Options Analysis: Data Processing and Simulation

Definition of Call and Put Options

Call Option is a financial derivative that gives the holder the right, but not the obligation, to buy an underlying asset at a specified price (strike price) within a specified time period.
Put Option gives the holder the right, but not the obligation, to sell an underlying asset at a specified price within a specified time period.

Advantages and Disadvantages of Call and Put Options

Advantages of Call Options

- Unlimited upside potential: Gains can grow indefinitely if the underlying asset price increases significantly.
- Limited loss: The maximum loss is limited to the premium paid.
- Leverage effect: Small investment controls a larger value of the underlying asset.

Disadvantages of Call Options

- Premium cost: If the strike price is not reached, the premium may be a sunk cost.
- Expiry: The option value becomes zero after expiration.

Advantages of Put Options

- Downside protection: Offers potential gains when the underlying asset price decreases.
- Limited risk (for buyers): Maximum loss is the premium paid.

Disadvantages of Put Options

- Limited upside: Gains are limited to the extent of the underlying asset's price drop.
- Time decay: Value reduces over time.

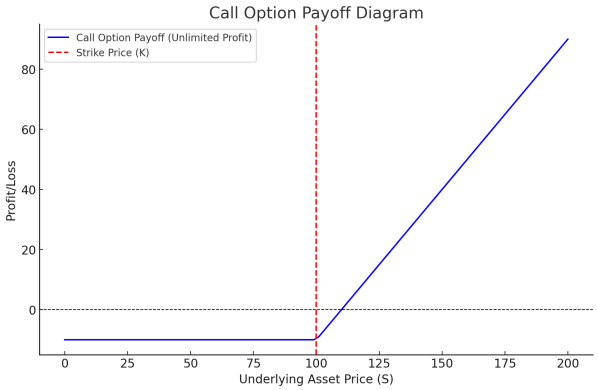

Figure 1: Enter Caption

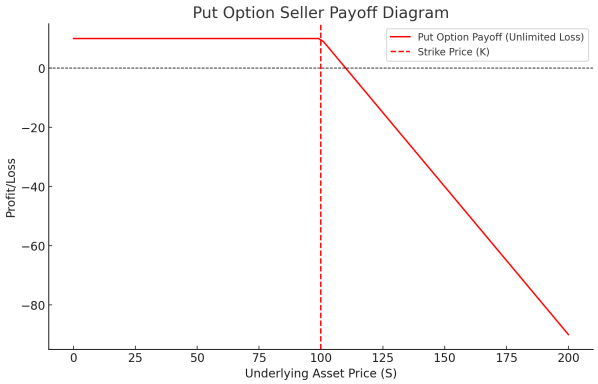

Figure 2: Enter Caption

Profit/Loss Diagrams of Call and Put Options

Call Option Profit/Loss Diagram

- If the asset price is below the strike price, the call option is worthless (out of the money).
- If the asset price is above the strike price, the call option gains intrinsic value (in the money).
- At the strike price, the call option is at the money (breakeven point).

Put Option Profit/Loss Diagram

- If the asset price is above the strike price, the put option is worthless (out of the money).
- If the asset price is below the strike price, the put option gains intrinsic value (in the money).
- At the strike price, the put option is at the money (breakeven point).

Hedging Strategies

Hedging is a risk management strategy that involves establishing a position opposite to an existing asset to mitigate potential losses.

- **Call Option Hedging**: If an investor holds an asset and fears a price drop, they can purchase put options as insurance.
- **Put Option Hedging**: If an investor has a short position, they can buy call options to limit losses.

Introduction to the Black-Scholes Model

Purpose: The Black-Scholes model is used to price European-style options, particularly call and put options on stocks.

Assumptions:

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

- S_t: Price of the underlying asset.
- μ: Drift rate.
- σ: Volatility.
- W_t: Brownian motion.

Mathematical Formulas:

$$C = S_0 N(d_1) - K e^{-rT} N(d_2)$$

$$P = K e^{-rT} N(-d_2) - S_0 N(-d_1)$$

Where:

$$d_1 = \frac{\ln(S_0/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}, \quad d_2 = d_1 - \sigma\sqrt{T}$$

Expectation Value and Option Payoff Distribution

The expectation value of an option is calculated based on the probability distribution of the underlying asset price:

$$E = \int_{-\infty}^{\infty} S(x) d(x) dx$$

Expectation Value of Call Option

$$E_{\text{Call}} = \int_{K}^{\infty} (S - K) f(S) dS$$

Expectation Value of Put Option

$$E_{\text{Put}} = \int_0^K (K - S) f(S) dS$$

Theoretical Explanation:

- Call Option Distribution: Gains grow when the price exceeds the strike price, right-skewed distribution.
- Put Option Distribution: Gains increase when the price drops below the strike price, left-skewed distribution.

Data Processing and Analysis

This section discusses the practical application of option analysis, focusing on data processing, modeling, and visualization. The insights are derived from simulated data and challenges encountered during real-world dataset analysis.

Challenges in Real-World Data

- **Data Incompleteness**: Many key fields, such as `Closing Price` and `Strike Price`, were missing or contained invalid values.
- **Data Irregularities**: Some files contained duplicate headers, empty rows, or inconsistent formatting, requiring extensive cleaning before analysis.
- **Zero Trading Volume**: A significant number of rows had zero trading volume, limiting the ability to derive meaningful insights.
- **Option Type Errors**: The `Call/Put` column included invalid or inconsistent entries, necessitating validation.

Detailed Data Cleaning and Refinement Steps

1. **Data Loading and Initial Inspection**:
   - Each file was read using Python's `pandas` library with `big5` encoding to handle text formatting.
   - Missing or invalid entries (e.g., `-`) were replaced with `NaN`.
   - Initial exploratory analysis identified structural issues.
2. **Value Validation and Filtering**:
   - Rows with non-numeric `Closing Price` or `Strike Price` were removed.
   - Only valid `Call` or `Put` entries in the `Call/Put` column were retained.
3. **Error Handling and Anomaly Detection**:
   - Automated scripts flagged duplicate headers and empty rows for manual review.

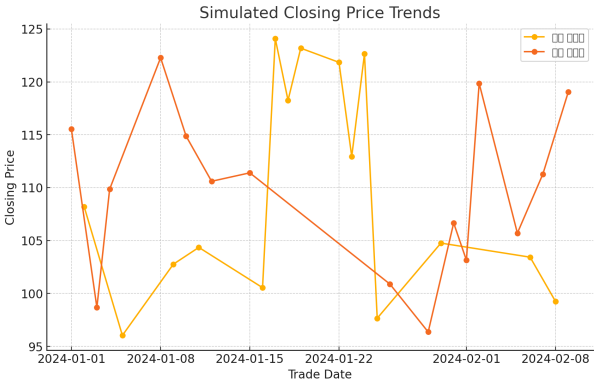

Figure 3: Enter Caption

- Invalid data patterns were logged and corrected.

4. **Black-Scholes Model Application**:
   - Cleaned data were used to calculate theoretical option prices.
   - Parameters included a risk-free rate (r), volatility (σ), and time to maturity (T).
   - **Data Integration**:
     - Cleaned datasets from all files were consolidated for comprehensive analysis.

Simulated Data Analysis

(a) **Data Generation**:
   - A simulated dataset of 30 entries was created, including `Trade Dates`, `Strike Prices`, `Call/Put` types, and `Closing Prices`.

(b) **Black-Scholes Model Application**:
   - Simulated data were used to calculate theoretical prices for validation.

(c) **Visualization**:
   - Trends in `Closing Prices` over time were plotted.
   - Scatter plots compared theoretical `BS Prices` with market `Closing Prices`.

Findings and Implications

- **Data Quality Impact**: The incomplete nature of real-world data necessitates robust preprocessing.
- **Simulation Value**: Simulated data provided a controlled environment to validate models and visualize market dynamics.
- **Model Insights**: The Black-Scholes model proved effective in understanding theoretical pricing.

Figure 4: Enter Caption

Recommendations for Future Work

(a) **Enhanced Data Validation**:

- Automate validation scripts to detect anomalies, such as duplicate headers.

(b) **Advanced Modeling Techniques**:

- Explore alternative pricing models like stochastic volatility models.

(c) **Improved Data Integration**:

- Implement tools to merge disparate datasets seamlessly.

(d) **Visualization Enhancements**:

- Incorporate multi-dimensional analysis, including volume and time decay effects.