# Quantitative Methods in Finance

Guessing game: Final Report

Julien Dos Reis

June 2023

Professor : [redacted]

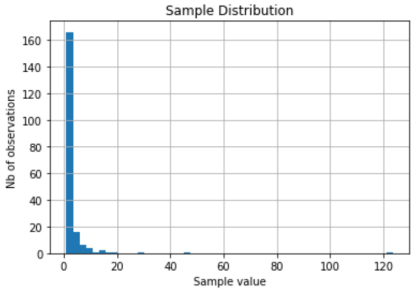
# 1 DGP Process

It refers to the underlying mechanism or model that generates the observed data. In other words, the DGP describes how the data points are generated or created based on certain assumptions or rules.

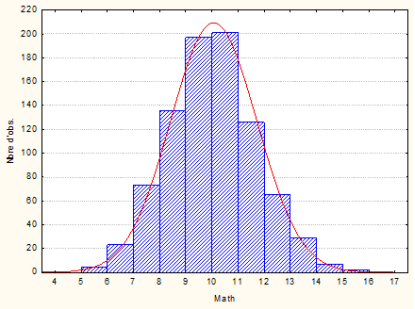
# 2 Presentation of the sample

We had the choice between 7 simple parametric distributions. The first part was to plot the distribution. Plotting distributions can provide valuable insights and intuitions about their characteristics. Visualizing distributions allows you to observe patterns.

To illustrate, here is the sample distribution and a normal distribution to compare:




Distribution of the sample




Normal law distribution

Comparison of the distributions

We can clearly see that our distribution doesn't follow a normal law. This give us a first direction on which distribution we should focus on, such as maybe pareto or exponential distributions which are close. It's important to note that while visualizing distributions provides intuitive insights, it should be supplemented with formal statistical tests and analyses to make accurate conclusions about the goodness of fit or to derive precise statistical measures.

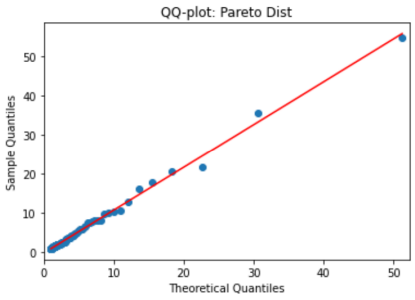
# 3 Statistical Study

First I decided to remove the outlier around 120. Knowing that the 75% quantile was around 2.35, this value was far too far from our sample. However, these kind of laws have different characteristics in their tails distribution. Exponential law are considered to have light tails compared to Pareto for example. I decided to keep the other outliers above 20 because understanding the tail behavior is crucial when applying it to real-world scenarios. It helps in assessing the likelihood of extreme events.

To support these hypotheses, as I said, we need to conduct statistical tests. I used the combiantion of QQ-plots and Kolmogorov-Smirnov test.

I first conducted these tests on Pareto and Exponential distribution and ended up having better fit on the QQ-plot with a p-value at 0.943 on the ks-test. In this case we fail to reject the null hypothesis that the obeserved sample deviates from the theoritical distribution. Pareto dsitribution following a power-law distribution, it is characterized by a heavy right tail which we can see in the QQ-plot below, in some of the highest quantiles.

Here is the QQ-plot and the parameters of pareto distribution:




QQ-plot of the sample fitted

| | | |
|---|---|---|
| pareto | loc | 0.16751 |
| pareto | scale | 0.840862 |
| pareto | shape | 1.34602 |

Parameters of the fitted sample

Sample fitted with Pareto distribution

# 4 Conclusion

To conclude, we will design the premium for one unit to be insured based on our historical sample.

We will use our parameters to do this since they represent the characteristics of the distribution. The shape controls for the tail behavior of the distribution. A shape of 1.34602 indicates a slow decline in the tail resulting in higher probability for extreme events. It is important because it reflects that a small portion of individuals contribute to a high portion of claims. Using a Value-at-Risk approach to fix this premium, with a confidence level of 95%, the result is 7.78. It means that there is a 5% chance that the loss experienced by the insurer will exceed 7.78. It seems high but it's caused by Pareto distribution characteristics with a highest risk of extreme values. With this level of VaR, while adding 5% business/managing costs, we finally get a premium of 8.60 in order to be covered.