Monte Carlo Analysis of SPY Returns and Black–Scholes Option Pricing

Geometric Brownian Motion, Statistical Validation, and Model Risk

Milan Fusco — Fall 2025

This notebook evaluates the adequacy of Geometric Brownian Motion (GBM) as a structural model for SPY equity returns and assesses the statistical reliability of Monte Carlo option pricing under the Black–Scholes framework.

All simulations are performed using `mcframework`, a purpose-built Monte Carlo engine featuring reproducible parallel random number generation via NumPy Philox streams.

---

0. Environment Setup

1. Data and Parameter Estimation

We retrieve one year of daily SPY closing prices (Dec 2024 – Dec 2025) and compute log returns:

$$r_t = \log\left(\frac{S_t}{S_{t-1}}\right)$$

Under Geometric Brownian Motion (GBM), price dynamics satisfy:

$$\frac{dS_t}{S_t} = \mu\,dt + \sigma\,dW_t$$

The closed-form solution is:

$$S_T = S_0 \exp\Big((\mu - \tfrac{1}{2}\sigma^2)T + \sigma W_T\Big)$$

Therefore, log returns are normally distributed:

$$\log\left(\frac{S_T}{S_0}\right) \sim \mathcal{N}\left((\mu - \tfrac{1}{2}\sigma^2)T,\ \sigma^2 T\right)$$

We estimate the real-world drift μ and volatility σ from historical log returns.

```
Spot Price  S₀ : $681.38
Annual Drift μ : 0.1384
Annual Vol   σ : 0.1969
Daily Mean     : 0.000549
Daily Std Dev  : 0.012401
N (trading days): 248
```

---

2. Geometric Brownian Motion — Path Simulation

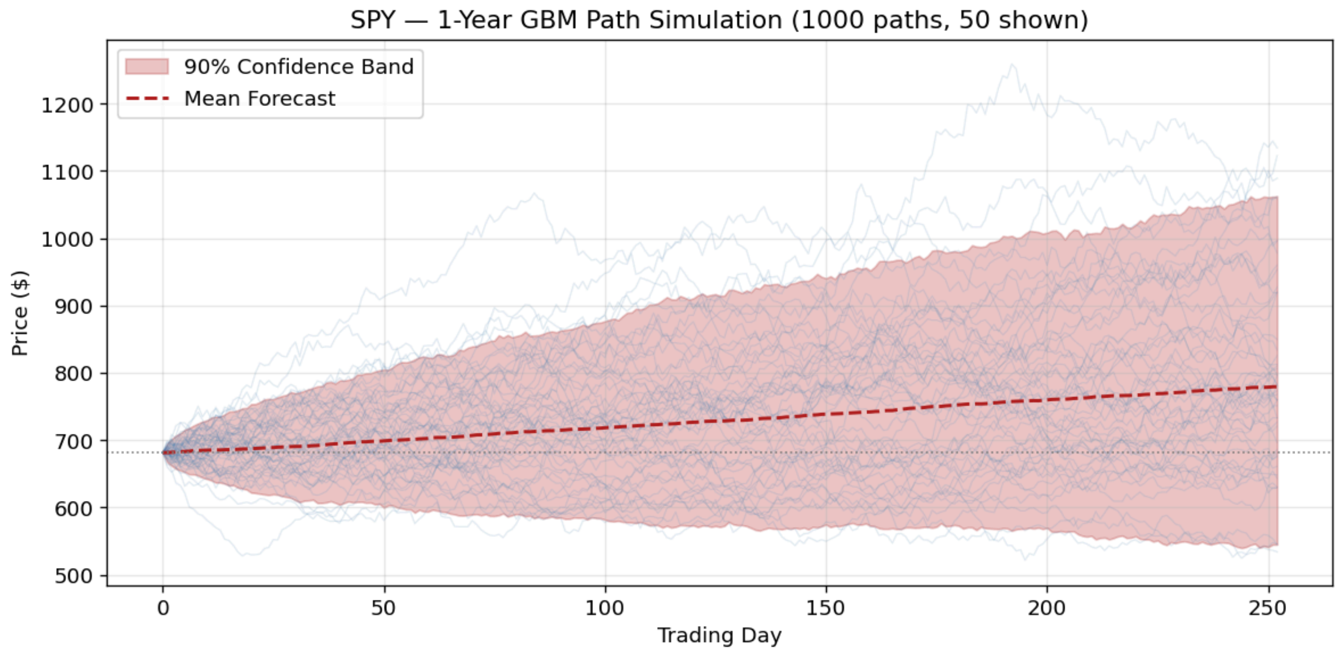
Simulations are executed using a reproducible parallel Monte Carlo engine built on:

- `numpy.random.SeedSequence`
- Independent Philox random number streams
- Deterministic stream splitting across workers
- Bitwise reproducibility regardless of parallelization (see docs for GPU caveats)

The GBM SDE and its closed-form solution:

$$\frac{dS_t}{S_t} = \mu\, dt + \sigma\, dW_t \quad \Longrightarrow \quad S_T = S_0 \exp\left[\left(\mu - \tfrac{1}{2}\sigma^2\right) T + \sigma W_T\right]$$

We use `BlackScholesPathSimulation` under the **real-world measure** (drift $= \mu$) to visualise the fan of plausible price trajectories.



4. One-Year Forecast Distribution (Real-World Measure)

`PortfolioSimulation` models terminal wealth V_T under GBM.

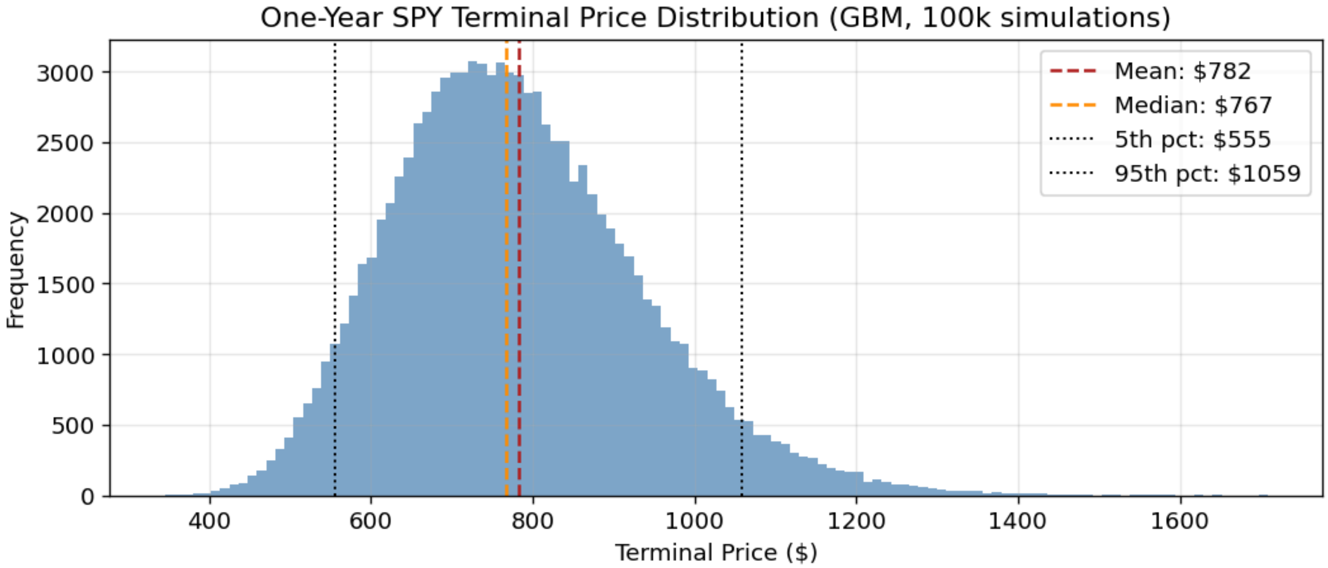
Terminal prices are simulated under the real-world drift μ.

The resulting distribution is lognormal, producing:

- Mean > Median (right-skewed asymmetry)
- Wide tail dispersion
- Probability estimates for upward price movement

Monte Carlo estimates include standard error and percentile-based Value-at-Risk metrics.

```
Mean terminal price  : $782.16
Median               : $767.26
5th percentile  (VaR): $555.33
95th percentile      : $1058.55
P(S_T > S₀)          : 0.73
```



---

4. Empirical Validation — Distributional Comparison

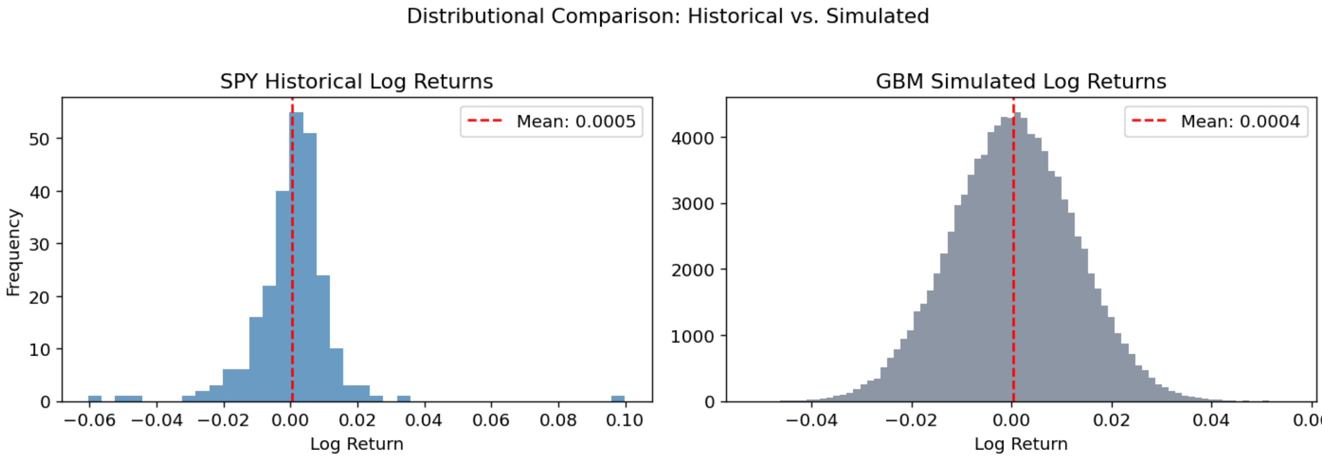
Although GBM matches the first two moments (mean and variance) by construction, its structural assumptions imply:

- Zero skewness
- Zero excess kurtosis
- Normally distributed log returns

We compare historical SPY log returns against simulated GBM returns.

```
=== Historical SPY Log Returns ===
  Mean     : 0.000549
  Std Dev  : 0.012376
  Skewness : 1.014
  Ex. Kurt : 19.182

=== GBM Simulated Log Returns ===
  Mean     : 0.000393
  Std Dev  : 0.012417
  Skewness : -0.009
  Ex. Kurt : 0.004
```



5. Hypothesis Testing

5a. Equality of Means — Two-sample z-test

Fail to reject equality of means, consistent with GBM's construction to match the first moment of historical returns.

```
Mean Test:   z = 0.198,   p = 0.843
```

5b. Equality of Variances — Levene's test

Reject equality of variances, consistent with the wider tails observed in historical returns.

```
Var  Test:   W = 23.826,   p = 0.000
```

5c. Normality of Historical Returns — Shapiro-Wilk

$p < 10^{-6} \Rightarrow$ reject normality of historical returns, consistent with fat tails observed in the histogram and excess kurtosis.

```
Normality:   W = 0.8014, p = 4.45e-17
```

Conclusion: Although the first moment is statistically consistent with calibration, historical returns strongly reject normality and exhibit variance behavior inconsistent with GBM.

The failure arises from higher-order moment structure, not from Monte Carlo noise.

6. Risk-Neutral Option Pricing

Under the risk-neutral measure $\mathbb{Q}$:

$$\frac{dS_t}{S_t} = r\,dt + \sigma\,dW_t^{\mathbb{Q}}$$

Under $\mathbb{Q}$, discounted asset prices are martingales. The real-world drift μ is replaced by the risk-free rate r to ensure no-arbitrage pricing.

Closed-form Black–Scholes solution:

$$C = S_0\Phi(d_1) - Ke^{-rT}\Phi(d_2)$$

$$d_1 = \frac{\ln(S_0/K) + \left(r + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}$$

$$d_2 = d_1 - \sigma\sqrt{T}$$

```
European Call  Ĉ = $70.94
  SE = 0.3119   95% CI: [$70.33, $71.55]

European Put   P̂ = $36.24
  SE = 0.1813   95% CI: [$35.88,  $36.60]

Put–Call Parity check:
  C − P = 34.6985
  S₀ − K·e^(−rT) = 34.8495
  Residual: 0.1510  (passes)
```

7. Greeks via Finite Differences

Option sensitivities are estimated using central finite differences.

Each Greek requires two simulations per parameter perturbation.

Computed sensitivities include:

- Delta: $\partial V/\partial S$
- Gamma: $\partial^2 V/\partial S^2$
- Vega: $\partial V/\partial \sigma$
- Theta: $\partial V/\partial T$
- Rho: $\partial V/\partial r$

Finite-difference estimates introduce additional Monte Carlo noise, but confidence intervals remain stable relative to price level.

```
European Call Greeks (ATM)
  Delta (dV/dS)             : 0.6396
  Gamma (d²V/dS²)           : 0.002763
  Vega  (dV/dσ, per 1% move) : 2.5153
  Theta (daily decay)       : -0.1204
  Rho   (dV/dr, per 1% move) : 3.6498
```

8. Conclusions

Key Findings

1. GBM reproduces empirical mean and variance by construction.
2. Historical SPY returns exhibit statistically significant skewness and extreme excess kurtosis inconsistent with GBM assumptions.
3. Monte Carlo pricing under the risk-neutral measure is numerically stable; sampling error is small relative to price magnitude.
4. Put–call parity holds within Monte Carlo tolerance, confirming internal consistency.
5. Structural model error dominates Monte Carlo sampling error.

Implication

Black–Scholes provides a coherent no-arbitrage pricing framework and GBM reproduces first-order and second-order moments by calibration.

However, its underlying GBM assumption inadequately captures real-world return dynamics. Significant positive skewness and extreme excess kurtosis in historical returns indicate heavy tails inconsistent with the Gaussian assumption.

Structural vs Sampling Error

Monte Carlo standard error decreases at rate $1/\sqrt{N}$.

However, no increase in simulation count can correct structural misspecification in the underlying return model.

Empirical skewness and heavy tails indicate that model risk, not Monte Carlo variance, is the dominant source of pricing error under GBM.

Future Directions

- Heston stochastic volatility
- GARCH dynamics
- Jump-diffusion (Merton)
- Calibration to implied volatility surface
- Out-of-sample validation

References

Black, F. & Scholes, M. (1973). *The pricing of options and corporate liabilities.*

Fusco, M. (2026). mcFramework.

NumPy, SciPy, Matplotlib, yfinance.