GRAM-CHARLIER PROCESSES AND EQUITY-INDEXED ANNUITIES

JEAN-PIERRE CHATEAU AND DANIEL DUFRESNE

ABSTRACT. A Gram-Charlier distribution has a density that is a polynomial times a normal density. The historical connection between actuarial science and the Gram-Charlier expansions goes back to the 19th century. A critical review of the financial literature on the Gram-Charlier distribution is made. Properties of the Gram-Charlier distributions are derived, including moments, tail estimates, moment indeterminacy of the exponential of a Gram-Charlier distributed variable, non-existence of a continuous-time Lévy process with Gram-Charlier increments, as well as formulas for option prices and their sensitivities. A procedure for simulating Gram-Charlier distributions is given. Multiperiod Gram-Charlier modelling of asset returns is described, apparently for the first time. Formulas for equity indexed annuities' premium option values are given, and a numerical illustration shows the importance of skewness and kurtosis of the risk neutral density.

1. INTRODUCTION

Gram-Charlier series are expansions of the form

$$f(x) \stackrel{?}{=} \phi(a,b;x)\left[1+c_1He_1\left(\frac{x-a}{b}\right)+c_2He_2\left(\frac{x-a}{b}\right)+\cdots\right], \tag{1}$$

where f is a probability density function,

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

is the usual normal density, and He_k is the Hermite polynomial of order k. Expression (1) is an orthogonal polynomial expansion for the ratio $f(x)/\phi(x)$; the expansion may or may not converge to the true value, explaining the question mark above the equal sign. In this paper we focus mostly on 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 $a, b, c_0, \ldots, c_N$, as is explained in detail below.

This paper has three main goals, (1) describe Gram-Charlier distributions from scratch, at the same time deriving new properties of those distributions; (2) apply those to options and equity indexed annuities, focusing on how the prices obtained vary with the skewness and kurtosis of returns; (3) define a Gram-Charlier process and study its basic properties. The formulas we give for 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.

Date: April 2, 2012.

Key words and phrases. Gram-Charlier distributions; Gram-Charlier process; log Gram-Charlier distribution; moment indeterminacy; Krein condition; skewness; kurtosis; equity indexed annuities; option prices; simulation.

Corresponding author, Centre for Actuarial Studies, University of Melbourne, Melbourne VIC 3010, Australia, dufresne@unimelb.edu.au.

Origins of Gram-Charlier series and distributions

The name "Gram-Charlier series" is not historically accurate; although Gram is indeed one of the originators of the series, he is not the first one, and the others include some of the greatest mathematicans, actuaries and statisticians of the 19th century; Charlier played a more minor role in the early 20th century (see Cramér [12]). We give some brief details on the history of Gram-Charlier series here, more can be found in the interesting papers by A. Hald [17],[18]. In 1811, Laplace [26] "derives the complete Gram-Charlier expansion in his discussion of a diffusion problem, using orthogonal polynomials proportional to the Hermite polynomials with argument $x\sqrt{2}$" ([18], p.140). The cumulants are implicitly present in the very early proofs of the central limit theorem, but a good definition of cumulants had to wait until Thiele gave one in 1899, and this was directly related to what later became known as the Gram-Charlier series. In 1855 and 1859, Chebyshev studied the least-squares approximation of discrete data and came up with formulas from which the Gram-Charlier series may be obtained (at the same time defining orthogonal polynomials later attributed to Hermite, denoted $He_n(x)$ in this paper). "It seems that the many-talented Danish philologist, politician, forester, statistician, and actuary L.H.F. Oppermann is the first to propose a system of skew frequency functions obtained by multiplying the normal density by a power series" ([18], p.144). Opperman did not publish this system, but his younger actuarial colleagues Thorvald Nicolai Thiele (1873) and Jargen Pedersen Gram (1879) did. Thiele later gave the modern definition of cumulants and made the connection between cumulants and the Gram-Charlier series clearer. (Among Thiele's many accomplishements there is also, as all actuaries know, a differential equation for actuarial reserves.)

Thiele was well aware of convergence issues with the Gram-Charlier series; Cramér gave the first proof of convergence under specific conditions in the 1920s, see his paper [12]. Hald [18] reflects on why R.A. Fisher disregarded Thiele's contribution to the discovery of the Gram-Charlier series (that are sometimes called "Gram-Charlier series type A").

Gram-Charlier distributions have the same form as a truncated Gram-Charlier series, that is, they have a density of the form

$$\phi(a,b,x)p\left(\frac{x-a}{b}\right),$$

where $p(\cdot)$ is a polynomial. Not all choices of $p(\cdot)$ lead to a proper probability density, the function $\phi(a,b,x)p(x)$ needs to be nonnegative and integrate to one. With respect to the latter condition it is useful to express the polynomial as

$$p(x) \;=\; c_0 + c_1 He_1(x) + \cdots + c_N He_N(x),$$

as is explained in Section 2.

Gram-Charlier distributions in option pricing

It has been observed that option prices have non-constant 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 (1) stochastic volatility models (where the parameter σ in Black-Scholes is replaced with a continuous-time stochastic process),

(2) GARCH time series and (3) Levy 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. The assumption is that the normalized log-price of the underlying security has a Gram-Charlier distribution under the risk-neutral measure. A majority of authors assume a density of the form

$$\phi(x)\left(1+\frac{\bar{s}}{6}He_3(x)+\frac{\bar{k}}{24}He_4(x)\right), \quad \phi(x) = \phi(0,1,x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, \tag{2}$$

for the normalized log-return. (In our notation this is a $\mathbf{GC}(0,0;0,0,\frac{\bar{s}}{6},\frac{\bar{k}}{24})$ distribution, see Section 2.) Here $He_k(x)$ is the Hermite polynomial of order k. The notation emphasizes that in this case the coefficient of $He_3(x)$ turns out to be the skewness coefficient divided by 6, and the coefficient of $He_4(x)$ the excess kurtosis coefficient divided by 24. The distribution of the 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 then allows non-zero skewness and excess kurtosis (unlike the normal distribution found in Black-Scholes). In (2) the parameters $(\bar{s},\bar{k})$ are restricted to a specific region (see Figure 1 at the end of this paper), because outside that region the function in (2) becomes negative for some values of x. This restriction of the four-parameter family of Gram-Charlier distributions was known at least as far back as Barton and Dennis [4] (early 1950s).

A great impetus for the use of series or, more often, truncated series expressions for option prices was Jarrow and Rudd [20]. That paper was however overly optimistic in stating that all Edgeworth series converge, a claim which is false, see Appendix. The paper by Jarrow and Rudd cannot be used as a justification that Edgeworth or Gram-Charlier series for option prices (that is, the term-by-term integrated series for the density times the discounted payoff) converge; in particular, using more terms does not necessarily mean greater accuracy. Under stringent assumptions there are rare cases where the inverted series, even though not convergent, is known to have an asymptotic property, see [11], p.229.

One of the earliest reference to Gram-Charlier distributions in option pricing appeared in 1995; Longstaff [29] does not derive any formula or compute option prices under that assumption. The following year Abken *et al.* [1] used a four-parameter Gram-Charlier density to price options, as did Backus *et al.* [2] in 1997, who moreover give explicit option price formulas.

The paper by Li [28] uses a truncated Edgeworth series (see [12] for more details) that happens to be precisely the four-parameter Gram-Charlier truncated series plus an extra term involving the square of the skewness coefficient. The goal of that paper is to approximate option prices when the risk-neutral density of the log-return is of the form

$$\frac{C_1}{(1+C_2(1+\mathrm{sgn}(x)C_3)x^{C_4})^{C_5}}, \quad x\in\mathbb{R}. \tag{3}$$

This is a type of generalized two-sided Pareto distribution, that had been studied before. By specifing the parameters or taking limits of some of them, this family of distributions (more precisely the set of its limit points) includes the Cauchy, Student t, double-exponential and normal distributions, among others. The problem is that, given a density as in (3) for the log-price, the no-arbitrage relationship

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

cannot hold, because the right hand side is infinite. Li fits the parameters $C_1, \ldots, C_5$ to observed call prices by least squares, not taking into account that call prices are infinite under the chosen distribution for the log-price density (3) . The Gram-Charlier or truncated Edgeworth series obtained are meaningless.

The paper by Knight and Satchell [25] appeared in 2001, though it had been a working paper since 1997. This paper assumes log-returns have a four-parameter Gram-Charlier distribution but sets the pricing of options in a model that includes consumption and utility functions. Option pricing formulas are derived but are not directly comparable to the ones in other papers (because of the different economic model). Knight and Satchell appear to be the first to find formulas for sensitivities to the parameters of the Gram-Charlier distribution. Jondeau and Rockinger [21] give a thorough description of the set of $(\bar{s}, \bar{k})$ that make (2) non-negative for all x and describe techniques for the estimation of the parameters from observed option prices.

In 2005 the article by Jurczenko *al.* [23] retraces a mistake in Corrado and Su [10] (noticed earlier in [6]) and specifies the martingale restriction (following from (4)) that the four-parameter Gram-Charlier density must satisfy in pricing options (previous authors had not taken it into account). The same year, Ki *et al.* [24] propose, as density for the log-price, a mixture of third-order Gram-Charlier functions of the form

$$J_{\xi,\eta}(x) \;=\; p\left[1+\frac{\xi}{6\alpha^6}(x^3-3\alpha^2 x)\right]\frac{e^{-\frac{x^2}{2\alpha^2}}}{\alpha\sqrt{2\pi}} + (1-p)\left[1+\frac{\xi}{6\beta^6}(x^3-3\beta^2 x)\right]\frac{e^{-\frac{x^2}{2\beta^2}}}{\beta\sqrt{2\pi}}$$

with

$$p \;=\; 1-\frac{9}{\eta^2}, \quad \alpha^2 \;=\; 1-\frac{1}{p}\sqrt{p(1-p)(\eta/3-1)}, \quad \beta^2 \;=\; 1-\frac{1}{1-p}\sqrt{p(1-p)(\eta/3-1)},$$

where ξ will be skewness and η will be kurtosis. They claim that "any flexible levels of skewness" and "any positive real kurtosis" can be achieved if the functions $J_{\xi,\kappa}$ are used. There are a couple of problems with this claim. First, skewness and kurtosis cannot be an arbitrary pair of numbers. From Hölder's inequality

$$[\mathbf{E}(|Y|^{q_1})]^{\frac{1}{q_1}} \;\le\; [\mathbf{E}(|Y|^{q_2})]^{\frac{1}{q_2}}, \quad 1 \le q_1 \le q_2,$$

and it follows that, letting $Y = (X-\mu)/\sigma$ in the above,

$$\bar{k} \;=\; \mathbf{E}Y^4 \;\ge [\mathbf{E}(Y^2)]^2 \;=\; 1, \quad \bar{k} \ge (\mathbf{E}|Y|^3)^{\frac{4}{3}} \;\ge\; |\mathbf{E}Y^3|^{\frac{4}{3}} \;=\; |\bar{s}|^{\frac{4}{3}},$$

and thus skewness and kurtosis must satisfy

$$\bar{k} \;\ge\; \max(1, |\bar{s}|^{\frac{4}{3}}).$$

A family of probability density functions that can have "any" skewness and kurtosis therefore does not exist. Second, and this is an even bigger problem here, the functions $J_{\xi,\kappa}$ violate the non-negativity condition for probablity density functions, unless $\xi = 0$. This is because both polynomials $x^3-\alpha^2 x$ and $x^3-\beta^2 x$ tend to $-\infty$ as x tends to $-\infty$, and to $+\infty$ as x tends to $+\infty$. Hence, if $\xi \neq 0$ there is a half-line $(-\infty, x_0)$ or (x_0, ∞) (depending on the sign of ξ) where $J_{\xi,\kappa}$ is negative. By the same reasoning any Gram-Charlier distribution must be of even order see Theorem 1 below). Another way to see that something is amiss with $J_{\xi,\kappa}$ is that its moment generating function is ([24], p.854)

$$\left(1+\frac{\xi s^3}{6}\right)\left[pe^{\frac{\alpha^2 s^2}{2}}+pe^{\frac{\beta^2 s^2}{2}}\right].$$

This expression is negative for some values of s whenever $\xi \neq 0$, while the moment generating function of a true probability density function cannot be negative. We conclude that the results in Ki *et al.* [24] are meaningless.

Corrado [9] suggests ways to apply the martingale condition (4). We return to that condition in Section 5. Chateau [7, 8] applies the four-moment Gram-Charlier distribution to the pricing of European futures put option embeddded in banks' loan commitments. The 2009 article provides the put closed-form solution under the martingale condition and positivity restriction as well as its sensitivity analysis (the greeks).

Our Gram-Charlier distribution with parameters $a, b, c_1, \ldots, c_N$, denoted $\mathbf{GC}(a, b : c_1, \ldots, c_N)$, has density

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

Unlike previous authors, we will not restrict our analysis to $c_1 = c_2 = 0, N = 4$ (more details in Section 2). The question whether (5) is non-negative for all x is an important one. Several authors have disregarded this issue and, in fact, some have come up with parameters that did not yield a true probability density function. One might argue that they were using the first few terms of an infinite series that converges to the true option price, and that the fact that the truncated expression for the density is not a valid density function is unavoidable. Series expressions for option prices are almost always that way, for instance the Laguerre series in Dufresne [14]. 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 talk of Gram-Charlier *distributions*, not *expansions*, and insist that the densities integrate to one and be non-negative. Our goal is to define a family of proper probability distributions, not to use truncated Gram-Charlier expansions to approximate unknown distributions.

The paper by León *et al* [27] 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

$$f_X(x) \;=\; \phi(a,b,x)p(x) \tag{6}$$

where the polynomial $p(x)$ is the square of another polynomial, $p(x) = q(x)^2$. This has the obvious advantage that the non-negativity restriction on $f_X(\cdot)$ is automatically satisfied. We discuss that subclass of "squared" Gram-Charlier distributions at some length in the Conclusion. (León *et al.* call those distributions "semi-non-parametric" (SNP) in [27].)

As far as we know, all previous authors have used a fixed number of parameters, most often four, to model the log-return distribution over the remaining life of any option, *i.e.* they used a single-period model. This has an obvious downside, in that it becomes tricky, if not impossible, to preserve consistency between the prices of options with different maturities. Lévy processes and stochastic volatility models driven by Brownian motion are much better in this respect. Section 3 unfortunately 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. A small disadvantage of such a model is that the order of the Gram-Charlier distribution of the multiperiod return has a larger number of parameters, though the model is still simpler than almost all the alternatives (Lévy and stochastic volatility models driven by Brownian motions). There is no problem computationally, since we give explicit formulas for options under Gram-Charlier distributions with an arbitrary number of parameters.

Layout of the paper

The four-parameter family of Gram-Charlier distributions $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$ has been studied in detail in the literature, as well as the subfamily consisting of polynomials $p(\cdot)$ that are squares of a polynomial $q(\cdot)$ (see Conclusion for more details); but a number of useful results are missing about the general Gram-Charlier distributions (5). 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). In Section 3 we show that there is no Lévy process with Gram-Charlier distributed increments, apart from Brownian motion, but we also 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. Next, Section 5 give formulas for European call and put prices when the log-price returns of the underlying has a general Gram-Charlier distribution; the previous literature only considered the $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$ family and the squared Gram-Charlier distributions. 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. A simple way to simulate Gram-Charlier distributions is described. We also derive formulas for the sentivitivies (greeks) of those option prices with respect to all parameters. Parts (c), (d), (e), (i), (j) (k) and (l) of Theorem 1, part (b) of Theorem 2 and Theorems 4, 5, 7 and 9 appear to be new, while Theorems 2(a), 3, 6 are for the first time formulated for general Gram-Charlier distributions (Theorems 2(a), 3, 6 are given, more or less explicitly, for the subclass of squared Gram-Charlier distributions in [27], and some of the greeks in Theorem 7 had been calculated for the $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$ distribution by previous authors).

The pricing of equity indexed annuities (EIAs in the sequel) has been studied by many authors, including Hardy [19], Gaillardetz and Lin [16], Boyle and Tian [5], Ballotta [3]. In Section 6 e derive formulas for EIA premium options under the assumption that log-returns have a general Gram-Charlier distribution, and then show how the pricing of EIAs is affected by the returns distribution for the index. More precisely, we want to see how prices vary when the skewness and kurtosis of the returns are varied. We focus on compound ratchet EIAs without life-of-contract guarantee.

Notation

Let us summarize some of the notation used in the rest of this paper. The density of a variable X is expressed as in (5), where e $p(x)$ is a polynomial and

$$\phi(a,b,x) \;=\; \frac{1}{b\sqrt{2\pi}} e^{-\frac{1}{2b^2}(x-a)^2}.$$

When $a = 0, b = 1$ this is written as

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

We also define

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

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

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

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

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

2. Gram-Charlier distributions

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

$$f(x) \;=\; \phi(x) \sum_{k=0}^{N} c_k\, He_k(x), \quad x \in \mathbb{R}, \tag{7}$$

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

$$\phi(x) \sum_{k=0}^{N} c_k\, He_k(x). \tag{8}$$

This will be called a Gram-Charlier distribution *with parameters $a, b, \tilde{c}$, with $\tilde{c} = (c_1, \ldots, c_N)$. 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 $\mathbf{GC}(a, b; (0, \ldots 0))$ (or $\mathbf{GC}(a, b; -)$) with order 0.*

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

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

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

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

The condition $c_0=1$ ensures that the function (8) integrates to one, since

$$\int_{-\infty}^{\infty}\phi(x)He_k(x)\,dx \;=\; \int_{-\infty}^{\infty}(-1)^k\frac{d^k}{dx^k}\phi(x)\,dx \;=\; 0, \quad k \;=\; 1,2,\ldots,$$

but there are no simple conditions that ensure that a polynomial remains non-negative 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

$$w_1(t,x) \;=\; \sum_{n=0}^{\infty}\frac{t^n}{n!}He_n(x).$$

From the definition of the Hermite polynomials,

$$\sum_{n=0}^{\infty}\frac{t^n}{n!}e^{-\frac{x^2}{2}}He_n(x) \;=\; \sum_{n=0}^{\infty}\frac{(-t)^n}{n!}\frac{d^n}{dx^n}e^{-\frac{x^2}{2}},$$

and thus

$$w_1(t,x) \;=\; e^{tx-\frac{t^2}{2}}.$$

Another one is

$$\begin{aligned} w_2(t,u,x,y) &= \sum_{k=0}^{\infty}\sum_{n=0}^{\infty}\frac{t^k}{k!}\frac{u^n}{n!}\int_{-\infty}^{\infty}e^{-\frac{x^2}{2}}He_k(x)He_n(x+y)\,dx \\ &= \int_{-\infty}^{\infty}e^{-\frac{x^2}{2}}e^{tx-\frac{t^2}{2}}e^{u(x+y)-\frac{u^2}{2}}\,dx \\ &= e^{-\frac{1}{2}(t^2+u^2)+uy}\int_{-\infty}^{\infty}e^{-\frac{x^2}{2}+x(t+u)}\,dx \\ &= \sqrt{2\pi}\,e^{u(t+y)}. \end{aligned}$$

Letting $y=0$ leads to

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{x^2}{2}}He_k(x)He_n(x)\,dx \;=\; \frac{\partial^k}{t^k}\frac{\partial^n}{u^n}e^{tu}\bigg|_{t=u=0} \;=\; \begin{cases} 0 & \text{if } k\neq n \\ n! & \text{if } k=n. \end{cases}$$

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

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

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

$$\int_{-\infty}^{\infty}e^{tx}\phi(x)He_k(x)\,dx \;=\; t^k e^{\frac{t^2}{2}}, \quad k=0,1,\ldots \tag{9}$$

A consequence is that $He_1(x)$, $He_2(x)$,... integrate to 0 when multiplied by $\phi(x)$:

$$\int_{-\infty}^{\infty} \phi(x) He_k(x)\, dx \;=\; 0, \quad k = 1, 2, \ldots \tag{10}$$

Let us calculate the moments of a distribution with density (7). First consider

$$\int_{-\infty}^{\infty} x^n \phi(x) He_k(x)\, dx \;=\; (-1)^k \int_{-\infty}^{\infty} x^n \left(\frac{d^k}{dx^k} \phi(x) \right) dx. \tag{11}$$

Integrate by parts repeatedly, first assuming $0 \le n < k$:

$$(-1)^{k-n} n! \int_{-\infty}^{\infty} \left(\frac{d^{k-n}}{dx^{k-n}} \phi(x) \right) dx \;=\; 0.$$

If $k = n$ then (11) equals $n!$. If $n > k$ then the result is

$$n(n-1)\cdots(n-k+1) \int_{-\infty}^{\infty} x^{n-k} \phi(x)\, dx.$$

The last integral is the moment of order $n-k$ of the standard normal distribution, which is well known to be 0 if $n-k$ is odd or

$$\frac{(n-k)!}{2^{\frac{n-k}{2}} \left(\frac{n-k}{2} \right)!}$$

if $n-k$ is even. Hence,

$$\int_{-\infty}^{\infty} x^n \phi(x) He_k(x)\, dx \;=\; \begin{cases} \dfrac{n!}{2^{\frac{n-k}{2}} \left(\frac{n-k}{2} \right)!} & \text{if } n-k \text{ is even and non-negative} \\ 0 & \text{if } n-k \text{ is odd and non-negative.} \end{cases}$$

Finally, the n-th moment of the distribution in (7) is

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

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. This is confirmed by the moment-generating function, which may be found from (9):

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

It can be checked that differentiating this expression n times and setting $t = 0$ gives the same expression found for the n-th moment of (7).

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

(a) $\mathbf{E}(Y-a)^n \;=\; b^n \sum_{k=0}^{N \wedge n} c_k \dfrac{n!}{2^{\frac{n-k}{2}} \left(\frac{n-k}{2} \right)!} \mathbf{1}_{\{n-k\,even\}}.$

(b) $\mathbf{E} e^{tY} \;=\; e^{at + \frac{b^2 t^2}{2}} \sum_{k=0}^{N} b^k c_k t^k.$

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

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

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

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

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

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

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

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

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

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

$$\mathbf{E}X^j = \mathbf{E}Y^j, \qquad j = 1, \ldots, K,$$

if, and only if, $c_j^X = c_j^Y$, $j = 1, \ldots, K$.

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

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

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 $N = 0, 2, 4, \ldots$.

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

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

$$g(y) \;=\; \frac{e^{-\frac{y}{2}}}{\sqrt{2\pi y}} \sum_{j=0}^{\frac{N}{2}} \alpha_j y^j, \tag{12}$$

where

$$\sum_{j=0}^{\frac{N}{2}} \alpha_j y^j \;=\; \sum_{j=0}^{\frac{N}{2}} c_{2j} He_{2j}(\sqrt{y}).$$

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 $\tilde{c} = \tilde{0} \in \mathbb{R}^N$ then the distribution is the standard normal, and this is a Gram-Charlier distribution. Since $c_N > 0$ and N is even, the polynomial

$$\sum_{k=1}^{N} He_k(x)$$

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

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

(recall that $c_0 He_0(x) = 1$ for all Gram-Charlier distributions). The set of valid vectors $\tilde{c}$ thus includes $(\epsilon,\ldots,\epsilon) \in \mathbb{R}^N$ 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

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

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

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

$$\psi(t) = e^{\frac{t^2}{2}}, \quad q_X(t) \;=\; \sum_{k=0}^{N} c_k^X\, t^k, \quad q_Y(t) \;=\; \sum_{k=0}^{N} c_k^Y\, t^k$$

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,

$$\begin{aligned}
\psi'(0)q_X(0) + \psi(0)q'_X(0) \;&=\; \psi'(0)q_Y(0) + \psi(0)q'_Y(0) \\
\vdots \qquad\qquad & \vdots \qquad\qquad \vdots \\
\sum_{j=0}^{K} \binom{K}{j} \psi^{(j)}(0) q_X^{(K-j)}(0) \;&=\; \sum_{j=0}^{K} \binom{K}{j} \psi^{(j)}(0) q_Y^{(K-j)}(0).
\end{aligned}$$

Suppose that $c_j^X = c_j^Y$, $j = 1,\ldots,K$. Then $q_X^{(j)}(0) = q_Y^{(j)}(0)$, $j = 1,\ldots,K$, and thus the first K moments of X and Y are the same. Conversely, suppose $\mathbf{E}X^j \;=\; \mathbf{E}Y^j$, $j =$

$1, \ldots, K$. 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

$$a = \mathbf{E}X \iff c_1^X = 0$$

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

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

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 $X \sim \mathbf{GC}(a, b; \tilde{c}^X)$, a, b arbitrary, this means that

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

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

$$e^{aqs+\frac{1}{2}b^2q^2s^2} \sum_{k=0}^{N} c_k b^k q^k s^k = e^{aqs+\frac{1}{2}b^2q^2s^2} \sum_{k=0}^{N} c_k \left(\frac{q}{|q|}\right)^k (b|q|)^k s^k.$$

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

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

The Hermite polynomials of odd order are odd functions, and so disappear from that expression, while the even order Hermite polynomials are even functions. The chi-square distribution with k degrees of freedom has density

$$\frac{2^{-\frac{k}{2}}}{\Gamma(\frac{k}{2})} y^{\frac{k}{2}-1} e^{-\frac{y}{2}}$$

□

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. Part (l) of the theorem above yields a curious result, namely an instance of the fact the "square root" of a distribution is often not unique. Since the parameters c_k of odd order k vanish, we see that the same density (12) is the density of the square of any variable having one of the possibly infinite number of Gram-Charlier densities that have parameters $c_2, c_4, \ldots, c_N$ (since the parameters $c_1, c_3, \ldots, c_{N-1}$ have no effect on the density of the square). For instance, suppose $X \sim \mathbf{GC}(0, 1; 0, 0, c_3, c_4)$; then, for any value of c_3 the distribution of X^2 is

$$g(y) = \frac{e^{-\frac{y}{2}}}{\sqrt{2\pi y}}[1 + c_4 He_4(\sqrt{y})] = \frac{e^{-\frac{y}{2}}}{\sqrt{2\pi y}}(1 + 3c_4 - 6c_4 y + c_4 y^2). \tag{13}$$

This is a combination of the χ_1^2, χ_3^2, χ_5^2 densities, with weights $1 + 3c_4$, $-6c_4$, $3c_4$, respectively. For any $0 < c_4 < \frac{1}{6}$ there is an infinite number of c_3's such that (13) is the density of the square of $X \sim \mathbf{GC}(0, 1; 0, 0, c_3, c_4)$.

Theorem 2. *(a) If* $X \sim \mathbf{GC}(0,1;c_1,\ldots,c_N)$ *then*

$$\begin{aligned}\mathsf{P}(X \le x) &= \Phi(x) - \phi(x)\sum_{k=1}^{N} c_k\, He_{k-1}(x)\\ \mathsf{P}(X > x) &= \Phi(-x) + \phi(x)\sum_{k=1}^{N} c_k\, He_{k-1}(x).\end{aligned}$$

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

$$\begin{aligned}\mathsf{P}(X > x) &\sim c_N\left(\frac{x-a}{b}\right)^{N-1}\phi\left(\frac{x-a}{b}\right) \quad \text{as } x \to \infty\\ \mathsf{P}(X < x) &\sim c_N\left(\frac{|x-a|}{b}\right)^{N-1}\phi\left(\frac{x-a}{b}\right) \quad \text{as } x \to -\infty.\end{aligned}$$

Proof. (a) This is a direct calculation:

$$\begin{aligned}\int_{-\infty}^{x}\left[\phi(y) + \phi(y)\sum_{k=1}^{N} c_k He_k(y)\right]dy &= \Phi(x) + \sum_{k=1}^{N} c_k \int_{-\infty}^{x}\phi(y)He_k(y)\,dy\\ &= \Phi(x) - \sum_{k=1}^{N} c_k \int_{-\infty}^{x} d\left[\phi(y)He_{k-1}(y)\right]\\ &= \Phi(x) - \phi(x)\sum_{k=1}^{N} c_k\, He_{k-1}(x).\end{aligned}$$

(b) For simplicity suppose $a = 0, b = 1$ and use L'Hospital's rule in

$$\frac{\mathsf{P}(X > x)}{x^{N-1}\phi(x)} = \frac{1}{x^{N-1}\phi(x)}\int_{x}^{\infty}\phi(y)\sum_{k=0}^{N} c_k\, He_k(y)\,dy.$$

We get

$$\lim_{x\to\infty}\frac{-\phi(x)\sum_{k=0}^{N} c_k\, He_k(x)}{-x^N\phi(x) + (N-1)x^{N-2}\phi(x)} = \lim_{x\to\infty}\frac{x^{-N}\sum_{k=0}^{N} c_k\, He_k(x)}{1 - \frac{N-1}{x^2}} = c_N,$$

since the leading term in $He_k(x)$ is x^k. The left-hand tail is handled similarly, or more simply, use the distribution of $-X$ in Theorem 1(k). □

Part (b) includes the well-known tail of the normal distribution as a particular case, and shows that, even though Gram-Charlier distributions may have non-zero skewness, in general the right and left tails are the same, *i.e.* $c_N x^{N-1}$ times the normal density. This asymptotic symmetry is not the best one could hope for when modelling skewed distributions such as asset log-returns. It may be seen as a good thing that the Gram-Charlier tails are thicker than the normal: for any $N > 0$, the ratio of the GC tail to the normal tail (the latter is the case $N = 0, c_N = 1$) tends to $+\infty$. However, this does not say much, as all the Gram-Charlier tails are still "thin" (because they are 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 [4], but a more detailed explanation is given in Jondeau and Rockinger [21]. The region is shown in Figure 1 (use the correspondance $\bar{s} = 6c_3$, $\bar{k} = 24c_4$ from Theorem 1(h)).

An important fact about this region is that it is not square; the possible excess kurtosis values depend on skewness, and conversely.

2.2. **The $\mathbf{GC}(a, b; c_1, c_2, c_3, c_4)$ family.** This 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 not been used in financial applications (the only case is León [27], who use the subfamily made up of polynomials $p(\cdot)$ that are squares of some second-degree polynomial). 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 by maximum likelihood, performing the optimization under the constraints

$$b \ > \ 0; \quad 1 + \sum_{k=1}^{4} c_k He_k(x) \ \geq \ 0 \quad \forall\, x \in \mathbb{R}.$$

An example is given in Section 7. Naturally there is a larger set of possible values of $\bar{s}$ and $\bar{k}$ for the six-parameter casethan for the four-parameter case.

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

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

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

$$P' \ = \ \frac{e^{qX}}{\mathbf{E}^{\mathsf{P}} e^{qX}}.P, \tag{14}$$

then $X \overset{\mathsf{P}'}{\sim} \mathbf{N}(\mu + \sigma^2 q, \sigma^2)$. The next result is an extension of the Cameron-Martin formula for Gram-Charlier distributions.

Theorem 3. *Suppose $X \overset{\mathsf{P}}{\sim} \mathbf{GC}(a, b; c_1, \ldots, c_N)$ and that P' is defined by* (14) *for $q \in \mathbb{R}$. Then $X \overset{\mathsf{P}'}{\sim} \mathbf{GC}(a + b^2 q, b; c'_1, \ldots, c'_N)$, where $c'_1, \ldots, c'_N$ are found from*

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

or, more precisely,

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

Proof. It is sufficient to calculate the mgf of X under P':

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

Example. If $X \overset{\mathsf{P}}{\sim} \mathbf{GC}(a, b; 0, 0, c_3, c_4)$ and $q = 1$ the change of measure (14) leads to

$$\begin{aligned}
c_1' &= \frac{3b^2c_3 + 4b^3c_4}{1 + b^3c_3 + b^4c_4}, \quad & c_2' &= \frac{3bc_3 + 6b^2c_4}{1 + b^3c_3 + b^4c_4} \\
c_3' &= \frac{c_3 + 4bc_4}{1 + b^3c_3 + b^4c_4}, \quad & c_4' &= \frac{c_4}{1 + b^3c_3 + b^4c_4}.
\end{aligned}$$

Hence, assuming $c_1 = c_2 = 0$ does not imply that c_1' and $c_2' = 0$ 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). This is one more reason to derive formulas for the general Gram-Charlier distributions.

Simulating Gram-Charlier distributions

Although we will not use simulation in our option pricing example below (because there is an explicit formula for the type of EIAs we look at), we present a simple technique for simulating any Gram-Charlier distribution, without having to invert its distribution function. When estimating $\mu = \mathbf{E} f(X_1, \ldots, X_s)$ by simulation, one generates n independent vectors

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

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

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

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

$$\begin{aligned}
\mathbf{E} f(X_1, \ldots, X_s) &= \int_{\mathbb{R}^s} f(x_1, \ldots, x_s) \prod_{k=1}^{s} [\phi(a, b, x_k) p(x_k)] \, dx_1 \ldots dx_s \\
&= \mathbf{E}^{\mathsf{Q}} \left[f(X_1, \ldots, X_s) \prod_{k=1}^{s} p(X_k) \right],
\end{aligned}$$

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

$$\mu_n = \frac{1}{n} \sum_{j=1}^{n} \left[f(X_1^{(j)}, \ldots, X_s^{(j)}) \prod_{k=1}^{s} p(X_k^{(j)}) \right].$$

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 mgfs. Suppose $X_j \sim \mathbf{GC}(a_j, b_j; c_1^{(j)}, \ldots, c_N^{(j)})$, $j = 1, 2$ are independent. Then

$$\mathbf{E}e^{z(X+Y)} = e^{(a_1+a_2)t+[(b^{(1)})^2+(b^{(2)})^2]\frac{t^2}{2}} \left(\sum_{k=0}^{N^{(1)}} c_k^{(1)} t^k \right) \left(\sum_{k=0}^{N^{(2)}} c_k^{(2)} t^k \right).$$

Expanding the product, this says that

$$X + Y \sim \mathbf{GC}(a_1 + a_2, \sqrt{(b^{(1)})^2 + (b^{(2)})^2}; c_1, \ldots, c_{N^{(1)}+N^{(2)}}), \qquad c_k = \sum_{j=0}^{k} c_j^{(1)} c_{k-j}^{(2)}. \quad (15)$$

The downside of this expression is that the number of parameters of the sum $X + Y$ is the sum of the parameters of X and Y. Nevertheless, it is possible to know the explicit distribution of

$$Z_n = X_1 + \cdots + X_n,$$

where the X's are independent and have a Gram-Charlier distribution, constituting a discrete-time Gram-Charlier process. If the X_K's have the same $\mathbf{GC}(a, b; c_1, \ldots, c_N)$ distribution then $\{Z_n, \, n \geq 0\}$ is a random walk. The derivation of the distribution of Z_n can be done recursively, using (15), or it can be done by finding the Taylor expansion of

$$\left(\sum_{k=0}^{N} c_k t^k \right)^n = \sum_{k=0}^{nN} c_k^{(n)} t^k = 1 + c_1 n t + (c_2 n + \tfrac{1}{2} c_1^2 n(n-1)) t^2 + \cdots + c_N^n t^{nN},$$

and thus

$$Z_n \sim \mathbf{GC}(an, b\sqrt{n}; c_1 n, c_2 n + \tfrac{1}{2} c_1^2 n(n-1), \ldots, c_N^n).$$

These computations are simple using symbolic mathematics software. An example is given at the end of Section 7.

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 4. *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 characteristic function is of the form

$$\mathbf{E}e^{itX} = e^{at+b^2\frac{t^2}{2}} \pi(s),$$

where $\pi(\cdot)$ is a polynomial. It is known (*e.g.* Sato [32], p.32) that the characteristic function of an infinitely divisible distribution has no zeros. This only leaves the possibility that $\pi(\cdot)$ is a constant, and this corresponds to the normal distribution. □

This result says that, as far as random walks are concerned, Gram-Charlier processes necessarily 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 $Y \sim \mathbf{GC}(a, b; \tilde{c})$, then $L = \exp(Y) \sim \mathbf{LogGC}(a, b; \tilde{c})$. The density of L is

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

This distribution has all moments finite, and they are given by Theorem 1(b). The log Gram-Charlier distribution shares one property with the lognormal, it is "moment indeterminate".

Theorem 5. *The log Gram-Charlier distribution is not determined by its moments. More precisely, there is a non-countable 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 [15], p.227); the trick works for arbitrary parameters but, to simplify the following, let $L \sim \mathbf{LogN}(0, 1)$, that is,

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

All the functions

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

are non-negative, 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 $L \sim \mathbf{LogGC}(0, 1; \tilde{c})$ (once again the same arguments work for other values of a and b), and consider

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

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_a is not necessarily non-negative for $-1 \leq \alpha \leq 1$. Two cases may arise. The first one is that the polynomial

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

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

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

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

$$s(y+2k\pi) \;=\; s(y), \quad s(-y) \;=\; -s(y).$$

(These two conditions are sufficient for

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

to hold for $k = 1, 2, \ldots$) 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 [33], p.941), which says that a continuous distribution on $\mathbb{R}_+$ with positive density $f(\cdot)$ is not determined by its moments if

$$-\int_0^{\infty} \frac{\log f(z^2)}{1+z^2}\,dz \;<\; \infty. \qquad \square$$

5. Option pricing

The formulas below hold for any vector $(c_1, \ldots, c_N)$ and thus extend those that have previously 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 [27]). Consider a market with a risky security S and a risk-free security with annual return r. Suppose the log return for period $[0, T]$ is denoted X_T and has a Gram-Charlier distribution under the risk-neutral measure (which we denote Q). The risky security has price S_0 at time 0, and at time T

$$S_T \;=\; S_0 e^{X_T}.$$

The market may have one or more periods.

Theorem 6. *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 Q the log-return of the risky security over $[0, T]$ is $X_T \overset{\mathrm{Q}}{\sim} \mathbf{GC}(a, b; c_1, \ldots, c_N)$, which satisfies the martingale condition*

$$e^{a+\frac{b^2}{2}} \sum_{k=0}^{N} b^k c_k \;=\; e^{(r-\delta)T}. \tag{16}$$

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

$$C_0 \;=\; S_0 e^{-\delta T}\Phi(d_1) - Ke^{-rT}\Phi(d_2) + Ke^{-rT}\phi(d_2)\sum_{k=1}^{N}[c_k^*\, He_{k-1}(-d_1) - c_k\, He_{k-1}(-d_2)], \tag{17}$$

where d_1, d_2 and $\{c_k^*\}$ are given by (18),(20),(22) *below. The price of the corresponding European put is*

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

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

$$C_0 \;=\; S_0 e^{-\delta T}\Phi(d_1) - Ke^{-rT}\Phi(d_2) + bKe^{-rT}\phi(d_2)[c_2 + (b-d_2)c_3 + (b^2 - bd_2 + d_2^2 - 1)c_4]$$
$$P_0 \;=\; Ke^{-rT}\Phi(-d_2) - S_0 e^{-\delta T}\Phi(-d_1) - bKe^{-rT}\phi(d_2)[c_2 + (b-d_2)c_3 + (b^2 - bd_2 + d_2^2 - 1)c_4].$$

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

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

This justifies (16). The price of the call is

$$\begin{aligned} C_0 &= e^{-rT}\mathbf{E}^{\mathsf{Q}}(S_T - K)_+ \;=\; C_0^+ - C_0^- \\ C_0^+ &= e^{-rT}\mathbf{E}^{\mathsf{Q}} S_T(\mathbf{1}_{\{S_T>K\}}), \quad C_0^- \;=\; Ke^{-rT}\mathbf{E}^{\mathsf{Q}}(\mathbf{1}_{\{S_T>K\}}). \end{aligned}$$

As usual the second part is easier to deal with:

$$C_0^- \;=\; Ke^{-rT}\mathsf{Q}(X_T > \log(K/S_0)).$$

The probability of the event $X_T > \log(K/S_0)$ can be calculated explicitly by recalling that $(X_T - a)/b \overset{\mathsf{Q}}{\sim} \mathbf{GC}(0,1;c_1,\ldots,c_N)$ and using Theorem 2:

$$\begin{aligned} \mathsf{Q}(X_T > \log(K/S_0)) &= \mathsf{Q}\left(\frac{X_T - a}{b} > \frac{1}{b}\left(\log\left(\frac{K}{S_0}\right) - a\right)\right) \\ &= \Phi(d_2) + \phi(d_2)\sum_{k=1}^{N} c_k\, He_{k-1}(-d_2), \end{aligned}$$

where

$$d_2 \;=\; \frac{1}{b}\left(\log\left(\frac{S_0}{K}\right) + a\right). \tag{18}$$

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

$$\mathsf{Q}' \;=\; \frac{S_T}{\mathbf{E}^{\mathsf{Q}} S_T}.\mathsf{Q} \;=\; \frac{e^{X_T}}{\mathbf{E}^{\mathsf{Q}} e^{X_T}}.\mathsf{Q}.$$

Then

$$\mathbf{E}^{\mathsf{Q}}(S_T\mathbf{1}_{\{S_T>K\}}) \;=\; (\mathbf{E}^{\mathsf{Q}} S_T)\mathbf{E}^{\mathsf{Q}'}(\mathbf{1}_{\{S_T>K\}}).$$

Since $X_T \overset{\mathsf{Q}'}{\sim} \mathbf{GC}(a+b^2, b; c_1', \ldots, c_N')$,

$$\begin{aligned} (\mathbf{E}^{\mathsf{Q}} S_T)\mathbf{E}^{\mathsf{Q}'}(\mathbf{1}_{\{S_T>K\}}) &= S_0 e^{(r-\delta)T}\mathsf{Q}'\left(\frac{X_T - a - b^2}{b} > \frac{1}{b}\left(\log\left(\frac{K}{S_0}\right) - a - b^2\right)\right) \\ &= S_0 e^{(r-\delta)T}\Phi(d_1) + S_0 e^{(r-\delta)T}\phi(d_1)\sum_{k=1}^{N} c_k'\, He_{k-1}(-d_1), \end{aligned} \tag{19}$$

where

$$c_k' \;=\; \frac{1}{\sum_{j=0}^{N} b^j c_j}\sum_{\ell=k}^{N}\binom{\ell}{k} b^{\ell-k} c_\ell, \quad k = 1,\ldots,N$$

$$d_1 \;=\; \frac{1}{b}\left[\log\left(\frac{S_0}{K}\right) + a + b^2\right] \;=\; d_2 + b. \tag{20}$$

Hence,

$$\begin{aligned} C_0 \;=\; S_0 e^{-\delta T}\Phi(d_1) - Ke^{-rT}\Phi(d_2) + S_0 e^{-\delta T}\phi(d_1)\sum_{k=1}^{N} c_k'\, He_{k-1}(-d_1) \\ - Ke^{-rT}\phi(d_2)\sum_{k=1}^{N} c_k\, He_{k-1}(-d_2). \end{aligned}$$

Using

$$\frac{S_0 e^{-\delta T}\phi(d_1)}{\sum_{k=0}^{N} b^k c_k} \;=\; S_0 e^{a+\frac{b^2}{2}-rT}\phi(d_1) \;=\; K e^{-rT}\phi(d_2) \tag{21}$$

(a consequence of (16) and (20)), it is possible to write

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

where

$$c^*_k \;=\; c'_k \sum_{k=0}^{N} b^k c_k \;=\; \sum_{\ell=k}^{N} \binom{\ell}{k} b^{\ell-k} c_\ell, \quad k = 1, \ldots, N. \tag{22}$$

This proves (17).

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

$$P_0 \;=\; C_0 - S_0 + K e^{-rT}.$$

If $N = 4$ then the summation in (17) becomes

$$\begin{aligned}
&\sum_{k=1}^{4} c^*_k\, He_{k-1}(-d_1) - \sum_{k=1}^{4} c_k\, He_{k-1}(-d_2) \\
&= \; c_1 + 2bc_2 + 3b^2c_3 + 4b^3c_4 + (c_2 + 3bc_3 + 6b^2c_4)(-d_1) + (c_3 + 4bc_4)(d_1^2 - 1) \\
&\qquad + c_4(-d_1^3 + 3d_1) - c_1 - c_2(-d_2) - c_3(d_2^2 - 1) - c_4(-d_2^3 + 3d_2) \\
&= \; (2b - d_1 + d_2)c_2 + (3b^2 - 3bd_1 + d_1^2 - d_2^2)c_3 \\
&\qquad + [4b^3 - 6b^2d_1 + 4b(d_1^2 - 1) - d_1^3 + 3d_1 + d_2^3 - 3d_2]c_4 \\
&= \; bc_2 + (b^2 - bd_2)c_3 + (b^3 - b^2d_2 + bd_2^2 - b)c_4. \qquad \square
\end{aligned}$$

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-Sholes formula. More precisely, in the Black-Scholes model (where $c_k = 0$ for all $k \geq 1$) we have $a = \mathbf{E}^{\mathrm{Q}} X_T$ and $b^2 = \mathbf{Var}^{\mathrm{Q}} X_T$, but this does not happen otherwise, first because of the martingale condition (16), second because of the result in part (j) of Theorem 1.

Our option price formulas include the one in Section 3.3 of [23] as a particular case. This is checked by setting

$$c_1 \;=\; c_2 \;=\; 0, \quad b \;=\; \sigma\sqrt{T}, \quad c_3 \;=\; \frac{\gamma_3}{3!}, \quad c_4 \;=\; \frac{\gamma_4}{4!}, \quad \omega \;=\; c_3b^3 + c_4b^4, \quad \delta \;=\; 0$$

and also using (16).

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.* [22], Rouah and Vainberg [30] and Chateau [8]. Below we give sensitivites of the option prices calculated above with respect to all the parameters for a general Gram-Charlier distribution, taking the martingale condition (16) 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 7. *Let C_0 be the price of the European call option described in Theorem 6. Then:*

$$\begin{aligned}
\Delta &= \frac{\partial C_0}{\partial S_0} = e^{-\delta T}\left(\Phi(d_1)+\phi(d_1)\sum_{k=1}^{N} c_k'\, He_{k-1}(-d_1)\right)\\
\gamma &= \frac{\partial^2 C_0}{\partial S_0^2} = \frac{e^{a+\frac{b^2}{2}-rT}}{bS_0}\phi(d_1)\sum_{k=0}^{N} c_k He_k(-d_2)\\
\rho &= \frac{\partial C_0}{\partial r} = KTe^{-rT}\left(\Phi(d_2)+\phi(d_2)\sum_{k=1}^{N} c_k He_{k-1}(-d_2)\right)\\
\kappa &= \frac{\partial C_0}{\partial b} = S_0 e^{-\delta T}\phi(d_1)\sum_{k=0}^{N}[c_k' - c_1' c_{k+1}' + (k+2)c_{k+2}']He_k(-d_1)\\
\frac{\partial C_0}{\partial c_j} &= Ke^{-rT}\phi(d_2)\left(\sum_{k=1}^{j-1} b^k He_{j-1-k}(-d_2) - b^j\sum_{k=1}^{N} c_k' He_{k-1}(-d_1)\right), \quad j\geq 1.
\end{aligned}$$

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 8. *For any integrable random variable U and any constant K,*

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

If U has a continuous density f_U, then

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

From $S_T = S_0 e^{X_T}$ and (19),

$$\begin{aligned}
\Delta &= \frac{\partial}{\partial S_0}[e^{-rT}\mathbf{E}^{\mathrm{Q}}(S_0e^{X_T}-K)_+] = e^{-rT}\mathbf{E}^{\mathrm{Q}}(e^{X_T}\mathbf{1}_{\{e^{X_T}>K/S_0\}}) = \frac{1}{S_0}e^{-rT}\mathbf{E}^{\mathrm{Q}}(S_T\mathbf{1}_{\{S_T>K\}})\\
&= \frac{1}{S_0}C_0^+ = e^{-\delta T}\left(\Phi(d_1)+\phi(d_1)\sum_{k=1}^{N} c_k'\, He_{k-1}(-d_1)\right).
\end{aligned}$$

This reduces to $e^{-\delta T}\Phi(d_1)$ when $c_k=0$, $k\geq 1$, as in the Black-Scholes model. From the lemma,

$$\gamma = \frac{\partial^2}{\partial^2 S_0}C_0 = e^{-rT}\frac{\partial^2}{\partial^2 S_0}\mathbf{E}^{\mathrm{Q}}(S_0U-K)_+ = \frac{K^2e^{-rT}}{S_0^3}f_U\Big(\frac{K}{S_0}\Big), \quad U = e^X,$$

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

$$f_U(u) = \frac{1}{bu}f_G\Big(\frac{1}{b}(\log u - a)\Big) = \frac{1}{bu}\phi(y)\sum_{k=0}^{N} c_k He_k(y)\Bigg|_{y=\frac{1}{b}(\log u - a)}.$$

Setting $u=K/S_0$, we obtain

$$\gamma = \frac{Ke^{-rT}}{bS_0^2}\phi(d_2)\sum_{k=0}^{N} c_k He_k(-d_2).$$

From (21) this is the same as

$$\gamma \;=\; \frac{e^{a+\frac{b^2}{2}-rT}}{bS_0}\phi(d_1)\sum_{k=0}^{N} c_k He_k(-d_2).$$

(In the Black-Scholes model the martingale condition (16) reduces to

$$e^{a+\frac{b^2}{2}} \;=\; e^{(r-\delta)T},$$

so the expression given above for γ agrees with the Black-Scholes' gamma, namely

$$\frac{e^{-\delta T}}{\sigma S_0\sqrt{T}}\phi(d_1).)$$

Next,

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

where $G \overset{\mathrm{Q}}{\sim} \mathbf{GC}(0,1,\tilde{c})$. For fixed G, the function $g(r) = h(G,r)$ is absolutely continuous and is thus expressible as the integral of its derivative:

$$\mathbf{E}^{\mathrm{Q}} h(G,r) \;=\; \mathbf{E}^{\mathrm{Q}} \int_0^r \frac{\partial}{\partial y} h(G,y)\,dy.$$

It is readily checked that the partial derivative inside the integral is integrable with respect to $d\mathrm{Q}\times dy$ over $\Omega\times[0,r]$, and thus the order of expectation and integration may be reversed. Thus,

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

Now, from (16),

$$e^{a+bG-rT} \;=\; \frac{e^{bG-\frac{b^2}{2}-\delta T}}{\sum_{k=0}^{N} b^k c_k} \tag{23}$$

does not depend on r, so

$$\rho \;=\; \mathbf{E}^{\mathrm{Q}}(KTe^{-rT}\mathbf{1}_{\{S_0 e^{a+bG-rT} > Ke^{-rT}\}}) \;=\; KTe^{-rT}\left[\Phi(d_2) + \phi(d_2)\sum_{k=1}^{N} c_k He_{k-1}(-d_2)\right].$$

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

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

Define

$$m(b) \;=\; \mathbf{E} e^{bG-\frac{b^2}{2}} \;=\; \sum_{k=0} b^k c_k, \quad m'(b) \;=\; \sum_{k=1} k b^{k-1} c_k \;=\; c_1^*.$$

In order to shorten the formulas we will write (from (16))

$$e^a \;=\; \frac{e^{(r-\delta)T-\frac{b^2}{2}}}{m(b)},$$

and thus find an expression for

$$\begin{aligned}
&S_0 e^{-\delta T}\mathbf{E}^{\mathrm{Q}}\left[\left(\frac{\partial}{\partial b}\frac{e^{bG-\frac{b^2}{2}}}{m(b)}\right)\mathbf{1}_{\{G>-d_2\}}\right] \\
&\qquad = \frac{S_0 e^{-\delta T-\frac{b^2}{2}}}{m(b)^2}\mathbf{E}^{\mathrm{Q}}\left[\Big(Ge^{bG}m(b) - e^{bG}(m'(b)+bm(b))\Big)\mathbf{1}_{\{G>-d_2\}}\right].
\end{aligned}$$

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

$$\mathbf{E}^{\mathsf{Q}}(e^{bG}\mathbf{1}_{\{G>-y\}}) \;=\; e^{\frac{b^2}{2}}\left[m(b)\Phi(y+b)+\phi(y+b)\sum_{k=1}^{N}c_k^* He_{k-1}(-y-b)\right].$$

To calculate

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

we need (see (22))

$$\frac{\partial c_k^*}{\partial b} \;=\; \frac{\partial}{\partial b}\sum_{\ell=k}^{N}\binom{\ell}{k}b^{\ell-k}c_\ell \;=\; (k+1)c_{k+1}^*, \quad k=1,\dots,N-1, \quad \frac{\partial c_N^*}{\partial b} \;=\; 0.$$

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

$$\begin{aligned}
\mathbf{E}^{\mathsf{Q}}(Ge^{bG}\mathbf{1}_{\{G>-y\}}) \;&=\; b\mathbf{E}^{\mathsf{Q}}(e^{bG}\mathbf{1}_{\{G>-y\}}) + e^{\frac{b^2}{2}}\Big[m'(b)\Phi(y+b)+m(b)\phi(y+b)\\
&\quad +\phi(y+b)\sum_{k=1}^{N-1}(k+1)c_{k+1}^*He_{k-1}(-y-b)+\phi(y+b)\sum_{k=1}^{N}c_k^*He_k(-y-b)\Big]\\
&=\; b\mathbf{E}^{\mathsf{Q}}(e^{bG}\mathbf{1}_{\{G>-y\}}) + e^{\frac{b^2}{2}}m(b)\Big[c_1'\Phi(y+b)\\
&\quad +\phi(y+b)\sum_{k=0}^{N-2}(k+2)c_{k+2}'He_k(-y-b)+\phi(y+b)\sum_{k=0}^{N}c_k'He_k(-y-b)\Big].
\end{aligned}$$

Setting $y=d_2$ and putting the pieces together yields the result. When $\tilde{c}=0$ this reduces to $S_0e^{-\delta T}\phi(d_1)$, 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 $\frac{\partial a}{\partial c_j}$; from (16),

$$\frac{\partial a}{\partial c_j}e^{a+\frac{b^2}{2}}\sum_{k=0}^{N}b^k c_k + b^j e^{a+\frac{b^2}{2}} \;=\; 0 \quad \text{or} \quad \frac{\partial a}{\partial c_j} \;=\; -\frac{b^j}{\sum_{k=0}^{N}b^k c_k} \;=\; -b^j e^{a+\frac{b^2}{2}+(\delta-r)T}.$$

Next,

$$\begin{aligned}
&e^{-rT}\frac{\partial}{\partial c_j}\mathbf{E}^{\mathsf{Q}}(S_0e^{a+bG}-K)_+\\
&\quad=\; e^{-rT}\frac{\partial}{\partial c_j}\int_{-d_2}^{\infty}(S_0e^{a+by}-K)\phi(y)\sum_{k=0}^{N}c_kHe_k(y)\,dy\\
&\quad=\; e^{-rT}\int_{-d_2}^{\infty}S_0\frac{\partial a}{\partial c_j}e^{a+by}\phi(y)\sum_{k=0}^{N}c_kHe_k(y)\,dy + e^{-rT}\int_{-d_2}^{\infty}(S_0e^{a+by}-K)\phi(y)He_j(y)\,dy.
\end{aligned} \tag{24}$$

The first integral reduces to

$$S_0\Delta\frac{\partial a}{\partial c_j}$$

(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,

$$\begin{aligned}\int_y^\infty g(x)\phi(x)He_j(x)\,dx &= (-1)^j\int_y^\infty g(x)\,d\phi^{(j-1)}(x)\\ &= (-1)^{j-1}g(y)\phi^{(j-1)}(y)+(-1)^{j-1}\int_y^\infty g'(x)\,d\phi^{(j-2)}(x)\\ &= (\ldots)\\ &= \phi(y)\sum_{k=0}^{j-1}g^{(k)}(y)He_{j-1-k}(y)+\int_y^\infty g^{(j)}(x)\phi(x)\,dx.\end{aligned}$$

Hence the second integral in (24) becomes

$$\begin{aligned}&S_0e^{a-bd_2-rT}\phi(d_2)\sum_{k=1}^{j-1}b^kHe_{j-1-k}(-d_2)+S_0e^{a-rT}\int_{-d_2}^\infty b^je^{by}\phi(y)\,dy\\ &= S_0e^{a-bd_2-rT}\phi(d_2)\sum_{k=1}^{j-1}b^kHe_{j-1-k}(-d_2)+b^jS_0e^{a-rT+\frac{b^2}{2}}\Phi(d_1).\end{aligned}$$

Putting all this together (using (21)) we find

$$\begin{aligned}\frac{\partial C_0}{\partial c_j} &= S_0e^{a-bd_2-rT}\phi(d_2)\sum_{k=1}^{j-1}b^kHe_{j-1-k}(-d_2)-b^jS_0e^{a+\frac{b^2}{2}-rT}\phi(d_1)\sum_{k=1}^{N}c'_kHe_{k-1}(-d_1)\\ &= S_0e^{a+\frac{b^2}{2}-rT}\phi(d_1)\left(\sum_{k=1}^{j-1}b^kHe_{j-1-k}(-d_2)-b^j\sum_{k=1}^{N}c'_kHe_{k-1}(-d_1)\right)\\ &= Ke^{-rT}\phi(d_2)\left(\sum_{k=1}^{j-1}b^kHe_{j-1-k}(-d_2)-b^j\sum_{k=1}^{N}c'_kHe_{k-1}(-d_1)\right).\end{aligned}$$

□

6. APPLICATION TO EQUITY INDEXED ANNUITIES (EIAS)

Hardy [19] describes the main types of EIAs, point-to-point, annual ratchet and high water mark, in which an embedded European call option has to be priced. We will use 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 ([19], p.248)

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

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 [19], p.249, the formula

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

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

$$R_t \overset{\mathrm{Q}}{\sim} \mathbf{GC}(a, b; \tilde{c}), \quad t = 1, \ldots, n.$$

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)

$$V_0 \;=\; \mathbf{E}^{\mathrm{Q}}(e^{-rn}B_n) \;=\; P\prod_{t=1}^{n}\left\{e^{-r}\mathbf{E}^{\mathrm{Q}}\left[1+\max\left(\alpha\left(\frac{S_t}{S_{t-1}}-1\right), 0\right)\right]\right\}. \tag{25}$$

As Hardy notes ([19], 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 9. *The no-arbitrage price of the EIA ratchet premium option described above is*

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

where C_0 is the price of a one-year call (see (17)*) with $S_0 = 1$ and $K = 1$. The first order sensitivities are*

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

with ξ replaced with one of S_0, r, b, c_j, and $\partial C_0/\partial \xi$ is given in Theorem 7. The second order sensitivity with respect to S_0 is

$$\frac{\partial^2 C_0}{\partial S_0^2} \;=\; \alpha^2 P(e^{-r} + \alpha C_0)^{n-2}\left(\frac{\partial C_0}{\partial S_0}\right)^2 + \alpha P(e^{-r} + \alpha C_0)^{n-1}\frac{\partial^2 C_0}{\partial S_0^2}.$$

7. Numerical Example

In this section, we 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 9 has a maturity of one year; we thus need the distribution of one-year returns. For illustrative purposes, we estimated the parameters of that distribution based on the S&P500 index over the period 1950-2011. (*N.B.* For pricing that is consistent with the 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.* [27] or Rompolis and Tzavalis [31] for more details. It is relatively easy to extract implied skewness and kurtosis values from short-term options on the S&P500 index, since they are numerous and very liquid. While long-term anticipation securities (LEAPS) on the S&P500 index do exist, they are not as numerous, only last up to 3 years and their liquidity is somewhat poor.)

The log-returns have the following:

mean	0.0687	
standard deviation	0.1685	
skewness	−0.8891	(26)
excess kurtosis	0.8903	

The pair of values (skewness, excess kurtosis) lies just outside the region of feasible $(\bar{s}, \bar{k})$ for the four-parameter Gram-Charlier distribution. We used maximum likelihood (with the constraint that the density be non-negative) to obtain the parameters of the $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$:

$$a = 0.0687, \quad b = 0.1685, \quad c_3 = -0.1150, \quad c_4 = 0.03598. \tag{27}$$

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

$$\bar{s} \;=\; 6c_3 \;=\; -0.6898, \quad \bar{k} \;=\; 24c_4 \;=0.8634.$$

This point is on the boundary of the feasible $(\bar{s}, \bar{k})$ region, see Figure 1.

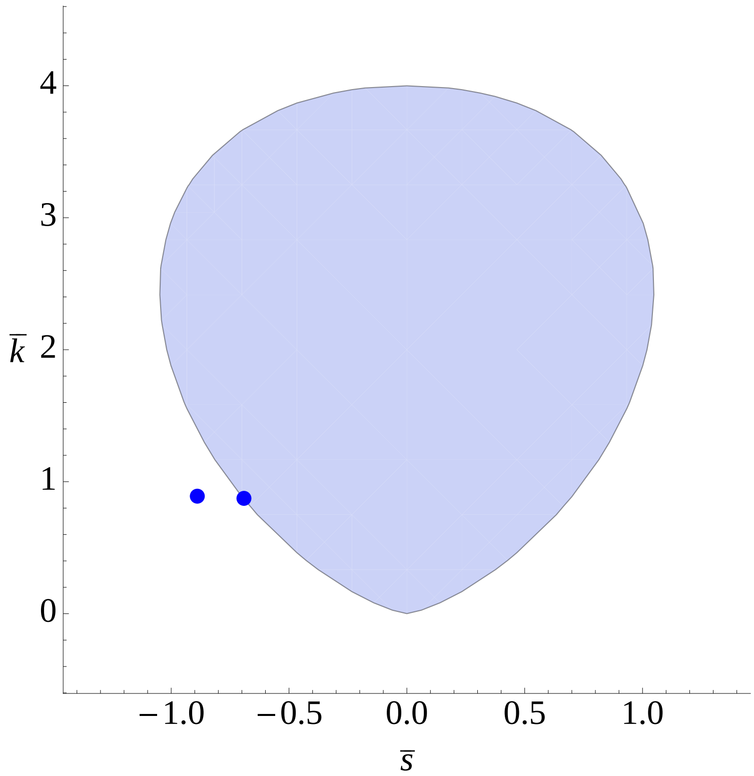

FIGURE 1. Feasible region for the skewness and excess kurtosis of the $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$ distributions.

Figure 2 shows the ratchet premium option values (Theorem 9) of an EIA with participation rate $\alpha = 0.6$ and single premium $P = 100$, as a function of skewness $\bar{s}$ and excess kurtosis $\bar{k}$. The risk-free rate is 3%, the dividend rate is 2%, and the term is 7 years. The martingale condition (16) must hold and so a is replaced with a' so that (16) 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 changing (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 $(c_3, c_4) = (0.0230,\ 0.00332567)$, the minimum at $(-0.0836,\ 0.163)$; these points correspond to $\bar{s}, \bar{k})$ equal to $(.138, .0798)$ and $(0.023, 0.003326)$, respectively) The dependence

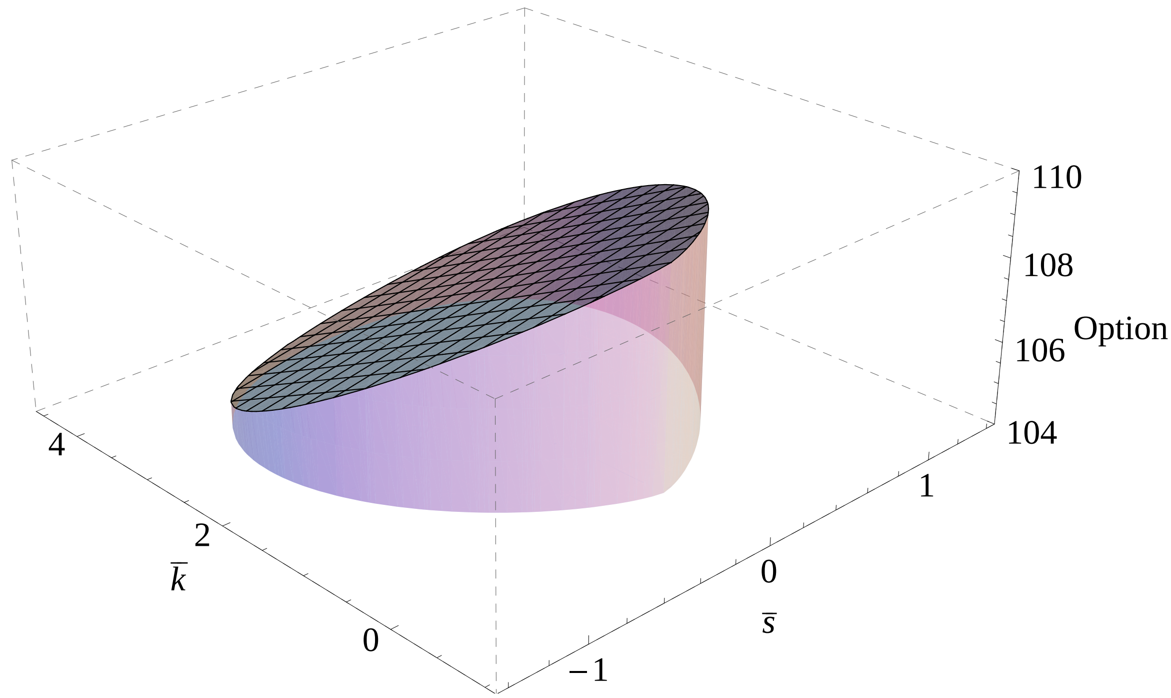

FIGURE 2. Ratchet premium option prices as a function of skewness and excess kurtosis, with $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$ distributions.

on $(\overline{s}, \overline{k})$ is nearly linear. This is easily explained. The price of a call is

$$C_0 \;=\; e^{-rT}\mathbf{E}^{\mathrm{Q}}(S_T - K) \;=\; e^{-rT}\int_{\frac{1}{b}(\log\frac{K}{S_0}-a')}^{\infty} (S_0 e^{a'+bx} - K)\phi(x)\left(\sum_{k=0}^{N} c_k He_k(x)\right) dx.$$

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

$$e^{-rT}\frac{\partial a'}{\partial c_j}\mathbf{E}^{\mathrm{Q}}(S_T \mathbf{1}_{S_T > K}) + e^{-rT}\int_{\frac{1}{b}(\log\frac{K}{S_0}-a')}^{\infty} (S_0 e^{a'+bx} - K)\phi(x) He_j(x)\, dx.$$

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 9).

The fitted parameters of the $\mathbf{GC}(a, b; c_1, c_2, c_3, c_4)$ distribution we found are:

$$a \;=\; 0.1174, b \;=\; 0.1595, c_1 \;=\; -0.3054, c_2 \;=\; 0.09542, c_3 \;=\; -0.12384, c_4 \;=\; 0.06120.$$

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 $\overline{s} = -0.5437$, $\overline{k} = 0.5092$. (*N.B.* The feasible region for $(\overline{s}, \overline{k})$ 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 $(\overline{s}, \overline{k})$ are just outside the region in Figure 1.)

Table 1 shows ratchet premium option values based on Gram-Charlier with four parameters that are lower than the Black-Scholes based ones. Row A is the Black-Scholes based data, row B is our estimation of the four-parameter Gram-Charlier (based on S&P500), row C is four-parameter Gram-Charlier with the largest negative skewness possible (see Figure 1), row D is similar but with the largest positive skewness, row E has maximum kurtosis but no skewness. Comparing the option values based on the four-parameter Gram-Charlier and those based on Black-Scholes, one notices that the maximum impact of excess kurtosis in the absence of skewness (4.3%) is larger than the maximum impact of skewness for a given level (2.4508) of excess kurtosis (3.5%).

	$\mathbf{a'}$	$\mathbf{b}$	$\mathbf{c_1}$	$\mathbf{c_2}$	$\mathbf{c_3}$	$\mathbf{c_4}$	$\bar{\mathbf{s}}$	$\bar{\mathbf{k}}$	$\mathbf{V_0(\alpha=0.6)}$	$\mathbf{\Delta(A)}$	$\mathbf{V_0(\alpha=0.419)}$	$\mathbf{\Delta(A)}$
A	-0.0042	0.1685	0	0	0	0	0	0	109.26	0	100.00	0
B	-0.0035	0.1685	0	0	-0.1150	0.0360	-0.6898	0.8634	107.60	-1.66	98.92	-1.08
C	-0.0034	0.1685	0	0	-0.1749	0.1021	-1.0493	2.4508	105.42	-3.84	97.50	-2.50
D	-0.0051	0.1685	0	0	0.1749	0.1021	1.0493	2.4508	107.39	-1.87	98.78	-1.22
E	-0.0043	0.1685	0	0	0	0.1667	0	4	104.59	-4.67	96.96	-3.04

	$\mathbf{a'}$	$\mathbf{b}$	$\mathbf{c_1}$	$\mathbf{c_2}$	$\mathbf{c_3}$	$\mathbf{c_4}$	$\bar{\mathbf{s}}$	$\bar{\mathbf{k}}$	$\mathbf{V_0(\alpha=0.6)}$	$\mathbf{\Delta(F)}$	$\mathbf{V_0(\alpha=0.441)}$	$\mathbf{\Delta(F)}$
F	0.0027	0.1595	0	0	0	.0	0	0	107.69	0	100.00	0
G	0.0451	0.1595	-.3054	.09542	-0.12384	.6120	-0.5437	0.5091	107.90	.21	100.14	0.14

TABLE 1. Comparaison of ratchet option values computed with the Black-Scholes formula and Gram-Charlier distributions. All rows except the last one are based on $\mathbf{GC}(a, b, 0, 0, c3, c4)$. Row G, based on $\mathbf{GC}(a, b, c1, c2, c3, c4)$, shows that the participation rate in the penultimate column changes with the value of b.

(See also [13] regarding the separate effects of skewness and excess kurtosis on option values.) When offering a lower participation rate of 41.9%, the insurers are just breaking even per $100 of premium when buying Black-Scholes based options; by way of contrast they save up to 3% when their coverage is based on Gram-Charlier based ratchet premium option values. Rows F and G relate to six-parameter Gram-Charlier ratchet premiums; the second one assumes the parameters we estimated, while the first one is the Black-Scholes case obtained by setting all c_k's equal to zero and using the Gram-Charlier value of b as σ in Black-Scholes (observe that in a six-parameter Gram-Charlier distribution the parameter b is not the standard deviation of log-returns, as it is in a $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$ distribution).

The pricing of this EIA spans seven periods, though only the one-year distribution parameters enter the formulas. For other types of options one might need the log-return for the whole period. For instance, it is a simple matter to find the risk-neutral distribution of the seven-year log-return in the six-parameter case above:

$$Z_7 \;=\; R_1 + \cdots + R_7 \overset{Q}{\sim} \mathbf{GC}(7a', b\sqrt{7}; c_1^{(7)}, \ldots, c_{28}^{(7)}),$$

where the parameters $\{c_k^{(7)}\}$ are obtained (as explained in Section 3) from the polynomial

$$\begin{aligned}
&(1 + c_1 t + c_2 t^2 + c_3 t^3 + c_4 t^4)^7 \\
&= 1 - 2.13757t + 2.62618t^2 - 3.08733t^3 + 3.44655t^4 - 3.22191t^5 + 2.71958t^6 - 2.22532t^7 \\
&+ 1.69033t^8 - 1.17306t^9 + 0.777132t^{10} - 0.493559t^{11} + 0.292173t^{12} - 0.162918t^{13} \\
&+ 0.0874727t^{14} - 0.0445158t^{15} + 0.0211729t^{16} - 0.00958571t^{17} + 0.00415085t^{18} \\
&- 0.0016759t^{19} + 0.000628657t^{20} - 0.000223827t^{21} + 0.0000744876t^{22} - 0.0000221865t^{23} \\
&+ 5.93937 * 10^{-6} t^{24} - 1.47127 * 10^{-6} t^{25} + 3.11682 * 10^{-7} t^{26} - 4.55625 * 10^{-8} t^{27} \\
&+ 3.2168 * 10^{-9} t^{28}.
\end{aligned}$$

8. DISCUSSION AND CONCLUSION

We have given a fairly extensive description of a general class of Gram-Charlier distributions, and derived formulas for European option prices and their partial derivatives with respect to all the parameters, as well as an explicit formula for an exponential change of measure. We have seen that EIA ratchet premium option prices are 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) distorts the pricing of EIAs and similar products. We have used maximum likelihood to compute the parameters of the Gram-Charlier distribution, using readily available software. Limiting the parameter space to the $\{c_k\}$ that yield a non-negative density is much less of a problem than it was a couple of decades ago. Readers should be aware that in the financial literature not all authors have taken the non-negativity restriction into account.

Estimation by moment matching is possible, and is especially easy for the $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$. In this case one immediately finds (a, b) from the mean and standard deviation of the data, and (c_3, c_4) follow from

$$c_3 = \frac{\overline{s}}{6}, \quad c_4 = \frac{\overline{k}}{24}.$$

This is not the procedure we recommend, as maximum likelihood is a better estimation method than moment matching. Whichever estimation method is used, the general Gram-Charlier distributions have limitations regarding which $(\overline{s}, \overline{k})$ are feasible. The skewness and kurtosis region widens as the order N of the Gram-Charlier increases. The Gram-Charlier distributions moreover have Gaussian tails, and this is where they may be inferior to some of the more sophisticated stochastic volatility models. The practitioner has to balance these limitations against the ease of use and interpretation of Gram-Charlier distributions.

The paper by León *et al.* [27] (cited in the introduction) is about a subclass of the Gram-Charlier distributions described in this paper; León *et al.* call their distributions "semi-non-parametric" 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

$$p(x) = \sum_{j=0}^{2m} \delta_j He_j(x). \tag{28}$$

However, most non-negative 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 $\mathbf{GC}(a, b; c_1, \ldots, c_{2m})$ distributions have $2m + 2$ parameters, but they are not "free", because the resulting density must be non-negative. It is then not *a priori* clear which of $\mathbf{SNP}(2m)$ or $\mathbf{GC}(a, b; c_1, \ldots, c_{2m})$ 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 result for $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$ are given in (27), and those for the order two SNP are

$$a = -0.2445, \quad b = 0.1940, \quad c_1 = 1.6140, \quad c_2 = 1.1733, \quad c_3 = 0.4253, \quad c_4 = 0.07320.$$

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

mean	0.0687
standard deviation	0.1671
skewness	−0.6289
excess kurtosis	2.5502

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 $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$ distribution. This is a puzzling result, that does not mean that $\mathbf{GC}(a, b; 0, 0, c_3, c_4)$ will always do better than **SNP**(2); it does however show that restricting the search to squares of polynomials does have consequences. Leon *et al.* [27] derive formulas for European option prices, taking the martingale restriction into account, based on expression (28); 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. Leon *et al.* [27] are correct in pointing out that there is an advantage in using $p(x) = q(x)^2$ as far as the non-negativity 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). Mathematica and other optimization packages easily accept the non-negativity restriction, though execution times can be longer. A possible advantage of our approach is that for the same number of parameters we end up with a lower order polynomial, which may well mean that the estimated density is less oscillatory (users of orthogonal expansions know from experience that higher order approximations become more and more oscillatory). A full comparison of the numerical and statistical advantages/disadvantages of our approach versus the one in [27] is an interesting avenue for further research.

In conclusion, Gram-Charlier distributions and processes have limitations compared to other models, but they do improve on the normal distribution, and their simplicity makes them a useful tool in asset modelling.

References

[1] Abken, P.A., Madan, D.B., and Ramamurtie, S. (1996). Estimation of risk-neutral and statistical densities by Hermite polynomial approximation. Federal Reserve Bank of Atlanta Working Paper 96-5.

[2] Backus, D., Foresi, S., Li, K. and Wu. L. (1997). Accounting for biases in Black-Scholes. Working paper, Stern School of Business, New York University.

[3] Ballotta, L. (2010). Efficient pricing of ratchet equity indexed annuities in a VG economy'. *North American Actuarial Journal* **14**(3): 355-368.

[4] Barton, D.E., and Dennis, K.E.R. (1952). The conditions under which Gram-Charlier and Edgeworth curves are positive definite and unimodal. *Biometrika* **39**: 425-427.

[5] Boyle, P., and Tian, W. (2008). The design of equity-indexed annuities *Insurance: Mathematics and Economics* **43**: 303-315.

[6] Brown, C., and Robinson, D. (1999). Option pricing and higher moments. *Proc. 10th Annual Asia-Pacific Research Symposium*, February, 36 pages.

[7] Chateau, J.-P. (2007). Beyond Basel-2 simplified standardized approach: Credit risk valuation of short-term loan commitments. *International Review of Financial Analysis* **16**(5): 412-433.

[8] Chateau, J.-P. (2009). Marking-to-model credit and operational risks of loan commitments: A Basel-2 advanced internal ratings-based approach. *International Review of Financial Analysis* **18**(5): 260-270.

[9] Corrado, C.J. (2007). The hidden martingale restriction in Gram-Charlier option prices. *Journal of Futures Markets* **27**(6): 517534.
[10] Corrado, C.J., and Su, T. (1996). Skewness and kurtosis in S&P 500 index returns implied by option prices. *Journal of Financial Research* **19**: 175-192.
[11] Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton.
[12] Cramér, H. (1972). On the history of certain expansions used in mathematical statistics. *Biometrika* **59**(1): 205-207.
[13] . Das, S.R., and Sundaram, R.K. (1999). Of smiles and smirks: A term structure perspective. *Journal of Financial and Quatitative Analysis* **34(2)**: 211-239.
[14] Dufresne, D. (2000). Laguerre series for Asian and other options. *Mathematical Finance* **10**(4): 407428.
[15] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications, Volume II.* Wiley.
[16] Gaillardetz, P., and Lin, X.S. (2006). Valuation of equity-linked insurance and annuity products with binomial models. *North American Actuarial Journal* **10**(4): 117-144.
[17] Hald, A. (1981). T. N. Thiele's Contributions to Statistics. *International Statistical Review* **49**: 1-20.
[18] Hald, A. (2000). T. N. The early history of the cumulants and the Gram-Charlier series. *International Statistical Review* **68**: 1-20.
[19] Hardy, M. (2003). *Investment Guarantees: Modeling and Risk Management for Equity-Linked Life Insurance*. Wiley.
[20] Jarrow R., and Rudd, A. (1982). Approximate option valuation for arbitrary stochastic processes. *Journal of Financial Economics* **10**: 347-369.
[21] Jondeau, E., and Rockinger, M. (2001). Gram-Charlier densities. *Journal of Economic Dynamics & Control* **25**: 1457-1483.
[22] Jurczenko, E., Maillet, B., and Négrea, B. (2002a). Skewness and kurtosis implied by option prices: A second comment. Discussion Paper of the LSE-FMG no. 419, July, 32 pages.
[23] Jurczenko E., Maillet, B, and Negrea, B. (2004). A note on skewness and kurtosis adjusted option pricing models under the martingale restriction. *Quantitative Finance* **4**: 479-488.
[24] Ki, H., Choi, B., Chang, K., and Lee, M. (2005). Option pricing under extended normal distribution. *Journal of Futures Markets* **25**(9): 845-871.
[25] Knight, J.L., and Satchell, S.E. (2001). *Return Distributions in Finance*. Butterworth Heineman.
[26] Laplace, P.S. (1811). Mémoire sur les intégrales définies et leur application aux probabilités, et spécialement à la recherche du milieu qu'il faut choisir entre les résultats des observations. *Mémoires de l'Académie des Sciences de Paris* **10**: 279-347. Available at http://gallica.bnf.fr/ark:/12148/bpt6k77600r/f362 (accessed 30 December 2011).
[27] León, A, Mencía, J, and Sentana, E. (2009). Parametric properties of semi-nonparametric distributions, with applications to option valuation. *Journal of Business and Economic Statistics* **27**(2): 176-192.
[28] Li, F. (2000). Option pricing: How flexible should the SPD be? *Journal of Derivatives* **7**(4): 49-65.
[29] Longstaff, S.A. (1995). Option pricing and the martingale restriction. *Review of Financial Studies* **8**(4): 1091-1124.
[30] Rouah, F.D., and Vainberg, G. (2007). *Option Pricing Models and Volatility Using Excel-VBA*. Wiley.
[31] Rompolis, L.S., and Tzavalis, E. (2007). Retrieving risk neutral densities based on risk neutral moments through a GramCharlier series expansion. *Mathematical and Computer Modelling* **46**: 225-234.
[32] Sato, K., (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.
[33] Stoyanov, J. (2000). Krein condition in probabilistic moment problems. *Bernoulli* **6**(5): 939-949.

Appendix A. The Jarrow and Rudd proof that Edgeworth series converge

The appendix of [20] is a proof that the term-by-term inversion of a Edgeworth series converges to the true distribution. This result has been quoted by several authors subsequently, though no specific example of a converging Edgeworth infinite series can be found in the option pricing litterature; instead, people always use the series truncated after three of four terms. The result in Jarrow and Rudd goes as follows. Suppose two distributions μ_1, μ_2 on $\mathbb{R}$, the first one called the reference distribution, with densities $f_1(\cdot)$, $f_2(\cdot)$ and characteristic functions

$$\psi_j(t) \;=\; \int_{\mathbb{R}} e^{itx}\, d\mu_j(x) \;=\; \int_{\mathbb{R}} e^{itx} f_j(x)\, dx, \quad j = 1, 2.$$

The idea of expressing f_2 in terms the derivatives of f_1 as shown below is not due to Schleher and Mitchell, as Jarrow and Rudd write, but goes at least as far back as Cramér ([12]). Schleher and Mitchell give no justification for the result that Jarrow and

Rudd claim to prove. The idea is to express the ratio of the two characteristic functions as a power series around the origin:

$$\frac{\psi_2(t)}{\psi_1(t)} \;=\; 1 + e_1 it + e_2(it)^2 + \cdots + e_N(it)^N + o(t^N).$$

(This is correct under the assumption that t is real and that the $N+1$-th absolute moments of μ_1 and μ_2 exist.) Each of the coefficients $\{e_k\}$ may be found explicitly in terms of the moments of order 1 to k of the two distributions. Then the above equation is rewritten as

$$\psi_2(t) \;=\; \psi_1(t)(1 + e_1 it + e_2(it)^2 + \cdots + e_N(it)^N) + \rho(t,N), \quad \rho(t,N) \;=\; o(t^N).$$

On the right-hand side each term $(it)^k\psi_1(t)$ corresponds to a derivative of f_1, via the usual formula

$$\int_{\mathbb{R}} e^{itx}(-1)^k\left(\frac{d^k f_1(x)}{dx^k}\right) dx \;=\; (it)^k\psi_1(t).$$

(Here one would need to assume that f_1 is sufficiently smooth.) Thus, taking inverse Fourier transforms,

$$f_2(x) \;=\; \sum_{k=0}^{N} e_k(-1)^k \frac{d^k f_1(x)}{dx^k} + \epsilon(x,N), \quad \epsilon(x,N) \;=\; \frac{1}{2\pi}\int_{\mathbb{R}} e^{-itx}\rho(t,N)\,dt.$$

Then Jarrow and Rudd write that if all the moments of μ_1 and μ_2 exist then "it can be shown that"

$$\lim_{N\to\infty} \sup_{x\in[-\infty,\infty]} |\epsilon(x,N)| \;=\; 0.$$

(The inclusion of $x = \pm\infty$ is probably a misprint.) The proof of this claim, they write, "is seen by noting that"

$$\text{“}\epsilon(x,N) \;=\; \sum_{k=N}^{\infty} e_k(-1)^k \frac{d^k f_1(x)}{dx^k}\text{”}. \tag{29}$$

But (29) is precisely what needs to be proved, and is left without a justification. This is a circular proof ("if A is true, then A is true").

The result cannot be true for a few obvious reasons. First, there is the moment uniqueness problem. The Edgeworth expansion is expressed in terms of the moments of μ_1 and μ_2 only. Therefore, if there is a different distribution μ_3 with the same moments as μ_2, then its Edgeworth expansion is the same as that for μ_2, so there is one Edgeworth series that needs to converge to two different values, which is not possible. (This is true in particular when one of the distributions is the lognormal or the distribution of a sum of lognormals, as in the case of Asian options.) A second impossibility lies in the fact that nowhere in the appendix of [20] is the support of μ_1 mentioned. For instance, $f_1(x)$ could be 0 for $x < 0$, and $f_2(x) > 0$ for all $x < 0$. The Edgeworth series would then be identically zero for $x < 0$, not the correct value $f_2(x)$. Another counterexample is the one where the target distribution has a discrete part.

JEAN-PIERRE CHATEAU, FACULTY OF BUSINESS ADMINISTRATION, UNIVERSITY OF MACAU, CHINA

E-mail address: jpchateau1@gmail.com

DANIEL DUFRESNE, CENTRE FOR ACTUARIAL STUDIES, UNIVERSITY OF MELBOURNE, AUSTRALIA

E-mail address: dufresne@unimelb.edu.au