The Alchemy of Probability Distributions: Beyond Gram-Charlier & Cornish-Fisher Expansions, and Skew-Normal or Kurtotic-Normal Distributions

William T. Shaw

Financial Mathematics Group, King's College, London, U.K.

Ian R.C. Buckley

Financial Mathematics Group, King's College, London, U.K.

February 18, 2007

Summary. We discuss the concept of distributional alchemy. This is defined by transmutation maps that are the functional composition of the cumulative distribution function of one distribution with the inverse cumulative distribution (quantile) function of another. We show that such maps can lead on the one hand to tractable mechanisms for the introduction of skewness into a symmetric distribution, without the pathology of Gram-Charlier expansions, and on the other hand to practical methods for converting samples from one distribution into those from another, without the limitations of Cornish-Fisher expansions. These maps have many applications in mathematical finance and statistics generally, including the assessment of distributional risk in pricing and risk calculations. We give examples of skew-uniform, skew-normal and skew-exponential distributions based on these techniques, suggest kurtotic variations, and also describe accurate methods for converting samples from the normal distribution into samples from the Student distributions or for converting one Student distribution into another.

Keywords: Distributional Alchemy, Gram Charlier, Cornish Fisher, Skew Uniform, Skew Normal, Skew T, Skew Student, Student Distribution, T Distribution, Kurtotic Uniform, Kurtotic Normal, Skew Exponential, Skew Kurtotic Normal

1. Introduction

Undergraduate students of probability usually learn the following two important facts about the cumulative distribution function of a continuous probability distribution. First, given a distribution function $F_X(\mathrm{x})$, a simple means of simulation is to set

$$X = F_X^{-1}(U) \tag{1}$$

where U is a sample from the uniform distribution on [0,1]. Second, if one makes a change of variables $Y = h(X)$, then the simplest and most reliable way of obtaining the density function for Y is to make the change of variables via the *distribution* function rather than the density.

This paper is about turning the second observation on its head and then using the resulting constructions to get more flexibility and power in the use of the first observation in simulation. That is, given a pair of distribution functions we shall attempt to infer the corresponding change of variables that links them. We shall in fact go further than this and also invent prosaic changes of variable, to be applied not to the random variable but to its ranks, in order to produce a modulation

E-mail: william.shaw@kcl.ac.uk

of a known distribution into another one of interest, such as a modulation to introduce skewness or indeed kurtosis.

The inverse cumulative distribution function (CDF): $F_X^{-1} = Q_{F_X}$ is the *quantile function* associated with the distribution. We do not have to use quantile functions for simulation, witness the use of Box-Muller, Polar-Marsaglia methods for the normal case (see, e.g. (25), and its extension to Student by Bailey (9)). A beautiful survey of a number of methods for sampling non-uniform deviates is given by Devroye (13). But it is very useful if we can employ quantile techniques, particularly if we are working with algorithms based on hypercube-filling quasi-Monte-Carlo (QMC) methods, or in particular copula methods. The issues that arise when using Box-Muller, such as the Neave effect, and bad interactions with low-discrepancy sequences, are eloquently discussed in Chapter 9 of the book by Jäckel (20).

This paper is not to be regarded as an advocate either way for the use or non-use of copula techniques, but a consideration of copula-based simulation gives us one clue as to why we might take the route described in this paper.

1.1. Quantile Functions and Copulas

Consider the case of two dimensions. If one wishes to engage in copula-based simulation with a copula that is based on an underlying bivariate distribution, then one first makes a sample (X_1, X_2) from the given bivariate (in general multivariate) distribution. Then one forms a sample from the associated copula:

$$\{U_1, U_2\} = \{F_{X_1}(X_1), F_{X_2}(X_2)\} \tag{2}$$

Then to get samples with marginals with any CDFs G_i:

$$\{Y_1, Y_2\} = \{G_1^{-1}[U_1], G_2^{-1}[U_2]\} = \{Q_{G_1}[U_1], Q_{G_2}[U_2]\} \tag{3}$$

This makes it clear that this is a context in which it is always helpful to know quantile functions.

1.2. The emergence of a transmutation mapping

In the analysis above note that we do not actually need the U_i, because when the copula comes from a "real" bivariate distribution, it would suffice to understand the *composite* mappings in the following:

$$\{Y_1, Y_2\} = \{G_1^{-1}[F_{X_1}(X_1)], G_2^{-1}[F_{X_2}(X_2)]\} \tag{4}$$

In fact, this is one of several motivations for considering composite maps of the form $y = G^{-1}[F(x)]$, where F, G are CDFs. Such a composite map essentially turns samples from one distribution, that of F, into samples from another one, that of G. We shall call this mapping a (sample) *transmutation mapping*, This is not a wholly new concept, but so far as the authors of this paper are aware an analysis of such transmutations, and of corresponding *rank transmutations*, e.g., $u \to G[F^{-1}(u)]$ have not been fully considered previously outside the asymptotic domain, and with one of the distributions being explicitly the normal distribution. There are good reasons why not! Sheer analytical tractability is, for some calculations, a major issue. However, we shall in some cases be able to use the brute force of symbolic computer algebra to overcome this. In other cases more elegant analysis will be possible. Before getting to this we need to first review the asymptotic case and mention some other motivations.

1.3. Cornish-Fisher and Gram-Charlier expansions

The idea of *approximating* one distribution in terms of another is a very old one and takes several forms, depending on whether one wishes to make the approximation explicit in terms of (a) samples; (b) the density function; (c) the distribution function, or perhaps something else. Case (a) give us the Cornish-Fisher (CF) expansions while case (b) gives us Gram-Charlier (GC) expansions. These are very well documented elsewhere. For example, the basic ideas are given in the widely available text by Abramowitz and Stegun (1), available on-line at (2). They pop up in a multitude of contexts, notably in the implementation of the addition of skewness and kurtosis to a normal or log-normal distribution. The basic idea of the GC methodology is to represent the density as a base (e.g. normal) density multiplied by an asymptotic series made up of special functions based on the base distribution multiplied by skewness, kurtosis and possibly other higher order coefficients based on moments. The CF methodology makes a parallel expansion on samples. However, there are a number of objections of either principle or practicality to the use of such methods. A non-exhaustive list as follows:

- There are many cases of interest where the moments needed to justify the CF or GC expansion do not exist, but the distribution of interest can still be expanded in terms of the target distribution. An elementary example would be the "Student" T distribution where the degrees of freedom n satisfies $n \leq 4$. The first moment needed to activate the GC/CF methodology is infinite but the expansion still exists.
- Increasing the accuracy of the series requires more high order moments than are less likely to exist (see the T again) or be non-robust functions of the data.
- In the case of the GC method for the density functions, the truncated series can give negative probability density functions, leading to:
- An arbitrary truncation decision needs to be made in order to cure the density negativity issues.
- Although perhaps less of a problem with modern computer algebra methods, the management of the series requires a detailed fluency with an associated series of special polynomials. The formulae in (1) give some indication of the thickness of the "Hermite function soup" associated with just the normal case.
- As one varies the base distribution the relevant special functions (and the identities that they obey) have to be re-established.

These considerations do not in themselves imply that considerations based on the GC or CF methodologies are wrong, but they do suggest that it may be worthwhile to seek other options. In the case of the introduction of skewness there are already other approaches that work in a closed-form and non-asymptotic representation. The elegant work of Azzalini and co-workers (5) is notable in this respect and will be discussed later in this paper. Readers should also see the extensive online bibliography helpfully provided by Azzalini (7).

1.4. Other motivations

Simulation based on copulas, the perception that there may be a more straightforward methodology than the asymptotic route, and the need for simple methods for introducing skewness, are not the only reasons for the work in this paper. There are other motivations, some of which are already well appreciated or indeed used by academics, practitioners, or both:

- Transmutation might help us generate new (hard) quantile functions from old (easy) ones for easy, QMC or copula applications;
- There is nothing special about the unit interval. We do not have to use the unit interval as a standard domain - we can change variables e.g. to Gaussian real line and use transmutation mappings for sampling.
- We can transmute a given sample to assess distributional risk in pricing/risk calculations, and avoid Monte Carlo noise in much as the same way as is employed in the construction of the "Greeks". The idea here is that having done a base case risk, VaR, pricing etc. calculation in, e.g. a normal framework, one could transmute the existing pre-calculated normal samples directly into something suitably fat-tailed to assess distribution or model risk.

It is also worth pointing out that the idea of using samples from one distribution to generate another is already well-established through another mechanism - the idea of *rejection.* This is a powerful method and is discussed, e.g. in (25). Our philosophy is based, rather, on using all the samples from one distribution in the construction of samples from another.

1.5. Plan of this paper

The plan of this work is as follows. In Section two we will give a proper definition of the transmutation maps and explain two ways in which they might be used. Section three gives examples of sample transmutation mappings computed from a pair of given distributions. Section four suggests some simple rank transmutation mappings that might be used to introduce skewness into a given base distribution, without some of the difficulties that arise with the GC method. Section five gives a detailed presentation of a structured set of mappings for the introduction of skewness and kurtosis and provides a Monte Carlo sampling method and a detailed analysis of the moment structure of a form of skew-kurtotic-normal distribution. Section six gives our conclusions and suggestions. Some of the transmutation examples will make use of the fact that the quantile function for the Student T distribution with even integer degrees of freedom is easily obtained by solving a simple polynomial equation of degree $n-1$. This was established in (28).

Acknowledgements

We are deeply grateful to Professors Adelchi Azzalini, Marc Genton and Samuel Kotz for a combination of useful comments on earlier versions of this paper and for elucidation of the history and current debate surrounding the skew-normal and related distributions.

2. Definition of transmutation mappings

In this section we write down definitions for the two cases of interest.

2.1. Sample Transmutation

Given one "base" distribution function, say $\Phi(x)$, possibly normal, and another distribution $F(x)$, we define a *sample transmutation mapping* T_S by the identity

$$F^{-1}[U] = T_S(\Phi^{-1}(U)),\ \text{i.e.,}\ T_S(z) = F^{-1}(\Phi(z)) = Q_F(\Phi(z)) \tag{5}$$

where $0 \leq U \leq 1$ and z is in its appropriate range (the real line in the normal case). So if we have the Φ quantile function we can get the Q_F quantile by post-applying T_S. This function "transmutes" samples from one distribution into samples from another. We might have a decent expression for Φ but Q_F may well be hard to determine. But we will utilize computer algebra methods for getting a series for the quantile, and we can take the functional composition of this with a series for Φ to get a good series for T_S. The creation of a transmutation map is not a new idea. It has previously found expression in the asymptotic setting via the use of Cornish-Fisher expansions, but we will now see how to use it in an essentially "exact" setting via the use of computer symbolic algebra.

2.2. Rank Transmutation

There is no particularly good reason why a transmutation mapping should be applied after applying a standard quantile rather than before. So we can define a corresponding *rank transmutation mapping* T_R by the following relationship

$$F^{-1}[U] = \Phi^{-1}(T_R(U)), \text{i.e., } T_R(u) = \Phi(F^{-1}(z)) = \Phi(Q_F(u)) \tag{6}$$

This will allow us to introduce modulations into a distribution in an exact way, and potentially avoid the use of asymptotic (Edgeworth/Gram-Charlier, or "EGC") methods and their problems. Note that equation (6) only makes sense if the two distributions have the same sample space.

2.3. Existing Examples of Exact Transmutation

We wish to emphasize that our approach has its roots in many existing constructions. Indeed, our method relies on turning around some methods of elementary probability. If one *postulates* a change of variable one ends up with a mapping on the CDFs which can in principle be reversed to extract the changes of variable. Here are a subset of known possibilities:

2.3.1. *Normal to χ_1^2*

Here we just make the mapping

$$Z \to W = Z^2 \tag{7}$$

to convert the normal distribution to a χ_1^2 or elementary gamma distribution. The mapping can be reconstructed from the two CDFs. This mapping is of course $2-1$.

2.3.2. *Exponential-Rayleigh*

If we take, for $x > 0,\ \theta > 0,\ \sigma > 0$.

$$F_1(x) = 1 - e^{-x/\theta}, \quad F_2(x) = 1 - e^{-x^2/(2\sigma^2)} \tag{8}$$

we have exponential and Rayleigh distributions. The corresponding quantile functions are well known and are

$$Q_1(u) = -\theta \log(1-u), \quad Q_2(u) = \sigma\sqrt{-2\log(1-u)} \tag{9}$$

so that all four maps, comprising sample transmutation and inversion, and rank transmutation and inversion, are all available in closed-form. The sample transmutations relating X (exponential) and Y (Rayleigh) are just:

$$Y = \sigma\sqrt{\frac{2X}{\theta}}, \quad X = \frac{\theta}{2}\frac{Y^2}{\sigma^2} \tag{10}$$

There is nothing new in these comments other than the observation that the changes of variable may be inferred from the CDFs and their inverses.

2.3.3. Beta-Student

The CDF for the Student distribution may be written as (see e.g. (28))

$$F_n(x) = \frac{1}{2}\left(1 + \mathrm{sgn}(x)(1 - \mathcal{I}_{n/(x^2+n)}\left(\frac{n}{2}, \frac{1}{2}\right)\right) \tag{11}$$

where $\mathcal{I}$ is the regularized β-functions. As usual $\mathrm{sgn}(x)$ is $+1$ if $x > 0$ and -1 if $x < 0$. The regularized beta function $\mathcal{I}_x(a, b)$ is given by

$$\mathcal{I}_x(a, b) = \frac{B_x(a, b)}{B(a, b)} \tag{12}$$

where $B(a, b)$ is the ordinary β-function and $B_x(a, b)$ is the incomplete form

$$B_x(a, b) = \int_0^x t^{(a-1)}(1 - t)^{(b-1)} dt \tag{13}$$

The quantile function for the Student distribution may be written as

$$F_n^{-1}(u) = \mathrm{sgn}\left(u - \frac{1}{2}\right)\sqrt{n\,(1/J - 1)},\ \text{where } J = I^{-1}_{\mathrm{If}\left[u<\frac{1}{2},2u,2(1-u)\right]}\left(\frac{n}{2}, \frac{1}{2}\right) \tag{14}$$

To see the transmutation to the beta distribution, we consider the CDF for $x < 0$, when it reduces to

$$F_n(x) = \frac{1}{2}\mathcal{I}_{n/(x^2+n)}\left(\frac{n}{2}, \frac{1}{2}\right) \tag{15}$$

The CDF $G(x)$ for the beta distribution with parameters α, β is just

$$G(y) = \mathcal{I}_y(\alpha, \beta) \tag{16}$$

So for example, if Y is a sample from a beta distribution with parameters $\alpha = n/2$, $\beta = 1/2$, then the transmutation map, which again is 2-1, tells us that

$$X = -\sqrt{n\left(\frac{1}{Y} - 1\right)} \tag{17}$$

is a sample from the negative portion of the Student distribution, and indeed its absolute value gives us a sample from the positive portion. Similarly if X is Student with n degrees of freedom, then

$$Y = \frac{n}{X^2 + n} \tag{18}$$

is distributed as $\mathrm{Beta}(n/2, 1/2)$. This allows numerical schemes for the Student distribution to also be based on sampling from a Beta distribution. However, the transmutation from the normal we shall present shortly is also a simple candidate for managing this.

2.3.4. Skew-normal by sample transmutation

The survey by Kotz and Vicari (22) on methods of skewing continuous distributions makes it clear that changes of variable, but usually expressed in terms of the random variable rather than its ranks, forms a key part of the research on skew distributions. Some key examples include Johnson's transformations (23), which take the form:

$$z = \nu + \delta f(y; \gamma, \sigma) \tag{19}$$

where f is monotone. Several candidates for f emerge and are summarized in (22).

3. Further Examples of sample transmutations

In this section we develop the methods needed to extract sample transmutation maps. In each case the idea will be to consider a base distribution for which the quantile function is available, or for which samples may be obtained easily by other methods. The examples we shall consider here are

- Base: normal; transmute to Student T_n for all real n;
- Base: T_4; transmute to Student T_n for small n.

We have several reasons for these choices. First the Student T is of general statistical interest but is highly pathological with regard to its moments. The k'th moment only exists if $n > k$ and is unstable when computed from a sample with n close to but above k. Second, particular choices of n in the neighbourhood of four are of considerable interest for financial applications. Next there is already a short expansion of Cornish-Fisher type available for the Student T with a normal base distribution, and we shall be able to (a) massively extend such expansions; (b) demonstrate that the existence of a high order expansion is independent of the existence of moments; (c) work out the coefficients of the relevant power series exactly rather than asymptotically in powers of n^{-1}. Second, given that we have an elementary exact quantile function for $n = 4$, then for other small n we might expect to get a more efficient representation of the samples by using a distribution that is closer than the $n = \infty$ or normal case. The idea of generating normal samples from those from a T is also an attractive one, and we shall pursue it given that the extraction of samples from a T when n is an even integer is straightforward, even for a moderately high (and hence close to normal) value of n.

The consideration of the Student is interesting from a financial risk management perspective, given that both the Student and normal distributions are of relevance to the calculation of VaR and its relations, such as coherent risk measures. The concept of transmutation from a normal is highly relevant, as the distributional risk may be estimated without recomputation of the samples, and hence one can avoid the Monte Carlo noise. This is much the same issue as arises in the calculation of elementary "Greeks"† in Monte Carlo sampling, where one should use the same sample paths, but shifted in a discrete way based on e.g. a variation of a starting value of S in order to get delta or gamma. If correlations are built based on a Gaussian copula, which is commonplace, then the resulting correlated samples may be squeezed by the transmutation maps to give, for example, a multivariate Student with any collection of marginal degrees of freedom‡. The underlying copula remains resolutely Gaussian - transmutation merely makes the calculations easier.

3.1. Transmuting the normal to the Student T

The following expansion of Cornish-Fisher type may be found in Abramowitz and Stegun (1; 2):

$$t = z + \frac{z^3 + z}{4n} + \frac{5z^5 + 16z^3 + 3z}{96n^2} + \frac{3z^7 + 19z^5 + 17z^3 - 15z}{384n^3} + \frac{79z^9 + 776z^7 + 1482z^5 - 1920z^3 - 945z}{92160n^4} + \dots \tag{20}$$

†These are the sensitivities of the value of a financial option with respect to parameter changes.

‡This gives us a pragmatic form of the multivariate T distribution that is a viable alternative to the "grouped T" of Demarta (12) and Daul *et al* (11). But note that this multivariate T constructed by transmutation of the marginals should not be seriously considered as a candidate for a canonical multivariate T. See Shaw and Lee (30) for a discussion of some new contenders for this title, in addition to the many already surveyed by Kotz and Nadarajah in their book (21).

This fourth order result may be used for simulation, but is not much use for small n. The limitations in the tails are discussed in detail in (28). The expansion eventually deteriorates in the tails whatever the value of n, though for larger n the issues are so far in the tails as to be of no practical consequence. The use of expansions of this type for approximation purposes is very old. For example, Goldberg and Levine (16) made use of the expansion as far as $O(n^{-2})$ for tabulation purposes in 1946. For our purposes we regard this expansion as an asymptotic form of a sample transmutation map, and we wish to see an efficient way of rebuilding it. We also want to try to reorganize this series. For example, we note that z appears in every order in n — can this sub-series and the corresponding series for z^3, z^5 etc. be added up? We shall see that the answer is "yes".

We consider the Student T distribution in the notation of Shaw (2006). Given $0 < u < 1$, we set

$$v = (u - \frac{1}{2})\frac{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)} = b_n(u - \frac{1}{2}) \tag{21}$$

which also serves to define b_n. The quantile function is then, given a rank u and hence a value of v, obtained by solving for x the equation

$$v = x\,{}_2F_1\left(\frac{1}{2}, \frac{n+1}{2}; \frac{3}{2}; -\frac{x^2}{n}\right) = \int_0^x ds \left(1 + \frac{s^2}{n}\right)^{-\frac{1}{2}(n+1)} \tag{22}$$

where ${}_2F_1$ is Gauss' hypergeometric function. The inversion of a series for such a CDF can be carried out step by step in any computer algebra system, following the methods described in on-line (29). In the specific computer algebra system *Mathematica* you can sometimes just ask for the inverse as a series. To give an idea of how this works, we set

```
F[x_, n_] := x*Hypergeometric2F1[1/2, (n + 1)/2, 3/2, -(x^2/n)];
```

and just ask for the inverse as a series as follows:

```
Map[Factor, InverseSeries[Series[F[x, n], {x, 0, 9}], v]]
```

which yields the output:

$$v + \frac{(n+1)v^3}{6n} + \frac{(n+1)(7n+1)v^5}{120n^2} + \frac{(n+1)\left(127n^2+8n+1\right)v^7}{5040n^3}$$
$$+\frac{(n+1)\left(4369n^3-537n^2+135n+1\right)v^9}{362880n^4} + O\left(v^{10}\right)$$

With a bit more work a corresponding tail series can be developed. See (28) for details. So we have the quantile function for general real n. This is one of several ways of direct sampling of the T. What we want to discuss here is the sample transmutation mapping, based on an expansion about the origin. Note that we could consider expansions in the tails but will focus here on the mid point of the distribution as it is a composite power series around $u = 1/2$ that reproduces, extends and re-sums the known expansions. Compare the following computer algebra program with the usual rash of high-order moments and Hermite functions that are traditionally employed. First we define some relevant functions for the normal distribution:

```
PhiMinusHalf[z_] := Erf[z/Sqrt[2]]/2;
v[pmh_,n_]:=(pmh*Sqrt[n*Pi]*Gamma[n/2])/Gamma[(n+1)/2]
```

Now we can define the transmutation as follows

```
QuantileF[v_, n_, truncation_] := InverseSeries[Series[F[x, n], {x, 0, truncation}], v]

Transmutation[z_, n_, tra_, trb_] := Module[{QF, toexp}, QF = QuantileF[v, n, tra];
 toexp = Normal[QF /. {v -> v[PhiMinusHalf[z], n], m -> n}]; Series[toexp, {z, 0, trb}]]
```

To check it is all working, we see if we can recover the known published expansion:

```
rawform = Transmutation[z, n, 10, 10];
CornishFisherExpansion = Map[Together, Normal[Series[rawform, {n, Infinity, 4}]]]
```

This produces the output

$$\begin{aligned} t = z + \frac{z^3+z}{4n} + \frac{5z^5+16z^3+3z}{96n^2} + \frac{3z^7+19z^5+17z^3-15z}{384n^3} \\ + \frac{79z^9+776z^7+1482z^5-1920z^3-945z}{92160n^4} + \ldots \end{aligned} \tag{23}$$

So we are on the right track. But in contrast with dealing with a human, we can also ask the computer to work out the following more detailed expansion:

```
rawform = Transmutation[z, n, 20, 20];
CornishFisherExpansion = Map[Together, Normal[Series[rawform, {n, Infinity, 9}]]]
```

with the result (adjusted here for spacing)

$$\begin{aligned} & z + \frac{z^3+z}{4n} + \frac{5z^5+16z^3+3z}{96n^2} + \frac{3z^7+19z^5+17z^3-15z}{384n^3} \\ & + \frac{79z^9+776z^7+1482z^5-1920z^3-945z}{92160n^4} \\ & + \frac{9z^{11}+113z^9+310z^7-594z^5-255z^3+5985z}{122880n^5} \\ & + \frac{1065z^{13}+15448z^{11}+48821z^9-82440z^7+616707z^5+6667920z^3+2463615z}{185794560n^6} \\ & + \frac{339z^{15}+6891z^{13}+41107z^{11}+113891z^9+1086849z^7+5639193z^5-18226215z^3-111486375z}{743178240n^7} \\ & + \frac{P_8(z)}{356725555200n^8} + \frac{P_9(z)}{1426902220800n^9} \end{aligned} \tag{24}$$

where

$$\begin{aligned} P_8(z) = {} & 9159z^{17} + 296624z^{15} + 3393364z^{13} + 16657824z^{11} \\ & + 27817290z^9 - 591760080z^7 - 9178970220z^5 - 42618441600z^3 - 14223634425z \end{aligned} \tag{25}$$

and

$$\begin{aligned} P_9(z) = {} & 63z^{19} - 7857z^{17} - 131468z^{15} - 5104636z^{13} - 115962198z^{11} \\ & - 1311524070z^9 - 8066259180z^7 - 5512748220z^5 + 294835704975z^3 + 1221207562575z \end{aligned} \tag{26}$$

We get a Cornish-Fisher expansion to a high order with little effort. For many applications even this many terms may be overkill. But note also we did not in fact have to expand in powers of n at all: we have a "raw form" as (just the first few terms are shown in this much detail):

```
rawseries = Series[rawform, {z, 0, 6}]
```

$$\frac{\sqrt{n}\Gamma\left(\frac{n}{2}\right)z}{\sqrt{2}\Gamma\left(\frac{n+1}{2}\right)}+\left(\frac{\sqrt{n}(n+1)\Gamma\left(\frac{n}{2}\right)^3}{12\sqrt{2}\Gamma\left(\frac{n+1}{2}\right)^3}-\frac{\sqrt{n}\Gamma\left(\frac{n}{2}\right)}{6\sqrt{2}\Gamma\left(\frac{n+1}{2}\right)}\right)z^3$$
$$+\left(\frac{\sqrt{n}\left(7n^2+8n+1\right)\Gamma\left(\frac{n}{2}\right)^5}{480\sqrt{2}\Gamma\left(\frac{n+1}{2}\right)^5}-\frac{\sqrt{n}(n+1)\Gamma\left(\frac{n}{2}\right)^3}{24\sqrt{2}\Gamma\left(\frac{n+1}{2}\right)^3}+\frac{\sqrt{n}\Gamma\left(\frac{n}{2}\right)}{40\sqrt{2}\Gamma\left(\frac{n+1}{2}\right)}\right)z^5+O\left(z^7\right) \tag{27}$$

Some of the complications and lack of precision in the Cornish-Fisher expansion arise from it having an unnecessary expansion of the gamma functions in inverse powers of n. We have literally added up this part of the expansion. So with some more computer algebra§ we can in fact write down a very detailed series without having to assume that n is large. To sort this out we recall the value of b_n, and define the quantity d_n as follows:

$$b_n=\frac{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)},\quad d_n=\frac{\sqrt{n}\Gamma\left(\frac{n}{2}\right)}{\sqrt{2}\Gamma\left(\frac{n+1}{2}\right)} \tag{28}$$

and note that the results may be expressed most succinctly by using the series representation of the inverse of the function defined in Eqn. (22).

$$x=v+\sum_{k=1}^{\infty}c_k v^{2k+1} \tag{29}$$

where the coefficients c_k were obtained in (28). They are given in a more simplified and useful form here as

$$c_k=\frac{(n+1)a_k}{n^k(2k+1)!} \tag{30}$$

$$\begin{aligned}
a_1&=1,\quad a_2=7n+1,\quad a_3=127n^2+8n+1,\quad a_4=4369n^3-537n^2+135n+1\\
a_5&=243649n^4-90488n^3+26238n^2-2504n+1\\
a_6&=20036983n^5-13250647n^4+5417578n^3-1115690n^2+95903n+1\\
a_7&=2280356863n^6-2236509264n^5+1239471171n^4-395850592n^3+69621693n^2-5211216n+1\\
a_8&=343141433761n^7-453764087825n^6+324622304493n^5-141775470493n^4\\
&\quad+38151577859n^3-5831289939n^2+388203631n+1\\
a_9&=65967241200001n^8-110906186539024n^7+98461432859068n^6-54955481849680n^5\\
&\quad+20054378275846n^4-4671822975280n^3+632882991292n^2-38001405808n+1\\
a_{10}&=15773461423793767n^9-32396923801365327n^8+34621962504766452n^7\\
&\quad-23655522692379732n^6+10950848950783482n^5-3445786392543594n^4\\
&\quad+709418470017684n^3-86442804846324n^2+4733368335639n+1
\end{aligned} \tag{31}$$

We now let

$$f_k=c_k d_n^{2k+1} \tag{32}$$

§Further *Mathematica* details are suppressed in order not to alienate some readers. The above is meant to give a flavour of how straightforward it is to generate such a basic series. The effort required to produce the more structured discussion given next is more substantial.

and suppose that the sample transmutation map is given in the form (it must be odd by symmetry since both distributions are even)

$$t = \sum_{k=1}^{\infty} g_k z^{2k+1} \tag{33}$$

Then some work with computer algebra reveals that we have the following sequence of values

$$\begin{aligned}
g_0 &= d_n \\
g_1 &= f_1 - \frac{d_n}{6} \\
g_2 &= \frac{d_n}{40} - \frac{f_1}{2} + f_2 \\
g_3 &= -\frac{d_n}{336} + \frac{19f_1}{120} - \frac{5f_2}{6} + f_3 \\
g_4 &= \frac{d_n}{3456} - \frac{583f_1}{15120} + \frac{29f_2}{72} - \frac{7f_3}{6} + f_4 \\
g_5 &= -\frac{d_n}{42240} + \frac{1573f_1}{201600} - \frac{437f_2}{3024} + \frac{91f_3}{120} - \frac{3f_4}{2} + f_5 \\
g_6 &= \frac{d_n}{599040} - \frac{2599f_1}{1900800} + \frac{15353f_2}{362880} - \frac{773f_3}{2160} + \frac{49f_4}{40} - \frac{11f_5}{6} + f_6 \\
g_7 &= -\frac{d_n}{9676800} + \frac{15459659f_1}{72648576000} - \frac{254339f_2}{23950080} + \frac{35227f_3}{259200} - \frac{3607f_4}{5040} + \frac{649f_5}{360} - \frac{13f_6}{6} + f_7 \\
g_8 &= \frac{d_n}{175472640} - \frac{6439f_1}{215255040} + \frac{34214503f_2}{14529715200} - \frac{49997f_3}{1140480} + \frac{67141f_4}{201600} - \frac{1265f_5}{1008} + \frac{299f_6}{120} - \frac{5f_7}{2} + f_8 \\
g_9 &= -\frac{d_n}{3530096640} + \frac{911145183f_1}{23712495206400} - \frac{123078503f_2}{261534873600} + \frac{231879881f_3}{18681062400} - \frac{1738547f_4}{13305600} + \frac{27841f_5}{40320} \\
&\quad - \frac{30433f_6}{15120} + \frac{79f_7}{24} - \frac{17f_8}{6} + f_9
\end{aligned} \tag{34}$$

This may seem like quite a complicated set of results and even invoke horror in the reader. But computationally it is all trivial¶. Given any n this set of coefficients may be evaluated just once, or precomputed and stored for a range of n, and then applied to a large set of sample values of z. Note that the series as given above is essentially correct to $O(z^{19})$ and when expanded in inverse powers of n is correct down to $O(n^{-9})$, which is five inverse powers of n more than the previously published results, so far as we are aware. However, there is however no need to make this last expansion as the coefficients of each power of z given are *exact*. If we compare the published result of Eqn. (20) with, for example the explicit form in Eqn. (27), which is the first three terms of the detailed result given by Eqns. (28-34), we can discover that the terms involving z in Eqn. (20) are given by the re-expansion

$$d_n = \frac{\sqrt{n}\Gamma\left(\frac{n}{2}\right)}{\sqrt{2}\Gamma\left(\frac{n+1}{2}\right)} = 1 + \frac{1}{4n} + \frac{1}{32}\left(\frac{1}{n}\right)^2 - \frac{5}{128}\left(\frac{1}{n}\right)^3 - \frac{21\left(\frac{1}{n}\right)^4}{2048} + O\left(\left(\frac{1}{n}\right)^5\right) \tag{35}$$

Readers may wish to experiment with this series. It is highly accurate and some experiments of our own have confirmed, for example, that one can simulate even a T_3 distribution very well with such

¶Perhaps not quite trivial; depending on the language being used it may be necessary to explicitly use *long* integer representations to treat some of these coefficients precisely – our own analysis is in *Mathematica*, which is immune to such difficulties because arbitrary precision arithmetic is employed.

a detailed series. Finally we note that if a unit variance expansion is needed then one must apply the further scaling

$$s = t\sqrt{\frac{n-2}{n}} \tag{36}$$

to produce samples from the Student distribution with unit variance.

3.2. Transmuting the T_4 to the T_n for n small

The normal samples employed in the discussion above can come from any sampling algorithm and do not have to be obtained by a quantile function for the normal distribution. If we want to make samples for n in the range $2 < n \leq 4$ there is a question as to how we might be most efficient if we do not care about starting from normal samples. As discussed in (28), there are closed-form expressions for the quantile functions for $n = 1, 2, 4$. As an example, we can consider using as a base distribution, not the normal, but the T_4! The analysis goes through exactly as above, and in Eqn. (33) if z is now a sample from a T_4, the corresponding coefficients are now given by:

$$\begin{aligned}
g_0 &= \frac{3d_n}{8} \\
g_1 &= \frac{27f_1}{512} - \frac{5d_n}{64} \\
g_2 &= \frac{21d_n}{1024} - \frac{135f_1}{4096} + \frac{243f_2}{32768} \\
g_3 &= -\frac{45d_n}{8192} + \frac{1017f_1}{65536} - \frac{2025f_2}{262144} + \frac{2187f_3}{2097152} \\
g_4 &= \frac{385d_n}{262144} - \frac{3355f_1}{524288} + \frac{22005f_2}{4194304} - \frac{25515f_3}{16777216} + \frac{19683f_4}{134217728} \\
g_5 &= -\frac{819d_n}{2097152} + \frac{40833f_1}{16777216} - \frac{97425f_2}{33554432} + \frac{362313f_3}{268435456} - \frac{295245f_4}{1073741824} + \frac{177147f_5}{8589934592} \\
g_6 &= \frac{3465d_n}{33554432} - \frac{117513f_1}{134217728} + \frac{1522065f_2}{1073741824} - \frac{2010015f_3}{2147483648} + \frac{5176629f_4}{17179869184} - \frac{3247695f_5}{68719476736} \\
&\quad + \frac{1594323f_6}{549755813888} \\
g_7 &= -\frac{7293d_n}{268435456} + \frac{649077f_1}{2147483648} - \frac{5452495f_2}{8589934592} + \frac{38236401f_3}{68719476736} - \frac{34499925f_4}{137438953472} + \frac{67768569f_5}{1099511627776} \\
&\quad - \frac{34543665f_6}{4398046511104} + \frac{14348907f_7}{35184372088832} \\
g_8 &= \frac{122265d_n}{17179869184} - \frac{1736163f_1}{17179869184} + \frac{36643305f_2}{137438953472} - \frac{163309797f_3}{549755813888} + \frac{772769889f_4}{4398046511104} \\
&\quad - \frac{527209155f_5}{8796093022208} + \frac{835956693f_6}{70368744177664} - \frac{358722675f_7}{281474976710656} + \frac{129140163f_8}{2251799813685248} \\
g_9 &= -\frac{255255d_n}{137438953472} + \frac{36209745f_1}{1099511627776} - \frac{117161025f_2}{1099511627776} + \frac{1287460965f_3}{8796093022208} - \frac{3827030571f_4}{35184372088832} \\
&\quad + \frac{13582221345f_5}{281474976710656} - \frac{7434564345f_6}{562949953421312} + \frac{9876830985f_7}{4503599627370496} - \frac{3658971285f_8}{18014398509481984} \\
&\quad + \frac{1162261467f_9}{144115188075855872}
\end{aligned} \tag{37}$$

It is easily verified that when $n = 4$ then $g_0 = 1$ and $g_i = 0$ for $i \geq 1$. An explicit representation of the first few terms is given by

$$t_n = \frac{3\sqrt{n}\sqrt{\pi}\Gamma\left(\frac{n}{2}\right)t_4}{8\Gamma\left(\frac{n+1}{2}\right)} + \left(\frac{9\sqrt{n}(n+1)\pi^{3/2}\Gamma\left(\frac{n}{2}\right)^3}{1024\Gamma\left(\frac{n+1}{2}\right)^3} - \frac{5\sqrt{n}\sqrt{\pi}\Gamma\left(\frac{n}{2}\right)}{64\Gamma\left(\frac{n+1}{2}\right)}\right)t_4^3$$
$$+ \left(\frac{81\sqrt{n}\left(7n^2+8n+1\right)\pi^{5/2}\Gamma\left(\frac{n}{2}\right)^5}{1310720\Gamma\left(\frac{n+1}{2}\right)^5} - \frac{45\sqrt{n}(n+1)\pi^{3/2}\Gamma\left(\frac{n}{2}\right)^3}{8192\Gamma\left(\frac{n+1}{2}\right)^3} + \frac{21\sqrt{n}\sqrt{\pi}\Gamma\left(\frac{n}{2}\right)}{1024\Gamma\left(\frac{n+1}{2}\right)}\right)t_4^5 + O\left(t_4^7\right) \tag{38}$$

The samples from the T_4 are of course best made by direct use of the quantile function as given in (28). If u is such that $0 < u < 1$ then we set, in order

$$\alpha = 4u(1-u),\ p = \frac{4}{\sqrt{\alpha}}\cos\left(\frac{1}{3}\arccos(\sqrt{\alpha})\right),\ t_4 = \mathrm{sign}(u-1/2)\sqrt{p-4} \tag{39}$$

We also remind the reader of the known quantile functions for $n = 2, 1$, which are given by

$$t_2 = \frac{2u-1}{\sqrt{2u(1-u)}},\quad t_1 = \tan(\pi(u-1/2)) \tag{40}$$

and one could treat sampling of the Student T for very low n as a problem in transmutation of the t_1 or t_2 samples. From a numerical point of view it should be noted that the convergence of a one-stop polynomial truncation of the power series based on, for example, a T_4 base is, perhaps surprisingly, less well behaved than that based on a normal base! This is not really a big problem. Rather, it motivates us to point out that a transmutation mapping can in principle be made about *any* base point. Our use of the origin is purely to illustrate the link with what has traditionally been done, in the case of the Student distribution, in terms of a series expansion based on a normal base.

3.2.1. Transmutation series about a different point

The fact that we are using series representations allows the consideration of the series being taken about a different base point. The points $z = \pm\infty$ come to mind as it is for these values that the polynomial truncations considered thus far break down. Such cases require a little more work to expand the base and target distributions around infinity, but the outcome is useful. For example, the T_4 to T_3 transmutation map may be written for large $t_4 > 0$ as

$$t_3 = \frac{\sqrt[3]{2}t_4^{4/3}}{\sqrt[6]{3}\sqrt[3]{\pi}} + \frac{20\sqrt[3]{\frac{2}{\pi}}}{9\sqrt[6]{3}t_4^{2/3}} - \frac{32^{2/3}\sqrt[6]{3}\sqrt[3]{\pi}}{5t_4^{4/3}} - \frac{145\sqrt[3]{\frac{2}{\pi}}}{81\sqrt[6]{3}t_4^{8/3}} + \frac{42^{2/3}\sqrt[3]{\pi}}{3^{5/6}t_4^{10/3}} + \frac{7672\sqrt[3]{\frac{2}{\pi}}}{2187\sqrt[6]{3}t_4^{14/3}} + \ldots \tag{41}$$

and can be used in the tail region to supplement the polynomial description. We have not yet found a corresponding tractable representation for the normal to T_n transmutation, though given the accuracy of the polynomial representation this is less of a problem.

4. Rank transmutation, skewness and kurtosis

To define a *rank transmutation mapping* in complete generality, suppose that we have two distributions with a common sample space, with CDFs F_1 and F_2. We can form

$$G_{R12}(u) = F_2\left(F_1^{-1}(u)\right),\ G_{R21}(u) = F_1\left(F_2^{-1}(u)\right) \tag{42}$$

and this pair of maps takes the unit interval $I = [0, 1]$ into itself and under suitable assumptions are mutual inverses and satisfy $G_{ij}(0) = 0, G_{ij}(1) = 1$. We shall optionally assume in addition that these rank transmutation maps are continuously differentiable. If not, a transmuted density may be discontinuous. We could consider all kinds of such maps arising from a particular choice of the F_i, but here, rather, we shall postulate some interesting forms. In general we require that rank transmutations are monotone. Our approach will be to first experiment fairly freely with some obvious transmutations, then to make a more detailed study of a composite skewness and kurtosis adjusting mapping and to explore its moment structure in some detail.

4.1. Existing representations of a skew-normal distribution

As we shall be touching on the notion of modulating a given distribution by the introduction of skewness, it is appropriate to comment on extensive work on this topic already existing, and in particular the elegant work carried out by Azzalini and co-workers. A survey of the history of continuous skewed distributions in general has been made recently by Kotz and Vicari (22). Another good entry point to the literature is the article (5), and an extensive bibliography has been made available at (7). Azzalini's own recent survey is available at (6), and forms part of a trio of articles with Genton (14; 15) that well demonstrates that this is a vigorous area of research. See also the 2006 paper by Arellano-Valle *et al* (3).

The essential idea of the framework developed by Azzalini in the univariate case is, using the normal as an example for the base case, to consider a distribution with density

$$f(x, \alpha) = 2\phi(x)\Phi(\alpha x) \tag{43}$$

where α is a perturbation parameter that may attain all real values. This produces an elegant representation for many cases of interest. It should be noted that it is generally assumed that the base density ϕ, whether normal or not, is symmetric about $x = 0$. This approach to skewing a distribution has been widely considered, but it is nevertheless interesting to consider whether there may be other natural options. We might, for example, wonder whether a formalism can be set up in which one can provide an easier description of the cumulative distribution function, or simpler Monte Carlo sampling, or removing the need to have the base distribution centred. We shall see that this can all be achieved using a different prescription. But we are emphatically not claiming that the rank transmutation approach developed below is superior, rather, it provides other options that may be more (or less) suitable in various circumstances. We also note that when $\alpha = 1$, then

$$f(x, 1) = 2\phi(x)\Phi(x) = \frac{d}{dx}(\Phi(x)^2) \tag{44}$$

and the Azzalini distribution has a closed-form CDF that is the distribution function of the maximum of two independent copies of the base distribution. Similarly when $\alpha = -1$ we obtain the distribution of the minimum. Despite our qualifications, we will note now some positive features of our own approach to be exhibited below:

- Our mappings will apply to *any* base distribution, whether symmetric, centred or even defined for negative arguments;
- Our mappings will be easily generalized to treat the introduction of some kurtosis;
- Our mappings are well-adapted to direct Monte Carlo simulation by the use of the quantile function of the base distribution;

- We will obtain the raw moments of the skewed/kurtotic distributions as simple linear functions of the transmutation parameters;
- We have explicit constructions for the CDFs in simple, *univariate* representation.

It is also important to realize that the ideas presented in our paper can be related to the pioneering work of F. de Helguero (18) in 1908. The work of De Helguero has been publicised by Azzalini (15), and as discussed in (15) consists of the multiplication of the normal distribution by a function based on a form of the skew-uniform distribution (the meaning of "skew-uniform" will be considered presently). The original Italian version of the paper has been made available on Azzalini's web site at (19). In modern notation, our understanding of de Helguero's work is that his form of the skew-normal density is, for $\alpha > 0$ (the case of negative α being obtained by reflection)

$$c\phi(x) * (1 + \frac{x}{\alpha}) \ , \quad \text{for } x > -\alpha \tag{45}$$

where ϕ is the normal density and c a normalizing constant given by

$$c^{-1} = \Phi(\alpha) + \frac{1}{\alpha}\phi(\alpha) \tag{46}$$

It is perhaps not a great leap from this idea to consider the case where we replace ϕ by the uniform distribution itself and consider a form of Eqn. (48) as defining modulations of the density of the ranks (i.e. of the CDF) of the normal, or indeed any other distribution. However, our representation is different from that given by either of Eqns. (46) or (48), but will contain the order statistic distribution of the maximum as given by Eqn. (47), as a special case.

4.2. Quadratic Transmutation

Possibly the simplest example of a rank transmutation is obtained by considering, for $|\lambda| \le 1$,

$$G_{\mathrm{R12}}(u) = u + \lambda u(1-u) \tag{47}$$

This has the consequence that the CDFs are related by

$$F_2(x) = (1+\lambda)F_1(x) - \lambda F_1(x)^2 \tag{48}$$

and *the sampling algorithm remains tractable* as the quantile functions are related by

$$F_2^{-1}(u) = F_1^{-1}\left(G_{\mathrm{R21}}(u)\right), \ G_{\mathrm{R21}}(u) = \frac{1+\lambda - \sqrt{(\lambda+1)^2 - 4\lambda u}}{2\lambda} \tag{49}$$

There are two important extremal cases. First, if $\lambda = -1$, then $G_{\mathrm{R12}}(u) = u^2$ and $F_2(x) = F_1(x)^2$ and we recognize that the distribution of F_2 corresponds to that of the maximum of two independent copies of the F_1 distribution. Correspondingly $\lambda = +1$ generates the distribution of the minimum. So this map has the same property as the Azzalini representation that the distributions of the max or min are recovered for certain values of the parameters. However, we note that the rank transmutation approach would also allow for a continuum of distributions containing as special cases the maximum or minimum of k independent copies and indeed other order statistics – it is just a matter of writing down an appropriate polynomial. So far as we can see, this is not possible within the Azzalini framework. However, the quadratic case does have a further nice property that we shall now discuss.

Note that no assumptions as to the symmetry of the underlying distribution are required. Indeed in the transmutation approach the underlying distribution need not be centred or even defined for $x < 0$, as exemplified below. However, *if* the F_1 distribution is symmetric about the origin, in the sense that

$$F_1(-x) = 1 - F_1(x) \tag{50}$$

we have the result that the distribution of the square of the transmuted random variable is identical to that of the distribution of the square of the original random variable. This follows from the following elementary algebra. Suppose that Z_λ has distribution function $F_2(x)$ with parameter λ. Let $W_\lambda = Z_\lambda^2$, then

$$P(W \leq y) = P(-\sqrt{y} \leq Z_\lambda \leq \sqrt{y}) = F_2(\sqrt{y}) - F_2(-\sqrt{y}) \tag{51}$$

Now we substitute the formula for F_2 and simplify the result using the symmetry. We see that

$$\begin{aligned}
P(W \leq y) &= F_2(\sqrt{y}) - F_2(-\sqrt{y}) \\
&= (1+\lambda)F_1(\sqrt{y}) - \lambda F_1(\sqrt{y})^2 - (1+\lambda)F_1(-\sqrt{y}) + \lambda F_1(-\sqrt{y})^2 \\
&= (1+\lambda)F_1(\sqrt{y}) - \lambda F_1(\sqrt{y})^2 - (1+\lambda)(1 - F_1(\sqrt{y})) + \lambda(1 - F_1(\sqrt{y}))^2 \\
&= (1+\lambda)F_1(\sqrt{y}) - \lambda F_1(\sqrt{y})^2 - (1+\lambda)(1 - F_1(\sqrt{y})) + \lambda(1 - 2F_1(\sqrt{y}) + F_1(\sqrt{y})^2) \\
&= 2F_1(\sqrt{y}) - 1
\end{aligned} \tag{52}$$

independently of λ. In particular, we note that, if the original distribution is symmetric, then quadratic rank transmutation preserves all even moments. This will not be true if we apply higher order powers, so while we might consider cubic, quartic and higher order transmutations, the quadratic case has this elegant property. While these considerations focus on the polynomial case and the quadratic in particular, we should also point out that frameworks for the preservation of the distribution of the square within a skewing methodology already exist. See for example, the papers by Roberts and Gesser (26), Gupta and Cheng (17), section two of the survey by Kotz and Vicari (22) and in particular the 2004 discussion by Wang *et al* (31), where the set of χ^2-preserving skewing maps is characterized; the quadratic map is one example of this. But we will also be interested later in modifying kurtosis with our framework so this is from our point of view a rather special situation.

4.3. The Skew-Uniform case

If we consider the uniform distribution on $[0, 1]$. Note that in our approach there is no requirement that the distribution be centred about a point of symmetry, and indeed no requirement that the distribution be symmetric. Then $F_1(x) = x$ and for $|\lambda| \leq 1$

$$F_2(x) = \begin{cases} 0 & \text{if } x < 0, \\ (1+\lambda)x - \lambda x^2 & \text{if } 0 \leq x \leq 1, \\ 1 & \text{if } x > 1. \end{cases} \tag{53}$$

and the corresponding density is given by

$$f_2(x) = \begin{cases} 0 & \text{if } x < 0, \\ (1+\lambda) - 2\lambda x & \text{if } 0 \leq x \leq 1, \\ 0 & \text{if } x > 1. \end{cases} \tag{54}$$

We obtain a "trapezoidal distribution" provided $|\lambda| \leq 1$. The details of a skew-uniform distribution have only been given relatively recently (24) using the Azzalini framework, in a detailed article by Nadarajah and Aryal. It should be noted that the representation here is identical, for the case $|\lambda| \leq 1$, to that given in (24). The work in (24) also cites the ease of Monte Carlo simulation obtained by simply solving the quadratic equation. This of course can be applied to an arbitrary base distribution in our representation.

4.3.1. Larger values of λ

The skew-uniform case and the comparison with the work by Azzalini, and Nadarajah and Aryal raises an amusing question as to what we might do with quadratic transmutation for values of λ greater than unity in magnitude. We need the transmutation map to take the unit interval into itself, so we can easily arrange this with a projection mechanism:

$$G_{\mathrm{R}12}(u) = \min[\max[u + \lambda u(1-u), 0], 1] \tag{55}$$

It is interesting to note that this gives a different, but equally valid, trapezoidal distribution when $|\lambda| > 1$. Figure 1 should be compared with that given in (24). In Figure 1 we show the family of distributions obtained by taking $-5/2 \leq \lambda \leq 5/2$ in steps of $1/2$. The boldest curve is the base case and increasing skewness indicated by thinner curves.

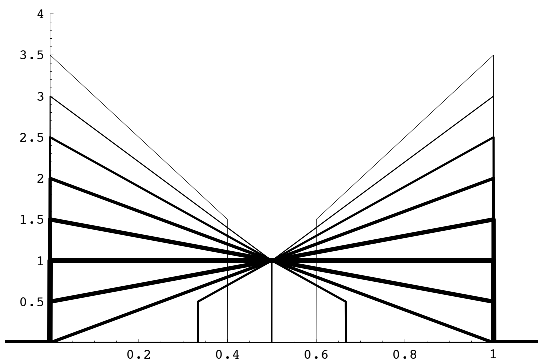

Fig. 1. Plots of the skew-uniform distribution via quadratic rank transmutation

4.4. The Skew-Exponential case

This might seem an odd case to consider but it helps to illustrate the fact that in the rank transmutation approach it is not necessary that the base distribution be centred, symmetric or even defined for negative values. Let us consider a base distribution with density, defined for $\beta > 0$,

$$f_1(x, \beta) = \begin{cases} 0 & \text{if } x < 0, \\ \beta e^{-\beta x} & \text{if } x \geq 0. \end{cases} \tag{56}$$

The corresponding CDF is clearly

$$F_1(x, \beta) = \begin{cases} 0 & \text{if } x < 0, \\ 1 - e^{-\beta x} & \text{if } x \geq 0. \end{cases} \tag{57}$$

After some trivial algebra we obtain the transmuted density in the form

$$f_2(x, \beta, \lambda) = \begin{cases} 0 & \text{if } x < 0, \\ \beta e^{-\beta x}(1-\lambda) + 2\lambda\beta e^{-2\beta x} & \text{if } x \geq 0. \end{cases} \tag{58}$$

With $\beta = 1$ and λ varying from -1 to $+1$ in steps of $1/3$ we obtain the pleasing set of curves shown in Figure 2.

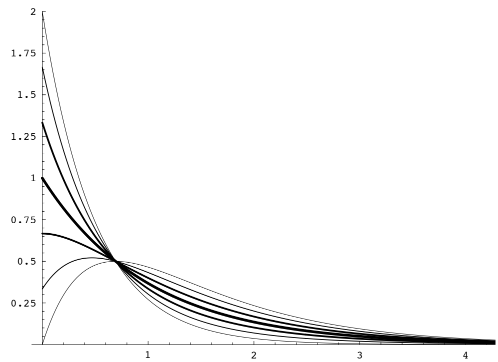

Fig. 2. Plots of the skew-exponential distribution via quadratic rank transmutation

4.5. The Skew-Normal case – a first look

In our scheme, a skew-normal distribution generated by quadratic transmutation is given by setting

$$F_1(z) = \Phi(z) = 1/2(1 + \mathrm{erf}[z/\sqrt{2}]) \tag{59}$$

and the density is then

$$f_2(x, \lambda) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}(1 + \lambda - 2\lambda\Phi(x)) \tag{60}$$

The first four moments of the transmuted distribution are then

$$E[X] = -\frac{\lambda}{\sqrt{\pi}}, \;\; E[X^2] = 1, \;\; E[X^3] = -\frac{5\lambda}{2\sqrt{\pi}}, \;\; E[X^4] = 3 \tag{61}$$

The variance is

$$\left(1 - \frac{\lambda^2}{\pi}\right) \tag{62}$$

The centred third and fourth moments are

$$E[(X - E[X])^3] = \frac{\lambda(\pi - 4\lambda^2)}{2\pi^{3/2}}, \quad E[(X - E[X])^4] = -\frac{3\lambda^4}{\pi^2} - \frac{4\lambda^2}{\pi} + 3 \tag{63}$$

and the skewness and excess kurtosis are given by

$$\gamma_1 = \frac{\lambda\left(1 - \frac{4\lambda^2}{\pi}\right)}{2\sqrt{\pi}\left(1 - \frac{\lambda^2}{\pi}\right)^{3/2}}, \quad \gamma_2 = \frac{2\lambda^2\left(\pi - 3\lambda^2\right)}{\left(\pi - \lambda^2\right)^2} \tag{64}$$

If we want to standardize this distribution to have zero mean and unit variance we form the following from an obvious linear transformation:

$$f_3(x,\lambda) = \sqrt{1-\lambda^2/\pi}\, f_2(x\sqrt{1-\lambda^2/\pi} - \frac{\lambda}{\pi}, \lambda) \tag{65}$$

We could, if we wanted to, do an series expansion of this around the standard normal distribution, in powers of both x and λ, in order to see the relation to an Gram-Charlier-type expansion, but, as with the Cornish-Fisher expansion on the quantile, this is now unnecessary!

We could also just as easily do skew-Student, skew-Cauchy or skew anything else, and the Monte Carlo sampling is precisely as tractable as it is for the base distribution via the quantile mechanism. These ideas should of course be compared with the work of many others, in particular A. Azzalini and co-workers, who use a different mechanism. See his web site (8) for details of those approaches. Our approach coincides with his in certain special cases (e.g. distributions of max/min are also contained) but is different in general.

4.6. Other types of rank transmutation maps

While our emphasis so far has been on perturbations of the symmetry, in order to introduce skewness, we can also consider other perturbations. If we stay within the polynomial structure we can consider maps of the form

$$G_{\mathrm{R12}}(u) = u + u(1-u)P(u) \tag{66}$$

where P is a polynomial with various parameters. It is of particular interest to place natural constraints on the structure of $P(u)$. In particular we might consider maps preserving particular values of u. So, for example, we have the following definition: a rank transmutation is said to be *median-preserving* if

$$G_{\mathrm{R12}}(1/2) = 1/2 \tag{67}$$

and for a polynomial map this would require that P has a zero at $u = 1/2$. We can also define a rank transmutation mapping to be *symmetric* if and only if

$$G_{\mathrm{R12}}(1-u) = 1 - G_{\mathrm{R12}}(u) \tag{68}$$

If the mapping is of the form given by P above this requires that

$$P(1-u) = -P(u) \tag{69}$$

The simplest possible type of such a mapping is obtained by choosing $P(u) = \gamma(u - 1/2)$ for some constant γ. This leads us to define a natural entity we shall term the *symmetric cubic rank transmutation mapping*. We could restrict the range of γ appropriately but will project to the unit interval to obtain a map valid for all γ as follows:

$$G_{\mathrm{R12}}(u) = \min[\max[u + \gamma u(1-u)(u-1/2), 0], 1] \tag{70}$$

It is a straightforward matter to work out the consequences of this transmutation on the standard distributions discussed above. In Figure 3 we show its effect on the uniform distribution or $-3 \le \gamma \le 3$ in steps of $1/2$. This is clearly a natural candidate for a kurtotic-uniform distribution.

If the underlying distribution is symmetric this mapping preserves this property. So a distribution with zero skew remains one with zero skew. But the kurtosis may be adjusted by this transmutation. So for the normal case we would obtain a kurtotic-normal distribution by the use of a symmetric cubic transformation. This is shown in Figure 4. It should be appreciated that this distribution eventually becomes bi-modal!

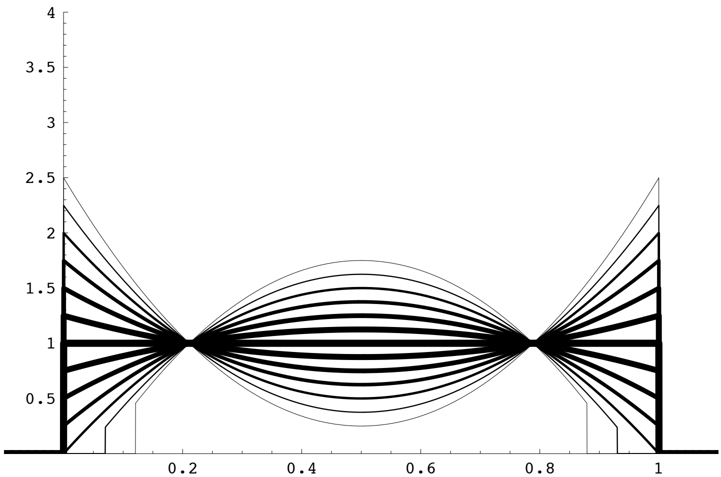

Fig. 3. Plots of the modulated uniform distribution via symmetric cubic rank transmutation

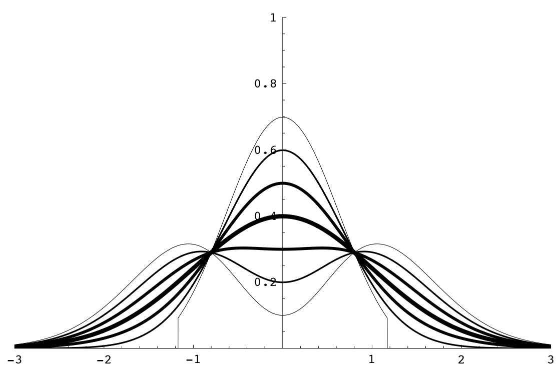

Fig. 4. Plots of the kurtotic normal distribution via symmetric cubic rank transmutation

5. A Structured Set of Skew-Kurtotic Transmutations

Following our initial experiments we shall now proceed more formally and attempt to standardize a family of polynomial rank transmutation maps. We shall also give a detailed discussion of the moment structure with a view to making it straightforward to calibrate the maps. For parameters α_1, α_2 we shall consider the polynomial family

$$P(z, \alpha_1, \alpha_2) = z - z(1-z)\left[\alpha_1 + \alpha_2(z - \frac{1}{2})\right] \tag{71}$$

The simplest expressions for the transmuted distribution are obtained when we restrict P to be a monotone increasing $1-1$ mapping of the unit interval into itself, i.e. no capping or flooring is needed. Non-negativity of P' at the end-points requires that

$$-1 - \frac{\alpha_2}{2} \le \alpha_1 \le 1 + \frac{\alpha_2}{2} \tag{72}$$

If $\alpha_2 = 0$ these conditions are sufficient given that P' is then linear and we then only need to check the end-points. If $\alpha_2 \neq 0$ we first note that non-negativity of P' at $z = 1/2$ requires that

$$\alpha_2 \le 4 \tag{73}$$

We need to establish under what conditions

$$\min_{0 \le z \le 1} P'(z) \ge 0 \tag{74}$$

Given that the above inequalities are assumed then this last relation is satisfied if $\alpha_2 < 0$ or $|\alpha_1| \ge 3|\alpha_2|/2$. If both $\alpha_2 > 0$ and $|\alpha_1| < 3|\alpha_2|/2$ we need to impose the further condition that

$$\alpha_1^2 \le 3\alpha_2 - \frac{3}{4}\alpha_2^2 \tag{75}$$

These conditions follow from observing that when $\alpha_2 \neq 0$, we may write

$$P'(z) = 3\alpha_2 \left(z - \frac{1}{2} + \frac{\alpha_1}{3\alpha_2}\right)^2 + 1 - \frac{\alpha_2}{4} - \frac{\alpha_1^2}{3\alpha_2} \tag{76}$$

These conditions may seem a trifle awkward but they are the guarantee of a globally valid density function, and the payback is the simplicity of the moment structure as we shall see shortly. As we have seen previously, this region, shown in Figure 5 in (α_1, α_2) space, can be extended by applying a floor and a cap, but within this region we have a simple polynomial mapping. The important thing is that the region contains a large open set around the origin, which is all that is needed for many practical purposes where the introduction of a modest amount of skewness and kurtosis is all that is required. The points on Figure 5 show some special cases, as follows:

- When $\alpha_2 = 0$, as previously discussed:
  - The distribution of the square is preserved if the original distribution is symmetric;
  - $P(z, 1, 0) = z^2$, which is the distribution of the maximum of two;
  - $P(z, -1, 0) = 1 - (1-z)^2$, which is the distribution of the minimum of two;
- $P(z, 3/2, 1) = z^3$, which is the distribution of the maximum of three;

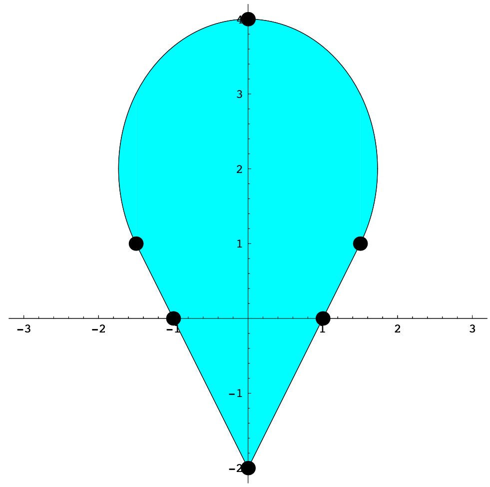

Fig. 5. Valid parameter set for unconstrained mapping in (α_1, α_2) space; special cases highlighted.

- $P(z, -3/2, 1) = 1 - (1-z)^3$, which is the distribution of the minimum of three.
- $P(z, 0, 4) = 4z^3 - 6z^2 + 3z$, with $P'(z) = 3(1-2z)^2$, which gives an extreme bimodal density vanishing at the median.
- $P(z, 0, -2) = 3z^2 - 2z^3$, with $P'(z) = 6z(1-z)$, which gives a mapping concentrating the density at the median, and is in fact the distribution of the middle of three independent samples.

So we can recover key members of the family of basic order statistics by making particular choices of the parameters, which also saturate the bounding inequalities if we do not wish to cap or floor the mappings onto the unit interval. All of these special cases, and the general formula can be directly interpreted as skew- and kurtotic- adjusted CDFs for the Uniform Distribution on $0 \leq z \leq 1$.

5.1. Monte Carlo Sampling Algorithm

In principle, this is a matter of solving a cubic equation, unless $\alpha_2 = 0$ when we have a quadratic. The solution of this was given by Tartaglia, as discussed in (27). In practice, for robust numerical use, it is a good idea to trap the special cases and treat them separately, in order that the procedure used for the general case does not become unstable. So the inverse of the rank transmutation mapping, given by solving

$$P(z, \alpha_1, \alpha_2) = u \tag{77}$$

for z, is taken to be based on the following *ordered* cases. By *ordered* we mean that e.g. the second case is only considered if the first one is not true - this means that the logic below may be used as pseudo-code.

$$z = \begin{cases} u & \text{if } \alpha_1 = \alpha_2 = 0, \\ (\alpha_1 - 1 + \sqrt{1 + \alpha_1 * (\alpha_1 + 4u - 2)})/(2\alpha_1) & \text{if } \alpha_2 = 0, \\ \sqrt[3]{u} & \text{if } \alpha_1 = 3/2 \text{ and } \alpha_2 = 1, \\ 1 - \sqrt[3]{1-u} & \text{if } \alpha_1 = -3/2 \text{ and } \alpha_2 = 1, \\ C(u, \alpha_1, \alpha_2) & \text{otherwise.} \end{cases} \tag{78}$$

where the C function denotes the general cubic solver for the other cases. This function is given by the following algorithm, using the notation and detailed implementations in Section 5.6 of (25). First we compute

$$Q = \frac{4\alpha_1^2 + 3(\alpha_2 - 4)\alpha_2}{36\alpha_2^2}, \; R = \frac{4\alpha_1^3 - 9\alpha_2(\alpha_2 + 2)\alpha_1 + 27(1-2u)\alpha_2^2}{108\alpha_2^3} \tag{79}$$

Then if $R^2 > Q^3$, the equation has one real and two complex roots. In this case we work out GC according to:

$$\begin{aligned} A &= -\text{sign}(R)\left(|R| + |\sqrt{R^2 - Q^3}|\right)^{1/3} \\ B &= \text{If } (A = 0 \text{ then } A, \text{ else } Q/A) \\ C &= A + B - \frac{1}{3}\left(\frac{a}{b} - \frac{3}{2}\right) \end{aligned} \tag{80}$$

Otherwise the cubic has three real roots, and one has to pick the right one (this applies when $\alpha_2 < 0$), and this is done by setting

$$\begin{aligned} \theta &= \arccos(R/\sqrt{Q^3}) \\ C &= -2\sqrt{Q}\cos(\frac{\theta - 2\pi}{3}) - \frac{1}{3}\left(\frac{a}{b} - \frac{3}{2}\right) \end{aligned} \tag{81}$$

An implementation of this is given in *Mathematica* in the Appendix. The final step is of course to apply the quantile function for the base case to the samples of z.

5.2. The Normal case

Now we consider the other important case where z is the CDF of the normal distribution. In this case we have that the transmuted density function takes the form

$$F_2(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} P'[\Phi[x], \alpha_1, \alpha_2] \tag{82}$$

Provided the inequalities are satisfied, it is a straightforward but lengthy‖ exercise to compute the first few moments when the base distribution is the standard normal distribution. The first five moments are:

$$E[X^k] = \begin{cases} \frac{1}{\sqrt{\pi}}\alpha_1 & \text{if } k = 1 \\ 1 + \frac{\sqrt{3}}{2\pi}\alpha_2 & \text{if } k = 2 \\ \frac{5}{2\sqrt{\pi}}\alpha_1 & \text{if } k = 3 \\ 3 + \frac{13}{2\pi\sqrt{3}}\alpha_2 & \text{if } k = 4 \\ \frac{43}{4\sqrt{\pi}}\alpha_1 & \text{if } k = 5 \end{cases} \tag{83}$$

The key central moments are

$$E[(X - \bar{X})^k] = \begin{cases} 1 + \frac{\sqrt{3}}{2\pi}\alpha_2 - \frac{1}{\pi}\alpha_1^2 & \text{if } k = 2 \\ \frac{2}{\pi^{3/2}}\alpha_1^3 - \frac{1}{2\sqrt{\pi}}\alpha_1 - \frac{3\sqrt{3}}{2\pi^{3/2}}\alpha_1\alpha_2 & \text{if } k = 3 \\ 3 - \frac{10}{\pi}\alpha_1^2 - \frac{3}{\pi^2}\alpha_1^4 + \frac{13}{2\sqrt{3}\pi}\alpha_2 + \frac{6}{\pi}\alpha_1^2\left(1 + \frac{\sqrt{3}}{2\pi}\alpha_2\right) & \text{if } k = 4 \end{cases} \tag{84}$$

‖Investigations with Mathematica and integration by parts are all that is needed.

5.2.1. Location of the median

Given that our methodology works transparently on the CDF we are also in a position to make a characterization of other measures of location. Consider, for example, the median, again with the normal background CDF. The location of the median cannot be given in an *explicit* form but it can be characterized very easily, and furthermore given explicitly for small perturbations. In the normal case the median $\hat{x}$ is given by the solution of the pair of equations

$$\begin{aligned} \hat{z} - \hat{z}(1-\hat{z})[\alpha_1 + \alpha_2(\hat{z} - 1/2)] &= \frac{1}{2} \\ \Phi(\hat{x}) &= \hat{z} \end{aligned} \tag{85}$$

As already noted, when $\alpha_1 = 0$ the median is preserved and we have $\hat{z} = \frac{1}{2}$ and $\hat{x} = 0$. In general the value of $\hat{z}$ can be found using the same cubic solver we use for Monte Carlo simulation and $\hat{x}$ found by standard methods. When $\alpha_2 = 0$ we can be more explicit and write

$$\hat{z} = \frac{1}{2} + \frac{\sqrt{1+\alpha_1^2} - 1}{2\alpha_1} \tag{86}$$

and of course $\hat{x} = \Phi^{-1}(\hat{z})$. If we have a situation where $\alpha_2 = 0$ and α_1 is small, we can go further with some manipulations of this formula composed with the normal quantile function (inverse CDF), and establish that for small α_1 the median is given by

$$\hat{x} \sim \frac{1}{2}\sqrt{\frac{\pi}{2}}\alpha_1 + O(\alpha_1^3) \tag{87}$$

More generally, when $\alpha_2 \neq 0$ but is also small, some further analysis shows that the shift in the median is given by

$$\hat{x} \sim \frac{1}{2}\sqrt{\frac{\pi}{2}}\frac{\alpha_1}{1 - \frac{\alpha_2}{4}} + O(\alpha_1^3) \tag{88}$$

By comparing this with the expression for the mean we can see that, at least for small values, α_1 modulates the separation of the median and the mean, which is another manifestation of the skewness.

5.3. Parameter estimation from data or calibration to a model

The question of how to estimate this extended set of parameters now arises, and is critical for practical applications. We shall distinguish between two cases:

(a) We have some data to which we wish to fit a distribution;
(b) We have analytical expressions for properties of another more complicated distribution that we wish to model by a skew-kurtotic transmuted form of a standard distribution.

It is important to realize that in practice, and particularly from a numerical standpoint, these may require different approaches. The first application is the classical statistical one and we shall discuss it in detail presently. The second application arises commonly in mathematical finance applications. It is often the case here that one has a distribution that is difficult to characterize in detail, but whose moment structure is nevertheless known very well. Examples include the distribution of the time average of some asset prices, or the average of several correlated assets, or a mixture of the two. Here, for example if the assets are log-normally distributed, it is difficult to give exact distributions for the various sums involved in the averages and hence the pricing of related options is awkward. However, a log-normal approximation may give reasonable answers, but we want to improve it by matching the skewness and kurtosis of the more complicated distribution. This is a good application for moment-matching as the third and fourth order moments are available exactly within the assumptions of the underlying price model. From these the skewness and kurtosis may be computed and then matched to the formulae we have given.

However, when one has real data, there are complications. While many moments of a sample may be calculated, these may be very unstable representations of the moments of the underlying distribution, even in the absence of small sample effects. A good illustration of this in general is a synthetic situation where a sample is created from a background distribution that is a Student T with n degrees of freedom. The theoretical kurtosis is

$$\gamma_2 = \frac{6}{n-4} \tag{89}$$

and diverges as $n \to 4$ from above. Yet one could always calculate a sample kurtosis. Or one only needs to observe that the higher the order of the moment, the greater its instability with respect to the introduction of outliers. While "genuine" outliers may cause one to revise one's distributional model, mathematical finance in particular is littered with plain bad data of no conceptual significance** and one really needs to consider a *robust* approach in dealing with estimation that involves not just location and scale but skewness and kurtosis.

The common traditional approach is of course to use maximum likelihood estimation. Of course, this does not necessarily give unbiased estimates but may be considered for the models proposed here. In the case of our particular form of the adjustments for skewness and kurtosis, it perhaps also makes sense to exploit the elegant form of the cumulative distribution function, which is a simple polynomial in the CDF of the base distribution, and is furthermore linear in the extra parameters α_1, α_2. On this basis we propose a parameter estimation procedure based on the CDF-fitting method, as proposed by Bandler *et al* in their 1994 paper (10). With a base distribution F given in standard form, e.g. $F(x) = \Phi(x)$ we wish to estimate location and scale parameters m, Σ and transmutation parameters α_1, α_2 so that the CDF

$$P(F(\frac{x-m}{\Sigma}), \alpha_1, \alpha_2) \tag{90}$$

is the best fit to the observed CDF. To this end we identify a sequence of (percentile) levels $u_i, i = 1, \ldots, N$ and calculate the corresponding data quantiles q_i. Working with a p-norm, we then seek to minimize an objective function

$$\mathcal{O}_p(m, \Sigma, \alpha_1, \alpha_2) = \sum_{i=1}^{N} |P(F(\frac{q_i - m}{\Sigma}), \alpha_1, \alpha_2) - u_i|^p \tag{91}$$

Bandler *et al* took $p = 1$. This and the "least squares" choice $p = 2$ represent preliminary proposals for parameter estimation routines, in that they are well adapted to our form of the CDF. In particular the minimization can be split into nested searches, where the minimization over α_i exploits the linearity of the CDF in these parameters. These details will be discussed elsewhere, along with likelihood- and moment-based estimation.

**Quoting a price temporarily in cents/pence rather than dollars/pounds comes readily to mind.

6. Discussion

Our conclusions are pretty self-evident so we shall not labour the matter. Transmutation maps are a powerful technique for:

- turning samples from one distribution into another;
- turning the ranks of one distribution into the ranks of another, e.g. to introduce skewness in a universal way.

These techniques are well adapted for quasi-Monte-Carlo and copula simulation methods, and may be extended to include a degree of kurtosis, in contrast to the traditional approach to distributional modulation. We have given explicit formulae to allow a skew-kurtotic-normal distribution to be simulated, and made preliminary proposals for parameter estimation. Clearly further work is needed to

- extend the scope of the sample transmutation maps;
- look at the rank transmutation analogues of the cases we have considered from the point of view of sample transmutation;
- look at the sample transmutation analogues of the cases we have considered from the point of view of rank transmutation;
- make more detailed comparisons with the Azzalini framework;
- look carefully at the details of the relationship with series of Gram-Charlier type;
- identify optimal parameter estimation methods.

However, initial results from our "alchemy" studies are very encouraging. The proposals for skewness adjustments are very simple and may be applied to any base distribution irrespective of whether it is symmetric or even defined for $x < 0$. The skewness adjustments may be extended to manage kurtosis adjustments as well. Our proposals also contain the basic order statistics (mix, min, middle) as special cases, and give elegant expressions for the CDFs of the relevant distributions within a univariate framework. We are also able to work out moments for the skew-kurtotic-normal developed within this framework, and these moments are all simple linear functions of the transmutation parameters. Our techniques are also very well adapted to Monte Carlo simulation as they make use of the quantile function of the base distribution composed with an elementary mapping.

References

[1] M. Abramowitz, I. A. Stegun Handbook of Mathematical Functions, Dover Edition, 1972.

[2] M. Abramowitz, I. A. Stegun Handbook of Mathematical Functions, on line at `www.math.sfu.ca/~cbm/aands/`

[3] R.B. Arellano-Valle, M. D. Branco, M.G. Genton, A unified view on skewed distributions arising form selections, *The Canadian Journal of Statistics* Vol 34, 4, pp. 581-601, 2006.

[4] A. Azzalini, A class of distributions which include the normal ones, *Scand. J. Stat.*, 12, 171-178.

[5] A. Azzalini, A. Capitanio, Distributions generated by perturbation of symmetry with emphasis on a multivariate skew-t distribution, J. Roy Statis. Soc. series B, 65, pp. 367-389, 2003.

[6] A. Azzalini, The skew-normal distribution and related multivariate families, *Scand. J. Stat.*, 32, 159-188, 2005.

[7] A. Azzalini, References on the skew-normal distribution and related ones
`http://tango.stat.unipd.it/SN/list-publ.pdf`

[8] A. Azzalini, Skew distributions web site, at
`http://tango.stat.unipd.it/SN/`

[9] R.W. Bailey, Polar generation of random variates with the t-distribution. *Mathematics of Computation*, 62, 206, pp. 779-781, 1994.

[10] J.W. Bandler, R.M. Biernacki, R.M. Cai, S.H. Chen A novel approach to statistical modelling using cumulative probability distribution fitting. *IEEE MTT S International microwave symposium digest*, Vol 1, p. 385., 1994.

[11] S. Daul, E. DeGiorgi, F. Lindskog and A.J. McNeil, The grouped t-copula with an application to credit risk. *RISK*, 16, 73 (2003).

[12] S. Demarta and A.J. Mcneil, The t copula and related copulas. *International Statistical Review*, Vol. 73, no. 1, 111. (2005).

[13] L. Devroye *Non-uniform random variate generation*, Springer 1986. Out of print - now available on-line from the author's web site at
`http://cg.scs.carleton.ca/~luc/rnbookindex.html`

[14] M.G. Genton, Discussion of The skew-normal, *Scand. J. Stat.*, 32, 189-198, 2005.

[15] A. Azzalini, Rejoinder [to Genton above], *Scand. J. Stat.*, 32, 199-200, 2005.

[16] H. Goldberg, H. Levine, Approximate formulas for the percentage points and normalization of t and x^2, *The Annals of Mathematical Statistics*, Vol. 17, 2, pp. 216-225, 1946.

[17] A.J. Gupta, J.T. Cheng, A class of multivariate skew-normal models, *Ann. Inst. Statist. Math.*, 56, No. 2 (2004), 305–315.

[18] F. de Helguero, Sulla rappresentazione analitica delle curve abnormali. In *Atti del IV Congresso Internazionale dei Matematici (Roma, 611 Aprile 1908)*, Vol. III (sez. III-B) (ed. G. Castelnuovo). R. Accademia dei Lincei, Roma, Printed in 1909.

[19] `http://azzalini.stat.unipd.it/SN/de_helguero-1908.pdf`

[20] P. Jäckel, *Monte Carlo methods in finance*, Wiley, 2002.

[21] S. Kotz and S. Nadarajah. *Multivariate t distributions and their applications*, Cambridge, 2004.

[22] S. Kotz, D. Vicari, Survey of developments in the theory of continuous skewed distributions, *Metron.* LXIII, 2, pp. 225-261, 2005.

[23] Johnson, N.L. Systems of frequency curves generated by methods of translation, *Biometrika* 36, 149-176, 1949.

[24] S. Nadarajah and G. Aryal, On the skew uniform distribution, *Random Oper. and Stoch. Eqn.* Vol. 12, 4, pp. 319-330.

[25] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in C++*, Cambridge, 2002.

[26] C. Roberts, S. Gesser A necessary and sufficient condition for the square of a random variable to be gamma. *Biometrika*, 53, 275-278, 1966.

[27] W. T. Shaw, *Complex Analysis with Mathematica*, Cambridge University Press, 2006.

[28] W. T. Shaw, Sampling Student's T distribution - use of the inverse cumulative distribution function. *Journal of Computational Finance*, Vol. 9, No. 4, 2006

`www.journalofcomputationalfinance.com/public/showPage.html?page=jcf_v9n4a2`

[29] `www.mth.kcl.ac.uk/~shaww/web_page/papers/Tsupp/InverseT.pdf`

[30] W.T. Shaw, K.T.A.Lee, Copula methods vs canonical multivariate distributions: the multivariate Student T distribution with general degrees of freedom. (KCL working paper, submitted).
`www.mth.kcl.ac.uk/~shaww/web_page/papers/MultiStudentc.pdf`

[31] J. Wang, J. Boyer, M.G. Genton, A note on an equivalence between chi-square and generalized skew-normal distributions. *Stat. Prob. Lett.* 66, 395-398.

Appendix

This appendix contains some sample computer code in *Mathematica*. The program to generate the inverse of the cubic rank transmutation is given, with $\alpha_1 \to a, \alpha_2 \to b$, by

```
Qfunc[a_, b_] := ((a/b - 3/2)^2 - 3(1/2 + 1/b - a/b))/9;

Rfunc[a_, b_, u_] := (2(a/b - 3/2)^3 - 9(a/b - 3/2)(1/2 + 1/b - a/b) - 27u/b)/54;

Sample[u_, a_, b_] := Which[
    b == 0 && a == 0, u,
    b == 0, ((a - 1) + Sqrt[1 + a*(a + 4u - 2)])/(2a),
    b == 1 && a == 3/2, u^(1/3),
    b == 1 && a == -3/2, 1 - (1 - u)^(1/3),
    True,
    Module[{R = Rfunc[a, b, u], Q = Qfunc[a, b], A, B, theta, rez},
      rez = Which[R^2 - Q^3 > 0, (
            A = -Sign[R](Abs[Sqrt[R^2 - Q^3]] + Abs[R])^(1/3);
            B = If[A == 0, 0, Q/A];
            A + B - (a/b - 3/2)/3), True,
          (theta = ArcCos[R/Sqrt[Q^3]];
            -2Sqrt[Q]Cos[(theta - 2Pi)/3] - (a/b - 3/2)/3)];
      rez]]
```

This can be used as is for exploring the transmuted quantile function, which is the composition of the above cubic inverse mapping followed by the quantile function of the base distribution.