NEW METHODS FOR SIMULATING THE STUDENT T-DISTRIBUTION - DIRECT USE OF THE INVERSE CUMULATIVE DISTRIBUTION

WILLIAM T. SHAW*

Abstract.
We explore the Student T-Distribution and present some new techniques for simulation. In particular, an explicit and accurate approximation for the inverse, F_n^{-1} of the CDF, F_n, is presented, as well as some simple exact and iterative techniques for defining this function. The methods presented are particularly well adapted to be used with copula and quasi-Monte-Carlo techniques.

Key words. Student, Distribution, T-distribution, Simulation, Monte Carlo, Inverse Cumulative Distribution Function, Inverse CDF.

1. Introduction. 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

$$\chi_n^2 = Z_1^2 + \cdots + Z_n^2$$

The density function of χ_n^2 is easily worked out, using moment generating functions, and is

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

and gives a random variable with a mean of n and a variance of $2n$. We now define a normal variable with a randomized variance in the form:

$$T = \frac{Z_0}{\sqrt{\chi_n^2/n}}$$

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

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

Then to get the *joint* density of T and χ_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 ν. We make the observation, based on a standard integral identity, that

$$\int_0^\infty f(t|\chi_n = \nu) q_n(\nu) d\nu = \frac{1}{\sqrt{n\pi}} \frac{\Gamma[\frac{n+1}{2}]}{\Gamma[\frac{n}{2}]} \frac{1}{(1 + t^2/n)^{\frac{n+1}{2}}} = f_n(t)$$

It is evident that a sample from this distribution can easily be obtained by using $n+1$ samples from the standard normal distribution. This is well known, as is the use of a normal variate divided by the square root of a scaled sample from the χ^2 distribution, that itself being obtained by other methods.

The distribution has the properties that the odd moments all vanish, provided n is large enough so that they are defined. Provided $n > 2$ the variance is well-defined and given by

$$\mathrm{Var[T]} = \mathbb{E}[T^2] = \frac{n}{n-2}$$

The fourth moment exists provided $n > 4$ and is given by

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

* Nomura Centre for Quantitative Finance, OCIAM, The Mathematical Institute, 24-29 St Giles, Oxford OX1 3LB, England. (shaww@maths.ox.ac.uk).

The fat-tailed behaviour of the distribution is characterized by the kurtosis relative to that of a normal distribution (3) by the formula

$$\frac{\mathbb{E}[T^4]}{\mathrm{Var}[T]^2} - 3 = \frac{6}{n-4}.$$

One of the ideas of this paper is to get a grip on the use of the elementary result:

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

to define a sample from the T-distribution directly, where U is uniform and F_n is the cumulative density function 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!).

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. The use of iterative and purely arithmetical schemes recalls the remark by J. von Neumann,

Anyone who considers arithmetic methods of producing random digits is, of course, in a state of sin.

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 is 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 [4] on an approximate method for $N^{-1}(u)$ (where $N(x)$ is the normal CDF), and although the methods presented here are different, the motivations are closely related. There are perhaps two schools of thought on how accurate such approximations need to be. Given all the uncertainties elsewhere in such problems, it is perhaps inappropriate to dwell too much on the number of significant figures obtained – to quote von Neumann again:

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

On the other hand one can take the view that one should at least try to eliminate uncertainty due to one's purely numerical considerations, and this author leans more to the latter view than to the former.

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 examaples 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 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 inverse CDFs for each marginal. In such an approach it is clearly helpful to have a grip on F^{-1}.

Many of the results described herein are discussed further, with sample code, in an on-line supplement. This is available as both a *Mathematica* notebook and a PDF file at
`http://www.maths.ox.ac.uk/~shaww/finpapers/InverseT.nb`
`http://www.maths.ox.ac.uk/~shaww/finpapers/InverseT.pdf`
Note that although a couple of the examples provided in this supplement rely on the special-function capabilities of *Mathematica*, most of it is devoted to supplying information and pseudo-code relevant to any implementation in any language.

2. The CDF for Student's T-distribution. The relevant CDF may be characterized in various different ways. Our universal starting point is the formula

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

To evaluate this and try to think about inversion, one of the most obvious things to do is making a trigonmetric substituation 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] and 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 represenations we start to make progress. $F_n(x)$ can be written in "closed form", albeit in terms of hypergeometric functions, for general n. The formula is just

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

This may also be thought of, in a way that makes it more accessible to more computer environments, in terms of β-functions, for we can rewrite the hypergeometric function to obtain

$$F_n[x] = \frac{1}{2}\left(1 + \text{sgn}(x)(1 - I_{(\frac{n}{x^2+n})}\left(\frac{n}{2},\frac{1}{2}\right)\right)$$

giving an expression in terms of regularized β-functions. The regularized beta function $I_x(a,b)$ employed here is given by

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

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

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

Having got such a representation, this may be formally inverted to give

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

This last result may indeed be useful in computer environments where internal representations of inverse beta functions are available. Alternatively, if the (forward) β-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 β-function then one can work directly with the formal inverse. As an example, we can use a representation in *Mathematica* [7] to visualize the inverse for various values of n. In Figure 2.1 we show the inverse 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

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

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

$$t = F_\infty^{-1}(u) = \sqrt{2}\text{erf}^{-1}(2(y - \frac{1}{2}))$$

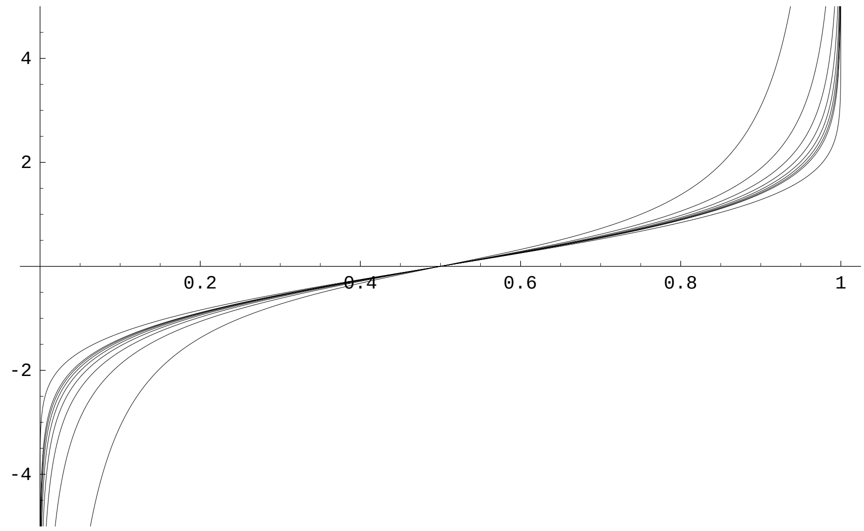

FIG. 2.1. *Inverse CDFs for the T-Distribution for $n = 1$ to 8 and $n = \infty$.*

The plots are constrained to the range $-5 \leq t \leq +5$. The plots show what we hope to see - as n decreases from infinity the image distribution becomes more fat-tailed, and the behaviour is monotone in n. The general formula for the inverse is also useful, but not that fast (cf using representations of 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.

However, this representation does not give us much insight into the structure of the function. 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 trigonmetric functions.

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	$\frac{1}{2} + \frac{x\left(x^2+6\right)}{2(x^2+4)^{3/2}}$
5	$\frac{1}{2} + \frac{1}{\pi}\tan^{-1}\left(\frac{x}{\sqrt{5}}\right) + \frac{\sqrt{5}x\left(3x^2+25\right)}{3\pi(x^2+5)^2}$
6	$\frac{1}{2} + \frac{x\left(2x^4+30x^2+135\right)}{4(x^2+6)^{5/2}}$

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.

2.1. 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:

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

where the coefficients are defined recursively by the relations

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

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

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 $0 < u < 1$:

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

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

$$\alpha = 1 - 4(u - \frac{1}{2})^2$$

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

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

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:

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

We now proceed to extract some solutions. The on-line supplement code to generate the resolvent polynomial equations for even $n \leq 20$ and exhibits them.

3. Some Simple Exact Solutions for the Inverse CDF. 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 $n = 6, 8, 10, \ldots$.

3.1. $n = 1$**.** As a simple reminder, the inverse CDF for the $n = 1$ case, the standard Cauchy distribution, is

$$x = \tan(\pi(u - \frac{1}{2}))$$

3.2. $n = 2$. This is now trivial as the resolvent polynomial is linear. After some simplification we obtain

$$x = \frac{2\sqrt{2}\left(u - \frac{1}{2}\right)}{\sqrt{1 - 4\left(u - \frac{1}{2}\right)^2}}$$

3.3. $n = 4$. 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

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

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 [5]). After some work along these lines and some simplification we obtain the solution in the form:

$$p = \frac{4}{\sqrt{\alpha}} \cos\left(\frac{1}{3}\tan^{-1}\sqrt{\left(\frac{1}{\alpha} - 1\right)}\right)$$

and where, as before,

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

3.4. $n \geq 6$. 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. 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 x^6, x^5, x^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 α[1]. 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

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

For $n = 8$ we have

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

For $n = 10$ we have

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

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. 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

[1] The author would be grateful to receive enlightenment on this matter from polynomial and Galois Theory experts!

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:

$$p_0 = 2(\frac{1}{\alpha} - 1) + n$$

and the result is extracted via

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

and $\alpha = 1 - 4(u - \frac{1}{2})^2$ as before.

3.5. Comments on the case $n = 4$. The new exact solution for F^{-1} presented above for the case $n = 4$ is easily applied to random sample from the uniform distribution to produce a simulation of the $n = 4$ distribution. 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 ∞ 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.

There is also some recent independent evidence supporting this as an interesting case for purely financial reasons. Recent work by Ferguson and Platen [6] also suggests that $n = 4$ is an accurate representation of index returns in a global setting, and propose models to underpin this idea.

General non-integer low values of n are also valuable on short time-scales. Work cited in [3] suggests that very short term returns exhibit power law decay in the pdf. For a T distribution the decay of the density is

$$O(t^{-n-1})$$

and the decay of the CDF is

$$O(t^{-n})$$

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 [3] take values in the range 2 to 6. So this leads us to consider not only small integer values of $n = 2, 3, 4, 5, 6$ but also non-integer n. The case of non-integer n is discussed later.

Some further examples of the use of the exact solutions and iteration schemes for even n are given in the on-line supplement.

3.6. 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 $n = 3, 5, 7, 9$. We have where

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

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

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

$$F_9(x) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}\left(\frac{x}{3}\right) + \frac{3x(35x^6+1155x^4+13797x^2+67797)}{35\pi\left(x^2+9\right)^4}$$

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

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

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

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

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

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

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

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

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

$$\theta_0 = \tan^{-1}\left(\frac{2\sqrt{\frac{2}{3}}\left(u - \frac{1}{2}\right)}{1 - 4\left(u - \frac{1}{2}\right)^2}\right)$$

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 behaviour 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 neigbouhood 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 defer further discussion of odd n.

4. The Power Series for the Inverse Cumulative Distribution. Having exhausted the cases that are exactly solvable in terms of elementary functions (the author would be very happy to be shown some other solvable cases!), and having dealt explictly with fast numerical schemes for low even integer n, 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 $u = 1/2$:

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

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:

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

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* [7]. The code for doing this is available in the on-line supplement. The first terms in the resulting power series are:

$$
\begin{aligned}
c_1 &= \frac{1}{6} + \frac{1}{6n} \\
c_2 &= \frac{7}{120} + \frac{1}{15n} + \frac{1}{120n^2} \\
c_3 &= \frac{127}{5040} + \frac{3}{112n} + \frac{1}{560n^2} + \frac{1}{5040n^3} \\
c_4 &= \frac{4369}{362880} + \frac{479}{45360n} - \frac{67}{60480n^2} + \frac{17}{45360n^3} + \frac{1}{362880n^4} \\
c_5 &= \frac{34807}{5702400} + \frac{153161}{39916800n} - \frac{1285}{798336n^2} + \frac{11867}{19958400n^3} - \frac{2503}{39916800n^4} + \frac{1}{39916800n^5} \\
c_6 &= \frac{20036983}{6227020800} + \frac{70691}{64864800n} - \frac{870341}{691891200n^2} + \frac{67217}{97297200n^3} - \frac{339929}{2075673600n^4} + \frac{37}{2402400n^5} \\
&\quad + \frac{1}{6227020800\, n^6} \\
c_7 &= \frac{2280356863}{1307674368000} + \frac{43847599}{1307674368000n} - \frac{332346031}{435891456000n^2} + \frac{843620579}{1307674368000n^3} - \frac{326228899}{1307674368000n^4} \\
&\quad + \frac{21470159}{435891456000n^5} - \frac{1042243}{261534873600n^6} + \frac{1}{1307674368000n^7} \\
c_8 &= \frac{49020204823}{50812489728000} - \frac{531839683}{1710035712000n} - \frac{32285445833}{88921857024000n^2} + \frac{91423417}{177843714048n^3} - \frac{51811946317}{177843714048000n^4} \\
&\quad + \frac{404003599}{4446092851200n^5} - \frac{123706507}{8083805184000n^6} + \frac{24262727}{22230464256000n^7} + \frac{1}{355687428096000n^8} \\
c_9 &= \frac{65967241200001}{121645100408832000} - \frac{14979648446341}{40548366802944000n} - \frac{26591354017}{259925428224000n^2} + \frac{73989712601}{206879422464000\, n^3} \\
&\quad - \frac{5816850595639}{20274183401472000n^4} + \frac{44978231873}{355687428096000n^5} - \frac{176126809}{5304600576000n^6} \\
&\quad + \frac{49573465457}{10137091700736000n^7} - \frac{4222378423}{13516122267648000n^8} + \frac{1}{121645100408832000n^9}
\end{aligned}
$$

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 $n \to \infty$ we obtain the series for the inverse error function with scaling of the arguments implied by the definition of v:

$$\sqrt{2}\mathrm{erf}^{-1}\left[\frac{x\sqrt{2}}{\sqrt{\pi}}\right] = x + \frac{x^3}{6} + \frac{7x^5}{120} + \frac{127x^7}{5040} + \frac{4369x^9}{362880} + \ldots$$

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.

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

But this result does give a power series about $u = \frac{1}{2}$. 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. Before doing so, however, we review another approach that is already rather well understood - the

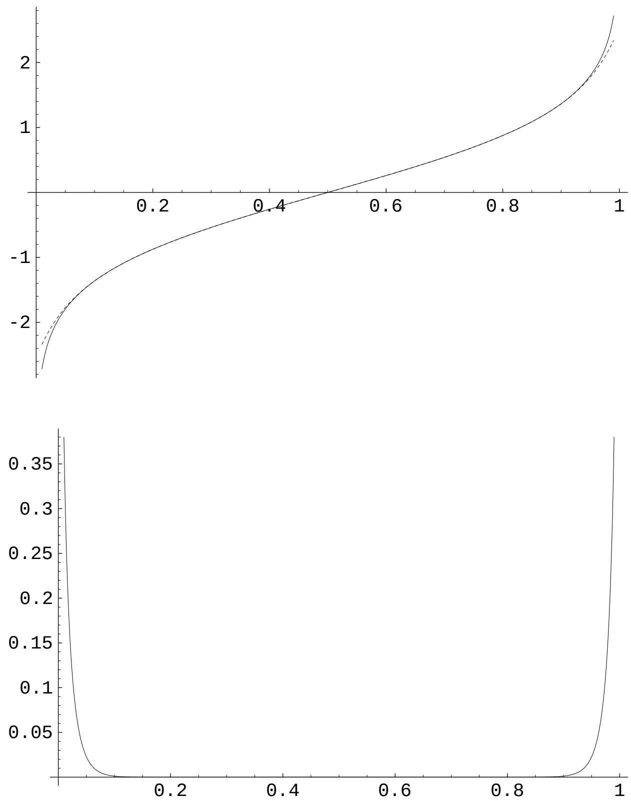

FIG. 4.1. *Top: Exact and Power Series (dashed) inverse CDF with* $n = 11$*, terms up to* c_9*. Bottom: Absolute Error.*

method based on taking n large and making corrections to the normal distribution. It will turn out that this will also push towards investigating the tails.

5. Large n and Cornish-Fisher Expansions. For a distribution that is asymptotically normal with respect to a parameter (here we consider $n \to \infty$) 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 me more explicit about the details given in [1, 2] as the CF expansion is given there rather non-explicitlyin terms of a slighly 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 first few relevant moments, $\mu_k = \mathbb{E}[T^k]$ are then

$$\mu_2 = \frac{n}{n-2}$$

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

$$\mu_4 = \frac{1.3.5\ n^3}{(n-2)(n-4)(n-6)}$$

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

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

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

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

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

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

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

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]

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

where z is a standard normal variable, the γ_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

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

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

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

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:

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

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

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

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}:

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

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 Σ 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:

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

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

$$u \longrightarrow z = N^{-1}(u) \longrightarrow s\ ,\quad t.$$

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 errors 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 5.1. Note that these are drawn using a high-precision formula for N^{-1}. If one use an approximation that is poor in the tails matters will be much worse.

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 behaviour 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 cofficients 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

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

The detailed calculations are laborious and not given.

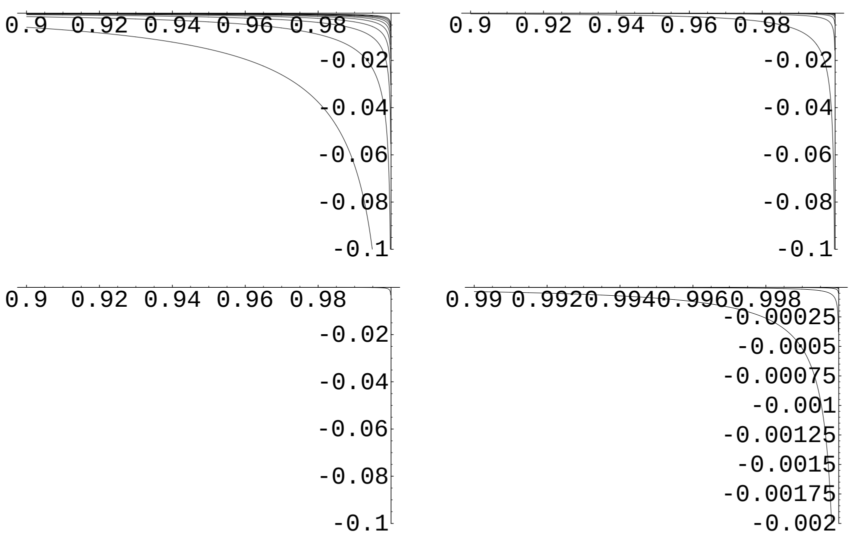

FIG. 5.1. *Errors in Large n expansion for $n = 10, 20, \ldots, 100$ with (top left), terms to n^{-1}; (top right), terms to n^{-2}; (bottom left), terms to n^{-4}; (bottom right) n^{-4} expanded plot.*

6. The Tails. 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 $|u - \frac{1}{2}| > 0.4$. We have an asymptotic series that works well for large n except in the fair tails. We see that to complete the 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

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

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

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

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

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

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

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

Further terms are given in the on-line supplement. We will now look at how well this works via some Case Studies

7. 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!

7.1. $n = 3$ **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 sorted out! The power series about $u = 1/2$ is given by

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

where

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

The corresponding tail series truncated at six terms is given by

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

where

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

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

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

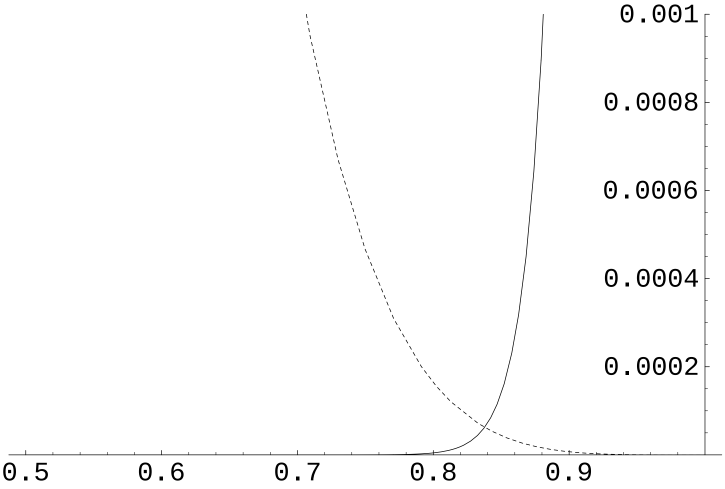

FIG. 7.1. *Absolute Errors in 9-term Power (solid) and 6-term Tail (dashed) $n = 3$ series.*

We now take a look at the results, using the method based on the inverse beta function as our benchmark. In Figure 7.1 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 taking more terms or perhaps refining by applying the cobwebbing method for $n = 3$ discussed previously.

7.2. $n = 11$ - a case of "modest" n. We repeat the above analysis with $n \to 11$. So for the power series

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

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

The tail series is now

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

and then

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

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

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

The results for the errors in the series and the tail are shown in Figure 7.2 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 7.3 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

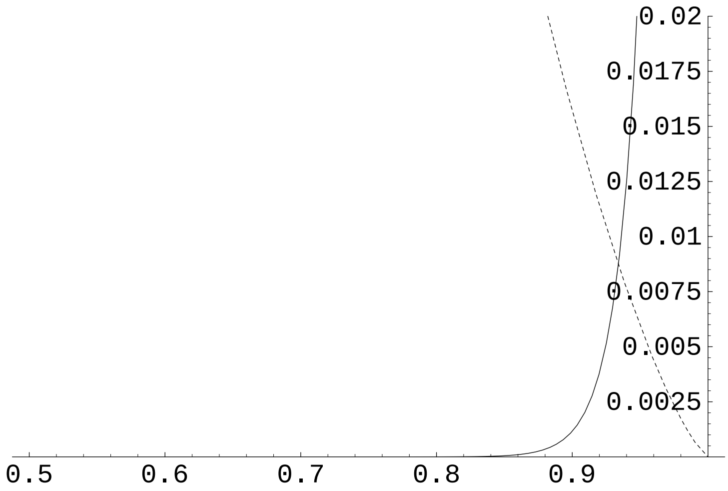

FIG. 7.2. *Absolute Errors in 9-term Power (solid) and 6-term Tail (dashed) $n = 11$ series.*

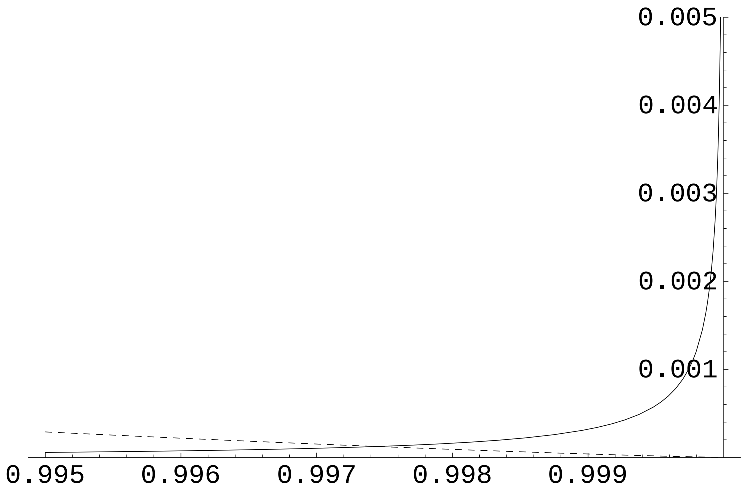

FIG. 7.3. *Absolute Errors in Cornish-Fisher (solid) Tail (dashed) series for $n = 11$ and last half percent.*

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 - 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!

7.3. Crossover. The fine details of the optimal crossover point between the central and tail power series are quite subtle, in that they will depend a lot on how many terms are taken in each series. But we can get a rough idea for scales by looking at the expansion parameters. In the power series we are expanding in terms of x^2/n, and for the inverse we are expanding in terms of n/x^2. So we would expect crossover points to lie at a point given by

$$u_{co} \sim F_n(\delta\sqrt{n})$$

where δ depends on how many terms are taken in each series. For the examples we have quoted, it appears that $\delta \sim 2/3$ is about right for small to modest n, and the crossover point moves into the tail of the distribution as n increases. In practice some experimentation is desirable. In the case of the CF expansion switching to the tail series, we can at least provide an upper bound. Recalling the tail formula

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

we can see that the power law behaviour dominates when the first term is larger than the second, i.e.,

$$x > \frac{n}{\sqrt{2}}$$

So this gives us an upper bound for where we *must* start to use the tail expansion, as

$$u_{co2} \sim F_n(n/\sqrt{2})$$

This is rather higher than the point we found in our study, so things go wrong before then. We can look at this again from the point of view of approximating the Student density with a Normal density.

$$(1 + \frac{t^2}{n})^{-\frac{n+1}{2}} = \exp\left\{\log\left((1 + \frac{t^2}{n})^{-\frac{n+1}{2}}\right)\right\} \sim \exp\left\{-\frac{(n+1)}{2}\left(\frac{t^2}{n} - \frac{t^4}{2n^2} + \dots\right)\right\}$$

We see again that the relevant expansion parameter, i.e. ratio of first to second term, is t^2/n so we might again expect the cross-over to be located at a point

$$u_{co3} \sim F_n(\epsilon\sqrt{n})$$

for some $\epsilon > \delta$. Again, some experimentation is desirable. The CF to tail crossover point we found is compatible with the choice $\delta \sim 1$. This should only be of concern if one wishes to correct the tail behaviour of the CF expansion back to strict T behaviour, including getting the power law in the tails. In cases of interest where one really cares about the tails, e.g. the high-freqeuncy applications, one is more likley to be interested in lower values of n, that we have already covered.

8. Summary. 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 β-functions, suitable for benchmark one-off computations;
- Exact solutions for the iCDF in terms of elementary functions for $n = 2, 4$, which are themselves of interest to "fat-tailed finance" applications;
- Fast iterative Newton-Raphson techniques the iCDF for even integer $n \leq 20$.
- 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 behaviour should exist and will fail with CF.

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 10-20 to use the power series and tail series developed here;
- If n is real and greater than about 10-20 to use the known Cornish-Fisher expansion given in [1, 2], using the generalized power series for the tail developed here if one wishes to recover the power law behaviour.

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.

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 neighbourhoods 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

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

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

Acknowledgements. The author wishes to thank Nomura International plc for their ongoing support of the Nomura Centre for Quantitative Finance in the Mathematical Institute, Oxford. The author was stimulated to consider new ways of managing the T by a lecture by and brief discussion with Prof. P. Embrechts during a visit to the Nomura Centre. The author is also grateful to Dr. A. Ilhan for explaining the importance of inverse CDFs to copula techniques.

REFERENCES

[1] M. Abramowitz, I. A. Stegun Handbook of Mathematical Functions, Dover Edition.
[2] M. Abramowitz, I. A. Stegun Handbook of Mathematical Functions, on line at http://www.math.sfu.ca/~cbm/aands/
[3] R. Gencay, M. Dacoronga, U. Muller, R.Olsen, O. Pictet *An Introduction to High-frequency Finance*, Academic Press, 2001.
[4] B. Moro, *The Full Monte*, *Risk*, 8(2): 57-58, Feb., 1995.
[5] W. T. Shaw, *Complex Analysis with Mathematica*, Cambridge University Press, Cambridge, UK, forthcoming 2005.
[6] K. Ferguson, E. Platen, *On the Distributional Characterization of Log-returns of a World Stock Index*, UTS Working Paper, March 2005, To appear in *Applied Mathematical Finance*.
[7] S. Wolfram, *The Mathematica Book, 5th ed*, Wolfram Media, 2004.