Justin La

Su24 ECE 131A Project

1a.

A five-sided dice will have a probability of 1/5 for each side.

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

1b.

3 of the 5 numbers are odd. Adding the probability of the odd numbers results in the probability of obtaining an odd number.

1c.

As the number of tosses increases, the estimated results approach the mathematical analysis. This demonstrates the Law of Large numbers.

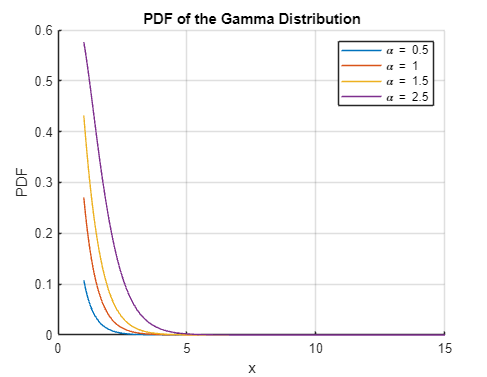
1d.

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

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

2a.

The gamma distribution is dependent on α, shape parameter, and λ, rate parameter.

A graph with colored lines and numbers

Description automatically generated

2b.

By setting the CDF of the geometric distribution equal to a variable and solving for x, we obtain the transformation. Plugging the provided parameter, ½ , we obtain the average number of comparisons, or the expected value/mean.

A paper with writing on it

Description automatically generated

2c

Setting k=0, the poisson random variable will become exponentially distributed.

A piece of paper with writing on it

Description automatically generated

Different α parameters will offset the mean further down.

A graph of a line

Description automatically generated

A graph of a function

Description automatically generated

A graph with a line

Description automatically generated

3a.

Each data set was plotted out. The age was parsed into bins of 5 years.

A graph of a blue bar

Description automatically generatedA graph of a number of blue rectangular bars

Description automatically generated

A graph with blue rectangular bars

Description automatically generated A graph of age and age

Description automatically generated with medium confidence

3b.

This is a list of all the probabilites for each condition given survived and not survived.

|  |  |
| --- | --- |
| 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 |

A graph of blue rectangular bars

Description automatically generatedA graph with blue rectangular bars

Description automatically generated

A graph of a bar graph

Description automatically generated with medium confidenceA graph of a bar graph

Description automatically generated

A graph of age and age

Description automatically generatedA graph of age and age

Description automatically generated

3c.

Using this given assumption, the probability of the union of the conditions, class, gender and age, given survived or not survived are just the individual conditions multiplied by each of the conditions given survived or not survived separately. This would assume that the conditions are all independent.

A mathematical equation with numbers

Description automatically generated

3d.

Using the results from 3c and bayes rule, we can predict whether a female whose age is under 40 and who is in first class will survive or not.

4a.

With a uniform continuous random variable taking values in the interval (3,7), the mean and variance are calculated below.

As more samples are used to calculate Zn, the distribution of Zn becomes more gaussian.

A graph of a graph

Description automatically generated A graph of a number of blue bars

Description automatically generated with medium confidence

A graph of a diagram

Description automatically generated A graph of a diagram

Description automatically generated

A graph of a diagram

Description automatically generatedA graph of a diagram

Description automatically generated

A graph of a graph of a graph

Description automatically generated A graph of a graph of a number of objects

Description automatically generated with medium confidence

A graph of a graph of a graph

Description automatically generated A graph of a tall tower with Burj Khalifa in the background

Description automatically generated

4b.

A piece of paper with writing on it

Description automatically generated

4d.

A graph of a graph

Description automatically generated A graph of a diagram

Description automatically generatedA graph of a diagram

Description automatically generated A graph of a diagram

Description automatically generated

A graph of a diagram

Description automatically generated A graph of a diagram

Description automatically generated

A graph of a graph of a number of lines with Burj Khalifa in the background

Description automatically generated with medium confidence A graph of a graph of a diagram with Burj Khalifa in the background

Description automatically generated with medium confidence

A graph of a graph of a graph

Description automatically generated A graph of a black column

Description automatically generated

A piece of paper with writing on it

Description automatically generated

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