1991

N.M. Temme

Asymptotic inversion of the incomplete beta function

Department of Analysis, Algebra and Geometry Report AM-R9106 July

CWI, nationaal instituut voor onderzoek op het gebied van wiskunde en informatica

CWI is the research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a non-profit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands organization for scientific research (NWO).

Asymptotic Inversion of the Incomplete Beta Function

N. M. Temme

CWI, P.O. Box 4079, 1009 AB Amsterdam, The Netherlands
e-mail: nicot@cwi.nl

The normalized incomplete beta function $I_x(a,b)$ is inverted for large values of the parameter a and b. That is, $x-$solutions of the equation

$$I_x(a,b)=p,\quad p\in[0,1]$$

are considered, especially for large values of a and b. The approximations are obtained by using uniform asymptotic expansions of the incomplete beta function, in which an error function or an incomplete gamma function is the dominant term. The inversion problem is started by inverting this dominant term. Further terms in the expansion are obtained by using standard perturbation methods, which are recently introduced in a paper describing a method for asymptotic inversion of the incomplete gamma functions. Numerical results indicate that for obtaining an accuracy of four correct digits the asymptotic method can already be used for $a+b\geq 5$.

1980 Mathematics Subject Classification: 33A15, 33A20, 41A60.

Keywords & Phrases: Incomplete beta function, inversion of beta distribution, asymptotic expansion.

1. Introduction

The incomplete beta function is defined by

$$I_x(a,b)=\frac{1}{B(a,b)}\int_0^x t^{a-1}(1-t)^{b-1}\,dt,\quad a>0,\quad b>0, \tag{1.1}$$

where $B(a,b)$ is Euler's (complete) beta function

$$B(a,b)=\int_0^1 t^{a-1}(1-t)^{b-1}\,dt=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}. \tag{1.2}$$

We consider the following inversion problem. Let $p\in[0,1]$ be given. We are interested in the $x-$value that solves the equation

$$I_x(a,b)=p, \tag{1.3}$$

Report AM-R9106
CWI
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

where a and b are fixed positive numbers. We are especially interested in solving (1.3) for large values of a and b.

This problem is of great importance in probability theory and mathematical statistics. The incomplete beta function is a standard probability function, with as special cases the (negative) binomial distribution, Student's distribution, and the $F-$(variance-ratio) distribution. Several approaches are available in the (statistical) literature, where often a first approximation of x, based on asymptotic estimates, is constructed, but this first approximation is not always reliable. Higher approximations may obtained by numerical inversion techniques, which require evaluation of the incomplete beta function. This may be rather time consuming, especially when a and b are large.

In the present method we also use an asymptotic result. The approximation is quite accurate, especially when the parameters a and b are large. It follows from numerical results, however, that a three term asymptotic expansion already gives an accuracy of 4 significant digits for $a + b \geq 5$, uniformly with respect to $p \in [0, 1]$.

The method of this paper is used earlier in [4] for the asymptotic inversion of the incomplete gamma functions. The present problem is more difficult, of course, since now two large parameters are considered. In fact we consider three asymptotic representations of the incomplete beta functionwith $a + b \to \infty$, valid in the following cases:

(i) $a = b + \beta$, where β stays fixed,

(ii) a/b and b/a are bounded away from zero,

(iii) at least one of the parameters a, b is large.

In the first two cases both parameters are large, in the third case we allow one parameter to be fixed or substantially smaller than the other one. In the first two cases the underlying beta distribution can be approximated by a normal (Gaussian) distribution, and we use an error function as main approximant. In the third case the distribution may be quite skew, and we consider an approximation in terms of the gamma distribution, with an incomplete gamma function as main approximant. It is possible to restrict ourselves to $a \geq b$, since we have the relation

$$I_x(a, b) = 1 - I_{1-x}(b, a). \tag{1.4}$$

This relation is used in the third case, where the only condition is that the sum $a + b$ should be large.

2. The nearly symmetric case.

We write $b = a + \beta$, where β is fixed. We obtain from (1.1)

$$I_x(a, a + \beta) = \frac{4^{-a}}{B(a, a + \beta)} \int_0^x [4t(1 - t)]^a \frac{(1 - t)^\beta \, dt}{t(1 - t)}.$$

We transform this to a standard form with a Gaussian character by writing

$$\begin{aligned} -\tfrac{1}{2}\zeta^2 &= \ln[4t(1 - t)], \quad 0 < t < 1, \quad \operatorname{sign}(\zeta) = \operatorname{sign}(t - \tfrac{1}{2}), \\ -\tfrac{1}{2}\eta^2 &= \ln[4x(1 - x)], \quad 0 < x < 1, \quad \operatorname{sign}(\eta) = \operatorname{sign}(x - \tfrac{1}{2}). \end{aligned} \tag{2.1}$$

Therefore

$$I_x(a, a+\beta) = \frac{4^{-a}}{B(a, a+\beta)} \int_{-\infty}^{\eta} e^{-\frac{1}{2}a\zeta^2} \frac{(1-t)^\beta}{t(1-t)} \frac{dt}{d\zeta}\, d\zeta.$$

We can write t as function of ζ:

$$t = \tfrac{1}{2}\left[1 \pm \sqrt{1-\exp(-\tfrac{1}{2}\zeta^2)}\right] = \tfrac{1}{2}\left[1 + \zeta\sqrt{[1-\exp(-\tfrac{1}{2}\zeta^2)]/\zeta^2}\right], \tag{2.2}$$

where the second square root is non-negative for real values of the argument. The same relation holds for x as function of η. It easily follows that

$$\frac{1}{t(1-t)}\frac{dt}{d\zeta} = \frac{-\zeta}{1-2t},$$

and that the following standard form (in the sense of [2]) can be obtained

$$I_x(a, a+\beta) = \sqrt{\frac{a}{2\pi}} \int_{-\infty}^{\eta} e^{-\frac{1}{2}a\zeta^2} f(\zeta)\, d\zeta, \tag{2.3}$$

with

$$f(\zeta) = \Phi(a)\phi(\zeta), \tag{2.4}$$

where

$$\Phi(a) = \frac{1}{\sqrt{a}} \frac{\Gamma(a+\frac{1}{2}\beta)}{\Gamma(a)} \frac{\Gamma(a+\frac{1}{2}\beta+\frac{1}{2})}{\Gamma(a+\beta)}, \quad \phi(\zeta) = [2(1-t)]^\beta \sqrt{\frac{\frac{1}{2}\zeta^2}{1-\exp(-\frac{1}{2}\zeta^2)}}. \tag{2.5}$$

This form of $\Phi(a)$ is obtained by using the duplication formula of the gamma function:

$$\sqrt{\pi}\,\Gamma(2z) = 2^{2z-1}\Gamma(z)\Gamma(z+\tfrac{1}{2}).$$

From the asymptotic expansion of the ratio of gamma functions (see for instance [1, formula 6.1.47]), we obtain

$$\Phi(a) \sim c_0 + c_1 a^{-1} + c_2 a^{-2} + \ldots \qquad (a \to \infty), \tag{2.6}$$

where

$$c_0 = 1, \quad c_1 = \tfrac{1}{8}(-2\beta^2 + 2\beta - 1), \quad c_2 = \tfrac{1}{128}(4\beta^4 + 8\beta^3 - 16\beta^2 + 4\beta + 1).$$

The function $\phi(\zeta)$ is analytic in a strip containing $\mathbb{R}$; the singularities nearest to the origin occur at $\pm 2\sqrt{\pi}\exp(\pm i\pi/4)$. The first coefficients of the Taylor expansion

$$\phi(\zeta) = d_0 + d_1\zeta + d_2\zeta^2 + d_3\zeta^3 + \ldots \tag{2.7}$$

are

$$d_0 = 1, \quad d_1 = -\frac{\beta\sqrt{2}}{2}, \quad d_2 = \frac{1}{8}(2\beta^2 - 2\beta + 1), \quad d_3 = -\frac{\beta\sqrt{2}}{24}(\beta^2 - 3\beta + 2),$$

$$d_4 = \tfrac{1}{384}(4\beta^4 - 24\beta^3 + 32\beta^2 - 12\beta + 1), \quad d_5 = -\frac{\beta\sqrt{2}}{960}(\beta^4 - 10\beta^3 + 25\beta^2 - 20\beta + 4).$$

From results in [2] it follows that the standard form (2.3) can be written in the form

$$I_x(a, a+\beta) = \tfrac{1}{2}\mathrm{erfc}(-\eta\sqrt{a/2}) - R_a(\eta), \tag{2.8}$$

where η is defined in (2.1), and erfc is the error function defined by

$$\mathrm{erfc}\ z = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-t^2}\, dt.$$

We try to solve equation (1.3) with the above representation of the incomplete beta function. First we solve the equation (1.3) in terms of η; afterwards, we determine x from the inverse relation of the second line in (2.1) (that is, (2.2) with t, ζ replaced by x, η, respectively). When a is large, we consider the error function in the equation

$$I_x(a, a+\beta) = \tfrac{1}{2}\mathrm{erfc}(-\eta\sqrt{a/2}) - R_a(\eta) = p \tag{2.9}$$

as the dominant term, and a first approximation η_0 of η is defined by the solution of the equation

$$\tfrac{1}{2}\mathrm{erfc}(-\eta_0\sqrt{a/2}) = p. \tag{2.10}$$

The exact solution of (1.3) (in terms of η) is written as

$$\eta = \eta_0 + \varepsilon, \tag{2.11}$$

and we try to determine ε. It appears that we can expand this quantity in the form

$$\varepsilon \sim \frac{\varepsilon_1}{a} + \frac{\varepsilon_2}{a^2} + \frac{\varepsilon_3}{a^3} + \ldots, \tag{2.12}$$

as $a \to \infty$. The coefficients ε_i can be expressed in terms of η_0 and β.

From (2.3), (2.9) and (2.10) we obtain

$$\frac{dp}{d\eta_0} = \sqrt{\frac{a}{2\pi}}\, e^{-\frac{1}{2}a\eta_0^2}, \quad \frac{dp}{d\eta} = \sqrt{\frac{a}{2\pi}}\, f(\eta) e^{-\frac{1}{2}a\eta^2}, \tag{2.13}$$

where f is given in (2.4). Upon dividing, we obtain

$$f(\eta)\frac{d\eta}{d\eta_0} = e^{\frac{1}{2}a(\eta^2 - \eta_0^2)}.$$

Substitution of (2.11) gives the differential equation

$$f(\eta_0 + \varepsilon)\left(1 + \frac{d\varepsilon}{d\eta_0}\right) = e^{a\varepsilon(\eta_0 + \frac{1}{2}\varepsilon)}.$$

We write η in place of η_0; that is, we try to find ε as function of η, that satisfies (see also (2.4))

$$\phi(\eta + \varepsilon)\Phi(a)\left(1 + \frac{d\varepsilon}{d\eta}\right) = e^{a\varepsilon(\eta + \frac{1}{2}\varepsilon)}. \tag{2.14}$$

When we have obtained ε from this equation (or an approximation), we use it in (2.11) to obtain the final value η.

The first coefficient ε_1 of (2.12) is obtained by comparing dominant terms in (2.14). Since $\Phi(a) = 1 + \mathcal{O}(a^{-1})$, we obtain

$$\varepsilon_1 = \frac{1}{\eta} \ln \phi(\eta). \tag{2.15}$$

This quantity is analytic (as a function of η) on $\mathbb{R}$; $\phi(\eta)$ is positive on $\mathbb{R}$, and $\phi(0) = 1$. Using (2.7) we obtain for small values of η

$$\varepsilon_1 = -\tfrac{1}{2}\beta\sqrt{2} + \tfrac{1}{8}(1-2\beta)\eta - \tfrac{1}{48}\beta\sqrt{2}\,\eta^2 - \tfrac{1}{192}\eta^3 - \tfrac{1}{3840}\beta\sqrt{2}\,\eta^4 + \dots .$$

Further terms ε_i can be obtained by using more terms in (2.6) and by expanding

$$\phi(\eta+\varepsilon) = \phi(\eta) + \varepsilon\phi'(\eta) + \tfrac{1}{2}\varepsilon^2\phi''(\eta) + \dots,$$

in which (2.12) is substituted to obtain an expansion in powers of a^{-1}. In this way we find

$$\varepsilon_2 = \frac{1}{2\eta\phi}(2\phi\varepsilon_1' + 2\phi'\varepsilon_1 + 2c_1\phi - \phi\varepsilon_1^2); \tag{2.16}$$

$$\begin{aligned}\varepsilon_3 = \frac{1}{8\eta\phi}(&8\phi\varepsilon_2' + 8\phi'\varepsilon_1\varepsilon_1' + 8c_1\phi\varepsilon_1' + 8\phi'\varepsilon_2 + 4\phi''\varepsilon_1^2 + \\ &8c_1\phi'\varepsilon_1 + 8c_2\phi - 8\phi\varepsilon_1\varepsilon_2 - 4\phi\varepsilon_2^2\eta^2 - 4\phi\varepsilon_2\eta\varepsilon_1^2 - \phi\varepsilon_1^4).\end{aligned} \tag{2.17}$$

The derivatives ϕ', ε', etc., are with respect to η, and all functions are evaluated at η. For small values of η we can use the Taylor series

$$\begin{aligned}\varepsilon_2 =& \frac{\beta\sqrt{2}}{12}(3\beta-2) + \frac{1}{128}(20\beta^2 - 12\beta + 1)\eta + \frac{\beta\sqrt{2}}{960}(20\beta-1)\eta^2 + \\ & \frac{1}{4608}(16\beta^2 + 30\beta - 15)\eta^3 + \frac{\beta\sqrt{2}}{53760}(21\beta + 32)\eta^4 + \\ & \frac{1}{368640}(-32\beta^2 + 63)\eta^5 - \frac{\beta\sqrt{2}}{25804480}(120\beta + 17)\eta^6 + \dots,\end{aligned}$$

$$\begin{aligned}\varepsilon_3 =& \frac{\beta\sqrt{2}}{480}(-75\beta^2 + 80\beta - 16) + \frac{1}{9216}(-1080\beta^3 + 868\beta^2 - 90\beta - 45)\eta + \\ & \frac{\beta\sqrt{2}}{53760}(-1190\beta^2 + 84\beta + 373)\eta^2 + \\ & \frac{1}{368640}(-2240\beta^3 - 2508\beta^2 + 2100\beta - 165)\eta^3 + \dots .\end{aligned}$$

3. The general error function case.

Let us write

$$a = r\sin^2\theta, \quad b = r\cos^2\theta, \quad 0 \le \theta \le \tfrac{1}{2}\pi.$$

Then (1.1) can be written as

$$I_x(a,b) = \frac{1}{B(a,b)} \int_0^x e^{r[\sin^2\theta \ln t + \cos^2\theta \ln(1-t)]} \frac{dt}{t(1-t)}. \tag{3.1}$$

We consider r as a large parameter, and θ bounded away from 0 and $\frac{1}{2}\pi$. The maximum of the exponential function occurs at $t = \sin^2\theta$. Hence, the following transformation brings the exponential part of the integrand into a Gaussian form:

$$-\tfrac{1}{2}\zeta^2 = \sin^2\theta \ln \frac{t}{\sin^2\theta} + \cos^2\theta \ln \frac{1-t}{\cos^2\theta}, \tag{3.2}$$

where the sign of ζ equals the sign of $t - \sin^2\theta$. The same transformation holds for $x \mapsto \eta$ if t and ζ are replaced with x and η, respectively. From (3.2) we obtain,

$$-\zeta\frac{d\zeta}{dt} = \frac{\sin^2\theta - t}{t(1-t)},$$

and we can write (3.1) in the standard form (confer (2.3)-(2.4))

$$I_x(a,b) = \sqrt{\frac{r}{2\pi}} \int_{-\infty}^{\eta} e^{-\frac{1}{2}r\zeta^2} f(\zeta)\, d\zeta, \tag{3.3}$$

with

$$f(\zeta) = \Phi(r)\phi(\zeta), \tag{3.4}$$

where

$$\Phi(r) = \frac{\Gamma^*(r)}{\Gamma^*(a)\Gamma^*(b)}, \quad \phi(\zeta) = \frac{\zeta \sin\theta\cos\theta}{t - \sin^2\theta}. \tag{3.5}$$

The function $\Gamma^*(z)$ is the slowly varying part of the Euler gamma function. That is,

$$\Gamma^*(z) = \sqrt{\frac{z}{2\pi}}\, e^z z^{-z}\Gamma(z), \quad z > 0, \tag{3.6}$$

with

$$\Gamma^*(z) \sim \sum_{n=0}^{\infty} (-1)^n \gamma_n z^{-n}, \quad \frac{1}{\Gamma^*(z)} \sim \sum_{n=0}^{\infty} \gamma_n z^{-n}, \quad z \to \infty. \tag{3.7}$$

The first few γ_n are

$$\gamma_0 = 1,\ \gamma_1 = -\tfrac{1}{12},\ \gamma_2 = \tfrac{1}{288},\ \gamma_3 = \tfrac{139}{51840}.$$

The analogue of the expansion (2.6) is now in terms of the large parameter r:

$$\Phi(r) \sim c_0 + c_1 r^{-1} + c_2 r^{-2} + \ldots \qquad (r \to \infty),$$

where

$$c_0 = 1, \quad c_1 = \frac{\sin^2\theta\cos^2\theta - 1}{3\sin^2 2\theta}, \quad c_2 = \frac{(\sin^2\theta\cos^2\theta - 1)^2}{18\sin^4 2\theta},$$

$$c_3 = -\frac{139(\sin^6\theta\cos^6\theta - \cos^6\theta - \sin^6\theta) + 15\sin^4\theta\cos^4\theta}{810\sin^6 2\theta}.$$

The first coefficients of the Taylor expansion

$$\phi(\zeta) = d_0 + d_1\zeta + d_2\zeta^2 + d_3\zeta^3 + \dots \tag{3.8}$$

are

$$d_0 = 1, \quad d_1 = -\tfrac{2}{3}\cot 2\theta, \quad d_2 = \frac{\sin^4\theta + \cos^4\theta + 1}{6\sin^2 2\theta}.$$

To solve equation (1.3) for large values of r, we use the method of the previous section. We write as in (2.8)

$$I_x(a,b) = \tfrac{1}{2}\mathrm{erfc}(-\eta\sqrt{r/2}\,) - R_a(\eta), \tag{3.9}$$

where the relation between x and η follows from (3.2). A first approximation η_0 follows from the equation

$$\tfrac{1}{2}\mathrm{erfc}(-\eta_0\sqrt{r/2}\,) = p, \tag{3.10}$$

and the terms ε_i in the expansion

$$\varepsilon \sim \frac{\varepsilon_1}{r} + \frac{\varepsilon_2}{r^2} + \frac{\varepsilon_3}{r^3} + \dots,$$

are the same as in (2.15)-(2.17), but with ϕ, c_1, c_2 of the present section. For small values of η we can expand

$$\begin{aligned}
\varepsilon_1 = \; & \frac{2s^2-1}{3sc} - \frac{5s^4-5s^2-1}{36s^2c^2}\eta + \frac{46s^6-69s^4+21s^2+1}{1620s^3c^3}\eta^2 \\
& - \frac{-2s^2-62s^6+31s^8+33s^4+7}{6480s^4c^4}\eta^3 \\
& + \frac{88s^6-52s^2-115s^8+46s^{10}-17s^4+25}{90720s^5c^5}\eta^4 + \dots, \\
\varepsilon_2 = & -\frac{52s^6-78s^4+12s^2+7}{405s^3c^3} + \frac{2s^2-370s^6+185s^8+183s^4-7}{2592s^4c^4}\eta \\
& - \frac{776s^2+10240s^6-13525s^8-533+5410s^{10}-1835s^4}{204120s^5c^5}\eta^2 \\
& + \frac{3747s^2+15071s^{12}-15821s^6+45588s^8-45213s^{10}-3372s^4-1579}{2099520s^6c^6}\eta^3 + \dots,
\end{aligned}$$

$$\begin{aligned}
\varepsilon_3 = \; & (3704s^{10}-9260s^8+6686s^6-769s^4-1259s^2+449)/(102060s^5c^5) \\
& -(750479s^{12}-151557s^2-727469s^6+2239932s^8 \\
& \quad -2251437s^{10}+140052s^4+63149)/(20995200s^6c^6)\eta \\
& +(729754s^{14}-78755s^2-2554139s^{12}+146879s^6-1602610s^8 \\
& \quad +3195183s^{10}+105222s^4++29233)/(367416000s^7c^7)\eta^2 + \dots\;,
\end{aligned}$$

where $s = \sin\theta, c = \cos\theta$.

The functions ε_i are now considered as functions of η_0 (instead of η), and we write

$$\eta \sim \eta_0 + \frac{\varepsilon_1}{r} + \frac{\varepsilon_2}{r^2} + \frac{\varepsilon_3}{r^3} + \cdots.$$

This approximation is substituted in the left hand side of (3.2), and we invert this equation to obtain t, or, equivalently, x.

4. The incomplete gamma function case.

In this section we consider the asymptotic condition that the sum $a + b$ should be large. We concentrate on the case $a \geq b$. In the other case we can solve (1.3) by using (1.4). From [3, formula (9.16)] it follows that we can write

$$I_x(a, b) = Q(b, \eta a) + R_{a,b}(\eta), \tag{4.1}$$

where η is given by a mapping $x \mapsto \eta$, which is defined by

$$\eta - \mu \ln \eta + A(\mu) = -\ln x - \mu \ln(1 - x), \tag{4.2}$$

and

$$\mu = \frac{b}{a}, \quad A(\mu) = (1 + \mu)\ln(1 + \mu) - \mu. \tag{4.3}$$

Q is the incomplete gamma function defined by

$$Q(\alpha, z) = \frac{1}{\Gamma(\alpha)} \int_z^\infty t^{\alpha-1} e^{-t}\, dt, \quad \alpha > 0. \tag{4.4}$$

Corresponding points in the mapping are

$$x = 0 \leftrightarrow \eta = +\infty, \quad x = \frac{1}{1+\mu} \leftrightarrow \eta = \mu, \quad x = 1 \leftrightarrow \eta = 0.$$

From (4.2) it follows that

$$\frac{dx}{d\eta} = \frac{\eta - \mu}{\eta} \frac{x(1 - x)}{(1 + \mu)x - 1}. \tag{4.5}$$

In [3] an asymptotic expansion of $R_{a,b}(\eta)$ in (4.1) is derived, which holds for $a \to \infty$, uniformly with respect to $x \in [0, 1]$ and $b \in [0, \infty)$.

We obtain the solution of equation (1.3) for large values of a, by first determining η_0, the solution of the reduced equation

$$Q(b, \eta_0 a) = p. \tag{4.6}$$

This involves an inversion of the incomplete gamma function, which problem is considered in [4], especially for large values of b. As in the previous sections, the exact solution of (1.3) is written as $\eta = \eta_0 + \varepsilon$, and we expand ε as in (2.12). We have (confer (2.13)),

$$\frac{dp}{d\eta_0} = -\frac{a^b}{\eta_0 \Gamma(b)} e^{a(-\eta_0 + \mu \ln \eta_0)}, \quad \frac{dp}{d\eta} = \frac{1}{B(a,b)x(1-x)} \frac{dx}{d\eta} e^{a[-\eta + \mu \ln \eta - A(\mu)]}.$$

Upon dividing these equations and using (4.5), we obtain

$$f(\eta)\frac{d\eta}{d\eta_0} = \frac{\eta}{\eta_0}e^{a[\eta-\eta_0-\mu\ln(\eta/\eta_0)]}, \tag{4.7}$$

with $f(\eta) = \phi(\eta)\Phi(a)$, and

$$\phi(\eta) = \frac{\eta-\mu}{1-x(1+\mu)}\,\frac{1}{\sqrt{1+\mu}}, \qquad \Phi(a) = \frac{\Gamma^*(a+b)}{\Gamma^*(a)},$$

where Γ^* is introduced in (3.6). By writing $\eta = \eta_0 + \varepsilon$, and writing η in place of η_0 (for the time being), (4.7) can be written as

$$\phi(\eta+\varepsilon)\Phi(a)\left(1+\frac{d\varepsilon}{d\eta}\right) = e^{a[\varepsilon-\mu\ln(1+\varepsilon/\eta)]}. \tag{4.8}$$

The analogue of the expansion (2.6) has coefficients

$$c_0 = 1, \quad c_1 = -\frac{\mu}{12(1+\mu)}, \quad c_2 = \frac{\mu^2}{288(1+\mu)^2},$$

$$c_3 = \frac{\mu(432+432\mu+139\mu^2)}{51840(1+\mu)^3}.$$

The analogue of (2.7) reads

$$\phi(\eta) = d_0 + d_1(\eta-\mu) + d_2(\eta-\mu)^2 + \ldots,$$

with coefficients

$$d_0 = 1, \quad d_1 = \frac{w+2}{3(w+1)w}, \quad d_2 = \frac{1}{12w^2}, \quad d_3 = \frac{8w^3+9w^2-9w-8}{540w(w+1)^3},$$

$$d_4 = \frac{15w^4-68w^3-182w^2-68w+15}{12960w^4(w+1)^4},$$

$$d_5 = -\frac{32w^5+265w^4+253w^3-253w^2-265w-32}{90720w^5(w+1)^5},$$

where

$$w = \sqrt{1+\mu}. \tag{4.9}$$

Substituting

$$\varepsilon \sim \frac{\varepsilon_1}{a} + \frac{\varepsilon_2}{a^2} + \frac{\varepsilon_3}{a^3} + \ldots, \tag{4.10}$$

into (4.8), we find the first coefficient:

$$\varepsilon_1 = \frac{\ln\phi(\eta)}{1-\mu/\eta},$$

a regular function at $\eta = \mu$, as follows from the expansion of $\phi(\eta)$ at this point. The next terms are

$$\varepsilon_2 = \frac{1}{2\phi\eta(\eta-\mu)}(2\phi\varepsilon_1'\eta^2 + 2\phi'\varepsilon_1\eta^2 + 2c_1\phi\eta^2 - \phi\mu\varepsilon_1^2 - 2\varepsilon_1\phi\eta),$$

$$\begin{aligned}\varepsilon_3 = \frac{1}{6\eta^2\phi^2(\eta-\mu)^2}(&-3\phi^2\varepsilon_1'^2\eta^4 - 6c_1\phi^2\eta^3 - 6\phi^2\varepsilon_1'\eta^3 + 6\eta^4 c_2\phi^2 + 6\varepsilon_1\phi^2\eta^2 - \\ &3\phi'^2\varepsilon_1^2\eta^4 - 3c_1^2\phi^2\eta^4 + 3\varepsilon_1^2\phi^2\eta^2 + 6\eta^4\phi'^2\varepsilon_1 + 2\mu\varepsilon_1^3\eta\phi^2 - \\ &6\varepsilon_1\phi^2\eta^2\mu\varepsilon_1' - 9\varepsilon_1^2\phi\eta^2\mu\phi' - 6\varepsilon_1\phi^2\eta^2\mu c_1 + 6\varepsilon_1^2\phi^2\eta\mu - \\ &12\phi'\varepsilon_1\eta^3\phi + \mu^2\phi^2\varepsilon_1^3 + 6\eta^4\varepsilon_2'\phi^2 + 3\eta^3\mu c_1^2\phi^2 - 6\eta^3\mu c_2\phi^2 + \\ &3\eta^3\mu\varepsilon_1'^2\phi^2 - 6\eta^3\mu\varepsilon_2'\phi^2 + 6c_1\eta^4\phi'\phi + 6\varepsilon_1'\eta^4\phi'\phi + \\ &3\eta^4\phi''\phi\varepsilon_1^2 - 3\eta^3\mu\phi''\phi\varepsilon_1^2 + 3\eta^3\mu\phi'^2\varepsilon_1^2),\end{aligned}$$

where the derivatives are respect with η. For small values of $|\eta - \mu|$ we can expand

$$\begin{aligned}\varepsilon_1 = &\frac{(w+2)(w-1)}{3w} + \frac{w^3 + 9w^2 + 21w + 5}{36w^2(w+1)}(\eta-\mu) - \\ &\frac{w^4 - 13w^3 + 69w^2 + 167w + 46}{1620(w+1)^2w^3}(\eta-\mu)^2 - \\ &\frac{7w^5 + 21w^4 + 70w^3 + 26w^2 - 93w - 31}{6480(w+1)^3w^4}(\eta-\mu)^3 - \\ &\frac{75w^6 + 202w^5 + 188w^4 - 888w^3 - 1345w^2 + 118w + 138}{272160(w+1)^4w^5}(\eta-\mu)^4 + \ldots,\end{aligned}$$

$$\begin{aligned}\varepsilon_2 = &\frac{(28w^4 + 131w^3 + 402w^2 + 581w + 208)(w-1)}{1620(w+1)w^3} - \\ &\frac{35w^6 - 154w^5 - 623w^4 - 1636w^3 - 3983w^2 - 3514w - 925}{12960(w+1)^2w^4}(\eta-\mu) - \\ &(2132w^7 + 7915w^6 + 16821w^5 + 35066w^4 + 87490w^3 + 141183w^2 + \\ &\quad 95993w + 21640)/[816480w^5(w+1)^3](\eta-\mu)^2 - \\ &(11053w^8 + 53308w^7 + 117010w^6 + 163924w^5 + 116188w^4 - 258128w^3 - \\ &\quad 677042w^2 - 481940w - 105497)/[14696640(w+1)^4w^6](\eta-\mu)^3 + \ldots,\end{aligned}$$

$$\begin{aligned}\varepsilon_3 = &-[(3592w^7 + 8375w^6 - 1323w^5 - 29198w^4 - 89578w^3 - 154113w^2 - \\ &\quad 116063w - 29632)(w-1)]/[816480w^5(w+1)^2] - \\ &\quad(412043w^9 + 2054169w^8 + 3803094w^7 + 3470754w^6 + 2141568w^5 - \\ &\quad 2393568w^4 - 19904934w^3 - 34714674w^2 - \\ &\quad 23128299w - 5253353)/[146966400w^6(w+1)^3](\eta-\mu) - \\ &\quad(116932w^{10} + 819281w^9 + 2378172w^8 + 4341330w^7 + 6806004w^6 + \\ &\quad 10622748w^5 + 18739500w^4 + 30651894w^3 + 30869976w^2 + \\ &\quad 15431867w + 2919016)/[146966400(w+1)^4w^7](\eta-\mu)^2 + \ldots,\end{aligned}$$

where w is given by (4.9).

Considering the functions ε_i as functions of η_0, we obtain using (4.10)

$$\eta \sim \eta_0 + \frac{\varepsilon_1}{a} + \frac{\varepsilon_2}{a^2} + \frac{\varepsilon_3}{a^3} + \dots,$$

which is substituted in the left hand side of (4.2). Solving for x, we finally obtain the desired approximation of the solution of equation (1.3).

In this section, the functions $\Phi(a), \phi(\eta), \varepsilon$ have expansions with coefficients c_i, d_i, ε_i in which the parameter $\mu = b/a$ may assume any value in $[0, \infty)$. This aspect demonstrates the uniform character (with respect to μ) of the present approach. In §2 large values of β are not allowed, and in §3 the value of θ should be bounded away from 0 and $\frac{1}{2}\pi$. Of course, the transformations and expansions of this sections are more complicated than those in the previous sections. Moreover, to start the inversion procedure, first equation (4.6) including an incomplete gamma function should be solved, whereas in the foregoing cases only an error function has to be inverted. See (2.10) and (3.9).

5. Numerical aspects

In numerical applications one needs the inversion of the mappings given in (2.1), (3.2), and (4.2). Only (2.1) can be inverted directly, as shown in (2.2). For small values of $|\zeta|$ we have

$$t = \tfrac{1}{2} + \tfrac{1}{4}\sqrt{2}\,\zeta - \tfrac{1}{32}\sqrt{2}\,\zeta^3 + \tfrac{5}{1536}\sqrt{2}\,\zeta^5 + \dots.$$

The inversion of (4.2) can be based on that of (3.2), with other parameters. We give some details on the inversion of (3.2).

For small values of $|\zeta|$ we have

$$t = s^2 + sc\zeta + \frac{1-2s^2}{3}\zeta^2 + \frac{13s^4 - 13s^2 + 1}{36sc}\zeta^3 + \frac{46s^6 - 69s^4 + 21s^2 + 1}{270s^2c^2}\zeta^4 + \dots,$$

where $s = \sin\theta, c = \cos\theta$. For larger values of $|\zeta|$, with $\zeta < 0$, we rewrite(3.2) in the form

$$t(1-t)^\alpha = u, \quad \alpha = \cot^2\theta, \quad u = \exp[(-\tfrac{1}{2}\zeta^2 + s^2\ln s^2 + c^2\ln c^2)/s^2],$$

and for small values of u we expand

$$t = u + \alpha u + \frac{3\alpha(3\alpha+1)}{3!}u^3 + \frac{4\alpha(4\alpha+1)(4\alpha+2)}{4!}u^4 + \frac{5\alpha(5\alpha+1)(5\alpha+2)(5\alpha+3)}{5!}t^5 + \dots.$$

A similar approach is possible for positive values of ζ, giving an expansion for t near unity. The approximations obtained in this way may be used for starting a Newton-Raphson method for obtaining more accurate values of t.

We have tested the inversion process of the incomplete beta function for several values of the parameters. We describe the testing for the method of §2. After obtaining the first approximation η_0 by inverting (2.10), we computed the values of $\varepsilon_1, \varepsilon_2, \varepsilon_3$, defined in (2.15) - (2.17) (with η replaced with η_0). The coefficient ε_3 is included only when η_0 is small enough for using the Taylor series given §2. Next, (2.11) gives the final approximation of η, which is used in the second line in (2.1), to obtain the approximation of x. Finally we verified (1.3), by computing the incomplete beta function with this value x. We used the continued fraction given in [1, formula 26.5.8].

In the tables 5.1, 5.2, 5.3 we show the relative errors $|(I_x(a,b)-p)/p|$, where x is obtained by the asymptotic inversion methods of §§2-4. As is expected, it follows that the larger values of β give less accuracy in the results in Table 5.1. The same holds for smaller values of θ in Table 5.2. From Table 5.3 it follows that the results are not influenced by large or small values of μ. This shows the uniform character of the method of §4. In fact, this method can be used in extreme situations: the ratio a/b may be very small and very large, and p may assume values quite close to zero or to unity.

References

[1] M. Abramowitz & I.A. Stegun (1964), *Handbook of mathematical functions with formulas, graphs and mathematical tables*, Nat. Bur. Standards Appl. Series, **55**, U.S. Government Printing Office, Washington, D.C.

[2] N.M. Temme (1982), The uniform asymptotic expansion of a class of integrals related to cumulative distribution functions, SIAM J. Math. Anal. **13**, 239-253.

[3] N.M. Temme (1987), Incomplete Laplace integrals: uniform asymptotic expansions with application to the incomplete beta function, SIAM J. Math. Anal. **18**, 1638-1663.

[4] N.M. Temme (1991), Asymptotic inversion of incomplete gamma functions, CWI Report AM-R9102; to appear in Mathematics of Computation.

p \ β	0	1	2	3
10^{-4}	$7.4\,10^{-5}$	$7.4\,10^{-5}$	$5.3\,10^{-3}$	$2.3\,10^{-2}$
0.1	$9.0\,10^{-7}$	$7.5\,10^{-6}$	$4.0\,10^{-4}$	$2.7\,10^{-3}$
0.3	$4.8\,10^{-7}$	$3.4\,10^{-6}$	$3.0\,10^{-4}$	$2.1\,10^{-3}$
0.5	$2.1\,10^{-7}$	$2.7\,10^{-6}$	$2.3\,10^{-4}$	$1.6\,10^{-3}$
0.7	$2.1\,10^{-7}$	$2.6\,10^{-6}$	$1.7\,10^{-4}$	$1.1\,10^{-3}$
0.9	$3.7\,10^{-6}$	$4.7\,10^{-6}$	$8.1\,10^{-5}$	$5.3\,10^{-4}$
0.9999	$9.1\,10^{-4}$	$9.1\,10^{-4}$	$9.1\,10^{-4}$	$9.2\,10^{-4}$

Table 5.1. Relative errors $|I_x(a, a+\beta) - p|/p$ for $a = 10$ and several values of p and β; the asymptotic inversion is based on the method of §2

p \ $\sin^2\theta$	0.5	0.4	0.3	0.2
10^{-4}	$5.4\,10^{-4}$	$1.1\,10^{-3}$	$2.1\,10^{-3}$	$3.9\,10^{-3}$
0.1	$1.4\,10^{-6}$	$1.4\,10^{-5}$	$3.4\,10^{-5}$	$2.3\,10^{-3}$
0.3	$2.9\,10^{-7}$	$1.1\,10^{-5}$	$3.6\,10^{-5}$	$1.2\,10^{-4}$
0.5	$1.5\,10^{-7}$	$7.9\,10^{-6}$	$2.5\,10^{-5}$	$9.5\,10^{-5}$
0.7	$1.2\,10^{-7}$	$5.0\,10^{-6}$	$1.5\,10^{-5}$	$6.0\,10^{-5}$
0.9	$3.7\,10^{-6}$	$5.3\,10^{-6}$	$9.1\,10^{-6}$	$7.0\,10^{-5}$
0.9999	$9.1\,10^{-4}$	$9.1\,10^{-4}$	$9.1\,10^{-4}$	$9.1\,10^{-4}$

Table 5.2. Relative errors $|I_x(a, b) - p|/p$ for $r = a + b = 10$ and several values of p and $\sin^2\theta = a/r$; the asymptotic inversion is based on the method of §3

p \ μ	0.1	0.5	2.0	10
10^{-4}	$6.8\,10^{-5}$	$1.2\,10^{-4}$	$1.7\,10^{-4}$	$1.3\,10^{-4}$
0.1	$1.2\,10^{-6}$	$2.7\,10^{-7}$	$3.4\,10^{-7}$	$3.6\,10^{-6}$
0.3	$1.8\,10^{-6}$	$9.1\,10^{-7}$	$6.3\,10^{-7}$	$5.4\,10^{-6}$
0.5	$6.6\,10^{-7}$	$2.0\,10^{-7}$	$6.6\,10^{-7}$	$2.2\,10^{-7}$
0.7	$1.7\,10^{-7}$	$3.6\,10^{-7}$	$1.3\,10^{-6}$	$4.2\,10^{-6}$
0.9	$2.0\,10^{-7}$	$1.2\,10^{-7}$	$3.0\,10^{-7}$	$6.3\,10^{-7}$
0.9999	$1.3\,10^{-7}$	$1.8\,10^{-8}$	$1.7\,10^{-8}$	$1.9\,10^{-8}$

Table 5.3. Relative errors $|I_x(a, b) - p|/p$ for $a = 10$ and several values of p and $\mu = b/a$; the asymptotic inversion is based on the method of §4