1995

Probit Analysis of Correlated Data: Multiple Observations Over Time at One Concentration

James E. Throne
USDA-ARS, Manhattan, KS, james.throne@ars.usda.gov

David K. Weaver
USDA-ARS

Victor Chew
USDA-ARS

James E. Baker
USDA-ARS

Probit Analysis of Correlated Data: Multiple Observations Over Time at One Concentration

JAMES E. THRONE,[1] DAVID K. WEAVER,[2] VICTOR CHEW,[3] AND JAMES E. BAKER[1]

Stored-Product Insects Research and Development Laboratory, USDA-ARS, 3401 Edwin Street, Savannah, GA 31405

J. Econ. Entomol. 88(5): 1510-1512 (1995)

ABSTRACT We present a method to analyze serial time-mortality data from bioassay experiments where successive observations are made on the same group of organisms exposed to 1 concentration of a stimulus (for example, a pesticide). Standard probit techniques are not applicable because such data are correlated. Methods are presented for calculating statistics for the time-mortality line by regressing complementary log-log, logit, or probit transformations of proportion of responders on untransformed time or logarithmic transformations of time; and for calculating confidence limits on lethal time values corresponding to given mortality levels. As a result we have developed a computer program to facilitate use of the method.

KEY WORDS bioassay, correlated data, probit analysis, serial time-mortality data

PROBIT ANALYSIS is used to analyze data from bioassay experiments (Finney 1964). In entomology, samples of insects are typically exposed to several concentrations of an insecticide to determine the concentration that will kill 50% of the insects within a given time span (for example, Cilek and Knapp 1993). Effects of time on percentage of kill at 1 concentration (serial time-mortality data) may be of interest when (1) materials are limited, as might occur in tests of insecticides on field strains where few insects are available, or when testing an experimental pesticide that is available in limited quantities; or (2) when speed of kill is important, as might occur with a pest that lays all of its eggs within a few days (like a short-lived stored-product insect) or in quarantine treatments. Standard probit analysis techniques are not applicable to serial time-mortality data because observations made on the same group of organisms at different times are correlated (Robertson and Preisler 1992).

Lampkin and Ogawa (1975) developed a method for calculating the slope and intercept of serial time-mortality data. Reports on results of probit-type analyses should include the standard errors of the slope, intercept, and lethal time or lethal concentration values, and a test for goodness-of-fit. Any tests comparing slopes, intercepts, or lethal time values should include confidence limits on the estimated statistics. Preisler and Robertson (1989) described a method to analyze bioassay data when response by the same groups of organisms was determined at several times and at several concentrations of the insecticide (time-dose-mortality data).

Here we present a method for analyzing correlated serial time-mortality data using the complementary log-log, logit, or probit transformation of proportion insects killed. We have developed a computer program that can be used to implement the method.

Materials and Methods

Probit Analysis. Probit analysis for correlated data differs from standard probit analysis because in addition to their variances, the covariances of the probits also must be estimated to account for correlation among observations. The usual data obtained in a serial time-mortality experiment are, the times at which observations were made (t), the cumulative number of insects that are dead at each observation time in the treatment (d_{trt}) and in the control (d_{cont}), the number of insects treated (n_{trt}) and in the control (n_{cont}), and the number of times observations were made (m). From these data, the number and proportion of insects that are killed in the treatment (k_{trt}, p_{trt}) or control (k_{cont}, p_{cont}) during a time interval are calculated. The proportion of insects killed during each time interval in the treatments is corrected for control mortality using Abbott's (1925) formula. If a corrected $p_{trt}(i) \leq 0$ for $i = 1, 2, \ldots m$, the correction suggested by McCullagh and Nelder (1989) is applied because the probit and logit transforms of $p_{trt}(i)$ are undefined if $p_{trt}(i) \leq 0$.

This article reports the results of research only. Mention of a proprietary product does not constitute an endorsement or a recommendation for its use by USDA.

[1]Current address: U.S. Grain Marketing Research Laboratory, USDA-ARS, 1515 College Avenue, Manhattan, KS 66502.

[2]Current address: Insect Attractants, Behavior, and Basic Biology Research Laboratory, USDA-ARS, 1700 S.W 23rd Drive, Gainesville, FL 32608.

[3]USDA-ARS, University of Florida, 412 Rolfs Hall, Gainesville, FL 32611.

Each insect is a multinomial trial that can result in one of m or $(m + 1)$ mutually exclusive and exhaustive outcomes, depending on whether all n_{trt} insects have or have not died at the last (mth) observation time, respectively. Thus, the unbiased estimates of the variances of the uncorrected $p_{trt}(i)$ and the covariances of the uncorrected $p_{trt}(i)$ and $p_{trt}(j)$ are defined by multinomial theory as:

$$\widehat{\text{var}}[p_{trt}(i)] = \frac{p_{trt}(i)[1 - p_{trt}(i)]}{n_{trt} - 1} \tag{1}$$

and

$$\widehat{\text{cov}}[p_{trt}(i), p_{trt}(j)] = \frac{-p_{trt}(i)p_{trt}(j)}{n_{trt} - 1}, \quad i \neq j, j = 1, 2, \ldots, m. \tag{2}$$

The equation for the corrected $p(i)$'s in terms of the uncorrected $p(i)$'s is linearized using 1st-order Taylor series expansion, and the variances and covariances of the corrected $p(i)$'s $[p_{corr}(i)]$ are obtained from those of the uncorrected $p(i)$'s using standard covariance propagation technique. The variances of the $p_{corr}(i)$ are:

$$\widehat{\text{var}}[p_{corr}(i)] = \left(\frac{1}{1 - p_{cont}(i)}\right)^2 \frac{p_{trt}(i)[1 - p_{trt}(i)]}{n_{trt} - 1} + \left(\frac{1 - p_{trt}(i)}{[1 - p_{cont}(i)]^2}\right)^2 \frac{p_{cont}(i)[1 - p_{cont}(i)]}{n_{cont} - 1} \tag{3}$$

and the covariances of $p_{corr}(i)$ and $p_{corr}(j)$ are:

$$\widehat{\text{cov}}[p_{corr}(i), p_{corr}(j)] = \left[\frac{1}{1 - p_{cont}(i)} \frac{1}{1 - p_{cont}(j)}\right] \cdot \frac{-p_{trt}(i)p_{trt}(j)}{n_{trt} - 1} + \frac{-[1 - p_{trt}(i)]}{[1 - p_{cont}(i)]^2} \frac{-[1 - p_{trt}(j)]}{[1 - p_{cont}(j)]^2} \cdot \frac{-p_{cont}(i)p_{cont}(j)}{n_{cont} - 1}, \quad i \neq j. \tag{4}$$

The estimate of the covariance matrix for the cumulative proportion treated insects dead at each time interval corrected for control mortality (p^*) is calculated from the variances and covariances of the $p_{corr}(i)$'s using standard covariance propagation technique so that:

$$\text{cov}(p^*) = \text{C cov}(p_{corr})\text{C}' \tag{5}$$

where **C** is an m x m matrix of zeros and ones that is used simply as a multiplier to accumulate the proportion insects dead at each observation time. For an experiment with 3 observation times, **C** is:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}. \tag{6}$$

The complementary log-log, logit, or probit transformations (z) of the cumulative proportion dead (p^*) are calculated, linearized using 1st-order Taylor series expansion, and their variances and covariances calculated using standard covariance propagation technique so that for the probit transformation:

$$\widehat{\text{var}}[z(i)] = \frac{\text{var}[p^*(i)]}{\phi[z(i)]^2} \tag{7}$$

and

$$\widehat{\text{cov}}[z(i), z(j)] = \frac{\text{cov}[p^*(i), p^*(j)]}{\phi[z(i)]\phi[z(j)]}, \quad i \neq j. \tag{8}$$

$\phi[z(i)]$ is the derivative of $p^*(i)$ with respect to $z(i)$ (Beyer 1987**).** If using the logit transformation, the variance of $z(i)$ is:

$$\widehat{\text{var}}[z(i)] = \frac{1}{n_{trt}p^*(i)[1 - p^*(i)]} \tag{9}$$

and the covariance of $z(i)$ and $z(j)$ is:

$$\widehat{\text{cov}}[z(i), z(j)] = \frac{1}{p^*(i)[1 - p^*(i)]} \frac{1}{p^*(j)[1 - p^*(j)]} \cdot \text{cov}[p^*(i), p^*(j)], \quad i \neq j. \tag{10}$$

If using the complementary log-log transformation, the variance of $z(i)$ is:

$$\widehat{\text{var}}[z(i)] = \frac{p^*(i)}{n_{trt}[1 - p^*(i)][\log_e(1 - p^*[i])]^2} \tag{11}$$

and the covariance of $z(i)$ and $z(j)$ is:

$$\widehat{\text{cov}}[z(i), z(j)] = \frac{1}{1 - p^*(i)} \frac{1}{1 - p^*(j)} \frac{1}{-\log_e[1 - p^*(i)]} \cdot \frac{1}{-\log_e[1 - p^*(j)]} \text{cov}[p^*(i), p^*(j)], \quad i \neq j. \tag{12}$$

Having estimated z and the covariance matrix of z, generalized weighted least-squares methods (Neter et al. 1990) can be used to estimate parameters for the regression of z on t. This method provides only initial estimates for the slope and intercept because the true value of cov(z) is unknown and its estimate is used in the least-squares regression. The final values of the slope and intercept are determined iteratively. Once the slope and intercept are calculated, the chi-square goodness-of-fit test is used to determine how well the regression line fits the observed data. A significant chi-square may indicate that the data are heterogeneous and should be corrected using a heterogeneity factor or that an alternative transformation would be more appropriate for the data (Finney 1964).

A lethal time value (the time required to kill a given proportion of the insects) is obtained by calculating z (using appropriate formulas for the complementary log-log transformation, logits, or probits)

corresponding with the proportion kill for which a lethal time value is desired, and calculating the lethal time value (LT) as:

$$\mathrm{LT} = \frac{z - \text{intercept}}{\text{slope}}. \qquad (13)$$

Fieller's theorem (Finney 1964) is used to calculate confidence limits on lethal times. If the chi-square is significant, all variances used to calculate confidence limits on lethal times must be multiplied by the heterogeneity factor, which is chi-square divided by m - 2 (Finney 1964). If the chi-square is not significant, a standard z of 1.96 is used to calculate confidence limits. If the chi-square is significant, a Student's t value with m - 2 df is used to calculate confidence limits.

A computer program to perform the probit analysis written in Mathematica language (Wolfram, Champaign, IL) can be obtained from J. E. T. All statistics required for complete reporting of probit-type analyses are provided by the program.

Discussion

The method described here is not intended to encourage the use of this experimental design as an alternative to making independent observations. If enough insects and insecticide are available, a well-designed study should include destructive sampling of insects at each observation time resulting in independent observations. Such an experiment would provide the best estimates of the slope and intercept of the regression line. However, when materials are limited, the method presented in this article will allow correct analysis of the data. Experiments that include multiple concentrations of pesticide and multiple observations over time should be analyzed by using the method of Preisler and Robertson (1989).

We have used our program successfully to analyze >25 sets of serial time-mortality data (for example, Baker et al. 1995). We fit all 6 possible equations (complementary log-log, logit, or probit transformation of proportion kill and logarithmic or no logarithmic transformation of time) to the data and choose the equation that gives the lowest chi-square to describe the data. When the intent is to compare regression lines, the same transformation must be used for all lines that will be compared with each other. The computer program provides an easy method for analyzing serial time-mortality data correctly and the resulting output provides all statistics necessary for complete reporting of probit analyses.

Acknowledgments

We thank Guy Hallman (USDA-ARS, Miami, FL), Bruce Mackey (USDA-ARS, Albany, CA), and Haiganoush Preisler (USDA-FS, Berkeley, CA) for their comments on an earlier version of the manuscript.

References Cited

Abbott, W. S. 1925. A method of computing the effectiveness of an insecticide. J. Econ. Entomol. 18: 265-267.

Baker, J. E., D. K. Weaver, J. E. Throne, and J. L. Zettler. 1995. Resistance to protectant insecticides in two field strains of the stored-product insect parasitoid *Bracon hebetor* (Hymenoptera: Braconidae). J. Econ. Entomol. 88: 512-519.

Beyer, W. H. 1987. CRC standard mathematical tables, 28th ed. CRC, Boca Raton, FL.

Cilek, J. E., and F. W. Knapp. 1993. Enhanced diazinon susceptibility in pyrethroid-resistant horn flies (Diptera: Muscidae): potential for insecticide resistance management. J. Econ. Entomol. 86: 1303-1307.

Finney, D. J. 1964. Probit analysis: a statistical treatment of the sigmoid response curve. Cambridge University Press, London.

Lampkin, H., and J. Ogawa. 1975. Estimation of distribution parameters in time mortality trials: an example of time mortality analysis. Can. J. Stat. 4C: 65-93.

McCullagh, P., and J. A. Nelder. 1989. Generalized linear models, 2nd ed. Chapman & Hall, New York.

Neter, J., W. Wasserman, and M. H. Kutner. 1990. Applied linear statistical models: regression, analysis of variance, and experimental designs, 3rd ed. Irwin, Homewood, IL.

Preisler, H. K., and J. L. Robertson. 1989. Analysis of time-dose-mortality data. J. Econ. Entomol. 82: 1534-1542.

Robertson, J. L., and H. K. Preisler. 1992. Pesticide bioassays with arthropods. CRC, Boca Raton, FL.

Received for publication 10 May 1994; accepted 22 May 1995.