TR/10/84 OCTOBER 1984

NULL DISTRIBUTION OF SOME GOODNESS OF FIT STATISTICS FOR LOGISTIC REGRESSION

by

Z. AL-SARRAF and D. H. YOUNG

W9259478

Null Distribution Properties Of Some Goodness Of Fit Statistics For Logistic Regression

by

Z. Al-Sarraf and D.H. Young

Summary

The null distribution moment and percentile properties of several goodness of fit statistics for logistic regression models are considered. Small sample approximations to the critical values of the statistics are evaluated for the case of a single explanatory variable with equally spaced values.

CONTENTS

1. Introduction

Let Y_1, Y_2 ,..., Y_g represent g independent binomial random variables where Y_t is the number of successes in a set of n_t independent trials. For the t^{th} group, we let P_t denote the unknown probability of success in each trial. This probability is assumed to depend on the values $x_{t1}, \ldots, x_{tk}$, of k explanatory variables which are measured for each group. A commonly used model for the above situation is the linear logistic regression model.

$$\log(P_t/Q_t) = \underset{\sim}{x}'_t \beta , \qquad t = 1,2,\ldots\ldots,g \tag{1.1}$$

where $Q_t = 1 - P_t$ $\underset{\sim}{x}'_t = (1, x_{t1}, \ldots, x_{tk})$ and $\beta' = (\beta_0, \beta_1, \ldots, \beta_k)$. A very broad area of application of this model is described by Cox (1970).

The parameters in β are usually all unknown and there are three commonly used methods for estimating them, namely maximum likelihood (ML), minimum chi-square (MC) and weighted least squares (WLS). The methods are defined here, and a more detailed discussion of them is given in section 2.

The ML estimate $\hat{\underset{\sim}{\beta}}_1$ is the value of β which maximises the kernel of the log-likelihood given by

$$L(\underset{\sim}{\beta}) = \sum_{t=1}^{g} \left\{ y_t \underset{\sim}{x}'_t \underset{\sim}{\beta} - n_t \log(1 + e^{\underset{\sim}{x}'_t \underset{\sim}{\beta}}) \right\}. \tag{1.2}$$

The MC estimate $\hat{\underset{\sim}{\beta}}_2$ is the value of $\underset{\sim}{\beta}$ which minimises

$$R(\underset{\sim}{\beta}) = \sum_{t=1}^{g} \frac{(y_t - n_t p_t)^2}{n_t P_t Q_t} \tag{1.3}$$

where from (1.1),

$$p_t = \exp(\underset{\sim}{x}'_t \underset{\sim}{\beta}) / \{1 + \exp(\underset{\sim}{x}'_t \underset{\sim}{\beta})\}. \tag{1.4}$$

The ML and MC estimates both require an iterative method of solution. A non-iterative solution can be found by weighted least squares, this method sometimes being referred to as minimum logit chi-square estimation. Defining the group empirical logits by

$$z_t = \log\{y_t / (n_t - y_t)\} , \qquad t = 1,2,\ldots,g \tag{1.5}$$

the WLS estimate $\hat{\underset{\sim}{\beta}}_3$ is the value $\underset{\sim}{\beta}$ which minimises

$$S(\underset{\sim}{\beta}) = \sum_{t=1}^{g} n_t p_t q_t (z_t - \underset{\sim}{x}'_t \underset{\sim}{\beta})^2 \tag{1.6}$$

where $P_t = y_t / n_t$ and $q_t = 1 - P_t$. Because z_t is undefined when $y_t = 0$ or $y_t - n_t$, modified empirical logits defined by

$$z_t^* = \log\{(y_t + \tfrac{1}{2})/(n_t - y_t + \tfrac{1}{2})\}, \qquad t = 1,2,\ldots,g \tag{1.7}$$

are often used, the factor ½ being selected to minimise the large sample bias in estimating $\underset{\sim}{x}'_t \underset{\sim}{\beta}$. An esimate of the large sample variance of $\underset{\sim}{z}^*_t$ is

$$v_t = \frac{(n_t + 1)(n_t + 2)}{n_t^3 (p_t + \frac{1}{n_t})(q_t + \frac{1}{n_t})}$$

(Gart and Zweifel (1967)), The modified WLS estimate $\hat{\beta}_4$ is therefore the value of $\underset{\sim}{\beta}$ which minimises

$$S^*(\underset{\sim}{\beta}) = \sum_{t=1}^{g} \frac{n_t^3 (p_t + n_t^{-1})(q_t + n_t^{-1})}{(n_t + 1)(n_t + 2)} (z_t^* - \underset{\sim}{x}'_t \underset{\sim}{\beta})^2 \tag{1.8}$$

The ML, MC and WLS methods of estimation for the logistic regression model were first considered by Berkson (1955), who showed that the WLS estimator gave smaller mean square errors of estimation for a number of parameter configurations when $g = 3$ and $n_t = 10$, $t = 1,2,3$. In a fairly large scale simulation investigation covering a much wider range of models and sample sizes, we have also found that the WLS estimator is often more efficient than the ML and MC estimators. The results of this investigation will be given in a separate report.

In applications, it is of course important to assess the goodness of fit of the logistic regression model. Two well-known statistics have been proposed for this purpose. The first statistic is $R(\hat{\underset{\sim}{\beta}}_1)$, the sum of squares of the residuals $R_t^{(1)}$ obtained from a ML fit where

$$R_t^{(1)} = (y_t - n_t \hat{p}_t^{(1)}) / \{n_t \hat{P}_t^{(1)} \hat{Q}_t^{(1)}\}^{\frac{1}{2}}, \qquad t = 1,2,\ldots,g \tag{1.9}$$

where $\hat{p}_t^{(1)}$ denotes the ML estimate of P_t. The second statistic is the likelihood ratio statistic for comparing the fit of the logistic regression model with that of the general alternative in which the $\{P_t\}$ can vary freely over the parameter space $0 \leq P_t \leq 1$, $t = 1,2,\ldots,g$. Denoting the statistic by $D(\hat{\beta}_1)$ we have

$$D(\hat{\beta}_1) = 2\sum_{t=1}^{g}\left\{y_t \log\left(\frac{p_t}{\hat{p}_t^{(1)}}\right) + (n_t - y_t)\log\left(\frac{q_t}{\hat{Q}_t^{(1)}}\right)\right\}. \tag{1.10}$$

Writing $D(\hat{\beta}_1) = \sum_{t=1}^{g}\{D_t^{(1)}\}^2$. individual group measures of fit are provided by

$$D_t^{(1)} = \pm 2^{\frac{1}{2}}\left\{y_t \log\left(\frac{p_t}{\hat{p}_t^{(1)}}\right) + (n_t - y_t)\log\left(\frac{q_t}{\hat{Q}_t^{(1)}}\right)\right\}^{\frac{1}{2}}, \tag{1.11}$$

Following Pregibon (1980), we shall refer to $D(\hat{\beta}_1)$ as the deviance statistic and to the $\{D_t^{(1)}\}$ as the deviance components. If the goodness of fit statistics are linked to the estimation procedures for fitting the model, it is natural to use $D(\hat{\beta}_1)$, but more natural to use $R(\hat{\beta}_2)$ than $R(\hat{\beta}_1)$.

Other goodness of fit statistics are clearly provided by $S(\hat{\beta}_3)$ and $S^*(\hat{\beta}_4)$. An alternative class of goodness of fit statistics based on the sample logits is obtained by replacing the sample proportions by fitted probabilities in the weights and is defined by

$$T(\hat{\beta}) = \sum \frac{n_t^3(\hat{P}_t + n_t^{-1})(\hat{Q}_t + n_t^{-1})}{(n_t+1)(n_t+2)}(z_t^* - x'_t\hat{\beta})^2 \tag{1.12}$$

where we use $\hat{\beta}$ to represent a general estimate of 3. The use of unmodified empirical logits is not considered because the associated statistic would be undefined when $y = 0$ or $y_t = n_t$.

All the goodness of fit statistics mentioned previously have limiting chi-square distributions if the assumed logistic regression model is correct. However, little work appears to have been done to investigate the small sample properties of the statistics and in particular to determine how rapidly the null distribution of the statistics approach their limiting forms. We have therefore performed a fairly large scale simulation investigation to examine the small sample properties of the statistics $D(\hat{\beta}_1)$ $R(\hat{\beta}_1)$., $R(\hat{\beta}_2)$, $S(\hat{\beta}_3)$, $S^*(\hat{\beta}_4)$, $T(\hat{\beta}_1)$ and $T(\hat{\beta}_4)$ under the logistic regression model

$$\log(P_t/Q_t) = \beta_0 + \beta_1(t-1), \quad t = 1,2,\ldots,g \tag{1.13}$$

which occurs when there is a single explanatory variable having equally spaced values- The results of the investigation are given in section 4.

The model given by (1.13) might arise in a time-series context with binary data sets collected at equally spaced time points, there being a trend in the probability of success. The model might also be appropriate in the context of a controlled experiment with binary responses, the {x.} representing equally spaced levels of a single test variable.

In section 3, asymptotic distribution results for the residuals and deviances, $R_t^{(1)}$ and $D_t^{(1)}$ respectively based on an ML fit are given for a general class of models. The results are used in section 5 to derive approximations to the critical values of goodness of fit statistics based on the extreme values of the $\{R_t^{(1)}\}$ and $\{D_t^{(1)}\}$. The approximations are then evaluated under the model given by (1.13).

2. Review of Estimation Procedures for the Linear Logistic Regression Model

In this section, we discuss the computational procedures that may be used for determining the ML, MC and WLS estimates of $\underset{\sim}{\beta}$. A new method is presented showing how the MC estimate can be calculated using GLIM.

2.1 Maximum Likelihood Estimates

From (1.2), the first derivatives of the log-likelihood are

$$\frac{\partial L(\underset{\sim}{\beta})}{\partial \beta_j} = \Sigma\, x_{tj}\left(y_t - n_t \frac{e^{\underset{\sim}{x}'_t \underset{\sim}{\beta}}}{1 + e^{\underset{\sim}{x}'_t \beta}} \right), \qquad J = 0,1,\ldots, \quad K. \tag{2.1}$$

Setting $\partial L(\beta)/\partial\beta_j = 0$ for $j = 0,1,\ldots,k$, the ML estimate β_1 is obtained as the solution of the $k+1$ equations

$$\sum_{t=1}^{g} x_{tj}(y_t - n_t \hat{P}_t^{(1)}) = 0, \qquad j = 0,1,\ldots, k. \tag{2.2}$$

Setting $\hat{y}_t^{(1)} = n_t\, \bar{P}_t^{(1)}$, $\underset{\sim}{y}' = (y_1, y_2, \ldots, y_g)$ and $\underset{\sim}{\hat{y}}'_i = (\hat{y}_1^{(1)}, \hat{y}_2^{(1)}, \ldots, \hat{y}_g^{(1)})$, the set of equations given by (2.2) may be written in matrix form as

$$\underset{\sim}{x}'(\underset{\sim}{y} - \underset{\sim}{y}') = 0 \tag{2.3}$$

The likelihood equations are nonlinear in $\underset{\sim}{\hat{\beta}_1}$ and are solved by an iterative method based on a Newton-Raphson approach For this we need the negative values of the second order derivatives of the log-likelihood which are given by

$$-\frac{\partial^2 L(\underset{\sim}{\beta})}{\partial \beta_j \partial \beta_{j'}} = \sum_{t=1}^{g} n_t x_{tj} x_{tj'} \frac{e^{\underset{\sim}{x}'_t \underset{\sim}{\beta}}}{(1+e^{\underset{\sim}{x}'_t \underset{\sim}{\beta}})^2} \tag{2.4}$$

for $j, j' = 0,1,...,k$. Thus

$$\left\{-\frac{\partial^2 L(\underset{\sim}{\beta})}{\partial \beta_J \partial \beta_{J'}}\right\}_{\underset{\sim}{\beta}=\hat{\underset{\sim}{\beta}}_1} = \sum_{t=1}^{g} n_t x_{tj} x_{tj'} \hat{P}_t^{(1)} \hat{Q}_t^{(1)} \tag{2.5.}$$

for $j, j' = 0,1,...,k$. In matrix form, we may write the $(k+1) \times (k+1)$ terms given by (2.5) as

$$-\frac{\partial^2 L(\underset{\sim}{\beta})}{\partial \hat{\underset{\sim}{\beta}}_1 \partial \hat{\underset{\sim}{\beta}}_1'} = \underset{\sim}{X}'\underset{\sim}{V}\underset{\sim}{X} \tag{2.6}$$

where

$$\underset{\sim}{V} = \begin{bmatrix} n_1 \hat{P}_1^{(1)} \hat{Q}_1^{(1)} & 0 & \cdots & 0 \\ 0 & n_1 \hat{P}_2^{(1)} \hat{Q}_2^{(1)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & & n_g \hat{P}_g^{(1)} \hat{Q}_g^{(1)} \end{bmatrix} = \text{diag}\left(\frac{\partial \hat{\underset{\sim}{y}}_1}{\partial \hat{\underset{\sim}{\beta}}}\right) \tag{2.7}$$

To find $\hat{\underset{\sim}{\beta}}_1$ by the Newton-Raphson method, if we let $\hat{\underset{\sim}{\beta}}_1^{(\ell)}$ denote the approximation to $\hat{\underset{\sim}{\beta}}_1$ at the ℓth stage of iteration, we have

$$\hat{\beta}_1^{(\ell+1)} = \hat{\beta}_1^{(\ell)} + \left\{-\frac{\partial^2 L(\underset{\sim}{\beta}_1^{(\ell)})}{\partial \hat{\underset{\sim}{\beta}}_1^{(\ell)} \partial \hat{\underset{\sim}{\beta}}_1^{(\ell)}},\right\}^{-1} \frac{\partial L(\underset{\sim}{\beta}_1^{(\ell)})}{\partial \hat{\underset{\sim}{\beta}}_\ell} \tag{2.8}$$

$$= \hat{\beta}_1^{(\ell)} + (\underset{\sim}{X}'\underset{\sim}{V}_\ell \underset{\sim}{X})^{-1} \underset{\sim}{X}'(\underset{\sim}{y} - \underset{\sim}{\hat{y}}_1^{(\ell)})$$

where $\underset{\sim}{v}_\ell$ and $\hat{\underset{\sim}{y}}_1$ denote $\underset{\sim}{v}$ and $\hat{\underset{\sim}{y}}_1$ evaluated at $\hat{\underset{\sim}{\beta}}_1^{(\ell)}$,

The above iterative process may be viewed as a method of iteratively reweighted least squares. Thus if we let

$$\underset{\sim}{z}_\ell = X\hat{\underset{\sim}{\beta}}_1^{(\ell)} + \underset{\sim}{v}_\ell^{-1}(y - \hat{\underset{\sim}{y}}_1^{(\ell)}) \tag{2.9}$$

denote a 'pseudo' observation vector at the ℓth stage, equation (2.8) may be written as

$$\hat{\underset{\sim}{\beta}}_1(\ell+1) = (\underset{\sim}{X}'\underset{\sim}{V}_\ell \underset{\sim}{X})^{-1} \underset{\sim}{X}' \underset{\sim}{V}_\ell \underset{\sim}{Z}_\ell \tag{2.10}$$

From generalised least squares theory, this formula corresponds to that for the best linear unbiased estimate of $\underset{\sim}{\beta}_1(\ell+1)$ in the linear model $\underset{\sim}{Z}_\ell = X\,\underset{\sim}{\beta}_1^{(\ell+1)} + \underset{\sim}{\varepsilon}$ where $\underset{\sim}{\varepsilon}$ has zero mean and known covariance matrix $\underset{\sim}{V}_\ell$.

At convergence we have $\underset{\sim}{z} = X\hat{\underset{\sim}{\beta}}_1 + \underset{\sim}{v}^{-1}(y - \underset{\sim}{\hat{y}}_1)$ and the ML estimate may be written as the solution of

$$\hat{\underset{\sim}{\beta}}_1 = (\underset{\sim}{X}'\underset{\sim}{V}\underset{\sim}{X})^{-1} \underset{\sim}{X}'\underset{\sim}{V}\underset{\sim}{Z} \tag{2.11}$$

It should be noted that this does not provide an explicit expression for $\hat{\underset{\sim}{\beta}}_1$ since $\underset{\sim}{V} = \underset{\sim}{V}(\hat{\underset{\sim}{\beta}}_1)$ and $\underset{\sim}{z} = \underset{\sim}{z}(\hat{\underset{\sim}{\beta}}_1.)$ are both functions of $\hat{\underset{\sim}{\beta}}_1$.

2.2 <u>Minimum Chi-Square Estimation</u>

The MC estimate $\hat{\underset{\sim}{\beta}}_2$ is the value of $\underset{\sim}{\beta}$ which minimises $R(\hat{\underset{\sim}{\beta}}_2)$ given by (1.3). We have

$$\frac{\partial R(\underset{\sim}{\beta})}{\partial \beta_j} = \sum_{t=1}^{g} \frac{(y_t - n_t P_t)}{n_t P_t Q_t} \frac{\partial P_t}{\partial \beta_j} \left\{ (y_t - n_t P_t)\left(\frac{1}{Q_t} - \frac{1}{P_t} \right) - 2n_t \right\} \tag{2.12}$$

where

$$\frac{\partial P_t}{\partial \beta_j} = \frac{x_{tj} \exp(\underset{\sim}{x}'_t \underset{\sim}{\beta})}{\{1 + \exp(\underset{\sim}{x}'_t \underset{\sim}{\beta})\}^2} = x_{tj} P_t Q_t. \tag{2.13}$$

Substitution of (2.13) in (2.12) and simplification gives

$$\frac{\partial R(\underset{\sim}{\beta})}{\partial \beta_j} = -\sum_{t=1}^{g} \frac{(p_t - P_t) n_t x_{tj} (Q_t p_t + P_t q_t}{P_t Q_t}.$$

Hence the MC estimates are given by the solution of the k +1 equations

$$\sum_{t=1}^{g} n_t x_t (p_t - \hat{P}_t^{(2)})(\hat{P}_t^{(2)} q_t + \hat{Q}_t^{(2)} p_t)(\hat{P}_t^{(2)} \hat{Q}_t^{(2)})^{-1} = 0 \tag{2.14}$$

j = 0,1,...,k, where

$$\hat{P}_t^{(z)} = \exp(\underset{\sim}{x}'_t \hat{\underset{\sim}{\beta}}_2)/\{1+\exp(\underset{\sim}{x}'_t \hat{\underset{\sim}{\beta}}_2) \qquad t = 1,2,\ldots,g. \tag{2.15}$$

The solution of (2.14) must be found by iteration. The following approach shows how GLIM can be used to calculate $\hat{\underset{\sim}{\beta}}_2$

We may write

$$R(\underset{\sim}{\beta}) = \sum_{t=1}^{g} \left\{ \frac{y_t^2}{n_t} e^{-\underset{\sim}{x}'_t \underset{\sim}{\beta}} + \frac{(n_t - y_t)^2}{n_t} e^{\underset{\sim}{x}'_t \underset{\sim}{\beta}} - \frac{2y_t(n_t - y_t)}{n_t} \right\}.$$

If we put

$$y_{t1} = y_t^2/n_t \quad , \qquad y_{t2} = (n_t - y_t)^2/n_t \tag{2.16}$$

$$\mu_{t1} = \exp(\underset{\sim}{x}'_t \underset{\sim}{\beta}), \qquad \mu_{t2} = \exp(-\underset{\sim}{x}'_t \underset{\sim}{\beta}) \tag{2.17}$$

then minimisation of $R(\underset{\sim}{\beta})$ is equivalent to minimisation of

$$R*(\underset{\sim}{\beta}) = \sum_{t=1}^{g} (y_{t1}\mu_{t1}^{-1} + y_{t2}\mu_{t2}^{-1}). \tag{2.18}$$

Now consider 2g independent random variables (Y_{t1}, Y_{t2}), $t = 1,2,\ldots,g$ where Y_{ti} . has an exponential distribution with mean μ_{ti}. For realised values y_{tj} $t = 1,2,,.,,g$, $j = 1,2$, the log-likelihood is

$$-\sum_{t=1}^{g} (y_{t1}\mu_{t1}^{-1} + y_{t2}\mu_{t2}^{-1}) = -R*(\underset{\sim}{\beta}) \tag{2.19}$$

since $\mu_{t1}\,\mu_{t2} = 1$ from (2.17). Thus minimisation of $R*(\underset{\sim}{\beta})$ is equivalent to maximisation of the log-likelihood treating the $\{y_{tj}\}$ as observations on independent exponentially distributed random variables.

To use GLIM, the data are entered as g pairs of vectors of observations, the vectors being $(y_{tl}, 1, xt_1, \ldots. x_{tk})$ and $(y_{t2}, -1, -x_{t1}, \ldots.. x_{tk})$ for the t^{th} pair. An exponential error distribution is declared and a logarithmic link function is used since $\log_{t1} = x'_t \underset{\sim}{\beta}$

2,3 Weighted Least Squares Estimation

The justification for the WLS and MWLS estimation procedures using the unmodified and modified empirical logits is as follows. If we let

$$z_t(a) = \log\left(\frac{y_t + a}{n_t - y_t + a}\right) = \log\left(\frac{p_t + \frac{a}{n_t}}{q_t + \frac{a}{n_t}}\right), \tag{2.20}$$

a Taylor series expansion about the value P_t gives

$$z_t(a) = \log\left\{\frac{(P_t + \frac{a}{n_t})}{(Q_t + \frac{a}{n_t})}\right\} + (P_t - p_t)\frac{(1+\frac{2a}{n_t})}{(p_t + \frac{a}{n_t})(Q_t + \frac{a}{n_t})}$$

$$+ \frac{(P_t - p_t)^2}{2}\frac{(1+\frac{2a}{n_t})(2P_t - 1)}{(P_t + \frac{a}{n_t})^2(Q_t + \frac{a}{n_t})^2} + \ldots \quad .$$

Using the results $E(p_t) = P$, $\text{var}(p_t) = p_t Q_t / n_t$ we obtain

$$E\{z_t(a)\} = \underset{\sim}{x}'_t \underset{\sim}{\beta} + \frac{(2P_t - 1)}{n_t P_t Q_t}\left(\frac{1}{2} - a\right) + 0\left(\frac{1}{n_t}\right) \tag{2.21}$$

showing that the bias of estimation of $\underset{\sim}{x'_t}\underset{\sim}{\beta}$ is $0(n_t^{-1})$ when the unmodified empirical logits are used $(a = 0)$ but is $o(n_t\text{-}1)$ when the modified empirical logits are used $(a=\frac{1}{2})$.

The approximate large sample variance of z_t is $(n_t P_t Q_t)^{-1}$. A least squares approach with empirical weights given by the reciprocals of the large sample variances evaluated at p_t leads to the criterion of minimisation of $S(\underset{\sim}{\beta})$ as defined by (1.6).The normal equations for obtaining the WLS estimate $\hat{\underset{\sim}{\beta}}_3$ are

$$\sum_{t=1}^{g} n_t p_t q_t x_{tj}(z_t - \underset{\sim}{x}'_t \hat{\underset{\sim}{\beta}}_3) = 0 \qquad J=0,1,\ldots,K \tag{2.22}$$

Similar arguments applied to the $\{z^*_t\}$ lead to the modified WLS estimation procedure based on minimisation of $S^*\underset{\sim}{\beta}$ given by (1.8).

3. Asymptotic Distribution Theory For The $\{R_t^{(1)}\}$ and $\{D_t^{(1)}\}$

The residuals $R_t^{(1)}$ and the deviance components $D_t^{(1)}$ from a maximum likelihood fit of a logistic regression model are often used as individual group measures of fit. In particular, normal probability plots of the ordered $R_t^{(1)}$ are commonly used to provide an informal graphical assessment of goodness of fit.

In this section, we derive the asymptotic joint distribution of the $\{R_t^{(1)}\}$ and $\{D_t^{(1)}\}$ under a general model in which the group probabilities of success are assumed to depend on ℓ unknown parameters, say $\theta_1, \theta_2, \ldots, \theta_\ell$. Putting $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_\ell)$ we now let $P_t(\underset{\sim}{\theta})$ denote the probability of success for each trial in the t^{th} group, $t = 1,2,\ldots,g$. We denote the total number of trials by $N = \sum_{t=1}^{g} n_t$ and let $\lambda_{t,N} = n_t/N$, $t = 1,2,\ldots,g$.

We shall now denote the number of successes in the t^{th} group by $Y_{t,N}$ and the sample proportion of successes in the t^{th} group by $P_{t,N} = Y_{t,N}/n_t$. The notational dependence on N is needed as we wish to establish limiting distribution results for the case when

$$\lim_{N\to\infty} \lambda_{t,N} = \lambda_t \qquad t = 1,2,\ldots,g \tag{3.1}$$

where the λ_t are fixed numbers satisfying $0 < \lambda_t < 1$, $t = 1,2,\ldots,g$.

We shall make use of the following theorem.

<u>Theorem 1</u> Let $\underset{\sim}{T}_N = (T_{1,N}, T_{2,N}, \ldots, T_{g,N})$ be a g-dimensional random variable such that

$$N^{\frac{1}{2}}(\underset{\sim}{T}_N - \underset{\sim}{\mu}) = \{N^{\frac{1}{2}}(T_{1,N} - \mu_1), \ldots, N^{\frac{1}{2}}(T_{g,N} - \mu_g)\} \tag{3.2}$$

has a limiting multivariate normal (MN) distribution with zero mean vector and covariance matrix $\underset{\sim}{X}$, that is

$$N^{\frac{1}{2}}(\underset{\sim}{T}_N - \underset{\sim}{\mu}) \xrightarrow{d} MN(\underset{\sim}{0}, \underset{\sim}{V}). \tag{3.3}$$

Let $h_1(x_1, x_2, \ldots, x_g), \ldots, h_\ell(x_1, x_2, \ldots, x_g)$ be ℓ functions of g variables where each function is totally differentiable. Then setting

$$\underset{\sim}{h}(\underset{\sim}{T}_N) = \begin{bmatrix} \underset{\sim}{h}_1(\underset{\sim}{T}_N) \\ \underset{\sim}{h}_2(\underset{\sim}{T}_N) \\ \\ \\ \underset{\sim}{h}_\ell(\underset{\sim}{T}_N) \end{bmatrix}, \tag{3.4}$$

we have

$$N^{\frac{1}{2}}\{\underset{\sim}{h}(\underset{\sim}{T}_N) - \underset{\sim}{h}(\underset{\sim}{\mu})\} \xrightarrow{d} MN(\underset{\sim}{0}, \underset{\sim}{H}\,\underset{\sim}{V}\,\underset{\sim}{H}') \tag{3.5}$$

where

$$H = ((H_{rs})) = (\partial h_r/\partial \mu_s)) \tag{3.6}$$

is an $\ell \times g$ matrix and

$$\partial h_r/\partial_{\mu S} = \{\partial h_r(x_1, x_2, ..., x_g)/\partial x_S\}_{\underset{\sim}{x}=\underset{\sim}{\mu}}. \tag{3.7}$$

We now apply theorem 1 taking $\underset{\sim}{T}_N = \underset{\sim}{P}_N$ where $\underset{\sim}{P}'_N = (P_{1,N}, P_{2,N}, \dots, p_{g,N})$ is the vector of sample proportions of successes. We have

$$E(P_{t,N}) = P_t \quad , \qquad var(p_{t,N}) = P_tQ_t/n_t \tag{3.8}$$

and

$$N^{\frac{1}{2}}(p_{t,N} - p_t) \xrightarrow{d} N(O, P_tQ_t/\lambda_t). \tag{3.9}$$

Since the $\{P_{T,N}\}$ are independently distributed, we have

$$N^{\frac{1}{2}}(\underset{\sim}{p}_N - \underset{\sim}{p}) \xrightarrow{d} MN(\underset{\sim}{O}, \underset{\sim}{V}) \tag{3.10}$$

where $\underset{\sim}{p}' = (P_1, P_2, ..., P_g)$ and

$$\underset{\sim}{V} = diag(P_1Q_1/\lambda_1 ..., P_gQ_g/\lambda_g) \ . \tag{3.11}$$

Using theorem 1, if $h_1(\underset{\sim}{p}_N), \dots, h_\ell(\underset{\sim}{p}_N)$ denote ℓ functions of the g variables pl,N ,......$p_{g,N}$ whose derivatives exist at $\underset{\sim}{p}_N = \underset{\sim}{p}$, then

$$N^{\frac{1}{2}}\{\underset{\sim}{h}(\underset{\sim}{p}_N) - \underset{\sim}{h}(\underset{\sim}{p})\} \xrightarrow{d} MN(\underset{\sim}{O}, \underset{\sim}{H}\,\underset{\sim}{V}\,\underset{\sim}{H}') \tag{3.12}$$

where $H = ((H_{rs}))$ is $\ell \times g$ and

$$H_{rs} = \left\{ \frac{\partial h_r(P_{1,N}, ..., P_{g,N})}{\partial p_{S,N}} \right\}_{\underset{\sim}{P}_N = \underset{\sim}{P}}, \qquad r = 1, ... , \ell, \quad s = 1, ..., g. \tag{3.13}$$

e above results apply to arbitrary functions $h(\cdot)$ which are differentiable at $\underset{\sim}{p}_N = \underset{\sim}{P}$. We now examine a special case which arises from consideration of the standardised residuals which we now denote by

$$R_{t,n} = \frac{Y_t - \_n_t P_t(\hat{\underset{\sim}{\theta}})}{\{n_t P_t(\hat{\underset{\sim}{\theta}}) Q_t(\hat{\underset{\sim}{\theta}})\}^{\frac{1}{2}}}, \qquad t = 1,2,\ldots,g$$

$$= \frac{N^{\frac{1}{2}} \lambda_{t,N}^{\frac{1}{2}} \{P_{t,N} - p_t(\hat{\underset{\sim}{\theta}})\}}{\{P_t(\hat{\underset{\sim}{\theta}}) Q_t(\hat{\underset{\sim}{\theta}})\}^{\frac{1}{2}}}, \qquad t = 1,2,\ldots,g. \tag{3.14}$$

Since the ML estimate $\hat{\underset{\sim}{\theta}} = \hat{\underset{\sim}{\theta}}$ (p_N) is a function of the sample proportions only, we write

$$\lambda_{t.N}^{-\frac{1}{2}} R_{t,N} = N^{\frac{1}{2}} h_t(\underset{\sim}{p}_N), \qquad t = 1,2,\ldots, g \tag{3.15}$$

where

$$h_t(P_{1,} P_2, \ldots, P_N) = \frac{P_{t.N} - P_t(\hat{\underset{\sim}{\theta}})}{\{P_t(\hat{\underset{\sim}{\theta}}) Q_t(\hat{\underset{\sim}{\theta}})\}^{\frac{1}{2}}}, \qquad t = 1,2,\ldots, \quad g. \tag{3.16}$$

We have

$$h_t(\underset{\sim}{p}) = h_t(\underset{\sim}{p}_N)_{\underset{\sim}{p}_N = \underset{\sim}{p}} = 0 \qquad t = 1,2,\ldots,g \tag{3.17}$$

Putting $\underset{\sim}{R}'_N = (R_{1,N}, R_{2,N} \ldots, R_{g,N})$ and $\underset{\sim}{\lambda}_N = \mathrm{diag}(\lambda_{1,N}, \lambda_{2,N}, \ldots\ldots \lambda_{g,N})$ use of (3.12) gives

$$\underset{\sim}{\lambda}_N^{-\frac{1}{2}} \underset{\sim}{R}_N \xrightarrow{d} MN(\underset{\sim}{O}, \underset{\sim}{H}\, \underset{\sim}{V}\, \underset{\sim}{H}') \tag{3.18}$$

where $\underset{\sim}{H} = ((H_{tu}))$ is $g \times g$ with

$$H_{tu} = \left\{ \frac{\partial h_t(\underset{\sim}{P}_N)}{\partial P_{u,N}} \right\}_{\underset{\sim}{P}_N = \underset{\sim}{P}}, \qquad u,t = 1,2,\ldots,g \ . \tag{3.19}$$

Now

$$\frac{\partial h_t(\underset{\sim}{P}_N)}{\partial p_{u,N}} = \{p_{t,N} - P_t(\hat{\underset{\sim}{\theta}})\} \frac{\partial}{\partial P_u} \{P_t(\hat{\underset{\sim}{\theta}}) Q_t(\hat{\underset{\sim}{\theta}})\}^{\frac{1}{2}}$$

$$+ \{P_t(\hat{\underset{\sim}{\theta}}) Q_t(\hat{\underset{\sim}{\theta}})\}^{-\frac{1}{2}} \frac{\partial}{\partial P_{u,N}} \{p_{t,N} - P_t(\hat{\underset{\sim}{\theta}})\}$$

so

$$
\begin{aligned}
\left\{\frac{\partial h_t(\underset{\sim}{P}_N)}{\partial P_{u,N}}\right\}_{\underset{\sim}{P}_N=\underset{\sim}{P}} &= (P_tQ_t)^{-\frac{1}{2}}\frac{\partial}{\partial P_{u,N}}\left\{P_{t,N}-P_t(\hat{\underset{\sim}{\theta}})\right\}_{\underset{\sim}{P}_N=\underset{\sim}{P}} \\
&= (P_tQ_t)^{-\frac{1}{2}}\left[\delta_{tu}-\sum_{j=1}^{\ell}\frac{\partial p_t}{\partial \theta_j}\,\frac{\partial \hat{\theta}_j}{\partial P_{u,N}}\right]_{\underset{\sim}{P}_N=\underset{\sim}{P}}
\end{aligned}
\tag{3.20}
$$

for u,t = 1,2,...,g where

$$\delta_{tu}=\begin{cases}1 & \text{if } t=u\\ 0 & \text{if } t\neq u\end{cases}.$$

If we set $P_t^{(j)}=\partial P_t/\partial\theta_j$ and $\hat{\underset{\sim}{\theta}}{}_j^{(t)}=\{\partial\hat{\theta}_j/\partial P_{t,N}\}_{\underset{\sim}{P}_N=\underset{\sim}{P}}$, we may write (3.20) as

$$
\begin{bmatrix}
\frac{\partial h_1}{\partial p_{1,N}} & \frac{\partial h_1}{\partial p_{2,N}} & \cdots & \frac{\partial h_1}{\partial p_{g,N}}\\
\frac{\partial h_2}{\partial p_{1,N}} & \frac{\partial h_2}{\partial p_{2,N}} & \cdots & \frac{\partial h_2}{\partial p_{g,N}}\\
\vdots & & & \\
\frac{\partial h_g}{\partial p_{1,N}} & \frac{\partial h_g}{\partial p_{2,N}} & \cdots & \frac{\partial h_g}{\partial p_{g,N}}
\end{bmatrix}
=
\begin{bmatrix}
(P_1Q_1)^{-\frac{1}{2}} & & & \underset{\sim}{O}\\
 & (P_2Q_2)^{-\frac{1}{2}} & & \\
 & \underset{\sim}{O} & \ddots & \\
 & & & (P_gQ_g)^{-\frac{1}{2}}
\end{bmatrix}
$$

$$
\cdot\left\{
\begin{bmatrix}
1 & & & \underset{\sim}{0}\\
 & 1 & & \\
\underset{\sim}{0} & & \ddots & \\
 & & & 1
\end{bmatrix}
-
\begin{bmatrix}
p_1^{(1)} & p_1^{(2)} & \cdots & p_1^{(\ell)}\\
p_2^{(1)} & p_2^{(1)} & \cdots & p_2^{(\ell)}\\
\vdots & & & \\
p_g^{(1)} & p_g^{(2)} & \cdots & p_g^{(\ell)}
\end{bmatrix}
\begin{bmatrix}
\hat{\theta}_1^{(1)} & \hat{\theta}_1^{(2)} & \cdots & \hat{\theta}_1^{(g)}\\
\hat{\theta}_1^{(1)} & \hat{\theta}_2^{(2)} & \cdots & \hat{\theta}_2^{(2)}\\
\vdots & & & \\
\hat{\theta}_\ell^{(1)} & \hat{\theta}_\ell^{(2)} & \cdots & \hat{\theta}_\ell^{(g)}
\end{bmatrix}
\right\}
$$

Putting $\underset{\sim}{\lambda}=\operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_g)$, we therefore have

$$\underset{\sim}{H}=(\underset{\sim}{\lambda}\,\underset{\sim}{v})^{-\frac{1}{2}}(\underset{\sim}{I}-\underset{\sim}{P}^{*1}\,\underset{\sim}{\theta}^{*}) \tag{3.21}$$

where

$$\underset{\ell\mathrm{x}g}{\underset{\sim}{P}^{*}} = ((P_t^{(j)})) \quad\text{and}\quad \underset{\sim}{\theta}^{*}=((\partial\hat{\theta}_j/\partial P_{u,N}))_{\underset{\sim}{P}_N=\underset{\sim}{P}}. \tag{3.22}$$

Noting that for a given model $\partial P_t / \partial\theta_j$ will be known, we see that to find $\underset{\sim}{H}$ we need to evaluate the matrix $\underset{\sim}{\theta}^*$ which consists of the derivatives $\partial\hat{\theta}_j / \partial p_u$ evaluated at $\underset{\sim}{P}_N = \underset{\sim}{P}'$. To do this, we consider the likelihood equations which have the form

$$\sum_{t=1}^{g} n_t \frac{\{P_{t,N} - P_t(\hat{\underset{\sim}{\theta}})\}}{p_t(\hat{\underset{\sim}{\theta}})Q_t(\hat{\underset{\sim}{\theta}})} \left\{\frac{\partial P_t(\underset{\sim}{\theta})}{\partial\theta_j}\right\}_{\underset{\sim}{\theta}=\hat{\underset{\sim}{\theta}}} = 0, \qquad j = 1, \ldots, \ell. \tag{3.23}$$

Differentiating the likelihood equations with respect to $p_{u,N}$ we get

$$\sum_{t=1}^{g} n_t \{P_{t,N} - P_t(\hat{\underset{\sim}{\theta}})\} \frac{\partial}{\partial P_{u,N}} \left[\frac{1}{P_t(\hat{\underset{\sim}{\theta}})Q_t(\hat{\underset{\sim}{\theta}})} \left\{\frac{\partial P_t(\underset{\sim}{\theta})}{\partial\theta_j}\right\}_{\underset{\sim}{\theta}=\hat{\underset{\sim}{\theta}}} \right]$$

$$+ \sum_{t=1}^{g} n_t \frac{1}{P_t(\hat{\underset{\sim}{\theta}})Q_t(\hat{\underset{\sim}{\theta}})} \left\{\frac{\partial P_t(\underset{\sim}{\theta})}{\partial\theta_j}\right\}_{\underset{\sim}{\theta}=\hat{\underset{\sim}{\theta}}} \left\{ - \sum_{k=1}^{\ell} \frac{\partial P_t(\underset{\sim}{\theta})}{\partial\hat{\theta}_k} \frac{\partial\hat{\theta}_k}{P_{u,N}} \right\}$$

$$+ \frac{n_u}{P_u(\hat{\underset{\sim}{\theta}})Q_u(\hat{\underset{\sim}{\theta}})} \left\{\frac{\partial P_u(\underset{\sim}{\theta})}{\partial\theta_j}\right\}_{\theta=\hat{\theta}} = 0 \tag{3.24}$$

for j = 1,2,..., ℓ, u = 1,2,...,g. Evaluating (3.24) at $\underset{\sim}{P}_N = \underset{\sim}{P}$ gives

$$\sum_{t=1}^{g} \frac{n_u}{P_u}{}_{(\underset{\sim}{\theta})Q_t(\underset{\sim}{\theta})} \frac{\partial P_t(\underset{\sim}{\theta})}{\partial\theta_j} \sum_{k=1}^{\ell} \frac{\partial P_t(\underset{\sim}{\theta})}{\partial\theta_k} \left\{\frac{\partial(\hat{\theta}_k)}{\partial P_{u,N}}\right\}_{\underset{\sim}{P}_N=\underset{\sim}{P}} = \frac{n_u}{P_u}{}_{(\underset{\sim}{\theta})Q_u(\underset{\sim}{\theta})} \frac{\partial P_u}{\partial\theta_j} \tag{3.25}$$

for j = 1,2,...,ℓ, u = 1,2,...,g, or equivalently $\underset{\sim}{P}_N = \underset{\sim}{P}$

$$\sum_{k=1}^{\ell} \theta_k^{(u)} \sum_{t=1}^{g} \frac{n_u}{P_u}{}_{(\underset{\sim}{\theta})Q_t(\underset{\sim}{\theta})} p_t^{(j)} p_t^{(k)} = \frac{n_u}{P_u}{}_{(\underset{\sim}{\theta})Q_u(\underset{\sim}{\theta})} \tag{3.26}$$

for j = 1,2,...,ℓ, u = 1,2,...,g. In matrix form these ℓ × g equations may be written as

$$\underset{\sim}{P}^* \underset{\sim}{V}^{-1} \underset{\sim}{P}^{*\prime} \underset{\sim}{\theta}^* = \underset{\sim}{P}^* \underset{\sim}{V}^{-1}$$

so

$$\underset{\sim}{\theta}^* = (\underset{\sim}{P}^* \underset{\sim}{V}^{-1} \underset{\sim}{P}^{*\prime})^{-1} \underset{\sim}{P}^* \underset{\sim}{V}^{-1} \tag{3.27}$$

and hence from (3.21),

$$\underset{\sim}{H}=(\lambda \underset{\sim}{V})^{-\frac{1}{2}}(\underset{\sim}{I}-\underset{\sim}{P}^{*\prime}(\underset{\sim}{P}^{*}\underset{\sim}{V}^{-1}\underset{\sim}{P}^{*\prime})^{-1}\underset{\sim}{P}^{*}\underset{\sim}{V}^{-1}.$$

Using the results that$(\underset{\sim}{\lambda}\underset{\sim}{V})'=\underset{\sim}{\lambda}\underset{\sim}{V}$ and $(\underset{\sim}{P}^{*}\underset{\sim}{V}^{-1}\underset{\sim}{P}^{*\prime})^{-1\prime}=(\underset{\sim}{P}^{*}\underset{\sim}{V}^{-1}\underset{\sim}{P}^{*\prime})^{-1\prime}$,
We have

$$\underset{\sim}{H}\underset{\sim}{V}\underset{\sim}{H}'=(\underset{\sim}{\lambda}\underset{\sim}{V})^{-\frac{1}{2}}\{\underset{\sim}{I}-\underset{\sim}{P}^{*}\underset{\sim}{V}^{-1}\underset{\sim}{P}^{*\prime})^{-1}\underset{\sim}{P}^{*}\underset{\sim}{V}^{-1}\}$$

$$.\underset{\sim}{V}\{\underset{\sim}{I}'-\underset{\sim}{V}^{-1\prime}\underset{\sim}{P}^{*\prime}(\underset{\sim}{P}^{*}\underset{\sim}{V}^{-1}\underset{\sim}{P}^{*\prime})^{-1}\underset{\sim}{P}^{*}\}(\underset{\sim}{\lambda}\underset{\sim}{V})^{-\frac{1}{2}}$$

$$=\underset{\sim}{\lambda}^{-1}-(\underset{\sim}{\lambda}\underset{\sim}{V})^{-\frac{1}{2}}\underset{\sim}{P}^{*\prime}(\underset{\sim}{P}^{*}\underset{\sim}{V}^{-1}\underset{\sim}{P}^{*\prime})^{-1}\underset{\sim}{P}^{*}(\underset{\sim}{\lambda}\underset{\sim}{V})^{-\frac{1}{2}} \tag{3.28}$$

We have therefore shown that the joint asymptotic distribution of the scaled standardised residuals $\{\lambda_{t,N}^{-\frac{1}{2}}R_{t,N}\}$ is multivariate normal with zero mean and covariance matrix given by (3.28).

The above asymptotic distribution result may be used to provide an approximation to the distribution of the standardised residuals $\{R_{t,N}\}$ for 'large' samples. Thus if we let $\underset{\sim}{V}_N = \mathrm{diag}((P_t\ Q_t/\lambda_{t,N}))$, we have

$$\underset{\sim}{R}_N \overset{\text{approx}}{\sim} MN(\underset{\sim}{O},\underset{\sim}{C}_{\underset{\sim}{R}_N}) \tag{3.29}$$

where $\underset{\sim}{C}_{\underset{\sim}{R}_N}$ denotes the approximate covariance matrix of $\underset{\sim}{R}_N$ and from (3.28) is given by

$$\underset{\sim}{C}_{\underset{\sim}{R}_N}=\underset{\sim}{\lambda}_N^{\frac{1}{2}}\{\underset{\sim}{\lambda}_N^{-1}-(\underset{\sim}{\lambda}_N\underset{\sim}{V}_N)^{-\frac{1}{2}}\underset{\sim}{P}^{*\prime}(\underset{\sim}{P}^{*}\underset{\sim}{V}_N^{-1}\underset{\sim}{P}^{*\prime})^{-1}\underset{\sim}{P}^{*}(\underset{\sim}{\lambda}_N\underset{\sim}{V}_N)^{-\frac{1}{2}}\}\underset{\sim}{\lambda}_N^{\frac{1}{2}}$$

$$=\underset{\sim}{I}_g-\underset{\sim}{V}_N^{-\frac{1}{2}}\underset{\sim}{P}^{*\prime}(\underset{\sim}{P}^{*}\underset{\sim}{V}_N^{-1}\underset{\sim}{P}^{*\prime})^{-1}\underset{\sim}{P}^{*}\underset{\sim}{V}_N^{-\frac{1}{2}}. \tag{3.30}$$

This covariance matrix has an interesting property which we will make use of later, namely

$$\mathrm{tr}(\underset{\sim}{C}_{\underset{\sim}{R}_N})=\mathrm{tr}(\underset{\sim}{I}_g)-\mathrm{tr}\{\underset{\sim}{V}_N^{-\frac{1}{2}}\underset{\sim}{P}^{*\prime}(\underset{\sim}{P}^{*}\underset{\sim}{V}_N^{-1}\underset{\sim}{P}^{*\prime})^{-1}\underset{\sim}{P}^{*}\underset{\sim}{V}_N^{-\frac{1}{2}}\}$$

$$=g-\mathrm{tr}\{\underset{\sim}{P}^{*}\underset{\sim}{V}_N^{-\frac{1}{2}}\underset{\sim}{V}_N^{-\frac{1}{2}}\underset{\sim}{P}^{*\prime}(\underset{\sim}{P}^{*}\underset{\sim}{V}_N^{-1}\underset{\sim}{P}^{*\prime})^{-1}\}$$

$$=g-\mathrm{tr}(\underset{\sim}{I}_\ell)$$

$$=g-\ell. \tag{3.31}$$

It follows that the sum of the large sample variances of the standardized residuals is equal to $g - \ell$, this result holding quite generally for any model assumed for the success probabilities $\{P_t\}$.

To utilise (3.30), the matrix $\underset{\sim}{P}^*$ is required which depends on the model specification for the $\{P_t\}$. For the linear logistic regression model we have

$$\frac{\partial P_t}{\partial \beta_j} = \frac{x_{tj} \exp(\underset{\sim}{x}'_t \underset{\sim}{\beta})}{\{1 + \exp(\underset{\sim}{x}'_t \underset{\sim}{\beta})\}^2} = x_{tj} P_t Q_t, \qquad j = 0,1,\ldots, k \tag{3.32}$$

giving

$$\underset{\sim}{P}^* = \begin{bmatrix} P_1Q_1 & P_2Q_2 & \cdots & P_gQ_g \\ P_1Q_1x_{11} & P_2Q_2x_{21} & \cdots & P_gQ_gx_{g1} \\ \cdot & & & \\ \cdot & & & \\ P_1Q_1x_{1k} & P_2Q_2x_{2k} & \cdots & P_gQ_gx_{gk} \end{bmatrix} \tag{3.33}$$

and

$$\underset{\sim}{P}^* \underset{\sim}{V}_N^{-1} \underset{\sim}{P}^{*\prime} = \begin{bmatrix} \sum_t w_t & \sum_t w_t x_{t1} & \cdots & \sum_t w_t x_{tk} \\ \sum_t w_t x_{t1} & \sum_t w_t x_{t1}^2 & \cdots & \sum_t w_t x_{t1} x_{tk} \\ \cdot & & & \\ \cdot & & & \\ \sum_t w_t x_{tk} & \sum_t w_t x_{t1} x_{tk} & \cdots & \sum_t w_t x_{tk}^2 \end{bmatrix} \tag{3.34}$$

where

$$w_t = \lambda_{t,N} P_t Q_t, \qquad t = 1,2,\ldots,g \tag{3.35}$$

When there is only one explanatory variable, we may write $x_{t1} = x_t$ $t = 1,2,\ldots,g$ and a straightforward calculation gives

$$(\underset{\sim}{P}^* \underset{\sim}{V}_N^{-1} \underset{\sim}{P}^{*\prime})^{-1} = \frac{1}{\sum_t w_t \sum w_t (x_t - \bar{x}_w)^2} \begin{bmatrix} \sum_t w_t x_t^2 & -\sum_t w_t x_t \\ -\sum_t w_t x_t & \sum_t w_t \end{bmatrix} \tag{3.36}$$

where $\bar{x}_w = \sum_t w_t x_t / \sum_t w_t$. The (i, j)th element in $\underset{\sim}{V}_N^{-\frac{1}{2}} \underset{\sim}{P}^{*\prime} (\underset{\sim}{P}^* \underset{\sim}{V}_N^{-1} \underset{\sim}{P}^{*\prime})^{-1} \underset{\sim}{P}^* \underset{\sim}{V}_N^{-\frac{1}{2}}$ is found to be

$$\frac{(w_i w_j)^{\frac{1}{2}} \sum_t w_t x_t^2 - (x_i + x_j) \sum_t w_t x_t + x_i x_j \sum_t w_t}{\sum_t w_t \sum_t w_t (x_t - \bar{x}_w)^2}$$

and the negative value of this element gives the asymptotic covariance between $R_{i,N}$ and $R_{j,N}$. The asymptotic value of the correlation coefficient between $R_{i,N}$ and $R_{j,N}$ is

$$\rho_{ij} = -(w_i w_j)^{\frac{1}{2}} \{\sum_t w_t x_t^2 - (x_i + x_j)\sum_t w_t x_t + x_i x_j \sum_t w_t\}/(C_i C_j)^{\frac{1}{2}} \tag{3.37}$$

where

$$C_i = \{\sum_t w_t \sum_t w_t (x_t - \bar{x}w)^2 - w_i(\sum_t w_t x_t^2 - 2x_i \sum_t w_t x_t + x_i^2 \sum_t w_t)\}^{\frac{1}{2}} \tag{3.38}$$

The asymptotic correlation structure of the residuals under the model given by (1.12) has been examined using (3.37) with $x_t = t-1$, $t = 1,2,...,g$, with $g = 3(1)10$. The results showed that the correlations were negative for most pairs of observed residuals but that small positive correlations occurred for residuals associated with groups having markedly different indices. These findings are illustrated in Table 1 for one particular parameter configuration.

Table 1

Correlation coefficients for the standardised residuals $\{R_t\}$ for the case $g = 10$, $t = 1, ...,10$, $\beta_0 = -2$, $\beta_1 - 0.2$ ($P_1 = 0.119$, $P_2 = 0.142$, $P_3 = 0.168$, $P_4 = 0.198$, $P_5 = 0.231$, $P_6 = 0.269$, $P_7 = 0.310$, $P_8 = 0.354$, $P_9 = 0.401$, $P_{10} = 0.450$)

	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}
R_1	-0.320	-0.276	-0.226	-0.172	-0.112	-0.048	0.025	0.109	0.213
R_2		-0.249	-0.207	-0.163	-0.114	-0.060	-0.001	0.068	0.153
R_3			-0.189	-0.154	-0.116	-0.074	-0.028	0.025	0.089
R_4				-0.144	-0.115	-0.089	-0.058	-0.023	0.018
R_5					-0.119	-0.105	-0.091	-0.076	-0.056
R_6						-0.122	-0.127	-0.132	-0.143
R_7							-0.167	-0.196	-0.236
R_8								-0.269	-0.343
R_9									-0.471

Turning to the deviance components, we write

$$D_t^2 = 2n_t\left[(\hat{P}_t + p_t - \hat{P}_t)\log\left(1 + \frac{p_t - \hat{P}_t}{\hat{P}_t}\right) + \{(1 - \hat{P}_t) - (p_t - \hat{P}_t)\}\log\left\{1 - \frac{(p_t - \hat{P}_t)}{1 - \hat{P}_t}\right\}\right]$$

and using a Taylor series expansion for the logarithmic terms we obtain after some simplification

$$D_t^2 = R_t^2(1+\varepsilon_t) \tag{3.39}$$

where

$$\varepsilon_t = 2\sum_{r=3}^{\infty} \frac{(p_t - \hat{P}_t)^{r-2}}{r(r-1)} \left\{ \frac{(1-\hat{P}_t)(-1)^r}{\hat{P}_t^{\,r-2}} + \frac{\hat{P}_t}{(1-\hat{P}_t)^{r-2}} \right\}, \tag{3.40}$$

Since $p_t - \hat{P}_t \xrightarrow{p} 0$ and $\hat{P}_t \xrightarrow{p} P_t$, we have $\varepsilon_t \xrightarrow{p} 0$ and hence the joint distribution of the $\{D_t\}$ is asymptotically the same as that of the joint distribution of the $\{R_t\}$.

4. Small Sample Properties Of The Goodness Of Fit Statistics Under The Logistic Regression Model

In section 2, we defined seven test statistics $R_{(\hat{\beta}_1)}$, $R_{(\hat{\beta}_2)}$, $D_{(\hat{\beta}_1)}$, $S_{(\hat{\beta}_3)}$, $S^*(\hat{\beta}_4)$, $T_{(\hat{\beta}_1)}$ and $T_{(\hat{\beta}_4)}$ which may be used for testing the goodness of fit of the logistic regression model. These may be considered as general purpose statistics since none were derived by considering specific alternatives to the logistic model . Each of the seven statistics has a limiting chi-square distribution under the logistic regression model. We give the proof for the first statistic only, this following straightforwardly from the results established in the previous section.

Putting $Z_{t,N} = \lambda_{t,N}^{-\frac{1}{2}} R_{t,N}$, $t = 1,\dots, g$ and $\underset{\sim}{Z}_N = (Z_{1,N}, Z_{2,N}, \dots, Z_{g,N})$,

we may write

$$R(\underset{\sim}{\hat{\beta}_1}) = \sum_{t=1}^{g} \lambda_{t,N} Z_{t,N}^2 = \underset{\sim}{Z}'_N \underset{\sim}{\lambda}_N \underset{\sim}{Z}_N \tag{4.1}$$

Hence

$$R(\underset{\sim}{\hat{\beta}_1}) \xrightarrow{d} \underset{\sim}{Z}' \underset{\sim}{\lambda} \underset{\sim}{Z} \tag{4.2}$$

where from (3.18)

$$\underset{\sim}{Z} \sim \underset{\sim}{M}\underset{\sim}{N}(0, \underset{\sim}{H}\underset{\sim}{V}\underset{\sim}{H}') \tag{4.3}$$

and $\underset{\sim}{H}\underset{\sim}{V}\underset{\sim}{H}'$ is given by (3.28). Now $\underset{\sim}{Z}'\underset{\sim}{\lambda}\underset{\sim}{Z}$ is distributed as x^2 iff $\underset{\sim}{\lambda}\,\underset{\sim}{H}\underset{\sim}{V}\underset{\sim}{H}'$ is idempotent. We have

$$\underset{\sim}{\lambda}\,\underset{\sim}{H}\,\underset{\sim}{V}\,\underset{\sim}{H}' = \underset{\sim}{I}_g - \underset{\sim}{\lambda}^{\frac{1}{2}}\,\underset{\sim}{V}^{-\frac{1}{2}}\,\underset{\sim}{P}^{*\prime}(\underset{\sim}{P}^{*}\underset{\sim}{V}^{-1}\,\underset{\sim}{P}^{*\prime})^{-1}\,\underset{\sim}{P}^{*}(\lambda\,\underset{\sim}{V})^{\frac{1}{2}}$$

$$= \underset{\sim}{I}_g - \underset{\sim}{A}, \quad \text{say.}$$

A simple calculation shows that $A^2 = A$ so A is idempotent and hence so is λHVH . We also have

$$\mathrm{tr}(\underset{\sim}{A}) = \mathrm{tr}\{\underset{\sim}{\lambda}^{\frac{1}{2}}\,\underset{\sim}{V}^{-\frac{1}{2}}\,\underset{\sim}{P}^{*\prime}(\underset{\sim}{P}^{*}\,\underset{\sim}{V}^{-1}\,\underset{\sim}{P}^{*\prime})^{-1}p^{*}(\underset{\sim}{\lambda}\,\underset{\sim}{V})^{-\frac{1}{2}}\}$$

$$= \mathrm{tr}\{p^{*}(\underset{\sim}{\lambda}\,\underset{\sim}{V})^{-\frac{1}{2}}\,\underset{\sim}{\lambda}^{\frac{1}{2}}\,\underset{\sim}{V}^{-\frac{1}{2}}\,\underset{\sim}{P}^{*\prime}(\underset{\sim}{P}^{*}\,\underset{\sim}{V}^{-1}\,\underset{\sim}{P}^{*\prime})^{-1}\}$$

$$= \mathrm{tr}\{\underset{\sim}{P}^{*}\,\underset{\sim}{V}^{-1}\,\underset{\sim}{P}^{*\prime}(\underset{\sim}{P}^{*}\,\underset{\sim}{V}^{-1}\,\underset{\sim}{P}^{*\prime})^{-1}\} = \mathrm{tr}(I_\ell).$$

Hence

$$\mathrm{tr}(\,\underset{\sim}{\lambda}\,\underset{\sim}{H}\,\underset{\sim}{V}\,\underset{\sim}{H}'\,) = \mathrm{tr}(\,\underset{\sim}{I}_g\,) - \mathrm{tr}(\,\underset{\sim}{I}_\ell\,) = g - \ell \tag{4.4}$$

showing that $\underset{\sim}{Z}'\underset{\sim}{\lambda}\underset{\sim}{Z} \sim \underset{\sim}{x}^2_{g-\ell}$. Hence $R_{(\hat{\underset{\sim}{\beta}}_1)}$ converges in distribution to $\underset{\sim}{x}^2_{g-\ell}$, this result holding for the general class of models considered in section 3.

The goodness of fit statistics under consideration have the desirable property that their asymptotic distributions do not depend on the unknown β and hence on the unknown $\{P_t\}$ when the logistic model holds. However for finite sample sizes this will not be the case and it is therefore important to investigate how rapidly the sampling distributions of the statistics approach the $x^2_{g\text{-}k\text{-}1}$ distribution. We have therefore performed a fairly large scale simulaton investigation for the case when the logi-istic regression model given by (1.13) holds, the group probabilities of success being

$$P_t = \exp\{\beta_0 + \beta_1(t\text{-}1)\}/[1 + \exp\{\beta_0 + \beta_1(t\text{-}1)\}] \qquad t = 1,\ldots,g\,. \tag{4.5}$$

The steps in the simulation consisted of:

(a) Specifying the model by fixing the values of β_0 ,β_1, and hence $\{P_t\}$.

(b) Generating and checking the binomial observations $\{y_t\}$ obtained by Monte Carlo sampling.

(c) Determining the ML, MC, WLS and modified WLS estimates of β_0, β_1.

(d) Constructing the empirical sampling distributions of the goodness of fit statistics and examining their moment and percentile properties.

Group numbers g = 5, 10 and samplesizes n = 25, 50 and 100 were used, A simulation run size of 1000 was used in each case. Checks on the generated binomial observations $\{y_t\}$ were made by applying the binomial dispersion test and the ordinary chi-square goodness of fit test to them. Three pairs of values (β_0, β_1) were examined for each value of g to give coverage for markedly different configurations for the group probabilities of success. The configurations used are shown in table 2.

Table 2

Parameter values (B_0, B_1) and group probabilities of success $\{P_t\}$ used in the simulation investigation.

					$\{P_t\}$		
g = 5							
(i)	β_0 = -2.0	β_1 = 0.4	0.119	0.168	0.232	0.310	0.401
(ii)	β_0 = H,0	β_1 = 0.5	0.269	0.378	0.500	0.623	0.731
(iii)	β_0 = 0.5	β_1 = 0.5	0.623	0.731	0.818	0.881	0.924
g = 10					$\{P_t\}$		
(iv)	β_0 = -2.0	β_1 = 0,2	0.119	0.142	0.168	0.198	0.231
			0.269	0.310	0.354	0.401	0.450
(v)	β_0 = -0,4	β_1 = 0.2	0.401	0.450	0.500	0.550	0.591
			0.646	0.690	0.731	0.769	0.802
(vi)	β_0 = 0.5	β_1 = 0.2	0.623	0.668	0.711	0.750	0.785
			0.818	0.846	0.870	0.891	0.908

In tables 3 to 9, values of the mean, variance, skewness and kurtosis coefficients for the empirical distributions of the seven goodness of fit statistics are given for the six parameter configurations shown in table 2. The statistics asymptotically have a chi-square distribution with g-2 degrees of freedom under the logistic regression models being considered, the values of the mean, variance, skewness and kurtosis coefficients of the asymptotic distribution being g-2, 2(g-2), 8/(g-2) and 12/(g-2) respectively.

Table 3

Moment measures for the empirical distributions of the $R^{(\hat{\beta}_{\underset{\sim}{1}})}$ statistic under six logistic regression models

Configuration		Mean	Variance	Skewness	Kurtosis
	(i)	3.001	5.838	2.65	3.86
n=25	(ii)	3-232	8.288	3.06	2.86
	(iii)	3.009	5.487	2.66	4.24
	(i)	3.115	5.888	1.69	1 .95
n=50	(ii)	3.011	5.426	1.67	2.04
	(iii)	2.951	5.877	2.92	4.27
	(i)	2.884	5.048	1.55	1 .56
n=100	(ii)	3.045	6.065	3.50	6.31
	(iii)	2.958	5.926	2.57	3.78
Asymptotic Values		3.000	6.000	2.67	4.00
	(iv)	7.992	14.53	0.76	1.38
n=25	(v)	8.103	16.11	0.93	1.11
	(vi)	8.065	15.65	0.84	1.20
	(iv)	8.124	16.29	1.13	1.70
n=50	(v)	8.119	15.24	0.74	0.85
	(vi)	7.982	15.81	0.85	1.12
	(iv)	8.186	16.19	1.34	2.27
n=100	(v)	7.696	14.86	0.96	1.20
	(vi)	8.099	15.46	0.63	0.47
Asymptotic Values		8.000	16.00	1.00	1.50

Table 4

Moment measures of the empirical distributions of the $D^{(\hat{\underset{\sim}{\beta}}_1)}$ statistic under six logistic regression models

Configuration		Mean	Variance	Skewness	Kurtosis
	(i)	3.159	7.037	2.93	4.05
n=25	(ii)	3.114	6.348	3.23	5.29
	(iii)	3.264	6.699	2.06	2.41
	(i)	3.181	6.393	1.90	2.38
m=50	(ii)	3.038	5.654	1.87	2.45
	(iii)	3.051	6.485	2.68	3.49
	(i)	2.907	5.185	1.56	1.54
n=100	(ii)	3.058	6.159	3.57	6.43
	(iii)	3.000	6.253	2.90	4.61
Asymptotic Values		3.000	6.000	2.67	4.00
	(iv)	8.351	17.04	0.80	1.11
n=25	(v)	8.286	17.46	1.00	1.18
	(vi)	8.532	18.27	0.81	1.21
	(iv)	8.259	16.90	1.07	1.51
n=50	(v)	8.217	15.98	0.78	0.89
	(vi)	8.183	17.78	1.18	1.94
	(iv)	8.266	16.88	1.37	2.24
n=100	(v)	7.739	15.16	0.95	1.16
	(vi)	8.190	16.01	0.65	0.53
Asymptotic Values		8.000	16.00	1.00	1.50

Table 5

Moment measures of the empirical distributions of the $R(\hat{\underset{\sim}{\beta}}_2)$ statistic under six logistic regression models

Configuration		Mean	Variance	Skewness	Kurtosis
	(i)	2.937	5.343	2.34	3.62
n=25	(ii)	3.131	5.771	2.06	2.75
	(iii)	2.940	4.990	2.49	3.99
	(i)	2.188	5.108	1.51	1.71
n=50	(ii)	2.845	5.420	2.84	4.04
	(iii)	2.922	5.257	1.80	2.30
	(i)	2.948	5.187	2.61	3.97
n=100	(ii)	3.034	6.047	2.23	2.89
	(iii)	2.931	5.120	2.29	3.56
Asymptotic Values		3.000	6.000	2.67	4.00
	(iv)	8.081	14.28	0.70	0.86
n=25	(v)	8.054	15.05	1.05	1 .42
	(vi)	7.987	14.42	0.84	1.50
	(iv)	7.747	13.57	0.69	1.10
n=50	(v)	8.245	15.81	0.91	1.14
	(vi)	7.963	14.82	0.99	1.46
	(iv)	8.105	15.83	0.85	1 .12
n=100	(v)	7.988	15.98	0.95	1.09
	(vi)	7.983	15.41	0.83	1 .00
Asymptotic Values		8.000	16.00	1.00	1.50

Table 6

Moment measures of the empirical distribution of the $S(\hat{\underset{\sim}{\beta}}_3)$ statistic under six logistic regression models

Configuration		Mean	Variance	Skewness	Kurtosis
	(i)	2.632	4.224	2.20	3.23
n=25	(ii)	3.023	5.021	1.71	2.19
	(iii)	2.376	3.519	2.29	3.61
	(i)	2.898	4.611	1.33	1.33
n-50	(ii)	2.801	5.053	2.48	3.34
	(iii)	2.753	4.349	1.51	1.83
	(i)	2.913	4.922	2.41	3.69
n=100	(ii)	3.012	5.878	2.15	2.75
	(iii)	2,880	4,758	2.01	3.02
Asymptotic Values		3.000	6.000	2.67	4.00
	(iv)	7.323	11.02	0.45	0.51
n=25	(v)	7.671	12.31	0.75	0.81
	(vi)	6.860	10.22	0.75	1.30
	(iv)	7.519	12.09	0.61	1.01
n=50	(v)	8.071	14.50	0.81	0.99
	(vi)	7.588	12.75	0.93	1.33
	(iv)	7.997	15.05	0.80	1.05
n=100	(v)	7.909	15.34	0.90	1.03
	(vi)	7.850	14,45	0.77	0.94
Asymptotic Values		8.000	16.00	1.00	1.50

Table 7

Moment measures of the empirical distributions of the $S^*(\hat{\beta}_4)$ statistic under six logistic regression models

Configuration		Mean	Variance	Skewness	Kurtosis
	(i)	2.802	4.764	2.03	2.96
n=25	(ii)	2.950	4.941	1.83	2.34
	(iii)	2.376	3.519	2.29	3.61
	(i)	2.897	4.696	1 .39	1.46
n=50	(ii)	2.755	4.955	2,63	3.65
	(iii)	2.842	4.901	1.66	1.96
	(i)	2.894	4.916	2.42	3.69
n=100	(ii)	2.985	5.778	2.15	2.75
	(iii)	2.879	4.862	2.14	3.25
Asymptotic Values		3.000	6.000	2.67	4.00
	(iv)	7.643	12.28	0.45	0.51
n=25	(v)	7.556	12.62	0.91	1 .15
	(vi)	7.626	12.86	0.70	1 .10
	(iv)	7.480	12.30	0.62	0.97
n=50	(v)	7.954	14.25	0.83	1 .01
	(vi)	7.695	13.45	0.88	1 .24
	(iv)	7.945	14.93	0.80	1 .05
n=100	(v)	7.841	15.13	0.90	1.04
	(vi)	7.819	14.50	0.78	0.94
Asymptotic Values		8.000	16.00	1 .00	1.50

Table 8

Moment measures of the empirical distributions of the $T(\hat{\beta}_1)$ statistic under six logistic regression models

Configuration		Mean	Variance	Skewness	Kurtosis
	(i)	3.418	10.53	9.38	18.79
n=25	(ii)	3.258	7.184	3.10	4.42
	(iii)	3.635	11 .18	7.05	11 .57
	(i)	3.208	7.269	4.65	8,82
n=50	(ii)	2.909	6.342	4.42	7.31
	(iii)	3.365	10.67	12.92	26.91
	(i)	3.059	6.385	4.93	9.08
n=100	(ii)	3.061	6.297	2.32	2.95
	(iii)	3.159	7.455	7.98	18.95
Asymptotic Values		3.000	6.000	2,67	4.00
	(iv)	9.434	32.21	4.16	7.97
n=25	(v)	8.643	24.10	3.48	6.33
	(vi)	9.846	36.66	2.63	3.68
	(iv)	8.413	21 .62	3.05	7.37
n=50	(v)	8.529	19.17	1 .47	2.39
	(vi)	9.114	31 .71	4.30	6.95
	(iv)	8.414	18,84	1.47	2.92
n=100	(v)	8.148	17.83	1.15	1 .38
	(vi)	8.451	19.88	1.27	1.66
Asymptotic Values		8.000	16.00	1.00	1.50

Table 9

Moment measures of the empirical distribution of the $T(\hat{\beta}_4)$ statistic under six logistic regression models

Configuration		Mean	Variance	Skewness	Kurtosis
	(i)	3.872	17.61	11.92	21.04
n=25	(ii)	3.374	8.384	3.61	5.18
	(iii)	4.323	21,44	8.70	14.27
	(i)	3.488	10.67	8.64	16.60
n=50	(ii)	2.967	7.148	6.26	11 .59
	(iii)	3.848	18.29	16.86	30.47
	(i)	3.208	8.287	11 .18	25.50
n=100	(ii)	3.088	6.502	2.36	2.99
	(iii)	3.415	10.92	19.76	48.02
Asymptotic Values		3.000	6.000	2.67	4.00
	(iv)	10.53	48.84	5.00	8.61
n=25	(v)	9.024	29.78	4.91	8.94
	(vi)	11 .49	61 .78	2.85	3.58
	(iv)	8.993	28.56	4.53	10.75
n=50	(v)	8.751	21.52	1 .97	3.88
	(vi)	10.12	49.02	6.44	10.16
	(iv)	8.135	21 .85	2.13	5.22
n=100	(v)	8.258	18.83	1.24	1.53
	(vi)	8.944	24.90	1.76	2.63
Asymptotic Values		8.000	16.00	1.00	1 .50

When the moment properties are compared with those of the approximating chi-square distribution, it is seen that the statistics fall into three groups, namely,

group1 : $R(\hat{\underset{\sim}{\beta}}_1)$, $R(\hat{\underset{\sim}{\beta}}_2)$, $D(\hat{\underset{\sim}{\beta}}_1)$

group2 : $S(\hat{\underset{\sim}{\beta}}_3)$, $S^*(\hat{\underset{\sim}{\beta}}_4)$

group3 : $T(\hat{\underset{\sim}{\beta}}_1)$, $T(\hat{\underset{\sim}{\beta}}_4)$.

For the statistics in group 1, the agreement between the means and variances of their sampling distributions and the coresponding moments of their asymptotic distribution is good, even for n as small as 25. The agreement for the skewness and kurtosis coefficients is less good, but this may be partly ascribed to the fairly large sampling errors in these coefficients for the run-size used in the investigation.

For the statistics in group 2, the agreement is less good, the means and variances being consistently smaller than those of the asymptotic distribution.

The sampling distributions of the statistics in group 3 clearly approach the asymptotic chi-square distribution much more slowly than the statistics in the other two groups. The means, and to a greater degree, the variances are larger than the corresponding moments of the chi-square distribution, but there is a marked improvement at $n = 100$.

Finally, it is seen that the chi-square approximations to the moments work best for configuration (ii) when $g = 5$ and configuration (v) when $g = 10$. In these configurations none of the $\{P_t\}$ are close to zero or to one.

Since the main application of the statistics under consideration is to test the goodness of fit of the logistic regression model, it is useful to assess the adequacy of approximations to their critical values based on the limiting chi-square distribution. In tables 10-14, the upper 10%, 5%, 2½% and 1% critical values of the statistics $R(\hat{\underset{\sim}{\beta}}_1)$, $D(\hat{\underset{\sim}{\beta}}_1)$, $R(\hat{\underset{\sim}{\beta}}_2)$, $S(\hat{\underset{\sim}{\beta}}_3)$ and $T(\hat{\underset{\sim}{\beta}}_1)$ are shown for the logistic regression models specified in table 2. Estimates of the actual significance levels associated with the chi-square approximating critical values are also given. The results for the statistic $S^*(\hat{\underset{\sim}{\beta}}_4)$ and $T(\hat{\underset{\sim}{\beta}}_4)$ are similar to those of $S(\hat{\underset{\sim}{\beta}}_3)$ and $T(\hat{\underset{\sim}{\beta}}_1)$ respectively and so are not presented here.

The findings are in agreement with those already noted for the moments. For the statistics $R(\hat{\underset{\sim}{\beta}}_1)$, $D(\hat{\underset{\sim}{\beta}}_1)$ and $R(\hat{\underset{\sim}{\beta}}_2)$, the agreement between the estimated actual significance levels and nominal significance levels is generally good even for the smallest sample size $n = 25$. Changes in the configurations for the $\{P_t\}$ have only marginal effects. For $S(\underset{\sim}{\hat{\beta}_3})$ the actual significance levels are appreciably less than the nominal levels for the smaller sample sizes, particularly for the configurations with small or large values for the $\{P_t\}$. For $T(\underset{\sim}{\hat{\beta}_1})$, it is seen that the chi-square approximating critical values give actual significance levels appreciably higher than the nominal values, particularly for configurations with small or large values for the $\{P_t\}$.

On the basis of these results, the use of the S or T classes of statistics cannot be recommended for sample sizes $n \leq 100$, if reasonable control of the significance level of the test is required.

Table 10

Upper critical values of the $R(\hat{\beta}_1)$ statistic under the logistic regression model (4.5). Actual significance levels associated with the chi-square approximating critical values are shown in parentheses

Configuration		α = 0.10	a = 0.05	α = 0.025	a = 0.01
	(i)	5.96(0.090)	7.57(0.048)	9.58(0.026)	11.08(0.10)
n=25	(ii)	6.16(0.097)	7.40(0.042)	9.25(0.025)	12.24(0.011)
	(iii)	5.98(0.090)	7.67(0.049)	9.03(0.021)	10.83(0.007)
	(i)	6.64(0.119)	7.82(0.051)	9.30(0.024)	10.79(0.070)
n=50	(ii)	6.21(0.095)	7.47(0.044)	8.74(0.019)	10.55(0.007)
	(iii)	6.16(0.097)	7.40(0.042)	9.25(0.025)	12.24(0.011)
	(i)	6.05(0.091)	7.37(0.039)	8.80(0.017)	10.24(0.003)
n=100	(ii)	5.93(0.091)	7.46(0.043)	8.92(0.022)	10,80(0.009)
	(iii)	6.08(0.093)	7.58(0.047)	9.17(0.018)	11.43(0.001)
Asymptotic		6.25	7.81	9.25	11.34
	(iv)	12.94(0.086)	15.13(0.042)	16.54(0.016)	18.03(0.006)
n=25	(v)	13.43(0.101)	15.62(0.053)	17.87(0.029)	20.57(0.013)
	(vi)	13.60(0.105)	15.36(0.045)	16,87(0.018)	19.93(0.009)
	(iv)	13.80(0.112)	15.57(0.053)	17.64(0.027)	20.47(0.012)
n=50	(v)	13.22(0.094)	15.62(0.054)	17.39(0.025)	19.52(0.008)
	(vi)	13.44(0.106)	15.35(0.047)	17.46(0.025)	19.41(0.007)
	(iv)	13.28(0.100)	15.67(0.034)	17.72(0,027)	19.93(0.010)
n=100	(v)	12.88(0.094)	15.17(0.045)	17.00(0.020)	19.27(0.006)
	(vi)	13.67(0.103)	15.83(0.059)	17.17(0.024)	19.00(0.007)
Asymptotic		13.36	15.51	17.53	20.09

Table 11

Upper critical values of the $D(\underset{\sim}{\hat{\beta}_1})$ statistic under the logistic regression model (4.5). Actual significance levels associated with the chi-square approximating critical values are shown in parentheses

Configuration		α = 0.10	α = 0.05	α=0.025	α = 0.01
	(i)	6.64(0.111)	8.26(0.063)	9.82(0.032)	13.13(0.016)
n=25	(ii)	6.48(0.111)	8.12(0.055)	9.71(0.030)	11.86(0.014)
	(iii)	6.59(0.118)	8.78(0.068)	10.38(0.038)	11.79(0.015)
	(i)	6.78(0.127)	8.14(0.055)	9.54(0.029)	11.08(0.10)
n=50	(ii)	6.17(0.098)	7.60(0.047)	8.85(0.023)	10.72(0.008)
	(iii)	6.48(0.111)	8.12(0.055)	9.71(0.030)	11.86(0.010)
	(i)	6.18(0.094)	7.54(0.041)	8.66(0.017)	10.36(0.004)
n=100	(ii)	5.95(0.091)	7.42(0.045)	8.83(0.022)	10.87(0.009)
	(iii)	6.34(0.102)	7.70(0.048)	9.18(0.024)	11,07(0,010)
Asymptotic		6.25	7.81	9.35	11.34
	(iv)	13.99(0.123)	16.25(0.065)	18.48(0.034)	20.57(0.015
n=25	(v)	13.64(0.111)	16.30(0.068)	18,10(0.037)	21.31(0.017)
	(vi)	14.36(0.131)	17.44(0.076)	19.73(0.048)	23.01(0.024)
	(iv)	13.40(0.103)	15.82(0.059)	18.15(0.030)	20.95(0.011)
n=50	(v)	13.61(0.108)	15.55(0.052)	17.30(0.023)	20.32(0.014)
	(vi)	13.65(0.108)	15.87(0.060)	18.25(0.028)	20.91(0.014)
	(iv)	13.03(0.092)	15.16(0.045)	17.92(0.029)	20.14(0.011)
n=100	(v)	13.30(0.100)	15.97(0.062)	17.87(0.029)	19.43(0.008)
	(vi)	13.58(0.108)	15.92(0.060)	17.86(0.030)	21.33(0.017)
Asymptotic		13.36	15.51	17.53	20.09

Table 12

Upper critical values of the $R_{(\hat{\underset{\sim}{\beta}}_2)}$ statistic under the logistic regression model (4.5), Actual significance levels associated with the chi-square approximating critical values are shown in parentheses

Configuration		α =0.10	α = 0.05	α = 0.025	α = 0.01
	(i)	6.17(0.099)	7.38(0.040)	8.47(0.016)	9.76(0.004)
n=25	(ii)	6.31(0.104)	7.49(0.041)	9.38(0.026)	11.60(0.011)
	(iii)	5.59(0.076)	7.05(0.034)	8.51(0.021)	10.31(0.006)
	(i)	6.23(0,100)	7.45(0.037)	8.41(0.014)	9.65(0.003)
n=50	(ii)	5.79(0.080)	7.41(0.044)	9.27(0.023)	10.67(0.006)
	(iii)	6.20(0.096)	7.14(0.039)	8,36(0.014)	10.10(0.004)
	(i)	5.75(0.084)	7.18(0.040)	8,58(0.017)	11.25(0.010)
n=100	(ii)	6.29(0.104)	7.75(0.048)	9.05(0.022)	11.47(0.011)
	(iii)	5.92(0.082)	7.08(0.039)	8.40(0.015)	10.14(0.006)
Asymptotic		6.25	7.81	9.35	11.34
	(iv)	13.05(0.091)	15.05(0.043)	17.20(0,023)	19,79(0.010)
n=25	(v)	12.93(0.091)	15.33(0.046)	17.90(0.027)	20.25(0.011)
	(vi)	13.13(0.091)	14.80(0.033)	16.26(0.016)	18.46(0.006)
	(iv)	12.47(0.073)	14.48(0.032)	16.28(0.018)	19.36(0.005)
n=50	(v)	13.37(0.102)	15.94(0.057)	17.33(0.025)	20.54(0.014)
	(vi)	13.18(0.093)	15.07(0.042)	16.26(0.017)	19.34(0.010)
	(iv)	13.41(0.105)	15.84(0.054)	17.59(0.028)	19.69(0.006)
n=100	(v)	13.34(0.100)	16.00(0.054)	17.96(0.031)	19.80(0.009)
	(vi)	13.25(0.099)	15.43(0,048)	17.10(0,020)	20.19(0.011)
Asymptotic		13.36	15.51	17.53	20.09

Table 13

Upper critical values of the $S(\hat{\underset{\sim}{\beta}}_3)$ statistic under the logistic regression model (4.5). Actual significance levels associated with chi-square approximating critical values are shown in parentheses

Configuration		α =0.10	α = 0.05	a = 0.025	α = 0.01
	(i)	5.52(0.066)	6.67(0.022)	7.47(0.008)	9.15(0.004)
n=25	(ii)	6.00(0.086)	7.10(0.039)	8.58(0.020)	11.05(0.009)
	(iii)	4.87(0.041)	5.94(0.016)	7.14(0.007)	8.52(0,001)
	(i)	6.02(0.087)	7.19(0.028)	7.97(0.011)	9.39(0,002)
n=50	(ii)	5.66(0.076)	7.20(0.041)	8.95(0.021)	10.41(0.005)
	(iii)	5.85(0.068)	6.70(0.023)	7.77(0.009)	8.74(0.002)
	(i)	5.71(0.083)	7.07(0.038)	8.30(0.015)	10.77(0.009)
n=100	(ii)	6.23(0.100)	7.69(0.047)	8.90(0.020)	11.30(0.010)
	(iii)	5.81(0.069)	6.93(0.035)	8.25(0.012)	10.01(0.004)
Asymptotic		6.25	7.81	9.35	11.34
	(iv)	11.82(0.056)	13.61(0.021)	15.03(0.006)	16.67(0.000)
n=25	(v)	12.24(0.068)	14.30(0.032)	15.92(0.017)	18.19(0.003)
	(vi)	11.09(0.031)	12.70(0.010)	13.63(0.005)	15.04(0.003)
	(iv)	11.99(0.060)	13.77(0.025)	15.35(0.014)	18.10(0.003)
n=50	(v)	13.09(0.091)	15.42(0.049)	16.64(0.020)	19.62(0.009)
	(vi)	12.51(0.072)	14.31(0.024)	15.45(0.013)	18.48(0.005)
	(iv)	13.20(0.091)	15.61(0.051)	17.30(0.020)	18.95(0.006)
n=100	(v)	13.20(0.091)	15.84(0.052)	17.48(0.025)	19.47(0.008)
	(vi)	12.90(0.089)	15.00(0.038)	16.40(0.019)	19.77(0.007)
Asymptotic		13.36	15.51	17.53	20.09

Table 14

Upper critical values of $T_{(\hat{\underset{\sim}{\beta}}_1)}$ statistic under the logistic regression model (4.5). Actual significance levels associated with the chi-square approximating critical values are shown in parentheses

Configuration		α = 0.10	α = 0.05	α=0.025	α = 0.01
	(i)	7.43(0.147)	8.80(0.084)	11.53(0.050)	14.75(0.026)
n=25	(ii)	6.71(0.115)	8.16(0.059)	10.52(0.035)	12.74(0.018)
	(iii)	7.48(0.145)	9.84(0.089)	13.08(0.057)	17.06(0.037)
	(i)	6.60(0.121)	7.96(0.056)	9.59(0.027)	12.13(0.015)
n=50	(ii)	5.88(0.090)	7.74(0.050)	10.10(0.034)	11.38(0.011)
	(iii)	6.99(0.139)	8.86(0.073)	10.77(0.043)	14.90(0.022)
	(i)	5.97(0.092)	7.66(0.045)	9.24(0.024)	12.56(0.013)
n=100	(ii)	6.51(0.111)	7.92(0.054)	9.24(0.025)	11.59(0.011)
	(iii)	6.57(0.116)	8.15(0.062)	9.48(0.028)	11.79(0.011)
Asymptotic		6.25	7.81	9.35	11.34
	(iv)	15.85(0.171)	19.91(0.107)	24.76(0.079)	29.16(0.050)
n=25	(v)	14.29(0.128)	17.72(0.078)	20.97(0.051)	25.91(0.033)
	(vi)	17.29(0.207)	22.46(0.144)	25.84(0.096)	31.06(0.064)
	(iv)	14.14(0.124)	16.95(0,068)	19.92(0.043)	22.72(0.022)
n=50	(v)	14.11(0.120)	16.99(0.073)	18.73(0.037)	23.04(0.020)
	(vi)	15.71(0.166)	19.28(0.106)	22.93(0,067)	31.85(0.043)
	(iv)	14.12(0.126)	16.33(0.059)	18.82(0.038)	21.76(0.015)
n=100	(v)	13.76(0.112)	16.07(0.061)	19.29(0.035)	21.38(0.017)
	(vi)	14.35(0.134)	17.07(0.079)	19.30(0.043)	22.34(0.021)
Asymptotic		13.36	15,51	17.53	20.09

5. The Extreme Standardised Residuals And Deviance Components

In a preliminary assessment of the adequacy of fit of a logistic regression model, the ordered values of the standardised residuals are usually computed and often plotted on normal probability paper. In particular, the extreme values

$$R_{max} = \max_{t=1,\ldots,g} R_t, \; R_{min} = \min_{t=1,\ldots,g} R_t, \;\; R_m = \max_{t=1,\ldots,g} \left|R_t\right| \quad (5.1)$$

would be examined to see if any outliers are present. If these statistics are to provide alternatives to the R goodness of fit statistic, approximations to at least the tail probabilities of their null distributions are required- The following non-rigorous approach which we illustrate for the statistic R_{max} appears to give useful approx— imations.

By the Bonferroni inequality we have

$$\sum_{t=1}^{g} P(R_t \geq r) - \sum_{t=1}^{g-1} \sum_{u=t+1}^{g} P(R_t \geq r, R_u \geq r) \leq P(R_{max} \geq r) \leq \sum_{t=1}^{g} P(R_t \geq r). \tag{5.2}$$

For the six logistic model configurations given in table 2, the large majority of pairs of residuals are negatively correlated and it seems reasonable to assume that this property will hold over a very wide range of configurations. We shall therefore assume that

$$\sum_{t=1}^{g-1} \sum_{u=t+1}^{g} P(R_t \geq r, R_u \geq r) < \sum_{t=1}^{g-1} \sum_{u=t+1}^{g} P(R_t \geq r) P(R_u \geq r). \tag{5.3}$$

This inequality may be replaced by the weaker inequality

$$\sum_{t=1}^{g-1} \sum_{u=t+1}^{g} P(R_t \geq r, R_u \geq r) < \frac{1}{2}(1 - g^{-1}) \left\{ \sum_{t=1}^{g} P(R_t \geq r) \right\}^2 . \tag{5.4}$$

For r large and hence $P(R_t \geq r)$ small, we use the approximation

$$P(R_{max} \geq r) \approx \sum_{t=1}^{g} P(R_t \geq r) \tag{5.5}$$

the error being less than $\frac{1}{2}(1 - g^{-1}) \left\{ \sum_{t=1}^{g} P(R_t \geq r) \right\}^2$ if the assumed inequality given by (5.3) is correct.

From (3.29), the residuals $\{R_t\}$ are approximately distributed as multivariate normal with zero means and covariance matrix given by (3.30). Since the elements in this matrix and in particular the variances depend on the $\{P_t\}$ which are unknown, we standardise the residuals using their average variance $1 - \ell/g$ and take

$$\left(\frac{g}{g-1}\right)^{\frac{1}{2}} R_t \overset{\text{approx}}{\sim} N(0,1). \tag{5.6}$$

Hence from (5.5) we obtain

$$P(R_{max} \geq r) \approx g\left[1 - \Phi\left\{\left(\frac{g}{g-\ell}\right)^{\frac{1}{2}} r\right\}\right] \tag{5.7}$$

where $\Phi(\bullet)$ is the c.d.f. of the N(0,1) distribution. If $r_{max}(\alpha)$ denotes the upper 100α % point of the distribution of R_{max} we therefore have the approximation

$$r_{max}(\alpha) \approx \left(\frac{g-\ell}{g}\right)^{\frac{1}{2}} \Phi^{-1}\left(1 - \frac{\alpha}{g}\right). \tag{5.8}$$

Similar arguments give the approximation

$$r_{min}(\alpha) \approx \left(\frac{g-\ell}{g}\right)^{\frac{1}{2}} \Phi^{-1}\left(1 - \frac{\alpha}{2g}\right) \tag{5.9}$$

to the lower 100α % point of distribution of R_{min} and

$$r_m(\alpha) \approx \left(\frac{g-\ell}{g}\right)^{\frac{1}{2}} \Phi^{-1}\left(1 - \frac{\alpha}{2g}\right) \tag{5.10}$$

to the upper 100α % point of the distribution of R_m, when the assumed logistic model is correct. Tables 15 and 17 give upper critical values of the statistics R_{max} and R_m respectively and table 16 gives the lower critical values for the statistic R_{min} for the logistic models specified in table 2. Estimates of the actual significance levels associated with the use of the approximate critical values given by (5.8), (5.10) and (5.9) as obtained by simulation are shown in brackets. The results indicate that the use of these simple approximations based on the use of average asymptotic variance to standardise the residuals will be satisfactory.

Since the $\{R_t\}$ and $\{D_t\}$ have the same asymptotic distribution, the approximations given by (5.8), (5.10) and (5.9) may also be used to approximate the upper critical values of the extremes Dmax and D_m and the lower critical values of D_{min} , respectively. Tables 18, 19 and 20 give the exact and approximate critical values of the statistics and the approximations again appear to give satisfactory results.

Table 15

Upper critical values of the R_{max} statistic under the logistic regression model (4.5). Actual significance levels associated with the approximating critical values given by (5.8) are shown in parentheses.

Configuration		α=0.10	α =0.05	α = 0.025	α =0.01
	(i)	1 .65(0.117)	1.95(0.073)	2.27(0.038)	2.45(0.030)
n=25	(ii)	1.67(0.119)	1.91(0.066)	2.15(0.039)	2.26(0.014)
	(iii)	1.52(0.073)	1.68(0.039)	1.86(0.013)	2.02(0.004)
	(i)	1.65(0.111)	1.90(0.070)	2.07(0.033)	2.34(0.016)
n=50	(ii)	1.61(0.110)	1.85(0.058)	2.05(0.033)	2.23(0.010)
	(iii)	1.58(0.095)	1.79(0.048)	1.96(0.021)	2.08(0.007)
	(i)	1.63(0.115)	1.91(0.072)	2.07(0.035)	2.33(0.016)
n=100	(ii)	1.61(0.111)	1.79(0.049)	2.02(0.026)	2.21(0.010)
	(iii)	1 .57(0.092)	1.74(0.042)	1.90(0.016)	2.14(0.006)
Approx .	(5.8)	1.60	1.80	2.00	2.23
	(iv)	2.11(0.108)	2.34(0.059)	2.63(0.030)	2.77(0.011)
n=25	(v)	2.00(0.080)	2.20(0.032)	2.43(0.018)	2.59(0.070)
	(vi)	1.90(0.045)	2.06(0.010)	2.19(0.001)	2.29(0.000)
	(iv)	2.11(0.108)	2.34(0.058)	2.55(0.028)	2.87(0.014)
n=50	(v)	2.02(0.087)	2.29(0.045)	2.47(0.023)	2.66(0.006)
	(vi)	2.01(0.072)	2.19(0.031)	2.33(0.012)	2,52(0.004)
	(iv)	2,13(0.116)	2.41(0.065)	2.69(0.040)	2.91(0.017)
n=100	(v)	2.00(0.075)	2.19(0.036)	2.39(0.014)	2.64(0.009)
	(vi)	2.01(0.080)	2.24(0.040)	2.41(0.015)	2.68(0.007)
Approx	(5.8)	2.08	2.30	2.51	2.76

Table 16

Lower critical values of the R_{min} statxistic under the logistic regression model (4.5). Actual significance levels associated with the approximating critical values given by (5.9) are shown in parentheses.

Configuration		α =0.10	α=0.05	α =0.025	α = 0.01
	(i)	-1.54(0.085)	-1.73(0.035)	-1.87(0.013)	-2.16(0.007)
n=25	(ii)	-1.60(0.101)	-1.82(0.053)	-2.00(0.023)	-2.33(0.012)
	(iii)	-1 .71(0.135)	-1.94(0.080)	-2.19(0.041)	-2.50(0,020)
	(i)	-1.65(0.111)	-1.83(0.055)	-2.09(0.029)	-2.38(0.016)
n=50	(ii)	-1.54(0.081)	-1.75(0.041)	-1.96(0.021)	-2.36(0.012)
	(iii)	-1.66(0.121)	-1.94(0.071)	-2.20(0.045)	-2.61(0.024)
	(i)	-1.61(0.108)	-1.81(0.051)	-2.00(0.023)	-2.17(0.008)
n=100	(ii)	-1.61(0.109)	-1.85(0.056)	-2.01(0.026)	-2.18(0.009)
	(iii)	-1 .64(0.110)	-1.84(0.058)	-1.99(0.023)	-2.34(0.012)
Approx.	(5.9)	-1.60	-1.80	-2.00	-2.23
	(iv)	-1.97(0.055)	-2.1(0.027)	-2.33(0.009)	-2.5(0.001)
n=25	(v)	-2.10(0.104)	-2.33(0.054)	-2.51(0.025)	-2.80(0.011)
	(vi)	-2.28(0.146)	-2.52(0.095)	-2,84(0.051)	-3.10(0.030)
	(iv)	-2.03(0.08)	-2.20(0.032)	-2.36(0.012)	-2.59(0.007)
n=50	(v)	-2.14(0.119)	-2.35(0.058)	-2.57(0.031)	-2.70(0.009)
	(vi)	-2,22(0,133)	-2.46(0.077)	-2.61(0,039)	-2.87(0.015)
	(iv)	-2.00(0.078)	-2.22(0.033)	-2.36(0.009)	-2.51(0.002)
n=100	(v)	-2.13(0.110)	-2.41(0.061)	-2.66(0.039)	-2.90(0.015)
	(vi)	-2.13(0.111)	-2.36(0.059)	-2.59(0.031)	-2.93(0.013)
Approx.	(5.9)	-2.08	-2.31	-2.51	-2.77

Table 17

Upper critical values of the R_m statistic under the logistic regression model (4.5). Actual significance levels associated with the approximating critical values given by (5.10) are shown in parentheses.

Configuration		α = 0.10	α =0,05	α = 0.025	α =0.01
	(i)	1.79(0.100)	1.99(0.048)	2.33(0.036)	2.45(0.017)
n=25	(ii)	1.83(0.105)	2.04(0.058)	2.23(0.033)	2.39(0.010)
	(iii)	1.82(0.108)	2.00(0.051)	2.21(0.030)	2.55(0.018)
	(i)	1.83(0.105)	2.01(0.055)	2.27(0.037)	2.61(0.015)
n=50	(ii)	1.73(0.085)	1.97(0.047)	2.14(0.025)	2.44(0.012)
	(iii)	1.82(0.103)	2.04(0.062)	2.29(0.032)	2.58(0.015)
	(i)	1.81(0.107)	2.02(0.055)	2.18(0.026)	2.43(0.012)
n=100	(ii)	1.79(0.096)	1.99(0.048)	2.11(0.020)	2.32(0.008)
	(iii)	1.78(0.093)	1.95(0.033)	2.17(0.022.)	2.49(0.011)
Approx.	(5.10)	1.80	2.00	2.18	2.40
	(iv)	2.23(0.082)	2.43(0.038)	2.63(0.023)	2.83(0.010)
n=25	(v)	2.29(0.092)	2.51(0.051)	2.72(0.027)	2.93(0.010)
	(vi)	2.27(0.088)	2.42(0.039)	2.65(0.023)	2.90(0.008)
	(iv)	2.24(0.088)	2.44(0.044)	2.65(0.022)	2.88(0.008)
n=50	(v)	2.27(0.086)	2,47(0.047)	2.66(0.023)	2,87(0,009)
	(vi)	2.27(0.094)	2,53(0.055)	2.7(0.026)	3.04(0.013)
	(iv)	2.29(0.097)	2.52(0.057)	2.73(0.032)	2.97(0.012)
n=100	(v)	2.26(0.087)	2.46(0.044)	2.72(0.030)	2.91(0.008)
	(vi)	2.33(0.110)	2.52(0.054)	2.73(0.032)	2.93(0.01)
Approx.	(5.10)	2.30	2.51	2.69	2.94

Table 18

Upper critical values of the D_{max} statistic under the logistic regression model (4.5). Actual significance levels associated with the approximating critical values given by (5.8) are shown in parentheses.

Configuration		α = 0.10	α=0.05	α =0.025	α =0.01
	(i)	1.58(0.098)	1.77(0.046)	1 .99(0.024)	2.24(0.011)
n=25	(ii)	1.56(0.086)	1.78(0.046)	1.97(0.023)	2.22(0.009)
	(iii)	1.84(0.177)	2.09(0.112)	2.37(0.061)	2.63(0.035)
	(i)	1.55(0.092)	1,81(0.051)	2.03(0.027)	2.30(0.012)
n=50	(ii)	1.60(0.103)	1.81(0,053)	2.10(0.031)	2.46(0.018)
	(iii)	1.76(0.144)	2.13(0.089)	2.34(0.063)	2.72(0.038)
	(i)	1.55(0.084)	1.77(0.044)	1 .94(0.019)	2.16(0.008)
n=100	(ii)	1.60(0.101)	1.85(0.059)	2.05(0.033)	2.25(0.012)
	(iii)	1.67(0.125)	1 .90(0.068)	2.15(0,039)	2.47(0.021)
Approx.		1.60	1.80	2.00	2.23
	(iv)	1.98(0.072)	2.19(0.034)	2.43(1.017)	2.62(0.01)
n=25	(v)	2.12(0.118)	2.46(0.073)	2.78(0.041)	3.04(0.28)
	(vi)	2.07(0.099)	2.29(0.049)	2.55(0.026)	2.74(0.10)
	(iv)	2.11(0.111)	2,32(0.053)	2.51(0.027)	2.82(0.014)
n=50	(v)	2.09(0.104)	2.36(0.056)	2.64(0.036)	2.98(0.018)
	(vi)	2.04(0.090)	2.34(0.056)	2.54(0.029)	2.68(0.008)
	(iv)	2.07(0.10)	2.34(0.056)	2.52(0.030)	2.67(0.000)
n=100	(v)	2.12(0.113)	2.35(0.058)	2.56(0.035)	2.77(0.011)
	(vi)	2.14(0.130)	2.41(0.069)	2.55(0.030)	2.78(0.012)
Approx.		2.08	2.30	2.51	2.76

Table 19

Lower critical values of the D_{min} statistic under the logistic regression model (4.5). Actual significance levels associated with the approximating critical values given by (5.9) are shown in parentheses.

Configuration		α = 0.10	α = 0.05	α =0.025	α = 0.0l
	(i)	-1.84(0.166)	-2.10(0.110)	-2.42(0.071)	-2.68(0.030)
n=25	(ii)	-1.57(0.096)	-1.90(0.057)	-2.14(0.036)	-2.45(0.019)
	(iii)	-1.54(0.096)	-1.79(0.047)	-1.96(0.020)	-2.20(0.009)
	(i)	-1.63(0.11)	-1.90(0.067)	-2.13(0,035)	-2.40(0.02)
n=50	(ii)	-1.67(0.121)	-1.90(0.070)	-2.18(0.037)	-2.51(0.023)
	(iii)	-1.53(0.083)	-1.76(0.045)	-2,05(0,030)	-2.30(0.012)
	(i)	-1.62(0.109)	-1.83(0.052)	-2.08(0.031)	-2.36(0.014)
n=100	(ii)	-1.57(0.087)	-1.78(0.044)	-2.04(0.025)	-2.29(0.012)
	(iii)	-1 .55(0.098)	-1.78(0.044)	-2.03(0.029)	-2.31(0.012)
Approx.		-1.60	-1 .80	-2.00	-2.23
	(iv)	-2.28(0.168)	-2.52(0.092)	-2,72(0.051)	-2.98(0.020)
n=25	(v)	-2.09(0.10)	-2.32(0.051)	-2.48(0.023)	-2.65(0.006)
	(vi)	-1.90(0.064)	-2.19(0.033)	-2.45(0.020)	-2.69(0.008)
	(iv)	-2.12(0.108)	-2.39(0.06)	-2.65(0.033)	-3.12(0.021)
n=50	(v)	-2.08(0.097)	-2.25(0.044)	-2.49(0.023)	-2.73(0.006)
	(vi)	-1 .94(0.059)	-2.12(0.028)	-2,37(0.015)	-2.55(0.004)
	(iv)	-2.07(0.097)	-2.27(0.041)	-2.48(0.021)	-2.87(0.011)
n=100	(v)	-2.11(0.104)	-2.31(0.050)	-2.53(0.027)	-2.83(0.011)
	(vi)	-2.05(0.091)	-2.28(0.045)	-2.55(0.027)	-2.95(0.012)
Approx.		-2.08	-2.31	-2.51	-2.77

Table 20

Upper critical values of the D_m statistic under the logistic regression model (4.5). Actual significance levels associated with the approximating critical values given by (5.10) are shown in parentheses.

Configuration		$\alpha = 0,10$	$\alpha = 0,05$	$\alpha = 0.025$	$\alpha = 0.01$
	(i)	2.00(0.139)	2.22(0.090)	2.52(0.055)	2.75(0.030)
n=25	(ii)	1.78(0.092)	2.01(0.053)	2.34(0.032)	2.55(0.014)
	(iii)	1.93(0.142)	2.16(0.075)	2.41(0.048)	2.63(0.020)
	(i)	1.82(0.105)	2.06(0.058)	2.29(0.035)	2.55(0.017)
n=50	(ii)	1.81(0.106)	2.11(0.062)	2.32(0.039)	2.55(0.022)
	(iii)	1.91(0.118)	2.22(0.085)	2.48(0.056)	2.73(0.031)
	(i)	1.79(0.089)	1.99(0.048)	2.16(0.025)	2.51(0.013)
n=100	(ii)	1.77(0.091)	2.01(0.051)	2.18(0.026)	2.41(0.012)
	(iii)	1.80(0.097)	2.05(0.060)	2.32(0.037)	2.55(0.015)
Approx.		1.80	2.00	2.18	2.40
	(iv)	2.40(0.124)	2.59(0.066)	2.80(0.037)	3.00(0.013)
n=25	(v)	2.36(0.114)	2.55(0.061)	2.84(0.038)	3.14(0.019)
	(vi)	2.23(0.079)	2.47(0.044)	2.67(0.023)	2.86(0.007)
	(iv)	2.32(0.105)	2.57(0.059)	2.90(0.038)	3.17(0.023)
n=50	(v)	2.27(0.096)	2.54(0.058)	2.75(0.031)	3.07(0.016)
	(vi)	2.22(0.084)	2.48(0.044)	2.62(0.016)	2.77(0.003)
	(iv)	2.28(0.095)	2.51(0.051)	2.67(0.025)	2.87(0.009)
n=100	(v)	2.31(0.104)	2.56(0.062)	2.74(0.031)	3.00(0.012)
	(vi)	2.31(0.104)	2.53(0.053)	2.74(0.029)	3.11(0.016)
Approx.		2.30	2.51	2.69	2.93

Acknowledgement

We are most grateful to Mr. Dennis Scrimshaw for his interest in this problem and the many helpful suggestions that he has provided.

References

Berkson, J. (1955) Maximum likelihood and minimum x^2 estimates of the logistic function . J. Amer. Statist. Assoc, 50, 130-162.

Cox, D.R. (1970) Analysis of Binary Data. Methuen, London.

Gart, J.J. and Zweifel, J.R. (1967). On the bias of various estimators of the logistic and its variance with application to quantal bioassay. Biometrika, 54, 181-187.

Pregibon, D. (1981). Logistic regression diagnostics. Annals of Statistics, 9, 705-724.