Option Pricing Beyond Black–Scholes: Empirical Evidence of Fat Tails

Tommaso Melotti

2025

1 Introduction

The Black–Scholes–Merton (BSM) model represents one of the cornerstones of modern financial mathematics. By providing a closed-form solution for option pricing, it revolutionized both theory and practice in derivatives markets. However, the model rests on a number of highly restrictive assumptions, such as constant volatility, frictionless trading, and normally distributed returns of the underlying asset. These simplifications make the framework mathematically elegant but poorly aligned with empirical evidence.

In this work, we first present the theoretical foundation of the BSM model, including its derivation through stochastic calculus and the hedging arguments that eliminate randomness under idealized conditions. We then contrast the theoretical assumptions with real-world market dynamics. Using long-term historical data on the S&P 500, we test whether returns are consistent with normality and highlight several deviations, such as the presence of fat tails and extreme events that the BSM framework essentially rules out.

The goal of this project is twofold: (i) to illustrate the mathematical and economic rationale behind the Black–Scholes approach, and (ii) to show its empirical limitations through data analysis. This dual perspective emphasizes the gap between elegant models and messy markets, motivating the need for richer stochastic processes in option pricing.

2 Characterization of the Black-Scholes model.

A trader who uses a theoretical pricing models is exposed to two type of risk, namely the risk of having the wrong inputs and the risk that the model itself is wrong because the model itself is based on unrealistic assumptions. To begin, we might list the most important assumptions when dealing with a traditional pricing model as the Black-Scholes.

- The markets are frictionless:
- No liquidity restrictions.

- Unlimited money can be borrowed or lent at the same interest rate.
- No transactions costs.
- No tax consequences.
- Interest rates are constant over the life of the contract.
- Volatility is constant over the life of the contract.
- Trading is continuous, this means that no gaps in the price of the underlying are allowed.
- The percent price changes in the underlying are normally distributed, in other words the prices of the underlying are lognormally distributed.

We can already see at a first glance that these assumptions are quite unrealistic in the real world, but we can be more precise. Let's first consider the Black-Scholes model assuming these assumptions are in order, moreover, let's assume that two assets are tradable in the market, the *non-risky asset B* and the *non-risky asset S*. We assume that the price process of B and S are as follows:

$$B_t = exp(rt) \tag{1}$$

$$S_t = S_0 exp(\mu t + \sigma W_t - \frac{1}{2}\sigma^2 t) \tag{2}$$

Where r is the short rate and W_t is a Standard Brownian Motion. It is trivial to verify that S_t is a Geometric Brownian Motion, in other words it solves the following SDE with initial condition $S_0 > 0$:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{3}$$

Denoting with $C(S,t)$ the price of a call at time t with underlying S we can apply Ito's formualae and deriving the following PDE:

$$dC = \left(\frac{\partial C}{\partial t} + \mu S_t \frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 C}{\partial S^2} \right) dt + \sigma S_t \frac{\partial C}{\partial S} dW_t \tag{4}$$

Let us now consider a **self-financing** strategy, we call x_t the position at time t on the non risky asset, and y_t the position at time t on the risky asset. By definition, the value at time t of our strategy is

$$V_t = x_t dBt + y_t dS_t \tag{5}$$

and by replacing dB_t with $rdB_t dt$ and dS_t with 1.3 we obtain.

$$dV_t = (rx_t B_t + y_t \mu S_t)dt + y_t \sigma S_t dW_t \tag{6}$$

We can now equate the corresponding terms in 1.4 with those in 1.6, obtaining

$$y_t = \frac{\partial C}{\partial S} \tag{7}$$

$$rx_tBt = \frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 C}{\partial S^2} \tag{8}$$

Now if we set $C_0 = V_0$ then by non arbitrage assumptions, it must be that $C_t = V_t$ for every t since the dynamics are exactly the same. By replacing 1.7 and 1.8 in the equation 1.5 we get the famous Black-Scholes PDE:

$$rS_t \frac{\partial C}{\partial S} + \frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 C}{\partial S^2} - rC = 0 \tag{9}$$

The derivation of the formula has used a *Dynamic Delta Hedging* strategy, indeed by defining y_t as the derivative of the Call with respect to the underlying S (i.e. the Delta) we are continuously hedging our replicating portfolio following the price process of S_t, thus "eliminating" the risk characterized by the randomness of the underlying.

This approach has been used long before Black and Scholes; the relatively new approach proposed by the two was simply a theoretical argument that fitted well with the economic assumption of an efficient market.

3 The Real World.

By now it is pretty clear that the assumption and the strategy proposed is not only unrealistic, but also impracticable and sometimes dangerous when facing the reality of the market. Let's start by saying that, nor the constant interest rate nor the constant volatility are plausible assumptions, indeed raising interest rates raises the forward price, which raises the values of the call and lowers the values of the puts (by a Put-Call parity argument); and regarding the volatility, it is well known that over the life of an option strategy, a trader will encounter periods of high volatility together with periods of low volatility; moreover, a continuous dynamic hedging is indeed impossible due to transaction costs and tax constraints.

But the assumption that most concern us is the fact that prices are assumed to follow a precise probability distribution, the continuous diffusion process governed by Brownian motion may be indeed a reasonable approximation of how prices change, but is clearly not perfect. Consider, for instance, that instead of a continuous diffusion process prices follow a jump-diffusion process, that is, a combination between continuous process and a jump process (note that this assumption can be easily supported by historical evidence). How is this likely to affect values generated by the model? To understand the effect if a gap would occur, consider a simple strategy like an at-the-money short straddle with the underlying at 100, (i.e. sell one call and one put, both ATM), if a gap happens with only a very short time remaining to maturity, there is no possibility of hedging, the operator would find himself short deeply in-the-money with a delta close to 1 without even considering the second orders effect due to the implied volatility and the gamma of the option. This rather simple, but practical, example shows how dangerous can be assume a continuous diffusion process approximation.

3.1 Are returns normally distributed?

In this section, we will try to investigate the real nature of markets from a probabilistic distribution point of view. The following arguments are at the core of what will come next, namely, we will prove via statistical observation that markets' returns are not normally distributed, indeed many large deviations have been observed, these events would have been considered impossible in a Black and Scholes world. The conclusions from this diagnosis will suggest the presence of *fat tails*, opening to new models that better fit the real behavior of the market.

Figure 1.1 is a histogram of daily Standard and Poor's 500 Index price changes from 1927-01-01 through 2024-12-30. One can instantly observe that, most of the changes are relatively small and close to zero. The empirical distribution seems to resemble the normal distribution, hence, we have computed the empirical mean (0.030%) and the standard deviation (1.19%), (see Figure 1.2) and fitted a normal distribution with those parameter. The red line that overlaid the histogram is the PDF obtained. One can point out that returns near zero are more frequent than predicted by the normal distribution, indeed the bars rise above the red curve, it seems that the fitted normal distribution underestimates the frequency of small changes. Moreover, although they are not particularly visible, there are several outliers that rise above the tails of the normal distribution, these outliers seem to suggest that the assumption of normally distributed returns underestimates the frequency of such values. Finally, in the midsection, one can observe that the opposite is true, that is, an overestimation of occurences.

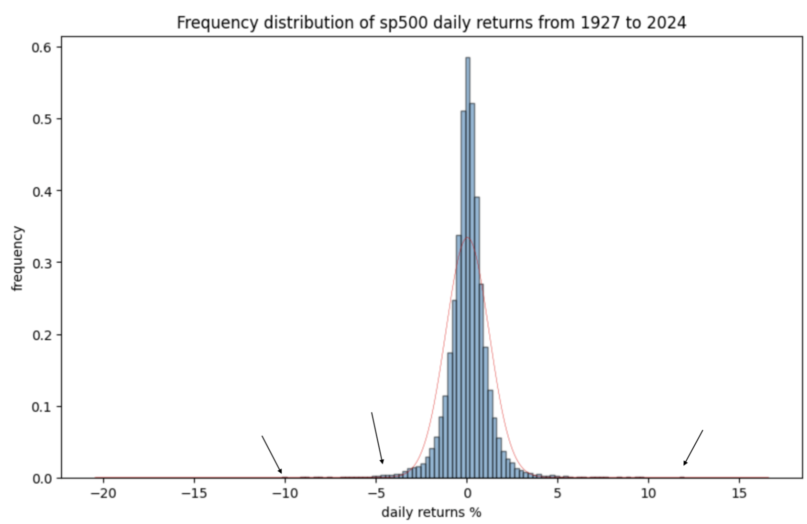

Figure 1: S&P500 daily returns and normal distribution

```
Mean of daily returns: Ticker
^GSPC    0.030993
dtype: float64%
Standard deviation of daily returns: Ticker
^GSPC    1.193248
dtype: float64%
```

Figure 2: Empirical Mean and Standard Deviation of the S&P500 daily returns

3.1.1 Black Monday - 1987

To further grasp the idea of how unrealistic the assumption of normally distributed returns is, we can consider the biggest move in the S&P over the same period, (Black Monday 19 October 1987). In one day the index lost 20.47%, assuming that the mean is 0.030% and the standard deviation is 1.19%, we can compute the number of standard deviation away from the mean as:

$$\frac{20.47\% - 0.03\%}{1.19\%} = -17.17 \tag{10}$$

Such an event, assuming a normal distribution, has a probability of 4.1×10^{-65}. These considerations underline an important aspect, the empirical distribution in figure 1.1 exhibit positive kurtosis. It has higher peaks, (high frequency of small returns), narrow midsection, and elongated tails.

This preamble underlines the necessity of further investigation to better understand the properties of the S&P500, this is why, in the next section, we will engage in several tests and check what picture emerges. [1]. It is important to note that, although the analysis is focused on the s&P500, similar conclusions can be derived by analyzing other indexes or underlyings products. In other words, we don't have picked the unique financial product that fits the hypothesis.

[1] For a broader diagnosis, see: *Nassim Taleb: Statistical Consequences of Fat Tails, Chapter 10*

3.2 If it's not normal, what is it?

The following section's goal is to present some statistical tools to further analyze the returns and successively exploit them to give a better representation of the true nature of markets.

3.2.1 "Maximum to sum plot"

The MS plot allows us to see the behavior of the Law Of Large Numbers for a given moment consider the contribution of the maximum observation to the total, and see how it behaves as n grows larger. Let $X_1, X_2, ..., X_n$ be a sequence of nonnegative i.i.d. random variables and define the partial sum as:

$$S_n(p) = \sum_{i=1}^{n} X_i^p \tag{11}$$

and the partial maximum as:

$$M_n(p) = max(X_1^p, X_2^p, ..., X_n^p) \tag{12}$$

then we can define the maximum to sum ratio as:

$$R_n^p = \frac{M_n^p}{S_n^p} \tag{13}$$

If for p=1,2,3,... $E[X^p] < +\infty$, then

$$R_n^p = \frac{M_n^p}{S_n^p} \longrightarrow^{a.s.} 0$$

3.2.2 "QQ-plot"

Typically the QQ-plot, (Quantile-Quantile plot) is helpful when one needs to compare the distribution of a datasetto a theoretical distribution by plotting the quantiles of the data against the quantiles of the distribution. If the data follows the theoretical distribution, the points will lies in a straight line near the theoretical quantiles. In our anlysis a QQ-plot with the exponential quantiles has been used. We will see later that the exponential distribution is somehow our benchmark for the *heavy tails* distributions.

3.2.3 "Zipf plot"

The Zipf plot is just a *log-log* plot of the empirical survival function [2] of our data. For a Pareto distribution we have a linear relationship between the logarithm of the data and the logarithm of the surviaval function, hence, we can hope to detect a similar behavior when plotting our data.

[2] By survial function we mean the probability of X being larger than k for k positive.

3.2.4 "Mean Excess Function Plot"

Let X be a random variable with $df = F$, then we can define the *Mean Excess Function* as:

$$e(u) = \mathbf{E}[X - u | X > u] = \frac{\int_v^\infty (t-u)dF(t)}{\int_v^\infty dF(t)} \tag{14}$$

Similarly we can define the empirical 'MEF' as:

$$e_n(u) = \frac{\sum_{i=1}^n (X_i - u)}{\sum_{i=1}^n \mathbf{1}_{\{X_i > u\}}} \tag{15}$$

Since the normal distribution has a fast decrease in the tail distribution, under normality assumption we would expect a decreasing function towards zero when u increases. The Mean Excess function measures the expected magnitude of outcomes over a certain threshold u, below some of the important Mean Excess Functions.

- MEF of a Pareto distribution:

$$e^{Par}(u) = \frac{x_m + u}{\alpha - 1},\ \alpha > 1$$

- MEF of an Exponential distribution:

$$e^{Exp}(u) = \frac{1}{\lambda}$$

- MEF of a Normal distribution:

$$e^{Normal}(u) = \sigma \frac{\phi(\frac{u-\mu}{\sigma})}{1 - \Phi(\frac{u-\mu}{\sigma})},\ \alpha > 1$$

We, thereafter, present a statistical diagnosis of the S&P500 returns, exploiting the tools that we have presented above. We started by defining the logarithm of the negative returns as positive quantities, (this practice is common in risk management, it is particularly useful when computing quantities such as Value At Risk and Expected Shortfall).

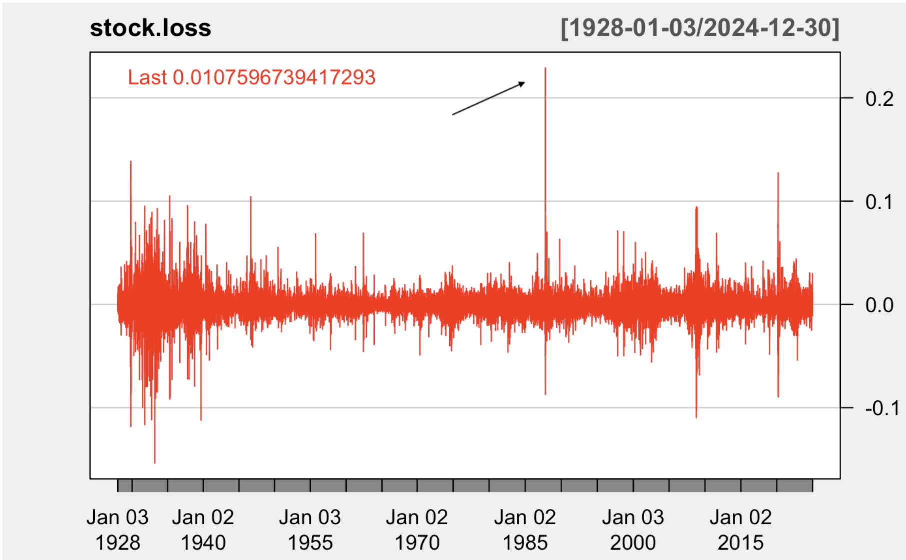

Figure 3: Log-losses of the S&P500.

In Figure 3, we can immediately observe the empirical behavior of market losses. By convention, losses are represented as positive quantities, which allows us to better highlight extreme events in the upper tail of the distribution. The prominent spike corresponds to the 1987 Black Monday crash, already discussed above. This visual evidence reinforces the idea that financial markets deviate substantially from the normality assumption: extreme losses are more frequent and more severe than what a Gaussian framework would predict.

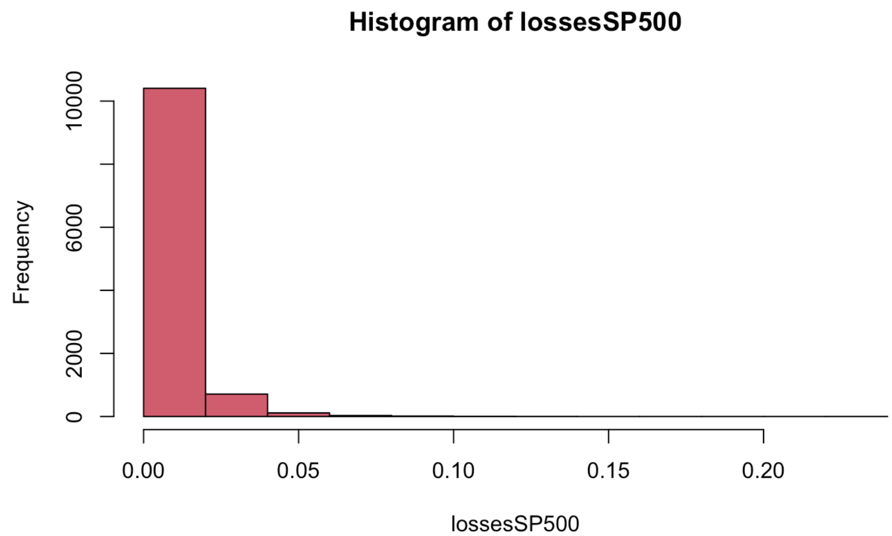

Figure 4: Histogram of the losses of the S&P500 index.

Figure 4 displays the histogram of S&P500 losses. The wide inter-range clearly indicates a high degree of variability in the data, while the asymmetry of the distribution reveals a tendency toward positive skewness. This suggests that extreme losses, although rare, extend further in one direction, reinforcing the departure from the Gaussian assumption.

To further investigate the behavior of the distribution's tail, we proceed with a Zipf plot, presented in the next figure.

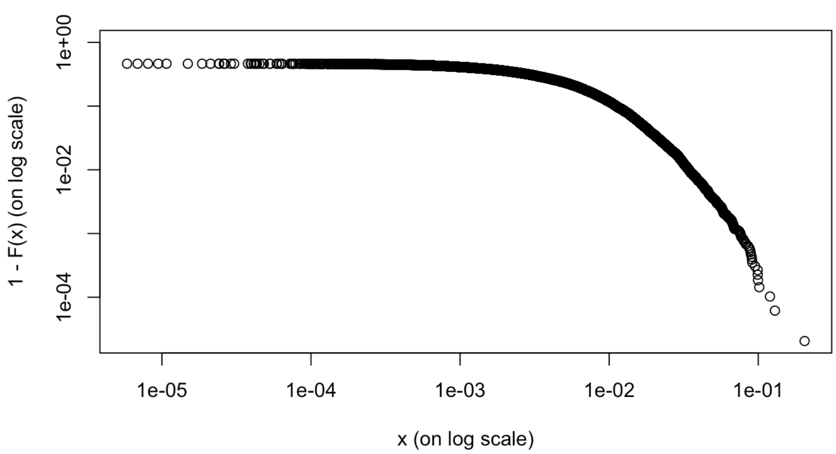

Figure 5: Zipf Plot of absolute negative returns.

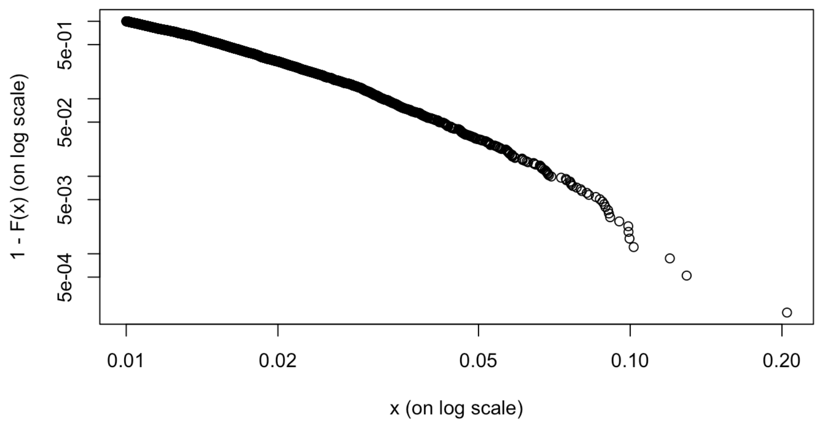

Figure 6: Zipf Plot of absolute negative returns over 0.01.

As previously mentioned, the Zipf plot is a log–log representation of the empirical survival function. In this case, we are interested in detecting possible Paretian behavior of the data. Recall that the survival function of a Pareto distribution is given by

$$F(x) = \left(\frac{x}{x_0}\right)^{-\alpha}. \tag{16}$$

Taking logarithms yields

$$\log(F(x)) = \alpha \log(x_0) - \alpha \log(x), \tag{17}$$

which can be rewritten as

$$\log(F(x)) = c - \alpha \log(x), \tag{18}$$

with $c = \alpha \log(x_0)$ constant. This implies a linear relationship between $\log(F(x))$ and $\log(x)$: for a distribution governed by a power-law decay, we therefore expect the Zipf plot to display an approximately linear, decreasing pattern.

Figure 5, which considers all absolute negative returns, does not exhibit this behavior, suggesting either that returns do not follow a strict power-law decay or that additional effects are present. However, Figure 6, where only losses greater than 1% are included, reveals a much clearer linear decay in the tail region. This indicates that the distribution of returns may be better interpreted as a mixture: smaller observations are consistent with a thin-tailed distribution, whereas the extreme losses display heavy-tailed, Pareto-like behavior.

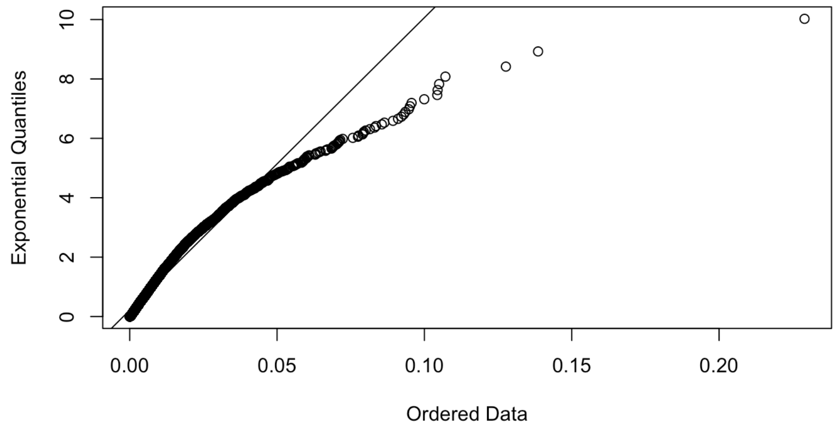

Figure 7: QQ-Plot: Exponential quantiles against empirical quantiles.

Figure 7 displays the empirical quantiles plotted against the theoretical quantiles of an exponential distribution. This comparison is particularly useful since, by definition, a distribution is considered heavy-tailed if its survival function decays more slowly than that of an exponential. The exponential law therefore serves as a natural benchmark to distinguish between light-tailed and heavy-tailed behavior.

Under this framework, the shape of the QQ-plot provides valuable insight: a convex pattern of the empirical quantiles relative to the exponential line would suggest a thin-tailed distribution, whereas concavity indicates the presence of heavy tails. In our case, the observed concavity strongly supports the hypothesis that returns are heavy-tailed.

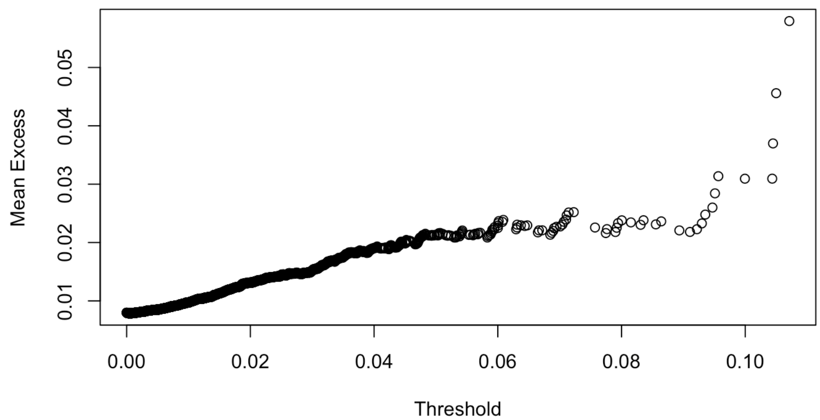

Figure 8: Mean Excess Function Plot: one can note the linear increase (way different from a decreasing behavior for a normal distribution), this suggest paretian behavior.

Figure 8 shows the Mean Excess Function (MEF) plot, which exhibits a clear linear upward trend. This behavior is consistent with the theoretical MEF of a Pareto-like distribution, as introduced in Section 3.2.4.

In contrast, a normal distribution would display a decreasing MEF, a pattern that is clearly not observed here.

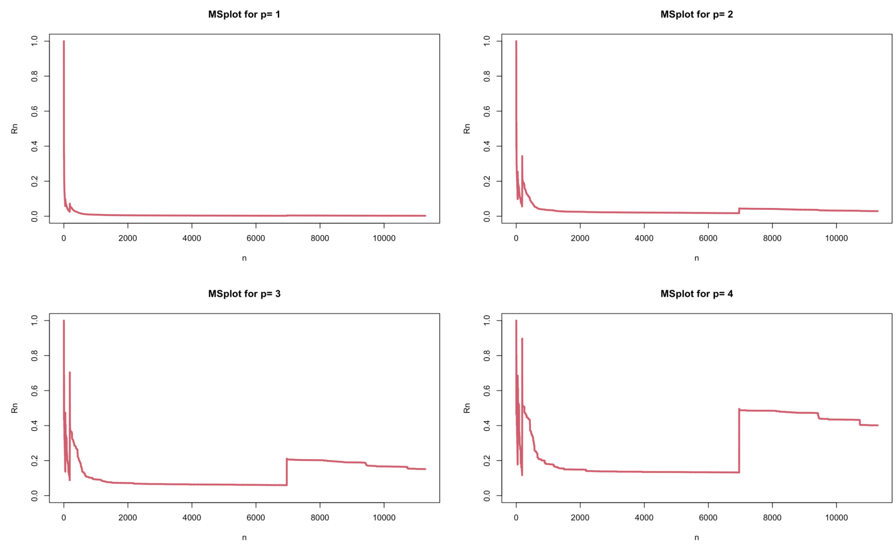

Figure 9: MS plot for p = 1, 2, 3, 4. We can safely claim that the 1^{st} and 2^{nd} moments are finite, however, the 4^{th} moment is infinite and the 3^{rd} may converge but from the plot we cannot conclude certainly.

As discussed in Section 3.2.1, if the p-th moment is finite, the ratio in the MS plot converges to zero as $n \to \infty$. In this case, the plot should display a decreasing function approaching zero, as observed in the first figure. Conversely, if the p-th moment is not finite, the ratio exhibits erratic behavior with sudden jumps and fails to converge, regardless of the number of observations considered. From a theoretical standpoint, we recall that a normal distribution possesses finite moments of all orders.

Figure 9 shows that the 1^{st} and 2^{nd} moments are finite, while the 4^{th} moment is infinite. The 3^{rd} moment may converge, but the plot does not provide sufficient evidence to conclude definitively. These findings are crucial when analyzing financial returns: in such cases, statistical inference based on the third moment or higher is unreliable and potentially misleading.

4 Conclusions from the Empirical Analysis

The empirical investigation carried out through different statistical tools consistently rejects the hypothesis of normality in financial returns. The Zipf plots have shown that while small losses resemble a thin-tailed distribution, extreme losses follow a linear decay consistent with power-law behavior. The QQ-plot against the exponential distribution has further confirmed the presence of concavity, indicating heavy-tailed dynamics. Similarly, the Mean Excess Function displayed a linear upward trend, in line with the theoretical behavior of Pareto-like distributions. Finally, the MS plot provided additional evidence, suggesting that only the first and second moments are finite, while higher-order moments may not exist.

Taken together, these results strongly support the presence of heavy tails in financial return distributions. This implies that extreme events are far more frequent than predicted under the Gaussian framework, and that models relying on higher-order moments (such as skewness or kurtosis under normality) are unreliable. Such findings motivate the exploration of alternative pricing models under fat-tailed assumptions, which are better suited to capture the true statistical properties of markets.