Quantitative Trading Module

Mathematical Structure and Limitations of Financial Time Series

Daniel Bloch

VinUnivesity & Paris Sorbonne

21st January 2026

Outline

- **Limitations of the First Moment**
  - Instability under Non-stationarity and Regimes
  - Low Signal-to-Noise Ratios and Jump Sensitivity
- **The Failure of Variance in Risk Assessment**
  - Violation of Homoscedasticity and GARCH limits
  - Symmetric Penalisation vs. Investor Preferences
- **Inadequacy of Linear Correlation**
  - Linearity Assumptions and False Negatives
  - Invisible Tail Dependence and Systemic Risk
- **Modern Alternatives and Design Principles**
  - Quantiles, Expected Shortfall, and Copulas
  - State-conditional Modelling and Path-sensitivity

Course Reference:
Futuretesting Quantitative Strategies
`http://ssrn.com/abstract=4647103`

Mathematical Properties of Asset Prices: Why Classical Statistics Fail in Markets

Defining the Mean: Population vs. Sample

The most fundamental statistical measure of a process is its average value, which can be viewed from two distinct perspectives:

- **Population Mean (First Moment):** For a stochastic process (X_t), the theoretical mean is defined as the expected value at time t:

$$\mu_t = E[X_t] = \int_{-\infty}^{\infty} x \, dF_t(x)$$

  where F_t is the marginal distribution.

- **Sample Mean (Empirical Estimate):** In practice, we estimate this using n historical observations:

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} X_{t_i}$$

Defining the Mean: Population vs. Sample

- **The Efficiency Gap:** While the sample mean is an unbiased estimator for stationary Gaussian processes, it becomes highly unreliable in the presence of the complex features previously discussed (heavy tails and long memory).

We now examine why these standard measures fail in financial applications.

Why the mean is ill-suited for financial time series

The first moment (mean) of financial time series suffers from:

Key implications include

- Instability under non-stationarity
- Low signal-to-noise ratios
- Susceptibility to jumps and heavy tails
- Loss of information from regime mixing
- Misalignment with economic or trading goals

Therefore, relying on the mean as a primary evaluative or predictive statistic in financial contexts is misguided. The mean fails to describe the true value or directional intent in a non-smooth, volatile, path-sensitive environment.

Non-stationarity and regime dependence

Let X_t be a financial return series. The mean is defined as $\mu = E[X_t]$. In financial markets, the mean return is not stable over time:

$$E[X_t] \neq E[X_{t+\tau}]$$

The mean return can shift due to:

- Changes in macroeconomic regimes
- Monetary policy adjustments
- Market sentiment transitions
- Structural breaks (e.g., crises, technological shocks)

Thus, estimating a global or historical mean is meaningless for forward-looking inference in markets with time-varying dynamics.

Low signal-to-noise ratio

Daily or intraday returns typically have a very low mean relative to their variance. For instance, the expected daily return on an index might be 0.04%, while the standard deviation is 1% giving a signal-to-noise ratio:

$$\frac{\mu}{\sigma} \approx 0.04$$

This means that the mean is easily overwhelmed by volatility, and estimating it requires massive sample sizes to achieve statistical significance. In finite samples, particularly under noise and jumps, the mean is unstable and statistically unreliable.

The Pareto Law: Modelling Extreme Events

The **Pareto distribution** is the canonical model for power-law behaviour, representing a world where a small number of events account for the majority of the total impact.

- **Probability Density:** For a threshold $x_m > 0$ and shape parameter $\alpha > 0$, the density is:

$$f(x) = \frac{\alpha x_m^\alpha}{x^{\alpha+1}}, \quad x \geq x_m$$

- **Tail Properties:** The 'heavy-headedness' of the tail is governed by α:
  - If $\alpha \leq 1$: The theoretical mean $E[X]$ is infinite.
  - If $\alpha \leq 2$: The theoretical variance is infinite.
- **Financial Context:** In these regimes, the empirical average does not 'settle down' quickly; it remains volatile and dominated by the largest observations in the sample.

This mathematical fragility explains why the sample mean is so easily distorted in financial markets.

Sensitivity to rare events and jumps

In heavy-tailed return distributions, the sample mean becomes sensitive to extreme observations:

The mean return can shift due to:

- $\mu = E[X_t]$ may not exist or may converge slowly if $X_t \sim Pareto(\alpha \leq 2)$
- Just one or two large jumps (e.g., earnings surprise, war announcement) can distort the mean disproportionately, making it an unreliable summary statistic.

Lack of economic meaning

In trading and investment, the central tendency often masks the critical drivers of performance:

What matters more than the average return:

- **Directionality:** The realization of a specific trend or structural move.
- **Timing:** The precision of a breakout or entry/exit relative to volatility.
- **Tail Events:** The magnitude of rare, extreme shocks (jumps).

A strategy that captures a few large moves may have a highly volatile return profile with near-zero mean, yet be economically useful. Conversely, a strategy that delivers a small, smooth mean return may fail catastrophically in a crisis. Hence, the mean fails to align with utility or economic objectives.

Structural Breaks and Regime Shifts

Financial markets are rarely governed by a single, constant mechanism. Instead, they often exhibit **regime shifts**, abrupt changes in the underlying stochastic process.

- **Definition:** A regime shift occurs when the parameters of the data-generating process (e.g., mean, volatility, or correlation) change significantly at specific points in time.
- **Economic Drivers:** These transitions are typically triggered by structural breaks, such as policy changes, geopolitical crises, or shifts in central bank mandates.
- **The Hidden State:** Mathematically, we often model these shifts using a latent (unobserved) state variable S_t, which determines the 'mode' the market is currently in.

We now observe how aggregating data across these different states can lead to a total loss of signal.

Aggregation across regimes leads to misleading inference

Suppose returns are regime-switching:

$$X_t \sim \begin{cases} \mu_1 + \epsilon_t \text{ if } S_t = 1 \\ \mu_2 + \epsilon_t \text{ if } S_t = 2 \end{cases}$$

where ϵ is white noise and S_t is a latent Markov state variable. Then the global mean is:

$$E[X_t] = p_1\mu_1 + p_2\mu_2$$

with probabilities p_1 and $p_2 = 1 - p_1$. This average obscures the structure of the underlying regimes. For example, if $\mu_1 > 0$ and $\mu_2 < 0$, then the global mean may be near zero, giving the illusion of no predictability, when in fact there may be strong signal conditional on regime.

Incompatibility with path-dependent outcomes

In many financial applications (e.g., options, risk management, drawdown control), the outcome depends on the entire path of the price process, not its mean. The sample mean does not tell us:

The mean return can shift due to:

- When large losses or gains occur
- Whether a position hits a stop-loss
- How clustered the returns are around certain events

Thus, the mean is blind to both timing and path dependence, which are critical in real-world trading and risk management.

Measuring Dispersion: Population vs. Sample Variance

Following the mean, the second moment, variance, is the standard proxy for risk, measuring the average squared deviation of the process from its mean.

- **Population Variance:** For a stochastic process (X_t), the theoretical variance is defined as:

  $$\sigma_t^2 = Var(X_t) = E[(X_t - \mu_t)^2]$$

  This measures the 'spread' of the probability distribution at time t.
- **Sample Variance:** The most common empirical estimator used by practitioners is:

  $$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_{t_i} - \hat{\mu})^2$$

Measuring Dispersion: Population vs. Sample Variance

- **Implicit Assumptions:** Relying on s^2 assumes that the variance is finite and that the process is sufficiently stable for the squared deviations to be representative of future risk.

However, in the presence of market complexities, this measure often becomes a deceptive indicator of true risk.

Why variance fails in financial time series

Variance is theoretically misaligned with the empirical characteristics of financial time series, which are:

The mean return can shift due to:

- Non-Gaussian and heavy-tailed,
- Non-stationary with regime-switching behavior,
- Marked by volatility clustering and discontinuities,
- Path-dependent and asymmetric in risk preference.

Hence, risk evaluation and model selection in finance should rely on distribution-sensitive, path-aware, and economically meaningful alternatives, rather than on variance alone.

Variance of a process

Definition

Let $X_t \in \mathbb{R}$ denote a financial time series, such as log-returns of an asset. The variance of X_t is defined as:

$$Var(X_t) = E[(X_t - \mu)^2] \text{ where } \mu = E[X_t]$$

While this measure is tractable and widely used, particularly under the assumption of Gaussian innovations and constant volatility, it becomes problematic when applied to real-world financial data, which deviate from these assumptions in several critical ways.

Assumption of homoscedasticity is violated

Constant variance assumption

Variance implicitly assumes that the second moment is time-invariant:

$$Var(X_t) = \sigma^2, \qquad \forall t.$$

Empirical reality

Financial time series exhibit *conditional heteroskedasticity*, meaning that volatility depends on past information:

$$Var(X_t \mid \mathscr{F}_{t-1}) = \sigma_t^2,$$

where $\mathscr{F}_{t-1}$ denotes the information set available at time $t-1$.

Assumption of homoscedasticity is violated

Characteristics

- Volatility clusters over time, alternating between calm and turbulent periods.
- Models such as GARCH attempt to capture this behaviour, but still assume smooth evolution.
- Abrupt regime shifts and structural breaks routinely violate these assumptions.

As a result, variance estimated under stationarity assumptions is unstable and often misleading for forward-looking risk assessment.

Non-existence or instability of the second moment

Empirical financial returns often exhibit power-law tails:

$$P(|X_t| > x) \sim x^{-\alpha}, \qquad \alpha \in (2, 4),$$

which is inconsistent with Gaussian assumptions.

Implications for variance

- If $\alpha \leq 2$, the second moment $E[X_t^2]$ does not exist.
- Even when $\alpha > 2$, variance estimates converge slowly and are dominated by rare extreme observations.
- A small number of tail events can cause large swings in estimated variance.

Consequently, variance is neither a robust nor a stable measure of risk in heavy-tailed environments, precisely where financial risk is most economically relevant.

Symmetric penalisation of deviations

Variance penalises deviations from the mean symmetrically:

$$Var(X_t) = E[(X_t - \mu)^2], \qquad \mu = E[X_t].$$

Positive and negative deviations enter the risk measure identically due to the squaring operation.

Mismatch with investor preferences

- Large gains and large losses are treated as equally undesirable.
- In practice, investors value upside risk and seek protection primarily against downside risk.
- Variance fails to align with economic utility or risk appetite.

Remedies: Semi-variance attempts to address this asymmetry:

$$SemiVar(X_t) = E\left[\left(\min(0, X_t - \mu)\right)^2\right],$$

but the standard variance itself remains blind to the sign of deviations.

Aggregation over regimes obscures local risk structure

Regime-switching dynamics

Let the return process follow a latent regime model:

$$X_t \sim \begin{cases} F_1, & \text{if } S_t = 1, \\ F_2, & \text{if } S_t = 2, \\ \vdots \end{cases}$$

where S_t is a latent Markov state variable.

Variance decomposition

The unconditional variance can be written as:

$$Var(X_t) = \sum_k P(S_t = k)\, Var(X_t \mid S_t = k) + \text{inter-regime drift terms.}$$

Aggregation over regimes obscures local risk structure

Interpretation

- The global variance is a mixture of regime-specific risks.
- High-volatility crisis regimes and low-volatility expansion regimes are averaged together.
- This aggregation masks local risk characteristics and leads to poorly calibrated decisions.

As a result, variance-based risk measures obscure precisely the state-dependent risks that matter most in practice.

Incompatibility with jump discontinuities

Jump-diffusion dynamics

Many financial return series are well modelled by jump-diffusion or pure-jump Lévy processes:

$$dX_t = \mu \, dt + \sigma \, dW_t + J_t \, dN_t,$$

where W_t is Brownian motion, N_t is a Poisson process, and J_t denotes the jump size.

Quadratic variation

The total quadratic variation of such a process is:

$$[X]_t = \int_0^t \sigma^2 \, ds + \sum_{s \le t} (\Delta X_s)^2,$$

which explicitly includes jump contributions.

Incompatibility with jump discontinuities

Interpretation

- Variance-based models derived from continuous diffusions ignore jump terms.
- This leads to systematic underestimation of true risk.
- Jump risk is episodic but economically dominant.

Hence, variance fails to capture the full risk profile of processes with discontinuities.

Poor alignment with financial objectives

Variance as an optimisation target

Variance minimisation is often used as a proxy for risk control, yet it does not correspond to core financial objectives.

Practitioner-relevant risk measures

In practice, risk and performance are evaluated using:

- **Sortino Ratio**: focuses on downside deviation,
- **Expected Shortfall (CVaR)**: captures tail risk,
- **Maximum Drawdown**: measures path-dependent losses.

- Variance treats all fluctuations symmetrically and averages across time.
- Low variance can coexist with catastrophic tail risk.
- Periods of artificially low variance often precede crises (e.g., 2006–2007).

Why correlation fails for non-smooth financial data

Definition

The Pearson correlation coefficient between two random variables X and Y is defined as:

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y}.$$

It measures *linear dependence* under finite first and second moments.

Implicit assumptions

Pearson correlation relies on:

- Approximate joint normality or symmetry,
- A stable linear relationship between X and Y,
- Stationarity of means and variances.

Why correlation fails for non-smooth financial data

Interpretation

- Financial time series routinely violate all these assumptions.
- Heavy tails, jumps, and regime changes distort correlation estimates.
- As a result, correlation is unstable and often misleading as a dependence measure.

Linearity assumption

What correlation can capture

Pearson correlation detects relationships of the form:

$$Y = aX + b + \epsilon,$$

where $a \in \mathbb{R}$ measures linear sensitivity, b is an intercept, and ϵ is mean-zero noise.

When linearity breaks down

- The dependence between X and Y is nonlinear or state-dependent,
- The data contain jumps, heavy tails, or structural breaks,
- Responses are asymmetric across market regimes.

Linearity assumption

False negatives

Any nonlinear relationship

$$Y = f(X) + \epsilon,$$

such as $f(x) = ax^2$ or $f(x) = \sin(x)$, can produce low or even zero Pearson correlation, despite strong dependence.

Zero correlation therefore does *not* imply absence of economic or predictive structure.

Instability under regime changes

Time-varying dependence

Let the joint dynamics of (X_t, Y_t) depend on a latent state S_t:

$$Cov(X_t, Y_t \mid S_t = k) = \sigma_{XY}^{(k)}.$$

Unconditional correlation

The observed correlation aggregates across regimes:

$$Cov(X_t, Y_t) = \sum_k P(S_t = k)\,\sigma_{XY}^{(k)} \; + \; \text{cross-regime terms.}$$

Instability under regime changes

Characteristics

- Dependence structures differ sharply between calm and stressed markets.
- Correlations often spike toward one during crises.
- A single unconditional estimate conceals these dynamics.

Correlation therefore provides a backward-looking and regime-averaged view of dependence, precisely when forward-looking, state-aware measures are required.

Beyond Linear Correlation

Pearson correlation is a linear measure that often fails to capture the true nature of joint risks.

- **The Limitation of Correlation:** Correlation summarises the *average* linear relationship across the entire distribution. It cannot distinguish between 'normal' co-movements and 'extreme' simultaneous crashes.
- **Mapping to Uniform Space:** To isolate the dependency structure from the individual distributions (marginals), we transform variables X and Y using their CDFs: $U = F_X(X)$ and $V = F_Y(Y)$. Both U and V are now Uniform(0,1).
- **Defining Extreme Co-movement:** We are specifically interested in the probability of one asset falling below its u-th quantile ($V \leq u$) *given* that the other has already fallen below its u-th quantile ($U \leq u$).

This conditional probability in the limit $u \to 0$ defines the concept of Tail Dependence.

Tail dependence is invisible to correlation

Definition

Tail dependence measures the probability of extreme co-movements:

$$\lambda_L = \lim_{u \to 0} P\big(Y \leq F_Y^{-1}(u) \mid X \leq F_X^{-1}(u)\big),$$

with an analogous definition for upper tails.

Key limitation of correlation

- Pearson correlation captures average co-movement, not extremes.
- Two assets can be uncorrelated in the centre yet strongly dependent in the tails.
- Gaussian copulas imply zero tail dependence by construction.

Tail dependence is invisible to correlation

Interpretation

- Financial contagion is a tail phenomenon.
- Portfolio diversification fails precisely when tail dependence dominates.
- Correlation-based risk models therefore underestimate systemic risk.

This mismatch was a central failure mode in pre-2008 risk modelling.

Breakdown under non-stationarity

Implicit stationarity assumption

Correlation assumes stable first and second moments:

$$\mu_X, \mu_Y, \sigma_X, \sigma_Y \text{ constant over time.}$$

Empirical violations

- Structural breaks due to policy changes or crises,
- Time-varying volatility and leverage effects,
- Changing market microstructure.

Breakdown under non-stationarity

Interpretation

- Estimated correlations depend strongly on the chosen sample window.
- Rolling correlations are unstable and difficult to interpret.
- Apparent diversification benefits vanish under regime shifts.

Correlation is therefore a fragile statistic in non-stationary financial environments.

Correlation is not a measure of causality or predictability

What correlation measures

Pearson correlation quantifies contemporaneous co-movement:

$$\rho_{X,Y} \neq 0 \quad \not\Rightarrow \quad X \text{ causes } Y.$$

Common drivers

- Two assets may be correlated due to shared exposure to latent factors.
- Correlation confounds direct dependence with indirect co-movement.
- Spurious correlations arise from common trends or volatility shocks.

Correlation is not a measure of causality or predictability

Interpretation

- High correlation does not imply predictability.
- Low correlation does not rule out lead–lag or nonlinear predictive structure.
- Trading strategies require directional, conditional information, not symmetric co-movement.

Using correlation as a proxy for economic linkage or forecasting power is therefore conceptually flawed.

Summary: why mean–variance–correlation fails

Conceptual misalignment

- Mean ignores timing, tails, and path dependence.
- Variance assumes symmetry, stationarity, and finite moments.
- Correlation captures only linear, average co-movement.

Empirical reality

Financial markets are:

- Heavy-tailed and jump-driven,
- Regime-dependent and non-stationary,
- Dominated by tail events and asymmetric payoffs.

Summary: why mean–variance–correlation fails

Practical consequence

Reliance on mean–variance–correlation leads to:

- False diversification,
- Underestimated risk,
- Fragile portfolios and models.

These limitations motivate the need for alternative, distribution-aware, and path-sensitive tools.

What replaces mean–variance–correlation?

Principle

Replace global, quadratic, and linear statistics with measures that are:

- Distribution-aware,
- State-conditional,
- Robust to tails and jumps,
- Aligned with economic objectives.

What replaces mean–variance–correlation?

Examples

- **Quantiles and expected shortfall** instead of mean and variance,
- **Drawdowns and path functionals** instead of second moments,
- **Copulas and tail dependence** instead of linear correlation,
- **Information-theoretic measures** for nonlinear dependence.

These tools preserve economically relevant structure that is systematically destroyed by averaging and squaring.

Design principles for modern financial models

Model construction

Effective financial models should:

- Condition on information sets and regimes,
- Respect heavy tails and discontinuities,
- Separate centre behaviour from tail risk,
- Explicitly model path dependence.

Risk evaluation

Risk should be assessed using:

- Tail-sensitive loss functions,
- Scenario and stress-based analysis,
- Forward-looking conditional measures.

This shift is not cosmetic: it changes portfolio construction, risk limits, and regulatory capital in material ways.

Takeaway

Core message

Mean, variance, and correlation are not wrong — they are *incomplete*.

- They work in linear, Gaussian, stationary worlds.
- Financial markets are nonlinear, heavy-tailed, and regime-driven.
- Using the wrong statistics produces a false sense of control.

Practical implication

Modern quantitative finance requires:

- State-conditional modelling,
- Tail-aware risk measures,
- Path-sensitive evaluation criteria.

The challenge is not computational — it is conceptual.

The end

Thank You !