Logistic Regression Analysis

C. Mitchell Dayton
University of Maryland, College Park

LOGISTIC REGRESSION ANALYSIS

C. Mitchell Dayton

Department of Measurement, Statistics & Evaluation

Room 1230D Benjamin Building

University of Maryland

September 1992

1. Introduction and Model

Logistic regression analysis (LRA) extends the techniques of multiple regression analysis to research situations in which the outcome variable is categorical. In practice, situations involving categorical outcomes are quite common. In the setting of evaluating an educational program, for example, predictions may be made for the dichotomous outcome of success/failure or improved/not-improved. Similarly, in a medical setting, an outcome might be presence/absence of disease. The focus of this document is on situations in which the outcome variable is dichotomous, although extension of the techniques of LRA to outcomes with three or more categories (e.g., improved, same, or worse) is possible (see, for example, Hosmer & Lemeshow, 1989, Chapter 8). In this section, we review the multiple regression model and, then, present the model for LRA.

The fundamental model underlying multiple regression analysis (MRA) posits that a continuous outcome variable is, in theory, a linear combination of a set of predictors and error. Thus, for an outcome variable, Y, and a set of p predictor variables, $X_1,\ldots,X_p$, the MRA model is of the form:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p + \varepsilon = \alpha + \sum_{j=1}^{p} \beta_j X_j + \varepsilon$$

where is the Y-intercept (i.e., the expected value of Y when all X's are set to 0), β_j is a multiple (partial) regression coefficient (i.e., the expected change in Y per unit change in X_j assuming all other X's are held constant) and ε is the error of prediction. If error is omitted, the resulting model represents the expected, or predicted, value of Y:

$$E(Y|X_1,\ldots,X_p) = Y' = \alpha + \sum_{j=1}^{p} \beta_j X_j$$

Note that $Y=Y' + \varepsilon$. Thus, we can interpret the MRA model as follows: each observed score, Y, is made up of an expected, or predictable, component, Y', that is a function of the predictor variables $X_1,\ldots,X_p$, and an error, or unpredictable component, ε , that represents error of measurement (i.e., unreliability) and/or error in the selection of the model (i.e., mis-specification).

The MRA model summarized above is applicable when the outcome variable, Y, is continuous, but is not appropriate for situations in which Y is categorical. For example, if Y takes on the value 1 for "success" and 0 for "failure," the multiple regression equation would not result in predicted values restricted to exactly 1 or 0. In fact, these predicted values would be spread out over an interval that contains uninterpretable values such as .5 or .33 and could even include negative values and/or values greater than 1. The model for logistic regression analysis, described below, is a more realistic representation of the situation when an outcome variable is categorical.

The model for logistic regression analysis assumes that the outcome variable, Y, is categorical (e.g., dichotomous), but LRA does not model this outcome variable directly. Rather, LRA is based on probabilities associated with the values of Y. For simplicity, and because it is the case most commonly encountered in practice, we assume that Y is dichotomous, taking on values of 1 (i.e., the positive outcome, or success) and 0 (i.e., the negative outcome, or failure). In theory, the hypothetical, population proportion of cases for which Y = 1 is defined as $\pi = P(Y=1)$. Then, the theoretical proportion of cases for which Y = 0 is $1 - \pi = P(Y = 0)$. In the absence of other information, we would estimate π by the sample proportion of cases for which Y = 1. However, in the regression context, it is assumed that there is a set of predictor variables, $X_1,\ldots,X_p$, that are related to Y and, therefore, provide additional information for predicting Y. For theoretical, mathematical reasons, LRA is based on a linear model for the natural logarithm of the odds (i.e., the log-odds) in favor of Y = 1 (see Appendix A for a further explanation of odds):

$$Log_e\left[\frac{P(Y=1|X_1,\ldots,X_p)}{1-P(Y=1|X_1,\ldots,X_p)}\right] = Log_e\left[\frac{\pi}{1-\pi}\right] =$$
$$= \alpha + \beta_1 X_1 + \ldots + \beta_p X_p = \alpha + \sum_{j=1}^{p} \beta_j X_j$$

Note that in the LRA model, π is a conditional probability of the form $P(Y=1| X_1,...,X_p)$. That is, it is assumed that "success" is more or less likely depending on combinations of values of the predictor variables. The log-odds, as defined above is also known as the logit transformation of π and the analytical approach described here is sometimes known as logit analysis.

The LRA model above is identical to the MRA model except that the log-odds in favor of Y = 1 replaces the expected value of Y. There are two basic reasons underlying the development of the model above. First, probabilities and odds obey multiplicative, rather than additive, rules. However, taking the logarithm of the odds allows for the simpler, additive model since logarithms convert multiplication into addition. And, second, there is a (relatively) simple exponential transformation for converting log-odds back to probability. In particular, the inverse transformation is the logistic function of the form:

$$P(Y=1|X_1,\ldots,X_p) = \frac{e^{\alpha+\sum_{j=1}^{p}\beta_j X_j}}{1+e^{\alpha+\sum_{j=1}^{p}\beta_j X_j}}$$

Due to the mathematical relationship, $e^a/(1+e^a) = 1/(1+e^{-a})$, the logistic function for LRA is sometimes presented in the form:

$$P(Y=1|X_1,\ldots,X_p) = \frac{1}{1+e^{-\alpha-\sum_{j=1}^{p}\beta_j X_j}}$$

Due to the mathematical relation, $1- e^a/(1+e^a) = 1/(1+e^a)$, the probability for a 0 response is:

$$P(Y=0|X_1,\ldots,X_p) = 1 - P(Y=1|X_1,\ldots,X_p) = \frac{1}{1+e^{\alpha+\sum_{j=1}^{p}\beta_j X_j}}$$

2. Fitting the LRA Model to Data

As in MRA, there are two important stages in the analysis of data. First, estimates for the parameters in the model must be obtained and, second, some determination must be made of how well the model actually fits the observed data. In MRA, the parameter estimates are obtained using the least-squares principle and assessment of fit is based on significance tests for the regression coefficients as well as on interpreting the multiple correlation coefficient. The process is analogous for LRA, although specific details of these stages of analysis are somewhat different as outlined in this section.

The parameters that must be estimated from the available data are the constant, α , and the logistic regression coefficients, β_j. Because of the nature of the model, estimation is based on the maximum likelihood principle rather than on the least-squares principle. In the context of LRA, maximum likelihood estimation (MLE) involves the following. First, we define the likelihood, L, of the sample data as the **product**, across all sampled cases, of the probabilities for success or for failure:

$$L = \prod_{i=1}^{n} P(Y_i|X_{i1},\ldots,X_{ip}) = \prod_{i=1}^{n}\left[\left(\frac{e^{\alpha+\sum_{j=1}^{p}\beta_j X_j}}{1+e^{\alpha+\sum_{j=1}^{p}\beta_j X_j}}\right)^{Y_i} \times \left(\frac{1}{1+e^{\alpha+\sum_{j=1}^{p}\beta_j X_j}}\right)^{1-Y_i}\right]$$

Note that Y is the 0/1 outcome for the i^{th} case and , $X_{i1},...,X_{ip}$ are the values of the predictor variables for the i^{th} case based on a sample of n cases. The use of Y_i and $1-Y_i$ as exponents in the equation above includes in the likelihood the appropriate probability term dependent upon whether $Y_i =1$ or $Y_i =0$ (note that $F^0=1$ for any expression, F). Using the methods of calculus, a set of values for α and the β_j can be calculated that **maximize** L and these resulting values are known as maximum likelihood estimates (MLE's). This maximization process is somewhat more complicated than the corresponding minimization procedure in MRA for finding least-square estimates. However, the general approach involves establishing initial guesses for the unknown parameters and then continuously adjusting these estimates until the maximum value of L is found. This iterative solution procedure is available in popular statistical procedures such as those found in SPSS and SAS.

To distinguish them from parameters, we denote the MLE's as a and b_j. Given that these estimates have been calculated for a real data set, tests of significance for individual logistic regression coefficients can be set up as in MRA. That is, for the hypothesis, H: $\beta_j = 0$, a statistic of the form $z = b_j / S_j$ can be calculated based on the estimated standard error, S_j, for b_j (SPSS and SAS report $\chi_1^2 = z^2$ and label these values as Wald statistics).

Similarly, the usefulness of the model as a whole can be assessed by testing the hypothesis that, simultaneously, all of the partial logistic regression coefficients are 0; i.e., H: $\beta_j = 0$ for all j. In effect, we can compare the general model given above with the restricted model $\text{Log}_e(\pi/(1-\pi)) = \alpha$. This test, that is equivalent to testing the significance of the multiple R in MRA, is based on a chi-squared statistic (SPSS labels this value as "Model Chi-Square").

Finally, different LRA models fitted to the same set of data can be compared statistically in a simple manner if the models are hierarchical. The hierarchy principle requires that the model with the larger number of predictors include among its predictors all of the predictors from the simpler model (e.g., predictors X_1 and X_2 in the simpler model and predictors X_1, X_2, X_3 and X_4 in the more complex model). Given this condition, the difference in model chi-squared values is (approximately) distributed as chi-squared with degrees of freedom equal to the difference in degrees of freedom for the two models (e.g., for the above example, the degrees of freedom would be 2). In effect, this procedure tests a conditional null hypothesis that, for the example, would be: H: $\beta_3 = \beta_4 = 0 | \beta_1, \beta_2$. That is, the values of the logistic regression coefficients associated with X_1 and X_2 are unrestricted, but the logistic regression coefficients associated with X_3 and X_4 are assumed by hypothesis to be 0. If the models are specified in a series of "blocks" in SPSS, an "Improvement" chi-square value is computed for each successive model and this can be used to test whether or not the additional predictors result in significantly better fit of the model to the data.

3. Interpreting the Fitted LRA Model

We may interpret the results from a LRA at three different levels. First, each term in the equation represents contributions to estimated log-odds. Thus, for each one unit increase (decrease) in X_j, there is predicted to be an increase (decrease) of b_j units in the log-odds in favor of Y = 1. Also, if all predictors are set equal to 0, the predicted log-odds in favor of Y = 1 would be the constant term, a. Second, since most people do not find it natural to think in terms of log-odds, the LRA equation can be transformed to odds by exponentiation:

$$\frac{\hat{\pi}}{1-\hat{\pi}} = e^{a+b_1X_1+\ldots+b_pX_p} = e^{a} \times e^{b_1X_1} \times \ldots \times e^{b_pX_p}$$

With respect to odds, the influence of each predictor is multiplicative. Thus, for each one unit increase in X_j, the predicted odds is increased by a factor of $\exp(b_j)$ If X is decreased by one unit, the multiplicative factor is $\exp(-b_j)$. Note that $\exp(c) = e^c$. Similarly, if all predictors are set equal to 0, the predicted odds are exp(a). Finally, the results can be expressed in terms of probabilities by use of the logistic function. With estimates substituted the equation above becomes:

$$P(Y = 1 | X_1, \ldots, X_p) = \frac{e^{a+\sum_{j=1}^{p} b_jX_j}}{1+e^{a+\sum_{j=1}^{p} b_jX_j}}$$

Although this form of the model may seem simplest because it results in predicted probabilities for Y = 1, there is no simple interpretation for the logistic regression coefficients in this form. However, for the case of a single predictor, X, a plot can be generated with X represented by the abscissa (i.e., horizontal axis) and P(Y|X) represented by the ordinate (i.e., vertical axis). The resulting curve is S-shaped and is illustrated in one of the exemplary analyses presented below.

4. Illustrative Examples

The database used for purposes of illustration is provided with SPSS and is based on 50 premature infants, some of whom did not survive infancy. The SPSS setup for this database, including the raw data, is contained in Appendix B. The outcome variable is dichotomous with "1" representing survival of the infant and "0" representing death of the infant. For the first exemplary analysis, a single independent variable, weight at birth in kilograms, was utilized. For the second analysis, three additional predictors were entered as described below.

(1) Single Predictor Example

Using birth weight as the predictor, the logistic regression equation for log-odds in favor of survival is estimated to be (the SPSS LOGISTIC REGRESSION printout for this block is provided in Appendix C):

$$\text{Log}_e\left[\frac{\hat{\pi}}{1-\hat{\pi}}\right] = a + bX = -3.6005 + 1.7406 \times Weight$$

To estimate odds per se, the equation is exponentiated:

$$\frac{\hat{\pi}}{1-\hat{\pi}} = e^{a+bX} = e^{-3.6005+1.7406\times Weight}$$

Finally, the probability of survival is obtained by applying the logistic transformation:

$$\hat{\pi} = \frac{e^{a+bX}}{1+e^{a+bX}} = \frac{e^{-3.6005+1.7406\times Weight}}{1+e^{-3.6005+1.7406\times Weight}} = \frac{1}{1+e^{3.6005-1.7406\times Weight}}$$

In the database, birth weight, in kilograms, ranged from 1.03 to 3.64 (i.e., from 2.27 to 8.02 pounds). Thus, for the lowest weight recorded, 1.03, the log-odds in favor of survival are -1.8077, the odds in favor of survival are .1640 and probability of survival is .14. At the other extreme, for birth weight of 3.64 kilograms, the log-odds in favor of survival are 2.7353, the odds in favor of survival are 15.4121 and the probability of survival is .94. The relation between birth weight and survival of infants is displayed in Figure 1. This graph was constructed by systematically varying birth weight from 1 to 4 kilograms (shown on the abscissa) and calculating the estimated probability of survival (shown on the ordinate). Note from reading the graph that infants with a birth weight of slightly over 2 kilograms (actually, 2.07) have an estimated odds of 1:1 in favor of survival (i.e., for π is estimated to equal .5 for a birth weight of 2.07 kilograms).

The estimated logistic regression coefficient for birth weight is 1.7406 and the exponential of this value is $e^{1.7406} = 5.70$. This indicates that for a one kilogram increase in birth weight, the odds in favor of survival are estimated to be increased by a multiplicative factor of 5.7. The reported standard error for b is .5786 and statistical significance can be assessed by the Wald chi-squared statistic, $(1.7406/.5786)^2 = 9.05$, that, with 1 degree of freedom, is significant at conventional levels (the empirical two-tailed p-value is reported to be .0026 in Appendix C). Thus, this study supports the conclusion that birth weight is a useful predictor of infant survival.

To aid in interpretation, a classification table can be constructed by predicting the survival or death of each infant based on whether or not the odds for survival are greater or less than 1.0, and comparing these predictions to the actual outcome for each infant. As shown in Appendix C, the percents of correct decisions are 56.5 for infants who survived, 74.1 for infants who died, and 66.0 overall. This overall result can be compared with a rate of 54.0% that would be obtained by simply predicting death as the outcome for every infant (i.e., since 27 of the 50 infants died, this is the prediction with the greater likelihood of being correct).

(2) Multiple Predictor Example

Building on the previous example, the following three predictors, in addition to birth weight, were included in the model: gestation age, sex of infant, and the APGAR score (a measure of infant well being). Using this set of four predictors, the logistic regression equation for the log-odds in favor of survival is estimated to be (note: the SPSS LOGISTIC REGRESSION printout for this block is provided in Appendix D):

$$Log_e\left[\frac{\hat{\pi}}{1-\hat{\pi}}\right] = -3.2696 + 3.0009\times Weight - .2870\times Age - .3809\times Sex + .0741\times APGAR$$

To estimate odds, the equation is exponentiated:

$$\frac{\hat{\pi}}{1-\hat{\pi}} = e^{-3.2696+3.0009\times Weight-.2870\times Age-.3809\times Sex+.0741\times APGAR}$$

Finally, the probability of survival is obtained by applying the logistic transformation:

$$\hat{\pi} = \frac{e^{-3.2696+3.0009\times Weight-.2870\times Age-.3809\times Sex+.0741\times APGAR}}{1+e^{-3.2696+3.0009\times Weight-.2870\times Age-.3809\times Sex+.0741\times APGAR}}$$

Unlike the case of a single predictor, there is no simple graphical display that can be used. However, log-odds, odds and probabilities can be calculated for various combinations of the predictors. In the present case, there is more interest in deciding whether or not the three predictors, in addition to birth weight, have improved the predictive efficiency of the model. The Wald chi-squared statistics, as taken from Appendix D, are non-significant for age, sex and APGAR (i.e., p-values of .23, .60 and .58, respectively), whereas the chi-squared value for weight is significant at the .05 level (i.e., p-value of .02). Thus, given that the other predictors remain in the model, removing weight as a predictor would result in significantly poorer predictive efficiency, although removing any of the other predictors does not have a significant impact.

A second, global way of comparing the single and multiple predictor models is based on the improvement chi-squared value that is reported as 1.638 based on 3 degrees of freedom This value is non-significant at conventional levels and, thus, we conclude that the predictive efficiency of the 4-predictor model is no greater than that of the 1-predictor model. It should be noted that the classification table results are exactly the same using four predictors as for using only a single predictor (e.g., 66.0% correct classifications overall).

Suggested References

Fienberg, S. E. (1980). The Analysis of Cross-Classified Categorical Data (Second Edition). Cambridge, MA: The MIT Press

Fleiss, J. (1981). Statistical Methods for Rates and Proportions (Second Edition). New York: Wiley

Hosmer, D. W. & Lemeshow, S. (1989). Applied Logistic Regression, New York: Wiley.

Appendix A: The Meaning of Odds

Odds represent the relative frequency with which different outcomes occur. Odds are sometimes expressed as a ratio of the form a:b. For example, odds of 3:1 in favor of the first outcome means that the first outcome occurs 3 times for each single occurrence of the second outcome. Similarly, odds of 5:2 means that the first outcome occurs 5 times for each 2 occurrences of the second outcome. Odds are directly related to probabilities and can be translated back and forth using these relations:

Probability = a/(a+b) when odds are expressed as a:b; or

Probability = Odds/(1 + Odds) when odds are expressed in decimal form (e.g., 3:1 becomes

3.0 and 5:2 becomes 2.5)

Odds = Probability/(1 - Probability)

Some examples:

The probability of rolling a 1 with a true die is 1/6 = .1667. The odds in favor of rolling a one are 1/6÷ 5/6 = .20, or 1:5. Viewed the other way, the odds against rolling a 1 with a true die are 5:1, or 5.0.

The odds in favor of not drawing a face card from a standard deck of playing cards (i.e., drawing a card other than a Jack, Queen, King or Ace) are 36:16 or 9:4. The corresponding probability is 9/(9+4) = 9/13 = .6923.

Appendix B: Exemplary Database (SPSS Syntaz)

```
TITLE 'RDS: Respiratory Distress Syndrome'.
DATA LIST /CASEID 1-2 SURVIVAL 4 TREATMNT 6 TIME 8-10(1)
WEIGHT 12-15(3) APGAR 17-18 SEX 20 AGE 22-23 PH 33-35(2) RESP 37.
RECODE SURVIVAL (1=2) (0=1).
VARIABLE LABELS SURVIVAL 'Infant Survival'
TREATMNT 'Treatment Administered'
TIME 'Time to Spontaneous Respiration'
WEIGHT 'Birthweight in Kilograms'
APGAR 'APGAR Score'
SEX 'Sex of Infant'
PH 'PH Level'
RESP 'Respiratory Therapy'.
MISSING VALUE RESP(9).
VALUE LABELS TREATMNT 1 'Tham' 0 'Sodium Bicarbonate'
/SEX 1 'Male' 0 'Female' /RESP 1 'Yes' 0 'No' 9 'No Answer'
/SURVIVAL 2 'Survive' 1 'Die'.
BEGIN DATA.
1 0 1 020 1050 5 0 28 89 49 709 0
2 0 1 020 1175 4 0 28 30 62 711 1
3 0 1 005 1230 7 0 29 53 47 724 9
4 0 1 40 1310 4 1 29 59 66 713 1
5 0 1 005 1500 8 1 32 78 49 723 1
6 0 1 100 1600 2 0 35 87 47 722 0
7 0 1 005 1720 9 0 35 59 76 703 0
8 0 1 000 1750 8 1 32 87 66 714 1
9 0 1 060 1770 1 1 32 41 86 710 1
10 0 1 020 2275 5 1 33 90 63 713 1
11 0 1 000 2500 9 1 38 43 64 705 1
12 0 0 70 1030 3 1 29 100 41 729 1
13 0 0 005 1100 7 1 32 31 26 736 1
14 0 0 000 1185 8 1 30 54 40 715 1
15 0 0 080 1225 1 1 28 41 48 723 0
16 0 0 100 1262 1 1 28 44 73 718 0
```

17 0 0 000 1295 7 0 32 60 77 724 1
18 0 0 020 1300 4 1 32 59 80 714 1
19 0 0 030 1550 4 1 31 54 54 725 1
20 0 0 000 1820 10 0 32 86 24 734 0
21 0 0 005 1890 8 1 31 51 36 725 1
22 0 0 050 1940 4 0 33 86 39 727 0
23 0 0 100 2200 2 0 36 94 -1 707 1
24 0 0 000 2270 8 0 38 98 52 717 0
25 0 0 000 2440 8 1 35 35 44 720 0
26 0 0 020 2560 6 1 36 95 46 716 1
27 0 0 000 2730 7 1 37 46 52 714 1
28 1 1 005 1130 7 1 30 92 46 726 9
29 1 1 000 1575 3 1 31 52 37 731 0
30 1 1 150 1680 9 0 35 48 42 728 0
31 1 1 040 1760 4 1 32 42 80 702 0
32 1 1 010 1930 1 1 31 80 68 709 0
33 1 1 000 2015 9 1 34 80 40 719 0
34 1 1 000 2090 9 0 33 74 43 718 0
35 1 1 000 2600 7 0 37 83 46 716 0
36 1 1 000 2700 9 1 37 52 43 737 1
37 1 1 100 2950 4 1 34 20 86 702 1
38 1 1 030 3160 1 0 37 71 62 913 0
39 1 1 005 3400 6 1 39 73 72 724 0
40 1 1 000 3640 9 1 39 90 67 717 0
41 1 1 000 2830 9 1 37 16 59 712 1
42 1 0 010 1410 8 0 30 72 110 685 1
43 1 0 020 1715 3 0 31 44 37 724 0
44 1 0 020 1720 7 0 34 26 82 695 1
45 1 0 040 2040 5 1 34 42 45 722 0
46 1 0 050 2200 9 1 32 63 34 922 1
47 1 0 020 2400 5 1 36 33 45 709 0
48 1 0 010 2550 6 1 36 41 52 724 0
49 1 0 005 2570 8 1 35 87 35 730 0
50 1 0 000 3005 8 1 38 72 43 723 0
END DATA.

Appendix C: Exemplary SPSS Analysis for a Single Predictor

Total number of cases: 50 (Unweighted)
Number of selected cases: 50
Number of unselected cases: 0

Number of selected cases: 50
Number rejected because of missing data: 0
Number of cases included in the analysis: 50

Dependent Variable Encoding:

Original Value	Internal Value
.00	0
1.00	1

Dependent Variable.. SURVIVE
Beginning Block Number 0. Initial Log Likelihood Function
-2 Log Likelihood 68.994376
* Constant is included in the model.

Beginning Block Number 1. Method: Enter
Variable(s) Entered on Step Number
1.. WEIGHT

Estimation terminated at iteration number 3 because

Log Likelihood decreased by less than .01 percent.

-2 Log Likelihood	56.975
Goodness of Fit	47.514
Cox & Snell - R^2	.214
Nagelkerke - R^2	.286

Chi-Square df Significance

	Chi-Square	df	Significance
Model	12.019	1	.0005
Block	12.019	1	.0005
Step	12.019	1	.0005

Classification Table[a]

			Predicted		
			SURVIVE		Percentage Correct
	Observed		Die	Survive	
Step 1	SURVIVE	Die	20	7	74.1
		Survive	10	13	56.5
	Overall Percentage				66.0

a. The cut value is .500

---------------------- Variables in the Equation ----------------------

Variable	B	S.E.	Wald	df	Sig	R	Exp(B)
WEIGHT	1.7406	.5786	9.0495	1	.0026	.3196	5.7009
Constant	-3.6005	1.1806	9.3010	1	.0023		

Appendix D: Exemplary SPSS Analysis for Multiple Predictors

```
Beginning Block Number 2. Method: Enter
```

Variable(s) Entered on Step Number
1.. GESTAGE
APGAR
SEX

Estimation terminated at iteration number 4 because
Log Likelihood decreased by less than .01 percent.

-2 Log Likelihood	55.338
Goodness of Fit	46.829
Cox & Snell - R^2	.239
Nagelkerke - R^2	.319

Chi-Square df Significance

	Chi-Square	df	Significance
Model	13.657	4	.0085
Block	1.638	4	.6509
Step	1.638	3	.6509

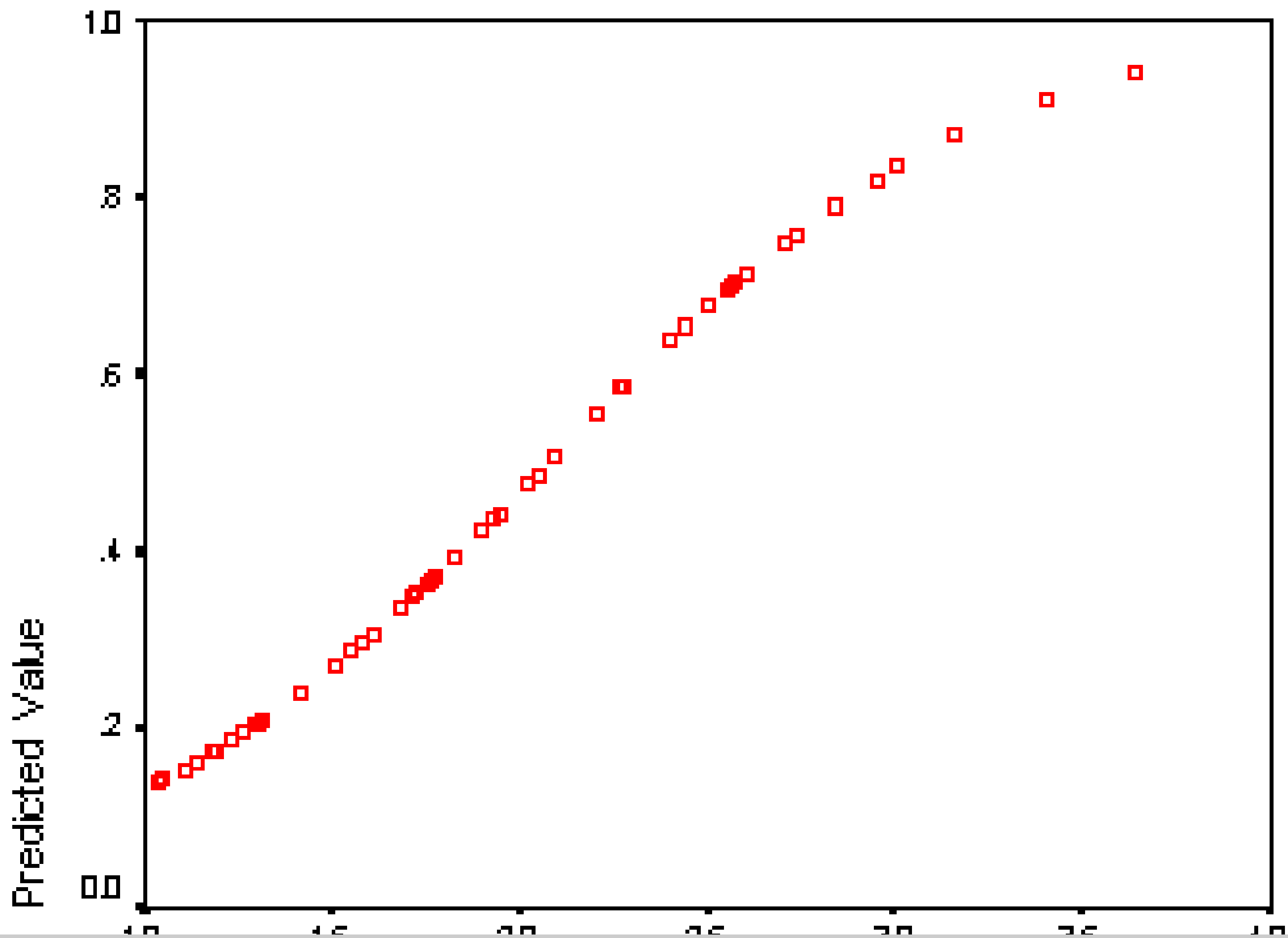
Classification Table[a]

Observed			Predicted SURVIVE Die	Predicted SURVIVE Survive	Percentage Correct
Step 1	SURVIVE	Die	20	7	74.1
		Survive	10	13	56.5
	Overall Percentage				66.0

a. The cut value is .500

--------------------- Variables in the Equation ----------------------

Variable	B	S.E.	Wald	df	Sig	R	Exp(B)
WEIGHT	3.0009	1.2716	5.5691	1	.0183	.2503	20.1027
GESTAGE	-.2879	.2379	1.4548	1	.2278	.0000	.7505
APGAR	.0741	.1333	.3096	1	.5779	.0000	1.0770
SEX	-.3809	.7350	.2685	1	.6043	.0000	.6833
Constant	3.2696	5.8122	.3165	1	.5737		



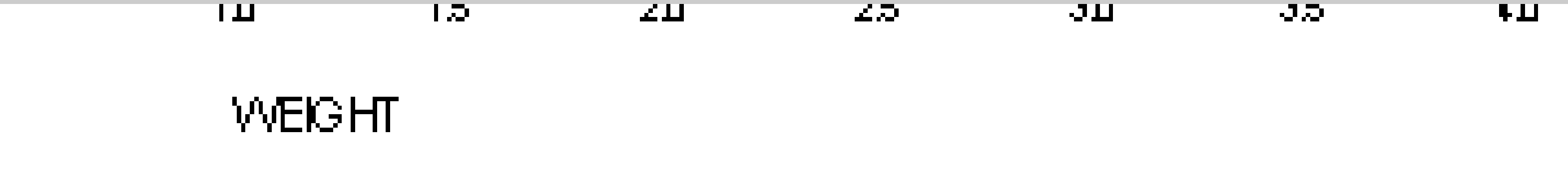


Figure 1