International Journal of Mathematical Archive-4(7), 2013, 66-73
Available online through www.ijma.info ISSN 2229 – 5046

SIZE AND POWER OF GOODNESS-OF-FIT TESTS FOR LOGISTIC REGRESSION MODEL WITH MISSING INTERACTION TERMS

B. Muniswamy[1*], Shibru Temesgen Wakweya[2] and B. Punyavathi[3]

[1]Department of Statistics, Andhra University, Visakhapatnam, (A.P.), India

[2]Department of Statistics, Addis Ababa University, Ethiopia

[3]Department of Statistics, Sri Venkateswara University, Tirupati, (A.P.), India

(Received on: 24-05-13; Received & Accepted on: 17-06-13)

ABSTRACT

An assessment of model fit and an evaluation of how well model-based predicted outcomes coincide with the observed data is an important component of any modeling procedure. When at least one continuous predictor is present, classical Pearson and deviance goodness-of-fit tests for logistic regression model are invalid. The Hosmer–Lemeshow test can be used in these situations. However, it does not have desirable power in many cases and provides no further information on the source of any detectable lack of fit. We propose a new method for goodness-of-fit testing that uses partitioning in the covariate space using the estimated probabilities from the assumed model. Properties of the proposed statistics are discussed, and a simulation study demonstrates increased power to detect omission of interaction terms in a variety of settings controlling type I error rates.

Keywords: *logistic regression, link function, maximum likelihood estimates, goodness of fit.*

1. INTRODUCTION

The specific form of the logistic regression model with unknown parameters $\beta_0, \beta_1, \dots, \beta_p$ is given as:

$$\pi(x) = \frac{Exp(X^T\beta)}{1 + Exp(X^T\beta)} \tag{1.1}$$

where $X = (x_0,\ x_1, \dots, x_p)^T$ and $\beta = (\beta_0,\ \beta_1, \dots, \beta_p)^T$

Parameter estimation in logistic regression using maximum likelihood and testing the significance of the regression coefficients using either Wald or score tests are usually key goals of the analysis (Cox and Snell, 1989). It is clear that, this maximum likelihood estimation depends on the link function and covariates included in the model. Significance testing of each coefficient provides information about the relationship between the covariate and response, relative to overall variability. Goodness-of-fit tests, on the other hand, reflect whether the predicted values are an accurate representation of the observed values.

The logistic regression model is being used with increasing rate in various fields in data analysis. In spite of such increase, there has been no commensurate increase in the use of commonly available methods for assessing the model adequacy (S.K Sarkar and H. Midi, 2010). The objective of this paper is to propose a better method of detecting omission of interaction terms from logistic regression model.

If the regression of the response variable on treatment and covariates is linear or exponential, *omission of important covariates or/and their interactions* only reduces the efficiency of the regression coefficient estimates, but has no effect on the consistency of the estimation. For logistic regression, this omission not only reduces the efficiency of the coefficient estimation, but also affects the consistency of the coefficient estimation. It leads to biased estimates of treatment effect, even in randomized experiments (Gail et al., 1984, 1988; Hauck et al., 1991; Robinson and Jewell, 1991). In case-control studies, the consistency of the estimators of the population odds ratio is still maintained if a correct logistic regression model is specified (Xie and Manski, 1988; Nagelkerke *et al*., 1995, 2005).

Corresponding author: *B. Muniswamy[1*]*
[1]Department of Statistics, Andhra University, Visakhapatnam, (A.P.), India

In Section 2, a general method for assessing goodness-of-fit test in logistic regression models, applicable to a wide range of continuous and categorical covariate configurations, is presented. Focus, however, is on the situation where there is a mix of continuous and discrete covariates and their interaction terms. Theoretical considerations of the proposed statistics are discussed. A simulation study comparing the proposed tests with existing tests is presented in Section 3.

The widely used chi-square and deviance statistic can be used as a measure of how far observed sample data deviate from a theoretical model providing expected counts in each of the distinct covariate patterns. The term used to represent a set of values for the explanatory variables for each subject is called a covariate pattern, which will be denoted by j $(j = 1, 2, ..., J)$. If $J=n$, we call it first type covariate pattern and if $J<n$ we call it second type covariate for our understanding. The deviance, D and Pearson Chi square, χ^2 can be respectively expressed as (Hosmer and Lemeshow, 1989):

$$D = \sum_{j=1}^{J} \left\{ 2 \left[y_j \, log \frac{y_j}{m_j \hat{\pi}(x_j)} + (m_j - y_j) \, log \frac{m_j - y_j}{m_j \left(1 - \hat{\pi}(x_j)\right)} \right] \right\}^2 \tag{1.2}$$

$$\chi^2 = \sum_{j=1}^{J} \frac{\left(y_j - \hat{\pi}(x_j)\right)^2}{\hat{\pi}(x_j)\left(1 - \hat{\pi}(x_j)\right)} \tag{1.3}$$

If the model is correct, the test statistic of Deviance has approximately a Chi-Square distribution with degrees of freedom J-$(p+1)$. Under the null model and type one covariate pattern (J=n), it can be shown that D degenerates to:

$$D = 2\sum_{j=1}^{J} \left\{ \hat{\pi}(x_j) log \left[\frac{\hat{\pi}(x_j)}{1 - \hat{\pi}(x_j)} \right] + log\left[1 - \hat{\pi}(x_j)\right] \right\} \tag{1.4}$$

Then from (1.4) it can be seen that D is completely independent of the observations and contains absolutely no information about the model fit. The Pearson chi-square statistic performs not that much better in this situation, for it can be shown that $\chi^2 \cong n$, the sample size also not being a helpful goodness of fit test, Jing Xu and MichaelLaValley(2011).

The Hosmer and Lemeshow goodness-of-fit statistic $\hat{C}$ is based on the grouping of estimated probabilities $[\hat{\pi}(x_1), \hat{\pi}(x_2), \dots \hat{\pi}(x_n)]$ obtained from the fitted logistic model in to G groups and defined as follows:

$$\hat{C} = \sum_{k=1}^{G} \frac{(O_k - n_k \bar{\pi}_k)^2}{n_k \bar{\pi}_k (1 = \bar{\pi}_k)}, where \sum_{k=1}^{G} n_k = n \; and \; n_k \; is \; the \; number \; of \; subjects \; in \; k^{th} \;\; group$$

The test statistic $\hat{C}$ may be sensitive to the cut-off points specified to form the groups and therefore, it can be unstable leading to different conclusions. Hosmer and Lemeshow (1989) also developed a statistic based on fixed cut points denoted by $\hat{H}$ in to say G groups. If the null hypothesis of the model fits the data holds, the distribution of both $\hat{C}$ and $\hat{H}$ for G groups are well approximated by the chi-square distribution with G-2 degrees of freedom.

Osius and Rojek (1992) has developed the test statistic which is obtained by modification of Pearson Chi-square test statistic, and it is approximately normally distributed when the number of subjects is large. The procedure to obtain this test statistic is described for the type two covariate pattern, $J<n$ as:

$$Let \; v_j = m_j \hat{\pi}_j (1 - \hat{\pi}_j), \;\; c_j = \frac{(1 - 2\hat{\pi}_j)}{v_j} \qquad j = 1, 2, \dots, J \;\; and \;\; A = \left(J - \sum_{j=1}^{J} \frac{1}{m_j}\right)$$

perform a weighted linear regression of c on x. Compute the standardized statistic, Z as follows: $Z = \frac{[\chi^2 - (J - p - 1)]}{\sqrt{A + RSS}}$,

where , χ^2 is the Pearson Chi-square statistics, RSS is the residual sum of squares from this regression and A is correction factor for the variance.

The Tsiatis goodness-of-fit statistic (Tsiatis, 1980) partitions the multidimensional space of covariates into some distinct regions instead of grouping observations by their predicted outcomes. Pulkstenis and Robinson (2002) presented a goodness-of-fit method which draws on the notable strengths of both the Hosmer–Lemeshow approach and Tsiatis approach. However, their statistics are constructed to deal with only the situations when both categorical and continuous covariates are present and the number of cross-classifications of categorical covariates is not too large.

A generalized logistic model proposed by Stukel (1988) provides a convenient model amenable to testing the adequacy of the fitted logistic model. The Stukel uses a logit function with two additional parameters and under null hypothesis it is well approximated by the chi-square distribution with 2 degrees of freedom. The test has not been implemented in any package; but it can be easily obtained from the following procedures.

(1) Save the fitted value from the assumed model, $\hat{\pi}_j$, j=1, 2, ..., J. (2) Compute the estimated logit

$\hat{g}(x_j) = log\left(\frac{\hat{\pi}(x_j)}{1-\hat{\pi}(x_j)}\right) = x_j^{'}\beta$. (3) Compute two new covariates Z_1 and Z_2 as follow:

$Z_{1j} = 0.5 * [\hat{g}(x_j)]^2 * I(\hat{\pi}(x_j) \geq 0.5)$ and $Z_{2j} = -0.5 * [\hat{g}(x_j)]^2 * I(\hat{\pi}(x_j) < 0.5)$ and (4) Perform the score test for t he addition Z_1 and Z_2 to the model.

Several other studies have provided insights in different aspects of goodness of fit tests for logistic regression model, some of them includes: Brown’s Score test (Brown 1982), le Cessie and van Howelingen’s test (le Cessie and Howelingen 1991), Copas’s (1983) residual sum of squares test, Xian-Jin et al.(2008), Rodney and Hosmer (2007), Hosmer et.al(2002)

2. THE PROPOSED METHOD

Suppose we have n pairs of observations(x_i, y_i), i = 1, 2, ..., n where y_i follows a Bernoulli distribution and x_i is a set of covariate value associated with y_i. The assumed logistic regression model is:

$$logit[\pi(x_j)] = x_i^{'}\beta, i = 1, 2, \dots, n \qquad (2.1)$$

where $x_i^{'} = [1, x_{1i}, x_{2i}, \dots, x_{pi},]$ represents a set of values of the p+1 covariates for the i^{th} subject and $\beta_{(p+1)X_1}$ denotes p+1 regression parameters. As discussed earlier, the estimates of the parameters are frequently obtained by the maximum likelihood method and denoted by $\hat{\beta}$. The assumed model (2.1) may or may not fit the data adequately as expected. We say that lack-of-fit is present in the model (2.1) if the “true” unknown model for the data is

$$logit[\pi(x_i)] = x_i'\beta + w_i'\Delta, \qquad i = 1, 2, \dots, n \qquad (2.2)$$

The value of $w_i'\Delta$ is not known to us. We propose to approximate the true model by a segmented logistic regression model. The approximately true model is constructed as follow: Group the probability of success into G mutually exclusive and exhaustive intervals:

$$S_g = (\pi_{g-1}, \pi_g), \; g = 1,2, \dots, G (G \geq 2), \text{with } \pi_G = 1 \text{ and } \pi_0 = 0$$

Accordingly group the data points into G mutually exclusive and exhaustive groups:

$$D_g = \{(x_{gj}, y_{gj}): j = 1,2, \dots, n_g, \pi(x_{gj}) \in S_g, \qquad g = 1,2, \dots, G\}$$

Where x_{gj} is the j^{th} set of covariate value in the g^{th} group. Of course, we do not know $\pi(x_{gj})$. Hence, we estimate $\pi(x_{gj})$ by the maximum likelihood method from the assumed model (2.1)

To approximate the unknown true model (2.2) the following family of segmented logistic regression model is used

$$logit[\pi(x_{gj})] = x_{gj}'\beta + z_{gj}'\alpha_g, g = 1, .., G; \; j = 1, \dots, n_g \qquad (2.3)$$

where $\sum_{g=1}^{G} n_g = n \;\; and \;\; z_{gj} = \begin{cases} x_{gj}x_{1j}, & if\ (x_{gj}, y_{gj}) \in D_g \\ o, & if\ (x_{gj}, y_{gj}) \notin D_g \end{cases}$

For the logistic regression model with only continuous covariates, *G=2 or 3*, seems to be sufficient. If the model contains both continuous and categorical covariates, we segment each category group defined by the categorical variables into two to three subgroups.

Let $L(X\,,Z)$ and $L(X)$ be the maximum likelihood obtained from the segmented model (2.3) and the assumed model (2.1) respectively. If the assumed model (2.1) is the true model, then $-2log\,L(X)\;-\;(-2log\,L(X\,,Z))$ follows an asymptotic Chi-square distribution with degrees of freedom (df) df = rank (XZ) − rank(X). This is an overall goodness-of-fit test for the assumed logistic regression model (2.1). Therefore, the test statistics, say T is the difference between -2times the log likelihood functions:$T = -2\log\Lambda$, Where $\Lambda = \frac{L(X)}{L(X,Z)}$

Consider a partition model for which models from different groups apart from having different intercepts, they all have the same regression coefficients

$$logit[\pi(x_{gj})] = x'_{gj}\beta + b_g\,,\;\; g = 1,..G\,;\;\; j = 1,\dots,n_g \tag{2.4}$$

Let $L(X,b)$ be the maximum likelihood of the model (2.4), we can write the overall lack of- fit test statistic as:

$$T = -2\log\Lambda_1 + -2\log\Lambda_2,\; where\;\; \Lambda_1 = \frac{L(X)}{L(X,b)} \text{ and } \Lambda_2 = \frac{L(X,b)}{L(X,Z)}$$

$$df_1 = rank(X,b) - rank(X\,),\qquad df_2 = rank(X,b) - rank(X\,,Z).$$

We call the first term the log- likelihood of between-group lack-of-fit and the second term the log- likelihood of within-group lack-of-fit. If the assumed model is correct, then lacks of these terms have approximately a Chi-square distribution with degrees of freedom df_1 and df_2 respectively. If the overall lack-of-fit test is significant, we can carry out the between group and within-group lack-of-fit tests to determine whether the lack-of-fit is due to between-group or within-group. For example, if the former is significant but the latter is not, we may conclude that the lack-of-fit may be due to difference of levels between groups and model (2.4) may be sufficient for fitting the data. However, if the latter is significant but the former is not, then it is an indication that model (2.4) is not sufficient to explain the data and we need a separate model for each group, that is, a segmented model.

3. SIMULATION STUDY

One option in investigating the power is to generate data under an alternative model, perform logistic regression on the generated data, and determine how often each goodness of fit test rejected the nul1 hypothesis of an adequate logistic model. Suppose that the assumed model is of the form:

$$H_0 : Logit(\pi(x)) = \beta_0 + \beta_1 x + \beta_2 D\;\; with\;\; \beta_0 = -0.5,\;\; \beta_1 = 0.4\,,\;\; \beta_2 = 0.1 \tag{3.1}$$

In all simulations we first generated a sample size of 50, 100, 200 and 400 from (3.1) with covariate x independently generated from uniform distribution over the interval (-3, 3), and the covariate D independently generated from a Bernoulli distribution with probability 0.5. And then we generated the outcome variable, y by comparing an independently generated U(0, 1) variate, u, to the true logistic probability using the rule $y = 1\; if\; u \geq \pi(x)\; and\;\; y = 0$ otherwise.

The six known goodness-of-fit test statistics used to compared with the proposed test statistic in the simulation study are Deviance, **DEV**; Pearson Chi-square, **PER** ; Hosmer and Lemeshow's decile of risk statistic, **HL**$\boldsymbol{\hat{C}}$; Hosmer and Lemeshow's predetermined cutoff point statistic, **HL**$\boldsymbol{\hat{H}}$; the Osius and Rojek, **OR**; and Stukel's test statistic, **ST**.

The proposed 4-group partition test partitions each of the two categorical group defined by D into two groups at cut off points set at the 50^{th} percentiles of the estimated probability from the assumed model and the six known tests mentioned were applied to the random samples generated and the proportion of rejections was calculated for each goodness of fit test. The results from the simulations for the assumed models when the sample sizes are 50,100,200 and 400 are summarized in table 3.1.The tabled value is the percent of times the *p*-value from the goodness-of-fit test was less than 0.05, with 1000 replications, this percent should range from 3.65% to 6.35% which is a 95% confidence interval at 5% level of significance and size 1000. Therefore, when the percent ranged from 3.65% to 6.35%, the goodness-of-fit test was interpreted to have Type I error rate close to the nominal level; when the percent was above 6.35%, the test was interpreted to have an inflated Type I error rate; when the percent was below 3.65% the test was interpreted to have a Type I error rate below the nominal level.

Table 3.1: Observed type I error rates for the proposed and known tests

	METHOD						
Size	DEV	PER	$HL\hat{C}$	$HL\hat{H}$	OR	ST	NEW
50	0.189	0.001	0.039	0.064	0.265	0.121	0.052
100	0.567	0.001	0.084	0.054	0.272	0.035	0.053
200	0.908	0.000	0.036	0.056	0.211	0.057	0.046
400	0.999	0.000	0.056	0.04	0.129	0.058	0.052

In this simulation study, the Deviance test cannot control type I error rate. When the assumed model is correct model, the chance of rejecting the null hypothesis increases as the sample size gets larger and larger, which indicate that the large sample size does not help Deviance test to control type I error rate. In addition, all type I error rates of Deviance test are larger than the upper bound of 95% confidence interval. Type I error rates of Pearson Chi-square test are smaller than the lower bound of 95% confidence interval under different sample sizes, even zero per cent chance to reject null hypothesis when the assumed model is true with n=200, 400. The ability of controlling type I error rate of Osius and Rojek's normal approximation test can be improved by larger sample size. Type I error rate of Osius and Rojek's normal approximately normal test are greater than the upper bound of 95% confidence interval. Based on Ttable 3.1, almost all type I error rates of Hosmer and Lemeshow'$\hat{C}$ test, Hosmer and Lemeshow's $\hat{H}$, Stukel's score test and the proposed test fall within 95% confidence interval, under different sample sizes.

Next we used seven different interaction models to study the power with omission of interaction terms from the model.

We generated the outcome variable from a model with:

$$Logit\big(\pi(x)\big) = \beta_0 + \beta_1 x + \beta_2 D + \beta_3 xD \tag{3.2}$$

where β_3 takes on the values $0.0, 0.1, 0.2, \ldots, 0.7$. This allows a series of models that increase in the strength of interactions between x and D. The proposed 4-group partition test and the six known tests were applied to the random sample generated from (3.2) to test the goodness-of-fit of the assumed model (3.1) at 5% level of significance. The simulated rejection rates of these tests were presented in Table 3.2 and graphically shown in Figure 3.1 to 3.4.

Table 3.2: Simulated rejection rates for the proposed and known tests

		β_3							
Size	GOF	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7
50	DEV	0.189	0.143	0.098	0.077	0.055	0.041	0.031	0.023
	PER	0.001	0.000	0.005	0.008	0.014	0.017	0.027	0.037
	$HL\hat{C}$	0.039	0.030	0.037	0.042	0.045	0.052	0.051	0.055
	$HL\hat{H}$	0.064	0.056	0.056	0.054	0.05	0.063	0.062	0.062
	OR	0.265	0.244	0.208	0.175	0.176	0.168	0.162	0.148
	ST	0.121	0.051	0.051	0.053	0.063	0.068	0.072	0.078
	NEW	0.052	0.136	0.149	0.171	0.197	0.228	0.251	0.285
100	DEV	0.567	0.447	0.330	0.242	0.173	0.131	0.093	0.066
	PER	0.001	0.001	0.005	0.005	0.009	0.016	0.024	0.031
	$HL\hat{C}$	0.084	0.055	0.055	0.064	0.069	0.065	0.066	0.067
	$HL\hat{H}$	0.054	0.048	0.041	0.059	0.061	0.076	0.076	0.070
	OR	0.272	0.223	0.200	0.176	0.156	0.149	0.149	0.150
	ST	0.035	0.043	0.041	0.040	0.056	0.100	0.113	0.125
	NEW	0.053	0.089	0.105	0.136	0.194	0.272	0.346	0.408
200	DEV	0.908	0.833	0.723	0.566	0.420	0.306	0.223	0.155
	PER	0.000	0.000	0.000	0.000	0.001	0.004	0.008	0.015
	$HL\hat{C}$	0.036	0.034	0.045	0.053	0.064	0.069	0.071	0.081
	$HL\hat{H}$	0.056	0.061	0.056	0.056	0.064	0.077	0.079	0.079
	OR	0.211	0.156	0.124	0.103	0.104	0.114	0.116	0.135
	ST	0.057	0.054	0.057	0.068	0.086	0.101	0.114	0.127
	NEW	0.046	0.052	0.096	0.182	0.289	0.413	0.532	0.655
400	DEV	0.999	0.990	0.971	0.921	0.820	0.666	0.507	0.382
	PER	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.003
	$HL\hat{C}$	0.056	0.050	0.056	0.056	0.063	0.084	0.095	0.122
	$HL\hat{H}$	0.040	0.042	0.065	0.081	0.090	0.105	0.117	0.151
	OR	0.129	0.108	0.092	0.097	0.112	0.128	0.155	0.182
	ST	0.058	0.058	0.064	0.075	0.099	0.119	0.151	0.182
	NEW	0.052	0.075	0.171	0.334	0.552	0.761	0.882	0.935

This simulation study results suggest that, the power of Deviance test is getting poorer and poorer when the assumed model departures further and further from the true model. For Pearson Chi-square test, bigger sample size and further departure from true model cannot enhance its power to detect the lack of fit. This simulation study showed that the Deviance test and Pearson chi-square test are not applicable to the type one covariate pattern (J=n). For Osius and Rojek's normal approximation test, larger sample sizes improve the ability to control the type one error rate. However, neither larger sample sizes nor further departure from the true model can improve the power of detecting lack-of-fit. Stukel's score test and Hosmer and Lemeshow's $\hat{C}$ and $\hat{H}$ tests have very similar performance: They control type I error rate, but the power of detecting lack of fit cannot be improved much by larger sample size and further departure away from the true model. The proposed test has the best performance among these tests at different sample sizes.

Figure 3.1: Plots of Simulated rejection rates for the proposed and known tests (n=50)

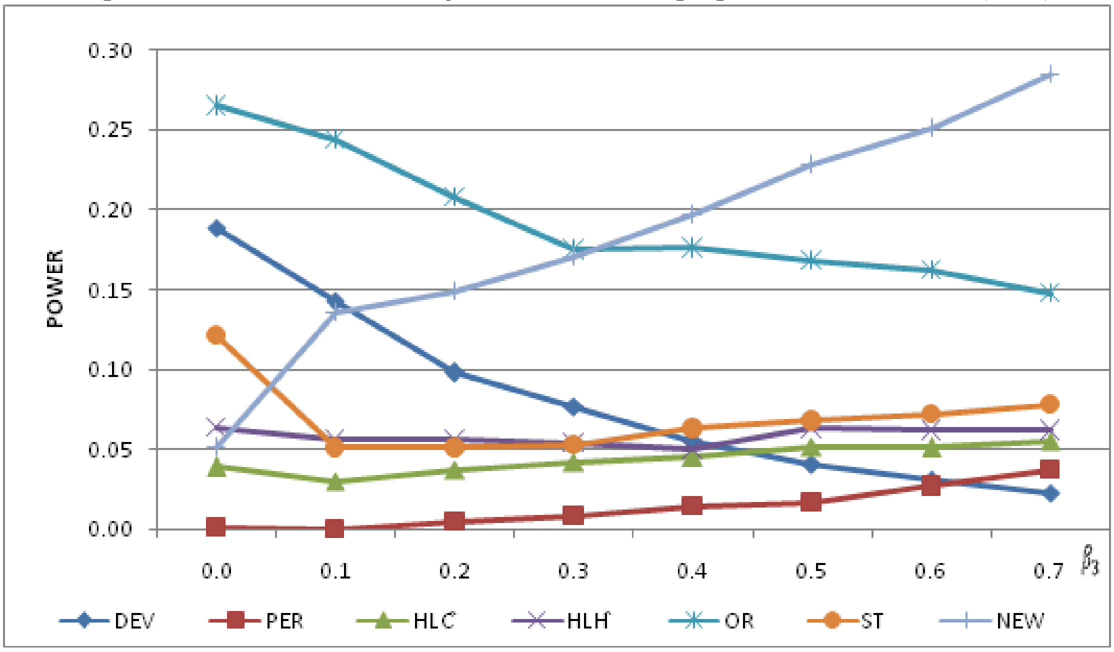


Figure 3.2: Plots of Simulated rejection rates for the proposed and known tests (n=100)

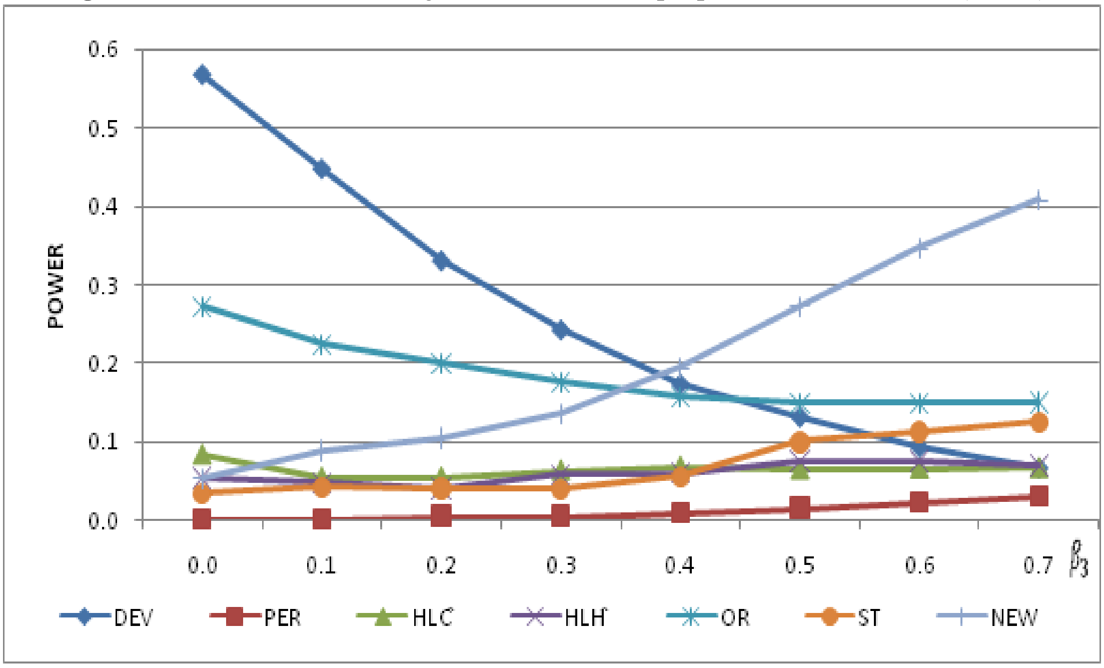


Figure 3.3: Plots of Simulated rejection rates for the proposed and known tests (n=200)

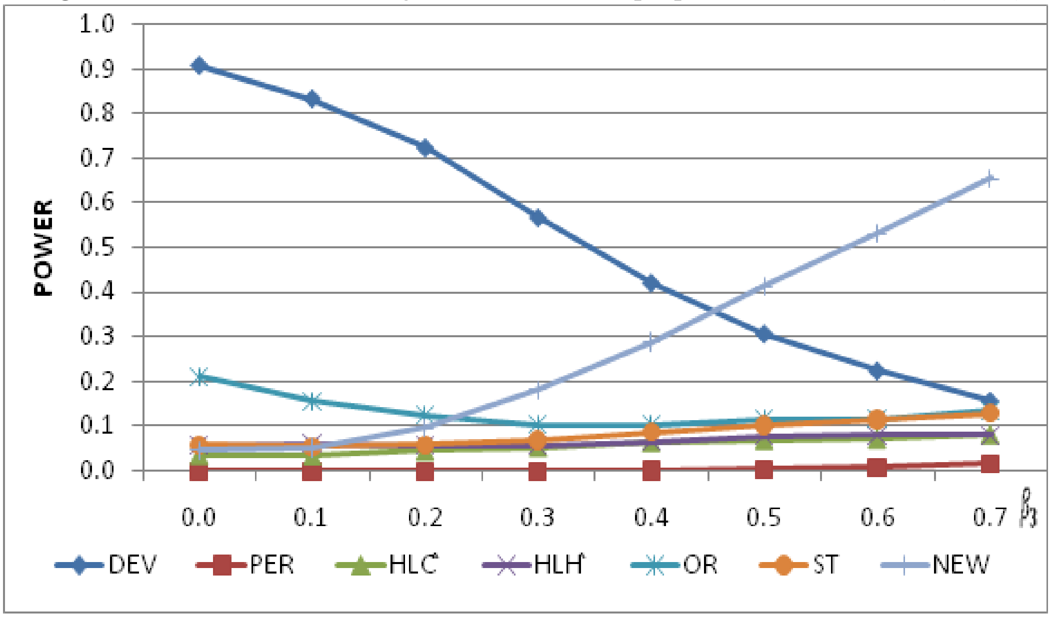
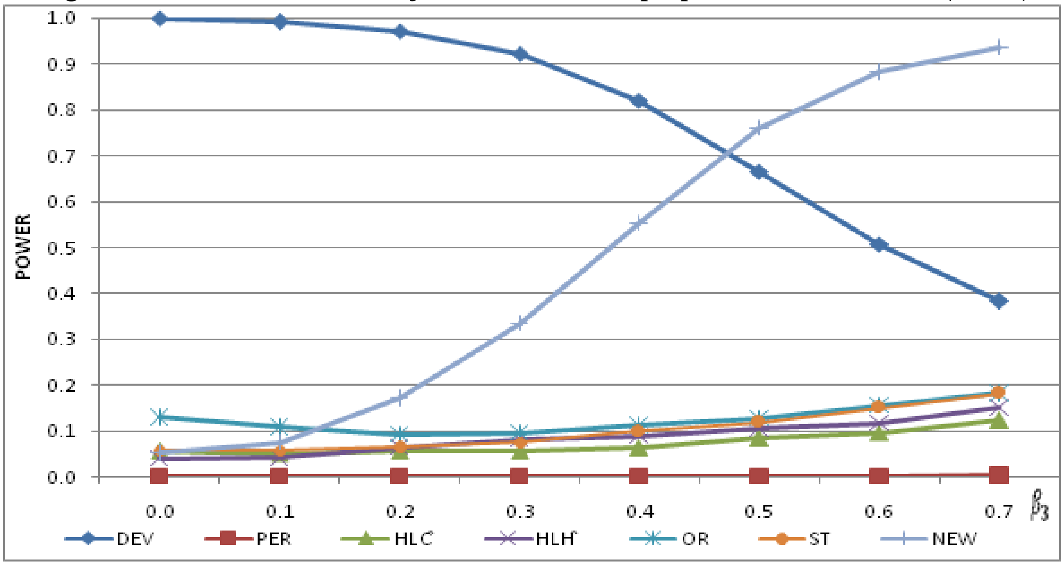


Figure 3.4: Plots of Simulated rejection rates for the proposed and known tests (n=400)



4. CONCLUSION

The results from this simulation study indicates that the proposed test has the best performance among these tests at different sample sizes and degree of departure from the true model under the arrangement created. The ability of controlling type I error rate and power of detecting lack of fit are improved by increasing the sample sizes. Overall, when assessing the goodness of fit of a logistic model in a situation where the predictor variable is a mixture of continuous and categorical, it would be useful to use the proposed method, the Stukel's test and Hosmer and Lemeshow's $\hat{C}$ statistic to detect omission of interaction terms from the model.

REFERENCE

1. Brown, C.C., On a goodness-of-fit test for the logistic model based on score statistics. *Communications in statistics* 11, (1982), 1087-1105.
2. Copas, J.B. Plotting P against x, *Applied Statistics* 32 (1983), 25-31.
3. Cox, D.R., Snell, E.J., The Analysis of Binary Data. Second ed. Chapman and Hall, London, UK, (1989).
4. Gail, M.H., Tan,W.-Y., Piantadosi, S., Tests for no treatment effect in randomized clinical trials when needed covariates are omitted. *Biometrika* 75 (1988), (1), 57–64.
5. Gail, M.H.,Wieand, S., Piantadosi, S., Biased estimates of treatment effect in randomized experiments with nonlinear regressions and omitted covariates. *Biometrika* 71 (1984), (3), 431–444.

6. Hauck, W.W., Neuhaus, J.M., Kalbfleisch, J.D., Anderson, S., A consequence of omitted covariates when estimating odds ratios. *J. Clinical Epidemiology*. 44 (1991), (1), 77–81.
7. Hosmer, D.W., Hjort, N.L., Goodness-of-fit processes for logistic regression: simulation results. *Statistics in Medicine.* 21, (2002), 2723–2738.
8. Hosmer, D.W., Hosmer, T., Le Cessie, S., Lemeshow, S., A comparison of goodness-of-fit tests for the logistic regression model. *Statistics in Medicine*.16, (1997), 965–980.
9. Hosmer, D.W., Lemeshow, S., Applied Logistic Regression. Wiley, NewYork, USA, (1989).
10. Jing Xu and MichaelLaValley, Pearson residual and efficiency of parameter estimates in generalized linear model, *Journal of Statistical Planning and Inference* 141 (2011) 1014–1020.
11. Kuss, O., Global goodness-of-fit tests in logistic regression with sparse data. *Statistics in Medicine*. 21, (2002), 3789–3801.
12. le Cessie, S., van Houwelingen, C., A goodness-of-fit test for binary regression models, based on smoothing methods. *Biometrics* 47, (1991), 1267–1282.
13. Nagelkerke, N., Moses, S., Plummer, F.A., Brunham, R.C., Fish, D., Logistic regression in case-control studies: the effect of using independent as dependent variables. *Statistics in Medicine*. 14, (1995), 769–775.
14. Nagelkerke, N., Smits, J., le Cessie, S., van Houwelingen, H., Testing goodness-of-fit of the logistic regression model in case-control studies using sample reweighting. *Statistics in Medicine.* 24, (2005), 121–130.
15. Nelder, J.A.,Wedderburn, R.W.M., Generalized linear models. *J. Roy. Statist. Soc. A* 135, (1972), 370–384.
16. Osius, G., Rojek, D., Normal goodness-of-fit tests for multinomial models with large degrees of freedom. *Journal of American Statistical, Association* 87 (420), (1992), 1145–1152.
17. Pulkstenis, E., Robinson, T.J., Two goodness-of-fit tests for logistic regression models with continuous covariates. *Statistics in Medicine*. 21, (2002), 79–93.
18. Robinson, L.D., Jewell, N.P., Some surprising results about covariate adjustment in logistic regression models. *Internat. Statist. Rev*. 59, (1991), 227–240.
19. Rodney X. Sturdivanta and David W. Hosmer Jr., A smoothed residual based goodness-of-fit statistic for logistic hierarchical regression models, *Computational Statistics and Data Analysis* 5, (2007), 1 3898 – 3912.
20. S.K Sarkar and H. Midi, Importance of Assessing the Model Adequacy of Binary Logistic Regression, Journal of Applied Science 10(6): (2010) 479-486.
21. Stukel, T.A., Generalized logistic models. *Journal of American Statistical, Association* 83, (1988), 426-431.
22. Tsiatis, A.A., A note on a goodness-of-fit test for the logistic regression model. *Biometrika* 67, (1980), (1), 250–251.
23. Xian-Jin Xie, Jane Pendergast, William Clarke, Increasing the power: A practical approach to goodness-of-fit test for logistic regression models with continuous predictors. *Computational Statistics and Data Analysis* 52(5): (2008), 2703-2713.
24. Xie,Y., Manski, C.F., The logit model and response-based samples. Sociol. Methods Res. 17, (1988), 283–302.

Source of support: Nil, Conflict of interest: None Declared