now P.825

FITTING NONLINEAR AND CONSTRAINED GEES WITH OPTIMIZATION SOFTWARE

BU-1501-M **February, 2000**

Martha Contreras
Louise M. Ryan

Keywords: none provided.

Abstract:
In this paper we present an estimation approach for solving nonlinear constrained GEEs that can be implemented using object-oriented software for nonlinear programming, such as *nlminb* is Splus or *fmincon* and *lsqnonlin* in Matlab. We show how standard Estimation Equation theory includes this method as a special case so that our estimated when unconstrained will remain consistent and asymptotically normal. To illustrate this method we fit a nonlinear dose response model with nonnegative mixed bound constraints to clustered binary data from a developmental toxicity study. Satisfactory confidence intervals are found using a nonparametric bootstrap (BCa-Method) when a common-correlation coefficient is assumed for all the dose groups and for the dose specific dose groups except for the first and last exposure dose groups.

Fitting Nonlinear and Constrained GEEs with Optimization Software

Martha Contreras[1] **and Louise M. Ryan** [2]

[1]Department of Biometry, Cornell University, 435 Warren Hall, Ithaca, New York 14853, U.S.A.

[2]Department of Biostatistics, Harvard School of Public Health and Dana Farber Cancer Institute 677 Huntington Avenue, Boston, Massachusetts 02115, U.S.A.

SUMMARY

In this paper we present an estimation approach for solving nonlinear constrained GEEs that can be implemented using object-oriented software for nonlinear programming, such as *nlminb* in Splus or *fmincon* and *lsqnonlin* in Matlab. We show how standard Estimating Equation theory includes this method as a special case so that our estimates when unconstrained will remain consistent and asymptotically normal. To illustrate this method, we fit a nonlinear dose response model with nonnegative mixed bound constraints to clustered binary data from a developmental toxicity study. Satisfactory confidence intervals are found using a nonparametric bootstrap (BCa-Method) when a common-correlation coefficient is assumed for all the dose groups and for the dose specific dose groups except for the first and last exposure dose groups.

1. Introduction

Generalized Estimating Equations (GEEs) provide a convenient approach to parameter

[1]Corresponding author's e-mail address: mpc14@cornell.edu

Keywords: BCa-method, Bootstrap, Developmental toxicity study, Nonlinear GEEs, Nonnegative constraints, Object-oriented programming language.

[2]

estimation (see, Wedderburn (1974), Hansen (1982), Liang and Zeger (1986), Zeger and Liang (1986), Godambe (1991), Hall (1993), Diggle, Liang and Zeger (1996), and the references found there in). Although standard statistical theory has broader extensions, in practice GEEs have been primarily applied in the context of generalized linear models mostly because this has been the setting accommodated by readily available GEE software. For example, the generalized linear models procedure *proc genmod* in SAS 6.12 includes the option to fit GEEs allowing for various user options; that is, it allows users to "mix and match" with choices of several link functions relating the mean response, μ_{ij} (i indexes the subject and j the observations per subject), to a linear function of covariates (logit, probit, complementary log-log, inverse) and variance functions, including the Gaussian ($var(Y_{ij}) = \sigma^2$), binomial ($var(Y_{ij}) = \mu_{ij}(1 - \mu_{ij})$), and Poisson ($var(Y_{ij}) = \mu_{ij}$). Users can choose between several standard correlation matrices (e.g. exchangeable, autoregressive) or define their own.

However, the use of GEEs, for broader settings than the generalized linear model setting, has also increased in the past years mostly due to an increased availability of *object-oriented* programming software. That is, with moderate programming, it is now straightforward to fit models where the mean is some nonlinear function of the covariates, $\mathbf{X}_{ij}$, and an unknown parameter vector, $\beta \in R^p$. Packages such as *Matlab* contain nonlinear root-finding procedures, *fsolve*, for instance, which may prove successful if there are no restrictions on the parameter space. However, if the parameter space is constrained, that is, β is in some subset of R^p, then we will see that it is possible to use routines such as *nlminb* in *Splus* and *fmincon* or *lsqnonlin* in *Matlab* which may yield satisfactory estimates(see *Splus*'s and *Matlab*'s online documentation for furthered details).

However, even though readily accessible software may be found to solve a giv-

en nonlinear programming problem, computational issues aside, the presence of constraints complicates statistical inference for GEEs. That is, asymptotic properties of the estimators have only been established for certain forms of the parameter space. For instance, Self and Liang (1987) and Geyer (1994) established large sample properties for the estimators when they are constrained to lie within a closed subset of the Eucledian space R^p, but not necessarily and unbounded subset of R^p as would be the case when the parameters are constrained to be nonnegative. In practice, employing bootstrap methods to obtain confidence intervals is computationally intensive and demands that software be numerically stable so as to produce reliable estimates. Geyer (1991) considered the parametric bootstrap when constructing confidence intervals for an inequality constrained likelihood problem using Fortran based routines such as NPSOL(Gill, Murray, Saunders (1987)). For furthered details about the bootstrap the reader is referred to Efron and Tibshirani (1993), for instance.

In this article, we describe an approach to fitting GEEs that include nonlinear predictors and whose parameters may lie in an unbounded subset of R^p and be of an arbitrary form. This method is based on recognizing that a root-finding estimation problem can also be regarded as an optimization problem by minimizing the inner product or the estimating equations or, more generally, calculating the least squares solution to the estimating equations. Once the inner product has been formed, then it can be specified as the objective function for a nonlinear programming code. This includes routines such as *lsqnonlin* which default the system of equations to a nonlinear least squares problem and *regularizes* or safe-guards the system against possible model misspecification via the Levenberg-Marquardt numerical updates. Moreover, this routine allows for the presence of a variety of constraints. We illustrate this approach by fitting the so called *multi-hit* dose response model (see Holland and Sielken (1993))

from the *teratology* literature which assumes a mean function of the form

$$\mu_{ij} = 1 - exp(-\sum_{j=0}^{q} \beta_j x_i^{j\beta_{j+1}}), \tag{1}$$

where q is the *power parameter* typically a fixed integer (*1, 2 or 3*) and the β_j's are constrained to be nonnegative. We will also see that with this estimation approach that one can simultaneously estimate the intra-subject correlation parameters, ϕ, which are furthered constrained to be in the unit interval $[0, 1]$. Even though, this is now a mixed-bound nonlinear programming problem, we found that *lsqnonlin* yielded satisfactory estimates, thus making the computational of the confidence intervals via the bootstrap possible.

2. GEEs and Estimation

In this section we provide the necessary background to the theory of GEEs and its relationship to our proposed methodology.

We start by supposing that $\mathbf{Y}_i$ denotes an $n_i \times 1$ vector of outcomes for subject i, $i = 1, ...I$, and letting $\mathbf{X}_i$ be the corresponding $n_i \times p$ matrix of covariates and $\mathbf{X}_{ij}$ the $p \times 1$ vector corresponding to the j^{th} row of $\mathbf{X}_i$. In the most familiar examples of GEEs, the mean of $\mathbf{Y}_{ij}$ is related to a linear function of the covariates $\mathbf{X}_{ij}$ through a *link function*, f. That is, $\mu_{ij} = E(Y_{ij}) = f(\mathbf{X}_{ij}\beta)$, where β is a $p \times 1$ vector of unknown regression coefficients and $\mathbf{X}_{ij}$ is the jth row of $\mathbf{X}_i$. Usually, the variance of Y_{ij} is chosen to be a suitable function of μ_{ij}, and the covariance matrix of $\mathbf{Y}_i$ is then written in the factorized form

$$\mathbf{V}_i = \mathbf{A}_i^{1/2} \mathbf{R}_i \mathbf{A}_i^{1/2}, \tag{2}$$

where $\mathbf{A}_i = diag(var(Y_{ij}))$ and $\mathbf{R}_i$ is a correlation matrix. In general, the correlation matrix $\mathbf{R}_i$ will also depend on unknown parameters, which will need to be estimated

from the data. Nonetheless, using the theory of Generalized Method of Moments or GMMs (Wedderburn (1974), Liang and Zeger (1986), Zeger and Liang (1986), Hansen (1982), and Hall (1993)), the GEE estimate is obtained by solving

$$\mathbf{U}(\boldsymbol{\beta}) := \sum_{i=1}^{I} \left(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}}\right)^T \mathbf{V}_i^{-1}(\mathbf{Y}_i - \boldsymbol{\mu}_i) = \mathbf{0}, \tag{3}$$

where $\boldsymbol{\mu}_i$ is the vector of means, $\boldsymbol{\mu}_i = (\mu_{i1}, ...\mu_{in_i})^T$. (Note that (3)) can be viewed as the *first order necessary conditions* for the corresponding least squares problem when the variance $\mathbf{V}$ does not depend on the unknown parameters).

There is a large literature on numerical methods for solving nonlinear equations of the form $\mathbf{U}(\boldsymbol{\beta}) = 0$, where $\boldsymbol{\beta}$ may or may not be constrained. Detailed discussion can be found in a variety of places, including the classic textbooks on iterative solutions to nonlinear equations by Ortega and Rheinboldt (1970), Dennis and Schnabel (1986), Fletcher (1987), and Björck (1996). Nonetheless, if constraints are involved then a way to incorporate them and use existing optimization software is to recognize that solving (3) can be translated to the problem of minimizing over $\boldsymbol{\beta}$ the quadratic form

$$||\mathbf{U}(\boldsymbol{\beta})||^2 := \mathbf{U}(\boldsymbol{\beta})^T\mathbf{U}(\boldsymbol{\beta}), \tag{4}$$

which is the inner product of the estimating equations. When no constraints are present, (4) is the default numerical method of Matlab's *fsolve* which is intended to solve $\mathbf{U}(\boldsymbol{\beta}) = 0$ in a numerically stable fashion.

The regularity conditions relating (3) to (4), when constraints are not involved, can be seen by considering the first order necessary conditions for (4). That is, differentiating (4) with regards to $\boldsymbol{\beta}$ and solving for $\boldsymbol{\beta}$ gives

$$\mathbf{J}^T(\boldsymbol{\beta})\mathbf{U}(\boldsymbol{\beta}) = \mathbf{0}, \tag{5}$$

where $\mathbf{J}$ is as the $\sum_i^I n_i \times p$ Jacobian matrix.

If at the computed solution, $\mathbf{U}(\boldsymbol{\beta}) = \mathbf{0}$ then this is certainly a solution to (5). However, the reverse need not necessarily hold unless we have that the Jacobian at the solution is invertible or full rank. In fact, we note that (5) is just a special case of the more general estimating equations or GMMs considered by Hall (1993) (see page 400, equation (3.2) with $W_n := I$, $G_n := \mathbf{J}$, and $g_n := \mathbf{U}$). Statistical properties such as consistency and asymptotic normality of the estimates have also been established in Hall (1993) for the case when the parameters are unconstrained.

From a more practical perspective, we note that (4) provides another framework for estimating the correlation parameters and for modeling them as a function of covariates. As we will see in the next section, the idea is to recognize that the correlation estimator, $\hat{\boldsymbol{\phi}}$, can also be expressed as the solution to a set of estimating equations that simultaneously with the parameters can be estimated. Further discussion about the idea of setting up additional equations to estimate the unknown correlation parameters can be found in Liang, Zeger and Qaqish (1992) who refer to this method of estimating the correlation parameters as GEE2.

3. Application

To illustrate the proposed estimation approach and to illustrate how the bootstrap was performed, we consider a problem from the teratology literature where the need to fit nonlinear models arises often. For instance, Chen and Kodell (1989), and Holland and Sielken (1993) have discussed the importance of having a flexible class of dose response models of the following form; that is,

$$\mu_i = h(\beta_0 + \beta_1 x_i^{\beta_2}), \tag{6}$$

where x_i is the dose level assigned to say the ith litter and $h(.)$ is some function that takes value between 0 and 1 typically a *cumulative distribution function* or a *logistic*

function. In particular, the model we consider is the *one-hit model*, or that with a power parameter $q = 1$ in (1),

$$\mu_i = 1 - \exp\left(-\beta_0 - \beta_1 x_i^{\beta_2}\right), \tag{7}$$

which assumes that the mean of the $\mathbf{Y}_{ij}$ is the same for all litters mates (Note that for (7) to be within the interval $[0, 1]$, it suffices to constraint β_0, β_1, and β_2 to be nonnegative.)

Model (7) is a popular one in risk assessment due to its flexibility in capturing nonlinear effects with relatively few parameters, see Holland and Sielken (1993), Chen and Kodell (1989) and Ryan (1992). In fact, we note that if the parameter $\beta_2 = 0$ in (7), then the argument of the *exponential* function is constant and does not depend on the dose level. However, if this parameter is near 1 then the exponential function is almost linear in the dose level. But, as we will see from our particular analysis using this model, β_2 is distinctly different from 0 or 1.

Since it is important to how we implemented the bootstrap, we provide details of the teratology study we used.

In a typical teratology study pregnant dams (usually mice, rats or sometimes rabbits) are randomized to a control group or one of 3 or 4 exposed groups. Dams are exposed to the test substance during the period of major oganogenesis when the developing offspring are likely to be most sensitive to insult. Just prior to normal delivery, the dams are sacrificed and the uterine contents examined for defects. A typical study might have 25 to 30 dams per group, with anywhere from 1 to 20 offspring per litter.

For the teratology example, Y_{ij} might represent the weight of the jth pup from

the *ith* litter, or alternatively, might be a binary indicator of whether or not the pup was defective (e.g. dead or malformed). Generally, x_i will denote the dose level for the *ith* litter, though it is also possible that pup-specific covariates might be included. The dots in Figure 1 provide a graphical representation of data from a study in DEHP, an industrial plasticizer. That is, each dot corresponds to the response rate for a particular litter, while the crosses show the overall response rate within each dose group.

place Figure 1 about here

The study involved a total of 131 dams, including 30 controls and 101 exposed to one of 4 different dose groups ranging from .044mg/kg to .292 mg/kg. Furthered details may be found in Chen and Kodell (1989). The lines shown in the Figure 1 correspond to various fitted values, to be discussed presently.

Since we are assuming that the mean of Y_{ij} is the same for all litter mates, fitting model (7) assuming an exchangeable correlation matrix involves solving

$$\mathbf{U}_1 := \sum_{i=1}^{I} (\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}})^T \ [\mu_i(1-\mu_i)]^{-1} \mathbf{R}_i^{-1}(\mathbf{Y}_i - \boldsymbol{\mu}_i) = \mathbf{0}, \tag{8}$$

where $I = 5$ or the total number of dose groups. For model (7), this becomes

$$\mathbf{U}_1 := \sum_{i=1}^{I} \begin{pmatrix} 1 \\ x_i^{\beta_2} \\ \beta_1 x_i^{\beta_2} log(x_i) \end{pmatrix} \mu_i^{-1} \mathbf{1}_i^T \mathbf{R}_i^{-1}(\mathbf{Y}_i - \boldsymbol{\mu}_i) = \mathbf{0}, \tag{9}$$

where $\mathbf{R}_i = (1-\phi_i)\mathbf{I}_{n_i} + \phi_i \mathbf{J}_{n_i}$, and ϕ_i denotes the correlation parameter for the i^{th} dose group which will be determined using the data from that dose group. In particular, for the case of DEHP, there are 5 dose groups altogether, so that we will have 5 correlation coefficients, $\phi_1, ..., \phi_5$.

Using results about the inverse of an exchangeable correlation matrix (see Searle

and Henderson (1979)), (7) is furthered simplified to

$$\mathbf{U}_1 := \sum_{i=1}^{I} \begin{pmatrix} 1 \\ x_i^{\beta_2} \\ \beta_1 x_i^{\beta_2} log(x_i) \end{pmatrix} \frac{(r_i - n_i\mu_i)}{\mu_i[1 + \phi_i(n_i - 1)]} = 0, \tag{10}$$

where $r_i = \sum_{j=1}^{n_i} Y_{ij}$ is the number of abnormal pups among the n_i in the ith litter or ith dose group.

We estimate the dose-specific correlation coefficient by solving

$$\mathbf{U}_{2i} := [\frac{(r_i - n_i\mu_i)^2}{n_i\mu_i(1 - \mu_i)[1 + \phi_i(n_i - 1)]} - 1] = 0, \tag{11}$$

for $i = 1, ..., 5$, where $\mathbf{U}_{2i}$ is derived as follows. Since the r_i are distributed as a Bernoulli random variable with mean $n_i\mu_i$, moment-based estimators of the model correlation coefficients can be obtained by exploiting the fact that $(r_i - n_i\mu_i)^2$ has expectation equal to $n_i\mu_i(1 - \mu_i)[1 + \phi_i(n_i - 1)]$. Thus, the dose-specific estimates are obtained by solving the estimating equation

$$\mathbf{U} := (\mathbf{U}_1, \mathbf{U}_{2i})^T = \mathbf{0}_{1\times 8}, \tag{12}$$

for the expanded 1×8 parameter vector $(\boldsymbol{\beta}, \boldsymbol{\phi})$ where $\boldsymbol{\phi}$ is 5×1.

When a common correlation coefficient is estimated, $\boldsymbol{\phi}$ is a scalar and then (12) is a 1×4 system; thus (11) is summed over all litters giving

$$\mathbf{U}_2 := \sum_{i=1}^{I} [\frac{(r_i - n_i\mu_i)^2}{n_i\mu_i(1 - \mu_i)[1 + \phi(n_i - 1)]} - 1] = 0, \tag{13}$$

for $I = 5$.

4. Bootstrap Results

Because we have fitted a model with constraints, the bootstrap is an appropriate approach to inference. That is, as discussed in the previous section, to more closely resemble a teratology study and to insure that we always had a control group when we performed the bootstrap, we held the litter size, n_i, fixed per dose group i and re-sampled within that dose group. We gathered 1000 bootstrap samples, and then computed $95^{th}\%$ confidence intervals for the estimates via the *BCa* method discussed in Efron and Tibshirani (1993).

Table 1 summarizes the results of fitting the estimating equations. The two versions included the common correlation parameter(ϕ_1 on the right side of Table 1) for all dose groups, and another that allowed the correlation to change with dose; that is, ϕ_1 on the right half of Table 1 is now the correlation coefficient estimate obtained using the data in the control group in the dose specific case. The headings of *Unbtstrp, Btstrp, l.b.,* and *u.b.* correspond to the original or no bootstrap estimate, the mean of the 1000 bootstrap estimates, and the lower and upper bound of the bootstrap confidence intervals, respectively.

place Table 1 about here

Our results indicate that it is possible to obtain satisfactory 95% confidence intervals via the bootstrap BCa method for the model parameters, the common correlation parameter, and for the dose-specific correlation parameters, ϕ_1, ϕ_3, and ϕ_4 (see Table 1, Figure 1, Figure 2, and Figure 3). However, as we can see from Figure 2, the correlation parameter for the second dose group, ϕ_2 (which corresponds to the group exposed to the smallest dose of DEHP), and the fifth dose group, ϕ_5 (which corresponds to the group exposed to the largest dose of DEHP), have highly skewed distributions which deposit most of their mass at the boundary point 0. Note, nonetheless, that the

mean curves or the model parameters for the dose-specific analysis and the common correlation coefficient analysis seem to equally well describe the mean of the data (see Figure 1).

place Figure 2 and Figure 3 about here

5. Discussion

In the teratology literature, it is argued that, because litter mates tend to respond more similarly than nonlitter mates, that it is important to use statistical methods that properly allow for within litter correlations (see Williams (1975)), our analysis indicates that it can be difficult to construct confidence intervals particularly for the dose-specific correlation coefficients. However, in our analysis, this was not the case we if we assumed that the litter mates responded similarly and used all the data to estimate a common correlation coefficient. Arguably so, furthered statistical methodology is needed to construct proper confidence intervals when the parameters lie on the boundary as is the case for the first and last group exposed to DEHP. At this point, we are not prepared to comment as to the generality of these findings.

From a computational view point, we have been concerned with the numerical solution to the following problem

$$\begin{aligned} &\text{solve} && \mathbf{U}(\boldsymbol{\beta}, \boldsymbol{\phi}) = 0 \\ &&& \text{where} \\ &&& \boldsymbol{\beta} \geq 0 \\ &&& \text{and} \\ &&& \boldsymbol{\phi} \in [0, 1] \quad . \end{aligned} \tag{14}$$

We found that although stating problem (14) as the inner product of the estimating equations, $\min_{\boldsymbol{\beta},\boldsymbol{\phi}} \mathbf{U}^T\mathbf{U}$ subject to the same constraints as (14), leads it amenable to

an optimization routine such as *fmincon* or *nlminb*, these are still difficult computing problems for which success is problem dependent. For the particular application we considered, we found consistent success when using *lsqnonlin* which by default converts a system of equations into the inner product of the equations and minimizes this problem via a *large scale* method. Large scale methods are also an option in *fmincon*, however, we found it not to be an easy option to invoke and instead *fmincon* defaulted to *medium-scale* methods yielding unsatisfactory results. For further details on *large-scale* optimization the reader is referred to Biegler, Coleman, Conn, Santosa (1997). At this stage, we are not prepare to comment further on this. However, there are some advantages to considering the systems of equations as a least squares question in that, it can be argued, that it is more robust to model misspecification since the Newton updates are now done via the Levenberg-Marquardt regularization method.

ACKNOWLEDGMENTS

The authors thank the reviewers whose input greatly contributed to the paper. The first author also thanks Charles E. McCulloch for valuable discussions. This work was supported by NIH grant CA48061, NSF grant DMS-9710081, and Hatch Project NYC-151413.

REFERENCES

Björck, Å. (1996). *Numerical Methods for Least Squares Problems.* Philadelphia: SIAM Publications.

Chen, J. J., and Kodell, R. L. (1989). Quantitative risk assessment for teratologic effects. *Journal of the American Statistical Association* **84**, 966-971.

Biegler, L., Coleman, T., Conn, A., and Santosa, F. (ed's). (1997). *Large-Scale Optimization with Applications*, Vols. I-III, IMA Volumes in Mathematics and Applications, Vol. 92-94, Springer-Verlag.

Dennis Jr., J.E. and Schnabel, R.B. (1983). *Numerical methods for unconstrained optimization and nonlinear equations.* Englewood Cliffs: Prentice Hall.

Diggle, P.J., Liang, K.Y, and Zeger, S.L. (1996). *Analysis of longitudinal data.* Oxford: Oxford University Press.

Efron, B., and Tibshirani, R. (1993). *An introduction to the bootstrap*, New York:Chapman and Hall.

Fletcher, R. (1987). *Practical methods of optimization* (2nd ed.), New York: John Wiley.

Geyer, C.J. (1991). Constrained maximum likelihood exemplified by isotonic convex logistic regression, *Journal of the American Statistical Association*, **86**, 415, 717-724.

Geyer, C.J. (1994). On the asymptotics of constrained M-estimation, *The annals of statistics*, **22**, 4, 1993-2010.

Gill, P.E., Murray, W., Saunders, M.A., and Wright, M.E. (1986). *User's guide for NPSOL (Version 4.0): A Fortran package for nonlinear programming*, Technical report SOL 86-2, Stanford University, Dept. of Operations Research.

Godambe, V.P. (Editor) (1991). *Estimating Functions.* Oxford: Oxford University Press.

Hall, A. (1993). Some aspects of generalized method of moments estimation. In *Handbook of statistics 11: Econometrics* (ed. G.S. Maddala, C.R. Rao, H.D. Vinod), pp. 393-417. Amsterdam: Elsevier Science Pub.

Hansen, L.P. (1982). Large sample properties of generalized method of moments estimators, *Econometrica*, **50**, 4, 1029-1054.

Henderson, H.V., and Searle, S.R. (1979). Dispersion matrices for variance components models. *Journal of the American Statistical Association* **74**, 366, 465-470.

Holland, C.D. and Sielken Jr., R.L. (1993). *Quantitative Cancer Modeling and Risk Assessment.* Prentice Hall, Englewood Cliffs, NJ.

Liang, K.Y and Zeger, S.L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73** 13-22.

Liang, K.Y, Zeger, S.L. and Qaqish, B. (1992). Multivariate regression analyses for categorical data. *Journal of the Royal Statistical Society B* **54**, 1, 3-40.

McCullagh, P. and Nelder, J.A. (1989). *Generalized linear models.* London: Chapman and Hall.

Ortega, J.M. and Rheinboldt, W.C. (1970). *Iterative solution of nonlinear equations in several variables.* New York: Academic Press.

Ryan, L.M. (1992). The use of generalized estimating equations for risk assessment in developmental toxicity. *Risk Analysis* **12**, 439-447.

Self, S.G., and Liang, K.Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions, *Journal of the American Statistical Association*, **82**, 605-610.

Wedderburn, R.W.M. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika* **61**, 3, 439.

Williams, D. A. (1975). The analysis of binary responses from toxicology experiments involving reproduction and teratogenicity. *Biometrics* **31**, 949-952.

Zeger, S.L. and Liang, K.Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* **42**, 121-130.

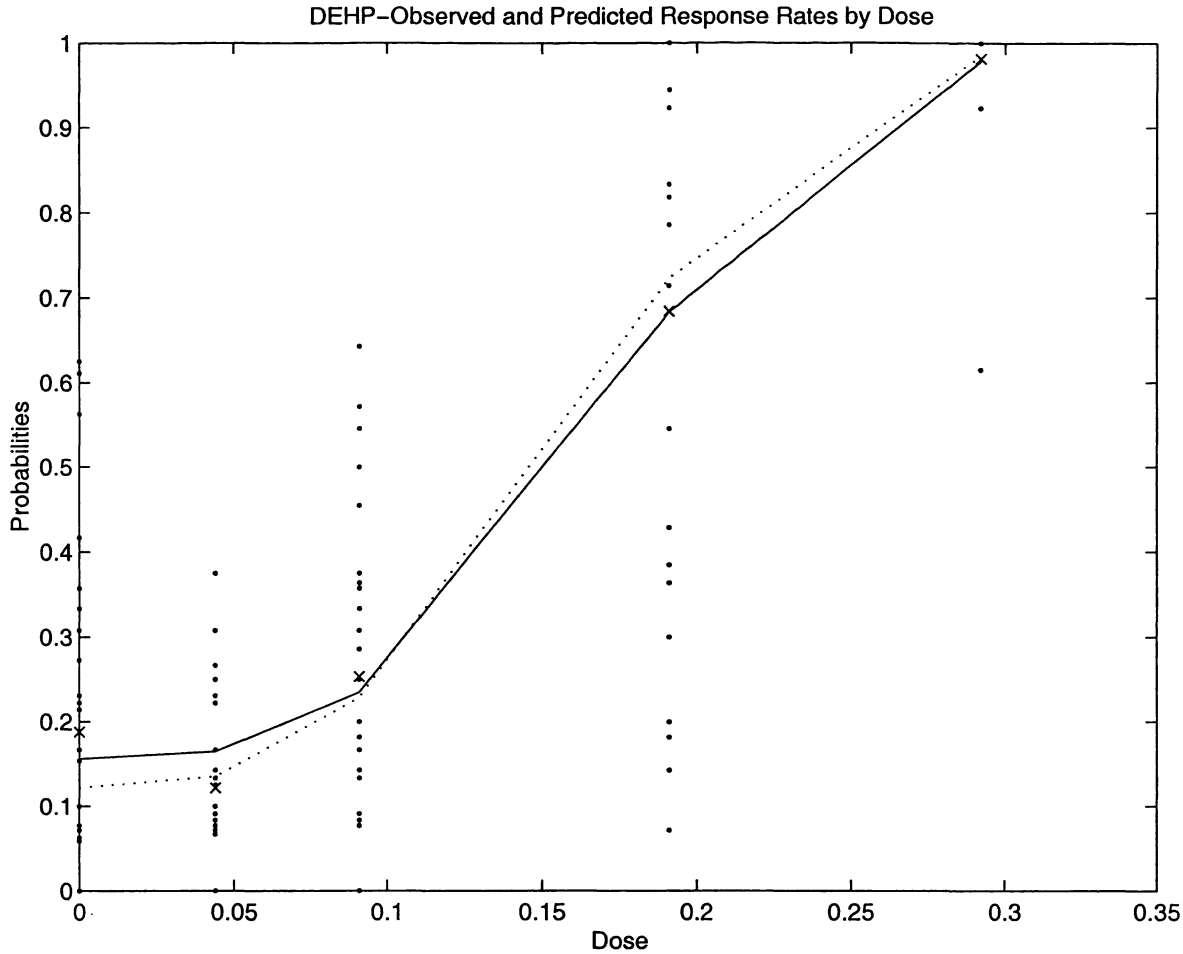

Figure 1: dots correspond to data, crosses to mean response rate within dose group, solid (fitted) line to common ϕ, and dashed (fitted) line to dose-specific ϕ

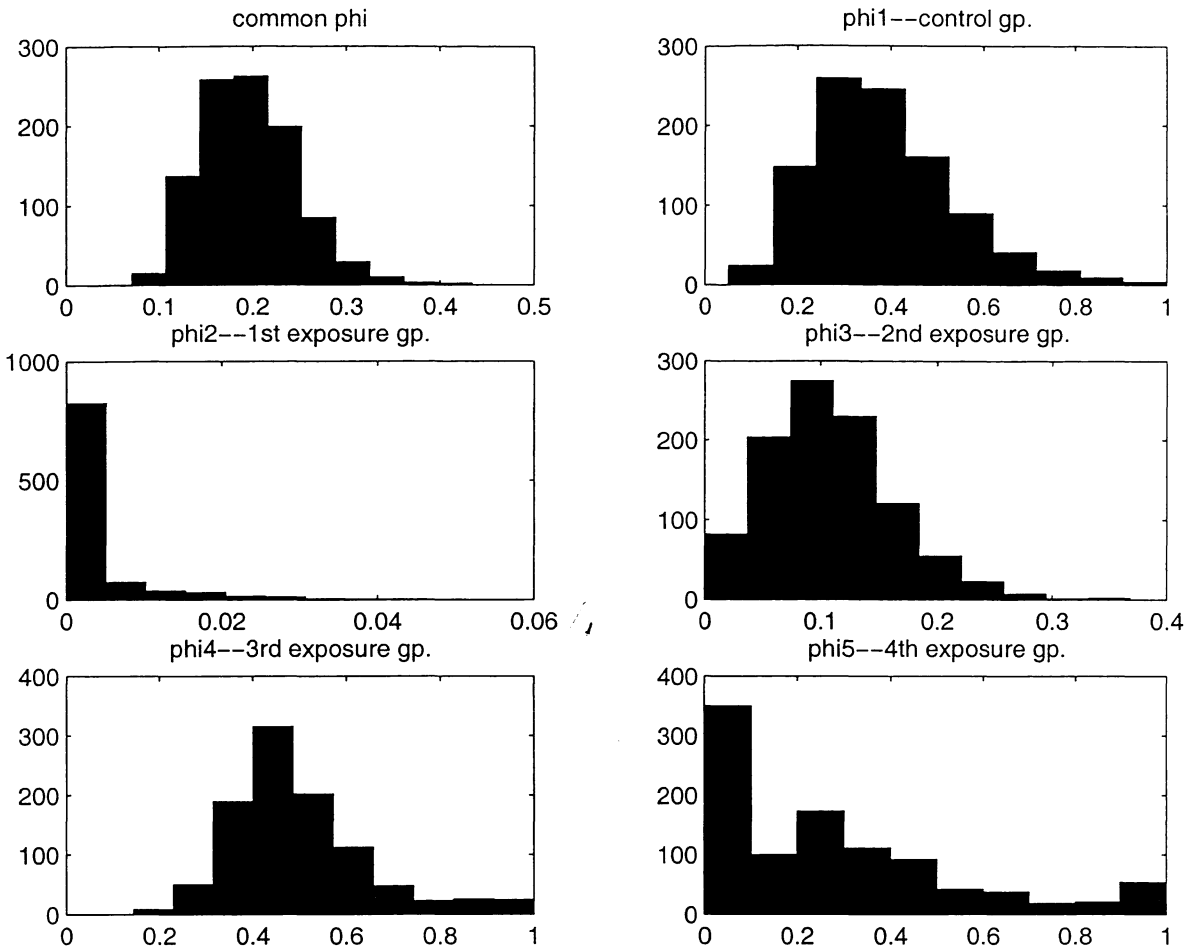

Figure 2: 1000 bootstrap estimates for the correlation coefficients

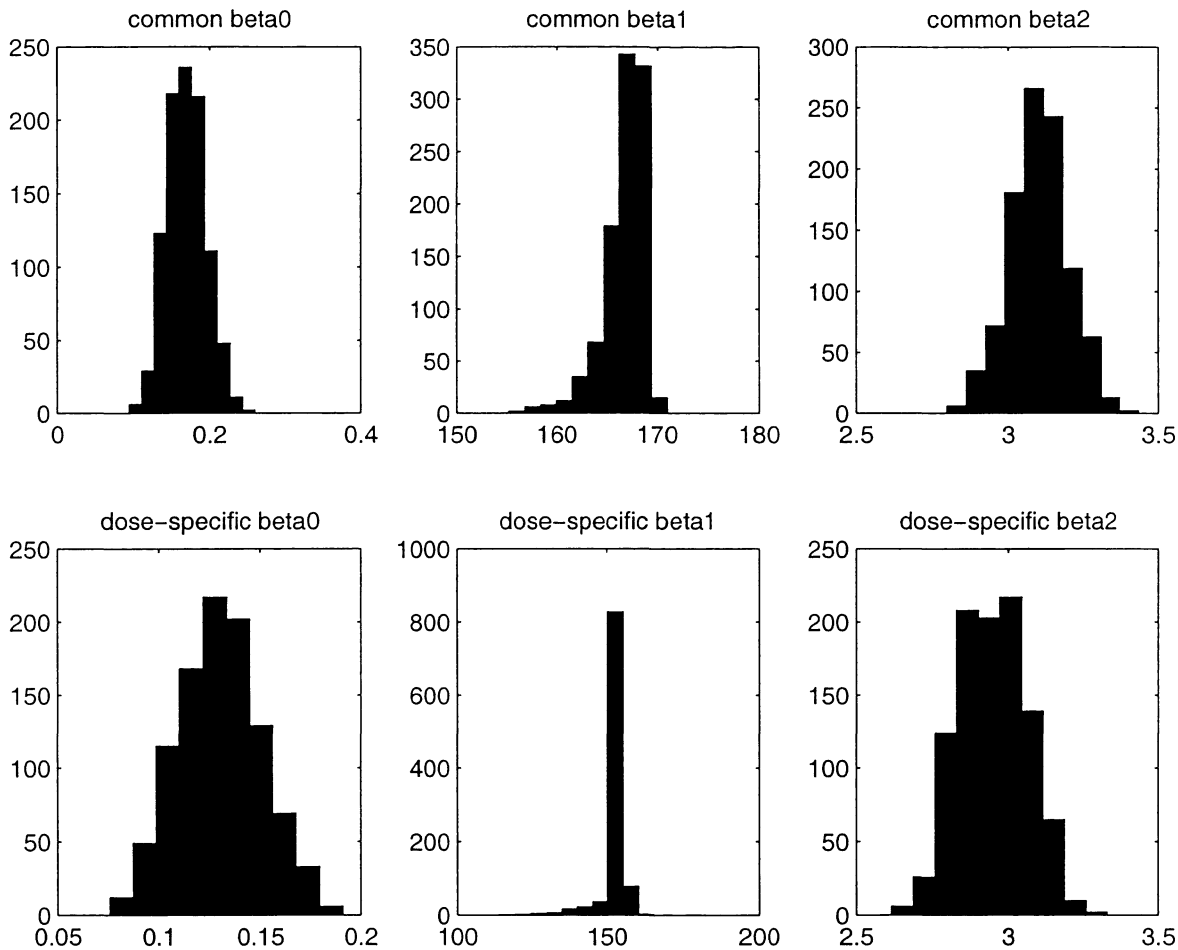

Figure 3: 1000 bootstrap estimates for the model parameters

Table 1

Model fitting for DEHP data

	Dose Specific ϕ				Common ϕ			
	Mean		Confidence Intervals		Mean		Confidence Intervals	
	Unbtstrp	Btstrp	l.b.	u.b.	Unbtstrp	Btstrp	l.b.	u.b.
β_0	.135	.129	.105	.173	.171	.169	.134	.216
β_1	153.498	152.523	139.824	155.488	166.500	166.706	158.283	168.329
β_2	2.99	2.94	2.836	3.195	3.106	3.103	2.945	3.279
ϕ_1	.350	.380	.167	.644	.198	.195	.132	.305
ϕ_2	.000	.002	.000	.006				
ϕ_3	.115	.107	.046	.236				
ϕ_4	.461	.498	.276	.758				
ϕ_5	.258	.271	.000	1.000				