Financial Risk and Return Distributions

Week 1 — Financial Management: Volatility, Risk, and AI

Yi-Hao Lai

2026 Spring

Today's Roadmap

1

The Story: Why This Matters

The VolTech Challenge

Scenario: A Japanese pension fund lost 18% in one week during COVID-19.

Their risk model predicted maximum weekly loss of **4.2%** at 99% confidence.

What went wrong?

The Core Question

Why do "once in 14,000 years" events happen every 3 years?

Today's Learning Objectives

By the end of this session, you will:

1. Calculate simple and log returns
2. Explain why normality fails for financial returns
3. Measure skewness and excess kurtosis
4. Read QQ plots fluently
5. Run the Jarque–Bera test in Python

2

Returns: The Language of Finance

Simple vs. Log Returns

Simple Return

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

Log Return

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) = \ln(1 + R_t)$$

Why Log Returns?

1. **Time-additive**: $r_{t:t+k} = r_{t+1} + r_{t+2} + \cdots + r_{t+k}$
   Simple returns need geometric compounding: $(1 + R_{t:t+k}) = \prod(1 + R_i)$
2. **Symmetric**: $+10\%$ and -10% log returns are symmetric around zero
3. **Approximation**: For small x, $\ln(1 + x) \approx x$, so the two are nearly identical for typical daily returns ($< 5\%$)

Remember

We analyze **returns**, not prices. Prices drift; returns fluctuate around a stable mean.

3

The Normal Distribution Problem

The Gaussian Model

If returns are normally distributed:

$$r_t \sim \mathcal{N}(\mu, \sigma^2)$$

The famous σ-rules:

Range	Coverage	1-sided tail
$\pm 1\sigma$	68.3%	15.87%
$\pm 2\sigma$	95.4%	2.28%
$\pm 3\sigma$	99.7%	0.13%
$\pm 5\sigma$	99.99994%	0.00003%

A 5σ event: once every 13,900 years of trading.

Reality Check

The Problem with Normality

The S&P 500 experienced **20+ daily moves exceeding** $\pm\mathbf{5}\sigma$ between 1990 and 2024.

Events that "should" happen once in millennia occur **regularly**.

Solution preview: The Student-t distribution ($\nu \approx 4$–8) provides heavier tails. We will use it as the error distribution in GARCH models (Week 3).

4

Fat Tails, Skewness, and Kurtosis

Skewness — The Third Moment

Skewness

$$S = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{r_t - \bar{r}}{\hat{\sigma}}\right)^3$$

- $S = 0$: symmetric (normal)
- $S < 0$: **negatively skewed** — large losses more common than large gains
- $S > 0$: positively skewed

Typical equity returns: $S \approx -0.3$ to -1.0

Excess Kurtosis — The Fourth Moment

Excess Kurtosis

$$K = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{r_t - \bar{r}}{\hat{\sigma}}\right)^4 - 3$$

- $K = 0$: mesokurtic (normal-like tails)
- $K > 0$: **leptokurtic** — fat tails, more extremes
- $K < 0$: platykurtic — thin tails

Typical daily stock returns: $K \approx 5$–20

Kurtosis measures **tail weight**, not "peakedness"

Why Do Fat Tails Arise?

1. **Volatility clustering**: calm $\rightarrow$ turbulent periods mix different-variance normals $\rightarrow$ fat unconditional tails (Week 3: GARCH)
2. **Leverage effect**: bad news $\uparrow$ vol more than good news (Week 4: GJR-GARCH)
3. **Herding**: collective panic amplifies downside
4. **Market microstructure**: halts, margin calls, forced liquidations create jumps

Key Insight

Time-varying volatility alone can produce fat tails, even if each period's returns are conditionally normal.

5

Testing for Normality

QQ Plot: Your Visual Diagnostic

A **QQ plot** compares sample quantiles against theoretical (normal) quantiles.

Reading guide:

- Points on line → normal
- S-shape → fat tails
- Curved left end → heavier left tail

Key Pattern

For financial returns, you will **always** see the S-shape. This is the single most important diagnostic plot in this course.

Jarque–Bera Test

JB Test Statistic

$$JB = \frac{T}{6}\left(S^2 + \frac{K^2}{4}\right) \sim \chi^2(2) \quad \text{under } H_0$$

Decision: Reject normality if $JB > 5.99$

In Practice

Daily equity returns produce JB in the **hundreds or thousands**. Normality is rejected with overwhelming certainty.

6

Python Live Demo

Step 1: Load Data and Compute Returns

```python
import numpy as np, pandas as pd
import yfinance as yf
from scipy import stats

sp500 = yf.download("^GSPC",
    start="2015-01-01", end="2024-12-31")
prices = sp500["Close"].squeeze()
log_returns = np.log(
    prices / prices.shift(1)).dropna()
```

Step 2: Descriptive Statistics and JB Test

```
print(f"Skewness:  {log_returns.skew():.4f}")
print(f"Kurtosis:  {log_returns.kurtosis():.4f}")

jb_stat, jb_p = stats.jarque_bera(
    log_returns.values)
print(f"JB stat:   {jb_stat:.2f}")
print(f"p-value:   {jb_p:.2e}")
```

Expected: Skewness ≈ -0.6, Kurtosis ≈ 12, JB $\gg 5.99$

Step 3: Histogram and QQ Plot

```python
import matplotlib.pyplot as plt
fig, axes = plt.subplots(1, 2,
    figsize=(14, 5))
# Left: histogram + normal overlay
axes[0].hist(log_returns, bins=100,
    density=True, alpha=0.7, color="#2D7D5E")
# Right: QQ plot
stats.probplot(log_returns.values,
    dist="norm", plot=axes[1])
```

Look for: tall center peak, heavy tails, S-shape in QQ

7

Application: Kenji's Portfolio

Tail Probability Comparison

Metric	Normal	Actual Data
Skewness	0	-0.73
Excess Kurtosis	0	13.4
$P(\text{loss} > 3\sigma)$	0.13%	1.7%
$P(\text{loss} > 5\sigma)$	0.00003%	0.12%

At 5σ: the normal model is off by a factor of **4,000**$\times$

This is why Kenji's VaR model failed.

AI Connection: Anomaly Detection

AI in Finance

Modern AI systems exploit the same insight — returns are *not* normal — to build better anomaly detectors.

- **Isolation Forests**: detect outliers without assuming any distribution
- **Autoencoders**: learn “normal” patterns and flag deviations
- No need to specify skewness or kurtosis — the model learns the distribution from data

We explore these in Week 8.

8

Key Takeaways

Five Things to Remember

1. **Returns, not prices**: Use log returns for statistical modeling — they are time-additive and approximately stationary
2. **Normality fails**: Fat tails and negative skewness are universal in financial returns
3. **Four moments**: Mean, variance, skewness, kurtosis — all four matter for risk assessment
4. **QQ plot + JB test**: First-line diagnostic tools for assessing distributional fit
5. **Model implications**: Normal-based risk models systematically underestimate tail risk by orders of magnitude

Mission 1: Distribution Diagnostics Report

Deliverables

1. Download S&P 500, Nikkei 225, FTSE 100 (2015–2024)
2. Compute 4 moments + JB test for each
3. Histogram + QQ plot for each
4. Comparison table across all indices
5. 200-word executive summary

Bonus: Empirical tail probability table (1σ to 5σ)

Next Week Preview

Week 2: Measuring Volatility

Kenji asks: “How volatile is our portfolio *right now*?”

We learn that you cannot measure risk with a single ruler.

Topics: Rolling windows, EWMA, realized volatility, VIX