Deriving fractional moments using the Moment Generating Function

Jelle Reisinger

(2780350)

30 June 2025

Bachelor Thesis Econometrics and Data Science

Thesis commission:

PhD candidate Gabriele Mingoli

Dr. yyyy (co-reader)

Abstract

The abstract should summarize the contents of the thesis. It should be clear, descriptive, self-explanatory and not longer than a third of a page. Please avoid using mathematical formulas as much as possible. Keywords might be given.

Keywords: Fractional moments, Moment Generating Function, Fractional Calculus.

1 Introduction

Statistical moments, defined as the n-th moment of a random variable X via its probability density function, are essential tools for characterizing data and its distribution. The first and second order moments, the mean and variance, respectively provide key insights into the average value and dispersion of a random variable. Higher-order moments offer further information about the shape and symmetry of the distribution. A less commonly discussed class of moments, however, are the fractional moments. The latter are defined in precisely the same manner as the integer moments, but now with $n \in \mathbb{R}$, or even $n \in \mathbb{C}$. From this point onward, when referring to fractional moments, we denote the order by $\alpha \in \mathbb{R}$ instead of n. While these moments may not find as much usage in comparison with the integer moments of a distribution, they can be rather valuable in specific applications.

Fractional moments play a significant role in various fields, including finance, economics, and statistics. One notable application in statistics is their use in approximating integer moments, as described by Novi Inverardi and Tagliani (2024). In such cases, evaluating a 'nearby' existing fractional moment can provide an approximation of the otherwise undefined integer moment Novi Inverardi and Tagliani (2024).

As an illustration, one can consider the student-t distribution with $\nu = 2$ degrees of freedom. This distribution only has central and raw moments of order k, where $0 < k < \nu$. This implies that its second central moment $(k = 2)$ does not exist. One could however consider the central moment of order k, where $k = 1.95$, given that this order of the moment does in fact exist and interpret its value as the variance of the distribution.

Fractional moments are also used in financial modelling, particularly in the context of Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models. These models are commonly applied to time series data such as financial returns and capture the dynamic volatility that changes over time. The GARCH model achieves this by modelling the volatility based on the returns and variances of previous time periods. It is useful to note that in the context of financial returns, the variable of interest often follows a distribution with heavy tails. As a consequence, not all relevant moments of integer order may exist. In such cases, fractional moments may serve as a viable alternative. Hansen and Tong (2024) propose a method for computing fractional absolute moments of the cumulative return, quantities that would otherwise be undefined using traditional integer-order techniques.

In addition, Hansen and Tong (2024) apply a Heterogenous Autoregressive Gamma Model (HARG) to model intraday stock data. The HARG model assumes the conditional distribution of the variable of interest, say X_t to be of a non-central Gamma distribution. This model is commonly used to model positive-valued time series data Gourieroux and Jasiak (2006). In particular, the volatility of the random variable is of interest, which by definition is non-negative.

In this context, X_t is the realized variance. Hansen and Tong (2024) study the conditional moments of the variance process. Specifically, they consider moments of order $\frac{1}{2}, \frac{3}{2}$ and 2 which correspond to the standard deviation, skewness and kurtosis respectively. Furthermore, the moment of order $-\frac{1}{2}$, the inverse volatility has also been computed. In this context, this moment

serves as an estimate of the Sharpe ratio, which is an index to measure the performance of some investment compared to a risk-free asset Sharpe (1994). These fractional moments offer valuable insights into financial return dynamics and are highly relevant for modelling and risk assessment. In a related application, Mikosch et al. (2013) apply fractional moments to financial time series data, focusing on the stationary solution of a stochastic recurrence equation - a recursive formula that describes the evolution of a time series. This equation generalizes common time series models in econometrics such as the ARCH and GARCH and is thus of great interest in the context of financial econometrics. Notably, the GARCH(1,1) model can be represented within this one-dimensional stochastic recurrence equation framework Mikosch et al. (2013).

Similar to the case study by Hansen and Tong (2024), conditional moments of fractional order are of great interest, as is often the case in the context of (G)ARCH models. The latter can again be explained by the fact that fractional moments offer a valuable alternative for characterizing the distributional properties of financial time series.

Gzyl et al. (2013) also introduce the use of Fractional moments in risk-models, specifically insurance models. In such models, often the probability density function of total ruin, the event where an insurance company's capital becomes negative, is unknown or difficult to compute. The method of maximum entropy has been developed as a way to approximate these unknown density functions. This method uses fractional moments as input, as they have been proven to be able to characterize its distribution Lin (1992). Compared to traditional techniques, such as inverse Laplace transforms, the maximum entropy approach requires fewer moments to achieve a reliable approximation, making it more computationally efficient and practical for real-world applications. In prior research by D'Amico et al. (2002) the same approach of using fractional moments as input for maximum entropy methods to characterize unknown distributions has been employed. One notable application is in option pricing in A Black-Scholes model. Making use of fractional moments, numerical approximations of theoretical entropies have been obtained. Remarkably, using only two fractional moments, the method achieved approximations with three-digit accuracy.

This approach outperformed traditional maximum entropy methods that rely on integer moments, both in terms of accuracy and computational efficiency. Moreover, the use of fractional moments helps to avoid numerical optimization issues that commonly arise when minimizing functions based on integer-moment constraints. Therefore, the results obtained by D'Amico et al. (2002) seem to correspond with the aforementioned results obtained by Gzyl et al. (2013). Namely, the implementation of fractional moments in the context of maximum entropy methods has great advantages when it comes to numerical stability and efficiency as compared to more traditional methods.

Beyond finance and risk modelling, fractional moments have found valuable applications in fields such as engineering and healthcare. As previously mentioned, Gzyl and Mayoral (2024) used fractional moments combined with the maximum entropy method to predict total ruin. This approach has been extended to estimating lifetime distributions, a general topic of interest in survival analysis Clark et al. (2003). Both probability density functions and survival functions were estimated using this technique. Once again, the prediction errors remained within three-digit accuracy, further supporting the method's consistency and precision across different fields of research.

Other examples in the field of engineering include optimizing signal processing and control systems as well as studying the response characteristics of random vibration systems. Wang et al. (2025) have shown that when using the concept of fractional moments for the latter, accuracy and stability is higher compared to traditional methods, such as Taylor expansions. Furthermore, in terms of simplicity and efficiency, the method of fractional moments is advantageous, as its computation steps are straightforward and avoid convergence issues, significantly reducing the resources required for computation. Implementing the usage of fractional moments has allowed Wang et al. (2025) to obtain both analytical, as well as numerical solutions to problems within

their research field, which again proves its viability.

Another noteworthy application of fractional moments is in the identification of distributions in complex, non-linear systems. Di Matteo et al. (2014) demonstrate how complex fractional moments can be used to solve differential equations such as the Kolmogorov or Fokker-Planck equation, which arise in the context of continuous-time Markov processes. By transforming the system into a more agreeable form using Mellin transforms, Di Matteo et al. (2014) show that solutions can be efficiently computed and later recovered using inverse transformations. A key advantage of this method is that it preserves the entire support of the probability distribution, which traditional integer moment methods fail to achieve.

The remainder of this paper is structured as follows. Section 2 discusses existing methods for computing fractional moments, outlining their respective advantages and disadvantages. This section concludes with an introduction to the novel approach proposed in this paper, the required techniques, and the limitations to be mindful of. Section 3 lays the mathematical foundation for this new method and provides a brief historical overview of the development of fractional derivatives. In section 4, some key definitions and properties of statistical moments and the moment generating function will be highlighted. What is more, the theory of fractional derivatives from section 3 will be incorporated in order to extend functionality of the moment generating function. Relevant properties are revisited in this new context, and potential analytical errors of the method are addressed. In section 5, simulation results will be analysed, which further clarify the numerical stability and potential errors of computing fractional moments via the moment generating function. Finally, in section 6, the practical implementation of this method in a financial setting will be implemented. The results will be compared against those obtained using traditional methods, with a focus on the implications of potential errors in real-world applications.

2 Methodology

2.1 Existing methods of obtaining fractional moments

The formal definition of the statistical moments mentioned in the previous section is as follows: The n-th moment of a random variable X, with probability density function (PDF) $f_X(x)$

$$\mathbb{E}[X^n] = \begin{cases} \int_{-\infty}^{\infty} (x-c)^n f_X(x) \, dx & \text{if } f_X(x) \text{ is continuous,} \\ \sum_i (x_i - c_i)^n f_X(x_i) & \text{if } f_X(x) \text{ is discrete.} \end{cases}, n \in \mathbb{N}$$

The traditional method of computing fractional moments is rather straightforward. Similarly to integer moments, one simply computes the summation (in the discrete case) or integral (in the continuous case) of $x^{\alpha} \cdot f_x(x)$, where $f_x(x)$ denotes the probability density function of the random variable X. In the context of fractional moments, $\alpha \in \mathbb{R}$ instead of $\mathbb{Z}$ (assuming that negative moments exist). Hansen and Tong (2024) introduce an alternative approach to computing fractional moments using the complex moment generating function (CMGF), which they apply in the context of the aforementioned GARCH models. One of their key expressions is given by:

$$\mathbb{E}|X - \xi|^{\alpha} = \frac{\Gamma(\alpha+1)}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-\xi z} M_X(z) + e^{\xi z} M_X(-z)}{z^{\alpha+1}} dt$$

where $z = s + it, s \in \mathbb{N}_+, \xi \in \mathbb{R}$ and α of course the order of the moment.

This formulation extends upon the traditional moment generating function (MGF) but avoids the process of taking derivatives, making it computationally efficient. The inclusion of the Gamma function is logical, as it extends the factorial function to real values, aligning well with the computation of fractional moments. Since this method relies on integration, rather than differentiation,

it avoids numerical issues that might arise when computing derivatives, such as obtaining rather great approximation errors.

2.2 Obtaining fractional moments by using the moment generating function

While the CMGF method provides an efficient and elegant alternative to the traditional method, this thesis explores a different approach: computing fractional moments directly by applying fractional derivatives to the MGF. The MGF is widely used for computing integer moments by differentiation around zero. Extending this approach to fractional orders requires us to take fractional derivatives. Thus, we need to define such fractional derivative operators. These fractional derivatives have a long history and often make use of the aforementioned Gamma function in combination with some integral. This means that, for continuous random variables, where we integrate the MGF, we will have to do double integration. A consequence might be that obtaining analytical expressions of these moments may not always be possible. A lot of alternative expressions of these fractional derivatives have been created, mostly based on different interpretations of the latter in the field of physics. In this thesis, we will focus on computing the MGF using the Riemann-Liouville derivative, the Caputo-Fabrizio derivative and the Grünwald-Letnikov fractional derivative. Each of these fractional derivatives in combination with the MGF might lead to different moments expressions for the same distribution and same fractional order of the moment. Thus, it is essential to compare each of these definitions with the traditional way of computing fractional moments, to derive their accuracy and conclude which approach is most suitable for fractional moment computation. Similar to the expression of the moments of a random variable, their errors may also be hard to derive analytically, depending on its distribution.

2.3 Order and methods of research

To realize the differences between the aforementioned fractional derivatives, the order of research will be as follows. First, a mathematical groundwork for the fractional derivatives will be laid, in which a number of their respective properties will be discussed. Next, we will revise some basic definitions of statistical moments and the moment generating function and see how some of these properties change when considering fractional moments in combination with the moment generating function. If possible, analytical expressions of the errors for each fractional derivative in combination with the MGF will be computed, which will conclude the theoretical research of this thesis. We will conduct a simulation using the programming language Julia, to now obtain numerical errors instead of analytical errors. The simulation will evaluate each MGF-based method using statistical measures, including minimum, maximum, and average errors. Finally, we will consider a practical case (WIP: TO BE DECIDED) in which all three derivatives will once again be compared.

3 Fractional Calculus

3.1 Overview and applications of fractional derivatives

Although not the main topic of interest of this thesis, it is useful to have some knowledge about the history and applications of fractional derivatives. The study of fractional derivatives has been relevant as early as the year 1695 when the concept of such a derivative was implicitly discussed by Leibniz and Bernoulli Katugampola (2014). Since then, numerous definitions of fractional derivatives have been developed. The best-known definition of the fractional derivative is the

Riemann-Liouville derivative, its upper derivative of a function $f(x)$ of order α is denoted as

$$\frac{d^n}{dx^n} \frac{1}{\Gamma(n-\alpha)} \int_a^x (x-t)^{n-\alpha-1} f(t)dt, \text{ where } n = \lceil \alpha \rceil$$

Kilbas et al. (2006). Michele Caputo (1967) defined a variation on this derivative, where instead of $\frac{d^n}{dx^n}$ in front of the integral, we have $\frac{d^n}{dt^n}$ inside the integral. Due to this adjustment, it is possible to have initial value conditions expressed as the traditional derivatives of integer-order, which made these fractional differential equation problems more intuitive. Other fractional derivatives, such as the Hadamard (1892) and Riesz (1949) derivative, have been defined to take advantage of particular beneficial properties. For example, each of the latter derivatives can be written as a Fourier transformation. As a consequence, the analytical expressions, can often be simplified. A rather unique derivative is the Grünwald-Letnikov derivative which, in contrast to all the aforementioned derivatives, is not based on integral. Instead, it generalizes the difference quotient, $\frac{f(x+h)-f(x)}{h}$, to fractional orders using binomial coefficients Atici et al. (2021). This variety of definitions emphasises how dependent fractional derivatives are on different physical interpretations and practical applications. Beyond their theoretical significance, fractional derivatives have been of significant importance in various scientific fields since the 19th century. Examples include fractional Fourier transformations, a generalization of the regular Fourier transformations Missbauer (2012), fractional diffusion equation models, describing the motion of particles in liquids as a consequence of thermal molecular motions Einstein (1905) and the fractional Schrödinger equation, a generalization of the Schrödinger equation, often used in quantum mechanics Laskin (2002). Their applications are less common in the fields of finance or economics, as fractional derivatives are mainly used to describe natural phenomena Boulaaras et al. (2023). Yet they still offer some great potential. (Symmetric) Levy flights make use of fractional derivatives in order to solve partial differential equations which describe random walk processes in time series Scalas et al. (2000). The development of fractional derivatives also led to the notion of fractional Brownian motions, a generalization of the Brownian motion Mandelbrot and Van Ness (1968). The latter is a continuous-time stochastic process which, similar to Levy flights, may be used to model random walk processes.

3.2 Formal definitions of fractional derivatives

In order to obtain the expressionS mentioned in section 2, some advanced tools are required. We can find these in the field of fractional calculus. We define the following:

Definition 3.2.1. Let D be the differential operator, such that $Df(x) = \frac{d}{dx}f(x)$. Then the fractional derivative of order α is defined as

$$D^\alpha f(x) = \frac{d^\alpha}{dx^\alpha} f(x)$$

.

In this definition, α can be any real number. When taking regular derivatives, $\alpha \in \mathbb{N}$. In most of our cases, we are interested in the instance where $\alpha \in \mathbb{R}_+$.

It is also possible to study derivatives of negative order, which can be used to obtain moments of negative order of a function, provided that such an order exists. A derivative of negative order is simply an integral of positive order. This is defined as follows:

Definition 3.2.2. Let I be the integral operator, such that $If(x) = \int f(x)dx$. Then the fractional integral of order α is defined as

$$(I^\alpha f)(x) = \frac{1}{(\alpha-1)!} \int (x-t)^{\alpha-1} f(t)dt$$

Cauchy (1823).

Combining the previous two definitions, we obtain the following, more general, definition.

Definition 3.2.3. The differintegral operator is defined as

$$R^\alpha f(x) = \begin{cases} I^{|\alpha|} f(x) & \text{if } \alpha < 0 \\ D^\alpha f(x) & \text{if } \alpha > 0 \\ f(x) & \text{if } \alpha = 0 \end{cases} , \text{ with } \alpha \in \mathbb{R}. \tag{1}$$

As mentioned in section 3.1, quite a number of different definitions have been proposed to compute a fractional derivative. Some of these definitions are rather similar, thus, in this paper, we will focus on some of the more well known fractional derivatives. We will start with the most famous fractional derivative, which laid the foundation of the study of fractional derivatives as early as in 1832.

Definition 3.2.4. The left-side Riemann-Liouville fractional derivative of order α is defined as:

$$D^\alpha_{a_+} f(x) = \frac{d^n}{dx^n} D^{-(n-\alpha)}_x f(x) = \frac{d^n}{dx^n} I^{n-\alpha}_x f(x) = \frac{d^n}{dx^n} \frac{1}{\Gamma(n-\alpha)} \int_a^x (x-t)^{n-\alpha-1} f(t) dt \tag{2}$$

Joseph Liouville (1832).

Here, $n = \lceil \alpha \rceil$, the ceiling function and $\Gamma(.)$ is the Gamma function, see section Appendix A.

Note that we just defined the left-side Riemann-Liouville fractional derivative, suggesting that there also exists a right side derivative. In the case of the latter, we would evaluate the associated integral the other way around. Namely, $D^\alpha_{b_-} f(x) = \frac{d^n}{dx^n} \frac{1}{\Gamma(n-\alpha)} \int_x^b (x-t)^{n-\alpha-1} f(t) dt$. We intend on using the left-side derivative, as is supported by Tarasov (2023). The reason is, due to the fact that many functions in probability theory, most importantly the cumulative distribution function, are defined as an integral from some constant to x.

Remark 3.2.1. For values $\alpha \in \mathbb{N}_+, n = \lceil \alpha \rceil = \alpha$, so $\Gamma(n-\alpha) = \Gamma(0)$, which is undefined. Thus for $\alpha \in \mathbb{N}_+$, we define: $D^\alpha f(x) = \frac{d^\alpha}{dx^\alpha} f(x)$, which is simply the regular expression for derivatives of integer order.

A modification of the Riemann-Liouville derivative is the Caputo-Fabrizio derivative, which is defined as follows:

Definition 3.2.5. The Caputo-Fabrizio fractional derivative of order $\alpha, \alpha \in [0,1)$ is defined as:

$$D^\alpha f(x) = \frac{1}{1-\alpha} \int_a^x \exp\left(\frac{-\alpha}{1-\alpha}(x-t)\right) f'(t) dt \tag{3}$$

Caputo and Fabrizio (2015). With $a \in [-\infty, x)$. The Caputo-Fabrizio is always defined as the integral from some constant to the variable x. This is another reason for choosing to work with the left-side Riemann-Liouville integral. In this way, it will be more straightforward to compare the two integrals. What is more, note that for the Caputo-Fabrizio derivative, the order $\alpha \in [0,1)$. This does not mean, however, that one can only compute fractional derivatives of order 1 or lower. There exists a rather convenient property of the differintegral operator which allows one to combine orders of derivatives, which will be discussed in a moment.

Lastly, we will define the Grünwald-Letnikov derivative, which is defined as follows:

Definition 3.2.6. The Grünwald-Letnikov fractional derivative of order α is defined as:

$$D^\alpha f(x) = \lim_{h \to 0} \frac{1}{h^\alpha} \sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} f(x - kh) \tag{4}$$

Where $\binom{\alpha}{k}$ is the binomial coefficient, with $0 \le k \le \alpha$, see section Appendix A.

Zhmakin (2022). It is immediately clear, observing the summmation symbol instead of the integral, that this derivative behaves quite differently from the two derivatives defined above. The Grünwald-Letnikov derivative is an extension on derivatives based of the concept of finite differences Flajolet and Sedgewick (1995).

We will consider a number of properties which come in useful when working with fractional derivatives.

Proposition 3.2.1. *The fractional derivatives above adhere to the following properties:*

(i) Linearity: Let $f(x), g(x)$ be functions and $a, b, x \in \mathbb{R}$. Then we have that $D^\alpha(af(x) + bg(x)) = aD^\alpha f(x) + bD^\alpha g(x)$.

(ii) $D^\alpha f(x) = f(x)$, for $\alpha = 0$

(iii) for sufficiently smooth functions f, we have that $D^{\alpha+\beta} f(x) = D^\alpha(D^\beta f(x)) = D^\beta(D^\alpha f(x))$, with $\alpha, \beta \in \mathbb{R}$. Note that, for definition 3.2.5, this property only holds for $\beta \in \mathbb{N}, \alpha \in [0, 1)$.

The proofs of these of the properties stated in this proposition can be found in Appendix B. Most of these proofs have been provided by myself, while some other proofs, which are outside the scope of thesis are based on other papers. Note that the third property is especially useful for the Caputo-Fabrizio derivative. This property allows one to take fractional derivatives of order greater than 1 by comparing fractional derivatives and regular integer derivatives.

We will now provide two numerical examples of these fractional derivatives. For simplicity, we will let $a = 0$:

Example 3.2.1.   (i) We consider the Riemann-Liouville derivative of order $\frac{3}{2}$ for some constant $c \in \mathbb{R}$:

$$D_{RL}^{\frac{3}{2}}(c) = \frac{d^2}{dx^2} \frac{1}{\Gamma(2 - \frac{3}{2})} \int_0^x (x - t)^{2 - \frac{3}{2} - 1} c \, dt$$

$$= \frac{d^2}{dx^2} \frac{c}{\sqrt{\pi}} \int_0^x (x - t)^{-\frac{1}{2}} dt = \frac{d^2}{dx^2} \frac{-2c}{\sqrt{\pi}} \sqrt{x - t} \Big|_0^x$$

$$= \frac{d^2}{dx^2} \frac{2c\sqrt{x}}{\sqrt{\pi}} = \frac{-c}{2x\sqrt{\pi x}} \neq 0.$$

As stated, the fractional derivative of a constant is not equal to zero when using the Riemann-Liouville derivative. This is also the case for the Grünwald-Letnikov derivative, but not for the Caputo-Fabrizio derivative.

(ii) We compute the semi-derivative of $\frac{x}{2}$ using the Caputo-Fabrizio derivative:

$$D_{CF}^{\frac{1}{2}}\left(\frac{x}{2}\right) = \frac{1}{1 - \frac{1}{2}} \int_0^x \exp\left(\frac{-\frac{1}{2}}{1 - \frac{1}{2}}(x - t)\right) \frac{1}{2} dt = \int_0^x \exp(t - x) dt$$

$$= \exp(t - x)\Big|_0^x = 1 - \exp(-x).$$

For $x \geq 0$, this expression is equal to the CDF of the exponential distribution with $\lambda = 1$. A remarkble result.

The observant reader might notice that no explicit examples of the Grünwald-Letnikov derivative have been provided. The latter is due to the fact that is rather difficult to obtain analytical expressions for this derivative. Thus, later on in this thesis, when computing fractional moments and their assiocated biases, the main focus for the Grünwald-Letnikov derivative will be on its numerical computations.

Remark 3.2.2. As shortly mentioned in the introduction, it is possible to generalize the order of derivatives even further, extending α to be in $\mathbb{C}$ instead of $\alpha \in \mathbb{R}$. This means that, when combining such derivatives with the moment generating function, we will obtain complex moments. Since statisical moments of complex order do almost not find any usage in applications, as they lack interpretability, they are not the main focus of this research. For the interested reader, Love (1971) has done some impressive research on the fundamentals of derivatives of complex order. The obtained expressions are somewhat similar to those of the fractional derivatives which have been discussed above.

4 Moments

Having defined the techniques required to compute fractional derivatives, we now turn to their application to the Moment Generating Function (MGF). Before proceeding, we will briefly review some key concepts related to statistical moments.

4.1 Moments

Definition 4.1.1. Let X be a random variable. The α-th moment of a X is defined as:

$$\mathbb{E}[X^\alpha] = \begin{cases} \int_{-\infty}^{\infty} (x - c)^\alpha f_X(x)\,dx & \text{if } X \text{ is continuous,} \\ \sum_i (x_i - c_i)^\alpha f_X(x_i) & \text{if } X \text{ is discrete.} \end{cases}$$

with, in the context of this paper $\alpha \in \mathbb{R}$. Here $f_X(x)$ is the probability density function if X is a continuous random variable, and the probability mass function if X is a discrete random variable.

If $c = \mu_x$, the first moment of $f(x)$, then our higher moments are called central moments. In the context of this research, we will focus on the case $c = 0$, corresponding to raw moments of a random variable X. This choice has been made as the Moment Generating Function, which we will soon define, only computes raw moments of higher order. A moment of order n is said to exist, if $\mathbb{E}[X^n] < \infty$.

The following proposition is rather standard and well known. Yet, since we will be working with a lot of fractional moments which replace non-existing integer moments, it does not hurt to reiterate these properties.

Proposition 4.1.1. *Let X be a random variable. Then:*

(i) If $\mathbb{E}[X^n]$ does not exist, then neither does $\mathbb{E}[X^k]$, for $n \leq k$.

(ii) If $\mathbb{E}[X^k]$ does exist, then all its lower-order moments, $\mathbb{E}[X^n]$, for $n \leq k$. Also exists.

The proof is rather straightforward and can be found in Appendix B.

4.1.1 Moments of negative order

Since the MGF can be extended to negative orders, we briefly examine how to compute raw moments of negative order using traditional methods. This will allow us to compare results between the two different methods. What is more, as mentioned in section 1, in specific cases, moments of negative order can be computed to obtain statistics such as the Sharpe Ratio. Thus, it is useful to understand how to compute moments of such orders. We will consider the continuous case:

$$\mathbb{E}[X^{-n}] = \int_{-\infty}^{\infty} x^{-n} f_X(x)dx = \int_{-\infty}^{\infty} \left(\frac{1}{x}\right)^n f_X(x)dx.$$

We can immediately observe a rather obvious problem. This integral is undefined at $x = 0$ and diverges in a neighbourhood around zero. Khuri and Casella (2002) have stated the following corollary for the existence of a moment with negative first order:

Corollary 4.1.1. *If $f_X(x)$ is a continuous pdf defined on $(-\infty, \infty)$, and if*

$$\lim_{x \to 0} \frac{f_X(x)}{|x|^\alpha} < \infty$$

for $\alpha > 0$, then

$$\mathbb{E}[X^{-1}] \text{ exists.}$$

Most common distribution functions do not adhere to this corollary, however, the Gamma function does:

Example 4.1.1. Let $f_X(x) \sim \Gamma(\alpha, \lambda) =$

$$\frac{x^{\alpha-1}e^{-\lambda x}\lambda^\alpha}{\Gamma(\alpha)}.$$

This PDF is defined on $(0, \infty)$. So the function is not defined on $\mathbb{R}$. This, however, is not a problem, as we can just evaluate the right limit. Since the Gamma function already uses α as a parameter, we will evaluate $\frac{f_X(x)}{|x|^\beta}, \beta > 0$:

$$\lim_{x \to 0+} \frac{f_X(x)}{|x|^\beta} = \lim_{x \to 0+} \frac{x^{\alpha-1-\beta}e^{-\lambda x}\lambda^\alpha}{\Gamma(\alpha)} = \lim_{x \to 0+} \frac{x^{\alpha-(1+\beta)}e^{-\lambda x}\lambda^\alpha}{\Gamma(\alpha)}.$$

Thus for $\alpha \geq \beta + 1, \lim_{x \to 0+} \frac{f_X(x)}{|x|^\beta} < \infty$. Therefore, the first negative moment of the Gamma distribution should exist.

We will compute the first negative moment:

$$\mathbb{E}[X^{-1}] = \int_0^\infty x^{-1} \frac{x^{\alpha-1}e^{-\lambda x}\lambda^\alpha}{\Gamma(\alpha)} dx = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-2}e^{-\lambda x} dx$$

Using the substitution $u = \lambda x, \frac{du}{dx} = \lambda, dx = \frac{du}{\lambda}$, we get:

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty \left(\frac{u}{\lambda}\right)^{\alpha-2} e^{-u} du = \frac{\lambda^\alpha}{\Gamma(\alpha)\lambda^{\alpha-1}} \int_0^\infty (\frac{u}{\lambda})^{\alpha-2} e^{-u} du.$$

This integral is equal to $\Gamma(\alpha - 1)$ (See Appendix A). So we get:

$$\mathbb{E}[X^{-1}] = \frac{\lambda^\alpha \Gamma(\alpha-1)}{\Gamma(\alpha)\lambda^{\alpha-1}} = \frac{\lambda \Gamma(\alpha-1)}{(\alpha-1)\Gamma(\alpha-1)} = \frac{\lambda}{(\alpha-1)}.$$

Thus, for $\alpha \neq 1, \mathbb{E}[X^{-1}] = \frac{\lambda}{(\alpha-1)}$. Fortunately, this is always the case, since we had just derived that the integral only converges when $\alpha \geq \beta + 1$, with $\beta > 0$. In other words, $\alpha > 1$. So this holds.

4.2 The Moment Generating Function

We now formally introduce the Moment Generating Function, one of the most significant subjects of this thesis.

Definition 4.2.1. The Moment Generating Function (MGF) of a variable X, is defined as

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

provided that $\mathbb{E}[e^{tX}] < \infty$, for all $t \in (-h, h)$, which contains 0, for some $h > 0$.

Remark 4.2.1. Deriving the expression $M_X(t) = \mathbb{E}[e^{tX}]$ is typically straightforward. Generally, it simply requires a handful of steps of analytic evaluation. This procedure is not that interesting nor relevant to this research. We provide a single explicit example on how to compute the Moment Generating Function for a specific distribution below. And for later cases, when we make use of an expression of the Moment Generating Function, we will simply refer to the distribution table in Appendix C.

Example 4.2.1. Let $f_X(x) \sim \Gamma(\alpha, \lambda)$ be the Gamma distribution with PDF:

$$f_X(x) = \frac{x^{\alpha-1} e^{-\lambda x} \lambda^{\alpha}}{\Gamma(\alpha)}.$$

Let $t < \lambda$ (in any other case, the integral diverges). The moment generating function $M_X(t)$ is given by:

$$M_X(t) = \int_0^\infty e^{tx} f_X(x) \, dx = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-1} e^{(t-\lambda)x} \, dx.$$

We make use of the substitution $u = -(t - \lambda)x$, $\frac{du}{dx} = (\lambda - t)$, $dx = \frac{du}{(\lambda - t)}$, so:

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty \left(\frac{u}{\lambda - t} \right)^{\alpha-1} e^{-u} \frac{du}{\lambda - t} = \frac{\lambda^\alpha}{\Gamma(\alpha)(\lambda - t)^\alpha} \int_0^\infty u^{\alpha-1} e^{-u} \, du.$$

This integral is just the definition of the Gamma Function, $\Gamma(\alpha)$, so we obtain:

$$= \frac{\lambda^\alpha \Gamma(\alpha)}{\Gamma(\alpha)(\lambda - t)^\alpha} = \left(\frac{\lambda}{\lambda - t} \right)^\alpha = M_X(t)$$

We will state the theorem which makes the MGF so useful. This theorem allows us to compute moments of higher order by taking derivatives of the given order instead of integrals.

Theorem 4.2.1. *If $M_X(t)$ exists on some interval $(-h, h)$, as defined before, we have that:*

$$\mathbb{E}[X^n] = M_X^{(n)}(0), \text{ for } n \in \mathbb{N}$$

Proof. The proof is important and fairly straightforward, thus it will be shown directly:

$$M_X^{(n)}(t) = \frac{d^n}{dt^n} \int_{-\infty}^\infty e^{tx} f_X(x) dx = \int_{-\infty}^\infty \frac{d^n}{dt^n} e^{tx} f_X(x) dx$$

(We can interchange differentiation and integration since all partial derivatives of $e^{tx} f(x)$ are continuous and the absolute value of the integral converges, as we assume the n-th moment exists, see Appendix A).

$$= \int_{-\infty}^\infty x^n e^{tx} f_X(x) dx, \text{ evaluating at } t = 0 := \int_{-\infty}^\infty x^n e^{0x} f_X(x) dx$$

$$= \int_{-\infty}^\infty x^n f_X(x) dx = \mathbb{E}[X^n]$$

The proof for the case that $f_X(x)$ is discrete is very similar. In that case, one would have to change the order of the derivative and summation, which has also been justified in Appendix A. $\square$

We introduce the following properties for the Moment Generating Function $M_X(t)$:

Proposition 4.2.1. *For X, Y random variables, we have that:*

(i) $M_X^{(0)}(t) = \mathbb{E}[e^{0X}] = 1$. *This property can be used to confirm that a given function is a valid probability density function (i.e., integrates to one).*

(ii) *Location scale-transform. Assuming $M_X(t)$ exists, for constant μ, σ, we have that:*

$$M_{\mu+\sigma X}(t) = e^{\mu t} \cdot M_X(\sigma t)$$

(iii) *If $X \perp Y$, then $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$.*

The proofs of the latter can be found in Appendix B.

There are several additional topics closely related to the MGF, including Fourier transforms, Laplace transforms, Wick rotations, and characteristic functions. While these subjects are relevant to the theoretical foundation of the MGF , they fall outside the scope of this paper and will therefore not be discussed. Readers interested in exploring these concepts further may find Kolmogorov and Fomin (1999) to be a valuable resource.

4.2.1 Computing negative moments using the Moment Generating Function

We have seen that under specific conditions, it is possible to compute negative, or so called inverse moments of distribution functions. This technique can also be applied to the moment generating function. In the 20-th century, Cressie et al. (1981) have published the following remarkable theorem:

Theorem 4.2.2. *Assuming the negative n-th raw moment exists, the negative n-th raw moment can be computed as follows:*

$$\mathbb{E}[X^{-n}] = \frac{1}{\Gamma(n)} \int_0^\infty t^{n-1} M_X(-t) dt$$

where n is a positive integer.

The proof of this Theorem can be found in Appendix B.

Since this is an extension on the regular functions of the MGF, this technique is of interest of this paper. Thus, it will be shortly be discussed. Let us again compute the first inverse moment of the Gamma distribution, but now by making use of the latter theorem for Moment Generating Functions!

Example 4.2.2. Let

$$f_X(x) \sim \Gamma(\alpha, \lambda) = \frac{x^{\alpha-1} e^{-\lambda x} \lambda^\alpha}{\Gamma(\alpha)}, M_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^\alpha$$

$$\mathbb{E}[X^{-1}] = \frac{1}{\Gamma(1)} \int_0^\infty t^{(1-1)} \left(\frac{\lambda}{\lambda - (-t)}\right)^\alpha dt = \int_0^\infty \left(\frac{\lambda}{\lambda + t}\right)^\alpha dt$$

$$= \lambda^\alpha \int_0^\infty (\lambda + t)^{-\alpha} dt, \text{ Let } u = \lambda + t, \frac{du}{dt} = 1, dt = du :$$

$$\lambda^\alpha \int_0^\infty u^{-\alpha} du = \lambda^\alpha \left. \frac{u^{1-\alpha}}{1-\alpha} \right|_0^\infty = \lambda^\alpha \left. \frac{(\lambda+t)^{1-\alpha}}{1-\alpha} \right|_0^\infty$$

$$= \lambda^\alpha \left(0 - \frac{\lambda^{1-\alpha}}{1-\alpha}\right) = \frac{-\lambda}{1-\alpha} = \frac{\lambda}{\alpha-1}.$$

Which corresponds with our result from example 4.1.1.

4.2.2 Extending the MGF to fractional order

Now that we have introduced the definitions of the MGF and discussed some key properties, we will combine these with the techniques developed in section 3. Therefore, we can at last obtain moments of fractional order using the MGF.

To avoid confusion regarding what fractional derivative is being used in combination with the MGF, we will from now on, work with the following notation:

Definition 4.2.2. We define the MGF of order $\alpha \in \mathbb{R}$ by $_{RL}M_X^{(\alpha)}$, $_{CF}M_X^{(\alpha)}$, $_{GL}M_X^{(\alpha)}$ for the MGF in combination with the Riemann-Liouville, Caputo-Fabrizio and Grünwald-Letnikov fractional derivative respectively.

Remark 4.2.2. The three properties mentioned in proposition 4.2.1 still hold for the MGF of fractional order. This is the case as the first property makes makes use of the derivative of order 0. Which has been defined to just be the original function itself as stated in the second property of proposition3.2.1. The other two properties do not involve any derivatives of any order. Thus, they are generally applicable to the MGF, regardless of its order or kind of derivative.

Before stating any results about the accuracy of these new MGF expressions, we first consider a numerical example to illustrate the interaction between fractional derivatives and the MGF.

Example 4.2.3. (i) We let $f_X(x) \sim Bernoulli(p)$, with $\mathbb{P}(X = 1) = p$ and with associated MGF expression: $M_X(t) = (1 - p) + p \cdot \exp(t)$, now we consider

$$_{CF}M_X^{(\frac{1}{2})} = \frac{1}{1 - \frac{1}{2}} \int_{-h}^{t} \exp\left(\frac{\frac{-1}{2}}{1 - \frac{1}{2}}(t - s)\right) f'(s)ds.$$

In this case, the domain of $f(t) = (-\infty, \infty)$, so we let $-h = -\infty$, and $f'(s) = p \cdot \exp(s)$, thus we obtain:

$$_{CF}M_X^{(\frac{1}{2})} = 2 \int_{-\infty}^{t} \exp\left((s - t)\right) \cdot (p \cdot \exp(s))ds.$$

$$= 2p \cdot exp(-t) \int_{-\infty}^{t} \exp(2s)ds$$

$$= p \cdot exp(-t) \left(\exp(2s)\Big|_{-\infty}^{t}\right) = p \cdot \exp(t)$$

Now, all that is left to do is set $t = 0$ and we obtain that $_{CF}M_X^{(\frac{1}{2})} = p$.

(ii) Computing $\mathbb{E}[X^{\frac{1}{2}}]$ in the traditional fashion, we obtain:

$$\mathbb{E}[X^{\frac{1}{2}}] = \sum_x x^{\frac{1}{2}}\mathbb{P}(X = x) = 0^{\frac{1}{2}} \cdot \mathbb{P}(X = 0) + 1^{\frac{1}{2}} \cdot \mathbb{P}(X = 1)$$

$$= 0 \cdot (1 - p) + 1 \cdot p = p$$

It is amazing and maybe even somewhat surprising that both expressions obtain the same result. Indeed, this result is actually more of a coincidence. It is important to note that this agreement of results is coincidental and specific to the chosen distribution. Namely, all raw higher moments of a Bernoulli random variable are p. If we had taken any other distribution in combination with a moment of fractional order, it becomes highly likely that the MGF returns a different value compared to the traditional method of computing moments. What is more, if we were to take

$$_{RL}M_X^{(\frac{1}{2})} = \frac{d}{dt}\frac{1}{\sqrt{\pi}}\int_{-h}^{t}(t - s)^{\frac{-1}{2}}f(s)ds$$

we obtain an integral which may diverge based on the choice of $-h$. These observations lead to the following theorem.

Theorem 4.2.3. *Consider the three MGF's as defined in definition 4.2.2. Assume $_{RL}M_X^{(\alpha)}$ and $_{GL}M_X^{(\alpha)}$ are well defined on some open interval $(-h, h)$, then the moment generating function expressions $_{RL}M_X^{(\alpha)}$ and $_{GL}M_X^{(\alpha)}$ accurately obtain moments of order $\alpha \in \mathbb{R}$*

The proof can be found in Appendix B. Unfortunately, this result does not hold for $_{CF}M_X^{(\alpha)}$, which leads to the following theorem.

Theorem 4.2.4. *Consider the three MGF's as defined in definition 4.2.2. Assume $_{CF}M_X^{(\alpha)}$ is well defined on some open interval $(-h, h)$, then the moment generating function expression $_{CF}M_X^{(\alpha)}$ inaccurately approximates moments of order $\alpha \in \mathbb{R}$ with approximation error given by*

$$\begin{cases} \displaystyle\int_{-\infty}^{\infty} x^\alpha f_X(x)dx - \int_{-\infty}^{\infty} \frac{x^{n+1}}{(1-\beta)x + \beta} f_X(x)dx & \text{if } X \text{ is continuous,} \\ \displaystyle\sum_i \left(x_i^\alpha - \frac{x_i^{n+1}}{(1-\beta)x_i + \beta} \right) f_X(x_i) & \text{if } X \text{ is discrete.} \end{cases}$$

with $\alpha \in \mathbb{R}$, $\beta = \alpha - n$ and $n = \lfloor \alpha \rfloor$.

The proof can be found in Appendix B.

Remark 4.2.3. In the case when $\alpha \in \mathbb{N}$, we have that $n = \alpha$, and thus $\beta = 0$, therefore, $_{CF}M_X^{(\alpha)}$ is accurate for integer orders. This is an expected result, as the MGF for integer moments is accurate and from section 3 we know that $D_{CF}^{\alpha+\beta} f(x) = D_{CF}^\alpha(D_{CF}^\beta f(x))$, with $\alpha \in \mathbb{N}$, $\beta \in [0, 1)$. In this case, let $\beta = 0$. So we get $D_{CF}^{\alpha+0} f(x) = D_{CF}^\alpha f(x)$ which is just a regular derivative of integer order.

4.2.3 Analysing the error of the Caputo-Fabrizio MGF

We now further inspect the error obtained in theorem 4.2.4. Note that in the following plots, we explicitly focus on the term

$$E(x, \alpha) = x^\alpha - \frac{x^{n+1}}{(1-\beta)x + \beta}$$

for simplicity. The expression we analyse, denoted $E(x, \alpha)$ is not the same as the expression obtained in theorem4.2.4. Yet, it is the part of the expression that involves the order α and is thus of great interest. What is more, considering the result obtained in theorem 4.2.4, it is easy to see that $_{CF}M_X^{(\alpha)}$ is accurate $\iff x^\alpha = \frac{x^{n+1}}{(1-\beta)x + \beta}$. Thus, we consider the function $E(x, \alpha)$, the function of their differences and analyse when this function is equal to 0. Errors of the Caputo-Fabrizio MGF as defined in theorem 4.2.4 will be explicitly numerically computed in the following section of this paper, where we will consider a multitude random variables, with different distributions. Plotting this expression for different orders of α, we obtain the following figure:
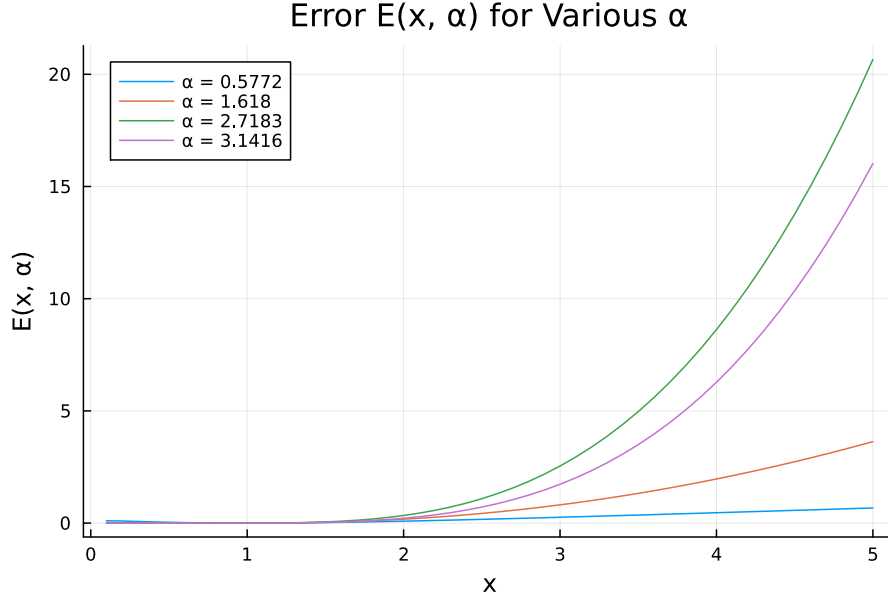
Figure 1: Error of Caputo-Fabrizio MGF

It can be observed from the figure above that, in general for greater orders α, the error tends to increase. However, the greatest order of α does not provide the greatest error. Instead it is the second greatest order of α that does so. This may be explained by the fact that for values of α in the middle of two integers, the value of $(1 - \beta)x + \beta$ will be the greatest. As a result, the fraction becomes small, so the error tends to increase.

This hypothesis is indeed confirmed in the following figure, which focuses on the growth rate of the error, for a fixed value of x and an increasing order α.

Figure 2: Error of Caputo-Fabrizio MGF, for fixed x

Figure 2 indeed seems to indicate that for some $\alpha \in (a, b)$, where $a, b \in \mathbb{Z}$ consecutive integers that $E(x, \alpha)$ attains it maximum value for some α 'slightly greater' than $\frac{a+b}{2}$. Figure 2

also reveals that the approximation error is considerably smaller for negative fractional moments compared to the positive ones. This suggests that, when negative moments are well-defined, the Caputo-Fabrizio MGF might still offer reasonable accuracy, despite its known limitations.

Note that within each integer interval, the function $E(x, \alpha)$ appears to be concave. Thus, as is supported by figure 2, it is only possible to obtain local maxima. Ideally, one would aim to analytically determine the optimal order $\hat{\alpha}$ such that minimizes $E(x, \hat{\alpha})$. However, due to the concavity of $E(x, \hat{\alpha})$, such analyticcal minimization is not feasible. Figure 2 seems to suggest that order of α near integer orders yields smaller approximation errors, which makes sensse intuitively. However, the latter is not rigorous evidence. Therefore, in the following section, we will analyse the numerical approximation errors of the Caputo-Fabrizio MGF by simulation. We will use numerical methods to try to obtain values α which minimize the expression from theorem 4.2.4.

5 Simulation

We have seen in the previous section how the error $E(x, \alpha) = x^\alpha - \dfrac{x^{n+1}}{(1 - \beta)x + \beta}$ changes with increasing values of α. Now, we focus on the entire expression, as stated in theorem 4.2.4.

5.1 Practical Issues

Computing these expressions poses some difficulties. In the first place, we want to compute the moment of fractional order α for a random variable X. Julia does not have any packages to compute raw moments of such an order, thus these moments will have to be created first. In order to accomplish the desired result, we directly compute $\int_{-\infty}^{\infty} x^\alpha f_X(x) dx$. If the support of the random variable is infinite, this integral is often not numerically well defined. Thus, in such cases we use a bounded support $(-1000, 1000)$ which provides a sufficient accurate approximation. What is more, for many fractional orders α in combination with negative values of x, x^α is of complex form. To avoid such expressions, given that the distribution is symmetric, we simply integrate the PDF over its non-negative support and factor the final value by 2. In this manner, it is still possible to obtain fractional moments of distributions such as the Normal distribution. If a distribution has a negative support and is not symmetric, it is possible to compute the complex or absolute moment of fractional order, which avoids any numerical issues. This approach, however, lacks interpretability, as the intuition behind absolute and complex moments is harder to grasp.

It now remains to compute the fractional moment using the Caputo-Fabrizio MGF. Unfortunately, it is rather difficult to compute the fractional derivative of an MGF expression numerically. The reason is, that programming packages often take (fractional) derivatives on a certain point x rather than over the entire function. Thus, we simply use the result obtained in theorem 4.2.4. That is, instead of the moment of order α, the Caputo-Fabrizio MGF computes the moment of order $n + 1$ divided by some variables dependent on α. More explicitly, we compute $_{CF}M_X^{(\alpha)}(0) = \int_{-\infty}^{\infty} \dfrac{x^{n+1}}{(1 - \beta)x + \beta} f_X(x)$. This integral faces the some problems as the aforementioned integral with x^α. What is more, for specific values of x and β, the denominator tends to go towards zero, which is numerically unstable. Thus, in such cases, we add a rather small ϵ in order to avoid this issue.

5.2 Accuracy Analysis

For this accuracy analysis we consider different parameter values and orders of α. We set $\alpha \in [0, 5]$ with steps of 0.01. This yields 500 errors per parameter-configuration, which is a reasonably reliable sample size. When possible, we extend α to the interval $(-1, 0)$. We consider three

15

distributions, namely the Exponential, Normal and Poisson distribution. This selection allows us to cover both continuous and discrete distributions, distributions with negative and positive supports and symmetric distributions.

5.2.1 Exponential Distribution

First we consider the Exponential distribution with PDF as defined in Appendix C. The support of the Exponential distribution is $[0, \infty)$, thus the integration process will avoid most numerical issues. Moreover, all relevant moments of the Exponential distribution are based on raw moments, which is precisely the kind of moments we are interested in.
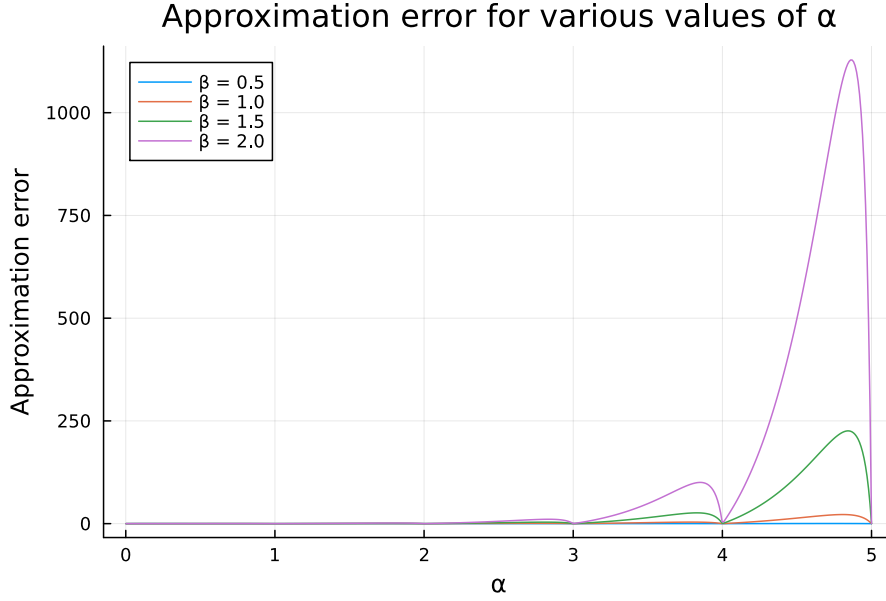We obtain the following figure:

Figure 3: Approximation error for the Exponential Distribution

From this figure, it is clear that for a greater parameter value β the approximation error increases. Namely, for moments of higher order ($\alpha \geq 3$), the approximation error increases rapidly. For smaller order α the approximation errors remain relatively low. In practice, moments of order $\alpha > 4$ are oftentimes not of interest. Thus, the greatest approximation errors will be avoided.

We obtain the following associated table with some core statistics:

β	minimum	maximum	mean	standard deviation	skewness	c_v
0.5	0.0	0.329	0.074	0.083	1.762	1.12
1.0	0.0	22.119	2.907	5.612	2.314	1.93
1.5	0.0	225.731	25.477	54.91	2.513	2.155
2.0	0.0	1128.26	115.489	265.154	2.655	2.296

Table 1: Exponential Distribution - Approximation Error Statistics

The results in table 1 align with the conclusions derived in figure 3. The greater the values of β and α, the greater the (average) approximation error. For all values of β, the minimum error is zero, which occurs when $\alpha \in \mathbb{N}$. An interesting observation is that for $\beta = 1.5$ the average approximation error exceeds the maximum approximation error for $\beta = 1.0$. For all values of

β, the skewness is positive, implying that the distribution of the approximation errors is right-skewed. Thus, clearly the distribution of the errors is not symmetric and therefore certainly not a Normal distribution. In the last column, the Coefficient of Variation has been reported, which has been given by the formula $\frac{\sigma}{\mu}$, where μ and σ denote the sample and sample standard deviation respectively **??**. This metric of measurement allows us to objectively compare the variability of the errors of each of the parameter configurations. The lower c_v, the lower the variability and thus the more consistent the dataset. This consistency is something we want to obtain, as it allows us to describe our data with greater confidence. From table 1, it is clear that the Coefficient of Variation increases as the parameter β increases. For $\beta = 0.5$ all result are reasonable. The Caputo-Fabrizio MGF in combination with the Exponential(0.5) could still be viable in practice. For Exponential(1), this conclusion may depend on the maximum order of α.

5.2.2 Normal Distribution

We now consider X to be Normally distributed with mean μ and standard deviation σ, as defined in Appendix C. This distribution is of particular interest as it has two parameters. Thus, we can analyse the differences of effect of both parameters on the approximation error. Therefore, while raw fractional moments of the Normal distribution are not commonly used in practice, we can still obtain valuable insights from this simulation. The Normal distribution is symmetric with domain $x \in \mathbb{R}$. So, we integrate the PDF over the interval $(0, 1000)$ to avoid numerical issues.
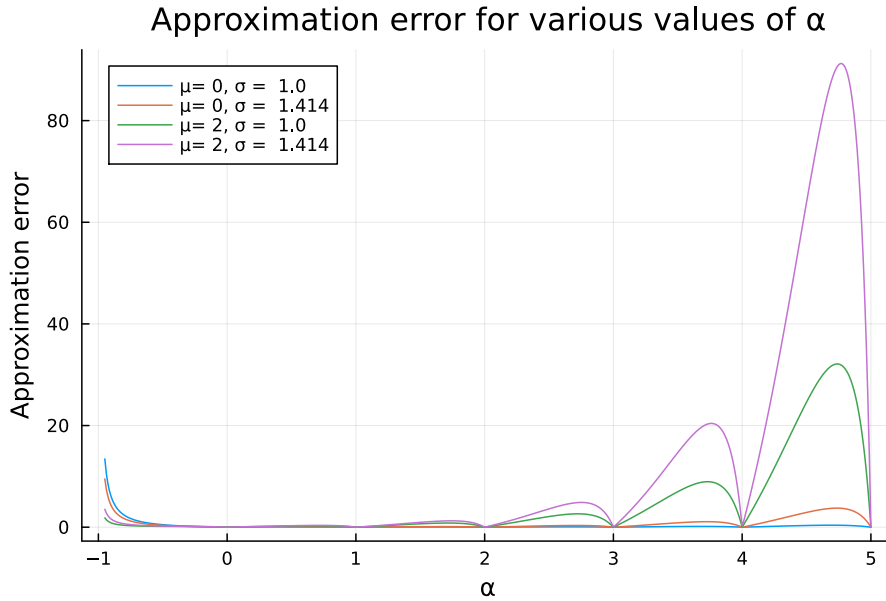
We obtain the following figure:

Figure 4: Approximation error for the Normal Distribution

As illustrated in figure 3, the approximation error increases with μ and σ. This result is consistent with the result obtained for the Exponential distribution. However, the magnitude of the errors of the Normal distribution is considerably smaller. Furthermore, the figure also displays the errors for order $\alpha \in (-0.95, 0)$. These errors tend to attain the same size as errors for order α near 3. Interestingly, for negative orders of α, the effect of the parameter values seems to reverse. Larger values of μ and σ are associated with smaller approximation errors.

We obtain the following associated table with some core statistics:

μ	σ	minimum	maximum	mean	standard deviation	skewness	c_v
0.0	1.0	0.0	13.439	0.298	1.07	7.898	3.591
0.0	1.414	0.0	9.505	0.7	1.141	2.824	1.629
2.0	1.0	0.0	32.123	4.517	8.055	2.269	1.783
2.0	1.414	0.0	91.231	11.411	22.404	2.428	1.963

Table 2: Normal Distribution - Approximation Error Statistics

The statistics in table 2 support the visuals conclusions drawn from figure 4. The average of the approximation error for $\mu = 0$ is acceptable. However, the size of its Coefficient of Variation compared to the other parameter configurations is remarkable. Indeed, configuration $(\mu, \sigma) = (0, 1)$ and configuration $(\mu, \sigma) = (0, \sqrt{2})$ have a rather similar standard deviation, however, but the latter has approximately twice the average error. Except for this case, c_v again tends to increase as the parameter values increase, suggesting higher variability in approximation errors. Similar to the Exponential distribution, for all parameter-configurations, the distribution of the errors is right-skewed. Thus, the approximation errors of a Normal distribution are not Normally distributed themselves! This result is confirmed in figure 5. Even in the histogram of the parameter-configuration with the greatest parameter values, more than 50 percent of the approximation errors cluster around zero, with only a small number of observations falling in the range $(40, 80)$. As before, the average errors and variability for large parameters may be too great to be reliable depending on the context of application.
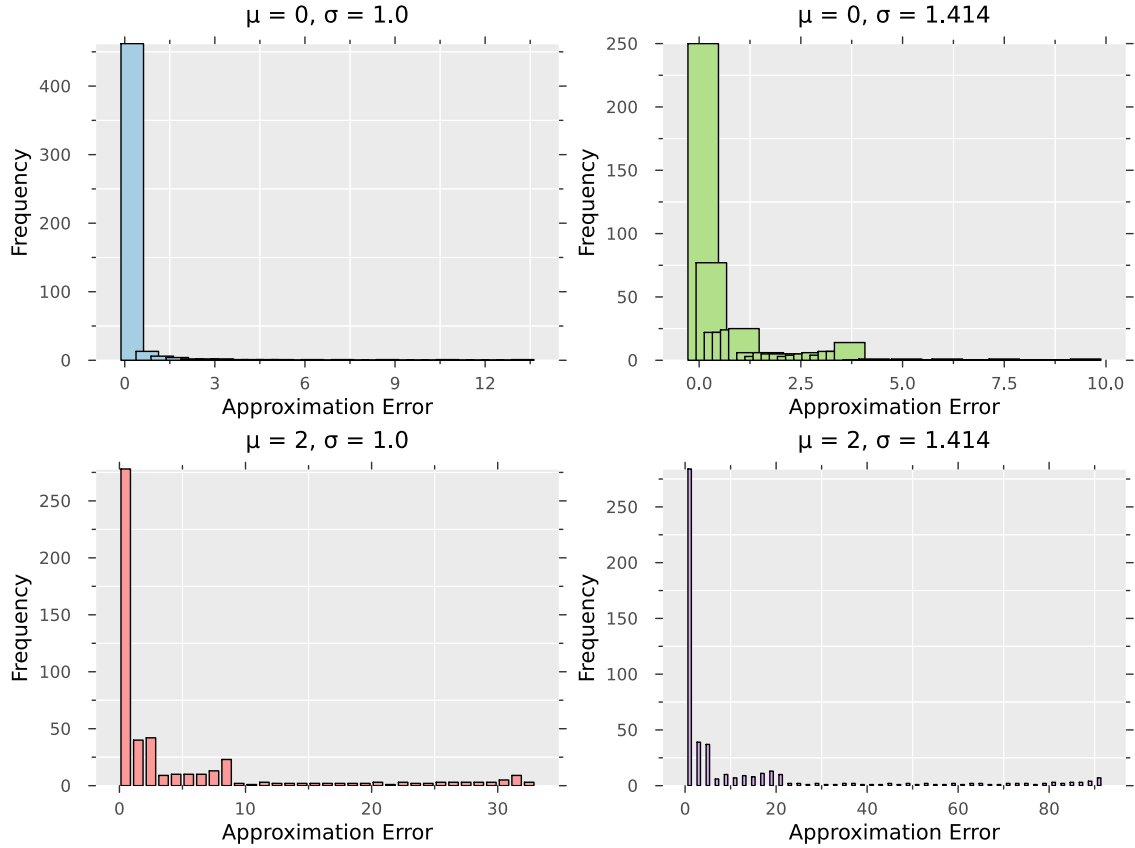

Figure 5: Error Histogram for the Normal Distribution

5.2.3 Poisson Distribution

Finally, we consider a discrete distribution to explore if these results differ from those of the continuous distributions. Specifically, we let $X \sim Poisson(\lambda)$, with the probaility mass function defined as in Appendix C. The Poisson distribution has positive support and its raw moments are given by:

$$\mathbb{E}[X^k] = \sum_{i=0}^{k} \lambda^i \begin{Bmatrix} k \\ i \end{Bmatrix}$$

with $k \in \mathbb{N}$ and where $\{\}$ denotes the Stirling numbers of the second kind (rather similar to the binomial coefficient) Haight (1967). This formula highlights that we cannot consider moment of negative order, as the expression relies on factorials and other combinatorial terms.
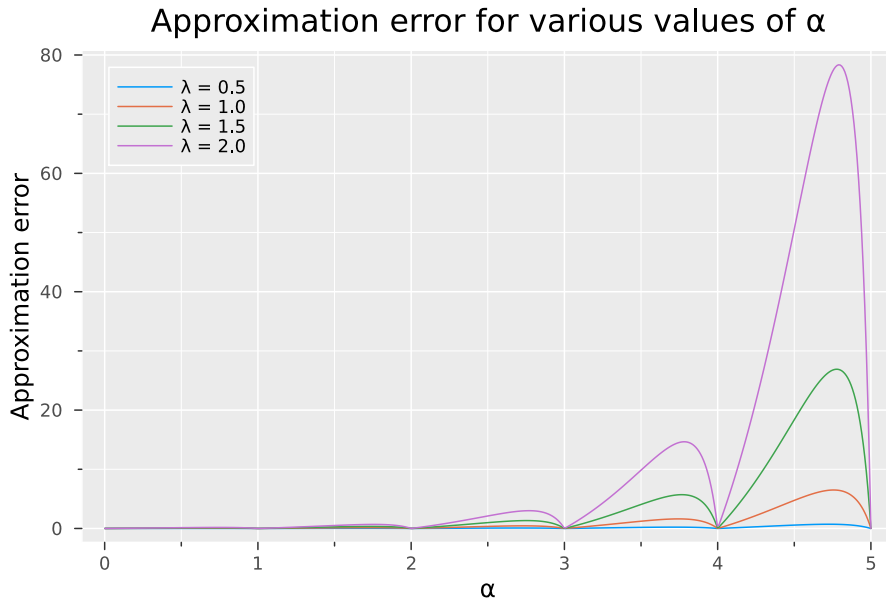
We obtain the following figure:

Figure 6: Approximation error for the Poisson Distribution

As shown in figure 6, the general trend is consistent with that observed for continuous functions. Larger values of λ lead to greater approximation errors. The magnitude of the errors is comparable to that of the Normal distribution, and increases significantly for $\alpha \geq 3$.

λ	minimum	maximum	mean	standard deviation	skewness	c_v
0.5	0.0	0.715	0.128	0.191	1.915	1.494
1.0	0.0	6.5	1.021	1.713	2.075	1.679
1.5	0.0	26.904	3.895	7.001	2.176	1.797
2.0	0.0	78.34	10.687	20.143	2.251	1.885

Table 3: Poisson Distribution - Approximation Error Statistics

The statistics in table 3 further support this conclusion As with previous distributions, the distribution of the errors is right-skewed and the Coefficient of Variation increases as the parameter, λ, grows. For $\lambda = 0.5$ and $\lambda = 1.0$, we can still obtain reliable computations. For $\lambda = 1.5$, the results are still potentially viable, depending on the application and the importance of higher-order

fractional moments. In order to minimize its associated standard deviation, one can consider a smaller maximum order of α. In contrast, the configuration $\lambda = 2.0$ is likely be too unreliable in practice, especially for order $\alpha > 3$.

5.3 General results

Based on the three distributions analysed, we arrive at the following general conclusions:

- As the parameter value of the underlying distribution increase, both the average and maximum approximation error tend to increase.

- For all distributions considered, higher parameter values, lead to greater standard deviations in the approximation errors.

- Furthermore, the Coefficient of Variationor also increases with larger parameter values, indicating that the standard deviation increases at a higher rate than the mean - a stronger result than the increase in the standard deviation alone.

- For all considered distributions and parameter values, the distribution of the approximation errors is right-skewed. Thus, we can conclude that the latter does not follow a Normal, or any other symmetric, distribution.

These findings raise important considerations regarding the reliability of the Caputo-Fabrizio MGF in practice. While high parameter values often result in a large Coefficient of Variation, suggesting that the quality of the approximation may behaves inconsistently, figure 5 offers a different perspective. Namely, in the case of the Normal distribution, at least 50 percent of the errors are clustered around zero. Most importantly, in this section, we are considering a sample of approximation errors of 500 fractional moments. When applying moments of fractional order in practice such as in Hansen and Tong (2024), Mikosch et al. (2013) or Gzyl et al. (2013), one usually computes no more than five moments. Considering the latter, metrics such as the standard deviation and Coefficient of Variation may not always be relevant or decisive when evaluating the usability of the Caputo-Fabrizio MGF. Therefore, rejecting a parameter configuration solely based on its high variability may be too strict. In general, as long as the fractional order of the moment remains below 3 or is close to an integer, the Caputo-Fabrizio MGF tends to yield sufficiently accurate results. This supports its viability in practical cases, such as in Hansen and Tong (2024), where only fractional moments of relatively low order are of interest.

6 Conclusion

The conclusion summarizes the results of your research. Also it gives possible directions of further research.

Acknowledgements

Place the acknowledgments section, if needed, after the main text, but before any appendices and the references. The section heading is not numbered.

A   (Appendix): Relevant functions and Identities

We define the (Euler-)Gamma function as follows:

Definition A.0.1. for $\Re(z) > 0$, we have the following: $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$

The Gamma function can be seen as an extension of the factorial function, for non-integers. This function is defined for complex numbers and all there subsets (so also real numbers), as long as the condition above holds. For positive integers values z, we have the following identity: $\Gamma(z) = (z-1)!$ Other important identities, not necessarily for z an integer, are:

- $\Gamma(z+1) = z\Gamma(z)$

- $\Gamma(2) = \Gamma(1) = 1$

- $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

Definition A.0.2. The falling factorial is defined as follows: $(x)_n = \prod_{k=0}^{n-1}(x-k)$, which is a polynomial

Definition A.0.3. For $0 \leq k \leq n$, the Binomial Coefficient is defined as follows: $\binom{n}{k}$, where $n, k \in \mathbb{N}$.

We can derive the following factorial identity, which is convenient to work with analytically: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. For numerically computing expressions containing the Binomial Coefficient, the following identity is computationally more efficient: $\binom{n}{k} = \frac{(n)_k}{k!}$. With $(n)_k$ as in A.0.2. Since we have established in A.0.1 that $\Gamma(z) = (z-1)!$, we can rewrite our factorial identify to:

$$\binom{n}{k} = \frac{\Gamma(n+1)}{\Gamma(k+1) \cdot \Gamma(n-k+1)} = \frac{n}{k} \cdot \frac{\Gamma(n)}{\Gamma(k) \cdot \Gamma(n-k+1)}$$

Definition A.0.4. The binomial series is a generalization of the binomial formula, namely:

$$(1+x)^\alpha = \sum_{k=0}^\infty \binom{\alpha}{k} x^k$$

Definition A.0.5. Vandermonde's identity: for non-negative integers, k, l, m, n, we have that

$$\sum_{k=0}^l \binom{m}{k} \cdot \binom{n}{l-k} = \binom{m+n}{l}$$

A modification on the latter identity has been called the Chu-Vandermonde identity. This is the same identity, but it his been proven that the identities still hold for complex values m, n as long as l is a positive integer (Askey (1975)).

For the interchange-ability of derivatives and integrals and sums, we can apply the following two theorems:

Theorem A.0.1. *Leibnitz's Rule: Let $f(x,\theta), a(\theta), b(\theta)$ be differentiable with respect to θ, then we have that:*

$$\frac{d}{d\theta} \int_{a(\theta)}^{b(\theta)} f(x,\theta)dx = f(b(\theta),\theta)\frac{db(\theta)}{d\theta} - f(a(\theta),\theta)\frac{da(\theta)}{d\theta} + \int_{a(\theta)}^{b(\theta)} \frac{\partial f(x,\theta)}{\partial \theta}dx.$$

For the special case, where $a(\theta), b(\theta)$ are constant we have that:

$$\frac{d}{d\theta} \int_a^b f(x,\theta)dx = \int_a^b \frac{\partial f(x,\theta)}{\partial d\theta}.$$

For the interchange-ability of derivatives and summations, the following theorem has been given by Casella and Berger (2002):

Theorem A.0.2. *Suppose that the series $\sum_{x=0}^{\infty} h(\theta, x)$ converges for all θ in an interval (a, b) of real numbers and*

(i) $\frac{\partial h(\theta, x)}{\partial \theta}$ *is continuous for all x*

(ii) $\sum_{x=0}^{\infty} \frac{\partial h(\theta, x)}{\partial \theta}$ *converges uniformly on every closed bounded subinterval of (a, b)*

Then:

$$\frac{d}{d\theta} \left(\sum_{x=0}^{\infty} h(\theta, x) \right) = \sum_{x=0}^{\infty} \frac{\partial h(\theta, x)}{\partial \theta}$$

B (Appendix): Proofs

B.1 Proofs section 3

Proof. Proof of 3.2.1

(i) We will proof for the Riemann-Liouville derivative, the proof for the Caputo-Fabrizio derivative is very similar and the Grünwald-Letnikov derivative is a direct consequence of the linearity of the sum.

$$D^\alpha(af(x) + bg(x)) = \frac{d^n}{dx^n}\frac{1}{\Gamma(n-\alpha)}\int_0^x (x-t)^{n-\alpha-1}(af(t) + bg(t))dt$$

$$= \frac{d^n}{dx^n}\left(\frac{a}{\Gamma(n-\alpha)}\int_0^x (x-t)^{n-\alpha-1}f(t)dt + \frac{b}{\Gamma(n-\alpha)}\int_0^x (x-t)^{n-\alpha-1}g(t)dt\right)$$

Where we simply split the integral and put the constants in front.

$$= \frac{d^n}{dx^n}\frac{a}{\Gamma(n-\alpha)}\int_0^x (x-t)^{n-\alpha-1}f(t)dt + \frac{d^n}{dx^n}\frac{b}{\Gamma(n-\alpha)}\int_0^x (x-t)^{n-\alpha-1}g(t)dt$$

As the regular derivative operator is linear.

$$= aD^\alpha f(x) + bD^\alpha g(x)$$

(ii) Intuitively, this makes perfect sense, as the 0-th derivative is just no derivative, so just the function $f(x)$. But for these derivatives, a little bit more effort is required to prove this rather obvious fact.
For the Grünwald-Letnikov derivative we get:

$$D^0 f(x) = \lim_{h\to 0}\frac{1}{h^0}\sum_{k=0}^\infty (-1)^k \binom{0}{k}f(x-kh) = \lim_{h\to 0}\frac{1}{1}\sum_{k=0}^\infty (-1)^k \frac{0!}{k!(0-k)!}f(x-kh).$$

The factorial identity of the binomial coefficient only holds for $0 \le k \le \alpha$. Since $\alpha = 0$ and k is always a positive integer lesser or equal to $\alpha, k = 0$. Thus, we get:

$$= \lim_{h\to 0}\sum_{k=0}^\infty (-1)^0 \frac{0!}{0!(0-0)!}f(x-0h) = \lim_{h\to 0} f(x-0h) = f(x).$$

For the Caputo-Fabrizio derivative, we obtain the following:

$$D^0 f(x) = \frac{1}{1-0}\int_0^x \exp\left(\frac{0}{1-0}(x-t)\right)f'(t)dt = \int_0^x f'(t)dt = f(x).$$

Finally, for the Riemann-Liouville derivative, we can simply make use of 3.2.3 and 3.2.1 to note that in this context $\alpha = 0$ is included in the natural integers. So $D^\alpha = \frac{d^\alpha}{dx^\alpha}f(x) = d^0 dx^0 f(x) = f(x)$ by the first fundamental theorem of calculus.

(iii) The proof for the Riemann-Liouville derivative is given by Koning (2015). And the proof for the Caputo-Fabrizio derivative is given by Losada and Nieto (2015).For the Grünwald-Letnikov derivative, we get:

$$D^\alpha(D^\beta f(x)) = \lim_{h\to 0}\frac{1}{h^\alpha}\sum_{k=0}^\infty (-1)^k \binom{\alpha}{k}(\frac{1}{h^\beta}\sum_{l=0}^\infty (-1)^l \binom{\beta}{l}f(x-lh-kh))$$

23

$$= \lim_{h \to 0} \frac{1}{h^{\alpha+\beta}} \sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} \sum_{l=0}^{\infty} (-1)^l \binom{\beta}{l} f(x - (k+l)h).$$

We substitute $m = k + l$ to deal with the dubble sums:

$$\lim_{h \to 0} \frac{1}{h^{\alpha+\beta}} \sum_{m=0}^{\infty} f(x - mh) \sum_{k=0}^{m} (-1)^k (-1)^{m-k} \binom{\alpha}{k} \binom{\beta}{m-k}$$

Now we make use of an identify from Appendix A to obtain:

$$= \lim_{h \to 0} \frac{1}{h^{\alpha+\beta}} \sum_{m=0}^{\infty} (-1)^m \binom{\alpha+\beta}{m} f(x - mh) = D^{\alpha+\beta} f(x).$$

It can be shown in an exactly similar way that the latter expression is equal to $D^{\beta}(D^{\alpha} f(x))$.

$\square$

B.2 Proofs section 4

Proof. Proof of proposition 4.1.1

(i) If X is a discrete random variable, we get:

$$\mathbb{E}[X^n] = \sum_i x_i^n f_X(x_i) = \infty, \mathbb{E}[X^k] = \sum_i x_i^k f_X(x_i)$$

$$= \sum_i x_i^n \cdot x_i^{k-n} f_X(x_i) \geq \sum_i x_i^n f_X(x_i) = \infty, \text{ as } k \geq n$$

We can prove this for continuous random variables in a similar manner.

(ii) This simply follows from the previous proposition, as the current proposition is just the contrapositive statement of the previous proposition.

$\square$

Proof. Proof of Proposition 4.2.1

(i) This is trivial. For X a continuous random variable, we get:

$$M_X^{(0)}(t) = \int_{-\infty}^{\infty} x^0 e^{tx} f_X(x) dx = \int_{-\infty}^{\infty} 1 e^{0x} f_X(x) dx = \int_{-\infty}^{\infty} f_X(x) dx.$$

Assuming that $f(x)$ is a PDF, this integrates to 1 by definition. If this integral is not equal to 1, this implies that $f(x)$ is not a PDF. The proof for the discrete case is the same but with a summation instead of an integral sign.

(ii)
$$M_{\mu+\sigma X}(t) = \mathbb{E}[e^{(\mu+\sigma X)t}] = \mathbb{E}[e^{\mu t} \cdot e^{\sigma X t}].$$

Since this is the expectation of x, every term that is not dependent on x can be taken out of the summation:
$$= e^{\mu t} \cdot \mathbb{E}[e^{\sigma X t}] = e^{\mu t} \cdot M_X(\sigma t)$$

(iii)

$$M_{X+Y}(t) = \mathbb{E}[e^{(X+Y)t}] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{(x+y)t} f_{X,Y}(x,y) dx dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{xt} \cdot e^{yt} f_{X,Y}(x,y) dx dy.$$

$f_{X,Y}(x,y)$ is the joint pdf for X, Y. But we know that the latter is equal to $f_X(x) \cdot f_Y(y)$, if X, Y are independent. Thus we get:

$$= \left(\int_{-\infty}^{\infty} e^{xt} f_X(x) dx \right) \left(\int_{-\infty}^{\infty} e^{yt} f_Y(y) dy \right) = M_X(t) \cdot M_Y(t).$$

$\square$

Proof. Proof of Theorem 4.2.2: Suppose for the moment that X is a positive random variable. Since $x \cdot f_X(x)$ is integrable for $x > 0$, we have:

$$\mathbb{E}(X) = \int_0^{\infty} x \, dF(x) = \int_0^{\infty} \int_0^{\infty} e^{tx} \, dt \, dF(x).$$

We can interchange the order of integration as follows:

$$\mathbb{E}(X) = \int_0^{\infty} e^{tx} \, dF(x) \, dt = \int_0^{\infty} M_X(-t) \, dt.$$

The interchange of the order of integration is subject to $\mathbb{E}(e^{-tX})$ being integrable from $t = 0$ to $t = \infty$.

Finally, by substituting X^{-1} for X, we find:

$$\mathbb{E}(X^{-1}) = \int_0^{\infty} M_X^{-1}(-t) \, dt.$$

There are two natural ways to generalize (1) to $\mathbb{E}(X^{-1})$; one way gives:

$$\mathbb{E}(X^{-n}) = \int_0^{\infty} \int_0^{\infty} \cdots \int_0^{\infty} M_X(-t_n) \, dt_n \cdots dt_2 dt_1, \tag{2}$$

while the second way gives:

$$\mathbb{E}(X^{-n}) = \frac{1}{\Gamma(n)} \int_0^{\infty} t^{n-1} M_X(-t) \, dt.$$

Cressie et al. (1981)

$\square$

Proof. Proof Theorem 4.2.3 We consider X to be a continuous variable. The three MGF expressions are accurate if $\mathbb{E}[X^\alpha] - M_X^{(\alpha)}(0) = 0$, in other words, if

$$\int_{-\infty}^{\infty} (x^\alpha - D^\alpha e^{tx}) \cdot f_X(x) dx = 0 \iff x^\alpha = D^\alpha e^{tx}$$

where D^α the differential operator of order α. Note that we are of course taking derivatives w.r.t. t, not x. The proof will be shown for the Grünwald-left derivative.

(i)

$$_{GL}M_X^{(\alpha)} = D_{GL}^\alpha(e^{tx}) = \lim_{h\to 0}\frac{1}{h^\alpha}\sum_{k=0}^{\infty}(-1)^k\binom{\alpha}{k}\exp(x(t-kh))$$

$$= \exp(xt)\lim_{h\to 0}\frac{1}{h^\alpha}\sum_{k=0}^{\infty}(-1)^k\binom{\alpha}{k}\exp(-xh)^k$$

$$= \exp(xt)\lim_{h\to 0}\frac{1}{h^\alpha}(1-\exp(-xh))^\alpha$$

(where we used the binomial series identity defined in A.0.4). Now we use the Taylor expansion of $\exp(-xh) = 1 - xh + \frac{(xh)^2}{2!} - \frac{(xh)^3}{3!} + \ldots - \ldots$. Since $h \to 0$, we get: $\exp(-xh) = 1 - xh + \mathcal{O}(h^2)$ (the rest of the terms are negligible). Thus we get:

$$\exp(xt)\lim_{h\to 0}\frac{1}{h^\alpha}(1-(1-xh+\mathcal{O}(n)))^\alpha = \exp(xt)\lim_{h\to 0}\frac{1}{h^\alpha}(h(x+\mathcal{O}(h)))^\alpha$$

$$= \exp(xt)\lim_{h\to 0}\frac{1}{h^\alpha}(h^\alpha((x+\mathcal{O}(h)))^\alpha) = \exp(xt)\lim_{h\to 0}(x+\mathcal{O}(h))^\alpha$$

$$= \exp(xt)x^\alpha.$$

Finally, we take a value of t around 0 and obtain $_{GL}M_X^{(\alpha)}(0) = x^\alpha$

(ii) We consider the Riemann-Liouville derivative:

$$D_{RL}^\alpha(e^{tx}) = \frac{d^n}{dt^n}\frac{1}{\Gamma(n-\alpha)}\int_{-\infty}^{t}(t-s)^{n-\alpha-1}e^{sx}ds$$

Koning (2015) has shown that this derivative is equal to $x^\alpha e^{tx}$. Thus, if we take a value of t around 0, we get: $_{RL}M_X^{(\alpha)}(0) = x^\alpha$

The proof also holds for the case when X is a discrete random variable. $\qquad\square$

Proof. Proof of Theorem 4.2.4 We compute

$$_{CF}M_X^{(\alpha)} = D_{CF}^\alpha(e^{tx}) = \frac{d^n}{dt^n}\frac{1}{1-\beta}\int_{-\infty}^{t}\exp\left(\frac{-\beta}{1-\beta}(t-s)\right)x\exp(xs)ds$$

$$= \frac{d^n}{dt^n}\frac{x}{1-\beta}\exp\left(\frac{-\beta t}{1-\beta}\right)\int_{-\infty}^{t}\exp(s\left(\frac{\beta}{1-\beta}+x\right))ds \text{ where } \beta = \alpha - n \text{ and } n = \lfloor\alpha\rfloor.$$

Now let $u = \frac{\beta}{1-\beta} + x$, so we get:

$$\frac{d^n}{dt^n}\frac{x}{1-\beta}\exp\left(\frac{-\beta t}{1-\beta}\right)\int_{-\infty}^{t}\exp(us)ds = \frac{d^n}{dt^n}\frac{x}{1-\beta}\exp\left(\frac{-\beta t}{1-\beta}\right)\frac{1}{u}exp(us)\Big|_{-\infty}^{t}$$

$$= \frac{x}{1-\beta}\exp\left(\frac{-\beta t}{1-\beta}\right)\cdot\frac{1}{u}\cdot\exp(ut) = \frac{x\exp(xt)}{(1-\beta)x+\beta}$$

Now we apply the derivative of the above expression, with respect to t, n times:

$$\frac{d^n}{dt^n}\frac{x\exp(xt)}{(1-\beta)x+\beta} = \frac{x^{n+1}\exp(xt)}{(1-\beta)x+\beta}$$

Now we take a value of t around 0 and obtain:

$$_{CF}M_X^{(\alpha)} = D_{CF}^{\alpha}(e^{tx}) = \frac{x^{n+1}}{(1-\beta)x + \beta}.$$

The latter expression is not equal to x^{α}, thus the error of $_{CF}M_X^{(\alpha)}$ is:

$$x^{\alpha} - \frac{x^{n+1}}{(1-\beta)x + \beta}$$

$\square$

C   Appendix C

Distribution	PDF / PMF	MGF $M_X(t)$	Restrictions
Bernoulli(p)	$f(x) = p^x(1-p)^{1-x}$	$M_X(t) = (1-p) + pe^t$	$0 \leq p \leq 1,\ x \in \{0,1\}$
Binomial(n,p)	$f(x) = \binom{n}{x}p^x(1-p)^{n-x}$	$M_X(t) = (1-p+pe^t)^n$	$n \in \mathbb{N},\ 0 \leq p \leq 1$
Discrete Uniform	$f(x) = \frac{1}{N}$	$M_X(t) = \frac{1}{N}\sum_{k=1}^{N} e^{tk}$	$N \in \mathbb{N}$
Geometric(p)	$f(x) = p(1-p)^{x-1}$	$\frac{pe^t}{1-(1-p)e^t},\ t < -\ln(1-p)$	$0 < p \leq 1,\ x \geq 1$
Poisson(λ)	$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$	$\exp(\lambda(e^t - 1))$	$\lambda > 0,\ x \geq 0$
Beta(α,β)	$f(x) = \frac{1}{B(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}$	Messy	$\alpha,\beta > 0,\ x \in (0,1)$
Chi-Squared(p)	$f(x) = \frac{1}{2^{p/2}\Gamma(p/2)}x^{p/2-1}e^{-x/2}$	$(1-2t)^{-p/2},\ t < 1/2$	$p > 0,\ x > 0$
Laplace(μ,σ)	$f(x) = \frac{1}{2\sigma}e^{-\frac{|x-\mu|}{\sigma}}$	$\frac{e^{\mu t}}{1-(\sigma t)^2},\ |t| < \frac{1}{\sigma}$	$\sigma > 0$
Exponential(β)	$f(x) = \frac{1}{\beta}e^{-x/\beta}$	$\frac{1}{1-\beta t},\ t < 1/\beta$	$\beta > 0,\ x > 0$
Gamma(α,β)	$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha}x^{\alpha-1}e^{-x/\beta}$	$(1-\beta t)^{-\alpha},\ t < 1/\beta$	$\alpha,\beta > 0,\ x > 0$
Logistic(μ,β)	$f(x) = \frac{e^{-(x-\mu)/\beta}}{\beta(1+e^{-(x-\mu)/\beta})^2}$	$e^{\mu t}\Gamma(1-\beta t)\Gamma(1+\beta t)$	$\beta > 0$

Name	PDF / PMF	MGF $M_X(t)$	Restrictions
Normal $\mathcal{N}(\mu, \sigma^2)$	$f(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$\exp\left(\mu t + \dfrac{\sigma^2 t^2}{2}\right)$	$\sigma > 0$
Uniform(a, b)	$f(x) = \frac{1}{b-a}$	$\frac{e^{tb} - e^{ta}}{t(b-a)}, t \neq 0$	$a < b,\ x \in [a, b]$

D   Appendix

Place any appendices after the acknowledgments, starting on a new page. The appendices are numbered **A**, **B**, **C**, and so forth. The appendices contain material that you would like to share with the reader but that would hinder the flow of reading. For instance long proofs of theorems, code of algorithms, data, etc.

References

Askey, R. (1975). *Orthogonal Polynomials and Special Functions*. Society for Industrial and Applied Mathematics.

Atici, F. M., S. Chang, and J. Jonnalagadda (2021). Grünwald-letnikov fractional operators: from past to present. *Fractional Differential Calculus 11*(1), 147–159.

Boulaaras, S., R. Jan, and V.-T. Pham (2023). Recent advancement of fractional calculus and its applications in physical systems. *The European Physical Journal Special Topics 232*(14), 2347–2350.

Caputo, M. (1967, 11). Linear models of dissipation whose q is almost frequency independent—ii. *Geophysical Journal International 13*(5), 529–539.

Caputo, M. and M. Fabrizio (2015, 04). A new definition of fractional derivative without singular kernel. *Prog Fract Differ Appl 1*, 73–85.

Casella, G. and R. L. Berger (2002). *Statistical inference*, Volume 2. Duxbury Pacific Grove, CA.

Cauchy, A.-L. (1823). *Summary of the lessons given at the Royal Polytechnic School on infinitesimal calculus, Volume I*. GTextes1. Paris: Impr. royale, Debure frères. Contains a table of contents, searchable online.

Clark, T. G., M. J. Bradburn, S. B. Love, and D. G. Altman (2003). Survival analysis part i: Basic concepts and first analyses. *British Journal of Cancer 89*(2), 232–238.

Cressie, N., A. S. Davis, J. L. Folks, and G. E. Policello (1981). The moment-generating function and negative integer moments. *The American Statistician 35*(3), 148–150.

Di Matteo, A., M. Di Paola, and A. Pirrotta (2014). Probabilistic characterization of nonlinear systems under poisson white noise via complex fractional moments. *Nonlinear Dynamics 77*(3), 729–738.

D'Amico, M., G. Fusai, and A. Tagliani (2002, May). Valuation of exotic options using moments. *Operational Research 2*(2), 157–186.

Einstein, A. (1905). On the movement of small particles suspended in stationary liquids required by the molecular-kinetic theory of heat. *Annalen der Physik*, 549–560.

Flajolet, P. and R. Sedgewick (1995). Mellin transforms and asymptotics: Finite differences and rice's integrals. *Theoretical Computer Science 144*(1), 101–124.

Gourieroux, C. and J. Jasiak (2006, 03). Autoregressive gamma process. *Journal of Forecasting 25*, 129–152.

Gzyl, H. and S. Mayoral (2024). Determining the lifetime distribution using fractional moments with maximum entropy. *Heliyon 10*(15), e35250.

Gzyl, H., P.-L. Novi-Inverardi, and A. Tagliani (2013). Determination of the probability of ultimate ruin by maximum entropy applied to fractional moments. *Insurance: Mathematics and Economics 53*(2), 457–463.

Hadamard, J. (1892). Essai sur l'étude des fonctions données par leur développement de taylor. *Journal de Mathématiques Pures et Appliquées 8*, 101–186.

Haight, F. A. (1967). *Handbook of the Poisson distribution [by] Frank A. Haight.* Operations Research Society of America. Publications in operations research, no. 11. New York: Wiley.

Hansen, P. and C. Tong (2024, 10). Fractional moments by the moment-generating function.

Joseph Liouville, École polytechnique (Palaiseau, E. (1832). Journal de l'École polytechnique. Accessed: 2025-03-30.

Katugampola, U. N. (2014). A new approach to generalized fractional derivatives.

Khuri, A. and G. Casella (2002). The existence of the first negative moment revisited. *The American Statistician 56*(1), 44–47.

Kilbas, A. A., H. M. Srivastava, and J. J. Trujillo (2006). Theory and applications of fractional differential equations. In *Theory and Applications of Fractional Differential Equations*.

Kolmogorov, A. N. and S. V. Fomin (1999). *Elements of the Theory of Functions and Functional Analysis*. Dover Publications. Translated by Leo F. Boron.

Koning, D. (2015). *Fractional Calculus*. Ph. D. thesis, University of Groningen.

Laskin, N. (2002, Nov). Fractional schrödinger equation. *Phys. Rev. E 66*, 056108.

Lin, G. D. (1992). Characterizations of distributions via moments.

Losada, J. and J. Nieto (2015, 04). Properties of a new fractional derivative without singular kernel. *Prog Fract Differ Appl 1*, 87–92.

Love, E. R. (1971). Fractional derivatives of imaginary order. *Journal of the London Mathematical Society s2-3*(2), 241–259.

Mandelbrot, B. B. and J. W. Van Ness (1968). Fractional brownian motions, fractional noises and applications. *SIAM Review 10*(4), 422–437.

Mikosch, T., G. Samorodnitsky, and L. Tafakori (2013). Fractional moments of solutions to stochastic recurrence equations. *Journal of Applied Probability 50*(4), 969–982.

Missbauer, A. (2012). Gabor frames and the fractional fourier transform.

Novi Inverardi, P. L. and A. Tagliani (2024). Probability distributions approximation via fractional moments and maximum entropy: Theoretical and computational aspects. *Axioms 13*(1).

Riesz, M. (1949). L'intégrale de riemann-liouville et le problème de cauchy. *Acta Mathematica 81*(1), 1–222.

Scalas, E., R. Gorenflo, and F. Mainardi (2000). Fractional calculus and continuous-time finance. *Physica A: Statistical Mechanics and its Applications 284*(1), 376–384.

Sharpe, W. F. (1994). The sharpe ratio. *The Journal of Portfolio Management 21*(1), 49–58.

Tarasov, V. E. (2023). Fractional probability theory of arbitrary order. *Fractal and Fractional 7*(2).

Wang, Y., R. Tani, and K. Uchida (2025). A numerical solution method for the fractional moment problem within engineering. *Engineering Computations*.

Zhmakin, A. (2022, 12). A compact introduction to fractional calculus.