

Research Article

Gram-Charlier Processes and Applications to Option Pricing

Jean-Pierre Chateau¹ and Daniel Dufresne²

¹Faculty of Business Administration, University of Macau, Macau

²Montreal, QC, Canada

Correspondence should be addressed to Daniel Dufresne; dufresne@unimelb.edu.au

Received 17 August 2016; Accepted 1 November 2016; Published 8 February 2017

Academic Editor: Aera Thavaneswaran

Copyright © 2017 Jean-Pierre Chateau and Daniel Dufresne. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A Gram-Charlier distribution has a density that is a polynomial times a normal density. For option pricing this retains the tractability of the normal distribution while allowing nonzero skewness and excess kurtosis. Properties of the Gram-Charlier distributions are derived, leading to the definition of a process with independent Gram-Charlier increments, as well as formulas for option prices and their sensitivities. A procedure for simulating Gram-Charlier distributions and processes is given. Numerical illustrations show the effect of skewness and kurtosis on option prices.

1. Introduction

Gram-Charlier series are expansions of the form

$$f(x) = \phi(a, b; x) \cdot \left[1 + c_1 He_1\left(\frac{x-a}{b}\right) + c_2 He_2\left(\frac{x-a}{b}\right) + \dots \right], \quad (1)$$

where

$$\phi(a, b; x) = \frac{1}{b\sqrt{2\pi}} e^{-(1/2b^2)(x-a)^2}, \quad x \in \mathbf{R}, \quad (2)$$

is the usual normal density and He_k is the Hermite polynomial of order k . The expression within square brackets in (1) is an orthogonal polynomial expansion for the ratio $f(x)/\phi(x)$; given an arbitrary function the expansion may or may not converge to the true value of $f(x)$. In this paper we focus on the properties of the Gram-Charlier distributions, obtained by truncating the series after a finite number of terms. What is obtained is a family of distributions parametrized by a, b, c_0, \dots, c_N , as is explained in detail below.

This paper has three main goals: (1) define and study the properties of the family of Gram-Charlier distributions; (2) define a Gram-Charlier process and derive its basic properties; (3) apply those to European options. The formulas we

give for European option prices and Greeks apply to Gram-Charlier distributions of any order, and we use four- and six-parameter Gram-Charlier distributions in our examples. Two numerical illustrations show how option prices are affected by the skewness and kurtosis of returns. This paper can be a reference for those using Gram-Charlier distributions in option pricing but also in statistics.

Most previous applications to option pricing have assumed that $c_1 = c_2 = 0$. We believe this restriction is not necessary in a general theory of Gram-Charlier distributions, as there may well be situations where the extra degrees of freedom given by c_1 and c_2 will be useful. We discuss it in detail at the end of Section 2.

It has been observed that option prices have nonconstant implied volatilities, meaning that log returns do not have a normal distribution under the risk-neutral measure. There is a wide literature on modelling log returns to fit observed option prices, the main alternatives to Brownian motion being stochastic volatility models (where the parameter σ in Black-Scholes is replaced with a continuous-time stochastic process), GARCH time series, and Lévy processes. Gram-Charlier distributions are mathematically simpler than the models just mentioned, while allowing a better fit to data than the normal distribution. Several authors have used Gram-Charlier distributions in option pricing, as a model for

log returns, among others [1–10]. The majority of previous authors assumed a density of the form

$$\phi(x) \left(1 + \frac{\bar{s}}{6} He_3(x) + \frac{\bar{k}}{24} He_4(x) \right), \quad (3)$$

$$\phi(x) = \phi(0, 1, x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2},$$

for the normalized log return. (In our notation this is a $\mathbf{GC}(0, 1; 0, 0, \bar{s}/6, \bar{k}/24)$ distribution; see Section 2.) The notation emphasizes that in this case the coefficient of $He_3(x)$ turns out to be Pearson's skewness coefficient \bar{s} divided by 6, and the coefficient of $He_4(x)$ the excess kurtosis coefficient \bar{k} divided by 24. The distribution of log returns is then a four-parameter Gram-Charlier distribution (since there are two other parameters, the mean and variance, besides \bar{s} and \bar{k}). This distribution allows nonzero skewness and excess kurtosis, unlike the normal distribution found in Black-Scholes. In (3) the parameters (\bar{s}, \bar{k}) are restricted to a specific region (see Figure 3), because outside that region the function in (3) becomes negative for some values of x (see Section 2.1).

Jurczenko et al. [8] specify the martingale restriction

$$S_0 = e^{-rT} \mathbf{E}^Q(S_T) \quad (4)$$

that the four-parameter Gram-Charlier density must satisfy in pricing options (previous authors had not taken it into account). The martingale condition for general Gram-Charlier distributions is described in Section 5. Our Gram-Charlier distribution with parameters a, b, c_1, \dots, c_N , denoted by $\mathbf{GC}(a, b; c_1, \dots, c_N)$, has density

$$\phi(a, b, x) \cdot \left[1 + c_1 He_1\left(\frac{x-a}{b}\right) + \dots + c_N He_N\left(\frac{x-a}{b}\right) \right]. \quad (5)$$

As stated above we do not set $c_1 = c_2 = 0$, and N can be any even positive integer. The question whether (5) is nonnegative for all x is an important one. Several authors have disregarded this issue and, in fact, some have come up with parameters that do not yield a true probability density function. One might argue that this is the price to pay for truncating an infinite Gram-Charlier series. However, if the same log return distribution is used to price many options, then a true probability density function is the only safe choice, because otherwise there might be inconsistencies among option prices. For instance, if the function used as density is negative over the interval (α, β) , then a digital option that pays off only when the log return is in that interval will have a negative price. In this paper we consider Gram-Charlier *distributions*, not *expansions*, and insist that the densities integrate to one and be nonnegative. Our goal is to define a family of proper probability distributions; nevertheless, our formulas do apply to truncated Gram-Charlier expansions as well.

The paper by León et al. [11] presents an alternative to the general Gram-Charlier distributions we study in this paper. Those authors consider the subclass of Gram-Charlier distributions consisting of densities

$$f_X(x) = \phi(a, b, x) p(x), \quad (6)$$

where the polynomial $p(x)$ is the square of another polynomial, $p(x) = q(x)^2$. This has the obvious advantage that the nonnegativity restriction on $f_X(\cdot)$ is automatically satisfied. We discuss that subclass of “squared” Gram-Charlier distributions in the Conclusion.

Almost all previous authors have used Gram-Charlier distributed log returns over a single time period. This has an obvious downside, in that it becomes tricky, if not impossible, to preserve consistency between the prices of options with different maturities. Section 3 shows that a Lévy process with Gram-Charlier increments does not exist; however, it also shows that the sum of independent Gram-Charlier distributed variables also has a Gram-Charlier distribution. This opens the way for multiperiod Gram-Charlier option pricing, using a discrete-time random walk model for which the log return over any period has a Gram-Charlier distribution. The Gram-Charlier distribution of the multiperiod return has a larger number of parameters, though the model is still simpler than almost any (if not all) alternative stochastic volatility models. There is no problem computationally, since we give explicit formulas for options under Gram-Charlier distributions with an arbitrary number of parameters.

The layout of the paper is as follows. In Section 2, we extend the study of Gram-Charlier distributions to all possible polynomials $p(\cdot)$ and derive their properties (moments, cumulants, moment determinacy, properties of the set of valid parameters, tail, and so on). Some of the formulas and properties in Theorem 2 appear to be new. In Section 3 we show that there is no Lévy process with Gram-Charlier distributed increments, apart from Brownian motion, and we define a discrete-time process with independent Gram-Charlier increments that is suitable for option pricing. In Section 4 we show that the log Gram-Charlier distribution is not determined by its moments, just like the lognormal. Next, Section 5 gives formulas for European call and put prices when the log-price returns of the underlying have a general Gram-Charlier distribution; the previous literature mostly considered the $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$ family and the squared Gram-Charlier distributions (an exception is [10], where it is assumed that $c_1 = c_2 = 0$). In particular, we derive a change of measure formula that extends the Cameron-Martin formula for the normal distribution; the latter is used in pricing European options in the Black-Scholes model. We also derive formulas for the sensitivities (Greeks) of those option prices with respect to all parameters. A technique for simulating Gram-Charlier distributions is described. Parts (c), (d), (i), (j) (k), and (l) of Theorem 2 and part (b) of Theorem 3 and Theorems 6, 7, 9, and 11 appear to be new. Theorem 3(a) is for the first time formulated for general Gram-Charlier distributions. Theorems 3(a), 4, and 8 are given, more or less explicitly, for the subclass of squared

Gram-Charlier distributions in [11], and some of the Greeks in Theorem 9 had been calculated for the $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$ distribution by previous authors.

In Section 6 we give two applications that show how option prices depend on skewness and kurtosis of the log returns (this of course cannot be done in the classical Black-Scholes setting, while it appears quite complicated to do so in stochastic volatility or Lévy driven models). The first example is equity indexed annuities (EIAs in the sequel) premium options. The pricing of EIAs has been studied by many authors, including Hardy [12], Gaillardetz and Lin [13], and Boyle and Tian [14]. The second example is lookback options, which illustrates the use of the simulation method in Section 2.5. Lookback options have been studied by numerous authors, but there is no closed form formula for the price of discretely monitored lookback options within the Black-Scholes model. We refer the reader to Kou [15]. In those examples all parameters are estimated by maximum likelihood, with the range of the parameters restricted to where they correspond to a true probability distribution. Moment estimation is of course possible when $c_1 = c_2 = 0$, but it is not trustworthy, because of the range restriction on the parameters c_3, c_4, \dots ; for instance, the empirical skewness and excess kurtosis have a positive probability of falling outside the feasible region (see Section 6). Theorem 2(f) shows that the first k moments of the Gram-Charlier distribution depend on more than k parameters, which probably rules out moment estimation, unless one a priori fixes two of the parameters (which again is not trustworthy). We apply maximum likelihood to the six-parameter $\mathbf{GC}(a, b; c_1, c_2, c_3, c_4)$ in Section 6, but we do realize that maximum likelihood estimation for higher order Gram-Charlier distributions poses computational problems, which are an interesting avenue for further research. (A referee pointed out that fixing a and b would make estimation easier; our guess is that if one wishes to simplify maximum likelihood estimation then fixing a and b might be a better idea than fixing c_1 and c_2 , though we have not looked at this in any depth. A two-step process might be imagined, whereby the data first give information about a and b , and then maximum likelihood is applied using that preliminary information. This would help by constraining the optimization to a smaller region, while agreeing with the intuitive idea that a and b are location and scale parameters, resp.)

Notation 1. The normal density function is denoted as

$$\begin{aligned} \phi(a, b, x) &= \frac{1}{b\sqrt{2\pi}} e^{-(1/2b^2)(x-a)^2}, \\ \phi(0, 1, x) &= \phi(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}, \end{aligned} \tag{7}$$

and its distribution function is

$$\Phi(x) = \int_{-\infty}^x \phi(y) dy. \tag{8}$$

Two equivalent versions of the Hermite polynomials may be found in the literature: for $n = 0, 1, 2, \dots$,

$$\begin{aligned} H_n(x) &= (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2} \\ He_n(x) &= (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}. \end{aligned} \tag{9}$$

The first one is common in mathematics and physics, but in probability and statistics there is an obvious advantage in using the second one. (The conversion formula is $He_n(x) = 2^{-n/2} H_n(x/\sqrt{2})$.) The first few Hermite polynomials are

$$\begin{aligned} He_0(x) &= 1, \\ He_1(x) &= x, \\ He_2(x) &= x^2 - 1, \\ He_3(x) &= x^3 - 3x, \\ He_4(x) &= x^4 - 6x^2 + 3, \\ He_5(x) &= x^5 - 10x^3 + 15x, \\ He_6(x) &= x^6 - 15x^4 + 45x^2 - 15. \end{aligned} \tag{10}$$

2. Gram-Charlier Distributions

For a fixed N , consider the class of distributions that have a pdf of the form

$$f(x) = \phi(x) \sum_{k=0}^N c_k He_k(x), \quad x \in \mathbf{R}, \tag{11}$$

with $c_N \neq 0$. Noting that the leading term of $He_k(x)$ is x^k , we conclude that N must necessarily be even, because if N were odd then the polynomial that multiplies $\phi(x)$ would take negative values for some x . For the same reason c_N cannot be negative.

Definition 1. Let $a \in \mathbf{R}$, $b > 0$, $c_k \in \mathbf{R}$, $c_0 = 1$, and $N \in \{0, 2, 4, \dots\}$. We write $Y \sim \mathbf{GC}(a, b; c_1, \dots, c_N)$ (or $Y \sim \mathbf{GC}(a, b; \vec{c})$) if the variable $(Y - a)/b$ has probability density function

$$\phi(x) \sum_{k=0}^N c_k He_k(x). \tag{12}$$

This will be called a *Gram-Charlier distribution* with parameters a, b, \vec{c} , with $\vec{c} = (c_1, \dots, c_N)$. The largest N such that $c_N > 0$ is called the *order* of the Gram-Charlier distribution. The normal distribution with mean a and variance b^2 is a $\mathbf{GC}(a, b; (0, \dots, 0))$ (or $\mathbf{GC}(a, b; -)$) with order 0.

The class of Gram-Charlier distributions just defined includes all distributions with density

$$\frac{1}{b} \phi\left(\frac{y-a}{b}\right) p(y), \tag{13}$$

where $p(y)$ is a polynomial of degree N , since $p(y)$ can be rewritten as a combination of

$$He_k\left(\frac{y-a}{b}\right), \quad k = 0, 1, \dots, N. \quad (14)$$

The condition $c_0 = 1$ ensures that function (12) integrates to one, since

$$\int_{-\infty}^{\infty} \phi(x) He_k(x) dx = (-1)^k \int_{-\infty}^{\infty} \frac{d^k}{dx^k} \phi(x) dx = 0, \quad (15)$$

$$k = 1, 2, \dots$$

There are no simple conditions that ensure that a polynomial remains nonnegative everywhere, though in some cases precise conditions on \tilde{c} are known; see below. If a vector \tilde{c} leads to a true Gram-Charlier pdf, then we will say that \tilde{c} is *valid*.

Generating functions are convenient when dealing with orthogonal polynomials. One is

$$w_1(t, x) = \sum_{n=0}^{\infty} \frac{t^n}{n!} He_n(x) = e^{tx - (t^2/2)}. \quad (16)$$

Another one is

$$w_2(t, u, x, y) = \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \frac{t^k}{k!} \frac{u^n}{n!} \int_{-\infty}^{\infty} e^{-x^2/2} He_k(x) He_n(x+y) dx \quad (17)$$

$$= \sqrt{2\pi} e^{u(t+y)}.$$

Letting $y = 0$ leads to

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} He_k(x) He_n(x) dx = \left. \frac{\partial^k}{t^k} \frac{\partial^n}{u^n} e^{tu} \right|_{t=u=0} \quad (18)$$

$$= \begin{cases} 0 & \text{if } k \neq n \\ n! & \text{if } k = n. \end{cases}$$

This proves the orthogonality of the Hermite polynomials and gives us the value of

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} He_n^2(x) dx, \quad (19)$$

which is essential in deriving Gram-Charlier series. Since $He_0 = 1$ this also implies

$$\int_{-\infty}^{\infty} \phi(x) He_k(x) dx = 0, \quad k = 1, 2, \dots \quad (20)$$

Another formula is the Laplace transform of $\phi(x)He_n(x)$, which may be found by integrating by parts n times:

$$\int_{-\infty}^{\infty} e^{tx} \phi(x) He_k(x) dx = t^k e^{t^2/2}, \quad k = 0, 1, \dots \quad (21)$$

Let us calculate the moments of a distribution with density (11). First consider

$$\int_{-\infty}^{\infty} x^n \phi(x) He_k(x) dx = (-1)^k \int_{-\infty}^{\infty} x^n \left(\frac{d^k}{dx^k} \phi(x) \right) dx. \quad (22)$$

Integrating by parts repeatedly yields

$$\int_{-\infty}^{\infty} x^n \phi(x) He_k(x) dx = \begin{cases} \frac{n!}{2^{(n-k)/2} ((n-k)/2)!} & \text{if } n-k \text{ is even and nonnegative} \\ 0 & \text{if } n-k \text{ is odd and nonnegative.} \end{cases} \quad (23)$$

Hence, the n th moment of the distribution in (11) is

$$\sum_{k=0}^{N \wedge n} c_k \frac{n!}{2^{(n-k)/2} ((n-k)/2)!} \mathbf{1}_{\{n-k \text{ even}\}}. \quad (24)$$

This says in particular that the parameter c_k only affects the moments of order k and higher of the $\mathbf{GC}(a, b; \tilde{c})$ distribution (see part (i) of Theorem 2). This is confirmed by the moment generating function, which may be found from (21):

$$\int_{-\infty}^{\infty} e^{tx} f(x) dx = e^{t^2/2} \sum_{k=0}^N c_k t^k. \quad (25)$$

It can be checked that differentiating this expression n times and setting $t = 0$ give the same expression found for the n th moment of (11).

Theorem 2. *Suppose $Y \sim \mathbf{GC}(a, b; \tilde{c})$, $\tilde{c} \in \mathbf{R}^N$ with $b > 0$, $c_0 = 1$, and $c_N > 0$. The order N of the distribution is necessarily even.*

(a)

$$\mathbf{E}(Y - a)^n = b^n \sum_{k=0}^{N \wedge n} c_k \frac{n!}{2^{(n-k)/2} ((n-k)/2)!} \mathbf{1}_{\{n-k \text{ even}\}}. \quad (26)$$

(b)

$$\mathbf{E}e^{tY} = e^{at + (b^2 t^2/2)} \sum_{k=0}^N b^k c_k t^k, \quad t \in \mathbf{R}. \quad (27)$$

(c) *The representation of the $\mathbf{GC}(a, b; \tilde{c})$ distribution in terms of the parameters a, b , and \tilde{c} is unique.*

(d) *All Gram-Charlier distributions are determined by their moments.*

(e) *The set of valid \tilde{c} in \mathbf{R}^N includes the origin, is not reduced to a single point, and is convex.*

(f) The first six moments of the $\mathbf{GC}(a, b; \bar{c})$ distribution are

$$\begin{aligned}
 m_1 &= a + bc_1, \\
 m_2 &= a^2 + b^2 + 2(abc_1 + b^2c_2), \\
 m_3 &= a^3 + 3ab^2 + 3bc_1(a^2 + b^2) + 6(ab^2c_2 + b^3c_3), \\
 m_4 &= a^4 + 4a^3bc_1 + 12a^2b^2c_2 + 6a^2b^2 + 12ab^3c_1 \\
 &\quad + 24ab^3c_3 + 12b^4c_2 + 24b^4c_4 + 3b^4, \\
 m_5 &= a^5 + 5a^4bc_1 + 20a^3b^2c_2 + 10a^3b^2 + 30a^2b^3c_1 \\
 &\quad + 60a^2b^3c_3 + 60ab^4c_2 + 120ab^4c_4 + 15ab^4 \\
 &\quad + 15b^5c_1 + 60b^5c_3 + 120b^5c_5, \\
 m_6 &= a^6 + 6a^5bc_1 + 30a^4b^2c_2 + 15a^4b^2 + 60a^3b^3c_1 \\
 &\quad + 120a^3b^3c_3 + 180a^2b^4c_2 + 360a^2b^4c_4 \\
 &\quad + 45a^2b^4 + 90ab^5c_1 + 360ab^5c_3 + 720ab^5c_5 \\
 &\quad + 90b^6c_2 + 360b^6c_4 + 720b^6c_6 + 15b^6.
 \end{aligned} \tag{28}$$

(g) The first six cumulants of the $\mathbf{GC}(a, b; \bar{c})$ distribution are

$$\begin{aligned}
 \kappa_1 &= a + bc_1, \\
 \kappa_2 &= b^2(1 - c_1^2 + 2c_2), \\
 \kappa_3 &= 2b^3(c_1^3 - 3c_1c_2 + 3c_3), \\
 \kappa_4 &= -6b^4(c_1^4 - 4c_1^2c_2 + 2c_2^2 + 4c_1c_3 - 4c_4), \\
 \kappa_5 &= 24b^5(c_1^5 - 5c_1^3c_2 + 5c_1c_2^2 + 5c_1^2c_3 - 5c_2c_3 - 5c_1c_4 \\
 &\quad + 5c_5), \\
 \kappa_6 &= -120b^6(c_1^6 - 6c_1^4c_2 + 9c_1^2c_2^2 - 2c_2^3 + 6c_1^3c_3 \\
 &\quad - 12c_1c_2c_3 + 3c_3^2 - 6c_1^2c_4 + 6c_2c_4 + 6c_1c_5 - 6c_6).
 \end{aligned} \tag{29}$$

(h) The following hold for any $\mathbf{GC}(a, b; \bar{c})$ distribution:

$$\begin{aligned}
 \text{mean:} & \quad a + bc_1 \\
 \text{variance:} & \quad b^2(1 - c_1^2 + 2c_2) \\
 \text{skewness coefficient:} & \quad \frac{2(c_1^3 - 3c_1c_2 + 3c_3)}{(1 - c_1^2 + 2c_2)^{3/2}} \\
 \text{excess kurtosis coefficient:} & \quad -\frac{6(c_1^4 - 4c_1^2c_2 + 2c_2^2 + 4c_1c_3 - 4c_4)}{(1 - c_1^2 + 2c_2)^2}.
 \end{aligned} \tag{30}$$

(i) Suppose $X \sim \mathbf{GC}(a, b; \bar{c}^X)$, $Y \sim \mathbf{GC}(a, b; \bar{c}^Y)$. Then the first K moments of X and Y are the same; that is,

$$\mathbf{E}X^j = \mathbf{E}Y^j, \quad j = 1, \dots, K, \tag{31}$$

if and only if $c_j^X = c_j^Y$, $j = 1, \dots, K$. This implies that if c_1 up to c_4 are equal to 0 then the distribution has zero skewness and excess kurtosis; hence, this shows how to construct an infinite number of distributions that share this property with the normal.

(j) Suppose $X \sim \mathbf{GC}(a, b; \bar{c}^X)$. Then

$$\begin{aligned}
 a = \mathbf{E}X &\iff c_1^X = 0, \\
 b^2 = \mathbf{E}(X - a)^2 &\iff c_2^X = 0,
 \end{aligned} \tag{32}$$

$$\{a = \mathbf{E}X, b^2 = \mathbf{Var} X\} \iff \{c_1^X = c_2^X = 0\}.$$

When $c_1^X = c_2^X = 0$ the skewness and excess kurtosis coefficients of X are $6c_3^X$ and $24c_4^X$, respectively, for any $N = 0, 2, 4, \dots$

(k) If $X \sim \mathbf{GC}(a, b; c_1, \dots, c_N)$ and q is a constant then $Y = qX \sim \mathbf{GC}(aq, b|q; c'_1, \dots, c'_N)$, where $c'_k = (\text{sign}(q))^k c_k$, $k \geq 1$. In particular, $-X \sim \mathbf{GC}(-a, b; -c_1, \dots, (-1)^{N-1}c_{N-1}, c_N)$.

(l) The law of the square of $X \sim \mathbf{GC}(0, 1; c_1, \dots, c_N)$ is a combination of chi-square distributions with $1, 3, \dots, N + 1$ degrees of freedom that has density

$$g(y) = \frac{e^{-y/2}}{\sqrt{2\pi y}} \sum_{j=0}^{n/2} \alpha_j y^j, \tag{33}$$

$$\text{where } \sum_{j=0}^{n/2} \alpha_j y^j = \sum_{j=0}^{n/2} c_{2j} \text{He}_{2j}(\sqrt{y}).$$

Proof. Parts (a) and (b) were proved above, and (c) follows directly from (b). To prove (d) it is sufficient to note the existence of $\mathbf{E}e^{tY}$ for t in an open neighbourhood of $s = 0$. For (e), if $\bar{c} = \bar{0} \in \mathbf{R}^N$ then the distribution is the standard normal, and this is a Gram-Charlier distribution. Since $c_N > 0$ and N is even, the polynomial

$$\sum_{k=1}^N \text{He}_k(x) \tag{34}$$

tends to ∞ when $|x|$ tends to ∞ . Hence, there is $\epsilon_0 > 0$ such that

$$1 + \epsilon \sum_{k=0}^N He_k(x) > 0, \quad x \in \mathbf{R}, \quad 0 \leq \epsilon \leq \epsilon_0 \quad (35)$$

(recall that $c_0 He_0(x) = 1$ for all Gram-Charlier distributions). The set of valid vectors \tilde{c} thus includes $(\epsilon, \dots, \epsilon) \in \mathbf{R}^N$ for each $0 \leq \epsilon \leq \epsilon_0$. If $\tilde{c}^{(j)}$, $j = 1, 2$, are valid then $p\tilde{c}^{(1)} + (1-p)\tilde{c}^{(2)}$ is also valid, for any $p \in (0, 1)$.

Part (f) is found by expanding the mgf in (b) as a series in s around the origin. Part (g) is found by expanding

$$t \mapsto \log \left(e^{at + (b^2 t^2/2)} \sum_{k=0}^N b^k c_k t^k \right). \quad (36)$$

The formulas in (h) follow (f) and (g).

For part (i), it is sufficient to consider the case $a = 0, b = 1$ only. Write

$$\begin{aligned} \psi(t) &= e^{t^2/2}, \\ q_X(t) &= \sum_{k=0}^N c_k^X t^k, \\ q_Y(t) &= \sum_{k=0}^N c_k^Y t^k \end{aligned} \quad (37)$$

and calculate the moments of X and Y by successively differentiating the mgf's of X and Y (note that $c_0^X = q_X(0) = c_0^Y = q_Y(0) = 1$). The first K moments of X and Y are the same if and only if

$$\begin{aligned} \psi'(0) q_X(0) + \psi(0) q_X'(0) &= \psi'(0) q_Y(0) + \psi(0) q_Y'(0), \\ &\vdots \\ \sum_{j=0}^K \binom{K}{j} \psi^{(j)}(0) q_X^{(K-j)}(0) &= \sum_{j=0}^K \binom{K}{j} \psi^{(j)}(0) q_Y^{(K-j)}(0). \end{aligned} \quad (38)$$

Suppose that $c_j^X = c_j^Y$, $j = 1, \dots, K$. Then $q_X^{(j)}(0) = q_Y^{(j)}(0)$, $j = 1, \dots, K$, and thus the first K moments of X and Y are the same. Conversely, suppose $\mathbf{E}X^j = \mathbf{E}Y^j$, $j = 1, \dots, K$. Then the first identity above implies that $c_1^X = q_X'(0) = q_Y'(0) = c_1^Y$, since $\psi(0)$ is not zero. The second identity implies that $c_2^X = c_2^Y$, and so on, up to $c_K^X = c_K^Y$.

Turning to (j), the first equivalence

$$a = \mathbf{E}X \iff c_1^X = 0 \quad (39)$$

is an immediate consequence of property (i) with $K = 1$. To prove the second equivalence, suppose $X \sim \mathbf{GC}(0, 1; \tilde{c})$, and let $Y \sim \mathbf{GC}(0, 1; -) = \mathbf{N}(0, 1)$. Then

$$\begin{aligned} \mathbf{E}X^2 &= \psi''(0) q_X(0) + 2\psi'(0) q_X'(0) + \psi(0) q_X''(0), \\ \mathbf{E}Y^2 &= \psi''(0) q_Y(0) + 2\psi'(0) q_Y'(0) + \psi(0) q_Y''(0). \end{aligned} \quad (40)$$

Here $q_X(0) = q_Y(0) = 1$ and $\psi'(0) = 0$; hence, $\mathbf{E}X^2 = \mathbf{E}Y^2$ if and only if $q_X''(0) = q_Y''(0)$. The last equality is $c_2^X = c_2^Y$. For $X \sim \mathbf{GC}(a, b; \tilde{c}^X)$, with a and b being arbitrary, this means that

$$\mathbf{E} \left(\frac{x-a}{b} \right)^2 = 1 \iff c_2^X = 0. \quad (41)$$

For (k), observe that if $q \neq 0$ then the mgf of Y is

$$\begin{aligned} e^{aq_s + (1/2)b^2 q^2 s^2} \sum_{k=0}^N c_k b^k q^k s^k \\ = e^{aq_s + (1/2)b^2 q^2 s^2} \sum_{k=0}^N c_k \left(\frac{q}{|q|} \right)^k (b|q|)^k s^k. \end{aligned} \quad (42)$$

Finally, turn to (l). Routine calculations show that the density of the square is

$$\begin{aligned} \frac{1}{2\sqrt{y}} (f_X(\sqrt{y}) + f_X(-\sqrt{y})) \\ = \frac{\phi(\sqrt{y})}{2\sqrt{y}} \sum_{k=0}^N c_k [He_k(\sqrt{y}) + He(-\sqrt{y})]. \end{aligned} \quad (43)$$

The Hermite polynomials of odd order are odd functions and so disappear from that expression, while the even order Hermite polynomials are even functions. \square

When $N = 0$ the $\mathbf{GC}(a, b; \tilde{c})$ distribution is the normal distribution with mean a and variance b^2 . However, part (b) of the theorem says that when $N > 0$ the parameters a and b^2 are not necessarily the mean and variance of the distribution. Simple calculations lead to the following result.

Theorem 3. (a) If $X \sim \mathbf{GC}(0, 1; c_1, \dots, c_N)$ then

$$\mathbf{P}(X \leq x) = \Phi(x) - \phi(x) \sum_{k=1}^N c_k He_{k-1}(x), \quad (44)$$

$$\mathbf{P}(X > x) = \Phi(-x) + \phi(x) \sum_{k=1}^N c_k He_{k-1}(x).$$

(b) The tails of the $\mathbf{GC}(a, b; c_1, \dots, c_N)$ distribution are

$$\mathbf{P}(X > x) \sim c_N \left(\frac{x-a}{b} \right)^{N-1} \phi \left(\frac{x-a}{b} \right) \quad \text{as } x \rightarrow \infty, \quad (45)$$

$$\mathbf{P}(X < x) \sim c_N \left(\frac{|x-a|}{b} \right)^{N-1} \phi \left(\frac{x-a}{b} \right) \quad \text{as } x \rightarrow -\infty.$$

The tails of the Gram-Charlier distributions are thicker than those of the normal distribution but are still "thin" because they are in the limit smaller than any exponential function $\exp(-\alpha x)$.

2.1. *The $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$ Family.* Here the exact region for the (c_3, c_4) that lead to a true probability distribution has been found. This goes back to Barton and Dennis [16], but a more detailed explanation is given in Jondeau and Rockinger [6]. The region is shown in Figure 3 (use the correspondence $\bar{s} = 6c_3, \bar{k} = 24c_4$ from Theorem 2(h)). An important fact about this region is that it is not rectangular; the possible excess kurtosis values depend on skewness, and conversely.

2.2. *$\mathbf{GC}(a, b; \bar{c})$ Distributions of Order 4 and Higher.* The $\mathbf{GC}(a, b; c_1, c_2, c_3, c_4)$ family has six parameters, rather than four, and thus has more degrees of freedom in fitting data; to the authors' knowledge the general six-parameter Gram-Charlier distribution has been used in financial applications by León et al. [11] only (those authors use the subfamily consisting of polynomials $p(\cdot)$ that are squares of some second-degree polynomial). Schlögl [10] fits the six- and eight-parameter families $\mathbf{GC}(a, b; 0, 0, c_3, c_4, c_5, c_6)$ and $\mathbf{GC}(a, b; 0, 0, c_3, c_4, c_5, c_6, c_7, c_8)$ to data.

The set of (c_1, c_2, c_3, c_4) that yield true probability distributions has not been identified, but it is possible to fit the six parameters a, b, c_1, c_2, c_3, c_4 by maximum likelihood.

2.3. *Why We Do Not Assume That $c_1 = c_2 = 0$.* Almost all previous authors have assumed that $c_1 = c_2 = 0$, because they used normalized data:

$$Y = \frac{X - \mathbf{E}X}{\sqrt{\mathbf{Var} X}} \quad (46)$$

(see part (j) of Theorem 2) as explained below. In this section we explain why it is important not to restrict Gram-Charlier distributions or series in that way. The first reason for not doing so is that enlarging the parameter space can only be a good thing. The second one is that in fitting those distributions to data there may be very real advantages in letting c_1 and c_2 be different from zero. The only downside of letting c_1 and c_2 take nonzero values, and it is of no real importance, is that c_3 and c_4 lose their simple relationship with skewness and excess kurtosis (see part (h) of Theorem 2). A third reason is that after an exponential change of measure a Gram-Charlier distribution will rarely have $c_1 = c_2 = 0$; see Section 2.4.

We now show, using an example that can be worked out explicitly, that it is not always best to use normalized data when fitting Gram-Charlier distributions, because choosing another affine transformation of the data may well yield a much better fit.

The "standard" Gram-Charlier expansion for a function $f(\cdot)$ is

$$\phi(x) \sum_{k=0}^{\infty} c_k He_k, \quad (47)$$

where

$$c_k = \frac{1}{k!} \int_{-\infty}^{\infty} He_k(x) f(x) dx = \frac{1}{k!} \mathbf{E}He_k(X). \quad (48)$$

A classical result about Hermite series, proved by Cramér [17], is that sufficient conditions for the Gram-Charlier expansion (47) to converge to

$$\frac{1}{2} (f(x-) + f(x+)) \quad (49)$$

for all $x \in \mathbf{R}$ are that (i) $f(\cdot)$ has finite variation in every bounded interval, and (ii) satisfies

$$\int_{-\infty}^{\infty} e^{x^2/4} f(x) dx < \infty. \quad (50)$$

If X has density $f(\cdot)$, then the last condition is $\mathbf{E}e^{X^2/4} < \infty$. This condition cannot be improved upon, in the sense that there are cases where the Gram-Charlier series defined above diverges, although $\mathbf{E}e^{\rho X^2} < \infty$ for all $\rho < 1/4$ (it will be shown below that one such case is the normal distribution with mean 0 and variance 2).

Let us first calculate

$$h_{q,k} = \mathbf{E}He_k(X), \quad X \sim \mathbf{N}(0, q^2). \quad (51)$$

Using the generating function (16) we find

$$h_{q,k} = \begin{cases} 0 & \text{if } k \text{ is odd} \\ \frac{k!}{2^{k/2} (k/2)!} (q^2 - 1)^{k/2} & \text{if } k \text{ is even.} \end{cases} \quad (52)$$

The Gram-Charlier expansion corresponding to the $\mathbf{N}(0, q^2)$ density is thus

$$\begin{aligned} \phi(x) \sum_{k=0}^{\infty} \frac{h_{q,k}}{k!} He_k(x) \\ = \phi(x) \sum_{n=0}^{\infty} \frac{1}{2^n n!} (q^2 - 1)^n He_{2n}(x). \end{aligned} \quad (53)$$

It is possible to determine whether this converges or not when $x = 0$; the Gram-Charlier series for $f(0)$ is

$$\frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} (-1)^n \frac{(2n)!}{2^{2n} (n!)^2} (q^2 - 1)^n. \quad (54)$$

From Stirling's formula,

$$\frac{(2n)!}{2^{2n} (n!)^2} \sim \frac{1}{\sqrt{\pi n}} \quad (55)$$

as n tends to infinity, which implies that there are three possibilities: expansion (54) (i) converges absolutely if $0 < q^2 < 2$, (ii) converges simply if $q^2 = 2$, and diverges if $q^2 > 2$. (The preceding calculations are from Cramér [17].)

Let us now consider a random variable $X_p, 0 \leq p \leq 1$, with density

$$g_p(x) = p\phi(x) + (1-p) \frac{1}{p} \phi\left(\frac{x}{p}\right). \quad (56)$$

In words, X_p has a $N(0, 1)$ distribution with probability p and a $N(0, p^2)$ distribution with probability $1 - p$. Define the normalized variable

$$Y_p = \frac{X_p}{\sqrt{\text{Var } X_p}} = \frac{X_p}{\sigma_p}, \tag{57}$$

$$\sigma_p^2 = \text{Var } X_p = p + p^2 - p^3.$$

The coefficients $c_{p,k}$ of the Gram-Charlier expansion for Y_p are

$$c_{p,k} = \frac{1}{k!} \mathbf{E} H e_k(Y_p) \tag{58}$$

$$= \frac{1}{k!} (p h_{1/\sigma_p, k} + (1-p) h_{p/\sigma_p, k}).$$

We have shown that (53) diverges if $q^2 > 2$. Here we see that $c_{p,k}$ involves $h_{q,k}$ with $q = 1/\sigma_p$ and $q = p/\sigma_p$; in the second case there is no problem, as $(p/\sigma_p)^2 < 2$ for all $p \in (0, 1]$. In the first case,

$$\frac{1}{\sigma_p^2} > 2 \iff p + p^2 - p^3 < \frac{1}{2}. \tag{59}$$

This condition is satisfied for p smaller than $p^* \approx 0.403$, and so the Gram-Charlier expansion for Y_p diverges for all $p < p^*$. This is a sad state of affair: a series designed to work for distributions that are “close to the normal” that fails for combinations of two normal densities!

There is an easy solution: use a different scaling for X_p . Rather than multiplying X_p by $1/\sigma_p$, use a factor r such that the series converges. In this example we know that the series converges if and only if $r < \sqrt{2}$. Say we choose $p = .3$; this implies

$$\frac{1}{\text{Var } X_p} = \frac{1}{0.3 + 0.3^2 - 0.3^3} = 2.75482 > 2. \tag{60}$$

Figure 1 shows how the true density for $X_{0.3}$ is approximated by order 10 and 20 Gram-Charlier truncated series if $r = 1/\sigma_{0.3} \doteq 1.66$ is chosen; the graph is not surprising, the infinite Gram-Charlier expansion diverges. Figure 2 shows the same except that r is set to 1.4. The latter is smaller than $\sqrt{2}$, so the infinite expansion converges. In this example the density to be approximated is symmetrical about 0, so c_1 is always 0, while

$$q = \frac{1}{\sigma_{0.3}} \implies c_2 = 0; \tag{61}$$

$$q = 1.4 \implies c_2 = -0.14426.$$

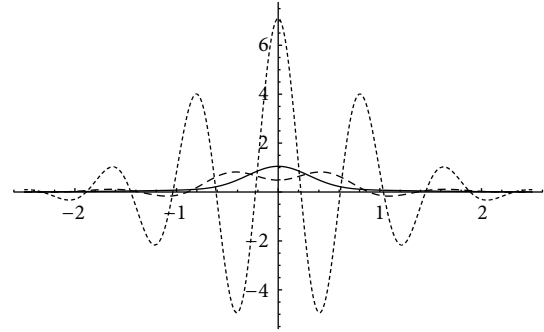


FIGURE 1: Gram-Charlier approximations to the density $g_{0.3}(x)$ (solid line) when $q = 1/\sigma_{0.3} \doteq 1.66$, resulting in $c_1 = c_2 = 0$. The curve with long dashes is the order 10 approximation, and the one with short dashes is the order 20 approximation.

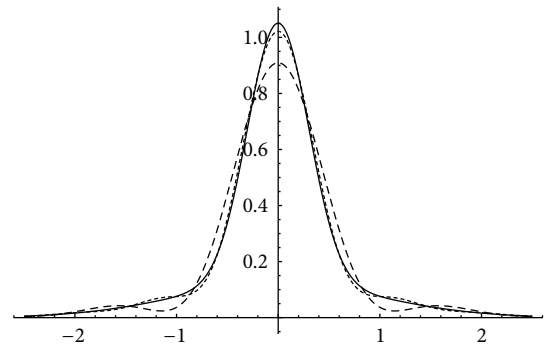


FIGURE 2: Gram-Charlier approximations to the density $g_{0.3}(x)$ (solid line) when $q = 1.4$. The curve with long dashes is the order 10 approximation, and the one with short dashes is the order 20 approximation.

2.4. *Exponential Change of Measure.* The one-dimensional Cameron-Martin formula may be stated as follows: if $X \sim N(\mu, \sigma^2)$ then for $q \in \mathbf{R}$ and $f \geq 0$,

$$\mathbf{E} e^{qX} f(X) = e^{\mu q + (1/2)\sigma^2 q^2} \mathbf{E} f(X + \sigma^2 q). \tag{62}$$

The same property may be expressed in terms of a change of measure. If $X \stackrel{\mathbf{P}}{\sim} N(\mu, \sigma^2)$ and a change of measure is defined by

$$\mathbf{P}' = \frac{e^{qX}}{\mathbf{E}^{\mathbf{P}} e^{qX}} \cdot \mathbf{P}, \tag{63}$$

then $X \stackrel{\mathbf{P}'}{\sim} N(\mu + \sigma^2 q, \sigma^2)$. The next result is an extension of the Cameron-Martin formula for the normal distribution to the Gram-Charlier distributions. There are different though equivalent formulas in Schlögl [10] (for the special case $c_1 = c_2 = 0$).

Theorem 4. *Suppose that $X \stackrel{\mathbf{P}}{\sim} \mathbf{GC}(a, b; c_1, \dots, c_N)$ and that \mathbf{P}' is defined by (63) for $q \in \mathbf{R}$. Then $X \stackrel{\mathbf{P}'}{\sim} \mathbf{GC}(a + b^2 q, b; c'_1, \dots, c'_N)$, where c'_1, \dots, c'_N are found from*

$$\sum_{k=0}^N b^k c'_k s^k = \frac{\sum_{k=0}^N b^k c_k (q + s)^k}{\sum_{j=0}^N b^j c_j q^j}, \quad s \in \mathbf{R}, \tag{64}$$

or, more precisely,

$$c'_k = \frac{1}{\sum_{j=0}^N b^j c_j q^j} \sum_{\ell=k}^N \binom{\ell}{k} b^{\ell-k} c_\ell q^{\ell-k}. \quad (65)$$

Proof. It is sufficient to calculate the mgf of X under \mathbf{P}' :

$$\begin{aligned} \mathbf{E}^{\mathbf{P}'} e^{sX} &= \frac{1}{\mathbf{E}^{\mathbf{P}} e^{qX}} \mathbf{E}^{\mathbf{P}} e^{(q+s)X} = (\dots) \\ &= e^{(a+b^2q)s+(1/2)b^2s^2} \sum_{k=0}^N b^k s^k \frac{1}{\sum_{j=0}^N b^j c_j q^j} \\ &\cdot \sum_{\ell=k}^N \binom{\ell}{k} b^{\ell-k} c_\ell q^{\ell-k}. \end{aligned} \quad (66)$$

□

Example 5. If $X \stackrel{\mathbf{P}}{\sim} \mathbf{GC}(a, b; 0, 0, c_3, c_4)$ and $q = 1$ the change of measure (63) leads to

$$\begin{aligned} c'_1 &= \frac{3b^2c_3 + 4b^3c_4}{1 + b^3c_3 + b^4c_4}, \\ c'_2 &= \frac{3bc_3 + 6b^2c_4}{1 + b^3c_3 + b^4c_4}, \\ c'_3 &= \frac{c_3 + 4bc_4}{1 + b^3c_3 + b^4c_4}, \\ c'_4 &= \frac{c_4}{1 + b^3c_3 + b^4c_4}. \end{aligned} \quad (67)$$

Hence, assuming $c_1 = c_2 = 0$ does not imply that c'_1 and c'_2 are also zero. In other words, the family $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$ is not closed under a change of measure that occurs naturally in option pricing (see Section 5).

2.5. Simulating Gram-Charlier Distributions. Simulation is required for many kinds of options, and it turns out that the Gram-Charlier distributions are very easy to generate, as we now show; there is no need to invert their distribution functions. When estimating some quantity $m = \mathbf{E}g(X_1, \dots, X_s)$ by simulation, one generates n independent vectors

$$(X_1^{(j)}, \dots, X_s^{(j)}), \quad j = 1, \dots, n, \quad (68)$$

with the same distribution. Suppose all the X 's are independent and have the same Gram-Charlier distribution with density

$$f_X(x) = \phi(a, b, x) p(x), \quad (69)$$

where the polynomial $p(x)$ is given in (5). Then

$$\begin{aligned} \mathbf{E}g(X_1, \dots, X_s) &= \int_{\mathbf{R}^s} g(x_1, \dots, x_s) \\ &\cdot \prod_{k=1}^s [\phi(a, b, x_k) p(x_k)] dx_1 \cdots dx_s \\ &= \mathbf{E}^{\mathbf{Q}} \left[g(X_1, \dots, X_s) \prod_{k=1}^s p(X_k) \right], \end{aligned} \quad (70)$$

where under the measure \mathbf{Q} the X 's have a normal distribution $\mathbf{N}(a, b^2)$. Hence, estimating $\mathbf{E}g(X_1, \dots, X_s)$ by simulation can be performed by generating ordinary normal random variables $(X_1^{(j)}, \dots, X_s^{(j)})$, $j = 1, \dots, n$, and then using the estimator

$$\widehat{m}_n = \frac{1}{n} \sum_{j=1}^n \left[g(X_1^{(j)}, \dots, X_s^{(j)}) \prod_{k=1}^s p(X_k^{(j)}) \right]. \quad (71)$$

This is an application of the likelihood ratio method.

3. Convolution of Gram-Charlier Distributions; Gram-Charlier Processes

The simplest way to find the distribution of the sum of two independent Gram-Charlier distributed variables is to multiply their moment generating functions. Suppose $X_j \sim \mathbf{GC}(a_j, b_j; c_1^{(j)}, \dots, c_N^{(j)})$, $j = 1, 2$ are independent. Then

$$\begin{aligned} \mathbf{E}e^{z(X+Y)} &= e^{(a_1+a_2)t+(b_1^2+b_2^2)(t^2/2)} \left(\sum_{k=0}^{N^{(1)}} c_k^{(1)} t^k \right) \left(\sum_{k=0}^{N^{(2)}} c_k^{(2)} t^k \right). \end{aligned} \quad (72)$$

Expanding the product, this says that

$$\begin{aligned} X + Y &\sim \mathbf{GC} \left(a_1 + a_2, \sqrt{b_1^2 + b_2^2}; c_1, \dots, c_{N^{(1)}+N^{(2)}} \right), \\ c_k &= \sum_{j=0}^k c_j^{(1)} c_{k-j}^{(2)}. \end{aligned} \quad (73)$$

It is then possible to know the explicit distribution of

$$Z_n = X_1 + \cdots + X_n, \quad (74)$$

where the X 's are independent and have a Gram-Charlier distribution, constituting a discrete-time Gram-Charlier process. If the X 's have the same $\mathbf{GC}(a, b; c_1, \dots, c_N)$ distribution then $\{Z_n, n \geq 0\}$ is a random walk. The derivation of the distribution of Z_n can be done recursively, using (73), or it can be done by finding the Taylor expansion of

$$\begin{aligned} \left(\sum_{k=0}^N c_k t^k \right)^n &= \sum_{k=0}^{nN} c_k^{(n)} t^k \\ &= 1 + c_1 n t + \left(c_2 n + \frac{1}{2} c_1^2 n(n-1) \right) t^2 + \cdots \\ &\quad + c_N^n t^{nN}, \end{aligned} \quad (75)$$

and thus

$$Z_n \sim \mathbf{GC}\left(an, b\sqrt{n}; c_1n, c_2n + \frac{1}{2}c_1^2n(n-1), \dots, c_N^n\right). \quad (76)$$

These computations are simple using symbolic mathematics software. An example is given at the end of Section 6.1.

The above raises the question of whether there is a continuous-time process that has Gram-Charlier distributed increments. There is such a process with normal increments (Brownian motion), and it is moreover a Lévy process.

Theorem 6. *The only Lévy process with Gram-Charlier distributed increments is Brownian motion.*

Proof. It is sufficient to show that, besides the normal distribution, any Gram-Charlier distribution cannot be infinitely divisible. If X has a Gram-Charlier distribution then its moment generating function is of the form

$$M(z) = \mathbf{E}e^{zX} = e^{az+b^2(z^2/2)}\pi(z), \quad (77)$$

where $\pi(\cdot)$ is a polynomial. Suppose $X_{n,1}, \dots, X_{n,n}$ are independent, have the same distribution, and add up to X (in law). Fubini's theorem implies that $\mathbf{E}e^{tX_{n,1}}$ is finite for all real t , and thus

$$M_n(z) = \mathbf{E}e^{zX_{n,1}} \quad (78)$$

is an analytic function of z in the whole complex plane. We then have the identity

$$M_n(z)^n = M(z), \quad z \in \mathbf{C}, \quad (79)$$

for every $n = 1, 2, 3, \dots$. This means that the function $\pi(z)^{1/n}$, which is well defined for all z that is not a zero of π , has an analytic continuation over the whole of \mathbf{C} for each $n \geq 1$, which is impossible, for instance, take n larger than the degree of π . \square

More precisely, this says that if we exclude Brownian motion, no increment of any Lévy process can have a Gram-Charlier distribution. Any Gram-Charlier process with independent increments must be discrete-time.

4. The Log Gram-Charlier Distribution

The distribution of the exponential of a Gram-Charlier distributed variable will naturally be called “log Gram-Charlier”, as we do for the lognormal: if $Y \sim \mathbf{GC}(a, b; \bar{c})$, then $L = \exp(Y) \sim \mathbf{LogGC}(a, b; \bar{c})$. The density of L is

$$f_L(z) = \frac{1}{z} f_Y(\log z), \quad z > 0. \quad (80)$$

This distribution has all moments finite, and they are given by Theorem 2(b). The log Gram-Charlier distribution shares one property with the lognormal, it is “moment indeterminate.”

Theorem 7. *The log Gram-Charlier distribution is not determined by its moments. More precisely, there is a noncountable number of other distributions that have the same moments as any particular log Gram-Charlier distribution.*

Proof. There is a well-known way to construct a family of distributions that have the same moments as the lognormal (Feller [18], p. 227); the trick works for arbitrary parameters but, for simplicity, let $L \sim \mathbf{LogN}(0, 1)$; that is,

$$f_L(z) = \frac{1}{z\sqrt{2\pi}} e^{-(1/2)\log^2 z}, \quad z > 0. \quad (81)$$

All the functions

$$f_\alpha(z) = f_L(z) (1 + \alpha \sin(2\pi \log z)), \quad -1 \leq \alpha \leq 1, \quad (82)$$

are nonnegative, integrate to 1, and have the same moments as L (as a result of the symmetry of the standard normal distribution, after an obvious change of variable). Now, suppose $L \sim \mathbf{LogGC}(0, 1; \bar{c})$ (once again the same arguments work for other values of a and b), and consider

$$g_\alpha(z) = f_L(z) + \alpha\phi(z) \sin(2\pi \log z). \quad (83)$$

By the same change of variables used for the lognormal one immediately finds that g_α integrates to one and has the same moments as L . The only difference with the case of the lognormal is that g_α is not necessarily nonnegative for $-1 \leq \alpha \leq 1$. Two cases may arise. The first one is that the polynomial

$$p(y) = \sum_{k=0}^N c_k He_k(y) \quad (84)$$

has no real zero. In that case its infimum is strictly greater than zero, and one can find a nonempty interval $I = (-\epsilon, \epsilon)$ such that g_α is nonnegative for all $\alpha \in I$. In the second case $p(\cdot)$ has one or more zeros and the previous argument breaks down but can be modified to yield the same conclusion, if the function $\sin(2\pi y)$ is replaced with

$$s(y) = \sin(2\pi y) \mathbf{1}_{\{y \notin J\}}, \quad (85)$$

where J is a collection of intervals that include the zeros of $p(\cdot)$, so defined that $s(\cdot)$ is not identically zero and satisfies

$$\begin{aligned} s(y + 2k\pi) &= s(y), \\ s(-y) &= -s(y). \end{aligned} \quad (86)$$

(These two conditions are sufficient for

$$\int_{-\infty}^{\infty} e^{ky} \phi(y) s(y) dy = 0 \quad (87)$$

to hold for $k = 1, 2, \dots$) The details are omitted. Another way to prove that the log Gram-Charlier distribution is moment indeterminate when $p(y)$ has no zero (and thus the density $f_L(z)$ does not take the value zero for any $z > 0$) is to use a Krein condition (Stoyanov [19], p. 941), which says that a continuous distribution on \mathbf{R}_+ with positive density $f(\cdot)$ is not determined by its moments if

$$-\int_0^\infty \frac{\log f(z^2)}{1+z^2} dz < \infty. \quad (88)$$

□

5. Option Pricing Formulas

The formulas below hold for any vector (c_1, \dots, c_N) and thus extend those that have been derived by previous authors for the case where the log return has a $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$ distribution or a “squared” Gram-Charlier distribution (León et al. [11]). Schlögl [10] derives a formula equivalent to (90) but for densities expressed as an infinite Gram-Charlier series with $c_1 = c_2 = 0$.

As previous authors have done, we consider a market with a risky security S and a risk-free security with annual return r and suppose that the log return for period $[0, T]$ (denoted as X_T) has a Gram-Charlier distribution under the risk-neutral (or “pricing”) measure, which we denote as \mathbf{Q} . The physical measure (usually denoted as \mathbf{P}) is not specified (nothing says that the log return has or does not have a Gram-Charlier distribution under the physical measure).

The market model may have one or more periods, but since we consider the pricing of ordinary European puts and calls only the distribution of the log return for the whole period $[0, T]$ is required. For other types of options, in particular the applications presented in the next section, it may be necessary to use the one-period returns separately, as is done in Theorem 11.

The risky security has price S_0 at time 0, and $S_T = S_0 e^{X_T}$.

Theorem 8. *Suppose that a risky security pays dividends at a constant rate δ over $[0, T]$ and that the risk-free rate of interest is r . Suppose also that under the risk-neutral measure \mathbf{Q} the log return of the risky security over $[0, T]$ is $X_T \stackrel{\mathbf{Q}}{\sim} \mathbf{GC}(a, b; c_1, \dots, c_N)$, which satisfies the martingale condition*

$$e^{a+(b^2/2)} \sum_{k=0}^N b^k c_k = e^{(r-\delta)T}. \quad (89)$$

Then the time-0 price of a European call option with maturity T and strike price K is

$$C_0 = S_0 e^{-\delta T} \Phi(d_1) - Ke^{-rT} \Phi(d_2) + Ke^{-rT} \phi(d_2) \cdot \sum_{k=1}^N [c_k^* He_{k-1}(-d_1) - c_k He_{k-1}(-d_2)], \quad (90)$$

where d_2 , d_1 , and $\{c_k^*\}$ are given by (97), (101), and (105). The price of the corresponding European put is

$$P_0 = Ke^{-rT} \Phi(-d_2) - S_0 e^{-\delta T} \Phi(-d_1) - Ke^{-rT} \phi(d_2) \cdot \sum_{k=1}^N [c_k^* He_{k-1}(-d_1) - c_k He_{k-1}(-d_2)]. \quad (91)$$

If $N = 4$, then the above simplify to

$$C_0 = S_0 e^{-\delta T} \Phi(d_1) - Ke^{-rT} \Phi(d_2) + bKe^{-rT} \phi(d_2) \cdot [c_2 + (b-d_2)c_3 + (b^2 - bd_2 + d_2^2 - 1)c_4], \quad (92)$$

$$P_0 = Ke^{-rT} \Phi(-d_2) - S_0 e^{-\delta T} \Phi(-d_1) - bKe^{-rT} \phi(d_2) \cdot [c_2 + (b-d_2)c_3 + (b^2 - bd_2 + d_2^2 - 1)c_4].$$

Proof. Absence of arbitrage implies that $\mathbf{E}^{\mathbf{Q}} S_T = e^{(r-\delta)T} S_0$, or

$$S_0 e^{a+(b^2/2)} \sum_{k=0}^N b^k c_k = e^{(r-\delta)T} S_0. \quad (93)$$

This justifies (89). The price of the call is

$$C_0 = e^{-rT} \mathbf{E}^{\mathbf{Q}} (S_T - K)_+ = C_0^+ - C_0^-$$

$$C_0^+ = e^{-rT} \mathbf{E}^{\mathbf{Q}} S_T (\mathbf{1}_{\{S_T > K\}}), \quad (94)$$

$$C_0^- = Ke^{-rT} \mathbf{E}^{\mathbf{Q}} (\mathbf{1}_{\{S_T > K\}}).$$

The second part is easier to deal with

$$C_0^- = Ke^{-rT} \mathbf{Q} \left(X_T > \log \left(\frac{K}{S_0} \right) \right). \quad (95)$$

The probability of the event $X_T > \log(K/S_0)$ can be calculated explicitly by recalling that $(X_T - a)/b \stackrel{\mathbf{Q}}{\sim} \mathbf{GC}(0, 1; c_1, \dots, c_N)$ and using Theorem 3:

$$\mathbf{Q} \left(X_T > \log \left(\frac{K}{S_0} \right) \right) = \mathbf{Q} \left(\frac{X_T - a}{b} > \frac{1}{b} \left(\log \left(\frac{K}{S_0} \right) - a \right) \right) \quad (96)$$

$$= \Phi(d_2) + \phi(d_2) \sum_{k=1}^N c_k He_{k-1}(-d_2),$$

where

$$d_2 = \frac{1}{b} \left(\log \left(\frac{S_0}{K} \right) + a \right). \quad (97)$$

To calculate C_0^+ use the exponential change of measure formula, defining

$$\mathbf{Q}' = \frac{S_T}{\mathbf{E}^{\mathbf{Q}} S_T} \cdot \mathbf{Q} = \frac{e^{X_T}}{\mathbf{E}^{\mathbf{Q}} e^{X_T}} \cdot \mathbf{Q}. \quad (98)$$

Then

$$\mathbf{E}^{\mathbf{Q}} \left(S_T \mathbf{1}_{\{S_T > K\}} \right) = \left(\mathbf{E}^{\mathbf{Q}} S_T \right) \mathbf{E}^{\mathbf{Q}'} \left(\mathbf{1}_{\{S_T > K\}} \right). \quad (99)$$

Since $X_T \stackrel{\mathbf{Q}'}{\sim} \mathbf{GC}(a + b^2, b; c'_1, \dots, c'_N)$,

$$\begin{aligned} \left(\mathbf{E}^{\mathbf{Q}} S_T \right) \mathbf{E}^{\mathbf{Q}'} \left(\mathbf{1}_{\{S_T > K\}} \right) &= S_0 e^{(r-\delta)T} \mathbf{Q}' \left(\frac{X_T - a - b^2}{b} \right) \\ &> \frac{1}{b} \left(\log \left(\frac{K}{S_0} \right) - a - b^2 \right) = S_0 e^{(r-\delta)T} \Phi(d_1) \\ &+ S_0 e^{(r-\delta)T} \phi(d_1) \sum_{k=1}^N c'_k He_{k-1}(-d_1), \end{aligned} \quad (100)$$

where

$$c'_k = \frac{1}{\sum_{j=0}^N b^j c_j} \sum_{\ell=k}^N \binom{\ell}{k} b^{\ell-k} c_\ell, \quad k = 1, \dots, N, \quad (101)$$

$$d_1 = \frac{1}{b} \left[\log \left(\frac{S_0}{K} \right) + a + b^2 \right] = d_2 + b.$$

Hence,

$$\begin{aligned} C_0 &= S_0 e^{-\delta T} \Phi(d_1) - K e^{-rT} \Phi(d_2) \\ &+ S_0 e^{-\delta T} \phi(d_1) \sum_{k=1}^N c'_k He_{k-1}(-d_1) \\ &- K e^{-rT} \phi(d_2) \sum_{k=1}^N c_k He_{k-1}(-d_2). \end{aligned} \quad (102)$$

Using

$$\frac{S_0 e^{-\delta T} \phi(d_1)}{\sum_{k=0}^N b^k c_k} = S_0 e^{a+(b^2/2)-rT} \phi(d_1) = K e^{-rT} \phi(d_2) \quad (103)$$

(a consequence of (89) and (101)), it is possible to write

$$S_0 e^{-\delta T} \phi(d_1) c'_k = K e^{-rT} \phi(d_2) c_k^*, \quad (104)$$

where

$$c_k^* = c'_k \sum_{k=0}^N b^k c_k = \sum_{\ell=k}^N \binom{\ell}{k} b^{\ell-k} c_\ell, \quad k = 1, \dots, N. \quad (105)$$

This proves (90).

The price of a European put can be found from the put-call parity identity

$$P_0 = C_0 - S_0 + K e^{-rT}. \quad (106)$$

Finally, if $N = 4$ then the summation in (90) becomes

$$\begin{aligned} \sum_{k=1}^4 c_k^* He_{k-1}(-d_1) - \sum_{k=1}^4 c_k He_{k-1}(-d_2) \\ = b c_2 + (b^2 - b d_2) c_3 + (b^3 - b^2 d_2 + b d_2^2 - b) c_4. \end{aligned} \quad (107)$$

This ends the proof. \square

The option price formulas are of the form ‘‘Black-Scholes plus correction term.’’ Observe, however, that the values of d_1 and d_2 are different from what they would be in the Black-Scholes formula. More precisely, in the Black-Scholes model (where $c_k = 0$ for all $k \geq 1$) we have $a = \mathbf{E}^{\mathbf{Q}} X_T$ and $b^2 = \mathbf{Var}^{\mathbf{Q}} X_T$, but this does not happen with Gram-Charlier distributed log returns, first because of the martingale condition (89) and second because of the result in part (j) of Theorem 2.

5.1. Sensitivities (‘‘Greeks’’). The previous literature includes formulas for sensitivities of Gram-Charlier option prices, but only in the case of the four-parameter $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$ Gram-Charlier distributions; see Jurczenko et al. [7], Rouah and Vainberg [20], and Chateau [3]. Below we give sensitivities of the option prices calculated above with respect to all the parameters for a general Gram-Charlier distribution, taking the martingale condition (89) into account. This means that a is a function of $\delta, b, \{c_k\}, r$ and T (we might have written $a = a(\delta, b, \{c_k\}, r, T)$). However, a is not a function of S_0 .

Theorem 9. *Let C_0 be the price of the European call option described in Theorem 8. Then*

$$\begin{aligned} \Delta &= \frac{\partial C_0}{\partial S_0} \\ &= e^{-\delta T} \left(\Phi(d_1) + \phi(d_1) \sum_{k=1}^N c'_k He_{k-1}(-d_1) \right), \\ \gamma &= \frac{\partial^2 C_0}{\partial S_0^2} = \frac{e^{a+(b^2/2)-rT}}{b S_0} \phi(d_1) \sum_{k=0}^N c_k He_k(-d_2), \\ \rho &= \frac{\partial C_0}{\partial r} \\ &= K T e^{-rT} \left(\Phi(d_2) + \phi(d_2) \sum_{k=1}^N c_k He_{k-1}(-d_2) \right), \end{aligned} \quad (108)$$

$$\begin{aligned} \kappa &= \frac{\partial C_0}{\partial b} = S_0 e^{-\delta T} \phi(d_1) \\ &\cdot \sum_{k=0}^N [c'_k - c'_1 c'_{k+1} + (k+2) c'_{k+2}] He_k(-d_1), \end{aligned}$$

$$\frac{\partial C_0}{\partial c_j} = K e^{-rT} \phi(d_2)$$

$$\cdot \left(\sum_{k=1}^{j-1} b^k He_{j-1-k}(-d_2) - b^j \sum_{k=1}^N c'_k He_{k-1}(-d_1) \right), \quad j \geq 1.$$

In the formula for κ the symbols c'_{N+1} and c'_{N+2} are equal to zero.

Proof. The following lemma is obtained by elementary calculations.

Lemma 10. For any integrable random variable U and any constant K ,

$$\frac{\partial}{\partial s} \mathbf{E}(sU - K)_+ = \mathbf{E}(U \mathbf{1}_{\{sU > K\}}). \quad (109)$$

If U has a continuous density f_U , then

$$\frac{\partial^2}{\partial s^2} \mathbf{E}(sU - K)_+ = \frac{K^2}{s^3} f_U\left(\frac{K}{s}\right). \quad (110)$$

From $S_T = S_0 e^{X_T}$ and (100),

$$\begin{aligned} \Delta &= \frac{\partial}{\partial S_0} \left[e^{-rT} \mathbf{E}^{\mathbf{Q}}(S_0 e^{X_T} - K)_+ \right] = \frac{1}{S_0} C_0^+ \\ &= e^{-\delta T} \left(\Phi(d_1) + \phi(d_1) \sum_{k=1}^N c_k' He_{k-1}(-d_1) \right). \end{aligned} \quad (111)$$

From the lemma,

$$\begin{aligned} \gamma &= \frac{\partial^2}{\partial S_0^2} C_0 = e^{-rT} \frac{\partial^2}{\partial S_0^2} \mathbf{E}^{\mathbf{Q}}(S_0 U - K)_+ \\ &= \frac{K^2 e^{-rT}}{S_0^3} f_U\left(\frac{K}{S_0}\right), \quad U = e^X, \end{aligned} \quad (112)$$

where $X_T \stackrel{\mathbf{Q}}{\sim} \mathbf{GC}(a, b; c_1, \dots, c_N)$. Now the density of U may be expressed in terms of the density of $G = (X_T - a)/b \stackrel{\mathbf{Q}}{\sim} \mathbf{GC}(0, 1; c_1, \dots, c_N)$:

$$\begin{aligned} f_U(u) &= \frac{1}{bu} f_G\left(\frac{1}{b}(\log u - a)\right) \\ &= \frac{1}{bu} \phi(y) \sum_{k=0}^N c_k He_k(y) \Big|_{y=(1/b)(\log u - a)}. \end{aligned} \quad (113)$$

Setting $u = K/S_0$ and then using (103), we obtain

$$\begin{aligned} \gamma &= \frac{K e^{-rT}}{b S_0^2} \phi(d_2) \sum_{k=0}^N c_k He_k(-d_2) \\ &= \frac{e^{a+(b^2/2)-rT}}{b S_0} \phi(d_1) \sum_{k=0}^N c_k He_k(-d_2). \end{aligned} \quad (114)$$

Next,

$$\rho = \frac{\partial}{\partial r} \mathbf{E}^{\mathbf{Q}}(S_0 e^{a+bG-rT} - K e^{-rT})_+ = \frac{\partial}{\partial r} \mathbf{E}^{\mathbf{Q}} h(G, r), \quad (115)$$

where $G \stackrel{\mathbf{Q}}{\sim} \mathbf{GC}(0, 1, \bar{c})$. For fixed G , the function $g(r) = h(G, r)$ is absolutely continuous and thus

$$\rho = \mathbf{E}^{\mathbf{Q}} \left[\frac{\partial}{\partial r} (S_0 e^{a+bG-rT} - K e^{-rT}) \right] \mathbf{1}_{\{S_0 e^{a+bG-rT} > K e^{-rT}\}}. \quad (116)$$

Now, from (89),

$$e^{a+bG-rT} = \frac{e^{bG-(b^2/2)-\delta T}}{\sum_{k=0}^N b^k c_k} \quad (117)$$

does not depend on r , so

$$\begin{aligned} \rho &= \mathbf{E}^{\mathbf{Q}} \left(K T e^{-rT} \mathbf{1}_{\{S_0 e^{a+bG-rT} > K e^{-rT}\}} \right) \\ &= K T e^{-rT} \left[\Phi(d_2) + \phi(d_2) \sum_{k=1}^N c_k He_{k-1}(-d_2) \right]. \end{aligned} \quad (118)$$

The same line of reasoning shows that the sensitivity to b is

$$\kappa = S_0 e^{-rT} \mathbf{E}^{\mathbf{Q}} \left[\left(\frac{\partial}{\partial b} e^{a+bG} \right) \mathbf{1}_{\{S_0 e^{a+bG} > K\}} \right]. \quad (119)$$

Define

$$\begin{aligned} m(b) &= \mathbf{E} e^{bG-(b^2/2)} = \sum_{k=0}^N b^k c_k, \\ m'(b) &= \sum_{k=1}^N k b^{k-1} c_k = c_1^*. \end{aligned} \quad (120)$$

In order to shorten the formulas we will write (from (89))

$$e^a = \frac{e^{(r-\delta)T-(b^2/2)}}{m(b)} \quad (121)$$

and thus find an expression for

$$\begin{aligned} S_0 e^{-\delta T} \mathbf{E}^{\mathbf{Q}} \left[\left(\frac{\partial}{\partial b} \frac{e^{bG-(b^2/2)}}{m(b)} \right) \mathbf{1}_{\{G > -d_2\}} \right] \\ = \frac{S_0 e^{-\delta T-(b^2/2)}}{m(b)^2} \end{aligned} \quad (122)$$

$$\begin{aligned} &\cdot \mathbf{E}^{\mathbf{Q}} \left[\left(G e^{bG} m(b) - e^{bG} (m'(b) + b m(b)) \right) \right. \\ &\left. \cdot \mathbf{1}_{\{G > -d_2\}} \right]. \end{aligned}$$

For arbitrary $y \in \mathbf{R}$,

$$\begin{aligned} \mathbf{E}^{\mathbf{Q}}(e^{bG} \mathbf{1}_{\{G > -y\}}) &= e^{b^2/2} \left[m(b) \Phi(y + b) \right. \\ &\left. + \phi(y + b) \sum_{k=1}^N c_k^* He_{k-1}(-y - b) \right]. \end{aligned} \quad (123)$$

To calculate

$$\mathbf{E}^{\mathbf{Q}}(G e^{bG} \mathbf{1}_{\{G > -y\}}) = \frac{\partial}{\partial b} \mathbf{E}^{\mathbf{Q}}(e^{bG} \mathbf{1}_{\{G > -y\}}) \quad (124)$$

we need (see (105))

$$\begin{aligned} \frac{\partial c_k^*}{\partial b} &= \frac{\partial}{\partial b} \sum_{\ell=k}^N \binom{\ell}{k} b^{\ell-k} c_\ell = (k+1) c_{k+1}^*, \\ &k = 1, \dots, N-1, \end{aligned} \quad (125)$$

$$\frac{\partial c_N^*}{\partial b} = 0.$$

Then, noting that $m'(b) = c_1^* = c_1' m(b)$ and $(\phi He_{k-1})' = -\phi He_k$,

$$\begin{aligned} \mathbf{E}^{\mathbf{Q}} \left(G e^{bG} \mathbf{1}_{\{G > -y\}} \right) &= b \mathbf{E}^{\mathbf{Q}} \left(e^{bG} \mathbf{1}_{\{G > -y\}} \right) + e^{b^2/2} m(b) \\ &\cdot \left[c_1' \Phi(y+b) + \phi(y+b) \right. \\ &\cdot \sum_{k=0}^{N-2} (k+2) c_{k+2}' He_k(-y-b) + \phi(y+b) \\ &\cdot \left. \sum_{k=0}^N c_k' He_k(-y-b) \right]. \end{aligned} \tag{126}$$

Setting $y = d_2$ and putting the pieces together yield the result. When $\tilde{c} = 0$ this reduces to $S_0 e^{-\delta T} \phi(d_1)$, which is the sensitivity of the Black-Scholes call price to $b = \sigma \sqrt{T}$.

Finally, turn to the sensitivities with respect to c_j , $j \geq 1$. First, let us derive $\partial a / \partial c_j$; from (89),

$$\begin{aligned} \frac{\partial a}{\partial c_j} e^{a+(b^2/2)} \sum_{k=0}^N b^k c_k + b^j e^{a+(b^2/2)} &= 0 \\ \text{or } \frac{\partial a}{\partial c_j} &= -\frac{b^j}{\sum_{k=0}^N b^k c_k} = -b^j e^{a+(b^2/2)+(\delta-r)T}. \end{aligned} \tag{127}$$

Next,

$$\begin{aligned} e^{-rT} \frac{\partial}{\partial c_j} \mathbf{E}^{\mathbf{Q}} \left(S_0 e^{a+bG} - K \right)_+ &= e^{-rT} \frac{\partial}{\partial c_j} \int_{-d_2}^{\infty} (S_0 e^{a+by} - K) \phi(y) \sum_{k=0}^N c_k He_k(y) dy \\ &= e^{-rT} \int_{-d_2}^{\infty} S_0 \frac{\partial a}{\partial c_j} e^{a+by} \phi(y) \sum_{k=0}^N c_k He_k(y) dy \\ &\quad + e^{-rT} \int_{-d_2}^{\infty} (S_0 e^{a+by} - K) \phi(y) He_j(y) dy. \end{aligned} \tag{128}$$

The first integral reduces to $S_0 \Delta \partial a / \partial c_j$ (see the derivation of Δ above). The second integral is evaluated by repeated partial integration: if $g(x)$ is continuously differentiable k times and does not grow too fast,

$$\begin{aligned} \int_y^{\infty} g(x) \phi(x) He_j(x) dx &= \phi(y) \sum_{k=0}^{j-1} g^{(k)}(y) He_{j-1-k}(y) \\ &\quad + \int_y^{\infty} g^{(j)}(x) \phi(x) dx. \end{aligned} \tag{129}$$

Hence the second integral in (128) becomes

$$\begin{aligned} S_0 e^{a-bd_2-rT} \phi(d_2) \sum_{k=1}^{j-1} b^k He_{j-1-k}(-d_2) &+ S_0 e^{a-rT} \int_{-d_2}^{\infty} b^j e^{by} \phi(y) dy \\ &= S_0 e^{a-bd_2-rT} \phi(d_2) \sum_{k=1}^{j-1} b^k He_{j-1-k}(-d_2) \\ &\quad + b^j S_0 e^{a-rT+(b^2/2)} \Phi(d_1). \end{aligned} \tag{130}$$

Putting all this together (using (103)) we get the formula for $\partial C_0 / \partial c_j$. \square

6. Two Applications

6.1. Equity Indexed Annuities. Hardy [12] describes the main types of equity indexed annuities (EIAs), point-to-point, annual ratchet and high water mark, in which an embedded European call option has to be priced. We consider compound ratchet EIAs without life-of-contract guarantee to explore the dependence of ratchet premium options on skewness and kurtosis of returns.

A single premium P is paid by the policy-holder, and the benefit under the ratchet premium contract is ([12], p. 248)

$$B_n = P \prod_{t=1}^n \left[1 + \max \left(\alpha \left(\frac{S_t}{S_{t-1}} - 1 \right), 0 \right) \right]. \tag{131}$$

Here, n is the term of the contract in years, α is the participation rate (a number between 0 and 1), and S_t is the value of the equity index, usually the S&P500 index.

In [12], p. 249, the formula

$$P \left[e^{-r} + \alpha \left(e^{-\delta} \Phi(d_1) - e^{-r} \Phi(d_2) \right) \right]^n \tag{132}$$

for B_n is proved for the value of the premium option under a compound annual ratchet contract in the Black-Scholes model. The same reasoning will now be applied when index log returns have a Gram-Charlier distribution. To find a formula for the price of this option, we assume that under the risk-neutral measure the one-year log returns are independent Gram-Charlier distributed random variables

$$R_t \stackrel{\mathbf{Q}}{\sim} \mathbf{GC}(a, b; \tilde{c}), \quad t = 1, \dots, n. \tag{133}$$

Independence of the variables across time implies that the value of the ratchet premium option is (if the annual risk-free rate of interest is r)

$$\begin{aligned} V_0 &= \mathbf{E}^{\mathbf{Q}} \left(e^{-rn} B_n \right) \\ &= P \prod_{t=1}^n \left\{ e^{-r} \mathbf{E}^{\mathbf{Q}} \left[1 + \max \left(\alpha \left(\frac{S_t}{S_{t-1}} - 1 \right), 0 \right) \right] \right\}. \end{aligned} \tag{134}$$

As Hardy notes ([12], p. 249), each factor in the product is the price of a one-year European call option on S_t , if initial index price and strike are both equal to one.

Theorem 11. *The no-arbitrage price of the EIA ratchet premium option described above is*

$$V_0 = P(e^{-r} + \alpha C_0)^n, \quad (135)$$

where C_0 is the price of a one-year call (see (90)) with $S_0 = 1$ and $K = 1$. The first-order sensitivities are

$$\frac{\partial V_0}{\partial \xi} = \alpha P(e^{-r} + \alpha C_0)^{n-1} \frac{\partial C_0}{\partial \xi}, \quad (136)$$

with ξ replaced with one of S_0, r, b, c_j , and $\partial C_0 / \partial \xi$ given in Theorem 9. The second-order sensitivity with respect to S_0 is

$$\begin{aligned} \frac{\partial^2 C_0}{\partial S_0^2} &= \alpha^2 P(e^{-r} + \alpha C_0)^{n-2} \left(\frac{\partial C_0}{\partial S_0} \right)^2 \\ &+ \alpha P(e^{-r} + \alpha C_0)^{n-1} \frac{\partial^2 C_0}{\partial S_0^2}. \end{aligned} \quad (137)$$

We now show the effect of skewness and kurtosis on the ratchet premium option values; this is done using the four-parameter distribution $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$, for which this is especially simple. We also compute prices using the six-parameter $\mathbf{GC}(a, b; c_1, c_2, c_3, c_4)$ distribution.

Our application involves an annual ratchet, and so the option C_0 in Theorem 11 has a maturity of one year; we thus need the distribution of one-year returns. For pricing that is consistent with the (derivatives) market what is needed is the one-year distribution of returns under the risk-neutral measure, which should be obtained from observed derivative prices; see León et al. [11] or Rompolis and Tzavalis [21] for more details. For illustrative purposes, we estimated the four parameters of that distribution by maximum likelihood ([6, 22]) to obtain the parameters of the $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$:

$$\begin{aligned} a &= 0.0687, \\ b &= 0.1685, \\ c_3 &= -0.1150, \\ c_4 &= 0.03598. \end{aligned} \quad (138)$$

This says that the skewness and excess kurtosis of the fitted distribution are

$$\begin{aligned} \bar{s} &= 6c_3 = -0.6898, \\ \bar{k} &= 24c_4 = 0.8634. \end{aligned} \quad (139)$$

This point is on the boundary of the feasible (\bar{s}, \bar{k}) region; see Figure 3.

We use the parameters a and b while varying the values of c_3, c_4 to show how EIA premiums vary with skewness and excess kurtosis. Figure 4 shows the ratchet premium option values (Theorem 11) of an EIA with participation rate $\alpha = 0.6$ and single premium $P = 100$, as a function of skewness \bar{s} and

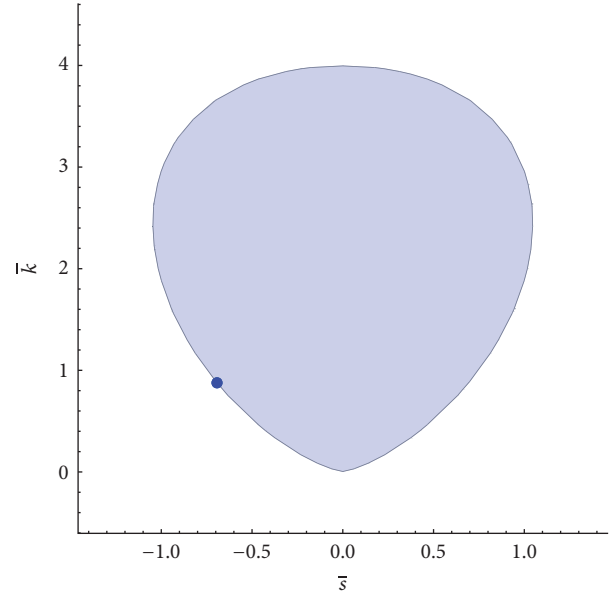


FIGURE 3: Feasible region for the skewness and excess kurtosis of the $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$ distributions and maximum likelihood estimate.

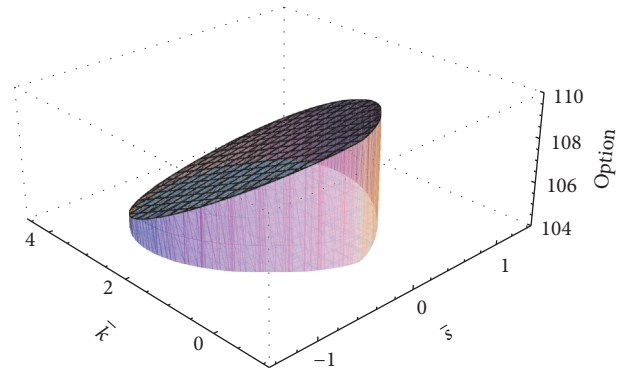


FIGURE 4: Ratchet premium option prices as a function of skewness and excess kurtosis, with $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$ distributions.

excess kurtosis \bar{k} . The risk-free rate is 3%, the dividend rate is 2%, and the term is 7 years. The martingale condition (89) must hold and so a is replaced with a' so that (89) is satisfied. (N.B. The Black-Scholes model corresponds to $c_3 = c_4 = 0$, in which case $a' = (r - \delta)T - b^2/2$; this is -0.004193 with the values of r, δ , and b we are using and $T = 1$.)

The effect of varying (c_3, c_4) is quite significant. The highest and lowest ratchet premium options in the graph are 109.30 to 104.24, which is the range [95%, 100%] as a proportion of the Black-Scholes premium option (\$109.26), which has skewness and excess kurtosis equal to zero. (The maximum value of the premium option is reached at $(c_3, c_4) = (0.0230, 0.00332567)$ and the minimum at $(-0.0836, 0.163)$; these points correspond to (\bar{s}, \bar{k}) equal to $(.138, .0798)$ and

(0.023, 0.003326), resp.) The dependence on (\bar{s}, \bar{k}) is nearly linear. This is easily explained. The price of a call is

$$\begin{aligned}
 C_0 &= e^{-rT} \mathbf{E}^Q(S_T - K) \\
 &= e^{-rT} \int_{(1/b)(\log(K/S_0) - a')}^{\infty} (S_0 e^{a' + bx} - K) \phi(x) \\
 &\quad \cdot \left(\sum_{k=0}^N c_k He_k(x) \right) dx.
 \end{aligned} \tag{140}$$

The sensitivity of C_0 to c_j ($j = 3, 4$) is thus

$$\begin{aligned}
 e^{-rT} \frac{\partial a'}{\partial c_j} \mathbf{E}^Q(S_T \mathbf{1}_{S_T > K}) \\
 + e^{-rT} \int_{(1/b)(\log(K/S_0) - a')}^{\infty} (S_0 e^{a' + bx} - K) \phi(x) \\
 \cdot He_j(x) dx.
 \end{aligned} \tag{141}$$

The first term above is relatively small, and the second one is almost constant, since a' does not change much with c_j . This carries over to the sensitivity of the ratchet premium option V_0 to changes in c_j (see the first formula in Theorem 11).

The parameters fitted by maximum likelihood for the $\mathbf{GC}(a, b; c_1, c_2, c_3, c_4)$ distribution are

$$\begin{aligned}
 a &= 0.1174, \\
 b &= 0.1595, \\
 c_1 &= -0.3053675695201066, \\
 c_2 &= 0.09542079373489153, \\
 c_3 &= -0.12383971126335243, \\
 c_4 &= 0.06120331530131559.
 \end{aligned} \tag{142}$$

The martingale condition implies $a' = 0.0451$, and the EIA premium option has value 107.90. The skewness and excess kurtosis of the fitted distribution are $\bar{s} = -0.5437$ and $\bar{k} = 0.5092$. (*N.B.* The feasible region for (\bar{s}, \bar{k}) in the $\mathbf{GC}(a, b; c_1, c_2, c_3, c_4)$ case is naturally larger than the one for the $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$; in this case (\bar{s}, \bar{k}) are just outside the region in Figure 3.) It was checked numerically that the parameters c_1 to c_4 lead to a nonnegative function.

Table 1 shows ratchet premium option values based on Gram-Charlier with four parameters that are lower than the Black-Scholes based ones. Rows A and F are Black-Scholes cases, row B is our estimation of the four-parameter Gram-Charlier (as explained above), row C is the four-parameter Gram-Charlier with the largest negative skewness possible (see Figure 1), row D is similar but with the largest positive skewness, and row E has maximum kurtosis but no skewness. Rows A to E have the same parameter b , the one we estimated from S&P500 for a four-parameter Gram-Charlier, while row F has the same b as the six-parameter Gram-Charlier we estimated. Skewness and kurtosis are shown, followed by the

break-even (or “fair”) α (“B-E α ” in the table) and then the ratchet premium option values for two specific choices for α . Percentage differences with the values in row A (resp., F) are also displayed (“ $\Delta A\%$ ” in the table). It is notable that in all cases the effect of moving away from Black-Scholes is more pronounced for the break-even participation rates than for the option values themselves. The scenario with the greatest impact (compared to Black-Scholes) is the one zero skewness and maximal excess kurtosis (row E). Insurers are unlikely to accept a cost of contract (V_0) greater than the received initial premium (\$100); they may thus wish to adjust the participation rate α down to the break-even rate (to wit, see Lin and Tan [23]). The Black-Scholes break-even participation rate in row A is 41.9%, but it increases from 44.3% up to 49.3% in rows B to E. These are significant increases, ranging from 5.5% to 17.5%. Rows F and G are based on two quite different risk-neutral probability distributions, but the ratchet premium options are surprisingly similar. The parameters we estimated may not be those that a practitioner would use, but what this example says is that skewness and kurtosis have a very real importance in the pricing of equity indexed annuities and that the fair participation rate is, at least here, more sensitive to nonzero skewness and excess kurtosis than are option values. (See also [24] regarding the separate effects of skewness and excess kurtosis on option values.)

Let us compare this with the approach in [11]. That paper is about a subclass of the Gram-Charlier distributions described in this paper; the authors call their distributions “seminonparametric” and have density (6), where $p(\cdot)$ is the square of the polynomial $q(\cdot)$. If $q(x)$ is of order m , then it is possible to express $p(x) = q(x)^2$ as

$$p(x) = \sum_{j=0}^{2m} \delta_j He_j(x). \tag{143}$$

However, most nonnegative polynomials cannot be expressed as squares of another polynomial, so the order- m SNP family is a strict subset of the order- $2m$ Gram-Charlier family. The order $2m$ SNP family has $2m + 2$ free parameters (the coefficients of an order- $2m$ polynomial plus the location and dispersion parameters a, b , subject to density integrating to one, which in effect removes one parameter); by comparison, the $\mathbf{GC}(a, b; c_1, \dots, c_{2m})$ distributions have $2m + 2$ parameters, but they are not “free” because the resulting density must be nonnegative. It is then not *a priori* clear which of $\mathbf{SNP}(2m)$ or $\mathbf{GC}(a, b; c_1, \dots, c_{2m})$ would do best. We looked at this problem with the S&P500 annual prices, comparing $\mathbf{SNP}(2)$ and $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$, which both have four parameters. The likelihood function is written the same way in both cases; there are just different restrictions on the parameters. The results for $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$ are given in (138), and those for the $\mathbf{SNP}(2)$ are

$$\begin{aligned}
 a &= -0.2445, \\
 b &= 0.1940, \\
 c_1 &= 1.6140,
 \end{aligned}$$

TABLE 1: Comparison of ratchet option values computed with the Black-Scholes formula and Gram-Charlier distributions. Rows A to F are based on the four-parameter $GC(a, b; 0, 0, c_3, c_4)$, while row G is based on a six-parameter $GC(a, b; c_1, c_2, c_3, c_4)$.

	a'	b	c_1	c_2	c_3	c_4	\bar{s}	\bar{k}	B-E α	$\Delta(A)\%$	$V_0 (\alpha = 0.6)$	$\Delta(A)\%$	$V_0 (\alpha = 0.419)$	$\Delta(A)\%$
A	-0.0042	0.1685	0	0	0	0	0	0	0.419	0	109.26	0	100.00	0
B	-0.0035	0.1685	0	0	-0.1150	0.0360	-0.6898	0.8634	0.443	5.5	107.60	-1.5	98.92	-1.1
C	-0.0034	0.1685	0	0	-0.1749	0.1021	-1.0493	2.4508	0.478	13.9	105.42	-3.5	97.50	-2.5
D	-0.0051	0.1685	0	0	0.1749	0.1021	1.0493	2.4508	0.446	6.3	107.39	-1.7	98.78	-1.2
E	-0.0043	0.1685	0	0	0	0.1667	0	4.0000	0.493	17.5	104.59	-4.3	96.96	-3.0
	a'	b	c_1	c_2	c_3	c_4	\bar{s}	\bar{k}	B-E α	$\Delta(F)\%$	$V_0 (\alpha = 0.6)$	$\Delta(F)\%$	$V_0 (\alpha = 0.441)$	$\Delta(F)\%$
F	0.0027	0.1595	0	0	0	.0	0	0	0.441	0	107.69	0	100.00	0
G	0.0451	0.1595	-.3054	.09542	-0.12384	0.06120	-0.5437	0.5091	0.438	-0.7	107.90	0.2	100.14	0.14

$$\begin{aligned}
 c_2 &= 1.1733, \\
 c_3 &= 0.4253, \\
 c_4 &= 0.07320.
 \end{aligned}
 \tag{144}$$

This translates into the following moments for the $SNP(2)$ fitted distribution:

$$\begin{aligned}
 \text{mean} & 0.0687 \\
 \text{standard deviation} & 0.1671 \\
 \text{skewness} & -0.6289 \\
 \text{excess kurtosis} & 2.5502.
 \end{aligned}
 \tag{145}$$

The excess kurtosis is strikingly far from the data's 0.8903. The likelihood function evaluated at that estimate was lower than the likelihood evaluated at the parameters of the $GC(a, b; 0, 0, c_3, c_4)$ distribution. This is a puzzling result that does not mean that $GC(a, b; 0, 0, c_3, c_4)$ will always do better than $SNP(2)$; it does however show that restricting the search to squares of polynomials does have consequences. León et al. [11] derive formulas for European option prices, taking the martingale restriction into account, based on expression (143); in this paper we do the same calculation for an arbitrary Gram-Charlier distribution. It is easy to see that the SNP class is not closed under convolution; that is, the distribution of the sum of two independent SNP variables is in general not an SNP distribution; the general formulas we derive are essential in order to add independent Gram-Charlier distributed variables (including those in the SNP subset), and therefore the general formulas are needed to define Gram-Charlier processes. León et al. [11] are correct in pointing out that there is an advantage in using $p(x) = q(x)^2$ as far as the nonnegativity constraint is concerned, but we believe that it is no more difficult to estimate the parameters of a general Gram-Charlier distribution, including the non-negativity constraint in the maximization procedure (rather than limiting the parameter space). A full comparison of the numerical and statistical advantages/disadvantages of our approach versus the one in [11] is an interesting avenue for further research.

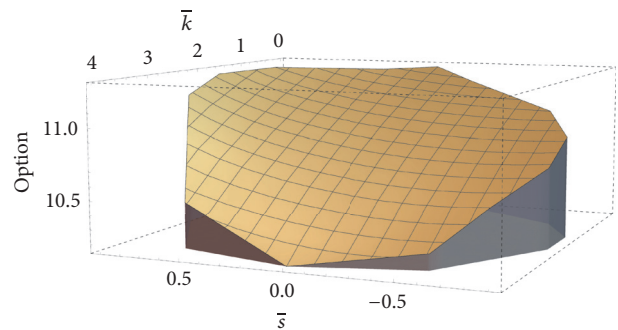


FIGURE 5: Lookback option prices as a function of skewness and excess kurtosis.

6.2. *Lookback Options.* Consider a European lookback call option with fixed strike, with payoff $(\bar{S} - K)_+$, where \bar{S} is the maximum of the stock price between time 0 and time T . We let the time points used to determine the maximum be equally spaced. There is no closed form formula for the distribution of the maximum price nor for the price of such a discretely monitored lookback option, so simulation is used. We generated normal random variables and used the technique described in Section 2.5. The assumptions are maturity $T = 1$, monthly monitoring, annual volatility 0.1685 (where this number comes from is explained in the next subsection). The interest rate is set to 0, initial stock price is \$100, and strike price is also \$100.

Figure 5 shows the option prices for values of skewness and excess kurtosis that can be achieved with the $GC(a, b; 0, 0, c_3, c_4)$ family for monthly rates of return. Roughly speaking, prices decrease when excess kurtosis increases, and also as kurtosis decreases, though the dependence on c_3 and c_4 becomes less linear as \bar{k} increases. When $c_3 = c_4 = 0$ (Gaussian case) the option price is 11.277. The price is a little higher (\$11.286) when $\bar{s} = 0.75$ and $\bar{k} = 1.0$. The lowest price (\$10.163) occurs when $\bar{s} = 0$ and $\bar{k} = 4$. Skewness and kurtosis thus have a significant effect on lookback option prices; the effect will of course vary depending on the parameters. (Observe that all those prices are quite far from the Black-Scholes continuous-monitoring lookback price, \$14.17.)

7. Discussion and Conclusion

This paper tries to provide a framework for applying Gram-Charlier series to option pricing. We have shown that option prices may be significantly affected by varying skewness or kurtosis away from their Gaussian values. The practical consequence is that using the Black-Scholes formula (*i.e.*, zero skewness and excess kurtosis) may distort option prices. Gram-Charlier distributions capture skewness and kurtosis and retain a lot of the tractability of the normal distribution. Estimating parameters remains a challenge for Gram-Charlier distributions involving more than four parameters.

Convergence of Gram-Charlier and other expansions in option pricing and other applied problems is an area for further research. The heuristic derivation of the Gram-Charlier/Edgeworth series of a density $g(\cdot)$ in terms of the normal density and its derivatives (see, e.g., [25], Section 3.2) may give the impression that the inverted series should converge or at least have some asymptotic property; this is not the case in general. Gram-Charlier (or Edgeworth) expansions do not always converge to the true probability distribution, as Section 2.3 indicates; see [26], Section 17, or [25] for more details. In the context of the Central Limit Theorem, the Edgeworth series for the density of the normalized sum of random variables have asymptotic properties, when the number of summands n tends to infinity. Details are given in [25]. Multivariate versions of the Gram-Charlier series exist, for instance, [10, 27, 28], apply such generalizations.

The Gram-Charlier type B series are based on the Poisson distribution, rather than the normal. There are several other expansions that can be used. For instance, series of Laguerre polynomials may be used for densities on the positive half-line (a convergent Laguerre series has some numerical success in pricing Asian options; see [29]). The Jacobi polynomials [30] are orthogonal over the interval $(-1, 1)$, but any other finite interval may be chosen by making the obvious change of variable, and this leads to series expansions for densities with compact support, based on the beta distribution. An application of log-transformed shifted Jacobi polynomials to arbitrary approximation of probability distributions on the positive half-line may be found in [31].

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

The authors thank the Committee for Knowledge Extension and Research (CKER) of the Society of Actuaries for partially funding this project.

References

- [1] P. A. Abken, D. B. Madan, and S. Ramamurtie, "Estimation of risk-neutral and statistical densities by *Hermite polynomial* approximation," Working Paper 96-5, Federal Reserve Bank of Atlanta, 1996.
- [2] D. Backus, S. Foresi, K. Li, and L. Wu, "Accounting for biases in Black-Scholes," Working Paper, Stern School of Business, New York University, New York, NY, USA, 1997.
- [3] J.-P. D. Chateau, "Marking-to-model credit and operational risks of loan commitments: a Basel-2 advanced internal ratings-based approach," *International Review of Financial Analysis*, vol. 18, no. 5, pp. 260–270, 2009.
- [4] J. D. Chateau, "Valuing european put options under skewness and increasing (excess) kurtosis," *Journal of Mathematical Finance*, vol. 4, no. 3, pp. 160–177, 2014.
- [5] C. J. Corrado, "The hidden martingale restriction in Gram-Charlier option prices," *Journal of Futures Markets*, vol. 27, no. 6, pp. 517–534, 2007.
- [6] E. Jondeau and M. Rockinger, "Gram-Charlier densities," *Journal of Economic Dynamics and Control*, vol. 25, no. 10, pp. 1457–1483, 2001.
- [7] E. Jurczenko, B. Maillet, and B. Négrea, "Skewness and kurtosis implied by option prices: a second comment," Discussion Paper of the LSE-FMG 419, 2002.
- [8] E. Jurczenko, B. Maillet, and B. Négrea, "A note on skewness and kurtosis adjusted option pricing models under the Martingale restriction," *Quantitative Finance*, vol. 4, no. 5, pp. 479–488, 2004.
- [9] J. L. Knight and S. E. Satchell, *Return Distributions in Finance*, Butterworth Heineman, 2001.
- [10] E. Schlögl, "Option pricing where the underlying assets follow a Gram/Charlier density of arbitrary order," *Journal of Economic Dynamics & Control*, vol. 37, no. 3, pp. 611–632, 2013.
- [11] A. León, J. Mencía, and E. Sentana, "Parametric properties of semi-nonparametric distributions, with applications to option valuation," *Journal of Business & Economic Statistics*, vol. 27, no. 2, pp. 176–192, 2009.
- [12] M. Hardy, *Investment Guarantees: Modeling and Risk Management for Equity-Linked Life Insurance*, John Wiley & Sons, New York, NY, USA, 2003.
- [13] P. Gaillardetz and X. S. Lin, "Valuation of equity-linked insurance and annuity products with binomial models," *North American Actuarial Journal*, vol. 10, no. 4, pp. 117–144, 2006.
- [14] P. Boyle and W. Tian, "The design of equity-indexed annuities," *Insurance: Mathematics & Economics*, vol. 43, no. 3, pp. 303–315, 2008.
- [15] S. G. Kou, "Discrete barrier and lookback options," in *Handbooks in Operations Research and Management Science*, vol. 15, chapter 8, pp. 343–373, 2007.
- [16] D. E. Barton and K. E. Dennis, "The conditions under which Gram-Charlier and Edgeworth curves are positive definite and unimodal," *Biometrika*, vol. 39, no. 3-4, pp. 425–427, 1952.
- [17] H. Cramér, *On Some Classes of Series Used in Mathematical Statistics*, Skandinaviske Matematiker Congress, Copenhagen, Denmark, 1925.
- [18] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 2, John Wiley & Sons, New York, NY, USA, 1971.
- [19] J. Stoyanov, "Krein condition in probabilistic moment problems," *Bernoulli*, vol. 6, no. 5, pp. 939–949, 2000.
- [20] F. D. Rouah and G. Vainberg, *Option Pricing Models and Volatility Using Excel-VBA*, John Wiley & Sons, New York, NY, USA, 2007.
- [21] L. S. Rompolis and E. Tzavalis, "Retrieving risk neutral densities based on risk neutral moments through a Gram-Charlier series expansion," *Mathematical and Computer Modelling*, vol. 46, no. 1-2, pp. 225–234, 2007.

- [22] E. B. Del Brio and J. Perote, "Gram-Charlier densities: maximum likelihood versus the method of moments," *Insurance: Mathematics and Economics*, vol. 51, no. 3, pp. 531–537, 2012.
- [23] X. S. Lin and K. S. Tan, "Valuation of equity-indexed annuities under stochastic interest rates," *North American Actuarial Journal*, vol. 7, no. 4, pp. 72–91, 2003.
- [24] S. R. Das and R. K. Sundaram, "Of smiles and smirks: a term structure perspective," *Journal of Financial and Quantitative Analysis*, vol. 34, no. 2, pp. 211–239, 1999.
- [25] J. E. Kolossa, *Series Approximation Methods in Statistics*, Springer, New York, NY, USA, 3rd edition, 2006.
- [26] H. Cramér, *Mathematical Methods of Statistics*, Princeton University, Princeton, NJ, USA, 1946.
- [27] E. B. Del Brio, T.-M. Níguez, and J. Perote, "Gram-Charlier densities: a multivariate approach," *Quantitative Finance*, vol. 9, no. 7, pp. 855–868, 2009.
- [28] E. B. Del Brio, T.-M. Níguez, and J. Perote, "Multivariate semi-nonparametric distributions with dynamic conditional correlations," *International Journal of Forecasting*, vol. 27, no. 2, pp. 347–364, 2011.
- [29] D. Dufresne, "Laguerre series for Asian and other options," *Mathematical Finance*, vol. 10, no. 4, pp. 407–428, 2000.
- [30] N. N. Lebedev, *Special Functions and Their Applications*, Dover, New York, NY, USA, 1972.
- [31] D. Dufresne, "Fitting combinations of exponentials to probability distributions," *Applied Stochastic Models in Business and Industry*, vol. 23, no. 1, pp. 23–48, 2007.



Hindawi

Submit your manuscripts at
<https://www.hindawi.com>

