

NASA CR-147,800



3 1176 00134 1008

NASA-CR-144880

19790013676

NASA Contractor Report 144880

BETA DISTRIBUTIONS: A COMPUTER PROGRAM FOR  
PROBABILITIES AND FRACTILE POINTS

James D. Brownlow and Ram Swaroop

System Development Corporation  
Edwards, California 93523

May 1979

LIBRARY COPY

MAY 3 1979

LANGLEY RESEARCH CENTER  
LIBRARY, NASA  
HAMPTON, VIRGINIA

Prepared for

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION  
Dryden Flight Research Center  
Edwards, California 93523

## INTRODUCTION

The family of beta distribution with density

$$f(x|a,b,\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \left(\frac{x-a}{b-a}\right)^{\alpha-1} \left(\frac{b-x}{b-a}\right)^{\beta-1} \left(\frac{1}{b-a}\right)$$

with range parameters  $a < x < b$  and shape parameters  $\alpha, \beta > 0$  is an extremely rich family of distributions. It includes J shaped, U shaped and unimodal distributions within finite bounds. In spite of such versatility and applications in physical and life sciences, tables and computer programs for probabilities for the beta distribution are generally not available. One might suspect that this is because the standard beta distribution ( $a=0, b=1$ ) probabilities and fractile points can be computed from F distribution tables which are generally available. However, this is not practical for even a moderate number of probabilities or fractile points, and is limited to integral values of  $\alpha$  and  $\beta$ . In practice beta probabilities and fractile points are often needed. The enclosed computer program meets this need.

The program presented here computes the probability that a beta random variable,  $X$  with shape parameters  $\alpha$  and  $\beta$ , on the range  $a$  to  $b$  will fall within an interval  $(x_1, x_2)$   $a \leq x_1 \leq x_2 \leq b$ . Also for any given probability  $p$  the program will compute the fractile  $x_p = F^{-1}(p)$ . Additionally routines to calculate the density function  $f(x)$  and the gamma function  $\Gamma(x)$  are presented.

## SYMBOLS

$\Gamma(x)$	Gamma Function $\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt$
$X, Y$	Random Variables
$a$	Real Number
$X, y$	Real Numbers
$F$	Distribution Function
$f, b$	Density Functions, or Real Numbers
$r$	Root of a Real Valued Function
$t$	Dummy Variable for Integration
$b, a$	Range Variables for Beta Distribution
$\alpha, \beta$	Shape Parameters for Beta Distribution
$n_1, n_2$	Degrees of Freedom for F-distribution
$\chi^2$	Chi Squared Test Value
$\Sigma$	Summation Notation

## FORTRAN FUNCTIONS

BDIST	Beta Distribution Function
BFRAC	Beta Fractile
F	Beta Density Function
G	Gamma Function



## PROCEDURE

The computation of probabilities and fractiles of beta distribution is based on the following procedure. The general beta density for random variable  $x$  is given by

$$f(x|\alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \left(\frac{x-a}{b-a}\right)^{\alpha-1} \left(\frac{b-x}{b-a}\right)^{\beta-1} \left(\frac{1}{b-a}\right)$$

The range of variable  $x$  is interval  $(a, b)$  and the shape parameters are  $\alpha$  and  $\beta > 0$ . The transformation  $y = (x-a)/(b-a)$  yields standard beta density with range  $(0, 1)$  and density

$$f(y|\alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1} (1-y)^{\beta-1}$$

which takes different shapes (see figure 1) depending, on the values of  $\alpha$  and  $\beta$ . These include J-shaped, U-shaped and symmetric and skewed unimodal distributions.

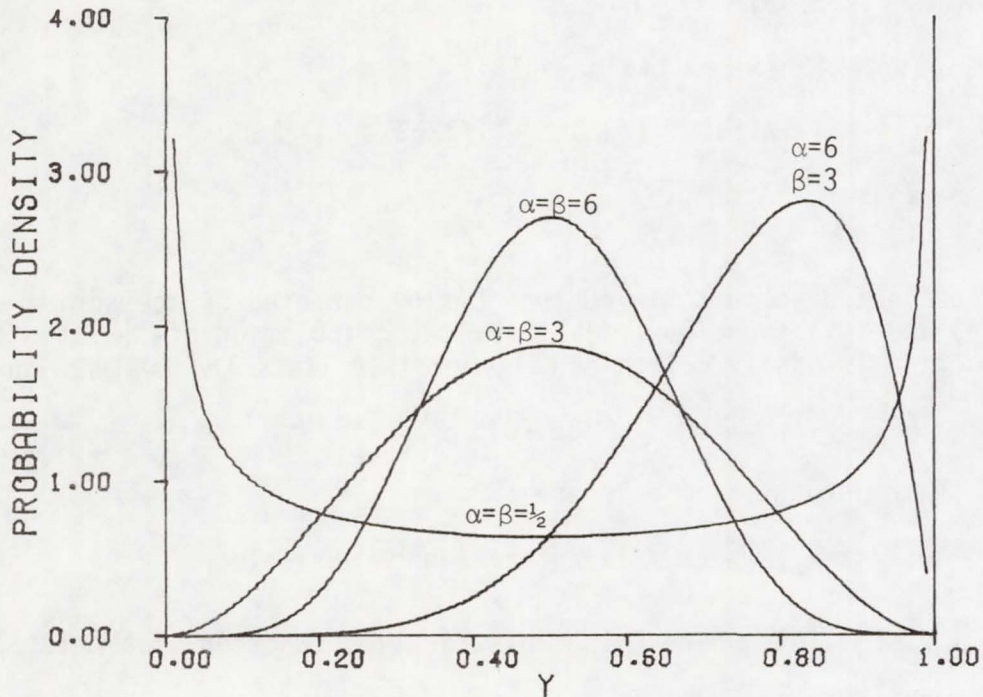


Figure 1.  $p(y) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1} (1 - y)^{\beta-1}$ , the probability density for various values of  $\alpha$  and  $\beta$ .

The beta probabilities are found by numerically integrating the beta density function. The technique for this is Gauss-Legendre quadrature (Ref. 1). The integral is translated to the interval (-1,1) and the Legendre polynomial of order 15 is used.

Having determined the integral of the beta density function, the p-th fractile point  $x_p$  of the beta distribution is obtained by solving the integral equation

$$F(x) = \int_{-\infty}^{x_p} f(t) dt = p$$

The distribution function  $F(x)$  is always a monotone-increasing function. As such, virtually any iterative algorithm will converge to the solution of the equation  $F(x)-p=0$ . The technique used in the programs presented here is the method of bisection (interval halving) (Ref. 2). The procedure terminates at  $x$  whenever  $|F(x_p)-p| < 10^{-6}$ .

The gamma function  $\Gamma(x)$  for argument  $x$  is evaluated by using sterling's approximation (Ref. 3)

$$(*) \quad \Gamma(x) = (2\pi/x) \exp(x \ln x - xg(x))$$

where

$$g(x) = 1 - 1/12x^2 + 1/360x^4 - 1/1260x^6 + 1/1680x^8$$

using  
(\*\*)

$$\Gamma(x) = \frac{\Gamma(x+n)}{x^{(n)}} \quad x^{(n)} = x(x+1)\cdots(x+n-1)$$

$n$  can be selected so that  $x+n > 10$  then  $\Gamma(x+n)$  computed by the series expansion, (\*) and  $\Gamma(x)$  found by (\*\*). From the calculation of  $\Gamma(x)$  the beta density function is easily determined. Appendix A lists the FORTRAN routines.

#### OUTPUTS

The four outputs available are:

$P(x_1 < X < x_2)$  where  $X$  is a beta random variable

$x_p = F^{-1}(p)$ , where  $F$  is a distribution function of a beta random variable

$f(x|a,b,\alpha,\beta)$ , the density function of a beta random variable

$\Gamma(x)$ , the value of the gamma function for positive argument,  $x$ .

These outputs and the parameters for each FORTRAN function are described



below.

### Beta Distribution Function

The distribution function value  $F(x)$  for the beta distribution on the range  $(a,b)$  with shape parameters  $\alpha$  and  $\beta$  is given by

$$\text{BPROB}(0., x, \text{ALPHA}, \text{BETA}, A, B).$$

The probability a beta random variable with parameters  $a,b,\alpha$ , and  $\beta$  takes a value between  $x_1$  and  $x_2$  is given by:

$$\text{BPROB}(x_1, x_2, \text{ALPHA}, \text{BETA}, A, B).$$

Example: The probability that a beta random variable with range  $(1., 3.5)$  and shape parameters  $\alpha=4.1$   $\beta=16.2$  takes a value less than 1.75 is

$$\text{BPROB}(1., 1.75, 4.1, 16.2, 1., 3.5) = .864 .$$

### Beta Distribution Fractile Points

For a given  $p$ , the  $100p$ -th fractile point from the beta distribution with parameters  $a,b,\alpha$ , and  $\beta$  is given by:

$$\text{BFRAC}(P, \text{ALPHA}, \text{BETA}, A, B)$$

Example 1: The 75-th percentile of the beta distribution on the range 240 to 1400, with shape parameters  $\text{ALPHA}=14.2$ ,  $\text{BETA}=34.7$  is

$$\text{BFRAC}(.75, 14.2, 34.7, 240., 1400.) = 559.307$$

Example 2: Appendix B presents a FORTRAN program which uses BFRAC to generate ten equiprobability points of a beta distribution on the range  $(0., 1.)$  with shape parameters  $\text{ALPHA}=4.$ ,  $\text{BETA}=16.$  Thirty beta random variables from this distribution are generated and a  $\chi^2$ -test is performed, at the 5 per cent level, to test the null hypothesis that these observations are, indeed, from the standard (i.e. range  $(0., 1.)$ ) beta distribution with shape parameters  $\text{ALPHA}=4.$ ,  $\text{BETA}=16.$  Not surprisingly, the  $\chi^2$ -test statistic accepts the null hypothesis.

### Beta Density Function

The probability density function  $f(x|\alpha,\beta,a,b)$  for a beta random variable on the range  $(a,b)$  with shape parameters  $\alpha$  and  $\beta$  is given by:

$$F(X, C, \text{ALPHA}, \text{BETA}),$$

where  $C = \Gamma(\text{ALPHA}+\text{BETA})/(\Gamma(\text{ALPHA}) \Gamma(\text{BETA}))$  must be passed to the function.  $C$  may be computed using the gamma function

$$C = \text{GAMMA}(\text{ALPHA} + \text{BETA}) / (\text{GAMMA}(\text{ALPHA}) * \text{GAMMA}(\text{BETA}))$$

Example: The beta density function on the range (.67 to 12.2) with shape parameters ALPHA=1.2, BETA=3.5 at x=5.2 is given by:

$$C = \text{GAMMA}(4.7) / (\text{GAMMA}(1.2) * \text{GAMMA}(3.5)), F(5.2, C, 1.2, 3.5) = 0.1045.$$

## APPENDIX A

This appendix presents the computer programs used for computing the beta-density probabilities and fractile points.



```

FUNCTION BFRAC(X, ALPHA, BETA,A,B)
C****
C****          ROUTINE TO DETERMINE FRACTILE POINTS OF
C****          THE BETA DISTRIBUTION, WITH PARAMETERS
C****          ALPHA AND BETA
C****
C****          INVERSE DISTRIBUTION VALUE FOUND BY
C****          HALVING-THE-INTERVAL TECHNIQUE..
C****
C****          BY BROWNLAW, SDC/ISI.
C****
C****          MAKE SURE PERCENTAGE POINT IS BETWEEN 0 AND 1..
C****
IF( X .LE. 1. .AND. X .GE. 0. ) GO TO 2
C****
PRINT 1,X
1  FORMAT(" X = ",E11.3," FOR BETA DIST*N, SO FRACTILE POINT COULD N
-OT BE FOUND")
C****
BFRAC = 0.
RETURN
2  KOUNT = 0
C****
TOP = 1. - 1.E-50
BOT = 1.E-50
C****
3  BFRAC = (TOP + BOT)/2.
F1 = BPROB(1.,BFRAC,ALPHA,BETA,0.,1.) - X
C****
C****          IF THE INTEGRAL OF THE BETA FUNCTION FROM ZERO TO
C****          BFRAC DIFFERS FROM X BY LESS THAN 1.E-9
C****          WE HAVE THE FRACTILE POINT..
C****
IF(ABS(F1) .LE. 1.E-9) GO TO 4
C****
C****          OTHERWISE RESET THE UPPER OR LOWER INTERVAL
C****          VALUE AND REPEAT THE PROCEDURE..
C****
IF( F1 .GT. 0.) TOP = BFRAC
IF( F1 .LT. 0.) BOT = BFRAC
C****
C****          IN ANY CASE NEVER DO MORE THAN 20 ITERATIONS
C****          TO FIND A VALUE ON (0.,1.) TO WITHIN 2**(-19).
C****
KOUNT = KOUNT + 1
IF(KOUNT .GT. 20) GO TO 4
C****
GO TO 3.
C****
C****          REDEFINE FRACTILE POINT FOR DISTRIBUTION ON THE
C****          RANGE A TO B..
C****
4  BFRAC = A + BFRAC*(B-A)
C****
C****
RETURN
END

```

```

FUNCTION BPROB(X1,X2,ALPHA,BETA,A,B)
C****
C****          ROUTINE TO INTEGRATE BETA DENSITY FUNCTION
C****          DEFINED ON (A,B) WITH PARAMETERS ALPHA, AND
C****          BETA. LIMITS OF INTEGRATION: X1 TO X2.
C****
C****          BY BROWNLAW, SDC/ISI, 1/79
C****
EXTERNAL F
C****
C****          F IS THE BETA DENSITY FUNCTION,
C****           $F(X,C,ALPHA,BETA) = C * ((X-A)/(B-A))^{ALPHA-1.} * ((B-X)/(B-A))^{BETA-1.} / (B-A)$ 
C****
C****           $C = GAMMA(ALPHA+BETA) / (GAMMA(ALPHA) * GAMMA(BETA))$ 
C****
C****          UNDER THE TRANSFORMATION  $Y = (X-A)/(B-A)$ ,
C****          Y HAS STANDARD DENSITY,
C****           $F(Y) = C * (Y^{ALPHA-1.} * (1.-Y)^{BETA-1.})$ 
C****
IF(A .LT. B) GO TO 2
PRINT 1,A,B
1  FORMAT(* IN ESTIMATING BETA PROBABILITY, THE LOWER BOUND,*,E12.4,
-/* IS NOT GREATER THAN THE UPPER BOUND,*,E12.4)
BPROB = 0.
RETURN
C****
2  IF(X1 .LT. X2) GO TO 4
PRINT 3,X1,X2
3  FORMAT(* IN ESTIMATING BETA PROBABILITY, THE LOWER LIMIT,*,E12.4,
-/* IS NOT GREATER THAN THE UPPER LIMIT,*,E12.4)
BPROB = 0.
RETURN
C****
4  IF(X1 .GE. A) GO TO 6
PRINT 5,X1,A
5  FORMAT(* IN ESTIMATING THE BETA PROBABILITY, THE LOWER LIMIT,*,
-E12.4,/,* IS LESS THAN THE LOWER BOUND OF THE DISTRIBUTION,*
-E12.4,* THE LOWER BOUND WILL BE USED*)
X1 = A
C****
6  IF(X2 .LE. B) GO TO 8
PRINT 7,X2,B
7  FORMAT(* IN ESTIMATING THE BETA PROBABILITY, THE UPPER LIMIT,*,
-E12.4,/,* IS GREATER THAN THE UPPER BOUND OF THE DISTRIBUTION*
-E12.4,* THE UPPER BOUND WILL BE USED.*)
X2 = B
C****
8  Y1 = (X1-A)/(B-A)
Y2 = (X2-A)/(B-A)
C****
C****          GAUSSLO DOES GAUSSIAN-LEGENDRE QUADRATURE TO
C****          ESTIMATE THE INTEGRAL OF THE BETA DENSITY FUNCTION.
C****
CALL GAUSSLO(F,ALPHA,BETA,Y1,Y2,ANS )
C****
BPROB = ANS

RETURN
END

```



```

SUBROUTINE GAUSSLQ(F,ALPHA,BETA,X1,X2,ANS)
C****
C****          ROUTINE TO DO GAUSSIAN-LEGENDRE QUADRATURE
C****          TO INTEGRATE BETA DENSITY FUNCTION FROM X1
C****          TO X2
C****
C****          F IS THE STANDARD BETA DENSITY FUNCTION...
C****          F(X,C,ALPHA,BETA)  WHERE C IS NORMALIZATION
C****          CONSTANT, C= GAMMA(ALPHA+BETA)/(GAMMA(ALPHA)
C****          *GAMMA(BETA))
C****
C****          EY BROWNLOW, SOC/ISI, 3/79
C****
      DIMENSION ROOTS(8),W(8)
      DOUBLE PRECISION ROOTS,W,ANSWER
      C = GAMMA(ALPHA + BETA)/(GAMMA(ALPHA)*GAMMA(BETA))
C****
      DATA ROOTS /
-      (.JDI          , 0.20119 40939 9743500,.39415 13470 7756300,
-      0.57097 21726 0853900,.72441 77313 6017000,.84820 65834 1042700,
-      0.93727 33924 0070600,.98799 25180 2048500/
C****
      DATA W /
-      0.20257 32419 2556100,.19843 14853 2711100,.18616 10001 1556200,
-      0.16626 92058 1699400,.13957 06779 2615400,.10715 92204 6717200,
-      0.07036 60474 8810800,.03075 32419 9611700/
C****
C****
C****          TRANSLATE THE INTERVAL (X1,X2) TO (-1,1)
C****
      ANSWER = W(1)* F( (X1+X2)/2.,C,ALPHA,BETA)
C****
C****
      DO 1 I=2,8
C****
      ANSWER=ANSWER + W(I)* F(((X2-X1)*ROOTS(I) + (X1+X2))/2.,C,
-          ALPHA,BETA)
C****
      ANSWER=ANSWER + W(I)* F(((X2-X1)*(-1.)*ROOTS(I) + (X1+X2))/2.,C,
-          ALPHA,BETA)
C****
      1 CONTINUE
C****
C****
      ANS=ANSWER*(X2-X1)/2.
C****
      RETURN
      END

```

```
FUNCTION F(X,C,ALPHA,BETA)
C****
C****          F IS THE BETA DENSITY FUNCTION...
C****
C****          C PASSED IN MUST HAVE THE VALUE
C****          C = GAMMA(ALPHA+3BETA)/(GAMMA(ALPHA)*GAMMA(BETA))
C****
C****          BY BROWNLW, SDC/ISI, 1/18/79
C****
C****          F = C*(X**(ALPHA-1.)*(1.-X)**(BETA-1.))
C****
RETURN
END
```



FUNCTION GAMMA(X)

```

C****
C****          WRITTEN BY BROWNLOW,SDC/ISI, JUNE 1978
C****
C****          GAMMA FUNCTION, REAL ARGUMENT MUST BE
C****          GREATER THAN ZERO.
C****
C****          G(X)= INTEGRAL (EXP(-T)*T**(X-1) DT
C****          EVALUATED BY EXPANSION OF STIRLING'S
C****          FORMULA
C****
C****          GAMMA(X) = GAMMA(X+N)/(X*(X+1)*...*(X+N))
C****          WHERE X+N IS CHOSEN TO GET MAXIMUM
C****          ACCURACY FROM THE ALGORITHM. SEE
C****          HENRICI, "COMPUTATIONAL ANALYSIS
C****          WITH THE HP-25 POCKET CALCULATOR"
C****
C****          A = X
C****          P = 1.
C****
1  IF( A .GT. 10. ) GO TO 2
   P = P*A
   A = A + 1.
   GO TO 1
C****
2  CONTINUE
C****
E= 1-(1./((12.*A*A))+1./((360.*A*A*A*A)- 1./((1260.*
. A*A*A*A*A*A) + 1./((1680.*A*A*A*A*A*A*A*A)
C****
C****          DATA PI/3.1415926538/
C****
C****          GAMMA = ( SQRT(2.*PI/A) * EXP( A*(ALOG(A)-E) ) )/P
C****
C****          RETURN
C****          END

```

APPENDIX B



This appendix presents the FORTRAN program which generates the ten equi-probability points of the standard beta distribution with shape parameters  $\alpha=4$ ,  $\beta=16$ . Additionally thirty random variables from this distribution are generated and a  $\chi^2$  goodness-of-fit test is performed.

```

PROGRAM MAIN(OUTPUT)
C****
C****          ROUTINE TO FIND ORDINATE VALUES
C****          OF THE BETA DISTRIBUTION WHICH
C****          DIVIDE THE (0., 1.) INTERVAL
C****          INTO TEN SUBINTERVALS OF EQUAL
C****          PROBABILITY..
C****
C****          BETA DISTRIBUTION WITH ALPHA = 4.
C****          BETA = 16.
C****
C****          BY BROWNLCW, SOC/ISI.
C****
DO 2 I=1,9
X = FLOAT(I)/10.
PRINT 1, BFRAC(X,4.,16.,0.,1.)
1  FORMAT(* *F10.5)
2  CONTINUE
C****
C****          AND NOW GENERATE 30 SAMPLES
C****          FROM A STANDARD BETA DISTRIBUTION
C****          WITH PARAMETERS ALPHA = 4.,
C****          AND BETA = 16.
C****
DO 4 I=1,30
PRINT 3, BETAGEN(4,16,0.,1.)
3  FORMAT(* *F5.3)
4  CONTINUE
END

```

```

FUNCTION BETAGEN(IALPHA, IBETA, A, B)
C****
C****          FUNCTION TO GENERATE A RANDOM
C****          VARIABLE FROM A BETA
C****          DISTRIBUTION ON THE RANGE
C****          (A, B) WITH SHAPE PARAMETERS
C****          IALPHA, AND IBETA.
C****
C****          METHOD IS FROM HASTINGS AND
C****          PEACOCK, "STATISTICAL DIST-
C****          RIBUTIONS", PAGE 72.
C****
      G1 = 0.
      DO 1 I=1, IALPHA
      G1 = G1 - ALOG( RANF(0) )
1     CONTINUE
C****
      G2 = 0.
      DO 2 I=1, IBETA
      G2 = G2 - ALOG( RANF(0) )
2     CONTINUE
C****
      BETAGEN = A + (G1/(G1 + G2))*(B-A)
C****
C****          NOTE THAT G1 AND G2 ARE, IN
C****          FACT, GAMMA RANDOM VARIABLES.
C****
      RETURN
      END

```

The equiprobability points for this distribution which partition the range (0,1) into ten intervals are: .09514, .12334, .14652, .16817, .18989, .21297, .23907, .27131, .31895, 1.00000.

The thirty beta random variables generated are: .096, .238, .147, .230, .234, .136, .202, .074, .338, .153, .216, .167, .370, .350, .167, .192, .171, .229, .144, .056, .257, .201, .137, .235, .181, .339, .156, .122, .235, .137.



Given the equi-probability points and the random sample found by using the computer programs given in this paper, the null hypothesis that the distribution from which these random variables are drawn is beta on (0,1) with shape parameters  $\alpha=4$ ,  $\beta=16$ . may be tested. Specifically,

<u>Class</u>	<u>Observed Freq.</u>	<u>Expected</u>
0.00000-0.09514	2	3
0.09514-0.12334	2	3
0.12334-0.14652	4	3
0.14652-0.16817	5	3
0.16817-0.18989	2	3
0.18989-0.21297	3	3
0.21297-0.23907	6	3
0.23907-0.27131	1	3
0.27131-0.31859	0	3
0.31859-1.00000	4	3

Each interval has equal probability (in this case 0.10) so with a random sample of size thirty we expect three observations in each interval.

The test statistic is given by

$$\chi^2 = \sum_{i=1}^{10} (O_i - E_i)^2 / E_i$$

where  $O_i$  is the observed number of occurrences in the  $i^{\text{th}}$  cell;  $E_i$  the expected number.

$$\begin{aligned} \text{hence } \chi^2 &= \frac{1}{3} (1^2 + 1^2 + 1^2 + 2^2 + \dots + 1^2) \\ &= \frac{1}{3} (31) = 10.333 \end{aligned}$$

$$\text{and } \chi^2_{.05}(9) = 16.919$$

thus we accept  $H_0$ : the distribution from which the sample was drawn is  $\beta(4,16)$

#### REFERENCES

1. Carnahan, Brice, et. al., Applied Numerical Methods, John Wiley & Sons, Inc., N.Y., 1969.
2. Franklin, Phillip, A Treatise on Advanced Calculus, Dover Publications N.Y., 1964.
3. Henrici, P., Computational Analysis on the HP-25 Pocket Calculator, John Wiley & Sons, Inc., N.Y., 1977.



1. Report No. NASA CR-144880	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle BETA DISTRIBUTIONS: A COMPUTER PROGRAM FOR PROBABILITIES AND FRACTILE POINTS		5. Report Date May 1979	
		6. Performing Organization Code	
7. Author(s) James D. Brownlow and Ram Swaroop		8. Performing Organization Report No.	
		10. Work Unit No.	
9. Performing Organization Name and Address System Development Corporation Edwards, California 93523		11. Contract or Grant No. NAS4-2334	
		13. Type of Report and Period Covered Contractor Report-Topical	
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546		14. Sponsoring Agency Code H-1082	
		15. Supplementary Notes NASA Technical Monitor: William R. Winter, Dryden Flight Research Center	
16. Abstract  A beta distribution is specified by range parameters $a < b$ , and two shape parameters $\alpha$ and $\beta > 0$ . The computer program presented calculates any desired probability and/or fractile point for specified values of $a$ , $b$ , $\alpha$ and $\beta$ . This program additionally computes gamma function values for integer and non integer arguments.			
17. Key Words (Suggested by Author(s)) Beta distribution Probabilities Fractile points		18. Distribution Statement  Unclassified-Unlimited  STAR Category: 65	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 20	22. Price* \$3.25

\*For sale by the National Technical Information Service, Springfield, Virginia 22161