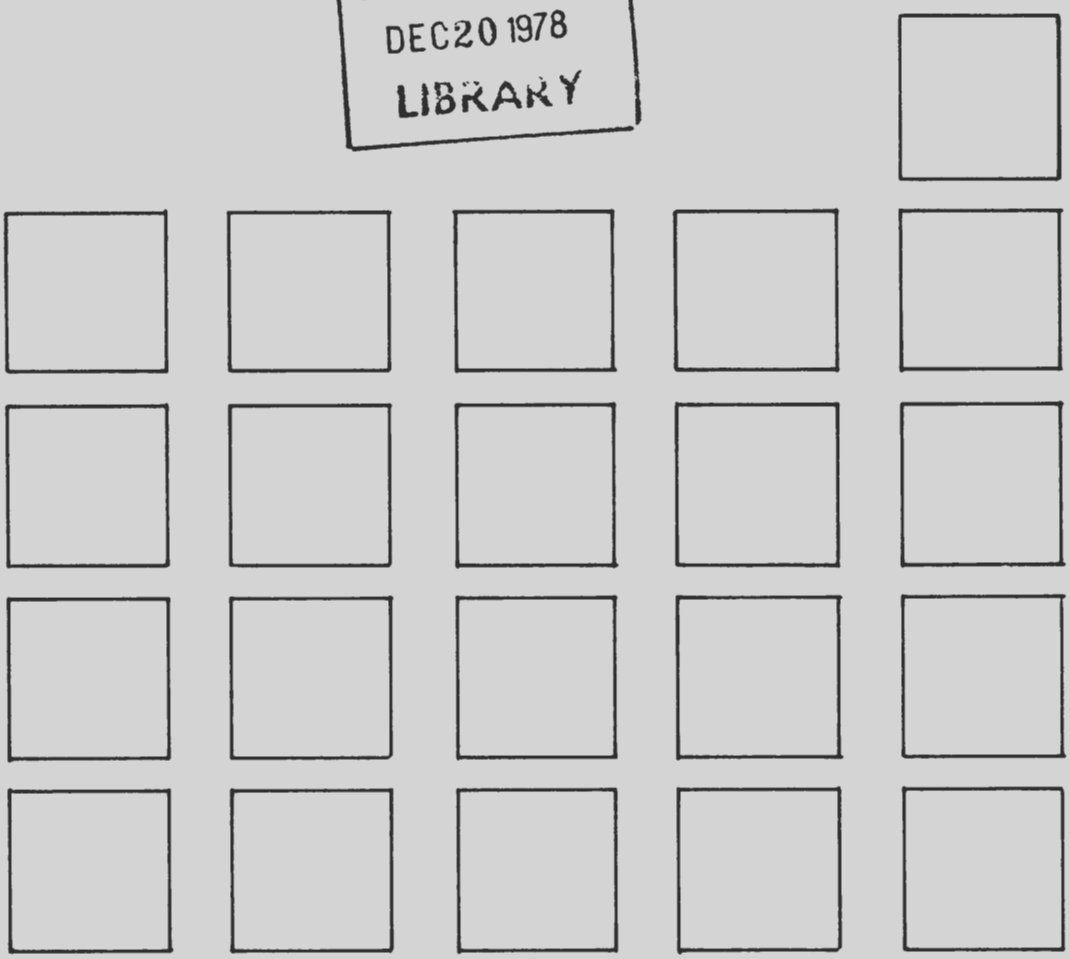
1978

Linear Regression Analysis Using a Programable Pocket Calculator

Paul Evenson

Linear Regression Analysis Using a Programable Pocket Calculator



Agricultural Experiment Station
South Dakota State University
Brookings, South Dakota

Linear Regression Analysis Using a Programable Pocket Calculator

By: Paul D. Evenson
Assoc. Professor
Plant Science and Statistics

Linear regression is an extremely useful "least squares" technique for fitting a linear equation to a set of data.

This program calculates means, sums of squares, and sums of cross-products of the dependent and independent values which are entered only once. It also calculates the slope and intercept of the line as well as the coefficients of determination and correlation. The program calculates a predicted dependent variable from a given independent variable and a predicted independent variable from a given dependent variable.

The program computes the reduction in sum of squares due to regression, residual sum of squares and degrees of freedom, variance about regression, standard error of the estimate, the standard deviation about the slope, and the t-test on the slope of the line.

Confidence limits are calculated about the slope, about a predicted $\hat{Y}$, about a predicted future $\hat{Y}$, and about a predicted future $\hat{\bar{y}}$ made up of n_1 observations. This program was written for a Hewlett-Packard 97 programable "pocket" calculator.

The symbol Y_i denotes the i^{th} dependent variable where i is the i^{th} pair and i = 1, 2, ..., n. The symbol X_i denotes the i^{th} independent variable in the i^{th} pair.

The main hypothesis tested in this analysis is H_o : B = 0, which says that the true slope of the regression line is equal to zero, or to say it another way, X has no influence on Y. The alternative hypothesis is H_1 : $B \neq 0$ or that X does influence Y. The null hypothesis (H_o : B = 0) can be tested by several different methods.

Formulas Used in Regression Analysis

1. Means, sum of squares and sum of cross-products of dependent and independent variables.

a. mean of $X = \dfrac{\Sigma X_i}{n} = \bar{x}$

b. sum of squares of $X = \Sigma X_i^2 - \dfrac{\left(\Sigma X_i\right)^2}{n} = \Sigma x^2$

c. mean of $Y = \frac{\Sigma Y_i}{n} = \bar{y}$

d. sum of squares of $Y = \Sigma Y_i^2 - \frac{\left(\Sigma Y_i\right)^2}{n} = \Sigma y^2$

e. sum of cross-products $= \Sigma X_i Y_i - \frac{\left(\Sigma X_i\right)\left(\Sigma Y_i\right)}{n} = \Sigma xy$

2. Slope, intercept, coefficients of determination and correlation.

a. slope $= b = \frac{\Sigma xy}{\Sigma x^2}$

b. intercept $= a = \bar{y} - b\bar{x}$

c. coefficient of determination $= r^2 = \frac{(\Sigma xy)^2}{(\Sigma x^2)(\Sigma y^2)}$

d. correlation coefficient $= r = \sqrt{r^2}$

3. Prediction of dependent variable from a given independent variable.

$\hat{Y} = a + bX$

4. Prediction of independent variable from a given dependent variable.

$X = \frac{\hat{Y} - a}{b}$

5. Reduction in sum of squares due to regression, residual sum of squares and degrees of freedom, variance about regression, standard error of the estimate, standard deviation about the slope, and t-test on slope.

a. reduction in sum of squares due to regression $= \frac{(\Sigma xy)^2}{\Sigma x^2}$

b. residual sum of squares (Res. S.S.) and degrees of freedom (d.f.):

$$\text{Res. SS} = \Sigma y^2 - \frac{(\Sigma xy)^2}{\Sigma x^2} \quad \text{and d.f.} = n - 2$$

c. variance about regression $= s^2_{y \cdot x} = \frac{\text{Res. SS}}{\text{d.f.}} = \frac{\Sigma y^2 - \frac{(\Sigma xy)^2}{\Sigma x^2}}{n - 2}$

d. standard error of the estimate $= s_{y \cdot x} = \sqrt{s^2_{y \cdot x}}$

e. standard deviation about the slope $= s_b = \frac{s_{y \cdot x}}{\sqrt{\Sigma x^2}}$

f. t-test on the slope $t = \frac{b}{s_b}$

6. Confidence limits about the slope =

$CL(b) = b \pm t_\alpha \cdot s_b$, $\ell_1 = b - t_\alpha \cdot s_b$ and $\ell_2 = b + t_\alpha \cdot s_b$

where t_α = Student's t at α-level of probability and n-2 df.

7. Confidence limits on $\hat{Y}$ with a given value of X =

$$CL(\hat{Y}) = a + bX \pm t_\alpha \cdot s_{y \cdot x} \sqrt{\frac{1}{n} + \frac{(X-\bar{x})^2}{\Sigma x^2}}$$

where t_α = Student's t at n-2 df at α - level of probability.

8. Confidence limits on a future $\hat{Y}$ with a given value of X =

$$CL(\text{future } \hat{Y}) = a + bX \pm t_\alpha \cdot s_{y \cdot x} \sqrt{1 + \frac{1}{n} + \frac{(X-\bar{x})^2}{\Sigma x^2}}$$

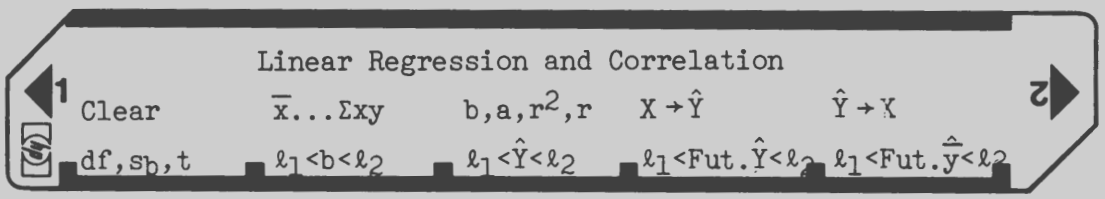
where t_α is the same as in 7.

9. Confidence limits on a future $\hat{\bar{y}}$ of sample size n_1 with a given value of X =

$$CL(\text{future } \hat{\bar{y}}) = a + bX \pm t_\alpha \cdot s_{y \cdot x} \sqrt{\frac{1}{n_1} + \frac{1}{n} + \frac{(X-\bar{x})^2}{\Sigma x^2}}$$

where t_α is the same as in 7.

User Instructions



STEP	INSTRUCTIONS	INPUT DATA/UNITS	KEYS	OUTPUT DATA/UNITS
1.	Set switch to "Norm"			
2.	Clear Registers		A	0.00
3.	Do 4-5 for each pair of data from i = 1,2,...n			
4.	Enter i<u>th</u> dependent variable	Y_i	↑	Y_i
5.	Enter i<u>th</u> independent variable	X_i	Σ+	i
6.	In case of error do 7-8 for i<u>th</u> incorrect entry			
7.	Enter k<u>th</u> dependent variable	Y_k	↑	Y_k
8.	Enter k<u>th</u> independent variable	X_k	Σ-	i-1
9.	Calculate means, sum of squares and sum of cross products		B	$\bar{x}$ Σx^2 $\bar{y}$ Σy^2 Σxy
10.	Calculate slope (b), intercept (a), coefficient of determination (r^2), and correlation coefficient (r).		C	b a r^2 r

(continued)

11.	$\hat{Y}$ from a given value of X	X	D		$\hat{Y}$
12.	X from a given value of $\hat{Y}$	$\hat{Y}$	E		X
13.	Calculate reduction in sum of squares due to		f	a	$(\Sigma xy)^2/\Sigma x^2$
	regression, $(\Sigma xy)^2/\Sigma x^2$; residual sum of squares,				Res. SS
	Res. SS; degrees of freedom, d.f.; variance				df
	about regression, $s^2_{y.x}$; standard error of the				$s^2_{y.x}$
	estimate, $s_{y.x}$; standard deviation about the				$s_{y.x}$
	slope, s_b; and t-test on the slope, $t=b/s_b$.				s_b
					t
14.	Calculate confidence limits about the slope	t_α	f	b	ℓ_1
					ℓ_2
15.	Calculate confidence limits about $\hat{Y}$ for a	X	f	c	$\hat{Y}$
	given value of X				$\pm$ value
					ℓ_1
					ℓ_2
16.	Calculate confidence limits about a future $\hat{Y}$	X	f	d	$\hat{Y}$
	for a given value of X				$\pm$ value
					ℓ_1
					ℓ_2
17.	Calculate confidence limits on future $\hat{\bar{y}}$ made	X	↑		X
	up of n_1 observations at a given value of X	n_1	f	e	$\hat{\bar{y}}$
					$\pm$ value
					ℓ_1
					ℓ_2

Example

Pair	Y	X
1	13.5	13.8
2	12.7	13.3
3	12.7	13.7
4	12.4	12.8
5	12.3	12.2
6	13.0	13.4
7	13.8	14.2
8	12.8	13.0

	Operation		Output
1.	Set switch to "Norm"		
2.	A	———	Clears Registers GSBA
3.	13.5 ↑	———	13.50 Ent ↑
4.	13.8 $\Sigma+$	———	13.80 $\Sigma+$
5.	12.7 ↑	———	12.70 Ent ↑
6.	13.3 $\Sigma+$	———	13.30 $\Sigma+$
7.	12.7 ↑	———	12.70 Ent ↑
8.	13.7 $\Sigma+$	———	13.70 $\Sigma+$

9.	12.4	↑			12.40	Ent ↑
10.	12.8	Σ+			12.80	Σ+
11.	12.3	↑			12.30	Ent ↑
12.	12.2	Σ+			12.20	Σ+
13.	13.0	↑			13.00	Ent ↑
14.	13.4	Σ+			13.40	Σ+
15.	13.8	↑			13.80	Ent ↑
16.	14.2	Σ+			14.20	Σ+
17.	12.8	↑			12.80	Ent ↑
18.	13.0	Σ+			13.00	Σ+
19.	B					GSBB
				$\bar{x}$	13.30	***
				Σx^2	2.78	***
				$\bar{y}$	12.90	***
				Σy^2	1.88	***
				Σxy	1.98	***
20.	C					GSBC
				b	0.71	***
				a	3.43	***
				r^2	0.75	***
				r	0.87	***
21.	13.0	D		X	13.00	GSBD
				$\hat{Y}$	12.69	***
22.	13.0	E		$\hat{Y}$	13.00	GSBE
				X	13.44	***
23.	f	a				GSBa
				$(\Sigma xy)^2/\Sigma x^2$	1.41	***
				Res. SS	0.47	***
				d.f.	6.00	***
				$s^2_{y\cdot x}$	0.08	***
				$s_{y\cdot x}$	0.28	***
				s_b	0.17	***
				t	4.24	***
24.	2.447	f	b		2.447	GSBb
				ℓ_1	.030	***
				ℓ_2	1.12	***

25.	13.0	f	c	X	13.00	GSBc
				$\hat{Y}$	12.69	***
				± value	0.27	***
				C.L.	12.41	***
					12.96	***
26.	13.0	f	d	X	13.00	GSBd
				$\hat{Y}$	12.69	***
				± value	0.74	***
				C.L.	11.95	***
					13.42	***
27.	13.0	↑		X	13.00	Ent ↑
28.	10.0	f	e	n_1	10.00	GSBe
				$\hat{\bar{y}}$	12.69	***
				± value	0.35	***
				C.L.	12.34	***
					13.03	***

Note: $t = 4.24$ which exceeds $t_{.05} = 2.447$ at 6 d.f. Also, C.L. about the slope (0.30, 1.12) do not include zero. Therefore, we reject $H_o : B = 0$ and accept $H_1 : B \neq 0$ and conclude that the independent variable does influence the dependent variable.

Example

	GSBA	12.80	ENT↑	13.00	GSBE	0.27	***
13.50	ENT↑	13.00	Σ+	13.44	***	12.41	***
13.80	Σ+		GSBB		GSBa	12.96	***
12.70	ENT↑	13.30	***	1.41	***	13.00	GSBd
13.30	Σ+	2.78	***	0.47	***	12.69	***
12.70	ENT↑	12.90	***	6.00	***	0.74	***
13.70	Σ+	1.88	***	0.08	***	11.95	***
12.40	ENT↑	1.98	***	0.28	***	13.42	***
12.80	Σ+		GSBC	0.17	***	13.00	ENT↑
12.30	ENT↑	0.71	***	4.24	***	10.00	GSBe
12.20	Σ+	3.43	***	2.447	GSBb	12.69	***
13.00	ENT↑	0.75	***	0.30	***	0.35	***
13.40	Σ+	0.87	***	1.12	***	12.34	***
13.80	ENT↑	13.00	GSBD	13.00	GSBc	13.03	***
14.20	Σ+	12.69	***	12.69	***		

Program

001	*LBL A	21 11
002	CLRG	16-53
003	P≠S	16-51
004	CLRG	16-53
005	CLX	-51
006	RTN	24
007	*LBL B	21 12
008	P≠S	16-51
009	RCL5	36 05
010	RCL4	36 04
011	RCL9	36 09
012	÷	-24
013	STO1	35 01
014	PRTX	-14
015	X²	53
016	RCL9	36 09
017	×	-35
018	-	-45
019	STO5	35 05
020	PRTX	-14
021	RCL7	36 07
022	RCL6	36 06
023	RCL9	36 09
024	÷	-24
025	STO2	35 02
026	PRTX	-14
027	X²	53
028	RCL9	36 09
029	×	-35
030	-	-45
031	STO7	35 07
032	PRTX	-14
033	RCL8	36 08
034	RCL6	36 06
035	RCL4	36 04
036	RCL9	36 09
037	÷	-24
038	×	-35
039	-	-45
040	STO8	35 08
041	PRTX	-14
042	RTN	24
043	*LBL C	21 13
044	RCL2	36 02
045	RCL8	36 08
046	RCL5	36 05
047	÷	-24
048	PRTX	-14
049	STO0	35 00
050	RCL1	36 01
051	×	-35
052	-	-45
053	PRTX	-14
054	STO3	35 03
055	RCL8	36 08
056	X²	53
057	RCL5	36 05
058	÷	-24
059	RCL7	36 07
060	÷	-24
061	PRTX	-14
062	∫X	54
063	PRTX	-14
064	STOI	35 46
065	RTN	24
066	*LBL D	21 14
067	RCL0	36 00
068	×	-35
069	RCL3	36 03
070	+	-55
071	PRTX	-14
072	RTN	24
073	*LBL E	21 15
074	RCL3	36 03
075	-	-45
076	RCL0	36 00
077	÷	-24
078	PRTX	-14
079	RTN	24
080	*LBL a	21 16 11
081	RCL7	36 07
082	RCL8	36 08
083	X²	53
084	RCL5	36 05
085	÷	-24
086	PRTX	-14
087	STOA	35 11
088	-	-45
089	PRTX	-14
090	STOB	35 12
091	RCL9	36 09
092	2	02
093	-	-45
094	PRTX	-14
095	STOC	35 13
096	÷	-24
097	PRTX	-14
098	STOD	35 14
099	∫X	54
100	PRTX	-14
101	STOE	35 15
102	RCL5	36 05
103	∫X	54
104	÷	-24
105	PRTX	-14
106	STO6	35 06
107	RCL0	36 00
108	X≠Y	-41
109	÷	-24
110	PRTX	-14
111	RTN	24
112	*LBL b	21 16 12
113	STO4	35 04
114	RCL6	36 06
115	×	-35
116	RCL0	36 00
117	X≠Y	-41
118	-	-45
119	PRTX	-14
120	LSTX	16-63
121	+	-55
122	LSTX	16-63
123	+	-55
124	PRTX	-14
125	RTN	24
126	*LBL c	21 16 13
127	STOD	35 14
128	RCL1	36 01
129	-	-45
130	X²	53
131	RCL5	36 05
132	÷	-24
133	RCL9	36 09
134	1/X	52
135	+	-55
136	∫X	54
137	RCLE	36 15
138	×	-35
139	RCL4	36 04
140	×	-35
141	RCLD	36 14
142	GSBD	23 14
143	X≠Y	-41
144	PRTX	-14
145	-	-45
146	PRTX	-14
147	LSTX	16-63
148	+	-55
149	LSTX	16-63
150	+	-55
151	PRTX	-14
152	RTN	24
153	*LBL d	21 16 14
154	STOD	35 14
155	RCL1	36 01
156	-	-45
157	X²	53
158	RCL5	36 05
159	÷	-24
160	RCL9	36 09
161	1/X	52
162	+	-55
163	1	01
164	+	-55
165	∫X	54
166	RCLE	36 15
167	×	-35
168	RCL4	36 04
169	×	-35
170	RCLD	36 14
171	GSBD	23 14
172	X≠Y	-41
173	PRTX	-14
174	-	-45
175	PRTX	-14
176	LSTX	16-63
177	+	-55
178	LSTX	16-63
179	+	-55
180	PRTX	-14
181	RTN	24
182	*LBL e	21 16 15
183	P≠S	16-51
184	STO0	35 00
185	P≠S	16-51
186	R↓	-31
187	STOD	35 14
188	RCL1	36 01
189	-	-45
190	X²	53
191	RCL5	36 05
192	÷	-24
193	RCL9	36 09
194	1/X	52
195	+	-55
196	P≠S	16-51
197	RCL0	36 00
198	1/X	52
199	P≠S	16-51
200	+	-55
201	∫X	54
202	RCLE	36 15
203	×	-35
204	RCL4	36 04
205	×	-35
206	RCLD	36 14
207	GSBD	23 14
208	X≠Y	-41
209	PRTX	-14
210	-	-45
211	PRTX	-14
212	LSTX	16-63
213	+	-55
214	LSTX	16-63
215	+	-55
216	PRTX	-14
217	RTN	24
218	R/S	51

Published in accordance with an Act passed in 1881 by the 14th Legislative Assembly, Dakota Territory, establishing the Dakota Agricultural College and with the Act of re-organization passed in 1887 by the 17th Legislative Assembly, which established the Agricultural Experiment Station at South Dakota State University.

300 printed at an estimated cost of 43 cents each—11-78gly—3033A