Analysis of Kalshi Prediction Markets

Vidit Pokharna

November 26, 2025

1. Introduction

Prediction markets provide a real time mechanism for aggregating beliefs about future events. Kalshi allows users to trade binary outcome contracts whose prices reflect implied probabilities of economic, financial, and political outcomes. This report studies one such market in detail, beginning with API level data collection and continuing through statistical analysis, decomposition, macroeconomic alignment, and machine learning based forecasting.

The objective is to construct a full pipeline for examining how a Kalshi market evolves over time, including identifying a high quality market, extracting hourly candles, analyzing relationships with Bitcoin spot and CME FedWatch data, and evaluating several predictive models.

Using metadata from the Kalshi REST API, more than seven thousand series and over two hundred markets were examined. Based on liquidity, depth, and history length, the contract **KXBTCMINY-25-2-DEC31-80000** was selected as the primary case study.

This report is structured as follows: Section 2 covers data processing and descriptive plots. Section 3 presents a seasonal decomposition. Section 4 examines connections to Bitcoin and interest rate expectations. Section 5 evaluates forecasting models. Section 6 concludes.

2. Data Collection and Preprocessing

Kalshi provides hourly candlestick data for each market including OHLC, volume, and timestamps. All data for the chosen market was collected using authenticated REST calls, converted into a timezone aware index, and sorted. Minor gaps were forward filled to preserve hourly granularity.

2.1. Data Source Selection

Data collection was attempted from both Kalshi and Polymarket prediction market platforms. Kalshi's REST API provided comprehensive hourly candlestick data including OHLC prices and volume with consistent formatting. Polymarket's CLOB API (`clob.polymarket.com`) was successfully integrated, but limitations included the absence of volume data in price history endpoints and variable response formats requiring extensive error handling. Given Kalshi's superior data completeness and consistency, the analysis focused on Kalshi markets, with the contract **KXBTCMINY-25-2-DEC31-80000** selected based on liquidity and historical depth criteria.

2.2. Data

Figure 1 shows the complete hourly price history. The contract price rose from roughly 0.10 in late October to near 0.84 in late November, indicating steadily increasing belief that Bitcoin would exceed 80,000 USD by year end.

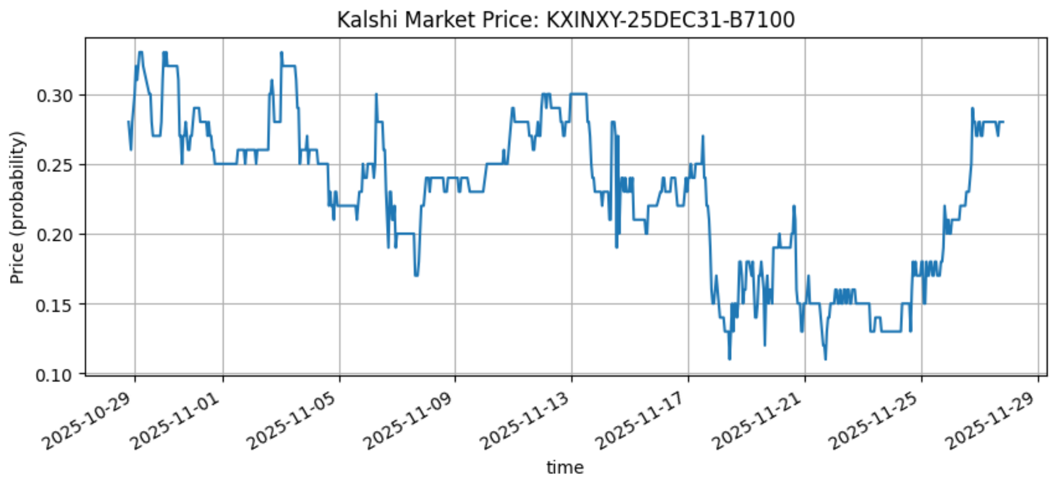

Figure 1: Kalshi market price for KXBTCMINY-25-2-DEC31-80000.

Volume dynamics in Figure 2 show several pronounced bursts aligning with major Bitcoin moves. The price–volume correlation was approximately 0.396, indicating that trading activity increases around repricing events rather than steadily influencing the trend.

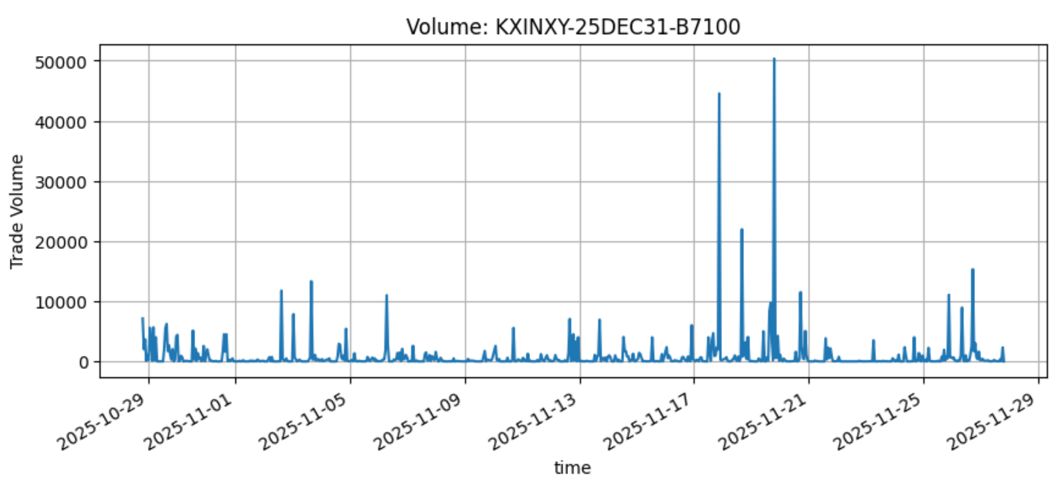

Figure 2: Hourly trade volume for the market.

Rolling volatility (Figure 3) reveals two clear volatility regimes corresponding to market repricing. These structural features support the suitability of smooth trend-based forecasting models.

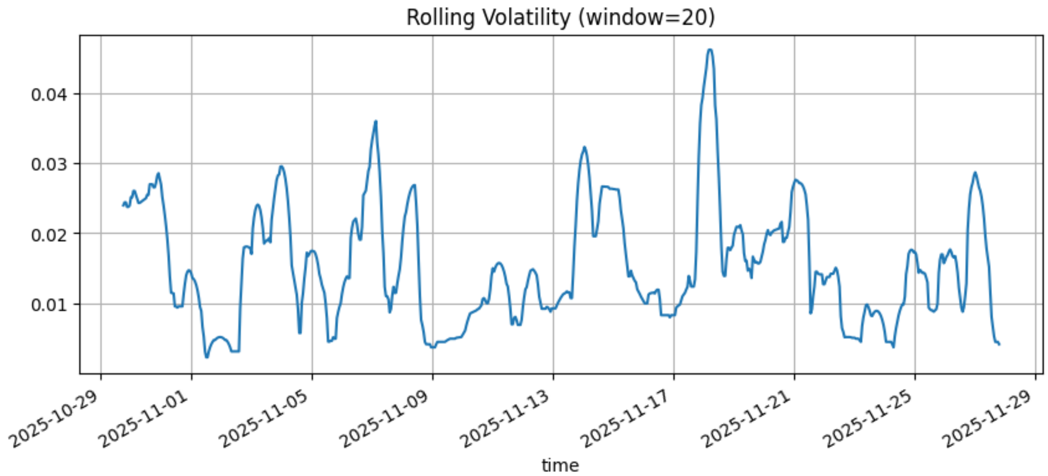

Figure 3: Rolling volatility of the contract using a 20 hour window.

3. Time Series Decomposition

A classical seasonal decomposition was used to separate the price into trend, seasonal, and residual components. Figure 4 displays the results. The trend dominates and increases smoothly from roughly 0.12 to 0.74. Seasonality has minimal amplitude, while the residual captures short bursts consistent with volatility spikes.

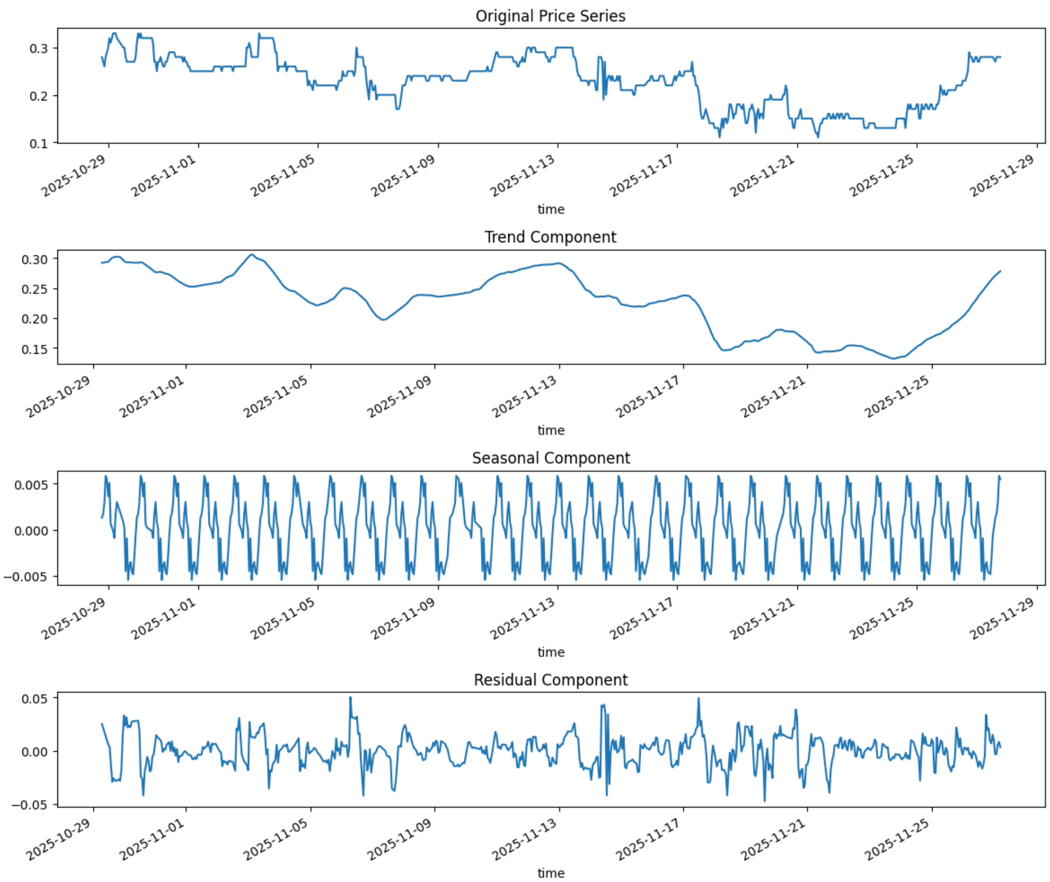

Figure 4: Seasonal decomposition of the Kalshi price series.

These results show that the series is largely trend driven with limited nonlinear structure. Consequently, simple models tend to outperform heavily regularized tree based methods.

4. Relationship to External Underlyings

Bitcoin spot data was aligned with the Kalshi series. The Pearson correlation between the Kalshi price and spot Bitcoin returns was roughly -0.85, consistent with the contract's payoff structure: downward Bitcoin moves reduce the likelihood of finishing above 80,000 USD.

A lead lag analysis, shown in Figure 5, found the strongest magnitude at lag +1 with correlation around -0.91. This indicates that the Kalshi market often incorporates information slightly earlier than Bitcoin spot.

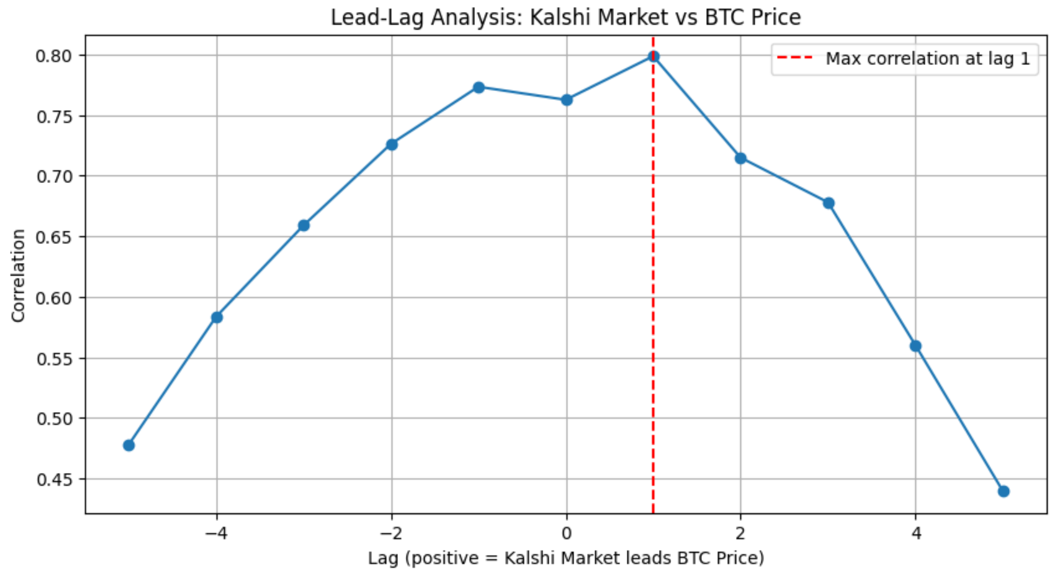

Figure 5: Lead lag analysis of Kalshi price versus Bitcoin spot returns.

Federal Reserve expectations were also examined. CME FedWatch data showed moderate correlations (up to 0.83 in the 150–200 bps bin), suggesting sensitivity to broader macro conditions, though these effects were weaker than Bitcoin based relationships.

5. Predictive Modelling

Several models were tested, including linear regression, a small neural network, and XGBoost, using lagged price and volume as features. An 80/20 chronological split was used.

Linear regression achieved a test R^2 of 0.93 and was the best performing method. The neural network achieved around 0.80, lagging during abrupt moves. XGBoost performed poorly, producing negative R^2 values even under conservative hyperparameters, due to overfitting in this low dimensional setting.

Overall, the series' smooth trend strongly favors linear forecasting approaches.

6. Conclusion

This report analyzed the Kalshi contract KXBTCMINY-25-2-DEC31-80000 across data extraction, decomposition, volatility, macro alignment, and predictive modelling. The market exhibited a dominant upward trend, limited seasonality, and a strong connection to Bitcoin price movements.

Lead lag analysis suggested that Kalshi prices often move ahead of Bitcoin spot, while FedWatch correlations highlighted sensitivity to interest rate expectations. Forecasting results demonstrated that linear models capture the structure effectively, whereas more complex models overfit.

The methodology used here offers a general framework for studying other Kalshi markets and understanding how prediction markets evolve over time.