Calm surface, fat tails

Taleb's fat-tail diagnostics applied to the S&P 500, 1950 to October 2026

Matheus Hoffmann ORCID

November 2026 · Preprint v0 (draft, not peer reviewed) · rev. f5fd12d

Abstract. On 30 October 2026 the S&P 500 had a daily volatility forecast of 0.76%, below its 1950 to 2026 average of 1.00%. A Gaussian model turns that number into a 99.99% one-day loss of 2.8%. We apply the diagnostics of Taleb's *Statistical Consequences of Fat Tails* to 19,330 daily returns and find a power-law left tail with exponent near 3 (3.75 after removing volatility clustering). One day, 19 October 1987, supplies 52% of the sample's fourth moment, so kurtosis is not a stable number. The max-to-sum test shows a finite mean and variance but no convergence of the third and fourth moments. Taleb's κ is 0.24 for daily returns. Sums of 30 days keep a kurtosis near 8, so time does not make the returns Gaussian. At today's volatility, a power-law tail gives a 99.99% one-day loss of 7.2% and an expected shortfall of 9.8%: 2.6 and 3.3 times the Gaussian numbers. The gap is small at 99% and grows with the confidence level. This is an educational study, not investment advice.

Keywords: fat tails; tail exponent; kappa metric; value at risk; S&P 500.

1 A calm Friday

On Friday 30 October 2026 the S&P 500 closed at 7,580. Its volatility forecast for the next trading day was 0.76%. That is lower than the average since 1950, which is 1.00%. A risk desk that trusts the Gaussian distribution reads this as good news. With $\sigma = 0.76\%$, a 1-in-100 day costs 1.8%, a 1-in-1,000 day costs 2.3%, and a 1-in-10,000 day costs 2.8%. On that reading, a 3% daily loss happens about once in a century.

The same index fell 22.9% on Monday 19 October 1987. Measured against the volatility forecast it had that morning, the fall was a 12-sigma event. A Gaussian assigns a 12-sigma loss a probability near 10^{-33}. A Gaussian with the full-sample volatility puts the return period of that day at 10^{114} years. The universe is about 1.4×10^{10} years old.

Taleb's *Statistical Consequences of Fat Tails* (Taleb 2020) argues that this is not bad luck but the wrong model. In a fat-tailed world, rare events dominate the statistics. The sample mean, variance and kurtosis converge slowly or not at all. Risk numbers fitted to the body of the distribution are blind to the tail. The book gives a set of diagnostics that test these claims on data. This paper applies them to the S&P 500 and asks one practical question:

> **How wrong is a Gaussian risk number for the S&P 500 today?**

We follow the diagnostics in the order that a sceptical reader would ask for them. Section 2 describes the data. Section 3 shows that kurtosis, the usual "fatness" number, is decided by one day. Section 4 asks which moments exist at all. Section 5 measures the tail exponent. Section 6 asks whether summing days makes the problem go away. Section 7 turns the tail into risk numbers for today and compares three indexes. Section 8 lists what the study cannot say.

2 Data and method

Prices. We use daily closes of the S&P 500 from 3 January 1950 to 30 October 2026, which give 19,330 daily log returns $r_t = \log P_t - \log P_{t-1}$. The data joins two public copies of Yahoo Finance's `^GSPC` series held on GitHub: one for 1950 to 2014 and one for 2015 to 2026. Both copies agree with the FRED series `SP500` to eight significant digits on every day where they overlap (1,001 and 413 days). For comparison we also use the Nasdaq 100 and the Nikkei 225 from January 2015 to 30 October 2026 (2,974 and 2,888 returns). The data and the script are in the `fat-tails/` directory of this repository.

Volatility. Daily volatility clusters: calm weeks follow calm weeks and wild weeks follow wild weeks. We estimate the volatility forecast for day t with an exponentially weighted moving average,

$$\sigma_t^2 = \lambda\,\sigma_{t-1}^2 + (1-\lambda)\,r_{t-1}^2, \qquad \lambda = 0.94,$$

the RiskMetrics convention. The *devolatilised* return $z_t = r_t/\sigma_t$ removes most of the clustering. The tail estimators in Section 5 assume independent observations, so they are more reliable on z_t than on r_t. We drop the first 250 days of z_t as a burn-in.

Diagnostics. We compute the metrics of (Taleb 2020) that apply to a single return series:

- the kurtosis and the share of the fourth moment that comes from the largest observations;
- the max-to-sum (MS) ratio $R_n^p = \max_{i\le n} |X_i|^p / \sum_{i\le n} |X_i|^p$ for $p = 1, \dots, 4$;
- the Hill estimator of the tail exponent α (Hill 1975) across thresholds, with the Zipf (log-log survival) and mean-excess plots;
- Taleb's κ metric, which measures how fast sums approach the Gaussian;
- value at risk (VaR) and expected shortfall (ES) under a Gaussian and under a power-law tail fitted with the Weissman extrapolation (Weissman 1978).

The book also discusses the Lindy effect. Lindy concerns the life expectancy of things such as firms and technologies, not daily returns, so we leave it out.

3 Kurtosis is decided by one day

The usual first check of fat tails is the kurtosis, $E[(X-\mu)^4]/\sigma^4$. A Gaussian has kurtosis 3. The S&P 500 daily returns since 1950 have kurtosis 28.2. That looks like a clear answer, but Taleb's point is that the number itself is not reliable when tails are fat.

The fourth power gives the largest days an enormous weight. The single largest day, 19 October 1987, contributes 52% of the sum $\sum r_t^4$. The ten largest days contribute 70%. Remove the one day and the kurtosis halves, from 28.2 to 14.4. Remove ten days and it falls to 9.8. A statistic that one observation out of 19,330 can halve does not describe the distribution. It describes the worst day in the sample.

Figure 1 shows the same instability in time. Year by year, kurtosis ranges from 2.6 (a Gaussian-looking year) to 57 (1987). The median year has 4.4. The right panel follows the kurtosis of the whole history as data arrives. It drifts down between shocks, as an estimator should when the true value is finite. Then a single day lifts it by a factor of ten. After 39 more years of data it has still not come back to its pre-1987 level. Under a Gaussian the curve would settle near 3 within a few hundred days.

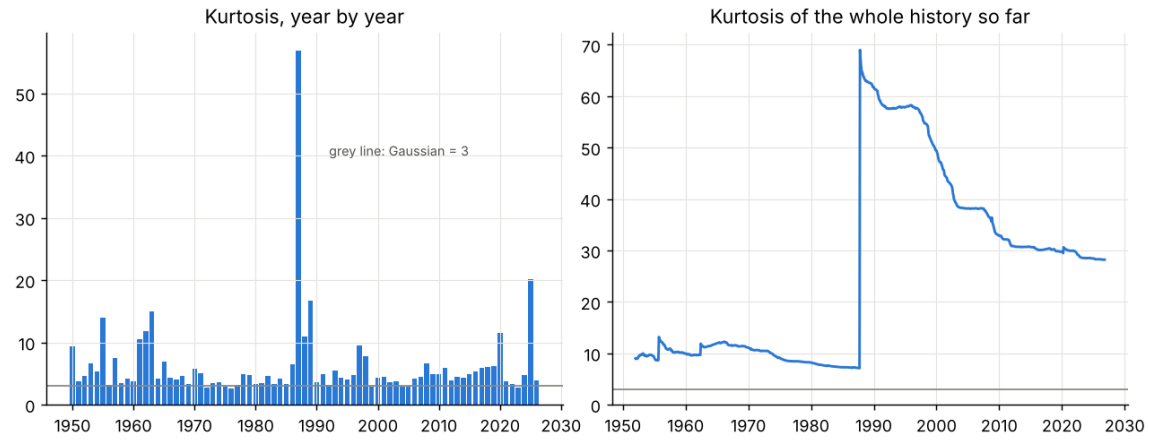


Figure 1: Left: kurtosis of each calendar year; the grey line is the Gaussian value 3. Right: kurtosis of all returns from 1950 up to each date. The jump in October 1987 never washes out.

The lesson is the first of the book: in a fat-tailed sample, a high-order moment tells you more about the largest observation than about the population. We need tools that look at the tail directly.

4 Which moments exist?

If kurtosis is unstable, perhaps the fourth moment does not exist. The MS plot tests this without assuming a model. For independent draws, $R_n^p \to 0$ as n grows if and only if $E|X|^p < \infty$ (Taleb 2020). When the moment is infinite, new record observations keep taking a large share of the sum, and the ratio keeps jumping back up.

Figure 2 gives a graded answer. For $p = 1$ the ratio is 0.002 at the end of the sample: the mean is finite. For $p = 2$ it is 0.027 and falling, with one visible jump in 1987: the variance is finite but its estimate still moves with single days. For $p = 3$ the ratio is 0.19, and for $p = 4$ it is 0.52. These do not converge. In the $p = 4$ panel, after 70 years and more than 19,000 observations, one day still holds half of the sum. The Gaussian sample reaches zero within a few hundred days in every panel.

This pattern is what a power law with tail exponent α between 2 and 4 predicts: moments of order below α are finite, and moments at or above α are infinite or so heavy that a sample cannot estimate them. Section 5 measures α directly.

5 How fat is the tail?

A power-law tail has the form $P(L > x) \approx Cx^{-\alpha}$ for large losses L. The exponent α is the one number that matters: the smaller it is, the fatter the tail. Moments of order $p \geq \alpha$ are infinite. A Gaussian has no finite α; its tail falls faster than any power.

The Hill estimator (Hill 1975) takes the k largest losses $X_{(1)} \geq \ldots \geq X_{(k+1)}$ and returns

$$\hat{\alpha}_k = \left(\frac{1}{k} \sum_{i=1}^{k} \log X_{(i)} - \log X_{(k+1)} \right)^{-1}.$$

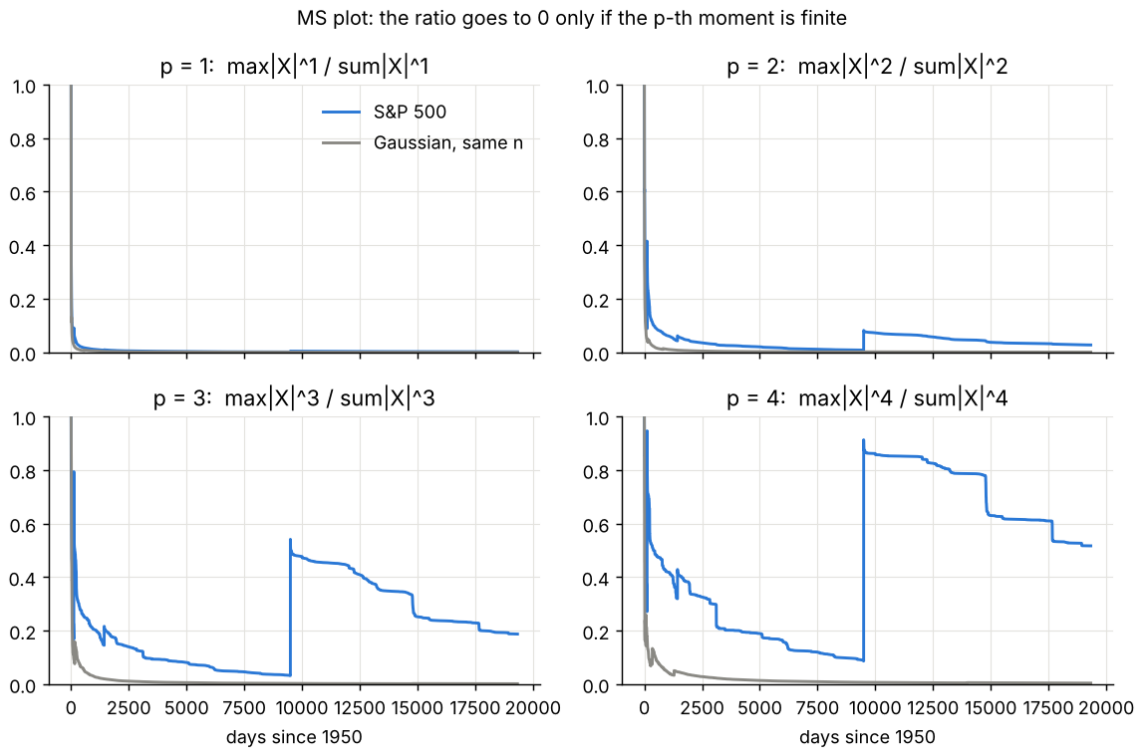


Figure 2: MS plot of S&P 500 daily returns for $p = 1$ to 4. Blue: the index. Grey: a Gaussian sample of the same length. The ratio tends to 0 only if the p-th moment is finite.

Its weakness is the choice of k. Too few points give noise, too many reach into the body of the distribution, where the power law does not hold. Taleb's advice is to plot $\hat{\alpha}_k$ against k and look for a stable region.

Figure 3 shows both versions. On raw losses, $\hat{\alpha}$ is 3.1 at the top 1% of days, 2.8 at 2.5% and 2.7 at 5%. On devolatilised losses it is higher, 4.0, 3.7 and 3.2, because dividing by σ_t removes the extra fatness that volatility clustering creates. Between the 1% and 5% thresholds the devolatilised estimate stays between 3.2 and 4.1. We use $\hat{\alpha} = 3.75$ at the 2.5% threshold ($k = 477$, standard error $\alpha/\sqrt{k} = 0.17$). The right tail (gains) is thinner, with $\hat{\alpha} \approx 5$: the S&P 500 falls faster than it rises.

Both estimates agree with the MS plot. With α near 3, the variance exists, the third moment is borderline, and the fourth is infinite.

Figure 4 shows the same fact in two pictures that need no estimator. In the Zipf plot, the index follows the Gaussian up to about 1.5 sigma. Beyond that point the data keeps a straight line on log-log axes, out past 12 sigma, while the Gaussian curve drops off the chart before 5 sigma. The mean-excess plot is the clearest test for a non-specialist. For a Gaussian, the further past a threshold you already are, the less extra loss you should expect: the grey curve falls. For the S&P 500 the curve turns up after 2 sigma. Once a day is bad, the expected extra loss *grows* with how bad it already is. That is the defining behaviour of a fat tail, and it is the behaviour that a risk model must capture.

6 Does time cure fat tails?

A common defence of Gaussian models is the central limit theorem. Daily returns are fat-tailed, the argument goes, but monthly returns are sums of about 21 daily ones, so they

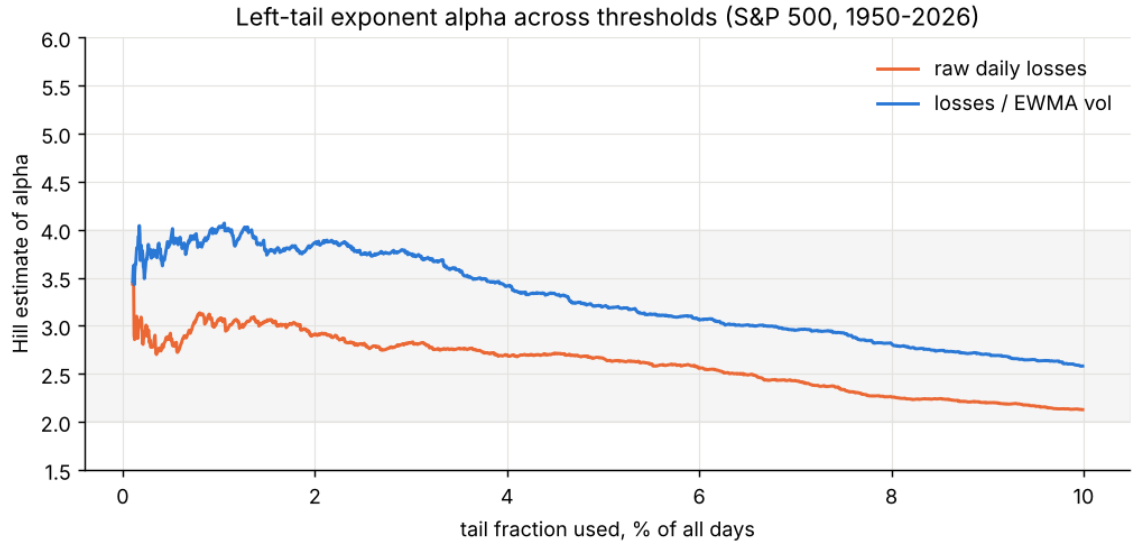


Figure 3: Hill estimate of the left-tail exponent as a function of the share of days used. Orange: raw daily losses. Blue: losses divided by the EWMA volatility forecast. The grey band marks $2 \leq \alpha \leq 4$.

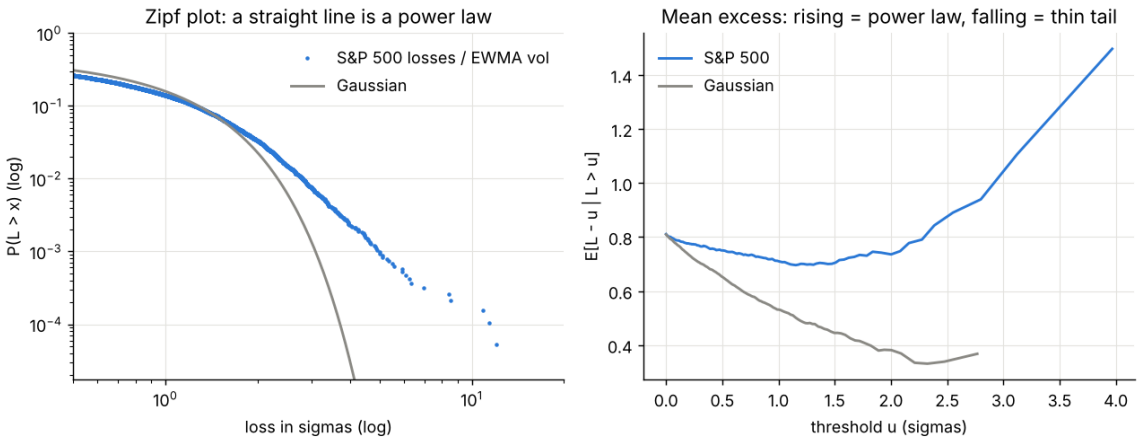


Figure 4: Left: Zipf plot of devolatilised losses, log of the survival probability against log of the loss. A power law is a straight line; the Gaussian (grey) bends down and vanishes near 4 sigma. Right: mean excess $E[L - u \mid L > u]$. It falls for a thin tail and rises for a power law.

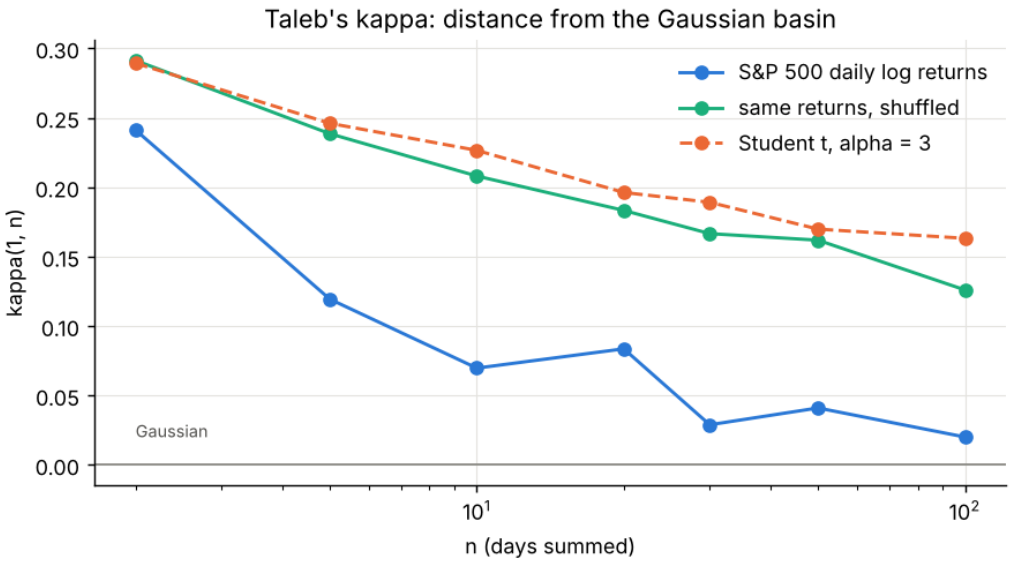


Figure 5: $\kappa(1, n)$ for sums of n days. Blue: S&P 500. Green: the same daily returns in random order, which keeps their distribution but removes volatility clustering. Orange: a Student t with $\alpha = 3$.

should be close to Gaussian. Taleb's κ metric tests how fast this happens (Taleb 2020). Let $M(n)$ be the mean absolute deviation of a sum of n returns. Then

$$\kappa(n_0, n) = 2 - \frac{\log n - \log n_0}{\log\left(M(n)/M(n_0)\right)}.$$

For a Gaussian, $M(n)$ grows like $\sqrt{n}$ and $\kappa = 0$. For a Cauchy distribution ($\alpha = 1$), $M(n)$ grows like n, there is no diversification at all, and $\kappa = 1$. Values in between measure how much more data a fat-tailed process needs to reach the precision that a Gaussian gives.

For two-day sums the S&P 500 has $\kappa(1, 2) = 0.24$, far outside the Gaussian 90% band of $[-0.05, 0.05]$ and close to a Student t with $\alpha = 3$ (0.29). Figure 5 then shows a surprise. For longer sums the index's κ falls towards zero: 0.07 at 10 days, 0.03 at 30 and 0.02 at 100. The shuffled series stays high, between 0.12 and 0.21 at 30 days, as a fat-tailed independent series should. Read naively, the figure says that monthly S&P 500 returns are nearly Gaussian.

They are not. Table 1 looks at the shape of the sums directly.

Table 1: Shape of n-day sums of S&P 500 returns. A Gaussian has kurtosis 3 and MAD/SD = 0.798. "Shuffled" uses the same daily returns in random order.

days summed	kurtosis	kurtosis, shuffled	MAD/SD	MAD/SD, shuffled
1	28.2	28.2	0.668	0.668
2	10.2	17.2	0.711	0.706
10	7.5	6.1	0.736	0.746
30	7.8	4.9	0.722	0.771
100	6.0	2.8	0.726	0.804

The shuffled sums behave as the central limit theorem says: kurtosis falls to 2.8 at 100 days, and the ratio of mean absolute deviation to standard deviation climbs to the Gaussian 0.798.

The real sums stop improving. Thirty-day sums keep a kurtosis of 7.8, and the MAD/SD ratio stays near 0.72 at every horizon. Volatility clustering explains this. A bad day tends to arrive with other bad days, so a month that contains one crash day often contains several, and the monthly return inherits the fat tail.

This also explains the low κ. κ compares how $M(n)$ grows with n. If the shape of the sum stops changing after a few days, $M(n)$ grows like $\sqrt{n}$ from then on, and $\kappa(1, n)$ drifts towards zero, although the sums are far from Gaussian. A low κ at long horizons here means "no further convergence", not "already Gaussian". This is the *preasymptotics* of the book's subtitle: the central limit theorem holds in the limit, and the limit is out of reach.

The practical consequence is direct. A common practice scales a daily Gaussian VaR by $\sqrt{t}$ to get a t-day VaR. That is wrong twice: the daily tail is not Gaussian, and summing days does not make it Gaussian.

7 Risk on 2 November 2026

We now return to the question of Section 1. Write tomorrow's loss as $L = \sigma \cdot Z$, where $\sigma = 0.76\%$ is today's volatility forecast and Z is a standardised loss. The Gaussian model takes $Z \sim N(0, 1)$. The fat-tailed model takes the empirical distribution of the devolatilised losses z_t and extends its tail with a power law of exponent $\hat{\alpha} = 3.75$. The Weissman estimator (Weissman 1978) gives the quantile at exceedance probability p as

$$\hat{q}_p = X_{(k+1)} \left(\frac{k}{np} \right)^{1/\hat{\alpha}},$$

and for a power law the expected shortfall is $\mathrm{ES}_p = \hat{q}_p \cdot \alpha/(\alpha - 1)$.

Table 2: One-day S&P 500 loss quantiles (VaR) and expected shortfall (ES) at the EWMA volatility forecast for 2 November 2026 ($\sigma = 0.76\%$).

one-day level	Gaussian VaR	power-law VaR	ratio	Gaussian ES	power-law ES	ratio
99%	1.76%	2.10%	1.2	2.01%	2.86%	1.4
99.9%	2.33%	3.87%	1.7	2.54%	5.28%	2.1
99.99%	2.81%	7.16%	2.6	2.99%	9.77%	3.3

Table 2 contains the paper's main result. At the 99% level the two models nearly agree; the Gaussian is 20% low. This is why Gaussian VaR survives in practice: it is tested at levels where it works. At 99.9%, a level that a daily risk report reaches about once in four years, the Gaussian is 40% too low. At 99.99% the power law gives a loss of 7.2%, 2.6 times the Gaussian 2.8%, and an expected loss beyond that point of 9.8%, 3.3 times the Gaussian. The error grows with the level because the two tails diverge: the Gaussian falls like $e^{-x^2/2}$, the power law like $x^{-3.75}$.

The history agrees with the power law. Since 1951, the S&P 500 had 19 days with a loss beyond 5 times the volatility forecast for that day. A Gaussian expects 0.005 such days in 75 years, one per 14,000 years. It had 9 days beyond 6 sigma, where a Gaussian expects 2×10^{-5}. On raw returns, the power-law fit gives the 1987 crash a return period of about 160 years: rare, but inside the span of recorded markets. The Gaussian gives 10^{114} years.

A calm reading today does not mean a thin tail. The fat-tailed model already uses today's low σ; the extra risk comes from Z, not from σ. A high σ does not explain the worst days either. On the morning of 19 October 1987 the forecast was 1.9%, high (the top 3% of days since 1950) but far from the 22.9% that followed.

7.1 Three indexes since 2015

Table 3: The same diagnostics on January 2015 to October 2026 for three indexes. "Top day" is the largest day's share of $\sum r^4$; $\hat{\alpha}$ is the Hill estimate on devolatilised losses at the 2.5% threshold; σ is the EWMA forecast for 2 November 2026; "Gauss" and "power" are the one-day 99.9% VaR under each model.

index	kurt.	top day	$\hat{\alpha}$	σ	Gauss	power
S&P 500	18.9	31%	2.8 ± 0.3	0.76%	2.33%	5.41%
Nasdaq 100	10.6	24%	3.2 ± 0.4	1.47%	4.55%	9.16%
Nikkei 225	10.7	28%	3.9 ± 0.5	1.68%	5.20%	8.73%

Table 3 repeats the analysis on a common, shorter window. Three points stand out. First, the S&P 500's tail is *fatter* in the last eleven years than over the full history: $\hat{\alpha}$ is 2.8 instead of 3.75. March 2020 (four of the six worst days, including −12.8% on 16 March) and 4 April 2025 (−6.2%) drive this. Fitted on this window, the 99.9% one-day loss at today's volatility is 5.4%, 2.3 times the Gaussian number. Second, the Nasdaq 100 and the Nikkei 225 are about twice as volatile as the S&P 500 today, so their Gaussian and power-law numbers are both higher. Third, all three tails are fat: no index has $\hat{\alpha}$ above 4, and on all three the largest single day carries a quarter or more of the fourth moment.

The window matters. Eleven years hold only about 70 days in the top 2.5%, so these $\hat{\alpha}$ values have standard errors of 0.3 to 0.5. This is itself one of Taleb's points: the tail is the part of the distribution about which a sample says least.

8 What this study cannot say

The diagnostics agree with each other. Kurtosis is unstable (Section 3), the third and fourth moments do not converge (Section 4), the left tail is a power law with α near 3 (Section 5), and summing days does not make returns Gaussian (Section 6). At today's volatility, the Gaussian understates the 1-in-10,000-day loss by a factor of 2.6 and the expected loss beyond it by 3.3 (Section 7). The study has five limits.

1. **The threshold of Hill.** $\hat{\alpha}$ depends on k. We report a range (1% to 5% of days) and a standard error, but a dozen extreme days drive the far tail. Fitting at the 1% or the 5% threshold instead of 2.5% gives $\hat{\alpha} = 4.0$ or 3.2, and moves the 99.99% VaR from 7.2% to 6.7% or 9.0%. Every choice stays at least 2.3 times the Gaussian 2.8%.
2. **The volatility model.** EWMA with $\lambda = 0.94$ is simple. It reacts to a crash only after the first bad day, and it does not model leverage (volatility rises more after falls than after gains). A GARCH-type model would change the devolatilised tail slightly.
3. **One-day horizon.** Table 2 covers one day. Section 6 shows why $\sqrt{t}$ scaling of these numbers to longer horizons is not valid. Longer horizons need their own tail fit.
4. **Data.** The prices come from two GitHub copies of Yahoo Finance. They agree with FRED on the days we could check, but we did not check every day before 2015. Closing prices miss intraday losses.

5. **Not a forecast.** The numbers describe the size of a rare loss, not its timing. A power-law tail says that a 7% day is a 1-in-10,000-day event at today's volatility. It does not say when.

The central claim does not depend on these choices. Every Hill estimate in this paper, for every index, threshold and window, lies between 2.4 and 4.4. In that range a Gaussian understates the far tail by a factor of two or more.

Statements

Not investment advice. This is an educational study of statistical methods. It does not recommend any trade or position.

Use of AI. Claude (Anthropic), an AI assistant, wrote the analysis code, ran it, and drafted this text at the author's request. The author is responsible for the content.

Reproducibility. `python3 fat-tails/python/fat_tails.py` regenerates every number, table and figure from the CSV files in `fat-tails/data/`.

References

Hill, Bruce M. 1975. "A Simple General Approach to Inference about the Tail of a Distribution." *The Annals of Statistics* 3 (5): 1163–74. https://doi.org/10.1214/aos/1176343247.

Taleb, Nassim Nicholas. 2020. "Statistical Consequences of Fat Tails: Real World Preasymptotics, Epistemology, and Applications." arXiv:2001.10488, v4 (Third Revised Edition, 2025). https://doi.org/10.48550/arXiv.2001.10488.

Weissman, Ishay. 1978. "Estimation of Parameters and Large Quantiles Based on the k Largest Observations." *Journal of the American Statistical Association* 73: 812–15. https://doi.org/10.1080/01621459.1978.10480051.