A GENERALIZED LIKELIHOOD RATIO APPROACH FOR CLUSTER-CORRELATED DATA FROM HUMAN FERTILITY STUDIES

By HAIBO ZHOU
University of North Carolina, Chapel Hill, U.S.A.
MINGE XIE
Rogers University Piscataway, U.S.A.
DOUGLAS G. SIMPSON
University of Illinois, Champaign, U.S.A.
and
CLARICE R. WEINBERG
National Institute of Environmental Health Sciences, U.S.A.

SUMMARY. Data from fertility studies are naturally cluster-correlated due to the fact that menstrual cycles belonging to the same couple are biologically related. We discuss a generalized likelihood ratio approach for a fertility model for the probability of conception, where the mean structure is given as the product of two independent components, and each of the components can be modeled as a function of a generalized linear model. We use the proposed method to construct confidence intervals for parameters that are restricted on [0, 1]. The proposed approach provides range-preserving, asymmetric confidence intervals for the fertility parameters. Analysis of a data set from a North Carolina fertility study reveals some interesting findings that can not be observed with the traditional methods. The analysis suggests that prenatal exposure to maternal smoking has a deleterious effect on cycle viability while late menarche has a beneficial effect. The data do not support the existence of a negative effect of frequent ejaculation on sperm potency.

1. Introduction

The association between the likelihood of conception and both endogenous and exogenous factors is of interest to many people, including reproductive scientists, and couples who wish to begin or avoid a pregnancy. Estimating and identifying effects of these factors on the fecundability, the probability of conception in a given menstrual cycle for a couple who is not using any contraceptive method, is of particular public health interests. Hypothesis testing and model selection in the traditional fertility models rely on the Wald statistics and the Chi-square distributed

Key words and phrases. Fertility, generalized likelihood ratio, latent variables, quantile, weighted $\chi^2$. 
likelihood ratio statistics (based on assuming that observations are independent). This is not satisfactory since the correlation among menstrual cycles belonging to the same couple is not accounted for. The confidence intervals using the Wald statistics are not range preserving to reparameterization and the asymptotic distribution of the likelihood ratio statistic is not exactly chi-square, due to the cluster-correlated nature of the data. Valid inference procedures are important in understanding the biological processes involved in human fertility. Our research is motivated by the need to carry out inference related to fertility parameters.

We begin with some basic biological background for human reproduction. For women of reproductive age, a single ovum is normally produced and released around the middle of each menstrual cycle in a process called ovulation. For pregnancy to occur, the ovum has to be fertilized by sperm that reach the woman's reproductive tract. These sperm can come from intercourse on different days around the time of ovulation. If the ovum is not fertilized during its viable period, it will degenerate and a new cycle will begin. If the ovum is fertilized, however, the specific act of intercourse during the cycle that caused a conception is usually indeterminate. In addition to the timing and frequency of intercourse, other endogenous and exogenous factors may also affect the probability of conception. Many of these factors are unobservable, but manifest themselves in dependency among outcomes within multiple cycles from the same couple. Other factors, such as age and exposure to a potential toxicant, can be observed directly. Valid inference procedure and model selection criteria for human fertility models should account for the dependency among outcomes within the same couple.

We develop a generalized likelihood ratio (GLR) analysis to carry out the inference for fertility parameters. The proposed GLR method, a generalization of the standard likelihood ratio test under independent settings, is suitable for batch-correlated observations. The statistic has an asymptotic distribution that is a weighted sum of $\chi^2$. The rest of the paper is organized as follows. In Section 2 we present the statistical models for fertility studies. In Section 3 we describe the basic asymptotic theory of the GLR statistics under these models and discuss how to construct a hypothesis test using the GLR statistics. In Section 4, we analyze the data from the North Carolina fertility study (Wilcox et al., 1988) to carry out the inference on fertility parameters and compare a sequence of models. We give final remarks in Section 5.

2. Marginal Modeling of Fertility

The data from the fertility studies we will consider are cluster-correlated data, where outcomes from the same couple are correlated. To fix notation, for each couple $i$, where $i = 1, 2, \ldots, N$, there are binary outcomes $Y_{ij}$ for $j = 1, 2, \ldots, n_i$ cycles, where $Y_{ij}$ is 1 if there is a conception signaled by the rise in pregnancy hormone, hCG, for cycle $j$, and 0 otherwise. We remark that there could be multiple $Y_{ij} = 1$ for couple $i$ if $n_i > 1$ and a conception failed prior to clinical detection. The outcome of each such cycle depends on a susceptibility factor (cycle viability), and
the aggregate of a set of unobservable Bernoulli trials (the effects of intercourse on various days). These two independent components are assumed to contribute to the probability of conception in a multiplicative way. For a given cycle \( j \) in couple \( i \), we assume there are independent competing risks of fertilization from the days on which there was intercourse, any of which could cause conception in a viable cycle. A menstrual cycle is viable if the ovum is viable and other factors, such as receptivity of the uterus, are such that conception is theoretically possible.

To keep track of when intercourse occurs, let \( x_{ijk} \) be an indicator variable such that \( x_{ijk} = 1 \) if couple \( i \) had intercourse on day \( k \) of cycle \( j \), and \( x_{ijk} = 0 \) otherwise. Here \( k \) is indexed relative to the day of ovulation (day \( k = 0 \)). It is negative prior to ovulation and positive afterwards. Let \( g_1(\alpha' u_{ij}) \in [0, 1] \) be the probability that cycle \( j \) for couple \( i \) is viable, where \( u_{ij} = (u_{ij1}, \ldots, u_{ijk}) \) is a covariate vector and \( \alpha' = (\alpha_1, \ldots, \alpha_q) \) is the corresponding regression parameter vector. Let \( g_2(\beta' v_{ijk}) \in [0, 1] \) be the probability that intercourse on the \( k \)th day would have succeeded if the other competing risks had not been present (no intercourse on any other day) and the cycle had been viable, where \( v_{ijk} = (v_{ijk1}, \ldots, v_{ikp}) \) is the associated covariate vector and \( \beta' = (\beta_1, \ldots, \beta_p) \) is the corresponding regression parameter vector. Conception occurs in a given cycle if and only if the cycle is viable and at least one of the days with intercourse causes fertilization. The functions \( g_1^{-1} \) and \( g_2^{-1} \) are assumed probability link functions such as the logit or probit.

The observed data are of the form \( \{Y_{ij}, X_{ij}, U_{ij}, V_{ij}\} \), where \( \{X_{ij}, U_{ij}, V_{ij}\} \) is a matrix containing observations \( \{x_{ijk}, u_{ij}, v_{ijk}\} \) for couple \( i \) in cycle \( j \). A marginal type model for clustered data, building on earlier work in fertility modeling (e.g., Barrett and Marshall, 1969; Schwartz, MacDonald and Heuchel, 1980; Royston, 1982; and Weinberg, Gladen, and Wilcox, 1994), has been proposed (Zhou and Weinberg, 1996) that specifies the marginal probability of conception in a given menstrual cycle \( j \) as

\[
\mu_{ij} = \Pr(Y_{ij} = 1|X_{ij}, U_{ij}, V_{ij}) = g_1(\alpha' u_{ij}) \left[ 1 - \prod_k (1 - g_2(\beta' v_{ijk}))^{x_{ijk}} \right].
\]

We will refer to the cycle viability term \( g_1(\alpha' u_{ij}) \) as the \( A \) component and the remaining day-specific probability term in (1) as the \( P \) component. A Bayesian approach is proposed recently (Dunson and Zhou, 2000) that extends the random effects model approach to the problem (Zhou et al. 1996, Zhou and Weinberg, 1999).

The definition of marginal mean for this model is different from the usual generalized linear model formulation. Indeed, the marginal mean is modeled as a product of two components, \( A \) and \( P \), where each component can be viewed as a function of a generalized linear model formulation. An important sub-model of (1) sets \( g_2(\beta' v_{ijk}) = p_k \), for \( k = -K, \ldots, 0, \ldots, K \). This model is used to study the length of the fertile interval (Wilcox, Weinberg and Baird, 1995). By estimating the probability \( (p_k) \) of conception associated with each day and its confidence interval, we can estimate how many consecutive days have \( p_k > 0 \) and therefore estimate the beginning and end of the fertile window.

Denote \( \Theta' = (\alpha', \beta') \). \( \Theta \) can be estimated by maximizing the weighted pseudo-
likelihood function

\[ L(\Theta) = \prod_{i=1}^{N} \left\{ \prod_{j=1}^{n_i} L_{ij}(\Theta) \right\}^{w_i}, \]

where

\[ L_{ij}(\Theta) = \left[ g_1(\alpha'u_{ij}) \left( 1 - \prod_{k} (1 - g_2(\beta'x_{ijk}))^{x_{ijk}} \right) \right]^{y_{ij}} \]

\[ \cdot \left[ 1 - g_1(\alpha'u_{ij}) \left( 1 - \prod_{k} (1 - g_2(\beta'x_{ijk}))^{x_{ijk}} \right) \right]^{1-y_{ij}}, \]

and \( w_i \) is a weight constant used to improve the efficiency of the estimator. Then the maximum pseudo-likelihood estimator \( \hat{\Theta} \) can be obtained by solving the following unbiased estimating equation:

\[ U(\hat{\Theta}) = \sum_{i=1}^{N} \psi_i (\hat{\Theta}) = 0, \quad (2) \]

where

\[ \psi_i(\Theta) = w_i \sum_{j=1}^{n_i} \frac{\partial}{\partial \Theta} \log [L_{ij}(\Theta)]. \]

The robust sandwich estimator of the covariance matrix of \( \hat{\Theta} \) is then

\[ \overline{\text{cov}}(\hat{\Theta}) = M_0^{-1} M_1 M_0^{-1T}, \quad (3) \]

where \( M_0 = \sum_{i=1}^{N} \frac{\partial}{\partial \Theta} \psi_i (\hat{\Theta}) \) and \( M_1 = \sum_{i=1}^{N} \psi_i (\hat{\Theta}) \psi_i^T (\hat{\Theta}) \). Inference on fertility parameters in the early published literature has generally relied on a naive variance estimator, \( \overline{\text{cov}}(\hat{\Theta}) = M_0^{-1} \), which is only valid when the observations are independent. Clearly, viewing a cluster-correlated sample as an independent sample would result in assuming the observed data has more information about the underlying model than it is truly has. Following Zhou and Weinberg, we use E-M algorithm (Dempster, Laird, and Rubin, 1977) that is based on a set of latent variables to compute \( \hat{\Theta} \). Let the indicator variable \( Z_{ij} \) denote whether the particular cycle \( ij \) is viable. For conception cycles, \( Z_{ij} \) is known to be 1. For nonconception cycles, \( Z_{ij} \) is unobservable. Similarly, we let \( S_{ijk} \) be a binary indicator that corresponds to whether pregnancy would have occurred had the cycle been viable and had there been no day other than day \( k \) with intercourse in cycle \( j \) for couple \( i \). \( S_{ijk} \) is an unobservable variable except for days on which there was no intercourse \( (S_{ijk} = 0) \), or for cycles where conception occurred and there was only one day with intercourse \( (S_{ijk} = 1) \). We estimate \( \Theta \) from the following complete-data likelihood function:

\[ \prod_{i=1}^{N} \left\{ \prod_{j=1}^{n_i} g_1(\alpha'u_{ij})^{Z_{ij}} \left( 1 - g_1(\alpha'u_{ij}) \right)^{1-Z_{ij}} \right\}. \]
The conditional expected value of $S_{ijk}$ and $Z_{ij}$ given the observed data are given in Zhou and Weinberg (1996). The E-M algorithm in our analysis converges fairly quickly. We programmed our analysis using S-plus on a low-end Sun workstation and it takes about 5 to 10 minutes for the E-M algorithm to converge.

Earlier analyses on fertility data had relied on either the Wald statistics or a naïve Chi-square distribution where all the observations are assumed independent to assess the covariates effects and their confidence intervals. It is well known that the Wald test based on the standard error is not transformation invariant, and the confidence intervals for $p_k$ based on the Wald statistics are restricted to be symmetric. Furthermore, the naïve Chi-square criterion may not be valid when there is dependence among individuals within a cluster. We next discuss a generalized likelihood ratio statistic which can be used for inference about parameters in the fertility study settings.

3. Generalized Likelihood Ratio Inference

Likelihood ratio inference when the underlying distribution function is misspecified has been considered by several authors in the standard i.i.d setting (e.g., Foutz and Srivastava, 1977 and Kent, 1982). Rotnitzky and Jewell (1990) studied the GLR statistics for cluster-correlated data under the GEE settings proposed by Liang and Zeger (1986) with a single generalized linear model. They established the asymptotic distribution of the GLR statistics under the “working independent” correlation matrix. The method has been used for binary and ordinal response regression (Xie, Simpson and Carroll, 1997). Our fertility model defined in (1) is different from the aforementioned references. The mean structure in our model is a product of two independent components ($A$ and $P$), and each of the components can be modeled as a function of a generalized linear model. In addition, a set of weights $w_i$ can be incorporated to improve the model efficiency. The two covariate vectors in (1) capture the cluster-level and individual-level information, respectively.

Partition the parameter $\Theta$ as $\Theta = (\gamma', \lambda')'$, where $\gamma = (\gamma_1, \cdots, \gamma_r)'$ is the subset of $\Theta$. We are interested in the hypothesis test $H_0 : \gamma = \gamma_0$, and $\lambda = (\lambda_1, \cdots, \lambda_{p+q-r})'$ is the remaining subset of parameters. Let $\hat{\Theta}_0 = (\gamma_0', \lambda'_0)$ be the estimator of $\Theta$ subject to $\gamma = \gamma_0$, then the generalized likelihood ratio statistic is defined as

$$T_L(\gamma) = 2 \left[ \log L(\hat{\theta}) - \log L(\hat{\theta}_0) \right].$$

If $Y_{ij}, j = 1, \cdots, n_i$ are independently distributed, $T_L(\gamma)$ is asymptotically distributed as chi-square with degree of freedom $r$ under very mild regularity conditions. In our model, the independence assumption is violated and $T_L(\gamma)$ is no longer asymptotically chi-square distributed. Define matrices $J(\Theta)$ and $H(\Theta)$ as

$$J(\Theta) = E \left\{ U(\Theta) U(\Theta)' \right\},$$

and
Following the derivations of Foutz and Srivastava (1977), Kent (1982), and Rotnitzky and Jewell (1990), we generate the following asymptotic results for the GLR statistics in our setting.

**Theorem 1.** Under mild regularity conditions and model specification (1) with \( \gamma = \gamma_0 \), \( T_L(\gamma) \) converges in distribution to \( \sum_{j=1}^{r} d_j \chi_j^2 \) as the number of clusters increases, where \( \chi_j^2 \), \( j = 1, \ldots, r \) are independent \( \chi^2 \) random variables, and \( d_1, \ldots, d_r \) are the eigenvalues of

\[
H(\Theta) = -E \left\{ \frac{\partial}{\partial \Theta} U(\Theta) \right\},
\]

where

\[
H(\Theta)_{\gamma, \lambda} = H(\Theta)_{\gamma} - H(\Theta)_{\gamma}^{-1}(\Theta)_{\gamma, \lambda} H(\Theta)_{\lambda}.
\]

The asymptotic distribution of the GLR statistics \( L_T(\gamma) \) in the theorem enables us to construct a confidence region for parameters of interest based on the generalized profile likelihood function. Specifically, we define the generalized 100\( \alpha \)% confidence region for \( \gamma \) as the set,

\[
\{ \gamma : L_T(\gamma) \leq Q_\alpha(\Theta) \},
\]

where \( Q_\alpha(\Theta) \) is the 100\( \alpha \)% quantile of the weighted chi-square distribution given in Theorem 1. We would replace \( Q_\alpha(\Theta) \) by \( Q_\alpha(\hat{\Theta}) \) in practice.

To utilize the above theorem in fertility studies, we need to estimate the matrices \( J(\Theta) \) and \( H(\Theta) \), and to calculate the quantiles of the weighted chi-square distribution. To estimate \( H(\Theta) \), one can simply use the empirical \( \hat{M}_0(\hat{\Theta}) \). For \( J(\Theta) \), we suggest \( (1 - r/N)\hat{M}_1(\hat{\Theta}) \), although in principle any consistent estimate would work. Several methods can be used to calculate the quantile of the weighted sum of chi-squares. For \( r = 1 \), the appropriate distribution is simply a scaled \( \chi^2 \) distribution. Boos (1992) discussed the difficulties in practical use of generalized likelihood ratio test when \( r > 1 \). Some authors (e.g., Rao and Scott, 1981; Robert, Rao and Kumar, 1987 and Rotnitzky and Jewell, 1990) approximated the weighted sum of chi-square distribution with a standard chi-square distribution. We used an iterative algorithm for calculating quantiles of the weighted Chi-square distribution.

### 4. North Carolina Fertility Study

We analyze the data from the Early Pregnancy Study (EPS) (Wilcox et al., 1988) conducted in North Carolina using the proposed method. We use the GLR statistics to construct an asymmetric confidence interval for the probability of conception due to a specific day’s intercourse, and to draw inference about some important covariates including women’s age, her age at menarche, whether she was exposed to her mother’s smoking while she was in utero, and the effect of frequent intercourse on the probability of conception.
The Early Pregnancy Study (Wilcox et al., 1988) enrolled couples with no known fertility problems who were ready to begin a pregnancy. The couples enrolled in the study when they stopped contraception. Women collected a first urine specimen each morning, and kept a daily diary recording whether they had had unprotected intercourse in the preceding 24 hours and whether they had had any menstrual bleeding during that time. Participation continued through 8 weeks into a pregnancy or for 6 months if no clinical pregnancy was identified. Most women were in their mid-twenties to early thirties. Urine specimens were assayed for levels of metabolites of ovarian hormones and these results were used to identify the day of ovulation (Baird et al., 1991). A special feature of the North Carolina study is that "success" as defined by the implantation, can be detected through bioassay on the daily urine specimen (Baird et al., 1991, Wilcox, Weinberg and Baird, 1995).

This is signaled by hCG, the human pregnancy hormone, which rises exponentially around the time of implantation. Following Wilcox, Weinberg and Baird (1995), the outcome variable in our analysis for a particular menstrual cycle is defined by whether there is a rise in the human gonadotropin (hCG) around the putative time of implantation. Hence, because we are treating both occult conceptions (lost prior to clinical diagnosis) and clinical pregnancies as success, there could be multiple successes (implantations) for a couple with multiple cycles in the study. The present analysis is based on 555 menstrual cycles from 198 women, with 177 successes (implantations) and 378 failures. A unit weight is assumed for each observation in the following analysis.

4.1. Estimating day-specific probability of conception. As we mentioned in Section 2, a sub-model of (1) with \( g_2(\beta^* v_{ijk}) = p_k \) allows the nonparametric modeling of the day-specific conception probability. The Wald statistic is especially limited in estimating the day-specific probability of conception as the probabilities are confined to \([0,1]\) and their confidence intervals need to be asymmetric. Earlier analysis of EPS data has shown that the fertile window is 6 days up to and including the day of ovulation in each menstrual cycle (Wilcox, Weinberg and Baird 1995). Empirical evidence from the EPS study did not attribute any pregnancy to intercourse outside this window. Applying generalized likelihood ratio inference to fertility models, we calculate the 95% confidence intervals for the conception probabilities associated with intercourse days in this window. We contrast the confidence intervals with those obtained from the Wald statistics.

To calculate the 95% confidence intervals, we fit the following simple nonparametric model. Denote \( p_k, k = -5, -4, \ldots, 0 \) as the probability of fertilization due to intercourse on the \( k \)th day in a viable cycle, where day 0 is the ovulation day. Take \( g_1(\alpha^* u_{ij}) = A \), and \( g_2(\beta^* v_{ijk}) = p_k \), then model (1) becomes

\[
\mu_{ij} = \Pr(Y_{ij} = 1 | X_{ij}, U_{ij}, V_{ij}) = A \left[ 1 - \prod_{k=-5}^{0} (1 - p_k)^{v_{ijk}} \right].
\]

This is equivalent to letting design matrix \( V \) to be a 555 \( \times \) 6 matrix, where each column is an indicator variable for intercourse on the corresponding day. The estimation procedure is the one outlined in Section 2. The cycle viability probability,
Table 1. Semiparametric Model for Intercourse-related Probabilities in a 6 Day Window. \( p_0 \) is the probability for the Day of Ovulation. \( T_{95} \) is the 95th percentile of the rescaled \( \chi^2 \). \( \chi_{36.1}^2 \) is the 95th percentile of the central chi-square with degree of freedom 1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate/SE</th>
<th>Eigenvalue</th>
<th>( T_{95} )</th>
<th>( \chi_{36.1}^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_{-5} )</td>
<td>0.25/0.22</td>
<td>0.80</td>
<td>3.08</td>
<td>3.84</td>
</tr>
<tr>
<td>( p_{-4} )</td>
<td>0.47/0.20</td>
<td>1.10</td>
<td>4.21</td>
<td>3.84</td>
</tr>
<tr>
<td>( p_{-3} )</td>
<td>0.31/0.24</td>
<td>1.04</td>
<td>4.00</td>
<td>3.84</td>
</tr>
<tr>
<td>( p_{-2} )</td>
<td>0.75/0.22</td>
<td>0.95</td>
<td>3.67</td>
<td>3.84</td>
</tr>
<tr>
<td>( p_{-1} )</td>
<td>0.82/0.21</td>
<td>1.05</td>
<td>4.03</td>
<td>3.84</td>
</tr>
<tr>
<td>( p_{0} )</td>
<td>0.84/0.25</td>
<td>1.02</td>
<td>3.90</td>
<td>3.84</td>
</tr>
</tbody>
</table>

\( A \), is estimated at 0.37 with estimated standard error of 0.04. Table 1 lists the fitted point estimates and standard errors for the \( p_k \)'s, their eigenvalues and asymptotic 95%th quantile based on Theorem 1 and the 95% quantile of the naive \( \chi^2 \). The corresponding 95% fitted generalized profile likelihood confidence intervals based on (4) and those based on the Wald statistics for \( p_k \) are plotted in Figures 1-2. Clearly except for \( p_{-4} \), all the 95% confidence intervals for other \( p_k \)'s based on the Wald standard errors extend outside the parameter space either above or below (Figure 1).

![Figure 1. Conditional conception probabilities (conditional on cycle viability) and their 95% profile likelihood confidence intervals based on Wald statistics for Model 4 in Table 2. Day 0 corresponds to the estimated day of ovulation.](image-url)
Figure 2. Conditional conception probabilities (conditional on cycle viability) and their 95% profile likelihood confidence intervals based on GLR statistics for Model 4 in Table 2. Day 0 corresponds to the estimated day of ovulation.

We compare a sequence of models with the GLR criteria to see if the two component model fits the data better than would a single constant. Table 2 provides 4 models we considered. The p-values listed in Table 2 are based on the asymptotic sum of weighted chi-squares. The first one is a simple grand mean model. The second one is a two components model where the success probability of each day’s intercourse act is a constant \( p \) and

\[
\mu_{ij} = \Pr(Y_{ij} = 1|X_{ij}, U_{ij}, V_{ij}) = A \left[ 1 - (1 - p)^{\sum_{s=-5}^{0} x_{ijs}} \right].
\]

The two-components model fit the data statistically better than the grand mean model. Notice from Figure 1 that the probability of success rises fairly linearly in time. Model 3 in Table 2 captures this by modeling \( p_k = (1 + \exp(-\alpha - \beta * k))^{-1} \), i.e. \( \text{logit}(p_k) \) is a simple linear function of the time index, \( k \). The hypothesis test for \( H_0 : \beta = 0 \) is significant with p-value at 0.018. Jumping to model 4, the same non-parametric model studied in Table 1, we see that with 5 more degrees of freedom than model 2, this model does not significantly improve the fit over model 2. Here we used the property of transformation invariance of the GLR statistics for different models. The fitted point estimate \( \beta \) in model 3 is 0.63 with standard error of 0.346. The 95% confidence interval based on the asymptotic normality of Wald statistic is
then (-0.07, 1.28) which give an equivocal result for testing $H_0 : \beta = 0$. Using GLR we obtain the 95% confidence interval (0.13, 1.42) which excludes 0 at level 0.05.

Table 2. Models for regression analysis of North Carolina EPS data. E is the prenatal smoking exposure status. $k$ here indexes the day relative to ovulation (day 0), where $k = -5, \cdots, -1, 0$. Ovulation day in each cycle is estimated by the pattern of change in the ratio of two hormone levels, estrogen and progesterone.

<table>
<thead>
<tr>
<th>Model</th>
<th>Components of the Model</th>
<th>Dev. d.f.</th>
<th>Model compared</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>A</td>
<td>694.9</td>
<td>A</td>
<td>0.022</td>
</tr>
<tr>
<td>2.</td>
<td>$p$</td>
<td>689.7</td>
<td>$p = k$</td>
<td>0.018</td>
</tr>
<tr>
<td>3.</td>
<td>$\logit^{-1}(\alpha + \beta \times k)$</td>
<td>684.1</td>
<td>$p = -5, \cdots, 0$</td>
<td>0.284</td>
</tr>
<tr>
<td>4.</td>
<td>$A$</td>
<td>663.1</td>
<td>$A$</td>
<td></td>
</tr>
</tbody>
</table>

4.2. Assessing covariates effects in the EPS study. We now incorporate covariates into the fertility model and use the GLR statistics to evaluate the effects of various covariates on the probability of conception. The covariates considered here are $PSMK$, which is a binary variable indicating whether the women was exposed to her mother’s smoking while she was in utero, $AGE$, which is the women’s current age, $MAGE$, which is the age at menarche (age at first menstrual period), and $DDAY$, where $DDAY_k = 1$ if there was consecutive intercourse on both day $k$ and $k - 1$ and $DDAY_k = 0$ otherwise. Each of these covariates could plausibly affect fertility. Components of cigarette smoke are known to be harmful to the oocytes in animal studies (e.g., Kristensen et al., 1995 and MacKenzie, and Angevine, 1981). Fecundability decreases as women age (Bongaarts, 1983). It is also suspected that too frequent intercourse may reduce the potency of the sperm (Auger et al., 1995). We model both $A$ and $P$ with logit link function, that is, covariate vectors $u_{ij}$ and $v_{ijk}$ in the respective $A$ and $P$ components in (1) is modeled as the following

$$g_1(\alpha_0 + \alpha_1^* \times u_{ij}) = \{1 + \exp(-\alpha_0 - \alpha_1^* \times u_{ij})\}^{-1},$$

$$g_2(\beta_0 + \beta_1 + k + \beta_2 \times v_{ijk}) = \{1 + \exp(-\beta_0 - \beta_1 \times k - \beta_2 \times v_{ijk})\}^{-1}$$

Using the GLR statistics, we tested effect of the covariates of interests in (5). It appears that PSMK is significant in $A$ and MAGE is significant in $P$. The GLR statistics for model without PSMK and MAGE is 684.10 v.s. 664.03 for the model with them. This results a p-value of 0.003 based on the GLR statistics for testing the null hypothesis $H_0 : \beta_{PSM} = \beta_{MAGE} = 0$. Since $DDAY$ would be expected to act through effects on sperm quality, we model it in $P$. The GLR failed to reject the hypothesis that $DDAY$ has no effects (difference in GLR statistics is only 0.1). Given PSMK and MAGE in the model, $AGE$ is marginally significant (p=0.081). Since participant’s ages are relatively close (most in mid twenties to thirty), and age is known to affect women’s fecundability (older women are less likely to conceive), it would be reasonable to keep $AGE$ in the model.

The estimated regression coefficients and their standard errors (SD) for the above model is as follows. For the $A$ component, we have $\beta_{PSM} = -0.777$ with
where \( SD = 0.229 \) and \( \beta_{AGE} = -0.054 \) with \( SD = 0.034 \). For the \( P \) component, we have \( \beta_{DAY} = 0.553 \) with \( SD = 0.29 \) and \( \beta_{MAGE} = 0.520 \) with \( SD = 0.317 \). This suggests that prenatal exposure to maternal smoking has a deleterious effect on cycle viability. This is consistent with animal experiments showing a loss of primordial oocytes in newborns prenatally exposed through the mother to components of cigarette smoke (e.g., Kristensen et al., 1995 and MacKenzie, and Angervine, 1981). In this way, subfertility could be seen as a subtle birth defect associated with maternal smoking. The data also provide some evidence for a negative trend with age, and a beneficial effect of late menarche. The results also show the increased chance of conception when the intercourse is close to the day of ovulation. The data, however, do not support the existence of a negative effect of frequent ejaculation on sperm potency.

5. Discussion

We have discussed and applied a generalized likelihood ratio statistic for inference and model selection in fertility models. The proposed method would be applicable in various medical research settings where the available data are naturally cluster- or batch-correlated. The GLR method is a dependent-data generalization of the standard likelihood ratio test, which is only valid in independent settings. The GLR method reduces to the standard likelihood ratio method if independence holds. Unlike the standard likelihood ratio statistics, the GLR statistic has some advantages, such as invariance under parameter transformation, over the Wald-type test. Unlike the usual likelihood ratio statistic, we have shown the GLR statistic follows a weighted sum of chi-squares distribution asymptotically. We used a marginal model for the batch-correlated binary outcomes where the probability of success is modeled as a product of two components with each component a function of probabilities that follow a generalized linear model. This model allows the distinction of two levels of covariate, the cycle level and the day level.

Application of the GLR method to a data set from the Early Pregnancy Study showed some very interesting results that can not be observed using existing methodology. The data suggest that prenatal exposure to maternal smoking has a deleterious effect on cycle viability. The data also provide some evidence for a negative trend with age, and a beneficial effect of late menarche. Although putting the menarche effect in the model component related to intercourse rather than in the component related to cycle viability seems fit data better, the data do not allow us confidently to choose the one pathway over the other. The data do not support the existence of a negative effect of frequent ejaculation on sperm potency. The Early Pregnancy Study is an unique study in that the day of ovulation is precisely identified with each woman’s daily urinary estrogen and progesterone metabolites through bioassay on the daily urine specimen (Baird et al., 1991). However, for studies where such detailed measurements is not available, one need to consider the potential measurement error in the identifying the day of ovulation. In this case a fertile window wider than the six-day window from the EPS study would
be reasonable. Likewise, conclusions in an analysis also depend on if there are any other measurement errors in other covariates, such as $x_{ijk}$ in Section 4.1. Great care should be exercised when collecting data. If one believes there is measurement error exist in the data, then some further modeling of the measurement process is needed in order to get the unbiased estimates.

We used an algorithm to evaluate the quantiles of the appropriate distributions. Considering the current advances in computing power, a feasible alternative approach might be to simulate the empirical weighted chi-square distribution quantile.

Acknowledgements. The authors thank Drs. D.D. Baird and A.J. Wilcox for sharing the Early Pregnancy Data.

References


By Haibo Zhou
Department of Biostatistics
University of North Carolina
Chapel Hill, NC 27599-7400
U.S.A.
E-mail: zhou@bios.unc.edu

Douglas G. Simpson
Department of Statistics
University of Illinois
Champaign, IL 61820
U.S.A.
E-mail: simpson@illinois.edu

By Minge Xie
Department of Statistics
Rutgers University, Busch Campus
Piscataway, NJ 08855
U.S.A.
E-mail: mxi@stat.rutgers.edu

Clarice R. Weinberg
Biostatistics Branch
National Institute of Environmental Health Sciences
RTP, NC 27709, U.S.A.
E-mail: weinberg@niehs.nih.gov