Modeling the Distribution of Stock Returns

Rachel Sebastian

Advisor: Dr. Matthew Wascher

University of Dayton

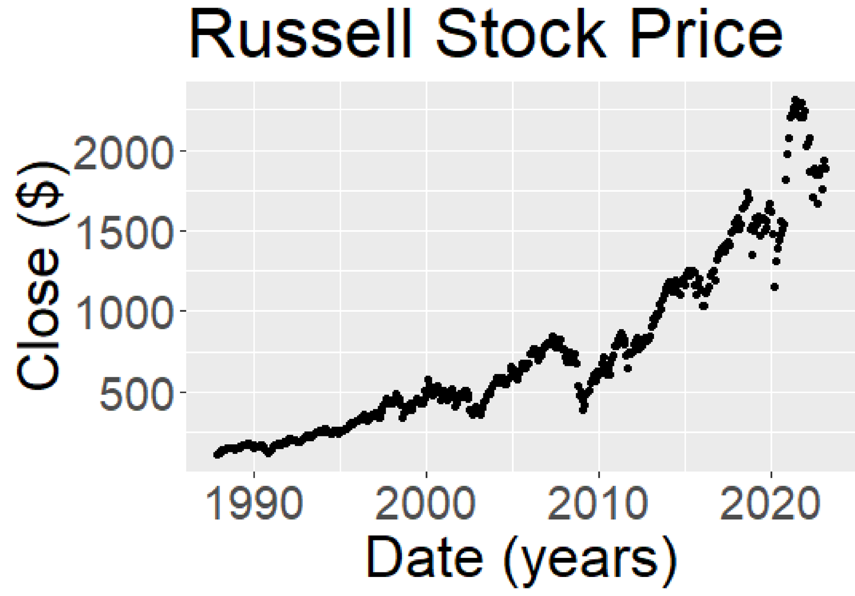
Introduction

- Research objective: determine what types of statistical models can characterize the distribution of stock returns
- Stock returns measure price changes and show the value of a company
- Distributions considered were normal, lognormal, Laplace, and Cauchy
- Several criteria were used to evaluate which distribution fits the data the best including distance metrics, quantile-quantile plots, and likelihood

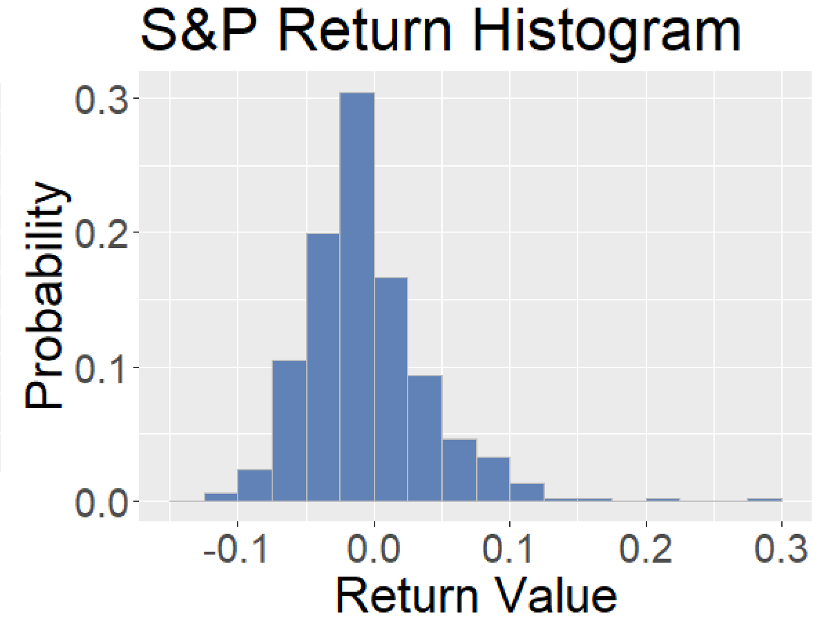
Data

- Data tested uses monthly prices from the S&P 500 and Russell 2000 indexes from 1980s-2023

Russell Stock Price



S&P Return Histogram



Price data is used to calculate monthly returns

- $Return = \frac{P_t - P_{t-1}}{P_{t-1}}$
- $Return\ Ratio = \frac{P_t}{P_{t-1}}$

Theory

The probability distribution functions for the distributions and reasoning is shown

Normal: previous literature

- $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$

Lognormal: positive skew

- $f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(\ln(x)-\mu)^2}$

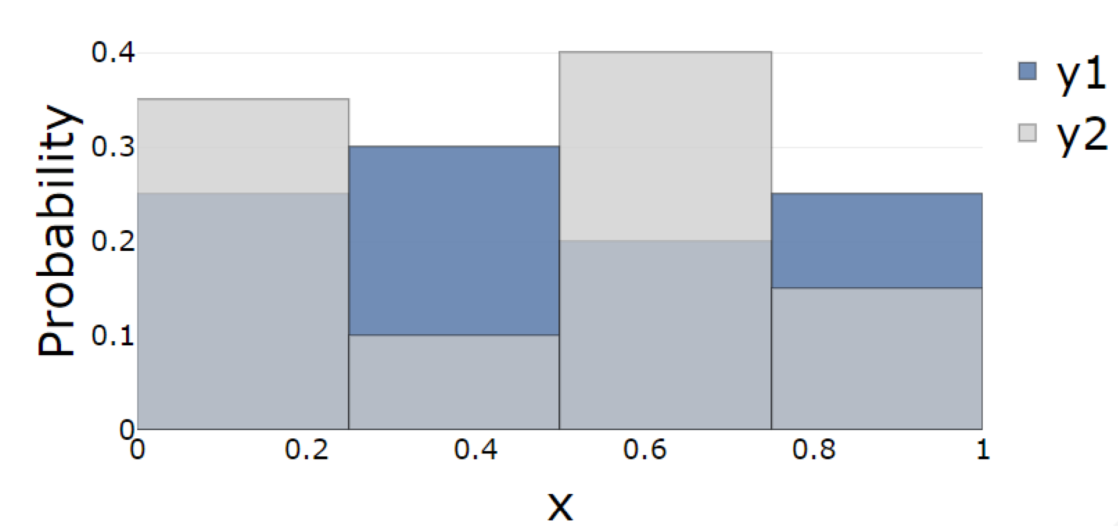
Laplace: higher peak and heavy tails

- $f(x) = \frac{1}{2b} e^{-\frac{|x-\mu|}{b}}$

Cauchy: higher peak and heavier tails

- $f(x) = \frac{1}{\pi\gamma\left[1+\left(\frac{x-x_0}{\gamma}\right)^2\right]}$

Distances



Earth Mover's Distance

- $EMD(P,Q) = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} f_{i,j} d_{i,j}}{\sum_{i=1}^{m}\sum_{j=1}^{n} f_{i,j}}$

Total Variation Distance

- $TV(P,Q) = \frac{1}{2}\sum_{s\in S}|P(s) - Q(s)|$

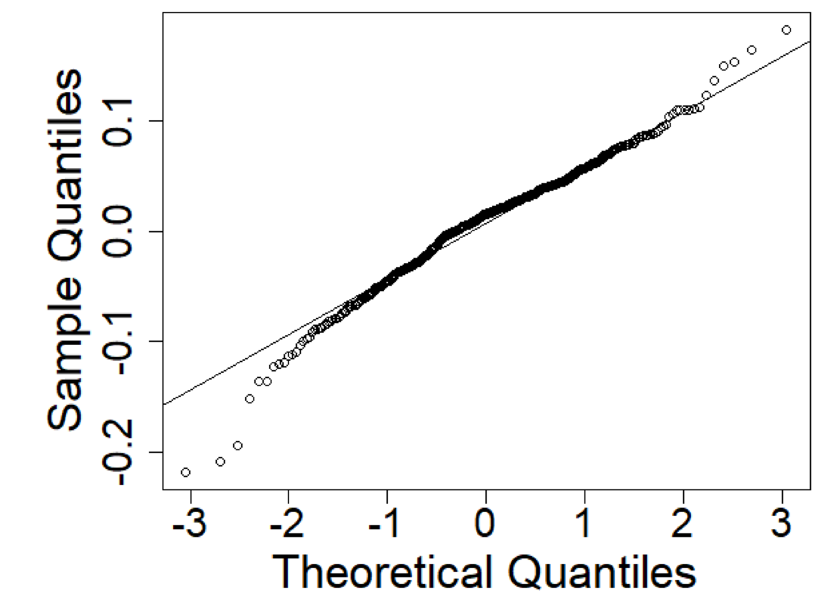
Kullback – Leibler Divergence

- $D_{KL}(P||Q) = \sum_{s\in S} P(s)\log(\frac{P(x)}{Q(x)})$

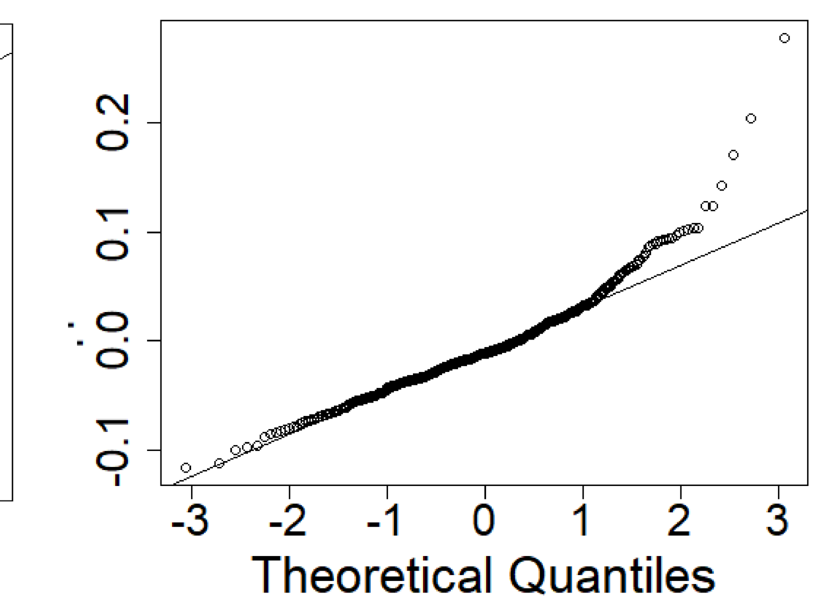
Maximum Likelihood Estimation estimates parameters using likelihood functions

Results

Russell Normal Q-Q Plot

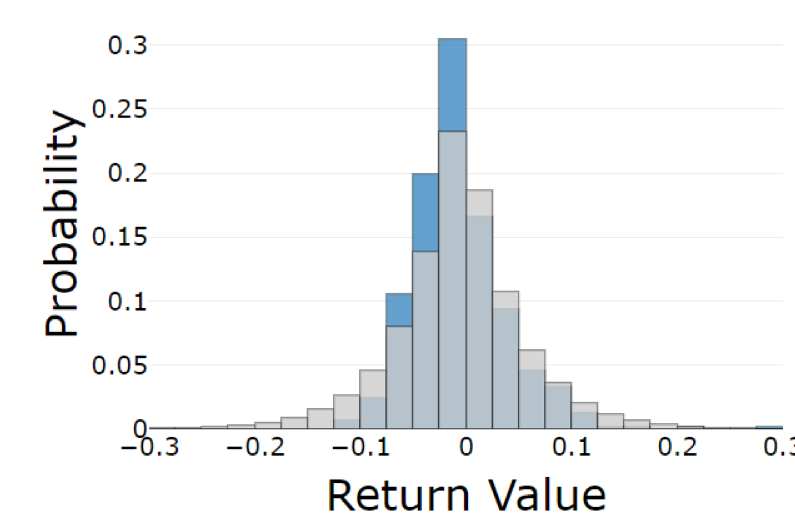


S&P Normal Q-Q Plot

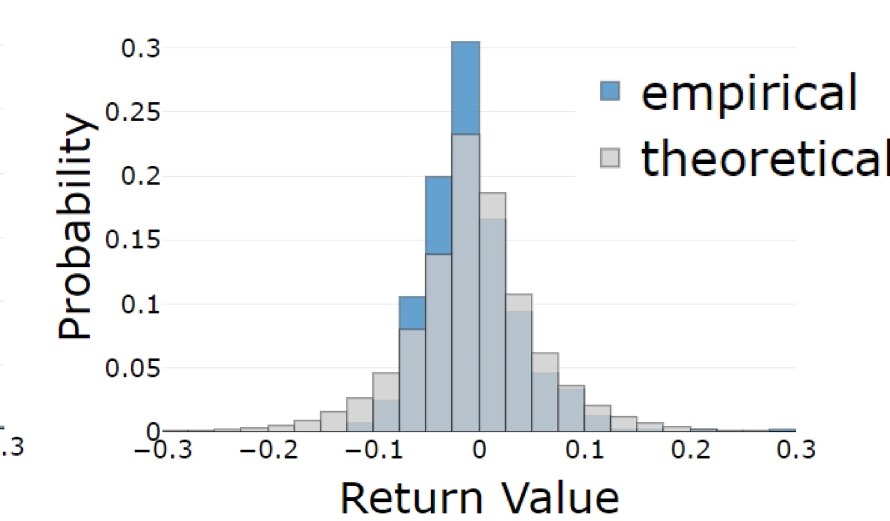


	EMD	TV	KL	MLE
Russell	Laplace	Laplace	Lognormal	Laplace
S&P	Normal	Normal	Lognormal	Laplace

Russell with Laplace



S&P 500 with Normal



Conclusion

- Using various measures of distance yields conflicting results
- Our models suggest none of the distributions characterize the return distributions well
- Next steps for this study could be to consider other distributions such as the Pareto and Power Law
- More stock return data will help conclude an overall best fit distribution for stock returns