Staff Papers Series

Staff Paper P77-12 May 1977

Discrete Dependent Variables: A Guide to Alternative Estimating Techniques With An Annotative Bibliography of Logit Analysis

Jean Kinsey

Department of Agricultural and Applied Economics

University of Minnesota
Institute of Agriculture, Forestry and Home Economics
St. Paul, Minnesota 55108

Discrete Dependent Variables: A Guide to Alternative Estimating Techniques With An Annotative Bibliography of Logit Analysis

by
Jean Kinsey

May 1977

Discrete Dependent Variables: A Guide to Alternative Estimating Techniques With An Annotative Bibliography of Logit Analysis

Introduction

With increasing frequency economists are being confronted with regression models which have discrete dependent variables e.g., the choice of occupation, college or travel mode. Ordinary least squares is not an appropriate estimating technique for these probability models because the error term is neither normally distributed nor does it have a constant variance. In addition, the ordinary least squares regression line cannot be confined to values consistent with measures of probability, that is, to values between zero and one. Consequently, several alternative estimating techniques have been developed. This paper briefly outlines the most well known of these techniques so that the reader may determine quickly which is the most likely to be appropriate for his/her model.

The limiting case of a discrete dependent variable is a binary (0, 1) dependent variable. In this case the problem is to estimate the probability that an observation belongs to one group or another or that a person will choose one option or another. These types of models are known as choice models or qualitative response models. The first section of this paper presents a general choice model in the form of a linear probability function, which is subsequently transformed into the appropriate specifications for discriminate, logit, probit and tobit analysis. The basic assumptions of each type of analysis is briefly outlined.

If the researcher determines that logit analysis is the most appropriate estimating technique for his/her model, the annotative bibliography in section two of this paper will provide him/her with a fairly broad (though not complete) selection of readings on the theoretical development and empirical applications of logit analysis. The bibliography is listed in alphabetical order by author. There is a separate index at the end which sorts the articles according to their primary emphasis, theoretical or empirical.

Section I

Regression With Discrete Dependent Variable(s)

Also known as:

- Choice Models
- Qualitative Response Models

The estimating technique you choose depends on:

1. The type of question being asked.
2. (In)dependence of alternatives in the dependent variable.
3. Reasonable assumptions on the distribution of explanatory variables and/or the error term.
4. The form and nature of the data.
5. Availability of computer programs for each technique.*

Begin with a linear probability function

(1) $Y = X' \beta + \varepsilon$ where:

Y is a vector of discrete observations

$$Y = \{y_1, \ldots y_i \ldots y_n\}$$

---

*The University of Minnesota has computer programs for all of the estimating techniques discussed herein. The logit program is called QUAIL and was developed at Berkeley. The probit program, which handles only a binary dependent variable and has no special name was developed at the University of Rochester. The tobit program is called LIMDEP and was developed at RAND.

X is a matrix of observations on explanatory variables. In a conditional probability model these variables are taken as given and are usually considered to be nonstochastic.

β is a vector of unknown parameters

ϵ is a vector of stochastic disturbance terms

The problem is to estimate the conditional probability that the observed values of X belong with a particular value of Y, i.e., y_i. ($y_i = 0$, 1 for the binary case, $y_i = 0, 1, 2, \ldots n$ for the multinomial case.) In other words, the expected value of Y, given the observations on X is equal to some function of $X' \beta$.

$$(2) \qquad E(Y \mid X) = F(X' \beta)$$

The probability that the expected value of Y is one when $Y = \{y_0, y_1\} = (0, 1)$ is:

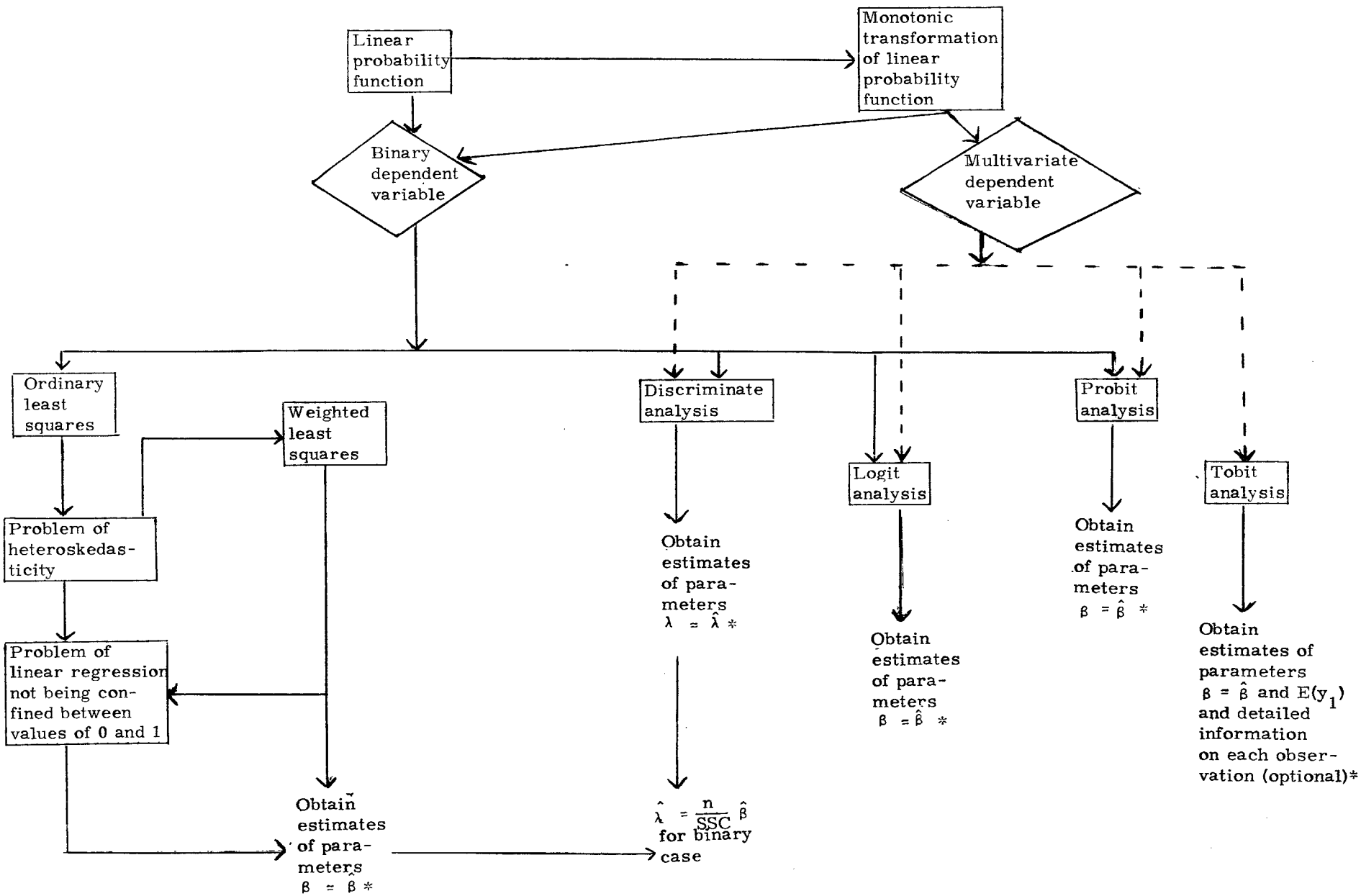
$$(3) \qquad P(Y = 1 \mid X) = F(X' \beta) = \beta_0 + \beta_1 x_{11} + \ldots + \beta_k X_{k1} + \epsilon_1$$

for a linear probability function which is estimated by ordinary or weighted least squares.

Discriminate analysis, logit, probit, and tobit are monotonic transformations of the basic linear probability function (equation 3).

A diagram of alternative estimating techniques is presented below. Each technique is discussed on the following pages with respect to the question each technique addresses, the mathematical form of the transformed linear probability function, the form of the dependent variable and the assumptions about the distribution of the explanatory variables.

Some Alternative Regression Estimating Techniques



*Summary statistics such as the mean and standard deviation of the explanatory variables, measures of the goodness of fit and standard errors of estimates are provided in varying forms by programs for each analytic technique.

Linear Probability Function

Estimated by ordinary or weighted least squares.

The question being asked (correctly or not) with these models is, "What is the probability that an individual will fall into one group or the other given his/her observed characteristics?"

For example: "What is the probability that a household with an observable set of socioeconomic characteristics and stock of durable goods will or will not purchase a new automobile in the current year?"

Or, "What is the probability that a given brand or breed of grass seed planted under a given set of environmental circumstances will grow to a certain height in a certain time period?"

The mathematical specification is:

$E(Y \mid X) = F(X' \beta)$ where $F(X' \beta)$ is assumed to be a linear function.

(4) $P_1 = P(Y = y_1 \mid X) = Y = X'\beta + \varepsilon$. (See equation 3)

where $Y = \{y_0, y_1\} = (0, 1)$

When the dependent variable is dichotomous heteroskedasticity occurs; the variance of the estimated ε_i, ($\hat{\varepsilon}_i$), depends on the values of X. Var $\hat{\varepsilon}_i = X_i' \beta \ (1 - X_i' \beta) = Ey_i \ (1 - Ey_i)$ and ε_i is not normally distributed. The ramifications of these violations of ordinary least squares assumptions is that the standard errors of the estimates of β are not consistent and the usual tests of significance cannot be performed.

Weighted least squares may be used to correct for heteroskedasticity by reestimating the model using $\hat{y}_i^* (1 - \hat{y}_i^*)$ as the diagonal elements in the variance-covariance matrix where $\hat{y}_i^*$ is the estimated value of y_i using ordinary least squares.

However, weighted least squares does not solve the problem of the regression line extending beyond the boundaries of 0 and 1.

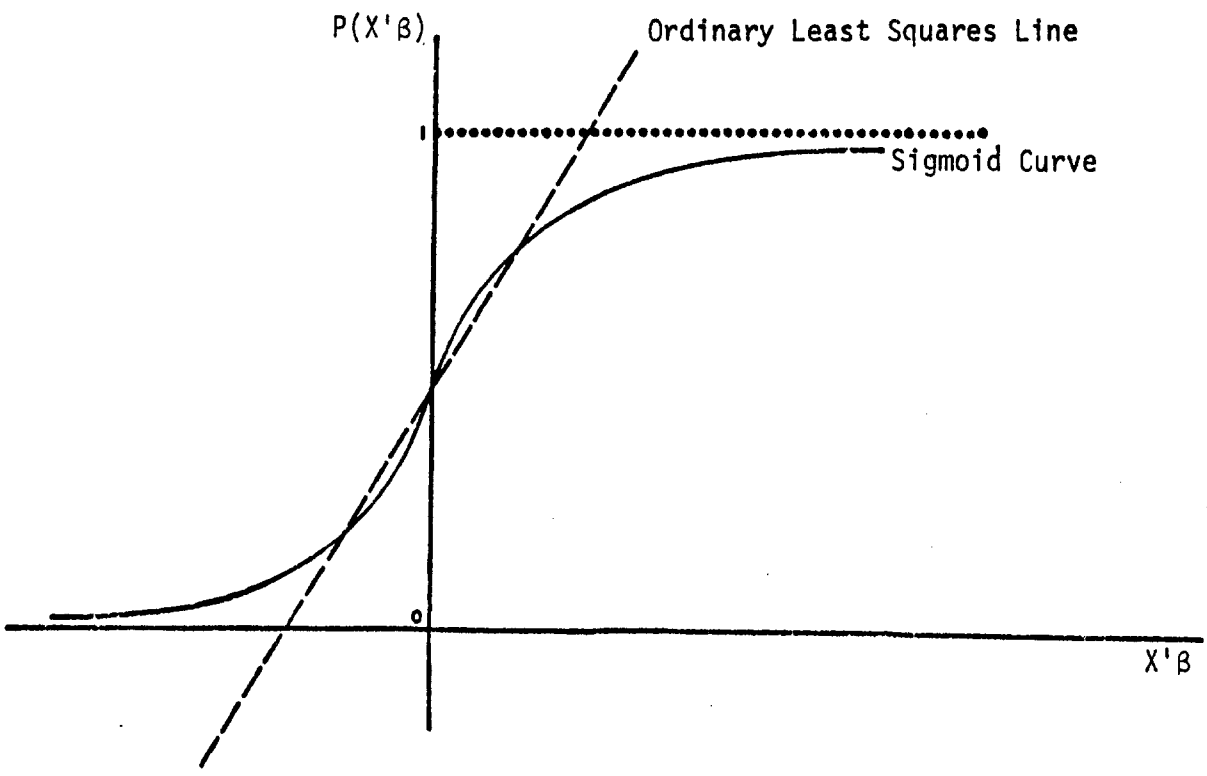


Figure 1. Comparison of Linear and Logistic Regression

The following estimating techniques were developed to cope with the limitations of ordinary or weighted least squares models. The appropriateness of each technique depends upon the type of question being asked and the form of the data. As a practical matter, the results are not expected to vary a great deal regardless of which technique is used.[1/]

Discriminate Analysis

The question being asked is "Which group does the subject belong to given its characteristics?" The purpose is to minimize the errors in classification. Group means (averages) rather than individual observations are the basis of the computations.

For example: Given a number of households with observable socioeconomic characteristics including the stock of debts and assets, which of these households are most likely to belong in each of the following categories: nondebtors, average debtors, overindebted, bankrupt.

The mathematical specification is:

(5) $Y = \lambda' X$ when $Y = \{y_0, \ldots y_i \ldots y_n\}$

(6) $y_i = \lambda_0 + \lambda_1 X_{1i} + \cdots \lambda_n X_{ni}$

$\lambda' X$ = the discriminate function where:

(7) $\lambda' = (\mu_j - \mu_i)^1 \Sigma^{-1}$ and

[1/] See bibliography in Section II, references by Aldrich and Cruddle; Brown, Moon and Zoloth; Kinsey.

μ_i = mean value of X for those in group i

Σ = variance-covariance matrix

If the discriminate function exceeds a critical value it is classified as being in group one, otherwise it is classified as being in group two.

It can be shown that estimates of λ, ($\hat{\lambda}$), are equal to $\frac{n}{SSE}\hat{\beta}$ where $\hat{\beta}$ is the ordinary least squares estimate of the coefficient of the linear probability function; n is the number of observations and SSE is the residual sum of squares from the ordinary least squares estimate, if the assumptions stated below hold.

$F_i(X'\beta)$ is assumed to have a multivariate normal distribution within each group and each group is assumed to have the same variance-covariance matrix; $\Sigma_i = \Sigma_j = \Sigma$. "Group" is identified as those observations on X which belong with each category of the dependent variable.

Logit Analysis

The question being asked is "What is the conditional probability that a randomly selected individual chooses one alternative or another, given the observed and unobserved characteristics of the individual and attributes of the alternative choice?"

For example: What is the probability that a household with an observable set of socioeconomic characteristics will choose one mode of transportation over another mode given the attributes of the two modes?

What in fact is estimated and reported is not the probability, but the coefficients of the explanatory variables which are associated with the

maximum likelihood log odds (equation 9) of the probability of one alternative being chosen over a second alternative. With this information, probabilities may be calculated.

Estimating the log odds implies a restrictive assumption which is built into logit programs, namely the axiom of the independence of irrelevant alternatives. This assumption precludes alternative choices in the dependent variable which are close substitutes for each other. The alternative choices should be unranked and independent.

The mathematical specification is:

$E(Y \mid y) = F(X'\beta)$ where $F(X'\beta)$ is a logistic function of the form $\frac{1}{1 + e^{-X'\beta}}$

The logistic transformation is the odds ratio

$\left(\frac{P_1}{1 - P_1}\right)$ where: $P_1 = Pr(Y=1)$ where $Y = \{0, 1\}$ or $P_i = Pr(Y=y_i)$ where $y_i = \{0, 1, \ldots N\}$

$$(8) \qquad (\text{Logit } P_i) \equiv \frac{1}{1 + e^{-X\beta}} = \frac{1}{1 + e^{(\beta_0 + \beta_1 X_{1i} + \ldots + \beta_n X_{ni})}} = \frac{P_i}{1-P_i}$$

$$(9) \qquad (\text{Logit } P_i) \equiv \log \left(\frac{P_i}{1 - P_i}\right) = X'\beta$$

X's are generally considered to be observations on nonstochastic variables which are independent of each other. If maximum likelihood estimates are used, both continuous and discrete explanatory variables may be used. If estimates are obtained by a minimum Chi Square method, the data must be grouped.

When using logit analysis to estimate utility functions, [$U = V(X, Y) + \varepsilon (X, Y)$] , the ε_i's (stochastic deviations from mean utilities) are assumed to be independently and identically distributed Weibull.

$$h(\varepsilon) = [e^{-\varepsilon_i - \varepsilon_i}]$$

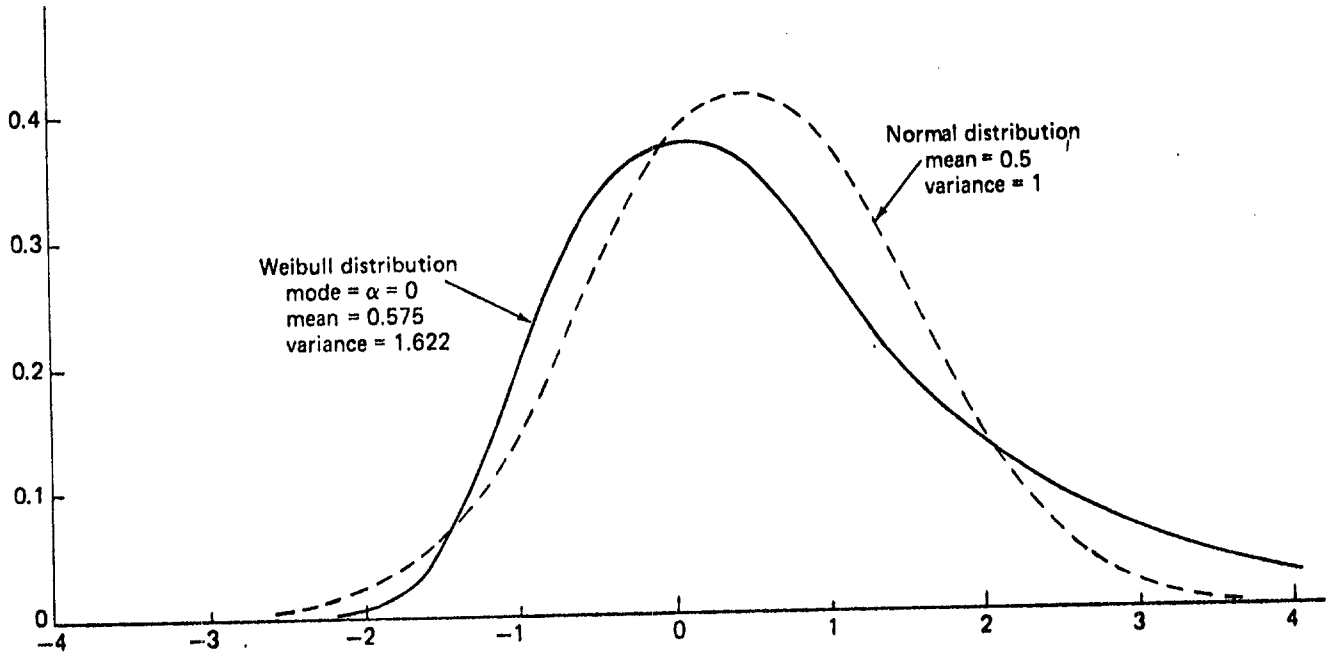


Figure 2. Comparison of Weibull and Normal Distributions (From Domencich and McFadden, Urban Travel Demand, North-Holland Publishing Co., 1975)

In some cases the error terms (unobserved characteristics) in logistic probability models are assumed to have a gamma distribution with mean =1, and variance = V.

$$h(\varepsilon) = \frac{V^{-Y_V}}{(\frac{1}{V})} \quad \varepsilon^{(Y_V - 1)} \quad e^{-\varepsilon/V} \qquad \text{for} \qquad 0$$

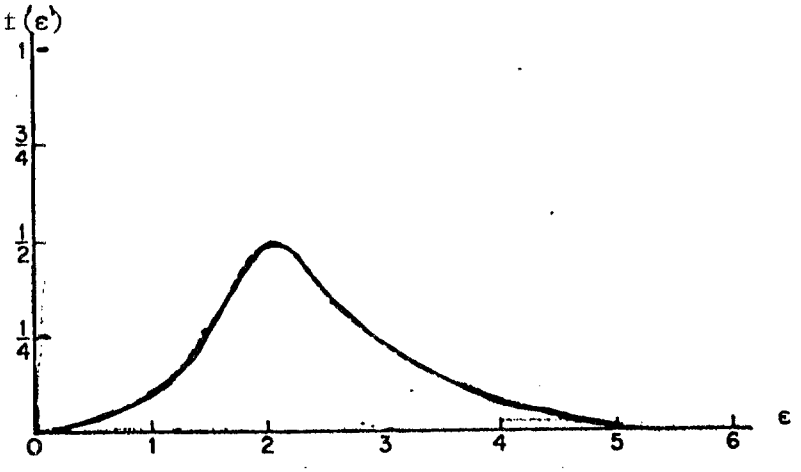


Figure 3. Gamma Distribution For More Than Three Degrees of Freedom

Probit Analysis

The question being asked is, "What is the probability of an alternative being selected by an individual given the observed characteristics of that individual?" Probit analysis is particularly appropriate for identifying the critical levels of X at which the value of Y switches from 0 to 1 (or from any given value to another). The alternative choices in Y can be ranked, i.e., they are ordinal.

For example: What is the probability that an individual with a set of observed characteristics will choose a particular type of milk to drink, i.e., whole milk, low fat milk, or skim milk?

The mathematical specification is:

$$E(Y \mid X) = F(X'\beta)$$

$F(X'\beta)$ is assumed to be a cumulative density function of a normal random variable, namely, the critical levels of X at which $E(Y|X)$ switches from one value of Y to another (V_i^*). X is generally assumed to be observations on nonstochastic variables, which are independent of each other.

The probit transformation is:

$$P_i = F(V_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{V_i} e^{-1/2\, t^2}\, dt \text{ where}$$

$$V_i = \beta_0 + \beta_1 X_{1i} + \ldots \beta_n X_{ni}$$

For a dichotomous dependent variable

$$Y = 0 \text{ if } V_i < V_i^*$$

$$Y = 1 \text{ if } V_i \geq V_i^*$$

where V_i^* is the critical value of V_i at which switching from one alternative to the other takes place.

V_i^* may be considered to be distributed N(0, 1) or N(5, 1). If $V_i^* \sim N(5, 1)$ then (Probit P_i) = $V_i + 5$. In some cases the 5 is added to avoid negative values for the Probit P_i.

Tobit Analysis

Tobit is a more general case of probit and is sometimes referred to as the Hybrid Probit Curve. For example: Probit analysis may be used to estimate the probability of an individual purchasing (y = 1) or not purchasing

(y = 0) a new automobile. Tobit analysis would estimate the same probability and in addition, estimate the amount spent for the automobile in the cases where y = 1.

Tobit was developed for estimating the probability of limited and nonlimited responses and, in addition, the value of the nonlimited response.

The mathematical specification is:

$$(\text{Tobit } P_i) = \Pr(y_i \leq V_i + \varepsilon) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{V_i} e^{-1/2t^2} dt$$

$$V_i = X'\beta$$

$$Z(\text{Tobit } P_i) = \frac{1}{\sqrt{2\pi}} e^{-1/2 (\text{Tobit } P_i)^2}$$

$y = L$ if $V_i - \varepsilon < L$

$= W$ if $V_i - \varepsilon \geq L$

$\varepsilon \sim N(0, \sigma)$

$(V_i - \varepsilon) = W$ which is the value of the nonlimiting response

L is the lower limit

Section II

Annotative Bibliography for Logit Analysis

This bibliography presents sources of information for the researcher who believes that logit analysis applies to his/her research and therefore wants to become familiar with the theoretical concepts and various empirical applications. One of the underlying purposes of this bibliography is to inform the reader of the wide variety of problems to which logit analysis has been applied. Most of the sources are available in books or journals; most papers from the Travel Demand Project at Berkeley are available free upon request.

Two sources are designated as good "single sources" meaning that each of them contains relatively complete theoretical and empirical developments.

Aldrich, John and Charles F. Cnudde
1975 "Probing the Bounds of Conventional Wisdom: A Comparison of Regression, Probit, and Discriminate Analysis," American Journal of Political Science, Volume XIX, Number 3, pp. 571-608.

The three estimating techniques designated in the title are discussed using political election data to illustrate their respective appropriateness to dependent variables which are continuous (interval), ordinal, or nominal (discrete). Particular attention is given to the theoretical and distributional assumptions of each model and their compatability with the substantive assumptions of the theory/hypotheses being investigated. Empirical models compare results

from ordinary least squares and probit analysis. A discriminate model is also estimated. Sample print-out from the 3 programs are provided.

This is a clearly written article; useful for basic understanding. Much of what is said about probit applies to logit analysis when comparing the three given alternatives estimating techniques.

Amemiya, Takeshi
1976 "The Maximum Likelihood, The Minimum Chi Square and the Nonlinear Weighted Least-Squares Estimator in General Qualitative Response Models," Journal of the American Statistical Association, 71:354, pp. 347-351.

This is a theoretical discussion of a general qualitative response model and of probit and logit models. All of these models may be estimated by various techniques, although maximum likelihood estimates are the most common. Amemiya shows the theoretical differences between various estimators. This article is useful for econometricians interested in developing or selecting alternative estimating packages. There is a good list of references for those who want to explore the use and development of logit and probit analysis within the field of biometrics.

Amemiya, Takeshi and Fredrick Nold
1973 A Modified Logit Model, Technical Report No. 113, Institute for Mathematical Studies in the Social Sciences, Stanford University, Stanford, California.
(Also Review of Economics & Statistics, 57, 1975, pp. 255-257)

This report poses a theoretical question with respect to the efficiency of the parameter estimates from logit models utilizing maximum likelihood estimates, and a single error term. An alternative logit model estimated by

weighted least squares is presented. An empirical study verifying the inefficiency of the usual logit model is included.

Ashton, Winefred D.
1972 The Logit Transformation, London: Charles Griffin and Co., Lmtd.

This book describes in mathematical detail the logistic function and how it is used to transform linear probability functions. The theoretical assumptions and properties of a logistic function are included. This is No. 32 of a series of books in Griffin's Statistical Monographs and Courses.

Berkson, Joseph
1953 "A Statistically Precise and Relatively Simple Method of Estimating the Bio Assay with Quantal Response, Based on the Logistic Function," Journal of the American Statistical Association, 48: 263, pp. 565-599.

This theoretical article compares two different estimators of logistic functions; maximum likelihood estimates (MLE) and Minimum Chi Square estimates. Berkson shows that MLE are not necessarily the better of the two estimates. He develops the Chi-Square estimate for a logit model.

Boskin, Michael J.
1974 "A Conditional Logit Model of Occupational Choice," Journal of Political Economy, 82:2 Part 1, pp. 389-397.

This is an empirical application in which the probability of choosing a particular occupation is measured as a function of the present value of potential earning, present value of time unemployed, and the ratio of training cost to net worth. Different equations were estimated for subpopulations divided by race and sex.

Brown, Randell S., Marilyn Moore, and Barbara Zoloth
1976 "Measuring Wage and Occupational Discrimination: A Comprehensive Approach," Madison, Wisconsin: Institute for Research on Poverty Discussion Paper Number 76-354.

This is an empirical application of multinominal logit analysis used to predict occupational attainment from a set of personal characteristics. The model allows the researchers to distinguish between the effects of occupational barriers and wage discrimination on wage differentials by first estimating the parameters of the model using a sample of mature men from the National Longitudinal Survey. The resulting parameter estimates are then used to simulate the occupational distribution of women that would exist if women faced the same employment possibilities as men. A comparison of estimators from logit analysis and discriminate analysis is provided; they are very similar.

Brownstone, David and High Wills
1975 "CDC XLOGIT 2.1 Users Guide": Working Paper No. 7412, 25 p. and "XLOGIT Programmers Guide 2.1": Working Paper No. 7411, 64p., Mimeos., Travel Demand Project, Institute of Transportation and Traffic Engineering, Berkeley, California; University of California.

Computer Program guides to the Berkeley CDC XLOGIT program. (These are helpful even if one is using a CLOGIT or QUAIL program).

Also: Berkman, Brownstone, McFadden, and Duncan have written an updated version of the XLOGIT program called QUAIL User's Manual; Working Page No. 7402. The users guide and the program are available from the same source cited above.

Domencich, Thomas A. and Daniel McFadden
1975 Urban Travel Demand: A Behavioral Analysis. Contributions to Economic Analysis, No. 93, New York: American Elsevier; Amsterdam and Oxford: North-Holland, 1975.

This book is a theoretical and empirical presentation of problems involving choice of discrete alternatives. The theory applies to several types of economic decisions other than the demand for various transportation modes. This book provides a prototype empirical study demonstrating the feasibility of the methods and the usefulness of the results. Logit models are developed from stochastic utility functions. Statistical estimation of logit and related types of analysis are explained and an empirical study is reported in which the probability of selecting travel mode, trip time, destination and frequency of travel was estimated. This is a good single source for information in logit theory and analysis. The development and explanation of the theory is more detailed, and in many instances clearer, than in other sources.

For more information see the book review by Kan Hua Young in the Journal of Economic Literature, December 1976, p. 1350.

Efron, Bradley
1975 "The Efficiency of Logit Regression Compared to Normal Discriminate Analysis", Journal of the American Statistical Association, 70:352, pp. 892-898.

This article is a mathematical presentation of a theoretical comparison of the efficiency of estimates obtained by discriminate analysis and logit analysis. The conclusion drawn is that logistic regression is less efficient but more robust due to less restrictive assumptions on the underlying density function.

Goldberger, Authur S.
1964 Econometric Theory, New York: John Wiley and Sons, Incorporated.

Goldberger discusses theoretically how linear probability models with dichatomous dependent variables may be estimated using weighted least squares. The reasons why ordinary least squares should not be used on such models are also presented. Logit analysis, as such, is not discussed.

Haggstrom, Gus
1974a "Logistic Regression and Discriminate Analysis," Memorandum, April 3, 1974, RAND Corporation, Santa Monica, California.

1974b "Postscript on Logistic Regression," Memorandum, April 9, 1974, RAND Corporation, Santa Monica, California.

1974c "Last Grasp on Logistic Regression," Memorandum, April 30, 1974, RAND Corporation, Santa Monica, California.

These three memorandums develop mathematically the theoretical likenesses of logit and discriminate analysis. A transformation of ordinary least squares (OLS) estimates to discriminate estimates is presented. A short empirical study compares coefficients estimated by OLS, transformed OLS, amd maximum likelihood techniques in a logit model.

Kinsey, Jean
1976 "The Effect of Debt on Household Welfare," Ph.D. dissertation (unpublished), Agricultural Economics Department, University of California, Davis, California.

This is an empirical application of binary logit analysis used to estimate perceived changes in household's financial well-being as a function of changes in debt-asset ratios and other financial and demographic variables. Included is a documented discussion of the theory of logit analysis and how it applies to choice theory and utility maximization and a comparison of maximum

likelihood (logit) and transformed ordinary least squares (discriminate) estimates. Tests for normality and the independence of irrelevant alternatives are reported.

Kulshreshtha, Surendra N.
1975 "Ownership of Farm Trucks for Hauling Grain: An Application of Multivariate Logit Analysis," Journal of Agricultural Economics, 57:2, pp. 302-308.

This is an empirical application of a bivariate logit model where the two dependent variables are size and age of trucks used for hauling grain. Each dependent variable consisted of three alternative choices. Six equations were estimated by weighted least squares. Logit functions were also used to predict the ownership of trucks. This application is not the same as the polytomous logit models as specified by Nerlove and Press or Schmidt and Strauss. It utilizes a logistic function but is different from the logit analysis discussed in other articles in this bibliography.

Ladd, George
1966 "Linear Probability Functions and Discriminate Functions," Econometrica, 34:4, pp. 873-885.

Although this article does not talk about logit analysis, it is useful to read in conjunction with the Haggstrom memos or McFadden (1974a). Ladd shows, theoretically, how linear probability functions estimated by ordinary least squares and discriminate analysis start at different places, follow different routes and end up at nearly the same place.

Luce, R. Duncan
1959 Individual Choice Behavior, A Theoretical Analysis, New York: J. Wiley and Sons, Incorporated.

This book develops the basic theory of choice which is the foundation for utility (choice) models later used by McFadden and estimated by logit analysis. One of the simplifying, but restricting, properties of multivariate logit programs is the property of the independence of irrelevant alternatives. Luce presents this property as an axiom of choice.

Mathematical set notation is used to present the theory but each section includes a clear verbal explanation.

Lundin, Anders
1975 "A Conditional Logit Analysis of Outdoor Recreation Demand in National Forests in California," Ph.D. dissertation (unpublished) Agricultural Economics Department, University of California - Berkeley.

This is an empirical application of multivariate logit analysis which also reviews choice theory and its link to utility maximization. The problem estimated was the probability that a randomly selected individual with a vector of socio-economic characteristics will choose a given campsite with its own vectar of attributes. A multistage decision model and its implications for substituability and cross elasticities is discussed.

McFadden, Daniel
1974a "A Comment on Discriminant Analysis 'Versus' Logit Analysis," Mimeo, Department of Economics, University of California, Berkeley, California, 22p.

This is a mathematical comparison of the theoretical likenesses of logit and discriminate analysis. There is a very readable discussion of underlying assumptions and experimental settings for which each of the methods is appropriate. In general logit analysis is deemed more appropriate for causal models and discriminate analysis more appropriate for cojoint models.

McFadden, Daniel
1974b "Conditional Logit Analysis of Qualitative Choice Behavior," in Frontiers of Econometrics, Paul Zarembka (Ed.), New York: Academic Press, pp. 105-142.

This article outlines the theory of selection probabilities as applied to a utility maximizing model. Assumptions which are inherent in the conditional logit transformation of this model are presented along with the maximum likelihood estimation procedure. Statistical properties are discussed and proofs are offered in an appendix.

An empirical study is included which estimates the conditional probability of the choice of mode of travel for a shopping trip, the distination of the trip and the frequency of shopping trips.

McFadden, Daniel
1975a "Comments on 'Estimation of a Stochastic Model of Reproduction: An Ecometric Approach" in Household Production and Consumption, Nestor E. Terleckyj (Ed.), Studies in Income and Wealth, Volume 40, New York: Columbia University Press, pp. 139-145.

This is an empirical application of binary logit analysis. The paper is offered as an alternative to methods used by James J. Heckman and Robert Willis to estimate the monthly conditional probability of conception (to be found in the same book). A brief theoretical argument is presented in favor of using logit analysis to estimate the choice of contraceptive technique. The empirical estimation presented is an estimation of the probability of a mother expecting one or more children as a function of the sex ratio of the current family, education of the wife, race and family income.

McFadden, Daniel
1975b "Economic Applications of Psychological Choice Models," Working Paper Number 7519, Travel Demand Forecasting Project, Institute of Transportation and Traffic Engineering, University of California, Berkeley, August, 1975, 27p.

In this paper, McFadden formulated theoretical models of the probability of economic choices based on psychophysical theories of judgment and choice developed in the past decade. Random utility maximization models are discussed in the context of choice theory. The independence and sequencing of choice probabilities are discussed and selection probability formulas for three types of choice models are presented in detail: the Luce model which assumes the independence of irrelevant alternatives allowing simple scalability of choices; the Tversky or "elimination by aspects" model; and an "elimination by strategy" model.

McFadden, Daniel
1975c "On Independence, Structure, and Simultaneity in Transportation Demand Analysis," Preliminary Draft of Working Paper Number 7511, Travel Demand Project, Institute of Transportation and Traffic Engineering, University of California, Berkeley, May 1975, 37p.

This is a theoretical presentation of the link between utility maximization, choice probabilities and logit analysis. The assumption of the independence of irrelevant alternatives is discussed and tested and a generalized multinomial logit structure which circumvents that assumption is presented. A section is devoted to the calculus of the probability for joint and independent choice structures, additively separable utility functions, conditional estimations and simultaneous models. (It is recommended that this paper be read before or simultaneously with "Conditional Logit Analysis of Qualitative Choice Behavior" McFadden, 1974b.)

McFadden, Daniel
1975d "The Revealed Preference of Public Bureaucracy: Theory," The Bell Journal of Economics, 6:3, pp. 401-416.

The theory of qualitative choice is presented in this article in the context of examining the performance of a public bureaucracy. The approach is to examine the outcomes of the organization's decisions and to pose the revealed preference question of whether there exists an implicit choice criterion such that the bureaucracy behaves as if it is attempting to follow this choice rule. The bureaucracy studied was the California Division of Highways which chooses highway routes.

The development includes a strict utility model and a multinomial logit model with all attendent assumptions.

The empirical followup to this article is in the same journal in Spring 1976.

McFadden, Daniel
1976a "Quantal Choice Analysis: A Survey," Working Paper No. 7625, Department of Economics and Urban Travel Demand Forecasting Project, Institute of Transportation Studies, University of California, Berkeley. (Forthcoming in Annals of Economic and Social Measurement).

This is a theoretical paper which discusses qualitative choices in the context of random and strict utility models. McFadden clearly points out the connections between these two models and their conceptual flexibility and computational difficulties. Various statistical problems and estimating techniques are presented; the reader is challenged to search for better specifications. Specifying logit models with simultaneous choices present other problems and these are also discussed. This paper will be most interesting to an econometrician or to someone who is using logit analysis and wants to anticipate statistical problems.

There is a list of available computer programs for logit analysis and an extensive reference list citing numerous related theoretical articles and some empirical applications. Twenty-one of the thirty-eight articles reviewed in this bibliography are included in McFadden's reference list.

McFadden, Daniel
1976b "The Revealed Preference of a Government Bureaucracy: Empirical Evidence," The Bell Journal of Economics, 7:1, pp. 55-72.

This is an empirical followup to a theoretical article in the same journal in Fall 1975. Logit analysis is used to estimate the probability of a certain highway route being chosen as a function of up to twenty variables. The interpretation and discussion of results is helpful for anyone actually using logit analysis.

McFadden, Daniel and Charles Manski
1976 "Partial Summary of Alternative Estimators and Sample Designs for Discrete Choice Analysis," Mimeo, Department of Economics, University of California - Berkeley, 23p.

This is a mathematical treatment of theoretical problems with maximum likelihood estimators (MLE) which are efficient under the assumptions of correct model specification and no data measurement errors. This paper is relevant to logit analysis because conditional logit programs typically employ MLE and there is little reason to believe that the correct specification of choice models is always obvious or that the measurement of variables is error free. A relatively robust estimator is suggested; a Modified M-Class estimator where the original likelihood function is weighted. A short Monte Carlo study is presented to compare MLE and MM-Class estimators.

McFadden, Daniel, William Tye, and Kenneth Train
1976 "Diagnostic Tests for the Independence from Irrelevant Alternatives Property of the Multinomial Logit Model," Working Paper No. 7616, Urban Travel Demand Forecasting Project, Institute of Transportation Studies, University of California, Berkeley, August 1976, 71p.

A theoretical and empirical paper which relates the independence from irrelevant alternatives (IIA) property of multinomial logit analysis to psychological concepts of simple scalability and order independence. The reader is alerted to possible sources of model misspecification by a discussion of conditions which violate the IIA property. Several tests are suggested to determine whether a given set of data or a given model violates the IIA assumption.

A multinomial logit model is estimated and tested for violations of the IIA assumption. In a lengthy appendix several multinomial logit models are fitted to data generated from "true" choice models. These models are used to determine whether diagnostic tests can detect whether or not the data were generated from models which deviate from multinomial logit models. The data concerned the probability of choosing various transportation modes.

Miklias, Walter, Kenneth L. Casavant, and Peter V. Garrod
1976 "Estimates of Demand for Transportation of Agricultural Commodities," American Journal of Agricultural Economics, 58:2, pp. 217-223.

This is an empirical study where binary logit analysis was used to estimate the elasticities and cross elasticities of demand for various transportation modes as a function of their price and the quality of service characteristics.

Morgan, James N.
1975 Five Thousand Families - Patterns of Economic Progress, Volume II, Ann Arbor, Michigan; Institute for Social Research, Survey Research Center, University of Michigan.

An appendix of this volume discusses, theoretically, the differences between logit and linear estimations of a probability function, with respect to the dependence of the probability estimates on the value of the explanatory variables.

Nerlove, Marc and S. James Press
1973 Univariate and Multivariate Log-Linear and Logistic Models, Santa Monica, California: The RAND Corporation, R-1306-EDA/NIH

This is a theoretical and empirical book. The basic theory of estimating regression models with discrete dependent variables by maximum likelihood techniques is presented. Logit models with multivariate (polytomous) dependent variables are developed in detail. Empirical applications to three different sets of data are presented. The first two sets present models with dichotomous dependent variables. These models are estimated by various techniques including ordinary least squares and the results are compared to logit estimates. The third set of data presents a model with a polytomous dependent variable.

A computer program for the analysis of jointly dependent qualitative data is outlined at the end of the book.

This is a particularly thorough treatment of logit models where multiplicative choices comprise the dependent variable. It is not the best book for intuative inspiration, but it is a good comprehensive single source on logit analysis.

Orcutt, Guy H., Martin Greenberger, John Korbel, and Alice M. Rivlin
1965 Microanalysis of Socioeconomic Systems: A Simulation Study. New York: Harper & Row.

This book begins with the basic assumption that aggregate time series data used to analyze economic relationships is not very useful in determining how the components--households, firms, labor unions, and governmental units--of the economic system make decisions and interact with other components. "Predictions about aggregates are needed, but they should be obtained by aggregating behavioral relationships of these elemental units" (p. 12).

The major task of the book is to develop simulation models of real systems in order to analyze economic behavior at a disaggregated level. A probabilistic approach to predicting behavior was adopted in an attempt to utilize the type of information available from decision-making units. The probability of an outcome is predicted given observed inputs and status variables. This approach provides the important theoretical link to other choice problems and to logit analysis.

Models are developed to simulate the decision-making units of the consumer sector of the economy. For example, models are developed to estimate the probability of birth, death, marriage and divorce. The second part of the book develops models for and examines labor force participation, debt and asset behavior of spending units and the demand for higher education. A computer program is also provided.

The building of complex economic models is one of the most interesting aspects of this book.

Radner, Roy and L. S. Miller
1970 "Demand and Supply in U.S. Higher Education: A Progress Report," The American Economic Review, 60:2, pp. 326-334.

This article consists of reports on two empirical models, one of which is a generalized logit model. Logit analysis was used to estimate the conditional probabilities of 1966 high school graduates selecting one of nine institutions of higher learning or not going on to college at all. The probabilities are a function of the ratio of the cost of the institution type to the family income and the product of an estimated SAT score times the average SAT of the people choosing the institutional type.

Schmidt, Peter and Robert P. Strauss
1975 "Estimation of Models with Jointly Dependent Qualitative Variables: A Simultaneous Logit Approach," Econometrica, 43:4, pp. 745-755.

This is an empirical application of multivariate logit analysis on a simultaneous model, designed to predict occupation and industry of a worker based on demographic characteristics. The models estimated are developed in mathematical detail, including the likelihood function. There is also a clear interpretation of the results.

Schmidt, Peter and Robert P. Strauss
1975 "The Prediction of Occupation Using Multiple Logit Models," International Economic Review, 16:2, pp. 471-486.

This is an empirical application of logit analysis with one dependent variable comprised of five discrete alternatives. Occupational attainment is regressed on five demographic characteristics. The models are identified in mathematical detail and the interpretation of the results is particularly interesting.

Tardiff, Tim
1976 "A Note on Goodness of Fit Statistics for Probit and Logit Models," Transportation 5:4, pp. 377-388.

This is a theoretical discussion of the mathematical properties of the Likelihood Ratio Index (LRI) as a measure of the goodness of fit for logit models. The LRI depends upon the sample proportions selecting each alternative in a given model; this prevents a valid comparison of indices eminating from different models. The proposed Adjusted Likelihood Ratio Index (ALRI) is more accurately compared to the R^2 measure obtained from ordinary least squares models since both measure the explanatory power of a model in addition to the variance explained by a constant. A simple procedure for converting the LRI to the ALRI is presented.

Theil, Henri
1970 "On the Estimation of Relationships Involving Qualitative Variables," American Journal of Sociology 76, July, pp. 103-154.

Theil discusses theoretically the linear logit relationship and the Chi Square test for the validity of the model specification including the case of multiple responses. He ties logit analysis to information theory. An appendix provides models for the theory discussed in the article.

This article provides more detail than the sections cited from Theil's textbook, Principles of Econometrics.

Theil, Henri
1971 Principles of Econometrics, New York: John Wiley & Sons, Inc., pp. 188-190, 628-635.

The designated sections of this econometrics text discuss the theory of multivariate normal distributions and the logit and probit monotonic transformations of a linear probability function. One example of logit analysis is given.

Zellner, A. and T. H. Lee
1965 "Joint Estimation of Relationships Involving Discrete Random Variables," Econometrica 33:2, pp. 382-393.

This article presents the theoretical development of probit and logit analysis and discusses the minimum Chi Square estimating techniques. (The linear probability model and the Gompit model are also presented.) When the dependent variable is really comprised of two sets of choices which are correlated with each other but modeled as separate dependent variables of

separate equations, Zellner and Lee propose a method of joint estimation utilizing Aitken's generalized least squares estimation. They suggest that this approach could be used with probit and logit models.

They supply an empirical study in which they used their joint estimation technique on a linear probability model to explain the proportions of households purchasing durable goods and the proportion using installment credit.

Author Index

THEORETICAL Readings on Logit Analysis and Choice Theory

Author Index (continued)

EMPIRICAL Applications Using Logit Analysis

Author Index (continued)

The author would like to thank the following people for their comments on earlier drafts of this paper: Andria Lubov, Malcolm Purvis, Terry Roe, Burt Sundquist, and Cameron Thraen. They bear no responsibility for the contents herein.