

# New Methods for Managing “Student’s” T Distribution

William T. Shaw\*

April 20, 2006

## Abstract

With the current interest in copula methods, and fat-tailed or other non-Normal distributions, it is appropriate to investigate technologies for managing marginal distributions of interest. We explore “Student’s” T distribution, survey its simulation, and present some new techniques for simulation. In particular, for a given real (not necessarily integer) value  $n$  of the number of degrees of freedom, we give a pair of power series approximations for the inverse,  $F_n^{-1}$  of the cumulative distribution function (CDF),  $F_n$ . We also give some simple and very fast exact and iterative techniques for defining this function when  $n$  is an even integer, based on the observation that for such cases the calculation of  $F_n^{-1}$  amounts to the solution of a reduced-form polynomial equation of degree  $n - 1$ . We also explain the use of Cornish-Fisher expansions to define the inverse CDF as the composition of the inverse CDF for the Normal case with a simple polynomial map. The methods presented are well adapted for use with copula and quasi-Monte-Carlo techniques.

## 1 Introduction

There is much interest in many areas of financial modelling on the use of copulas to glue together marginal univariate distributions where there is no easy canonical multivariate distribution, or one wishes to have flexibility in the mechanism for combination. One of the more interesting marginal distributions is “Student’s” T distribution. This statistical distribution was published by W. Gosset in 1908 [23]. His employer, Guinness Breweries, required him to publish under a pseudonym, so he chose “Student.” This distribution is familiar to many through its applications to small-sample statistics in elementary discussions of statistics. It is parametrized by its mean, variance (as in the Normal case) and a further variable  $n$  indicating the number of “degrees of freedom” associated with the distribution. As  $n \rightarrow \infty$  the Normal distribution is recovered, whereas for finite  $n$  the tails of the density function decay as an inverse power of order  $(n + 1)$ , and is therefore fat-tailed relative to the Normal case. For current purposes, its fat-tailed behavior compared to the Normal distribution is of considerable interest. Recent work by Ferguson and Platen [9] suggests, for example, that the “T” (with  $n \sim 4$ ) is an accurate representation of index returns in a global setting, and propose models to underpin this idea. That returns are in general leptokurtic, in the sense that they have positive excess kurtosis (see below for definitions) has been known for over four decades – see e.g. the work in the 1960s by Mandelbrot [16] and Fama [7].

The idea of this note is to examine the univariate T distribution in a way that makes its application to current financial applications straightforward. The idea is to present several options for how to sample from a T distribution in a way that may be useful for

- Managing the simulation of T-distributed marginals in a copula framework for credit or other applications;
- Simulation of fat-tailed equity returns;
- Simulation of anything with a power-law tail behavior

---

\*Department of Mathematics King’s College, The Strand, London WC2R 2LS, England; e-mail: william.shaw@kcl.ac.uk

We should note in connection with the first item that the “T” has a clear canonical multivariate distribution only when all marginals have the same degrees of freedom (see Section 2.3, but also [8]). Throughout this paper we shall use the abbreviation “PDF” for the probability density function, “CDF” for the cumulative distribution function and “iCDF” for its inverse. Historically the iCDF has also been known as the “quantile function”.

## 1.1 Plan of this article

The plan of this work is as follows:

- In Section 2 we define the PDF and give some basic results, establishing two ways of simulation without the iCDF, for  $n$  an integer, and summarize Bailey’s method [3] for sampling without the iCDF. In order to be self-contained, we also explain the link between iCDFs and copula theory;
- In Section 3 we establish *exact* formulae for the CDF and iCDF for general real  $n$ , and explore these functions;
- In Section 4 we show that the calculation of the iCDF for even integer  $n$  is done by solving a sparse polynomial of degree  $n - 1$ , give exact solutions for  $n = 2, 4$  and iterative solutions for even  $n \geq 6$ ;
- In Section 5 we develop the central power series for the iCDF valid for general *real*  $n \geq 1$  i.e.  $n$  is not necessarily an integer;
- In Section 6 we develop the tail power series for the iCDF;
- In Section 7 we explore the use of Cornish-Fisher expansions;
- In Section 8 we present some case studies and error data, and in particular information for when to switch methods;
- In Section 9 we give a pricing example that may be useful as an elementary benchmark.

We summarize our results in Section 10. This work is supplemented by on-line supplementary material available from the links at

[www.mth.kcl.ac.uk/~shaww/web\\_page/papers/Tsupp/](http://www.mth.kcl.ac.uk/~shaww/web_page/papers/Tsupp/)

A catalogue of the contents is given at the end of this paper.

## 2 Definitions and observations related to the T

We shall begin by defining the Student’s T distribution in a way that makes manifest one method of its simulation. We let  $Z_0, Z_1, \dots, Z_n$  be standard Normal random variables and set

$$\chi_n^2 = Z_1^2 + \dots + Z_n^2 \quad (1)$$

The density function of  $\chi_n^2$  is easily worked out, using moment generating functions (see e.g. Sections 7.2 and 8.5 of Stirzaker [22], and a summary of the calculation in the on-line supplement), and is

$$q_n(z) = \frac{1}{2\Gamma(\frac{n}{2})} e^{-z/2} \left(\frac{z}{2}\right)^{\frac{n}{2}-1} \quad (2)$$

$\chi_n^2$  is a random variable with a mean of  $n$  and a variance of  $2n$ . We now define a “normal variable with a randomized variance”<sup>1</sup> in the form:

$$T = \frac{Z_0}{\sqrt{\chi_n^2/n}} \quad (3)$$

---

<sup>1</sup>This view of the T is a useful and extensible concept developed by Embrechts.

To obtain the density  $f(t)$  of  $T$  we note that

$$f(t|\chi_n^2 = \nu) = \sqrt{\frac{\nu}{2\pi n}} e^{-\frac{t^2\nu}{2n}} \quad (4)$$

Then to get the *joint* density of  $T$  and  $\chi_n^2$  we need to multiply by  $q_n(\nu)$ . Finally, to extract the univariate density for  $T$ , which we shall call  $f_n(t)$ , we integrate out  $\nu$ . The density  $f_n(t)$  is then given by

$$\int_0^\infty f(t|\chi_n = \nu)q_n(\nu)d\nu \equiv \int_0^\infty \frac{d\nu}{2\Gamma(\frac{n}{2})} \sqrt{\frac{\nu}{2\pi n}} \left(\frac{\nu}{2}\right)^{\left(\frac{n}{2}-1\right)} e^{-\left(\frac{\nu}{2} + \frac{t^2\nu}{2n}\right)} \quad (5)$$

and by the use of the following standard integral, which is just a rescaling of the variables in the integral *defining* the  $\Gamma$ -function (see formula 6.1.1 of [1, 2]):

$$\int_0^\infty x^a e^{-bx} = b^{-a-1}\Gamma(a+1) \quad (6)$$

with the choices  $a = \frac{n}{2} - \frac{1}{2}$ ,  $b = \frac{1}{2}\left(1 + \frac{t^2}{n}\right)$ ,  $x = \nu$ , we obtain the formula

$$f_n(t) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \frac{1}{\left(1 + t^2/n\right)^{\frac{n+1}{2}}} \quad (7)$$

The number  $n$ , which is often, and especially in the case of small-sample statistics, regarded as an integer, is called the “degrees of freedom” of the distribution. It is evident that a sample from this distribution can easily be obtained by using  $n + 1$  samples from the standard Normal distribution, provided  $n$  is an integer. This is well known, as is the use of a Normal variate divided by the square root of a scaled sample from the  $\chi^2$  distribution, that itself being obtained by other methods. For example, when  $n$  is an even integer the  $\chi^2$  distribution, then regarded as a gamma distribution with parameter  $n/2$  can itself be sampled efficiently based on taking logs of the product of  $n/2$  uniform deviates. See for example Chapter 7 of [20]. In this paper we shall *not* treat  $n$  as necessarily being an integer, though we shall also develop special and highly efficient methods for treating the T distribution directly in the case of  $n$  an even integer. An excellent survey of the classical methods for simulation is given by Devroye’s 1986 text [5] in Section IX.5.

General non-integer low values of  $n$  may well be of interest in financial analysis for short time-scales. Work cited in [10] suggests that very short term returns exhibit power law decay in the PDF. For a T distribution the decay of the PDF is

$$O(t^{-n-1}) \quad (8)$$

and the decay of the CDF is

$$O(t^{-n}) \quad (9)$$

so that if the power decay index in the CDF is  $q$  we take a value of  $n = q$ . The values of  $q$  reported in [10] take values in the range 2 to 6. So this leads us to consider not only small integer values of  $n : 2 \leq n \leq 6$  but also non-integer  $n$ .

## 2.1 Optimal simulation without the iCDF - Bailey’s method

The use of the obvious sampling techniques described above was essentially rendered obsolete by the discovery by Bailey in 1994 [3] that the T distribution could be sampled by a very elegant modification to the well-known Box-Muller method, and its polar variant, for the Normal distribution (see e.g. Section 7.2 of [20]). Although Bailey’s method does not supply a pair of *independent* T deviates, it otherwise works in the same way for the Student T case, and moreover is fine with non-integer degrees of freedom. The “Box-Muller” version of the algorithm is given as Theorem 2 of [3], but the more interesting polar algorithm is perhaps more pertinent and may be summarized as follows:

1. Sample two uniform variates  $u$  and  $v$  from  $[0, 1]$  and let  $U = 2u - 1$ ,  $V = 2v - 1$ ;
2. Let  $W = U^2 + V^2$ . If  $W > 1$  return to step 1 and resample;

$$3. T = U \sqrt{n(W^{-2/n} - 1)/W}.$$

This wonderful algorithm also has the manifest limit that step 3 produces the result  $T = U \sqrt{(-2 \log W)/W}$  as  $n \rightarrow \infty$ , which is the well known polar formula for the Normal case.

Bailey’s method is very useful for certain types of finance calculations. In particular, if one is using a polar method for generating Normal deviates for use in a VaR calculation, the *same* underlying random variables may simultaneously be used to compute the VaR with Normal replaced by Student T with one or more values of  $n$ , so that the difference is less subject to Monte Carlo noise. This is the same simple idea of using the same sample to compute Greeks by simple differencing in a Monte Carlo derivative valuation exercise, except here the “Greeks” would represent distributional risk.

## 2.2 Moments

The T distribution has the property that, by its symmetry, the odd moments all vanish, provided  $n$  is large enough so that they are defined. In general we can calculate the absolute moments  $\mathbb{E}[|T|^k]$  by evaluating the integral

$$\mathbb{E}[|T|^k] \equiv \frac{2}{\sqrt{n\pi}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \int_0^\infty \frac{t^k}{(1 + \frac{t^2}{n})^{(n+1)/2}} dt \quad (10)$$

Counting powers shows that this integral converges provided  $n > k$  and yields, in general, (see the definitions and results on the beta function given in Section 6.2 of [1, 2])

$$\mathbb{E}[|T|^k] = \frac{n^{\frac{k}{2}} \Gamma(\frac{k+1}{2}) \Gamma(\frac{n-k}{2})}{\sqrt{\pi} \Gamma(\frac{n}{2})} \quad (11)$$

For example, the variance exists provided  $n > 2$  and Eq. (11) simplifies to

$$\text{Var}[T] = \mathbb{E}[T^2] = \frac{n}{n-2} \quad (12)$$

The fourth moment exists for  $n > 4$  and Eq. (11) simplifies to

$$\mathbb{E}[T^4] = \frac{3n^2}{(n-2)(n-4)} \quad (13)$$

The leptokurtic behavior of the distribution is characterized by the excess kurtosis,  $\gamma_2$ , relative to that of a Normal distribution by the formula

$$\gamma_2 = \frac{\mathbb{E}[T^4]}{\text{Var}[T]^2} - 3 = \frac{6}{n-4} \quad (14)$$

These results and values for higher moments are used in Section 7.

## 2.3 The role of the iCDF in financial modelling

The main idea of this paper is to get a grip on the use of the basic result:

$$T = F_n^{-1}(U) \quad (15)$$

to define a sample from the T distribution directly, where  $U$  is uniform and  $F_n$  is the CDF for the T distribution with  $n$  degrees of freedom. Throughout this paper we use the  $F^{-1}$  notation to denote the functional inverse and not the arithmetical reciprocal, and we shall refer to it as the iCDF.

There are several good reasons for wanting to do this. First - can we be more efficient? We shall answer this question very directly for the case of low even  $n$ , for which cases we can find fast iterative algorithms relying on purely arithmetical operations and square roots, and for  $n = 2, 4$  *exact closed form* solutions needing at most the evaluation of trigonometric functions. These are of particular interest both in themselves and for seeding iterative schemes.

Second, if instead of Monte-Carlo techniques we wish to use quasi-Monte-Carlo (QMC) methods, for example to simulate a basket of size  $m$ , then it is useful to have a direct mapping from a hypercube of

dimension  $m$  (on which the QMC method is often defined), rather than, e.g. as in the case of Box-Muller or Polar-Marsaglia methods for the Normal case, one of dimension  $2m$ , or with the default sampling implied by our definition,  $m * (n + 1)$ . There may be a clear efficiency gain to be made by having an explicit representation of  $F_n^{-1}(U)$ , provided  $F_n^{-1}$  is not expensive to calculate. This is one of the motivations for the work by Moro [18] on an approximate method for  $N^{-1}(u)$  (where  $N(x) \equiv F_\infty(x)$  is the Normal CDF), and although the methods presented here are different, the motivations are closely related. There are various schools of thought on how accurate such approximations need to be. Given the many uncertainties elsewhere in financial problems, some may feel (this author does not) that it is perhaps inappropriate to dwell too much on the number of significant figures obtained – to quote J. von Neumann:

*There’s no sense in being precise when you don’t even know what you’re talking about.*

We will instead take the view that one should at least try to eliminate uncertainty due to one’s purely numerical considerations, and to characterize the errors involved. As far as this author has been able to ascertain, the main numerical analysis of the problem of finding the iCDF for the T has so far been given by Hill [12] in 1970.

## 2.4 Copulas and comments

There is currently considerable interest in the use of non-Normal marginal distributions combined to give exotic multivariate distributions. For continuous distributions, there are very few tractable cases where one can write down a useful distribution. The clear examples are the “natural” forms for the multivariate Normal and “multivariate T”, where in the latter case all the marginals have the same degrees of freedom (i.e. same  $n$ ). The problem is now routinely treated by the use of a copula function to characterize the links between the marginal distributions, with the marginals themselves specified independently. In a completely general setting, with arbitrary choices of copula and marginal distributions, a natural route is to first generate a correlated sample from a unit hypercube of dimension  $m$  based on the copula (working sequentially from the first to the  $m$ ’th value using conditional distributions), and then to apply the iCDFs for each marginal. In such an approach it is clearly helpful to have a grip on  $F^{-1}$ . Copula simulation based on this “conditional sampling” is explained in detail in [17] and also Section 6.3 of [4], with applications to Clayton, Gumbel and Frank copulas.

The same need for the iCDF of the marginals occurs does not go away when the choice of copula is such that the simulation of a correlated sample from the hypercube becomes very straightforward. When we model the correlations via the Normal copula we have the following elementary algorithm, as given by [4, 17] (see also the presentation by Duffie [6]):

1. Simulate correlated Normal variables  $(X_1, \dots, X_n)$  using Cholesky or diagonalization method;
2. Let  $U_i = N(X_i)$ , where  $N$  is the Normal CDF;
3. Feed  $U_i$  to the marginal iCDFs to get the sample  $Y_i = F^{-1}(U_i)$ .

Steps one and two simulate the Normal copula directly. It is clear that in such an approach we can use whatever iCDFs we choose at step three, in particular T distributions with many different degrees of freedom are straightforward provided we have the iCDF! There is a further major drop in complexity if we can furthermore filter the iCDF as the composition of the iCDF for the Normal case followed by a further map  $G$ . That is, if we can write

$$Y_i = F^{-1}(U_i) = G(N^{-1}(U_i)) \quad (16)$$

then steps two and three can be coalesced into the single step

$$Y_i = G(X_i) \quad (17)$$

The map  $G$  can sometimes be computed quickly and approximately using Cornish-Fisher methods and this will be discussed in Section 7, where  $G$  is given by the mapping of Eq. (75), or, with explicit maintenance of a unit variance, Eq. (76). So with the Normal copula and T marginals the simulation may become particularly straightforward.

If one prefers instead to try to work with a “canonical” multivariate distribution rather than some arbitrary copula one faces the issue of simply trying to write down the appropriate structure. The issue with

the multivariate T and the degrees of freedom having to be the same for all marginals is readily appreciated by writing down the canonical result that does exist when all  $k$  marginals have the *same* degrees of freedom  $n$ . If the correlation matrix of the marginals is given by  $R$ , then a canonical density function is given as a function of the vector  $\underline{t}$  of possible values by (see e.g. the work by Genz et al. [11], the references contained therein, and also [24])

$$\frac{\Gamma(\frac{n+k}{2})}{\Gamma(\frac{n}{2})\sqrt{|R|}(n\pi)^k} \left(1 + \frac{\underline{t}^T R^{-1} \underline{t}}{n}\right)^{-\frac{(n+k)}{2}} \quad (18)$$

It is clear that in this case the generalization from univariate to multivariate proceeds just as in the Normal case. The difficulty is that in the general case with marginals of differing degrees of freedom (i.e. different  $n$  for different elements of  $\underline{t}$ ) it is far from clear what to write down. As well, there is the issue that Eq. (18) is not in fact the only possible choice when all degrees of freedom are the same! Some of the possibilities are discussed in the book devoted to the matter by Kotz and Nadarajah [15], who also cite recent work [8] suggesting a distribution that copes with differing marginal degrees of freedom.

The other thing we must make clear is that this paper is about using T distributed marginals, potentially with *any* choice of copula, and potentially many different values for the degrees of freedom in the marginals, in a simulation process. We are *not* discussing the so-called T-copula, based on the multivariate T distribution above and where all marginals have the same degrees of freedom. This is an entirely different matter. The T-copula and its simulation are discussed in [4, 17], and the simulation is as above for the Normal case except that (a) between steps one and two one applies Eq. (3) with the same  $\chi^2$  sample for all the components, and, (b), the CDF applied in step two is then the T CDF.

### 3 The CDF for Student’s T distribution

The relevant CDF may be characterized in various different ways. Our universal starting point is the formula

$$F_n(x) = \int_{-\infty}^x f_n(t) dt = \frac{1}{\sqrt{n\pi}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \int_{-\infty}^x \frac{1}{(1+t^2/n)^{\frac{n+1}{2}}} dt \quad (19)$$

To evaluate this and try to think about inversion, one of the most obvious things to do is to make a trigonometric substitution of the obvious form,  $t = \sqrt{n} \tan \theta$ . We can then obtain the integral as a collection of powers of trigonometric functions. The resulting trigonometric expressions are well known and given by expressions 26.7.3 and 26.7.4 of Abramowitz and Stegun [1] (on-line at [2]). This author at least has not found such representations helpful in considering direct analytical inversion.

Can we get “closed-form” expressions? If we avoid the trigonometric representations we start to make progress.  $F_n(x)$  can be written in “closed form”, albeit in terms of hypergeometric functions, for general  $n$ . For example, integration in Mathematica [25] leads to the formula

$$F_n(x) = \frac{1}{2} + \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})} x {}_2F_1\left(\frac{1}{2}, \frac{n+1}{2}; \frac{3}{2}; -\frac{x^2}{n}\right) \quad (20)$$

This is fine, but as  $x$  appears in two places it does not make inversion at all obvious! The CDF may also be thought of, in a way that makes it both more obvious how to do the inversion, and also more accessible to more computer environments, in terms of  $\beta$ -functions, for we can rewrite the hypergeometric function to obtain (see Section 26.7.1 of [1, 2], bearing in mind the conversion from one- to two-sided results):

$$F_n(x) = \frac{1}{2} \left(1 + \operatorname{sgn}(x) \left(1 - I_{\frac{n}{x^2+n}}\left(\frac{n}{2}, \frac{1}{2}\right)\right)\right) \quad (21)$$

giving an expression in terms of regularized  $\beta$ -functions. As usual  $\operatorname{sgn}(x)$  is  $+1$  if  $x > 0$  and  $-1$  if  $x < 0$ . The regularized beta function  $I_x(a, b)$  employed here is given by

$$I_x(a, b) = \frac{B_x(a, b)}{B(a, b)} \quad (22)$$

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

$$B_x(a, b) = \int_0^x t^{(a-1)}(1-t)^{(b-1)} dt \tag{23}$$

Having got such a representation, this may be formally inverted to give the formula for the iCDF:

$$F_n^{-1}(u) = \operatorname{sgn}\left(u - \frac{1}{2}\right) \sqrt{n \left( \frac{1}{I_{\operatorname{If}[u < \frac{1}{2}, 2u, 2(1-u)]}\left(\frac{n}{2}, \frac{1}{2}\right)} - 1 \right)} \tag{24}$$

This last result may indeed be useful in computer environments where internal representations of inverse beta functions are available. Alternatively, if only the (forward)  $\beta$ -function is available, one can make the inversion by employing a stepping or bisection approach. The use of direct Newton-Raphson methods is awkward, because of the fact that the derivative of CDF is the PDF, which clearly becomes very small in the tails. This causes the iteration produced by such methods to be highly unstable.

If one can access an accurate representation of the inverse  $\beta$ -function then one can work directly with the formal inverse as given by Eq. (24). As an example, we can use a representation in *Mathematica* [25] to visualize the inverse for various values of  $n$ . In Figure 2.1 we show the iCDF for the cases  $n = 1, 2, 3, 4, 5, 6, 7, 8, \infty$  on the same plot. The uppermost plot in the region  $u > 0.5$  is that for  $n = 1$ , a.k.a. the Cauchy distribution with inverse CDF also given by

$$t = F_1^{-1}(u) = \tan\left(\pi\left(u - \frac{1}{2}\right)\right) \tag{25}$$

The lowest plot in the region  $u > 0.5$  is the special case of the Normal distribution,  $n = \infty$ , where we have

$$t = F_\infty^{-1}(u) = \sqrt{2}\operatorname{erf}^{-1}\left(2\left(y - \frac{1}{2}\right)\right) \tag{26}$$

The plots are constrained to the range  $-5 \leq t \leq +5$ . The plots show what we hope to see - as  $n$  decreases

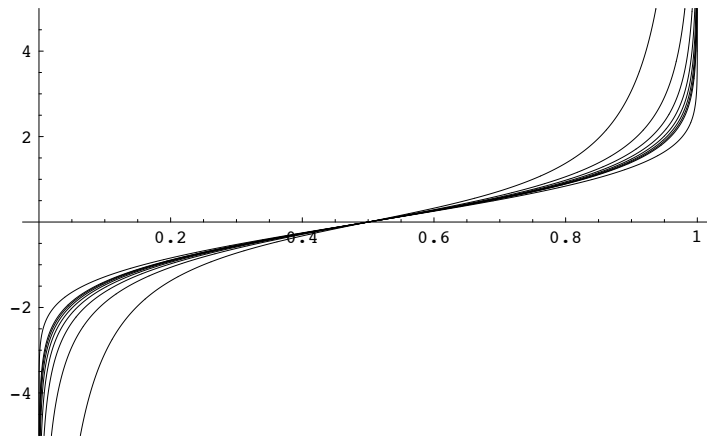


Figure 1: iCDFs for the T distribution for  $n = 1$  to 8 and  $n = \infty$ .

from infinity the image distribution becomes more fat-tailed, and the behavior is monotone in  $n$ . The general formula for the inverse is also useful, but not that fast (cf using representations of  $\operatorname{erf}^{-1}$  to do the Normal distribution), but may be useful to generate one-off large and accurate look-up tables. The on-line supplement contains an implementation of the inverse beta representation and shows how the graphic above was generated. It also shows how to generate lookup tables for the quantiles of the T distribution. One such table has been created using values of  $n$  in the range  $1 \leq n \leq 25$  in steps of 0.1, for values of  $U$  in the range  $0 < U < 1$  in steps of 0.001, with more detail in the tails. It is available as a standard comma-separated variable (CSV) file at

[www.mth.kcl.ac.uk/~shaww/web\\_page/papers/Tsupp/tquantiles.csv](http://www.mth.kcl.ac.uk/~shaww/web_page/papers/Tsupp/tquantiles.csv)

However, the generation of lookup tables aside, this representation does not give us much insight into the structure of the iCDF. Nor does it tell us whether there are any simpler representations, perhaps for particular values of  $n$ . Nor is it much use in computing environments where relatively exotic special functions are not provided. A raw version of C/C++ without function libraries comes to mind. So for our immediate purposes it will be useful to look at some cases of  $F_n(x)$  for small  $n$  very explicitly. We tabulate the cases  $n = 1$  to  $n = 6$  explicitly in terms of rational and trigonometric functions.

$$\begin{aligned}
 n & F_n(x) \\
 1 & \frac{1}{2} + \frac{1}{\pi} \tan^{-1}(x) \\
 2 & \frac{1}{2} + \frac{x}{2\sqrt{x^2+2}} \\
 3 & \frac{1}{2} + \frac{1}{\pi} \tan^{-1}\left(\frac{x}{\sqrt{3}}\right) + \frac{\sqrt{3}x}{\pi(x^2+3)} \\
 4 & \frac{1}{2} + \frac{x(x^2+6)}{2(x^2+4)^{3/2}} \\
 5 & \frac{1}{2} + \frac{1}{\pi} \tan^{-1}\left(\frac{x}{\sqrt{5}}\right) + \frac{\sqrt{5}x(3x^2+25)}{3\pi(x^2+5)^2} \\
 6 & \frac{1}{2} + \frac{x(2x^4+30x^2+135)}{4(x^2+6)^{5/2}}
 \end{aligned} \tag{27}$$

This establishes the general pattern. We can see that odd  $n$  contains a mixture of algebraic and trigonometric functions, but the case of  $n$  even is always algebraic. We now turn to this case to explore in more detail.

## 4 The case of even $n$

We have seen some simple examples above. The CDF for the case of *any* even  $n$  can be written in the form:

$$\frac{1}{2} + x \left( \frac{x^2}{n} + 1 \right)^{\frac{1-n}{2}} \left( \sum_{k=0}^{\frac{n}{2}-1} x^{2k} a(k, n) \right) \tag{28}$$

where the coefficients are defined recursively by the relations

$$a(0, n) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)} \tag{29}$$

$$a(k, n) = \frac{(n-2k)a(k-1, n)}{(2k+1)n} \tag{30}$$

This may be proved by elementary differentiation and noting the the recurrence relation causes cancellations of all non-zero powers of  $x$  in the numerator of the resulting expression. The equation that we have to solve is, given  $0 < u < 1$ :

$$x \left( \frac{x^2}{n} + 1 \right)^{\frac{1-n}{2}} \left( \sum_{k=0}^{\frac{n}{2}-1} x^{2k} a(k, n) \right) = u - \frac{1}{2} \tag{31}$$

To treat this problem we set  $p = n + x^2$ . This allows us to multiply up by the denominator and by then squaring both sides we obtain a *polynomial* equation in  $p$  that now has to be solved. We call this, with a minor abuse of historical terminology, the *resolvent polynomial* equation. The resolvent polynomials all involve a characterization of  $u$  in the form

$$\alpha = 4u(1-u) \tag{32}$$



and have an intriguing structure, as we shall now see. Given the solution,  $p$ , of the resolvent polynomial equation, the solution for  $x$  is given by

$$x = \text{sign}\left(u - \frac{1}{2}\right)\sqrt{p - n} \quad (33)$$

While it is difficult to characterize the case of general even  $n$ , and indeed it does not appear to be helpful to do so, the first few yield interesting results:

$$\begin{aligned} n = 2 : \alpha p - 2 &= 0 \\ n = 4 : \alpha p^3 - 12p - 16 &= 0 \\ n = 6 : \alpha p^5 - 135p^2 - \frac{1215}{4}p - \frac{2187}{2} &= 0 \\ n = 8 : \alpha p^7 - 2240p^3 - 7168p^2 - 35840p - 204800 &= 0 \\ n = 10 : \alpha p^9 - \frac{196875p^4}{4} - \frac{1640625p^3}{8} - \frac{10546875p^2}{8} \\ &\quad - \frac{615234375p}{64} - \frac{2392578125}{32} = 0 \end{aligned} \quad (34)$$

The on-line supplement shows how to generate, and exhibits, the resolvent polynomial equations for even  $n \leq 20$ . We now look at their solutions.

#### 4.1 Some Simple Exact Solutions for the iCDF

The cases when  $n = 1$  and  $n = \infty$  are well known as the Cauchy distribution and Normal distribution. It should be clear from the table of resolvent polynomial equations that  $n = 2, 4$  can be solved exactly and we also have a new way of investigating the cases  $n = 6, 8, 10, \dots$ . As a simple reminder, the inverse CDF for the  $n = 1$  case, the standard Cauchy distribution, is

$$x = \tan\left(\pi\left(u - \frac{1}{2}\right)\right) \quad (35)$$

##### 4.1.1 $n = 2$

This is now trivial as the resolvent polynomial is linear. After some simplification we obtain

$$x = \frac{2u - 1}{\sqrt{2u(1 - u)}} \quad (36)$$

This result was certainly known by Hill by 1970 [12]. Hill noted the invertibility of the case  $n = 2$  (but not, apparently, the general polynomial structure this was part of) and also started the development of the tail series discussed later in this paper (though to rather fewer terms). Hill’s paper is also available on-line at [13]. The invertibility of this case is also given as Theorem 3.3 of [5]. The  $n = 2$  Student distribution has also very recently been promoted as a pedagogical tool by Jones [14], who also noted the simple iCDF formula, but its financial applications are perhaps limited due to the problem of infinite variance. Further interesting properties of this distribution have been discussed by Nevzorov *et al* [19].

##### 4.1.2 $n = 4$

The resolvent polynomial equation is now a cubic in reduced form (no quadratic term). A cubic in reduced form may be solved by exploiting the identity

$$(p - A - B) * (p - A\omega - B\omega^2) * ((p - A\omega^2 - B\omega) \equiv p^3 - 3ABp - A^3 - B^3 \quad (37)$$

where  $\omega = e^{\frac{2\pi i}{3}}$  is the standard cube root of unity. We just have to solve some auxiliary equations for  $A$  and  $B$ . This is just a modern formulation of the solution due to Tartaglia (see [21]). After some work along these lines and some simplification we obtain the solution in the form:

$$p = \frac{4}{\sqrt{\alpha}} \cos\left(\frac{1}{3} \arccos(\sqrt{\alpha})\right) \quad (38)$$

and where, as before,

$$x = \text{sign}\left(u - \frac{1}{2}\right)\sqrt{p-4}, \quad \alpha = 4u(1-u) \quad (39)$$

Once one has the solution in the form of Eq. (38) it is possible to give an easier justification of it. If we let  $p = \frac{4}{\sqrt{\alpha}} \cos y$ , then the  $n = 4$  part of Eq. (34) becomes the condition

$$4 \cos^3 y - 3 \cos y \equiv \cos(3y) = \sqrt{\alpha} \quad (40)$$

and the result of Eq. (38) is immediate!

The exact solution for  $F_4^{-1}$  presented above for the case  $n = 4$  is easily applied to random samples from the uniform distribution to produce a simulation of the  $n = 4$  distribution. But there is more reason to consider this case than the mere “doability” of the inversion. The case  $n = 4$  corresponds to a case of finite variance and *infinite* kurtosis. In fact, as we decrease  $n$  from  $\infty$  and consider it as a real number, it is the point at which the kurtosis becomes infinite. It is therefore an interesting case from a risk management point of view, in that it represents a good alternative base case to consider other than the Normal case. So perhaps VaR simulations might be tested in the log-Student- $(n = 4)$  case as well as in the log-Normal case. As discussed in the introduction, recent work by Ferguson and Platen [9] also suggests that  $n = 4$  is an accurate representation of index returns in a global setting.

#### 4.1.3 $n \geq 6$

In this case we obtain a quintic, septic, nonic equation and so on, that in general cannot be solved in closed form by elementary methods. However, now we are armed with simple polynomial equations, we can employ efficient iterations schemes such as Newton-Raphson (note that this was not a good idea for the original distribution function due to the smallness of its derivative, i.e. the density, especially in the tails). This author has not investigated the Galois groups of these polynomials for further analytical insight<sup>2</sup>. The solution of the quintic example, given that it is in principal quintic form, can be carried out in terms of hypergeometric functions, but this turns out to be slower than the iterative methods discussed below. By the *principal quintic form* we mean a quintic with no terms in  $p^4, p^3$ . Similarly, the polynomial of degree 7 has no terms in  $p^6, p^5, p^4$ , and so on. In the case of the cubic this allows us to proceed straight to the solution. In the higher order cases the author does not know in general what interesting simplifications might be obtained from the fact that the resolvent polynomials are rather sparse, and depending only on  $u$  through the highest order term and then through the factor  $\alpha$ . But what we *can* say is that this sparseness in the polynomial coefficients allows a Newton-Raphson iterative scheme to proceed very efficiently, as there are fewer operations to be carried out than in the case of a general polynomial problem.

Elementary algebra makes it easy to define the associated iteration schemes. In the case  $n = 6$  the relevant Newton-Raphson iteration takes the form

$$p_{k+1} = \frac{2(8\alpha p_k^5 - 270p_k^2 + 2187)}{5(4\alpha p_k^4 - 216p_k - 243)} \quad (41)$$

For  $n = 8$  we have

$$p_{k+1} = \frac{2}{7} \left( 3p_k + \frac{640(p_k(p_k(p_k+4)+24)+160)}{p_k(\alpha p_k^5 - 960p_k - 2048) - 5120} \right) \quad (42)$$

For  $n = 10$  we have

$$p_{k+1} = \frac{8p_k}{9} + \frac{218750(4p_k(p_k(2p_k(p_k+5)+75)+625)+21875)}{9(8p_k(p_k(8\alpha p_k^6 - 175000p_k - 546875) - 2343750) - 68359375)} \quad (43)$$

The relevant expressions for the cases  $n = 12, 14, 16, 18, 20$  are given in the on-line supplement together with code to generate them for any even  $n$ .

<sup>2</sup>The author would be grateful to receive enlightenment from Galois theory experts.

#### 4.1.4 Seeding the iterations

These iteration schemes need to be supplemented by a choice of starting value. A straightforward choice is to use the exact solution for  $n = 2$ , for which the value of  $x^2$  will be slightly higher than than for a higher value of  $n$ . In this case, unwinding the transformation, the starting value of the iteration may be taken to be:

$$p_0 = 2\left(\frac{1}{\alpha} - 1\right) + n \tag{44}$$

and the result is extracted via

$$x = \text{sign}\left(u - \frac{1}{2}\right)\sqrt{p - n} \tag{45}$$

and  $\alpha = 1 - 4\left(u - \frac{1}{2}\right)^2$  as before. More exotic seeding schemes that lead to faster evaluation are available in the on-line supplement. We will only summarize the idea here. These all exploit the fact that the form of the cubic problem for  $n = 4$  gives a clue how the solution to the other cases *scales*. Some numerical experimentation shows that in general one can write  $p = \frac{n}{\alpha^{2/n}} \times x$ , and while this author cannot determine a nice formula for  $x$  for even  $n \geq 6$ , the solution for  $x$  is always a slowly varying and bounded function of  $\alpha$  of order unity. When  $n = 2$  we have  $x = 1$  and when  $x = 4$  it is as given by Eq. (38). When  $n = 6$  the equation for  $x$  becomes, with  $b = \alpha^{1/3}$  and  $0 \leq b \leq 1$ ,

$$x^5 - \frac{5}{8}x^2 - \frac{15b}{64}x - \frac{9b^2}{64} = 0 \tag{46}$$

The solution to this equation<sup>3</sup> varies smoothly and monotonically from  $x = 1$  when  $b = 1$  down to  $\frac{5^{1/3}}{2}$  when  $b = 0$ , and a good seed can be built from interpolation on this basis. Similar methods apply for higher  $n$  as discussed in the supplement.

The combination of exact solutions and iterative Newton-Raphson methods has been compared with the inverse beta function method in Mathematica, and in the on-line supplement it is checked that the two methods agree for  $n = 2, 4, 6, 8, 10$  with a difference of less than  $10^{-11}$  with a default termination criteria for the iteration where needed.

## 4.2 Low odd $n$

We now turn to the more awkward case of low odd  $n$ . There is no problem with  $n = 1$ , but the general issues involved are well exemplified by the first few cases  $n = 3, 5, 7$ . We have:

$$F_3(x) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1}\left(\frac{x}{\sqrt{3}}\right) + \frac{\sqrt{3}x}{\pi(x^2 + 3)} \tag{47}$$

$$F_5(x) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1}\left(\frac{x}{\sqrt{5}}\right) + \frac{\sqrt{5}x(3x^2 + 25)}{3\pi(x^2 + 5)^2} \tag{48}$$

$$F_7(x) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1}\left(\frac{x}{\sqrt{7}}\right) + \frac{\sqrt{7}x(15x^4 + 280x^2 + 1617)}{15\pi(x^2 + 7)^3} \tag{49}$$

If we consider  $n = 3$ , we wish to solve the equation

$$\pi\left(u - \frac{1}{2}\right) = \tan^{-1}\left(\frac{x}{\sqrt{3}}\right) + \frac{\sqrt{3}x}{(x^2 + 3)} \tag{50}$$

for  $x$  in terms of  $u$ . Equivalently, we can take the trigonometric form

$$\pi\left(u - \frac{1}{2}\right) = \theta + \sin \theta \cos \theta \tag{51}$$

where  $x = \sqrt{3} \tan \theta$ . Neither of these representations offer any immediate analytical insight nor are they helpful for Newton-Raphson solution. However, it does suggest that a simpler numerical scheme may be helpful. The latter representation may be written in the form

$$\theta = G(\theta) = \pi\left(u - \frac{1}{2}\right) - \frac{1}{2} \sin(2\theta) \tag{52}$$

---

<sup>3</sup>It would be interesting to know the Galois group of Eq. (46) in particular as  $b$  varies.

This may be made the basis of an elementary “cobwebbing” scheme based on the iteration

$$\theta_k = G(\theta_{k-1}) \tag{53}$$

with a suitable choice of starting point. As before, this can be based on the  $n = 2$  case, and we take

$$\theta_0 = \tan^{-1} \left( \frac{1}{\sqrt{6}} \frac{(u - \frac{1}{2})}{u(1-u)} \right) \tag{54}$$

This can be coded up rapidly with a suitable termination criteria and it works reasonably well. We also note that the convergence condition  $|G'(\theta)| < 1$  within the range of interest is satisfied except at  $\theta = 0, \pm\pi/2$  but the iteration at zero terminates immediately in any case. The convergence is slowest in a punctured neighbourhood of  $\theta = 0, u = 1/2$ , and there are also issues in the far tails. The remedy is a better choice of starting value with good behavior near the slow-convergence points but we shall defer the discussion of this until after we have discussed the power and asymptotic series. The power series we shall derive provides a much better starting value for any iteration scheme in the neighborhood of  $u = 1/2$ . We shall also have to confront the fact that when we go to  $n = 5$  the cobwebbing idea breaks down as the derivative exceeds unity in magnitude in a significant range of  $x$ . So we do not proceed further with the discussion of special methods for odd integer  $n$ . Devroye’s book [5] has an interesting discussion of the  $n = 3$  case in Exercise II.2.4.

## 5 The central power series for the iCDF

We now turn attention to the case of general (and not necessarily integer)  $n$ . We need to solve the following equation for  $x$ , where we note that it is easier to work from the mid-point  $u = 1/2$ :

$$u - \frac{1}{2} = \frac{1}{\sqrt{n\pi}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \int_0^x \frac{1}{(1 + s^2/n)^{\frac{n+1}{2}}} ds \tag{55}$$

This tells us that  $x$  is manifestly an odd function of  $u - 1/2$ . Absorbing the normalizing factor and exploiting the oddness, we work with the problem in the power series form:

$$x = F_n^{-1}(u) = v + \sum_{k=1}^{\infty} c_k v^{2k+1}, \quad v = (u - 1/2)\sqrt{n\pi} \frac{\Gamma[\frac{n}{2}]}{\Gamma[\frac{n+1}{2}]} \tag{56}$$

The integrand may be worked out as a power series, integrated term by term, and then we substitute our power series assumption for  $x$ . This results in an increasingly unpleasant non-linear iteration but is one that is easily managed in a symbolic computation environment such as *Mathematica* [25]. The code for doing this is available in the on-line supplement. The first nine values of the coefficients are:

$$\begin{aligned} c_1 &= \frac{1}{6} + \frac{1}{6n} \\ c_2 &= \frac{7}{120} + \frac{1}{15n} + \frac{1}{120n^2} \\ c_3 &= \frac{127}{5040} + \frac{3}{112n} + \frac{1}{560n^2} + \frac{1}{5040n^3} \\ c_4 &= \frac{4369}{362880} + \frac{479}{45360n} - \frac{67}{60480n^2} + \frac{17}{45360n^3} + \frac{1}{362880n^4} \\ c_5 &= \frac{34807}{5702400} + \frac{153161}{39916800n} - \frac{1285}{798336n^2} + \frac{11867}{19958400n^3} - \frac{2503}{39916800n^4} + \frac{1}{39916800n^5} \\ c_6 &= \frac{20036983}{6227020800} + \frac{70691}{64864800n} - \frac{870341}{691891200n^2} + \frac{67217}{97297200n^3} - \frac{339929}{2075673600n^4} + \frac{37}{2402400n^5} \\ &+ \frac{1}{6227020800 n^6} \end{aligned} \tag{57}$$

$$\begin{aligned}
 c_7 &= \frac{2280356863}{1307674368000} + \frac{43847599}{1307674368000n} - \frac{332346031}{435891456000n^2} + \frac{843620579}{1307674368000n^3} - \frac{326228899}{1307674368000n^4} \\
 &+ \frac{435891456000n^5}{21470159} - \frac{261534873600n^6}{1042243} + \frac{1307674368000n^7}{32285445833} \\
 c_8 &= \frac{50812489728000}{404003599} - \frac{1710035712000n}{123706507} - \frac{88921857024000n^2}{24262727} + \frac{91423417}{177843714048n^3} - \frac{51811946317}{177843714048000n^4} \\
 &+ \frac{4446092851200n^5}{65967241200001} - \frac{8083805184000n^6}{14979648446341} + \frac{22230464256000n^7}{26591354017} + \frac{355687428096000n^8}{73989712601} \\
 c_9 &= \frac{121645100408832000}{5816850595639} - \frac{40548366802944000n}{44978231873} - \frac{259925428224000n^2}{176126809} + \frac{206879422464000}{n^3} \\
 &- \frac{20274183401472000n^4}{49573465457} + \frac{355687428096000n^5}{4222378423} - \frac{5304600576000n^6}{1} \\
 &+ \frac{10137091700736000n^7}{13516122267648000n^8} + \frac{1}{121645100408832000n^9}
 \end{aligned}$$

and so on. The coefficients  $c_{10}$  through  $c_{30}$  are given in the on-line supplement, together with the code to generate them. C/C++ programmers should note that the supplement contains both exact and numerical representations - the latter being more suitable for coding up in such a language. It is easy to check that this series works in the case of the known exact solutions. For example, letting  $n \rightarrow \infty$  we obtain the series for the inverse error function with scaling of the arguments implied by the definition of  $v$ :

$$\sqrt{2}\operatorname{erf}^{-1}\left(\frac{x\sqrt{2}}{\sqrt{\pi}}\right) = x + \frac{x^3}{6} + \frac{7x^5}{120} + \frac{127x^7}{5040} + \frac{4369x^9}{362880} + \dots \tag{58}$$

Less obvious (and best checked symbolically) is the emergence of the series for the tangent function to deal with the Cauchy distribution in the case  $n = 1$ , as well as the exact cases  $n = 2, 4$ .

How good are these expansions considered truncated to give simple polynomials? Given that we have dealt with cases of low  $n$ , let’s consider the case  $n = 11$ . It turns out that the error gets smaller as  $n$  gets larger, as well as decreasing the more terms one takes in the series. Let us also consider a rather modest truncation using *only* the terms given above, so that we go as far as  $v^{19}$ . The results are shown in Figure 2. This is reasonably pleasing. One can easily build in more terms and get fast results in compiled code - we are only working out polynomials and the Gamma functions can be tabulated in advance for a large range of  $n$  and then Stirling’s formula applied for large  $n$ .

$$v = (u - 1/2)\sqrt{n\pi} \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n+1}{2})} = (u - 1/2)\sqrt{2\pi} \left(1 + \frac{1}{4n} + \frac{1}{32} \left(\frac{1}{n}\right)^2 - \frac{5}{128n^3} - \frac{21}{2048n^4} + \dots\right) \tag{59}$$

But this result does give a *power series* about  $u = \frac{1}{2}$  whose radius of convergence is  $1/2$ . We know that there will be a divergence as we approach  $u = 0, 1$  so a polynomial approximation can only take us so far. We need to look separately at the tails, and will now proceed to do so.

## 6 The tail power series for the iCDF

We have considered several approaches so far. We have a small number of exact solutions and some fast iterative methods that work over the whole range for small to moderate  $n$ . We have a power series that works for all  $n$  but needs many terms to work well in the approximate region  $|u - \frac{1}{2}| > 0.4$ . To complete the power series analysis we need to understand the tails better. We proceed as before, but work from  $u = 1$  as a base point. All results can by symmetry be applied to the series about  $u = 0$ . We let

$$(1 - u)\sqrt{n\pi} \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n+1}{2})} = w = \int_x^\infty \frac{1}{(1 + \frac{s^2}{n})^{\frac{n+1}{2}}} ds \tag{60}$$

The integral may be evaluated in terms of a series of inverse powers of  $x$ , the first few terms of the resulting equation being

$$w = \left(\frac{1}{x}\right)^n n^{\frac{n}{2}-\frac{1}{2}} - \frac{(n+1)\left(\frac{1}{x}\right)^{n+2} n^{\frac{n}{2}+\frac{3}{2}}}{2(n+2)} + \frac{(n+1)(n+3)\left(\frac{1}{x}\right)^{n+4} n^{\frac{n}{2}+\frac{5}{2}}}{8(n+4)} + \dots \tag{61}$$

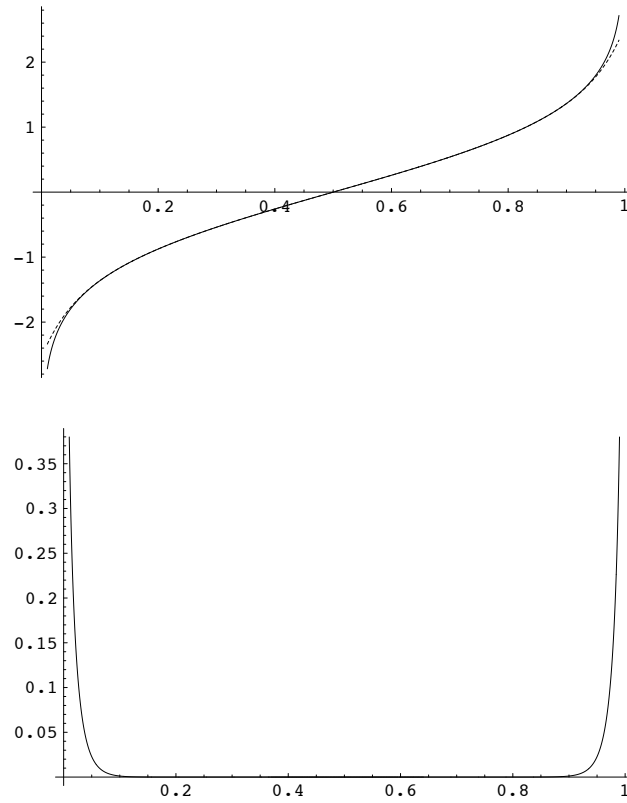


Figure 2: Top: Exact & central series to  $c_9$  (dashed) iCDF,  $n = 11$ . Bottom: absolute error.

We now proceed as before, postulating an appropriate series for  $x$  as a function of  $w$ . This time a little experimentation is needed to get the right form for evaluation. After some trial and error, we find that the right *ansatz* for the series is given by

$$x = \sqrt{n} (\sqrt{nw})^{-1/n} \left( 1 + \sum_{k=1}^{\infty} (\sqrt{nw})^{\frac{2k}{n}} d(k) \right) \quad (62)$$

We now substitute this expression into our equation relating  $x$  to  $w$  and proceed as before, extracting each term through an increasingly non-linear recursion using symbolic computation methods. The first few terms

in the series are

$$\begin{aligned}
d_1 &= -\frac{(n+1)}{2(n+2)} \\
d_2 &= -\frac{n(n+1)(n+3)}{8(n+2)^2(n+4)} \\
d_3 &= -\frac{n(n+1)(n+5)(3n^2+7n-2)}{48(n+2)^3(n+4)(n+6)} \\
d_4 &= -\frac{n(n+1)(n+7)(15n^5+154n^4+465n^3+286n^2-336n+64)}{384(n+2)^4(n+4)^2(n+6)(n+8)} \\
d_5 &= -\frac{n(n+1)(n+3)(n+9)(35n^6+452n^5+1573n^4+600n^3-2020n^2+928n-128)}{1280(n+2)^5(n+4)^2(n+6)(n+8)(n+10)} \\
d_6 &= -\frac{n(n+1)(n+11)P_6(n)}{46080(n+2)^6(n+4)^3(n+6)^2(n+8)(n+10)(n+12)} \\
P_6(n) &= 945n^{11} + 31506n^{10} + 425858n^9 + 2980236n^8 + 11266745n^7 + 20675018n^6 + 7747124n^5 \\
&\quad - 22574632n^4 - 8565600n^3 + 18108416n^2 - 7099392n + 884736
\end{aligned}$$

Further terms are given in the on-line supplement. Before analyzing the error characteristics of the tail series, and its combination with the central power series, we explore another approach that will also turn out to make a useful combination with the tail series.

## 7 Large $n$ and Cornish-Fisher expansions

For a distribution that is asymptotically Normal with respect to a parameter (here we consider  $n \rightarrow \infty$ ) we can make use of the Cornish-Fisher (“CF”) expansion. Indeed, this can be generalized to non-Normal target distributions but here we explicitly consider the purely Normal case. Results for the basic Cornish-Fisher expansion are of course well known and are quoted in sections 26.2.49-51 of Abramowitz and Stegun [1, 2], who also quote direct asymptotic expansions for the  $T$  distribution in section 26.7.5. Our purpose here is first to explain the relationship between (a), the CF expansions quoted in [1, 2]; (b), the  $T$  expansion also quoted in [1, 2]; (c) our power series quoted above. At first sight they can all be written in terms in powers of  $n^{-1}$ , but they all look different! As well as this reconciliation it may be helpful to be more explicit about the details given in [1, 2] as the CF expansion is given there rather non-explicitly in terms of a slightly unusual representation of the Hermite polynomials. Finally we need to take account of some issues raised by asymptotic expansions in the tails of the distribution.

In order to make the discussion self-contained we begin by defining the central moments and cumulants. In the introduction we already wrote down expressions for the mean (zero) and variance and noted that all the odd moments are zero. The even moments,  $\mu_k = \mathbb{E}[T^k]$  are then given by simplifying Eq. (11) and the first few are

$$\mu_2 = \frac{n}{n-2} \quad (63)$$

$$\mu_4 = \frac{1.3 n^2}{(n-2)(n-4)} \quad (64)$$

$$\mu_6 = \frac{1.3.5 n^3}{(n-2)(n-4)(n-6)} \quad (65)$$

and the form of these expressions indicates the general pattern. These moments get folded into the associated moment generating function (MGF)

$$\phi(t) = 1 + \frac{1}{2!}t^2\mu_2 + \frac{1}{4!}t^4\mu_4 + \frac{1}{6!}t^6\mu_6 + \dots \quad (66)$$

The associated *cumulant* generating function is given by the series expansion of the log of the MGF:

$$\log \phi(t) = \sum_{m=0}^{\infty} \frac{1}{m!} \kappa_m t^m \quad (67)$$

and we can deduce quickly that

$$\begin{aligned} \kappa_2 &= \mu_2 \\ \kappa_4 &= \mu_4 - 3\mu_2^2 \\ \kappa_6 &= \mu_6 - 15\mu_2\mu_4 + 30\mu_2^3 \end{aligned}$$

and so on. For the first terms of the CF expansion we need the quantities

$$\gamma_2 = \frac{\kappa_4}{\kappa_2^2} = \frac{\mu_4}{\mu_2^2} - 3 = \frac{6}{(n-4)} \quad (68)$$

$$\gamma_4 = \frac{\kappa_6}{\kappa_2^3} = \frac{\mu_6}{\mu_2^3} - 15\frac{\mu_4}{\mu_2^2} + 30 = \frac{240}{(n-4)(n-6)} \quad (69)$$

and so on.

For a distribution associated with a random variable  $S$  that is asymptotically Normal, and with zero mean and *unit* variance, and with vanishing odd moments, the CF expansion takes the simplified form [1]

$$s = z + [\gamma_2 h_2(z)] + [\gamma_4 h_4(z) + \gamma_2^2 h_{22}(z)] + \dots \quad (70)$$

where  $z$  is a standard Normal variable, the  $\gamma_i$  are as above, and

$$\begin{aligned} h_2(z) &= \frac{1}{24} He_3(z) = \frac{1}{24} z(z^2 - 3) \\ h_4(z) &= \frac{1}{720} He_5(z) = \frac{1}{720} z(z^4 - 10z^2 + 15) \\ h_{22}(z) &= -\frac{1}{384} (3He_5(z) + 6He_3(z) + 2He_1(z)) = -\frac{1}{384} z(3z^4 - 24z^2 + 29) \end{aligned}$$

defines the first few terms in the expansion in terms of Hermite polynomials  $He_n(z)$ . These are related to the standard Hermite “H” functions by  $He_n(z) = 2^{-(n/2)} H(z/\sqrt{2})$ .

We can now write down the Cornish-Fisher expansion for our case of interest (where we work with a unit variance variable). To the order we have calculated, it becomes

$$s = z + \frac{(z^2 - 3)z}{4(n-4)} + \frac{(z^4 - 10z^2 + 15)z}{3((n-4)(n-6))} - \frac{3(3z^4 - 24z^2 + 29)z}{32(n-4)^2} + \dots \quad (71)$$

We should now expand this in inverse powers of  $n$  to get the right asymptotic result:

$$s = z + \frac{1}{4n} z(z^2 - 3) + \frac{1}{96n^2} z(5z^4 - 8z^2 - 69) + \dots \quad (72)$$

Note carefully what we have calculated: this is the asymptotic relationship between a Normal variable  $z$  and a T-like variable  $s$  that has a T distribution scaled to *unit* variance. To get the asymptotic relationship between a Normal variable  $z$  and a variable  $t$  that has a T distribution with variance  $n/(n-2)$  we need to multiply this last asymptotic expansion by the expansion of the standard deviation:

$$\sqrt{\frac{n}{n-2}} = \sqrt{\frac{1}{1-2/n}} = 1 + \frac{1}{n} + \frac{3}{2n^2} + \dots \quad (73)$$

and this gives us the desired asymptotic series for a T-distributed variable  $t$  in terms of a Normal variable  $z$ :

$$t = z + \frac{1}{4n} z(z^2 + 1) + \frac{1}{96n^2} z(5z^4 + 16z^2 + 3) + \dots \quad (74)$$



This can now be recognized as the first three terms of the expression given in section 26.7.5 of [1, 2], which goes to order  $n^{-4}$ :

$$\begin{aligned}
 t = z + \frac{1}{4n}z(z^2 + 1) + \frac{1}{96n^2}z(5z^4 + 16z^2 + 3) + \frac{1}{384n^3}z(3z^7 + 19z^5 + 17z^3 - 15z) \\
 + \frac{1}{92160n^4}z(79z^9 + 776z^7 + 1482z^5 - 1920z^3 - 945z) + \dots
 \end{aligned}
 \tag{75}$$

However, in practice it is the corresponding formula for  $s$  that is likely to be more useful as we can directly multiply this series by the standard deviation  $\Sigma$  we wish to use, and then add back the appropriate mean parameter  $m$ . Borrowing the above high order form from [1] and taking out the series expansion of the standard deviation gives us the unit-variance expansion:

$$\begin{aligned}
 s = z + \frac{1}{4n}z(z^2 - 3) + \frac{1}{96n^2}z(5z^4 - 8z^2 - 69) + \frac{1}{384n^3}z(3z^6 - z^4 - 95z^2 - 267) \\
 + \frac{1}{92160n^4}z(79z^8 + 56z^6 - 5478z^4 - 25200z^2 - 67905) + \dots
 \end{aligned}
 \tag{76}$$

Whichever representation is to be used, we note that these expansions suggest for large  $n$  that we merely need to sample a Normal distribution, for example by a good approximation to  $N^{-1}$  applied to a uniform distribution, and then “stretch” the sample by these asymptotic formulae, that are just simple polynomials. In other words, we build  $F_n^{-1}(u)$  as

$$u \rightarrow z = N^{-1}(u) \rightarrow s \text{ or } t.
 \tag{77}$$

In practice, how well does this work? Armed with a good implementation of the exact result for all  $n$  and of  $N^{-1}$  via the inverse error function we can plot the errors with ease. It turns out that the errors are small except in the tails. In fact, no matter how large  $n$  is, the asymptotic series does eventually draw away from the exact solution. The effect is mitigated by taking more powers of  $n^{-1}$ , in that the problematic region is confined more to the far tail. The effects are shown in Figure 3. Note that these are drawn using a high-precision formula for  $N^{-1}$  based on the arbitrary precision implementation of the inverse error function in Mathematica [25]. If one use an approximation that is poor in the tails matters will be much worse.

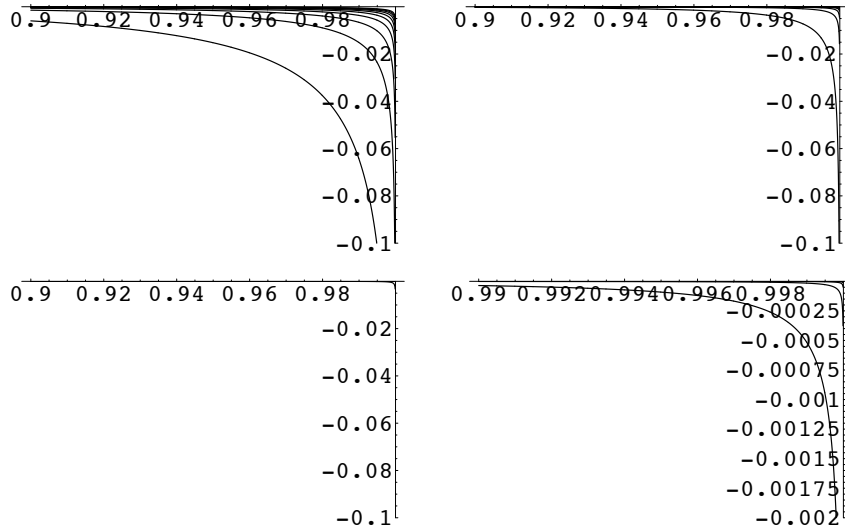


Figure 3: Errors in Large  $n$  expansion for  $n = 10, 20, \dots, 100$  with (top left), terms to  $n^{-1}$ ; (top right), terms to  $n^{-2}$ ; (bottom left), terms to  $n^{-4}$ ; (bottom right)  $n^{-4}$  expanded plot.

What should we take from this? Clearly, it is desirable to use the fourth order result. The error in the CDF for  $n = 10$  becomes or order  $10^{-3}$  as we pass through the 99.9 per cent quantile, and improves as  $n$  increases so this might be considered acceptable by some. One could also take the view that we introduced

the use of the T distribution *precisely* so we could get power-law behavior in the tails, so the fact the far-tail misbehaves with these asymptotic expansions might be deemed unacceptable. One could also take the view that one wants power-law behavior for a while but that it should eventually die off faster. Within this framework there is no difficulty with using such asymptotic results for  $n > 10$ .

So to summarize, these asymptotic results based on Cornish-Fisher expansions are good for larger  $n$  except in the far tail. Care needs to be taken to scale for the appropriate variance. The  $N^{-1}$  used needs to be good in the tails otherwise the tail errors will be made worse still.

How are these asymptotic results related to our power series, where we have *exact* values for the coefficients of powers of  $u - \frac{1}{2}$ ? This is actually a rather messy calculation. To match up the series we have to take the asymptotic results discussed here (i.e. the results from [1]) and expand  $z$  in terms of  $u - \frac{1}{2}$ . Then we must take the power series coefficients and correct them by the expansion for  $v$  in inverse powers of  $n$ . The relevant scaling is given by

$$v = (u - 1/2)\sqrt{n\pi} \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n+1}{2})} = (u - 1/2)\sqrt{2\pi} \left( 1 + \frac{1}{4n} + \frac{1}{32} \left( \frac{1}{n} \right)^2 - \frac{5}{128n^3} - \frac{21}{2048n^4} + \dots \right) \quad (78)$$

The detailed calculations are laborious and not given.

## 8 Case studies

In order to understand the methods we have presented a couple of examples will be presented. Note that there is now nothing special about the use of integer  $n$  - we pick  $n = 3, 11$  as examples of small and “modest”  $n$ . In the examples that we consider only the series as far as given explicitly in this paper will be used. The on-line supplement allows many more terms to be generated with correspondingly better accuracy, and a detailed study of the errors for a high order combination of central and tail power series and CF expansions will be given in Section 8.3.

### 8.1 $n = 3$ revisited

Prior to the development of both our power series, the case  $n = 3$  had been left in a slightly unsatisfactory state. Given that we had exact and simple solutions for  $n = 2, 4$  this needs to be sorted out! The power series about  $u = 1/2$  is given by

$$x = v \left( 1 + \frac{2v^2}{9} + \frac{11v^4}{135} + \frac{292v^6}{8505} + \frac{3548v^8}{229635} + \frac{273766v^{10}}{37889775} + \frac{15360178v^{12}}{4433103675} + \frac{214706776v^{14}}{126947968875} \right. \\ \left. + \frac{59574521252v^{16}}{71217810538875} + \frac{15270220299064v^{18}}{36534736806442875} + O(v^{20}) \right)$$

where

$$v = \frac{\sqrt{3}}{2} \pi \left( u - \frac{1}{2} \right) = 2.720699046 \left( u - \frac{1}{2} \right) \quad (79)$$

The corresponding tail series truncated at six terms is given by

$$x = \sqrt{3} \left( \sqrt{3}w \right)^{-1/3} \left( 1 + \sum_{k=1}^6 \left( \sqrt{3}w \right)^{\frac{2k}{3}} d(k) \right) \quad (80)$$

where

$$w = \frac{\sqrt{3}}{2} \pi (1 - u) = 2.720699046 (1 - u) \quad (81)$$

and the vector of coefficients  $d(k)$  is given by the list

$$\left\{ -\frac{2}{5}, -\frac{9}{175}, -\frac{92}{7875}, -\frac{1894}{606375}, -\frac{19758}{21896875}, -\frac{2418092}{8868234375} \right\} \quad (82)$$

We now take a look at the results, using the method based on the inverse beta function as our benchmark. In Figure 4 we show the absolute errors associated with the power series and tail series based on just 9 and 6 terms in the power and tail series. It is quite clear that acceptable results for many purposes can be obtained with a crossover at about  $u = 0.84$ , when the absolute error is  $O(10^{-4})$ . These results can be improved by

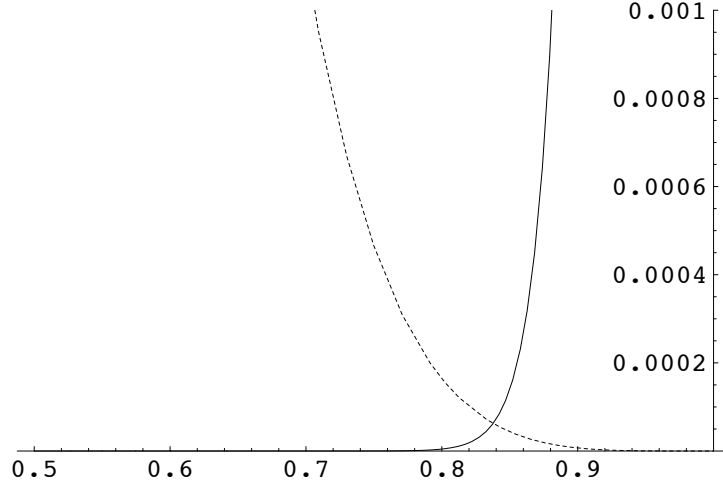


Figure 4: Absolute Errors in 9-term central (solid) and 6-term tail (dashed)  $n = 3$  series.

taking more terms or perhaps refining by applying the cobwebbing method for  $n = 3$  discussed previously.

### 8.2 $n = 11$ - a case of “modest” $n$

We repeat the above analysis with  $n \rightarrow 11$ . So for the power series

$$v = \frac{63\sqrt{11}}{256}\pi(u - \frac{1}{2}) = 2.564169909(u - \frac{1}{2}) \tag{83}$$

$$x = v \left( 1 + \frac{2v^2}{11} + \frac{39v^4}{605} + \frac{184v^6}{6655} + \frac{951v^8}{73205} + \frac{285216v^{10}}{44289025} + \frac{20943909v^{12}}{6333330575} + \frac{606462424v^{14}}{348333181625} + \frac{4679034804v^{16}}{5010638843375} + \frac{6917399415188v^{18}}{13613905737449875} + O(v^{20}) \right)$$

The tail series is now

$$w = \frac{63\sqrt{11}}{256}\pi(1 - u) = 2.564169909(1 - u) \tag{84}$$

and then

$$x = \sqrt{11} \left( \sqrt{11}w \right)^{-1/11} \left( 1 + \sum_{k=1}^6 (\sqrt{11}w)^{\frac{2k}{11}} d(k) \right) \tag{85}$$

where the vector of coefficients  $d(k)$  is given by the list

$$\left\{ -\frac{6}{13}, -\frac{77}{845}, -\frac{6424}{186745}, -\frac{3657753}{230630075}, -\frac{4839824}{599638195}, -\frac{331986068799}{76199023629625} \right\} \tag{86}$$

The results for the errors in the series and the tail are shown in Figure 5 and indicate a cross-over at about 0.94. This *is* a case where more terms might be desirable. Alternatively, let’s revisit the Cornish-Fisher expansion. With  $n = 11$ , we plot in Figure 6 the absolute error in the fourth-order Cornish-Fisher expansion (solid line) in the region  $0.995 < u < 1$ , together with the absolute error (dashed line). The range of the plot is capped at 0.005. It is quite clear that the CF method starts to go wrong in this last half percentile

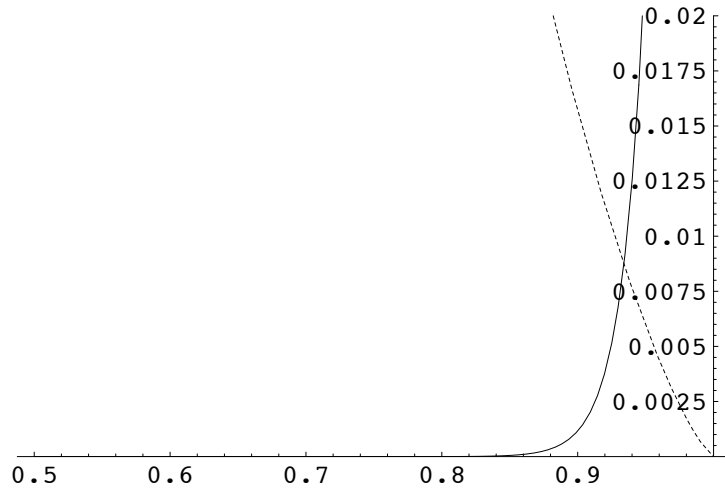


Figure 5: Absolute Errors in 9-term central (solid) and 6-term tail (dashed)  $n = 11$  series.

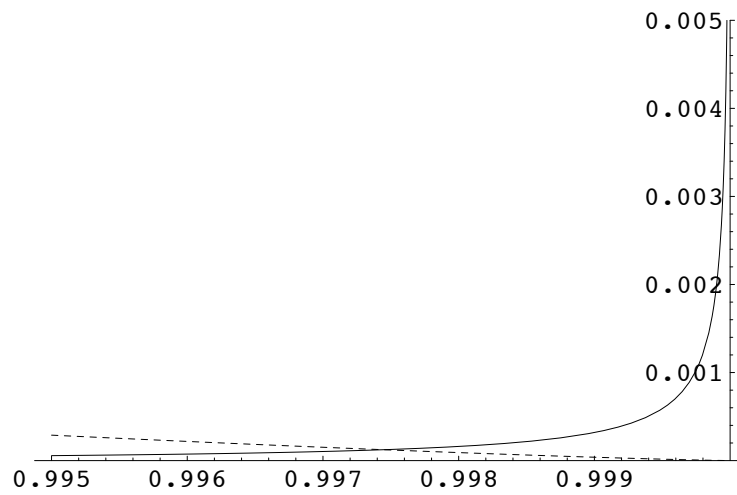


Figure 6: Absolute Errors in CF (solid) and tail (dashed) series for  $n = 11$  and last half percent.

- the tails do go wrong. We should also be clear about the nature of the effect. Rather like the Gibbs effect in Fourier Analysis, the problem never really goes away. Rather, it just moves to the edges of the interval. A careful calculation shows that the error in the Cornish-Fisher fourth order expansion is about 3 at  $u = 1 - 10^{-13}$ . It is a matter of judgement as to whether one wishes to get things that right at that level of unlikelihood!

### 8.3 Error analysis and crossover

Here we look carefully at the errors in the various approaches we have investigated, grouped by method. In all cases the benchmark is the inverse beta function solution for the iCDF and its implementation in Mathematica.

#### 8.3.1 Errors for exact solutions

In the special cases  $n = 1, 2, 4$  where we have an exact analytical result, the errors are mathematically zero but in practice are given by the machine-precision errors arising from the use of the trigonometric and square root functions employed. In practice these can be ignored.

#### 8.3.2 Errors for Newton-Raphson methods

As discussed in Section 4.2, these were found to be less than  $10^{-11}$ , based on a comparison with an implementation in Mathematica of both the iterative methods and the inverse beta function also available in Mathematica. One can trade error for speed by having a looser termination criteria for the Newton-Raphson iteration.

#### 8.3.3 Errors and crossover for combined central and tail series approach

The fine details of the errors and the resulting optimal crossover point between the central and tail power series are quite subtle, in that they will depend on how many terms are taken in each series. For serious applications the following table may be useful, and is based on a 24-term central power series and a 12-term tail series. In each case the absolute errors are measured by reference to the exact inverse  $\beta$ -function representation, computed using high-precision routines in Mathematica. Sensible interpolation can be applied for real intermediate values. For each value of  $n$  we give the crossover (the point at which the absolute errors of the central and tail series coincide and are at a maximum), the absolute error at the crossover, and the value of the iCDF at the crossover.

| $n$ | crossover | max error            | iCDF(co) |
|-----|-----------|----------------------|----------|
| 1   | 0.750     | $< 10^{-14}$         | 1.0      |
| 1.5 | 0.794     | $< 10^{-12}$         | 1.115    |
| 2   | 0.823     | $6 \times 10^{-11}$  | 1.197    |
| 3   | 0.860     | $8 \times 10^{-9}$   | 1.315    |
| 4   | 0.883     | $1.4 \times 10^{-7}$ | 1.400    |
| 5   | 0.8985    | $1.0 \times 10^{-6}$ | 1.464    |
| 6   | 0.9103    | $4 \times 10^{-6}$   | 1.520    |
| 7   | 0.9195    | $1.2 \times 10^{-5}$ | 1.567    |
| 8   | 0.9268    | $2.7 \times 10^{-5}$ | 1.567    |
| 9   | 0.9328    | $5.5 \times 10^{-5}$ | 1.645    |
| 10  | 0.9379    | $\sim 10^{-4}$       | 1.678    |
| 15  | 0.9548    | $7.5 \times 10^{-4}$ | 1.810    |
| 20  | 0.9644    | $2.4 \times 10^{-3}$ | 1.905    |

This table shows that the combination of the central and tail power series can be used very effectively for the low values of  $n$  that appear to be most relevant to mathematical finance applications. This combination of methods becomes rather less accurate as  $n$  increases through 10 and is probably not appropriate for  $n > 20$ , depending on accuracy criteria.

### 8.3.4 Errors and crossover for combined Cornish-Fisher expansion and tail series approach

The last combination of methods for which we analysis the errors is the use of a fourth-order Cornish-Fisher expansion combined with the tail power series as before. In assessing this approach, we emphasize that we are using a high-precision implementation of the iCDF for the Normal distribution, in the form of Mathematica’s inverse error function. One must also consider errors introduced by the use of a further approximation of the Normal iCDF in any given implementation. In the following table we start at  $n = 5$  and work up.

| $n$ | crossover   | max error             | iCDF(co) |
|-----|-------------|-----------------------|----------|
| 5   | 0.829       | $1.5 \times 10^{-5}$  | 1.050    |
| 6   | 0.8913      | $9 \times 10^{-6}$    | 1.378    |
| 7   | 0.9286      | $7.2 \times 10^{-6}$  | 1.651    |
| 8   | 0.9511      | $7.0 \times 10^{-6}$  | 1.874    |
| 9   | 0.9656      | $7.5 \times 10^{-6}$  | 2.066    |
| 10  | 0.9755      | $8.2 \times 10^{-6}$  | 2.240    |
| 15  | 0.99523     | $1.25 \times 10^{-5}$ | 2.97     |
| 20  | 0.999051    | $1.62 \times 10^{-5}$ | 3.574    |
| 30  | 0.999961    | $2.24 \times 10^{-5}$ | 4.57     |
| 40  | 0.9999984   | $2.8 \times 10^{-5}$  | 5.408    |
| 50  | 0.99999993  | $3.2 \times 10^{-5}$  | 6.1248   |
| 60  | 0.999999997 | $4.0 \times 10^{-5}$  | 6.777    |

This interesting table reminds us that no matter how large the value of  $n$ , the CF expansion eventually breaks down in the tails. Nevertheless, it also suggests says that for  $n > 60$  one might consider using the CF expansion everywhere, unless one is using very large sample sizes, since the tail region where the CF expansion breaks down is unlikely to be probed. It also suggest that the double power series method should be switched to the CF-tail series method for  $n \gtrsim 7$ .

These analyses support the view that between the various methods, we have good accuracy over a wide range of  $n$ .

## 9 A simple benchmark calculation with T marginals

In order to provide an implementation benchmark, we give a very simple example that can be computed very quickly. The example chosen has the merit that although it is not completely trivial it still has a semi-analytical solution for the zero-correlation case, so we have some check on the calculation as well. We will also be able to illustrate the use, for the correlated case, of both a Normal copula with T marginals (with a variation using a Cornish-Fisher expansion) and a Frank 2-copula with T marginals, illustrating the freedom afforded by having explicit functions for the iCDF. We consider two assets  $S_i$ ,  $i = 1, 2$ , with zero risk-neutral drift whose terminal distribution at a future time  $T$  is given by

$$S_i(T) = S_i(0) \exp\{\sqrt{T}\sigma_i X_i\} \quad (87)$$

where  $X_i$  both have a zero mean, *unit variance* T distribution with degrees of freedom  $n_i$ . The contract to be priced has a payoff at time  $T$  that is to be some function of the maximum of the asset returns from time zero to time  $T$ , i.e., a function of:

$$M_T = \text{Max}[\exp\{\sqrt{T}\sigma_1 X_1\}, \exp\{\sqrt{T}\sigma_2 X_2\}] \quad (88)$$

To keep the number of parameters down and focus purely on the distributional effects we shall set  $\sqrt{T}\sigma_i = 1$ . The maximum at  $T$  is then just

$$M_T = \text{Max}[\exp\{X_1\}, \exp\{X_2\}] = \exp\{\text{Max}[X_1, X_2]\} \quad (89)$$

So to keep matters simple, we shall focus on the computation of the valuation of a contract that is the log of the maximum, whose payoff is:

$$P_T = \log M_T = \text{Max}[X_1, X_2] \quad (90)$$

Although this may seem rather a contrived example, the construction of payoffs from such *order statistics* is in general a common financial calculation, and in general we may be interested in, e.g. the  $k$ th of  $m$  sorted values or other quantities associated with financial entities. With any such contract, it is easy to construct an analytical formula for  $\mathbb{E}[P_T]$  in the case where the components are *independent*. We do this explicitly for our two-dimensional example. First note if  $X_1$  has CDF  $G(x)$  and  $X_2$  CDF  $H(x)$  then the CDF,  $F(x)$  of  $P_T$  is, assuming independence.

$$F(x) = P(\text{Max}[X_1, X_2] \leq x) = P(X_1 \leq x \cap X_2 \leq x) = P(X_1 \leq x)P(X_2 \leq x) = G(x)H(x) \quad (91)$$

Second, the expectation of  $P_T$  can be written entirely in terms of its distribution function  $F(x)$  as

$$\begin{aligned} \mathbb{E}[P_T] &= \int_{-\infty}^{\infty} yf(y)dy = \int_0^{\infty} -y(1 - F(y))'dy + \int_{-\infty}^0 yF'(y)dy \\ &= \int_0^{\infty} (1 - F(y))dy - \int_{-\infty}^0 F(y)dy \end{aligned} \quad (92)$$

where the last step is a simple integration by parts, assuming good behaviour of  $F$  at  $\pm\infty$ . Now if we combine Eq. (91) and Eq. (92), and further note that  $G(-y) = 1 - G(y)$ , similarly for  $H$ , we are lead after some simplification to the desired result, that

$$\mathbb{E}[P_T] = \int_0^{\infty} [G(y) + H(y) - 2G(y)H(y)]dy \quad (93)$$

Finally, given that  $X_i \sim \sqrt{\frac{n_i-2}{n_i}}T_i$ , we have

$$G(y) = F_{n_1}\left(y\sqrt{\frac{n_1}{n_1-2}}\right) \quad H(y) = F_{n_2}\left(y\sqrt{\frac{n_2}{n_2-2}}\right) \quad (94)$$

It is rather amusing to note that the integral in Eq. (93), when combined with the assumptions of Eq. (94), can often be done in closed form in terms of standard elliptic  $E$  and  $K$  functions, and further simplifies to a multiple of  $\pi$  when  $n_1 = n_2$ . We have the following table of results from exact integration within Mathematica, for the zero-correlation results:

| $n_i$ | exact integral  | numerical value |
|-------|---|-----------------|
| 4, 4  | $\frac{15\pi}{64\sqrt{2}}$                              | 0.520650        |
| 6, 6  | $\frac{2835\pi}{16384}$                                 | 0.543604        |
| 8, 8  | $\frac{75075\sqrt{3}\pi}{524288\sqrt{2}}$               | 0.550961        |
| 4, 6  | $\frac{1}{8}(21E(1/2) - 13K(1/2))$                      | 0.532569        |
| 4, 8  | $\frac{1}{128}\sqrt{\frac{3}{2}}(178E(2/3) - 83K(2/3))$ | 0.536663        |
| 6, 8  | NA  | 0.547352        |

These results are perhaps slightly surprising in that the trend is for the expected value of the maximum to decrease as the distribution gets more fat-tailed. We need to note that we are rescaling to ensure that the distributions have unit variance always, even as the tails decay more slowly. These results may be useful in testing any implementation of a method for sampling from the  $T$ , and we now look at some of the methods we have discussed.

## 9.1 Simulated results - zero correlation

The integrals above can be calculated by Monte Carlo methods using several of the methods discussed here. In the zero correlation case there is no need to introduce a copula, so that we may make a choice to use Bailey’s method, or any of the representations of the iCDFs. First of all, picking the latter, so as to illustrate the more novel techniques developed here, we calculate Monte-Carlo estimates of the expectation in the form (note the allowance for getting the variance to unity):

$$\mathbb{E}[P_T] \sim \frac{1}{N_{MC}} \sum_{k=1}^{N_{MC}} \text{Max} \left[ \sqrt{\frac{n_1 - 2}{n_1}} F_{n_1}^{-1}(u_k), \sqrt{\frac{n_2 - 2}{n_2}} F_{n_2}^{-1}(v_k) \right] \quad (95)$$

where the  $(u_k, v_k)$  are random uniform samples from  $[0, 1]$ . For example, in the interesting case  $n_1 = n_2 = 4$ , this estimate simplifies to

$$\mathbb{E}[P_T] \sim \frac{1}{N_{MC}\sqrt{2}} \sum_{k=1}^{N_{MC}} \text{Max} \left[ F_4^{-1}(u_k), F_4^{-1}(v_k) \right] \quad (96)$$

A simulation of this with  $N_{MC} = 10^7$  pairs of uniform deviates yielded the result 0.5204 with a standard error of 0.00027, based on a compiled implementation of the exact solution of Eq. (38). So the exact and Monte Carlo solution differ by less than one standard error, which is very satisfactory.

Going back to Bailey’s method, we can work with the algorithm discussed in Section 2.1. Let us denote the result of applying the algorithm with uniform deviate inputs  $u, v$  as  $\text{Bailey}_n(u, v)$ . Then, for example, simulating the  $n_1 = n_2 = 6$  case, the corresponding Bailey Monte Carlo estimate is given by

$$\mathbb{E}[P_T] \sim \frac{\sqrt{2}}{N_{MC}\sqrt{3}} \sum_{k=1}^{N_{MC}} \text{Max} \left[ \text{Bailey}_6(u_{k,1}, v_{k,1}), \text{Bailey}_6(u_{k,2}, v_{k,2}) \right] \quad (97)$$

Using  $N_{MC} = 10^7$  quadruples of uniform deviates yielded the result of 0.54361 with a standard error of 0.00027, which is also consistent with our exact solution.

## 9.2 Simulated results - Normal copula and $T_8$ marginals

It might seem rather odd to use T marginals with a Normal copula rather than just a T copula. However, our idea is to illustrate the fact that the T iCDFs can be used with *any* copula, and also to give an example of what happens when the full machinery is replaced with a Cornish-Fisher (CF) method. We shall work with the  $n_1 = n_2 = 8$  case in order to give the CF method a hope of providing reasonable results, but will be able to see the impact, if any, of the tail error in the raw CF method. With just two assets the sampling with the Normal copula method with  $n = 8$  T *unit variance* marginals and correlation  $\rho$  can be simplified to the following sampling scheme, where the  $Z_i$  are *independent* samples from  $N(0, 1)$ , and  $N$  is the Normal CDF:

$$Y_1 = Z_1, \quad Y_2 = \rho Z_1 + \sqrt{1 - \rho^2} Z_2 \quad (98)$$

$$X_1 = \frac{\sqrt{3}}{2} F_8^{-1}[N(Y_1)], \quad X_2 = \frac{\sqrt{3}}{2} F_8^{-1}[N(Y_2)] \quad (99)$$

We do the calculation first (the “Full version”) with (a) polar sampling of the  $Z_i$ , (b) a high-precision implementation of  $N$  in Mathematica, (c) our polynomial Newton-Raphson implementation of  $F_8^{-1}$ . Then we shall do the CF approximation, where the function  $\frac{\sqrt{3}}{2} F_8^{-1}(N(z))$  is approximated by the unit variance CF expansion of Eq. (76) with  $n = 8$ . We shall use samples of  $\rho$  between  $-1$  and  $+1$  in steps of 0.25, and use the same random seed for the full simulation and the CF approximation.

Note that we expect to recover the exact solution given above when  $\rho = 0$ , and in the case  $\rho = 1$  expect to get zero, since the  $X_i$  are then identical and have zero expectation. The case  $\rho = -1$  can also be calculated analytically, since in this case the simulated variables will be pairs of samples of opposite sign, so that the maximum is just the absolute value. The value of the expectation should therefore be  $\sqrt{3}/2$  times  $\mathbb{E}[|T|]$  with  $n = 8$ . From Eq. (11) we can work out that the final answer should be  $5/8\sqrt{3}/2 \sim 0.7655$ . In the following table the results are shown as a function of  $\rho$  with  $10^6$  estimates. The maximum standard error over all cases is about 0.001. Values comparable with exact solutions are shown in bold.



| $\rho$      | Full version   | CF method      |
|-------------|----------------|----------------|
| -1.0        | <b>0.7658</b>  | <b>0.7659</b>  |
| -0.75       | 0.7197         | 0.7198         |
| -0.5        | 0.6700         | 0.6701         |
| -0.25       | 0.6115         | 0.6116         |
| <b>0.0</b>  | <b>0.5490</b>  | <b>0.5491</b>  |
| +0.25       | 0.4789         | 0.4789         |
| +0.5        | 0.3934         | 0.3934         |
| +0.75       | 0.2797         | 0.2797         |
| <b>+1.0</b> | <b>-0.0015</b> | <b>-0.0015</b> |

This table indicates that all is behaving as expected, and gives us confidence in the simulation methods. We also note that for this particular example, the “maximum product” is rather more prone to correlation risk than distributional risk. In particular, as expected, we get rather higher values for strongly negatively correlated assets.

### 9.3 Simulated results - Frank 2-copula and $T_8$ marginals

Finally, in order to assess “copula risk” we reprice this log-maximum once more using the Frank 2-copula with parameter  $\alpha$ , and marginals with  $n = 8$  as before. This illustrates the flexibility in the choice of copula when one has the iCDF for the marginals. The Frank  $m$ -copula is discussed in detail in [4], where methods for the estimation of the parameter  $\alpha$  are given in Section 5.3.1, and the use of such a copula with major indices is also argued for in Section 2.3.4 of [4].

Another reason for picking the Frank copula for study is that with just two assets we can take  $\alpha$  to have either sign and furthermore a simple explicit formula can be given for the correlated samples from the uniform distribution. We shall regard  $\alpha$  just as some form of proxy for correlation. In [4] it is also shown that the general conditional sampling approach can be reduced to a simple iterative scheme. When  $m = 2$  the simulation of a pair of correlated uniform deviates under the Frank copula reduces to the following algorithm. Let  $(v_1, v_2)$  be pair of independent samples from a uniform distribution on  $[0, 1]$ . Then set  $u_1 = v_1$  and

$$u_2 = -\frac{1}{\alpha} \log \left\{ 1 + \frac{v_2(1 - e^{-\alpha})}{v_2(e^{-\alpha u_1} - 1) - e^{-\alpha u_1}} \right\} \tag{100}$$

Then the Monte Carlo estimate for the value is

$$\mathbb{E}[P_T] \sim \frac{1}{N_{MC}} \frac{\sqrt{3}}{2} \sum_{k=1}^{N_{MC}} \text{Max} \left[ F_8^{-1}(u_{1,k}), F_8^{-1}(u_{2,k}) \right] \tag{101}$$

In the following table we present the results from this formula with  $-12 \leq \alpha \leq +12$  in steps of 4, with  $N_{MC} = 10^6$  samples. The maximum standard error is again about 0.001.

| $\alpha$ | Frank value   |
|----------|---------------|
| -12      | 0.7653        |
| -8       | 0.7374        |
| -4       | 0.6878        |
| <b>0</b> | <b>0.5510</b> |
| +4       | 0.3753        |
| +8       | 0.2654        |
| +12      | 0.2057        |

Again we achieve plausible results in the appropriate range. This simple example is consistent with the view that the choice of parameters associated with any given copula is at least as important as the choice of copula itself, or indeed the marginals. This toy contract is indeed a correlation-dominated entity, however the correlation is defined.

## 10 Conclusions and further work

We have explored the inverse CDF (iCDF) for the Student T distribution and presented the following:

- A clear description of the iCDF in terms of inverse  $\beta$ -functions, suitable for benchmark one-off computations;
- Exact solutions for the iCDF in terms of elementary functions for  $n = 1, 2, 4$ , which are themselves of interest to “fat-tailed finance” applications;
- Fast iterative Newton-Raphson techniques the iCDF for even integer  $n$ , with details for  $n \leq 20$ ;
- A power series for the iCDF valid for general real  $n$  accurate except in the tails;
- A generalized power series for the tails that is good for low to modest  $n$ ;
- A summary of known results on the Cornish-Fisher expansions valid for modest to infinite  $n$ ;
- The limitations of Cornish-Fisher expansions in the far tails, which is where the power-law behavior should exist and will fail with CF;
- An example of using the iCDFs to price a simple contract under various assumptions for the correlation structure.

Between them these results allow either slow and very precise or fast and reasonably accurate methods for the iCDF for all  $n$  and  $u$ . Although this is something of a patch-work of methods the best methods would appear to be:

- If  $n$  is a low even integer use one of the exact or iterative polynomial methods developed here;
- If  $n$  is real and less than about 7 to use the power series and tail series developed here;
- If  $n$  is real and greater than about 7 to use the known Cornish-Fisher expansion given in [1, 2], with the generalized power series for the tail developed here above the crossover point until  $n \sim 60$ , at which point, except for very large simulations, the CF method alone will suffice.

The author is emphatically not claiming that these suggestions are the last word on the matter - indeed it is hoped that the methods shown here stimulate discussion and improvements. In practice if it is a matter of just having indicative results for a variety of fat-tailed distributions with a finite variance, the exact or iterative solutions for  $n = 4, 6, 8, 10$  may often suffice. Applications to “high-frequency finance” requiring a specific value of  $n \leq 7$  are well catered for by the pair of power series. If one does not need to use the iCDF at all, then Bailey’s method will suffice.

It should also be clear that the power series methods employed here can be applied to any pdf that can be characterized by a series in neighborhoods of  $u = 1/2$  and  $u = 0, 1$ . A novel feature of the analysis given here is the use of symbolic computation to do the nasty inversion of a general power series, term by term, that would otherwise be intractable beyond a handful of terms. This is easily generalized. A case of interest would be a generalized skew-T distribution with a pdf

$$f_{m,n,\alpha}(x) = f_m(x)F_n(\alpha x). \quad (102)$$

The central power series for this can clearly be computed - further work on this case will be reported elsewhere.

## A guide to the on-line supplements

This paper is supported by various material downloadable from the author’s web site in the directory

[www.mth.kcl.ac.uk/~shaww/web\\_page/papers/Tsupp/](http://www.mth.kcl.ac.uk/~shaww/web_page/papers/Tsupp/)

and there are four documents to download at the time of finalizing this paper:

1. A Mathematica notebook showing how many of the calculations were done and the graphics generated. Note that much of what is in this file can be regarded as pseudo-code for other languages and there are some sections specifically for C/C++ applications. The file is at  
[www.mth.kcl.ac.uk/~shaww/web\\_page/papers/Tsupp/InverseT.nb](http://www.mth.kcl.ac.uk/~shaww/web_page/papers/Tsupp/InverseT.nb)  
and can be read using the free *MathReader* application.
2. I have also posted a PDF of the above notebook available from the link  
[www.mth.kcl.ac.uk/~shaww/web\\_page/papers/Tsupp/InverseT.pdf](http://www.mth.kcl.ac.uk/~shaww/web_page/papers/Tsupp/InverseT.pdf)
3. A lookup table of quantiles of the T distribution (i.e. values of  $F_n^{-1}(u)$ ) for  $1 \leq n \leq 25$  in steps of 0.1 is provided in CSV form at:  
[www.mth.kcl.ac.uk/~shaww/web\\_page/papers/Tsupp/tquantiles.csv](http://www.mth.kcl.ac.uk/~shaww/web_page/papers/Tsupp/tquantiles.csv)
4. A note on the Excel<sup>TM</sup> spreadsheet function TINV, including its limitations and how to make sense of it is available at  
[www.mth.kcl.ac.uk/~shaww/web\\_page/papers/Tsupp/TINV.pdf](http://www.mth.kcl.ac.uk/~shaww/web_page/papers/Tsupp/TINV.pdf)

## Acknowledgements

This work was originally stimulated by Prof. P. Embrecht’s 2004 visit to Oxford to deliver the 2004 Nomura lecture. My understanding of copula methods has benefited greatly from a lecture given by Aytac Ilhan, and I am grateful to Walter Vecchiato for his help on references on this matter. I also wish to thank the Editor and anonymous referees for helpful comments on the initial version of this paper.

## 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  
<http://www.math.sfu.ca/~cbm/aands/>
- [3] R.W. BAILEY *Polar generation of random variates with the t-distribution*. Mathematics of Computation, 62, 206, 779-781, 1994.
- [4] U. CHERUBINI, E. LUCIANO, W. VECCHIATO *Copula methods in finance*, Wiley, 2004.
- [5] L. DEVROYE *Non-uniform random variate generation*, Springer 1986. Out of print - now available on-line from the author’s web site at  
<http://cgm.cs.mcgill.ca/~luc/rnbookindex.html>
- [6] D. DUFFIE, *Default Correlation, Lecture 3, Clarendon Lectures in Finance*, Oxford June 2004. Available from the link at:  
[http://www.finance.ox.ac.uk/clarendon\\_lectures2004.htm](http://www.finance.ox.ac.uk/clarendon_lectures2004.htm)
- [7] E.F. FAMA, *The behavior of stock prices*, *Journal of Business* 37, 34-105, 1965.
- [8] H.-B. FANG, K.-T. FANG, S. KOTZ, *The meta-elliptical distribution with given marginals*, *Journal of Multivariate Analysis*, 82, 1-16, 2002.
- [9] K. FERGUSON, E. PLATEN, *On the Distributional Characterization of daily Log-returns of a World Stock Index*, *Applied Mathematical Finance*, 13 (1), 19-38, March 2006.
- [10] R. GENÇAY, M. DACORONGA, U. MULLER, R. OLSEN, O. PICTET, *An Introduction to High-frequency Finance*, Academic Press, 2001.

- [11] A. GENZ, F. BRETZ, Y. HOCHBERG, *Approximations to multivariate  $t$  integrals with applications to multiple comparison procedures*. In *Recent Developments in Multiple Comparison Procedures*, Institute of Mathematical Statistics LNMS 47, 24-32, 2004.
- [12] G. W. HILL, *Algorithm 396, Student’s  $t$ -Quantiles*, *Communications of the ACM*, 13(10): 619-620, Oct., 1970.
- [13] G. W. HILL, *Algorithm 396, Student’s  $t$ -Quantiles*, on-line version at <http://portal.acm.org/citation.cfm?id=355600>
- [14] M.C. JONES, *Student’s simplest distribution*, *The Statistician (Journal of the Royal Statistical Society Series D)*, 51, 1, 41-49.
- [15] S. KOTZ AND S. NADARAJAH. *Multivariate  $t$  distributions and their applications*, Cambridge, 2004.
- [16] B. MANDELBROT, *The variation of certain speculative prices*, *Journal of Business* 36, 394-419, 1963.
- [17] D. MENEGUZZO, W. VECCHIATO, *Copula sensitivity of collateralized debt obligations and basket default swaps pricing and risk monitoring*, *Journal of Futures Markets*, 24, 1, 37-70.
- [18] B. MORO, *The Full Monte*, *Risk*, 8(2): 57-58, Feb., 1995.
- [19] V.B. NEVZOROV, N. BALAKRISHNAN, M. AHSANULLAH, *The Statistician (Journal of the Royal Statistical Society Series D)*, 52, 3, 395-400, 2003.
- [20] W.H. PRESS, S.A. TEUKOLSKY, W.T. VETTERLING, B.P. FLANNERY, *Numerical Recipes in C++*, Cambridge, 2002.
- [21] W. T. SHAW, *Complex Analysis with Mathematica*, Cambridge University Press, 2006.
- [22] D. STIRZAKER, *Elementary Probability*, Cambridge University Press, 1994.
- [23] STUDENT (A.K.A. WILLIAM GOSSET), *The Probable Error of a Mean*. *Biometrika* 6, 1-25, 1908.
- [24] Y. TONG, *The multivariate normal distribution*, Springer, 1990.
- [25] S. WOLFRAM, *The Mathematica Book, 5th ed*, Wolfram Media, 2004.