International Journal of Mathematics and Statistics Studies
Vol.9, No.1, pp.29-40, 2021
Print ISSN: 2053-2229 (Print), Online ISSN: 2053-2210 (Online)

ASSESSMENT THE DISTRIBUTION OF THE INFORMATION MATRIX TEST BY DIFFERENT LEVEL OF ELEMENTS OF THE MATRIX TO AVOID THE SINGULARITY PROBLEM

Nuri H. Salem Badi
Faculty of Science, Statistical Department, University of Benghazi, Benghazi Libya
Emails: nuri.badi@uob.edu.ly

ABSTRACT: *The behaviour of the distribution of goodness-of-fit tests is an important statistical problem especially in dichotomous data. Many authors, discussed methods to examine and compare between different goodness-of-fit tests, where they show the Information matrix test (IMT) and the Information Matrix Diagonal test (IMT_{DIAG}) statistics have reasonable power even for a logistic model with very sparse data. However, many issues related to the elements of the log-likelihood functions and covariance matrix which may be quite close to zero, these elements leading to singularity and have an effect on the behaviour of the distribution of statistics. In this paper, we are interested to investigate the behaviour of the distribution of the \$IMT\$ statistic and IMT_{DIAG} statistic, by using different levels of elements of the matrix. Moreover, the mean and the variance of the distribution of statistics should be examined, by the simulation to compare between different levels of elements of the matrix to find a reasonable new form of \$IMT\$ to avoid the Singularity Problem forecasting.*

KEYWORDS: Binary outcome, logistic model, goodness-of-fit tests, Information matrix test, Singularity Problem

INTRODUCTION

The Information Matrix test IMT and IMT_{DIAG} are a tests for general misspecification, proposed by White (1982). The two well-known expressions for the information matrix coincide only if the correct model has been specified and the IMT takes advantage of this fact. The IMT and IMT_{DIAG} avoid the grouping necessary for tests like the Hosmer-Lemeshow test. Many researchers (Lancaster (1984), Newey (1985), Davidson and Mackinnon (1984) and Davidson and Mackinnon (1988)), pointed out the behaviour of the asymptotic distribution of IMT and IMT_{DIAG} statistic. Chesher (1984), discussed the information matrix test and showed that it is useful with binary data models. Kuss (2002), made comparisons between some goodness-of-fit tests in logistic regression models with sparse data. The results of his simulation showed that the IMT_{DIAG} has reasonable power compared with other tests. However, Kuss (2002), did not give information about the asymptotic distribution of the IMT and IMT_{DIAG} statistic. Also he did not focus exclusive in the case mi = 1. Badi (2017), discussed and investigated the behavior of asymptotic distribution of goodness-of-fit tests which found some of goodness-of-fit tests affected by assumption on covariance matrix. Although the IMT and IMT_{DIAG} are extensively discussed in the econometrics literature, it is less well known in the biostatistics literature.

Vol.9, No.1, pp.29-40, 2021
Print ISSN: 2053-2229 (Print), Online ISSN: 2053-2210 (Online)

There are several forms of the IMT, some of which give rather unstable behaviour. A complication in this analysis is that the test statistic is parameter dependent and must be evaluated at the maximum likelihood estimation (M LE) of the parameters of the fitted model. As such we need to the limiting values of these parameters under what may well be a wrong model. In fact, the previous work discussed by Badi (2021), which pointed out, the behaiviour of the dispersion Matrix of the IMT under wrong logistic model which computed empirical variance of IMT, the results appeared some elements of the covariance matrix leading to singularity. The idea of the information matrix test is to compare $E\left(\frac{-\partial^2 \ell}{\partial\theta\partial\theta^T}\right)$ *and* $E\left(\frac{\partial \ell}{\partial\theta}\,\frac{\partial \ell}{\partial\theta^T}\right)$, as these differ when the model is mis-specified but not when the model is correct. The idea of the IMT_{DIAG} is to reduce the elements of the matrix, just consider the diagonal elements to avoid singularity problem. In this paper, we consider the mean and the variance of IMT and use different forms of IMT by delete elements which may be affected on the behavior of the distribution of statistic to find the best form of IMT to avoid the singularity problem.

Fisher Information Matrix for Logistic Regression Model

We consider binary regression, where the outcome for individual $i,\ i = 1, \dots, n,$ is a random variable $Y_i = 1 \epsilon \{0,1\}$. Also $\Pr(Y_i|x_i) = \pi_i = \pi(\beta^T x_i)$ where x_i is a $p \times 1$ dimensional vector of covariates and β is a p-dimensional vector of parameters. It will be convenient to write $\alpha_i = \beta^T x_i$ *and* ℓ_i to be the contribution to the log-likelihood ℓ from unit i. We have:

$$\ell(\beta) = \sum_{i=1}^{n} \ell_i(\beta) = \sum_{i=1}^{n} Y_i \log \pi_i + (1 - Y_i)\log(1 - \pi_i)$$

The p-dimensional likelihood equations $\partial\ell/\partial\beta = 0$ can be written:

$$\frac{\partial \ell}{\partial \beta} = \sum_{i=1}^{n} \left[\frac{(Y_i - \pi_i)}{\pi_i(1-\pi_i)}\right] \frac{\partial \pi_i}{\partial a_i} x_i = 0$$

We can also derive the p × p matrix $\partial^2\ell/\partial\beta\,\partial\beta^T = 0$ as:

$$\sum_{i=1}^{n} \left[\frac{(Y_i - \pi_i)}{\pi_i(1-\pi_i)}\,\frac{\partial^2 \pi_i}{\partial {a^2}_i} - \frac{(Y_i - \pi_i)^2}{{\pi_i}^2(1-\pi_i)^2}\left(\frac{\partial \pi_i}{\partial a_i}\right)^2\right] x_i {x_i}^T$$

In case of logistic regression model, let us consider the standard logistic regression model and for simplicity consider the case

$$\pi_i = expit\,(a_i)\ , \qquad i = (1,2,\dots,n)$$

Vol.9, No.1, pp.29-40, 2021
Print ISSN: 2053-2229 (Print), Online ISSN: 2053-2210 (Online)

where $a_i = \alpha + \beta_1 x_{1i}$. To some writing in the following we write x_{1i} as x_i the dimension of x_i is clear from the context. The first derivatives of the log likelihood are

$$\frac{\partial \ell}{\partial \alpha} = \sum_{i=1}^{n} (y_i - \pi_i) \ , \quad and \quad \frac{\partial \ell}{\partial \beta_1} = \sum_{i=1}^{n} x_i (y_i - \pi_i)$$

So, we then have

$$\frac{\partial^2 \ell}{\partial \alpha^2} = \frac{\partial}{\partial \alpha}\left[\frac{\partial \ell}{\partial \alpha}\right] = -\sum_{i=1}^{n} \left[\frac{\partial}{\partial \alpha}\left(\frac{exp\ (\alpha + \beta_1 x_i)}{1 + exp\ (\alpha + \beta_1 x_i)}\right)\right] = -\sum_{i=1}^{n} \pi_i (1 - \pi_i).$$

Similarly, the second derivative with β_1 is

$$\frac{\partial^2 \ell}{\partial \beta_1^2} = \sum_{i=1}^{n} x_i^2\ \pi_i (y_i - \pi_i)$$

And also, we have

$$\frac{\partial^2 \ell}{\partial \alpha \partial \beta_1} = -\sum_{i=1}^{n}\ x_i\ \pi_i (y_i - \pi_i)$$

Then, the Fisher's information matrix in this case is

$$I_n = \begin{pmatrix} \sum_{i=1}^{n} \pi_i (1 - \pi_i) & \sum_{i=1}^{n}\ x_i\ \pi_i (y_i - \pi_i) \\ \sum_{i=1}^{n}\ x_i\ \pi_i (y_i - \pi_i) & \sum_{i=1}^{n} x_i^2\ \pi_i (y_i - \pi_i) \end{pmatrix}$$

it is evaluated at the M LE $\hat{\beta}$.

Information Matrix test (IMT) and (IMT_{DIAG})

The idea of the IMT is to compare two different estimators of the information matrix to assess model fit. The IM T provides a unified framework for specification goodness of fit tests for a wide variety of distribution, multivariate or univariate, discrete or continuous. Lancaster (1984), pointed out, can be estimated the covariance matrix of IMT, dependent upon the IMT of White (1982), which can be estimated without the computation of analytic third derivatives of the density function. Newey (1985), discussed that, the IMT is sensitive to non-normality. Moreover, he proposed a simple computation procedure which employs the outer product of the gradient (OP G) covariance matrix estimator of IMT statistic. However, Davidson and Mackinnon (1984), argue that, such a procedure maybe gives unreliable inferences, related to the stochastic nature of the covariance matrix estimator which uses high sample moments to estimate high population moments. Davidson and Mackinnon (1988), purposed a simple calculation procedure for the test statistic, for general binary data models, which employs the ML covariance matrix estimator instead the OP G estimator. White (1982), introduced the test statistic as

$$d_g = (y, \theta) = \frac{\partial \ell(y)}{\partial \theta_r} \frac{\partial \ell(y)}{\partial \theta_s} + \frac{\partial^2 \ell(y)}{\partial \theta_r \partial \theta_s}$$

Where g ranges over appropriately chosen elements of the matrix and y will stand in place of the data: $g = 1, \ldots, q \leq \frac{1}{2} p(p+1)$, where $p = \dim(\theta)$ and $r, s = 1, \ldots, p$. The IMT statistic is based on the q-vector

$$D_g(\hat{\theta}_n) = \frac{1}{\sqrt{n}} \sum_{i}^{n} d_g(y_i, \hat{\theta}_n)\,; \;\; 1 \leq g \leq q$$

Where $\hat{\theta}_n$ is the $M\,LE$ under $\ell\,(.)$, where $y_1, y_2, \ldots, y_n$ are the data. We assume that the y_i are independent and identically distribution. Badi (2021), discussed the covariance matrix of the IMT which described the behaiviour of the variance under missing covariate, and he computed a new form of Empirical Variance of the IMT. The results as below: We have that

$$d_g = (y_i - \pi_i)(1 - 2\pi_i) \begin{pmatrix} 1 \\ x_i \\ x_i^2 \end{pmatrix} \text{ and } \qquad \nabla \ell_i = (y_i - \pi_i) \begin{pmatrix} 1 \\ x_i \end{pmatrix}$$

so, the variance is

$$Var\,(d_g) = E(d_g d_g^T) - E(d_g)E(d_g^T)$$

The idea of the IM_{DIAG} test and IM test are the same, the only difference is that for the former the elements of z_i are just the diagonal elements of $x_i x_i^T$, so z_i is the p dimensional vector: $z_i^T = (x_{i1}^2, x_{i2}^2, \ldots, x_{ip}^2)$ To explain the difference in size of vector z_i in the two cases of IM test and IM_{DIAG} test, let us consider a simple example. Suppose we have a symmetric matrix with elements $x_i x_i^T$ and 3 × 3 dimension as:

$$\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix}$$

where, $x_{rs} = x_{ri}x_{si}$. Then in the case of the IM test, the dimension of vector z_i^T is 1 × 6 and elements are :

$$z_i^T = (x_{11}, x_{22}, x_{33})$$

The IM TDIAG Under missing covariates

The behaviour of the IMT under missing covariates logistic model was discussed by Badi (2021), we consider the same idea for the calculation of IMT_{DIAG}. We know that the IMT_{DIAG} approach has the same idea as the Information matrix test, but compares just the diagonal elements of the two form of the information matrix. So, z_i is (p+1)×1 -dimensional vector of the diagonal elements $x_i x_i^T$ matrix. Therefore, the IMT_{DIAG} has the same behaviour of case IMT statistic, but the vector z_i has different dimension and different elements. We consider IMT_{DIAG} Under missing covariates so, we have true model with two covariates X_1 and X_2 and fitting the model with X_1 then,

$$d_g(y_i, \theta) = (Y_i - \pi_i)(1 - 2\pi_i)\begin{pmatrix} 1 \\ X_i^2 \end{pmatrix}$$

Now, we consider the Variance of IMT_{DIAG} for Logistic Regression Model, in this case we have:

$$d_g = (y_i - \pi_i)(1 - 2\pi_i)\begin{pmatrix} 1 \\ x_i^2 \end{pmatrix}$$

So to calculate the variance V, we need to calculate $Var\ (d_g)$ and $Cov(d_g, \nabla\ell)$, we can see $Var(\nabla\ell)$ has the same expression which used in case of IMT. Firstly, we will work out $Var\ (d_g)$, we have

$$dd^T = (y - \pi)^2(1 - 2\pi)^2\begin{pmatrix} 1 & x_i^2 \\ x_i^2 & x_i^4 \end{pmatrix}$$

taking expectation $E_{Y|X}$ we obtain

$$E(dd^T) = E_X\left[(\pi_t(1 - 2\pi) + \pi^2)(1 - 2\pi)^2\begin{pmatrix} 1 & X^2 \\ X^2 & X^4 \end{pmatrix}\right]$$

Secondly, we need to calculate $Cov(d_g, \nabla\ell)$ and $E(\nabla\ell) = 0$ at the least false value, and we have

$$Cov(d_g, \nabla\ell) = E(d\nabla\ \ell^T) = E_X\left[(\pi_t(1 - 2\pi) + \pi^2)((1 - 2\pi))\begin{pmatrix} 1 & X \\ X^2 & X^3 \end{pmatrix}\right]$$

Empirical Variance of IM TDIAG

Now, we need calculate the estimated variance $(\hat{V})$ of d as estimated to $V(d)$. One candidate would be to compute

$$d_i = (y_i - \hat{\pi}_i)(1 - 2\hat{\pi}_i)\begin{pmatrix} 1 \\ x_i^2 \end{pmatrix} \quad i = 1,2,\ldots,n$$

and,

$$\nabla \ell_i = (y_i - \hat{\pi}_i)\begin{pmatrix} 1 \\ x_i \end{pmatrix} \quad i = 1,2,\ldots,n$$

where, $\hat{\pi}_i$ is the fitted value from the model with just x_1 . Now compute

$$\hat{F}_n = \frac{1}{n}\sum_{i=1}^{n} dd^T - \left(\frac{1}{n}\sum_{i=1}^{n} d_i\right)\left(\frac{1}{n}\sum_{i=1}^{n} d_i^T\right)$$

and

$$\hat{E}_n = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{\pi})^2\begin{pmatrix} 1 & x_i \\ x_i & x_i^2 \end{pmatrix}$$

$$\hat{D}_n = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{\pi})^2(1 - 2\hat{\pi})\begin{pmatrix} 1 & x_i \\ x_i^2 & x_i^3 \end{pmatrix}$$

Then,

$$\hat{V} = \hat{F}_n - \hat{D}_n\hat{E}_n^{-1}\hat{D}_n^T$$

Then use it as an estimate of V. As we computed the final form of the variance of IMT_{DIAG}, we can see clearly it is dependent on E(d). The first two elements of E(d), which may be quite close to zero under true model and use the least false value which discussed and calculated by Matthews and Badi (2015) and Badi (2021), the

$$E(\pi_t - \pi) = E\big((\pi_t - \pi)X\big) =$$

related to the log likelihood functions. So, these elements leading to singularity of the estimated covariance matrix, and has effect on the behaviour of the distribution of the IMT, this results discussed by Badi (2021). Although the IMT_{DIAG} one of the approaches to avoid the singular problem which use just the diagonal elements, but this problem still present related to the first element of E(d). The important point here is the investigation a new form of IM T which deleted some elements of the E(d) to avoid the singularity problem.

Information Matrix Test with Different Level of Elements (IMT_{DE})

We know that the parameters estimators, under the null hypotheses H_0 where there is no mis-specification, will be consistent, asymptotically normal and asymptotically efficient estimators. Under the alternative hypothesis H_1 when the model is mis-specified, however, this estimator will be biased and inconsistent. The constructing of the IMT is based on $\widehat{d}$, so, to develop the test the probability limit of d required, and the mean and the variance of the asymptotic distribution of $\hat{d}^T\hat{V}^{-1}\hat{d}$ should also be examined. For more information, about asymptotic distribution of statistics see Hausman (1978), Chesher (1983) and Davidson and Mackinnon (1992) which discussed a new different form of the information matrix test. Moreover, Stomberg and White (2000), provide considerable Monte Carlo evidence on the finite sample performance of several alternative forms of IMT. Our purpose in this paper is to develop the form of IMT statistic which is asymptotically distributed χ_R^2 distribution under H_0, when the model is correctly specified, and non- central $\chi_R^2\,(\lambda)$ distribution under H_1 , when the model is mis-specified, in this case $d \sim N(\mu, V)$, then $d^T V^{-1} d \sim \chi^2$ and

$$E(d^T V^{-1} d) = E(V^{-1} d d^T)$$

$$= E(V^{-1}(d-\mu)(d-\mu)^T) + \mu^T V^{-1} \mu = rank(V) + \mu^T V^{-1} \mu$$

Note that in this case χ^2 has mean $R + \lambda$ and variance $2(R + 2\lambda\,)$, where R is the rank of V and $\lambda\, = \mu^T V^{-1} \mu$. So, the main point is avoiding the singularity problem that discussed in previous sections, which is related with the log likelihood function. The basic idea is to consider a version of the IMT based on a reduced set of the elements of d . Therefore, we removed the elements which are related to the log likelihood function. To illustrate our idea let us consider an example as we discussed in previous if we have fitted the model with one covariate then, we have

$$E(d) = E_X \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} = E_X \begin{pmatrix} (\pi_t - \pi)(1-2\pi) \\ (\pi_t - \pi)(1-2\pi)x_i \\ (\pi_t - \pi)(1-2\pi)x_i^2 \end{pmatrix}$$

So, as we discussed we need to remove the elements d_1 and d_2 from d , and then we will use only just d_3 to compute the statistic. in this case $d = d_3$ and the statistic is $n d_3^2 V^{-1}$. This approach we call the IMT_{DE}, and we will evaluate the IMT_{DE} statistic by simulation to examine the behaviour of its asymptotic distribution.

Simulation Study

In this part of simulation, we are interested to examine the asymptotic distribution of IMT statistic in case when all the elements of (d) are used, and also we need to investigate the properties of the IMT_{ED} and how the reduced elements improve and the asymptotic distribution of the IMT_{DE} as chi-square distribution with mean [rank (V)] and variance [2 rank (V)], if the fitted model is correct. Also, we investigate the asymptotic distribution of IMT under mis-specified model to focus on the behaviour of the asymptotic distribution of IMT, which is in this case is distributed non central chi-square distribution with mean is [rank (V) + λ] and variance [2 rank (V)+4 λ] where $\lambda = E(d)^T V^{-1} E(d)$. Moreover, examine the effect of elements of variance matrix by likelihood function

Design of Simulation

This simulation designed to examine the asymptotic distribution of IM and IMT_{DE}, we will consider two cases of simulation under true model and under mis-specified model. If we have true logistic regression model with two covariates

$$\pi_i = expit(\alpha + \beta_1 x_{i1} + \beta_1 x_{i2}).$$

Firstly, we will focus on asymptotic distribution of IMT when the true model is fitted. Secondly, investigate the asymptotic distribution of IMT when the missing covariate logistic model has been fitted:

$$\pi_i = expit(\alpha + \beta_1 x_{i1}).$$

- We consider x_{i1} and x_{i2} as a draw from bivariate normal distribution X ∼ N_2 (0, Ω).
- We consider the 2 × 2 covariance matrix is.

$$\Omega = \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

- Use the variance $\sigma^2 = 0.2$ and ρ = 0.1.

- We choose different component of parameters under fitted true model as (α_t, β_t) : (0,1), (0.8, 0.5), (3.5,2.3).

- Under fitted missing covariates model, we choose different component of parameters α_t= (0,0.9,0.7), β_{t1}=(1,1.3,1.5) and β_{t2} =(0.6,1.2,2).

RESULTS AND DISCUSSION

Results and Discussion in Case of Correctly Specified Model:

In this simulation we consider to compute the IM T with two cases of dispersion matrix V and V_E to investigate the behavior of IM T and IMT_{DE} under effects of theoretical variance which computed by alternative formulae and empirical variance, and comparing the results. The results of simulation reported in several tables. These tables show the mean and the variance of IMT and IMT_{DE} by each found the theoretical and empirical variance. That is The IMTE denote to the statistic computed by empirical variance and IMTV denote to use theoretical variance,

$$IMTE = \bar{d}^T \widehat{var}(\bar{d})^{-1} \bar{d}$$

and

$$IMTV = \bar{d}^T var(\bar{d})^{-1} \bar{d}$$

where, $\hat{d} = (\hat{d}_1, \hat{d}_2, \hat{d}_3)^T$ and $d = (d_1, d_2, d_3)^T$ i.e. full matrix. Also, IMTE1 and IMTV1 denote to the statistic when, $\hat{d} = (\hat{d}_1, \hat{d}_2)^T$ and $d = (d_1, d_2)^T$, i.e reduced the first element. Finally, IMTE2 and IMTV2 denoted to the statistic when, $\hat{d} = (\hat{d}_1)^T$, and $d = (d_1)^T$, i.e. reduced the two first elements. Also, (α_t) and (β_{t1}) denote to the true parameters of πt and $S.D(\pi_t)$ is the standard deviation of π_t over the distribution of the covariates. Table 1 and Table 2, shows the results in case of sample size n = 500 and n = 5000 respectively. If we maintain the IMT is asymptotically distributed as χ^2_R distribution, with $df = R$ where, R is the rank of V , so, the statistics IMTE or IMTV should have mean R = 3 and variance 2R = 6, the statistics IMTE1 or IMTV1 has mean R = 2 and variance

Vol.9, No.1, pp.29-40, 2021
Print ISSN: 2053-2229 (Print), Online ISSN: 2053-2210 (Online)

2R = 4 and the last statistics, IMTE2 or IMTV2 have mean R = 1 and variance 2R = 2. Generally, we can see clearly, that the properties of χ2 distribution do not apply for both IMTE and IMTV for most sets of parameters and different sample sizes. The variance shows by far the more erratic behaviour. If we look at the second proposed statistic IMTE1 or IMEV1, the properties of χ2 still do not apply, but, the departures are less than problem for IMTE, IMTV and it is looks better. The final proposed statistic, which is our proposed IMT_{DE}, the new form of the IMT denoted in this simulation by IMTE2 and IMTV2, shows reasonable properties, the mean and the variance appeared very close to the properties of χ2 distribution across all cases.

If we consider the results by the sample size, we can see that, when the sample size is larger, the results appear much better. In case of sample size n = 500, IMTDE in some cases appeared slightly affected, especially when using the empirical variance, and for small values of the S.D of (π_t). If we make a comparison between the IMT, computed by empirical variance and theoretical variance, the results reported that, in large sample size n = 5000 have the same behaviour. Finally, we can say although there are slight effects in some cases related to the small value of $S.D(\pi_t)$, the new form of statistic IMTDE works well and has reasonable behaviour in most of the cases investigated. Moreover, we can say that the IMTDE statistic appeared to have an asymptotic χ2 distribution without strange behaviour, at least with request to the mean and variance.

Table 1: Simulation results of mean and variance of IMT, when the model is correctly specified and *df* = 3, 2, 1 related to the three cases of IMT respectively, with sample size n = 500.

α_t	β_{t1}	π_t	$S.D\pi_t$	-	IMTE	IMTE1	IMTE2	IMTV	IMTV1	IMTV2
0	1	0.48	0.07	**Mean**	3.57	2.29	1.15	1874.6	1.99	1.01
				var	8.96	5.91	2.96	3962969	4.813	2.42
0.8	0.5	0.48	0.07	**Mean**	2.95	2.18	1.12	18028.2	3.23	0.99
				var	7.69	5.06	2.85	6196039	32.24	2.11
			0.02	**Mean**	12.06	4.04	2.05	20.45	2.21	1.17

3.5	2.3	0.95		**var**	613.9	50.88	23.29	2581.01	10.06	4.14

Table 2: Simulation results of mean and variance of IMT, when the model is correctly specified and *df* = 3, 2, 1 related to the three cases of IMT respectively, with sample size n = 5000.

α_t	β_{t1}	π_t	$S.D\pi_t$	-	IMT E	IMT E1	IMT E2	IMTV	IMTV1	IMTV2
0	1	0.50	0.10	**Mean**	3.32	2.06	1.03	127.34	1.98	0.99
				var	8.54	4.48	2.27	104530	4.08	2.05
0.8	0.5	0.70	0.04	**Mean**	2.91	2.02	1.02	1626.3	2.17	1.04
				var	6.47	4.12	2.12	2505339	5.82	2.18
3.5	2.3	0.95	0.05	**Mean**	3.05	2.16	1.11	5.07	2.11	1.04
				var	13.55	5.30	2.91	56.48	5.10	2.28

Results and Discussion in Case of Mis-specified Model:

Now, we will discuss the results under H_1, when the model is mis-specified. We used the same assumptions which we discussed in previous section, but in this case $\beta_{t2} \neq 0$, and we choose different cases of parameters (β_{t2} = 0.6, 1.2, 2). Table 3 and Table 4 shows the results in two case of sample size n = 500 and n = 5000 respectively.

We can see that, the tables reported, the IMTV and IMTE more strange behaviour due to overlap with elements of the log likelihood function and appeared are far away from the properties of χ2 distribution. The second proposed test, IMTV1 and IMT E1 appeared better with slightly effect especially with IMT E1, which is appeared more sensitive. The same behaviour which we found in case of the model under H_0 , our proposed method IMT_{DE} appeared more stable. However, the assumption that its distribution closely follows a non-central χ2 is not well supported.

Table 3: Simulation results of mean and variance of IMT, when the model is mis-specified and *df* = 3, 2, 1 related to the three cases of IMT respectively, with sample size n = 500.

α_t	β_{t1}	β_{t2}	π_t	$S.D\pi_t$	-	IMT E	IMT E1	IMT E2	IMTV	IMTV1	IMTV2
0	1	0.6	0.80	0.07	**Mean**	4.72	2.49	1.17	84.71	2.01	0.973
					var	29.73	8.55	3.37	5223	5.63	1.96
0.8	1.3	1.2	0.94	0.05	**Mean**	10.94	3.82	1.79	35.70	2.17	0.96
					var	424.1	39.46	12.66	1102	11.31	2.00
0.7	1.5	2	0.97	0.04	**Mean**	18.85	5.18	2.36	23.96	2.56	0.98
					var	1905	115	26.32	3322	18.85	2.30

Table 4: Simulation results of mean and variance of IMT, when the model is mis-specified and *df* = 3, 2, 1 related to the three cases of IMT respectively, with sample size n = 5000.

α_t	β_{t1}	β_{t2}	π_t	$S.D\pi_t$	-	IMT E	IMT E1	IMT E2	IMTV	IMTV1	IMTV2
0	1	0.6	0.50	0.10	**Mean**	3.30	2.07	1.02	11.57	2.05	1.01
					var	8.90	4.45	2.12	618.9	4.35	2.02
0.8	0.5	1.2	0.70	0.04	**Mean**	3.67	2.02	1.12	6.14	1.83	0.99
					var	14.37	5.52	2.88	131.4	3.74	1.93
3.5	2.3	2	0.95	0.05	**Mean**	3.98	2.34	1.22	6.09	2.19	1.07
					var	20.10	6.68	3.65	98.12	5.68	2.18

CONCLUSION

The goal of the present work in this paper is to examine the singularity problem of IMT and investigate the distribution under correct and missing covariate logistic model. New form of the information matrix test IMT_{DE} was examine by simulation which reduced the elements of d to remove overlap with elements of the log-likelihood function. In fact, although there is slightly different results when using the empirical covariance matrix with sample size n = 500, the IMT_{DE} appeared reasonable asymptotic distribution behaviour in large sample size n = 5000 and the properties very close to the χ2 -distribution under H_0. However, the form of the distribution under H_1 is less clear. According to these results, it would be helpful to try an alternative approach**.**

REFERENCES

Badi, N.H.S. (2017). *Asymptomatic Distribution of Goodness-of-Fit Tests in Logistic Regression Mode*. Open Journal of Statistics, **7**, 434-445.

Badi, N.H.S. (2021). *The Behaviour of the Dispersion Matrix of the Information Matrix Test under Wrong Logistic Regression Model*. Open Access Library Journal, **8**: e7153.

Chesher, A. (1983). *The information matrix test: Simplified calculation via a score test interpretation.* Economics Letters, **13**, 45-48.

Chesher, A. (1984). *Testing for Neglected Heterogeneity*. Econometrica, **21**, 3789- 3801.

Davidson, R. and Mackinnon, J. G. (1984). *Convenient Specification Tests for Logit and Probit Model* Journal of Econometrics, **25**, 241-262.

Davidson, R. and Mackinnon, J. G. (1988). *The Calculation of the Information Matrix Test for Binary Data Models.* The Manchester school, **54**, 370-376.

Davidson, R. and Mackinnon, J. G (1992). *Anew Form of the Information Matrix Test*. Econometrica, **60**,145-157.

Hausman ,J. A. (1978). *Specification Tests in Econometrics*. The Econometrix So- ciety, **46**, 1251-1271.

Kuss, O. (2002). *Global goodness-of-fit tests in logistic regression with sparse data*. Statistics in Medicine, **31**, 459-468.

Matthews, J. N. S. and Badi, N. H. (2015). *Inconsistent treatment estimates from mis-specified logistic regression analyses of randomized trials*. Statistics in Medicine, **34**:2681-2694.

Lancaster, T. (1984). *Covariance Matrix of the Information Matrix Test*. Econometrica, **4**, 1051-1053.

Newey, W. K. (1985). *Maximum Likelihood Specification Testing and Conditional Moment Test*. Econometrica, **53**, 1047-1070.

Stomberg, C. and White, H. (2000). *Bootstrapping The information matrix test.* University of California, San Diego, https://escholarship.org/uc/item/***15845***1cr.

White, H. (1982). *Maximum Likelihood Estimation of Misspecified Models*. Econometrica, **50**, 1-25.