Financial Data Analytics: with Machine Learning, Optimization and Statistics

Dependence-Aware VaR for Equity Put Sellers

Connor Radkov and Josh Marino

The Problem

- Black-Scholes assumes independence, but stocks often fall together in crashes
- This leads to underpricing of risk when selling puts on multiple stocks
- We simulate realistic joint movements using vine copulas
- Our model creates a full loss distribution across a portfolio of puts
- We compare this to Black-Scholes to see what percentage of risk is actually covered
- Helps sellers see if Black-Scholes premium are sufficient to cover extreme losses.

Data Collection and Pre-Processing

- **Historical Data Download:**
  We download one year of historical adjusted closing prices (01/01/2024 - 12/31/2024) for each stock (AAPL, MSFT, GOOG) using yfinance.
- **Price Extraction:**
  For each stock, we took the closing price on the final trading day of 2024. This becomes the starting point (S) for calculating the Black-Scholes prices.

Data Collection and Pre-Processing

- **Volatility Estimation:** (Chapter 6, pg 259)
  - **Log Returns:** We calculate daily log returns as:

$$\text{log return}_t = \ln\left(\frac{S_t}{S_{t-1}}\right)$$

- **Daily Variance and Annualized Volatility:** (Chapter 3, pg 67) We compute the variance of the log returns and annualize it by multiplying by 252 (approximate number of trading days in a year). The annualized volatility is then:

$$\sigma = \sqrt{\text{daily variance} \times 252}$$

Data Collection and Pre-Processing

- **Black–Scholes Put Pricing:** (Chapter 5, pg 165)
  - **Setting the Strike:** We set the strike price (K) to 90% of the current price (a typical out-of-the-money level for selling puts).
  - **Formula:** Using the Black–Scholes formula for a put:

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

$$\text{Put Premium} = Ke^{-rT}N(-d_2) - SN(-d_1)$$

  - This gives the premium received by the seller.

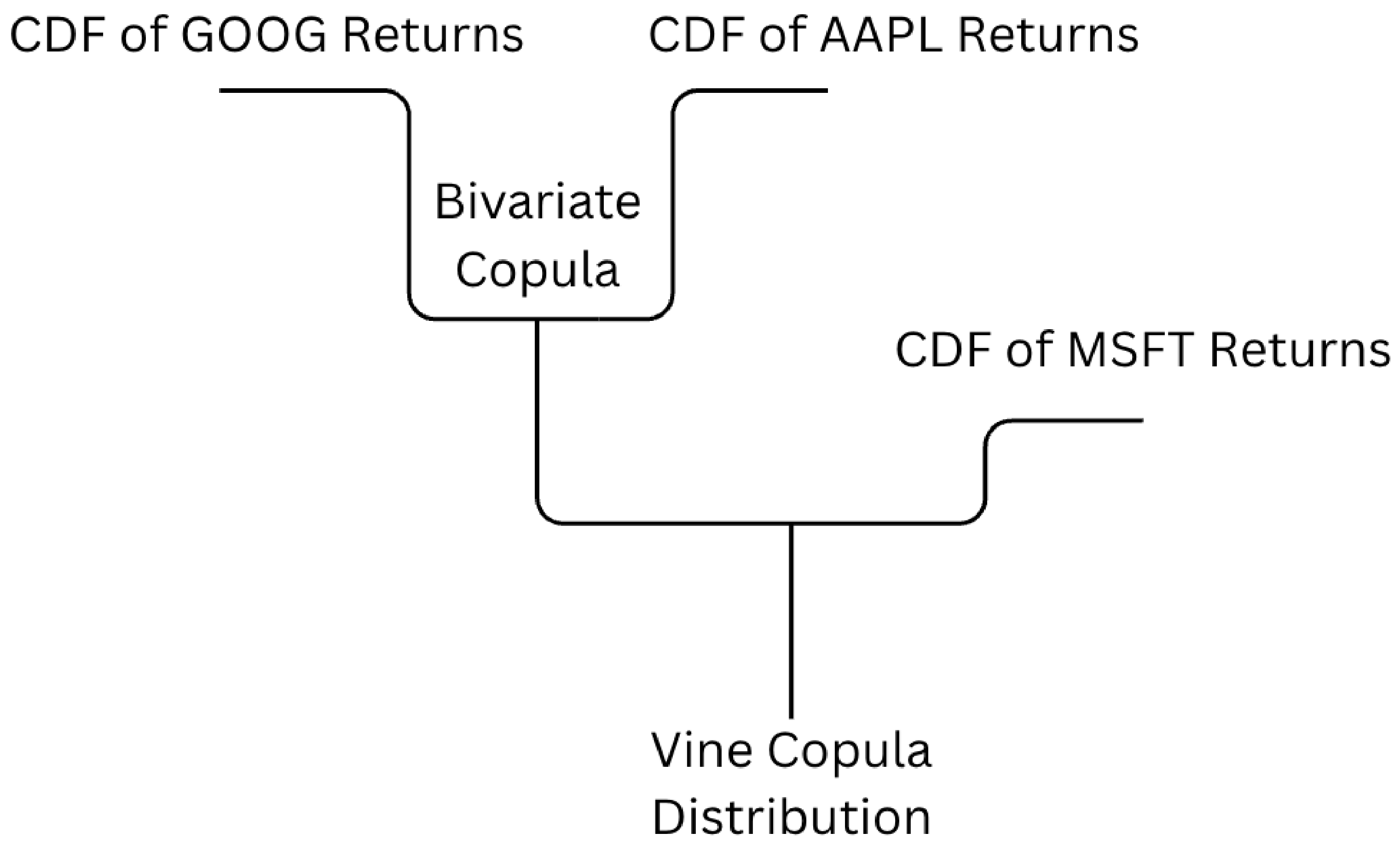
Copulas

- **Definition:** A **copula** is a multivariate cumulative distribution function for which the marginal CDF of each variable is uniform on the interval [0, 1].

- **Sklar's Theorem:** Given some marginal distributions and their copula, the copula can be used to model the behavior of their joint probability distribution.

"For any random variables $X_1,\ldots,X_\mathrm{p}$ with joint CDF $F(x_1,\ldots,x_\mathrm{p})$ and marginal CDFs $F_j(\mathrm{x}) = P(X_j \leq x)$, there exists a copula such that
$F(x_1,\ldots,x_p) = C(F_1(x_1),\ldots,F_p(x_p))$. Furthermore, if each $F_j(x)$ is continuous, then C is unique"

Vine Copula, Visualized



Capturing Cross-Sectional Dependencies with a Vine Copula (Chapter 8, pg 402)

Uniform Transformation:
Each return series is transformed into uniform data on the [0,1] interval by ranking the data and scaling:

$$u_i = \frac{\mathrm{rank}(r_i) - 0.5}{N}$$

- This is necessary because copula models work on uniform margins.

Vine Copula Fitting:
We fit a vine copula to the uniformized returns. The vine copula captures the dependency (correlation and tail behavior) between the returns of the stocks.

Simulation of Future Price Trajectories

- **Daily Uniform Simulation:**
  For each trading day over the option's life, we use the fitted vine copula to simulate joint uniform observations (1000 simulations per day).
- **Conversion to Daily Returns:**
  The simulated uniform values are transformed back into daily returns via the inverse normal transform:

$$\text{simulated daily return} = m + \sigma_{\text{daily}} \times \Phi^{-1}(u)$$

- **Compounding Returns:**
  For each simulation, the log returns are summed and exponentiated to provide the cumulative return of the stock:

$$\mathrm{R} = \sum_{t=0}^{n-1} ln(\tfrac{s_{t+1}}{s_t}) = ln(\tfrac{s_n}{s_0})$$
$$\text{Cumulative Return} = s_0 e^R$$

Loss Distribution and VaR Calculation

- **Loss Calculation (Seller's Perspective):**
  For each stock's simulated future price:

  - **If $S_{\text{simulated}} < K$:** The put is exercised, and the loss is:

    $$\text{loss} = (K - S_{\text{simulated}})$$

  - **If $S_{\text{simulated}} \geq K$:** The put expires worthless and the seller keeps the premium, so the "loss" is:

    $$\text{loss} = \ 0$$

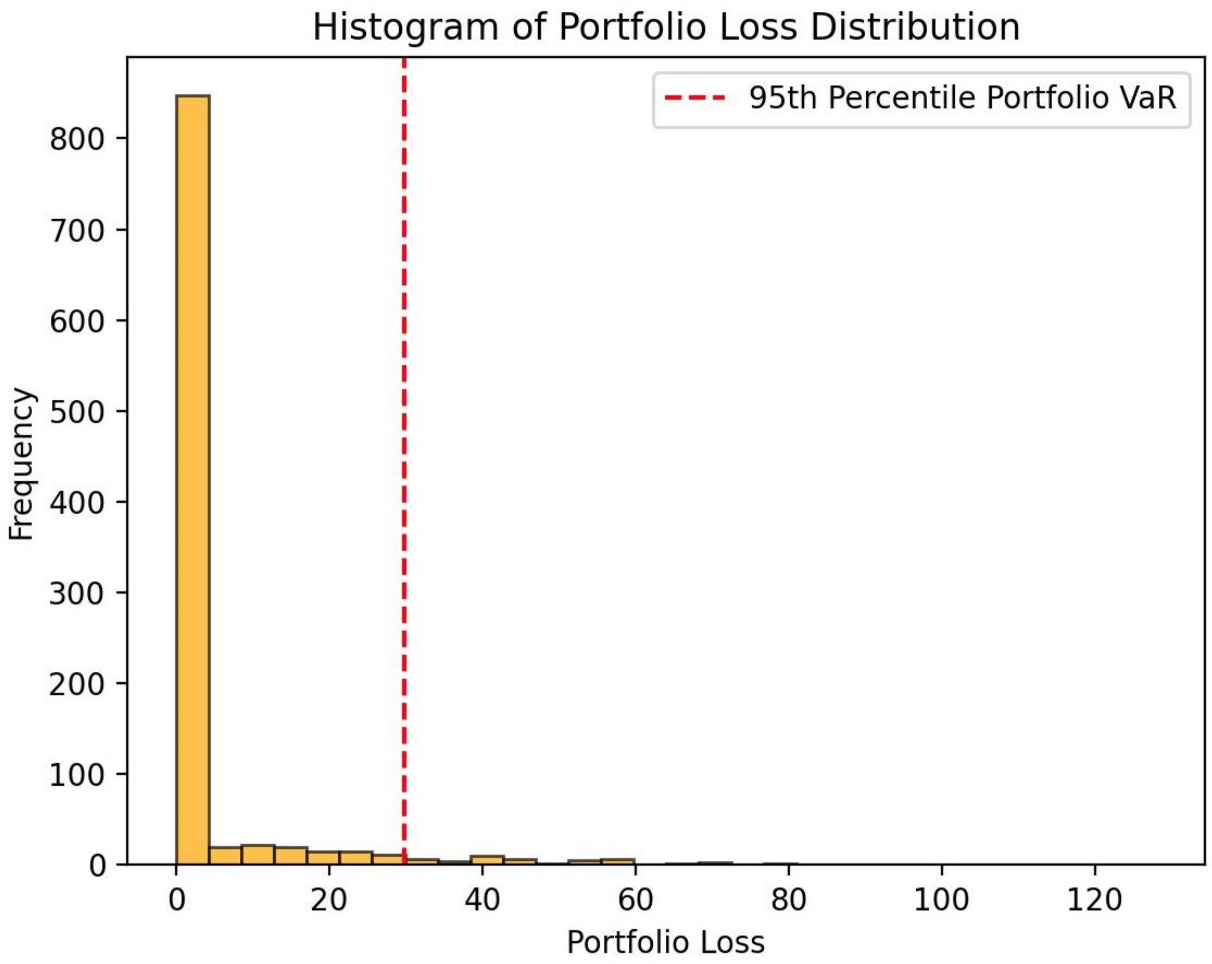
Loss Distribution and VaR Calculation

- **VaR (Value-at-Risk):** (Chapter 8, pg 345)
  The 95th percentile of the loss distribution is calculated for each stock. This represents statistically extreme losses (at 95% confidence) and indicates the capital the seller needs to cover that loss. Defined mathematically as the *p*-quantile loss distribution.
- **Total Capital Requirement:**
  The individual VaRs can be modeled to provide an estimate of the total capital required to cover potential losses across all positions for regulated sellers.

$$\mathrm{V@R_L}(p) = F_{\mathrm{L}}^{-1}(p) = \min\{l \in \mathbb{R} : F_{\mathrm{L}}(l) \geq p\}.$$

Program Results

Black-Scholes Price	Vine Copula Price
AAPL: $4.40	N/A
MSFT: $5.64	N/A
GOOG: $5.46	N/A
Total: **$15.50**	Total: **$29.73**



Findings

- The vine copula suggests that losses between the three stocks are likely **dependent**.
- From the vine copula simulations, we found that the 95% VaR is **higher** than the sum of the BS premiums. This means that the BS premium would not be sufficient to cover extreme losses.
- Based on the copula results, the Black-Scholes premium would only cover the 81.2% VaR.
- So, the seller may need additional capital available to cover extreme losses.

Remarks

1. Data Collection and Pre-Processing
2. Capturing Cross-Sectional Dependencies with a Vine Copula
3. Simulation of Future Price Trajectories
4. Loss Distribution and VaR Calculation
5. Results
6. Findings