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 The Heston Model

1.1 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:

$$\mathrm{d}S_t = \mu S_t \mathrm{d}t + \sqrt{v_t} S_t \mathrm{d}W_t^S \tag{1.1.1}$$

$$\mathrm{d}X_t = \mathrm{d}\log(S_t) = \left(\mu - \frac{1}{2}v_t\right)\mathrm{d}t + \sqrt{v_t}\mathrm{d}W_t^S \tag{1.1.2}$$

$$\mathrm{d}v_t = \kappa(\theta - v_t)\mathrm{d}t + \sigma\sqrt{v_t}\mathrm{d}W_t^v \tag{1.1.3}$$

$$\mathbb{E}(\mathrm{d}W_t^S \mathrm{d}W_t^v) = \rho \mathrm{d}t \tag{1.1.4}$$

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:

$$\begin{aligned} v_t \mid v_0 &\sim c \cdot \chi^{2'}_{\nu}(\Lambda) \\ c &= \sigma^2 \left(1 - \exp(-\kappa t)\right)(4\kappa)^{-1} \\ \nu &= \frac{4\kappa\theta}{\sigma^2} \\ \Lambda &= \frac{v_0}{c}\exp(-\kappa t) \end{aligned} \tag{1.1.5}$$

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

1.2 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

$$\begin{aligned}
A &= \mu \cdot \tau \cdot t \cdot \mathrm{i} \\
d &= \sqrt{(\rho\sigma\mathrm{i}t - \kappa)^2 - \sigma^2(-\mathrm{i}t - t^2)} \\
g &= \frac{\kappa - \rho\sigma\mathrm{i}t - d}{\kappa - \rho\sigma\mathrm{i}t + d} \\
B &= \frac{\theta\kappa}{\sigma^2}\left(\tau(\kappa - \rho\sigma\mathrm{i}t - d) - 2\log\left[\frac{1 - g\exp(-d\tau)}{1 - g}\right]\right) \\
\gamma &= \frac{2\kappa\theta}{\sigma^2} \\
C &= \log\left(\left[\frac{2\kappa}{\sigma^2}\right]^{\gamma} \cdot \left\{\frac{2\kappa}{\sigma^2} - \frac{\kappa - \rho\sigma\mathrm{i}t - d}{\sigma^2} \cdot \frac{1 - \exp(-d\tau)}{1 - g\exp(-d\tau)}\right\}^{-\gamma}\right)
\end{aligned} \tag{1.2.1}$$

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 1.1). 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 (1.2.1) 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

$$C_{unc} = \gamma\log\left(\frac{2\kappa}{\sigma^2}\right) - \gamma\log\left(\frac{2\kappa}{\sigma^2} - \frac{\kappa - \rho\sigma t\mathrm{i} - d}{\sigma^2}\frac{1 - \exp(-d\tau)}{1 - g\exp(-d\tau)}\right) \tag{1.2.2}$$

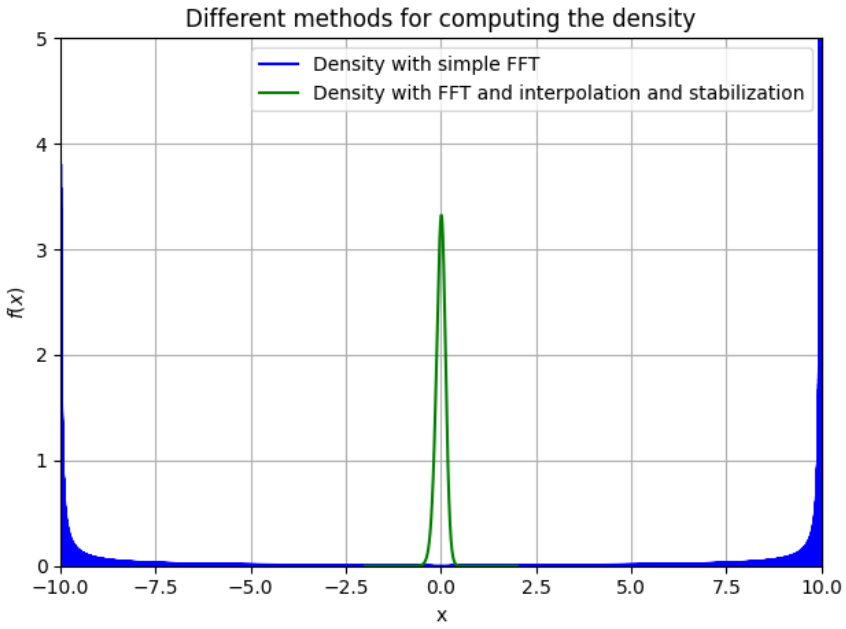

Figure 1.1: Comparison of different methods for the inverse Fourier transformation of the characteristic function of the Heston model ($\mu = 0$, $\kappa = 3$, $\theta = 0.19$, $\sigma = 0.4$, $\rho = -0.7$, $\tau = \frac{1}{12}$). Grid points: $N = 2^{15}$

1.3 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 (1.1.5) using a mixture of distributions: a Dirac distribution and a noncentral Gaussian distribution. For sufficiently large values of v_t,

$$v_{t+\Delta} = a(b + Z_v)^2 \tag{1.3.1}$$

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

$$v_{t+\Delta} = \Psi^{-1}(U_v, p, \beta) \tag{1.3.2}$$

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

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 (1.3.2) is used; otherwise, (1.3.1) 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 $\xi_1 = \xi_2 = 0.5$.

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

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 sec-

ond (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 i-th intraday log return $r_{t,i}$ for day t is computed as

$$r_{t,i} = \log(P_{t,i/N}) - \log(P_{t,(i-1)/N})$$

where p represents the natural logarithm of the price, and N denotes the number of return observations within a trading day. The opening log-price on day t is given by $p_{t,0}$, while the closing log-price is $p_{t,1}$. Using five-minute returns, a standard 6.5-hour trading session results in $N = 78$. Based on this, the daily realized variance is computed as

$$\hat{\sigma}_t^2 = \sum_{i=1}^{N} r_{t,i}^2$$

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 as

$$\hat{\gamma}_1 = \frac{\sqrt{N} \cdot \sum_{i=1}^{N} r_{t,i}^3}{\hat{\sigma}_t^3}$$
$$\hat{\gamma}_2 = \frac{N \cdot \sum_{i=1}^{N} r_{t,i}^4}{\hat{\sigma}_t^4}$$

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:

$$[X]_t = X_t^2 - 2\int_0^t X_u \, \mathrm{d}X_u \tag{2.3.1}$$

and the quadratic covariation process of two semimartingales X and Y:

$$[X, Y]_t = X_t Y_t - \int_0^t X_u \, \mathrm{d}Y_u - \int_0^t Y_u \, \mathrm{d}X_u \tag{2.3.2}$$

For the log-return process R_t,

$$R_t = \log(P_t) - \log(P_0)$$

equations (2.3.1) and (2.3.2) can be approximated as follows:

$$[R]_T \approx \sum_{i=1}^{N}(R_i - R_{i-1})^2$$
$$[R, R^2]_T \approx \sum_{i=1}^{N}(R_i - R_{i-1})(R_i^2 - R_{i-1}^2)$$
$$[R^2]_T \approx \sum_{i=1}^{N}(R_i^2 - R_{i-1}^2)^2$$

From these, the moments follow:

$$\mathbb{E}(R_T^3) = \frac{3}{2}\mathbb{E}([R, R^2]_T)$$
$$\mathbb{E}(R_T^4) = \frac{3}{2}\mathbb{E}([R^2]_T)$$

The estimation of low-frequency variance follows the same approach as Andersen & Bollerslev (1998) and Amaya et al. (2015). Neuberger & Payne (2021) propose a novel method that only requires log prices to be martingales and stationary, meaning there is no drift. Under these conditions, they define new moment measures that approximate the standard definitions:

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

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

$$
\begin{aligned}
y_t^{(j)} &= \sum_{u=1}^{T} \frac{x^{(j)}(R_{t-1}(u))}{T} \quad \text{for } j = 1, 2L \\
x^{(1)} &= e^r - 1
\end{aligned} \tag{2.3.3}
$$

where equation (2.3.3) 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.

2.4 The First 4 Moments of the Heston Model

Care must be taken with notation, as symbols such as μ and σ previously denoted the mean and variance but are now used to represent the drift and volatility of the Heston model (see Equations (1.1.1) and (1.1.3)). 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:

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

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, re-

spectively.

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

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

3 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.

3.1 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):

$$\begin{aligned} H_{n+1} &= x \cdot H_n(x) - H_n'(x) \\ H_n'(x) &= n \cdot H_{n-1}(x) \end{aligned} \tag{3.1.1}$$

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:

$$\begin{aligned} f(x)_{GC,A} \approx \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \\ \times \left[1 + \frac{\kappa_3}{6\sigma^3} H_3\left(\frac{x-\mu}{\sigma}\right) + \frac{\kappa_4}{24\sigma^4} H_4\left(\frac{x-\mu}{\sigma}\right)\right] \end{aligned} \tag{3.1.2}$$

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.

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 (3.1.2) can be rewritten in terms of the skewness γ_1 and excess kurtosis γ_2^*, defining $z = \frac{x-\mu}{\sigma}$:

$$f(x)_{GC,A} \approx \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)\left[1 + \frac{\gamma_1}{6}H_3(z) + \frac{\gamma_2^*}{24}H_4(z)\right] \tag{3.1.3}$$

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 3.1 and 3.2.

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 3.1 and Figures 3.3 and 3.4. 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 $(\gamma_1, \gamma_2^*) = (0, 2)$.

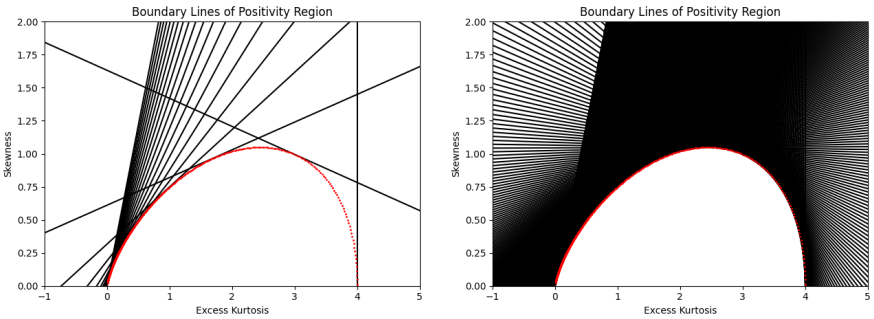

Figure 3.1: Boundary lines of the positivity region of the Gram-Charlier Expansion. The left image shows 20 lines, the right image shows 1000 lines. The red dots are the boundary points. The boundary is symmetric to the x-axis.

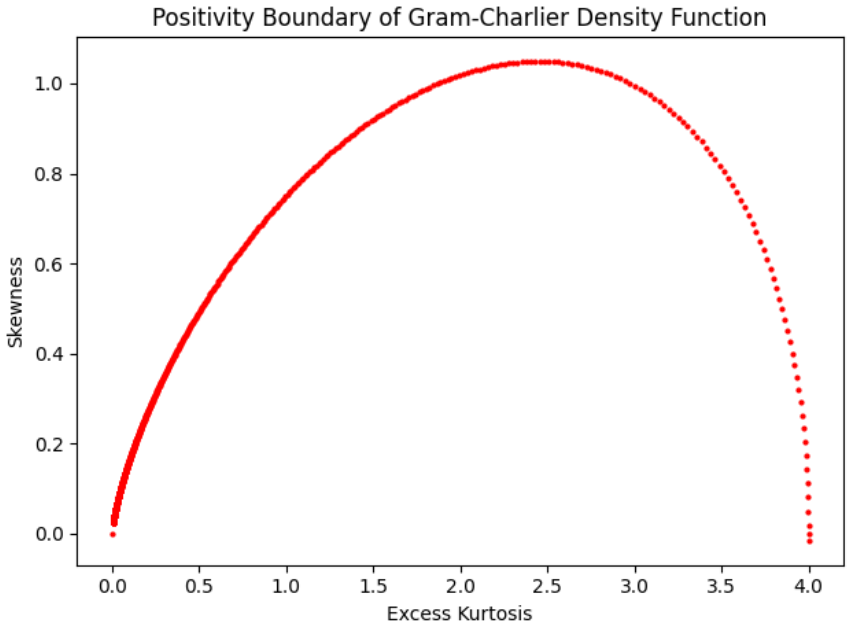

Figure 3.2: Approximation (1000 steps) of the positivity boundary of the Gram-Charlier Expansion. For simplicity, only the part above the x-axis is shown. The boundary is symmetric to the x-axis.

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 3.1: 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.

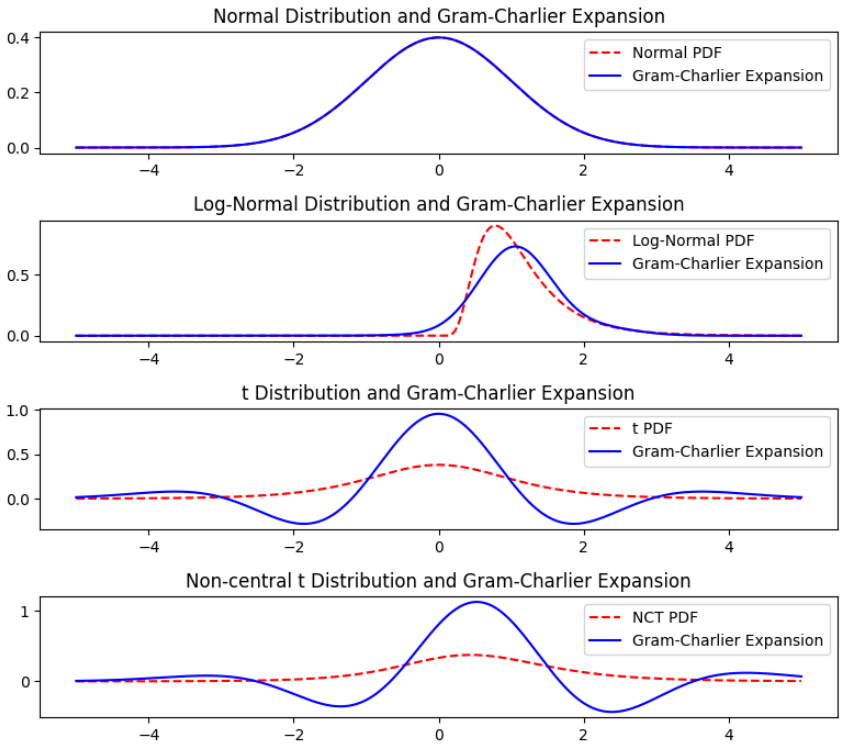

Figure 3.3: Gram-Charlier Expansion of different distributions

3.2 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

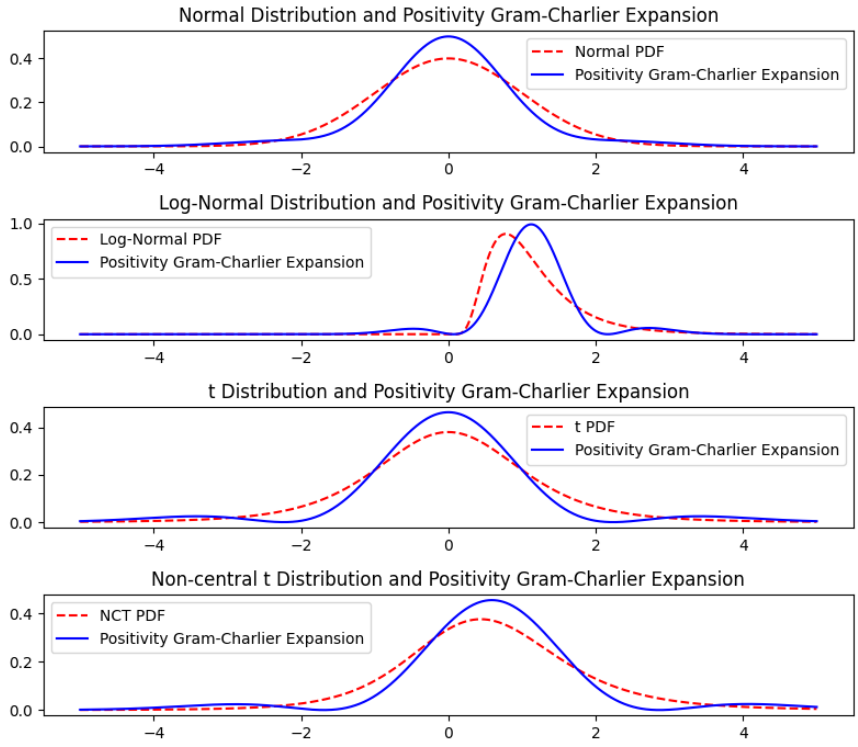

Figure 3.4: Gram-Charlier Expansion with positivity constraints of different distributions

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:

$$f(x)_{EW} \approx \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \times \left[1 + \frac{\kappa_3}{6\sigma^3} H_3\left(\frac{x-\mu}{\sigma}\right) + \frac{\kappa_4}{24\sigma^4} H_4\left(\frac{x-\mu}{\sigma}\right) + \frac{\kappa_3^2}{72\sigma^6} H_6\left(\frac{x-\mu}{\sigma}\right)\right] \tag{3.2.1}$$

Following the approach of Jondeau & Rockinger (2001) for the Gram-Charlier expansion, we determine the positivity boundary for the Edgeworth expansion by rewriting Equation (3.2.1) using $z = \frac{x-\mu}{\sigma}$:

$$f(x)_{EW} \approx \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)\left[1 + \frac{\gamma_1}{6}H_3(z) + \frac{\gamma_2^*}{24}H_4(z) + \frac{\gamma_1^2}{72}He_6(z)\right] \tag{3.2.2}$$

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

$$\begin{aligned} 0 &= 1 + \frac{\gamma_1}{6}He_3(z) + \frac{\gamma_2^*}{24}He_4(z) + \frac{\gamma_1^2}{72}He_6(z) \\ \gamma_1 &= \pm\sqrt{-\frac{72}{He_6(z)} - 3\gamma_2^*\frac{He_4(z)}{He_6(z)} + 36\frac{He_3(z)^2}{He_6(z)^2} - 6\frac{He_3(z)}{He_6(z)}} \end{aligned} \tag{3.2.3}$$

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 3.5 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 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 3.6 to 3.9.

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

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

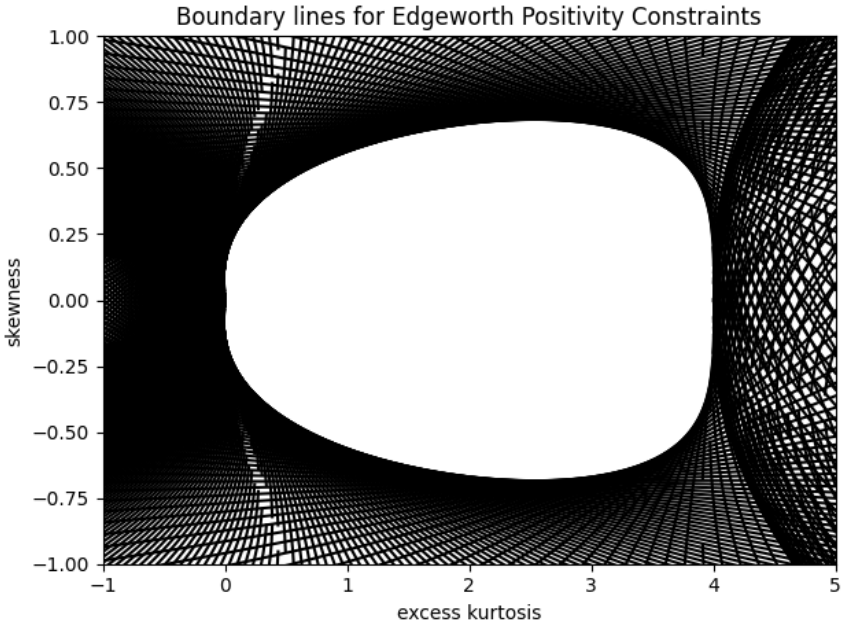

Figure 3.5: Boundary Lines for Edgeworth Expansion

compute the upper half and mirror it afterward.

2. Filtering out non-relevant points: Figure 3.7 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 3.8)

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 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

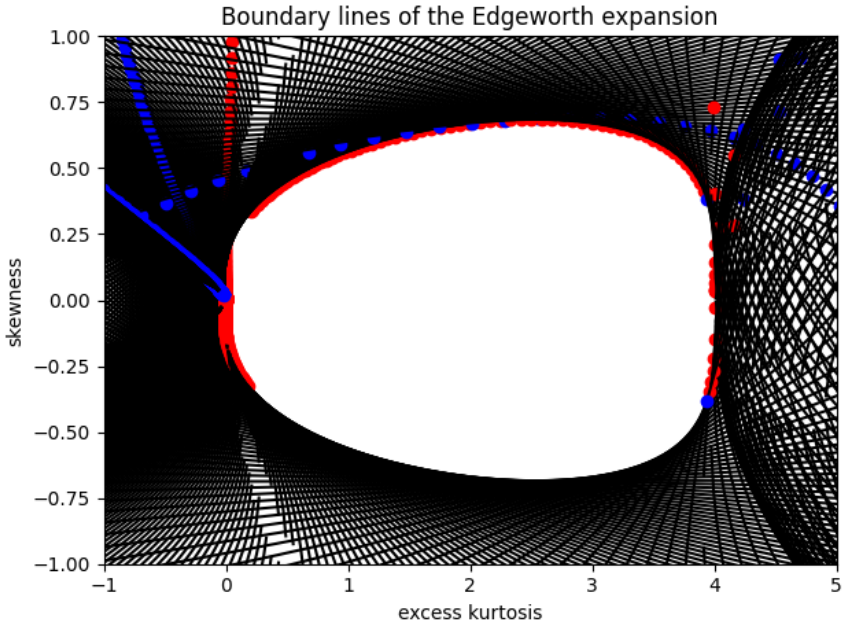

Figure 3.6: Intersections of Boundary Lines for Edgeworth Expansion, red is first intersection, blue is second intersection

positivity boundary (see Figure 3.9).

As with the Gram-Charlier expansion, we compare the constrained and unconstrained parameters for four distributions: standard normal, lognormal, t-distribution, and non-central t-distribution (see Table 3.1). The results, shown in Figures 3.10 and 3.11, 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.

3.3 Cornish-Fisher Expansion

The Cornish-Fisher expansion, introduced by Cornish and 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):

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

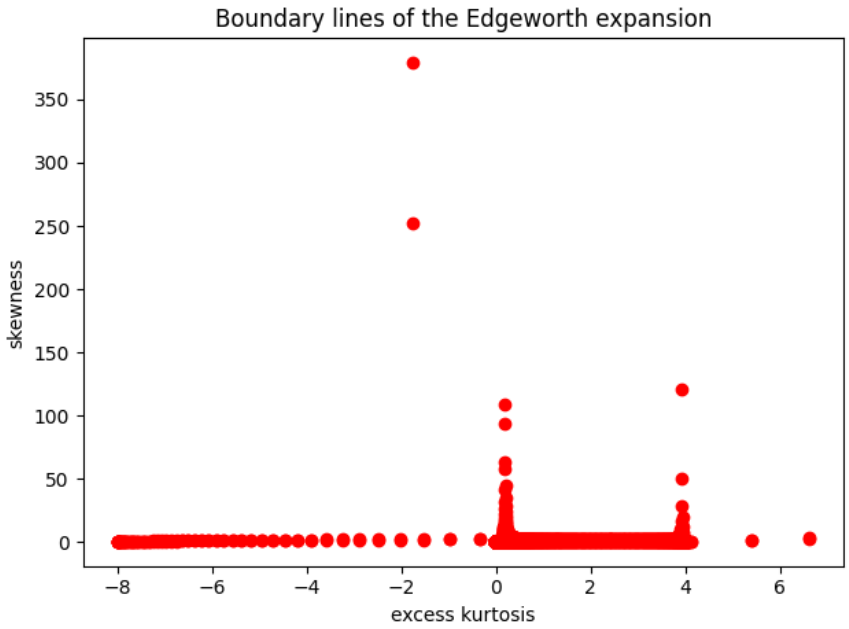

Figure 3.7: Intersections of Boundary Lines for Edgeworth Expansion (zoomed out), upper half

To obtain the probability density function (PDF), the quantiles x_p can be numerically computed and differentiated.

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:

$$s^* = \frac{M_3}{M_2^{3/2}}$$

$$k^* = \frac{M_4}{M_2^2} - 3$$

$$M_1 = 0 \tag{3.3.1}$$

$$M_2 = 1 + \frac{1}{96}k^2 + \frac{25}{1296}s^4 - \frac{1}{36}ks^2 \tag{3.3.2}$$

$$M_3 = s - \frac{76}{216}s^3 + \frac{85}{1296}s^5 + \frac{1}{4}ks - \frac{13}{144}ks^3 + \frac{1}{32}k^2s \tag{3.3.3}$$

$$M_4 = 3 + k + \frac{7}{16}k^2 + \frac{3}{32}k^3 + \frac{31}{3072}k^4 - \frac{7}{216}s^4 - \frac{25}{486}s^6 + \frac{21665}{559872}s^8$$

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-

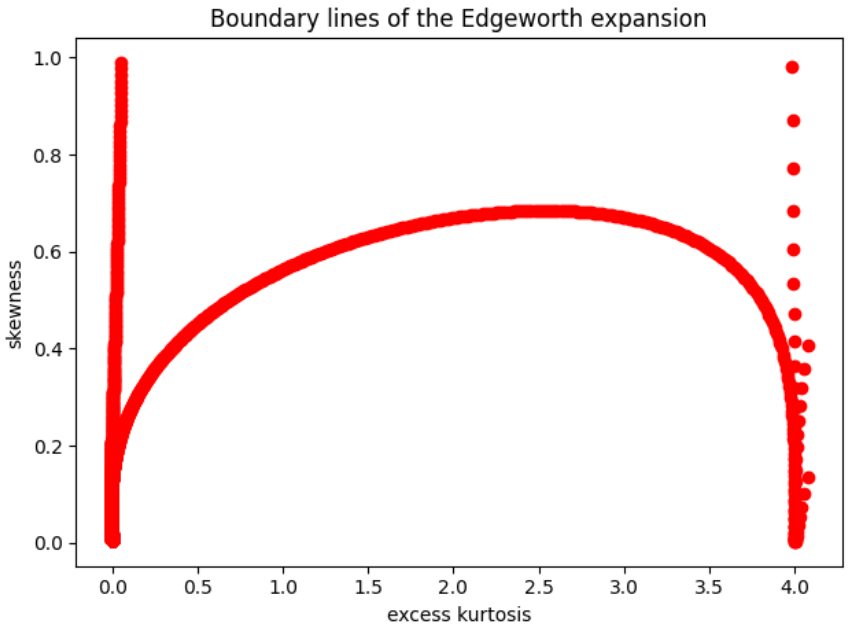

Figure 3.8: Intersections of Boundary Lines for Edgeworth Expansion (zoomed in), upper half

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 (see Figures 3.12 and 3.13).

3.4 Saddlepoint Approximation

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:

$$f(x)_{SP} \approx \frac{1}{\sqrt{2\pi \cdot K''(s)}} \exp(K(s) - s \cdot x) \tag{3.4.1}$$

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

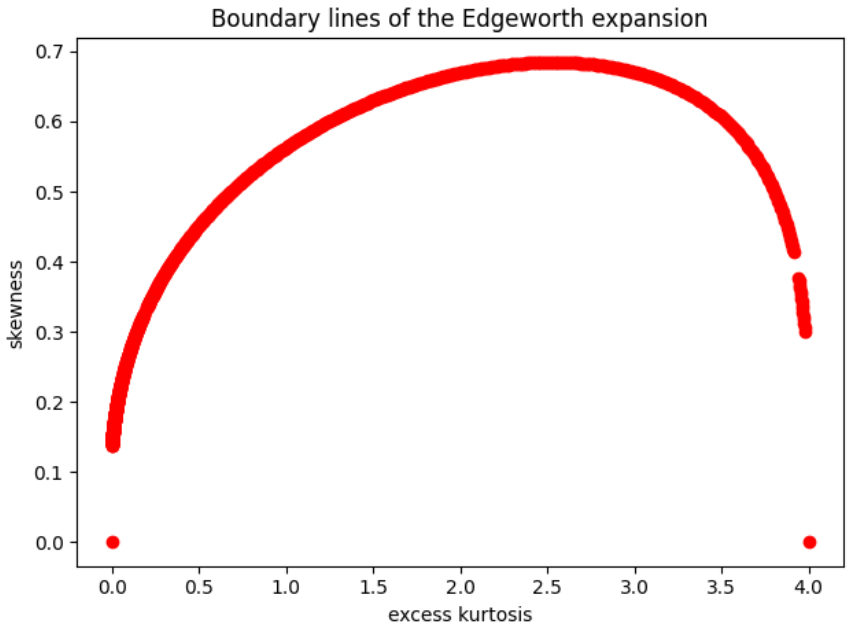

Figure 3.9: Final Boundary Points for Edgeworth Expansion, upper half

By definition, the cumulant generating function $K(s)$ can be approximated as:

$$K(s) \approx \kappa_1 s + \frac{\kappa_2 s^2}{2} + \frac{\kappa_3 s^3}{6} + \frac{\kappa_4 s^4}{24} \tag{3.4.2}$$

with its first and second derivatives:

$$K'(s) = \kappa_1 + \kappa_2 s + \frac{\kappa_3 s^2}{2} + \frac{\kappa_4 s^3}{6}$$
$$K''(s) = \kappa_2 + \kappa_3 s + \frac{\kappa_4 s^2}{2}$$

To find s, we solve the equation $K'(s) = x$ under different conditions:

1. If $\kappa_4 = 0$, $\kappa_3 = 0$, $\kappa_2 = 0$, and $\kappa_1 \neq 0$: No real solution exists.

2. If $\kappa_4 = 0$, $\kappa_3 = 0$, and $\kappa_2 \neq 0$: A direct solution is given by:

$$s = \frac{z - \kappa_1}{\kappa_2}$$

3. If $\kappa_4 = 0$ and $\kappa_3 \neq 0$: The equation becomes quadratic, yielding two solutions:

$$s = \frac{\pm\sqrt{-2\kappa_1\kappa_3 + 2\kappa_3 z + \kappa_2^2} - \kappa_2}{\kappa_3}$$

4. If $\kappa_4 \neq 0$, then three solutions exist, two of which are complex, and one is real

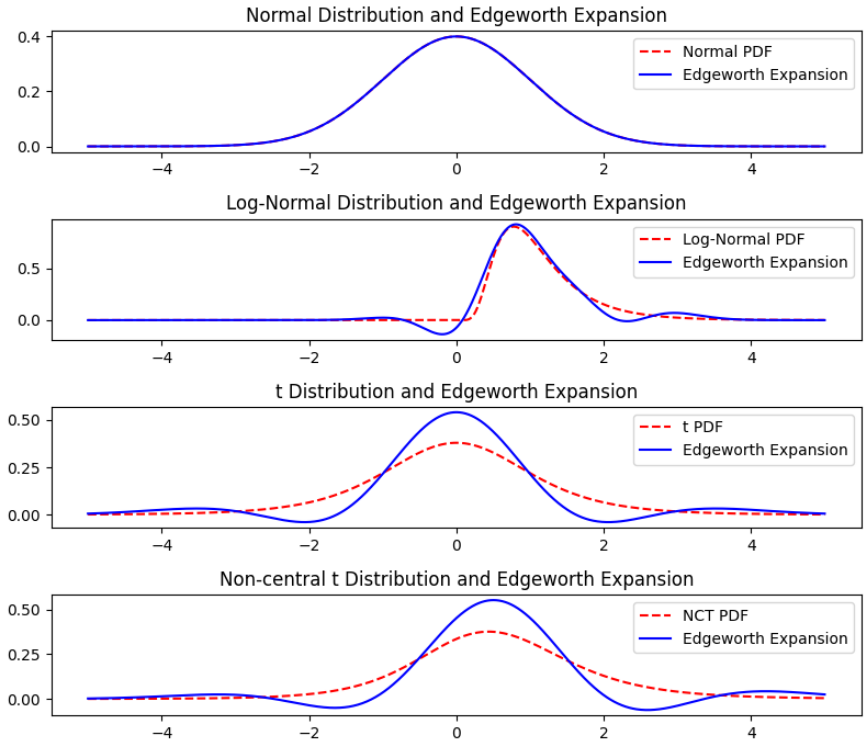

Figure 3.10: Edgeworth Expansion of different distributions

which is given by

$$s = \frac{1}{3\sqrt[3]{2}\kappa_4}\sqrt[3]{\Delta} - \frac{\sqrt[3]{2}\,(18\kappa_4\kappa_2 - 9\kappa_3^2)}{3\kappa_4\sqrt[3]{\Delta}} - \frac{\kappa_3}{\kappa_4}$$

$$\Delta := \sqrt{(-162\kappa_4^2\kappa_1 + 162\kappa_4^2 x + 162\kappa_4\kappa_3\kappa_2 - 54\kappa_3^3)^2 + 4(18\kappa_4\kappa_2 - 9\kappa_3^2)^3}$$
$$- 162\kappa_4^2\kappa_1 + 162\kappa_4^2 x + 162\kappa_4\kappa_3\kappa_2 - 54\kappa_3^3$$

The Saddlepoint Approximation successfully approximates the chosen distributions (see Table 3.1), as shown in Figure 3.14. The key advantages of this method are the high accuracy: The Saddlepoint method provides excellent approximations, even in cases where other expansions (such as Gram-Charlier or Edgeworth) fail. Furthermore there are no issues with negativity: The Saddlepoint Approximation does not produce negative densities. This is because the exp(·) term in Equation (3.4.1) 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.

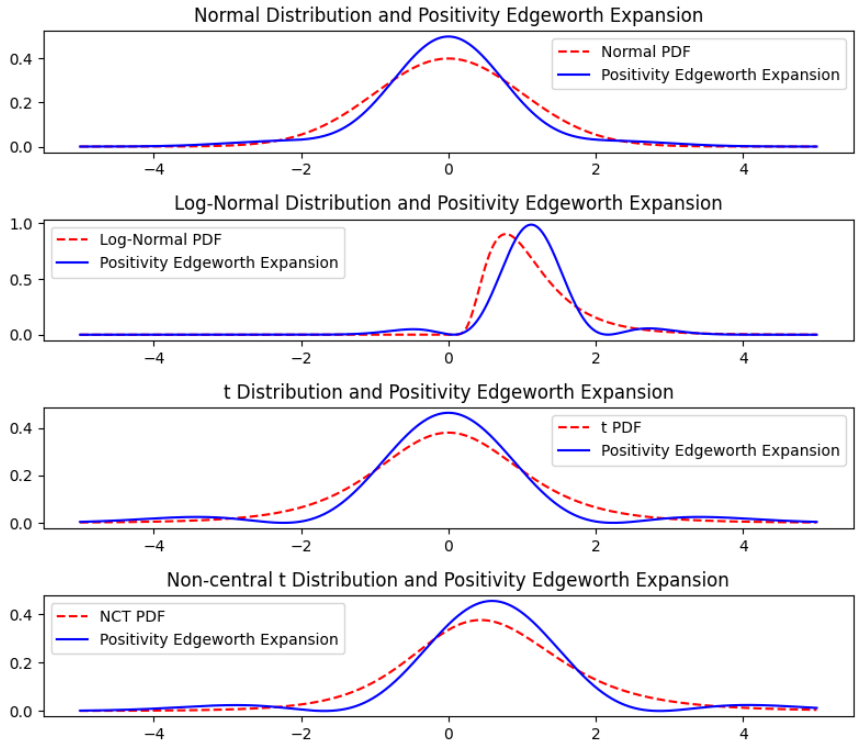

Figure 3.11: Edgeworth Expansion with positivity constraints of different distributions

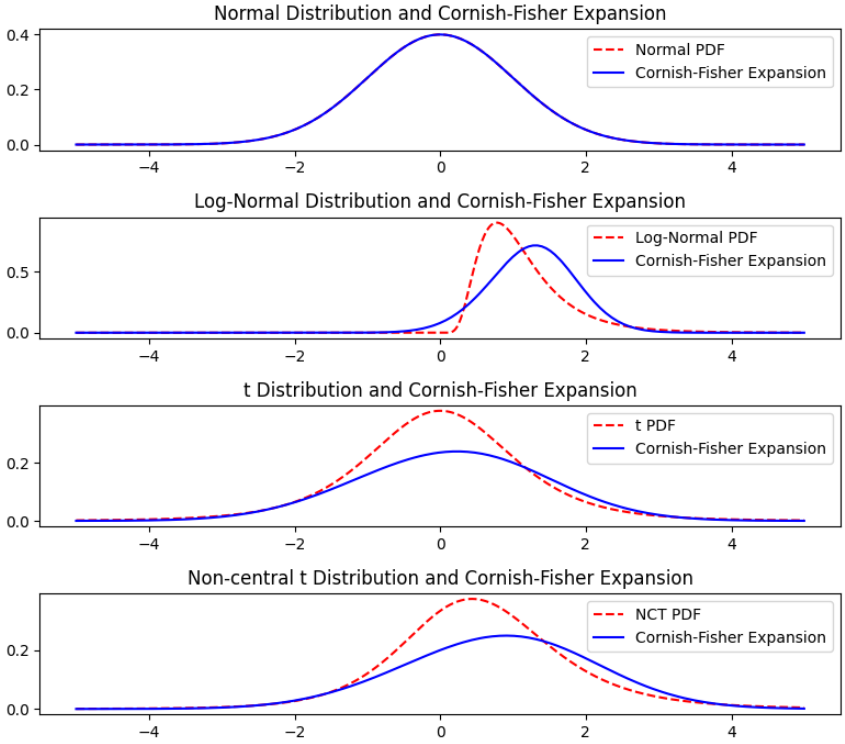

Figure 3.12: Cornish-Fisher Expansion of different distributions

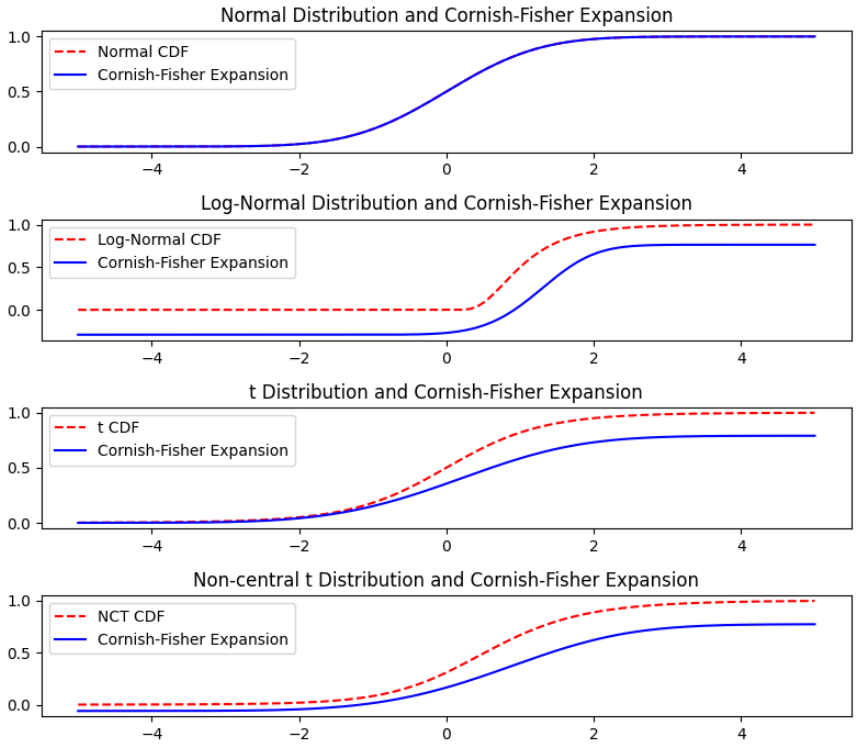

Figure 3.13: CDF of Cornish-Fisher Expansion of different distributions

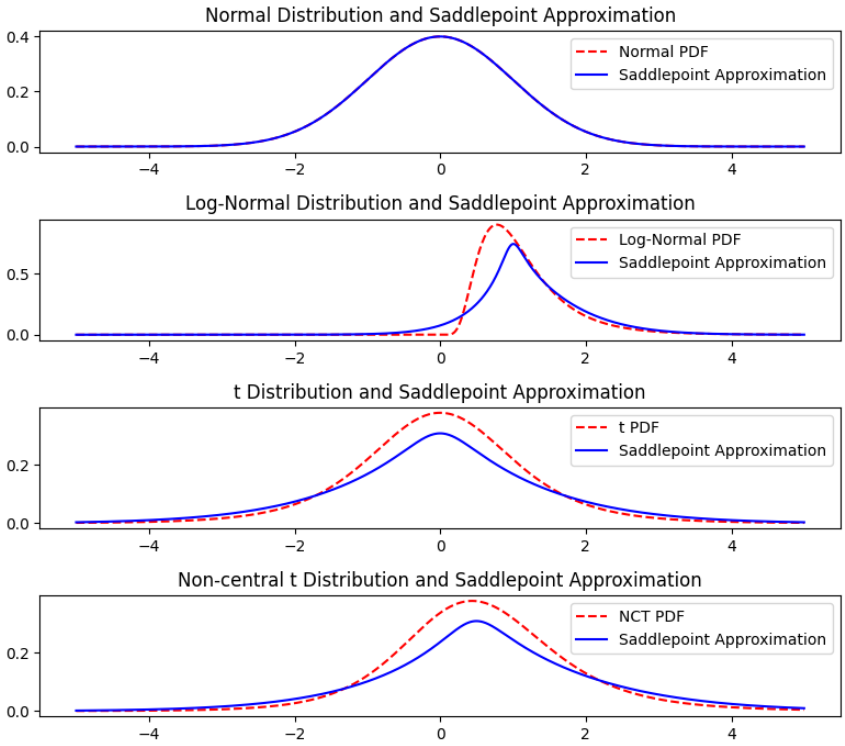

Figure 3.14: Saddlepoint Approximation of different distributions

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

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. (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.

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

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., 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

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

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.

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

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?"

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

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

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

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

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