Communications of the Korean Statistical Society Vol. 15, No. 6, 2008, pp. 993–1002

Use of Pseudo-Likelihood Estimation in Taylor's Power Law with Correlated Responses

Bumhee Park[1)], Heungsun Park[2)]

Abstract

Correlated responses have been widely analyzed since Liang and Zeger (1986) introduced the famous Generalized Estimating Equations(GEE). However, their variance functions were restricted to known quantities multiplied by scale parameter. In so many industries and academic/research fields, power-of-the-mean variance function is one of the common variance function. We suggest GEE-type pseudolikelihood estimation based on the power-of-the-mean variance using existing software and investigate it's efficiency for different working correlation matrices.

Keywords: Generalized estimating equations; GEE; power-of-the-mean; Taylor's power law; linear mixed model.

1. Introduction

In many disciplines such as biology, immunology, insectology, zoo-ecology and econometrics, the variances increases as the mean does according to the power of the mean. This powef-of-the-mean(POM) variance function is one of the common variance structures in heteroscedasticity (Carroll and Ruppert, 1988). Moreover, it is not only a general variance function for clinical data (Davidian and Giltinan, 1995; Carroll *et al.*, 1995), but also a wide spread animal behavior pattern in analyzing population dynamics of pest or animal, which is named Taylor's Power Law (TPL: Taylor, 1961; Perry, 1981; Park and Cho, 2004), since Taylor (1961) introduced the famous behavial pattern simply denoted by

$$V = a\mu^b, \tag{1.1}$$

where V is the variance, μ is the mean of subject counts and a, b are the species-specific unknown parameters.

Even though Taylor (1961) and many his successors have introduced various estimation methods for a and b, their approaches were limited to using simple linear regression with pairs of sample mean/variance, which is

$$\log\left(s_i^2\right) = \log a + b\log(\bar{x}_i) \tag{1.2}$$

1) Graduate student, Department of Radiology and Nuclear Medicine, Research Institute of Radiological Science, Yonsei University, Seoul 120-749, Korea.

2) Professor, Department of Statistics, Hankuk University of Foreign Studies, Yongin 449-791, Korea. Correspondence: hspark@hufs.ac.kr

out of different quadrats, $(\bar{x}_i, s_i^2)$, $i = 1, 2, \ldots, q$, for q quadrats.

However, the estimates a and b have been occasionally different for different geological regions, other environmental or physiological status (Southwood, 1978; Davis and Pedigo, 1989), which contradicts to the fact that the parameters are species-specific. This problem causes researchers to avoid to use TPL unless they have the same estimates and so the estimated TPL parameters were restrictly used on limited places or situations.

Park and Cho (2004) suggested quasilikelihood and variance function(QVF) method to estimate POM parameters(or TPL parameters), where they adopted quasilikelihood estimation (Wedderburn, 1974) in order to use covariate information within power-of-the-mean variance structure. The main sketch of QVF estimation is as follows:

Suppose $\boldsymbol{x}_i$ is a $p \times 1$ regressor vector and $\boldsymbol{\beta}$ is the corresponding coefficient vector. The mean count of animal/pest in the i^{th} observation, μ_i is given by

$$\log(\mu_i) = \boldsymbol{x}_i^T \boldsymbol{\beta} \tag{1.3}$$

and with a light of TPL the variance can be expressed as

$$V(\mu_i) = a\mu_i^b = a\left(\exp(\boldsymbol{x}_i^T \boldsymbol{\beta})\right)^b. \tag{1.4}$$

Given a and b, $\boldsymbol{\beta}$ can be estimated by setting the quasi-score to zero:

$$\boldsymbol{U}(\boldsymbol{\beta}) = \boldsymbol{D}^T \boldsymbol{V}^{-1}(\boldsymbol{Y} - \boldsymbol{\mu}) = \boldsymbol{0}, \tag{1.5}$$

where $\boldsymbol{D}$ is an $n \times p$ derivative matrix of $\boldsymbol{\mu}$ defined as $D_{ij} = \partial\mu_i/\partial\beta_j$, $\boldsymbol{V}$ is the $n \times n$ variance matrix, $\boldsymbol{V} = \mathrm{diag}\{v(\mu_i), \ldots, v(\mu_n)\}$, with $v(\mu_i) = a\mu_i^b$ and $\mu_i = \exp(\boldsymbol{x}_i^T \boldsymbol{\beta})$.

Also, once an estimate of $\boldsymbol{\beta}$ is obtained, the coefficients of the variance structure, a and b, can be estimated by using a pseudolikelihood. Park and Cho (2004) devised SAS macro(SAS institute) including two GENMOD procedures to estimate β, a and b simultaneously; nevertheless, as quasilikelihood assumes the independence of responses, their approach can be used only for independent responses.

In order to extend Park and Cho (2004)'s method to correlated count data in the same analogy, we use generalized estimating equation for estimating β and the multivariate pseudolikelihood for estimating a and b. The purpose of this paper is to compare QVF estimator and a newly presented GEE-type estimator for a and b in correlated count responses and to study their robustness property as working correlation changes.

In the next section, we use generalized estimating equation and multivariate pseudolikelihood(MPL) for estimating POM parameters in correlated data. Section 3 shows results of simulation for comparing it with QVF method between GVF and QVF. Conclusion and final remarks are in Section 4.

2. GEE and Variance Function(GVF) Model

Researchers are more often interested in analyzing temporal/spatial correlated data. Liang and Zeger (1986) formalized an approach using generalized estimating equations (GEE) to extend quasilikelihood scoring equation to a multivariate version. A basic feature of GEE is that the joint distribution of a subject response vector y_i, for the i^{th}

subject, $i = 1, 2, \ldots, k$, does not to be specified. Instead, the marginal mean and variance need to be specified with within-subject correlation assumption such as

$$\sum_{i=1}^{k} D_{1i}^T V_{1i}^{-1}(y_i - \mu_i) = 0, \tag{2.1}$$

for k subjects where $D_{1i} = \partial\mu_i/\partial\beta$ and

$$V_{1i} = \phi A_i^{\frac{1}{2}} R(\alpha) A_i^{\frac{1}{2}}, \tag{2.2}$$

where $A_i(\theta) = \text{diag}\{v(\mu_{i1}, \theta), v(\mu_{i2}, \theta), \ldots, v(\mu_{in_i}, \theta)\}$ and $R(\alpha)$ is a marginal correlation matrix.

For the estimation of β's in μ_i, Eq. (2.1) is solved; and they proposed method of moment(MOM) estimators for the nuisance parameters ϕ and α given β.

Liang *et al.* (1992) generalized the MOM estimation of ϕ and α by suggesting another estimating equations such as

$$\sum_{i=1}^{k} D_{2i}^T V_{2i}^{-1}(s_i - \sigma_i) = 0, \tag{2.3}$$

where s_i =vec$\{(y_i - \mu_i)(y_i - \mu_i)^T\}$, $\sigma_i = E(s_i)$ and $D_{2i} = \partial\sigma_i/\partial\alpha$. Eq. (2.3) can be expressed as a form of joint estimating equation with Eq. (2.1),

$$\sum_{i=1}^{k} D_i^T V_i^{-1} \begin{pmatrix} y_i - \mu_i \\ s_i - \sigma_i \end{pmatrix} = 0, \tag{2.4}$$

where now

$$D_i = \begin{pmatrix} \dfrac{\partial\mu_i}{\partial\beta^T} & 0 \\ 0 & \dfrac{\partial\sigma_i}{\partial\alpha^T} \end{pmatrix} \tag{2.5}$$

and

$$V_i = \begin{pmatrix} \text{var}(y_i) & 0 \\ 0 & \text{var}(s_i) \end{pmatrix} = \begin{pmatrix} V_{1i} & 0 \\ 0 & V_{2i} \end{pmatrix}. \tag{2.6}$$

This method is referred as GEE1 (Liang *et al.*, 1992). In a context of power-of-the-mean variance function, we can formalize the variance as $v(\mu_{ij}, \theta) = a(\mu_{ij})^b$ with $\theta = (ab)^T$. In short, for the correlated count data, the power-of-the-mean variance function can be estimated by GEE1 with $v(\mu_{ij}, \theta) = a(\mu_{ij})^b$.

The characteristic feature of GEE1 is to solve quadratic estimating equations for variance parameter, θ, correlation parameter, α. Eq. (2.3) is easily shown as a multivariate version of pseudolikelihood score function because it corresponds to the first derivative of the quadratic exponential family loglikelihood deduced from gaussian working assumption (Liang *et al.*, 1992).

Pseudolikelihood is to use normal likelihood even though the underlying distribution is not normal. Since Whittle (1961) introduced the concept in time series analysis, pseudolikelihood esimation is preferred to the method-of-moment for estimating variance parameters in GEE type estimating equations not only because it has smaller asymptotic variance (Crowder, 1985) but also it minimizes some object function referred as pseudolikelihood (Crowder, 1995) and provides the robustness to β estimation against working correlation misspecification (Wang and Carey, 2003, 2004).

Therefore, we suggest a GEE1 type estimating equation for β, α, a and b such as in Eq. (2.4) adopting pseudolikelihood method, which is referred to as Generalized estimating equation and Variance Function(GVF) estimation in this paper contrasting with Quasilikelihood and Variance Function(QVF) estimation in Park and Cho (2004). Good thing about GVF is to use the commercial software, NLINMIX macro(SAS Institute.) with a little extra programming modifications.

In origin, NLINMIX macro was developed for the purpose of estimating nonlinear mixed models by Littell *et al.* (1996) and consists of two iterative steps: (1) creating pseudo-data from modified residuals and (2) calling PROC MIXED procedure designed for linear mixed models.

Suppose y_i be a response vector and ϵ_i be an error. The nonlinear model is defined as

$$y_i = f(x_i, \beta) + \epsilon_i \tag{2.7}$$

and by Taylor's expansion at $\hat{\beta}$

$$y_i \simeq f\left(x_i, \hat{\beta}\right) + \left(\frac{\partial f\left(\hat{\beta}\right)}{\partial \beta}\right)\left(\beta - \hat{\beta}\right) + \epsilon_i, \tag{2.8}$$

which means

$$y_i - f\left(x_i, \hat{\beta}\right) + \left(\frac{\partial f\left(\hat{\beta}\right)}{\partial \beta}\right)\hat{\beta} \simeq \left(\frac{\partial f(\hat{\beta})}{\partial \beta}\right)\beta + \epsilon_i \tag{2.9}$$

and it turns into a linear model,

$$y_i^* = X_i^* \beta + \epsilon_i^*, \tag{2.10}$$

where $y_i^* = y_i - f(x_i, \hat{\beta}) + (\partial f(\hat{\beta})/\partial\beta)\hat{\beta}$, $X_i^* = (\partial f(\hat{\beta})/\partial\beta)$ and ϵ_i^* is an error term for the pseudo response y_i^*.

The final estimates $\hat{\beta}$, $\hat{\alpha}$ and $\hat{\theta} = (\hat{a}\ \hat{b})^T$ in NLINMIX macro solve a GEE as follows:

$$\sum_i^k X_i^{*T} V_i^{-1}\left(y_i^* - X_i^*\hat{\beta}\right) = 0, \tag{2.11}$$

$$\sum_i^k \frac{\partial l(y_i^*, \beta, \theta, \alpha)}{\partial(\alpha\ \theta)^T}\left(\hat{\beta}, \hat{\alpha}, \hat{\theta}\right) = 0. \tag{2.12}$$

Because PROC MIXED provides maximum likelihood estimators, θ, β, α, under normal assumption, these estimating equations are equivalent to

$$\sum_{i}^{k} X_i^{*T} V_i^{-1} \left(y_i - f\left(x_i, \hat{\beta}\right)\right) = 0, \tag{2.13}$$

$$\sum_{i}^{k} \frac{\partial l\left(y_i, \beta, \theta, \alpha\right)}{\partial \left(\alpha\ \theta\right)^T} \left(\hat{\beta}, \hat{\alpha}, \hat{\theta}\right) = 0, \tag{2.14}$$

which is GEE1 as in Eq. (2.3) using multivariate pseudo-likelihood estimation for α and θ. The followings are the algorithm of the suggested GVF modeling:

[STEP 1] Choose an initial value of b $(b = b_0\ )(i.e.\ b_0 = 1.0)$

[STEP 2] Suppose $E(y\,|\,x) = f(x, \beta)$, $V(y\,|\,x) = a(f(x, \beta))^{b_0}$ and correlation structure $R(\alpha)$. And evaluate $\hat{\beta}$, $\hat{a}$ by using NLINMIX macro and record log-likelihood, $L(b_0)$ corresponding to this evaluation.

[STEP 3] Increase b_0 within specific interval and redo [STEP 2] to record log-likelihood, $L(b_0)$.

[STEP 4] Iterate [STEP 2]$\sim$[STEP 3] for the various b_0 from 0.5 to 2.5.

[STEP 5] After fitting $L(b_0)$'s onto quadratic curve, obtain $\hat{b}$ maximizing $L(b_0)$.

[STEP 6] By using $\hat{b}$ from [STEP 5], assume that $E(y\,|\,x) = f(x, \beta)$, $V(y\,|\,x) = a(f(x, \beta))^{\hat{b}}$ and obtain $\hat{\beta}$, $\hat{\alpha}$, $\hat{a}$ using NLINMIX macro.

Within these steps, [STEP 4] is a profile-likelihood type approach in order to maximize the pseudo-likelihood given the different b_0 values. And the range of b_0, $[0.5, 2.5]$, seems reasonable because in most areas the common over-dispersion is proportional to $\mu^{1.2} \sim \mu^{2.5}$.

3. Simulation

Since Liang and Zeger (1986) insisted robustness of β estimation with different working correlation structures, there has been many simulation studies associated with working correlation in GEE.

Liang *et al.* (1992) illustrated GEE1 and GEE2 have β consistency regardless of working correlation like as Liang and Zeger's GEE. But they also suggested that unbalanced repeated responses and misspecified covariance structure led β estimation to less efficient. In other studies, if working correlation cannot be expected correctly, to use independence working correlation can be more effective for β estimation (Pepe and Anderson, 1994; Fitzmaurice, 1995). Mancl and Leroux (1996) found the choice of $R(\alpha)$, working correlation, cluster size and correlation degree might have an impact on β efficiency, so that it is recommended to use Independence structure as a safe strategy in GEE (Sutradhar and Das, 1999). As for pseudolikelihood estimation combined with GEE1 as in GVF, Wang and Carey (2003) compared MOM with pseudolikelihood and recommended to use

Table 3.1: The relative efficiency of GVF to QVF for estimating β_0 with different working correlation matrices($\beta_0 = 1.0,\ \beta_1 = 4.0,\ a = 1.5,\ b = 1.5$).

sub.	QVF	GVF			
		IND	CS	AR(1)	TOEP(2)
		The true correlation is independent			
25	1.000	5.250	5.250	5.236	5.250
100	1.000	4.728	4.728	4.682	4.682
500	1.000	4.859	4.835	4.859	4.859
		The true correlation is CS with 0.3			
25	1.000	4.259	4.269	4.279	4.279
100	1.000	3.927	3.927	3.892	3.892
500	1.000	4.023	4.004	3.967	3.986
		The true correlation is CS with 0.8			
25	1.000	2.910	3.734	3.660	3.573
100	1.000	2.861	3.613	3.580	3.485
500	1.000	2.871	3.527	3.475	3.457

Table 3.2: The relative efficiency of GVF to QVF for estimating β_1 with different working correlation matrices($\beta_0 = 1.0,\ \beta_1 = 4.0,\ a = 1.5,\ b = 1.5$).

sub.	QVF	GVF			
		IND	CS	AR(1)	TOEP(2)
		The true correlation is independent			
25	1.000	11.887	11.863	11.838	11.863
100	1.000	10.480	10.480	10.396	10.396
500	1.000	10.875	10.829	10.829	10.829
		The true correlation is CS with 0.3			
25	1.000	9.745	9.725	10.526	9.704
100	1.000	8.760	8.760	8.688	8.688
500	1.000	9.094	9.017	8.902	8.940
		The true correlation is CS with 0.8			
25	1.000	5.934	7.398	7.184	7.378
100	1.000	5.705	7.037	6.915	6.983
500	1.000	5.779	6.923	6.737	6.961

pseudolikelihood with AR(1) structure for more effective β estimation. In spite of many simulation studies for GEE, GEE1, or pseudolikelihood, the prime interests lie on the β estimation not on the variance parameters, a and b.

In this simulation, we compare GVF estimator with QVF estimator for variance parameters specially focused on power-of-the-mean variance structure with correlated count data. Additionally, we investigate the robustness of variance parameter estimation as working correlation changes.

For simulation, we generated lognormal poisson mixture responses, y_{ij} (j^{th} observation of the i^{th} subject) having mean, $\exp(1.0+4.0x_{ij})$ and variance, $a(\exp(1.0+4.0x_{ij}))^b$, with (a, b)=$(1.5, 1.5)$ and $x_{ij} = j$ for each subject. To make observations correlated, Compound Symmetry(CS) correlation structured is considered within a subject ($\rho =$ 0, 0.3, 0.8). Madsen and Dalthorp (2007)'s algorithm to build correlated count data was used.

The simulation result on Table 3.1~3.2 illustrate that GVF is more effective than

Table 3.3: The relative efficiency of GVF to QVF for estimating a with different working correlation matrices($\beta_0 = 1.0,\ \beta_1 = 4.0,\ a = 1.5,\ b = 1.5$).

sub.	QVF	GVF			
		IND	CS	AR(1)	TOEP(2)
		The true correlation is independent			
25	1.000	0.705	0.700	0.408	0.394
100	1.000	1.000	0.786	0.162	0.161
500	1.000	1.000	0.183	0.230	0.030
		The true correlation is CS with 0.3			
25	1.000	1.000	0.639	0.549	0.671
100	1.000	1.000	0.597	0.208	0.271
500	1.000	1.000	0.169	0.173	0.048
		The true correlation is CS with 0.8			
25	1.000	1.000	0.513	0.500	0.984
100	1.000	1.000	0.261	0.326	0.415
500	1.000	1.000	0.056	0.054	0.079

Table 3.4: The relative efficiency of GVF to QVF for estimating b with different working correlation matrices($\beta_0 = 1.0,\ \beta_1 = 4.0,\ a = 1.5,\ b = 1.5$).

sub.	QVF	GVF			
		IND	CS	AR(1)	TOEP(2)
		The true correlation is independent			
25	1.000	11.456	11.456	12.291	11.419
100	1.000	1.000	5.772	5.562	5.602
500	1.000	1.000	0.088	0.771	0.771
		The true correlation is CS with 0.3			
25	1.000	1.000	9.163	11.004	9.256
100	1.000	1.000	3.680	5.975	3.758
500	1.000	1.000	0.483	0.492	0.497
		The true correlation is CS with 0.8			
25	1.000	1.000	5.555	5.759	6.395
100	1.000	1.000	0.915	8.085	1.142
500	1.000	1.000	0.087	3.586	0.112

QVF because GVF permits the within-subject correlation whereas QVF deals with only independent observation. In addition, GVF estimator is better than QVF regardless of the chosen type of working correlation. It is noteworthy that, in spite of the same independence assumption, GVF-IND in a multivariate approach outperforms QVF in univariate version and IND structure is nearly effective as other non-diagonal correlation matrices such as AR(1), TOEP(2) and CS.

Moreover, Table 3.1$\sim$3.2 illustrates not only that the efficiency of GVF over QVF for estimating mean parameter but also that the independence working correlation is no longer effective than other working structures, which contradicts to the results by several authors (Pepe and Anderson, 1994; Fitzmaurice, 1995; Mancl and Leroux, 1996; Sutradhar and Das, 1999), in which they used MOM estimators are not pseudolikelihood.

Another interesting aspect is that GVF-IND is outperformed by other correlation matrices if responses are highly correlated($\rho = 0.8$). Therefore we could say that GVF-IND is used for low or moderate correlation, while other non-diagonal correlation structure

Table 4.1: The estimates of parameters in GVF modeling for epileptic seizure count data.

parameter	GVF			
	IND	CS	AR(1)	TOEP(2)
Intercept	0.557	0.559	0.482	0.498
Trt	−0.196	−0.195	−0.225	−0.220
Base	0.026	0.026	0.026	0.027
Age	0.017	0.017	0.019	0.018
a	0.954	1.404	0.970	0.893
b	1.862	1.812	1.849	1.884

should be considered with highly correlated data for better β-estimation.

On the contrary, Table 3.3 $\sim$ 3.4 show that QVF is better than GVF for estimating a variance parameter, a, while MSE($\hat{b}_{GVF}$) is larger than MSE($\hat{b}_{QVF}$). However, the differences are statistically negligible(significance level $= 0.05$) and as the number of subjects increases, QVF becomes more effective then GVF even in estimating b, so that we may conclude that QVF is better than GVF for estimating a, b in POM variance function whether the responses are correlated or not.

4. Example

Breslow and Clayton (1993) reanalyzed the same data that Thall and Vail (1990) did with introducing the their GLMM method. The clinical trial of 59 epileptics who were under a new drug(trt $=1$) and a placebo(trt $=0$) as an adjuvant chemotherapy. Baseline data included the number of epileptic seizures recorded in the preceding 8-week period. The ages of subjects are also considered. A longitudinal study was performed with 2-weeks intervals, and we reanalyze this data in the light of the power-of-the-variance modeling with correlated responses. The general model considered here is

$$\log \mu_{ijk} = \beta_0 + \beta_1(\mathrm{Trt})_i + \beta_2(\mathrm{Base})_{ij} + \beta_3(\mathrm{Age})_{ij}, \tag{4.1}$$

where y_{ijk} represents the epileptic seizure counts for k^{th} visit of i subject in j^{th} group. And the relevant variance function is

$$V(y_{ijk}) = a(\exp(\mu_{ijk}))^b. \tag{4.2}$$

Therefore, Table 4.1 presents the result obtained with GVF with different working correlation matrices(AR(1), CS, TOEP(2)).

5. Conclusion

We compared GVF estimators with QVF estimators for POM variance function with correlated or independent count data. Whereas GVF is more effective in estimating mean parameter, β, QVF is superior to GVF even for the correlated case in estimating variance parameters. For the robustness of GEE1 estimators combined with pseudolikelihood estimators not method-of-moment estimators in POM variance structure, working correlation does not influence the mean parameter, β for low or moderate correlated responses, but for highly correlated case, GVF-CS/AR(1)/TOEPF(2) outperformed GVF-IND as well as QVF. Finally, for variance parameter, a and b, QVF estimator is no less effective than GVF.

When the different a, b set-ups such as (a=1.0, 1.5, 1.8, b=1.0, 1.5, 1.8) were applied, the similar results were obtained but we eliminated the tables for avoiding redundancy.

In conclusion, we recommended to use GVF estimation with any non-diagonal working correlations for β estimation and highly correlated data whereas GVF-IND for low or moderate correlation. Concerning variance parameter, a and b ($v_{ij} = a(\mu_{ij})^b$), QVF developed for independent outcomes may be used even for highly correlated data.

References

Breslow, N. E. and Clayton, D. G. (1993). Approximate inference in generalized linear mixed models, *Journal of the American Statistical Association,* **88**, 9–25.

Carroll, R. J. and Ruppert, D. (1988). *Transformation and Weighting in Regression,* Chapman & Hall/CRC, London, p.270.

Carroll, R. J., Ruppert, D. and Stefanski, L. A. (1995). *Measurement Error in Nonlinear Models,* Chapman & Hall/CRC, London.

Crowder, M. (1985). Gaussian estimation for correlated binomial data, *Journal of the Royal Statistical Society, Series A,* **47**, 229–237.

Crowder, M. (1995). On the use of a working correlation matrix in using generalised linear models for repeated measures, *Biometrika,* **82**, 407–410.

Davidian, M. and Giltinan, D. M. (1995). *Nonlinear Models for Repeated Measurement Data,* Chapman & Hall/CRC, London, p.164.

Davis, P. M. and Pedigo, L. P. (1989). Analysis of spatial patterns and sequential count plans for stalk, *Environmental Entomology,* **18**, 504–509.

Fitzmaurice, G. M. (1995). A caveat concerning independence estimating equations with multivariate binary data, *Biometrics,* **51**, 309–317.

Liang, K. Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models, *Biometrika,* **73**, 13–22.

Liang, K. Y., Zeger, S. L. and Qaqish, B. (1992). Multivariate regression analyses for categorical data, *Journal of the Royal Statistical Society, Series B,* **54**, 3–40.

Littell, R. C., Milliken, G. A., Stroup, W. W. and Wolfinger, R. (1996). *SAS SYSTEM for Mixed Models,* SAS Institute. Inc., Cary, NC, USA.

Madsen, L. and Dalthorp, D. (2007). Generating correlated count data, *Environmental and Ecological Statistics,* **14**, 129–148.

Mancl, L. A. and Leroux, B. G. (1996). Efficiency of regression estimates for clustered data, *Biometrics,* **52**, 500–511.

Park, H. and Cho, K. (2004). Use of covariates in Taylor's power law for sequential sampling in pest management, *Journal of Agricultural, Biological and Environmental Statistics,* **9**, 462–478.

Pepe, M. S. and Anderson, G. L. (1994). A cautionary note on inference for marginal regression models with longitudinal data and general correlated response data, *Communications in Statistics - Simulation and Computation,* **23**, 939–951.

Perry, J. N. (1981). Taylor's power law for dependence of variance on mean in animal population, *Applied Statistics,* **30**, 254–263.

Southwood, T. R. E. (1978). *Ecological Methods (2nd Ed.),* Chapman & Hall/CRC, London, p.391.

Sutradhar, B. C. and Das, K. (1999). On the efficiency of regression estimators in generalised linear models for longitudinal data, *Biometrika,* **86**, 459–465.

Taylor, L. R. (1961). Aggregation, variance and the mean, *Nature,* **189**, 732–735.

Thall, P. F. and Vail, S. C. (1990). Random effects models for serial observations with overdispersion, *Biometrics,* **40**, 961–971.

Wang, Y. G. and Carey, V. (2003). Working correlation structure misspecification, estimation and covariate design: Implications for generalised estimating equations performance, *Biometrika,* **90**, 29–41.

Wang, Y. G. and Carey, V. (2004). Unbiased estimating equations from working correlation models for irregularly timed repeated measures, *Journal of the American Statistical Association,* **99**, 845–853.

Wedderburn, R. W. M. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss—Newton method, *Biometrika,* **61**, 439–447.

Whittle, P. (1961). Gaussian estimation in stationary time series, *Bulletin of the International Statistical Institute,* **39**, 1–26.

[Received August 2008, Accepted October 2008]