Problem 1

Example 1: Images are not represented in a matrix that is defined in the question. In our matrix, each row are independent as they are samples. Images rows are related and they are adjacent pixels and they will convey information about the image.

Example 2: Text sequences are also able to be represented as a matrix used in NLP. The text sequences are usually one-hot encoded to a dictionary which means the vector does not convey any ordering of the original document and hence can not be represented as a data matrix.

Problem 2

a) $E(X-u)^2 = EX^2 - 2uEX + u^2 = EX^2 - 2u^2 + u^2 = EX^2 - u^2$
This equation calculates the variance of a R.V, expressed as a combination of the expectation of the R.V and its mean.

b) Let $X_1, ..., X_n$ be random samples from a distribution:

$$\hat{u} = \frac{X_1 + ... + X_n}{n}$$
$$E[\hat{u}] = \frac{1}{n}(E[X_1] + ... + E[X_n])$$
$$E[\hat{u}] = \frac{1}{n}(u + ... + u) = u$$

This equation shows that the estimator of sample mean is an unbiased estimator of the population mean.

c) Using the assumption from above, assume that the R.V is normally distributed, we know that the distribution of the sum of R.V is the distribution of the R.V. Hence we can compute it's variance.

$$Var[\hat{u}] = Var[\frac{X_1 + ... + X_n}{n}]$$
$$Var[\hat{u}] = \frac{1}{n^2}Var[X_1 + ... + X_n]$$
$$Var[\hat{u}] = \frac{1}{n^2}(\sigma^2 + ... + \sigma^2) = \frac{n}{n^2}(\sigma^2 + ... + \sigma^2) = \frac{\sigma^2}{n}$$

This equation shows the variance of the random sample from a distribution.

Problem 3

a) Bernoulli:

- Probability Mass Function:
  $f(x) = p^x(1-p)^{1-x}; x = 0, 1; p \in (0, 1)$
- Cumulative Distribution Function:

$$\begin{cases} 0 & \text{for } x < 0 \\ 1-p & \text{for } x \in [0,1) \\ 1 & \text{otherwise} \end{cases}$$

- Mean:
  p
- Variance:
  $p(1-p)$

b) Binomial:

- Probability Mass Function:
  $f(x) = \binom{n}{x}p^x(1-p)^{1-x}; x = 1, ..., n$
- Cumulative Distribution Function:
  $F(x) = \sum_{i=0}^{x} \binom{n}{i}p^i(1-p)^{n-i}$
- Mean:
  np
- Variance:
  $np(1-p)$

c) Uniform:

- Probability Mass Function:
  $f(x) = \frac{1}{b-a}$
- Cumulative Distribution Function:
  $F(x) = \frac{x-a}{b-a}$
- Mean:
  $\frac{a+b}{2}$
- Variance:
  $\frac{(b-a)^2}{12}$

d) Normal:

- Probability Mass Function:
  $f(x;u,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-(x-u)^2}{2\sigma^2}}$
- Cumulative Distribution Function: (assume standard normal)
  $F(x;u,\sigma^2) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{\frac{-t^2}{2}}dt$
- Mean:
  u
- Variance:
  σ^2

e) Exponential:

- Probability Mass Function:
  $f(x;\lambda) = \lambda e^{-\lambda x}; x \geq 0$
- Cumulative Distribution Function:
  $F(x;\lambda) = 1 - e^{-\lambda x}$
- Mean:
  $\frac{1}{\lambda}$
- Variance:
  $\frac{1}{\lambda^2}$

f) Power Law:

- Probability Mass Function:
  $f(x;\alpha,x_{min}) = \frac{\alpha-1}{x_{min}}(\frac{x}{x_{min}})^{-\alpha}$
- Cumulative Distribution Function:
  $F(x;\alpha,x_{min}) = (\frac{x}{x_{min}})^{-\alpha-1}$
- Mean:
  $\int_{x_{min}}^{\infty} xf(x;\alpha,x_{min})dx$, could be infinite.
- Variance:
  $\int_{x_{min}}^{\infty}(x-E(X))^2 f(x;\alpha,x_{min})dx$, could be infinite.
- Notes:
  The mean and variance of a Power Law (Pareto Distribution) is meaningless as it is a heavily tailed distribution. Hence mean and variance is only finite if the shape is sufficiently large.

Problem 4

The link to the code can be found on the github page: *https://github.com/feedlord18/CSDS313.* Please let me know if the repo is not available.

a) Mean and Variance of 10 Samples and 100 Samples of All 6 Distributions:

```
samp_10 =

   1.0e+03 *

    0.0042    0.0400    0.0639   -0.0088    0.0054    0.0188
    0.0145    1.2985    3.3082    0.0663    0.0241    0.2869


samp_100 =

    0.4200    4.0000    6.3870   -0.8843    0.5425    1.8756
    0.2436    2.5200    3.9024    3.0011    0.2858    1.9988
```

Figures Generated From Matlab

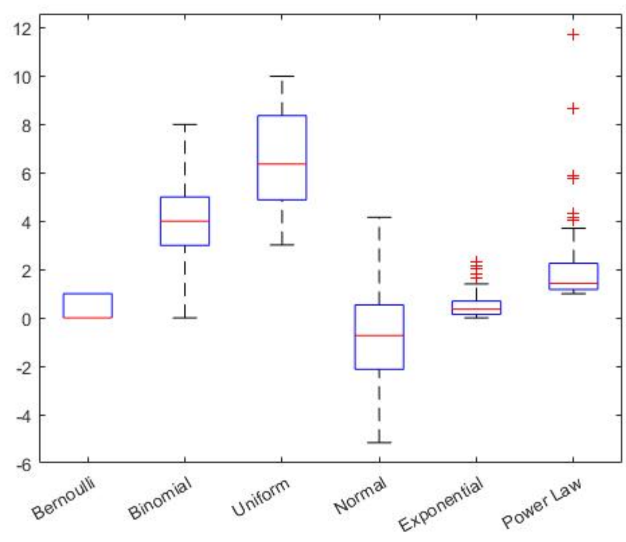

Figure 1: Box Plot of All 6 Sampled Distributions.

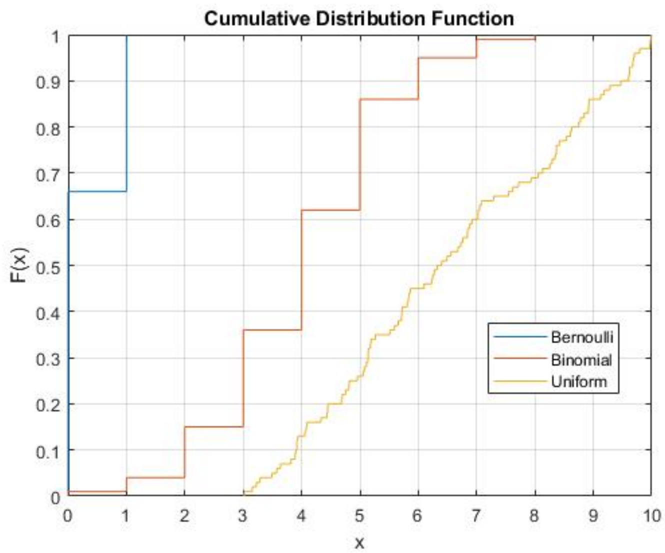

Figure 2: CDF of First Three Distributions.

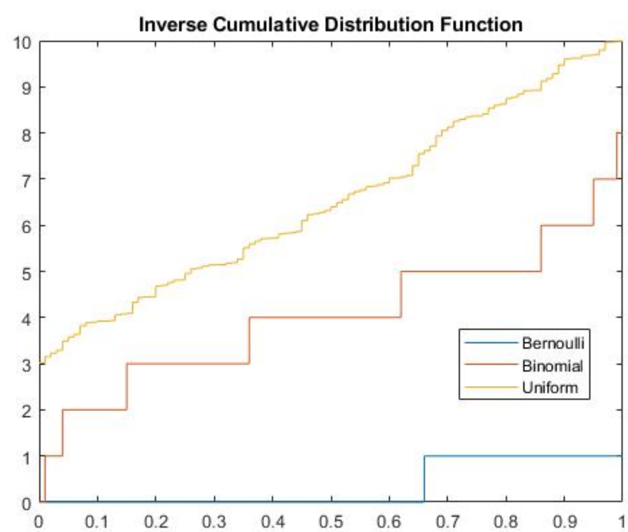

Figure 3: Inverse CDF of First Three Distributions.

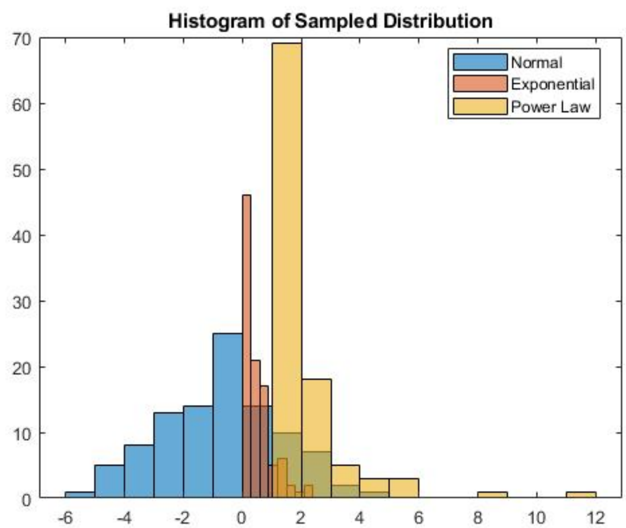

Figure 4: Histogram of Last Three Distributions.