Complementary Log Regression for Generalized Linear Models

WALTER W. PIEGORSCH*

Use and implementation of the complementary log regression model are discussed, integrating various separate applications of the model under the form of a generalized linear model. Some motivation is drawn from cases where an underlying random variable is reduced to a dichotomous form. Estimation and testing are facilitated by recognizing the complementary log as a specific link function within a generalized linear framework. Testing for goodness of link via efficient scores is also discussed.

KEY WORDS: Binomial model; Data truncation; Extended link family; Goodness-of-link testing; Logistic regression; Nonlinear regression.

1. GENERALIZED LINEAR MODELS FOR DICHOTOMOUS DATA

In many experimental settings, interest is directed at the effects of a set of explanatory variables, $x_1, \ldots, x_K$, upon an observed, dependent variate Y. A commonly employed model in these settings relates the mean response $E(Y)$ and the explanatory variables in a linear fashion:

$$E(Y) = \eta = \beta_0 + \beta_1 x_1 + \cdots + \beta_K x_K.$$

When the dependent variate is dichotomous, $E(Y)$ is simply the probability of response p. The associated linear model can be generalized to

$$g\{E(Y)\} = g(p) = \beta_0 + \beta_1 x_1 + \cdots + \beta_K x_K, \quad (1.1)$$

or simply $g(p) = \eta$, for some function $g(\cdot)$. Since it links the random and systematic components of the linear model, g is known as the link function (McCullagh and Nelder 1989, sec. 2.2).

The commonly seen link functions in this setting are the logit, $g(p) = \log\{p/(1-p)\}$; the probit, $g(p) = \Phi^{-1}(p)$; and the complementary log-log, $g(p) = \log\{-\log(1-p)\}$ (Santner and Duffy 1989, sec. 5.1). All three share the feature that they map the unit interval onto the real line. The logit link also shares theoretical connections with the natural parameter for the binomial model, $\theta = \log\{p/(1-p)\}$, and provides a useful interpretation in certain applications as the log odds of success (Cox 1970, sec. 2.3). Its use in (logistic) regression has become quite popular in recent years (Santner and Duffy 1989, sec. 5).

The above three link functions may be inappropriate, however, for certain experimental applications. In some of these cases, a useful alternative link is the complementary log

$$g(p) = -\log(1-p), \quad (1.2)$$

which maps the unit interval onto the positive real line. Notice that the inverse function is $p = 1 - \exp(-\eta)$ for $\eta > 0$. For $\eta \leq 0$, one could define the inverse link as simply $p = 0$, so that the inverse function is continuous and nondecreasing, $\forall\eta$. Thus the inverse link can be viewed as a form of distribution function, corresponding to an exponential random variable with unit mean. This connection between inverse links and distribution functions is common in binary regression models (Santner and Duffy 1989, sec. 5.1): The inverse probit clearly corresponds to a standard normal distribution, while the inverse logit corresponds to a standard logistic distribution with density function $e^x/(1+e^x)^2$ (Hastings and Peacock 1975, sec. 17).

The complementary log link has been applied in a wide variety of experimental settings. Applications have been described for epidemiological investigations where risk ratios are of interest (Wacholder 1986), for tests of simple independent action in 2×2 tables of proportions (Wahrendorf, Zentgraf, and Brown 1981), for calculations of carcinogenic drug potency (Bernstein, Gold, Ames, Pike, and Hoel 1985) and other aspects of carcinogenic dose response (Guess and Crump 1978; Whittemore 1983), and for multihit models of disease response (Cox 1962; Gart 1991). Cornell and Speckman (1967) discuss further applications. The complementary log link can also result in cases of data truncation, that is, when polytomous or continuous data are truncated into a dichotomous response. The goal herein is to provide a brief exposition on the use of (1.2), integrating estimation and fitting of the model within the generalized linear framework. Procedures for fitting (1.2) to data are discussed briefly in Section 2. Section 3 presents the details for the truncation phenomenon and includes an example. Section 4 deals with possible strategies for testing the complementary log link using goodness-of-link testing (Pregibon 1985).

2. FITTING THE COMPLEMENTARY LOG

Model (1.1) has a generalized linear form, and using (1.2), it may be fit via iteratively reweighted least squares (Green 1984). Guess and Crump (1978) discussed statistical inferences under the inverse model $p = 1 - \exp\{-\eta\}$ within the context of dose-response experi-

*Walter W. Piegorsch is a Mathematical Statistician in the Statistics and Biomathematics Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709. He thanks Norman Kaplan, Barry H. Margolin, Clarice R. Weinberg, and two anonymous reviewers for their helpful suggestions during the preparation of this article, and Arnold R. Brody and Lilla H. Hill for providing the data in Table 2.

mentation (i.e., η is a function of dose d). These approaches require the use of a computer and can be performed directly in most programming languages. Alternatively, the calculations can be performed with the GLIM statistical system (Payne 1987). GLIM provides likelihoods, maximum likelihood (ML) estimates, $\hat{\beta}$, for the parameters, and estimates of the asymptotic covariance matrix for the ML estimates. Conditions for existence and uniqueness of the ML estimates under (1.1) are given by Silvapulle (1981). For the complementary log link in (1.2), these ensure existence and uniqueness of $\hat{\beta}$ iff the sets

$$\mathcal{S}_0 = \left\{ \sum_{i \in \mathcal{J}_0} k_i x_i | k_i > 0 \right\},$$

and

$$\mathcal{S}_1 = \left\{ \sum_{i \in \mathcal{J}_1} k_i x_i | k_i > 0 \right\}$$

exhibit a nonempty intersection, where

$$\mathcal{J}_\omega = \{i | Y_i = \omega;\ i = 1, \ldots, n\}, \qquad \omega = 0, 1,$$

and $x_i = [1\ x_{1i} \ldots x_{Ki}]'$, such that $\beta' x_i > 0$.

Analyses are facilitated by Wacholder (1986), who provides simple user-defined GLIM programs ("macros") to assist the user in fitting generalized linear models that include the link in (1.2). [He referred to use of (1.2) as a "health ratio" model, since $1 - p$ is often the probability of avoiding an unhealthy outcome.] To satisfy the inherent requirement that $\hat{\eta} = \hat{\beta}' x > 0$, the fitting algorithm truncates any iterated η values that drop below zero to some small positive number, say 10^{-7}. Iteration then continues from that point. It has been noted (Baumgarten, Seliske, and Goldberg 1989), however, that Wacholder's macros may fail to achieve the true ML estimates, due to sensitivity in the iterative procedure to poor initial values. To alleviate this problem, more stringent criteria for convergence (e.g., lower relative or absolute errors in differences of iterative deviances or parameter estimates) are recommended (Baumgarten et al. 1989). All calculations for the various activities described in the following sections will be performed in GLIM.

3. DATA TRUNCATION LEADING TO THE COMPLEMENTARY LOG LINK

Consider first an experiment generating a discrete-valued response U_i, which is the number of occurrences of some phenomenon. This might be the number of tumors seen in a certain organ of an experimental animal, or the number of cells in a tissue or culture responding to a chemical stimulus (Collings, Margolin, and Oehlert 1981). If the observing mechanism or technique is such that only the occurrence of a nonnull state is recorded (e.g., "no tumors" versus "some tumors"), the data will be truncated into a dichotomous response. The observed variable becomes:

$$\begin{aligned} Y_i &= 1 \qquad \text{if } U_i > 0 \\ &= 0 \qquad \text{if } U_i = 0. \end{aligned}$$

Sobel and Elashoff (1975) have referred to this sampling scheme as (binomial) "group testing" (also see Chen and Swallow 1990). When interest centers on the nonresponse $Pr[y_i = 0]$, the data are often referred to as "Hansen frequencies" (Bishop et al. 1975, Ex. 14.6-2), based on E. W. Hansen's work in the behavioral sciences (Hansen 1966).

For the underlying response we often take the U_i as independently distributed Poisson random variates with mean, say, $\eta_i > 0$. Then, clearly, $Y_i \sim$ indep. $b(1, 1 - \exp\{-\eta_i\})$. This construction, based on Poisson occurrence rates, was discussed in detail by Cochran (1950), who had in mind application to bacterial concentrations in suspension and the planning of dilution experiments. He suggested that the concept was fairly well known, starting with the work of McCrady (1915) on the concentration of organisms in liquids. Of interest also is the seminal paper by Fisher (1922) on foundations of mathematical statistics: One of Fisher's in-depth examples involved estimation under (1.2) for serial dilution assays. Fisher described an experiment in which a soil or water sample was taken, and a series of dilutions of the sample was made to determine the presence or absence of some microbial contaminant. If Ω is the mean frequency of microbes in the initial ($i = 1$) sample, and if dilution proceeds by powers of a, then based on Poisson occurrence rates the expected proportion of plates containing any microbes at the ith dilution is

$$1 - \exp\{-\Omega/a^{i-1}\}, \qquad i = 1, \ldots, n$$

(Fisher 1922, sec. 3). Of interest is estimation of Ω. This is easily expressed as a generalized linear model, since the experimental scenario involves the predictor $\eta_i = \Omega/a^{i-1}$. That is, if $Y_i = 1$ when any microbial contamination is detected at the ith dilution ($Y_i = 0$ otherwise), we set $K = 1$ with known, zero intercept ($\beta_0 = 0$) and single predictor variable $x_{1i} = 1/a^{i-1}$. Attention is then directed at estimating $\beta_1 = \Omega$.

In general, if a set of explanatory variables, $x_{1i}, \ldots, x_{Ki}$, are associated with Y_i, a generalized linear model could be fit under this model construction using the complementary log link and $\eta_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_K x_{Ki}$.

Next, consider a continuous data setting where the underlying variable V_i is the time to failure of an experimental unit, indexed by $i = 1, \ldots, N$. Denote the cumulative distribution function associated with V_i by $F_i(t)$. Often, the observations will be limited to the number of units that fail by a specified time $t = \tau$. If the actual failure times occurring prior to τ are recorded, this is simply Type I censoring (Lawless 1982, sec. 1.4). If, however, it is only known whether or not the ith unit failed by time τ, we observe

$$\begin{aligned} Y_i &= 1 \qquad \text{if } i\text{th unit fails by time } \tau \\ &= 0 \qquad \text{otherwise.} \end{aligned}$$

Thus $Y_i \sim$ indep. $b(1, F_i\{\tau\})$.

In particular, for $V_i \sim$ indep. $\exp(\eta_i)$, $F_i(\tau) = 1 - \exp(-\tau\eta_i)$. Given a set of explanatory variables, $x_{1i}, \ldots, x_{Ki}$, for the ith observation, a generalized linear model can be fit using $\eta_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_K x_{Ki}$. The appropriate link is an extension of the complementary log function:

$$g(p) = -\tau^{-1} \log(1 - p), \tag{3.1}$$

where (1.2) obtains for $\tau = 1$.

Example 1. Layman, Agyras, and Glynn (1986) studied the analgesic effect of iontophoretic treatment with the nerve conduction-inhibiting chemical vincristine on elderly patients complaining of postherpetic neuralgia. Patients were interviewed $\tau = 6$ weeks after undergoing treatment to determine if any improvement in the neuralgia was evident. Assuming the actual time to cessation of pain (in weeks) follows an exponential distribution with possibly differing hazards for treated and untreated subjects, the complementary log model could be used to examine whether or not vincristine treatment was a contributing factor in the cessation of pain. The exponential assumption on time to pain cessation is a simple first choice for this setting. An extension could involve, for example, a two-parameter exponential model with a "guarantee time" parameter ϕ, prior to which no cessation is assumed to occur. Clearly, $\phi < \tau$. For known ϕ, the analysis proceeds as above, except that τ is replaced by $\tau - \phi$. For unknown ϕ, Lawless (1982, sec. 3.5) suggests use of methods for known ϕ applied over a range of plausible ϕ values, since numerical instabilities can occur when estimating ϕ under more advanced criteria.

Table 1 presents the data with the responses $Y_i = 1$ if pain cessation was reported by $\tau = 6$ weeks ($Y_i = 0$ otherwise); the variable x_{1i} indicates treatment, and three additional explanatory variables were also recorded: x_{2i} = age at treatment, x_{3i} = sex, and x_{4i} = pretreatment duration of complaint (months); $i = 1, \ldots, 18$.

Applying the link function in (3.1) with $\tau = 6$ and under the requirement that $\hat{\eta}_i > 0\ \forall i$, one finds that there was an analgesic effect attributable to vincristine treatment for these subjects; the likelihood ratio (LR) for testing the treatment term (i.e., $H_0{:}\beta_1 = 0$ versus $H_0{:}\beta_1 \neq 0$) is 11.84 with 1 df. The three additional variables—age, sex, and complaint duration—were not seen to significantly improve the fit after inclusion of the treatment term in the model.

Under the reduced model with linear predictor $\beta_0 + \beta_1 x_1$, the data naturally form a 2×2 table:

		x_1	
		0	1
Y	0	8	3
	1	0	7

Fitting the simple complementary log (1.2) to these data produces GLIM estimates $\hat{\beta}_0 = 1.9 \times 10^{-8}$ (with asymptotic s.e. $= 1.2 \times 10^{-4}$) and $\hat{\beta}_1 = 1.2038$ (s.e. = .4767). It is easy to identify the source of these values under (1.2): at $x_i = 0$, $\hat{p} = \frac{0}{8}$, while at $x_i = 1$, $\hat{p} = \frac{7}{10}$. Write (1.2) as $-\log(1 - \hat{p}(x_1)) = \hat{\beta}_0 + \hat{\beta}_1 x_1$. Then, to solve for $\hat{\beta}_0$, set $x_1 = 0$ and take

$$\hat{\beta}_0 = -\log(1 - 0) = 0.$$

To solve for $\hat{\beta}_1$, set $x_1 = 1$, keep $\hat{\beta}_0 = 0$ from above, and find

$$\hat{\beta} = -\log(1 - 0.7) - 0 = -\log(0.3) = 1.204.$$

Contrastingly, under the logistic model $p = \exp\{\beta_0 + \beta_1 x_1\}/(1 + \exp\{\beta_0 + \beta_1 x_1\})$, setting $x_1 = 0$ produces $p = \exp\{\beta_0\}/(1 + \exp\{\beta_0\})$, which is estimated as zero from 2×2 table. Thus modeling p via the logit link forces the estimate of β_0—which corresponds under the logit model to the empirical log odds of $Y = 1$ at $x_1 = 0$ (Santner and Duffy 1989, sec. 5.1)—to diverge to $-\infty$.

Table 1. Data From Neuralgia Study

Y	x_1	x_2	x_3	x_4
1	1	76	M	36
1	1	52	M	22
0	0	80	F	33
0	1	77	M	33
0	1	73	F	17
0	0	82	F	84
0	1	71	M	24
0	0	78	F	96
1	1	83	F	61
1	1	75	F	60
0	0	62	M	8
0	0	74	F	35
1*	1	78	F	3
1	1	70	F	27
0	0	72	M	60
1	1	71	F	8
0	0	74	F	5
0	0	81	F	26

NOTE: Data from Layman et al. (1986). Variables are explained in the text. Patient identified by asterisk died of unrelated causes prior to the six-week observation; improvement in pain was seen throughout and after treatment, so Y was set to 1.

4. TESTING GOODNESS OF LINK

To test the goodness of fit of link functions, Pregibon (1985) reviewed a general methodology that involves embedding the hypothesized link into an extended family of link functions, say $g(p; \gamma)$. Examples of such families exist for binomial models (Aranda-Ordaz 1981; Guerrero and Johnson 1982; Pregibon 1980) and may involve tolerance distributions (Copenhaver and Mielke 1977; Prentice 1976) or relative risk formulations (Breslow and Storer 1985). Morgan (1988) provides a useful review of such extended families. Some work has also considered goodness of link for ordinal regression (Genter and Farewell 1985).

Typically, these link families include as a member the popular logit link, $g(p) = \log\{p/(1 - p)\}$. A simple one-parameter family that includes both the logit ($\gamma = 0$) and complementary log($\gamma = 1$) link functions is

$$g(p; \gamma) = \log\{p + \gamma(1 - p)\} - \log(1 - p) \tag{4.1}$$

(Whittemore 1983) for $\gamma > -p/(1-p)$. Interest can be further restricted, if desired to $\gamma \geq 0$. This is a useful family, since the logistic can be a reasonable alternative to complementary log regression. Consider the problem of assessing possible synergistic activity between two stimuli in a 2×2 table of proportions. The complementary log link would be employed when the null model corresponds to simple independent action between the stimuli (Wahrendorf et al. 1981). The logistic link with additive contributions for the two stimuli corresponds to the null model of no three-way interaction in the corresponding $2 \times 2 \times 2$ contingency table, as developed by Bartlett (1935). Distinguishing between the two links and their corresponding models is of interest in certain experimental settings (Piegorsch, Weinberg, and Haseman 1986).

The extended link in (4.1) was considered by Whittemore (1983) to compare the relative fit of the logit model to other exponential forms for carcinogenicity data. In that context, the complementary log link resulting when $\gamma = 1$ in (4.1) can be viewed as a single-hit model of cancer initiation (Morgan 1988). Alternative, simplified families that include both logit and complementary log links include the forms.

$$\gamma \log(p) - (1-\gamma)\log(1-p)$$

(logit: $\gamma = \frac{1}{2}$; complementary log: $\gamma = 0$), or

$$\gamma \log(p) - \log(1-p)$$

(logit: $\gamma = 1$; complementary log: $\gamma = 0$). In both cases, one could require $0 \leq \gamma \leq 1$, although the links are valid functions for any real γ when $0 < p < 1$.

Following Pregibon (1985), testing for the complementary log link within the extended family $g(p; \gamma)$ involves the new linear predictor,

$$\eta = \beta_0 + \beta_1 x_1 + \cdots + B_K x_K + \delta z, \quad (4.2)$$

where $\delta = \gamma_0 - \gamma$, and $z = g'(p; \gamma_0)$, that is, $\partial g(p; \gamma)/\partial\gamma$ evaluated at $\gamma = \gamma_0$. Construction of the z variable is based on an assumption of local linearity (in γ) of $g(p; \gamma)$ near γ_0; see Pregibon (1980, 1985). [Notice the similarity to the misspecification problem in normal linear models; cf. Seber (1977, sec. 6.1) and White (1982).] Under (4.1), $\partial g(p; \gamma)/\partial\gamma = (1-p)/[p + \gamma(1-p)]$. Hypothesizing a complementary log model involves $\gamma_0 = 1$, so that $z = 1 - p$.

Since z usually involves unknown parameters, an initial fit of the data to $x_1, \ldots, x_K$ using the complementary log link is required to form an estimate for z, denoted here by $\hat{z}$. This can be, for example, $g'(\hat{p}; \gamma_0)$, where $\hat{p}$ is the ML estimate of p under $H_0{:}\gamma = \gamma_0$. [Further iteration to improve the estimate of z is performed only if interest exists in calculating the correct ML estimates of the β_j or δ; see Pregibon (1980).] Testing $\gamma = \gamma_0$ is then tantamount to testing $\delta = 0$. Pregibon (1985) specifically suggests use of the score test of $\delta = 0$, since it is in fact identical to the score test of $\gamma = \gamma_0$. Code for constructing score tests in GLIM is given by Pregibon (1982). He notes that the number of GLIM instructions necessary for calculation of the score statistic, say C^2, is no larger than that necessary for calculation of the corresponding LR statistic.

Under the hypothesis $H_0{:}\delta = 0$, C^2 is asymptotically distributed as χ^2 with 1 df. Application of the score test for $\delta = 0$ to (4.1) using C^2 is therefore straightforward. Some caveats regarding its use are in order, however. For instance, the nature of the link family, or of the design matrix associated with the predictor variables, may hinder the method by effectively aliasing $\hat{z}$ with some linear combination of the predictor variables. This can occur if the null model leads to fitted values for $\hat{p}$ that are effectively constant (Pregibon 1980), aliasing $\hat{z} = 1 - \hat{p}$ with the grand mean, or if the linear predictor only exhibits one significantly important *qualitative* predictor variable (e.g, a dichotomous indicator). Indeed, the latter problem occurs with the neuralgia data from Example 1, since the variables x_2, x_3, x_4 do not improve the fit of the model over and above the treatment indicator x_1. Thus $\hat{z} = 1 - \hat{p}$ is found to add insignificantly to the model fit as well, since it effectively takes on only two values ($\hat{z}$ would be perfectly aliased if the variables x_2, x_3, x_4 were dropped entirely from the model). Hence, calculation of C^2 for these data is effectively meaningless. Indeed, many models may describe data adequately under a dichotomous predictor (compare, however, the failure of the logit model described at the end of Example 1); it is only when more than one factor influences the outcome that distinguishing among two or more links becomes important.

Example 2. As a further illustration of the goodness-of-link methodology, consider the data in Table 2. These data represent results from a laboratory experiment where rodent alveolar/bronchiolar tissue was exposed in vitro to asbestos dust, to examine the inflammatory effects of asbestos exposure to mammalian lung tissue. The alveolar/bronchiolar cells were pretreated with an assumed-benign radioactive marker that is released after asbestos dust exposure, indicating whether any cell "activated," that is, generated an inflammatory response. Of interest was the effect of time (as hours after exposure, t) on the inflammatory response; additional explanatory variables of interest included cell type (epithelial or interstitial) and original cell location (terminal bronchial airway, alveolar duct, or bifurcation duct), which were arranged in a 2×3 factorial design.

Each experimental unit—a Petri plate—of cells was examined t hours after exposure for evidence of radioactivation. Unfortunately, the actual number of cells U_i activating on the ith plate could not be recorded. Only the presence or absence of *any* radioactivation was recorded per plate. Thus we observe a data truncation into the observed binomial responses

$$\begin{aligned} Y_i &= 1 \quad \text{if } i\text{th plate indicates any radioactivation} \\ &= 0 \quad \text{if not.} \end{aligned}$$

As noted in Section 3, the complementary log becomes a viable candidate for the link function, assuming a

Table 2. Rodent Tissue Radioactivation Data From in Vitro Exposure to Asbestos Dust

Time t (hr)	Cell location	Cell type	Plates examined	Plates responding
0	T	E	16	4
19	T	E	10	3
24	T	E	10	10
31	T	E	10	10
0	T	N	16	4
19	T	N	10	3
24	T	N	10	7
31	T	N	10	10
0	A	E	16	4
19	A	E	10	4
24	A	E	10	10
31	A	E	10	10
0	A	N	16	6
19	A	N	10	3
24	A	N	10	8
31	A	N	10	8
0	B	E	16	1
19	B	E	10	3
24	B	E	10	8
31	B	E	10	8
0	B	N	16	2
19	B	N	10	2
24	B	N	10	7
31	B	N	10	5

NOTE: Cell location is denoted by T (terminal bronchial airway), A (alveolar duct), or B (bifurcation duct). Cell type is denoted by E (epithelial) or N (interstitial).

Poisson distribution on the U_i, under this form of data truncation.

Fitting the complementary log link (1.2) to these data under the requirement that $\hat{\eta}_i > 0\ \forall_i$ yields a significant effect due to time: The LR statistic for the linear time effect is 68.85 on 1 df. Cell location apparently affects the nature of the radioactivated response as well; the LR statistic for the location factor is 8.53 on 2 df. No other effects (including the location $\times$ type interaction) were seen to improve the fit after the time and location terms were included in the linear predictor.

To assess the adequacy of the complementary log link for the data in Table 2, one can apply the goodness-of-link methodology described previously. Hypothesizing the family (4.1) with $\gamma_0 = 1$ yields a null model corresponding to (1.2). The z variable is simply $1 - p$, which is estimated as $\hat{z} = 1 - \hat{p}$ from a complementary log fit of the full model: time, location, type, and the location $\times$ type interaction. One then includes the additional term $\delta\hat{z}$ in a GLIM analysis (Pregibon 1980, 1982) to assess the null hypothesis $H_0{:}\delta = 0$, with $\delta = \gamma_0 - \gamma$, versus any departure. The resulting score statistic is $C^2 = 11.02$ on 1 df. Clear departure from H_0 is evidenced.

A similar goodness-of-link analysis using the logit as the null form under (4.1) provides a score statistic of $C^2 = 2.98$ on 1 df. This suggests that the logit link may be an appropriate altenative under which to formulate further inferences. Fitting the logit link to these data results in similar statements on the predictor variables as noted under (1.2): The LR statistic for the linear time effect is 76.77 on 1 df, whereas that for the (cell) location factor is 10.22 on 2 df. No other effects were seen to significantly improve the fit after the time and location terms were included in the linear predictor.

[*Received January 1990. Revised August 1990.*]

REFERENCES

Aranda-Ordaz, F. J. (1981), "On Two Families of Transformations to Additivity for Binary Response Data," *Biometrika*, 68, 357–363.

Bartlett, M. A. (1935), "Contingency Table Interaction," *Journal of the Royal Statistical Society*, Suppl. 2, 248–252.

Baumgarten, M., Seliske, P., and Goldberg, M. S. (1989), "Warning re. the Use of GLIM Macros for the Estimation of Risk Ratio," *American Journal of Epidemiology*, 130, 1065.

Bernstein, L., Gold, L. S., Ames, B. N., Pike, M. C., and Hoel, D. G. (1985), "Some Tautologous Aspects of the Comparison of Carcinogenic Potency in Rats and Mice," *Fundamental and Applied Toxicology*, 5, 79–86.

Bishop, Y. M. M., Fienberg, S. E., and Holland, P. W. (1975), *Discrete Multivariate Analysis: Theory and Practice*, Cambridge: M.I.T. Press.

Breslow, N. E., and Storer, B. E. (1985), "General Relative Risk Functions for Case-Control Studies," *American Journal of Epidemiology*, 122, 149–162.

Chen, C. L., and Swallow, W. H. (1990), "Using Group Testing to Estimate a Proportion, and to Test the Binomial Model," *Biometrics*, 46, 1035–1046.

Cochran, W. G. (1950), "Estimation of Bacterial Densities by Means of the Most Probable Number," *Biometrics*, 6, 105–116.

Collings, B. J., Margolin, B. H., and Oehlert, G. W. (1981), "Analysis for Binomial Data, With Applications to the Fluctuations Test for Mutagenicity," *Biometrics*, 37, 775–794.

Copenhaver, T. W., and Mielke, P. W. (1977), "Quantit Analysis: A Quantal Assay Refinement," *Biometrics*, 33, 175–186.

Cornell, R. G., and Speckman, J. A. (1967), "Estimation for a Simple Exponential Model," *Biometrics*, 23, 717–737.

Cox, D. R. (1962), "Further Results on Testing Separate Families of Hypotheses," *Journal of the Royal Statistical Society*, Ser. B, 24, 406–424.
——— (1970), *Analysis of Binary Data*, London: Chapman and Hall.
Fisher, R. A. (1922), "On the Mathematical Foundations of Theoretical Statistics," *Philosophical Transactions of the Royal Society*, 222, 309–368.
Gart, J. J. (1991), "An Application of Score Methodology: Confidence Intervals and Tests of Fit for One-Hit Curves," in *Handbook of Statistics* (vol. 8), eds. C. R. Rao and R. Chakraborty, Amsterdam: North-Holland, pp. 395–406.
Genter, F. C., and Farewell, V. T. (1985), "Goodness-of-link Testing in Ordinal Regression Models," *Canadian Journal of Statistics*, 13, 37–44.
Green, P. J. (1984), "Iteratively Reweighted Least Squares for Maximum Likelihood Estimation" (with discussion), *Journal of the Royal Statistical Society*, Ser. B, 46, 149–192.
Guerrero, V. M., and Johnson, R. A. (1982), "Use of the Box–Cox Transformation With Binary Response Models," *Biometrika*, 69, 309–314.
Guess, H. A., and Crump, K. S. (1978), "Maximum Likelihood Estimation of Dose-Response Functions Subject to Absolutely Monotonic Constraints," *The Annals of Statistics*, 6, 101–111.
Hansen, E. W. (1966), "The Development of Initial and Infant Behavior in the Rhesus Monkey," *Behavior*, 27, 107–149.
Hastings, N. A. J., and Peacock, J. B. (1975), *Statistical Distributions*, New York: John Wiley.
Lawless, J. F. (1982), *Statistical Models and Methods for Lifetime Data*, New York: John Wiley.
Layman, P. R., Agyras, E., and Glynn, C. J. (1986), "Iontophoresis of Vincristine Versus Saline in Post-herpetic Neuralgia: A Controlled Trial," *Pain*, 25, 165–170.
McCrady, M. H. (1915), "The Numerical Interpretation of Fermentation-Tube Results," *Journal of Infectious Disease*, 17, 183–212.
McCullagh, P., and Nelder, J. A. (1989), *Generalized Linear Models* (2nd ed.), London: Chapman and Hall.
Morgan, B. J. T. (1988), "Extended Models for Quantal Response Data," *Statistica Neerlandica*, 42, 253–272.
Payne, C. D. (ed.) (1987), *The GLIM System—Release 3.77—Manual* (2nd ed.), Oxford: Numerical Algorithms Group.
Piegorsch, W. W., Weinberg, C. R., and Haseman, J. K. (1986), "Testing for Simple Independent Action Between Two Factors for Dichotomous Response Data," *Biometrics*, 42, 413–419.
Pregibon, D. (1980), "Goodness of Link Test for Generalized Linear Models," *Applied Statistics*, 29, 15–23, 14.
——— (1982), "Score Tests in GLIM With Applications," in *GLIM82: Proceedings of the International Conference on Generalized Linear Models*, ed. R. Gilchrist, New York: Springer-Verlag, pp. 87–97.
——— (1985), "Link Tests," in *Encyclopedia of Statistical Sciences* (vol. 5), eds. S. Kotz, N. L. Johnson, and C. B. Read, New York: John Wiley, pp. 82–85.
Prentice, R. L. (1976), "A Generalization of the Probit and Logit Methods for Dose Response Curves," *Biometrics*, 32, 761–768.
Santner, T. J., and Duffy, D. E. (1989), *The Statistical Analysis of Discrete Data*, New York: Springer-Verlag.
Seber, G. A. F. (1977), *Linear Regression Analysis*, New York: John Wiley.
Silvapulle, M. J. (1981), "On the Existence of Maximum Likelihood Estimates for the Binomial Response Models," *Journal of the Royal Statistical Society*, Ser. B, 43 310–313.
Sobel, M., and Elashoff, R. M. (1975), "Group Testing With a New Goal, Estimation," *Biometrika*, 62, 181–193.
Wacholder, S. (1986), "Binomial Regression in GLIM: Estimating Risk Ratios and Risk Differences," *American Journal of Epidemiology*, 123, 174–184.
Wahrendorf, J., Zentgraf, R., and Brown, C. C. (1981), "Optimal Designs for the Analysis of Interactive Effects of Two Carcinogens or Other Toxicants," *Biometrics*, 37, 45–54.
White, H. (1982), "Maximum Likelihood Estimation of Mis-specified Models," *Econometrica*, 50, 1–26.
Whittemore, A. S. (1983), "Transformations to Linearity in Binary Regression," *SIAM Journal of Applied Mathematics*, 43, 703–710.