


1/47

SOFTWARE RELEASE NOTICE

01. SRN Number: PA-SRN-153		
02. Project Title: Summary of Information Relevant to Specification of the Critical Group and Reference Biosphere		Project No. 5707-771-720
03. SRN Title:     STEPWISE		
04. Originator/Requestor: Patrick LaPlante		Date: 7/30/97
05. Summary of Actions		
<input checked="" type="checkbox"/> Release of new software <input type="checkbox"/> Release of modified software: <input type="checkbox"/> Enhancements made <input type="checkbox"/> Corrections made <input type="checkbox"/> Change of access software <input checked="" type="checkbox"/> Software Retirement		
<i>PLW 12/4/2001</i>		
06. Persons Authorized Access		
Name	RO/RW	A/C/D
Patrick LaPlante	RO	A
James Weldy	RO	A
Bob Baca	RO	A
07. Element Manager Approval: Bob Baca <i>RF/Baca</i>		Date: <i>7/30/97</i>
08. Remarks: STEPWISE performs linear multiple regression for raw and ranked data using stepwise selection procedure. The output includes standard ANOVA tables, PRESS plot, residual plots, and scatter plots for parameters selected in the final model. Version obtained is early version that was provided by R. Cady at NRC. It produces the same results as 1980 user manual. Originally developed at Kansas State University and assembled by Sandia National Lab, the code is currently under the control of Martin Marietta under DOE contract. The most current version that they have has undergone a QA review and documentation but was not obtained by CNWRA at this time due to the amount of time and number of formal contacts that would be necessary to obtain it.		

2/47

SOFTWARE SUMMARY FORM

01. Summary Date: 7/10/97	02. Summary prepared by (Name and phone) Patrick A. LaPlante 301-881-0289	03. Summary Action:  New	
04. Software Date: 5/19/93 (Source code file date)	05. Short Title:  STEPWISE		
06. Software Title:      Stepwise Regression with PRESS and Rank Regression		07. Internal Software ID:	
08. Software Type:  <input type="checkbox"/> Automated Data System  <input checked="" type="checkbox"/> Computer Program  <input type="checkbox"/> Subroutine/Module	09. Processing Mode:  <input type="checkbox"/> Interactive  <input checked="" type="checkbox"/> Batch  <input type="checkbox"/> Combination	10. APPLICATION AREA a. General: <input checked="" type="checkbox"/> Scientific/Engineering <input checked="" type="checkbox"/> Auxiliary Analyses <input checked="" type="checkbox"/> Total System PA <input checked="" type="checkbox"/> Subsystem PA <input type="checkbox"/> Other  b. Specific: Statistical	
11. Submitting Organization and Address:  CNWRA/SwRI 6220 Culebra Rd San Antonio, TX 78238		12. Technical Contact(s) and Phone:  Patrick A. LaPlante (301) 881-0289	
13. Narrative: STEPWISE performs linear multiple regression for raw and ranked data using stepwise selection procedure. The output includes standard ANOVA tables, PRESS plot, residual plots, and scatter plots for parameters selected in final model. Version obtained is early version that came from R. Cady at NRC. This version produces the same results as shown in 1980 manual. Originally developed at Kansas State University and assembled by Sandia National Lab, the code is currently under the control of Martin Marietta under DOE contract. The most current version that they have has undergone a QA review and documentation but was not obtained by CNWRA at this time due to the amount of time and number of formal contacts that would be necessary to obtain it.			
14. Computer Platform: IBM	15. Computer Operating System: DOS	16. Programming Language(s): FORTRAN	17. Number of Source Program Statements: 3148 lines including comments.
18. Computer Memory Requirements: Unknown (source code designed for VAX and compiled for 386 PC). Memory requirements considered minimal by todays standards.	19. Tape Drives: None	20. Disk/Drum Units: None	21. Graphics: None (output file read by ASCII editor)
22. Other Operational Requirements: 386 processor w/ math coprocessor or higher. Might not run on Pentium unless recompiled. Will run in OS/2 window as well.			
23. Software Availability: <input type="checkbox"/> Available <input checked="" type="checkbox"/> Limited <input type="checkbox"/> In-House ONLY		24. Documentation Availability: <input checked="" type="checkbox"/> Available <input type="checkbox"/> Inadequate <input type="checkbox"/> In-House ONLY	
Software Custodian:  Date: 7/30/97			

3/47

# CENTER FOR NUCLEAR WASTE REGULATORY ANALYSES

## SOFTWARE CONTROL CHECKLIST

Name of Software: STEPWISE

Version: N/A

Primary User: Pat LaPlante

<input type="checkbox"/>	SOFTWARE REQUIREMENTS DESCRIPTION Documentation	<input type="checkbox"/>
<input type="checkbox"/>	DESIGN AND DEVELOPMENT Documentation (Scientific Notebook)	<input type="checkbox"/>
<input type="checkbox"/>	DESIGN VERIFICATION Computer runs uniquely identified Software analysis tools have been applied and discrepancies resolved Design Verification Report	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<input checked="" type="checkbox"/>	INSTALLATION TESTING Installation test documentation-Scientific Notebook No. 178 Discrepancy resolution	<input checked="" type="checkbox"/> <input type="checkbox"/>
<input checked="" type="checkbox"/>	CONFIGURATION CONTROL Software Summary Form User's Manual - portions enclosed - resides w/P. LaPlante Technical Description Source Code - in this folder Version Control Software Release Notice	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input checked="" type="checkbox"/>
<input type="checkbox"/>	SOFTWARE PROBLEM REPORTING AND RESOLUTION Software Problem and Change Request	<input type="checkbox"/>
<input type="checkbox"/>	SOFTWARE VALIDATION Software Validation Test Plan Software Validation Test Report Software Validation Review	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/>	SOFTWARE RETIREMENT Software Release Notice	<input type="checkbox"/>

4/47

TO Mail List:#DIRS-MGRS  
To: Larry McKague at CNWRA  
To: English Percy at CNWRA  
To: Mark Jarzempa at CNWRA  
To: James Weldy at CNWRA-SUN  
To: Patrick LaPlante  
CC: Linda Hearon  
CC: RBrient@swri.edu at Internet  
BCC: Bruce Mabrito  
From: Bruce Mabrito  
Subject: Control of Software MCNP and STEPWISE  
07-30-97 08:14 PM

The following two scientific and engineering codes have been put under control in the CNWRA QA Records Room: MCNP Version 4A and STEPWISE (No Version Number Available).

Both of these codes were acquired by the CNWRA; installation testing has been performed and adequately documented on both; there is a technical description of each code in the respective software folders and references to the users' manuals; and both Software Release Notices have been signed by the appropriate Element Manager (MCNP, signed by N. Sridhar; STEPWISE, signed by R. Baca).

MCNP refers to "A General Monte Carlo N-Particle Transport Code" (primary CNWRA users determined to be Mark Jarzempa and James Weldy) and STEPWISE refers to "linear multiple regression for raw and ranked data using stepwise selection procedures" (primary CNWRA users determined to be Pat LaPante, James Weldy, and Bob Baca). A copy of this message will be inserted into the respective software QA folder and an electronic copy of each code is in that folder. Bruce Mabrito

5/27/97 Installation Testing for "STEPWISE" Software

*My*

NO version number

Participants: P. LaPlante

O. OACA-20-5708-771-720

Author/Primary Participant: P. LaPlante

S/Notebook No. 178

6/19/97

Overall Objective: Verify proper installation and operation of the STEPWISE ("Stepwise Regression with PRESS and Rank Regression") software following the procedure outlined in TOP-018 Section 5.6. This procedure involves running tests <sup>problems</sup> and comparing results with those published in the user manual.

Work Plan: The STEPWISE software was originally developed at Kansas State University and was subsequently provided to Sandia National Laboratories who added additional features to the code as described in the user guide (Fman et. al, 1980). The STEPWISE program is designed to perform single and multiple regression statistics on data with up to 129 independent variables. The code performs regression on raw or rank transformed data and allows <sup>either</sup> a stepwise or backward selection technique to be used. The code outputs summary statistics as well as tables of correlation coefficients (correlation matrices) and plots of residuals and PRESS (predicted regression error sum of squares) plots. The user manual provides 2 examples ~~problems~~ which involve stepwise, backward selection ~~and~~ on raw and ranked data. These examples will be ~~run~~ copied directly from the user manual.

6/47

5/27/97

4/6

(i.e., the STEPWISE input decks) and run with the installed version of STEPWISE. Output of results will be compared with the output printouts provided in the user manual to check for deviations. Any discrepancies will be documented. Primary statistics to check on the output will include: summary statistics, results for each model, including the sum of squares (SS), the mean square (MS), F statistic, F test significance,  $R^2$ , variables selected for model, regression coefficients, standardized regression coefficients, partial sum of squares, and F test values, PRESS value and the summary plot of PRESS values for each model. A visual inspection of plots will also be done to check for differences.

If no differences are detected in the output, these results will indicate the code has been installed correctly and is functioning as intended.

6/9/97

4/6

Results:

The input deck was prepared in an ascii file to match the input deck provided for the examples in the STEPWISE manual. The manual's input deck is shown on page 30, and the deck that was used for the testing is shown on page 31. Subsequent pages show the code output from the manual on the even numbered pages and the testing output on the odd numbered pages - for easy comparison. On the testing output, check

6/6/97

Setup information for STEPWISE INPUT DECK (Cards not used so few  
as purpose cards = lines)

Hj

## VI. DECK SETUP FOR STEPWISE

On the following page is an example of how to setup the cards for use in running STEPWISE. Cards 1 - 15, 22 - 32, and 46 - 64 all begin in card column one, while the Fortran statements in cards 16 - 21 start in column 7. Cards 35 - 45 are data cards in 5F6.0 format. Card number 1 is the JOB card and needs to be changed only to show the user's name and the box number. Note also that due to large core memory requirements of STEPWISE, that the extended core parameter must be set at 1200. Card number 2 must contain the user's social security number, division number and charge number. Cards 3 - 14 are control cards and will remain the same for all runs. The only exception being the use of TAPE 19 and TAPE 20 control cards that are explained in Section V. Cards 16 - 21 provide a dummy subroutine for transformations with any desired transformations (see page 17) following card number 20. Cards 23 - 32 are parameter cards (see pages 2-17). If the data is on cards, it will appear starting on card 33 with one vector of observations per card(s), followed by the END OF DATA card. Cards 47 - 53 reprocess the data with a backward solution. Note the absence of the LABEL card. The labels used in the first pass of the data will be used for this second analysis. Cards 54 - 64 reprocess the data with a stepwise solution and RANK regression.

Card No.

## DECK SETUP FOR STEPWISE

1	STE,T10,ECL200.
2	ACCOUNT,S509423684,D1223,G13,ACL,3000,RP,KUNC.
3	FTN,R=2,B=TRANZ.
4	REWIND,TRANZ.
5	ATTACH,OBJECT,STEPWISE-RLI.

Files for Installation files

8/47

6/9/97

## Results (cont)

works indicates these results that were compared with the manual output. The manual copy is difficult to read, however all compared values could be deciphered from the available copy.

## Discussion of results:

All information that was checked from the testing output file compared exactly with that presented in the manual except in 2 instances where deviations were considered to be of no significance to current anticipated use of the code. These 2 instances include the addition of confidence interval output (noted on page 55) which was not found in the stepwise manual. This indicates that this feature was added after the manual was published. The determination of confidence intervals for regression estimates is a simple calculation and if such output is used, it will be possible to confirm correctness w/ hand calculations. The 2nd instance where a deviation was noted is a difference in the 'alpha hat' numbers determined for rank regression as shown on page 61. The deviation from the manual example is by .0005 and since the alpha level used to determine <sup>6/9/97</sup> in significance (t-tests) tests is .05 (in) and .1 (out) this difference is thought to be insignificant and should not affect results. Furthermore, the current intended use will not involve rank regression - so this part of the code is not being exercised.



Input deck for examples in STEPWISE Manual

9/47

6/6/97  
M

6/6/97  
M

23 TITLE, EXAMPLE OF STEPWISE OPTION (DATA FROM DRAPER AND SMITH,  
 24 p. 365 - 402)  
 25 DATA,5,0,1.  
 26 LABEL(1)=X1,X2,X3,X4,Y  
 27 MODEL,5=1+2+3+4.  
 28 STEPWISE,SIGIN=.05,SIGOUT=.10  
 29 PRESS  
 30 OUTPUT,ALL  
 31 PLT RESIDUALS  
 32 END OF PARAMETERS  
 (5F6.0)

33	71.	26.	6.	60.	78.5
34	1.	29.	15.	52.	74.3
35	11.	56.	8.	20.	104.3
36	11.	31.	8.	47.	87.6
37	7.	52.	6.	33.	95.9
38	11.	55.	9.	22.	109.2
39	3.	71.	17.	6.	102.7
40	1.	31.	22.	44.	72.5
41	2.	54.	18.	22.	93.1
42	21.	47.	4.	26.	115.9
43	1.	40.	23.	34.	83.8
44	11.	66.	9.	12.	113.3
45	10.	68.	8.	12.	109.4

A(1) B(1)

46 END OF DATA  
 47 TITLE,EXAMPLE OF BACKWARD OPTION (DATA FROM DRAPER & SMITH,  
 48 p. 364 - 402)  
 49 DATA,5,0,2.  
 50 MODEL,5=1+2+3+4.  
 51 BACKWARD,SIG=.05  
 52 OUTPUT,CORR,STEPS,RESIDUALS  
 53 PRESS  
 54 END OF PARAMETERS  
 55 TITLE,EXAMPLE OF RANK REGRESSION WITH STEPWISE OPTION (DATA FROM  
 56 DRAPER/SMITH)  
 DATA,5,0,2.  
 LABEL(1)=RANK(X1),RANK(X2),RANK(X3),RANK(X4),RANK(Y)

57 MODEL,5=1+2+3+4.  
 58 STEPWISE,SIGIN=.05,SIGOUT=.10  
 59 PRESS  
 60 RANK REGRESSION  
 61 OUTPUT,CORR,STEPS,RESIDUALS  
 62 END OF PARAMETERS  
 63 (END OF INFORMATION -- MULTI-PUNCH 6 7 8 9 IN COL 1)  
 64 (END OF INFORMATION -- MULTI-PUNCH 6 7 8 9 IN COL 1)

Files for Installation test

10/47

Input deck for examples (created for testing)

06/97  
up

```

E.EXE - exampl.in
File Edit Options Help
TITLE, EXAMPLE OF STEPWISE OPTION (DATA FROM DRAPER AND SMITH, PG 365-482)
DATA, 5, 0, 1.
LABEL (1)=X1, X2, X3, X4, Y
MODEL, 5=1+2+3+4
STEPWISE, SIGIN=.05, SIGOUT=.10
PRESS
OUTPUT, ALL
PLOT RESIDUALS
END OF PARAMETERS
(5F6.0)
  7. 26.  6.  60.  78.5
  1. 29. 15.  52.  74.3
 11. 56.  8.  28. 104.3
 11. 31.  8.  47.  87.6
  7. 52.  6.  33.  95.9
 11. 55.  9.  22. 109.2
  3. 71. 17.  6. 102.7
  1. 31. 22. 44.  72.5
  2. 54. 18. 22.  93.1
 21. 47.  4.  26. 115.9
  1. 40. 23. 34.  83.8
 11. 66.  9.  12. 113.3
 10. 68.  8.  12. 109.4
END OF DATA
TITLE, EXAMPLE OF BACKWARD OPTION (DATA FROM DRAPER AND SMITH, P 364-482)
DATA, 5, 0, 2.
MODEL, 5=1+2+3+4.
BACKWARD, SIG=0.85
OUTPUT, CORR, STEPS, RESIDUALS
PRESS
END OF PARAMETERS
TITLE, EXAMPLE OF RANK REGERSSION WITH STEPWISE OPTION (DATA FROM DRAPER/SMITH)
DATA, 5, 0, 2.
LABEL (1)=RANK(X1), RANK(X2), RANK(X3), RANK(X4), RANK(Y)
MODEL, 5=1+2+3+4.
STEPWISE, SIGIN=.05, SIGOUT=.10
PRESS
RANK REGRESSION
OUTPUT, CORR, STEPS, RESIDUALS
END OF PARAMETERS

```

FROM

Files for Installation test

32

0/6/97 by

Stephanie Murrell

1/1/97

VARIABLE NAME	MEAN	STDEV	VARIANCE	STDEV	S.E. = $\frac{ST. DEV.}{\sqrt{n}}$	C.V. = $\frac{ST. DEV.}{\bar{x}} \times 100\%$
X1	7.64254	1.64254	2.69794	1.64254	0.41063	21.46%
X2	24.1783	2.41783	5.84583	2.41783	0.60446	25.03%
X3	11.7437	1.17437	1.37914	1.17437	0.29359	25.03%
X4	10.177	1.0177	1.0357	1.0177	0.25442	25.03%
X5	10.6438	1.06438	1.1328	1.06438	0.2661	25.03%

STATISTICS

TITLE: EXAMPLE OF STEPWISE OPTION DATA FROM JABER AND SMITH, P. 104-105  
 JABER LABORATORIES AND STEPWISE REGRESSION FROM KANSAS STATE UNIVERSITY

CORRELATION MATRIX

X1	1	0.957102	0.906408	0.823019	0.734203
X2	0.957102	1	0.923019	0.871021	0.781021
X3	0.906408	0.923019	1	0.871021	0.781021
X4	0.823019	0.871021	0.871021	1	0.781021
X5	0.734203	0.781021	0.781021	0.781021	1

NAME	X1	X2	X3	X4	X5
------	----	----	----	----	----

TITLE: EXAMPLE OF STEPWISE OPTION DATA FROM JABER AND SMITH, P. 104-105  
 JABER LABORATORIES AND STEPWISE REGRESSION FROM KANSAS STATE UNIVERSITY

CORRELATION MATRIX

X1	1	0.9571	0.9064	0.8230	0.7342
X2	0.9571	1	0.9230	0.8710	0.7810
X3	0.9064	0.9230	1	0.8710	0.7810
X4	0.8230	0.8710	0.8710	1	0.7810
X5	0.7342	0.7810	0.7810	0.7810	1

NAME	X1	X2	X3	X4	X5
------	----	----	----	----	----

PARTIAL CORRELATIONS AFTER X4 IS ADDED TO MODEL

$$r_{15.4} = \frac{.7307 - (-.2454)(-.8213)}{\sqrt{(1 - (.2454)^2)(1 - (.8213)^2)}} = .9568$$

= HIGHEST PARTIAL CORR. → X1 WILL BE THE SECOND VARIABLE ADDED AFTER X4

FOR ALL THE INPUT DATA

HIGHEST CORR. → X4 IS FIRST INDEPENDENT VARIABLE ADDED TO THE MODEL

$$r_{25.4} = \frac{-.8163 - (-.9730)(-.8213)}{\sqrt{(1 - (.9730)^2)(1 - (.8213)^2)}} = .1304$$

$$r_{35.4} = \frac{-.5397 - (.0295)(-.8213)}{\sqrt{(1 - (.0295)^2)(1 - (.8213)^2)}} = -.8952$$

$$r_{12} = \frac{.7307 - (-.2454)(-.8213)}{\sqrt{(1 - (.2454)^2)(1 - (.8213)^2)}} = .9568$$

= HIGHEST PARTIAL CORR.  $\Rightarrow$   $Y_1$  WILL BE THE SECOND VARIABLE ADDED AFTER  $Y_4$

1/6/97  
My

Testing Disturb

E:EXE - exampl.out  
File Edit Options Help

SANDIA LABORATORIES <>> STEPWISE REGRESSION <>> FROM KANSAS STATE UNIVERSITY

VARIABLE NAME	VARIABLE NUMBER	MEAN	VARIANCE	STD. DEV.	STD. ERR.	C.V.
X1	1	7.46154	34.6026 ✓	5.88239	1.63148	78.84
X2	2	48.1538 ✓	242.141 ✓	15.5609	4.31581	32.31
X3	3	11.7692 ✓	41.0256 ✓	6.40513 ✓	1.77646	54.42
X4	4	38.8088	288.167 ✓	16.7382	4.64234	55.79
Y	5	95.4231	226.314 ✓	15.0437	4.17236	15.77

13 OBSERVATIONS  
♀ TITLE, EXAMPLE OF STEPWISE OPTION (DATA FROM DRAPER AND SMITH, PG 365-402) PAGE 3

SANDIA LABORATORIES <>> STEPWISE REGRESSION <>> FROM KANSAS STATE UNIVERSITY

SUM OF SQUARES MATRIX

X1	1	4.152E+02 ✓				
X2	2	2.511E+02 ✓	2.986E+03			
X3	3	-3.726E+02 ✓	-1.665E+02	4.923E+02		
X4	4	-2.908E+02 ✓	-3.841E+03	3.800E+01	3.362E+03 ✓	
Y	5	7.760E+02 ✓	2.293E+03	-6.182E+02	-2.482E+03	2.716E+03 ✓

NO. 1 2 3 4 5  
NAME X1 X2 X3 X4 Y  
♀ TITLE, EXAMPLE OF STEPWISE OPTION (DATA FROM DRAPER AND SMITH, PG 365-402) PAGE 4

SANDIA LABORATORIES <>> STEPWISE REGRESSION <>> FROM KANSAS STATE UNIVERSITY

CORRELATION MATRIX

X1	1	1.0000				
X2	2	0.2286	1.0000			
X3	3	-0.8241	-0.1392 ✓	1.0000		
X4	4	-0.2454 ✓	-0.9738	0.8295	1.0000 ✓	
Y	5	0.7387	0.8163	-0.5347 ✓	-0.8213 ✓	1.0000

NO. 1 2 3 4 5  
NAME X1 X2 X3 X4 Y  
♀ TITLE, EXAMPLE OF STEPWISE OPTION (DATA FROM DRAPER AND SMITH, PG 365-402) PAGE 5

SANDIA LABORATORIES <>> STEPWISE REGRESSION <>> FROM KANSAS STATE UNIVERSITY

INVERSE OF CORR MATRIX

VA 1 0.000E+00

12/47  
33

6/6/97

34

8 for use Manual example of part

TITLE: EXAMPLE OF STEPWISE REGRESSION DATA FROM DEAR AND SMITH, P. 350-351  
 ANALYSIS: REGRESSION FOR VARIABLE 5 → THE DEPENDENT VARIABLE IS INPUT VARIABLE 5

NO. 1.000000  
 NO. 2.000000  
 NO. 3.000000

ANOVA TABLE

SOURCE	D.F.	SS	MS	F	P
REGRESSION	1	1871.4962	1871.4962	1871.4962	0.000000
RESIDUAL	14	1407.4638	100.5331		
TOTAL	15	3278.9600			

MEAN IS 1.000000  
 INTERCEPT IS 17.56793  
 STANDARD ERROR OF INTERCEPT IS 0.742201

VARIABLE NUMBER	VARIABLE NAME	REGRESSION COEFFICIENT	STANDARD ERROR OF REGRESSION COEFFICIENT	PARTIAL SS	F	P
4	X4	-0.73018101	0.521165	1871.4962	1871.4962	0.000000

UNIT FREE BETA, I.E.  $\beta_4^u = \hat{\beta}_4 \sqrt{\frac{\sum (X_{4i} - \bar{X}_4)^2}{\sum (Y_i - \bar{Y})^2}}$

REGRESSION EQUATION:  $\hat{Y} = 17.56793 - 0.730181 X_4$

THE CONSTANT TERM

... THE  $\beta^u$ 'S CAN BE USED FOR ORDERING THE RELATIVE IMPORTANCE WHEREAS THE USUAL  $\hat{\beta}$ 'S CANNOT BE USED FOR ORDERING

USED FOR TESTING  $H_0: \beta_4 = 0$

13/1/97

REGRESSION EQUATION:  $Y = 117.36793 - .73816181 X_4$

THE CONSTANT TERM

6/6/97  
My

```

E.EXE - exampl.out
File Edit Options Help
TITLE, EXAMPLE OF STEPWISE OPTION (DATA FROM DRAPER AND SMITH, PG 365-402) PAGE 6
SANDIA LABORATORIES <>< STEPWISE REGRESSION <>< FROM KANSAS STATE UNIVERSITY
ANOVA TABLE
ANALYSIS OF REGRESSION FOR VARIABLE 5---Y
(TABLE 1)

```

SOURCE	D.F.	SS	MS	F	SIGNIFICANCE
REGRESSION	1 ✓	1831.8962 ✓	1831.8962 ✓	22.798528 ✓	0.0006
RESIDUAL	11 ✓	883.86692 ✓	80.351538 ✓		
TOTAL	12 ✓	2715.7631 ✓			

```

R**2 IS 0.67454 ✓
INTERCEPT IS 117.56793 ✓
STANDARD ERROR OF INTERCEPT IS 5.26221 ✓

```

VARIABLE NUMBER	VARIABLE NAME	REGRESSION COEFFICIENTS	STANDARDIZED REGRESSION COEFFICIENTS	PARTIAL SSQ	T-TEST VALUES	R**2 DELETES	ALPHA HATS
4	X4	-.73816181 ✓	-0.821305 ✓	1831.8962 ✓	-4.7748 ✓	0.0006	0.0006

```

UNIQUE SEQUENCE NUMBER FOR THIS ANOVA = 181
PRESS IS 1194.2 ✓
TITLE, EXAMPLE OF STEPWISE OPTION (DATA FROM DRAPER AND SMITH, PG 365-402) PAGE 7
SANDIA LABORATORIES <>< STEPWISE REGRESSION <>< FROM KANSAS STATE UNIVERSITY
INVERSE OF CORR MATRIX

```

	NO.	1	4
X1	1	1.064E+08	
X4	4	2.612E-01	1.064E+08

```

NAME X1 X4
TITLE, EXAMPLE OF STEPWISE OPTION (DATA FROM DRAPER AND SMITH, PG 365-402) PAGE 8
SANDIA LABORATORIES <>< STEPWISE REGRESSION <>< FROM KANSAS STATE UNIVERSITY
ANOVA TABLE
ANALYSIS OF REGRESSION FOR VARIABLE 5---Y
(TABLE 1)

```

Testing output

14/1/97  
35

Files for Installation test

6/6/97  
15/1/97

36

Stepwise Manual Example output

ANALYSIS OF VARIATION FOR DEPENDENT VARIABLE Y  
 SOURCE OF VARIATION FOR INDEPENDENT VARIABLES SELECTED IN STEP 1

1	1.000000	1.000000
2	0.000000	0.000000
3	0.000000	0.000000
4	0.000000	0.000000
5	0.000000	0.000000
6	0.000000	0.000000
7	0.000000	0.000000
8	0.000000	0.000000
9	0.000000	0.000000
10	0.000000	0.000000
11	0.000000	0.000000
12	0.000000	0.000000
13	0.000000	0.000000
14	0.000000	0.000000
15	0.000000	0.000000
16	0.000000	0.000000
17	0.000000	0.000000
18	0.000000	0.000000
19	0.000000	0.000000
20	0.000000	0.000000
21	0.000000	0.000000
22	0.000000	0.000000
23	0.000000	0.000000
24	0.000000	0.000000
25	0.000000	0.000000
26	0.000000	0.000000
27	0.000000	0.000000
28	0.000000	0.000000
29	0.000000	0.000000
30	0.000000	0.000000
31	0.000000	0.000000
32	0.000000	0.000000
33	0.000000	0.000000
34	0.000000	0.000000
35	0.000000	0.000000
36	0.000000	0.000000
37	0.000000	0.000000
38	0.000000	0.000000
39	0.000000	0.000000
40	0.000000	0.000000
41	0.000000	0.000000
42	0.000000	0.000000
43	0.000000	0.000000
44	0.000000	0.000000
45	0.000000	0.000000
46	0.000000	0.000000
47	0.000000	0.000000
48	0.000000	0.000000
49	0.000000	0.000000
50	0.000000	0.000000

TITLE: EXAMPLE OF STEPWISE REGRESSION DATA FROM DEAPER AND SMITH, P. 204-205  
 DATA LABORATORY FOR THE BMS REGRESSION FROM MATH 1000  
 ANALYSIS OF REGRESSION FOR VARIABLE Y  
 (TABLE 1)

SOURCE	D.F.	SS	MS	F	PROB > F
REGRESSION	2	2461.061	1230.5305	174.4257	
RESIDUAL	10	74.75212	7.475212		
TOTAL	12	2535.813			

VARIABLE NUMBER	VARIABLE NAME	REGRESSION COEFFICIENT	STANDARD ERROR OF REGRESSION COEFFICIENT	T-STAT	PROB >  T	LOWER 95% CI	UPPER 95% CI
1	X1	1.439958	0.087112	16.530	<.0001	1.265734	1.614182
2	X2	-0.613956	0.087112	-7.048	<.0001	-0.790110	-0.437802

ROOT MEAN SQUARE ERROR OF ESTIMATE IS 2.73394

UNTIL THE SEQUENCE NUMBER FOR THIS ANOVA = 10  
 DELETED = 121

THIS IS THE VALUE THAT R<sup>2</sup> WILL DECREASE TO IF X<sub>2</sub> IS NOW DELETED FROM THE MODEL, I.E.

SIGNIFICANCE LEVEL HAS TO BE < SIGN = .05 IF THE VARIABLE TO ENTER THE MODEL

$$.5339 = .97247 - \frac{1190.9146}{2715.7631}$$

FINAL REGRESSION EQUATION:  $\hat{Y} = 105.09758 + 1.439958 X_1 - .613956 X_2$

15/1/97

MODEL, F.E.  
0.5339 = 0.97247 - 1190.9246  
2715.7631

FINAL REGRESSION EQUATION:  $\hat{Y} = 103.09738 + 1.4399583 * X_1 - .61395363 * X_4$

01/01/07

Resting output

E:EXE - exampl.out

File Edit Options Help

NAME X1 X4 TITLE, EXAMPLE OF STEPWISE OPTION (DATA FROM DRAPER AND SMITH, PG 365-402) PAGE 8

SANDIA LABORATORIES <>> STEPWISE REGRESSION <>> FROM KANSAS STATE UNIVERSITY

ANOVA TABLE  
ANALYSIS OF REGRESSION FOR VARIABLE 5---Y  
(TABLE 1)

SOURCE	D.F.	SS	MS	F	SIGNIFICANCE
REGRESSION	2 ✓	2641.0010 ✓	1320.5005 ✓	176.62696 ✓	0.0000
RESIDUAL	10 ✓	74.762112 ✓	7.4762112 ✓		
TOTAL	12 ✓	2715.7631 ✓			

R\*\*2 IS 0.97247 ✓  
INTERCEPT IS 103.09738 ✓  
STANDARD ERROR OF INTERCEPT IS 2.12398 ✓

VARIABLE NUMBER	VARIABLE NAME	REGRESSION COEFFICIENTS	STANDARDIZED REGRESSION COEFFICIENTS	PARTIAL SSQ	T-TEST VALUES	R**2 DELETES	ALPHA HATS
1	X1	1.4399583 ✓	0.563052 ✓	809.1048 ✓	10.4031 ✓	0.6745 ✓	0.0000
4	X4	-.61395363 ✓	-0.603107 ✓	1190.9246 ✓	-12.6212 ✓	0.5339 ✓	0.0000

UNIQUE SEQUENCE NUMBER FOR THIS ANOVA = 102

PRESS IS 121.22 ✓

TITLE, EXAMPLE OF STEPWISE OPTION (DATA FROM DRAPER AND SMITH, PG 365-402) PAGE 9

SANDIA LABORATORIES <>> STEPWISE REGRESSION <>> FROM KANSAS STATE UNIVERSITY

1/6/07  
37



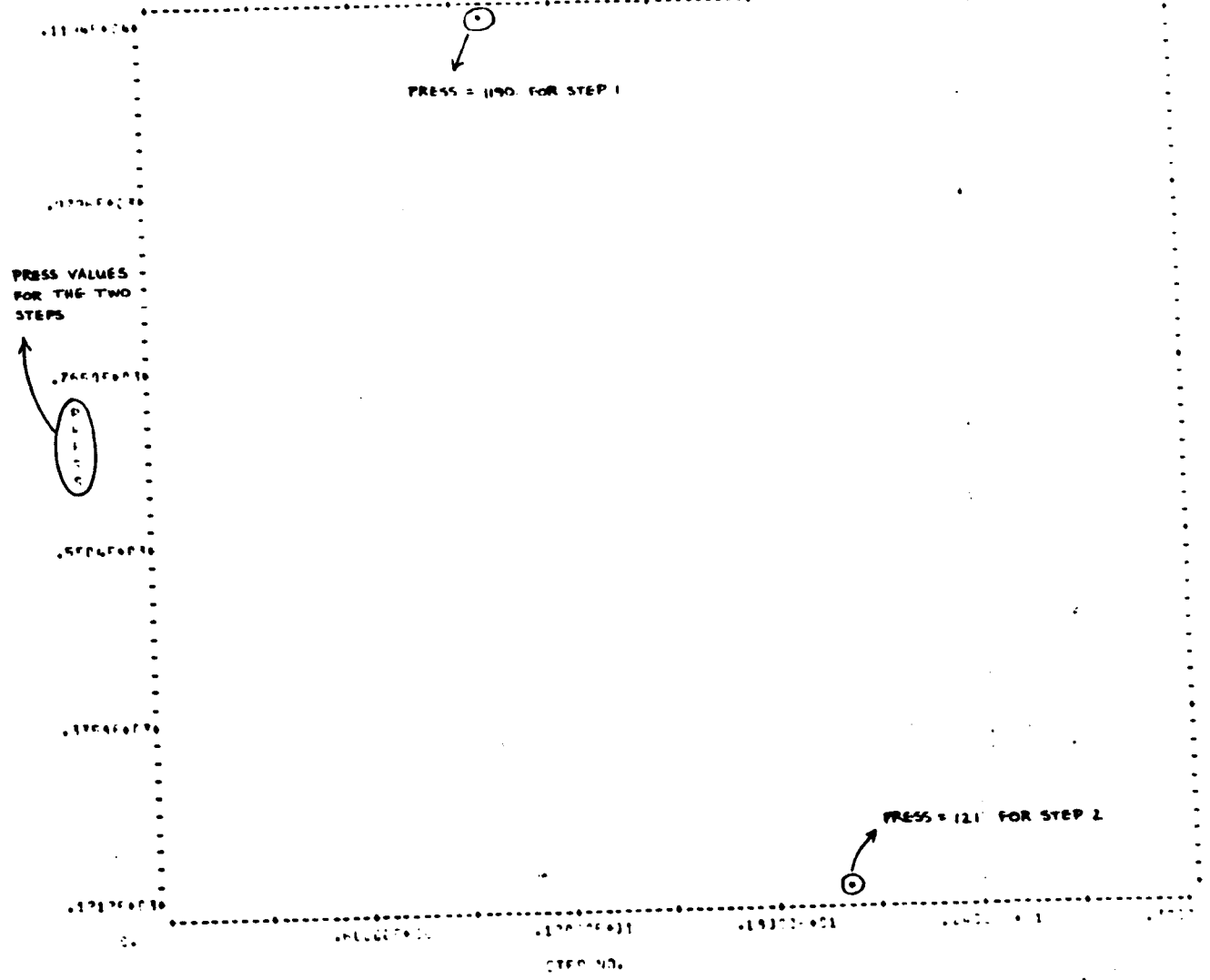
Files for Installation test

6/19/97

38

My

Separate manual Example A4 part



17/1

E.EXE - exampl.out

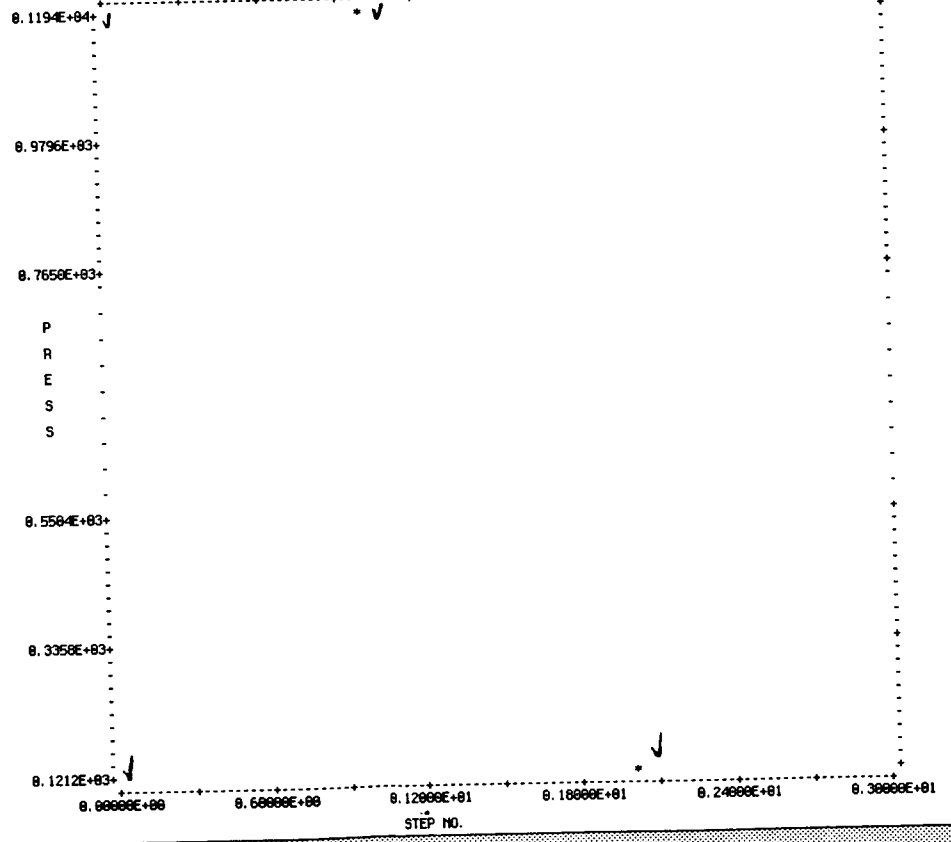
File Edit Options Help

PRESS IS 121.22

TITLE, EXAMPLE OF STEPWISE OPTION (DATA FROM DRAPER AND SMITH, PG 365-402)

PAGE 9

SANDIA LABORATORIES <>> STEPWISE REGRESSION <>> FROM KANSAS STATE UNIVERSITY



6/19/97  
My

Testing Output

18/1/97

39

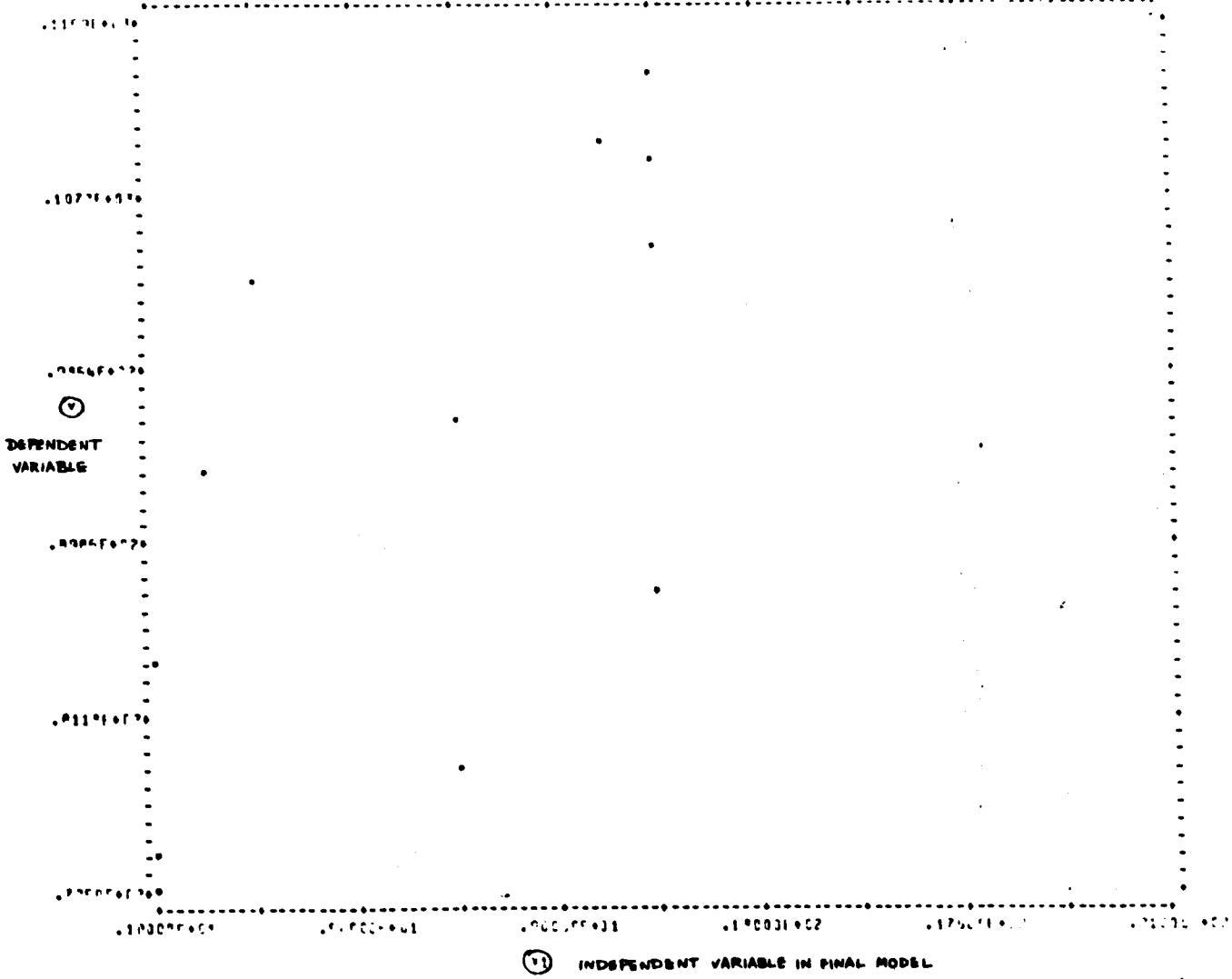
Files for Installation test

40

6/19/47

MS

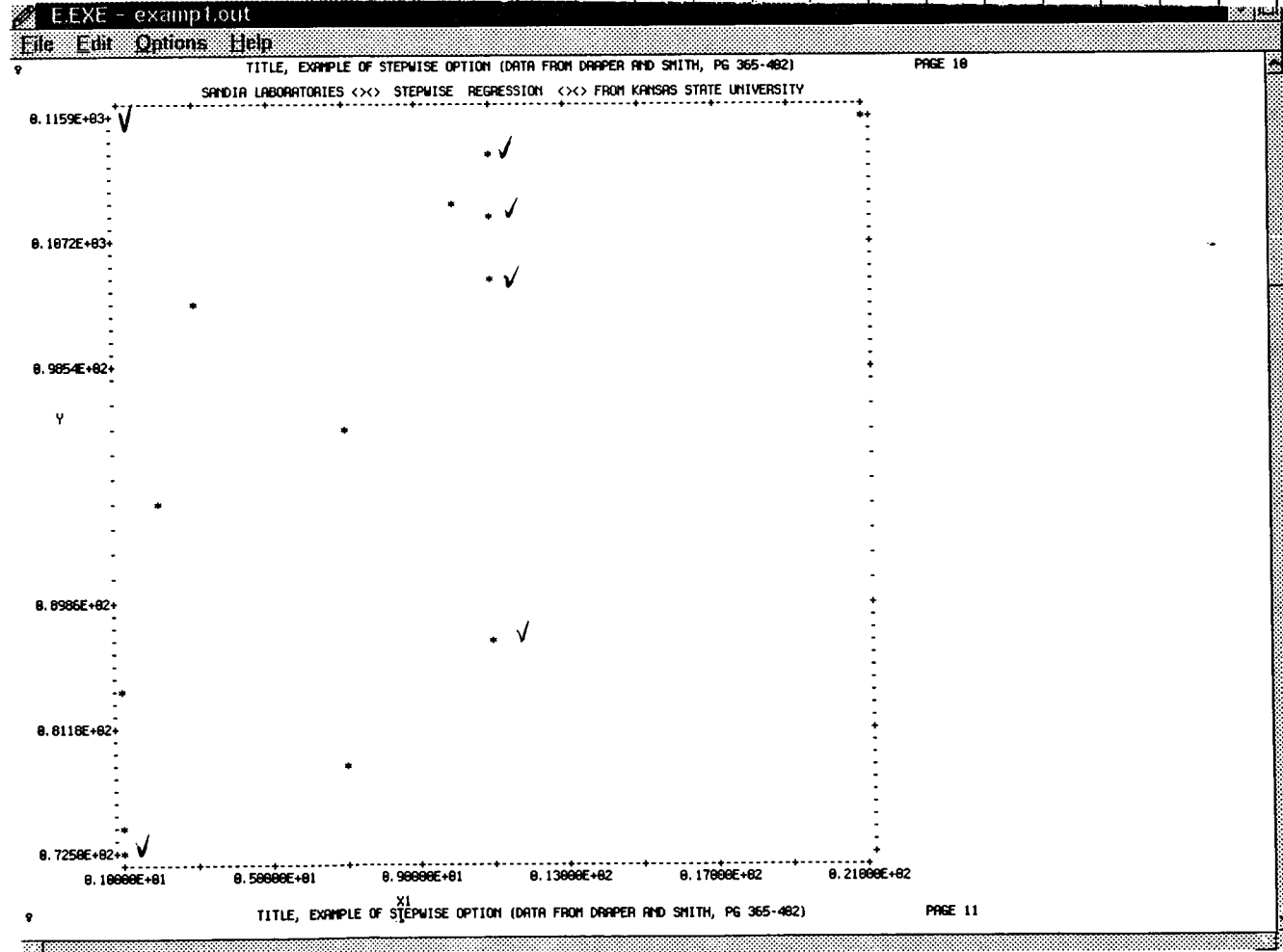
Storage Manual Example Output



19/47

(1) INDEPENDENT VARIABLE IN FINAL MODEL

My  
6/9/97



Testing output

20/9/97  
41

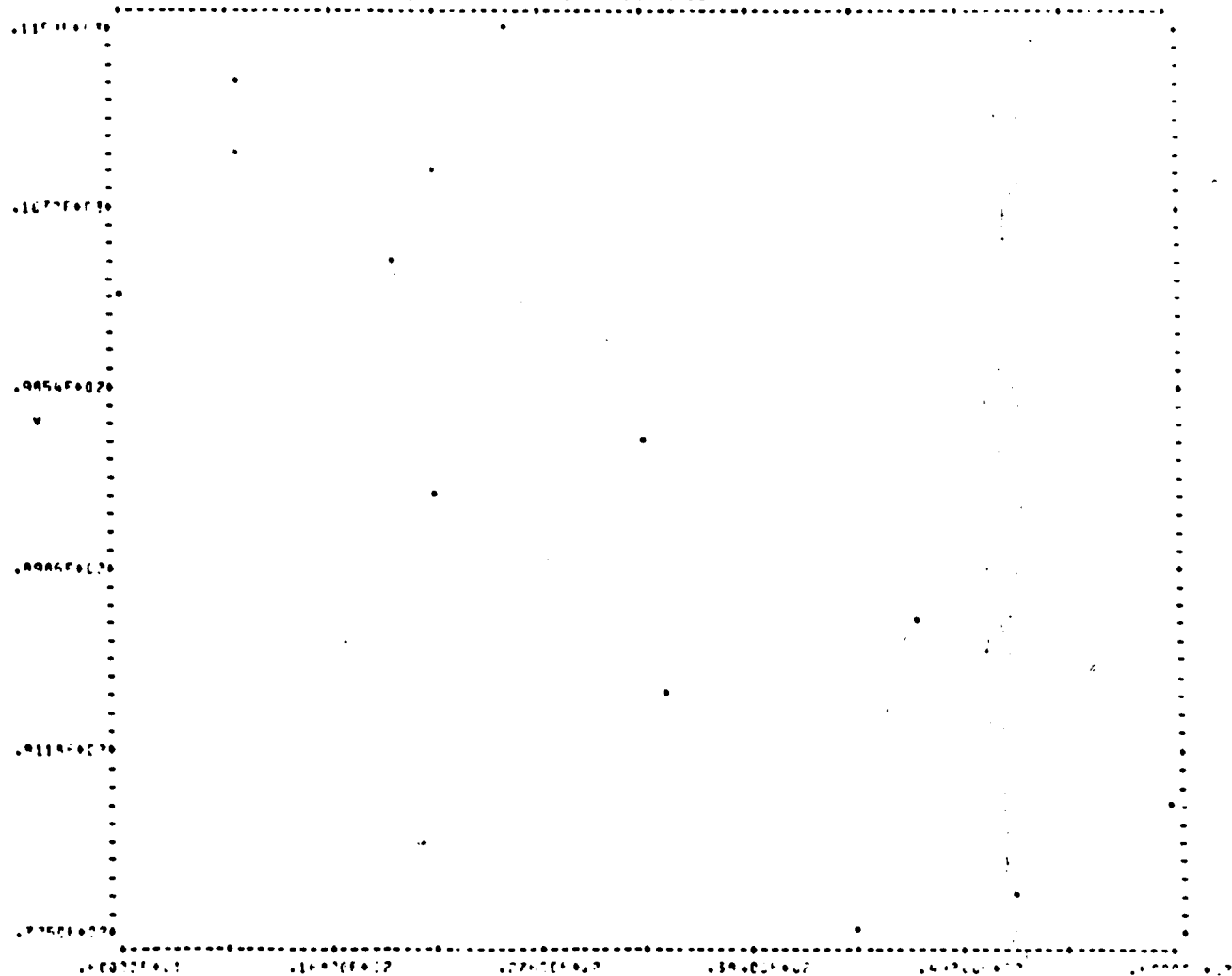
Files for Installation test

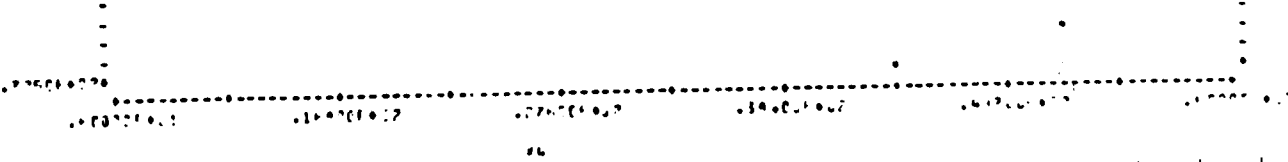
42

10/9/97  
MS

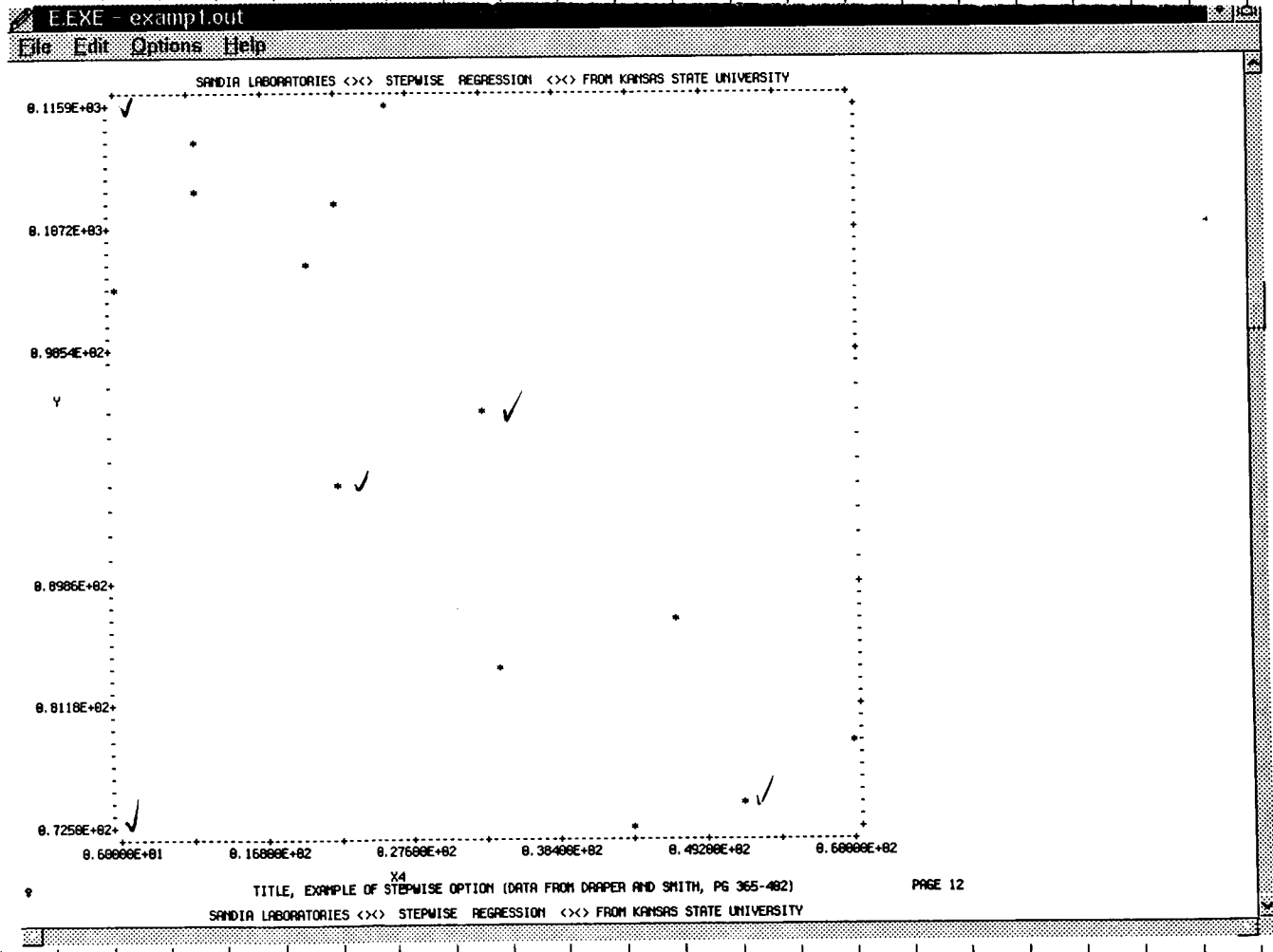
Stopover Naval Grounds O'Hart

2/9/97





6/15/87  
AB



Testin Output

23/43  
1/47

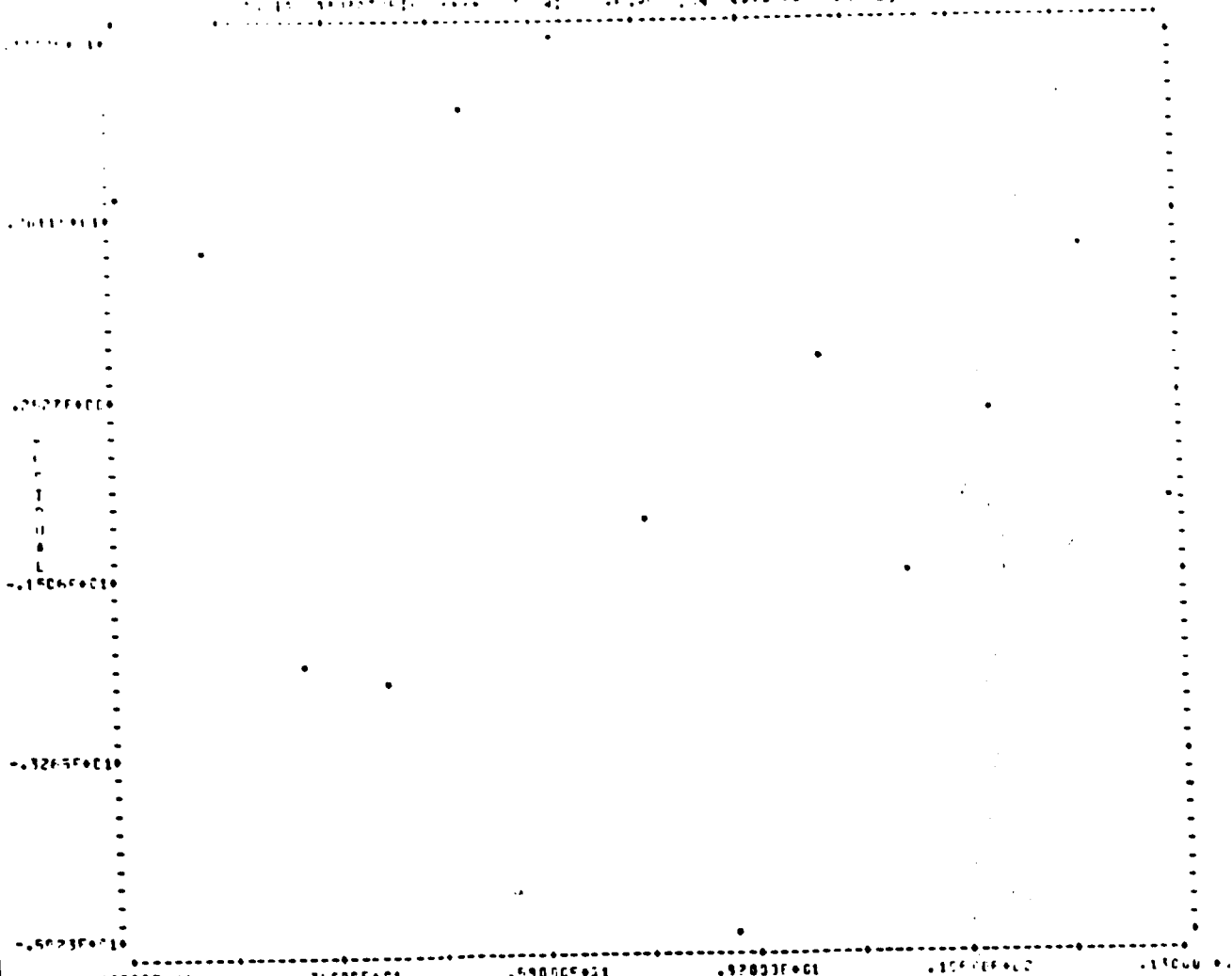
Files for Installation test

44

6/19/47  
pp

Stoppage Mineral Sample Sheet

23  
47/47



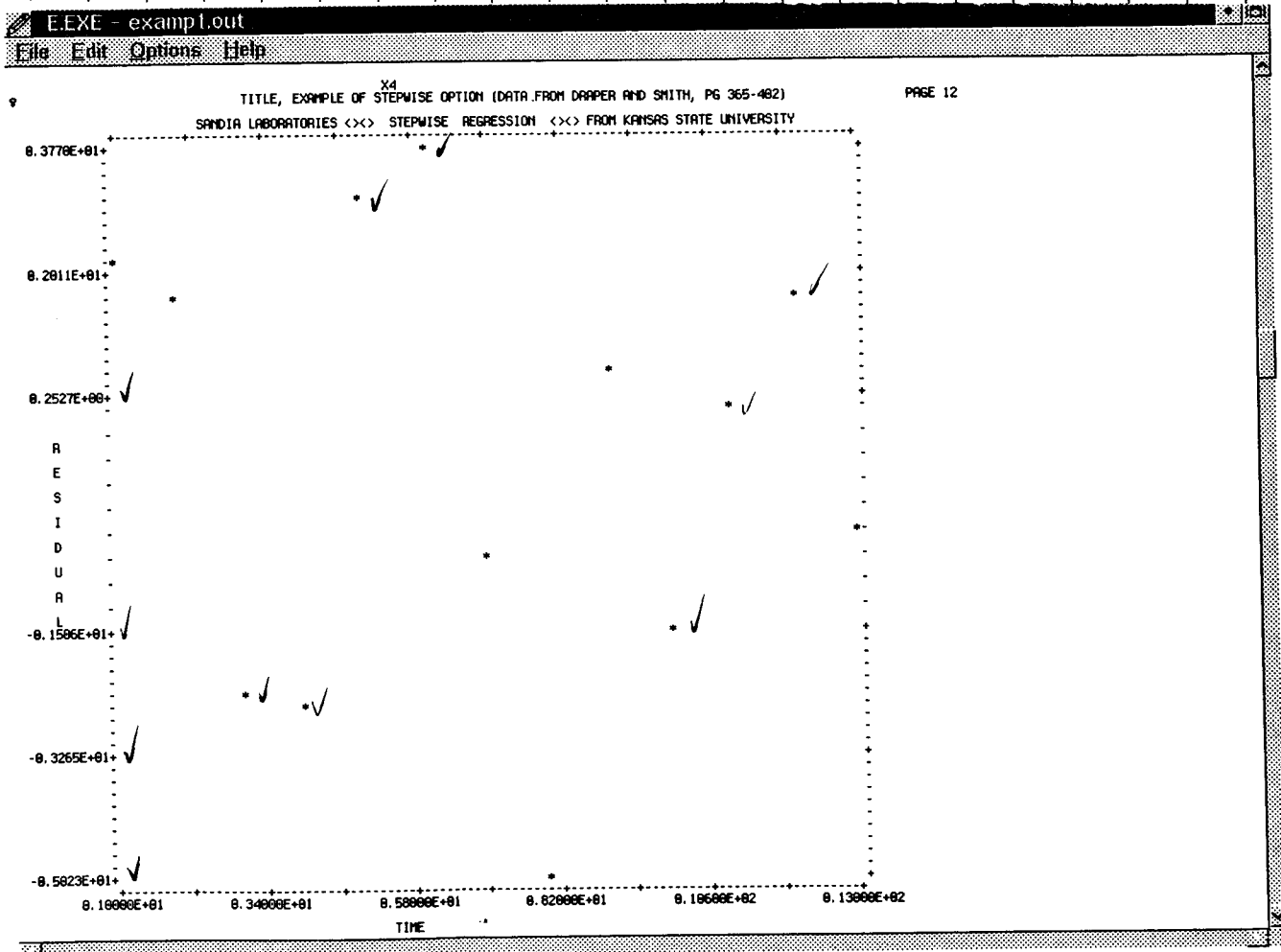
(TIME)

ORDER IN WHICH DATA WAS READ IN

.50000E+01    .10000E+01    .20000E+01    .30000E+01    .40000E+01    .50000E+01    .60000E+01    .70000E+01    .80000E+01    .90000E+01    .10000E+02

ORDER IN WHICH DATA WAS READ IN

6/19/97  
 NY



Test's output

24/45  
 6/19/97



Files for Installation Test

46  
6/19/74

Stepwise Manual Example 21-10d

TITLE:EXAMPLE OF BACKWARD OPTION (DATA FROM DRAPER AND SMITH, P. 144-147)  
SANDIA LABORATORIES <<<> STEPWISE REGRESSION <<<> FROM KANSAS STATE UNIVERSITY

INPUT CHECK ON DATA

V1	V2	V3	V4	V
1	2	3	4	5
7.000000	26.00000	6.000000	41.00000	74.50000

NO. RAW DATA INPUT = 17  
NO. TRANSFORMED OBSERVATIONS = 17  
NO. OF OBSERVATIONS DROPPED = 0

TITLE:EXAMPLE OF BACKWARD OPTION (DATA FROM DRAPER AND SMITH, P. 144-147)  
SANDIA LABORATORIES <<<> STEPWISE REGRESSION <<<> FROM KANSAS STATE UNIVERSITY

VARIABLE NAME	VARIABLE NUMBER	MEAN	VARIANCE	STD. DEV.	STD. COV.	STD. CORR.
V1	1	7.68196	36.4026	6.03364	1.63164	0.4000
V2	2	48.1530	242.161	15.5630	6.11501	0.1731
V3	3	11.7497	41.0256	6.40513	1.77667	0.2792
V4	4	30.0000	220.157	14.8346	4.66776	0.3171
V	5	45.6211	224.314	15.0037	6.17210	0.4077

13 OBSERVATIONS

TITLE:EXAMPLE OF BACKWARD OPTION (DATA FROM DRAPER AND SMITH, P. 144-147)  
SANDIA LABORATORIES <<<> STEPWISE REGRESSION <<<> FROM KANSAS STATE UNIVERSITY

CORRELATION MATRIX

V1	1	1.0000				
V2	2	.7206	1.0000			
V3	3	-.4261	-.1302	1.0000		
V4	4	-.2654	-.0790	.0906	1.0000	
V	5	.7407	.4163	-.4447	-.4211	1.0000

NO.	1	2	3	4	5
NAME	V1	V2	V3	V4	V

RETURN TO PROMPT FOR INPUT

FLAG INDICATING HIGHLY CORRELATED INPUT VARIABLES.  
IF MATRIX IS SINGULAR A MESSAGE WILL BE PRINTED AND PROCESSING STOPPED

25  
1/47

FLAG INDICATES HIGHLY CORRELATED INPUT VARIABLES.  
 IF MATRIX IS SINGULAR A MESSAGE WILL BE PRINTED AND PROCESSING STOPPED

6/9/97  
 AS

Testing output

E.EXE - exampl.out

File Edit Options Help

TITLE, EXAMPLE OF BACKWARD OPTION (DATA FROM DRAPER AND SMITH, P 364-482) PAGE 2

SANDIA LABORATORIES <<> STEPWISE REGRESSION <<> FROM KANSAS STATE UNIVERSITY

VARIABLE NAME	VARIABLE NUMBER	MEAN	VARIANCE	STD. DEV.	STD. ERR.	C.V.
X1	1	7.46154	34.6825	5.88239	1.63148	78.84
X2	2	48.1538	242.141	15.5689	4.31531	32.31
X3	3	11.7692	41.8256	6.46513	1.77546	54.42
X4	4	38.8988	288.167	16.7382	4.64234	55.79
Y	5	95.4231	226.314	15.0437	4.17238	15.77

13 OBSERVATIONS

TITLE, EXAMPLE OF BACKWARD OPTION (DATA FROM DRAPER AND SMITH, P 364-482) PAGE 3

SANDIA LABORATORIES <<> STEPWISE REGRESSION <<> FROM KANSAS STATE UNIVERSITY

CORRELATION MATRIX

	1	2	3	4	5
X1	1.0000				
X2	0.2286	1.0000			
X3	-0.8241	-0.1392	1.0000		
X4	-0.2454	-0.9736	0.8295	1.0000	
Y	0.7387	0.8163	-0.5347	-0.8213	1.0000

NO. 1 2 3 4 5

NAME X1 X2 X3 X4 Y

ROUNDING ERROR IN ROW 3

269 47  
 1/97

6/14/91  
mg

Stepwise Method Example Output

TITLE: SAMPLE OF HACKWARE OPTION (DATA FROM DEARLE AND SMITH, P. 100)  
SANDIA LABORATORIES STEPWISE REGRESSION FROM RANKED DATA (LIVE) 17

ROW TABLE  
ANALYSIS OF REGRESSION FOR VARIABLE 5---9  
TABLE 11

SOURCE	D.F.	SS	MS	F	SIGNIFICANT
REGRESSION	4	2667.8996	666.9749	111.67917	.0000
RESIDUAL	12	47.963679	3.996973		
TOTAL	17	2715.8631			

CONST IS .98274  
INTERCEPT IS 42.405369  
STANDARD ERROR OF INTERCEPT IS 70.0710

VARIABLE NUMBER	VARIABLE NAME	REGRESSION COEFFICIENTS	STANDARDIZED REGRESSION COEFFICIENTS	PARTIAL SS	T-TEST VALUES @ D.F.	PROB DELT	R-SQ
1	X1	1.5511026	.676517	25.9533	2.0927	.0724	.3270
2	X2	.51016758	.527766	2.9725	.7669	.4611	.2213
3	X3	.10170940	.043700	.1091	.1150	.9127	.0000
4	X4	-.14406107	-.150787	.2470	-.2087	.4424	.0000

UNIQUE SEQUENCE NUMBER FOR THIS ANALYSIS = 101

PROB IS 110.

SINCE NO VARIABLES WERE FORCED TO STAY IN THE MODEL, THE VARIABLE WITH THE LARGEST  $\hat{R}^2 > \text{SIG} = .05$  WILL BE DELETED.

27/47

6/11/97

Testing output

TITLE, EXAMPLE OF BACKWARD OPTION (DATA FROM DRAPER AND SMITH, P 364-482) PAGE 4  
SANDIA LABORATORIES <>> STEPWISE REGRESSION <>> FROM KANSAS STATE UNIVERSITY

ANOVA TABLE  
ANALYSIS OF REGRESSION FOR VARIABLE 5---Y  
(TABLE 1)

SOURCE	D.F.	SS	MS	F	SIGNIFICANCE
REGRESSION	4	2567.8994	641.97485	111.47917	0.0000
RESIDUAL	8	47.863639	5.9829549		
TOTAL	12	2715.7631			

R<sup>2</sup> IS 0.98238 ✓  
 INTERCEPT IS 62.485369 ✓  
 STANDARD ERROR OF INTERCEPT IS 70.0710 ✓

VARIABLE NUMBER	VARIABLE NAME	REGRESSION COEFFICIENTS	STANDARDIZED REGRESSION COEFFICIENTS	PARTIAL SSQ	T-TEST VALUES	R <sup>2</sup> DELETES	ALPHA HATS
1	X1	1.5511826	0.686512	25.9509	2.0827	0.9728	0.0708
2	X2	0.51816758	0.527786	2.9725	0.7049	0.9813	0.5099
3	X3	0.18198948	0.043398	0.1891	0.1358	0.9823	0.8959
4	X4	-0.14486183	-0.168287	0.2470	-0.2832	0.9823	0.8441

UNIQUE SEQUENCE NUMBER FOR THIS ANOVA = 183

PRESS IS 110.35 ✓ TITLE, EXAMPLE OF BACKWARD OPTION (DATA FROM DRAPER AND SMITH, P 364-482) PAGE 5  
SANDIA LABORATORIES <>> STEPWISE REGRESSION <>> FROM KANSAS STATE UNIVERSITY

ANOVA TABLE  
ANALYSIS OF REGRESSION FOR VARIABLE 5---Y  
(TABLE 1)

SOURCE	D.F.	SS	MS	F	SIGNIFICANCE
--------	------	----	----	---	--------------

28/49  
1/47

6/19/67  
M

Stepwise Manual Example of First

TITLE: EXAMPLE OF FORWARD OPTION (DATA FROM GRABER AND SMITH, P. 365-462)  
SANDIA LABORATORIES <<< STEPWISE REGRESSION <<<< FROM KANSAS STATE UNIVERSITY

ANOVA TABLE  
ANALYSIS OF REGRESSION FOR VARIABLE: 5---V  
TABLE 11

SOURCE	D.F.	SS	MS	F	SIGNIFICANCE
REGRESSION	3	2667.7960	889.2653	155.93160	.0000
RESIDUAL	12	47.072729	3.9227274		
TOTAL		2715.7631			

R<sup>2</sup> IS .99234  
INTERCEPT IS 71.669307  
STANDARD ERROR OF INTERCEPT IS 14.1474

VARIABLE NUMBER	VARIABLE NAME	REGRESSION COEFFICIENTS	STANDARDIZED REGRESSION COEFFICIENTS	PARTIAL SS	T-TEST VALUES	NO D.F.	CRITICAL VALUES
1	X1	1.4519190	.547737	820.9176	17.4116	12	.6961
2	X2	.41416976	.490616	76.7716	2.7419	11	.6975
3	X3	-.2364422	-.769193	9.9319	-1.3472	10	.6977

ALPHA = .05  
CRITICAL VALUES  
17.4116  
2.7419  
-1.3472  
MOST VARIABLE TO BE DELETED

UNIQUE SEQUENCE NUMBER FOR THIS ANOVA = 104

PRESS IS 45.6

TITLE: EXAMPLE OF BACKWARD OPTION (DATA FROM GRABER AND SMITH, P. 365-462)  
SANDIA LABORATORIES <<< STEPWISE REGRESSION <<<< FROM KANSAS STATE UNIVERSITY

ANOVA TABLE  
ANALYSIS OF REGRESSION FOR VARIABLE: 5---V  
TABLE 11

SOURCE	D.F.	SS	MS	F	SIGNIFICANCE
REGRESSION	2	2657.4544	1328.7272	229.58376	.0000
RESIDUAL	13	47.094687	3.6226682		
TOTAL		2715.7631			

R<sup>2</sup> IS .97968  
INTERCEPT IS 52.577769  
STANDARD ERROR OF INTERCEPT IS 7.29617

VARIABLE NUMBER	VARIABLE NAME	REGRESSION COEFFICIENTS	STANDARDIZED REGRESSION COEFFICIENTS	PARTIAL SS	T-TEST VALUES	NO D.F.	CRITICAL VALUES
1	X1	1.46693857	.576137	860.4719	17.1047	13	.6659
2	X2	.66225369	.445017	1207.7423	14.4426	12	.6673

ALPHA = .05  
CRITICAL VALUES  
17.1047  
14.4426  
BOTH < SIG = .05  
NO MORE DELETED

UNIQUE SEQUENCE NUMBER FOR THIS ANOVA = 105

4/1/62

UNIQUE SEQUENCE NUMBER FOR THIS ANOVA =	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	

E.EXE - exam1.out

File Edit Options Help

SANDIA LABORATORIES <<> STEPWISE REGRESSION <<> FROM KANSAS STATE UNIVERSITY

ANOVA TABLE  
ANALYSIS OF REGRESSION FOR VARIABLE 5---Y  
(TABLE 1)

SOURCE	D.F.	SS	MS	F	SIGNIFICANCE
REGRESSION	3	2667.7983	889.26345	166.83168	0.0000
RESIDUAL	9	47.972729	5.3303633		
TOTAL	12	2715.7631			

R-squared IS 0.98234  
INTERCEPT IS 71.648387  
STANDARD ERROR OF INTERCEPT IS 14.1424

VARIABLE NUMBER	VARIABLE NAME	REGRESSION COEFFICIENTS	STANDARDIZED REGRESSION COEFFICIENTS	PARTIAL SSQ	T-TEST VALUES	R-squared DELETES	ALPHA HATS
1	X1	1.4519388 ✓	0.567737 ✓	828.9974 ✓	12.4188 ✓	0.6881 ✓	0.0000
2	X2	0.41618976 ✓	0.438414 ✓	26.7894 ✓	2.2418 ✓	0.9725 ✓	0.0517
4	X4	-.23654822 ✓	-0.263183 ✓	9.9318 ✓	-1.3658 ✓	0.9787 ✓	0.2854

UNIQUE SEQUENCE NUMBER FOR THIS ANOVA = 104

PRESS IS 85.351

TITLE, EXAMPLE OF BACKWARD OPTION (DATA FROM DRAPER AND SMITH, P 364-482)

PAGE 6

---

SANDIA LABORATORIES <<> STEPWISE REGRESSION <<> FROM KANSAS STATE UNIVERSITY

ANOVA TABLE  
ANALYSIS OF REGRESSION FOR VARIABLE 5---Y  
(TABLE 1)

SOURCE	D.F.	SS	MS	F	SIGNIFICANCE
REGRESSION	2 ✓	2657.8586 ✓	1328.9293 ✓	229.58378 ✓	0.0000 ✓
RESIDUAL	10 ✓	57.904483 ✓	5.7904483 ✓		
TOTAL	12 ✓	2715.7631 ✓			

R-squared IS 0.97868  
INTERCEPT IS 52.577349  
STANDARD ERROR OF INTERCEPT IS 2.28617

VARIABLE NUMBER	VARIABLE NAME	REGRESSION COEFFICIENTS	STANDARDIZED REGRESSION COEFFICIENTS	PARTIAL SSQ	T-TEST VALUES	R-squared DELETES	ALPHA HATS
1	X1	1.4683857 ✓	0.574137 ✓	848.4319 ✓	12.1847 ✓	0.6663 ✓	0.0000
2	X2	0.66225849 ✓	0.685817 ✓	1287.7823 ✓	14.4424 ✓	0.5339 ✓	0.0000

UNIQUE SEQUENCE NUMBER FOR THIS ANOVA = 105

PRESS IS 93.883

TITLE, EXAMPLE OF BACKWARD OPTION (DATA FROM DRAPER AND SMITH, P 364-482)

PAGE 7

---

SANDIA LABORATORIES <<> STEPWISE REGRESSION <<> FROM KANSAS STATE UNIVERSITY

0.1183E+03

4/19/97  
Testing 0 input

30/1/97

Files for Installation for

52

6/9/91  
MG

Steamer Manual Sample Sheet

.1107E+06

⊙

STEP 1 IN BACKWARD

.1051E+06

.1003E+06

D  
P  
S

.9535E+05

⊙

STEP 5

.9035E+05

.8535E+05

⊙

STEP 2

B.

.8035E+05

.7535E+05

.7035E+05

.6535E+05

STEP NO.

31/47

0.00000E+00

STEP 1

0.

0.00000E+00

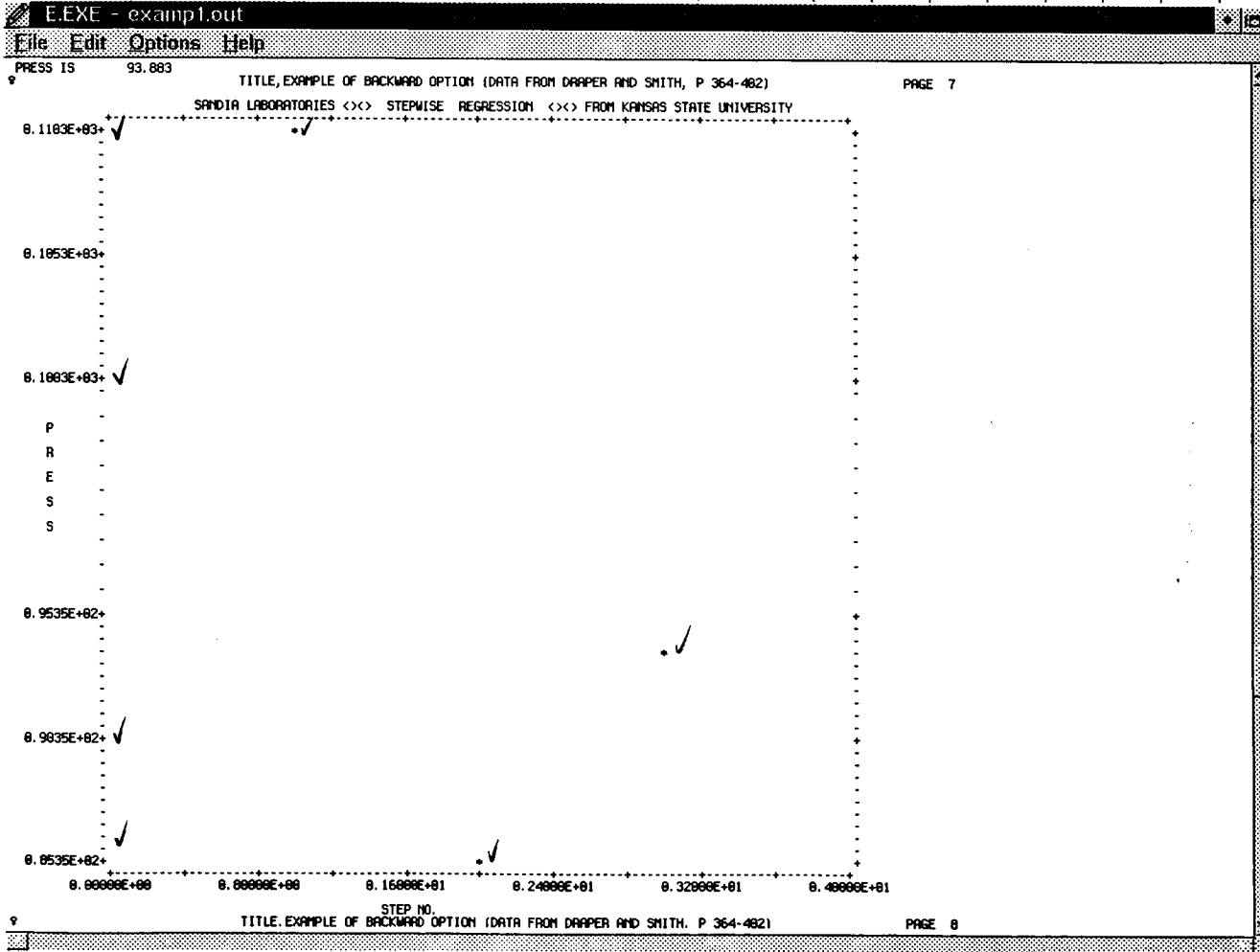
0.16000E+01

0.32000E+01

0.48000E+01

STEP NO.

2/19/97  
MS



testis about

32/197



files for Installation test

54

6/15/67

W

Stepwise Manual Sample Output

TITLE: EXAMPLE OF BACKWARD OPTION DATA FROM DEPT. OF STATISTICS - KANSAS STATE UNIVERSITY  
SANDIA LABORATORIES \*\*\* STEPWISE REGRESSION \*\*\* FROM KANSAS STATE UNIVERSITY

TABLE OF RESIDUALS FOR VARIABLE 1 \*\*\*\*\*

TIME	CUMULATED VALUE	PROBABILITY OF REJ.	RESIDUAL
1	78.4000	93.3763	-1.5766
2	76.2999	75.9577	1.2643
3	106.130	135.915	-1.51674
4	97.8000	93.2595	-1.65444
5	75.9000	97.2325	-1.39751
6	109.700	105.157	4.84751
7	117.707	106.837	-1.30707
8	72.5000	76.5756	-2.0756
9	98.1000	91.7755	1.07551
10	115.900	116.539	1.16539
11	81.0000	93.5457	1.26647
12	117.700	117.637	1.96275
13	109.600	117.239	-2.93766

END OF ANALYSIS FOR 2ND PASS ON THE DATA

SANDIA LABORATORIES \*\*\* STEPWISE REGRESSION PROGRAM \*\*\* COURTESY OF DEPT. OF STATISTICS - KANSAS STATE UNIVERSITY

TITLE: EXAMPLE OF BACK REGRESSION WITH THE STEPWISE OPTION DATA FROM DEPT. OF STATISTICS

DATA: 5.0.0.

INPUT CHECK OF PARAMETERS

LIST CONTROL CARD

NUMBER OF VARIABLES READ IN = 5

NO. OF TRANSFORMED VARIABLES = 0

DATA DISPOSITION IS 2

LABEL(1)=RANK(1), RANK(2), RANK(3), RANK(4), RANK(5) -> THESE LABELS WILL REPLACE

LIST CONTROL CARD

MODEL: 5-1-2-3-4-5

THE LABELS USED ON THE PREVIOUS TWO PASSES

LIST CONTROL CARD

STEPWISE: SIGN=.25, SIGOUT=.10

LIST CONTROL CARD

PGSS

LIST CONTROL CARD

BACK REGRESSION -> A REGRESSION ANALYSIS WILL BE PERFORMED ON THE RANKS OF THE DATA.

LIST CONTROL CARD

OUTPUT: CORR, STEPS, RESIDUALS. -> SAME OUTPUT AS 2ND PASS.

LIST CONTROL CARD

END OF PARAMETERS

LIST CONTROL CARD

3/1/67

LABEL(1)=RANK(X1),RANK(X2),RANK(X3),RANK(X4),RANK(Y) → THESE LABELS WILL REPLACE  
 THE LABELS USED ON THE PREVIOUS TWO PASSES  
 MODEL,5=1+2+3+4.  
 STEPWISE,SIGIN=.05,SIGOUT=.10  
 PRESS  
 RANK REGRESSION → A REGRESSION ANALYSIS WILL BE PERFORMED ON THE RANKS OF THE DATA.  
 OUTPUT,CORR,STEPS,RESIDUALS. → SAME OUTPUT AS 2ND PASS.  
 END OF PARAMETERS

(STAT CONTROL CARD)  
 (STAT CONTROL CARD)  
 (STAT CONTROL CARD)  
 (STAT CONTROL CARD)  
 (STAT CONTROL CARD)  
 (STAT CONTROL CARD)

E.EXE - exampl.out

File Edit Options Help

TITLE,EXAMPLE OF BACKWARD OPTION (DATA FROM DRAPER AND SMITH, P 364-402) PAGE 0

SANDIA LABORATORIES <>> STEPWISE REGRESSION <<<> FROM KANSAS STATE UNIVERSITY

TABLE OF RESIDUALS FOR VARIABLES 5----Y

TIME	OBSERVED VALUE	PREDICTED VALUE	RESIDUAL	95 PER CENT CONF INT (MEAN)	95 PER CENT CONF INT (INDIV)
1	78.5800	88.8748	-1.57489	77.3858	82.7612
2	74.3800	73.2589	1.04908	70.5871	75.9947
3	104.300	105.815	-1.51474	103.966	107.664
4	87.6800	89.2585	-1.65848	86.6196	91.8974
5	95.9600	97.2925	-1.39251	95.7421	98.8429
6	109.200	105.152	4.04751	103.333	106.972
7	102.700	104.082	-1.38285	100.777	107.227
8	72.5000	74.5754	-2.07542	71.9422	77.2086
9	93.1800	91.2755	1.82451	89.0861	93.5449
10	115.900	114.538	1.36246	110.561	110.514
11	63.0000	60.5357	3.26433	70.2356	82.8357
12	113.300	112.437	0.862756	110.060	114.815
13	109.400	112.293	-2.89344	109.812	114.736

SANDIA LABORATORIES <>> STEPWISE REGRESSION PROGRAM <<<> COURTESY OF DEPT. OF STATISTICS - KANSAS STATE UNIVERSITY

TITLE,EXAMPLE OF RANK REGRESSION WITH STEPWISE OPTION (DATA FROM DRAPER/SMITH)

DATA,5,0,2. (STAT CONTROL CARD)

INPUT CHECK OF PARAMETERS

NUMBER OF VARIABLES READ IN = 5

NO. OF TRANSFORMED VARIABLES = 0

DATA DISPOSITION IS 2

LABEL(1)=RANK(X1),RANK(X2),RANK(X3),RANK(X4),RANK(Y) (STAT CONTROL CARD)

MODEL,5=1+2+3+4. (STAT CONTROL CARD)

STEPWISE,SIGIN=.05,SIGOUT=.10 (STAT CONTROL CARD)

PRESS (STAT CONTROL CARD)

RANK REGRESSION (STAT CONTROL CARD)

OUTPUT,CORR,STEPS,RESIDUALS (STAT CONTROL CARD)

END OF PARAMETERS (STAT CONTROL CARD)

TITLE,EXAMPLE OF RANK REGRESSION WITH STEPWISE OPTION (DATA FROM DRAPER/SMITH) PAGE 1

SANDIA LABORATORIES <>> STEPWISE REGRESSION <<<> FROM KANSAS STATE UNIVERSITY

INPUT CHECK ON DATA

RANK(X1)	RANK(X2)	RANK(X3)	FIRST OBSERVATION RANK(X4)	RANK(Y)
1	2	3	4	5
6.5000000	1.0000000	2.5000000	13.0000000	3.0000000

NO. RAW DATA INPUT = 13

NO. TRANSFORMED OBSERVATIONS = 13

NO. OF OBSERVATIONS DROPPED = 0

TITLE,EXAMPLE OF RANK REGRESSION WITH STEPWISE OPTION (DATA FROM DRAPER/SMITH) PAGE 2

This output is not shown in the manual example output, thus indicating the installed version is later than the manual version. Since this feature appears to have been added. If these numbers will be used for future work the a pop sample will be checked w/ hand calculator.

6/9/97 leg

testis output

34/47 55

Files for Installation for

6/19/47

56

TITLE: EXAMPLE OF RANK REGRESSION WITH THE STEPWISE OPTION (DATA FROM GRAPE) WITH  
SANDIA LABORATORIES <<<> STEPWISE REGRESSION <<<> FROM KANSAS STATE UNIVERSITY

INPUT CHECK ON DATA

RANK(X1)	RANK(X2)	RANK(X3)	RANK(X4)	RANK(X5)
1	2	3	4	5
6.500000	1.000000	2.930000	14.300000	7.300000

NOTE: THESE ARE THE RANKS ASSIGNED TO THE FIRST OBSERVATION  
RANKS ARE ASSIGNED BY VARIABLE NUMBER

EXACT OBSERVATION (X1)

AVERAGE RANK RESULTING FROM TIED INPUT OBSERVATIONS.

NO. RAW DATA INPUT = 13  
NO. TRANSFORMED OBSERVATIONS = 13  
NO. OF OBSERVATIONS JUMPED = 0

TITLE: EXAMPLE OF RANK REGRESSION WITH THE STEPWISE OPTION (DATA FROM GRAPE) WITH  
SANDIA LABORATORIES <<<> STEPWISE REGRESSION <<<> FROM KANSAS STATE UNIVERSITY

VARIABLE NAME	VARIABLE NUMBER	MEAN	VARIANCE	STD. DEV.	STD. ERR.	T.V.
RANK(X1)	1	7.00000	16.5617	4.0695	1.5743	16.49
RANK(X2)	2	7.00000	15.1250	3.8899	1.1746	16.55
RANK(X3)	3	7.00000	14.9157	3.8622	1.1714	16.17
RANK(X4)	4	7.00000	15.0483	3.8787	1.1771	16.04
RANK(X5)	5	7.00000	15.1657	3.8944	1.1801	16.61

13 OBSERVATIONS

AVERAGE RANK ASSIGNED FOR 13 OBSERVATIONS IS 7.

WITHOUT TIED OBSERVATIONS THE VALUES IN THESE COLUMNS WILL ALL BE THE SAME WITHIN A COLUMN

TITLE: EXAMPLE OF RANK REGRESSION WITH THE STEPWISE OPTION (DATA FROM GRAPE) WITH  
SANDIA LABORATORIES <<<> STEPWISE REGRESSION <<<> FROM KANSAS STATE UNIVERSITY

CORRELATION MATRIX

RANK(X1)	1	1.0000				
RANK(X2)	2	.9901	1.0000			
RANK(X3)	3	-.7146	.0577	1.0000		
RANK(X4)	4	-.8320	-.3903	-.0416	1.0000	
RANK(X5)	5	.7912	.7379	-.4640	-.7571	1.0000

RANK CORRELATIONS

NO. 1 2 3 4 5

NAME RANK(X1) RANK(X2) RANK(X3) RANK(X4) RANK(X5)

Stepwise Manual Example of Rank

6/19/47

NO. 1 2 3 4 5  
 NAME RANK(X1) RANK(X2) RANK(X3) RANK(X4) RANK(Y)

```

E.EXE - exampl.out
File Edit Options Help
TITLE, EXAMPLE OF RANK REGRESSION WITH STEPWISE OPTION (DATA FROM DRAPER/SMITH)

DATA, 5, 0, 2. (STAT CONTROL CARD)

INPUT CHECK OF PARAMETERS
NUMBER OF VARIABLES READ IN = 5
NO. OF TRANSFORMED VARIABLES = 0
DATA DISPOSITION IS 2

LABEL(1)=RANK(X1), RANK(X2), RANK(X3), RANK(X4), RANK(Y) (STAT CONTROL CARD)
MODEL, 5=1+2+3+4. (STAT CONTROL CARD)
STEPWISE, SIGIN=.05, SIGOUT=.10 (STAT CONTROL CARD)
PRESS (STAT CONTROL CARD)
RANK REGRESSION (STAT CONTROL CARD)
OUTPUT, CORR, STEPS, RESIDUALS (STAT CONTROL CARD)
END OF PARAMETERS (STAT CONTROL CARD)
TITLE, EXAMPLE OF RANK REGRESSION WITH STEPWISE OPTION (DATA FROM DRAPER/SMITH) PAGE 1
SANDIA LABORATORIES <>> STEPWISE REGRESSION <>> FROM KANSAS STATE UNIVERSITY
INPUT CHECK ON DATA

RANK(X1) RANK(X2) RANK(X3) FIRST OBSERVATION RANK(X4) RANK(Y)
1 2 3 4 5
6.5000000 ✓ 1.0000000 ✓ 2.5000000 ✓ 13.0000000 ✓ 3.0000000 ✓
NO. RAW DATA INPUT = 13
NO. TRANSFORMED OBSERVATIONS = 13 ✓
NO. OF OBSERVATIONS DROPPED = 0 ✓
TITLE, EXAMPLE OF RANK REGRESSION WITH STEPWISE OPTION (DATA FROM DRAPER/SMITH) PAGE 2
SANDIA LABORATORIES <>> STEPWISE REGRESSION <>> FROM KANSAS STATE UNIVERSITY

VARIABLE VARIABLE MEAN VARIANCE STD. DEV. STD. ERR. C.V.
NAME NUMBER
RANK(X1) 1 7.00000 ✓ 14.5417 ✓ 3.81335 ✓ 1.05763 ✓ 54.48 ✓
RANK(X2) 2 7.00000 ✓ 15.1250 ✓ 3.88989 ✓ 1.07064 ✓ 55.56 ✓
RANK(X3) 3 7.00000 ✓ 14.9167 ✓ 3.86221 ✓ 1.07110 ✓ 55.17 ✓
RANK(X4) 4 7.00000 ✓ 15.0033 ✓ 3.88373 ✓ 1.07115 ✓ 55.48 ✓
RANK(Y) 5 7.00000 ✓ 15.1667 ✓ 3.89444 ✓ 1.08012 ✓ 55.63 ✓

13 OBSERVATIONS
TITLE, EXAMPLE OF RANK REGRESSION WITH STEPWISE OPTION (DATA FROM DRAPER/SMITH) PAGE 3
SANDIA LABORATORIES <>> STEPWISE REGRESSION <>> FROM KANSAS STATE UNIVERSITY
CORRELATION MATRIX ← checked ok
RANK(X1) 1 1.0000
RANK(X2) 2 0.3381 1.0000
RANK(X3) 3 -0.7186 0.0527 1.0000
RANK(X4) 4 -0.3320 -0.9963 -0.8086 1.0000
RANK(Y) 5 0.7912 0.7373 -0.4488 -0.7521 1.0000
  
```

6/19/97  
 808  
 Test from output

30/4/97  
 57

Files for *[illegible]*

6/9/97  
Mg

58

TITLE: EXAMPLE OF RANK REGRESSION WITH THE STEPWISE OPTION (DATA FROM DRAPER/SMITH) PAGE 9  
 JANJIA LABORATORIES <>> STEPWISE REGRESSION <>> FROM KANSAS STATE UNIVERSITY

ANOVA TABLE  
 ANALYSIS OF REGRESSION FOR VARIABLE 5---RANK(1)  
 (TABLE 1)

SOURCE	D.F.	SS	MS	F	SIGNIFICANCE
REGRESSION	1	113.93123	113.93123	10.411433	.0013
RESIDUAL	11	68.368764	6.1880698		
TOTAL	12	182.00000			

R\*\*2 IS .62600  
 INTERCEPT IS 1.3038395  
 STANDARD ERROR OF INTERCEPT IS 1.40743

VARIABLE NAME	PARAMETER NAME	REGRESSION COEFFICIENTS	STANDARD ERROR COEFFICIENTS	PARTIAL CORRELATION	T-TEST VALUE	P-VALUE
1	RANK(1)	.68802292	.791199	113.9312	9.2989	0.0000

UNIQUE SEQUENCE NUMBER FOR THIS ANOVA = 106  
 RANK FIT GIVES A RAW DATA NORMALIZED R\*\*2 = .64292846  
 COEFFICIENT OF INTERPOLATION = .64403359E-01  
 PRESS IS 30.347

THE USER SHOULD NOTE THAT THIS R\*\*2 VALUE IS CALCULATED AS

$$\frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (\hat{y}_i - \bar{y})^2 + \sum (y_i - \hat{y}_i)^2}$$

AND SINCE THE  $\hat{y}_i$  VALUES HAVE BEEN OBTAINED THROUGH INTERPOLATION THIS CALCULATION WILL NOT NECESSARILY RESULT IN THE SAME R\*\*2 VALUE AS THE USUAL CALCULATION

$$\frac{\sum (y_i - \bar{y})^2}{\sum (y_i - \bar{y})^2}$$

SINCE THE  $\hat{y}_i$  VALUES ARE NOT THE LEAST SQUARES ESTIMATES, THE QUANTITIES  $\sum (\hat{y}_i - \bar{y})^2$  AND  $\sum (y_i - \hat{y}_i)^2$  ARE NOT ORTHOGONAL. THAT IS, THE SUM OF CROSS-PRODUCTS  $\sum (y_i - \hat{y}_i)(\hat{y}_i - \bar{y})$  IS NON-ZERO. IF THE SUM OF CROSS-PRODUCTS IS NEAR ZERO, THEN THE COEFFICIENT OF INTERPOLATION WILL BE NEAR ZERO. IF IT IS LARGE, THEN THE COEFFICIENT WILL BE NEAR ONE. SEE THE TEXT FOR FURTHER DISCUSSION.

Stepwise Manual would print

37/47

SINCE THE  $\hat{y}_i$  VALUES ARE NOT THE LEAST SQUARES ESTIMATES, THE QUANTITIES  $\sum (\hat{y}_i - \bar{y})^2$  AND  $\sum (\hat{y}_i - \hat{y}_i)^2$  ARE NOT ORTHOGONAL. THAT IS, THE SUM OF CROSS-PRODUCTS  $\sum (\hat{y}_i - \hat{y}_i)(\hat{y}_i - \bar{y})$  IS NON-ZERO. IF THE SUM OF CROSS-PRODUCTS IS NEAR ZERO, THEN THE COEFFICIENT OF INTERPOLATION WILL BE NEAR ZERO. IF IT IS LARGE, THEN THE COEFFICIENT WILL BE NEAR ONE. SEE THE TEXT FOR FURTHER DISCUSSION.

6/9/97  
AT

Testing output

E.EXE - exampl.out

File Edit Options Help

NO.	1	2	3	4	5
NAME	RANK(X1)	RANK(X2)	RANK(X3)	RANK(X4)	RANK(Y)

TITLE, EXAMPLE OF RANK REGRESSION WITH STEPWISE OPTION (DATA FROM DRAPER/SMITH) PAGE 4

SANDIA LABORATORIES <>> STEPWISE REGRESSION <>> FROM KANSAS STATE UNIVERSITY

ADV TABLE  
ANALYSIS OF REGRESSION FOR VARIABLE 5---RANK(Y)  
(TABLE 1)

SOURCE	D.F.	SS	MS	F	SIGNIFICANCE
REGRESSION	11 ✓	113.93123 ✓	113.93123 ✓	18.411433 ✓	0.0013 ✓
RESIDUAL	11 ✓	68.868768	6.188698 ✓		
TOTAL	12 ✓	182.80000 ✓			

R\*\*2 IS 0.62688  
INTERCEPT IS 1.3438395  
STANDARD ERROR OF INTERCEPT IS 1.48783

VARIABLE NUMBER	VARIABLE NAME	REGRESSION COEFFICIENTS	STANDARDIZED REGRESSION COEFFICIENTS	PARTIAL SSD	T-TEST VALUES	R**2 DELETES	ALPHA PARTS
1	RANK(X1) 0.80882292 ✓	0.791199 ✓	113.9312 ✓	4.2989 ✓	0.0008 ✓	0.0013	

UNIQUE SEQUENCE NUMBER FOR THIS ANOVA = 186

RANK FIT GIVES A RAW DATA NORMALIZED R\*\*2 = 0.64292846

COEFFICIENT OF INTERPOLATION = 0.68489359E-01

PRESS IS 90.347

*unexplained diff  
is .0007  
6/9/97*

38/47

59

6/1/97  
M

60

Stepwise Manual source data

TITLE: EXAMPLE OF RANK REGRESSION WITH THE STEPWISE OPTION DATA FROM DRAPER/SMITH  
 ANJIA L. CHARLES (C) STEPWISE REGRESSION (C) FROM KANSAS STATE UNIVERSITY

PAGE 5

ANOVA TABLE  
 ANALYSIS OF REGRESSION FOR VARIABLE 3---RANKED  
 (TABLE 1)

SOURCE	D.F.	SS	MS	F	SIGNIFICANCE
REGRESSION	2	12.92243 ✓	6.461215 ✓	42.700700 ✓	.0000 ✓
RESIDUAL	13	19.07717 ✓	1.467474 ✓		
TOTAL	15	32.00000 ✓			

R<sup>2</sup> IS .49518  
 INTERCEPT IS 6.509944  
 STANDARD ERROR OF INTERCEPT IS 1.31214

VARIABLE NUMBER	VARIABLE NAME	REGRESSION COEFFICIENTS	STANDARDIZED REGRESSION COEFFICIENTS	PARTIAL SSG	T-TEST VALUES	R <sup>2</sup> DELETES	ALPHA HATS
1	RANKED1	.6219140 ✓	.604501 ✓	55.7422 ✓	5.6873 ✓	.2656 ✓	.0007
4	RANKED4	-.550024 ✓	-.550024 ✓	49.9716 ✓	-5.8670 ✓	.2600 ✓	.0011

UNIQUE SEQUENCE NUMBER FOR THIS ANOVA = 107

RANK FIT GIVES A RAW DATA NORMALIZED R<sup>2</sup> = .5244414

COEFFICIENT OF INT. CORRELATION = .19334400E-01

FRESH IS 35.410

39  
6/1/97

6/19/97  
RS

Facts output

TITLE, EXAMPLE OF RANK REGRESSION WITH STEPWISE OPTION (DATA FROM DRAPER/SMITH) PAGE 5  
SANDIA LABORATORIES <>> STEPWISE REGRESSION <>> FROM KANSAS STATE UNIVERSITY

ANOVA TABLE  
ANALYSIS OF REGRESSION FOR VARIABLE 5---RANK(Y)  
(TABLE 1)

SOURCE	D.F.	SS	MS	F	SIGNIFICANCE
REGRESSION	2	162.92283 ✓	81.461413 ✓	42.709984 ✓	0.0000 ✓
RESIDUAL	18	19.077175 ✓	1.977175 ✓		
TOTAL	12	182.00000 ✓			

R\*\*2 IS 0.89518  
INTERCEPT IS 6.5099988  
STANDARD ERROR OF INTERCEPT IS 1.31214

VARIABLE NUMBER	VARIABLE NAME	REGRESSION COEFFICIENTS	STANDARDIZED REGRESSION COEFFICIENTS	PARTIAL SSO	T-TEST VALUES	R**2 DELETES	ALPHA HATS
1	RANK (X1)	0.62154188 ✓	0.688681 ✓	59.9822 ✓	5.6873 ✓	0.5656 ✓	0.8002
4	RANK (X4)	-0.55154162 ✓	-0.550824 ✓	48.9916 ✓	-5.0676 ✓	0.6268 ✓	0.8085

explained diff  
of ~ .0005  
by 6/19/97

UNIQUE SEQUENCE NUMBER FOR THIS ANOVA = 107  
RANK FIT GIVES A RAW DATA NORMALIZED R\*\*2 = 0.9284814  
COEFFICIENT OF INTERPOLATION = 0.14538880E-01  
PRESS IS 35.860

TITLE, EXAMPLE OF RANK REGRESSION WITH STEPWISE OPTION (DATA FROM DRAPER/SMITH) PAGE 6  
SANDIA LABORATORIES <>> STEPWISE REGRESSION <>> FROM KANSAS STATE UNIVERSITY

6/19/97



Files for Installation

6/9/97  
My 62

Sheppard Manual Sample Output

TITLE: EXAMPLE OF RANK REGRESSION WITH THE STEPWISE OPTION: DATA FROM CARPENTER WITH

TABLE OF RESIDUALS FOR VARIABLES 5-----RANKED

TIME	RANK OF Y	PREDICTED RANK OF Y	RANK RESIDUAL	RANK Y	RANK Y-82	RANK Y-81
1	3.C	7.17000	-.379073	79.5361	79.5361	79.5361
2	2.C	11.13659	-.456617	76.1800	76.1800	76.1800
3	9.C	17.4330	-1.43300	104.900	104.900	104.900
4	9.C	6.96071	-1.36223	87.5801	87.5801	87.5801
5	7.C	6.11769	-.457113	85.3301	85.3301	85.3301
6	10.C	17.0127	-.27076F-32	109.201	109.201	109.201
7	8.C	7.36417	-1.36617	102.70F	102.70F	102.70F
8	1.0	7.21767	-1.21767	72.5301	72.5301	72.5301
9	6.0	6.96269	-.373133E-01	91.1001	91.1001	91.1001
10	11.C	10.7297	-.270765	115.980	115.980	115.980
11	6.C	2.78071	1.21177	83.8001	83.8001	83.8001
12	12.C	11.6577	-.342666	113.101	113.101	113.101
13	11.C	15.1035	-.496521	109.600	109.600	109.600

RESIDUAL SUM OF SQUARES ON RANK DATA = 700.000

TERMINATION ON END OF FILE - UNIT 5

SEE MAN AND COMPUR (1979) FOR AN EXPLANATION OF HOW THESE RESIDUALS ARE CALCULATED.

4/1/97

19/97  
MS

Testing output

E:EXE - exampl.out  
File Edit Options Help

TABLE OF RESIDUALS FOR VARIABLES 5---RANK(Y)

TIME	RANK OF Y	PREDICTED RANK OF Y	RANK RESIDUAL	RAW Y	RAW YHAT	RAW RESIDUAL
1	3.0	3.379984	-0.379979	78.5000	80.5139	-2.01389
2	2.0	1.134584	0.865417	74.3000	72.7422	1.55775
3	9.0	10.3390	-1.03902	104.300	109.366	-5.06598
4	9.0	6.96923	-1.96923	67.6000	95.8138	-8.21384
5	7.0	6.13769	0.862313	95.9000	93.4855	2.41448
6	10.0	10.0027	-0.270074E-02	109.200	109.281	-0.541747E-03
7	8.0	9.06617	-1.06617	182.700	104.624	-1.92421
8	7.0	9.06617	-1.23767	72.5000	75.2962	-2.79029
9	1.0	2.23767	0.373130E-01	93.1000	92.6948	0.205221
10	6.0	5.96269	2.27075	115.900	109.346	6.55415
11	13.0	10.7293	2.27075	63.8000	77.6147	6.18533
12	4.0	2.78921	1.21079	113.300	111.964	1.33648
13	12.0	11.6573	0.342666	109.400	109.221	0.179384
13	11.0	10.1035	0.896521			

RESIDUAL SUM OF SQUARES ON RAW DATA = 200.052 ✓  
(STAT CONTROL CARD)

THE ABOVE CARD IS AN UNDEFINED PARAMETER CARD AND IS TAKEN AS A COMMENT  
TERMINATION ON END OF FILE - UNIT 5

4/2/97

6/9/97  
My

Conclusions:

The results of the installation testing indicate that the code is producing the same output as shown in the manual when example problems are run. 2 minor deviations were noted and considered to not impact results. These tests confirm that the code has been installed correctly (and is running as intended by the authors) per the procedures 4/9/97 procedures in TOP-018.

Files for Installation to

SAND 79 1472

Unlimited Release

1980

CODE 6

44/47

Example pages from Manual  
for STEPWISE/ code.

X+1

16

6/12/97

44/47

# Stepwise Regression With PRESS and Rank Regression (Program User's Guide)

# MASTER

Ronald L. Iman, James M. Davenport, Elizabeth L. Frost, Michael J. Shortencarier

Printed January 1980



Sandia Laboratories

SAND79-1472  
Unlimited Release  
Printed January 1980

45/47

Stepwise Regression with PRESS and Rank Regression  
(Program User's Guide)

Ronald L. Iman  
Statistics and Computing Division 1223  
Sandia Laboratories  
Albuquerque, New Mexico 87185

James M. Davenport  
Department of Mathematics  
Texas Tech University  
Lubbock, Texas 79409

Elizabeth L. Frost  
Statistics and Computing Division 1223  
Sandia Laboratories  
Albuquerque, New Mexico 87185

Michael J. Shortencarier  
Statistics and Computing Division 1223  
Sandia Laboratories  
Albuquerque, New Mexico 87185

848 0767

ABSTRACT

This report contains a description of a stepwise multiple regression program. This program provides for either a forward stepwise or backward elimination solution to multiple regression problems. The program also provides PRESS values that can be used for subset selection as well as providing options for regression analysis on the ranks of the data and for a weighted regression analysis on either raw or rank transformed data. This document has been written and designed for users of this STEPWISE program.

DISCLAIMER

This document contains information which is classified "Secret" under Executive Order 11652, dated August 14, 1972, and is being released to the public pursuant to the provisions of Executive Order 11652, dated August 14, 1972, which provides that information which is classified "Secret" and which is not required to be kept secret in the interest of national defense shall be declassified and released to the public as soon as practicable. This document is being released to the public as a matter of course.

46/47

## STEPWISE REGRESSION

### I. Introduction

The description contained herein details how to use the STEPWISE multiple regression program which was originally obtained through the courtesy of Dr. K. E. Kemp, Department of Statistics, Kansas State University. The program incorporates the most desirable features of various available programs into one program.

The features of this program are these: (1) fit any simple or multiple regression model with up to 179 independent variables, (2) find the "best" subset by backward elimination procedure, i.e., fit the full model specified and then progressively drop variables until the only variables left are those that are significant at some specified level of significance, or (3) find the "best subset using a stepwise procedure, i.e., start with the independent variable<sup>X</sup> in the model that explains the most variation in the dependent variable<sup>Y</sup> and then add variables from those in the model statement in order of their additional contribution to the model as measured by their partial correlation with the dependent variable. However, at each step, those variables already in the model are examined regarding their current contribution to the regression sum of squares and level of significance, the variable contributing least (smallest t-statistic) will be dropped. The procedure continues until only variables (that are not significant at a specified level) remain out of the model, (4) the user may force certain variables to appear in the final model in either the backward or stepwise solution, (5) the user may also request residual plots for the final model.

The following features have been added to the program at Sandia: (6) a regression analysis on the ranks of the data with all the previously mentioned features is also available when requesting RANK Regression, (7) obtain PRESS\* values that can be used in subset selection (the PRESS value is associated with each model fitted), (8) the regression coefficients of the fitted models from any one of the models considered within execution of STEPWISE can be saved on disk, (9) the observed values and predicted

\*See explanation on page 9.

values from rank regression analysis can also be saved on disk for additional investigation outside the STEPWISE program, (10) weighted regression analysis on the raw data or rank transformed data.

47/47

The reader who is familiar with SAND76-0364 (Iman, 1976) and has used STEPWISE on previous occasions is advised to read this report to note the extensive additions.

## II. OUTPUT

The program will output means, variances, standard deviations, standard errors and coefficients of variation for each variable read in or generated via transformations. As options, the user may request the table of correlation coefficients, sum of squares and cross-products matrix, inverse correlation matrix, and residuals  $y$ -values and  $\hat{y}$ 's, PRESS values and a plot of these PRESS values. The program will also print an analysis of variance table for a regression model and a table of statistics regarding the regression coefficients. Residual plots may be requested, and if they are, will be printed on the line printer.

See pg 10  
3rd xcr

## III. INPUT - PARAMETER CARDS

All parameter cards must start in column 1 and may be placed in any order. An explanation and illustration of each of the parameter cards follows.

### A. TITLE card (optional).

The "TITLE" card may contain any alphameric data that may be meaningful to the user. The word "TITLE" identifies the card, and its contents are printed at the top of each page of output. Only one "TITLE" card may be used.

Example:

```
TITLE, THIS IS A SAMPLE TITLE CARD
```

### B. DATA card (required).

The "DATA" card has 3 arguments specified as follows: