國立臺灣科技大學

財務金融研究所碩士班

碩士學位論文

學號：M11218014

比特幣選擇權隱含風險中立機率密度

之平滑尾部導取

Extracting Smooth Tails of Option Implied Risk-neutral Densities

in the Bitcoin Market

研究生：王士誠

指導教授：薛博今 博士

中　華　民　國　一　一　四　年　五　月

摘要

本研究旨在探討風險中立機率密度（Risk-neutral Density, RND）尾部配適方法對比特幣選擇權市場報酬率預測能力之影響，並提出創新的單點加斜率配適法，與文獻中常用的雙點配適法 (Birru & Figlewski, 2012) 進行實證比較。自 Breeden 與 Litzenberger (1978) 開創性研究以來，從選擇權價格中提取 RND 已成為探索市場預期的重要方法，然而在處理極端市場條件下的尾部風險估計仍存挑戰。

本研究以全球最大比特幣選擇權交易平台 Deribit 於 2021 年至 2024 年的交易數據為基礎，結合廣義柏拉圖分布（Generalized Pareto Distribution, GPD），建立完整的尾部延伸模型。透過比較距到期日 1 天與 7 天之選擇權樣本，檢驗 RND 統計特徵對現貨報酬率的預測能力。實證結果顯示，在計算效率方面，單點加斜率配適法較雙點配適法平均提升 10.95%；在日報酬率預測中，單點配適法建構之模型以偏態（Skewness）、中位數（Median）及前期報酬率為顯著預測變數（R-squared = 0.0130），表現優於雙點配適法（R-squared = 0.0098），此結果與 Conrad 等人 (2013) 及 Liu 與 Tsyvinski (2021) 關於偏態與動量效應的研究相呼應。

在週報酬率預測方面，兩種方法表現接近，最佳預測變數組合為超額峰態（Excess Kurtosis）、中位數（Median）及加密貨幣恐懼與貪婪指數。超額峰態的預測力呼應 Amaya 等人 (2015) 研究結果，而恐懼與貪婪指數的顯著影響則與 He 等人 (2023) 的發現一致。樣本外預測結果顯示，無論在日報酬率或週報酬率預測中，兩種方法均產生正值的樣本外 R 平方，且單點配適法表現優於雙點配適法，顯示我們的預測模型具實質經濟價值 (Campbell & Thompson, 2008)。

本研究主要貢獻在於提出單點加斜率配適法，此方法提升 RND 尾部配適的效率與穩定性，更驗證了 RND 統計特徵在高波動性市場中的應用價值，為投資組合風險管理及交易策略優化提供具體指引。

Abstract

This study investigates the impact of risk-neutral density (RND) tail estimation methods on return predictability in the Bitcoin options market and proposes an innovative single point with the slope estimation method, comparing it with the conventional two-point estimation approach documented in the literature (Birru & Figlewski, 2012). Since the pioneering work of Breeden and Litzenberger (1978), extracting RNDs from option prices has become an important technique for exploring market expectations, yet challenges remain in estimating tail risks under extreme market conditions.

Based on trading data from Deribit, the world's largest Bitcoin options trading platform, spanning from 2021 to 2024, this study integrates the Generalized Pareto Distribution (GPD) to develop a comprehensive tail extension model. By comparing options samples with 1-day and 7-day maturities, we examine the predictive power of RND statistical features on spot returns. Empirical results indicate that the single point with the slope estimation method improves computational efficiency by 10.95% compared to the two-point estimation method. For daily return prediction, the regression model constructed using the single point with the slope estimation method identifies skewness, median, and lagged returns as significant predictive variables (R-squared = 0.0130), outperforming the two-point estimation method (R-squared = 0.0098). These findings align with research on skewness and momentum effects by Conrad et al. (2013) and Liu and Tsyvinski (2021).

For weekly return prediction, both methods perform similarly, with the optimal combination of predictive variables being excess kurtosis, median, and the Cryptocurrency Fear and Greed Index. The predictive power of excess kurtosis echoes the findings of Amaya et al. (2015), while the significant influence of the Fear and Greed

Index is consistent with discoveries by He et al. (2023). Out-of-sample prediction results demonstrate that both methods yield positive out-of-sample R-squared values for both daily and weekly return predictions, with the single point with the slope estimation method outperforming the two-point estimation method, indicating that our predictive models possess substantial economic value (Campbell & Thompson, 2008).

The primary contribution of this research lies in the development of the single point with the slope estimation method, which enhances the efficiency and stability of RND tail estimation while validating the application value of RND statistical features in highly volatile markets, providing concrete guidance for portfolio risk management and trading strategy optimization.

Acknowledgement

Contents

List of Figures

List of Tables

1. Introduction

1.1 Research Background and Motivation

The evolution of financial market risk assessment tools has consistently been one of the core topics in financial research. In recent years, with the rapid expansion of the cryptocurrency ecosystem, Bitcoin, as the largest cryptocurrency by market capitalization, has witnessed explosive growth in its derivatives market (Akyildirim et al., 2020). According to the latest data from The Block (*The Block*, 2025), Bitcoin futures and options trading volume exceeded $21 trillion in 2024, with options trading volume growing at an annual rate of 130%, far surpassing traditional financial derivatives markets. This phenomenon reflects investors' growing demand for cryptocurrency risk management tools while providing researchers with a unique perspective to explore price discovery mechanisms in emerging markets (Zulfiqar & Gulzar, 2021).

The cryptocurrency market exhibits characteristics significantly different from traditional financial markets: 24/7 trading, extreme volatility, decentralized structure, and unique investor composition. Particularly in terms of volatility, Bitcoin's historical annualized volatility frequently exceeds 100%, far higher than the approximately 15-20% volatility level of traditional stock indices (Liu & Tsyvinski, 2021). This extreme volatility makes risk management particularly important in this market and presents unique challenges for interpreting the implied information in option prices.

As a derivative financial instrument, options contain rich market information in their prices. Call options grant holders the right to purchase the underlying asset at a predetermined price at a specific future time; put options grant holders the right to sell the underlying asset at a predetermined price at a specific future time. Due to the non-linear payoff structure of options, their prices reflect not only the market's expectations of future trends in the underlying asset but also market participants' assessment of volatility risk (Hull, 2021). The information embedded in option prices reflects not only the market consensus on risk expectations but also serves as an important window for exploring market microstructure and investor behavior (Bakshi et al., 2003). Although the traditional Black-Scholes model (1973) provides a theoretical foundation for option pricing, its assumption of normal distribution fails to adequately capture the fat-tailed distribution and skewness characteristics common in financial markets. Particularly in high-volatility markets such as cryptocurrencies, the frequency of abnormal returns and extreme events is significantly higher than in traditional assets, making accurate estimation of market-implied risk distributions more critical (Chordia et al., 2021).

In option pricing theory, the Black-Scholes model (1973) laid an important foundation for modern option pricing. The model assumes that the underlying asset price follows geometric Brownian motion and introduces the no-arbitrage principle to derive the theoretical price of options. However, observed option prices in practice often deviate from theoretical values, with these deviations reflecting the market's actual assessment of future risks. In particular, the implied volatility obtained by inverting the Black-Scholes model exhibits systematic differences across options with different strike prices, forming the so-called "volatility smile" or "volatility skew" phenomenon (Rubinstein, 1994). This phenomenon indicates that the market's expectation of the future price distribution of the underlying asset is not the log-normal distribution assumed by the Black-Scholes model, but rather has a more complex form, especially in the tail regions.

Extracting the risk-neutral density (RND) from option prices is a more comprehensive approach that captures the market's complete expectation of future price distributions. According to Breeden and Litzenberger (1978), deriving RND from option prices primarily involves three steps. First, using observed market option prices, implied volatility is obtained through the Black-Scholes model, with each strike price option corresponding to an implied volatility value. However, since only a finite number of strike prices exist in the market, appropriate models are needed to fit the implied volatility across different strike prices to establish a complete implied volatility curve. There are multiple methods for fitting implied volatility. Hagan & West (2006) research indicates that quadratic splines can avoid the risk of overfitting while maintaining curve smoothness, which is very important for ensuring model robustness and reliability, especially in situations of high market volatility. Haslip & Kaishev (2014) research shows that when dealing with complex derivative financial products such as lookback options, quadratic splines combined with Fourier transforms can provide efficient and accurate pricing results, achieving a good balance between computational efficiency and precision, though with the risk of discontinuous first derivatives. Research by Bliss & Panigirtzoglou (2004), Figlewski (2008), and Monteiro et al (2008) used cubic spline functions to interpolate implied volatility. The cubic spline method has the advantages of high computational efficiency and relatively simple implementation, providing reasonable fitting results under most market conditions. However, cubic splines only guarantee the continuity of the first derivative, while the second derivative may be discontinuous at nodes, which may not yield ideal fitting effects when dealing with extreme volatility. Additionally, when extracting RND, since the second derivative needs to be calculated, cubic splines may lead to insufficiently smooth density functions. Therefore, our proposed approach references the research method of Reinke (2020) adopting quartic spline functions with a single knot for implied volatility curve fitting. This method has higher smoothness compared to quadratic and cubic spline functions, ensuring continuity of both first and second derivatives while providing more accurate volatility curve estimates to construct more precise RND curves. Finally, based on the theory proposed by Breeden and Litzenberger (1978), the RND is derived by taking the second derivative of European call option prices with respect to strike prices. This method of extracting RND from option prices is known as the Model-free Approach because it does not rely on specific assumptions about the underlying asset price dynamics but directly extracts RND from market prices. The advantage of this method lies in its ability to fully capture the market's expectations of future price distributions, including higher-order moments such as skewness and kurtosis.

However, due to liquidity limitations in the options market, especially in deep out-of-the-money regions, RND extracted directly from market prices often lacks sufficient information in the tail regions. To construct complete RNDs, previous research has developed two major categories of RND tail fitting methods: non-parametric and parametric approaches. Non-parametric methods are applied in the field of RND estimation due to their freedom from specific model assumptions. Bondarenko (2000) proposed a non-parametric method for deriving RND from option prices, with its main advantage being minimal assumptions about the underlying asset, computational simplicity, and effective control of overfitting problems, yielding reasonable results even with small samples. Empirical research shows that daily changes in RND correlate with index performance. Grith et al. (2012) discussed in detail the application of kernel smoothing, spline functions, and other non-parametric techniques in RND estimation in the Handbook of Computational Finance. The authors emphasized the flexibility of non-parametric methods to capture complex distributional features implied by the market, such as skewness, excess kurtosis, and multimodal distributions that traditional parametric models struggle to describe. Monteiro & Santos (2022) pointed out that in kernel-based non-parametric RND estimation, local constraints cannot ensure that the integral of the second derivative function equals one. To address these issues, the authors imposed broader no-arbitrage constraints, using the Heston model and hypergeometric functions for simulation, applied to VIX and S&P 500 index RND estimation. Dong et al. (2024) proposed the innovative Implied Willow Tree (IWT) method, which directly reconstructs the complete risk-neutral process from cross-maturity market option price data without relying on any preset parametric models.

Parametric methods mostly use Extreme Value Theory (EVT) as an important approach, utilizing distribution functions from extreme value theory to extend the RND tails. Figlewski (2008) first proposed using the Generalized Extreme Value Distribution (GEV) with two points fitting, selecting two joining points in each of the left and right tails, requiring continuity conditions for density function values and cumulative density values at these points. Birru and Figlewski (2012) further improved this method by replacing GEV with the Generalized Pareto Distribution (GPD). The research found that regardless of high or low volatility periods, RND exhibits significant left skewness, contrasting with the log-normal distribution assumed by the Black-Scholes model. Additionally, to avoid arbitrage opportunities between markets, the mean of the RND should equal the futures price of the index for the same period. The research confirmed that even within the shortest time intervals during extreme market turbulence, the two values remain very close, indicating that the market pricing mechanism remains effective in maintaining no-arbitrage relationships. During crisis periods, changes in investors' attitudes toward risk are directly reflected in option prices and manifested in the extracted RND. These findings provide important insights into understanding investor behavior and market pricing mechanisms under extreme market stress, and also demonstrate the accuracy and contribution of GPD in tail risk estimation. McNeil and Frey (2000) proposed a method combining GARCH models with extreme value theory (EVT) to estimate tail risk in financial time series. Another common method is the mixed distribution approach, which decomposes RND into core and tail parts, fitting them with different distribution functions. Glatzer and Scheicher (2005) proposed a method combining a mixture of log-normal distributions for the core part with GPD distributions for the tails, which performed well in applications in the Eurozone bond market. Markose and Alentorn (2011) proposed parameterizing RND tails using the Generalized Extreme Value Family, a method that effectively captures market expectations of extreme events. A third method is density function extrapolation, which extrapolates tails based on the known shape of the RND. Monteiro et al. (2008) proposed using cubic spline functions with non-negativity constraints to extract risk-neutral probability density from option prices and adopted exponential functions for extrapolating RND tails. This method is computationally simple but may not fully capture complex tail features. Recently, parametric methods have also developed approaches that do not require extreme value theory. Orosi (2015) proposed a novel parametric method for estimating risk-neutral probability density functions. By setting appropriate functional forms and imposing constraints on model parameters, this method can produce risk-neutral density estimates with good properties. Uberti (2023) proposed a new semi-parametric estimation method combining the stability of parametric methods with the flexibility of non-parametric methods. Y. Li et al. (2024) proposed a new parametric estimation method using a Lognormal-Weibull mixture model, which provides more accurate predictive performance when measuring skewness and analyzing RNDs with multiple peaks, offering a new research direction for RND estimation and demonstrating the continuous refinement trend in risk-neutral density estimation methodology.

Our proposed approach chooses to use parametric methods to estimate RND tails and studies the methods proposed by Birru and Figlewski, using GEV (Figlewski, 2008) and GPD (Birru & Figlewski, 2012) to construct complete RND curves. This is primarily because Figlewski's method (Figlewski, 2008) successfully addresses the key challenge of extreme risk estimation in financial markets by accurately extending tails using distribution functions based on extreme value theory, enhancing the accuracy of risk assessment. Subsequently, GPD (Birru & Figlewski, 2012) achieves a balance between theoretical rigor and practical applicability, not only conforming to financial theoretical foundations such as no-arbitrage pricing but also demonstrating empirical analytical utility in extreme market environments like the 2008 financial crisis. Third, the two points fitting method ensures the continuity of the density function and its cumulative distribution at joining points, making the overall RND curve smoother and more natural. The progression from initially adopting GEV (Figlewski, 2008) to subsequently improving to GPD (Birru & Figlewski, 2012) demonstrates the method's good flexibility and scalability, allowing adjustments and optimizations based on different market characteristics. This method effectively captures market expectations of extreme risks and panic sentiment, providing valuable market prediction value for investors, risk managers, and policy makers.

However, we note that this method has significant limitations, especially when applied to highly volatile cryptocurrency markets: First, the two points fitting method requires simultaneously satisfying continuity conditions at two joining points, which not only increases computational complexity but may also affect the stability of estimation results, causing kinks in the fitting; Second, in the cryptocurrency market, the application of RND is still in its early stages, and its characteristics of dramatic price fluctuations, investor composition, and behavioral patterns pose challenges to traditional RND estimation methods (López-Cabarcos et al., 2021); Furthermore, existing research has less explored the performance differences of RND characteristics across different prediction periods.

Based on the above research gaps, our proposed approach introduces the innovative single point with the slope fitting method, requiring continuity conditions for both the cumulative distribution function value and the density function slope at a single joining point, using GPD (Birru & Figlewski, 2012; Y. He et al., 2022) for fitting. Compared to the two points fitting method, our proposed method has the following advantages: First, it reduces the constraints of fitting conditions, improving the applicability of the method in extreme market situations; Second, it simplifies the calculation process, enhancing estimation efficiency; Third, by maintaining the continuity of the density function slope, it ensures the smoothness of the fitting results, avoiding unnatural kinks at joining points.

Higher-order moments of RND, especially skewness and excess kurtosis, play key roles in predicting financial market returns. Research by Bali and Murray (2013) and Conrad et al. (2013) both found a significant negative relationship between risk-neutral skewness and future stock returns, consistent with the theoretical expectation that investors prefer positive skewness. Kim and Park (2018) further confirmed that even after controlling for multiple firm characteristic variables, the negative correlation between option-implied skewness and subsequent stock returns still holds. In commodity futures markets, research by Fuertes et al. (2022) showed that trading strategies constructed using risk-neutral skewness (RNSK) can generate significant excess returns, particularly when futures markets are in contango. Cortés et al. (2020) research on the oil market found that the log-SNP distribution can more accurately capture the RND characteristics of oil prices, with skewness and kurtosis containing important information related to market expectations—negative skewness indicates a higher probability of expected price declines, while high narrow peaks suggest the possibility of extreme price changes. Chen et al. (2018) further found that higher-order moments of risk-neutral probability density functions have significant explanatory power in predicting crash risk and risk premiums, with skewness highly positively correlated with risk premiums and kurtosis positively correlated with foreign exchange swap spreads. Recently, Böök et al. (2025) proposed a new method for estimating the volatility smile, deriving conditional volatility, skewness, and kurtosis indicators from S&P 500 index short-term options data. These indicators perform excellently in predicting U.S. stock risk premiums, outperforming equivalent indicators calculated based on historical returns in both in-sample and out-of-sample tests. While the predictive power of these higher-order moments has been empirically supported in traditional financial markets, research in the more volatile cryptocurrency market remains limited. Particularly in crypto assets like Bitcoin, investor structure and behavioral patterns differ significantly from traditional markets, with market sentiment shifts occurring more rapidly and extremely. This research will estimate the higher-order moments of Bitcoin options RND to understand the pricing mechanisms and risk characteristics of the cryptocurrency market and attempt to find forward-looking indicators for predicting returns.

To validate the effectiveness of our proposed method, this research uses Bitcoin options trading data from the Deribit platform from January 2020 to April 2024, examining options samples with one day and seven days to expiration. The empirical analysis primarily focuses on the predictive power of RND statistical characteristics (such as skewness, excess kurtosis, etc.) for Bitcoin returns, comparing the results of the single point with the slope fitting method with the traditional two points fitting method.

In terms of daily return prediction, we find that models constructed using the single point with the slope fitting method identify skewness, median, and lagged returns as significant predictive variables, with explanatory power (R-squared = 0.0130) superior to the two points fitting method (R-squared = 0.0098). Among these, the skewness coefficient shows a negative correlation with returns, consistent with the findings of Conrad et al. (2013) in the stock market, indicating that the market prices negative risk higher than positive risk. The significant predictive power of lagged returns supports Liu and Tsyvinski's (2021) argument about the existence of momentum effects in the cryptocurrency market.

In weekly return prediction, the performance of both methods tends to be consistent, with the best combination of predictive variables being excess kurtosis, median, and the Crypto Fear and Greed Index. Excess kurtosis shows a positive correlation with returns, consistent with the findings of Amaya et al. (2015), indicating that the pricing of extreme risk is reflected in future returns. The significant influence of the Crypto Fear and Greed Index aligns with the research of He et al. (2023), confirming the important role of market sentiment in the price formation process of cryptocurrencies.

To test the robustness of the prediction models, we conducted out-of-sample prediction tests using the rolling window method for parameter estimation and prediction, a method widely adopted in financial forecasting research (Campbell & Thompson, 2008; Welch & Goyal, 2008). The results show that both methods produce positive out-of-sample R-squared values for both daily and weekly return predictions, with the single point with the slope fitting method outperforming the two points fitting method. According to Campbell and Thompson's (2008) research, a positive out-of-sample R-squared indicates that the prediction model has substantial economic value and can provide useful reference information for investment decisions.

The methodological innovation of this research is not merely to improve computational efficiency. More importantly, our single point with the slope fitting method can more accurately capture the extreme risk characteristics implied by the Bitcoin options market, providing investors with more reliable risk assessment tools. Research by Chang et al. (2013) and Conrad et al. (2013) has confirmed that higher-order moments of RND (such as skewness and kurtosis) have significant predictive power for asset returns.

However, whether these conclusions apply to the structurally different cryptocurrency market and different prediction time scales remains a question that requires in-depth exploration.

Furthermore, the cryptocurrency market has unique investor composition and behavioral patterns. Research shows that compared to traditional financial markets, cryptocurrency investors are more susceptible to emotional and irrational factors (M. He et al., 2023). Therefore, incorporating investor sentiment indicators (such as the Crypto Fear and Greed Index) into RND characteristic prediction models may reveal market dynamics that traditional financial theories struggle to explain.

The main contributions of this research can be divided into two aspects:

1. Methodological breakthrough: Our proposed single point with the slope fitting method reduces the calculation of joining reference points compared to the traditional two points fitting method, and ensures the smoothness of tail fitting through the continuity condition of the first derivative, significantly enhancing the stability and efficiency of tail fitting while maintaining theoretical consistency. Empirical results show that the single point with the slope fitting method rarely produces kinks when fitting RND tails, and the computational speed is on average 10.95% faster than the two points fitting method.

2. Multi-time scale prediction framework: By comparing daily and weekly return prediction models, this research reveals the differential predictive capabilities of RND characteristics across different time scales. We find that in daily return prediction, skewness (Conrad et al., 2013), median, and lagged returns (Liu & Tsyvinski, 2021) have significant predictive power; while in weekly return prediction, excess kurtosis (Amaya et al., 2015), median, and the Crypto Fear and

Greed Index (M. He et al., 2023) become key predictive factors. The findings of this research fill the research gap in the existing literature regarding return prediction across different time scales in the cryptocurrency domain.

As institutional investors continue to enter the cryptocurrency market, the demand for professional risk management tools will continue to grow. This research not only responds to this market demand but also opens new directions for academic research in the cryptocurrency derivatives market. By establishing a more robust RND estimation framework, we hope to promote a deeper understanding of the nature of risk in cryptocurrency markets and lay a solid foundation for subsequent research.

1.2 Research Structure

This research is divided into six chapters, with the structure as follows:

"Chapter 1: Introduction": Explains the research background, motivation, and contributions.

"Chapter 2: Literature Review": Explores the theoretical development of risk-neutral density (RND), estimation methods, and related empirical research.

"Chapter 3: Data Sources": Introduces research data sources and market overview.

"Chapter 4: Research Methodology": Details the research methods, including risk-neutral density calculation, tail fitting methods, and empirical regression model design.

"Chapter 5: Empirical Results Analysis": Analyzes regression results and compares performance differences between different fitting methods.

"Chapter 6: Conclusion and Recommendations": Summarizes the empirical results of this research and outlines directions for improvement in subsequent research.

2. Literature Review

2.1 Fundamental Theory of Risk-neutral Density

The concept of Risk-neutral Density (RND) was first proposed by Breeden and Litzenberger (1978). They demonstrated that market expectations regarding the future price distribution of underlying assets could be extracted from option prices. This pioneering research opened a new direction for option pricing theory. Shimko (1993) further proposed a technique of converting market option prices into implied volatility space for interpolation. This method leverages the characteristic that implied volatility curves are smoother and more continuous than price data, helping to improve the accuracy of RND estimation.

Christoffersen et al. (2013) conducted an in-depth exploration of the predictive power of option-implied information, discovering that risk-neutral skewness can effectively predict the direction and magnitude of future returns. Chang et al. (2013) approached from another angle, studying the relationship between risk-neutral kurtosis and market returns, indicating that higher risk-neutral kurtosis often presages greater future market volatility.

In extreme market scenarios, Birru and Figlewski (2012) conducted research on S&P 500 index options during the 2008 financial crisis, finding that RND shapes undergo significant changes as market stress increases. To more accurately describe such extreme situations, they adopted the Generalized Pareto Distribution (GPD) to refine RND tail estimation, a method that demonstrates excellent performance when handling extreme market conditions (Hosking & Wallis, 1987). Jackwerth (2020) further analyzed the risk-neutral probabilities of S&P 500 index options, discovering that markets often require a certain amount of time to fully reflect the impact of major events, providing important empirical evidence for understanding market information efficiency.

2.2 Estimation Methods for Risk-neutral Density

Figlewski (2008) conducted in-depth research on U.S. market portfolios, proposing a comprehensive framework for risk-neutral density estimation, categorizing it into parametric and non-parametric methods, providing important references for practical applications. McNeil and Frey (2000) proposed an innovative method combining GARCH models with Extreme Value Theory (EVT) to estimate tail risk in financial time series. They particularly emphasized that the Generalized Pareto Distribution (GPD) has unique advantages in modeling extreme events in financial time series.

Orosi (2015) proposed a novel parametric method for estimating risk-neutral probability density functions. By setting appropriate functional forms and imposing constraints on model parameters, this method can produce risk-neutral density estimates with good properties. He et al. (2022) recommended using GPD to estimate RND tails, a method particularly effective when dealing with extreme market conditions. Recently, Uberti (2023) proposed a new semi-parametric estimation method combining the stability of parametric methods with the flexibility of non-parametric methods, offering a new research direction for RND estimation and demonstrating the continuous refinement trend in risk-neutral density estimation methodology.

Ammann and Feser (2019) conducted in-depth research on robust estimation methods for risk-neutral moments, proposing improved methods that effectively reduce estimation bias in situations of market noise, extreme price movements, or insufficient liquidity. Hayashi (2020) proposed an innovative method for analyzing risk-neutral probability density functions from volatility smiles, a method that can completely avoid approximation errors that might arise from numerical methods.

Recently, Dong et al. (2024) proposed the innovative Implied Willow Tree (IWT) method, addressing research gaps in reconstructing risk-neutral stochastic processes. Unlike traditional methods that only reconstruct risk-neutral probability density functions, this research directly reconstructs the complete risk-neutral process from cross-maturity market option price data without relying on any preset parametric models. Empirical evidence demonstrates the effectiveness of this method in pricing American options and Asian options, as well as its good capability in handling noise in the original data.

2.3 Research on Risk-neutral Distribution Characteristics

2.3.1 Characteristics and Applications of Higher-Order Moments

Higher-order moments of risk-neutral distributions, particularly skewness and kurtosis, play crucial roles in financial market research. Studies by Bali and Murray (2013) and Conrad et al. (2013) demonstrate that risk-neutral skewness exhibits a significant negative relationship with future stock returns, consistent with the theoretical expectation that investors prefer positive skewness. Empirical research by Kim and Park (2018) further confirms that option-implied skewness is significantly negatively correlated with subsequent stock returns, a relationship that persists even after controlling for multiple firm characteristic variables.

Mei et al. (2017) investigate the impact of realized skewness and realized kurtosis on stock market volatility prediction. Out-of-sample prediction results indicate that both have significant negative effects on future volatility, with realized skewness outperforming realized kurtosis in medium to long-term prediction horizons, while showing limited effectiveness in short-term predictions.

Fuertes et al. (2022) find that risk-neutral skewness (RNSK) plays an important role in commodity futures price prediction. Strategies that involve buying futures with positive RNSK values and selling those with negative RNSK values generate significant excess returns, particularly when futures markets are in contango, providing new perspectives for asset allocation and risk management in commodity futures markets.

Cortés et al. (2020) examine the efficacy of using semi-parametric methods to derive implied risk-neutral density functions from West Texas Intermediate crude oil options. Results indicate that, compared to traditional log-normal distributions, log-SNP distributions more accurately capture the RND of oil prices, with skewness and kurtosis containing important information related to oil market expectations. When the market exhibits negative skewness, it indicates a higher probability of expected declines in the underlying price; high leptokurtosis suggests a higher probability of extreme price changes.

In recent research, Böök et al. (2025) propose a new method for estimating volatility smiles, deriving robust conditional volatility, skewness, and kurtosis indicators from S&P 500 index short-term options data. These indicators demonstrate excellent performance in predicting U.S. equity risk premiums and market higher-order moments, outperforming equivalent indicators calculated from historical returns in both in-sample and out-of-sample tests.

2.3.2 Tail Risk Characteristics and Extreme Value Theory

Regarding foundational research on extreme value theory, the threshold exceedance model proposed by Balkema and de Haan (1974) establishes an important foundation for subsequent research on extreme events in financial markets. Their research demonstrates that when samples exceed a sufficiently high threshold, their distribution converges to the Generalized Pareto Distribution, a finding with critical significance for estimating tail risks in financial markets.

Wang 和 Yen (2018) demonstrate that option-implied tail risk indicators constructed based on extreme value theory effectively predict future movements of underlying assets. Empirical results reveal that tail risks implied by S&P 500 index and VIX options have significant predictive power for future returns, particularly pronounced during economic recessions. Chen et al. (2018) conduct an in-depth investigation into the relationship between crash risk and risk-neutral probability density functions, finding that higher-order moments of risk-neutral probability density functions possess significant explanatory power in predicting and explaining crash risk and risk premiums. Skewness exhibits a high positive correlation with risk premiums, while kurtosis shows a positive correlation with foreign exchange swap spreads, with higher kurtosis representing an increased probability of extreme market events.

Lehnert (2022) explores the relationship between risk-neutral skewness and hedge fund tail risk. Empirical results indicate that short-selling behavior in index options markets leads to a negative relationship between risk-neutral market skewness and returns. This research challenges the traditional view of option-implied skewness as a downside risk indicator, providing new perspectives for understanding market risk preferences.

Conrad et al. (2013) further extend methods combining GARCH models with mixed normal distributions, particularly suitable for capturing asymmetric volatility characteristics in financial markets. Neumann and Skiadopoulos (2013) conduct in-depth research on higher-order risk-neutral moments of S&P 500 options, finding that market expectations regarding volatility, skewness, and kurtosis exhibit significant predictability, especially during periods of greater market volatility.

2.4 Empirical Applications in Financial Markets

2.4.1 Traditional Financial Markets

Mohrschladt and Schneider (2021) analyze the differences in implied volatility between in-the-money and out-of-the-money options using high-frequency trading data, finding that in-the-money options contain important market information. Research by Li et al. (2024) indicates that risk-neutral skewness has significant predictive power for future stock returns, particularly pronounced during economic recessions, a finding consistent with the theoretical predictions of Cujean and Hasler (2017).

The latest research by Feng et al. (2024) reveals a strong association between market sentiment and option-implied volatility, especially during periods of higher market uncertainty. Köse et al. (2024) study the influence of trader behavior on option price formation from a market microstructure perspective, finding that institutional investors' trading behavior significantly impacts RND shape.

Regarding return prediction, empirical research by Amaya et al. (2015) finds that realized skewness has significant explanatory power for cross-sectional stock returns, with this predictive capability remaining significant after controlling for other risk factors. Bali and Zhou (2016) approach from the perspective of risk and uncertainty, finding a significant association between market uncertainty and expected returns, particularly during periods of higher macroeconomic uncertainty. The pioneering research by Campbell and Thompson (2008) proposes a framework for evaluating the economic value of predictive models and confirms that even seemingly minor predictive capabilities can yield considerable economic benefits.

Jondeau et al. (2020) investigate the predictive power of individual stock skewness for index futures returns. Empirical results demonstrate that this indicator exhibits significant effectiveness in predicting S&P 500 index futures returns. This predictive relationship persists even after controlling for liquidity risk and economic cycle factors, and individual skewness also performs excellently in out-of-sample prediction of index futures returns.

2.4.2 Cryptocurrency Market

In the cryptocurrency market, Zulfiqar and Gulzar (2021) find that options trading provided by major cryptocurrency exchanges offers investors more diverse hedging instruments. Baur and Smales (2022) conduct in-depth research on trading behavior in the Bitcoin futures market, discovering that leveraged fund traders play a key role in the Bitcoin futures market, not only holding the largest positions but typically maintaining a net short position. These traders demonstrate an ability to accurately predict the largest market fluctuations.

López-Cabarcos et al. (2021) conduct an in-depth analysis of the relationship between investor sentiment and price volatility in the Bitcoin market, finding that social media sentiment indicators have significant predictive power for short-term price movements. This research emphasizes the importance of sentiment factors in cryptocurrency market pricing. Research by Chordia et al. (2021) indicates that the RND of the Bitcoin options market often exhibits significant left skewness and excess kurtosis. Akyildirim et al. (2020) further find that the cryptocurrency market exhibits higher volatility when investor fear sentiment rises, showing significant correlation with traditional market volatility indicators.

In research on the dynamic characteristics of cryptocurrency markets, Liu and Tsyvinski (2021) study the risk and return characteristics of cryptocurrency markets, finding significant momentum effects. Li et al. (2021) further confirm that the MAX momentum effect is particularly significant in cryptocurrency markets. Specifically, cryptocurrencies that have performed best in the past tend to maintain better performance in future periods, a phenomenon similar to the momentum effect in traditional financial markets but with greater intensity.

Liu and Chen (2024) find that cryptocurrencies with larger market capitalizations exhibit left-skewed characteristics, while those with smaller market capitalizations show right skewness, and asymmetric risk is negatively correlated with future returns, meaning cryptocurrencies with lower skewness typically generate higher returns. Liu et al. (2023) use machine learning methods to predict cryptocurrency returns, confirming that lagged returns possess strong predictive power.

3. Data Sources

3.1 Data Sources

This research utilizes historical trading data provided by Deribit exchange (*Deribit*, 2025) the world's largest Bitcoin options trading platform, as the research sample. The research period covers January 2020 to April 2024, with collected data including daily trading volume, closing prices, implied volatility, spot prices, and futures prices among other trading information. The selection of Deribit exchange as the research subject is primarily based on its leadership position in the global cryptocurrency options market.

According to statistics from cryptocurrency information service provider The Block (*The Block*, 2025), among the three major trading platforms (Deribit, OKX, and Binance), Deribit not only possesses the largest trading volume, but its open interest also accounts for over 80% of the overall market share (as shown in Figure 3-1). This advantageous position mainly stems from its long operating history and robust market development strategy.

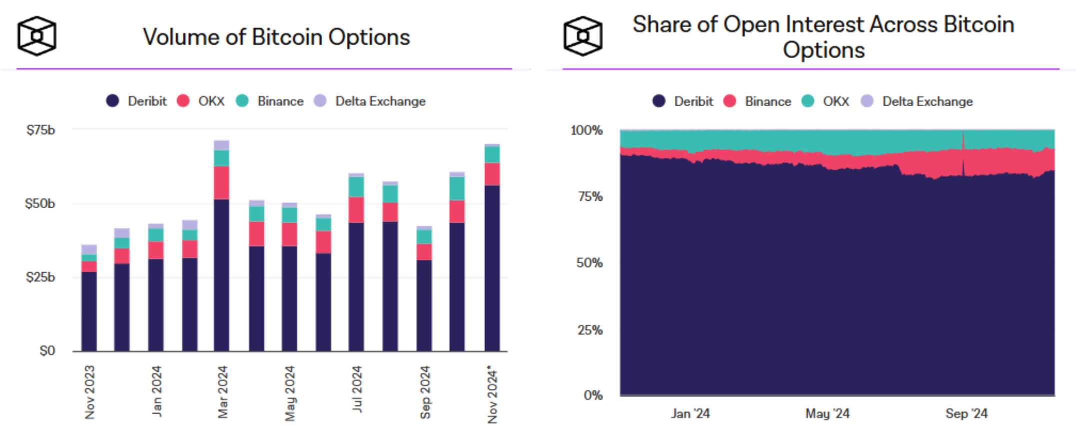

Figure 3-1: Statistics on Bitcoin Options Market Trading Volume and Open Interest
(Data Source: The Block Official Website)

Deribit was established in 2016 with headquarters in the Netherlands. Its name is a combination of "Derivatives" and "Bitcoin," making it the world's first professional exchange to launch cryptocurrency options products.

In terms of product design, Bitcoin European cash-settled options provided by Deribit adopt a 24/7 trading mechanism (*Deribit Options*, 2025), with expiration times uniformly set at 08:00 Coordinated Universal Time (UTC+0). Regarding expiration date selection, the exchange offers diversified product combinations, including: short-term contracts (1-day, 2-day, 3-day), medium-term contracts (1-week, 2-week, 3-week), month-end expiration long-term contracts (January, February, April, May, July, August, October, November), and quarterly expiration long-term contracts (March, June, September, December).

This diversified product design not only satisfies the trading needs of different investors but also helps enhance market liquidity and price discovery efficiency. Particularly after October 2020, Deribit added daily and weekly expiration products, significantly increasing market participation, with trading volume growing markedly as a result.

3.2 Overview of Bitcoin Options Trading Market

This research will conduct empirical analysis through historical Bitcoin options trading data, exploring the relationship between implied volatility in the options market and expected price movements, and calculating the Risk-neutral Density (RND) function to analyze its changes under different market atmospheres.

In terms of transaction counts (Figure 3-2) and trading volume (Figure 3-3), a significant growth trend has been observed since October 2020. This turning point coincides with Deribit's product line expansion, including the addition of daily and weekly expiration products, demonstrating that the product diversification strategy has effectively enhanced market activity. Furthermore, observing market performance in the second half of 2023, trading volume again showed substantial growth momentum, particularly after September 2023, with call option trading volume reaching a historical high in February 2024 alongside the strong rise in Bitcoin spot prices. This phenomenon reflects market participants' continued optimism about Bitcoin's future, with investors preferring to use options products for speculation or hedging operations. Regarding market depth, the continuous rise in trading volume since the second half of 2023 not only reflects increased market participation but also indicates that the liquidity and depth of the Bitcoin options market have matured. This improvement in market structure helps reduce transaction costs, enhance price discovery efficiency, and consequently attract more institutional investors. The Bitcoin options market demonstrates clear growth momentum and structural improvement, providing a rich and representative data foundation for subsequent research.

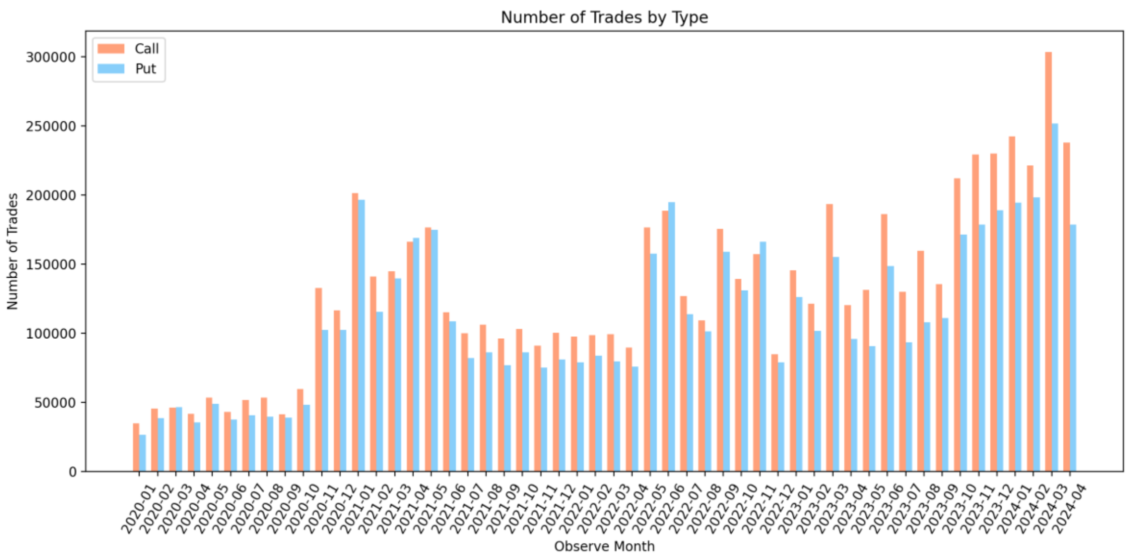

Figure 3-2: Monthly Transaction Counts of Bitcoin Call and Put Options from January 2020 to April 2024

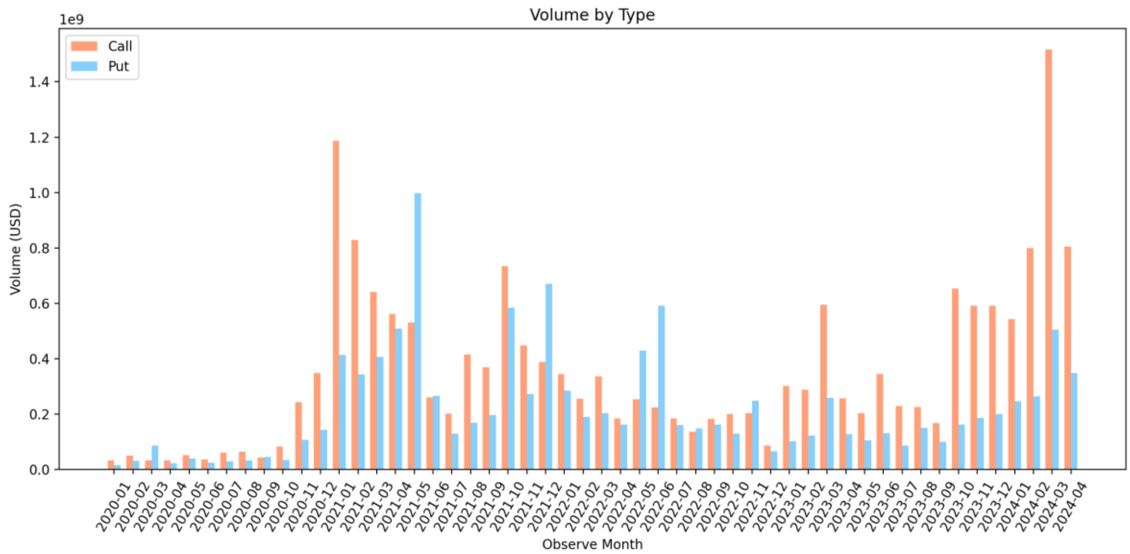

Figure 3-3: Monthly Trading Volume of Bitcoin Call and Put Options from January 2020 to April 2024

This research analyzes trading patterns in the Bitcoin options market through total trading volume heat maps. Regarding the call options market (Figure 3-4), trading activities are notably concentrated around the at-the-money area, with the most active trading occurring at Moneyness $\left(\frac{StrikePrice(K)}{SpotPrice(S)}\right)$ between 0.9 and 1.1. Slightly out-of-the-money (1.0-1.1) call options recorded the highest trading volume of 2,072 million USD, indicating investors' preference for using out-of-the-money call options with greater leverage effects for trading. In contrast, deep out-of-the-money (greater than 1.5) call options show relatively sparse trading volume, indicating market participants' lack of interest in call options with excessively high strike prices.

Regarding the term structure, short-term call options market exhibits higher trading activity, particularly call options with 31 to 90 days to expiration maintaining considerable transaction volumes across various ratio intervals. Notably, extremely short-term (within 14 days) call options still maintain substantial trading volume near at-the-money, reflecting significant short-term speculative trading demand in the market. On the other hand, long-term (over 180 days) call options trading is relatively thin, except for an abnormal trading volume of up to 392 million contracts in the deep out-of-the-money area (greater than 2.0), which is likely related to specific investment strategies or institutional hedging needs.

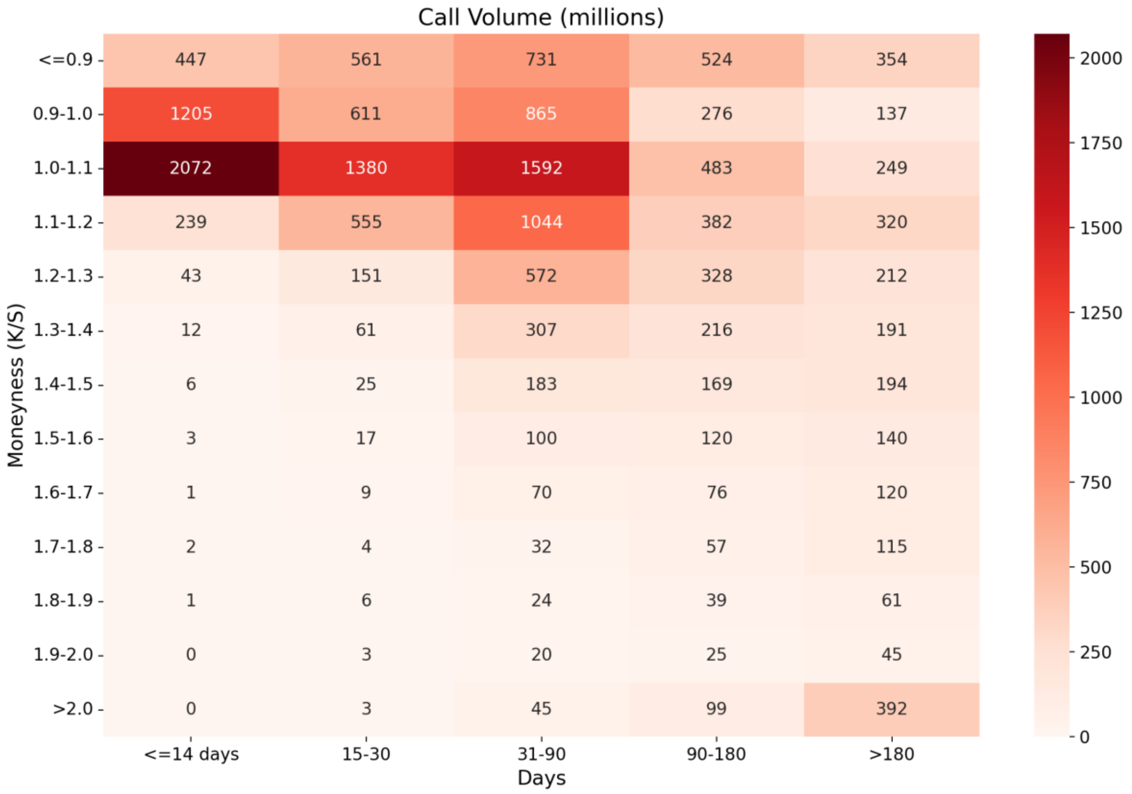

Figure 3-4: Heat Map of Total Bitcoin Call Options Trading Volume from January 2020 to April 2024

The put options market (Figure 3-5) exhibits distinctly different trading characteristics. Trading activities are highly concentrated near at-the-money (0.9-1.0), with the highest single block trading volume reaching 1,676 million USD, far exceeding other areas. Deep out-of-the-money (less than or equal to 0.7) put options show relatively sparse trading volume, indicating relatively limited market demand for hedging against significant Bitcoin price declines. Regarding expiration period distribution, short-term put options (within 14 days and 15-90 days) are most active near at-the-money, while medium to long-term (90-180 days) put options show significant trading volume (821 million USD) in the deep in-the-money (greater than 1.1) area. The trading volume distribution of extremely long-term (over 180 days) put options is relatively even, maintaining a certain level of transaction volume across various ratio intervals.

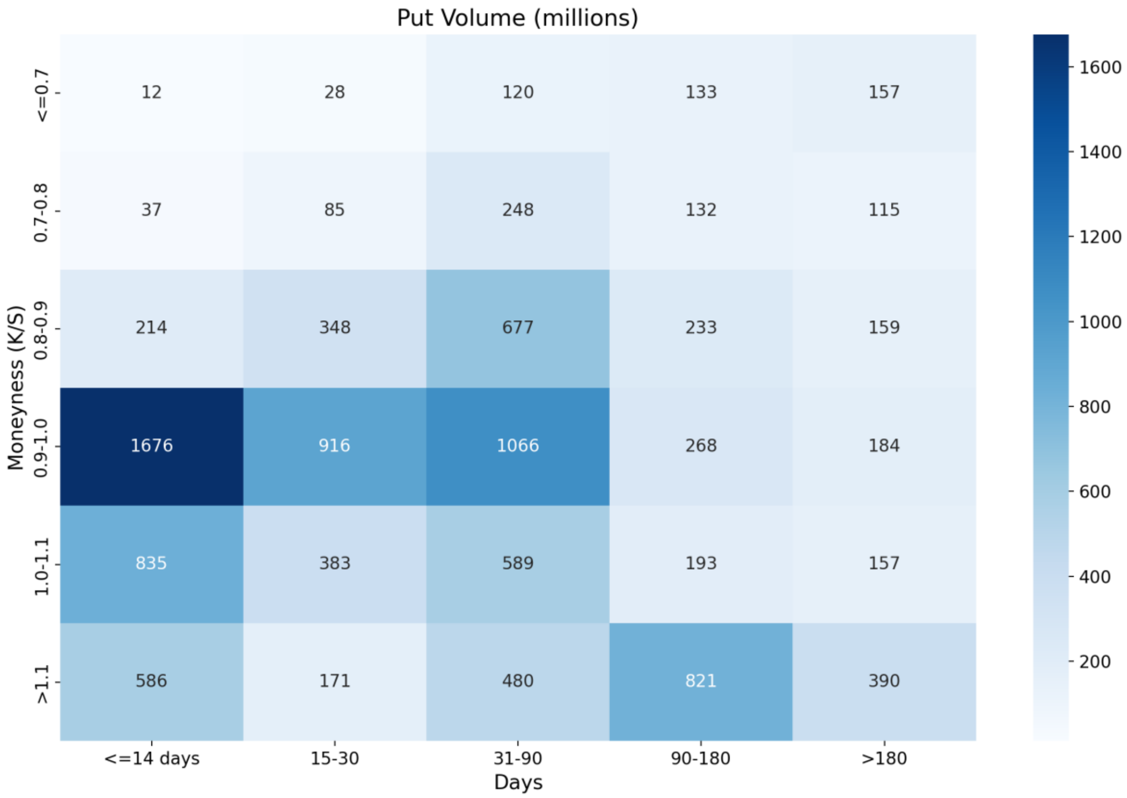

Figure 3-5: Heat Map of Total Bitcoin Put Options Trading Volume from January 2020 to April 2024

The Bitcoin options market has developed considerable depth and liquidity, with trading activities primarily concentrated around the at-the-money area, reflecting market participants' preference for more efficient speculative or hedging operations. Short-term options are generally more favored, indicating market participants' tendency to adopt short-term trading strategies. Moreover, the phenomenon of overall call options trading volume exceeding put options may reflect the market's bullish outlook on Bitcoin price trends. Abnormal trading volumes in specific areas (such as deep out-of-the-money call options and deep in-the-money put options) may be closely related to specific investment strategies or institutional hedging needs. The aforementioned trading characteristics collectively constitute the basic profile of the current Bitcoin options market, which, although having reached a considerable scale, still primarily features short-term speculative trading patterns.

4. Research Methodology

4.1 Method for Calculating Risk-neutral Probability Density Base on Theory

In the following text, symbols C, P, S, K, and T represent standard option meanings: C is call option price, P is put option price, S is underlying asset current price, K is strike price, r is risk-free rate, T is days to option expiration. This research will also use $f(K)$ to represent the Risk-neutral Probability Density Function (RND) and $F(K) = \int_{-\infty}^{K} f(z)dz$ to represent the Risk-neutral Distribution Function.

The call option price is the expected payoff before its expiration day T, discounted back to the present value. Under risk-neutral conditions, this expected price can be calculated based on risk-neutral probability and discounted using the risk-free rate, as follows:

$$C = \int_{K}^{\infty} e^{-rT} (S_T - K) f(S_T) dS_T \tag{1}$$

Next, by taking the first partial derivative of call option price C with respect to strike price K, we can derive the risk-neutral distribution function $F(K)$, as follows:

$$\begin{aligned}
\frac{\partial C}{\partial K} &= \frac{\partial}{\partial K}\left[\int_{K}^{\infty} e^{-rT} (S_T - K) f(S_T) dS_T\right] \\
&= e^{-rT}\left[-(K-K)f(K) + \int_{K}^{\infty} -f(S_T) dS_T\right] \\
&= -e^{-rT}\int_{K}^{\infty} f(S_T) dS_T \\
&= -e^{-rT}[1 - F(K)]
\end{aligned}$$

Rearranging terms, we obtain the risk-neutral distribution function $F(K)$:

$$F(K) = e^{rT}\frac{\partial C}{\partial K} + 1 \tag{2}$$

Then, by taking another partial derivative of equation (2) with respect to strike price K, we can derive the RND at strike price K:

$$f(K) = \frac{\partial}{\partial K}\left[e^{rT}\frac{\partial C}{\partial K} + 1\right] = e^{rT}\frac{\partial^2 C}{\partial K^2} \tag{3}$$

In actual options trading markets, since strike prices are in discrete form, we can use observed option prices and apply Finite Difference Methods (FDM) to obtain approximate solutions to equations (2) and (3). Assuming that at time T to expiration, there are N options with different strike prices in the market, where K1 represents the lowest strike price and K_N represents the highest strike price. We will use three options with strike prices K_{n-1}, K_n and K_{n+1} to calculate the approximate value centered at K_n, as follows:

$$F(K_n) \approx e^{rT}\left[\frac{C_{n+1} - C_{n-1}}{X_{n+1} - X_{n-1}}\right] + 1 \tag{4}$$

$$f(K_n) \approx e^{rT}\frac{C_{n+1} - 2C_n + C_{n-1}}{(\Delta X)^2} \tag{5}$$

Equations (1) to (5) explain how to theoretically derive the RND between strike prices K_2 and K_{N-1} from a set of call option prices C. Similar derivation methods can also be applied to calculate RND from put option prices P. For put options, the equivalent expressions corresponding to equations (2) to (5) are as follows:

$$F(K) = e^{rT}\frac{\partial P}{\partial K} \tag{6}$$

$$f(K) = e^{rT} \frac{\partial^2 P}{\partial K^2} \tag{7}$$

$$F(K_n) \approx e^{rT} \left[\frac{P_{n+1} - P_{n-1}}{X_{n+1} - X_{n-1}}\right] \tag{8}$$

$$f(K_n) \approx e^{rT} \frac{P_{n+1} - 2P_n + P_{n-1}}{(\Delta X)^2} \tag{9}$$

In this research, ΔX is a fixed constant value used to construct artificially spaced option prices to fill gaps between discrete strike prices in the market. This approach addresses the problem of sparse or uneven trading data and ensures consistent spacing between strike prices, facilitating numerical calculations through finite difference methods and improving the accuracy of estimation results.

4.2 Practical Methods for Calculating Risk-neutral Probability Density

The method introduced in the previous section assumes the existence of a set of option prices that perfectly conform to theoretical pricing relationships (Equation (1)). However, when applied to option prices traded in actual markets, several important issues and challenges arise. First, market imperfections in observed option prices must be carefully addressed, otherwise the derived RND may exhibit unacceptable characteristics, such as negative values in certain regions. Second, appropriate methods must be found to complete the tails of the RND outside the range from K_2 to K_{N-1}. This section introduces methods proposed in this research and reviewed literature for calculating RND from market option prices, and explains the techniques we adopt here.

4.2.1 Bitcoin-Applicable Black-Scholes Model

In traditional financial markets, option pricing models (such as the Black-Scholes model) typically use risk-free interest rates as parameters, which are generally represented by the yields of low-risk assets like government bonds. However, in cryptocurrency markets like Bitcoin, the applicability of risk-free interest rates is limited and therefore not widely used. This is because the Bitcoin market lacks a unified risk-free asset. Due to the decentralized nature of cryptocurrency markets, there are no widely accepted risk-free assets like government bonds, making it difficult to determine a single universal risk-free rate, thus limiting its applicability in this market.

Additionally, Bitcoin price volatility is far higher than traditional assets. This highly volatile characteristic has a more significant impact on option prices than risk-free interest rates, causing traders to focus more on changes in implied volatility rather than risk-free rates. Furthermore, in cryptocurrency markets, the interest rate environment may be influenced by exchange rules and market supply and demand, not necessarily related to traditional risk-free rates, making traditional interest rate indicators difficult to reflect the actual situation in cryptocurrency markets. Moreover, the cost of holding Bitcoin differs from the cost of holding traditional currencies or assets, including security aspects and technological risks, which are difficult to quantify through risk-free interest rates, further limiting the applicability of risk-free rates in Bitcoin option pricing.

This research uses Bitcoin option trading prices from the Deribit exchange, which has adopted a more suitable model (priced in Bitcoin) to calculate option prices, adapting to the characteristics of the cryptocurrency market and meeting trading market needs. To meet research requirements, this study observes the traditional Black-Scholes model (Equation (10)) and compares it with the calculation formula provided by Deribit exchange (Equation (11)). It can be seen that multiplying the Deribit exchange quote $C_{Deribit}$ by the Bitcoin spot price S_0 yields the Bitcoin option price denominated in US dollars.

$$C_{BS} = S_0 \times N(d_1) - Ke^{-rT} \times N(d_2)$$

$$\Rightarrow C_{BS} = S_0 \times \left[N(d_1) - \frac{Ke^{-rT}}{S_0} \times N(d_2) \right] \; and \; F = S_0 e^{rT}$$

$$\Rightarrow \frac{C_{BS}}{S_0} = N(d_1) - \frac{K}{F} \times N(d_2) \tag{10}$$

$$C_{Deribit} = N(d_1) - \frac{K}{F} \times N(d_2) \tag{11}$$

$$where \; d_1 = \frac{\ln\left(\frac{F}{K}\right) + \left(\frac{\sigma^2}{2}\right) \times T}{\sigma \times \sqrt{T}}$$

$$d_2 = d_1 - \sigma \times \sqrt{T}$$

Where C_{BS} is the Black-Scholes call option price (denominated in USD), $C_{Deribit}$ is the Deribit exchange Bitcoin call option price (denominated in Bitcoin), S_0 is the Bitcoin spot price, N(x) is the cumulative distribution function of the normal distribution, K is the strike price, r is the risk-free interest rate, T is the option's time to expiration, F is the Bitcoin futures price, ln is the natural logarithm, σ is the annualized standard deviation.

4.2.2 Calculation of Bitcoin Option Implied Volatility

In the Bitcoin options market, traders are predominantly risk-seeking and tend to operate with out-of-the-money (OTM) options, primarily due to their lower cost, high leverage effect, and particularly strong sensitivity to volatility. For buyers, given the extremely high volatility of the Bitcoin market itself, these options are highly attractive to speculators and high-risk-preferring investors. Despite the higher probability of these options expiring worthless, traders are still willing to take on such risks. For sellers, since Bitcoin price volatility is significantly higher than traditional financial markets, the premium levels for OTM options are usually higher, further enhancing the incentives for seller participation, making OTM options a core tool for many sellers to create stable cash flow. In summary, OTM options have better trading volume and liquidity, and their prices can more efficiently reflect market sentiment. This research will use OTM option trading data to calculate implied volatility; however, to avoid anomalies caused by unreasonable trades in deeply out-of-the-money areas, options data with strike prices below \$10 will be excluded.

Shimko (1993) proposed converting market option prices to implied volatility (IV) before interpolation, as implied volatility curves are typically smoother and more continuous than price data, making them suitable for interpolation and smoothing processes. The interpolated curves are then converted back to call option prices to calculate RND. This method does not rely on option prices conforming to Black-Scholes model assumptions, but rather uses the Black-Scholes formula merely as a calculation tool to convert data into a form more suitable for smoothing.

Figlewski (2008) proposed a method aimed at resolving abnormal fluctuations in implied volatility data near at-the-money options, particularly the jump phenomenon in call and put option prices near the at-the-money point. Such jumps may lead to non-smooth implied volatility curves, thereby affecting the stability of RND calculation results. Following this method, if a strike price is between 0.9 and 1.1 times the futures price, the average of call and put implied volatilities is taken as the data point. The formula is as follows:

$$IV_{mix}(K) = 0.5 \times IV_{call}(K) + 0.5 \times IV_{put}(K)$$

$$K \in [0.9 \times F, 1.1 \times F]$$

Figure 4-1 presents the relationship between Bitcoin option implied volatility and strike price, showing that the market has a higher preference for OTM options. At the same time, significant fluctuations in implied volatility can be observed near the at-the-money point. To reduce the dramatic volatility of implied volatility at the at-the-money position, this research takes the average value of call and put implied volatilities for data within the range of 0.9 to 1.1 times the futures price as the basis for subsequent smoothing. As shown in Figure 4-2, the green marked points represent the averaged implied volatility, which effectively reduces the fluctuation amplitude of call and put implied volatility near the at-the-money point. For data outside this range, the implied volatility data of OTM options are directly adopted. After the above processing, the final constructed implied volatility data are shown in Figure 4-3.

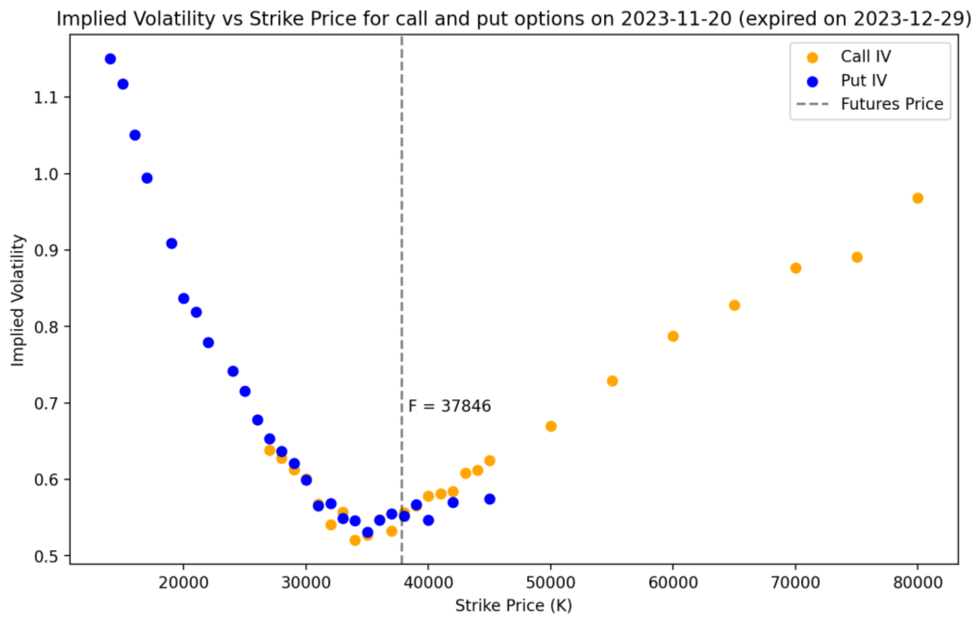

Figure 4-1: Bitcoin Option Implied Volatility Distribution on November 20, 2023 (Expiring December 29, 2023)

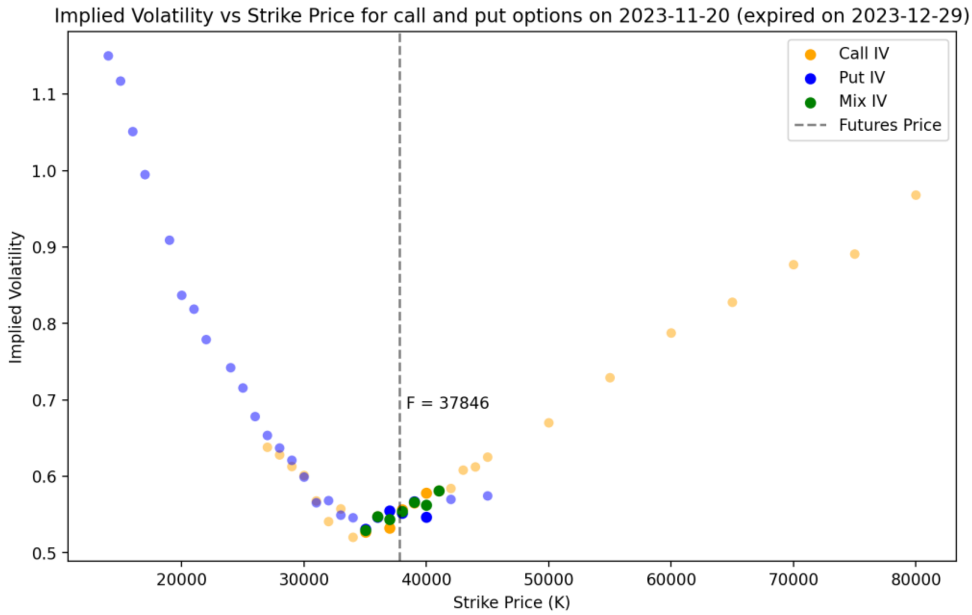

Figure 4-2: Bitcoin Option Implied Volatility Distribution on November 20, 2023 (Expiring December 29, 2023)

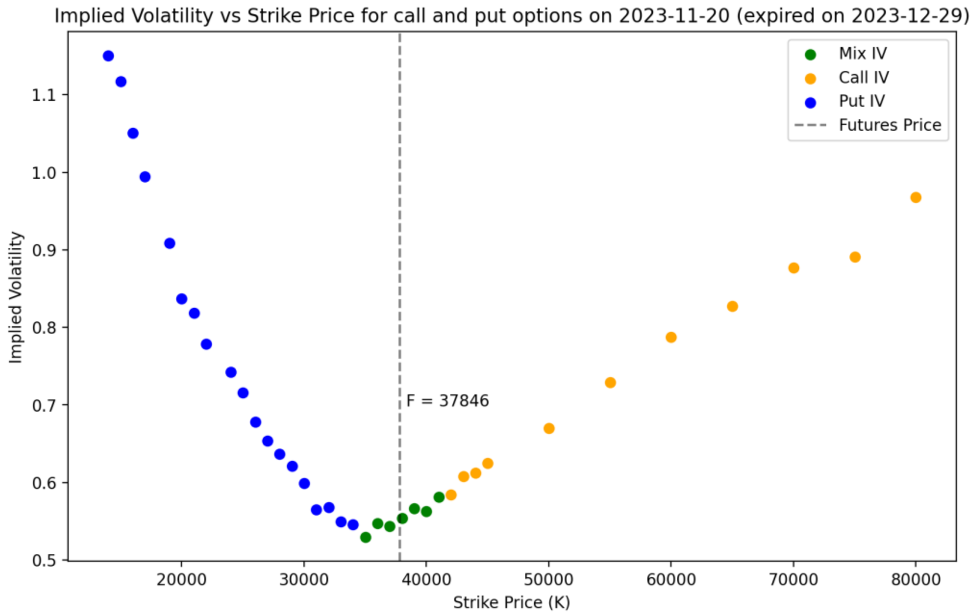

Figure 4-3: Bitcoin Option Implied Volatility Distribution on November 20, 2023 (Expiring December 29, 2023)

4.2.3 Fitting of Bitcoin Option Implied Volatility Curve

To more precisely fit the processed implied volatility data, this research adopts a 4th-order spline function with a single knot for curve fitting. The knot is placed at the futures price, a design that allows greater flexibility at this key position while maintaining the overall continuity of the curve. Using a 4th-order spline function ensures that the fitted curve has third-order continuous differentiability (C^3 continuity), effectively capturing subtle changes in the implied volatility curve while avoiding over-fitting problems.

The mathematical representation of the 4th-order spline function is as follows:

$$S(x) = \begin{cases} \sum_{i=0}^{4} a_i\,(x - x_0)^i, & x < k \\ \sum_{i=0}^{4} b_i\,(x - x_0)^i, & x \geq k \end{cases}$$

Where k is the knot position (i.e., the futures price), x_0 is the reference point, and a_i and b_i are coefficients to be determined. At the knot, the function must satisfy the following continuity conditions:

$$\begin{cases} \sum_{i=0}^{4} a_i(k - x_0)^i = \sum_{i=0}^{4} b_i(k - x_0)^i \\ \sum_{i=1}^{4} i\,a_i(k - x_0)^{i-1} = \sum_{i=1}^{4} i\,b_i(k - x_0)^{i-1} \\ \sum_{i=2}^{4} i(i-1)a_i(k - x_0)^{i-2} = \sum_{i=2}^{4} i(i-1)b_i(k - x_0)^{i-2} \\ \sum_{i=3}^{4} i(i-1)(i-2)a_i(k - x_0)^{i-3} = \sum_{i=3}^{4} i(i-1)(i-2)b_i(k - x_0)^{i-3} \end{cases}$$

These conditions ensure the continuity of the function value and its first, second, and third derivatives at the knot.

Setting the sole knot at the futures price has important economic significance, as this position typically corresponds to at-the-money options. This knot placement divides the curve into two segments, corresponding to areas above and below the futures price, allowing the fitted curve to more accurately reflect volatility characteristics near the at-the-money point. This segmented fitting method is particularly suitable for handling the asymmetric features that may appear in option implied volatility before and after the at-the-money position.

At the implementation level, this research uses the LSQUnivariateSpline method from Python's SciPy package for curve fitting. This method employs the least squares approach for parameter estimation, effectively handling non-uniformly distributed data points and achieving segmented fitting through specified internal knots. Through the combination of the least squares method and knot placement, LSQUnivariateSpline provides a flexible and stable mathematical tool capable of fitting smooth and accurate implied volatility curves in non-uniformly distributed data (as shown in Figure 4-4), providing a solid foundation for subsequent risk-neutral density function (RND) calculations. The least squares method formula is as follows:

$$\underset{\{a_i\},\{b_i\}}{min} \sum_{j=1}^{n} [y_j - S(x_j)]^2$$

Where y_j is the actual observed value, $S(x_j)$ is the observed value of the spline function at x_j, $\{a_i\},\{b_i\}$ is the set of parameters to be estimated for the spline function, and n is the total number of data points.

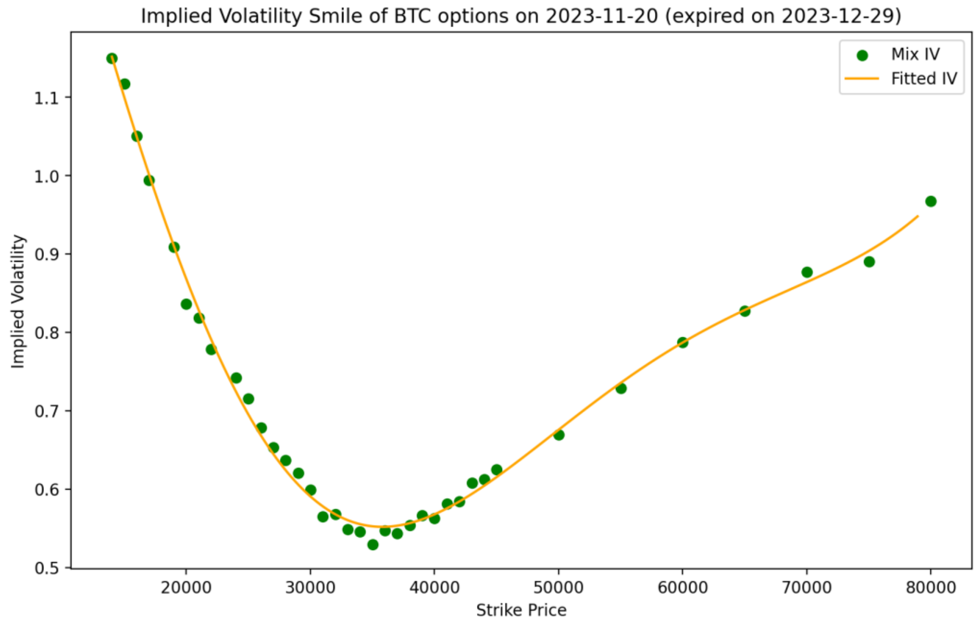

Figure 4-4: Bitcoin Option Implied Volatility Fitted Curve on November 20, 2023 (Expiring December 29, 2023)

4.2.4 Calculation of Bitcoin Option Risk-Neutral Probability Density Function

After completing the fitting of the implied volatility curve, the calculation of the risk-neutral probability density follows. First, this research uses the fitted implied volatility curve, combined with the pricing model adopted by the Deribit exchange (Formula (11)), to calculate theoretical call option prices at different strike prices, with the calculation results shown in Figure 4-5.

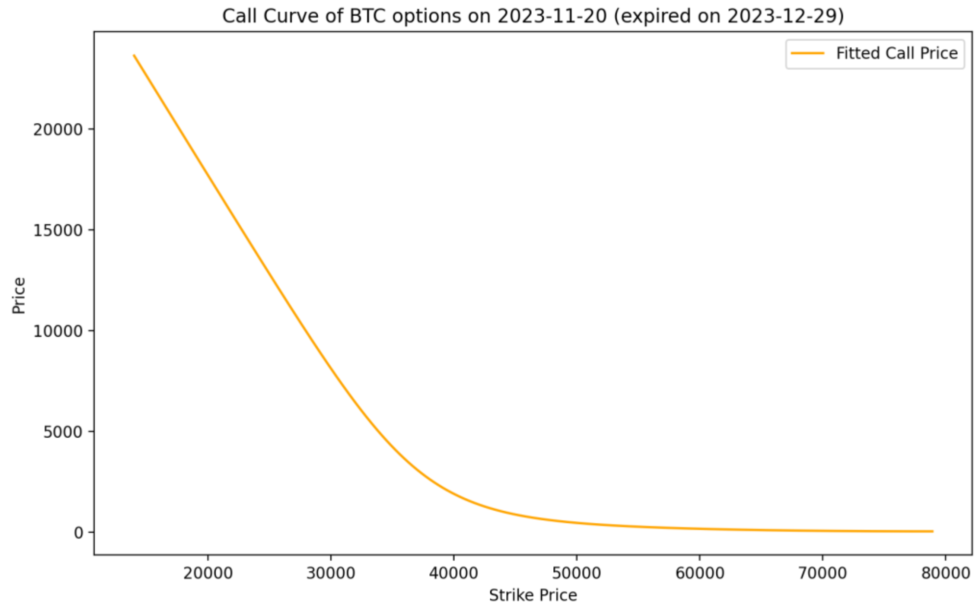

Figure 4-5: Bitcoin Option Theoretical Call Option Prices on November 20, 2023 (Expiring December 29, 2023)

After obtaining the theoretical call option prices, this research employs the finite difference method (central difference method) for discrete data differentiation to calculate the risk-neutral probability density. Compared to forward or backward differences, the central difference method can effectively reduce truncation errors (Formulas (4) and (5)). To ensure the stability and accuracy of numerical calculations, this research adopts an equidistant partitioning approach in setting the price spacing, with ΔX set to 0.1. Choosing a smaller price spacing not only provides more refined density estimation results, but the equidistant partitioning approach also helps improve the stability of numerical differentiation calculations.

After completing the risk-neutral probability density calculation, to ensure the reliability and reasonableness of the data, this research focuses on the cumulative distribution function (CDF) and right cumulative probability for data validation. In terms of data integrity validation, this research first ensures that both cumulative distribution function values and right cumulative probability values do not contain missing values, with observations containing missing values excluded from the analysis scope.

Furthermore, to maintain theoretical consistency of calculation results, this research further restricts these two probability values to be strictly between 0 and 1, while excluding boundary values equal to 0 or 1, to avoid extreme cases affecting subsequent analysis.

$$\begin{cases} F(K)\ exists\ (non-missing) \\ F_R(K)\ exists\ (non-missing) \\ 0 < F(K) < 1 \\ 0 < F_R(K) < 1 \end{cases}$$

Where $F(K)$ is the cumulative distribution function, $F_R(K)$ is the right cumulative probability, and $F_R(K) = 1 - F(K)$ 。

The validation criteria are set based on three aspects of consideration. First, from the perspective of theoretical consistency, these criteria ensure that the estimation results conform to the basic properties of cumulative distribution functions in probability theory, while also satisfying the basic requirements of probability density functions. Second, in terms of numerical stability, this filtering approach can avoid computational problems in subsequent analyses due to extreme or abnormal values, effectively improving the reliability of the overall estimation results. Finally, from a practical application perspective, removing abnormal values that might lead to misinterpretation better ensures that results accurately reflect market participants' true price expectations.

The calculated risk-neutral probability density function and its cumulative distribution function are shown in Figure 4-6 and Bitcoin Option Risk-neutral Probability Cumulative Distribution Function on November 20, 2023 (Expiring December 29, 2023)Figure 4-7:

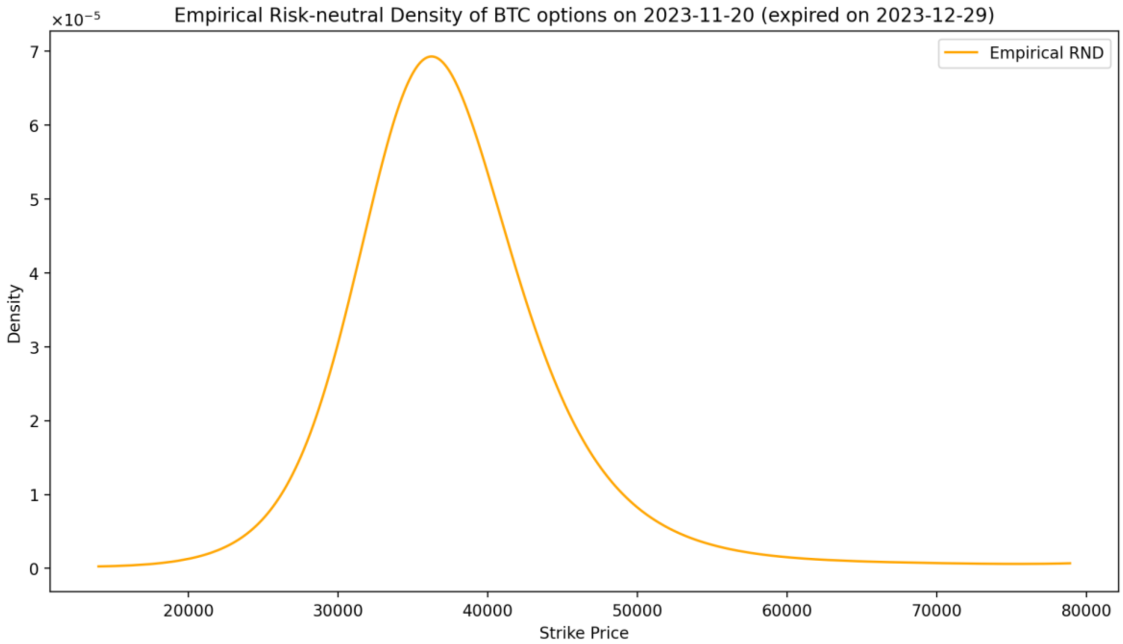

Figure 4-6: Bitcoin Option Risk-neutral Probability Density Function on November 20, 2023 (Expiring December 29, 2023)

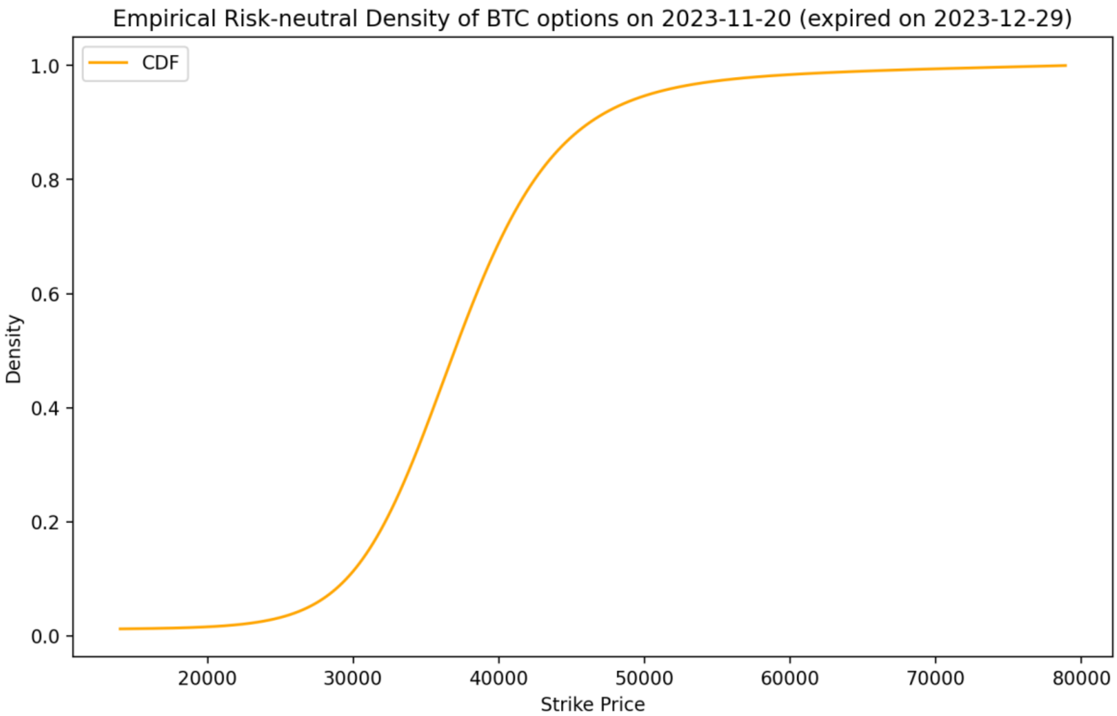

Figure 4-7: Bitcoin Option Risk-neutral Probability Cumulative Distribution Function on November 20, 2023 (Expiring December 29, 2023)

4.3 Fitting Risk-neutral Probability Density Tail Distribution

4.3.1 Two Points Fitting Method with GEV

The risk-neutral probability density (RND) extracted from market option prices can only cover the range of effective trading strike prices. To fully describe market expectations, the tails of the RND need to be extended. Figlewski (2008) proposed using the Generalized Extreme Value Distribution (GEV) to fit the tails of the RND. This method requires setting three conditions for each tail to ensure the continuity of tail fitting:

$$\text{Right tail conditions: } \begin{cases} F_{GEVR}\big(K(\alpha_{1R})\big) = \alpha_{1R} \\ f_{GEVR}\big(K(\alpha_{1R})\big) = f_{EMP}\big(K(\alpha_{1R})\big) \\ f_{GEVR}\big(K(\alpha_{2R})\big) = f_{EMP}\big(K(\alpha_{2R})\big) \end{cases}$$

$$\text{Left tail conditions: } \begin{cases} F_{GEVL}\big(-K(\alpha_{1L})\big) = 1 - \alpha_{1L} \\ f_{GEVR}\big(-K(\alpha_{1L})\big) = f_{EMP}\big(K(\alpha_{1L})\big) \\ f_{GEVR}\big(-K(\alpha_{2L})\big) = f_{EMP}\big(K(\alpha_{2L})\big) \end{cases}$$

Where F_{GEVR} and f_{GEVR} are the cumulative distribution function and probability density function of the right tail GEV, respectively, F_{GEVL} is the cumulative distribution function of the left tail GEV, f_{EMP} is the empirical risk−neutral probability density function calculated in this research, and $K(\alpha)$ is the strike price corresponding to the αquantile of the Empirical RND. The fitting conditions from Figlewski (2008) can be summarized as:

(1) At the first joining point, the cumulative probability of the GEV tail must equal the cumulative probability of the RND

(2) At the first joining point, the GEV density function value must equal the density function value of the RND

(3) At the second joining point, the GEV density function value must equal the density function value of the RND

4.3.2 GPD Distribution Theory

This research adopts the Generalized Pareto Distribution (GPD) for tail fitting. The central position of GPD in extreme value theory stems from the groundbreaking research by Balkema and Haan (1974) , who proved that when observations exceed a sufficiently high threshold, their excess distribution asymptotically converges to GPD. This property makes GPD particularly suitable for describing tail events. The choice of GPD over GEV is primarily based on its parameter structure advantage: GPD only requires setting two parameters: scale parameter (σ) and shape parameter (ξ), which is more concise compared to GEV's three parameters: location parameter (μ), scale parameter (σ), and shape parameter (ξ). This concise parameter structure not only enhances computational efficiency but also reduces the risk of over-fitting. Additionally, research by Hosking and Wallis (1987) further verified the threshold stability property of GPD, making it more advantageous in practical applications. According to extreme value theory, GPD and GEV share the same shape parameter in the tail, meaning that both distributions have the same asymptotic properties when describing extreme events. However, GPD is more intuitive in practical applications, especially in the field of financial risk management, as shown in the research by McNeil and Frey (2000). In terms of empirical research, Birru and Figlewski (2012) further confirmed that the performance of GPD and GEV in tail fitting is quite similar, and both outperform the lognormal distribution.

The mathematical expression of the GPD cumulative distribution function (CDF) is as follows:

$$F_{GPD}(x;\sigma,\xi) = \begin{cases} 1-(1+\xi\frac{x}{\sigma})^{-\frac{1}{\xi}}, & \xi \neq 0 \\ 1-exp\,(-\frac{x}{\sigma}), & \xi = 0 \end{cases}$$

The mathematical expression of the GPD probability density function (PDF) is as follows:

$$f_{GPD}(x;\sigma,\xi) = \begin{cases} \frac{1}{\sigma}(1+\xi\frac{x}{\sigma})^{-\frac{1}{\xi}-1}, & \xi \neq 0 \\ \frac{1}{\sigma}exp\,(-\frac{x}{\sigma}), & \xi = 0 \end{cases}$$

Where $\sigma > 0$ is the scale parameter, used to control the degree of dispersion of the distribution, with larger σ values indicating greater variability in the data. The shape parameter ξ determines the type and tail characteristics of the distribution:

(1) When $\xi > 0$: The distribution is a Pareto Distribution with heavy-tailed characteristics; the distribution has infinite support, with domain $[0,\infty)$; the tail decays more slowly

(2) When $\xi = 0$: The distribution degenerates to an Exponential Distribution with a fixed decay rate; it is the simplest continuous memoryless distribution; the tail decays at a moderate speed

(3) When $\xi < 0$: The distribution belongs to the Beta Family with finite support characteristics; the distribution function is only defined on the interval $\left[0,-\frac{\sigma}{\xi}\right)$; it is less commonly used in financial market applications because asset returns typically do not have a clear upper limit

4.3.3 Two Points Fitting Method with GPD

Birru and Figlewski (2012) proposed a two points fitting method for GPD tails, using two joining points to compare the density function values of GPD and RND. The tail fitting conditions are set as follows:

(1) At the first joining point, the GPD density function value must equal the RND density function value

(2) At the second joining point, the GPD density function value must equal the RND density function value

The mathematical expressions are as follows:

Right tail conditions:

$$\begin{cases} f_{GPD}(K(\alpha_{1R})) = f_{EMP}(K(\alpha_{1R})) & (PDF\ condition) \\ f_{GPD}(K(\alpha_{2R})) = f_{EMP}(K(\alpha_{2R})) & (PDF\ condition) \end{cases}$$

Left tail conditions:

$$\begin{cases} f_{GPD}(K(\alpha_{1L})) = f_{EMP}(K(\alpha_{1L})) & (PDF\ condition) \\ f_{GPD}(K(\alpha_{2L})) = f_{EMP}(K(\alpha_{2L})) & (PDF\ condition) \end{cases}$$

The scale parameter and shape parameter of GPD are solved by minimizing the following objective functions:

Right tail parameter minimization objective function:

$$\min_{\sigma,\xi}\{[f_{GPD}(K(\alpha_{1R})) - f_{EMP}(K(\alpha_{1R}))]^2 + [f_{GPD}(K(\alpha_{2R})) - f_{EMP}(K(\alpha_{2R}))]^2\}$$

Left tail parameter minimization objective function:

$$\min_{\sigma,\xi}\{[f_{GPD}(K(\alpha_{1L})) - f_{EMP}(K(\alpha_{1L}))]^2 + [f_{GPD}(K(\alpha_{2L})) - f_{EMP}(K(\alpha_{2L}))]^2\}$$

Where α_{1L} is preset to 0.05; α_{2L} is preset to 0.02; α_{1R} is preset to 0.95; α_{2R} is preset to 0.98. The empirical RND curve and its CDF completed using the two points fitting method are shown in Figure 4-8 and Figure 4-9.

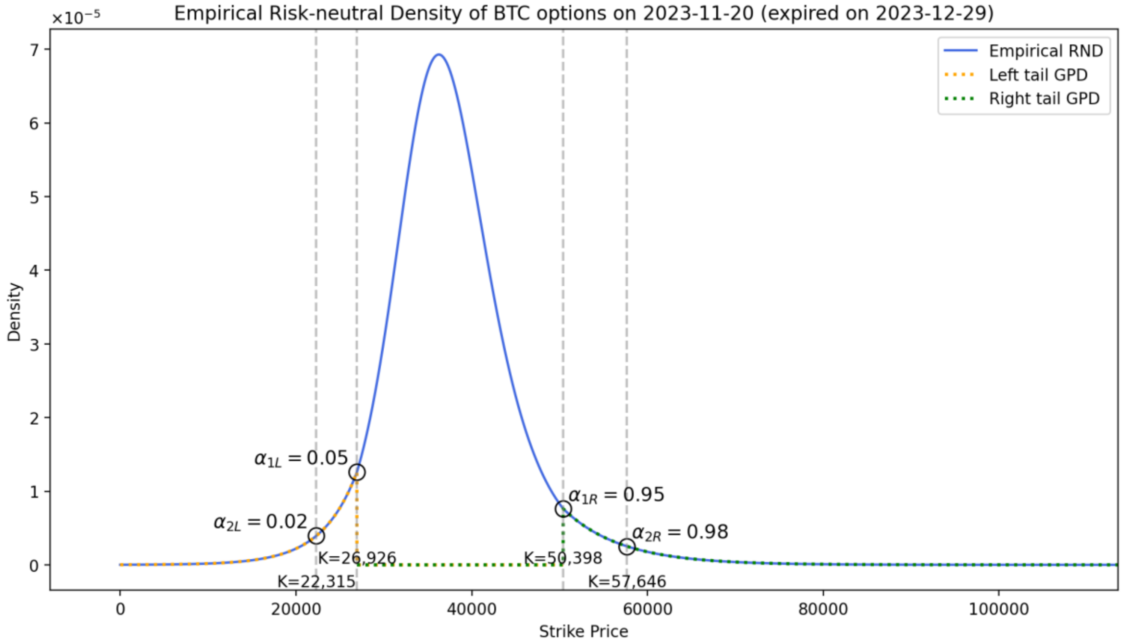

Figure 4-8: Bitcoin Option Empirical RND and GPD Tail Fitting on November 20, 2023 (Two Points Fitting Method)

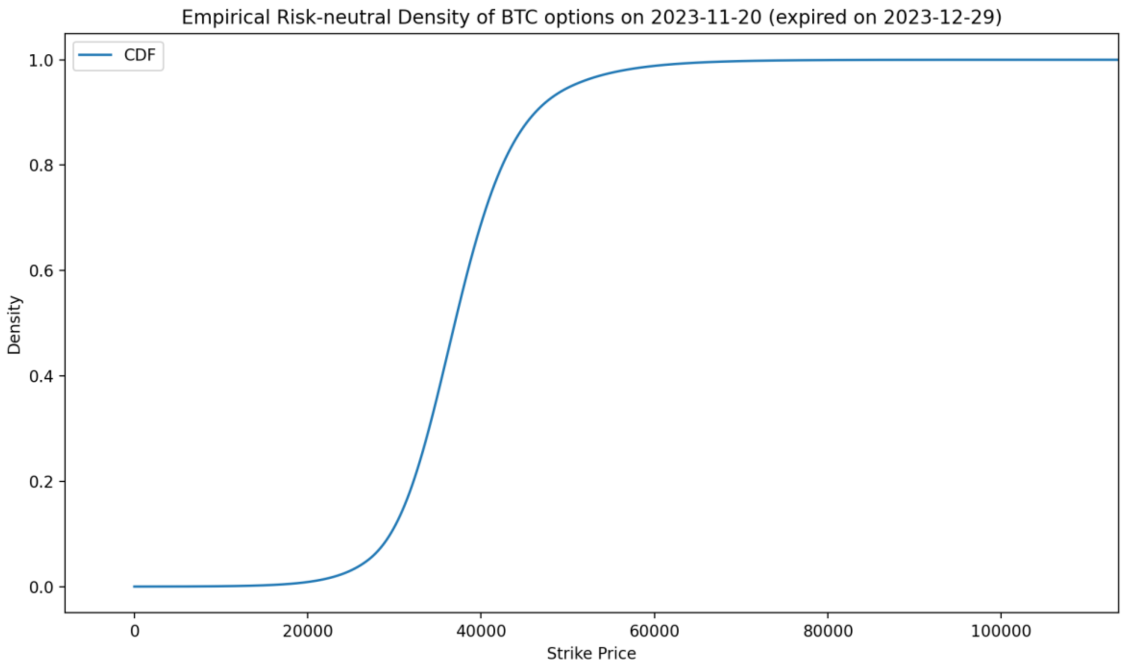

Figure 4-9: CDF of Bitcoin Option Empirical RND on November 20, 2023 (Two Points Fitting Method)

4.3.4 Single Point with the Slope Fitting Method with GPD

This research proposes using a single point with the slope fitting method with GPD (hereinafter referred to as the "single point method"). Our proposed method not only considers the fitting of CDF values but also adds continuity conditions for the slope of the density function. The main advantage of this method is that it can simultaneously ensure the continuity and smoothness of the density function while simplifying the fitting process and improving computational efficiency. The tail fitting conditions set in this research are as follows:

(1) At the joining point, the cumulative probability of GPD must equal the cumulative probability of RND

(2) At the joining point, the slope of the GPD density function must equal the slope of the RND density function

The mathematical expressions are as follows:

Right tail conditions:

$$\begin{cases} F_{GPD}(K(\alpha_{1R})) = \alpha_{1R} & (CDF\ condition) \\ \dfrac{f_{GPD}(K(\alpha_{1R}) + \Delta x) - f_{GPD}(K(\alpha_{1R}))}{\Delta x} = \dfrac{f_{EMP}(K(\alpha_{1R}) + \Delta x) - f_{EMP}(K(\alpha_{1R}))}{\Delta x} & (Slope\ condition) \end{cases}$$

Left tail conditions:

$$\begin{cases} F_{GPD}(-K(\alpha_{1L})) = \alpha_{1L} & (CDF\ condition) \\ \dfrac{f_{GPD}(-K(\alpha_{1L}) + \Delta x) - f_{GPD}(-K(\alpha_{1L}))}{\Delta x} = \dfrac{f_{EMP}(K(\alpha_{1L}) + \Delta x) - f_{EMP}(K(\alpha_{1L}))}{\Delta x} & (Slope\ condition) \end{cases}$$

The scale parameter and shape parameter of GPD are solved by minimizing the following objective functions:

Right tail parameter minimization objective function:

$$\min_{\sigma,\xi}\left\{\left[F_{GPD}\left(K(\alpha_{1R})\right)-\alpha_{1R}\right]^2 + \left[\frac{f_{GPD}(K(\alpha_{1R})+\Delta x)-f_{GPD}(K(\alpha_{1R}))}{\Delta x} - \frac{f_{EMP}(K(\alpha_{1R})+\Delta x)-f_{EMP}(K(\alpha_{1R}))}{\Delta x}\right]^2\right\}$$

Left tail parameter minimization objective function:

$$\min_{\sigma,\xi}\left\{\left[F_{GPD}(-K(\alpha_{1L}))-\alpha_{1L}\right]^2 + \left[\frac{f_{GPD}(-K(\alpha_{1L})+\Delta x)-f_{GPD}(-K(\alpha_{1L}))}{\Delta x} - \frac{f_{EMP}(K(\alpha_{1L})+\Delta x)-f_{EMP}(K(\alpha_{1L}))}{\Delta x}\right]^2\right\}$$

Where Δx is a fixed constant value used to construct artificially spaced option prices, preset to 0.1; α_{1L} is preset to 0.05; α_{1R} is preset to 0.95.

This research performs fitting separately for the left and right tails. For the left tail, we select the point with a cumulative probability of 5% as the joining point; for the right tail, we select the point with a cumulative probability of 95% as the joining point. This design ensures the smoothness of tail fitting while maintaining the continuity of the overall distribution. To minimize fitting errors, this research adopts the least squares method for parameter estimation, solving for the scale parameter and shape parameter of GPD through numerical optimization methods. The completed empirical RND curve and its CDF are shown in Figures 4-10 and 4-11.

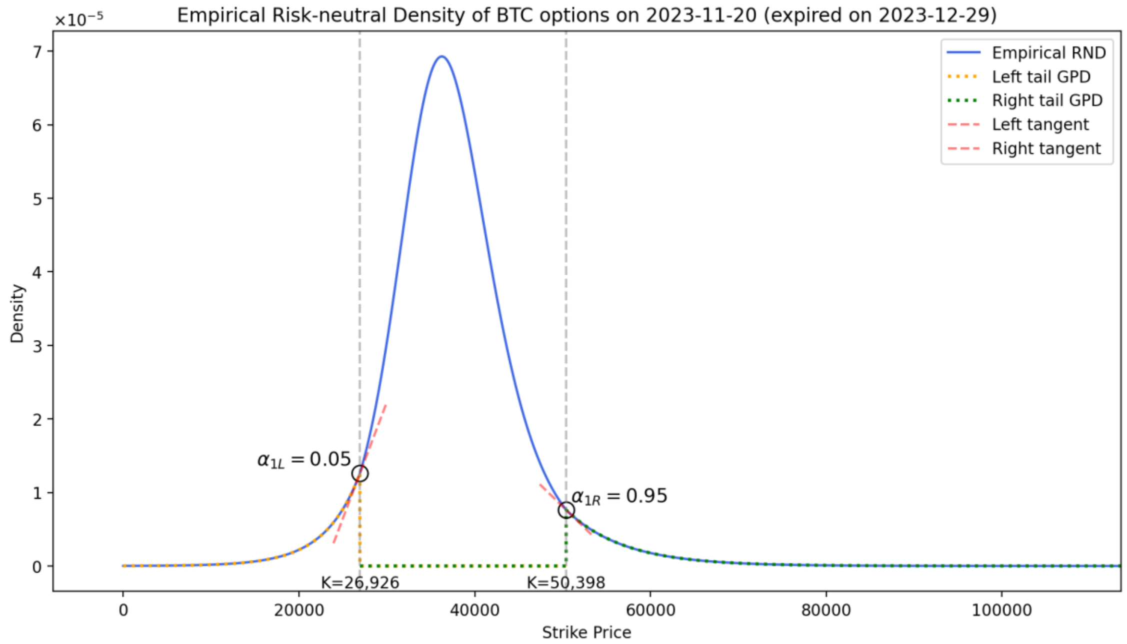

Figure 4-10: Bitcoin Option Empirical RND and GPD Tail Fitting on November 20, 2023 (Single Point Fitting Method)

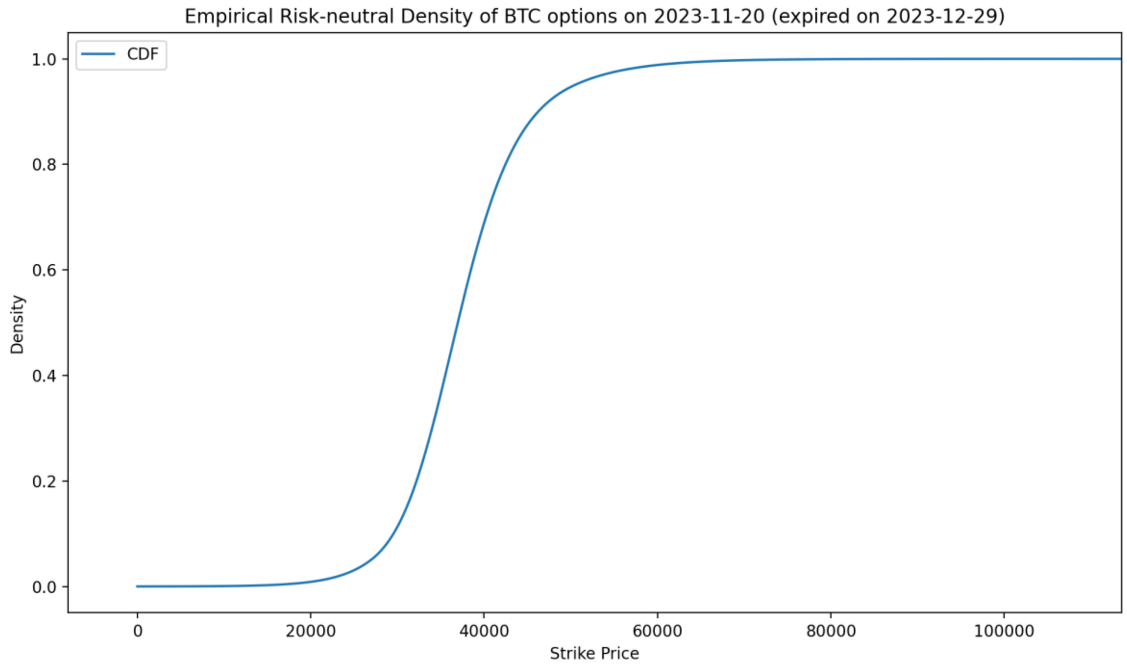

Figure 4-11: CDF of Bitcoin Option Empirical RND on November 20, 2023 (Single Point Fitting Method)

4.4 Statistical Characteristics of Risk-neutral Probability Density Function

This section will introduce in detail the statistical characteristics of the risk-neutral probability density (RND), including moments of various orders and their derived statistics, to analyze market participants' expectations regarding future Bitcoin price movements.

4.4.1 Definition and Implications of Moments

Moments are important statistical measures for describing the characteristics of probability distributions and can be divided into raw moments and central moments. For discrete strike prices K, the nth-order raw moment is defined as:

$$m_n' = E[K^n] = \sum_{i=1}^{N} K_i^n f(K_i)\,\Delta K$$

And the nth-order central moment is defined as:

$$m_n = E[(K-\overline{K})^n] = \sum_{i=1}^{N} (K_i-\overline{K})^n f(K_i)\,\Delta K$$

Where $f(K_i)$ is the risk−neutral probability density function (RND); ΔK is the price interval, set to 0.1; $\overline{K}$ is the expected value, i.e., the first−order raw moment m_1'; N is the number of observations.

4.4.2 Statistical Characteristics of Risk-neutral Probability Density Function

Recent research shows that option-implied statistical characteristics not only effectively describe market expectations but also possess significant predictive power. Chang et al. (2013) found that risk-neutral skewness can effectively predict stock returns, especially during periods of high market volatility; while Neumann and Skiadopoulos (2013) pointed out that changes in risk-neutral kurtosis often lead market trends, providing important signals for investment decisions. This research will calculate the following four main statistical characteristics:

1. Mean

The mean is the first-order raw moment, reflecting the overall market expectation for the future price level of the underlying asset. Bali and Murray (2013) pointed out that under risk-neutral pricing theory, the mean of the RND should equal the forward price discounted by the risk-free rate, providing an important arbitrage constraint. The formula is as follows:

$$\overline{K} = m_1' = E[K] = \sum_{i=1}^{N} K_i f(K_i) \Delta K$$

2. Standard Deviation

The standard deviation is the square root of the second-order central moment, measuring the degree of price dispersion. Christoffersen et al. (2013) found that option-implied standard deviation has stronger predictive power than historical volatility, especially in emerging markets. Standard deviation reflects market expectations for price volatility, with higher values indicating greater uncertainty among market participants regarding future price movements. The formula is as follows:

$$\sigma = \sqrt{m_2} = \sqrt{E[(K - \overline{K})^2]} = \sqrt{\sum_{i=1}^{N} (K_i - \overline{K})^2 f(K_i) \Delta K}$$

3. Skewness

Skewness is the standardized third-order central moment, describing the asymmetry of the distribution:

$$\text{Skewness} = \frac{m_3}{\sigma^3} = \frac{E[(K-\overline{K})^3]}{\sigma^3} = \frac{\sum_{i=1}^{N}(K_i-\overline{K})^3 f(K_i)\,\Delta K}{\sigma^3}$$

The skewness coefficient has important implications in financial markets. When positive skewness is observed, it indicates that the price distribution has a longer right tail, implying that market participants expect a higher probability of significant upward movements, reflecting overall optimistic market sentiment. Conversely, negative skewness indicates that the price distribution has a longer left tail, representing a greater perceived downside risk in the market, usually reflecting higher hedging demand among market participants. Research by Conrad et al. (2013) shows that risk-neutral skewness not only reflects market sentiment but also contains investors' expectations for extreme events; Li et al. (2024) empirical research shows that the dynamic changes in the skewness coefficient can often serve as a leading indicator of market sentiment shifts, with its changing trends providing important reference value for investment decisions.

4. Excess Kurtosis

Excess Kurtosis is the standardized fourth-order central moment minus 3 (the kurtosis value of the normal distribution), used to describe the tail characteristics of the distribution. It is calculated as:

$$\text{Excess Kurtosis} = \frac{m_4}{\sigma^4} - 3 = \frac{E[(K-\overline{K})^4]}{\sigma^4} - 3 = \frac{\sum_{i=1}^{N}(K_i-\overline{K})^4 f(K_i)\,\Delta K}{\sigma^4} - 3$$

In practical applications, excess kurtosis is an important indicator for assessing extreme market risk. Amaya et al. (2015) pointed out that excess kurtosis can effectively capture extreme market risk, with its predictive power being particularly significant during financial crises. When positive excess kurtosis is observed, it indicates that the distribution has more pronounced fat-tail characteristics compared to the normal distribution, meaning that the probability of extreme events occurring is higher than expected under a normal distribution.

4.4.3 Market Implications of Statistical Characteristics

The statistical characteristics of RND not only provide a mathematical description of the distribution but also contain rich market information. From the perspective of price expectations, the mean of the distribution reflects the consensus expectation of the market for future price levels. If it deviates significantly from the actual market price, it may suggest potential arbitrage opportunities. Such price deviations often provide profit opportunities for arbitrageurs while also contributing to the self-adjusting mechanism of market prices.

In terms of risk assessment, standard deviation, as a traditional indicator of volatility risk, combined with the extreme risk information provided by excess kurtosis, allows investors to assess market risk more comprehensively. Especially during periods of increased market volatility, changes in these indicators often provide timely risk warning signals, helping investors adjust their portfolios in a timely manner to respond to market changes. Market sentiment monitoring can be achieved through the changing trends in skewness, reflecting shifts in market sentiment and providing a quantitative basis for dynamic adjustment of investment strategies.

This research will use regression analysis methods to empirically test the predictive power of RND characteristics for spot returns. By analyzing the correlation between these statistical characteristics and future market trends, we will be able to more deeply evaluate their application value in the Bitcoin market, providing a more solid quantitative foundation for investment decisions.

4.5 Empirical Analysis

4.5.1 Variable Selection and Theoretical Basis for Regression Models

To explore the predictive power of Bitcoin option-implied risk-neutral probability density (RND) for underlying asset price movements, this research adopts multiple regression analysis for empirical testing. The dependent variable (Y) is set as the logarithmic return of Bitcoin, while the independent variables (X) gradually incorporate statistical characteristics of RND such as Mean, Standard Deviation, Skewness, Excess Kurtosis, Median, as well as market sentiment indicators like the Cryptocurrency Fear and Greed Index and the Chicago Board Options Exchange Volatility Index (VIX), along with historical returns from the previous 1 to 4 periods, to construct the most explanatory predictive model.

Research by Bali and Zhou (2016) found that statistical characteristics of RND, especially skewness and kurtosis, can effectively predict cross-sectional changes in asset returns. Empirical results show that these characteristics not only reflect market participants' risk preferences but also contain important pricing information. Additionally, Amaya et al. (2015) pointed out that the excess kurtosis of RND is particularly effective in predicting extreme market risk, a finding that is especially important in high-volatility markets such as cryptocurrencies.

López-Cabarcos et al. (2021) studied the relationship between Bitcoin volatility, stock market performance, and investor sentiment. Empirical results show that during periods of market stability, VIX returns and investor sentiment do have a significant impact on Bitcoin volatility. Furthermore, Akyildirim et al. (2020), using high-frequency data, explored the dynamic relationship between cryptocurrencies and market panic indicators. The research found that there exists a time-varying positive correlation between cryptocurrencies and market panic indicators such as VIX and VSTOXX, and this correlation significantly increases with financial market pressure. M. He et al. (2023) conducted research on cryptocurrency return prediction, using the daily updated and easily accessible Cryptocurrency Fear and Greed Index as a predictor. Empirical results show that this index has significant predictive power both in-sample and out-of-sample, with predictive ability for returns of individual cryptocurrencies and market indices over prediction periods from one day to one week.

Liu and Tsyvinski (2021) studied the risk and return characteristics of the cryptocurrency market. Results show that there is a significant momentum effect in the cryptocurrency market, with Bitcoin's current period returns significantly predicting returns for the next 1 to 6 days. Secondly, Y. Li et al. (2021) explored the MAX effect (maximum daily returns effect) in the cryptocurrency market. The research found that there is a positive MAX momentum effect in the cryptocurrency market, with cryptocurrencies having higher extreme daily returns tending to produce higher returns in the future. Furthermore, Liu et al. (2023), using machine learning methods to predict cryptocurrency returns, found that among numerous predictive variables, the previous 1-day return has the strongest predictive power, even exceeding the combined effect of all other variables.

4.5.2 Regression Model Design

Campbell & Thompson (2008) introduced the concept of "economic significance threshold," showing that by gradually introducing predictive variables and setting strict statistical significance standards, one can effectively distinguish variables with substantial predictive power. In the empirical asset pricing research by Gu et al. (2020), a "staged variable introduction framework" was proposed specifically for high-dimensional data, which can alleviate overfitting problems compared to models that introduce all variables at once.

This research will adopt a multi-level regression analysis method, gradually expanding from univariate to four-variable models, to systematically explore the predictive power of various risk-neutral probability density characteristics for future returns. The following details the design of regression models at each level:

1. Univariate Regression Model

The univariate regression model is mainly used to test the explanatory power of individual variables for future returns. Its basic form is:

$$R_t = \beta_0 + \beta_1 Variable_{i,t-1} + \varepsilon_t, i \in \{1, 2, \dots, 11\}$$

Where $R_t = ln(\frac{Close_t}{Close_{t-1}})$ is the return of Bitcoin price in the next period (T). If the option sample expires in 7 days, then the return is calculated using the closing price from the current day to the closing price 7 days later (expiration date).

$Variable_i$ will be sequentially replaced with the following variables for univariate regression analysis: Mean, Standard Deviation (Std), Skewness, Excess Kurtosis, Median, Cryptocurrency Fear and Greed Index, Chicago Board Options Exchange Volatility Index (VIX), T-1 Return, T-2 Return, T-3 Return, T-4 Return.

2. Bivariate Regression Model

Considering the importance of Skewness in option pricing theory, this research designs bivariate models with Skewness as a fixed factor. The model is set as follows:

$$R_t = \beta_0 + \beta_1 Skewness_{t-1} + \beta_2 Variable_{i,t-1} + \varepsilon_t, i \in \{1, 2, \ldots, 10\}$$

$Variable_i$ will be sequentially replaced with the following variables, paired with Skewness for bivariate regression analysis: Mean, Standard Deviation (Std), Excess Kurtosis, Median, Cryptocurrency Fear and Greed Index, Chicago Board Options Exchange Volatility Index (VIX), T-1 Return, T-2 Return, T-3 Return, T-4 Return.

3. Three-Variable Regression Model

The three-variable model further incorporates Excess Kurtosis as a fixed factor, forming:

$$R_t = \beta_0 + \beta_1 Skewness_{t-1} + \beta_2 ExcessKurtosis_{t-1} + \beta_3 Variable_{i,t-1} + \varepsilon_t, i \in \{1, 2, \ldots, 9\}$$

$Variable_i$ will be sequentially replaced with the following variables, paired with Skewness and Excess Kurtosis for three-variable regression analysis: Mean, Standard Deviation (Std), Median, Cryptocurrency Fear and Greed Index, Chicago Board Options Exchange Volatility Index (VIX), T-1 Return, T-2 Return, T-3 Return, T-4 Return.

4. Four-Variable Regression Model

Building on the three-variable model, the four-variable model adds the Standard Deviation (Std) variable. The complete model is as follows:

$$R_t = \beta_0 + \beta_1 Skewness_{t-1} + \beta_2 ExcessKurtosis_{t-1} + \beta_3 Std_{t-1} + \beta_4 Variable_{i,t-1} + \varepsilon_t,$$

$i \in 1,2, \ldots ,8$

$Variable_i$ will be sequentially replaced with the following variables, paired with Skewness, Excess Kurtosis, and Standard Deviation for four-variable regression analysis: Mean, Median, Cryptocurrency Fear and Greed Index, Chicago Board Options Exchange

Volatility Index (VIX), T-1 Return, T-2 Return, T-3 Return, T-4 Return.

4.5.3 Validation of the Effectiveness of Our Proposed Method

Our proposed method of single point compared to the two points method used in existing literature has the advantage of computational efficiency in practical applications. To objectively validate the effectiveness of this method, this research adopts three quantitative indicators for evaluation: Mean Squared Error (MSE), Out-of-sample R-squared (R_{OS}^2), and computational efficiency, to comprehensively compare the accuracy and practicality of predictive models constructed by the two methods.

1. Mean Squared Error (MSE) Calculation

Mean Squared Error is a commonly used indicator for evaluating the accuracy of predictive models (Orosi, 2015). Its calculation formula is as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

Where y_i is the actual observed value (i.e., the actual Bitcoin return), $\hat{y}_i$ is the model's predicted value, and n is the sample size. A smaller MSE value indicates smaller prediction errors and higher prediction accuracy.

2. Out-of-sample R-squared (R_{OS}^2) Calculation

Referring to the research methods of Campbell and Thompson (2008) and Welch and Goyal (2008), this research adopts out-of-sample R-squared (R_{OS}^2) to evaluate the predictive power of the model. R_{OS}^2 measures the improvement in prediction of the predictive model relative to the historical average benchmark model. Its calculation formula is as follows:

$$R_{OS}^2 = 1 - \frac{\sum_{t=s_0+1}^{T}(R_{t+1} - \hat{R}_{t+1})^2}{\sum_{t=s_0+1}^{T}(R_{t+1} - \bar{R}_{t+1})^2}$$

Where R_{t+1} is the actual return, $\hat{R}_{t+1}$ is the predicted value from the predictive model, $\bar{R}_{t+1}$ is the historical average return (benchmark model), s_0 is the initial in-sample period length, and T is the total sample size. When R_{OS}^2 is greater than zero, it indicates that the predictive model outperforms the historical average benchmark model; the larger the R_{OS}^2 value, the more significant the improvement in prediction.

This research adopts a rolling window approach for out-of-sample prediction, with the initial in-sample period set to 80% of the total sample size, and gradually advancing forward for prediction. By comparing the R_{OS}^2 values of the single point method and the two points method, the difference in out-of-sample predictive power between the two methods can be objectively evaluated.

3. Computational Efficiency Comparison

In addition to prediction accuracy, this research also values the practical applicability of the method, especially its computational efficiency when processing large amounts of data. To objectively evaluate the computational performance of the two methods, this research selects ten representative option expiration dates, calculates the 7-day risk-neutral probability density for each expiration date, and uses both the single point method and the two points method to fit the tails with the Generalized Pareto Distribution (GPD). To ensure the reliability and stability of the results, this research performs ten repeated computations for each method, records their execution times, and calculates the average, thereby comprehensively evaluating the differences in computational efficiency between the two methods in practical applications.

5. Analysis of Empirical Results

This chapter will analyze the empirical results in detail. Section 5.1 compares the characteristics of the single point method and the two points method, including overall fitting effect and computational efficiency. Section 5.1.1 demonstrates the fitting performance of both methods in different market scenarios through specific samples, while Section 5.1.2 empirically compares the computational efficiency of both methods to evaluate their practical applicability. Sections 5.2 and 5.3 then conduct in-depth analyses of option products with 1 day and 7 days to expiration, respectively, and explore the differences in predictive effects between the two fitting methods across different expiration periods.

This research selects Bitcoin option products with 1 day and 7 days to expiration for empirical analysis, primarily based on the following reasons. The Bitcoin market features 24-hour continuous trading, and its trading volume has continued to climb in recent years, with daily average trading volume reaching tens of billions of dollars, indicating sufficient market liquidity. With abundant liquidity, the price discovery mechanism becomes more effective, allowing market participants to quickly react to new information, making option prices better reflect market expectations in real-time. Additionally, the Bitcoin market responds extremely quickly to new information with high price adjustment efficiency. Compared to traditional financial markets, cryptocurrency market traders are more technology-oriented investors who can quickly receive and process market information. Therefore, studying short-term products can more accurately capture market participants' real-time assessment of and response to risk.

5.1 Analysis of Empirical Sample Fitting Effects

5.1.1 Comparison Between Single Point Method and Two Points Method

To further compare the fitting effects of the single point method and the two points method, this research selects samples from observation dates July 10, 2022 (expiration date July 11, 2022) and September 27, 2023 (expiration date September 28, 2023) for analysis. From the empirical results, the single point method demonstrates better stability, with smoother and more continuous fitting curves, indicating that this method has stronger adaptability when handling different market scenarios.

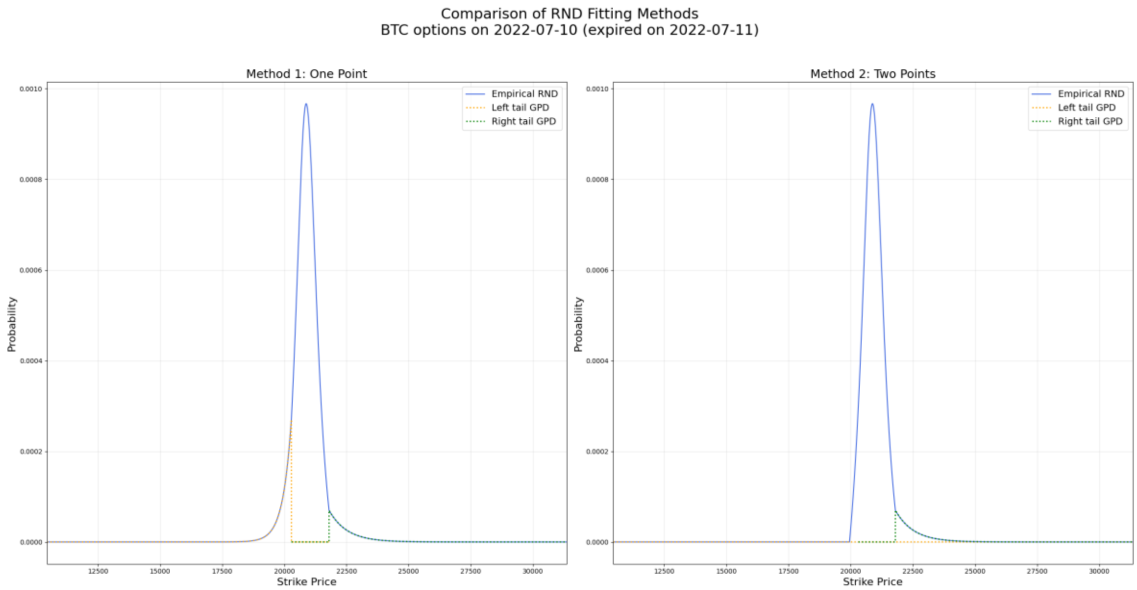

Figure 5-1: Comparison of Bitcoin Option GPD Tail Fitting on July 10, 2022

(Left: Single Point Method; Right: Two Points Method)

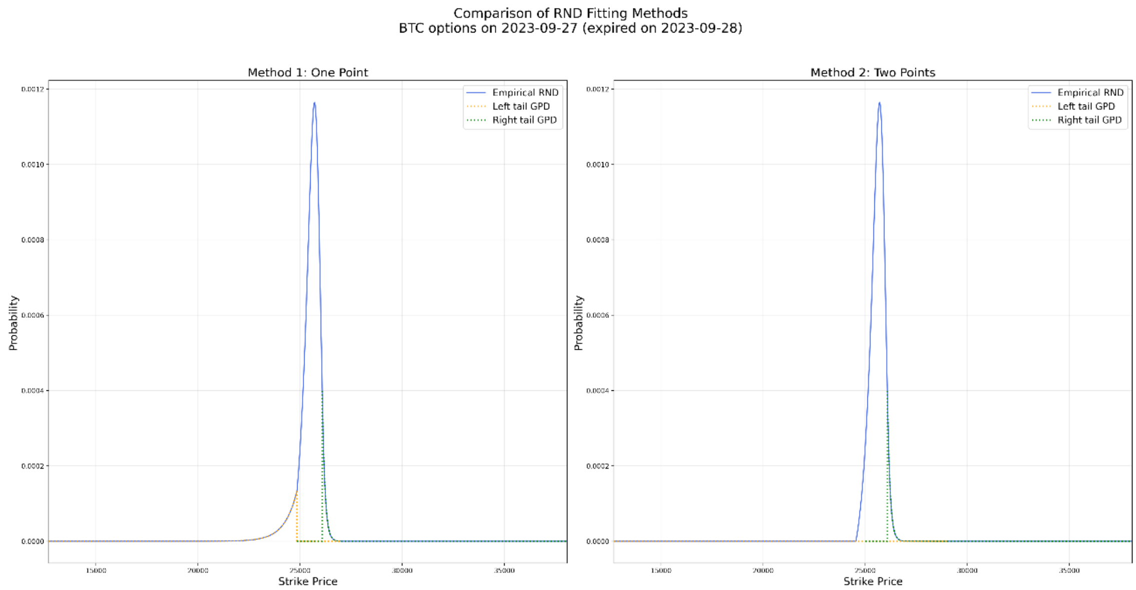

Figure 5-2: Comparison of Bitcoin Option GPD Tail Fitting on September 27, 2023
(Left: Single Point Method; Right: Two Points Method)

Observing the single point method fitting results on the left side of the figures, it is evident that regardless of whether in the 2022 or 2023 samples, the RND curve can be successfully fitted, maintaining good continuity at the joining points. Especially in the tail regions, the fitting curve of the single point method shows a gradually decreasing trend, conforming to the basic properties of probability density functions. This result indicates that the single point method has better robustness when handling extreme values.

In contrast, although the two points method fitting results on the right side can also produce reasonable fitting curves in most cases, fitting failures or curve discontinuities may occur in certain market scenarios. The main reason is that the two points method needs to simultaneously satisfy the continuity conditions at two joining points. When the market fluctuates dramatically or the price distribution is extremely skewed, it becomes more difficult to find parameter combinations that simultaneously satisfy dual conditions, thereby affecting the quality of the fitting results.

From a practical application perspective, the single point method has higher computational efficiency and is less likely to encounter situations where fitting cannot be performed, an advantage that is particularly evident when analyzing large samples. Moreover, the fitting results of the single point method are more stable, which facilitates the subsequent calculation and analysis of statistical characteristics. Based on these advantages, this research believes that when constructing the risk-neutral probability density function of Bitcoin options, the single point method has more practical value than the two points method.

5.1.2 Empirical Comparison of Computational Efficiency

To evaluate the practical applicability of the single point method and the two points method, this research conducts an empirical comparison of computational efficiency. Under the same hardware equipment, tests were performed on option products with 10 expiration dates between September and December 2023. Each experiment generated 10 weekly return risk-neutral probability density functions (RNDs) with GPD distribution tail fitting, and the results are shown in Table 5-1.

From the empirical results, the single point method demonstrates a clear advantage in computational efficiency. The average execution time for the single point method is 309.86 seconds, with the shortest and longest execution times being 297.58 seconds and 313.52 seconds, respectively. In comparison, the average execution time for the two points method is 347.95 seconds, with the shortest and longest execution times being 346.49 seconds and 349.10 seconds, respectively. Overall, the single point method saves approximately 10.95% of computation time compared to the two points method.

The main reason for this difference in computational efficiency lies in the complexity of the algorithms. The two points method needs to simultaneously satisfy the continuity conditions at two joining points, requiring consideration of more constraints in the optimization process, thereby increasing the computational burden. In contrast, the single point method only needs to handle a single joining point, making its optimization process more direct and allowing it to converge more quickly to suitable parameter combinations. Notably, the single point method not only has faster computation speed but also better stability in execution time compared to the two points method. The standard deviation of execution time for the single point method is smaller, indicating more consistent computational performance. This characteristic is particularly important when conducting large-scale empirical analyses, helping to improve research efficiency and reduce waste of computational resources.

Based on the aforementioned empirical results, this research believes that the single point method has significant advantages in practical applications, especially in scenarios requiring the processing of large samples or high real-time analysis, where its higher computational efficiency will provide better application flexibility. This finding has important practical implications for the subsequent development of automated trading strategies or risk monitoring systems.

The following subsections will conduct empirical analyses on option products with 1 day and 7 days to expiration, respectively, exploring the predictive effects of the two fitting methods across different expiration periods.

Table 5-1: Comparison of Computational Efficiency for Bitcoin Option GPD Tail Fitting
(Left: Single Point Method; Right: Two Points Method)

Execution Time:	2025/2/5 00:48
Execution Conditions:	Each time generates 10 weekly return RNDs with GPD distribution tail fitting.
Option Expiration Dates:	2023/9/22, 2023/9/29, 2023/10/13, 2023/10/20, 2023/10/27, 2023/11/10, 2023/11/17, 2023/11/24, 2023/12/15, 2023/12/22

Single Point Method		Two Points Method	
1st Execution Time (sec)	297.58	1st Execution Time (sec)	346.49
2nd Execution Time (sec)	300.23	2nd Execution Time (sec)	346.95
3rd Execution Time (sec)	313.52	3rd Execution Time (sec)	347.78
4th Execution Time (sec)	313.51	4th Execution Time (sec)	347.68
5th Execution Time (sec)	312.25	5th Execution Time (sec)	347.81
6th Execution Time (sec)	311.37	6th Execution Time (sec)	347.94
7th Execution Time (sec)	313.52	7th Execution Time (sec)	348.39
8th Execution Time (sec)	311.86	8th Execution Time (sec)	349.10
9th Execution Time (sec)	312.36	9th Execution Time (sec)	348.59
10th Execution Time	312.42	10th Execution Time	348.80
Shortest Execution Time (sec)	297.58	Shortest Execution Time (sec)	346.49
Longest Execution Time (sec)	313.52	Longest Execution Time (sec)	349.10
Average Execution Time (sec)	309.86	Average Execution Time (sec)	347.95

5.2 Empirical Regression Analysis of Products with 1 Day to Expiration

5.2.1 Tail GPD Single Point Method

This section uses option products expiring daily from January 10, 2021, to April 30, 2024, deriving RND from the observation date one day before expiration, and constructs complete RND functions using the single point method. Statistical characteristics such as mean, standard deviation, skewness, and kurtosis are then calculated as explanatory variables. Using the next period's Bitcoin spot return as the explained variable, multi-level regression analysis is conducted to observe whether RND has predictive effects. The descriptive statistics of the variables are shown in Table 5-2, with a total of 832 samples. Observing Skewness and Excess Kurtosis, it can be found that the RND functions constructed using the single point method have outliers in skewness and excess kurtosis, which in turn affect the mean and standard deviation of these variables.

Table 5-2: Descriptive Statistics of RND Statistical Characteristics and Bitcoin Returns for Products with 1 Day to Expiration (Single Point Method)

	Count	Mean	Std	Min	25%	Median	75%	Max
T Return (Y)	832	-0.0002	0.0336	-0.1670	-0.0155	-0.0003	0.0156	0.1353
Mean	832	35891.0208	14144.2251	0.0000	25047.4826	34274.1105	46035.6010	72614.4638
Std	832	1784.3452	1422.5234	0.0000	821.3941	1463.6223	2291.7552	12224.7221
Skewness	832	0.4054	7.2140	-12.5062	-0.7712	0.2451	1.0921	193.7010
Excess Kurtosis	832	56.6637	1392.5124	-140.8251	-0.2814	1.6822	4.2421	40084.7494
Median	832	36972.0948	14057.6155	7860.0000	25800.8750	35687.3500	46933.4000	72597.4000
Fear and Greed Index	832	46.9892	22.4770	6.0000	26.0000	49.0000	68.2500	95.0000
VIX	832	20.0917	5.2866	12.0700	16.1875	19.2250	23.0300	37.2100
T-1 Return	832	-0.0001	0.0337	-0.1670	-0.0155	-0.0003	0.0156	0.1353
T-2 Return	832	-0.0001	0.0337	-0.1670	-0.0153	-0.0003	0.0161	0.1353
T-3 Return	832	-0.0001	0.0337	-0.1670	-0.0153	-0.0003	0.0160	0.1353
T-4 Return	832	-0.0003	0.0337	-0.1670	-0.0157	-0.0004	0.0156	0.1353

Following the regression analysis method set in Section 4.5, the univariate regression analysis results are shown in

Table 5-3. It can be observed that Mean, Skewness, Median, and T-4 Return are significant, mainly concentrated in RND distribution characteristics. The predictive ability of market sentiment indicators such as the Cryptocurrency Fear and Greed Index and Volatility Index (VIX) is extremely low, indicating that technical indicators may be more valuable than market sentiment indicators.

Table 5-3: Univariate Regression Analysis Results for Products with 1 Day to Expiration (Single Point Method)

	Coefficient	p value	Significance	R-squared
Mean	-0.0866	0.0124	**	0.0075
Std	-0.0247	0.4761		0.0006
Skewness	0.0614	0.0770	*	0.0038
Excess Kurtosis	0.0515	0.1379		0.0027
Median	-0.0707	0.0416	**	0.0050
Fear and Greed Index	-0.0023	0.9462		0.0000
VIX	-0.0010	0.9777		0.0000
T-1 Return	-0.0402	0.2471		0.0016
T-2 Return	0.0276	0.4261		0.0008
T-3 Return	0.0093	0.7893		0.0001
T-4 Return	0.0617	0.0752	*	0.0038

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

In the bivariate regression analysis, with Skewness as a fixed variable, the results are shown in

Table 5-4. It can be observed that the addition of Median and T-4 Return can give the regression model stable predictive ability, and the explanatory power R-squared also increases, indicating that RND median and short-term momentum strategies may provide additional information for predicting returns. The results of the three-variable and four-variable regression analyses are shown in Appendix Table 1 and Appendix Table 2.

Table 5-4: Bivariate Regression Analysis Results for Products with 1 Day to Expiration (Single Point Method)

	Skewness_Coef	Skewness_p	Skewness_Sig	Variable_Coef	Variable_p	Variable_Sig	R-squared
Mean	0.0502	0.1510		-0.0796	0.0229	**	0.0100
Std	0.0607	0.0803	*	-0.0230	0.5064		0.0043
Excess Kurtosis	0.1211	0.2595		-0.0631	0.5565		0.0042
Median	0.0640	0.0647	*	-0.0730	0.0352	**	0.0091
Fear and Greed Index	0.0618	0.0758	*	-0.0063	0.8551		0.0038
VIX	0.0614	0.0771	*	0.0009	0.9792		0.0038
T-1 Return	0.0581	0.0956	*	-0.0347	0.3194		0.0050
T-2 Return	0.0624	0.0724	*	0.0298	0.3906		0.0046
T-3 Return	0.0615	0.0765	*	0.0100	0.7722		0.0039
T-4 Return	0.0602	0.0821	*	0.0606	0.0802	*	0.0074

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

After multi-level regression analysis, this section finally adopts a regression model with three variables: Skewness, Median, and T-4 Return. The model results are shown in Table 5-5, indicating that Skewness and T-4 Return have positive effects on Bitcoin spot return prediction, while Median has a negative effect; the Mean Squared Error (MSE) shows the high volatility of Bitcoin. This section also uses this model as a basis to attempt to add a fourth variable to find a more explanatory model. However, none of the added variables are statistically significant (

Appendix Table 3), indicating that this model has already demonstrated relatively stable predictive ability.

Table 5-5: Regression Analysis Results for Products with 1 Day to Expiration (Single Point Method)

	Coefficient	p value	Significance	R-squared	MSE
Skewness	0.0629	0.0690	*	0.0130	0.9858
Median	-0.0746	0.0311	**		
T-4 Return	0.0626	0.0705	*		

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

5.2.2 Tail GPD Two Points Method

This section uses option products expiring daily from January 10, 2021, to April 30, 2024, deriving RND from the observation date one day before expiration, and constructs complete RND functions using the two points method. Statistical characteristics such as mean, standard deviation, skewness, and kurtosis are then calculated as explanatory variables. Using the next period's Bitcoin spot return as the explained variable, multi-level regression analysis is conducted to observe whether RND has predictive effects. The descriptive statistics of the variables are shown in Table 5-6, with a total of 831 samples, one less than the single point method. This is because the two points method encountered fitting problems, indicating that the single point method is more stable. Observing Skewness and Excess Kurtosis, it can be found that the statistical characteristic data calculated using the two points method is less affected by outliers.

Table 5-6: Descriptive Statistics of RND Statistical Characteristics and Bitcoin Returns for Products with 1 Day to Expiration (Two Points Method)

	Count	Mean	Std	Min	25%	Median	75%	Max
T Return (Y)	831	-0.0003	0.0336	-0.1670	-0.0156	-0.0004	0.0155	0.1353
Mean	831	36049.7626	13917.9793	771.3413	25159.9872	34158.9645	46118.6419	72614.4638
Std	831	1733.1035	1173.4002	179.8905	854.8143	1502.1723	2276.4792	10729.4925
Skewness	831	0.3140	1.7155	-12.5062	-0.5639	0.5011	1.1725	16.9925
Excess Kurtosis	831	2.9968	9.4980	-140.8251	-0.7554	1.3626	3.6404	61.3908
Median	831	37010.3730	14027.0369	15890.6000	25859.8500	35717.4000	46934.7000	72597.4000
Fear and Greed Index	831	47.0253	22.4806	6.0000	26.0000	49.0000	68.5000	95.0000
VIX	831	20.0818	5.2947	12.0700	16.1800	19.2000	23.0300	37.2100
T-1 Return	831	-0.0001	0.0337	-0.1670	-0.0156	-0.0003	0.0157	0.1353
T-2 Return	831	-0.0001	0.0337	-0.1670	-0.0153	-0.0003	0.0160	0.1353
T-3 Return	831	-0.0001	0.0337	-0.1670	-0.0153	-0.0003	0.0160	0.1353
T-4 Return	831	-0.0003	0.0337	-0.1670	-0.0157	-0.0004	0.0157	0.1353

Following the regression analysis method set in Section 4.5, the univariate regression analysis results are shown in Table 5-7. It can be observed that Mean, Median, and T-4 Return are significant.

Table 5-7: Univariate Regression Analysis Results for Products with 1 Day to Expiration (Two Points Method)

	Coefficient	p value	Significance	R-squared
Mean	-0.0707	0.0415	**	0.0050
Std	-0.0179	0.6054		0.0003
Skewness	-0.0250	0.4718		0.0006
Excess Kurtosis	0.0297	0.3922		0.0009
Median	-0.0703	0.0428	**	0.0049
Fear and Greed Index	-0.0022	0.9494		0.0000
VIX	-0.0009	0.9792		0.0000
T-1 Return	-0.0402	0.2476		0.0016
T-2 Return	0.0269	0.4383		0.0007
T-3 Return	0.0096	0.7830		0.0001
T-4 Return	0.0618	0.0748	*	0.0038

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

In the bivariate regression analysis, with Skewness as a fixed variable, the results are shown in

Table 5-8. It can be observed that after adding a second variable, Skewness becomes insignificant in all cases, indicating that the statistical characteristic data of RND functions constructed using the two points method is less stable in predicting Bitcoin spot returns. The results of the three-variable and four-variable regression analyses are shown in Appendix Table 4 and

Appendix Table 5.

Table 5-8: Bivariate Regression Analysis Results for Products with 1 Day to Expiration (Two Points Method)

	Skewness_Coef	Skewness_p	Skewness_Sig	Variable_Coef	Variable_p	Variable_Sig	R-squared
Mean	-0.0327	0.3485		-0.0741	0.0336	**	0.0061
Std	-0.0232	0.5069		-0.0153	0.6629		0.0009
Excess Kurtosis	-0.0232	0.5046		0.0283	0.4168		0.0014
Median	-0.0307	0.3768		-0.0727	0.0367	**	0.0059
Fear and Greed Index	-0.0249	0.4735		-0.0008	0.9805		0.0006
VIX	-0.0251	0.4715		0.0015	0.9648		0.0006
T-1 Return	-0.0339	0.3380		-0.0466	0.1877		0.0027
T-2 Return	-0.0236	0.4977		0.0256	0.4613		0.0013
T-3 Return	-0.0248	0.4754		0.0091	0.7942		0.0007
T-4 Return	-0.0221	0.5253		0.0608	0.0804	*	0.0043

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

To compare with the regression model constructed using the single point method, this section also selects Skewness, Median, and T-4 Return as the three variables for the model. The model results are shown in Table 5-9, indicating that Skewness is not significant, and the regression model constructed using the two points method does not have stable predictive ability.

Table 5-9: Regression Analysis Results for Products with 1 Day to Expiration (Two Points Method)

	Coefficient	p value	Significance	R-squared	MSE
Skewness	-0.0278	0.4233		0.0098	0.9890
Median	-0.0742	0.0330	**		
T-4 Return	0.0625	0.0718	*		

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

5.2.3 Comparative Analysis of the Two Methods

This research compares the single point method and the two points method for constructing RND for Bitcoin option products with 1 day to expiration, and conducts empirical regression analysis. The empirical results show that there are significant differences between the two methods in variable selection and predictive effects.

First, in terms of sample size, the single point method (832) has one more effective sample than the two points method (831). Although this difference is small, it reflects that the two points method may face technical limitations of being unable to fit in certain market scenarios, indicating that the single point method has better stability in practical applications.

For the return prediction regression model with 1 day to expiration, this research finally selects a three-variable regression model, including Skewness, Median, and T-4 Return. The statistical characteristics calculated by the two methods have different impacts on model performance. The model performance comparison is shown in Table 5-10. Under the single point method, all three variables reach statistically significant levels. As the results show, Skewness has a positive effect on returns, which contrasts with the research results of Bali and Murray (2013) and Conrad et al. (2013). This difference may stem from structural differences between the cryptocurrency market and traditional stock markets, as well as differences in investor behavior patterns. T-4 Return has significant predictive power for Bitcoin daily returns, which is consistent with the research of Liu and Tsyvinski (2021) and Liu et al. (2023), indicating that there is a significant momentum effect in the cryptocurrency market, with previous returns effectively predicting future return performance. Median has a negative effect. In contrast, under the two points method, the Skewness variable is not significant, with only Median and T-4 Return maintaining significant levels, indicating that the single point method can more effectively capture the impact of market skewness characteristics on price trends.

In terms of overall model predictive effect, the explanatory power of the single point method (R-squared = 0.0130) is better than that of the two points method (R-squared = 0.0098), and the Mean Squared Error (MSE) is also lower (0.9858 < 0.9890). This result shows that the single point method not only performs better in variable significance but also has better prediction accuracy than the two points method.

To further evaluate the predictive ability of the two methods, this research uses 60% of the total sample size as the initial sample length to calculate the Out-of-sample R-squared (R_{OS}^2). The results show that the out-of-sample R-squared for the single point method is 0.0134, while that for the two points method is 0.0121. This indicates that in out-of-sample prediction, the out-of-sample R-squared value of the single point method is slightly higher than that of the two points method, meaning that its out-of-sample predictive ability is relatively better. This result is consistent with the in-sample analysis, further supporting the advantages of the single point method in practical applications.

Table 5-10: Comparison of Regression Analysis Results for Products with 1 Day to Expiration
(Left: Single Point Method; Right: Two Points Method)

	單點配適法						雙點配適法					
	Coef.	p value	Sig.	R-squared	MSE	R_{OS}^2	Coef.	p value	Sig.	R-squared	MSE	R_{OS}^2
Skewness	0.0629	0.0690	*	0.0130	0.9858	0.0134	-0.0278	0.4233		0.0098	0.9890	0.0121
Median	-0.0746	0.0311	**				-0.0742	0.0330	**			
T-4 Return	0.0626	0.0705	*				0.0625	0.0718	*			

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

In summary, this research finds that when predicting Bitcoin daily returns, the single point method has the following advantages over the two points method: (1) higher stability in sample fitting; (2) more effective capture of market skewness characteristics; (3) better model predictive ability, showing relative advantages in both in-sample and out-of-sample evaluations.

5.3 Empirical Regression Analysis of Products with 7 Days to Expiration

5.3.1 Tail GPD Single Point Method

This section uses option products expiring daily from January 15, 2021, to April 19, 2024, deriving RND from the observation date seven days before expiration, and constructs complete RND functions using the single point method. Statistical characteristics such as mean, standard deviation, skewness, and kurtosis are then calculated as explanatory variables. Using the next period's Bitcoin spot return as the explained variable, multi-level regression analysis is conducted to observe whether RND has predictive effects. The descriptive statistics of the variables are shown in Table 5-11, with a total of 119 samples. The means of Skewness and Excess Kurtosis are both greater than 0, and there are relatively few extreme values.

Table 5-11: Descriptive Statistics of RND Statistical Characteristics and Bitcoin Returns for Products with 7 Days to Expiration (Single Point Method)

	Count	Mean	Std	Min	25%	Median	75%	Max
T Return (Y)	119	-0.0070	0.0978	-0.3516	-0.0511	-0.0071	0.0423	0.3071
Mean	119	36529.4634	13781.4125	16591.1342	26070.6956	35496.4638	45453.7462	69393.9514
Std	119	3599.4222	2428.7791	710.3376	1727.0200	2868.8427	5085.2385	16566.9193
Skewness	119	0.0719	0.6877	-1.4832	-0.2484	0.0372	0.2935	4.2156
Excess Kurtosis	119	2.1573	2.9048	0.4791	1.2927	1.6335	1.9781	26.4692
Median	119	36497.0748	13738.8186	16678.0000	26066.9000	35379.1000	45497.1000	69137.5000
Fear and Greed Index	119	46.5378	22.4287	9.0000	25.0000	48.0000	69.0000	93.0000
VIX	119	20.0029	5.1284	12.2800	16.2950	18.8100	22.8100	32.0200
T-1 Return	119	-0.0060	0.0980	-0.3516	-0.0480	-0.0065	0.0463	0.3071
T-2 Return	119	-0.0061	0.0980	-0.3516	-0.0480	-0.0065	0.0463	0.3071
T-3 Return	119	-0.0055	0.0977	-0.3516	-0.0439	-0.0065	0.0463	0.3071
T-4 Return	119	-0.0055	0.0977	-0.3516	-0.0439	-0.0065	0.0463	0.3071

Following the regression analysis method set in Section 4.5, the univariate regression analysis results are shown in Table 5-12. It can be observed that Mean, Std, and Median are significant, while the individual predictive ability of market sentiment indicators such as the Cryptocurrency Fear and Greed Index and Volatility Index (VIX) remains extremely low.

Table 5-12: Univariate Regression Analysis Results for Products with 7 Days to Expiration (Single Point Method)

	Coefficient	p value	Significance	R-squared
Mean	-0.1559	0.0904	*	0.0243
Std	-0.1547	0.0931	*	0.0239
Skewness	-0.0607	0.5122		0.0037
Excess Kurtosis	-0.1145	0.2148		0.0131
Median	-0.1553	0.0917	*	0.0241
Fear and Greed Index	0.0277	0.7648		0.0008
VIX	-0.0505	0.5858		0.0025
T-1 Return	-0.0398	0.6677		0.0016
T-2 Return	0.0126	0.8920		0.0002
T-3 Return	0.0043	0.9626		0.0000
T-4 Return	-0.0849	0.3587		0.0072

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

In conducting multiple regression analyses with two, three, and four variables, we found that most explanatory variables did not exhibit significant predictive effects, as documented in Appendix Table 6 to Appendix Table 8. This phenomenon indicates that merely increasing the number of variables cannot effectively enhance the model's predictive capability and may instead lead to overfitting problems.

After iteratively testing various variable combinations, our research discovered that when predicting Bitcoin weekly returns, the pairing of Excess Kurtosis and Median demonstrated superior predictive performance. Building on this foundation, we further incorporated market sentiment indicators by adding the Cryptocurrency Fear and Greed Index to the model, which exhibited significant predictive power. Through careful selection of variable combinations, rather than indiscriminately increasing the number of variables, our research ultimately identified a prediction model with both statistical significance and economic meaning. The regression model results are presented in Table 5-13.

Table 5-13: RND Regression Analysis Results for Options with 7 Days to Expiration (Single Point Method)

	Coefficient	p value	Significance	R-squared	MSE
Excess Kurtosis	-0.1620	0.0874	*	0.0666	0.9256
Median	-0.2706	0.0144	**		
Fear and Greed Index	0.2171	0.0561	*		

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

5.3.2 Tail GPD Two Points Method

This section employs options contracts expiring daily between January 15, 2021, and April 19, 2024, using observations 7 days prior to expiration to derive the RND. We construct the complete RND function using the two points method, then calculate statistics including mean, standard deviation, skewness, and kurtosis as explanatory variables. Using subsequent Bitcoin spot returns as the dependent variable, we conduct multi-level regression analyses to examine whether the RND possesses predictive power. The descriptive statistics of the variables are presented in

Table 5-14, with a total sample size of 119. Both Skewness and Excess Kurtosis have means greater than 0, and the descriptive statistics are highly similar to those of the single point method, indicating that for weekly returns, the two methods do not exhibit substantial differences.

Table 5-14: Descriptive Statistics of RND Characteristics and Bitcoin Returns for Options with 7 Days to Expiration (Two Points Method)

	Count	Mean	Std	Min	25%	Median	75%	Max
T Return (Y)	119	-0.0070	0.0978	-0.3516	-0.0511	-0.0071	0.0423	0.3071
Mean	119	36529.5600	13781.3082	16592.5497	26071.3029	35496.4638	45453.7462	69393.9514
Std	119	3598.8321	2429.3086	708.7932	1725.0112	2868.8427	5085.2385	16566.9193
Skewness	119	0.0723	0.6879	-1.4832	-0.2469	0.0372	0.2912	4.2156
Excess Kurtosis	119	2.1404	2.9086	0.4327	1.2819	1.6078	1.9556	26.4692
Median	119	36497.0748	13738.8186	16678.0000	26066.9000	35379.1000	45497.1000	69137.5000
Fear and Greed Index	119	46.5378	22.4287	9.0000	25.0000	48.0000	69.0000	93.0000
VIX	119	20.0029	5.1284	12.2800	16.2950	18.8100	22.8100	32.0200
T-1 Return	119	-0.0060	0.0980	-0.3516	-0.0480	-0.0065	0.0463	0.3071
T-2 Return	119	-0.0061	0.0980	-0.3516	-0.0480	-0.0065	0.0463	0.3071
T-3 Return	119	-0.0055	0.0977	-0.3516	-0.0439	-0.0065	0.0463	0.3071
T-4 Return	119	-0.0055	0.0977	-0.3516	-0.0439	-0.0065	0.0463	0.3071

Following the regression analysis methodology specified in Section 4.5, the univariate regression analysis results are presented in Table 5-15. We observe that Mean, Standard Deviation, and Median exhibit statistical significance.

Table 5-15: Univariate Regression Analysis Results for Options with 7 Days to Expiration (Single Point Method)

	Coefficient	p value	Significance	R-squared
Mean	-0.1559	0.0904	*	0.0243
Std	-0.1547	0.0930	*	0.0239
Skewness	-0.0608	0.5115		0.0037
Excess Kurtosis	-0.1157	0.2100		0.0134
Median	-0.1553	0.0917	*	0.0241
Fear and Greed Index	0.0277	0.7648		0.0008
VIX	-0.0505	0.5858		0.0025
T-1 Return	-0.0398	0.6677		0.0016
T-2 Return	0.0126	0.8920		0.0002
T-3 Return	0.0043	0.9626		0.0000
T-4 Return	-0.0849	0.3587		0.0072

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

The two points method, when conducting multiple regression analyses with two, three, and four variables, encounters the same issues as the single point method, with most explanatory variables lacking significant predictive effects. These results are documented in

Appendix Table 9 to

Appendix Table 11.

To facilitate comparison with the regression model constructed using the single point method, this section also selects Excess Kurtosis, Median, and the Cryptocurrency Fear and Greed Index as the three variables for inclusion in the model. The model results are presented in Table 5-16, with variable significance and model predictive capability closely resembling those of the single point method model.

Table 5-16: RND Regression Analysis Results for Options with 7 Days to Expiration (Two Points Method)

	Coefficient	p value	Significance	R-squared	MSE
Excess Kurtosis	-0.1620	0.0874	*	0.0666	0.9256
Median	-0.2706	0.0144	**		
Fear and Greed Index	0.2171	0.0561	*		

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

5.3.3 Comparative Analysis of the Two Methods

For Bitcoin options contracts with 7 days to expiration, our research employed both the single point method and the two points method to construct RNDs and conducted empirical regression analyses for comparison. The empirical results indicate that both methods can identify statistically significant variable combinations when predicting weekly returns.

Regarding descriptive statistics, the RND characteristics calculated by both methods are remarkably similar, demonstrating that both the single point method and the two points method effectively capture market information for weekly return prediction. This result contrasts with the differences observed in daily return prediction, suggesting that the longer prediction horizon for weekly returns may reduce the relative impact of tail fitting methods on overall prediction outcomes.

In terms of model predictive performance, our research also selected a three-variable regression model, with both methods incorporating Excess Kurtosis, Median, and the Cryptocurrency Fear and Greed Index as the optimal predictor combination. The models exhibit identical explanatory power (R-squared = 0.0666) and prediction error (MSE = 0.9256), as summarized in Table 5-17. The significant predictive power of Excess Kurtosis aligns with findings by Amaya et al. (2015), who noted that excess kurtosis effectively captures extreme market risks, with its predictive ability being particularly significant during financial crises. Our research also found that the Cryptocurrency Fear and Greed Index demonstrates significant predictive power for weekly returns, consistent with He et al. (2023), who discovered that this index possesses significant predictive ability both in-sample and out-of-sample, predicting cryptocurrency returns over periods ranging from one day to one week. Additionally, López-Cabarcos et al. (2021) indicated that investor sentiment indicators have significant predictive power for short-term Bitcoin price movements.

To further examine the out-of-sample predictive capability of the models, our research used 80% of the total sample size as the initial sample length to calculate the out-of-sample R-squared (R_{OS}^2). Empirical results show that the single point method's out-of-sample R-squared value is 0.3342, slightly higher than the two points method's 0.3335. Though the difference is minimal, it still indicates that the single point method performs marginally better in out-of-sample prediction. Notably, the out-of-sample R-squared values for both methods are positive and substantial, indicating that our proposed prediction models possess economic value compared to simple historical average models.

Table 5-17: Comparison of RND Regression Analysis Results for Options with 7 Days to Expiration

(Left: Single Point Method; Right: Two Points Method)

	單點配適法						雙點配適法					
	Coef.	p value	Sig.	R-squared	MSE	R_{OS}^2	Coef.	p value	Sig.	R-squared	MSE	R_{OS}^2

Excess Kurtosis	-0.1620	0.0874	*	0.0666	0.9256	0.3342	-0.1620	0.0874	*	0.0666	0.9256	0.3335
Median	-0.2706	0.0144	**				-0.2706	0.0144	**			
Fear and Greed Index	0.2171	0.0561	*				0.2171	0.0561	*			

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

In summary, our research finds that when predicting weekly returns, the single point method not only achieves comparable in-sample predictive performance to the two points method but also slightly outperforms in out-of-sample prediction. Combined with the aforementioned computational efficiency advantages, this demonstrates that the single point method indeed possesses greater development potential for practical applications.

5.4 Summary

This chapter compares the performance differences between the single point method and the two points method in constructing risk-neutral probability density functions through empirical analysis. Our research finds that the two methods exhibit different characteristics and advantages across different prediction horizons.

For daily return prediction, the single point method clearly outperforms the two points method. First, regarding sample completeness, the single point method (832 samples) preserves more valid samples than the two points method (831 samples), demonstrating superior stability when handling extreme market scenarios. Second, in terms of variable predictive capability, the regression model constructed using the single point method achieves statistical significance for three variables—Skewness, Median, and the fourth lagged return (T-4 Return)—with model explanatory power (R-squared = 0.0130) superior to that of the two points method (R-squared = 0.0098). This result indicates that the single point method more effectively captures short-term market volatility characteristics.

For weekly return prediction, the performance of both methods is more comparable. Empirical results show that both methods identify Excess Kurtosis, Median, and the Cryptocurrency Fear and Greed Index as the optimal predictor combination, with identical model explanatory power (R-squared = 0.0666) and prediction error (MSE = 0.9256). This phenomenon may reflect that in longer prediction horizons, the impact of tail fitting methods on overall prediction outcomes is relatively diminished.

Our out-of-sample prediction results demonstrate that both methods generate positive out-of-sample R-squared values for both daily and weekly return predictions, with the single point method consistently yielding higher values than the two points method. This indicates that compared to historical average models, our proposed prediction models possess substantial economic value.

Regarding computational efficiency, the single point method demonstrates clear advantages. When tested on identical hardware, the single point method's average execution time is 309.86 seconds, approximately 10.95% less than the two points method's 347.95 seconds. More importantly, the single point method exhibits a smaller standard deviation in execution time, indicating more stable computational performance—a particularly important characteristic for large-scale empirical analyses.

Synthesizing our research findings, the single point method not only performs better in daily return prediction but also offers significant advantages in computational efficiency. Although the performance difference between the two methods is less pronounced for weekly return prediction, considering the computational efficiency advantages of the single point method, our research concludes that in practical applications, the single point method has greater development potential than the two points method, especially in scenarios requiring processing of large sample sizes or higher time sensitivity. This finding has important practical implications for subsequent development of automated trading strategies or risk monitoring systems.

Furthermore, our research also discovers that the statistical characteristics of risk-neutral probability density functions exhibit different predictive capabilities across different time scales. For short-term (daily return) prediction, skewness and historical returns demonstrate stronger predictive power, while for medium-term (weekly return) prediction, excess kurtosis and market sentiment indicators play more important roles.

6. Conclusions and Recommendations

6.1 Research Conclusions

This study proposes an innovative approach for risk-neutral density (RND) tail fitting in the Bitcoin options market, comparing the single point method with the two points method to examine their differences in return prediction capability. Through empirical analysis, our research yields the following main conclusions:

First, regarding the comparison of tail fitting methods, the single point method demonstrates significant advantages. Compared to the two points method, the single point method exhibits greater stability and computational efficiency when handling extreme market scenarios. Empirical results show that the single point method reduces average execution time by approximately 10.95% compared to the two points method, while also offering advantages in sample completeness, effectively avoiding fitting failures caused by severe market fluctuations.

For predictive capability of daily and weekly returns, the two methods exhibit certain differences in performance. In predicting daily returns, the RND function constructed using the single point method more accurately captures market skewness and demonstrates statistical significance for variables including Skewness (Bali & Murray, 2013; Chang et al., 2013; Conrad et al., 2013), Median, and the fourth lagged return (T-4 Return) (Liu & Tsyvinski, 2021) in the regression model. Its explanatory power (R-squared = 0.0130) and predictive accuracy both surpass those of the two points method (R-squared = 0.0098). This indicates that the single point method offers greater application value in studying short-term market fluctuations.

However, for weekly return prediction, the performance of both methods is more comparable. Both the single point method and the two points method identify Excess Kurtosis (Amaya et al., 2015), Median, and the Cryptocurrency Fear and Greed Index (M. He et al., 2023) as the optimal predictor combination, with identical model explanatory power (R-squared = 0.0666) and prediction error (MSE = 0.9256). This reflects that in longer prediction horizons, the impact of tail fitting methods on overall prediction outcomes is relatively diminished, while market sentiment indicators play a more crucial role. The research results demonstrate that the statistical characteristics of risk-neutral probability density functions exhibit different predictive capabilities across different time scales. For short-term (daily return) prediction, skewness and historical returns demonstrate stronger predictive power, while for medium-term (weekly return) prediction, excess kurtosis and market sentiment indicators play more important roles.

Our out-of-sample prediction results show that for weekly return prediction, both the single point method and the two points method generate positive out-of-sample R-squared values (0.3342 and 0.3335, respectively), indicating that compared to historical average models, our proposed prediction models possess substantial economic value (Campbell & Thompson, 2008). This finding has important implications for practical investment, demonstrating that risk-neutral probability density characteristics extracted from options prices indeed contain valuable forward-looking market information.

Finally, our empirical results further validate the importance of RND statistical characteristics in market prediction. Higher-order moments such as skewness and excess kurtosis not only reflect market participants' expectations of extreme events but also provide valuable market sentiment information (Chordia et al., 2021), offering quantitative foundations for investment decisions and risk management. Particularly in cryptocurrency markets such as Bitcoin, due to their high volatility and unique market participant structure, the information embedded in these statistical characteristics may be more abundant than in traditional financial markets.

6.2 Research Recommendations

Based on our research findings, we propose the following recommendations for future research and practical applications:

1. Methodological Application and Improvement

The single point method demonstrates significant advantages in stability and computational efficiency, making it particularly suitable for high-frequency trading or research scenarios requiring processing of large sample sizes. We recommend that future research further optimize the algorithmic structure of the single point method to enhance its adaptability in extreme market scenarios.

2. Expansion of Research Scope

This study focuses on the Bitcoin options market. Future research could extend the scope to other cryptocurrencies or traditional financial markets to examine the applicability and performance differences of the single point method and the two points method across different market structures. For example, using Ethereum options as the research subject or analyzing stock index options markets to explore the predictive capability of RND statistical characteristics in different markets.

3. Integration with Market Microstructure Factors

This study primarily focuses on the implied information in options prices. Future research could further incorporate market microstructure factors (such as trading volume, bid-ask spread, etc.) to analyze their impact on RND statistical characteristics. Particularly in highly volatile cryptocurrency markets, microstructure factors may play important roles in market expectations and price formation mechanisms.

Appendix

Appendix Table 1: Three-Variable Regression Analysis Results for Products with 1 Day to Expiration (Single Point Method)

	Skewness_Coef	Skewness_p	Skewness_Sig	Kurtosis_Coef	Kurtosis_p	Kurtosis_Sig	Variable_Coef	Variable_p	Variable_Sig	R-squared
Mean	0.0844	0.4362		-0.0360	0.7387		-0.0783	0.0263	**	0.0101
Std	0.1232	0.2516		-0.0661	0.5387		-0.0239	0.4914		0.0047
Median	0.1055	0.3262		-0.0439	0.6833		-0.0718	0.0392	**	0.0093
Fear and Greed Index	0.1220	0.2566		-0.0636	0.5538		-0.0068	0.8444		0.0042
VIX	0.1211	0.2600		-0.0631	0.5571		0.0002	0.9961		0.0042
T-1 Return	0.1107	0.3049		-0.0555	0.6060		-0.0333	0.3398		0.0053
T-2 Return	0.1233	0.2513		-0.0643	0.5492		0.0300	0.3866		0.0051
T-3 Return	0.1201	0.2638		-0.0620	0.5641		0.0093	0.7899		0.0043
T-4 Return	0.1229	0.2518		-0.0662	0.5369		0.0610	0.0786	*	0.0079

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Appendix Table 2: Four-Variable Regression Analysis Results for Products with 1 Day to Expiration (Single Point Method)

	Skewness_Coef	Skewness_p	Skewness_Sig	Kurtosis_Coef	Kurtosis_p	Kurtosis_Sig	Std_Coef	Std_p	Std_Sig	Variable_Coef	Variable_p	Variable_Sig	R-squared
Mean	0.0827	0.4478		-0.0343	0.7513		0.0063	0.8655		-0.0807	0.0341	**	0.0101
Median	0.1034	0.3373		-0.0411	0.7034		0.0111	0.7750		-0.0769	0.0496	**	0.0094
Fear and Greed Index	0.1234	0.2515		-0.0661	0.5386		-0.0235	0.5088		-0.0016	0.9644		0.0048
VIX	0.1233	0.2516		-0.0662	0.5383		-0.0240	0.4910		-0.0012	0.9716		0.0048
T-1 Return	0.1127	0.2968		-0.0584	0.5878		-0.0252	0.4691		-0.0343	0.3269		0.0059
T-2 Return	0.1252	0.2442		-0.0671	0.5326		-0.0228	0.5113		0.0292	0.4005		0.0056
T-3 Return	0.1224	0.2552		-0.0651	0.5452		-0.0233	0.5046		0.0072	0.8356		0.0048
T-4 Return	0.1248	0.2451		-0.0688	0.5217		-0.0213	0.5401		0.0600	0.0837	*	0.0083

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Appendix Table 3: Four-Variable Regression Analysis Results Based on the Three-Variable Model (Daily Return Single Point Method)

	Skewness_Coef	Skewness_p	Skewness_Sig	Median_Coef	Median_p	Median_Sig	T-4 Return_Coef	T-4 Return_p	T-4 Return_Sig	Variable_Coef	Variable_p	Variable_Sig	R-squared
Mean	0.0177	0.7255		0.1761	0.3957		0.0633	0.0672	*	-0.2567	0.2202		0.0148
Std	0.0637	0.0663	*	-0.0823	0.0342	**	0.0635	0.0670	*	0.0169	0.6627		0.0132
Excess Kurtosis	0.1070	0.3185		-0.0733	0.0350	**	0.0628	0.0697	*	-0.0467	0.6641		0.0132
Fear and Greed Index	0.0612	0.0771	*	-0.0966	0.0209	**	0.0576	0.0998	*	0.0397	0.3479		0.0140
VIX	0.0625	0.0709	*	-0.0830	0.0250	**	0.0610	0.0788	*	-0.0237	0.5215		0.0135
T-1 Return	0.0598	0.0855	*	-0.0736	0.0335	**	0.0628	0.0695	*	-0.0329	0.3433		0.0141
T-2 Return	0.0640	0.0647	*	-0.0751	0.0302	**	0.0618	0.0743	*	0.0293	0.3970		0.0139
T-3 Return	0.0631	0.0684	*	-0.0750	0.0305	**	0.0631	0.0685	*	0.0141	0.6835		0.0132

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Appendix Table 4: Three-Variable Regression Analysis Results for Products with 1 Day to Expiration (Two Points Method)

	Skewness_Coef	Skewness_p	Skewness_Sig	Kurtosis_Coef	Kurtosis_p	Kurtosis_Sig	Variable_Coef	Variable_p	Variable_Sig	R-squared
Mean	-0.0310	0.3751		0.0250	0.4717		-0.0730	0.0366	**	0.0067
Std	-0.0214	0.5418		0.0285	0.4133		-0.0157	0.6545		0.0017
Median	-0.0291	0.4034		0.0238	0.4941		-0.0712	0.0412	**	0.0064
Fear and Greed Index	-0.0233	0.5050		0.0283	0.4172		0.0005	0.9896		0.0014
VIX	-0.0232	0.5069		0.0283	0.4177		0.0001	0.9988		0.0014
T-1 Return	-0.0320	0.3677		0.0249	0.4755		-0.0447	0.2081		0.0033
T-2 Return	-0.0219	0.5307		0.0280	0.4209		0.0254	0.4661		0.0021
T-3 Return	-0.0230	0.5092		0.0287	0.4107		0.0102	0.7691		0.0015
T-4 Return	-0.0203	0.5595		0.0281	0.4198		0.0607	0.0809	*	0.0051

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Appendix Table 5: Four-Variable Regression Analysis Results for Products with 1 Day to Expiration (Two Points Method)

	Skewness_Coef	Skewness_p	Skewness_Sig	Kurtosis_Coef	Kurtosis_p	Kurtosis_Sig	Std_Coef	Std_p	Std_Sig	Variable_Coef	Variable_p	Variable_Sig	R-squared
Mean	-0.0355	0.3182		0.0241	0.4896		0.0269	0.5032		-0.0863	0.0318	**	0.0072
Median	-0.0350	0.3249		0.0220	0.5277		0.0358	0.3963		-0.0914	0.0305	**	0.0073
Fear and Greed Index	-0.0215	0.5402		0.0287	0.4102		-0.0170	0.6400		0.0050	0.8909		0.0017
VIX	-0.0213	0.5460		0.0285	0.4136		-0.0157	0.6542		-0.0008	0.9813		0.0017
T-1 Return	-0.0301	0.3996		0.0251	0.4721		-0.0168	0.6307		-0.0451	0.2040		0.0036
T-2 Return	-0.0201	0.5675		0.0282	0.4175		-0.0153	0.6623		0.0251	0.4703		0.0023
T-3 Return	-0.0213	0.5439		0.0288	0.4084		-0.0148	0.6742		0.0088	0.8023		0.0017
T-4 Return	-0.0186	0.5957		0.0283	0.4165		-0.0147	0.6740		0.0605	0.0823	*	0.0053

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Appendix Table 6: Two-Variable Regression Analysis Results for Products with 7 Days to Expiration (Single Point Method)

	Skewness_Coef	Skewness_p	Skewness_Sig	Variable_Coef	Variable_p	Variable_Sig	R-squared
Mean	-0.0401	0.6654		-0.1505	0.1066		0.0259
Std	0.0185	0.8601		-0.1636	0.1212		0.0242
Excess Kurtosis	0.0353	0.7822		-0.1390	0.2782		0.0138
Median	-0.0429	0.6430		-0.1502	0.1062		0.0259
Fear and Greed Index	-0.1021	0.3462		0.0803	0.4581		0.0084
VIX	-0.0750	0.4298		-0.0667	0.4830		0.0079
T-1 Return	-0.0538	0.5779		-0.0249	0.7963		0.0043
T-2 Return	-0.0653	0.4901		0.0248	0.7932		0.0043
T-3 Return	-0.0620	0.5077		0.0115	0.9021		0.0038
T-4 Return	-0.0515	0.5811		-0.0789	0.3982		0.0098

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Appendix Table 7: Three-Variable Regression Analysis Results for Products with 7 Days to Expiration (Single Point Method)

	Skewness_Coef	Skewness_p	Skewness_Sig	Kurtosis_Coef	Kurtosis_p	Kurtosis_Sig	Variable_Coef	Variable_p	Variable_Sig	R-squared
Mean	0.0685	0.5935		-0.1555	0.2228		-0.1591	0.0884	*	0.0385
Std	0.0644	0.6194		-0.0822	0.5435		-0.1410	0.2082		0.0273
Median	0.0660	0.6065		-0.1560	0.2212		-0.1591	0.0876	*	0.0386
Fear and Greed Index	-0.0054	0.9709		-0.1264	0.3322		0.0621	0.5716		0.0165
VIX	0.0163	0.9069		-0.1225	0.3708		-0.0355	0.7259		0.0148
T-1 Return	0.0820	0.5680		-0.1766	0.2032		-0.0750	0.4713		0.0182
T-2 Return	0.0312	0.8120		-0.1370	0.2893		0.0150	0.8744		0.0140
T-3 Return	0.0357	0.7840		-0.1393	0.2833		-0.0017	0.9859		0.0138
T-4 Return	0.0552	0.6700		-0.1524	0.2374		-0.0908	0.3330		0.0218

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Appendix Table 8: Four-Variable Regression Analysis Results for Products with 7 Days to Expiration (Single Point Method)

	Skewness_Coef	Skewness_p	Skewness_Sig	Kurtosis_Coef	Kurtosis_p	Kurtosis_Sig	Std_Coef	Std_p	Std_Sig	Variable_Coef	Variable_p	Variable_Sig	R-squared
Mean	0.0648	0.6165		-0.1846	0.2507		0.0612	0.7640		-0.2019	0.2372		0.0392
Median	0.0615	0.6348		-0.1861	0.2477		0.0631	0.7574		-0.2031	0.2339		0.0394
Fear and Greed Index	-0.0131	0.9283		-0.0336	0.8126		-0.1941	0.1101		0.1350	0.2543		0.0384
VIX	0.0427	0.7614		-0.0622	0.6664		-0.1431	0.2038		-0.0413	0.6830		0.0288
T-1 Return	0.1105	0.4462		-0.1196	0.4106		-0.1406	0.2105		-0.0743	0.4741		0.0317
T-2 Return	0.0603	0.6497		-0.0803	0.5561		-0.1410	0.2103		0.0145	0.8779		0.0275
T-3 Return	0.0645	0.6251		-0.0823	0.5478		-0.1410	0.2102		-0.0007	0.9940		0.0273
T-4 Return	0.0841	0.5219		-0.0957	0.4819		-0.1407	0.2094		-0.0904	0.3338		0.0353

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Appendix Table 9: Two-Variable Regression Analysis Results for Products with 7 Days to Expiration (Two Points Method)

	Skewness_Coef	Skewness_p	Skewness_Sig	Variable_Coef	Variable_p	Variable_Sig	R-squared
Mean	-0.0401	0.6654		-0.1505	0.1066		0.0259
Std	0.0185	0.8601		-0.1636	0.1212		0.0242
Excess Kurtosis	0.0353	0.7822		-0.1390	0.2782		0.0138
Median	-0.0429	0.6430		-0.1502	0.1062		0.0259
Fear and Greed Index	-0.1021	0.3462		0.0803	0.4581		0.0084
VIX	-0.0750	0.4298		-0.0667	0.4830		0.0079
T-1 Return	-0.0538	0.5779		-0.0249	0.7963		0.0043
T-2 Return	-0.0653	0.4901		0.0248	0.7932		0.0043
T-3 Return	-0.0620	0.5077		0.0115	0.9021		0.0038
T-4 Return	-0.0515	0.5811		-0.0789	0.3982		0.0098

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Appendix Table 10: Three-Variable Regression Analysis Results for Products with 7 Days to Expiration (Two Points Method)

	Skewness_Coef	Skewness_p	Skewness_Sig	Kurtosis_Coef	Kurtosis_p	Kurtosis_Sig	Variable_Coef	Variable_p	Variable_Sig	R-squared
Mean	0.0685	0.5935		-0.1555	0.2228		-0.1591	0.0884	*	0.0385
Std	0.0644	0.6194		-0.0822	0.5435		-0.1410	0.2082		0.0273
Median	0.0660	0.6065		-0.1560	0.2212		-0.1591	0.0876	*	0.0386
Fear and Greed Index	-0.0054	0.9709		-0.1264	0.3322		0.0621	0.5716		0.0165
VIX	0.0163	0.9069		-0.1225	0.3708		-0.0355	0.7259		0.0148
T-1 Return	0.0820	0.5680		-0.1766	0.2032		-0.0750	0.4713		0.0182
T-2 Return	0.0312	0.8120		-0.1370	0.2893		0.0150	0.8744		0.0140
T-3 Return	0.0357	0.7840		-0.1393	0.2833		-0.0017	0.9859		0.0138
T-4 Return	0.0552	0.6700		-0.1524	0.2374		-0.0908	0.3330		0.0218

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

Appendix Table 11: Four-Variable Regression Analysis Results for Products with 7 Days to Expiration (Two Points Method)

	Skewness_Coef	Skewness_p	Skewness_Sig	Kurtosis_Coef	Kurtosis_p	Kurtosis_Sig	Std_Coef	Std_p	Std_Sig	Variable_Coef	Variable_p	Variable_Sig	R-squared
Mean	0.0648	0.6165		-0.1846	0.2507		0.0612	0.7640		-0.2019	0.2372		0.0392
Median	0.0615	0.6348		-0.1861	0.2477		0.0631	0.7574		-0.2031	0.2339		0.0394
Fear and Greed Index	-0.0131	0.9283		-0.0336	0.8126		-0.1941	0.1101		0.1350	0.2543		0.0384
VIX	0.0427	0.7614		-0.0622	0.6664		-0.1431	0.2038		-0.0413	0.6830		0.0288
T-1 Return	0.1105	0.4462		-0.1196	0.4106		-0.1406	0.2105		-0.0743	0.4741		0.0317
T-2 Return	0.0603	0.6497		-0.0803	0.5561		-0.1410	0.2103		0.0145	0.8779		0.0275
T-3 Return	0.0645	0.6251		-0.0823	0.5478		-0.1410	0.2102		-0.0007	0.9940		0.0273
T-4 Return	0.0841	0.5219		-0.0957	0.4819		-0.1407	0.2094		-0.0904	0.3338		0.0353

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level

References

Akyildirim, E., Corbet, S., Lucey, B., Sensoy, A., & Yarovaya, L. (2020). The relationship between implied volatility and cryptocurrency returns. *Finance Research Letters*, *33*, 101212. https://doi.org/10.1016/j.frl.2019.06.010

Amaya, D., Christoffersen, P., Jacobs, K., & Vasquez, A. (2015). Does realized skewness predict the cross-section of equity returns? *Journal of Financial Economics*, *118*(1), 135–167. https://doi.org/10.1016/j.jfineco.2015.02.009

Ammann, M., & Feser, A. (2019). Robust estimation of risk-neutral moments. *Journal of Futures Markets*, *39*(9), 1137–1166. https://doi.org/10.1002/fut.22020

Bakshi, G., Kapadia, N., & Madan, D. (2003). Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options. *The Review of Financial Studies*, *16*(1), 101–143. https://doi.org/10.1093/rfs/16.1.0101

Bali, T. G., & Murray, S. (2013). Does Risk-Neutral Skewness Predict the Cross Section of Equity Option Portfolio Returns? *The Journal of Financial and Quantitative Analysis*, *48*(4), 1145–1171.

Bali, T. G., & Zhou, H. (2016). Risk, Uncertainty, and Expected Returns. *The Journal of Financial and Quantitative Analysis*, *51*(3), 707–735.

Balkema, A. A., & Haan, L. de. (1974). Residual Life Time at Great Age. *The Annals of Probability*, *2*(5), 792–804. https://doi.org/10.1214/aop/1176996548

Baur, D. G., & Smales, L. A. (2022). Trading behavior in bitcoin futures: Following the “smart money.” *Journal of Futures Markets*, *42*(7), 1304–1323. https://doi.org/10.1002/fut.22332

Birru, J., & Figlewski, S. (2012). Anatomy of a meltdown: The risk neutral density for the S&P 500 in the fall of 2008. *Journal of Financial Markets*, *15*(2), 151–180. https://doi.org/10.1016/j.finmar.2011.09.001

Black, F., & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, *81*(3), 637–654.

Bliss, R. R., & Panigirtzoglou, N. (2004). Option-Implied Risk Aversion Estimates. *The Journal of Finance*, *59*(1), 407–446. https://doi.org/10.1111/j.1540-6261.2004.00637.x

Bondarenko, O. (2000). *Recovering Risk-Neutral Densities: A New Nonparametric Approach* (SSRN Scholarly Paper No. 246063). Social Science Research Network. https://doi.org/10.2139/ssrn.246063

Böök, A., Imbet, J. F., Reinke, M., & Sala, C. (2025). The Forecasting Power of Short-Term Options. *The Journal of Derivatives*, *32*(3), 80–116. https://doi.org/10.3905/jod.2025.1.221

Breeden, D. T., & Litzenberger, R. H. (1978). Prices of State-Contingent Claims Implicit in Option Prices. *The Journal of Business*, *51*(4), 621–651.

Campbell, J. Y., & Thompson, S. B. (2008). Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average? *The Review of Financial Studies*, *21*(4), 1509–1531. https://doi.org/10.1093/rfs/hhm055

Chang, B. Y., Christoffersen, P., & Jacobs, K. (2013). Market skewness risk and the cross section of stock returns. *Journal of Financial Economics*, *107*(1), 46–68. https://doi.org/10.1016/j.jfineco.2012.07.002

Chen, R.-R., Hsieh, P., & Huang, J. (2018). Crash risk and risk neutral densities. *Journal of Empirical Finance*, *47*, 162–189. https://doi.org/10.1016/j.jempfin.2018.03.006

Chordia, T., Lin, T.-C., & Xiang, V. (2021). Risk-Neutral Skewness, Informed Trading, and the Cross Section of Stock Returns. *Journal of Financial and Quantitative Analysis*, *56*(5), 1713–1737. https://doi.org/10.1017/S0022109020000551

Christoffersen, P., Jacobs, K., & Chang, B. Y. (2013). Chapter 10—Forecasting with Option-Implied Information. In G. Elliott & A. Timmermann (Eds.), *Handbook of Economic Forecasting* (Vol. 2, pp. 581–656). Elsevier. https://doi.org/10.1016/B978-0-444-53683-9.00010-4

Conrad, J., Dittmar, R. F., & Ghysels, E. (2013). Ex Ante Skewness and Expected Stock Returns. *The Journal of Finance*, *68*(1), 85–124. https://doi.org/10.1111/j.1540-6261.2012.01795.x

Cortés, L. M., Mora-Valencia, A., & Perote, J. (2020). Retrieving the implicit risk neutral density of WTI options with a semi-nonparametric approach. *The North American Journal of Economics and Finance*, *54*, 100862. https://doi.org/10.1016/j.najef.2018.10.010

Cujean, J., & Hasler, M. (2017). Why Does Return Predictability Concentrate in Bad Times? *The Journal of Finance*, *72*(6), 2717–2758. https://doi.org/10.1111/jofi.12544

Deribit. (2025). https://www.deribit.com/

Deribit Options. (2025). https://www.deribit.com/

Dong, B., Xu, W., & Cui, Z. (2024). Implied Willow Tree. *The Journal of Derivatives*, *31*(4), 44–74. https://doi.org/10.3905/jod.2024.1.200

Feng, Y., He, M., & Zhang, Y. (2024). Market Skewness and Stock Return Predictability: New Evidence from China. *Emerging Markets Finance and Trade*, *60*(2), 233–244. https://doi.org/10.1080/1540496X.2023.2217327

Figlewski, S. (2008). *Estimating the Implied Risk Neutral Density for the U.S. Market Portfolio* (SSRN Scholarly Paper No. 1256783). Social Science Research Network. https://papers.ssrn.com/abstract=1256783

Fuertes, A.-M., Liu, Z., & Tang, W. (2022). Risk-neutral skewness and commodity futures pricing. *Journal of Futures Markets*, *42*(4), 751–785. https://doi.org/10.1002/fut.22308

Glatzer, E., & Scheicher, M. (2005). What moves the tail? The determinants of the option-implied probability density function of the DAX index. *Journal of Futures Markets*, *25*(6), 515–536. https://doi.org/10.1002/fut.20157

Grith, M., Härdle, W. K., & Schienle, M. (2012). Nonparametric Estimation of Risk-Neutral Densities. In J.-C. Duan, W. K. Härdle, & J. E. Gentle (Eds.), *Handbook of Computational Finance* (pp. 277–305). Springer. https://doi.org/10.1007/978-3-642-17254-0_11

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies*, *33*(5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009

Hagan, P. S., & and West, G. (2006). Interpolation Methods for Curve Construction. *Applied Mathematical Finance*, *13*(2), 89–129. https://doi.org/10.1080/13504860500396032

Haslip, G. G., & and Kaishev, V. K. (2014). Lookback option pricing using the Fourier transform B-spline method. *Quantitative Finance*, *14*(5), 789–803. https://doi.org/10.1080/14697688.2014.882010

Hayashi, F. (2020). Analytically Deriving Risk-Neutral Densities from Volatility Smiles in Delta. *The Journal of Derivatives*, *27*(4), 6–12. https://doi.org/10.3905/jod.2020.1.099

He, M., Shen, L., Zhang, Y., & Zhang, Y. (2023). Predicting cryptocurrency returns for real-world investments: A daily updated and accessible predictor. *Finance Research Letters*, *58*, 104406. https://doi.org/10.1016/j.frl.2023.104406

He, Y., Peng, L., Zhang, D., & Zhao, Z. (2022). Risk Analysis via Generalized Pareto Distributions. *Journal of Business & Economic Statistics*, *40*(2), 852–867. https://doi.org/10.1080/07350015.2021.1874390

Hosking, J. R. M., & Wallis, J. R. (1987). Parameter and Quantile Estimation for the Generalized Pareto Distribution. *Technometrics*, *29*(3), 339–349. https://doi.org/10.2307/1269343

Hull, J. (2021). *Options, Futures, and Other Derivatives: Global Edition*. Pearson Deutschland. https://elibrary.pearson.de/book/99.150005/9781292410623

Jackwerth, J. (2020). What Do Index Options Teach Us About COVID-19? *The Review of Asset Pricing Studies*, *10*(4), 618–634. https://doi.org/10.1093/rapstu/raaa012

Jondeau, E., Wang, X., Yan, Z., & Zhang, Q. (2020). Skewness and index futures return. *Journal of Futures Markets*, *40*(11), 1648–1664. https://doi.org/10.1002/fut.22112

Kim, T. S., & Park, H. (2018). Is stock return predictability of option-implied skewness affected by the market state? *Journal of Futures Markets*, *38*(9), 1024–1042. https://doi.org/10.1002/fut.21921

Köse, N., Yildirim, H., Ünal, E., & Lin, B. (2024). The Bitcoin price and Bitcoin price uncertainty: Evidence of Bitcoin price volatility. *Journal of Futures Markets*, *44*(4), 673–695. https://doi.org/10.1002/fut.22487

Lehnert, T. (2022). Is Risk-Neutral Skewness an Indicator of Downside Risk? Evidence from Tail Risk Taking of Hedge Funds. *The Journal of Derivatives*, *29*(3), 65–84. https://doi.org/10.3905/jod.2022.1.148

Li, X., Wu, Z., Zhang, H., & Zhang, L. (2024). Risk-neutral skewness and stock market returns: A time-series analysis. *The North American Journal of Economics and Finance*, *70*, 102040. https://doi.org/10.1016/j.najef.2023.102040

Li, Y., Nolte, I., & Pham, M. C. (2024). Parametric risk-neutral density estimation via finite lognormal-Weibull mixtures. *Journal of Econometrics*, *241*(2), 105748. https://doi.org/10.1016/j.jeconom.2024.105748

Li, Y., Urquhart, A., Wang, P., & Zhang, W. (2021). MAX momentum in cryptocurrency markets. *International Review of Financial Analysis*, *77*, 101829. https://doi.org/10.1016/j.irfa.2021.101829

Liu, Y., & Chen, Y. (2024). Skewness risk and the cross-section of cryptocurrency returns. *International Review of Financial Analysis*, *96*, 103626. https://doi.org/10.1016/j.irfa.2024.103626

Liu, Y., Li, Z., Nekhili, R., & Sultan, J. (2023). Forecasting cryptocurrency returns with machine learning. *Research in International Business and Finance*, *64*, 101905. https://doi.org/10.1016/j.ribaf.2023.101905

Liu, Y., & Tsyvinski, A. (2021). Risks and Returns of Cryptocurrency. *The Review of Financial Studies*, *34*(6), 2689–2727. https://doi.org/10.1093/rfs/hhaa113

López-Cabarcos, M. Á., Pérez-Pico, A. M., Piñeiro-Chousa, J., & Šević, A. (2021). Bitcoin volatility, stock market and investor sentiment. Are they connected? *Finance Research Letters*, *38*, 101399. https://doi.org/10.1016/j.frl.2019.101399

Markose, S., & Alentorn, A. (2011). The Generalized Extreme Value Distribution, Implied Tail Index, and Option Pricing. *The Journal of Derivatives*, *18*(3), 35–60. https://doi.org/10.3905/jod.2011.18.3.035

McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance*, *7*(3), 271–300. https://doi.org/10.1016/S0927-5398(00)00012-8

Mei, D., Liu, J., Ma, F., & Chen, W. (2017). Forecasting stock market volatility: Do realized skewness and kurtosis help? *Physica A: Statistical Mechanics and Its Applications*, *481*, 153–159. https://doi.org/10.1016/j.physa.2017.04.020

Mohrschladt, H., & Schneider, J. C. (2021). Option-implied skewness: Insights from ITM-options. *Journal of Economic Dynamics and Control*, *131*, 104227. https://doi.org/10.1016/j.jedc.2021.104227

Monteiro, A. M., & Santos, A. A. F. (2022). Option prices for risk-neutral density estimation using nonparametric methods through big data and large-scale problems. *Journal of Futures Markets*, *42*(1), 152–171. https://doi.org/10.1002/fut.22258

Monteiro, A. M., Tütüncü, R. H., & Vicente, L. N. (2008). Recovering risk-neutral probability density functions from options prices using cubic splines and ensuring nonnegativity. *European Journal of Operational Research*, *187*(2), 525–542. https://doi.org/10.1016/j.ejor.2007.02.041

Neumann, M., & Skiadopoulos, G. (2013). Predictable Dynamics in Higher-Order Risk-Neutral Moments: Evidence from the S&P 500 Options. *The Journal of Financial and Quantitative Analysis*, *48*(3), 947–977.

Orosi, G. (2015). Estimating Option-Implied Risk-Neutral Densities: A Novel Parametric Approach. *The Journal of Derivatives*, *23*(1), 41–61. https://doi.org/10.3905/jod.2015.23.1.041

Reinke, M. (2020). Risk-Neutral Density Estimation: Looking at the Tails. *The Journal of Derivatives*, *27*(3), 99–125. https://doi.org/10.3905/jod.2019.1.090

Rubinstein, M. (1994). Implied Binomial Trees. *The Journal of Finance*, *49*(3), 771–818. https://doi.org/10.2307/2329207

Shimko, D. (1993). Bounds of probability. *Risk*, *6*(4), 33–37.

The Block. (2025). The Block. https://www.theblock.co/data/crypto-markets/options

Uberti, P. (2023). A theoretical generalization of the Markowitz model incorporating skewness and kurtosis. *Quantitative Finance*, *23*(5), 877–886. https://doi.org/10.1080/14697688.2023.2176250

Wang, Y.-H., & Yen, K.-C. (2018). The information content of option-implied tail risk on the future returns of the underlying asset. *Journal of Futures Markets*, *38*(4), 493–510. https://doi.org/10.1002/fut.21887

Welch, I., & Goyal, A. (2008). A Comprehensive Look at The Empirical Performance of Equity Premium Prediction. *The Review of Financial Studies*, *21*(4), 1455–1508. https://doi.org/10.1093/rfs/hhm014

Zulfiqar, N., & Gulzar, S. (2021). Implied volatility estimation of bitcoin options and the stylized facts of option pricing. *Financial Innovation*, *7*(1), 67. https://doi.org/10.1186/s40854-021-00280-y