Y(y) = \frac{e^{-3y}}{2} \Rightarrow \bar{y} = \int_0^\infty \frac{y e^{-3y}}{2} dy$, using u-sub let $u = y$, $dv = e^{-3y}$, $v = \frac{e^{-3y}}{-3}$

$\Rightarrow \frac{1}{2}\int_0^\infty \frac{e^{-3y}}{2} dy = \frac{e^{-3y}}{-3} \cdot y\Big|_0^\infty - \int_0^\infty \frac{e^{-3y}}{-3} dy = \frac{1}{2}\left(\frac{e^{-3y}}{-9} \cdot y\Big|_0^\infty\right) - \frac{1}{2}\left(\frac{e^{-3y}}{-9}\Big|_0^\infty\right)$

$= (0 - 0) - \frac{1}{2}\left(0 - \frac{1}{9}\right) = \boxed{\frac{1}{18} = \bar{y}}$

c.) Calculate $E(x^2)$, $E(y^2)$, & $E(XY)$

$E[g(x)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx \Rightarrow E(x^2) = \int_0^\infty x^2 \frac{e^{-2x}}{3} dx$, IBP, let $u = x^2$, $dv = \frac{e^{-2x}}{3}$, $du = 2x$, $v = \frac{e^{-2x}}{-6}$

$\Rightarrow x^2 \frac{e^{-2x}}{6}\Big|_0^\infty - \int_0^\infty 2x \frac{e^{-2x}}{-6} dx$, IBP again, $u = 2x$, $dv = \frac{e^{-2x}}{-6}$, $du = 2$, $v = \frac{e^{-2x}}{12}$

$\Rightarrow 2x \frac{e^{-2x}}{12}\Big|_0^\infty - \int_0^\infty 2\frac{e^{-2x}}{12} dx \Rightarrow \left(x^2 \frac{e^{-2x}}{6}\right)\Big|_0^\infty - \left[\left(2x\frac{e^{-2x}}{12}\right)\Big|_0^\infty - \left(\frac{e^{-2x}}{12}\Big|_0^\infty\right)\right]$

$= 0 - \left[(0-0) - \left(0 - \frac{1}{12}\right)\right] = \boxed{\frac{1}{12}}$

Similarly, $E[y^2] \Rightarrow \int_0^\infty y^2 \frac{e^{-3y}}{2} dy$, IBP, let $u = y^2$, $dv = \frac{e^{-3y}}{2}$, $du = 2y$, $v = \frac{e^{-3y}}{-6}$

$\Rightarrow y^2 \frac{e^{-3y}}{6}\Big|_0^\infty - \int_0^\infty 2y \frac{e^{-3y}}{-6} dy$, IBP again, let $u = 2y$, $dv = \frac{e^{-3y}}{-6}$, $du = 2$, $v = \frac{e^{-3y}}{18}$

$\Rightarrow y^2 \frac{e^{-3y}}{6}\Big|_0^\infty - \left[2y\frac{e^{-3y}}{18}\Big|_0^\infty - \int_0^\infty \frac{2e^{-3y}}{18} dy\right]$

$\Rightarrow y^2 \frac{\cancel{e^{-3y}}}{6}\Big|_0^\infty - \left[2y\frac{\cancel{e^{-3y}}}{18}\Big|_0^\infty - \frac{e^{-3y}}{27}\Big|_0^\infty\right] \Rightarrow (0-0) - \left[(0-0) - \left(0 - \frac{1}{27}\right)\right] = \frac{1}{27}$

2.10 Cont.

(c. cont.) $E(XY) = \int_0^\infty \int_0^\infty xy e^{-2x} e^{-3y} \, dx\, dy$

$\Rightarrow x e^{-2x} dx$, IBP, $u = x,\ dv = e^{-2x}$, $du = dx,\ v = \frac{e^{-2x}}{-2}$ $\Rightarrow$ $\frac{xe^{-2x}}{-2}\Big|_0^\infty - \int_0^\infty \frac{e^{-2x}}{-2} dx$

$\Rightarrow (0-0) - \left(\frac{e^{-2x}}{4}\Big|_0^\infty\right) \Rightarrow (0-0) - (0 - 1/4) = 1/4$

$\Rightarrow \int_0^\infty \frac{1}{4} y e^{-3y} dy$, IBP, let $u = y,\ dv = e^{-3y}$, $du = dy,\ v = \frac{e^{-3y}}{-3}$

$\Rightarrow \frac{1}{4}\left[\frac{ye^{-3y}}{-3}\Big|_0^\infty - \int_0^\infty \frac{e^{-3y}}{-3} dy\right] \Rightarrow \frac{1}{4}\left[(0-0) - \left(0 - \frac{e^{-3y}}{9}\Big|_0^\infty\right)\right] = \frac{1}{4} \cdot \frac{1}{9} = \frac{1}{36}$

d.) calculate autocorrelation matrix of the random vector $[X\ Y]^T$

$\Rightarrow R_X = \begin{bmatrix} E[X^2] & E[XY] \\ E[YX] & E[Y^2] \end{bmatrix} = \boxed{\begin{bmatrix} 1/12 & 1/36 \\ 1/36 & 1/27 \end{bmatrix} = R_X}$

e.) calculate variances σ_x^2 & σ_y^2 & covariance C_{XY}

$\Rightarrow$ using expression $var(X) = \sigma_x^2 = E[X^2] - E[X]^2 = E[X^2] - \bar{X}$

$\Rightarrow \sigma_x^2 = \frac{1}{12} - \left(\frac{1}{12}\right)^2 = \frac{12}{144} - \frac{1}{144} = \boxed{\frac{11}{144} = \sigma_x^2}$

Similarly, $\sigma_y^2 = \frac{1}{27} - \left(\frac{1}{18}\right)^2 = \frac{1}{27} - \frac{1}{324} = \boxed{\frac{11}{324} = \sigma_y^2}$

covariance $C_{XY} = E[XY] - \bar{X}\bar{Y} = \frac{1}{36} - \left(\frac{1}{12} \cdot \frac{1}{18}\right) = \frac{1}{36} - \frac{1}{216} = \frac{6}{216} - \frac{1}{216} = \boxed{\frac{5}{216} = C_{XY}}$

f.) calculate autocovariance of random vector $[X\ Y]^T$

$\Rightarrow C_X = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{yx} & \sigma_y^2 \end{bmatrix} = \begin{bmatrix} 11/144 & 5/216 \\ 5/216 & 11/324 \end{bmatrix}$

G.) calculate correlation coefficient between X & Y

$\Rightarrow \rho = \frac{C_{XY}}{\sigma_x \cdot \sigma_y} = \frac{5/216}{\sqrt{\left(\frac{11}{144}\right)\left(\frac{11}{324}\right)}} = \frac{5}{216} \cdot \frac{216}{11} = \boxed{5/11 = \rho}$

2.13 /6 consider equation $Z = X + V$

a.) plot the pdf of $(Z|X)$ as a function of X for $Z = 0.5$

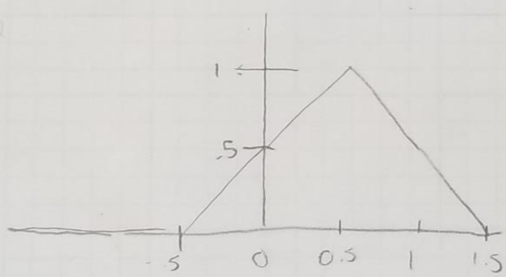
$\Rightarrow Z = 0.5 = X + V$
$\Rightarrow V = 0.5 - X$

using given pdfs

X	$P_v(V)$
$-.5$	0
0	0.5
0.5	1
1	0.5
1.5	0

$X \in [-.5, 1.5]$

$\Rightarrow$



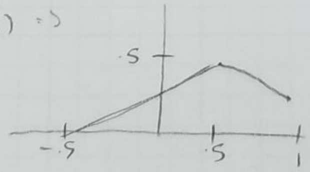
b.) for $Z = 0.5$, find $E(X|Z=0.5) = \int_{-\infty}^{\infty} x P(X|Z=0.5)\,dx$

$\Rightarrow P(X|Z=0.5) = \frac{P(X)P(Z=0.5|X)}{\alpha}$, where α is a constant, where $\alpha = \int_{-\infty}^{\infty} P(x)p(Z=0.5|x)\,dx$

$\Rightarrow P(X|Z=0.5) = \frac{1}{\alpha}\begin{cases} \frac{1}{2}(x+\frac{1}{2}), & x \in [-.5, .5] \\ \frac{1}{2}(-x+3/2), & x \in [.5, 1.5] \\ 0 & \text{otherwise} \end{cases}$

where α is area under $P(X|Z=0.5)$

graphically, $P(X|Z=.5) \Rightarrow$



$\Rightarrow P(X|Z=.5) = 7/16$

$\Rightarrow E(X|Z=0.5) = \frac{16}{7}\int_{-.5}^{.5} x \, \frac{1}{2}(x+\frac{1}{2})\,dx + \frac{16}{7}\int_{.5}^{1} x \, \frac{1}{2}(-x+3/2)\,dx$

$\Rightarrow \frac{16}{7}\int_{-.5}^{.5} \frac{x^2}{2} + \frac{x}{4}\,dx = \left[\left(\frac{x^3}{6} + \frac{x^2}{8}\right)\Big|_{-.5}^{.5}\right]\frac{16}{7} = .041667 \cdot \frac{16}{7}$

$\Rightarrow \frac{16}{7}\int_{.5}^{1} -\frac{x^2}{2} + \frac{3x}{4}\,dx = \frac{16}{7}\left[\frac{-x^3}{6} + \frac{3x^2}{8}\right]\Big|_{.5}^{1} = \frac{16}{7} \cdot .135$

$\Rightarrow \frac{16}{7}(.041667 + .13541) = \boxed{.4017 = E(X|Z=.5)}$

$\boxed{\text{Most Probable val @ } X = .5}$

2.13 con. b cont.)

Median value $\Rightarrow \frac{16}{7}\int_{-\infty}^{m} p(x|z=.5)\,dx = 0.5$

In first part $\frac{16}{7}\int_{-0.5}^{m} \frac{1}{2}(x+\frac{1}{2})\,dx = 0.5$

$$\frac{16}{7}\left[\frac{x^2}{4}+\frac{x}{4}\right]\Big|_{-.5}^{m} = 0.5$$

$$\Rightarrow \frac{m^2}{4}+\frac{m}{4}-\left(\frac{(-0.5)^2}{4}+\frac{-.5}{4}\right)=\frac{7}{16}\cdot .5 \Rightarrow \boxed{m=0.435}$$

Problem 7 consider 2 cont. random variables X & Y, where $y = \ln(x)$, $X > 0$. Derive analytical closed-form expressions for each of the following:

a.) $p(y)$ if $p(x) = \mathcal{U}[a,b]$ (i.e. if X has a uniform pdf for $0 < a \leq x \leq b$)

$\Rightarrow g(x) = \ln(x)$, $p(x) = \frac{1}{b-a} \in x[a,b]$

$h(y) = \text{inverse } g(x) = e^y$

Known that $p_y(y) = p_x(h(y))|h'(y)| = \boxed{\frac{1}{b-a} \cdot e^y = p_y(y)}$

b.) $p_y(y)$ if $p(\frac{1}{x}) = \mathcal{U}[c,d]$ $\Rightarrow p(x) = x \cdot \frac{1}{d-c}$

$\Rightarrow g(x) = \ln(x) \Rightarrow h(y) = e^y$

$p(y) = p_x(h(y))|h'(y)| = \frac{e^y}{d-c} \cdot e^y = \boxed{\frac{e^{2y}}{d-c} = p(y)}$

c.) $p(x)$ if $p(y) = \mathcal{U}[l,m]$, $g(x) = \ln(x)$

$\Rightarrow p(x) = p_y(g(x))|g'(x)| = \boxed{\frac{1}{m-l} \cdot \left|\frac{1}{x}\right| = p(x)}$

d.) $p(x)$ if $p(y) = \mathcal{N}(\mu_y, \sigma_y^2)$, $g(x) = \ln(x)$, $g'(x) = \frac{1}{x}$

for Gaussian, $p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{-(x-\bar{x})^2}{2\sigma^2}\right] = p_y(g(x))|g'(x)|$

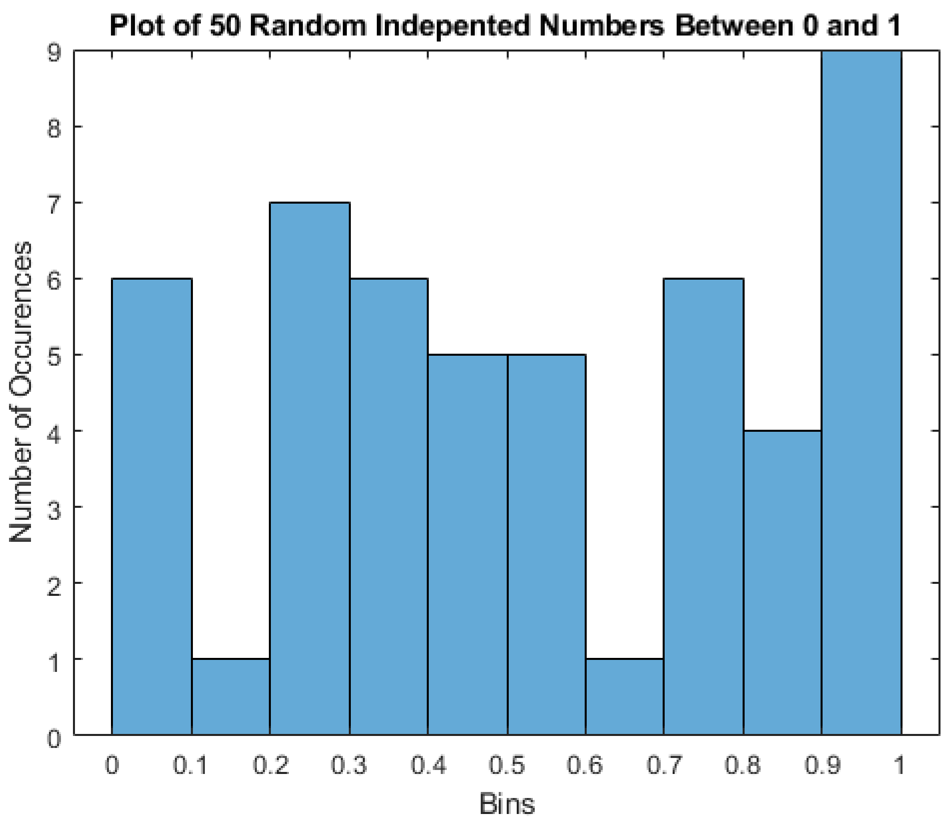
$\Rightarrow \boxed{p(x) = \frac{1}{|x|\sigma_y\sqrt{2\pi}} \exp\left[\frac{-(\ln(x) - \mu_y)^2}{2\sigma_y^2}\right]}$

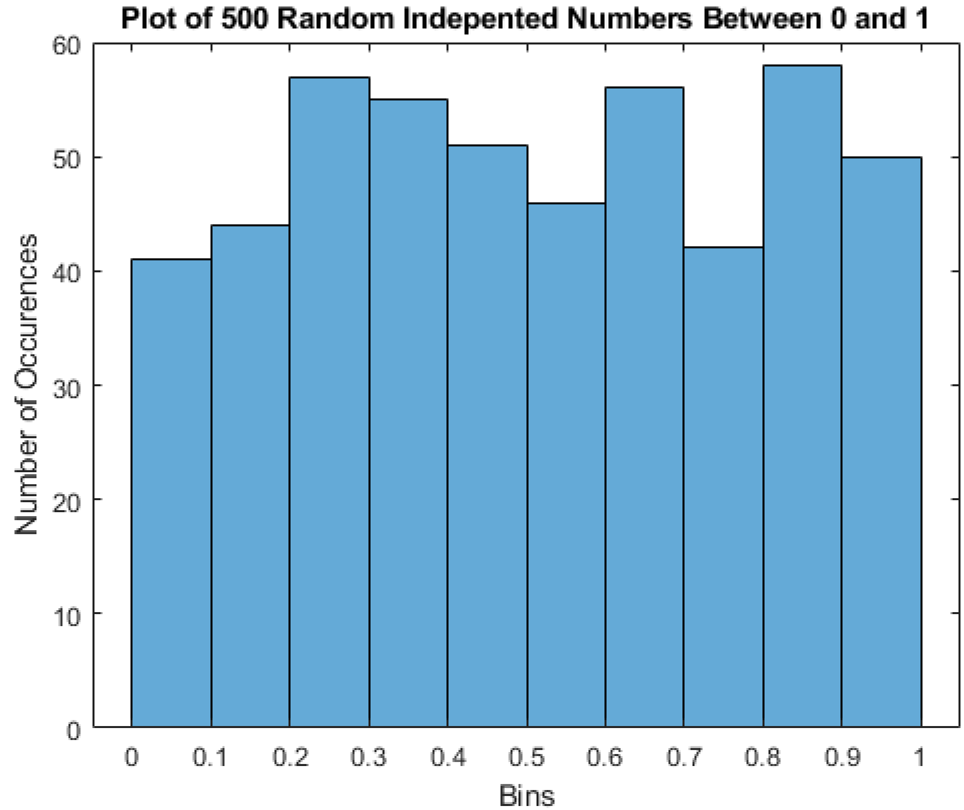
Question 2, 2.16

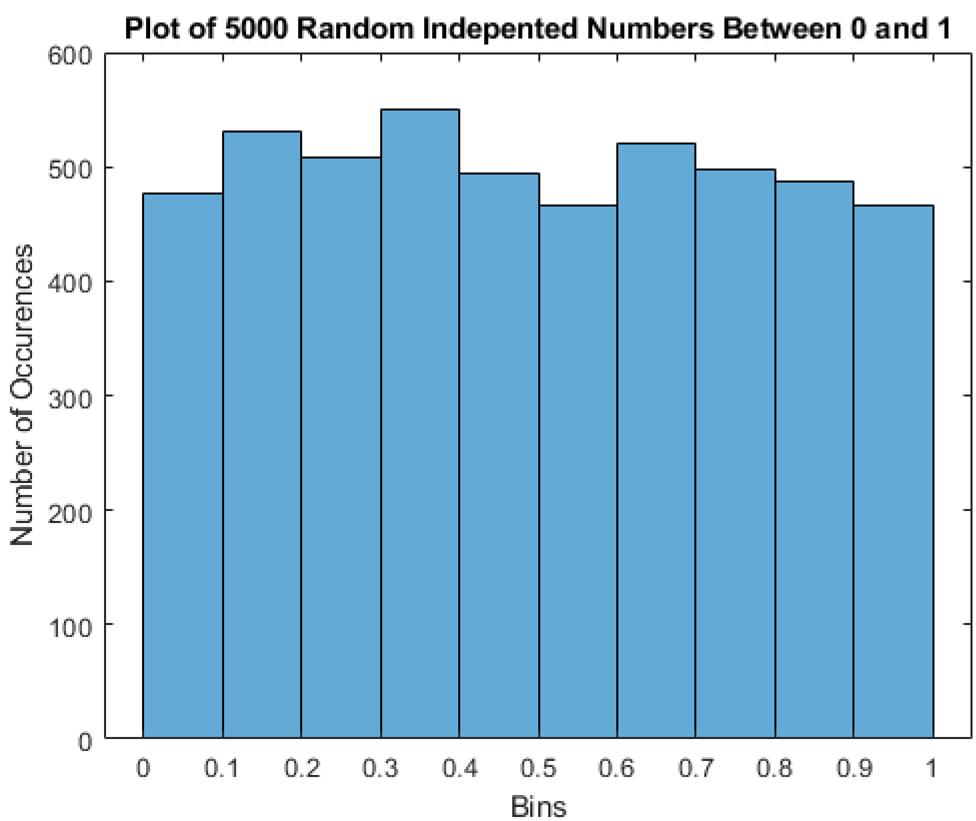
```
clear;close all;clc

nums = rand(50,1); %Generate 50 random numbers between 0 and 1 from
 the uniform dist.
figure(1)
histogram(nums,10)
title('Plot of 50 Random Independented Numbers Between 0 and 1')
xlabel('Bins')
ylabel('Number of Occurences')

mean(nums) % Calculate mean of generated random numbers

std(nums) %Calculate standard deviation of random numbers

nums500 = rand(500,1);
figure(2)
histogram(nums500,10)
title('Plot of 500 Random Independented Numbers Between 0 and 1')
xlabel('Bins')
ylabel('Number of Occurences')

mean(nums500)

std(nums500)

nums5000 = rand(5000,1);
figure(3)
histogram(nums5000,10)
title('Plot of 5000 Random Independented Numbers Between 0 and 1')
xlabel('Bins')
ylabel('Number of Occurences')

mean(nums5000)

std(nums5000)
```

```
ans =

    0.5300


ans =

    0.3009


ans =

    0.5101
```

```
ans =

    0.2845

ans =

    0.4946

ans =

    0.2863
```

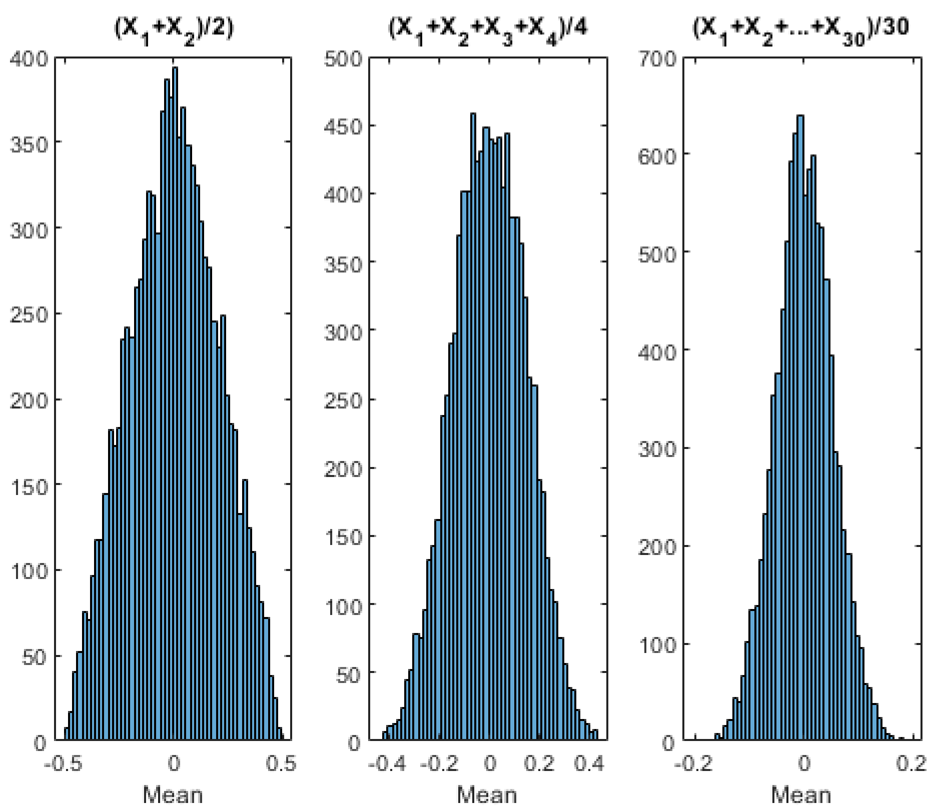



Plot of 500 Random Independed Numbers Between 0 and 1
Number of Occurences
Bins
0
10
20
30
40
50
60
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


Plot of 5000 Random Independed Numbers Between 0 and 1
Number of Occurences
Bins
0
100
200
300
400
500
600
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Problem 2.16

```
clear;close all;clc
for i = 1:10000
    num = -.5 + rand(30,1);
    A(i) = (num(1) + num(2))/2;
    B(i) = (sum(num(1:4)))/4;
    C(i) = (sum(num(1:30)))/30;
end
figure
subplot(1,3,1)
histogram(A,50)
title('(X_1+X_2)/2)')
xlabel('Mean')
subplot(1,3,2)
histogram(B,50)
title('(X_1+X_2+X_3+X_4)/4')
xlabel('Mean')
subplot(1,3,3)
histogram(C,50)
title('(X_1+X_2+...+X_{30})/30')
xlabel('Mean')
```



Published with MATLAB® R2019b