TECHNISCHE UNIVERSITÄT DRESDEN

„Friedrich List“ Faculty of Transport and Traffic Sciences, Chair of Econometrics and Statistics

Diploma Thesis

Evaluating Expansion Methods for Realized Moments in Heston Model Simulations

Henry Haustein
Student number: 4685025

Supervised by

Haozhe Jiang

Dresden, XX.XX.XXXX

Contents

List of Figures

List of Tables

1 Literature Review

This chapter provides a brief overview of various approaches found in the literature that are not used in this work. The exclusion of these methods is generally due to the availability of more advanced techniques or because their results have not been sufficiently successful. The methods that are employed in this study are discussed in detail in the following chapters.

1.1 Realized Moments

For the pricing of financial derivatives, it is crucial to know the moments of returns, particularly those of monthly or quarterly returns (Barro, 2006). However, estimating the moments of such low-frequency returns can be challenging due to the limited number of observations available (Neuberger and Payne, 2021). Today, financial markets operate continuously, making it possible to obtain daily or even minute-level returns without difficulty. For example, the German stock index DAX is calculated every second (Frankfurt, n.d.). There are several approaches to estimating the moments of monthly or quarterly returns based on the moments of daily returns. One such method is proposed by Amaya et al. (2015). In this approach, the variance of daily returns is estimated using the sum of squared returns. This idea is not new and was first introduced by Andersen & Bollerslev (1998). Building upon this approach, the daily realized skewness and kurtosis can be computed using cubed and quartic returns, respectively. Thia estimator is consistent, but it does not caputure skewness coming from the leverage effect (Gallo et al., 2024). Zhang et al. (2005) reports that this approach can be highly biased and the bias depends on the sampling frequency. Liu et al. (2014) propose a new estimator on the basis of the Amaya et al. estimator which is robust to the microstructure noise at ultra-high frequency level. To transition from daily realized moments to weekly or monthly moments, a moving average approach is applied. Choe & Lee (2014) use variation processes to estimate low-frequency moments, specifically the quadratic variation of a semimartingale X. From these, the higher moments of R follow as expectation of the quadratic covariation of R and R^2 or the quadratic variation of R^2. The estimation of low-frequency variance follows the same approach as Andersen & Bollerslev (1998) and Amaya et al. (2015).

In addition to these methods, Neuberger and Payne developed an approach for esti-

mating realized moments. Since this method is utilized in this study, we will examine it in more detail in Chapter 2.

1.2 Expansion Methods

The Gram-Charlier expansion and the Edgeworth expansion are among the most well-known methods. They allow for the approximation of the density of a normal distribution while incorporating additional terms for skewness and kurtosis. Since these methods play a central role in this study, we will examine them in detail in Chapter 4.

In addition to these expansion methods, there are also the Cornish-Fisher expansion and the Saddlepoint approximation.

The Cornish-Fisher expansion, introduced by Cornish & Fisher (1938), is an asymptotic expansion that approximates the quantiles of a probability distribution based on its cumulants.

Given that z_p is the p-quantile of a normal distribution with mean μ and variance σ^2, the p-quantile of a random variable X, denoted as x_p, can be approximated as follows (only the first terms shown, as it is common practice) (Abramowitz and Stegun, 1968, p. 935):

$$x_p \approx z_p + \frac{\gamma_1}{6} He_2(z_p) + \frac{\gamma_2^*}{24} He_3(z_p) - \frac{\gamma_1^2}{36}(2 \cdot He_3(z_p) + He_1(z_p))$$

To obtain the probability density function (PDF), the quantiles x_p can be numerically computed and differentiated. $He_n(x)$ are the Hermite polynomials, and γ_1 and γ_2^* are the skewness and excess kurtosis of the approximated distribution, respectively.

Aboura & Maillard (2016) point out that the parameters γ_1 and γ_2^* do not correspond to the skewness and excess kurtosis of the approximated distribution. Instead, they denote these parameters as $s = \gamma_1$ and $k = \gamma_2^*$ and provide equations to compute the actual skewness s^* and excess kurtosis k^* of the approximated distribution. Later, Maillard (2018) published the inverse transformation, allowing one to compute the parameters s and k given the actual skewness and excess kurtosis. A corresponding table can be found in the appendix of his paper.

Aboura & Maillard (2016) also investigate the domain of validity for the Cornish-Fisher expansion. Their findings suggest that the expansion is valid for a wide range of parameters—even an excess kurtosis above 40 and skewness exceeding ± 3 are possible. However, when operating outside the validity domain, the issue is not immediately apparent in the probability density function itself. Instead, it becomes visible in the quantiles, which can turn negative.

The Saddlepoint Approximation, introduced by Daniels (1954), provides an accurate

method for approximating probability densities. While Daniels initially derived the density function, the cumulative distribution function (CDF) was later introduced by Lugannani & Rice (1980). This method is based on the moment generating function (MGF) and offers a highly precise approximation formula. Given that $M(t)$ is the moment generating function and $K(t) = \log(M(t))$ is the cumulant generating function, the approximation of the density function $f(x)$ is given by:

$$f(x)_{SP} \approx \frac{1}{\sqrt{2\pi \cdot K''(s)}} \exp(K(s) - s \cdot x) \tag{1.2.1}$$

where s is the solution of the equation $K'(s) = x$.

By definition, the cumulant generating function $K(s)$ can be approximated with a Taylor series in s which can then be used to calculate the Saddelpoint Approximation analytically.

The key advantages of this method are the high accuracy: The Saddlepoint method provides excellent approximations, even in distribution tail areas (Du, 2010; Reid, 1988). Furthermore there are no issues with negativity: The Saddlepoint Approximation does not produce negative densities. This is because the $\exp(\cdot)$ term in Equation (1.2.1) is always positive, and the denominator involves a square root, which is either positive or complex. If complex, the approximation does not exist, rather than producing invalid results.

2 The First 4 Moments

2.1 Moments and central moments

For a random variable X, the expectation, also referred to as the first moment, is given by

$$\mu = \mathbb{E}(X)$$

This expectation is estimated by the sample mean of the observed values x:

$$\hat{\mu} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Variance serves as a measure of the dispersion of the random variable X. In the special case where $\mu = 0$, the variance simplifies to

$$\sigma^2 = \mathbb{E}(X^2)$$

and is also referred to as the second moment. If $\mu \neq 0$, the variance is defined as

$$\begin{aligned} \sigma^2 &= \mathbb{E}((X-\mu)^2) \\ &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \end{aligned} \tag{2.1.1}$$

which is also known as the centered second moment. The variance is estimated using

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

where the denominator $n-1$ represents Bessel's correction, which improves the estimation of variance (Radziwill, 2017). Analogously, the r-th moment is given by

$$\mathbb{E}(X^r)$$

and the corresponding centered r-th moment is

$$\mathbb{E}((X-\mu)^r)$$

Skewness measures the asymmetry of a distribution and is defined as the standardized third moment:

$$\begin{aligned}\gamma_1 &= \frac{\mathbb{E}((X-\mu)^3)}{\sigma^3} \\ &= \frac{\mathbb{E}(X^3) - 3\mathbb{E}(X)\mathbb{E}(X^2) + 2\mathbb{E}(X)^3}{\sigma^3}\end{aligned} \tag{2.1.2}$$

Different methods exist for estimating skewness, such as those proposed by Joanes & Gill (1998):

$$\begin{aligned}b_1 &= \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^3}{\left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^{3/2}} \\ g_1 &= \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^3}{\left[\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^{3/2}} \\ G_1 &= \frac{n^2}{(n-1)(n-2)}b_1 = \frac{\sqrt{n(n-1)}}{n-2}g_1 \\ \hat{\gamma}_1 &= \frac{n}{(n-1)(n-2)}\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{\hat{\sigma}}\right)^3\end{aligned}$$

where b_1 and g_1 are estimators of the population skewness, while G_1 and $\hat{\gamma}_1$ estimate the skewness of a sample. The estimator G_1 is implemented in statistical software such as Excel, SAS, and SPSS (Doane and Seward, 2011). Kurtosis measures the tailedness of a distribution and is defined as the standardized fourth moment:

$$\begin{aligned}\gamma_2 &= \frac{\mathbb{E}((X-\mu)^4)}{\sigma^4} \\ &= \frac{\mathbb{E}(X^4) - 4\mathbb{E}(X^3)\mathbb{E}(X) + 6\mathbb{E}(X^2)\mathbb{E}(X)^2 - 3\mathbb{E}(X)^4}{\sigma^4}\end{aligned} \tag{2.1.3}$$

Various estimation methods for kurtosis exist, such as those presented by Joanes & Gill (1998):

$$\begin{aligned}g_2 &= \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^4}{\left[\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^2} \\ \hat{\gamma}_2 &= \frac{n(n+1)}{(n-1)(n-2)(n-3)}\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{\hat{\sigma}}\right)^4\end{aligned}$$

where g_2 estimates the kurtosis of a population and G_2 estimates the kurtosis of a sample. A commonly used alternative is excess kurtosis, which is obtained by subtracting

3:

$$\gamma_2^* = \gamma_2 - 3$$
$$G_2 = \frac{n-1}{(n-2)(n-3)}[(n+1)g_2 + 6]$$

This adjustment is motivated by the fact that for a standard normally distributed variable X, the kurtosis is $\gamma_2 = 3$ and the excess kurtosis is $\gamma_2^* = 0$. In general, due to the high powers involved in the definitions of skewness and kurtosis, these estimators are highly sensitive to outliers. In the following, the term moments will be used broadly to include related measures such as variance, skewness, and kurtosis.

2.2 Cumulants

Throughout this work, the concept of cumulants will also be used, which provide an alternative representation of moments. The r-th cumulant is defined as the coefficient of t^r in the logarithm of the moment generating function of X. The moment generating function of X is given by

$$M_X(t) = \mathbb{E}(e^{tX})$$

The cumulant generating function of X is then defined as

$$K_X(t) = \log(M_X(t))$$

From this definition, the first four cumulants are given by

$$\kappa_1 = \mu \tag{2.2.1}$$
$$\kappa_2 = \sigma^2 \tag{2.2.2}$$
$$\kappa_3 = \gamma_1 \sigma^3 \tag{2.2.3}$$
$$\kappa_4 = \gamma_2^* \sigma^4 \tag{2.2.4}$$

2.3 Estimating the Moments of Low-Frequency Data using High-Frequency Data

In contrast to the previously published methods Neuberger & Payne (2021) propose a novel method that only requires log prices to be martingales. Under these conditions,

they define new moment measures that approximate the standard definitions:

$$
\begin{aligned}
var^L(r) &= \mathbb{E}(x^{(2L)}(r)), \text{ where } x^{(2L)}(r) = 2(e^r - 1 - r) \\
var^E(r) &= \mathbb{E}(x^{(2E)}(r)), \text{ where } x^{(2E)}(r) = 2(re^r - e^r + 1) \\
skew(r) &= \frac{\mathbb{E}(x^{(3)}(r))}{var^L(r)^{3/2}}, \text{ where } x^{(3)}(r) = 6((e^r + 1)r - 2(e^r - 1)) \\
kurt(r) &= \frac{\mathbb{E}(x^{(4)}(r))}{var^L(r)^2}, \text{ where } x^{(4)}(r) = 12(r^2 + 2(e^r + 2)r - 6(e^r - 1))
\end{aligned}
$$

where r represents the log-return process:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right)$$

The long-horizon returns process R is defined as

$$R_t(T) = \ln\left(\frac{P_t}{P_{t-T}}\right)$$

This establishes a connection between low-frequency moments and high-frequency log returns:

$$
\begin{aligned}
var^L(R(T)) &= T \cdot var^L(r) \\
skew(R(T)) &= \left(skew(r) + 3\frac{cov(y^{(1)}, x^{(2E)}(r))}{var^L(r)^{3/2}}\right) T^{-1/2} \\
kurt(R(T)) &= \left(kurt(r) + 4\frac{cov(y^{(1)}, x^{(3)}(r))}{var^L(r)^2} + 6\frac{cov(y^{(2L)}, x^{(2L)}(r))}{var^L(r)^2}\right) T^{-1}
\end{aligned}
$$

where

$$
\begin{aligned}
y_t^{(j)} &= \sum_{u=1}^{T} \frac{x^{(j)}(R_{t-1}(u))}{T} \quad \text{for } j = 1, 2L \\
x^{(1)} &= e^r - 1
\end{aligned}
\tag{2.3.1}
$$

where equation (2.3.1) originates from Neuberger (2012). In Neuberger (2012), the aggregation property is also introduced: if g is a real function, X is a process, and for times $0 \leq s \leq t \leq u \leq T$,

$$\mathbb{E}_s(g(X_u - X_s)) = \mathbb{E}_s(g(X_u - X_t)) + \mathbb{E}_t(g(X_t - X_s))$$

then the pair (g, X) satisfies the aggregation property. This property is used, for instance, in estimating low-frequency variance by summing high-frequency variance. Fukasawa & Matsushita (2021) build on this idea and derive formulas for realized cumulants that also satisfy the aggregation property.

3 The Heston Model

3.1 Model Description

The Heston model, introduced by Heston (1993), is a stochastic volatility model in which volatility is not constant, as in the Black-Scholes-Merton model, but instead follows a random process. The dynamics of the model are given by the following system of stochastic differential equations:

$$\mathrm{d}S_t = \mu S_t \mathrm{d}t + \sqrt{v_t} S_t \mathrm{d}W_t^S \tag{3.1.1}$$

$$\mathrm{d}X_t = \mathrm{d}\log(S_t) = \left(\mu - \frac{1}{2}v_t\right)\mathrm{d}t + \sqrt{v_t}\mathrm{d}W_t^S \tag{3.1.2}$$

$$\mathrm{d}v_t = \kappa(\theta - v_t)\mathrm{d}t + \sigma\sqrt{v_t}\mathrm{d}W_t^v \tag{3.1.3}$$

$$\mathbb{E}(\mathrm{d}W_t^S \mathrm{d}W_t^v) = \rho \mathrm{d}t \tag{3.1.4}$$

where κ, θ, and σ are strictly positive parameters. The terms $\mathrm{d}W_t^S$ and $\mathrm{d}W_t^v$ represent the increments of Brownian motions with correlation ρ. The variable S_t denotes the price of an asset, such as a stock, bond, or foreign exchange rate. The process X_t represents the logarithm of the price process S_t, while v_t denotes the instantaneous variance process. The parameter μ represents the drift of the price process.

The variance process follows a Cox-Ingersoll-Ross (CIR) process (Cox et al., 1985) with mean reversion κ, long-run variance θ, and volatility σ. The conditional transition probability of v_t given v_0 is proportional to a noncentral chi-squared distributed random variable:

$$\begin{aligned} v_t \mid v_0 &\sim c \cdot \chi^{2'}_{\nu}(\Lambda) \\ c &= \sigma^2 \left(1 - \exp(-\kappa t)\right)(4\kappa)^{-1} \\ \nu &= \frac{4\kappa\theta}{\sigma^2} \\ \Lambda &= \frac{v_0}{c}\exp(-\kappa t) \end{aligned} \tag{3.1.5}$$

where $\chi^{2'}$ denotes a noncentral chi-squared distribution with ν degrees of freedom and noncentrality parameter Λ (Okhrin et al., 2022).

3.2 Characteristic Function and Density of the Heston Model

If a random variable X has a density function $f(x)$, its characteristic function $\phi(t)$ is given by

$$\phi(t) = \mathbb{E}(\exp(\mathrm{i}tX)) = \int_{-\infty}^{\infty} e^{\mathrm{i}tx} f(x)\mathrm{d}x$$

The characteristic function always exists, even if the probability density function does not. Once the characteristic function is known, the density function can be recovered via the inverse Fourier transform:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\mathrm{i}tx} \phi(t)\mathrm{d}t$$

Gatheral (2011) derives the characteristic function of the Heston model as

$$\phi(t) = \exp(A + B + C)$$

where

$$\begin{aligned}
A &= \mu \cdot \tau \cdot t \cdot \mathrm{i} \\
d &= \sqrt{(\rho\sigma\mathrm{i}t - \kappa)^2 - \sigma^2(-\mathrm{i}t - t^2)} \\
g &= \frac{\kappa - \rho\sigma\mathrm{i}t - d}{\kappa - \rho\sigma\mathrm{i}t + d} \\
B &= \frac{\theta\kappa}{\sigma^2}\left(\tau(\kappa - \rho\sigma\mathrm{i}t - d) - 2\log\left[\frac{1 - g\exp(-d\tau)}{1 - g}\right]\right) \\
\gamma &= \frac{2\kappa\theta}{\sigma^2} \\
C &= \log\left(\left[\frac{2\kappa}{\sigma^2}\right]^{\gamma} \cdot \left\{\frac{2\kappa}{\sigma^2} - \frac{\kappa - \rho\sigma\mathrm{i}t - d}{\sigma^2} \cdot \frac{1 - \exp(-d\tau)}{1 - g\exp(-d\tau)}\right\}^{-\gamma}\right)
\end{aligned} \tag{3.2.1}$$

where $\tau = T - t$ represents the time horizon.

A straightforward inverse Fourier transform is not suitable due to numerical instabilities at the boundaries (see Figure 3.1). The simple method centers $\phi(t)$ and normalizes $f(x)$ using the step size Δt to ensure correct amplitudes. An alternative approach employs a boundary correction, which improves the reconstruction of the density. Additionally, results are smoothed using cubic spline interpolation.

Equation (3.2.1) has the drawback that it can lead to overflow errors, as the term $\left(\frac{2\kappa}{\sigma^2}\right)^{\gamma}$ may become excessively large, making the logarithm intractable. To mitigate this issue, the equation is reformulated as

$$C_{unc} = \gamma\log\left(\frac{2\kappa}{\sigma^2}\right) - \gamma\log\left(\frac{2\kappa}{\sigma^2} - \frac{\kappa - \rho\sigma t\mathrm{i} - d}{\sigma^2}\frac{1 - \exp(-d\tau)}{1 - g\exp(-d\tau)}\right) \tag{3.2.2}$$

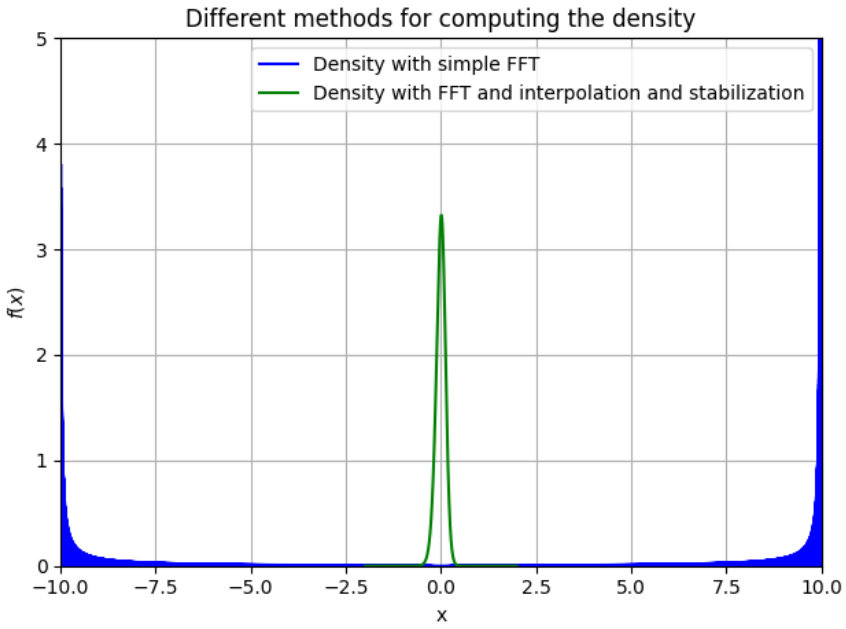


Figure 3.1: Comparison of different methods for the inverse Fourier transformation of the characteristic function of the Heston model ($\mu = 0$, $\kappa = 3$, $\theta = 0.19$, $\sigma = 0.4$, $\rho = -0.7$, $\tau = \frac{1}{12}$). Grid points: $N = 2^{15}$

3.3 Simulating the Heston Model

The Heston model is a continuous-time model, and for simulation purposes, time must be discretized. Small timesteps require significant computational power, while large timesteps may lead to zero or even negative volatility if the so-called Feller condition is not satisfied (Albrecher et al., 2007). The Feller condition holds if $2\kappa\theta > \sigma^2$, but Eraker et al. (2003) show that for S&P 500 index options, this condition is violated. Many other studies confirm this result for various markets and asset classes (e.g., Chang et al., 2021; Hu and Liu, 2022).

The most intuitive approach to time discretization is the Euler–Maruyama scheme:

$$X_{t+\Delta} = X_t - \frac{1}{2} v_t \Delta + \sqrt{v_t}\sqrt{\Delta} Z_{X,t}$$
$$v_{t+\Delta} = v_t + \kappa(\theta - v_t)\Delta + \sigma\sqrt{v_t}\sqrt{\Delta} Z_{v,t}$$

where $Z \sim \mathcal{N}(0,1)$, $\Delta = T/n$, with T representing the total time and n the number of time steps. The correlation between $Z_{X,t}$ and $Z_{v,t}$ can be simulated as follows (L. B. G. Andersen, 2007):

$$Z_{v,t} = \Phi^{-1}(U_1)$$
$$Z_{X,t} = \rho Z_{v,t} + \sqrt{1-\rho^2}\Phi^{-1}(U_2)$$

where U_1 and U_2 are independent random variables uniformly distributed on $[0, 1]$, and Φ^{-1} is the inverse cumulative distribution function of the standard normal distribution.

This discretization highlights the issue of potential negative volatility (Okhrin et al., 2022):

$$\begin{aligned}\mathbb{P}(v_{t+\Delta} < 0 \mid v_t > 0) &= \mathbb{P}\left(Z_{v,t} < \frac{-v_t - \kappa(\theta - v_t)\Delta}{\sigma\sqrt{v_t}\sqrt{\Delta}}\right) \\ &= \Phi\left(Z_{v,t} < \frac{-v_t - \kappa(\theta - v_t)\Delta}{\sigma\sqrt{v_t}\sqrt{\Delta}}\right) \\ &> 0\end{aligned}$$

where Φ denotes the cumulative distribution function of the standard normal distribution. Various methods exist to address this issue, such as the absorption method (replacing v_t with $v_t^+ = \max(0, v_t)$) or the reflection method (replacing v_t with $|v_t|$). However, these modifications alter the underlying process, causing the moments of volatility to deviate from their theoretical values (Okhrin et al., 2022; Tsoskounoglou, 2024).

In the study by Okhrin et al. (2022), Andersen's Quadratic Exponential (QE) scheme is found to perform best in terms of both speed and accuracy. Andersen (2007) approximates the noncentral chi-square distribution in (3.1.5) using a mixture of distributions: a Dirac distribution and a noncentral Gaussian distribution. For sufficiently large values of v_t,

$$v_{t+\Delta} = a(b + Z_v)^2 \tag{3.3.1}$$

where $Z_v \sim \mathcal{N}(0, 1)$. For smaller values of v_t,

$$v_{t+\Delta} = \Psi^{-1}(U_v, p, \beta) \tag{3.3.2}$$

$$\Psi^{-1}(u, p, \beta) = \begin{cases} 0 & 0 \le u \le p \\ \beta^{-1} \ln\left(\frac{1-p}{1-u}\right) & p < u \le 1 \end{cases}$$

The parameters a, b, p, and β are estimated using moment matching and U_v is drawn from a uniform distribution. The transition between the two approximation schemes is determined by a threshold ψ_c: if the ratio $\psi = s^2/m^2$ (where m and s^2 are the mean and variance of $v_{t+\Delta}$, respectively) exceeds ψ_c, then (3.3.2) is used; otherwise, (3.3.1) is applied.

For the price process, Andersen (2007) proposes the following discretization scheme:

$$\begin{aligned}
\ln(S_{t+\Delta}) &= \ln(S_t) + K_0 + K_1 v_t + K_2 v_{t+\Delta} + \sqrt{K_3 v_t + K_4 v_{t+\Delta}} \cdot Z \\
K_0 &= -\frac{\rho\kappa\theta}{\sigma}\Delta \\
K_1 &= \xi_1\Delta\left(\frac{\kappa\rho}{\sigma} - \frac{1}{2}\right) - \frac{\rho}{\sigma} \\
K_2 &= \xi_2\Delta\left(\frac{\kappa\rho}{\sigma} - \frac{1}{2}\right) + \frac{\rho}{\sigma} \\
K_3 &= \xi_1\Delta(1-\rho^2) \\
K_4 &= \xi_2\Delta(1-\rho^2)
\end{aligned}$$

where Z is standard normally distributed, and ξ_1 and ξ_2 are constants. The paper suggests $\xi_1 = \xi_2 = 0.5$.

3.4 The First 4 Moments of the Heston Model

We now focus on the moments of the Heston Model but care must be taken with notation, as symbols such as μ and σ usually denoting the mean and variance are now used to represent the drift and volatility of the Heston model (see Equations (3.1.1) and (3.1.3)). The noncentral moments are denoted by μ_1 through μ_4, while the central and standardized moments are denoted by ζ_1 through ζ_4.

Fortunately, moments for returns and other related measures have been derived. Okhrin et al. (2022) provide the unconditional noncentral moments for the log-return $r_t = \log(S_t) - \log(S_0)$:

$$\begin{aligned}
\mu_1 &= \left(\mu - \frac{\theta}{2}\right)t \\
\mu_2 &= \frac{1}{4\kappa^3}\Bigg\{\exp(-\kappa t)\Bigg[\exp(\kappa t)\Big\{\kappa^3 t\Big[t(\theta-2\mu)^2 + 4\theta\Big] - 4\kappa^2\rho\sigma t\theta \\
&\qquad + \kappa\sigma\theta(4\rho + \sigma t) - \sigma^2\theta\Big\} + \sigma\theta(\sigma - 4\kappa\rho)\Bigg]\Bigg\}
\end{aligned}$$

The expressions for μ_3 and μ_4 are too lengthy to be included here but can be found in Okhrin et al. (2022). The corresponding unconditional central and standardized

moments are then given by:

$$
\begin{aligned}
\zeta_1 &= \mu_1 \\
\zeta_2 &= \mathbb{E}\left[(r_t - \mu_1)^2\right] \\
&= \frac{\theta}{4\kappa^3}\Bigg[-4\kappa^2\rho\sigma t + 4\kappa^3 t + \sigma\exp(-\kappa t)(\sigma - 4\kappa\rho) \\
&\qquad + 4\kappa\sigma\rho + \kappa\sigma^2 t - \sigma^2\Bigg] \\
\zeta_3 &= \mathbb{E}\left[\left(\frac{r_t - \mu_1}{\zeta_1^{1/2}}\right)^3\right] \\
&= \frac{3\kappa\sigma\theta\exp(\kappa t/2)(\sigma - 2\kappa\rho)}{\zeta_2^{3/2}}\Bigg[4\kappa^2\Big\{\exp(\kappa t)(\rho\sigma t + 1) + \rho\sigma t - 1\Big\} \\
&\qquad - 4\kappa^3 t\exp(\kappa t) \\
&\qquad - \kappa\sigma\Big\{\exp(\kappa t)(8\rho + \sigma t) - 8\rho + \sigma t\Big\} \\
&\qquad + 2\sigma^2\Big(\exp(\kappa t) - 1\Big)\Bigg] \\
\zeta_4 &= \mathbb{E}\left[\left(\frac{r_t - \mu_1}{\zeta_1^{1/2}}\right)^4\right]
\end{aligned}
$$

The expression for ζ_4 is omitted here but can be found in Okhrin et al. (2022).

Zhao et al. (2013) and Zhang et al. (2017) analyze moments for the continuously compounded return $R_t^T = \ln\left(\frac{S_T}{S_t}\right)$ and derive the conditional central variance:

$$
\begin{aligned}
\mathbb{E}_t\Big(R_t^T - \mathbb{E}_t(R_t^T)\Big)^2 = \frac{1}{4}\mathrm{Var}_t\left(\int_t^T v_s\,\mathrm{d}s\right) + SW_{t,T} \\
- \mathbb{E}_t\left(\int_t^T \sqrt{v_s}\,\mathrm{d}B_s\left[\int_t^T v_s\,\mathrm{d}s - \mathbb{E}_t\left(\int_t^T v_s\,\mathrm{d}s\right)\right]\right)
\end{aligned}
$$

where v_t represents the variance at time t, SW is the variance swap rate (the expectation of realized variance), and B_t is a Brownian motion. Zhang et al. (2017) further

elaborate:

$$\begin{aligned}\mathbb{E}_t\Big(R_t^T - \mathbb{E}_t(R_t^T)\Big)^2 = &\int_t^T \mathbb{E}_t(v_s)\,\mathrm{d}s \\ &- \rho\sigma \int_t^T \frac{1-\exp(-\kappa(T-s))}{\kappa}\mathbb{E}_t(v_s)\,\mathrm{d}s \\ &+ \frac{1}{4}\left(\sigma^2 \int_t^T \frac{\Big(1-\exp(-\kappa(T-s))\Big)^2}{\kappa^2}\mathbb{E}_t(v_s)\,\mathrm{d}s\right)\end{aligned}$$

Using the expected instantaneous variance $\mathbb{E}_t(v_s) = \theta + (v_t - \theta)\exp(-\kappa(s-t))$, Mathematica yields the following results:

$$\int_t^T \mathbb{E}_t(v_s)\,\mathrm{d}s = \frac{v_t - \theta + \exp\big(\kappa(t-T)\big)(-v_t+\theta) - t\theta\kappa + T\theta\kappa}{\kappa}$$

$$\begin{aligned}\rho\sigma \int_t^T \frac{1-\exp\big(-\kappa(T-s)\big)}{\kappa}\mathbb{E}_t(v_s)\,\mathrm{d}s = \frac{\exp(-T\kappa)\,\rho\sigma}{\kappa^2}\Bigg[&\exp(t\kappa)\Big(-v_t + 2\theta + (t-T) \\ &\times \Big[(v_t-\theta)\,\kappa\Big]\Big) \\ &+ \exp(T\kappa)\Big(v_t + \theta(-2 - t\kappa + T\kappa)\Big)\Bigg]\end{aligned}$$

$$\begin{aligned}\sigma^2 \int_t^T \frac{\big(1-\exp\big(-\kappa(T-s)\big)\big)^2}{\kappa^2}\mathbb{E}_t(v_s)\,\mathrm{d}s = \frac{\exp(-2T\kappa)\,\sigma^2}{2\kappa^3}\Bigg[&\exp(2t\kappa)(-2v_t+\theta) \\ &+ 4\exp\big((t+T)\kappa\big)\Big(\theta + (t-T)(v_t-\theta)\kappa\Big) \\ &+ \exp(2T\kappa)\Big(2v_t + \theta(-5 - 2t\kappa + 2T\kappa)\Big)\Bigg]\end{aligned}$$

The third conditional central moment is given by

$$\mathbb{E}_t(R_t^T - \mathbb{E}_t(R_t^T))^3 = \mathbb{E}_t(X_T^3) - \frac{3}{2}\mathbb{E}_t(X_T^2 Y_T) + \frac{3}{4}\mathbb{E}_t(X_T Y_T^2) - \frac{1}{8}\mathbb{E}_t(Y_T^3)$$

where X_T and Y_T are defined as

$$\begin{aligned}X_T &= \int_t^T \sqrt{v_s}\,\mathrm{d}B_s^S \\ Y_T &= \int_t^T (v_s - \mathbb{E}_t(v_s))\,\mathrm{d}s = \sigma\int_t^T \frac{1-\exp(-\kappa(T-s))}{\kappa}\sqrt{v_s}\,\mathrm{d}B_s^v\end{aligned}$$

where B^S and B^v are the Brownian motions associated with price and volatility, respectively.

Dunn et al. (2014) derive the unconditional noncentral moments for the return $Q_{t+1} =$

$\frac{S_{t+1}}{S_t}$:

$$
\begin{aligned}
\mathbb{E}(Q_{t+1}) &= \mu_1 = 1 + \mu \\
\mathbb{E}(Q_{t+1}^2) &= \mu_2 = (\mu + 1)^2 + \theta \\
\mathbb{E}(Q_{t+1}^3) &= \mu_3 = (\mu + 1)^3 + 3\theta + 3\mu\theta \\
\mathbb{E}(Q_{t+1}^4) &= \mu_4 \\
&= \frac{1}{\kappa(\kappa - 2)} \Bigg(\kappa^2\mu^4 + 4\kappa^2\mu^3 + 6\kappa^2\mu^2\theta - 2\kappa\mu^4 \\
&\quad + 6\kappa^2\mu^2 + 12\kappa^2\mu\theta + 3\kappa^2\theta^2 - 8\kappa\mu^3 \\
&\quad - 12\kappa\mu^2\theta + 4\kappa^2\mu + 6\kappa^2\theta - 12\kappa\mu^2 \\
&\quad - 24\kappa\mu\theta - 6\kappa\theta^2 - 3\sigma^2\theta + \kappa^2 \\
&\quad - 8\kappa\mu - 12\kappa\theta - 2\kappa \Bigg)
\end{aligned}
$$

Using Equations (2.1.1), (2.1.2), and (2.1.3), the central and standardized moments follow as

$$
\begin{aligned}
\zeta_1 &= 1 + \mu \\
\zeta_2 &= \theta \\
\zeta_3 &= 0 \\
\zeta_4 &= 3\frac{\kappa^2\theta - 2\kappa\theta - \sigma^2}{\kappa\theta(\kappa - 2)}
\end{aligned}
$$

4 Expansion Methods

Expansion methods are series approximations of a probability density function. In general, these approximations are not true densities, as they can take negative values for certain parameter choices. However, for some parameter sets, they do define valid probability densities. In the following sections, we will explore the conditions under which expansion methods yield proper densities and how to transform a parameter set that does not correspond to a valid density into one that does.

4.1 Gram-Charlier Expansion

The Gram-Charlier expansion was introduced by Gram (1883) and Charlier (1914). There are two types of expansions: Gram-Charlier A and Gram-Charlier B, which are defined as

$$f_{GC,A} \approx f(x) + \sum_{k=3}^{n} a_k f^{(k)}(x)$$

$$f_{GC,B} \approx \psi(x) \sum_{m=0}^{n} b_m g_m(x)$$

Although these expansions can be applied to any density function f and ψ, for Gram-Charlier Type A, the density f is typically the standard normal distribution

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

and for Gram-Charlier Type B, ψ corresponds to the probability mass function of the Poisson distribution (Mitropol'skii, 2020):

$$\psi(x) = \frac{\lambda^x}{x!} \exp(-\lambda)$$

The term $f^{(k)}$ represents the k-th derivative of the density function f. There exist polynomials H_k that satisfy

$$f^{(k)}(x) = (-1)^k f(x) H_k(x)$$

These polynomials, known as Hermite polynomials, were studied by Laplace (1812, 1811), Chebyshev (1860), and Hermite (1864). They have the following properties (Abramowitz and Stegun, 1968, p. 771ff):

$$\begin{aligned} H_{n+1} &= x \cdot H_n(x) - H_n'(x) \\ H_n'(x) &= n \cdot H_{n-1}(x) \end{aligned} \tag{4.1.1}$$

Using these recurrence relations, the first few Hermite polynomials can be computed as

$$\begin{aligned} H_{n+1}(x) &= x \cdot H_n(x) - nH_{n-1}(x) \\ H_0(x) &= 1 \\ H_1(x) &= x \\ H_2(x) &= x^2 - 1 \\ H_3(x) &= x^3 - 3x \\ H_4(x) &= x^4 - 6x^2 + 3 \\ H_5(x) &= x^5 - 10x^3 + 15x \\ H_6(x) &= x^6 - 15x^4 + 45x^2 - 15 \end{aligned}$$

The coefficients a_k in the expansion can be expressed in terms of the moments r_k of the density f. This leads to the first terms of the Gram-Charlier A expansion:

$$\begin{aligned} f(x)_{GC,A} \approx \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \\ \times \left[1 + \frac{\kappa_3}{6\sigma^3} H_3\left(\frac{x-\mu}{\sigma}\right) + \frac{\kappa_4}{24\sigma^4} H_4\left(\frac{x-\mu}{\sigma}\right)\right] \end{aligned} \tag{4.1.2}$$

where μ, σ^2, κ_3, and κ_4 represent the first four cumulants of the target distribution. Based on Equations (2.2.1) and (2.2.2), μ and σ^2 correspond to the first two cumulants κ_1 and κ_2.

The first terms of the Gram-Charlier B expansion are given by

$$\begin{aligned} f(x)_{GC,B} \approx \frac{\lambda^x}{x!} \exp(-\lambda) \\ \times \left(1 + \frac{\mu_2 - \lambda}{\lambda^2}\left[\frac{x^{[2]}}{2} - \lambda x^{[1]} + \frac{\lambda^2}{2}\right]\right. \\ \left. + \frac{\mu_3 - 3\mu_2 + 2\lambda}{\lambda^3}\left[\frac{x^{[3]}}{6} - \frac{\lambda}{2}x^{[2]} + \frac{\lambda^2}{2}x^{[1]} - \frac{\lambda^3}{6}\right]\right) \end{aligned}$$

where μ_i are the central moments of the target distribution, and $x^{[i]} = x(x-1)\dots(x-i+1)$ (Mitropol'skii, 2020). However, since Gram-Charlier Type B only allows discrete values for x, it cannot be applied to continuous distributions. Therefore, we focus exclusively on Gram-Charlier Type A in this work.

4.2 Gram-Charlier Expansion with Positivity Constraints

The Gram-Charlier expansion is not an asymptotic expansion, as it does not allow for a well-defined approximation error. The Edgeworth expansion, however, is an asymptotic expansion (Cramér, 1999, Section 17.6) and is therefore preferred. An asymptotic expansion consists of a series of functions f_n that, after a finite number of terms, approximate a function at a specific point ξ (often infinite) as x approaches ξ:

$$f_{n+1}(x) = o(f_n(x)) \quad x \to \xi$$

Furthermore, the Gram-Charlier expansion can take negative values, which is not permissible for a probability density function. Jondeau & Rockinger (2001) analyzed the conditions under which the expansion remains a valid density. By using Equations (2.2.3) and (2.2.4), Equation (4.1.2) can be rewritten in terms of the skewness γ_1 and excess kurtosis γ_2^*, defining $z = \frac{x-\mu}{\sigma}$:

$$f(x)_{GC,A} \approx \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)\left[1 + \frac{\gamma_1}{6}H_3(z) + \frac{\gamma_2^*}{24}H_4(z)\right] \tag{4.2.1}$$

To determine when the Gram-Charlier expansion remains a valid density, the following condition must hold:

$$\begin{aligned}
1 + \frac{\gamma_1}{6}He_3(z) + \frac{\gamma_2^*}{24}He_4(z) &= 0 \\
\frac{\gamma_1}{6}He_3(z) &= -1 - \frac{\gamma_2^*}{24}He_4(z) \\
\gamma_1 \cdot He_3(z) &= -6 - \frac{\gamma_2^*}{4}He_4(z) \\
\gamma_1 &= -\frac{6}{He_3(z)} - \frac{He_4(z)}{4 \cdot He_3(z)}\gamma_2^* \\
\gamma_1 &= \frac{z^4 - 6z^2 + 3}{12z - 4z^3} \cdot \gamma_2^* + \frac{24}{12z - 4z^3}
\end{aligned}$$

This leads to a boundary condition for the skewness and excess kurtosis, illustrated in Figures 4.1 and 4.2.

Using a bisection algorithm and a logistic mapping, Jondeau & Rockinger (2001) construct a piecewise linear boundary, ensuring that any unconstrained pair $(\tilde{\gamma_1}, \tilde{\gamma_2^*}) \in \mathbb{R}^2$ is mapped to a constrained pair within the positivity region $\mathcal{D}$. Finding a closed-

form expression for the boundary is computationally difficult; even with 24 hours on a high-performance computer using Python's SymPy library, no explicit solution was found.

A comparison between constrained and unconstrained parameters for four distributions—standard normal, lognormal, t-distribution, and non-central t-distribution—is presented in Table 4.1 and Figures 4.3 and 4.4. The results highlight the necessity of positivity constraints, which, however, distort the expansion, particularly for cases where no constraints were originally required (e.g., the standard normal distribution). The logistic mapping shifts the excess kurtosis from 0 to 2, mapping the unconstrained pair $(\tilde{\gamma_1}, \tilde{\gamma_2^*}) = (0, 0)$ to the constrained pair $(\gamma_1, \gamma_2^*) = (0, 2)$.

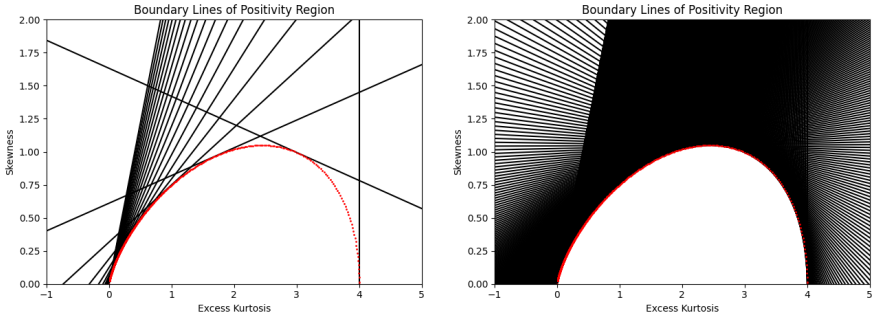


Figure 4.1: Boundary lines of the positivity region of the Gram-Charlier Expansion. The left image shows 20 lines, the right image shows 1000 lines. The red dots are the boundary points. The boundary is symmetric to the x-axis.

Distribution	μ	σ^2	κ_3	γ_1	κ_4	γ_2^*
$\mathcal{N}(\mu = 0, \sigma 1)$	0	1	0	0	0	0
$\mathcal{L}(\mu = 0, \sigma = 0.5)$	1.1331	0.3647	0.3855	1.7502	0.7845	5.8984
$\mathcal{T}(\nu = 5)$	0	1.6667	0	0	16.6667	6
$\mathcal{NCT}(\nu = 5, \mu = 0.5)$	0.5947	1.7297	1.5357	0.6751	21.5969	7.2189

Table 4.1: Distribution parameters and theoretical moments and cumulants. $\mathcal{N}$ stands for the Normal distribution, $\mathcal{L}$ for the Lognormal distribution, $\mathcal{T}$ for the Student's t-distribution, and $\mathcal{NCT}$ for the Non-Central t-distribution.

4.3 Edgeworth Expansion

Originally introduced by Edgeworth 1907, who suggested computing the approximation up to the sixth term to mitigate the issues associated with asymptotic expansions. A

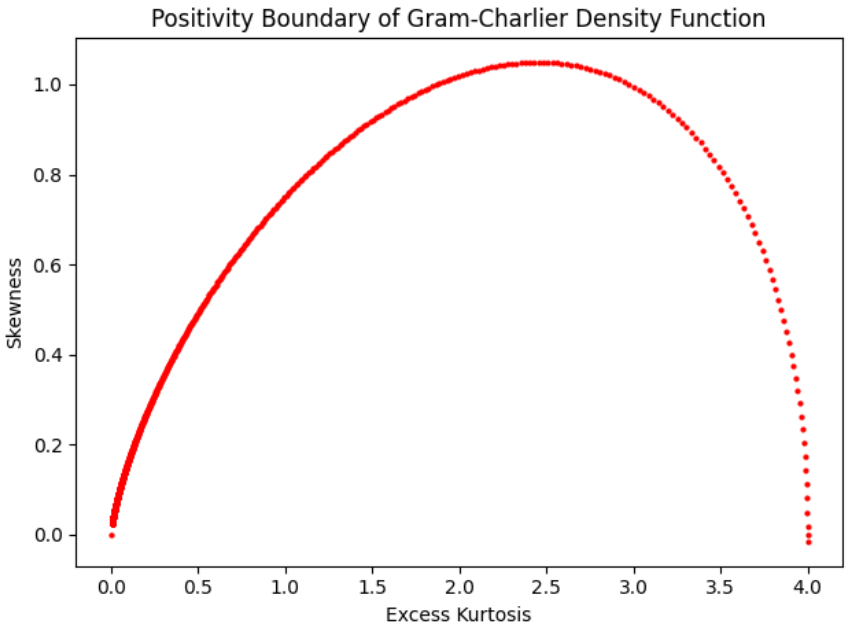


Figure 4.2: Approximation (1000 steps) of the positivity boundary of the Gram-Charlier Expansion. For simplicity, only the part above the x-axis is shown. The boundary is symmetric to the x-axis.

formal representation of the first terms can be found in Brenn & Anfinsen (2017).

$$f(x)_{EW} \approx \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \times \left[1 + \frac{\kappa_3}{6\sigma^3} H_3\left(\frac{x-\mu}{\sigma}\right) + \frac{\kappa_4}{24\sigma^4} H_4\left(\frac{x-\mu}{\sigma}\right) + \frac{\kappa_5}{120\sigma^5} H_5\left(\frac{x-\mu}{\sigma}\right) + \frac{\kappa_6 + 10\kappa_3^2}{720\sigma^6} H_6\left(\frac{x-\mu}{\sigma}\right)\right]$$

Since this work only considers the first four cumulants, the expression simplifies to:

$$f(x)_{EW} \approx \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \times \left[1 + \frac{\kappa_3}{6\sigma^3} H_3\left(\frac{x-\mu}{\sigma}\right) + \frac{\kappa_4}{24\sigma^4} H_4\left(\frac{x-\mu}{\sigma}\right) + \frac{\kappa_3^2}{72\sigma^6} H_6\left(\frac{x-\mu}{\sigma}\right)\right] \tag{4.3.1}$$

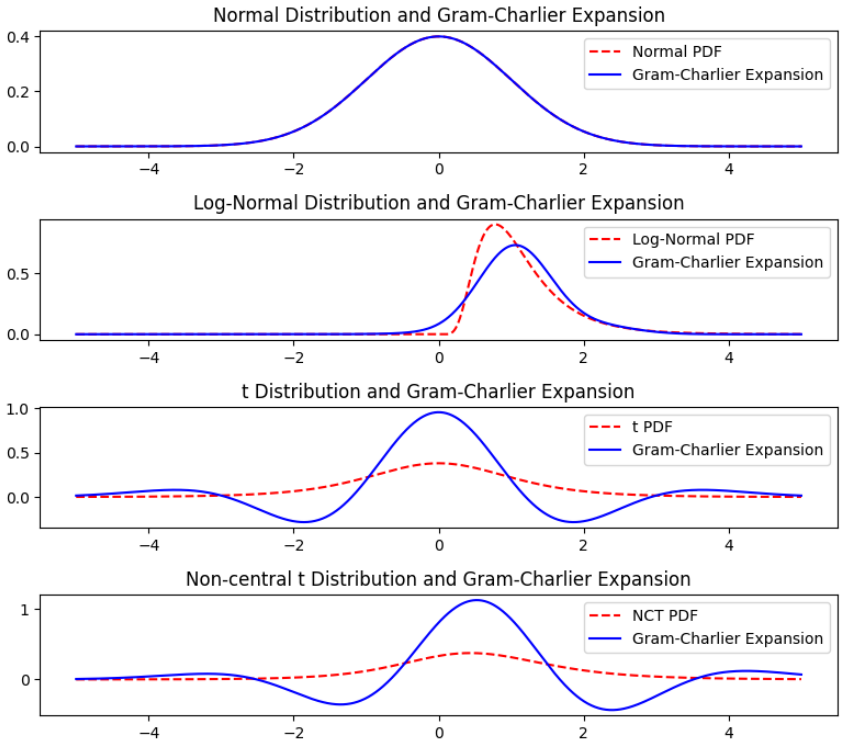


Figure 4.3: Gram-Charlier Expansion of different distributions

4.4 Edgeworth Expansion with Positivity Constraints

Following the approach of Jondeau & Rockinger (2001) for the Gram-Charlier expansion, we determine the positivity boundary for the Edgeworth expansion by rewriting Equation (4.3.1) using $z = \frac{x-\mu}{\sigma}$:

$$f(x)_{EW} \approx \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)\left[1 + \frac{\gamma_1}{6}H_3(z) + \frac{\gamma_2^*}{24}H_4(z) + \frac{\gamma_1^2}{72}He_6(z)\right] \tag{4.4.1}$$

To ensure the expansion remains a valid density, we solve for skewness γ_1 in the equation:

$$0 = 1 + \frac{\gamma_1}{6}He_3(z) + \frac{\gamma_2^*}{24}He_4(z) + \frac{\gamma_1^2}{72}He_6(z) \tag{4.4.2}$$

$$\gamma_1 = \pm\sqrt{-\frac{72}{He_6(z)} - 3\gamma_2^*\frac{He_4(z)}{He_6(z)} + 36\frac{He_3(z)^2}{He_6(z)^2} - 6\frac{He_3(z)}{He_6(z)}}$$

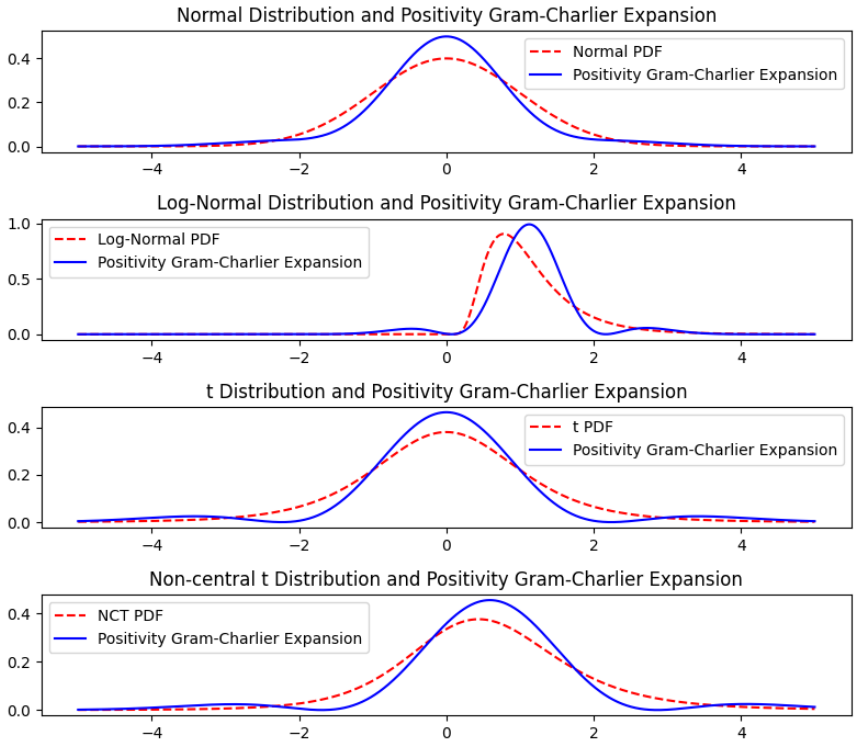


Figure 4.4: Gram-Charlier Expansion with positivity constraints of different distributions

This equation is valid as long as $H_6(z) \neq 0$, which occurs at six points:

$$z_{1/2} = \pm \sqrt{5 - \frac{5^{2/3}\left(1+i\sqrt{3}\right)}{\sqrt[3]{2\left(2+i\sqrt{6}\right)}} - \frac{\left(1-i\sqrt{3}\right)\sqrt[3]{5\left(2+i\sqrt{6}\right)}}{2^{2/3}}} = \pm 0.6167$$

$$z_{3/4} = \pm \sqrt{5 - \frac{5^{2/3}\left(1-i\sqrt{3}\right)}{\sqrt[3]{2\left(2+i\sqrt{6}\right)}} - \frac{\left(1+i\sqrt{3}\right)\sqrt[3]{5\left(2+i\sqrt{6}\right)}}{2^{2/3}}} = \pm 1.8892$$

$$z_{5/6} = \pm \sqrt{5 + \frac{10^{2/3}}{\sqrt[3]{2+i\sqrt{6}}} + \sqrt[3]{10\left(2+i\sqrt{6}\right)}} = \pm 3.3243$$

The positivity boundary for the Edgeworth expansion is determined by plotting the lines $\gamma_1(\gamma_2^*, z)$ for many values of z, skipping over the six singularities. Figure 4.5 shows that the positivity region for the Edgeworth expansion is smaller than for the Gram-Charlier expansion. For $\gamma_1 = 0$, the excess kurtosis is constrained between 0 and 4, which matches the boundary of the Gram-Charlier expansion since both expansions are identical when skewness is zero.

As with the Gram-Charlier expansion, the positivity boundary is given by the envelope

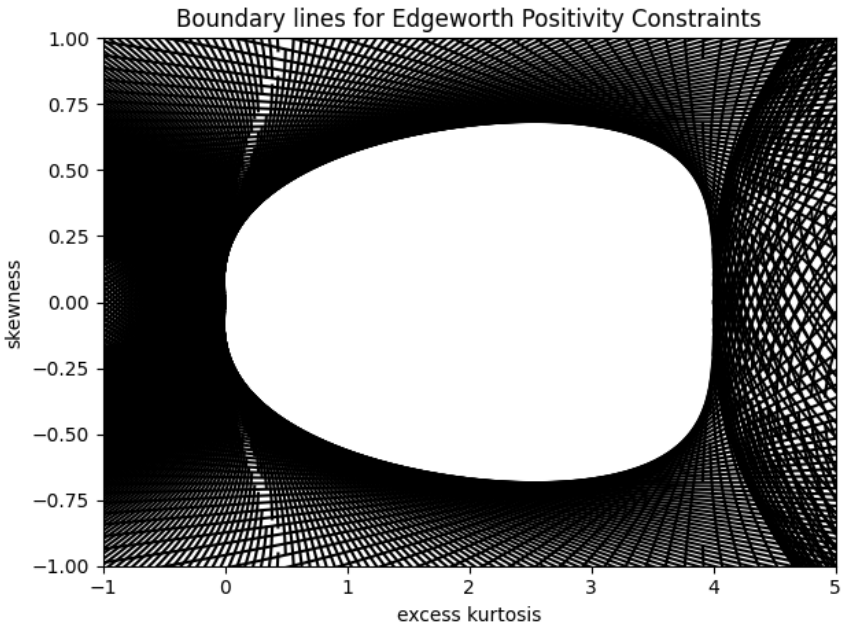


Figure 4.5: Boundary Lines for Edgeworth Expansion

of the lines $\gamma_1(\gamma_2^*, z)$. We obtain this by computing the intersections of pairs of parabolic equations. The equations are to long to be displayed here, you can find them in the corresponding GitHub repository[1]. These intersections are shown in Figures 4.6 to 4.9.

Since each $\gamma_1(\gamma_2^*, z)$ equation is a parabola, the positivity region can be computed by:

1. Ignoring the second intersection for each z-value (since the first intersection defines the boundary). The boundary is symmetric around the x-axis, so we only compute the upper half and mirror it afterward.

2. Filtering out non-relevant points: Figure 4.7 shows many extraneous intersection points. Based on previous results, we restrict the solutions to $\gamma_2^* \in (-0.1, 4.1)$ and $|\gamma_1| \in [0, 1)$. (see Figure 4.8)

3. Removing points from non-relevant boundary lines: The lower-boundary artifacts arise from z-values smaller than the third singularity; these are removed. The upper-boundary artifacts come from $|z|$ around 1.8 and 1.67, and are filtered out by removing all z in the ranges $(1.8 - 0.035, 1.8 + 0.035)$ and $(1.67 - 0.015, 1.67 + 0.015)$.

4. Computational accuracy is a concern since some points lie slightly outside the expected range (e.g., one point is at $(\gamma_2^*, \gamma_1) = (4.01, 0.222)$). Testing this point

[1] https://github.com/henrydatei/heston-moments-pdf

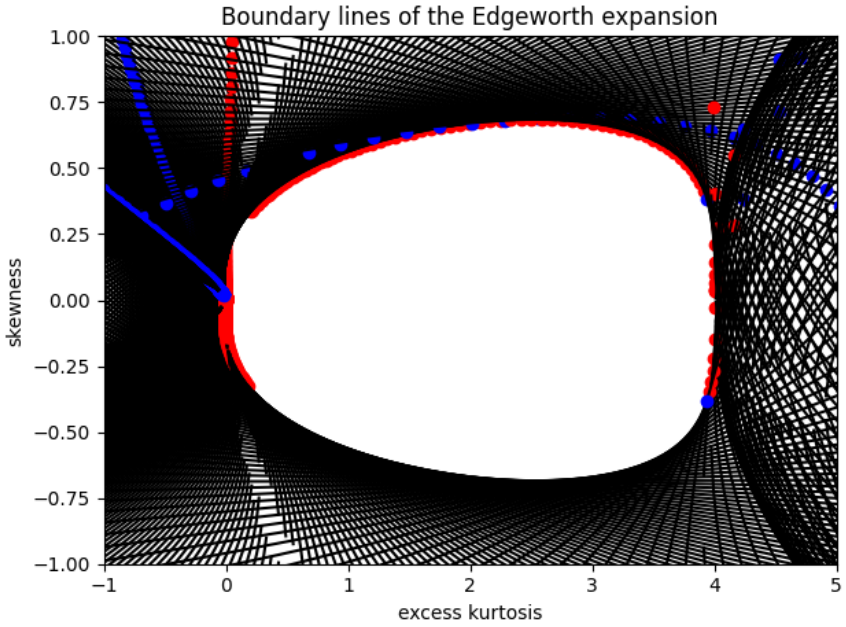


Figure 4.6: Intersections of Boundary Lines for Edgeworth Expansion, red is first intersection, blue is second intersection

in the Edgeworth expansion shows that it does not satisfy the positivity condition:

$$1 + \frac{0.222}{6} He_3(z) + \frac{4.01}{24} He_4(z) + \frac{0.222^2}{72} He_6(z) < 0$$

gives a solution: $-1.84611 < z < -1.75826$. This might be due to Python's float datatype which maps to IEEE-754 double precision with 64 bits where 52 bits are used for the fraction which equals about 16 decimal digits (Foundation, n.d.; Leonardo.Z, 2013) To be conservative, we remove any intersection with $\gamma_2^* \notin [0, 4]$.

5. Adding the points $(0, 0)$ and $(4, 0)$ completes the positivity region. A linear interpolation between the closest computed boundary points results in the final positivity boundary (see Figure 4.9).

As with the Gram-Charlier expansion, we compare the constrained and unconstrained parameters for four distributions: standard normal, lognormal, t-distribution, and noncentral t-distribution (see Table 4.1). The results, shown in Figures 4.10 and 4.11, lead to the same conclusion as in the Gram-Charlier case: Positivity constraints are necessary, but they also distort the expansion and even when constraints are not needed (e.g., for the standard normal distribution), applying the positivity constraints artificially increases excess kurtosis from 0 to 2.

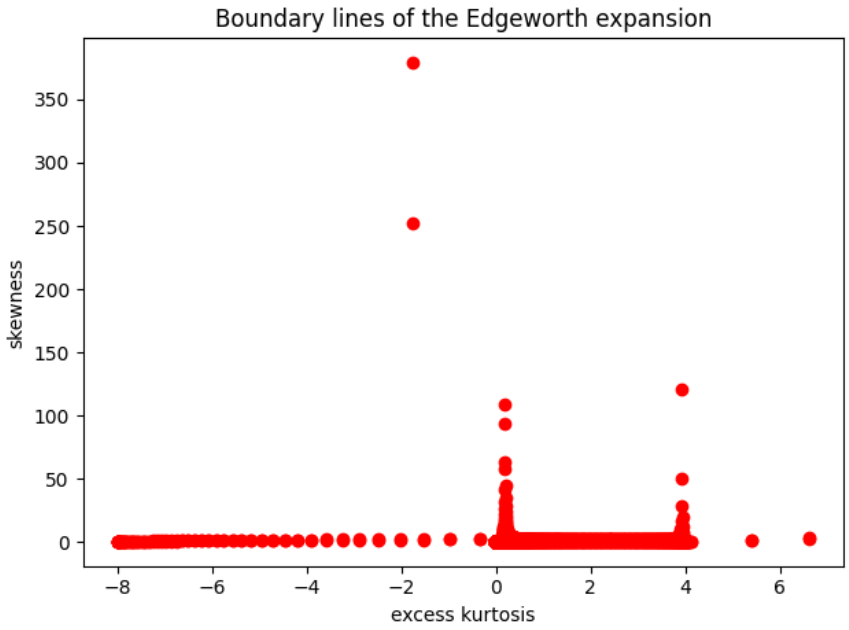


Figure 4.7: Intersections of Boundary Lines for Edgeworth Expansion (zoomed out), upper half

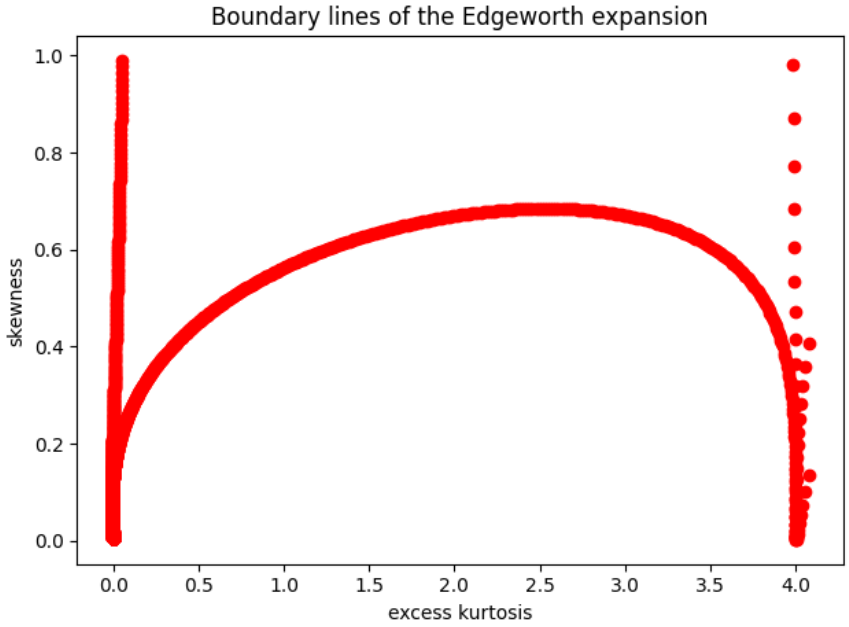


Figure 4.8: Intersections of Boundary Lines for Edgeworth Expansion (zoomed in), upper half

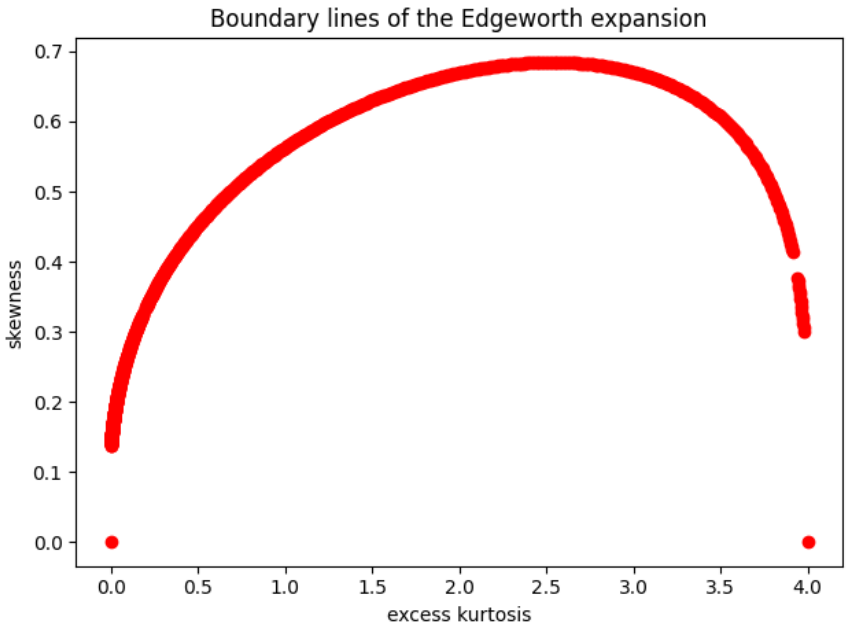


Figure 4.9: Final Boundary Points for Edgeworth Expansion, upper half

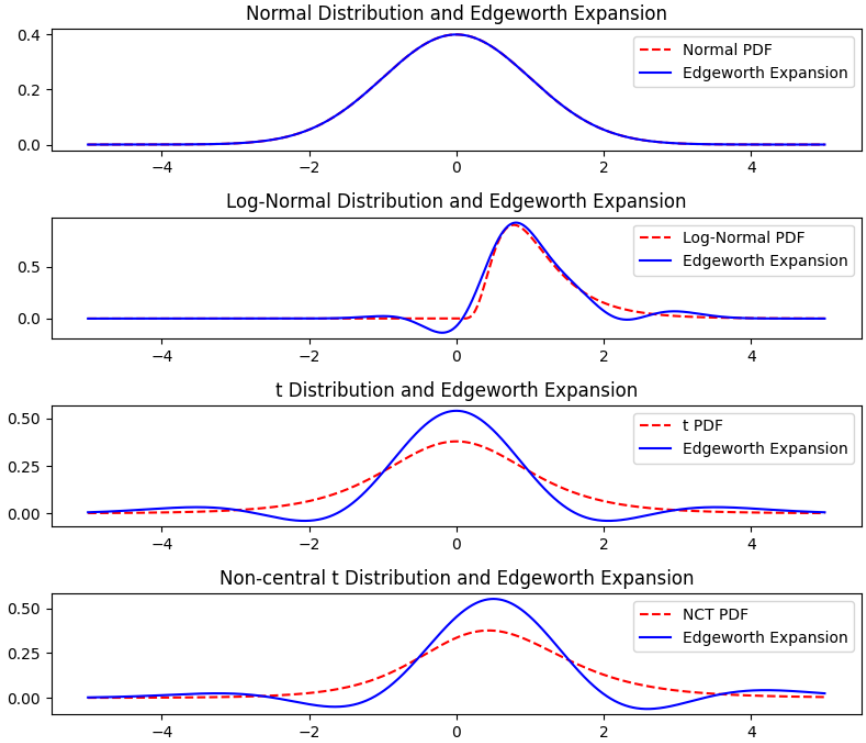


Figure 4.10: Edgeworth Expansion of different distributions

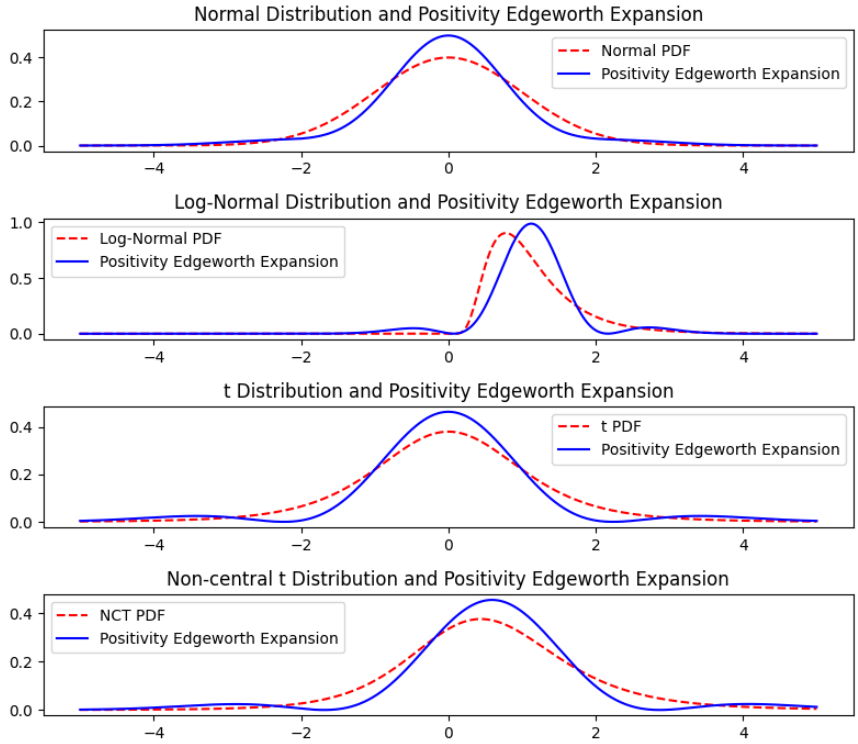


Figure 4.11: Edgeworth Expansion with positivity constraints of different distributions

5 Implementation of the simulation study

5.1 Carrying out the Simulation

All the methods introduced in the previous chapters were implemented in Python. To analyze their behavior in a realistic setting, we conducted a simulation of $n = 1$ paths of the Heston model using the Quadratic Exponential (QE) scheme by Andersen (2007). Since the results are averaged over time, simulating multiple paths was deemed unnecessary. The simulation was based on interday 5-minute price data, with 79 observations per trading day, 22 trading days per month, 12 months per year, and a total simulation period of 15 years ($T = 15$), leading to 312,840 prices per path. A burn-in period of three years was applied to eliminate biases stemming from the initial price $S_0 = 100$ and initial volatility v_0.

The simulation covered a broad parameter space inspired by the estimates of Eraker (2004). Specifically, v_0 ranged from 0.01 to 0.5 in steps of 0.05, κ from 0.01 to 1 in steps of 0.1, θ from 0.01 to 1 in steps of 0.05, and σ from 0.01 to 1 in steps of 0.05. The drift parameter μ took only two values, 0 and 0.05, while the correlation ρ varied from -0.9 to 0.9 in steps of 0.1. In total, the simulation encompassed 1,440,000 parameter combinations.

Preliminary tests revealed that the QE scheme occasionally produces numerical errors, particularly when the Feller condition is strongly violated. A closer inspection of these cases showed that the issue manifests as an excessive number of price values clustering at the same level. To systematically detect these errors, the frequency of the most common price value was recorded for the first simulated path of each scenario. In simulations without errors, this frequency remained in the single-digit range, whereas in faulty simulations, it surged into the hundreds or thousands, indicating numerical instability.

For each simulated path, we computed the first four realized moments, skewness, and kurtosis following Neuberger & Payne (2021), as well as the first four cumulants based on Fukasawa & Matsushita (2021). Since this process is computationally intensive, requiring approximately 100 seconds per simulation on a 1.8 GHz Dual-Core Intel Core i5, parallelization was employed to improve efficiency. To manage large-scale data output, each computing core stored its results in a separate CSV file, minimizing write conflicts. The collected data was then merged into a central database for further

analysis.

Additionally, for each simulation, we computed the theoretical mean, variance, skewness, and kurtosis of the returns based on the closed-form expressions derived in Okhrin et al. (2023). Since these quantities depend only on the model parameters, they could be computed efficiently using basic arithmetic operations, allowing their calculation to be performed directly within the database.

5.2 Calculating Results

For each simulation, the Gram-Charlier expansion, the Gram-Charlier expansion with positivity constraints, the Edgeworth expansion, the Edgeworth expansion with positivity constraints, the Cornish-Fisher expansion, and the Saddlepoint approximation are computed. Each expansion method is applied to both the first four realized moments and the first four cumulants.

To evaluate the accuracy of these expansions, the theoretical density is derived from the characteristic function and compared to the densities obtained from the expansions using the Kolmogorov-Smirnov test (Kolmogorov, 1993), the Cramér-von Mises test (Anderson, 1962; Cramér, 1928; Von Mises, 1928), and the Anderson-Darling test (Anderson and Darling, 1954). While the Kolmogorov-Smirnov test measures the overall fit between two distributions, the Cramér-von Mises test and the Anderson-Darling test place greater emphasis on the tails of the distribution. This is particularly relevant for financial markets, where heavy tails arise due to crises and market crashes. The Cramér-von Mises test is less sensitive to the tails than the Anderson-Darling test, providing a balanced assessment of goodness-of-fit.

Since these tests require samples drawn from the cumulative distribution function (CDF), a numerical CDF is constructed from the density by performing cumulative summation and normalizing all values so that the last element equals 1.

Additionally, to further analyze the tails of the distributions, the Hill estimator is used to estimate the tail index α (Hill, 1975). The tail index is a shape parameter that characterizes the power-law behavior of a distribution, where larger values of α indicate thinner tails (Danielsson et al., 2016; Fischler, 2017).

As previously noted in the discussion on the Heston model, the formula in Equation (3.2.1) can lead to overflow errors, as Python's floating-point arithmetic is limited to values in the range 10^{-308} to 10^{308}. To address this, the reformulated expression in Equation (3.2.2) is used for computations.

6 Results

6.1 Findings from the Simulations

As observed in preliminary tests, the QE scheme produces numerical errors, which manifest as repeated price values within the simulation. In a correctly executed simulation, the most frequently occurring price in the first simulated path appears only a few times. However, in cases where errors occur, this number increases to the hundreds or even thousands.

Figure 6.1 illustrates the distribution of this phenomenon. The peak at 1 indicates that most simulations were performed correctly. However, the presence of cases where the same price appears up to 250,000 times suggests significant numerical instabilities. A closer examination of these faulty simulations reveals that once a numerical error occurs for the first time, the remaining time series retains the same price value throughout. In extreme cases where the same price appears 250,000 times, the error must have occurred very early in the simulation.

To ensure the reliability of the analysis, simulations exhibiting numerical errors will be excluded, as they are not correctly executed. However, since it is possible for identical prices to occur naturally in correct simulations, a threshold must be determined that retains as many valid simulations as possible while filtering out erroneous ones.

Figure 6.2 illustrates the cumulative percentage of simulations that do not exceed a given threshold for the maximum number of identical price occurrences. If the threshold is set to 1 -— meaning only simulations where no price ever appears more than once are retained —- 68.40% of the simulations remain. Increasing the threshold reduces the number of discarded simulations, but at the cost of potentially including simulations with numerical errors. However, the impact is minimal, as the increase in retained simulations is negligible beyond a certain point.

For the final analysis, all simulations where any price appears more than once will be excluded to ensure data integrity.

To analyze the source of these numerical errors, we examine the proportion of correct simulations as a function of each individual Heston model parameter (see Figure 6.3). The results show that as σ increases, the error rate rises, whereas increasing θ and κ reduces the frequency of errors. In contrast, the parameters μ, ρ, and v_0 appear to

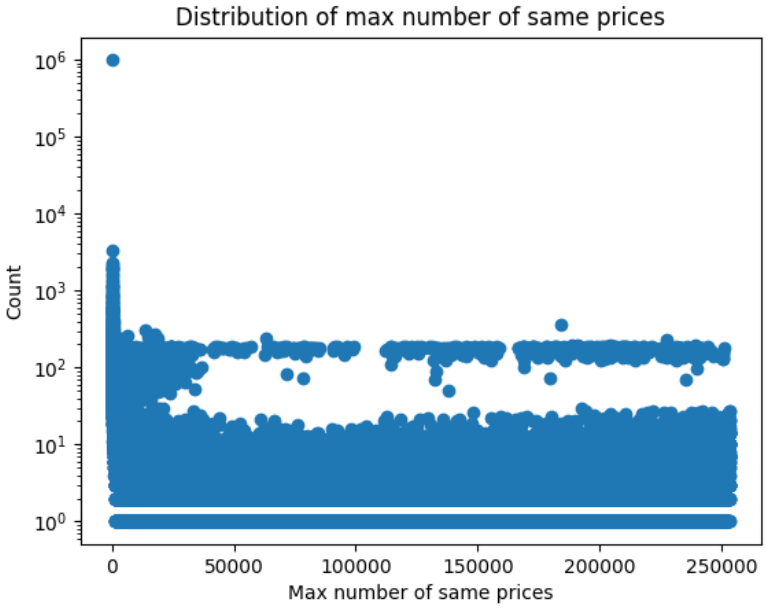


Figure 6.1: Distribution of the maximum number of the same prices in the first path of the simulation

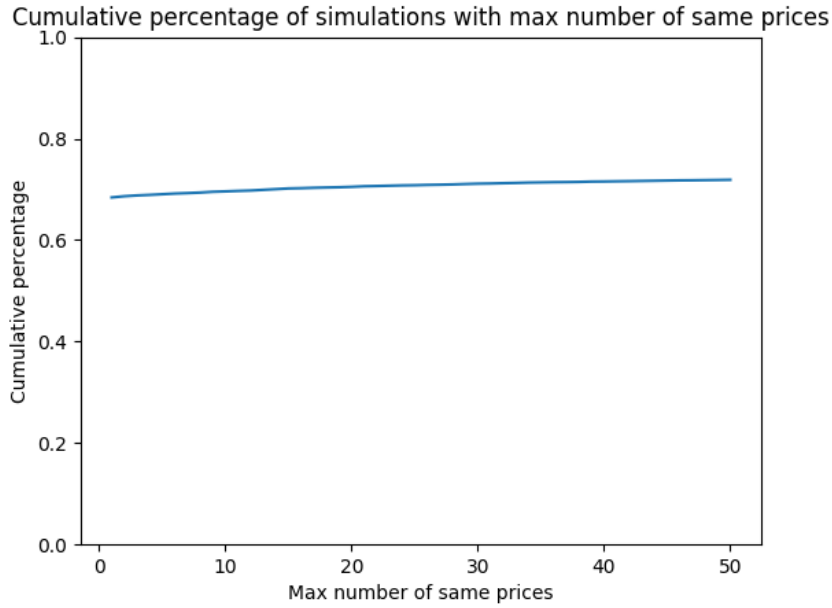


Figure 6.2: Cumulative percentage of simulations that do not exceed the maximum number of same prices

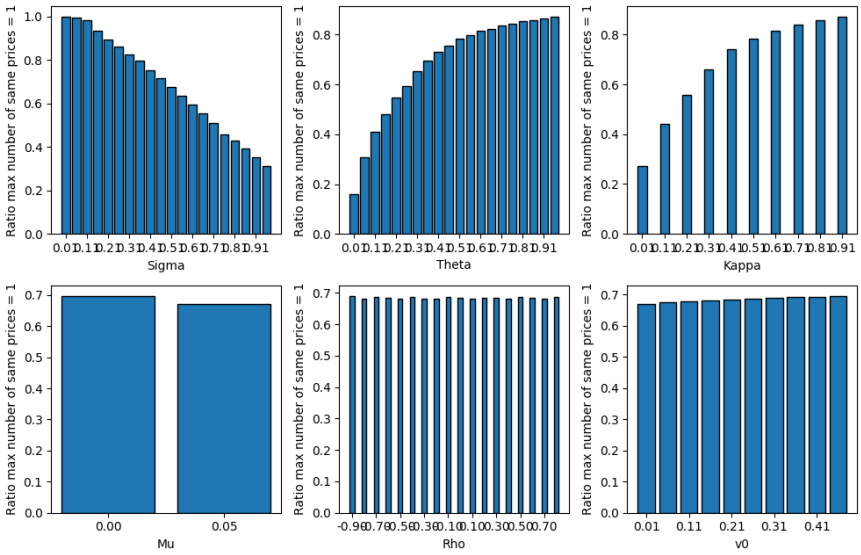


Figure 6.3: Ratio of simulations that do not exceed the maximum number of same prices in relation to the parameters of the Heston model

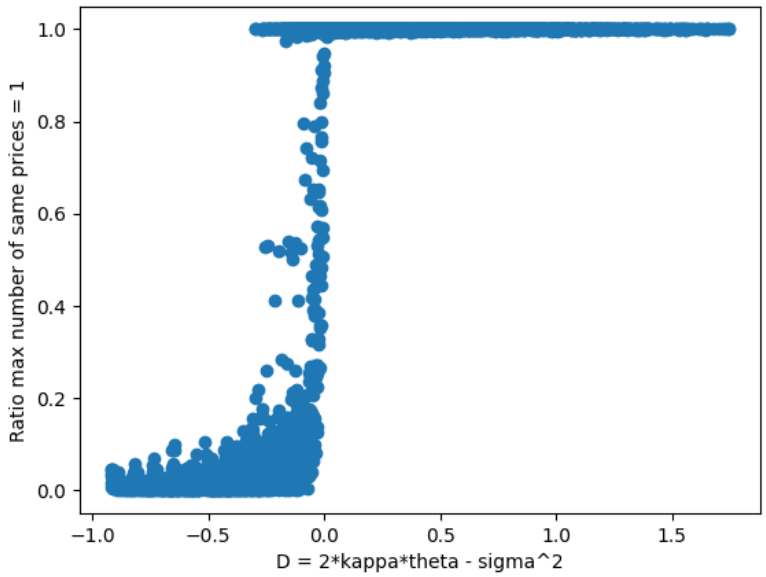
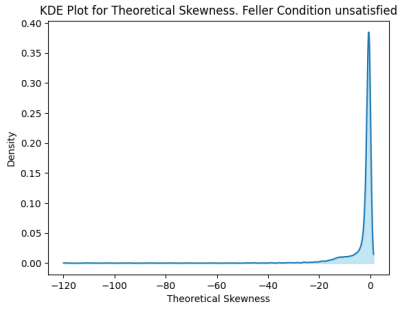
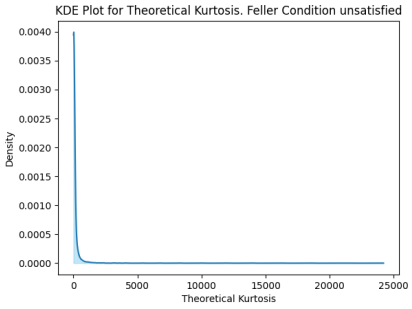


Figure 6.4: Ratio of simulations that do not exceed the maximum number of same prices in relation to the value D



(a) Theretical Skewness



(b) Theoretical Kurtosis

Figure 6.5: Kernel density estimation of the theoretical skewness and kurtosis for simulations where the Feller condition is not fulfilled

have no impact on the occurrence of errors. This suggests a strong relationship between numerical errors and the Feller condition.

To further investigate this relationship, we analyze the error frequency in relation to the value $D = 2\kappa\theta - \sigma^2$. If $D > 0$, the Feller condition is satisfied; the more negative D, the more strongly it is violated. Figure 6.4 confirms that once the Feller condition is no longer met, the error rate increases significantly. When the condition is satisfied, 99.90% of the simulations are correct, whereas when it is violated, only 25.48% are correct. Overall, 57.80% of the simulations satisfy the Feller condition.

Following this analysis, we evaluate which expansion methods best approximate the theoretical distribution under different conditions. First, we examine how well the Gram-Charlier expansion performs for given levels of skewness and kurtosis, as this is the simplest method. In the second part of the analysis, we compare different expansion methods in terms of their general ability to approximate the distribution, their accuracy in capturing the tails, and their sensitivity to the Feller condition.

For the final evaluation, we focus only on simulations with $\mu = 0$, as a large proportion of simulations with $\mu = 0.05$ contain errors. Additionally, we restrict our analysis to simulations where the Feller condition is met, since violations of the condition lead to extreme values for theoretical skewness and kurtosis (see Figure 6.5).

6.2 Investigating the Results for the Gram-Charlier-Expansion with respect to Skewness und Kurtosis

Figure 6.6 shows the distribution of p-values from the Kolmogorov-Smirnov test comparing the Gram-Charlier expansion based on cumulants to the theoretical density. All simulations where the same price appeared more than once have been removed.

A p-value above 5% indicates that there is no statistically significant difference between

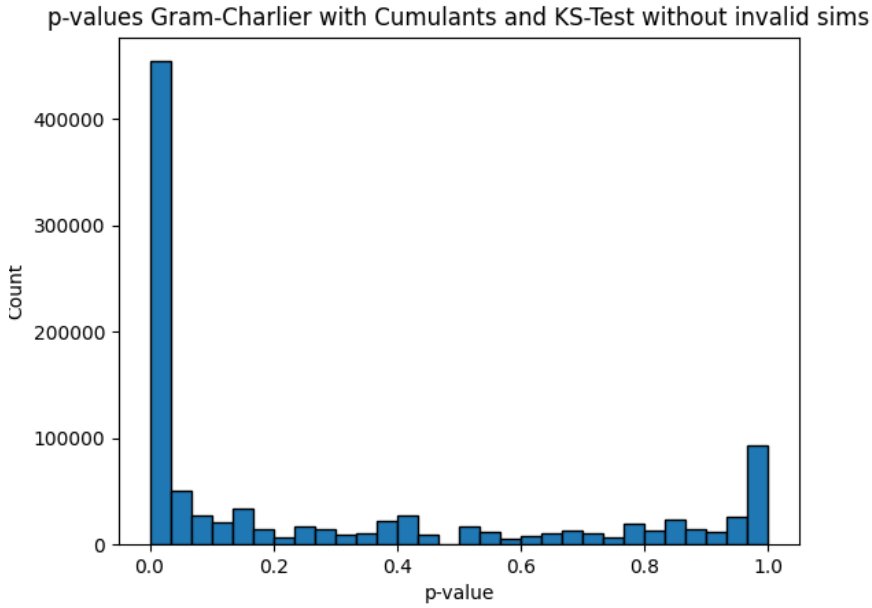


Figure 6.6: Distribution of p values of the Kolmogorov-Smirnov-Test for the Gram-Charlier Expansion with Cumulants vs the theoretical density without invalid simulations

the two distributions, meaning that the Gram-Charlier expansion based on cumulants provides an adequate approximation of the theoretical density. However, the results reveal that for some parameter combinations, the approximation does not hold; overall, 52.48% of the simulations fail to meet this threshold.

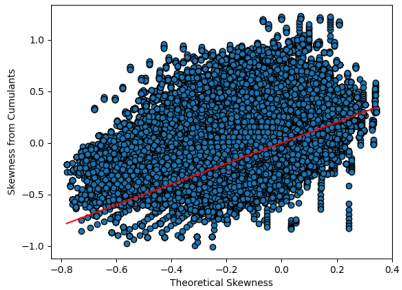
For results on other expansion methods, see Table 6.1.

	Gram-Charlier	**GC+**	**Edgeworth**	**EW+**	**Cornish-Fisher**	**Saddlepoint**
Cumulants	52.48%	47.63%	52.14%	47.64%	35.85%	45.36%
Moments	37.35%	21.24%	36.02%	23.27%	44.90%	40.92%

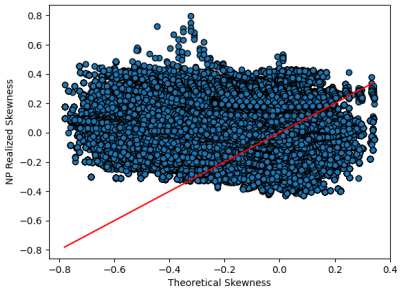
Table 6.1: Percentage of simulations where the p-value of the Kolmogorov-Smirnov-Test against the theoretical density is above 5%. Invalid simulations are excluded. GC+ and EW+ stand for the Gram-Charlier Expansion with positivity constraint and the Edgeworth Expansion with positivity constraint, respectively.

Furthermore, Table 6.1 shows that expansion methods perform better when based on cumulants rather than moments. Among the examined methods, the Cornish-Fisher expansion and the Saddlepoint approximation, which were briefly discussed in Chapter 1, perform the worst. Consequently, these methods, along with all moment-based expansion approaches, will not be considered in further analyses.

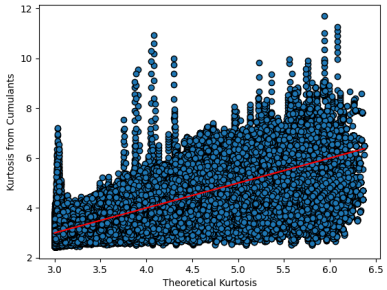
The superior performance of the cumulant-based approach can be attributed to its ability to better capture the theoretical skewness and kurtosis of the Heston model compared to the moment-based approach. This is clearly illustrated in Figure 6.7.



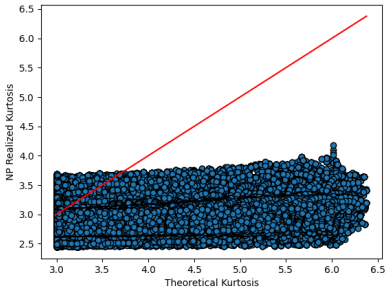
(a) Theretical Skewness vs Skewness from Cumulants



(b) Theoretical Skewness vs Skewness from Neuberger & Payne (2021)



(c) Theoretical Kurtosis vs Kurtosis from Cumulants



(d) Theoretical Kurtosis vs Kurtosis from Neuberger & Payne (2021)

Figure 6.7: Comparison of the theoretical skewness and kurtosis with the skewness and kurtosis from the Cumulants and the real skewness and kurtosis from Neuberger & Payne (2021) for simulations where the Feller condition is fulfilled. Invalid simulations are excluded.

The skewness and kurtosis values proposed by Neuberger & Payne (2021) are derived directly from raw moments, as described in their original paper.

We now examine whether there are correlations between the Heston model parameters and the p-values from the Kolmogorov-Smirnov test. Figure 6.8 presents a pair plot illustrating all parameter combinations (excluding μ) and the p-values of the Kolmogorov-Smirnov test for the Gram-Charlier expansion based on cumulants compared to the theoretical density. The more yellow the region, the higher the proportion of p-values above 5%, indicating a better approximation of the theoretical density by the Gram-Charlier expansion with cumulants.

The color gradients reveal which parameters appear to have an impact. Specifically, σ, κ, and θ seem to influence the quality of the approximation, while ρ appears to have little to no effect. Regardless of how ρ is chosen while holding other parameters fixed, the color distribution remains unchanged. The same applies to v_0, which is expected since v_0 represents the initial volatility and a burn-in period of three years (out of the 15-year simulation period) was applied to ensure that the unconditional moments and

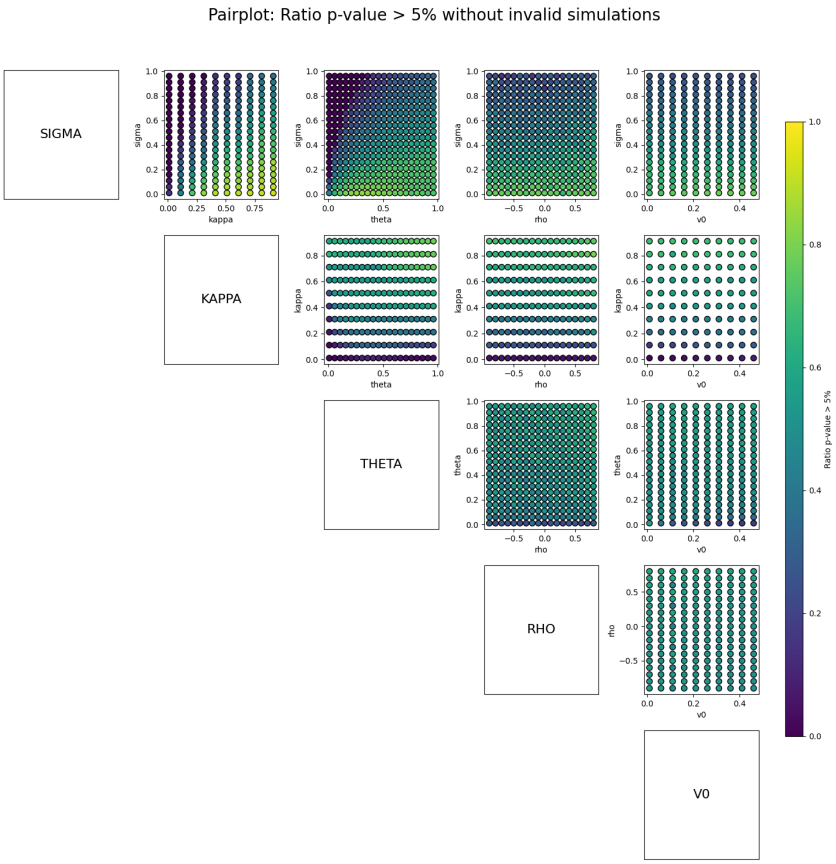


Figure 6.8: Pairplot for each pair of parameters for the Heston Model and the percentage of p values of the Kolmogorov-Smirnov-Test for the Gram-Charlier Expansion with Cumulants vs the theoretical density above 5%. Invalid simulations are excluded and $\mu = 0$.

cumulants are not affected by the initial value.

For σ, the results suggest that lower values lead to better approximations, while higher values of κ improve the fit. However, the influence of θ appears more complex, as there is no clear trend. Generally, higher κ values seem to favor higher θ values, while at low σ values, a low θ is preferable. Figure 6.9 provides a 3D visualization of the relationship between σ, κ, and θ and their impact on the Kolmogorov-Smirnov p-values.

It is challenging to derive general rules for which parameter combinations lead to good approximations, primarily due to the limitations of graphical visualizations, which cannot effectively represent high-dimensional relationships. To overcome this, we employ classification methods that do not rely on visual representation, such as logistic regression and random forests.

For this analysis, all valid simulations were used and split into 80% training data and 20% test data. The independent variables X are the Heston model parameters, while the dependent variable y indicates whether the Kolmogorov-Smirnov p-value exceeds

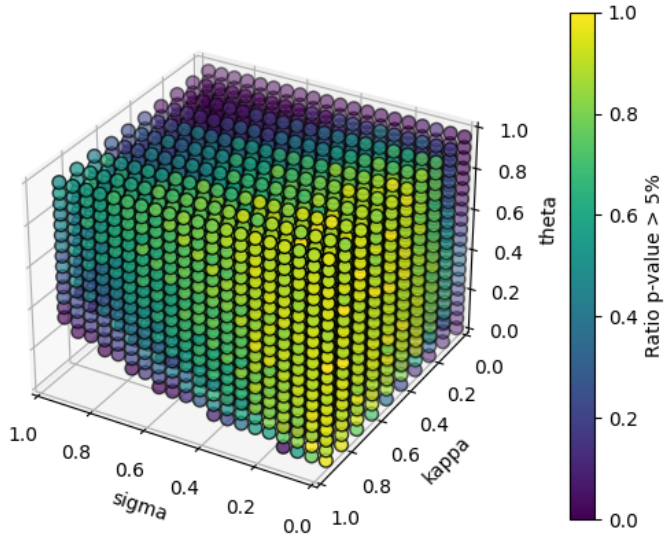


Figure 6.9: Percentage of simulations where the p-value of the Kolmogorov-Smirnov-Test against the theoretical density is above 5%. Invalid simulations are excluded. $\mu = 0$

5%. The dataset is imbalanced, with 60.98% of simulations yielding a p-value above 5%. The logistic regression model achieves an accuracy of 79.53%, while the random forest model, consisting of 100 decision trees, reaches 91.59% accuracy. The precision for $y = 1$ is 82% for logistic regression and 92% for the random forest model.

Precision is particularly relevant in this context, as it represents the proportion of true positives among all predicted positives. A low precision would imply that parameter combinations unsuitable for the Gram-Charlier approximation are mistakenly classified as suitable. In a financial context, where such approximations might be used to price securities, incorrect classifications could lead to substantial financial losses.

Examining the logistic regression coefficients (Table 6.2) and the feature importance scores from the random forest model (Figure 6.10) reveals that σ is by far the most influential parameter. Both θ and κ are also important, but the models disagree on which of the two has a stronger impact. The role of ρ is also unclear—it is insignificant in logistic regression but influential in the random forest model. Conversely, μ and v_0 play no significant role. All parameters included in the models are statistically significant.

Parameter	**Coefficient**	**Standard Error**	z	$\mathbb{P} > \|z\|$
const	-0.8708	0.010	-84.332	0.000
μ	-1.2841	0.111	-11.568	0.000
σ	-7.2015	0.017	-429.506	0.000
κ	4.4584	0.014	318.689	0.000
θ	3.0120	0.013	237.146	0.000
ρ	0.4219	0.005	78.359	0.000
v_0	-0.0728	0.019	-3.765	0.000

Table 6.2: Coefficients of the logistic regression

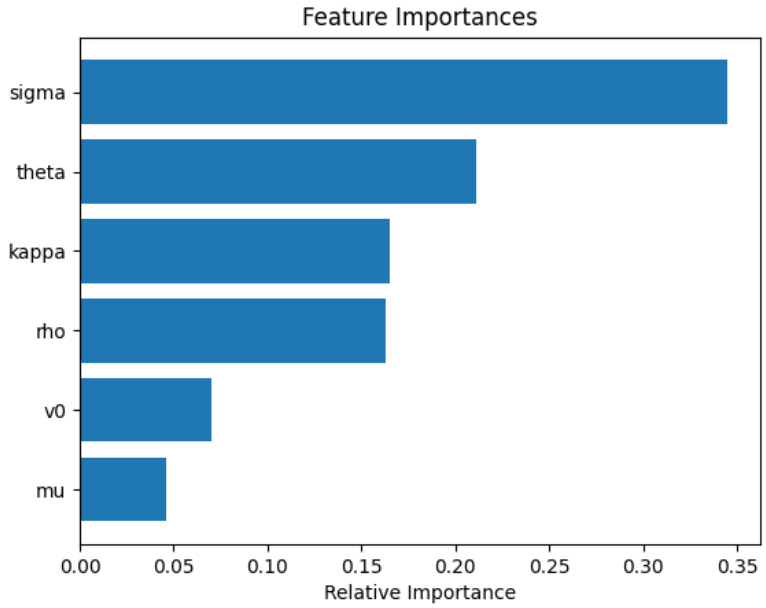
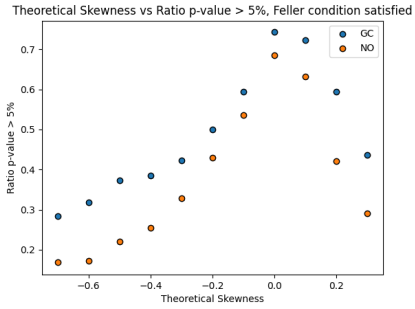
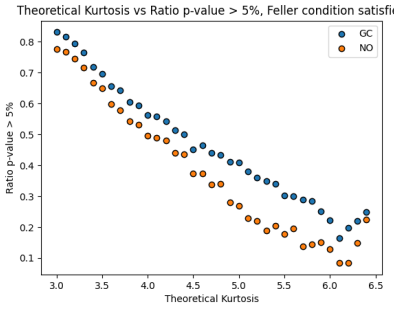


Figure 6.10: Feature Importances of the Random Forest



(a) Theretical Skewness



(b) Theoretical Kurtosis

Figure 6.11: Ratio of simulations where the p-value of the Kolmogorov-Smirnov-Test against the theoretical density is above 5% in relation to the theoretical skewness and kurtosis. GC stands for the Gram-Charlier-Expansion while NO stands for the Normal Distribution. Feller condition is fulfilled.

6.3 Comparing the Expansion Methods

6.3.1 General Approximation

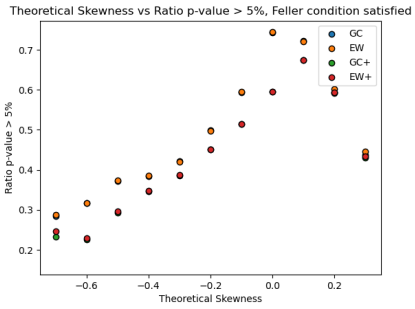
So far, we have analyzed the Gram-Charlier expansion from multiple perspectives. However, an important question remains: Are expansion methods even worth using? To answer this, we first examine whether adding the third and fourth cumulants improves the results, specifically in relation to the theoretical skewness and kurtosis.

As a baseline model, we use the normal distribution, which relies only on the first two cumulants—the mean (μ) and variance (σ^2). This allows us to compare whether incorporating higher-order cumulants provides a meaningful advantage. The results of this comparison are shown in Figure 6.11.

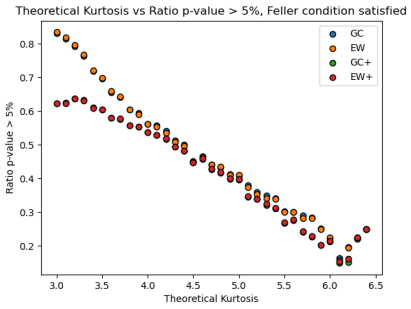
The findings indicate that using a more complex expansion method than the normal distribution is beneficial, as the approximation improves significantly for given levels of skewness and kurtosis.

A similar pattern emerges when comparing the Gram-Charlier expansion and the Edgeworth expansion, along with their positivity-constrained variants. The results, shown in Figure 6.12, indicate that positivity-constrained variants perform significantly worse. This decline in performance is due to the distortion effect introduced by the positivity constraint.

Moreover, no specific Heston model parameter dependencies were found where positivity-constrained expansions outperform their unconstrained counterparts. However, this is largely because the chosen parameter combinations for the simulations never result in theoretical skewness or kurtosis values that exceed the validity range of the expansion methods. As a result, the positivity constraints never actually prevent negative density values, which would otherwise invalidate the expansion as a proper density function.



(a) Theretical Skewness



(b) Theoretical Kurtosis

Figure 6.12: Ratio of simulations where the p-value of the Kolmogorov-Smirnov-Test against the theoretical density is above 5% in relation to the theoretical skewness and kurtosis. GC stands for the Gram-Charlier-Expansion, EW for the Edgeworth-Expansion. A + behind the Expansion denotes the variant with positivity constraints. Feller condition is fulfilled.

To investigate cases where the Edgeworth expansion provides a better approximation, we search the database for parameter combinations where:

1. The p-value from the Kolmogorov-Smirnov test is higher for the Edgeworth expansion than for the Gram-Charlier expansion.
2. The p-value for the Edgeworth expansion is above 5%, while the p-value for the Gram-Charlier expansion is below 5%.

A total of 2,866 such parameter combinations were identified out of 501,713 possible cases (after filtering out invalid simulations), representing only 0.57% of all simulations. This suggests that the Edgeworth expansion is not a better alternative to the Gram-Charlier expansion for the Heston model.

6.3.2 Approximation of the Tails

To compare the tails of the distributions, we use the Anderson-Darling test. Additionally, we attempted a comparison using the Hill estimator to estimate the tail index. However, it was not possible to identify a plateau in the Hill plot, which suggests that the tails are not sufficiently fat, as seen in distributions like the Pareto distribution. This is a critical requirement for the Hill estimator to be applicable.

One might expect that using a test that is more sensitive to the tails, such as the Anderson-Darling test, would yield different results compared to previous tests. However, the comparison between Figure 6.13 and Figure 6.8 reveals practically no difference. This suggests that the Gram-Charlier expansion based on cumulants provides a good approximation of the tails as well.

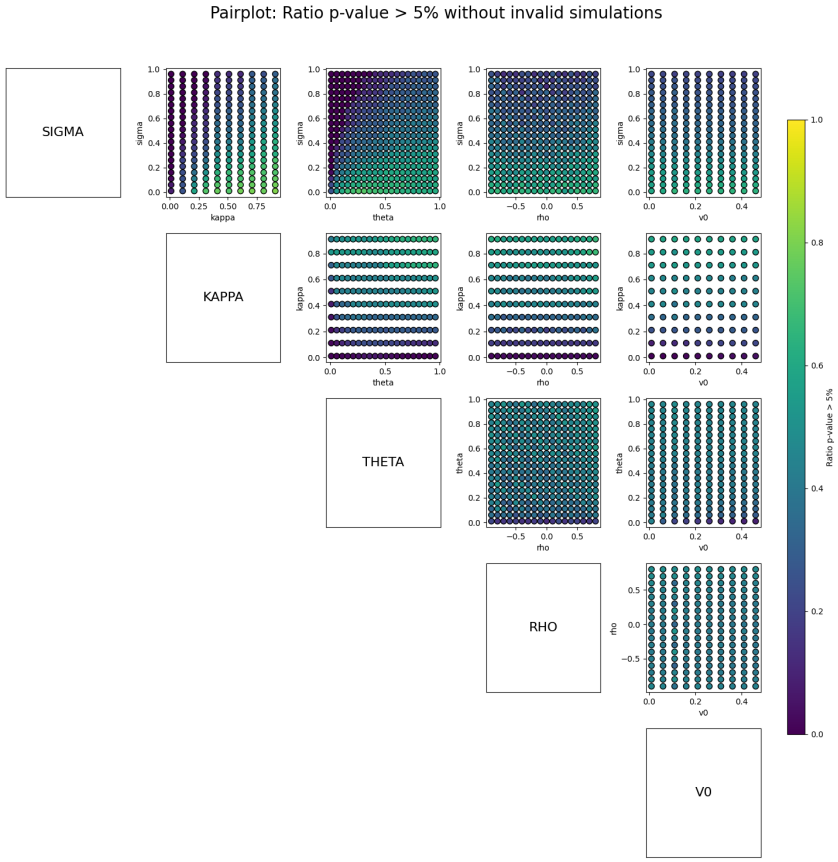


Figure 6.13: Pairplot for each pair of parameters for the Heston Model and the percentage of p values of the Anderson-Darling-Test for the Gram-Charlier Expansion with Cumulants vs the theoretical density above 5%. Invalid simulations are excluded and $\mu = 0$.

Bibliography

Aboura, S., & Maillard, D. (2016). Option Pricing Under Skewness and Kurtosis Using a Cornish–Fisher Expansion. *Journal of Futures Markets*, *36*(12), 1194–1209. https://doi.org/10.1002/fut.21787

Abramowitz, M., & Stegun, I. A. (1968). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. U.S. Government Printing Office.

Albrecher, H., Mayer, P., Schoutens, W., & Tistaert, J. (2007). The little Heston trap. *Wilmott*, (1), 83–92.

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

Andersen, L. B. G. (2007). Efficient Simulation of the Heston Stochastic Volatility Model. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.946405

Andersen, T. G., & Bollerslev, T. (1998). Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts. *International Economic Review*, *39*(4), 885–905. https://doi.org/10.2307/2527343

Anderson, T. W. (1962). On the Distribution of the Two-Sample Cramer-von Mises Criterion. *The Annals of Mathematical Statistics*, *33*(3), 1148–1159. https://doi.org/10.1214/aoms/1177704477

Anderson, T. W., & Darling, D. A. (1954). A Test of Goodness of Fit. *Journal of the American Statistical Association*, *49*(268), 765–769. https://doi.org/10.1080/01621459.1954.10501232

Barro, R. J. (2006). Rare Disasters and Asset Markets in the Twentieth Century*. *The Quarterly Journal of Economics*, *121*(3), 823–866. https://doi.org/10.1162/qjec.121.3.823

Brenn, T., & Anfinsen, S. N. (2017). A revisit of the Gram-Charlier and Edgeworth series expansions.

Chang, Y., Wang, Y., & Zhang, S. (2021). Option Pricing under Double Heston Jump-Diffusion Model with Approximative Fractional Stochastic Volatility. *Mathematics*, *9*(2), 126. https://doi.org/10.3390/math9020126

Charlier, C. V. L. (1914). Contributions to the mathematical theory of statistics 5. Frequency curves of type A in heterograde statistics. *Meddelanden fran Lunds Astronomiska Observatorium Serie I*, *57*, 1–17
ADS Bibcode: 1914MeLuF..57....1C.

Chebyshev, P. (1860). Sur le développement des fonctions à une seule variable. *Œuvres*, *1*, 193–200.

Choe, G. H., & Lee, K. (2014). High Moment Variations and Their Application. *Journal of Futures Markets*, *34*(11), 1040–1061. https://doi.org/10.1002/fut.21635

Cornish, E. A., & Fisher, R. A. (1938). Moments and Cumulants in the Specification of Distributions. *Revue de l'Institut International de Statistique / Review of the International Statistical Institute*, *5*(4), 307–320. https://doi.org/10.2307/1400905

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, *53*(2), 385–407. https://doi.org/10.2307/1911242

Cramér, H. (1928). On the composition of elementary errors: First paper: Mathematical deductions. *Scandinavian Actuarial Journal*, *1928*(1), 13–74. https://doi.org/10.1080/03461238.1928.10416862

Cramér, H. (1999). *Mathematical Methods of Statistics*. Princeton University Press.

Daniels, H. E. (1954). Saddlepoint Approximations in Statistics. *The Annals of Mathematical Statistics*, *25*(4), 631–650.

Danielsson, J., Ergun, L. M., De Haan, L., & De Vries, C. G. (2016). Tail Index Estimation: Quantile Driven Threshold Selection. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2717478

de Laplace, P.-S. (1812). Théorie analytique des probabilités. *Œuvres complètes*, *7*(2), 194–203.

Doane, D. P., & Seward, L. E. (2011). Measuring Skewness: A Forgotten Statistic? *Journal of Statistics Education*, *19*(2). https://doi.org/10.1080/10691898.2011.11889611

Du, X. (2010). System reliability analysis with saddlepoint approximation. *Structural and Multidisciplinary Optimization*, *42*(2), 193–208. https://doi.org/10.1007/s00158-009-0478-x

Dunn, R., Hauser, P., Seibold, T., & Gong, H. (2014). Estimating Option Prices with Heston ' s Stochastic Volatility Model.

Edgeworth, F. Y. (1907). On the Representation of Statistical Frequency by a Series. *Journal of the Royal Statistical Society*, *70*(1), 102–106. https://doi.org/10.2307/2339504

Eraker, B. (2004). Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices. *The Journal of Finance*, *59*(3), 1367–1403. https://doi.org/10.1111/j.1540-6261.2004.00666.x

Eraker, B., Johannes, M., & Polson, N. (2003). The Impact of Jumps in Volatility and Returns. *The Journal of Finance*, *58*(3), 1269–1300. https://doi.org/10.1111/1540-6261.00566

Fischler, M. (2017). Answer to "Definition of Tail-index of a probability distribution".

Foundation, P. (n.d.). 15. Floating-Point Arithmetic: Issues and Limitations — Python 3.12.4 documentation.

Frankfurt, B. (n.d.). So funktioniert die Börse.

Fukasawa, M., & Matsushita, K. (2021). Realized cumulants for martingales. *Electronic Communications in Probability*, *26*(none). https://doi.org/10.1214/21-ECP382

Gallo, G., Okhrin, O., & Storti, G. (2024). Dynamic tail risk forecasting: What do realized skewness and kurtosis add? https://doi.org/10.48550/arXiv.2409.13516

Gatheral, J. (2011). *The Volatility Surface: A Practitioner's Guide*. John Wiley & Sons.

Gram, J. (1883). Ueber die Entwickelung reeller Functionen in Reihen mittelst der Methode der kleinsten Quadrate. *crll*, *1883*(94), 41–73. https://doi.org/10.1515/crll.1883.94.41

Hermite, C. (1864). Sur un nouveau développement en série de fonctions. *Œuvres*, *2*, 293–308.

Heston, S. L. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies*, *6*(2), 327–343.

Hill, B. M. (1975). A Simple General Approach to Inference About the Tail of a Distribution. *The Annals of Statistics*, *3*(5). https://doi.org/10.1214/aos/1176343247

Hu, G., & Liu, Y. (2022). The Pricing of Volatility and Jump Risks in the Cross-Section of Index Option Returns. *Journal of Financial and Quantitative Analysis*, *57*(6), 2385–2411. https://doi.org/10.1017/S0022109022000333

Joanes, D. N., & Gill, C. A. (1998). Comparing measures of sample skewness and kurtosis. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *47*(1), 183–189. https://doi.org/10.1111/1467-9884.00122

Jondeau, E., & Rockinger, M. (2001). Gram–Charlier densities. *Journal of Economic Dynamics and Control*, *25*(10), 1457–1483. https://doi.org/10.1016/S0165-1889(99)00082-2

Kolmogorov, A. N. (1993). Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'Istituto Italiano degli Attuari*, (4), 83–91.

Laplace, P.-S. (1811). Mémoire sur les intégrales définies et leur application aux probabilités, et spécialement a la recherche du milieu qu'il faut choisir entre les resultats des observations. *Mémoires de la Classe des Sciences Mathématiques et Physiques de l'Institut Impérial de France*, *11*, 297–347.

Leonardo.Z. (2013). Answer to "How can I make numbers more precise in Python?"

Liu, Z., Wang, K., & Liu, J. (2014). Realized Skewness at High Frequency and the Link to a Conditional Market Premium. https://doi.org/10.2139/ssrn.2224216

Lugannani, R., & Rice, S. (1980). Saddle point approximation for the distribution of the sum of independent random variables. *Advances in Applied Probability*, *12*(2), 475–490. https://doi.org/10.2307/1426607

Maillard, D. (2018). A User's Guide to the Cornish Fisher Expansion. https://doi.org/10.2139/ssrn.1997178

Mitropol'skii, A. (2020). Gram-Charlier series.

Neuberger, A. (2012). Realized Skewness. *Review of Financial Studies*, *25*(11), 3423–3455. https://doi.org/10.1093/rfs/hhs101

Neuberger, A., & Payne, R. (2021). The Skewness of the Stock Market over Long Horizons (S. Van Nieuwerburgh, Ed.). *The Review of Financial Studies*, *34*(3), 1572–1616. https://doi.org/10.1093/rfs/hhaa048

Okhrin, O., Rockinger, M., & Schmid, M. (2022). Simulating the Cox–Ingersoll–Ross and Heston processes: Matching the first four moments. *Journal of Computational Finance*. https://doi.org/10.21314/JCF.2022.022

Okhrin, O., Rockinger, M., & Schmid, M. (2023). Distributional properties of continuous time processes: From CIR to bates. *AStA Advances in Statistical Analysis*, *107*(3), 397–419. https://doi.org/10.1007/s10182-022-00459-3

Radziwill, N. M. (2017). *Statistics (the Easier Way) with R*. Lapis Lucera.

Reid, N. (1988). Saddlepoint Methods and Statistical Inference. *Statistical Science*, *3*(2), 213–227. https://doi.org/10.1214/ss/1177012906

Tsoskounoglou, A. (2024). Simulating the Heston model with Quadratic Exponential.

Von Mises, R. (1928). *Wahrscheinlichkeit Statistik und Wahrheit*. Springer Berlin / Heidelberg.

Zhang, J. E., Zhen, F., Sun, X., & Zhao, H. (2017). The Skewness Implied in the Heston Model and Its Application. *Journal of Futures Markets*, *37*(3), 211–237. https://doi.org/10.1002/fut.21801

Zhang, L., Mykland, P. A., & Aït-Sahalia, Y. (2005). A Tale of Two Time Scales. *Journal of the American Statistical Association*, *100*(472), 1394–1411. https://doi.org/10.1198/016214505000000169

Zhao, H., Zhang, J. E., & Chang, E. C. (2013). The Relation between Physical and Risk-neutral Cumulants. *International Review of Finance*, *13*(3), 345–381. https://doi.org/10.1111/irfi.12013

Declaration of independence

I hereby declare that this thesis was written independently and without the use of any other resources than those stated. Any ideas taken literally or analogously from other sources are identified as such. I further declare that I have not submitted or will not submit this thesis as an examination paper to any other institution.

Dresden, XX.XX.XXXX

Henry Haustein