MIDDLE EAST TECHNICAL UNIVERSITY

Dep. Of Electrical and Electronics Engineering
EE230 - Probability & Random Variables Term Project
(Spring, 2019-2020)

Modeling the Randomness of Photons

Mustafa Barış Emektar 2304533

Part 1: Modeling

1.

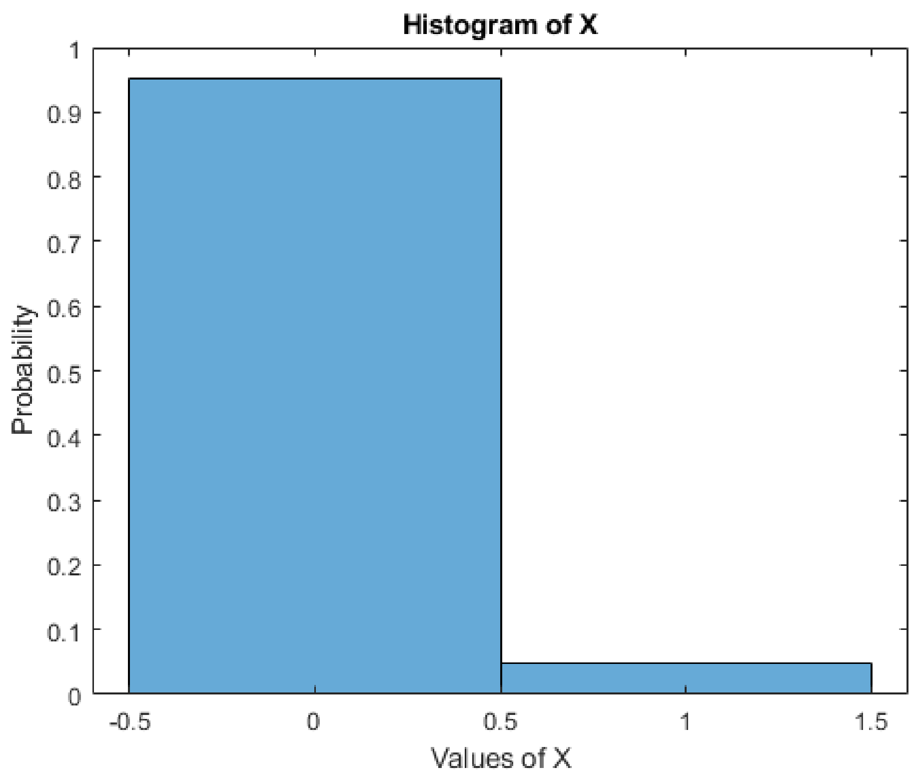
a.

$$P(x) = \begin{cases} a\Delta, & x = 1 \\ 1 - a\Delta, & x = 0 \\ 0, & otherwise \end{cases}$$

$$E(x) = \sum_X (x * P_x(x)) = 1 * a\Delta + 0(1 - a\Delta) = a\Delta$$

b.

Does this histogram look similar to the PMF of X?



```
num_trials = 10^5 ;
a = 10 ;
delta = 0.005 ;
X = zeros(10^5,1);
for trial=1:num_trials
    if rand() < (a*delta)
       X(trial) = 1;
    end
end
histogram(X,"Normalization","Pro
bability");
xlabel('Values of X');
ylabel('Probability');
title('Histogram of X');
```

Comment:

Yes. Histogram looks similar to PMF of X where PMF of X is Px(0)=0.95, Px(1)=0.05

2.

a.

$$a = 10, \qquad \Delta = 0.005, \qquad t = 1$$

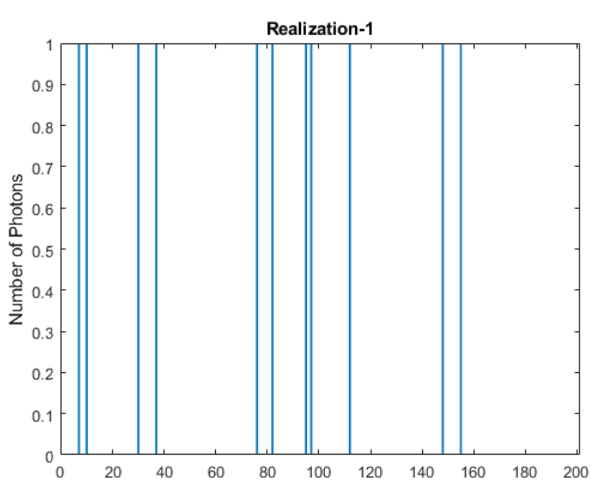
$$Ber(P) \sim X_i \;\; P_{xi}(x) = \begin{cases} a\Delta, & x = 1 \\ 1 - a\Delta, & x = 0 \end{cases}$$

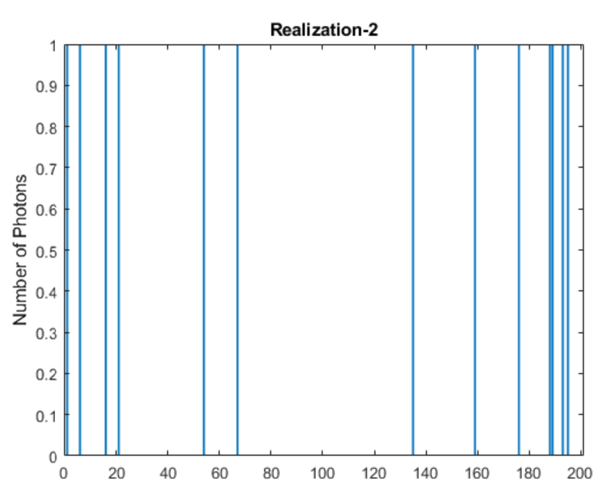
$t = n\Delta,$

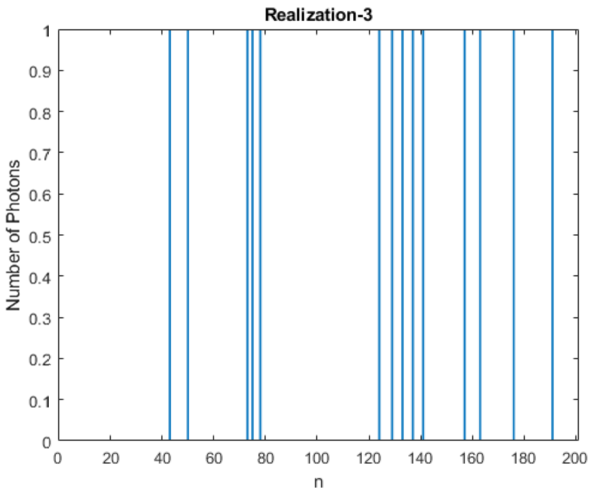
$1 = n * 0.005,$ $\qquad$ *i goes from 1 to 200*

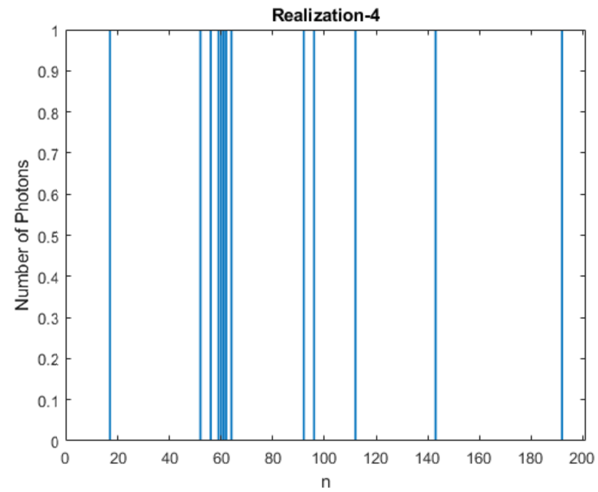
$n = 200$

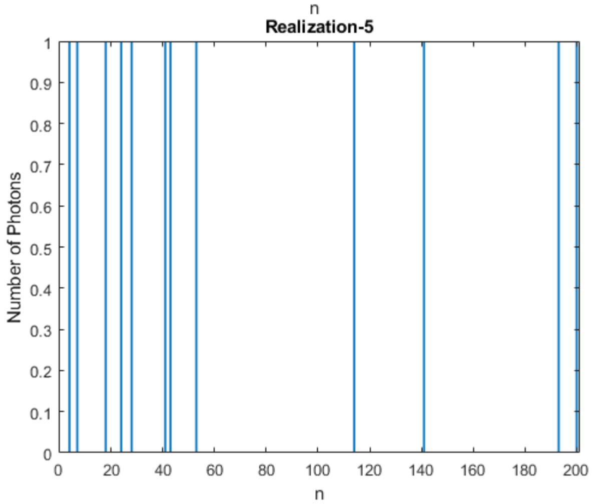
Do you think there is anything common in these plots? Comment.

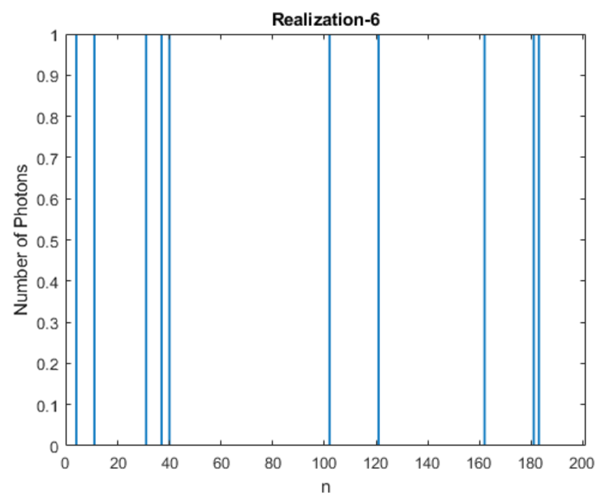


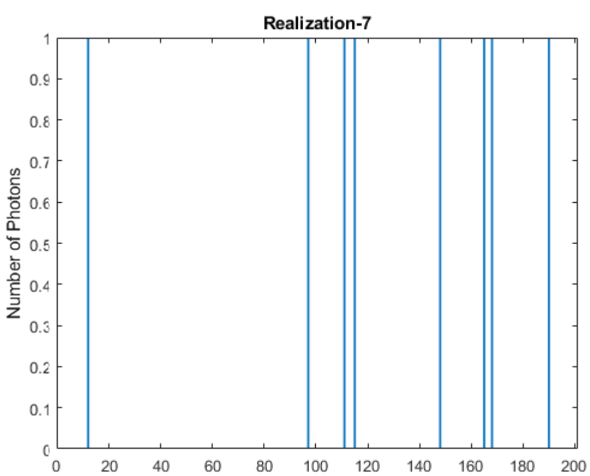


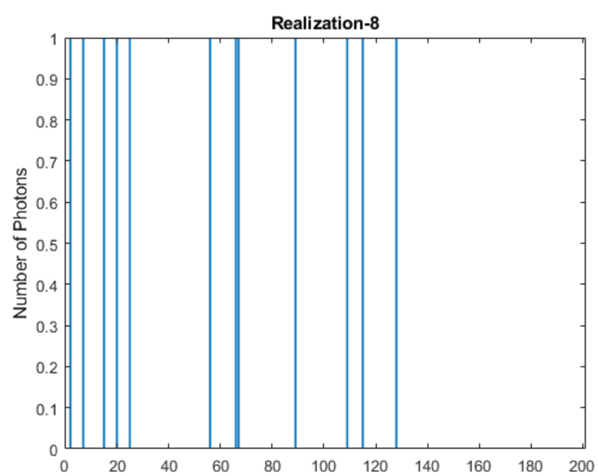

Realization-3
Number of Photons
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
0 20 40 60 80 100 120 140 160 180 200
n

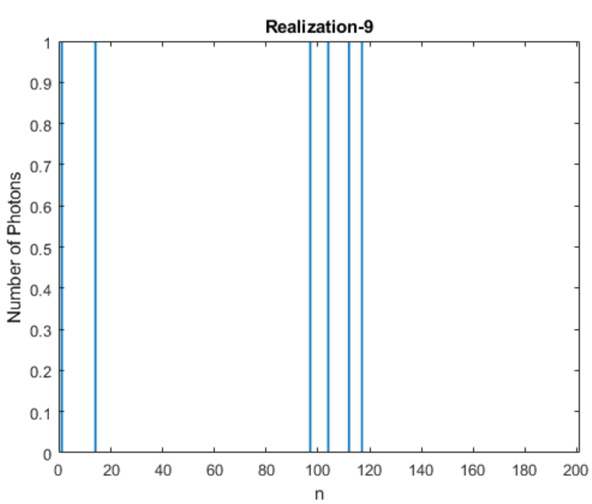

Realization-4
Number of Photons
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
0 20 40 60 80 100 120 140 160 180 200
n

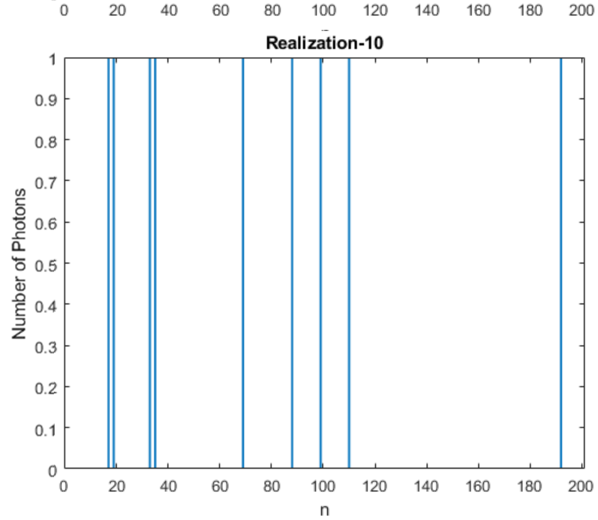

Realization-5
Number of Photons
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
0 20 40 60 80 100 120 140 160 180 200
n


Realization-6
Number of Photons
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
0 20 40 60 80 100 120 140 160 180 200
n


Realization-7
Number of Photons
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
0 20 40 60 80 100 120 140 160 180 200


Realization-8
Number of Photons
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
0 20 40 60 80 100 120 140 160 180 200


Realization-9
Number of Photons
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
0 20 40 60 80 100 120 140 160 180 200
n


Realization-10
Number of Photons
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
0 20 40 60 80 100 120 140 160 180 200
n

```
k=1
a = 10 ;
delta = 0.005 ;
n = 1/delta;
X = zeros(200,1);
 for trial=1:200

        if rand() < (a*delta)
          X(trial) =  1;
        end
end

bar(X);
xlabel('n');
ylabel('Number of Photons');
title('Realization-10');
```

Comment:

Common thing in those plots is that number of photons that arrived in 200 Bernoulli trials are around 10 which is the mean of 200 Bernoulli trials with a=10 and delta=0.005.

b. Since we can assume Y as sum of independent Bernoulli random variables, Y is a Binomial Random Variable.

$$P_Y(y) = \binom{n}{y} (a\Delta)^y (1 - a\Delta)^{1-y}$$

By linearity of expectation, we know that expectation of sum is sum of expectations.

We also know Y is sum of independent Bernoulli random variables X_{i}.

$$Y = \sum_{i=1}^{n} Xi \;\rightarrow \mathrm{E}\left[\sum_{i=1}^{n} Xi\right] = \sum_{i=1}^{n} E(xi)$$

(Sum of expectations of Independent Trials)

$$\Rightarrow E[Y] = \sum_{i=1}^{n} E(xi) = \sum_{i=1}^{n} P = n * P$$

(each trial is identical)

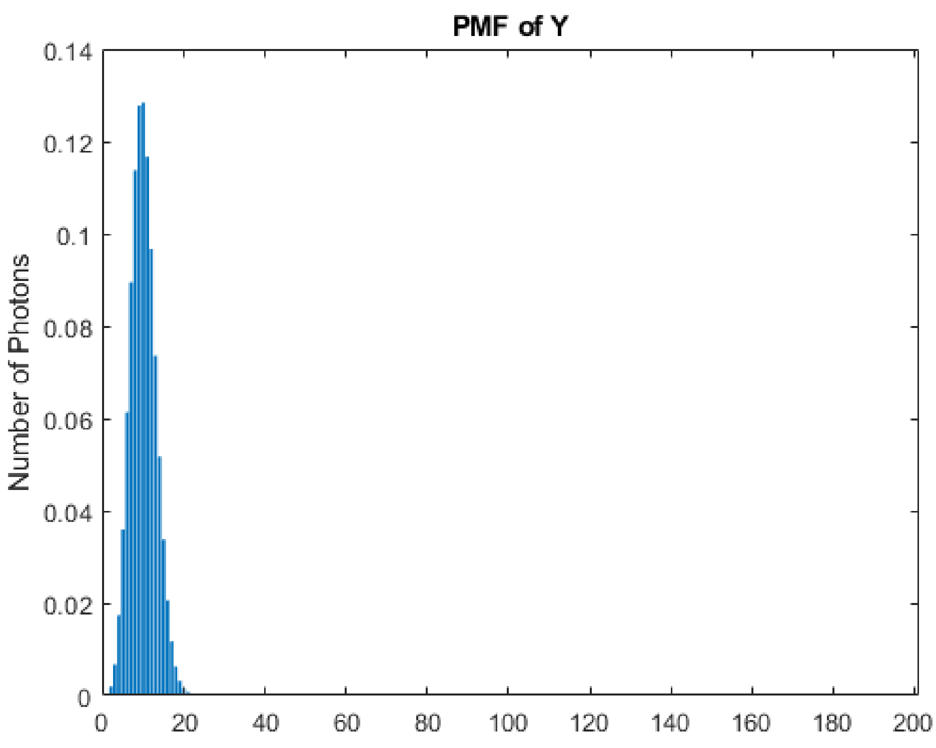
$$\Rightarrow E[Y] = n * P = n * a\Delta = \mathrm{n} * \Delta \mathrm{a} = \mathrm{at} \;\Rightarrow\; \boldsymbol{E[Y] = \mathrm{at}}$$

What does this expected value mean in terms of the photons hitting the sensor?

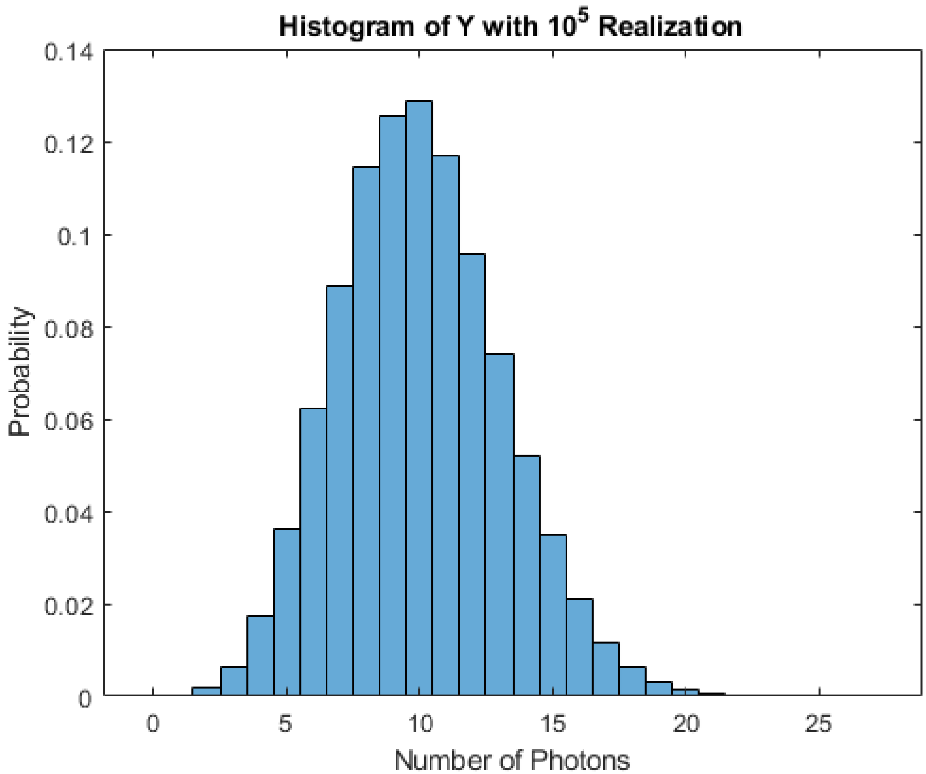
Comment:

Since E[Y]=a*t, if we increase the exposure time "t" where t=n*delta, number of the photons hitting the sensor will increase, therefore mean will increase.
Or if the constant "a" increases, probability that a photon hitting sensor will increase, therefore mean will increase also.

c. Verify that the normalized histogram look similar to the PMF of Y.




```
a = 10 ;
delta = 0.005 ;
n = 1/delta;
X = 1;
Px = zeros(200,1);
for trial=1:200
    Px(X) = nchoosek(n,X)*((a*delta).^X)*((1-
a*delta).^(200-X));
    X=X+1;
end
bar(Px);
xlabel('Probability');
ylabel('Number of Photons');
title('PMF of Y');
```




```
num_trials = 10^5 ;
a = 10 ;
delta = 0.005 ;
X = zeros(200,1);
Y = zeros(10^5,1);
for c=1:num_trials
       for trial=1:200
        if rand() < (a*delta)
          X(trial) = 1;
        end
    end
    for d=1:200
    Y(c) = Y(c) + X(d);
    end
    X = zeros(200,1);
end
histogram(Y,"Normalization","Probability")
xlabel('Number of Photons');
ylabel('Probability');
title('Histogram of Y with 10^5 Realization');
```

Comment:
As it can be seen from plots normalized histogram with 10^5 realizations looks like PMF of Y.

3.

a.

$$P_Y(y) = \binom{n}{y}(a\Delta)^y(1-a\Delta)^{1-y}$$

$$P_Y(y) = \frac{n!}{y! * (n-y)!} * (a\Delta)^y * (1-a\Delta)^{n-y}$$

$$= \frac{n!}{y! * (n-y)!} * (a\Delta)^y * (1-\frac{at}{n})^{n-y}$$

$$\left(Since\ t = n\Delta \Rightarrow\ \Delta = \frac{t}{n}\right)$$

$$P_Y(y) = \frac{(at)^y}{y!} * \frac{n!}{(n-y)!} * \frac{1}{n^y} * (1-\frac{at}{n})^{n-y}$$

$$\left(\frac{n!}{(n-y)!} * \frac{1}{n^y} * \left(1-\frac{at}{n}\right)^{n-y} = e^{-at} \quad as\ n \rightarrow\ \infty\right)(*)$$

$$P_y(y) = \frac{(at)^y}{y!} * e^{-at}$$

$$(*) \qquad \frac{n!}{(n-y)!} * \frac{1}{n^y} = \frac{n*(n-1)*(n-1)*\ldots*(n-y+1)*(n-y)!}{(n-y)!*n^y}$$

$$= \frac{n*(n-1)*(n-1)*\ldots*(n-y+1)}{n^y} = \frac{n}{n} = \frac{n-2}{n} = \cdots = \frac{n-y+1}{n}$$

$$= 1 * \left(1-\frac{1}{n}\right) * \left(1-\frac{2}{n}\right) * \ldots * \left(1-\frac{y+1}{n}\right)$$

$$P_Y(y) = \frac{(at)^y}{y!} * \frac{n!}{(n-y)!} * \frac{1}{n^y} * \left(1-\frac{at}{n}\right)^{n-y}$$

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = \left(1-\frac{1}{n}\right) * \left(1-\frac{2}{n}\right) * \ldots * \left(1-\frac{y+1}{n}\right)$$

$$\lim_{n\to\infty} P_Y(y) = \frac{(at)^y}{y!} * \lim_{n\to\infty}\left(1-\frac{1}{n}\right) * \ldots * \left(1-\frac{y+1}{n}\right) * \lim_{n\to\infty}\left(1-\frac{at}{n}\right)^{-y}$$

$$P_Y(y) = \frac{(at)^y}{y!} * \lim_{n\to\infty}\left(1-\frac{at}{n}\right)^n$$

$$\lim_{n\to\infty}\left(1-\frac{at}{n}\right)^n = z$$ (assign this expression to z)

$$ln(z) = \lim_{n\to\infty}\left(1-\frac{at}{n}\right) = \lim_{n\to\infty}\frac{ln(1-\frac{at}{n})}{\frac{1}{n}}$$

L'Hospitals Rule ⇒

$$\lim_{n\to\infty}\frac{\left(\frac{1}{1-\frac{at}{n}}\right)\left(\frac{at}{n^2}\right)}{\frac{-1}{n^2}} = \lim_{n\to\infty}\frac{ln(-at)}{a-\frac{at}{n}} = -at$$

$$\Rightarrow ln(z) = -at \quad \Rightarrow \quad \mathrm{z} = e^{-at}$$

$$P_Y(y) = \frac{(at)^y}{y!} * \lim_{n\to\infty}\left(1-\frac{at}{n}\right)^n$$

$$P_Y(y) = \frac{(at)^y}{y!} * e^{-at}$$

What does the PMF of Y converge to in this case?

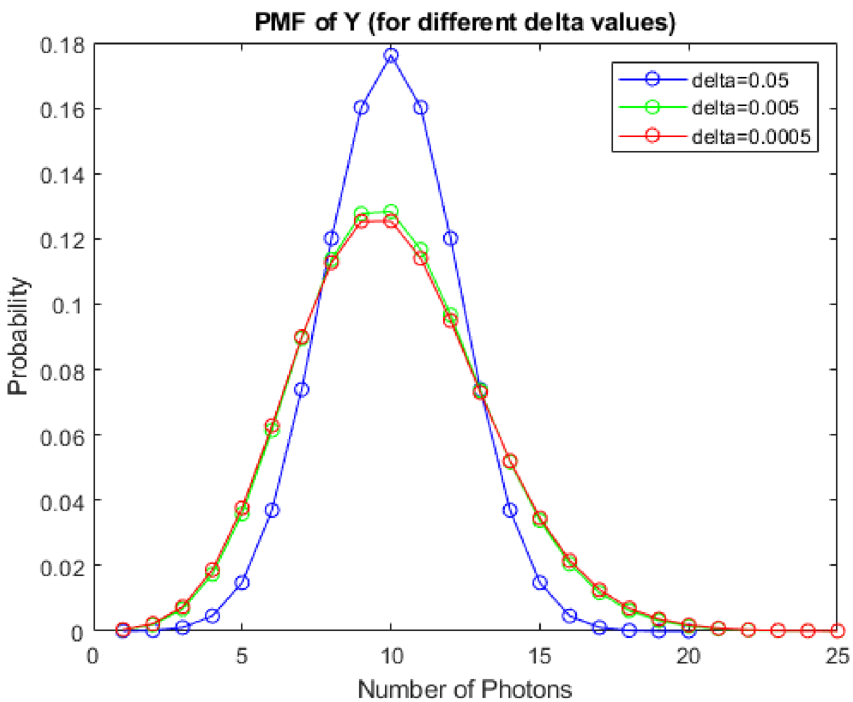
Comment:
As n goes to infinity Binomial random variable Y converge to a Poisson random variable. So PMF of Y converges to Poisson distribution.

$$E[Y] = \sum_{y=0}^{n} y * P_y(y) = \sum_{y=0}^{n} y * e^{-at} * \frac{(at)^y}{y!} = \sum_{y=1}^{n} y * e^{-at} * \frac{(at)^y}{y(y-1)!}$$

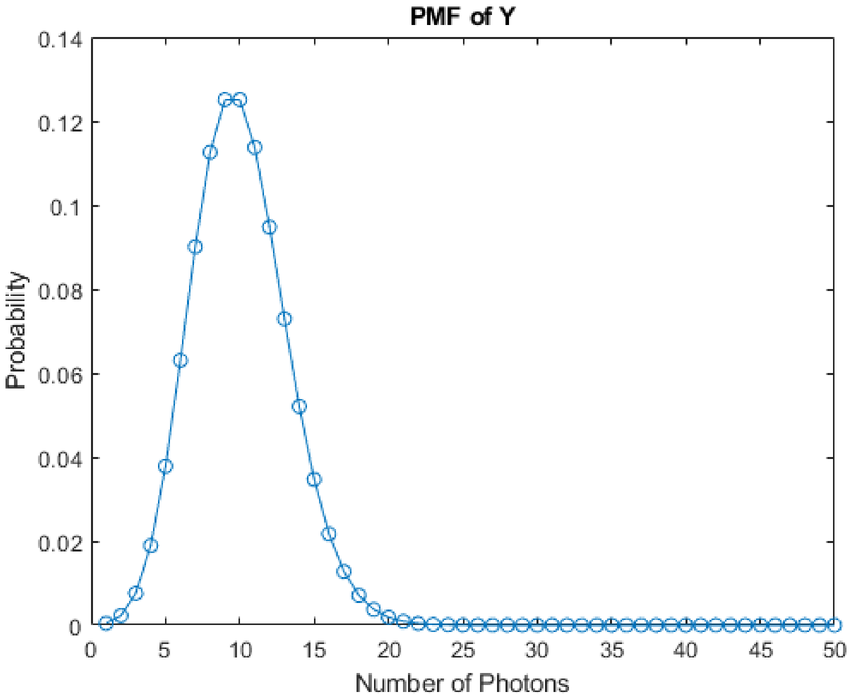
$$= \sum_{y=1}^{n} e^{-at} * \frac{(at)^y}{(y-1)!} = \sum_{y=1}^{n} e^{-at} * \frac{(at)^{k+1}}{k!}$$

$$= (at)\sum_{y=1}^{n} e^{-at} * \frac{(at)^k}{k!} = at \Rightarrow \boldsymbol{E[Y] = at}$$

b. Compare these plots and comment.




```
a = 10 ;
delta = 0.05 ;
n = 1/delta;
X = 1;
Px = zeros(n,1);
grid on
for trial=1:n
    Px(X) =
nchoosek(n,X)*((a*delta).^X)*((1-
a*delta).^(n-X));
    X=X+1;
end
plot(Px,'b-o');
hold on
a = 10 ;
delta = 0.005 ;
n = 1/delta;
X = 1;
Px = zeros(n,1);
for trial=1:n
    Px(X) =
nchoosek(n,X)*((a*delta).^X)*((1-
a*delta).^(n-X));
    X=X+1;
end
plot(Px,'g-o');
hold on
a = 10 ;
delta = 0.0005 ;
n = 1/delta;
X = 1;
Px = zeros(n,1);
for trial=1:n
    Px(X) =
nchoosek(n,X)*((a*delta).^X)*((1-
a*delta).^(n-X));
    X=X+1;
end
plot(Px,'r-o');
xlim([0 25])
xlabel('Number of Photons');
ylabel('Probability');
title('PMF of Y (for different
delta values)');
legend('delta=0.05','delta=0.005'
,'delta=0.0005');
```




```
a = 10 ;
n = 50;
Y = 1;
t=1;
Py = zeros(n,1);
grid on

for trial=1:n
    Py(Y) = [(a*t).^Y]*(exp(-
a*t)) /factorial(Y);
    Y=Y+1;
end

plot(Py,'-o');

xlabel('Number of Photons');
ylabel('Probability');
title('PMF of Y');
```

Comment:
As delta decreases, we can see that Binomial random variable Y gets smoother and converge to a Poisson random variable.

c. By Binomial,

$$P_Y(y) = \binom{n}{y}(a\Delta)^y(1-a\Delta)^{1-y}$$

$$P(Y<2) = P(Y=0) + P(Y=1)$$

$$= 1*1*(1-a\Delta)^n + n(a\Delta)(1-a\Delta)^{n-1}$$

$$= 1(1-a\Delta)^{t/\Delta} + at(1-a\Delta)^{t/\Delta} * (1-a\Delta)^{-1}$$

$$P(Y<2) = (1-a\Delta)^{\frac{t}{\Delta}} * [1 + \frac{at}{1-a\Delta}]$$

$$P(Y<2) = (1-a\Delta)^{\frac{t}{\Delta}-1} * [1 - a\Delta + at]$$

By Poisson distribution,

$$P_Y(y) = \frac{at^y}{y!}(a\Delta)^y e^{-at}$$

$$P(Y < 2) = (at + 1) * e^{-at}$$

$$P(Y < 2) = (11)e^{-10} = 4.99399 * 10^{-4} \quad \Rightarrow \quad By\ Poisson$$

$$a = 10, \qquad \Delta = 0.005, \qquad t = 1$$

$$P(Y < 2) = 4.04028 * 10^{-4} \quad \Rightarrow \quad By\ Binomial$$

$$P(Y < 2) = 11 * e^{-10} = 4.99399 * 10^{-4} \quad \Rightarrow \quad By\ Poisson$$

$$a = 10, \qquad \Delta = 0.005, \qquad t = 1$$

$$P(Y < 2) = (1 - 0.005)^{-4}[1 - 0.005 + 10] \quad \Rightarrow \quad By\ Binomial$$

$$P(Y < 2) = 4.89253 * 10^{-4} \quad \Rightarrow \quad By\ Binomial$$

$$P(Y < 2) = 11 * e^{-10} = 4.99399 * 10^{-4} \quad \Rightarrow \quad By\ Poisson$$

Compare the results and comment on their agreement.

Comment:
For huge delta values, n is small and difference between Binomial and Poisson is significant. But as delta decreases, n increases; Binomial and Poisson values gets closer and closer just like we showed that as n goes to infinity Binomial converges to Poisson.

Part 1: Applications

4.

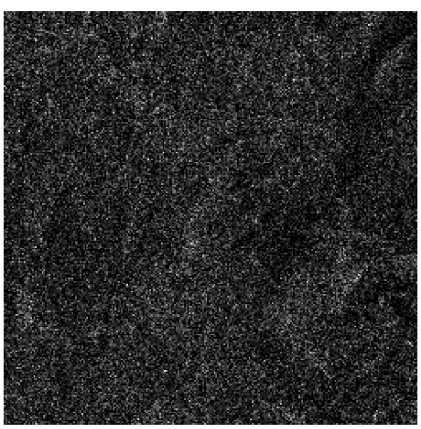
a.

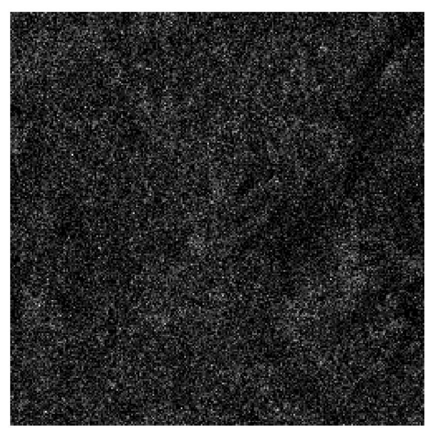
```
P = zeros (250,250);
Y = zeros(250,250);

for i=1:250

    for j=1:250

        lambda = I(i,j);
        Y(i,j) = random('Poisson',lambda);

    end

end

imshow(Y,[min(Y(:)) max(Y(:))])
```

b. Compare the resulting images and comment on how lower levels of light effects the quality of the image obtained. In particular, clearly explain why you obserce a grainy structure at lower levels of light, and try to relate this to ratio of the mean of the photon counts(for each pixel) to its standard deviati

```
P = zeros (250,250);

Y = zeros(250,250);

for i=1:250

    for j=1:250

        lambda = I(i,j)/10;
        Y(i,j) = random('Poisson',lambda);

    end

end

imshow(Y,[min(Y(:)) max(Y(:))])
```

```
P = zeros (250,250);
 Y = zeros(250,250);
 for i=1:250
    for j=1:250
        lambda = I(i,j)/100;
        Y(i,j) = random('Poisson',lambda);
    end
end
imshow(Y,[min(Y(:)) max(Y(:))])
```





```
P = zeros (250,250);
Y = zeros(250,250);
for i=1:250
    for j=1:250
        lambda = I(i,j)/1000;
        Y(i,j) = random('Poisson',lambda);
    end
end
imshow(Y,[min(Y(:)) max(Y(:))])
```

Comment:

Standard deviation represents how values are spread across a mean. That is if standard deviation is higher, we have data that are much more spread. As we decrease the brightness, standard deviation also decreases and differences between photon counts on each pixel decreases significantly.

Therefore, each pixel starts to look like same. Since information is carried by photons to sensor, decrease in photon numbers increases the noises and diminishes the differences between pixels, therefore we observe grainy structure.

c. To prevent obtaining such noisy images at low-light levels, today's cameras resort two common things:

1-Using flash.
2-Increasing the exposure time when the flash is off.

Explain how these two approaches help to deal with the shot noise. Also try to comment on what other things you could possibly change on the camera to better deal with shot noise and why these changes can be undesirable.

Comment:

By using flash, we increase the number of photons that reflects from light-colored objects which allows us to distinguish them from dark colored objects in a night. Therefore, we increase the mean and standard deviation, which allows us to have clear photo. By increasing exposure time when flash is off, we increase the number of n in t=n*delta. By doing this we decrease the randomness in photograph, longer the exposure closer it will be to its mean. Aperture is the opening in a lens through which light passes to enter the camera. The higher the f-number, the darker the image projected on the image sensor, and the darker the resulting photograph. On the other hand, the lower the f-number, the brighter the image projected on the image sensor, and the brighter the resulting photograph. But increasing the number of photons that enters the sensor too much in day light can cause similar problems when environment is dark; that is we have small standard deviation.

An Example of Exposure Time

```
P = zeros (250,250);
Y = zeros(250,250);
K = zeros(250,250);
    for i=1:250
        for j=1:250
            lambda = I(i,j)/5000;
            Y(i,j) =
random('Poisson',lambda);
        end
    end
imshow(Y,[min(Y(:)) max(Y(:))])
```

```
P = zeros (250,250);
Y = zeros(250,250);
K = zeros(250,250);
for m=1:100
    for i=1:250
        for j=1:250
            K(i,j) = K(i,j) + Y(i,j);
            lambda = I(i,j)/10000;
            Y(i,j) =
random('Poisson',lambda);
        end
    end
end
K = K/100;
imshow(K,[min(K(:)) max(K(:))])
```

5.

a.

$$P_{X_1,X_2,\dots,X_N}(X_1,X_2,\dots,X_N) = \prod_{k=1}^{\infty} P_{X_i}(X_i,\lambda) = \prod_{k=1}^{\infty} e^{-at} * \frac{\lambda^{X_i}}{X_i!}$$

$$ln\left(P_{X_1,X_2,\dots,X_N}(X_1,X_2,\dots,X_N)\right) = \sum_{i=1}^{n} \ln\left(e^{-at} * \frac{\lambda^{X_i}}{X_i!}\right)$$

$$= \sum_{i=1}^{n} [(-\lambda) + X_i \ln(\lambda) - \ln(X_i!)]$$

$$= -\lambda * \mathrm{n} + \ln\left[\sum_{i=1}^{n} X_i\right] - \ln\left(X_1!\, X_2! \ldots X_n!\right)$$

$$\frac{d\left[ln\left(P_{X_1,X_2,\ldots,X_N}(X_1, X_2, \ldots, X_N; \lambda)\right)\right]}{d\lambda} = -n + \frac{1}{\lambda} * \sum_{i=1}^{n} X_i = 0$$

We took derivative and set it equal t ozero find λ valve which maximizes the function $P_{X_1,X_2,\ldots,X_N}(X_1, X_2, \ldots, X_N)$

$$\frac{d\left[ln\left(P_{X_1,X_2,\ldots,X_N}(X_1, X_2, \ldots, X_N; \lambda)\right)\right]}{d\lambda^2} = \frac{1}{\lambda^{*2}} * \sum_{i=1}^{n} X_i = 0$$

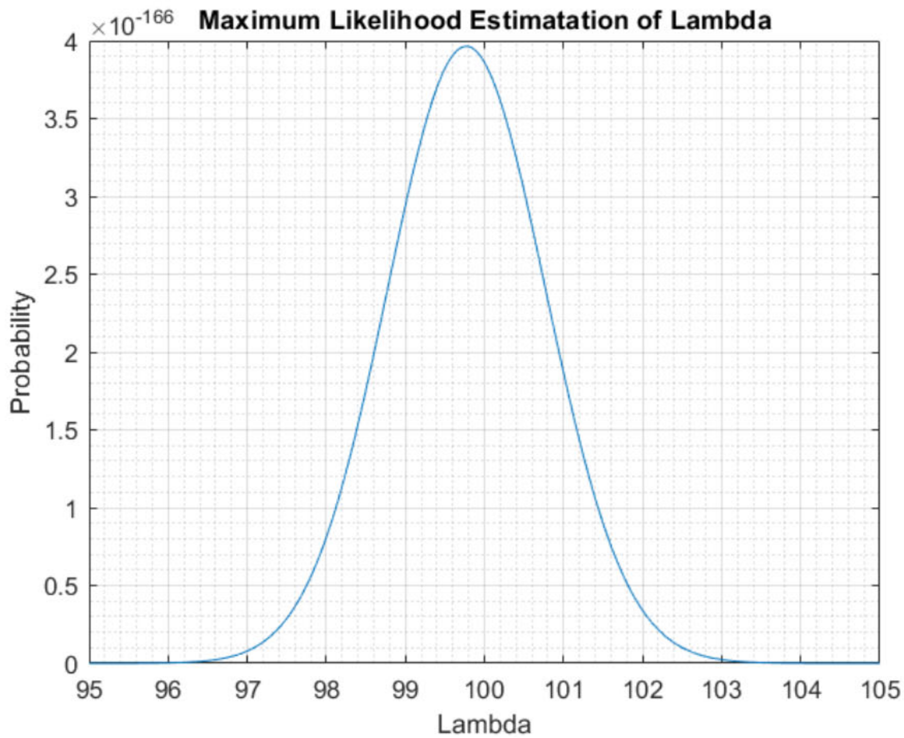
$$\lambda \Rightarrow \lambda^* \qquad \lambda^* \; is\ maximum$$

If you use this approach, explain why taking the logarithm does not change the optimal value of λ.

b. Comment on the resulting estimation error. Also verify that the value of λ corresponding to the maximum value of this function is your maximum likelihood estimate.

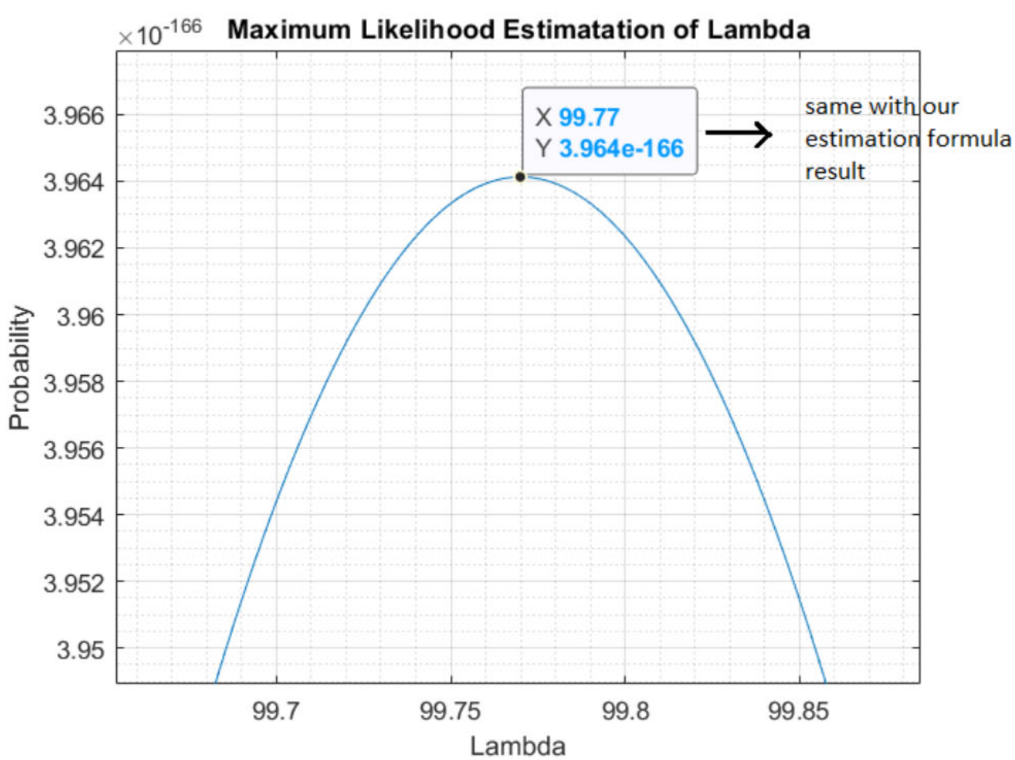
Comment :

We can use logarithmic likelihood instead of normal likelihood since the behavior of our probability graph does not change due to nature of logarithm. The functions have the same extreme points (for greater than zero). Monotone increasing functions stay monotone increasing after taking logarithm or non-monotone functions stay non-monotone, as well.



According to our esmation formula for lambda in part a:

```
Command Window

estimation =

   99.7700
```



```
x = zeros (100,1);
PX = zeros (100000,1);
lambda = 100;
estimation = 0;
lambdas = zeros(100000,1)
product = ones(100000,1)
z= 1;
for x=95:0.0001:105
if z == 100001
break
end
lambdas(z,1)=x;
z = z+1;
end
j =1;
for i=1:100000
if j == 100001
break
end
X(i,1) =
random('Poisson',lambda);
j= j+1;
end
%------------
for n=1:100
estimation = estimation +
(1/100)*X(n,1); end
estimation
error = estimation-lambda
for a=1:100000
for b=1:100
PX(X(b,1)) = (lambdas(a,1).^X(b,1))*exp(-
lambdas(a,1))/factorial(X(b,1));
product(a,1)=product(a,1)*PX(X(b,1));
end
end
plot(lambdas,product)
xlabel('Lambda');
ylabel('Probability');
title('Maximum Likelihood Estimatation of
Lambda');
grid on
grid minor
```

Comment:

As input, we have given lambda as 100, and created Xi depending on Poisson with lambda=100. Due to randomness, lambda which maximizes the Px1,...,xn(x1,...,xn; lambda) is not exactly 100, but it is around 100 and it is same with our estimation in part a.