

Analysis of Quasi-Likelihood Models using SAS/IML

Il-Do Ha ¹

Abstract

The quasi-likelihood models which greatly widened the scope of generalized linear models are widely used in data analysis where a likelihood is not available. Since a quasi-likelihood may not appear to be an ordinary likelihood for any known distribution in the natural exponential family, to fit the quasi-likelihood models the standard statistical packages such as GLIM, GENSTAT, S-PLUS and so on may not directly applied. SAS/IML is very useful for fitting of such models. In this paper, we present simple SAS/IML(version 6.11) program which helps to fit and analyze the quasi-likelihood models applied to the leaf-blotch data introduced by Wedderburn(1974), and the problem with deviance useful generally to model checking is pointed out, and then its solution method is mention through the data analysis based on this quasi-likelihood models checking.

Key Words and Phrases: deviance, deviance residual, generalized linear models, Pearson residual, quasi-likelihood, variance function.

1. Introduction

The generalized linear models(GLMs) developed by Nelder and Wedderburn(1972) are a class of statistical models with distributions in natural exponential family (or GLM family) that generalizes classical normal linear models to include many other models which have been found useful in statistical data analysis. These other models include log-linear Poisson models for counts data, logit and probit Binomial models for proportions(or ratio of counts)data, and models for Gamma data with constant coefficient of variation rather than constant variance.

Further, Wedderburn's(1974) quasi-likelihood models(QLMs) which greatly widened the scope of generalized linear models by replacing the assumption of GLM family distribution by a much weaker assumption in which only the first two moments

¹Department of Statistics, Kyungsan University, Kyungpook, 712-240, Korea.

are defined are widely used in data analysis where a likelihood is not available. However to fit QLMs, differently from GLMs, the standard statistical packages such as GLIM, GENSTAT, and S-PLUS etc. may not directly applied because the quasi-likelihood(QL) may not appear to be an ordinary likelihood for any known distribution in the GLM family. As an alternative, SAS/IML may be useful for fitting of the QLMs.

In general the deviance residual rather than the Pearson residual, either unstandardized or standardized, is used for model checking procedures since its distributional properties are closer to the residuals arising in linear regression models, see Pierce and Schafer(1986), and McCullagh and Nelder(1989, pp.398). But the problem with the deviance in case that comes largely from the zero cell observations also induces the large deviance residuals corresponding to them, though the residuals may not be really large, while the Pearson statistic may be insensitive to this observations, and so may the Pearson residuals.

In this paper, to deal with the above-mentioned things, simple SAS/IML (version 6.11) program that helps to fit and analyze the QLMs applied to the leaf-blotch data introduced by Wedderburn(1974) is presented, and the problem with the deviance is pointed out, and then its solution method is mentioned through the data analysis based on this models checking. In Section 2, we review the QLMs and describe the fitting procedure which can be easily programed as SAS/IML, in Section 3 mention the model checking methods, in Section 4 the data applied to QLMs are analyzed through the model checking by using SAS/IML and the problem with the deviance is pointed out and examined, and conclusion is given by Section 5, and then in appendix the SAS/IML program is presented.

2. Quasi-likelihood and estimation procedure

Suppose that the i -th observation's response variable $Y_i(i = 1, \dots, n)$ have independent with the first two moments

$$E(Y_i) = \mu_i, \quad Var(Y_i) = \phi V(\mu_i), \quad (1)$$

where μ_i is some known mean function of a set of unknown parameters β_1, \dots, β_p , ϕ is unknown dispersion parameter, and $V(\cdot)$ is some known variance function. Then Wedderburn(1974)'s QL, strictly the quasi-log-likelihood, q_i for observation i is defined by

$$q_i(\mu_i; y_i) = \int_{y_i}^{\mu_i} \frac{y_i - t_i}{\phi V(t_i)} dt_i, \quad (2)$$

or equivalently by

$$\partial q_i(\mu_i; y_i) / \partial \mu_i = (y_i - \mu_i) / \phi V(\mu_i). \quad (3)$$

Note that the QL will be a true likelihood if there is a distribution of GLM type having $Var(Y_i) = \phi V(\mu_i)$.

In typical applications μ_i is determined by known covariates, x_{i1}, \dots, x_{ip} say, possibly through the following model equations

$$\eta_i = g(\mu_i) = x_i^t \beta \quad (i = 1, \dots, n), \tag{4}$$

where η_i is the linear predictor, $g(\cdot)$ is some known link function, $x_i^t = (x_{i1}, \dots, x_{ip})$ is the $1 \times p$ vector of the covariates for observation i , and $\beta = (\beta_1, \dots, \beta_p)^t$ is the $p \times 1$ vector of the model parameters. Models of the general form given by (1) and (4) are called by the name QLMs.

Denote $q_j'(\beta)$ by $\sum_{i=1}^n \partial q_i / \partial \beta_j$. Then Wedderburn's QL equations for β are given by

$$q_j'(\beta) = \sum_{i=1}^n \{W_i u_i x_{ij}\} / \phi = 0 \quad (j = 1, \dots, p), \tag{5}$$

where $W_i = \{V(\mu_i)[g'(\mu_i)]^2\}^{-1}$ is the weight function, $g'(\mu_i) = \partial g(\mu_i) / \partial \mu_i$, and $u_i = (y_i - \mu_i)g'(\mu_i)$. Since the equation (5) are non-linear functions of β , solving (5) for maximum quasi-likelihood(MQL) estimate $\hat{\beta}$ of β requires Fisher scoring approach. In fact, (5) have the same estimating equations as GLMs. The p-dimensional quasi-score function becomes

$$s(\beta) = (X^t W u) / \phi, \tag{6}$$

where X is the $n \times p$ model matrix composed of the elements x_{ij} , W is the $n \times n$ diagonal weight matrix with elements W_i on the main diagonal, u is the $n \times 1$ vector with the elements u_i . The $p \times p$ covariance matrix of $s(\beta)$, which is also the negative expected value of $\partial s(\beta) / \partial \beta$ (or the expected quasi-information matrix), is

$$I(\beta) = (X^t W X) / \phi. \tag{7}$$

Starting with an initial estimate $\hat{\beta}_0$ sufficiently close to $\hat{\beta}$, Fisher scoring iterations are given by

$$\hat{\beta}^{(k+1)} = \hat{\beta}^{(k)} + I^{-1}(\hat{\beta}^{(k)})s(\hat{\beta}^{(k)}), \quad k = 0, 1, 2, \dots, \tag{8}$$

where $I(\hat{\beta}^{(k)})$ and $s(\hat{\beta}^{(k)})$ are $I(\beta)$ and $s(\beta)$ evaluated at $\beta = \hat{\beta}^{(k)}$, respectively. Note that the dispersion parameter ϕ cancels out in the term $I(\hat{\beta}^{(k)})s(\hat{\beta}^{(k)})$. In fact, (8) can be expressed by

$$\hat{\beta}^{(k+1)} = (X^t \hat{W}^{(k)} X)^{-1} X^t \hat{W}^{(k)} \hat{z}^{(k)}, \quad k = 0, 1, 2, \dots, \tag{9}$$

where $\hat{W}^{(k)}$ and $\hat{z}^{(k)}$ are W and the adjusted dependent vector $z (= \eta + u)$ evaluated at $\beta = \hat{\beta}^{(k)}$, respectively. Since (9) give iteratively weighted least squares(IWLS), the MQL estimate $\hat{\beta}$ can be easily obtained from (9). A simple initial estimate $\hat{\beta}_0$ consists of using the data themselves $y = (y_1, \dots, y_n)^t$ as the initial estimate $\mu^{(0)}$ of

$\mu = (\mu_1, \dots, \mu_n)^t$. Thus the iterations (8) or (9) can begin with the following initial estimate

$$\hat{\beta}^{(0)} = (X^t W^{(0)} X)^{-1} X^t W^{(0)} z^{(0)}, \quad (10)$$

where $W^{(0)}$ and $z^{(0)}$ are W and z evaluated at $\mu = \mu^{(0)} (= y)$, respectively. At the initial step, adjustment may be required to the data, for example, the initial value $g(y)$ of z when g is the log link may be adjusted to replacing $y_i = 0$ by $y_i = 10^{-6}$.

We then can easily program as SAS/IML for the following estimation procedure to fit QLMs:

Step 0: Use (10) as an initial estimate $\hat{\beta}^{(0)}$ of $\hat{\beta}$.

Step 1: Calculate for $i = 1, \dots, n$

$$\hat{\mu}_i^{(0)} = g^{-1}(x_i^t \hat{\beta}^{(0)}).$$

Step 2: Calculate $\hat{W}^{(0)}$ and $\hat{z}^{(0)}$ in (9).

Step 3: Calculate the next estimate $\hat{\beta}^{(1)}$ to $\hat{\beta}$, as (9);

$$\hat{\beta}^{(1)} = (X^t \hat{W}^{(0)} X)^{-1} X^t \hat{W}^{(0)} \hat{z}^{(0)}$$

Step 4: Repeat from step 1 to step 3, replacing $\hat{\beta}^{(0)}$ with $\hat{\beta}^{(1)}$. Continue repeating this until convergence is (hopefully) achieved. One stop if $\hat{\beta}^{(0)}$ and $\hat{\beta}^{(1)}$ is close to zero.

McCullagh(1983) showed that under regular conditions, the MQL estimator $\hat{\beta}$ is consistent and asymptotically Normal with

$$Cov(\hat{\beta}) \simeq I^{-1}(\hat{\beta}) = \phi(X^t W X)^{-1}. \quad (11)$$

Clearly the MQL estimator is not affected by the value of ϕ , so that it can be calculated as if ϕ was known to be 1. But, to obtain its standard error(SE) some estimate of ϕ is required. Wedderburn suggested a moment estimate given by

$$\hat{\phi} = \frac{1}{n-p} \sum_{i=1}^n (y_i - \hat{\mu}_i) / V(\hat{\mu}_i) = X^2 / (n-p), \quad (12)$$

where X^2 is the Pearson statistic. Thus the SE of $\hat{\beta}_j (j = 1, \dots, p)$ is obtained from (j, j) -th main diagonal elements of (11) with (12), *i.e.*

$$SE(\hat{\beta}_j) = \{\hat{\phi}(X^t \hat{W} X)^{-1}\}_{(j,j)}^{1/2}, \quad (13)$$

where \hat{W} is W evaluated at $\beta = \hat{\beta}$.

3. Model Checking

The measure of overall discrepancy or goodness of fit for QLMs, similarly in GLMs, is the deviance, more strictly the quasi-deviance, $D = \sum_{i=1}^n d_i$, where d_i is the deviance component of the i -th observation, *i.e.*

$$d_i(y_i; \hat{\mu}_i) = 2\phi\{q_i(y_i; y_i) - q_i(\hat{\mu}_i; y_i)\} = 2 \int_{\hat{\mu}_i}^{y_i} \frac{y_i - t_i}{V(t_i)} dt_i, \quad (14)$$

or the Pearson X^2 statistic defined by (12). These two statistics follow asymptotic $\phi\chi^2$ distribution with degrees of freedom(df) $n - p$ under regular conditions, and so the greater D or X^2 for df the poorer the fit, *i.e.* lack of fit may exist, and also the estimation of ϕ based on the deviance may be possible, *i.e.* $\hat{\phi} = D/(n - p)$.

Reiduals can be used to explore the adequacy of fit of a model, in respect of variance function, link function and terms in the linear predictor, etc. In general, the residual used widely to checking of GLMs type, including QLMs, is deviance r_{d_i} or Pearson r_{p_i} defined by the signed square root of deviance component or the square root of Pearson statistic component, for the i -th observation, respectively, *i.e.*

$$r_{d_i} = \text{sign}(y_i - \hat{\mu}_i)\sqrt{d_i} \quad (15)$$

and

$$r_{p_i} = \frac{y_i - \hat{\mu}_i}{\sqrt{V(\hat{\mu}_i)}}. \quad (16)$$

Also, standardized deviance and Pearson residual are given by

$$r_{d_i}' = \frac{r_{d_i}}{\sqrt{\hat{\phi}(1 - h_{ii})}} \quad (17)$$

and

$$r_{p_i}' = \frac{r_{p_i}}{\sqrt{\hat{\phi}(1 - h_{ii})}}, \quad (18)$$

respectively, where h_{ii} is the i -th diagonal element of $n \times n$ projection matrix H defined by

$$H = W^{1/2}X(X^tWX)^{-1}X^tW^{1/2}.$$

Usually, examining of the following residual plots may help the model checking and indicate further model selection:

- (i) A link function or linear predictor can be checked as plot of standardized deviance (or Pearson) residual against $\hat{\eta}$ or against some function of fitted values.
- (ii) A variance function can be checked as plot of absolute standardized deviance(or Pearson) residual against some function of fitted values.
- (iii) A half-Normal or a full Normal plot makes normality

of standardized deviance (or Pearson) residual a check.

Note that the null pattern of (i) and (ii) shows no trend, in particular for (ii), an ill-chosen variance function will result in a trend in the mean and smoothing or Spearman rank correlation coefficient invariant under monotone transformations may help to see the trend more clearly, and *log* may be used as the function of the fitted values in (i) and (ii) because it is usually helpful in spreading out the points on the horizontal scale, and the null pattern of (iii) shows approximately linear, and that this two Normal plots provide the checking, in particular the half-Normal plot, of whether extreme values would be outliers in a sample of a given size.

In case of no showing special departure in examination of the (i) – (iii) for a model, the model will be reasonable, for detail see McCullagh and Nelder(1989, Chap.12).

4. Example

The data in Table 1, introduced and analyzed by Wedderburn(1974) and then reanalyzed by McCullagh and Nelder(1989, sec. 9.2.4), concerns the incidence of leaf-blotch on 10 different varieties of barley grown at 9 different sites in a variety trial in 1965. The response Y_{ij} , which is the percentage leaf area affected, can be considered by a continuous proportion in the interval $[0,1]$, for convenience of analysis. Intuitively, the responses will follow Bernoulli distribution, but since the column or row mean in Table 1 is increasing with the number of site or variety the site or variety effect may affect the variance component, *i.e.* dispersion parameter ϕ of Y_{ij} . Thus as a first step for this data analysis, they considered the following QL model

$$\text{Model 1 : } \eta_{ij} = \log\left(\frac{\mu_{ij}}{1 - \mu_{ij}}\right) = m + \alpha_i + \beta_j, \quad \text{Var}(Y_{ij}) = \phi\mu_{ij}(1 - \mu_{ij})$$

Table 1. Incidence of leaf-blotch on 10 varieties of barley grown at 9 sites: response is the percentage of leaf affected

Site	Variety										Mean
	1	2	3	4	5	6	7	8	9	10	
1	0.05	0.00	0.00	0.10	0.25	0.05	0.50	1.30	1.50	1.50	0.52
2	0.00	0.05	0.05	0.30	0.75	0.30	3.00	7.50	1.00	12.70	2.56
3	1.25	1.25	2.50	16.60	2.50	2.50	0.00	20.00	37.50	26.25	11.03
4	2.50	0.50	0.01	3.00	2.50	0.01	25.00	55.00	5.00	40.00	13.35
5	5.50	1.00	6.00	1.10	2.50	8.00	16.50	29.50	20.00	43.50	13.36
6	1.00	5.00	5.00	5.00	5.00	5.00	10.00	5.00	50.00	75.00	16.60
7	5.00	0.10	5.00	5.00	50.00	10.00	50.00	25.00	50.00	75.00	27.51
8	5.00	10.00	5.00	5.00	25.00	75.00	50.00	75.00	75.00	75.00	40.00
9	17.50	25.00	42.50	50.00	37.50	95.00	62.50	95.00	95.00	95.00	61.50
mean	4.20	4.77	7.34	9.57	14.00	21.76	24.17	34.81	37.22	49.33	20.72

$$(i = 1, \dots, 9; j = 1, \dots, 10),$$

where $\mu_{ij} = E(Y_{ij})$, α_i is the site effect, β_j is the variety effect, and $\alpha_1 = \beta_1 = 0$. Then the resulting QL and deviance for a single observation y , by (2) and (14), are given by

$$\phi q(\mu; y) = y \log\left(\frac{\mu}{1-\mu}\right) + \log(1-\mu), \quad (0 < \mu < 1, 0 \leq y \leq 1) \quad (19)$$

and

$$d(y; \hat{\mu}) = 2\left\{y \log\left(\frac{y}{\hat{\mu}}\right) + (1-y) \log\left(\frac{1-y}{1-\hat{\mu}}\right)\right\}, \quad (20)$$

respectively. The analysis of the Model 1 for this data is possible by using statistical package GLIM, GENSTAT, S-PLUS and SAS genmode procedure, etc. But, we here use SAS/IML(version 6.11) to fit and analyze the Model 1, through the estimating procedure described in Section 2. The deviance results in 6.126 on df 72 and Pearson statistic does in 6.392. Thus the estimate of ϕ is $\hat{\phi} = 6.392/72 = 0.089$. Since the data do not involve counts there is no reason to expect ϕ to be near 1.0.

In fact, however, as is shown in Figure 1 which provides checking of the variance function by using the standardized Pearson residual giving results similar to the standardized deviance in the Model 1, since the variance function in the Model 1 is increasing with the mean and since the Figure 1 shows a positive trend (also, Spearman rank correlation between the absolute standardized Pearson residuals and \log fitted values is $\hat{\rho}=0.52$ with p -value=0.0001), the current chosen variance function is not a satisfactory description of variability in this data. Therefore, the variance function in following QL model suggested by Wedderburn may be a natural choice:

$$\text{Model 2 : } \eta_{ij} = \log\left(\frac{\mu_{ij}}{1-\mu_{ij}}\right) = m + \alpha_i + \beta_j, \quad \text{Var}(Y_{ij}) = \mu_{ij}^2(1-\mu_{ij})^2$$

$$(i = 1, \dots, 9; j = 1, \dots, 10),$$

where $\mu_{ij} = E(Y_{ij})$ and $\alpha_1 = \beta_1 = 0$. Then QL and deviance for a single observation y in the Model 2, are given by

$$q(\mu; y) = (2y-1) \log\left(\frac{\mu}{1-\mu}\right) - \frac{y}{\mu} - \frac{1-y}{1-\mu}, \quad (0 < \mu < 1, 0 \leq y \leq 1) \quad (21)$$

and

$$d(y; \hat{\mu}) = 2\left\{(2y-1) \log\left(\frac{y(1-\hat{\mu})}{(1-y)\hat{\mu}}\right) + \frac{y-2y\hat{\mu}+\hat{\mu}}{\hat{\mu}(1-\hat{\mu})} - 2\right\}, \quad (y \neq 0), \quad (22)$$

respectively. The estimating procedure for Model 2 take a particularly simple form because the weighting matrix is constant *i.e.* identity. However, the fitting of Model 2 would not directly applied by the statistical packages but require GLIM or GENSTAT macros, since the QL (21) does not appear to be an ordinary likelihood for

any known distribution in the GLM family. We here fit and then analyze the Model (2) by using SAS/IML for the estimation procedure mentioned in Section 2. To do this, in initial step zero data in the Table 1 was replaced by 10^{-6} and then the convergence is achieved at 13-th iteration. The SAS/IML(version 6.11) program associated with the Model 2 fitting is given in appendix, including several useful calculations.

We now will check the Model (2) with methods based on residuals. Here, as measure for overall goodness of fit will prefer Pearson rather than deviance because, differently from (20) resulting in the Model 1, the deviance which is a total sum of terms each of (22) can not be defined in the usual way for the zero data in Table 1, without any adjustment and because the Pearson statistic may be insensitive to this zero data. We then will deal with this problems, with results calculated from SAS/IML for the Model 2 applied to the given data in Table 1. The Model 2 seems to be reasonable for given data in Tabel 1 since the Pearson statistic is 71.2 at df 72. However, with any adjustment the deviance also may provide a measure for overall goodness of fit. The deviance is approximate to infinite as the data is close to zero, but when zero data is replaced by 10^{-6} the deviance is 133.18, and so the deviance components which are corresponded to zero observations (variety 1, 2, 3, 7 at site 2, 1, 1, 3, respectively) give large values such as 8.86, 9.95, 12.56, 22.54, respectively. Excepting these 4 components, the deviance becomes 79.3, which makes no much difference Pearson statistic 71.2. Even when this zero data is replaced by 10^{-8} or 10^{-10} , it still becomes 79.3 except the corresponding components and also the Pearson statistic gives very insensitive result for this adjustment. Thus, we here can see facts that the problem with the deviance in case that comes largely from the $Y_{ij} = 0$ cell may be solved by considering the components of the deviance rather than just the deviance alone and that in such case also the Pearson rather than the deviance may prefer.

Further, we will more concretely check the Model (2) by using residual plots mentioned in the Section 3. We here will use the Pearson residual rather than the deviance for convenience sake because of control problem of large deviance residuals due to the zero data. From Figure 2, the variance function in the Model 2 is somewhat insensitive with the mean and the Figure 2, compared with the Figure 1, seems to be appear no trend (also, the Spearman rank correlation is $\hat{\rho} = -0.06$ with p -value=0.577), and so the new variance function may be suitable. The plot of standardized Pearson residual against the fitted linear predictor $\hat{\eta}$ shows form and trend similar to the Figure 2, see McCulluagh and Nelder(1989, pp.332). The half-Normal plot of Figure 3 shows approximately linear, though 4 large positive standardized Pearson residuals exist. These correspond, in decreasing order, to variety 4 at site 3 (3.37), variety 2 at site 6 (2.30), variety 5 at site 7 (2.81), and variety 6 at site 8 (2.51). we can see that there is no further evidence of systematic departures from the Model 2.

Table 2. the estimated variety effects and standard errors for the fitted Model 2 to the leaf-blotch data.

Variety									
1	2	3	4	5	6	7	8	9	10
0.000	-0.467	0.079	0.954	1.353	1.329	2.340	3.263	3.135	3.887
(0.000)	(0.471)	(0.471)	(0.471)	(0.471)	(0.471)	(0.471)	(0.471)	(0.471)	(0.471)

Thus, we can make an analysis conclusion with fitted Model 2 for given data in the Table 1. Tabel 2 gives the estimated variety effects and standard errors, fitted for the Model 2. Clearly there are differences between varieties; varieties 1-3 are most resistant to leaf-blotch and varieties 4-6 less so, while the remaining varieties 7-10 are much more susceptible. In other words, the proportion of leaf-blotch on 10 varieties increases with changing from variety 1 to variety 10.

5. Conclusion

We illustrated with an example that when a QL did not appear to be an ordinary likelihood for any known distribution in GLM family, SAS/IML could be usefully used to fit and analyze the QLMs, and also the associated SAS/IML(version 6.11) program was presented in appendix. In doing this models checking as the deviance, when zero cell observations are exist it may be a good idea to consider the components of the deviance rather than just the deviance alone, or the Pearson rather than the deviance may prefer. These facts were examined through the real data analysis based on the model checking in Section 4, but will be showed through simulation study. In addition, SAS/IML may be used to fit a broder class of models, such as not only QLMs for joint modelling of mean and dispersion introduced by Pregibon(1984) but also hierarchical generalized linear models(HGLMs) delveloped by Lee and Nelder(1996).

References

1. Lee, Y and Nelder, J. A. (1996). hierarchical generalized linear models (with discussions), *Journal of the Royal Statistical Society B*, 58, 619-678.
2. McCullagh, P. (1983). Quasi-likelihood functions, *Annals of Statistics*, 11, 59-67.
3. McCullagh, P. and Nelder, J. A. (1989). *Generalized linear models*, 2nd edition, London: Chapman and Hall.
4. Nelder, J. A. and Wedderburn, R. W. M. (1972), Generalized linear models, *Journal of the Royal Statistical Society A*, 135, 370-384.

5. Pierce, D. A. and Schafer, D. W. (1986), Residuals in generalized linear models, *Journal of the American Statistical Association*, 81, 977-986.
6. Pregibon, D. (1984), Review of generalized linear models, *Annals of Statistics*, 12, 1589-1596.
7. Wedderburn, R. W. M. (1974), Quasi-likelihood functions, generalized linear models and the Gauss-Newton method, *Biometrika*, 61, 439-447.

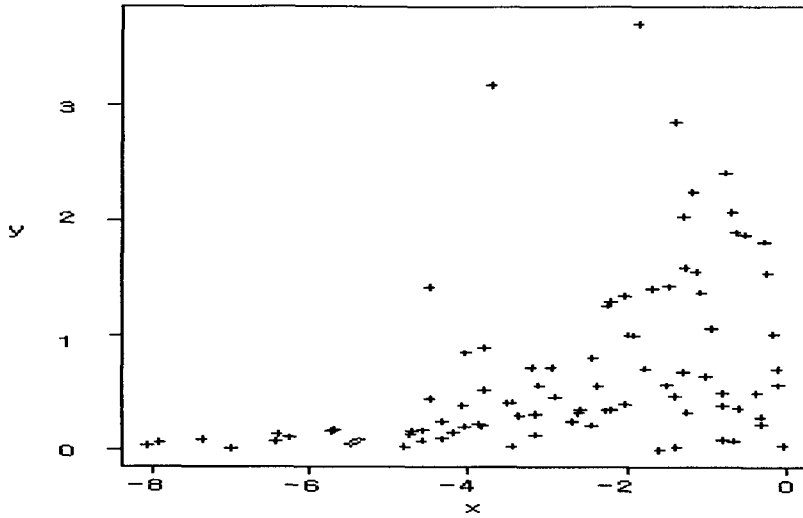


Figure 1. standardized Pearson residuals (y) plotted against the log fitted values (x) for the fitted Model 1 to the leaf-blotch data.

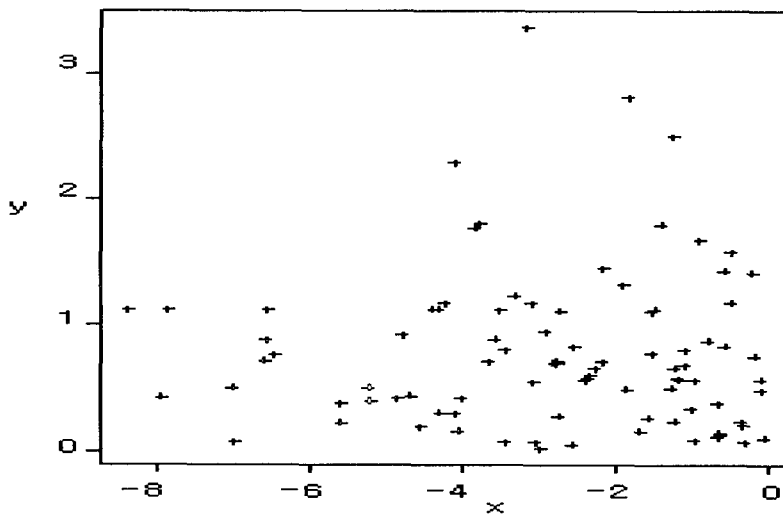


Figure 2. standardized Pearson residuals (y) plotted against the log fitted values (x) for the fitted Model 2 to the leaf-blotch data.

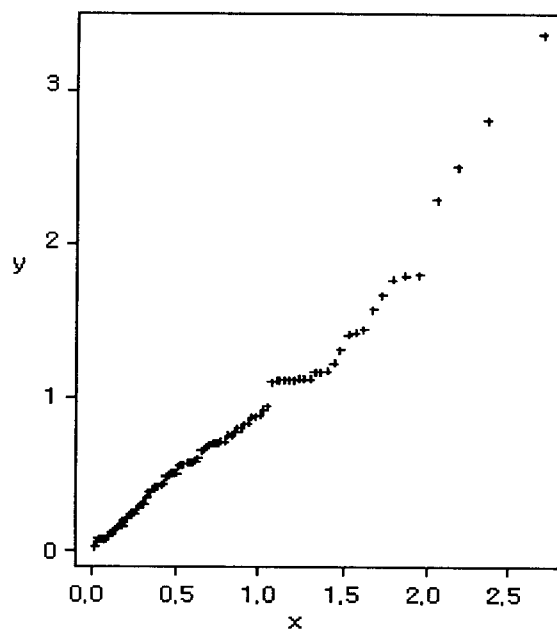


Figure 3. half-Normal plot: ordered absolute standardized Pearson residuals (y) plotted against the expected order statistics of a Normal sample (x) for the fitted Model 2 to the leaf-blotch data.

Appendix

```

proc iml workspace=2000;
  reset spill nolog ;

  /**** Construction of y and model matrix X with const, site, variety *****/
  y={0.05, 0.00, 0.00, 0.10, 0.25, 0.05, 0.50, 1.30, 1.50, 1.50,
    0.00, 0.05, 0.05, 0.30, 0.75, 0.30, 3.00, 7.50, 1.00, 12.70,
    1.25, 1.25, 2.50, 16.60, 2.50, 2.50, 0.00, 20.00, 37.50, 26.25,
    2.50, 0.50, 0.01, 3.00, 2.50, 0.01, 25.00, 55.00, 5.00, 40.00,
    5.50, 1.00, 6.00, 1.10, 2.50, 8.00, 16.50, 29.50, 20.00, 43.50,
    1.00, 5.00, 5.00, 5.00, 5.00, 5.00, 10.00, 5.00, 50.00, 75.00,
    5.00, 0.10, 5.00, 5.00, 50.00, 10.00, 50.00, 25.00, 50.00, 75.00,
    5.00, 10.00, 5.00, 5.00, 25.00, 75.00, 50.00, 75.00, 75.00, 75.00,
    17.50, 25.00, 42.50, 50.00, 37.50, 95.00, 62.50, 95.00, 95.00, 95.00};
  const=J(90,1,1);
  j0=J(10,8,0);
  j1=J(10,1,1);j2=J(10,1,2);j3=J(10,1,3);j4=J(10,1,4);
  j5=J(10,1,5);j6=J(10,1,6);j7=J(10,1,7);j8=J(10,1,8);
  j=(j1//j2//j3//j4//j5//j6//j7//j8);
  de=DESIGN(j);
  site=(j0//de);
  v1=J(1,9,0);v2=I(9);
  v3=(v1//v2);
  variety=REPEAT(v3,9,1);
  X=(const||site||variety);
  /*****Initial step *****/
  n=NROW(X);p=NCOL(X);
  y=y/100;
  y0=J(n,1,0);
  do i=1 to n;
    if y[i,1]=0 then do;
      y0[i,1]=0.000001;
    end;
    else do;
      y0[i,1]=y[i,1];
    end;
  end;
  z0=log10(y0/(1-y0));W0=I(n);
  beta_h=INV(X'*W0*X)*X'*W0*z0;
  print, 'intial=iter0' beta_h;
  /***** Iteration step *****/
  START imp(X,y,beta_h0,beta_h);
  n=NROW(X);p=NCOL(X);
  xbh=X*beta_h0;mu1=exp(xbh);mu=mu1/(1+mu1);
  u=(y-mu)/(mu*(1-mu));W=I(n);
  /*dft=X'*W*u;print,'q-like eqs' dft; */
  z=xbh+u;
  beta_h=INV(X'*W*X)*X'*W*z;
  FINISH;
  iter=1;
  do until (maxerr <= 0.001 | iter > 30); /** Stopping rule **/
  RUN imp(x,y,beta_h,beta_h1);
  /*print, iter beta_h1;*/
  err=abs( (beta_h-beta_h1)/beta_h );
  err1=err;
  r=rank(err);

```

```

err[r,1]=err1;
maxerr=err[p,1];
beta_h=beta_h1;
iter=iter+1;
end;
/***** Output of converged values *****/
print 'convergence is achieved at', iter;
xbh=X*beta_h1;mu1=exp(xbh);mu=mu1/(1+mu1);
u=(y-mu)/(mu*(1-mu));W=I(n);
dft=X'*W*u;
print,'q-like eqs' dft;
dinv=INV(X'*W*X);
se_beh=sqrt(vecdiag(dinv));t=beta_h1/se_beh;
print, iter beta_h1 se_beh t;
/**** Calculation of quasi-deviance (component) and Pearson statistic *****/
j=J(n,1,1);h1=y0*(1-mu);h2=(1-y0)*mu;
cqdevi1=(2*y0-1)*log(h1/h2);
cqdevi2= ( (y0-2*mu*y0 +mu)/(mu*(1-mu)) ) -2;
cqdevi=cqdevi1+cqdevi2;
cqdevi=2*cqdevi;

qdevi1=j'*( (2*y0-1)*log(h1/h2) );
qdevi2= j'*( (y0-2*mu*y0 +mu)/(mu*(1-mu)) ) -2*n;
qdevi=qdevi1+qdevi2;
qdevi=2*qdevi;

pearson=j'*( (y-mu)**2/(mu*(1-mu))**2 );
df=n-p;pie=pearson/df;piee=qdevi/df;
print,qdevi pearson df pie piee;
/**** Calculation of deviance and Pearson residual *****/
sign=y-mu;
do i=1 to n;
  if sign[i,1] > 0.0 then sign[i,1]=1;
  if sign[i,1] < 0.0 then sign[i,1]=-1;
  if sign[i,1] = 0.0 then sign[i,1]=0;
end;
dres=sign*sqrt(cqdevi);pres=(y-mu)/(mu*(1-mu));
H=(W#0.5)*X*dinv*X'*(W#0.5);hh=vecdiag(H);
spres=pres/sqrt(1-hh);sdres=dres/sqrt(1-hh);
eta=xbh;lmu=log(mu);
print, y0 mu lmu eta cqdevi spres sdres;

reset log;
quit;

```