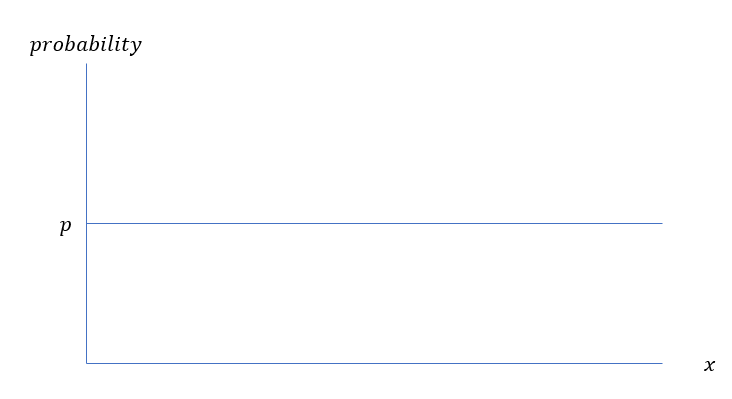
**Statistics Final Project**

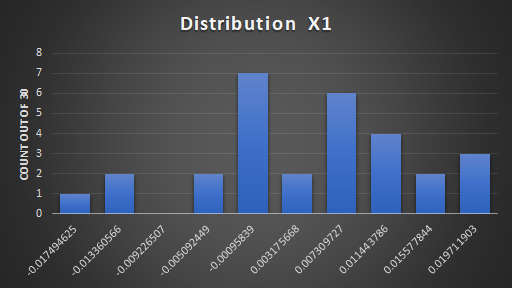
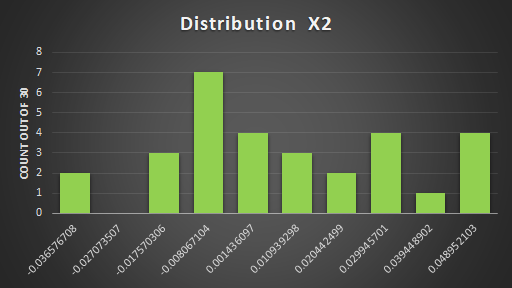
1. Data points chosen by uniformly sampling across the historical opening prices over the past 2 years. That is the probability of any data point chosen from the historical data points defined as . Then we have:



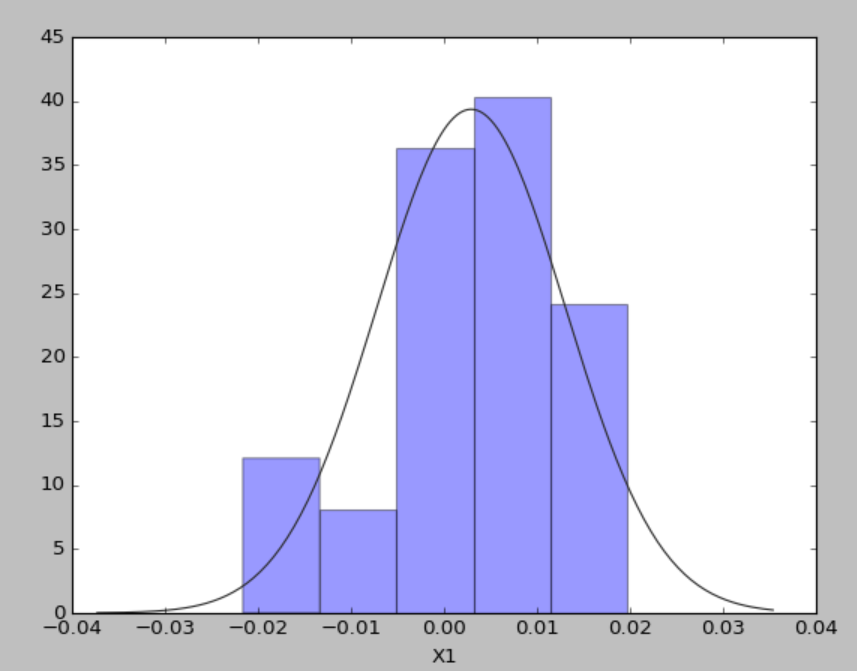
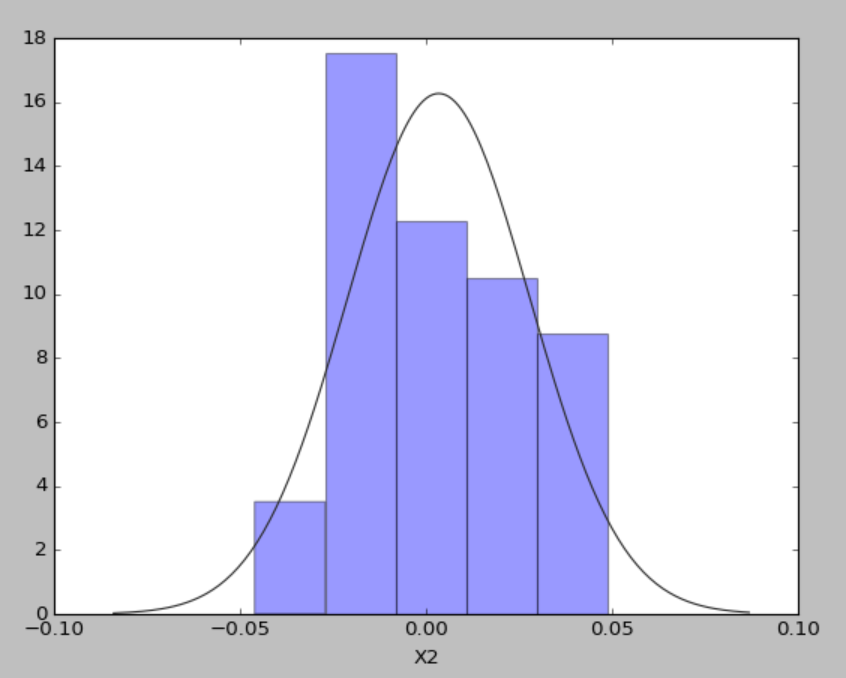
1. Let and be the mean of samples and respectively. is the daily percentage change in opening prices of stock where:

Our claim is that there is no difference between sample means and

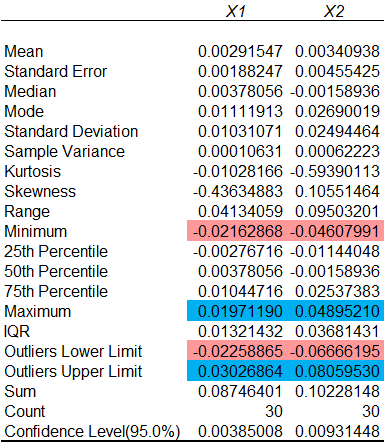
1. We plot the histogram of each (default to 10 bins for each histogram)

We can also compare the histogram against a t-distribution PDF as following:

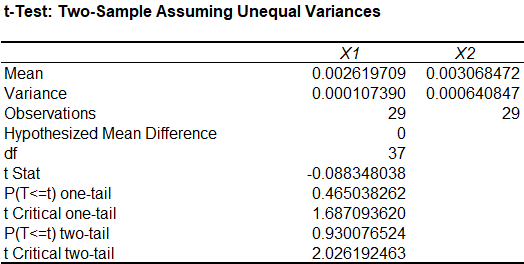
We can see that both distribution do not follow typical normal or t-distributions and therefore median/IQR present a more descriptive measure of them. Nevertheless, we present all of the standard descriptive statistics as follows:



We define outliers as data points that fall below or above . From the descriptive statistics above, we see that there are no data points from the sample that are considered outliers.

1. We next run hypothesis test for 2 population means. Since the samples are obtained using simple random sampling, the sample sizes are large and sample size is of population size → we conclude that the samples can be considered as independent

Since variances of the sample are significantly different, we use Excel’s **2-Sample Assuming Unequal Variance T-Test data analysis tool** and the results are as follows:



Since the 2-tail p-value = 0.93 > 0.025, we cannot reject the null hypothesis and conclude that the sample means are not significantly different.

* One possible reason for the null hypothesis could be passive investing of ETFs that buy most names in the NASDAQ composite index. Before the opening bell, outside of idiosyncratic effects of individual stocks, the big NASDAQ (we chose TSLA and MSFT, both big stocks) stocks all gap up/down with the overall index by the same amount
* TSLA and MSFT can be seen as high beta names in the NASDAQ composite index as both stocks outperformed greatly over the past 2 years

