Technical Report: Cointegration-Based Pairs Trading in Indian Equities

Course: Applied Stochastic Processes
Date: November 12, 2025

---

1. Introduction

Pairs trading is a market-neutral strategy built upon identifying cointegrated asset pairs — those whose individual prices behave like random walks but whose **linear combination forms a stationary, mean-reverting spread**. This report details the application of this principle on Indian pharmaceutical equities, combining stochastic financial theory with practical quantitative techniques to engineer, validate, and backtest a statistically grounded trading strategy.

---

2. Model Foundations and Approach Integration

Random Walk Assumption for Prices

Individual stock prices P_t are modeled as **random walks**:

$$P_t = P_{t-1} + \varepsilon_t$$

where ε_t are independent noise innovations. This captures the nonstationarity and unpredictability that challenge direct asset forecasting, motivating the search for stable inter-asset relationships.

Cointegration Model for Pairs Selection

Modeling the relationship between two stock prices P_A and P_B, an **Ordinary Least Squares (OLS) regression** on log-prices is performed:

$$\log P_A = \alpha + \beta \log P_B + \varepsilon_t$$

where:

- β is the **hedge ratio**, quantifying the relative exposure between assets.
- ε_t is the **spread residual** — a candidate mean-reverting process.

By estimating β, the strategy constructs a hedge that neutralizes systemic trends, facilitating analysis of spread stationarity.

Stationarity Confirmation with ADF Test

The spread $S_t = \log P_A - \beta \log P_B$ is tested for **stationarity** using the **Augmented Dickey-Fuller (ADF) test**:

$$\Delta S_t = \alpha + \gamma S_{t-1} + \sum_{i=1}^{p} \delta_i \Delta S_{t-i} + u_t$$

- Null hypothesis $H_0: \gamma = 0$ implies nonstationary spread.

- Alternative $H_1: \gamma < 0$ implies stationary spread.

Rejection of H_0 with a statistically significant test statistic (e.g., $t = -3.45 < -2.89$ critical value at 5%) and small p-value (<0.05) validates cointegration suitability.

Spread Dynamics

With stationarity confirmed, the spread is modeled :

$$dS_t = \theta(\mu - S_t)dt + \sigma dW_t$$

where:

- μ is the long-run mean spread,
- θ quantifies the **speed of mean reversion**,
- σ is volatility,
- dW_t is Brownian motion noise.

This continuous-time model bears direct implications for the timing and size of trading signals.

---

3. Threshold Optimization

Z-Score Construction

The spread is normalized via rolling mean $\bar{S}$ and std deviation σ_S to form the **z-score**:

$$Z_t = \frac{S_t - \bar{S}}{\sigma_S}$$

Signaling captures standardized deviation magnitude conducive to mean reversion.

Threshold Selection: Empirical Quantiles

Traditional pairs trading fixes entry and exit signals at arbitrary z-score thresholds, often ± 1.5. This work employs **data-driven quantile thresholds**:

- **Long Entry:** $Z_t < q_{0.05}$
- **Short Entry:** $Z_t > q_{0.95}$
- **Exit:** Crossing the median (near zero)

  From this we have got:
  Lower Threshold value = -1.86
  Upper Threshold value = 2.64
  Exit Threshold value = 0.83

4. Implementation & Backtesting

- **Data:** Downloaded daily close prices for 10 pharma stocks (2017-2020).

- **Pair Selection:** OLS regression and cointegration tests identify CIPLA–LUPIN as the optimal pair.
- **Backtest Window:** June 2020 – June 2023 out-of-sample trading with simulated position management.

---

5. Results and Discussion

- The CIPLA–LUPIN spread was validated as stationary (ADF test statistic -3.45, p-value < 0.05).
- Made from 2.5% percent in 3-year period.(2021-23)

---

6. Conclusion

By integrating random walk assumptions, Engle-Granger cointegration, ADF stationarity testing, OU spread modeling, and Sharpe ratio threshold optimization, this strategy forms a cohesive framework for **statistically robust, market-neutral pairs trading**. The systematic use of these models throughout data selection, signal construction, and portfolio management bridges stochastic theory and quantitative finance practice, delivering a replicable and interpretable trading methodology.