

Comparison of Probit and Logistic Regression Models in the Analysis of Dichotomous Outcomes

¹Amrutha Jose, ¹Mariyamma Philip, ²Lavanya Tumkur Prasanna and ³M. Manjula

¹Department of Biostatistics, National Institute of Mental Health and Neurosciences (NIMHANS), Bengaluru, India

²Lund University, Sweden

³Department of Clinical Psychology, National Institute of Mental Health and Neurosciences (NIMHANS), Bengaluru, India

Article history

Received: 29-01-2020

Revised: 23-03-2020

Accepted: 25-03-2020

Corresponding Author:

Amrutha Jose

National Institute of Mental

Health and Neurosciences

(NIMHANS),

Bengaluru, India

Email: amruthaannjose@gmail.com

Abstract: Probit and logistic regression models are members of the family of generalized linear models, used for estimating the functional relationship between the dichotomous dependent and independent variables. The current study is designed to find the performance of logistic and probit regression models in different conditions under multivariate normality. The objective of the study is to compare the performance of probit and logistic regression models under multivariate normal. A Monte Carlo simulation study was done in which artificial datasets were generated under multivariate normality. Datasets were generated by employing the latent variable approach, under different variance-covariance matrices, varying sample sizes and prevalences. For each of the combinations, 1000 simulations were carried out. Probit and logistic regression analyses were performed and compared using parameter estimates, standard error, Likelihood Ratio test, RMSEs, null and residual deviances, different pseudo R² measures, AIC, BIC and Correct Percent Prediction. A live data set was also used to compare the efficiency of the models. It was evident from AIC, BIC and RMSE values that logit and probit models fit the dataset equally well in all the combinations of sample size, correlation structure and proportion of outcome. However, sensitivity, specificity and CPP values showed that the logit model predicts the outcome better than the probit model in most of the situations. The results showed that the probit and logit models perform equally well under multivariate normality.

Keywords: Logistic Regression, Probit Regression, Monte Carlo Simulation, Self-Efficacy, Children

Introduction

Most of the outcome variables in Biomedical research are categorical in nature. Many times continuous variables will be categorized and is known as discretization. Even though studies indicate that dichotomization of the variables is a bad practice and results in loss of statistical power (Altman and Royston, 2006; Naggara *et al.*, 2011), it is quite a common practice. For example, the Hamilton Rating Scale for Depression - 17 (HRSD-17), contains 17 items to be rated (Hamilton, 1960). It is used to diagnose whether a subject is having depression or not. The items in the questionnaire are scored on 3 point to 5 point scale. The subject is considered as normal if the HRSD score is between 0 and 7 and there is evidence of depression otherwise

(Hamilton, 1960). In such a scenario where the outcome variable is dichotomous, binary logistic and probit regression models are the frequently used statistical methods for predicting the outcome variable based on a set of independent variables (Chai and Draxler, 2014).

The underlying dependent variable, Y^* in binary logit and probit models can be defined as:

$$Y^* = \sum_{k=1}^K \beta_k x_k + \varepsilon \quad (1)$$

where, K is the number of parameters involved in the model. In practice, Y^* is an unobserved variable ranging from $-\infty$ to $+\infty$ which makes the observed Y , a dichotomous outcome variable (Cakmakyan and Goktas, 2013; Alsoruj *et al.*, 2018).

Logit and probit models are members of the family of Generalized Linear Models (GLM) and are commonly used to predict the categorical dependent variable based on a number of covariates or independent variables using the link functions - logit and probit respectively. Both models have been addressed in the literature and are used for the same purpose. However, studies are seldom available evaluating the efficiency of the models under different parameter estimates such as varying sample sizes and different proportions (prevalence) of outcome with a range of correlations between dependent and independent variables, which will provide evidence to propose appropriate methods under various situations.

In this paper, an effort is made to summarize the methodology of the binary logit model, the binary probit model, fit indices for both models with an application to a live data situation. A Monte Carlo simulation study was also conducted in which artificial datasets were generated under multivariate normality using a latent variable approach. Various conditions were imposed in terms of variance-covariance matrices assuming specific correlations between dependent and independent variables. Datasets were generated under different sample sizes as well as varying prevalences. For each of the combinations, 1000 simulations were carried out and the results of the simulation study were also documented.

Binary Outcome Models – Logit and Probit

Suppose that there is a variable y^* ranging from $-\infty$ to $+\infty$ that generates the observed outcome variable, y . y^* is assumed to be linearly related to the observed x 's through the structural model:

$$y_i^* = x_i \beta + \varepsilon_i \quad (2)$$

The variable y^* is linked to the observed binary variable y by the equation:

$$y_i^* = \begin{cases} 1, & y_i^* > \tau \\ 0, & y_i^* \leq \tau \end{cases} \quad (3)$$

where, τ is the threshold or cut point value.

If y_i^* crosses the threshold τ (i.e., $y_i^* > \tau$), then $y=1$, otherwise $y=0$. The link between y^* and the observed y is illustrated in figure 1 for the model $y^* = \alpha + \beta x + \varepsilon$.

In figure 1, y^* is taken along the vertical axis, with the threshold τ indicated by a horizontal dashed line. The distribution of y^* is shown by the bell-shaped curves which should be thought of as coming out of the figure into a third dimension (Long, 1997; Cakmakyan and Goktas, 2013).

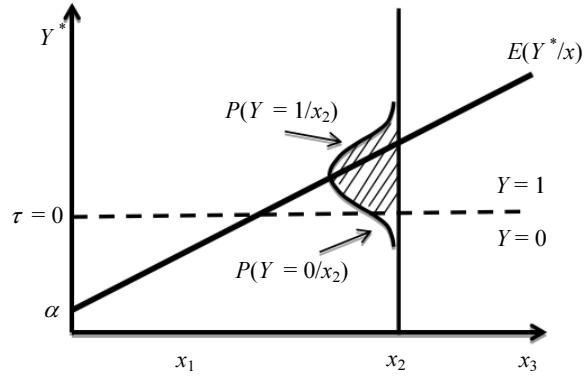


Fig. 1: The relationship between latent variable Y^* and the probability of observed values in the binary response models

Since the outcome variable is dichotomous, linear regression methods that use ordinary least squares for estimation cannot be applied. Maximum likelihood estimation which assumes some specific distribution of the errors should be used, instead. The most commonly used statistical models for predicting dichotomous outcome variable are the logit model which assumes logistic errors and probit model which assumes normal errors (Long, 1997).

In both probit and logit models, the measurement equation is $Y = X\beta + \varepsilon$. The logit models use the standard logistic probability distribution function and the event probability can be estimated as:

$$E(Y = 1 | x) = \pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \quad (4)$$

The link function in the logit model is:

$$\ln \left[\frac{\pi(x)}{1 - \pi(x)} \right] = X\beta \quad (5)$$

The error distribution in the logit model is assumed to be a standard logistic distribution with mean, $E(\varepsilon/x) = 0$ and variance, $\text{Var}(\varepsilon/x) = \pi^2/3 \approx 3.29$ (Long, 1997). Similarly, the event probability in the probit model can be estimated as:

$$E(Y = 1 | x) = \pi(x) = \Phi(X'\beta)$$

where, Φ is the cumulative standard normal distribution. The link function in the probit model is $\Phi^{-1}(\pi(x))$ and the error distribution is assumed to be normally distributed with $E(\varepsilon/x) = 0$ and $\text{Var}(\varepsilon/x) = 1$.

Measures of Model Fit

Measures of model fit are used to compare the competing models and to select a final model. In the linear regression model, the coefficient of determination (R^2) is the standard measure of fit. When the outcome variable is dichotomous, there are numerous fit indices such as pseudo R^2 but none of these measures gives a clear interpretation in terms of explained variation (Long, 1997). Another way of selecting the final model is by using the Likelihood Ratio (LR) test where it is tested that all of the slope coefficients are simultaneously equal to zero or not. Information criteria measures such as Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) are also popular for comparing models especially when the models are non-nested. Root Mean Square Errors (RMSEs), Null and Residual deviances, Sensitivity, Specificity and Accuracy or Correct Percentage Prediction (CPP) values are also commonly used measures to compare binary outcome models.

Pseudo- R^2 Measures

In logistic regression, there are a number of pseudo- R^2 measures and are used to indicate how well a model fits a set of data. But there is no consensus on which measure is the best. Also, different pseudo- R^2 measures have different values for the same model and most of them are not seen as a number between 0 and 1. The pseudo- R^2 measures considered in the present study were as follows:

- R^2_{McF} (McFadden's Pseudo- R^2): It is also known as Likelihood Ratio Index and is defined as:

$$R^2_{McF} = 1 - \frac{\ln \hat{L}(M_\beta)}{\ln \hat{L}(M_\alpha)} \quad (6)$$

where, M_α be the model with just the intercept, M_β be the model with the regressors included, $\ln \hat{L}(M_\alpha)$ is the log-likelihood of M_α which is analogous to the total sum of squares and $\ln \hat{L}(M_\beta)$ is the log-likelihood of M_β (analogous to the residual sum of squares) (McFadden, 1973). R^2_{McF} can never exactly equal to 1.

- R^2_{McFAdj} (McFadden's Adjusted Pseudo- R^2): It penalizes a model for including too many predictors and is defined as:

$$R^2_{McFAdj} = 1 - \frac{\ln \hat{L}(M_\beta) - k}{\ln \hat{L}(M_\alpha)} \quad (7)$$

where, k is the number of predictors in the model.

- R^2_{ML} (Maximum Likelihood pseudo- R^2 or Cox-Snell R^2): It is also called the geometric mean squared improvement per observation and is defined as:

$$R^2_{ML} = 1 - \left[\frac{L(M_\beta)}{L(M_\alpha)} \right]^{2/N} \quad (8)$$

- $R^2_{C&U}$ (Cragg and Uhler's pseudo- R^2 or Nagelkerke R^2): As the fit of M_β approaches M_α , R^2_{ML} approaches 0. Maddala (1986) showed that R^2_{ML} only reaches a maximum of $1 - [L(M_\alpha)]^{2/N}$

Cragg and Uhler (1970) suggest the normed measure:

$$R^2_{C&U} = \frac{R^2_{ML}}{\max R^2_{ML}} = \frac{1 - [L(M_\alpha) / L(M_\beta)]^{2/N}}{1 - [L(M_\alpha)]^{2/N}} \quad (9)$$

$R^2_{C&U}$ is an adjustment for R^2_{ML} , which makes it possible for the R^2 to have a maximum value of 1.

R^2_{McF} , R^2_{McFAdj} , R^2_{ML} and $R^2_{C&U}$ are defined in terms of the likelihood function and can be applied to any model estimated by ML (Long, 1997). All else being equal, models with a larger value of the log-likelihood are preferred and R^2_{McF} provides a convenient way to compare log likelihoods across different models (Long, 1997).

Scaled Deviance

Scaled deviance or simply deviance is a fit index which is analogous to the residual sum of squares within the framework known as the generalized linear model. The deviance compares a given model to the full model M_F . The full model has one parameter for each observation and can reproduce perfectly the observed data. Since the observed data are perfectly predicted, the likelihood of M_F is 1 and the log-likelihood is 0 (Long, 1997).

To test that M_F significantly improves the fit over M_β , then the deviance is defined as:

$$D(M_\beta) = 2 \ln L(M_F) - 2 \ln L(M_\beta) = -2 \ln L(M_\beta) \quad (10)$$

The deviance function is very useful for comparing two models when one model has parameters that are a subset of the second model. The null deviance shows how well the response variable is predicted by a model that includes only the intercept whereas residual deviance shows how well the response variable is predicted by the model with the inclusion of independent variables.

Likelihood Ratio (LR) Test

LR test is used to test the hypothesis that all of the slope coefficients are simultaneously equal to zero and its test statistic is defined as:

$$LR = 2 \ln L(M_\beta) - 2 \ln L(M_\alpha) \quad (11)$$

Under H_0 , LR follows Chi-square distribution with degrees of freedom equal to the difference between the numbers of parameters in the two models (Long, 1997).

AIC

Another way to assess the performance of a model is by calculating fit indices such as AIC and BIC. AIC is a way of selecting a model from a set of models and is defined as:

$$AIC = -2 \ln L(M_\beta) + 2k \quad (12)$$

where, $L(M_\beta)$ is the likelihood of the model and ' k ' is the number of parameters in the model. The term $-2 \ln L(M_\beta)$ ranges from 0 to $+\infty$ with smaller values indicating a better fit. As the number of parameters in the model increases, $-2 \ln L(M_\beta)$ becomes smaller since more parameters make what is observed more likely. Hence, the constant $2k$ is added to $-2 \ln L(M_\beta)$ as a penalty. A smaller AIC value indicates a better fit. i.e., the chosen model is the one that minimizes the Kullback-Leibler distance between the model and the truth (Burnham and Anderson, 2003). AIC is often used to compare models across different samples or to compare non-nested models that cannot be compared with the LR test.

Corrected AIC (AICc)

AIC can severely violate the principle of parsimony in extreme circumstances. The failure of AIC to select an adequately parsimonious model can be a problem when the number of parameters in the model under consideration is more than (roughly) 30% of the sample size. It could be argued that a good model selection criterion should work even if the researcher tries a 'bad' (e.g., over-parameterized) model. In order to overcome this deficiency, Hurvich and Tsai (1989) introduced a corrected version of AIC, which includes a correction for small sample sizes as:

$$AIC_c = AIC + \left[\frac{2k(k+1)}{n-k-1} \right] \quad (13)$$

where, 'AIC' is the standard AIC, ' k ' is the number of parameters in the model and ' n ' is the number of observations (Hurvich and Tsai, 1989).

BIC and sBIC

It is also called Schwartz Information Criteria (SIC) or Schwarz BIC (sBIC). It is proposed by Raftery (1996) as a measure to assess the overall fit of a model to allow comparison of both nested and non-nested models (Raftery, 1996). BIC can be defined as:

$$BIC = -2 \ln L(M_\beta) + k \cdot \ln(n)$$

where, ' k ' is the number of regressors and ' n ' is the sample size. Sample size adjusted BIC (sBIC) can be computed as:

$$sBIC = -2 \ln L(M_\beta) + 4 \cdot \log\left(\frac{n+2}{24}\right) \quad (14)$$

and which can be used as a better measure to evaluate the model fit between the models. Similar to AIC, smaller BIC value indicates better fit and the difference in the BICs from two models indicates which model is more likely to have generated the observed data.

Root Mean Square Error (RMSE)

The RMSE is frequently used as a standard statistical metric to measure model performance by calculating the differences between predicted and observed values. These individual differences are also called residuals and the RMSE serves to aggregate them into a single measure of predictive power. It can be defined as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (15)$$

where, y_i and \hat{y}_i are observed and predicted values respectively. i.e., RMSE is the square root of the mean squared errors which make an excellent general purpose error metric for numerical predictions.

Sensitivity, Specificity and Correct Percent Prediction (CPP/Count R²)

In order to check the model adequacy, various evaluation measures based on estimated probabilities such as accuracy, precision (positive predictive value), recall (sensitivity), specificity and negative predictive value can be used. In the aforementioned evaluation measures, an aspect of the model's effectiveness in assigning each observation to the correct categories is measured.

AUC

A Receiver Operating Characteristic (ROC) curve is a standard technique of indicating the performance of a classification model by plotting sensitivity against 1-specificity for many possible thresholds. The higher the Area Under the ROC Curve (AUC) indicates a better discrimination power of the model. AUC is also referred to as the Index of Accuracy or Concordance Index.

Residuals

There are many types of residuals such as ordinary residual, Pearson residual and Studentized residual. They all reflect the differences between fitted and observed values and are the basis of varieties of diagnostic methods (Zhang, 2016). The default residual for the generalized linear model is Pearson residual.

Live dataset

In order to compare logit and probit models, data from the archives of the Department of Biostatistics, NIMHANS, Bengaluru has been used. The study is a descriptive study which involved 175 adolescents pursuing their high school (8th to 10th standards) or pre-university courses (I and II PUC) from schools and colleges in Bengaluru. The sampling technique used was a two-stage sampling procedure. In the first stage, schools/colleges were randomly selected from the list of schools/colleges that gave permission for the study. From the selected schools and colleges, classes were chosen randomly from different sections available and the students were chosen based on the cluster sampling method. The whole class chosen randomly was included in the study and those students, who gave assent and obtained consent from parents, were finally form the sample. The inclusion criteria were: (1) students pursuing high school and pre-university courses, (2) students having knowledge of English and (3) students belonging to Indian nationality. The questionnaires had been administered in groups. One whole class was taken at once for the assessment.

Description of Tools

Socio-demographic datasheet: It included information about the socio-demographic characteristics, such as age, gender, birth date, place of living, family type, siblings, education, stream (for PUC students) and help sought with any mental health professional.

Difficulties in Emotion Regulation Questionnaire (Gratz and Roemer, 2004): DERS is a brief 36 item questionnaire consists of 6 subscales. Two subscales viz. - impulsivity and level of emotional awareness were used for the study.

Cognitive Emotional Regulation Questionnaire (CERQ): The CERQ is a 36-item questionnaire consisting of nine subscales (Garnefski et al., 2002). CERQ positive contains positive refocusing, planning and positive reappraisal and putting into perspective while CERQ negative contains self-blame, other blame, rumination and catastrophizing. CERQ has good factorial validity and high reliabilities can be used for both adolescents and young adults (anyone above the age of 12 years).

BAR Emotional Quotient Inventory Youth Version: The scale has been considered as one of the best measures of emotional competence among youth (Bar-On and Parker, 2000). It provides a combined social-emotional and personality attribute assessment of self-reported emotion-related functioning. The five subscales are interpersonal, intrapersonal, stress management scale, adaptability scale and general mood scale.

Distress Tolerance Scale (DTS): This scale measures the participants' perceived ability to experience and endure negative emotional states. The scale consists of 4 subdomains namely, appraisal,

tolerance, absorption and regulation (Sinha et al., 2007). Higher scores on DTS indicate greater positive affect, less affective distress and liability.

Interpersonal Reactivity Index: This is a 28 item, five-point Likert scale consisting of response categories ranging from "not at all like me" to "very much like me". This scale comprises of 4 subscales, namely, a) Empathic concern scale b) Perspective-taking scale c) personal distress and d) fantasy/imaginal involvement, are not of concern to the current study.

Self-Efficacy questionnaire - Children: This scale contains, 24 items, representative of 3 domains, namely: (a) Academic self-efficacy (b) Emotional self-efficacy and (c) Social self-efficacy (Muris, 2001). This is a 5 point Likert scale having scores ranging from "not at all" to "very well". This scale is considered as a psychometrically sound measure of self-efficacy among adolescents.

Scale for Assessing Academic Stress (SAAS): It is a 30 item self-report measure developed for grade 8 to 12 standard students of English medium schools with students belonging to middle and higher socioeconomic status (Sinha et al., 2007). It assesses five major indicators of academic stress (cognitive, affective, physical, social/interpersonal and motivation) which would finally yield a total stress score.

Statistical Analysis

The variable, self-efficacy score among the study subjects was normally distributed. It was categorized into a dichotomous variable by considering the probability density function of the variable as well as by seeking the expert's opinion. The stepwise logistic and probit regression analyses were conducted to predict the self-efficacy among 175 adolescents using the type of education (High school/PUC), appraisal, Distress Tolerance Scale (DTS) score, empathy concern, perspective-taking, school stress (in terms of cognitive, affective and social as well as motivation), CERQ positive and emotional intelligence as predictors which were significantly contributing to the prediction of self-efficacy in the univariate models.

Results and Discussion

The predictors such as school/college, appraisal, empathy concern, perspective-taking, motivation and CERQ positive were finally included in both logit and probit models. Table 1 shows the comparison of logit and probit models for the current dataset in terms of coefficient estimates, its standard errors and significance, null and residual deviances, information criteria measures (such as AIC and BIC), RMSE, sensitivity, specificity and correct percent prediction, LR test statistic and its p-values, pseudo-R² measures (such as R²_{McF}, R²_{ML} and R²_{C&U}) and also in terms of marginal effects.

Table 1: Output of live dataset

Parameters	Logit model			Probit model		
	Estimate	SE	p-value	Estimate	SE	p-value
Final model:						
Intercept	9.542	2.573	<0.001	5.560	1.464	<0.001
School/College	-1.155	0.391	0.003	-0.718	0.230	0.001
Appraisal	-0.076	0.047	0.104	-0.046	0.028	0.102
Empathy concern	-0.129	0.054	0.016	-0.070	0.032	0.026
Perspective-taking	-0.151	0.060	0.012	-0.089	0.035	0.012
Motivation	0.192	0.075	0.010	0.110	0.044	0.012
CERQ Positive	-0.063	0.024	0.008	-0.037	0.014	0.007
Null deviance	235.55			235.55		
Residual deviance	195.76			196.10		
AIC	209.76			210.10		
BIC	231.91			232.26		
RMSE	0.435			0.436		
Sensitivity	81.91			81.91		
Specificity	50.00			47.14		
Correct Percent Prediction	69.14			68.00		
LR statistic	39.793			39.449		
(p-value)	(<0.001)			(<0.001)		
Pseudo R²						
McFadden	0.169			0.167		
ML	0.203			0.202		
C&U	0.275			0.273		
Marginal Effects:						
Intercept	1.816			1.772		
School/College	-0.220			-0.229		
Appraisal	-0.014			-0.015		
Empathy concern	-0.025			-0.022		
Perspective-taking	-0.029			-0.028		
Motivation	0.037			0.035		
CERQ Positive	-0.012			-0.012		

The null and residual deviances of logit model were 235.55 and 195.76 respectively while those of probit models were 235.55 and 196.10 respectively. The AIC and BIC values also were almost same in logit (AIC_{logit} , BIC_{logit} : 209.76, 231.91) and probit (AIC_{probit} , BIC_{probit} : 210.10, 232.26) models indicating equal fit of both the models to the data set. The RMSE values were also similar for both the models ($RMSE_{logit}$: 0.435, $RMSE_{probit}$: 0.436) which showed that both the models are equally fitting the data. The specificity and correct percent prediction in logit models were 50.00%, 69.14% respectively and that of probit models were 47.14%, 68.00% respectively showed a better fit to logit model even though the sensitivity of the models was the same (81.91%).

The LR test and the pseudo-R² measures were also showed more preference to logit model than probit model in the current data set. The marginal effects were also computed and were almost similar in both the models indicating equal fit of both the models to the current data set.

Simulation Study

A computer simulation is an endeavour to establish a reality or theoretical circumstance on a computer

which can be utilized to contemplate and perceive how the framework functions and verifies the models and assumptions in the data. By changing various parameters in the simulation, forecasts might be made about the characteristics of the framework by recapturing the parameters used to generate the dataset. Hence, a simulation study is designed to find if there exists any precedence or difference between logistic and probit regression models in fitting under certain situations. The objective of the study was to compare the performance of probit and logistic regression models under multivariate normality.

Methodology

The study was a simulation study in which the data were generated artificially under the assumption of multivariate normality. The models used in the study are Generalized Linear Models especially binary logistic and probit regression models. The efficiency of both the models was tested under various situations namely different correlation structures between the dependent variable and independent variables, varying sample sizes and different proportions/prevalences of the outcome. For that, artificial datasets of a dependent and three-independent variables were randomly

generated using a multivariate normal distribution with an appropriate mean vector, $\mu = (12, 10, 0, 0)$.

Various correlation structures of High ($R^2_{ols} = 0.90$), Moderate ($R^2_{ols} = 0.40$), Low ($R^2_{ols} = 0.15$) and No ($R^2_{ols} = 0.01$) were imposed in terms of different variance-covariance matrices assuming specific bivariate correlation between dependent and independent variables. The different variance-covariance matrices used to generate the specified correlation structures in the dataset were:

$$\sum_{High} = \begin{bmatrix} 25 & 9 & 2.5 & -3.5 \\ 9 & 16 & 0 & 0 \\ 2.5 & 0 & 1 & 0 \\ -3.5 & 0 & 0 & 1 \end{bmatrix}$$

$$\sum_{Moderate} = \begin{bmatrix} 25 & 8 & 1.5 & -1.5 \\ 8 & 16 & 0 & 0 \\ 1.5 & 0 & 1 & 0 \\ -1.5 & 0 & 0 & 1 \end{bmatrix}$$

$$\sum_{Low} = \begin{bmatrix} 25 & 3 & 1 & -0.5 \\ 3 & 16 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ -0.5 & 0 & 0 & 1 \end{bmatrix}$$

$$\sum_{No} = \begin{bmatrix} 25 & 0.2 & 0.01 & -0.005 \\ 0.2 & 16 & 0 & 0 \\ 0.01 & 0 & 1 & 0 \\ -0.005 & 0 & 0 & 1 \end{bmatrix}$$

These matrices were determined in such a way that they were positive definite and correlations between independent variables were zeros in order to avoid the multicollinearity problem in the datasets. Covariance values between independent and dependent variables were appropriately chosen to produce the pre-specified correlation among them. The covariance between the variables would be identical to their correlations if the data is generated under the multivariate standard

normal distribution (Cakmakyan and Goktas, 2013; Alsoraji et al., 2018).

After the data generation, two of the independent variables were converted into categorical variables using an appropriate cut-off value ($p = 0.4$; i.e., 40% of the subjects those are having high values in the variable are classified into exposed group keeping in mind that the baseline distribution is normal). The dependent variable was then transformed into a dichotomous variable for 17 different proportions of the outcome variable, i.e., prevalences of 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85% and 90% assuming that the underlying distribution is normal. In order to examine the effect of sample size, the data were generated under 10 different sample sizes: 20, 30, 50, 100, 150, 200, 500, 1000, 2000 and 5000. Thus, a total of 680 different combinations were considered in the study and the summary of the data generation process is given in Table 2. The data generation process was repeated 1000 times for each of the combinations which resulted in a total of 6,80,000 datasets.

Probit and logistic regression analyses were performed and the parameter, standard error and probability estimates were obtained. In order to compare the models, different estimation and tests of significance procedures were done. The performances of the models were computed and compared using AIC, corrected AIC (AICc), BIC and sBIC statistics with the assistance of Likelihood Ratio (LR) test. The null and residual deviances and different pseudo R^2 measures (R^2_{McF} , R^2_{McFAdj} , R^2_{ML} and $R^2_{C&U}$) were computed and the models were compared. In addition to that, RMSEs in the estimation, as well as Correct Percentage Prediction (CPP) were also used to compare the performance of the regression methods. Student's t-tests were used to test whether there is any statistically significant difference between logit and probit models in terms of goodness-of-fit under various situations. Also, R^2_{ols} 's were calculated from linear regression models for the dependent and independent variables in order to cross-check whether the pre-specified correlation structure exists in the dataset.

Table 2: Data generation and Classification in the Simulation study
 Dataset ($4 \times 10 \times 17 = 680$ combinations)

Variance-Covariance matrix (4 levels)	Sample size (10 levels)	Proportions (in %) (17 levels)	Replications
High	20, 30, 50,	10, 15, 20, 25,	1000
Moderate	100, 150,	30, 35, 40, 45,	
Low	200, 500,	50, 55, 60, 65,	
No	1000, 2000, 5000	70, 75, 80, 85, 90	

Simulation Study Results

Figure 2 gives the AIC values of probit and logit models. The results of all the 680 combinations are available. In order to accommodate the information in the table, only 9 different prevalence measures (i.e., 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% and 90%) out of 17 and 6 different sample sizes (i.e., 50, 100, 150, 200, 500 and 1000) out of 10 are displayed. Both probit and logit models showed higher AIC values when the prevalence is 0.5 and decreased towards both lower and higher prevalence. For a higher prevalence rate with smaller sample size, the logit model fitted

better than the probit model even though the underlying distribution is normal. This scenario can be seen with all the correlation structures.

It can be seen from figure 2 that as the degree of correlation between DV and IVs increases, AIC values show a decreasing trend indicating better fit for models of having higher correlation levels. But, it is evident that both the models fitted equally well in all the situations. Figure 3 gives an idea about the performance of the model based on sample size in terms of AICc values. It is evident from the study results that both the models perform equally well with respect to varying sample sizes.

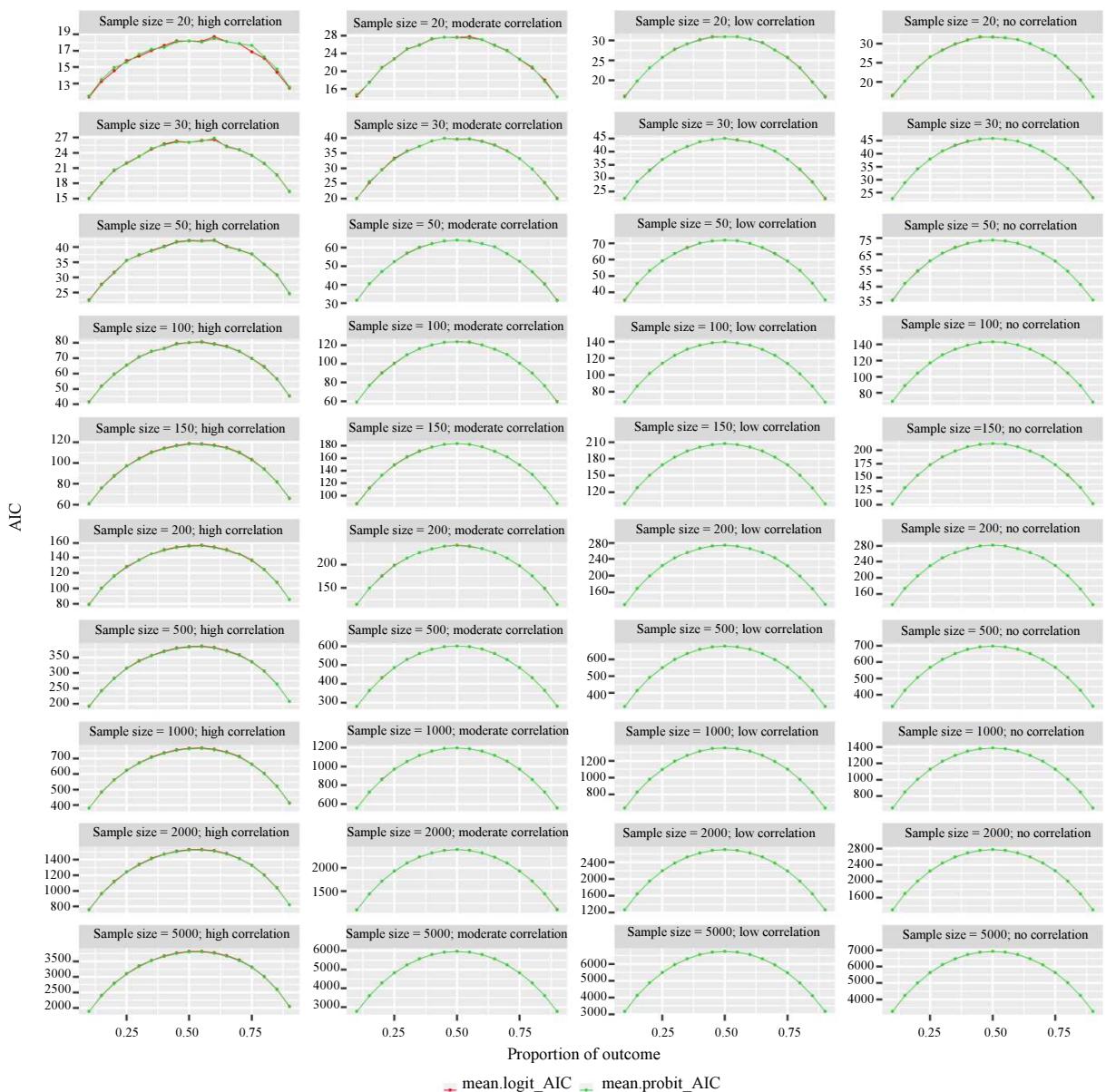


Fig 2: Comparison of Logit and Probit models using AIC values

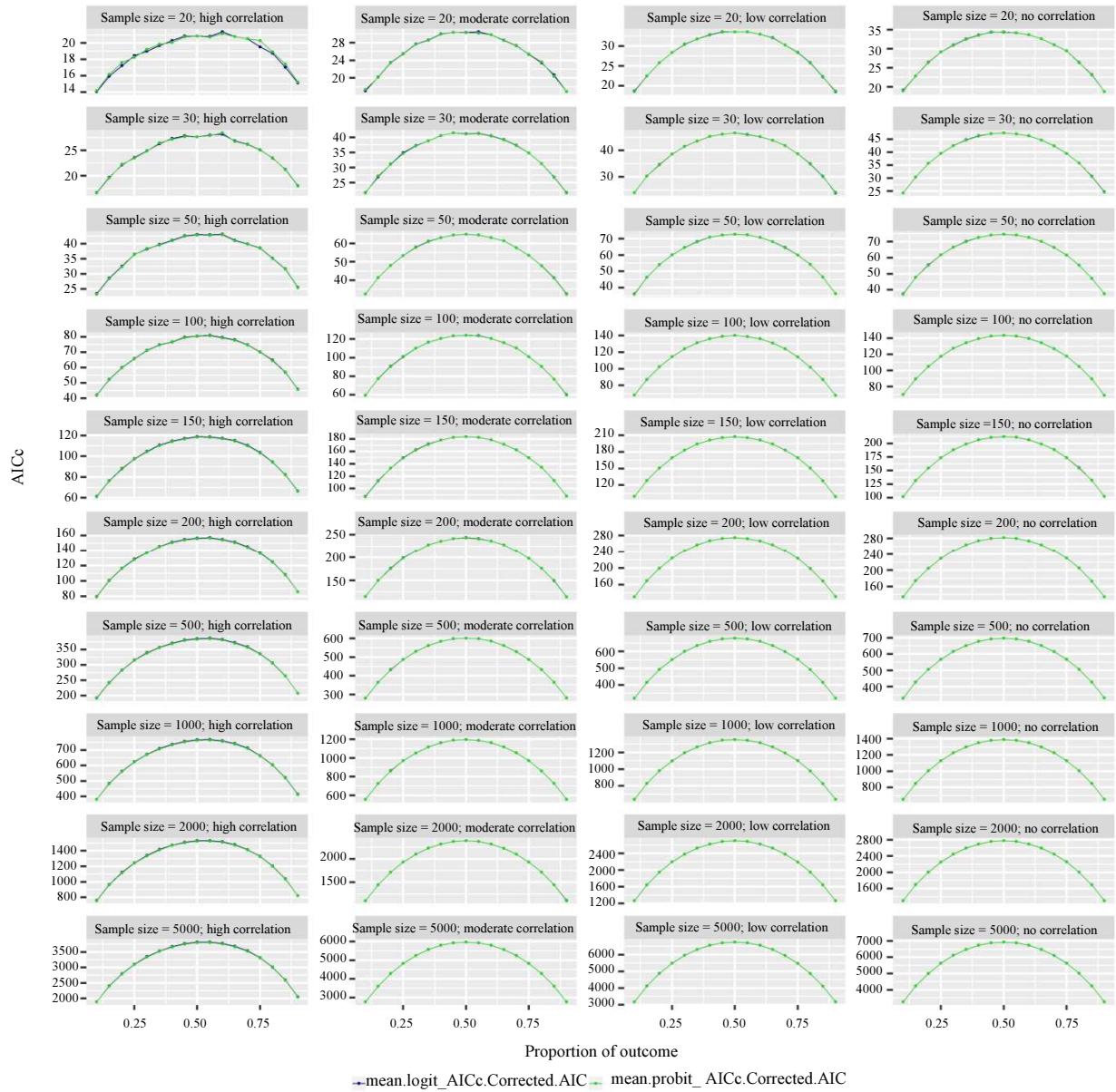


Fig. 3: Comparison of Logit and Probit models using AICc values

The estimates of the models i.e., intercept (B_0) and the regression coefficients (B_1 , B_2 and B_3) are compared in Tables 3 to 6 respectively. As the logit and probit models can't be directly comparable in terms of their original estimates, the marginal estimates of the coefficients are computed and are displayed in Tables 7 to 10. The study results from Tables 11 and 12 showed that for smaller sample sizes, prediction of outcome in terms of both sensitivity and specificity respectively is higher in logit model compared to probit, even though the underlying distribution is

normal. Also, it is clear that the logit model fits better in most of the situations especially when the prevalence is closer to 50%. It is evident from the results in Table 13 that the logit model has higher overall prediction (in terms of accuracy) than probit model in most of the situations. Please see Tables 3 to 13 which are included as the supplementary material.

BIC values and Likelihood ratio test p-values in figures 4 and 5 respectively indicated that both models perform equally well in all the situations. The LR test showed higher test statistic values at the prevalence of

50% and decreased towards both the extremes. The p-values showed higher levels of significance under the situations of higher sample size and also when the prevalence was closer to 50%. As the correlation increases, the LR test showed higher test statistic values with smaller p-values indicating a good fit.

It is evident from Table 14 that the RMSEs were similar for both logistic and probit regression models under most of the situations except for smaller sample sizes. For smaller samples, the logit model showed a

better fit to the data than the probit model. The comparison of both the models in terms of discrimination power of the model (AUC) is included in Table 15. The models are compared in terms of Pseudo R² values such as R²_{McF}, R²_{McFAdj}, R²_{ML}, R²_{C&U} in Tables 16 to 19 respectively. The null and residual deviances are also computed and compared between the models in Tables 20 and 21 respectively. Please see Tables 14 to 21 which are included as the supplementary material.

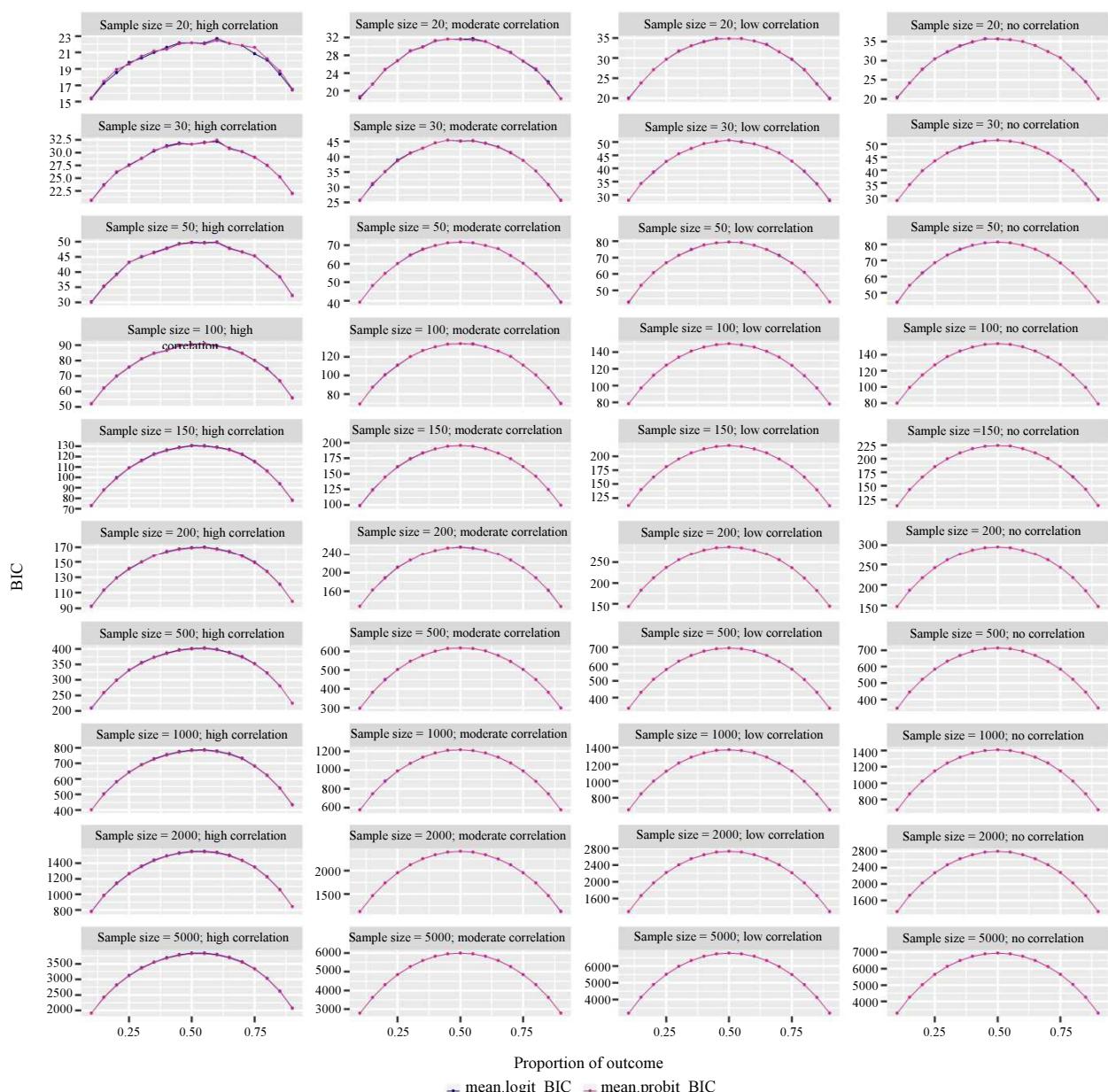


Fig. 4: Comparison of Logit and Probit models using BIC values

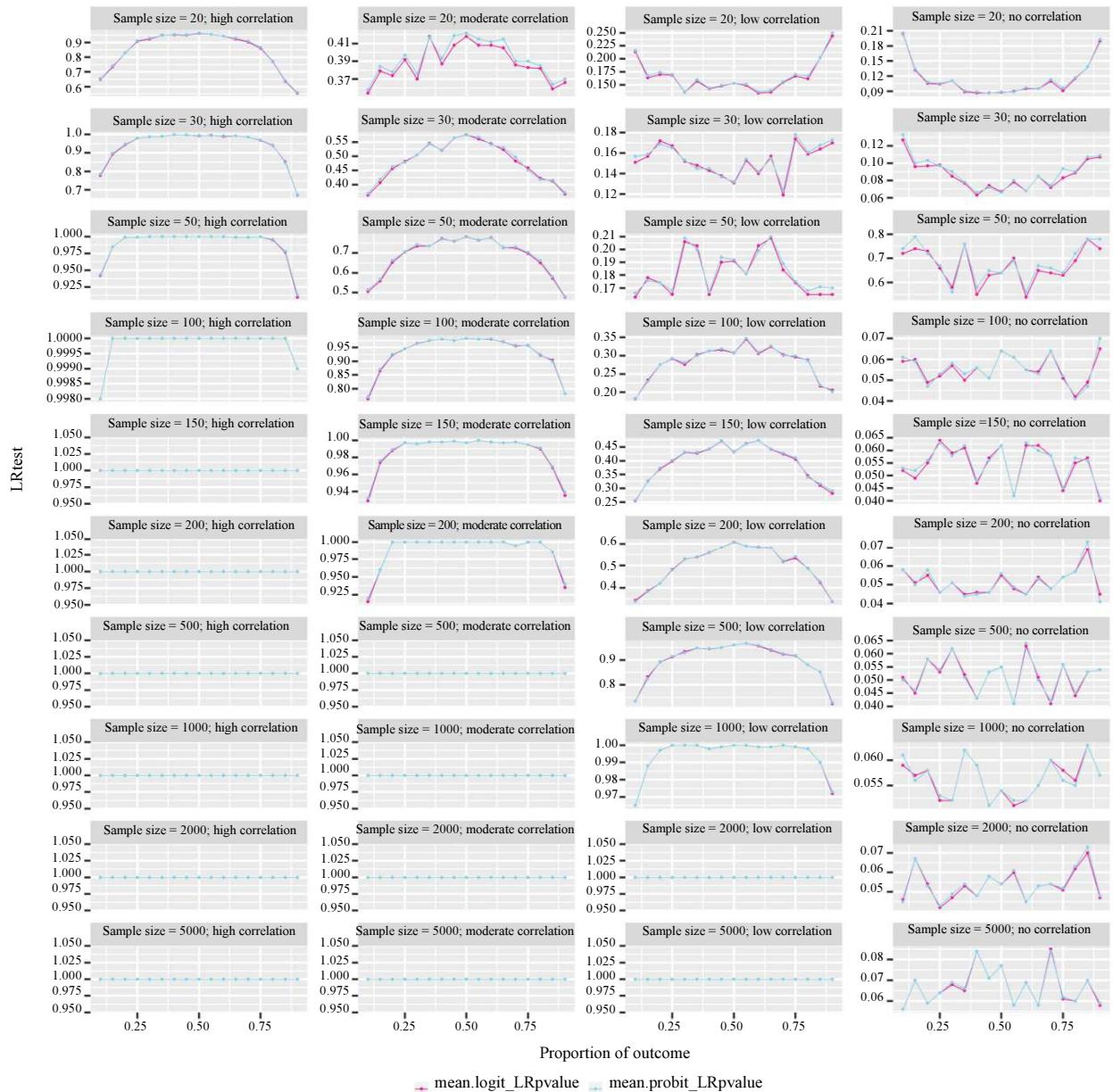


Fig. 5: Comparison of Logit and Probit models using LR test p values

Conclusion

The logit and probit models were considered in the study under various conditions of varying correlation structures, different prevalences and sample sizes. These models were compared using indices of AIC/AICC, BIC, sensitivity, specificity, correct prediction percentage, RMSEs, null deviance, residual deviance, LR tests, etc. The indices AIC/AICC, BIC, LR test, pseudo R²s showed that both models fitted equally well in all the situations. Clinicians are more interested in the correct prediction of the outcome

among the study participants than any of the model fit indices. From the point of view, based on sensitivity, specificity, correct prediction percentage as well as RMSEs, it can be concluded that logit model fits better than probit model in most of the situations even though the underlying distribution is normal.

Strengths and Limitations of the Study

Even though there are simulation studies available in the literature, this is the first study which has considered many combinations in the simulation in

order to assess the efficiency of the models. 680 different combinations of simulation parameters in terms of correlation, sample size and prevalence of the outcome are considered in the present study. In comparison with other similar studies available in the literature, a live dataset also considered and the models are compared in the current study. Above all, in order to compare the performance of the models many possible measures are estimated compared to other studies in terms of fit indices, prediction, discrimination and model fitting.

The inferences made in this study are only based on the data in which the underlying distribution is normal. Further studies have to be conducted for getting concrete evidence about which model fits better under different situations such as various underlying skewed distributions namely lognormal distribution, exponential distribution, etc. and also by introducing different types and levels of skewness and kurtosis.

Financial Support and Sponsorship

This research received no specific grant from any funding agency.

Authors' Contributions

Amrutha Jose: Study conception and design, generation of simulated datasets, analysis and interpretation of data, writing and drafting the manuscript.

Mariyamma Philip: Study design, Generation of simulated datasets, analysis and interpretation of data, helped to evaluate and edit the manuscript, drafting of the manuscript and Critical revision.

Lavanya Tumkur Prasanna: Acquisition of Real life data, Helped to evaluate and edit the manuscript and Critical revision.

M. Manjula: Acquisition of real life data, Interpretation of data, helped to evaluate and edit the manuscript, drafting of the manuscript and critical revision.

Conflicts of Interest

There are no conflicts of interest.

Software Details

All the analyses and simulation were performed using R software (version 3.3.1). The important R packages used were: MASS, moments, zoo, lmtest, pscl, matrixStats, MVN, sgeostat, nlme, stats, dplyr, ggplot2, cowplot, grid, gridExtra and ggforce.

Supplementary Material

Supplementary material includes Tables 3 to 21.

References

- Alsoraji, A.H., S. Binhimd and M.K. Abd Elaal, 2018. A comparison of univariate probit and logit models using simulation. *Applied Math. Sci.*, 12: 185-204.
DOI: 10.12988/ams.2018.818
- Altman, D.G. and P. Royston, 2006. The cost of dichotomising continuous variables. *Brit. Med. J.*, 332: 1080-1080.
DOI: 10.1136/bmj.332.7549.1080
- Bar-On, R. and J.D.A. Parker, 2000. BarOn Emotional Quotient Inventory: Youth Version. 1st Edn., Multi-Health System, Incorporated, ISBN-10: 015802558X, pp: 86.
- Burnham, K.P. and D.R. Anderson, 2003. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. 2nd Edn., Springer, New York, ISBN-10: 0387953647, pp: 488.
- Cakmakyan, S. and A. Goktas, 2013. A comparison of binary logit and probit models with a simulation study. *J. Soc. Economic Stat.*, 2: 1-17.
- Chai, T. and R.R. Draxler, 2014. Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)? - Arguments against avoiding RMSE in the literature. *Geoscientific Model Dev.*, 7: 1247-1250.
DOI: 10.5194/gmd-7-1247-2014
- Garnefski, N., V. Kraaij and P. Spinhoven, 2002. Manual for the use of the cognitive emotion regulation questionnaire. Leiderdorp, The Netherlands.
- Gratz, K.L. and L. Roemer, 2004. Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure and initial validation of the difficulties in emotion regulation scale. *J. Psychopathol. Behav. Assess.*, 26: 41-54.
DOI: 10.1023/B:JOBA.0000007455.08539.94
- Hamilton, M., 1960. A rating scale for depression. *J. Neurol. Neurosurg. Psychiatry*, 23: 56-62.
DOI: 10.1136/jnnp.23.1.56
- Hurvich, C.M. and C.L. Tsai, 1989. Regression and time series model selection in small samples. *Biometrika*, 76: 297-307.
- Long, J.S., 1997. Regression Models for Categorical and Limited Dependent Variables. 1st Edn., Sage publications, California, pp: 416.
- Maddala, G.S., 1986. Limited-Dependent and Qualitative Variables in Econometrics. 1st Edn., Cambridge University Press, Cambridge, ISBN-10: 0521338255, pp: 401.
- McFadden, D., 1973. Conditional logit analysis of qualitative choice behavior.

Muris, P., 2001. A brief questionnaire for measuring self-efficacy in youths. J. Psychopathol. Behav. Assess., 23: 145-149.

Naggara, O., J. Raymond, F. Guilbert, D. Roy and A. Weill *et al.*, 2011. Analysis by categorizing or dichotomizing continuous variables is inadvisable: An example from the natural history of unruptured aneurysms. Am. J. Neuroradiol., 32: 437-440.
 DOI: 10.3174/ajnr.A2425

Raftery, A.E., 1996. Approximate Bayes factors and accounting for model uncertainty in generalised linear models. Biometrika, 83: 251-266.

Sinha, U.K., V. Shrama and M.K. Nepal, 2007. Development of a scale for assessing academic stress: A preliminary report. J. Institute Med., 23: 105-112.

Zhang, Z., 2016. Residuals and regression diagnostics: Focusing on logistic regression. Annals Translat. Med., 4: 195-195. DOI: 10.21037/atm.2016.03.36

List of Abbreviations

AIC	Akaike Information Criteria
AICc	Corrected AIC
AUC	Area Under the ROC Curve
BIC	Bayesian Information Criteria
CERQ	Cognitive Emotional Regulation Questionnaire
CPP	Correct Percentage Prediction
DERS	Difficulties in Emotion Regulation Questionnaire
DTS	Distress Tolerance Scale
HRSD	Hamilton Rating Scale for Depression
LR	Likelihood Ratio
PUC	Pre-university course
R ² _{C&U}	Cragg and Uhler's pseudo-R ²
R ² _{McF}	McFadden's pseudo-R ²
R ² _{McFAdj}	McFadden's adjusted pseudo-R ²
R ² _{ML}	Maximum Likelihood pseudo-R ²
RMSE	Root Mean Square Errors
SAAS	Scale for Assessing Academic Stress
sBIC	Sample size adjusted BIC

Supplementary Material

Table 3: Comparison of Logit and Probit models using Intercept (B_0) values

B0_ESTIMATES

Var-Cov matrix	Sample size	R square (OLS)	Proportion of outcome																	
			0.10		0.20		0.30		0.40		0.50		0.60		0.70		0.80		0.90	
			Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit		
High	50	0.9403	8.5581	2.4344	0.7865	0.0611	-2.3123	-1.1621	-3.0307	-1.5716	-2.9464	-1.7603	-2.8435	-1.9090	-3.3594	-2.4827	-2.3264	-2.5703	-71.3961	-24.8904
	100	0.9411	7.6576	2.3966	0.7216	0.0956	-1.1112	-0.6197	-1.9535	-1.0116	-2.7377	-1.5628	-3.2769	-1.9008	-3.0269	-1.9898	0.5216	-1.2687	-4.3067	-3.5479
	150	0.9420	5.7973	1.9867	0.4044	0.1868	-0.9946	-0.5514	-1.9314	-1.0907	-2.5480	-1.4538	-3.1493	-1.8035	-3.3740	-2.0246	-1.4172	-1.7104	4.4150	-0.5965
	200	0.9421	3.9679	1.5104	0.2248	0.1109	-0.9755	-0.5400	-1.7837	-1.0049	-2.5613	-1.4574	-3.1552	-1.8026	-3.5957	-2.0685	-2.7694	-2.0381	2.8693	-0.9443
	500	0.9422	2.2001	1.0932	0.2107	0.1218	-0.9116	-0.4992	-1.7835	-1.0053	-2.4736	-1.4090	-3.1000	-1.7713	-3.5669	-2.0442	-4.0106	-2.3211	-0.9517	-1.7926
	1000	0.9425	1.9418	1.0101	0.2149	0.1272	-0.8881	-0.4865	-1.7588	-0.9898	-2.4231	-1.3801	-3.0053	-1.7170	-3.5082	-2.0110	-4.0484	-2.3300	-3.4751	-2.3997
Moderate	50	0.3306	-8.0854	-2.4584	-1.3709	-0.5574	-1.4494	-0.7855	-1.7909	-1.0652	-2.1853	-1.3034	-2.6924	-1.6048	-3.0376	-1.9923	-3.4482	-2.2687	-11.2381	-5.1293
	100	0.3371	-1.2485	-0.2029	-0.5721	-0.2755	-1.0845	-0.6269	-1.5757	-0.9376	-2.1226	-1.2723	-2.6279	-1.3184	-1.8602	-3.7692	-2.2049	-3.8799	-2.5326	
	150	0.3380	-0.2087	0.1138	-0.4824	-0.2415	-1.0726	-0.6174	-1.5181	-0.9032	-2.0236	-1.2187	-2.4791	-1.4879	-3.0652	-1.8187	-3.7664	-2.1820	-4.5381	-2.6510
	200	0.3355	0.2556	0.2635	-0.4411	-0.2178	-1.0246	-0.5885	-1.5836	-0.9446	-1.9802	-1.1933	-2.5022	-1.5055	-3.0739	-1.8239	-3.7943	-2.1981	-4.8499	-2.7230
	500	0.3390	0.3747	0.2943	-0.4455	-0.2193	-1.0321	-0.5967	-1.4953	-0.8946	-1.9830	-1.1967	-2.4597	-1.4822	-3.0240	-1.7951	-3.7489	-2.1706	-4.8329	-2.6586
	1000	0.3407	0.4010	0.3079	-0.4321	-0.2097	-0.9891	-0.5718	-1.4861	-0.8899	-1.9435	-1.1746	-2.4478	-1.4750	-3.0056	-1.7869	-3.7241	-2.1563	-4.8801	-2.6808
Low	50	0.0750	-2.5296	-0.2883	-0.2036	0.0428	-0.5128	-0.2652	-0.6720	-0.4038	-1.0672	-0.6534	-1.5694	-0.9567	-2.0513	-1.2393	-2.6366	-1.5698	-3.6508	-2.1857
	100	0.0714	0.4098	0.5127	0.1871	0.1735	-0.1669	-0.0855	-0.6052	-0.3668	-1.0068	-0.6210	-1.4892	-0.9164	-2.0231	-1.2230	-2.6305	-1.5441	-3.6045	-2.0254
	150	0.0710	0.8221	0.6117	0.2889	0.2112	-0.1773	-0.0917	-0.5722	-0.3477	-1.0693	-0.6627	-1.4571	-0.8983	-1.9291	-1.1698	-2.5763	-1.5151	-3.6010	-2.0118
	200	0.0733	0.9560	0.6346	0.3704	0.2611	-0.1796	-0.0928	-0.5771	-0.3520	-0.9869	-0.6121	-1.4326	-0.8844	-1.9292	-1.1713	-2.5238	-1.4873	-3.4289	-1.9211
	500	0.0721	1.0203	0.6733	0.3228	0.2331	-0.1619	-0.0834	-0.6008	-0.3682	-1.0126	-0.6299	-1.4276	-0.8833	-1.9390	-1.1783	-2.5460	-1.5006	-3.4658	-1.9311
	1000	0.0710	1.0728	0.7001	0.3465	0.2469	-0.1629	-0.0839	-0.5809	-0.3561	-0.9847	-0.6127	-1.3910	-0.8609	-1.9180	-1.1664	-2.5148	-1.4829	-3.4416	-1.9196
No	50	0.0002	0.1364	0.7755	1.1314	0.8000	0.8393	0.5202	0.3763	0.2358	-0.1012	-0.6020	-0.5598	-0.3467	-0.9097	-0.5595	-1.3139	-0.8513	-1.1300	-1.1023
	100	0.0000	1.8066	1.1746	1.3598	0.8240	0.7681	0.4763	0.3895	0.2437	-0.0324	-0.2091	-0.3934	-0.2477	-0.9395	-0.5784	-1.4632	-0.8845	-2.0302	-1.2599
	150	0.0000	2.1220	1.2473	1.3184	0.8013	0.7769	0.4817	0.3620	0.2261	-0.0772	-0.4082	-0.4339	-0.2701	-0.9045	-0.5588	-1.4516	-0.8780	-2.2593	-1.3202
	200	0.0005	2.1801	1.2706	1.3895	0.8418	0.8161	0.5048	0.3537	0.2211	-0.0538	-0.0335	-0.4532	-0.2822	-0.8744	-0.5399	-1.4564	-0.8793	-2.2354	-1.3044
	500	0.0001	2.1667	1.2651	1.3152	0.8009	0.8319	0.5150	0.3504	0.2190	-0.0472	-0.0295	-0.4512	-0.2816	-0.8967	-0.5542	-1.4580	-0.8822	-2.2652	-1.3154
	1000	0.0000	2.1467	1.2554	1.3548	0.8233	0.8158	0.5052	0.3613	0.2259	-0.0673	-0.0421	-0.4398	-0.2746	-0.9014	-0.5569	-1.4563	-0.8815	-2.2380	-1.3023

Table 4: Comparison of Logit and Probit models using B₁ values

B1_ESTIMATES

Var-Cov matrix	Sample size	Proportion of outcome																	
		0.10		0.20		0.30		0.40		0.50		0.60		0.70		0.80		0.90	
		Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit
High	50	2.9521	1.0481	1.0529	0.4750	0.9565	0.4506	0.5493	0.2994	0.4755	0.2750	0.4594	0.2592	0.6670	0.3323	0.7427	0.3438	4.5073	1.5405
	100	0.4814	0.2575	0.4185	0.2347	0.4235	0.2395	0.4165	0.2374	0.4060	0.2335	0.4013	0.2300	0.3867	0.2203	0.3710	0.2108	0.7467	0.3299
	150	0.3870	0.2154	0.3939	0.2206	0.3998	0.2258	0.3984	0.2281	0.3872	0.2231	0.3793	0.2177	0.3673	0.2097	0.3580	0.2026	0.3498	0.1957
	200	0.3756	0.2079	0.3862	0.2157	0.3952	0.2233	0.3885	0.2222	0.3877	0.2228	0.3787	0.2175	0.3640	0.2076	0.3520	0.1988	0.3452	0.1903
	500	0.3570	0.1958	0.3738	0.2085	0.3784	0.2141	0.3802	0.2176	0.3751	0.2160	0.3665	0.2107	0.3546	0.2025	0.3391	0.1911	0.3292	0.1803
	1000	0.3521	0.1921	0.3668	0.2046	0.3752	0.2123	0.3760	0.2152	0.3714	0.2139	0.3625	0.2084	0.3498	0.1999	0.3374	0.1901	0.3242	0.1774
Moderate	50	4.1893	1.4481	0.2504	0.1433	0.2257	0.1325	0.2248	0.1337	0.2187	0.1306	0.2256	0.1340	0.2423	0.1420	0.2509	0.1438	0.9410	0.3647
	100	0.2478	0.1335	0.2207	0.1257	0.2094	0.1234	0.2067	0.1234	0.2073	0.1243	0.2086	0.1247	0.2109	0.1243	0.2243	0.1279	0.2449	0.1323
	150	0.2370	0.1273	0.2166	0.1233	0.2072	0.1220	0.2040	0.122										

Table 5: Comparison of Logit and Probit models using B_2 values

B2_ESTIMATES

Var-Cov matrix	Sample size	Proportion of outcome																	
		0.10		0.20		0.30		0.40		0.50		0.60		0.70		0.80		0.90	
		Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit
High	50	23.9127	7.6624	8.8689	3.5032	7.2698	4.0161	5.3174	2.4848	5.4989	2.4870	4.0527	2.0248	5.4599	2.5222	4.5932	2.1600	27.4874	9.1295
	100	8.9146	3.2307	3.7293	1.9103	1.7053	3.0068	3.1197	1.7494	3.2354	1.7681	1.6772	2.6985	1.5492	2.4714	1.4161	5.7568	2.3524	
	150	5.8190	2.3598	2.9290	1.6223	2.8763	1.6475	2.9313	1.6705	2.8424	1.6277	2.7325	1.5719	2.5630	1.4803	2.4527	1.4008	2.6177	1.3684
	200	4.6709	2.0505	2.7971	1.5787	2.8427	1.6255	2.8019	1.6062	2.8214	1.6139	2.6844	1.5452	2.5342	1.4621	2.3754	1.3562	2.5081	1.3378
	500	2.7301	1.4764	2.6802	1.5181	2.7261	1.5623	2.7711	1.5907	2.7274	1.5664	2.6315	1.5176	2.4611	1.4226	2.3223	1.3272	2.2952	1.2499
	1000	2.6472	1.4360	2.6505	1.5019	2.7009	1.5480	2.7439	1.5755	2.6916	1.5477	2.5796	1.4891	2.4237	1.4020	2.3163	1.3233	2.2494	1.2270
Moderate	50	4.0392	1.1074	2.1625	0.9413	1.3257	0.7203	1.0993	0.6472	1.0507	0.6264	1.0263	0.6126	1.0502	0.6162	1.2742	0.6915	2.0835	0.8424
	100	3.1187	1.1602	1.1868	0.6450	1.0448	0.6110	1.0203	0.6092	1.0088	0.6064	0.9992	0.5998	0.9832	0.5811	1.0199	0.5816	1.2840	0.6398
	150	1.8172	0.7945	1.0906	0.6096	1.0198	0.5969	0.9766	0.5837	0.9622	0.5808	0.9588	0.5767	0.9499	0.5615	1.0134	0.5785	1.1293	0.5916
	200	1.3457	0.6567	1.0577	0.5945	0.9959	0.5825	0.9748	0.5839	0.9439	0.5705	0.9509	0.5728	0.9901	0.5859	1.0120	0.5764	1.1360	0.5985
	500	1.1607	0.6005	1.0354	0.5830	0.9898	0.5814	0.9527	0.5723	0.9440	0.5714	0.9409	0.5674	0.9461	0.5604	0.9971	0.5689	1.0667	0.5612
	1000	1.1236	0.5832	1.0296	0.5800	0.9752	0.5736	0.9495	0.5708	0.9331	0.5655	0.9409	0.5677	0.9540	0.5656	0.9909	0.5656	1.0808	0.5697
Low	50	5.0390	1.5128	1.0831	0.4873	0.7679	0.4298	0.5729	0.3468	0.5855	0.3593	0.5896	0.3596	0.5821	0.3476	0.6780	0.3701	1.2494	0.5075
	100	1.3893	0.5366	0.7277	0.3898	0.5718	0.3386	0.5491	0.3353	0.5220	0.3228	0.5462	0.3361	0.5580	0.3351	0.6113	0.3487	0.7826	0.3854
	150	0.9207	0.4157	0.6092	0.3419	0.5798	0.3446	0.5379	0.3295	0.5575	0.3460	0.5483	0.3380	0.5535	0.3330	0.5890	0.3366	0.6754	0.3481
	200	0.7865	0.3890	0.5957	0.3342	0.5694	0.3387	0.5363	0.3292	0.5322	0.3304	0.5357	0.3306	0.5569	0.3357	0.5806	0.3322	0.6220	0.3210
	500	0.6944	0.3504	0.6078	0.3431	0.5650	0.3376	0.5419	0.3336	0.5462	0.3399	0.5443	0.3368	0.5658	0.3413	0.5770	0.3310	0.6383	0.3297
	1000	0.6622	0.3357	0.5894	0.3330	0.5593	0.3345	0.5407	0.3332	0.5324	0.3316	0.5248	0.3349	0.5314	0.3376	0.6240	0.3326		
No	50	1.3492	0.3445	0.1821	0.0515	0.2999	0.0108	0.0212	0.0122	0.0295	0.0185	0.0178	0.0121	-0.0062	0.0026	-0.1830	-0.0654	-1.2420	-0.3229
	100	0.1999	0.0594	0.0219	0.0104	0.0413	0.0238	0.0018	0.0005	0.0136	0.0085	-0.0098	-0.0059	0.0083	0.0059	-0.0199	-0.0091	-0.1747	-0.0557
	150	0.0374	0.0173	0.0097	0.0044	0.0146	0.0080	-0.0062	-0.0042	0.0224	0.0140	-0.0014	-0.0006	0.0052	0.0040	0.0005	0.0013	-0.0334	-0.0102
	200	0.0505	0.0233	-0.0011	-0.0012	0.0001	-0.0004	0.0253	0.0154	0.0085	0.0052	-0.0036	-0.0021	-0.0118	-0.0067	0.0068	0.0047	-0.0607	-0.0247
	500	0.0147	0.0069	0.0183	0.0100	-0.0010	-0.0009	0.0046	0.0028	0.0082	0.0051	0.0006	0.0004	0.0011	0.0065	0.0040	-0.0028	-0.0009	
	1000	0.0134	0.0064	0.0032	0.0016	0.0051	0.0030	-0.0022	-0.0014	0.0122	0.0076	0.0023	0.0015	0.0028	0.0018	0.0120	0.0070	-0.0098	-0.0045

Table 6: Comparison of Logit and Probit models using B_3 values

B3_ESTIMATES

Var-Cov matrix	Sample size	Proportion of outcome																	
		0.10		0.20		0.30		0.40		0.50		0.60		0.70		0.80		0.90	
		Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit
High	50	-29.4335	-9.6474	-10.5888	-4.3781	-9.2740	-4.0653	-6.2203	-3.0506	-6.9029	-3.2412	-6.4128	-3.0409	-11.1364	-4.4980	-14.0298	-5.0277	-35.3134	-11.2086
	100	-10.1259	-3.7809	-4.3486	-2.3245	-9.9152	-2.2236	-4.0747	-2.3013	-4.2149	-2.3395	-4.2294	-2.3665	-4.8734	-2.5272	-8.9644	-3.5462	-16.9479	-5.4959
	150	-7.2458	-2.9653	-3.8726	-2.1443	-3.7492	-2.1426	-3.7968	-2.1757	-3.8239	-2.2030	-3.8599	-2.2199	-4.1504	-2.2912	-6.8418	-2.9837	-14.2894	-4.7336
	200	-5.7138	-2.5439	-3.6685	-2.0548	-3.7079	-2.1186	-3.7120	-2.1368	-3.7943	-2.1828	-3.7783	-2.1806	-3.8402	-2.1921	-5.3014	-2.5446	-12.4495	-4.2733
	500	-3.7751	-1.9915	-3.5310	-1.9879	-3.5560	-2.0391	-3.6338	-2.0950	-3.6738	-2.1221	-3.6596	-2.1165	-3.6834	-2.1171	-3.8272	-2.1285	-8.0203	-3.1381
	1000	-3.5849	-1.9149	-3.4800	-1.9612	-3.5357	-2.0276	-3.5998	-2.0766	-3.6430	-2.1068	-3.6435	-2.1088	-3.6369	-2.0920	-3.7601	-2.1017	-5.3364	-2.4452
Moderate	50	-11.9732	-3.9839	-1.2737	-0.6855	-1.0402	-0.6119	-1.0407	-0.6136	-1.0532	-0.6295	-1.0540	-0.6540	-1.2598	-0.7110	-1.4982	-0.8232	-2.5642	
	100	-1.3078	-0.6465	-1.0485	-0.5947	-0.9921	-0.5863	-1.0017	-0.6001	-0.9699	-0.5835	-0.9879	-0.5905	-1.0148	-0.5954	-1.2353	-0.6721	-2.8831	-1.0807
	150	-1.1556	-0.6059	-1.0308	-0.5844	-0.9738	-0.5756	-0.9842	-0.5912	-0.9641	-0.5817	-0.9846	-0.5907	-1.0078	-0.5902	-1.0626	-0.5981	-1.7842	-0.7787
	200	-1.1171	-0.5865	-1.0226	-0.5823	-0.9897	-0.5850	-0.9401	-0.5663	-0.9539	-0.5768	-0.9630	-0.5777	-1.0133	-0.5950	-1.0488	-0.5874	-1.4303	-0.6922
	500	-1.0900	-0.5753	-0.9904	-0.5646	-0.9448	-0.5596	-0.9415	-0.5678	-0.9432	-0.5710	-0.9401	-0.5649	-0.9750	-0.5732	-1.0286	-0.5800	-1.1651	-0.6022
	1000	-1.0817	-0.5715	-0.9812	-0.5603	-0.9538	-0.5662	-0.9378	-0.5685	-0.9442	-0.5677	-0.9689	-0.5702	-1.0257	-0.5787	-1.1215	-0.5826		
Low	50	-0.3384	-0.1955	-0.2487	-0.1471	-0.2679	-0.1612	-0.2828	-0.1725	-0.3299	-0.2025	-0.3065	-0.1856	-0.3451	-0.1967	-0.5716	-0.2621	-2.6743	-0.7955
	100	-0.2908	-0.1552	-0.2810	-0.1619	-0.2589	-0.1553	-0.2775	-0.1704	-0.2829	-0.1747	-0.2759	-0.1685	-0.2914	-0.1729	-0.3369	-0.1884	-0.6321	-0.2593
	150	-0.3074	-0.1588	-0.2798	-0.1603	-0.2903	-0.1742	-0.2827	-0.1739	-0.2527	-0.1565	-0.2726	-0.1705	-0.2959	-0.1763	-0.3015	-0.1697	-0.4207	-0.2044
	200	-0.2903	-0.1507	-0.2750	-0.1577	-0.2723	-0.1636	-0.2727	-0.1682	-0.2894	-0.1791	-0.27							

Table 8: Comparison of Logit and Probit models using marginal estimates of B_1

B1_MARGINAL ESTIMATES

Var-Cov marix	Sample size	Proportion of outcome																	
		0.10		0.20		0.30		0.40		0.50		0.60		0.70		0.80		0.90	
		Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit
High	50	0.0203	0.0180	0.0352	0.0299	0.0403	0.0351	0.0449	0.0380	0.0462	0.0383	0.0443	0.0366	0.0402	0.0335	0.0320	0.0272	0.0205	0.0186
	100	0.0199	0.0185	0.0318	0.0287	0.0402	0.0348	0.0435	0.0367	0.0454	0.0375	0.0441	0.0363	0.0396	0.0329	0.0316	0.0269	0.0197	0.0181
	150	0.0199	0.0192	0.0315	0.0285	0.0394	0.0338	0.0441	0.0368	0.0449	0.0369	0.0433	0.0355	0.0390	0.0324	0.0315	0.0270	0.0199	0.0185
	200	0.0198	0.0195	0.0317	0.0286	0.0392	0.0337	0.0435	0.0362	0.0453	0.0372	0.0438	0.0359	0.0390	0.0324	0.0316	0.0273	0.0198	0.0188
	500	0.0197	0.0200	0.0315	0.0284	0.0391	0.0333	0.0435	0.0361	0.0449	0.0367	0.0436	0.0355	0.0393	0.0324	0.0315	0.0272	0.0198	0.0191
	1000	0.0198	0.0201	0.0313	0.0282	0.0391	0.0333	0.0435	0.0359	0.0449	0.0366	0.0435	0.0353	0.0392	0.0322	0.0315	0.0272	0.0198	0.0194
Moderate	50	0.0183	0.0180	0.0290	0.0239	0.0349	0.0263	0.0390	0.0281	0.0397	0.0281	0.0392	0.0282	0.0361	0.0272	0.0290	0.0239	0.0179	0.0172
	100	0.0177	0.0177	0.0279	0.0227	0.0345	0.0254	0.0384	0.0270	0.0401	0.0277	0.0390	0.0274	0.0348	0.0257	0.0284	0.0231	0.0176	0.0176
	150	0.0177	0.0178	0.0281	0.0227	0.0348	0.0255	0.0387	0.0270	0.0400	0.0275	0.0382	0.0266	0.0351	0.0256	0.0276	0.0223	0.0179	0.0179
	200	0.0173	0.0175	0.0280	0.0226	0.0351	0.0256	0.0392	0.0273	0.0399	0.0273	0.0388	0.0270	0.0348	0.0254	0.0278	0.0224	0.0178	0.0179
	500	0.0176	0.0177	0.0280	0.0225	0.0348	0.0252	0.0387	0.0267	0.0402	0.0273	0.0386	0.0266	0.0349	0.0253	0.0286	0.0225	0.0176	0.0176
	1000	0.0177	0.0177	0.0279	0.0223	0.0347	0.0251	0.0387	0.0267	0.0399	0.0271	0.0385	0.0265	0.0348	0.0251	0.0280	0.0224	0.0176	0.0177
Low	50	0.0071	0.0069	0.0111	0.0088	0.0145	0.0103	0.0148	0.0100	0.0154	0.0103	0.0148	0.0101	0.0134	0.0095	0.0106	0.0085	0.0071	0.0069
	100	0.0065	0.0065	0.0107	0.0083	0.0125	0.0087	0.0145	0.0096	0.0153	0.0099	0.0148	0.0098	0.0142	0.0099	0.0105	0.0082	0.0064	0.0063
	150	0.0067	0.0066	0.0110	0.0085	0.0132	0.0091	0.0145	0.0095	0.0149	0.0096	0.0146	0.0096	0.0131	0.0090	0.0103	0.0079	0.0069	0.0069
	200	0.0064	0.0064	0.0099	0.0077	0.0131	0.0090	0.0144	0.0094	0.0153	0.0098	0.0144	0.0094	0.0126	0.0087	0.0105	0.0081	0.0064	0.0064
	500	0.0066	0.0065	0.0105	0.0080	0.0131	0.0089	0.0145	0.0094	0.0150	0.0096	0.0144	0.0093	0.0131	0.0089	0.0105	0.0080	0.0064	0.0063
	1000	0.0065	0.0063	0.0106	0.0080	0.0131	0.0089	0.0143	0.0093	0.0150	0.0095	0.0144	0.0093	0.0131	0.0090	0.0105	0.0080	0.0065	0.0064
No	50	0.0004	0.0003	0.0002	0.0001	0.0009	0.0006	0.0011	0.0007	0.0020	0.0013	0.0017	0.0011	0.0002	0.0001	0.0009	0.0007	0.0004	0.0004
	100	0.0005	0.0005	0.0009	0.0007	0.0012	0.0008	0.0007	0.0004	0.0009	0.0006	0.0008	0.0005	0.0011	0.0007	0.0005	0.0004	0.0000	0.0001
	150	0.0006	0.0005	0.0010	0.0008	0.0012	0.0008	0.0012	0.0008	0.0010	0.0006	0.0009	0.0006	0.0010	0.0007	0.0006	0.0006	0.0004	0.0004
	200	0.0003	0.0003	0.0006	0.0004	0.0006	0.0004	0.0010	0.0006	0.0009	0.0006	0.0014	0.0009	0.0011	0.0007	0.0006	0.0005	0.0007	0.0007
	500	0.0003	0.0003	0.0007	0.0005	0.0005	0.0003	0.0010	0.0007	0.0010	0.0007	0.0010	0.0009	0.0006	0.0006	0.0007	0.0005	0.0004	0.0004
	1000	0.0005	0.0005	0.0008	0.0006	0.0008	0.0005	0.0012	0.0007	0.0010	0.0007	0.0010	0.0009	0.0006	0.0010	0.0007	0.0005	0.0005	0.0005

Table 9: Comparison of Logit and Probit models using marginal estimates of B_2

B2_MARGINAL ESTIMATES

Var-Cov marix	Sample size	Proportion of outcome																	
		0.10		0.20		0.30		0.40		0.50		0.60		0.70		0.80		0.90	
		Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit
High	50	0.5437	0.2589	0.4358	0.2768	0.3834	0.2887	0.4561	0.3158	0.5191	0.3295	0.3882	0.2802	0.2969	0.2384	0.2396	0.1889	0.2941	0.1644
	100	0.4096	0.2340	0.2777	0.2294	0.2848	0.2473	0.3260	0.2699	0.3548	0.2797	0.3263	0.2636	0.2756	0.2317	0.2111	0.1816	0.1913	0.1397
	150	0.2952	0.1995	0.2334	0.2088	0.2835	0.2468	0.3234	0.2685	0.3294	0.2691	0.3122	0.2566	0.2718	0.2288	0.2163	0.1874	0.1461	0.1286
	200	0.2403	0.1840	0.2291	0.2091	0.2825	0.2457	0.3140	0.2620	0.3303	0.2694	0.3109	0.2551	0.2718	0.2285	0.2136	0.1864	0.1431	0.1313
	500	0.1506	0.1510	0.2263	0.2068	0.2820	0.2433	0.3174	0.2639	0.3268	0.2663	0.3136	0.2559	0.2725	0.2276	0.2156	0.1887	0.1377	0.1326
	1000	0.1485	0.1502	0.2264	0.2068	0.2814	0.2425	0.3173	0.2631	0.3258	0.2650	0.3094	0.2523	0.2714	0.2260	0.2159	0.1891	0.1374	0.1345
Moderate	50	0.3788	0.1922	0.2422	0.1473	0.1992	0.1392	0.1911	0.1362	0.1905	0.1351	0.1795	0.1295	0.1578	0.1190	0.1415	0.1127	0.1447	0.0942
	100	0.2088	0.1373	0.1497	0.1159	0.1724	0.1260	0.1902	0.1335	0.1946	0.1352	0.1864	0.1315	0.1629	0.1202	0.1289	0.1051	0.0912	0.0836
	150	0.1303	0.1055	0.1412	0.1121	0.1714	0.1246	0.1856	0.1292	0.1902	0.1311	0.1828	0.1279	0.1599	0.1174	0.1333	0.1075	0.0847	0.0831
	200	0.1004	0.0926	0.1389	0.1104	0.1688	0.1222	0.1870	0.1299	0.1887	0.1294	0.1831	0.1277	0.1678	0.1229	0.1335	0.1076	0.0857	0.0852
	500	0.0904	0.0885	0.1389	0.1100	0.1714	0.1234	0.1866	0.1287	0.1916	0.1306	0.1846	0.1277	0.1634	0.1188	0.1340	0.1075	0.0834	0.0830
	1000	0.0884	0.0868	0.1392	0.1100	0.1721	0.1231	0.1867	0.1286	0.1908	0.1297	0.1853	0.1280	0.1663	0.1238	0.1072	0.0851	0.0849	0.0849
Low	50	0.3420	0.1645	0.1369	0.0855	0.1346	0.0909	0.1181	0.0797	0.1262	0.0843	0.1213	0.0828	0.1041	0.0750	0.0889	0.0684	0.0829	0.0557
	100	0.1092	0.0757	0.1063	0.0768	0.1102	0.0758	0.1206	0.0794	0.1195	0.0778	0.1196	0.0796	0.1071	0.0750	0.0888	0.0692	0.0624	0.0579
	150	0.0759	0.0637	0.0914	0.0691	0.1136	0.0778	0.1201	0.0786	0.1300	0.0841	0.1226	0.0808	0.1084	0.0753	0.0875	0.0678	0.0567	0.0551
	200	0.0668	0.0621	0.0900	0.0679	0.													

Table 11: Comparison of Logit and Probit models using Sensitivity

SENSITIVITY																			
Var-Cov marix	Sample size	Proportion of outcome																	
		0.10		0.20		0.30		0.40		0.50		0.60		0.70		0.80		0.90	
		Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit
High	50	0.4960	0.4961	0.7150	0.7060	0.8640	0.8590	0.9490	0.9470	0.9820	0.9810	0.9930	0.9920	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	100	0.5180	0.5160	0.8200	0.8160	0.9240	0.9240	0.9710	0.9720	0.9870	0.9880	0.9980	0.9980	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	150	0.5050	0.4980	0.8630	0.8630	0.9470	0.9500	0.9820	0.9800	0.9970	0.9970	0.9980	0.9980	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	200	0.5810	0.5740	0.8490	0.8470	0.9560	0.9550	0.9880	0.9870	0.9990	0.9990	0.9990	0.9990	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	500	0.6260	0.6110	0.9050	0.9030	0.9760	0.9770	0.9960	0.9960	1.0000	1.0000	0.9990	0.9990	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	1000	0.6820	0.6660	0.9300	0.9340	0.9790	0.9800	0.9990	0.9990	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Moderate	50	0.5560	0.5790	0.4150	0.4200	0.5030	0.4950	0.6650	0.6600	0.8490	0.8460	0.9090	0.9070	0.9690	0.9690	0.9920	0.9920	0.9990	0.9990
	100	0.5070	0.5450	0.3330	0.3330	0.5060	0.5020	0.7250	0.7210	0.8670	0.8640	0.9410	0.9400	0.9820	0.9820	0.9960	0.9960	1.0000	1.0000
	150	0.4090	0.4590	0.3100	0.3010	0.5740	0.5670	0.7330	0.7300	0.8980	0.8950	0.9680	0.9680	0.9900	0.9900	0.9970	0.9980	1.0000	1.0000
	200	0.3770	0.4520	0.2980	0.2900	0.5540	0.5500	0.7530	0.7530	0.8780	0.8770	0.9610	0.9610	0.9910	0.9910	0.9960	0.9960	0.9990	0.9990
	500	0.2000	0.2320	0.3630	0.3420	0.6270	0.6240	0.8250	0.8270	0.9350	0.9370	0.9750	0.9750	0.9950	0.9950	1.0000	1.0000	1.0000	1.0000
	1000	0.1280	0.1230	0.3950	0.3770	0.6830	0.6800	0.8670	0.8680	0.9530	0.9520	0.9860	0.9860	0.9960	0.9960	1.0000	1.0000	1.0000	1.0000
Low	50	0.8400	0.8580	0.6950	0.7080	0.4760	0.4810	0.4550	0.4550	0.6360	0.6370	0.8480	0.8440	0.9480	0.9490	0.9940	0.9940	0.9970	0.9970
	100	0.9200	0.9380	0.7260	0.7600	0.4980	0.5080	0.4160	0.4150	0.6850	0.6850	0.9110	0.9090	0.9790	0.9790	0.9940	0.9950	1.0000	1.0000
	150	0.9380	0.9610	0.7590	0.7890	0.4650	0.4780	0.3950	0.3950	0.6710	0.6690	0.9070	0.9080	0.9850	0.9860	0.9980	0.9990	1.0000	1.0000
	200	0.9590	0.9750	0.8030	0.8270	0.4700	0.4840	0.3600	0.3620	0.7080	0.7080	0.9330	0.9310	0.9910	0.9920	1.0000	1.0000	1.0000	1.0000
	500	0.9860	0.9960	0.8190	0.8670	0.3260	0.3450	0.3130	0.3110	0.7470	0.7470	0.9250	0.9250	0.9960	0.9960	1.0000	1.0000	1.0000	1.0000
	1000	0.9980	1.0000	0.8180	0.8780	0.1950	0.2160	0.3430	0.3400	0.7750	0.7760	0.9560	0.9560	0.9960	0.9970	1.0000	1.0000	1.0000	1.0000
No	50	0.9290	0.9340	0.7980	0.8030	0.5450	0.5530	0.4190	0.4200	0.6050	0.6050	0.8600	0.8600	0.9500	0.9520	0.9920	0.9940	0.9960	0.9970
	100	0.9780	0.9860	0.9210	0.9300	0.7080	0.7190	0.3980	0.3960	0.5930	0.5930	0.9030	0.9040	0.9850	0.9850	0.9990	0.9990	0.9990	0.9990
	150	0.9950	0.9970	0.9700	0.9810	0.7960	0.8110	0.4760	0.4740	0.5610	0.5590	0.9080	0.9070	0.9950	0.9950	1.0000	1.0000	1.0000	1.0000
	200	0.9950	0.9960	0.9860	0.9890	0.8730	0.8890	0.4500	0.4540	0.5590	0.5600	0.9570	0.9570	0.9980	0.9980	1.0000	1.0000	1.0000	1.0000
	500	1.0000	1.0000	0.9990	0.9990	0.9780	0.9820	0.6620	0.6640	0.5350	0.5350	0.9950	0.9950	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	1000	1.0000	1.0000	1.0000	1.0000	0.9990	0.9990	0.8490	0.8490	0.5280	0.5280	0.9990	0.9990	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 12: Comparison of Logit and Probit models using Specificity

SPECIFICITY																			
Var-Cov marix	Sample size	Proportion of outcome																	
		0.10		0.20		0.30		0.40		0.50		0.60		0.70		0.80		0.90	
		Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit
High	50	0.7050	0.7000	0.8860	0.8840	0.9700	0.9680	0.9870	0.9870	0.9970	0.9960	0.9970	0.9970	1.0000	1.0000	0.9990	0.9990	1.0000	1.0000
	100	0.6770	0.6740	0.8940	0.8940	0.9780	0.9780	0.9930	0.9930	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	150	0.6440	0.6410	0.9390	0.9400	0.9790	0.9790	0.9950	0.9950	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	200	0.6920	0.6880	0.9340	0.9340	0.9860	0.9870	0.9980	0.9980	0.9990	0.9990	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	500	0.7650	0.7490	0.9530	0.9530	0.9960	0.9960	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	1000	0.7920	0.7770	0.9640	0.9650	0.9930	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Moderate	50	0.6570	0.6690	0.5110	0.5170	0.6340	0.6280	0.7980	0.7980	0.9110	0.9130	0.9580	0.9560	0.9850	0.9850	0.9980	0.9980	0.9990	0.9990
	100	0.5700	0.5940	0.4650	0.4610	0.6580	0.6520	0.8240	0.8220	0.9280	0.9280	0.9730	0.9720	0.9940	0.9940	0.9990	0.9990	1.0000	1.0000
	150	0.4600	0.5080	0.4400	0.4280	0.6600	0.6550	0.8360	0.8380	0.9420	0.9410	0.9810	0.9810	0.9950	0.9950	1.0000	1.0000	1.0000	1.0000
	200	0.4480	0.5170	0.4740	0.4630	0.6990	0.6960	0.8620	0.8590	0.9480	0.9470	0.9910	0.9910	0.9980	0.9980	1.0000	1.0000	1.0000	1.0000
	500	0.2730	0.2980	0.4900	0.4710	0.7570	0.7530	0.8900	0.8900	0.9770	0.9760	0.9910	0.9910	0.9980	0.9980	1.0000	1.0000	1.0000	1.0000
	1000	0.2060	0.1900	0.5150	0.5760	0.9150	0.9130	0.9750	0.9750	0.9950	0.9950	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Low	50	0.8590	0.8770	0.6910	0.7070	0.4940	0.4990	0.4920	0.4910	0.6920	0.6890	0.8800	0.8800	0.9610	0.9610	0.9570	0.9570	0.9960	0.9960
	100	0.9300	0.9430	0.7340	0.7680	0.5170	0.5320	0.4650	0.4630	0.7280	0.7290	0.9210	0.9190	0.9880	0.9880	1.0000	1.0000	1.0000	1.0000
	150	0.9400	0.9600	0.7670	0.7970	0.4580	0.4720	0.4180	0.4180	0.7430	0.7420	0.9420	0.9430	0.9900	0.9900	0.9990	0.9990	1.0000	1.0000
	200	0.9620	0.9760	0.8090	0.8320	0.4770													

Table 14: Comparison of Logit and Probit models using RMSE

Var-Cov matrix	Sample size	Proportion of outcome																	
		0.10		0.20		0.30		0.40		0.50		0.60		0.70		0.80		0.90	
		Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit
High	50	12.1123	12.1130	12.1810	12.1808	12.2173	12.2173	12.2927	12.2925	12.3837	12.3836	12.4866	12.4866	12.6229	12.6227	12.6918	12.6918	12.8627	12.8627
	100	12.1051	12.1050	12.1826	12.1826	12.2602	12.2602	12.3385	12.3384	12.4260	12.4259	12.5422	12.5419	12.6136	12.6133	12.7534	12.7530	12.8515	12.8514
	150	12.1135	12.1134	12.1818	12.1818	12.2748	12.2748	12.3273	12.3270	12.4226	12.4224	12.5226	12.5221	12.6111	12.6106	12.7430	12.7426	12.8751	12.8749
	200	12.1130	12.1129	12.1728	12.1729	12.2665	12.2665	12.3550	12.3549	12.4176	12.4174	12.5321	12.5318	12.6210	12.6205	12.7564	12.7559	12.8700	12.8698
	500	12.1242	12.1242	12.1720	12.1721	12.2638	12.2638	12.3486	12.3486	12.4193	12.4190	12.5276	12.5273	12.6452	12.6447	12.7466	12.7467	12.8576	12.8573
	1000	12.1280	12.1280	12.1898	12.1898	12.2571	12.2571	12.3472	12.3472	12.4283	12.4281	12.5262	12.5258	12.6292	12.6286	12.7454	12.7448	12.8704	12.8701
Moderate	50	12.1123	12.1122	12.1989	12.1988	12.3058	12.3055	12.4211	12.4211	12.4547	12.4548	12.5586	12.5586	12.6710	12.6712	12.7566	12.7568	12.8751	12.8752
	100	12.1706	12.1706	12.2338	12.2339	12.3039	12.3038	12.3879	12.3880	12.4680	12.4681	12.5987	12.5987	12.6720	12.6721	12.7701	12.7703	12.8932	12.8933
	150	12.1309	12.1309	12.2294	12.2294	12.2932	12.2931	12.3966	12.3965	12.4957	12.4957	12.5892	12.5892	12.6888	12.6889	12.7570	12.7570	12.8861	12.8862
	200	12.1426	12.1426	12.2298	12.2298	12.3124	12.3125	12.3954	12.3954	12.4749	12.4748	12.5682	12.5682	12.6709	12.6709	12.7600	12.7600	12.8734	12.8734
	500	12.1546	12.1546	12.2483	12.2483	12.3128	12.3128	12.4136	12.4136	12.4869	12.4869	12.5816	12.5816	12.6749	12.6749	12.7899	12.7896	12.8986	12.8986
	1000	12.1607	12.1607	12.2291	12.2291	12.3196	12.3196	12.4115	12.4114	12.4858	12.4858	12.5869	12.5869	12.6767	12.6767	12.7800	12.7800	12.8912	12.8912
Low	50	12.1371	12.1372	12.2443	12.2444	12.2895	12.2896	12.3878	12.3877	12.5180	12.5180	12.5835	12.5835	12.6948	12.6949	12.8187	12.8189	12.8676	12.8680
	100	12.1281	12.1281	12.2654	12.2655	12.3602	12.3602	12.4211	12.4211	12.5177	12.5177	12.5892	12.5892	12.6862	12.6862	12.7950	12.7951	12.8921	12.8922
	150	12.1526	12.1527	12.2657	12.2657	12.3636	12.3636	12.4233	12.4233	12.5203	12.5203	12.6005	12.6006	12.7098	12.7099	12.7726	12.7727	12.9093	12.9094
	200	12.1776	12.1776	12.2500	12.2500	12.3497	12.3497	12.4402	12.4402	12.5433	12.5433	12.6063	12.6063	12.6995	12.6995	12.8125	12.8125	12.9044	12.9044
	500	12.1695	12.1695	12.2588	12.2589	12.3511	12.3511	12.4465	12.4465	12.5287	12.5287	12.6207	12.6207	12.7212	12.7212	12.8062	12.8062	12.9045	12.9045
	1000	12.1713	12.1713	12.2565	12.2565	12.3414	12.3414	12.4408	12.4408	12.5318	12.5318	12.6206	12.6206	12.7035	12.7035	12.8057	12.8057	12.9077	12.9077
No	50	12.1129	12.1130	12.2312	12.2313	12.3582	12.3582	12.4239	12.4239	12.5155	12.5155	12.6123	12.6124	12.6594	12.6594	12.8188	12.8190	12.8889	12.8891
	100	12.1665	12.1666	12.2705	12.2705	12.3460	12.3459	12.4367	12.4367	12.5250	12.5250	12.6124	12.6124	12.7242	12.7242	12.7930	12.7931	12.8813	12.8814
	150	12.1619	12.1619	12.2679	12.2679	12.3393	12.3393	12.4511	12.4511	12.5304	12.5304	12.5984	12.5984	12.7184	12.7184	12.8003	12.8003	12.8893	12.8893
	200	12.1735	12.1735	12.2515	12.2515	12.3755	12.3755	12.4243	12.4243	12.5233	12.5233	12.6236	12.6236	12.7053	12.7053	12.8302	12.8302	12.9192	12.9193
	500	12.1707	12.1707	12.2645	12.2645	12.3512	12.3512	12.4349	12.4349	12.5369	12.5369	12.6257	12.6257	12.7279	12.7279	12.8133	12.8133	12.9087	12.9087
	1000	12.1788	12.1788	12.2677	12.2677	12.3598	12.3598	12.4501	12.4501	12.5367	12.5367	12.6407	12.6407	12.7163	12.7163	12.8200	12.8200	12.9064	12.9064

Table 15: Comparison of Logit and Probit models using AUC

Var-Cov matrix	Sample size	Proportion of outcome																	
		0.10		0.20		0.30		0.40		0.50		0.60		0.70		0.80		0.90	
		Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit
High	50	0.9441	0.9444	0.9319	0.9321	0.9252	0.9249	0.9223	0.9221	0.9164	0.9161	0.9127	0.9125	0.9160	0.9158	0.9155	0.9156	0.9260	0.9264
	100	0.9339	0.9342	0.9254	0.9255	0.9199	0.9198	0.9186	0.9185	0.9133	0.9131	0.9110	0.9108	0.9103	0.9102	0.9087	0.9087	0.9146	0.9148
	150	0.9309	0.9311	0.9224	0.9225	0.9179	0.9179	0.9138	0.9137	0.9106	0.9105	0.9086	0.9085	0.9062	0.9079	0.9080	0.9133	0.9136	0.9136
	200	0.9306	0.9308	0.9206	0.9207	0.9179	0.9179	0.9127	0.9127	0.9107	0.9106	0.9084	0.9083	0.9063	0.9062	0.9058	0.9058	0.9130	0.9132
	500	0.9271	0.9272	0.9188	0.9189	0.9142	0.9142	0.9116	0.9116	0.9088	0.9088	0.9055	0.9055	0.9037	0.9037	0.9029	0.9030	0.9100	0.9101
	1000	0.9255	0.9256	0.9178	0.9178	0.9139	0.9139	0.9108	0.9108	0.9078	0.9077	0.9052	0.9052	0.9023	0.9023	0.9085	0.9086	0.9120	0.9120
Moderate	50	0.8349	0.8367	0.7940	0.7945	0.7772	0.7770	0.7698	0.7696	0.7647	0.7642	0.7701	0.7701	0.7780	0.7778	0.7778	0.7963	0.7963	0.8325
	100	0.8074	0.8081	0.7751	0.7752	0.7615	0.7614	0.7570	0.7568	0.7560	0.7558	0.7562	0.7561	0.7606	0.7605	0.7787	0.7788	0.8070	0.8076
	150	0.8011	0.8018	0.7719	0.7720	0.7586	0.7585	0.7535	0.7533	0.7499	0.7498	0.7515	0.7515	0.7577	0.7576	0.7688	0.7689	0.7998	0.7995
	200	0.7931	0.7934	0.7681	0.7682	0.7580	0.7579	0.7519	0.7518	0.7475	0.7474	0.7497	0.7496	0.7575	0.7575	0.7669	0.7671	0.7976	0.7976
	500	0.7884	0.7886	0.7638	0.7638	0.7514	0.7514	0.7460	0.7459	0.7452	0.7451	0.7450	0.7450	0.7518	0.7518	0.7636	0.7636	0.7866	0.7867
	1000	0.7866	0.7866	0.7611	0.7611	0.7503	0.7503	0.7450	0.7450	0.7431	0.7444	0.7444	0.7444	0.7499	0.7499	0.7618	0.7618	0.7851	0.7852
Low	50	0.7513	0.7523	0.6884	0.6885	0.6774	0.6774	0.6587	0.6583	0.6604	0.6604	0.6643	0.6643	0.6737	0.6737	0.6738	0.6738	0.6894	0.6894
	100	0.6877	0.6878	0.6575	0.6574	0.6387	0.6385	0.6344	0.6345	0.6330	0.6329	0.6348	0.6347	0.6432	0.6432				

Table 17: Comparison of Logit and Probit models using R^2_{McFadden}

R-SQUARE (Adjusted McFadden)

Var-Cov marix	Sample size	Proportion of outcome																	
		0.10		0.20		0.30		0.40		0.50		0.60		0.70		0.80		0.90	
		Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit
High	50	0.2242	0.2275	0.3212	0.3241	0.3532	0.3500	0.3695	0.3722	0.3595	0.3622	0.3381	0.3409	0.3248	0.3273	0.2705	0.2726	0.1588	0.1613
	100	0.3196	0.3220	0.3821	0.3843	0.4031	0.4053	0.4158	0.4179	0.4058	0.4080	0.3943	0.3965	0.3730	0.3751	0.3337	0.3357	0.2567	0.2582
	150	0.3514	0.3538	0.4014	0.4036	0.4195	0.4216	0.4226	0.4250	0.4182	0.4207	0.4078	0.4102	0.3857	0.3881	0.3589	0.3605	0.2989	0.3001
	200	0.3712	0.3736	0.4097	0.4117	0.4315	0.4335	0.4301	0.4324	0.4282	0.4304	0.4171	0.4196	0.3973	0.3995	0.3664	0.3681	0.3181	0.3195
	500	0.3985	0.4010	0.4295	0.4314	0.4410	0.4431	0.4462	0.4486	0.4418	0.4443	0.4285	0.4309	0.4114	0.4137	0.3846	0.3864	0.3509	0.3523
	1000	0.4081	0.4103	0.4350	0.4369	0.4473	0.4493	0.4504	0.4526	0.4451	0.4476	0.4341	0.4365	0.4148	0.4171	0.3941	0.3959	0.3611	0.3626
Moderate	50	-0.0532	-0.0505	0.0134	0.0147	0.0295	0.0303	0.0413	0.0418	0.0409	0.0413	0.0409	0.0414	0.0337	0.0344	0.0133	0.0147	-0.0624	-0.0596
	100	0.0458	0.0473	0.0749	0.0755	0.0833	0.0839	0.0900	0.0902	0.0916	0.0917	0.0883	0.0886	0.0821	0.0827	0.0776	0.0785	0.0443	0.0462
	150	0.0817	0.0832	0.0989	0.0994	0.1032	0.1035	0.1065	0.1067	0.1056	0.1058	0.1046	0.1048	0.1020	0.1023	0.0945	0.0953	0.0799	0.0810
	200	0.0948	0.0955	0.1098	0.1104	0.1151	0.1153	0.1157	0.1159	0.1133	0.1134	0.1136	0.1138	0.1145	0.1149	0.1075	0.1081	0.1016	0.1028
	500	0.1308	0.1318	0.1309	0.1314	0.1288	0.1291	0.1286	0.1288	0.1294	0.1276	0.1278	0.1297	0.1300	0.1303	0.1309	0.1287	0.1296	0.1296
	1000	0.1418	0.1429	0.1366	0.1372	0.1351	0.1354	0.1343	0.1344	0.1334	0.1335	0.1334	0.1335	0.1346	0.1349	0.1376	0.1382	0.1407	0.1417
Low	50	-0.1749	-0.1734	-0.1094	-0.1089	-0.0798	-0.0796	-0.0786	-0.0785	-0.0729	-0.0727	-0.0742	-0.0742	-0.0836	-0.0833	-0.1096	-0.1091	-0.1791	-0.1777
	100	-0.0801	-0.0796	-0.0420	-0.0418	-0.0333	-0.0333	-0.0271	-0.0271	-0.0258	-0.0258	-0.0264	-0.0264	-0.0296	-0.0296	-0.0413	-0.0411	-0.0780	-0.0778
	150	-0.0433	-0.0430	-0.0195	-0.0194	-0.0117	-0.0116	-0.0094	-0.0094	-0.0093	-0.0093	-0.0087	-0.0087	-0.0118	-0.0117	-0.0207	-0.0207	-0.0411	-0.0408
	200	-0.0248	-0.0247	-0.0111	-0.0110	-0.0021	-0.0021	-0.0015	-0.0015	0.0006	0.0006	-0.0008	-0.0007	-0.0036	-0.0036	-0.0076	-0.0075	-0.0263	-0.0261
	500	0.0067	0.0068	0.0129	0.0129	0.0144	0.0144	0.0146	0.0146	0.0155	0.0155	0.0151	0.0151	0.0154	0.0155	0.0123	0.0123	0.0064	0.0064
	1000	0.0164	0.0165	0.0194	0.0194	0.0200	0.0200	0.0199	0.0199	0.0202	0.0202	0.0198	0.0198	0.0204	0.0204	0.0198	0.0198	0.0171	0.0171
No	50	-0.2178	-0.2165	-0.1359	-0.1357	-0.1132	-0.1131	-0.1026	-0.1026	-0.0991	-0.0990	-0.1032	-0.1031	-0.1101	-0.1100	-0.1366	-0.1363	-0.2084	-0.2075
	100	-0.1074	-0.1072	-0.0712	-0.0712	-0.0561	-0.0561	-0.0513	-0.0513	-0.0496	-0.0496	-0.0521	-0.0521	-0.0562	-0.0562	-0.0697	-0.0696	-0.1074	-0.1071
	150	-0.0735	-0.0734	-0.0462	-0.0461	-0.0369	-0.0369	-0.0344	-0.0344	-0.0324	-0.0324	-0.0337	-0.0337	-0.0381	-0.0381	-0.0469	-0.0469	-0.0738	-0.0737
	200	-0.0538	-0.0537	-0.0346	-0.0345	-0.0285	-0.0285	-0.0260	-0.0260	-0.0249	-0.0249	-0.0260	-0.0260	-0.0282	-0.0282	-0.0342	-0.0341	-0.0548	-0.0547
	500	-0.0212	-0.0212	-0.0136	-0.0136	-0.0112	-0.0112	-0.0103	-0.0103	-0.0100	-0.0100	-0.0102	-0.0102	-0.0115	-0.0115	-0.0140	-0.0140	-0.0213	-0.0213
	1000	-0.0106	-0.0106	-0.0069	-0.0069	-0.0057	-0.0057	-0.0052	-0.0052	-0.0051	-0.0051	-0.0051	-0.0051	-0.0056	-0.0056	-0.0069	-0.0069	-0.0107	-0.0107

Table 18: Comparison of Logit and Probit models using R^2_{ML}

R-SQUARE (CS)

Var-Cov marix	Sample size	Proportion of outcome																	
		0.10		0.20		0.30		0.40		0.50		0.60		0.70		0.80		0.90	
		Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit
High	50	0.2918	0.2933	0.3989	0.4007	0.4606	0.0769	0.4943	0.4962	0.4940	0.4959	0.4728	0.4747	0.4421	0.4438	0.3691	0.3704	0.2598	0.2609
	100	0.2627	0.2638	0.3787	0.3801	0.4436	0.4451	0.4790	0.4805	0.4806	0.4822	0.4636	0.4652	0.4232	0.4247	0.3486	0.3499	0.2321	0.2328
	150	0.2544	0.2556	0.3716	0.3730	0.4376	0.4390	0.4680	0.4736	0.4754	0.4659	0.4586	0.4135	0.4152	0.3448	0.3459	0.2287	0.2293	
	200	0.2516	0.2528	0.3671	0.3684	0.4367	0.4381	0.4648	0.4665	0.4728	0.4743	0.4558	0.4576	0.4127	0.4143	0.3388	0.3400	0.2246	0.2253
	500	0.2425	0.2437	0.3612	0.3624	0.4273	0.4288	0.4616	0.4632	0.4680	0.4697	0.4486	0.4504	0.4067	0.4083	0.3321	0.3333	0.2188	0.2195
	1000	0.2406	0.2416	0.3587	0.3599	0.4264	0.4278	0.4596	0.4613	0.4654	0.4672	0.4477	0.4495	0.4033	0.4050	0.3321	0.3332	0.2173	0.2181
Moderate	50	0.1520	0.1534	0.1880	0.1890	0.2054	0.2062	0.2195	0.2199	0.2202	0.2206	0.2191	0.2196	0.2083	0.2090	0.1877	0.1888	0.1482	0.1498
	100	0.1208	0.1216	0.1581	0.1586	0.1800	0.1805	0.1952	0.1954	0.1999	0.2001	0.1938	0.1940	0.1789	0.1794	0.1605	0.1612	0.1201	0.1212
	150	0.1122	0.1131	0.1510	0.1514	0.1735	0.1738	0.1874	0.1876	0.1899	0.1900	0.1853	0.1855	0.1724	0.1727	0.1478	0.1485	0.1113	0.1119
	200	0.1048	0.1052	0.1464	0.1469	0.1722	0.1725	0.1846	0.1848	0.1854	0.1856	0.1822	0.1825	0.1717	0.1721	0.1446	0.1451	0.1088	0.1095
	500	0.0994	0.0999	0.1396	0.1401	0.1620	0.1623	0.1751	0.1752	0.1801	0.1802	0.1740	0.1741	0.1628	0.1630	0.1393	0.1398	0.0983	0.0988
	1000	0.0971	0.0977	0.1362	0.1366	0.1606	0.1733	0.1735	0.1768	0.1769	0.1724	0.1725	0.1599	0.1602	0.1370	0.1375	0.0964	0.0970	0.0208
Low	50	0.0868	0.0876	0.0852	0.0856	0.0956	0.0957	0.0881	0.0882	0.0923	0.0924	0.0932	0.0938	0.0918	0.0921	0.0851	0.0855	0.0846	0.0853
	100	0.0475	0.0478	0.0561	0.0562	0.0571	0.0571	0.0607	0.0607	0.0614	0.0614	0.0617	0.0611	0.0611	0.0611	0.0567	0.0568	0.0488	0.0489
	150	0.0380	0.0381	0.0458	0.0459	0.0505	0.0506	0.0506	0.0506	0.0512	0.0521	0.0518	0.0518	0.0530	0.0530	0.0505	0.0506	0.0445	0.0446
	200	0.0334	0.0335	0.0380	0.0380	0													

Table 20: Comparison of Logit and Probit models using Null Deviance

		Proportion of outcome																	
Var-Cov matrix	Sample size	0.10		0.20		0.30		0.40		0.50		0.60		0.70		0.80		0.90	
		Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit
High	50	32.2117	32.1196	49.8748	49.8748	60.8055	60.8055	66.9806	66.9806	68.9483	68.9483	66.8996	66.8996	60.8076	60.8076	49.8544	49.8544	32.0760	32.0760
	100	64.4609	64.4609	99.6598	99.6598	122.0475	122.0475	134.1724	134.1724	138.2680	138.2680	134.1801	134.1801	121.9441	121.9441	99.6980	99.6980	64.1290	64.1290
	150	97.4109	97.4109	149.9290	149.9290	183.1909	183.1909	201.7315	201.7315	207.5950	207.5950	201.4324	201.4324	182.9426	182.9426	149.9502	149.9502	97.2204	97.2204
	200	129.6533	129.6533	200.0871	200.0871	244.1599	244.1599	268.7358	268.7358	276.8885	276.8885	268.9660	268.9660	244.0867	244.0867	199.5753	199.5753	128.8636	128.8636
	500	323.7834	323.7834	499.4844	499.4844	610.4052	610.4052	672.5030	672.5030	692.7626	692.7626	672.5753	672.5753	611.3042	611.3042	499.7756	499.7756	323.8324	323.8324
	1000	650.2299	650.2299	999.1904	999.1904	1221.3415	1221.3415	1345.6950	1345.6950	1385.9120	1385.9120	1345.7407	1345.7407	1222.0145	1222.0145	999.6068	999.6068	651.4052	651.4052
Moderate	50	32.1196	32.1196	49.8748	49.8748	60.8055	60.8055	66.9806	66.9806	68.9448	68.9448	66.9689	66.9689	60.8688	60.8688	49.7304	49.7304	31.8717	31.8717
	100	64.2248	64.2248	99.7446	99.7446	121.9100	121.9100	134.3004	134.3004	138.2575	138.2575	134.2260	134.2260	122.0359	122.0359	99.8610	99.8610	64.5183	64.5183
	150	97.0194	97.0194	149.4993	149.4993	182.9689	182.9689	201.5228	201.5228	207.6051	207.6051	201.5236	201.5236	182.8031	182.8031	150.3184	150.3184	97.3419	97.3419
	200	129.0748	129.0748	199.5712	199.5712	244.2232	244.2232	268.8746	268.8746	276.8659	276.8659	268.0051	268.0051	244.1470	244.1470	199.5870	199.5870	129.4707	129.4707
	500	324.8849	324.8849	499.5845	499.5845	610.7554	610.7554	672.5899	672.5899	692.7651	692.7651	672.9564	672.9564	610.0978	610.0978	500.3212	500.3212	325.1309	325.1309
	1000	650.2104	650.2104	1000.2849	1000.2849	1221.1841	1221.1841	1345.5199	1345.5199	1385.9304	1385.9304	1345.5060	1345.5060	1222.6573	1222.6573	1000.2190	1000.2190	650.5622	650.5622
Low	50	31.7463	31.7463	49.8653	49.8653	60.9021	60.9021	66.9426	66.9426	68.9541	68.9541	66.9775	66.9775	60.6692	60.6692	49.8935	49.8935	31.7921	31.7921
	100	64.8023	64.8023	99.6952	99.6952	121.6991	121.6991	131.2649	131.2649	138.2511	138.2511	134.2599	134.2599	121.9209	121.9209	99.3889	99.3889	64.5870	64.5870
	150	97.3320	97.3320	149.5112	149.5112	183.0748	183.0748	201.4854	201.4854	207.5937	207.5937	201.5822	201.5822	182.9757	182.9757	149.6003	149.6003	97.1209	97.1209
	200	129.1909	129.1909	199.1282	199.1282	243.7828	243.7828	268.7979	268.7979	276.8892	276.8892	268.7892	268.7892	243.9569	243.9569	199.1783	199.1783	129.6601	129.6601
	500	324.7672	324.7672	501.2406	501.2406	610.8314	610.8314	672.8439	672.8439	692.7777	692.7777	672.8286	672.8286	610.9082	610.9082	499.9065	499.9065	324.2700	324.2700
	1000	650.3308	650.3308	1001.1373	1001.1373	1221.2144	1221.2144	1345.8467	1345.8467	1385.9413	1385.9413	1345.4445	1345.4445	1220.9287	1220.9287	1000.6388	1000.6388	650.1909	650.1909
No	50	31.8114	31.8114	49.8339	49.8339	60.8890	60.8890	66.9588	66.9588	68.9631	68.9631	66.8917	66.8917	60.7832	60.7832	49.7263	49.7263	32.2584	32.2584
	100	64.7857	64.7857	99.4428	99.4428	122.1953	122.1953	134.2240	134.2240	138.2679	138.2679	134.1798	134.1798	121.7164	121.7164	99.4868	99.4868	63.8456	63.8456
	150	96.7598	96.7598	149.5316	149.5316	183.1381	183.1381	201.5096	201.5096	207.5937	207.5937	201.7880	201.7880	183.2711	183.2711	149.5131	149.5131	96.9427	96.9427
	200	129.0858	129.0858	199.8466	199.8466	244.1384	244.1384	269.0012	269.0012	276.8791	276.8791	268.8408	268.8408	243.8574	243.8574	200.1458	200.1458	129.3113	129.3113
	500	324.1980	324.1980	500.4951	500.4951	611.1514	611.1514	672.5177	672.5177	692.7805	692.7805	672.8250	672.8250	610.2319	610.2319	500.4340	500.4340	325.1131	325.1131
	1000	648.7476	648.7476	999.5336	999.5336	1221.0092	1221.0092	1345.7500	1345.7500	1385.9266	1385.9266	1345.8474	1345.8474	1221.9949	1221.9949	999.8199	999.8199	649.1526	649.1526

Table 21: Comparison of Logit and Probit models using Residual Deviance

		Proportion of outcome																		
Var-Cov matrix	Sample size	0.10		0.20		0.30		0.40		0.50		0.60		0.70		0.80		0.90		
		Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	Logit	Probit	
High	50	14.5075	14.4039	23.6417	23.4952	29.3548	29.5555	32.2519	32.0705	34.1611	33.9751	34.2749	34.0990	31.0405	30.8901	26.2725	26.1671	16.6004	16.5804	
	100	33.5980	33.4456	51.5404	51.3193	62.8510	62.5840	68.3786	68.0934	72.1604	71.8503	71.2635	70.9645	66.4213	66.3270	66.1621	66.1383	37.2819	37.2819	
	150	52.9905	52.7535	79.7111	79.3806	96.3304	95.9500	106.4963	106.0089	110.7752	110.2554	109.2825	108.8085	102.3256	101.8973	101.8973	101.8973	57.9557	57.8347	
	200	71.3238	70.9976	108.0944	107.6961	128.0856	128.3226	128.3226	128.3226	143.1443	142.5174	142.5174	141.6740	141.6740	137.0791	136.5355	116.3935	116.0444	77.4948	77.4948
	500	184.6064	183.8052	274.9121	273.9622	331.1793	329.9124	362.4131	360.8600	376.6665	374.9949	374.3712	372.7629	349.7734	348.3860	297.4948	296.5888	200.0733	199.6273	
	1000	374.6985	373.2976	554.5062	552.6554	665.0821	662.5877	729.6326	726.5838	758.9964	755.6245	751.5639	748.2921	702.2720	705.0916	593.8741	406.0390	405.0386	405.0386	
Moderate	50	23.6416	23.5575	39.1484	39.0858	48.9514	48.8241	54.1821	54.1018	54.2272	54.1949	48.8085	48.7604	54.2272	54.1949	38.9515	38.9515	23.6053	23.5156	
	100	51.1370	51.0445	82.2628	82.2015	101.7495	101.6863	112.2048	112.1812	115.5982	115.5839	112.3315	112.0218	101.9551	101.8209	82.0210	81.5224	51.4034	51.4034	
	150	78.9782	78.8293	124.6924	124.6178	154.0878	154.0267	170.0245	175.6749	175.6464	170.4038	170.4038	154.1544	154.0875	126.0942	125.9718	79.4778	79.3697		
	200	106.7542	106.6558	167.6555	167.5324	206.1123	206.0562	227.7605	227.6994	235.5017	235.4628	228.3627	228.3011	206.1810	206.0877	168.1062	167.9943	106.2312	106.0848	
	500	312.5539	312.5304	484.7927	484.7781	592.0584	592.0545	652.9942	652.9911	672.0199	652.6917	652.6854	652.6854	5						