

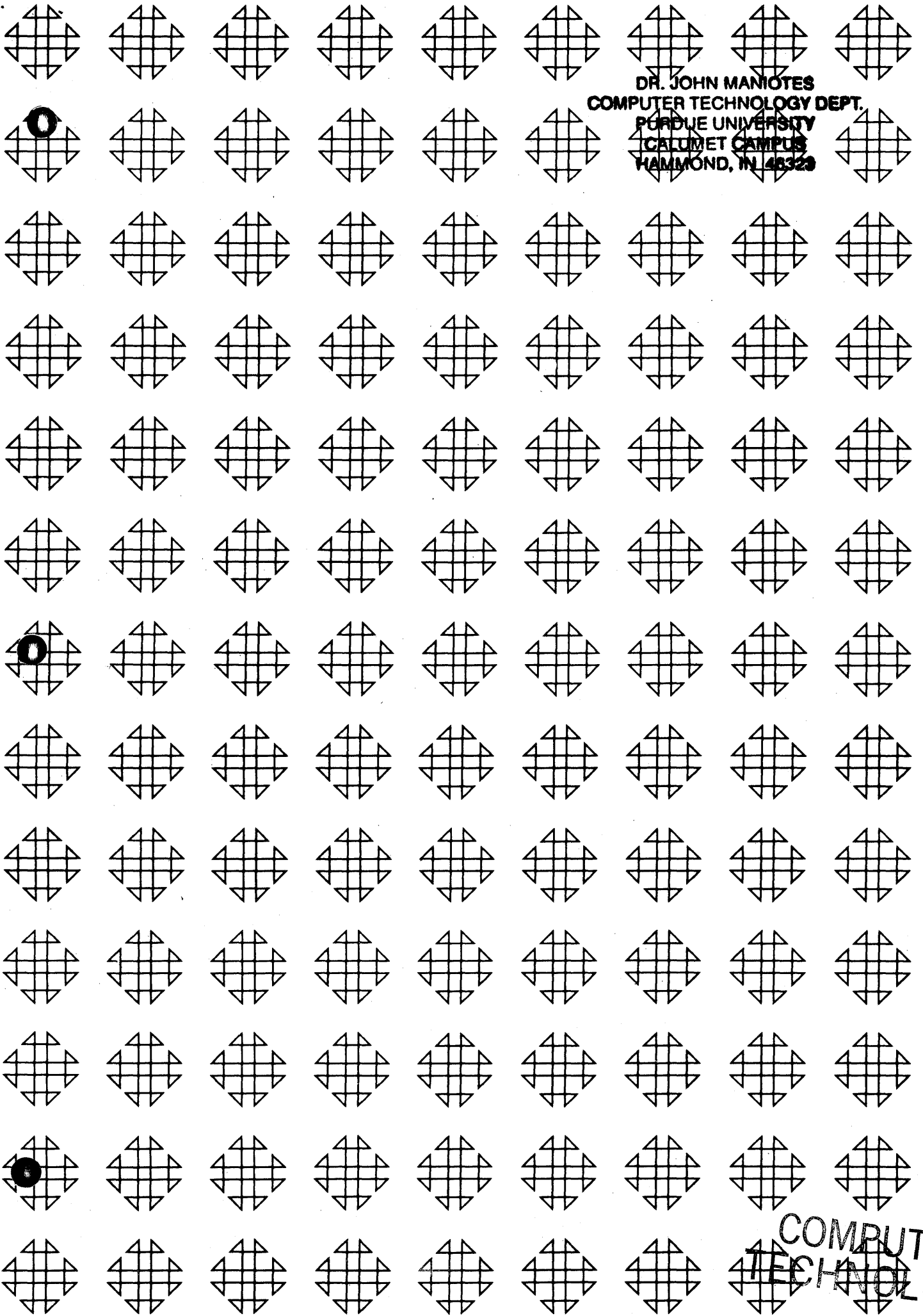
DR. JOHN MANIOTES
COMPUTER TECHNOLOGY DEPT.
PURDUE UNIVERSITY
CALUMET CAMPUS
HAMMOND, IN 46322

1620 GENERAL PROGRAM LIBRARY

STATISTICS I (CARD)

6.0.0891

COMPUTER
TECHNOLOGY



277011AM 11/10/81
1430 Y00100M 11/10/81
YTI0000000000000000
000000000000000000
000000000000000000

Disclaimer

Although this program was tested by its author prior to submission, no warranty, express or implied, is made by the author, 1620 USERS Group, or IBM as to the accuracy and functioning of the program and related program material and no responsibility is assumed by the author, 1620 USERS Group, or IBM in connection therewith.

1620 USERS GROUP PROGRAM REVIEW AND EVALUATION

(fill out in typewriter or pencil, do not use ink)

Program No. _____

Date _____

Program Name: _____

1. Does the abstract adequately describe what the program is and what it does? Yes ___ No ___
Comment _____
2. Does the program do what the abstract says? Yes ___ No ___
Comment _____
3. Is the Description clear, understandable, and adequate? Yes ___ No ___
Comment _____
4. Are the Operating Instructions understandable and in sufficient detail? Yes ___ No ___
Comment _____
Are the Sense Switch options adequately described (if applicable)? Yes ___ No ___
Are the mnemonic labels identified or sufficiently understandable? Yes ___ No ___
Comment _____
5. Does the source program compile satisfactorily (if applicable)? Yes ___ No ___
Comment _____
6. Does the object program run satisfactorily? Yes ___ No ___
Comment _____
7. Number of test cases run _____. Are any restrictions as to data, size, range, etc. covered adequately in description? Yes ___ No ___
Comment _____
8. Does the Program Meet the minimal standards of the 1620 Users Group? Yes ___ No ___
Comment _____
9. Were all necessary parts of the program received? Yes ___ No ___
Comment _____
10. Please list on the back any suggestions to improve the usefulness of the program. These will be passed onto the author for his consideration.

Please return to:

Mr. Richard L. Pratt
Data Corporation
7500 Old Xenia Pike
Dayton, Ohio 45432

Your Name _____

Company _____

Address _____

User Group Code _____

THIS REVIEW FORM IS PART OF THE 1620 USER GROUP ORGANIZATION'S PROGRAM REVIEW AND EVALUATION PROCEDURE. NONMEMBERS ARE CORDIALLY INVITED TO PARTICIPATE IN THIS EVALUATION.

STATISTICS I

ROBERT J. ROBINSON

MARQUETTE UNIVERSITY

MILWAUKEE, WISCONSIN

NOVEMBER 11, 1961

Modifications or revisions to this program, as they occur, will be announced in the appropriate Catalog of Programs for IBM Data Processing Systems. When such an announcement occurs, users should order a complete new program from the Program Information Department.

Program: STATISTICS I
Date: September 21, 1961
Programmer: Robert J. Robinson (Marquette University)
Description:

This program accepts up to 50 variables, and any number of observations, including missing data, and computes the standard deviation, standard error, coefficient of variation, number of observations, sum, mean, and sum of squared of deviation for each variable, and T-ratios and the degree of freedom between all variables.

The standard deviation is computed by one of the two formulas, depending on the number of observations (N_i) for the variable concerned:

If $N_i \leq 30$

$$\sigma_i = \sqrt{\frac{(x_i - \bar{x}_i)^2}{N_i - 1}}$$

If $N_i > 30$

$$\sigma_i = \sqrt{\frac{(x_i - \bar{x}_i)^2}{N_i}}$$

The standard error is computed from:

$$\sigma_{\bar{x}_i} = \frac{\sigma_i}{\sqrt{N_i}}$$

STATISTICS I

The computation of the T-ratios is optioned by setting a sense switch. If the computation is optioned, the T-ratio and the degrees of freedom are computed by one of the following formula sets, depending on the relation between the number of observations (N_i and N_j) of the two variables being compared:

If $N_i = N_j$,

$$T_{ij} = (X_i - X_j) \sqrt{\frac{N_j(N_j - 1)}{(\sum(\bar{x}_i - x_i)^2) + (\sum(\bar{x}_j - x_j)^2)}}$$

$$D = 2(N_j - 1)$$

If $N_i \neq N_j$,

$$T_{ij} = (X_i - X_j) \sqrt{\frac{(N_i N_j) (N_i + N_j - 2)}{(N_i + N_j) \cdot (\sum(\bar{x}_i - y_i)^2 + (\sum(\bar{x}_j - x_j)^2))}}$$

$$D = N_i + N_j - 2$$

Fortran

Coding
Language:

Input
Format:

Input is via cards. Data consists of two fixed point constants, N and J, which specify the number of rows in the input matrix (ie, the number of observations for the variables with the greatest number of rows), and the number of variables. These values are then followed by the elements of the matrix, punched row-wise in floating point form. Missing data is indicated by punching 9.E40 in the proper sequence on the card.

STATISTICS I

Thus, if the data to be used were:

	X_1	Y_2	X_3	X_4
1	4.3	6.8	-10.9	9×10^3
2	-6.85	0.0	—	45.68
3	2.96	48.2	63.	-10.1
4	—	—	—	4.95
5	—	5.9	—	—

The data cards could be punched as follows:

Card 1:	$\frac{N}{5}$	$\frac{j}{4}$	$\frac{X11}{4.3}$	$\frac{Y12}{6.8}$	$\frac{Y13}{10.9}$	$\frac{X14}{9.E3}$	$\frac{Y21}{-6.85}$
Card 2:	$\frac{X22}{0.0}$	$\frac{Y23}{9.E40}$	$\frac{X24}{15.38}$	$\frac{X31}{2.96}$	$\frac{Y32}{48.2}$	$\frac{X33}{63.}$	
Card 3:	$\frac{X34}{-10.1}$	$\frac{Y41}{9.E40}$	$\frac{Y42}{9.E40}$	$\frac{X43}{9.E40}$	$\frac{Y44}{4.95}$		
Card 4:	$\frac{X51}{9.E40}$	$\frac{Y52}{5.9}$	$\frac{X53}{9.E40}$	$\frac{X54}{9.E40}$	$\frac{X55}{9.E40}$		

any number of elements may be punched on any card, but only the first 72 columns are to be used, and an element may not be split into two parts. (ie, cannot punch the elements X_{44} as 4.9 on the one card and the 5 on the next card).

Output
Format:

Output consists of two lines of print (or two cards) for each variable (ie, 2j lines of print). The first line contains the following items:

Variable number, I
(ie, 3 if output is for variable X_3)

$$\sum_{j=1}^{N_I} (\bar{x}_I)_i$$

(ie, the sum of the variable X_I)

$$\bar{y}_I$$

(mean of Variable X_I)

$$\sum_{i=1}^{N_I} (\bar{x}_I - (X_I)_i)^2$$

(the sum of squared deviations)

```

DIMENSIONX(50),SUM(50),SMEAN(50),SUMDV(50),FNS(50),SUMSQ(50)
90 READ,N,J
  N2 = N
  DO11 = 1,J
  SUMDV(1)=0.0
  FNS(1)=0.0
  SUMSQ(1)=0.0
  1SUM(1)=0.0
  3DO 12 I=1,J
  READ,X(1)
  IF(X(1)-9.E40)11,12,11
  11FNS(1)=FNS(1)+1.0
  SUM(1)=SUM(1)+X(1)
  SUMSQ(1)=SUMSQ(1)+(X(1)**2)
  12CONTINUE
  N=N-1
  IF(N) 99,5,3
  5DO 6 I=1,J
  SUMDV(1)=(SUMSQ(1)-(SUM(1)**2/FNS(1)))
  6 SMEAN(1)=SUM(1)/FNS(1)
  8DO 10 I=1,J
  IF(FNS(1)-30.)17,17,19
  17 STDEV=SQR(SUMDV(1))/(FNS(1)-1.0)
  GO TO 18
  19 STDEV=SQR(SUMDV(1)/FNS(1))
  18 STERR=STDEV/SQR(FNS(1))
  COEFV=STDEV/SMEAN(1)
  IF (SENSE SWITCH 2) 13,14
  13 PRINT,I,SUM(1),SMEAN(1),SUMDV(1)
  PRINT,STDEV,STERR,COEFV,FNS(1)
  GO TO 10
  14 PUNCH,I,SUM(1),SMEAN(1),SUMDV(1)
  PUNCH,STDEV,STERR,COEFV,FNS(1)
  10 CONTINUE
  IF (SENSE SWITCH 1) 20,30
  20 IF (SENSE SWITCH 2) 91,92
  91 PRINT
  GO TO 93
  92 PUNCH
  93 J2=J-1
  DO 26 L=1,J2
  SMENA=SMEAN(L)
  SUMDA=SUMDV(L)
  FNSA=FNS(L)
  21 K=L+1
  DO 25 I=K,J
  DIFF= SMENA-SMEAN(I)
  IF(FNSA-FNS(I)) 23,22,23
  22 TRATO =DIFF*(SQR((FNS(I)*(FNS(I)-1.0)))/(SUMDA+SUMDV(I))))
  DEGFR=2.0*(FNS(I)-1.0)
  GO TO 24
  23 PART1=(FNSA*FNS(I))*(FNSA+FNS(I)-2.0)
  PART2=(FNSA+FNS(I))*(SUMDA+SUMDV(I))
  TRATO =DIFF*(SQR(PART1/PART2))

```

```

DEGFR=(FNSA +FNS(1)-2.0)
24 IF (SENSE SWITCH 2) 27,28
27 PRINT,L,I
  PRINT,TRATO,FNSA,FNS(1), DEGFR
  GO TO 25
28 PUNCH,L,I
  PUNCH,TRATO,FNSA,FNS(1),DEGFR
25 CONTINUE
26 CONTINUE
  30 PAUSE
  GO TO 90
  END
END OF LISTING

```

STATISTICS I

The second line contains:

\sqrt{I}	$\sqrt{X_I}$	C	N_I
(Standard Deviation of variable X_I)	(Standard Error of the mean, for variable X_I)	(Coefficient of variation, for variable X_I)	(number of significant items for variable X_I)

If the T-test is optioned (sense switch 1 on) the following answers will be output when the above output is completed.

(j-1) + (j-2) + ... + 1 sets of two lines each will be printed (or punched). The first line of each pair is:

I	J
(designates variable X_I)	(designates variable X_J)

The second line is:

T_{IJ}	N_I	N_J	D
(T-ratio between X_I and X_J)	(number of significant items for variable X_I)	(same, for variable X_J)	(degrees of freedom between variables X_I and X_J)

Restrictions:

Maximum number of variables is 50. If the data is drawn from the same related group the formulas in this program are not valid. If the mean of a variable is zero, the coefficient of variation has no meaning as calculated. An "error E 7" will be printed prior to printing the answers for the variable whose mean is zero. All other calculations (including those for that variable) are valid.

Accuracy
and Speed:

Speed not yet determined. Accuracy to 7 places.

Operating
Instructions:

I. Switch Settings

Sense Switch 1 ON Compute T-ratios.

Sense Switch 1 OFF Do not compute T-ratios.

Sense Switch 2 ON Output via typewriter.

Sense Switch 2 OFF Output via card punch.

STATISTICS I

O/flow switch	set to program
Parity switch	set to stop
I/O switch	set to stop

- II. Set margins at 11 and 96. Set tabs at 28, 45, 62, 79.
- III. Put 1620 in manual mode (push reset).
- IV. Ready read hopper with the program followed by data cards.
- V. Press "Load" key.

Program will load, and typewriter will type "load data".

Merely press "start" key on 1620 console.
- VI. When "reader no feed" light comes on, press "reader start" key on card reader.

Computation will than begin, and typing will proceed. After the "stop" types out, if more data sets are to be run, the following procedure can be used:

Press "reset" and "insert". Type 49 07500, and press "release" and "start". Then place data deck in read hopper and press "reader start". Proceed with the above instructions, starting with item VI.

Equipment
Required:

IBM 1620 card system with 40K memory. (May be used on a 20K machine, but it will be necessary to change the array sizes in the demension statement and re-compile).

Hash
Total:

2390⁸82701699275232⁸
(using card hash total program developed at Marquette University).