On the Elliptic Distribution and Its Application to Leptokurtic Financial Data

Stephen H-T. Lihn*

e-mail: stevelihn@gmail.com

Abstract: A novel distribution based on elliptic curves is developed to capture major leptokurtic features in the return distribution of financial assets. This distribution is based on the Weierstrass equation and depressed cubic polynomial; therefore, it is intuitive, mathematically elegant, analytically tractable, and easy to calculate. It fits well to the historical daily log-return distributions of currencies, commodities, Treasury yields, VIX, and, most difficult of all, DJIA whose kurtosis is above 20. Various asymptotic behaviors are studied that encompasses a wide range of kurtosis from 2.3 to 35. The formal expansion near O of elliptic curves also suggests a possible different perspective on the tail behavior of the financial data. The numerical methods are built into an R package. A comprehensive sqlite database stores pre-calculated statistical attributes that cover a large parameter space. This allows practitioners to quickly examine any time series of interest. The distribution serves as a viable alternative to other fat-tailed distributions used in financial and risk modeling.

MSC 2010 subject classifications: Primary 60E05, 14H52, 62P05; secondary 65C50.
Keywords and phrases: distribution, elliptic curves, timeseries, fat tail, leptokurtic.

Contents

*Mr. Lihn is currently managing director of equity quantitative strategy at Novus Partners, Inc.. LinkedIn: https://www.linkedin.com/pub/stephen-horng-twu-lihn/0/71a/65

1. Introduction

One of the primary interests of the author has been to find a satisfactory probability distribution that can describe the heavily leptokurtic distributions of daily return data in the financial markets[11, 12]. For instance, many daily time series of equity indices and commodities have kurtosis of more than 10. One can easily collect 80 years of DJIA daily closing prices, whose log-returns have kurtosis of more than 20[1]. In the past, the pursues have been based on the method of

[1]In this paper, a normal distribution has kurtosis of 3. DJIA historical prices since 1928 are widely available. The `ecd` package comes with sample data of DJIA.

mixing Gaussian distributions of varying variances to produce a leptokurtic distribution, which has been known since 1970's [14]. However, these subordinated distributions either failed to capture the leptokurtic tails, or too complicated for practitioners to handle, especially when skew is involved. One candidate to describe such fat-tail data is the Levy alpha stable distribution (SaS). But the computation of its probability distribution density (PDF) is non-trivial; and, due to the infinite moments and very fat tails, the tails of SaS have to be truncated, which adds significant complexity (For instance, see Chapter 7 of Rachev, Fabozzi, and Menn, 2005[15]). The second popular choice is the Generalized Hyperbolic (GHYP) family of distributions by Barndorff-Nielsen [1]. As Fergusson and Platen [6] have shown, the maximum likelihood estimates on the log-return distribution of the world stock indices appear to cluster in the neighborhood of those of the Student's t-distribution, which is a special case of GHYP. Only the Student's t-distribution qualifies to fit world stock indices after 99% significance level applied. The best fits require the degrees of freedom to be near or less than 4.0, where the kurtosis ceases to exist.

I've also studied two subordinated distributions before - one based on log-normal cascade [11], and the other based on Poisson distribution [12]. The later in fact achieved high quality fits to the real-world data, but not without significant tweaks on the structure of the subordinator. As the saying goes, if you add enough parameters, you can fit pretty much everything. The mathematical elegance faded, and it became difficult to obtain insights. More importantly, if the distribution itself is too complicated (and/or hard to calculate), the option pricing model adds even more complexity to it, and the end result will be hopelessly difficult to understand.

Based on these experiences and observations, the proposed distribution must be simple and elegant, ideally a two-parameter family, with several single-parameter sub-families, in order to pass the smoke test, so to speak. In this paper, I propose a novel probability distribution based on elliptic curves. This distribution is constructed from a very simple starting point - the Weierstrass equation in the depressed cubic polynomial form[2] ,

$$y^2 = x^3 + Ax + B. \tag{1.1}$$

Its elegant origin makes it analytically tractable, intuitive and easy to calculate. It can generate many interesting shapes with a wide ranging kurtosis from 2.3 to as high as 35. It is asymptotic to the normal distribution and the Laplace distribution with the standard cusp distribution in the center. Its flexibility is not less than GHYP. It fits well to the historical daily log-return distributions of currencies, commodities, Treasury yields, VIX, and DJIA. The new concept of elliptic tail, PDF $\sim \exp\left(-x^{2/3}\right)$, that captures the outward-bending tails nicely (See Panel (2) of Figure 2.1), sets it apart from the tail behavior of GHYP. Hopefully with these nice features, it can become a useful tool for the financial professionals to describe the kind of data they encounter everyday.

[2]See Section III.1, p.45 of Silverman 2008[17]. Or see Tate 1974[16].

The volatility smile (or skew/smirk) is a challenging area in quantitative finance (See Gatheral 2006 [7] and Bergomi's "Smile Dynamics" series[2]), an aeea that I am interested in solving with elliptic distribution. Many believes the smile is caused by fat-tailed distribution (and/or jump process), yet up to date, there is no suitable fat-tailed distribution that can explain the observed data. The dynamics of the volatility surface continues to be an intensive research topic in quantitative finance. My belief is that the prices of an option chain must be related to the distribution of its underlying asset. If the elliptic distribution is believed to be the "correct" distribution for the underlying, it must be capable of explaining the volatility structure of the options built on top of it. In the subsequent papers that follow this, I'd also like to address this topic.

The past two decades have witnessed tremendous progress in the study of elliptic curves. It is related to an impressive number of ideas and problems in other parts of mathematics, such as number theory, group theory, modular form, and cryptography. The famous Last Theorem of Fermat, for example, was proven by Wiles [19] as corollary to a deep, but beautiful theorem about elliptic curves. Many techniques, methods, and results from the study of elliptic curves have been used to other areas of mathematics, sometimes with astonishing consequences. In recent years, Elliptic Curve Cryptography (ECC) is found to be more secure compared to the traditional RSA cryptography of the same key sizes. Is it possible that the fat-tail nature of financial markets is also deeply rooted in elliptic curves? If the answer is positive, then our understanding of economics and finance will be profoundly impacted.

I use the acronym ECD for the elliptic distribution in this paper. Almost every formula developed in this paper has been validated rigorously by `GNU Maxima`, and tested in `R` with hundreds of test cases. Calculations and charts are generated by the `ecd` package in CRAN. Feedback and collaboration are welcomed.

2. The Development of ECD

2.1. Basic Notations

The basic notations are defined in this section. Additional investigations about the distribution will be elaborated in subsequent sections. Our elliptic curves take the following form:

$$\frac{(x-\mu)^2}{\sigma^2} = -y^3 - \left(\gamma + \beta\left(\frac{x-\mu}{\sigma}\right)\right) y + \alpha. \tag{2.1}$$

It has been rearranged, compared to Eq. (1.1), in order to make it suitable in the context of a probability distribution. First, x and y are swapped. Second, y is inverted to $-y$. Third, the skew term, βxy, is added. Fourth, x is enriched via $x \mapsto \frac{x-\mu}{\sigma}$, where the location parameter μ and the volatility parameter σ are added to manage the location and scale. The structure of $\frac{x-\mu}{\sigma}$ makes the distribution qualified as a location-scale distribution family. The location

parameter μ is a proxy of the mean log-return. In the option pricing model, it is fixed by the "risk-neutral" condition[3].

In most cases of theoretical study, it is simpler to set $\mu = 0, \sigma = 1$. Alternatively, by setting $\hat{x} = \frac{x-\mu}{\sigma}$, we arrive at the so-called "unit distribution", in which the curve becomes

$$\hat{x}^2 = -y^3 - (\gamma + \beta\,\hat{x})\,y + \alpha. \tag{2.2}$$

The β term is the skewness parameter. It is a new element to our elliptic curves. It is the coefficient of the $\hat{x}y$ term. When β=0, the distribution is symmetric; otherwise, it is asymmetric. Since most financial time series have some degrees of skewness, we need β to fit the real-world data. In Section 3.6, I will show that β enters the elliptic distribution elegantly (it is not a hack). However, for theoretical study of variance and kurtosis, it is often more convenient to study the symmetric distribution.

There are places where y has multiple roots[4]. In ECD, this issue is addressed by requiring that the elliptic curve $y(x)$ is constructed from the smallest real root of Eq. (2.1) for each $x \in [-\infty, +\infty]$. This leads to the definition of the probability density functon (PDF) of an elliptic distribution,

$$P(x) = \frac{1}{C} e^{y(x)}, \text{ where } C = \int_{-\infty}^{\infty} e^{y(x)} dx \tag{2.3}$$

and $y(x)$ is parametrized by $y(x; \alpha, \gamma, \sigma, \beta, \mu)$. C is the normalization constant to maintain the unity of the density function, $\int_{-\infty}^{\infty} P(x)\,dx = 1$. The CDF is obviously $\Phi(x) = E\left[\mathbf{1}_{\{X \le x\}}\right] = \int_{-\infty}^{x} P(x)dx$; while at times it may be more suitable to study the tail behavior with the complementary CDF (CCDF), $1 - \Phi(x) = \int_{x}^{\infty} P(x)dx$. As mentioned before, for theoretical study, it is convenient to set $\mu = 0, \sigma = 1$ and β=0, which reduces the parametrization to the simple form of $y(x; \alpha, \gamma)$. In such cases, we can use the elegant notation $\texttt{ECD}(\alpha, \gamma)$ to represent the distribution.

The slope of log PDF carries important information about the distribution, which is (assuming $\mu = 0, \sigma = 1$)

$$\frac{d}{dx} \log P(x) = \frac{dy}{dx} = -\frac{2x + \beta y}{3y^2 + \gamma + \beta x}. \tag{2.4}$$

Since $y(x) = \log P(x) + C$, Eq. (2.4) is basically the ODE form of $P(x)$, albeit a very complicated one, by which $P(x)$ can be solved, analytically or numerically, with boundary conditions of a known C, and/or a known $P(x = \mu)$, and $P(\pm\infty) \to -\infty$.

Furthermore, the curvature of log PDF is the derivative of the slope: d^2y/dx^2. The typical shape of an elliptic distribution is shown in Figure 2.1. From Eq. (2.4), we can define the "elliptic points", x_e, as where $y(x)$ changes from concave

[3] The detail will be described in the subsequent paper.

[4] See the chart on catelog of elliptic curves at https://en.wikipedia.org/wiki/Elliptic_curve

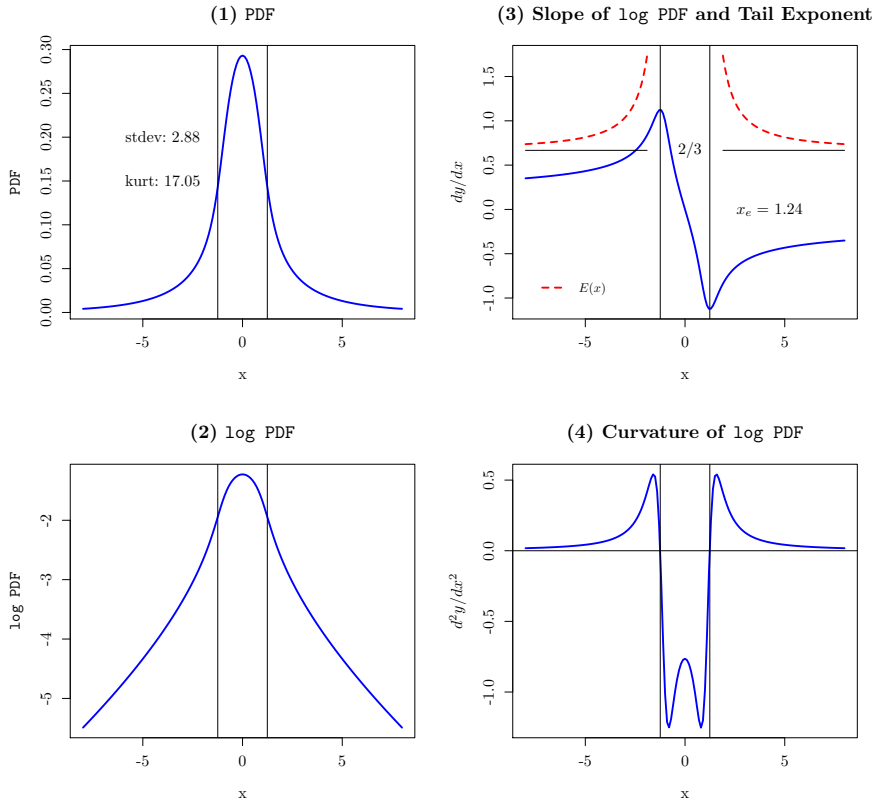

FIG 2.1. *The shape of a typical elliptic distribution. The PDF, log PDF, slope, and curvature of* ECD$(1, 2)$. *In Panel (2), notice the curve is bending outward in the tails of the log PDF. In Panel (3), the ellipticity,* $x_e = 1.24$, *is marked by two black vertical lines; and the convergence of tail exponent* $E(x)$ *to 2/3 is drawn in red dash lines.*

to convex. This occurs at the inflection points when the curvature crosses zero,

$$\frac{d^2y}{dx^2}(x_e) = 0. \tag{2.5}$$

There are always a positive solution x_e^+ and a negative solution x_e^- for x_e. The region between them, $x \in [x_e^-, x_e^+]$, looks like a normal distribution which is concave; while the regions outside are convex and eventually approach the "elliptic tails" $y(x) \sim -x^{2/3}$ asymptotically as $x \to \pm\infty$. (See Section 2.3 for more detail.) From the analyses of several financial time series in Section 8, we will see that the region $x \in [x_e^-, x_e^+]$ is very similar to the standard deviation. Therefore, we can define a new quantity called "ellipticity" that describes the proximity where elliptic influence begins to dominate,

$$\text{ellipticity} = \frac{x_e^+ - x_e^-}{2}. \tag{2.6}$$

Since the PDF is the exponential of $y(x)$, the convex nature of the elliptic tails is also called "logarithmically convex"[5]. Logarithmical convexity and elliptic tails are the main differentiators between the elliptic distribution and the normal distribution since the later is always concave and never turns convex in $y(x) \sim -x^2$ for any x. For a hyperbolic tail, $y(x) \sim -x$, there is also no convexity since $d^2y/dx^2 \to 0$ as $|x| \to \infty$. We can define the tail exponent as

$$E(x) = \frac{d\log(-y)}{d\log(x)} = \frac{x}{y}\frac{dy}{dx}, \text{ where } |x| \geq x_e. \tag{2.7}$$

This is a measurement of how fast the curve is converging to the expected tail exponent. When $y(x)$ is approaching the elliptic tail -$x^{2/3}$, $E(x)$ is approaching $\frac{2}{3}$. This is illustrated in Panel (3) of Figure 2.1. For a normal distribution, $E(x)$ is 2. For a hyperbolic tail, $E(x)$ is 1.

The standard measures of elliptic curves, the discriminant Δ and the j-invariant (See p. 45 of Silverman 2008[17]) can be defined in the symmetric case ($\sigma = 1$ and β=0),

$$\Delta(\alpha, \gamma) = -16(4\gamma^3 + 27\alpha^2), \tag{2.8}$$

$$j(\alpha, \gamma) = -1728\frac{(4\gamma)^3}{\Delta(\alpha, \gamma)}. \tag{2.9}$$

There are several special situations related to them, as shown in Figure 2.2. First, the zero discriminant condition defines a line in the 4th quadrant of (α, γ) plane, where the elliptic curves are singular (at $x = \mu$). This line is called "the critical line", which is defined as

$$\gamma_c(\alpha) = -\left(\frac{27}{4}\alpha^2\right)^{\frac{1}{3}}, \text{ where } \alpha \geq 0. \tag{2.10}$$

[5]See wikipedia at http://en.wikipedia.org/wiki/Logarithmically_convex_function.

The distributions on this line form a special single-parameter sub-family called "the cusp distribution", which starts with the standard cusp distribution at $\alpha = \gamma = 0$, and approaches the asymptotic limit of Laplace distribution at $\alpha \to \infty$. Depending on the application, the singular points may or may not cause you problems. You should take a note of it.

Second, the zero j-invariant condition defines a horizontal line on the (α, γ) plane, that is, $L_{j0} = \{\gamma = 0, \alpha \in [-\infty, +\infty]\}$. The positive side of the $j = 0$ line lies the important high-kurtosis distribution family. Third, the $j = 1728$ condition defines a vertical line on the (α, γ) plane, that is, $L_{j1728} = \{\alpha = 0, \gamma \in [-\infty, +\infty]\}$. The positive side of the $j = 1728$ line lies the most smooth asymptotic path to the normal distribution, while the negative side produces the distribution family closely resembling the hyperbolic distribution. Fourth, the region below the $j = 0$ line and above the critical line, $\gamma > \gamma_c$, should be excluded from the distribution family because those curves have jumps. These are very important lines. They separate the (α, γ) parameter space into different regions. Figure 2.2 illustrates these regions.

In the rest of the paper, when we draw the contours on the (α, γ) plane, we prefer to adjust γ by the discriminant so that the critical line is a straight 45° line. This allows us to transform the (α, γ) plane from the Cartesian coordinate to the polar coordinate, which is mathematically more powerful in some respects. Polar coordinate also gives better visual insight to the contours. Therefore, the discriminant adjusted γ, denoted as γ_Δ, is defined here as

$$\gamma_\Delta = \text{sgn}(\gamma) \left(\frac{4}{27} |\gamma|^3 \right)^{\frac{1}{2}}. \tag{2.11}$$

2.2. Quantile Function

The quantile function, $w(u) : \inf\{x \in R,\, u \le \Phi(x)\}$, is the inverse of the CDF, $\Phi(x)$. It is an essential function for QQ-plot. The `ecd` package implements the quantile function by ways of numerical sampling and interpolation. Since $\Phi(x)$ is a smooth function of x most of the time, and it is fairly fast to calculate $\Phi(x)$ by integrating via quadpack, numerical sampling is performed on the CDF segment by segment. These segments are units of stdev, up to 4 stdev. Special handling is required around elliptic points (See Eq. (2.5)) because there are kinks when kurtosis is high. A separate numerical sampling is employed for the tail regions, typically defined as $\Phi(x) < 1\%$ and $\Phi(x) > 99\%$. The procedure exploits the asymptotic behavior of the elliptic tails so that it can be precise up to, say, 40 stdev. The tail behavior is studied in depth in Section 3.

On the other hand, Steinbrecher and Shaw has shown $w(u)$ satisfies the following ODE [18],

$$\frac{d^2 w}{du^2} = -\frac{d\, y(w)}{dw} \left(\frac{dw}{du} \right)^2, \tag{2.12}$$

where $\frac{d\, y(w)}{dw}$ is the slope from Eq. (2.4) expressed in terms of w (that is, x). This representation opens up possibility of numerical solution techniques that are

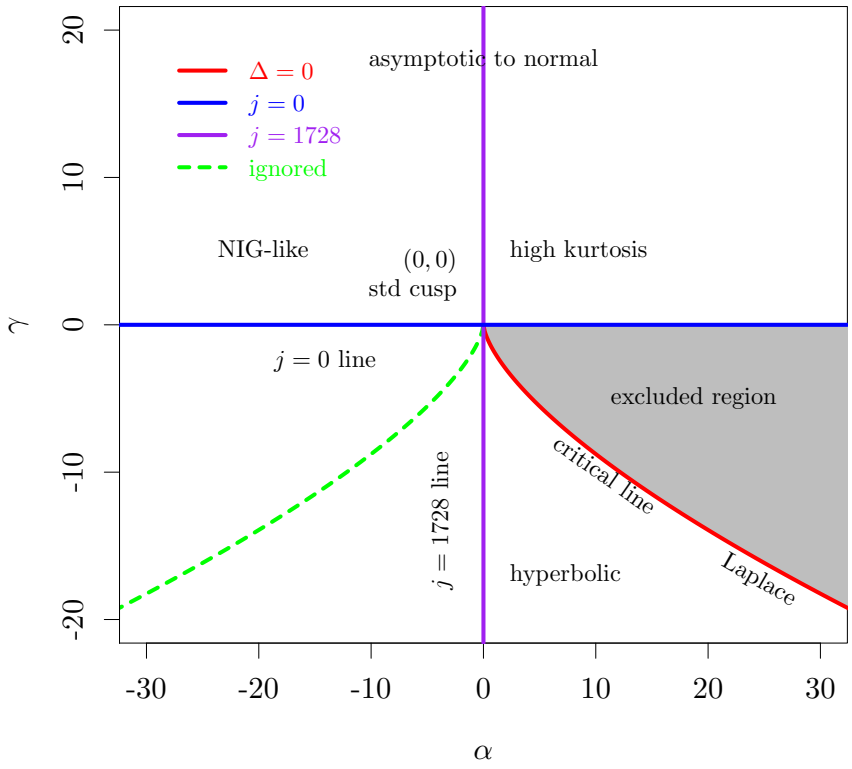

FIG 2.2. *The (α, γ) parameter space. The critical line ($\Delta = 0$) is drawn in red. The excluded region is in grey. The horizontal $j = 0$ line is in blue, and the vertical $j = 1728$ line is in purple.*

complementary to numerical interpolation methods based on the CDF. Further work is needed to explore the optimal routes for solving such ODE, and to establish whether a power series method is viable.

2.3. Asymptotic Behavior at $x = \pm\infty$

The origin O of the elliptic curves is at $x = \pm\infty$ and $y \to -\infty$, where the PDF vanishes, $P(x = \pm\infty) \to 0$. It is obvious from Eq. (2.1), the distribution has an elegant asymptotic behavior,

$$\log P(x)\big|_{x\to\pm\infty} \sim -(\frac{x}{\sigma})^{\frac{2}{3}}. \tag{2.13}$$

The $\frac{2}{3}$ exponent is important. It is the marker of the elliptic tails. The tails with $\frac{2}{3}$ exponent are fatter than that of the GHYP (and/or Laplace distributions). The asymptotic behavior of the GHYP family is driven by that of the modified Bessel function of the third kind (See Appendix B of [4]),

$$\log P_{ghyp}(x) \sim -(\frac{x}{\sigma}) - \frac{1}{2}\log(\frac{x}{\sigma}). \tag{2.14}$$

However, this doesn't mean ECD is drastically different from GHYP. For most levels of kurtosis below 15, ECD can produce similar log PDF as HGYP. See Section 7 for more detail.

Also notice that Eq. (2.13) has the same form as Eq. (3.2) and none of α, γ, β shows up in Eq. (2.13). It indicates all elliptic distributions converge to the same tail behavior as the standard cusp distribution, which will be examined in Section 3.

2.4. Symmetric When $\beta = 0$

When $\beta = 0$, the distribution is symmetric. All the odd moments are zero. Figure 2.3 demonstrates the various shapes of symmetric distribution and how the shape of log PDF changes with different α and γ . They are chosen to reflect the range of kurtosis observed in typical financial applications. ECD can generate distributions from zero excess kurtosis to very high kurtosis, (as high as 35). More numeric results will be presented later when kurtosis is studied in greater details.

2.5. Asymmetric When $\beta \neq 0$

When β is not zero, the distribution is skewed. All the odd moments are non-zero. The skewness is positive when $\beta > 0$ and negative when $\beta < 0$. The mean is also shifted accordingly. The skewness can be quite large as demonstrated in Figure 2.4.

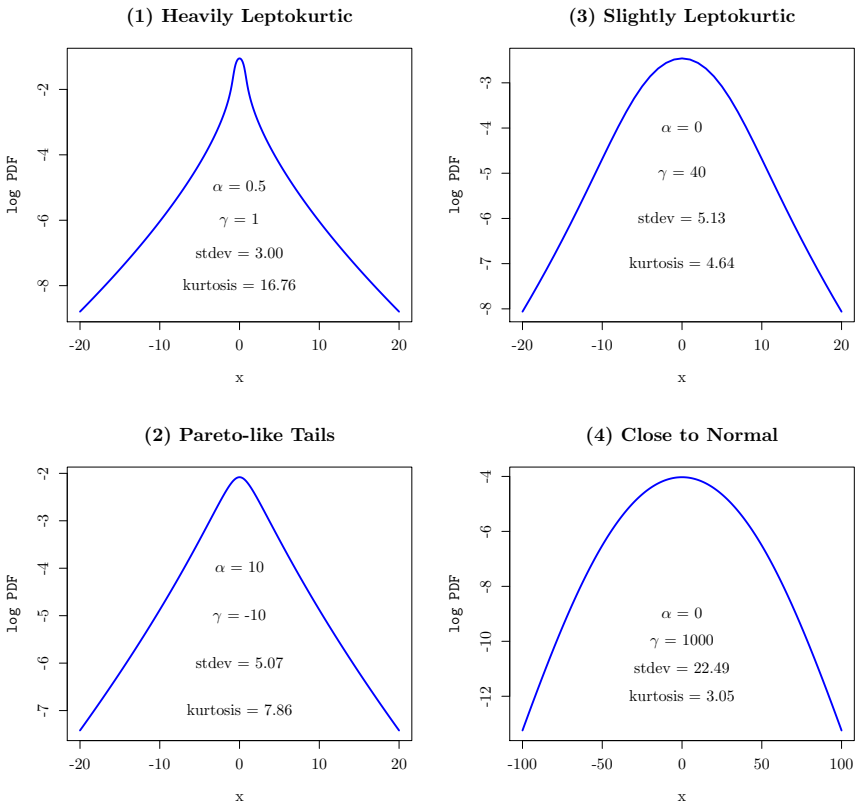

FIG 2.3. *Demo of* log *PDF*(x) *of various symmetric shapes. Notice that kurtosis can become very large. Ability to generate large kurtosis is a major strength of ECD.*

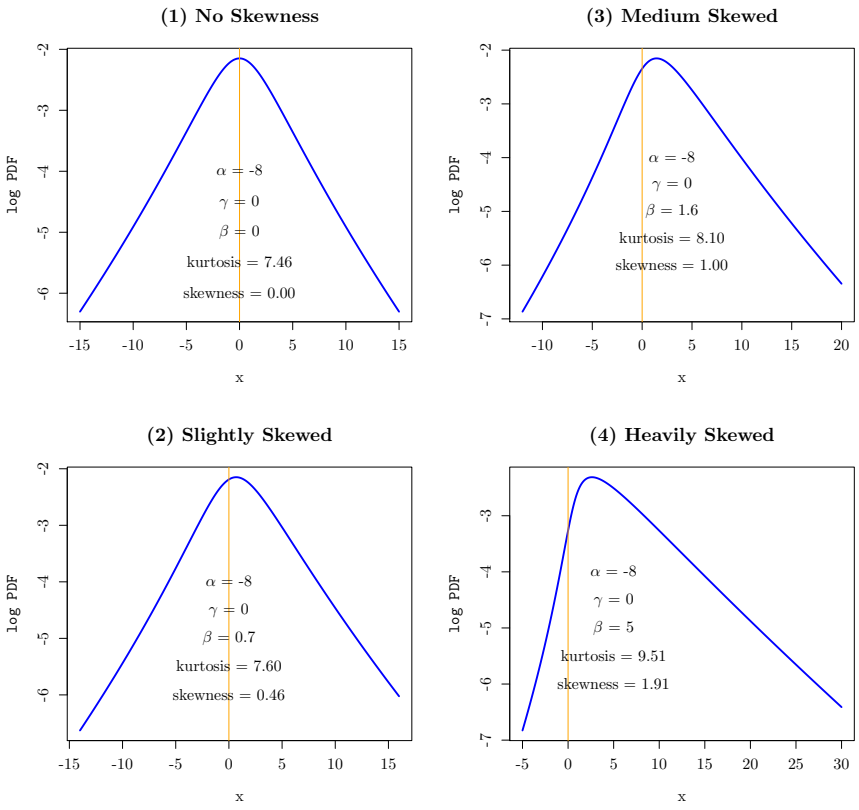

FIG 2.4. *Demo of* log *PDF*(x) *of various asymmetric shapes. Notice that the shape can become lopsided, largely due to the flexibility of the* βxy *term.*

2.6. Analytic Form of $y(x;\alpha,\gamma)$

From the solutions of a depressed cubic polynomial, the analytic form of $y(x;\alpha,\gamma)$ can be derived as following,

$$\begin{array}{l} f_1 = \frac{1}{2}(x^2-\alpha) \\ f_{2C} = 27x^4 - 54\alpha x^2 + 4\gamma^3 + 27\alpha^2 \\ f_2 = \frac{1}{2\sqrt{27}}\sqrt{f_{2C}} \\ \text{sgn} = -1 \text{ if } f_{2C} \geq 0 \text{ and } |f_1| < f_2 \text{ ; else } 1 \\ f_3 = (\text{sgn} \times (f_1 - f_2))^{1/3} \\ y(x;\alpha,\gamma) = \text{sgn} \times (-f_3 + \frac{\gamma}{3f_3}). \end{array} \tag{2.15}$$

This formula works nearly everywhere except when $\gamma = 0$, where f_3 becomes zero and so is the j-invariant. A special treatment is needed - however, it is just $y(x) = \text{sgn}(\alpha - x^2)\left|\alpha - x^2\right|^{1/3}$. Also notice that there are cases where f_2 can be complex ($f_{2C} < 0$) even though $y(x;\alpha,\gamma)$ always comes out real numbers; therefore, certain care must be given when you convert between complex numbers and real numbers in your program. In those cases, you can avoid complex number calculation with the trigonometric solution.

2.7. Chebyshev Trigonometry

Another approach of solving a depressed cubic polynomial is based on Chebyshev trigonometry[6]. In order to capture the broadest solution, let's extend α,γ to $\tilde{\alpha} = \alpha - \frac{(x-\mu)^2}{\sigma^2}$ and $\tilde{\gamma} = \gamma + \beta\left(\frac{x-\mu}{\sigma}\right)$. And the extended discriminant is $\tilde{\Delta} = -16(4\tilde{\gamma}^3 + 27\tilde{\alpha}^2)$. That is, nonzero x can be viewed as moderating (α,γ) to a different location $(\tilde{\alpha},\tilde{\gamma})$, but we still have the same form of a depressed cubic,

$$y^3 + \tilde{\gamma}y - \tilde{\alpha} = 0. \tag{2.16}$$

There are four scenarios:

(1) When $\tilde{\gamma} = 0$, that is, on the $j = 0$ line, the solution is simply $y(x) = \text{sgn}(\tilde{\alpha})\left|\tilde{\alpha}\right|^{1/3}$. the trigonometric system can't handle such scenario (because $V \to \infty$).

(2) When $\tilde{\Delta} < 0$ and $\tilde{\gamma} > 0$, that is, the entire upper $(\tilde{\alpha},\tilde{\gamma})$ plane, we have

$$y(x) = -2\sqrt{\frac{\tilde{\gamma}}{3}}\sinh\left(\frac{1}{3}\text{arcsinh}(V)\right), \text{ where } V = -\frac{3\tilde{\alpha}}{2\tilde{\gamma}}\sqrt{\frac{3}{\tilde{\gamma}}}. \tag{2.17}$$

(3) When $\tilde{\Delta} < 0$ and $\tilde{\gamma} < 0$, that is, the lower left outer region on the $(\tilde{\alpha},\tilde{\gamma})$ plane, we have

$$y(x) = 2\,\text{sgn}(\tilde{\alpha})\sqrt{\frac{-\tilde{\gamma}}{3}}\cosh\left(\frac{1}{3}\text{arccosh}(V)\right), \text{ where } V = -\frac{3\left|\tilde{\alpha}\right|}{2\tilde{\gamma}}\sqrt{\frac{-3}{\tilde{\gamma}}}. \tag{2.18}$$

[6] See http://en.wikipedia.org/wiki/Cubic_function#Trigonometric_.28and_hyperbolic.29_method

(4) When $\tilde{\Delta} \geq 0$ (which implies $\tilde{\gamma} < 0$), that is, the lower central region on the $(\tilde{\alpha}, \tilde{\gamma})$ plane, we have

$$y(x) = -2\sqrt{\frac{-\tilde{\gamma}}{3}} \cos\left(\frac{1}{3}\arccos(V)\right), \text{ where } V = \frac{3\tilde{\alpha}}{2\tilde{\gamma}}\sqrt{\frac{-3}{\tilde{\gamma}}}. \tag{2.19}$$

Notice that, in scenario (4), since the cos(...) term is always between -1 and 1, the scale of $y(x)$ is controlled by $-|\tilde{\gamma}|^{1/2}$. What $\tilde{\alpha}$ is doesn't actually matter that much to the scale of $y(x)$. In addition, it should be pointed out that scenarios (3) and (4) are the same due to the analytic continutity on the complex plane between trigonometric and hyperbolic functions. There is a relation rarely mentioned in textbook,

$$\cos\left(\frac{1}{3}\arccos(z^2-1)\right) = \cosh\left(\frac{1}{3}\operatorname{arccosh}(z^2-1)\right), \text{ where } z \text{ is real.} \tag{2.20}$$

The beauty of the trigonometric solution is that, in all four scenarios, no complex number calculation is involved, especially taking cubic roots, which is somewhat convoluted (I am referring to f_3 in Eqs. (2.15). In addition, for ECD, the output has to yield the smallest real root for $y(x)$ in a definite way. These trigonometric equations are particularly good for such purpose. There is no ambiguity.

3. The Cusp Distribution

In this section, various types of cusp distribution are studied. The cusp distributions are on the critical line, where the discriminant is zero with $\alpha \geq 0$ and $\gamma \leq 0$. On one end of the critical line where $\alpha = 0$, we have the "standard cusp distribution", which is symbolic of the tails of all elliptic distributions; on the other end of the critical line where α is large, the cusp distributions evolve into the Laplace distribution. These are very interesting aspects associated with the critical line.

3.1. The Standard Cusp Distribution

The standard cusp distribution $\texttt{ECD}(0,0)$ is one of the few distributions in the family that have known analytic solutions. It has a very special place in ECD because it is representative of the general shape of the tails for the entire elliptic distribution family, following the result of the expansion around the origin O. I use the Λ subscript to represent it here. This distribution is the result of $\alpha = \gamma = \beta = \mu = 0$ and $\sigma = 1$. Its elliptic curve is simply

$$y_\Lambda(x) = -x^{2/3}. \tag{3.1}$$

This curve is historically called "semicubical parabola" [7]. $C_\Lambda = \int_{-\infty}^{\infty} e^{y_\Lambda(x)}\, dx$ has a closed form $3\sqrt{\pi}/2$. Therefore, its PDF is

$$P_\Lambda(x) = \frac{2}{3\sqrt{\pi}}\, e^{-x^{2/3}}, \tag{3.2}$$

which is shown in Panel (1) of Figure 3.2. It has a cusp at $x = 0$. The PDF also satisfies the following differential equation

$$x^{\frac{1}{3}}\, \frac{dP_\Lambda(x)}{dx} + \frac{2}{3} P_\Lambda(x) = 0, \tag{3.3}$$

which can be solved if the boundary condition is given, $P_\Lambda(x=0) = \frac{2}{3\sqrt{\pi}}$.

3.2. Moments and Kurtosis of Standard Cusp

All the central moments can be calculated via the change of variable $x = z^3$. The odd moments are zero, and the even moments are

$$\mu_\Lambda(n) = \int_{-\infty}^{\infty} x^n\, P_\Lambda(x)\, dx = \frac{4}{\sqrt{\pi}} \int_0^{\infty} z^{3n+2}\, e^{-z^2}\, dz. \tag{3.4}$$

And the closed form for the even moments is

$$\mu_\Lambda(n) = \frac{2}{\sqrt{\pi}} \Gamma(\frac{3}{2}(n+1)) = \frac{(3n+1)!!}{2^{3n/2}}. \tag{3.5}$$

Specifically, the second moment is $\mu_\Lambda(2) = 105/8$ and the fourth moment is $\mu_\Lambda(4) = 135135/64$. Therefore, its kurtosis is

$$K_\Lambda = \frac{\mu_\Lambda(4)}{\mu_\Lambda(2)^2} = \frac{429}{35} \approx 12.257, \tag{3.6}$$

which is called the "standard kurtosis" of the elliptic distribution family. Kurtosis higher than this level is called high kurtosis. For people familiar with financial time series, the kurtoses of the daily log-returns of many stock market indices and commodities are in this range. So this distribution should have an important role for financial application. Another important reference of kurtosis below this level is 6 from the Laplace distribution. Therefore, between 6 and 12.257 is called medium kurtosis.

Since all the moments exist, the characteristic function can be composed as[8]

$$\varphi_\Lambda(t;\mu,\sigma) = E\left[e^{itX}\right] = e^{it\mu}\left[1 + \sum_{n=2,4,\ldots}^{\infty} (-1)^{\frac{n}{2}}\, \frac{\mu_\Lambda(n)\,(\sigma t)^n}{n!}\right]. \tag{3.7}$$

[7] http://en.wikipedia.org/wiki/Semicubical_parabola

[8] It is preferred to have μ and σ in the characteristic function, in order to compare it with what's on Wikipedia. See examples in http://en.wikipedia.org/wiki/Characteristic_function_(probability_theory)#Examples

When $\mu = 0$, the characteristic function is a real-value transcendental function. Likewise, the moment generating function (MGF) is composed by the sum of moments,

$$M_\Lambda(t;\mu,\sigma) = E\left[e^{tX}\right] = e^{t\mu}\left[1 + \sum_{n=2,4,\ldots}^{\infty} \frac{\mu_\Lambda(n)\,(\sigma t)^n}{n!}\right]. \tag{3.8}$$

However, strictly speaking, MGF diverges due to the elliptic tail exponent $\frac{2}{3}$ less than 1. MGF is the crucial element in the option pricing model. The issue of MGF explosion and tail truncation will be studied in great details in the subsequent paper on option pricing model.

3.3. The CDF and Tail Behavior of Standard Cusp

Since the standard cusp doesn't depend on any parametrization, it is a good place to explore the tail behavior, which is necessary for the implementation of the quantile function. The complimentary CDF (CCDF) of the cusp distribution allows us to study the right tail without worrying about the sign of x. It is defined as

$$\texttt{CCDF} = 1 - \Phi_\Lambda(x) = \frac{2}{3\sqrt{\pi}} \int_x^\infty e^{-x^{2/3}}\, dx, \text{ where } x > 0. \tag{3.9}$$

With a change of variable, $t = x^{2/3}$, The CCDF is transformed to the equivalent of a gamma distribution with shape $\frac{3}{2}$ and scale 1, [9]

$$1 - \Phi_\Lambda(x) = \frac{1}{2} \int_{x^{2/3}}^\infty \Gamma(x; \frac{3}{2}, 1)\, dx, \text{ where } x > 0. \tag{3.10}$$

Therefore, the tail behavior of the ECD is closely related to that of the gamma distribution. This integral can be expressed in terms of the complimentary error function (erfc),

$$1 - \Phi_\Lambda(x) = \frac{1}{\sqrt{\pi}}\, x^{\frac{1}{3}}\, e^{-x^{2/3}} + \frac{1}{2}\text{erfc}\left(x^{\frac{1}{3}}\right). \tag{3.11}$$

The tail region is defined as $\Phi(x) < 1\%$ on the left tail; or $1-\Phi(x) < 1\%$ on the right tail. This is approximately 4-5 standard deviations away from the mean. (One standard deviation is 3.62 for standard cusp.) The asymptotic expansion of incomplete gamma function (See §8.11(i) of NIST DLMF) yields

$$1 - \Phi_\Lambda(x) = \frac{1}{\sqrt{\pi}}\, x^{\frac{1}{3}} e^{-x^{2/3}} \left(1 + \sum_{k=1}^{N-1} \frac{\Gamma(\frac{3}{2})}{\Gamma(\frac{3}{2} - k)\, x^{2k/3}} + R_N(x)\right), \text{ when } x \to \infty, \tag{3.12}$$

[9] Gamma distribution with scale one is also called "incomplete gamma function".

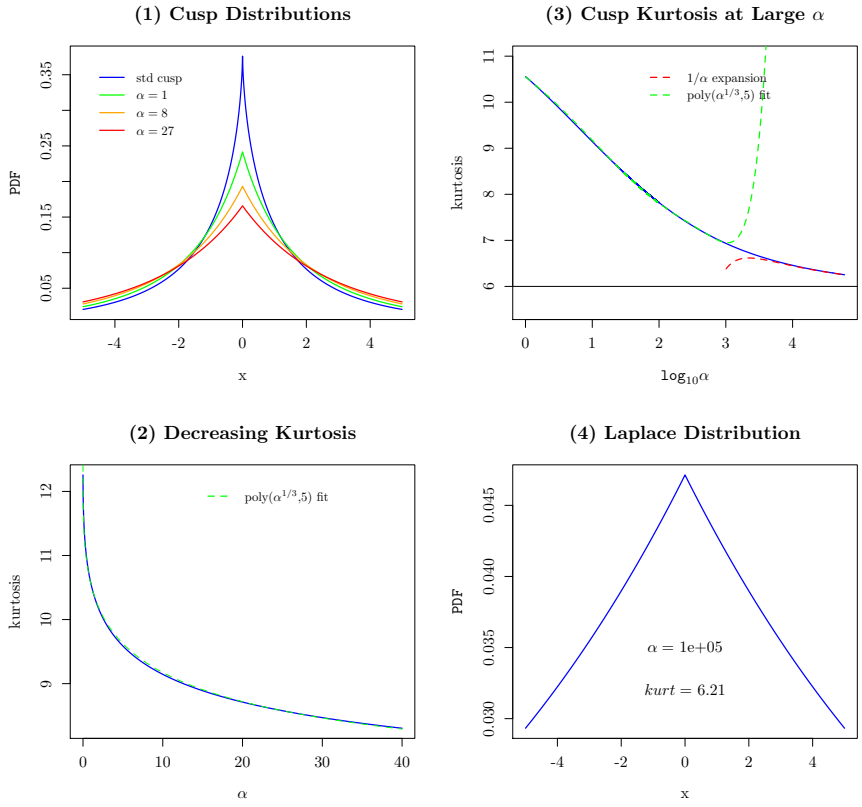

FIG 3.1. *The cusp distributions. (1) The* log *PDF*(x) *of a few cusp distributions. (2) The rapidly decreasing kurtosis of cusp distributions for small* α*. (3) The asymptotic behavior of kurtosis when* $\alpha \to \infty$*. (4) Illustration of a cusp distribution at* $\alpha = 10^5$ *where the kurtosis is 6.21, very close to the theoretical value of 6 for a Laplace distribution.*

where N is a large integer and $R_N(x)$ is the residual error term $O(x^{2N/3})$. Numerical test shows 10^{-4} precision for $x = -20, N = 10$. Taking the logarithm, the dominant tail behavior is

$$\log\left(1 - \Phi_\Lambda(x)\right) = -x^{\frac{2}{3}} + \frac{\log x}{3} - \frac{\log \pi}{2} + \log(1 + ...). \tag{3.13}$$

That is, the log CCDF decreases like $-x^{2/3}$ in the tails, which is then moderated by the $\frac{1}{3}\log|x|$ term. As shown in Panel (4) of Figure 3.2, the log CCDF is quasi-linear from 4 stdev to 30 stdev with a linear slope of ~ 0.62 per stdev. Numerically it can be fitted reasonably accurate by a polynomial of less than 6 orders. This fit gives us a precise mapping up to 30 stdev, which can be used to invert the CDF for the quantile function. This is how the quantile function is implemented in the `ecd` package. For financial applications, the probability of an event occurring one day in 100 years is approximately $\log \Phi(x) = -\log(250 \times 100) \sim -10$. This is the minimum precision required for the quantile function. The range shown in Panel (4) is twice of this level.

3.4. The Discriminant

The discriminant and j-invariant are two fundamental quantities associated with elliptic curves. Contour of discriminant on the (α, γ_Δ) plane is shown in Figure 3.3. On the upper plane, the discriminant is negative. The contour lines form many confocal circles due to

$$\Delta(\alpha, \gamma_\Delta) = -432\left(\mathrm{sgn}(\gamma_\Delta)\,\gamma_\Delta^2 + \alpha^2\right). \tag{3.14}$$

Elliptic curves in this region have single real roots. As the discriminant approaches $-\infty$, the distribution is asymptotic to a normal distribution.

On the lower plane, the critical line ($\Delta = 0$) is where many contour lines converge on the right side of the chart. In the region of positive discriminant, γ_Δ is negative and the contour lines are hyperbolic. Elliptic curves in this region can have three real roots in parts of the curve. The distributions have wedge-like shapes (in log PDF) similar to that of hyperbolic distributions. The closer to the critical line, the sharper the wedge is at the peak.

3.5. General Cusp Distributions and Laplace Distribution

In this section, we will discuss some properties of the general "cusp distribution" family. The standard cusp distribution is a special case of this family. All the cusp distributions are located on the critical line defined by Eq. (2.10) in the fourth quadrant of the (α, γ) plane[10]. They are singular at $x = 0$. This special group of distributions forms the first single-parameter sub-family of elliptic distribution.

[10] In the polar coordinate that follows, this is $\theta = \frac{7}{4}\pi, R \in [0, \infty)$.

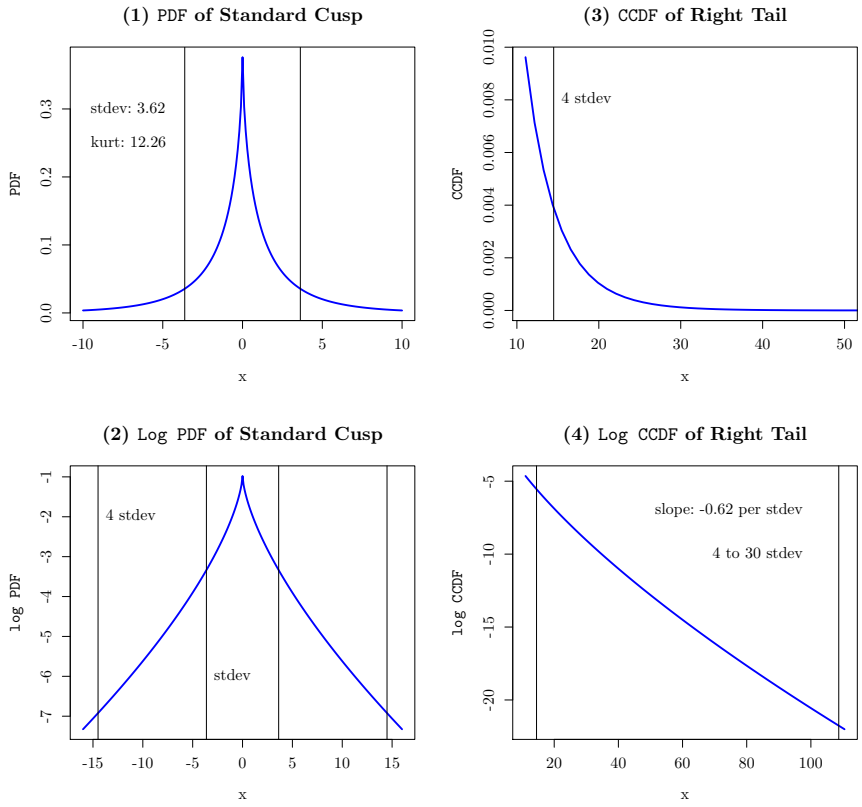

FIG 3.2. *The Shape and Tail Behavior of Standard Cusp. (1) PDF with vertical lines indicating one stdev. (2) Log PDF with vertical lines indicating one and 4 stdev. (3) CCDF of right tail with vertical line indicating 4 stdev. (4) Log CCDF of right tail from 4 to 30 stdev.*

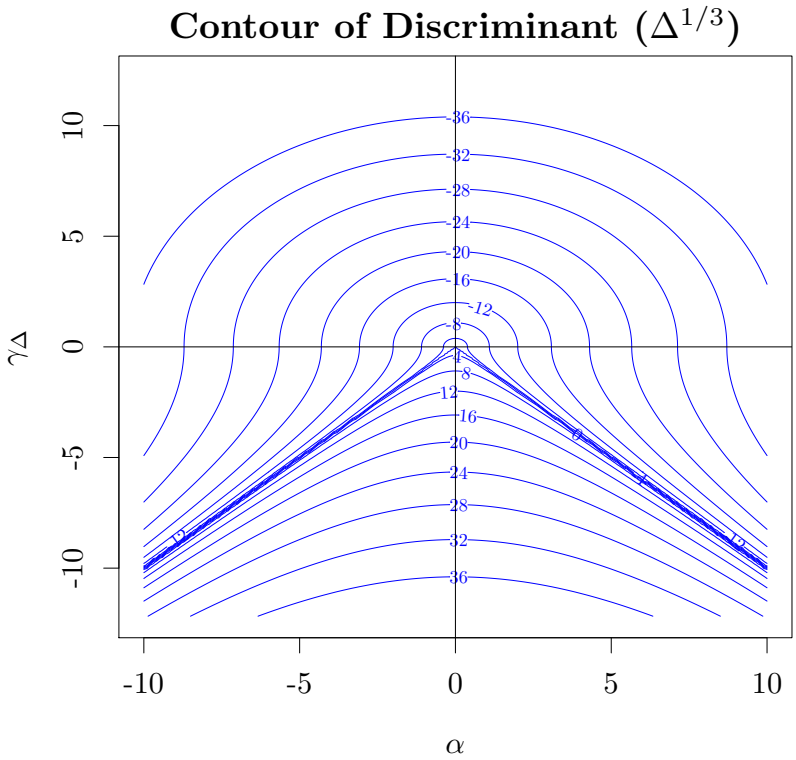

FIG 3.3. *Contour of Discriminant. The critical line ($\Delta = 0, \alpha > 0$) is where many contour lines converge on the right side of the chart.*

Here we illustrate a method to identify all the cusp distributions by examining the slope dy/dx. By definition, a cusp distribution has a singular point where the slope is undefined, that is, the denominator of dy/dx is zero,

$$3y^2 + \gamma + \beta x = 0 \tag{3.15}$$

This can be evaluated at the peak of the distribution, which is defined by the numerator of dy/dx being zero,

$$2x + \beta y = 0 \tag{3.16}$$

Putting them back to the Weierstrass equation, Eq. (2.1), we can solve the location of the peak and the relationship between the parameters. When $\beta = 0$, we first get the location of the peak, $x = 0$, from Eq. (3.16), Then we get the peak density $y^2(0) = -\gamma/3$, which implies $\gamma < 0$; and $y(0) = \frac{3}{2}\alpha/\gamma < 0$. Finally we get $\frac{27}{4}\alpha^2 = -\gamma^3$, which is identical to $\Delta = 0$. The asymmetric case can be solved numerically by following the same procedure. It is important to have a defined procedure to identify where the peak is since the peak area tends to be numerically unstable and requires special handling.

We prefer to specify each cusp distribution by a positive α. Based on $\Delta = 0$, Eq. (2.19) can be simplified quite well for the symmetric case. With the change of variable $z^2 = x^2/(\alpha\sigma^2)$, we have

$$y(z) = -(4\alpha)^{1/3} \cos\left(\frac{1}{3}\arccos\left(z^2 - 1\right)\right), \text{ where } z^2 \le 2. \tag{3.17}$$

Its Taylor expansion is (assuming $z > 0$)[11]

$$y(z) \approx -\left(\frac{\alpha}{2}\right)^{\frac{1}{3}}\left(1 + \frac{\sqrt{2}\,z}{\sqrt{3}} - \frac{z^2}{9} + \frac{5\,z^3}{2^{3/2}\,3^{7/2}} - \frac{4z^4}{243} + ...\right). \tag{3.18}$$

We see that the cusp shows up via the linear z term. We also have $y(0) = -(\alpha/2)^{1/3}$, that is, the peak is negative and moving lower as α increases. Panel (2) of Figure 3.1 shows the kurtosis decreases rapidly as α increases from zero. As you can see, the Laplace distribution, $\sim \exp(-|x|/b)$, is a special case of the general cusp distribution when z is small or α is very large ($\alpha \gg (x/\sigma)^2$). It should also be pointed out that this quasi-linear segment of the general cusp distributions only exists when $z^2 \le 2$, that is, $x^2 \le 2\alpha\sigma^2$. It disappears when $\alpha \to 0$.

The stable growth of an economic system requires a power law (Gabaix 2009[9]), which in turn requires a large linear segment by a large α. As α increases, the cusp distribution will approach the Laplace distribution asymptotically. Replacing z with x, we have

$$y(x) \approx -\left(\frac{\alpha}{2}\right)^{\frac{1}{3}} - \left(\frac{2}{\alpha}\right)^{\frac{1}{6}}\frac{x}{\sqrt{3}\sigma} + ..., \text{ when } \alpha \to \infty. \tag{3.19}$$

[11] Eq. (3.17) can cover the entire z-axis by extending to hyperbolic functions on the complex plane. But Taylor expansion can't be extended naturally because $z^2 = 2$ is singular.

The linear coefficient b is equal to $\sqrt[6]{\alpha}\sigma\,\frac{\sqrt{3}}{\sqrt[6]{2}}$. We see that σ is merged with α asymptotically in a peculiar way.

Next, we will explore the behavior of the kurtosis on the critical line. Panel (2) of Figure 3.1 shows the kurtosis drops rapidly for small α. Attempt of using Taylor expansion to analyze this decreasing curve has failed; however, it is found that the kurtosis curve can be fitted quite well empirically by a polynomial of $\alpha^{1/3}$. Within order of 5 [12], a reasonably high precision fit can be obtained as shown in green dash line of Panel (2) and (3). This fit can go as high as $\alpha \approx 3000$.

At this level of large α, we can study the variance and kurtosis asymptotically by the following Taylor expansion method. The dimensionless moments are rearranged as

$$\mu_n = C\int_0^\infty z^n\,\exp\left(y(z)-L(z)\right)e^{L(z)}dz,\ \text{where}\ L(z) = -\left(\frac{\alpha}{2}\right)^{\frac{1}{3}}\left(1+\frac{\sqrt{2}\,z}{\sqrt{3}}\right). \tag{3.20}$$

C is a constant that normalizes μ_0 to one, whose exact value is not needed here. Next, Taylor expansion on z is applied to exp(...) at $z=0$. The integral is then broken up to a sum of Laplace integrals. The kurtosis and variance at large α are expanded by $1/\alpha$. The results are

$$\frac{\text{var}}{\sigma^2} = \frac{\mu_2\alpha}{\mu_0} = 10 + 3\,2^{\frac{2}{3}}\,\alpha^{\frac{1}{3}} + \frac{13}{3\,2^{\frac{2}{3}}\,\alpha^{\frac{1}{3}}} - \frac{17\,2^{\frac{2}{3}}}{9\,\alpha^{\frac{2}{3}}} + \frac{445}{108\,\alpha} + ..., \tag{3.21}$$

$$\text{kurtosis} = \frac{\mu_4\mu_0}{\mu_2^2} = 6 + \frac{2^{\frac{10}{3}}}{\alpha^{\frac{1}{3}}} - \frac{2^{\frac{8}{3}}}{\alpha^{\frac{2}{3}}} - \frac{44}{3\,\alpha} + \frac{19\,2^{\frac{10}{3}}}{3\,\alpha^{\frac{4}{3}}} - \frac{257\,2^{\frac{2}{3}}}{3\,\alpha^{\frac{5}{3}}} - \frac{15140285}{27\,\alpha^2} + ... \tag{3.22}$$

The kurtosis 6 of a Laplace distribution appears naturally. The red dash line in Panel (3) of Figure 3.1 shows the result of Eq. (3.22), which is precise for $\alpha > 10^{3.5}$ as the kurtosis drops below 6.5. Panel (4) shows a cusp distribution at $\alpha = 10^5$. Using MPFR, the kurtosis 6.214 calculated numerically matches the sum in Eq. (3.22). The variance 231.1 also matches the sum in Eq. (3.21).

One interesting phenomena is that, because kurtosis stays above 6 on the critical line, it pushes the spirals of low kurtosis regime towards infinity, as shown in Figure 5.2. The contour lines are straight lines, almost parallel to the critical line, in the third and fourth quadrants.

3.6. Asymmetric Standard Cusp Distribution

When $\beta \neq 0$, the elliptic curve of the standard cusp distribution becomes asymmetric,

$$x^2 = -y^3 - \beta xy. \tag{3.23}$$

This curve also has analytic solution. The cusp is still located at the peak of the distribution, $y(0) = 0$; and $y(x) \leq 0$ everywhere. Notice that $\beta \mapsto -\beta$ is equivalent to $x \mapsto -x$. Therefore it is sufficient to study $\beta > 0$. Using the

[12] In R, it is [`a3<-a^(1/3); lm(k ~ poly(a3,5,raw=TRUE))`].

trigonometric solutions from Section 2.7, we have $\tilde{\alpha} = -x^2$, $\tilde{\gamma} = \beta x$, and $\tilde{\Delta} = -432x^3\,(x - x_0)$, where $x_0 = -\frac{4\beta^3}{27}$. With $V = \left|\frac{x}{x_0}\right|^{\frac{1}{2}}$ and $W = 2\left|\frac{\beta x}{3}\right|^{\frac{1}{2}}$, there are three scenarios:

(1) When $\tilde{\Delta} < 0$ and $\tilde{\gamma} \geq 0$ (that is, if $\beta > 0$, then $x \geq 0$), we have the solution of the right tail,

$$y^+(x,\beta) = -W\,\sinh\left(\frac{1}{3}\operatorname{arcsinh}(V)\right). \tag{3.24}$$

(2) When $\tilde{\Delta} < 0$ and $\tilde{\gamma} < 0$ (that is, if $\beta > 0$, then $x < x_0$), we have the solution of the left tail,

$$y^-(x,\beta) = -W\,\cosh\left(\frac{1}{3}\operatorname{arccosh}(V)\right). \tag{3.25}$$

(3) When $\tilde{\Delta} \geq 0$ and $\tilde{\gamma} < 0$ (that is, if $\beta > 0$, then $x_0 \leq x < 0$), we have

$$y^-_{(3)}(x,\beta) = -W\,\cos\left(\frac{1}{3}\arccos(V)\right). \tag{3.26}$$

According to analytic continuity in the complex plane, (2) and (3) can be viewed as identical. That is, $\cosh\left(\frac{1}{3}\operatorname{arccosh}(V)\right) = \cos\left(\frac{1}{3}\arccos(V)\right)$, if V is evaluated as a complex number, $V + 0i$. [13]

Due to the facts that $y^+(-x,\beta) = y^+(x,\beta)$ and $y^-(-x,\beta) = y^-(x,\beta)$, the moments can be expressed as integrals of $x \in [0,\infty)$,

$$\mu_n = \frac{1}{C}\int_0^\infty x^n\left(e^{y^+(x,\beta)} + (-1)^n e^{y^-(-x,\beta)}\right)dx, \tag{3.27}$$

where $C = \int_0^\infty \left(e^{y^+(x,\beta)} + e^{y^-(-x,\beta)}\right)dx$. The quantities, $B^\pm = \int_0^\infty e^{y^\pm(x,\beta)}dx$, follow a peculiar conservation law: $B^+ - B^- = \beta$. It is very precise numerically, but yet to have analytic proof. Such precision also indicates the elegance of the skew term βxy.

Generally speaking, the effect of β to the even moments and C is small. The first order effect is less than 5% for C and 2% for stdev. Eq. (3.27) appears quite complicated, but it produces rather simple numerical results for the first moment (μ_1) and skewness (S). The regression analysis in the range of $\beta \in (0.1, 0.8)$ yields

$$\begin{gathered}\mu_1 \approx 1.2353\,\beta,\\ S \approx 1.2113\,\beta,\\ \mu_1 \approx \operatorname{sgn}(S)\,(0.00178 + 0.9875\,|S| + 0.0405\,|S|^2).\end{gathered} \tag{3.28}$$

[13] However, for numerical implementation, if one chooses to use complex number, the issue of branch cut discontinuities in the complex plane for inverse hyperbolic/trignometric functions must be handled properly. $\cosh\left(\frac{1}{3}\operatorname{arccosh}(x)\right)$ is also associated with Chebychev polynomial $T_n(x)$ where $n = \frac{1}{3}$.

As you can see, both μ_1 and skewness are almost linear to β in this range, and μ_1 is almost equal to skewness, all within 3% error. The implication is that, in the high frequency data where $\sigma$is small and the negative skewness is obvious, such as in an equity index, the skewness and μ_1 are synonym to each other and will overtake the second moment.

4. The J-Invariant and the Polar Coordinate

The j-value in Eq. (2.9) is an invariant for elliptic curves, it does not depend on the choice of inflection (elliptic) point if the equation is written in the Weierstrass or Legendre form. Over an algebraically closed field, two elliptic curves are isomorphic if and only if they have the same j-invariant. The contour of j-invariant on the (α, γ_Δ) plane is presented in Panel (1) of Figure 4.1. The contour levels are straight lines when presented in the units of γ_Δ instead of γ. This inspires the definition of the polar coordinate (R, θ) as

$$R\, e^{i\theta} = \alpha + i\gamma_\Delta. \tag{4.1}$$

These straight lines are defined by the polar angle θ alone, with the j-invariant in the form of

$$\frac{j(\alpha,\gamma_\Delta)}{1728} = \frac{\mathrm{sgn}(\gamma_\Delta)\,\gamma_\Delta^2}{\mathrm{sgn}(\gamma_\Delta)\,\gamma_\Delta^2+\alpha^2} = \frac{j(\theta)}{1728} = \begin{cases} \sin^2\theta, & \gamma_\Delta > 0; \\ \left(1-\cot^2\theta\right)^{-1}, & \gamma_\Delta < 0. \end{cases} \tag{4.2}$$

Each θ forms a distribution sub-family that shares the mathematical property called "isomorphism".

4.1. Isomorphic Mapping

All the distributions with the same θ are isomorphic. The isomorphic mapping between two elliptic curves, say (α, γ) and (α', γ'), is stated as [14]

$$(\alpha', \gamma') = \left(\alpha\lambda^6, \gamma\lambda^4\right) \Leftrightarrow (x', y') = \left(x\lambda^3, y\lambda^2\right). \tag{4.3}$$

It follows that $R' = R\lambda^6$. So λ is not a new dimension. It is simply the scaling of R. This speaks for the elegance of the polar coordinate. Also on the right hand side, $x\lambda^3$ is just the scaling of the volatility parameter, $\sigma = \lambda^{-3}$. So the isomorphism can be carried out from within our existing parametrization framework. This is good.

Panel (2) of Figure 4.1 shows the j-invariant on the unit circle in the polar coordinate ($R = 1$), Eq. (4.2). Notice that $\pm\alpha$ map to the same j. This isomorphism is facilitated by $\lambda = i$. However, because λ is imaginary, the mapping alters the structure of the roots and the distribution is changed in a very fundamental way (remember we are taking the smallest real root). So I would not consider such isomorphism a valid mapping in the context of elliptic distribution.

[14] See Proposition 1.7, p.50 of Silverman 2008[17].

The isomorphism allows us to "map out" all other distributions from the ones on the unit circle. Assume we have $y_0(x)$ for $\texttt{ECD}(\alpha_0, \gamma_0)$ on the unit circle at angle θ, that is, $\alpha_0 = \cos\theta, \gamma_{\Delta 0} = \sin\theta$. Now there is another elliptic distribution with $R \neq 1$ on the same j-invariant line. It can be parametrized as $\texttt{ECD}(\alpha_0 R, \gamma_0 R^{2/3})$. And its elliptic curve $y(x)$ can be constructed according to isomorphism,

$$y(x) = R^{\frac{1}{3}}\, y_0\left(xR^{-\frac{1}{2}}\right) = R^{\frac{1}{3}}\, y_0\left(x;\, \sigma_0 = \sqrt{R}\right). \tag{4.4}$$

If $y_0(x)$ is known, there is no need to recompute $y(x)$ (at least in theory). This is the power of isomorphism - once an elliptic curve on the unit circle is fully understood, all the distributions on the same j-invariant line are known. However, this doesn't mean they are the same distributions. The PDF is $\sim e^{y(x)}$. The $R^{1/3}$ factor multiplied to $y_0(...)$ changes the exponential integral on the PDF, which in turn changes the statistical properties of the new distribution.[15] In fact, an important effect of moving R via isomorphic mapping is to alter the variance and kurtosis. It is very similar to a diffusion process, which I call "isomorphic diffusion". I plan to address this interesting topic in a separate paper.

In the Introduction, I've emphasized that ECD should have several well-known single-parameter sub-families. Indeed, more than you'd expect, every angle θ forms its own sub-family with $R \in (0, \infty)$. When $R = 0$, it is the standard cusp distribution. So the standard cusp distribution has a very special place in elliptic distribution - It is an singularity in isomorphism. The general cusp distribution family discussed in Section 3.5 is just the sub-family at $\theta = \frac{7}{4}\pi$. And we will study two sub-families, the $j = 0$ and $j = 1728$ lines, in Sections 4.3 and 4.4.

4.2. Solving $y(0; R, \theta)$ Isomorphically

The analytic solution of $y(0)$ is an important subject for numerical computation and it can be solved elegantly by isomorphism. The peak of PDF, $P(x = 0)$, is $e^{y(0)}$ before normalization. Therefore, if $|y(0)|$ becomes very large, the computation associated with $P(x)$ can exhaust the digital precision available to the integration library. The intregral $\int e^{y(x)}dx$ can have convergence issue if the absolute tolerance (`abs.tol`) is not set properly. This can happen when R is very large (for instance, when studying asymptotic behavior).

From Eq. (4.4), the isomorphic mapping leads to

$$y(0; R, \theta) = R^{\frac{1}{3}} y(0; R = 1, \theta), \tag{4.5}$$

where $y(0; R = 1, \theta)$ is $y(0)$ on the unit circle. From the trigonometric equations

[15] The $\sqrt{R}$ term that x is divided by in Eq. (4.4) is just σ. This should be fairly obvious to the reader.

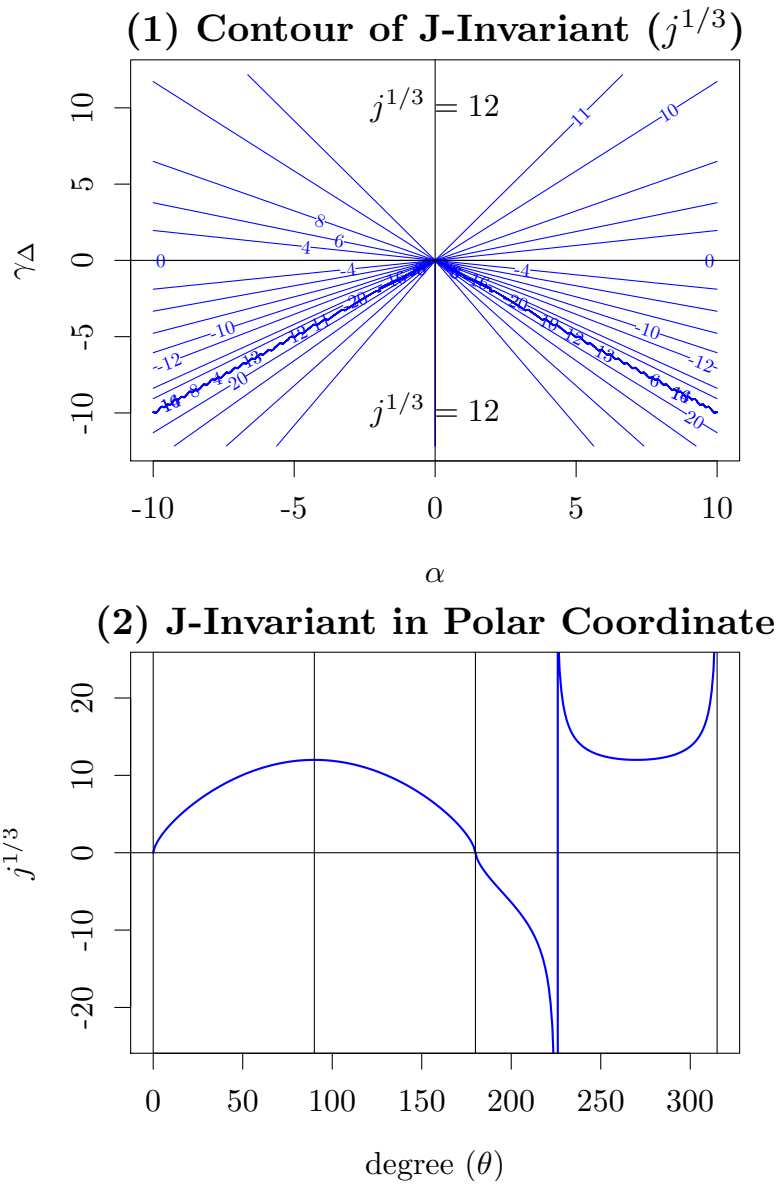

FIG 4.1. *(1) Contour of J-Invariant on the (α, γ_Δ) plane. Since j-invariant tends to be quite large, the cubic root of j-invariant is shown in the plot. Notice j-invariant becomes infinity when the discriminant is zero. (2) J-Invariant on the unit circle in the polar coordinate. Its structure becomes complicated when $\theta > 180°$.*

at $x = 0$, we arrive at the following analytic form,

$$y(0; R = 1, \theta) = Y(\theta) = \begin{cases} -2^{\frac{2}{3}} \sin(\theta)^{\frac{1}{3}} \sinh\Theta, & \theta \in [0, \pi]; \\ -2^{\frac{2}{3}} \sin(\theta)^{\frac{1}{3}} \cosh\Theta', & \theta \in (\pi, \frac{7}{4}\pi); \\ -2^{-\frac{1}{3}} \cos(\theta)^{\frac{1}{3}}, & \theta = \frac{7}{4}\pi; \end{cases} \tag{4.6}$$

where $\Theta = \frac{1}{3}\log\left(\tan(\frac{\theta}{2})\right)$ and $\Theta' = \frac{1}{3}\text{arccosh}\left(\cot(\theta)\right)$[16]. The angular function $Y(\theta)$ defined here is range-bound: For $\theta \in [0, \pi]$, $Y(\theta)$ is 1 at $\theta = 0$; 0 at $\theta = \frac{\pi}{2}$; -1 at $\theta = \pi$; and predominantly linear around $\theta = \frac{\pi}{2}$. For $\theta \in (\pi, \frac{7}{4}\pi)$, $Y(\theta)$ is like a half circle with upward opening, bounded between -1.44 and -0.71. In Section 6, the first part of $Y(\theta)$ will show up again from the asymptotic Taylor expansion on the upper (α, γ) plane. The solution developed here confirms the exact form of $y(0)$, not just the first-order approximation. Furthermore, the $y(0)$ formula will be used in Section 6.4 to estimate the magnitude of the normalization constant C and the absolute tolerance required for the integration library to converge properly.

4.3. Distributions on the J=0 Line

When $\gamma = 0$, the distribution $\texttt{ECD}(\alpha, 0)$ is reduced to $x^2 = -y^3 + \alpha$, and the solution of $y(x)$ has a simple form of

$$y_{j0}(x; \alpha) = \text{sgn}(\alpha - x^2) \left|\alpha - x^2\right|^{1/3}. \tag{4.7}$$

When $\alpha < 0$, $y_{j0}(x) = -\left(x^2 + |\alpha|\right)^{1/3}$ which has a sharp round top near $x = 0$, but not a cusp, due to the smoothing effect of $|\alpha|$ inside the cubic root. The peak value is negative, $y_{j0}(0) = -|\alpha|^{1/3}$. The ellipticity is at $|x| = \sqrt{3\alpha}$ where the convexity changes sign. As shown on the left sides of Panel (1) and (2) of Figure 4.2, the kurtosis decreases and the stdev increases with larger $|\alpha|$. Heuristically speaking, as $|\alpha| \to \infty$, $y_{j0}(x)$ is dominated by the x^2 term asymptotically, and it becomes a normal distribution.

When $\alpha > 0$, the distribution has a special shape that looks like a nipple. This is caused by the flip of sign in $\text{sgn}(\alpha - x^2)$ at the elliptic points, $|x_e| = \sqrt{\alpha}$, where $y_{j0}(x_e) = 0$. The peak value is at $y(0) = \alpha^{1/3}$. This nipple-shape subfamily of distributions consists of the high kurtosis solutions crucial for the financial applications. In Panel (1) of Figure 4.2, we see that the maximum kurtosis $K_{max} = 35.05$ occurs at $\alpha_{max} = 2.94$. This maximum-kurtosis distribution marks the highest modeling potential of ECD. The special symbols, K_{max} and α_{max}, are designated for its quantity and location. Many high-kurtosis daily log-return distributions require parametrization in this neighborhood. Panel (3) illustrates the nipple-shape log PDF of this "maximum" distribution. Panel (4) illustrates its PDF that looks like bullet, very different from a normal distribution.

[16] $\frac{\sin\theta}{\cos\theta+1} = \tan(\frac{\theta}{2})$ and $\text{arcsinh}(z) = \log\left(z + \sqrt{z^2+1}\right)$.

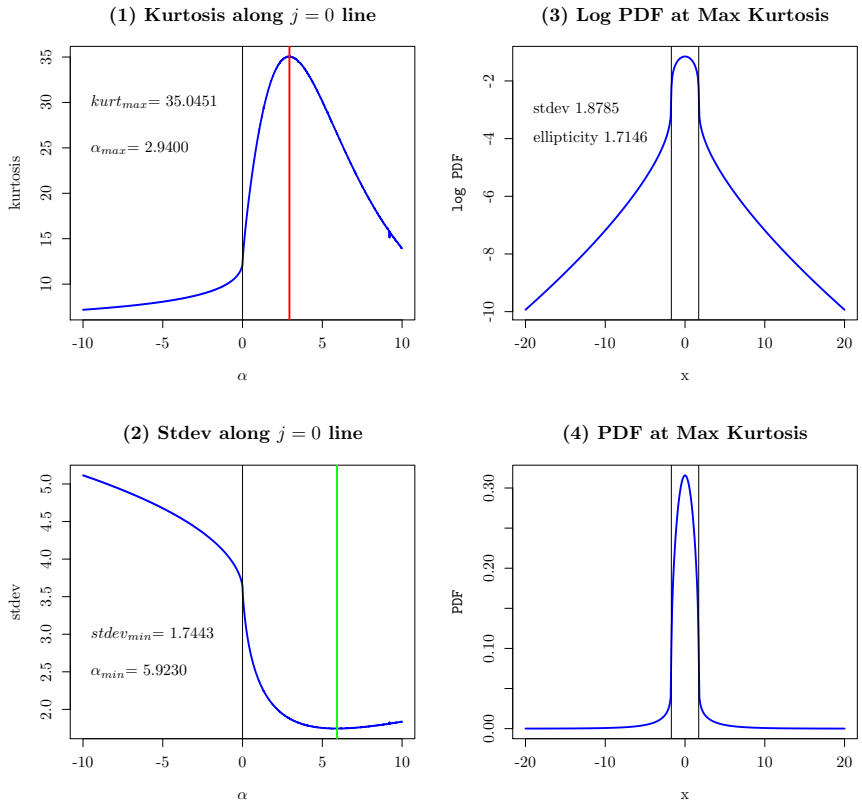

FIG 4.2. *The kurtosis and stdev along the $j = 0$ line. (1) Notice the maximum kurtosis $K_{max} = 35.05$ occurs at $\alpha_{max} = 2.94$. (2) In the range of $\alpha \in (-5, 5)$, the decreasing stdev compresses the probability towards the peak as α increases. (3) The log PDF of the K_{max} distribution has a nipple shape in the peak, then decays quickly at ellipticity, $\sqrt{|\alpha_{max}|}$, indicated by the vertical lines. (4) The bullet-like PDF of the K_{max} distribution. A shape typical of a daily log-return distribution when its kurtosis is near or greater than 20.*

To study the kurtosis asymptotically at large α, we first transform $y_{j0}(x)$ dimensionless with the change of variable $x^2 = \alpha z^2$ (that is, α merged into a role of variance). This leads to

$$y_{j0}(z) = \mathrm{sgn}(\alpha)\,|\alpha|^{\frac{1}{3}}\,\left(1 - \mathrm{sgn}(\alpha)\,z^2\right)^{1/3}. \tag{4.8}$$

The connection to a normal distribution becomes obvious in its Taylor expansion,

$$y_{j0}(z)\,|\alpha|^{-\frac{1}{3}} = \mathrm{sgn}(\alpha) - \frac{z^2}{3} - \mathrm{sgn}(\alpha)\,\frac{z^4}{9} - \frac{5\,z^6}{81} + ... \tag{4.9}$$

The dimensionless moments can be rearranged as

$$\mu_n = C' \int_0^\infty z^n\,\exp\left(y_{j0}(z) - G(z)\right) e^{G(z)} dz, \text{ where } G(z) = -\,|\alpha|^{\frac{1}{3}}\,\frac{z^2}{3} \tag{4.10}$$

C' is a constant that normalizes μ_0 to one. Next, Taylor expansion on z is applied to exp(...) at $z = 0$. The integral is then broken up to a sum of Gaussian integrals. The power expansion by $1/\alpha$ yields both the variance and kurtosis as

$$\frac{\mathrm{var}}{\sigma^2} = \frac{\mu_2\alpha}{\mu_0} = \frac{63}{8} + \frac{3\,\alpha^{\frac{2}{3}}}{2} - \frac{9\,\alpha^{\frac{1}{3}}}{2} - \frac{39}{8\,\alpha^{\frac{1}{3}}} - \frac{117}{32\,\alpha^{\frac{2}{3}}} - \frac{3}{4\alpha} + ..., \tag{4.11}$$

$$\mathrm{kurtosis} = \frac{\mu_4\mu_0}{\mu_2^2} = 3 - \frac{6}{\alpha^{\frac{1}{3}}} + \frac{15}{\alpha^{\frac{2}{3}}} - \frac{345}{4\,\alpha^{\frac{4}{3}}} + \frac{75}{8\,\alpha^{\frac{5}{3}}} + ... \tag{4.12}$$

The precision can be verified by $\texttt{ECD}(-10^3, 0)$ and $\texttt{ECD}(10^3, 0)$. $\texttt{ECD}(-10^3, 0)$ has variance of 203.332 and kurtosis of 3.742 by way of numeric integral. The variance expansion yields 203.327; the kurtosis expansion yields 3.741. On the other hand, $\texttt{ECD}(10^3, 0)$ has variance of 112.351 and kurtosis of 2.542 by way of numeric integrals. The variance expansion yields 112.350; the kurtosis expansion yields 2.541. This is very good. Notice that the variance scales like $\alpha^{\frac{2}{3}}\sigma^2 = R^{\frac{2}{3}}\sigma^2$. This is a general asymptotic result, see Eq. (6.12).

The kurtosis 3 of a normal distribution comes out naturally from Eq. (4.12). The second term $-6\alpha^{-\frac{1}{3}}$ indicates the kurtosis converges to 3 from above and below 3. In particular, for positive α, the kurtosis drops from the maximum of 35 rapidly to below 3, then comes back to 3 slowly. This is illustrated by the purple line in Panel (1) of Figure 6.1. The speed of convergence is slow due to the power of $\alpha^{-\frac{1}{3}}$. α has to be greater than 10^3 to approach the asymptotic region.

4.4. Distributions on the J=1728 Line

When $\alpha = 0$, the distribution $\texttt{ECD}(0, \gamma)$ is reduced to $x^2 = -y^3 - \gamma y$. It has the analytical solution of

$$\begin{array}{l} f_1 = \frac{1}{2}x^2 \\ f_{2C} = \frac{1}{4}x^4 + \frac{1}{27}\gamma^3 \\ f_2 = \sqrt{f_{2C}} \\ f_{sgn} = -1 \text{ if } f_{2C} \geq 0 \text{ and } |f_1| < f_2 \text{ ; else } 1 \\ f_3 = [f_{sgn}(f_1 - f_2)]^{1/3} \\ y_{j1728}(x; \alpha, \gamma) = f_{sgn}(-f_3 + \frac{\gamma}{3f_3}). \end{array} \tag{4.13}$$

When $\gamma > 0$, f_3 is real for all x; and $y_{j1728}(0) = 0$ all the time. This leads to an important property: all the curves on the positive $j = 1728$ line anchor at $(0, 0)$. But when $\gamma < 0$, the peak value is at $y(0) = -\sqrt{|\gamma|}$. The negative peak increasing with $|\gamma|$ is a different behavior compared to the positive side, where the peak is at zero all the time. However, f_3 can become complex for small x when $\gamma < 0$, which makes taking cubic root cumbersome. It becomes hard to gain more analytic insight.

Figure 4.3 illustrates how kurtosis and stdev evolves along $j = 1728$ line. The maximum kurtosis $K_{max}^{j1728} = 13.66$ occurs at $\gamma_{max}^{j1728} = 1.488$. The kurtosis decreases monotonically in both directions as $|\gamma| \to \infty$. In Section 6, we will see that the kurtosis converges to that of a normal distribution most smoothly and orderly along the positive $j = 1728$ line. (This is illustrated by the blue line in Panel (1) of Figure 6.1.) So here we will take a close look at this sub-family of distributions using the trigonometric solution.

When $\gamma > 0$, from Eq. (2.17), we have $\tilde{\alpha} = -x^2 < 0$, $\tilde{\gamma} = \gamma > 0$, and $\tilde{\Delta} = -16(4\tilde{\gamma}^3 + 27\tilde{\alpha}^2) < 0$. Therefore, only one scenario is applicable,

$$y_{j1728}(x) = -2\sqrt{\frac{\gamma}{3}} \sinh\left(\frac{1}{3}\operatorname{arcsinh}(V)\right), \text{ where } V = x^2\sqrt{\frac{27}{4\gamma^3}} > 0. \tag{4.14}$$

To make it dimensionless, we use the change of variable $x^2 = \gamma z^2$, (that is, γ merged into a role of variance) this leads to

$$y_{j1728}(z) = -2\sqrt{\frac{\gamma}{3}} \sinh\left(\frac{1}{3}\operatorname{arcsinh}\left(z^2\sqrt{\frac{27}{4\gamma}}\right)\right). \tag{4.15}$$

The connection to a normal distribution becomes obvious in its Taylor expansion,

$$y_{j1728}(z) = -z^2 + \frac{z^6}{\gamma} - \frac{3\,z^{10}}{\gamma^2} + \frac{12\,z^{14}}{\gamma^3} + ... \tag{4.16}$$

The convergence to normal is speedy. For each power of γ, z adds power of 4. The dimensionless moments can be rearranged as

$$\mu_n = C' \int_{-\infty}^{\infty} z^n \exp\left(y_{j1728}(z) + z^2\right) e^{-z^2} dx, \tag{4.17}$$

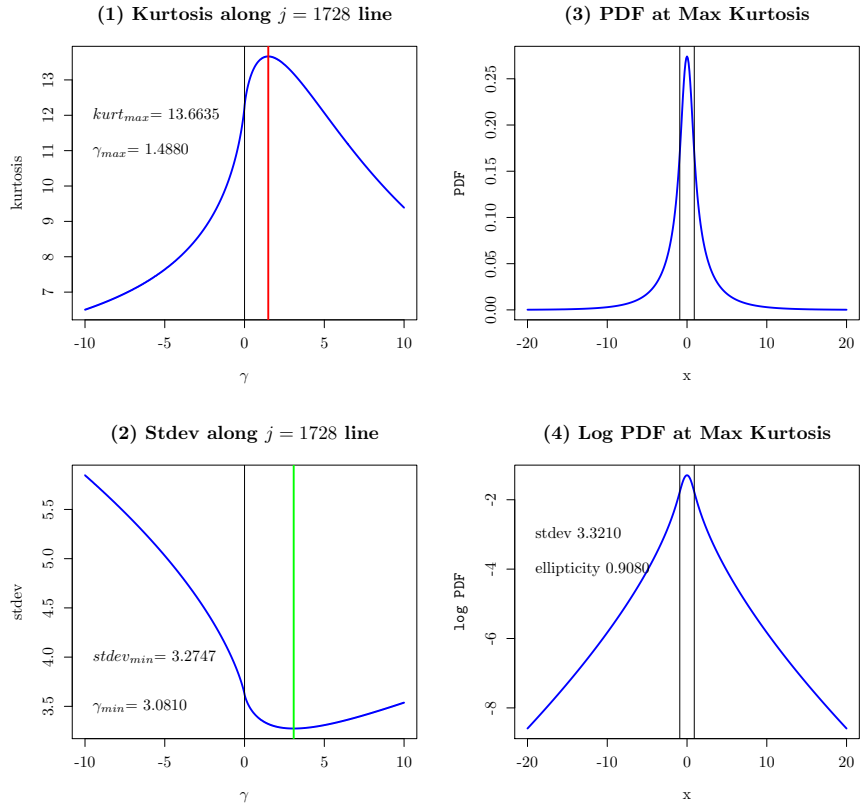

FIG 4.3. *The kurtosis and stdev along the* $j = 1728$ *line. (1) Notice the maximum kurtosis* $K_{max}^{j1728} = 13.6635$ *occurs at* $\gamma_{max}^{j1728} = 1.488$. *(2) The stdev forms a bottom near* $\gamma = 3.08$. *(3) The PDF of the distribution at* K_{max}^{j1728}. *(4) The log PDF of the distribution at* K_{max}^{j1728} *has a triangle shape with an extensive linear slope in both tails.*

where C' is a constant that normalizes μ_0 to one. Next, Taylor expansion on z is applied to exp(...) at $z = 0$. The integral is then broken up to a sum of Gaussian integrals. The power expansion by $1/\gamma$ yields both the variance and kurtosis as

$$\frac{\text{var}}{\sigma^2} = \frac{\mu_2\gamma}{\mu_0} = \frac{45}{8} + \frac{\gamma}{2} + \frac{135}{4\,\gamma} - \frac{13365}{32\,\gamma^2} + \frac{103275}{8\,\gamma^3} - \frac{403920675}{512\,\gamma^4} + ..., \tag{4.18}$$

$$\text{kurtosis} = \frac{\mu_4\mu_0}{\mu_2^2} = 3 + \frac{45}{\gamma} + \frac{4725}{4\,\gamma^2} - \frac{150255}{16\,\gamma^3} - \frac{41802075}{64\,\gamma^4} + ... \tag{4.19}$$

The kurtosis 3 of a normal distribution comes out naturally from Eq. (4.19). The variance is eventually dominated by $\gamma/2$. The precision of two formulai can be verified by $\texttt{ECD}(0, 100)$ whose variance is 55.933 and kurtosis is 3.554 by way of numeric integral. The variance expansion yields 55.926; the kurtosis expansion yields 3.552. They match very well.

5. Standard Deviation, Kurtosis, and Ellipticity

In this section, the contours of standard deviation, kurtosis, and ellipticity on the (α, γ) plane are presented. Interestingly, the contours form beautiful spirals. Especially for kurtosis contour, the spirals convey important asymptotic behaviors.

5.1. Beautiful Contours and Spirals

Figure 5.1 shows the contour plot of standard deviation on the (α, γ) plane with $\sigma = 1$. Since the standard deviation is always proportional to σ, it is fairly easy to put σ back. There are some fine structures in the high-kurtosis region, i.e. $R \lesssim 10$ in the first quadrant, as shown in Panel (2). Once R is reasonably large, the standard deviation is increasing monotonicly with R at every isomorphic direction (that is, at a fixed angle θ).

Figure 5.2 shows the contour plot of kurtosis on the (α, γ) plane. The "high-kurtosis region" can be defined as kurtosis higher than that of the standard cusp distribution ($K_\Lambda \sim 12.257$). This region is concentrated in the first quadrant on the (α, γ) plane as shown in Panel (2). The maximum kurtosis $K_{max} = 35.05$ occurs at $\alpha_{max} = 2.94$. The spirals are all centered around this point although their analytic formula is still unknown. Another interesting phenomena is that the spirals becomes nearly parallel straight lines in the third and fourth quadrants. These lines are parallel to the critical line because the kurtosis decreases very slowly on the critical line and never drops below 6.

The contour plot of ellipticity is shown in Figure 5.3. Ellipticity characterizes the inflection points of an elliptic distribution. Its meaning is different from that of the standard deviation. It marks the points that separate the normal region (concave) from the elliptic (convex) region. So the larger the ellipticity the more it is like a normal distribution. On the contrary, zero ellipticity is the mark of the cusp distributions in which the entire curve is convex.

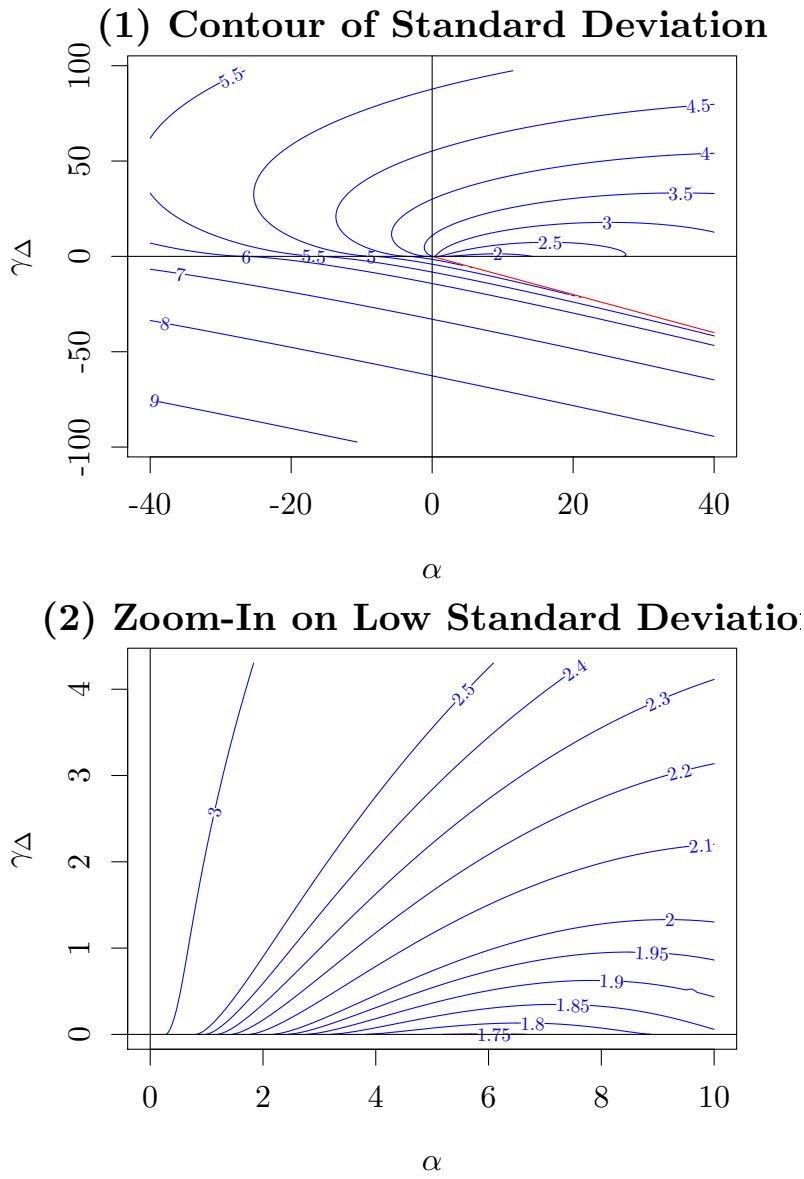

FIG 5.1. *Contour of Standard Deviation. Notice there is an interesting spiral shape. (1) The panel shows a large parameter space with R up to 40. (2) Zooms-in a smaller region wtih* $R \lesssim 10$ *in the first quadrant. This region is also called the "high-kurtosis region".*

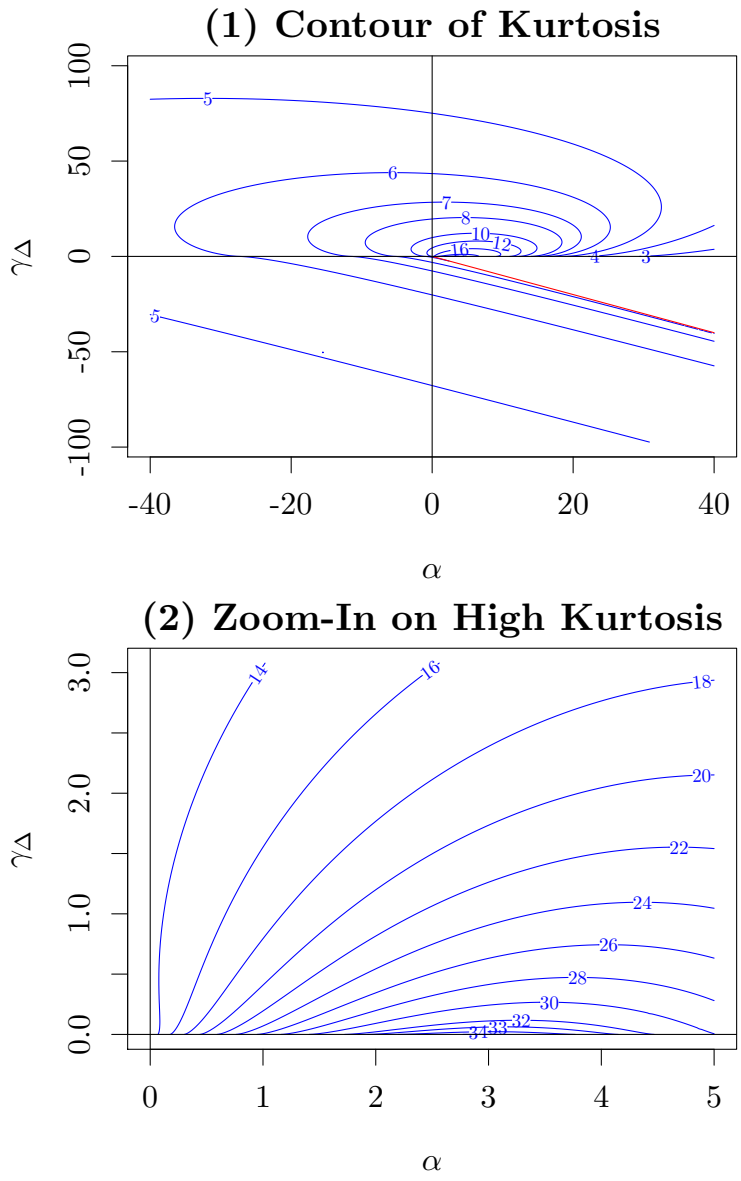

FIG 5.2. *Contour of Kurtosis. Notice there is an interesting spiral shape. (1) The panel shows a large parameter space with R up to 40. (2) Zooms in on the region of highest kurtosis, located at $R < 5$ in the first quadrant.*

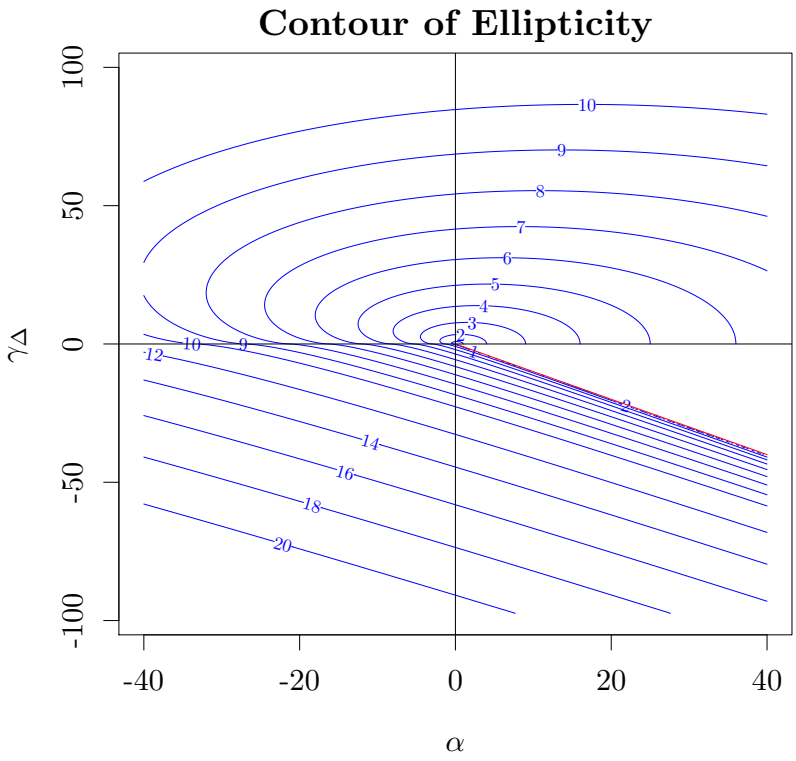

FIG 5.3. *Contour of Ellipticity. Notice the spirals start from zero on the critical line. On the positive $j = 0$ line, the ellipticity is $\sqrt{\alpha}$; while on the negative $j = 0$ line, the ellipticity is $\sqrt{3\alpha}$.*

5.2. *Kurtosis Peak at* ECD$(2.94, 0)$

The spirals of kurtosis contour converges to a point of maximum kurtosis on the $j = 0$ line. Numerical analysis indicates this point is approximately at $\alpha_{max} = 2.94, \gamma = 0$, with kurtosis of $K_{max} = 35.05$. The shape of elliptic curve associated with this highest kurtosis is rather unusual. The log PDF shows a large triangle base with a nipple shaped peak and a round top at the tip of the nipple, as shown in Panel (3) of Figure 4.2. This shape leads to a bullet-like PDF in Panel (4) of of Figure 4.2. Rarely any financial asset can have such a high kurtosis in its daily log-return distribution and still survive for a long history, so we are less concerned about its specifics. However, this point serves as an important reference for analytical purpose as well as for the data fitting procedure.

If we draw a vertical line from $(2.94, 0)$ to $(2.94, \infty)$, this line cuts across all possible kurtosis values from 35 to 3. This property is valuable for an automated data fitting program. When you have a data set with a certain kurtosis and variance, how do you figure out where to start on the (α, γ) plane? One way to get started is to locate the γ on this vertical line with the same kurtosis (or the asymptotic kurtosis, which will be explained later). There is a contour line associated with this kurtosis as shown in Figure 5.2. On this contour line, the distributions have different shapes. You choose the shape closest to your data, either via QQ-plot or some measure of least mean square (LMS) estimation, and that is a good starting point to fit the data.

There is a novel concept associated with such procedure. You can think of kurtosis as the radius parameter of an abstract polar coordinate, and there is an imaginary angle parameter linked to the contour of a given kurtosis. Different angle gives you different shape - a nipple shape, a round shape, a triangular shape, or a cusp. Therefore, the angle parameter is the shape parameter. It is just that such imaginary "polar coordinate" is quite difficult to formulate analytically.

6. Asymptotic to Normal Distribution

By inspecting the kurtosis contour in Figure 5.2, it seems to indicate that the kurtosis spirals approach 3 when they are further away from $(0, 0)$ on the (α, γ) plane, but that region is well beyond 40. So the conformity to normal distribution is a subject of asymptotic behavior in the parameter space. I've made some preliminary exploration of such asympotic behavior on the $j = 0$ and $j = 1728$ lines. In this section, I will use several approaches to study this subject more broadly.

6.1. *Heuristic Expansion*

The $j = 1728$ ***Line.***

First, we look at the asymptotic behavior at the $j = 1728$ line heuristically. When $\alpha = 0, \gamma \to \infty$, and keep $\gamma\sigma^2 =$ constant, Eq. (2.1) becomes

$$\frac{x^2}{\sigma^2} = -y^3 - \gamma y. \tag{6.1}$$

Assume our focus is near the peak where $y^3 \lll \gamma y$, we get the normal distribution,

$$y = -\frac{x^2}{\gamma\sigma^2} \tag{6.2}$$

Therefore, in such limiting case, γ is merged into the role of the variance of a normal distribution, the same role as σ^2. When $\gamma < 0$, we have to expand $y = y(0) + a_2 x^2$ where $y(0) = -\sqrt{|\gamma|}$. And there is a factor of 2 added to the variance, that is, $y = -\frac{x^2}{2|\gamma|\sigma^2}$.

The $j = 0$ Line.

Second, we look at the asymptotic behavior at the $j = 0$ line heuristically. When $\gamma = 0, \alpha \to \infty$, and keep $\alpha\sigma^2 =$ constant, Eq. (2.1) becomes

$$\frac{x^2}{\sigma^2} = -y^3 + \alpha, \tag{6.3}$$

and we have

$$y^3 = \alpha\left(1 - \frac{x^2}{\alpha\sigma^2}\right). \tag{6.4}$$

Taking the cubic root, and perform Taylor expansion, we get

$$y = \alpha^{1/3}\left(1 - \frac{x^2}{3\alpha\sigma^2} + ...\right) \approx \alpha^{1/3} - \frac{x^2}{3\alpha^{2/3}\sigma^2} \tag{6.5}$$

Therefore, in such limiting case, $\alpha^{2/3}$ is merged into the role of the variance of a normal distribution, the same role as σ^2.

The $\theta = \frac{5}{4}\pi$ Line.

Thirdly, when $\gamma = -(27\alpha^2/4)^{1/3}, \alpha \to -\infty$, and keep $\alpha\sigma^2 =$ constant, Eq. (2.1) becomes

$$\frac{x^2}{\sigma^2} = -y^3 + \left(\frac{27\alpha^2}{4}\right)^{\frac{1}{3}} y + \alpha, \tag{6.6}$$

and we have

$$y \approx 2^{2/3}a^{1/3} - \frac{2^{2/3}}{9}\frac{x^2}{\alpha^{2/3}\sigma^2}. \tag{6.7}$$

Again similarly, in this case, $\alpha^{2/3}$ is merged into the role of the variance of a normal distribution, the same role as σ^2. This case is provided because its Δ is also zero, but it is not on the critical line (The ignored line in Figure 2.2).

6.2. Taylor Expansion in Cartesian Coordinate

With these empirical results, we now develop a formal Taylor expansion in the Cartesian coordinate of (α, γ). Assume we can expand $y(x)$ into $y(x) = y(0) + a_2 x^2 + \ldots$, put this into the cubic Eq. (2.1), factoring out x^2 terms, and we get

$$a_2\sigma^2 = -\left(3y(0)^2 + \gamma\right)^{-1}. \tag{6.8}$$

So the task of solving a_2 is the same as solving $y(0)$. And ECD is asymptotic to the normal distribution when $3y(0)^2 + \gamma$ approaches ∞.

Based on the analytic solution of $y(x)$ in Eq. (2.15), we can get a general form of $y(0)$ when $\Delta < 0$,

$$y(0) = \frac{\phi}{2^{1/3}} - \frac{2^{1/3}\gamma}{3\phi}, \tag{6.9}$$

and $\phi = \left(\sqrt{-\Delta/432} + \alpha\right)^{1/3}$. ϕ has the same dimension as $\alpha^{1/3}$ and $\gamma^{1/2}$. Setting $\alpha = 0$, it becomes Eq. (6.2). Setting $\gamma = 0$, it becomes Eq. (6.5). These are validations for the expansion. However, this expansion has a singularity when $\gamma = 0$ and $\alpha < 0$, where ϕ becomes zero. In such case, you just have to count on Eq. (4.7) and Eq. (6.5). For $\Delta > 0$ region, there is no simple way to perform Taylor expansion and estimate variance.

6.3. Taylor Expansion in Polar Coordinate

Similar Taylor expansion technique can be applied to the trigonometric solution, Eq. (2.17) for $\theta \in [0, \pi)$, using polar coordinate (R, θ). This leads to elegant solutions of[17]

$$\begin{gathered} y(0) = \text{-}2^{\frac{2}{3}}\, R^{\frac{1}{3}} \sin(\theta)^{\frac{1}{3}}\, \sinh\Theta, \\ a_2\sigma^2 = \text{-}\frac{2^{\frac{2}{3}}}{3}\, R^{-\frac{2}{3}} \sin(\theta)^{\frac{1}{3}}\, \cosh\Theta, \\ \text{where } \Theta = \tfrac{1}{3}\log\left(\tan(\tfrac{\theta}{2})\right). \end{gathered} \tag{6.10}$$

Dimensional analysis validates $\phi \sim R^{1/3}$. We see that $y(0)$ is just $R^{\frac{1}{3}}\, Y(\theta)$, coinciding with Eq. (4.6). And the normal variance in terms of highest order of R is

$$\text{var}_N = -(2a_2)^{-1} = \tfrac{R^{2/3}\sigma^2}{A_1(\theta)}, \text{ where } A_1(\theta) = \tfrac{2^{\frac{5}{3}}}{3} \sin(\theta)^{\frac{1}{3}}\, \cosh\Theta. \tag{6.11}$$

Here the angular dependency is factored into the angular function $A_n(\theta)$. The $A_1(\theta)$ function is range bound in $\theta \in [0, \pi)$, with min 0.67 and max 1.06, and a shape like sine function. Mirroring Eq. (4.11), the variance should have the general Taylor expansion form of

$$\text{var}_N(R, \theta) = R^{\frac{2}{3}}\sigma^2 \sum_{n=1}^{\infty} \frac{R^{-\frac{n-1}{3}}}{A_n(\theta)}, \tag{6.12}$$

[17] $\frac{\sin\theta}{\cos\theta+1} = \tan(\frac{\theta}{2})$

where we've solved $A_1(\theta)$ and the first few terms of $A_n(0), A_n(\frac{\pi}{2}), A_n(\pi)$ via `Maxima`. This brings clarity for the asymptotic behavior on the entire upper plane.

The next question is - How large does α or γ (or R) have to be to approach a normal-like distribution? I'd like to answer this question here in terms of asymptotic behavior of the kurtosis - assuming when the kurtosis approaches 3, we are close to a normal distribution. Panel (1) of Figure 6.1 shows asymptotic behaviors of kurtosis in four directions, from $\theta = 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$. The kurtosis drops below 3 then comes back to 3 along the positive $j = 0$ line. On the positive $j = 1728$ line, the kurtosis decreases to 3 in the most smooth and orderly manner. On both the negative $j = 0$ line and the negative $j = 1728$ line, the kurtosis drops to 3, but not as fast as in previous two directions. It should also be pointed out that, on the negative $j = 1728$ line, the kurtosis calculation requires MPFR due to increasing negative $y(0)$ causing integration over very small numbers difficult for double precision. The difficulty of calculating kurtosis also has to do with its closeness to the critical line where the asymptoticity is towards Laplace distribution instead of normal distribution.

In the first quadrant, $\theta \in [0, \frac{\pi}{2})$, the kurtosis drops below 3 at certain $R(\theta)$. A normal-like distribution with negative excess kurtosis is useful to describe the effect of finite sample size. This crossover is shown in Panel (2) of Figure 6.1. The sample size of real-world data is always limited in certain ways. For example, even 90 years of DJIA daily data will only have 30 data points if sampled at 3-year intervals (~1024 days). Long-duration time series also have strong survivorship bias. At large time scales, only time series of small and/or negative excess kurtosis can survive for long history. These elliptic distributions are useful to provide a strong "cut-off" that mimics the real-world data.

6.4. The Convergence In Numerical Integration

In the numerical implementation[18], the first challenge is to determine the constant C in

$$P(x) = \frac{1}{C} e^{y(x)}, \text{ where } C = \int_{-\infty}^{\infty} e^{y(x)} dx \tag{6.13}$$

Estimating the order of magnitude of C can improve the speed and accuracy of convergence significantly. The proper estimation depends on the knowledge of both $y(0)$ and the estimated variance. $y(0)$ can be calculated via Eq. (6.10) without further statistical dependency. The variance however requires some analytical estimation before any statistics (moments) can be calculated subsequently (otherwise, it becomes catch-22).

In the first quadrant of the upper (α, γ) plane, the challenge is that, both $y(0)$ and the variance are growing positively as R increases. And $y(0)$ causes C

[18]Note: Reader not concerning numerical implementation can skip this section.

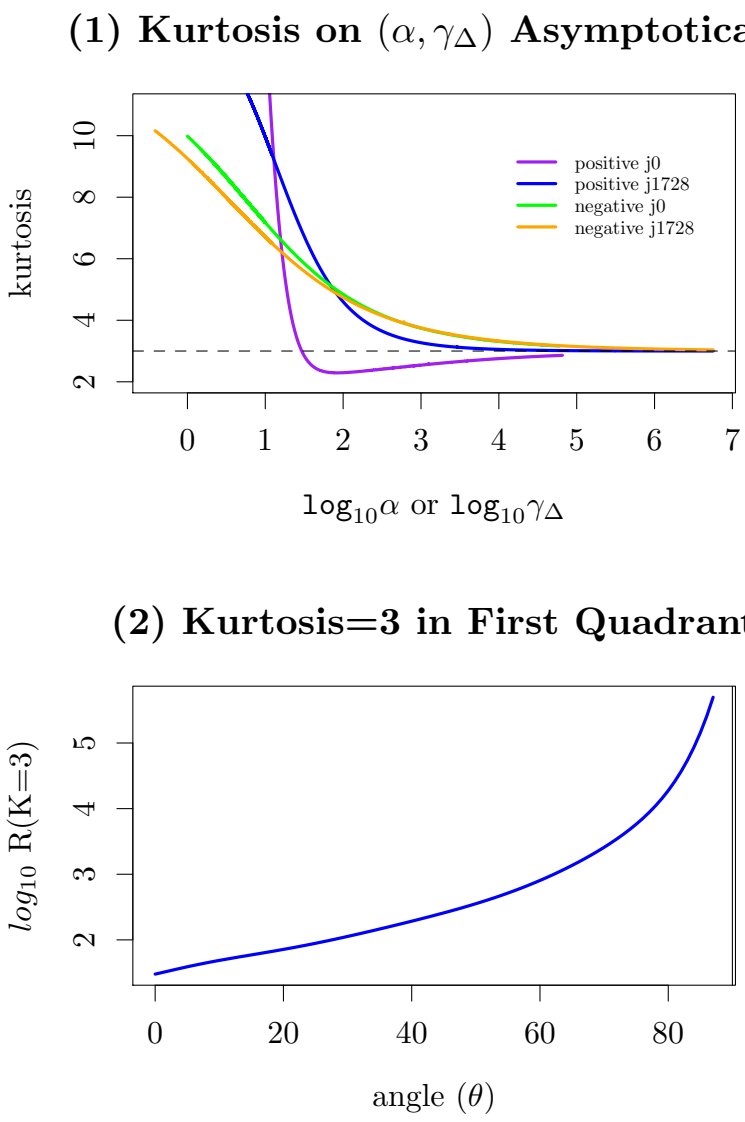

FIG 6.1. *(1) The asymptotic behavior of kurtosis towards normal distribution in the outer parts of* (α, γ_Δ) *plane, from four directions,* $\theta = 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$. *(2) Location* $(R(\theta))$ *where kurtosis drops below 3 in the first quadrant.*

to grow exponentially. [19] From the asymptotic analysis, C can be approximated by

$$C \approx \sqrt{2\pi \ \mathrm{var}_N}\, e^{y(0)}, \text{ where } R \gg 1, \text{ and } \theta \in [0, \pi). \tag{6.14}$$

This allows the program to estimate the range of absolute tolerance when performing the integral.

In the lower (α, γ) plane, the numerical challenge is the opposite. Since $y(0)$ is negative and grows in magnitude as R increases, $y(0)$ causes C to decrease exponentially. The numerical integration is losing accuracy at large R if using double precision based library. One must switch to MPFR in order to regain the numerical precision for infinitesimally small numbers.

7. Comparison to GHYP Distributions

In this section, the similarities and differences between ECD and GHYP family of distributions are explored. Since GHYP has been used in financial modeling for fat tails for many years, the new alternative has to be at least equally good if not more. Figure 7.1 shows 6 simulations at different kurtosis levels. In particular, the well-known variants - HYP, NIG, VG - are chosen (Panel (2,3,4)). As you can see, ECD can fit GHYP fairly good, except the very far end of the tails. This is due to the fundamental difference of asymptotic behavior between the standard cusp and modified Bessel function of the third kind (Appendix B of Breymann 2013[4]). However, ECD's flexible parametrization, especially asymptotically, can push the difference in the tails far beyond the observed quantile. In general, the elliptic tails tend to be slightly more convex than HGYP. The major difference comes when kurtosis is above 20, as shown in Panel (6). GHYP forms a triangular peak, while ECD forms a nipple peak as illustrated in Panel (3,4) of Figure 4.2.

7.1. Normal-Like Distribution

Panel (1) of Figure 7.1 shows an example of a normal-like distribution from GHYP with kurtosis of 4. In Section 6, it has been shown that ECD approaches the normal distribution asymptotically as $R \to \infty$. This simulation shows that the ECD parametrization leans towards the second quadrant (that is, large negative α and positive γ) in order to get a good fit. This is because the shape of ECD in the second quadrant approaches normal distribution in a most smooth manner.

7.2. Hyperbolic Distribution

The Hyperbolic distribution (HYP) has a round top and linear tails. It is moderately leptokurtic with the kurtosis less than or equal to 6.0, as illustrated in

[19] In addition, when θ is small, the nipple shape of $y(x)$ and $dy/dx \to \pm\infty$ around elliptic points adds more complexity to the error tolerance of an automated integration algorithm.

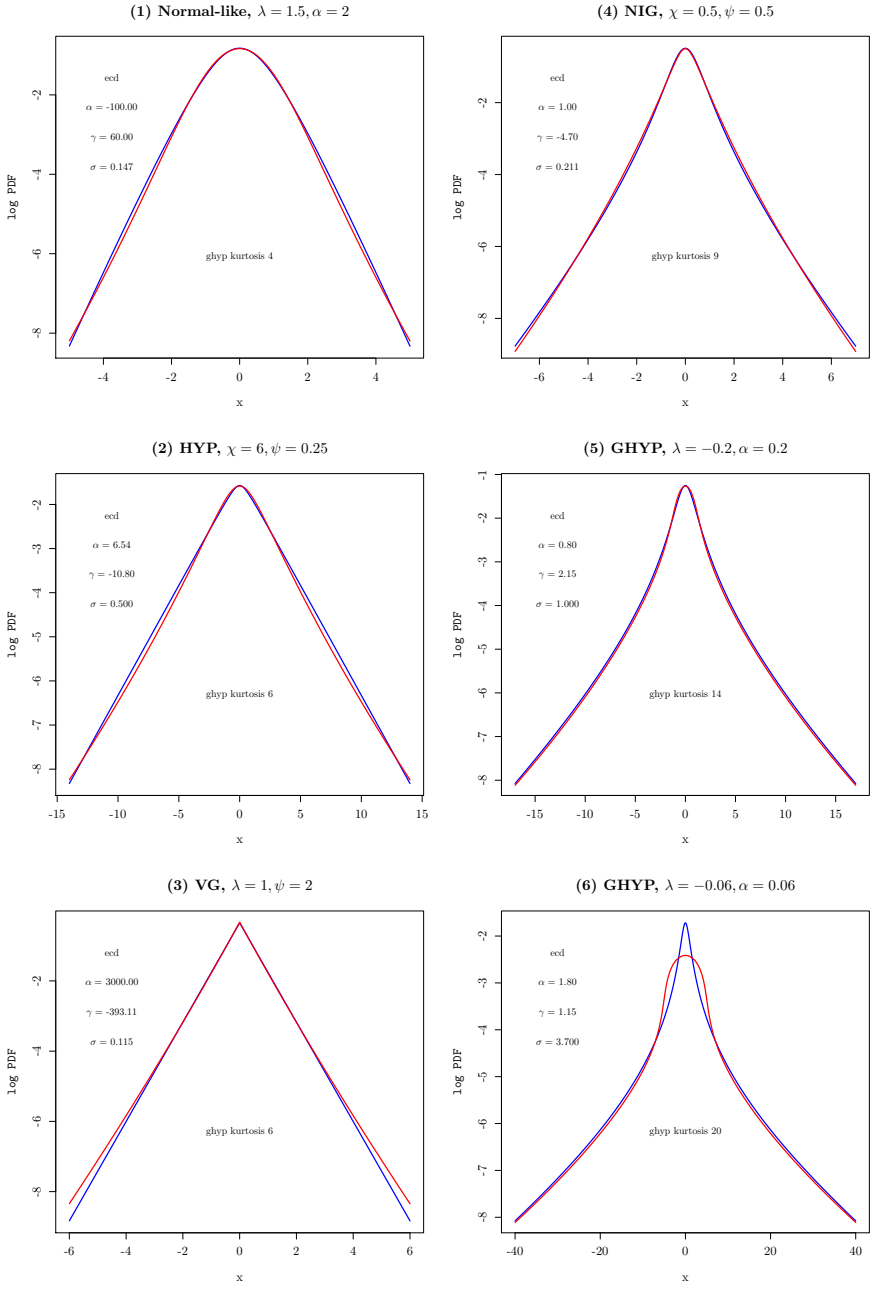

FIG 7.1. *Comparison of ECD to GHYP. Blue lines are GHYP and red lines are ECD fit. The plots are arranged in increasing order of kurtosis.*

Panel (2) of Figure 7.1. It can be characterized by ECD reasonably well with parametrization typically between the negative $j = 1728$ line and the critical line ($\theta > \pi$). HYP has very linear tails while ECD fit tends to be slightly convex, although the difference is very small.

ECD can be generalized to cover hyperbolic distribution by extending the basic polynomial (Eq. (2.1)) from curves of genus 1 to rational curves of degree 2. Using the parameter convention on Wikipedia[20], $y(x)$ for the hyperbolic distribution satisfies

$$\left(\alpha^2 + \beta^2\right) x^2 = -y^2 + 2\beta xy - (\alpha\delta)^2 . \tag{7.1}$$

We can see the two polynomials are very similar except that ECD has a leading y^3 term and the hyperbolic distribution has a leading y^2 term. Adding y^2 term to Eq. (2.1) seems a natural extension of ECD. For example, we can have

$$\frac{(x-\mu)^2}{\sigma^2} = -y^3 - \delta y^2 - \left(\gamma + \beta\left(\frac{x-\mu}{\sigma}\right)\right) y + \alpha. \tag{7.2}$$

When δ is large, the y^2 term will overtake the y^3 term in the finite range of x. This seems a plausible proposal; however, this is actually redundant. In Eq. (2.3), the PDF is divided by the normalization constant C. This makes y translationally invariant, that is, y and $y+c$ are the same for any constant c. So the δy^2 term can be cancelled out by a properly chosen c, and be transformed into translation of (α, γ). Therefore, there must be an asymptotic limit in ECD that can approximate a hyperbolic distribution.

7.3. Variance Gamma Distribution

Variance Gamma (VG) distribution is very similar to the Laplace limit of the general cusp distributions. As demonstrated in Panel (3) of Figure 7.1, it requires very large α (and small σ) to simulate the linear tails in VG. There is still a small difference in the far end of the tails.

7.4. Normal Inverse Gaussian Distribution

Normal Inverse Gaussian (NIG) distribution has a round top (unlike VG) and median-high kurtosis. Panel (4) of Figure 7.1 shows ECD fits NIG particularly well. There is almost no visible difference in the demonstrated range.

7.5. High-Kurtosis GHYP Distribution

Panel (5) and (6) shows two examples of GHYP with high kurtosis. In Panel (5), the kurtosis is slightly above 12, and ECD fits GHYP very well. The

[20] http://en.wikipedia.org/wiki/Hyperbolic_distribution

parametrization is in the first quadrant near $(0, 0)$. In Panel (6), the kurtosis is very high, near 20, the difference between GHYP and ECD becomes obvious. ECD parametrization is close to $(2.94, 0)$. The nipple shape sets in and forms a round top; while GHYP has a distinct sharp peak.

8. Fitting Financial Data in Single Period

In this section, I will apply the elliptic distribution to the daily log-return time series of various financial assets, from exchange rates, commodities, volatility index, to stock market indices, over long history (20-80 years). During the long history, an asset class typically went through major financial and political events causing large disruptions in prices. Such disruptions were reflected in large jumps, fat tails, and large kurtosis (6-30). Tail events can be devastating to ignorant financial institutions and individual investors who use wrong kind of distributions to analyze tail risks. We will explore how to use this distribution to fit various sets of data, get intuition on the range of ECD parameters, and have a feel of how good this distribution behaves in real world.

8.1. Largest Tail Event and Asymptotic Kurtosis

Kurtosis may be one of the most common ways to illustrate the fat tails. However, there are many different ways to interpret kurtosis. Before we dive into fitting the distribution, we must first introduce a new concept called "asymptotic kurtosis along the tail". Such asymptotic kurtosis (or you can call it "tail-truncated kurtosis") follows the same procedure of calculating central moments, then dividing the 4th moment by square of variance, except the integral of PDF is taken between the range of given quantiles - that is, the far ends of the tails are truncated. The reason for such truncation is that, in the real-world data, kurtosis has a subtle instability caused by the largest tail event(s). This is observed among many heavily leptokurtic time series. Such instability can be worked around by this concept of asymptotic kurtosis, making the outcome of the data fitting more predictable.

First, the asymptotic moments $\mu_i(q)$ are defined as, given the tail quantile q,

$$\begin{gathered} x_1 = \Phi^{-1}(q), \\ x_2 = \Phi^{-1}(1-q), \\ \mu_i(q) = \int_{x_1}^{x_2} x^n P(x)\, dx, \text{ where } n = 1, 2, 3, 4, ... \end{gathered} \tag{8.1}$$

Then the asymptotic statistics is defined as

$$\begin{gathered} \text{var}(q) = \mu_2(q) - \mu_1^2(q), \\ \text{skewness}(q) = \big(\mu_3(q) - 3\mu_1(q)\mu_2(q) + 2\mu_1^3(q)\big) \,/\, \text{var}(q)^{3/2}. \end{gathered} \tag{8.2}$$

$$\begin{gathered} \text{kurtosis}(q) = K(q) = \\ \big(\mu_4(q) - 4\mu_1(q)\mu_3(q) + 6\mu_1^2(q)\mu_2(q) - 3\mu_1^4(q)\big) \,/\, \text{var}(q)^2. \end{gathered} \tag{8.3}$$

This forms the theoretical "asymptotic statistics along the tail". By definition, when $q \to 0$, the asymptotic statistics becomes the true theoretical statistics of the distribution. In this paper, I will focus on the behavior of asymptotic kurtosis along the tail since it is the most unstable quantity among the first 4 moments. As shown in Figure 8.1, with limited number of years, even the largest one-day events (green dotted line) can hardly show us where the true kurtosis is (the green bar). What's interesting though is that, in a fairly large range of q, $K(q)$ has a quasi-linear relationship with $\log(q)$. This allows us to extrapolate a "reasonable" trend of asymptotic kurtosis along the tail. By comparing the trend between the theoretical distribution and the statistics from the data, it tells us whether we are getting a good fit or not, which in return validates the relation between observed data and theory.

In order to compare data with theory intuitively, the quantiles (q) are represented by numbers of tail events (n), that is, days. For instance, the largest jump in DJIA (1987 Black Friday) is one out of 37840 days, therefore, $n = 1$ is equal to $q = 2.6 \times 10^{-5}$. Take this day out, the statistics of the remaining data is the asymptotic statistics for excluding one tail event, and n is incremented to $n = 2$. Subsequently you can calculate nth asymptotic statistics by excluding $n - 1$ tail events, where $n = 1..32$. Each statistics (variance, skewness, and kurtosis) forms a curve $f(\log(n))$, which can be fit linearly or quadratically. Asymptotic kurtosis is a linear fit.

Formally, the nth asymptotic statistics for data is calculated as following. Assume $\{r_i, i = 1..N\}$ is the daily log-returns with N days in scope, they are sorted by their absolute values such that $|r_{i+1}| \leq |r_i|$. The asymptotic moments $\mu_i(n)$ are calculated from the truncated sample set $\{r\}_n$,

$$\begin{gathered} \{r\}_n = \{r_k, k = n..N\},\ n = 1..32, \\ \mu_i(n) = \frac{1}{N-n+1} \sum_{k=n}^{N} r_k^i, \text{ where } i = 1, 2, 3, 4, ... \end{gathered} \tag{8.4}$$

From the sample moments $\mu_i(n)$, asymptotic variance, skewness, and kurtosis ($K(n)$) can be calculated accordingly. q and n can be matched up by $q(n) = n/N$.

The six panels of Figure 8.1 shows the asymptotic kurtosis of the six time series we are about to fit. The asymptotic kurtosis of data $K(n)$ is shown in red color. The theoretical asymptotic statistics from the fit $K(q)$ is in green dotted line. The green bar is where the ECD kurtosis should be, as $q \to 0$ or $n \to \infty$ (but will never reach in reality). The quasi-linearity of the trend (blue line) should be obvious to the readers.

In the case of DJIA, after removing the largest event (1987 Black Friday), the asymptotic kurtosis of data drops from 35 to 13. This level of kurtosis is more reasonable. The data can be fit by a ECD distribution with kurtosis of 16 whose asymptotic kurtosis matches well with that of data. Similar improvement is found in all other cases (except VIX). The kurtosis caused by their largest tail events (often just the last one point) jump far above the trend lines. Such jumps distort the kurtosis quite a bit and is an indication that the asymptotic kurtosis is necessary in order to make sense of the finite data set with fat tails. By

matching the observed and theoretical asymptotic kurtosis, we can understand how good the fit is within the ECD framework. Furthermore, if the asymptotic variance and skewness can also be aligned, that is even better.

In VIX, the kurtosis is not high, and the ECD fit is relatively easy. In such case, I choose to cap the theoretical kurtosis by the observed kurtosis. Use of asymptotic statistics is not as important for VIX as in other cases.

8.2. Regression Methodology

It is a challenge to perform regression fit with so many parameters in ECD. The `optimx` and `spg` packages in R are used to perform nonlinear programming, which minimizes the diff function. The diff function is a least-mean-square combination of the following deviations between ECD fit and data.

1. Deviation of stylized statistics: variance, skewness, kurtosis; (In addition, we use asymptotic variance, skewness and kurtosis as an aide for optimal fit.)
2. Deviation of PDF, $P(x)$, for x within 3 standard deviations;
3. Deviation of log PDF; (This is close to maximum log-likelihood.)
4. Deviation between QQ-plot and the 45° line.

Each item can be given different weights to accommodate varying behaviors of the underlying data. Sometimes fitting moments are as good as fitting QQ-plot. But in other cases, one has to choose between better fits to kurtosis or QQ-plot. In these circumstances, weights can be used to influence the output of `optimx`. After the machine fit, ECD parameters can be fine-tuned manually.

The fits are presented in a standard format of four panels: (1) PDF fit; (2) log PDF fit; (3) CDF fit; (4) QQ plot fit. The PDF and CDF fits show how good the theoretical distribution describes the peak region of the population. The log PDF fit shows how good the distribution describes the tails. The QQ plot describes the quantile-to-quantile comparison between the theoretical distribution and observed data, which is a stringent test on the theoretical distribution. In addition, the asymptotic statistics (kurtosis, skewness, and variance) are also examined to remove the distortion caused by the largest tail event(s), as discussed in Section 8.1.

8.3. Swiss Franc (CHF/USD)

The first financial data set is the daily log-returns of Swiss Franc (CHF) to US dollar (USD) exchange rate from 1975 to 2015. This currency is considered a "strong currency", appreciating year over year consistently, as indicated by its negative mean. Up to just a few years, the kurtosis of its log-returns was medium, less than 8, and was one of the lowest. But after 2008, there were several important developments that proved devastating. On Sep 7, 2011, Swiss National Bank (SNB) declared Swiss Franc would peg to Euro at 1.2. Therefore, its exchange rate to USD was literally that of Euro. This date was also the end

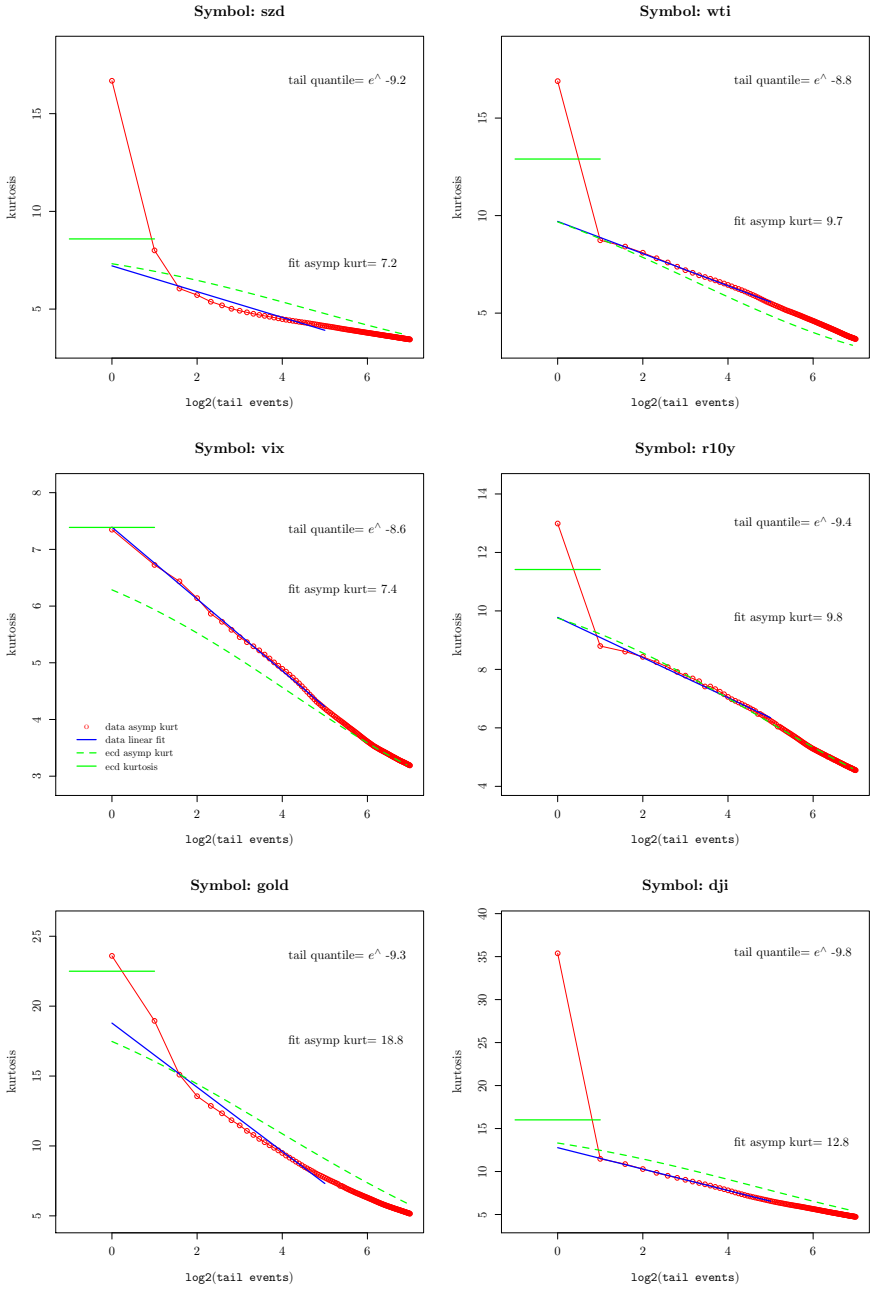

FIG 8.1. *Kurtosis instability and adjustment at the tails. Asymptotic kurtosis follows a quasi-linear relation to the logarithm of tail quantiles, measured by the number of largest tail events. Data is in red. Theory is in green dotted line. The green bar is the ECD kurtosis.*

of gold's 11-year bull market. On Jan 15, 2015, the peg was suddenly scrapped by SNB as Euro suffered free fall selloff against USD due to Greek bailout issues. The exchange rate appreciated 20% overnight and became highly volatile. Investors who purchased Swiss Franc denominated bonds suffered huge losses. The kurtosis jumped to above 16. This demonstrates that even an asset that has been perceived as stable historically can make surprising moves that put its long-term kurtosis well above 12, our barometer of "highly leptokurtic". Figure 8.2 shows the ECD fit for CHF/USD exchange rate. It is impressive that the QQ-plot follows the 45° line closely except the far ends of the tails.

8.4. The Volatility Index (VIX)

The volatility index is a mean-reverting process. It swings between the range of 10 to 90 between bull markets and bear markets. Figure 8.3 is the fit for VIX from 1990 to 2011. The data is medium leptokurtic with excess kurtosis slightly more than 4.0. The fitting is a relatively easy case. The shape of log PDF looks like a wedge, a typical shape of hyperbolic tails. It is also well known that the log-returns of VIX is positively skewed, that is, volatility begets more volatility. The skewness is described very well by the ECD fit. It should be noted that VIX is the only time series that we study here without using asymptotic kurtosis to make large adjustments. That is, all the large jumps in the tails can be described by the linear trend of the asymptotic kurtosis (See Panel (2) of Figure 8.1). It is very impressive that the QQ-plot follows the 45° line precisely.

8.5. Gold

Gold is studied as the representative of commodities and precious metals. This group of assets tends to be very volatile, therefore, the log-returns are highly leptokurtic. Large jumps are common once every few years.

Historical gold prices are based on the London PM fixing, quoted in USD, from 1972 to 2015. Figure 8.4 shows the fit of the log returns of gold prices. The data is highly leptokurtic with kurtosis of more than 20. It is obvious that the tails are bending outward in log PDF on both sides. Its (α, γ) falls squarely in the high-kurtosis region. Notice that gold has a small positive skewness, which is different from the equity indices. The QQ-plot is a straight 45° line. And both the asymptotic kurtosis and theoretical kurtosis fits very well between ECD and data. It appears ECD is a perfect model for gold prices.

8.6. WTI Oil

Another important commodity is the crude oil. Half of the CRB index is based on the movements of oil prices. Figure 8.5 presents the fit of the log returns of WTI oil prices. Its (α, γ) is in the first quadrant but further away from $(0, 0)$. Notice the straight line in the QQ-plot. The asymptotic kurtosis fits reasonably well between ECD and data. ECD has a lot of explanatory power for commodity prices.

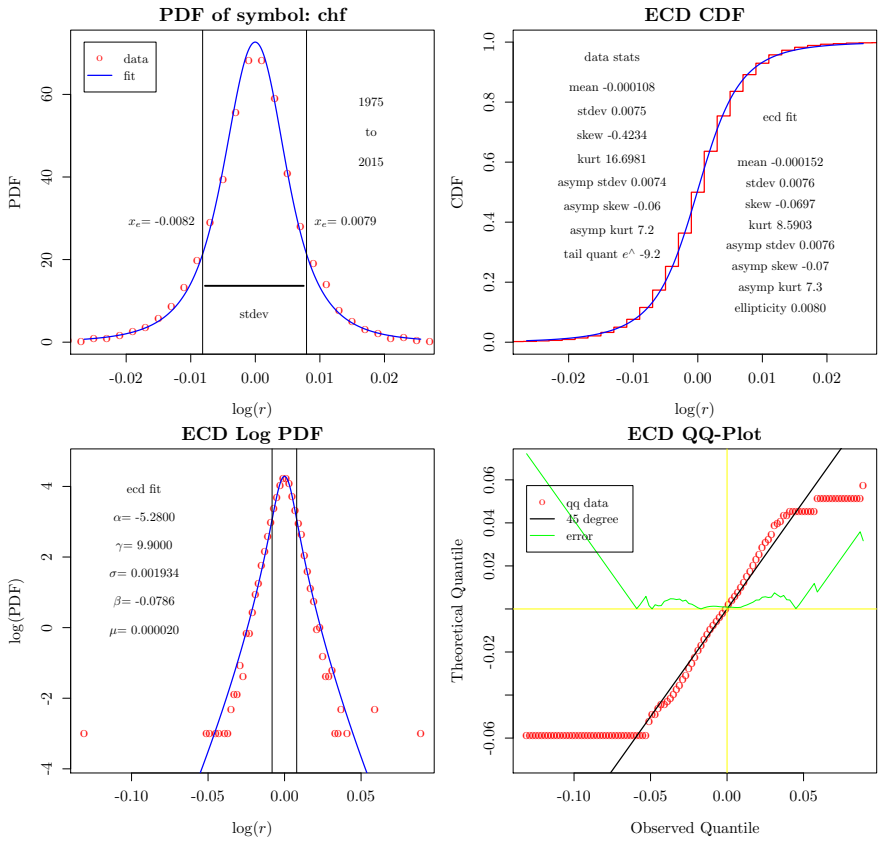

FIG 8.2. *The daily log-returns of CHF/USD exchange rate from 1975 to 2015. The tails are fit very well except the few farthest points.*

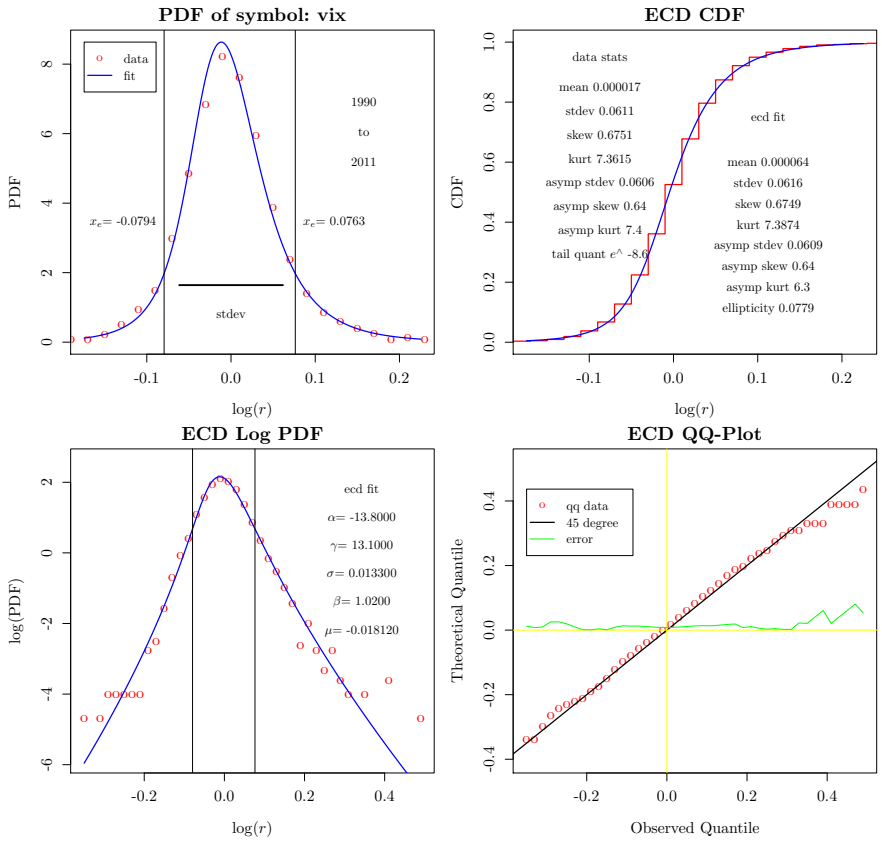

FIG 8.3. *The daily log-returns of VIX (1990-2011). It is very impressive that the QQ-plot follows the* 45° *line precisely.*

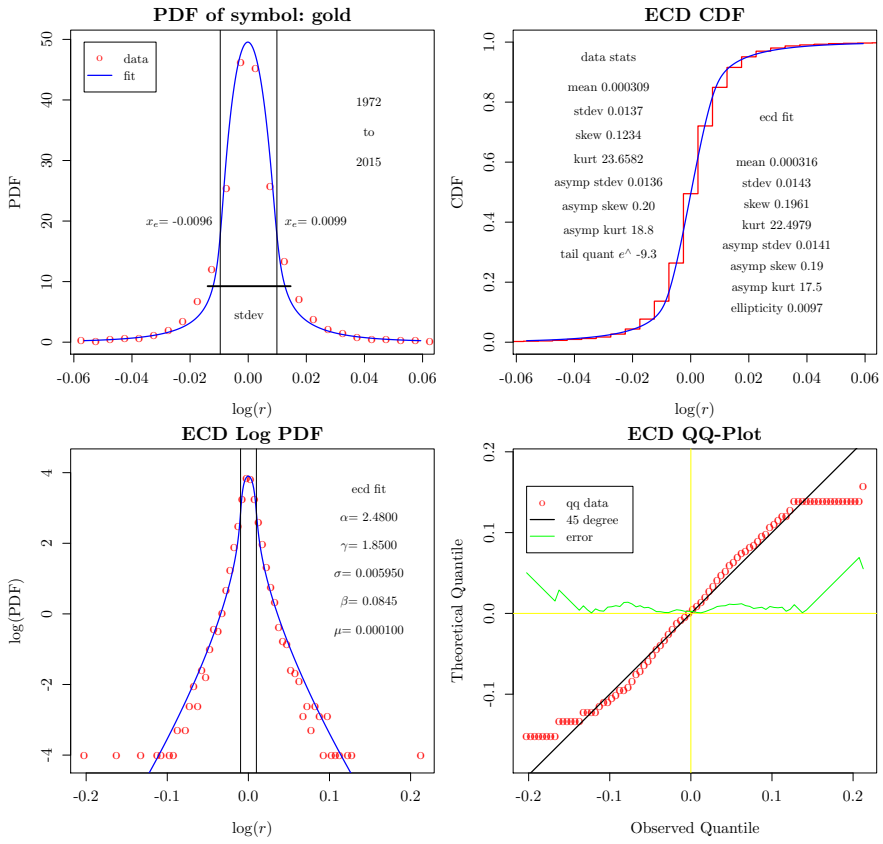

FIG 8.4. *Gold data fit (1972-2015). Both the asymptotic kurtosis and theoretical kurtosis fits very well between ECD and data.*

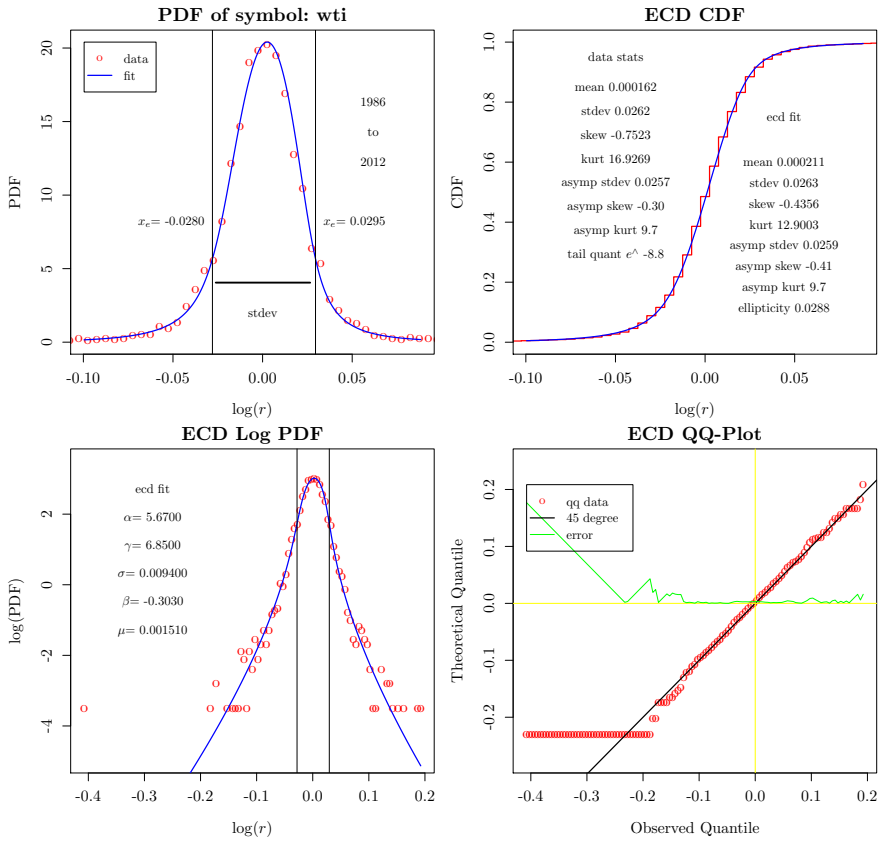

FIG 8.5. *WTI oil data fit (1986-2012). Overweight on fitting QQ-plot; underweight on kurtosis. The QQ-plot and the asymptotic kurtosis are fit very well.*

8.7. 10-Year Treasury Yield (R10Y)

Historical data of the Treasury yield is published on Federal Reserve's website. 10-year treasury yield (R10Y) is the representative of the long-term risk-free interest rate. Figure 8.6 shows the fit for R10Y from 1962 to 2012. The data is highly leptokurtic with excess kurtosis of more than 10.

Although its kurtosis is similar to that of WTI, its shape of distribution is subtly different. R10Y has a more concentrated triangular peak, while WTI has a rounded peak. Such difference is reflected in the ECD parameters. R10Y has a negative α, while WTI has a positive α larger than α_{max} (2.94).

8.8. Dow Jones Industrial Average (DJIA)

DJIA is one of the longest-running market benchmarks with world-wide attention. It is arguably one of the most difficult time series due to its very high kurtosis of 35. The high kurtosis was caused by several market crashes and large price swings during recessions.

Figure 8.7 shows the fit for DJIA daily log-returns from 1935 to 2010. One can see that ECD fit handles it reasonably well. Its (α, γ) falls squarely in the high-kurtosis region. It is somewhat difficult to get to kurtosis of 35, which is about the theoretical limit of ECD. But with the guide of asymptotic kurtosis, a good fit can be accomplished with 25% lower kurtosis. The point on the far left is the Black Monday of 1987, a one-day crash of 22.6%. The asymptotic statistics can correctly isolate the impact of this one day event.

9. Summary

In this paper, I have developed the novel elliptic distribution. The mathematical framework is elegant. It is capable of fitting a wide range of leptokurtic financial data equal to or better than many known distributions. The straight line in the QQ-plot is particularly impressive. The distribution family inherits a rich structure from elliptic curves, e.g. the isomorphic mapping. The parametrization has three distinct asymptotic behaviors - the cusp distribution at $R = 0$; approaching the normal distribution as $R \to \infty$; and approaching Laplace distribution on the critical line. Most importantly, it introduces a new concept of tail behavior unique to the elliptic curves. It is very promising to be a useful statistical tool for the financial professionals.

There are still a lot of statistical properties to be researched on the elliptic distribution. Is it close as a group under addition? How to construct a multivariate distribution? What is the stochastic process associated with it? How does it scale with longer time intervals? How does it apply to volatility forecasting models (e.g. GARCH)? Can it help to explain the kurtosis term structure?

On the toptic of an option pricing model, it is certainly one of the most interesting aspects in quantitative finance. There is likely a general approach based on elliptic distribution; while there could be a special approach that focuses on

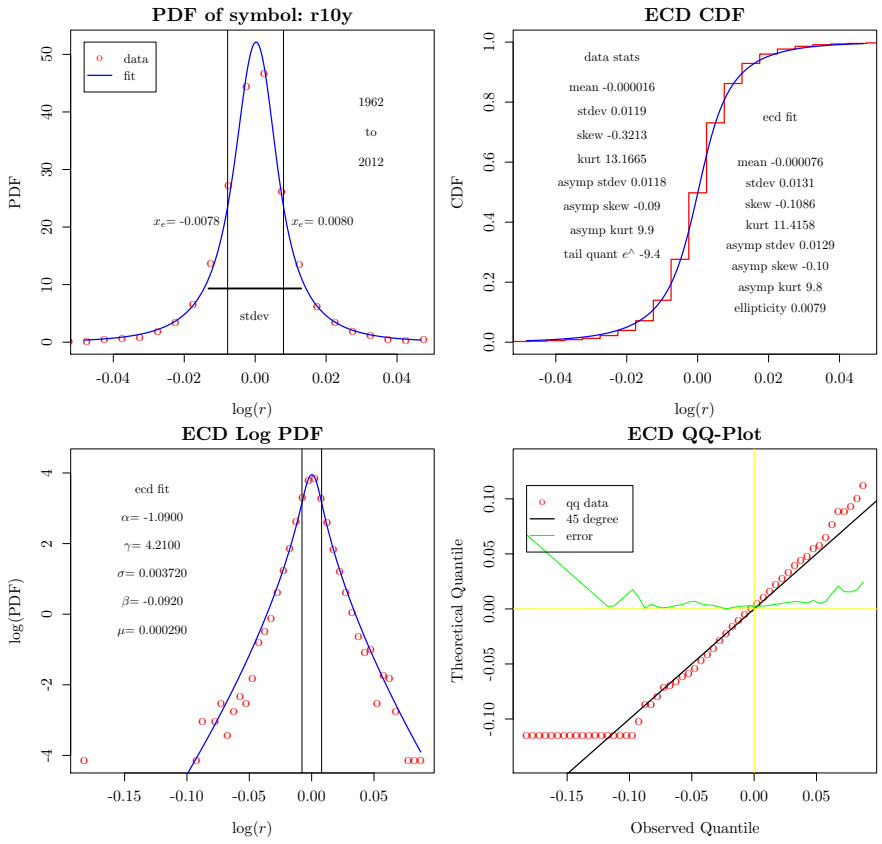

FIG 8.6. *The daily log-return of 10-year Treasury yield (1962-2012). The QQ-plot and the asymptotic kurtosis are fit very well.*

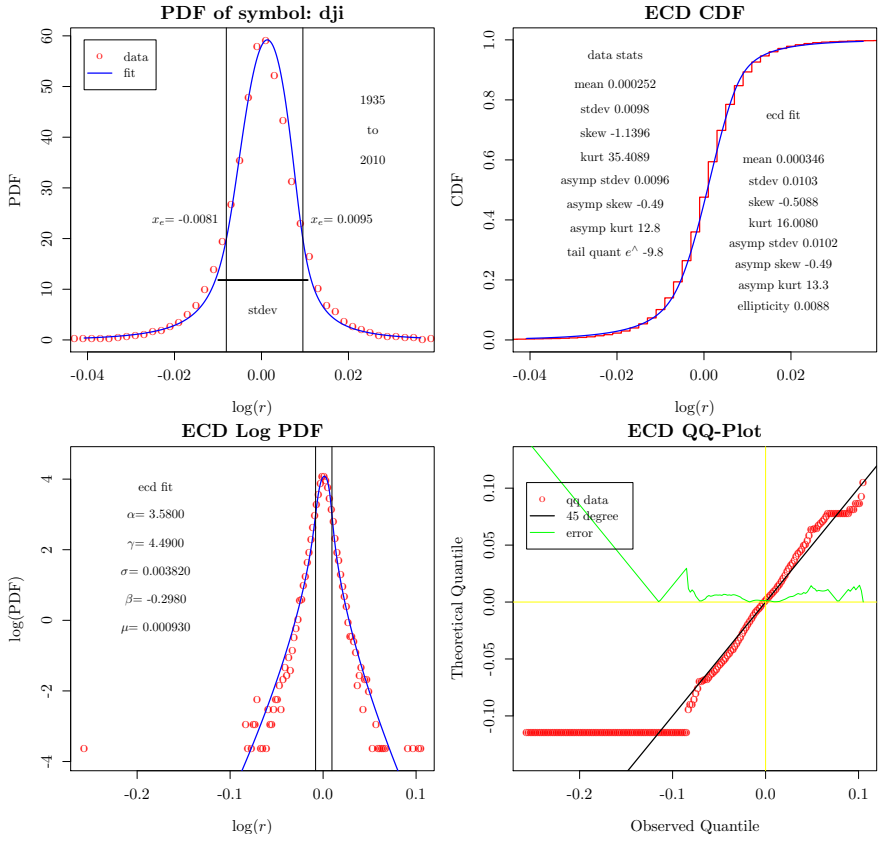

FIG 8.7. *The daily log-returns of DJIA (1935 to 2010). ECD handles the fit reasonably well except the very far end of the tails in QQ-plot. The point on the far left is the Black Monday of 1987, a one-day crash of 22.6%. The asymptotic statistics can correctly isolate the impact of such one day event.*

the big picture of tail behaviors. I will follow up with a subsequent paper from the second perspective.

In 2008, I presented a working paper on a subordinated distribution to B. B. Mandelbrot. He explained to me why he didn't think mixing normal distributions together could be able to describe the kind of fractal and scaling behaviors observed in the financial market. I am indebted to his insight that kept me pondering on this subject these years. Elliptic distribution may be very far from what he had in mind; nevertheless, it is an elegant framework that works really well with real-world data.

I would like to thank the authors of the `ghyp` package. The `ecd` package emulates the wonderful S4 structure laid out by the `ghyp` package. I am also indebted to the `devtools` and `knitr` packages, which make package development and paper authoring much more efficiently. Finally, this work can't be accomplished without the tolerance of my family members. I am very grateful to them.

References

[1] Barndorff-Nielsen, Halgreen (1977). "Infinite divisibility of the hyperbolic and generalized inverse Gaussian distributions". Wahrscheinlichkeitstheorie und verwandte Gebiete 38, p. 309-311.

[2] Bergomi, Lorenzo (2004-2009). "Smile Dynamics" Series on SSRN. http://papers.ssrn.com/sol3/cf_dev/AbsByAuth.cfm?per_id=1177893

[3] Black, Fischer; Myron Scholes (1973). "The Pricing of Options and Corporate Liabilities". Journal of Political Economy 81 (3): 637–654.

[4] Breymann, Wolfgang, and Luthi, David (2013), "ghyp: A package on generalized hyperbolic distributions". Vignette of `ghyp` package on CRAN.

[5] Das, S. R. and Sundaram, R. K. (1999). "Of Smiles and Smirks: A Term Structure Perspective". Journal of Financial and Quantitative Analysis, 34(2), 211-240.

[6] Fergusson, K. and Platen, E. (2006). "On the distributional characterization of log-returns of a world stock index". Applied Mathematical Finance, 13(1):19(38), March 2006.

[7] Gatheral, J (2006). "The Volatility Surface: A Practitioner's Guide". Wiley, ISBN: 0471792519.

[8] Gatheral, J, Jaisson, T., and Rosenbaum, M. (2014). "Volatility is rough". SSRN: 2509457.

[9] Gabaix, X. (2009). "Power Laws in Economics and Finance," Annual Review of Economics, 1, p. 255-93.

[10] Mantegna, R.N. and Stanley, H.E. (1994). "Stochastic processes with ultraslow convergence to a Gaussian: The truncated Lévy flight," Physical Review Letters 73: 2946-2949.

[11] Lihn, S. H.-T. (2008). "Analytic Study of Skew Lognormal Cascade Distribution," SSRN: 1273087.

[12] Lihn, S. H.-T. (2012). "On a Poisson Subordinated Distribution for Precise Statistical Measurement of Leptokurtic Financial Data". SSRN: 2032762.

[13] O'Hagan A. and Leonard T. (1976). "Bayes estimation subject to uncertainty about parameter constraints". Biometrika, 63.
[14] Praetz, P. D. (1972). "The Distribution of Share Price Changes". Journal of Business 45, 49-55.
[15] Rachev, S. T., Fabozzi, F. J., Menn, C. (2005). "Fat-Tailed and Skewed Asset Return Distributions : Implications for Risk Management, Portfolio Selection, and Option Pricing". Wiley, ISBN: 0471718866.
[16] Tate, John T. (1974). "The Arithmetic of Elliptic Curves". Inventiones Mathematicae, Volume 23, Issue 3-4, p. 179-206, Springer-Verlag.
[17] Silverman, J. H. (2008). "The Arithmetic of Elliptic Curves". Second Edition, ISBN 978-0-387-09493-9, Springer.
[18] Steinbrecher, G. and Shaw, W. T. (2008). "Quantile Mechanics".
[19] Wiles, A. (1995). "Modular elliptic curves and Fermat's Last Theorem". Annals of Mathematics, 141 (3) p. 443–551.