Justin La

# Su24 ECE 131A Project

1a

| Toss | Probability |
|---|---|
| 10 | 0.80 |
| 50 | 0.68 |
| 100 | 0.55 |
| 500 | 0.61 |
| 1000 | 0.60 |

1b.

$S = \{1,3,5\}$

$$P = \frac{1}{5} + \frac{1}{5} + \frac{1}{5} = \frac{3}{5} = 0.6$$

1c.

As the number of tosses increases, the estimated results approach the mathematical analysis.

1d.

$$P = [\frac{2}{7}, \frac{2}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}]$$

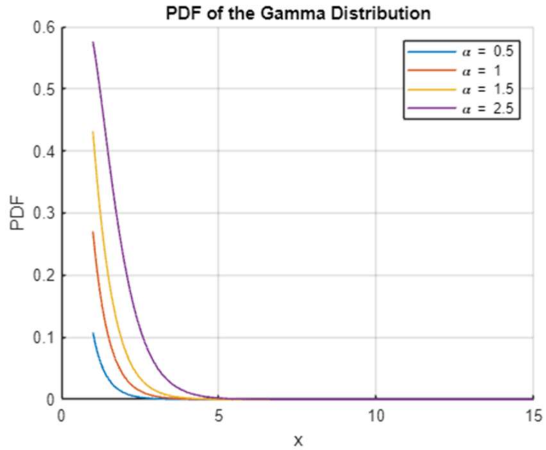
| Toss | Probability |
|---|---|
| 10 | 0.20 |
| 50 | 0.48 |
| 100 | 0.60 |
| 500 | 0.5580 |
| 1000 | 0.5460 |

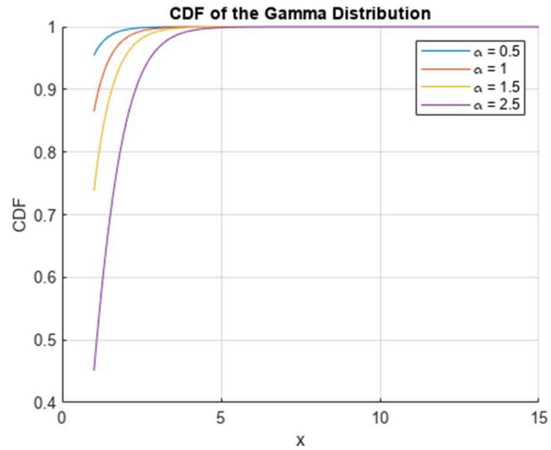
$S = \{1,3,5\}$

$$P = \frac{2}{7} + \frac{1}{7} + \frac{1}{7} = \frac{4}{7} = 0.5714$$

As the number of tosses increases, the estimated results approach the mathematical analysis.

2a.




2b.

2b $F_X(k) = P(X \le k) = 1-(1-p)^k$

$U = F_X(x) = 1-(1-p)^k$

$1-U = (1-p)^x$

$X = \dfrac{\log(1-U)}{\log(1-p)}$ $\quad p=\frac{1}{2}$

$X = \dfrac{\log(1-U)}{\log(\frac{1}{2})} = \dfrac{\log(U)}{\log(2)}$

$E(X) = \frac{1}{p} = \frac{1}{\frac{1}{2}} = 2$

2c

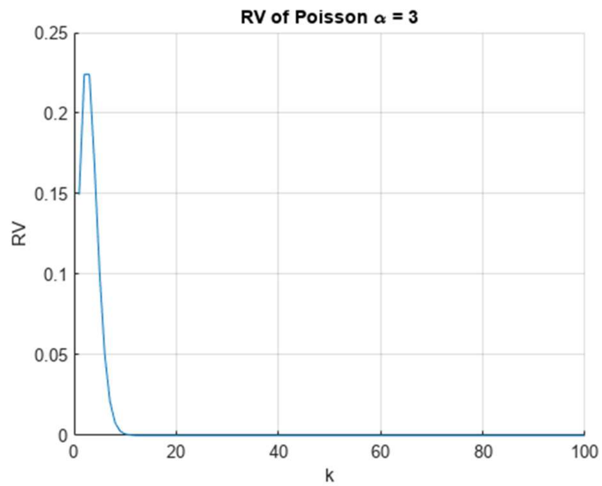
$p(N=k) = \dfrac{\alpha^k}{k!}e^{-\alpha}$

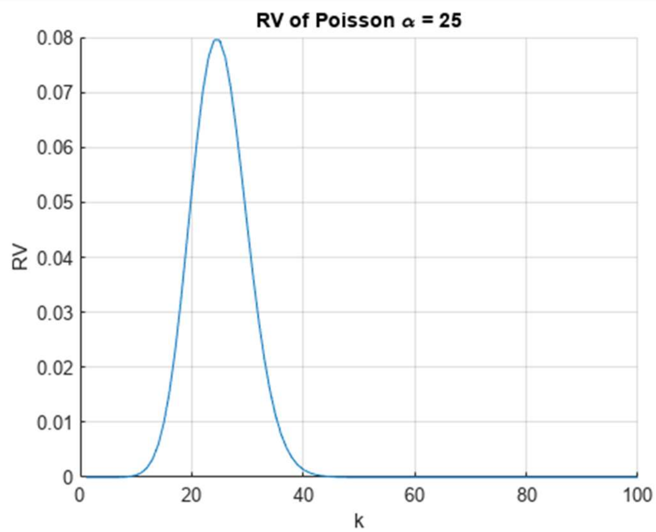
$p(N=0) = \dfrac{\alpha^0}{0!}e^{-\lambda t} = e^{-\lambda t}$ $\quad f(x) = \lambda e^{-\lambda t}$

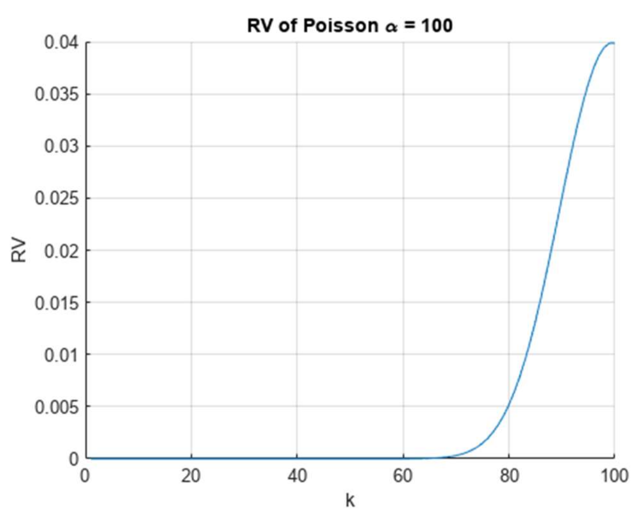
if $\lambda$ is exponential

where $\lambda t = \alpha$

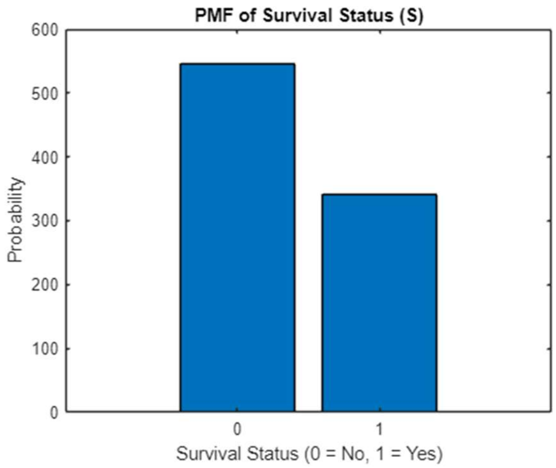
then poisson becomes exponential

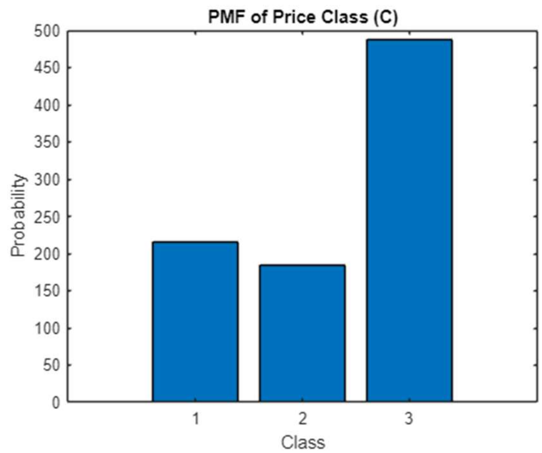

RV of Poisson α = 3
0.25
0.2
0.15
0.1
0.05
0
RV
0
20
40
60
80
100
k

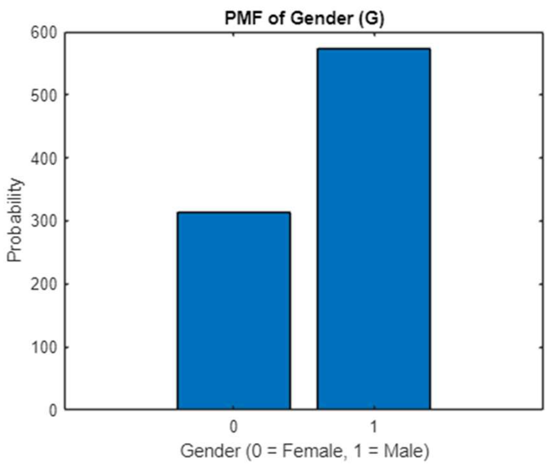

RV of Poisson α = 25
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
RV
0
20
40
60
80
100
k

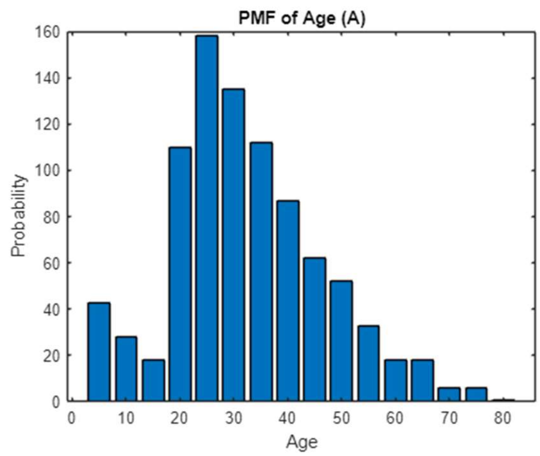

RV of Poisson α = 100
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
RV
0
20
40
60
80
100
k

3a.



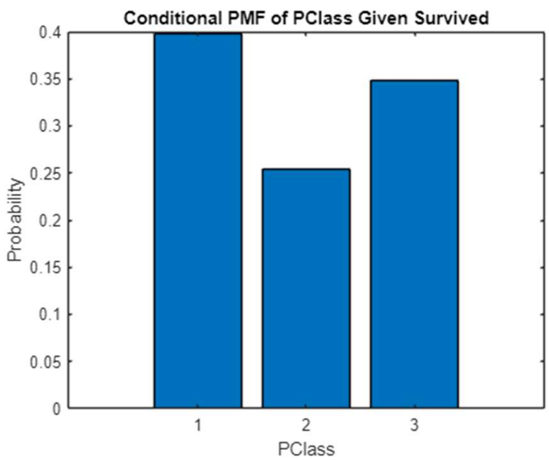



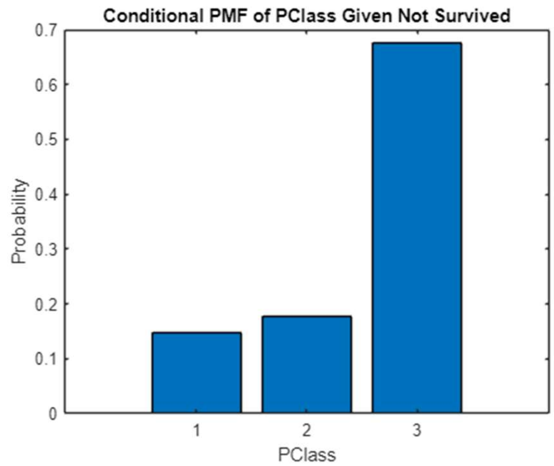

3b.

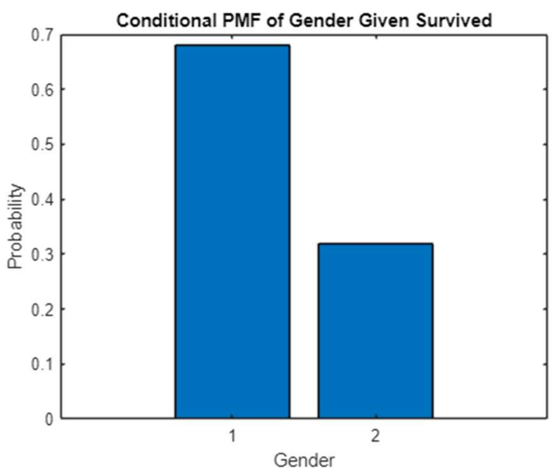
| Condition (Survived) | Probability |
|---|---|
| P(S=1) | 0.3856 |
| P(C=1 \| S=1) | 0.3977 |
| P(C=2 \| S=1) | 0.2544 |
| P(C=3 \| S=1) | 0.3480 |
| P(G=0 \| S=1) | 0.6813 |
| P(G=1 \| S=1) | 0.3187 |
| P(0≤A<5 \| S=1) | 0.0965 |
| P(5≤A<10 \| S=1) | 0.0234 |
| P(10≤A<15 \| S=1) | 0.0322 |
| P(15≤A<20 \| S=1) | 0.1199 |
| P(20≤A<25 \| S=1) | 0.1550 |
| P(25≤A≤<30 \| S=1) | 0.1491 |
| P(30≤A≤<35 \| S=1) | 0.1404 |
| P(35≤A≤<40 \| S=1) | 0.0936 |

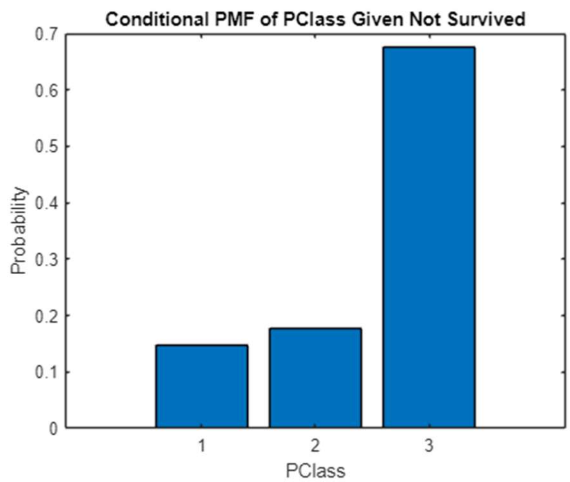
| | |
|---|---|
| P(40≤A≤<45 \| S=1) | 0.0614 |
| P(45≤A≤<50 \| S=1) | 0.0643 |
| P(50≤A≤<55 \| S=1) | 0.0292 |
| P(55≤A≤<60 \| S=1) | 0.0205 |
| P(60≤A≤<65 \| S=1) | 0.0117 |
| P(65≤A≤<70 \| S=1) | 0 |
| P(70≤A≤<75 \| S=1) | 0 |
| P(75≤A≤<80 \| S=1) | 0.0029 |

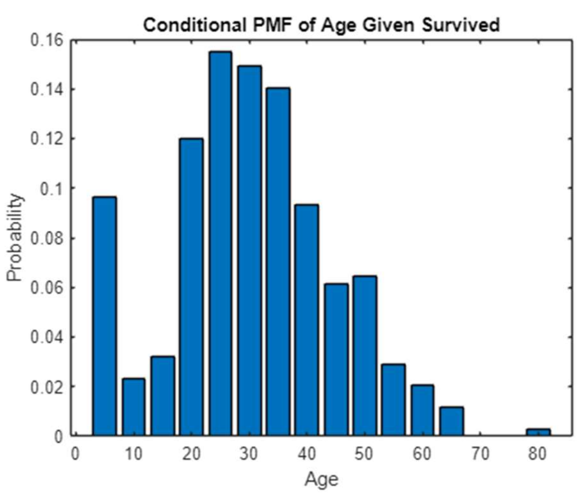
| Condition (Not Survived | Probability |
|---|---|
| P(S=0) | 0.6144 |
| P(C=1 \| S=0) | 0.1468 |
| P(C=2 \| S=0) | 0.1780 |
| P(C=3 \| S=0) | 0.6752 |
| P(G=0 \| S=0) | 0.1486 |
| P(G=1 \| S=0) | 0.8514 |
| P(0≤A<5 \| S=0) | 0.0294 |
| P(5≤A<10 \| S=0) | 0.0294 |
| P(10≤A<15 \| S=0) | 0.0183 |
| P(15≤A<20 \| S=0) | 0.1596 |
| P(20≤A<25 \| S=0) | 0.1945 |
| P(25≤A≤<30 \| S=0) | 0.1706 |
| P(30≤A≤<35 \| S=0) | 0.0954 |
| P(35≤A≤<40 \| S=0) | 0.0954 |
| P(40≤A≤<45 \| S=0) | 0.0679 |
| P(45≤A≤<50 \| S=0) | 0.0477 |
| P(50≤A≤<55 \| S=0) | 0.0294 |
| P(55≤A≤<60 \| S=0) | 0.0239 |
| P(60≤A≤<65 \| S=0) | 0.0220 |
| P(65≤A≤<70 \| S=0) | 0.0092 |
| P(70≤A≤<75 \| S=0) | 0.0073 |
| P(75≤A≤<80 \| S=0) | 0 |


Conditional PMF of PClass Given Survived
Probability
0 0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4
1 2 3
PClass

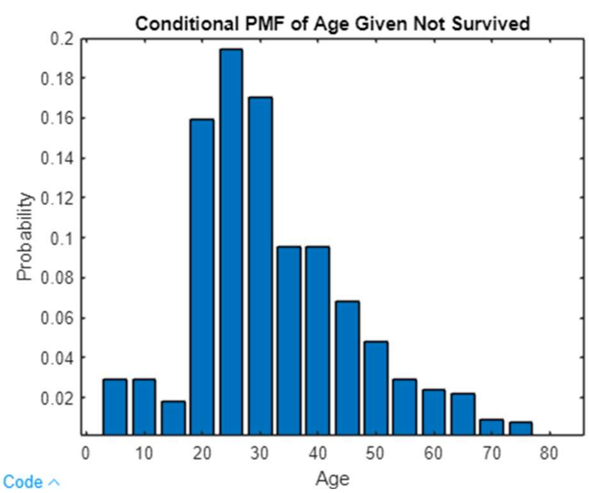

Conditional PMF of PClass Given Not Survived
Probability
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7
1 2 3
PClass


Conditional PMF of Gender Given Survived
Probability
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7
1 2
Gender


Conditional PMF of PClass Given Not Survived
Probability
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7
1 2 3
PClass


Conditional PMF of Age Given Survived
Probability
0 0.02 0.04 0.06 0.08 0.1 0.12 0.14 0.16
0 10 20 30 40 50 60 70 80
Age

Code ^


Conditional PMF of Age Given Not Survived
Probability
0.02 0.04 0.06 0.08 0.1 0.12 0.14 0.16 0.18 0.2
0 10 20 30 40 50 60 70 80
Age

3c.

$$P(S = 1, C = 1, G = 0, A \le 40) = P(C = 1 \mid S = 0) * P(G = 0 \mid S = 0) * P(A \le 40 \mid S = 0)$$
$$= 0.2194$$

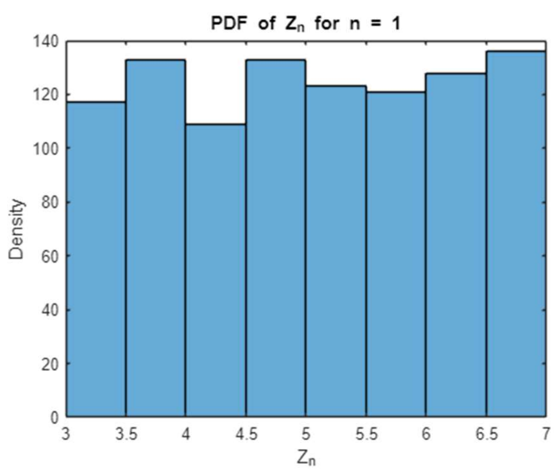
$$P(S = 0, C = 1, G = 0, A \le 40) = P(C = 1 \mid S = 0) * P(G = 0 \mid S = 0) * P(A \le 40 \mid S = 0)$$
$$= 0.0173$$

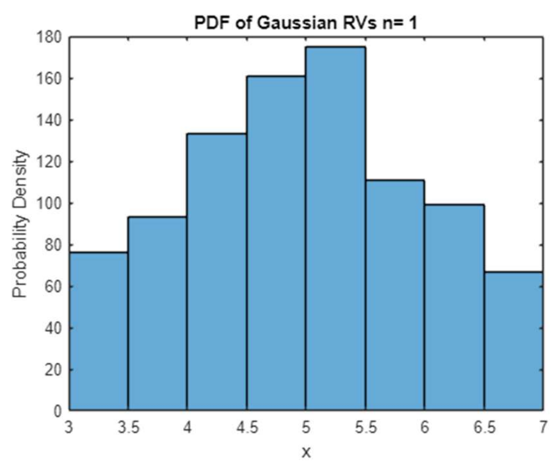
3d.

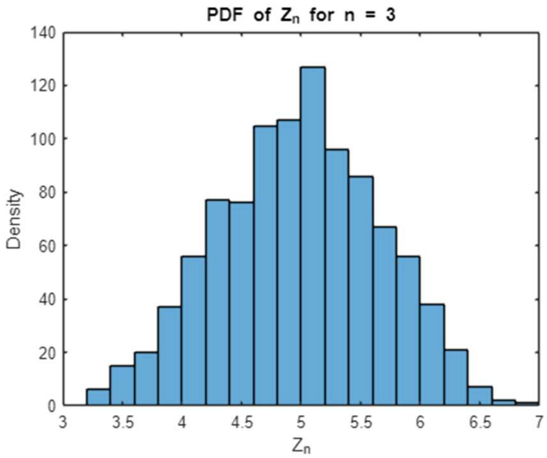
$$P(C = 1, G = 0, A \le 40 \mid \mathrm{S} = 0) = \frac{P(C = 1 \mid S = 0) * P(G = 0 \mid S = 0) * P(A \le 40 \mid S = 0)}{P(S = 0)}$$
$$= 0.0281$$

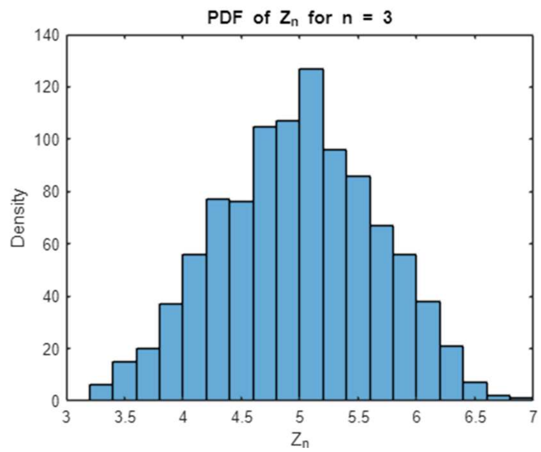
$$P(C = 1, G = 0, A \le 40 \mid \mathrm{S} = 1) = \frac{P(C = 1 \mid S = 1) * P(G = 0 \mid S = 1) * P(A \le 40 \mid S = 1)}{P(S = 1)}$$
$$= 0.5691$$

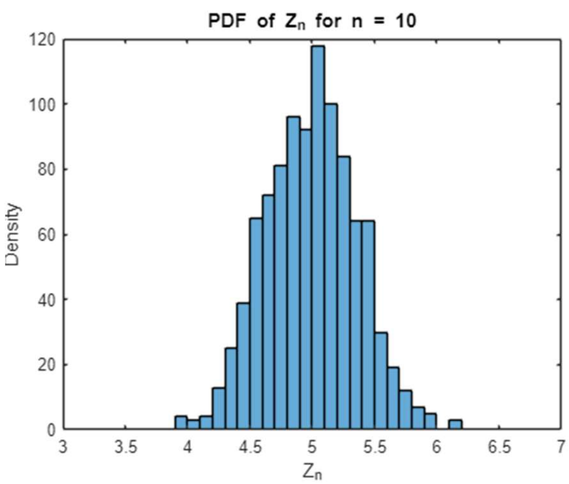
4a.

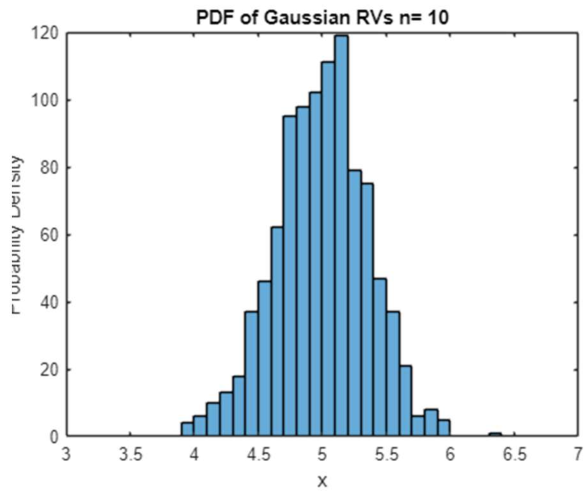


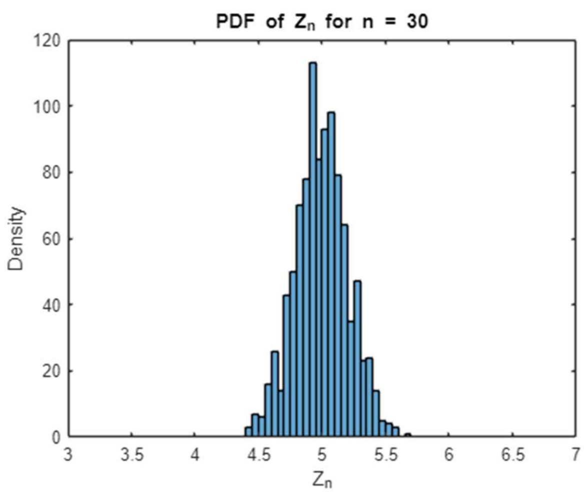



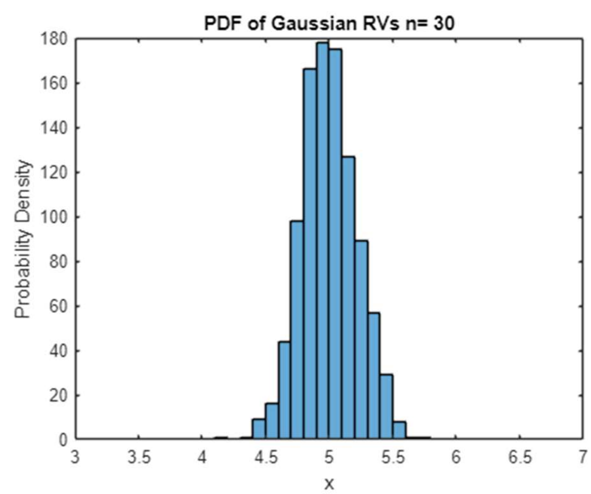

PDF of $Z_n$ for n = 3
Density
$Z_n$

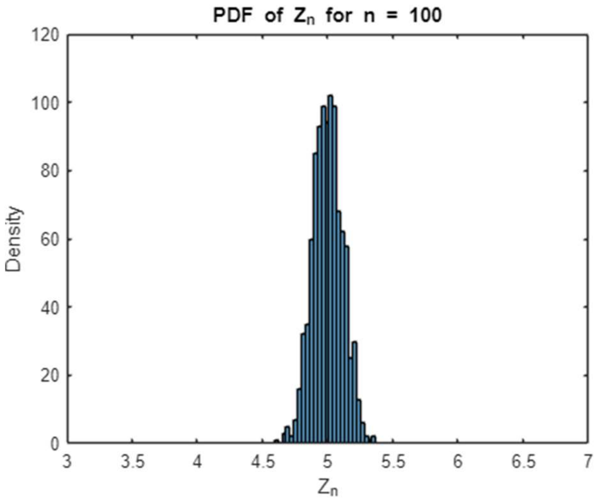

PDF of $Z_n$ for n = 3
Density
$Z_n$

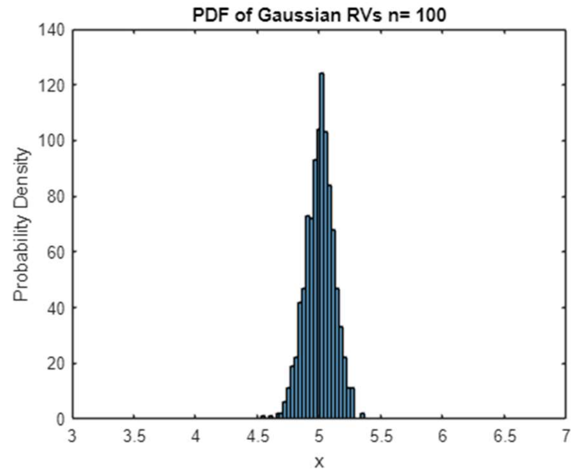

PDF of $Z_n$ for n = 10
Density
$Z_n$


PDF of Gaussian RVs n= 10
x


PDF of $Z_n$ for n = 30
Density
$Z_n$


PDF of Gaussian RVs n= 30
Probability Density
x




4b.

$Z_n = \frac{1}{n}\sum_{i=1}^{n} x_i$

$a = 3$
$b = 7$

$S_n = \sum_{i=1}^{n} x_i$

$E[x_i] = \mu = \frac{a+b}{2} = 5$ $\quad \sigma^2 = VAR(x_i) = \frac{(b-a)^2}{12} = \frac{16}{12} = \frac{4}{3}$

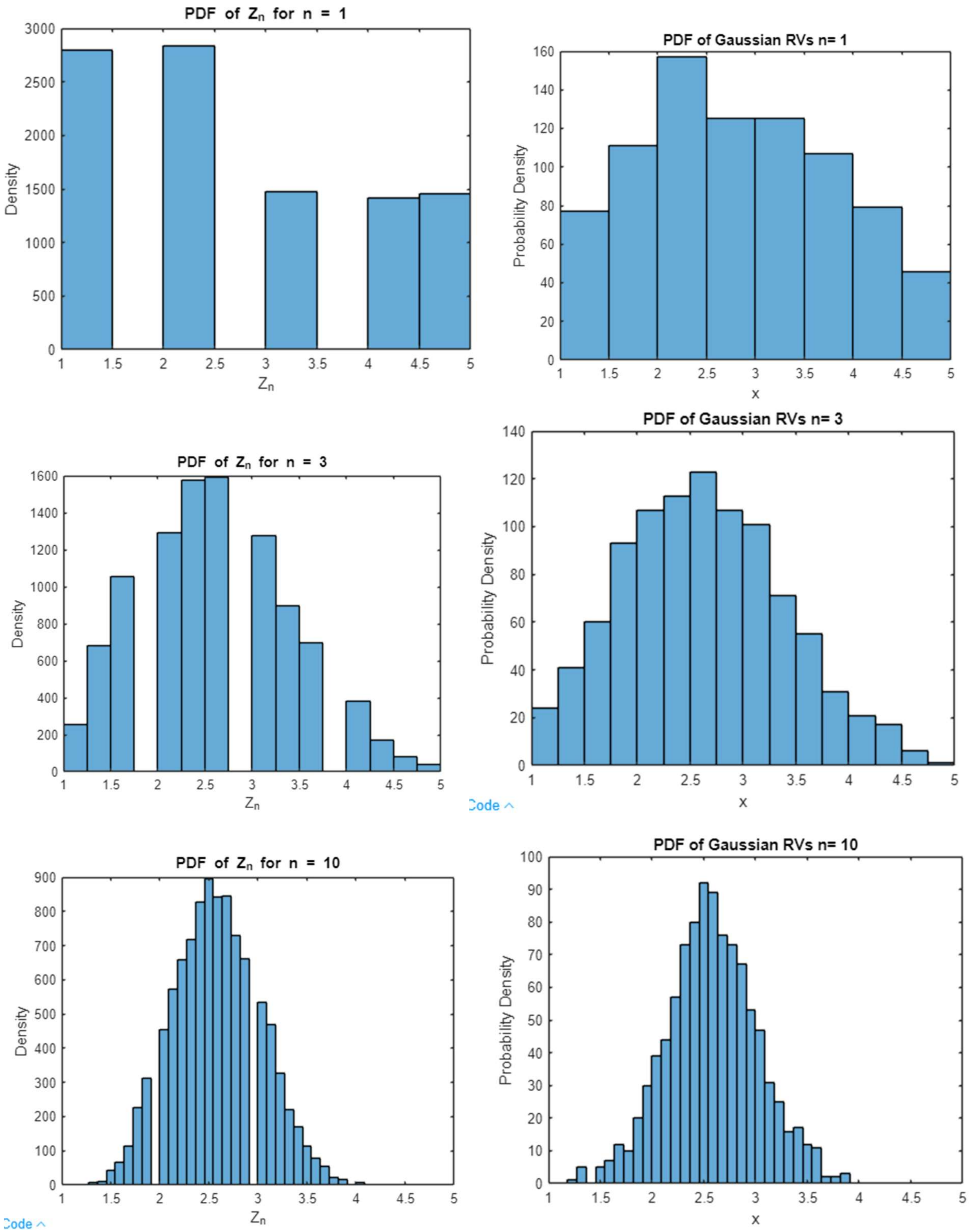
$E[S_n] = n\mu$ $\quad \cancel{E[Z_n] = \frac{1}{\sigma\sqrt{n}} E[S_n - n\mu]}$

$E[S_n] = n\mu \quad VAR[S_n] = n\sigma^2$

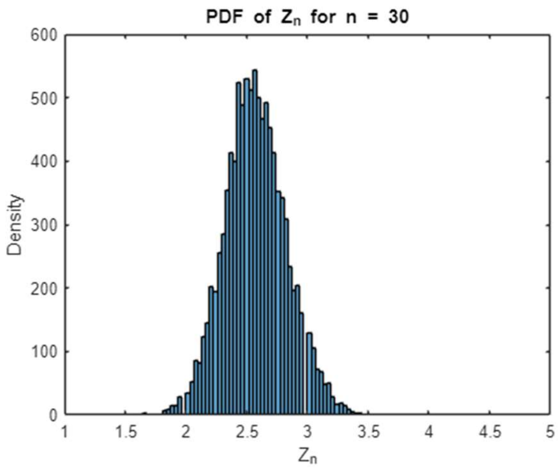
$E[Z_n] = E[\frac{1}{n}S_n] = \frac{1}{n}n\mu = \mu$

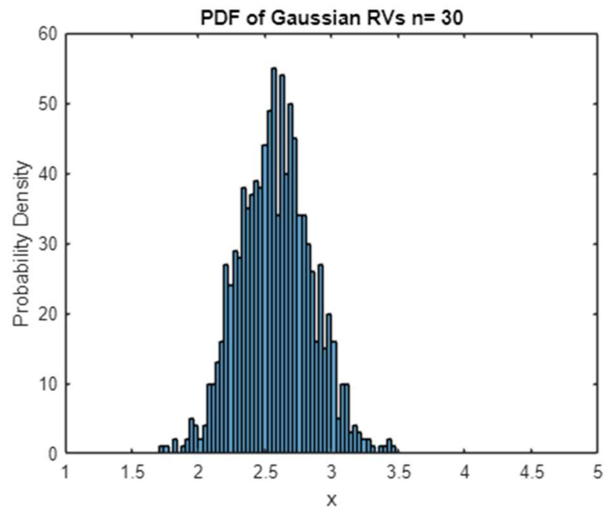
$VAR[\frac{1}{n}S_n] = \frac{1}{n^2}VAR[S_n] = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n} \quad \frac{4/3}{n}$

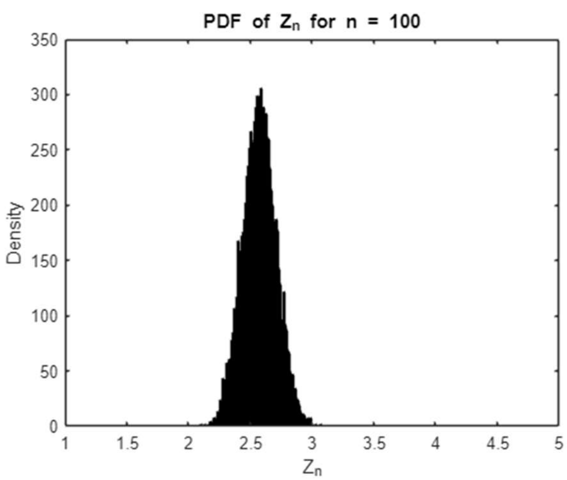
4d.

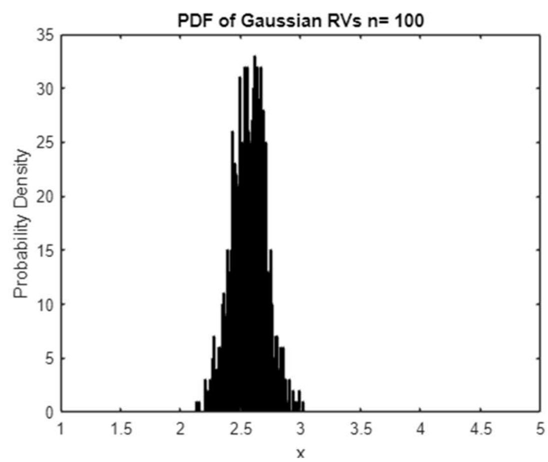


PDF of Zn for n = 1
Density
Zn
PDF of Gaussian RVs n= 1
Probability Density
x
PDF of Zn for n = 3
Density
Zn
PDF of Gaussian RVs n= 3
Probability Density
x
PDF of Zn for n = 10
Density
Zn
PDF of Gaussian RVs n= 10
Probability Density
x




$P = [2/7, 2/7, 1/7, 1/7, 1/7]$

$E[x] = \sum_{i=1}^{n} x_i P(X = x_i)$

$= 1 \cdot \frac{2}{7} + 2 \cdot \frac{2}{7} + 3 \cdot \frac{1}{7} + 4 \cdot \frac{1}{7} + 5 \cdot \frac{1}{7} = \frac{18}{7} = 2.57$

$\cancel{VAR}(x) = E(x^2) = \sum_{i=1}^{n} x^2 P(x = x_i)$

$E[x^2] = 1^2 \frac{2}{7} + 2^2 \frac{2}{7} + 3^2 \frac{1}{7} + 4^2 \frac{1}{7} + 5^2 \frac{1}{7}$

$E(x^2) = \frac{60}{7} = 8.57$

$VAR(x) = E(x^2) - (E(x))^2 = 8.57 - 2.57^2$

$= 1.96$

$E[Z_n] = \mu = 2.57$

$VAR[Z_n] = \frac{\sigma^2}{n} = \frac{1.96}{n}$ $\qquad \sigma = \sqrt{\frac{1.96}{n}}$

Appendix (Matlab Code)

```
%%1

t = 10;

num = randi([1, 5], 1, t);

prob_odd_count_10 = sum(mod(num,2))/t


t = 50;

num = randi([1, 5], 1, t);

prob_odd_count_50 = sum(mod(num,2))/t


t = 100;

num = randi([1, 5], 1, t);

prob_odd_count_100 = sum(mod(num,2))/t


t = 500;

num = randi([1, 5], 1, t);

prob_odd_count_500 = sum(mod(num,2))/t


t = 1000;

num = randi([1, 5], 1, t);

prob_odd_count_1000 = sum(mod(num,2))/t


%1,3,5

math_analy_prob_odd = 3/5


%d
```

```
P = [2/7, 2/7, 1/7, 1/7, 1/7];

t = 10;
outcomes = randsample(1:5, t, true, P);
d_prob_odd_count_10 = sum(mod(outcomes,2))/t

t = 50;
outcomes = randsample(1:5, t, true, P);
d_prob_odd_count_50 = sum(mod(outcomes,2))/t

t = 100;
outcomes = randsample(1:5, t, true, P);
d_prob_odd_count_100 = sum(mod(outcomes,2))/t

t = 500;
outcomes = randsample(1:5, t, true, P);
d_prob_odd_count_500 = sum(mod(outcomes,2))/t

t = 1000;
outcomes = randsample(1:5, t, true, P);
d_prob_odd_count_1000 = sum(mod(outcomes,2))/t

%P = [2/7, 2/7, 1/7, 1/7, 1/7];
% 1,2,3
d_math_analy_prob_odd = (2+1+1)/7
```

```
%%2

x = linspace(1, 15, 1000);

figure;

hold on;

title('PDF of the Gamma Distribution');

xlabel('x');

ylabel('PDF');

grid on;


lambda = 0.5;

alpha_values = [0.5, 1, 1.5, 2.5];



alpha = alpha_values(1);

pdf = gampdf(x, alpha, lambda); % Compute PDF

plot(x, pdf, 'DisplayName', ['\alpha = ', num2str(alpha)]);

alpha = alpha_values(2);

pdf = gampdf(x, alpha, lambda); % Compute PDF

plot(x, pdf, 'DisplayName', ['\alpha = ', num2str(alpha)]);

alpha = alpha_values(3);

pdf = gampdf(x, alpha, lambda); % Compute PDF

plot(x, pdf, 'DisplayName', ['\alpha = ', num2str(alpha)]);

alpha = alpha_values(4);

pdf = gampdf(x, alpha, lambda); % Compute PDF

plot(x, pdf, 'DisplayName', ['\alpha = ', num2str(alpha)]);

legend show;
```

```
hold off;


figure;
hold on;
title('CDF of the Gamma Distribution');
xlabel('x');
ylabel('CDF');
grid on;
alpha = alpha_values(1);
cdf = gamcdf(x, alpha, lambda); % Compute CDF
plot(x, cdf, 'DisplayName', ['\alpha = ', num2str(alpha)]);
alpha = alpha_values(2);
cdf = gamcdf(x, alpha, lambda); % Compute CDF
plot(x, cdf, 'DisplayName', ['\alpha = ', num2str(alpha)]);
alpha = alpha_values(3);
cdf = gamcdf(x, alpha, lambda); % Compute CDF
plot(x, cdf, 'DisplayName', ['\alpha = ', num2str(alpha)]);
alpha = alpha_values(4);
cdf = gamcdf(x, alpha, lambda); % Compute CDF
plot(x, cdf, 'DisplayName', ['\alpha = ', num2str(alpha)]);
legend show;
hold off;


%c
t = 100
k = linspace(1, t, t);
```

```
alpha_values = [3, 25, 100];

figure;
hold on;
title('RV of Poisson \alpha = 3');
xlabel('k');
ylabel('RV');
grid on;
alpha = alpha_values(1);
Poisson_RV = (alpha.^k).*exp(-alpha)./factorial(k);
plot(k, Poisson_RV, 'DisplayName', ['\alpha = ', num2str(alpha)]);
hold off;

figure;
hold on;
title('RV of Poisson \alpha = 25');
xlabel('k');
ylabel('RV');
grid on;
alpha = alpha_values(2);
Poisson_RV = (alpha.^k).*exp(-alpha)./factorial(k);
plot(k, Poisson_RV, 'DisplayName', ['\alpha = ', num2str(alpha)]);
hold off;

figure;
hold on;
```

```
title('RV of Poisson \alpha = 100');

xlabel('k');

ylabel('RV');

grid on;

alpha = alpha_values(3);

Poisson_RV = (alpha.^k).*exp(-alpha)./factorial(k);

plot(k, Poisson_RV, 'DisplayName', ['\alpha = ', num2str(alpha)]);

hold off;


%3

data = readtable('modified_titanic.xlsx');

S = data.Survived;

C = data.Pclass;

G = data.Sex;

A = data.Age;


n=887;


pmf_S = histcounts(S);

figure;

x = 0:1:1;

bar(x,pmf_S);

title('PMF of Survival Status (S)');

xlabel('Survival Status (0 = No, 1 = Yes)');

ylabel('Probability');
```

```
pmf_C = histcounts(C);

figure;

bar(pmf_C);

x = 1:3:3;

title('PMF of Price Class (C)');

xlabel('Class');

ylabel('Probability');


pmf_G = histcounts(G);

figure;

x = 0:1:1;

bar(x,pmf_G);

title('PMF of Gender (G)');

xlabel('Gender (0 = Female, 1 = Male)');

ylabel('Probability');


[pmf_A,edges] = histcounts(A);

figure;

bar(edges(2:length(edges)),pmf_A);

title('PMF of Age (A)');

xlabel('Age');

ylabel('Probability');


%Survived

SCount = 0;

for i = 1:n
```

```
    if data.Survived(i) == 1
        SCount = SCount + 1;
    end
end
P_SCount = SCount/n

PClass1_Count=0;
for i = 1:n
    if data.Survived(i) == 1 && data.Pclass(i) == 1
        PClass1_Count = PClass1_Count + 1;
    end
end
P_PClass_Count(1) = PClass1_Count/SCount

PClass2_Count=0;
for i = 1:n
    if data.Survived(i) == 1 && data.Pclass(i) == 2
        PClass2_Count = PClass2_Count + 1;
    end
end
P_PClass_Count(2) = PClass2_Count/SCount

PClass3_Count=0;
for i = 1:n
    if data.Survived(i) == 1 && data.Pclass(i) == 3
        PClass3_Count = PClass3_Count + 1;
```

```
  end
end
P_PClass_Count(3) = PClass3_Count/SCount

GCount=0;
for i = 1:n
  if data.Survived(i) == 1 && data.Sex(i) == 0
    GCount = GCount + 1;
  end
end
P_GCount(1) = GCount/SCount

GCount=0;
for i = 1:n
  if data.Survived(i) == 1 && data.Sex(i) == 1
    GCount = GCount + 1;
  end
end
P_GCount(2) = GCount/SCount



max_age = 80;
age_bin = 5;
P_ACount(max_age/age_bin) = 0;
for j = 1:max_age/age_bin
```

```
    ACount = 0;
    min_age_bin = j*5-5;
    max_age_bin = j*5;
    for i = 1:n
        if data.Survived(i) == 1 && data.Age(i) > min_age_bin && data.Age(i) <= max_age_bin
            ACount = ACount + 1;
        end
    end
    P_ACount(j) = ACount/SCount;
end
P_ACount

%Not Survived
SCount_n = 0;
for i = 1:n
    if data.Survived(i) == 0
        SCount_n = SCount_n + 1;
    end
end
P_SCount_n = SCount_n/n

PClass1_Count_n=0;
for i = 1:n
    if data.Survived(i) == 0 && data.Pclass(i) == 1
        PClass1_Count_n = PClass1_Count_n + 1;
    end
```

```
end
P_PClass_Count_n(1) = PClass1_Count_n/SCount_n

PClass2_Count_n=0;
for i = 1:n
   if data.Survived(i) == 0 && data.Pclass(i) == 2
      PClass2_Count_n = PClass2_Count_n + 1;
   end
end
P_PClass_Count_n(2) = PClass2_Count_n/SCount_n

PClass3_Count_n=0;
for i = 1:n
   if data.Survived(i) == 0 && data.Pclass(i) == 3
      PClass3_Count_n = PClass3_Count_n + 1;
   end
end
P_PClass_Count_n(3) = PClass3_Count_n/SCount_n

GCount_n=0;
for i = 1:n
   if data.Survived(i) == 0 && data.Sex(i) == 0
      GCount_n = GCount_n + 1;
   end
end
P_GCount_n(1) = GCount_n/SCount_n
```

```
GCount_n=0;
for i = 1:n
  if data.Survived(i) == 0 && data.Sex(i) == 1
    GCount_n = GCount_n + 1;
  end
end
P_GCount_n(2) = GCount_n/SCount_n



P_ACount_n(max_age/age_bin) = 0;
max_age = 80;
age_bin = 5;
for j = 1:max_age/age_bin
  ACount_n = 0;
  min_age_bin = j*5-5;
  max_age_bin = j*5;
  for i = 1:n
    if data.Survived(i) == 0 && data.Age(i) > min_age_bin && data.Age(i) <= max_age_bin
      ACount_n = ACount_n + 1;
    end
  end
  P_ACount_n(j) = ACount_n/SCount_n;
end
P_ACount_n
```

```
figure;
x = 1:1:3;
bar(x,P_PClass_Count);
title('Conditional PMF of PClass Given Survived');
xlabel('PClass');
ylabel('Probability');

figure;
x = 1:1:3;
bar(x,P_PClass_Count_n);
title('Conditional PMF of PClass Given Not Survived');
xlabel('PClass');
ylabel('Probability');

figure;
x = 1:1:2;
bar(x,P_GCount);
title('Conditional PMF of Gender Given Survived');
xlabel('Gender');
ylabel('Probability');

figure;
x = 1:1:2;
bar(x,P_GCount_n);
title('Conditional PMF of Gender Given Not Survived');
xlabel('Gender');
```

```
ylabel('Probability');

figure;
bar(edges(2:length(edges)),P_ACount);
title('Conditional PMF of Age Given Survived');
xlabel('Age');
ylabel('Probability');

figure;
bar(edges(2:length(edges)),P_ACount_n);
title('Conditional PMF of Age Given Not Survived');
xlabel('Age');
ylabel('Probability');

%c
P_S1_Alteq40 = sum(P_ACount(1:8))
P_S1_C1 = P_PClass_Count(1)
P_S1_G0 = P_GCount(1) %Female
P_S1_C1_G0_Alteq40 = P_S1_C1*P_S1_G0*P_S1_Alteq40

P_S0_Alteq40 = sum(P_ACount_n(1:8))
P_S0_C1 = P_PClass_Count_n(1)
P_S0_G0 = P_GCount_n(1) %Female
P_S0_C1_G0_Alteq40 = P_S0_C1*P_S0_G0*P_S0_Alteq40

%d
```

```
P_S1_given_C1_G0_Alteq40 = P_S1_C1_G0_Alteq40/P_SCount

P_S0_given_C1_G0_Alteq40 = P_S0_C1_G0_Alteq40/P_SCount_n



%Q4a

n_values = [1,3,10,30,100];

samples = 1000;

Zn(samples) = 0;

for k = 1:length(n_values)

  n = n_values(k);

  Zn(samples) = 0;

  for i = 1:samples

    Xi = 3 + 4*rand(n, 1);

    Zn(i) = 1/n*sum(Xi);

  end

  figure;

  histogram(Zn);

  title(['PDF of Z_n for n = ', num2str(n)]);

  xlabel('Z_n');

  ylabel('Density');

  xlim([3 7]);

end


%b

VAR(length(n_values))= 0;

for k = 1:length(n_values)
```

```
    n = n_values(k);
    VAR(k) = 1.33/n;
end

%c
mu = 5;
samples = 1000;
for k = 1:length(n_values)
    n = n_values(k);
    sigma = sqrt(1.33/n);
    X = mu + sigma * randn(samples, 1);
    figure;
    histogram(X);
    title(['PDF of Gaussian RVs n= ', num2str(n)]);
    xlabel('x');
    ylabel('Probability Density');
    xlim([3 7]);
end

%Q4d - redo abc, fair 5-sided die that is described in Problem 1(d).
P = [2/7, 2/7, 1/7, 1/7, 1/7];

n_values = [1,3,10,30,100];
samples = 10000;
Zn(samples) = 0;
for k = 1:length(n_values)
```

```
  n = n_values(k);
  Zn(samples) = 0;
  for i = 1:samples
    Xi = randsample(1:5, n, true, P);
    Zn(i) = 1/n*sum(Xi);
  end
  figure;
  histogram(Zn, 'BinWidth', 1/(n+1));
  title(['PDF of Z_n for n = ', num2str(n)]);
  xlabel('Z_n');
  ylabel('Density');
  xlim([1 5]);
end

%b
VAR(length(n_values))= 0;
for k = 1:length(n_values)
  n = n_values(k);
  VAR(k) = 1.96/n;
end

%c
mu = 18/7;
samples = 1000;
for k = 1:length(n_values)
  n = n_values(k);
```

```
    sigma = sqrt(96/(49*n));
    X = mu + sigma * randn(samples, 1);
    figure;
    histogram(X, 'BinWidth', 1/(n+1));
    title(['PDF of Gaussian RVs n= ', num2str(n)]);
    xlabel('x');
    ylabel('Probability Density');
    xlim([1 5]);
end
```