

A Bayesian Estimator of the Intraclass Correlation Coefficient from Correlated Binary Responses

Marwa Ahmed¹, Mohamed Shoukri^{2,3}

¹*Cairo University*, ²*King Faisal Specialist Hospital and Research Center and*
³*University of Western Ontario*

Abstract: Clustered binary samples arise often in biomedical investigations. An important feature of such samples is that the binary responses within clusters tend to be correlated. The Beta-Binomial model is commonly applied to account for the intra-cluster correlation – the correlation between responses within the clusters – among dichotomous outcomes in cluster sampling. The intraclass correlation coefficient (ICC) quantifies this correlation or level of similarity. In this paper, we propose Bayesian point and interval estimators for the ICC under the Beta-Binomial model. Using Laplace's method, the asymptotic posterior distribution of the ICC is approximated by a normal distribution. The posterior mean of this normal density is used as a central point estimator for the ICC, and 95% credible sets are calculated. A Monte Carlo simulation is used to evaluate the coverage probability and average length of the credible set of the proposed interval estimator. The simulations indicate that for the situation when the number of clusters is above 40, the underlying mean response probability falls in the range of [0.3;0.7], and the underlying ICC values are ≤ 0.4 , the proposed interval estimator performs quite well and attains the correct coverage level. Even for number of clusters as small as 20, the proposed interval estimator may still be useful in the case of small ICC (≤ 0.2).

Key words: Beta-Binomial model, correlated binary responses, interval estimation, intraclass correlation coefficient, Laplace's method, posterior point estimation.

1. Introduction

The intraclass correlation coefficient (ICC) ρ has wide applications in biology, epidemiology and medical research. In family studies, it is used to measure the degree of intra-family resemblance with respect to characteristics such as blood pressure, weight and height. Moreover, it is used as means of investigating the

heritability of certain traits, whether such traits are continuous or dichotomous. An extensive literature, summarized in Shoukri and Ward (1985) and Donner (1986) exists for the statistical analysis of ICC for continuous response variables using the frequentist approach. Recently, Bayesian techniques have been developed; for example Spiegelhalter (2001). For binary traits statistical analysis of the ICC were extensively discussed by Mak (1988), Donner, Klar and Eliasziw (1995), and Gao, Klar and Donner (1997). Less developed is Bayesian statistical analysis of ICC for binary outcomes, which are of considerable practical importance in many toxicological and psychological studies. Within the Bayesian paradigm, the application of multivariate normal theory of continuous outcomes to binary variables is not strictly valid, and appropriate modeling strategy is needed. Turner, Omar and Thompson (2006) obtain the posterior distribution of ρ under the hierarchical logistic model and the beta binomial model using MCMC (Markov Chain Monte Carlo) methods. The posterior median is used as a point estimator and 95% interval estimates are calculated.

In this paper we develop a Bayesian estimator of the ICC under the Beta-Binomial model of correlated binary responses. In particular, we derive the approximate/asymptotic posterior distribution of the ICC, and use it to obtain the point and interval estimators. The remainder of the paper is organized as follows: Section 2 describes the underlying Beta-Binomial model. The proposed Bayesian estimator and its approximate posterior distribution are presented in Section 3. To study the coverage probability and average length of the credible set of the proposed interval estimator, a simulation study is given in Section 4. This is followed by an example in Section 5, and discussion in Section 6.

2. The Beta-Binomial Model

Consider a random sample of k clusters each of size n_i ($i = 1, 2, \dots, k$). We assume further that all subunits within the i th cluster, Y_{ij} ($j = 1, 2, \dots, n_i$), are binary taking one of two possible values of Y_{ij} , success or failure (coded as one and zero, respectively). The Y_{ij} 's are conditionally independent with probability $P(Y_{ij} = 1|p_i) = p_i$, where $0 < p_i < 1$. Let $Y_i = \sum_{j=1}^{n_i} Y_{ij}$ denote the total number of successes in the i th cluster. Thus, the conditional distribution of $Y_i|p_i$ is Binomial(n_i, p_i). In addition, it is assumed that the p_i 's are identically independently distributed as Beta(a, b) with mean $\pi = a/(a + b)$ and variance $v = ab/(a + b)^2(a + b + 1) = \pi(1 - \pi)/(a + b + 1)$. The unconditional distribution of Y_i is well known as the Beta-Binomial(n_i, a, b) with the following density

$$P(Y_i = y_i) = \binom{n_i}{y_i} \frac{B(a + y_i, n_i + b - y_i)}{B(a, b)},$$

$$y_i = 0, 1, \dots, n_i.$$

The correlation between any pair of responses within the same cluster $\text{corr}(Y_{ij}, Y_{il}), j \neq l$, is given by the intracluster correlation coefficient $\rho = 1/(1 + a + b)$ (Moore, 1987). Prentice (1986) showed that the Beta-Binomial distribution is valid for ρ values falling in the range

$$\min\left(\frac{\pi}{n_{\max} - \pi - 1}, \frac{1 - \pi}{n_{\max} + \pi - 2}\right) < \rho < 1,$$

where n_{\max} is the size of the largest cluster.

A positive correlation means that responses within a cluster are more alike. An example for positive correlation is in toxicological experiments designed to study teratogenic effects of chemical compounds on animals. Fetuses within a litter tend to respond more similarly than fetuses from different litters, a phenomenon known as *litter effect*. Less commonly, the observations within a cluster are negatively correlated. This occurs for example in family studies where children may be competing for maternal care. The focus of this paper is on applications where the ICC is positive.

Following the logic of Moore (1987), the distribution of the p_i 's may be reparameterized in terms of π and ρ , i.e. $p_i \sim \text{Beta}(\pi, \rho)$.

3. A Bayesian Estimator of the ICC

The Beta-Binomial model in (2.1) is a popular example for hierarchical models. In the first stage, $Y_i|p_i$ are independent Binomial(n_i, p_i). In the second stage, $p_i|\pi, \rho$ are i.i.d. Beta(π, ρ). For a full Bayesian analysis, the hyperparameters, π and ρ , are in turn drawn from a hyperprior distribution. In this paper, it is assumed that they are distributed as Uniform(0, 1), and that they are a priori-independent of each other. The joint posterior distribution is obtained via Bayes theorem (Carlin and Louis, 2000, p. 19):

$$\begin{aligned} G(\pi, \rho|y) &= \int_p f(p, \pi, \rho|y) dp \\ &= \int_p \frac{f(p, \pi, \rho, y)}{f(y)} dp \\ &= \int_p \frac{f(y|p)f(p|\pi, \rho)f(\pi, \rho)}{f(y)} dp \end{aligned}$$

The marginal posterior distribution of ρ is then derived via integrating p out, that is

$$g(\rho|y) = \int_\pi \int_p \frac{f(y|p)f(p|\pi, \rho)f(\pi, \rho)}{f(y)} dp d\pi.$$

The evaluation of the previous integrals is intractable and computationally difficult. An alternative is to use asymptotic techniques to obtain approximations of the posterior density. In this paper, when the number of clusters k gets large and the cluster size n remains small, $g(\pi, \rho|y)$ is approximated using Laplace's method. Following the logic of Kass and Steffey (1989), under mild regularity conditions, $g(\pi, \rho|y)$ is asymptotically Bivariate Normal with mean and variance given by the posterior mode $\tilde{\lambda}$ and the inverse of the negative Hessian of the log posterior evaluated at the mode

$$\tilde{\Sigma} = (-D^2 \log(Lf(\pi, \rho))^{-1}|_{\tilde{\lambda}},$$

where $L = \prod_{i=1}^k \int f(y_i|p_i)f(p_i|\pi, \rho)dp_i$, and $f(\pi, \rho)$ is the prior introduced on $\lambda = (\pi, \rho)$. When the Uniform prior on λ is adopted, which is the case here, $\tilde{\lambda}$ and $\tilde{\Sigma}$ can be replaced by the MLE's, $\hat{\pi}$ and $\hat{\rho}$, and the inverse of the observed information matrix. That is asymptotically, $g(\pi, \rho|y) \sim \text{Bivariate Normal}(\hat{\lambda}, \hat{\Sigma})$, where, $\hat{\lambda} = (\hat{\pi}, \hat{\rho})$ and $\hat{\Sigma} = (\hat{I}_{obs})^{-1}$, the inverse of the observed information matrix evaluated at $\hat{\pi}$ and $\hat{\rho}$. The elements of the observed information matrix are given by

$$\hat{I}_{ij} = \left[-\frac{\partial^2}{\partial \lambda_i \partial \lambda_j} \log(L) \right]_{\lambda=\hat{\lambda}}.$$

For the case of the Beta-Binomial model,

$$L = \prod_{i=1}^k \binom{n_i}{y_i} \frac{B(a + y_i, n_i + b - y_i)}{B(a, b)},$$

where $a = \pi(1 - \rho)/\rho$, and $b = (1 - \pi)(1 - \rho)/\rho$.

By properties of the Bivariate Normal distribution, the marginal posterior distribution $g(\rho|y)$ is asymptotically Normal $(\hat{\rho}, \sigma_*^2)$, where σ_*^2 is the corresponding diagonal element of $\hat{\Sigma}$. Griffiths (1973) describes an approach for ML estimation of the Beta-Binomial distribution with $n = n_i$ ($i = 1, 2, \dots, k$). The ML estimators of π and θ , where $\theta = \rho/(1 - \rho)$, are the solution of the following two equations (Griffiths, 1973):

$$\begin{aligned} \frac{\partial \ln L}{\partial \pi} &= \sum_{i=0}^{n-1} \left(\frac{S_n - S_i}{\pi + i\theta} - \frac{S_{n-1-i}}{1 - \pi + i\theta} \right) = 0 \\ \frac{\ln L}{\partial \theta} &= \sum_{i=1}^{n-1} i \left(\frac{S_n - S_i}{\pi + i\theta} + \frac{S_{n-i-i}}{1 - \pi + i\theta} - \frac{S_n}{1 + i\theta} \right) = 0, \end{aligned}$$

where, $S_i = \sum_{y=0}^I f_y$, ($i = 0, 1, 2, \dots, n$) and f_y is the observed frequency ($y = 0, 1, 2, \dots, n$).

Equations (3.1) may be solved iteratively using numerical algorithms such as Newton-Raphson method (Press *et al.*, 2007, chap. 9) or the Jenkins-Traub algorithm (Jenkins and Traub, 1970). The second derivatives can be easily obtained to find the information matrix of π and θ , $\hat{I}(\pi, \theta)$, and are given by:

$$\begin{aligned}\frac{\partial^2 \ln L}{\partial \pi^2} &= \sum_{i=0}^{n-1} \left(\frac{S_i - S_n}{(\pi + i\theta)^2} - \frac{S_{n-1-i}}{(1 - \pi + i\theta)^2} \right) \\ \frac{\partial^2 \ln L}{\partial \pi \partial \theta} &= \sum_{i=0}^{n-1} i \left(\frac{S_i - S_n}{(\pi + i\theta)^2} + \frac{S_{n-1-i}}{(1 - \pi + i\theta)^2} \right) \\ \frac{\partial^2 \ln L}{\partial \theta^2} &= \sum_{i=0}^{n-1} i^2 \left(\frac{S_i - S_n}{(\pi + i\theta)^2} - \frac{S_{n-1-i}}{(1 - \pi + i\theta)^2} + \frac{S_n}{(1 + i\theta)^2} \right)\end{aligned}$$

$Var(\hat{\theta})$ is then the corresponding diagonal element of $[\hat{I}(\pi, \theta)]^{-1}$. The MLE of ρ is given by

$\hat{\rho} = \hat{\theta}/(1 + \hat{\theta})$. By the Delta method (Kendall and Stuart, 1986, p. 324),

$$var(\hat{\rho}) = \frac{1}{(1 + \hat{\theta})^4} var(\hat{\theta}).$$

Thus, the asymptotic posterior distribution of ρ is

$$\rho|y \sim N \left(\frac{\hat{\theta}}{1 + \hat{\theta}}, \frac{1}{(1 + \hat{\theta})^4} var(\hat{\theta}) \right).$$

The posterior mean, mode or median can be used as a point estimator of ρ . In this paper, a squared error loss function is adopted. Therefore, the Bayes rule (the point estimator that minimizes the posterior risk) is the posterior mean. A $100(1 - \alpha)\%$ credible set (Bayesian confidence interval) for ρ is given by the 2.5 and 97.5 per cent quantiles of this posterior distribution.

4. Simulation Study

In order to evaluate the accuracy of the Bayesian credible set, we need to conduct a large scale Monte Carlo simulation. Since there are many parameters involved (n, k, π, ρ), a theoretical evaluation is difficult to conduct. Therefore, the simulation approach is adopted, to study the coverage probability and average length of the credible set of the proposed estimator. For the case of cluster size $n = 2$, the situations were considered where the number of clusters k equals 20, 40, 100 and 200, the underlying mean response probability p equals 0.1, 0.2, 0.3, 0.4, 0.6, and 0.7 and the ICC, ρ , takes one of four values 0.1, 0.2, 0.4, 0.6. A

Table 1: Performance of the proposed credible set based on Monte Carlo simulation: the estimated coverage probability for the ICC at nominal level 0.95, average length of the credible set (in parentheses), and the proportion of invalid samples [in brackets]

π		0.1				0.2				0.3			
k	ρ	0.1	0.2	0.4	0.6	0.1	0.2	0.4	0.6	0.1	0.2	0.4	0.6
20		0.93	0.94	0.97	0.98	0.94	0.94	0.90	0.87	0.95	0.93	0.86	0.73
		(1.07)	(1.05)	(1.01)	(0.91)	(0.94)	(0.94)	(0.89)	(0.75)	(0.87)	(0.87)	(0.80)	(0.60)
		[0.70]	[0.57]	[0.43]	[0.42]	[0.34]	[0.24]	[0.12]	[0.12]	[0.11]	[0.06]	[0.03]	[0.03]
40		0.97	0.95	0.90	0.82	0.96	0.94	0.86	0.73	0.92	0.92	0.90	0.75
		(0.84)	(0.84)	(0.81)	(0.70)	(0.66)	(0.67)	(0.65)	(0.51)	(0.62)	(0.63)	(0.59)	(0.45)
		[0.46]	[0.35]	[0.18]	[0.12]	[0.10]	[0.04]	[0.02]	[0.01]	[0.01]	[0.00]	[0.00]	[0.00]
100		0.97	0.94	0.88	0.74	0.92	0.93	0.93	0.79	0.94	0.94	0.94	0.71
		(0.49)	(0.53)	(0.54)	(0.43)	(0.42)	(0.44)	(0.44)	(0.34)	(0.40)	(0.40)	(0.39)	(0.26)
		[0.15]	[0.06]	[0.01]	[0.00]	[0.01]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
200		0.91	0.92	0.92	0.74	0.94	0.94	0.94	0.78	0.96	0.94	0.95	0.75
		(0.34)	(0.38)	(0.40)	(0.30)	(0.30)	(0.31)	(0.31)	(0.23)	(0.28)	(0.29)	(0.27)	(0.19)
		[0.02]	[0.01]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]

π		0.4				0.6				0.7			
k	ρ	0.1	0.2	0.4	0.6	0.1	0.2	0.4	0.6	0.1	0.2	0.4	0.6
20		0.92	0.93	0.87	0.69	0.92	0.98	0.92	0.78	0.94	0.92	0.92	0.84
		(0.85)	(0.84)	(0.75)	(0.55)	(0.86)	(0.85)	(0.61)	(0.61)	(0.87)	(0.87)	(0.83)	(0.69)
		[0.02]	[0.01]	[0.00]	[0.00]	[0.02]	[0.01]	[0.00]	[0.00]	[0.10]	[0.05]	[0.02]	[0.03]
40		0.95	0.93	0.92	0.74	0.94	0.93	0.93	0.85	0.93	0.92	0.92	0.90
		(0.61)	(0.61)	(0.56)	(0.41)	(0.61)	(0.61)	(0.57)	(0.46)	(0.62)	(0.63)	(0.61)	(0.52)
		[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.01]	[0.00]	[0.00]	[0.00]
100		0.95	0.94	0.95	0.77	0.96	0.94	0.95	0.85	0.94	0.93	0.94	0.94
		(0.39)	(0.39)	(0.36)	(0.26)	(0.39)	(0.39)	(0.37)	(0.28)	(0.40)	(0.40)	(0.39)	(0.34)
		[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
200		0.94	0.95	0.94	0.81	0.94	0.95	0.95	0.90	0.95	0.95	0.94	0.95
		(0.28)	(0.27)	(0.26)	(0.19)	(0.28)	(0.27)	(0.26)	(0.21)	(0.28)	(0.29)	(0.27)	(0.24)
		[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]

fully factorial combination of these three factors was used, giving a total of 96 combinations. For each combination, 2000 valid samples were generated. In situations where the parameter π was near the edge, i.e., outside the range [0.3; 0.7], the percentage of samples that gave invalid solutions was quite high, sometimes reaching 70%. Whenever a generated sample led to an invalid point estimate of π or ρ (i.e., when either parameter falls out of range), the sample was discarded and replaced by a new one until a total of 2000 valid samples were obtained. As for the limits of the interval estimates, if they exceeded one or fell below the minimum possible value given in equation (2.2), then they were replaced by the appropriate extreme value. For each combination of the simulation factors, the coverage probability and the average length of the credible set were calculated. A SAS program was used to generate the data from the Beta-Binomial distribution¹. The random numbers were generated in two steps. First, the probability

¹SAS Institute Inc., (2004). Version 8.2. Cary, NC, USA.

of success, p_i , were generated from the Beta distribution. In the second step, the y_i 's were generated from a Binomial(2, p_i). The NSolve function for numerically solving sets of simultaneous equations in Mathematica software was used to find the roots of the ML equations².

Table 1 shows for 95% nominal credible sets, the estimated coverage probability, the average length of the resulting credible sets, and the percentage of invalid samples. First, note that the estimated coverage probability is above or less than the nominal level by no more than 3% in 67 out of the 96 cases, i.e., 69.79% of the time. This ratio reaches 97.22% for the case where, simultaneously, the number of clusters increases ($m \geq 40$), ICC values are ≤ 0.4 , and π is in the interval [0.3; 0.7]. Even for number of clusters as small as 20, the proposed interval estimator may still be useful in the case of small ICC (≤ 0.2). On the other side, the worst situation occurs when $\rho = 0.6$. Here, the coverage probability does not attain its nominal level.

Examining the average length of the confidence interval, it can be seen that, for all other conditions fixed, the average confidence interval length tends to decrease as either the number of clusters increases, or when π increases.

For $\pi \leq 0.4$, all invalid samples in this simulation study are the result of the case when $f_2 = 0$. The number of invalid samples increases as π decreases. Recall the following relationship between $Prb[y_{i1} = 1, y_{i2} = 1]$ and p_i , and consequently π , (Shoukri and Pause, 1998, p. 66):

$$Prb[y_{i1} = 1, y_{i2} = 1] = p_i^2 + \rho p_i(1 - p_i).$$

A small π/p_i will result in a small $Prb[y_{i1} = 1, y_{i2} = 1]$, and thus the chances of obtaining $f_2 = 0$ will increase. For any fixed π , the percentage of invalid samples decreases as either the value of ρ increases or the sampled number of clusters, m , increases. This can be explained also by equation (4.1), where it is clear that there is a proportional relationship between ρ and $Prb[y_{i1} = 1, y_{i2} = 1]$. It also makes sense to say that as more clusters are sampled, the chances of obtaining two positive responses within any cluster will increase. Therefore, as m increases, the probability of obtaining $f_2 = 0$ will decrease. By a similar logic, for $\pi \geq 0.6$, the invalid samples are the result of the case when $f_1 = 0$.

Based on this simulation study, it is recommended to use the proposed estimator when the sampled number of clusters is at least 40, when it is believed that the ICC values are ≤ 0.4 , and the probability of positive response π is in the range [0.3;0.7]. A limitation of this simulation study is that it is restricted to the case of cluster size 2.

²Wolfram Research Inc. (2005). Version 5.2. Champaign, IL, USA.

5. Example

To illustrate the application of the proposed estimator, an example is considered. In ophthalmologic studies, the eye is the unit for statistical analysis rather than the individual. Typically, an individual contributes two eyes worth of information whose values might be correlated. Each person is a cluster of size two. To obtain valid inference, the statistical analysis must account for the effect of the intracluster correlation. Berson, Rosner and Simonoff (1980) and Rosner (1982) describe an ophthalmologic study conducted at the Massachusetts Eye and Ear Infirmary from 1970 to 1979 of an outpatient population of 216 persons aged 20-39 with retinitis pigmentosa (RP). This population was classified into the genetic types of autosomal recessive RP (AR), autosomal dominant RP (DOM), sex-linked RP (SL), and isolate RP (ISO) for a study of differences among these four groups on certain measurements. The details of the study design and the procedures for genetic classification are given by Berson, Rosner and Simonoff (1980). In the current paper, the binary outcome of interest is the best corrected Snellen visual acuity (VA). Any eye is considered affected if VA is 20/50 or worse, and normal if VA is 20/40 or better. The number of affected eyes for the given sample is presented in Table 2. Hence, $S_0 = f_0 = 92$, $S_1 = 129$, and $S_2 = 216$.

Table 2: Distribution of the number of affected eyes

f_0	f_1	f_2	total
92	37	87	216

The goal is to estimate the ICC for this outpatient population using the Bayesian estimator described in section 3. The MLE estimators of π and θ are the solution of the following two equations:

$$\begin{aligned} \frac{S_2 - S_0}{\pi} - \frac{S_1}{1 - \pi} + \frac{S_2 - S_1}{1 - \pi + \theta} &= 0 \\ \frac{S_2 - S_1}{\pi + \theta} + \frac{S_0}{1 - \pi + \theta} - \frac{S_2}{1 + \theta} &= 0 \end{aligned}$$

Equations (5.1) can be easily solved using a rootfinding routine in mathematical or statistical computer software such as Mathematica. Using the Nsolve function of Mathematica, the solution of equations (5.1) is found to be $\hat{\pi} = 0.488426$ and $\hat{\theta} = 1.917357$. The covariance matrix is given by:

$$Var(\hat{\pi}, \hat{\theta}) = \begin{bmatrix} 0.00095853 & 0.00010274 \\ 0.00010274 & 0.19060900 \end{bmatrix}.$$

Therefore, $var(\hat{\theta}) = 0.190609$, and the MLE of ρ is $\hat{\rho} = \hat{\theta}/(1 + \hat{\theta}) = 0.657224$. By the Delta method, $var(\hat{\rho}) = 0.00263139$. The asymptotic posterior distribution of ρ/y is $N(0.657224, 0.00263139)$. The proposed point estimate of the ICC is given by the posterior mean, which is equal 0.657224. A 95% credible set of the ICC is $[0.55668148, 0.7577665]$. Rosner (1982) analyzed the same data set using the classical approach, and similar results were obtained. The effective number of units within a cluster, denoted by e , is used in his analysis. Easy manipulation shows that the ICC is related to e through the the following function: $\rho = 2/e - 1$. The MLE of e obtained by Rosner is $\hat{e} = 1.207$. Therefore, the MLE of the ICC is $\rho_{Rosner} = 0.657001$.

6. Discussion

We have presented a Bayesian estimator for the ICC under the Beta-Binomial model. The approach is based on Laplace's method to approximate the posterior distribution and moments of the ICC given the data. The technique is asymptotic, that is for sufficiently large number of clusters. An advantage of adopting a Bayesian perspective is the possibility of incorporating prior beliefs on likely values of the ICC. The approach is flexible to accommodate informative and noninformative priors for π and ρ . In the current study a Uniform prior has been adopted. When a nonuniform prior of π and ρ is used, then the mean and variance of the Normal approximation are given by the posterior mode and the inverse of the negative Hessian of the log posterior evaluated at the mode (Kass and Steffey, 1989). As noted by Turner, Omar and Thompson (2006), ICC values are generally reported unaccompanied by confidence intervals, which makes them of limited value as in estimation of any other parameter. The current paper provides point and interval estimators for a single ICC. The proposed approach is simple, relatively easy to implement, and is not computer intensive, which makes it useful for practitioners from the biomedical fields. This is in contrast to already existing Bayesian approaches depending on simulations and MCMC methods. It is clear that the technique used here to find the MLE's of the Beta-Binomial distribution (Griffiths, 1973) assumes equal cluster sizes. When the clusters are not of equal sizes, two nonlinear equations need to be solved numerically to obtain the posterior estimators of ρ and π .

References

- Berson, E. L., Rosner, B., and Simonoff, E. (1980). An outpatient population of retinitis pigmentosa and their normal relatives: risk factors for genetic typing and detection derived from their ocular examinations. *American Journal of Ophthalmology* **89**, 763-775.

- Carlin, B. P., and Louis, T. A. (2000). *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman and Hall.
- Donner, A. (1986). A review of inference procedures for the intraclass correlation coefficient in the one-way random effects model. *International Statistical Review* **54**, 67-82.
- Donner, A. A., Klar, N., and Eliasziw, M. (1995). Statistical methodology for estimating twin similarity with respect to a dichotomous trait. *Genetic Epidemiology* **12**, 267-277.
- Gao, X. J., Klar, N. and Donner, A. (1997). Comparison of methods for analyzing binary data arising from two-sample twin studies. *Genetic Epidemiology* **14**, 349-363.
- Griffiths, D. A. (1973). Maximum Likelihood estimation for the beta-binomial distribution and an application to the household distribution of the total number of cases of a disease. *Biometrics* **29**, 637-648.
- Jenkins, M. A. and Traub, J. F. (1970). A three-stage variables-shift iteration for polynomial zeros and its relation to generalized Rayleigh iteration. *Numerische Mathematik* **14**, 252-263.
- Kass, R. E. and Steffey, D. (1989). Approximate Bayesian inference in conditionally independent hierarchical models (parametric empirical Bayes models). *Journal of the American Statistical Association* **84**, 717-726.
- Kendall, M. and Stuart A. (1986). *The Advanced Theory of Statistics, Volume 1: Distribution Theory*. Macmillan.
- Mak, T. K. (1988). Analysing intraclass correlation for dichotomous variables. *Applied Statistics* **37**, 344-352.
- Moore, D. F. (1987). Modelling the extraneous variance in the presence of extra-binomial variation. *Applied statistics* **36**, 8-14.
- Prentice, R. L. (1986). Binary regression using an extended beta-binomial distribution, with discussion of correlation induced by covariate measurement errors. *Journal of the American Statistical Association* **81**, 321-327.
- Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (2007). *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press.
- Rosner, B. (1982). Statistical methods in ophthalmology: an adjustment for the intraclass correlation between eyes. *Biometrics* **38**, 105-114.
- Shoukri, M. M., and Pause, C. A. (1998). *Statistical Methods for Health Sciences*, 2nd edition. CRC Press.
- Shoukri, M. M., and Ward, R. H. (1985). Estimation of intraclass correlation. *Communications in Statistics, Theory and Methods* **13**, 1239-1255.

Spiegelhalter, D. J. (2001). Bayesian methods for cluster randomized trials with continuous responses. *Statistics in Medicine* **20**, 435-52.

Turner, R. M., Omar, R. Z., and Thompson, S. G. (2006). Constructing intervals for the intraclass correlation coefficient using Bayesian modeling, and application in cluster randomized trials. *Statistics in Medicine* **25**, 1443-1456.

Received May 10, 2008; accepted July 14, 2008.

Marwa Ahmed
Department of Statistics
Cairo University
P.O.Box 197
Dokki, Cairo 12311, Egypt
althinker@yahoo.com

Mohamed Shoukri
Department of Biostatistics, Epidemiology and Scientific Computing
King Faisal Specialist Hospital and Research Center
MBC # 3 , P.O. Box 3354
Riyadh 11211, Saudi Arabia
shoukri@kfshrc.edu.sa

Also: Department of Epidemiology and Biostatistics
Schulich School of Medicine
University of Western Ontario
London, Ontario, Canada