Table of Contents

ECE537: Lab 2 Report
1.Simulating Bivariate Gaussian Distributions
1.1 Numerical Simulation
1.2 Summary of Results
2. Empirical Verification of the Law of Large Numbers and the Central Limit Theorem
2.1 Numerical Simulation
2.2 Summary of Results
3. Code

ECE537: Lab 2 Report

*It is recommended to access this report by opening the `html` file on the browser or by clicking **here**.*

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 correlated joint random variables. In the second part, we will be empirically verifying the central limit theorem and the law of large numbers using uniform (univariate) random variables by testing for the relevant convergence cirtieria.

Throughout this lab, the **Distributions.jl** package in Julia has been utilized to be able to use the probabllity constructs in code.

```
using Distributions , StatsBase , StatsPlots , LinearAlgebra ,
LaTeXStrings , PlutoUI , Measures
```

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 (symmetric, 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:

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

1.1 Numerical Simulation

We will now define and sample various bivariate normal distributions and inspect their coverage and densities through scatter plots. Below is a slider for the number of samples we wish to take per distribution in this section.

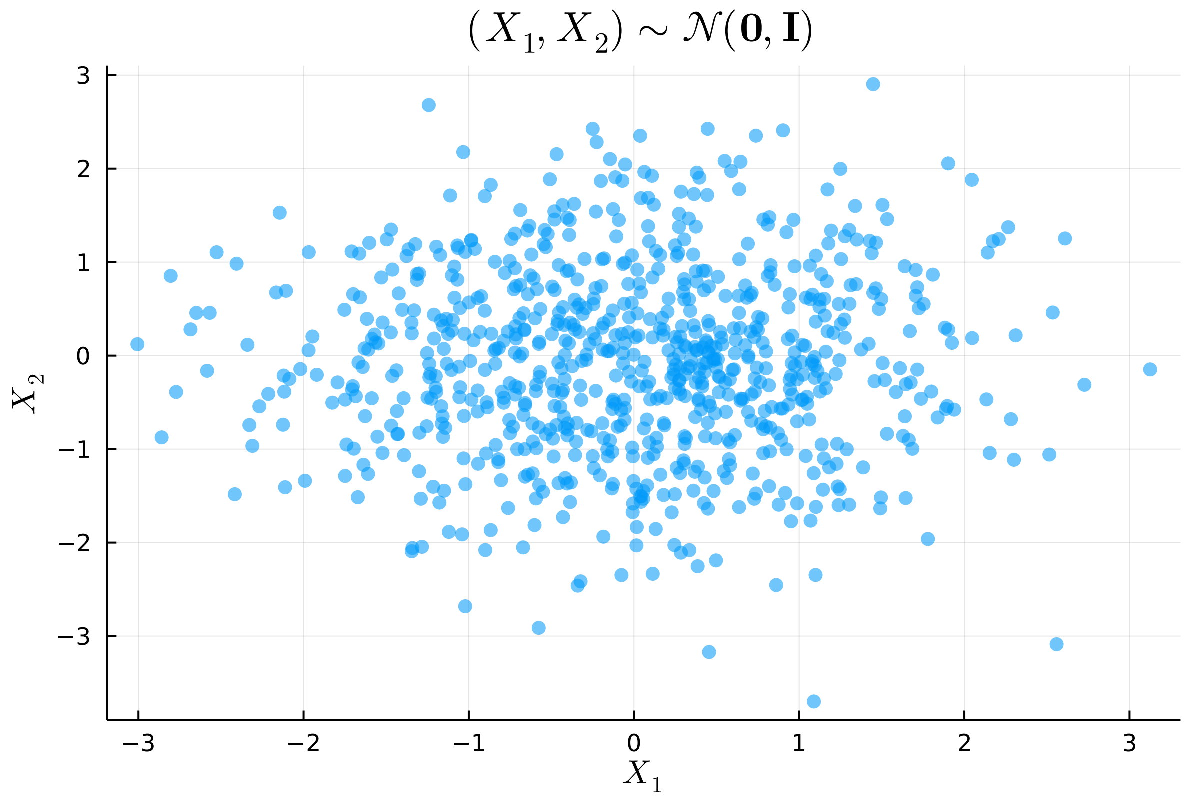
N_1 = 800

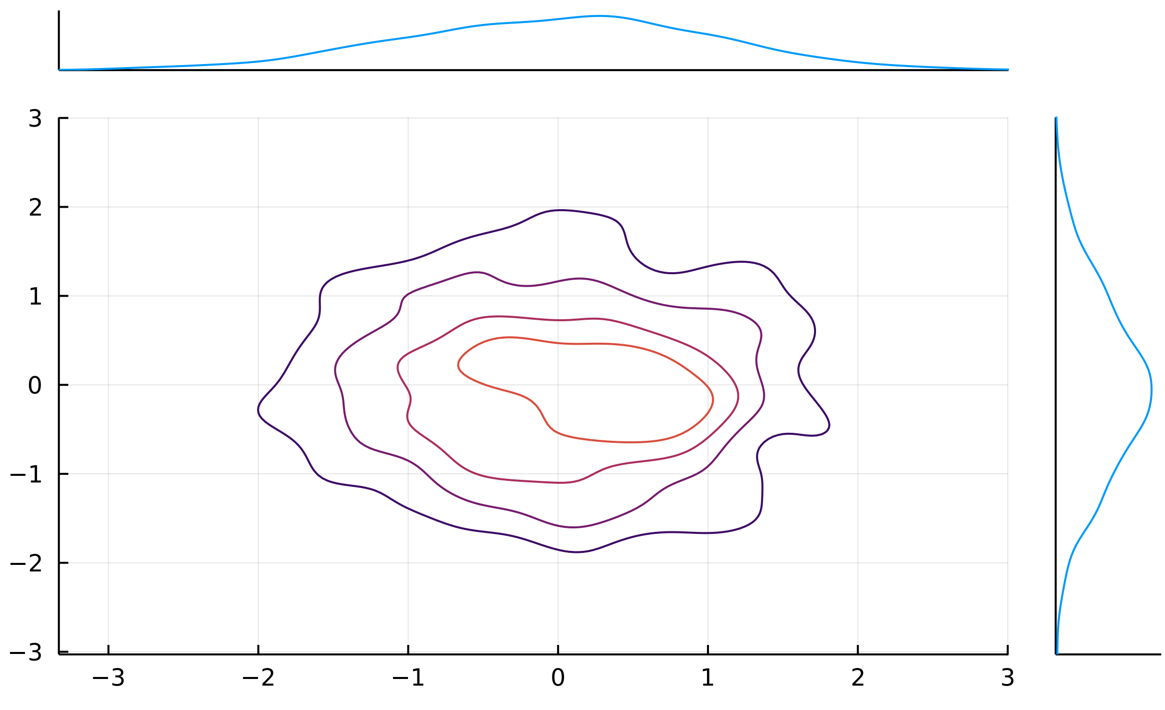
We will start by defining an uncorrelated bivariate normal distribution with zero mean and unit variance, $X \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. After sampling the distribution N_1 times, we also show the scatter and contour plots for the variate.

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

```
X = MvNormal(zeros(2), I(2));
```

```
Xsamples = rand(X, N₁) |> matrixtotuple;
```





Contours and Marginal Densities of N(0,I)

On first look, and especially after visually adjusting the aspect ratios, the X_1 and X_2 marginal densities look nearly identical in trend and suggest that they are the same as the partitioned distributions of the bivariate. We will verify and formally present the this idea in the next section.

Now, we will define a collection of bivariate normal distributions with varying degree of correlation to see its effects qualitatively.

Since, the covariance matrix depends on the correlation coefficient, ρ and the standrd devaiations σ_1, σ_2, we will define a generic function, $\mathbf{\Sigma}(\sigma_1, \sigma_2, \rho)$, for generating valid covariance matrices based on these parameters.

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

```
function Σ(σ₁, σ₂, ρ)
    @assert abs(ρ) ≤ 1

    # add epsilon along diagonal for numerical stability during cholesky decomposition
    ϵ = 1e-6
    return ϵ*I(2) + [σ₁^2     ρ*σ₁*σ₂;
                     ρ*σ₁*σ₂  σ₂^2]
end
```

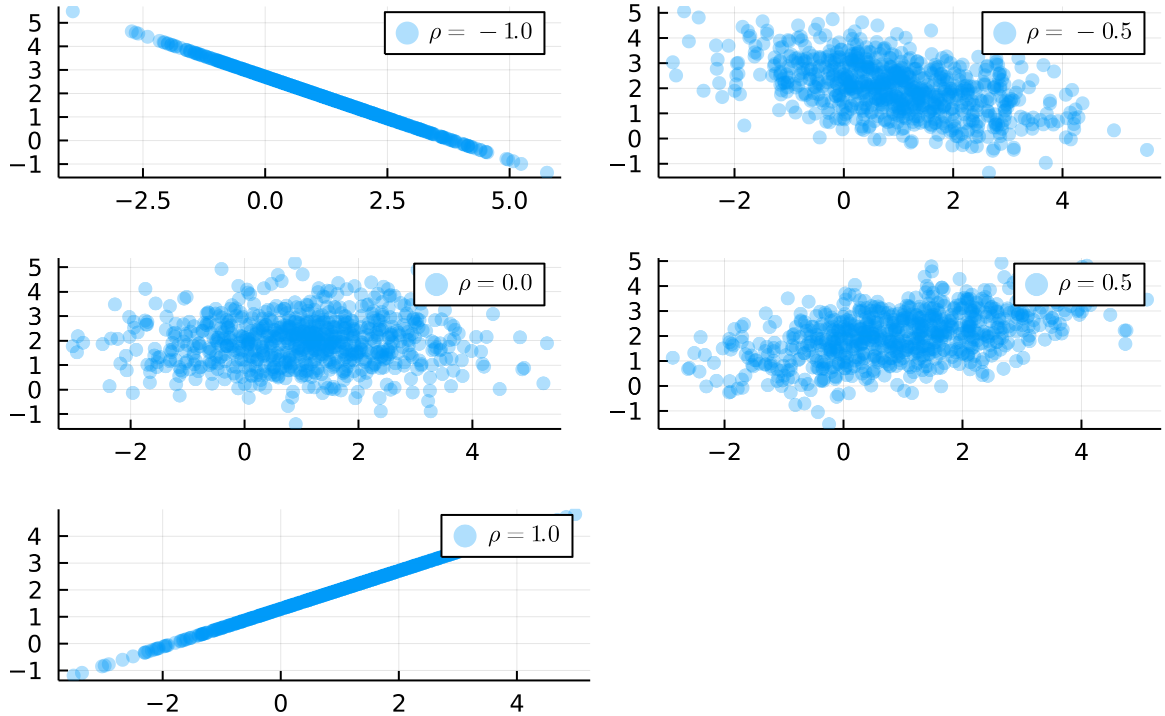
To see the effects of correlation in bivariate distributions, we will fix the mean across all distributions and only vary ρ. Therefore, we define 5 distributions with mean $\boldsymbol{\mu} = [1 \quad 2]^\intercal$, variance $\sigma_1^2 = 2$, $\sigma_2^2 = 1$, and correlation coefficient ranging from $\rho = -1.0$ to $\rho = 1.0$ in increments of 0.5. These distributions are stored in the $\mathbf{X}_\rho$ vector.

```
μ = [1; 2];
```

```
𝑋ᵨ = [MvNormal(μ, Σ(√2, 1, ρ)) for ρ ∈ -1:0.5:1];
```

```
𝑋ᵨsamples =
  [[(1.88967, 1.3734), (0.0444359, 2.67393), (2.77892, 0.739851), (1.16727, 1.88472), (0.
```

```
𝑋ᵨsamples = [rand(𝑋, N₁) |> matrixtotuple for 𝑋 ∈ 𝑋ᵨ]
```



Correlated Bivariate Normals, X_ρ

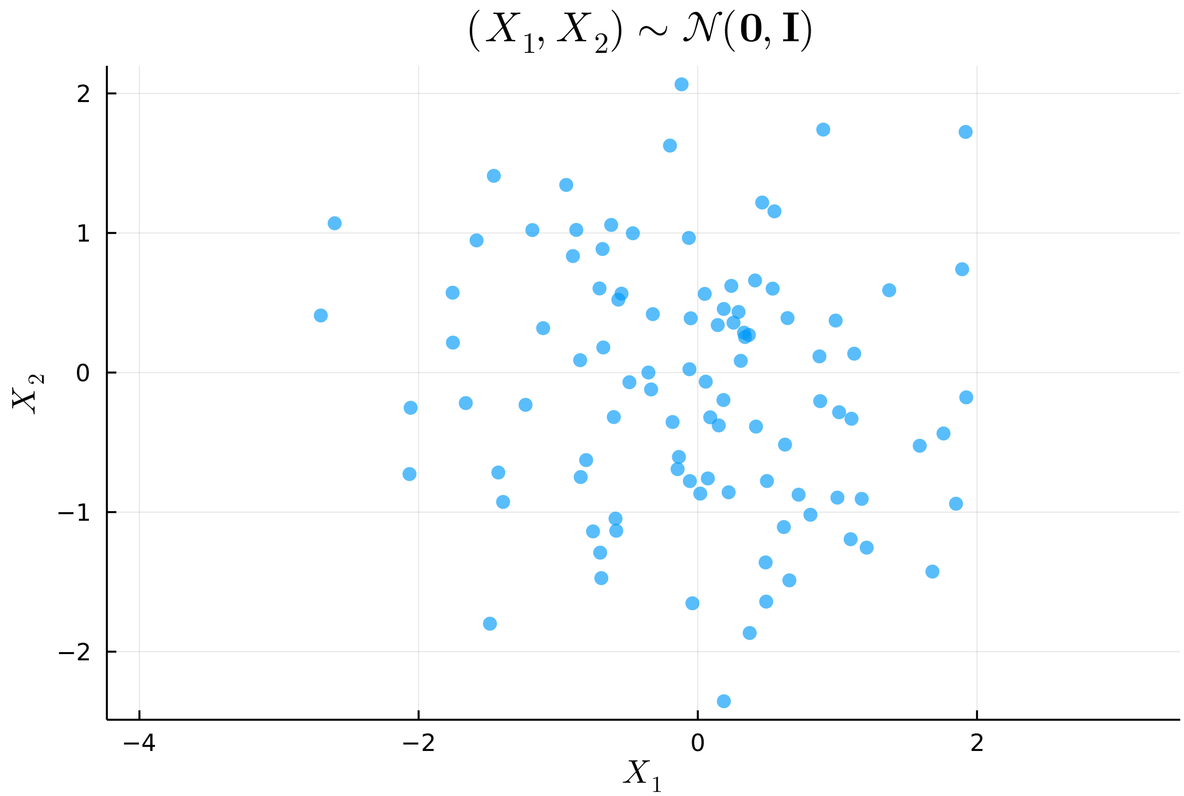
1.2 Summary of Results

Here we will test for a fixed number of samples, $N = 100$, and observe the characteristics of the bivariate distributions and if they match our expectations.

```
N = 100; # fixed number of samples
```

The scatter plot for $X \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ is shown below.

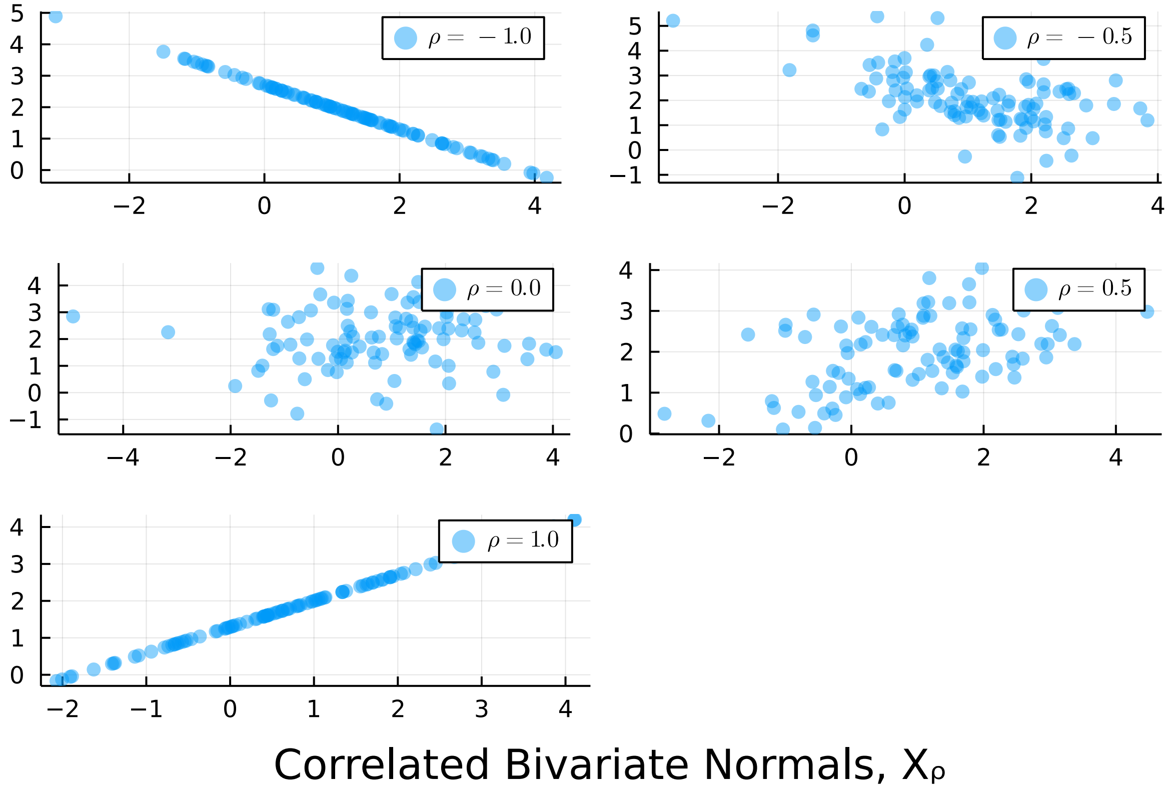
```
𝑋fixedsamples = rand(𝑋, N) |> matrixtotuple;
```



As can be seen, both X_1 and X_2 show equal spread around zero and will thus also have nearly identical marginal densities f_{X_1} and f_{X_2}.

Similarly, we will sample the correlated collection of bivariate normals and display their scatter plots below.

```
Xᵨfixedsamples = [rand(X, N) |> matrixtotuple for X ∈ Xᵨ];
```



Thus, we can observe in the plots for $\mathbf{X}_\rho$ above, and in section 1.1 as well, that the correlation coefficient captures the degree to which the joint variates have an affine relationship. For $\rho = -1$, it is exactly the case that $X_2 \propto -X_1$ about the mean, and the opposite, i.e. $X_2 \propto X_1$, for $\rho = 1$. For $\rho = 0$, there is no single affine relationship explaining the spread of the bivariate distribution and the random variates are thus uncorrelated.

Note, that for any partition of a multivariate normal distribution $\mathbf{Z} \sim \mathcal{N}(\boldsymbol{\mu}_\mathbf{Z}, \boldsymbol{\Sigma}_\mathbf{Z})$ will have the case that

$$\mathbf{Z} = (\mathbf{X}, \mathbf{Y}) \sim \mathcal{N}\left(\begin{bmatrix}\boldsymbol{\mu}_\mathbf{X} \\ \boldsymbol{\mu}_\mathbf{Y}\end{bmatrix}, \begin{bmatrix}\boldsymbol{\Sigma}_\mathbf{XX} & \boldsymbol{\Sigma}_\mathbf{XY} \\ \boldsymbol{\Sigma}_\mathbf{YX} & \boldsymbol{\Sigma}_\mathbf{YY}\end{bmatrix}\right)$$

where,

$$\boldsymbol{\mu}_\mathbf{X} = E[\mathbf{X}], \boldsymbol{\Sigma}_\mathbf{XX} = E[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^\intercal], \text{ and } \boldsymbol{\Sigma}_\mathbf{XY} = E[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})(\mathbf{Y} - \boldsymbol{\mu}_\mathbf{Y})^\intercal]$$

Similar definitions follow for $\mathbf{Y}$, where $\boldsymbol{\Sigma}_\mathbf{YX} = \boldsymbol{\Sigma}_\mathbf{XY}^\intercal$.

The marginal distributions, as we visually guessed before, are

$$\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}_\mathbf{X}, \boldsymbol{\Sigma}_\mathbf{XX}), \text{ and } \mathbf{Y} \sim \mathcal{N}(\boldsymbol{\mu}_\mathbf{Y}, \boldsymbol{\Sigma}_\mathbf{YY}).$$

In higher dimensions, these will also be joint normal distributions. To verify this general result for the simpler bivariate case, we will choose $\mathbf{X}_\rho$ with $\rho = 0.5$ above and extract the marginal density for X_1 which should approximately be the univariate $\mathcal{N}(\mu_1 = 1, \sigma_1^2 = 2)$.

```
X₁samples = [X[1] for X ∈ Xρfixedsamples[4]];
```

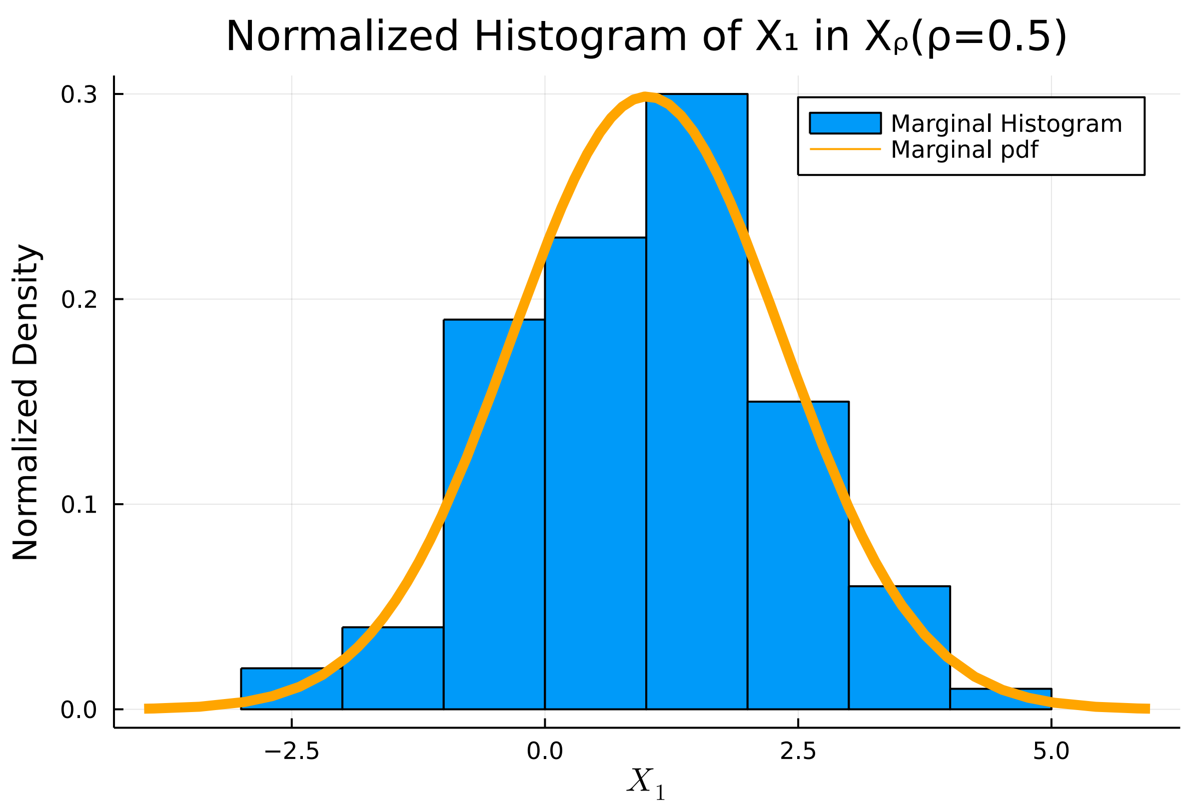
```
1.0080591755323156
```

```
mean(X₁samples) # ≈ 1
```

```
1.7824667087891368
```

```
var(X₁samples; corrected=false) # ≈ 2
```

```
X₁hist = fit(Histogram, X₁samples, nbins=10);
```



Note that the mean and variance are sufficient statistics to construct the pdf of a collection of samples with an underlying normal distribution, and therefore, we have empiricially verified this result for the bivariate case.

2. Empirical Verification of the Law of Large Numbers and the Central Limit Theorem

In this section, we will begin by verifing the strong law of large numbers, which says that for the sample mean, or the unbiased estimator for $E[X]$, $M_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ of an independent and identically distributed (iid) sequence of random variables with the pdf of X, converges almost surely to the theoretical mean, i.e. $M_n \xrightarrow{a.s} E[X], n \to \infty$.This means that the probability associated with the set of realizations of the random process M_n that converge to $E[X]$ is equal to 1.

Then we will empirically verify the central limit theorem for the same sequence $X_1, X_2, \ldots$ of iid random variables above with mean μ and variance σ^2. The theorem states that a process $Z_n = \frac{1}{\sigma\sqrt{n}}\sum_{i=1}^{n}(X_i - \mu)$, will converge in distribution to the univariate normal with zero mean and unit variance, i.e. $Z_n \xrightarrow{a.s} \mathcal{N}(0,1), n \to \infty$.

We will use a sequaence of iid uniform random variables $X \sim \mathcal{U}(0,1)$ for undertaking both verifications.

2.1 Numerical Simulation

Below is a slider for the number of samples we wish to take per random process in this section.

N_2 = 20

We begin by defining N_2 number of uniform random variables, U_i, over the support $[0,1]$, which can be viewed as a sequence of iid random variables to construct a random process.

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

```julia
U(n) = [Uniform(0, 1) for n ∈ 1:n]
```

```julia
Uₙ = U(N₂);
```

We then define the $n-$degree sum, S_n, of the sequence U_n,

$$S_n = \sum_{i=1}^{n} U_i.$$

Then, we would like to check if the value of the unbiased mean estimator $M_n = \frac{S_n}{n}$ converges to $E[X]$ as n increases.

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

```
S(Uₙ) = rand.(Uₙ) |> sum
```

```
Mₙ = 0.37687672655687604
```

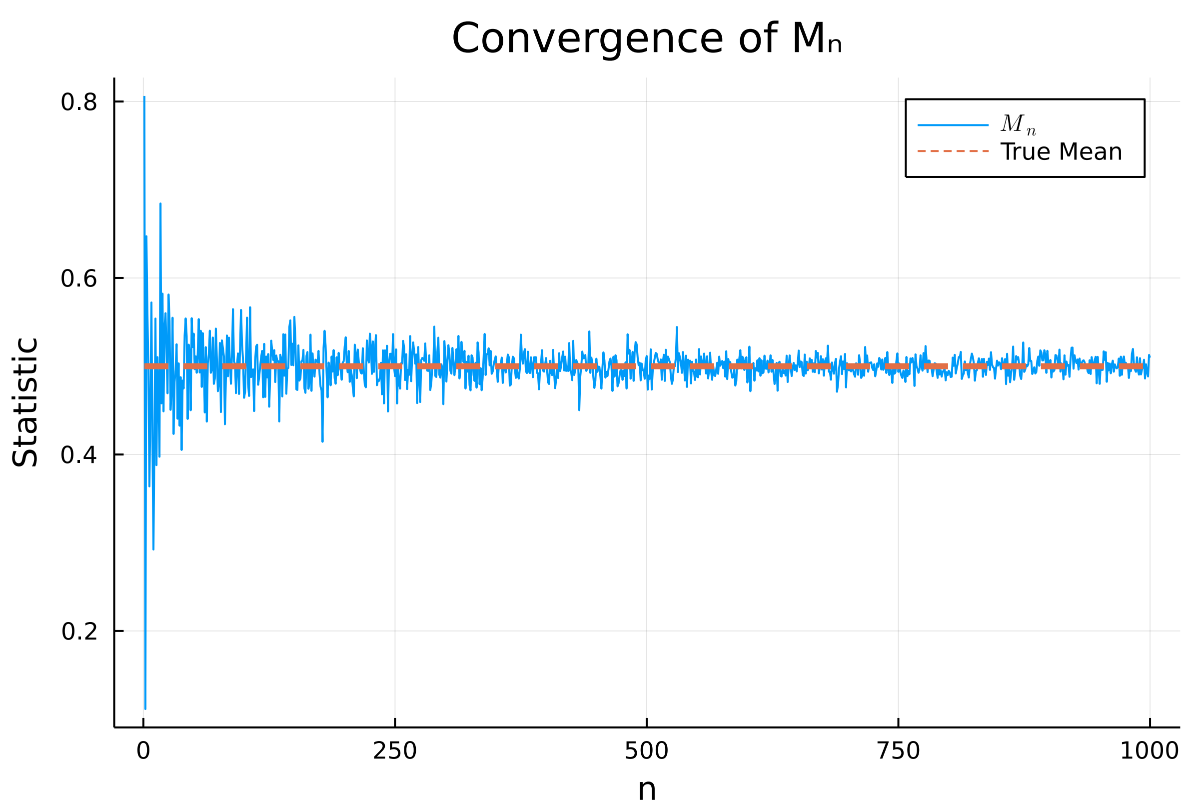
```
Mₙ = S(Uₙ)/length(Uₙ)
```

2.2 Summary of Results

Here we will empirically test for the convergence of M_n and Z_n. We begin by defining M_n by composing S_n and U_n and dividing by the number of iid variables we consider in the sequence, n. Then, we test for convergence by plotting M_n for $1 \leq n \leq 1000$ and observe that it is indeed showing convergence to the true mean of the underlying uniform distribution.

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

```
M(n) = (S ∘ U)(n)/n # composing the S and U definitions over n
```



We can now use the definition of the process M_n to further define Z_n as

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{\sqrt{n}}{\sigma}(M_n - \mu).$$

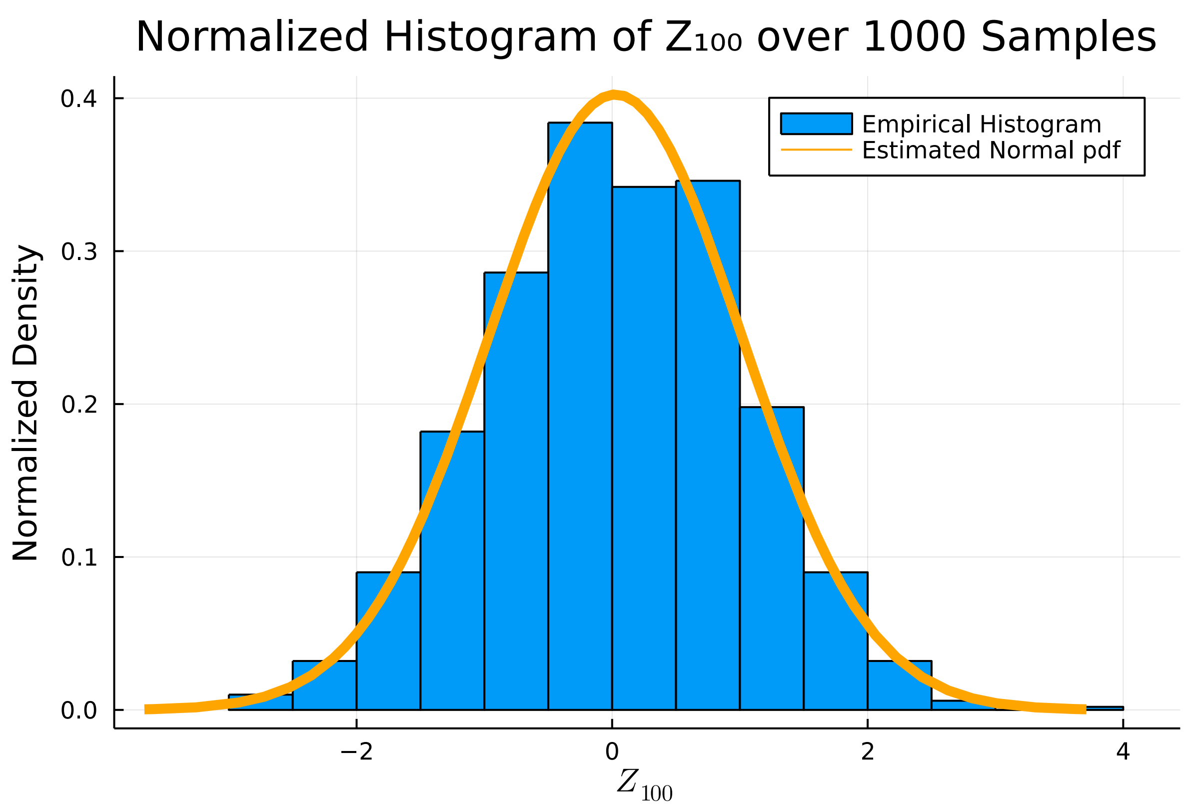
```
Z (generic function with 1 method)
```

```
Z(n, μ, σ) = (M(n) - μ) * (√n)/σ
```

Now, we will test for the convergence of the process Z_n by sampling the Z_{100} random variable $N = 1000$ times. We will plot the normalized histogram and also the best-fit normal distribution and check if it is close to the theoretical convergence limit $\mathcal{N}(0, 1)$.

```
Z₁₀₀samples = [Z(100, 0.5, 1/√12) for n ∈ 1:1000];
```

```
Z₁₀₀hist = fit(Histogram, Z₁₀₀samples, nbins=15);
```




```
0.02547664637533764
```

```
mean(Z₁₀₀samples) # ≈ 0
```

```
0.9825487944827361
```

```
var(Z₁₀₀samples; corrected=false) # ≈ 1
```

Note that the normal distribution constructed using the mean and uncorrected variance of the Z_{100} samples corresponds to the normal maximum likelihood estimate (MLE) of the data, which is close to the theoretical limit of converging to the underlying zero mean and unit variance normal variate structure of Z_n for $n \to \infty$.

3. Code

Note that this lab report can be run on the cloud and viewed as is on the github repository page **here**. All code for the notebook can be accessed **here**.