MATHEMATICS OF COMPUTATION
Volume 65, Number 215
July 1996, Pages 1283–1288

AN ASYMPTOTIC EXPANSION FOR THE INCOMPLETE BETA FUNCTION

B.G.S. DOMAN

ABSTRACT. A new asymptotic expansion is derived for the incomplete beta function $I(a, b, x)$, which is suitable for large a, small b and $x > 0.5$. This expansion is of the form

$$I(a, b, x) \quad \sim \quad Q(b, -\gamma \log x) + \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{\gamma} \sum_{n=0}^{\infty} T_n(b, x)/\gamma^{n+1},$$

where Q is the incomplete Gamma function ratio and $\gamma = a + (b-1)/2$. This form has some advantages over previous asymptotic expansions in this region in which T_n depends on a as well as on b and x.

1. INTRODUCTION

The incomplete beta function $I(a, b, x)$ is defined by [1, p.269, Eq. 6.6.2 and p.944, Eq. 26.5.1]

$$(1.1) \quad I(a, b, x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^x t^{a-1}(1-t)^{b-1} dt, \qquad a > 0, b > 0, \quad 0 \le x \le 1.$$

While best known for its applications in Statistics, it is also widely used in many other fields.

Owing to the wide variation in behavior in different regions of the parameter space, efficient code to evaluate $I(a, b, x)$ involves a number of different subroutines for different parts of this parameter space [3]. In this paper we shall confine our interest to a subdomain of the parameter space in which a is large, b is small and x is close to 1.0. Indeed if $b < 1.0$, then $I(a, b, x)$ varies most rapidly as x approaches 1.0. This region has to be treated very carefully. Asymptotic expansions suitable for this subdomain have been derived by Molina [8] and Temme [9]. Molina's result was rederived by Wise [10, 11] and DiDonato and Morris [3]. These asymptotic expansions have the form $\Sigma A_n/\gamma^n$, where γ is either a [7] or $a + (b-1)/2$ [3, 8, 10], and in which the expansion coefficients A_n depend on all three parameters a, b and x. The expansion to be described here has the same general form, but the expansion coefficients A_n depend only on b and x. The advantage of this new expansion is that it is cleaner and that an algorithm based on it can be more easily tuned for particular accuracy requirements and for particular parameter ranges.

Received by the editor March 16, 1995 and, in revised form, June 26, 1995.
1991 *Mathematics Subject Classification.* Primary 33B20; Secondary 65D20.
Key words and phrases. Gamma function ratio, incomplete Beta function, Chi-square distribution, Student's distribution, F distribution.

The derivation here starts in the same way as that for the expansion derived by Wise [10] and can in fact be derived from it. First we show a simple derivation of an asymptotic expansion for the ratio of two Gamma functions derived first by Fields [5], see also Frenzen [6] and Luke [7, pp.33-34]. The method shown here in §2, while equivalent to that of Luke, is a special case of the method applied to the incomplete beta function in §3.

2. Asymptotic expansion of $\Gamma(a+b)/\Gamma(a)$

In this section we shall derive an asymptotic expansion which we shall need later on, which provides an efficient method of calculating $\Gamma(a+b)/\Gamma(a)$ when $a \gg b$. We start from the Beta function $B(a,b)$,

$$(2.1) \qquad B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} = \int_0^1 t^{a-1}(1-t)^{b-1}dt = \int_0^\infty e^{-at}(1-e^{-t})^{b-1}dt$$

$$= \int_0^\infty e^{-\gamma t} t^{b-1} \left(\frac{\sinh(t/2)}{t/2}\right)^{b-1} dt,$$

where $\gamma = a + (b-1)/2$. We now expand $[\sinh(t/2)/(t/2)]^{b-1}$ in powers of t^2 and use Watson's Lemma to obtain the asymptotic expansion

$$(2.2) \qquad \frac{\Gamma(a)}{\Gamma(a+b)} \sim \frac{1}{\gamma^b} \sum_{n=0}^\infty c_n \frac{\Gamma(b+2n)}{\Gamma(b)} \left(\frac{1}{\gamma}\right)^{2n},$$

where c_n are the expansion coefficients of $[\sinh(t/2)/(t/2)]^{b-1}$. The coefficients c_n can be expressed in terms of the generalized Bernoulli polynomials [7, p.34], $c_n = B_{2n}^{1-b}((1-b)/2)/(2n)!$. They can be evaluated using the recurrence relations of the generalized Bernoulli polynomials, by using computer algebra to work out the expansion of $[\sinh(t/2)/(t/2)]^{b-1}$, or following DiDonato and Morris [3], by differentiating

$$(2.3) \qquad [\sinh(t/2)/(t/2)]^{b-1} = \left(\sum_0^\infty h_n t^{2n}\right)^{b-1} = \sum_0^\infty c_n t^{2n},$$

multiplying by $\sinh(t/2)/(t/2)$ in series form and equating powers of t. This leads to a recursion formula for c_n . Actual expressions in terms of b can easily be obtained by computer algebra. If we write $d_n = 12^n c_n/z$, where z is $b-1$, the first few d_n are

$$
\begin{aligned}
d_1 &= 1/2, \\
d_2 &= z/8 - 1/20, \\
d_3 &= z^2/48 - z/40 + 1/105, \\
d_4 &= z^3/384 - z^2/160 + 101z/16800 - 3/1400, \\
d_5 &= z^4/3840 - z^3/960 + 61z^2/33600 - 13z/8400 + 1/1925, \\
d_6 &= z^5/46080 - z^4/7680 + 143z^3/403200 - 59z^2/112000 \\
&\quad + 7999z/19404000 - 691/5255250, \\
d_7 &= z^6/645120 - z^5/76800 + 41z^4/806400 - 11z^3/96000 \\
&\quad + 5941z^2/38808000 - 2357z/21021000 + 6/175175, \\
d_8 &= z^7/10321920 - z^6/921600 + 37z^5/6451200 - 73z^4/4032000 \\
&\quad + 224137z^3/6209280000 - 449747z^2/10090080000 \\
&\quad + 52037z/1681680000 - 10851/1191190000.
\end{aligned} \tag{2.4}
$$

3. Asymptotic expansion of $I(a, b, x)$

Following Wise [10] and DiDonato and Morris [3], we transform the expression for $I(a, b, x)$ in (1.1) in the same way as in (2.1) to obtain

$$
I(a, b, x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_{-\log x}^{\infty} e^{-\gamma t} t^{b-1} \left(\frac{\sinh(t/2)}{t/2} \right)^{b-1} dt, \tag{3.1}
$$

where as before $\gamma = a + (b-1)/2$. From (3.1), Wise [10] and DiDonato and Morris [3] proceeded by expanding $[\sinh(t/2)/(t/2)]^{b-1}$ as in §2 to obtain

$$
I(a, b, x) \sim \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)\gamma^b} \sum_0^{\infty} \Gamma(b+2n) Q(b+2n, -\gamma \log x) c_n / \gamma^{2n}, \tag{3.2}
$$

where $Q(c, z)$ is the incomplete Gamma function ratio [2] and [1, p.260 and p.941], and the coefficients c_n are defined in §2.

In their subroutine BGRAT, DiDonato and Morris [3] use (3.2) directly, with $Q(b + 2n, -\gamma \log x)$ and c_n being determined recursively.

We can proceed by using the recurrence relations for $Q(c, z)$ to express $Q(b + 2n, -\gamma \log x)$ in terms of $Q(b, -\gamma \log x)$. This gives

$$
I(a, b, x) \sim Q(b, -\gamma \log x) + R(a, b, x), \tag{3.3}
$$

where we have used (2.2) to cancel out the factors multiplying Q. The other term $R(a, b, x)$ is a double summation over n and the $2n$ residual terms obtained by expressing $Q(b + 2n, -\gamma \log x)$ in terms of $Q(b, -\gamma \log x)$. We can obtain the asymptotic expansion we require by reordering this sum. We shall however proceed differently. First we write (3.1) in the form

$$
\begin{aligned}
I(a, b, x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \Bigg[&\int_{-\log x}^{\infty} e^{-\gamma t} \left(\left(2 \sinh(t/2)\right)^{b-1} - t^{b-1} \right) dt \\
&+ \int_{-\log x}^{\infty} e^{-\gamma t} t^{b-1} dt \Bigg].
\end{aligned} \tag{3.4}
$$

The first of these integrals is integrated by parts twice to give

$$\begin{aligned}(3.5)\qquad &\frac{1}{\gamma^2}\int_{-\log x}^{\infty} e^{-\gamma t}\frac{d^2}{dt^2}\left([2\sinh(t/2)]^{b-1}-t^{b-1}\right)dt\\ &+\frac{x^\gamma}{\gamma}\left[[2\sinh(t/2)]^{b-1}-t^{b-1}+\frac{1}{\gamma}\frac{d}{dt}\left([2\sinh(t/2)]^{b-1}-t^{b-1}\right)\right]\Bigg|_{t=-\log x}.\end{aligned}$$

In the integral in (3.5) we now subtract the second term in the expansion of $[2\sinh(t/2)]^{b-1}$ and add a corresponding integral so that the integral in (3.5) becomes

$$\begin{aligned}(3.6)\qquad &\int_{-\log x}^{\infty} e^{-\gamma t}\frac{d^2}{dt^2}([2\sinh(t/2)]^{b-1}-t^{b-1}-c_1t^{b+1})\,dt\\ &+\frac{\Gamma(b+2)}{\Gamma(b)}c_1\int_{-\log x}^{\infty} e^{-\gamma t}t^{b-1}\,dt.\end{aligned}$$

The first of these integrals is then integrated by parts twice producing two further integrated terms evaluated at $t=-\log x$ and an integral of a fourth derivative. In this integral, a further term from the expansion of $[2\sinh(t/2)]^{b-1}$, c_2t^{b+3} is subtracted from the differentiated part and a corresponding integral added on separately. This procedure is continued indefinitely. The separate integrals starting from the ones on the right of (3.4) and (3.6) add together to give $Q(b,-\gamma\log x)$ as in (3.3) so that

$$(3.7)\qquad I(a,b,x)\quad\sim\quad Q(b,-\gamma\log x)+\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^\gamma\sum_{n=0}^{\infty}T_n(b,x)/\gamma^{n+1},$$

where

$$\begin{aligned}(3.8)\qquad T_n(b,x)&=\frac{d^n}{dt^n}\left([2\sinh(t/2)]^{b-1}-\sum_{m=0}^{n/2}c_mt^{2m+b-1}\right)\Bigg|_{t=-\log x}\\ &=\frac{d^n}{dt^n}\left(\sum_{m=n/2+1}^{\infty}c_mt^{2m+b-1}\right)\Bigg|_{t=-\log x},\end{aligned}$$

where $n/2$ in the summation is to be interpreted as the largest integer $\le n/2$ as in integer division. The quantities T_n satisfy the simple recurrence formulae

$$(3.9)\qquad T_{2n+1}=\frac{d}{dt}T_{2n},\qquad\qquad T_{2n}=\frac{d}{dt}T_{2n-1}-c_nt^{b-1}\frac{\Gamma(2n+b)}{\Gamma(b)}.$$

We can express $T_n(b,x)$ directly in terms of b and x, for example, $T_0(b,x)=(1/\sqrt{x}-\sqrt{x})^{b-1}-(-\log x)^{b-1}$. However, for x close to 1.0, evaluation of T_n in this way can lead to large rounding errors on subtraction, and so $T_n(b,x)$ is better evaluated from its power series expansion in t. The radius of convergence for this expansion is $t\approx 4.35464$, which corresponds to $x=0.0128$. Therefore, for $x>0.5$, t is well within this radius of convergence and so $T_n(b,x)$ converges quite rapidly. It follows also that the largest errors in T_n occur for the smallest x-values.

From (2.4) we see that for $0<b<1$, c_1 is negative, so that the first two terms in the series of (3.7) are negative; c_2 is positive so that the next two terms are positive, and so on. Thus, if we take an even number of terms in (3.7), we would expect the error to not exceed the contribution of the first two neglected terms.

We should note that the expansion (3.7) remains valid for $b>1$ and may well form the basis for a useful algorithm for evaluating $I(a,b,x)$ for $b\ll a$.

4. Implementation

To design an algorithm based on (3.7) and (3.8), it is useful to have an approximate idea of the relative magnitudes of $Q(b, -\gamma \log x)$ and $R(a, b, x)$. From the asymptotic expansion of $Q(b, z)$, [1, p.263] , we can see that in the region where this expansion is applicable, $R(a, b, x)/Q(b, -\gamma \log x)$ is $\sim c_1(\log x)^2$, and thus that the magnitude of this ratio decreases as x increases. Extensive computer experiments show that even in the region where the asymptotic approximation for Q is not valid, this magnitude continues to decrease as x increases. For $x > 0.5$, the ratio $R/Q \lesssim (b-1)/49.95$. Thus, for T_0 sufficiently many terms should be taken to ensure convergence to the required accuracy for the smallest value of x required. For T_n when $n > 0$, the required accuracy will be determined by the value of γ being considered. It is relatively straightforward to obtain bounds for T_n when $n > 0$, and therefore to determine how many of the T_n to take.

In a simple implementation in Ada in 15-digit floating-point arithmetic, with eight terms in T_0 and T_1, seven terms in T_2 and T_3, etc., up to one term in T_{14} and T_{15}, that is including all the terms up to c_8, in the expansions of T_0 up to T_{15}, a maximum relative error of $3.8E-14$ was obtained for the sum of the T_n terms in tests over 2,530 points in the parameter subdomain $15.0 < a < 39.0$, $0 < b < 1$ and $0.5 < x < 1$. This maximum error was obtained at the smallest values of a, b and x. It was found that the error decreased as a increased, as b increased and as x increased. This error would contribute an amount not exceeding $7.7E-16$ to the relative error of the algorithm as a whole. Inclusion of the terms involving c_9 in T_0 to T_{11} reduced the greatest relative error in the sum of the T_n terms to $\approx 1.2E-15$. The error in the above tests was determined by comparing the approximate value with values for $I(a, b, x)$ and Q, determined from series expansions evaluated using an accurate arithmetic package produced by Doman, Pursglove and Coen [4]. The Ada implementation of this algorithm was timed in comparison with an Ada implementation of BGRAT [3]. It was found that the part calculating the correction term $R(a, b, x)$ took about $1/15$ th of the time taken by the recursion part of BGRAT.

In practice, because of the small size of R relative to Q, the actual error produced by this algorithm is dominated by the error arising from Q. In the asymptotic region for $Q(a, z)$, it can be seen from the asymptotic expansion [1, p.263] that if Q' is the rate of change of Q with respect to z, $Q'/Q \sim 1$ in magnitude. The relative error $\delta Q/Q$ will then be $(Q'/Q)\delta(\gamma \log x)$, or γ times the absolute error of $\log x$. This can be the major contribution to the total error of the algorithm. For example, for $a = 55.1$, $b = 0.5$ and $x = 0.5$, an absolute error in $\log x$ of $\approx 1.0E-16$ would give a relative error $\delta Q/Q$ of $\approx 5.5E-15$. The actual relative error obtained from the algorithm of $\approx 1.067E-14$ is of the order one would expect from this argument. One can also see that as a increases this effect will become more pronounced.

A simple implementation of this algorithm in Ada is available by E-Mail to anyone who is interested.

Acknowledgement

The author is grateful to the the Teaching Company Directorate and to N.A. Software Ltd. for their support in this work, particularly in the development of the accurate arithmetic package.

References

1. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1970.
2. A. R. Didonato and A. H. Morris, *Computation of the Incomplete Gamma Function Ratios and their Inverse*, ACM Trans. Math. Software **12** (1986), 377-393.
3. ———, *Significant Digit Computation of the Incomplete Beta Function Ratios*, ACM Trans. Math. Software **18** (1992), 360-373.
4. B. G. S. Doman, C. J. Pursglove and W. M. Coen, *A Set of Ada Packages for High Precision Calculations*, ACM Trans. Math. Software **21** (1995), 416-431.
5. J. L. Fields, *A Note on the Asymptotic Expansion of the Ratio of Two Gamma Functions*, Proc. Edinburgh Math. Soc. **15** (1966), 43-55. MR **34:**379
6. C. L. Frenzen, *Error Bounds for Asymptotic Expansions of the Ratio of Two Gamma Functions*, SIAM J. Math. Anal. **18** (1987), 890-896. MR **88d:**33001
7. Y. L. Luke, *The Special Functions and their Approximations*, Vol. I, Academic Press, New York, 1969. MR **39:**3039
8. E. C. Molina, *Expansions for Laplacian Integrals in Terms of Incomplete Gamma Functions*, International Congress of Mathematicians, Zurich, Bell System Technical Journal **11** (1932), 563-575 and Monograph B704.
9. N. M. Temme, *Incomplete Laplace Integrals: Uniform Asymptotic Expansion with Application to the Incomplete Beta Function*, SIAM J. Math. Anal **18** (1987), 1638-1663. MR **89f:**41036
10. M. E. Wise, *The use of the Negative Binomial Distribution in an Industrial Sampling Problem*, Suppl. J. Roy. Statist. Soc. **8** (1946), 202-211. MR **9:**49c
11. ———, *The Incomplete Beta Function as a Contour Integral and a Quickly Converging Series for its Inverse*, Biometrika **37** (1950), 208-218. MR **12:**724e

Department of Mathematical Sciences, University of Liverpool, PO Box 147, Liverpool L69 3BX, England

E-mail address: `doman@liv.ac.uk`