Honeywell Bull

APPLICATIONS LIBRARY GUIDE
VOLUME II – STATISTICS

SERIES 6000/600

TIME-SHARING

APPLICATIONS

Ref. : 19.53.107 A

Honeywell Bull

SERIES 6000/600

TIME-SHARING APPLICATIONS LIBRARY GUIDE VOLUME II—STATISTICS

SUBJECT:

Descriptions, Sample Problems, and Solutions of Statistical Time-Sharing Programs.

SPECIAL INSTRUCTIONS:

This manual, in conjunction with *Series 6000/600 Time-Sharing Applications Library Guide Volume I—Mathematics* (Order No. DA43) and *Volume III—Industry* (Order No. DA45), supersedes *GE-600 Line Time-Sharing Library Programs*, Document No. CPB-1694. The contents of CPB-1694 have been divided into three volumes. Sixty-five programs have been added to the new set of manuals. They are listed on the back of this page.

DATE:

June, 1971

ORDER NUMBER: DA44

Rev. 0

Ref. : 19.53.107 A

ADDITIONAL PROGRAMS INCORPORATED IN THE MAY 1971 EDITIONS

Series 6000/600 Time-Sharing Applications Library Guide Volume II - Statistics, Order No. DA44

ANOVA
ANVA3
FACTAN
FLAT
FORIR
LINREG
MREG1
POISON
POLFT
PROBC
RANDX
RNDNRM
STAT01
STAT02
STAT04
STAT05
STAT06
STAT08
STAT09
STAT11
STAT12
STAT13
STAT14
STAT15
STAT16
STAT18
STAT33
UNIFM
URAN
XNOR1
XNORM

Series 6000/600 Time-Sharing Applications Library Guide Volume I - Mathematics, Order No. DA43

4SQRS
ARCTAN
COMP1
EIGSR
EUALG
FRESNL
GAHER
GALA
GAUSSN
GCDN
JACELF
LINSR
ORTHP
POLYC
QUADEQ
ROMBINT
STIRLING
ZCOP
ZCOP2
ZORP2

Series 6000/600 Time-Sharing Applications Library Guide Volume III - Industry, Order No. DA45

BUSINESS AND FINANCE
- DEPREC
- RETURN

MANAGEMENT SCIENCE AND OPTIMIZATION
- GASPIIA
- PERT
- SMOOTH
- TCAST

ENGINEERING
- ACNET
- NLNET
- PVT

GEOMETRIC AND PLOTTING
- PLOT

EDUCATION AND TUTORIAL
- DRIVES
- EXPERN
- PREPRS

UTILITY AND MISCELLANEOUS
- CONVRT

PREFACE

This manual describes and tells how to use the statistical time-sharing programs available with the Series 6000 and 600 information processing systems. The programs are listed alphabetically in the Table of Contents.

The writeup about each program includes the purpose of the program; the language in which the program is written; the method of approach, if applicable; instructions for its use; restrictions of the program, if any; and sample problems and their solution. In the sample solutions, all information that the user types is underlined.

The instructions in this manual assume that the programs are available in the user master catalog LIBRARY, and are accessible with READ or EXECUTE permission.

In the sample solution printouts the programs had already been accessed using the GET command, and/or copied onto the current file using the OLD or LIB command.

Time-sharing programs for mathematics and other classifications are also available. Individual manuals are published for these categories as follows:

> Series 6000/600 Time-Sharing Applications Library Guide Volume I - Mathematics, Order No. DA43
>
> Series 6000/600 Time-Sharing Applications Library Guide Volume III - Industry, Order No. DA45

The Industry manual is organized into sections by type as follows:

- BF - Business and Finance
- MS - Management Science and Optimization
- EN - Engineering

PREFACE (Cont.)

GP - Geometric and Plotting

ED - Education and Tutorial

DE - Demonstration

UM - Utility and Miscellaneous

Each section is paginated with the 2-letter identifier that is shown above.

A complete listing of the programs in the library is available by listing the LIBRARY program CATALOG. A copy of this program follows the Table of Contents for your information.

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont.)

TABLE OF CONTENTS (Cont.)

TABLE OF CONTENTS (Cont.)

CATALØG ØF SERIES 6000/600 T-S LIBRARY PRØGRAMS

```
FØRMAT INDICATØR:
FIRST LETTER              FØLLØWING LETTERS
------------              ----------------
F  FØRTRAN-SØURCE         P     (ØR BLANK) PRØGRAM
Ø  FØRTRAN-ØBJECT         S     SUBRØUTINE(S)
C  CARDIN                 F     FUNCTIØN(S)
B  BASIC-SØURCE           P-S   PRØGRAM WITH EXTRACTABLE SUBRØUTINE
E  EDITØR(ASCII)
```

```
SUBJECTS                                          DØCUMENTATIØN MANUAL
--------                                          --------------------

MATHEMATICS (MA)                  ..............ØRDER # DA43
     INTEGRATIØN
     DIFFERENTIATIØN, DIFFERENTIAL EQ.
     INTERPØLATIØN
     PØLYNØMIALS
     LINEAR EQUATIØNS
     MATRICES
     NØN-LINEAR EQUATIØNS
     SPECIAL FUNCTIØN EVALUATIØN
     LØGIC AND NUMBER THEØRY
STATISTICS (ST)                   ..............ØRDER # DA44
     CURVE FITTING AND REGRESSIØN
     ANALYSIS ØF VARIANCE
     PRØBABILITY DISTRIBUTIØNS
     CØNFIDENCE LIMITS
     HYPØTHESIS TESTING
     DESCRIPTIVE STATISTICS
     RANDØM NUMBER GENERATIØN
     MISCELLANEØUS STATISTICS
BUSINESS AND FINANCE (BF)         ..............ØRDER # DA45
MANAGEMENT SCIENCE AND ØPTIMIZATIØN (MS)
     LINEAR PRØGRAMMING
     INTEGER PRØGRAMING
     NØN-LINEAR ØPTIMIZATIØN
     NETWØRK ANALYSIS
     FØRECASTING
     SIMULATIØN
ENGINEERING (EN)
GEØMETRIC AND PLØTTING (GP)
EDUCATIØN AND TUTØRIAL (ED)
DEMØNSTRATIØN (DE)
UTILITY AND MISCELLANEØUS (UM)
```

THE DØCUMENTATIØN FØR THESE PRØGRAMS IS AVAILABLE IN THREE MANUALS:
SEE ØRDER # DA43 FØR PRØGRAMS IN MATHEMATICS
ØRDER # DA44 FØR PRØGRAMS IN STATISTICS
ØRDER # DA45 FØR PRØGRAMS IN ALL ØTHER CATEGØRIES.

SUBRØUTINES THAT ARE CALLED BY A PRØGRAM AND MUST BE EXECUTED WITH IT ARE LISTED IN BRACKETS AT THE END ØF THE DESCRIPTIØN.

THESE PRØGRAMS HAVE ALL BEEN REVIEWED AND TESTED BUT NØ RESPØNSIBILITY CAN BE ASSUMED.

```
*******************MA--MATHEMATICS**************************************

***INTEGRATION***
CLCINT     FF     INTEGRATION BY SIMPSON'S RULE
FINT       FF     EVALUATE FOURIER INTEGRALS BY FILON'S FORMULA
GAHER      FF     GAUSS-HERMITE QUADRATURE
GALA       FF     GAUSS-LAGUERRE QUADRATURE
GAUSSN     FF     EVALUATE DEFINITE DOUBLE OR TRIPLE INTEGRALS
GAUSSQ     FF     GAUSSIAN QUADRATURE
NCOATES    FP-S   NEWTON-COATES QUADRATURE
NUMINT     B      GAUSSIAN QUADRATURE
ROMBINT    FP-S   ROMBERG INTEGRATION

***DIFFERENTIATION, DIFFERENTIAL EQ.***
AMPBX      FS     ADAMS-MOULTON FOR 1ST-ORDER DIFF. EQNS [RKPBX]
FDRVUL     FF     DIFFERENTIATE TABULATED FUNCTION, UNEQUAL SPACING
HDRVEB     FF     DIFFERENTIATE TABULATED FUNCTION, EQUAL SPACING
RKPBX      FS     RUNGE-KUTTA FOR 1ST-ORDER DIFF. EQNS

***INTERPOLATION***
TNT1       FF     SINGLE LAGRANGIAN INTERPOLATION [TLU1]
INT2       FF     DOUBLE LAGRANGIAN INTERPOLATION [TLU1]
TNT2A      FF     VARIABLE DOUBLE LINEAR INTERPOLATION [TLU1]

***POLYNOMIALS***
BICOF      FS     CALCULATE BINOMIAL COEFFICIENTS
CLPLY      FF     EVALUATE REAL POLY AT REAL ARGUMENT
DVALG      FS     POLYNOMIAL DIVISION
EUALG      FS     G.C.D. OF TWO POLYNOMIALS [DVALG]
MTALG      FS     MULTIPLY POLYNOMIALS
PLMLT      FS     REAL POLY COEFFICIENTS RECONSTRUCTED FROM REAL ROOTS
POLRTS     FP     SOLUTION OF POLY BY BAIRSTOWS METHOD
POLYC      FS     REAL POLY COEFFICIENTS RECONSTRUCTED FROM COMPLEX ROOTS
POLYV      FS     EVALUATE REAL POLY AT COMPLEX ARGUMENT
QUADEQ     B      SOLUTION TO QUADRATIC EQUATIONS
ROOTER     B      SOLUTION OF POLY BY BAIRSTOWS METHOD
ZCOP       FP     ROOTS OF POLYNOMIAL WITH COMPLEX COEFF.
ZCOP2      FS     ROOTS OF POLYNOMIAL WITH COMPLEX COEF. [ZCOP2]
ZORP       FP     ROOTS OF REAL POLY
ZORP2      FS     ROOTS OF REAL POLY

***LINEAR EQUATIONS***
GJSIMEQ    FS     SOLVE LINEAR SYSTEMS BY GAUSS-JORDAN
GSEIDEL    FP-S   SOLVE LINEAR SYSTEMS BY GAUSS-SEIDEL
LINEQ      FS     SOLVE LINEAR SYSTEMS BY GAUSSIAN ELIMINATION
LINSR      FP     SOLVE LINEAR SYSTEMS BY GAUSSIAN ELIMINATION [LINEQ]
SIMEQN     B      SOLVE LINEAR SYSTEMS BY MATRIX INVERSION

***MATRICES***
DETE       FF     EVALUATE DETERMINANT OF REAL MATRIX
DOMEIG     FP-S   DOMINANT EIGENVALUES OF REAL MATRIX
EIG1       FS     EIGENVALUES OF SYM MATRIX BY JACOBI METHOD
EIGSR      FP     EIGENVALUES AND VECTORS OF REAL SYM. MATRIX [EIG1]
MTINV      FS     MATRIX INVERSION BY PIVOTS
MTMPY      FS     MATRIX MULTIPLICATION
SPEIG1     FS     SPECIAL EIGEN PROBLEMS [EIG1]
SYMEIG     FP     EIGENVALUES OF SYM MATRIX BY JACOBI METHOD

***NON-LINEAR EQUATIONS***
BROWN      FS     SOLN OF SIMULTANEOUS SYSTEMS BY BROWN METHOD
SECANT     FS     SOLN OF SIMULTANEOUS SYSTEMS BY SECANT METHOD [MTINV]
SOLN       FF     ZERO OF AN ARBITRARY FUNCTION
ZEROES     B      ZERO,MAX,MIN OF FUNCTION
```

SPECIAL FUNCTIØN EVALUATIØN

ARCTAN	FF	ARCTANGENT IN RADIANS ØF Y/X
BESL	FS	BESSEL FUNCTIØN [GAMF]
CØMP1	FF	EVALUATES REAL HYPERBØLIC TRIG FUNCTIØNS
CØMP2	FS	CØMPLEX MULT. AND DIVISIØN
CØMP3	FS	EVALUATES VARIØUS FUNCTIØNS FØR CØMPLEX ARGUMENT [CØMP2]
ERRF	FF	ERRØR FUNCTIØN
ERRINV	FF	INVERSE ERRØR FUNCTIØN
FRESNL	FS	EVALUATES FRESNAL INTEGRALS
GAMF	FF	GAMMA FUNCTIØN
JACELF	FS	EVALUATES JACØBIAN ELLIPTIC FUNCTIØNS SN, CN, DN
ØRTHP	FF	EVALUATE ØRTHØGØNAL PØLYNØMIALS
STIRLING	FP-S	N FACTØRIAL BY STIRLINGS APPRØXIMATIØN
TMFCEV	B	EVALUATE DAMPED ØR UNDAMPED FØURIER SERIES

LØGIC AND NUMBER THEØRY

4SQRS	B	WRITES INTEGERS AS SUM ØF SQUARES ØF FØUR INTEGERS
GCDN	FS	G.C.D. ØF N INTEGERS

*******************ST--STATISTICS***************************************

CURVE FITTING AND REGRESSIØN

CFIT	FP	LEAST SQRS. PØLY. WITH RESTRAINTS
CURFIT	B	FITS SIX DIFFERENT CURVES BY LEAST SQRS
FØRIR	FP	LEAST SQUARES ESTIMATE ØF FINITE FØURIER SERIES MØDEL
FØURIER	B	CØEFF ØF FØURIER SERIES TØ APPRØX A FUNCTIØN
LINEFIT	FS	LEAST SQRS LINE
LINREG	B	LST.SQRS. BY LINEAR, EXPØNENTIAL, ØR PØWER FUNCTIØN
LSPCFP	FP	LEAST SQRS PØLYNØMIAL FIT
LSQMM	FS	GENERALIZED PØLY FIT BY LEAST SQRS ØR MIN-MAX
MREG1	FP	MULTIPLE LINEAR REGRESSIØN
MULFIT	B	MULTIPLE LINEAR FIT WITH TRANSFØRMATIØNS
ØRPØL	FP	LEAST SQRS FIT WITH ØRTHØGØNAL PØLYS
PØLFIT	B	LEAST SQRS PØLYNØMIAL FIT
PØLFT	FP	LEAST SQRS PØLYNØMIAL FIT
SMLRP	FP	MULTIPLE LINEAR REGRESSIØN
SMLRPØBJ	Ø	ØBJECT FILE FØR SMLRP

ANALYSIS ØF VARIANCE

ANØVA	FP	ØNE ØR TWØ WAY ANALYSIS ØF VARIANCE
ANVA1	FP	ØNEWAY ANALYSIS ØF VARIANCE
ANVA3	FP	THREE WAY ANALYSIS ØF VARIANCE
ANVA5	FP	MULTIPLE VARIANCE ANALYSIS
KRUWAL	FP	KRUSKAL-WALLIS 2-WAY VARIANCE [XINGAM]
ØNEWAY	B	ØNEWAY ANALYSIS ØF VARIANCE
STAT13	B	ANALYSIS ØF VARIANCE TABLE, 1-WAY RANDØM DESIGN
STAT14	B	ANALYSIS ØF VARIANCE TABLE FØR RANDØMIZED BLØCK DESIGN
STAT15	B	ANALYSIS ØF VARIANCE TABLE FØR SIMPLE LATIN-SQ DESIGN
STAT16	B	ANALYSIS ØF VARIANCE TABLE, GRAECØ-LATIN SQUARE DESIGN
STAT18	B	ANALYSIS ØF VARIANCE TABLE, YØUDEN SQUARE DESIGN
STAT33	B	ANALYSIS ØF VARIANCE TABLE, 1-WAY RANDØM DESIGN

PRØBABILITY DISTRIBUTIØNS

ANPF	FF	NØRMAL PRØBABILITY FUNCTIØN [ERRF]
BETA	FF	BETA DISTRIBUTIØN
BINDIS	B	BINØMIAL PRØBABILITIES
EXPLIM	B	EXPØNENTIAL DISTRIBUTIØNS
PØISØN	FF	PØISSØN DISTRIBUTIØN FUNCTIØN
PRØBC	FP	PRØBABLITIES ØF CØMBINATIØNS ØF RANDØM VARIABLES
PRØVAR	B	NØRMAL AND T-DISTRIBUTIØN
TDIST	FF	T-DISTRIBUTIØN [BETA]
XINGAM	FF	INCØMPLETE GAMA FUNCTIØN

CØNFIDENCE LIMITS

BAYES	B	DIFFERENCE ØF MEANS IN NØN-EQUAL VARIANCE
BICØNF	B	CØNF. LIMITS FØR PØPULATIØN PRØPØRTIØN (BINØMIAL)
BINØM	FP	BINØMIAL PRØBABILITIES AND CØNFIDENCE BANDS
CØLINR	B	CØNFIDENCE LIMITS ØN LINEAR REGRESSIØNS
CØNBIN	B	CØNF. LIMITS FØR PØPULATIØN PRØPØRTIØN (NØRMAL)
CØNDIF	B	DIFFERENCE ØF MEANS IN EQUAL VARIANCE
CØNLIM	B	CØNF. LIMITS FØR A SAMPLE MEAN
STAT05	B	CØNFIDENCE INTERVAL FØR MEAN BY SIGN TEST
STAT06	B	CØNFIDENCE LIMITS,WILCØXØN SIGNED RANK SUM TEST

HYPØTHESIS TESTING

BITEST	B	TEST ØF BINØMIAL PRØPØRTIØNS
CHISQR	FS	CHI-SQUARE CALCULATIØNS
CØRREL	FP	CONTINGENCY CØEFFICIENT [XINGAM]
CØRRL2	FP	CØRRELATIØN CØEFFICIENT [TDIST;BETA]
KØKØ	FP	KØLMØGØRØV-SMIRNØV TWØ SAMPLE TEST [XINGAM]
SEVPRØ	B	CHI-SQUARE
STAT01	B	MEAN,STD ØF MEAN, ... , T-RATIØ,2 GRØUPS, PAIRED
STAT02	B	MEANS, VARIANCES, AND T-RATIØ 2 GRØUPS, UNPAIRED DATA
STAT04	B	CHI-SQUARE AND PRØBABILITIES, 2X2 TABLES
STAT08	B	CØMPARES TWØ GRØUPS ØF DATA USING THE MEDIAN TEST
STAT09	B	CØMPARE 2 DATA GRØUPS,MANN-WHITNEY 2-SAMPLE RANK TEST
STAT11	B	SPEARMAN RANK CØRRELATIØN CØEF. FØR 2 SERIES ØF DATA
STAT12	B	CØMPUTES CØRRELATIØN MATRIX FØR N SERIES ØF DATA
TAU	FP	KENDALL-RANK CØRRELATIØN

DESCRIPTIVE STATISTICS

MANDSD	B	FIND MEAN, VARIANCE, STD
STAT	FP	FIND SEVERAL STATISTICS FØR SAMPLE DATA [ANPF;ERRF]
STATAN	B	FIND VARIØUS STATISTICAL MEASURES
TESTUD	B	SAMPLE STATISTICS
UNISTA	B	DESCRIPTIØN ØF UNI-VARIANT DATA

RANDØM NUMBER GENERATIØN

FLAT	ØF	UNIFØRM RANDØM NUMBER GENERATØR
FLATSØRC	C	CARDIN SØURCE FILE FØR FLAT
RANDX	FF	RANDØM #'S, UNIFØRM DIST. BETWEEN 0 AND 1
RNDNRM	FF	CALCULATES NØRMAL RANDØM NUM. [FLAT]
UNIFM	ØF	UNIFØRM RANDØM NUMBER GENERATØR
UNIFMSØR	C	CARDIN SØURCE FILE FØR UNIFM
URAN	ØF	UNIFØRM RANDØM NUMBER GENERATØR
URANSØRC	C	CARDIN SØURCE FILE FØR URAN
XNØR1	FF	NØRMAL RANDØM NUMBERS, VARIABLE MEAN, STD [RANDX]
XNØRM	FF	NØRMAL RANDØM NUMBERS, MEAN 0, STD 1. [RANDX]

MISCELLANEØUS STATISTICS

FACTAN	FP	FACTØR ANALYSIS
STADES	E	EXPLANATIØN ØF CØLINR,CURFIT,MULFIT,UNISTA

********************BF--BUSINESS AND FINANCE****************************

ANNUIT	B	ANNUITIES,LØANS,MØRTGAGES
BLDGCØST	B	ANALYZE BUILDING CØSTS
DEPREC	B	CALCULATES DEPRECIATIØN BY FØUR METHØDS
SAVING	B	SAVINGS PLAN CALCULATIØNS
RETURN	B	CØMPUTES ANNUAL RETURNS FØR A SECURITY FRØM ANNUAL DATA
TRUINT	B	INTEREST RATE CALCULATIØNS

********************MS--MANAGEMENT SCIENCE AND ØPTIMIZATIØN************

LINEAR PRØGRAMMING

LINPRØ	B	LINEAR PRØGRAMMING
LNPRØG	FP	LINEAR PRØGRAMMING

```
***INTEGER PRØGRAMMING***
INTO1       FP    ZIØNTS' MØDIFICATIØN ØF BALAS' ZERØ-ØNE ALGØRITHM
INTLP       FP    GØMØRY'S PURE AND MIXED INTEGER PRØGRAMMING

***NØN-LINEAR ØPTIMIZATIØN***
DAVIDØN     B     DAVIDØN'S UNCØNSTRAINED ØPTIMIZATIØN
LØGIC3      FP    UNCØNSTRAINED ØPTIMIZATIØN
MAXØPT      FP    UNCØNSTRAINED ØPTIMIZATIØN

***NETWØRK ANALYSIS***
CPM         FP    CRITICAL PATH METHØD
KILTER      FP    'ØUT ØF KILTER' ALGØRITHM (MINIMUM CØST CIRCULATIØN)
MAXFLØW     FP    MAXIMUM FLØW THRU NETWØRK
PERT        B     SIMPLE ANALYSIS ØF A PERT NETWØRK
SHØRTEST    FP    SHØRTEST PATH - MIN SPANNING TREE

***FØRECASTING***
TCAST       FP    TIME SERIES FØRECASTING [TCAST1;TCAST2]
TCAST1      Ø     ØVERLAY MØDULE ØF TCAST
TCAST2      Ø     ØVERLAY MØDULE ØF TCAST
SMØØTH      FS    TRIPLE SMØØTHING ØF A TIME SERIES

***SIMULATIØN***
GASPDATA    E     DATA FILE FØR SAMPLE PRØGRAM GASPSAMP
GASPIIA     FS    'GASP' SIMULATIØN SYSTEM
GASPSAMP    FP    SAMPLE PRØGRAM FØR GASPIIA [GASPIIA;GASPDATA]
********************EN--ENGINEERING**************************************
ACNET       FP    FREQUENCY RESPØNSE ØF A LINEAR CIRCUIT
BEMDES      B     STEEL BEAM SELECTIØN
GCVSIZ      B     GAS CØNTRØL VALVE CØEFF.
LCVSIC      B     LIQUID CØNTRØL VALVE CØEFF.
LPFILT      B     DESIGN LØW PASS FILTERS
NLNET       FP    GENERAL STEADY-STATE CIRCUIT ANALYSIS
ØTTØ        B     ØTTØ CYCLE ØF ENGINE
PVT         FP    FINDS MØLAR VØLUME ØF A GAS GIVEN TEMPERATURE AND PRES.
SCVSIZ      B     STEAM CØNTRØL VALVE CØEFF.
SECAP       B     STEEL SECTIØN CAPACITIES
********************GP--GEØMETRIC AND PLØTTING***************************
CIRCLE      B     DIVIDES A CIRCLE INTØ N EQUAL PARTS
PLØT        FS    PLØTS UP TØ 9 CURVES SIMULTANEØUSLY
PLØTTØ      B     SIMULTANEØUSLY PLØTS 1 TØ 6 FUNCTIØNS
PØLPLØ      FP    PLØTS EQNS IN PØLAR CØØRDINATES
SPHERE      B     SØLVES ANY SPHERICAL TRIANGLE
TRIANG      B     SØLVES FØR ALL PARTS ØF ANY TRIANGLE
TWØPLØ      B     SIMULTANEØUSLY PLØTS 2 FUNCTIØNS
XYPLØT      B     PLØTS SINGLE-VALVED FUNCTIØNS
********************ED--EDUCATIØN AND TUTØRIAL***************************
DRIVES      Ø     DRIVER FØR EXPER, A CØMPUTER ASSISTED INST. LANG.
EXPERN      E     EXPER TUTØRIALS IN EXPER (N=1 TØ 5) [PREPRS;DRIVES]
PREPRS      Ø     PREPRØCESSØR FØR EXPER, A CØMPUTER ASSISTED INST. LANG.
********************UM--UTILITY AND MISCELLANEØUS************************
BLKJAK      B     THE CØMPUTER DEALS BLACKJACK
********************DE--DEMØNSTRATIØN************************************
CATALØG     E     CATALØG ØF SERIES 6000/600 T/S LIBRARY (THIS FILE)
CØNVRT      B     CØNVERTS MEASUREMENTS FRØM ØNE SCALE TØ ANØTHER
DBLSØRT     FS    SØRT TWØ ARRAYS
SGLSØRT     FS    SØRT AN ARRAY
TLU1        FS    TABLE SEARCH
TPLSØRT     FS    SØRT THREE ARRAYS

***END ØF CATALØG***
```

ANOVA

This FORTRAN program performs either a one-way or a two-way analysis of variance.

INSTRUCTIONS

Data for this program can be entered via the teleprinter keyboard or by using a data file. In either case, the free field format is used for input.

The following definitions apply:

1. COLUMNS - The number of columns in the experimental design, a column being a level of a factor in the experiment.
2. ROWS - The number of rows in the design, i.e., the second variable of the design.
3. OBSERVATIONS - The number of replicates for each experimental condition.

The program generates instructions for entering the data: after requesting the number of columns, rows, and observations, the program requests the observed experimental data. This data is entered starting with row 1, column 1 replicates; row 2, column 1 replicates; and so forth. The last set of data is the last row, last column replicates.

As output, the program generates the appropriate Analysis of Variance table. Each source of variations enumerated with the corresponding sum of squares, degrees of freedom, and mean square value. Since the experimental model (fixed, mixed, or random) is not known by the program, the "F" ratios are not calculated. An additional output is the mean value of every row and column.

RESTRICTIONS

1. The number of rows and columns must be greater than or equal to 1 and less than or equal to 10.
2. The number of observations (replicates) must be greater than or equal to 1 and less than or equal to 19. All cells must have the same number of replicates.

SAMPLE PROBLEM

Generate an Analysis of Variance table using the following data (consisting of 2 rows, 2 columns, and 12 replicates per cell):

Row 1, Column 1
107, 126, 122, 129, 117, 128, 103, 117, 132, 139, 122, 121

Row 1, Column 2
99, 95, 79, 94, 122, 117, 99, 102, 110, 116, 121, 96

Row 2, Column 1
86, 89, 96, 101, 81, 99, 113, 79, 89, 82, 91, 74

Row 2, Column 2
104, 107, 93, 92, 82, 87, 100, 80, 102, 103, 85, 69

```
*RUN
IS DATA TO BE READ FROM A FILE (YES,NO OR STOP)
= NO
 COLUMNS  ROWS  OBSERVATIONS
= 2,2,12
SEND DATA BY COLUMNS WITHIN ROWS--ALL OBSERVATIONS
IN COLUMN 1-ROW 1, THEN COLUMN 2-ROW 1----UP TO
THE LAST COLUMN IN THE LAST ROW.

= 107 126 122 129 117 128 103 117 132 139 122 121
=  99  95  79  94 122 117  99 102 110 116 121  96
=  86  89  96 101  81  99 113  79  89  82  91  74
= 104 107  93  92  82  87 100  80 102 103  85  69

  SOURCE...   SUMS OF SQUARES..   DEG. F.    MEAN SQUARES....

ROWS.......       5.83002209E+03        1          5.83002209E+03
COLUMNS....       7.44188797E+02        1          7.44188797E+02
INTERACTION       1.17018620E+03        1          1.17018620E+03
WITHIN.....       5.82258337E+03       44          1.32331440E+02
TOTAL......       1.35669805E+04       47          2.88659157E+02
ROW  1 MEAN VALUE        1.13041666E+02
ROW  2 MEAN VALUE        9.10000000E+01
COLUMN  1 MEAN VALUE        1.05958333E+02
COLUMN  2 MEAN VALUE        9.80833330E+01
IS DATA TO BE READ FROM A FILE (YES,NO OR STOP)
= STOP

PROGRAM STOP AT 90
*
*LIST ADATA

2,2,12
107 126 122 129 117 128 103 117 132 139 122 121
 99  95  79  94 122 117  99 102 110 116 121  96
 86  89  96 101  81  99 113  79  89  82  91  74
104 107  93  92  82  87 100  80 102 103  85  69

READY

*RUN
IS DATA TO BE READ FROM A FILE (YES,NO OR STOP)
= YES
NAME OF THE DATA FILE
= ADATA

  SOURCE...   SUMS OF SQUARES..   DEG. F.    MEAN SQUARES....

ROWS.......       5.83002209E+03        1          5.83002209E+03
COLUMNS....       7.44188797E+02        1          7.44188797E+02
INTERACTION       1.17018620E+03        1          1.17018620E+03
WITHIN.....       5.82258337E+03       44          1.32331440E+02
TOTAL......       1.35669805E+04       47          2.88659157E+02
ROW  1 MEAN VALUE        1.13041666E+02
ROW  2 MEAN VALLE        9.10000000E+01
COLUMN  1 MEAN VALUE        1.05958333E+02
COLUMN  2 MEAN VALUE        9.80833330E+01
IS DATA TO BE READ FROM A FILE (YES,NO OR STOP)
= STOP

PROGRAM STOP AT 90
*
```

ANPF

This FORTRAN function computes the probability that the normally distributed random variable with a given mean and standard deviation lies between two specified numbers.

INSTRUCTIONS

The calling sequence for ANPF is:

Y = ANPF(X1, X2, XMU, SIGMA)

where,

- Y is the probability.
- X1 is the lower bound.
- X2 is the upper bound.
- XMU is the mean.
- SIGMA is the standard deviation.

NOTE:

For X1 = -6., X2 = +6., XMU = 0. and SIGMA = 1., a value of 1.0 is returned. For values of X1 and X2 larger than 9 underflow may occur.

RESTRICTION

The subprogram ERRF must be used with this subprogram, as shown in the sample problem.

SAMPLE SOLUTION

```
*LIST

10 PRINT:"WHAT ARE X1,X2,XMU,SIGMA"
20 READ:X1,X2,XMU,SIGMA
30 Y=ANPF(X1,X2,XMU,SIGMA)
40 PRINT:"Y=",Y
50 STØP;END

READY

*RUN *;ANPF;ERRF
WHAT ARE X1,X2,XMU,SIGMA
= -.5 .5 0 .9
Y=    4.2148518E-01

PRØGRAM STØP AT 50
*
```

ANVA1

This FORTRAN program performs a standard one way analysis of variance between groups of data. A maximum of 1500 data points and 100 groups can be processed.

INSTRUCTIONS

The user is given the option of entering his data directly from the keyboard or from a data file. The format of the data on a file must be as described below for keyboard input.

The program first requests the number of groups of data, and then for each group requests the number of data points in the group and their values.

Output consists of an ANOVA listing plus tests. This listing is as follows for k groups and N points.

SOURCE	Degrees of freedom	Sum of Squares	Mean Square	F Ratio
Between Groups	k - 1	Q_1	$Q_1/k - 1$	$\frac{Q_1(N-k)}{Q_2(k-1)}$
Within Groups (residual)	N - k	Q_2	$Q_2/N - k$	
TOTAL	N - 1	Q_3		

OVERALL MEAN

where,

Q_1 is the sum of squares associated with group to group variation.

Q_2 is the unexplained variation after group variation has been removed.

$S^2 = Q_2/_{N-k}$ is regarded as the estimate of residual variance.

The routine then requests a

CONFIDENCE LEVEL FOR MEANS =

For example, for a 95% band, the fraction .95 is entered. A two tail "T" value is used and generated internally by the program.

The output following this is a four column listing of group number, mean value for the group, standard deviation and half confidence band width for the sample mean.

The standard deviation is the "biased" form

$$SD = \sqrt{\frac{\Sigma (X - \bar{X})^2}{n}}$$

The program then types -

DO YOU WISH TO LOOK AT CONFIDENCE BANDS FOR GROUP MEAN DIFFERENCES, YES OR NO =

If NO, the routine requests the next problem.

If YES, the program prints -

GROUP NO., GROUP NO., CONFIDENCE =

and responds with the DIFFERENCE between the appropriate means and the confidence band half width. This procedure will continue until the user responds with 0, 0, 0. The program will then request the next problem.

SAMPLE SOLUTION

```
*RUN

   ØNE WAY ANALYSIS ØF VARIANCE

IS DATA TØ BE READ FRØM A FILE, TYPE YES,NØ ØR STØP-
= NØ

NUMBER ØF GRØUPS
= 4
NUMBER, VALUES FØR GRØUP    1
= 3 42 0 63
NUMBER, VALUES FØR GRØUP    2
= 4 45 64 33 29
NUMBER, VALUES FØR GRØUP    3
= 4 44 82 64 74
NUMBER, VALUES FØR GRØUP    4
= 4 109 120 116 97

   SØURCE              DF        SS              MS              F RATIØ
BETWEEN GRØUPS          3   1.2961983E+04   4.3206610E+03   1.2149874E+01
RESIDUAL               11   3.9117500E+03   3.5561364E+02
TØTAL                  14   1.6873733E+04

ØVERALL MEAN =   6.5466666E+01

CØNFIDENCE LEVEL FØR MEANS
= .95

           GRØUP       MEAN             SD              CØNFIDENCE
             1    3.5000000E+01    2.6191602E+01    2.3911325E+01
             2    4.2750000E+01    1.3608361E+01    2.0707815E+01
             3    6.6000000E+01    1.4212670E+01    2.0707815E+01
             4    1.1050000E+02    8.7321246E+00    2.0707815E+01
```

```
DØ YØU WISH TØ LØØK AT CØNFIDENCE BANDS FØR
GRØUP MEAN DIFFERENCES, YES ØR NØ
= YES

GRØUP NØ., GRØUP NØ., CØNFIDENCE
= 2 1 .95
 DIFFERENCE =  7.7500000E+00 PLUS/MINUS  3.1631709E+01

GRØUP NØ., GRØUP NØ., CØNFIDENCE
= 3 1 .96
 DIFFERENCE =  3.1000000E+01 PLUS/MINUS  3.3463970E+01

GRØUP NØ., GRØUP NØ., CØNFIDENCE
= 3 1 .95
 DIFFERENCE =  3.1000000E+01 PLUS/MINUS  3.1631709E+01

GRØUP NØ., GRØUP NØ., CØNFIDENCE
= 3 1 .80

 DIFFERENCE =  3.1000000E+01 PLUS/MINUS  1.9619683E+01

GRØUP NØ., GRØUP NØ., CØNFIDENCE
= 0 0 0

IS DATA TØ BE READ FRØM A FILE, TYPE YES,NØ ØR STØP
= STØP

PRØGRAM STØP AT 1110
*
```

ANVA3

This FORTRAN program performs a 3-way analysis of variance.

INSTRUCTIONS

Data for this program can be entered via the teleprinter keyboard or by using a data file (see sample problem). In either case, the free-field format is used for input.

The program initially requests the values for NA, NB, NC, and N where:

NA = number of A levels
NB = number of B levels
NC = number of C levels
N = number of observations per cell
(NA x NB x NC x N = total number of observations)

The program then requests the input data for the total number of observations. This input data can be considered a 4-dimensional matrix of the form:

X (NA, NB, NC, N)

where:

N varies first
NC varies second
NB varies third
NA varies last

[For the sample problem where NA = 3, NB = 2, NC = 4 and N = 3, the input data would be X (1, 1, 1, 1) = 52, X (1, 1, 1, 2) = 44, X (1, 1, 1, 3) = 43, X (1, 1, 2, 1) = 57, X (1, 1, 2, 2) = 69, X (1, 1, 2, 3) = 53, X (1, 1, 3, 1) = 65, . . . , X (3, 2, 4, 3) = 30.]

After this data has been entered, the program gives the user the option of printing out the input. This data is printed out as observations per cell.

The program then permits the user the option of printing out the cell means for each level of data and all combinations of levels in addition to U, the overall mean. Following this, the analysis of variance table is computed and printed. The table includes the sum of squares, degrees of freedom, mean square, and F ratio for each source of data. The remainders and totals are also printed.

RESTRICTIONS

1. Each cell must have the same number of observations.
2. The total number of observations (NA x NB x NC x N) must not exceed 400.
3. The following conditions must be observed:

   NA ≤ 10
   NB ≤ 10
   NC ≤ 10
   NA x NB x NC ≤ 400 (that is, the combination of NA, NB, and NC multiplied together cannot exceed 400. Therefore, if NA = 10 and NB = 10 then NC cannot be greater than 4.)

REFERENCE

Gunther, W. C. Analysis of Variance, Prentice Hall Inc., 1964.

Scheffe, Henry, The Analysis of Variance, John Wiley and Sons Inc., 1959.

SAMPLE PROBLEM

Perform a 3-way analysis of variance using the following data.

NOTE:

This data consists of 3 A levels, 2 B levels, 4 C levels, and 3 observations per cell:

CELL 1	52, 44, 43
CELL 2	57, 69, 53
CELL 3	65, 76, 79
CELL 4	74, 70, 104
CELL 5	48, 44, 51
CELL 6	48, 51, 42
CELL 7	58, 36, 73
CELL 8	68, 61, 77
CELL 9	38, 38, 21
CELL 10	62, 43, 39
CELL 11	64, 46, 63
CELL 12	73, 65, 50
CELL 13	32, 18, 13
CELL 14	43, 25, 34
CELL 15	37, 51, 48
CELL 16	51, 50, 74
CELL 17	44, 29, 27
CELL 18	32, 13, 27
CELL 19	27, 50, 50
CELL 20	43, 44, 50
CELL 21	20, 8, 1
CELL 22	26, 20, 11
CELL 23	32, 16, 48
CELL 24	39, 34, 30

RUN this program for the above data and print out the level means.

SAMPLE SOLUTION

These sample solutions illustrate both methods of data input.

```
*RUN

DO YOU WANT TO USE A FILE FOR INPUT (YES OR NO)
= NO
ENTER NA,NB,NC,N
= 3,2,4,3
= 52,44,43 57,69,53 65,76,79 74,70,104 48,44,51
= 48,51,42 58,36,73 68,61,77 38,38,21 62,43,39
= 64,46,63 73,65,50 32,18,13 43,25,34 37,51,48
= 51,50,74 44,29,27 32,13,27 27,50,50 43,44,50
= 20, 08,1 26,20,11 32,16,48 39,34,30
DO YOU WANT A PRINTOUT OF THE INPUT
= YES

                    PRINT-OUT OF INPUT BY CELLS
                            CELL    1
5.20000E+01 4.40000E+01 4.30000E+01
                            CELL    2
5.70000E+01 6.90000E+01 5.30000E+01
                            CELL    3
6.50000E+01 7.60000E+01 7.90000E+01
                            CELL    4
7.40000E+01 7.00000E+01 1.04000E+02
                            CELL    5
4.80000E+01 4.40000E+01 5.10000E+01
                            CELL    6
4.80000E+01 5.10000E+01 4.20000E+01
                            CELL    7
5.80000E+01 3.60000E+01 7.30000E+01
                            CELL    8
6.80000E+01 6.10000E+01 7.70000E+01
                            CELL    9
3.80000E+01 3.80000E+01 2.10000E+01
                            CELL   10
6.20000E+01 4.30000E+01 3.90000E+01
                            CELL   11
6.40000E+01 4.60000E+01 6.30000E+01
                            CELL   12
7.30000E+01 6.50000E+01 5.00000E+01
                            CELL   13
3.20000E+01 1.80000E+01 1.30000E+01
                            CELL   14
4.30000E+01 2.50000E+01 3.40000E+01
                            CELL   15
3.70000E+01 5.10000E+01 4.80000E+01
                            CELL   16
5.10000E+01 5.00000E+01 7.40000E+01
                            CELL   17
4.40000E+01 2.90000E+01 2.70000E+01
                            CELL   18
3.20000E+01 1.30000E+01 2.70000E+01
                            CELL   19
2.70000E+01 5.00000E+01 5.00000E+01
                            CELL   20
4.30000E+01 4.40000E+01 5.00000E+01
                            CELL   21
2.00000E+01 8.00000E+00 1.00000E+00
                            CELL   22
2.60000E+01 2.00000E+01 1.10000E+01
                            CELL   23
3.20000E+01 1.60000E+01 4.80000E+01
                            CELL   24
3.90000E+01 3.40000E+01 3.00000E+01
```

```
            NA= 3  NB=  2  NC= 4  N= 3
DO YOU WANT A PRINTOUT OF CELL MEANS (YES OR NO)
= YES

             A LEVEL MEANS

 XA( 1) = 6.01250E+01
 XA( 2) = 4.49167E+01
 XA( 3) = 3.00417E+01

             B LEVEL MEANS

 XB( 1) = 5.06667E+01
 XB( 2) = 3.93889E+01

             C LEVEL MEANS

 XC( 1) = 3.17222E+01
 XC( 2) = 3.86111E+01
 XC( 3) = 5.10556E+01
 XC( 4) = 5.87222E+01

             U= 4.50278E+01

             AB LEVEL MEANS

 XAB(  1) = 6.55000E+01
 XAB(  2) = 5.01667E+01
 XAB(  3) = 3.63333E+01
 XAB(  4) = 5.47500E+01
 XAB(  5) = 3.96667E+01
 XAB(  6) = 2.37500E+01

             AC LEVEL MEANS

 XAC(  1) = 4.70000E+01
 XAC(  2) = 2.66667E+01
 XAC(  3) = 2.15000E+01
 XAC(  4) = 5.33333E+01
 XAC(  5) = 4.10000E+01
 XAC(  6) = 2.15000E+01
 XAC(  7) = 6.45000E+01
 XAC(  8) = 5.15000E+01
 XAC(  9) = 3.71667E+01
 XAC( 10) = 7.56667E+01
 XAC( 11) = 6.05000E+01
 XAC( 12) = 4.00000E+01

             BC LEVEL MEANS

 XBC(  1) = 3.73333E+01
 XBC(  2) = 2.61111E+01
 XBC(  3) = 4.38889E+01
 XBC(  4) = 3.33333E+01
 XBC(  5) = 5.77778E+01
 XBC(  6) = 4.43333E+01
 XBC(  7) = 6.36667E+01
 XBC(  8) = 5.37778E+01
```

```
            ABC LEVEL MEANS

 XABC(   1) = 4.63333E+01
 XABC(   2) = 5.96667E+01
 XABC(   3) = 7.33333E+01
 XABC(   4) = 8.26667E+01
 XABC(   5) = 4.76667E+01
 XABC(   6) = 4.70000E+01
 XABC(   7) = 5.56667E+01
 XABC(   8) = 6.86667E+01
 XABC(   9) = 3.23333E+01
 XABC(  10) = 4.80000E+01
 XABC(  11) = 5.76667E+01
 XABC(  12) = 6.26667E+01
 XABC(  13) = 2.10000E+01
 XABC(  14) = 3.40000E+01
 XABC(  15) = 4.53333E+01
 XABC(  16) = 5.83333E+01
 XABC(  17) = 3.33333E+01
 XABC(  18) = 2.40000E+01
 XABC(  19) = 4.23333E+01
 XABC(  20) = 4.56667E+01
 XABC(  21) = 9.66667E+00
 XABC(  22) = 1.90000E+01
 XABC(  23) = 3.20000E+01
 XABC(  24) = 3.43333E+01

            A THREE FACTØR ANALYSIS ØF VARIANCE PRØGRAM
            WITH AN EQUAL NUMBER ØF ØBSERVATIONS PER CELL
```

SØURCE	S.S.	D.F.	MEAN SQUARE	F RATIØ
A	1.08605E+04	2	5.43026E+03	4.93723E+01
B	2.28939E+03	1	2.28939E+03	2.08153E+01
C	7.95750E+03	3	2.65250E+03	2.41167E+01
AB	1.55278E+01	2	7.76389E+00	7.05898E-02
AC	4.99583E+02	6	8.32639E+01	7.57041E-01
BC	3.21667E+01	3	1.07222E+01	9.74871E-02
ABC	6.41917E+02	6	1.06986E+02	9.72725E-01
REM	5.27933E+03	48	1.09986E+02	
TØT	2.75759E+04	71		

```
PRØGRAM STØP AT 0
*RUN

DØ YØU WANT TØ USE A FILE FØR INPUT (YES ØR NØ)
= YES
NAME THE DATA FILE
= ADATA
DØ YØU WANT A PRINTØUT ØF THE INPUT
= YES
```

```
                         PRINT-ØUT OF INPUT BY CELLS
                                   CELL     1
5.20000E+01  4.40000E+01  4.30000E+01
                                   CELL     2
5.70000E+01  6.90000E+01  5.30000E+01
                                   CELL     3
6.50000E+01  7.60000E+01  7.90000E+01
                                   CELL     4
7.40000E+01  7.00000E+01  1.04000E+02
                                   CELL     5
4.80000E+01  4.40000E+01  5.10000E+01
                                   CELL     6
4.80000E+01  5.10000E+01  4.20000E+01
                                   CELL     7
5.80000E+01  3.60000E+01  7.30000E+01
                                   CELL     8
6.80000E+01  6.10000E+01  7.70000E+01
                                   CELL     9
3.80000E+01  3.80000E+01  2.10000E+01
                                   CELL    10
6.20000E+01  4.30000E+01  3.90000E+01
                                   CELL    11
6.40000E+01  4.60000E+01  6.30000E+01
                                   CELL    12
7.30000E+01  6.50000E+01  5.00000E+01
                                   CELL    13
3.20000E+01  1.80000E+01  1.30000E+01
                                   CELL    14
4.30000E+01  2.50000E+01  3.40000E+01
                                   CELL    15
3.70000E+01  5.10000E+01  4.80000E+01
                                   CELL    16
5.10000E+01  5.00000E+01  7.40000E+01
                                   CELL    17
4.40000E+01  2.90000E+01  2.70000E+01
                                   CELL    18
3.20000E+01  1.30000E+01  2.70000E+01
                                   CELL    19
2.70000E+01  5.00000E+01  5.00000E+01
                                   CELL    20
4.30000E+01  4.40000E+01  5.00000E+01
                                   CELL    21
2.00000E+01  8.00000E+00  1.00000E+00
                                   CELL    22
2.60000E+01  2.00000E+01  1.10000E+01
                                   CELL    23
3.20000E+01  1.60000E+01  4.80000E+01
                                   CELL    24
3.90000E+01  3.40000E+01  3.00000E+01

          NA= 3   NB=  2   NC= 4   N= 3
DØ YØU WANT A PRINTØUT ØF CELL MEANS (YES OR NO)
= NØ
```

A THREE FACTOR ANALYSIS OF VARIANCE PROGRAM
WITH AN EQUAL NUMBER OF OBSERVATIONS PER CELL

SOURCE	S.S.	D.F.	MEAN SQUARE	F RATIO
A	1.08605E+04	2	5.43026E+03	4.93723E+01
B	2.28939E+03	1	2.28939E+03	2.08153E+01
C	7.95750E+03	3	2.65250E+03	2.41167E+01
AB	1.55278E+01	2	7.76389E+00	7.05898E-02
AC	4.99583E+02	6	8.32639E+01	7.57041E-01
BC	3.21667E+01	3	1.07222E+01	9.74871E-02
ABC	6.41917E+02	6	1.06986E+02	9.72725E-01
REM	5.27933E+03	48	1.09986E+02	
TØT	2.75759E+04	71		

```
PRØGRAM STØP AT 0
*LIST ADATA

3 2 4 3
52 44 43
57 69 53
65 76 79
74 70 104
48 44 51
48 51 42
58 36 73
68 61 77
38 38 21
62 43 39
64 46 63
73 65 50
32 18 13
43 25 34
37 51 48
51 50 74
44 29 27
32 13 27
27 50 50
43 44 50
20 8 1
26 20 11
32 16 48
39 34 30

READY

*
```

ANVA5

This FORTRAN program performs the calculations required for a 2- to 5- factor analysis of variance.

INSTRUCTIONS

After the program has been compiled, the routine types -

IS DATA TO BE READ FROM A FILE, TYPE YES, NO OR STOP =

If the response is NO, the program requests the necessary input.

If the response is YES, the program prints -

TYPE FILENAME, 1 TO 8 CHARACTERS =

The data on the file must be in the same format as data input from keyboard. The first record contains the number of levels for each factor entered in decreasing order from factor five to factor one. If less than five factors are to be used, a 1 must be entered for the unused factors. That is, for a 3-way factor analysis of variance, input 1, 1, n_3, n_2, n_1. The succeeding records contain the data for each cell. The cell data is provided in such a way that the index for the first variate varies most rapidly. For example, for a 2 factor analysis of variance, data for the cells is requested as follows -

```
NO., DATA FOR CELL (1  1  1  1  1 ) =
NO., DATA FOR CELL (1  1  1  1  2 ) =
   .     .     .   .  .
NO., DATA FOR CELL (1  1  1  1  n1 ) =

NO., DATA FOR CELL (1  1  1  2  1 ) =
NO., DATA FOR CELL (1  1  1  2  2 ) =
   .     .     .   .  .
NO., DATA FOR CELL (1  1  1  2  n1 ) =

   .     .     .   .  .

NO., DATA FOR CELL (1  1  1  n2  1 ) =
NO., DATA FOR CELL (1  1  1  n2  2 ) =
   .     .     .   .  .
NO., DATA FOR CELL (1  1  1  n2  n1 ) =
```

If data is read from a file, the mean and standard deviation for each cell is computed and printed in tabular form. For keyboard input, the mean and standard deviation is computed and printed immediately after the input for each cell.

After the output for the last cell, the overall (GRAND) mean is printed. Following this the analysis of variance is listed including the source, sum of squares (SS), degrees of freedom (DF) and mean square (M-SQUARE) associated with each source.

In case of unequal replications, the degree of freedom for the error term (SSE) and the total term (SST) could be noninteger since each cell is assumed to have a "weighted" number of observations.

If there is only one observation for each cell, no error term is printed, but the interaction terms can be used if they are felt to be confounded with the residual error.

RESTRICTIONS

The total number of observations per cell cannot exceed 100. The total number of cells cannot exceed 500. That is, if n_1, n_2, n_3, n_4, n_5 are levels for each succeeding factor, $n_1 n_2 n_3 n_4 n_5 \leq 500$.

It is recommended that the same number of observations be provided for each combination of factor levels (referred to as a "cell"). For an unequal number of observations, the result can be considered as an approximation.

SAMPLE PROBLEM

Consider a 3 factor problem with 3 observations per cell, where there are 3 "A" levels, 2 "B" levels and 4 "C" levels. A, B, C are associated with the first, second and third factor respectively.

	C Levels	A LEVELS 1		2		3	
B Levels		1	2	1	2	1	2
	1	52 44 43	48 44 51	38 38 21	32 18 13	44 29 27	20 8 1
	2	57 69 53	48 51 42	62 43 39	43 25 34	32 13 27	26 20 11
	3	65 76 79	58 36 73	64 46 63	37 51 48	27 50 50	32 16 48
	4	74 70 104	68 61 77	73 65 50	51 50 74	43 44 50	39 34 30

SAMPLE SOLUTION

The data was entered into the file DATA.

```
*RUN

TWØ-FIVE VARIABLE FACTØRIAL PRØGRAM

IS DATA TØ BE READ FRØM A FILE,TYPE  YES,NØ ØR STØP
= YES
TYPE FILENAME, 1 TØ 8 CHARACTERS
= DATA

  E    D    C    B    A      MEAN            SD
  1    1    1    1    1   4.633333E+01   4.027686E+00
  1    1    1    1    2   3.233333E+01   8.013878E+00
  1    1    1    1    3   3.333333E+01   7.586539E+00
  1    1    1    2    1   4.766667E+01   2.867444E+00
  1    1    1    2    2   2.100000E+01   8.041559E+00
  1    1    1    2    3   9.666667E+00   7.845735E+00
  1    1    2    1    1   5.966667E+01   6.798694E+00
  1    1    2    1    2   4.800000E+01   1.003328E+01
  1    1    2    1    3   2.400000E+01   8.041559E+00
  1    1    2    2    1   4.700000E+01   3.741657E+00
  1    1    2    2    2   3.400000E+01   7.348469E+00
  1    1    2    2    3   1.900000E+01   6.164414E+00
  1    1    3    1    1   7.333333E+01   6.018494E+00
  1    1    3    1    2   5.766667E+01   8.259676E+00
  1    1    3    1    3   4.233333E+01   1.084231E+01
  1    1    3    2    1   5.566667E+01   1.519503E+01
  1    1    3    2    2   4.533333E+01   6.018492E+00
  1    1    3    2    3   3.200000E+01   1.306395E+01
  1    1    4    1    1   8.266667E+01   1.517308E+01
  1    1    4    1    2   6.266667E+01   9.533567E+00
  1    1    4    1    3   4.566667E+01   3.091208E+00
  1    1    4    2    1   6.866667E+01   6.548968E+00
  1    1    4    2    2   5.833333E+01   1.108553E+01
  1    1    4    2    3   3.433333E+01   3.681790E+00

 GRAND MEAN=   4.502778E+01

SØURCE          SS            DF     M-SQUARE

 A           1.086053E+04      2   5.430264E+03

 B           2.289390E+03      1   2.289390E+03

 C           7.957500E+03      3   2.652500E+03

 AB          1.552694E+01      2   7.763470E+00

 AC          4.995825E+02      6   8.326375E+01

 BC          3.216576E+01      3   1.072192E+01
 ABC         6.419174E+02      6   1.069862E+02

 SSE         5.279342E+03   48.0   1.099863E+02

 SST         2.757595E+04   71.0   3.883937E+02
```

```
IS DATA TØ BE READ FRØM A FILE, TYPE  YES,NØ ØR STØP
= STØP

PRØGRAM STØP AT 1700
*LIST DATA

1  1  4  2  3
3  52  44  43
3  38  38  21
3  44  29  27
3  48  44  51
3  32  18  13
3  20  8  1
3  57  69  53
3  62  43  39
3  32  13  27
3  48  51  42
3  43  25  34
3  26  20  11
3  65  76  79
3  64  46  63
3  27  50  50
3  58  36  73
3  37  51  48
3  32  16  48
3  74  70  104
3  73  65  50
3  43  44  50
3  68  61  77
3  51  50  74
3  39  34  30

READY

*
```

BAYES

This BASIC program analyzes the difference between the arithmetic means of two sets of data whose variances may not be equal.

METHOD

This program is described in A Time-Sharing Program from Bayesian Statistics by R. C. Hewitt (T. I. S. 69CMS505). It is especially useful for those cases where the assumption of equal variance is not desired. In cases where populations variances are known, or assumed to be equal, but the sample variances are different, the program CONDIF may be used. If both the sample and population variances are equal, then the programs differ only in that BAYES uses more accurate approximations to the normal distribution.

INSTRUCTIONS

To use this program, the data is to be entered starting on line 100. The data can be entered in two ways. The first way is if the data has already been summarized. In this case only one line is to be used; that is, type only the following:

```
100 DATA H1, N1, M1, S1, H2, N2, M2, S2
RUN
```

where,

- H1 is 0 and is a dummy used only so the input is the same as for CONDIF.
- N1 is the number of elements in the first sample.
- M1 is the arithmetic mean of sample 1.
- S1 is the standard deviation of sample 1 (based on a divisor of N1).
- H2, N2, M2, S2 are the same for the second sample.

If the data has not been summarized, then it can be entered without calculating any sample characteristics. This is done by typing the data on at least two lines as:

```
100  DATA  N1,Y(1),Y(2),Y(3),Y(4)
101  DATA  Y(5),......,Y(N1)
102  DATA  N2,Z(1),Z(2),Z(3),Z(4)
103  DATA  Z(5),.........,Z(N2)
RUN
```

where N is the number of data points for sample 1 and Y(I), I = 1 to N1 is the data for sample 1, N2 is the number of data points for sample 2 and Z(I), I = 1 to N2 is the data for sample 2. It is important for entering data by the second method to make sure data is entered on both lines 100 and 101.

The output of the program needs some explanation. First, sample statistics (mean, standard deviation, variance number of data, and difference of means) are printed. Following this are given the "highest posterior density intervals" which are the shortest intervals which contain 100α % of the area under the probability density curve for the difference between means. That is, the program calculated the probability density function for the difference between means, then chose two values for the difference between means, $\theta_2 - \theta_1$, symmetric about the expected value, $Y_2 - Y_1$, such that the integral of the probability density function over the interval equalled 100α% of the integral of the same function over all possible values of $\theta_2 - \theta_1$. These intervals are numerically the same as the Behrens-Fisher confidence level intervals although the interpretation is different.

Certain values for the highest posterior density are written into the program, e.g., 50%, 75%, 90%; these may be changed by entering a different set with the data as:

200 DATA L1, L2, L3, ...

Each level (L1, L2, L3, ...) must be a number between zero and one, i.e., 50% level would be entered as 0.5.

SAMPLE SOLUTION

```
*100 DATA 0,20,67,4.5,0,25,59,3.9
*RUN


         STATISTIC                    SAMPLE 1          SAMPLE 2

SAMPLE MEAN                                 67                59
SAMPLE STD DEVIATION                       4.5               3.9
SAMPLE SIZE                                 20                25
SAMPLE VARIANCE                          20.25             15.21

DIFF OF MEANS                                8


HIGHEST POSTERIOR DENSITY INTERVALS
DIFFERENCE BETWEEN MEANS:

 HPD LEVEL       LOWER LIM         UPPER LIM

          50                 0         8.885841
          75                 0         9.521447
          90                 0         10.19852
          95                 0         10.64231
        99.9                 0         12.68091
       99.99                 0         13.72642
      99.999                 0         14.73568

READY
*100 DATA 34,1,2,3,4,5,6,7,8,9,0,1,2,3,4,5,6,7,8,9,0,1,2,3,4,5
*101 DATA 6,7,8,9,0,1,2,3,4
*102 DATA 27,4,5,6,1,2,3,8,9,0,3,4,5,6,7,8,0,9,8,1,2,3,5,4,3,2,3,4
*RUN
```

STATISTIC	SAMPLE 1	SAMPLE 2
SAMPLE MEAN	4.264706	4.259259
SAMPLE STD DEVIATIØN	2.800303	2.604733
SAMPLE SIZE	34	27
SAMPLE VARIANCE	7.841695	6.784636

DIFF ØF MEANS .0054466

HIGHEST PØSTERIØR DENSITY INTERVALS
DIFFERENCE BETWEEN MEANS:

HPD LEVEL	LØWER LIM	UPPER LIM
50	0	.4846222
75	0	.826492
90	0	1.187369
95	0	1.421475
99.9	0	2.464242
99.99	0	2.974863
99.999	0	3.451164

READY
*

BETA

This FORTRAN function evaluates the cumulative probabilities and percentage points of the Beta distribution.

METHOD

The incomplete Beta function is given by

$$B(\rho) = \int_0^{\rho} t^{\alpha-1}(1-t)^{\beta-1}\,dt \Big/ \int_0^{1} t^{\alpha-1}(1-t)^{\beta-1}\,dt$$

$$\begin{cases} 0 \le t < 1 \\ \alpha > 0 \\ \beta > 0 \end{cases}$$

The denominator of this expression is occasionally known as the Beta function and can be shown to be

$$\frac{\Gamma(\alpha)\,\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

Fortunately both integrals can be easily integrated with reasonable accuracy by Gaussian quadrature techniques.

Noting that $B(\rho)$ is single valued and monotonically increasing, Newton's method should work satisfactorily on the function

$$F(\rho) = B(\rho) - Q$$

in determining ρ given Q.

This results in the recursion

$$\rho_{i+1} = \rho_i + \left\{ \frac{Q\int_0^{1} t^{\alpha-1}(1-t)^{\beta-1}dt - \int_0^{\rho_i} t^{\alpha-1}(1-t)^{\beta-1}dt}{\rho_i^{\alpha-1}(1-\rho_i)^{\beta-1}} \right\}$$

This distribution has expected value $\frac{\alpha}{\alpha+\beta}$ and variance $\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$.

In the recursion process, $\frac{\alpha}{\alpha+\beta}$ is used as a starting value and experience has shown that 5 iterations are satisfactory.

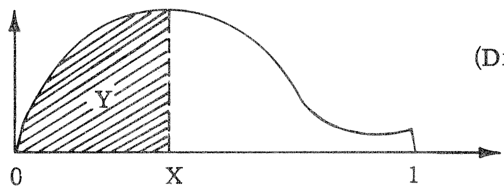
INSTRUCTIONS

The calling sequence is

Y = BETA(IND,X,A,B)

1. IND = 0 corresponds to the evaluation of the cumulative probability, Y, given the percentage point, X where

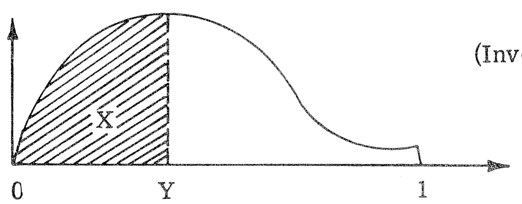
   Y = Pr(BETA ≤ X)



(Direct Beta Distribution)

   IND = 1 corresponds to the evaluation of the percentage point, Y, given the cumulative probability, X, where

   X = Pr(BETA ≤ Y)



(Inverse Beta Distribution)

2. X is as defined in (1) above.

3. A is a parameter corresponding to α as described in the method.

4. B is a parameter corresponding to β as described in the method.

5. Y is the result returned as defined in (1) above.

RESTRICTIONS

It is recommended that for reasonable numerical accuracy

a. $1 \leq A \leq 40$
b. $1 \leq B \leq 40$
c. $0 \leq X \leq 1$

Also results appear to agree to 3 or 4 digits with most tables if

$05 \leq BETA(\rho) \leq .5$

SAMPLE SOLUTION

```
10 PRINT:"WHAT ARE IND,X,A,B";READ:IND,X,A,B
20 Y=BETA(IND,X,A,B)
30 PRINT:"RESULTS",Y,X
40 STØP;END

READY

*RUN *;BETA
WHAT ARE IND,X,A,B
= 0 .45 2 3
RESULTS     6.0901874E-01     4.5000000E-01

PRØGRAM STØP AT 40
*RUN *;BETA
WHAT ARE IND,X,A,B
= 1 .60901874 2 3
RESULTS     4.5000000E-01     6.0901874E-01

PRØGRAM STØP AT 40
*
```

BICONF

This BASIC program determines the confidence limits for a population proportion based on the exact binomial distribution.

INSTRUCTIONS

Enter values for X, N, and C when requested by the program.

X - Successes.

N - Sample size.

C - Confidence Coefficient in percent.

Additional instructions may be found in the listing.

SAMPLE PROBLEM

A polling agency makes a sample of 200 voters in a certain city and it is found that 110 of these people intend to vote for Candidate A. Therefore, the best estimate that can be made from this sample is that 55 percent of the entire population intend to vote for Candidate A.

If the agency wants to publish a prediction, with a 95 percent chance that they will be correct that the actual percentage of the entire population will be within certain bounds, what limits should they choose?

SAMPLE SOLUTION

```
*RUN


DØ YØU WANT INSTRUCTIØNS (1=YES, 0=NØ) ?0
WHAT ARE X,N,C ?55,100,95

    BEST ESTIMATE ØF PØPULATIØN PRØPØRTIØN (PCT) =     55

    THE     95 PERCENT CØNFIDENCE LIMITS ØN THE PØPULATIØN
    PRØPØRTIØN (PCT) ARE  44.7279 AND   64.968


READY
*
```

BINDIS

This BASIC program computes the binomial distribution function for problems of this type: In a series of N independent trials, each with probability of success P, what is the probability that there will be exactly X successes?

INSTRUCTIONS

Enter values for P, N, and M, when requested by the program.

NOTE:

> This program will determine all the values of X for which the probability is significantly large [up to 7 standard deviations away from P * N.]

The number M is a multiplier of the Binomial function useful in scaling the function to a conveniently graphable range. [M does not affect the values of the binomial function, but is multiplied and printed out in a separate column. Each value of B * M is rounded off to the nearest integer.]

To terminate the program type "STOP" at any request for information.

Additional instructions may be found in the listing.

SAMPLE PROBLEM

Determine the probability that in 30 rolls of a die it will come up 3 exactly 6 times (or 7, or 8, or 9, etc).

To use this program enter the following data: P = 1/6 = .166667

N = 30

M = 100

SAMPLE SOLUTION

```
*RUN


DØ YØU WANT INSTRUCTIØNS (1=YES, 0=NØ) ?0

WHAT ARE P,N,M ?.166667,30,100

BINØMIAL DISTRIBUTIØN:  N=   30   P=   0.166667

          X          X/N          B(X;P,N)          B*   100

          0            0          0.0042127              0
          1    0.0333333          0.0252761              3
          2    0.0666667          0.0733008              7
          3          0.1          0.1368285             14
          4    0.1333333          0.1847189             18
          5    0.1666667          0.1921081             19
          6          0.2          0.1600905             16
          7    0.2333333          0.1097766             11
          8    0.2666667          0.0631217              6
          9          0.3          0.0308596              3
         10    0.3333333           0.012961              1
         11    0.3666667          0.0047131              0
         12          0.4          0.0014925              0
         13    0.4333333          0.0004133              0
         14    0.4666667          0.0001004              0
         15          0.5        2.14133 E-5              0
         16    0.5333333        4.01500 E-6              0
         17    0.5666667        6.61296 E-7              0
         18          0.6        9.55208 E-8              0
         19    0.6333333        1.20658 E-8              0

WHAT ARE P,N,M ?STØP

READY
*
```

BINOM

This FORTRAN program computes binomial probabilities (discrete or cumulative) and also derives confidence bands for the parameter.

METHOD

The routine offers the user three options:

Option 1: The direct calculation of ρ (x) and $\sum \rho$ (x) given n, x and θ where

n	is the number of trials
x	is the number of successes
θ	is the probability of a success
$\rho(x) =$	$\binom{n}{x} \theta^x (1 - \theta)^{n-x}$

Option 2: Given n, x and the probability P desired for your confidence band, the program calculates $\bar{\theta}$ and $\underline{\theta}$ such that Prob($\underline{\theta} \le \theta \le \bar{\theta}$) = P, that is, a two tail confidence band, and $\underline{\underline{\theta}}$ such that Prob($\underline{\underline{\theta}} \le \theta$) = P, that is, a one tail confidence band where θ is the empirical value of your parameter. These confidence bands are derived from the F-distribution.

Option 3: Given n, x and the probability P desired for your confidence band, the program calculates $\underline{\theta}$ such that Prob($\underline{\theta} \le \theta$) = P, that is, a one tail confidence band, where newly occurring successes are used to adjust the "a priori" confidence bands for θ. The "a priori" alpha and beta must be supplied in this case. These confidence bands are derived from the Beta distribution.

INSTRUCTIONS

The codes for the options are as follows:

P - PROBABILITY EVALUATION
2 - CONFIDENCE BAND FOR PARAMETER
3 - BAYESIAN CONFIDENCE BAND

If option 1 is chosen the program requests:

PARAMETER, TOTAL NO., NO. POINTS, DATA =

where,

- PARAMETER is the probability of a success
- TOTAL NO. is the total number of trials
- NO. POINTS is the total number of cases
- DATA is the total number of successes for each case

If options 2 or 3 are chosen the program requests:

NO. TRIALS, EVENTS, CONFIDENCE (0. - 1.) =

where,

- NO. TRIALS is the total number of trials
- EVENTS is the number of events
- CONFIDENCE (0. - 1.) is the probability desired for your confidence band

In option 3 the program also requests:

A PRIORI ALPHA, BETA =

RESTRICTIONS

The number of trials must be less than 100. If using option 3, the values of "a priori" Alpha and Beta must be less than 80.

SAMPLE SOLUTION

```
*RUN BINØM ;BETA


BINØMIAL DISTRIBUTIØN PRØGRAM

ØPTIØN CØDE

1 - PRØBABILITY EVALUATIØN
2 - CØNFIDENCE BAND FØR PARAMETGR
3 - BAYESIAN CØNFIDENCE BAND


READ IN CØDE
= 1

PARAMGTER,TØTAL NØ.,NØ. PØINTS, DATA
= .8262112  50   10 30 31 32 33 34 35 36 37 38 39
NUMBER   PRØBABILITY   CUM. PRØBABILITY
   39      0.09530720     0.14741392
   38      0.06515380     0.08226612
   37      0.04005998     0.04220614
   36      0.02226974     0.01993641
   35      0.01124235     0.00869406
   34      0.00517292     0.00352114
   33      0.00217619     0.00134495
   32      0.00083921     0.00050574
   31      0.00029730     0.00020844
   30      0.00009693     0.00011151
```

```
READ IN CØDE
= 2

NØ. TRIALS, EVENTS, CØNFIDENCE(0.-1.)
= 7 6 .95

LØWER LIMIT =   4.212836E-01
UPPER LIMIT =   0.996390E+00
ØNE-TAIL LIMIT(LØWER BØUND) =   4.792970E-01

READ IN CØDE
= 3

NØ. TRIALS, EVENTS, CØNFIDENCE(0.-1.)
= 50 46 .95

A PRIØRI ALPHA, BETA
= 25 25

LØWER LIMIT =   6.334549E-01

READ IN CØDE
= 0

PRØGRAM STØP AT 690
*
```

BITEST

This BASIC program performs a statistical test of a binomial proportion.

INSTRUCTIONS

Enter values for X, N, and P when requested by the program.

NOTE:

X - Successes in sample

N - Sample size

P - Population proportions

Additional instructions may be found in the listing.

SAMPLE PROBLEM

Consider a city in which 75 percent of the population intends to vote for Candidate A (and the rest for some other candidate). From a survey of 100 people picked at random, what is the probability that 60 percent or less (i.e. 60 people) are planning to vote for Candidate A?

Let a "success" be a person in the sample who intends to vote for Candidate A. Therefore, the input to the program will be:

X(Number of successes in sample)	= 60
N(Sample size)	= 100
P(The true proportion of the population intending to vote for A)	= .75

Obviously the accuracy of a smaller sample (say 20 people instead of 200) is much less. This is demonstrated by the second of the 2 sample solutions.

SAMPLE SOLUTION

```
*RUN


DØ YØU WANT INSTRUCTIØNS (1=YES, O=NØ) ?0
WHAT ARE X,N,P ?60,100,.75

IN SAMPLES ØF SIZE    100 RANDØMLY SELECTED FRØM A
BINØMIAL PØPULATIØN HAVING A TRUE PRØPØRTIØN ØF 0.75
THE PRØBABILITY ØF A SAMPLE HAVING     60 ØR LESS
SUCCESSES IS    0.000687


READY
*RUN


DØ YØU WANT INSTRUCTIØNS (1=YES, O=NØ) ?0
WHAT ARE X,N,P ?12,20,.75

IN SAMPLES ØF SIZE     20 RANDØMLY SELECTED FRØM A
BINØMIAL PØPULATIØN HAVING A TRUE PRØPØRTIØN ØF 0.75
THE PRØBABILITY ØF A SAMPLE HAVING     12 ØR LESS
SUCCESSES IS    0.101812


READY
*
```

CFIT

This FORTRAN program finds, by the method of least squares, the polynomial of degree N, where K ≤ N ≤ K+M whose graph contains the points (A(1), B(1)) ... (A(K1, B(K)) and approximates the points (X(1), Y(1)) ... (X(M), Y(M)) and W(I) is the weight corresponding to (X(I), Y(I)). It accepts NR points Z(1) ... Z(NR) and evaluates the polynomial at these points.

METHOD

The method used is one developed by J.E.L. Peck as described in The Communications of the ACM, January 1962. A subroutine developed by G.E. Forsyth is used to find the polynomial of degree K-1 whose graph contains the points (A(1), B(1)) through (A(K), B(K)). Then after adjusting the weights and ordinates, the subroutine is used again to find the polynomial fitting the original and approximate weight points.

INSTRUCTIONS

To use this program, enter the data as requested. A maximum of 100 points can be approximated and a maximum of 200 points can be tested.

This program may also be used to find the polynomial of degree N fitting N+1 points by setting M equal to zero and entering 1.0 for X(1), Y(1), and Z(1). If at any time M is non-zero, then at least one of the weights must be zero.

SAMPLE PROBLEM

Find the 4th order polynomial that contains the points (0, 0) and (5, 5) and approximate the points:

X(I):	1.0	.5	1.5	2.0	3.0	2.5	3.5	4.0
Y(I):	-1.0	-1.2	-1.5	-1.1	- .5	-1.1	- .2	1.5
weight:	1.0	3.0	3.0	1.0	1.0	1.0	3.0	1.0

Then test the fit for the points: 0, 1.5, 2.0, 2.5, 3.0, 5.0

SAMPLE SOLUTION

```
*RUN CFIT

 INSTRUCTIØNS ? (YES ØR NØ) WHICH ?
= YES
CFIT FINDS THE PØLYNØMIAL ØF DEGREE N,K<=N<K+M
WHØSE GRAPH CØNTAINS PØINTS (A(1),B(1))...(A(K),B(K))
AND APPRØXIMATES (X(1),Y(1))...(X(M),Y(M)) WITH WEIGHTS
W(1)...W(M).IT ACCEPTS NR PØINTS Z(1)...Z(NR) AND
EVALUATES THE PØLYNØMIAL AT THESE PØINTS.THEN THE
CØEFFICIENTS ØF THE PØLYNØMIAL ARE ØUTPUT.
THE NUMBER ØF A-B PAIRS?
= 2
ENTER THE A(I),B(I) PAIRS
= 0,0
= 5.0,5.0
THE NUMBER ØF X-Y PAIRS?
= 8
THE X(I)-Y(I) PAIRS AND THEIR WEIGHTS
= 1.0,-1.0,1.0
= .5,-1.2,3.0
= 1.5,-1.5,3.0
= 2.0,-1.1,1.0
= 3.0,-.5,1.0
= 2.5,-1.1,1.0
= 3.5,-.2,3.0
= 4.0,1.5,1.0
WHAT IS THE DEGREE ØF THE PØLYNØMIAL YØU WANT TØ FIT?
= 4
IF THE INDEPENDENT VARIABLES(THE A AND X VALUES) ARE
TØ BE NØRMALIZED,EITHER TYPE IN THEIR MEAN AND
STANDARD DEVIATIØN(NØRMALIZING CØNSTANT) ØR ENTER
THEM BØTH AS ZERØ AND THEY WILL BE CØMPUTED BY THE
PRØGRAM. IF NØRMALIZATIØN IS NØT DESIRED, ENTER A
MEAN ØF ZERØ AND A NØRMALIZING CØNSTANT ØF 1.
MEAN ?
= 0
NØRMALIZING CØNSTANT ?
= 0
HØW MANY PØINTS DØ YØU WANT TØ TEST?
= 6
ENTER THE Z(I) PØINTS
= 0,1.5,2.0,2.5,3.0,5.0
 VALUE ØF X          VALUE ØF PØLYNØMIAL AT X
 0.                -0.24974589E+01
 0.15000000E+01     0.10544062E-03
 0.20000000E+01     0.49003530E+00
 0.25000000E+01     0.80866930E+00
 0.30000000E+01     0.95600748E+00
 0.50000000E+01    -0.16759780E+00
 PØLYNØMIAL CØEFFICIENTS
C(  1)=     -0.24974589E+01
C(  2)=      0.21789303E+01
C(  3)=     -0.34259162E+00
C(  4)=      0.
C(  5)=      0.

PRØGRAM STØP AT 13310
*
```

CHISQR

This FORTRAN program is used to obtain a measure of the discrepancy between an empirical distribution and a known theoretical distribution such as the Normal Distribution by means of the Chi-Square Test.

METHOD

Compute the Chi-Square Test according to the formula:

$$x^2 = \sum_{j=1}^{k} \frac{(f_{oj} - f_{ej})^2}{f_{ej}}$$

where,

f_o = observed frequency

f_e = expected frequency.

INSTRUCTIONS

Enter data as requested. For additional instructions run the program.

SAMPLE SOLUTION

```
*RUN


CHISQR           REVISED  10/13/69   GVK

INSTRUCTIØNS ? (IF DESIRED PRINT Y,ØTHERWISE N)
= N
 SUPPLY VALUES FØR THE FØLLØWING  VARIABLES : -
 NN,KALL,TØTAL, SMALL
= 3
= 1
= 53
= 3.0
 SUPPLY NN VALUES FØR  ØBS



= 23
= 21
= 24



 SUPPLY NN VALUES FØR  EXPT



= 17
= 14
= 15

    I       ØBS FR        EXP FR       ((Ø-E)**2)/)

    1       23.000       17.000       855.587
    2       21.000       14.000       700.594
    3       24.000       15.000       747.725
 CHI SQR =    2.3039060E+03



PRØGRAM STØP AT 62
*
```

COLINR

This BASIC program defines confidence limits on simple linear regressions.

NOTE: Under certain conditions this program has produced inaccurate results.

INSTRUCTIONS

Enter input data using the following format:

```
100   DATA  116,  132,  104,  139
200   DATA  105,  120,   85,  121
```

Then type: RUN

NOTE:

Lines number 0-699 are free for use as data statements, as required. All values for one variable are given in data statements, followed by all data for a second variable, and so on. One convenient way to arrange input is to put all values of variable 1 at lines 100-199, all for variable 2 at lines 200-299, etc. If separate tapes are made for each variable, the full use of input flexibility can be obtained.

Additional instructions may be obtained by listing the program STADES.

SAMPLE PROBLEMS

Determine the confidence limits using the following observed data:

```
Y   =   116, 132, 104, 139

X   =   105, 120, 85, 121
```

The second sample uses the following data:

```
X = -3, -2, -1, -0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

Y = 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0, -1, -2, -3
```

SAMPLE SOLUTION

```
*100 DATA 116,132,104,139
*200 DATA 105,120, 85,121
*RUN


COLINR


S I M P L E   L I N E A R   R E G R E S S I Ø N

          EQUATIØN:    Y = A + B*X

HØW MANY ØBSERVATIØNS ØN EACH VARIABLE ?4

VARIABLE                MEAN          VARIANCE       STD DEVIATIØN

   X                   107.75         283.5833          16.83993
   Y                   122.75         248.9167          15.77709

 INDEX (R↑2)         EXPL VAR        UNEXPL VAR       STD ERRØR

   0.9423758         175.9298         10.75772          3.279896

PARAMETER              VALUE          95 PCT CØNFIDENCE LIMITS

   A                 24.75228        -49.42544            98.93
   B                0.9094916        0.2272877          1.591696

ESTIMATED VALUES ØF Y (FRØM THE REGRESSIØN) AND CØNFIDENCE
LIMITS FØR INDIVIDUAL VALUES ØF Y, FØR EACH VALUE ØF X:

       X-ACTUAL         Y-ACTUAL        Y-CALC          95 PCT CØNFIDENCE LIMITS

         105              116          120.2489          97.923          142.5748
         120              132          133.8913        110.1265          157.6561
          85              104          102.0591        74.93341          129.1847
         121              139          134.8008        110.7876           158.814


READY
*
```

```
*100 DATA -3, -2, -1, -0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
*200 DATA 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0, -1, -2, -3
*RUN


COLINR



S I M P L E   L I N E A R   R E G R E S S I O N

          EQUATION:    Y = A + B*X

HOW MANY OBSERVATIONS ON EACH VARIABLE ?14

VARIABLE                 MEAN          VARIANCE     STD DEVIATION

   X                       3.5            17.5            4.1833
   Y                       3.5            17.5            4.1833

 INDEX (R↑2)          EXPL VAR        UNEXPL VAR       STD ERROR

         1              16.25              0               0

PARAMETER             VALUE          95 PCT CONFIDENCE LIMITS

   A                         7              7               7
   B                        -1             -1              -1

ESTIMATED VALUES OF Y (FROM THE REGRESSION) AND CONFIDENCE
LIMITS FOR INDIVIDUAL VALUES OF Y, FOR EACH VALUE OF X:

      X-ACTUAL       Y-ACTUAL      Y-CALC          95 PCT CONFIDENCE LIMITS-

          10            -3            -3              -3              -3
           9            -2            -2              -2              -2
           8            -1            -1              -1              -h
           7             0             0               0               0
           6             1             1               1               1
           5             2             2               2               2
           4             3             3               3               3
           3             4             4               4               4
           2             5             5               5               5
           1             6             6               6               6
           0             7             7               7               7
          -1             8             8               8               8
          -2             9             9               9               9


          -3            10            10              10              10


READY
*
```

CONBIN

This BASIC program determines confidence limits for a population proportion based on the normal curve approximation.

NOTE:

For small sample sizes use the program BICONF.

INSTRUCTIONS

To use this program enter values for X and N when requested by the program.

NOTE:

X = number of successes in sample
N = sample size

SAMPLE PROBLEM

A polling agency makes a sample of 200 voters in a certain city and it is found that 110 of these people intend to vote for Candidate A. Suppose the agency is willing to settle for accuracy of only 50 percent (or 75, or 90, or 95, etc.) in their prediction, what limits should they choose?

Suppose that they decide to make a larger survey for greater accuracy. This time they poll 2000 people and find that 1150 people (or 57.5 percent) are planning to vote for Candidate A. What limits should they then pick for a 50 percent (or 75, or 90, or 95, etc.) chance of accuracy?

SAMPLE SOLUTION

```
*RUN


CØNBIN



CØNFIDENCE LIMITS FØR A PØPULATIØN PRØPØRTIØN BASED ØN
THE NØRMAL CURVE APPRØXIMATIØN. (FØR SMALL SAMPLE SIZES
USE THE PRØGRAM BICØNF).  WHAT ARE X (NUMBER ØF SUCC-
ESSES IN SAMPLE), N (SAMPLE SIZE) ?110,200

   BEST ESTIMATE ØF PØPULATIØN PRØPØRTIØN (PCT) =    55


CØNFIDENCE LIMITS ØN PØPULATIØN PRØPØRTIØN:

  CØNF LEVEL        LØWER LIM        UPPER LIM

          50           52.377           57.623
          75           50.703           59.297
          90           48.964           61.036
          95           47.855           62.145
          99           45.689           64.311
        99.9           43.174           66.826
       99.99           41.064           68.936
```

```
READY
*RUN


CØNBIN


CØNFIDENCE LIMITS FØR A PØPULATIØN PRØPØRTIØN BASED ØN
THE NØRMAL CURVE APPRØXIMATIØN. (FØR SMALL SAMPLE SIZES
USE THE PRØGRAM BICØNF).  WHAT ARE X (NUMBER ØF SUCC-
ESSES IN SAMPLE), N (SAMPLE SIZE) ?1150,2000

   BEST ESTIMATE ØF PØPULATIØN PRØPØRTIØN (PCT) = 57.5

CØNFIDENCE LIMITS ØN PØPULATIØN PRØPØRTIØN:

  CØNF  LEVEL       LØWER LIM         UPPER LIM

          50         56.729            58.271
          75         56.203            58.797
          90         55.657            59.343
          95         55.308            59.692
          99         54.628            60.372
        99.9         53.838            61.162
       99.99         53.174            61.826

READY
*
```

CONDIF

This BASIC program computes confidence limits for the difference between two population means, based on data supplied for two samples, one from each population.

INSTRUCTIONS

To use this program type the data in the following format:

```
200 DATA H1,N1,M1,S1,  H2,N2,M2,S2
THEN RUN
```

NOTE:

H1 = SIZE OF POPULATION 1 [LET H1 EQUAL ZERO IF POPULATION IS INFINITE]

N1 = SIZE OF SAMPLE 1

M1 = ARITHMETIC MEAN OF SAMPLE 1

S1 = STANDARD DEVIATION OF SAMPLE 1 [BASED ON DIVISOR OF N1]

H2, N2, M2, S2 ARE THE SAME FOR SAMPLE 2.

Additional instructions may be found in the listing.

SAMPLE PROBLEM

Determine the confidence limits for these 2 samples

```
(1)  0, 21, 30.9048, 29.2752
(2)  0, 21, 44, 26.8239
```

SAMPLE SOLUTION

```
*200 DATA 0,21,30.9048,29.2752,0,21,44,26.8239
*RUN


        STATISTIC                  SAMPLE 1         SAMPLE 2

SAMPLE MEAN                         30.9048               44
SAMPLE VARIANCE                    857.0373         719.5216
SAMPLE STD DEVIATIØN                29.2752          26.8239
SAMPLE SIZE                              21               21
PØPULATIØN SIZE                    INFINITE         INFINITE
ESTIM PØPN STD DEV                 29.99815         27.48632
STD ERRØR ØF MEAN                  6.546134         5.998006

DIFF BETWEEN MEANS                       -13.0952
STD ERRØR ØF DIFF                         8.87851
DEGR ØF FREEDØM (DIFF)                    39.7

CØNFIDENCE LIMITS ØN DIFFERENCE BETWEEN MEANS:

CØNF LEVEL      LØWER LIM       UPPER LIM

        50      -19.13895       -7.051454
        75      -23.46018       -2.730221
        90      -28.04783        1.857428
        95      -31.04327        4.852865
        99      -37.11432        10.92392
      99.9      -44.63356        18.44316
     99.99      -51.46557        25.27517
    99.999      -57.95332        31.76292

READY
*
```

CONLIM

This BASIC program computes confidence limits for an unknown population mean based on random sample data.

INSTRUCTIONS

To use this program enter the input data using this format:

```
50   DATA   [SIZE OF POPULATION]
     [OMIT THIS INPUT IF INFINITE POP'N]
100  DATA  X[1], X[2],.....,X[N]

   WHERE THE X[I] ARE THE SAMPLE OBSERVATIONS.
```

Additional instructions may be found in the listing.

SAMPLE PROBLEM

Analyze the data shown in the following sample solution.

SAMPLE SOLUTION

```
*100 DATA 2,3,6,9,8,11,13,16,23,45,67,89,35,67,13,43,21,67,98,7,6
*RUN




CONLIM



VALUES OF SAMPLE STATISTICS:

   SIZE OF SAMPLE                           21
   SAMPLE MEAN VALUE                  30.90476
   VARIANCE OF SAMPLE                 857.0386
   SAMPLE STD DEVIATION               29.27522
   ESTIMATED POPN STD DEV             29.99817
   STANDARD ERROR OF MEAN             6.546138

CONFIDENCE LIMITS ON POPULATION MEAN:

  CONF LEVEL      LOWER LIM         UPPER LIM

          50       26.40788          35.40164
          75       23.14931          38.66021
          90       19.61534          42.19419
          95       17.25186          44.55766
          99       12.28898          49.52055
        99.9       5.752042          56.05748
       99.99     -0.6232884          62.43281
      99.999      -7.094567          68.90409

READY
*
```

CORREL

This FORTRAN program measures the extent of association or relation between two sets of attributes.

METHOD

Let $X_1, X_2, X_3, \ldots X_n$ and $Y_1, Y_2, Y_3, \ldots Y_m$ be two sets of attributes. Denote the frequency of the mutual occurrence of X_i and Y_j by (X_i, Y_j) and for a given sample construct the contingency table, which will be the input to the program. The Contingency Coefficient, C, is a measure of the correlation between the two sets of attributes.

TABLE I

	X_1	X_2	.	.	.	X_n
Y_1	(X_1Y_1)	(X_2Y_1)	.	.	.	(X_nY_1)
Y_2	(X_1Y_2)	(X_2Y_2)	.	.	.	(X_nY_2)
.	.	.	.	.	.	.
Y_m	(Y_1Y_m)	(X_2Y_m)	.	.	.	(X_nY_m)

$$C = \sqrt{\frac{X^2}{N + X^2}}$$

where,

$$X^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

$$N = \sum_{j=1}^{m} \sum_{i=1}^{n} (X_iY_j)$$

and X^2 is computed in the usual manner from (TABLE I).

INSTRUCTIONS

Enter data as requested.

RESTRICTION

The subprogram XINGAM must be used with this program as shown in the sample problem.

SAMPLE PROBLEM

To determine the contingency coefficient of the following two sets of attributes: X_1, X_2, X_3, X_4 and Y_1, Y_2, Y_3 with the contingency table.

	X_1	X_2	X_3	X_4
Y_1	23	40	16	2
Y_2	11	75	107	14
Y_3	1	31	60	10

SAMPLE SOLUTION

```
*RUN CØRREL;XINGAM
INPUT NUMBER ØF RØWS AND CØLUMNS
= 3 4

INPUT CØNTINGENCY TABLE BY RØWS
= 23 40 16 2   11 75 107 14   1 31 60 10


CHI SQUARE =        0.69389324E+02


YØUR CØNTINGENCY CØEFFICIENT:C =        0.38864750E+00


C IS SIGNIFICANT AT THE      0.99999999E-03  LEVEL


PRØGRAM STØP AT 450
*
```

CORRL2

This FORTRAN program measures the association or correlation between two sets of N elements each by the Spearman-Rank Correlation Coefficient.

METHOD

Given two sets of values $x_1', x_2', \ldots, x_N'$ and $y_1', y_2', \ldots, y_N'$ replace each element of the two sets by its rank x_i or y_i and form the sets $x_1, x_2, \ldots, x_N$ and $y_1, y_2, \ldots, y_N$. The Spearman Rank Correlation Coefficient r_s is as follows:

$$r_s = \frac{\Sigma x^2 + \Sigma y^2 - \Sigma d_i^2}{2\sqrt{\Sigma x^2 \cdot \Sigma y^2}}$$

where

$$d_i = x_i - y_i$$

$$\Sigma x^2 = \frac{N^3 - N}{12} - \sum_{i=1}^{N} T_{x_i}$$

$$\Sigma y^2 = \frac{N^3 - N}{12} - \sum_{i=1}^{N} T_{y_i}$$

$T_{x_i} = \frac{t^3 - t}{12}$ where t is the number of observations tied at a given x rank.

$T_{y_i} = \frac{t^3 - t}{12}$ where t is the number of observations tied at a given y rank.

To find the significance level at which we may reject the null hypothesis

H_0: The two sets of values are uncorrelated.

We consider $TAU = r_s\sqrt{\frac{N - 2}{1 - r_s^2}}$.

For N sufficiently large this has a t distribution with df = N - 2.

INSTRUCTIONS

Enter data as requested.

RESTRICTION

The subprograms TDIST and BETA must be used with this program as shown in the sample problem.

SAMPLE SOLUTION

```
*RUN CØRRL2;TDIST;BETA


NUMBER ØF ELEMENTS IN EACH ARRAY
= 12


INPUT X,Y ARRAYS
* 42 46 39 37 65 88 86 56 62 92 54 81
* 0 0 1 1 3 4 5 6 7 8 8 12


THE SPEARMAN CØRRELATIØN CØEFFICIENT: RS=      0.61512278E+00


THIS VALUE ØF RS IS SIGNIFICANT AT THE      0.17790236E-01  LEVEL


PRØGRAM STØP AT 600
*
```

CURFIT

This BASIC program determines which of 6 curves best fits the supplied data.

INSTRUCTIONS

Enter data on lines numbered 41-699. All values for the dependent variable (Y) are given in data statements followed by all data for the independent variable (X). One convenient arrangement might be lines 100-199 for one variable and lines 200-299 for the other variable.

NOTE:

This program accepts up to 200 observations on 2 variables.

Additional instructions may be obtained by listing the program STADES.

SAMPLE PROBLEM

Find which of 6 curves presents the closest fit to the following data points:

X	Y	X	Y
1	1.0	7	4.0
2	1.5	8	4.5
3	2.0	9	5.0
4	2.5	10	5.5
5	3.0	11	6.0
6	3.5	12	6.5

Supply details for curves 1 and 3.

SAMPLE SOLUTION

```
*100 DATA 1,1.5,2,2.5,3,3.5,4,4.5,5,5.5,6,6.5
*200 DATA 1,2,3,4,5,6,7,8,9,10,11,12
*RUN


PLEASE SPECIFY THE NUMBER ØF VALUES (N) GIVEN AS DATA
FØR THE TWØ INPUT VARIABLES, AND THE ØUTPUT CØDE (D).
(D=1 IF ØUTPUT IS TØ BE IN ØRDER ØF INCREASING VALUES
ØF THE INDEPENDENT VARIABLE, ELSE D=0).   N,D =  ?12,0


 L E A S T   S Q U A R E S   C U R V E   F I T

CURVE TYPE          INDEX  ØF             A               B
                  DETERMINATIØN

1. Y=A+(B*X)              1              0.5             0.5
2. Y=A*EXP(B*X)     0.9373864      1.178574       0.1572145
3. Y=A*(X↑B)        0.9932813      0.9051706      0.7719925
4. Y=A+(B/X)         0.612308         5.1356      -5.358062
5. Y=1/(A+B*X)      0.7573067      0.7589013     -0.0608532
6. Y=X/(A+B*X)      0.9684889      0.9424266      0.1196433


DETAILS FØR ?1

    1. Y=A+(B*X)                  IS A LINEAR FUNCTIØN.  THE RESULTS
       ARE AS FØLLØWS:

       X-ACTUAL        Y-ACTUAL      Y-CALC        PCT DIFFER

           1               1             1             0
           2             1.5           1.5             0
           3               2             2             0
           4             2.5           2.5             0
           5               3             3             0
           6             3.5           3.5             0
           7               4             4             0
           8             4.5           4.5             0
           9               5             5             0
          10             5.5           5.5             0
          11               6             6             0
          12             6.5           6.5             0
```

```
DETAILS FØR ?3

   3. Y=A*(X↑B)                    IS A PØWER FUNCTIØN.  THE RESULTS
      ØF A LEAST-SQUARES FIT ØF ITS LINEAR TRANSFØRM
      ARE AS FØLLØWS:

      X-ACTUAL          Y-ACTUAL     Y-CALC         PCT DIFFER

          1                 1        0.9051706              10.4
          2               1.5        1.545693      -2.9
          3                 2        2.113803      -5.3
          4               2.5        2.639467      -5.2
          5                 3        3.135668      -4.3
          6               3.5        3.609586        -3
          7                 4        4.065742      -1.6
          8               4.5        4.507224      -0.1
          9                 5        4.936266                1.2
         10               5.5         5.35455                2.7
         11                 6        5.763387                4.1
         12               6.5        6.163825                5.4


DETAILS FØR ?S

READY
*
```

EXPLIM

This BASIC program computes confidence limits for an unknown population mean based on random sample data, and given that the population is known to be exponentially distributed.

INSTRUCTIONS

To use this program, enter data using this format:

100 DATA X[1], X[2], X[3], _ _ _, X[N]

Where the X[N] are random sample observations.

Additional instructions may be found in the listing.

NOTE

Sample statistics will be given for any size sample; confidence limits will be calculated for samples of size eleven or greater.

SAMPLE PROBLEM

Analyze the data which is entered in the following sample solution.

SAMPLE SOLUTION

```
*100 DATA 3.4,3.2,3.5,3.6,2.9,3.4,3.1,3,3.4,3.2,4,2.8
*RUN

EXPLIM

VALUES ØF SAMPLE STATISTICS
------ -- ------ ----------

SIZE ØF SAMPLE                  12
SAMPLE MEAN                     3.291667
SAMPLE VARIANCE                 0.1099248
SAMPLE STD DEVIATIØN            0.3315491

CØNFIDENCE LIMITS ØN PØPULATIØN MEAN
---------- ------ -- ---------- ----

CØNF LEVEL           LØWER LIM            UPPER LIM

  0.5                  2.755198             4.087561
  0.6                  2.648272             4.348012
  0.7                  2.533869             4.696124
  0.8                  2.402533             5.225544
  0.9                  2.231834             6.268296
  0.95                 2.102222             7.581048
  0.98                 1.969338             10.01902
  0.99                 1.887827             12.83937
   0.995               1.818288             17.35297
```

Rev. January 1971

FACTAN

This FORTRAN program analyzes intercorrelation within a set of variables through factor analysis. The significant factor loadings are obtained by a principle component analysis, and the rotations are performed using the varimax method. In this way, the minimum number of independent dimensions needed to account for most of the variance in the original set of variables is determined.

METHOD

1) Principle Component Analysis

   Given N sets of test scores with M tests, the MxM correlation matrix R is constructed, where the diagonal elements are defined as 1 (as suggested by Kaiser). The factor loadings, V_i, $i = 1, \ldots, M$, are the normalized eigenvectors of R. In accordance with Kaiser, those factor loadings V_i are considered significant whose corresponding eigenvalues, λ_i, are less than 1. The variance accounted for by factor loading v_i is proportioned to λ_i.

2) Varimax

   The initial factor loadings are rotated to obtain a simple structure as defined by Thurstone. The rotation is performed analytically by maximizing the varimax criterion:

$$V = \sum_i \left\{ \left[M \sum_j \left(b_{ij}^2 / h_i^2 \right)^2 - \left(\sum_i b_{ij}^2 / h_i^2 \right)^2 \right] / M^2 \right\}$$

where:

M = the number of variables (i. e. , number of tests)

h_i^2 = communality of test i

b_{ij} = the new factor loading for variable i on factor j ;

$i = 1, 2, \ldots M$

$j = 1, 2, \ldots P$

where P is the number of significant factor loadings.

INSTRUCTIONS

After compilation, the program requests:

INPUT NUMBER OF TESTS AND NUMBER OF SUBJECTS =

The response is the number of variables and the number of observations, if raw data is being input. When the correlation matrix is being input, the user types in the number of tests for both of these variables.

The program then prints:

ARE YOU INPUTTING RAW DATA =

If the response is YES, the routine requests:

INPUT TEST SCORES FOR SUBJECT i =

If the response is NO, the routine requests:

INPUT THE ELEMENTS ON AND ABOVE THE DIAGONAL IN YOUR CORRELATION MATRIX =

The routine outputs the significant factor loadings before rotations. If input was raw data, the correlation matrix is also printed.

The routine then requests:

INPUT ERROR CRITERION =

The errors criterion, E, is a value such that the iteration on the varimax criterion V stops when the difference between 2 successive values of V is less than E in absolute value, i. e.

$$\left| V_{ij} - V_{ij} + 1 \right| < E$$

The routine then outputs the varimax criterion at each iteration step and the significant factor loadings after rotation.

RESTRICTIONS

M must be less than or equal to 24.

This program calls the LIBRARY routine EIG1.

REFERENCE

H. H. Harman, Modern Factor Analysis,
University of Chicago Press, (1960).

SAMPLE PROBLEM

Perform a factor analysis on the data whose correlation matrix is

1.	.14	.987	.168	.931	.804	.597	.98
-	1.	.16	.93	.491	.693	.877	.331
-	-	1.	.185	.927	.807	.608	.927
-	-	-	1.	.489	.671	.835	.347
-	-	-	-	1.	.962	.848	.984
-	-	-	-	-	1.	.95	.903
-	-	-	-	-	-	1.	.743
-	-	-	-	-	-	-	1.

SAMPLE SOLUTION

```
*RUN FACTAN;EIG1
INPUT NUMBER ØF TEST AND NUMBER ØF SUBJECTS
= 8,8
ARE YØU INPUTTING RAW DATA
= NØ
INPUT THE ELEMENTS ØN AND ABØVE THE
DIAGØNAL IN YØUR CØRRELATIØN MATRIX

= 1. .14 .987 .168 .931 .804 .597 .98
= 1. .16 .93 .491 .693 .877 .331
= 1. .185 .927 .807 .608 .972
= 1. .489 .671 .835 .347
= 1. .962 .848 .984
= 1. .95 .903
= 1. .743
= 1.



 0.74433926E+00= VARIANCE ACCØUNTED FØR BY LØADING     1

FACTØR LØADING      1  IS
 0.85314583E+00  0.63343935E+00  0.85737874E+00  0.63304069E+00
 0.98346706E+00  0.99197998E+00  0.92607763E+00  0.93857161E+00



 0.24096077E+00= VARIANCE ACCØUNTED FØR BY LØADING     2

FACTØR LØADING      2  IS
-0.51819443E+00  0.76199648E+00 -0.49968376E+00  0.73550999E+00
-0.17583970E+00  0.77429643E-01  0.36622528E+00 -0.34179850E+00

 INPUT ERRØR CRITERIØN
= .0001


THE VARIMAX CRITERIØN =
   2.0124442E+01
```

```
THE VARIMAX CRITERIØN =
   2.0124440E+01


THE RØTATED LØADINGS ARE

 0.99650201E+00  0.88596644E-01  0.98947305E+00  0.10331124E+00
 0.90932857E+00  0.77249351E+00  0.55423267E+00  0.96663111E+00

 0.58025334E-01  0.98693191E+00  0.75664948E-01  0.96490526E+00
 0.41382185E+00  0.62711506E+00  0.82738550E+00  0.25172823E+00

PRØGRAM STØP AT 1020
*
```

FLAT (FLATSORC)

The file FLAT contains three FORTRAN compatible, GMAP coded routines for calculating random numbers having a uniform (rectangular) distribution. A combination mixed congruential[1] method is used. The file FLATSORC is the CARDIN listing for FLAT.

INSTRUCTIONS

The calling sequence is:

A = FLAT (B) or A = UNIFM1(B) or A = UNIFM2 (B, C, D)

where:

A is the computed random number.
B is an arbitrary starting number.
C is the mean.
D is the width of the interval.

FLAT and UNIFM1 are identical. They assume a mean of .5 and an interval width of 1. The starting number B is used to initialize the calculations of the random number. Subsequent calls to any of the routines use the previous calculated random number in place of B.

METHOD

The subprogram maintains, in storage: four multipliers, λ_i; four "old" random integers, N_i; and the current value of a random valued index, I (= 0, 1, 2, 3), which was constructed on the previous call to the subprogram. This current value of the index is used to choose one of the four congruential generators to produce the next random number, a revised value for N_I, and a revised value for the index.

The equations are:

$$N_i = (N_i\lambda_i) \text{ MOD } 2^{35}$$

$$R = N_i \text{ normalized to the proper interval}$$

$$I = [(N_i\lambda_i) \text{ MOD } 8 - (N_i\lambda_i) \text{ MOD } 2^{36}]/2^{36}$$

where:

$$\lambda_0 = (273673163155)_8$$

$$\lambda_1 = (74052161255)_8$$

$$\lambda_2 = (1050005)_8$$

$$\lambda_3 = (10405)_8$$

[1] G. M. Roe, "Design for a Random Number Generator," General Electric Company, Report No. 69-C-337, Nov. 1969.

The period of each component sequence is 2^{33}. The period of the shuffled sequence is a random variable with an exponential distribution function and a mean value of 2^{132}.

NOTE:

For other methods of evaluating UNIFM1 and UNIFM2, see the programs URAN and UNIFM.

SAMPLE PROBLEM

Calculate 10 random numbers with uniform distribution in the interval (0, 1) and 10 random numbers with uniform distribution in the interval (-1, 1). This sample demonstrates calls to FLAT, UNIFM1 and UNIFM2.

SAMPLE SOLUTION

```
*LIST

10 B=5.778
20 DO 10 I=1,10
30 C=FLAT(B)
40 D=UNIFM2(B,0.,2.)
50 E=UNIFM1(B)
60 10 PRINT 30,C,D,E
70 30 FØRMAT(3F12.8)
80 STØP;END

READY

*RUN *;FLAT
 0.13633348  0.31908965  0.72209641
 0.47026765 -0.98096177  0.63681064
 0.67571584  0.16795669  0.39364031
 0.65854639  0.49066935  0.09086038
 0.60899374  0.45487808  0.87867872
 0.23968677 -0.19357704  0.19482627
 0.55464186  0.58482587  0.87449226
 0.87252840  0.66247866  0.49738025
 0.17166802  0.44481332  0.14288863
 0.08899599  0.92853701  0.41821570

PRØGRAM STØP AT 80
*
```

FORIR

This FORTRAN program determines the least square estimates of the finite Fourier series model:

$$Y = \frac{A_o}{2} + \sum_{j=1}^{M} [A_j \text{COS}(jZ) + B_j \text{SIN}(jZ)]$$

If the independent data points are assumed to be:

$$Z = 0, \left(\frac{1}{2N}\right) 2\pi, \left(\frac{2}{2N}\right) 2\pi, \ldots, \left(\frac{2N-1}{2N}\right) 2\pi$$

or generally:

$$Z_j = \left[\frac{X_j - X_o}{X_{2N} - X_o}\right] 2\pi \qquad \text{For } j = 0, 1, 2, \ldots, 2N-1$$

then the model may be considered to be defined over the 2N equally spaced points X_o, X_1, X_2, ..., X_{2N-1}

The quantity A_j COS (jZ) + B_j SIN (jZ) is known as the jth harmonic. The quantity $A_j^2 + B_j^2$ is sometimes referred to as the power spectral density of the jth harmonic.

Therefore, this program can be used to investigate harmonics for equally spaced values of

$$Y_0, Y_1, Y_2, \ldots, Y_{2N-1}$$

INSTRUCTIONS

Data for this program can be entered via the teleprinter keyboard or from a previously created permfile. In either case, the free field format is used for input.

The program requests the total number of data points (2N) and the dependent data points when required.

After receiving the requested input, the program prints a four column table of output. These columns are defined as: the harmonic, the A_j terms, the B_j terms, and the error sum of squares removed. The fourth column is the quantity:

$$N \left(\frac{A_k^{\ 2}}{2} \right) \qquad \text{for } k = 0 \text{ and}$$

$$N \left(A_k^{\ 2} + B_k^{\ 2} \right) \qquad \text{for } k = 1, 2, 3, \ldots, N\text{-}1$$

which is proportional to the power spectral density of the jth harmonic. Also, this column has statistical significance since it is the sum of the squares removed by the A_j and B_j terms, which contribute two degrees of freedom for performing the standard analysis of variance tests.

The program permits the user to select any combination of harmonics to compute the following sum over the selected points:

$$Y = \underset{i}{\ } A_i \ \mathrm{COS} \left(i \frac{X}{2N} \right) 2 \ + B_i \ \mathrm{SIN} \left(i \frac{X}{2N} \right) 2$$

(and $Y = \frac{A_o}{2}$ if the zeroth harmonic is desired)

The user selects by typing answers to the questions DESIRED NUMBER OF HARMONICS TO TRY? and WHICH ONES?. After the user selects the desired combination of harmonics, the program requests the independent data point, X, by typing INPUT?. A predicted value is computed and printed when the user types the value of X. A response of **1.E35** to INPUT? permits the user to select a new combination of harmonics. A response of 0 (zero) to the question DESIRED NUMBER OF HARMONICS TO TRY? transfers control to the beginning of the program.

RESTRICTIONS

The maximum number of dependent data points is 100.

SAMPLE PROBLEM

Determine the finite Fourier Series model for the step function:

1, 1, 1, 1, 2, 2, 2, 2

Using all the harmonics generated, compute the predicted values for X = 1, 2, . . . 12. For the same values of X, compute the predicted values for harmonics 0 and 1.

SAMPLE SOLUTION

```
*RUN


FINITE FØURIER SERIES

IS DATA TØ BE READ FRØM A FILE, TYPE YES,NØ ØR STØP
= NØ

NUMBER, DATA PØINTS
= 8  1 1 1 1 2 2 2 2

      HARMØNIC    CØS TERMS          SINE TERMS        ERRØR SS REMØVED
          0    3.0000000E+00    0.                    1.8000000E+01
          1   -2.4999993E-01   -6.0355341E-01         1.7071067E+00
          2    2.4214387E-08   -1.8626451E-08         3.7331249E-15
          3   -2.4999997E-01   -1.0355345E-01         2.9289321E-01
          4   -2.3841858E-07    0.                    1.1368684E-13

DESIRED NUMBER ØF HARMØNICS TØ TRY
= 5
WHICH ØNES
= 0 1 2 3 4
INPUT
= 1
PREDICTED VALUE    1.0000000E+00
INPUT
= 2
PREDICTED VALUE    9.9999987E-01
INPUT
= 3
PREDICTED VALUE    9.9999999E-01
INPUT
= 4
PREDICTED VALUE    1.9999998E+00
INPUT
= 5
PREDICTED VALUE    2.0000001E+00
INPUT
= 6
PREDICTED VALUE    1.9999997E+00
INPUT
= 7
PREDICTED VALUE    2.0000002E+00
INPUT
= 8
PREDICTED VALUE    9.9999999E-01
INPUT
= 9
PREDICTED VALUE    1.0000000E+00
INPUT
= 10
PREDICTED VALUE    9.9999990E-01
INPUT
= 11
PREDICTED VALUE    9.9999995E-01
INPUT
= 12
PREDICTED VALUE    1.9999997E+00
INPUT
= 1E35
```

```
DESIRED NUMBER ØF HARMØNICS TØ TRY
= 2
WHICH ØNES
= 0 1
INPUT
= 1
PREDICTED VALUE   8.9644663E-01
INPUT
= 2
PREDICTED VALUE   8.9644658E-01
INPUT
= 3
PREDICTED VALUE   1.2499999E+00
INPUT
= 4
PREDICTED VALUE   1.7499999E+00
INPUT
= 5
PREDICTED VALUE   2.1035533E+00
INPUT
= 6
PREDICTED VALUE   2.1035534E+00
INPUT
= 7
PREDICTED VALUE   1.7500001E+00
INPUT
= 8
PREDICTED VALUE   1.2500001E+00
INPUT
= 9
PREDICTED VALUE   8.9644665E-01
INPUT
= 10
PREDICTED VALUE   8.9644658E-01
INPUT
= 11
PREDICTED VALUE   1.2499999E+00
INPUT
= 12
PREDICTED VALUE   1.7499999E+00
INPUT
= 1E35

DESIRED NUMBER ØF HARMØNICS TØ TRY
= 0

IS DATA TØ BE READ FRØM A FILE, TYPE YES,NØ ØR STØP
= STØP

PRØGRAM STØP AT 200
*
```

FOURIER

This BASIC program produces the fourier coefficients for a given periodic function of the form:

$$F(X) = A(0)/2 + \sum_{N=1}^{M} (A(N) * COS(N*X) + B(N) * SIN(N*X))$$

which approximates a set of data points. The data points can be supplied either as discrete points equally spaced from 0 to 2π or as a set of defining functions.

INSTRUCTIONS

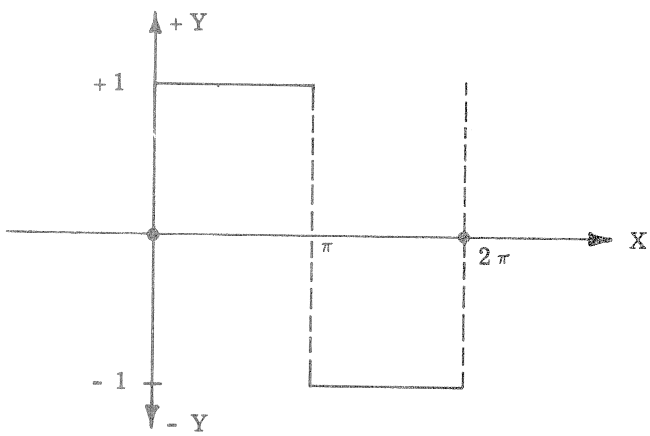
If the data to be approximated consists of discrete points, they are entered in data statements at line numbers 6000 to 6998. If the data consists of defining functions, use lines 2000 to 4000 to define Y as a function of X, for example:

```
*2000 IF X < 3.1416 THEN 2030
*2010 LET Y = X13.1416
*2020 GO TO 2040
*2030 LET Y = X13.1416-1.0
```

Additional instructions can be found by running the program.

SAMPLE PROBLEM 1

Find the 9th order periodic function which best approximates the square wave:



by using 20 discrete points.

SAMPLE SOLUTION 1

```
FØURIER

*RUN FØURIER

------------------------FØURIER SERIES PRØGRAM----------------------


DØ YØU DESIRE IINSTRUCTIØNS---1=YES,O=NØ ?1

INSTRUCTIØNS:

THIS PRØGRAM PRØDUCES THE FØURIER CØEFFICIENTS FØR A GIVEN PERIØDIC
FUNCTIØN ØF THE FØRM:

                M
                -
                \
   F(X)=A(O)/2+ /[A(N)*CØS(N*X)+B(N)*SIN(N*X)]
                -
               N=1

 GIVEN THE FØLLØWING DATA:

      1. A STATEMENT ØR SERIES ØF STATMENTS STARTING AT LINE NUMBER
         2000 THAT DEFINE 'Y' AS A FUNCTIØN ØF 'X' ØVER THE INTERVAL
         ZERØ TØ TWØ PI. FØR EXAMPLE:

           *2000 IF X < 3.1416 THEN 2030
           *2010 LET Y=X/3.1416
           *2020 GØ TØ 2040
           *2030 LET Y=X/3.1416-1.0
           *2040 REM
           *RUN

          ØR A SERIES ØF STATEMENTS ØF THE FØRM:
            *6000 DATA Y(1),Y(2),........Y(N)
            *RUN

           WHERE Y IS A VECTØR ØF TABULATED FUNCTIØN VALUES
           TAKEN AT EVEN INTERVALS FRØM ZERØ TØ TWØ PI.

      2. THE NUMBER ØF PØINTS TØ BE TAKEN ØVER THE INTERVAL
```

```
        (REQUESTED DURING PROGRAM EXECUTION)

     3. THE DESIRED ORDER OF THE FOURIER COEFFICIENTS
          (REQUESTED DURING PROGRAM EXECUTION)

NOTE: THE ORDER OF THE SERIES MUST BE GREATER THAN OR EQUAL
      TO ZERO

      THE NUMBER OF POINTS MUST BE GREATER THAN TWICE THE ORDER OF
      THE FOURIER SERIES

      B(0) IS ALLWAYS ZERO.



READY
*6000 DATA 0,1,1,1,1,1,1,1,1,1
*6010 DATA -1,-1,-1,-1,-1,-1,-1,-1,-1,0
*RUN

FOURIER


------------------------FOURIER SERIES PROGRAM------------------------



DO YOU DESIRE INSTRUCTIONS---1=YES,0=NO ?0
ARE YOU USING A FUNCTION TO SUPPLY DATA POINTS (TYPE '1') OR
A SET OF DATA POINTS (TYPE '2') ?2
HOW MANY DATA POINTS ?20
WHAT IS THE MAXIMUM ORDER OF THE HARMONICS TO BE FITTED ?9


FOURIER SERIES COEFFICIENTS:

     A(I)             B(I)                 I

       .2                 0                0
  -5.71346E-07      1.26275                1
       .2           1.90018E-07            2
  -4.70089E-07      .392522                3
       .2           2.76197E-07            4
  -2.96533E-07      .1999999               5
       .2           3.15528E-07            6
  -1.56479E-07      .1019051               7
       .2           2.57586E-07            8
  -4.02283E-10      .0316769               9
```

```
DØ YØU WISH TØ SEE A TABLE ØF PREDICTED  VS. ACTUAL VALUES
(1-YES, 0=NØ) ?1
```

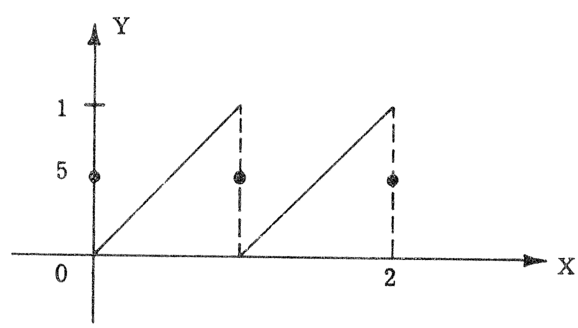
X	PRED-Y	ACTUAL-Y	ERRØR -
0	.8999985	0	.8999985
.3141593	1.1	1	.0999998
.6283186	.8999998	1	.1000002
.9424779	1.1	1	.0999997
1.256637	.8999998	1	.1000002
1.570796	1.1	1	.0999997
1.884956	.8999997	1	.1000003
2.199115	1.1	1	.0999998
2.513274	.8999996	1	.1000004
2.827434	1.1	1	.1
3.141593	.9000006	1	.0999994
3.455752	-.8999991	-1	.1000009
3.769911	-1.1	-1	.0999996
4.084071	-.8999999	-1	.1000001
4.39823	-1.1	-1	.0999998
4.712389	-.8999998	-1	.1000002
5.026549	-1.1	-1	.0999999
5.340708	-.8999997	-1	.1000003
5.654867	-1.1	-1	.1000002
5.969026	-.9000001	-1	.0999999
6.283186	.8999997	0	.8999997

```
TØTAL ERRØR=     3.699998
MEAN ERRØR=      .1849999

READY
```

SAMPLE PROBLEM 2

Find the 10th order periodic function which best approximates the saw tooth:



SAMPLE SOLUTION 2

```
*2000 IF X=0.0 THEN 2510
*2010 IF X=3.141593 THEN 2510
*2020 IF X=2.*3.141593 THEN 2510
*2030 IF X>3.141593 THEN 2540
*2040 LET Y=X/3.141593
*2050 RETURN
*2510 LET Y=.5
*2520 RETURN
*2540 LET Y=(X-3.141593)/3.141593
*2550 RETURN
*RUN


FØURIER


-----------------------FØURIER SERIEØ PRØGRAM------------------


DØ YØU DESIRE INSTRUCTIØNS---1=YES,0=NØ ?0
ARE YØU USING A FUNCTIØN TØ SUPPLY DATA PØINTS (TYPE '1') ØR
A SET ØF DATA PØINTS (TYPE '2') ?1
WHAT IS THE MAXIMUM ØRDER ØF THE HARMØNICS TØ BE FITTED ?10


FØURIER SERIES CØEFFICIENTS:

     A(I)              B(I)              I

    .9950789            0                0
   -.0048659          4.50002E-05        1
   -.0048587         -.318279            2
   -.0048718          .0001172           3
   -.0048672         -.159102            4
   -.0048733          .0001944           5
   -.0048694         -.1060252           6
   -.0048741          .0002722           7
   -.0048708         -.0794738           8
   -.0048745          .0003502           9
   -.0048716         -.0635325          10
```

DØ YØU WISH TØ SEE A TABLE ØF PREDICTED VS. ACTUAL VALUES
(1=YES, 0=NØ) ?1

X	PRED-Y	ACTUAL-Y	ERRØR
0	.4488421	.5	.0511579
.3141593	.0171184	.1	.0828816
.6283186	.2389015	.2	.0389015
.9424779	.2802065	.3	.0197935
1.256637	.4074454	.4	.0074455
1.570796	.5026014	.5	.0026014
1.884956	.587217	.6	.012783
2.199115	.7250958	.7	.0250959
2.513274	.7555218	.8	.0444782
2.827434	.9889879	.9	.088988
3.141593	.4975609	.9999999	.502439
3.455752	.0158364	.0999999	.0841635
3.769911	.2395453	.1999999	.0395454
4.084071	.2797097	.2999999	.0202902
4.39823	.407841	.3999999	.0078411
4.712389	.5022012	.4999999	.0022013
5.026549	.587618	.5999999	.0123819
5.340708	.7246103	.6999999	.0246104
5.654867	.7561849	.7999999	.043815
5.969026	.9877437	.8999999	.0877438
6.283186	.4488412	.9999999	.5511586

TØTAL ERRØR= 1.750316

MEAN ERRØR= .0833484

READY
*

KOKO

This FORTRAN program, given two independent samples, decides, using the Kolmogoroo-Smirnov (one-tailed) test, if the values of the population from which one of the samples was drawn are stochastically larger than the values of the population from which the other sample was drawn.

METHOD

Setting up a grid along the X axis and evaluating the cumulative distribution of both populations at each step of the grid we take the maximum positive deviation, D, between the two cumulative step functions.

To test the Null hypothesis that

H_0: Samples S_2 and S_1 are drawn from the same population (against the alternate hypothesis that the population from which S_2 was drawn is stochastically larger than population from which S_1 was drawn).

We consider the statistic -

$$Z = 4D^2 \frac{n_1 n_2}{n_1 + n_2}$$

n_1 is number of elements in S_1

n_2 is number of elements in S_2

which is approximated by the chi-square distribution with df = 2.

In using this routine the second sample is checked to see if its distribution is stochastically larger than that of the first sample.

INSTRUCTIONS

After compilation the program requests -

HOW MANY IN EACH SET =

The response to this is the number of points in sample one (n_1), followed by the number of points in sample 2 (n_2).

The routine then requests -

SEND X POINTS =

where input consists of the first sample, and

SEND Y POINTS =

where input consists of the second sample.

SCALING FACTOR =

and

CELL SIZE =

are next requested where CELL SIZE is the length of a step in your grid.

If either n_1 or n_2 is .GE. 40, output consists of the significance level at which the null hypothesis is rejected and the statistic D.

If both n_1 and n_2 are .LT. 40, the statistic K_D, which is sometimes used in this case, is also included in the output.

The program then requests -

NEW CELL SIZE =

which allows the user to try a smaller cell size. If the response to this is a positive integer the program will recompute the output quantities. If the response is 0, execution will be terminated.

RESTRICTIONS

1. n_1 and n_2 must be .LE. 1000.
2. The cell size must be a positive integer.
3. If MAX is the largest value in your input data for sample one and sample two, and MIN is the smallest value, the scaling factor, SF, must be such that SF*(MAX - MIN) is large enough to allow a grid with a sufficient number of cells to be constructed between SF*MAX and SF*MIN.
4. If both n_1 and n_2 are .LT. 40, the significance level is a conservative approximation. The actual level should be determined by locating D or K_D in a Kolmogorov-Smirnov table.
5. The subprogram XINGAM must be used with this program, as shown in the sample execution.

SAMPLE SOLUTION

```
*RUN KØKØ;XINGAM
HØW MANY IN EACH SET
= 54, 44


SEND X PØINTS
= 1 4 4 4 7 7 7 7 7 7 7 10 10 10 10 10 10 10 10 10 10 10
= 13 13 13 13 13 13 13 13 13 13 13 13
= 16 16 16 16 16 16 16 16 16 16 16 16 16 16
= 19 19 19 19 19 19 10


SEND Y PØINTS
= 1 1 1 1 1 1 1 1 1 1 1 4 4 4 4 4 4 4 7 7 7 7 7 7 7 7
= 10 10 10 13 13 13 13 13 16 16 16 16 16
= 19 19 19 19 19


SCALING FACTØR
= 1


CELL SIZE
= 3


THIS TEST IS SIGNIFICANT AT THE        0.34155697E-03 SIGNIFICANCE LEVEL


D=        0.40572390E+00


NEW CELL SIZE
= 2


THIS TEST IS SIGNIFICANT AT THE        0.34155697E-03 SIGNIFICANCE LEVEL


D=        0.40572390E+00


NEW CELL SIZE
= 0

PRØGRAM STØP AT 710
*
```

KRUWAL

This FORTRAN program performs a Kruskal-Wallis[1] one way analysis of variance by ranks. This technique is used for deciding whether k independent samples are from different populations. The Kruskal-Wallis method tests the null hypothesis -

H_o: The K samples come from the same source.

METHOD

Each of the observations in the entire set is replaced by a rank and a statistic H is calculated using the formula

$$H = \frac{12}{N(N+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j} - 3(N+1)$$

where,

k = number of samples

n_j = number of cases in j^{th} sample

$N = \Sigma n_j$; the number of cases in all samples combined

R_j = sum of ranks in j^{th} sample

H approximates a chi-square distribution with df = k-1 for sample sizes sufficiently large.

INSTRUCTIONS

Enter data as requested by the program. This routine calls the function XINGAM, which must be executed together with KRUWAL, for example

RUN KRUWAL; XINGAM

RESTRICTIONS

If the number of samples is less than 4 and the number of elements in each sample is less than 5, the significance level is a rough approximation. In this case the actual significance level should be found in a table for the H-statistic.

The maximum number of entries for any sample is 400. The sum of the numbers of entries for the k samples cannot exceed 1500.

[1] Siegel, Sidney, Nonparametric Statistics, McGraw-Hill Book Co. 1956.

SAMPLE SOLUTION

```
*RUN KRUWAL;XINGAM
INPUT NUMBER ØF SAMPLES TØ BE CØNSIDERED
= 3

INPUT THE NUMBER ØF ENTRIES IN
SAMPLE     1 FØLLØWED BY SAMPLE     1
= 5 96 128 83 61 101

INPUT THE NUMBER ØF ENTRIES IN
SAMPLE     2 FØLLØWED BY SAMPLE     2
= 5 82 124 132 135 109

INPUT THE NUMBER ØF ENTRIES IN
SAMPLE     3 FØLLØWED BY SAMPLE     3
= 4,115 149 166 147
YØUR ØRDERED DATA IS
    6.1000000E+01   8.2000000E+01   8.3000000E+01   9.6000000E+01
    1.0100000E+02   1.0900000E+02   1.1500000E+02   1.2400000E+02
    1.2800000E+02   1.3200000E+02   1.3500000E+02   1.4700000E+02
    1.4900000E+02   1.6600000E+02


THE NUMBER ØF SAMPLES =
            3


THE KRUSKAL-WALLIS H-STATISTIC=
    6.4057140E+00

THE NULL HYPØTHESIS CAN BE REJECTED AT THE
 0.40645912E-01 SIGNIFICANCE LEVEL


PRØGRAM STØP AT 990
*
```

LINEFIT

This FORTRAN subroutine finds the least squares line to describe a set of data points.

METHOD

The equation of the line is calculated from the following formulas:

$$\text{Slope} = \frac{N\Sigma xy - \Sigma x \Sigma y}{N\Sigma x^2 - (\Sigma x)^2}$$

$$\text{Y-intercept} = \frac{\Sigma y \Sigma x^2 - \Sigma x \Sigma xy}{N \Sigma x^2 - (\Sigma x)^2}$$

INSTRUCTIONS

The calling sequence for the entry LINEFIT is:

CALL LINEFIT (IN, LAST, X, Y, SLOPE, YCUT)

where,

- IN is the first location of the two arrays X and Y.
- LAST is the last location of the two arrays X and Y.
- X is the array of independent variables.
- Y is the array of dependent variables.
- SLOPE is returned as the slope of the least squares line.
- YCUT is the dependent variable intercept for the least squares line.

SAMPLE SOLUTION

```
*

10 DIMENSIØN X(10),Y(10)
20 PRINT:"NUMBER ØF PØINTS"
30 READ:LAST
40 PRINT:"LIST THE PAIRS"
50 READ:(X(I),Y(I),I=1,LAST)
60 CALL LINEFIT(1,LAST,X,Y,SLØPE,YCUT)
70 PRINT:"THE SLØPE ØF THE LEAST SQUARES LINE IS"
80 PRINT:SLØPE
90 PRINT:"THE DEPENDENT VARIABLE INTERCEPT FØR THE LEAST SQUARE LINE IS
100 PRINT:YCUT
110 STØP:END

READY

*RUN *:LINEFIT
NUMBER ØF PØINTS
= 7
LIST THE PAIRS
= 4,5
= 9,2
= 5.6,2
= 9,8
= 4,7
= 9,0
= 4,3
THE SLØPE ØF THE LEAST SQUARES LINE IS
  -2.8351297E-01
THE DEPENDENT VARIABLE INTERCEPT FØR THE LEAST SQUARE LINE IS
   5.6635251E+00

PRØGRAM STØP AT 110
*
```

LINREG

This BASIC program produces an equation that best fits input points described by **X and Y** coordinates. This fit is made in the least-squares sense.

INSTRUCTIONS

To use this linear regression program enter data using the following format:

1 DATA X [1], Y [1], X [2], Y [2],......., X [N], Y [N]
[Where X [1], Y [1] is the first point, X [2],
Y 2 is the second, and so on until all points have been entered. Additional data statements 2-199 may be used as needed.]

Then type RUN.

The program will request the type of curve fit desired (linear, exponential or power).

Additional instructions may be found in the listing.

SAMPLE PROBLEM

Determine what curve best fits the following bivariate data:

X	Y	X	Y	X	Y
1.0	38.10	3.0	11.78	5.0	5.30
1.5	14.67	3.5	1.67	5.5	1.67
2.0	12.76	4.0	5.35	6.0	8.92
2.5	13.15	4.5	14.60	6.5	15.67

Insert the following data. This may be done by either preparing a data tape or via the teleprinter.

1. DATA 1, 38. 1, 1. 5, 14. 67, 2, 12. 76, 2. 5, 13. 15, 3, 11. 78
2. DATA 3. 5, 1. 67, 4, 5. 35, 4. 5, 14. 6, 5, 5. 3, 5. 5, 1. 67, 6
3. DATA 8. 92, 6. 5, 15. 67

SAMPLE SOLUTION

```
*1 DATA 1,38.1, 1.5,14.67, 2,12.76, 2.5,13.15, 3,11.78
*2 DATA 3.5,1.67, 4,5.35, 4.5,14.6, 5,5.3, 5.5,1.67, 6
*3 DATA 8.92, 6.5,15.67
*RUN


WHAT TYPE ØF CURVE DØ YØU WANT TØ FIT:
LINEAR=1, EXPØNENTIAL=2, PØWER=3 ?1

LINEAR:  Y=A+B*X  WITH  A=   22.54395 AND B=-2.81972

CØEFFICIENT ØF CØRRELATIØN =    -.5272391
CØEFFICIENT ØF DETERMINATIØN =    .2779811

DØ YØU WANT TØ SEE A CØMPARISØN ØF THE ACTUAL Y'S AND
THE ESTIMATED Y'S. 1=YES, 0=NØ ?1

X-ACTUAL        Y-ACTUAL        Y-ESTIM         DIFFER          PCT-DIFF

         1          38.1        19.72423       -18.37577       -48.23037
       1.5         14.67        18.31437        3.644369        24.84232
         2         12.76        16.90451        4.144509        32.48048
       2.5         13.15        15.49465        2.344649        17.83003
         3         11.78        14.08479        2.304789        19.56527
       3.5          1.67        12.67493        11.00493        658.9778
         4          5.35        11.26507        5.915069         110.562
       4.5          14.6        9.855209       -4.744791       -32.49857
         5           5.3        8.445349        3.145349        59.34622
       5.5          1.67         7.03549         5.36549        321.2868
         6          8.92         5.62563        -3.29437        -36.9324
       6.5         15.67         4.21577       -11.45423       -73.09656

READY
*RUN


WHAT TYPE ØF CURVE DØ YØU WANT TØ FIT:
LINEAR=1, EXPØNENTIAL=2, PØWER=3 ?2

EXPØNENTIAL:  Y=A*EXP(B*X)  WITH  A=   19.64085 AND B=  -.2183258

CØEFFICIENT ØF CØRRELATIØN =    -.4255844
CØEFFICIENT ØF DETERMINATIØN =    .1811221

DØ YØU WANT TØ SEE A CØMPARISØN ØF THE ACTUAL Y'S AND
THE ESTIMATED Y'S. 1=YES, 0=NØ ?1

X-ACTUAL        Y-ACTUAL        Y-ESTIM         DIFFER          PCT-DIFF

         1          38.1        15.78856       -22.31144       -58.56021
       1.5         14.67        14.15577        -.514226        -3.50529
         2         12.76        12.69185       -.0681543       -.5341249
       2.5         13.15        11.37931        -1.77069       -13.46532
         3         11.78        10.20251       -1.577489       -13.39125
       3.5          1.67        9.147412        7.477412        447.7492
         4          5.35        8.201426        2.851426        53.29769
       4.5          14.6         7.35327       -7.246729       -49.63513
         5           5.3        6.592827        1.292827        24.39296
       5.5          1.67        5.911026        4.241026        253.9536
         6          8.92        5.299733       -3.620267       -40.58595
       6.5         15.67        4.751658       -10.91834       -69.67672
```

```
READY
*RUN


WHAT TYPE ØF CURVE DØ YØU WANT TØ FIT:
LINEAR=1, EXPØNENTIAL=2, PØWER=3 ?3

PØWER:  Y=A*(X↑B)   WITH  A=   23.52734 AND B=  -.8424159

CØEFFICIENT ØF CØRRELATIØN =    -.5332986
CØEFFICIENT ØF DETERMINATIØN =    .2844074

DØ YØU WANT TØ SEE A CØMPARISØN ØF THE ACTUAL Y'S AND
THE ESTIMATED Y'S. 1=YES, 0=NØ ?1
```

X-ACTUAL	Y-ACTUAL	Y-ESTIM	DIFFER	PCT-DIFF
1	38.1	23.52734	-14.57266	-38.3
1.5	14.67	16.71979	2.049786	14
2	12.76	13.12141	.3614054	2.8
2.5	13.15	10.87281	-2.277189	-17.4
3	11.78	9.324773	-2.455227	-20.9
3.5	1.67	8.189195	6.519195	390.4
4	5.35	7.317924	1.967924	36.8
4.5	14.6	6.626683	-7.973317	-54.7
5	5.3	6.063863	.7638626	14.4
5.5	1.67	5.596023	3.926023	235.1
6	8.92	5.200508	-3.719491	-41.8
6.5	15.67	4.861403	-10.8086	-69.1

```
READY
```

LSPCFP

This FORTRAN program fits least-square polynomials to bivariate data, using orthogonal polynomials. Limits are 10th degree fit and a maximum of 500 data points. The limit on degree of the fitted polynomial will vary depending on the number of data points and their size.

The program allows the user to specify the lowest degree polynomial to be fit and any higher degree polynomials. It also allows the user to see a complete residual summary.

INSTRUCTIONS

The user must prepare a data file, the first line of which must be a problem identification (up to 60 alphanumeric characters). The rest of the file should include the bivariate data with one set of points per line. Example:

```
*  10 SAMPLE PROBLEM

*  20 11.42 40.68

*  30 39.74 41.73

   .   .    .
   .   .    .
   .   .    .
```

The last line of the data file must be as follows

```
1.0E-5  1.0E-5
```

To use this program type:

```
RUN
```

The following message will appear:

```
ENTER DATA FILE 'NAME'
=
```

After the = sign, type the name of your data file. Then you will see

ENTER 'IP', 'LP', 'NR'

=

where,

IP = Lowest degree of polynomial to be calculated

LP = Highest degree of polynomial to be calculated

NR = Residual flag:

0 or blank = no residual output

1 = print residual output

These values should be typed after the = sign and separated by a comma or a blank.

SAMPLE PROBLEM

Fit the following data with a 4th degree polynomial

X	Y	X	Y	X	Y	X	Y
11.42	40.68	229.79	47.38	352.73	57.49	549.17	67.44
39.74	41.73	238.50	48.13	367.83	57.38	586.59	68.89
58.55	41.33	251.88	49.18	394.21	58.61	608.38	69.90
72.87	41.67	258.11	49.62	401.45	59.11	651.47	70.70
96.09	42.71	279.54	50.98	420.68	61.03	682.83	71.96
138.14	43.48	294.88	52.25	438.52	62.74	717.71	73.04
159.33	44.56	303.49	53.44	463.63	64.39	753.67	73.22
180.46	45.79	326.11	55.08	487.41	65.92	789.44	73.87
202.18	46.82	336.36	56.31	509.92	65.98		

SAMPLE SOLUTION

NOTE:

The input data was stored in the file LSQ DATA.

```
*LIST LSQDATA

10   S A M P L E   P R Ø B L E M
20   11.42   40.68
30   39.74   41.73
40   58.55   41.33
50   72.87   41.67
60   96.09   42.71
70   138.14  43.48
80   159.33  44.56
90   180.46  45.79
100  202.18  46.82
110  229.79  47.38
120  238.50  48.13
130  251.88  49.18
140  258.11  49.62
150  279.54  50.98
160  294.88  52.25
170  303.49  53.44
180  326.11  55.08
190  336.36  56.31
200  352.73  57.49
210  367.83  57.38
220  394.21  58.61
230  401.45  59.11
240  420.68  61.03
250  438.52  62.74
260  463.63  64.39
270  487.41  65.92
280  509.92  65.98
290  549.17  67.44
300  586.59  68.89
310  608.38  69.90
320  651.47  70.70
330  682.83  71.96
340  717.71  73.04
350  753.67  73.22
360  789.44  73.87
370  1.0E-5  1.0E-5

READY

*RUN
ENTER DATA FILE 'NAME'
= LSQDATA
ENTER 'IP', 'LP', 'NR'
= 4,4,0
```

```
          LEAST SQUARES PØLYNØMIAL CURVE FIT

S A M P L E   P R Ø B L E M

  NØ. ØBSERVATIØNS        35
  PØLY. DEGREE             4

          CØEFFICIENTS

    0        0.41917E+02
    1       -0.30363E-01
    2        0.36349E-03
    3       -0.53705E-06
    4        0.24074E-09

ERR**2 =  0.10207E+02           STD-ERR =  0.58330E+00

PRØGRAM STØP AT 1270
*
```

LSQMM

This FORTRAN subroutine approximates by

$$Y_i \approx f(X) \approx \sum_{j=1}^{n} A_j \varphi_j(X_i)$$

The set of data (Y_i, X_i), $i = 1, \ldots, M$ in either the weighted least squares or the min-max sense.[1]

INSTRUCTIONS

The calling sequence is:

CAL LSQMM(PHI, Y, A, RW, M, N, NT, NS, AM, IDIMM, IDIMN)

where,

PHI is the two dimensional array, PHI(M, N), of coordinate functions which are supplied by the user. The k^{th} column of PHI contains the k^{th} coordinate function evaluated at each of the data points.

(i. e., $PHI(J, I) = \varphi_i(X_j)$)

Y is the one dimensional array, Y(M), containing the dependent variables. A is the one dimensional array, A(N), containing the coefficients A(J) of the function F(X).

RW is the name of an array containing the residuals R(I) = Y(I) - F(I), the weights W(I), and temporary storage to save the vertical weights while doing the horizontal iterations. It contains at least 3*M locations.

M is the number of data points.

N is the number of coefficients, i.e., number of coordinate functions.

NT is the maximum number of vertical iterations. For least squares fit, NT = 1.

For least squares fit and when there is no division of the data points in min-max., NS = M. Otherwise, NS is the array containing the index values of the ends of the sections when using min-max. fit.

AM is a two dimensional array, AM(N, N), used internally to contain the matrix of the system of linear equations.

IDIMM is the first dimension of PHI, i.e., PHI(IDIMM, N).

IDIMN is the first dimension of AM, i.e., AM(IDIMN, N).

[1] M. A. Martin, Digital Filters for Data Processing, General Electric Company, Publication T. I. S. 62SD484, Appendix A.

F. E. Lilley, Notes on LSQMM Subroutine, General Electric Company, Publication PIR 5540-19.

METHOD

If the user wishes to minimize $\sum_{n=1}^{M} w_n \left[w_n - F(X_n) \right]^2$, $w_i > 0$, modify the PHI and Y arrays as follows:

$$PHI = \begin{bmatrix} \sqrt{w_1}\,PHI(1,1) & \sqrt{w_1}\,PHI(1,2) & \ldots & \sqrt{w_1}\,PHI(1,N) \\ \sqrt{w_2}\,PHI(2,1) & \ldots & \ldots & \ldots \\ \vdots & & & \\ \sqrt{w_M}\,PHI(M,1) & \ldots & \ldots & \sqrt{w_M}\,PHI(M,N) \end{bmatrix}$$

$$Y = \begin{bmatrix} \sqrt{w_1}\,(Y1) \\ \sqrt{w_2}\,Y(2) \\ \sqrt{w_3}\,Y(3) \\ \vdots \\ \sqrt{w_M}\,Y(M) \end{bmatrix}$$

RESTRICTIONS

The library subroutine program LINEQ must be used with this subroutine (see sample problem).

SAMPLE PROBLEM

Problem I:

Find the second degree polynomial F = A(3)*X**2 + A(2)*X + A(1), which best fits the following data in the least squares sense.

X = -4, , -3., -2., -1., 0., 1., 2., 3., 4.
Y = 2., -3., -6., -7., -6., -3., 2., 9., 18.

where M = 9, N = 3, NT = 1

SOLUTION: F = X**2 + 2.*X. -6.

Problem II:

Find the third degree polynomial F = A(4)*X**3 + A(3)*X**2 + A(2)*X + A(1), which best fits the following data in a min-max. sense with no division of data points.

X = 0., .1, .2, .3, .4, .5, .6, .7, .8, .9, 1.
Y = 1., 1., 1., 1., 1., 1., 0., 0., 0., 0., 0.

where M = 11, N = 4, NT = 6

SOLUTION: F = 11.11*X**3-18.33*X**2 + 6.39*X + .683

Problem III:

Find a function of the form F = A(3)*EXP(X + Z) + A(2)*SIN(X*Z) + A(1)*EXP(X) *SIN(Z), which best fits the following data in at least squares sense.

X	Y	Z
-4.	1.66698	1.
-3.	1.12278	2.
-2.	5.53939	3.
-1.	27.9934	4.
0.	151.4132	5.
1.	1098.34	6.
2.	8116.23	7.
3.	59930.7	8.
4.	442427.0	9.

where M = 9, N = 3, NT = 1

SOLUTION: F = 3.*EXP(X + Z) + 2.*SIN(X*Z) + EXP(X)*SIN(Z)

SAMPLE SOLUTION

The following main program was coded to solve these three problems:

```
*LIST MAIN

CØMMØN PHI(50,15),X(50),Z(50),A(15),RW(150)
CØMMØN AM(15,15),Y(50),YA(50)
1 PRINT:"M,N,NT"
READ:M,N,NT
NS=M
PRINT:"X ARRAY"
READ:(X(I),I=1,M)
PRINT:"Y ARRAY"
READ:(Y(I),I=1,M)
CALL PHI1(M,N)
CALL LSQMM(PHI,Y,A,RW,M,N,NT,NS,AM,50,15)
DØ 40 I=1,M
YA(I)=0.
DØ 30 J=1,N
30 YA(I)=YA(I)+A(J)*PHI(I,J)
40 CØNTINUE
PRINT 100
100 FØRMAT(5X,"X",10X,"F(X)",10X,"Y-F(X)",10X,"A(N)"/)
DØ 50 I=1,N
50 PRINT 60,X(I),YA(I),RW(I),A(I)
60 FØRMAT(4E13.4)
K=N+1
DØ 70 I=K,M
70 PRINT 60,X(I),YA(I),RW(I)
GØ TØ 1
END

READY

*
```

For problems I and II the routine PHI1 was coded to evaluate the **polynomial terms at the** x_i's.

```
*LIST PHI

10 SUBRØUTINE PHI1(M,N)
20 CØMMØN PHI(50,15),X(50)
30 DØ 10 I=1,M
40 10 PHI(I,1)=1.
50 IF(N-2)40,15,15
60 15 DØ 30 I=1,M
70 DØ 20 J=2,N
80 20 PHI(I,J)=X(I)**(J-1)
90 30 CØNTINUE
100 40 RETURN
110 END

READY

*
```

Solution to problem I and II

```
*RUN MAIN;PHI1;LSQMM;LINEO
M,N,NT
= 9,3,1
X ARRAY
= -4.,-3.,-2.,-1.,0.,1.,2.,3.,4.
Y ARRAY
= 2.,-3.,-6.,-7.,-6.,-3.,2.,9.,18.
       X              F(X)            Y-F(X)           A(N)
 -0.4000E+01     0.2000E+01      0.               -0.6000E+01
 -0.3000E+01    -0.3000E+01      0.                0.2000E+01
 -0.2000E+01    -0.6000E+01      0.                0.1000E+01
 -0.1000E+01    -0.7000E+01      0.
  0.            -0.6000E+01      0.
  0.1000E+01    -0.3000E+01      0.
  0.2000E+01     0.2000E+01      0.
  0.3000E+01     0.9000E+01      0.
  0.4000E+01     0.1800E+02      0.
M,N,NT
= 11,4,6
X ARRAY
=0.,.1,.2,.3,.4,.5,.6,.7,.8,.9,1.
Y ARRAY
= 1,1,1,1,1,1,0,0,0,0,0
       X              F(X)            Y-F(X)           A(N)
  0.             0.6833E+00      0.3167E+00        0.6833E+00
  0.1000E+00     0.1150E+01     -0.1500E+00        0.6389E+01
  0.2000E+00     0.1317E+01     -0.3167E+00       -0.1833E+02
  0.3000E+00     0.1250E+01     -0.2500E+00        0.1111E+02
  0.4000E+00     0.1017E+01     -0.1667E-01
  0.5000E+00     0.6833E+00      0.3167E+00
  0.6000E+00     0.3167E+00     -0.3167E+00
  0.7000E+00    -0.1667E-01      0.1667E-01
  0.8000E+00    -0.2500E+00      0.2500E+00
  0.9000E+00    -0.3167E+00      0.3167E+00
  0.1000E+01    -0.1500E+00      0.1500E+00
M,N,NT
= 
*
```

For problem III the routine PHI2 was coded to evaluate the φ functions.

```
*LIST PHI2

10 SUBROUTINE PHI1(M,N)
20 COMMON PHI(50,15),X(50),Z(50)
30 DO 5 I=1,9
40 5 Z(I)=I
50 DO 10 I=1,M
60 PHI(I,1)=EXP(X(I)+Z(I))
70 PHI(I,2)=SIN(X(I)*Z(I))
80 10 PHI(I,3)=EXP(X(I)*SIN(Z(I)))
90 RETURN
100 END

READY

*
```

Solution to problem III

```
*RUN MAIN;PHI2;LSQMM;LINEQ
M,N,NT
= 9,3,1
X ARRAY
= -4.,-3.,-2.,-1.,0.,1.,2.,3.,4.
Y ARRAY
= 1.66698, 1.12278, 5.53939, 27.9934, 151.4132
= 1098.34, 8116.23, 59930.7, 442427.0
     X             F(X)          Y-F(X)          A(N)
 -0.4000E+01    0.1667E+01    -0.5287E-04    0.1000E+01
 -0.3000E+01    0.1123E+01    -0.4596E-04    0.2000E+01
 -0.2000E+01    0.5540E+01    -0.4083E-03    0.3001E+01
 -0.1000E+01    0.2799E+02    -0.1241E-02
  0.            0.1514E+03    -0.4768E-03
  0.1000E+01    0.1098E+04    -0.3311E-02
  0.2000E+01    0.8116E+04     0.5493E-03
  0.3000E+01    0.5993E+05     0.
  0.4000E+01    0.4424E+06     0.1953E-01
M,N,NT
= X
*
```

MANDSD

This BASIC program calculates the mean, variance, and standard deviation for each of several sets of individual values or frequency distributions.

INSTRUCTIONS

To use this program enter data for each set of values as follows:

```
1 DATA  N, X[1],  X[2],  X[3],....., X[N]
```

Where the N values of the set are X [1] thru X [N]. If needed, additional data statements may be used to give the entire list of values. Additional cases may be given in subsequent data statements in the same format.

The input for grouped values has the following format:

```
1 DATA 0, N,  X[1],  F[1],  X[2],  F[2],...,  X[N],  F[N]
```

Where the initial zero signals grouped data, the N is the number of different values to be given, and the F [I] are the number of times the X [I] occur. Data statements following may be used to extend the list as necessary, and blocks of grouped data may be intermixed freely with straight lists described above.

As an example, suppose we were interested in the mean and standard deviation of the numbers 1, 5, 4, 2, 6, 7, 4, 7 and also for the distribution consisting of 5-1'S, 3-4'S, 6-7'S, and 2-11'S. These two cases could be run by typing the following:

```
1 DATA   8,1,5,4,2
2 DATA   6,7,4,7
3 DATA   0,4,1,5,4,3
4 DATA   7,6,11,2
RUN
```

Or equivalently:

```
1 DATA 8,1,5,4,2,6,7,4,7,0,4,1,5,4,3,7,6,11,2
RUN
```

NOTE:

Statement numbers 1 thru 299 are available for continuation of the data field. Additional instructions may be found in the listing.

SAMPLE PROBLEM

Determine the statistical characteristics of the following data points:

261.4	252.1	255.5	258.3	253.2
270.8	268.3	249.6	256.3	266.4
265.4	250.3	280.9	259.3	
261.4	272.3	270.3	270.1	
258.1	262.8	263.2	259.3	

To use this program to obtain the data's statistical characteristics, prepare the following data tape:

```
1 DATA    22, 261.4, 270.8, 265.4, 261.4, 258.1, 252.1, 268.3, 250.3
2 DATA    272.3, 262.8, 255.5, 249.6, 280.9, 270.3, 263.2, 258.3, 256.3
3 DATA    259.3, 270.1, 259.3, 253.2, 266.4
```

The first data value (22) represents the total number of values.

SAMPLE SOLUTION

```
*1 DATA 22, 261.4, 270.8, 265.4, 261.4, 258.1, 252.1, 268.3, 250.3
*2 DATA     272.3, 262.8, 255.5, 249.6, 280.9, 270.3, 263.2, 258.3
*3 DATA     256.3, 259.3, 270.1, 259.3, 253.2, 266.4
*RUN


MANDSD



ARITHMETIC MEAN, VARIANCE, AND STANDARD DEVIATION


INDIVIDUAL SET NUMBER     1

SAMPLE VALUES:

   261.4     270.8     265.4     261.4     258.1     252.1     268.3     250.3
   272.3     262.8     255.5     249.6     280.9     270.3     263.2     258.3
   256.3     259.3     270.1     259.3     253.2     266.4

MAXIMUM LIKELIHOOD ESTIMATES
OF POPULATION PARAMETERS

       NUMBER OF VALUES =     22
        ARITHMETIC MEAN =     262.0591
     STANDARD DEVIATION =     7.783971
        SAMPLE VARIANCE =   60.5902

UNBIASED ESTIMATES
OF POPULATION PARAMETERS

        ARITHMETIC MEAN =     262.0591
     STANDARD DEVIATION =     7.967148
               VARIANCE =     63.47545



READY
*
```

SAMPLE PROBLEM

Assume that each data point has a frequency associated with it. See the following printout for the frequencies used.

To use this program prepare the following data tape:

```
1 DATA     0, 22, 261.4, 3, 270.8, 4, 265.4, 2, 261.4, 7, 258.1, 3
2 DATA     252.1, 5, 268.3, 7, 250.3, 1, 272.3, 9, 262.8, 2, 255.5, 8
3 DATA     249.6, 5, 280.9, 9, 270.3, 3, 263.2, 8, 258.3, 2, 256.3, 3
4 DATA     259.3, 5, 270.1, 3, 259.3, 2, 253.2, 5, 266.4, 6
```

The first data value (0) indicates grouped data. The second data value (22) indicates the number of groups. The frequency associated with each data value is punched directly after that data value.

SAMPLE SOLUTION

```
*1 DATA 0,22, 261.4,3, 270.8,4, 265.4,2, 261.4,7, 258.1,3
*2 DATA       252.1,5, 268.3,7, 250.3,1, 272.3,9, 262.8,2, 255.5,8
*3 DATA       249.6,5, 280.9,9, 270.3,3, 263.2,8, 258.3,2, 256.3,3
*4 DATA       259.3,5, 270.1,3, 259.3,2, 253.2,5, 266.4,6
*RUN


MANDSD

ARITHMETIC MEAN, VARIANCE, AND STANDARD DEVIATIØN


FØR GRØUPED DATA SET            1

      X-VALUE          FREQUENCY

       261.4               3
       270.8               4
       265.4               2
       261.4               7
       258.1               3
       252.1               5
       268.3               7
       250.3               1
       272.3               9
       262.8               2
       255.5               8
       249.6               5
       280.9               9
       270.3               3
       263.2               8
       258.3               2
       256.3               3
       259.3               5
       270.1               3
       259.3               2
       253.2               5
       266.4               6

MAXIMUM LIKELIHØØD ESTIMATES
ØF PØPULATIØN PARAMETERS

      NUMBER ØF VALUES =   102
       ARITHMETIC MEAN =    263.5235
    STANDARD DEVIATIØN =    8.530582
       SAMPLE VARIANCE =    72.77083

UNBIASED ESTIMATES
ØF PØPULATIØN PARAMETERS

       ARITHMETIC MEAN =    263.5235
    STANDARD DEVIATIØN =    8.572709
              VARIANCE =    73.49134


READY
*
```

MREG1

This FORTRAN program fits the model

$$Y-\overline{Y} = (X_1 - \overline{X}_1)\,\beta 1 + (X_2 - \overline{X}_2)\,\beta_2 + \ldots + (X_m - \overline{X}_m)\,\beta_m$$

To a series of n discrete observations where: X_1, X_2, X_3, ..., X_m represents the independent variables; Y represents an observed data point; $\overline{X}_1$, $\overline{X}_2$, $\overline{X}_3$, ..., $\overline{X}_m$ represent the mean value of the independent variables; and Y represents the mean value of the observed data points.

INSTRUCTIONS

Data for this program can be entered via the teleprinter keyboard or by using the FORTRAN disk file capability (see sample problem). In either case, the free field format is used for input.

Instructions for using this program can be listed by typing YES in answer to the question DO YOU DESIRE USER INSTRUCTIONS? (see output of sample solution). After the data in matrix form has been entered, the program allows the user to make corrections to this data and provides the necessary instructions for the corrections.

After the corrections, if any, are made, the values for $\overline{Y}$ and $\overline{X}_1$, $\overline{X}_2$, $\overline{X}_3$, ..., $\overline{X}_m$ are generated. The program permits the user to select which variates (X_1, X_2, X_3, ..., X_m) are to be considered by typing answers to the questions NUMBER OF VARIATES TO BE CONSIDERED? and WHICH ONES?. The user has the option of printing or suppressing the calculated values of Y.

The program generates a standard of analysis which includes the sum of squares removed, the least squares estimates of the coefficients (1, 2, 3, ..., m), the coefficient 'T' confidence band width associated with each coefficient estimate (obtained by multiplying the least square coefficient by a 'T' value with n-m-1 degrees of freedom), and the 'F' ratio estimate (by comparing this with an F table, usually found in the appendices of a statistical text, the statistical significance of the regression can be determined.)

The program prints the upper right triangular portion of the inverse of the sums and cross-products matrix which provides the variances-covariances of the least square coefficients when multiplied by the mean error sum of squares. The first row of values is associated with $(_1)$, $(_{1,2})$, ..., $(_{1,\ m})$, the second row is associated with $(_2)$, $(_{2,\ 3})$, ..., $(_{2,\ m})$, and so forth until the last row in which the element is associated with $(_m)$.

The program permits the user to choose a new combination of variates (X_1, X_2, X_3, ..., X_m) and restart the program from that point. The program may be terminated by responding 0 to the question: Number of variates to consider.

RESTRICTIONS

1. The number of observed data points (n) must be less than or equal to 90.

2. The independent data matrix (n, m) must not have n greater than 90 or m greater than 10.

3. The independent data (X_1, X_2, X_3, . . . , X_m) may be calculated data from any type of function, provided the columns of the independent data matrix are not equal or proportional. (If the generated output contains abnormally large numbers, i. e. , numbers with a magnitude of 10^{76} or greater, the columns of the independent data matrix were probably equal or proportional.)

SAMPLE PROBLEM

Run this program for the following set of data:

Y	X	
7.0	1.0	2.0
4.0	3.0	0.4
6.0	-1.0	-0.5
6.5	-0.5	-2.0

Use both X variates to calculate the standard analysis of variance and print the calculated values of Y.

SAMPLE SOLUTION

```
*RUN
DØ YØU WANT INSTRUCTIØNS, TYPE YES ØR NO
= YES


MULTIPLE REGRESSION IS PRIMARILY AN ATTEMPT TØ CURVE FIT ØBSERVED
DATA BY THE MØDEL

          Y-YMEAN=A(1)(X(1)-X1MEAN)+...+A(P)(X(P)-XPMEAN)

WHERE

YMEAN=MEAN ØF ØBSERVED DATA.
XIMEAN=MEAN ØF EACH INDEPENDENT VARIABLE.
THE A(I) ARE UNKNØWN COEFFICIENTS
DETERMINED BY THE LEAST SQUARES PRØCESS.

THE DATA IS ENTERED IN MATRIX FØRM

          Y1        X11       X12...X1M
          Y2        X21       X22...X2M
          .         .         .      .
          YN        XN1       XN2...XNM

Y1 IS THE ØBSERVED DATA WHILE X11,X12,X1M ARE THE
CORRESPONDING INDEPENDENT VARIABLES.

FØR THE DATAFILE ØPTIØN BUILD THE FILE WITH LINE NUMBERS IN
THE SAME SEQUENCE AS REQUESTED IN TERMINAL ØPTIØN
```

```
RESTRICTIONS:
         M MUST BE .LE. 10 AND N MUST BE .LE. 90.
         READ IS TYPED IN THE FREE FIELD FORMAT,
         A BLANK OR COMMA ENDS THE FIELD.

NOW YOU TRY IT.

IS DATA READ FROM A FILE, TYPE YES OR NO
= YES
NAME OF FILE IS
= FILE2
  7.000000     1.000000     2.000000
  4.000000     3.000000     0.400000
  6.000000    -1.000000    -0.500000
  6.500000    -0.500000    -2.000000

ARE ANY OF THE ABOVE Y(N),X(N,M) ELEMENTS TYPED INCORRECTLY?
TYPE YES OR NO
ANY CORRECTIONS
= NO


                    MULTIPLE  REGRESSION  PROGRAM

                        YMEAN          XMEAN
                    5.875000E+00
                                    6.250000E-01
                                   -2.500001E-02
NUMBER OF VARIATES TO CONSIDER
= 2
WHICH ONES
= 1,2
WANT TO SEE PREDICTED VALUES, TYPE YES OR NO
= YES


                     CALCULATED      OBSERVED
                    6.477283E+00   7.000000E+00
                    4.344426E+00   4.000000E+00
                    6.842831E+00   6.000000E+00
                    5.835460E+00   6.500000E+00


   ................................................................
   .            .DEGREE OF FREE.  SUM OF SQUARES . VARIANCE ESTIMATE.
   . RECRESSION.       2       .  3.643661E+00  .    1.821831E+00   .
   . REMAINDER .       1       .  1.543839E+00  .    1.543839E+00   .
   . TOTAL     .       3       .  5.187500E+00  .                   .
   ................................................................


   LEAST SOUARE COEFFICIENTS        COEFFICIENT T CONFIDENCE BAND
          -7.215877E-01                       4.699703E-01
           4.310510E-01                       5.044787E-01


F RATIO(  2,  1  DEGREES OF FREEDOM)=  1.180065E+00
```

```
          VARIANCE-COVARIANCE-MATRIX
 1.430668E-01
-8.104141E-02
 1.648480E-01
NUMBER ØF VARIATES TØ CØNSIDER
= 0

PRØGRAM STØP AT 2045
*
```

```
*LIST FILE2

10 4
20 2
30 7,1,2
40 4,3,.4
50 6,-1,-.5
60 6.5,-.5,-2

READY

*
```

MULFIT

This BASIC program provides a statistical analysis for up to 61 observations on up to 6 variables.

INSTRUCTIONS

To use this program enter the input data using line number 0-699. All values for 1 variable are given in data statements, followed by all data for a second variable, etc. One convenient way to arrange the input variables is as follows:

LINES	VARIABLE
100-199	First
200-299	Second
300-399	Third
400-499	Fourth
500-599	Fifth
600-699	Sixth

If no transformed variables are given a linear function is fitted for the first data variable in terms of the subsequent variables. By using transforms, any function of any or all of the input variables may be used as any variable in the analysis. Transforms are entered before the run, at lines 1000-1099, in the following way:

```
    10XX  LET X(J)=ANY FUNCTIØN ØF V(1), V(2),....., V(R)

WHERE X(J) IS THE J-TH VARIABLE IN THE ANALYSIS (J=1 IS THE
DEPENDENT VARIABLE), AND THE V(I) ARE THE INPUT DATA VARIABLES.
THUS, ØNE MIGHT ENTER DATA FØR THREE VARIABLES V(1), V(2),
AND V(3), BUT RUN THE ANALYSIS WITH X(1)=V(2) AS THE DEPENDENT
VARIABLE, AND THREE 'INDEPENDENT' VARIABLES, BY ENTERING
TRANSFØRMATIØNS LIKE THE FØLLØWING AT LINES 1000-1003:

    1000 LET X(1)=V(2)
    1001 LET X(2)=LØG(V(1)+V(3))
    1002 LET X(3)=V(1)+X(2)*EXP(-1-V(3))
    1003 LET X(4)=(V(3)-1)↑2
```

Additional instructions may be obtained by listing STADES.

SAMPLE PROBLEM

Secure a statistical analysis for the observed data consisting of 30 data values for each of 4 variables.

NOTE:

Commas are not needed after each line because this program is written in BASIC.

The data for the first variable is entered on line 100, the second variable on 200, etc.

The run fits the equation Y = A + B·X1 + C·X2 + D·X3. The coefficients for the fitted equation are A = .936716, B = -.00142542, C = -.0040351 and D = .193222.

SAMPLE SOLUTION

```
100 DATA .43,.55,.55,.55,.54,.52,.62,.60,.50,.55,.52,.53,.40
101 DATA .46,.57,.59,.49,.45,.48,.47,.46,.61,.59,.60,.57
102 DATA .64,.62,.54,.66,.64
200 DATA 75,66,67,70,64,65,67,63,70,65,73,73,88,75,90,77
201 DATA 72,70,63,66,67,57,55,51,53,58,51,51,54,54
300 DATA 9.4,7.8,8.1,8.3,8.6,8.0,8.8,8.7,8.7,8.9,8.8,8.0,8.,8.7
301 DATA 8.6,8.7,8.1,8.5,8.6,8.3,8.9,8.9,8.6,8.7,8.4,1.2,8.4
302 DATA 8.3,7.8,8.5
400 DATA 2.04,1.36,1.36,1.36,1.36,1.36,1.36,1.36,1.59,1.59,1.59
401 DATA 1.59,1.59,1.59,1.13,1.13,1.59,1.59,1.59,1.59,1.59,1.02
402 DATA 1.02,1.02,1.13,1.13,1.13,1.36,1.13,1.13

READY

*RUN


MULFIT


DØ YØU WANT INSTRUCTIØNS (0=NØ, 1=YES).... WHICH  ?1
N=# VALUES, R=# IN VARIABLES, S=# ØUT VARIABLES, D IS A CØDE
VALUE (D: 1 = TRANSFØRMS HAVE BEEN SPECIFIED, 0 = ØTHERWISE.)
 N, R, S, D =  ?30,4,0,0

      M U L T I V A R I A T E   C U R V E   F I T


VARIABLE        REGR CØEFF     MEAN VALUE      STD DEV

 1 (CØNSTANT =    .9367094 )       .5433333         .0663996

         2        -.0014253        65.66666         9.910756
         3        -.0040348        8.243333         1.356636
         4        -.1932227        1.379333          .241536


   STANDARD DEVIATIØN ØF RESIDUALS =     .0343456
     INDEX ØF DETERMINATIØN (R-SQ) =     .7324458
       ZERØ-CHECK ØN MEAN RESIDUAL = -9.95894E-08
```

A C T U A L   V S   C A L C U L A T E D

ACTUAL	CALCULATED	DIFFERENCE	PCT DIFFER
.43	.3977073	-.0322927	-8.1
.55	.5483825	-.0016175	-.2
.55	.5457467	-.0042533	-.7
.55	.5406637	-.0093363	-1.7
.54	.5480054	.0080054	1.4
.52	.5490009	.0290009	5.2
.62	.5429224	-.0770776	-14.1
.6	.5490272	-.0509728	-9.2
.5	.4946086	-.0053914	-1
.55	.5009284	-.0490716	-9.7
.52	.4899291	-.0300709	-6.1
.53	.4931569	-.0368431	-7.4
.4	.4717768	.0717768	15.2
.46	.4874819	.0274819	5.6
.57	.5553876	-.0146124	-2.6
.59	.5735136	-.0164864	-2.8
.49	.4941788	.0041788	.8
.45	.4954156	.0454156	9.1
.48	.5049895	.0249895	4.9
.47	.5019239	.0319239	6.3
.46	.4980777	.0380777	7.6
.61	.622468	.012468	2
.59	.6265292	.0365292	5.8
.6	.6318271	.0318271	5
.57	.6089323	.0389323	6.3
.64	.6308561	-.0091439	-1.4
.62	.611783	-.008217	-1.3
.54	.5677453	.0277453	4.8
.66	.6099279	-.0500721	-8.2
.64	.6071035	-.0328965	-5.4

READY
*

ONEWAY

This BASIC program performs a one way analysis of variance with equal sample sizes.

INSTRUCTIONS

To use this program enter data using the following format:

```
100   DATA  M,N
200   DATA  X11,X12,X13,...,X1N
201   DATA  X21,X22,X23,...,X2N
 ETC UNTIL
2XX   DATA  XM1,XM2,XM3,...,XMN
```

where:

```
  M = number of samples
  N = common sample size
XIJ = observations
```

Additional instructions maybe found in the listing.

SAMPLE PROBLEM

Determine what the probability is that the variability in the observed data is due to pure chance.

As Manager of the Materials Laboratory, you have been following with interest the preparation of tensile test specimens on three rather exotic alloys being developed for high-temperature use. The materials are difficult to fabricate, but the men have finally succeeded in getting four test bars ready for stretching in the tensile test machines. Since these alloys differ only in the relative amounts of a single ingredient, the main purpose for the test is to discover the effect of the ingredient on the tensile strength, if indeed it has any at all. Results of the twelve tests are on the following page.

(Tensile Values in 1000 lbs./sq. inch)

Alloy A	Alloy B	Alloy C
181.4	166.3	176.4
168.6	172.1	183.2
174.0	179.2	189.4
183.2	159.7	177.5

What's the probability that the observed variability is due to pure chance rather than the amount of the added ingredient?

SAMPLE SOLUTION

```
*100 DATA 3,4
*200 DATA 181.4,168.6,174,183.2
*201 DATA 166.3,172.1,179.2,159.7
*202 DATA 176.4,183.2,189.4,177.5
*RUN


A N A L Y S I S   O F   V A R I A N C E
          3 SAMPLES OF SIZE      4

SAMPLE            SAMPLE TOTAL   SAMPLE MEAN

            1          707.2          176.8
            2          677.3        169.325
            3          726.5        181.625

MEAN SQUARE (BETWEEN SAMPLES) =    153.6328
MEAN SQUARE (WITHIN SAMPLES)  =    50.16146
CALCULATED VALUE OF F-RATIO   =    3.062766
CORRESPONDING NORMAL DEVIATE  =    1.305796

 THE PROBABILITY OF AN F-RATIO THIS LARGE
 OCCURRING BY CHANCE ALONE IS   0.096


READY
*
```

ANALYSIS OF RESULT

The evidence as to the effect of varying amounts of the alloying ingredient is not too clear. If the ingredient had no effect at all, and we ran ten experiments like this, we'd expect one of the ten to show as much diversity of sample means as we actually found. Maybe the test method should be looked at for improvement; there's quite a bit of variability among specimens of the same alloy.

ORPOL

This FORTRAN program determines the least squares coefficients of the model

$$Y_i = \sum_j C_j \varphi_j (X_i)$$

to fit this expression over the discrete set of data points X_i where $i = 1, 2, 3, \ldots, N$. The φ_j (X) are the jth degree polynomials which satisfy the criterion

$$\sum_X \rho(X) \varphi_j (X) \varphi_k (X) = 0 \quad \text{For } j \neq k$$

where ρ (X) are weights associated with each one of the data points X_i, where $i = 1, 2, 3, \ldots, N$.

INSTRUCTIONS

Data for this program can be entered via the teletypewriter keyboard or from a file (see Sample Solution). In either case, the free field format is used for input.

The following data is requested by the program whenever required:

1. Number of points, maximum degree
2. The dependent data points
3. The independent data points
4. The weights

The output generated at this point is as follows:

1. The Dependent Data Mean which is equal to

$$\frac{\sum_{i=1}^{N} \rho (X_i) Y_i}{N}$$

2. A table of five columns defined as: degree of the polynomial (j), ALPHA (α), BETA (β), COEFF, and SSR where:

   a. j, α, and β refer to each orthogonal polynomial. The jth polynomial is defined over the X data point by

$$\varphi_j (X) = X \varphi_{j-1} (X) - \alpha_j \varphi_{j-1} (X) - \beta_j \varphi_{j-2} (x)$$

where:

$$\varphi_0 (X) = 1$$
$$\varphi_1 (X) = X - \alpha_1$$
$$\varphi_2 (X) = X (X - \alpha_1) - \alpha_2 (X - \alpha_1) - \beta_2$$

The columns of α_j and β_j can therefore generate the value of these polynomials recursively.

b. The COEFF column corresponds to the C_j associated with each $\varphi_j (X)$.

c. The SSR column describes the total sum of the squares removed by the particular coefficient. The larger the SSR value, the more significant is the corresponding coefficient.

The program will then request the number of polynomials the user determines to be important and the degree of these polynomials by asking NUMBER and POLYNOMIALS. Following this the program will output REGULAR POLYNOMIAL IN DECREASING DEGREE. These values, read row-wise, correspond to the coefficients of the polynomials to which the data has been fit. The program will then request a set of independent data points for which a predicted value is to be computed and the number of such points by typing NUMBER, SET OF VALUES. For the sample solution where 5 polynomials (2, 4, 6, 8, and 10th degree) were selected, the predicted value for each of the independent data points, X, was computed as follows:

$$Y = C_2 \varphi_2 (X) + C_4 \varphi_4 (X) + C_6 \varphi_6 (X) + C_8 \varphi_8 (X) + C_{10} \varphi_{10} (X)$$

Output will consist of INPUT values and corresponding output from the orthogonal polynomials (OR OUTPUT) and the regular polynomials (REG OUTPUT). The routine now types FINISHED, YES OR NO. If NO, the routine returns to the request NUMBER AND POLYNOMIALS. If YES, the program returns to the request IS DATA TO BE READ FROM A FILE, TYPE YES, NO OR STOP.

RESTRICTIONS

The maximum degree is 30.

The maximum number of data points is 100.

The weights must be positive.

The maximum degree specified must be less than the number of data points.

REFERENCE

Hamming, Numerical Methods for Scientists and Engineers, McGraw-Hill, 1962, pp. 223-246.

SAMPLE PROBLEM

Curve fit the following data to orthogonal polynomials of maximum degree 10 and, after determining the significant polynomials, find the predicted values for X = 1, 4, 6, and 4.5:

X	Y	WEIGHTS
1	1	1
2	2	1
3	3	1
4	4	1
5	5	1
6	6	1
7	5	1
8	4	1
9	3	1
10	2	1
11	1	1

SAMPLE SOLUTION

This sample solution illustrates both methods of input, first from the terminal and then from the data file F1 which had been previously created.

```
ØRTHØGØNAL PØLYNØMIAL CURVE-FITTING


IS DATA TØ BE READ FRØM A FILE, TYPE YES,NØ ØR STØP
= NØ

NUMBER ØF PØINTS, MAX DEGREE
= 11 10
TYPE IN DEPENDENT DATA
= 1 2 3 4 5 6 5 4 3 2 1
TYPE IN INDEPENDENT DATA
= 1 2 3 4 5 6 7 8 9 10 11
TYPE IN WEIGHTS
= 1 1 1 1 1 1 1 1 1 1 1
DEPENDENT DATA MEAN =   3.272727E+00

DEGREE        ALPHA           BETA            CØEFF           SSR
   1      6.000000E+00   0.              0.              0.
   2      6.000000E+00   1.000000E+01   -1.748252E-01   2.622378E+01
   3      6.000000E+00   7.800000E+00   -4.631286E-10   1.325021E-15
   4      6.000000E+00   7.200000E+00    5.827505E-03   1.398601E+00
   5      6.000000E+00   6.666666E+00    5.043470E-10   6.348973E-14
   6      6.000000E+00   6.060607E+00   -4.901964E-04   3.208560E-01
   7      6.000000E+00   5.349650E+00    9.790116E-11   5.788681E-14
   8      6.000000E+00   4.523076E+00    7.832094E-05   1.324996E-01
   9      6.000000E+00   3.576472E+00    2.103909E-10   2.397700E-12
  10      6.000000E+00   2.507738E+00   -3.857962E-05   1.050824E-01

NUMBER AND PØLYNØMIALS
= 5 2 4 6 8 10
```

```
REGULAR PØLYNØMIAL IN DECREASING DEGREE
   -3.8579616E-05    2.3147770E-03    -6.0349540E-02     8.9681065E-01
   -8.3742828E+00    5.1078316E+01    -2.0485381E+02     5.2956509E+02
   -8.3669358E+02    7.2143513E+02    -2.5199560E+02

NUMBER, SET ØF VALUES
= 4 1 4 6 4.5

       INPUT             ØR ØUTPUT         REG ØUTPUT
    1.0000000E+00     9.9999980E-01     9.9999619E-01
    4.0000000E+00     3.9999982E+00     3.9996986E+00
    6.0000000E+00     5.9999970E+00     5.9990883E+00
    4.5000000E+00     4.3513668E+00     4.3500404E+00

FINISHED, YES ØR NØ
= NØ


NUMBER AND PØLYNØMIALS
= 2 2 4

REGULAR PØLYNØMIAL IN DECREASING DEGREE
    0.               0.               0.                0.
    0.               0.               5.8275053E-03    -1.3986013E-01
    9.3822833E-01   -1.1888109E+00    1.4545450E+00

NUMBER, SET ØF VALUES
= 4 1 4 6 4.5

       INPUT             ØR ØUTPUT         REG ØUTPUT
    1.0000000E+00     1.C699299E+00     1.0699298E+00
    4.0000000E+00     4.2517482E+00     4.2517475E+00
    6.0000000E+00     5.4405594E+00     5.4405582E+00
    4.5000000E+00     4.7489073E+00     4.7489064E+00

FINISHED, YES ØR NØ
= YES

IS DATA TØ BE READ FRØM A FILE, TYPE YES,NØ ØR STØP
= YES

TYPE IN FILENAME
= F1
DEPENDENT DATA MEAN =   3.272727E+00
```

```
DEGREE        ALPHA            BETA             COEFF           SSR
  1       6.000000E+00    0.              0.              0.
  2       6.000000E+00    1.000000E+01  -1.748252E-01   2.622378E+01
  3       6.000000E+00    7.800000E+00  -4.631286E-10   1.325021E-15
  4       6.000000E+00    7.200000E+00   5.827505E-03   1.398601E+00
  5       6.000000E+00    6.666666E+00   5.043470E-10   6.348973E-14
  6       6.000000E+00    6.060607E+00  -4.901964E-04   3.208560E-01
  7       6.000000E+00    5.349650E+00   9.790116E-11   5.788681E-14
  8       6.000000E+00    4.523076E+00   7.832094E-05   1.324996E-01
  9       6.000000E+00    3.576472E+00   2.103909E-10   2.397700E-12
 10       6.000000E+00    2.507738E+00  -3.857962E-05   1.060824E-01

NUMBER AND POLYNOMIALS
= 5 2 4 6 8 10

REGULAR POLYNOMIAL IN DECREASING DEGREE
  -3.8579616E-05   2.3147770E-03  -6.0349540E-02   8.9681065E-01
  -8.3742828E+00   5.1078316E+01  -2.0485381E+02   5.2956509E+02
  -8.3669358E+02   7.2143513E+02  -2.5199560E+02

NUMBER, SET OF VALUES
= 4 1 4 6 4.5

        INPUT            OR OUTPUT       REG OUTPUT
   1.0000000E+00    9.9999980E-01    9.9999619E-01
   4.0000000E+00    3.9999982E+00    3.9996986E+00


   6.0000000E+00    5.9999970E+00    5.9990883E+00
   4.5000000E+00    4.3513668E+00    4.3500404E+00

FINISHED, YES OR NO
= YES

IS DATA TO BE READ FROM A FILE, TYPE YES,NO OR STOP
= STOP

PROGRAM STOP AT 280
*LIST F1

11,10
1 2 3 4 5 6 5 4 3 2 1
1 2 3 4 5 6 7 8 9 10 11
1 1 1 1 1 1 1 1 1 1 1

READY

*
```

POISON

This FORTRAN program computes the Poisson distribution function.

$$F(A;\mu) = Pr(X \le A) = \sum_{i=0}^{A} \frac{\mu^i e^{-\mu}}{i!}$$

i. e. , the probability that a random variable X, having a Poisson distribution with parameter μ, has a value less than A. The random variable must be a positive integer.

INSTRUCTIONS

The calling sequence is:

Y = POISON (XMU, A)

where:

XMU Is the parameter μ in the Poisson distribution,

Y Is the accumulative probability from 0 to A.

RESTRICTIONS

A must be a positive integer less than or equal to 20.

SAMPLE PROBLEM

Calculate $F(A;\mu)$ for $\mu = 10$ and values of A from 1 to 20.

SAMPLE SOLUTION

```
*LIST

10 XMU=10
20 DØ 10 I=1,20
30 A=I
40 Y=PØISØN(XMU,A)
50 10 PRINT 20, I,Y
60 20  FØRMAT(" F(10:",I2,")=",F10.8)
70 STØP
80 END

READY

*RUN *;PØISØN
F(10: 1)=0.00049940
F(10: 2)=0.00276940
F(10: 3)=0.01033606
F(10: 4)=0.02925269
F(10: 5)=0.06708597
F(10: 6)=0.13014143
F(10: 7)=0.22022063
F(10: 8)=0.33281966
F(10: 9)=0.45792967
F(10:10)=0.58303975
F(10:11)=0.69677625
F(10:12)=0.79155658
F(10:13)=0.86446463
F(10:14)=0.91654123
F(10:15)=0.95125934
F(10:16)=0.97295847
F(10:17)=0.98571660
F(10:18)=0.99282313
F(10:19)=0.99656541
F(10:20)=0.99846482

PRØGRAM STØP AT 70
*
```

POLFIT

This BASIC program fits least-square polynomials to bivariate Data, using an orthogonal polynomial method. Limits are 11-th degree fit and a max. of 100 data points. Program allows user to specify the lowest degree polynomial to be fit, and then fits the polynomials in order of ascending degree. At each stage, the index of determination is printed, and the user has the choice of going to the next higher degree fit, seeing either of two summaries of fit at that stage, or of stopping the program.

NOTE: This program may produce invalid results beyond a 5th degree fit.

INSTRUCTIONS

To use this program type:

```
10    DATA N, D
          [WHERE N = Number of data points to be read
              and D = Initial [lowest] degree to be fit]
100   DATA X[1], Y[1], X[2], Y[2], ...., X[N], Y[N]
          [Continuation on lines 101-299 as needed]
RUN
```

Additional instructions may be found in the listing.

SAMPLE PROBLEM

With the following bivariate data:

X	Y	X	Y	X	Y
1	38.1	3	11.78	5	5.3
1.5	14.67	3.5	1.67	5.5	1.67
2	12.76	4	5.36	6	8.91
2.5	13.15	4.5	14.6	6.5	15.67

To fit this data to an equation of the general form

$$(Y = A + BX + CX^2 + DX^3 + \ldots + LX^{10})$$

Enter the following data:

```
10    DATA 12, 1
100   DATA 1, 38.1, 1.5, 14.67, 2, 12.76, 2.5, 13.15, 3, 11.78
101   DATA 3.5, 1.67, 4, 5.35, 4.5, 14.6, 5, 5.3, 5.5, 1.67, 6
102   DATA 8.91, 6.5, 15.67
```

NOTE:

You must get the Index of Determination for all degrees first and then the detail summary on the degree with the best fit. If the degree with the best fit was not the last one, terminate that run, enter a new Line 10 with the correct degree, and then rerun the program.

The program compensates for degrees of freedom lost in the fit. Also the Index of Determination is calculated from the difference value.

SAMPLE SOLUTION

```
*TAPE
READY

10 DATA 12,1
100 DATA 1,38.1,1.5,14.67,2,12.76,2.5,13.15,3,11.78
101 DATA 3.5,1.67,4,5.35,4.5,14.6,5,5.3,5.5,1.67,6
102 DATA 8.91,6.5,15.67

*RUN


POLFIT


L E A S T - S Q U A R E S   P Ø L Y N Ø M I A L S

VERSIØN ØF 9/16/69

       NUMBER ØF PØINTS =    12
        MEAN VALUE ØF X = 3.75
        MEAN VALUE ØF Y =    11.96917
         STD ERRØR ØF Y =   9.641686

   NØTE:  CØDE FØR 'WHAT NEXT?' IS:

           0 = STØP PRØGRAM
           1 = CØEFFICIENTS ØNLY
           2 = ENTIRE SUMMARY
           3 = FIT NEXT HIGHER DEGREE
```

```
PØLYFIT ØF DEGREE    1 INDEX ØF DETERM =    0.278089     WHAT NEXT ?3

PØLYFIT ØF DEGREE    2 INDEX ØF DETERM =    0.685453     WHAT NEXT ?3

PØLYFIT ØF DEGREE    3 INDEX ØF DETERM =    0.711898     WHAT NEXT ?3

PØLYFIT ØF DEGREE    4 INDEX ØF DETERM =    0.818617     WHAT NEXT ?3

PØLYFIT ØF DEGREE    5 INDEX ØF DETERM =  0.8213     WHAT NEXT ?3

PØLYFIT ØF DEGREE    6 INDEX ØF DETERM =    0.864561     WHAT NEXT ?3

PØLYFIT ØF DEGREE    7 INDEX ØF DETERM =    0.917285     WHAT NEXT ?3

PØLYFIT ØF DEGREE    8 INDEX ØF DETERM =  0.9653     WHAT NEXT ?3

PØLYFIT ØF DEGREE    9 INDEX ØF DETERM =    0.967539     WHAT NEXT ?2
          TERM   CØEFFICIENT

            0        -1938.412
            1         7077.586
            2        -10333.11
            3         8097.644
            4        -3792.742
            5          1109.82
            6        -204.2445
            7         22.92412
            8        -1.430534
            9        0.0379561

    X-ACTUAL        Y-ACTUAL        Y-CALC          DIFF           PCT-DIFF

            1           38.1       38.0685      0.0315032        0.082754
          1.5          14.67      14.94878     -0.2787763       -1.864877
            2          12.76      11.66829       1.091711         9.35622
          2.5          13.15      15.48094      -2.330942       -15.05685
            3          11.78       8.95546        2.82454        31.53987
          3.5           1.67      2.973923      -1.303923       -43.84521
            4           5.35      6.581619      -1.231619       -18.71301
          4.5           14.6      11.74251       2.857492         24.3346
            5            5.3      7.457718      -2.157718       -28.93268
          5.5           1.67     0.3255463       1.344454        412.9839
            6           8.91      8.768875      0.1411248        1.609384
          6.5          15.67      14.86574       0.804262        5.410172

          STD ERRØR ØF ESTIMATE FØR Y =    4.073958

    WHAT NEXT ?0


READY
*
```

POLFT

This FORTRAN program fits least-squares polynomials to bivariate data, using an orthogonal polynomial method.

INSTRUCTIONS

Data for this program can be entered via teleprinter keyboard or by using the FORTRAN data file capability (see sample problem). In either case, the free field format is used for input data.

The program permits the new user to specify the lowest polynomial to be considered and then fits polynomials in ascending degree order. After each fit, the index of determination is printed and the user has a choice of the following

- 0 = input the next set of data
- 1 = print coefficients of present degree polynomial
- 2 = print entire summary of present fit
- 3 = fit next higher degree
- 4 = stop the program

The program gives the user the option to have instructions printed out. If the answer is NO, the program asks if data will be read from a previously defined file. (A sample data file is listed in the sample problem.) If the answer is NO, the program asks for pertinent data.

RESTRICTIONS

Maximum degree fit is 11.
Maximum number of paired X, Y data points is 100.
The degree of each polynomial fit must be less than the number of data points.

SAMPLE PROBLEM

Find an acceptable polynomial fit for the following data using both the keyboard input and data file input capabilities:

X	1	2	3	4	5	6	7	8	9	10
Y	-5	-17	5	145	535	1355	2833	5245	8915	14215

SAMPLE SOLUTION

```
*LIST PØLDAT

100 10,1
110 1,-5
120 2,-17
130 3,5
140 4,145
150 5,535
160 6,1355
170 7,2833
180 8,5245
190 9,8915
200 10,14215

READY

*RUN
DØ YØU WISH USER INSTRUCTIØNS, TYPE YES ØR NØ
= YES

THIS PRØGRAM FITS LEAST-SQUARES PØLYNØMIALS TØ BIVARIATE
DATA, USING AN ØRTHØGØNAL PØLYNØMIAL METHØD.  LIMITS ARE
 11-TH DEGREE FIT AND A MAX ØF 100 DATA PØINTS.  THE PRØGRAM
ALLØWS THE USER TØ SPECIFY THE LØWEST DEGREE PØLYNØMIAL TØ BE
FIT, AND THEN FITS THE PØLYNØMIALS IN ØRDER ØF ASCENDING
DEGREE.  AT EACH STAGE, THE INDEX ØF DETERMINATIØN IS
PRINTED, AND THE USER HAS THE CHØICE ØF GØING TØ THE NEXT
HIGHEST DEGREE FIT, SEEING EITHER ØF TWØ SUMMARIES ØF FIT
AT THAT STAGE, ØR ENTERING NEW READ DATA.  THE USER
SHØULD PRØVIDE THE FØLLØWING INFØRMATIØN WHEN IT IS RE-
QUESTED:

      1.  THE NUMBER ØF DATA PØINTS AND THE INITIAL (LØWEST)
          DEGREE PØLYNØMIAL TØ BE FIT.
      2.  THE DATA PØINTS IN X(1),Y(1),X(2),Y(2), ... ,X(N),
          Y(N) ØRDER.

NØTE:  EACH VALUE ENTERED MUST BE FØLLØWED BY A CØMMA (,).
DØ YØU WANT TØ USE DATA FILES, TYPE YES ØR NØ
= YES
NAME ØF DATA FILE
= PØLDAT
DØ YØU WANT DATA SHIFTED TØ ZERØ MEAN, YES ØR NØ
= NØ


L E A S T   S Q U A R E S   P Ø L Y N Ø M I A L S


      NUMBER ØF PØINTS =10
      MEAN VALUE ØF X  =  5.500000E+00
      MEAN VALUE ØF Y  =  3.322600E+03
      STD ERRØR ØF  Y  =  4.816941E+03


NØTE:  CØDES FØR 'WHAT NEXT?' ARE:

      0 = INPUT THE NEXT SET ØF DATA
      1 = CØEFFICIENTS ØNLY
      2 = ENTIRE SUMMARY
      3 = FIT NEXT HIGHER DEGREE
      4 = STØP
```

```
PØLYFIT ØF DEGREE  1   INDEX ØF DETERM =    7.384701E-01
WHAT NEXT
= 3

PØLYFIT ØF DEGREE  2   INDEX ØF DETERM =    9.783259E-01
WHAT NEXT
= 3

PØLYFIT ØF DEGREE  3   INDEX ØF DETERM =    9.996845E-01
WHAT NEXT
= 3

PØLYFIT ØF DEGREE  4   INDEX ØF DETERM =    1.000000E+00
WHAT NEXT
= 2

TERM        CØEFFICIENT
  0         5.0000541E+00
  1         9.0001956E+00
  2         3.0000537E+00
  3         6.0000079E+00
  4         2.0000004E+00

X-ACTUAL             Y-ACTUAL      Y-CALC          DIFF          PCT-DIFF
 1.000000E+00  -5.000000E+00  -5.000095E+00   9.542704E-05  -1.908504E-03
 2.000000E+00  -1.700000E+01  -1.700018E+01   1.797676E-04  -1.057445E-03
 3.000000E+00   5.000000E+00   4.999767E+00   2.327561E-04   4.655339E-03
 4.000000E+00   1.450000E+02   1.449997E+02   2.841949E-04   1.959969E-04
 5.000000E+00   5.350000E+02   5.349996E+02   3.585815E-04   6.702463E-05
 6.000000E+00   1.355000E+03   1.355000E+03   4.577637E-04   3.378331E-05
 7.000000E+00   2.833000E+03   2.832999E+03   6.103516E-04   2.154436E-05
 8.000000E+00   5.245000E+03   5.244999E+03   7.324219E-04   1.396419E-05
 9.000000E+00   8.915000E+03   8.914999E+03   9.765625E-04   1.095415E-05
 1.000000E+01   1.421500E+04   1.421500E+04   1.098633E-03   7.728687E-06

          STD ERRØR ØF ESTIMATE FØR Y =    8.466561E-04
WHAT NEXT
= 0
DØ YØU WANT TØ USE DATA FILES, TYPE YES ØR NØ
= NØ
ENTER NUMBER ØF PØINTS, INITIAL DEGREE ØF PØLYNØMIAL
= 10,4
ENTER DATA PØINTS IN X(1),Y(1), ... ,X(N),Y(N) ØRDER
= 1,-5,2,-17,3,5,4,145,5,535
= 6,1355,7,2833,8,5245,9,8915,10,14215
DØ YØU WANT DATA SHIFTED TØ ZERØ MEAN, YES ØR NØ
= NØ


L E A S T   S Q U A R E S   P Ø L Y N Ø M I A L S

      NUMBER ØF PØINTS =10
      MEAN VALUE ØF X  =   5.500000E+00
      MEAN VALUE ØF Y  =   3.322600E+03
      STD ERRØR ØF  Y  =   4.816941E+03
```

```
NØTE:  CØDES FØR 'WHAT NEXT?' ARE:

       0 = INPUT THE NEXT SET ØF DATA
       1 = CØEFFICIENTS ØNLY
       2 = ENTIRE SUMMARY
       3 = FIT NEXT HIGHER DEGREE
       4 = STØP

PØLYFIT ØF DEGREE  4  INDEX ØF DETERM =    1.000000E+00
WHAT NEXT
= 1

TERM        CØEFFICIENT
  0         5.0000541E+00
  1         9.0001956E+00
  2         3.0000537E+00
  3         6.0000079E+00
  4         2.0000004E+00
WHAT NEXT
= 4

PRØGRAM STØP AT 2255
*
```

PROBC

This program calculates the probabilities of simple combinations of random variables and handles continuous, discrete and mixed cases.

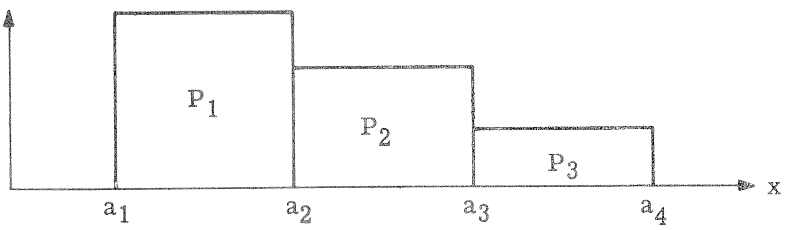
INSTRUCTIONS

The routine allows the user to form five basic binary combinations of random variables: addition, subtraction, multiplication, division and involution. Also, four elementary transformations are provided: sine, cosine, natural logarithm and exponential.

The calculated probability distribution function is presented in interval form. For example, for the random variable x:

$$Pr(a_1 < x < a_2) = P_1$$
$$Pr(a_2 < x < a_3) = P_2$$
$$Pr(a_3 < x < a_4) = P_3$$

The representation can be pictured as



i.e., within the endpoints the distribution is assumed to be uniform. The endpoints or "grid" are entered as: 4, a_1, a_2, a_3, a_4 (the number of endpoints, followed by the points in algebraically increasing order).

During execution the program requests the following:

WANT TO SEE CODE, YES OR NO =

A response of YES will result in the printing of the various options available in combining variables and modifying the intervals. It may be desired to halve the intervals for the operator and/or operand grid to enhance the resulting probability calculations. Codes (10-12) provide for this. In the case of transformations, it is only necessary to halve the operator distribution. Option 13, RESTART, destroys the previous output grid and returns to (1).

Following a response of NO or the output of the operation code, the program requests -

(1) NUMBER, POINTS FOR FIRST DENSITY =

and

PROBABILITIES =

The option of saving the resulting distributions for later combining of independent random variables is then provided by -

(2) WANT TO SAVE DISTRIBUTION, YES, NO OR STOP =

The routine then requests the operation code -

(3) READ OPERATION CODE (1 - 13) =

If the response involves a binary operation (1 - 5) a request for the next density follows -

NEXT DENSITY FROM STORAGE, YES OR NO =

NO requires keyboard input of appropriate information. YES, implies that at a prior time the distribution was saved by a response of YES to (2).

NEW OUTPUT POINTS, YES OR NO =

requires a response of YES at least once to define the grid points.

A response to (3) involving the codes (10 - 12) dealing with the intervals will result in the output of the number of grid points corresponding to both distributions and the maximum cumulative probability difference encountered between the current and previously calculated distributions.

The program then requests -

WANT TO SEE DISTRIBUTION, YES OR NO =

If YES, the resulting probabilities are output, followed by CHECK SUM and the mean (first moment about the origin), variance (second moment about the mean - σ^2), the coefficient of skewness (third moment about the mean divided by σ^3), and the coefficient of kurtosis (fourth moment about the mean divided by σ^4).

If NO, only the CHECK SUM and MOMENTS are printed.

The routine then cycles to (2). A response of STOP to (2) will stop the program.

SAMPLE PROBLEM

Two wooden rods are to be joined together and then a piece is to be cut off from the new rod. The first rod has a length of 0' to 1' with probability .2, 1' to 2' with probability .6 and 2' to 3' with probability .2; the second rod has a length of 1' to 2' with probability .6 and 2' to 3' with probability .4; the piece to be cut off has a length of 0' to .25' with certainty. What is the probability function for the length of the final rod?

SAMPLE SOLUTION

```
*RUN

PRØGRAM TØ CØMBINE RANDØM VARIABLES

WANT TØ SEE CØDE, YES ØR NØ
= YES


ØPERATIØN CØDE
ADD, 1
SUBTRACT, 2
MULTIPLY , 3
DIVIDE, 4
INVØLUTIØN, 5
SINE, 6
CØSINE, 7
LØG, 8
EXP, 9
HALVE BØTH GRIDS, 10
HALVE ØPERATØR GRID, 11
HALVE ØPERAND GRID, 12
RESTART, 13

NUMBER, PØINTS FØR FIRST DENSITY
= 4  0  1  2  3

PRØBABILITIES
= .2  .6  .2

WANT TØ SAVE DISTRIBUTIØN,YES,NØ ØR STØP
= NØ


READ ØPERATIØN CØDE(1-13)
= 1

NEXT DENSITY FRØM STØRAGE, YES ØR NØ
= NØ

NUMBER, PØINTS FØR NEXT DENSITY
= 3  1  2  3

PRØBABILITIES
= .6  .4

NEW ØUTPUT PØINTS, YES ØR NØ
= YES

NUMBER, ØUTPUT PØINTS
= 41  0  .25  .5  .75  1  1.25  1.5  1.75  2  2.25  2.5  2.75
= 3  3.25  3.5  3.75  4  4.25  4.5  4.75  5  5.25  5.5  5.75
= 6  6.25  6.5  6.75  7  7.25  7.5  7.75  8  8.25  8.5  8.75
= 9  9.25  9.5  9.75  10
```

```
WANT TØ SEE DISTRIBUTIØN, YES OR NO
= NØ

CHECK SUM =  1.000000

MEAN,VARIANCE,SKEWNESS,KURTOSIS
 3.400000E+00   9.681260E-01   5.038453E-02   2.659762E+00

WANT TØ SAVE DISTRIBUTIØN,YES,NØ OR STOP
= NØ

READ ØPERATIØN CODE(1-13)
= 2

NEXT DENSITY FROM STØRAGE, YES ØR NØ
= NØ

NUMBER, PØINTS FØR NEXT DENSITY
= 2  0  .25

PRØBABILITIES
= 1

NEW ØUTPUT PØINTS, YES ØR NØ
= NØ

WANT TØ SEE DISTRIBUTIØN, YES OR NO
= YES

*****RESULTING PRØBABILITIES*****

      START            STØP            PRØBABILITY
   0.               7.5000000E-01     0.
   7.5000000E-01    1.0000000E+00     7.4999999E-03
   1.0000000E+00    1.2500000E+00     1.5000000E-02
   1.2500000E+00    1.5000000E+00     1.5000000E-02
   1.5000000E+00    1.7500000E+00     1.5000000E-02
   1.7500000E+00    2.0000000E+00     4.2500000E-02
   2.0000000E+00    2.2500000E+00     6.9999999E-02
   2.2500000E+00    2.5000000E+00     6.9999999E-02
   2.5000000E+00    2.7500000E+00     6.9999999E-02
   2.7500000E+00    3.0000000E+00     8.4999999E-02
   3.0000000E+00    3.2500000E+00     9.9999999E-02
   3.2500000E+00    3.5000000E+00     9.9999999E-02
   3.5000000E+00    3.7500000E+00     9.9999999E-02
   3.7500000E+00    4.0000000E+00     7.7499999E-02
   4.0000000E+00    4.2500000E+00     5.4999999E-02
   4.2500000E+00    4.5000000E+00     5.4999999E-02
   4.5000000E+00    4.7500000E+00     5.4999999E-02
   4.7500000E+00    5.0000000E+00     3.2499999E-02
   5.0000000E+00    5.2500000E+00     9.9999999E-03
   5.2500000E+00    5.5000000E+00     9.9999999E-03
   5.5000000E+00    5.7500000E+00     9.9999999E-03
   5.7500000E+00    6.0000000E+00     4.9999999E-03

CHECK SUM =  1.000000

MEAN,VARIANCE,SKEWNESS,KURTOSIS
 3.275000E+00   9.837506E-01   4.919180E-02   2.669969E+00

WANT TØ SAVE DISTRIBUTIØN,YES,NO OR STOP
= STØP

PRØGRAM STØP AT 1460
*
```

PROVAR

This BASIC program performs several types of calculations for normal and student's-T distributions. Given one- or two-tailed probabilities, it will produce the corresponding limits; given a variate or a set of limits, it will produce the corresponding probabilities.

INSTRUCTIONS

To use this program enter data as follows:

```
100 DATA   DF, MEAN, SIGMA, (BLOCK1), (BLOCK2),.....
```

Where df is degrees of freedom (setting df=zero signals normal distribution), mean and sigma are the mean and standard deviation of the distribution, and the data blocks may be any mixture of the 4 types shown below:

BLOCK(INPUT)	OUTPUT PRODUCED
1,V	PROB OF X EXCEEDING THE VALUE V
2,L,U	PROB OF X INSIDE LIMITS L AND U
3,P	VALUE WHICH IS EXCEEDED WITH PROB P
4,P	LIMITS CONTAINING X WITH PROB P

For example, the line of data shown below would produce as output the probability of a normally distributed variable with mean 100 and standard deviation 10 being in the range 95-105, and also the value of the variable that will be exceeded only 2.5 percent of the time:

```
100 DATA    0,100,10,2,95,105,3,0.025
```

Data for additional cases may be continued on lines 101-199 as needed.

Additional instructions may be obtained by listing the program.

SAMPLE PROBLEM

Given: Monthly sales (in pounds) will be distributed normally with a mean of 155,000. 99 percent of the time actual sales will be within 18,000 pounds of the mean. The standard deviation can be approximated by 6,000 pounds. Based on the above information the following problems can be solved:

1. What is the probability that sales volume will exceed 165,000 pounds? 170,000 pounds?
2. What is the probability that sales volume will be at least 145,000 pounds?
3. What is the probability that sales volume will be between 150,000 and 160,000 pounds?
4. What sales volume level will fall below 25 percent of the time?
5. What sales volume range can be expected 50 percent of the time?

To use this program prepare the following data lines:

```
100 DATA 0, 155000, 6000, 1, 165000, 1, 170000, 1, 145000, 2, 150000
101 DATA 160000, 3, .75, 4, .5
```

Question 2 is really the same as Question 1 and, therefore, uses type 1 input.

In order to answer Question 4 above, the value which is exceeded 75 percent of the time was requested.

SAMPLE SOLUTION

```
*100 DATA 0,155000,6000,1,165000,1,170000,1,145000,2,150000
*101 DATA 160000,3,.75,4,.5
*RUN


    CALCULATIØNS FØR A NØRMAL DISTRIBUTIØN
    HAVING A MEAN ØF    155000 AND A SIGMA ØF   6000

    CASE   THE PRØBABILITY ØF A VARIATE:

     1     EXCEEDING    165000 IS  0.0476
     2     EXCEEDING    170000 IS 0.00621
     3     EXCEEDING    145000 IS  0.9524
     4     BETWEEN    150000 AND    160000 IS 0.59586
     5     EXCEEDING    150953 IS 0.75
     6     BETWEEN    150953 AND    159047 IS  0.5

READY
*
```

SAMPLE PROBLEM

To illustrate the use of the Student-t portion of this program, calculate several values of t which could be used for testing the significance between 2 sample means. Consider the following conditions (all of which are based on single-tail area):

1. 20 degrees of freedom, alpha of .05
2. 17 degrees of freedom, alpha of .025
3. 32 degrees of freedom, alpha of .03

To use this program type the following data lines:

```
100 DATA 20, 0, 1, 3, .05

100 DATA 17, 0, 1, 3, .025

100 DATA 32, 0, 1, 3, .03
```

Standard Student-t curves are based on a mean of 0 and a sigma of 1.

Since the degrees of freedom were different for each case the program was run 3 times.

To get the two-tail areas, simply multiply by 2 the answers given by the program.

This program will provide a critical value for 3 percent, something most tables do not provide.

```
*100 DATA 20,0,1,3,.05
*101
*RUN


    CALCULATIØNS FØR A STUDENT'S T-DISTRIBUTIØN
    HAVING A MEAN ØF       0 AND A SIGMA ØF      1
    AND HAVING    20 DEGREES ØF FREEDØM.

    CASE   THE PRØBABILITY ØF A VARIATE:

     1     EXCEEDING     1.724243 IS 0.05

READY
*100 DATA 17,0,1,3,.025
*RUN


    CALCULATIØNS FØR A STUDENT'S T-DISTRIBUTIØN
    HAVING A MEAN ØF       0 AND A SIGMA ØF      1
    AND HAVING    17 DEGREES ØF FREEDØM.

    CASE   THE PRØBABILITY ØF A VARIATE:

     1     EXCEEDING     2.108943 IS   0.025

READY
*100 DATA 32,0,1,3,.03
*RUN


    CALCULATIØNS FØR A STUDENT'S T-DISTRIBUTIØN
    HAVING A MEAN ØF       0 AND A SIGMA ØF      1
    AND HAVING    32 DEGREES ØF FREEDØM.

    CASE   THE PRØBABILITY ØF A VARIATE:

     1     EXCEEDING     1.949741 IS 0.03

READY
*
```

RANDX

This FORTRAN function generates a random number from a uniform distribution between zero and one.

INSTRUCTIONS

The calling sequence is:

Y = RAND(X)

Where:

Y is the value of the random number and represents a sample from the uniform distribution between zero and one.

X must be negative the first time the routine is called. It is ignored when positive or zero. The routine uses the previous random number on subsequent references. Therefore, X should be set positive or zero for additional random numbers.

To generate any sequence of random numbers starting with A, Set X = -A. (A must be greater than 0, but not greater than 1.)

SAMPLE PROBLEM

Find 10 random numbers from a uniform distribution between zero and one.

SAMPLE SOLUTION

```
*LIST

10       Y=RAND(-1.0)
20       PRINT 1,Y
30    1  FØRMAT(15H RANDØM NUMBER=,1PE20.7)
40       DØ 2 L=2,10
50       Y=RAND(0.0)
60    2  PRINT 1,Y
70       STØP
80       END

READY

*RUN *;RANDX
RANDØM NUMBER=          6.8890101E-01
RANDØM NUMBER=          6.7469373E-01
RANDØM NUMBER=          5.3704079E-01
RANDØM NUMBER=          2.0333507E-01
RANDØM NUMBER=          9.7009331E-01
RANDØM NUMBER=          3.9913800E-01
RANDØM NUMBER=          8.6720850E-01
RANDØM NUMBER=          4.0229724E-01
RANDØM NUMBER=          8.9781811E-01
RANDØM NUMBER=          6.9887078E-01

PRØGRAM STØP AT 70
*
```

RNDNRM

This FORTRAN function computes pseudo-random numbers having a normal distribution (mean = 0 and STD = 1). The algorithm uses an acceptance-rejection scheme applied to a transformed variable.[1]

INSTRUCTIONS

The calling sequence is:

R = RNDNRM (S)

Where:

R is the random number

S is an arbitrary number used to initiate the calculations. Changing S after the first call to RNDNRM will have no effect upon the random number sequence.

This function calls the uniform random number generator FLAT, which must therefore be included in the RUN list.

NOTE:

This method is about three times as fast per normal deviate as DNORM.

SAMPLE PROBLEM

Find 10 normally distributed random numbers.

[1] Roe, G. M., "A Fast Normal Random Number Generator", General Electric Company, Report No. 70-C-151, May 1970.

SAMPLE SOLUTION

```
*LIST

10 S=1014
20 DØ 10 I=1,10
30 X=RNDNRM(S)
40 10 PRINT 40,X
50 40 FØRMAT(F13.8)
60 END

READY

*RUN *;RNDNRM;FLAT
 -0.02045153
  1.72119142
 -0.88185998
  1.22851484
  0.51439027
 -0.91936707
  0.16635922
  0.57126343
  0.41624698
 -0.46230396

PRØGRAM STØP AT 0
*
```

SEVPRO

This BASIC program applies a CHI-SQUARE test to several sample proportions.

INSTRUCTIONS

To use this program simply supply values for N and S in the following format:

```
100    DATA  N1,S1, N2,S2, N3,S3, .....
RUN
```

where:

N1 = size of sample

S1 = the number of successes in sample 1.

Repeat for the total number of samples.

Additional instructions may be found in the listing.

SAMPLE PROBLEM

Statistically test the following data for significance.

In response to heavy demand for a particular model portable radio, four separate assembly lines have been in operation for the last two weeks. While all are identical operations for all intents and purposes, there are unavoidable differences in equipment, operator experience and so on. The reject rate has been running fairly high, and each line is blaming it on the others. The Quality Engineer decided to check all rejects for one day to discover whether the quality was significantly different for the four lines. This data appears on the following page.

Assembly Line	Total Units Assembled	Number Rejected	Percent Rejected
1	1217	45	3.7
2	948	49	5.2
3	1165	33	·2.8
4	1121	44	3.9

Line 2 seemed higher, and line 3 lowest in reject rate, but such a difference could be the result of just chance. The engineer decided to make a statistical test for significance.

SAMPLE SOLUTION

```
*100 DATA 1217,45,948,49,1165,33,1121,44
*RUN


CHI-SQUARE TEST ØF SEVERAL PRØPØRTIØNS:

 SAMPLE        SUCCESSES / TØTAL          % SUCCESSES

      1          45 /          1217          3.698
      2          49 /           948          5.169
      3          33 /          1165          2.833
      4          44 /          1121          3.925

CHI-SQUARED FØR THESE DATA    =     7.819924
CØRRESPØNDING NØRMAL DEVIATE =     1.718657

 THE PRØBABILITY ØF THIS VALUE ØF CHI-SQUARE
 BEING EXCEEDED BY CHANCE ALØNE IS    0.043


READY
*
```

ANALYSIS OF RESULT

It looks as though the 4 lines are not alike on quality. Lines 1 and 4 are much alike, but Line 2 seems to have problems. The test of the four proportions simply tells us that it's rather improbable (only 1 chance in 24 or so) that chance could account for this much variability, assuming the quality level was actually the same on all lines.

SMLRP (SMLRPOBJ)

SMLRP (Stepwise Multiple Linear Regression Program) computes in a stepwise manner the least squares "best" value for the coefficients of an equation of the form:

$$Y = B_0 + B_1X_1 + B_2X_2 + \ldots + B_nX_x$$

where Y is the dependent variable; X_1, X_2, . . . , X_n are the independent variables; and B_0, B_1, B_2, . . . , B_n are the coefficients to be determined. The program also provides statistical measures of the reliability of the computed coefficients.

SMLRPOBJ is the compiled object version of this same program.

PROGRAM FEATURES

- Program handles up to 20 variables.
- Ability to identify the problem data file at run time.
- Stepwise addition and deletion of independent variables based on the critical F-statistic.
- Optional printout of means, standard deviation, and correlation matrix.
- User selection of dependent and independent variables at run time.
- Optional stepwise results.
- Free-field data input.
- Variable transformations on input data.
- Ability to force variables into the regression.
- Optional residual summary.
- Option of entering the critical F-value or the confidence level.
- User provided variable labels.

METHOD

The program performs the stepwise regression analysis, using the matrix of simple correlation coefficients. Any combination of one dependent variable and one or more independent variables may be selected for analysis from the input or transformed variables.

In the stepwise procedure, independent variables are added one at a time to the regression, giving the following intermediate equations:

$$Y = B_0 + B_1X_1$$

$$Y = B'_0 + B'_1X_2 + B'_2X_2$$

$$Y = B''_0 + B''_1X_1 + B''_2X_2 + B''_3X_3$$

In this manner, it is possible to obtain valuable statistical information at each step of the calculation.

At each step, those independent variables not included in the regression are inspected to find the one that will give the greatest reduction in the variation of Y. This variable is then tested for significance; i.e., the computed F-ratio of the variable is compared to the supplied critical F-value. If the computed value is greater than the critical value, the variable is considered significant and is added to the regression solution.

After each new variable has been added, those variables already in the regression are inspected to see if any of them can now be deleted because their contribution to the reduction in the variation of Y is no longer significant. Those variables which have a computed F-ratio less than the critical value are considered insignificant and are deleted from the regression solution.

This process is continued until no more variables can be added or deleted. Thus, the final regression solution contains only those variables that are statistically significant (except forced variables).

Under user control, it is possible to force one or more of the independent variables into the regression solution. These variables are added to the fit, before any other variables, with no regard for statistical significance. The remaining independent variables are then added to the fit in the normal stepwise procedure as outlined in the preceding paragraphs. However, once a variable has been forced into the regression, it will not be deleted in the following steps even though it is not statistically significant.

DATA PREPARATION

Before running the program, the user should go into the EDIT Subsystem and enter the problem data into one of his own files in free-field format.

NOTE:

Line numbers must not be used, which is the reason for using the EDIT Subsystem.

General rules for free-field input are:

1. Data entries are separated by one or more blanks, or a comma.

2. Blanks preceding a data field are ignored. The first blank following a data field terminates the field. The combination of a blank followed by a comma is treated as a null field and must be avoided.

3. To repeat a numeric value, enter the repetition count followed by an asterisk (*). The asterisk is then followed by the numeric value to be repeated, i.e., 4*3.2.

4. The following input formats are acceptable for the number twenty-five:

   (a) 25
   (b) 25.0
   (c) 2.5E1
   (d) .25E+2
   (e) 250E-1

The first line of the problem data file contains the total number of input variables, which is limited to 20. The observed data follows next and is entered by observations (row-wise); that is, each record contains an entry for X_1, X_2, X_3, . . . , X_n. Each observation must begin on a new line. All the data values for an observation are assumed to be contained on a single line unless that line ends with a comma followed by a carriage return. When it does, the next line is considered a continuation of the first line.

Since there is no special end of data flag, the user should ensure that no extraneous information follows the last observation for the current problem. The maximum number of observations is limited to 999.

See the following sections for optional transformations and variable labels.

After all the data has been entered and checked for accuracy, the file should be saved.

DATA TRANSFORMATIONS

The program provides the option of performing transformations on the raw variable data. The user performs these transformations by entering in the data file a series of pseudo-codes which logically produce the desired transformations on the specified variables.

To use the transformation option, the second line of the data file must contain the alphabetic word "TRNF" followed by the total number of variables after all transformations have been performed.

The transformation codes are entered one per line following the line containing "TRNF", and are of the following general form:

OP I J K

where OP is the three character alphabetic transformation code, I and J are the indices of the variables to be operated upon, and K is the index of the resulting variable. Most transformation codes operate on only one variable, and in these cases a dummy index of '1' must be used in the position of the index that will not be used.

The index of a given variable is simply its position with respect to the other variables. For example, the third variable has an index of 3. The program provides 10 pseudo variables with indices of 21-30 that may be used for storage of intermediate results. However, the final index of a transformed variable may not exceed 20. A description of the transformation codes follows:

1. Two-Operand Codes

CODE				MEANING
ADD	I	J	K	$X_i + X_j \longrightarrow X_k$
SUB	I	J	K	$X_i - X_j$ -- X_k
MPY	I	J	K	$X_i * X_j$ -- X_k
DVD	I	J	K	X_i/X_j -- X_k

2. One-Operand Codes

CODE				MEANING
RCP	I	K	1	$1.0/X_i$ -- X_k
MOV	I	K	1	X_i -- X_k
EXP	I	K	1	e^{X_i} -- X_k
TEN	I	K	1	10^{X_i} -- X_k
LGN	I	K	1	LOG_e (Xi) -- X_k
LOG	I	K	1	LOG_{10} (Xi) -- X_k
SIN	I	K	1	Sine (Xi) -- X_k
COS	I	K	1	Cosine (Xi) -- X_k
TNH	I	K	1	Hyperbolic Tangent (Xi) -- X_k
SQR	I	K	1	Xi -- X_k

3. Terminating Code

END 1 1 1 last line of transformation codes.

The observed variable data or variable labels start on the next line following the "END" code.

A special transformation code permits the use of constants in the transformation manipulations. This special code has the form:

CON I constant value 1

where I is the index (21-30) of one of the pseudo variables. The appended power of ten exponent (E) is not permitted for the constant value.

An example of transformation coding would be to evaluate the expression:

$$X_{13} = X_1 - 3(X_3)^2 + LN\ (X_2 - X_4)$$

The following lines would be required:

TRNF	13			
CON	21	-3	1	-3.0 -- X_{21}
MPY	3	3	22	$(X_3)^2$ -- X_{22}
MPY	21	22	22	$-3(X_3)^2$ -- X_{22}
SUB	2	4	13	(X_2-X_4) -- X_{13}
LGN	13	13	1	$LN\ (X_2-X_4)$ -- X_{13}
ADD	22	13	13	$-3(X_3)^2 + LN\ (X_2-X_4)$ -- X_{13}
ADD	1	13	13	$X_1 - 3(X_3)^2 + LN(X_2-X_4)$ -- X_{13}
END	1	1	1	

VARIABLE LABELS

As an option, the user may provide meaningful labels or names for each of the final variables. If present, these labels are included in the printed results for each variable. The labels are entered in the data file following the transformation coding (if present) and preceding the observed data.

The label names may consist of one to four characters, the first of which must be alphabetic. The remaining characters of the label may be any character with the exception of the comma (,), quote mark ("), exclamation mark (!), or embedded blank.

First, a line containing the entry "LABL" and the number of labels is made (e.g., LABL 11). This is followed by the line of label names, which must end with the entry "ELAB" (e.g., X1, X2, X3, Y4, Y5, ELAB). This line may be continued to the next by ending it with a comma.

The number of labels may not exceed the number of final variables. However, fewer labels than variables are acceptable, in which case the labels for the remaining variables are blank.

OPERATING INSTRUCTIONS

1. The user should first establish and save the data file. The program can then be executed by running either the source version SMLRP or the object version SMLRPOBJ. For example, to run the object version, the user could type:

```
GET    LIBRARY/SMLRPOBJ,R

RUN    SMLRPOBJ
```

2. The program responds by asking the user to enter the name of his data file. The program then prints the number of initial variables, the final number of transformed variables, and the number of observations.

3. The user is then asked to enter "YES" or "NO" (Y or N is sufficient) for the optional printout of the mean, standard deviation, minimum value, and maximum value for each variable.

4. Next, "YES" or "NO" is requested for the printout of the lower half of the simple correlation matrix.

5. The user is next requested to enter the number of independent (X) variables to be considered, then the index of the dependent (Y) variable, followed by the indices (in any order) of all the independent (X) variables. The index of any variable is simply its position with respect to the other variables. If the index of any independent variable is entered as a negative value, that variable will be "forced" into the regression. This option should be used with discretion, and only those variables known to be significant by some prior knowledge or theory should be forced. If any input errors are detected, such as an index larger than the total number of variables, an incorrect number of indices, or a conflict between the index of the Y variable and that of one of the X variables; the input is requested again.

6. Next, a request is made to enter a "F" or "CL", indicating whether the user is going to supply the critical F-value of the confidence level. The indicated value is then requested. If the confidence level is entered, the program will compute the critical F-value at each step of the regression.

7. The user is then asked to answer "YES" or "NO" to whether full results are to be printed at each step of the regression.

8. The program then proceeds with the stepwise regression analysis. First, a line is printed listing the data file name, the identification of the dependent variable, the confidence level (CL), the degrees of freedom (DØF), and the residual sum of squares (RSS). Initially RSS equals the total sums of squares of the dependent variable.

At each step of the regression, the following summary is printed: The number of the step, the identification and label of the independent variable preceded by a "+", "-" or "*" (indicating the variable was added, deleted or forced), DØF, the coefficient or multiple determination (R^2), and the standard error of the estimate (SEE).

Depending on the option taken, the regression results are printed at each step of the analysis or at its completion. The following items are listed for each variable included in the regression: Variable identification and label, regression coefficient, F-ratio and beta weight (coefficient).

9. The program then asks if the residual summary is desired. If the answer is "YES", the following results are printed for each observation: The observed Y-value, the calculated Y-value, the residual or error, the percent error, and the cumulative sum of the squared errors (C-ER 2).

10. The user is then asked to indicate which one of the following actions the program is to take: (1) request a new data file, (2) select a new set of variables from the current data file, (3) alter the critical F-value or confidence level, (4) terminate program execution.

SAMPLE PROBLEM

A listing of the data and program output for a sample regression problem is shown on the following pages. The sample problem contains 16 observations of 11 variables. The transformation coding does nothing constructive and is included only for the purpose of illustration. The regression analysis was run using variable 10 as the dependent variable and variables 1 - 9 as the independent variables. Variable 9 was forced into the regression. The confidence level was entered as 0.90.

The regression results show that 90.35% of the total variation in the dependent variable was explained by variables 1, 2, 3, 5, and 9. Looking at variable 9, it is apparent that there was no need to force this variable into the regression. The F-ratio of 9 was 9.92, well above the critical F-value of 3.29. The remaining unexplained variation (RSS) was 43.60 and the estimate of the error about the regression line was 2.09. Variables 4, 6, 7, and 8 were not significant at the 0.90 confidence level.

When the critical F-value was reduced to 0.1, all the independent variables were included in the regression. R^2 increased from 0.904 to 0.959, while the standard error of the estimate decreased from 2.09 to 1.77. From inspecting the F-ratios, it is apparent that variables 6 and 8 contributed very little to the overall fit.

As an interesting exercise, the user is urged to run this same problem, including all the independent variables in the regression by using a critical F-value of 0.1. After the results have been printed, enter a confidence level of 0.90. The final results will be different from those obtained by initially entering a confidence level of 0.90. In fact, two more significant variables will be included in the results. Although the sample data is "manufactured", this situation does occur with real live data and points out that backward deletion of variables is sometimes better than forward addition.

```
*LIST LØDAT

11
TRNF    11
CØN  21 -10.25 1
MPY  11 21 23
DVD 23 21 23
MØV  23 11 1
END  1 1 1
LABL 11
XV1 XV2 XV3 XV4 XV5 XV6 XV7 XV8 XV9 YV10 YV11 ELAB
-1 -1 -1 1 1 1 1 1 1 52.5 6.12
1 -1 -1 1 1 1 -1 -1 1 44.4 9.54
-1 1 -1 1 1 1 -1 1 -1 40.6 8.32
1 1 -1 1 1 1 1 -1 -1 36.0 9.88
-1 -1 1 1 1 1 1 -1 -1 43.0 13.31
1 -1 1 1 1 1 -1 1 -1 36.0 16.52
-1 1 1 1 1 1 -1 -1 1 39.3 9.97
1 1 1 1 1 1 1 1 1 38.0 11.17
0 0 0 0 0 0 0 0 0 38.0 12.58
0 0 0 0 0 0 0 0 0 37.4 12.22
2 0 0 4 0 0 0 0 0 35.5 13.76
-2 0 0 4 0 0 0 0 0 46.6 12.37
0 2 0 0 4 0 0 0 0 37.5 9.21
0 -2 0 0 4 0 0 0 0 52.6 9.67
0 0 2 0 0 4 0 0 0 37.9 13.19
0 0 -2 0 0 4 0 0 0 39.5 5.57

READY


*RUN
STEPWISE MULTIPLE LINEAR REGRESSIØN PRØGRAM


ENTER DATA FILE 'NAME'
= LØDAT
 11 INITIAL VARIABLES
 11 TRANSFØRMED VARIABLES
 16 ØBSERVATIØNS
MEANS
= YES

VAR  LABEL        MEAN     STD-DEV          MIN          MAX
  1    XV1        0.        1.0328  -2.0000E+00   2.0000E+00
  2    XV2        0.        1.0328  -2.0000E+00   2.0000E+00
  3    XV3        0.        1.0328  -2.0000E+00   2.0000E+00
  4    XV4        1.0000    1.2649   0.           4.0000E+00
  5    XV5        1.0000    1.2649   0.           4.0000E+00
  6    XV6        1.0000    1.2649   0.           4.0000E+00
  7    XV7        0.        0.7303  -1.0000E+00   1.0000E+00
  8    XV8        0.        0.7303  -1.0000E+00   1.0000E+00
  9    XV9        0.        0.7303  -1.0000E+00   1.0000E+00
 10    YV10      40.9250    5.4894   3.5500E+01   5.2600E+01
 11    YV11      10.8375    2.8704   5.5700E+00   1.6520E+01
CØRR. MATRIX
= YES
```

```
CØRRELATIØN MATRIX

        [  1]     [  2]     [  3]     [  4]     [  5]     [  6]     [  7]     [  8]
   [  9]     [ 10]     [ 11]
[  1]
        1.000
[  2]
        0.        1.000
[  3]
        0.        0.        1.000
[  4]
        0.        0.        0.        1.000
[  5]
        0.        0.        0.       -0.333     1.000
[  6]
        0.        0.        0.       -0.333    -0.333     1.000
[  7]
        0.        0.        0.        0.        0.        0.        1.000
[  8]
        0.        0.        0.        0.        0.        0.        0.        1.000
[  9]
        0.        0.        0.        0.        0.        0.        0.        0.
        1.000
[10]
       -0.508    -0.614    -0.240     0.033     0.340    -0.148     0.153     0.073
        0.309     1.000
[11]
        0.274    -0.159     0.727     0.293    -0.240    -0.248    -0.123    -0.018

       -0.357    -0.426     1.000
ENTER NØ. ØF 'X' VARS, THEN INDEX ØF 'Y' VAR
FØLLØWED BY INDICES ØF ALL 'X' VARS
= 9,10,-9,1,2,3,4,5,6,7,8
ENTER F ØR CL FØR: CRITICAL F-VALUE ØR CØNFIDENCE LEVEL
= CL
ENTER CØNFIDENCE LEVEL
= 0.90
STEPWISE RESULTS?
= NØ

LØDAT    : 1   DEP-VAR=10:   YV10   CL=0.90   DØF= 15   RSS=       45
.0100                               2

STEP    VAR:LABEL      F-CRIT    DØF      R-SQ           SEE
  1     * 9:  XV9       3.10      14    0.0957         5.4035
  2     + 2:  XV2       3.14      13    0.4724         4.2829
  3     + 1:  XV1       3.18      12    0.7305         3.1862
  4     + 5:  XV5       3.23      11    0.8460         2.5155
  5     + 3:  XV3       3.29      10    0.9035         2.0880

VAR   LABEL        CØEFFICIENT        STD-ERR       F-RATIØ        BETA-WT
  1     XV1         -2.700000         0.5220         26.75         -0.5080
  2     XV2         -3.262500         0.5220         39.06         -0.6138
  3     XV3         -1.275000         0.5220          5.97         -0.2399
  5     XV5          1.475000         0.4262         11.98          0.3399
  9     XV9          2.325000         0.7382          9.92          0.3093
CØNS                39.449998
```

```
LQDAT   : 1  DEP-VAR=10:  YV10  CL=0.90  DØF= 10  RSS=        43.5975

RESIDUALS?
= NØ
ENTER A 1,2,3 ØR 4FØR:  1-SELECT NEW DATA FILE
2-SELECT NEW VARIABLES, 3-SELECT NEW 'F' ØR 'CL',
4-END ØF RUN
= 3
ENTER F ØR CL
= F
ENTER CRITICAL-F
= 0.01
STEPWISE RESULTS?
= NØ

LQDAT   : 2  DEP-VAR=10:  YV10  CL=0.     DØF= 10  RSS=        43.5975

STEP   VAR:LABEL    F-CRIT   DØF      R-SQ         SEE
  6     + 4:  XV4     0.01     9     0.9275      1.9081
  7     + 7:  XV7     0.01     8     0.9509      1.6654
  8     + 8:  XV8     0.01     7     0.9563      1.6605
  9     + 6:  XV6     0.01     6     0.9585      1.7686


VAR  LABEL       CØEFFICIENT      STD-ERR     F-RATIØ       BETA-WT
  1    XV1        -2.700000       0.4421       37.29        -0.5080
  2    XV2        -3.262500       0.4421       54.45        -0.6138
  3    XV3        -1.275000       0.4421        8.32        -0.2399
  4    XV4         0.837500       0.4421        3.59         0.1930
  5    XV5         1.837500       0.4421       17.27         0.4234
  6    XV6         0.250000       0.4421        0.32         0.0576
  7    XV7         1.150000       0.6253        3.38         0.1530
  8    XV8         0.550000       0.6253        0.77         0.0732


  9    XV9         2.325000       0.6253       13.83         0.3093
CØNS              37.999999

LQDAT   : 2  DEP-VAR=10:  YV10  CL=0.     DØF=  6  RSS=        18.7675

RESIDUALS?
= YES

ØBS      Y-ØBS        Y-CALC        ERRØR      %-ERR       C-ER:2
  1     52.5000      52.1875       0.3125      0.60        0.0977
  2     44.4000      43.3875       1.0125      2.33        1.1228
  3     40.6000      38.7125       1.8875      4.88        4.6855
  4     36.0000      34.5125       1.4875      4.31        6.8982
  5     43.0000      43.8875      -0.8875     -2.02        7.6858
  6     36.0000      37.2875      -1.2875     -3.45        9.3435
  7     39.3000      39.7125      -0.4125     -1.04        9.5136
  8     38.0000      37.7125       0.2875      0.76        9.5963
  9     38.0000      38.0000       0.0000      0.00        9.5963

 10     37.4000      38.0000      -0.6000     -1.58        9.9563
 11     35.5000      35.9500      -0.4500     -1.25       10.1588
 12     46.6000      46.7500      -0.1500     -0.32       10.1813
 13     37.5000      38.8250      -1.3250     -3.41       11.9369
 14     52.6000      51.8750       0.7250      1.40       12.4625
 15     37.9000      36.4500       1.4500      3.98       14.5650
 16     39.5000      41.5500      -2.0500     -4.93       18.7675
ENTER A 1,2,3 ØR 4
= 4

PRØGRAM STØP AT 4040
*
```

STADES

This ASCII file provides a series of instructions for the following statistical programs:

Program	Purpose
UNISTA	Description of Univariate Data
CURFIT	Fits Six Different Curves (Least squares)
MULFIT	Multiple Linear Fit - with Transformations
COLINR	Confidence Limits on Simple Linear Regression

INSTRUCTIONS

This program does not provide a solution to problems; it simple provides the time sharing user with a series of instructions which define the input to the above programs.

PROGRAM LISTING

The following is a copy of the printout that will appear at your teleprinter.

```
*LIST

                    S T A D E S

THESE ARE INSTRUCTIONS FOR THE FOLLOWING STATISTICAL PROGRAMS:

    PROGRAM                    PURPOSE

    UNISTA       DESCRIPTION OF UNIVARIATE DATA
    CURFIT       FITS SIX DIFFERENT CURVES (LEAST SQUARES)
    MULFIT       MULTIPLE LINEAR FIT - WITH TRANSFORMATIONS
    COLINR       CONFIDENCE LIMITS ON SIMPLE LINEAR REGRESSION

THE INPUT FORMAT IS IDENTICAL FOR THESE FOUR PROGRAMS.
LINE NUMBERS 41-699 ARE FREE FOR USE AS DATA STATEMENTS, AS REQUIRED.
ALL VALUES FOR ONE VARIABLE ARE GIVEN IN DATA STATEMENTS, FOLLOWED BY
ALL DATA FOR A SECOND VARIABLE, AND SO ON.
ONE WAY TO ARRANGE INPUT IS TO PUT ALL VALUES OF VARIABLE 1 AT LINES
100-199, ALL FOR VARIABLE 2 AT LINES 200-299, ETC.
IF SEPARATE TAPES ARE MADE FOR EACH VARIABLE, THE FULL USE OF INPUT
FLEXIBILITY CAN BE OBTAINED.

UNISTA ACCEPTS UP TO 300 OBSERVATIONS ON ONE VARIABLE.

CURFIT ACCEPTS UP TO 200 OBSERVATIONS ON TWO VARIABLES.  THE DATA FOR
       THE DEPENDENT VARIABLE (Y) IS SUPPLIED FIRST, FOLLOWED BY THE
       DATA FOR THE INDEPENDENT VARIABLE (X).

MULFIT ACCEPTS UP TO 61 OBSERVATIONS ON UP TO 6 VARIABLES.  THE ORDER
       OF VARIABLES IN THE INPUT IS IMPORTANT ONLY IF TRANSFORMATIONS
       ARE NOT BEING USED.  IF NO TRANSFORMED VARIABLES ARE GIVEN, THEN
       THE PROGRAM FITS A LINEAR FUNCTION FOR THE FIRST DATA VARIABLE
       IN TERMS OF THE FOLLOWING VARIABLES.  BY USING TRANSFORMS, ANY
       FUNCTION OF ANY (OR ALL) OF THE INPUT VARIABLES MAY BE USED AS
       ANY VARIABLE IN THE ANALYSIS.
       TRANSFORMS ARE ENTERED BEFORE THE RUN, AT LINES 1000-1099 AS:

       10XX LET X(J) = ANY FUNCTION OF V(1), V(2),...,V(R)

       WHERE X(J) IS THE J-TH VARIABLE IN THE ANALYSIS (J=1 IS THE
       DEPENDENT VARIABLE), AND THE V(I) ARE THE INPUT DATA VARIABLES.
       THUS, ONE MIGHT ENTER DATA FOR THREE VARIABLES V(1), V(2),
       AND V(3), BUT RUN THE ANALYSIS WITH X(1)=V(2) AS THE DEPENDENT
       VARIABLE, AND THREE 'INDEPENDENT' VARIABLES, BY ENTERING
       TRANSFOROATIONS LIKE THE FOLLOWING AT LINES 1000-1003:

       1000 LET X(1)=V(2)
       1001 LET X(2)=LOG(V(1)+V(3))
       1002 LET X(3)=V(1)+X(2)*EXP(-1-V(3))
       1003 LET X(4)=(V(3)-1)*2

       NOTE:  IF TRANSFORMS ARE USED, THE PROGRAM EXPECTS AT LEAST
       THREE SETS OF INPUT VARIABLES.  IF THERE ARE ONLY TWO SETS OF
       VALID DATA, A SET OF ZEROS MAY BE INPUT FOR THE THIRD VARIABLE.
       HOWEVER, THIS DUMMY SET SHOULD NOT BE USED IN DEFINING THE
       TRANSFORMATIONS.

COLINR ACCEPTS UP TO 250 OBSERVATIONS ON TWO VARIABLES.
       IT ASSUMES THAT THE FIRST DATA VARIABLE IS THE DEPENDENT ONE,
       AND ALSO ALLOWS FOR THE ADDITION OF ADDITIONAL VALUES OF THE
       INDEPENDENT VARIABLE IN THE DATA LIST, AS IN CURFIT.

                          E N D

READY

*
```

STAT

This FORTRAN program processes a series of data points and computes several fundamental statistics of the data.

METHOD

The number of data points and print increment are input prior to the date. All input is free field.

The output consists of:

1. number of observations
2. mean
3. median
4. range
5. biased variance
6. biased standard deviation
7. unbiased variance
8. unbiased standard deviation
9. skewness
10. kurtosis
11. upper confidence limit on mean
12. lower confidence limit on mean
13. upper confidence limit on variance
14. lower confidence limit on variance

Following this are 5 columns of output in increments as specified: input data, ordered data, cumulative percentages, cumulative percentages of normal distribution, difference of the last two columns. A test such as the "Kolmogorov-Smirnov one sample test" could be applied to the last column to test for normality.

The latter output can be reduced by selecting an increment other than 1.

INSTRUCTIONS

After the program has been compiled, the routine types -

IS DATA TO BE READ FROM A FILE, TYPE YES, NO OR STOP =

If the user types NO, the program prints -

TYPE NUMBER OF DATA POINTS (N), AND THE PRINT INCREMENT =

After this request is fulfilled, the program prints -

TYPE IN DATA POINTS =

After the data points are input, the output, as described above, is printed. If the response is YES, the program prints -

TYPE FILENAME, 1 TO 8 CHARACTERS =

After this request is fulfilled, the output, as described above, is printed.

The data on the file must be in the same order and format as the data input from the keyboard.

The program will request another set of data by printing IS DATA TO BE READ FROM A FILE, TYPE YES, NO, OR STOP. If user wishes to terminate, he types STOP, otherwise he supplies the necessary input.

RESTRICTION

The subprogram ANPF and ERRF must be used with this program as shown in the sample problem.

SAMPLE SOLUTION

```
*RUN STAT;ANPF;ERRF



IS DATA TØ BE READ FRØM A FILE, TYPE YES,NØ ØR STØP
= NØ
TYPE NUMBER ØF DATA PØINTS (N), AND THE PRINT INCREMENT
= 5 1
TYPG IN N DATA PØINTS

= 4 2 3 5 1

              DESCRIPTIVE STATISTICS FØR A SINGLE SAMPLE

NUMBER ØF ØBSERVATIØNS ........................................   5

MEAN = XBR = SUM(X(I))/N .................................. 3.000000E+00

MEDIAN .................................................... 3.000000E+00

RANGE ..................................................... 4.000000E+00

UNADJUSTED (BIASED) VARIANCE =
 SUM((X(I)-XBR)**2)/N = UVR ............................... 2.000000E+00

UNADJUSTED STANDARD DEVIATIØN = SQRT(UVR) = S ............. 1.414214E+00

ADJUSTED (UNBIASED) VARIANCE = UVR*N/(N-1) = AVR .......... 2.500000E+00

ADJUSTED STANDARD DEVIATIØN = SQRT(AVR) ................... 1.581139E+00

SKEWNES = SUM((X(I)-XBR)**3)/(N*UVR**1.5) ................. 0.

KURTØSIS = SUM((X(I)-XBR)**4)/(N*UVR**2) .................. 1.700000E+00
```

```
UPPER CØNFIDENCE LIMIT ØN MEAN =
       XBR          + T(N-1 D.F.)*   S/SQRT(N-1)
    3.000000E+00                     7.071068E-01

LØWER CØNFIDENCE LIMIT ØN MEAN =
       XBR          - T(N-1 D.F.)*   S/SQRT(N-1)
    3.000000E+00                     7.071068E-01

UPPER CØNFIDENCE LIMIT ØN VARIANCE =
 N*S**2 DIVIDED BC CHI-SQUARE (RIGHT
 TAIL CRITICAL REGIØN) WITH N-1 D.F.,  N*S**2 ............ 1.000000E+01

LØWER CØNFIDENCE LIMIT ØN VARIANCE =
 N*S**2 DIVIDED BY CHI-SQUARE (LEFT
 TAIL CRITICAL REGIØN) WITH N-1 D.F.,  N*S**2 ............ 1.000000E+01

                ****************************************
```

	DATA	ØRDERED DATA	CUMULATIVE PERCENTAGE	CUMULATIVE NØRMAL	DIFFERENCE
1	4.000000E+00	1.000000E+00	16.667	7.865	-8.8017
2	2.000000E+00	2.000000E+00	33.333	23.975	-9.3583
3	3.000000E+00	3.000000E+00	50.000	50.000	0.0000
4	7.000000E+00	4.000000E+00	66.667	76.025	9.3583
5	1.000000E+00	5.000000E+00	83.333	92.135	8.8017

```
IS DATA TØ BE READ FRØM A FILE, TYPE YES,NØ ØR STØP
= STØP

PRØGRAM STØP AT 710
*
```

STAT01

This BASIC program computes the means, standard error of means, mean difference, standard error of difference, and T-ratio for two groups of paired data. The groups have an equal variance.

INSTRUCTIONS

Within lines 900 to 997, use DATA statements to enter the data as ordered pairs.

SAMPLE PROBLEM

Analyze the two groups of data

(1, 2, 3, 4, 5, 6) and (2, 3, 5, 5, 6, 7)

SAMPLE SOLUTION

```
*900 DATA 1,2, 2,3, 3,5, 4,5, 5,6, 6,7
*RUN


GRØUP          NUMBER          MEAN          STD DEVIATIØN     STD ERRØR MEAN
  1              6            3.5               1.870829           .7637626

GRØUP          NUMBER          MEAN          STD DEVIATIØN     STD ERRØR MEAN
  2              6            4.666667          1.861899           .7601169

 MEAN DIFFERENCE         VARIANCE ØF DIFF,      STD ERRØR ØF DIFF,
 -1.166667                   .1666667                 .1666667

T-RATIØ =    -7 ØN     5 DEGREES ØF FREEDØM.


READY
*
```

STAT02

This BASIC program computes the means, variances, and T-ratio for two groups of unpaired data. The groups may have unequal variances.

INSTRUCTIONS

Within lines 900 to 998, use DATA statements to enter the data. Flag the end of the first group with a 999999. Then enter the second group, again ending with a 999999.

SAMPLE PROBLEM

Analyze the two groups

(73, 43, 47, 53, 58, 47, 52, 38, 61, 56, 56, 34, 55, 65, 75)

and

(51, 41, 43, 41, 47, 32, 24, 43, 53, 52, 57, 44, 57, 40, 68)

SAMPLE SOLUTION

```
*900 DATA 73,43,47,53,58,47,52,38,61,56,56,34,55,65,75,999999
*910 DATA 51,41,43,41,47,32,24,43,53,52,57,44,57,40,68,999999
*RUN


           GRØUP          NUMBER            MEAN       VARIANCE      STD. DEV.

               1              15            54.2       134.0286       11.57707
               2              15            46.2       116.0286       10.77166

MEAN DIFF.        VAR. DIFF.      STD. DEV. DIFF.
               8     16.67048          4.082949

T RATIØ              1.959368     ØN    27.85566 DEGREES ØF FREEDØM.

READY
*
```

STAT04

This BASIC program computes chi-square values for any number of 2 x 2 tables.

INSTRUCTIONS

Within lines 900 to 997, use DATA statements to enter the tables by horizontal rows; first Table 1, then Table 2, etc...

SAMPLE PROBLEM

Perform a chi-square analysis on the following 2 x 2 tables

$$\begin{pmatrix} 6 & 10 \\ 8 & 16 \end{pmatrix} \begin{pmatrix} 14 & 8 \\ 41 & 8 \end{pmatrix} \begin{pmatrix} 8 & 16 \\ 12 & 17 \end{pmatrix}$$

SAMPLE SOLUTION

```
*900 DATA6,10,8,16, 14,8,41,8, 8,16,12,17
*RUN


TABLE                           CHI SQUARE

            6          10
            8          16        .0732601

EXACT PRØBABILITY IS   .78316

TABLE                           CHI SQUARE

           14           8
           41           8        3.491921

EXACT PRØBABILITY IS   .05856

TABLE                           CHI SQUARE

            8          16
           12          17        .3618251

EXACT PRØBABILITY IS   .55491


READY
*
```

STAT05

This BASIC program calculates a confidence interval for the mean of a set of data using the sign test.

INSTRUCTIONS

Within lines 900 to 9000, use DATA statements to enter the data. The program can handle up to 125 data points. Specify either the confidence level (as 95%) or the critical value (from a table of critical values, see Kurtz, "Basic Statistics"). The program uses fractional counts for the critical value option.

SAMPLE PROBLEM

Find the sign test confidence interval for the mean at the 95% level for:

12, 11, 12, 12, 10, 9, 12, 10, 9, 12, 13, 13, 9, 10, 12, 14

SAMPLE SOLUTION

```
*900 DATA 12,11,12,12,10,9,12,10,9,12,13,13,9,10,12,14
*RUN


ENTER '1' TØ USE CØNFIDENCE LEVEL, '2' TØ USE CRITICAL VALUE ?1

ENTER THE CØNFIDENCE LEVEL ?95

THE CRITICAL VALUE FØR THIS LEVEL IS     3


THE CØNFIDENCE INTERVAL BY THE SIGN TEST:
LØWER LIMIT =     10 UPPER LIMIT =     12


READY
*
```

STAT06

This BASIC program calculates the confidence limits for a set of data using the Wilcoxon signed rank sum procedure with fractional counts.

INSTRUCTIONS

Within lines 900 to 998, use DATA statements to enter the data. The first datum is N, the number of data values in the set. The second is the critical value C, (from a table of critical values, see Kurtz, "Basic Statistics"). Then enter the set of data itself.

SAMPLE PROBLEM

Using C = 13.8, analyze: 20.1, 21.0, 20.4, 18.1, 19.0, 17.8, 20.3, 19.2, 21.5, 19.7, 20.0, 18.2.

SAMPLE SOLUTION

```
*900 DATA 12,13.8
*910 DATA 20.1,21.0,20.4,18.1,19.0,17.8
*920 DATA 20.3,19.2,21.5,19.7,20.0,18.2
*RUN


CØNFIDENCE INTERVAL BY SIGNED RANK SUM, FRACTIØNAL COUNT.
LØWER LIMIT IS 18.9     UPPER LIMIT IS   20.35

READY
*
```

STAT08

This BASIC program compares two groups of data using the median test.

INSTRUCTIONS

Starting on line 900, use DATA statements to enter the data in the following order. First enter the number of entries in the first group, and the number of entries in the second group. On consecutive lines enter the first group itself, followed by the second group. The program prints out the chi-square statistic of the 2 x 2 table with 1 degree of freedom.

SAMPLE PROBLEM

Compare (1, 2, 3, 4, 5, 6, 7) and (1, 2, 5, 6, 7, 10, 10, 16).

SAMPLE SOLUTION

```
*900 DATA 7,8
*910 DATA 1,2,3,4,5,6,7
*920 DATA 1,2,5,6,7,10,10,16
*RUN


TWO SAMPLE MEDIAN TEST.
GROUP 1        4      3
GROUP 2        3      5
CHI-SQUARE =       .0585938

READY
*
```

STAT09

This BASIC program compares two groups of data by the Mann-Whitney two sample rank test.

INSTRUCTIONS

Starting on line 900, use DATA statements to enter the data in the following order. First enter the values of M, N, and C. M and N are the two sample sizes, and C is the critical value for the Mann-Whitney test with fractional counts. If M or N is greater than 30, the DIM statement in line 90 should be changed. On consecutive lines, after M, N, and C, enter the first series of test data, followed by the second series.

SAMPLE PROBLEM

Compare (190, 160, 160, 140) and (117, 120, 120, 145, 147, 150) with a critical value of 2.5.

SAMPLE SOLUTION

```
*900 DATA 4,6,2.5
*910 DATA 190,160,160,140
*920 DATA 117,120,120,145,147,150
*RUN


CONFIDENCE INTERVAL BY RANK SUM TEST.
LOWER LIMIT =           2.5    UPPER LIMIT =    57.5

READY
*
```

STAT11

This BASIC program computes the Spearman rank correlation coefficient for two series of data.

INSTRUCTIONS

Within lines 900 to 9000, use DATA statements to enter the elements of the series by groups, i.e., X1, Y1, X2, Y2, etc.

SAMPLE PROBLEM

Analyze the groups

1, 2, 3, 1, 4, 5, 7, 8

and

3, 1, 2, 4, 1, 6, 8, 9

SAMPLE SOLUTION

```
*900 DATA 1,3,  2,1,  3,2
*910 DATA 1,4,  4,1,  5,6
*920 DATA 7,8,  8,9
*RUN


SPEARMAN RANK CØRRELATIØN CØEFFICIENT
  R =     .5903614


READY
*
```

STAT12

This BASIC program computes the correlation matrix for N series of data.

INSTRUCTIONS

Starting on line 900, use DATA statements to enter the data in the following order:

1. The number of series
2. The number in each series
3. The data by group (not by series)

RESTRICTIONS

To use more than 25 series, change the dimension statement in line 100. Each series may have an arbitrary number of elements.

SAMPLE PROBLEM

Compute the correlation matrix for the series:

(1, 2, 3, 4, 5), (2, 4, 6, 8, 10), (1, 1, 2, 5, 6)

SAMPLE SOLUTION

```
*900 DATA 3,5
*910 DATA 1,2,1
*911 DATA 2,4,1
*912 DATA 3,6,2
*913 DATA 4,8,5
*914 DATA 5,10,6
*RUN


  VARIABLE             MEAN          VARIANCE        STD. DEV.
         1                3               2.5        1.581139
         2                6                10        3.162278
         3                3               5.5        2.345208

THE CØRRELATIØN MATRIX

         1                1          .9438798

                          1          .9438798

                                            1


READY
*
```

STAT13

This BASIC program computes the analysis of variance table for a one-way, completely randomized design. It is the same as STAT33 except for the data entry method.

INSTRUCTIONS

Starting on line 900, use DATA statements to enter the data in the following order:

1. The total number of observations
2. The number of different treatments
3. The number of observations in each treatment
4. The observations themselves, first for treatment 1, then treatment 2, etc.

If the number of observations for any treatment is greater than 20, or the number of treatments is greater than 10, change the DIM statement in line 100.

SAMPLE PROBLEM

Compute the analysis of variance table for:

Treatments

	83	84	86	89	90
	85	85	87	90	92
		85	87	90	
		86	87	91	
		86	88		
Observations		87	88		
			88		
			88		
			88		
			89		
			90		

SAMPLE SOLUTION

```
*900 DATA 25,5
*905 DATA 2,6,11,4,2
*910 DATA 83,85
*920 DATA 84,85,85,86,86,87
*930 DATA 86,87,87,87,88,88,88,88,88,89,90
*940 DATA 89,90,90,91
*950 DATA 90,92
*RUN


ANØVA TABLE:

ITEM                        SS                DF              MS

GRAND TØTAL             191791                25
GRAND MEAN             191668.8                1
TREATMENTS             99.02344                4          24.75586
ERRØR                  23.13672               20          1.156836

F =    21.39963 ØN     4 AND    20 DEGREES ØF FREEDØM.

EXACT PRØB. ØF F=    21.39963 WITH (      4 ,     20  ) D.F. IS
 1.00000E-05


READY
*
```

STAT14

This BASIC program produces an analysis of a variance table, and F-ratios for treatments and blocks of a randomized complete block design.

INSTRUCTIONS

Starting on line 900, use DATA statements to enter the data in the following order. First enter T, the number of treatments, then B, the number of blocks. On consecutive lines, enter the observations for each treatment. The observations for treatment 1 are entered first, beginning with the observation for block 1, then block 2, etc., up to block B. Continue on the next line with the observations for treatment 2, etc.

SAMPLE PROBLEM

Analyze

treatments \ blocks	1	2	3	4
1	4	1	-1	0
2	1	1	-1	-2
3	0	0	-3	-2
4	0	-5	-4	-4

SAMPLE SOLUTION

```
*900 DATA 4,4
*910 DATA 4,1,-1,0
*920 DATA 1,1,-1,-2
*930 DATA 0,0,-3,-2
*940 DATA 0,-5,-4,-4
*RUN


DØ YØU WANT INSTRUCTIØNS? ENTER 1 FØR YES ØR 0 FØR NØ
?0
ANØVA TABLE:
   ITEM                SS               DF          MS


GRAND TØTAL             95               16
GRAND MEAN           14.0625              1
TREATMENTS           30.6875              3          10.22917
BLØCKS               38.6875              3          12.89583
ERRØR                11.5625              9          1.284722

F-RATIØ FØR TREATMENTS =    7.962162 , ØN    3 AND    9 DEGREES ØF FREEDØM

EXACT PRØB. ØF F=   7.962162 WITH (     3 ,      9  ) D.F. IS    .00706

F-RATIØ FØR BLØCKS =    10.03784 , ØN    3 AND    9 DEGREES ØF FREEDØM.

EXACT PRØB. ØF F=   10.03784 WITH (     3 ,      9  ) D.F. IS    .00357

READY
*
```

STAT15

This BASIC program computes the analysis of variance table for a simple Latin square design.

INSTRUCTIONS

Starting on line 900, use DATA statements to enter the data. The first datum is the number of rows, then the matrix of treatment assignments by rows, and finally the matrix of data by rows.

SAMPLE PROBLEM

Compute the analysis of variance table for the design whose treatment and data matrices are given below.

$$\begin{bmatrix} 3 & 4 & 1 & 2 \\ 2 & 3 & 4 & 1 \\ 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{bmatrix} \qquad \begin{bmatrix} -1 & -2 & 0 & -5 \\ 1 & -1 & -2 & 0 \\ 4 & 1 & -3 & -4 \\ 0 & 1 & 0 & -4 \end{bmatrix}$$

SAMPLE SOLUTION

```
*900 DATA 4
*910 DATA 3,4,1,2
*911 DATA 2,3,4,1
*912 DATA 1,2,3,4
*913 DATA 4,1,2,3
*920 DATA -1,-2,0,-5
*921 DATA 1,-1,-2,0
*922 DATA 4,1,-3,-4
*923 DATA 0,1,0,-4
*RUN

ITEM            SUM-SQR          DEG. FREE.       MEAN-SQR         F-RATIØ
----            -------          ----------       --------         -------
RØWS              6.1875                  3          2.0625         2.302326
CØLS             38.6875                  3        12.89583         14.39535
TREATS           30.6875                  3        10.22917          11.4186
ERRØR              5.375                  6        .8958333

EXACT PRØBABILITY ØF F =    2.302326   WITH (    3  ,     6   )   D.F.
IS   .18512


EXACT PRØBABILITY ØF F =    14.39535   WITH (    3  ,     6   )   D.F.
IS   .01776


EXACT PRØBABILITY ØF F =  11.4186   WITH (    3  ,     6   )   D.F.
IS   .02309


READY
*
```

STAT16

This BASIC program computes the analysis of variance table for a simple Graeco-Latin square design.

INSTRUCTIONS

Starting on line 1000, use DATA statements to enter the data in the following order:

1. N = the number of treatments
2. The matrix giving the Latin treatment assignments (numbered from 1 to N only)
3. The matrix giving the Graeco treatments
4. The matrix of data.

All matrices should be entered by rows. In the present version of this program, N must be no greater than 10.

SAMPLE PROBLEM

Latin Treatment Assignments

1	2	3	4
2	1	4	3
3	4	1	2
4	3	2	1

Graeco Treatment Assignments

4	3	2	1
2	1	4	3
1	2	3	4
3	4	1	2

Data

27	47	35	42
47	85	23	47
65	49	23	62
12	14	19	23

SAMPLE SOLUTION

```
*1000 DATA 4
*1001 DATA 1,2,3,4,  2,1,4,3,  3,4,1,2,  4,3,2,1
*1005 DATA 4,3,2,1,  2,1,4,3,  1,2,3,4,  3,4,1,2
*1011 DATA 24,47,35,42
*1012 DATA 47,85,23,47
*1013 DATA 65,49,23,62
*1014 DATA 12,14,19,23
*RUN
```

ITEM	SUM-SQR	DEG. FREE.	MEAN-SQR	F-RATIO
RØWS	2940.188	3	980.0625	8.961425
COLS	1258.188	3	419.3958	3.834841
TREAT L	318.6875	3	106.2292	.9713306
TREAT G	1208.688	3	402.8958	3.68397
ERRØR	656.1875	6	109.3646	

```
EXACT PRØBABILITY ØF F =     8.961425   WITH (     3  ,     6   ) D.F.
IS   .01311

EXACT PRØBABILITY ØF F =     3.834841   WITH (     3  ,     6   ) D.F.
IS   .07606

EXACT PRØBABILITY ØF F =     .9713306   WITH (     3  ,     6   ) D.F.
IS   .46695

EXACT PRØBABILITY ØF F =  3.68397   WITH (     3  ,     6   ) D.F.
IS   .08186


READY
*
```

STAT18

This BASIC program computes the analysis of variance table and the F-ratio for treatments for a Youden square design. Sum-of-squares for treatments is adjusted because of incompleteness.

INSTRUCTIONS

Starting on line 900, use DATA statements to enter the data in the following order:

1. N = the number of rows and treatments
2. K = the number of columns and replications of each treatment in the experiment
3. The M(I, H) matrix by rows. M(I, H) = J if treatment J appears in row I, Column H. This matrix is N by K.
4. The N(I, J) matrix. N(I, J) = 1 if treatment J appears in row I, and is 0 otherwise. This matrix is N by N.
5. The matrix, X, of observations.

Enter all matrices by rows. In the present version of this program, N must be no greater than 10.

SAMPLE PROBLEM

N = 4, K = 3.

$$M(I,H) = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 1 & 2 \\ 2 & 3 & 4 \\ 3 & 4 & 1 \end{bmatrix}, \quad N(I,J) = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}$$

$$X = \begin{bmatrix} 2 & 1 & 0 \\ -2 & 2 & 2 \\ -1 & -1 & -3 \\ 0 & -4 & 2 \end{bmatrix}$$

SAMPLE SOLUTION

```
*900 DATA 4,3
*910 DATA 1,2,3
*911 DATA 4,1,2
*912 DATA 2,3,4
*913 DATA 3,4,1
*920 DATA 1,1,1,0
*921 DATA 1,1,0,1
*922 DATA 0,1,1,1
*923 DATA 1,0,1,1
*930 DATA 2,1,0
*931 DATA -2,2,2
*932 DATA -1,-1,-3
*933 DATA 0,-4,2
*RUN


ANØVA TABLE:

    ITEM                SS              DF        MS

GRAND TØTAL              48              12
GRAND MEAN         .3333333               1
TREATMENTS         31.08333               3        10.36111
RØWS               13.66667        3      ...SS NØT ADJUSTED...
CØLUMNS            1.166667               2        .5833333
ERRØR                  1.75               4           .4375

TREATMENT F-RATIØ =    23.68254 , ØN     3 AND     4 DEGREES ØF FREEDØM.

EXACT PRØB. ØF F=   23.68254 WITH (      3 ,      4  ) D.F. IS    .00727

IF MSC/MSE =    1.333333  IS NØT SIGNIFICANT, IT MAY BE
DESIRABLE TØ PØØL CØLUMN AND ERRØR SS TØ ØBTAIN AS AN
ERRØR MS ESTIMATE    .5833334 WITH    5 DEGREES ØF FREEDØM.

READY
*
```

STAT33

This BASIC program computes the analysis of variance table for a one-way, completely randomized design. It is the same as STAT13, except for the data entry method. STAT13 requires each observation to be listed. STAT33 allows grouping of identical observations within each treatment. The observation value is entered, followed by the number of times that value occurs in a treatment.

INSTRUCTIONS

Starting on line 900, use DATA statements to enter the data in the following order:

1) A, the total number of observations

2) M, the number of different treatments

3) N(1), , , , , N(M), where N(J) is the number of observations in treatment J

4) V(1), , , , , V(M), where V(J) is the number of different values of the observations in treatment J

5) The observations themselves, first for treatment 1, then for treatment 2, etc. For each treatment these are listed as pairs of numbers. The first number is an actual observation value taken on B. The second number is the number of observations in that treatment, having that value.

If any N(J) >20, change the dimensions in line 100.

If M >10, change the dimensions in line 100.

SAMPLE PROBLEM

Compute the analysis of variance table for:

Treatments

Observations	83	84	86	89	90
	85	85	87	90	92
		85	87	90	
		86	87	91	
		86	88		
		87	88		
			88		
			88		
			88		
			89		
			90		

SAMPLE SOLUTION

```
*900 DATA 25,5
*905 DATA 2,6,11,4,2
*907 DATA 2,4,5,3,2
*910 DATA 83,1, 85,1
*920 DATA 84,1, 85,2, 86,2, 87,1
*930 DATA 86,1, 87,3, 88,5, 89,1, 90,1
*940 DATA 89,1, 90,2, 91,1
*950 DATA 90,1, 92,1
*RUN


ANØVA TABLE:

ITEM                    SS              DF          MS

GRAND TØTAL          191791             25
GRAND MEAN           191668.8            1
TREATMENTS           99.02344            4        24.75586
ERRØR                23.13672           20        1.156836

F =    21.39963 ØN     4 AND    20 DEGREES ØF FREEDØM.

EXACT PRØB. ØF F=    21.39963 WITH (      4 ,      20   ) D.F. IS
1.00000E-05


READY
*
```

STATAN

This BASIC program performs a statistical analysis on Data for one variable. It computes 34 different measures for an array of weighted (as with frequencies) or unweighted values of the variable. It also gives a 10-class frequency distribution summary, and a recapitulation of the input data in terms of deviations from the mean and as an ordered array.

INSTRUCTIONS

To use this program supply data in either of these 2 formats as determined by the problem:

1) FOR UNWEIGHTED VALUES:

```
10 DATA 0, X(1), X(2), X(3),......
```

2) FOR DATA WITH WEIGHTS OR FREQUENCIES:

```
10 DATA 1, X(1), F(1), X(2), F(2), X(3), F(3),......
```

Where the initial 0 or 1 signals the presence or the absence of weights. Lines 11 through 99 are available for additional input data.

NOTES:

This program produces output corresponding to the listing in the National Bureau of Standards Handbook No. 101.

There can be no more than 99 values entered. If the values are weighted the maximum amount of data cannot exceed 99 values and 99 weights.

Additional instructions may be obtained from the listing.

SAMPLE PROBLEM

Suppose you had obtained the following data points as the result of an experiment:

261.4	252.1	255.5	258.3	253.2
270.8	268.3	249.6	256.3	266.4
265.4	250.3	280.9	259.3	
261.4	272.3	270.3	270.1	
258.1	262.8	263.2	259.3	

To use this program to get the data's statistical characteristics you would prepare the following data tape:

```
10 DATA 0, 261.4, 270.8, 265.4, 261.4, 258.1, 252.1, 268.3, 250.3
11 DATA 272.3, 262.8, 255.5, 249.6, 280.9, 270.3, 263.2, 258.3, 256.3
12 DATA 259.3, 270.1, 259.3, 253.2, 266.4
```

Where the first data value (0) indicates unweighted data.

SAMPLE SOLUTION

```
*10 DATA 0,261.4,270.8,265.4,261.4,258.1,252.1,268.3,250.3,272.3
*11 DATA 262.8,255.5,249.6,280.9,270.3,263.2,258.3,256.3,259.3
*12 DATA 270.1,259.3,253.2,266.4
*RUN



COMPUTATIONS ON THE DATA ARRAY:

NUMBER OF VALUES =    22
NUMBER OF NONZERO WEIGHTS =    22
SUM OF WEIGHTS =    22
SUM OF VALUES =   5765.3
WEIGHTED MEAN =    262.0591
UNWEIGHTED MEAN =    262.0591
MINIMUM VALUE=   249.6
MAXIMUM VALUE=   280.9
RANGE = 31.3
WEIGHTED SUM OF SQUARES = 1512182
VARIANCE = 63.4711
STANDARD DEVIATION =    7.966875
STANDARD DEVIATION OF MEAN =    1.698544
COEFFICIENT OF VARIATION =    3.040107
STUDENT'S T =    154.2846
MEAN SQUARE SUCCESSIVE DIFFERENCES =    145.4
(MEAN SQ SUCC DIFF)/(VARIANCE) =    2.290806
MEDIAN =    261.4
NUMBER OF RUNS UP AND DOWN =    11
EXPECTED NUMBER OF RUNS =    14.33333
STD DEV OF NUMBER OF RUNS =    1.894436
(ACTUAL RUNS - EXP RUNS)/(STD DEV) =    1.759538

FREQUENCY DISTRIBUTION (TEN EQUAL CLASSES):
    3     2     3     4     2     3     3     1     0     1

COMPUTATIONS ON DEVIATIONS FROM MEAN:

NUMBER OF + SIGNS IN DEVIATIONS =    10
NUMBER OF - SIGNS IN DEVIATIONS =    12
NUMBER OF RUNS (SIGN CHANGES + 1) =    11
EXPECTED NUMBER OF RUNS =    11.90909
STD DEVIATION OF NUMBER OF RUNS =    2.268828
(ACTUAL RUNS - EXP RUNS)/(STD DEV) =    0.4006875
TREND VALUE = -0.0576464
STD DEV OF TREND =    0.0612895
(TREND)/(STD DEV) = -0.9405591
BETA ONE =    0.1486213
BETA TWO =    2.679281
MEAN DEVIATION = 6.35537
```

SAMPLE SOLUTION

RECAPITULATION ØF INPUT:

VALUE	DEVIATIØNS	WEIGHTS	ØRDERED ARRAY
261.4	-0.6590729	1	249.6
270.8	8.740925	1	250.3
265.4	3.340927	1	252.1
261.4	-0.6590729	1	253.2
258.1	-3.959076	1	255.5
252.1	-9.959074	1	256.3
268.3	6.240925	1	258.1
250.3	-11.75908	1	258.3
272.3	10.24092	1	259.3
262.8	0.7409248	1	259.3
255.5	-6.559074	1	261.4
249.6	-12.45907	1	261.4
280.9	18.84093	1	262.8
270.3	8.240925	1	263.2
263.2	1.140926	1	265.4
258.3	-3.759075	1	266.4
256.3	-5.759075	1	268.3
259.3	-2.759075	1	270.1
270.1	8.040924	1	270.3
259.3	-2.759075	1	270.8
253.2	-8.859074	1	272.3
266.4	4.340927	1	280.9

READY
*

TAU

This FORTRAN program computes the Kendall Rank Correlation Coefficient or the Kendall Partial Rank Correlation Coefficient.

The Kendall Rank Correlation Coefficient is a measure of the association between two sets of ordered data with N elements in each set.

The Kendall Partial Rank Correlation Coefficient is a measure of the association between two sets of ordered data with N elements in each set while a third set of ordered data containing N elements is held constant.

METHOD

I. Kendall Rank Correlation Coefficient, τ_{xy}

Let x and y be two sets of ordered data each containing N elements. The values in each set are replaced by their ranks and a statistic S is calculated. This statistic reflects the magnitude of the difference between the rankings.

$$\tau_{xy} = \frac{S}{\sqrt{\frac{N}{2}(N-1) - T_x} \cdot \sqrt{\frac{N}{2}(N-1) - T_y}}$$

where N = number of elements in each set.

$T_x = \frac{1}{2}\Sigma\, t(t-1)$, t = the number of tied observations in each group of ties on the x variable.

$T_y = \frac{1}{2}\Sigma\, t(t-1)$, t = the number of tied observations in each group of ties on the y variable.

II. Kendall Partial Rank Correlation Coefficient, $\tau_{xy.z}$

Let x, y and z be three sets of ordered data each containing N elements. The correlation between x and y, holding z constant, is

$$\tau_{xy.z} = \frac{\tau_{xy} - \tau_{zy} \cdot \tau_{xz}}{\sqrt{(1 - \tau_{zy}^2)(1 - \tau_{zx}^2)}}$$

INSTRUCTIONS

After compilation the program requests -

DO YOU WISH TO USE THE PARTIAL COEFFICIENTS =

If the response is NO, the program requests -

NUMBER OF ELEMENTS IN EACH ARRAY =

followed by -

INPUT THE X, Y ARRAYS =

The response to this is the elements in the X array followed by the elements in the Y array.

The output consists of the Kendall Rank Correlation Coefficient, τ_{xy}.

If the response is YES, the program requests -

NUMBER OF ELEMENTS IN EACH ARRAY =

followed by -

INPUT THE X, Y, Z ARRAYS =

where X and Y are the variables whose relation is to be determined and Z is the variable held constant. The response to this is the elements in the X array, followed by the elements in the Y array, followed by the elements in the Z array.

The output consists of the Kendall Partial Rank Correlation Coefficient, $\tau_{xy.z}$.

RESTRICTION

N.LE. 500.

SAMPLE SOLUTION

```
*RUN
DØ YØU WISH TØ USE THE PARTIAL CØEFFICIENT
= NØ


NUMBER ØF ELEMENTS IN EACH ARRAY
= 12


INPUT THE X,Y ARRAYS
= 42 46 39 37 65 88 86 56 62 92 54 81
= 0  0  1  1  3  4  5  6  7  8  8 12


TAU =        0.38770175E+00


PRØGRAM STØP AT 470
*RUN
DØ YØU WISH TØ USE THG PARTIAL CØEFFICIENT
= YES


NUMBER ØF ELEMENTS IN EACH CRRAY
= 12


INPUT THE X,Y,Z ARRAYS
=  82  98  87  40 116 113 111  83  85 126
= 106 117  42  46  39  37- 65  88  86  56  62  92  54  81
=   0   0   1   1   3   4   5   6   7   8   8  12


TAU(XY.Z) =        0.61357089E+00


PRØGRAM STØP AT 470
*
```

TDIST

This FORTRAN function evaluates cumulative probabilities and percentage points of the T-distribution (2-tail).

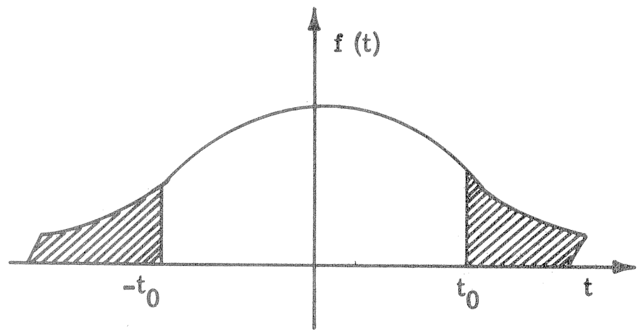
METHOD

The Student's "T" distribution is given in density form by:

$$f(t) = \frac{\Gamma\left[(r+1)/2\right]}{\sqrt{Hr}\;\Gamma(r/2)} \quad \frac{1}{(1+t^2/r)^{(r+1)/2}} \qquad -\infty < t < \infty$$

The parameter "r" is known as the degrees of freedom.

An article in the journal "Technometrics" by C. Y. Kramer (1966) described a method of approximating T. The program uses this method to compute the probability "outside" the region $-t_0 < t < t_0$.



The inverse problem of calculating t_0 given the probability in the "tails" was attacked by an iterative numerical technique known as "Steffensen's Method." This can be found in "Elements of Numerical Analysis" by P. Henrici, John Wiley and Sons, 1964 (pp. 90-95).

The technique is as follows:

Given $\quad t = t + F(t) - p = G(t)$

where F(t) is the two-tailed "T" probability and p is the desired probability, then generate the recursion sequence

$$t_{n+1} = G'(t_n)$$

where

$$G'(t) = \begin{cases} t - \dfrac{\left[G(t)-t\right]^2}{N(t)}, & N(t) \neq 0 \\ t, & N(t) = 0 \end{cases}$$

and $N(t) = G(G(t)) - 2G(t) + t$.

INSTRUCTIONS

This routine calls the function BETA, which must be executed jointly with TDIST.

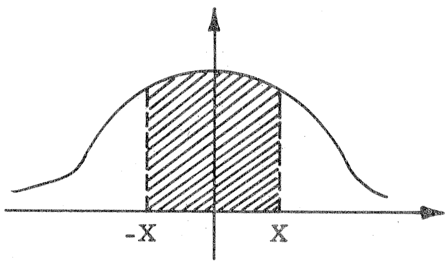
The calling sequence is:

Y = TDIST(IND,A,X)

where,

- IND = 0 corresponds to the evaluation of the cumulative probability, Y, given the percentage point, X, where

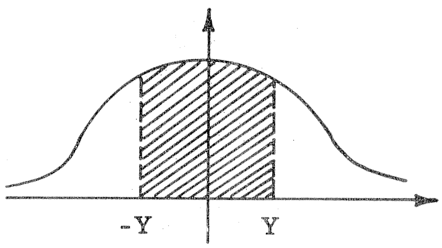
  $Y = Pr(|T| \leq X)$

  (Direct T-distribution)

  IND = 1 corresponds to the evaluation of the percentage point, Y, given the cumulative probability, X, such that

  $X = Pr(|T| \leq Y)$

  (Inverse T-distribution)

- A, floating point, is the number of degrees of freedom.
- X is as defined above.
- Y is the result returned as defined above.

RESTRICTIONS

- It has been observed that accuracy degenerates rapidly for percentages
- 99 or greater. $0 < A < 120$

SAMPLE SOLUTION

```
10 PRINT:"WHAT ARE IND,A,X";READ:IND,A,X
20 Y=TDIST(IND,A,X)
30 PRINT:"RESULTS",Y,X
40 STØP;END

READY

*RUN *;TDIST;BETA
WHAT ARE IND,A,X
= 0 1 .45
RESULTS    2.5964531E-01    4.5000000E-01

PRØGRAM STØP AT 40
*RUN *;TDIST;BETA
WHAT ARE IND,A,X
= 1 1 .25964531
RESULTS    4.5000004E-01    2.5964531E-01

PRØGRAM STØP AT 40
*
```

TESTUD

This BASIC program tests an unknown population mean using sample statistics.

INSTRUCTIONS

To use this program simply supply values for the 5 variables N, M, S, W, and X.

where:

```
N = Sample size
M = Sample mean
S = Sample standard deviation
W = Population size (0 if infinite)
X = The population mean to be tested
```

Additional instructions may be found in the listing.

SAMPLE PROBLEM

Determine the probability that the annual fallout will exceed or equal a critical amount.

During the early stages of a project to develop high-energy fuels for gas turbine engines, the question of liability for damages to agricultural crops in the vicinity of the outdoor test site arose. As the exhaust products were toxic to plant growth if applied in concentrations exceeding about 20 pounds per acre, a sampling experiment was established to measure the fallout at various distances from the test site. The greatest potential liability seemed to be from farmland located about a mile and a half downwind from the prevailing direction of local winds, since crop losses for any reason would doubtless be blamed on the "poison gases" which the local people were already grumbling about. By annualizing the results from the first eight test runs, a sample of eight readings averaging 13.4 pounds per acre with a standard deviation of 5.1 p.p.a. was available for the lawyers' consideration. Their question was simply this: is this evidence sufficient to deny a claim that the fallout actually will equal or exceed the critical value of p.p.a. ?

SAMPLE SOLUTION

```
*RUN


INSTRUCTIØNS ? (1=YES, 0=NØ) ..... WHICH ?0

N, M, S, W, X = ?8,13.4,5.1,0,20

BASED ØN STUDENT'S T-DISTRIBUTIØN WITH     7  DEGREES ØF FREEDØM,
THE PRØBABILITY ØF FINDING A SAMPLE MEAN THIS MUCH LESS THAN THE
PØPULATIØN MEAN IS 0.00059


READY
*
```

ANALYSIS OF RESULT

In terms of the statistical model, we conclude that such sample results as we observed would be extremely rare if the population mean were 20 p.p.a. It seems fairly safe to say that the annual fallout will be less than this critical amount.

UNIFM (UNIFMSOR)

UNIFM is a file containing two FORTRAN compatible, GMAP coded functions for calculating random numbers having a uniform (rectangular) distribution. The shift and add method is used. The UNIFMSOR file is the CARDIN source listing for UNIFM.

INSTRUCTIONS

The calling sequence is:

A = UNIFM1 (B) or A = UNIFM2 (B, C, D)

where:

A is the computed random number.
B is an arbitrary starting number.
C is the mean.
D is the width of the interval.

UNIFM1 assumes a mean of .5 and an interval width of 1. The starting number, B, is used to initialize the calculations of the random number. Subsequent calls to either UNIFM1 or UNIFM2 use the previous calculated random number in place of B.

METHOD

The equation[1] is:

X(I +1) = (2**7+1) * X(I) + C

where

C = (.5 + SQRT (3)/6)*2**35.

The actual number returned by UNIFM1 is:

R(I) = (X(I+1) MODULO 2**35) * (2**(-35)).

NOTE:

It has been shown that the shift and add method does not satisfy all the statistical tests for randomness. For other methods of evaluating UNIFM1 and UNIFM2, see the programs URAN and FLAT.

---

[1] A. Rotenberg, "A New Pseudo-Random Number Generator," J. ACM 7, P. 75-77, (1960)

R. R. Coveyou, "Serial Correlation in the Generation of Pseudo-Random Numbers," ibid p. 72-74.

SAMPLE PROBLEM

Calculate 10 random numbers with uniform distribution in the interval (0, 1) and 10 random numbers with uniform distribution in the interval (-1, 1).

SAMPLE SOLUTION

```
*LIST

10 B=2.583
20 DØ 10 I=1,10
30 A=UNIFM1(B)
40 C=UNIFM2(B,0.0,2.)
50 10 PRINT 20,A,C
60 20 FØRMAT(2F12.8)
70 STØP;END

READY

*RUN *;UNIFM
 0.12969759  0.03932790
 0.82532462 -0.48889613
 0.75487493 -0.66491617
 0.40158292  0.18574443
 0.26919080  0.02857788
 0.13194861  0.62009207
 0.28461419  0.00781051
 0.79245318 -0.96972769
 0.74123907 -0.18296778
 0.48725314  0.28866061

PRØGRAM STØP AT 70
*
```

UNISTA

This BASIC program provides a description of univariate data with up to 300 observations on one variable.

NOTE: This program may require excessive running time for large problems.

INSTRUCTIONS

Enter data on lines numbered 41-699.

Additional instructions may be obtained by listing the program STADES.

SAMPLE PROBLEM

Determine statistical characteristics for the following data points:

261.4	252.1	255.5	258.3	253.2
270.8	268.3	249.6	256.3	266.4
265.4	250.3	280.9	259.3	
261.4	272.3	270.3	270.1	
258.1	262.8	263.2	259.3	

SAMPLE SOLUTION

```
100 DATA 261.4, 270.8, 265.4, 261.4, 258.1, 252.1, 268.3, 250.3, 272.3
101 DATA 262.8, 255.5, 249.6, 280.9, 270.3, 263.2, 258.3, 256.3, 259.3
102 DATA 270.1, 259.3, 253.2, 266.4
*RUN


UNISTA


TYPICAL INTERVAL FØR FREQUENCY DISTRIBUTIØNS: L,U = ?260,280


S U M M A R Y   S T A T I S T I C S

    NUMBER ØF VARIATES =   22
       ARITHMETIC MEAN =   262.0591
    STANDARD DEVIATIØN =   7.784011
              VARIANCE =   60.59082
     CØEFF ØF VAR (PCT) =   2.97
     STANDARD SKEWNESS =    0.383
       STANDARD EXCESS =   -0.176

Ø R D E R   S T A T I S T I C S

      SMALLEST VARIATE =   249.6
          LØWER DECILE =   250.84
        FIRST QUARTILE =   256.1
                MEDIAN =   261.4
        THIRD QUARTILE =   268.75
          UPPER DECILE =   271.85
      LARGEST VARIATE =   280.9

           TØTAL RANGE =  31.3
          DECILE RANGE =   21.01
   SEMI-QUARTILE RANGE =   6.324999
     BØWLEY'S SKEWNESS =    0.162
      PEARSØN SKEWNESS =    0.254
```

SAMPLE SOLUTION

F R E Q U E N C Y   D I S T R I B U T I Ø N

FRØM	UP TØ BUT NØT INCLUDING	FREQUENCY	PERCENT FREQUENCY
220	240	0	0
240	260	10	45.455
260	280	11	50
280	300	1	4.545
300	320	0	0

C U M U L A T I V E   D I S T R I B U T I Ø N

VALUE	NUMBER LESS THAN VALUE	% LESS THAN VALUE	VARIATE SUM % LESS THAN VALUE
240	0	0	0
260	10	45.455	44.265
280	21	95.455	95.128
300	22	100	100

Ø R D E R E D   A R R A Y

249.6	258.1	262.8	270.1
250.3	258.3	263.2	270.3
252.1	259.3	265.4	270.8
253.2	259.3	266.4	272.3
255.5	261.4	268.3	280.9
256.3	261.4		

READY
*

URAN (URANSORC)

URAN is a file containing three FORTRAN compatible, GMAP coded routines for calculating random numbers having a uniform (rectangular) distribution. A mixed congruential method is used. The URANSORC file is the CARDIN listing for URAN.

INSTRUCTIONS

Two of the routines are coded as functions. The calling sequence for these is

A = UNIFM1 (B) or A = UNIFM2 (B, C, D)

where,

A is the computed random number.

B is an arbitrary starting number.

C is the mean.

D is the width of the interval.

UNIFM1 assumes a mean of .5 and an interval width of 1. The starting number B is used to initialize the calculations of the random number. Each time the routine is called, the contents of B are changed. The random number sequence can be restarted by reinitializing B.

The third routine is called by:

CALL URAN (B, N. RANDOM)

where,

B is the starting number as above.

N is the number of random values to be generated.

RANDOM is the name of a vector in which the random numbers will be stored. URAN assumes a mean of .5 and a range of 1.

METHOD

The method used is the mixed congruential

$$R_{n+1} = ((\alpha+1) R_n + B) \text{ MOD } 2^{**}35$$

with $\alpha = 2^9$ and $B = 262035034724_8$

These values have been shown experimentally[1] to yeild good results for long strings of random numbers (greater than 5000). However, local non-randomness may occur. This method has a period of 2**35.

NOTE: For other methods of evaluating UNIFM1 and UNIFM2, see the programs UNIFM and FLAT.

SAMPLE PROBLEM

Calculate 10 random numbers with uniform distribution in the interval (0, 1). Also calculate 10 random numbers uniformly distributed in (-1, 1).

SAMPLE SOLUTION

The 10 random numbers in (0, 1) are calculated by using UNIFM1 and URAN demonstrating the recall capability.

```
*LIST

010 DIMENSIØN R(10),S(10),T(10)
020 B=2.583
030 DØ 10 I=1,10
040 10 R(I)=UNIFM1(B)
050 DØ 20 I=1,10
60 20 S(I)=UNIFM2(B,0.,2.)
070 B=2.583
080 CALL URAN(B,10,T)
090 PRINT 30,(R(I),S(I),T(I),I=1,10)
100 30 FØRMAT(3F12.8)
110 STØP;END

READY

*RUN *;URAN
 0.00518162  0.03672557  0.00518162
 0.35370346  0.23128421  0.35370346
 0.14540854  0.03987048  0.14540854
 0.29011743 -0.15537452  0.29011743
 0.52577567 -0.31605968  0.52577567
 0.41845382 -0.74754387  0.41845382
 0.36234745 -0.09893599  0.36234745
 0.57977710  0.63690507  0.57977710
 0.12118834  0.12337043  0.12118834
 0.86515171  0.68010121  0.86515171

PRØGRAM STØP AT 110
*
```

[1] A. M. Olson, "A Statistical Examination of Some Pseudo-Random Number Generators Having a Uniform Frequency Function," General Electric Company, TIS No. R63ASD3, 23 Aug. 1963.

XINGAM

This FORTRAN function evaluates the cumulative probabilities and percentage points of the Gamma distribution.

METHOD

The probability density function for the random variable x having a Gamma distribution with parameters $\alpha > 0$ and $B > 0$* is

$$f(x) = \frac{1}{\Gamma(\alpha)\, B^{\alpha}} x^{\alpha-1} e^{-x/B} \quad , 0 < x < \infty$$

$$= 0 \quad \text{elsewhere.}$$

The cumulative probability, i. e., x is in the interval $[0, x_0]$ with $x_0 > 0$, is

$$\Pr(x \le x_0) = \int_0^{x_0} f(x)\, dx$$

INSTRUCTIONS

The calling sequence for this function is

Y = XINGAM (IND, ALPHA, BETA, XP)

where:

1. ALPHA and BETA are the parameters α and B as described under METHOD
2. if IND = 0 then the function calculates

   $Y = \Pr(x \le XP)$
3. if IND = 1 then the function calculates Y such that

   $XP = \Pr(x \le Y)$

RESTRICTIONS

1. $0 < \text{ALPHA} \le 20$
2. $0 < \text{BETA}$
3. if IND = 1 then $XP \ge .005$ and $0 < \text{ALPHA} \le 10$

* Wall, H. S., Analytic Theory of Continued Fractions

SAMPLE SOLUTION

```
*10 PRINT:"WHAT ARE IND,ALPHA,BETA,XP";READ:IND,ALPHA,BETA,XP
*20 Y=XINGAM(IND,ALPHA,BETA,XP)
*30 PRINT:"RESULTS",Y,XP
*40 STØP;END
*RUN *;XINGAM
WHAT ARE IND,ALPHA,BETA,XP
= 0 1 3 .75
RESULTS    2.2119923E-01    7.5000000E-01

PRØGRAM STØP AT 40
*RUN *;XINGAM
WHAT ARE IND,ALPHA,BETA,XP
= 1 1 3 .22119923
RESULTS    7.4999976E-01    2.2119923E-01

PRØGRAM STØP AT 40
*
```

XNOR1

This FORTRAN function generates a random number from a normal distribution for a specified mean and standard deviation.

INSTRUCTIONS

The calling sequence is:

Y = XNORM1 (X, A, B)

where,

Y is the value of the random number.

A is the mean.

B is the standard deviation.

X must be negative the first time the routine is called and is ignored when positive or zero. The routine uses the previous random number on subsequent references; therefore, X should be set positive or zero for additional random numbers. To generate any sequence of random numbers starting with A, set X = -A. (A must be greater than 0, but not greater than 1.)

RESTRICTIONS

The library subroutine program RANDX must be used with this program (see sample solution.)

SAMPLE PROBLEM

Find 10 random numbers from a normal distribution with a mean of zero and a standard deviation of one.

SAMPLE SOLUTION

```
*LIST

10      Y=XNØRM1(-1.,0.,1.)
20      PRINT 1,Y
30    1 FØRMAT(16H RANDØM NUMBER =,1PE20.7)
40      DØ 2 L=2,10
50      Y=XNØRM1(0.,0.,1.)
60    2 PRINT 1,Y
70      STØP
80      END

READY

*RUN *;XNØR1;RANDX
RANDØM NUMBER =        -8.6320875E-01
RANDØM NUMBER =         1.7534616E+00
RANDØM NUMBER =         9.1031063E-01
RANDØM NUMBER =         1.0807851E+00
RANDØM NUMBER =         1.1289800E-01
RANDØM NUMBER =        -1.0634889E+00
RANDØM NUMBER =         1.4758431E-01
RANDØM NUMBER =        -1.5049016E-01
RANDØM NUMBER =        -1.3519395E+00
RANDØM NUMBER =        -6.3745698E-01

PRØGRAM STØP AT 70
*
```

XNORM

This FORTRAN function generates a random number from a normal distribution with a mean of zero and a standard deviation of one.

INSTRUCTIONS

The calling sequence is:

Y = XNORM(X)

where,

Y is the value of the random number.

X must be negative the first time the routine is called and is ignored when positive or zero. The routine uses the previous random number on subsequent references; therefore, X should be set positive or zero for additional random numbers. To generate any sequence of random numbers starting with A, set X = -A. (A must be greater than 0, but not greater than 1.)

RESTRICTIONS

The library subprogram RANDX must be used with this subprogram. (See sample solution.)

SAMPLE PROBLEM

Find 10 random numbers from a normal distribution with a mean of zero and a standard deviation of one.

SAMPLE SOLUTION

```
*LIST
10      Y=XNØRM(-1.0)
20      PRINT 1,Y
30    1 FØRMAT(16H RANDØM NUMBER =,1PE20.7)
40      DØ 2 L=2,10
50      Y=XNØRM(0.0)
60    2 PRINT 1,Y
70      STØP
80      END

READY

*RUN *;XNØRM;RANDX
RANDØM NUMBER =        -8.6320875E-01
RANDØM NUMBER =         1.7534616E+00
RANDØM NUMBER =         9.1031063E-01
RANDØM NUMBER =         1.0807851E+00
RANDØM NUMBER =         1.1289800E-01
RANDØM NUMBER =        -1.0634889E+00
RANDØM NUMBER =         1.4758431E-01
RANDØM NUMBER =        -1.5049016E-01
RANDØM NUMBER =        -1.3519395E+00
RANDØM NUMBER =        -6.3745698E-01

PRØGRAM STØP AT 70
*
```

HONEYWELL INFORMATION SYSTEMS

Technical Publications Remarks Form*

TITLE: SERIES 6000/600 TIME-SHARING APPLICATIONS LIBRARY GUIDE VOLUME II – STATISTICS

ORDER NO: DA44, Rev. 0

DATED: JUNE, 1971

ERRORS IN PUBLICATION:

SUGGESTIONS FOR IMPROVEMENT TO PUBLICATION:

(Please Print)

FROM: NAME ____________________ DATE ____________

COMPANY ____________________

TITLE ____________________

ADDRESS ____________________

____

*Your comments will be promptly investigated by appropriate technical personnel, action will be taken as required, and you will receive a written reply. If you do not require a written reply, please check here. ☐

CUT ALONG LINE

CUT ALONG LINE

FOLD ALONG LINE

POSTAGE STAMP NOT NECESSARY IF MAILED IN THE UNITED STATES

BUSINESS REPLY MAIL: FIRST CLASS PERMIT NO. 39531, NEWTON HIGHLANDS, MASS.

POSTAGE WILL BE PAID BY:

HONEYWELL INFORMATION SYSTEMS

60 WALNUT STREET
WELLESLEY HILLS, MASS. 02181

ATTN: U.S. GROUP PUBLICATIONS, MS 050

FOLD ALONG LINE

Honeywell Bull

Honeywell Bull

HONEYWELL INFORMATION SYSTEMS