Incomplete Gamma Functions for Evaluating Erlang Process Probabilities

By John R. B. Whittlesey

1. Introduction. A number of important statistical distributions can be expressed in terms of the normalized incomplete gamma function [1], [2]

$$(1) \qquad G_a(x) = \Gamma(a, x)/\Gamma(a) = \frac{1}{\Gamma(a)} \int_x^\infty e^{-t} t^{a-1} \, dt = 1 - \gamma(a, x)/\Gamma(a)$$

which is also sometimes referred to as the partial or incomplete Poisson sum:

$$(2) \qquad G_a(x) = E_{a-1} = \sum_{n=0}^{a-1} e^{-x} x^n / n! = \sum_{n=0}^{a-1} H_n(x)$$

when a is an integer.

Single and double precision FORTRAN subroutines have now been written to evaluate $G_a(x)$ and $H_a(x)$ for all real values of the parameters equal to, or greater than zero [3], [4].

The purpose of this paper is to give brief details of these subroutines and to show how they can be used for evaluating the probabilities from two distributions of importance in Operations Research: namely, the generalized Poisson [5], [6], and the Morse-Jewell distribution [7], [8], which are respectively the *synchronous* and *asynchronous* counting distributions for an Erlang process, [9].

2. Calculating the Poisson Term $H_a(x)$. In the double-precision subroutine, $H_a(x)$ is calculated logarithmically by the formula

$$(3) \qquad \begin{aligned} \log H_a(x) &= -x + a \log x - \log \Gamma(a + 1) \\ &= (a - x) + a \log(x/a) - \tfrac{1}{2} \log(2\pi a) - \text{Series } B \end{aligned}$$

for all values of $a \geqq 8$. Here

$$(4) \qquad \text{Series } B = \frac{B_1}{1 \cdot 2a} - \frac{B_2}{3 \cdot 4a^3} + \cdots,$$

comes from the extended form of Stirling's approximation [10] in which B_1, B_2, $\cdots$ are the Bernoulli numbers $\frac{1}{6}$, $\frac{1}{30}$, up through $B_7 = \frac{7}{6}$. For $a < 8$, $H_a(x)$ is calculated from

$$(5) \qquad H_a(x) = H_w(x) w(w - 1)(w - 2) \cdots (a + 1) x^{a-w}$$

where $w = (8 + \theta)$ with $\theta = a - [a] = $ the fractional part of a, when a is not an integer.

In the single-precision subroutine, the generalization to non-integral values of a, $(a \leqq 40)$ is accomplished by means of the expression

$$(6) \qquad H_a(x) = \frac{e^{-x} x^\theta}{\Gamma(1 + \theta)} \prod_{0 \leqq n < [a]} x/(a - n)$$

where $T(1 + \theta)$ is evaluated by using one of Hastings' polynomial approximations [11].

3. Evaluating the Incomplete Gamma Function. Below are listed four methods for evaluating $G_a(x)$. Three of them make use of the Poisson term, $H_a(x)$. The fourth evaluates $G_a(x)$ for very large values of the parameters.

A. *Convergent series* [12]:

$$G_a(x) = 1 - H_a(x) \sum_{n=0}^{\infty} T_n$$

where

$$T_0 = 1, \quad \text{and} \quad T_n = \frac{x}{a + n} T_{n-1} \quad \text{for} \quad n > 0.$$

B. *Asymptotic series* [12], [13]:

$$G_a(x) = \frac{a}{x} H_a(x) \sum_{n=0} S_n$$

where

$$S_0 = 1 \quad \text{and} \quad S_n = \frac{a - n}{x} S_{n-1}$$

so long as $| S_n/S_{n-1} | < 1$ except when a is an integer.

C. *Method of continued fractions* [14]:

$$G_a(x) = H_a(x) \left(\cfrac{a_1}{b_1 + \cfrac{a_2}{b_2 + \cfrac{a_3}{b_3 + \cdots}}} \right)$$

where

$$a_1 = a(a - 1)/x^2$$
$$a_n = (n - 1)(a - n)/x^2 \quad \text{for} \quad n > 1$$

and

$$b_n = (x + 2n - a)/x \quad \text{for} \quad n > 0.$$

D. *Gaussian limit:* For very large values of the parameters, a Gaussian approximation to the normalized incomplete gamma function is achieved by means of a series of successive approximations to z, the Gaussian variable of integration. First let $z_0 = (x/a)^{1/3}$ after Wilson and Hilferty [15]. The moments for this variable are given by Abdel-Aty [16]. When written in terms of the transformed parameter $b = 9a$, they become:

(7)
$$\mu_1 = 1 - \frac{1}{b} + \frac{10}{3b^3} + \frac{11}{3b^4} + \cdots$$

$$(8) \qquad \mu_2 = \frac{1}{b} - \frac{13}{3b^3} - \frac{10}{b^4} + \cdots$$

$$(9) \qquad \mu_3 = \frac{4}{b^3} + \frac{16}{b^4} + \cdots$$

$$(10) \qquad \mu_4 = -\frac{2}{b^3} - \frac{16}{b^4} + \cdots$$

The second approximation is

$$(11) \qquad z_1 = (z_0 - \mu_1)/\sqrt{\mu_2}$$

taking terms through $O(b^{-3})$.

Finally

$$(12) \qquad z_2 = z_1 - \frac{1}{3b}\left\{\frac{2(z_1^2 - 1)}{b^{1/2}} - \frac{(z_1^3 - 3z_1)}{4}\right\}$$

is obtained by the method of Severo and Zelen [17] which uses μ_3 and μ_4 through $O(b^{-3})$ in the Gram-Charlier and Taylor series expansions to obtain a small further correction to z_1. Figure 1 shows the error curve for this approximation when $a = 30000$.

An estimate of the maximum error from equation (12) can be gained by inserting the true value for the median

$$(13) \qquad x_m = a - \tfrac{1}{3} + \tfrac{8}{135}b^{-1} + O(b^{-2})$$

(from Pearson [18]) into the expression for z_2. This gives

$$(14) \qquad z_2(x_m) = \tfrac{1}{5}b^{-3/2} + O(b^{-5/2}).$$

The maximum error in $G_a(x)$ is then approximated by

$$(15) \qquad z_2(x_m) \cdot \text{Gaussian ordinate (at } z = 0)$$

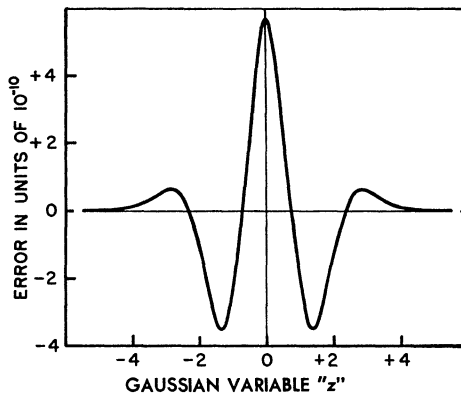which for $a = 30000$ equals $.57 \cdot 10^{-9}$, in good agreement with the maximum error from Figure 1.

FIG. 1.—Error curve for Gaussian approximation when $a = 30000$.

4. Choice Amongst Methods. Figure 2 shows the regions selected for methods **A**, **B**, **C**, and **D**. Below are given the values for the constants A_1, A_2, $\cdots$ A_7 and the reasons for their choice.

A_1) Methods **B** and **C** are both capable of giving correct results in the region $x > a > A_4$, $(z > A_7)$. A_1 depends on their relative efficiency in this region and was set equal to 200 in both the single and double-precision subroutines.

A_2) is determined by the point at which the terminating series for integer values of a, method **B**, requires, on average, more terms than the non-terminating convergent series **A** for values of $x < a$. In the single-precision subroutine (working to 6 significant figures) it was set at $A_2 = 20$; in the double-precision subroutine (working to 10 significant figures) at $A_2 = 75$.

A_3) When x is much smaller than a, there exists the possibility that method **B** will arrive at an overflow condition before it terminates at the ath term. This will happen on the IBM-7090 when $S_{a-1} = (a - 1)!/x^{a-1}$ is allowed to exceed $10^{38.5}$. Overflow in this situation is adequately guarded against by substituting method **A** for method **B** whenever x is less than $a(a - 1)/A_3$, $A_3 = 500$, $a \leqq A_2$.

A_4) When a is a small non-integer, the asymptotic series of method **B** may begin to diverge prematurely. To prevent this, A_4 is set at 13 in the single-precision subroutine, and in the double-precision at $A_4 = 23$.

A_5) and A_6) determine the region for the Gaussian approximation. The number of terms required for the most applicable of methods **A**, **B**, and **C** decreases with increasing $|z|$, but increases with a. To obtain 6-place accuracy A_6 should exceed 300, and for 9-place accuracy it should be greater than about 30000. A_5 was set equal to 5.5 to prevent floating-point overflow in the convergent (double-precision) expansion for $\mathrm{Erf}(z)$.

A_7) For small positive values of z, method **C** begins to become inaccurate as

REGIONS OF x

REGIONS OF a	$0 < x < 1$	$1 \leqq x \leqq A_1$		$x > A_1$	
		$x < a$	$x \geqq a$	$x < a$	$x \geqq a$
INTEGER a		A, B^*	B	$G_a(x) = 0$	
— $0 < a \leqq A_2$ — NON-INTEGER a		A	B, C^*		
$a > A_2$	$G_a(x) = 1$	A	B	A, D^{**}	B, C^* D^{**}

Fig. 2.—Choice of method as a function of regions of x and a.

* Method **B** used rather than method **A** when $x > a(a - 1)/A_3$

Method **C** used rather than method **B** when $(a + x) < A_4$ and when $x > A_1$, $z > A_7(a)$, or $z > A_5$

** Method **D** used when $a > A_6$ and $|z| < A_5$.

a grows large. To exclude this region in the double-precision routine, A_7 was set equal to the smallest among the following expressions in a:

$$0.7 + .037 \cdot a^{1/2}$$
$$3.0 + 7.0 \cdot 10^{-5} \cdot a$$
$$3.5 + 3.3 \cdot 10^{-5} \cdot a$$

as a function of a. Method **B** was then used in place of method **C** whenever $z < A_7$, $x > a$, $a \leqq A_6$.

5. Probabilities for an Erlang Process. The probabilities of exactly n events from a generalized Poisson distribution [6], [19] are given by

$$(16) \qquad P_n(k, x) = G_{nk+k} - G_{nk}$$

and for the Morse-Jewell distribution by [9]

$$(17) \qquad Q_n(k, x) = \Delta_n^2 \left(n(G_{nk} + H_{nk}) - \frac{x}{k} G_{nk} \right) = -\Delta_n^2 \frac{1}{k} \int^x G_{nk}(t) \, dt$$

where Δ_n^2 is the second-order central difference operator with respect to the index n, and $G_a = G_a(x)$ and $H_a = G_{a+1} - G_a$.

A FORTRAN subroutine, $QPROB$, for evaluating $Q = Q_n(k, X)$ is given in Figure 3. $D2$ produces the second-order differences as U goes from 1 to -1. For details of $GAMA$ see reference [3].

In the limit when k and x go to infinity as a fixed ratio

$$(18) \qquad Q_n = \begin{cases} 1 - \left| \dfrac{x}{k} - n \right|, & \text{for} \quad \left| \dfrac{x}{k} - n \right| < 1 \\ 0, & \text{elsewhere.} \end{cases}$$

For the generalized Poisson distribution, the limiting probabilities are approximated by:

$$(19) \quad P_n = \begin{cases} \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{z} e^{-t^2/2} \, dt, & \text{with} \quad a = nk + k, \quad \text{for} \quad n \leqq \left(\dfrac{x}{k} - .5 \right) \\ \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{z}^{\infty} e^{-t^2/2} \, dt, & \text{with} \quad a = nk, \quad \text{for} \quad n > \left(\dfrac{x}{k} - .5 \right) \end{cases}$$

ITERATE

```
ENTER        INITIALIZE      DO 100  J=1, 3        EXIT
QPROB                        A=(V+U) * W          RETURN
             V = n           CALL GAMA (A,X,G,H,ER)
             W = k           GU=(V+U)*(G+H)-X*G/W
             U = I.          D2=3.*U**2-2.
             Q = 0.          U=U-I.
                         100 Q=Q+D2*GU
```
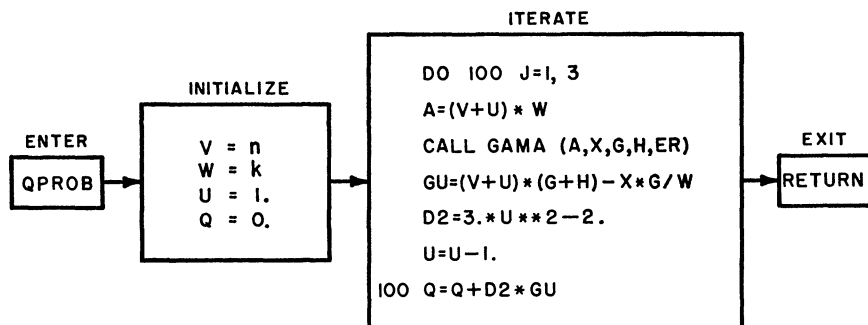
FIG. 3.—FORTRAN subroutine.

TABLE 1

x/k		k = 5	k = 10	k = 25	k = 50	k = 100	k = 200	k = 500	k = 1000
.25	Q	3.1	4.8	—	—	—	—	—	—
	P	2.0	3.6	7.8	—	—	—	—	—
.50	Q	1.6	2.4	4.1	6.6	—	—	—	—
	P	2.8	4.4	4.4	6.2	—	—	—	—
.75	Q	*	*	1.9	2.6	3.8	5.8	—	—
	P	2.4	2.9	3.7	5.0	4.9	6.2	—	—
.95	Q	*	*	*	*	1.4	1.7	2.3	2.8
	P	1.7	2.9	3.4	3.8	4.1	4.4	—	—
1.05	Q	*	*	*	*	1.4	1.7	2.2	2.8
	P	1.4	2.3	3.8	4.8	4.8	4.8	—	—
1.25	Q	*	*	1.6	2.2	3.1	4.7	—	—
	P	*	1.5	2.9	5.1	4.9	5.7	—	—
1.50	Q	*	1.5	2.4	3.5	5.5	—	—	—
	P	*	*	1.9	3.0	5.3	—	—	—
1.75	Q	*	*	1.3	1.8	2.5	3.6	6.3	—
	P	*	1.7	3.0	4.3	6.0	5.5	6.8	—
1.95	Q	*	*	*	*	*	1.4	1.8	2.2
	P	*	1.8	3.7	4.2	4.5	—	—	—
2.05	Q	*	*	*	*	*	1.4	1.8	2.2
	P	*	1.5	3.0	5.2	5.0	—	—	—
2.25	Q	*	*	*	1.6	2.3	3.3	5.6	—
	P	*	*	2.0	3.4	6.0	5.5	6.6	—
2.50	Q	*	1.3	1.9	2.8	4.1	6.5	—	—
	P	*	*	1.3	2.0	3.3	5.8	—	—
2.75	Q	*	*	*	1.4	2.0	2.8	4.8	—
	P	*	*	2.1	3.4	5.1	5.7	6.6	7.5
2.95	Q	*	*	*	*	*	1.3	1.6	2.0
	P	*	1.4	2.7	4.4	4.9	—	—	—

The first line of each entry gives minus the $\log_{10}$ of the Max. (over all n) of $|Q_n - \lim Q_n|$. For a given k, the differences $|Q_n - \lim Q_n|$ are least for x/k near an odd half-integer or zero. The second line gives $-\log_{10} \text{Max} \cdot_n |P_n - \lim P_n|$. Points where the differences exceed .05 are marked with asterisks (*), points where they are smaller than 10^{-8} are marked with dashes (—).

where z is obtained in terms of a and x from equation (12).

Actually, in the extreme limit, the synchronous Erlang counting probabilities become

$$(20) \qquad P_n' = \begin{cases} 1, & \text{for} \quad 0 \leqq \left(\dfrac{x}{k} - n\right) < 1 \\ 0, & \text{elsewhere} \end{cases}$$

since identical gaps of length k/x are then being counted.

Table 1 evaluates the goodness of fit of these limiting forms for various values of k and x/k.

UCLA Medical Center
Los Angeles 24, California

1. S. H. Khamis, "Incomplete gamma function expansions of statistical distribution functions," *Bull. Internat. Statist. Inst.*, v. 37, 1960, p. 385–396.

2. F. A. Haight, *Mathematical Theories of Road Traffic*, Academic Press, New York, 1963, (in preparation).

3. J. R. B. Whittlesey, "Normalized incomplete gamma function with Poisson term (UR GAMA)," *SHARE Dist.* ✻1177, IBM, New York, 1961.

4. J. R. B. Whittlesey, "Gamma(a, x)/Gamma(a) + Poisson term in double-precision (UR GAM2)," *SHARE Dist.* ✻1299, IBM, New York, 1962.

5. L. A. Goodman, "On the Poisson-gamma distribution problem," *Ann. Inst. Statist. Math. Tokyo.*, v. 11, 1959, p. 101–105.

6. F. A. Haight, "The generalized Poisson distribution," *Ann. Inst. Statist. Math. Tokyo.*, v. 3, 1952, p. 123–125.

7. P. M. Morse, *Queues, Inventories, and Maintenance*, John Wiley and Sons, New York, 1958.

8. W. S. Jewell, "The properties of recurrent-event processes," *Operations Res.*, v. 8, 1960, p. 446–472.

9. J. R. B. Whittlesey & F. A. Haight, "Counting distributions for an Erlang process," *Ann. Inst. Statist. Math. Tokyo.*, v. 13, 1961, p. 91–103.

10. T. J. I'a. Bromwich & T. M. Macrobert, *An Introduction to the Theory of Infinite Series*, Macmillan & Co., Ltd., London, 1955, p. 330.

11. C. Hastings, Jr., *Approximations for Digital Computers*, Princeton University Press, New Jersey, 1955, p. 158.

12. Bateman Manuscript Project, *Higher Transcendental Functions*, v. 2, McGraw-Hill Book Co., Inc., New York, 1953, p. 135.

13. A. Erdélyi, *Asymptotic Expansions*, Dover Publications, Inc., New York, 1956, p. 1–12.

14. H. S. Wall, *Analytic Theory of Continued Fractions*, D. Van Nostrand Co., Inc., 1948, p. 15, 21, 356.

15. E. B. Wilson & M. M. Hilferty, "The distribution of Chi-square," *Proc. Nat. Acad. Sci.*, v. 17, 1931, p. 684–688.

16. S. H. Abdel-Aty, "Approximate formulae for the percentage points and the probability integral of the non-central Chi-square distribution," *Biometrika*, v. 41, 1954, p. 538–540.

17. N. C. Severo & M. Zelen, "Normal approximation to the Chi-square and non-central F-probability functions," *Biometrika*, v. 47, 1960, p. 411–413.

18. K. Pearson, *Tables of the Incomplete Γ-Function*, Cambridge University Press, 1951, p. xvi.

19. F. A. Haight, B. F. Whisler, & W. W. Mosher, Jr., "New statistical method for describing highway distribution of cars," *Proc. Highway Res. Board*, v. 40, 1961, p. 557–564.