Testing of Gender Differences on Sib-Sib Correlations for Binary Traits: Likelihood Based Inference with Application to Arterial Blood Pressures Data

Mohamed Shoukri1,2, K Collison1,3 and F Al-Mohanna1,2,3

1National Biotechnology Center, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia
2Department of Cell Biology, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia
3College of Medicine Al-Faisal University, Riyadh, Saudi Arabia

Abstract

Estimation of measures of familial aggregation is considered the first step in establishing whether a specified disease has a genetic component. Population-based family study designs are usually used to estimate correlations among siblings. When the trait of interest is quantitative (e.g. blood pressure, body mass index, blood glucose level) testing the effect of gender differences on sib-sib correlations is achieved using the likelihood method of estimation under the assumption of multivariate normality. When the trait of interest is measured on the binary scale testing the equality of a brother-brother and sister-sister correlation is more complex. In this paper we develop likelihood-based inference procedures for this purpose which may be applied to nuclear family data.

Keywords: Family studies; Multivariate beta-binomial distribution; Likelihood estimation; Score test

Introduction

Population-based family studies have been used in genetic epidemiology to assess the association of environmental risk factors with disease and to quantify the aggregation of cases within families. These types of studies integrate statistical methods and classical epidemiology to analyze the correlations among family members who share the same genetic and environmental background. There are several advantages in using family study designs. The use of extended pedigrees or even nuclear families enhances the statistical power for gene discovery. Clinical characteristics common to family members may also be used to increase information by defining subgroups of families for analysis such as in the investigation of familial aggregation of components of metabolic syndrome [1]. Another important feature of family studies in contrast to studies of unrelated individuals is the issue of internal control. The analysis of traits of interest for family members should account for both genetic background factors and environmental exposures. A common example is the case of monozygotic twins where maximum genetic control is achieved. It is well-known that nuclear family members tend to have relatively similar environmental conditions, diet, and perhaps levels of physical activity. The familial aggregation of much chronic and infectious disease is also well documented. For example, results from recent studies have shown that pathogens causing Th1 diseases are passed from parents to child. Information accessed on December 23-2012 from (http://bacteriality.com/2008/07/31/hpv/) reveals that some of the chronic viral species that cause inflammatory illness can remain alive in breast milk and thus be passed from mother to child through breast feeding. Growing evidence suggests that the Th1 pathogens, rather than genetic mutation, are the driving force behind this familial aggregation. Although their role is uncertain, researchers have also found a relationship between bacterial infection and cancer [2]. This chain of reasoning provides a possible explanation for the aggregation of cancer in families. A study conducted in 2010 using the PET scanner to examine the prevalence of plaque in brains (which is the hallmark of Alzheimer’s disease) found that a child’s level of plaque is consistent with the corresponding levels of their fathers and in particular of their mothers, even years before the child’s diagnosis [3].

In many population-based family studies, interest is focused in detecting gender differences in the risk of developing a chronic disease. For example, a recent study [4] aiming at examining sex-specific associations between cardiovascular risk factors and type 2 diabetes mellitus showed that there are gender-related dissimilarities that are apparently involved in disease development. Another study conducted on a sample of families from South Australia [5] found that men and women face different challenges in the management of diabetes and its associated complications.

One of the major limitations of the above studies is that the comparisons between males and females were based on parallel group designs, and consequently suffer from the lack of control over possible confounding. Another limitation is that the lack of a reference population makes the problem of statistical inference (estimation and hypothesis testing) less meaningful. A further methodological challenge that faces researchers is that estimates of trait correlations, specifically the intra-class correlations for males and females, are themselves correlated. Although studies on comparing sib-sib correlations have been of frequent interest [6,7] comparisons among these correlations have been usually made descriptively. When traits are measured on the continuous scale, Donner et al. [8] developed several procedures for comparing the sib-sib correlations among males and females, including likelihood-based tests, while assuming that the underlying mechanism generating the data is multivariate normal. When the trait of interest is measured the binary scale, efficient methods for comparing sib-

*Corresponding author: Mohamed Shoukri, National Biotechnology Center-King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia, Tel. +966-509491454; E-mail: shoukri@kfshrc.edu.sa

Received January 16, 2014; Accepted February 22, 2014; Published February 28, 2014


Copyright: © 2014 Shoukri M, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
sib correlations characterizing males and female have not yet been developed.

This paper has a threefold objective. First, we develop a multivariate probability distribution for the vector of binary observations based on a random sample of independent sib-ships. The vector will be split into two sub-clusters, separating female responses from male responses. Secondly we construct the likelihood function of the sample as based on the joint distribution of the created sub-clusters. This allows us to develop score and Wald chi-square-tests of significance that compare the levels of similarity among males and females from the same family. Finally, we illustrate our procedures using published arterial blood pressures data.

**Models**

Suppose we have a random sample of k sib-ships, where each sib-ship constitutes a cluster. Let \(y_i=(y_{ij1}, y_{ij2}, \ldots, y_{ijb}, y_{ijsi}, y_{ij2}, \ldots, y_{ij2, si})\) denote the vector of observations from the \(i\)th cluster, where \(b\) = number of brothers in the \(i\)th family, \(s_i\) = number of sisters in the \(i\)th family, \(b = \sum_{i=1}^{k} b_i\) = number of brothers in the sample of \(k\) families, \(s = \sum_{i=1}^{k} s_i\) = number of sisters in the sample of \(k\) families, and \(N=b+s\) = number of siblings in the \(k\) families. It is clear then that each cluster (sib-ship) constitutes a cluster. Let \(y_{ij}=(y_{ij1}, y_{ij2}, \ldots, y_{ijb}, y_{ijsi})\) denote the probability that a randomly selected brother from the \(i\)th family is classified as having the trait of interest, and let \(\varphi_i=(\mu_i, \rho_b, \rho_s, \rho_{12})\) denote the vector of parameters for the \(i\)th family. The resulting representation is given by:

\[
E(y_{ib})=b \mu_i \\
\sigma_{yi}^2 = Var(y_{ib}) = \mu_i (1-\mu_i) [1 + (b_i-1) \rho_b]
\]

Similarly, the unconditional distribution of \(X_{is} = \sum_{j=1}^{b_i} y_{ij} \) will be that of a beta-binomial distribution with:

\[
E(x_{is})=s \mu_i \\
\sigma_{xi}^2 = Var(x_{is}) = \mu_i (1-\mu_i) [1 + (s_i-1) \rho_s]
\]

Note that, because of the exchangeability condition, the unconditional distribution of \(Y_{ib} = \sum_{j=1}^{b_i} y_{ij} \) is that of a beta-binomial distribution with:

\[
E(y_{ib})=b \mu_i \\
\sigma_{yi}^2 = Var(y_{ib}) = \mu_i (1-\mu_i) [1 + (b_i-1) \rho_b]
\]

Details may be found in references [9-11]. The beta-binomial probability distributions of \(y_{ib}\) and \(x_{is}\) are given respectively as:

\[
p(y_{ij}=(0|\mu_i, \rho_b, \rho_s, \rho_{12})) = \mu_i (1-\mu_i) [1 + (b_i-1) \rho_b] \\
p(x_{ij}=(0|\mu_i, \rho_b, \rho_s, \rho_{12})) = \mu_i (1-\mu_i) [1 + (s_i-1) \rho_s]
\]

We first construct a bivariate distribution based on the marginal distributions. This allows us to develop score and Wald chi-square-tests of significance that compare the levels of similarity among males and females from the same family. However a different approach is needed to construct the bivariate distribution of the sibling scores characterized by the interclass correlation. This approach, developed by Sarmanov [12] and Lancaster [13], is known as Positive Dependence by Expansion (PDE). Danaher [14] proposed a simplified and flexible form of the distribution based on Lancaster’s representation.

We are interested in testing the following hypotheses:

1. \(H_0 : \rho_b=\rho_s\)
2. \(H_0 : \rho_s=0\)

As a first step, we follow Lancaster [13] in constructing a bivariate distribution by joining the marginal distributions given in (5) and (6). The resulting representation is given by:

\[
p(x_{ij}, y_{ij}) = p(x_{ij}) p(y_{ij}) \left[ 1 + \rho \left( \frac{y_{ij} - b_i \mu_i}{\sigma_{yi}} \right) \left( \frac{x_{ij} - s_i \mu_i}{\sigma_{xi}} \right) \right] \\
\]

Direct computations show that the common intraclass correlation among siblings in the same cluster is:

\[
Cov(y_{ij}, y_{ij}) = \rho_b
\]

We can show that the population common intraclass correlation among brothers in the same sibship-cluster is:

\[
Cov(y_{ij}, y_{ij}) = \rho_b
\]

and the common intraclass correlation among sisters in the other sub-cluster is:

\[
Cov(x_{ij}, y_{ij}) = \rho_s
\]

We further define the interclass correlation among brothers and sisters as:

\[
\varphi_i=(\mu_i, \rho_b, \rho_s, \rho_{12})
\]

\[
\Phi= (\mu_i, \rho_b, \rho_s, \rho_{12})
\]

**Methods**

Our inferences on the parameters of interest are based on the
likelihood principle. To test the hypothesis $H_0: \rho_{12}=0$ against the alternative $H_1: \rho_{12}>0$, we assume that a random sample of $k$ sib-ships is available. The log-likelihood function of the sample is given by:

$$l = \sum_{i=1}^{k} \log p(x_i, y_i)$$

$$= \sum_{i=1}^{k} \left( \log p(x_i) + \log p(y_i) + \log \left[ 1 + \rho_i H(x_i, y_i) \right] \right),$$

Where, $H(x_i, y_i) = \begin{pmatrix} y_i - \mu_{x_i} & x_i - \mu_{y_i} \\ \sigma_{x_i} & \sigma_{y_i} \end{pmatrix} \begin{pmatrix} y_i - \mu_{x_i} \\ \sigma_{x_i} \end{pmatrix}$. The score function $u = \frac{\partial l}{\partial \rho_{12}} \rho_{12} = 0$ is given by:

$$u = \sum_{i=1}^{k} \begin{pmatrix} y_i - \mu_{x_i} \\ \sigma_{x_i} \end{pmatrix} \begin{pmatrix} y_i - \mu_{y_i} \\ \sigma_{y_i} \end{pmatrix} = \sum_{i=1}^{k} u_i.$$

By the central limit theorem and for fixed $b_i$ and $s_i$, the distribution of each component $u_i$ tends uniformly to the standard normal distribution under $H_0$ as $k \to \infty$. Moreover, we can show that under $H_0$ that the statistic $s_i^2 = u_i$ will be asymptotically distributed as chi-square with one degree of freedom [15]. This statistic is the locally most powerful one-sided test of $H_0: \rho_{12}=0$ against $H_1: \rho_{12}>0$. Full details of the proof can be found [15, 16]. Moran [17] showed that if the hypothesis the asymptotic properties of this test statistic will be preserved. Such consistent estimators can be either the maximum likelihood (MLE) or the moment estimators.

In Table 1 we provide estimates of the sample sizes (number of sib-sips) needed to detect the departure from $H_0: \rho_{12}=0$ in the direction of ℓ with respect to the five parameters. The two matrices are given as:

$I_{21}$ is a $2 \times 2$ and symmetric, $I_{22} = I_{22}^T$ is a $2 \times 3$ matrix and $I_{22}$ is a $3 \times 3$ diagonal matrix. The elements of the covariance matrix are given in closed form as:

$$\sum_{i=1}^{k} = \begin{pmatrix} I_{11} & I_{12} \\
I_{21} & I_{22} \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\
\Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

Here, $\Sigma$ is the variance-covariance matrix of $(\hat{\mu}_x, \hat{\mu}_y, \hat{\sigma}_x, \hat{\sigma}_y, \hat{\rho}_{12})$. To find the elements of $\Sigma$ we use the method of matrix partitioning [18].

Maximum Likelihood Estimation

The MLE’s of the model parameters are obtained by simultaneously solving the likelihood equations:

$$\frac{\partial l}{\partial \mu_i} = 0, \quad \frac{\partial l}{\partial \sigma_i} = 0, \quad \frac{\partial l}{\partial \rho_i} = 0, \quad \frac{\partial l}{\partial \rho_{12}} = 0.$$

We obtain the variance-covariance matrix $\Sigma$ of the MLE’s by inverting the matrix of the negative of the second partial derivatives of $\ell$ with respect to the five parameters. The two matrices are given as:

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\
\Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

Table 1: Sample size requirements for Type I error rate 5% and power 80%. To test the hypothesis $H_0: \rho_{12}=0$. |

<table>
<thead>
<tr>
<th>$\rho_s \rho_b b\times s$</th>
<th>$B=2$</th>
<th>$B=5$</th>
<th>$B=10$</th>
<th>$B=2$</th>
<th>$B=5$</th>
<th>$B=10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>.2 .2 .1</td>
<td>715</td>
<td>667</td>
<td>663</td>
<td>612</td>
<td>612</td>
<td>612</td>
</tr>
<tr>
<td>.2 .5 .1</td>
<td>715</td>
<td>613</td>
<td>613</td>
<td>612</td>
<td>612</td>
<td>612</td>
</tr>
<tr>
<td>.5 .5 .2</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>.8 .8 .8</td>
<td>45</td>
<td>41</td>
<td>41</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>.8 .8 .2</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
</tbody>
</table>

Table 3: Power Calculations for testing $H_0: \rho_{12}=0$. |

<table>
<thead>
<tr>
<th>$\rho_2 \rho_s \rho_{12} b\times s$</th>
<th>$k=25$</th>
<th>$k=50$</th>
<th>$k=100$</th>
</tr>
</thead>
<tbody>
<tr>
<td>.5 .5 .5</td>
<td>0.051</td>
<td>0.051</td>
<td>0.051</td>
</tr>
<tr>
<td>.5 .5 .5</td>
<td>0.336</td>
<td>0.334</td>
<td>0.334</td>
</tr>
<tr>
<td>.8 .8 .8</td>
<td>0.661</td>
<td>0.661</td>
<td>0.660</td>
</tr>
<tr>
<td>.8 .8 .8</td>
<td>0.830</td>
<td>0.830</td>
<td>0.830</td>
</tr>
</tbody>
</table>

Table 4: Power calculations for testing $H_0: \rho_{12}=0, \mu_1=\mu_2=0.5$. Under unbalanced design.
and correlation is zero [19]. This orthogonality also implies that the maximum likelihood estimators of these parameters are asymptotically independent. Therefore, as in Donner et al. [8] we can test the equality of $\rho_b$ and $\rho_s$ independent of the values of $\mu_b$ and $\mu_s$. For the bivariate beta-binomial distribution, the orthogonality condition is not satisfied and therefore we propose an omnibus test in the form:

$$H_o : \mu_b = \mu_s \cap \rho_b = \rho_s$$

(8)

This hypothesis takes into account the correlations among all the estimated parameters. Let $\psi = (\mu_b, \mu_s, \rho_b, \rho_s)$ and consider the affine transformation.

$$H_o: \psi = 0$$

The matrix A has 2 rows and 4 columns and is specified as:

$$A = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

To test the stated hypothesis, an omnibus test statistic is constructed, using the asymptotic distributional properties of the MLE of $\psi$. From Serfling [20], the MLE $\hat{\psi}$ has the asymptotic distribution:

$$\sqrt{n} (\hat{\psi} - \psi) \sim N(0, V)$$

where $V$ is the variance-covariance matrix of $\hat{\psi}$ and is obtained by deleting the 5th row and 5th column of $\Sigma$. Letting $H = AV$, the question reduces to testing $H_o, H = 0$:

$$\hat{H} = A\hat{\psi}$$

is therefore distributed as $\sqrt{n} H \sim \chi^2(4)$. Hence, from Graybill [21] the quadratic form:

$$Q(\epsilon) = \epsilon^T (AV)^T \epsilon$$

is asymptotically non-central chi-square with 2 degrees of freedom and non-centrality parameter $\epsilon = KH^T (AV AT)^{-1} H$

Moreover, $\epsilon=0$ if and only if $H_o$ is true. Hence referring $Q(0)$ to the table of chi-square distribution with 2 degrees of freedom, $H_o; H=0$ is rejected if $Q(0)$ exceeds the tabulated value of a chi-square with 2 degrees of freedom at the chosen level of significance.

Example: Mial and Oldham's blood pressure data

The data used for illustration here are obtained from a survey that aimed at assessing the levels of similarity in systolic and diastolic blood pressure among family members living within 25 miles of Rhonda Fach Valley in South Wales and published by Miall and Oldham [22] previously analyzed [23,24]. Observations were made on parents and their offspring, with each observation consisting of systolic and diastolic blood pressures measured to the nearest 5 mmHg. However among 250 sampled families, only 204 contained information on brothers and sisters. Furthermore, because of the impossibly low systolic blood pressure (15 mmHg) for one daughter, another family was omitted leaving 203 families for the analysis. Since these data were given on a continuous scale, we dichotomized the observations such that for an individual whose systolic/diastolic blood levels above 130/80, the assigned binary score was 1, otherwise, set as 0. The results of the data analysis are summarized in Tables 5 and 6. Table 5 shows the maximum likelihood of the model parameters, together with their standard errors. Table 6 displays the variances and covariances among the estimated parameters using the expression in the Appendix.

The null hypothesis $H_o, \rho_{b,s}=0$ was tested against the one-sided alternative $H_o, \rho_{b,s}>0$ is rejected as $z=4.48$ (p-value=.034). The Wald one degree of freedom chi-square test $w=(0.195/0.095)^2=4.21$, leading to the same conclusion as the score test. To test the null hypothesis (8) we use the statistic given in (9). Direct computation shows that $Q(0)=3.246$, (p=0.097). Therefore, we conclude that there is no sufficient evidence to support the claim of gender differences in the distribution of hypertension based on this data set. An equally important hypothesis to be tested is whether gender has influence on the dependence structure. That is we need to test whether within gender correlations are the same as across gender correlation. This hypothesis can be easily formulated as:

$$H_o: \rho_{b,s}=\rho_{b,s}$$

We may then formulate the simple contrast $\hat{T} = \hat{\rho}_b - \hat{\rho}_s - 2\hat{\rho}_{bs}$. Under the null hypothesis, $T$ is asymptotically unbiased, that is $E(T)=0$, and $\text{var}(T) = \text{var}(\hat{\rho}_b) + \text{var}(\hat{\rho}_s) + 2\text{cov}(\hat{\rho}_b, \hat{\rho}_s) + 4\text{var}(\hat{\rho}_{bs}) - 4\text{cov}(\hat{\rho}_b, \hat{\rho}_{bs}) - 4\text{cov}(\hat{\rho}_s, \hat{\rho}_{bs})$.

Therefore, asymptotically $G = T^2 / \text{var}(T)$ has a chi-square distribution with one-degree of freedom. From the data, G=0.185, and a p-value=0.911. Therefore, based on this data we conclude that the correlations within gender are the same across genders.

Remarks

Originally the Miall and Oldham's data are measured on the continuous scale. Under the assumption of multivariate normality Mian and Shoukri [24] used the MLE to produce the following estimates for the within gender and across gender correlations. We summarize the results in Table 7. For testing $H_o: \rho_{b,s}=0$, one degree of freedom chi-square test statistic is 61, with p-value<0.00001. Similar to the above approach, $H_o: \rho_{b,s}=0$ is rejected (p-value<0.00001). Similarly $H_o: \rho_{b,s}=\rho_{b,s}$ has a chi-square value=26.48, with p-value<0.00001.

It is clear that the dichotomization resulted in a reduction in

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate ± SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_b$</td>
<td>294 ± 0.028</td>
</tr>
<tr>
<td>$\mu_s$</td>
<td>200 ± 0.077</td>
</tr>
<tr>
<td>$\rho_b$</td>
<td>217 ± 0.026</td>
</tr>
<tr>
<td>$\rho_s$</td>
<td>274 ± 0.071</td>
</tr>
<tr>
<td>$\rho_{bs}$</td>
<td>.195 ± .095</td>
</tr>
</tbody>
</table>

Table 5: Estimates of the model of parameters ± standard error.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Systolic</th>
<th>Diastolic</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_b$</td>
<td>0.146 ± 0.073</td>
<td>0.163 ± 0.073</td>
</tr>
<tr>
<td>$\rho_s$</td>
<td>0.32 ± 0.069</td>
<td>0.248 ± 0.070</td>
</tr>
<tr>
<td>$\rho_{bs}$</td>
<td>0.178 ± 0.054</td>
<td>0.215± 0.052</td>
</tr>
</tbody>
</table>

Table 6: Variance-Covariance matrix of the estimates.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Systolic</th>
<th>Diastolic</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_b$</td>
<td>0.146 ± 0.073</td>
<td>0.163 ± 0.073</td>
</tr>
<tr>
<td>$\rho_s$</td>
<td>0.32 ± 0.069</td>
<td>0.248 ± 0.070</td>
</tr>
<tr>
<td>$\rho_{bs}$</td>
<td>0.178 ± 0.054</td>
<td>0.215± 0.052</td>
</tr>
</tbody>
</table>

Table 7: Estimates of Correlation parameters for original data.
the efficiency of the maximum likelihood estimates of the sibling correlations, which would result in a substantial loss of power of detecting departure from the null hypotheses of interest. This issue has been a subject of discussion by many authors [25,26]. Loss of power and sensitivity to the choice of the cut-off point are the price to pay due to discretization. However, selecting a cut-off point is not a matter of concern to statisticians but is based on clinical expertise. For example, components of what is known as metabolic syndrome (obesity, triglyceride, high density lipoprotein, blood pressures level, and blood glucose levels) are all measured on the continuous scale. However, communicating the clinical diagnoses of the components of the syndrome are based on the cutoff points recommended by the WHO, or the International Diabetes Federation. In this paper we used the WHO definition of hypertension 130/80 when we dichotomized the blood pressures data.

Discussion

Estimation of measures of family resemblance is considered the first step prior to investigating whether the variation in the distribution of the trait of interest is may be attributed to genetic factors. Similarly, detection of gender differences may be important to identify sex-linked traits. Establishing a statistical significance may provide the quantitative bases to study the distribution of the traits at the molecular level. The major contribution of this paper is the application of likelihood methods to a constructed bivariate beta-binomial distribution. This has allowed us to establish a score test for the goodness of fit of the model. A second finding is that testing for gender differences in the sib-sib correlations can be established in a relatively simple way, e.g. without computing the more complicated likelihood ratio test. Our limited scale computations showed that we need to sample a large number of families to retain reasonable power for the test statistic. Moreover, we showed that the power of the test of significance for the interclass correlation is quite insensitive to variations in the sub-cluster sizes. The implication is that as long as we have a sufficient number of families in the sample, the actual sib-ship sizes become less important. The model (7) is quite flexible. For example it easily allows for inclusion of covariates measured at the sub-cluster level. This can be done by employing a suitable transformation on the response probabilities similar to the case of a non-linear mixed model for binary responses. One important assumption of the present model is that it assumes that the correlation parameters are constant in the sampled population. This assumption may not be tenable in cases where some siblings are reared together and some reared apart. It should also be noted that there is a large number of statistical models used to fit clustered binary data, to name but a few, the Generalized Estimating Equations (GEE) which is a semi parametric approach, and the General Linear Mixed Models (GLIMMIX). These models are geared towards estimation of the correlation coefficients, treating the correlation structure as nuisance. The application of the GEE can be problematic for the analysis of family data. In fact Crowder [27] demonstrated that the parameters involved in working correlation matrix are subject to “uncertainty of definition which can lead to a breakdown of the asymptotic properties of the estimators”. On the other hand, the GLIMMIX does not readily produce estimates of correlations at each level of hierarchy. More seriously, there are some concerns regarding the approximation of the variance covariance matrix of the estimated parameters. In genetic epidemiology, clustering of traits is usually measured by a set of familial correlations, and such correlations become the population target parameters of interest. The model developed in this paper is constructed to address these issues. Finally, it should be noted however, that further research is needed to investigate the asymptotic properties of the test statistics that we developed when the sub-cluster sizes are much larger than the number of clusters, a problem of common occurrence in community-based studies.

Acknowledgement

The authors are thankful to the constructive comments made by two anonymous reviewers.

References


26. http://biostat.mc.vanderbilt.edu/twiki/bin/view/Main/CatContinuous

APPENDIX: Elements of the variance-covariance matrix of the MLE’s of the parameters of the bivariate-beta binomial distribution.

The following are the negative of the second partial derivatives of the log-likelihood function with respect to the five parameters:

\[
\begin{align*}
  i_{\mu_b}^2 &= \sum_{i=1}^{k} \sum_{j=0} \left( (\mu_b + j \theta_b)^{-2} + \sum_{j=0}^{b_i^* - y_i} (1 - \mu_b + j \theta_b)^{-2} \right) \\
  &+ \sum_{i=1}^{k} H^2(y_i, x_i) b_i^2 w_i^2 (x_i - s_i \mu_s)^2 \\
  i_{\mu_b \theta_b} &= -\sum_{i=1}^{k} \sum_{j=0} \left( (\mu_b + j \theta_b)^{-2} + \sum_{j=0}^{b_i^* - y_i} j (1 - \mu_b + j \theta_b)^{-2} \right) \\
  i_{\mu_b \theta_s} &= \sum_{i=1}^{k} \left[ H^2(\rho_{12}) b_i (x_i - s_i \mu_s) (\sigma_b \sigma_s)^{-1} \right] \\
  i_{\mu_s \theta_b} &= \sum_{i=1}^{k} \left[ H^2(\rho_{12}) b_i (x_i - s_i \mu_s) (\sigma_b \sigma_s)^{-1} \right] \\
  i_{\mu_s \theta_s} &= \sum_{i=1}^{k} \left[ H^2(\rho_{12}) b_i (x_i - s_i \mu_s) (\sigma_b \sigma_s)^{-1} \right] \\
  i_{\rho_{12}^2} &= \sum_{i=1}^{k} \left[ H^2(\rho_{12}) b_i (x_i - s_i \mu_s) (\sigma_b \sigma_s)^{-1} \right] \\
  i_{\theta_b \theta_s} &= \sum_{i=1}^{k} \left[ H^2(\rho_{12}) b_i (x_i - s_i \mu_s) (\sigma_b \sigma_s)^{-1} \right] \\
  \frac{\partial^2 l}{\partial \mu_b \partial \theta_b} &= \frac{\partial^2 l}{\partial \mu_b \partial \theta_s} = \frac{\partial^2 l}{\partial \theta_b \partial \theta_s} = \frac{\partial^2 l}{\partial \theta_b \partial \rho_{12}} = \frac{\partial^2 l}{\partial \theta_s \partial \rho_{12}} = 0
\end{align*}
\]

The other second partial derivations are obtained by symmetry, on replacing \((b, y)\) with \((s, x)\).

Elements of the information matrix are:

\[
\begin{align*}
  i_{\mu_b}^2 &= \sum_{i=1}^{k} \sum_{j=0} \left( (\mu_b + j \theta_b)^{-2} + \sum_{j=0}^{b_i^* - y_i} (1 - \mu_b + j \theta_b)^{-2} \right) \\
  &+ \sum_{i=1}^{k} H^2(y_i, x_i) b_i^2 w_i^2 (x_i - s_i \mu_s)^2 \\
  i_{\mu_b \theta_b} &= -\sum_{i=1}^{k} \sum_{j=0} \left( (\mu_b + j \theta_b)^{-2} + \sum_{j=0}^{b_i^* - y_i} j (1 - \mu_b + j \theta_b)^{-2} \right) \\
  i_{\mu_b \theta_s} &= \sum_{i=1}^{k} \left[ H^2(\rho_{12}) b_i (x_i - s_i \mu_s) (\sigma_b \sigma_s)^{-1} \right] \\
  i_{\mu_s \theta_b} &= \sum_{i=1}^{k} \left[ H^2(\rho_{12}) b_i (x_i - s_i \mu_s) (\sigma_b \sigma_s)^{-1} \right] \\
  i_{\mu_s \theta_s} &= \sum_{i=1}^{k} \left[ H^2(\rho_{12}) b_i (x_i - s_i \mu_s) (\sigma_b \sigma_s)^{-1} \right] \\
  i_{\rho_{12}^2} &= \sum_{i=1}^{k} \left[ H^2(\rho_{12}) b_i (x_i - s_i \mu_s) (\sigma_b \sigma_s)^{-1} \right] \\
  i_{\theta_b \theta_s} &= \sum_{i=1}^{k} \left[ H^2(\rho_{12}) b_i (x_i - s_i \mu_s) (\sigma_b \sigma_s)^{-1} \right] \\
  \frac{\partial^2 l}{\partial \mu_b \partial \theta_b} &= \frac{\partial^2 l}{\partial \mu_b \partial \theta_s} = \frac{\partial^2 l}{\partial \theta_b \partial \theta_s} = \frac{\partial^2 l}{\partial \theta_b \partial \rho_{12}} = \frac{\partial^2 l}{\partial \theta_s \partial \rho_{12}} = 0
\end{align*}
\]
\[
I = \begin{bmatrix}
i_{\mu_b}^2 & i_{\mu_b} & 0 & i_{\mu_b\rho_{12}} \\
i_{\mu_b} & i_{\mu_b} & 0 & i_{\mu_b\rho_{12}} \\
i_{\mu_b}^2 & 0 & i_{\mu_b\rho_{12}} & 0 \\
i_{\mu_b}^2 & 0 & 0 & i_{\mu_b\rho_{12}} \\
i_{\theta_b}^2 & 0 & 0 & i_{\rho_{12}}^2
\end{bmatrix}
\]

When the matrix is partitioned:

\[
I_{11} = \begin{bmatrix}
i_{\mu_b}^2 & i_{\mu_{bs}} \\
i_{\mu_{sb}} & i_{\mu_b}^2
\end{bmatrix}
\]

\[
I_{12} = \begin{bmatrix}
i_{\mu_b} & 0 & i_{\mu_b\rho_{12}} \\
p_{\mu_s} & 0 & i_{\mu_b\rho_{12}}
\end{bmatrix} = I_{21}'
\]

\[
I_{22} = \begin{bmatrix}
i_{\theta_b}^2 & 0 & 0 \\
i_{\theta_s}^2 & 0 & 0 \\
i_{\rho_{12}}^2 & 0 & 0
\end{bmatrix}
\]

The variance covariance matrix is given in a partitioned form.

\[
I_{11}^{-1} = \Delta^{-1} \begin{bmatrix}
i_{\mu_b}^2 & -i_{\mu_{bs}} \\
-i_{\mu_{sb}} & i_{\mu_b}^2
\end{bmatrix}
\]

Where \(\Delta = i_{\mu_b}^2 i_{\mu_b} - (i_{\mu_{sb}})^2\).

\[
I_{22}^{-1} = \begin{bmatrix}
\lambda_1 & 0 & 0 \\
0 & \lambda_2 & 0 \\
0 & 0 & \lambda_3
\end{bmatrix}
\]

Where

\[
\lambda_1^{-1} = i_{\theta_b}^2, \quad \lambda_2^{-1} = i_{\theta_s}^2, \quad \lambda_3^{-1} = i_{\rho_{12}}^2
\]

\[
\Sigma_{11} = \begin{bmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{bmatrix}
\]

Where

\[
a_{11} = \Delta_{11}^{-1} (i_{\mu_b}^2 - \lambda_2 i_{\mu_s} - \lambda_3 i_{\mu_b\rho_{12}})
\]

\[
a_{12} = a_{21} = \Delta_{11}^{-1} (\lambda_3 i_{\mu_s \rho_{12}} - i_{\mu_{bs}})
\]

\[
a_{22} = \Delta_{11}^{-1} (i_{\mu_b}^2 - \lambda_1 i_{\mu_b \theta_b} - \lambda_3 i_{\mu_b \rho_{12}})
\]

and,

\[
\Delta_{11} = (i_{\mu_b}^2 - \lambda_2 i_{\mu_s} - \lambda_3 i_{\mu_b \rho_{12}}) (i_{\mu_b}^2 - \lambda_1 i_{\mu_b \theta_b} - \lambda_3 i_{\mu_b \rho_{12}}) - (\lambda_3 i_{\mu_b \rho_{12}} i_{\mu_s \rho_{12}} - i_{\mu_{bs}})^2
\]
Note that, since \( \rho_b = \frac{\theta_b}{1 - \theta_b} \), the variance of \( \hat{\rho} \) is obtained by the delta method as

\[
\text{var}(\hat{\rho}_b) = (1 - \hat{\theta}_b)^{-2} \text{var}(\hat{\theta}_b)
\]

where \( \hat{\theta} \) is the MLE of \( \theta \), and \( \text{var}(\hat{\theta}) \) is the asymptotic variance of \( \hat{\theta} \) obtained from the inverted observation matrix.