Jump Detection & Option Pricing Models

Jacob Aylmer, Harshil Cherukuri, Dhruv Kanchi,
Om Mehta, William Parente

12-10-2023

Introduction & Overview

- The jump is incorporated into the model so we can capture the way the market reacts to volatile unexpected shifts in prices.
- We plan to explore this relationship in depth, and provide visual representations such as histograms to show the impact of jumps in stock prices on the prices of their options.
- Based on a variety of different factors such as time of the occurrence of the jump, size of the jump, and volume of the jump, we will create a clustering technique to categorize the jumps.
- We have created 2 models and are in the process of creating a third.

Methodology

- Calibrate parameters for the Merton Jump Diffusion process daily
  - Jump rate: λ
  - Jump size: θ
- Analyze intricate patterns and categorize jumps with a nuanced approach, considering various factors such as time, size, and volume
- Delve into the intricate relationships between daily jumps and option prices for a thorough understanding
- Conduct comprehensive sensitivity analysis to evaluate the robustness and resilience of the Jump-Diffusion model
- Various techniques to detect jumps include:
  - Returns exceeding 2 standard deviations from the mean
  - Utilizing the classical jump detection model by Lee-Mykland
  - Employing machine learning algorithms specifically designed for identifying jumps

Motivation and Previous Literature Review

- Robert C. Merton published a revolutionary paper in this field in 1976
- Caused jump diffusion processes to become a classic tool to evaluate and model different financial instruments
- Merton published a series of papers and books since the 1970s that develop the ideas presented in this paper
- Analyze the influence this paper had on other paper had on other papers relating to this topic in the industry, considering its groundbreaking nature

"A Jump-Diffusion Model for Option Pricing"

- Addresses Leptokurtic feature of stock returns
- Higher peak and heavier tails, skewed left
- Volatility smile
- Black-Scholes - implied volatility constant
- Double Exponential Jump-Diffusion model improves upon Merton Model

"Jump Diffusion Models for Option Pricing vs. the Black-Scholes Model"

- Standardizes a similar approach to traditional option pricing but made some small tweaks.
- Addresses the jump factor by using parameter *Se* instead of the traditional S into Black-Scholes.
- λ represents the jump parameter.
- The stock price itself, multiplied by the exponential, is given by Kou's equation seen earlier.

"Jump Detection in Financial Time Series using Machine Learning"

- Hybrid model implementing aspects of classical jump detection models such as Barndorff-Nielsen, Shephard, Lee-Mykland, and Yacine Ait-Sahalia & Jean Jacod
- 15-minute interval data due to cost of data collection, the accuracy of the data, and limited historical data availability
- Tail risk protection strategy

"The Jump-Risk Premia"

- Demonstrates integration techniques through applications
- Models markets during times of high volatile
- Identifies jumps via machine learning
- Provides context as to the relationship between discontinuous jumps and risk premiums

Data Description

- 15 Minute Intraday Data from Refinitiv Hanlon Lab
- 2010 & 2011 Time Series Data
- SPY Shape:
  - 20,150 Rows
  - 8 Columns
- SPY Options Shape:
  - 65,478,041 Rows
    - 2,587,557 rows after eliminating options not traded
  - 12 Columns

Head SPY

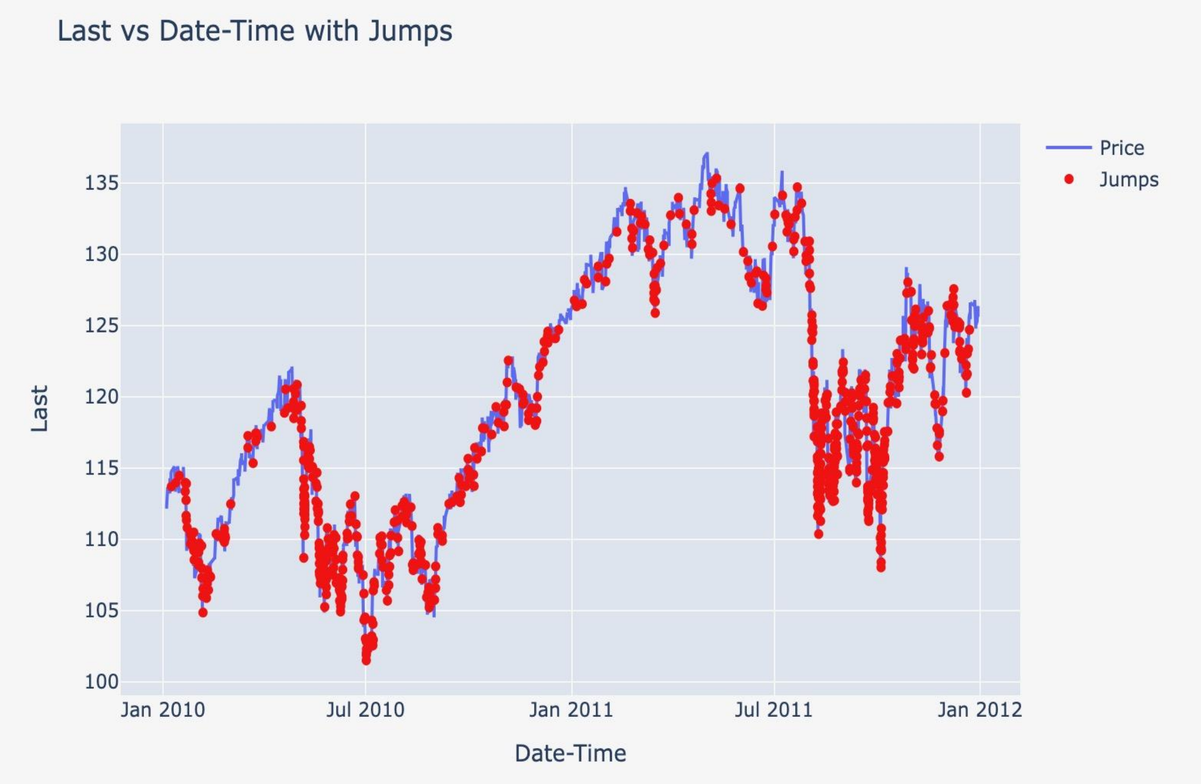
	#RIC	Date-Time	Open	High	Low	Last	Volume	No. Trades
0	SPY	2010-01-04T08:00:00.000000000-05	112.12	112.180000	111.440000	112.140	1765282.0	428.0
1	SPY	2010-01-04T08:15:00.000000000-05	112.14	112.160000	112.090000	112.150	682644.0	674.0
2	SPY	2010-01-04T08:30:00.000000000-05	112.16	112.230000	112.140000	112.220	988952.0	873.0
3	SPY	2010-01-04T08:45:00.000000000-05	112.22	112.330000	112.220000	112.320	378532.0	1167.0
4	SPY	2010-01-04T09:00:00.000000000-05	112.33	112.460000	112.330000	112.390	554833.0	1641.0
5	SPY	2010-01-04T09:15:00.000000000-05	112.39	112.440000	111.440000	112.360	684084.0	2020.0
6	SPY	2010-01-04T09:30:00.000000000-05	112.37	112.750000	112.330000	112.730	12623404.0	31193.0
7	SPY	2010-01-04T09:45:00.000000000-05	112.73	112.940002	111.440000	112.830	9225048.0	25754.0
8	SPY	2010-01-04T10:00:00.000000000-05	112.83	112.960000	112.730000	112.900	8984505.0	20777.0
9	SPY	2010-01-04T10:15:00.000000000-05	112.91	113.100000	112.839996	113.085	4809410.0	13115.0

Head SPY Options

	#RIC	Date-Time	Open	High	Low	Last	Volume	No. Trades	Root	Strike	Expiry	Option Type
0	SPYA161010000.U	2010-01-04T09:30:00.000000000-05	12.50	12.75	12.50	12.75	19.0	3.0	SPY	100.0	2010-01-16	call
1	SPYA161010200.U	2010-01-04T09:30:00.000000000-05	10.84	10.84	10.84	10.84	40.0	1.0	SPY	102.0	2010-01-16	call
2	SPYA161010400.U	2010-01-04T09:30:00.000000000-05	8.48	8.48	8.47	8.47	2.0	2.0	SPY	104.0	2010-01-16	call
3	SPYA161010500.U	2010-01-04T09:30:00.000000000-05	7.65	7.88	7.65	7.88	84.0	10.0	SPY	105.0	2010-01-16	call
4	SPYA161010600.U	2010-01-04T09:30:00.000000000-05	6.80	6.89	6.80	6.83	13.0	4.0	SPY	106.0	2010-01-16	call
5	SPYA161010800.U	2010-01-04T09:30:00.000000000-05	4.82	5.01	4.81	5.00	23.0	7.0	SPY	108.0	2010-01-16	call
6	SPYA161010900.U	2010-01-04T09:30:00.000000000-05	3.80	4.15	3.80	4.15	22.0	8.0	SPY	109.0	2010-01-16	call
7	SPYA161011000.U	2010-01-04T09:30:00.000000000-05	3.00	3.34	3.00	3.30	318.0	23.0	SPY	110.0	2010-01-16	call
8	SPYA161011100.U	2010-01-04T09:30:00.000000000-05	2.38	2.55	2.29	2.51	3458.0	99.0	SPY	111.0	2010-01-16	call
9	SPYA161011200.U	2010-01-04T09:30:00.000000000-05	1.71	1.83	1.56	1.81	7475.0	204.0	SPY	112.0	2010-01-16	call

Simple Detection Method

- Identified 776 jumps in SPY data
  - Jump: Return >2 standard deviations from mean
  - 382 upward jumps
  - 394 downward jumps
- Mean up jump: 0.747%
- Mean down jump: -0.726%
- Largest jump: -4.26%
  - Flash Crash: May 6, 2010
    - @ 2:30 PM - 2:43 PM



Classical Method: Lee - Mykland Model

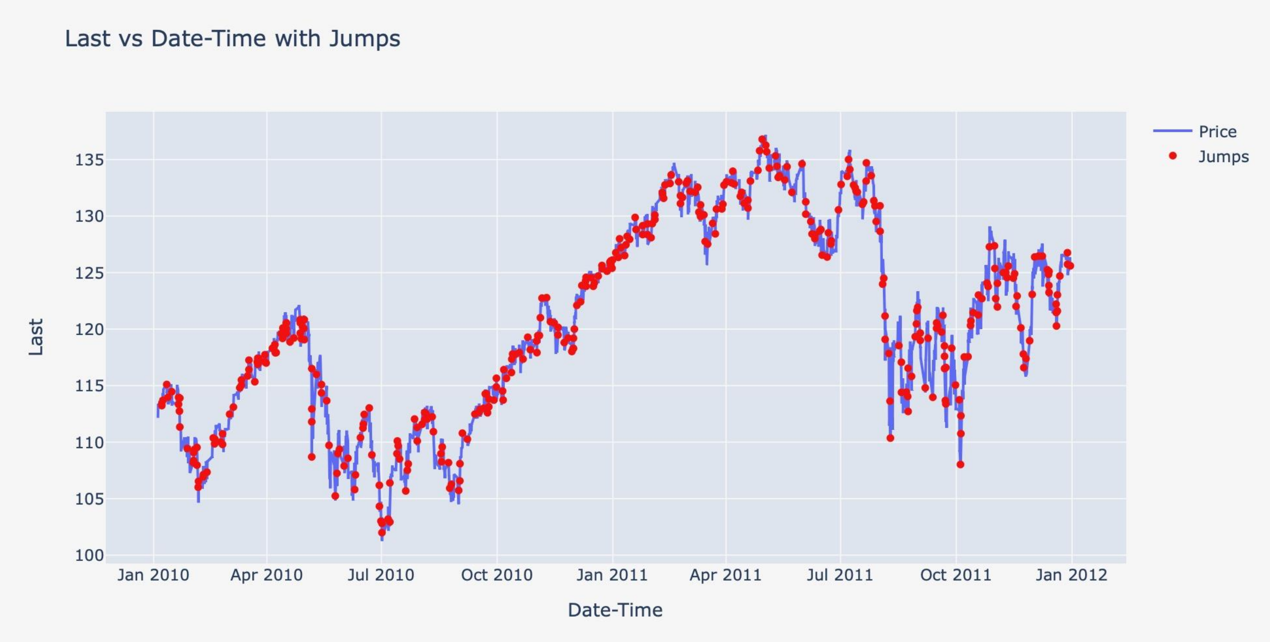
Definition 1. *The statistic $\mathcal{L}(i)$, which tests at time t_i whether there was a jump from t_{i-1} to t_i, is defined as*

$$\mathcal{L}(i) \equiv \frac{\log S(t_i)/S(t_{i-1})}{\widehat{\sigma(t_i)}}, \tag{7}$$

where

$$\widehat{\sigma(t_i)}^2 \equiv \frac{1}{K-2} \sum_{j=i-K+2}^{i-1} |\log S(t_j)/S(t_{j-1})||\log S(t_{j-1})/S(t_{j-2})|. \tag{8}$$

- Identified 378 jumps in SPY data
  - 172 upward jumps
  - 206 downward jumps
- Formula given for sigma and test statistic
- Statistically significant results indicate jump
- Next steps:
  - Calibrate model to get parameters based on the SPY data
    - Rolling window



Detection Result Comparison and Reflection

	# Jumps Identified	# Up Jumps	Mean Up Jumps	# Down Jumps	Mean Down Jumps
Simple Model	776	382	0.747%	394	-0.726%
Lee-Mykland Model	378	172	0.869%	206	-0.797%

Potential changes:

- Simple model:
  - Increase standard deviations away from mean to reduce number of jumps identified
- Lee-Mykland model:
  - Adjust **α** to require more or less significant test statistic

Next Steps

- After coding the simple model and the Lee-Mykland model, our next goal is to create a machine learning based algorithm for identifying jumps.
  - Will be reliant on the creation of synthetic data
  - Able to naturally handle data of varying dimensionality by jointly modeling :
    - state of each datapoint
    - dimension of each datapoint
- Define/derive dimension based on:
  - destroying forward noising process
  - time-reversed generative process
  - novel evidence lower bound training objective for approximation
- Simulating our learned approximation to the time-reversed generative process then provides an effective way of sampling data of varying dimensionality by jointly generating state values and dimensions.
- Approach is demonstrated on molecular and video datasets of varying dimensionality
- Also looking to:
  - Calibration
  - Application to Pricing Models