A new approach to modeling positive random variables with repeated measures

João Victor B. de Freitas a, Juvêncio S. Nobre b, Marcelo Bourguignon c and Manoel Santos-Neto d

a Departamento de Estatística, Instituto de Matemática, Estatística e Computação Científica, Universidade Estadual de Campinas, Campinas, Brazil; b Departamento de Estatística e Matemática Aplicada, Universidade Federal do Ceará, Fortaleza, Brazil; c Departamento de Estatística, Universidade Federal do Rio Grande do Norte, Natal, Brazil; d Departamento de Estatística, Universidade Federal de Campina Grande, Campina Grande, Brazil

ABSTRACT

In many situations, it is common to have more than one observation per experimental unit, thus generating the experiments with repeated measures. In the modeling of such experiments, it is necessary to consider and model the intra-unit dependency structure. In the literature, there are several proposals to model positive continuous data with repeated measures. In this paper, we propose one more with the generalization of the beta prime regression model. We consider the possibility of dependence between observations of the same unit. Residuals and diagnostic tools also are discussed. To evaluate the finite-sample performance of the estimators, using different correlation matrices and distributions, we conducted a Monte Carlo simulation study. The methodology proposed is illustrated with an analysis of a real data set. Finally, we create an R package for easy access to publicly available the methodology described in this paper.

1. Introduction

Repeated measures study, in special, longitudinal data studies are very common in several areas of knowledge, such as medicine, economy, biology, and so on. In the modeling of such experiments is necessary to consider and model the structure of dependence in experimental units [10]. The first proposal to modeling such experiments was based under the normality assumption [39], as multivariate models, linear mixed models [15,20] and non-linear mixed models [22]. But they do not always present a good alternative. An alternative of flexibilization is to make use of generalized estimating equations (GEE) [21] based on generalized linear models [24]. Outside the exponential dispersion family, there are also some approaches in the literature such as the GEE Simplex model [40,41], GEE Birnbaum-Saunders [44] and the GEE Beta Prime model developed in this paper. A good review of the various approaches can be seen in [39], for example.
In particular, for modeling positive continuous data, few alternatives are available in
the literature, generally restricting to the use of the gamma (GA), inverse gaussian (IG),
log-normal (LN) and Birnbaum–Saunders (BS) distributions (see [11,12,38,51] and [32],
for example). Recently, [2] proposed a new parameterization in terms of mean and precision
parameter of the beta prime (BP) distribution [19]. Under this parameterization, the
probability density function (pdf) is given by
\[
f_Y(y|\mu, \phi) = \frac{y^{\mu(1+\phi)-1}}{B(\mu(1+\phi), \phi + 2)} \frac{(1+y)^{\phi+2}}{(1+y)^{\mu(1+\phi)+\phi+2}} 1_{\mathbb{R}^+}(y),
\]
where $\mu \in \mathbb{R}^+$ is the location parameter (mean), $\phi \in \mathbb{R}^+$ is a precision parameter,
\(B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a + b)\) is the beta function and $\Gamma(a) = \int_0^\infty w^{a-1}e^{-w}dw$ is the
gamma function. It follows from (1) that:
\[
\mathbb{E}(Y) = \mu \quad \text{and} \quad \text{Var}(Y) = \frac{\mu(1+\mu)}{\phi} = \frac{V(\mu)}{\phi},
\]
where $V(\mu)$ acts as a ‘variance function’. This function assumes a quadratic form of the
mean and it is convenient for modeling data with right asymmetry, which can be seen as
an alternative to generalized linear models when data present skewness.

In [2], they emphasize that the BP distribution has common properties with GA and
reparameterized BS [34] such as unimodality in the density function and quadratic vari-
ance function. It follows from (1) that the log-likelihood function based on a single
observation has the form:
\[
\ell(\mu, \phi) = [\mu(1+\phi) - 1] \ln(y) - [\mu(1+\phi) + \phi + 2] \ln(1+y) - \ln[\Gamma(\mu(1+\phi))] - \ln[\Gamma(\phi + 2)] + \ln[\Gamma(\mu(1+\phi) + \phi + 2)].
\]

In the literature, several proposals are for modeling positive longitudinal data, such as
generalized linear mixed models [3], LN model [38] and [4] GEE’s for GLM’s and more
recently GEE using BS [44]. However, the BP distribution has some advantages over the
popular GA and IG models. For example, the BP hazard rate function can be upside-
down bathtub or increasing depending on the parameters values. While most classical
two-parameter distributions such as Gamma distribution have monotone hazard rate func-
tions. Besides that, the skewness and kurtosis of the BP distribution can be much larger
than those of the GA and IG distributions, which maybe more appropriate in certain prac-
tical situations (see Table 1). The BS distribution has been widely used today to model
positive continuous data, mainly in reliability applications to model failure times. (This
distribution was developed initially to model the fatigue life of a material, subject to cyclic
stress). Also, in the context of GEEs, which are approximations, this distribution was pro-
posed for median and shape modeling, which is a different approach to BP that models
mean and precision. Therefore, the main objective of this paper is to extend the proposal
of [2] for modeling longitudinal data, i.e. propose a regression model for positive continu-
ous data with repeated measurements, allowing skewness, based on the BP distribution
and GEE approach.

This paper is organized as follows: in Section 2, we introduce a new regression model
for positive continuous data with repeated measurements based on the BP distribution.
In Section 3, we discuss Wald and score testing of coefficients and information criteria. Diagnostic measures are discussed in Section 4. In Section 5, some numerical results of the estimators are presented. In Section 6, we provide application to real data set. Concluding remarks are given in the final section.

2. GEE for BP regression models with repeated measures

Consider that the longitudinal data consists of $n$ subjects, such that for subject $i (i=1, \ldots, n)$, there are $t$ observations where $y_{ij}$ indicates the $j$th response $(j=1, \ldots, t)$ of subject $i$. Let $y_i = (y_{i1}, \ldots, y_{it}^*)^\top$ be the individual profile of the $i$th subject with the mean vector $\mu_i = (\mu_{i1}, \ldots, \mu_{it})^\top$ where $\mu_{ij}$ is the corresponding $j$th mean. Note that we are assuming, the balanced case (for unbalanced data the algebra is essentially the same). Assume that the marginal distribution of $y_{ij}$ follows a BP distribution, i.e. $y_{ij} \sim \text{BP}(\mu_{ij}, \phi)$, where $\phi$ is equal for the $N$ observations, where $N = nt$. Assume a model for the mean $\mu_{ij}$ in the form:

$$g(\mu_{ij}) = \eta_{ij} = x_{ij}^\top \beta,$$

where $\beta = (\beta_1, \ldots, \beta_p)^\top$ is a vector of parameters $(p < n)$ and $g(\cdot)$ is called the link function which is assumed to be monotone and at least doubly differentiable, $x_{ij} = (x_{ij1}, \ldots, x_{ijp})^\top$ is a fixed and known covariate vector on the $j$th observation of the $i$th subject. For the construction of GEE, consider the vectors $u_i = u_i(y_i, \beta) = (u_{i1}, \ldots, u_{it})^\top$, $i = 1, \ldots, n$, based on the score function for $\mu_{ij}$ (see supplemental material of this paper), with zero mean, mutually independent and satisfying the properties of regular estimating functions [7]. Notice that we are working with the case of constant precision, but we can extend this model by placing a regression structure on $\phi$.

In the context of BP regression models with repeated measures, we use $u_i = y_i^* - \mu_i^*$, where $y_i^* = (y_{i1}^*, \ldots, y_{it}^*)^\top$, $\mu_i^* = (\mu_{i1}^*, \ldots, \mu_{it}^*)^\top$,

$$y_{ij}^* = \ln \left( \frac{y_{ij}}{1 + y_{ij}} \right), \quad \mu_{ij}^* = \psi[\mu_{ij}(1 + \phi)] - \psi[\mu_{ij}(1 + \phi) + \phi + 2]$$

and

$$\psi(\lambda) = \frac{\partial \ln [\Gamma(\lambda)]}{\partial \lambda},$$

where $\psi(\cdot)$ is denoted as digamma function, with $i = 1, \ldots, n$ and $j = 1, \ldots, t$.

For the subject $i$, we have $X_i = (x_{i1}, \ldots, x_{it})^\top$, $A_i = -\mathbb{E}\left( \frac{\partial u_i}{\partial \mu_i} \right)^\top = \mathbb{E}\left( \frac{\partial \mu_i}{\partial \mu_i} \right)^\top = (1 + \phi)G_iA_i$, and $\Sigma_i = \text{Cov}(u_i) = A_i^{1/2} R(\alpha) A_i^{1/2}$, where $\eta_i = (\eta_{i1}, \ldots, \eta_{it})^\top$, $A_i = \text{diag}(\nabla(u_{ij})) = \text{diag}(\nabla(a_{ij}))$, $a_{ij} = \psi^{(1)}[\mu_{ij}(1 + \phi)] - \psi^{(1)}[\mu_{ij}(1 + \phi) + \phi + 2]$, $G_i = \text{diag}$.
reweighed least squares by regressing the modified response vector $H$ of interest can be expressed as follows:

$$
\text{imation:}
$$

matrices, and

$$
\alpha
$$

with block-diagonal weight matrix

$$
\text{has asymptotic distribution } \chi^2,
$$

... 

$$
\text{mth iteration.} [16] \text{ showed that under certain regularity conditions (see } [36] \text{ for example):}
$$

$$
\sqrt{n}(\hat{\beta} - \beta) \overset{D}{\rightarrow} \mathcal{N}_p (0, J^{-1}),
$$

where $J^{-1} = \lim_{n \to \infty} nS^{-1}(\hat{\beta})X^\top \hat{\Sigma}^{-1} \text{Cov}(u) \hat{\Sigma}^{-1} \hat{X}S^{-1}(\hat{\beta})$. The sandwich variance estimator $\hat{J}^{-1}$ of $\hat{\beta}$ (this is referred to as robust estimator in the literature) can be consistently estimated [21] by

$$
\hat{J}^{-1} = S^{-1}(\hat{\beta})X^\top \hat{\Sigma}^{-1} \hat{u} \hat{u}^\top \hat{\Sigma}^{-1} \hat{X}S^{-1}(\hat{\beta}),
$$

where

$$
S(\beta) = \mathbb{E} \left[ \frac{\partial \Psi_1(\beta)}{\partial \beta^\top} \right] = -X^\top W X,
$$

which is known as the sensitivity matrix.

3. **Hypothesis test and model selection**

Let $v$ be a $(q \times 1)$ vector of the $q$ components of interest of $\beta$. Then the hypothesis tests of interest can be expressed as follows: $H_0 : v = v_0$ versus $H_a : v \neq v_0$. An adaptation of the Wald statistic [48] to GEE (see [13] for example) is given by

$$
Q_w = (\hat{v} - v_0)^\top \hat{J}_u^{-1}(\hat{v} - v_0),
$$

where $\hat{J}_u^{-1}$ is the $(q \times q)$ submatrix of the robust estimator $\hat{J}^{-1}$ given in (6). Considering the convergence given in (5), when $n \to \infty$, we have, under the null hypothesis, that (7) has asymptotic distribution $\chi^2_q$. [13] also present Rao score statistics [31] for GEE models.

Several proposals appear in the literature as measures to select the best model in the context of GEE’s, among them we mention the classic ones Quasi-likelihood Information
Criterion (QIC) proposed by [26] based on a modification in the Akaike Information Criterion (AIC) [1], the Rotnitzky–Jewell criterion (RJ) proposed by [33] and the generalized error sum of squares (ESS) [37]. [49] presents the measures cited and some extensions based on them, in addition, it presents simulations to check its performance and notes that the RJ criterion does not perform well for the scenarios with either continuous or binary responses. The criterion based on the ESS does not perform well for longitudinal data with a normal response. However, the performance improvement with binary response. The QIC presents problems when is used a correlation structure independent, but presents advantageous performance when the correlation structure is exchangeable or AR-1. Having said that, we will now talk about the construction of the QIC and two measures based on it.

The QIC can be used both for the selection of covariates and for the choice of the working correlation matrix (see [8] and [9] for example). Another attractive recently proposed model selection measure is the Extended Quasi-likelihood Information Criterion (EQIC) proposed by [50]. These authors using the extended quasi-likelihood [23], based on the deviance function under the independence. EQIC can be used for both covariate and variance function selection [49, 50], so it is a good propose for comparing models of different distributions. The structure (model) that obtains the smaller QIC (EQIC) is better according to this criterion.

4. Diagnostic analysis

Through diagnostic analysis we can verify possible deviations from the initially assumed assumptions for a model, as well as find possible extreme observations that can influence the results of the model fit, such as hypothesis tests and estimates for regression parameters.

In the context of GEEs, the misspecification of the work correlation matrix still gives us consistent estimates for \( \beta \), but can impact the estimation efficiency. That said, the diagnostic analysis together with the measures of goodness of fit provide us with tools to select the best working correlation matrix improving the efficiency of the estimators for \( \beta \).

For GEE, [46], based on the GLM’s measures, extend the concepts of leverage, case deletion and residual analysis based on the proposals of [29] and [42]. [47] presents an extension of the idea of local influence [6] using the generalized local influence [4]. Among other studies cited in the literature dealing with diagnostic techniques for GEE we have [5, 18, 26] and [25], for example.

4.1. Leverage analysis

The hat matrix, based on propose of [27], is given by

\[
H_i = \hat{W}_i^{1/2}X_i \left( X_i^\top \hat{W}_i X_i \right)^{-1} X_i^\top \hat{W}_i^{1/2},
\]

where \( H = \text{diag}(H_1, \ldots, H_n) \), i.e. \( H_i \) has the role of orthogonal projection matrix of local vectors in \( \mathbb{R}^n \) in the subspace generated by the columns of matrix \( \hat{W}_i^{1/2}X_i \). A point is said to be leverage if it is a remote point with respect to the values of \( W^{1/2}X \). We can use the values of \( h_{ij} \), the diagonal values of \( H \), to evaluate whether the observations are leverage points, or \( \sum_{j=1}^{\tau} h_{ij}/t \) to evaluate whether the \( i \)th subject is leverage. A simple visual inspection of leverage plots reveals which points are highlighted relative to others.
4.2. Residual analysis

The vector of ordinary residuals is given by [46]

\[ r_i^* = \hat{W}_i^{1/2} (z_i - \hat{\eta}_i) = \hat{W}_i^{1/2} \Lambda_i^{-1} \hat{u}_i. \]

Venezuela et al. [46] showed that \( \text{Cov}(r_i^*) \approx (I_t - H_i) \), where \( I_t \) is an \((t \times t)\) identity matrix. Since the elements of \( r_i^* \) can have different variances, we can use the standard residuals which are given by

\[ r_{ij} = \frac{r_{ij}^*}{\sqrt{1 - h_{ij}}}, \]

where \( r_{ij}^* \) is the ordinary residual based on the \( j \)th response of the \( i \)th subject. We can use \( r_{ij} \) to highlight possible influence observations or evaluate the linearity of effects.

4.3. Case-deletion analysis

The assessment of the influence of a cluster of observations on the overall model fit may be carried out with a measure of the standardized influence of the subset \( G \) of observations on the linear predictor [28,29]. Based on the idea of [27], of 1-step approximation, one can approximate Cook’s Distance as follows:

\[ DC_{ij} = \frac{1}{p} \left( \hat{\beta}_G - \hat{\beta}_{G(ij)} \right) ^\top X^\top \hat{W} X \left( \hat{\beta}_G - \hat{\beta}_{G(ij)} \right) \approx r_{ij}^2 \frac{h_{ij}}{p(1 - h_{ij})}, \]

which has similar expression for linear models. We can use the normalized Cook’s distance: \( DC_{ij} / \sum_{i,j} DC_{ij} \), in order to obtain values in scale \((0, 1)\) to improve the comparison between points, specially in difference class of models. The diagnostic measures presented above are based on an approximation to \( \hat{\beta} \) and the quality of these measurements depends on the accuracy of the proposed estimator.

4.4. Local influence

The local influence method proposed by [6] consists in evaluate, through an appropriate measure of influence, the robustness of estimates provided by the model thought the effect of minor perturbations in the model itself or in the data. The initial propose is to work with likelihood displacement, but the models based on GEE are not likelihood models. [4] generalized the likelihood displacement to any fit function, which is given by

\[ FD(\omega) = 2 \left\{ F(\hat{\beta}) - F(\hat{\beta}_\omega) \right\}, \]

where \( F \) is a fit function, assumed doubly differentiable and whose estimator is \( \beta \), denoted by \( \hat{\beta} \), the solution of

\[ \frac{\partial F(\beta)}{\partial \beta} = 0. \]

Here, \( \hat{\beta} \) and \( \hat{\beta}_\omega \) are the estimated value for the postulated original and perturbed model, respectively, with \( FD(\omega) \geq 0 \) and \( \omega = (\omega_1, \ldots, \omega_m)^\top \) is a perturbation vector, where the
size $m$ depends on the proposed perturbation scheme and $\omega \in \Omega \subset \mathbb{R}^M$. Based on [6], the idea is to study the local behavior of $\text{FD}(\omega)$ for any value of $\omega$ in a neighborhood of $\omega_0$, which represents the null perturbation vector, such that $\mathcal{F}(\hat{\beta}, \omega) = \mathcal{F}(\hat{\beta}) \Rightarrow \text{FD}(\omega_0) = 0$.

Venezuela et al. [47] extend the proposal of [4] to GEE, replacing likelihood equations with estimation functions, in such a way that given a perturbed estimating equation $\Psi(\hat{\beta} | \omega) = 0$, there is a null perturbation vector so that $\Psi(\beta | \omega_0) = \Psi(\beta)$. Venezuela et al. [47] propose a local influence measure for GEE given by the eigenvector corresponding to the largest eigenvalue of the matrix:

$$\hat{B}_G = -\Delta S^{-1} \Delta,$$

where

$$\Delta = \frac{\partial \Psi_1(\beta | \omega)}{\partial \omega} \quad \text{and} \quad S = \mathbb{E}(\hat{F}) = \frac{\partial \Psi_1(\beta)}{\partial \beta},$$

which are evaluated at $\beta = \hat{\beta}$ and $\omega = \omega_0$. In this paper, we calculate the matrix $\Delta$ for six different perturbation schemes based on the proposals of [47] namely: case-weight perturbation, response perturbation, single-covariate perturbation, precision parameter perturbation and working correlation matrix perturbation. These results are presented with details in supplemental material.

In general, highlighted observations in case-weight can be interpreted as a perturbation in variance of each experimental unit, in particular for Normal linear models [43], as cited by [47]. Perturbation in the response variable can be seen as an alternative way of identifying outliers [35]. The single-covariate perturbation scheme helps to evaluate the influence of each continuous covariate in the estimating process. Perturbation in the precision parameter indicates how sensitive the model is in relation to the homoscedasticity assumption. Working correlation matrix perturbation can indicate, for example, whether we need to use another correlation structure.

### 4.4.1. Case-weight perturbation

First, we consider an arbitrary attribution of weights for the estimating equation. For this perturbation scheme, we find:

$$\Psi_1(\beta | \omega) = X^T W \Lambda^{-1} \text{diag}(\omega) u,$$

where $\omega = (\omega_1^T, \ldots, \omega_n^T)^T$, with $\omega_i = (\omega_{i1}, \ldots, \omega_{it})^T, i = 1, \ldots, n$. Here, the nonperturbation vector, $\omega_0$, assumes $\omega_{ij} = 1$, with $i = 1, \ldots, n$ and $j = 1, \ldots, t$. For this scheme, the matrix $\hat{B}_G$ defined in (8) is given by $\text{diag}(u) \Lambda^{-1} W^T X (X^T W X)^{-1} X^T W \Lambda^{-1} \text{diag}(u)$, which is evaluated at $\hat{\beta}$.

### 4.4.2. Response perturbation

Consider the additive perturbation scheme in $y_{ij}$ given by

$$y_{oij} = y_{ij} + \omega_{ij} \sqrt{\text{Var}(y_{ij})},$$

where the nonperturbation vector assumes $\omega_{ij} = 0$, i.e. $\omega_0 = 0$. Analyzing the GEE defined in (4), the only part that depends on $y_{ij}$ is the vector $u$. Therefore, considering $u_\omega$, the vector...
$u$ perturbed in response variable. The perturbed estimating equation is given by

$$\Psi_1(\beta | \omega) = X^T W \Lambda^{-1} u_0,$$

where $u_0 = (u_{o1}, \ldots, u_{on})^T$, with $u_{oi} = (u_{oi1}, \ldots, u_{oit})^T, i = 1, \ldots, n$. In this case, the matrix $\Delta$ is given by $X^T W \Lambda^{-1} B$, where $B = \partial u_0 / \partial \omega^T = \text{diag}\{\partial u_{o11} / \partial \omega_{11}, \ldots, \partial u_{on} / \partial \omega_{nt}\}$. The matrix $\hat{B}_G$ under response perturbation is given by

$$B \Lambda^{-1} W^T X (X^T WX)^{-1} X^T W \Lambda^{-1} B,$$

which is evaluated at $\hat{B}$ and $\omega_0$. For GEE using BP, we have to (see supplemental material):

$$\frac{\partial u_{oij}}{\partial \omega_{ij}} = \frac{s_{ij}}{V(y_{oij})},$$

where $s_{ij} = \sqrt{\nu(\mu_{ij})/\phi}$ and $V(y_{oij}) = y_{oij}(1 + y_{oij})$, thereby we have to

$$B = S \nu^{-1},$$

where $\nu = \text{diag}(\nu_1, \ldots, \nu_n), S = \text{diag}(S_1, \ldots, S_n), V_i = \text{diag}(V(y_{oi1}), \ldots, V(y_{oit}))$ and $S_i = \text{diag}(s_{i1}, \ldots, s_{it}), i = 1, \ldots, n$.

### 4.4.3. Single-covariate perturbation

Thomas and Cook [43] proposed an additive scheme of perturbation in the $k$th column of $X$, let is name this column of $x_k = (x_{1k}, x_{2k}, \ldots, x_{nk})^T$, with $k = 1, \ldots, p$. Therefore the elements of perturbation vector $x_k$ is given by

$$x_{oijk} = x_{ijk} + \omega_{ij} s_{x_k},$$

where $s_{x_k}$ is the standard deviation of $x_k$ and nonperturbation vector $\omega_0$ assumes $\omega_{ij} = 0$, with $i = 1, \ldots, n$ and $j = 1, \ldots, t$. In this perturbation scheme, according to (4), all components are modified when we perturb the $k$th covariate. Therefore, the perturbed estimating equation is given by

$$\Psi_1(\beta | \omega) = X_{o}^T \Lambda_0 \Sigma_{\omega}^{-1} u_0,$$

where the index $\omega$ indicates that the matrix depends on the perturbation scheme defined in (9). According to [14], the derivative of $\Psi_1(\beta | \omega)$ with respect to the $\omega^T$ is given by

$$\Delta = X_{o}^T \Lambda_0 \left[ \Sigma_{\omega}^{-1} \frac{\partial u_0}{\partial \omega^T} + \Sigma_{\omega}^{-1} \text{diag}(u_0) \right] + \left[ X_{o}^T \frac{\partial \Lambda_0}{\partial \omega^T} + \frac{\partial X_{o}^T}{\partial \omega^T} \Lambda_0 \right] \Sigma_{\omega}^{-1} \text{diag}(u_0),$$

where the derivative of $X_{o}^T$ with respect to the $\omega^T$ is a $(p \times N)$ of zeros except the $k$th line that is $s_{x_k}$, and

$$\frac{\partial \Sigma_{\omega}^{-1}}{\partial \omega^T} = -\Sigma_{\omega}^{-1} \Sigma_{\omega}^T, \quad \frac{\partial \Lambda_0}{\partial \omega^T} = \left( \frac{\partial \Lambda_{o1}}{\partial \omega_{1}}, \ldots, \frac{\partial \Lambda_{on}}{\partial \omega_{n}} \right),$$

$$\frac{\partial \Sigma_{\omega}}{\partial \omega^T} = \left( \frac{\partial \Sigma_{o1}}{\partial \omega_{1}}, \ldots, \frac{\partial \Sigma_{on}}{\partial \omega_{n}} \right),$$

$$\frac{\partial u_0}{\partial \omega^T} = \left( \frac{\partial u_{o1}}{\partial \omega_{1}}, \ldots, \frac{\partial u_{on}}{\partial \omega_{n}} \right).$$

(10)
According to estimating equation in (4), we have to
\[ \Lambda_{oij} = (1 + \phi)G_{oij}A_{oij}, \quad \Sigma_{oij} = A_{oij}^{1/2} R(\alpha) A_{oij}^{1/2} \quad \text{and} \quad u_{oij} = y_i^* - \mu_{oij}^*, \]
being \( G_{oij} = \text{diag}(\partial g^{-1}(\eta_{oij})/\partial \eta_{oij1}, \ldots, \partial g^{-1}(\eta_{oij})/\partial \eta_{oijt}) \), \( A_{oij} = \text{diag}(a_{oij1}, \ldots, a_{oijt}) \) and \( \mu_{oij}^* = \text{diag}(\mu_{oij1}^*, \ldots, \mu_{oijt}^*)^T \), where \( a_{oij} = \psi^{(1)}[\mu_{oij}(1 + \phi)] - \psi^{(1)}[\mu_{oij}(1 + \phi) + \phi + 2] \) and \( \mu_{oij}^* = \psi[\mu_{oij}(1 + \phi)] - \psi[\mu_{oij}(1 + \phi) + \phi + 2] \), with \( i = 1, \ldots, n \) and \( j = 1, \ldots, t \). Therefore, we get (see supplemental material):
\[
\frac{\partial \Lambda_{oij}}{\partial \omega_i^T} = (1 + \phi) s_k \beta_k \left[ V_{oij} G_{oij}^2 + G_{oij} A_{oij} \right], \quad \frac{\partial u_{oij}}{\partial \omega_i} = -(1 + \phi) s_k \beta_k G_{oij} A_{oij}, \quad \text{and}
\]
\[
\frac{\partial \Sigma_{oij}}{\partial \omega_i^T} = \frac{1}{2} s_k \beta_k \left[ A_{oij}^{1/2} R(\alpha) A_{oij} + A_{oij} R(\alpha) A_{oij}^{1/2} \right],
\]
where \( V_{oij} = \text{diag}(\partial^2 g^{-1}(\eta_{oij})/\partial \eta_{oij1}^2, \ldots, \partial^2 g^{-1}(\eta_{oij})/\partial \eta_{oijt}^2) \), \( A_{oij} = A_{oij}^{-1/2} G_{oij} V_{oij} \) and \( V_{oij} = \text{diag}(V_{oij1}, \ldots, V_{oijt}) \), where
\[
V_{oij} = (1 + \phi) \left\{ \psi^{(2)}[\mu_{oij}(1 + \phi)] - \psi^{(2)}[\mu_{oij}(1 + \phi) + \phi + 2] \right\},
\]
with the polygamma function \( \psi^{(2)}(\cdot) \) given by
\[
\psi^{(2)}(\lambda) = \frac{\partial^2 \psi^{(1)}(\lambda)}{\partial \lambda^2} = \frac{\partial^3 \ln[G(\lambda)]}{\partial \lambda^3} = \frac{\Gamma^{(3)}(\lambda) \Gamma(\lambda) - \Gamma^{(2)}(\lambda) \Gamma^{(1)}(\lambda)}{\Gamma(\lambda)^2} - \psi(\lambda)^2,
\]
where \( \psi^{(1)}(\cdot) \) is denoted as trigamma function and is given by
\[
\psi^{(1)}(\lambda) = \frac{\partial \psi(\lambda)}{\partial \lambda} = \frac{\partial^2 \ln[G(\lambda)]}{\partial \lambda^2} = \frac{\Gamma^{(2)}(\lambda) \Gamma(\lambda) - \Gamma^{(1)}(\lambda)}{\Gamma(\lambda)^2}.
\]

### 4.4.4. Precision parameter perturbation

Venezuela et al. [47] proposed a perturbation scheme for precision parameter, so that it is not constant throughout the observations and nor between individuals, i.e. \( \phi_{oij} = \phi_{-oij} \), in this scheme, the nonperturbation vector \( \omega_0 = I_N \) and perturbed estimating equation is given by
\[
\Psi_1(\beta|\omega) = \Psi_1(\beta|\omega) = X^T \Lambda_\omega \Sigma_\omega^{-1} u_\omega.
\]

Thus, according to [14], the derivative of \( \Psi_1(\beta|\omega) \) with respect to the \( \omega^T \) can be expressed by
\[
\Delta = X^T \Lambda_\omega^T \left[ \Sigma_\omega^{-1} \frac{\partial u_\omega}{\partial \omega_i} + \frac{\partial \Sigma_\omega^{-1}}{\partial \omega_i^T} \text{diag}(u_\omega) \right] + X^T \Lambda_\omega \Sigma_\omega^{-1} \text{diag}(u_\omega),
\]
where the derivatives of \( \Sigma_\omega^{-1}, \Lambda_\omega \) and \( u_\omega \) all with respect to \( \omega^T \) are given in (11), respectively. According to estimating equation defined in (4), we have to
\[
\Lambda_{oij} = G_{oij}^* A_{oij}, \quad \Sigma_{oij} = A_{oij}^{1/2} R(\alpha) A_{oij}^{1/2} \quad \text{and} \quad u_{oij} = y_i^* - \mu_{oij}^*,
\]
where \( G_{oij} = \text{diag}(1 + \phi_{oij1}, \ldots, 1 + \phi_{oijt}), \ A_{oij} = \text{diag}(a_{oij1}, \ldots, a_{oijt}) \) and \( \mu_{oij}^* = \text{diag}(\mu_{oij1}^*, \ldots, \mu_{oijt}^*)^T \), where \( a_{oij} = \psi^{(1)}[\mu_{ij}(1 + \phi_{oij})] - \psi^{(1)}[\mu_{ij}(1 + \phi_{oij} + \phi + 2)] \) and
\[ \mu_{oij} = \psi[\mu_{ij}(1 + \phi_{oij})] - \psi[\mu_{ij}(1 + \phi_{oij}) + \phi_{oij} + 2], \] 
with \( i = 1, \ldots, n \) and \( j = 1, \ldots, t \). 
Therefore, we get (see supplemental material):

\[ \frac{\partial A_{oi}}{\partial \omega_i} = -\phi^{-1}G_i(\Phi_{oi} - 1)^2 [\Phi_{oi}W_{oi}^* + A_{oi}], \quad \frac{\partial u_{oi}}{\partial \omega_i} = \phi^{-1}(\Phi_{oi} - 1)^2 W_{oi}^{**}, \] 
and

\[ \frac{\partial \Sigma_{oi}}{\partial \omega_i} = -\frac{1}{2} \phi^{-1} \left[ A_{oi}^{1/2} R(\alpha) A_{oi} + A_{oi} R(\alpha) A_{oi}^{1/2} \right], \]

where \( A_{oi} = A_{oi}^{-1/2}(\Phi_{oi} - 1)^2 W_{oi}^* \), \( W_{oi}^* = \text{diag}(W_{o11}^*, \ldots, W_{omi}^*) \) and \( W_{oi}^{**} = \text{diag}(W_{o11}^{**}, \ldots, W_{omi}^{**}) \), with

\[ W_{oi}^* = \left\{ \mu_{ij}\psi^{(2)}[\mu_{ij}(1 + \phi_{oi})] - (1 + \mu_{ij})\psi^{(2)}[\mu_{ij}(1 + \phi_{oi}) + \phi_{oi} + 2] \right\}, \]

and

\[ W_{oi}^{**} = \left\{ \mu_{ij}\psi^{(1)}[\mu_{ij}(1 + \phi_{oi})] - (1 + \mu_{ij})\psi^{(1)}[\mu_{ij}(1 + \phi_{oi}) + \phi_{oi} + 2] \right\}, \]

for \( i = 1, \ldots, n \) and \( j = 1, \ldots, t \).

### 4.4.5. Working correlation matrix perturbation

Let \( R(\alpha) \) be a working correlation matrix characterized by a \( \binom{n}{2} \)-vector \( \alpha = (\alpha_1, \ldots, \alpha_{(n-1)})^\top \). For example, if the structure of correlation matrix is exchangeable, then \( \alpha_{jl} = \alpha \), if it is first-order autoregressive (AR(1)), then \( \alpha_{jl} = \alpha_{|j-l|} \), with \( j, l = 1, \ldots, t \) and \( j < l \).

As each unit has its own working correlation matrix, a perturbation scheme in the correlation vector \( \alpha \) can be given by [47]

\[ \alpha_{oij(lj)} = \frac{\alpha_{jl}}{\omega_{i(lj)}}, \]

with \( i = 1, \ldots, n \), \( j < l \) and \( j, l = 1, \ldots, t \). For this perturbation scheme \( \omega = (\omega_{11}, \ldots, \omega_{n1}, \omega_{1n}, \ldots, \omega_{n(n-1)})^\top \) is a perturbation vector and \( \omega_0 \) is a vector with ones. The perturbation estimating equation is given by

\[ \Psi_1(\beta, \omega) = X^\top \Delta \Sigma_\omega^{-1} u. \]

The each column of matrix \( \Delta \) can be expressed by

\[ \frac{\partial \Psi_1(\beta, \omega)}{\partial \omega_{(jl)}} = X^\top \Delta \frac{\partial \Sigma_\omega^{-1}}{\partial \omega_{(jl)}} u, \]

where the derivative of \( \Sigma_\omega^{-1} \) with respect to \( \omega_{(jl)} \) is similar to Equation (13). The \( i \)th diagonal block of \( \Sigma_\omega \) is \( \Sigma_{oi} = \sqrt{V(u_i)} R(\alpha_{oi}) \sqrt{V(u_i)} \), with \( \alpha_{oi} = (\alpha_{oi(12)}, \ldots, \alpha_{oi((t-1)t)}) \) and \( i = 1, \ldots, n \). We also have to:

\[ \frac{\partial \Sigma_{oi}}{\partial \omega_{(il)}} = \sqrt{V(u_i)} \frac{\partial R(\alpha_{oi})}{\partial \omega_{(il)}} \sqrt{V(u_i)}, \]

where \( \partial R(\alpha_{oi})/\partial \omega_{(il)} \) is a \( (t \times t) \) symmetric matrix with null diagonal and \( il \) and \( lj \) elements equal to \(-\alpha_{jl}\), with \( i = 1, \ldots, n \) and \( j, l = 1, \ldots, t \).
Theoretical results of this paper have been implemented in the R software [30] through the creation of a package called geeBP. The geeBP package provides functions for fitting and analyzes GEE with Beta prime models. The current version can be downloaded from GitHub using the command in R:

```
devtools::install_github("jvbfreitas/geeBP")
```

The function `geeBP()` fits a model and creates an object of class `geeBP` which can be applied in the functions `diag.geeBP()` (Cook’s distance and $H$ matrix), `LI.BP()` (local influence), `summary()` and `residuals()` (Standard and Pearson residuals).

5. Monte Carlo simulation studies

5.1. Parameter recovery

In this section, we present an simulation study to analyze the consistency of the estimates using: Relative bias (RB), standard deviation (SD) and root-mean-square error (RMSE). The Monte Carlo experiments were carried out using

$$y_{ij} \sim D(\mu_{ij}, \phi) \quad \text{and} \quad \ln(\mu_{ij}) = 3.21 + 1.74x_{ij} - 0.78w_{ij},$$

where $D$ is the marginal distribution considered (BP), the covariate $x_{ij}$ was generated from a standard uniform and $w_{ij}$ from a Gamma(1,2), with $i = 1, \ldots, n$, $j = 1, \ldots, t$ and the working correlation matrix is AR-1. The simulated values were obtained via Gaussian copulas, due to the dependency between the observations of the same experimental units. The simulated values for each explanatory variable were kept constant during the simulations. The number of Monte Carlo iterations per scenario (or case) was set equal to 2000. This study aimed to evaluate the impact of the sample size, precision, and correlation in the parameters estimates. All simulations were performed using the R software [30] and regression coefficients were estimated for each sample using the function `geeBP()` . The results of the simulation can easily be obtained from https://santosneto.shinyapps.io/GEEBP/.

5.1.1. Sample size effect

In this scenario, we consider $\phi = 10$, $\alpha = 0.5$, $n = 30, 50, 100$ and $t = 4, 6, 10$. Table 2 allows you to see what happens to RB, SD and RMSE when $n$ and $t$ grow. Note that, in general, the RB decreases with the growth of $n$ and $t$ (mainly for $n$). Now for SD and RMSE we note that these quantities decreasing with the increase of $n$ and $t$. In this way we show that the estimators are consistent.

5.1.2. Precision effect

What effect does increasing the precision have on estimates? To answer, we consider $n = 50$, $t = 4$ and $\alpha = 0.5$ fixed and $\phi = 2, 10$ and 100. In Table 3, note that the performance of the estimators improves with increasing precision. For example, for $\hat{\beta}_0$, the RB suffers a 98% cut in its value.

5.1.3. Correlation effect

The correlation effect on estimates of parameters was measured considering a low ($\alpha = 0.2$), moderate ($\alpha = 0.5$) and high ($\alpha = 0.5$) correlation. The values of $n$, $t$ and $\phi$...
considered were 50, 10 and 10, respectively. We conclude that the correlation effect is stronger on estimates of SD and RMSE. For these quantities, you can see a reduction in its estimates with the growth of the degree of correlation. For example, for $\hat{\beta}_2$, there was a 33.5% reduction in the value of the RMSE (see Table 4).

### 5.2. Performance with beta prime data

In order to verify the performance of the GEE BP model and show the poor performance of GEE GA and IG models when fitting data generated from a BP we generate 2000 replicas of the model: $\ln(\mu_{ij}) = 4 - 0.6x_{ij}$, where $x_{ij} = j$, with $n = 40$, $j = 1, \ldots, 8$, $\alpha = 0.5$, AR-1 working correlation matrix and $\phi = 1$.

We note from Table 5 that both GA and BP estimated $\beta_0$ with low bias ($< 0.04$) and $\beta_1$ with low bias as well ($< 0.01$), although we note that in particular the BP model has the lowest standard deviation indicating that this model estimates more precisely, which is evidenced through the RMSE.
Table 5. Estimates of the RB, SD and RMSE under GEE BP, GA and IG models.

| Model | RB($\beta_0$) | RB($\beta_1$) | SD($\beta_0$) | SD($\beta_1$) | RMSE($\beta_0$) | RMSE($\beta_1$) |
|-------|----------------|----------------|---------------|---------------|----------------|----------------|
| BP    | -0.0077        | -0.0028        | 0.1145        | 0.0202        | 0.1186         | 0.1006         |
| GA    | -0.0033        | 0.0014         | 0.1509        | 0.0322        | 0.1514         | 0.1057         |
| IG    | 0.0473         | -0.0518        | 0.7666        | 0.1071        | 0.7895         | 0.1302         |

6. Real-world data analysis

In this section, we apply the methodology considered in previous sections to one real dataset. The required numerical evaluations for the data analysis were carried out by using the R software. Consider an experiment, analyzed by [46], carried out in School of Physical Education and Sport of the University of São Paulo (USP), whose data were obtained from Centre for Applied Statistics, part of the Institute of Mathematics and Statistics, at the USP. The purpose of the study was to evaluate the process of learning synchronizing tasks developed by volunteers, analyzing the variable absolute error. For this, each one of the 40 volunteers performed the same task 80 times, divided in 8 blocks of 10 attempts. In each attempt, was observed the time difference between the moment the individual received the stimulus and the moment of their motor reaction (in ms). The response variable was defined, for each block, as the mean of the absolute values of observed time differences over the 10 attempts. This database was previously analyzed by [46] using a Normal GEE model, but this distribution is initially proposed to model real data and a distribution with positive support is more appropriate to model such data. Also, we noticed through their analysis of Cook’s distance that there are points which influence the modeling and that indicate the need for a more robust model. The data can be found in Appendix A of [45].

It should be noticed that the previous analyses made for this experiment were based on normal models. The purpose of this data analysis is to remodel this data set with more adequate distributions (continuous positive), among them BP, and show that it presents more reliable results according to the diagnostic analysis compared to the classic Gamma and IG, for example.

Figure 1. Individual profiles with average profile for absolute time difference (a) and sample variogram for absolute time difference (b).
Figure 2. Half-normal plots for adjusted models.

Table 6. QIC and EQIC for the BP, GA and IG models under the unstructured and AR-1 structure.

| Measure | BP       | GA       | IG       |
|---------|----------|----------|----------|
|         | AR-1     | Unstructured | AR-1     | Unstructured | AR-1     | Unstructured |
| EQIC    | -222.62  | -223.62  | 2738.12  | 2735.33  | 2749.64 | 2721.47    |
| QIC     | 259.64   | 269.21   | 294.16   | 289.43   | 300.34 | 301.06     |
Through the profile plot (Figure 1), we observed a possible linear relation between the logarithm of absolute difference and the blocks. In the sample variogram plot (Figure 1), we have indications that the correlation between blocks decreases with respect to lag of the time, because the spline approaches the dotted line referring to the sample variance $\sigma^2_s$ related to the standardized response variable. The analysis of the sample variogram consists of verifying whether its values in each lag are close to $\sigma^2_s$ or not, the closer they are, the lower the correlation (due to the definition of the sample variogram), in this case the use of the spline serves as tool to study the behavior of the sample variogram around this variance, for more details and explanation, see [10] for example. Therefore, there is evidence that a model with AR-1 structure for intra-class dependence is reasonable, but unstructured structure will be considered as well.

We then initially fit six models to this data. The GEE models with BP, GA and IG distributions with AR-1 and unstructured correlation matrix and the following structure:

$$\ln(\mu_{ij}) = \beta_0 + \beta_1 \text{block}_j,$$

where $\text{block}_j$ is the $j$th block, with $i = 1, \ldots, 40$ and $j = 1, \ldots, 8$. Note that only in order to have another candidate model, we also considered the model based on the IG distribution in this application, although it was not considered in the simulation.

![Normalized Cook's distance for adjusted models.](image)

**Figure 3.** Normalized Cook's distance for adjusted models.
The results of half-normal plots with simulated envelope of 95% confidence bands are shown in Figure 2. We still have that the GEE BP model presented the best fit analyzing the half-normal plots in comparison to GA and IG models.

In Table 6, we observe that through the EQIC we have an indication that the BP model presents a better fit, because considering either of the two working correlation matrices it has the lowest EQIC in all cases, based on this also, for the AR-1 type correlation the Gamma model is better than the IG and the inverse is observed for the unstructured correlation matrix. Analyzing the QIC, across the working correlation matrices for each model, we have an indication that the AR-1 structured is the most appropriate for BP and Inverse Gaussian models, respectively, already for GA the best structure is the unstructured.

In Figure 3, we plot the normalized Cook’s distance versus Subjects observations and the highlighted points were named by (Subject, Block). Note that the observations highlighted in the GA and IG models have less influence on the BP model. In Figure 4, we plot the normalized $|d_{\text{max}}|$, i.e. $|d_{\text{max}}|_{ij}/ \sum_{i,j} |d_{\text{max}}|_{ij}$ for case-weight perturbation and we have the same comment made for the normalized Cook’s distance, showing even more the good fit of GEE BP against GEE Gamma and GEE IG models.

The results of the fitted models follow in Table 7. Note that the parameters for the models presented $p$-value, using Wald statistic, less than 0.01, but BP and IG model have much smaller standard errors than GA.

![Normalized $d_{\text{max}}$ versus Index for case-weight perturbation of GEE IG model.](image1)

![Normalized $d_{\text{max}}$ versus Index for case-weight perturbation of GEE Gamma model.](image2)

![Normalized $d_{\text{max}}$ versus Index for case-weight perturbation of GEE BP model.](image3)

**Figure 4.** Normalized $d_{\text{max}}$ versus Index for case-weight of adjusted models.
Table 7. Estimates, standard errors (SE) and significant tests for parameters of GEE BP and GA model.

| Parameter | Beta prime (AR-1) | Gamma (unstructured) | IG (unstructured) |
|-----------|-----------------|---------------------|-----------------|
|           | Estimate   | SE   | p-value (Wald) | Estimate   | SE   | p-value (Wald) | Estimate   | SE   | p-value (Wald) |
| Intercept | 3.91       | < 0.01 | < 0.01         | 3.94       | 0.08  | < 0.01         | 3.89       | < 0.01 | < 0.01         |
| Block     | −0.05      | < 0.01 | < 0.01         | −0.05      | 0.01  | < 0.01         | −0.05      | < 0.01 | < 0.01         |
| \hline
| \( \hat{\phi} \) | 4.73       | ×    | ×              | 4.97       | ×    | ×              | 202.75     | ×    | ×              |
| \( \hat{\alpha} \) | 0.52       | ×    | ×              | ×          | ×    | ×              | ×          | ×    | ×              |

Figure 5. Standard residuals versus Subjects and Local influence for GEE BP model with AR-1 structure.

That said, through the goodness-of-fit measures, the Cook’s distance and case-weight perturbation, we selected the GEE BP model as the best. In Figure 5, we plot standard residuals and the normalized \(|d_{\text{max}}|\) for the response, precision parameter and working correlation matrix perturbations, and highlighted observations associated to subjects 5 and 17. We decide to make a confirmatory analysis by withdraw subjects 5 and 17, and we observed that the inferential results did not undergo major changes, but in the working correlation matrix perturbation new observations are highlighted. Analyzing the dataset, in the second block the experimental units 17 and 5 were the second and fourth lowest values, respectively. It is also worth noting that in relation to other local influence schemes, the IG and Gamma models have points considered influential in the response variable perturbation.
7. Discussion

In this paper, we have developed a new regression model used to modeling asymmetric positive continuous data to take account of possible serial dependence between repeated observations. We developed the modeling only mean. The proposed model may serve as a good alternative to the GA and IG regression models for modeling asymmetric positive continuous data with repeated observations. Diagnostic tools have been obtained to detect locally influential data. Simulations show the good performance of this new model.

Through the analysis of synchronizing tasks data, we found that the mean can be a function of blocks and show that the BP models have a good performance in relation to the its competitors: Gamma and IG. The diagnostic analysis developed in this paper shows that the GEE BP model proves to fit well to the data, which is mainly evidenced by the simulated envelopes, also, the confirmatory analysis revealed its stability.

As limitations of the proposed model we mention that it does not consider the heterogeneity of the precision parameter, a GEE BP model for joint modeling of mean and precision is being developed by the authors as an extension to the model proposed in this paper. Another study of interest would be to consider semi-parametric predictors for the mean of the response variable.

Acknowledgments

We are very grateful to editors and reviewers for their careful reading of our manuscript and constructive comments.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Research carried out using the computational resources of the Center for Mathematical Sciences Applied to Industry (CeMEAI) funded by FAPESP (grant 2013/07375-0).

ORCID

João Victor B. de Freitas http://orcid.org/0000-0001-6025-8087
Juvêncio S. Nobre http://orcid.org/0000-0002-7321-3221
Marcelo Bourguignon http://orcid.org/0000-0002-1182-5193
Manoel Santos-Neto http://orcid.org/0000-0002-9007-0680

References

[1] H. Akaike, Information theory and an extension of the maximum likelihood principle, in Selected papers of Hirotugu Akaike, Springer, 1998, pp. 199–213.
[2] M. Bourguignon, M. Santos-Neto and M. de Castro, A new regression model for positive random variables with skewed and long tail, METRON. 79 (2021), 33–55. https://doi.org/10.1007/s40300-021-00203-y,
[3] N.E. Breslow and D.G. Clayton, Approximate inference in generalized linear mixed models, J. Am. Stat. Assoc. 88 (1993), pp. 9–25.
[4] N.G. Cadigan and P.J. Farrell, *Generalized local influence with applications to fish stock cohort analysis*, J. R. Stat. Soc. Ser. C (Appl. Stat.) 51 (2002), pp. 469–483.

[5] Y.C. Chang, *Residuals analysis of the generalized linear models for longitudinal data*, Stat. Med. 19 (2000), pp. 1277–1293.

[6] R. Cook, *Assessment of local influence*, J. R. Stat. Soc. Ser. B. (Methodol.) 48 (1986), pp. 133–169.

[7] M. Crowder, *Assessment of local influence*, J. R. Stat. Soc. Ser. B. (Methodol.) 48 (1986), pp. 591–597.

[8] J. Cui, *Qic program and model selection in gee analyses*, Stata. J. 7 (2007), pp. 209–220.

[9] J. Cui and G. Qian, *Selection of working correlation structure and best model in gee analyses of longitudinal data*, Commun. Stat. Simul. Comput. 36 (2007), pp. 987–996.

[10] P.J. Diggle, P.J. Heagerty, K.Y. Liang and S. Zeger, *Analysis of Longitudinal Data*, Oxford University Press, Oxford, 2002.

[11] D. Firth, *Multiplicative errors: log-normal or gamma?*, J. R. Stat. Soc. Ser. B. (Methodol.) 50 (1988), pp. 266–268.

[12] J.L. Folks and R.S. Chhikara, *The inverse gaussian distribution and its statistical application review*, J. R. Stat. Soc. Ser. B. (Methodol.) 40 (1978), pp. 263–275.

[13] J.W. Hardin and J.M. Hilbe, *Generalized Estimating Equations*, 2nd ed., Chapman and Hall/CRC, 2012.

[14] D.A. Harville, *Matrix Algebra from a Statistician's Perspective*, 1st ed., Springer-Verlag, New York, 1997.

[15] C.R. Henderson, *Estimation of variance and covariance components*, Biometrics. 9 (1953), pp. 226–252.

[16] B. Jørgensen and R. Labouriau, *Exponential families and theoretical inference*, Lecture notes, Department of Statistics, University of British Columbia, 1994.

[17] B. Jørgensen, S. Lundbye-Christensen, P.X.K. Song and L. Sun, *State-space models for multivariate longitudinal data of mixed types*, Can. J. Stat. / La Rev Canad Stat 24 (1996), pp. 385–402.

[18] K.M. Jung, *Local influence in generalized estimating equations*, Scand. J. Stat. 35 (2008), pp. 286–294.

[19] E. Keeping, *Introduction to Statistical Inference*, New York, Van Nostrand, 1962.

[20] N.M. Laird and J.H. Ware, *Random-effects models for longitudinal data*, Biometrics. 38 (1982), pp. 963–974.

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

[22] M.J. Lindstrom and D.M. Bates, *Nonlinear mixed effects models for repeated measures data*, Biometrics 46 (1990), pp. 673–687.

[23] J. Nelder and D. Pregibon, *An extended quasi-likelihood function*, Biometrika. 74 (1987), pp. 221–232.

[24] J. Nelder and R. Wedderburn, *Generalized linear models*, J. R. Stat. Soc. Ser. A (General). 135 (1972), pp. 370–384.

[25] S. Oh, K. Carriere and T. Park, *Model diagnostic plots for repeated measures data using the generalized estimating equations approach*, Comput. Stat. Data Anal. 53 (2008), pp. 222–232.

[26] W. Pan, *Akaike's information criterion in generalized estimating equations*, Biometrics. 57 (2001), pp. 120–125.

[27] D. Pregibon, *Logistic regression diagnostics*, Ann. Statist. 9 (1981), pp. 705–724.

[28] J.S. Preisser, K. By, J. Perin and B.F. Qaqish, *Deletion diagnostics for alternating logistic regressions*, Biom. J. 54 (2012), pp. 701–715. https://onlinelibrary.wiley.com/doi/abs/10.1002/bimj.201200002.

[29] J.S. Preisser and B.F. Qaqish, *Deletion diagnostics for generalised estimating equations*, Biometrika. 83 (1996), pp. 551–562.

[30] R Core Team: *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria (2019). https://www.R-project.org/.

[31] C.R. Rao, *Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation*, in Mathematical Proceedings of the Cambridge Philosophical Society, Vol. 44. Cambridge University Press, 1948, pp. 50–57.
[32] J.R. Rieck and J.R. Nedelman, A log-linear model for the Birnbaum–Saunders distribution, Technometrics 33 (1991), pp. 51–60.

[33] A. Rotnitzky and N.P. Jewell, Hypothesis testing of regression parameters in semiparametric generalized linear models for cluster correlated data, Biometrika. 77 (1990), pp. 485–497.

[34] M. Santos-Neto, F.J. Cysneiros, V. Leiva and M. Barros, Reparameterized Birnbaum–Saunders regression models with varying precision, Electron. J. Stat. 10 (2016), pp. 2825–2855.

[35] B. Schwarzmann, A connection between local-influence analysis and residual diagnostics, Technometrics. 33 (1991), pp. 103–104.

[36] P.K. Sen, J.M. Singer and A.C.P. de Lima, From Finite Sample to Asymptotic Methods in Statistics, 1st ed., Cambridge University Press, Oct. 2009. Cambridge Series in Statistical and Probabilistic Mathematics (Book 29).

[37] J. Shults and N.R. Chaganty, Analysis of serially correlated data using quasi-least squares, Biometrics. 54 (1998), pp. 1622–1630.

[38] J.M. Singer, J.S. Nobre and H.C. Sef, Regression models for pretest/posttest data in blocks, Stat. Modelling. 4 (2004), pp. 324–338.

[39] J.M. Singer, F.M. Rocha and J.S. Nobre, Graphical tools for detecting departures from linear mixed model assumptions and some remedial measures, Int. Stat. R. 85 (2017), pp. 290–324.

[40] P.X.K. Song, Z. Qiu and M. Tan, Modelling heterogeneous dispersion in marginal models for longitudinal proportional data, Biom. J. J. Math. Methods Biosci. 46 (2004), pp. 540–553.

[41] P.X.K. Song and M. Tan, Marginal models for longitudinal continuous proportional data, Biometrics. 56 (2000), pp. 496–502.

[42] M. Tan, Y. Qu and M.H. Kutner, Model diagnostics for marginal regression analysis of correlated binary data, Commun. Stat. – Simul. Comput. 26 (1997), pp. 539–558.

[43] W. Thomas and R.D. Cook, Assessing influence on predictions from generalized linear models, Technometrics. 32 (1990), pp. 59–65.

[44] A.B. Tsuyuguchi, G.A. Paula and M. Barros, Analysis of correlated Birnbaum–Saunders data based on estimating equations, TEST. 29 (2019), pp. 1–21.

[45] M.K. Venezuela, Modelos lineares generalizados para análise de dados com medidas repetidas, Ph.D. diss., Universidade de São Paulo, 2003.

[46] M.K. Venezuela, D.A. Botter and M.C. Sandoval, Diagnostic techniques in generalized estimating equations, J. Stat. Comput. Simul. 77 (2007), pp. 879–888.

[47] M.K. Venezuela, M.C. Sandoval and D.A. Botter, Local influence in estimating equations, Comput. Stat. Data. Anal. 55 (2011), pp. 1867–1883.

[48] A. Wald, Tests of statistical hypotheses concerning several parameters when the number of observations is large, Trans. Am. Math. Soc. 54 (1943), pp. 426–482.

[49] M. Wang, Generalized estimating equations in longitudinal data analysis: a review and recent developments, Adv. Stat. 2014 (2014).

[50] Y.G. Wang and L.Y. Hin, Modeling strategies in longitudinal data analysis: covariate, variance function and correlation structure selection, Comput. Stat. Data. Anal. 54 (2010), pp. 3359–3370.

[51] B.L. Wiens, When log-normal and gamma models give different results: a case study, Am. Stat. 53 (1999), pp. 89–93.