Insurance Risk | Project #1

Robert Mead

7/5/2021

Introduction

An insurance company has \$1,000,000 in assets and 1,000 customers paying a \$5,000 premium at the beginning of each year. The scope of the project is to find the probability of bankruptcy for the insurance company and estimate the amount of money the bank will make in a specific time frame. The Pareto Distribution, $f(x)$, is used to characterize the size of the insurance claim that a customer has. Similarly, company assets can be characterized by a piece wise function $Z(t)$. To encapsulate a full understanding of the Pareto Distribution and its effects on many observations there will be 10,000 simulations of the distribution to provide clarity in how the analytical and experimental statistics compare to each other. Through the extensive simualtion this will provide an opportunity to determine the likelihood of the insurance company from going bankrupt.

The Pareto Distribution,$f(x)$ and the piece wise function,$Z(t)$ are below.

$$f(x) = \begin{cases} \frac{ab^a}{(x+b)^{(a+1)}} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$Z(1) = 1,000,000 \tag{1}$$

$$Z(t) = \begin{cases} max[Z(t-1) + premiums - claims, 0] & t > 1 \\ 0 & t < 1 \end{cases} \tag{2}$$

Analytical Calculations

Expected Value

Given the probability density function of the Pareto Distribution enables calculation of the expected value, $E[X]$, the variance, $Var[X]$, the skewness and excess kurtosis. Analytically calculating these statistics from the distribution will provide a reference point for the accuracy of the 10,000 simulations.

To calculate the expected value, or the first moment, of the probability density function of the Pareto Distribution the following steps are provided below:

$$
\begin{aligned}
E[X] &= \int_0^\infty \frac{xab^a}{(x+b)^{(a+1)}}\,dx && (3)\\
&= ab^a \int_0^\infty x(x+b)^{-(a+1)}\,dx && (4)\\
&= ab^a \int_0^\infty x(x+b)^{-a-1}\,dx && (5)\\
&= ab^a\left(\frac{-x(x+b)^{-a}}{a} + \int_0^\infty \frac{(x+b)^{-a}}{a}\,dx\right) && (6)\\
&= ab^a\left(\lim_{n\to\infty}\frac{-x(x+b)^{-a}}{a} + \lim_{n\to\infty}\left(\frac{(x+b)^{-a+1}}{a(-a+1)}\right)\right)\Big|_0^n && (7)\\
&= ab^a\left(\lim_{n\to\infty}\frac{-x(x+b)^{-a})(-a+1)+(x+b)^{-a+1}}{a(-a+1)}\right)\Big|_0^n && (8)\\
&= ab^a\left(\frac{-n(n+b)^{-a})(-a+1)+(n+b)^{-a+1}}{a(-a+1)} + \frac{0(0+b)^{-a})(-a+1)+(0+b)^{-a+1}}{a(-a+1)}\right) && (9)\\
&= ab^a\left(\frac{(0+b)^{-a+1}}{a(-a+1)}\right) && (10)\\
&= ab^a\left(\frac{b^{-a+1}}{a(a-1)}\right) && (11)\\
&= \frac{ab^a b^{-a+1}}{a(a-1)} && (12)\\
&= \frac{b}{a-1} && (13)
\end{aligned}
$$

Therefore, the expected value of the Pareto Distribution is $\frac{b}{a-1}$. When the parameters a and b are given with $a = 5$ and $b = 200,000$ the expected value of the Pareto Distribution can be calculated so that the $E[X] = 50,000$. Calculations completed in R can confirm that the expected value of the Pareto Distribution is 50,000.

```
#packages
library('cubature')

#Expected Value
moment_1 <- function(x) {((5)*(200000^5)*(x))/((x+200000)^(6))}
expected_value <- adaptIntegrate(moment_1,lowerLimit = 0,upperLimit = Inf)
expected_value$integral
```

```
## [1] 50000
```

Variance

The difference of the second moment and the first moment squared are used to calculate the variance of the probability density function of the Pareto Distribution. The steps provided will show how to derive the variance of the Pareto Distribution:

$$Var[X] = E[X^2] - E[X]^2$$

In order to calculate the variance, the second moment is required. The derivation of the second moment of the Pareto Distribution is worked out below:

$$
\begin{aligned}
E[X^2] &= \int_0^\infty \frac{x^2 a b^a}{(x+b)^{(a+1)}}\, dx && (14)\\
&= ab^a \int_0^\infty x^2 (x+b)^{-(a+1)}\, dx && (15)\\
&= ab^a \int_0^\infty x^2 (x+b)^{-a-1}\, dx && (16)\\
&= ab^a \left(\frac{-x^2(x+b)^{-a}}{a} + \int_0^\infty \frac{-2x(x+b)^{-a}}{a(-a+1)}\, dx \right) && (17)\\
&= ab^a \left(\frac{-x^2(x+b)^{-a}}{a} - \left(\frac{2x(x+b)^{-a+1}}{a(-a+1)} - \int_0^\infty \frac{2(x+b)^{-a}}{a(-a+1)})\, dx \right) \right) && (18)\\
&= ab^a \left(\lim_{n\to\infty} \frac{-x^2(x+b)^{-a}}{a} - \lim_{n\to\infty} \left(\frac{2x(x+b)^{-a+1}}{a(-a+1)} - \frac{2(x+b)^{-a+2}}{a(-a+1)(-a+2)} \right) \right) \Big|_0^n && (19)\\
&= ab^a \left(\frac{2b^{-a+2}}{a(a-1)(a-2)} \right) && (20)\\
&= \frac{2b^2}{(a-1)(a-2)} && (21)
\end{aligned}
$$

Now, being able to substitute the second moment into the equation will enable for the complete calculation of the variance of the Pareto Distribution.

$$
\begin{aligned}
Var[X] &= E[X^2] - E[X]^2 && (22)\\
&= \frac{2b^2}{(a-1)(a-2)} - \left(\frac{b}{a-1} \right)^2 && (23)\\
&= \frac{2b^2}{(a-1)(a-2)} - \frac{b^2}{(a-1)^2} && (24)\\
&= \frac{2b^2(a-1) - b^2(a-2)}{(a-1)^2(a-2)} && (25)\\
&= \frac{ab^2}{(a-1)^2(a-2)} && (26)
\end{aligned}
$$

Therefore, the variance of the Pareto Distribution, given the parameters $a = 5$ and $b = 200,000$ we can calculate that the variance is $Var[X] = E[X^2] - E[X]^2$ which is $Var[X] = 4.16 \times 10^9$. Calculations completed in R can confirm that the variance of the Pareto Distribution is 4.16×10^9

```
#packages
library('cubature')

#Variance
moment_2<- function(x) {(((5)*(200000^5)*(x^2))/(x+200000)^(5+1))-(((5)*(200000^5)*(x))/(x+200000)^(5+1
variance <- adaptIntegrate(moment_2,lowerLimit = 0,upperLimit = Inf)
theovar <- variance$integral - expected_value$integral^2
theovar
```

```
## [1] 4166656566
```

Skewness

To analytically calculate the skewness of the Pareto Distribution would require to find the third moment. The calculations for the third moment are intensive and require integration by parts three times over. The remainder of the calculations for the skewness and excess kurtosis will be completed through R coding. The skewness of the Pareto Distribution is as follows when $a = 5$ and $b = 200,000$.

```
#packages
library('cubature')

#Skewness
moment_3 <- function(x) {((5)*(200000^5)*(x^3))/((x + 200000)^(5+1))}
skewness <- adaptIntegrate(moment_3,lowerLimit = 0, upperLimit = Inf)
skewness$integral
```

```
## [1] 2e+15
```

Excess Kurtosis

The kurtosis describes the thickness of the tail in the Pareto Distribution. To compute the excess kurtosis, the code shows that first the fourth moment needs to be computed. The fourth moment of the Pareto Distribution is the calculation for kurtosis. Then, to classify the excess kurtosis the fourth moment is subtracted by 3.

```
moment_4 <- function(x) {((5)*(200000^5)*(x^4))/((x + 200000)^6)}
kurtosis <- adaptIntegrate(moment_4,lowerLimit = 0,upperLimit = Inf)
kurtosis$integral
```

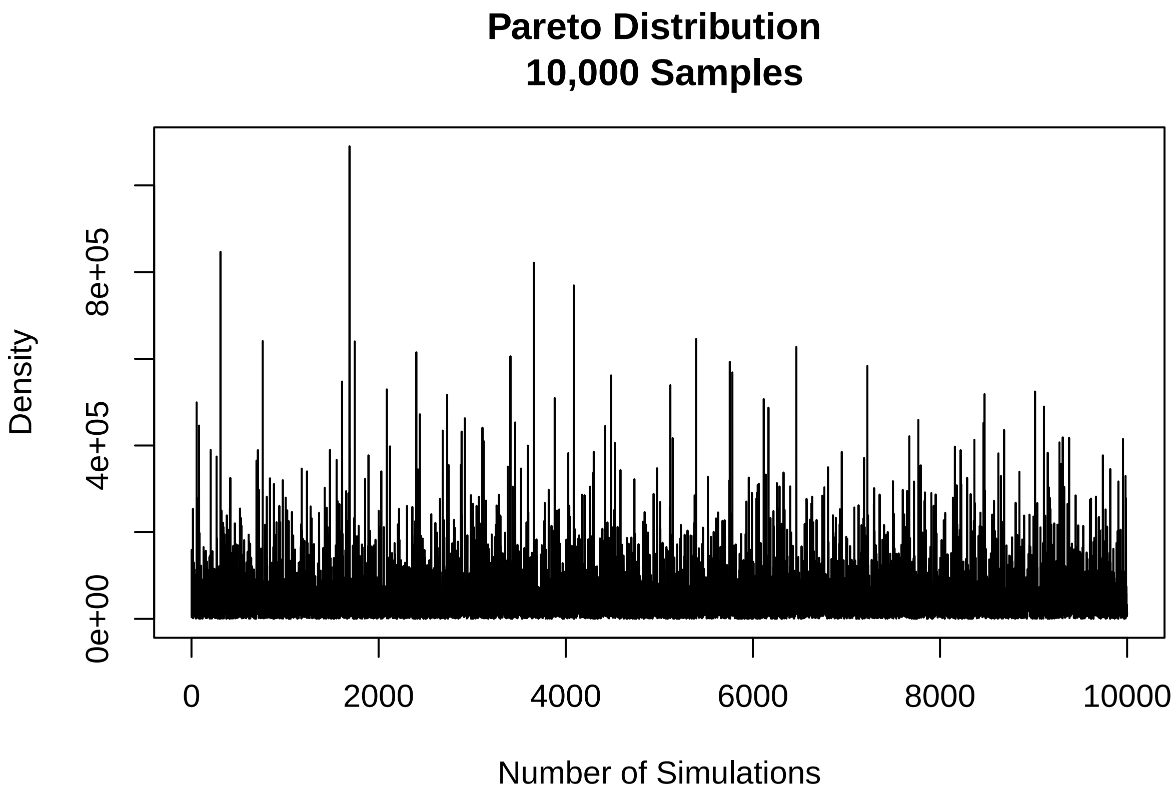
```
## [1] 1.6e+21
```

```
excess_kurtosis <- kurtosis$integral - 3
excess_kurtosis
```

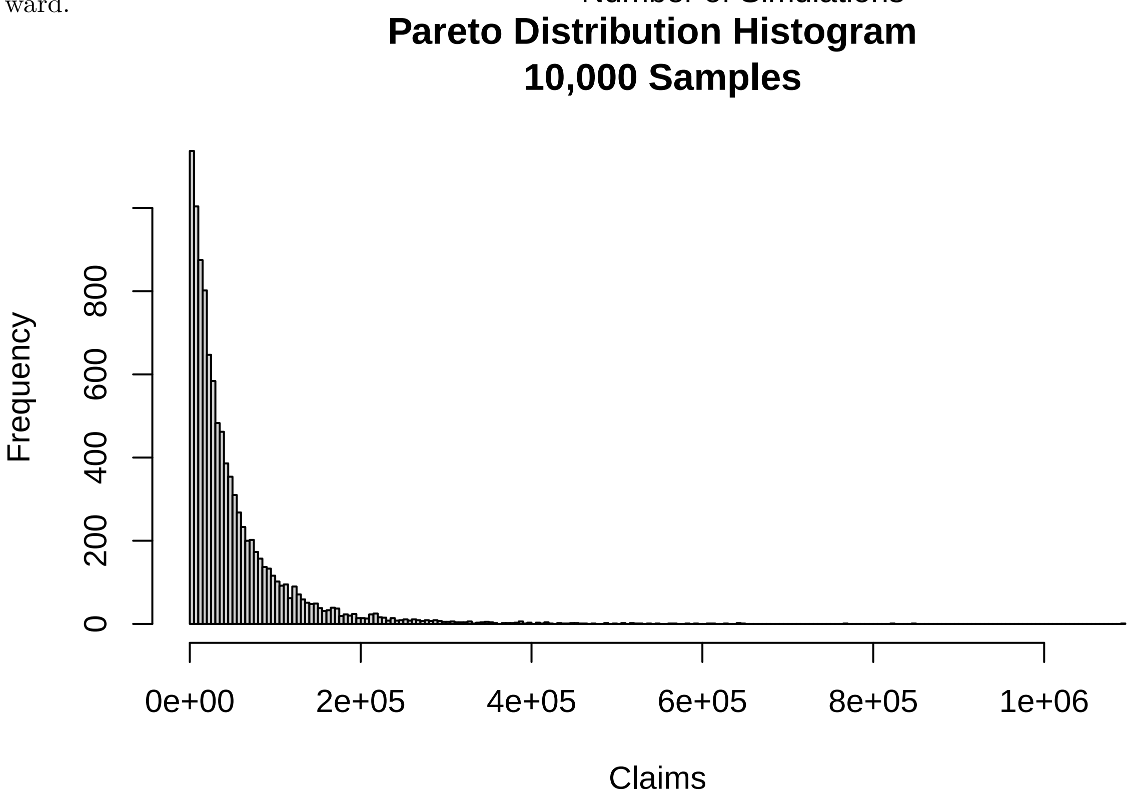
```
## [1] 1.6e+21
```

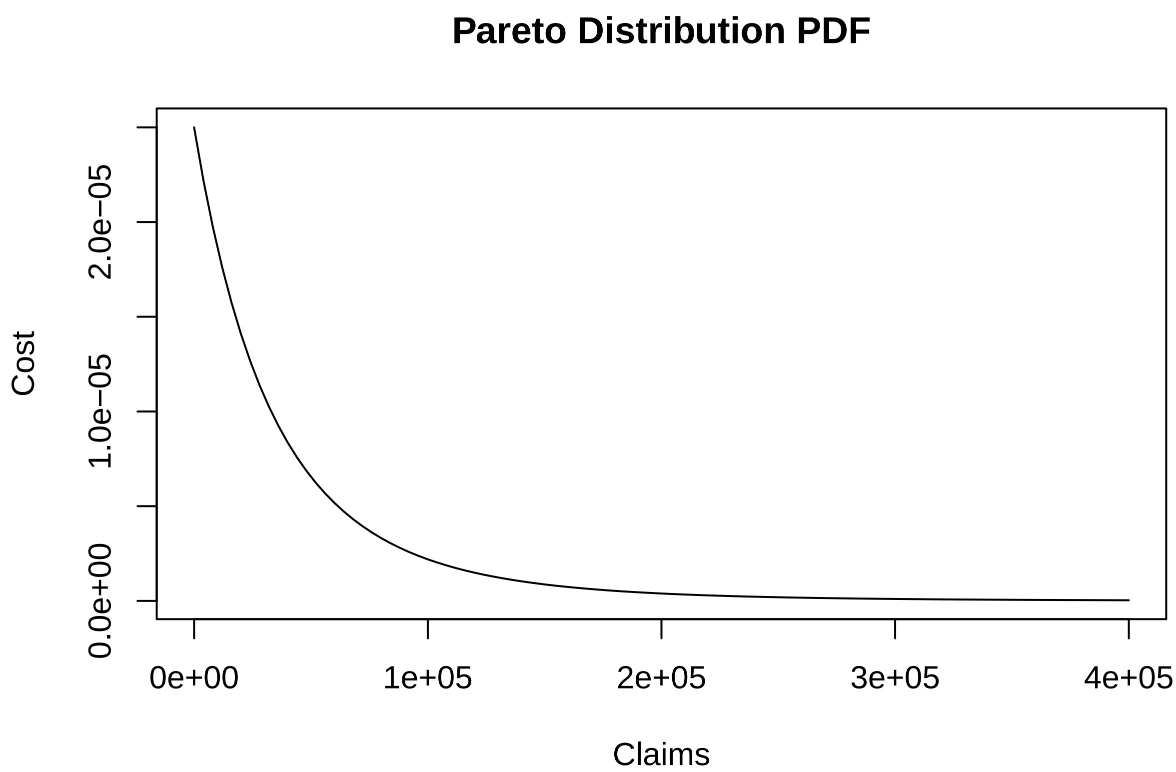
Simulations

Creating 10,000 samples of the Pareto Distribution using the following code below and the packages *fBasics* and *actuar* allowed for graphics that made for a better depiction of the 10,000 samples. The histogram of the 10,000 samples follows that of the density function of the Pareto Distribution. The histogram visualization and the Pareto Distribution are similar, in that, they are skew right and most of the data is left-


Pareto Distribution
10,000 Samples
Density
0e+00
4e+05
8e+05
0
2000
4000
6000
8000
10000
Number of Simulations

ward.


Pareto Distribution Histogram
10,000 Samples
Frequency
0
200
400
600
800
0e+00
2e+05
4e+05
6e+05
8e+05
1e+06
Claims


Pareto Distribution PDF
Cost
0.0e+00
1.0e−05
2.0e−05
0e+00
1e+05
2e+05
3e+05
4e+05
Claims

Experimental Statistics Analyzing the 10,000 samples that were created enable a comparison of the theoretical expected value, variance, skewness and excess kurtosis to determine how accurate the sampling is in comparison to the theoretical statistics. The values below elaborate on the values of the experimental mean, variance, skewness and excess kurtosis.

The experimental mean of the 10,000 samples is stated below.

```
## [1] 50033.12
```

```
## [1] 50033.12
```

The experimental variance of the 10,000 samples is stated below.

```
## [1] 47979.28
```

```
## [1] 6477823101
```

```
## [1] 4175811675
```

```
## [1] 4176229298
```

The experimental skewness of the 10,000 samples is stated below.

```
## [1] 1.873789e+15
```

The experimental excess kurtosis of the 10,000 samples is stated below.

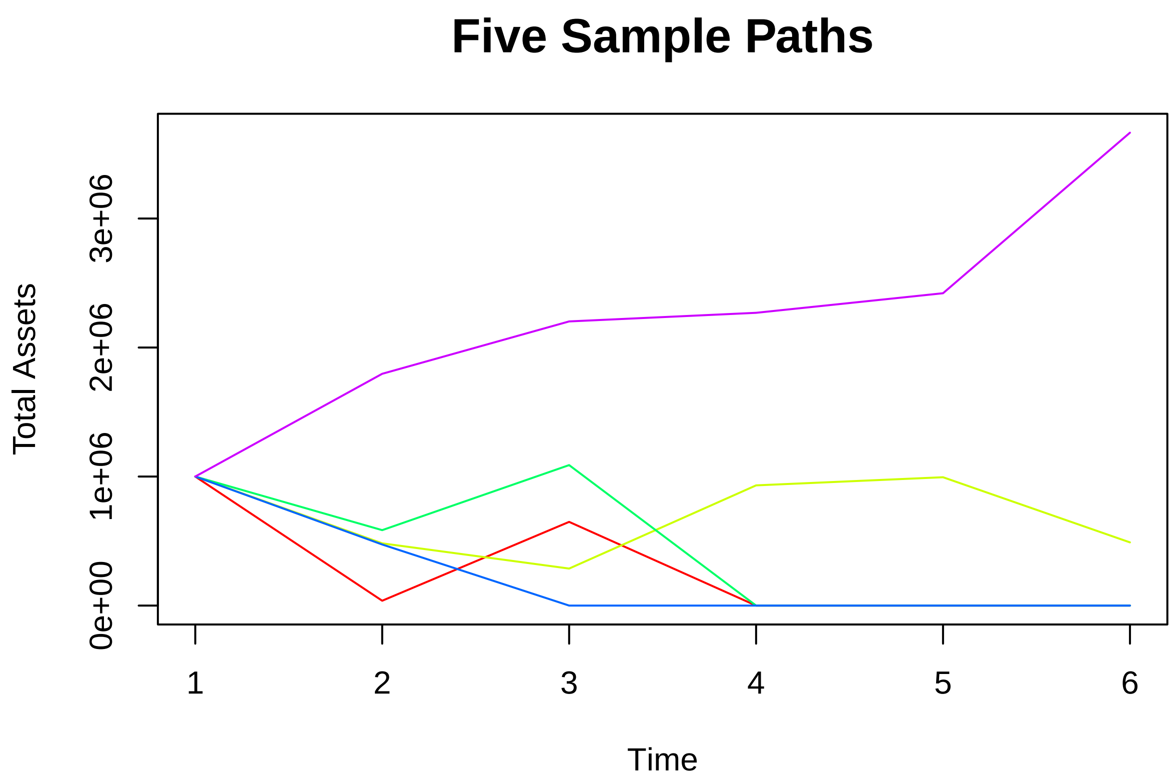
```
## [1] 1.62027e+21
```

The relationship the experimental and theoretical mean, variance, skewness and excess kurtosis will be discussed. In 10,000 samples, the experimental mean and theoretical mean are practically similar where the theoretical mean was 50,000 and the experimental mean of the 10,000 samples was 50,997.34 - values that are close to each other. The theoretical variance and experimental variance were also similar, but not terribly close in value. The theoretical variance was 4,166,656,566 and the experimental variance was 4,701,382,046 - differing by almost 600,000,000. The theoretical skewness and experimental skewness were similar in value. The theoretical skewness was 200,000,000,000,000 and the experimental skewness was 204,756,200,000,000. The excess kurtosis were the two theoretical and experimental values that were not similar, where the theoretical value was measured to have an excess kurtosis of $1.6 \cdot 10^{21}$ and the theoretical excess kurtosis was measured to be $7.536 \cdot 10^{20}$.

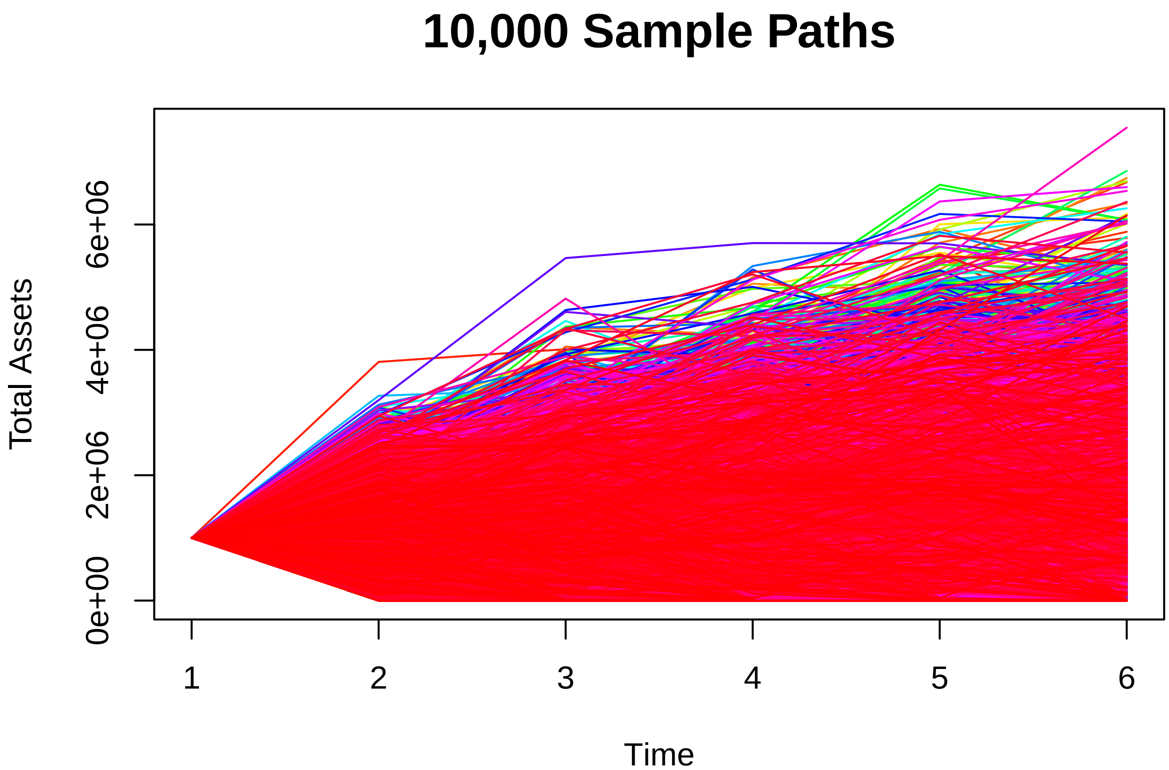
Overall, the theoretical and experimental values are consistent with each other, except for the theoretical and experimental excess kurtosis.

Probability of Bankrupcy

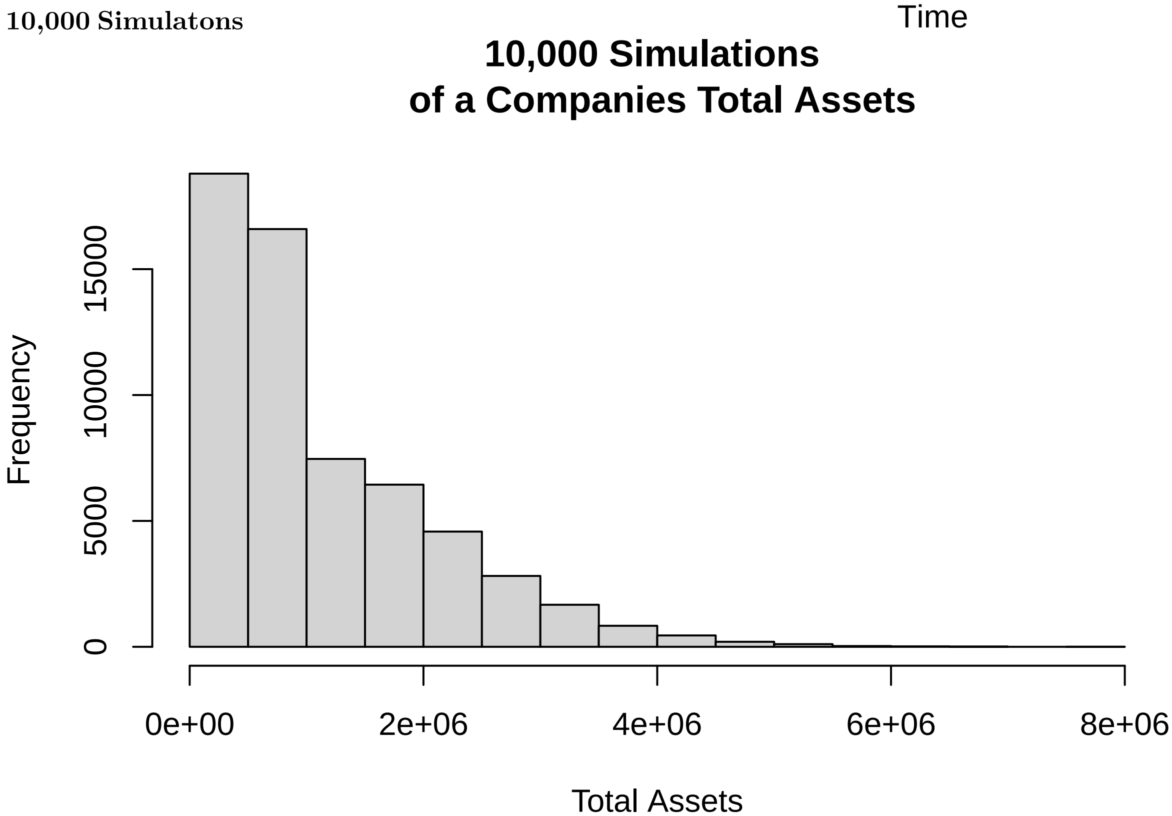
Five Simulations The companies assets can be characterized through the piecewise function $Z(t)$ with an initial value of $Z(1) = 1,000,000$. The function $Z(t)$ can help determine the companies assets at the end of each year through five years. Below are five sample paths of the Companies Assets through five years. The sample paths are able to be simualted due to the fact that the function that characterizes insurance claims is practically the same theoretically and analytically. This was proven in the first section.



The five sample paths represent the insurance companies total assets during a five year span. As you can see from the graph above the five sample paths yield varying paths with different terminating values at the end of the five years. Some of the paths lead to total loss of the assets having the company go bankrupt, while other paths lead to loss in assets from the initial value of \$1,000,000, and other paths leading to a gain in total assets that are larger than the initial value of \$1,000,000.



10,000 Simulatons



Probability of Bankruptcy The probability of bankruptcy is calculated when the simulated value reaches the x-axis. Having the unique simulation reach zero at any point in the five years represents that the insurance company has gone bankrupt since their total assets will be zero. To calculate the likelihood that the insurance company goes bankrupt is by completing the simple probability of the number of simulations that reach zero over the total number of simulations.

Looking at the 10,000 simulations is an impossible way to calculate the number of simulations that reached zero, so organizing the information as a histogram allows for a better visualization as to how many simulations have hit bankruptcy.

The probability that the insurance company goes bankrupt in the 60,000 total simulations is 0.02375.

Expected Value E[Z(5)]

```
## [1] 1188835
```

The expected value of the total assets of the insurance company in it's fifth year would be calculated above.

Conclusion

The simulations conducted for the Pareto Distribution provide stability in the simulations that are completed for the insurance company and determining the value of their assets through five years. Having the theoretical and experimental statistics remain consistent after 10,000 simulations allowed for the continuation of the simulation of the total assets the bank had through five years.

In simulating the total assets for five years using the piece wise function, this showed how likely it is that an insurance company would go bankrupt. The probability was computed by taking the ratio of simulations that reach zero within the five year span to the total number of simulations.

Finally, computing the expected value of the total assets after five years shows how much the insurance company has in total assets.