

## A new asymptotic representation and inversion method for the Student's $t$ distribution

A. Gil<sup>a</sup> and J. Segura<sup>b</sup> and N. M. Temme<sup>c</sup>

<sup>a</sup>Departamento de Matemática Aplicada y CC. de la Computación. ETSI Caminos. Universidad de Cantabria. 39005-Santander, Spain; <sup>b</sup>Departamento de Matemáticas, Estadística y Computación. Facultad de Ciencias. Universidad de Cantabria. 39005-Santander, Spain; <sup>c</sup>IAA, 1825 BD 25, Alkmaar, The Netherlands. Former address: Centrum Wiskunde & Informatica (CWI), Science Park 123, 1098 XG Amsterdam, The Netherlands.

### ARTICLE HISTORY

Compiled November 18, 2022

### ABSTRACT

Some special functions are particularly relevant in applied probability and statistics. For example, the incomplete beta function is the cumulative central beta distribution. In this paper, we consider the inversion of the central Student's- $t$  distribution which is a particular case of the central beta distribution. The inversion of this distribution function is useful in hypothesis testing as well as for generating random samples distributed according to the corresponding probability density function. A new asymptotic representation in terms of the complementary error function will be one of the important ingredients in our analysis. As we will show, this asymptotic representation is also useful in the computation of the distribution function. We illustrate the performance of all the obtained approximations with numerical examples.

### KEYWORDS

Incomplete beta function; Asymptotic expansions; Student's- $t$  distribution

## 1. Introduction

There is a very close relationship between some special functions and some of the most popular distribution functions in statistics. For example, the incomplete beta function [1, §8.17] is the central beta distribution. Particular cases include other well-known distributions such as the geometric, binomial, negative binomial or the central Student's- $t$  distribution. Therefore, standard methods for the computation and inversion of special functions [2] can also be applied to evaluate and invert distribution functions. The problem of inversion appears, for example, when computing percentage points of the distribution functions; also, it is closely related to the generation of random variates from a continuous probability density function needed, for example, in Monte Carlo or quasi-Monte Carlo methods.

We considered the central beta distribution  $I_x(p, q)$  in a previous publication [3]. For the asymptotic analysis of  $I_x(p, q)$ , we considered large values of  $p$  and  $q$ . In this paper, we focus on the computation and inversion of the central Student's- $t$  distribution, which has multiple applications in science and engineering (for an application in

physics, see for example [4]). For this distribution, the parameters  $p$  or  $q$  of the central beta distribution take the value  $\frac{1}{2}$  (see Eq. (5)). Therefore, the case of the Student's- $t$  distribution requires a particular study. An asymptotic representation in terms of the complementary error function will play a key role in our analysis. The performance of the approximations obtained will be illustrated with numerical examples.

There is a vast literature on Student's  $t$  distribution. An extensive overview can be found at [5]; see also the references contained therein. For an historical account and the origin of this distribution, we refer to [6]. For a generalization from the viewpoint of special functions, see [7].

Some useful expressions for the analysis of the central Student's- $t$  distribution are:

a) Probability density function:

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

where  $t \in \mathbb{R}$  and  $B(p, q)$  is the Beta integral;  $n > 0$ , not necessarily an integer.

b) Cumulative distribution function:

$$F_n(x) = \int_{-\infty}^x f_n(t) dt, \quad x \in \mathbb{R}. \quad (2)$$

c) Incomplete beta function:

$$I_x(p, q) = \frac{1}{B(p, q)} \int_0^x t^{p-1} (1-t)^{q-1} dt, \quad B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}. \quad (3)$$

From the integral representation, we have:

$$I_x(p, q) = 1 - I_{1-x}(q, p). \quad (4)$$

d) Cumulative distribution function in terms of the incomplete beta function:

$$F_n(x) = \begin{cases} \frac{1}{2} + \frac{1}{2} I_{\frac{x^2}{n+x^2}}(\frac{1}{2}, \frac{1}{2}n) = 1 - \frac{1}{2} I_{\frac{n}{n+x^2}}(\frac{1}{2}n, \frac{1}{2}), & \text{if } x \geq 0, \\ \frac{1}{2} - \frac{1}{2} I_{\frac{x^2}{n+x^2}}(\frac{1}{2}, \frac{1}{2}n) = \frac{1}{2} I_{\frac{n}{n+x^2}}(\frac{1}{2}n, \frac{1}{2}), & \text{if } x \leq 0. \end{cases} \quad (5)$$

e) Cumulative distribution function in terms of the Gauss hypergeometric functions:

$$F_n(x) = \begin{cases} \frac{1}{2} + \frac{x}{\sqrt{n} B(\frac{1}{2}, \frac{1}{2}n)} {}_2F_1\left(\frac{1}{2}, \frac{1}{2}n + \frac{1}{2}; \frac{3}{2}; -\frac{x^2}{n}\right), & \text{if } x \in \mathbb{R}, \\ 1 - \frac{(1-y)^{\frac{1}{2}n} y^{-\frac{1}{2}}}{n B(\frac{1}{2}, \frac{1}{2}n)} {}_2F_1\left(1, \frac{1}{2}; \frac{1}{2}n + 1; -\frac{n}{x^2}\right), & \text{if } x \geq 0, \\ \frac{(1-y)^{\frac{1}{2}n} y^{-\frac{1}{2}}}{n B(\frac{1}{2}, \frac{1}{2}n)} {}_2F_1\left(1, \frac{1}{2}; \frac{1}{2}n + 1; -\frac{n}{x^2}\right), & \text{if } x \leq 0, \end{cases} \quad (6)$$

where  $y = x^2/(n + x^2)$ . The first formula in (6) is given in [8], the other ones follow from well-known relations between the incomplete beta function and the hypergeometric function; see [1, §8.17(ii)].

## 2. Asymptotic expansion of the Student's $t$ cumulative distribution function

The second and third representation in (6) can be used for large values of  $x$  by using the standard power series of the hypergeometric functions. It is not necessary that  $x^2 > n$ , but a condition  $x^2/n = \mathcal{O}(1)$  is needed. To obtain a large- $n$  asymptotic representation, whether or not  $x$  is large, we use a method that we have used for other cumulative distribution functions; see [9, Chapter 36].

We use in (1) the substitution  $t = s\sqrt{n}$ . This gives

$$\begin{aligned} F_n(x) &= \frac{1}{\sqrt{n} B(\frac{1}{2}, \frac{1}{2}n)} \int_{-\infty}^x \left(1 + \frac{t^2}{n}\right)^{-\frac{1}{2}n - \frac{1}{2}} dt \\ &= \frac{1}{B(\frac{1}{2}, \frac{1}{2}n)} \int_{-\infty}^{x/\sqrt{n}} (1 + s^2)^{-\frac{1}{2}n - \frac{1}{2}} ds. \end{aligned} \quad (7)$$

We write  $u^2 = \ln(1 + s^2)$ , with the condition  $\text{sign}(u) = \text{sign}(s)$ , and obtain

$$F_n(x) = \frac{1}{B(\frac{1}{2}, \frac{1}{2}n)} \int_{-\infty}^{\xi} e^{-\frac{1}{2}nu^2} g(u) du, \quad (8)$$

where

$$g(u) = \sqrt{\frac{u^2}{1 - e^{-u^2}}}, \quad \xi^2 = \ln\left(1 + \frac{x^2}{n}\right), \quad \text{sign}(\xi) = \text{sign}(x). \quad (9)$$

Using the method of [9, §36.1] we find that  $F_n(x)$  can be written in the form

$$F_n(x) = \frac{1}{2} \text{erfc}\left(-\xi\sqrt{n/2}\right) - \frac{e^{-\frac{1}{2}n\xi^2}}{\sqrt{2\pi n} \beta(n)} B_n(\xi), \quad \beta(n) = \sqrt{\frac{n}{2\pi}} B\left(\frac{1}{2}, \frac{1}{2}n\right), \quad (10)$$

where we have introduced the complementary error function

$$\operatorname{erfc} z = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-t^2} dt. \quad (11)$$

The function  $B_n(\xi)$  has the asymptotic expansion

$$B_n(\xi) \sim \sum_{k=0}^{\infty} \frac{C_k(\xi)}{n^k}, \quad n \rightarrow \infty, \quad \xi \in \mathbb{R}, \quad (12)$$

where the coefficients follow from the recursive scheme

$$C_k(\xi) = \frac{g_k(\xi) - g_k(0)}{\xi}, \quad g_{k+1}(u) = \frac{d}{du} \frac{g_k(u) - g_k(0)}{u}, \quad (13)$$

with  $g_0(u) = g(u)$  defined in (9). The first coefficients are

$$\begin{aligned} C_0(\xi) &= \frac{g(\xi) - 1}{\xi}, \\ C_1(\xi) &= -\frac{4g(\xi)^3 - 4g(\xi)\xi^2 + \xi^2 - 4}{4\xi^3}, \\ C_2(\xi) &= \frac{96g(\xi)^5 - 128g(\xi)^3\xi^2 + 32g(\xi)\xi^4 - \xi^4 + 8\xi^2 - 96}{32\xi^5}. \end{aligned} \quad (14)$$

The function  $\beta(n)$  has the expansion

$$\beta(n) \sim \sum_{k=0}^{\infty} \frac{D_k}{n^k}, \quad n \rightarrow \infty, \quad D_k = g_k(0). \quad (15)$$

The first coefficients are

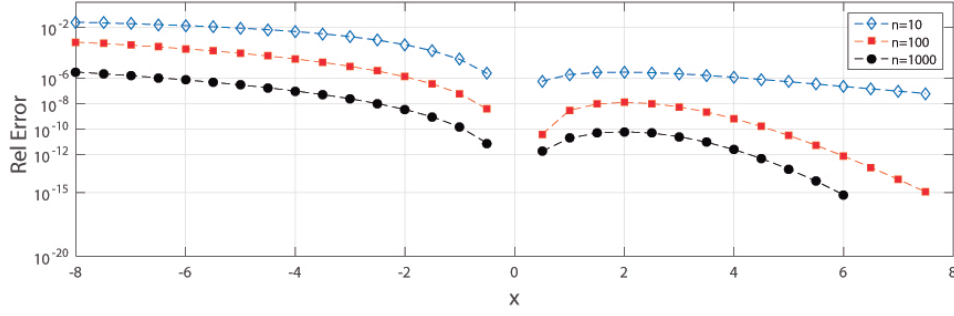
$$D_0 = 1, \quad D_1 = \frac{1}{4}, \quad D_2 = \frac{1}{32}, \quad D_3 = -\frac{5}{128}. \quad (16)$$

These coefficients can be expressed in terms of the coefficients  $a_k$  of  $g(u) = \sum_{k=0}^{\infty} a_k u^k$ .

We have (see [9, Remark 36.2])

$$D_k = \left(\frac{1}{2}\right)^k 2^k a_{2k}, \quad k = 0, 1, 2, \dots \quad (17)$$

Examples of the performance of the expansion (10) for three values of  $n$  ( $n = 10, 100, 1000$ ) are shown in Figure 1. Five  $C_k$  coefficients in the series (12) have been considered in the computations. The relative errors were obtained comparing the expansion (10) with the values of the distribution function given in (6), which is computed with Maple.



**Figure 1.** Relative errors obtained when computing the Student's  $t$  cumulative distribution function using the expansion (10).

### 3. Inversion of the Student's $t$ cumulative distribution function

The inversion problem is: find  $x$  that satisfies the equation

$$F_n(x) = p, \quad 0 < p < 1. \quad (18)$$

We consider three different approaches, the first one is for small values of  $p - \frac{1}{2}$ , which will give small values of  $x$ . The second method is for small values of  $p$ , which gives large negative values  $x$ . Thirdly we use the uniform approximation of §2, which will be valid for a large range of  $x$ , including the values near  $x = 0$ .

#### 3.1. Inversion for small values of $p - \frac{1}{2}$

We use the first representation in (6), and write the inversion problem in the form

$$x \sum_{k=0}^{\infty} c_k x^{2k} = q, \quad q = \left(p - \frac{1}{2}\right) \sqrt{n} B\left(\frac{1}{2}, \frac{1}{2}n\right), \quad (19)$$

and the  $c_k$  follow from the coefficients of the hypergeometric function:

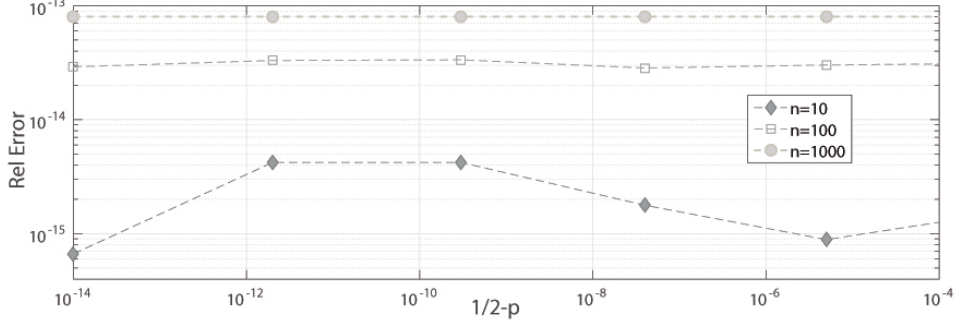
$$c_k = \frac{\left(\frac{1}{2}\right)_k \left(\frac{1}{2}n + \frac{1}{2}\right)_k (-1)^k}{k! \left(\frac{3}{2}\right)_k n^k}. \quad (20)$$

The solution of the equation in (18) has the expansion

$$x = q \sum_{k=0}^{\infty} x_k q^{2k}, \quad (21)$$

and the first coefficients are

$$\begin{aligned} x_0 &= 1, \quad x_1 = \frac{n+1}{6n}, \quad x_2 = \frac{(n+1)(7n+1)}{120n^2}, \\ x_3 &= \frac{(n+1)(127n^2+8n+1)}{5040n^3}, \quad x_4 = \frac{(n+1)(4369n^3-537n^2+135n+1)}{362880n^4}. \end{aligned} \quad (22)$$



**Figure 2.** Performance of the expansion for small values of  $p - \frac{1}{2}$  for three values of  $n$ . The five coefficients given in (22) for the expansion have been considered in the computations.

We see that the shown coefficients are bounded for large values of  $n$ . Also,  $q = (p - \frac{1}{2}) \sqrt{2\pi} \beta(n)$  (see (10)) with an expansion of  $\beta(n)$  given in (15). This shows that the inversion considered here for small values of  $(p - \frac{1}{2})$  is rather well conditioned for large values of  $n$ . In Figure 2 we show examples of the performance of the expansion (using the four terms given in (22)) for three values of  $n$ . The values of  $p$  considered in the calculations are  $p = \frac{1}{2} - \Delta$ , for  $\Delta = 10^{-14}, 2 \times 10^{-12}, 3 \times 10^{-10}, 4 \times 10^{-8}, 5 \times 10^{-6}, 6 \times 10^{-4}$ . The results obtained with the expansion have been compared with the Matlab function for the inversion of the central Student's-t distribution (function `tinu`). As can be seen, a relative error better than  $10^{-13}$  is obtained in all cases.

**Remark 1.** By using the definition in (2), it is easily verified that the function  $F_n(x)$  becomes for  $n = 1$

$$F_1(x) = \frac{1}{2} + \frac{1}{\pi} \arctan x, \quad (23)$$

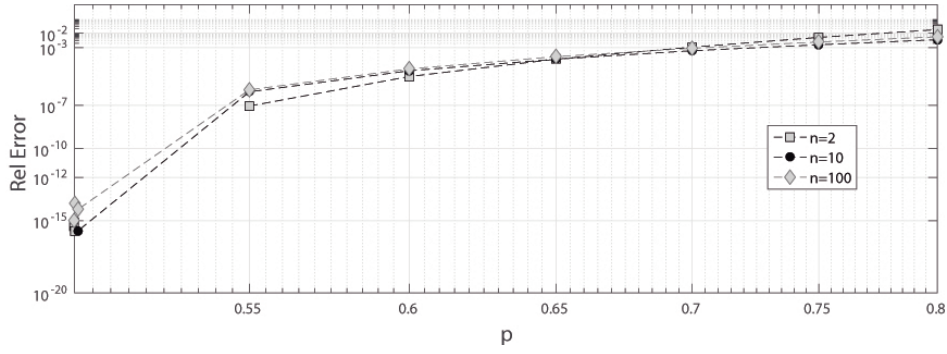
and for the inversion problem  $F_n(x) = p$  we have  $q = (p - \frac{1}{2}) \pi$ . The equation to be inverted becomes  $\arctan x = q$ , with solution  $x = \tan q$ . The shown coefficients in (22) correspond with those of the expansion  $x = q + \frac{1}{3}q^3 + \frac{2}{15}q^5 + \dots$ . This expansion converges for  $0 < p < 1$ .

### 3.2. High-order iteration

It should be mentioned that it is also possible to obtain numerical approximations for the inversion problem for small values of  $p - \frac{1}{2}$  using the fixed point iterations giving sharp error bounds for the central beta distribution described in [3]. For example, iterating the fixed point iterations  $y = g(y)$  or  $y = h(y)$  where

$$g(y) = (2(p - \frac{1}{2})B(\frac{1}{2}, \frac{1}{2}n)(\frac{1}{2} - (\frac{1}{2} + \frac{1}{2}n)y))^2 (1 - y)^{-n},$$

$$h(y) = 225 \frac{\left(p - \frac{1}{2}\right)^2 B\left(\frac{1}{2}, \frac{1}{2}n\right)^2 (1 - y)^{-n}}{(y^2(n^2 + 4n + 3) + 5y(n + 1) + 15)^2}, \quad (24)$$



**Figure 3.** Relative errors obtained when using (25) to approximate the solution to the inversion problem (18) for  $n = 2, 10, 100$ .

starting from  $y = 0$ , approximations to the values of  $x$  in (18) are obtained with  $x = \sqrt{\frac{ny}{1-y}}$ . As an example, using 2 iterations of the fixed point iteration  $y = h(y)$  for  $n = 10$ ,  $p = 0.5 + 10^{-5}$  we find  $x = 0.00002569978035$  with a relative error  $5 \times 10^{-12}$ . An explicit expression for the  $x$  value obtained in the second iteration is

$$x = 120 \left( \frac{nZ(1 - (1/4)Z)^{-n}}{Z^4 A_4 + Z^3 A_3 + Z^2 A_2 + Z A_1 + 57600} \right)^{1/2}, \quad (25)$$

where  $Z = (2(p - \frac{1}{2})B(\frac{1}{2}, \frac{1}{2}n))^2$  and

$$\begin{aligned} A_1 &= -14400(1 - (1/4)Z)^{-n} + 9600n + 9600, \\ A_2 &= 880n^2 + 2720n + 1840, \\ A_3 &= 40n^3 + 200n^2 + 280n + 120, \\ A_4 &= n^4 + 8n^3 + 22n^2 + 24n + 9. \end{aligned} \quad (26)$$

This approximation can be also used for not so small values of  $p - \frac{1}{2}$ , as can be seen in Figure 3.

A comparison of (25) with the expansion given in Section 3.1 for few values of  $p$  close to  $\frac{1}{2}$  is given in table 1. The five coefficients given in (22) have been considered in the computations. In the table, we show the accuracy for two values of  $n$  ( $n = 10, 100$ ). As can be seen, the results obtained with the methods are very similar. Also, we have checked that the average CPU time spent by the methods (implemented in Fortran 90) are very close. As an example, the computations for  $p = \frac{1}{2} + 10^{-4}$ ,  $n = 10$  took  $1.37 \mu s$  and  $1.52 \mu s$ , when using (21) and (25), respectively. For  $n = 100$  (using the same value of  $p$ ), the computations took  $0.81 \mu s$  and  $0.96 \mu s$ , respectively. As can be seen, both methods are very efficient.

$p, n$	Rel. Error	Rel. Error
	Method in Section 3.2	Method in Section 3.1
$p = \frac{1}{2} + 10^{-12}, n = 10$	$2.22 \times 10^{-15}$	$2.00 \times 10^{-15}$
$p = \frac{1}{2} + 10^{-10}, n = 10$	$1.55 \times 10^{-15}$	$1.55 \times 10^{-15}$
$p = \frac{1}{2} + 10^{-8}, n = 10$	$4.44 \times 10^{-16}$	$4.44 \times 10^{-16}$
$p = \frac{1}{2} + 10^{-6}, n = 10$	$1.11 \times 10^{-16}$	$2.22 \times 10^{-16}$
$p = \frac{1}{2} + 10^{-4}, n = 10$	$4.44 \times 10^{-16}$	$6.66 \times 10^{-16}$
$p = \frac{1}{2} + 10^{-12}, n = 100$	$3.26 \times 10^{-14}$	$3.24 \times 10^{-14}$
$p = \frac{1}{2} + 10^{-10}, n = 100$	$3.26 \times 10^{-14}$	$3.24 \times 10^{-14}$
$p = \frac{1}{2} + 10^{-8}, n = 100$	$3.15 \times 10^{-14}$	$3.13 \times 10^{-14}$
$p = \frac{1}{2} + 10^{-6}, n = 100$	$3.22 \times 10^{-14}$	$2.98 \times 10^{-14}$
$p = \frac{1}{2} + 10^{-4}, n = 100$	$2.77 \times 10^{-14}$	$3.02 \times 10^{-14}$

**Table 1.** Comparison of the accuracy obtained when using (25) and (21) to approximate the solution to the inversion problem (18).

### 3.3. Inversion for small values of $p$

For this case we use<sup>1</sup> the representation in the third line of (6) and introduce

$$\eta = \frac{n}{x^2}, \quad \delta = \left( pnB\left(\frac{1}{2}, \frac{1}{2}n\right) \right)^{\frac{2}{n}}. \quad (27)$$

Then we can write the equation  $F_n(x) = p$  in the form

$$\eta(1 + \eta)^{\frac{1}{n}-1} S(\eta)^{\frac{2}{n}} = \delta, \quad (28)$$

where  $S(\eta)$  is the standard power series of the hypergeometric function in the third line of (6).

We see that for small values of  $\delta$  the wanted variable  $\eta$  behaves as  $\eta \sim \delta$ , and substituting the expansion

$$\eta = \sum_{k=1}^{\infty} \eta_k \delta^k, \quad (29)$$

we find the following first few coefficients

$$\begin{aligned} \eta_1 &= 1, \quad \eta_2 = \frac{n+1}{n+2}, \quad \eta_3 = \frac{(n+1)(2n^2+9n+6)}{2(n+2)^2(n+4)}, \\ \eta_4 &= \frac{3n^6+35n^5+134n^4+328n^3+1174n^2+2100n+1152}{3n(n+2)^3(n+4)(n+6)}. \end{aligned} \quad (30)$$

When we have computed  $\eta$ ,  $x$  follows from (27):  $x = -\sqrt{n/\eta}$ .

---

<sup>1</sup>The approach of this section is similar to that of one of the inversion methods discussed in [10].



$p$	$x$	Rel. Error
$1 \times 10^{-50}$	-256452.5718769479	$6.5 \times 10^{-15}$
$2 \times 10^{-40}$	-23927.87084268530	$3.2 \times 10^{-15}$
$3 \times 10^{-30}$	-2297.706518629116	$1.0 \times 10^{-13}$
$4 \times 10^{-20}$	-223.234400503956	$1.1 \times 10^{-09}$
$5 \times 10^{-10}$	-21.62201646469524	$1.2 \times 10^{-05}$

**Table 2.** Inversion values  $x$  obtained (for  $n = 10$ ) using the series in (29) for small values of  $p$ . The coefficients given in (30) have been considered in the calculations. Relative errors in comparison to Maple, are also shown in the table.

Again, we see that the coefficients satisfy  $\eta_k = \mathcal{O}(1)$  for large values of  $n$ , which also happens for all coefficients that we have evaluated. The first 10 coefficients satisfy  $\eta_k = 1 + \mathcal{O}(1/n)$  as  $n \rightarrow \infty$ .

When  $n$  is large, the only problem is that  $\delta$  tends to 1, When  $n$  is large, the only problem is that  $\delta$  tends to 1, which is a bad condition for convergence of the series in (29).

For example, when  $p = 10^{-8}$  and  $n = 10$ , we have  $\delta \doteq 0.038193186$ . With 5 terms in the series in (29) we find  $\eta \doteq 0.039576861$ , giving  $x \doteq -15.8956879$ . Then,  $F_n(x) = 9.9999981 \times 10^{-9}$ , with relative error  $1.92 \times 10^{-7}$ . For  $n = 25$  we find  $\delta \doteq 0.281$ ,  $x \doteq -8.0759$  and a relative error 0.011. Other examples of the performance of the series for small values of  $p$  are given in Table 2, where we show the values and relative errors obtained for  $n = 10$  and few values of  $p$ . The four coefficients given in (30) have been considered in the calculations. Since the Matlab function `tin` seems to fail for very small values of  $p$ , the tests have been performed comparing with Maple. Testing is done by checking whether the composition of the function with its inverse is the identity: using the inversion value  $x_{inv}$  obtained with the method, we compute  $F_n(a, x_{inv})$  given in Eq. (6). Next, we calculate the relative error  $|1 - F_n(a, x_{inv})/p|$ .

**Remark 2.** When we take  $n = 1$  and use (23), the inversion problem becomes

$$x = -\cot(p\pi), \quad \eta = \tan^2(p\pi), \quad \delta = p^2\pi^2, \quad 0 < p < 1. \quad (31)$$

The expansion in (29) becomes for  $p < \frac{1}{2}$  and  $\delta < \frac{1}{4}\pi^2$

$$\tan^2(p\pi) = \sum_{k=1}^{\infty} \eta_k \delta^k, \quad \eta_1 = 1, \quad \eta_2 = \frac{2}{3}, \quad \eta_3 = \frac{17}{45}. \quad (32)$$

These values of  $\eta_k$  correspond with those given in (30) for  $n = 1$ .

### 3.4. Inversion by using the uniform expansion

We use the representation given in (10) and first try to find  $\xi$ , then  $x$  follows from the relation in (9). We assume that  $n$  is large and use the asymptotic method as described in [9, §42.1] and in our papers [3], [11], [12].

Let  $\xi_0$  satisfy the equation

$$\frac{1}{2}\text{erfc}\left(-\xi_0\sqrt{n/2}\right) = p. \quad (33)$$

Then we assume for  $\xi$  the expansion

$$\xi \sim \xi_0 + \frac{\xi_1}{n} + \frac{\xi_2}{n^2} + \frac{\xi_3}{n^3} + \dots, \quad (34)$$

where  $\xi_k$  have to be determined. When we have this approximation for  $\xi$  we compute  $x$  from (9).

From (18), (33) and (8) we find

$$\frac{dp}{d\xi_0} = \sqrt{\frac{n}{2\pi}} e^{-\frac{1}{2}n\xi_0^2}, \quad \frac{dp}{d\xi} = \sqrt{\frac{n}{2\pi}} \frac{g(\xi)}{\beta(n)} e^{-\frac{1}{2}n\xi^2}, \quad (35)$$

where  $\beta(n)$  is defined in (10).

Dividing the two derivatives, we find

$$g(\xi) \frac{d\xi}{d\xi_0} = \beta(n) e^{\frac{1}{2}n(\xi^2 - \xi_0^2)}, \quad (36)$$

We substitute the expansion given in (34), use  $\beta(n) = 1 + \mathcal{O}(1/n)$ , and obtain, considering equal large-order terms of  $n$ , the next term in the expansion:

$$g(\xi_0) = e^{\xi_0\xi_1} \implies \xi_1 = \frac{1}{\xi_0} \ln g(\xi_0). \quad (37)$$

Because

$$g(u) = 1 + \frac{1}{4}u^2 + \frac{1}{96}u^4 + \mathcal{O}(u^6), \quad (38)$$

it follows that  $\xi_1$  is well defined when  $\xi_0$  tends to zero (that is, when  $p \sim \frac{1}{2}$ ).

We can find higher-order terms  $\xi_j, j \geq 2$ , of the expansion in (34) using more coefficients in the asymptotic expansion of  $\beta(n)$  (see (16)). Also, we need the expansion

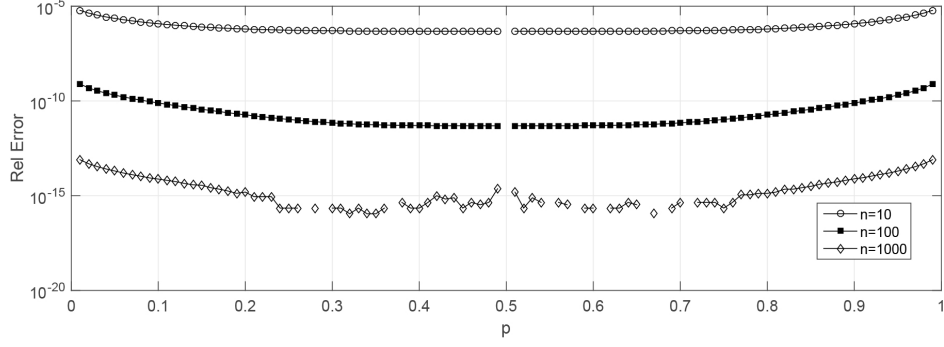
$$g(\xi) = g(\xi_0) + (\xi - \xi_0)g'(\xi_0) + \frac{1}{2}(\xi - \xi_0)^2g''(\xi_0) + \dots \quad (39)$$

By using (36) and algebraic manipulations we find a few other coefficients:

$$\begin{aligned} \xi_2 &= -\left(2g\xi_1^2 + 4(g - \xi_1g')\xi_1 + \xi_1g - 4g'\right)/(4\xi^2g), \\ \xi_3 &= \left(2\xi^2g^2\xi_1^3 + \left(2\xi^3gg'' - 2\xi^3g'^2 - 6\xi^2gg' + 8\xi g^2\right)\xi_1^2 + \right. \\ &\quad \left.(12g + \xi^2g - 16\xi g' + 4\xi^2g'')g\xi_1 + \right. \\ &\quad \left.\xi g^2 + 4\xi gg'' + 2\xi g'^2 - \xi^2gg' - 12gg'\right)/(4\xi^4g^2), \end{aligned} \quad (40)$$

where  $\xi = \xi_0$  and  $g, g'$  and  $g''$  are evaluated at  $\xi_0$ .

For small values of  $\xi_0$  (that is, when  $p \sim \frac{1}{2}$ ), we need expansions. We have



**Figure 4.** Performance of the asymptotic inversion method using four terms of the expansion for small values of  $\xi_0$  given in (41). Three values of  $n$  are considered in the computations.

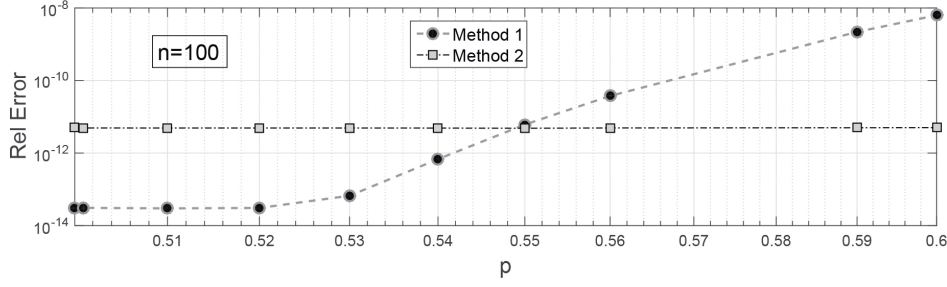
$$\begin{aligned}
\xi_1 &= \frac{1}{4}\xi_0 - \frac{1}{48}\xi_0^3 + \frac{1}{5760}\xi_0^7 - \frac{1}{362880}\xi_0^{11} + \frac{1}{19353600}\xi_0^{15} + \dots, \\
\xi_2 &= \frac{1}{32}\xi_0 - \frac{5}{192}\xi_0^3 + \frac{7}{2560}\xi_0^5 + \frac{1}{2560}\xi_0^7 - \frac{407}{5806080}\xi_0^9 - \frac{13}{1451520}\xi_0^{11} + \dots, \\
\xi_3 &= -\frac{5}{128}\xi_0 - \frac{11}{1536}\xi_0^3 + \frac{63}{10240}\xi_0^5 - \frac{823}{2580480}\xi_0^7 - \frac{5291}{23224320}\xi_0^9 + \dots, \\
\xi_4 &= -\frac{21}{2048}\xi_0 + \frac{37}{2048}\xi_0^3 + \frac{179}{81920}\xi_0^5 - \frac{22711}{10321920}\xi_0^7 + \dots, \\
\xi_5 &= \frac{399}{8192}\xi_0 + \frac{219}{32768}\xi_0^3 - \frac{3679}{327680}\xi_0^5 + \dots, \\
\xi_6 &= \frac{869}{65536}\xi_0 - \frac{6877}{131072}\xi_0^3 + \dots
\end{aligned} \tag{41}$$

**Example 3.1.** We summarise the algorithmic steps for the inversion method. We take  $n = 10$ ,  $p = 0.44$  and two terms in the expansion in (34).

- (1) Compute  $\xi_0$  from equation (33). We have  $\xi_0 \doteq -0.047746$ .
- (2) Compute  $\xi_1$  from equation (37) by using  $g(u)$  defined in (9):  $\xi_1 \doteq -0.011933$ .
- (3) Compute  $\xi$  by using (34) with the information now available:  $\xi \sim \xi_0 + \xi_1/n \doteq -0.048934$ .
- (4) Compute  $x$  from equation (9). Because  $\xi < 0$ ,  $x$  should be negative:  $x \doteq -0.1548354$ .
- (5) Verification: compute  $F_n(x)$  by using the first line in (6):  $F_{10}(x) \doteq 0.4400158$ . Relative error: 0.000036.

A test of the performance of the asymptotic inversion method for three values of  $n$  using the expansion for small values of  $\xi_0$  given in (41), can be seen in Figure 4. The terms of the expansion  $\xi_1, \xi_2, \xi_3, \xi_4$  given in (41) have been considered in the computations. For computing the inverse of the complementary error function needed to compute  $\xi_0$  in (33) we use the function **inverfc** given in [13]. The relative errors obtained in comparison to the Matlab function **tinu** are shown in the figure.

It is interesting to compare the performance of the asymptotic inversion method with the methods given in Section 3.1 (for values of  $p$  close to  $\frac{1}{2}$ ) and Section 3.3 (for values of  $p$  close to 0). In Figure 5 we show the performance of the method given in Section 3.1 and the asymptotic inversion method for a few values of  $p$  and  $n = 100$ . Four terms in the expansions have been considered in the calculations. As can be seen,



**Figure 5.** Comparison of the performance of the method given in Section 3.1 (Method 1) and the asymptotic inversion method (Method 2) for few values of  $p$  and  $n = 100$ .

$p, n$	Rel. Error	Rel. Error
	Method in Section 3.3	Method in Section 3.4
$p = 10^{-10}, n = 10$	$6.7 \times 10^{-7}$	$4.4 \times 10^{-3}$
$p = 10^{-8}, n = 10$	$3.9 \times 10^{-6}$	$1.5 \times 10^{-3}$
$p = 10^{-6}, n = 10$	$8.0 \times 10^{-6}$	$3.4 \times 10^{-4}$
$p = 10^{-4}, n = 10$	$7.2 \times 10^{-4}$	$1.9 \times 10^{-5}$
$p = 10^{-3}, n = 10$	$5.3 \times 10^{-3}$	$7.3 \times 10^{-6}$

**Table 3.** Comparison of the accuracy obtained when using the methods given in Sections 3.3 and 3.4 to approximate the solution to the inversion problem (18).

the expansion in Section 3.1 performs better for the values of  $p$  closer to  $\frac{1}{2}$  (as expected) but for  $p > 0.55$  the better accuracy is obtained with the asymptotic inversion method. On the other hand, table 3 shows a comparison of the performance of the method given in Section 3.3 and the asymptotic inversion method. Relative errors obtained with the two methods are given for few values of  $p$  and  $n = 10$ . The results given in the table show that for the values of  $p$  closer to 0, the relative errors are smaller with the method given in Section 3.3; however as  $p$  increases, the asymptotic inversion method becomes the most accurate method.

#### 4. Concluding remarks

We have presented approximations for the inversion problem (18) of the central Student- $t$  distributions. To obtain the approximations, different methods have been considered depending on the values of  $p$ . In particular, one of the key elements in our analysis was the use of an asymptotic representation of the distribution function in terms of the complementary error function. Numerical tests have shown that the approximations obtained are, in all cases, accurate. Also, they are easy to compute, which is an important advantage when using the inverse to generate random variates distributed according to central Student- $t$  probability density function.

## Acknowledgements

We thank the reviewer for his/her constructive and helpful remarks. NMT thanks CWI, Amsterdam, for scientific support.

## Funding

Financial support from Ministerio de Ciencia e Innovación, Spain, project PGC2018-098279-B-I00 (MCIN/ AEI /10.13039/501100011033/ FEDER) is acknowledged.

## References

- [1] Paris RB. Chapter 8, Incomplete gamma and related functions. In: NIST Handbook of Mathematical Functions. Washington, DC: U.S. Dept. Commerce; 2010. p. 173–192. [Http://dlmf.nist.gov/8](http://dlmf.nist.gov/8).
- [2] Gil A, Segura J, Temme NM. Numerical methods for special functions. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM); 2007.
- [3] Gil A, Segura J, Temme NM. Efficient algorithms for the inversion of the cumulative central beta distribution. *Numer Algorithms*. 2017;74(1):77–91.
- [4] Röver C. Student- $t$  based filter for robust signal detection. *Phys Rev D*. 2011;84:122004. Available from: <https://link.aps.org/doi/10.1103/PhysRevD.84.122004>.
- [5] Johnson NL, Kotz S, Balakrishnan N. Continuous Univariate Distributions, volume 2. New York: Wiley-Interscience; 2nd edition; 1995.
- [6] Zabell SL. On Student’s 1908 article “The probable error of a mean”. *J Amer Statist Assoc*. 2008;103(481):1–20. With comments and a rejoinder by the author; Available from: <https://doi.org/10.1198/016214508000000030>.
- [7] Koepf W, Masjed-Jamei M. A generalization of Student’s  $t$ -distribution from the viewpoint of special functions. *Integral Transforms Spec Funct*. 2006;17(12):863–875. Available from: <https://doi.org/10.1080/10652460600856419>.
- [8] Amos D. Representations of the central and non-central  $t$  distributions. *Biometrika*. 1964; 51:451–458.
- [9] Temme NM. Asymptotic methods for integrals. Singapore: World Scientific; 2015. Series in Analysis, Vol. 6.
- [10] Hill GW. Algorithm 396, Student’s  $t$ -quantiles. *Comm ACM*. 1970;13(10):619–620.
- [11] Gil A, Segura J, Temme NM. On the computation and inversion of the cumulative non-central beta distribution. *Appl Math Comput*. 2019;361:74–86.
- [12] Gil A, Segura J, Temme NM. Asymptotic inversion of the binomial and negative binomial cumulative distribution functions. *Electron Trans Numer Anal*. 2020;52:270–280.
- [13] Gil A, Segura J, Temme NM. GammaCHI: a package for the inversion and computation of the gamma and chi-square cumulative distribution functions (central and noncentral). *Comput Phys Commun*. 2015;191:132–139.