Assignment 1, IME692, Fall 2021

1. Consider the handwritten digit recognition dataset uploaded with this assignment (train3.csv). This dataset is from the handwritten ZIP codes on envelopes from US postal mail. We have selected the data corresponding to handwritten digit '3' for this assignment. Each observation for digit '3' was digitized into 16x16 grayscale image, and 658 such images are mentioned in the file. Each line in the file consists of the 256 grayscale values for an image.

   Select first 130 observations from the dataset and compute their first two principal components. Plot the images of the mean vector $\bar{\mathbb{x}}$ along with the first two PCA eigenvectors and mention the corresponding eigenvalues.

2. Consider following joint distribution for the random variables X and Y.

	Y=0	Y=1
X=0	1/3	1/3
X=1	0	1/3

   Calculate the following quantities: (a) $\mathbb{H}(y|x)$, (b) $\mathbb{H}(x,y)$, and (c) $\mathbb{I}(x,y)$

3. Consider the following bi-variate normal distribution $X = \binom{X_1}{X_2} \sim N_2\left(\binom{2}{1}, \begin{bmatrix} 2 & 0.4 \\ 0.4 & 2 \end{bmatrix}\right)$. Plot the constant density contours of this bivariate distribution. Mention the directions and length of semi-major and semi-minor axes for the elliptical contours.