ST501 – 001 Fall 2019 Final Project

Antonio Jones

Hannah Park

Tharun Polamarasetty

We have neither given nor received unauthorized aid on this assignment.

GOOGLE

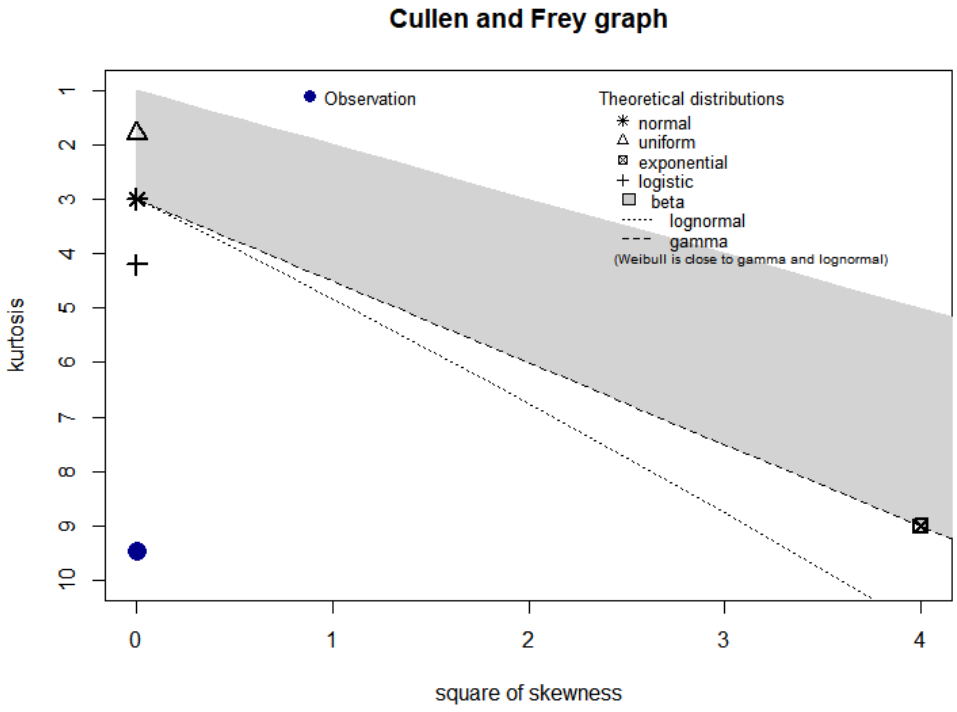
1(a)

$\mathbf{Q_{0.05}(R)}$ = -0.02593215

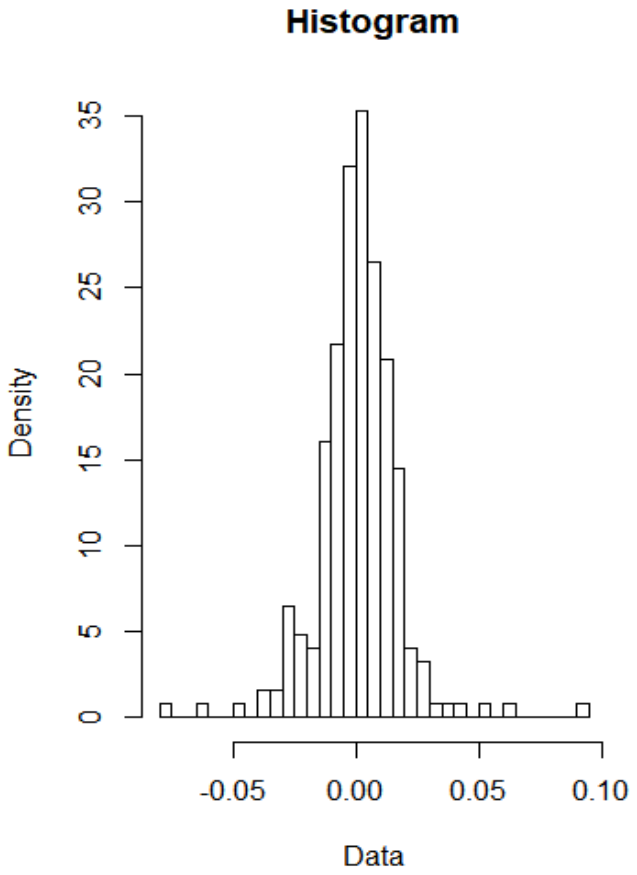
$\mathbf{E[R \mid R < Q_{0.05}]}$ = -0.03802518

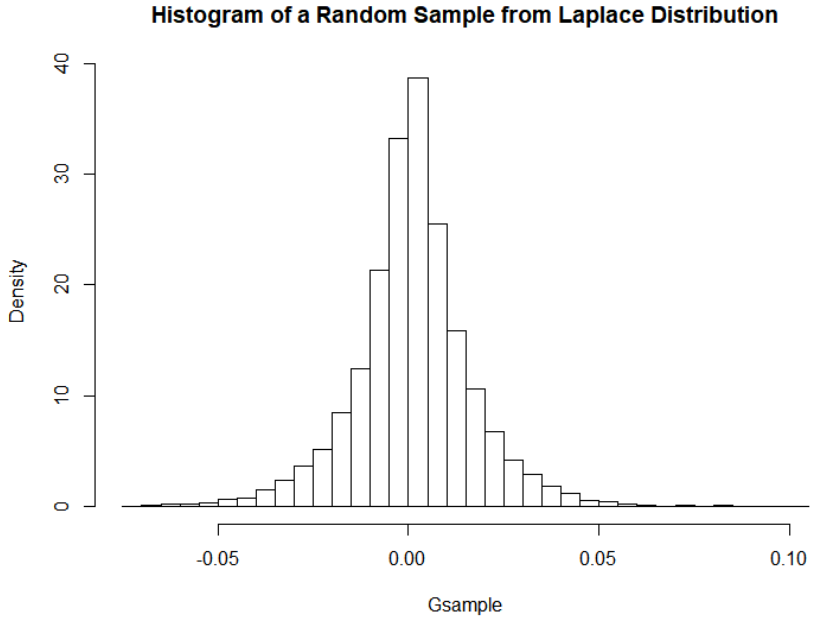
1(b)

According to the Cullen and Frey graph created with the descdist function in R, the distribution of our Google data does not match exactly with any of the common theoretical distributions plotted on the same graph. Of the families of distributions listed on the graph, Logistic comes the closest. However, fitting the data to Logistic distribution gives a small p-value (p = 0.487).



Instead, we found that the distribution of the Google data is better described by the **Laplace Distribution**. Below are the histograms of the log-return values of Google's adjusted closing price (top) and of a random sample from the Laplace Distribution (bottom). We can visually see that the two histograms are similar. Numerically validating our guess using the fitdist and ad.test functions in r, fitting the log-return values of the Google's adjusted closing price to Laplace distribution gave a p-value of p = 0. 8944.


Histogram
Density
0
5
10
15
20
25
30
35
-0.05
0.00
0.05
0.10
Data


Histogram of a Random Sample from Laplace Distribution
Density
0
10
20
30
40
-0.05
0.00
0.05
0.10
Gsample

1(c)

Parameter	Estimate
Minimum	-0.07798239
Maximum	0.09185175
Median	0.001176109
Mean	0.0008878837
Estimated SD	0.01660324
Estimated Skewness	0.07202605
Estimated Kurtosis	9.46819

p-value = 0.8944

1(d)

Proof of E[R | R < c] = $\int_{-\infty}^{c} \frac{r f_R(r)}{F_R(c)} dr$**:**

For any constant c:

$$\begin{aligned}
\mathrm{E}(R|R<c) &= \int_{-\infty}^{\infty} r f_{R|R<c}(r) dr \\
&= \int_{-\infty}^{\infty} r \cdot \frac{f_R(r) \cdot 1_{R<c}}{P(R<c)} dr, \text{ where } (1_{R<c}) \text{ represents a } \textit{Delta Function} \\
&= \frac{1}{P(R<c)} \int_{-\infty}^{c} r f_R(r) dr \\
&= \int_{-\infty}^{c} \frac{r f_R(r)}{F_R(c)} dr
\end{aligned}$$

PDF of Laplace Distribution: $f_R(r) = \frac{1}{2\sigma} e^{-\left|\frac{r-\pi}{\sigma}\right|}$

From Wolodzko, T. (2019). Additional Univariate and Multivariate Distributions. Package 'extraDistr': Laplace.

Comparing Quantiles and Estimates from 2(a) to those estimated using MC method:

	Google log-return adjusted closing price (from 2(a))	Monte Carlo
$Q_{0.05}(R)$	-0.02593215	-0.02513841
$E[R \mid R < Q_{0.05}]$	-0.03802518	-0.03640205

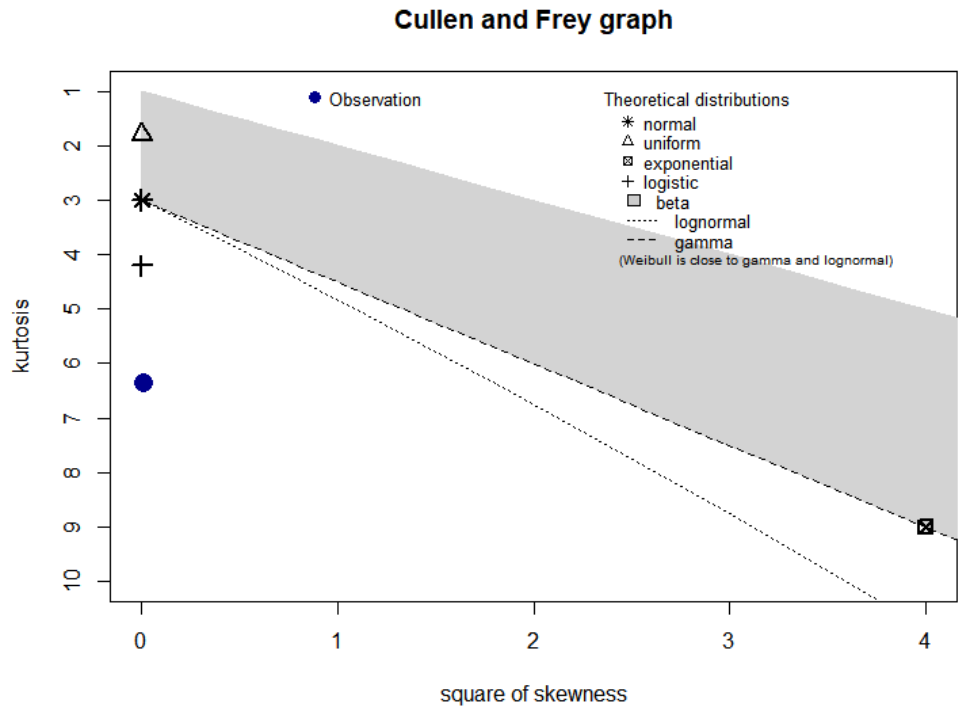
AMAZON

2(a) $\mathbf{Q_{0.05}(R)}$ = -0.03182392

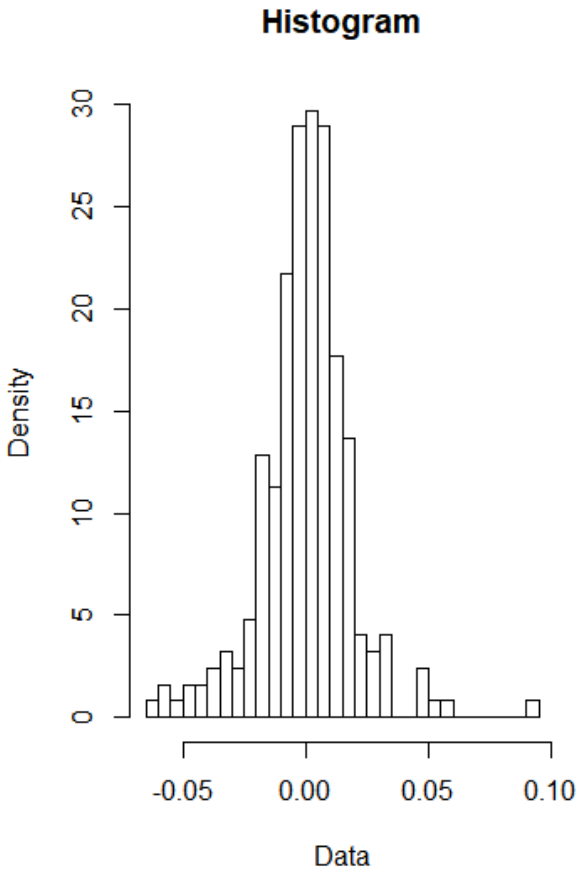
$\mathbf{E[R \mid R < Q_{0.05}]}$ = -0.04471782

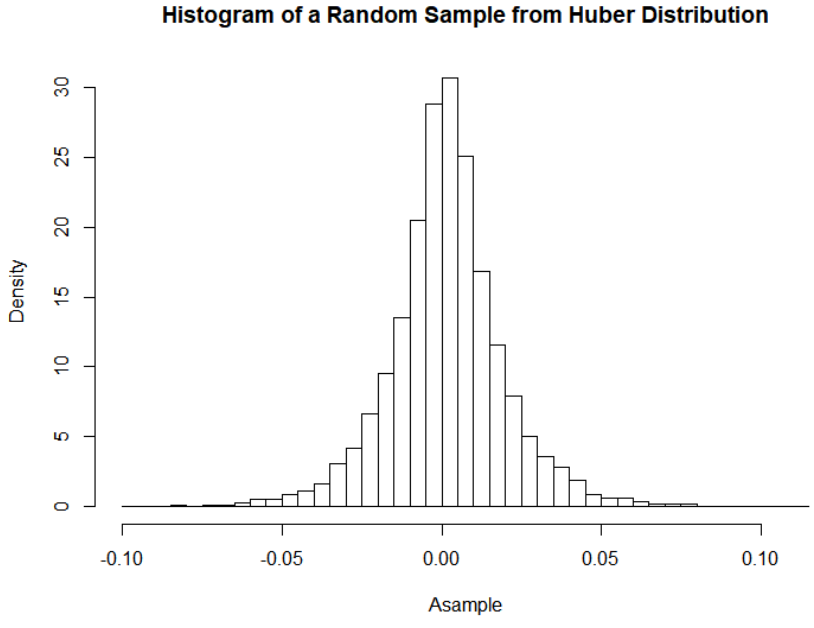
2(b)

According to the Cullen and Frey graph created with the descdist function in R, the distribution of our Amazon data does not match exactly with any of the common theoretical distributions plotted on the same graph. Of the families of distributions listed on the graph, Logistic comes the closest. However, fitting the data to Logistic distribution gives a small p-value (p = 0.3383).



Instead, we found that the distribution of the Amazon data is better described by the **Huber Distribution**. Below are the histograms of the log-return values of Amazon's adjusted closing price (top) and of a random sample from the Huber Distribution (bottom). We can visually see that the two histograms are similar. Numerically validating our guess using the fitdist and ad.test functions in r, fitting the log-return values of the Amazon's adjusted closing price to Huber distribution gave a p-value of p = 0. 8542.


Histogram
Density
0
5
10
15
20
25
30
-0.05
0.00
0.05
0.10
Data


Histogram of a Random Sample from Huber Distribution
Density
0
5
10
15
20
25
30
-0.10
-0.05
0.00
0.05
0.10
Asample

2(c)

Parameter	Estimate
Minimum	-0.06044689
Maximum	0.09025402
Median	0.001339325
Mean	0.0003499504
Estimated SD	0.01858348
Estimated Skewness	0.1048055
Estimated Kurtosis	6.372742

p-value = 0.8542

2(d)

PDF of Huber Distribution:

$$f_R(r) = \frac{1}{2\sqrt{2\pi}\left(\Phi(k) + \phi\left(k \Big/ k - \frac{1}{2}\right)\right)} e^{-\rho_k(r)}$$

where

$$\rho_k(r) = \frac{1}{2}r^2 \qquad |r| \leq k$$

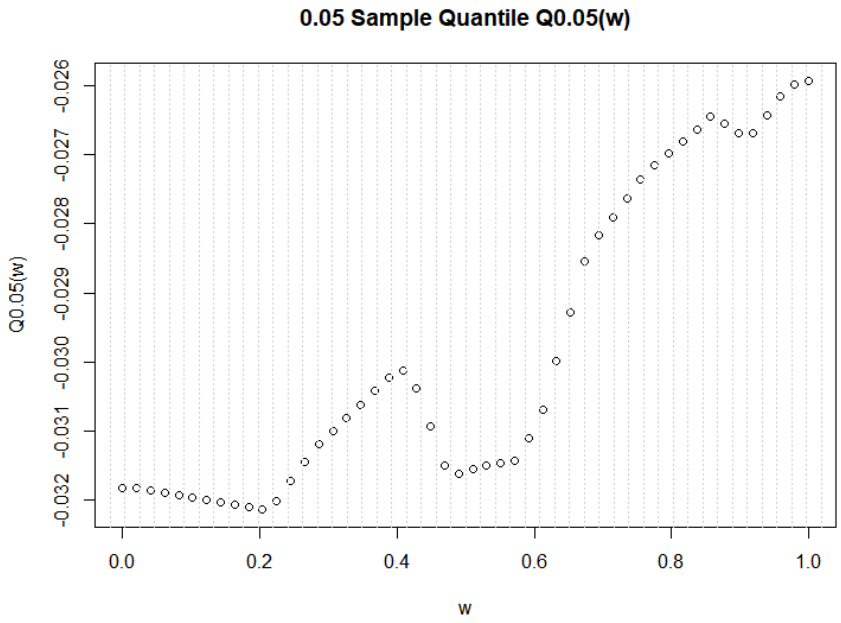
$$\rho_k(r) = k|r| - \frac{1}{2}k^2 \quad |r| > k$$

From Wolodzko, T. (2019). Additional Univariate and Multivariate Distributions. Package 'extraDistr': Huber.

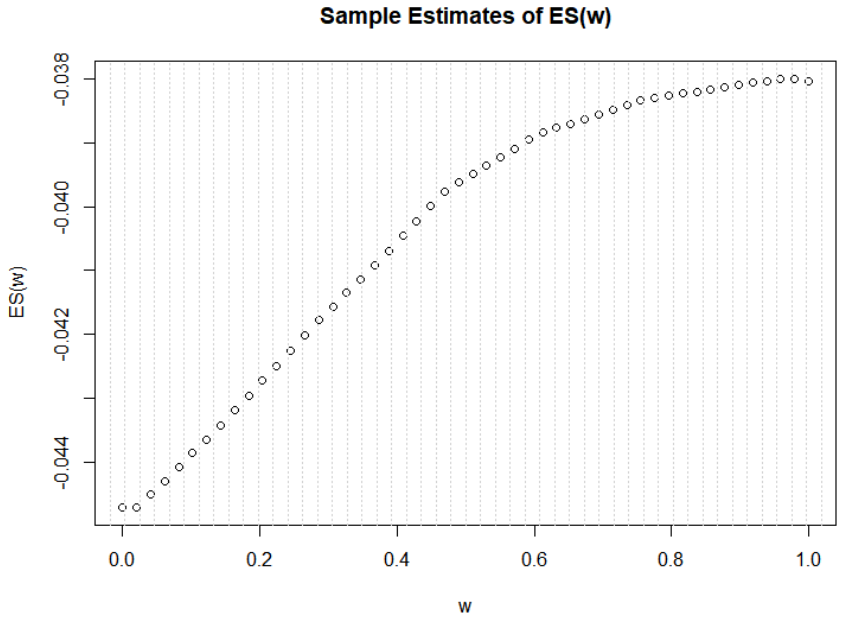
Comparing Quantiles and Estimates from 2(a) to those estimated using MC method:

	Amazon log-return adjusted closing price (from 2(a))	Monte Carlo
$Q_{0.05}(R)$	-0.03182392	-0.02837518
$E[R \mid R < Q_{0.05}]$	-0.04471782	-0.04084089

3(a)



3(b)



3(c)

Based on a grid search of 50 values w ϵ [0/49, 49/49], ES(w) is minimized when w = 1/49 and maximized when w = 48/49.