ASYMPTOTIC INVERSION OF THE BINOMIAL AND NEGATIVE BINOMIAL CUMULATIVE DISTRIBUTION FUNCTIONS[*]

A. GIL[†], J. SEGURA[‡], AND N. M. TEMME[§]

Abstract. The computation and inversion of the binomial and negative binomial cumulative distribution functions play a key role in many applications. In this paper, we explain how methods used for the central beta distribution function (described in Gil, Segura, and Temme, [Numer. Algorithms, 74 (2017), pp. 77–91]) can be utilized to obtain asymptotic representations of these functions and also for their inversion. The performance of the asymptotic inversion methods is illustrated with numerical examples.

Key words. binomial cumulative distribution function, negative binomial cumulative distribution function, asymptotic representation, asymptotic inversion methods

AMS subject classifications. 33B20, 41A60

1. Introduction. The binomial and negative binomial distribution functions are used in many areas of science and engineering. In particular, the generation of random binomial variables plays a key role in simulation algorithms such as, for example, the stochastic spatial modeling of chemical reactions [4]. On the other hand, the negative binomial distribution is, for example, widely used in genomic research to model gene expression data arising from RNA-sequences; see, for example, [3, 5].

The binomial cumulative distribution function is defined by

$$(1.1) \qquad P(n,p,x) = \sum_{k=0}^{x} \binom{n}{k} p^k (1-p)^{n-k}, \quad 0 \le p \le 1,$$

with x and n positive integers, $x \le n$. The complementary function is

$$Q(n,p,x) = \sum_{k=x+1}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1 - P(n,p,x).$$

The negative binomial cumulative distribution function (also called Pascal distribution) is given by

$$P^{NB}(r,p,x) = \sum_{k=0}^{x} \binom{k+r-1}{r-1} p^r (1-p)^k, \quad 0 \le p \le 1,$$

with x and r positive integers. The complementary function, denoted by $Q^{NB}(r,p,x)$, satisfies $Q^{NB}(r,p,x) = 1 - P^{NB}(r,p,x)$. The definition of the negative binomial distribution can be extended to the case where the parameter r takes positive real values. In this case, the distribution is called Pólya distribution.

These functions are particular cases of the cumulative central beta distribution. This distribution function (also known as the incomplete beta function) is defined by

$$(1.2) \qquad I_y(a,b) = \frac{1}{B(a,b)} \int_0^y t^{a-1}(1-t)^{b-1}\, dt,$$

[†]Departamento de Matemática Aplicada y CC. de la Computación. ETSI Caminos. Universidad de Cantabria. 39005-Santander, Spain (amparo.gil@unican.es).
[‡]Departamento de Matemáticas, Estadistica y Computación. Universidad de Cantabria, 39005 Santander, Spain.
[§]IAA, 1825 BD 25, Alkmaar, The Netherlands. Former address: Centrum Wiskunde & Informatica (CWI), Science Park 123, 1098 XG Amsterdam, The Netherlands.

where we assume that a and b are real positive parameters and $0 \le y \le 1$. $B(a,b)$ is the Beta function

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

The relation between the binomial and the central beta distribution functions is the following:

$$(1.3) \qquad P(n,p,x) = I_{1-p}(n-x, x+1), \quad Q(n,p,x) = I_p(x+1, n-x).$$

In order to avoid a loss of significant digits by cancellation, it is always convenient to compute the smallest of the two functions ($P(n,p,x)$ or $Q(n,p,x)$). For this, one can use the transition point for the function $I_x(p,q)$, which is given by $x_t \approx p/(p+q)$. In the case of the binomial distribution, we will have $p_t \approx (x+1)/(n+1)$. Then, if $p > p_t$ ($p < p_t$) it is preferable to evaluate $P(n,p,x)$ ($Q(n,p,x)$).

For the negative binomial, we have

$$(1.4) \qquad P^{NB}(r,p,x) = I_p(r, x+1), \quad Q^{NB}(r,p,x) = I_{1-p}(x+1, r).$$

In this case, the transition point will be given by $p_t \approx r/(r+x+1)$. When $p < p_t$ ($p > p_t$) it is convenient to evaluate $P(n,p,x)$ ($Q(n,p,x)$).

In this paper, we explain that the methods used for the central beta distribution function (described in [2]) can be applied to obtain asymptotic representations of the binomial and negative binomial cumulative distribution functions and also for inverting these functions.

The inversion problem is, however, now slightly different: In [2] we considered the problem of finding y from the equation $I_y(a,b) = \alpha$. In the present case, the problem of inverting the binomial cumulative distribution function can be stated as follows: Given $\alpha \in (0,1]$, $p \in (0,1)$, and n (in the asymptotic problem a large positive integer), find the smallest positive integer x such that

$$(1.5) \qquad \alpha \le P(n,p,x) = \sum_{k=0}^{x} \binom{n}{k} p^k (1-p)^{n-k}.$$

When we assume $x \in [1,n]$, we cannot take α smaller than the sum of the first two terms of the sum at the right-hand side. However, the sum of these two terms becomes very small when n is large.

In the definitions of the finite sum in (1.1) and the subsequent equations, x should be an integer, but in the representations in (1.3) and (1.4), x may be real. In the inversion procedure we first assume that x is a real parameter and later round x to the smallest integer larger than x.

We give in detail the results for the binomial cumulative distribution function, and in the final section we will redefine some parameters to obtain the results for the negative binomial cumulative distribution function.

2. Results for the binomial distribution function. In Appendix A, we summarize earlier results for the incomplete beta function. We use them for the present case, where we need to change some notation. With the notation

$$\nu = n+1, \quad \xi = \frac{x+1}{\nu}, \quad 1-\xi = \frac{n-x}{\nu},$$

and from (1.4) and (A.8) (with $a = x+1$ and $b = n-x$), it follows that the representation of both binomial distributions $P(n,p,x)$ and $Q(n,p,x)$ in terms of the complementary error

function is

$$P(n, p, x) = I_{1-p}(n - x, x + 1) = \tfrac{1}{2}\mathrm{erfc}\left(+\eta\sqrt{\nu/2}\right) + R_\nu(\eta),$$

(2.1)

$$Q(n, p, x) = I_p(x + 1, n - x) = \tfrac{1}{2}\mathrm{erfc}\left(-\eta\sqrt{\nu/2}\right) - R_\nu(\eta),$$

where the function $R_\nu(\eta)$ has the asymptotic expansion given in (A.9). The expansion can be obtained by using a recursive scheme given in (A.10) in terms of a function $f(\eta)$ that arises when a change of the variable of integration is used; see (A.1), (A.2), with the final result in (A.4). In the present case we use

(2.2) $$f(\zeta) = \frac{\lambda\zeta}{t - \xi}, \quad f(\eta) = \frac{\lambda\eta}{p - \xi}, \quad \lambda = \sqrt{\xi(1 - \xi)},$$

where ζ is defined in (A.2) (t is a variable of integration in (A.1)), and the definition of η becomes

(2.3) $$-\tfrac{1}{2}\eta^2 = \xi \log\frac{p}{\xi} + (1 - \xi) \log\frac{1 - p}{1 - \xi}, \quad \mathrm{sign}(\eta) = \mathrm{sign}(p - \xi).$$

REMARK 2.1. The choice of the sign follows from the change of variables in Appendix A. We know that when $p \downarrow 0$, the binomial distributions approach the values $P(n, p, x) \to 1$, $Q(n, p, x) \to 0$. From equation (2.3) we see that the corresponding η in the complementary error function tends to infinity when $p \downarrow 0$, and when we take $\eta \to -\infty$, we have $\tfrac{1}{2}\mathrm{erfc}(\eta\sqrt{\nu/2}) \to 1$, which is the wanted limit for $P(n, p, x)$. We see that this corresponds to the choice $\mathrm{sign}(\eta) = \mathrm{sign}(p - \xi)$. The result for $p \to 1$ follows similarly, in which case we need positive values of η.

Other representations that follow from (2.1) and (A.4) are

$$Q(n, p, x) = \frac{F_\nu(\eta)}{F_\nu(\infty)}, \quad F_\nu(\eta) = \sqrt{\frac{\nu}{2\pi}} \int_{-\infty}^{\eta} e^{-\frac{1}{2}\nu\zeta^2} f(\zeta)\, d\zeta,$$

(2.4)

$$P(n, p, x) = \frac{G_\nu(\eta)}{F_\nu(\infty)}, \quad G_\nu(\eta) = \sqrt{\frac{\nu}{2\pi}} \int_{\eta}^{\infty} e^{-\frac{1}{2}\nu\zeta^2} f(\zeta)\, d\zeta,$$

where $f(\zeta)$ is given in (2.2).

Here and in the representation of the incomplete beta function in (A.4), a function $F_\nu(\infty)$ occurs, which is defined in (A.5). It has the large-ν asymptotic expansion stated in (A.5). The first coefficients are given in (A.6).

2.1. Some expansions. An expansion of η in (2.3) in terms of powers of $q = (p - \xi)/\lambda^2$ with $\lambda = \sqrt{\xi(1 - \xi)}$ reads

$$\eta = q\lambda \left(1 - \tfrac{1}{3}(1 - 2\xi)q + \tfrac{1}{36}\left(7 - 19\xi + 19\xi^2\right)q^2 + \mathcal{O}\left(q^3\right)\right).$$

The limiting values (for fixed $\xi \in (0, 1)$) are

(2.5) $$\lim_{p\downarrow 0} \eta = -\infty, \quad \lim_{p\uparrow 1} \eta = +\infty.$$

We can also consider η as a function of ξ. The limiting values (for fixed $p \in (0, 1)$) are

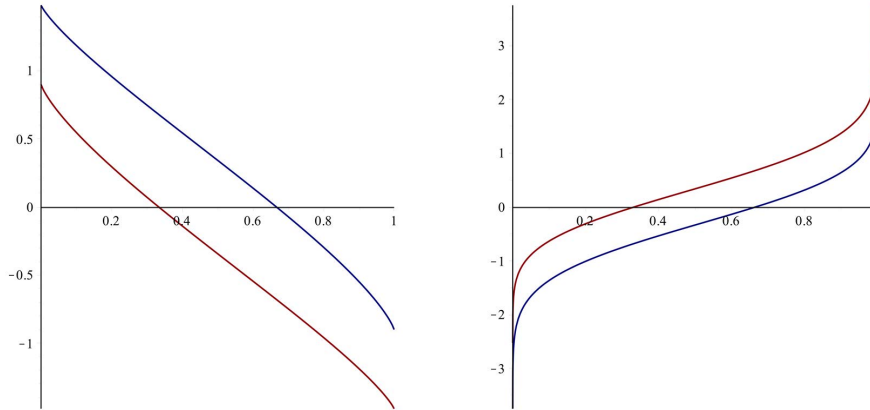(2.6) $$\lim_{\xi\downarrow 0} \eta = \sqrt{-2\log(1 - p)}, \quad \lim_{\xi\uparrow 1} \eta = -\sqrt{-2\log p}.$$

FIG. 2.1. *Left: the function η defined in (2.3) as a function of $\xi \in (0, 1)$ for two values of p: $p = 1/3$ (lower curve) and $p = 2/3$ (upper curve). The function η has a zero at $\xi = p$. Right: the function η defined in (2.3) as a function of $p \in (0, 1)$ for two values of ξ: $\xi = 1/3$ (upper curve) and $\xi = 2/3$ (lower curve). The function η has a zero at $p = \xi$.*

At the left-hand side of Figure 2.1 we display two curves of η as a function of ξ for two values of p: $p = 1/3$ (upper curve) and $p = 2/3$ (lower curve). The function η has a zero at $\xi = p$. At $\xi = 0$ and $\xi = 1$, the values of η follow from (2.6). At the right-hand side of Figure 2.1 we give a similar illustration of η as a function of p for two values of ξ: $\xi = 1/3$ (lower curve) and $\xi = 2/3$ (upper curve). The function η has a zero at $p = \xi$. At $p = 0$ and $p = 1$, we have $\eta \to \pm\infty$; see (2.5).

For the inversion procedure it is convenient to state the expansion of ξ in terms of powers of η:

$$(2.7) \qquad \xi = p - p(1-p) \sum_{k=1}^{\infty} a_k \widetilde{\eta}^k, \quad \widetilde{\eta} = \frac{\eta}{\sqrt{p(1-p)}}.$$

The first coefficients are

$$a_1 = 1, \qquad a_2 = \tfrac{1}{6}(2p - 1), \qquad a_3 = \tfrac{1}{72}(2p^2 - 2p - 1),$$
$$a_4 = -\tfrac{1}{540}(2p^3 - 3p^2 - 3p + 2), \qquad a_5 = \tfrac{1}{17280}(4p^4 - 8p^3 - 48p^2 + 52p - 23).$$

We also have

$$p = \xi + \lambda^2 \sum_{k=1}^{\infty} b_k \widehat{\eta}^k, \quad \widehat{\eta} = \frac{\eta}{\lambda}, \quad \lambda = \sqrt{\xi(1-\xi)},$$

with the first coefficients

$$b_1 = 1, \qquad b_2 = \tfrac{1}{3}(1 - 2\xi), \qquad b_3 = \tfrac{1}{36}(13\xi^2 - 13\xi + 1),$$
$$b_4 = -\tfrac{1}{270}(2\xi - 1)(23\xi^2 - 23\xi - 1),$$
$$b_5 = \tfrac{1}{4320}(313\xi^4 - 626\xi^3 + 339\xi^2 - 26\xi + 1).$$

With these coefficients we can find the coefficients of the expansion

$$f(\eta) = \frac{\lambda\eta}{p - \xi} = \sum_{k=0}^{\infty} c_k \widehat{\eta}^k,$$

and the first coefficients are

$$c_0 = 1, \qquad c_1 = \tfrac{1}{3}(2\xi - 1), \qquad c_2 = \tfrac{1}{12}(\xi^2 - \xi + 1),$$
$$c_3 = -\tfrac{1}{135}(2\xi - 1)(\xi - 2)(\xi + 1), \quad c_4 = \tfrac{1}{864}(\xi^2 - \xi + 1)^2.$$

3. Inverting the binomial distribution function using the error function. We consider the inversion as described in (1.5), assuming that $\nu = n + 1$ is a large parameter. The inversion procedure is based on finding η from the equation (see (2.1))

$$\tag{3.1} \tfrac{1}{2}\mathrm{erfc}\left(\eta\sqrt{\nu/2}\right) + R_\nu(\eta) = \alpha, \quad \alpha \in (0, 1),$$

and with η fixed, we compute ξ, and then $x = \nu\xi - 1$ (rounded to an integer). We consider p and n as fixed given quantities.

The starting point for the inversion is to consider the error function in (3.1) as the main term in the representation. We compute η_0, the solution of the reduced equation

$$\tag{3.2} \tfrac{1}{2}\mathrm{erfc}\left(\eta_0\sqrt{\nu/2}\right) = \alpha.$$

A simple and efficient algorithm for computing the inverse of the complementary error function is included, for example, in the package described in [1]. Using this $\eta = \eta_0$ in (2.3), we compute ξ, either by using the series expansion in (2.7) or a numerical iteration procedure.

REMARK 3.1. When α or $1 - \alpha$ is very small, the value of $|\eta_0|$ may be very large, although a large value of ν may control this. Referring to the limits shown in (2.6) for a given p, we observe that if the value of η_0 satisfies $\eta_0 < -\sqrt{-2\log(1 - p)}$ or $\eta_0 > \sqrt{-2\log p}$, then a corresponding value of $\xi \in (0, 1)$ cannot be found.

Next we try to find a better approximation of η and assume that we have an expansion of the form

$$\tag{3.3} \eta \sim \eta_0 + \frac{\eta_1}{\nu}.$$

We can find the coefficient η_1 by using a perturbation method. We have from (3.2)

$$\tag{3.4} \frac{d\alpha}{d\eta_0} = -\sqrt{\frac{\nu}{2\pi}}\, e^{-\frac{1}{2}\nu\eta_0^2}.$$

To proceed, we consider $P(n, p, x) = I_{1-p}(n - x, x + 1) = \alpha$ and use the representation in (2.4). This yields

$$\tag{3.5} \frac{d\alpha}{d\eta} = -\frac{1}{F_\nu(\infty)}\sqrt{\frac{\nu}{2\pi}}\, e^{-\frac{1}{2}\nu\eta^2} f(\eta),$$

with $f(\eta)$ given in (2.2) and η given in (3.3). We obtain from (3.4) and (3.5) that

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

The coefficient η_1 in (3.3) depends on η_0, and we can substitute this approximation, compare equal powers of ν, and find η_1. It follows that

$$\tag{3.6} \eta_1 = \frac{1}{\eta_0}\log f(\eta_0).$$

This quantity is defined for $\eta_0 \to 0$ because of the expansion in (A.7).

For small values of η_0 (that is, when $\xi \sim p$, see (2.3)), we need an expansion of η_1 in terms of powers of η_0. We have

$$\eta_1 = -\frac{1 - 2\xi}{3\lambda} - \frac{5\xi^2 - 5\xi - 1}{36\lambda^2}\eta_0 + \frac{(2\xi - 1)(23\xi^2 - 23\xi - 1)}{1620\lambda^3}\eta_0^2$$
$$- \frac{31\xi^4 - 62\xi^3 + 33\xi^2 - 2\xi + 7}{6480\lambda^4}\eta_0^3 + \cdots, \quad \text{where } \lambda = \sqrt{\xi(1 - \xi)}.$$

REMARK 3.2. The asymptotic estimates in this section are uniformly valid for values $\xi \in [\delta, 1 - \delta]$, where δ is a small fixed positive number. This corresponds to the result of the expansion of the incomplete beta function; see (A.9).

3.1. The algorithmic steps of the inversion procedure. In the following steps, the algorithm for inverting the binomial distribution using the error function is summarized.
1. First obtain a value for η (η_0) from (3.2).
2. With this value η_0, obtain a first approximation ξ_0 of ξ by solving equation (2.3) either by a numerical iterative procedure or, when η_0 is small, by using the expansion in (2.7).
3. Evaluate η_1 by using (3.6), where $f(\eta_0) = \eta_0 \sqrt{\xi_0(1 - \xi_0)}/(p - \xi_0)$; see (2.2).
4. Next compute $\eta = \eta_0 + \eta_1/\nu$.
5. With this new value of η, obtain a further approximation of ξ by solving equation (2.3) either by a numerical iterative procedure or, when η is small, by using the expansion in (2.7).
6. Compute $x = \xi\nu - 1$, and round it to the nearest larger integer; this gives the final x.

4. Numerical examples. As a first example to find x from $\alpha \leq P(n, p, x)$, we take $n = 50$, $p = 0.4$, and $\alpha = 0.51$. With $\nu = 51$, we compute $\eta_0 \doteq -0.0035103$ by using (3.2). This gives $\xi \doteq 0.40172$ by (2.7) and $\eta_1 \doteq -0.13454$ by (3.6). Then $\eta \sim \eta_0 + \eta_1/\nu \doteq -0.0061484$. The new value of ξ follows from (2.7), $\xi \doteq 0.40301$. This gives $x \doteq 19.554$ and $I_{1-p}(n - x, x + 1) \doteq 0.510043$. Comparing this with $\alpha = 0.51$, the absolute error is 0.000043. The computations are done in Maple with the setting Digits $= 16$. The integer value of x is 20.

When we take the same values of α and p, and $n = 1500$, we find $x \doteq 599.94236$, with $P(n, p, x) \doteq 0.51000026659$, yielding an absolute error of 2.6×10^{-7}. Rounding x to the nearest integers we find $P(n, p, 599) \doteq 0.490189$ and $P(n, p, 600) \doteq 0.511212$.

A more extensive test of the performance of the expansion is provided in Figure 4.1. In the plots we show the relative errors when the approximation (3.3) has been considered in the inversion process for $p \in (0, 1)$ and with two different values of α ($\alpha = 0.35, 0.85$) and n ($n = 100, 1000$). As expected, a higher accuracy is obtained for the larger of the two n-values.

The efficiency of the computation also improves as n increases. This is not always the case in other existing algorithms for the inversion of the binomial distribution: for example, the CPU time in the computation of $0.96 \leq P(n, 0.5, x)$ for $n = 10000$ using the Matlab function **binoinv** is approximately 100 times higher than the same computation for $n = 100$. On the other hand, the algorithm implemented in R (the function **qbinom**) for the inversion of the binomial distribution seems to be much more efficient than the Matlab function (according to our tests, the difference in CPU times is only a factor 2 when computing with $n = 100$ and $n = 10000$), but, as before, there is no improvement in the efficiency of the computation as n increases.
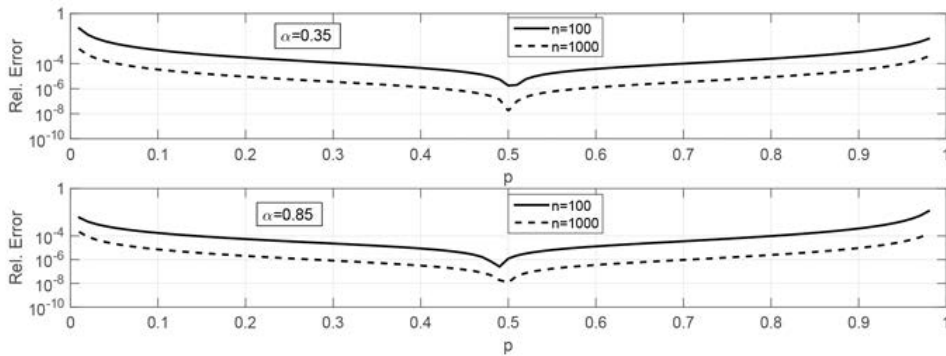
FIG. 4.1. *Inversion of the binomial distribution: performance of the expansion* (3.3) *for* $p \in (0, 1)$ *and two different values of* α *and* n.

5. Results for the negative binomial distribution function.

We recall the relations for the negative binomial distribution function

$$P^{NB}(r, p, x) = \sum_{k=0}^{x} \binom{k + r - 1}{r - 1} p^r (1 - p)^k = I_p(r, x + 1), \quad 0 \le p \le 1.$$

Comparing this with the representation of $P(n, p, x)$ in (1.3), we see that we can redefine the parameters: we change p into $1 - p$, and write

$$\nu = r + x + 1, \quad \xi = \frac{r}{\nu}, \quad 1 - \xi = \frac{x + 1}{\nu}.$$

The representation of the two negative binomial distributions in terms of the complementary error function is as in (2.1):

(5.1)
$$P^{NB}(r, p, x) = I_p(r, x + 1) = \tfrac{1}{2}\mathrm{erfc}\left(-\eta\sqrt{\nu/2}\right) - R_\nu(\eta),$$
$$Q^{NB}(r, p, x) = I_{1-p}(x + 1, r) = \tfrac{1}{2}\mathrm{erfc}\left(+\eta\sqrt{\nu/2}\right) + R_\nu(\eta),$$

where

(5.2)
$$-\tfrac{1}{2}\eta^2 = \xi \log \frac{p}{\xi} + (1 - \xi) \log \frac{1 - p}{1 - \xi}, \quad \mathrm{sign}(\eta) = \mathrm{sign}(p - \xi).$$

In the analysis of $P(n, p, x)$, the function $R_\nu(\eta)$ has not been used, and we refer to Appendix A to see its role in the asymptotic expansion of the incomplete beta function $I_x(a, b)$. The asymptotic expansion of $P^{NB}(r, p, x)$ for large ν follows from the expansion of the incomplete beta function $I_p(r, x + 1)$.

6. Inverting the negative binomial distribution function using the error function.

We consider the inversion problem in the form: with a given positive integer r, $p \in (0, 1)$, and $\alpha \in (0, 1)$, find the smallest integer x such that

$$\alpha \le P^{NB}(r, p, x).$$

In particular, we assume that r is large.

We use the representation in (5.1) and start with solving the equation

$$(6.1) \qquad \tfrac{1}{2}\mathrm{erfc}\left(-\eta\sqrt{\nu/2}\right) = \alpha.$$

Because the sought value of x is also part of ν, we have to modify the analysis for $P(n, p, x)$. We write the solution in the form

$$(6.2) \qquad -\eta\sqrt{\nu/2} = z, \quad z = \text{inverse erfc}(2\alpha), \quad \eta = -z\sqrt{2/\nu} = -z\sqrt{2\xi/r},$$

because $\nu = r/\xi$. To find the corresponding ξ from equation (5.2), we write this equation in the form

$$(6.3) \qquad \psi(\xi) = -\tfrac{1}{2}\rho^2, \quad \rho = -z\sqrt{2/r} = \eta/\sqrt{\xi},$$

where

$$(6.4) \qquad \psi(\xi) = -\frac{1}{2\xi}\eta^2 = \frac{1-\xi}{\xi}\log\frac{1-p}{1-\xi} + \log\frac{p}{\xi}, \quad \frac{d}{d\xi}\psi(\xi) = -\frac{1}{\xi^2}\log\frac{1-p}{1-\xi}.$$

The solution ξ of the equation $\psi(\xi) = -\tfrac{1}{2}\rho^2$ should satisfy $\mathrm{sign}(p-\xi) = \mathrm{sign}(\eta)$.

The limiting values of the function $\psi(\xi)$ are

$$\lim_{\xi\downarrow 0}\psi(\xi) = -\infty, \quad \lim_{\xi\uparrow 1}\psi(\xi) = \log p,$$

and for η we have

$$\lim_{\xi\downarrow 0}\eta = \sqrt{-2\log(1-p)}, \quad \lim_{\xi\uparrow 1}\eta = -\sqrt{-2\log p}.$$

So, when $\alpha < \tfrac{1}{2}$, that is, the solution should satisfy $p < \xi$, we can always find a solution of the equation $\psi(\xi) = -\tfrac{1}{2}\rho^2$ for $\xi \in (0, p)$. When $\tfrac{1}{2} < \alpha < 1$, there is a solution for $\xi \in (p, 1)$ when $\log p < -\tfrac{1}{2}\rho^2$. For large values of r this may be satisfied, but if not, then we cannot use the error function equation in (6.1) to find a value of ξ. For $p \to 1$, we have $P^{NB}(r, p, x) \to 1$, and the interval $(\log p, 0)$ becomes very small.

For small values of ρ, the solution of the equation in (6.3) can be expanded in the form

$$(6.5) \qquad \xi = p - p(1-p)\sum_{k=1}^{\infty} r_k\widetilde{\rho}^k, \quad \widetilde{\rho} = \frac{\rho}{\sqrt{1-p}},$$

and the first coefficients are

$$r_1 = 1, \quad r_2 = \tfrac{1}{6}(5p - 4), \quad r_3 = \tfrac{1}{72}\left(47p^2 - 74p + 26\right),$$
$$r_4 = \tfrac{1}{540}\left(268p^3 - 627p^2 + 453p - 92\right),$$
$$r_5 = \tfrac{1}{17280}\left(6409p^4 - 19868p^3 + 21792p^2 - 9608p + 1252\right).$$

We also have

$$p = \xi + \xi(1-\xi)\sum_{k=1}^{\infty} s_k\widehat{\rho}^k, \quad \widehat{\rho} = \frac{\rho}{\sqrt{1-\xi}},$$

and the first coefficients are

$$s_1 = 1, \quad s_2 = \tfrac{1}{3}(1 - 2\xi), \quad s_3 = \tfrac{1}{36}(13\xi^2 - 13\xi + 1),$$
$$s_4 = \tfrac{1}{270}(1 - 2\xi)(23\xi^2 - 23\xi + 1),$$
$$s_5 = \tfrac{1}{4320}(313\xi^4 - 626\xi^3 + 339\xi^2 - 26\xi + 1).$$
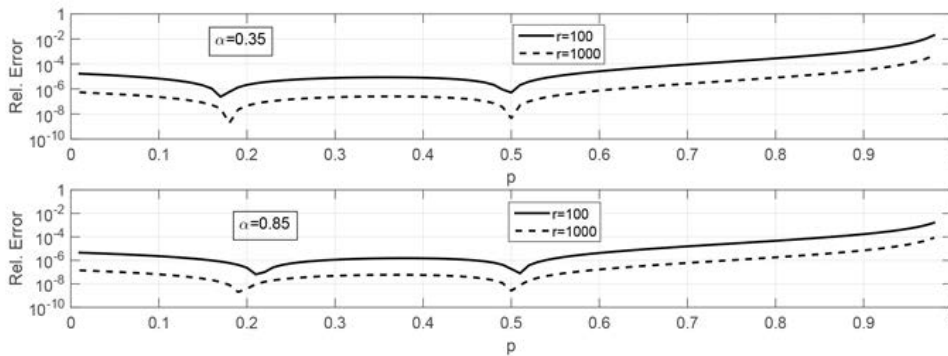
FIG. 6.1. *Inversion of the negative binomial distribution: performance of the expansion* (3.3) *for* $p \in (0, 1)$ *and two different values of* α *and* r.

The inversion method proceeds as in the case for $P(n, p, x)$ with minor modifications.

1. Compute z and ρ from (6.2) and (6.3).
2. Compute ξ from (6.4) by solving $\psi(\xi) = -\frac{1}{2}\rho^2$ by iteration or by using the expansion (6.5) when ξ is small. Call this first approximation ξ_0 and $x_0 = r/\xi_0 - r - 1$.
3. The corresponding η_0 follows from equation (6.3): $\eta_0 = \rho\sqrt{\xi_0}$.
4. Compute

$$(6.6) \qquad \eta_1 = \frac{1}{\eta_0} \log f(\eta_0), \quad f(\eta) = \frac{\eta\sqrt{\xi_0(1 - \xi_0)}}{p - \xi_0}.$$

5. Compute $\eta = \eta_0 + \eta_1/\nu$ with $\nu = r + x_0 + 1$.
6. The new value ξ follows from the expansion given in (6.5) when ξ is small (or by solving $\psi(\xi) = -\frac{1}{2}\rho^2$ by iteration), with $\rho = \eta/\sqrt{\xi_0}$.
7. Finally, $x = r/\xi - r - 1$, rounded to the integer just larger than this value.

As an example to find the smallest integer x from $\alpha \le P^{NB}(r, p, x)$, we take $r = 50$, $p = 0.4$, and $\alpha = 0.51$. The value z of (6.2) is $z \doteq -0.0177264$ and $\rho \doteq 0.00354528$. Using (6.5) we obtain $\xi_0 \doteq 0.398903$. Then (see (6.3)) $\eta_0 = \rho\sqrt{\xi_0} \doteq 0.00223916$ and (6.6) gives $\eta_1 \doteq -0.137068$, with $x_0 = r/\xi_0 - r - 1 \doteq 74.34369$ and $\nu \doteq 125.344$. The approximation of $\eta = \eta_0 + \eta_1/\nu$ becomes $\eta \doteq 0.001145617$ and $\rho = \eta/\sqrt{\xi_0} \doteq 0.00181387$. The corresponding ξ follows from the expansion in (6.5), which gives $\xi \doteq 0.399438$, and finally $x = r/\xi - r - 1 \doteq 74.1757$. When we compute $P^{NB}(r, p, x)$ with these values, we obtain $P^{NB}(r, p, x) \doteq 0.509992$. Comparing this with α, we observe an absolute error of 0.79×10^{-5}. The computations are done by Maple with the setting Digits = 16.

When we take the same values of α and p, and $r = 1500$, we find $x \doteq 2250.71$ with $P^{NB}(r, p, x) \doteq 0.50999995$, yielding an absolute error of 0.48×10^{-7}.

A more detailed example of the performance of the asymptotic inversion of the negative binomial distribution is provided in Figure 6.1. In the plots we display the relative errors (obtained by comparing to the values of the incomplete beta function $I_p(r, x + 1)$) when the approximation in (3.3) has been used in the inversion process. The results obtained for $p \in (0, 1)$ and two different values of α ($\alpha = 0.35, 0.85$) and r ($r = 100, 1000$) are shown for comparison. The expansion (6.5) has been considered in all cases to obtain the value ξ_0.

Appendix A. Summary of the asymptotic results for the incomplete beta function.
We collect results from [2, 7], [8, Section 38.4], with a slightly different notation. We write

$$\nu = a + b, \quad \xi = \frac{a}{\nu}, \quad b = \nu(1 - \xi).$$

Then (1.2) can be written as

$$(A.1) \qquad I_x(a, b) = \frac{1}{B(a, b)} \int_0^x e^{\nu(\xi \log t + (1-\xi) \log(1-t))} \frac{dt}{t(1 - t)}.$$

We consider ν a large parameter and ξ bounded away from 0 and 1. The maximum of the exponential function occurs at $t = \xi$. We use the transformation

$$(A.2) \qquad -\tfrac{1}{2}\zeta^2 = \xi \log \frac{t}{\xi} + (1 - \xi) \log \frac{1 - t}{1 - \xi},$$

where the sign of ζ equals the sign of $t - \xi$. The same transformation holds for $x \mapsto \eta$ if t and ζ are replaced by x and η, respectively. That is,

$$(A.3) \qquad -\tfrac{1}{2}\eta^2 = \xi \log \frac{x}{\xi} + (1 - \xi) \log \frac{1 - x}{1 - \xi}.$$

When taking the square root of η we assume that $\mathrm{sign}(\eta) = \mathrm{sign}(x - \xi)$, this means that $\mathrm{sign}(\eta) = \mathrm{sign}\,(x - a/(a + b))$.

Using (A.2) we obtain

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

and we can write (A.1) in the form

$$(A.4) \qquad I_x(a, b) = \frac{F_\nu(\eta)}{F_\nu(\infty)}, \quad F_\nu(\eta) = \sqrt{\frac{\nu}{2\pi}} \int_{-\infty}^\eta e^{-\frac{1}{2}\nu\zeta^2} f(\zeta) \, d\zeta,$$

where

$$(A.5) \qquad f(\zeta) = \frac{\zeta\lambda}{t - \xi}, \quad F_\nu(\infty) = \frac{\Gamma^*(a)\Gamma^*(b)}{\Gamma^*(a + b)} \sim \sum_{k=0}^\infty \frac{F_k}{\nu^k}, \quad \lambda = \sqrt{\xi(1 - \xi)}.$$

The function $\Gamma^*(x)$, the slowly varying part of the Euler gamma function, is defined by

$$\Gamma^*(x) = \frac{\Gamma(x)}{\sqrt{2\pi/x}\, x^x e^{-x}}, \quad x > 0.$$

The first coefficients F_k are

$$(A.6) \qquad \begin{aligned} &F_0 = 1, \quad F_1 = \frac{1 - \xi + \xi^2}{12\lambda^2}, \quad F_2 = \frac{(1 - \xi + \xi^2)^2}{288\lambda^4}, \\ &F_3 = -\frac{139\xi^6 - 417\xi^5 + 402\xi^4 - 109\xi^3 + 402\xi^2 - 417\xi + 139}{51840\lambda^6}. \end{aligned}$$

The first coefficients of the Taylor expansion

$$(A.7) \qquad f(\zeta) = a_0 + a_1\zeta + a_2\zeta^2 + a_3\zeta^3 + \cdots$$

are

$$a_0 = 1, \quad a_1 = \frac{2\xi - 1}{3\lambda}, \quad a_2 = \frac{1 - \xi + \xi^2}{12\lambda^2}.$$

When we replace in (A.4) the function $f(\zeta)$ by 1, the integral becomes the complementary error function defined by

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

As explained in [6], we can write

$$(A.8) \qquad I_x(a,b) = \tfrac{1}{2}\operatorname{erfc}\left(-\eta\sqrt{\nu/2}\right) - R_\nu(\eta), \quad \nu = a + b,$$

where the relation between x and η follows from (A.3), and $R_\nu(\eta)$ has the expansion

$$(A.9) \qquad R_\nu(\eta) \sim \frac{1}{F_\nu(\infty)} \frac{e^{-\frac{1}{2}\nu\eta^2}}{\sqrt{2\pi\nu}} \sum_{k=0}^{\infty} \frac{C_k(\eta)}{\nu^k}, \quad \nu \to \infty,$$

and $F_\nu(\infty)$ is defined in (A.5). This expansion is uniformly valid for $\xi = a/(a+b) \in [\delta, 1-\delta]$, where δ is a small fixed positive number.

The coefficients $C_k(\eta)$ can be obtained from the scheme

$$(A.10) \qquad C_k(\eta) = \frac{f_k(\eta) - f_k(0)}{\eta}, \quad f_k(\zeta) = \frac{d}{d\zeta} \frac{f_{k-1}(\zeta) - f_{k-1}(0)}{\zeta},$$

$k = 0, 1, 2, \ldots$, with $f_0 = f$ defined in (A.5).

REFERENCES

[1] A. GIL, J. SEGURA, AND N. M. TEMME, *GammaCHI: A package for the inversion and computation of the gamma and chi-square cumulative distribution functions (central and noncentral)*, Comput. Phys. Commun., 191 (2015), pp. 132–139.

[2] ———, *Efficient algorithms for the inversion of the cumulative central beta distribution*, Numer. Algorithms, 74 (2017), pp. 77–91.

[3] X. LI, D. WU, N. G. F. COOPER, AND S. N. RAI, *Sample size calculations for the differential expression analysis of RNA-seq data using a negative binomial regression model*, Stat. Appl. Genet. Mol. Biol., 18 (2019), Art. 20180021, 17 pages.

[4] T. MARQUEZ-LAGO AND K. BURRAGE, *Binomial tau-leap spatial stochastic simulation algorithm for applications in chemical kinetics*, J. Chem. Phys., 127 (2007), Art. 104101, 9 pages.

[5] D. MCCARTHY, Y. CHEN, AND G. SMYTH, *Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation*, Nucleic Acids Res., 40 (2012), pp. 4288–4297.

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

[7] ———, *Asymptotic inversion of the incomplete beta function*, J. Comput. Appl. Math., 41 (1992), pp. 145–157.

[8] N. M. TEMME, *Asymptotic Methods for Integrals*, World Scientific, Hackensack, 2015.