

This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: Annals of Economic and Social Measurement, Volume 4, number 3

Volume Author/Editor: NBER

Volume Publisher: NBER

Volume URL: <http://www.nber.org/books/aesm75-3>

Publication Date: July 1975

Chapter Title: Qualitative Response Models

Chapter Author: Takeshi Amemiya

Chapter URL: <http://www.nber.org/chapters/c10405>

Chapter pages in book: (p. 363 - 372)

QUALITATIVE RESPONSE MODELS

BY TAKESHI AMEMIYA*

This article gives a systematic discussion of various qualitative response models, with a special emphasis on multi-response and multivariate models. While some new models (notably, multivariate polytomous probit models) are defined, old models are given new interpretations. The article discusses the relative merits of two basically differing ways to formulate multivariate models: the one that specifies marginal probabilities first and the one that specifies conditional probabilities first.

1. INTRODUCTION

In this paper I will consider various ways to define qualitative response models. Qualitative response models may be defined generically as models that involve one or more discrete random variables whose conditional probability distribution given the values of the independent variables is specified up to a finite number of unknown parameters. These models have been extensively discussed in the biometric literature. See, for example, Cox [1970] or Finney [1971] for a survey and references. However, since most of the papers have been concerned with the univariate dichotomous case, I believe a systematic account of polytomous and multivariate models attempted here will serve a useful purpose.

I take for granted the merits of the normal or logistic transformation when I extend univariate dichotomous models to either polytomous or multivariate ones. Hence, all of the models discussed in this paper fall into the general category of normal or logistic models. Ideally, one should specify a model on the basis of realistic behavior assumptions about the respondent. The difficulty of this approach is, however, that it often leads to an estimation problem which is computationally intractable. In the univariate dichotomous case, probit and logit models have proved useful in explaining real data in addition to being computationally manageable. Whether or not the same holds true for some of the polytomous and multivariate models defined in this paper remains to be seen.

While I define a number of new models, I also discuss models that have been proposed by others giving them new interpretations and illuminating their differences. The univariate dichotomous model is well-known, but I outline it at the outset to provide a background for the subsequent development. The problem of estimation is mentioned only briefly, as I have discussed it fully in the context of the most general model (Amemiya [1974a and b]).

2. UNIVARIATE CASE

Models for a dichotomous and a trichotomous variable are presented in turn. The analysis of a trichotomous variable necessitates the adoption of one

*This work was supported by National Science Foundation Grant GS-39906 at the Institute for Mathematical Studies in the Social Sciences, Stanford University. I am grateful to Fred Nole for improving the style of the paper. The earlier version of the paper, entitled "A Note on the Regression Analysis of Polychotomous Dependent Variables," was reported at the NSF-NBER Conference on Decision Rules and Uncertainty on March 23, 1974 at the University of California, Berkeley.

of several types of models and raises many interesting problems. I will not discuss a general polytomous model because all the results for a trichotomous variable can be easily extended to the polytomous case and no new problems emerge.

A. Dichotomous Variable

The dichotomous random variable y_t , $t = 1, 2, \dots, T$, takes the values 0 and 1 with the probabilities determined by

$$(2.1) \quad P(y_t = 1) = F(\beta'x_t)$$

where x_t is a vector of known constants and β is a vector of unknown parameters.

The most commonly used forms of function F are

$$(2.2) \quad \text{Normal } \Phi(\beta'x_t)$$

where Φ is the standard normal distribution function and

$$(2.3) \quad \text{Logistic } L(\beta'x_t)$$

where $L(x) = [1 + \exp(-x)]^{-1}$. These two distributions have been successfully used in many empirical applications. For theoretical and empirical reasons for using these functions and their relative merits, see Berkson [1951], Cox [1966], and Finney [1971]. The logistic is a good approximation to the normal distribution, and the estimates of β obtained by using the two distributions are often very close except for a multiplicative factor. A full discussion of the properties of the logistic distribution can be found in Johnson and Kotz [1970].

The linear function $F(\beta'x_t) = \beta'x_t$ has also been frequently used, especially in economic applications. Its major deficiency is that its range, unlike the normal or logistic transformation, is not constrained to lie between 0 and 1. This and other difficulties encountered when using the linear probability model are pointed out by Goldberger [1964] and Nerlove and Press [1973]. In this paper I will confine my attention to the normal and logistic models.

Example 1. When the dosage x_t of an insecticide is given to the t -th insect, it dies ($y_t = 1$) if its tolerance u_t is less than x_t . If one assumes that u_t is distributed as normal with mean μ and variance σ^2 , one has $P(y_t = 1) = \Phi[(x_t - \mu)/\sigma]$, which is model (2.2). If one assumes that u_t is distributed as logistic with mean μ and variance σ^2 , one has $P(y_t = 1) = L[\pi(x_t - \mu)/\sigma\sqrt{3}]$, which is model (2.3).

Example 2. A coal miner develops breathlessness ($y_t = 1$) when his tolerance u_t is less than an unknown constant γ . If one assumes that u_t is distributed as normal or logistic with mean $\alpha_0 + \alpha_1x_t$ and variance σ^2 where x_t is the coal miner's age, then one again has model (2.2) or (2.3).

Example 3. A consumer buys a car ($y_t = 1$) when his net utility u_t of buying a car exceeds 0. As one assumes that u_t is distributed as normal or logistic with mean $\alpha_0 + \alpha_1x_t$ and variance σ^2 , where x_t is the consumer's income, model (2.2) or (2.3) results.

B. Trichotomous Variable

The following notation is henceforth used:

$$P_i = P[y_t = i], \quad i = 0, 1, 2.$$

1. Ordered Response

The ordered normal model is defined as

$$(2.4) \quad \begin{cases} P_2 = \Phi(\beta'x) \\ P_2 + P_1 = \Phi(\beta'x + \alpha) \\ P_0 = 1 - \Phi(\beta'x + \alpha), \end{cases}$$

where α is a positive constant. Note that the subscript t has been suppressed from x in the above.

The ordered logistic model is similarly defined as

$$(2.5) \quad \begin{cases} P_2 = L(\beta'x) \\ P_2 + P_1 = L(\beta'x + \alpha) \\ P_0 = L(-\beta'x - \alpha). \end{cases}$$

Example 4. (Gurland, Lee, and Dahm [1960]). When the dosage x_t of an insecticide is given to the t -th insect, it dies ($y_t = 2$) if its tolerance $u_t < x_t$ and it becomes moribund ($y_t = 1$) if $u_t - \gamma < x_t$. The assumption that u_t is distributed as normal with mean μ and variance σ^2 leads to model (2.4) and the assumption that u_t is distributed as logistic leads to model (2.5).

Other examples of univariate polytomous models with ordered response may be found in Aitchison and Silvey [1957] and Ashford [1959].

2. Unordered Response

I will mention three types of models for this situation. One can obviously think of other models.

First, assume that given any pair of responses the selection is made according to model (2.3). Then we have

$$(2.6) \quad P_1(P_1 + P_0)^{-1} = L(\beta'_1x)$$

$$(2.7) \quad P_2(P_2 + P_0)^{-1} = L(\beta'_2x)$$

$$(2.8) \quad P_1(P_1 + P_2)^{-1} = L(\beta'_3x).$$

Probabilities P_0, P_1, P_2 are uniquely determined from (2.6) and (2.7) so that one must have $\beta_3 = \beta_1 - \beta_2$. This leads to the unordered logistic model

$$(2.9) \quad \begin{cases} P_2 = D^{-1} \exp(\beta'_2x) \\ P_1 = D^{-1} \exp(\beta'_1x) \\ P_0 = D^{-1} \end{cases}$$

where $D = 1 + \exp(\beta'_1x) + \exp(\beta'_2x)$. This model was suggested by Cox [1966] and a similar model was applied to explaining the selection of highway routes in California by McFadden [1968].

The second model assumes that selection is made sequentially. That is, one first determines whether $y_i = 2$ or $\neq 2$, and then, given $y_i \neq 2$, one determines whether $y_i = 1$ or 0 . Using normal model (2.2) for each selection leads to the following sequential unordered normal model

$$(2.10) \quad \begin{cases} P_2 = \Phi(\beta'_2 x) \\ P_1 = [1 - \Phi(\beta'_2 x)]\Phi(\beta'_1 x) \\ P_0 = [1 - \Phi(\beta'_2 x)][1 - \Phi(\beta'_1 x)]. \end{cases}$$

The sequential unordered logistic model can be similarly defined. An advantage of sequential models is that the likelihood function can be maximized by maximizing the likelihood of the dichotomous case repeatedly.

Computational ease is a major consideration in defining the preceding two models. However, one may specify a model solely on the basis of theoretical consideration of the behavior of the respondent. For example, if the responses are the outcome of the free choice of the respondent, it is natural to assume that $P(y_i = i) = P(U_i(x) > U_j(x), j \neq i)$ where U_i , given a vector of exogenous variables x , is the random utility associated with the i -th choice. Such a model has been proposed by Quandt [1968] and Aitchison and Bennett [1970]. The difficulty of this approach is that if a realistic distribution of U_i is assumed the estimation problem becomes intractable, whereas if a convenient distribution is chosen for U_i the model becomes as arbitrary as the preceding two models. In fact, McFadden [1974] has shown that model (2.9) follows from the maximum utility model if $\{U_i\}$ are assumed to be independent each following the double exponential distribution.

C. Partition of a Probability Space

It is useful to associate the values taken by a polytomous variable with the partition of a probability space. That is, one can define

$$(2.11) \quad P(y_i = k) = P(S_k)$$

where $\cup_k S_k$ may be taken as a subset of Euclidean space with P generated by a set of random variables. All the models considered thus far may be interpreted in this way.

One can characterize the ordered models simply as the case where the basic probability space is the real line. If one uses an ordered model when the true model is unordered, one could get into serious trouble, whereas the loss in using an unordered model in the ordered situation is only in efficiency since consistent estimates are still obtained.

3. MULTIVARIATE CASE

There are two essentially different ways to specify multivariate models:
 (A) First specify marginal distributions and then specify joint distributions;
 (B) Specify conditional distributions.

A. Marginal Distribution

Following the discussion of Section 2C, let us assume that random variables y_{1t} and y_{2t} are distributed according to

$$(3.1) \quad P(y_{it} = k) = P(S_k^i), \quad i = 1, 2.$$

Then one can define

$$(3.2) \quad P(y_{1t} = k, y_{2t} = h) = P(S_k^1 \otimes S_h^2)$$

where the probability on the product space should be appropriately chosen. Note that (3.1) does not uniquely determine the probability on the product space. One can find a natural way to extend marginal probabilities to a joint probability if each marginal probability is generated by a multivariate normal distribution, but the extension is in general difficult with other distributions. In the univariate case logistic and normal models give similar results and therefore logistic models may be preferred because they require simpler computation. However, in the multivariate case normal models have an advantage because the multivariate normal distribution has nice properties, whereas there is no simple bivariate logistic distribution with a correlation coefficient which can freely vary. (See Gumbel [1961].)

The dichotomous normal model (2.2) or the ordered normal model (2.4) can be easily extended to the multivariate case in the way indicated by (3.2).

Example 5. (Ashford and Sowden [1970].) Let y_i be as defined in Example 2 with u_i normally distributed. Similarly assume that a coal miner develops wheeze ($z_i = 1$) when his tolerance v_i against wheeze is less than a constant δ where v_i is distributed normally with mean $\beta_0 + \beta_1 x_i$ and a constant variance. Then one can naturally specify the joint probabilities of y and z by assuming that u and v have a bivariate normal distribution with correlation ρ .

The sequential unordered normal model (2.10) may be extended to the bivariate case as follows. For each $i = 1, 2$, rewrite (2.10) as

$$(3.3) \quad \begin{cases} P(y_i = 2) = P(u_i < \beta'_{i2}x) \\ P(y_i = 1) = P(u_i > \beta'_{i2}x, v_i < \beta'_{i1}x) \\ P(y_i = 0) = P(u_i > \beta'_{i2}x, v_i > \beta'_{i1}x) \end{cases}$$

where I have suppressed subscript t and assumed that for each i , u_i and v_i are mutually independent and distributed as standard normal. One simple way to extend this to the bivariate case is to assume $Eu_i u_2 = \rho_u$ and $Ev_i v_2 = \rho_v$ where ρ_u and ρ_v are parameters to be estimated.

The log likelihood function of this model is given by

$$(3.4) \quad \begin{aligned} \log L = \sum_t \{ & y_{22} \log F_u + (y_{21} + y_{20}) \log [\Phi(\beta'_{12}x) - F_u] \\ & + (y_{12} + y_{02}) \log [\Phi(\beta'_{22}x) - F_u] \\ & + (y_{11} + y_{10} + y_{01} + y_{00}) \log [1 - \Phi(\beta'_{12}x) - \Phi(\beta'_{22}x) + F_u] \\ & + y_{11} \log F_v + y_{10} \log [\Phi(\beta'_{11}x) - F_v] + y_{01} \log [\Phi(\beta'_{21}x) - F_v] \} \end{aligned}$$

$$\begin{aligned}
& + y_{00} \log [1 - \Phi(\beta'_{11}x) - \Phi(\beta'_{21}x) + F_u] + y_{12} \log \Phi(\beta'_{11}x) \\
& + y_{02} \log [1 - \Phi(\beta'_{11}x)] + y_{21} \log \Phi(\beta'_{21}x) \\
& + y_{20} \log [1 - \Phi(\beta'_{21}x)]
\end{aligned}$$

where $F_u = F(\beta'_{12}x, \beta'_{22}x, \rho_u)$, F being the bivariate standard normal distribution function, $F_v = F(\beta'_{11}x, \beta'_{21}x, \rho_v)$, and

$$\begin{aligned}
y_{kh} &= 1 && \text{if } y_1 = k \text{ and } y_2 = h \\
&= 0 && \text{otherwise.}
\end{aligned}$$

Thus we know that so far as β_{12} , β_{22} and ρ_u are concerned, the maximization of $\log L$ above is identical with the maximum likelihood estimation of the bivariate dichotomous normal model (see Example 5). But for the estimation of β_{11} and β_{21} , the last four terms constitute the additional terms that contain information and must be taken into account if fully efficient estimates are to be obtained. In the minimum chi-square approach (Amemiya [1972]), the last four terms can be taken into account by estimating $P(y_1 = 1)$ with

$$\frac{\sum_t (y_{12} + y_{11} + y_{10})}{\sum_t (y_{12} + y_{11} + y_{10} + y_{02} + y_{01} + y_{00})}$$

and a similar equation for $P(y_2 = 1)$.

B. Conditional Distribution

In the univariate case there was no essential difference between Example 2 of a coal miner developing breathlessness and Example 3 of a consumer buying a car. However, a bivariate model of a consumer buying a car and a house cannot be specified in the same way as Example 5 of a coal miner developing breathlessness and wheeze. The reason is that the probability of buying a car depends on whether or not the consumer buys a house because the latter decision changes the level of the independent variable "income," whereas breathlessness or wheeze does not directly affect each other as neither affects the independent variable "age". Therefore, in the former case it is better to specify conditional probabilities and in the latter case marginal probabilities.

Example 6. Let $y_1 = 1$ if a consumer buys a car and $y_2 = 1$ if he buys a house. We assume that the conditional probabilities are given by

$$(3.5) \quad P(y_1 = 1|y_2) = L(\beta'_1x + \beta_{12}y_2)$$

and

$$(3.6) \quad P(y_2 = 1|y_1) = L(\beta'_2x + \beta_{21}y_1)$$

where β'_1x and β'_2x are linear functions of the consumer's income. Define

$$P_{ij} = P(y_1 = i, y_2 = j).$$

Then, from (3.5) we have

$$(3.7) \quad P_{11} = \exp(\beta'_1x + \beta_{12})P_{01}$$

$$(3.8) \quad P_{10} = \exp(\beta'_1 x) P_{00}$$

and from (3.6) we have

$$(3.9) \quad P_{11} = \exp(\beta'_2 x + \beta_{21}) P_{10}$$

$$(3.10) \quad P_{01} = \exp(\beta'_2 x) P_{00}$$

We can evaluate P_{11}/P_{00} either from (3.7) and (3.10) or from (3.8) and (3.9). Therefore, we must have $\beta_{12} = \beta_{21}$.¹ Thus, we can write the joint probability as

$$(3.11) \quad P(y_1, y_2) = D^{-1} \exp(\beta'_1 x y_1 + \beta'_2 x y_2 + \beta_{12} y_1 y_2)$$

where

$$D = 1 + \exp(\beta'_1 x) + \exp(\beta'_2 x) + \exp(\beta'_1 x + \beta'_2 x + \beta_{12}).$$

Example 6 can be generalized to the case of three dichotomous variables as follows. We have

$$(3.12) \quad \begin{cases} P(y_1 = 1 | y_2, y_3) = L(\beta'_1 x + \beta_{12} y_2 + \beta_{13} y_3 + \beta_{123} y_2 y_3) \\ P(y_2 = 1 | y_1, y_3) = L(\beta'_2 x + \beta_{21} y_1 + \beta_{23} y_3 + \beta_{213} y_1 y_3) \\ P(y_3 = 1 | y_1, y_2) = L(\beta'_3 x + \beta_{31} y_1 + \beta_{32} y_2 + \beta_{312} y_1 y_2) \end{cases}$$

Putting each of y 's equal to 0 in turn in the above and using the results of Example 6 we have

$$(3.13) \quad \beta_{12} = \beta_{21}, \quad \beta_{13} = \beta_{31}, \quad \beta_{23} = \beta_{32}.$$

Define $P_{ijh} = P(y_1 = i, y_2 = j, y_3 = h)$. Then we have

$$(3.14) \quad \frac{P_{111}}{P_{011}} \cdot \frac{P_{011}}{P_{000}} = \frac{P_{111}}{P_{101}} \cdot \frac{P_{101}}{P_{000}} = \frac{P_{111}}{P_{110}} \cdot \frac{P_{110}}{P_{000}}$$

where each of the six probability ratios can be written as an exponential function using (3.12). But, because of (3.13), the three terms in (3.14) differ only in the β coefficient with three subscripts. Therefore (3.14) implies

$$(3.15) \quad \beta_{123} = \beta_{213} = \beta_{312}.$$

Now, it is easy to show that the joint probability

$$(3.16) \quad P(y_1, y_2, y_3) = D^{-1} \exp(\beta'_1 x y_1 + \beta'_2 x y_2 + \beta'_3 x y_3 + \beta_{12} y_1 y_2 + \beta_{13} y_1 y_3 + \beta_{23} y_2 y_3 + \beta_{123} y_1 y_2 y_3),$$

where D is chosen so as to make the sum of probabilities to add up to 1, implies the conditional probability (3.12) with the constraints (3.13) and (3.15). But, since

¹ After this paper was written, I was made aware of a paper by Schmidt and Strauss [1974] which notes this equality.

the conditional probabilities determine at most one set of joint probabilities,² (3.12) implies (3.16).

The generalization to the case of J dichotomous variables each taking values 0 and 1 can be easily inferred from above. Thus, the conditional probabilities

$$(3.17) \quad P(y_j = 1 | y_i, i \neq j) = L \left(\beta'_j x + \sum_{i \neq j} \beta_{ji} y_i + \sum_{i \neq j} \sum_{h \neq j} \beta_{jih} y_i y_h + \dots \right. \\ \left. + \beta_{j,1,\dots,j-1,j+1,\dots,j} y_1 \dots y_{j-1} y_{j+1} \dots y_j \right)$$

imply the joint probabilities

$$(3.18) \quad P(y_1, y_2, \dots, y_j) = D^{-1} \exp \left[\sum_j \beta'_j x y_j + \sum_{j < i} \beta_{ji} y_j y_i + \right. \\ \left. \sum_{j < i < h} \beta_{jih} y_j y_i y_h + \dots + \beta_{12,\dots,j} y_1 y_2 \dots y_j \right],$$

where D is chosen so as to make the sum of the probabilities equal to 1.

The further generalization to the case of J polytomous variables is simple as one can describe an n -response variable by $n - 1$ dichotomous variables. Let y^i be the i -th polytomous variable taking values $1, 2, \dots, n_i$ and define the dichotomous variable $y_j^i, j = 1, 2, \dots, n_i - 1$, by $y_j^i = 1$ if $y^i = j$ and 0 otherwise. Then, the result of the preceding paragraph holds by treating y_j^i as one of the y 's as long as we add the restriction that (3.17) and (3.18) hold only when only one of y_j^i is equal to 1 for every i .

The joint probabilities of the form of (3.18) without the x variable were considered by Goodman in a series of papers. For example, see Goodman [1970]. Mantel [1966] and Cox [1972] provide pedagogical discussions of the same subject. Nerlove and Press [1973] were the first to consider the inclusion of the x variable in the probabilities. These authors have shown that (3.18) implies (3.17) with the constraint corresponding to (3.13) and (3.15). I have indicated in this paper that (3.17) implies the constraint and hence (3.18). Equation (3.18) contains as many β parameters (counting $\beta'_j x$ as one) as the number of probabilities minus 1. One may wish to put certain of these parameters equal to 0 *a priori* to obtain more efficient estimators of the remaining parameters.

In case (A) the maximization of each marginal probability will give consistent estimators of the regression parameters which are computationally simpler than the full maximum likelihood estimators. Similarly, in case (B) the maximization of each conditional probability will give computationally simpler consistent estimates. I will give a heuristic proof of this below.

² Let $y_j, j = 1, 2, \dots, J$, be a discrete random variable taking values $1, 2, \dots, N_j$. Suppose the conditional probability $P(y_j | y_1, \dots, y_{j-1}, y_{j+1}, \dots, y_J)$ is given for every j and is positive. Then there is at most one set of joint probabilities consistent with the given conditional probabilities. This can be proved as follows: Let $J = 2$ and define $P_{ij} = P(y_1 = i, y_2 = j)$. Then, given P_{11} , conditional probabilities at most uniquely determine $P_{ij} = \alpha_{ij} P_{11}$. But there is only one value of P_{11} which satisfies $\sum P_{ij} = 1$. Next, suppose the proposition is true for $J - 1$. Define the composite random variable y which takes $\prod_{j=1}^{J-1} N_j$ number of values with probabilities determined by the joint probabilities of y_1, y_2, \dots, y_{j-1} . Then, by the assumption, $P(y | y_j)$ and $P(y_j | y)$ are at most uniquely determined. Hence the desired result, using the result for $J = 2$.

For simplicity consider the bivariate dichotomous case. Suppose one maximizes

$$(3.19) \quad L = \frac{1}{T} \sum_i [y_i \log P + (1 - y_i) \log (1 - P)]$$

with respect to θ where P depends on θ and y_2 . Write $L = L_0 + L_1$ where

$$L_0 = \frac{1}{T} \sum_0 [y_i \log P_0 + (1 - y_i) \log (1 - P_0)]$$

$$L_1 = \frac{1}{T} \sum_1 [y_i \log P_1 + (1 - y_i) \log (1 - P_1)]$$

where subscripts 0 and 1 indicate whether y_2 is 0 or 1. Therefore,

$$(3.20) \quad \begin{aligned} \text{plim } L &= \text{plim } L_0 + \text{plim } L_1 \\ &= P^*(y_2 = 0)[P_0^* \log P_0 + (1 - P_0^*) \log (1 - P_0)] \\ &\quad + P^*(y_2 = 1)[P_1^* \log P_1 + (1 - P_1^*) \log (1 - P_1)] \end{aligned}$$

where superscript * indicates that the probability is evaluated at the true value θ^* of θ . Thus, clearly the above is maximized at $\theta = \theta^*$, implying the consistency of the estimator. Its asymptotic variance-covariance matrix is given by

$$(3.21) \quad \left\{ -E \Sigma \left[\frac{1}{P(1-P)} \frac{\partial P}{\partial \theta} \frac{\partial P}{\partial \theta'} \right] \right\}^{-1}$$

where the expectation is taken with respect to y_2 .

4. CONCLUSIONS

In this paper I have attempted to give a systematic discussion of various qualitative response models, with a special emphasis on multivariate models. I have defined some new models (notably, multivariate polytomous probit models) and give new interpretations on old models. I have contrasted two basically differing ways to formulate multivariate models: the one that specifies marginal probabilities first and the one that specifies conditional probabilities first. Depending upon the nature of the problem at hand, one model is more appropriate than the other. An interesting topic of research seems to be to investigate how good one model performs when the other is the true model. For each model one can define a simplified maximum likelihood estimator maximizing a marginal probability or a conditional probability as the case may be. Thus, another interesting problem is to compare the efficiency of the simplified estimator with the full maximum likelihood estimator in each model. One needs also to investigate the computational feasibility of the maximum likelihood estimator in multivariate probit models when the number of dependent variables exceeds 2. If multivariate logistic distributions with convenient properties cannot be found, one may have to search for another useful approximation to the multivariate normal distribution.

It is possible that the Nerlove and Press model, setting aside its intrinsic merits, is useful even when it is merely regarded as a substitute for the multivariate probit model.

Stanford University

REFERENCES

- Aitchison, J. and J. Bennett (1970), "Polychotomous Quantal Response by Maximum Indicant," *Biometrika* 57, 253-262.
- Aitchison, J. and S. D. Silvey (1957), "The Generalization of Probit Analysis to the Case of Multiple Responses," *Biometrika* 44, 131-140.
- Amemiya, T. (1972), "Bivariate Probit Analysis: Minimum Chi-Square Methods," Technical Report No. 76, Institute for Mathematical Studies in the Social Sciences, Stanford University (to appear in the *Journal of the American Statistical Association*).
- Amemiya, T. (1974a), "The Maximum Likelihood Estimator vs. the Minimum Chi-Square Estimator in the General Qualitative Response Model," Technical Report No. 136, Institute for Mathematical Studies in the Social Sciences, Stanford University.
- Amemiya, T. (1974b), "The Equivalence of the Nonlinear Weighted Least Squares Method and the Method of Scoring in the General Qualitative Response Model," Technical Report No. 137, Institute for Mathematical Studies in the Social Sciences, Stanford University.
- Ashford, J. R. (1959), "An Approach to the Analysis of Data for Semi-Quantal Responses in Biological Assay," *Biometrics* 15, 573-581.
- Ashford, J. R. and R. S. Sowden (1970), "Multivariate Probit Analysis," *Biometrics* 26, 535-546.
- Berkson, J. (1951), "Why I Prefer Logits to Probits," *Biometrics* 7, 327-339.
- Cox, D. R. (1966), "Some Procedures Connected With the Logistic Qualitative Response Curve," *Research Papers in Statistics*, ed. by F. N. David, John Wiley, London, 55-71.
- Cox, D. R. (1970), *Analysis of Binary Data*, Methuen, London.
- Cox, D. R. (1972), "The Analysis of Multivariate Binary Data," *Applied Statistics* 21, 113-120.
- Finney, D. J. (1971), *Probit Analysis*, Third Edition, University Press, Cambridge.
- Goldberger, A. S. (1964), *Econometric Theory*, John Wiley, New York.
- Goodman, L. A. (1970), "The Multivariate Analysis of Qualitative Data: Interactions Among Multiple Classifications," *Journal of the American Statistical Association* 65, 226-256.
- Gumbel, E. J. (1961), "Bivariate Logistic Distributions," *Journal of the American Statistical Association* 56, 335-349.
- Gurland, J., I. Lee, and P. A. Dahm (1960), "Polychotomous Quantal Response in Biological Assay," *Biometrics* 16, 382-398.
- Johnson, N. L. and S. Kotz (1970), *Continuous Univariate Distributions* 2, Houghton Mifflin, Boston.
- Mantel, N. (1966), "Models for Complex Contingency Tables and Polychotomous Response Curves," *Biometrics* 22, 83-110.
- McFadden, D. (1968), "The Revealed Preferences of a Government Bureaucracy," Department of Economics, University of California at Berkeley.
- McFadden, D. (1974), "Conditional Logit Analysis of Qualitative Choice Behavior," *Frontiers in Econometrics*, ed. by P. Zarembka, Academic Press, New York.
- Nerlove, M. and S. J. Press (1973), "Univariate and Multivariate Log-Linear and Logistic Models," R-1306-EDA/NIH, Rand Corporation, Santa Monica.
- Schmidt, P. and R. P. Strauss (1974), "Estimation of Models with Jointly Dependent Qualitative Variables: A Simultaneous Logit Approach," mimeographed.
- Quandt, R. (1968), "Estimation of Modal Splits," *Transportation Research* 2, 41-50.