Calibration Explained

What the Sunday gate does

Every Sunday evening before any picks ship, the runner re-executes the full canonical 2019–2024 walk-forward backtest on the frozen reference dataset and confirms three published numbers reproduce within tolerance:

Metric	Expected	Tolerance
Sharpe Ratio	0.717	±0.010
CAGR	15.4%	±0.5%
Max Drawdown	22.6%	±0.5%

If any of those three drift outside tolerance, **picks do not ship that week.** The runner instead generates an Integrity Letter that goes to both Premium and Free subscribers, explaining what failed and what's being investigated.

Why bother doing this every week?

Three reasons.

1. Code drift. The same Python source code can produce different output if the underlying libraries change. A pandas version bump that subtly changes how rolling-window calculations handle NaN edges can shift Sharpe by 0.02. The calibration gate catches this immediately — before subscribers see picks built on a silently-different engine.

2. Data drift. Our production picks use live yfinance data. yfinance occasionally revises historical data (corporate action adjustments, split corrections, dividend reinstatements). If a 2023 close gets revised this week, the model's training data is now subtly different than what produced the canonical 0.717. The gate catches this too — and you, the subscriber, deserve to know if it's happening.

3. Trust. A signal you can verify is a signal you can use. Most weekly research subscriptions ask you to take it on faith that the methodology is still working. The calibration gate makes that verifiable each week. If we drift, we tell you. If we shipped picks anyway when we shouldn't, that's a credibility crack we can't recover from.

What "within tolerance" means in practice

The frozen 2019–2024 backtest is run end-to-end every Sunday — fresh model fit, fresh walk-forward, fresh metric calculation. The output is compared to the published canonical values. If Sharpe comes out at 0.716 instead of 0.717, that's within ±0.010 tolerance and the gate passes. If it comes out at 0.682, that's outside tolerance — gate fails, picks held.

The tolerance levels were set during system design to allow for the small numerical noise that natural floating-point variation can produce, while still being tight enough to catch any real regression.

What happens when the gate fails

You receive an Integrity Letter. It contains: - The exact metric that drifted - The actual value vs. expected value - What we know so far about the cause - When we expect to resume

What it does NOT contain: this week's picks. We won't ship picks the engine can't justify.

The Integrity Letter is part of the subscription. Some weeks the Letter IS the deliverable. That's by design.

How often will this realistically fail?

We don't know yet — Live launched May 26, 2026. Historically, in development: - The 2019–2024 calibration gate fires correctly on every test run since the system was built - A yfinance data revision in early 2025 caused a Sharpe drift of 0.04 (above tolerance) before it was reconciled — that's a real-world example of what the gate catches

Realistically, we expect a few Integrity Letters per year. That's the system working.

The deeper point

A model's calibration result is a contract between the methodology and reality. Most weekly subscriptions ship picks regardless of whether that contract still holds. We don't. The gate is the differentiator.

If you ever wonder whether The Signal Live is still doing what it says it does — the calibration report at the top of every Monday's email is your answer.