QUESTION 1) The code is present in .py file

QUESTION 2) The code is present in .py file

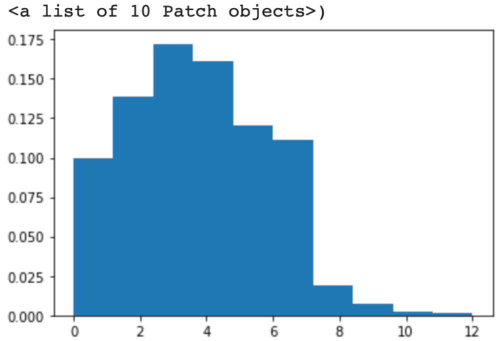
CF FOR TRAINING SET - 50.46528436113467
CF FOR TESTING SET - 268.7927887195447

QUESTION 3)

A)

```
[ ] from scipy.stats import poisson
    import numpy as np
    import matplotlib.pyplot as plt
    from scipy import stats
    A = np.load("PoissonX.npy")
    plt.hist(A, density = True, histtype = 'bar')
```

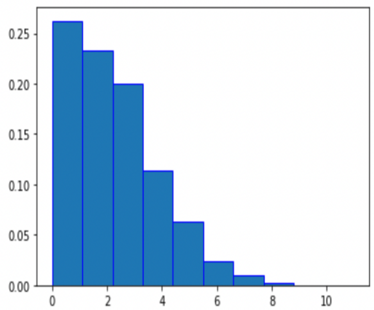
```
(array([0.09958333, 0.13908333, 0.17216667, 0.16083333, 0.12075    ,
        0.11141667, 0.019     , 0.00691667, 0.00241667, 0.00116667]),
 array([ 0. ,  1.2,  2.4,  3.6,  4.8,  6. ,  7.2,  8.4,  9.6, 10.8, 12. ]),
 <a list of 10 Patch objects>)
```



1) probability distributions of Poisson random variable with mu = 2.5

```python
import numpy as np
from scipy.stats import poisson
import matplotlib.pyplot as plt
from scipy import stats
a = poisson.rvs(mu = 2.5, size = 10000)
plt.hist(a, density=True, edgecolor='blue')
```

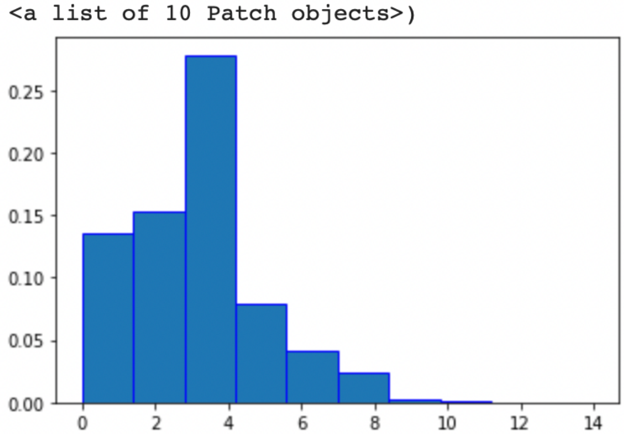
```
(array([0.26245455, 0.23336364, 0.19936364, 0.11345455, 0.06290909,
        0.02436364, 0.01018182, 0.00245455, 0.00027273, 0.00027273]),
 array([ 0. ,  1.1,  2.2,  3.3,  4.4,  5.5,  6.6,  7.7,  8.8,  9.9, 11. ]),
 <a list of 10 Patch objects>)
```



2) probability distributions of Poisson random variable with mu = 3.1

```python
import numpy as np
from scipy.stats import poisson
import matplotlib.pyplot as plt
from scipy import stats
a = poisson.rvs(mu = 3.1, size = 10000)
plt.hist(a, density=True, edgecolor='blue')
```

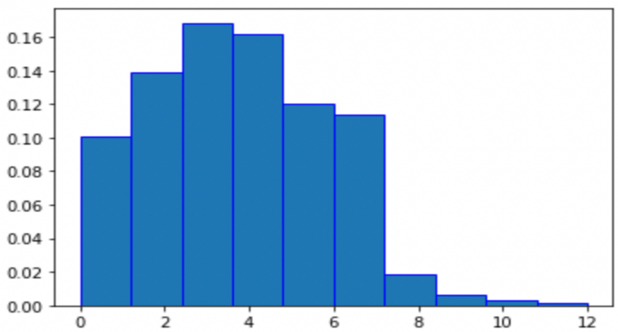
```
(array([1.36071429e-01, 1.52500000e-01, 2.78214286e-01, 7.92142857e-02,
        4.13571429e-02, 2.37142857e-02, 2.07142857e-03, 1.00000000e-03,
        7.14285714e-05, 7.14285714e-05]),
 array([ 0. ,  1.4,  2.8,  4.2,  5.6,  7. ,  8.4,  9.8, 11.2, 12.6, 14. ]),
 <a list of 10 Patch objects>)
```



3) probability distributions of Poisson random variable with mu = 3.7

```python
import numpy as np
from scipy.stats import poisson
import matplotlib.pyplot as plt
from scipy import stats
a = poisson.rvs(mu = 3.7, size = 10000)
plt.hist(a, density=True, edgecolor='blue')
```

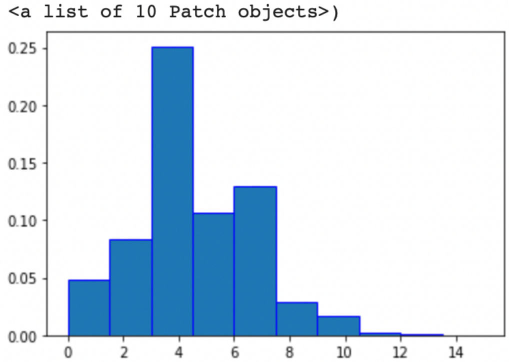
```
(array([0.10091667, 0.13883333, 0.1685    , 0.16208333, 0.11991667,
        0.114     , 0.0185    , 0.00625   , 0.00275   , 0.00158333]),
 array([ 0. ,  1.2,  2.4,  3.6,  4.8,  6. ,  7.2,  8.4,  9.6, 10.8, 12. ]),
 <a list of 10 Patch objects>)
```



4) probability distributions of Poisson random variable with mu = 4.3

```python
import numpy as np
from scipy.stats import poisson
import matplotlib.pyplot as plt
from scipy import stats
a = poisson.rvs(mu = 4.3, size = 10000)
plt.hist(a, density=True, edgecolor='blue')
```

```
(array([4.80000000e-02, 8.34666667e-02, 2.51066667e-01, 1.06000000e-01,
        1.29800000e-01, 2.88666667e-02, 1.68000000e-02, 1.86666667e-03,
        6.00000000e-04, 2.00000000e-04]),
 array([ 0. ,  1.5,  3. ,  4.5,  6. ,  7.5,  9. , 10.5, 12. , 13.5, 15. ]),
 <a list of 10 Patch objects>)
```



Based on visual inspection of the idealized distributions with the empirical distribution, mu = 3.7 is most consistent with the data.

B) i. For the values of x with large magnitude corresponding value of y tends to be larger than others.
   ii. Here for the values of x with small magnitude corresponding uncertainty of y tends to be larger.

QUESTION 4)

On the rest of the pages of pdf

Question no 4

4 a)

$\Rightarrow$

$$\nabla_x = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} \qquad a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

i) $$x^T a = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix}$$

$$= \begin{bmatrix} x_1 a_1 & x_2 a_2 & x_3 a_3 & \cdots & x_n a_n \end{bmatrix}$$

ii) $$a^T x = \begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$$

$$= \begin{bmatrix} a_1 x_1 & a_2 x_2 & a_3 x_3 & \cdots & a_n x_n \end{bmatrix}$$

i) $\nabla_x (x^T a)$

$$\Rightarrow \nabla_x [x_1 a_1 \; x_2 a_2 \; x_3 a_3 \; \cdots \; x_n a_n]$$

$$\Rightarrow \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{bmatrix} [x_1 a_1 \; x_2 a_2 \; x_3 a_3 \; \cdots \; x_n a_n]$$

$$\Rightarrow \left[\frac{d(x_1 a_1)}{dx_1} \quad \frac{d(x_2 a_2)}{dx_2} \; \cdots \; \frac{d(x_n a_n)}{dx_n} \right]$$

$$= [a_1 \; a_2 \; a_3 \; \cdots \; a_n]$$

$$= \underline{\underline{a}}$$

=

ii) $\nabla_x (a^T x)$

$$\begin{bmatrix} \frac{\partial}{\partial x_1} \\ \frac{d}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{bmatrix} [x_1 a_1 \; x_2 a_2 \; \cdots \; x_n a_n]$$

$$= \left[\frac{d(x_1 a_1)}{dx_1} \quad \frac{d(x_2 a_2)}{dx_2} \; \cdots \; \frac{d(x_n a_n)}{dx_n} \right]$$

$$= [a_1 \; a_2 \; a_3 \; \cdots \; a_n]$$

$$= a_1$$

$\therefore \quad \nabla_x (x^T a) = \nabla_x (a^T x) = a$

b) $\nabla_x (x^T A x) = (A + A^T) x$

for any column vector $x \in R^n$ and any $n \times n$ matrix A.

Here we know

$$\frac{d}{dx}(u \cdot Av) = \frac{\partial u^T A v}{dx} = \frac{du}{\partial x} Av + \frac{\partial v}{\partial x} A^T u$$

$\left(\frac{\partial u}{dx}, \frac{\partial v}{dx} \text{ in denominator layout} \right)$

Here,

$u = v = x$

$$LHS = \frac{d}{dx}(x^T A x) = \frac{\partial x^T A x}{\partial x} = \frac{\partial x}{\partial x} A x + \frac{\partial x}{dx} A^T x$$

$$= Ax + A^T x$$

$$= (A + A^T) x$$

$\therefore$ LHS = RHS

$\therefore$ Hence proved

c) $\nabla_x(x^T A x) = 2Ax$

for any column vector $x \in R^n$ and any symmetric $n \times n$ matrix A.

$\Rightarrow$ Here for A being a symmetric matrix

$A^T = A$

$\therefore$ By using

$\nabla_x(x^T A x) = (A + A^T)x$

$= (2A)x$

$\therefore$ LHS = RHS

Hence proved

d) $\nabla_x \left[(Ax+b)^T (Ax+b) \right] = 2A^T (Ax+b)$

for any column vector $x \in R^n$, any symmetric $n \times n$ matrix A, and any constant column vector $b \in R^n$,

$\Rightarrow$ Here by using,

$$\frac{d(uv)}{dx} = u\frac{dv}{dx} + v\frac{du}{dx}$$

Here,
$(Ax+b)^T = u$

$(Ax+b) = v$

$$(Ax+b)^T \frac{d(Ax+b)}{dx} + (Ax+b)\frac{d(Ax+b)^T}{dx}$$

$$\Rightarrow (Ax+b)\ A + (Ax+b)\ A$$

As A is a symmetric matrix

$$A = A^T$$

$$\Rightarrow A^T \left(2(Ax+b)\right)$$

$$= \underline{\underline{2A^T (Ax+b)}}$$

$\therefore$ LHS = RHS Hence proved