NASA TECHNICAL MEMORANDUM

NASA TM X-1592

NASA TM X-1592

GPO PRICE $ ________

CFSTI PRICE(S) $ ________

Hard copy (HC) 3.00

Microfiche (MF) [redacted]

ff 653 July 65

FACILITY FORM 602

(ACCESSION NUMBER)	(THRU)
13 (PAGES)	1 (CODE)
(NASA CR OR TMX OR AD NUMBER)	19 (CATEGORY)

A COMPUTER PROGRAM FOR REVERSION OF THE CUMULATIVE BINOMIAL PROBABILITY DISTRIBUTION

by Darl D. Bien

Lewis Research Center
Cleveland, Ohio

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • MAY 1968

NASA TM X-1592

A COMPUTER PROGRAM FOR REVERSION OF THE CUMULATIVE BINOMIAL PROBABILITY DISTRIBUTION

By Darl D. Bien

Lewis Research Center
Cleveland, Ohio

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151 – CFSTI price $3.00

ABSTRACT

A technique for reversion of the cumulative binomial probability distribution and an efficient and rapid computer program using the technique are presented. Reversion of the binomial series is useful in calculating the element success probability in a specified parallel system with redundancy which will yield the specified system success probability (reliability). The program is not limited to reliability problems but can also be used in a variety of scientific and engineering problems involving the reversion of the binomial series.

STAR Category 19

A COMPUTER PROGRAM FOR REVERSION OF THE CUMULATIVE BINOMIAL PROBABILITY DISTRIBUTION

by Darl D. Bien

Lewis Research Center

SUMMARY

A technique for reversion of the cumulative binomial probability distribution is presented. Reversion of the binomial series is useful in calculating element success probability in a parallel system with redundancy. That is, if a given parallel system and its attendant success probability (reliability) are specified, calculation of the required element success probability is of interest.

An efficient and rapid computer program was written in FORTRAN IV language and is presented along with instructions for its use. The program is not limited to reliability problems but can also be used in a variety of scientific and engineering problems involving the reversion of the binomial series.

INTRODUCTION

The binomial probability distribution has increasing application in present day scientific and engineering problems. The cumulative binomial distribution is

$$C(X \geq N_S | N, P) = \sum_{X=N_S}^{X=N} \frac{N!}{X!(N - X)!} P^X (1 - P)^{N-X}$$

It can be used to compute success probability C of a system of N parallel elements where the individual element success probability P is specified along with N_S, the number of elements required to succeed.

Frequently, the reversion of the cumulative binomial distribution is of interest. That is, if a given parallel system and its attendant success probability are specified, calculation of the required element success probability is desired. This type of problem arises, for example, in life testing procedures, sampling plans, and confidence interval calculations.

In general terms, the problem is to calculate the root P of the cumulative binomial series, given C, N, and N_S. Various methods have been used to evaluate P. Linear interpolation in tables of the cumulative binomial distribution (such as refs. 1 to 6) may be sufficient. Under certain conditions the binomial distribution can be approximated by other distributions which are also tabulated. Tables of the incomplete Beta function, the F distribution, and the normal distribution are examples. These methods result in limited accuracy and/or require lengthy computation.

A table of the reverse cumulative binomial distribution is presented in reference 7, but its range is limited. This report presents a computer program which rapidly calculates the root P of the general cumulative binomial series over a range of parameters exceeding the range of reference 7.

SYMBOLS

$C(X \geq N_S \mid N, P)$	probability of N_S or more successes, given the values of N and P
N	number of elements in parallel or number of trials
N_S	specified number of successes
P	element or trial success probability
X	variable number of successes

Subscripts:

j	iteration number
0	initial value

EQUATIONS

The binomial probability C of exactly N_S successes in N independent trials is given by

$$C(X = N_S \mid N, P) = \frac{N!}{N_S!\,(N - N_S)!} P^{N_S}(1 - P)^{N-N_S} \tag{1}$$

where P is the probability of success for an individual trial. The cumulative binomial probability of N_S or more successes in N independent trials is thus

$$C(X \geq N_S | N, P) = \sum_{X=N_S}^{X=N} \frac{N!}{X!\,(N - X)!} P^X (1 - P)^{N-X} \tag{2}$$

The computer program presented herein solves for P in equation (2) for specified values of C, N, and N_S.

Equation (2) specifies the upper tail of the cumulative binomial probability distribution. The lower tail can be computed from equation (2) by the following relation:

$$C(X < N_S | N, P) = 1 - C(X \geq N_S | N, P) \tag{3}$$

For certain values of N_S, such as $N_S < N/2$, the following relations result in a more efficient calculation than equation (2):

$$C(X < N_S | N, P) = \sum_{X=N-N_S+1}^{X=N} \frac{N!}{X!\,(N - X)!} (1 - P)^X P^{N-X} \tag{4}$$

and

$$C(X \geq N_S | N, P) = 1 - \sum_{X=N-N_S+1}^{X=N} \frac{N!}{X!\,(N - X)!} (1 - P)^X P^{N-X} \tag{5}$$

METHOD OF SOLUTION

The method consists of specifying C, N, and N_S in equation (2) and finding the value of P which satisfies the equation. An initial guess for P, designated P_0, is used in equation (2), and the numerical value of the summation is compared with the specified value of C. If the absolute value of the difference between the summation and the specified value of C is greater than 1×10^{-6}, a new estimate for P is used in equation (2). This process is continued until the value of the right side of equation (2) converges to the specified left-side value.

The Newton-Raphson method of finding the root P was chosen because it offers rapid convergence and because the required derivative of the cumulative binomial series is simple. The $(j+1)^{th}$ estimate for the root of an equation from the j^{th} estimate of that root with the Newton-Raphson iterative technique is stated in equation form as

$$P_{j+1} = P_j - \frac{F(P_j)}{\frac{\partial F(P_j)}{\partial P_j}} \tag{6}$$

(ref. 8). For the problem under consideration

$$F(P) = \sum_{X=N_S}^{X=N} \frac{N!}{X!\,(N - X)!} P^X (1 - P)^{N-X} - C(X \geq N_S | N, P) \tag{7}$$

and

$$\frac{\partial F(P)}{\partial P} = \frac{N!}{(N_S - 1)!\,(N - N_S)!} P^{N_S-1} (1 - P)^{N-N_S} \tag{8}$$

The Newton-Raphson technique requires a judicious initial guess for the root P. In this analysis, P_0 is chosen so that convergence to the root P is guaranteed. The binomial series is a monotonic function; therefore, convergence will always occur if P_0 is chosen at the point of maximum slope on the curve of C as a function of P (fig. 1). This point is determined by setting the second partial derivative of F(P) with respect to P equal to zero. When the second partial derivative

$$\frac{\partial^2 F(P)}{\partial P^2} = \frac{N!}{(N_S - 1)!\,(N - N_S)!} P^{N_S-2} (1 - P)^{N-N_S-1} \left[(N_S - 1) - P(N - 1)\right] \tag{9}$$

is set equal to zero, the result is

$$P_0 = \frac{N_S - 1}{N - 1} \tag{10}$$

The steps in the solution for the case $C(X \geq 6 | N = 10, P) = 0.90$ are illustrated in figure 1. In most cases the rapidity of the convergence by the Newton-Raphson technique to the desired accuracy (1×10^{-6}) is such that fewer than five iterations are required.

FORTRAN PROGRAM

The Newton-Raphson method of solution was programmed in FORTRAN IV language and required a storage capacity of less than 600 36-bit words on the Lewis IBM 7094 computer. The three inputs to the program are $C(X \geq N_S | N, P)$, N, and N_S; their FORTRAN names are, respectively, C, N, and NS (see Program Listing in the appendix). All three inputs are placed on a single card in the following format:

C: E10.8
N: I4
NS: I4

There is no limit to the number of cards which can be used to produce output for various sets of input.

The computer printout is in tabular form with five column headings: C, N, NS, P, and ERROR IN P. The first three columns are merely the input. The solution P is printed in column four; this is the j^{th} iteration on P, which satisfies the requirement that the series sum using P_j differ from the input C by less than 1×10^{-6}. ERROR IN P is simply the absolute value of the next increment in P, that is, $|P_{j+1} - P_j|$.

The program has been run over a wide range of C, N, and NS. For high values of N, intermediate steps of the calculation procedure may result in extremely small numerical values. If the computer used for the calculation does not signal an underflow/overflow condition, erroneous results could be produced for these extreme cases. For values of N as large as 500, however, no difficulties were encountered. Use of the program for values of input outside this range, therefore, should be made with an awareness of the underflow/overflow problem.

The program sums the cumulative binomial series from the mean term (largest term) outward in each direction until each additive term drops below 10^{-12}. This procedure saves the time required to compute those terms in the tail of the distribution which have a negligible effect on the cumulative sum and avoids the underflow problem encountered when computing a term like P^N where P may equal 0.5 and N may equal 500.

The FORTRAN IV listing and a sample set of output are given in the appendix. The running time for this sample, including input/output time, was 0.10 minute on the Lewis IBM 7094 computer.

Lewis Research Center,
National Aeronautics and Space Administration,
Cleveland, Ohio, April 17, 1968,
120-33-05-02-22.

APPENDIX - FORTRAN PROGRAM LISTING AND SAMPLE OUTPUT

Program Listing

```
                                                              JANUARY 24, 1968
C      REVBIN
C ----THIS PROGRAM USES THE NEWTON-RAPHSCN TECHNIQUE TO REVERSE THE BINOMIAL
C ----SERIES.  INPUTS ARE C,N,NS AND THE OUTPUT IS P IN THE FOLLOWING EQUATION.
C
C                           X=N    N     X       N-X
C                       C = SUM     C   P   (1-P)
C                           X=NS     X
C
C ----PRCGRAM CEVELOPEC BY DARL BIEN 1/24/68 AT NASA-LEWIS RESEARCH CENTER
C
      WRITE (6,31)
   31 FORMAT (1H1,49X,21HX=N   N      X      N-X /46X,22HC = SUM   C    P (1
     1-P) /50X,9HX=NS    X //,35X,1HC,14X,1HN,9X,2HNS,12X,1HP,12X,
     210HERROR IN P//)
      I=0
      M=0
      K=0
    1 READ (5,101) C,N,NS
  101 FORMAT (E10.8,2I4)
      AN = N
      ANS = NS
      IF (N-NS)1,50,51
   50 P=C**(1./AN)
      DELP = 0.
      GO TC 40
   51 IF (NS-1) 1,52,1C
   52 P=1.-(1.-C)**(1./AN)
      DELP = 0.
      GO TC 40
C ----INITIAL GUESS ON P SUCH THAT SECOND DERIVATIVE OF C WRT P EQUALS 0
   10 P = (ANS - 1.)/(AN - 1.)
   11 Q = 1. - P
      NLIMIT = N - NS
      ANLIMT = NLIMIT
      NL1=N-1
C
    2 NP =AN*P
      ANP = NP
      NLNP = N-NP
      ANLNP = NLNP
      NSLNP=NS-NP
      NPLNS=NP-NS
      TERM1 = 1.0
C ----CALCULATION OF MEAN (NP) TERM
      DO 6 K1 = 1,NLNP
      AK1= K1
    6 TERM1=TERM1*(ANP+AK1)/AK1*P**(ANP/ANLNP)*Q
C ----IF NP LESS THAN NS, START SERIES WITH NS TERM INSTEAD
      IF(NP-NS) 15,9,9
   15 DO 17 K11=1,NSLNP
```

```
      AK11=K11
   17 TERM1=TERM1*(ANLNP-AK11+1.)/(ANP+AK11)*P/Q
    9 CGUESS = TERM1
      TERM2 = TERM1
      DRCP = 1.0 E-12
C
      IF (NP-NS) 29,29,43
C ----NS+1 TERM, NS+2 TERM,....., N TERM
   29 DO 37 K21 = NS,NL1
      AK21 = K21
      TERM2=TERM2*(AN-AK21)/(AK21+1.0)*P/Q
      IF (TERM2-DROP) 14,14,37
   37 CGUESS =CGUESS + TERM2
   14 GO TC 16
C ----TERMS GREATER THAN NP
   43 DO 7 K2 =NP,NL1
      AK2=K2
      TERM2=TERM2*(AN-AK2)/(AK2+1.0)*P/Q
      IF (TERM2-DROP) 22,22,7
    7 CGUESS = CGUESS + TERM2
C ----TERMS LESS THAN NP (INCLUDING NS TERM)
   22 TERM3=TERM1
      DO 8 K3=1,NPLNS
      AK3=K3
      TERM3=TERM3*(ANP-AK3+1.)/(ANLNP+AK3)*Q/P
      IF (TERM3-DROP) 16,16,8
    8 CGUESS = CGUESS + TERM3
C ----CALCULATION OF DERIVATIVE OF BINOMIAL SERIES
   16 DERIV = ANS
      DO 4 J=1,NLIMIT
      AJ=J
    4 DERIV=DERIV*(ANS+AJ)/AJ*P**((ANS-1.)/ANLIMT)*Q
C
C ----THE FORM OF NEWTCN-RAPHSON SOLUTION IS P2=P1-Y(P1)/Y'(P1)
C ----DELP  IS THE CORRECTION TO BE MADE ON P IF ACCURACY OF C IS NOT SATISFIED
      DELP = CGUESS/DERIV - C/DERIV
      ACCRYC = ABS(CGUESS/C-1.)
C ----RELATIVE ERROR OF THE FORM (CGUESS-C)/C IS COMPUTED ABOVE AS ACCRYC.  IF
C ----LESS THAN 1.0 E-6, PRINT RESULT--IF NOT GET NEW GUESS FOR P AND REPEAT
      IF (ACCRYC-.000001)40,5,5
    5 P = P - DELP
      Q = 1. - P
      GO TC 2
   40 ERROR=ABS(DELP)
      I = I+1
      M = M+1
      IF (I-5) 205,205,201
  201 WRITE (6,21)
   21 FORMAT (/)
      M = M+1
      I = 1
  205 IF (K) 210,210,203
  210 IF (M-53)200,200,110
  203 IF (M-57)200,200,110
  110 WRITE (6,41)
   41 FORMAT(1H1,34X,1HC,14X,1HN, 9X,2HNS,12X,1HP,12X,10HERROR IN P//)
      I = 1
      M = 0
      K = K+1
  200 WRITE (6,20) C,N,NS,P,ERROR
   20 FORMAT (     30X,F10.8,7X,I4,7X,I4,8X, F8.6,8X,E10.2/)
C----READ ANOTHER SET OF DATA
      GO TC 1
      END
```

Sample Output

```
          X=N   N       X       N-X
      C = SUM    C    P (1-P)
          X=NS    X
```

C	N	NS	P	ERROR IN P
0.90000000	10	6	0.732682	0.30E-07
0.90000000	10	7	0.812438	0.75E-08
0.90000000	10	8	0.884175	0.37E-08
0.90000000	10	9	0.945471	0.11E-07
0.90000000	10	10	0.989519	0.
0.90000000	50	30	0.675475	0.26E-07
0.90000000	50	35	0.767509	0.19E-07
0.90000000	50	40	0.855019	0.56E-08
0.90000000	50	45	0.935740	0.
0.90000000	50	50	0.997895	0.
0.90000000	100	60	0.656162	0.19E-08
0.90000000	100	70	0.750898	0.19E-07
0.90000000	100	80	0.842332	0.11E-07
0.90000000	100	90	0.928702	0.10E-07
0.90000000	100	100	0.998947	0.
0.90000000	500	300	0.626771	0.14E-07
0.90000000	500	350	0.724691	0.67E-07
0.90000000	500	400	0.821096	0.85E-07
0.90000000	500	450	0.915134	0.18E-07
0.90000000	500	500	0.999789	0.
0.95000000	10	6	0.777559	0.60E-07
0.95000000	10	7	0.849971	0.55E-06
0.95000000	10	8	0.912736	0.37E-07
0.95000000	10	9	0.963229	0.
0.95000000	10	10	0.994884	0.
0.95000000	50	30	0.698616	0.67E-07
0.95000000	50	35	0.787896	0.60E-07
0.95000000	50	40	0.871443	0.37E-08
0.95000000	50	45	0.946429	0.10E-06
0.95000000	50	50	0.998975	0.
0.95000000	100	60	0.673014	0.16E-06
0.95000000	100	70	0.765991	0.60E-07
0.95000000	100	80	0.854755	0.26E-07
0.95000000	100	90	0.937075	0.19E-08
0.95000000	100	100	0.999487	0.
0.95000000	500	300	0.634565	0.56E-08
0.95000000	500	350	0.731838	0.17E-06
0.95000000	500	400	0.827167	0.63E-07
0.95000000	500	450	0.919465	0.18E-07
0.95000000	500	500	0.999897	0.

REFERENCES

1. Anon.: Tables of the Cumulative Binomial Probability Distribution. Harvard Univ. Press, 1955.
2. Anon.: Tables of the Binomial Probability Distribution. Appl. Math. Series No. 6, National Bureau of Standards, Jan. 27, 1950.
3. Anon.: Tables of the Cumulative Binomial Probabilities. Ordnance Corps, Washington, D.C., 1952.
4. Romig, Harry G.: 50-100 Binomial Tables. John Wiley & Sons, Inc., 1953.
5. Robertson, William H.: Tables of the Binomial Distribution Function for Small Values of p. Rep. SCR-143, Sandia Corp., Jan. 1960.
6. Weintraub, Sol: Tables of the Cumulative Binomial Probability Distribution for Small Values of p. Free Press of Glencoe, New York, 1963.
7. Leone, Fred C.; Hayman, George E.; Chu, John T.; and Topp, Chester W.: Percentiles of the Binomial Distribution, Rep. 1030, Case Inst. Tech., Stat. Lab. (AFOSR TN 60-620), June 1960.
8. McCracken, Daniel D.; and Dorn, William S.: Numerical Methods and FORTRAN Programming. John Wiley & Sons, Inc., 1964, pp. 133-135.

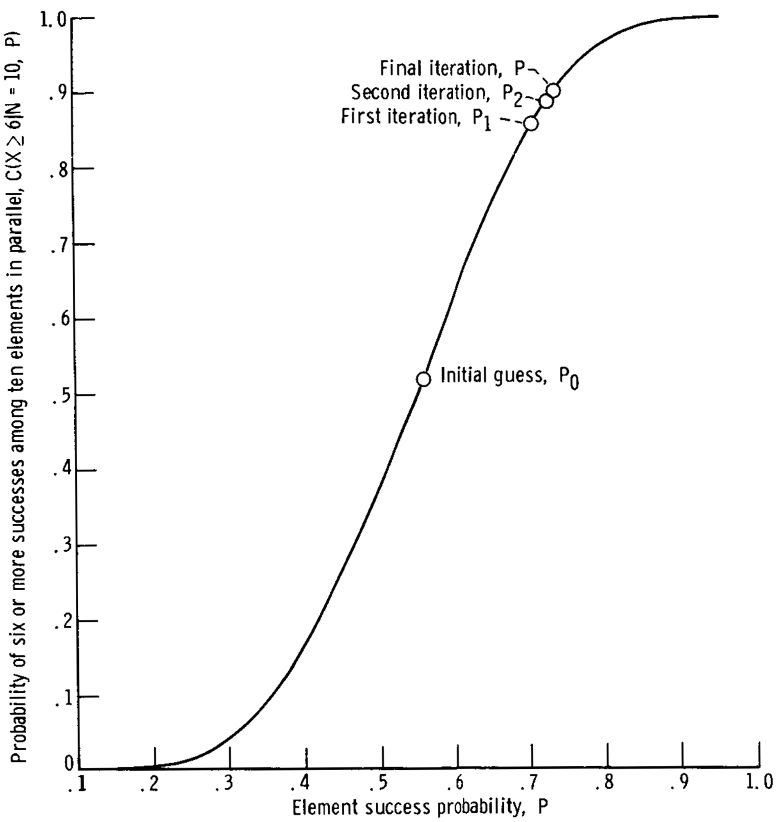

Figure 1. - Element success probability and cumulative probability for six or more successes among ten elements.