Hypothesis Testing: Market Making on CSI 300 ETF via Signal-Gated DRL

Jiaxing Wei
jwei2002@uw.edu

1 Hypothesis Testing Design

Following the criteria for a rigorous, systematic trading hypothesis, this section defines the statistical tests designed to validate the Signal-Gated DRL framework on the 510300 ETF. Given the high-frequency nature of the data, which typically exhibits fat tails and serial correlation, we move beyond simple t-tests to incorporate robust econometric evaluations.

1.1 H1: Predictive Efficacy and Forecast Robustness of SGUs

- **Subject**: The standalone Signal Generating Units (XGBoost for volatility range and LSTM for trend).
- **Independent Variables**: LOB-derived features including order flow imbalance, historical realized price ranges, and volume-weighted average price (VWAP).
- **Dependent Variables**: Actual modified realized price range (y_i^{RR}) and financial pseudo-returns (y_i^{TR}).
- **Anticipated Outcome**: SGUs will generate signals that are significantly more accurate than a naive drift forecast, even after adjusting for the high noise-to-signal ratio inherent in A-share ETF microstructure.
- **Validation Method**: To account for heteroscedasticity and autocorrelation (HAC) in high-frequency residuals, we employ **Newey-West adjusted standard errors**. The HAC covariance matrix estimator is defined as $\hat{\Sigma}_{HAC} = (X'X)^{-1}\mathbf{S}(X'X)^{-1}$, where $\mathbf{S}$ is split to ensure proper alignment:

$$\begin{aligned}\mathbf{S} = &\sum_{t=1}^{T} \hat{\epsilon}_t^2 \mathbf{x}_t \mathbf{x}_t' + \\ &\sum_{j=1}^{L} w_j \sum_{t=j+1}^{T} \hat{\epsilon}_t \hat{\epsilon}_{t-j} (\mathbf{x}_t \mathbf{x}_{t-j}' + \mathbf{x}_{t-j} \mathbf{x}_t')\end{aligned} \tag{1}$$

where $w_j = 1 - \frac{j}{L+1}$ represents the Bartlett kernel weights for lag L. Additionally, we use the **Diebold-Mariano (DM) test**:

$$DM = \frac{\bar{d}}{\sqrt{\hat{V}(\bar{d})/T}} \sim N(0,1) \tag{2}$$

where $d_t = L(e_{t,\text{naive}}) - L(e_{t,\text{SGU}})$ is the loss differential between the naive baseline and the SGU prediction.

- **Significance**: We reject the null hypothesis of equal forecast accuracy if the DM statistic > 1.96 ($p < 0.05$).

1.2 H2: Nonlinear Execution Intensity and Microstructure Facts

- **Subject**: The Limit Order Book (LOB) execution dynamics near price limits.
- **Independent Variables**: Distance to the $\pm 10\%$ price limit ($D_{limit} = |M_t - P_{limit}|$).
- **Dependent Variables**: Realized fill rates $\lambda(\delta)$.
- **Anticipated Outcome**: The execution intensity $\lambda(\delta)$ deviates from the standard exponential decay as $D_{limit} \to 0$, showing a cliff effect where liquidity vanishes.
- **Validation Method**: We use the **Durbin-Watson (DW) statistic** to detect first-order autocorrelation in the residuals of the fitted intensity model $\lambda(\delta) = Ae^{-k\delta}$:

$$d = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2} \tag{3}$$

where e_t is the residual at time t.

- **Significance**: A value of d significantly lower than 2 indicates positive serial correlation, confirming that a static exponential model fails to capture the persistent microstructure dependencies near price limits.

1.3 H3: Comparative Superiority and Risk-Adjusted Returns

- **Subject**: DRL Agent performance vs. GLFT benchmark.
- **Independent Variables**: Strategy configuration.
- **Dependent Variables**: Terminal Wealth (W_T) and Sharpe Ratio (SR).
- **Anticipated Outcome**: The DRL agent achieves a significantly higher PnL-to-MAP ratio and Terminal Wealth than GLFT.

- **Validation Method**: Since HFT returns are non-Gaussian and exhibit volatility clustering, we employ **Moving Block Bootstrapping (MBB)** to construct robust confidence intervals. For a sequence of returns $\{r_t\}$, we compute the bootstrap distribution of the Sharpe Ratio difference $\Delta SR^* = SR^*_{DRL} - SR^*_{GLFT}$:

$$\text{Var}^*(\Delta SR) = \frac{1}{B-1}\sum_{b=1}^{B}(\Delta SR^*_b - \overline{\Delta SR}^*)^2 \tag{4}$$

  We also report the asymptotic standard error for the estimated Sharpe Ratio:

$$SE(\widehat{SR}) = \sqrt{\frac{1 + \widehat{SR}^2/2}{T}} \tag{5}$$

- **Significance**: The DRL agent is deemed superior if the 95% MBB confidence interval for ΔSR excludes zero.

1.4 H4: Economic Rationality of Policy Skewing

- **Subject**: The learned reward-driven policy π_θ.
- **Independent Variables**: Inventory level I_t.
- **Dependent Variables**: Quoting offsets $A_t = [Q_{ask} - M, M - Q_{bid}]$.
- **Anticipated Outcome**: The agent prioritizes spread capturing over trend chasing by enforcing a linear absolute inventory penalty: $R_{t+1} = \text{PnL}_{t+1} - \lambda|I_{t+1}|$.
- **Validation Method**: Analysis of Partial Dependence Plots (PDP) to estimate the marginal effect of I_t on A_t:

$$\hat{h}_S(x_S) = \frac{1}{n}\sum_{i=1}^{n} f(x_S, \mathbf{x}_{C,i}) \tag{6}$$

- **Significance**: The policy is verified if $\frac{\partial \text{AskOffset}}{\partial I} < 0$ and $\frac{\partial \text{BidOffset}}{\partial I} > 0$, demonstrating that the agent has successfully learned to skew quotes to mitigate directional risk.

2 Statistical Robustness and HFT Considerations

To address the challenges of high-frequency data, we acknowledge that standard parametric tests may yield inflated t-statistics due to the high sample size (T). Therefore, we prioritize the **Model Confidence Set (MCS)** to identify the superior strategy among a group of candidates. The MCS procedure sequentially eliminates inferior models based on a loss function $L_{i,t}$, until the remaining set

$\mathcal{M}^*$ contains only models with equal predictive power at a given confidence level:

$$t_{ij} = \frac{\bar{d}_{ij}}{\sqrt{\widehat{\text{Var}}(\bar{d}_{ij})}} \tag{7}$$

where $d_{ij,t} = L_{i,t} - L_{j,t}$ represents the loss differential. This avoids the p-hacking pitfalls associated with massive datasets where even economically negligible differences appear statistically significant.

2.1 Microstructural Adaptation: BTC vs. CSI 300 ETF

While the core First-Passage Time (FPT) execution model is preserved, the transition from the original Bitcoin environment to the CSI 300 ETF requires specific adjustments to account for distinct microstructural regimes:

- **Finite Horizon and Session Boundary Effects**: Unlike the 24/7 continuous trading of BTC, the 510300 ETF operates within two discrete 2-hour sessions. We modify the reward logic to account for the heightened inventory risk as the market approaches the 15:00 CST close. This necessitates a time-dependent weighting of the absolute inventory penalty $\lambda|I_{t+1}|$ to ensure the agent avoids a liquidity trap at the end of the trading day.
- **FPT Validity near Price Limits**: In the A-share market, the $\pm10\%$ price limit imposes a hard boundary on the FPT model. In the original BTC framework, execution intensity is assumed to decay with distance from the mid-price. For the ETF, we test the hypothesis that the FPT model remains valid but requires a volatility-gating mechanism; specifically, once the price is locked at a limit, the execution intensity λ drops to zero on the side of the limit, regardless of the quoting offset.
- **Tick Size and Spread Saturation**: The 510300 ETF often experiences spread saturation, where the bid-ask spread is locked at a single tick (0.001) for extended periods. This differs from the more volatile BTC spreads. Our testing framework evaluates whether the DRL agent's continuous action space A_t allows it to exploit sub-tick order flow signals that are invisible to benchmarks utilizing crude discretization.

2.2 Robustness to Intraday Seasonality

To ensure that performance metrics are not skewed by high-volatility opening auctions or midday pauses, we apply **Newey-West adjustments** to all strategy regressions. This correction accounts for the significant intraday autocorrelation in returns:

$$\begin{aligned} \mathbf{S} = &\sum_{t=1}^{T} \hat{\epsilon}_t^2 \mathbf{x}_t \mathbf{x}_t' + \\ &\sum_{j=1}^{L} \left(1 - \frac{j}{L+1}\right) \sum_{t=j+1}^{T} \hat{\epsilon}_t \hat{\epsilon}_{t-j} (\mathbf{x}_t \mathbf{x}_{t-j}' + \mathbf{x}_{t-j} \mathbf{x}_t') \end{aligned} \tag{8}$$

This ensures that the reported p-values reflect the true significance of the DRL agent's alpha beyond the expected intraday seasonality of the China A-share market.