Demonstration of fat-tailed non-Gaussian shape of price difference logarithm's distribution

1 Introduction

In financial markets, it is often assumed that asset returns' logarithm,

$$r_t = ln(\frac{P_t}{P_{t-1}}) \tag{1}$$

where r_t is the logarithmic return at a certain time t, P_t is the closing price at time t, and P_{t-1} is the closing price at the previous time step (day); are normally distributed, as in a geometric Brownian motion.

However, this assumption has been increasingly challenged, as real market data frequently exhibits **fat tails**, indicating a higher probability of extreme events than predicted by a Gaussian model, closely linked to volatility. In this section, there is a will to demonstrate the natural emergence of non-Gaussian fat-tailed distributions in the logarithm of daily price differences of several relevant indexes, futures and stocks: the S&P 500 *(ˆGSPC))*, Gold futures *(GC=F)*, Crude Oil futures *(CL=F)*, and JPMorgan Chase & Co. *(JPM)*.

Four methods are used to illustrate this phenomenon, showing the discrepancies between the idealized Gaussian model and the behavior of the real financial data. The Gaussian probability distribution is characterized by:

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\sigma^2\pi}} \exp -\frac{(x-\mu)^2}{2\sigma^2} \tag{2}$$

where μ is the mean and σ^2 is the variance.

2 Method 1: Logarithmic Returns Histogram vs. Gaussian Fit

By plotting the logarithmic returns histogram and comparing it with a Gaussian distribution (with the same mean and standard deviation), one can see some irregularities in the data with respect to the idealised model.

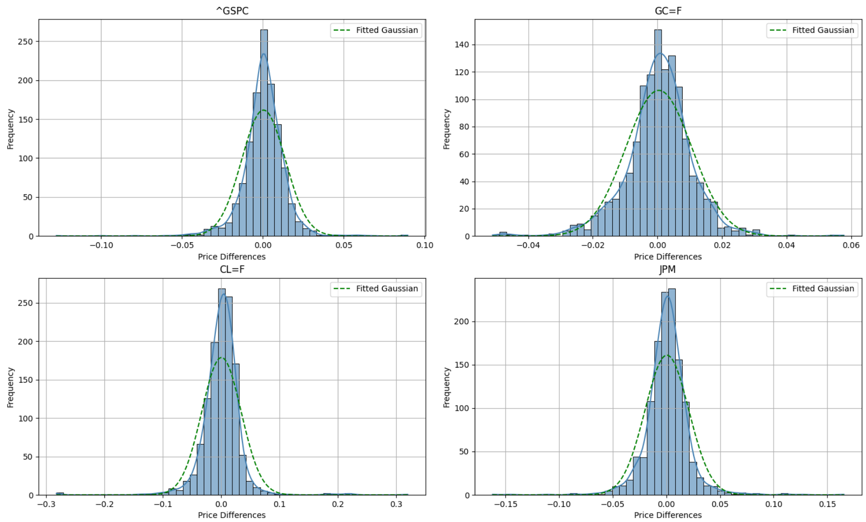

Figure 1: Histogram of logarithmic returns computed from historical price data retrieved via Yahoo Finance, covering the period from 2020 to 2025 for the four selected financial assets. The blue line is an estimated probability density, and the green one is the Gaussian probability distribution based on the data

This empirical mismatch indicates the presence of large jumps or abrupt price changes that the normal distribution severely underestimates, represented in the heavy tails behaviour.

3 Method 2: Q-Q Plot Against the Normal Distribution

Quantile-Quantile (Q-Q) plots are a graphical method for comparison between two probability distributions by plotting their quantiles [1] against each other, in this case the empirical returns against those of a standard normal distribution.

If the data was to behave as the normal distribution, the points in the Q-Q plot would lie on the straight 45° line. However, there is a significant deviation at both ends of the plot indicating the tendency to a fatter tails than the Gaussian benchmark.

This visual cue provides strong evidence of non-normality and heavy-tailed behaviour in the data.

[1] Cut points dividing the range of a probability distribution into continuous intervals of equal probability.

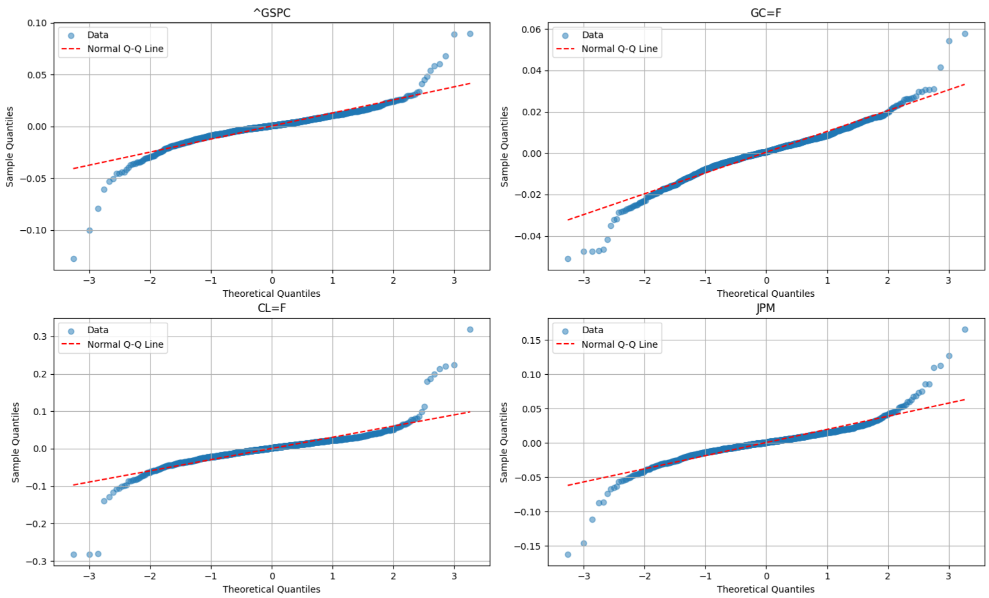

Figure 2: Quantile-quantile (Q-Q) plot of empirical logarithmic returns against the theoretical quantiles of a normal distribution. Deviations from the reference line in red indicate the presence of heavy tails and non-Gaussian behaviour in the selected financial assets.

4 Method 3: Kurtosis Analysis

Kurtosis is a measure of the combined weight of a distribution's tails relative to the mean. A standard normal distribution has a kurtosis of 3 and is referred to as *mesokurtic*. The closer the kurtosis is to 3, the more does the dataset approach a Gaussian behaviour. Distributions with kurtosis greater than 3 exhibit heavy tails and are classified as *leptokurtic*. In contrast, distributions with kurtosis less than 3 are labelled *platykurtic* and are characterized by light tails, therefore a reduced probability of extremes values i.e. volatility.

The analysis on the four assets,

$$\hat{}GSPC\text{: } 14.15 \quad GC{=}F\text{: } 3.59 \quad CL{=}F\text{: } 23.73 \quad JPM\text{: } 12.51$$

shows that all four tend to exceed the 3 threshold kurtosis, often well above, confirming the fat-tailed nature of return logarithmic distributions, quantifying the deviation from normality, with the exception of Gold futures, which is closer to it.

5 Method 4: Fitting a Cauchy Distribution

The Cauchy distribution is a distribution with undefined mean and variance, heavy-tailed, and is characterized by the probability density function:

$$f(x; x_0, \gamma) = \frac{1}{\pi\gamma\left[1 + \left(\frac{x - x_0}{\gamma}\right)^2\right]} \tag{3}$$

where x_0 is the median and γ is the median absolute deviation (related to the width of the distribution).

By plotting both the Cauchy and the Gaussian distributions, one can compare the suitability of the two models.

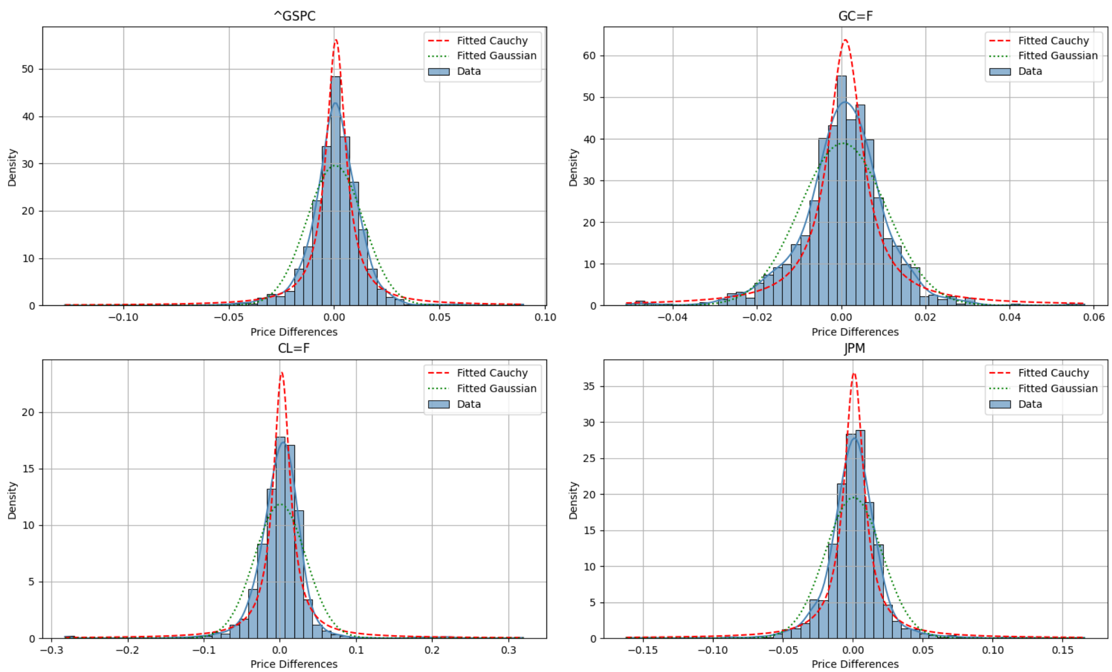

Figure 3: Histogram of log returns with fitted Cauchy distribution for the four selected stocks in red, an estimation of the probability density of the data in blue, and the Gaussian distribution in green.

Visually and numerically, the Cauchy fit aligns much better with the observed distribution of returns than the Gaussian model, providing a better appreciation of both the fat tails behaviour of the data and the peak representing the mean value.

Conclusion

Across the four methods, it is observed that the logarithm of daily price differences in financial data consistently exhibit fat-tailed, non-Gaussian behaviour. This has significant implications for risk management, option pricing, and financial modelling in general. Assuming Gaussian behaviour underestimates volatility. This may be better described by heavy-tailed distributions, such as the Cauchy or other Lévy stable distributions, rather than the conventional Gaussian assumption.