Optimal Hedging with Advanced Delta Modelling

Table of Contents

1.Geometric Brownian motion asset simulation

Here it will involve three methods: Euler-Maruyama scheme, Milstein scheme, and Sobol sequence with the Brownian Bridge method.

1.1Euler-Maruyama scheme

Started with the general form of a stochastic differential equation (SDE):

$$dY_t = a(t, Y_t)dt + b(t, Y_t)dX_t$$

Where:

Y_t: stochastic process

$a(t, Y_t)$: drift term

$b(t, Y_t)$: diffusion term

X_t: standard Brownian motion (Wiener process)

Euler-Maruyama Scheme:

It is a discrete approximation. Given a time step δt, the simulation is

$Y_{t+\delta t} = Y_t + \ a(t, Y_t)\delta t + \ b(t, Y_t)dX_t$

where dX_t is the Brownian increment

$dX_t = X_{t_{i+1}} - X_{t_i} \sim N(0, \delta t)$, is typically simulated as: $dX_t = \sqrt{\delta t}\phi_i$, $\phi_i \sim N(0,1)$

For geometric Brownian motion, the SDE formula is $dS_t = \mu S_t dt + \sigma S_t dX_t$,

so the Euler scheme formula is $S_{t+\delta t} = S_t(1 + \mu\delta t + \sigma\sqrt{\delta t}\phi_i)$ (1)

1.2 Milstein scheme

It includes a second-order term from the Itô -Taylor expansion to improve accuracy.

The formula is: $Y_{t+\delta t} = Y_t + \ a(t, Y_t)\delta t + \ b(t, Y_t)dX_t + \frac{1}{2}b(t, Y_t)\frac{\partial}{\partial Y}b(t, Y_t)((dX_t)^2 - \delta t)$

For geometric Brownian motion:

$b(t, Y_t) = \ \sigma S_t$,

$\frac{\partial}{\partial Y}b(t, Y_t) = \frac{\partial}{\partial S_t}(\sigma S_t) = \sigma$

So the Milstein scheme formula becomes:

$$S_{t+\delta t} = S_t(1 + \mu\delta t + \sigma\sqrt{\delta t}\phi_i + \frac{1}{2}\sigma^2(\delta t\phi_i^2 - \delta t)) \qquad (2)$$

1.3 Properties

Euler-Maruyama Scheme:

It is of simple implementation, and minimal computational cost

It has strong order of convergence: O($\sqrt{\delta t}$)

It fails to capture curvature of diffusion term

It has weak order of convergence 1

Milstein scheme:

It requires computing derivative of the diffusion term

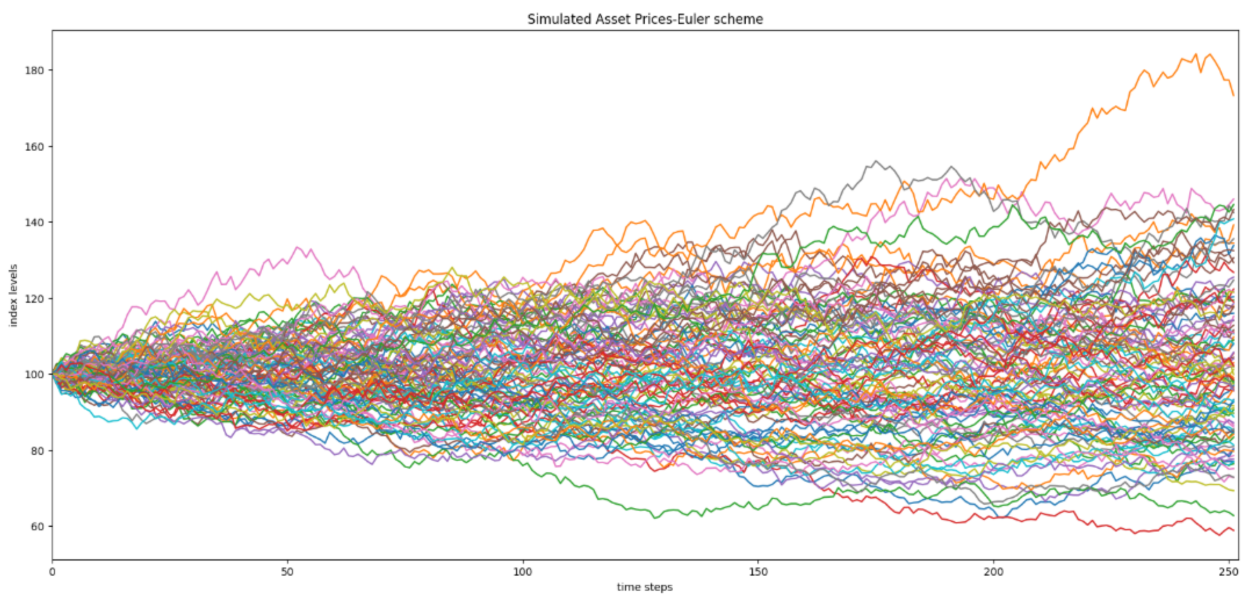
It has strong order of convergence O(δt), more accuracy than Euler

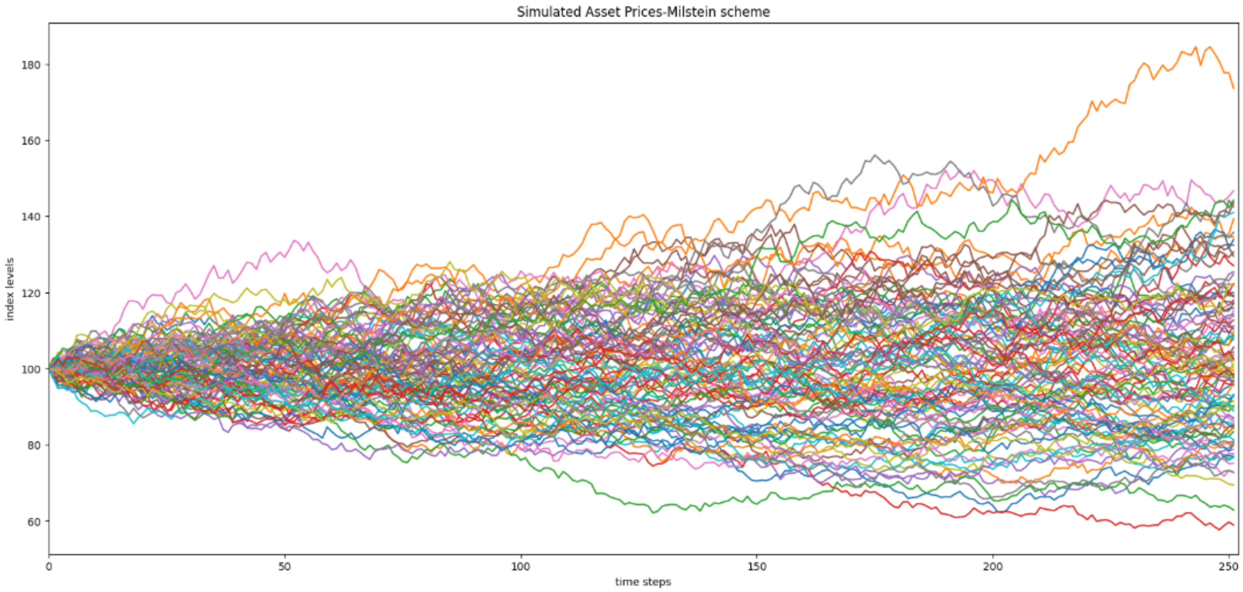
It captures Itô-Taylor expansion up to O(δt)

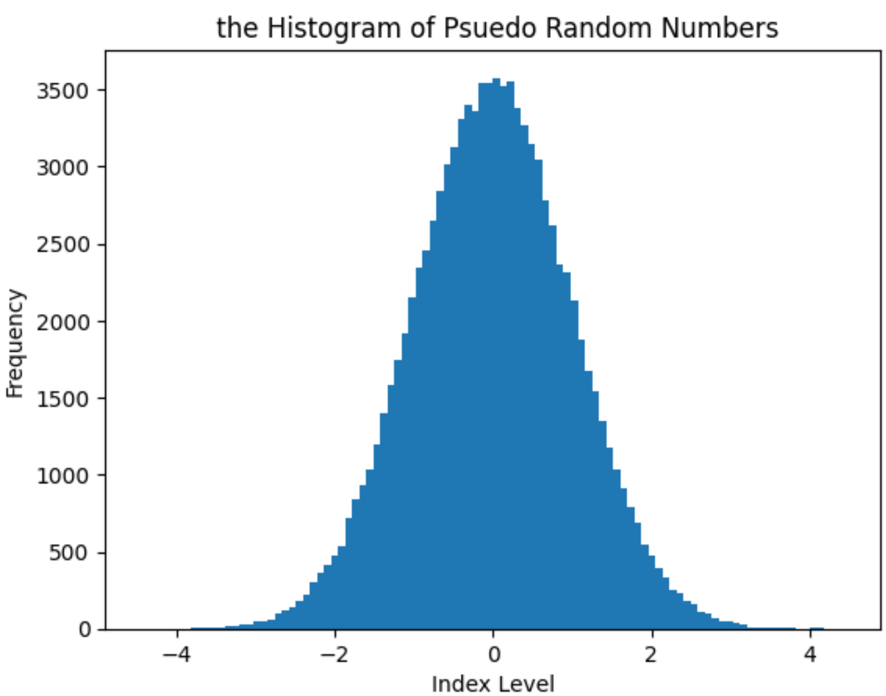
It has weak order of convergence 1

1.4 Path simulation and observations

Using formula (1) and (2) for implementation.

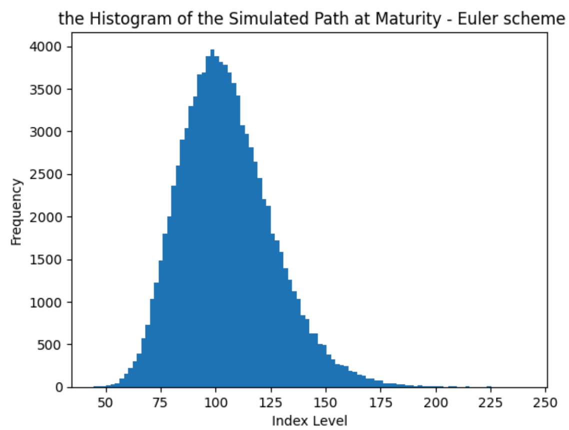


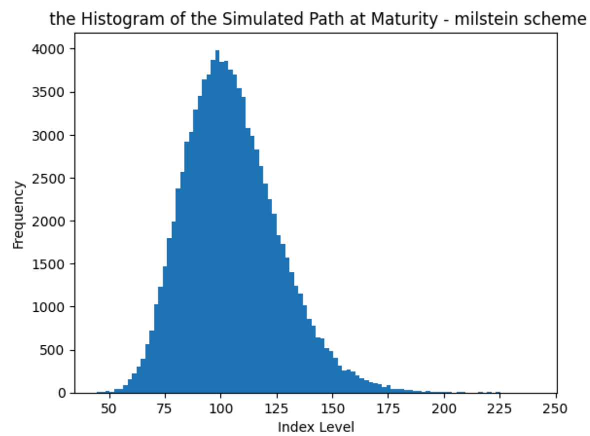





```
pseudorandom.mean(), pseudorandom.std()
```

```
(-0.005543773317243638, 1.0019601254713244)
```





Observations:

a) The path simulations and terminal distributions of Euler and Milstein schemes do not show remarkable distinctions. It means that the time discretization is fine enough that the higher-order term in Milstein does not contribute much in this case.
b) The histogram of pseudo random numbers, and the statistical mean and standard deviation show a standard normal distribution.
c) Both histograms of simulated path at maturity indicate approximate lognormal, as expected for geometric Brownian motion.

1.5 Sobol sequence with the Brownian Bridge method

1.5.1 Sobol sequence

Sobol sequence are low-discrepancy sequences. Unlike pseudo-random numbers, which mimic randomness, Sobol numbers are quasi-random, namely they follow a specific pattern to fill space more uniformly. Sobol sequences minimize discrepancy, avoiding clustering and gaps seen in pseudo-random numbers. They have better convergence properties. Its convergence rate is O($\frac{(logN)^d}{N}$), which is faster than O($\frac{1}{\sqrt{N}}$) for pseudo-random.

It is built by using direction numbers and bitwise XOR operations on binary fractions.

Direction Numbers:$v_k^{(j)}$, where j is dimension, j=1, 2, … ,d.

They are chosen based on primitive polynomials over GF(2). They also define how bits contribute to the Sobol numbers.

Generating the sequence: for the n-th point in dimension j,

compute: $x_n^{(j)} = n_1 v_1^{(j)} \oplus n_2 v_2^{(j)} \oplus ... \oplus n_k v_k^{(j)} \oplus ...$

where $n = n_1 n_2 n_3$ ...is the binary representation of n, and $\oplus$ denotes the bitwise XOR operation.

The result is a sequence of points $\left\{x_n = \left(x_n^{(1)}, x_n^{(2)}, \dots, x_n^{(d)}\right)\right\}$ in $[0,1)^d$, where each coordinate is a binary fraction.

Sobol sequence scramble the sequences and use Power-of-2 Sample Sizes: $N = 2^m$ for best results.

1.5.2 Brownian Bridge method

The Brownian bridge is a continuous-time stochastic process derived from Brownian motion by conditioning on already computed points. Instead of simulating a Brownian motion (Wiener process) path incrementally from start to end, the Brownian Bridge constructs the path recursively and conditionally by filling in the intermediate values.

Standard Brownian motion X_t at discrete time points $t_i = i\delta t$, where $\delta t = \frac{T}{n}$, i= 0, 1, 2, …, n

$X_{t_{i+1}} = X_{t_i} + \sqrt{t_{i+1} - t_i} \cdot \phi_i$, namely $dX_t = \sqrt{\delta t}\phi_i$

Where $X_{t_0} = 0$

$\phi_i \sim N(0,1)$, are independent standard normal variates

The increments $X_{t_{i+1}} - X_{t_i}$ are normally distributed

Brownian bridge path construction

The Brownian bridge first fixes the final value, and then recursively fills in the intermediate points.

Final value: $X_T = \sqrt{T} \cdot \phi$, $\phi \sim N(0, 1)$ (3)

Recursive refinement:

For any intermediate time t_j between t_i and t_k ($t_i < t_j < t_k$), the value of X_{t_j} is conditioned on X_{t_i} and X_{t_k}:

Conditional mean:

$$\mathbb{E}\left[X_{t_j} \middle| X_{t_i}, X_{t_k}\right] = X_{t_i} + \frac{t_j - t_i}{t_k - t_i}\left(X_{t_k} - X_{t_i}\right)$$

Conditional variance:

$$\mathbb{V}\left[X_{t_j} \middle| X_{t_i}, X_{t_k}\right] = \frac{(t_j - t_i)(t_k - t_j)}{t_k - t_i}$$

$$X_{t_j} = \mathbb{E}\left[X_{t_j} \middle| X_{t_i}, X_{t_k}\right] + \sqrt{\mathbb{V}\left[X_{t_j} \middle| X_{t_i}, X_{t_k}\right]} \cdot \phi_j \qquad (4)$$

where $\phi_j \sim N(0,1)$

This method assumes the underlying process is Gaussian. Its variance reduction is achieved by conditioning on already known points. It works well with quasi-random numbers (e.g. Sobol).

1.5.3 Path simulation

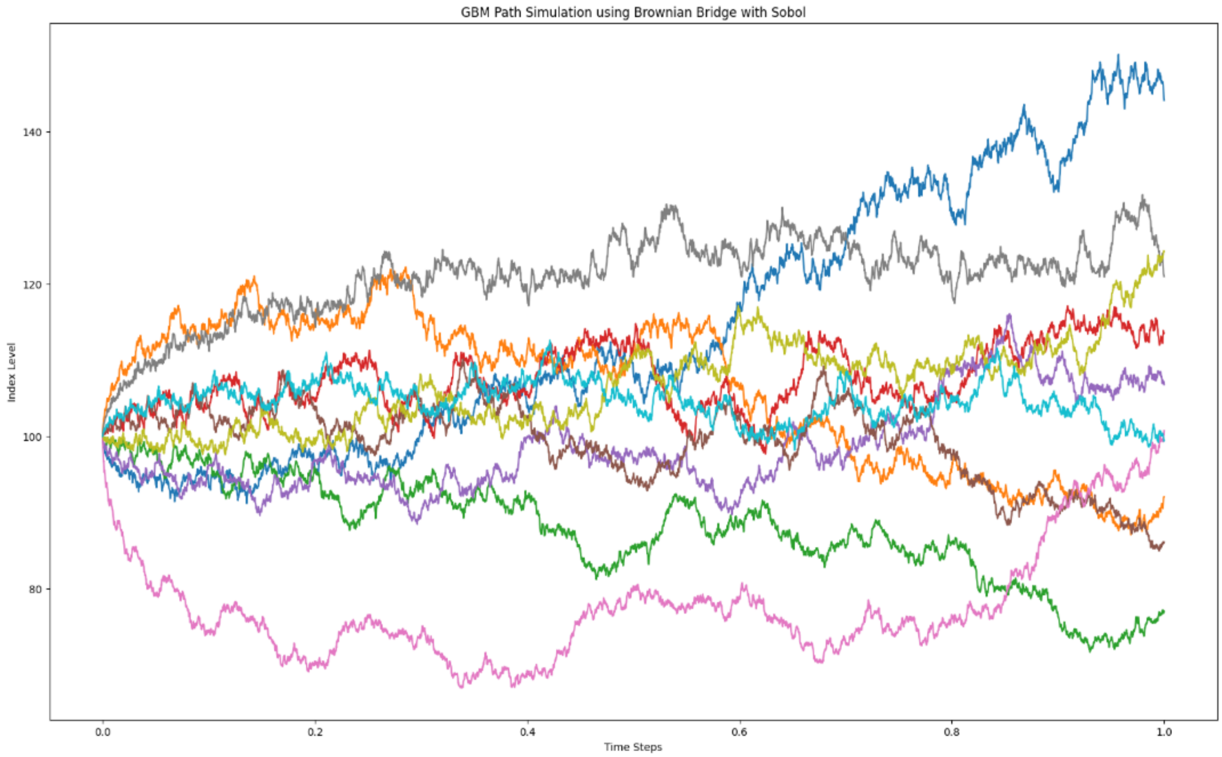
For implementations of Brownian bridge with Sobol, the iterative method is chosen, instead of recursive one, since iterative method is more robust and it shows the dyadic construction sequence more explicit.

It uses formula (3) to calculate terminal values, and uses formula (4) to calculate Brownian bridge process. ϕ_j in both formulae are generated by Sobol sequences.

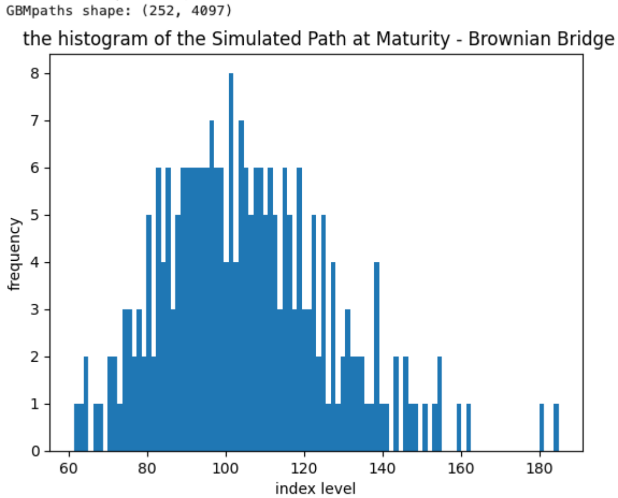
Based on geometric Brownian motion: $\frac{dS_t}{S_t} = \mu dt + \sigma dX_t$,

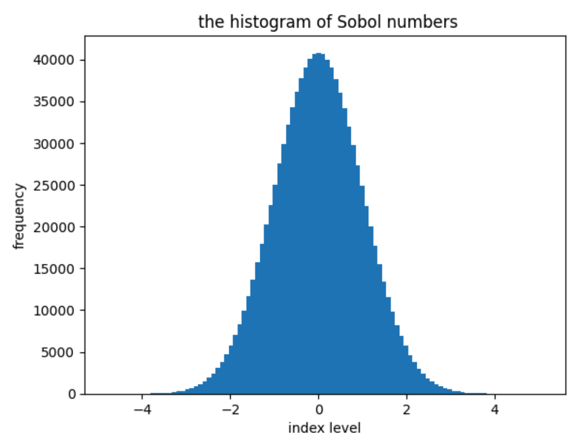
$$S_t = S_0 e^{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma X_t} \qquad (5)$$

It uses formula (5) to simulate path S_t, where X_t is the Brownian bridge process.




```
sobolseq_z shape: (252, 4096)
BrownianBridgePaths shape: (252, 4097)
drift shape: (1, 4097)
diffusion shape: (252, 4097)
GBMpaths shape: (252, 4097)
```






```
sobol_z.mean(), sobol_z.std()
```

```
(2.6132861627827364e-06, 1.0000899454729617)
```

Observations:

a) The simulated paths show less erratic and explosive in early steps, namely they are more guided toward the known endpoint. Because the Brownian bridge spreads variance more evenly across the path, which leads to better convergence in Monte Carlo simulations, especially when combined with Sobol sequences.

b) The Sobol sequence shows good mean, quite close to zero. It also shows a tighter clustering around mean, consistent with its features of low-discrepancy.

c) The histogram at maturity displays a slightly different shape, with more jagged and not smooth, resulted from a smaller number of samples.

2.Hedge with actual volatility & hedge with implied volatility

2.1 Hedge with actual volatility

2.1.1 Mathematical derivatives of the known total P&L

We need to construct a portfolio and calculate the mark-to market profits.

Assumption:

a) the option is vanilla call option: $V_i(S,t;\ \sigma_i)$, $V_a(S,t;\sigma_a)$

b) the underlying asset follows geometric Brownian motion: $dS = \mu S dt + \sigma_a S dX_t$

c) assets follow risk-neutral drift: $\mu = r$

d) no dividends D=0

the portfolio is set up by buying one unit of option V_i and short Δ_a units of stocks.

At time t, the hedging portfolio:

a) option: V_i

b) stocks: $-\Delta_a S$

c) cash flows: $-V_i + \Delta_a S$

At time $t + dt$, the portfolio value becomes:

a) option: $V_i + dV_i$

b) stocks: $-\Delta_a S - \Delta_a dS$

c) cash flows: $(-V_i + \Delta_a S)(1 + rdt)$

so, the mark-to-market profit over time dt:

$dV_i - \Delta_a dS + (-V_i + \Delta_a S) rdt$

$= dV_i - \Delta_a dS - r(V_i - \Delta_a S)dt$ (6)

The option would be correctly valued at V_a, and then we have:

$dV_a - \Delta_a dS - r(V_a - \Delta_a S)dt = 0$ (7)

(6)-(7): $dV_i - dV_a - r(V_i - \mathrm{V}_a)dt$

$= e^{rt} d(e^{-rt}(V_i - \mathrm{V}_a))$

This is the mark-to-market profit from time t to $t + dt$.

The present value of the profit at time t_0:

$e^{-r(t-t_0)} e^{rt} d(e^{-rt}(V_i - \mathrm{V}_a))$

$= e^{rt_0} d(e^{-rt}(V_i - \mathrm{V}_a))$

Then, the total profit from t_0 to maturity T is calculated by $e^{rt_0} \int_{t_0}^{T} d\,(e^{-rt}(V_i - \mathrm{V}_a))$ (8)

Considering boundary condition: at expiration T, $V_i(T) - \mathrm{V}_a(T) = 0$

So, for the formula (8), the final result is $V_a(0) - \mathrm{V}_i(0)$

Conclusion: if we hedge with actual volatility, the total profit is a guaranteed amount, equal to the difference of option values, calculated by Black Scholes model with actual volatility and implied volatility respectively.

In order to do simulation, we use Itô-lemma to write the mark-to-market profit over time step.

Starting from formula (6): $dV_i - \Delta_a dS - r(V_i - \Delta_a S)dt$

$dV_i = \frac{\partial V_i}{\partial t} dt + \frac{\partial V_i}{\partial S} dS + \frac{1}{2} \frac{\partial^2 V_i}{\partial S^2} dS^2$

$= \Theta_i dt + \Delta_i dS + \frac{1}{2} \sigma_a^2 S^2 \Gamma_i dt \quad (dS^2 = \sigma_a^2 S^2 dt)$ (9)

Substitute (9) into (6): $\Theta_i dt + \Delta_i dS + \frac{1}{2} \sigma_a^2 S^2 \Gamma_i dt - \Delta_a dS - r(V_i - \Delta_a S)dt$

$= \Theta_i dt + (\Delta_i - \Delta_a)dS + \frac{1}{2} \sigma_a^2 S^2 \Gamma_i dt - r(V_i - \Delta_a S)dt$ (10)

Since, $dS = \mu S dt + \sigma_a S dX_t$,

the Black Scholes model: $\Theta_i + \frac{1}{2} \sigma_i^2 S^2 \Gamma_i + rs\Delta_i - rV_i = 0$

substitute these two formulae into (10):

$= \frac{1}{2}(\sigma_a^2 - \sigma_i^2) S^2 \Gamma_i dt + (\Delta_i - \Delta_a)[\sigma_a S dX_t + (\mu - r)S dt]$

Since $\mu = r$

$= \frac{1}{2}(\sigma_a^2 - \sigma_i^2) S^2 \Gamma_i dt + (\Delta_i - \Delta_a)\sigma_a S dX_t$

The formula shows the diffusion term, which means that how the profit is achieved is random.

2.1.2 Simulations of P&L

If hedging with actual volatility, the mark-to-market profit over a time step dt is calculated with formula $\frac{1}{2}(\sigma_a^2 - \sigma_i^2)S^2\Gamma_i dt + (\Delta_i - \Delta_a)\sigma_a S dX_t$

Based on the formula of Black Scholes, for vanilla call option, with no dividend, D=0:

$\Delta = N(d_1)$, where N is the cumulative distribution function of standard normal distribution

$N(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{\frac{-s^2}{2}} ds$

$d_1 = \frac{ln\frac{S}{E} + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}$

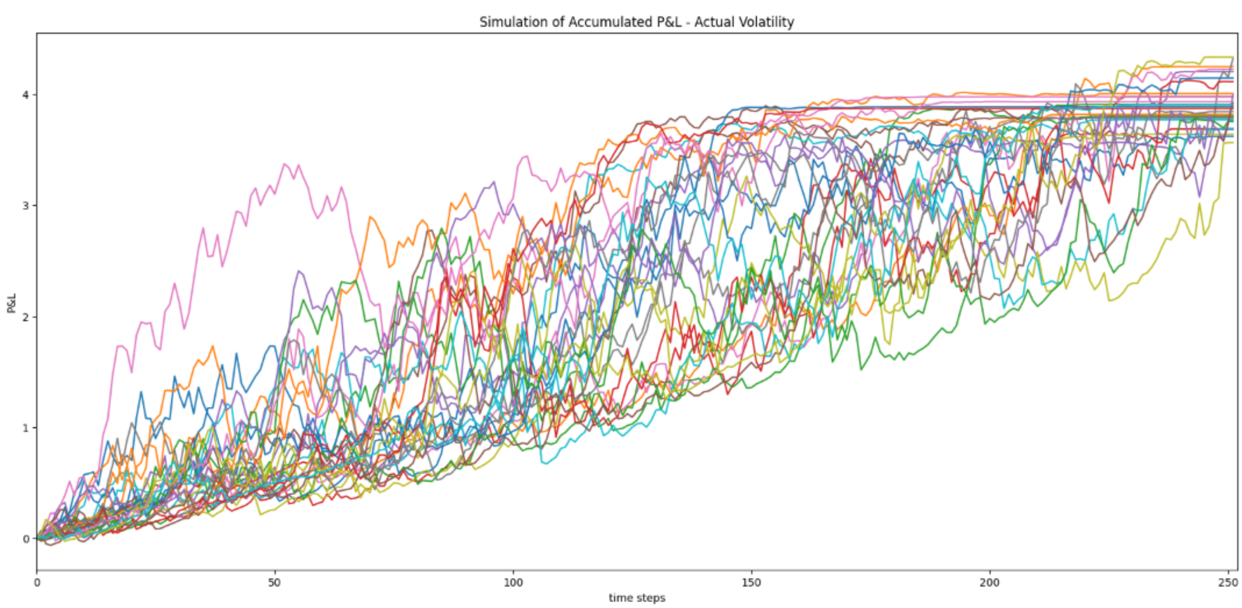
$\Gamma = \frac{N'(d_1)}{\sigma S\sqrt{T-t}}$,

where N' is the probability density function of the standard normal distribution

$N'(d_1) = \frac{e^{\frac{-1}{2}d_1^2}}{\sqrt{2\pi}}$

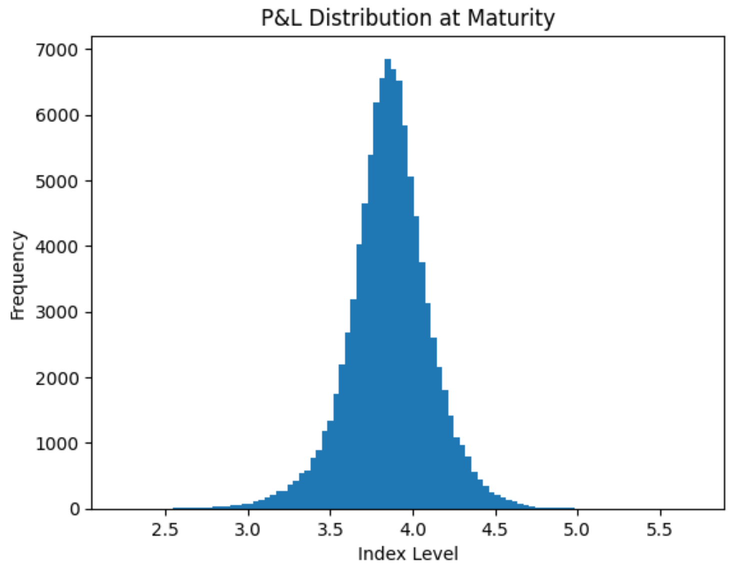
$dX_t = \sqrt{\delta t}\phi_i$, $\phi_i \sim N(0,1)$

The simulated paths show as follow:



Scenario: S_0=100, strike=100, rate=0.05, actual volatility= 0.3, implied volatility= 0.2, time period=1, time steps=252

We can see that the paths of the profit show random fluctuation, but trends are upward and the final profits converge.




```
PandL[-1,:].mean(), PandL[-1,:].std()
```

```
(3.8577294305214482, 0.2515301515573105)
```

The distribution of the profit at maturity shows clear. Its mean is 3.857729, which means the expectation of final profit by delta hedging with actual volatility is 3.857729.

Also, the calculation shows that $V_a(0) - \mathrm{V}_i(0) = 3.780671$, according to the Black Scholes formula.

2.2 Hedge with implied volatility

2.2.1 Mathematical derivatives of the uncertain path-dependent total P&L

The mark-to-market profit over time step dt is $dV_i - \Delta_i dS - r(V_i - \Delta_i S)dt$ (11)

using Itô-lemma $dV_i = \frac{\partial V_i}{\partial t}dt + \frac{\partial V_i}{\partial S}dS + \frac{1}{2}\frac{\partial^2 V_i}{\partial S^2}dS^2$

since, $dS = \mu S dt + \sigma_a S dX_t$, get $dS^2 = \sigma_a^2 S^2 dt$

so, $dV_i = \Theta_i dt + \Delta_i dS + \frac{1}{2}\sigma_a^2 S^2 \Gamma_i dt$

substitute into formula (11), get $\Theta_i dt + \frac{1}{2}\sigma_a^2 S^2 \Gamma_i dt - r(V_i - \Delta_i S)dt$ (12)

the Black Scholes formula is $\Theta_i + \frac{1}{2}\sigma_i^2 S^2 \Gamma_i + rs\Delta_i - rV_i = 0$

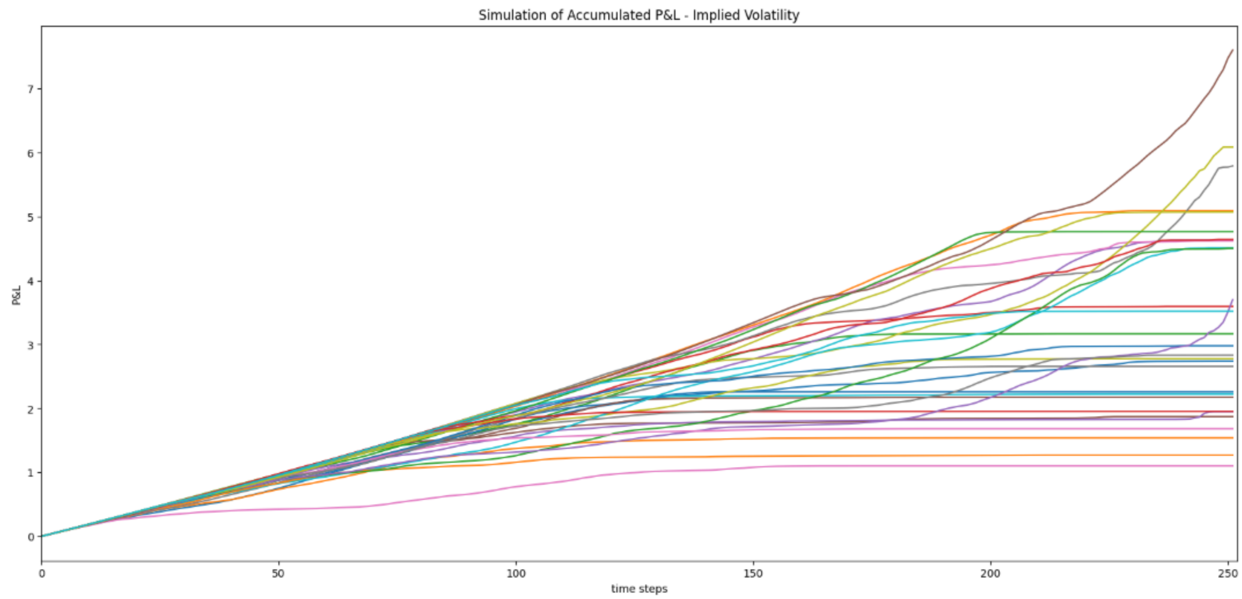
substitute into formula (12), and get the mark-to-market profit over time step dt is $\frac{1}{2}(\sigma_a^2 - \sigma_i^2)S^2\Gamma_i dt$

So, the accumulated profit at present time is $\frac{1}{2}(\sigma_a^2 - \sigma_i^2)\int_{t_0}^{T} e^{-r(t-t_0)}\, S^2 \Gamma_i dt$

If the forecast $\sigma_a^2 > \sigma_i^2$ is correct, the profit obtained by hedging with implied volatility is always positive, but it is path-dependent.

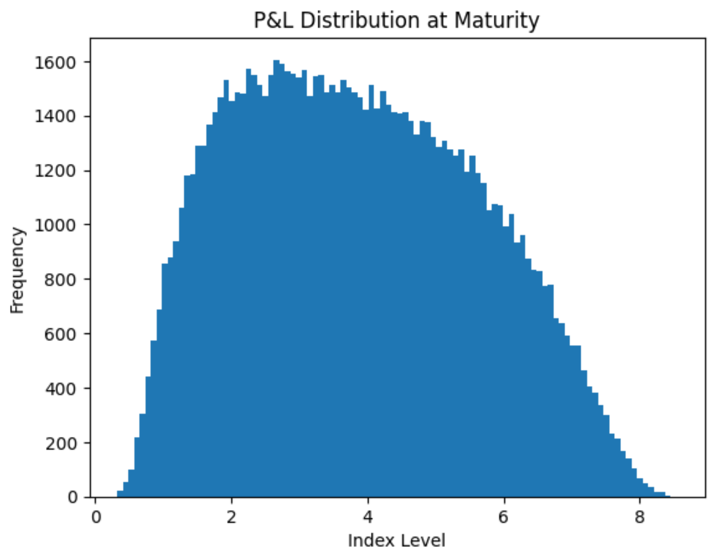
2.2.2 Simulations of P&L

Using the mark-to-market profit over time step dt : $\frac{1}{2}\ (\sigma_a^2 - \sigma_i^2) S^2 \Gamma_i dt$ to do simulations. The simulated paths show as follow:



Scenario: S_0=100, strike=100, rate=0.05, actual volatility= 0.3, implied volatility= 0.2, time period=1, time steps=252

The distribution of profits at maturity shows more erratic with higher deviation.

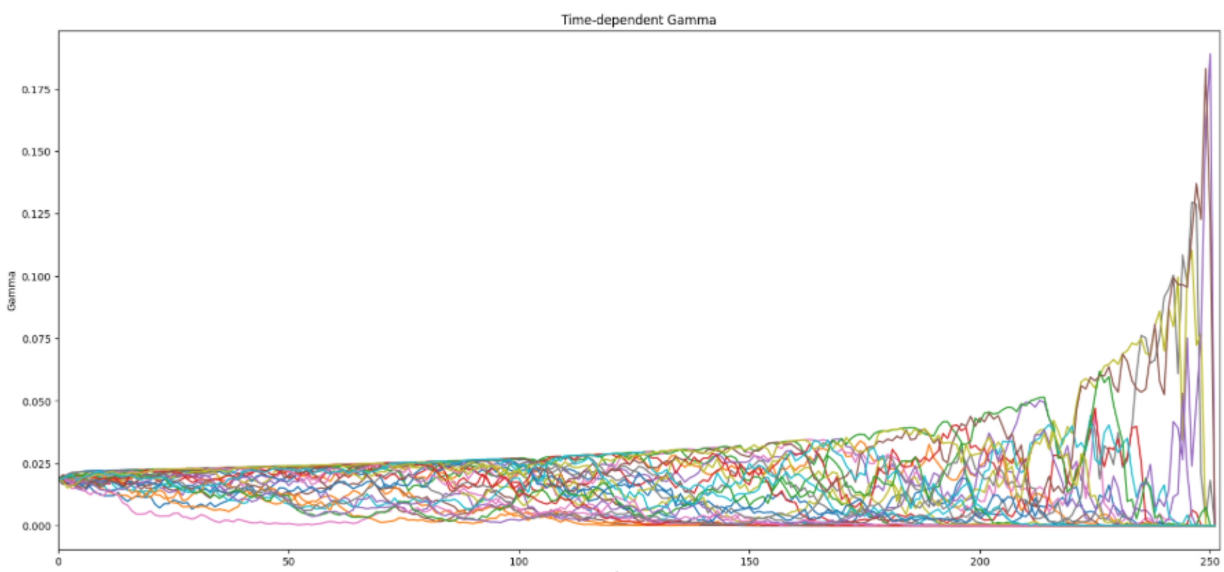



```
PandL[-1,:].mean(), PandL[-1,:].std()
```

```
(3.8572020270436904, 1.757039616629354)
```

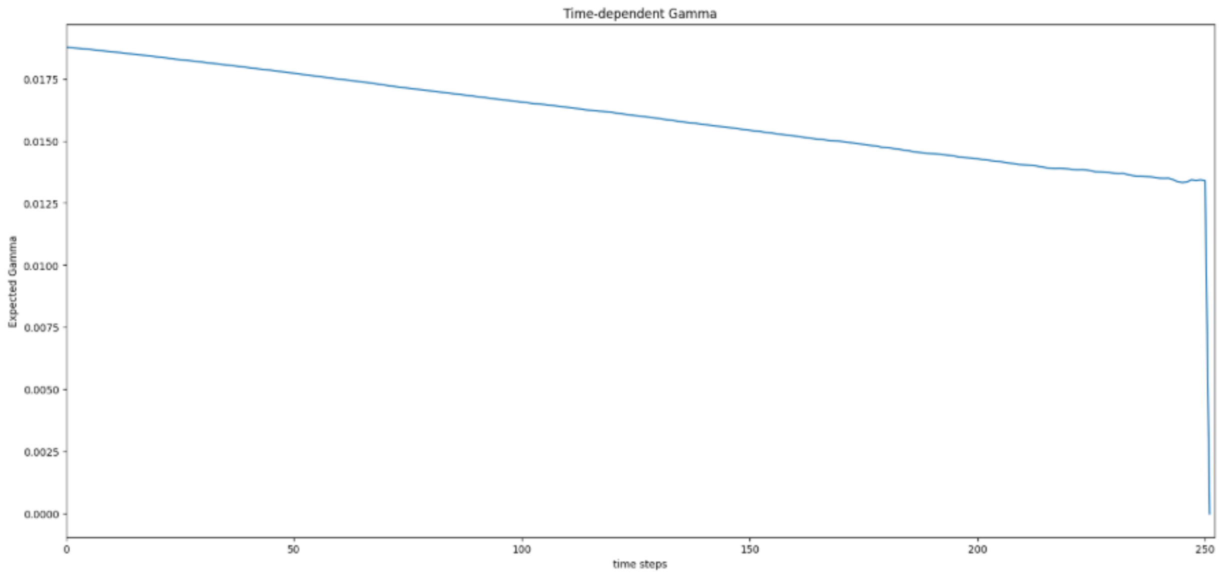
2.3 Time-dependent Gamma Analysis

2.3.1 the Impact of Gamma for Hedging with Actual Volatility

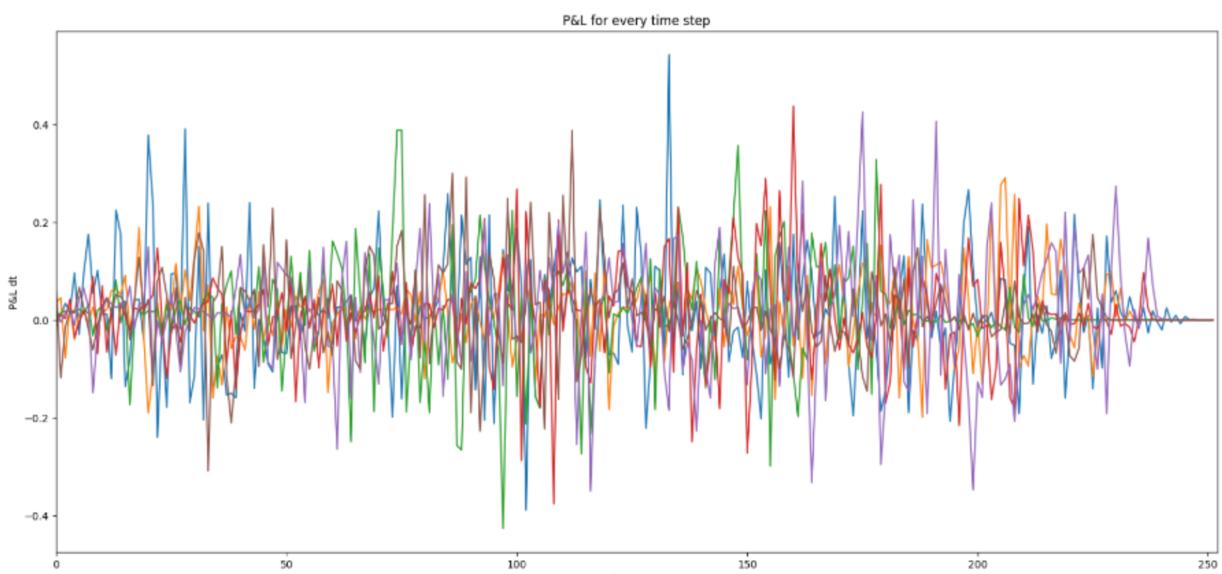


Scenario: S_0=100, strike=100, rate=0.05, actual volatility= 0.3, implied volatility= 0.2, time period=1, time steps=252

It shows the initial 30 paths of gamma evolution. These paths show gamma increasing as time steps approach expiration. It accelerates more rapidly as expiration approaches, and shows high variability near the maturity. The simulated paths that show the highest terminal gamma are likely those where the price is very close to the strike at expiration.

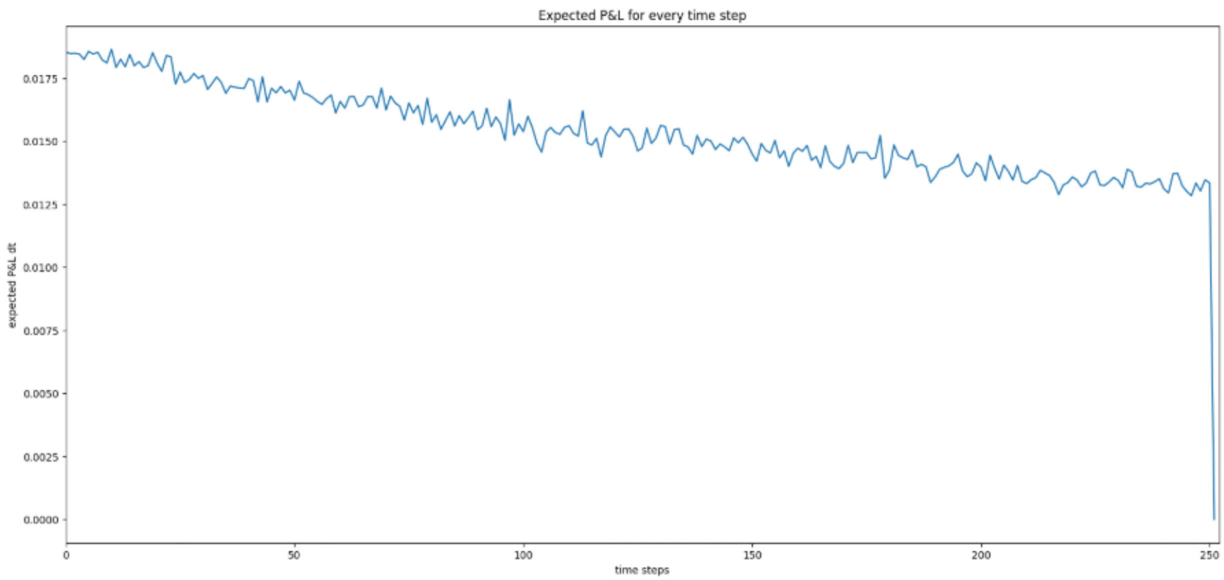


It shows the expectation of 100000 paths of gamma evolution. It decreases as time to expiration declines, and there is the noise, a sharp drop, around the maturity. The curve appears mostly smooth, and has some small irregularities.

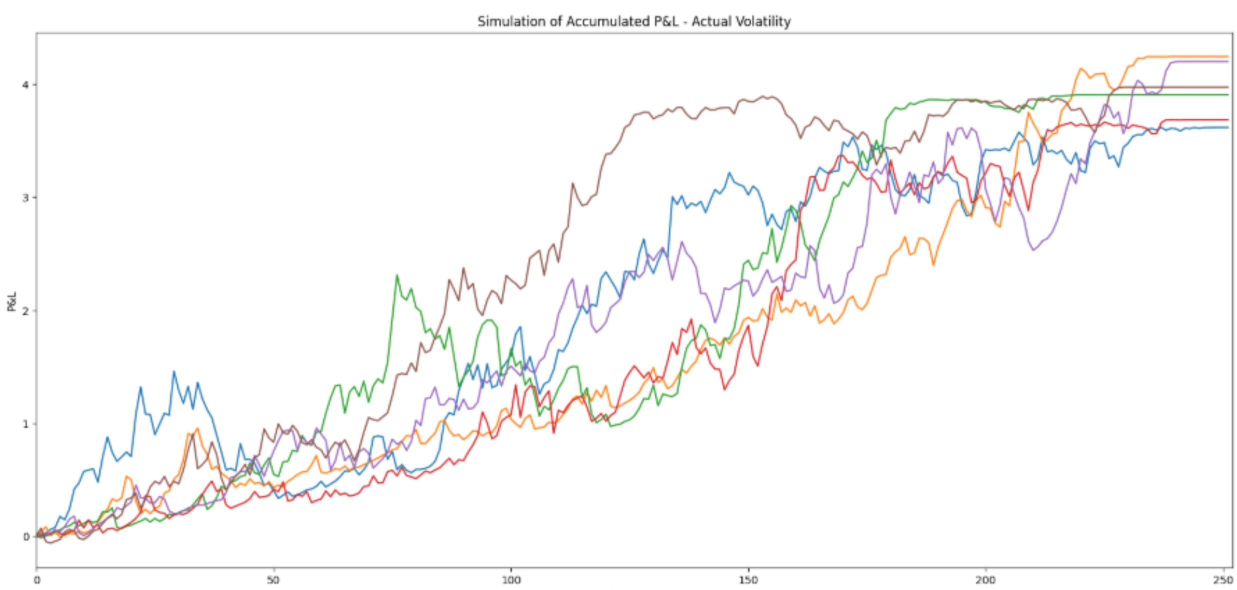


Scenario: S_0=100, strike=100, rate=0.05, actual volatility= 0.3, implied volatility= 0.2, time period=1, time steps=252

It shows initial 6 paths of mark-to-market profit for every time step. It shows significant volatility in the individual P&L path for every time step. It is consistent with the feature of hedging with actual volatility that the path to achieve the final profit is random.

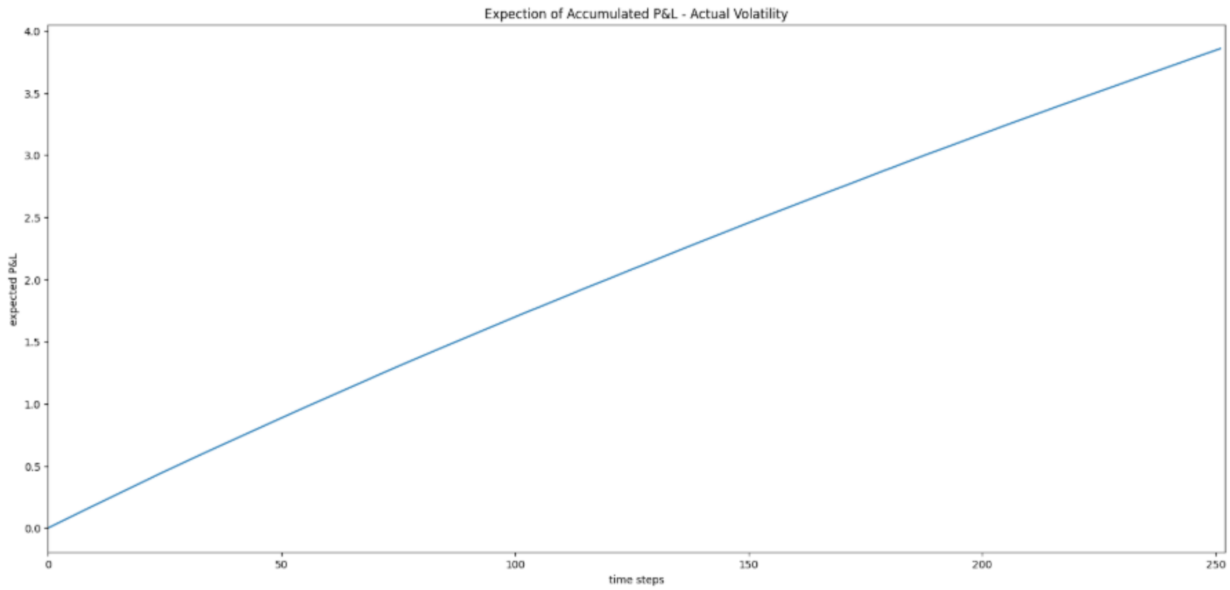


It shows the expectation of P&L for every time step. It also goes down gradually with the decrease of time to expiration, and display the noise around the maturity. This trend is comparable to the expectation of gamma evolution, although it is with a series of zigzags whereas the gamma evolution appears mostly smooth.

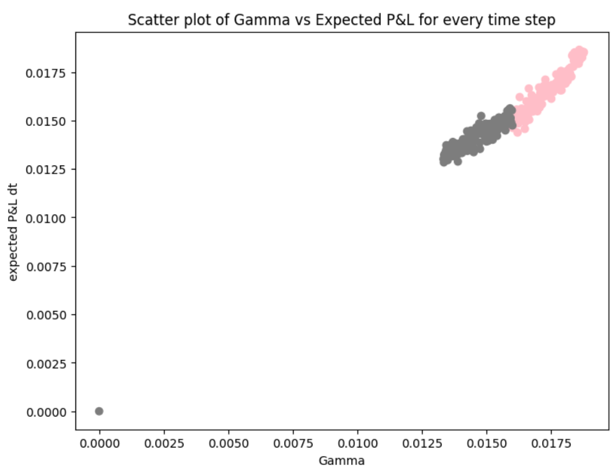


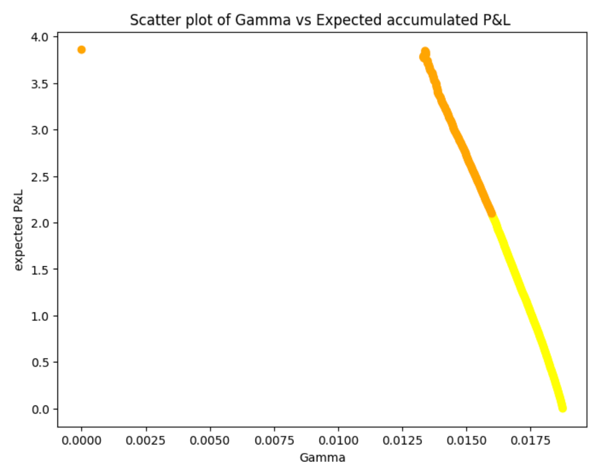
Scenario: S_0=100, strike=100, rate=0.05, actual volatility= 0.3, implied volatility= 0.2, time period=1, time steps=252

It shows the initial 6 simulated paths of accumulated P&L.



It shows the expectation of accumulated P&L. It is smooth and upward, reflecting that hedging with actual volatility provides the guaranteed final profit.





Graphs above show the scatter plot of expectation of gamma evolution against expectation of P&L for time step and expectation of accumulated P&L.

Pink and yellow indicate first half-time steps, and grey and orange show second half time steps.

The relationship between gamma evolution and P&L for time step is positive, and the trajectory declines as the expiration approaches, which are consistent with their tendencies respectively showed previously. Additionally, their correlated coefficient is 0.976933.

The relationship between gamma evolution and accumulated P&L is negative, and the trajectory goes up to the approximate 3.8 as the expiration approaches.it matches their tendencies respectively showed previously. And their correlated coefficient is -0.899201.

Finally, they are the regression analysis.

P&L for every time step = a + b * Γ_t +ε

Accumulated P&L = a + b * Γ_t +ε

ε is residual term

The following list the results:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.954
Model:                            OLS   Adj. R-squared:                  0.954
Method:                 Least Squares   F-statistic:                     5232.
Date:                Sat, 26 Apr 2025   Prob (F-statistic):         1.22e-169
Time:                        04:33:54   Log-Likelihood:                 1619.3
No. Observations:                 252   AIC:                            -3235.
Df Residuals:                     250   BIC:                            -3228.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0004      0.000      1.697      0.091   -5.68e-05       0.001
x1             0.9381      0.013     72.334      0.000       0.913       0.964
==============================================================================
Omnibus:                        0.647   Durbin-Watson:                   1.404
Prob(Omnibus):                  0.724   Jarque-Bera (JB):                0.579
Skew:                          -0.117   Prob(JB):                        0.749
Kurtosis:                       2.994   Cond. No.                         524.
==============================================================================
```

the coefficient of Γ_t with P&L for every time step is 0.938, and p-value shows its significance. Also, R^2 is 0.954. These statistical values show Γ_t dominates the P&L for every time step when hedging with actual volatility.

The formula for mark-to-market profit for every time step is

$\frac{1}{2}(\sigma_a^2 - \sigma_i^2)S^2\Gamma_i dt + (\Delta_i - \Delta_a)\sigma_a S dX_t$

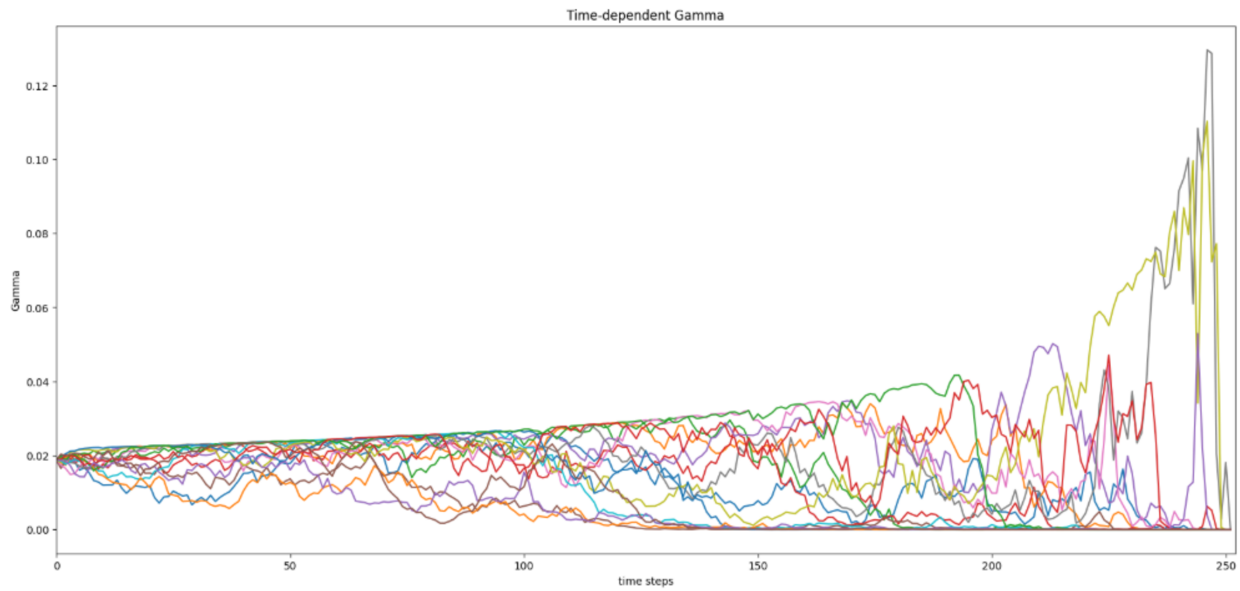
The theoretical coefficient is $\frac{1}{2}(\sigma_a^2 - \sigma_i^2)S^2 dt$ = 0.5*(0.3^2-0.2^2)*100^2*(1/252)=0.992

The coefficient from regression analysis corresponds to formula. Therefore, the results of regression support the P&L decomposition.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.809
Model:                            OLS   Adj. R-squared:                  0.808
Method:                 Least Squares   F-statistic:                     1056.
Date:                Sat, 26 Apr 2025   Prob (F-statistic):           1.00e-91
Time:                        04:38:10   Log-Likelihood:                -176.00
No. Observations:                 252   AIC:                             356.0
Df Residuals:                     250   BIC:                             363.1
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         10.3805      0.259     40.145      0.000       9.871      10.890
x1          -523.3155     16.105    -32.495      0.000    -555.034    -491.597
==============================================================================
Omnibus:                      447.113   Durbin-Watson:                   0.821
Prob(Omnibus):                  0.000   Jarque-Bera (JB):           168971.254
Skew:                          -9.585   Prob(JB):                         0.00
Kurtosis:                     128.399   Cond. No.                         524.
==============================================================================
```

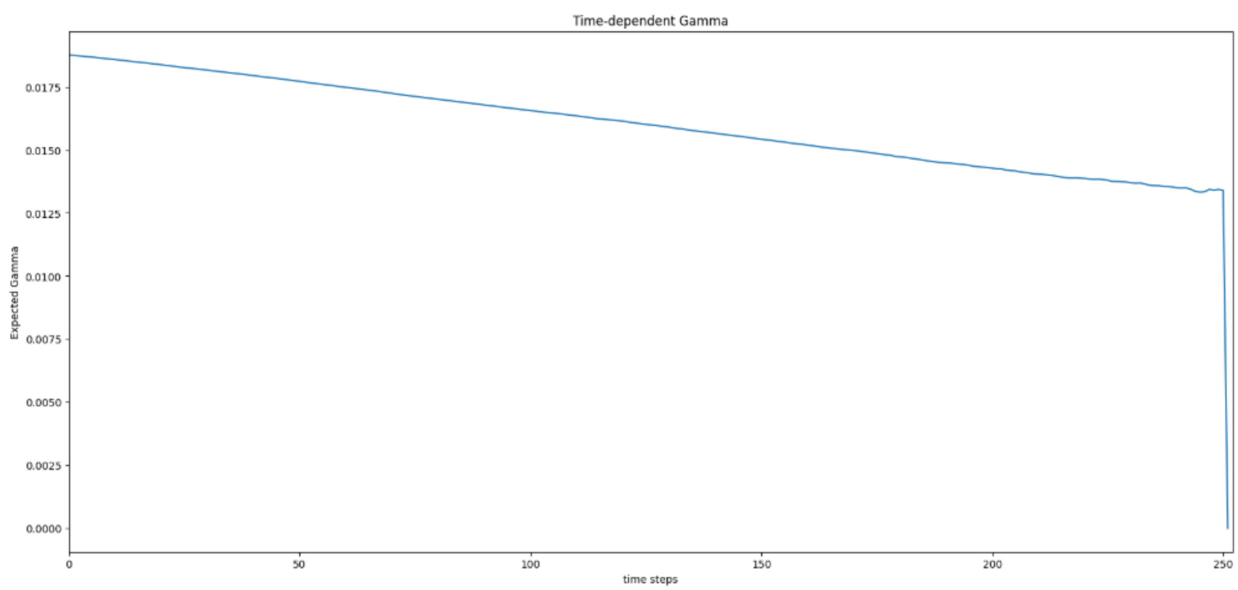
The coefficient of Γ_t with accumulated P&L is -523.316, and p-value shows its significance. Also, R^2 is 0.809. The result is compatible to the scatter plot and the correlated coefficient. However, given the DW test result, it is not a proper regression. Conclusion: according to the analysis above, P&L for every time step determined by gamma exposure when hedging with actual volatility.

2.3.2 the Impact of Gamma for Hedging with Implied Volatility

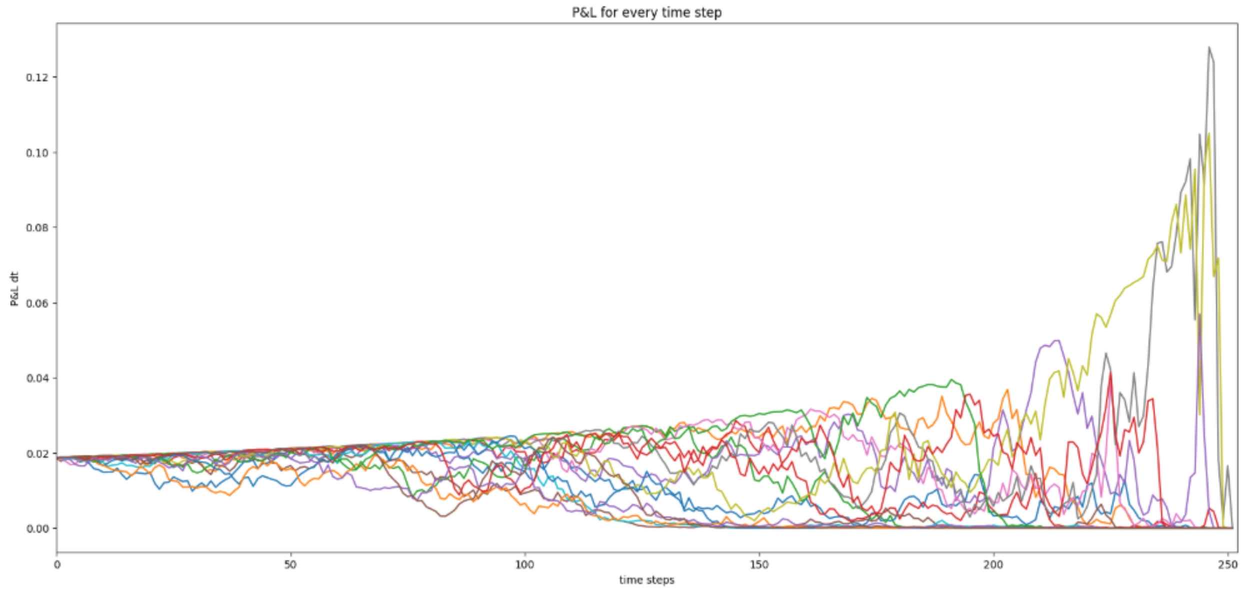


Scenario: S_0=100, strike=100, rate=0.05, actual volatility= 0.3, implied volatility= 0.2, time period=1, time steps=252

It shows the initial 16 paths of gamma evolution. Its shape is comparable to that hedging with actual volatility.

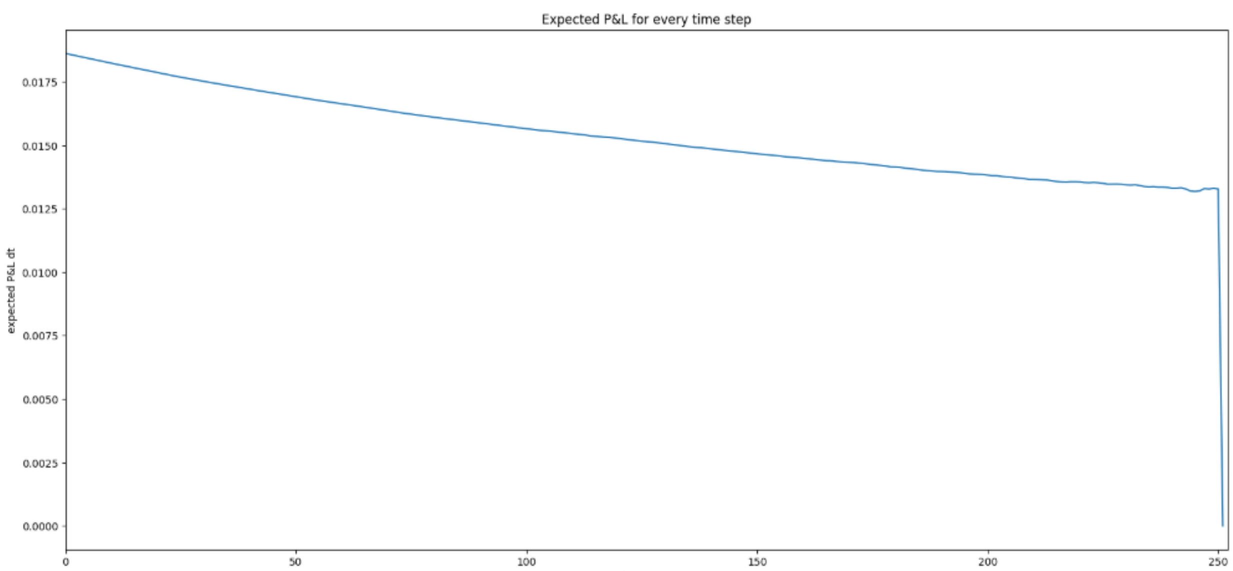


It shows the expectation of 100000 paths of gamma evolution. Its shape is comparable to that when hedging with actual volatility.



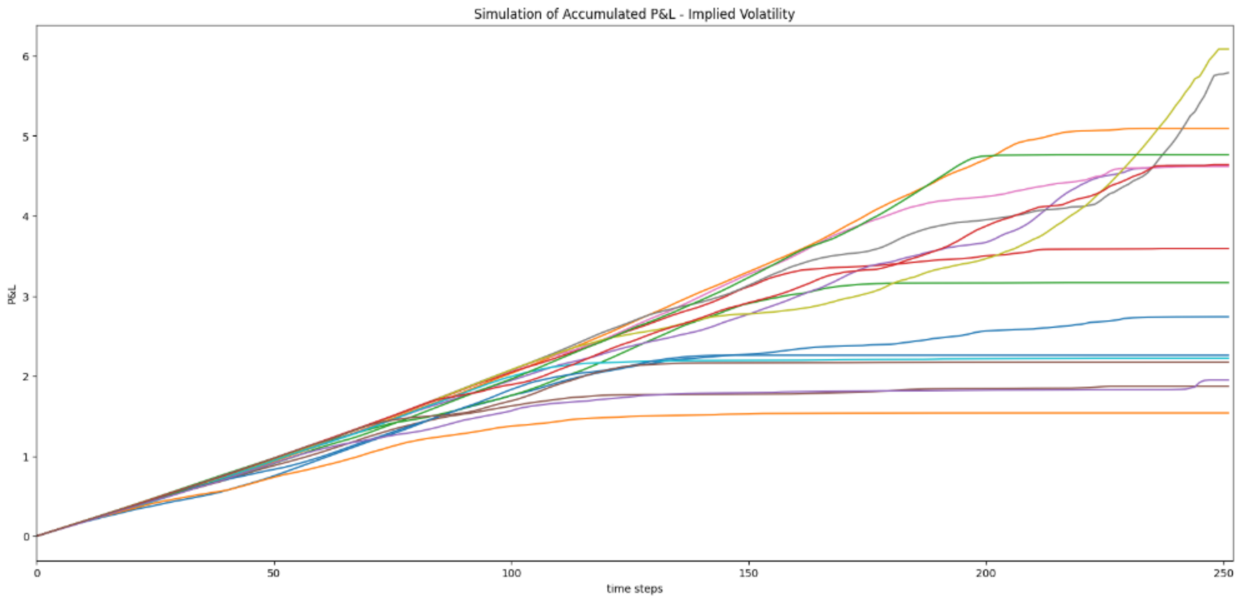
Scenario: S_0=100, strike=100, rate=0.05, actual volatility= 0.3, implied volatility= 0.2, time period=1, time steps=252

It shows initial 16 paths of mark-to-market profit for every time step. Its shape is similar to paths of gamma evolution.



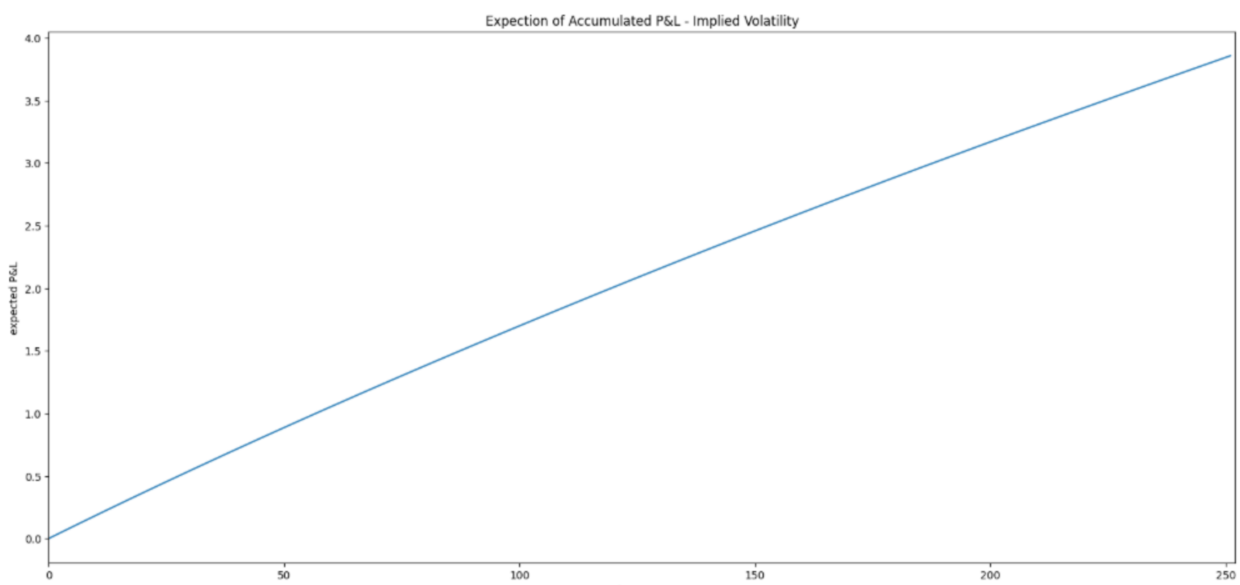
It shows the expectation of P&L for every time step. It is mostly smooth, comparable to the expectation of gamma evolution. It differs from that when hedging with actual volatility, which shows a series of zigzags.

Since when hedging with implied volatility, the formula of mark-to-market profit for every time step is $\frac{1}{2}\,(\sigma_a^2 - \sigma_i^2)S^2\Gamma_i dt$, and it only has drift term, no diffusion term. This smooth path corresponds to the formula.

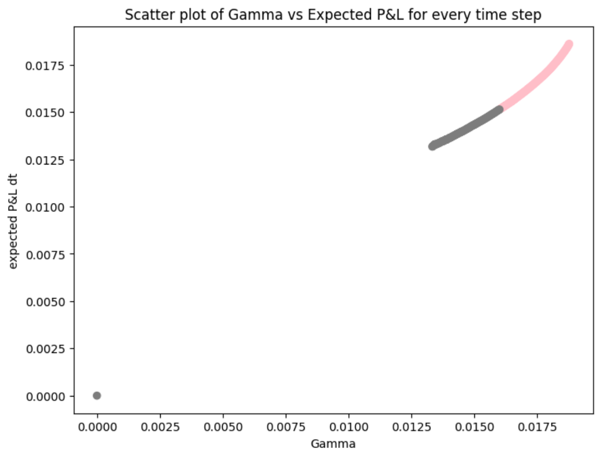


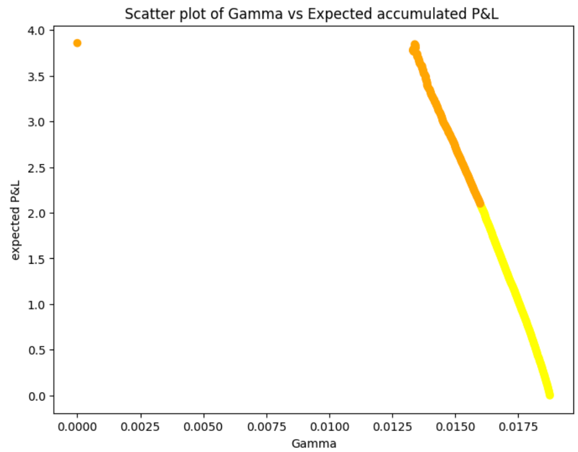
Scenario: S_0=100, strike=100, rate=0.05, actual volatility= 0.3, implied volatility= 0.2, time period=1, time steps=252

It shows the initial 16 simulated paths of accumulated P&L. These paths show accumulation steadily.



It shows the expectation of accumulated P&L. it is the same as the one when hedging with actual volatility.





Graphs above show the scatter plot of expectation of gamma evolution against expectation of P&L for time step and expectation of accumulated P&L.

Pink and yellow indicate first half-time steps, and grey and orange show second half time steps.

The scatter plot displays the smooth trajectory, rather than extended pattern when hedging with actual volatility. Since when hedging with implied volatility, theoretical formula is $\frac{1}{2}\,(\sigma_a^2 - \sigma_i^2)S^2\Gamma_i dt$, which has drift term, no diffusion term.

Finally, they are the regression analysis.

P&L for every time step = a + b * Γ_t +ε

Accumulated P&L = a + b * Γ_t +ε

ε is residual term

The following list the results:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.985
Model:                            OLS   Adj. R-squared:                  0.985
Method:                 Least Squares   F-statistic:                 1.694e+04
Date:                Sun, 27 Apr 2025   Prob (F-statistic):          1.06e-231
Time:                        00:12:33   Log-Likelihood:                 1766.6
No. Observations:                 252   AIC:                            -3529.
Df Residuals:                     250   BIC:                            -3522.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0003      0.000      2.610      0.010    7.43e-05       0.001
x1             0.9411      0.007    130.172      0.000       0.927       0.955
==============================================================================
Omnibus:                       23.153   Durbin-Watson:                   0.038
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               27.661
Skew:                           0.804   Prob(JB):                     9.85e-07
Kurtosis:                       2.782   Cond. No.                         524.
==============================================================================
```

The coefficient of Γ_t with P&L for every time step is 0.941, and p-value shows its significance. Also, R^2 is 0.985. These statistical values show Γ_t dominates the P&L for every time step when hedging with implied volatility. The theoretical coefficient is 0.992. The coefficient from regression analysis corresponds to that calculated based on formula. Since the theoretical formula is with no diffusion term, the R^2 is better than that hedging with actual volatility.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.808
Model:                            OLS   Adj. R-squared:                  0.808
Method:                 Least Squares   F-statistic:                     1055.
Date:                Sun, 27 Apr 2025   Prob (F-statistic):           1.06e-91
Time:                        00:12:39   Log-Likelihood:                -176.02
No. Observations:                 252   AIC:                             356.0
Df Residuals:                     250   BIC:                             363.1
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         10.3793      0.259     40.137      0.000       9.870      10.889
x1          -523.2219     16.106    -32.486      0.000    -554.942    -491.501
==============================================================================
Omnibus:                      447.005   Durbin-Watson:                   0.821
Prob(Omnibus):                  0.000   Jarque-Bera (JB):           168712.977
Skew:                          -9.581   Prob(JB):                         0.00
Kurtosis:                     128.302   Cond. No.                         524.
==============================================================================
```

The coefficient of Γ_t with accumulated P&L is -523.222, and p-value shows its significance. Also, R^2 is 0.808. The result is compatible to the scatter plot and the correlated coefficient. However, given the DW test result, it is not a proper regression. Conclusion: according to the analysis above, P&L for every time step determined by gamma exposure when hedging with implied volatility.

2.4 Impact of $r^2 - \sigma_{imp}^2 dt$

Starting from the Itô-lemma, $dV = \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S} dS + \frac{1}{2}\frac{\partial^2 V}{\partial S^2} dS^2$

$dV - \Delta dS = \theta dt + \frac{1}{2}\Gamma dS^2$

$d(V - \Delta S) = \theta dt + \frac{1}{2}\Gamma dS^2$ (13)

θdt: time decay

$\frac{1}{2}\Gamma dS^2$: gain from curvature

Based on Black Scholes formula: $\theta_i + \frac{1}{2}\sigma_i^2 S^2 \Gamma_i + rs\Delta - rV = 0$

$\theta_i = -\frac{1}{2}\sigma_i^2 S^2 \Gamma_i - rs\Delta + rV$

$\theta_i = -\frac{1}{2}\sigma_i^2 S^2 \Gamma_i$

Ignoring -$rs\Delta, rV$ terms, since they are relatively small compared to the gamma and theta effects, especially for short-dated options or in low interest rate environments, the volatility component dominates the theta calculation.

The formula θ_i provides connection between gamma and theta, namely convexity must be paid for through time decay.

Substitute θ_i into the formula (13): $d(V - \Delta S) = \frac{1}{2}\Gamma dS^2 - \frac{1}{2}\sigma_i^2 S^2 \Gamma_i dt$

$= \frac{1}{2}\Gamma S^2 \left[(\frac{dS}{S})^2 - \sigma_i^2 dt\right]$

$(\frac{dS}{S})^2 = r^2$

So, $d(V - \Delta S) = \frac{1}{2}\Gamma S^2 [r^2 - \sigma_i^2 dt]$

We can see the mark-to-market profit is proportional to the difference between realized and implied variance.

Accumulated P&L is $\sum_t^T \frac{1}{2}\Gamma_t S_t^2 \left[{r_t}^2 - \sigma_{t,imp}^2 dt\right]$

3. Minimum Variance Delta

3.1 Theoretical formula

According to Black Scholes theory, Δ_{BS} hedge for stock price movements, with assumption that volatility stays constant. However, the implied volatility changes associated with the underlying asset. The minimum variance delta Δ_{MV} is the hedge ratio that minimizes the variance of $dV - \Delta_{MV} dS$, and it improves upon the traditional Δ_{BS} by incorporating implied volatility movement varying with the underlying asset.

V is the option price observed from market, and based on the Black Scholes formula, we can write $V = f_{BS}(S, \sigma_{imp})$

$dV = \frac{\partial V}{\partial S} dS + \frac{\partial V}{\partial \sigma_{imp}} d\sigma_{imp} + O(dS^2)$

$dV = \Delta_{BS} dS + \upsilon_{BS} d\sigma_{imp} + \epsilon$

$dV - \Delta_{MV} dS = (\Delta_{BS} - \Delta_{MV}) dS + \upsilon_{BS} d\sigma_{imp} + \epsilon$

$\mathrm{E}[dV - \Delta_{MV} dS | dS] = (\Delta_{BS} - \Delta_{MV}) dS + \upsilon_{BS} \mathrm{E}[d\sigma_{imp} | dS] + \mathrm{E}[\epsilon | dS]$

LHS is the minimum variance of hedging error, which is equal to zero.

Therefore, $\Delta_{MV} = \Delta_{BS} + \upsilon_{BS}\frac{\mathrm{E}[d\sigma_{imp}|dS]}{dS} + \frac{\mathrm{E}[\epsilon|dS]}{ds}$

$= \Delta_{BS} + \upsilon_{BS}\frac{\mathrm{E}[d\sigma_{imp}]}{dS} + \frac{\mathrm{E}[\epsilon]}{ds}$

Where $\frac{\mathrm{E}[\epsilon]}{ds}$ is the residual term. When S follows a diffusion process, with no jumps, $dS \rightarrow 0$, then $\frac{\mathrm{E}[\epsilon]}{ds} \rightarrow 0$

Finally, $\Delta_{MV} = \Delta_{BS} + \upsilon_{BS}\frac{\mathrm{dE}[\sigma_{imp}]}{dS}$ (14)

Where $\upsilon_{BS}\frac{\mathrm{dE}[\sigma_{imp}]}{dS}$ is a correction term, reflecting volatility dynamics.

3.2 Empirical analysis formula

As for empirical analysis, we introduce an assumption of empirical structure for $\frac{\mathrm{dE}[\sigma_{imp}]}{dS}$ (Hull&White, 2017)

$$\frac{\mathrm{dE}[\sigma_{imp}]}{dS} = \frac{1}{S_t\sqrt{T}}\left(a + b\Delta_{BS} + c\Delta_{BS}^2\right)$$

Substitute into (14): $\Delta_{MV,t} = \Delta_{BS,t} + \frac{\upsilon_{BS,t}}{S_t\sqrt{T}}\left(a + b\Delta_{BS,t} + c\Delta_{BS,t}^2\right)$

Since $\mathrm{d}V \approx \Delta_{MV} dS$

So, $dV = \Delta_{BS,t}dS + \frac{dS}{S_t}\frac{\upsilon_{BS,t}}{\sqrt{T}}\left(a + b\Delta_{BS,t} + c\Delta_{BS,t}^2\right) + \varepsilon_t$

$$\left(dV - \Delta_{BS,t}dS\right)\frac{S_t}{dS}\frac{\sqrt{T}}{\upsilon_{BS,t}} = a + b\Delta_{BS,t} + c\Delta_{BS,t}^2 + \varepsilon_t$$

It is the regression formula.

We can rewrite it into two parts:

First, calculate the dependent variable part, $y = \left(dV - \Delta_{BS,t}dS\right)\frac{S_t}{dS}\frac{\sqrt{T}}{\upsilon_{BS,t}}$

then, do regression for $y = a + b\Delta_{BS,t} + c\Delta_{BS,t}^2 + \varepsilon_t$

namely, the correction term can be calculated:

$$\Delta_{MV,t} - \Delta_{BS,t} = \frac{\upsilon_{BS,t}}{S_t\sqrt{T}}\left(\hat{a} + \hat{b}\Delta_{BS,t} + \hat{c}\Delta_{BS,t}^2\right)$$

3.3 Data preparation

Since, it is challenge to obtain the data of full option chains daily from open source, I will generate synthetic option prices data based on historical SPY prices as the underlying price, VIX index as the implied volatility base, 3-month treasury bill secondary market rate as the risk-free rate[1], and then create moneyness and time-to-maturity two-dimensional bucketing system.

3.3.1 Approach to Generating Synthetic Historical Option Dataset

There are some considerations of generation of synthetic historical option price data. Initially, I created the strike grid by setting discrete strike ratio arbitrarily and use VIX index as implied volatility directly, which resulted in the binary behavior of distribution of deltas.

In order to create more balanced distributions of delta values, reflecting real-world market behaviors, volatility surface and term structure were involved. Generally, there are two methods, parametric models, such as SABR, SVI, Heston, and empirical approaches. With my preference of better theoretical control and interpretation, I incorporated the parametric volatility surface modeling to account for a skew in moneyness and term structures.

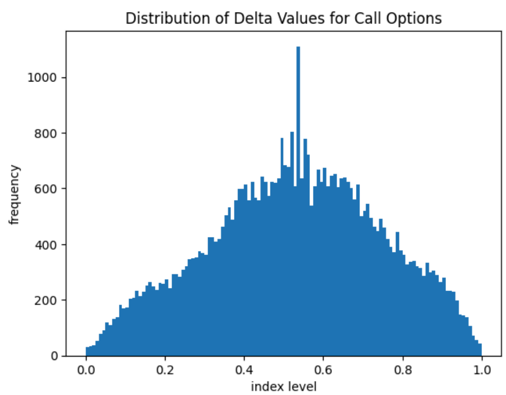
Based on my purpose, SVI model was chosen in the generation of synthetic historical option dataset, which offers the direct way to generate a smooth volatility surface, and is easier to implement.

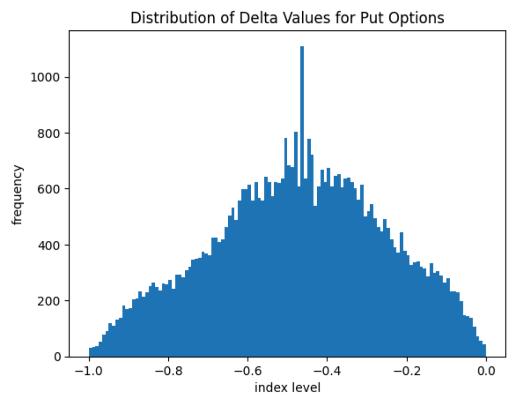
Another consideration is that the generation of the strike price grid used log-moneyness spacing with forward prices, rather than spot prices, since it is consistent with volatility surface dynamics. I have also tried to use spot prices, but it gives rise to heavily skewed distribution of delta values, particularly the long-dated options skewed toward, whereas forward-centered strike price grid can provide symmetric log-moneyness.

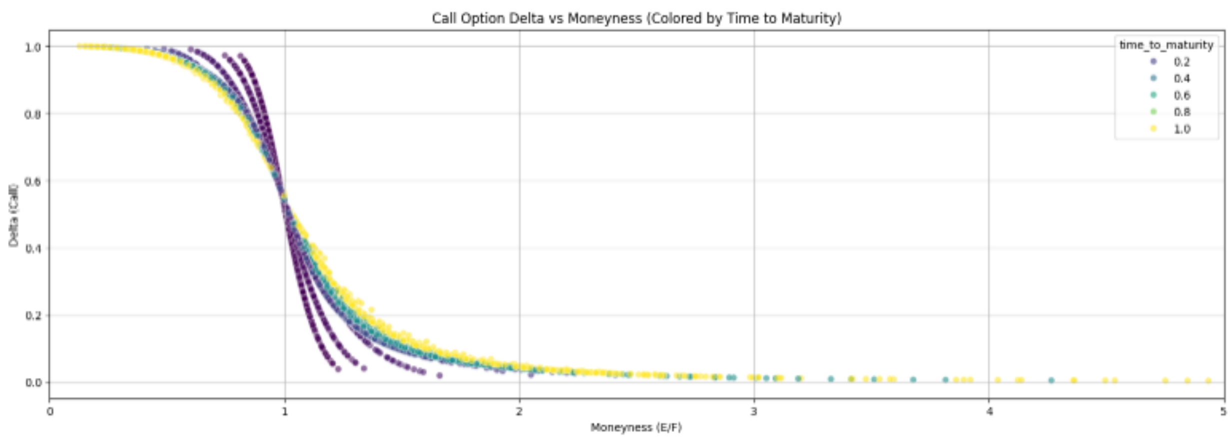
3.3.2 Statistics of the Dataset

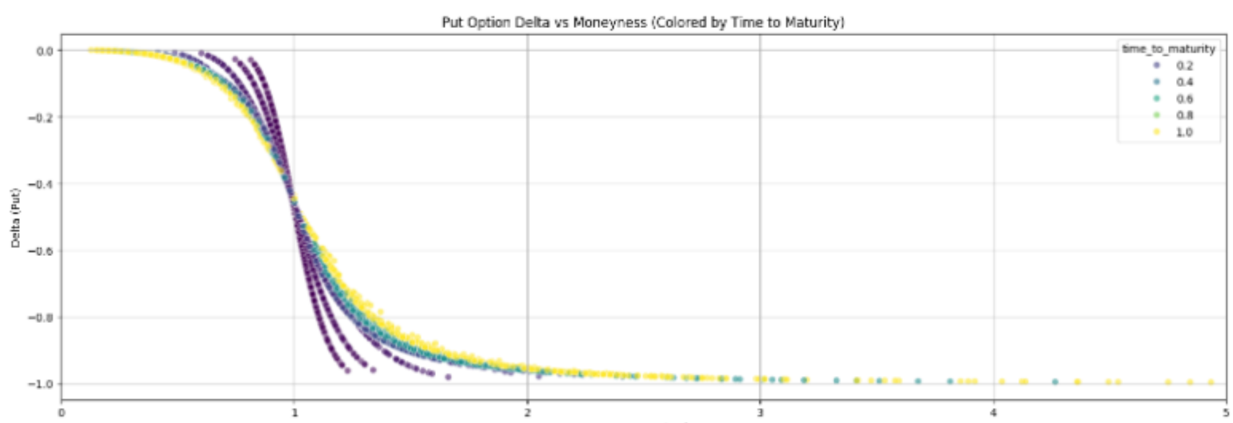
It created 95634 option data, and 47817 are call options, and 47817 are put options. The distribution of delta values and the relations of deltas and moneyness are showed as follow.

[1] Data source: pandas datareader, https://www.cboe.com/tradable_products/vix/ , https://fred.stlouisfed.org/series/DTB3









	Call options				Put options			
	Underlying price	Option price	Implied vol (SVI)	Delta	Underlying price	Option price	Implied vol (SVI)	Delta
Count	47817	47817	47817	47817	47817	47817	47817	47817

Mean	565.170	55.07	0.565	0.529	565.170	66.309	0.565	-0.471
Std	28.072	50.78	0.197	0.217	28.072	94.827	0.197	0.217
Min	496.480	0.05	0.247	0.001	496.480	0.060	0.247	-0.999
25%	544.510	23.06	0.398	0.378	544.510	22.992	0.398	-0.622
50%	563.980	37.59	0.521	0.539	563.980	37.810	0.521	-0.461
75%	590.300	68.47	0.800	0.686	590.300	71.572	0.800	-0.314
Max	612.930	432.71	0.800	0.999	612.930	3365.337	0.800	-0.001

	Call options			Put options		
	ITM	ATM	OTM	ITM	ATM	OTM
	14517(30.36%)	18432(38.55%)	14868(31.09%)	14868(31.09%)	18432(38.55%)	14517(30.36%)
Delta distribution						
Mean	0.779	0.534	0.276	-0.221	-0.466	-0.724
Std	0.091	0.071	0.110	0.091	0.071	0.110
Min	0.628	0.354	0.001	-0.372	-0.646	-0.999
25%	0.703	0.483	0.193	-0.297	-0.517	-0.807
50%	0.766	0.540	0.293	-0.234	-0.460	-0.707
75%	0.850	0.591	0.367	-0.150	-0.409	-0.633
Max	0.999	0.695	0.477	-0.001	-0.305	-0.523

Distributions of other parameters are listed in the Appendix.

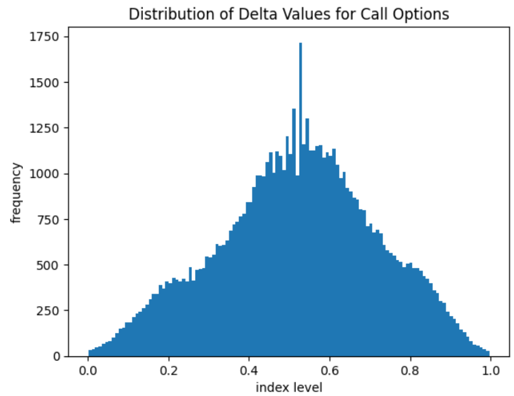
The datasets span the full deltas spectrum, with a bell-shaped, slightly skewed distribution concentrated around 0.5. The proportions of moneyness provide balanced segmentation mix, which is suitable for robust tests.

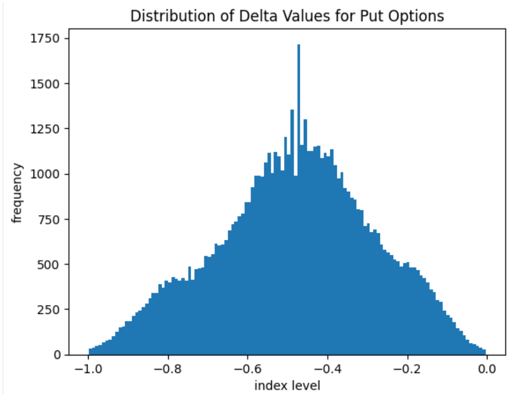
3.4 Result discussion

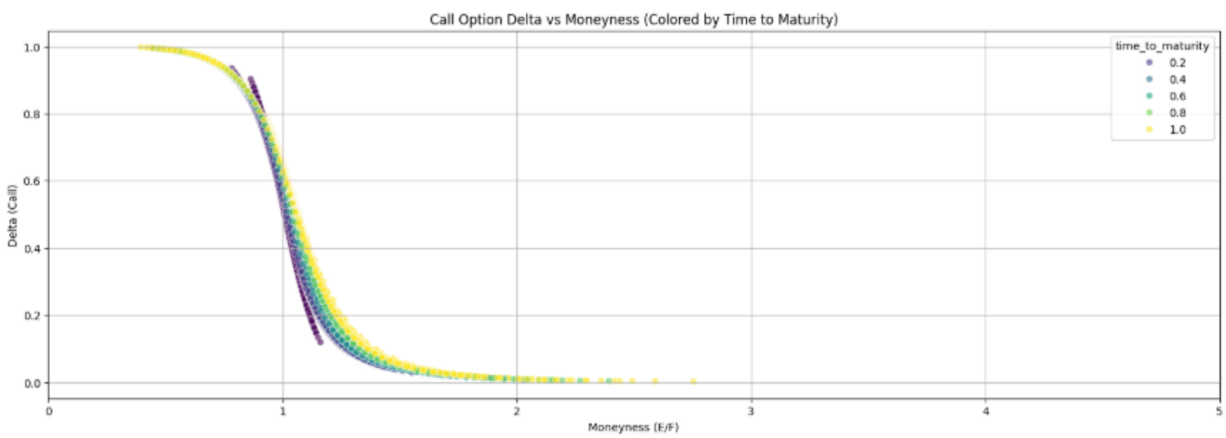
However, based on this dataset, the results have problems. Correction term and the coefficients a, b, c do not have correct magnitudes. Also, the sign of c is not consistent. Therefore, I did filtrations on dataset. Conditions are: delta [0.15, 0.85], underlying price difference >0.05, moneyness: [0.7, 1.3], implied vol_svi: [0.1, 0.6], vega [1, 60]. However, it brings another problem that it does not left sufficient data amount. Therefore, I regenerate the dataset, tuning parameters in SVI equations, term structures, and spike price generation grid.

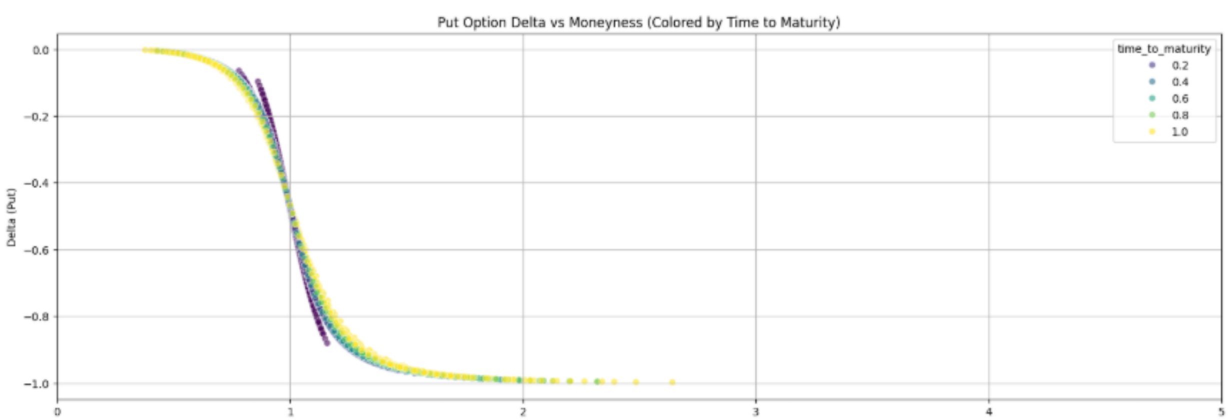
New dataset has 69069 call option data, and 69069 put option data.

The distribution of delta values and the relations of deltas and moneyness are showed as follow.









The graph of mean correction term across delta buckets shows reasonable around [0.35, 0.75] range, close to an inverted parabola-like function shape. However, the [0.75, 0.95] range may have problems.

Unfortunately, it does not solve the problem. Correction term and the coefficients a, b, c still have incorrect magnitudes.

Appendix

Original dataset

Call options

	underlying_price	strike_price	time_to_maturity	days_to_expire	implied_vol_svi	risk_free_rate	option_price	delta	vega	d1
count	47817.000000	47817.000000	47817.000000	47817.000000	47817.000000	47817.000000	47817.000000	47817.000000	47817.000000	47817.000000
mean	565.170474	11145.714676	0.294218	74.142857	0.565436	4.630158	55.070866	0.528522	80.420890	0.084291
std	28.072129	26461.478886	0.325852	82.114767	0.196795	0.429537	50.780138	0.216860	46.072254	0.661568
min	496.480000	423.746688	0.019841	5.000000	0.247131	4.170000	0.053183	0.000768	0.938507	-3.167943
25%	544.510000	703.841959	0.039683	10.000000	0.397717	4.220000	23.064583	0.377841	43.956496	-0.311157
50%	563.980000	1223.563825	0.166667	42.000000	0.521310	4.480000	37.590671	0.539163	69.392100	0.098325
75%	590.300000	4582.586198	0.500000	126.000000	0.800000	5.140000	68.467611	0.686082	101.989671	0.484776
max	612.930000	381063.617102	1.000000	252.000000	0.800000	5.260000	432.710861	0.999466	244.494329	3.271718

Put options

	underlying_price	strike_price	time_to_maturity	days_to_expire	implied_vol_svi	risk_free_rate	option_price	delta	vega	d1
count	47817.000000	47817.000000	47817.000000	47817.000000	47817.000000	47817.000000	47817.000000	47817.000000	47817.000000	47817.000000
mean	565.170474	11145.714676	0.294218	74.142857	0.565436	4.630158	66.308619	-0.471478	80.420890	0.084291
std	28.072129	26461.478886	0.325852	82.114767	0.196795	0.429537	94.827005	0.216860	46.072254	0.661568
min	496.480000	423.746688	0.019841	5.000000	0.247131	4.170000	0.059833	-0.999232	0.938507	-3.167943
25%	544.510000	703.841959	0.039683	10.000000	0.397717	4.220000	22.991970	-0.622159	43.956496	-0.311157
50%	563.980000	1223.563825	0.166667	42.000000	0.521310	4.480000	37.810137	-0.460837	69.392100	0.098325
75%	590.300000	4582.586198	0.500000	126.000000	0.800000	5.140000	71.571856	-0.313918	101.989671	0.484776
max	612.930000	381063.617102	1.000000	252.000000	0.800000	5.260000	3365.336595	-0.000534	244.494329	3.271718

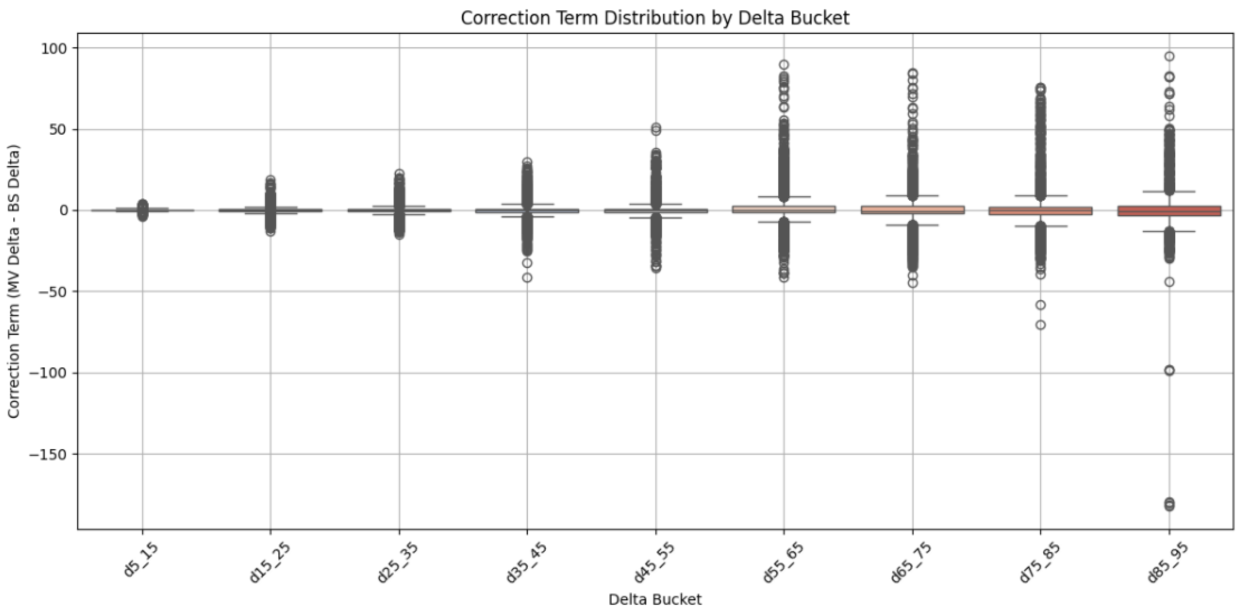
New dataset

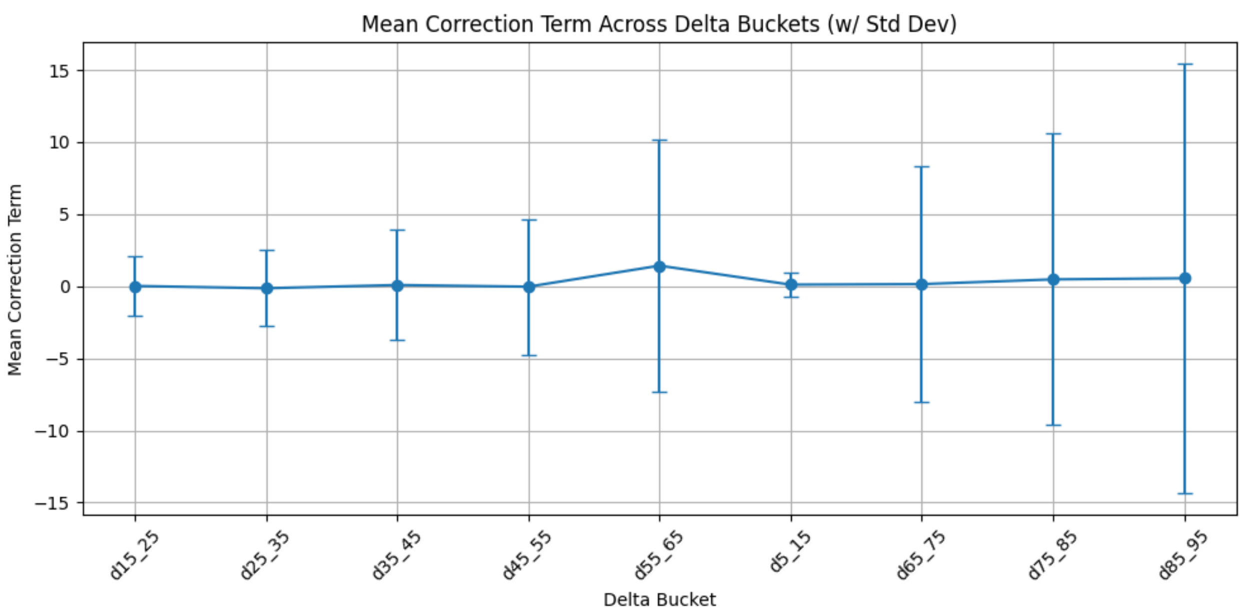
Call options

	underlying_price	strike_price	time_to_maturity	days_to_expire	implied_vol_svi	risk_free_rate	option_price	delta	vega	d1	Forward_price	moneyness
count	69069.000000	69069.000000	69069.000000	69069.000000	69069.000000	69069.000000	69069.000000	69069.000000	69069.000000	69069.000000	69069.000000	69069.000000
mean	565.170474	578.901312	0.398526	100.428571	0.325396	0.046302	37.648892	0.520498	106.247258	0.056199	565.274464	1.024278
std	28.072039	68.062184	0.337966	85.167440	0.148750	0.004295	28.853382	0.196441	51.598145	0.578624	28.071032	0.113015
min	496.480000	195.975388	0.039683	10.000000	0.156513	0.041700	0.082516	0.001946	3.070086	-2.886832	496.488255	0.394564
25%	544.510000	538.073350	0.083333	21.000000	0.193936	0.042200	19.511543	0.391663	63.548522	-0.274988	544.581199	0.967061
50%	563.980000	573.601624	0.250000	63.000000	0.276409	0.044800	30.021899	0.526744	95.560949	0.067087	564.039080	1.011390
75%	590.300000	612.354792	0.750000	189.000000	0.463400	0.051400	46.284981	0.656065	143.437968	0.401748	590.495274	1.064879
max	612.930000	1367.715790	1.000000	252.000000	0.600000	0.052600	308.721865	0.997198	244.506663	2.770091	613.188711	2.753672

Put options

	underlying_price	strike_price	time_to_maturity	days_to_expire	implied_vol_svi	risk_free_rate	option_price	delta	vega	d1
count	69069.000000	69069.000000	69069.000000	69069.000000	69069.000000	69069.000000	69069.000000	69069.000000	69069.000000	69069.000000
mean	565.170474	578.901312	0.398526	100.428571	0.325396	0.046302	40.715911	-0.479502	106.247258	0.056199
std	28.072039	68.062184	0.337966	85.167440	0.148750	0.004295	38.345551	0.196441	51.598145	0.578624
min	496.480000	195.975388	0.039683	10.000000	0.156513	0.041700	0.175535	-0.998054	3.070086	-2.886832
25%	544.510000	538.073350	0.083333	21.000000	0.193936	0.042200	19.654493	-0.608337	63.548522	-0.274988
50%	563.980000	573.601624	0.250000	63.000000	0.276409	0.044800	30.101322	-0.473256	95.560949	0.067087
75%	590.300000	612.354792	0.750000	189.000000	0.463400	0.051400	47.189100	-0.343935	143.437968	0.401748
max	612.930000	1367.715790	1.000000	252.000000	0.600000	0.052600	815.195008	-0.002802	244.506663	2.770091





	a	b	c
count	2461.000000	2461.000000	2461.000000
mean	-425.532891	882.749306	-197.242855
std	26977.267043	64103.171493	39483.003421
min	-357658.378703	-869762.525918	-463101.007519
25%	-1104.324022	-3201.156398	-4616.556170
50%	3.767346	-68.191594	129.794246
75%	666.486051	4384.999766	4047.717568
max	376705.101911	814189.913865	501972.187560