The Korean Communications in Statistics
Vol. 13 No. 1, 2006, pp. 49-60

Information Theoretic Standardized Logistic Regression Coefficients with Various Coefficients of Determination

Chong Sun Hong[1] and Hyeon Sang Ryu[2]

Abstract

There are six approaches to constructing standardized coefficient for logistic regression. The standardized coefficient based on Kruskal's information theory is known to be the best from a conceptual standpoint. In order to calculate this standardized coefficient, the coefficient of determination based on entropy loss is used among many kinds of coefficients of determination for logistic regression. In this paper, this standardized coefficient is obtained by using four kinds of coefficients of determination which have the most intuitively reasonable interpretation as a proportional reduction in error measure for logistic regression. These four kinds of the sixth standardized coefficient are compared with other kinds of standardized coefficients.

Keywords : Entropy loss; Information theory; Inherent prediction error; Proportional reduction.

1. Introduction

Whereas there is only one definition for the standardized coefficient in ordinary least squares regression analysis (OLS), there is no widely accepted definition for a standardized coefficient in the following logistic regression model: for $i=1,\cdots,n$,

$$logit(Y_i) = \ln[\Pr(Y_i=1)/\Pr(Y_i=0)] = \beta_0+\beta_1X_{i1}+\beta_2X_{i2}+\cdots+\beta_kX_{ik}\,. \qquad (1.1)$$

Mayer and Younger (1976) described that reasons for the use of standardized coefficients in logistic regression are essentially the same as those in OLS. Menard (2004) mentioned that the first reason for using standardized coefficients is that, for variables with no natural metric, a scale-free standardized coefficient may be more meaningful than an unstandardized one. Second reason is that when variables are measured in different units of measurement, standardized coefficients are useful for comparing the relative influence of different predictors within an OLS or logistic regression model (Agresti and Finlay, 1997). In other words, the magnitudes of the standardized coefficients are not affected by the scales of measurement of the various model variables and thus may be useful in ascertaining the relative importance of the effects of predictor variables not affected by the scales of measurement (Freund and Littell, 2000).

Six approaches to obtaining standardized logistic regression coefficients have been

1) Professor, Department of Statistics, SungKyunKwan University, Seoul, KOREA.
Correspondence : cshong@skku.ac.kr.
2) Research Fellow, Department of Statistics, SungKyunKwan University, Seoul, KOREA.

summarized by Menard (2004). The first approach to a standardized coefficient in logistic analysis was proposed by Goodman (1972), who divided each unstandardized coefficient by its estimated standard deviation. Since this is a test statistic (called the Wald statistic), rather than a standardized coefficient, Menard (2004) did not consider Goodman's version of a standardized coefficient. The second procedure, suggested by Agresti (1996) and mentioned also by Menard (1995), is to standardize only the predictors as the following:

$$b_A^* = (b)(s_X) , \tag{1.2}$$

where b is the sample estimate of the unstandardized logistic regression coefficient and s_X is the sample standard deviation of the predictor X. A third procedure, implemented in SAS (see, e.g., Menard, 1995), is to standardize both the predictors and the dependent variable, using $\pi/\sqrt{3}$, the standard deviation of the standard logistic distribution, as the estimate for the standard deviation of the dependent variable, Y:

$$b_S^* = (b)(s_X)/(\pi/\sqrt{3}) . \tag{1.3}$$

A fourth approach, suggested by Long (1997), is to standardize the predictors and the dependent variable based on the standard deviation of the standard logistic distribution plus the standard deviation of the standard normal distribution (the latter equal to one by definition):

$$b_L^* = (b)(s_X)/[(\pi/\sqrt{3})+1] . \tag{1.4}$$

These standardized coefficients in (1.2), (1.3) and (1.4) can be characterized as partially standardized coefficients because they incorporate the variance in the predictor. None of these approaches really takes into account the empirical variation in the dependent variable. To incorporate information about the actual variation in the dependent variable, the standard deviation of $logit(Y)$ can be estimated indirectly, using the formular borrowed from OLS: $b^* = (b)(s_X)/s_Y$. Since $R^2 = s_{\hat{Y}}^2/s_Y^2$ (Kvälseth, 1985), where $s_{\hat{Y}}$ is the standard deviation of the predicted value of Y, one obtains that $s_Y = s_{\hat{Y}}/R$ for OLS. With similar arguments, the standard deviation of $logit(Y)$, $s_{logit(Y)}$, may be estimated as $s_{logit(\hat{Y})}/R$ for logistic regression. In order to obtain the standardized coefficient one can multiply by $R/s_{logit(\hat{Y})}$ instead of dividing by $s_{logit(Y)}$. Menard (1995) suggested a fifth approach to standardization, defining the fully standardized logistic regression coefficient as

$$b_M^* = (b)(s_X)(R)/s_{logit(\hat{Y})} . \tag{1.5}$$

The coefficient calculated from Equation (1.5) may be described as a fully standardized logistic regression coefficient. The interpretation of b_M^* for logistic regression is the same as for the standardized coefficient in OLS: an one standard deviation increase in the predictor, X, is expected to increase by b_M^* standard deviations in the dependent variable, $logit(Y)$, holding all other predictors constant.

The second to fifth approaches to calculating standardized logistic regression coefficients include multiplication by the sample standard deviation of the predictor. These permit us to interpret the standardized coefficients in terms of the dependent variable when the predictor variables do not have a common metric. Just as with standardized coefficients in OLS regression, these four approaches may produce a different order than that suggested by the unstandardized coefficients for the relative strength of the predictors. However, all four approaches incorporating the standard deviation of the predictors produce the same order of relative influence of the predictors.

A sixth approach is derived from the information theory. An approach to measuring the relative importance of a variable, based on the amount of explained variance attributable to each predictor, was suggested by Kruskal (1987). Menard (2004) mentioned that relative influence refers to the information conveyed by the standardized coefficient, while relative importance refers to the information conveyed by Kruskal's approach.

A detail procedure for obtaining the sixth standardized coefficient and its extension to logistic regression are explained in Section 2. Menard (2004) used the coefficient of determination, R^2, to get the sixth standardized coefficient. We knew that there are twelve kinds of R^2s for logistic regression (see Mittlbock and Schemper, 1996; Menard, 2000). Based on eight criteria proposed by Kvälseth (1985), Menard (2000) and Mittlbock and Schemper (1996) preferred two R^2s (R_l^2 and R_o^2 in (2.2) and (2.4), respectively), which have the most intuitively reasonable interpretation as a proportional reduction in error measure. And Liao and McGee (2003) proposed two adjusted R^2s ($R_{l,adj}^2$ and $R_{o,adj}^2$ in (2.5) and (2.6), respectively), since both R_l^2 and R_o^2 could overestimate the measure of association. These four kinds of coefficient of determination are introduced in Section 2.

In this paper, these four kinds of the coefficient of determination are used to obtain the sixth standardized logistic regression coefficient, so their results are compared and their properties are discussed.

2. Information Theoretic Standardized Coefficient

When a predictor variable X_k is added to the model equation in (1.1), the change in R^2 could be defined as $\delta_k R^2$ (Agresti and Finlay, 1997). For any given ordering j of K predictor variables ($j=1, 2, \cdots, J$, where $J=K!$), the sum of the changes in the explained variance associated with the variables X_k, $k=1,2,\cdots,K$ will sum to the explained variance, i.e., $\Sigma_k(\delta_k R^2)=R^2$. By taking the average over all possible orders, $J^{-1}\Sigma_j\ (\delta_k R^2)$, one obtains the average contribution of X_k to R^2. Kruskal (1987) called $J^{-1}\Sigma_j\ (\delta_k R^2)$ as an relative importance of the predictor X_k. In OLS (see Tatsuoka, 1971), $R^2=\ \Sigma_k(b_k^* r_k)$, where b_k^* is the OLS sample estimate for the standardized regression coefficient for the predictor X_k, and r_k is the sample (Pearson) correlation coefficient between the dependent variable Y and the predictor X_k. Since $b_k^*\ r_k$ could be described as the direct contribution to explained variance of the variable X_k, one might obtain that for $k=1,\cdots,K$, $j=1,\cdots,J(K!)$,

$$R^2 = \Sigma_k(b_k^* r_k) = \Sigma_k(\delta_k R^2) = \Sigma_k[J^{-1}\Sigma_j(\delta_k R^2)]. \tag{2.1}$$

For logistic regression, McFadden (1974) developed the likelihood ratio R^2 based on entropy loss. If LL_0 is the log-likelihood for the null model with no predictors, and LL_M is the log-likelihood for the full model with all of the predictors of interest, then the likelihood ratio R^2 for logistic regression is defined as the following:

$$R_l^2 = 1 - LL_M / LL_0 . \tag{2.2}$$

In OLS, we can decompose the information index over all ordering of the predictors. The average relative importance of a predictor, $J^{-1}\Sigma_j(\delta_k R^2)$, can then be obtained by calculating models with all possible orderings of the predictors, and the sum of the relative importances of all of the predictors, $\Sigma_k[J^{-1}\Sigma_j(\delta_k R^2)]$, adds up to R^2. Now given a measure of the relative importance, $J^{-1}\Sigma_j(\delta_k R^2)$ and the correlation coefficient r_k, an information theoretic fully standardized coefficient can be calculated based on equating $J^{-1}\Sigma_j(\delta_k R^2) = b_k^* r_k$. Now for logistic regression, one could get a measure of the relative importance $J^{-1}\Sigma_j(\delta_k R_l^2)$ and the correlation coefficient r_{lk} which is a square root of R_l^2 with a single predictor X_k. Substituting r_{lk} for r_k and R_l^2 for R^2 in the $J^{-1}\Sigma_j(\delta_k R^2) = b_k^* r_k$, and dividing both sides by r_{lk}, the information theoretic standardized logistic regression coefficient may be estimated from a relative importance dividing a correlation coefficient such as

$$b_I^* = [J^{-1}\Sigma_j(\delta_k R_l^2)] / r_{lk} , \tag{2.3}$$

where the sign of b_I^* should be the same as the sign of the unstandardized coefficient b_k.

Now we introduce other coefficients of determination for logistic regression than R_l^2 in the sixth standardized logistic regression coefficient defined in (2.3). Although there is only one generally accepted definition of coefficient of determination in OLS, there are twelve kinds of coefficients of determination for logistic regression (see Mittlbock and Schemper, 1996; Menard, 2000) Based on eight criteria proposed by Kvälseth, (1985), Menard (2000) and Mittlbock and Schemper (1996) preferred R_l^2 in (2.2) and the following R_o^2 based on proportional reduction in (2.4):

$$R_o^2 = 1 - \Sigma_{i=1}^n (Y_i - \hat{p}_i)^2 / \Sigma_{i=1}^n (Y_i - \overline{Y})^2 , \tag{2.4}$$

where $\hat{p}_i$ is a maximum likelihood estimate of $p_i = P(Y_i = 1)$.

And Liao and McGee (2003) proposed two adjusted coefficients of determination ($R_{l,adj}^2$ and $R_{o,adj}^2$ in (2.5) and (2.6), respectively), since both R_l^2 and R_o^2 could overestimate the measure of association.

$$R_{l,adj}^2 = 1 - \widehat{IPE}_l^M / \widehat{IPE}_l^0 \tag{2.5}$$

$$R_{o,adj}^2 = 1 - \widehat{IPE}_o^M / \widehat{IPE}_o^0 , \tag{2.6}$$

where the estimates of Inherent Prediction Error (IPE) are $\widehat{IPE}_l^M = -n^{-1}LL_M - B_l(\hat{p})$ and $\widehat{IPE}_o^M = -n^{-1}\Sigma_i(Y_i - \hat{p}_i)^2 - B_o(\hat{p})$ for the full model, and both $\widehat{IPE}_l^0$ and $\widehat{IPE}_o^0$ are for the null model, and where both biases $B_l(\hat{p})$ and $B_o(\hat{p})$ are defined by Liao and McGee (2003). The adjusted coefficients of determination $R^2_{l,adj}$ and $R^2_{o,adj}$ succeed advantageous properties of R_l^2 and R_o^2, respectively, and all of these coefficients of determination R_l^2, R_o^2, $R^2_{l,adj}$ and $R^2_{o,adj}$ satisfy eight criteria of Kvälseth (1985). Also based on our experience with extended simulated data, the proposed adjusted coefficients are robust when irrelevant predictors are added or when the sample size changes. An R function to obtain R_l^2, R_o^2, $R^2_{l,adj}$ and $R^2_{o,adj}$ is provided on the Web site http://www.geocities.com/jg_liao/software. In this work, these three kinds of R^2s (R_o^2, $R^2_{l,adj}$ and $R^2_{o,adj}$) are substituted for R_l^2 to obtain the sixth standardized logistic regression coefficients in (2.3).

3. Numerical Examples

3.1 The prevalence of marijuana use

To illustrate the calculation of the different standardized logistic regression coefficients, an example has been taken from Menard (1995), in which logistic regression was used to analyze the relationship of the prevalence of marijuana use to three predictors: exposure to delinquent friends (Exposure), belief that it is wrong to violate the law (Belief), and gender (Gender). This data set is not identical with that of Menard (2004), but most analysis results of this work are much similar.

The calculations of the information theoretic fully standardized coefficients b_I^* with R_l^2, R_o^2, $R^2_{l,adj}$ and $R^2_{o,adj}$ are presented in <Table 3.1> and <Table 3.2>, respectively. <Figure 3.1> represents, with respect to R_l^2, R_o^2, $R^2_{l,adj}$ and $R^2_{o,adj}$, the relative proportions of both the relative absolute values of b_I^* and the relative importances $(1/6)\Sigma_j(\delta_k R^2)$ of each three predictors, whose sum is identical with R^2. We could see that the relative importances of each three predictors are much different from the relative absolute values of b_I^*, since the correlation coefficients have different values.

We find that the relative magnitudes of the absolute values of the information theoretic fully standardized coefficients b_I^* based on R_l^2 and R_o^2 have little different values. the relative magnitudes of the absolute values of b_I^* based on R_l^2 (R_o^2) are very much close with those with $R^2_{l,adj}$ ($R^2_{o,adj}$). Their relative magnitudes of b_I^* with R_l^2, R_o^2, $R^2_{l,adj}$ and $R^2_{o,adj}$ of three predictor variables are also shown in <Figure 3.2>.

The changes $\delta_k R^2$ in R^2 should be positive, but some of these terms may be negative and could not be interpreted as proportions (Bock, 1975). And Menard (2000) warned that R_o^2 may decrease when additional variables are added to the model. In this example, the changes $\delta_k R_{l,adj}^2$ and $\delta_k R_{o,adj}^2$ in $R_{l,adj}^2$ and $R_{o,adj}^2$ for Gender variable are found to be negative in fourth row of <Table 3.2> when Gender variable whose relative influence is too small is added to the model.

<Table 3.1> Calculation of b_I^* (with $R_l^2 = 0.3055$, $R_o^2 = 0.3800$)

Equation (Order of variable entry into equation)	Change in R_l^2 : $\delta_k R_l^2$				Change in R_o^2 : $\delta_k R_o^2$			
	Exposure	Belief	Gender	R_l^2	Exposure	Belief	Gender	R_o^2
1: (Exposure, Belief, Gender)	0.2697	0.0303	0.0055	0.3055	0.3420	0.0310	0.0071	0.3800
2: (Exposure, Gender, Belief)	0.2697	0.0324	0.0034	0.3055	0.3420	0.0332	0.0049	0.3800
3: (Belief, Exposure, Gender)	0.1341	0.1659	0.0055	0.3055	0.1581	0.2149	0.0071	0.3800
4: (Belief, Gender, Exposure)	0.1393	0.1659	0.0003	0.3055	0.1646	0.2149	0.0006	0.3800
5: (Gender, Exposure, Belief)	0.2706	0.0324	0.0025	0.3055	0.3435	0.0332	0.0034	0.3800
6: (Gender, Belief, Exposure)	0.1393	0.1638	0.0025	0.3055	0.1646	0.2121	0.0034	0.3800
Relative Importance	0.2038	0.0985	0.0033	0.3055	0.2524	0.1232	0.0044	0.3800
Correlation Coefficient	0.5193	0.4073	0.0496	-	0.5848	0.4636	0.0582	-
b_I^*	0.3924	-0.2417	0.0660	-	0.4317	-0.2657	0.0755	-

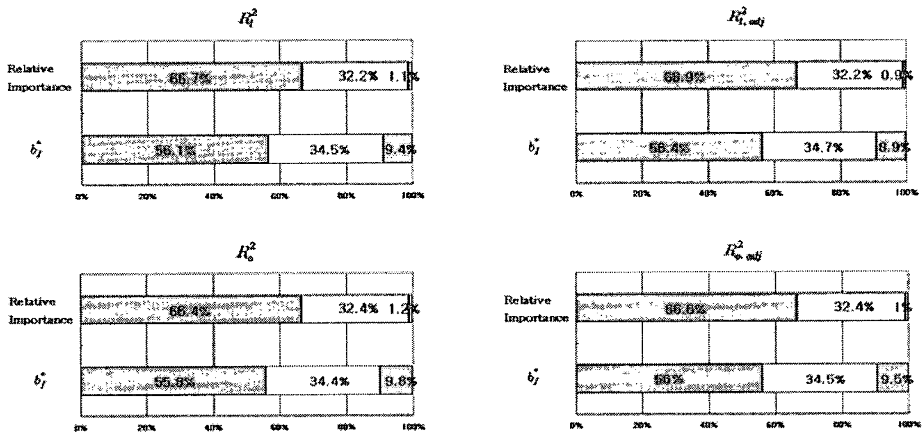
<Table 3.2> Calculation of b_I^* (with $R_{l,adj}^2 = 0.3035$, $R_{o,adj}^2 = 0.3782$)

Equation (Order of variable entry into equation)	Change in $R_{l,adj}^2$: $\delta_k R_{l,adj}^2$				Change in $R_{o,adj}^2$: $\delta_k R_{o,adj}^2$			
	Exposure	Belief	Gender	$R_{l,adj}^2$	Exposure	Belief	Gender	$R_{o,adj}^2$
1: (Exposure, Belief, Gender)	0.2688	0.0297	0.0049	0.3035	0.3414	0.0304	0.0064	0.3782
2: (Exposure, Gender, Belief)	0.2688	0.0318	0.0028	0.3035	0.3414	0.0326	0.0042	0.3782
3: (Belief, Exposure, Gender)	0.1335	0.1651	0.0049	0.3035	0.1575	0.2143	0.0064	0.3782
4: (Belief, Gender, Exposure)	0.1386	0.1651	-0.0003	0.3035	0.1641	0.2143	-0.0001	0.3782
5: (Gender, Exposure, Belief)	0.2698	0.0318	0.0019	0.3035	0.3431	0.0326	0.0026	0.3782
6: (Gender, Belief, Exposure)	0.1386	0.1630	0.0019	0.3035	0.1641	0.2156	0.0026	0.3782
Relative Importance	0.2030	0.0978	0.0027	0.3035	0.2519	0.1226	0.0037	0.3782
Correlation Coefficient	0.5185	0.4064	0.0431	-	0.5843	0.4629	0.0506	-
b_I^*	0.3915	-0.2406	0.0617	-	0.4312	-0.2649	0.0728	-

<Table 3.3> summarizes the results of calculating the standardized logistic regression coefficients. In <Table 3.3>, it is evident that the five standardized coefficients (b_A^*, b_S^*, b_L^*, and b_M^*, and b_I^* based on R_l^2) produce the same order of magnitude for the effects of the

three predictors: Exposure, Belief, and Gender. This unsurprising reasons are discussed by Menard (2004). However, from the results (<Table 3.7> and <Figure 3.4>) of another example in Section 3.2, we can find that these standardized coefficients do not produce the same order of magnitude, which will be discussed later.

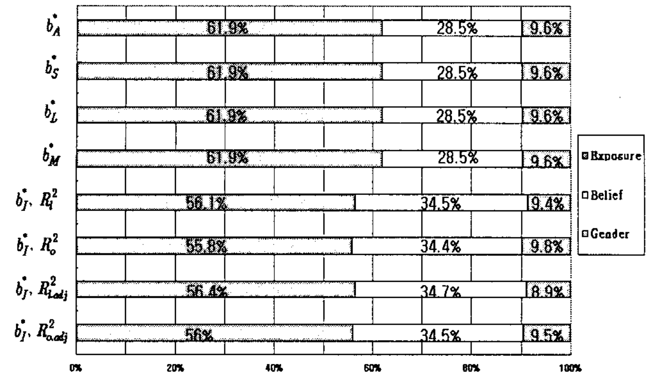
Menard (2004) mentioned that the absolute values of standardized coefficients, b_A^*, b_S^*, b_L^*, and b_M^* of three predictor variables are decreasing. But it is apparently confirmed that their relative magnitudes of these standardized coefficients are identical for three predictor variables from <Table 3.3> and <Figure 3.2>.



<Figure 3.1> Relative Proportions of the Information Theoretic Fully Standardized Coefficients and Relative Importances

<Table 3.3> Standardized Logistic Regression Coefficients

Types		Standardized Coefficients		
		Exposure	Belief	Gender
b_A^*		1.4458	-0.6644	0.2252
b_S^*		0.7970	-0.3663	0.1242
b_L^*		0.5138	-0.2361	0.0800
b_M^*		0.4308	-0.1980	0.0671
b_I^*	R_I^2	0.3924	-0.2417	0.0660
	R_o^2	0.4317	-0.2657	0.0755
	$R_{I,adj}^2$	0.3915	-0.2406	0.0617
	$R_{o,adj}^2$	0.4312	-0.2649	0.0728



<Figure 3.2> Relative Comparison of Standardized Coefficients

3.2 A study of carriers of muscular dystrophy

Another example is illustrated with data from a study of carriers of muscular dystrophy (Freund and Littell, 2000). Two groups of women, one consisting of known carriers of the disease and the other a control group, were examined for three (four in Freund and Littell, 2000) types of protein in their blood. it is known that proteins may be used as a screening tool to identify carriers. The objective is to determine the effectiveness of these proteins to identify carriers of the disease. The variables in the data are Carrier (0 for control and 1 for carriers), P1 (measurement of protein type 1), P2 (protein type 2), and P3 (protein type 4). In this work, we omit a predictor variable which measures protein type 3, since the score chi-square test statistic of protein type 3 is not significant.

<Table 3.4> Logistic Regression Results

Predictor	Standard deviation	Coefficient	Wald statistic Chi-square	p-value
P1	125.4351	0.0276	5.0431	0.0247
P2	13.4360	0.0956	5.7476	0.0165
P3	69.1495	0.0269	6.8513	0.0089

<Table 3.4> presents the results for the logistic regression analysis of a data of carriers of muscular dystrophy with three predictors, P1, P2 and P3. The standard deviation of each predictor in <Table 3.4> will be used in calculating standardized coefficients b_A^*, b_S^*, b_L^* and b_M^*.

The calculations of the information theoretic fully standardized coefficients b_I^* with R_l^2, R_o^2, $R_{l,adj}^2$ and $R_{o,adj}^2$ are presented in <Table 3.5> and <Table 3.6>. The properties of the values and relative magnitudes of b_I^* in <Table 3.5> and <Table 3.6> are similar with those in Section 3.1. We find that all changes, $\delta_k R^2$, in R^2 for three predictors are positive in

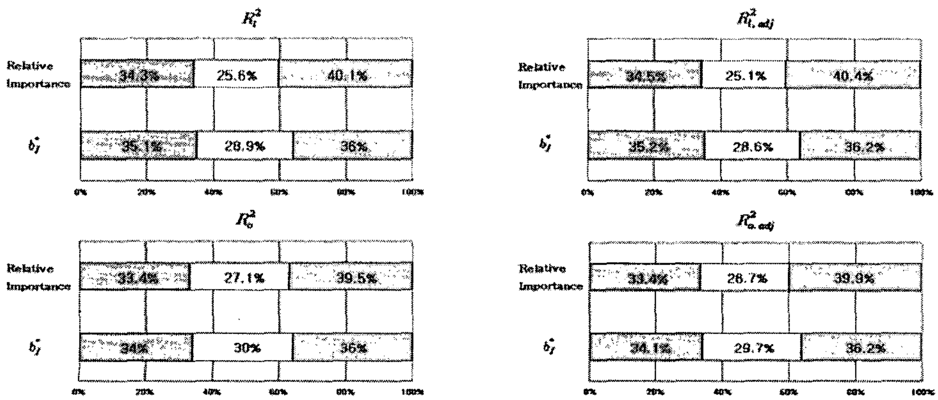
<Table 3.5> and <Table 3.6>.

<Table 3.5> Calculation of b_I^* (with $R_l^2 = 0.5829$, $R_o^2 = 0.6280$)

Equation (Order of variable entry into equation)	Change in R_l^2 : $\delta_k R_l^2$				Change in R_o^2 : $\delta_k R_o^2$			
	P1	P2	P3	R_l^2	P1	P2	P3	R_o^2
1: (P1, P2, P3)	0.2840	0.1883	0.1106	0.5829	0.3302	0.1975	0.1003	0.6280
2: (P1, P3, P2)	0.2840	0.0776	0.2213	0.5829	0.3302	0.0793	0.2185	0.6280
3: (P2, P1, P3)	0.2403	0.2321	0.1106	0.5829	0.2477	0.2800	0.1003	0.6280
4: (P2, P3, P1)	0.1281	0.2321	0.2228	0.5829	0.1077	0.2800	0.2404	0.6280
5: (P3, P1, P2)	0.1368	0.0776	0.3686	0.5829	0.1336	0.0793	0.4151	0.6280
6: (P3, P2, P1)	0.1281	0.0863	0.3686	0.5829	0.1077	0.1053	0.4151	0.6280
Relative Importance	0.2002	0.1490	0.2337	0.5829	0.2095	0.1702	0.2483	0.6280
Correlation Coefficient	0.5329	0.4817	0.6071	-	0.5746	0.5291	0.6443	-
b_I^*	0.3757	0.3093	0.3850	-	0.3646	0.3217	0.3854	-

<Table 3.6> Calculation of b_I^* (with $R_{l,adj}^2 = 0.5434$, $R_{o,adj}^2 = 0.6004$)

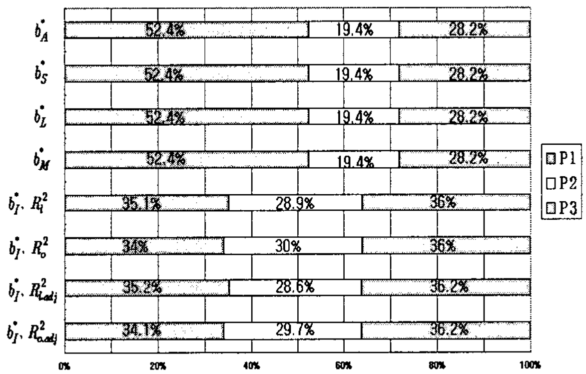
Equation (Order of variable entry into equation)	Change in $R_{l,adj}^2$: $\delta_k R_{l,adj}^2$				Change in $R_{o,adj}^2$: $\delta_k R_{o,adj}^2$			
	P1	P2	P3	$R_{l,adj}^2$	P1	P2	P3	$R_{o,adj}^2$
1 : (P1, P2, P3)	0.2706	0.1757	0.0970	0.5434	0.3201	0.1883	0.0921	0.6004
2 : (P1, P3, P2)	0.2706	0.0653	0.2074	0.5434	0.3201	0.0694	0.2109	0.6004
3 : (P2, P1, P3)	0.2278	0.2186	0.0970	0.5434	0.2377	0.2706	0.0921	0.6004
4 : (P2, P3, P1)	0.1150	0.2186	0.2098	0.5434	0.1003	0.2706	0.2295	0.6004
5 : (P3, P1, P2)	0.1244	0.0653	0.3536	0.5434	0.1245	0.0694	0.4065	0.6004
6 : (P3, P2, P1)	0.1150	0.0748	0.3536	0.5434	0.1003	0.0936	0.4065	0.6004
Relative Importance	0.1872	0.1364	0.2197	0.5434	0.2005	0.1603	0.2396	0.6004
Correlation Coefficient	0.5202	0.4675	0.5946	-	0.5657	0.5202	0.6376	-
b_I^*	0.3599	0.2917	0.3695	-	0.3544	0.3082	0.3758	-



<Figure 3.3> Relative Proportions of the Information Theoretic Fully Standardized Coefficients and Relative Importances

<Table 3.7> Standardized Logistic Regression Coefficients

Types		Standardized Coefficients		
		P1	P2	P3
b_A^*		3.4620	1.2845	1.8601
b_S^*		1.9085	0.7081	1.0254
b_L^*		1.2303	0.4565	0.6610
b_M^*		0.6108	0.2266	0.3282
b_I^*	R_l^2	0.3757	0.3093	0.3850
	R_o^2	0.3646	0.3217	0.3854
	$R_{l,adj}^2$	0.3599	0.2917	0.3695
	$R_{o,adj}^2$	0.3544	0.3082	0.3758



<Figure 3.4> Relative Comparison of Standardized Coefficients

<Figure 3.3> represents the relative proportions of both the relative values of b_I^* and the relative importances $(1/6)\Sigma_j(\delta_k R^2)$ of each three predictors, with respect to R_l^2, R_o^2, $R_{l,adj}^2$ and $R_{o,adj}^2$. We could see, in this example, that the relative importances of each three predictors are close to the relative values of b_I^*, since the correlation coefficients have almost same values, which is a contrast to that in Section 3.1 (compare <Figure 3.3> with <Figure 3.1>).

<Table 3.7> summarizes the results of calculating eight kinds of the standardized logistic regression coefficients. Their relative magnitudes of b_I^* of three predictor variables are also shown in <Figure 3.4>. We find that in <Table 3.7> and <Figure 3.4>, it is evident that eight standardized coefficients (b_A^*, b_S^*, b_L^*, b_M^* and four kinds of b_I^*s) do not produce the same order of magnitude for the effects of the three predictors: the most relative important variable is P1 variable based on b_A^*, b_S^*, b_L^* and b_M^* while, based on four kinds of b_I^*s, the most relative important variable is P3 variable.

4. Conclusion

Menard (2004) mentioned that the absolute values of standardized coefficients, b_A^*, b_S^*, b_L^* and b_M^* of three predictor variables are decreasing, which are shown at <Table 3.3> and < Table 3.7>. It is confirmed that their relative magnitudes of these standardized coefficients are identical for three predictor variables from <Figure 3.2> and <Figure 3.4>, since these coefficients incorporate the standard deviation of the predictors so that their magnitudes have the same order of relative influence of the predictors.

The information theoretic fully standardized coefficient b_I^* based on R_l^2 was proposed by Menard (2004). In this paper, other three kinds of coefficients of determination R_o^2, $R_{l,adj}^2$ and $R_{o,adj}^2$ are considered additionally to get the information theoretic fully standardized coefficients b_I^*. We might find that the relative proportions of the absolute values of b_I^* of each predictor based on R_l^2 and R_o^2 are little different. The relative proportions of the absolute values of b_I^* based on R_l^2 are much similar with those with $R_{l,adj}^2$ and its proportions based on R_o^2 are also much similar with those with $R_{o,adj}^2$.

The changes $\delta_k R^2$ in R^2 may be negative and could not be interpreted as proportions (Bock, 1975, Menard, 2000). The changes, $\delta_k R_{l,adj}^2$ and $\delta_k R_{o,adj}^2$, in $R_{l,adj}^2$ and $R_{o,adj}^2$ are also found to be negative. This case is found in <Table 3.2> obtained from an illustrated example in Section 3.1.

Menard (2004) discussed that the five standardized coefficients (b_A^*, b_S^*, b_L^*, b_M^* and b_I^* based on R_l^2) produce the same order of magnitude for the effects of the three predictors in his example. However, we also find that these standardized coefficients, b_A^*, b_S^*, b_L^*, b_M^* and four kinds of b_I^*, may not produce the same order of magnitude in <Table 3.7> and <Figure 3.4> obtained from an example in Section 3.2. Therefore, we may conclude that the other standardized coefficients should be obtained and compared to analyze the logistic regression with the sixth standardized coefficient which is known to be the best from a conceptual standpoint.

References

[1] Agresti, A. (1996). *An Introduction to Categorical Data Analysis*, Wiley, New York.

[2] Agresti, A. and Finlay, B. (1997). *Statistical Methods for the Social Sciences* (3rd ed.). Dellen/Macmillan, San Francisco.

[3] Bock, D.R. (1975). *Multivariate Statistical Methods in Behavioral Research*. McGraw-Hill Series in Psychology.

[4] Freund, R.J. and Littell, R.C. (2000). *SAS System for Regression* (3rd ed.). Sas Institute:

Wiley Inter-Science.

[5] Goodman, L.A. (1972). A Modified Multiple Regression Approach to the analysis of Dichotomous Variables. *American Sociological Review,* Vol. 37, 28-46.

[6] Kruskal, W. (1987). Relative Importance by Averaging Over Orderings. *The American Statistician*, Vol. 41, 6-10.

[7] Kvälseth, T, O. (1985). Cautionary Note About R^2. *The American Statistician*, Vol. 39, 279-285.

[8] Liao, J.G. and McGee, D. (2003). Adjusted Coefficients of Determination for Logistic Regression. *The American Statistician*, Vol. 57, 161-165.

[9] Long, J.S. (1997). *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks, CA: Sage.

[10] Mayer, L.S. and Younger, M.S. (1976). Estimation of Standardized Regression Coefficients. *Journal of the American Statistical Association*, Vol. 71, 154-157.

[11] McFadden, D. (1974). The Measurement of Urban Travel Demand. *Journal of Public Economics*, Vol. 3, 303-328.

[12] Menard, S. (1995). *Applied Logistic Regression Analysis.* Thousand Oaks, CA: Sage.

[13] Menard, S. (2000). Coefficients of Determination for Multiple Logistic Regression Analysis. *The American Statistician*, Vol. 54, 17-24.

[14] Menard, S. (2004). Six Approaches to Calculating Standardized Logistic Regression Coefficients. *The American Statistician*, Vol. 58, 218-223.

[15] Mittlbock, M. and Schemper, M. (1996). Explained Variation for Logistic Regression. *Statistics in Medicine*, Vol. 15, 1987-1997.

[16] Tatsuoka, M.M. (1971). *Multivariate analysis: Techniques for Educational and Psychological Research*, Wiley, New York.

[Received September 2005, Accepted December 2005]