Brazilian Journal of Probability and Statistics (2001), 15, pp. 15–32.
©Associação Brasileira de Estatística

SALAMANDER MATING EXPERIMENT – A BAYESIAN DATA ANALYSIS

Atanu Biswas[1] and **Kalyan Das**[2]

[1] *Applied Statistics Unit, Indian Statistical Institute 203 B. T. Road, Calcutta - 700 035, India. E-mail: atanu@isical.ac.in*

[2] *Department of Statistics, University of Calcutta 35 Ballygunge Circular Road, Calcutta - 700 019, India. E-mail: desk@cal3.vsnl.net.in*

Summary
The well-designed salamander mating experiment was reported by McCullagh and Nelder (1989). Different authors tried to analyze the salamander mating data from different considerations. The present work initially reviews the earlier works, and then provides a generalized mixed model which takes into account different random effects. The data is then analyzed through the Bayesian approach. The different random components are also estimated along with the mating probabilities.

Key Words: Generalized linear mixed model; Gibbs sampler; importance sampling; latent variable; penalized quasi-likelihood method; rejection sampling.

1 Introduction

Generalized linear mixed models (GLMM) are of considerable interest today because of their tremendous application in biomedical, environmental and other related fields. Traditionally, in such models, under the assumption of normality of random components it becomes extremely difficult to work with the full likelihood. This is because when the outcomes are binary (or count) in nature, a full maximum likelihood analysis based on their joint marginal distribution requires numerical integration techniques for calculation of the loglikelihood, score equations and information matrix. To date it has been proved intractable for more complicated problems involving irreducibly high-dimensional integrals. By way of illustration one can consider the famous salamander mating experiment (see McCullagh and Nelder (1989)) where the responses are binary and a mixed model

is considered where the associated study design is crossed. The scientific question addressed in this study is whether geographically isolated populations of salamanders develop barriers to successful mating. A second question may be to study the existence of heterogeneity among individuals in mating success and if so whether it is greater for males or females.

To avoid the numerical integration, one can consider the Bayesian methodology by taking repeated samples from the posterior distribution using importance or Gibbs (Besag et al., 1991, Zeger and Karim, 1991) sampling techniques. Bayesian approach is welcomed due to its flexibility for full assessment of the uncertainty in the estimated random effects and functions of model parameters.

The objective in this paper is first of all to extend the models that have earlier been considered by many authors in the context of salamander mating experiment and then to carry out the Gibbs sampling technique for obtaining repeated samples from the posterior distributions. In fact in Section 2 we discuss initially the salamander data and then consider various models used by different authors at subsequent stages. In Section 3, we describe our Bayesian model and explain its advantages over others. Unlike the other approaches we have developed the mating probabilities cellwise, i.e., corresponding to every combination of the salamanders permitted by the design. Finally in Section 4, we go for intensive computation and state our results.

2 Review of the literature

The salamander mating experiment, conducted in 1986 by S. Arnold and P. Verrell of the Department of Ecology and Evolution at the University of Chicago and reported by McCullagh and Nelder (1989), is one of the very few excellent statistically designed experiments in ecology and environmental studies available in the literature. In this experiment salamanders of same species but from two geographically isolated areas (Rough Butt (RB) and Whiteside (WS)) were considered. Three experiments were conducted, one in Summer of 1986 and two in Fall of 1986. From each of the above two areas 20 animals were used in the experiment, of which 10 male and 10 female salamanders, and also for each of ten animals two groups of size 5 were considered for mating. Thus in all there are 8 groups. Animals used in the first Fall experiment were identical to those in Summer experiment.

In fact, matings were performed in the laboratory according to the design used by Arnold and Verrell. The details of the design are available in McCullagh and Nelder (1989). In this experiment, instead of the maximum 16 possible between group crossings, only 8 between group crossings were permitted. According to the design each female in a group is paired 6 times, 3 from own population and 3 from others. This is true for each male also. Total number of matings per (female) group is $6 \times 5 = 30$. Thus, altogether there are 120 matings in each experiment. The 120 binary responses are

not all independent as (i) they represent 6 repeated crosses on each of 20 females and (ii) matings for different females are correlated if they share the same male.

Some obvious question arise out of this experiment. Besides the two questions addressed in the introduction there is one important one: How to make valid inferences of the four mating probabilities (π_{RR},π_{RW},π_{WR},π_{WW}) specific to a male salamander and a female in a particular season? (π_{RW}, for example, is the mating probability between a Rough Butt female (RBF) and a Whiteside male (WSM)). Again these mating probabilities may be influenced by some other conditions like summer/fall etc.

From the *biological* view point the present salamander data and data from some subsequent experiments are discussed in various papers mainly by the biological scientists. Quite naturally they concentrated on different biological aspects of salamander matings in different papers according to their interest. An interested reader can go through Tilley, Verrell and Arnold (1990), Verrell (1991), Herring and Verrell (1996), Vinnedge and Verrell (1998).

Eversince the wellknown salamander mating data has been published, quite a good number of statisticians looked at the analysis from different angle. Some interesting analysis of such data has been carried out by McCullagh and Nelder (1989), Karim and Zeger (1992), Breslow and Clayton (1993) and McCulloch (1994).

Let y_{ijk} be an indicator variable denoting the outcome of the mating experiment between i-th female and j-th male in the k-th set up, $i = 1, \cdots, m$; $j = 1, \cdots, n$; $k = 1, \cdots, r$, and $(i, j, k) \in d$, the design set. In this particular experiment $m = n = 20$. Note that all the possible combinations of (i, j, k) are not permitted by the design. y_{ijk} takes the value 1 or 0 according as a mating success or failure. We assume that all effects are additive on the logit scale, i.e.,

$$\text{logit } P\left[y_{ijk} = 1 \middle| b_i^f, b_j^m\right] = x_{ijk}\beta + z_{ik}b_i^f + w_{jk}b_j^m, \tag{2.1}$$

where x_{ijk}, z_{ik} and w_{jk} are associated covariates, and b_i^f, b_j^m are female and male random effects respectively.

Model A : (McCullagh and Nelder (1989)).

Here $r = 1$, i.e., analysis is done separately for each experiment. It is assumed that

1. $\pi_{RR}, \cdots, \pi_{WW}$ are same for each experiment.

2. b_i^f has mean 0 and variance σ^{f^2} for each i and independent of b_j^m which has mean 0 and variance σ^{m^2} for each j.

McCullagh and Nelder (1989) analyzed the Summer '86 data. Here $y = (\cdots, y_{ij}, \cdots)'$ is a 120×1 vector. Note that

$$\begin{aligned} E(y_{ij}) &= \pi_{RR} \quad \text{if} \quad i \in RBF,\ j \in RBM, \\ &= \pi_{RW} \quad \text{if} \quad i \in RBF,\ j \in WSM, \\ &= \pi_{WR} \quad \text{if} \quad i \in WSF,\ j \in RBM, \\ &= \pi_{WW} \quad \text{if} \quad i \in WBF,\ j \in WBM, \end{aligned}$$

with observed frequencies 22, 20, 7 and 21. They obtained the approximated covariance matrix as

$$\text{Disp } (y) = V_0(\pi) + \sigma^{f^2} V_1(\pi) + \sigma^{m^2} V_2(\pi),$$

where

$$V_0(\pi) = \text{ diag } (\pi_{RR}(1-\pi_{RR})I_{30}, \cdots, \pi_{WW}(1-\pi_{WW})I_{30}).$$

A quasi-likelihood approach is applied to y and estimates of $(\pi_{RR}, \pi_{RW}, \pi_{WR}, \pi_{WW})$ are obtained.

Model B : (Karim and Zeger (1992)).

Karim and Zeger (1992) considered the case $r = 2$, replication of the same experiment in Summer and Fall is considered. Here

$$x_{ijk} = (1, x_{i..}, x_{.j.}, x_{ij.}, x_{..k}),$$

where

$x_{i..}$	=	indicator of a female salamander which takes value 0 or 1 according as RBF or WSF,
$x_{.j.}$	=	indicator of a male salamander which takes value 0 or 1 according as RBM or WSM,
$x_{ij.}$	=	indicator which takes value 0 or 1 according as WSF and WSM or not,
$x_{..k}$	=	0 or 1 according as Summer or Fall.

Suppose both z_{ik} and w_{jk} contain $x_{..k}$ as a component letting $z_{ik} = (1, x_{..k})$ and $w_{jk} = (1, x_{..k})$. It is assumed that

$$\begin{aligned} b_i^f &\sim N_2\left(0, \Sigma^f\right) \quad \forall\ i, \quad \text{independent of} \\ b_j^m &\sim N_2\left(0, \Sigma^m\right) \quad \forall\ j. \end{aligned}$$

Note that, in Summer, $b_{1i}^f \sim N(0, \sigma_{11}^f)$, and in Fall, $b_{1i}^f + b_{2i}^f \sim N(0, \sigma_{11}^f + \sigma_{22}^f + 2\sigma_{12}^f = \sigma_0^{f^2})$. This is true for b_j^m also. Here $\pi_{RR}, \cdots, \pi_{WW}$ are different for two seasons. For example,

$$\begin{aligned}
\pi_{WR}^{fall} &= \int \frac{e^{\beta_0+\beta_1+\beta_4+\left(b_{1i}^f+b_{2i}^f\right)+\left(b_{1j}^m+b_{2j}^m\right)}}{1+e^{\beta_0+\beta_1+\beta_4+\left(b_{1i}^f+b_{2i}^f\right)+\left(b_{1j}^m+b_{2j}^m\right)}} dF\left(b^f\right) dF\left(b^m\right) \\
&\cong \frac{e^{(\beta_0+\beta_1+\beta_4)a(\sigma^2)}}{1+e^{(\beta_0+\beta_1+\beta_4)a(\sigma^2)}},
\end{aligned}$$

where $a(\sigma^2)$ is a suitably chosen function of $\sigma^2 = \sigma_0^{f^2} + \sigma_0^{m^2}$ (see Zeger, Liang and Albert (1988)). In fact Karim and Zeger (1992) obtained $a(\sigma^2) = [\{(16\sqrt{3})/(15\pi)\}^2\sigma^2 + 1]^{-1/2}$.

Karim and Zeger (1992) carried out Bayesian analysis with a noninformative prior on the dispersion matrix to estimate the underlying parameters. Difficulty arises as posterior distribution of b^f (or b^m) is not in a closed form. Sample from approximate distribution was obtained by rejection sampling (see Ripley (1987)).

Model C : (Breslow and Clayton (1993)).

Here all the 3 experiments are considered in (2.1). Here it is assumed that

$$\begin{aligned}
b^f &\sim N_{60}\left(0, \Sigma^{*f}\right), \\
b^m &\sim N_{60}\left(0, \Sigma^{*m}\right),
\end{aligned}$$

where

$$\begin{aligned}
\Sigma^{*f} &= \begin{pmatrix} \sigma_{11}^f I_{20} & \sigma_{12}^f I_{20} & 0 \\ \sigma_{21}^f I_{20} & \sigma_{22}^f I_{20} & 0 \\ 0 & 0 & \sigma_{22}^f I_{20} \end{pmatrix} \\
\Sigma^{*m} &= \begin{pmatrix} \sigma_{11}^m I_{20} & \sigma_{12}^m I_{20} & 0 \\ \sigma_{21}^m I_{20} & \sigma_{22}^m I_{20} & 0 \\ 0 & 0 & \sigma_{22}^m I_{20} \end{pmatrix}.
\end{aligned}$$

Analysis is done as in McCullagh and Nelder (1989) using quasi-likelihood based on approximate mean of y and Disp (y) or by penalized quasi-likelihood (PQL) method (see Breslow and Clayton (1993) for details).

Model D : (McCulloch (1994)).

McCulloch (1994) considered a slightly different model and obtained the maximum likelihood estimates of the underlying parameters using EM

algorithm (see Dempster et al. (1977)) separately for each of the three experiments. Here Gibbs sampler is used to obtain ML estimates rather than using a Bayesian framework. In the present paper we try to analyze the salamander data using a model similar to McCulloch (1994), but in a Bayesian framework and considering the three experiments simultaneously (alongwith the dependence of two experiments as the same animals were used in these Summer and first Fall experiments).

McCullagh and Nelder (1989) considered a single experiment and a simple model in a frequentist's view point where besides the population of the corresponding famale and male no other factor is considered to estimate the mating probabilities. Karim and Zeger (1992) considered two experiments in a more general model in a Bayesian framework where b_i^f (and b_j^m) are two component vectors and the mating probabilities are function of the season. Breslow and Clayton (1993), although they have considered all the three experiments, followed the approach of McCullagh and Nelder (1989). McCulloch (1994) used the model we use in this paper, but he followed the frequentist's approach and eventually estimated the mating probabilities ignoring the other effects.

3 Proposed Bayesian model and its analysis

Let us think of modeling the data as

$$P\left(y_{ijk}=1\left|b_i^f,b_j^m\right.\right)=\Phi\left(\frac{x_{ijk}\beta+z_{ik}b_i^f+w_{jk}b_j^m}{\sigma}\right), \tag{3.1}$$

where $\Phi(\cdot)$ is the distribution function of a standard normal distribution.

Conceptually (3.1) indicates a suitable choice of the conditional probability that takes care of effects due to i-th female (salamander), j-th male and other covariates. Our approach is general in the sense, we consider a model which incorporates all the information available from the experiment and finally estimate the mating probabilities taking into account the effects of season, experiment number, particular female and male salamanders.

Biologically (3.1) means that for a given pair (i,j), if matings are attempted several times in the laboratory, it's not that matings become always successful or always failure. The successful chance of mating depends on the courtship behavior (urge for mating), understanding (unobserved or latent character) between them.

Mathematically, this does mean, conditionally given $b=(b_1^{f'},\cdots,b_{20}^{f'}, b_1^{m'},\cdots,b_{20}^{m'})$,

$$\begin{aligned} y_{ijk} &= 1 \quad \text{if and only if} \quad y_{ijk}^*>0,\\ &= 0 \quad \text{if and only if} \quad y_{ijk}^*\le 0, \end{aligned}$$

where y^*_{ijk} is a latent variable subject to

$$y^*_{ijk} = x_{ijk}\beta + z_{ik}b^f_i + w_{jk}b^m_j + \epsilon_{ijk}, \tag{3.2}$$

such that

$$\begin{aligned} \epsilon_{ijk} &\sim N(0,\sigma^2) \;\; \forall\, i,j,k, \\ b^f_i &\sim N_q\left(0, D^f\right) \;\; \forall\, i, \\ b^m_j &\sim N_q\left(0, D^m\right) \;\; \forall\, j, \end{aligned}$$

independently of each other. Thus conditionally given b, for each i, j, k,

$$y^*_{ijk} \sim N\left(x_{ijk}\beta + z_{ik}b^f_i + w_{jk}b^m_j, \sigma^2\right),$$

independently of each other where z_{ik} and w_{jk} are q dimensional row vectors.

In fact (3.1) represents a random effect model. Such random effect models have been widely used to account for the dependence within clusters (e.g., Laird and Ware, 1982; Lindstorm and Bates, 1988; Zeger and Karim, 1991). The model (3.1) can be viewed as an extension of the binary regression model with latent variables considered by Albert and Chib (1993) by incorporating random effects to model dependency between observations.

We assume the following specific prior: β has a uniform prior over $\mathcal{R}^p$; $(\sigma^2)^{-1}$ is Gamma(0.001,0.001). The conjugate prior for D^{f-1} $[D^{m-1}]$ is the Wishart with ρ_f $[\rho_m]$ degrees of freedom and scale matrix $(\rho_f R_f)^{-1}$ $\left[(\rho_m R_m)^{-1}\right]$. It is known that (see, e.g., Carlin and Louis, 1996, p. 168) a good default specification is one for which ρ_f $[\rho_m]$ is taken as small as possible and R_f (R_m) is taken to be a prior guess at D^f $[D^m]$. Let us take $\rho_f = \rho_m = 2$ and $R_f = R_m = I_q$, which is a plausible choice in the absence of prior information about D^f and D^m (see Gilks et al., 1996, p. 411; McCulloch and Rossi, 1994). Note that in the absence of subjective prior information, the choice of uniform prior on β, independent of D^f, D^m and σ^2, leads to a proper conditional posterior distribution of β (see Natarajan and Kass, 2000). Note that in the present context q is 6.

Our objective is to carry out posterior analysis and then to estimate the marginal mating probabilities for the three experiments separately.

Now the conditional joint density of (y^*_{ijk}, y_{ijk}) is

$$\begin{aligned} & \pi\left(y^*_{ijk}, y_{ijk} \middle| b, \beta, \sigma^2, D^f, D^m\right) \\ = \;& \pi\left(y^*_{ijk} \middle| y_{ijk} = 0, b, \beta, \sigma^2, D^f, D^m\right) P\left(y_{ijk} = 0 \middle| b, \beta, \sigma^2, D^f, D^m\right) \end{aligned}$$

or

$$=\pi\left(y_{ijk}^{*}\left|y_{ijk}=1,b,\beta,\sigma^{2},D^{f},D^{m}\right.\right)P\left(y_{ijk}=1\left|b,\beta,\sigma^{2},D^{f},D^{m}\right.\right) \qquad (3.3)$$

Thus we can write the conditional joint density as

$$\begin{aligned}
&\pi\left(y^{*},y\left|b,\beta,\sigma^{2},D^{f},D^{m}\right.\right)\\
=\;& c\prod_{i}\prod_{j}\prod_{k}\left[1\left(y_{ijk}=1\left|b_{i}^{f},b_{j}^{m}\right.\right)1\left(y_{ijk}^{*}>0\left|b_{i}^{f},b_{j}^{m}\right.\right)\right.\\
&\left.+1\left(y_{ijk}=0\left|b_{i}^{f},b_{j}^{m}\right.\right)1\left(y_{ijk}^{*}\leq 0\left|b_{i}^{f},b_{j}^{m}\right.\right)\right]\\
&\times\phi\left(y_{ijk}^{*}\left|x_{ijk}\beta+z_{ik}b_{i}^{f}+w_{jk}b_{j}^{m},\sigma^{2}\right.\right), \qquad (3.4)
\end{aligned}$$

where $1(\cdot)$ is an indicator variable, c being a generic constant. $\phi(\cdot)$ is the density function of some normal random variable (or vector) with some parameters specified within the parenthesis.

The joint conditional density of $b(=(b^{f'},b^{m'})')$, $b^{f'}=(b_1^f,\cdots,b_{20}^f)$, $b^{m'}=(b_1^m,\cdots,b_{20}^m)$, is

$$\begin{aligned}
\pi\left(b\left|\sigma^{2},D^{f},D^{m}\right.\right) &= \pi\left(b^{f}\left|\sigma^{2},D^{f},D^{m}\right.\right)\pi\left(b^{m}\left|\sigma^{2},D^{f},D^{m}\right.\right)\\
&= \prod_{i=1}^{20}\pi\left(b_{i}^{f}\left|\sigma^{2},D^{f},D^{m}\right.\right)\prod_{j=1}^{20}\pi\left(b_{j}^{m}\left|\sigma^{2},D^{f},D^{m}\right.\right)
\end{aligned}$$

with

$$\begin{aligned}
\pi\left(b_{i}^{f}\left|\sigma^{2},D^{f},D^{m}\right.\right) &= \phi\left(b_{i}^{f}\left|0,D^{f}\right.\right),\\
\pi\left(b_{j}^{m}\left|\sigma^{2},D^{f},D^{m}\right.\right) &= \phi\left(b_{j}^{m}\left|0,D^{m}\right.\right). \qquad (3.5)
\end{aligned}$$

To use the powerful Gibbs sampler technique we need to find the conditional posterior densities. From (3.4),

$$\left[\beta\left|b,\sigma^{2},D^{f},D^{m},y^{*},y\right.\right]=N_{p}\left(\widehat{\beta}^{*},\frac{(X'X)^{-1}}{\sigma^{2}}\right), \qquad (3.6)$$

with

$$\widehat{\beta}^{*}=(X'X)^{-1}X'\left(y^{*}-\sum_{d}Zb_{i}^{f}-\sum_{d}Wb_{j}^{m}\right),$$

where d is the design set, X is a 360×5 matrix such that

$$X = \begin{pmatrix} \vdots \\ x_{ij1} \\ x_{ij2} \\ x_{ij3} \\ \vdots \end{pmatrix},$$

with x_{ijk} is a 1×5 vector, y^* is a 360×1 vector as

$$y^* = \begin{pmatrix} \vdots \\ y^*_{ij1} \\ y^*_{ij2} \\ y^*_{ij3} \\ \vdots \end{pmatrix},$$

where only those combinations (120 in total) of (i, j) are considered which are allowed by the design of Arnold and Verrell. Z (and similarly W) is 360×6 matrix with a typical 3×6 component

$$z_i = \begin{pmatrix} z'_{i1} \\ z'_{i2} \\ z'_{i3} \end{pmatrix},$$

with $z_{i1} = (1, x_{..1}, 0, 0, 0, 0)'$, $z_{i2} = (0, 0, 1, x_{..2}, 0, 0)'$ and $z_{i3} = (0, 0, 0, 0, 1, x_{..3})'$. Again

$$\left[\sigma^2 \,\middle|\, b, \beta, D^f, D^m, y^*, y\right] = c\left(\sigma^2\right)^{-\frac{360+2.002}{2}} e^{-\frac{1}{2}(0.001+a)/\sigma^2}, \tag{3.7}$$

which is inverse gamma, where

$$a = \left(y^* - \sum_d Zb_i^f - \sum_d Wb_j^m - X\beta\right)' \left(y^* - \sum_d Zb_i^f - \sum_d Wb_j^m - X\beta\right).$$

For $i = 1, \cdots, 20$, we have

$$\begin{aligned} &\left[b_i^f \,\middle|\, b^m, b_l^f, l \neq i, \beta, \sigma^2, D^f, D^m, y^*, y\right] \\ &= N_q\left(\left(z'_{i.}z_{i.} + \sigma^2 D^{f^{-1}}\right)^{-1} z'_{i.}\left(y^*_{i.} - x_{i.}\beta - \sum_{d_{i.}} w_{i.}b_j^m\right),\right. \\ &\left.\left(z'_{i.}z_{i.} + \sigma^2 D^{f^{-1}}\right)^{-1}\right), \end{aligned} \tag{3.8}$$

where $y_{i.}$ is a 18×1 vector and $x_{i.}$ is a 18×5 matrix, $z_{i.}$ and $w_{i.}$ are 18×6 matrices each comprising of 6 blocks of typical three-row component, and the 6 blocks corresponds to the 6 male salamanders designed to mate with the i-th female salamander. This is represented by the index set $d_{i.} = \{i_1, i_2, \cdots, i_6\}$, where the i_1-th, i_2-th, $\cdots$, i_6-th males are mated with the i-th female. Similarly, for $j = 1, \cdots, 20$, we have

$$\begin{aligned}
&\left[b_j^m \,\middle|\, b^f, b_h^m, h \neq j, \beta, \sigma^2, D^f, D^m, y^*, y\right] \\
&= N_q\left(\left(w_{.j}^{'} w_{.j} + \sigma^2 D^{m^{-1}}\right)^{-1} w_{.j}^{'}\left(y_{.j}^* - x_{.j}\beta - \sum_{d_{.j}} w_{.j} b_j^m\right),\right. \\
&\left.\left(z_{.j}^{'} z_{.j} + \sigma^2 D^{f^{-1}}\right)^{-1}\right). \qquad (3.9)
\end{aligned}$$

Further

$$\begin{aligned}
&\left[D^f \,\middle|\, b^f, b^m, \beta, \sigma^2, D^m, y^*, y\right] \\
&= c\left|B^f\right|^{20/2}\left|D^f\right|^{-\frac{20+q+1+\rho_f}{2}} e^{-\frac{1}{2}\text{ tr } D^{f^{-1}}(\rho_f R_f + B^f)}, \qquad (3.10)
\end{aligned}$$

with $B^f = \sum_{i=1}^{20} b_i^f b_i^{f'}$, which is an inverted Wishart. Similarly,

$$\begin{aligned}
&\left[D^m \,\middle|\, b^f, b^m, \beta, \sigma^2, D^f, y^*, y\right] \\
&= c\,|B^m|^{20/2}\,|D^m|^{-\frac{20+q+1+\rho_m}{2}} e^{-\frac{1}{2}\text{ tr } D^{m^{-1}}(\rho_m R_m + B^m)}, \qquad (3.11)
\end{aligned}$$

with $B^m = \sum_{i=1}^{20} b_i^m b_i^{m'}$. Finally,

$$\begin{aligned}
&\left[y_{ijk}^* \,\middle|\, b^f, b^m, \beta, \sigma^2, D^f, D^m, y\right] \\
&= \frac{\phi\left(y_{ijk}^* \,\middle|\, x_{ijk}\beta + z_{ik} b_i^f + w_{jk} b_j^m, \sigma^2\right)}{1 - \Phi\left(\frac{x_{ijk}\beta + z_{ik} b_i^f + w_{jk} b_j^m}{\sigma^2}\right)} \times 1\left(y_{ijk}^* \leq 0\right) \text{ if } y_{ijk} = 0, \\
&= \frac{\phi\left(y_{ijk}^* \,\middle|\, x_{ijk}\beta + z_{ik} b_i^f + w_{jk} b_j^m, \sigma^2\right)}{\Phi\left(\frac{x_{ijk}\beta + z_{ik} b_i^f + w_{jk} b_j^m}{\sigma^2}\right)} \times 1\left(y_{ijk}^* > 0\right) \text{ if } y_{ijk} = 1. \qquad (3.12)
\end{aligned}$$

To implement Gibbs sampler (see Geman and Geman, 1984, Gelfand et al., 1990), we consider initial guesses $(\beta^0, \sigma^{2^0}, D^{f^0}, D^{m^0})$ (may be MLE) and simulate from the conditional distributions (3.6)-(3.12). Thus after a large number (t) of iterations we obtain a sample from $[\beta^t, \sigma^{2^t}, D^{f^t}, D^{m^t}|y]$ and hence $\beta^{(t)}$ is same as that of drawing a sample from $[\beta|y]$.

4 Numerical computations

In this section we present the detailed computational results. To implement the Gibbs sampler technique we have taken a 1000 updates burn in followed by 5000 final samples of the experiment. Standard approaches can be considered for assessing convergence of the Gibbs sampler (e.q., Cowles and Carlin, 1994; Gelman and Rubin, 1992). In fact, in the present paper the convergence of Gibbs sampler is ensured by the basic approach of Gelman and Rubin (1992). Starting from some initial values of the parameters we generate 4 chains of values, each chain being generated starting from an over-dispersed distribution and with a sample size of 6000. We delete 1000 replications as "burning" samples to minimize the effect of initial values and retain the values of the next 5000 replications to approximate the posterior distributions. To monitor the convergence we focus our attention to the components of β and σ^2. Following Gelman and Rubin (1992) we compute the between and within chain mean squares of the retained values, say B and W respectively, for each of the parameters. Then we find

$$s^2 = (5000-1)W/5000, \quad \nu = s^2 + (4 \times 5000)^{-1}B,$$

and finally the potential scale reduction factor $r = \nu/W$. The potential scale reduction factors are nearly 1, and this suggests that the desired convergence is achieved in the Gibbs sampler.

Now we observe the following.

1. Posterior means of the components of β: (1.1143, -3.1629, -1.2846, 4.3936, -0.1920).
2. Posterior variances of the components of β: (0.3802, 0.1956, 0.5630, 1.2002, 0.1863).
3. Posterior mean of σ^2: 2.0051.
4. Posterior variance of σ^2: 0.2443.

The posterior means (and variances) of the components of b_i^f, $i = 1, \cdots, 20$, b_j^m, $j = 1, \cdots, 20$, D^f, D^m, and the estimated mating probabilities in all the three experiments for all the experimented pairs according to the design of the experiment is provided in Tables 1-4 and Figures 1-3. In Tables 1-4, the two components in each cell correspond to the posterior mean and variance respectively.

Table 1
Posterior means (and variances) of the components of b_i^f, $i = 1, \cdots, 20$.

b_1^f	0.1149	0.0082	0.0533	0.0645	-0.0046	0.1231
	0.2329	0.0562	0.0506	0.0444	0.0342	0.0367
b_2^f	0.1931	-0.0013	-0.1784	0.0626	-0.0074	-0.1244
	0.2462	0.0950	0.0642	0.0636	0.0437	0.0518
b_3^f	0.1956	-0.0038	0.0477	0.0409	0.0221	-0.0921
	0.3433	0.0844	0.0615	0.0522	0.0311	0.0403
b_4^f	0.2092	0.0444	0.0625	0.0686	-0.0002	0.1342
	0.3505	0.0830	0.0556	0.0424	0.0512	0.0400
b_5^f	0.1734	-0.0156	-0.2012	0.0932	-0.0161	-0.0713
	0.3414	0.0967	0.0734	0.0586	0.0342	0.0530
b_6^f	0.3001	0.0109	-0.0948	0.1066	-0.0204	0.0477
	0.2124	0.0844	0.0533	0.0441	0.0271	0.0463
b_7^f	-0.1740	0.0019	-0.0963	-0.0611	-0.0070	0.0528
	0.1812	0.0470	0.0503	0.0429	0.0358	0.0388
b_8^f	-0.2850	0.0272	0.1835	-0.0902	-0.0418	0.2164
	0.3126	0.0777	0.0645	0.0520	0.0419	0.0659
b_9^f	-0.2154	-0.0262	-0.1339	-0.0919	0.0048	0.0468
	0.2118	0.0511	0.0670	0.0360	0.0323	0.0352
b_{10}^f	-0.2475	-0.0687	-0.0960	-0.0795	0.0194	-0.1823
	0.2842	0.1021	0.0622	0.0516	0.0358	0.0620
b_{11}^f	-0.0384	-0.0319	-0.0218	-0.0318	-0.0095	-0.0510
	0.2422	0.0566	0.0675	0.0421	0.0244	0.0553
b_{12}^f	-0.0471	-0.0092	-0.0351	-0.0170	0.0157	-0.0483
	0.2389	0.0704	0.0531	0.0520	0.0401	0.0603
b_{13}^f	-0.0852	0.0960	0.3200	-0.0308	-0.0152	0.2420
	0.3422	0.1880	0.1000	0.0740	0.0631	0.0736
b_{14}^f	-0.1328	0.0668	0.3114	-0.0642	-0.0018	0.2318
	0.2130	0.1261	0.0857	0.0693	0.0486	0.0658
b_{15}^f	0.0162	-0.0114	-0.0015	-0.0126	0.0061	-0.0490
	0.1820	0.0582	0.0710	0.0403	0.0318	0.0803
b_{16}^f	-0.4237	-0.0738	-0.2005	-0.2015	-0.0160	-0.0940
	0.4414	0.1143	0.0814	0.0967	0.0509	0.0355
b_{17}^f	0.0787	-0.0346	0.1086	0.0423	0.0139	0.1561
	0.2820	0.0554	0.0621	0.0500	0.0471	0.0693
b_{18}^f	0.1563	-0.0489	-0.2314	0.0700	0.0275	-0.3205
	0.4328	0.1258	0.0863	0.0530	0.0554	0.0733
b_{19}^f	0.1902	0.0555	0.1011	0.0983	0.0442	0.1343
	0.4733	0.0710	0.0926	0.0594	0.0352	0.0444
b_{20}^f	0.1312	0.0029	-0.2128	0.0341	0.0126	-0.1046
	0.1745	0.0844	0.0521	0.0352	0.0276	0.0476

Table 2
Posterior means (and variances) of the components of b_j^m, $j = 1, \cdots, 20$.

b_1^m	0.0186	0.0046	0.0632	0.0010	0.0075	0.0560
	0.0186	0.0272	0.0201	0.0240	0.0229	0.0198
b_2^m	0.0383	-0.0006	0.0340	0.0476	0.0123	-0.0054
	0.0350	0.0240	0.0374	0.0286	0.0276	0.0283
b_3^m	0.0179	0.0276	0.0885	0.0200	0.0082	0.0501
	0.0435	0.0222	0.0295	0.0283	0.0311	0.0270
b_4^m	0.0304	-0.0002	-0.0257	0.0090	-0.0028	-0.0263
	0.0424	0.0258	0.0251	0.0260	0.0242	0.0220
b_5^m	0.0234	-0.0116	0.0545	0.0289	0.0114	0.0284
	0.0485	0.0255	0.0334	0.0197	0.0220	0.0226
b_6^m	-0.0952	0.0255	-0.0513	-0.0831	-0.0008	0.0307
	0.0492	0.0271	0.0431	0.0254	0.0217	0.0240
b_7^m	0.0044	0.0288	-0.0443	0.0046	0.0235	0.0584
	0.0301	0.0211	0.0392	0.0196	0.0173	0.0285
b_8^m	-0.0562	0.0067	-0.1086	-0.0454	0.0083	-0.0194
	0.0384	0.0184	0.0392	0.0224	0.0214	0.0200
b_9^m	0.0355	-0.0176	-0.1055	0.0050	0.0028	-0.0689
	0.0555	0.0244	0.0318	0.0277	0.0217	0.0262
b_{10}^m	-0.0859	0.0026	-0.0738	-0.0834	-0.0284	-0.0605
	0.0356	0.0286	0.0382	0.0273	0.0311	0.0209
b_{11}^m	-0.0294	0.0235	-0.1268	-0.0228	0.0003	-0.0836
	0.0587	0.0360	0.0372	0.0304	0.0246	0.0303
b_{12}^m	-0.0116	-0.0008	-0.1256	-0.0219	-0.0038	0.0081
	0.0425	0.0245	0.0411	0.0233	0.0164	0.0281
b_{13}^m	-0.0335	0.0132	-0.0277	-0.0295	0.0130	0.0972
	0.0506	0.0318	0.0396	0.0265	0.0211	0.0268
b_{14}^m	-0.0518	0.0178	-0.0723	-0.0316	0.0034	-0.0451
	0.0289	0.0245	0.0358	0.0339	0.0322	0.0261
b_{15}^m	-0.0444	-0.0084	-0.0755	-0.0565	-0.0100	-0.0233
	0.0322	0.0200	0.0397	0.0227	0.0225	0.0254
b_{16}^m	0.0633	0.0118	0.0551	0.0078	0.0007	0.0557
	0.0363	0.0281	0.0379	0.0245	0.0257	0.0313
b_{17}^m	0.0619	-0.0157	-0.0365	0.0143	0.0223	0.0016
	0.0274	0.0243	0.0276	0.0236	0.0289	0.0247
b_{18}^m	0.0245	0.0280	-0.0050	0.0353	0.0128	-0.0133
	0.0335	0.0218	0.0317	0.0187	0.0218	0.0255
b_{19}^m	0.0227	0.0266	0.1566	0.0050	-0.0125	0.1253
	0.0426	0.0332	0.0475	0.0242	0.0205	0.0314
b_{20}^m	0.0241	0.0243	0.0388	0.0155	0.0127	0.0502
	0.0325	0.0231	0.0251	0.0222	0.0246	0.0192

Table 3
Posterior means (and variances) of the components of D^f.

0.0814 0.0089					
0.0044 0.0013	0.0416 0.0015				
-0.0006 0.0012	0.0055 0.0005	0.0336 0.0005			
-0.0008 0.0003	0.0012 0.0002	-0.0006 0.0001	0.0229 0.0003		
0.0016 0.0006	-0.0003 0.0001	-0.0002 0.0002	-0.0005 0.0001	0.0240 0.0002	
0.0003 0.0003	0.0017 0.0002	-0.0008 0.0001	0.0004 0.0001	-0.0013 0.0001	0.0181 0.0002

Table 4
Posterior means (and variances) of the components of D^m.

0.0476 0.0022					
0.0001 0.0007	0.0333 0.0004				
-0.0007 0.0005	0.0006 0.0003	0.0358 0.0007			
-0.0034 0.0004	-0.0008 0.0003	-0.0006 0.0003	0.0285 0.0003		
-0.0033 0.0008	0.0024 0.0005	0.0008 0.0002	0.0032 0.0005	0.0310 0.0012	
0.0013 0.0001	0.0014 0.0002	-0.0020 0.0004	-0.0013 0.0001	-0.0006 0.0001	0.0243 0.0002

From the computation we see that the effect of the salamanders (females and males) considerably vary among themselves. Thus to draw any meaningful inference one has to pay attention to the individual female- and male-effects. Again Figures 1-3 show that the mating probabilities are different for different pairs even if the females and/or males are taken from a specified population (RBF or WSF), and thus it would be unreasonable to consider only four mating probabilities as in Models A, C and D. Although Model B considers effects of different experiments among the mating probabilities, in practice there are individual female- and male-effects which

should be taken into account. These female- and male-effects are provided in Tables 1 and 2. From those tables we see that the individual effects vary considerably among themselves.

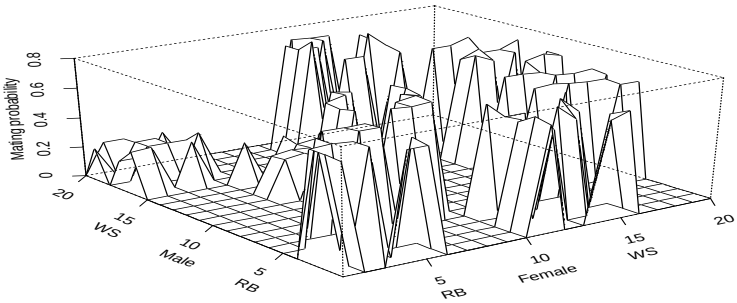

Figure 1: Estimated mating probabilities for Summer 1986 experiment

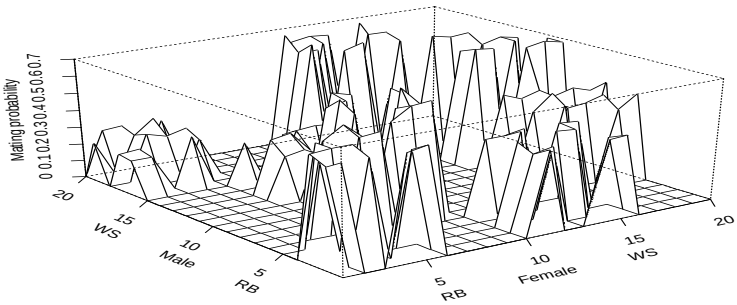

Figure 2: Estimated mating probabilities for Fall 1986 first experiment

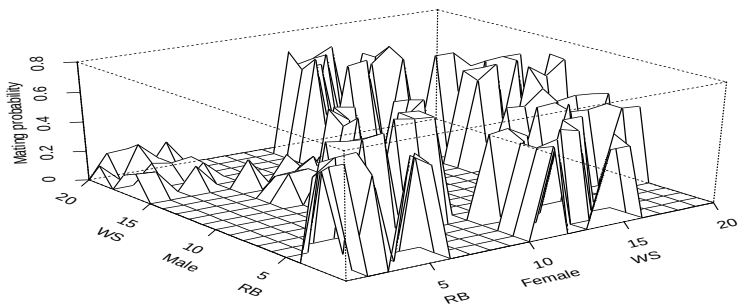

Figure 3: Estimated mating probabilities for Fall 1986 second experiment

Acknowledgements

The authors wish to thank the referee for his valuable comments which led to some improvement over an earlier version of the manuscript.

(Received December, 1999. Revised May, 2001.)

References

Albert, J. H. and Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association,* **88**, 669-679.

Besag, J., York, J. and Mollie, A. (1991). Bayesian image restoration, with two applications in spatial statistics (with discussion). *Annals of the Institute of Statistical Mathematics,* **43**, 1-59.

Breslow, N. E. and Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association,* **88**, 9-25.

Carlin, B. P. and Louis, T. A. (1996). *Bayes and empirical Bayes methods for data analysis.* London: Chapman and Hall.

Cowles, M. K. and Carlin, B. P. (1994). Markov Chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association,* **89**, 883-904.

Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum likelihood estimation from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society* B, **39**, 1-22.

Gelfand, A. E., Hills, S. E., Racine-Poon, A. and Smith, A. F. M. (1990). Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association,* **85**, 972-985.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science,* **7**, 457-472.

Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transitions on Pattern Analysis and Machine Intelligence,* **6**, 721-741.

Gilks, W. R., Richardson, S. and Spiegelhalter, D. J. (1996). *Markov Chain Monte Carlo in practice.* London: Chapman and Hall.

Herring, K. and Verrell, P. (1996). Sexual incompatibility and geographical variation in mate recognition systems: tests in the salamander *Desmognathus ochrophaeus. Animal Behaviour,* **52**, 279-287.

Karim, M. R. and Zeger, S. L. (1992). Generalized linear models with random effects; salamander mating revisited. *Biometrics,* **48**, 631-644.

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

Lindstorm, M. J. and Bates, D. M. (1988). Newton-Raphson and EM algorithms for linear mixed-effects models for repeated measure data. *Journal of the American Statistical Association,* **83**, 1014-1022.

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

McCulloch, C. E. (1994). Maximum likelihood variance components estimation for binary data. *Journal of the American Statistical Association,* **89**, 330-335.

McCulloch, R. and Rossi, P. E. (1994). An exact likelihood analysis of the multinomial probit model. *Journal of Econometrics,* **64**, 207-240.

Natarajan, R. and Kass, R. E. (2000). Reference Bayesian methods for generalized linear mixed models. *Journal of the American Statistical Association,* **95**, 227-237.

Ripley, B. (1987). *Stochastic simulation.* New York: Wiley.

Tilley, S. G., Verrell, P. A. and Arnold, S. J. (1990). Correspondence between sexual isolation and allozyme differentiation: a test in the salamander *Desmognathus ochrophaeus. PNAS* U.S.A., **87**, 2715-2719.

Verrell, P. (1991). Male mating success in the mountain dusky salamander, *Desmognathus ochrophaeus*: are small, young, inexperienced males at a disadvantage? *Ethology, 88*, 277-286.

Vinnedge, B. and Verrell, P. (1998). Variance in male mating success and female choice for persuasive courtship displays. *Animal Behaviour,* **56**, 443-448.

Zeger, S. L. and Karim, M. R. (1991). Generalized linear models with random effects; a Gibbs sampling approach. *Journal of the American Statistical Association,* **86**, 79-86.

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