

Introduction to Random Variables

Part 1.

Using MATLAB, generate a vector \mathbf{X} that contains 100,000 random numbers whose elements are normally distributed with mean $\mu = 0$ and variance $\sigma^2 = 1$. (Hint: Use subroutine "randn")

a) Plot the histogram of these 100,000 elements.



b) Calculate the sample mean and variance of the 100,000 values in \mathbf{X} . The sample mean and variance of a list of numbers $x_1, x_2, x_3, \dots, x_N$ are respectively

$$\mu_s = \frac{1}{M} \sum_{n=1}^M x_n \quad \text{and} \quad \sigma_s^2 = \frac{1}{M-1} \sum_{n=1}^M (x_n - \mu_s)^2.$$

Compare them with the theoretical values.



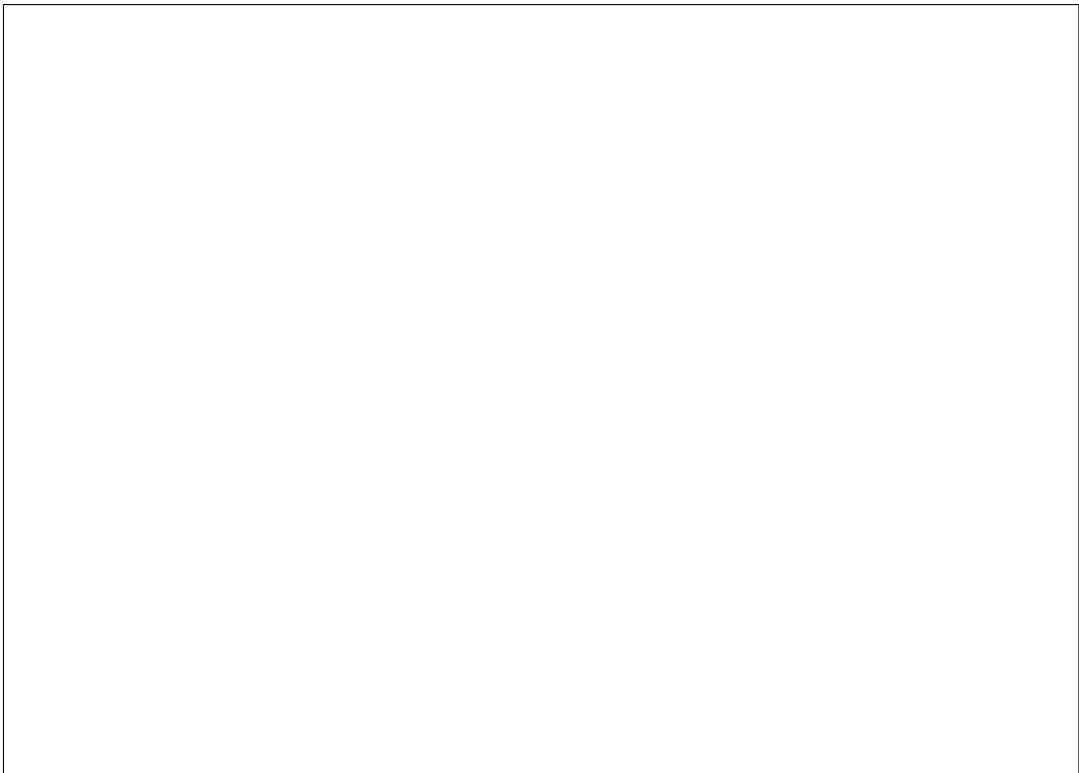
Part 2.

Generate a vector \mathbf{Y} that contains 100,000 random numbers whose elements are normally distributed with mean $\mu = 0$ and variance $\sigma^2 = a$. Suppose we used the transformation rule

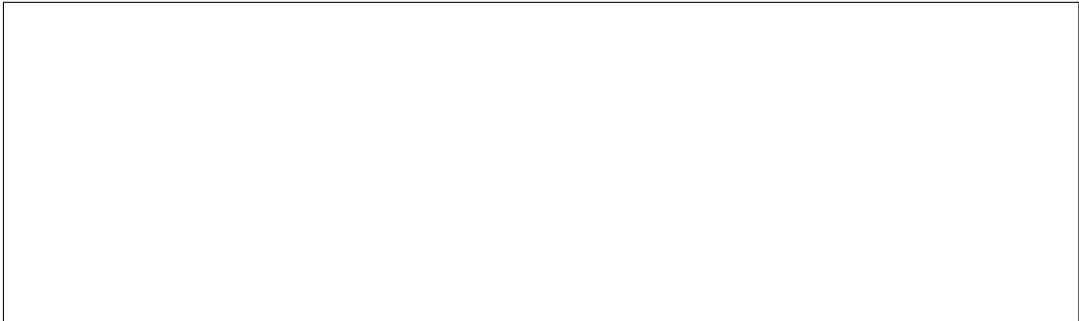
$$y_i = \ln x_i$$

where x_i and y_i are elements in \mathbf{X} and \mathbf{Y} respectively, and $i = 1, 2, \dots, 100,000$

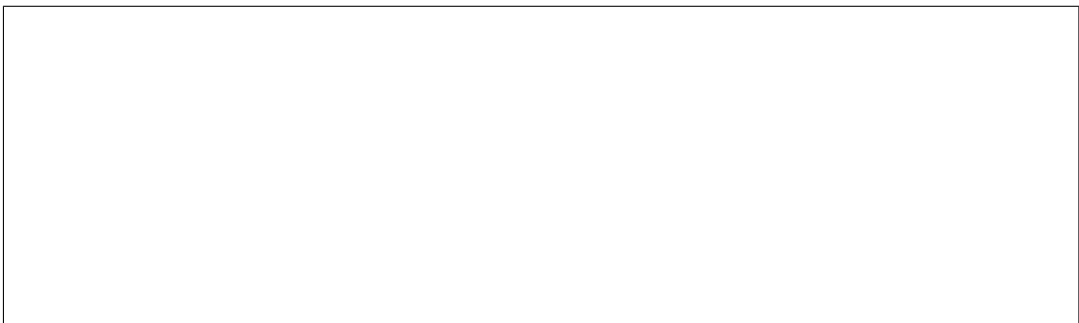
a) For $a = 1$, plot the histogram of the data in vector \mathbf{X} . What is the distribution of the new random numbers in vector \mathbf{X} ?



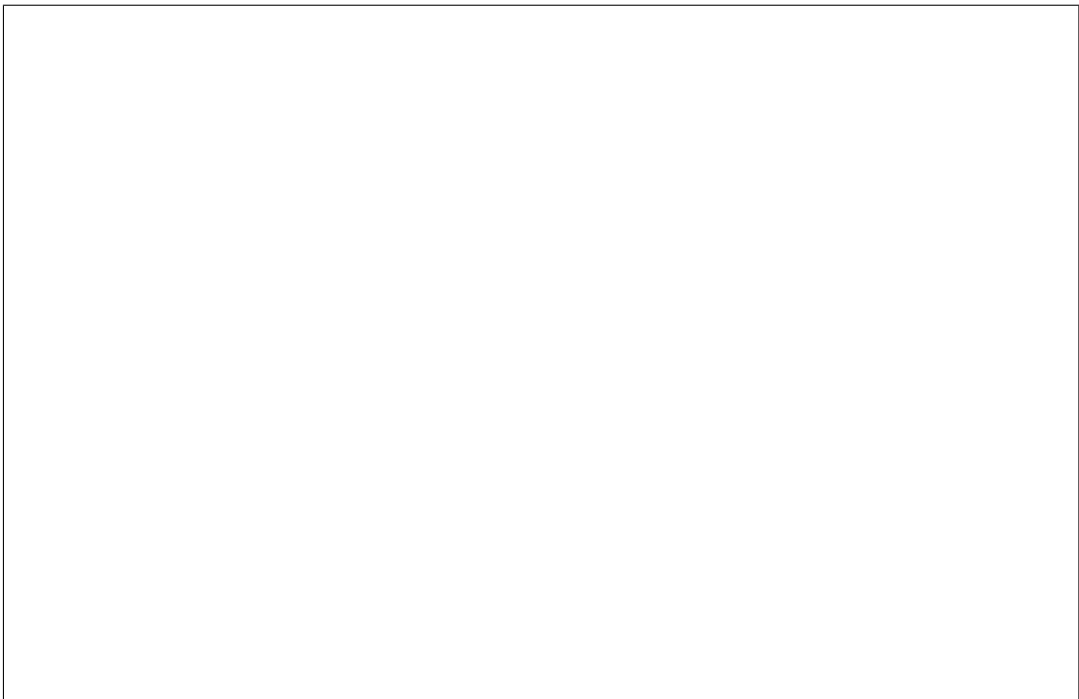
b) Calculate the median of the values in \mathbf{X} and \mathbf{Y} . The median, M , of any random variable \mathbf{X} is defined by $P(\mathbf{X} \geq M) = 0.5$.



c) Calculate the sample mean of the values in \mathbf{X} and \mathbf{Y} . Discuss any similarities and differences in your two answers in this part and part **b**.



d) Plot the sample mean μ_s as a function of a , where $a = 0, 0.1, 0.2, \dots, 10$. What can you conclude from the plot of the sample mean?



Note: This part will not be marked !

The theoretical mean of the random variable \mathbf{X} can be calculated as follow:

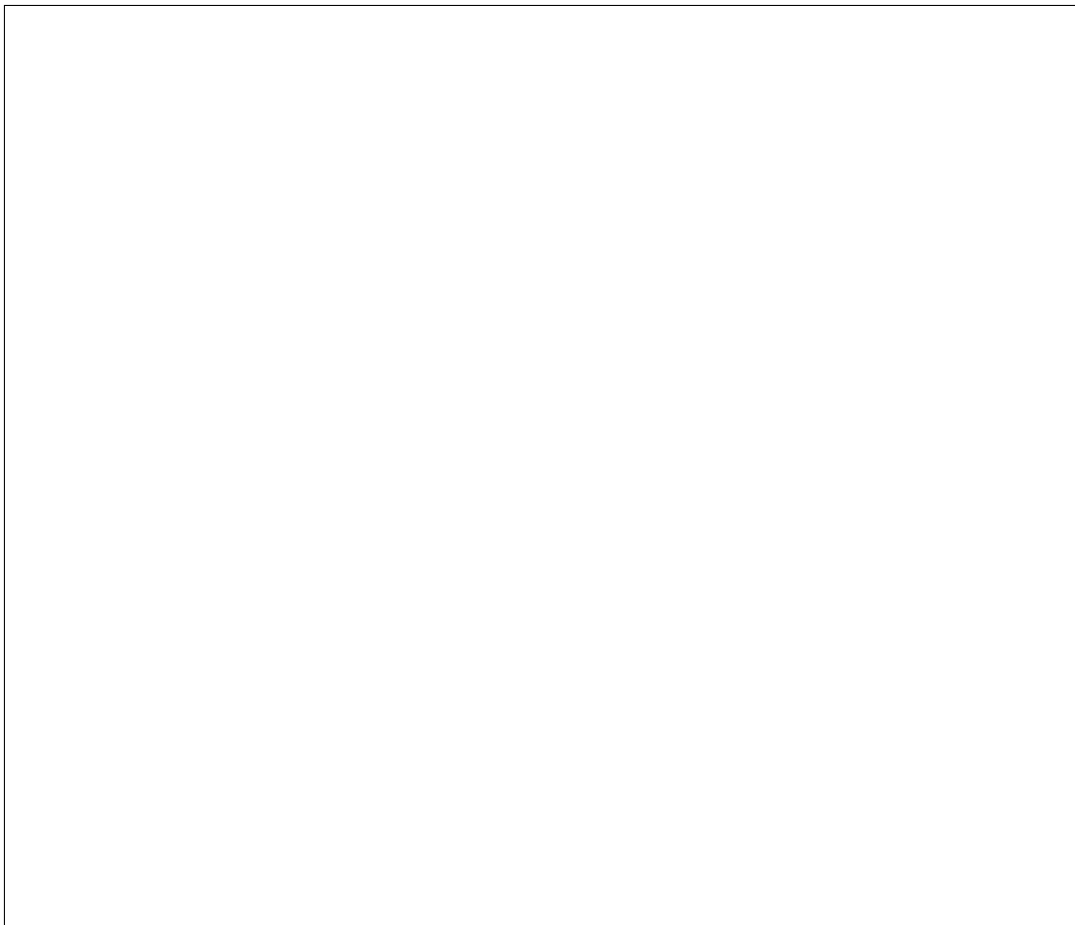
$$E[X] = \int_{-\infty}^{+\infty} x f_{\mathbf{X}}(x) dx = \int_{-\infty}^{+\infty} e^y f_{\mathbf{Y}}(y) dy = \int_{-\infty}^{+\infty} e^y \frac{e^{-\frac{y^2}{2a}}}{\sqrt{2\pi a}} dy = e^{\frac{a}{2}} \underbrace{\int_{-\infty}^{+\infty} \frac{e^{-\frac{(y-a)^2}{2a}}}{\sqrt{2\pi a}}}_{=1} dy = e^{\frac{a}{2}}$$

You should expect an exponentially graph for the simulation part, which is the case in the plot shown above.

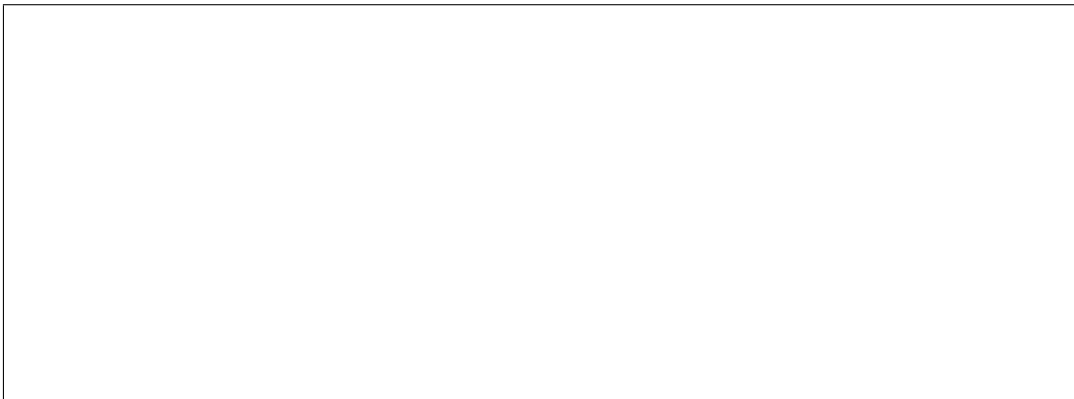
Part 3.

Generate a vector \mathbf{X} that contains 100,000 random numbers whose elements are normally distributed with mean $\mu = 5$ and variance $\sigma^2 = 10$. Note that subroutine `rand` in MATLAB generates only Gaussian random numbers with mean $\mu = 0$ and variance $\sigma^2 = 1$.

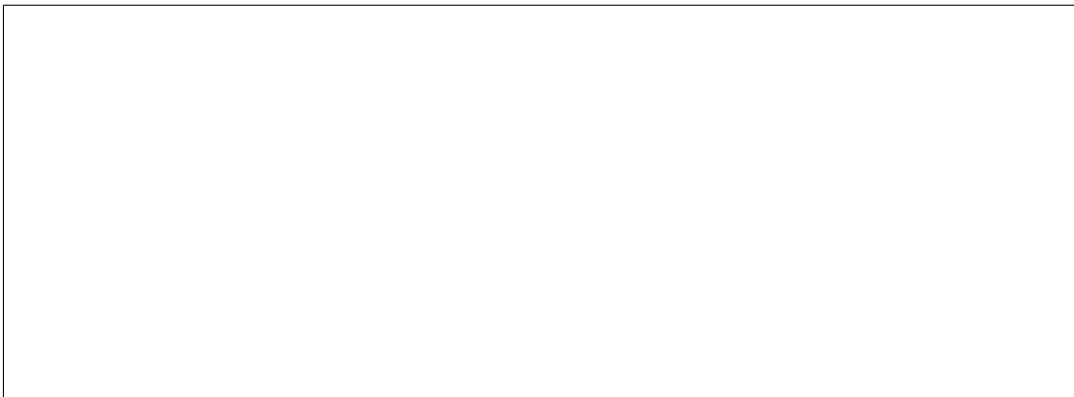
a) Plot the histogram of the generated Gaussian random numbers in \mathbf{X} .



b) Calculate the sample mean and variance of the Gaussian random numbers in \mathbf{X} . Compare it to the theoretical values.



c) Calculate $P(\mathbf{X} \leq \pi)$.



Part 4.

In electrical engineering it is customary to work with the \mathbf{Q} -function which is which is defined by

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{t^2}{2}} dt$$

Write MATLAB code to compute the \mathbf{Q} -function and plot it.



Part 5.

Generate two vectors \mathbf{X} and \mathbf{Y} which contains 100,000 random numbers whose elements are normally distributed with $\mu = 0$ and $\sigma^2 = 1$. Suppose we want to generate a new vector \mathbf{Z} that contains 100,000 of complex Gaussian random numbers using the transformation rule

$$z_i = x_i + \mathbf{j}y_i$$

where x_i, y_i and z_i are elements in \mathbf{X} , \mathbf{Y} and \mathbf{Z} respectively, and $\mathbf{j} = \sqrt{-1}$.

a) Plot a scatter-plot of the generated random numbers in \mathbf{Z} .



b) Generate a new vector \mathbf{R} whose elements are defined by the following transformation

$$r_i = |z_i|$$

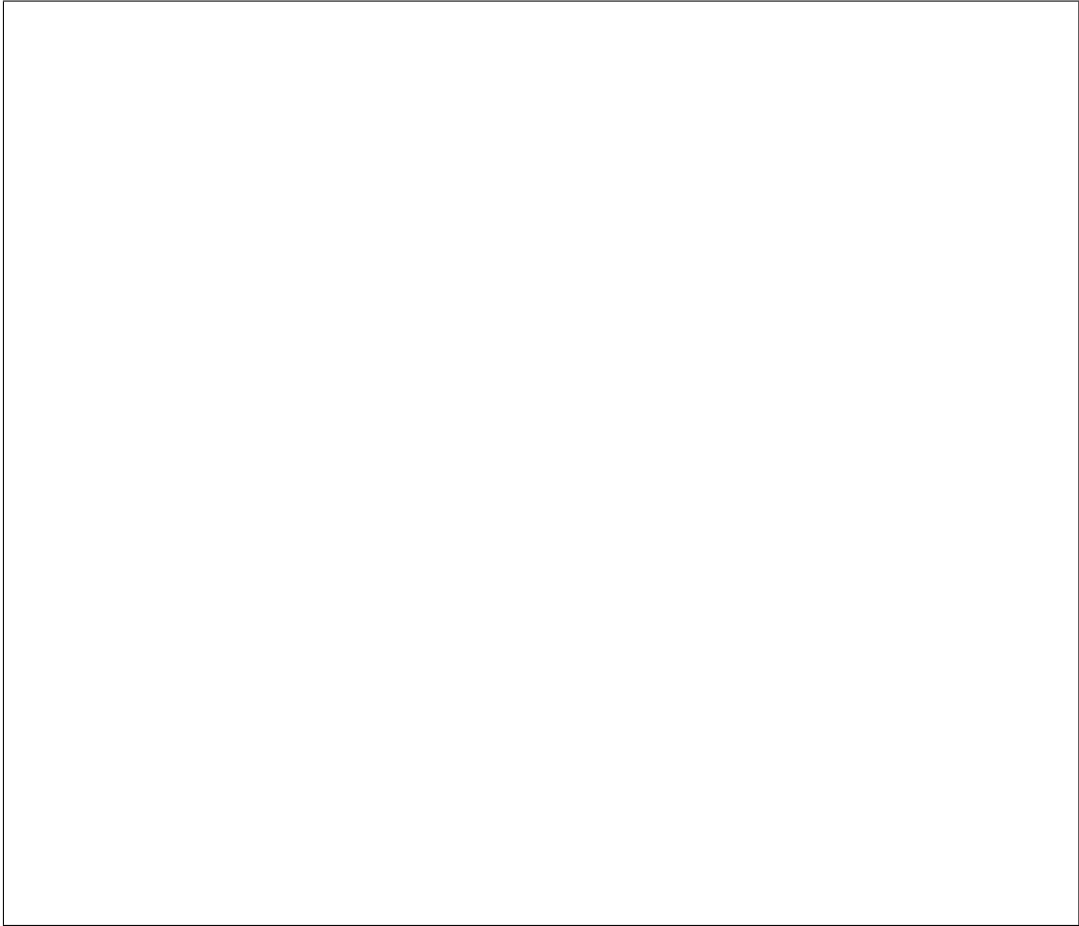
(i.e. the amplitude of the complex random number z_i). Plot the histogram of the new random numbers in \mathbf{R} . What is the distribution of the new random numbers in \mathbf{R} (i.e. Gaussian, Exponential, etc...).



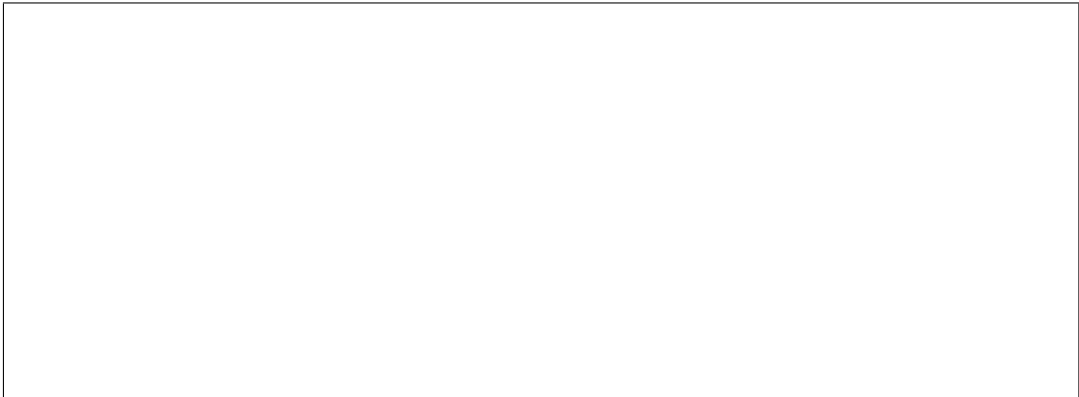
c) Generate a new vector Θ whose elements are defined by the following transformation

$$\theta_i = \tan^{-1}\left(\frac{y_i}{x_i}\right)$$

(i.e. the phase of the complex random number x_i). Plot the histogram of the data in vector Θ . What is the distribution of the new random numbers in Θ .



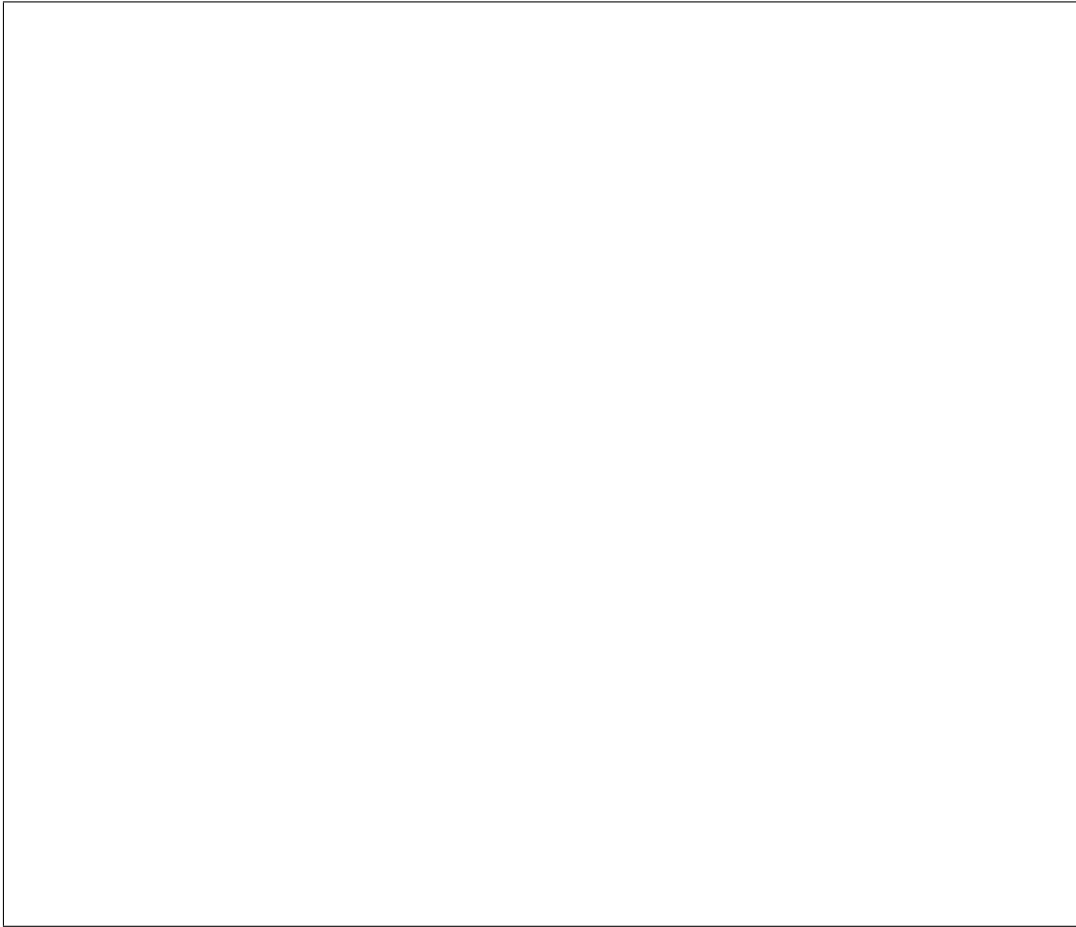
d) Using MATLAB, calculate the sample mean μ_s and variance σ_s^2 of the values in **R**. Compare the results with the theoretical values.



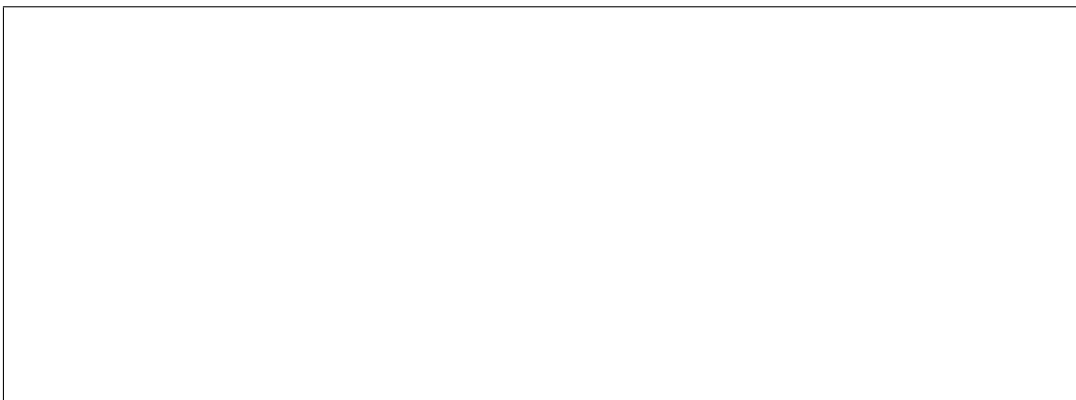
Part 6.

The binomial and Poisson are the most important discrete random variables. The "binopdf" and "poisspdf" are the MATLAB functions for the probability mass functions of these random variables.

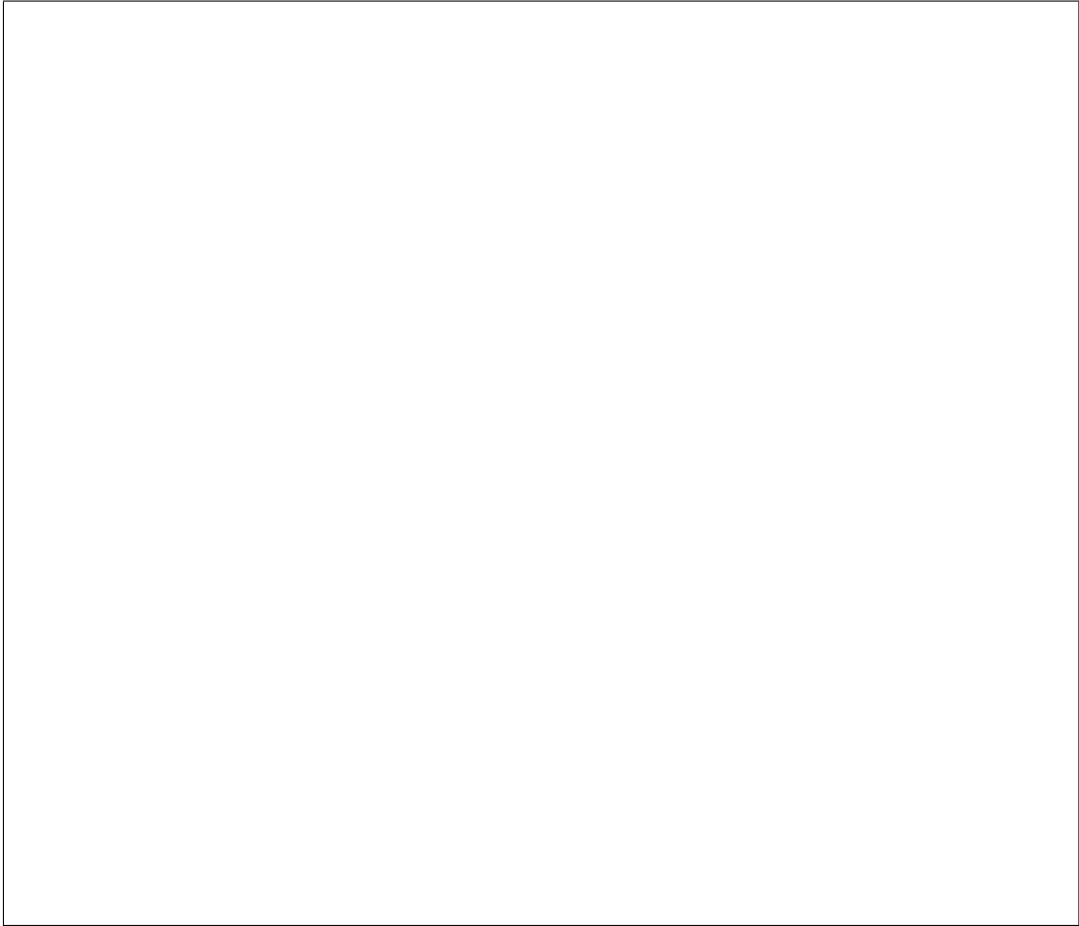
a) Plot the probability mass function of a binomial random variable \mathbf{X} with success probability $p = 0.2$ and $n = 10$.



b) The cumulative distribution function is useful for computing probabilities that a binomial random variable lies in a range. For the same random variable in a, compute the probability that \mathbf{X} lies between 2 and 8, i.e. $P(2 \leq \mathbf{X} \leq 8)$.

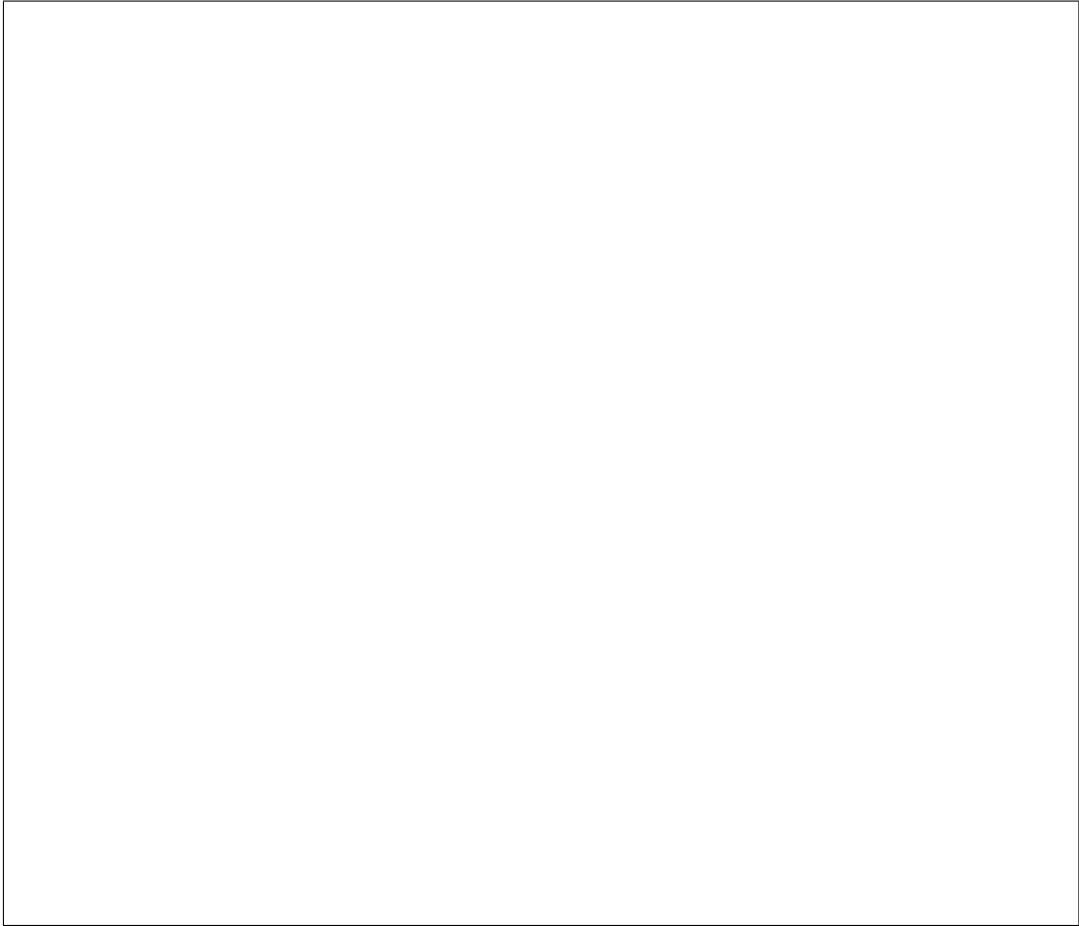


c) Plot the CDF of \mathbf{X} , i.e. $F_{\mathbf{X}}(x)$.

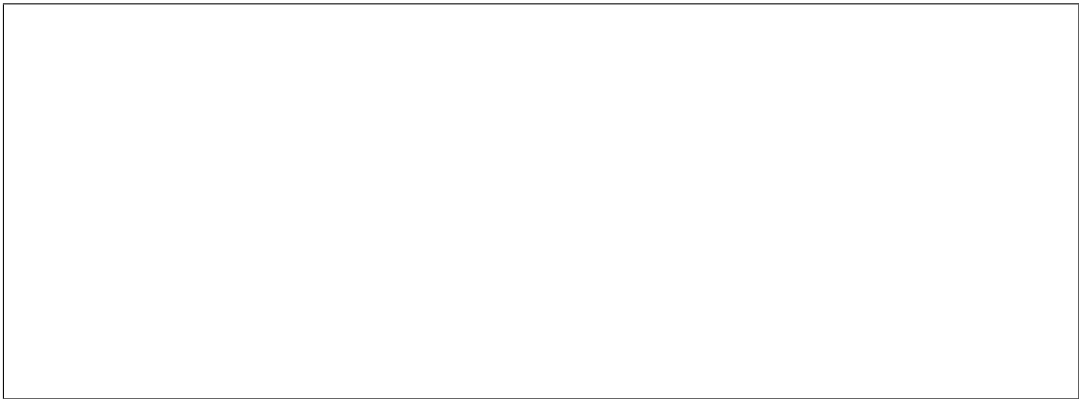


d) As $n \rightarrow \infty$, while keeping the product np fixed, call it λ , what distribution do you expect the random variable \mathbf{X} to have?. What is the probability mass function (PMF) of \mathbf{X} when $n \rightarrow \infty$ (i.e. $P(\mathbf{X} = k)$)?

(i) Plot the PMF of \mathbf{X} for $p = 10^{-5}$ and $n = 1000$.



(ii) Calculate the mean and variance of \mathbf{X} . Compare them with the theoretical values.



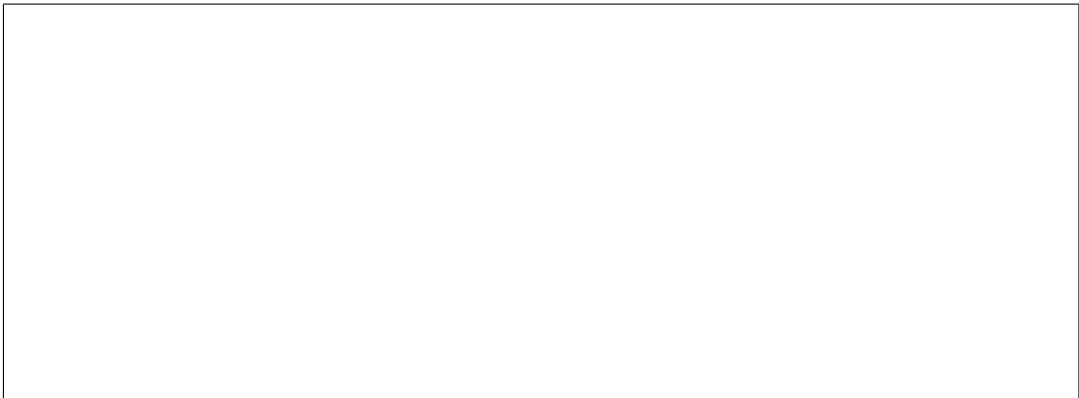
Part 7. (A little experiment with IMAGES)

a) Suppose $F_{\mathbf{X}}(x)$ and $f_{\mathbf{X}}(x)$ are respectively the CDF and pdf of a continuous random variable \mathbf{X} . Let $\mathbf{Y} = F_{\mathbf{X}}(x)$. Determine $f_{\mathbf{Y}}(y)$, i.e. the PDF of \mathbf{Y} .



b) Download one of the images from "www.site.uottawa.ca/~wabed" and display it in MATLAB. (Hint: Use subroutines "imread" and "imshow").

c) Plot the relative frequency of the pixel values in the image. Notice that gray pixel values range between $[0, 255]$ where gray= 0 and white= 255. (Warning: MATLAB stores the pixel values in the image when using "imread" in a "uint8 array". In order to be able to calculate the relative frequency of pixel values you have to convert them to "double array". To do that, use the subroutine "double").



d) Treat the relative frequency above as the PMF of a random variable \mathbf{X} . Calculate the CDF of \mathbf{X} and plot it. (Hint: Use "for" statement)



e) Generate a new image by changing the value x of each pixel in the above image to $[F_{\mathbf{x}}(x) \times 255]$, where $[.]$ is the round operation. Display the new modified image, and plot the histogram of it. Explain what you see!

