ECE537: Lab 2 Report

In the first part of the lab, we will be creating and analyzing joint Gaussian distributions as a part of which we will be extracting marginal densities of the joint random variables. In the second part, we will be simulating the central limit theorem and the law of large numbers with uniform (univariate) random variables and test for the relevant convergence cirtieria.

```
using Random, Distributions, StatsBase, StatsPlots, LinearAlgebra, DataFrames,
LaTeXStrings, PlutoUI
```

```
PlotlyBackend()
```

1.Simulating Bivariate Gaussian Distributions

The multivariate normal (or Gaussian) distribution is a multidimensional generalization of the normal distribution. The probability density function of a n-dimensional multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and (positive definite) covariance matrix $\boldsymbol{\Sigma}$ is:

$$f_{\mathbf{X}}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{(2\pi)^n \det(\boldsymbol{\Sigma})}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\intercal \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right).$$

We will proceed with using its implementation provided as the `MvNormal` distribution struct or type, with $n = 2$ to make it a bivariate distribution.

For with individual means μ_1, μ_2, standard deviations σ_1, σ_2, and correlation coefficient ρ, the bivariate normal distribution can be defined as:

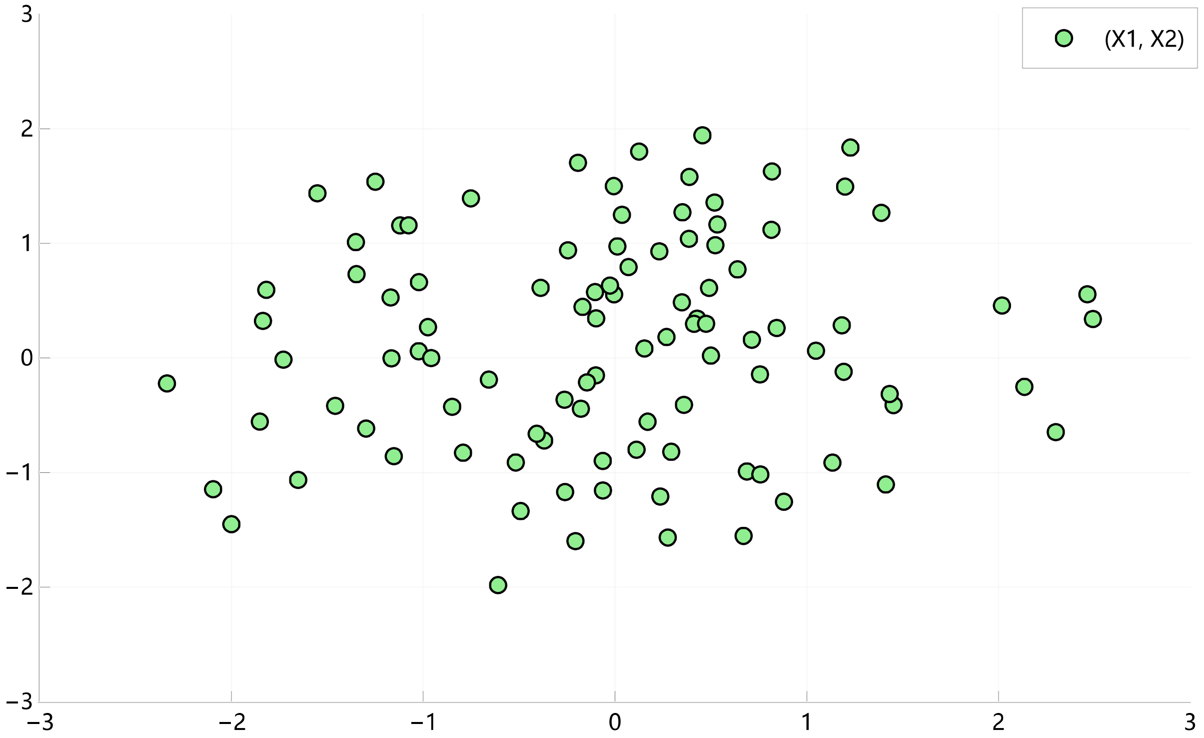
$$X \sim \mathcal{N}\left(\boldsymbol{\mu} = \begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix}, \boldsymbol{\Sigma} = \begin{bmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{bmatrix}\right).$$

For the determinstic question, we can have the question 1 as follow.

Let $X = (X_1, X_2)$ consist of two Gaussian distributed random variables X_1 and X_2

Q1(a)

Assume $E(X_1) = E(X_2) = 0$ and $\sigma^2(X_1) = \sigma^2(X_2) = 1$. Let $\rho = 0$. Take $N = 100$ samples from the distribution of X and plot the scattergram (samples shown in the $X_1 - X_2$ plane).Explain your observation.



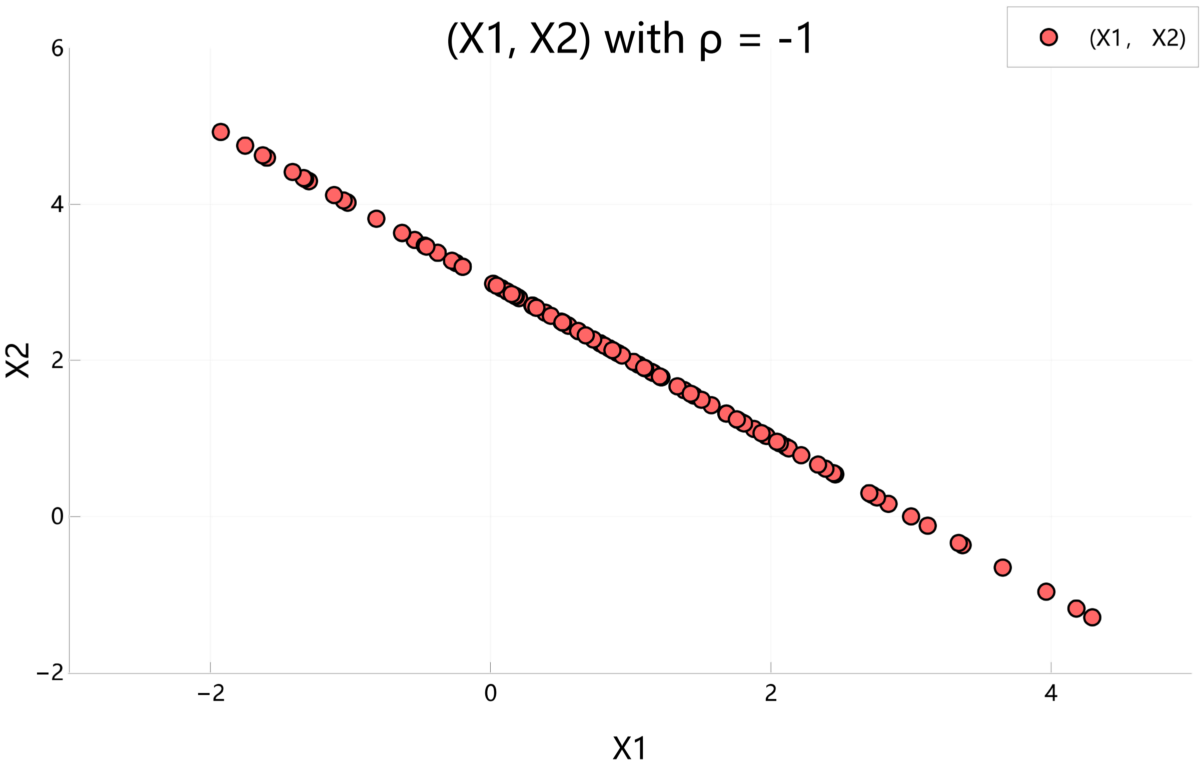
After run the cell a lot of times, we can find it centered in the original point$(0, 0)$, that is corresponding to what we expect.

Q1(b)

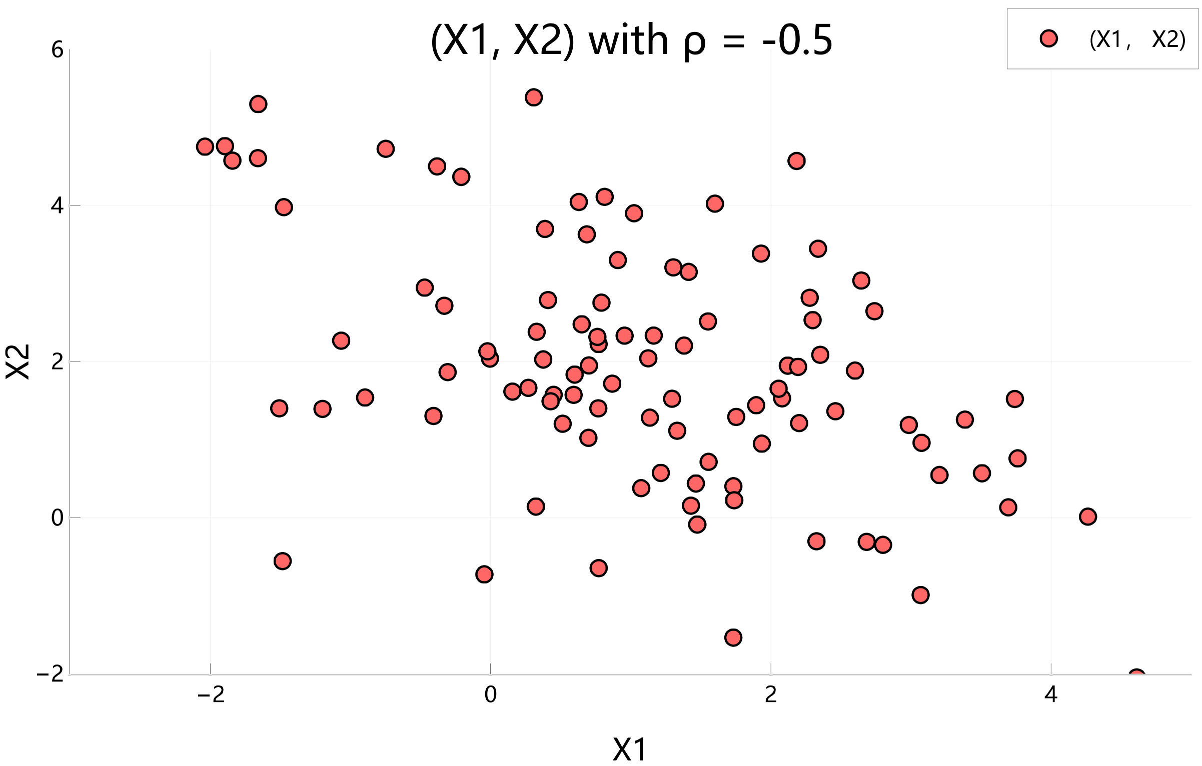
Let $E(X_1) = 1$ and $E(X_2) = 2$ and $\sigma^2(X_1) = 2$ and $\sigma^2(X_2) = 2$. For $\rho = -1, -0.5, 0, 0.5, 1$ take $N = 100$ samples from the distribution of X and plot the scattergram (five different plos). Explain your observation.

```
plot_X1X2 (generic function with 1 method)
```

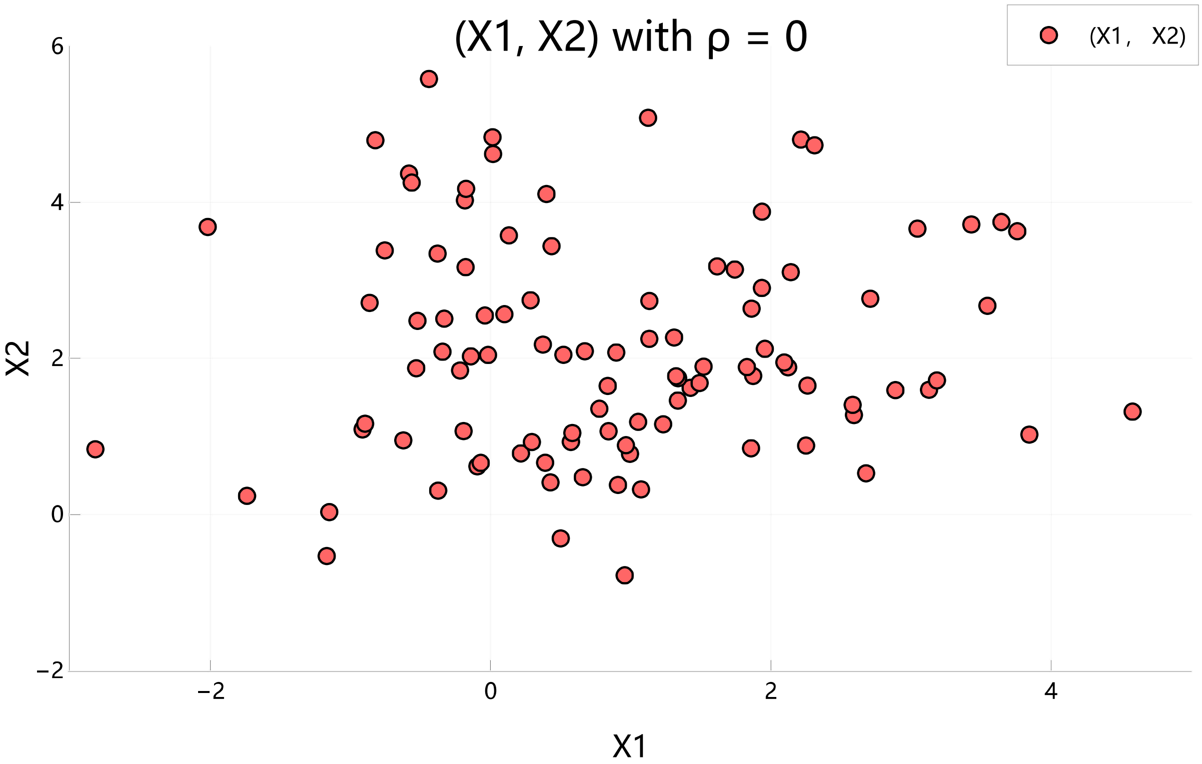
```
function plot_X1X2(E1, E2, sigma1, sigma2, rho, N)
    mvnorm = MvNormal([E1, E2], [sigma1^2 rho * sigma1 * sigma2; rho * sigma1 * sigma2 sigma2^2])
    samples = rand(mvnorm, N)
    return samples
end
```

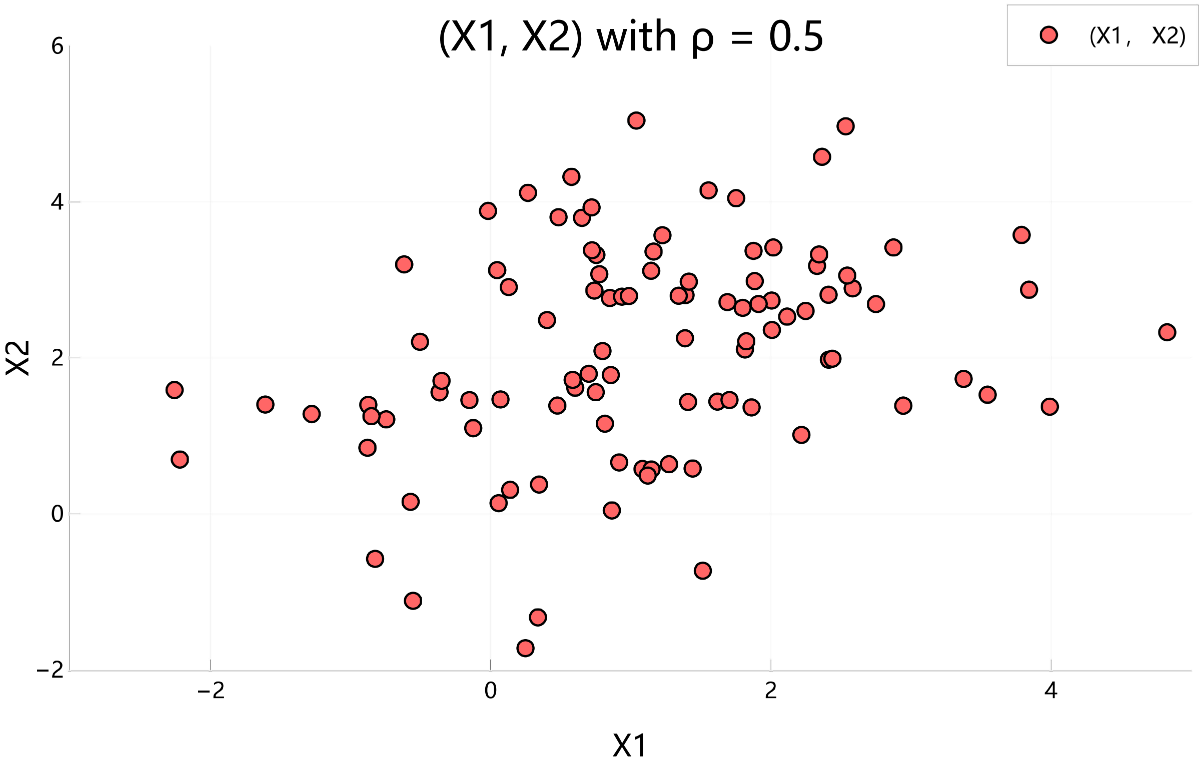


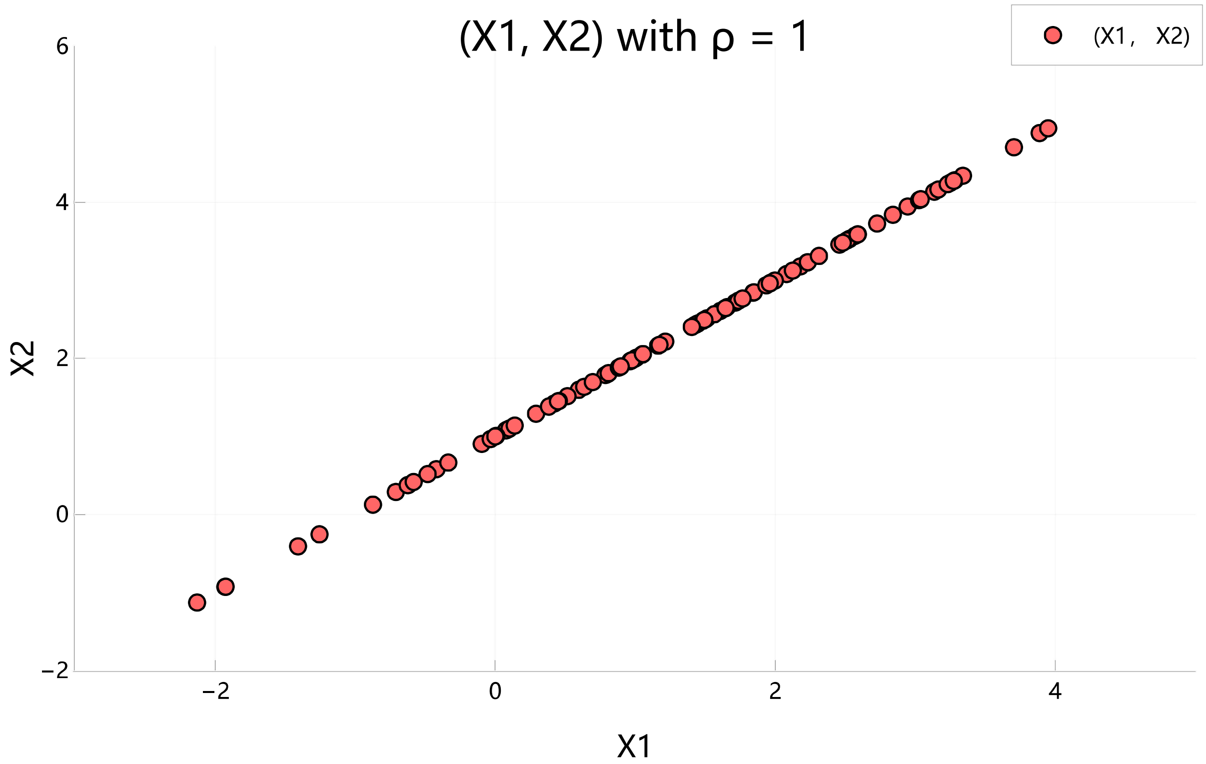
(X1, X2) with ρ = -1
(X1, X2)
X2
X1
6
4
2
0
−2
−2
0
2
4



(X1, X2) with ρ = -0.5
(X1, X2)
X2
X1
6
4
2
0
−2
−2
0
2
4


(X1, X2) with ρ = 0
(X1, X2)
X2
X1
6
4
2
0
−2
−2
0
2
4

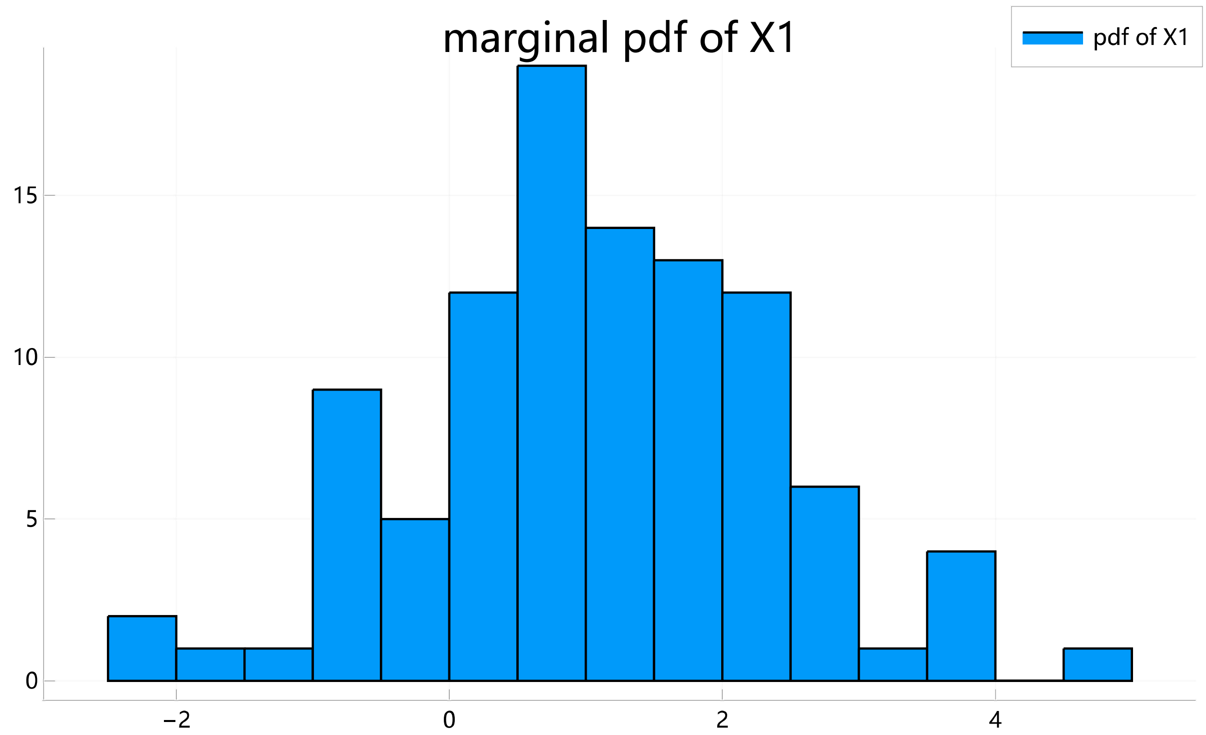

(X1, X2) with ρ = 0.5
(X1, X2)
X2
X1
6
4
2
0
−2
−2
0
2
4



From the figure, we can find that from -1 to 1, the angle of the figure has become from left to right.

Q1(c)

Use the samples of X in part(b) with $\rho = 0.5$ to find the marginal PDF of X_1 and plot as a histogram. Estimate the expected value and the variance of the marginal PDF.




```
the estimated expected value of marginal PDF is 1.097
the estimated variance of marginal PDF is 1.727
```

2.simulating the central limit theorem

Question:

Let $X-i$ be a uniform random variable distributed in the interval $[0,1]$. Define

$$S_n = \sum_{i=1}^{n} X_i$$

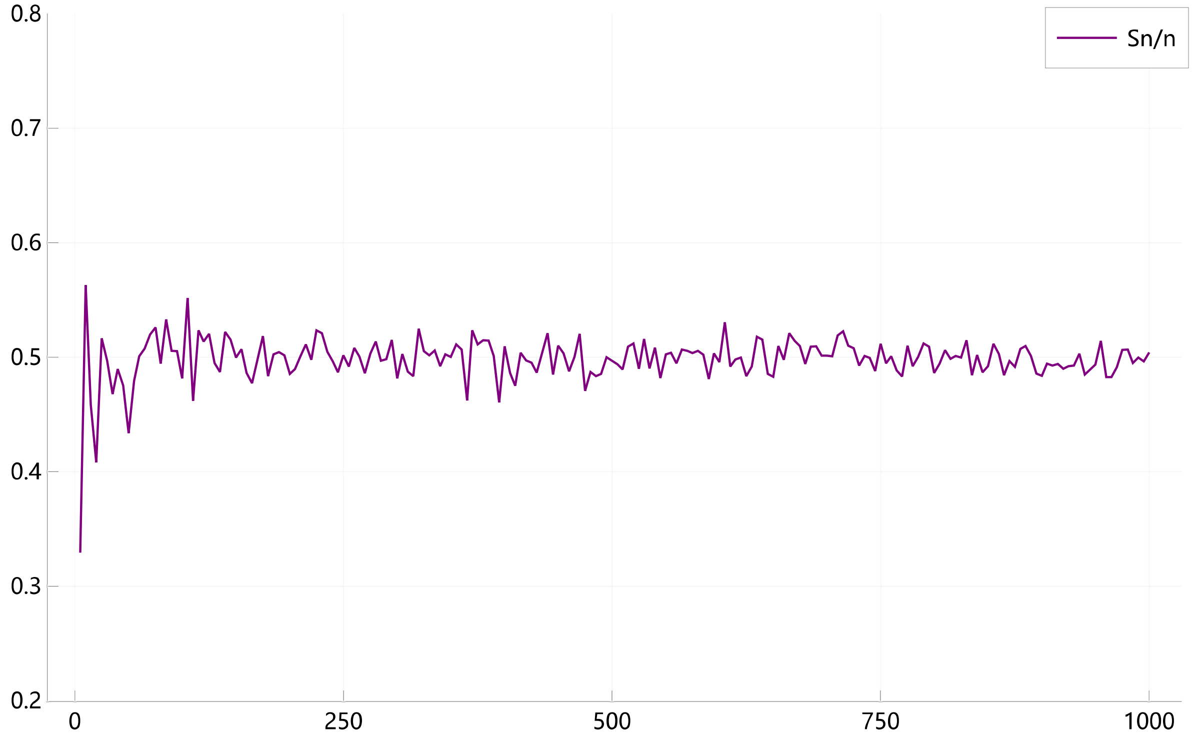
where X_i's are independent.

Q2(a)

For $n=1,\cdots,1000$ plot, $\frac{1}{n}S_n$ as a function of n. Does $\frac{1}{n}S_n$ converge? What is the limit?

```
sample_S (generic function with 1 method)
```

```
function sample_S(n)
    rv = 0
    for i = 1:n
        rv = rv + rand()
    end
    return rv
end
```



From the plot, we can find the $\frac{1}{n}S_n$ converge to 0.5, when $n \to \infty$.

Q2(b)

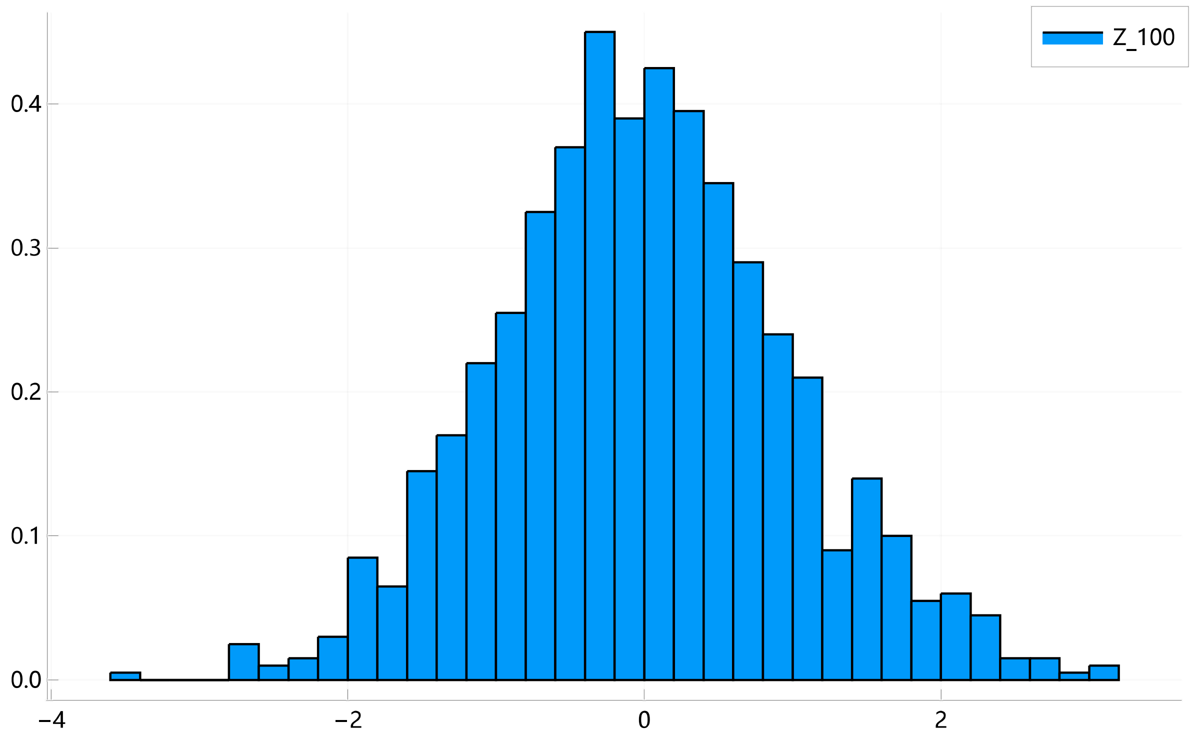
Let $n = 100$, Define

$$Z_{100} = \frac{S_{100} - 50}{\sqrt{100/12}}$$

Note that Z_{100} is a random variable. Generate 1000 samples of Z_{100} and plot the histogram of Z_n.

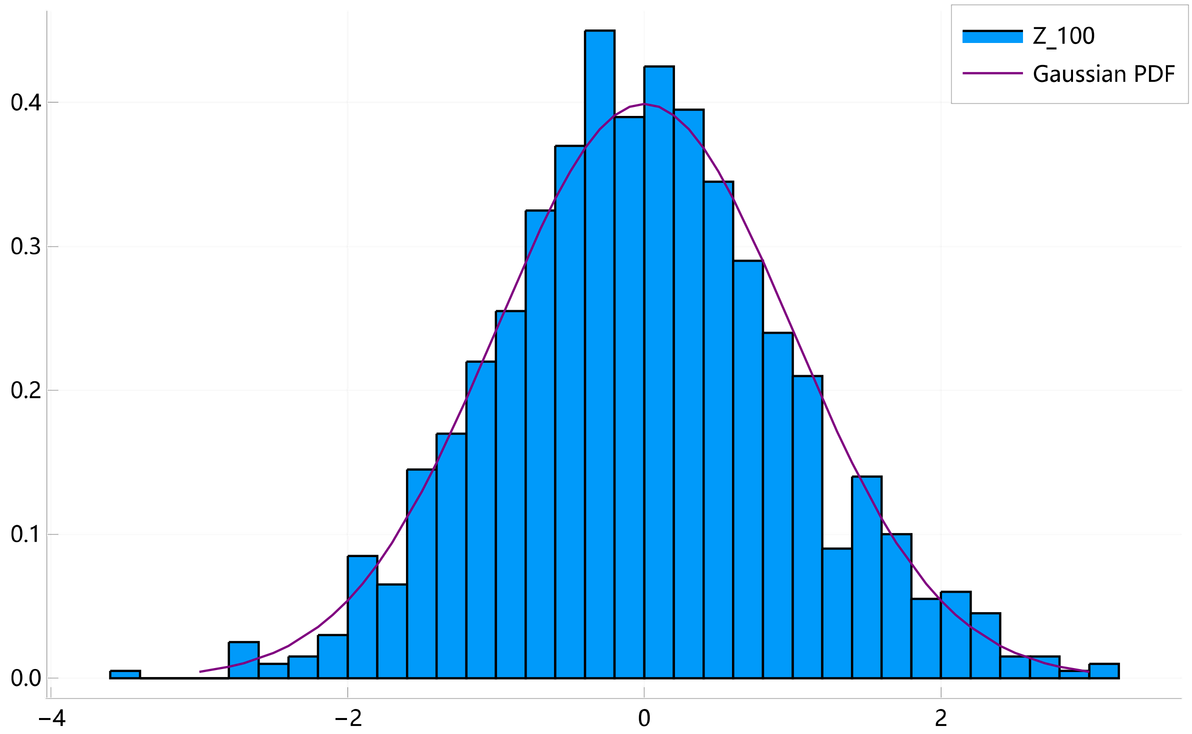
sample_Z100 (generic function with 1 method)

```
function sample_Z100()
    return (sample_S(100) - 50) / (100 / 12)^0.5
end
```



Q2(c)

Overlap a Gaussian pdf with the histogram in part (b).



Q2(d)

Use your simulated data to estimate the expected value and the variance of Z_n and compare with the theoretical values.

Since all the X_i are independent, so we can have the deriavtions as follow.

$$E(Z_{100}) = E(\frac{S_{100} - 50}{\sqrt{100/12}}) = \frac{E(S_{100}) - 50}{\sqrt{100/12}}$$

where

$$E(S_{100}) = \sum_{i=1}^{100} E(X_i) = 100E(X) = 50$$

finally, we can have

$$E(Z_{100}) = 0$$

the same way for checkout the variance, we can have

$$var(Z_{100}) = \frac{var(S_{100})}{100/12}$$

where

$$var(S_{100}) = 100var(X) = \frac{100}{12}$$

finally, we can have

$$var(Z_{100}) = 1$$

```
begin
    Est_mean_Z = mean(Z_N2_b)
    println("the estimated expected value of Z is $(round(Est_mean_Z, digits=3))")
    Est_var_Z = var(Z_N2_b)
    println("the estimated variance of Z is $(round(Est_var_Z, digits=3))")
end
```

```
the estimated expected value of Z is -0.006
the estimated variance of Z is 0.98
```