1. ***L_2* Norm**
   a. If we assume toward a contradiction that there exists a vector *x* whose L_2 norm is less than zero, we then assume that there exists a value, or set of values, that can squared, and the squares can be summed up and the square root can be taken, to yield a negative number. There exist no values whose square is negative, and there exist no values whose square root is negative, thus leading to a contradiction, so it must be the case that the L_2 norm must be positive for all vectors *x*.
   b. If we assume, toward a contradiction, that there exists a nonzero vector *x* whose L_2 norm is equal to zero, we assume that there exists a value or a set of values whose sum of squares is equal to zero. This is a contradiction, as there is no value other than zero whose square is zero. As proven in part (a), any other real value's square is a positive number, meaning that there exists no possible value that could negate a positive square to bring the sum of squares to zero. Thus it must be the case that the only vector whose L_2 norm equals zero is the zero vector.
   c. Let us assume, toward a contradiction, that the scalar multiple property of the L_2 norm does not hold:

$$|ax|_2 \neq |a||x|_2$$

$$\sqrt{\sum_{i=1}^{n} (ax_i)^2} \neq |a|\sqrt{\sum_{i=1}^{n} x_i^2}$$

$$\sqrt{\sum_{i=1}^{n} a^2x_i^2} \neq |a|\sqrt{\sum_{i=1}^{n} x_i^2}$$

$$\sqrt{a^2\sum_{i=1}^{n} x_i^2} \neq |a|\sqrt{\sum_{i=1}^{n} x_i^2}$$

$$|a|\sqrt{\sum_{i=1}^{n} x_i^2} \neq |a|\sqrt{\sum_{i=1}^{n} x_i^2}$$

   Through the mathematical properties of the summation and square root operators, we have been able to manipulate the expression to be an identity. Thus assuming the scalar multiple property of the magnitude of a vector to not be true has led us to a contradiction, so it must be that case that such a property does hold.
   d. We can prove that the norm of a vector sum is less than the sum of its constituent vectors' norms using the Triangle Inequality and expanding the norm of the sum:

$$\begin{aligned}|x + y|^2 &= x \cdot x + x \cdot y + x \cdot y + y \cdot y \\ &= |x|^2 + 2|x||y| + |y|^2 \\ &= (|x| + |y|)^2 \\ &\Rightarrow |x + y| \leq |x| + |y|\end{aligned}$$

2. **Random Events**

A good example of a random event is price movement and direction of price movement for any given company's will be the next morning on the New York Stock Exchange. This can be considered a random event, because although a stock's price is influenced by many real factors, such as the company's balance sheet/earnings report, overall market sentiment, inflation of the US dollar, related legislation/federal regulations, and the state of the US macroeconomy, there are so many factors that it is extremely difficult if not impossible to try to model a security's valuation as a system or algorithm. In such cases it can be considered appropriate to model an event as a random event.

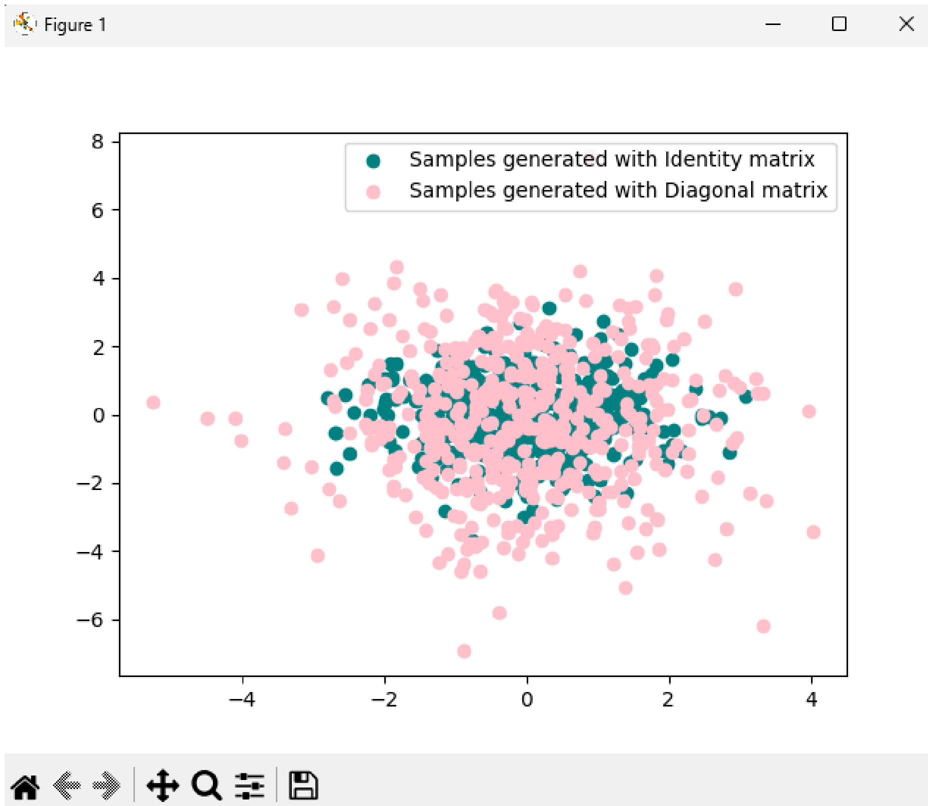
Since the variable is the change in the stock's price, the sample space is all real numbers (i.e., an unbounded continuous sample space). I wanted to say that the sample space is a continuous interval of real numbers lower bounded by the negative value of the stock's price; this would make sense because it would be the delta value that signals the rare event of a stock becoming worthless overnight. While intuitive, we saw in 2020 that oil futures were able to go negative as holders of these securities became so desperate to get rid of these assets that they were willing to pay people to take them off their hands.

A good distribution may vary on a case-by-case basis, as every security is different. In general, though, I think a positively skewed Gaussian distribution would be a good representative distribution. Gaussian because the central limit theorem posits that more observations will cause a distribution to take the form of a normal distribution, positive skew because US asset markets tend to naturally go up in prices in an ordinary economy.

3. **Sampling from an arbitrary Gaussian distribution**

Python script is attached to assignment submission.

The samples generated on my first runtime are shown below:



4. **Conditional probability and independence**

The conditional probability of P(X|Y=0) is the probability of the assignment of X to a particular value given the assignment event of Y = 0 has already happened. Conditional probability is defined as the joint probability divided by the conditioned event:

P(X|Y=0) = P(X,Y=0) / P(Y=0)

For instance, we can compute the conditional probability P(X=0 | Y=0) first by finding the marginal probability of P(Y=0) by summing up the joint probabilities corresponding to Y=0, which we find to be 0.3. We then divide the joint probability of X=0 and Y=0, which is 0.1, by 0.3, which comes out to P(X=0 | Y=0) = 0.33.

For X, Y to be independent, the following condition must be satisfied:

P(X,Y) = P(X)P(Y)

Looking at the joint probability P(X=0, Y=0), if X and Y were to be independent, we would expect the joint probability to be P(X=0)*P(Y=0) = 0.35 * 0.3 = 0.105. This does match the given table of joint probabilities, meaning that X and Y are not independent.