The Evaluation of integrals of the form $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(s,t)\exp\left\{-\frac{1}{2(1-\rho^2)}\left[s^2+t^2-2\rho st\right]\right\}dsdt$: Application to Correlated bivariate logistic-Gaussian Models.

Edmund Essah Ameyaw* Paul Bezandry† Victor Apprey‡ John Kwagyan§

Abstract

The Logistic-Gaussian distribution is used in statistical applications to account for clustering among binary outcomes. However, its extension and applicability to bivariate outcomes is limited. We developed a model for correlated bivariate binary data that incorporated the Logistic-Gaussian distribution. A bivariate normally distributed variate is decomposed into a product of two univariate normally distributed variate and applied to the development of a correlated bivariate logistic Gaussian model. Bivariate response probabilities in terms of random effects models are formulated, and maximum marginal likelihood estimation procedures based on Gauss-Hermite quadrature are used. Application to the analysis of vision loss in diabetic retinopathy is discussed.

Key Words: Correlated Data; Logistic -Gaussian Distribution; Maximum Marginal Likelihood; Bivariate Binary Outcomes; Gauss-Hermite quadrature.

1. Introduction

Bivariate outcomes involve the analysis of two response variables for the purpose of determining the empirical relationship between them.(Babie,2009). Clustered bivariate binary outcome has its applications in various disciplines. In public health the bivariate outcome of HIV and HCV could be assessed in a cluster of geographical regions (Del Fava et al., 2011). In toxicology, a bivariate outcome of fetal weight (high or low) and a binary malformation status can be observed on offspring of mice clustered in litters (Catalan, 1997). In economics, bivariate outcomes of economic recession and economic growth rate cycle can be measured across various US states. In psychology, the bivariate outcome of social anxiety and persecutory ideation can be assessed on a range of psychological factors (Freeman, 2008).

Although clustered bivariate outcomes are common, there are limited statistical approaches to evaluate data of such nature. Some of the traditional approaches include: Dale's model (Dale, 1986) which uses the proportional odds model on the distribution of a bivariate-ordered response vector; alternating logistic regression. Carey (1993) which models association among responses in terms of pairwise odds ratios and the Del Fava model which uses generalized mixed model (Del Fava et al., 2014).

Bonney (2003) introduced the disposition model to account for correlation of binary outcomes within clusters, taking explanatory variables into consideration. We will adopt the disposition model for correlated outcomes and extend it to bivariate outcomes.

*Department of Mathematics, Howard University
†Department of Mathematics, Howard University
‡College of Medicine, Howard University.
§College of Medicine, Howard University.

2. Modelling Consideration

Suppose that data consist of N clusters each of size $n_i, i = 1, ...N$. Let $Y_i = (y_{1i}.....y_{ni})$ be a vector of binary outcome on the i^{th} cluster . Assume further that a pair of observed response within the same cluster satisfy the following relation:

$$\frac{P(Y_j = 1, Y'_{j'} = 1)}{P(Y_j = 1)P(Y_{j'} = 1)} = \frac{1}{\alpha_i}, \alpha_i > 0, j \neq j', j, j' = 1, 2, ...n,$$

where α_i is assumed common for all pairs. Clearly, α_i = 1 implies independence of the observations. Thus α_i is a measure of departure from independence.

Let us further assume that

$$\delta_j = P(Y_j = 1|Y_{j'} = 1), j \neq j', j, j' = 1, 2, ...n$$

With the above definition, Bonney(2003) showed that the joint distribution of the i^{th} cluster Y_i is given by

$$P(Y_i) = P(Y_1 = y_{1,.....,}Y_n = y_n) = (1 - \alpha)\prod_{j=1}^{n}(1 - y_j) + \alpha\prod_{j=1}^{n}\delta^{y_j}(1 - \delta_j)^{1-y_j} . \quad (1)$$

and parametrized δ_j in terms of covariates as

$$\log it(\delta_{ij}) = \boldsymbol{\beta}\boldsymbol{X}_{ij}$$

where $\boldsymbol{X}_{ij}$ is a set of covariates.

Kwagyan(2001,2016) reparametrized δ_{ij} to include a random effect term to account for excess heterogeneity across clusters as

$$logit(\delta_{ij}) = \beta\mathbf{X} + a_i\,, a_i \sim N(0, \sigma^2)$$

and showed that the joint marginal distribution Y_i for N clusters is

$$P(Y_i|\theta) = \prod_{i=1}^{N}\left\{(1 - \alpha)\prod_{j=1}^{n_i}(1 - y_{ij}) + \alpha\int_{-\infty}^{\infty}\left[\prod_{j=1}^{n_i}\delta_{ij}^{y_{ij}}(1 - \delta_{ij})^{1-y_{ij}}f(a_i\boldsymbol{\sigma}^2)\right]da_i\right\} \quad (2)$$

and the log likelihood, log L of (2) is approximated using the Gauss-Hermite quadrature as

$$l = LogL = \sum_{i=1}^{N}\log\left\{(1 - \alpha)\prod_{j=1}^{n_i}(1 - y_{ij}) + \frac{\alpha}{\sqrt{\pi}}\sum_{m=1}^{M}w_m\left[\prod_{j=1}^{n_i}\delta_{ijm}^{y_{ij}}(1 - \delta_{ijm})^{1-y_{ij}})\right]\right\}$$

where

$$\delta_{ijm} = \frac{1}{1 + exp[-(\beta\mathbf{X}_{ij} + \sqrt{2}\sigma_{v_m}]}$$

and (w_n, v_m) are quadrature weights and nodes, respectively,

3. Extension to Correlated Bivariate Binary Outcomes

We extend the model discussed in section 2 to bivariate binary outcomes. Let $(\mathbf{Y}_1, \mathbf{Y}_2)$ such that $\mathbf{Y}_1 = (y_{11i}, y_{12i}, ..., y_{1n_i i})$, $\mathbf{Y}_2 = (y_{21i}, y_{22i}, ..., y_{2n_i i})$ be a pair of bivariate binary outcomes on the i^{th} cluster with size n_i, Let $\mathbf{X} = (X_1, ..., X_p)$ be a set of covariates measured on the i^{th} cluster. Let $\delta_{ij1} = P(Y_{ij1} = 1|Y_{ij'1} = 1)$ and $\delta_{ij2} = P(Y_{ij2} = 1|Y_{ij'2} = 1)$, then δ_{ij1} and δ_{ij2} are modelled in terms of $\mathbf{X}$ and random effects as

$$logit(\delta_{ij1}|a_i) = \boldsymbol{\beta} X_1 + a_i, \;\; a_i \sim \;\; N(0, \sigma_a)$$
$$logit(\delta_{ij2}|b_i) = \boldsymbol{\beta} X_2 + b_i, \;\; b_i \sim \;\; N(0, \sigma_b)$$

Suppose $\mathbf{Y}_1$ and $\mathbf{Y}_2$ are correlated such that

$$\begin{pmatrix} a_i \\ b_i \end{pmatrix} \sim BVN \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_a^2 & \rho_{ab}\sigma_a\sigma_b \\ \rho_{ab}\sigma_a\sigma_b & \sigma_b^2 \end{pmatrix} \right]$$

The diagonal σ_a^2 and σ_b^2, account for the within-cluster correlation for a particular outcome, whereas the parameter ρ_{ab} accounts for the association between the two outcomes.

If we let $\boldsymbol{\theta} = \{\alpha, \beta, \rho_{ab}, \sigma_a, \sigma_b\}$ be the parameters to be estimated, then the conditional distribution of $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2)$, given a_i and b_i is

$$P(\mathbf{Y}_1, \mathbf{Y}_2|a_i, b_i; \boldsymbol{\theta}) = \prod_{k=1}^{2} \left\{ (1-\alpha_i) \prod_{j=1}^{n_i} (1 - y_{ijk}) + \prod_{j=1}^{n_i} \delta_{ijk}^{y_{ijk}} (1-\delta_{ijk})^{1-y_{ijk}} \right\}$$

The (marginal) distribution of $\mathbf{Y}$ for the i^{th} cluster is found by integrating out of the conditional distribution with respect to the unobserved variables a_i and b_i and is given by

$$P(\mathbf{Y}_1\mathbf{Y}_2|\boldsymbol{\theta}) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \{P(\mathbf{Y}_1\mathbf{Y}_2|a_i, b_i; \theta) f((a_i, b_i; \rho_{ab}\sigma_a\sigma_b)\} \, da_i db_i$$

where

$$f(a_i, b_i; \rho_{ab,}\sigma_a, \sigma_b) = C \exp \left\{ -\frac{1}{1-\rho_{ab}^2} \left[\left(\frac{a_i}{\sqrt{2}\sigma_a} \right)^2 - 2\rho_{ab} \left(\frac{a_i}{\sqrt{2}\sigma_a} \right) \left(\frac{b_i}{\sqrt{2}\sigma_b} \right) + \left(\frac{b_i}{\sqrt{2}\sigma_b} \right)^2 \right] \right\} \tag{4}$$

and

$$C = \frac{1}{2\pi\sigma_a\sigma_b\sqrt{(1-\rho_{ab}^2)}}$$

$$-\infty < a_i, b_i < \infty$$

Equation (3) can be written as

$$P(\mathbf{Y}_{1,}\mathbf{Y}_2|\boldsymbol{\theta}) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \{[L_{1i} + \alpha B(a_i)] \, [L_{2i} + \alpha B(b_i)] \, f((a_i, b_i; \rho_{ab,}\sigma_a, \sigma_b)\} \, da_i db_i \tag{5}$$

$L_{1i} = (1-\alpha)\prod_{j=1}^{n_i}(1-y_{ij1})$, $L_{2i} = (1-\alpha)\prod_{j=1}^{n_i}(1-y_{ij2})$,
$B(a_i) = \prod_{j=1}^{n_i}\delta_{ij1}^{y_{ij1}}(1-\delta_{ij1})^{1-y_{ij1}}$ and $B(b_i) = \prod_{j=1}^{n_i}\delta_{ij2}^{y_{ij2}}(1-\delta_{ij2})^{1-y_{ij2}}$

The density function equation (4) can be written as

$$f(a_i, b_i; \rho_{ab}\sigma_a, \sigma_b) = C\exp - \left[\left(\frac{a_i}{\sqrt{2}\sigma_a}\right)^2 + \frac{1}{1-\rho_{ab}^2}\left(-\rho_{ab}\frac{a_i}{\sqrt{2}\sigma_a} + \frac{b_i}{\sqrt{2}\sigma_b}\right)^2\right]$$

$$-\infty < a_i, b_i < \infty$$

3.1 Transformation and Decomposition

We introduce the transformation $(a_i, b_i) \rightarrow (T_i, Z_i)$ such that

$$T_i = \frac{a_i}{\sqrt{2}\sigma_a}$$
$$Z_i = -\frac{\rho_{ab}}{\sqrt{1-\rho_{ab}^2}}T_1 + \frac{1}{\sqrt{1-\rho_{ab}^2}}\frac{b_i}{\sqrt{2}\sigma_b}$$

This implies

$$b_i = \sqrt{2}\sigma_b\rho_{ab}T_i + \sqrt{2}\sigma_b\sqrt{1-\rho_{ab}^2}Z_i)$$
$$a_i = T_i\sqrt{2}\sigma_a$$

The Jacobian of the transformation,J,

$$J = \begin{bmatrix} \sqrt{2}\sigma_a & 0 \\ \sqrt{2}\sigma_b\rho_{ab} & \sqrt{2}\sigma_b\sqrt{1-\rho_{ab}^2} \end{bmatrix} = 2\sigma_a\sigma_b\sqrt{1-\rho_{ab}^2}$$

And so in terms of the transformation variables T_i, Z_i the density equation (4) is

$$f(T_i, Z_i) = \frac{1}{\pi}exp\left\{-(T_i^2 + Z_i^2)\right\} = \frac{1}{\sqrt{\pi}}exp\left[-T_i^2\right] \cdot \frac{1}{\sqrt{\pi}}exp\left[-Z_i^2\right]$$
$$f(T_i, Z_i) = f_T(T_i) \cdot f_Z(Z_i)$$

where

$$f_T(T_i) = \frac{1}{\sqrt{\pi}}exp\left[-T_i^2\right], \quad -\infty < T_i < \infty$$
$$f_Z(Z_i) = \frac{1}{\sqrt{\pi}}exp\left[-Z_i^2\right], \quad -\infty < Z_i < \infty$$

replacing $f((a_i, b_i; \rho_{ab,}\sigma_a, \sigma_b)$ by the transformed form $f(T_i)$ and $f(Z_i)$ in equation (5) as

$$P(\mathbf{Y}_{1,}\mathbf{Y}_2|\boldsymbol{\theta}) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\{[L_{1i}+\alpha B(T_i)]\,[L_{2i}+\alpha B(Z_i)]\,f((T_i)f(Z_2)\}\,dT_i dZ_i$$

which can be written as

$$P(\mathbf{Y}_{1,}\mathbf{Y}_2|\boldsymbol{\theta}) = \left\{L_{1i}+\alpha_i\int_{-\infty}^{+\infty}[B(T_i)f((T_i)]\,dT_i\right\}\left\{L_{2i}+\alpha_i\int_{-\infty}^{+\infty}[B(Z_i)f(Z_i)]\,dZ_i\right\} \tag{6}$$

When the dimension of the random effects is one or two, numerical integration techniques can be implemented reasonably easily and will be used. Since the integrals are over normal densities, Gaussian quadrature is used. Using an M-point Gaussian quadrature; an integral of the form $\int f(x)\phi(x)dx$, where $\phi(x)$ is the bivariate normal density is approximated by the weighted sum;

$$\int e^{-x^2}f(x)dx \approx \sum_{\mathbf{m=1}}^{\mathbf{M}}\mathbf{w_m}f(x_m)$$

where x_m are the Gaussian quadrature points and W_m the associated weights. The terms x_m and w_m are available from (Abramowitz and Stegun, 1972). Equation (6) becomes

$$P(\mathbf{Y}_{1,}\mathbf{Y}_2|\boldsymbol{\theta}) \approx \left[L_{1i}+\frac{\alpha_i}{\sqrt{\pi}}\sum_{\mathbf{m=1}}^{\mathbf{M}}\mathbf{w_m}B(T_m)\right]\left[L_{2i}+\frac{\alpha_i}{\sqrt{\pi}}\sum_{\mathbf{n=1}}^{\mathbf{M}}\mathbf{w_n}B(Z_n)\right] \tag{7}$$

$B(T_m) = \prod_{\mathbf{j=1}}^{\mathbf{n_i}}\boldsymbol{\delta}_{\mathbf{ij1m}}^{\mathbf{y_{ij1}}}(\mathbf{1}-\boldsymbol{\delta}_{\mathbf{ij1m}})^{\mathbf{1-y_{ij1}}}$ and $B(Z_n) = \prod_{\mathbf{j=1}}^{\mathbf{n_i}}\boldsymbol{\delta}_{\mathbf{ij2n}}^{\mathbf{y_{ij2}}}(\mathbf{1}-\boldsymbol{\delta}_{\mathbf{ij2n}})^{1-y_{ij2}}$

where

$$\delta_{ij1m}(\beta,\sigma_a) = \frac{1}{1+exp[-(\beta X+T_m\sqrt{2}\sigma_a]}$$

$$\delta_{ij2n}(\beta,\sigma_b) = \frac{1}{1+exp[-(\beta X+\sqrt{2}\sigma_b\rho_{ab}T_m+\sqrt{2}\sigma_b\sqrt{1-\rho_{ab}^2}Z_n)]}$$

For N clusters, equation (7) becomes

$$P(\mathbf{Y}_{1,}\mathbf{Y}_2|\boldsymbol{\theta}) \approx \prod_{i=1}^{N}\left\{\left[L_1+\frac{\alpha_i}{\sqrt{\pi}}\sum_{\mathbf{m=1}}^{\mathbf{M}}\mathbf{w_m}B(T_m)\right]\left[L_2+\frac{\alpha_i}{\sqrt{\pi}}\sum_{\mathbf{n=1}}^{\mathbf{M}}\mathbf{w_n}B(Z_n)\right]\right\}$$

taking log of the likelihood, l

$$l = \log P(\mathbf{Y}_{1,}\mathbf{Y}_2|\boldsymbol{\theta}) \approx \sum_{\mathbf{i=1}}^{\mathbf{N}}\log\left\{\left[L_1+\frac{\alpha_i}{\sqrt{\pi}}\sum_{\mathbf{m=1}}^{\mathbf{M}}\mathbf{w_m}B(T_m)\right]\left[L_{2i}+\frac{\alpha_i}{\sqrt{\pi}}\sum_{\mathbf{n=1}}^{\mathbf{M}}\mathbf{w_n}B(Z_n)\right]\right\} \tag{8}$$

4. Application to Analysis of Vision Loss on Diabetic Retinopathy

Data was collected from 7151 participants with type 2 diabetes in the ACCORD EYE substudy. Participants were randomized to 8 treatment groups-[Intensive Glycemia/Lipid Fibrate],[Intensive Glycemia/Lipid Placebo],[Intensive Gylcemia/Intensive Blood Pressure],[Intensive Gylcemia/Standard Blood Pressure],[Standard Glycemia/Lipid Fibrate],[Standard Glycemia/Lipid Placebo],[Standard Gylcemia/Intensive Blood Pressure] and [Standard Gylcemia/Standard Blood Pressure](ACCORD Eye Study Group, 2010). In this analysis, we consider as a cluster, the treatment group, and assess aggregation of the vision loss adjusting for measured risk factors. The participants within a treatment group has correlated outcomes which are influenced part or wholly by the treatment as well as the variables on the individual respondents. The following covariates are available: Smoking Status, LSmoker is coded 1 if smoked more than 100 cigarettes during lifetime and 0 otherwise. Years of diabetes, (Ydiab), number of years participants has been suffering from diabetes. Neuropathy is coded 1, if participant had nerve pains and 0 otherwise. The bivariate outcome variables are vision loss in the Left, Y_L, or Right, Y_R and coded 1 for vision loss and 0 otherwise. The set of possible outcomes are $(Y_L, Y_R) = \{(0,0), (0,1), (1,0), (1,1)\}$. Assuming a single random effect model the general model for predicting an individual bivariate response to the Left and Right vision loss, accounting for potential treatment to treatment heterogeneity while adjusting covariates effects is given as

$$
\begin{aligned}
&logit(\delta_{ijL}|a_i) = \beta_{0L} + \beta_{1L} * (LSmoker)_L + \beta_{2L} * (Neuropathy)_L + \beta_{3L} * (Ydiab)_L + a_i \\
&logit(\delta_{ijR}|b_i) = \beta_{0R} + \beta_{1R} * (LSmoker)_R + \beta_{2R} * (Neuropathy)_R + \beta_{3R} * (Ydiab)_R + b_i \\
&0 \leq \alpha_1 \leq 1,\ 0 \leq \alpha_2 \leq 1,\ 0 \leq \rho_{ab} \leq 1
\end{aligned}
$$

Suppose Y_L and Y_R are correlated such that

$$
\begin{pmatrix} a_i \\ b_i \end{pmatrix} \sim BVN\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_a^2 & \rho_{ab}\sigma_a\sigma_b \\ \rho_{ab}\sigma_a\sigma_b & \sigma_b^2 \end{pmatrix}\right]
$$

The following describes specific fitted models

1. Model I assumes, no correlation within treatment, no correlation among treatments and no association between Left vision loss and Right vision loss. Thus if $\alpha_1 = \alpha_2 = 1$, $\sigma_a = \sigma_b = 0$ and $\rho_{ab} = 0$, then

$$
\begin{aligned}
&logit(\delta_{ij1}|a_i) = \beta_{0L} + \beta_{1L} * (LSmoker)_L + \beta_{2L} * (Neuro)_L + \beta_{3L} * (Ydiab)_L \\
&logit(\delta_{ij2}|b_i) = \beta_{0R} + \beta_{1R} * (LSmoker)_R + \beta_{2R} * (Neuro)_R + \beta_{3R} * (Ydiab)_R
\end{aligned}
$$

2. Model II assumes, correlation among treatment, no correlation within treatment and no association between Left vision loss and Right vision loss. Thus if $\alpha_1 = \alpha_2 = 1$, $\sigma_a = \sigma_b = free$ and $\rho_{ab} = 0$, then

$$
\begin{aligned}
&logit(\delta_{ijL}|a_i) = \beta_{0L} + \beta_{1L} * (LSmoker)_L + \beta_{2L} * (Neuropathy)_L + \beta_{3L} * (Ydiab)_L + a_i \\
&logit(\delta_{ijR}|b_i) = \beta_{0R} + \beta_{1R} * (LSmoker)_R + \beta_{2R} * (Neuropathy)_R + \beta_{3R} * (Ydiab)_R + b_i \\
&a_i \sim\ N(0, \sigma_a),\ b_i \sim\ N(0, \sigma_b)
\end{aligned}
$$

3. Model III assumes, correlation among treatment, correlation within treatment and no association between Left vision loss and Right vision loss. Thus if $\alpha_1 = \alpha_2 = 1$, $\sigma_a = \sigma_b = free$ and $\rho_{ab} = 0$, then

$$logit(\delta_{ijL}|a_i) = \beta_{0L} + \beta_{1L} * (LSmo\,\mathrm{ker})_L + \beta_{2L} * (Neuropathy)_L + \beta_{3L} * (Ydiab)_L + a_i$$
$$logit(\delta_{ijR}|b_i) = \beta_{0R} + \beta_{1R} * (LSmo\,\mathrm{ker})_R + \beta_{2R} * (Neuropathy)_R + \beta_{3R} * (Ydiab)_R + b_i$$
$$a_i \sim\ N(0, \sigma_a),\ b_i \sim\ N(0, \sigma_b)$$

4. Model IV: assumes, correlation among treatment, correlation within treatment and association between Left vision loss and Right vision loss. Thus if $\alpha_1 = \alpha_2 = 1$, $\sigma_a = \sigma_b = free$ and $\rho_{ab} = free$, then

$$logit(\delta_{ijL}|a_i) = \beta_{0L} + \beta_{1L} * (LSmo\,\mathrm{ker})_L + \beta_{2L} * (Neuropathy)_L + \beta_{3L} * (Ydiab)_L + a_i$$
$$logit(\delta_{ijR}|b_i) = \beta_{0R} + \beta_{1R} * (LSmo\,\mathrm{ker})_R + \beta_{2R} * (Neuropathy)_R + \beta_{3R} * (Ydiab)_R + b_i$$

$$\begin{pmatrix} a_i \\ b_i \end{pmatrix} \sim BVN\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_a^2 & \rho_{ab}\sigma_a\sigma_b \\ \rho_{ab}\sigma_a\sigma_b & \sigma_b^2 \end{pmatrix}\right]$$

Table1, shows the results from the four different submodel. Computations of the proposed models II-IV were performed using computer programs we developed which was linked with likelihood optimization routines using the R-package (Version 3.2.5). For our proposed models, we run the analyses using quadrature points M=8, 9, 10 and 11. The results did not change much for quadrature points, M $>$ 9 and so for computations, M = 9 was employed to complete the analysis. In all of the models we have considered, the algorithm converged in less than 38 iterations.

The odds of a positive response for a Life time smokers to be vision loss is estimated to be exp(0.439) = 1.551 times higher for vision to be loss than non-life smokers in all models. Neuropathy was not significant in all the models. With respect to the participants years of diabetes, it was significant for all three models, indicating that individuals with more years of diabetes are more likely to experience a vision loss. The approach seems to under estimated the parameters in the dependent model. This may be due to the fact that the data were not randomly sampled from the population but from treatment groups where there is a high risk of aggregation of the occurrence of vision loss.Within treatment group dependence is described by the magnitude of the relative dependence parameters, $\alpha_{1,}\alpha_2$ and excess treatment heterogeneity by the magnitude of, σ_a, σ_b the variance component parameters.

The significance of the individual estimates is judged by t tests based on the standard errors. Dependence is described by the magnitude of the relative disposition α_1, α_2. In addition the degree of heterogeneity is measured by the magnitude of σ_a^2,σ_b^2 the variance of the random effect distribution.

Model III and Model IV seems to fits the data better than model II. The deviance (-2Log) obtained from the independent model (Model II) was 1725.607, the deviance under the partial correlated bivariate model 1715.921 and the deviance under the dependence model is 1715.921. For this model, the maximum likelihood estimate of the relative dependence parameter, $\alpha_{1,} = \alpha_2$ was exp(0.121)=0.887 and a computed standard error (0.086) for both the partial dependence and the dependence model . This suggests that the data was sampled from a population where the aggregation of the treatment group is higher than that from the general population. Similarly, the estimate of $\sigma_a = \sigma_b$, the variance component parameters is 0.681 with standard error of 0.086 for model III and Model IV. Thus the data further suggests some degree of heterogeneity of outcomes across treatment groups.

The odds of a positive response for a Life time smokers to vision loss is estimated to be exp(0.439) = 1.551 times higher for vision to be loss than non-life smokers in all models.

Neuropathy was not significant in all the models. With respect to the participants years of diabetes, it was significant for all three models, indicating that individuals with more years of diabetes are more likely to experience a vision loss. The approach seems to under estimated the parameters in the dependent model. This may be due to the fact that the data were not randomly sampled from the population but from treatment groups where there is a high risk of aggregation of the occurrence of vision loss. In summary, we conclude that vision loss aggregates in the treatment group sampled

5. Conclusions and Future Work

This paper has been concerned with development of a likelihood formulation for clustered bivariate binary data. The development, albeit being straight forward and based on simple analytic formulation, is novel and well suited for areas of application including public health and biomedical research. We demonstrated that the proposed logistic-Gaussian random effects model provides a useful tool for analysing clustered bivariate binary data. The advantage of the proposed model is that, it accounts for within cluster (treatment) dependence, provides a good portrayal of cluster (treatment) differences. In many applications, the use of regressor variables is inevitable. Regression parametrization of the response probability is modelled in terms of random effects. The choice of a particular model, for a given dataset should be guided by the purpose of the analysis. For example the partial correlated bivariate binary model provides a good portrayal of cluster differences while controlling for within cluster aggregation. We have also shown that the parameters in the models can be estimated by maximum likelihood methods. Iterative procedures to produce estimates are derived from maximum likelihood methods (ML). Numerical approximation methods based on Gaussian quadratures was utilized for estimation in the random effect models where closed form results are intractable. Other approaches such as the Gibbs sampling method, Monte Carlo techniques can be considered in future studies.

REFERENCES

ACCORD Eye Study Group. (2010). Effects of medical therapies on retinopathy progression in type 2 diabetes. The New England journal of medicine, 363(3), 233.

Babie, E. (2009). The practice of social research. Wadsworth Publishing.

Carey et.al (1993) .Carey, V., Zeger, S.L. and Diggle, P. (1993). Modeling multivariate binary data with alternating logistic regressions. Biometrica, 80(30), 517-526.

Catalan, P. J. (1997). Bivariate modelling of clustered continuous and ordered categorical outcome. Statistics in Medicine .Vol 16.

Connolly, M.A. and Liang K.Y. (1988). Conditional logistic regression models for correlated binary data.

Bonney, G.E. (2003). Disposition to a correlated binary outcome and its regression analysis *Journal of Statistics,Biometry and Genetics,* 1, 1-30.

Cox, D.R. (1958). The regression analysis of binary sequences. *J. Roy. Statist. Soc., B,* 20, 215-242.

Dale, J. R. (1986). Global cross-ratio models for bivariate, discrete, ordered responses. Biometrics 42,909-917.

Del Fava E, Kasim A, Usman M, Shkedy Z, Hens N, Aerts M, Bollaerts K, Scalia Tomba G, Vickerman P, Sutton AJ, Wiessing L, and Kretzschmar M (2011) Joint Modeling of HCV and HIV Infections among Injecting Drug Users in Italy Using Repeated Cross-Sectional Prevalence Data. *Statistical Communications in Infectious Diseases*, 3, 1–24.

Del Fava E, Kasim A, Shkedy Z, Hens N,Mehreteab Aregay,and Geert Molenberghs (2014) Modelling multivariate, overdispersed binomial data with additive and multiplicative random effects.Statistical Modelling 2014; 14(2): 99–133

Kwagyan, J. (2001). Further investigations of the disposition models for correlated binary outcomes. PhD dissertation, The Temple University Graduate School, USA.

D. Freeman, M. G. (August 2008,). What makes one person paranoid and another person anxious?. The differential prediction of social anxiety and persecutory ideation in an exp

Kwagyan, J. (2016). A Logistic-Gaussian Model for Clustered Binary Data with Excess Zero Clusters. *Statistical Methods in Medical Research.*

Li, J. (2006). *Analysis of Longitudinal data with missing values.* PhD dissertation, University of California, Los Angeles USA.

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.F. (1992). Multivariate regression analysis for categorical data. *J. of the Roy. Statistical Soc. B,* 54,3-40.

Lipsitz, S.R., Laird, N.M., and Harrington, D.P. (1991) Generalized estimating equations for correlated binary data: using the odds ratio as a measure of association.*Biometrika*, 78, 153-160.

McCullagh,P. and Nelder, J.A. (1989). *Generalized Linear Models,* 2 ed. Chapman and Hall, London.

Molenberghs, G. and Verbeke, G. (2005). *Models for discrete longitudinal data.* Springer, New York USA.

Paige L. Williams, H. G. (2002). Topics in Modelling of Clustered Data. Boca Raton : Chapman and Hall.

Prentice R.L. (1988). Correlated binary regression with covariates specific to each binary observation.*Biometrics*, 44,1033-1048.

Qaqish, B.F., and Liang, K.Y. (1992). Marginal models for correlated binary responses with multiple classes and multiple levels of nesting. *Biometrics*, 48, 939-950.

Zeger, C.L., Liang, K.Y. and Albert, P.S. (1988). Models for longitudinal data: A generalized estimating equation approach. *Biometrics*, 44(4), 1049-1060.

Table 1: Estimates and Standard Errors of the Regression Analysis of Left and Right Vision Loss (ACCORD EYE Data)

	Model I		Model II		Model III		Model IV	
Parameter	Estimates	S.E	Estimates	S.E	Estimates	S.E	Estimates	S.E
$\sigma_a = \sigma_b$			0.773	0.075*	0.681	0.085*	0.681	0.121*
$\alpha_1 = \alpha_2$					0.121	0.086*	0.121	0.086*
ρ_{ab}							1.000	0.251*
LEFT VISION LOSS								
Constant	-4.090	0.142*	-3.017	0.214*	-2.919	0.218*	-2.919	0.226*
Lifetime Smoker	0.466	0.126*	0.439	0.184*	0.437	0.184*	0.437	0.184*
Neuropathy	0.154	0.140	0.169	0.203	0.165	0.203	0.165	0.203
Years of Diabetes	0.047	0.008*	0.046	0.011*	0.045	0.011*	0.045	0.011*
RIGHT VISION LOSS								
Constant	-4.090	0.142*	-3.017	0.216*	-2.919	0.217*	-2.919	0.227*
Lifetime Smoker	0.466	0.126*	0.439	0.184*	0.437	0.184*	0.437	0.183*
Neuropathy	0.154	0.140	0.169	0.203	0.165	0.203	0.165	0.203
Years of Diabetes	0.047	0.008*	0.046	0.011*	0.045	0.011*	0.045	0.011*
-2*Log(Likelihood)	907.36014		1725.607		1715.921		1715.921	

*indicates statistical significance. Original data was modified to illustrate Model II-IV

Model I: No correlation within treatment, no correlation among treatments and no association between Left vision loss and Right vision loss.

Model II: Correlation among treatment, no correlation within treatment and no association between Left vision loss and Right vision loss

Model III: Correlation among treatment, correlation within treatment and no association between Left vision loss and Right vision loss

Model IV: Correlation among treatment, correlation within treatment and association between Left vision loss and Right vision loss