COMPUTING PRESSURE REDUCTION CONSTANTS USING INDICATOR VARIABLES

Bernard Augustine (1)
Suitland House
4500 Silver Hill Road
Suitland, MD 20746

ABSTRACT

Since computers are an integral part of a weather station, more standard procedures can be automated. A procedure is now available for calculating sea-level pressure for manned and unmanned weather stations. Each weather station has a unique equation that is used to compute reduction constants. Sea-level pressure is calculated by multiplying station pressure by the reduction constant (2). This paper discusses an alternative method for computing reduction constants.

1. INTRODUCTION

There are two reasons for writing this paper:

(a) To discuss the indicator variable technique for developing models;

(b) To use this technique to develop a single regression model that can compute a reduction constant for more than one weather station.

2. METHODOLOGY

The independent variables used in regression models are quantitative; i.e., they are actual numerical values. However, many variables cannot be defined numerically (e.g., marital status, sex, race, color, shape, religion, etc.). These are <u>qualitative variables</u> and serve to identify class membership. It is necessary to examine a method to quantify the levels of a qualitative variable for regression analysis. The qualitative variables are commonly referred to as indicator variables and take the values 0 and 1. If a qualitative variable has n levels, it is represented by n-1 indicator variables. For example, if three weather stations are grouped geographically, there will be two indicator variables each being assigned the value 0 or 1. If this rule is violated, a computational difficulty - multicollinearity - occurs. It is beyond the scope of this paper to discuss multicollinearity, but its importance cannot be over emphasized. Serious errors will produce unwanted results.

3. EXAMPLE

The following three weather stations are grouped geographically:

(a) Big Piney, Wyoming (BPI)

(b) Cody, Wyoming (COD)

(c) Wendover, Utah (ENV)

The indicator variables x_1 and x_2 are defined as:

$$x_1 = \begin{cases} 1 \text{ if the station is BPI} \\ 0 \text{ otherwise} \end{cases}$$

$$x_2 = \begin{cases} 1 \text{ if the station is COD} \\ 0 \text{ otherwise} \end{cases}$$

To illustrate the indicator variable process, consider the data in Table 1 where:

1) T is the independent variable (12 hour mean temperature)

2) x_1 and x_2 are the indicator variables

3) y^1 is the dependent variable (reduction constant)

TABLE 1

T	x_1	x_2	y
-50	1	0	0.3294
-40	1	0	0.3228
-30	1	0	0.3168
-20	1	0	0.3113
-10	1	0	0.3064
0	1	0	0.3020
10	1	0	0.2977
20	1	0	0.2936
30	1	0	0.2895
40	1	0	0.2848
50	1	0	0.2810
60	1	0	0.2779
70	1	0	0.2751
80	1	0	0.2727
90	1	0	0.2707
-50	0	1	0.2365
-40	0	1	0.2324
-30	0	1	0.2284
-20	0	1	0.2250
-10	0	1	0.2217
0	0	1	0.2186
10	0	1	0.2157
20	0	1	0.2128
30	0	1	0.2100
40	0	1	0.2070
50	0	1	0.2041
60	0	1	0.2016
70	0	1	0.1995
80	0	1	0.1977
90	0	1	0.1959
100	0	1	0.1950
-30	0	0	0.1860

-20	0	0	0.1826
-10	0	0	0.1797
0	0	0	0.1773
10	0	0	0.1751
20	0	0	0.1727
30	0	0	0.1704
40	0	0	0.1685
50	0	0	0.1670
60	0	0	0.1656
70	0	0	0.1642
80	0	0	0.1628
90	0	0	0.1620
100	0	0	0.1613

It is desired to find the best set of independent variables to predict the reduction constants. Therefore, an interaction effect must be included in the model. The interaction effect involves a cross-product term between the independent variables and the indicator variables. In this case, the cross-product terms are Tx_1, Tx_2, T^2x_1, and T^2x_2. If the cross-product terms are not significant, they are not included in the model. The regression analysis is done by using the statistical computer package SAS (Statistical Analysis System).

The results of fitting the model are:

$$y = b_0 - b_1T + b_2T^2 + b_3x_1 + b_4x_2 - b_5Tx_1 - b_6Tx_2 - b_7T^2x_2 \qquad (1.0)$$

where:

$b_0 = 0.17721023$ $b_1 = 0.00027178$ $b_2 = 1.1887078 \times 10^{-6}$

$b_3 = 0.12486728$ $b_4 = 0.04124339$ $b_5 = 0.00019448$

$b_6 = 4.7676655 \times 10^{-5}$ $b_7 = 4.0631989 \times 10^{-7}$

To solve reduction constants for station BPI, use:

$$y = b_0 - b_1T + b_2T^2 + b_3x_1 - b_5Tx_1 \quad \text{because } x_2 = 0 \qquad (1.1)$$

To solve reduction constants for station COD, use:

$$y = b_0 - b_1T + b_2T^2 + b_4x_2 - b_6Tx_2 - b_7T^2x_2 \quad \text{because } x_1 = 0 \qquad (1.2)$$

To solve reduction constants for station ENV, use:

$$y = b_0 - b_1T + b_2T^2 \quad \text{because both } x_1 \text{ and } x_2 = 0 \qquad (1.3)$$

Table 2 shows the results of five other groups of weather stations. Included are the station name, call letters, significant variables and their regression coefficients.

The constant b_0 is the intercept.

TABLE 2

INDICATOR VARIABLES	STATION NAME	CALL LETTERS
x_1	Clayton, New Mexico	CAO
x_2	Deming, New Mexico	DMN
x_3	Douglas, Arizona	DUG
x_4	Winslow, Arizona	INW
x_5	Junction, Texas	JCT
x_6	Marfa, Texas	MRF
x_7	Prescott, Arizona	PRC
x_8	Truth or Consequences	TCS
x_9	Caliente, Arizona	P38
	Page, Arizona	PGA

Parameter	Regression Coefficients
Intercept	0.18606296
T	-0.00044666
T^2	1.8289148E-06
x_1	0.02453614
x_2	-0.00635929
x_3	-0.01649633
x_4	0.02090837
x_5	-0.11830744
x_6	0.01594893
x_7	0.02810267
x_8	0.01837490
x_9	-0.00278536
Tx_1	0.00011796
Tx_2	0.00017853
Tx_3	0.00020567
Tx_4	0.00013467
Tx_5	0.00035055
Tx_6	0.00012151
Tx_7	0.00014068
Tx_8	0.00013679
Tx_9	0.00016779
T^2x_1	-9.4330045E-07
T^2x_2	-1.1593344E-06
T^2x_3	-1.3199238E-06
T^2x_4	-1.0491946E-06
T^2x_5	-1.5973964E-06
T^2x_6	-7.8462579E-07

T^2x_7	-1.2479958E-06
T^2x_8	-1.0407030E-06
T^2x_9	-9.6294037E-07

INDICATOR VARIABLES	STATION NAME	CALL LETTERS
x_1	Greenville, Maine	6B2
x_2	Johnbury, Vermont	9B2
x_3	Worcester, Massachusetts	ORH
	Rome, Georgia	RMG

Parameter	Regression Coefficient
Intercept	0.02505938
T	-3.8256313E-05
T^2	1.2106106E-07
x_1	0.01604374
x_2	0.01604374
x_3	0.01548032
Tx_1	-1.7049487E-05
Tx_2	-2.2802562E-06
Tx_3	-1.7665513E-05

INDICATOR VARIABLES	STATION NAME	CALL LETTERS
x_1	Blue Canyon, California	BLU
x_2	Meacham, Oregon	MEH
x_3	Mount Shasta, California	MHS
x_4	Sandberg, California	SDB
	Sexton Summit, Oregon	SXT

Parameter	Regression Coefficient
Intercept	0.16275247
T	-0.00030873
T^2	0.00000073
x_1	0.06578299
x_2	0.00936228
x_3	-0.01417618
x_4	0.02709380
Tx_1	-0.00009081
Tx_3	0.00008338
T^2x_1	-0.00000030
T^2x_4	0.00000045

INDICATOR VARIABLES	STATION NAME	CALL LETTERS
x_1	Aniak, Alaska	ANI
x_2	Big Delta, Alaska	BIG
x_3	Cape Decision, Alaska	CDE
x_4	Nenana, Alaska	ENN
x_5	Fort Yukon, Alaska	FYU
x_6	Gulkana, Alaska	GKN
x_7	Illiama, Alaska	ILI
x_8	Middleton Island, Alaska	MDO
	Port Heiden, Alaska	PTH

Parameter	Regression Coefficient
Intercept	0.00375817
T	-5.0274260E-06
T^2	1.6202833E-08
x_1	-0.00033310
x_2	0.04926333
x_3	-0.00178684
x_4	0.01102643
x_5	0.01406149
x_6	0.06186686
x_7	0.00262816
x_8	-0.00116935
Tx_2	-0.00010611
Tx_3	1.4689853E-06
Tx_4	-1.5963286E-05
Tx_5	-2.0043523E-05
Tx_6	-0.00012322
Tx_7	-3.6784392E-06
Tx_8	1.0552560E-06
T^2x_2	2.8712619E-07
T^2x_4	4.8038654E-08

T^2x_5	6.2099541E-08
T^2x_6	2.6046168E-07

INDICATOR VARIABLES	STATION NAME	CALL LETTERS
x_1	Chamberlain, South Dakota	CHB
x_2	Warroad, Minnesota	D45
x_3	Marseilles, Illinois	MMO
x_4	Poplar Bluff, Missouri	PO2
x_5	Devils Lake, North Dakota	P11
x_6	Roseglen, North Dakota	P24
x_7	Medicine Lodge, Kansas	P28
x_8	Pequot Lake, Minnesota	P39
x_9	Redig, South Dakota	REJ
x_{10}	Sidney, Nebraska	SNY
x_{11}	Valentine, Nebraska	VTN
	Elkhart, Kansas	1K5

FOOTNOTES AND REFERENCES

1. Bernard Augustine is a graduate of Marshall University, Huntington, West Virginia. He has been with the Communications Division since 1967. Mr. Augustine has a wide variety of experience with the National Weather Service. Before coming to the Washington, D.C., area, he worked in weather stations at Barter Island and Nome, Alaska; Beckley and Huntington, West Virginia; and Norfolk, Virginia.

2. Augustine, Bernard G, 1981: "Computer Generated Sea-Level Pressure," National Weather Digest, 6:2, p.34.

Parameter	Regression Coefficient
Intercept	0.12261314
T	-0.00016899
T^2	3.1476108E-07
x_1	-0.05258632
x_2	-0.07984645
x_3	-0.09382217
x_4	-0.10797337
x_5	-0.06307131
x_6	-0.03947410
x_7	-0.06161888
x_8	-0.07129986
x_9	0.00256523
x_{10}	0.05959343
x_{11}	-0.01600669
Tx_1	6.7954851E-05
Tx_2	9.9562546E-05
Tx_3	0.00011157
Tx_4	0.00012494
Tx_5	8.0444190E-05
Tx_6	5.2752462E-05
Tx_7	7.3588753E-05
Tx_8	8.9498260E-05
Tx_{10}	-8.5144597E-05
Tx_{11}	1.9748260E-05