RECIDIVISM AND LIKELIHOOD FUNCTIONS

A Reply to Stollmack

MICHAEL D. MALTZ
University of Illinois, Chicago Circle
RICHARD McCLEARY
Arizona State University
STEPHEN P. POLLOCK
University of Michigan

S tollmack (1978) has made a thorough analysis of our paper (Maltz and McCleary, 1977), and has provided us with additional citations of papers which use the same (Anscombe, 1961) or related (Blumenthal and Marcus, 1975) methods. Since our paper was published, we have found other, even earlier, work that would be of interest to those concerned with this general class of methods. As far as we can determine, Boag (1949) was the first to use this approach, for estimating the proportion of patients cured of cancer. Berkson and Gage (1952), Cutler and Axtell (1963), and Haybittle (1965) extended these methods. In addition, Carr-Hill and Carr-Hill (1972) and Greenberg (1978) have employed similar models in analyzing recidivism data.

Stollmack has a number of criticisms of our paper. They include:

(1) Our method does not "allow for a finer discrimination" between programs than the Stollmack-Harris method, nor did we provide an alternate methodology;

(2) Our likelihood function does not reduce to his when r = 1; see equations [1] and [2] in his comments;

(3) We should have based our model on sampling without replacement;

(4) An alternative estimation procedure used by Blumenthal and Marcus (1975) can be used;

(5) A two-parameter Weibull distribution fits the Illinois data better than the model we used.

We respond to these criticisms below:

(1) Our method does not permit finer discrimination that the Stollmack-Harris method, nor will it necessarily do so when we have developed the appropriate tests. In fact, we do not feel that fine discrimination is the ultimate goal of a method or the major issue in selecting one method over another. Their method may distinguish finely between two programs in terms of how *fast* its participants fail, when *all* are assumed to fail. This may not be as important as distinguishing between two programs in terms of how *many* in each program are expected to fail. Even if the discrimination between programs for the latter model of recidivism is coarser than for the former, it will often be more useful.

Stollmack is correct in stating that we did not provide an alternative methodology. We provided another conceptualization of the basic process in which not everyone in a program fails at the same rate, as in their paper. Furthermore, our methodology is not complete since, as we stated, "we are working on the development of statistical tests of significance" (p. 432) and "we are looking at other models of the recidivism process" (p. 433). In other words, we never promised you a rose garden, but described some promising buds.

(2) When r = 1, the difference between the two likelihood functions is a multiplicative constant. In the two most common uses of likelihood functions, maximum likelihood estimation and Bayesian estimation, the constant is unimportant: Box and Tiao (1973: 11) state: "The likelihood function is defined up to a multiplicative constant, that is, multiplication by a constant leaves the likelihood unchanged. . . . It is only the relative value of the likelihood which is of importance."

The reason that Blumenthal and Marcus (1975) included the multiplicative factor in their likelihood function is that their multiplicative factor is *not* constant. They use their procedure to estimate n, the total number in the population, whereas n is given in all cases of interest to evaluators.

(3) The point about sampling with and without replacement is a subtle one. Stollmack is correct in his assertion about this issue if we assume that the two subpopulations (successes and failures) are identifiable a priori as are, say, red and white marbles in an urn. However, in our case, the individuals are not identifiable beforehand; rather than considering there to be a mixture with proportions of r failures and 1-r successes, we consider that *each individual* has probability r of failing. This conception, we feel, is closer to the actual situation than the former one.

(4) Stollmack points out that the Blumenthal and Marcus (1975) method can be adapted to use in the type of problem we consider. This is true, although theirs is a more complicated approach: they suggest using a Bayesian procedure to estimate n and a maximum likelihood procedure for estimating the failure rate. We have also developed Bayesian procedures, which are discussed below.

(5) The fact that an alternative method "obtained a superior fit to the Illinois data," as Stollmack points out, is not entirely relevant. We specifically stated (p. 433) that we were basing our model on the process being studied; our goal was not to achieve the best fit of *any* model to the data, but the best fit of a model of the recidivism process to the data.

Furthermore, one can argue about the value of a best fit to the data in the region for which data exist, rather than looking for a best predictor of future data. The figures in our comment (Maltz and McCleary, 1978) on Miley's (1978) paper move in this direction.

There are other technical problems related to the methodology we described which should be addressed. The distribution we use is what Feller (1966: 127) calls a "defective distribution." (We hasten to point out that this does not imply that the method is "defective," but rather that the distribution does not have certain well-behaved properties at infinity.) As a consequence, one cannot use the second derivative of the likelihood function at its maximum to estimate the variance and correlation coefficient of the two parameters.[1]

For example, Figure 1 portrays the likelihood function for the first set of data given in Boag (1949),[2] in the vicinity of its maximum. As can be seen, a "ridge" runs off to the northwest. Figure 2 shows the bivariate normal probability density function with the variance and correlation coefficient estimated using the second derivatives. Figure 3 shows the difference between the two. As can be seen, the northeast ridge shows no signs of abating while the difference between the two curves in the other quadrants dies out quickly.
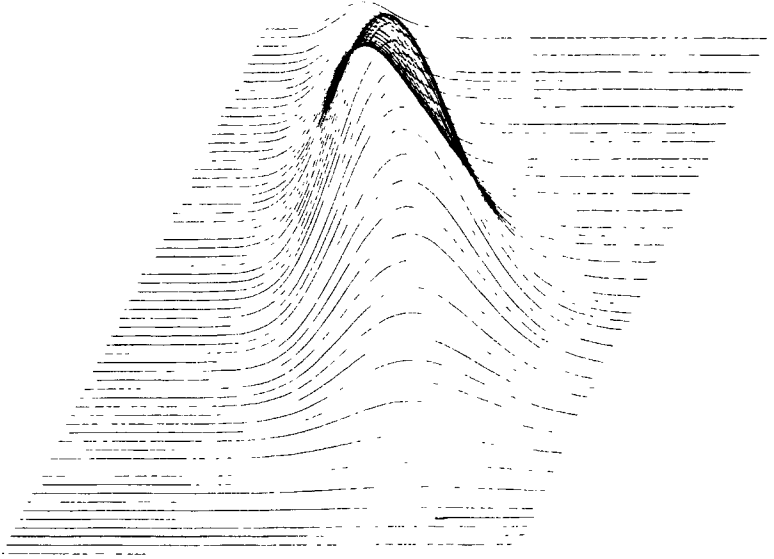
Figure 1: Likelihood Function for the Boag (1949) Data

One can easily overcome the difficulties that this distribution presents by using Bayesian estimation procedures, which we are presently doing (Maltz and Pollock, forthcoming). We use probabilities which restrict the range of the parameters (e.g., $0 < r < 1$). Bayesian techniques are also useful from the standpoint of developing statistical tests; we are presently developing such hypothesis testing procedures. Furthermore, we are working on methods using split populations in which both populations fail, but in which the two failure rates are quite dissimilar. These methods use well-behaved distribution factors.

CONCLUSIONS

The split-population model is a frequently used analytic convention in applied social sciences. Whenever a single-population assumption fails to fit a social phenomenon, analysts have employed the next
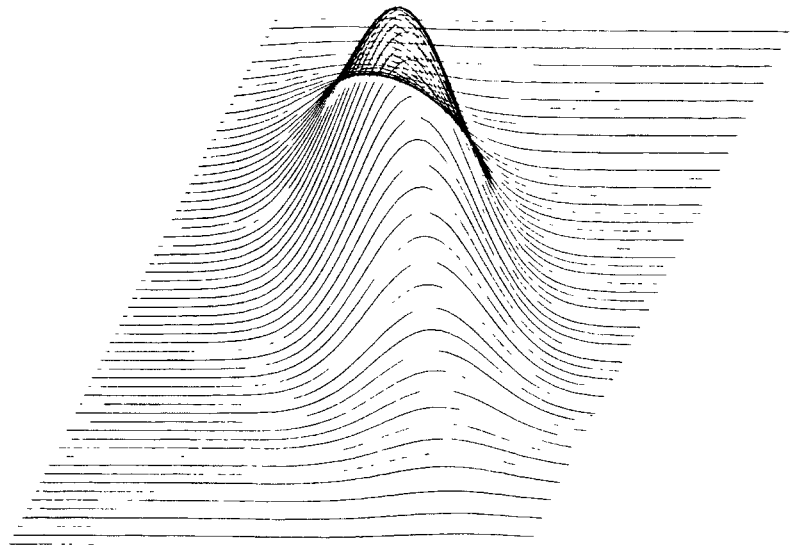
Figure 2: **Asymptotic Bivariate Normal Approximation to the Likelihood Function of Figure 1**

simplest assumption: that is, that a population dichotomy results from the phenomenon. The mover-stayer model used in demographic studies is an example of this device in another context. Thus, if we have made any contribution to the measurement of recidivism at all, it has been to point out that single-population models do not seem to fit the broad social phenomenon of recidivism.

We thank Stollmack for his comments and hope that we have clarified some of the more technical aspects of the method we proposed. We did not mean to imply that our method is some magical breakthrough that holds for all cases; our purpose was to provoke a debate of the subject of measurement models. Judging from the response of our paper, this purpose has been realized. And regarding the use of our particular method, we are not so pessimistic as Stollmack and are continuing to work on its development.
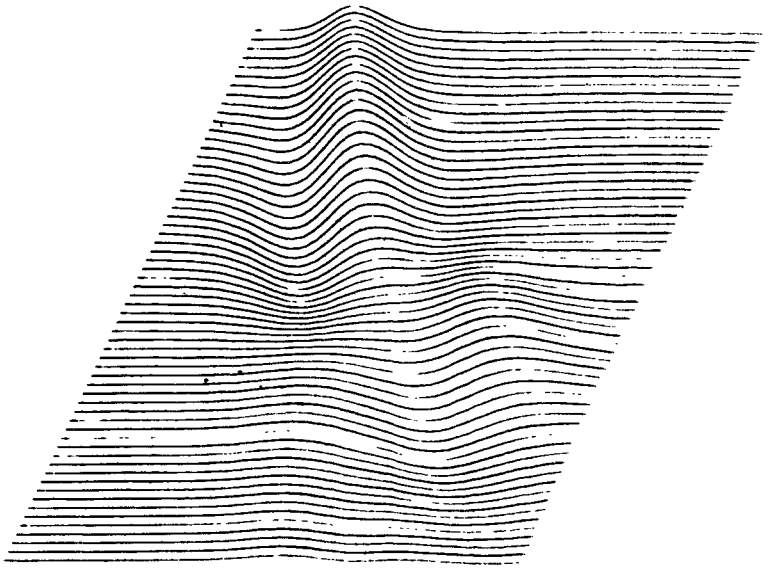
Figure 3: Difference Between Figures 1 and 2

NOTES

1. In our response to Miley's (1978) paper, we *incorrectly* used the asymptotic properties of the likelihood function to develop a lower bound on the variance for $\hat{r}$ and $\hat{a}$. Boag (1949: 25) committed the same error, as did Lloyd and Joe (1979) in their extension of our 1977 paper. We would caution those using the tests of significance developed by Lloyd and Joe not to place too much reliance on them, since the magnitude of error in the estimates of the covariance matrix is not known and may vary considerably depending on the data.

We have subsequently compared the "standard deviations" shown in the figures in our response (Maltz and McCleary, 1978) to Miley with (more accurate) Bayesian estimates. They are not very different from the actual values—for that particular data set.

2. Although Boag employed a likelihood function based on a lognormal distribution, we used a likelihood function based on an exponential distribution. Our likelihood function could be depicted (as a function of the two variables r and a), whereas his likelihood function could not.

REFERENCES

ANSCOMBE, F. J. (1961) "Estimating a mixed-exponential response law." J. of the Amer. Statistical Association 56: 493-502.

BERKSON, J. and R. P. GAGE (1952) "Survival curve for cancer patients following treatment." J. of the Amer. Statistical Association 47: 501-515.

BLUMENTHAL, S. and R. MARCUS (1975) "Estimating population size with exponential failure." J. of the Amer. Statistical Association 70: 913-922.

BOAG, J. W. (1949) "Maximum likelihood estimates of the proportion of patients cured by cancer therapy." J. of the Royal Statistical Society, Series B. 11: 15-53.

BOX, G.E.P. and G. C. TIAO (1973) Bayesian Inference in Statistical Analysis. Reading, MA: Addison-Wesley.

CARR-HILL, G. A. and R. A. ČARR-HILL (1972) "Reconviction as a process." Brit. J. of Criminology 12: 35-43.

CUTLER, S. J. and L. M. AXTELL (1963) "Partitioning of a patient population with respect to different mortality risks." J. of the Amer. Statistical Association 58: 701-712.

FELLER, W. (1966) An Introduction to Probability Theory and Its Applications, Vol. II. New York: John Wiley.

GREENBERG, D. F. (1978) "Recidivism as radioactive decay." J. of Research in Crime and Delinquency 15: 124-125.

HAYBITTLE, J. L. (1965) "A two-parameter model for the survival curve of treated cancer patients." J. of the Amer. Statistical Association 60: 16-26.

LLOYD, M. R. and G. W. JOE (1979) "Recidivism comparisons across groups: methods of estimation and tests of significance for recidism rates and asymptotes." Evaluation Q. 3, 1.

MALTZ, M. D. and R. McCLEARY (1978) "Rejoinder on 'Stability of parameter estimates in the split population exponential distribution.' " Evaluation Q. 2: 650-654.

——— (1977) "The mathematics of behavioral change: recidivism and construct validity." Evaluation Q. 1: 421-438.

MALTZ, M. D. and S. M. POLLOCK (forthcoming) "Bayes estimates for singly-censored data from a mixed failure distribution."

MILEY, A. D. (1978) "Stability of parameter estimates in the split population exponential distribution." Evaluation Q. 2: 646-649.

STOLLMACK, S. (1979) "Comments on 'The mathematics of behavioral change.' " Evaluation Q. 3, 1.

——— and C. M. HARRIS (1974) "Failure rate analysis applied to recidivism data." Operations Research 22: 1192-1205.

Michael D. Maltz is Associate Professor of Criminal Justice and Systems Engineering at the University of Illinois, Chicago Circle. He is author of Evaluation of Crime Control Programs *and numerous reports and papers on criminal justice evaluation and measurement problems.*

Richard McCleary is an Assistant Professor of Criminal Justice and Sociology at Arizona State University.

Stephen M. Pollock is Professor of Industrial and Operations Engineering and a Research Scientist in the Institute of Public Policy Studies at the University of Michigan. He has been extensively involved in the development and use of mathematical models to aid in public and private sector decision making.