Pay Me My Money Down

Math Mini Project 2 - Semester 2

Rishi Laddha

Note - Part of this project (Section - 1 (Nuts and Bolts), Question - 3) was done in collaboration with Yash Sarvesh Srivastava. This note is to give due credit to his contribution.

---

1. Monte Carlo Basics

Theoretical calculation of expectation for a compound Poisson random variable.
Answer: Let Y be the compound Poisson random variable

$$
\begin{aligned}
E(Y) &= \sum_{n=0}^{\infty} E(Y|N=n)P(N=n) \\
&\implies \sum_{n=0}^{\infty} nE(X)P(N=n) \\
&\implies \mu_X \sum_{n=0}^{\infty} nP(N=n) \\
&\implies E(Y) = \mu_X \, \mu_N
\end{aligned}
$$

Magician and Beer Problem

Answer:

For the case when the magician works for 3 hours a day:

```
% Monte Carlo Simulation of Compound Poisson process
clc
clear all
format long
```

```
time = 3;

lambda = 5;
N = 10000; % Number of Monte Carlo Steps
beer_price = 350;
beer = zeros(N,1);
for i=1:N
    n = poissrnd(lambda*time); %total no of people who are coming
    if n > 0
        coins = zeros(n,1);
        for j=1:n
            U = rand(1);
            if U<=0.4
                coins(j) = 5;
            elseif U>=0.8
                coins(j) = 20;
            else
                coins(j) = 10;
            end
        end
    end
    if sum(coins)>=beer_price
        beer(i) = 1;
    else
        beer(i) = 0;
    end
end
l_hat = mean(beer)% l_hat = P(X3 >= beer_price)
```

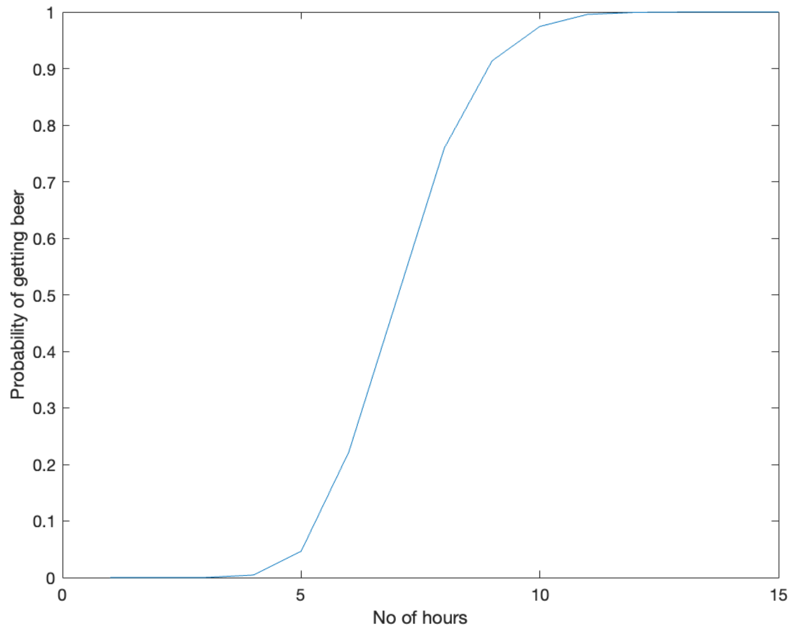
Output: 2.000000000000000e-04. This signifies that the probability of the magician being able to afford the beer after 3 hours of work is **very less**.

Plotting probability of purchasing beer vs number of hours worked:

```
% Monte Carlo Simulation of Compound Poisson process
clc
clear all
format long

arr=[];
time=[1 2 3 4 5 6 7 8 9 10 11 12 13 14 15];

for t = 1:15
lambda = 5;
N = 10^3; % Number of Monte Carlo Steps
beer_price = 350;
beer = zeros(N,1);
for i=1:N
    n = poissrnd(lambda*t); %total no of people who are coming
    if n > 0
        coins = zeros(n,1);
        for j=1:n
            U = rand(1);
            if U<=0.4
                coins(j) = 5;
            elseif U>=0.8
                coins(j) = 20;
            else
                coins(j) = 10;
```

```
            end
        end
    end
    if sum(coins)>=beer_price
        beer(i) = 1;
    else
        beer(i) = 0;
    end
end
l_hat = mean(beer)% l_hat = P(X3 >= beer_price)
arr = [arr l_hat]

end
% relErr_hat = std(beer) / (l_hat * sqrt(N)); % relative error of l_hat by crude

disp arr_plot;
plot(arr,time);
xlabel("No of hours");
ylabel("probability of getting beer");
```

Graph:



The probability of being able to afford the beer increases with an increase in the number of hours that the magician performs. Here is a more analytical proof:

Previously, we arrived at $E(Y) = \mu_X \times \mu_N$. Here, $E(Y)$ corresponds to the expected amount of money that the magician earns at the end of his performance. As seen in the relation, $E(Y)$ is directly related to μ_N. μ_N represents the mean amount of time that the magician performs for.

Let the random variable A correspond to the number of people donating to the magician. Then,

$A \sim Poiss(\lambda \times t)$, where t is the number of hours that the magician performs for. As μ_N increases, t increases. An increase in t corresponds to an increase in the number of people donating to the magician, which leads to higher earnings. As the earnings increase, the chance of the magician being able to afford the beer increases.

---

2. Insurance Problem

Section 1 - The Nuts and Bolts

a. Question 1 - Identify the distribution of Y_j.
Answer - Compound Poisson Distribution. The primary distribution is **Poisson** (for N_j), after which we look at claims (X_i) which are distributed in a Bernoulli distribution. Hence, it is a Compound Poisson Distribution.

b. Question 2 - Compute $E(Z)$ and $Var(Z)$.
Answer -

$$\begin{gathered}
E(Y) = \sum_{n=0}^{\infty} E(Y|N=n)P(N=n) \\
\implies \sum_{n=0}^{\infty} nE(X)P(N=n) \\
\implies \mu_X \sum_{n=0}^{\infty} nP(N=n) \\
\implies E(Y) = \mu_X \, \mu_N \\
\implies E(Y) = \lambda \mu_X
\end{gathered}$$

Now, computing $E(Z)$

$$\begin{gathered}
E(Z) = E(Y_1) + E(Y_2) + E(Y_3) + E(Y_4) \\
\implies \mu_X \, (\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4) \\
\implies (\frac{2}{3} + \frac{2}{3})(2+3+1+3) \\
\implies \frac{4}{3} \times 9 \\
\implies E(Z) = 12
\end{gathered}$$

For $Var(Z)$,

$$
\begin{gathered}
Var(Y) = E_N(Var(Y|N)) + Var_N(E(Y|N)) \\
\implies E_N(NVar(X) + Var(NE(X)) \\
\implies Var(X)E(N) + (E(X))^2 Var(N) \\
\implies \mu_N \sigma_X^2 + \mu_X^2 \sigma_N^2 \\
\implies \lambda(E(X^2) - (E(X))^2) + \mu_X^2 \sigma_N^2 \\
\implies \frac{2}{9}\lambda + \frac{16}{9}\lambda = \frac{18}{9}\lambda = 2\lambda = 2 \times 9 = 18 \\
\implies Var(Z) = 18
\end{gathered}
$$

c. Compute $P(Y_2 > 5)$ and $P(Y_3 > 5)$ analytically (without a computer simulation). Subsequently, comment on the discrepancy between the two results (if any).

Answer:

Because Y_j ~ Compound Poisson(λ), we perform the computation in the following two steps:

1. We list a n-fold PMF for $\sum_{i=1}^{n} X_i$ ~ $bin(n, p)$.
2. We scale the n-fold PMF by the Poisson distributed probability weights with λ_j.

<u>**Let us first compute $P(Y_2 > 5)$, for which $\lambda = 3$.**</u>

One way of computing $P(Y_2 > 5)$ is to compute $1 - P(Y_2 \leq 5)$. We shall be employing this method for the purpose of this solution.

For the sake of computation, we consider an approved claim of \$200,000 ($p_2 = \frac{1}{3}$) to be success and \$100,000 ($p_1 = \frac{2}{3}$) to be failure.

Let us also define the notation that will be followed throughout the computation. For X_i ~ $bin(n, p)$:

$$
P^{(n)}(\sum_{i=1}^{n} X_i) =^n C_k \times p^k \times (1-p)^{n-k}
$$

Where,

$n \rightarrow$ total number of claims

$\sum_{i=1}^{n} X_i \rightarrow$ total amount claimed by claimants.

$k \rightarrow$ number of successes i.e. number of people claiming (and receiving $200,000)

$p \rightarrow$ success probability.

And, for **Poisson Distributed Probability Weights,**

$$P(Y = y) = \frac{\lambda^n}{n!} e^{-\lambda}$$

Where,

$n \rightarrow$ total number of claims

$\lambda \rightarrow$ rate at which claims are received per policy period

$y \rightarrow$ total amount claimed by claimants = $\sum_{i=1}^{n} X_i$

Let us now begin the computation.

For $\sum_{i=1}^{n} X_i = Y_2 = 1$,

$$P^{(1)}(1) = \frac{2}{3}$$

Now, scaling by Poisson Distributed Probability Weight

$$P(Y = 1) = \frac{2}{3} \times \lambda \times e^{-\lambda} = 0.0995$$

For $\sum_{i=1}^{n} X_i = Y_2 = 2$,

Case - 1: **1 person** claims a total of $200,000.

$$P^{(1)}(2) = \frac{1}{3}$$

Case - 2: **2 people** claim a total of $200,000.

$$P^{(2)}(2) =^2 C_0 \times (\frac{1}{3})^0 \times (\frac{2}{3})^2 = 0.44$$

Now, scaling by Poisson Distributed Probability Weights

$$P(Y=2) = \frac{1}{3} \times \lambda \times e^{-\lambda} + 0.44 \times \frac{\lambda^2 \times e^{-\lambda}}{2} = 0.1483$$

For $\sum_{i=1}^{n} X_i = Y_2 = 3$,

Case - 1: **2 people** claims a total of $300,000.

$$P^{(2)}(3) =^2 C_1 \times (\frac{1}{3})^1 \times (\frac{2}{3})^1 = 0.44$$

Case - 2: **3 people** claim a total of $300,000.

$$P^{(3)}(3) =^3 C_0 \times (\frac{1}{3})^0 \times (\frac{2}{3})^3 = 0.296$$

Now, scaling by Poisson Distributed Probability Weights

$$P(Y=3) = 0.44 \times \frac{\lambda^2 \times e^{-\lambda}}{2} + 0.296 \times \frac{\lambda^3 \times e^{-\lambda}}{6} = 0.1649$$

For $\sum_{i=1}^{n} X_i = Y_2 = 4$,

Case - 1: **2 people** claims a total of $400,000.

$$P^{(2)}(4) =^2 C_2 \times (\frac{1}{3})^2 \times (\frac{2}{3})^0 = 0.11$$

Case - 2: **3 people** claim a total of $400,000.

$$P^{(3)}(4) =^3 C_1 \times (\frac{1}{3})^1 \times (\frac{2}{3})^2 = 0.44$$

Case - 3: **4 people** claim a total of $400,000.

$$P^{(4)}(4) =^4 C_0 \times (\frac{1}{3})^0 \times (\frac{2}{3})^4 = 0.197$$

Now, scaling by Poisson Distributed Probability Weights

$$P(Y=4) = 0.11 \times \frac{\lambda^2 \times e^{-\lambda}}{2} + 0.296 \times \frac{\lambda^3 \times e^{-\lambda}}{6} = 0.1563$$

For $\sum_{i=1}^{n} X_i = Y_2 = 5$,

Case - 1: **3 people** claims a total of $500,000.

$$P^{(3)}(5) = ^3 C_2 \times (\frac{1}{3})^2 \times (\frac{2}{3})^1 = 0.22$$

Case - 2: **4 people** claim a total of $500,000.

$$P^{(4)}(5) = ^4 C_1 \times (\frac{1}{3})^1 \times (\frac{2}{3})^3 = 0.395$$

Case - 3: **5 people** claim a total of $500,000.

$$P^{(5)}(5) = ^5 C_0 \times (\frac{1}{3})^0 \times (\frac{2}{3})^5 = 0.132$$

Now, scaling by Poisson Distributed Probability Weights

$$P(Y=5) = 0.22 \times \frac{\lambda^3 \times e^{-\lambda}}{6} + 0.395 \times \frac{\lambda^4 \times e^{-\lambda}}{24} + 0.132 \times \frac{\lambda^5 \times e^{-\lambda}}{120} = 0.1294$$

For $\sum_{i=1}^{n} X_i = Y_2 = 0$,

$$P(Y=0) = e^{-\lambda} = 0.049787$$

Finally, calculating $P(Y > 5)$:

$$\begin{gathered} P(Y_2 > 5) = 1 - P(Y_2 \leq 5) \\ \implies 1 - [P(Y=0) + P(Y=1) + P(Y=2) + P(Y=3) + P(Y=4) + P(Y=5)] \\ \implies 1 - 0.0751784 \\ \implies P(Y_2 > 5) = 0.2482 \end{gathered}$$

Answer: $P(Y_2 > 5) = 0.2482$

Now, computing $P(Y_3 > 5)$, for which $\lambda = 1$

After performing similar steps for the third quarter, we arrive at the following results:

$$\begin{aligned} P(Y = 0) &= 0.3678794412 \\ P(Y = 1) &= 0.2452529608 \\ P(Y = 2) &= 0.2043774673 \\ P(Y = 3) &= 0.09991787291 \\ P(Y = 4) &= 0.05071589004 \\ P(Y = 5) &= 0.02008450173 \end{aligned}$$

Hence,

$$\begin{aligned} P(Y_3 > 5) &= 1 - P(Y \leq 5) \\ \implies P(Y_3 > 5) &= 1 - 0.98822 = 0.01177 \end{aligned}$$

Answer: $P(Y_3 > 5) = 0.01177$

Section 2 - Crank up the Monte Carlo Engine

a. Question 1 - Estimate $P(Y_2 > 5)$ and $P(Y_3 > 5)$ using the crude Monte Carlo simulation. Compare your simulation results with the analytical results that you obtained in the previous section. Comment on your comparisons.

Answer:

For $P(Y_2 > 5)$:

```
lambda=3;
N = 10^6;
quartertotalclaim = 5;
claimprob = zeros(N,1);
for i = 1:N
    n=poissrnd(lambda);
    claimconcat = zeros(n,1);
    if n>=0
```

```
        for j = 1:n
            U = rand(1);
            if U < 2/3
                claimconcat(j) = 1;
            else
                claimconcat(j) = 2;
            end
        end
    end

    if sum(claimconcat)>quartertotalclaim
        claimprob(i)=1;
    else
        claimprob(i)=0;
    end

end

ans = mean(claimprob)
relErr_hat = std(claimprob) / (ans*sqrt(N))
```

```
ans =

   0.248847000000000


relErr_hat =

   0.001737393592405
```

Analytical result - 0.2482

Simulation based result - 0.2488

Discrepancy = $0.2488 - 0.2482 = 0.0006$

For $P(Y_3 > 5)$

```
lambda=1;
N = 10^6;
quartertotalclaim = 5;
claimprob = zeros(N,1);
for i = 1:N
    n=poissrnd(lambda);
    claimconcat = zeros(n,1);
    if n>0

        for j = 1:n
            U = rand(1);
            if U < 2/3
                claimconcat(j) = 1;
            %elseif U >= 0.4 && U < 0.8
                %claimconcat(j) = 10;
            else
```

```
                claimconcat(j) = 2;
            end
        end
    end

    if sum(claimconcat)>quartertotalclaim
        claimprob(i)=1;
    else
        claimprob(i)=0;
    end

end

ans = mean(claimprob)
relErr_hat = std(claimprob) / (ans*sqrt(N))
```

```
ans =

   0.011719000000000


relErr_hat =

   0.009183223685576
```

Analytical result - 0.01177

Simulation based result - 0.01172

Discrepancy = $0.2488 - 0.2482 = 0.0006$

b. Question 2 - Let the total annual income on the sale of insurance premiums be $ 1, 000, 000. What is the risk of yearly loss for the company in terms of $P(Z > 1,000,000)$? You may provide your analysis of the risk by using an appropriate Monte Carlo simulation.

Answer:

```
clc
clear all
lambda=[2 3 1 3];
N = 10^4;

claimprob = [];
for i = 1:N
    total=0;
    for x=1:4
        claimconcat = [];
        temp=lambda(x);
        n=poissrnd(temp);
        if n>0
            for j = 1:n
                U = rand(1);
```

```
                if U < 2/3
                    claimconcat(j) = 1;
                else
                    claimconcat(j) = 2;
                end
            end
        end
        total=total+sum(claimconcat);
    end

    if total>10
        claimprob(i)=1;
    end

end

ans = mean(claimprob)
```

```
ans =

   0.610600000000000
```

Based on the output, we can see that there is a more than 60% chance that the insurance company will incur a loss.
