UM - 2

FREQUENCY ANALYSIS

SATISH CHANDRA
DIRECTOR

STUDY GROUP

S M SETH

N K GOEL

NATIONAL INSTITUTE OF HYDROLOGY

JAL VIGYAN BHAVAN

ROORKEE-247667

1984-85

CONTENTS

LIST OF TABLES

ABSTRACT

The user's manual gives the details of a computer programme for best fit distribution using normalization procedures and chi-square criterion. Various normalization procedures which have been used in the programme are a. Inverse Pearson type III transformation, 2. Log normal transformation based on theoretical relationships, 3. Log transformation, 4. Inverse log Pearson type III transformation and 5. Square root transformation. The best fit normalization procedure corresponding to each season is chosen based on chi-square statistic. The programme is capable of comparing any combination of above mentioned normalization procedures. The same programme can be used for any type of seasonal hydrologic data i.e. daily, pentad, weekly, ten daily, monthly or annual. The data should be continuous and observations across the years should be independent.

The computer programme is written in Fortran IV. The definitions of the terminology related to normalization procedures and the chi-square statistic have been given in the manual. The input data specifications and output description are explained. The user's manual also contains the test data and example calculations.

1.0 INTRODUCTION

One of the major problems faced in hydrology is the estimation of design flood from fairly short data. If the length of data is more say 1000 years then the same data can be used to find out design flood but the length of data generally available is very less. So the sample data is used to fit frequency distribution which in turn is used to extrapolate from recorded events to design events either graphically or by estimating the parameters of frequency distribution.

Graphical method has the advantage of simplicity and visual presentation. But the main disadvantage is that different engineers will fit different curves.

Hydrologic data in its original form are rarely normally distributed. Many of the hydrologic data are positively skewed and so the application of normal distribution to the original data is not appropriate. However, transformation methods are availaoble to transform the data to normal distribution. Various normalization procedures are available in practice. Many transformation procedures require the use of computer as the transformation procedure involved is cumbersome to apply with a calculator. Since the sample size of the hydrologic data generally available is limited, it is desirable to transform the data to normality using many transformation procedures and select the best fitting normalization procedure based on some fitting cirterion.

Log normal transformation, inverse Pearson type III transformation, square root transformation, inverse log Pearson type III transformation and Box-Cox (Power) transformation etc. are some of the commonly used transformations for normalization.

Chi-square test, Kolmorov-Smirnov test and Cramer-Von Mises test are some of the well known tests to judge the goodness of fit.

A computer programme developed by Dr M Krishnaswamy at I I T Kanpur has been modified and improved after incorporating suitable changes.

1.1 Purpose and Capabilities of the Programme

The purpose of the programme is to find out best fitting distribution after testing various normalization procedures on the basis of chi-square statistic for different seasons/months of the year. The programme compares the following normalization procedures:

(a) Normal distribution by method of moments

(b) Inverse Pearson type III transformation

(c) Log normal transformation (parameters are obtained on the basis of theoretical relationships)

(d) Log transformation(parameters are estimated by method of moments)

(e) Inverse log Pearson type III transformation

(f) Square root transformation

The programme gives the following statistics for the desired transformations:

(i) Mean of the transformed series

(ii) Standard deviation of the transformed series

(iii) Coefficient of skewness of the transformed series

(iv) Chi-square statistic for the transformation

(v) Number of degrees of freedom

The programme can be applied to data sets other than weekly/monthly/seasonal time series.

The number of data points should not be more than 500. However, if the number of data points is more, the dimension statements of the programme can be changed accordingly.

1.2 Definitions of the Terminology Related to Topic and Components

Various transformations used for normalization in the programme are described below:

(a) Normal distribution by method of moments:

The data as such is compared with the normal distribution.

(b) Inverse Pearson type III transformation

The following equation is used to normalize the data:

$$Y = \left(\left(\frac{C_S}{2}\left(\frac{X-\mu}{\sigma}\right)+1\right)^{1/3} - 1\right)\frac{6}{C_S} + \frac{C_S}{6} \quad ...(1)$$

where:

X : Original series

μ : Mean of the original series

σ : Standard deviation of the original series

C_S : Coefficient of skewness of the original series

Y : Pearson type III transformed series

(c) Log normal distribution (parameters estimation on the basis of theoretical relationships)

Parameters of the log transformed series are calculated on the basis of following theoretical relationships:

$$\mu_y = \log_e(\mu_x) - 0.5 \log_e((\sigma_x/\mu_x)^2 + 1) \quad ...(2)$$

$$\sigma_y = (\log_e((\sigma_x/\mu_x)^2 + 1))^{1/2} \quad ...(3)$$

where:

μ_x : Mean of the original series

σ_x : Standard deviation of the original series

μ_y : Mean of log (base e) transformed series

σ_y : Standard deviation of log(base e) transformed series

The parameters so obtained are used for the calculation of chi-square statistic.

(d) Log transformation

$$Y = \log_e X \qquad \ldots (4)$$

where :

Y : Log (base e) transformed series

X : Original series

(e) Inverse log Pearson type III transformation

In inverse log Pearson type III transformation, log transformed series is used instead of original series. Rest of the procedure for transformation is similar to that for inverse Pearson type III transformation.

(f) Square root transformation

In this procedure square root of the original series is used as transformed series. The following equation is used.

$$Y = (X)^{1/2} \qquad \ldots (5)$$

where :

Y : Square root transformed series

X : Original series

1.3 Scope

The programme can be used to analyse any type of seasonal hydrologic data i.e.daily, pentad, weekly, ten daily, monthly, annual, annual peaks etc. The programme gives the statistical parameters e.g.mean, standard deviation, coefficient of skewness, chi-square value and number of degrees of freedom for different(desired) normalization procedures for various seasons of the year.

1.4 Hardware and Software Requirements

FORTRAN compiler and simple FORTRANinstructions are required to run the programme.

2.0 SPECIFIC METHOD

(i) Method of computation

Steps for the computation of chi-square statistic are given in the following section.

(a) The probability limits for each class are calculated. Equal probability is assigned to each class. The probability limit for the last class is 0.9999. If the number of classes is six then probability limits for the classes will be as given under:

Class	Lower limit	Upper limit
1	0.0	1/6
2	1/6	2/6
3	2/6	3/6.
4	3/6	4/6
5	4/6	5/6
6	5/6	0.9999

(b) Standard normal variates corresponding to class limits are calculated.

(c) The theoretical frequency for each class is calculated. The theoretical frequency of each class will be equal to the number of data points in a particular season divided by the number of classes.

(d) The programme sorts out the data for a particular season, e.g. in case of weekly data, the programme will sort out the data for week one. In case of monthly data, the programme will sort out the data for month one.

(e) The programme arranges the seasonal data in descending order.

(f) Various transformations are tested on this arranged seasonal data.

(g) For the first desired transformation, the programme calculates the mean, standard deviation and coefficients of skewness using the following equations:

$$\mu_i = \sum_{m=1}^{N} X_{i,m}/N \qquad \text{...(6)}$$

$$\sigma_i = \left(\sum_{m=1}^{N} (X_{i,m} - \mu_i)^2 / N-1 \right)^{1/2} \qquad \text{...(7)}$$

$$C_{s_i} = \sum_{m=1}^{N} (X_{i,m} - \mu_i)^3 / (N-1)(N-2)\sigma_i^3) \qquad \text{...(8)}$$

where:

X_i : Transformed data for the i^{th} season

μ_i : Mean of the transformed data

N : Total number of years in the record

σ_i : Unbiased estimate of the standard deviation

C_{s_i} : Unbiased estimate of the skew coefficient

i : Season number

m : Year number

(h) After computing statistical parameters of the arranged and transformed seasonal data, the programme calculates the chi-square for the transformation.

(i) Steps through g to h are repeated for other desired transformations. The transformation giving the least chi-square value is the best normalizing procedure for that season.

Steps through d to i are repeated for other seasons.

(ii) Test of goodness of fit

The chi-square test, devised by K Pearson, is used to compare various transformations/normalization procedures. The chi-square statistic is given by:

$$\chi^2 = \sum_{j=1}^{K} \frac{(0_j - E_j)^2}{E_j} \qquad \text{...(9)}$$

where :

K : Number of classes

O_j : Observed frequency in the j^{th} class

E_J : Expected frequency in the j^{th} class

The chi-square test prescribes the critical values χ^2_c for a given significance level and number of degrees of freedom so that for $\chi^2 \leq \chi^2_c$, the null hypothesis of a good fit is accepted and for $\chi^2 > \chi^2_c$ the same hypothesis is rejected. Percentile values ($\chi^2_{\alpha,\nu}$) for the chi-square distribution with ν degrees of freedom(shaded area= α) have been given in table 1.

(iii) Selection of n umber of classes

The observed data can be classified into mutually exclusive and exhaustive class intervals to test the goodness of fit of various transformations.

There is no specified rule for the number of classes to be adopted. If too many classes are chosen, some of them would have few or no frequencies and the resulting frequency distribution would be irregular. Similarly the observed data would be very much compressed and the large number of proportion of the frequencies will fall in one or two classes, resulting in the loss of information, if the number of classes is too less.

The number of classes should not be less than 6 and more than 20, though these rules do not have any theoretical basis. The number of classes should be selected in such a way that theoretical frequency of each class is not less than 5. The number of class intervals should be selected in such a way that the main characteristic features of the observed distribution are emphasized and chance variations are obscured.

(iv) Number of degrees of freedom

The term degrees of freedom is defined as a comparison between the data, independent of other comparisons in the analysis. Each observation in a data of size n can be compared with(n-1) other observations and hence there are (n-1) degrees of freedom. The degrees of freedom is equal to (K-m-1). Here K is the number of classes and m is the number of parameters. Since various transformations are compared with normal distribution which is a two parameter distribution, the number of degrees of freedom is always (K-3).

2.1 General Description/Programme Description

The computer programme for best fit distribution consists of one main routine and ten subroutines. The subroutines have been described below:

(i) SUBROUTINE NDTRI (P,X,C, IER)

This purpose of this subroutine is to calculate standard normal variate corresponding to a given probability. The calling arguments are:

P : Input probability

X : Output argument such that P=Y=the probability that the random variable is less than or equal to 0.0

C : Output density function F(X)
IER = -1; if P is not in the interval between (0,1)
IER= 0; if there is no error

(ii) SUBROUTINE MSS(X, AMEAN, STDEV, SKEW)

This subroutine calculates mean, standard deviation and coefficient of skewness of the given series. Various calling arguments are:

X : Given series

AMEAN : Mean of the series

STDEV : Standard deviation of the series

SKEW : Coefficient of skewness of the series

(iii) SUBROUTINE CSS (X, AMEAN, STDEV, SKEW, THF, T)

This subroutine calculates chi-square value and degrees of freedom for a transformation. Various calling arguments are:

X : Given series

AMEAN : Mean of the series, calculated from subroutine MSS

STDEV : Standard deviation of the series, calculated from the subroutine MSS

SKEW : Coefficient of skewness of the series calculated from subroutine MSS

THF : Theoretical frequency of each class

T : Standard normal variate

(iv) SUBROUTINE NORMAL (CX,L,THF,T,AMEAN,STDEV, SKEW)

This subroutine analyses the series for normal distribution . The calling arguments are:

CX : Input series

L : Number of years

THF : Theoretical frequency

T : Standard normal variate

AMEAN : Mean of the transformed series

STDEV : Standard deviation of the transformed series

SKEW : Coefficient of skewness of the transformed series

(v) SUBROUTINE PT3 (CX, L,THF,T,AMEAN,STDEV,SKEW)

This subroutine analyses inverse Pearson type III transformation.

(vi) SUBROUTINE LNC(CX,L,THF,T,AMEAN,STDEV,SKEW)

This subroutine analyses log normal distribution for which parameters are calculated on the basis of theoretical relationships.

(vii) SUBROUTINE LN(CX,L,THF,T,AMEAN,STDEV,SKEW)

This subroutine analyses log transformation. The parameters are calculated by method of moments.

(viii) SUBROUTINE LP3 (CX,L,THF,T,AMEAN,STDEV,SKEW)

This subroutine analyses inverse log Pearson type III transformation.

(ix) SUBROUTINE SQRTT (CX,L,THF,T,AMEAN,STDEV,SKEW)

This subroutine analyses square root transformation.

(x) SUBROUTINE SORTX(N,X)

This subroutine arranges the data in descending order.

The calling arguments are:

N : Total number of observations in the series

X : Series to be arranged in descending order

Note: Various calling arguments in subroutine PT3,LNC,LN, LP3,and SQRTT are same.

2.2 Data Requirement

Continuous data of hydrologic variable is required.

2.3 Analysis

Annual peak discharge (Cumecs) data of river Narmada at Mortakka has been analysed.

The statistical parameters of original, log transformed (parameters estimation by method of moments) and inverse log Pearson type III transformed series are given below:

(i) Original series:

Mean = 25935.750 Chi-square = 3.2500

Standard deviation = 11615.848 Number of degrees of freedom =3

Coefficient of skewness = 1.044

(ii) Log transformed series:

Mean = 10.072 Standard deviation = 0.433

Coefficient of skewness = 0.105 Chi-square = 0.2500

Number of degrees of freedo m = 3

(iii) Inverse log Pearson type III transformed series:

Mean	=	0.001
Standard deviation	=	1.00
Coefficient of skewness	=	0.030
Chi-square	=	0.2500
Number of degrees of freedom	=	3

(Original series, log transformed series and inverse log Pearson type III transformed series have been chosen just to illustrate the calculations)

The chi-square values for normal distribution, Pearson type III distribution, log normal distribution(parameters on the basis of theoretical relationships), log normal distribution, log Pearson type III distribution and square root distribution are 3.2500, 1.0000, 1.0000, 0.2500, 0.2500 and 0.6250 respectively. From table 1, the critical value of chi-square statistic at 95% probability level and for 3 degrees of freedom is 7.81. On the basis of chi-square statistic all the distributions are fitting to annual peak discharge data of river Narmada at Mortakka. However, the chi-square is minimum in case of log normal distribution and log Pearson type III distribution. Since log normal distribution is a special case of log Pearson type III distribution, it can be assumed that the data follows log Pearson type III distribution.

2.4 Advantages and Limitations

2.4.1 Advantages

Different seasons of year have different statistical parameters and probability distributions. The programme is capable of analysing different seasons of the year.

2.4.2 Limitations

(a) The number of classes should be chosen in such a way that atleast 5 observations are there in each class.

(b) The length of data should not be less than 30 years (approx.).

(c) The data should be continuous without any gap.

(d) The observations for a season should be independent.

(e) The drawbacks of chi-square index as a criterion for best fit distribution are associated with the programme also.

2.5 Programme Details

The flow chart for the main programme and subroutines is given in Appendix I.

The source programme listing, input specifications, and output description are given in Appendix II, III and IV respectively. Test data (input data file and output file are in Appendix V, VI and VII respectively).

3.0 RECOMMENDATIONS

The programme for best fit distribution using normalization procedures and chi-square criterion can be used for any type of seasonal hydrologic data e.g.daily, pentad, weekly, ten daily, monthly or annual. The data should be continuous and observations across the years should be independent. The programme has been implemented and tested on VAX-11/780 computer system of National Institute of Hydrology ,Roorkee. The programme can be run on any other system also with simple FORTRAN instructions and with minor or no modifications.

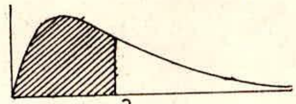
REFERENCES

1. Krishnaswamy, M.(1979), " Computer Programme for Best Fit Distribution Using Normalization Procedures and Chi-square Criterion", Indian Institute of Technololgy, Kanpur(Unpublished)

2. Haan, C.T.(1977),"Statistical Methods in Hydrology", Iowa State University Press, Ames, IOWA.

TABLE 1

Percentile Values ($\chi^2_{\alpha,\nu}$) for the Chi-Square Distribution with ν Degrees of Freedom (shaded area = α).



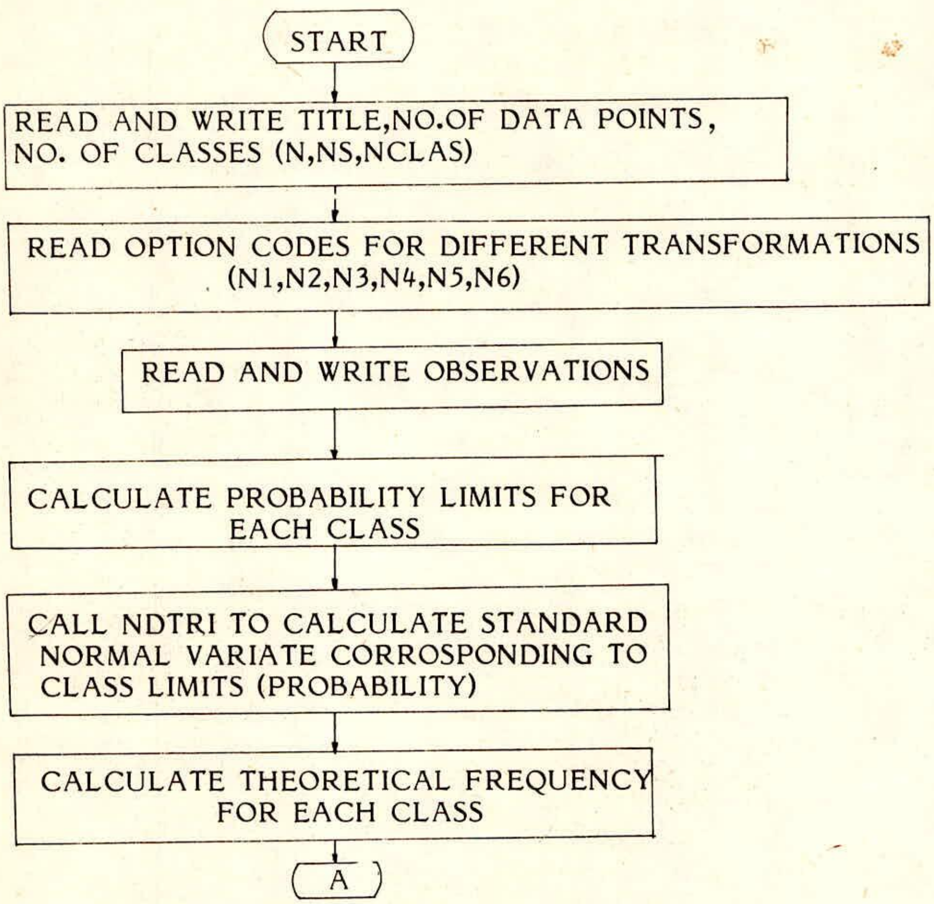
$\chi^2_{\alpha,\nu}$

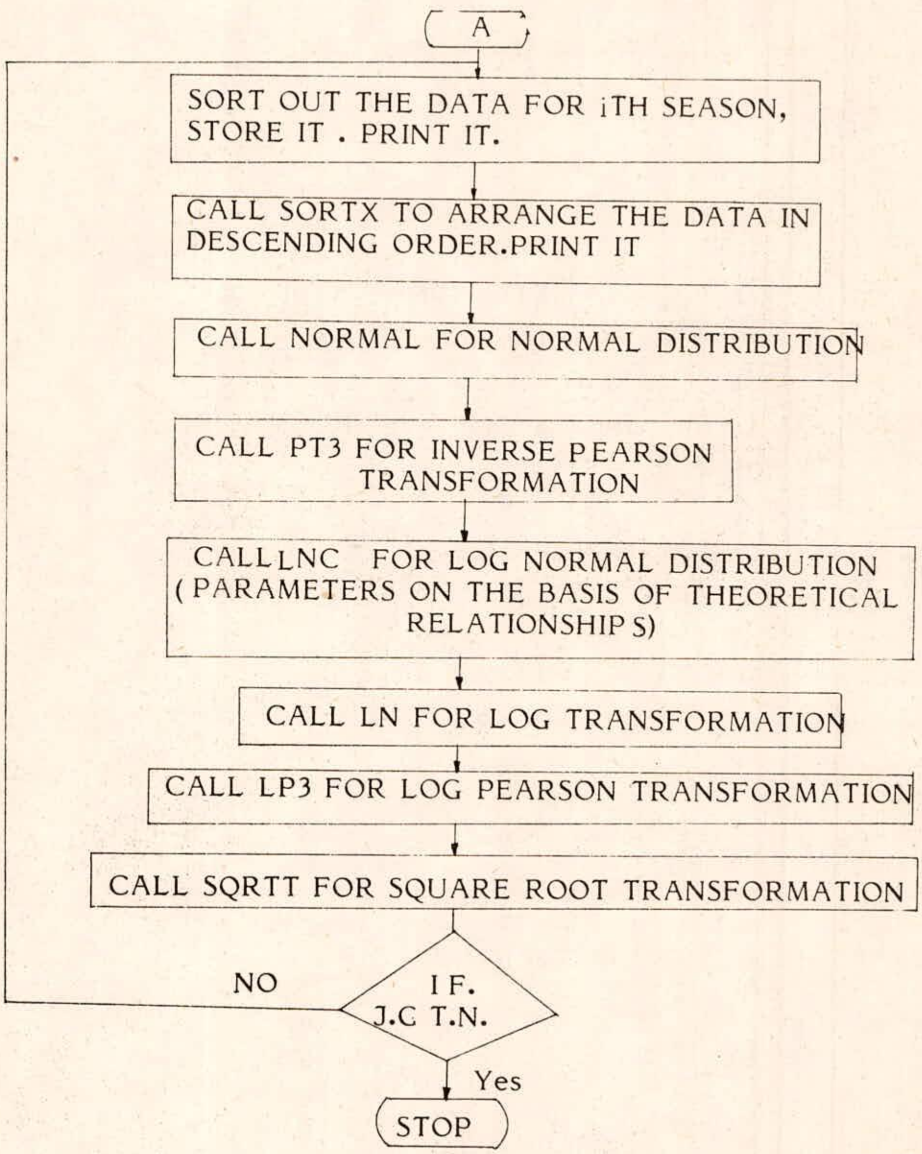
ν	$\chi^2_{.995}$	$\chi^2_{.99}$	$\chi^2_{.975}$	$\chi^2_{.95}$	$\chi^2_{.90}$	$\chi^2_{.75}$	$\chi^2_{.50}$
1	7.88	6.63	5.02	3.84	2.71	1.32	.455
2	10.6	9.21	7.38	5.99	4.61	2.77	1.39
3	12.8	11.3	9.35	7.81	6.25	4.11	2.37
4	14.9	13.3	11.1	9.49	7.78	5.39	3.36
5	16.7	15.1	12.8	11.1	9.24	6.63	4.35
6	18.5	16.8	14.4	12.6	10.6	7.84	5.35
7	20.3	18.5	16.0	14.1	12.0	9.04	6.35
8	22.0	20.1	17.5	15.5	13.4	10.2	7.34
9	23.6	21.7	19.0	16.9	14.7	11.4	8.34
10	25.2	23.2	20.5	18.3	16.0	12.5	9.34
11	26.8	24.7	21.9	19.7	17.3	13.7	10.3
12	28.3	19.8	23.3	21.0	18.5	14.8	11.3
13	29.8	27.7	24.7	22.4	29.6	16.0	12.3
14	31.3	29.1	26.1	23.7	21.1	17.1	13.3
15	32.8	30.6	27.5	25.0	22.3	18.2	14.3
16	34.3	32.0	28.8	26.3	23.5	19.4	15.3
17	35.7	33.4	30.2	27.6	24.8	20.5	16.3
18	37.2	34.8	31.5	28.9	26.0	21.6	17.3
19	38.6	36.2	32.9	30.1	27.2	22.7	18.3
20	40.0	37.6	34.2	31.4	28.4	23.8	19.3
21	41.4	38.9	35.5	32.7	29.	24.9	20.3
22	42.8	40.3	36.8	33.9	30.8	26.0	21.3
23	44.2	41.6	38.1	35.2	32.0	27.1	22.3
24	45.6	43.0	39.4	36.4	33.2	28.2	23.3
25	46.9	44.3	40.6	37.7	34.4	29.3	24.3
26	48.3	45.6	41.9	38.9	35.6	30.4	25.3
27	49.6	47.0	43.2	40.1	36.7	31.5	26.3
28	51.0	48.3	44.5	41.3	37.9	32.6	27.3
29	52.3	49.6	45.7	42.6	39.1	33.7	28.3
30	53.7	50.9	47.0	43.8	40.3	34.8	29.3
40	66.8	63.7	59.3	55.8	51.8	45.6	39.3
50	79.5	76.2	71.4	67.5	63.2	56.3	49.3
60	92.0	88.4	83.3	79.1	74.4	67.0	59.3
70	104.2	100.4	95.0	90.5	85.5	77.6	69.3
80	116.3	112.3	106.6	101.9	96.6	88.1	79.3
90	128.3	124.1	118.1	113.1	107.6	98.6	89.3
100	140.2	135.8	129.6	124.3	118.5	109.1	99.3

ν	$\chi^2_{.25}$	$\chi^2_{.1}$	$\chi^2_{.05}$	$\chi^2_{.025}$	$\chi^2_{.01}$	$\chi^2_{.005}$
1	.102	.0158	.0039	.001	.0002	.0000
2	.575	.211	.103	.050	.0201	.0100
3	1.21	.584	.352	.216	.115	.072
4	1.92	1.06	.711	.484	.297	.207
5	2.67	1.61	1.15	.831	.554	.412
6	3.45	2.20	1.64	1.24	.872	.676
7	4.25	2.83	2.17	1.69	1.24	.989
8	5.07	3.49	2.73	2.18	1.65	1.34
9	5.90	4.17	3.33	2.70	2.09	1.73
10	6.74	4.87	3.94	3.25	2.56	2.16
11	7.58	5.58	4.57	3.82	3.05	2.60
12	8.44	6.30	5.23	4.40	3.57	3.07
13	9.30	7.04	5.89	5.01	4.11	3.57
14	10.2	7.79	6.57	5.63	4.66	4.07
15	11.0	8.55	7.26	6.26	5.23	4.60
16	11.9	9.31	7.96	6.91	5.81	5.14
17	12.8	10.1	8.67	7.56	6.41	5.70
18	13.7	10.9	9.39	8.23	7.01	6.26
19	14.6	11.7	10.1	8.91	7.63	6.84
20	15.5	12.4	10.9	9.59	8.26	7.43
21	16.3	13.2	11.6	10.3	8.90	8.03
22	17.2	14.0	12.3	11.0	9.54	8.64
23	18.1	14.8	13.1	11.7	10.2	9.26
24	19.0	15.7	13.8	12.4	10.9	9.89
25	19.9	16.5	14.6	13.1	11.5	10.5
26	20.8	17.3	15.4	13.8	12.2	11.2
27	21.7	18.1	16.2	14.6	12.9	11.8
28	22.7	18.9	16.9	15.3	13.6	12.5
29	23.6	19.8	17.7	16.0	14.3	13.1
30	24.5	20.6	18.5	16.8	15.0	13.8
40	33.7	29.1	26.5	24.4	22.2	20.7
50	42.9	37.7	34.8	32.4	29.7	28.0
60	52.3	46.5	43.2	40.5	37.5	35.5
70	61.7	55.3	51.7	48.8	45.4	43.3
80	71.1	64.3	60.4	57.2	53.5	51.2
90	80.6	73.3	69.1	65.6	61.8	59.2
100	90.1	82.4	77.9	74.2	70.1	67.3

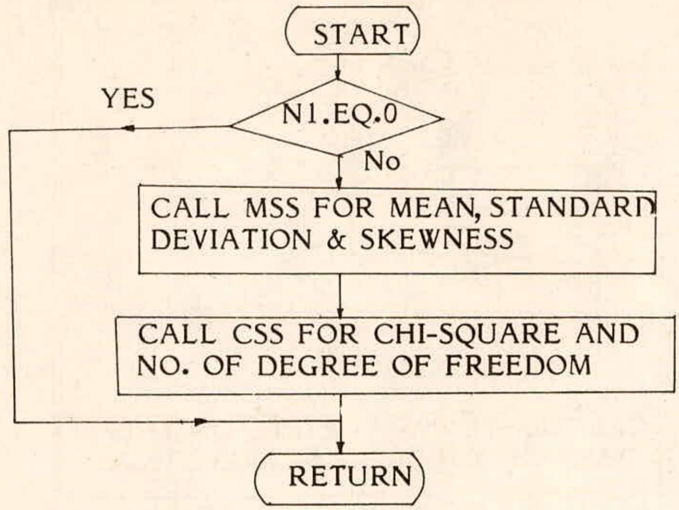
APPENDIX I

FLOW CHART FOR MAIN PROGRAMME

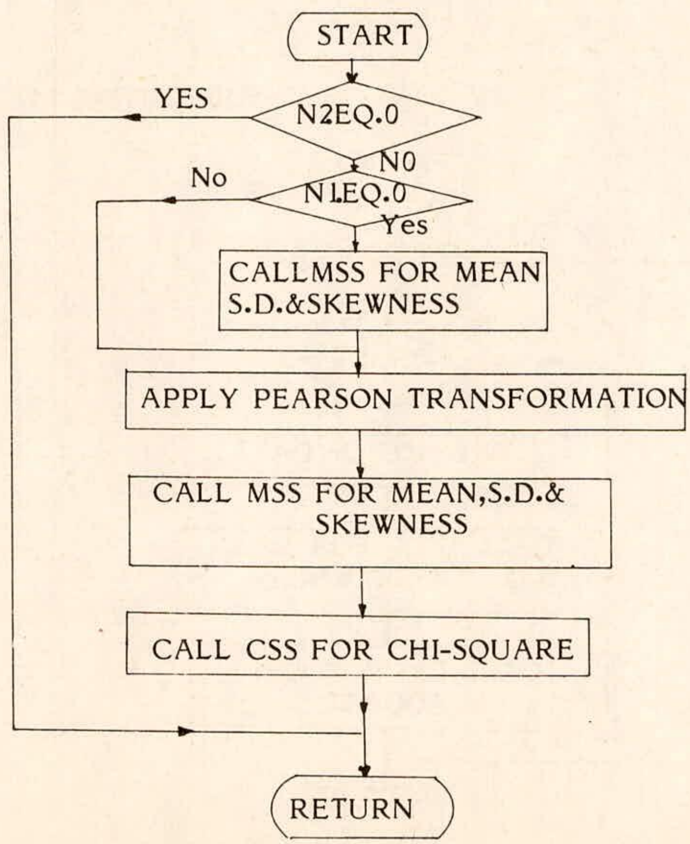



A
SORT OUT THE DATA FOR iTH SEASON, STORE IT . PRINT IT.
CALL SORTX TO ARRANGE THE DATA IN DESCENDING ORDER.PRINT IT
CALL NORMAL FOR NORMAL DISTRIBUTION
CALL PT3 FOR INVERSE PEARSON TRANSFORMATION
CALLLNC FOR LOG NORMAL DISTRIBUTION (PARAMETERS ON THE BASIS OF THEORETICAL RELATIONSHIPS)
CALL LN FOR LOG TRANSFORMATION
CALL LP3 FOR LOG PEARSON TRANSFORMATION
CALL SQRTT FOR SQUARE ROOT TRANSFORMATION
NO
IF. J.C T.N.
Yes
STOP

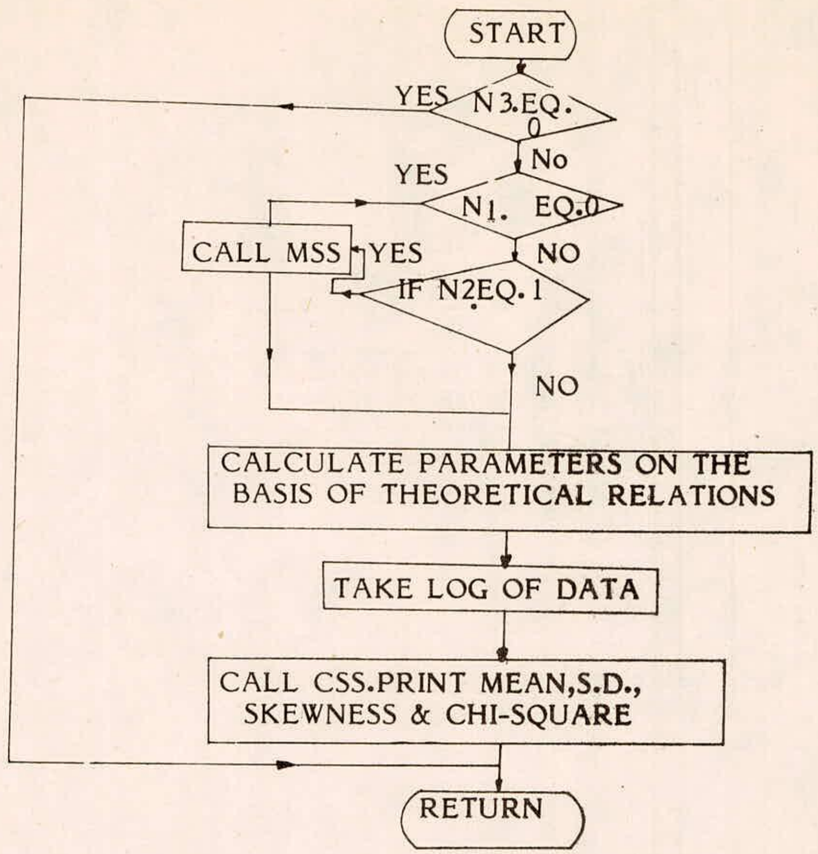
FLOW CHART FOR SUBROUTINE NORMAL



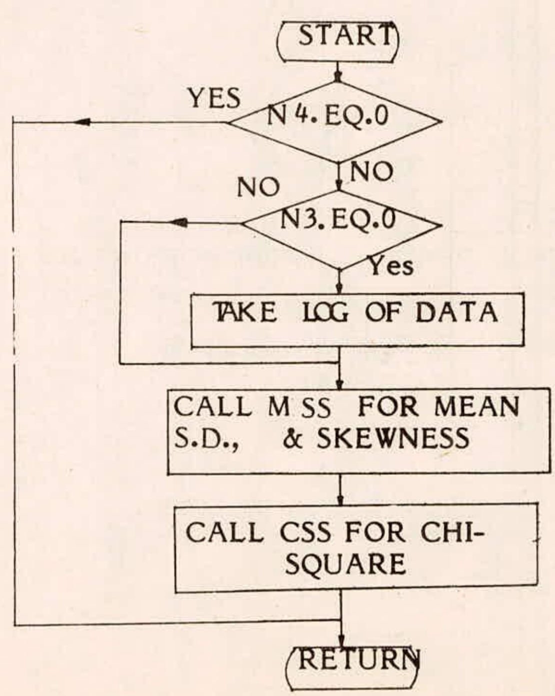
FLOW CHART FOR SUBROUTINE PT3



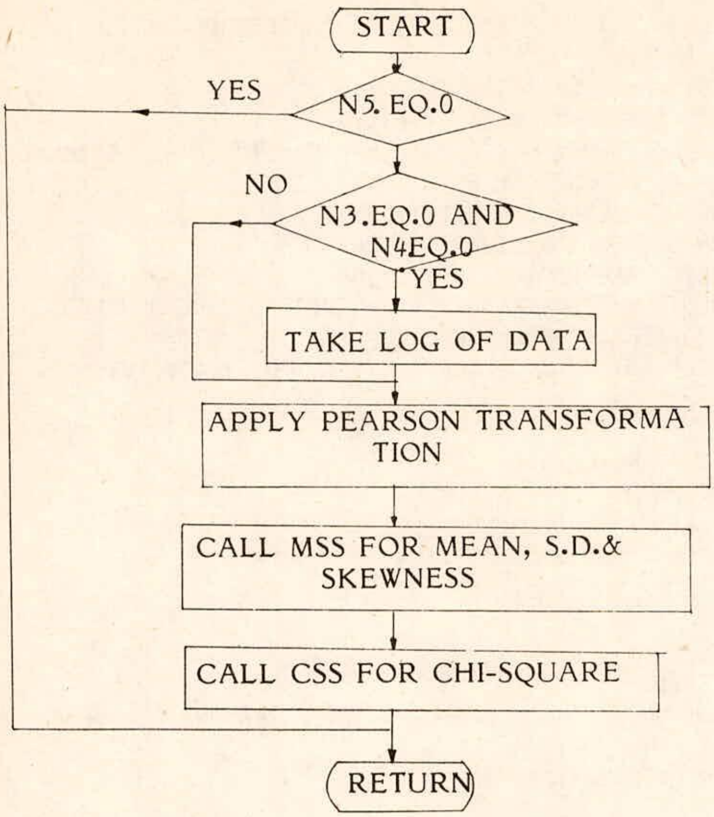
FLOW CHART FOR SUBROUTINE LNC



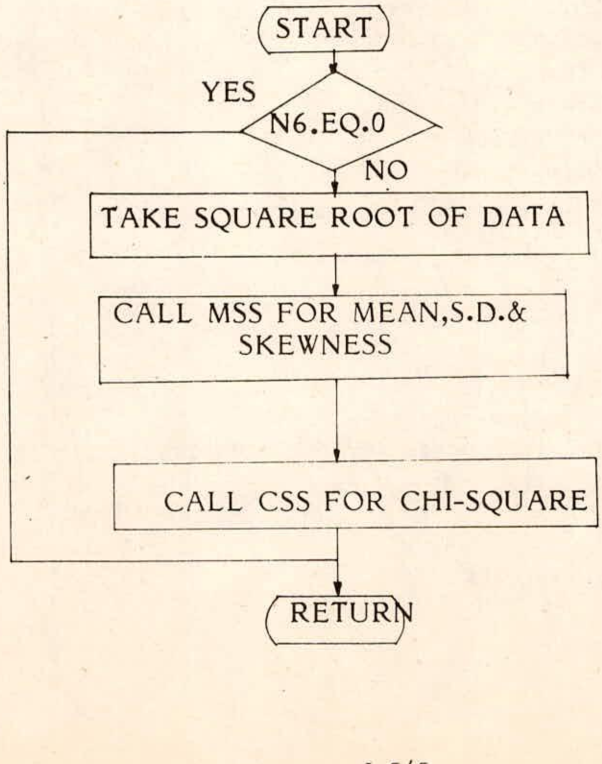
FLOW CHART FOR SUBROUTINE LN



FLOW CHART FOR SUBROUTINE LP3



FLOW CHART FOR SUBROUTINE SQRTT



APPENDIX II

SOURCE PROGRAMME

```
C       PROGRAM TO TEST VARIOUS NORMALIZATION PROCEDURES ON THE
C       BASIS OF CHI SQUARE
C       N1 STANDS FOR NORMAL DISTRIBUTION
C       N2 STANDS FOR PEARSON TYPE 3 DISTRIBUTION
C       N3 STANDS FOR LOGNORMAL DISTRIBUTION PARAMETERS ARE ESTIMATED
C       ON THE BASIS OF THEORETICAL RELATIONS
C       N4 STANDS FOR LOGNORMAL DISTRIBUTION
C       N5 STANDS FOR LOGPEARSON DISTRIBUTION
C       N6 STANDS FOR SQUARE ROOT DISTRIBUTION
C       IF ANY OF THE TRANSFORMATION IS NOT REQUIRED GIVE 0 CORROS
C       PONDING TO THAT TRANSFORMATION
        DIMENSION TITLE(80),KX(500),THF(100),FL(100),T(100),CX(500)
        1,CX1(500),CX2(500)
        COMMON/BL1/L
        COMMON/BK1/N1,N2,N3,N4,N5,N6
        COMMON/BL2/NCLAS
        REAL KX
        OPEN(UNIT=1,FILE='GOEL.DAT',STATUS='OLD')
        OPEN(UNIT=2,FILE='GOEL.OUT',STATUS='NEW')
 80     FORMAT(80A1)
 81     FORMAT(1X,80A1)
 14     FORMAT(1X,12F7.1)
1001    FORMAT(' TOTAL NO. OF VALUES=',I6/' NO.OF SEASONS PER YEAR=',I6
        1/' NCLAS=',I6)
1002    FORMAT(11H INPUT DATA/12(1H*)/)
        READ(1,80) TITLE
        WRITE(2,81) TITLE
118     READ(1,*) N,NS,NCLAS
C       N IS THE TOTAL NO. OF OBSERVATIONS
C       NS IS THE NO. OF SEASONS PER YEAR
C       NCLAS IS THE NO. OF CLASSES REQUIRED FOR CHI
C       SQUARE CALCULATION; THIS DEPENDS UPON THE NO.
C       OF VALUES PER YEAR
        READ(1,*)N1,N2,N3,N4,N5,N6
        WRITE(2,1001) N,NS,NCLAS
        READ(1,*) (KX(I),I=1,N)
        WRITE(2,1002)
        WRITE(2,14) (KX(I),I=1,N)
        NPT=NCLAS+1
        L=N/NS
        FL(1)=0.0
        FL(NPT)=0.9999
        DO 5 I=2,NCLAS
 5      FL(I)=FLOAT(I-1)/FLOAT(NCLAS)
C       CALCULATE STANDARISED VARIATES CORRESPONDING TO LIMITS OF
C       SELECTED CLASS INTERVALS
        DO 555 I=1,NPT
        FFL=FL(I)
        CALL NDTRI(FFL,AX,C,IER)
 555    T(I)=AX
```

```
C       CALCULATE THEORETICAL FREQUENCY FOR EACH CLASS
        DO 68 I=1,NCLAS
68      THF(I)=FLOAT(L)/FLOAT(NCLAS)
        J=1
 100    DO 15 I=1,L
        II=(I-1)*NS+J
        CX(I)=KX(II)
        CX2(I)=CX(I)
 15     CX1(I)=CX(I)
        WRITE(2,40)J
40      FORMAT(20X,'ANALYSIS FOR SEASON',4X,I5/20X,19(1H*)/)
        WRITE(2,10)(CX(I),I=1,L)
 10     FORMAT(3X,10F7.1)
        CALL SORTX(L,CX)
        WRITE(2,20)
20      FORMAT(3X,'SHORTED RECORDED DATA')
        WRITE(2,10)(CX(I),I=1,L)
        CALL NORMAL(CX,L,THF,T,AMEAN,STDEV,SKEW)
        CALL PT3(CX,L,THF,T,AMEAN,STDEV,SKEW)
        CALL LNC(CX1,L,THF,T,AMEAN,STDEV,SKEW)
        CALL LN(CX1,L,THF,T,AMEAN,STDEV,SKEW)
        CALL LP3(CX1,L,THF,T,AMEAN,STDEV,SKEW)
        CALL SQRTT(CX2,L,THF,T,AMEAN,STDEV,SKEW)
        J=J+1
        IF(J.GT.NS)GO TO 200
        GO TO100
 200    STOP
        END
C       NORMAL DISTRIBUTION
        SUBROUTINE NORMAL(CX,L,THF,T,AMEAN,STDEV,SKEW)
        DIMENSION CX(500),THF(100),T(100)
        COMMON/BK1/N1,N2,N3,N4,N5,N6
        IF(N1)500,100,500
500     WRITE(2,1003)
1003    FORMAT(33H ANALYSIS FOR NORMAL DISTRIBUTION)
        CALL MSS(CX,AMEAN,STDEV,SKEW)
        CALL CSS(CX,AMEAN,STDEV,SKEW,THF,T)
100     RETURN
        END
C       PEARSON TYPE 3 DISTRIBUTION BY BEARD NORMALIZATION
        SUBROUTINE PT3(CX,L,THF,T,AMEAN,STDEV,SKEW)
        COMMON/BK1/N1,N2,N3,N4,N5,N6
        DIMENSION CX(500),THF(100),T(100)
        IF(N2.EQ.0)GOTO 100
        IF(N1.EQ.0)GO TO 200
        GO TO 300
200     CALL MSS(CX,AMEAN,STDEV,SKEW)
300     WRITE(2,25)
25      FORMAT(1X,'PEARSON TYPE 3 DISTRIBUTIN')
        DO 725 J=1,L
```

```
      CX(J)=(CX(J)-AMEAN)/STDEV
      CX(J)=(SKEW*CX(J))/2.0+1.0
      IF(CX(J))507,507,508
 507  CX(J)=(-1.)*(ABS(CX(J)))**(1./3.)
      GO TO 509
 508  CX(J)=CX(J)**(1./3.)
 509  CX(J)=(6./SKEW)*(CX(J)-1.)+(SKEW)/6.
 725  CONTINUE
      CALL MSS(CX,AMEAN,STDEV,SKEW)
      CALL CSS(CX,AMEAN,STDEV,SKEW,THF,T)
 100  RETURN
      END
C     LOGNORMAL DISTRIBUTION PARAMETERS ESTIMATED BY CHOW'S METHOD
      SUBROUTINE LNC(CX,L,THF,T,AMEAN,STDEV,SKEW)
      COMMON/BK1/N1,N2,N3,N4,N5,N6
      DIMENSION CX(500),THF(100),T(100)
      IF(N3.EQ.0)GO TO 100
      IF(N1.EQ.0)GO TO 300
      IF(N2.EQ.1)GO TO 300
      GOTO 500
 300  CALL MSS(CX,AMEAN,STDEV,SKEW)
 500  WRITE(2,25)
 25   FORMAT(1X,'CHOW METHOD FOR LOGNORMAL DISTRIBUTION')
      VAR=(STDEV/AMEAN)**2+1.0
      AMEAN=ALOG(AMEAN)-0.5*ALOG(VAR)
      STDEV=ALOG(VAR)
      STDEV=SQRT(STDEV)
      DO 60 I=1,L
      IF(CX(I).EQ.0.0)CX(I)=1.
 60   CX(I)=ALOG(CX(I))
      CALL MSS(CX,AMEAN1,STDEV1,SKEW1)
      CALL CSS(CX,AMEAN,STDEV,SKEW1,THF,T)
 100  RETURN
      END
C     LOG TRANSFORMATION
      SUBROUTINE LN(CX,L,THF,T,AMEAN,STDEV,SKEW)
      COMMON/BK1/N1,N2,N3,N4,N5,N6
      DIMENSION CX(500),THF(100),T(100)
      IF(N4.EQ.0)GO TO 100
      IF(N3.EQ.0)GO TO 200
      GO TO 500
 200  DO 20 I=1,L
      IF(CX(I).EQ.0.)CX(I)=1.
 20   CX(I)=ALOG(CX(I))
 500  WRITE(2,1012)
 1012 FORMAT(1X,19H LOG TRANSFORMATION)
      CALL MSS(CX,AMEAN,STDEV,SKEW)
      CALL CSS(CX,AMEAN,STDEV,SKEW,THF,T)
 100  RETURN
      END
```

```
C       LOG PEARSON TYPE 3 DISTRIBUTION
        SUBROUTINE LP3(CX,L,THF,T,AMEAN,STDEV,SKEW)
        COMMON/BK1/N1,N2,N3,N4,N5,N6
        DIMENSION CX(500),THF(100),T(100)
        IF(N5.EQ.0)GO TO 100
        IF(N3.EQ.0.AND.N4.EQ.0)GOTO 200
        GO TO 500
200     DO 20 I=1,L
        IF(CX(I).EQ.0.)CX(I)=1.0
20      CX(I)=ALOG(CX(I))
        CALL MSS(CX,AMEAN,STDEV,SKEW)
500     DO 728 J=1,L
       CX(J)=(CX(J)-AMEAN)/STDEV
        CX(J)=(SKEW*CX(J))/2.0+1.0
        IF(CX(J))607,607,608
607     CX(J)=(-1.)*(ABS(CX(J)))**(1./3.)
        GO TO 609
608     CX(J)=CX(J)**(1./3.)
609     CX(J)=(6./SKEW)*(CX(J)-1.)+(SKEW)/6.
728     CONTINUE
        WRITE(2,1023)
 1023   FORMAT(1X,'LOG PEARSON TYPE 3 DISTRIBUTION')
        CALL MSS(CX,AMEAN,STDEV,SKEW)
        CALL CSS(CX,AMEAN,STDEV,SKEW,THF,T)
 100    RETURN
        END
C       SQUARE ROOT TRANSFORMATION
        SUBROUTINE SQRTT(CX,L,THF,T,AMEAN,STDEV,SKEW)
        COMMON/BK1/N1,N2,N3,N4,N5,N6
        DIMENSION CX(500),THF(100),T(100)
        IF(N6)500,100,500
500     DO 123 I=1,L
        IF(CX(I).LE.0.)CX(I)=0.1
123     CX(I)=SQRT(CX(I))
        WRITE(2,1020)
1020    FORMAT(26H SQUAREROOT TRANSFORMATION)
        CALL MSS(CX,AMEAN,STDEV,SKEW)
        CALL CSS(CX,AMEAN,STDEV,SKEW,THF,T)
100     RETURN
        END
C       CALCULATE AND TYPE SEASONAL STATISTICAL PARAMETERS
        SUBROUTINE MSS(X,AMEAN,STDEV,SKEW)
        DIMENSION X(500)
        COMMON/BL1/L
        SUM=0.
        DO 2 I=1,L
 2      SUM=SUM+X(I)
        AMEAN=SUM/L
        SUM1=0.
        SUM2=0.
```

```
      DO 4 I=1,L
      SUM1=SUM1+(X(I)-AMEAN)**2
4     SUM2=SUM2+(X(I)-AMEAN)**3
      STDEV=SQRT(SUM1/(L-1))
      SKEW=(L*SUM2)/((L-1)*(L-2)*STDEV*STDEV*STDEV)
      RETURN
      END
C     CALCULATE CHI SQUARE AND NO. OF DEGREE OF FREEDOM
      SUBROUTINE CSS(X,AMEAN,STDEV,SKEW,THF,T)
      COMMON/BL2/NCLAS
      COMMON/BL1/L
      DIMENSION X(500),THF(100),T(100),FREQ(100)
      DO 78 I=1,NCLAS
78    FREQ(I)=0.
      DO 59 I=1,L
      AMINP=AMEAN+STDEV*T(2)
      M=1
61    IF(X(I)-AMINP)62,62,63
62    FREQ(M)=FREQ(M)+1.0
      GO TO 59
63    IF(M-NCLAS)501,62,501
501   M=M+1
      AMINP=AMEAN+STDEV*T(M+1)
      GO TO 61
59    CONTINUE
C     CALCULATE CHI-SQUARE STATISTIC AND NO. OF DEGREES OF FREEDOM
      CHIS=0.
      DO 69 I=1,NCLAS
69    CHIS=CHIS+((ABS(FREQ(I)-THF(I))))**2/THF(I)
      NDF=NCLAS-3
      WRITE(2,1004)AMEAN,STDEV,SKEW,CHIS,NDF
 1004 FORMAT(7X,'MEAN',2X,'STD. DEVIATION',2X,'COEF.OF SKEW',2X
     1,'CHI SQUARE',2X,'DEGREE OF FREEDOM'/5X,F9.3,2X,F9.3
     2,6X,F8.3,5X,F8.4,4X,I5)
      RETURN
      END
C
C.................................................................
C
C     SUBROUTINE NDTRI
C
C     PURPOSE
C       COMPUTES X=P**(-1)(Y), THE ARGUMENT X SUCH THAT Y=P(X)
C       =THE PROBABILITY THAT THE RANDOM VARIABLE U,DISTRIBUTED
C       NORMALLY(0,1), IS LESS THAN OR EQUAL TO X. F(X), THE
C       ORDINATE OF THE NORMAL DENSITY, AT X, IS ALSO COMPUTED.
C
C     USAGE
C       CALL NDTRI(P,X,C,IER)
C
```

```
C       DESCRIPTION OF PARAMETERS
C         P    -INPUT PROBABILITY
C         X    -OUTPUT ARGUMENT SUCH THAT P=Y=THE PROBABILITY THAT
C                THE RANDOM VARIABLE IS LESS THAN OR EQUAL TO X
C         C    -OUTPUT DENSITY,F(X)
C         IER -OUTPUT ERROR CODE
C               =-1  IF P IS NOT IN THE INTERVAL (0,1),INCLUSIVE
C              X=C=.99999E+37 IN THIS CASE
C               =C  IF THERE IS NO ERROR
C                      SEE REMARKS BELOW
C       REMARKS
C         MAXIMUM ERROR IS 0.00045
C         IF P=0,X IS SET TO -(10)**74.D IS SET TO C
C         IF P=1,X IS SET TO  (10)**74.D IS SET TO C
C       SUBROUTINES AND SUBPROGRAMS REQUIRED
C         NONE
C       METHOD
C          BASED ON APPROXIMATIONS IN C.HASTINGS,''APPROXIMATIONS
C          FOR DIGITAL COMPUTERS'',PRICETON UNIV.PRESS,PRINCETON,
C          N.J.,1955.SEE EQUATION 26.2.23,HAND BOOK OF MATHEMAICAL
C          FUNCTIONS,ABRAMOWITZ AND STEGUN,DOVER PUBLICATIONS,INC.,
C          NEW YORK.
        SUBROUTINE NDTRI(P,X,D,IE)
        IE=C
        X=.99999E+37
        D=X
        IF(P)1,4,2
1       IE=-1
        GO TO 12
2       IF(P-1.0)7,5,1
4       X=-0.999999E+37
5       D=0.0
        GO TO 12
7       D=P
        IF(D-0.5)9,9,8
8       D=1.0-D
9       T2=ALOG(1.0/(D*D))
        T=SQRT(T2)
        X=T-(2.515517+0.802853*T+0.010328*T2)/(1.0+1.432788*T+
        10.189269*T2+0.001308*T*T2)
        IF(P-0.5)10,10,11
10      X=-X
11      D=0.3989423*EXP(-X*X/2.0)
12      RETURN
        END
        SUBROUTINE SORTX (N,X)
C   SORTS IN DECREASING ORDER, X(I)=LARGEST
        DIMENSION X(100)
        K=N-1
        DO 2 L=1,K
```

```
M=N-L
DO 2 J=1,M
IF (X(J)-X(J+1)) 1,1,2
XT=X(J)
X(J)=X(J+1)
X(J+1)=XT
CONTINUE
RETURN
END
```

APPENDIX III

INPUT SPECIFICATIONS

Input cards/lines have been divided in job cards and data cards.

(a) Job Cards

Card	Variable	Description	Format
FIRST	TITLE	Title of the problem	A
SECOND	N	Total number of observations	FREE
	NS	Number of seasons in a year	
	NCLAS	Number of classes for the calculation of chi-square	
THIRD	N1	Option code for normal distribution	FREE
	N2	Option code for inverse Pearson type III transformation	
	N3	Option code for log normal distribution(parameters on the basis of theoretical relationship)	
	N4	Option code for log transformation	
	N5	Option code for inverse log Pearson type III transformation	
	N6	Option code for square root transformation	

If any of the transformation is not required, 0 is given corresponding to its option code, otherwise 1 is given.

(b) Data Cards

Observations are punched till end in the free format.

APPENDIX IV

OUTPUT DESCRIPTION

The following statistics are given for the desired transformations for all the seasons:

	Statistics	Format
a.	Mean of the transformed series	F8.3
b.	Standard deviation of the transformed series	F8.3
c.	Coefficient of skewness of the transformed series	F8.3
d.	Chi-square for the transformation	F8.4
e.	Number of degrees of freedom	I5

APPENDIX V

TEST DATA

The programme was run on 32 years annual peak discharge data (1951-1982) of river Narmada at Mortakka for different options. The values of N,NS, NCLAS, N1,N2,N3,N4,N5 and N6 are given and described below:

N = 32

NS = 1(Since data is for annual peaks)

NCLAS = 6 (for NCLAS = 6, the theoretical frequency of each class will be 5.33)

N1 = 1

N2 = 1

N3 = 1

N4 = 1

N5 = 1

N6 = 1

(The best fitting distribution is required, as such values of N1,N2,N3,N4,N5 and N6 have been given as 1)

APPENDIX VI

TEST INPUT

```
ANNUAL PEAK DISCHARGES FOR NARMADA AT MORTAKKA(1951-82)
32,1,6
1,1,1,1,1,1
11127   13631   19521   33915   20746   11982   25023   13005
30372   20540   55323   31604   16135   23438   18591   11338
19690   31604   27935   41691   18101   47851   54063   36562
33278   17713   24354   29564   26232   22751   25662   16602
```

APPENDIX VII
TEST OUTPUT

```
 ANNUAL PEAK DISCHARGES FOR NARMADA AT MORTAKKA(1951-82)
TOTAL NO. OF VALUES=    32
NO.OF SEASONS PER YEAR=     1
NCLAS=    6
INPUT DATA
**********

 11127. 13631. 19521. 33915. 20746. 11982. 25023. 13005. 30372. 20540.
 55323. 31604. 16135. 23438. 18591. 11338. 19690. 31604. 27935. 41691.
 18101. 47851. 54063. 36562. 33278. 17713. 24354. 29564. 26232. 22751.
 25662. 16602.
                  ANALYSIS FOR SEASON        1
                  ******************

   11127. 13631. 19521. 33915. 20746. 11982. 25023. 13005. 30372. 20540.
   55323. 31604. 16135. 23438. 18591. 11338. 19690. 31604. 27935. 41691.
   18101. 47851. 54063. 36562. 33278. 17713. 24354. 29564. 26232. 22751.
   25662. 16602.
  SHORTED RECORDED DATA
   55323. 54063. 47851. 41691. 36562. 33915. 33278. 31604. 31604. 30372.
   29564. 27935. 26232. 25662. 25023. 24354. 23438. 22751. 20746. 20540.
   19690. 19521. 18591. 18101. 17713. 16602. 16135. 13631. 13005. 11982.
   11338. 11127.
ANALYSIS FOR NORMAL DISTRIBUTION
      MEAN  STD. DEVIATION  COEF.OF SKEW  CHI SQUARE  DEGREE OF FREEDOM
    25935.750  11615.848        1.044        3.2500          3
PEARSON TYPE 3 DISTRIBUTIN
      MEAN  STD. DEVIATION  COEF.OF SKEW  CHI SQUARE  DEGREE OF FREEDOM
      0.014      0.967          0.303        1.0000          3
CHOW METHOD FOR LOGNORMAL DISTRIBUTION
      MEAN  STD. DEVIATION  COEF.OF SKEW  CHI SQUARE  DEGREE OF FREEDOM
     10.072      0.428          0.105        1.0000          3
 LOG TRANSFORMATION
      MEAN  STD. DEVIATION  COEF.OF SKEW  CHI SQUARE  DEGREE OF FREEDOM
     10.072      0.433          0.105        0.2500          3
LOG PEARSON TYPE 3 DISTRIBUTION
      MEAN  STD. DEVIATION  COEF.OF SKEW  CHI SQUARE  DEGREE OF FREEDOM
      0.001      1.000          0.030        0.2500          3
SQUAREROOT TRANSFORMATION
      MEAN  STD. DEVIATION  COEF.OF SKEW  CHI SQUARE  DEGREE OF FREEDOM
     157.411     34.568         0.577        0.6250          3
```