Alex Nguyen

Computer Project # 1 Report

Date: February 11, 2021

Problem 1

First, I generated N random experiments for n i.i.d. Bernoulli random variables to be summed. This gave me several sets of n Bernoulli random variables.

$$S_n = \sum_{i=1}^{n} X_i = X_1 + X_2 + ... + X_n$$

Now, I had a sequence of N Bernoulli random variable sums which was used to construct the probability density function (pdf) using a histogram. Then, the approximate sum of a Gaussian random variable was obtained using the equation for the central limit theorem. The following equation gives a zero-mean, unit-variance Gaussian (Normal) random variable for a sample sequence.

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

Finally, we were asked to produce the same approximate Gaussian (Normal) random variable with different n's. I chose new n samples to be summed which was less than the one specified in the problem.

Below, I've included the `MATLAB` generated plots. The first plot contains a pmf for a Bernoulli random variable. The second plot is the pdf for the sum of several sets of $n = 100$ i.i.d. Bernoulli random variables. The third plot is the approximate sum variable as a Gaussian (Normal) random variable using $n = 100$ i.i.d. Bernoulli random variables. The last plot is the same Gaussian random variable figure but with two other n sample values to be summed. These different n values are compared with each other and to the theoretical distribution.

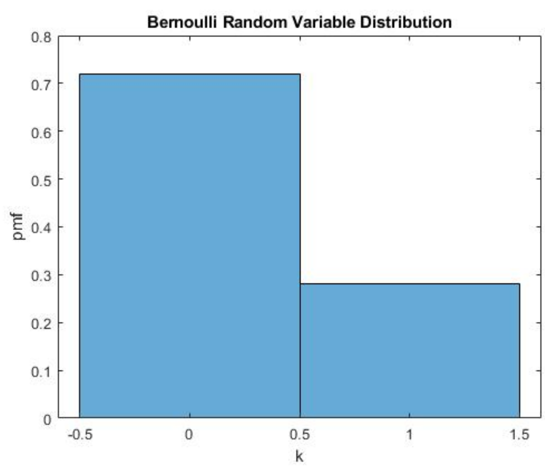

Figure 1: Bernoulli Random Variable PMF

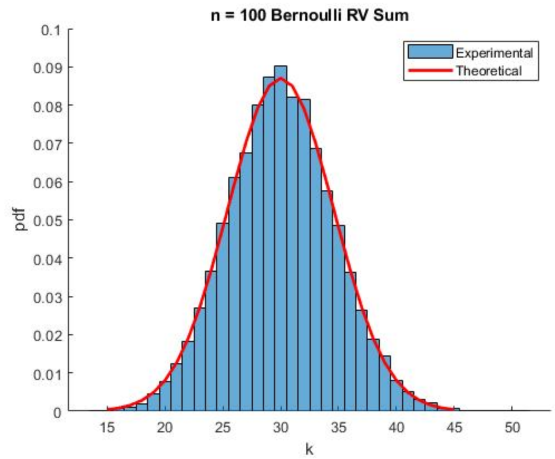

Figure 2: Bernoulli Random Variable Probability Density Function (PDF) of the Sum

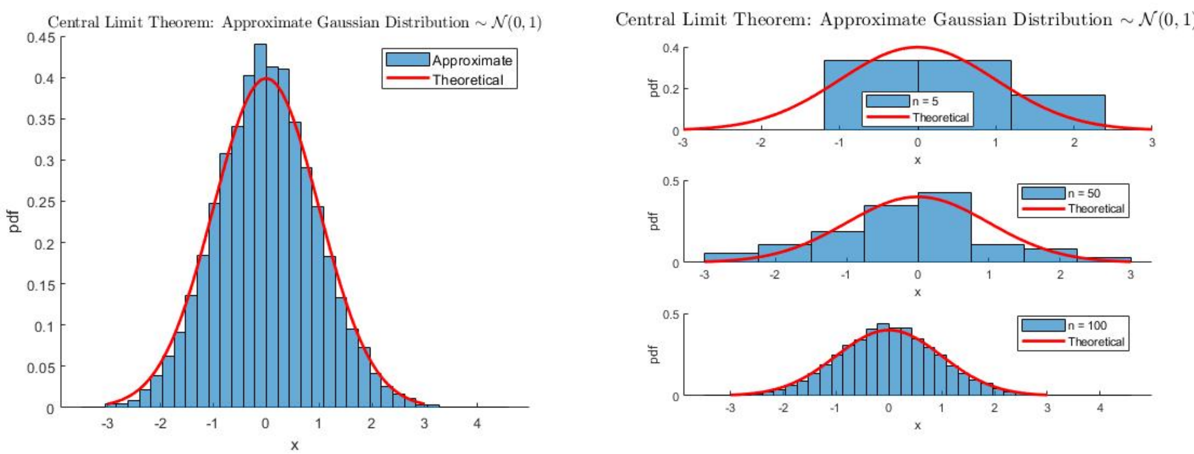

Figure 3: Approximate Gaussian Distribution for n Bernoulli RVs

Problem 2

This section's approach is almost identical to the last section's. First, I generated N random experiments for n i.i.d. Poisson random variables to be summed. This gave me several sets of n Poisson random variables.

$$S_n = \sum_{i=1}^{n} X_i = X_1 + X_2 + ... + X_n$$

Now, I had a sequence of N Poisson random variable sums which was used to construct the probability density function (pdf) using a histogram. Then, the approximate sum of a Gaussian random variable was obtained using the equation for the central limit theorem. The following equation gives a zero-mean, unit-variance Gaussian (Normal) random variable for a sample sequence.

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

Finally, we were asked to produce the same approximate Gaussian (Normal) random variable with different n's. I chose new n samples to be summed which was less than the one specified in the problem.

Below, I've included the `MATLAB` generated plots. The first plot contains a pmf for a Poisson random variable. The second plot is the pdf for the sum of several sets of $n = 100$ i.i.d. Poisson random variables. The third plot is the approximate sum variable as a Gaussian (Normal) random variable using $n = 100$ i.i.d. Poisson random variables. The last plot is the same Gaussian random variable figure but with two other n sample values to be summed. These different n values are compared with each other and to the theoretical distribution.

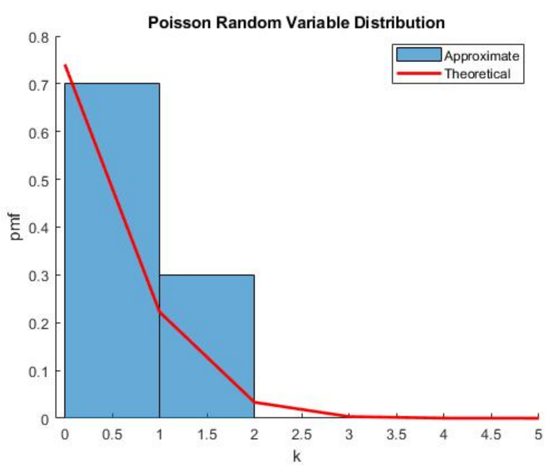

Figure 4: Poisson Random Variable PMF

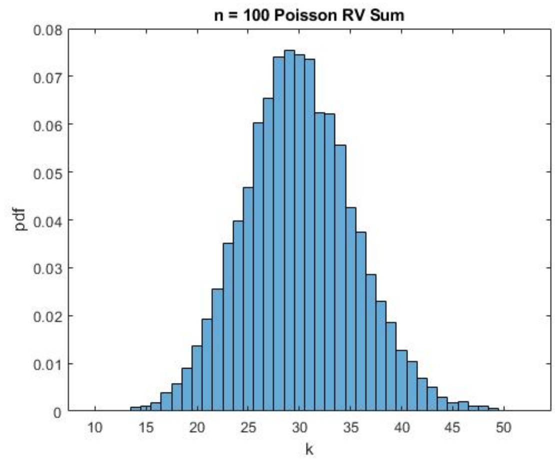

Figure 5: Poisson Random Variable Probability Density Function (PDF) of the Sum

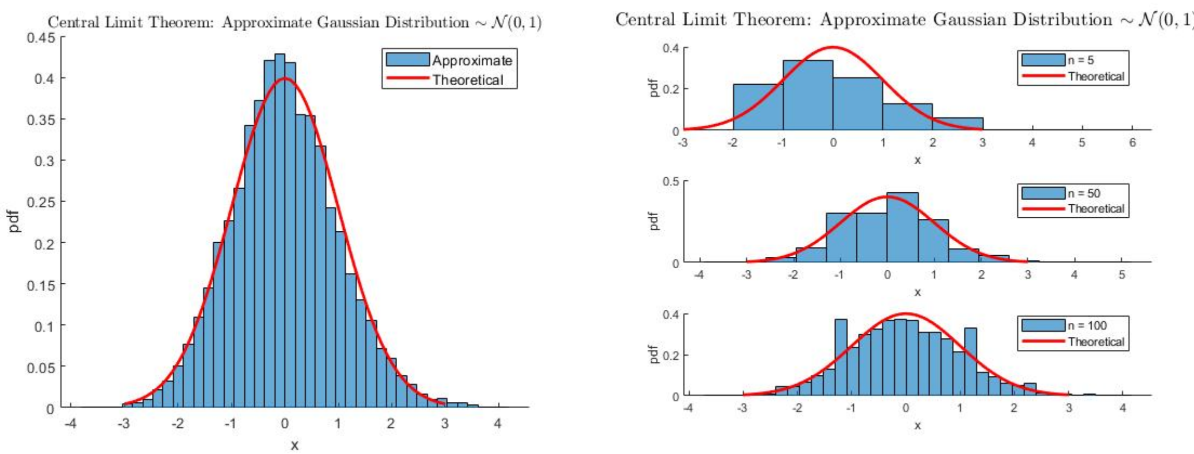

Figure 6: Approximate Gaussian Distribution for n Poisson RVs

Problem 3

First, I created N sample realizations of uniformly distributed random variables. This sequence was then used to generate a zero-mean, unit-variance Gaussian (Normal) distribution which could be shifted to create a two-mean, three-variance Gaussian (Normal) distribution.

The zero-mean, unit-variance Gaussian (Normal) distribution is generated by an independent Rayleigh (R) and uniform random variable (Θ). The two random variables were transformed transformed from the typical Gaussian (Normal) distribution then multiplied together.

$$f_{R,\Theta}(r,\theta) = \frac{1}{2\pi} r e^{-r^2/2} = f_R(r) f_\Theta(\theta)$$

First, we generate R^2 as an exponential random variable with parameter $\lambda = \frac{1}{2}$. Second, we generate Θ as a uniformly distributed random variable the region $[0,\, 2\pi]$. Then, we substitute these values into the equation below to obtain an independent zero-mean, unit-variance Gaussian (Normal) distribution.

$$X = R\cos(\Theta)$$

The figures generated below follow the same methodology. The first plot shows the two-mean, three-variance pdf for a Gaussian (Normal) random variable which was shifted from a zero-mean, unit-variance Gaussian (Normal) distribution. The second plot shows the empirical CDF of the shifted ($\mathcal{N}(2,3)$) Gaussian (Normal) distribution which was generated from a uniformly distributed random variable from (0, 1).

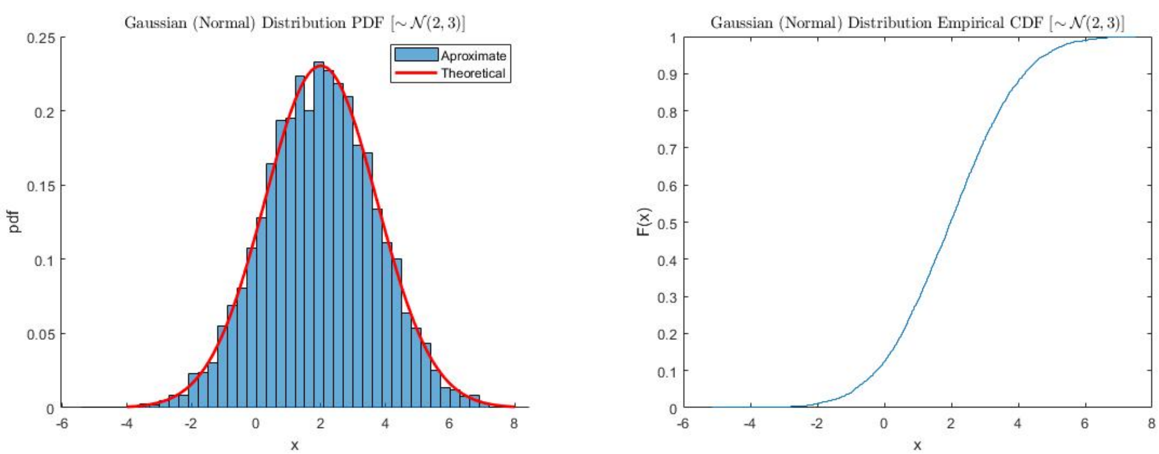

Figure 7: Two-Mean, Three-Variance Gaussian Distributions Generated from $\mathcal{U}(0,1)$

Problem 4

The approach used the definition of the Weak Law of Large Numbers which is the following:

$$\lim_{x \to \infty} P[|M_n - \mu| < \epsilon] = 1$$

Put another way, the sample mean will converge in probability to the true mean with high likelihood for a large enough fixed value of n samples. Each particular sequence of sample means approach μ and stays there.

The results below are for scenarios of n sample realizations which varied from a small to large number. The accuracy of approximating the sample mean as the true mean increased with the number of samples n. This trend can verify the law of large numbers which used two different random variables: Bernoulli and binomial random variables.

```
Bernoulli Random Variable with Expected Value, E[X] = 0.40:
   i) Low Number of Trials: Mn = 0.200 (N = 5)
  ii) Medium Number of Trials: Mn = 0.430 (N = 100)
 iii) High Number of Trials: Mn = 0.399 (N = 10000)

Binomial Random Variable with Expected Value, E[X] = 1.80:
   i) Low Number of Trials: Mn = 1.600 (N = 5)
  ii) Medium Number of Trials : Mn = 1.910 (N = 100)
 iii) High Number of Trials: Mn = 1.797 (N = 10000)

The law of large numbers (LLN) states the sample mean (Mn) will converge in probability to the expected
the number of samples in a sequence increases. Above, the two examples illustrate this phenomena. Thus,
```

Problem 5

The approach used the standard expectation definition for variance.

$$\sigma_x^2 = E[X^2] - (E[X])^2$$

Rearranging this equation, we can find the expected value (or mean) of X^2.

$$E[X^2] = \sigma_x^2 + (E[X])^2$$

Here, the random variable is a zero-mean, unit-variance Gaussian random variable which meant (theoretically) the the mean of $E[X^2]$ should be one. The estimated mean of X^2 approached one every simulation but never exactly equaled exactly one. The parameter used to estimate $E[X^2]$ varied to show convergence with larger number of sample realizations. Below are the results of running one particular experiment.

```
Estimated Expected value of X^2 (or E[X^2]) with mu = 1:
   i) Low Number of Samples: E[X^2] = 2.1795 (N = 10)
  ii) Medium Number of Samples: E[X^2] = 0.9658 (N = 5000)
 iii) High Number of Samples: E[X^2] = 1.0020 (N = 100000)
```