International Journal of Business and Economics, 2005, Vol. 4, No. 3, 223-230

Estimating Function Approach to a Time-Dependent Recovery Model for Survival Rates in a Business District

Yunn-Kuang Chu
Department of Statistics, National Taichung Institute of Technology, Taiwan

Jeryan Lin*
Department of Statistics, Feng Chia University, Taiwan

Abstract

Burnham and Rexstad (1993) presented a band recovery model that incorporates time-dependent recovery rates and survival rates. In this paper, an estimating function approach is proposed to derive estimators of recovery rates and survival rates. Explicit expressions are given for the derived estimators and their associated asymptotic variances. Examples, including a case for the survival of stalls in the Feng Chia University business district night market, are used to demonstrate the proposed estimating function approach.

Key words: band recovery model; survival rate; stall; business district

JEL classification: C13; M31

1. Introduction

In general, banding studies involve catching, banding, and releasing a sample from a population at a regular interval, usually a calendar year in the bird banding field. Bird banding studies, tracking and estimating both migratory and resident bird populations, are perhaps the most extensive area of application. Indeed, probably over 80 million birds have been banded in North America alone. Interest in this specific subject was engendered by Seber (1970). In addition to biological and ecological sciences, capture-recapture models and estimating function approaches can be applied to other disciplines, for examples, assessing software reliability (Yip, 1995; Voas and McGraw, 1997), introducing precision and rigor in property audits, and estimating the size of an informal venture capital market in small business economics (Kristek, 2002; Mason and Harrison, 2000).

Burnham and Rexstad (1993) proposed an extension of the band recovery model originally presented by Brownie et al. (1985) that incorporates

Received April 2, 2005, revised October 24, 2005, accepted February 9, 2006.

*Correspondence to: Department of statistics, Feng Chia University, 100 Wenhwa Road, Taichung 40724, Taiwan. E-mail: jylin@fcu.edu.tw. This research was supported by the National Science Council, R.O.C., NSC92-2118-M035-002 and Feng Chia University, FCU-RD-93-GB32. The authors would like to thank an anonymous referee and the editor for their insightful comments which significantly enrich this paper.

time-dependent recovery rates and heterogeneous survival rates. They stated that:

> The basic approach to such band-recovery model is that each banded cohort in the analysis of band-recovery data is modeled as a multinomial with individuals classified into cells. The cells are modeled on the basis of probability of surviving the intervening years since banding and the probability of being recovered given death during a particular year.

The authors assumed an ultrastructure form of survival and recovery rates and applied a nonlinear estimation procedure in SAS® (PROC NLIN) to estimate band recovery parameters. In this paper, we apply an estimating function approach to derive estimators of survival and recovery rates without assuming any special forms. Explicit expressions are given for the derived estimators and their associated asymptotic variances. Without using computer program or statistical software packages, we simply compute estimates of band recovery and survival parameters. A real data example for the survival rate of small stalls in a local business district is presented to demonstrate the proposed estimating function approach.

The band recovery approach was originally developed to estimate the survival rate of animals. However, this study proposes application to a business market. In a competitive business district, stores or arcade stalls may fail, be shut down, or for other reasons disappear from the district. In this event, other stores jump in and try to survive the continuing business battle. In the band recovery model we present, we consider a survey framework for collecting data used to estimate survival. Initially, we take a random sample or complete census of existing stalls; the surveyed stalls are then "released" back into in the market. At the beginning of each subsequent survey, we observe which stalls survive, which vanish (are "recovered"), and which vanished stalls are replaced.

The band recovery model incorporates two types of parameters: average survival rates S_j and recovery rates f_j. Recovery rates may be time dependent and survival rates may be heterogeneous. The structure of the model $\{f_t, S_h\}$ is presented in Table 1 with the following notation:

k = number of surveys during which releases are observed
l = number of surveys during which recoveries are observed ($l \geq k$)
N_i = number of new stalls released by the i th survey ($i = 1,...,k$)
R_{ij} = number of stalls recovered by the j th survey by stalls released in the i th survey ($i = 1,...,k$ and $j = 1,...,l$)
f_j = band recovery rate of each stall still operating by survey j
S_j = stall population mean of finite conditional survival rate in survey j ($j = 1,...,l-1$).

In general, band recoveries (banded replacement stalls) continue to be reported until the banding experiment ends. A common objective in most studies is the estimation of time-specific survival rates. Often one has $k = l$ as in the triangular data array shown in Table 1.

Table 1. The Basic Structure of the Band Recovery Model $\{f_t, S_h\}$

Number	Recovery R_{ij} and Expected Number Recovered					
Released	Year 1	Year 2	$\cdots$	Year j	$\cdots$	Year l
N_1	R_{11}	R_{12}		R_{1j}		R_{1l}
	N_1f_1	$N_1S_1f_2$	$\cdots$	$N_1S_1 \cdots S_{k-1}f_j$	$\cdots$	$N_2S_1 \cdots S_{l-2}f_l$
N_2		R_{22}		R_{2j}		R_{2l}
		N_2f_2	$\cdots$	$N_2S_1 \cdots S_{k-2}f_j$	$\cdots$	$N_2S_1 \cdots S_{l-2}f_l$
$\vdots$			$\ddots$	$\vdots$	$\vdots$	$\vdots$
N_k				R_{kj}		R_{kl}
				N_kf_j	$\cdots$	$N_kS_1 \cdots S_{l-k}f_l$

2. Inference via the Estimating Function Approach

In this section, we derive estimates of average survival rates. In general, there are numerous assumptions involved in making inference from banding data. We describe ours in terms of arcade stalls: (1) there is no band loss, (2) survival rates are not affected by banding, (3) the fate of each banded stall is independent of other stalls, (4) the fate of a banded stall in recovery surveys is a multinomial random variable, (5) all banded stalls of an identifiable class have the same annual survival and recovery rates, and (6) annual survival and recovery rates may vary by survey.

For convenience, we let:

$$\tau_j = \min(k-1, l-1), \quad 1 \le j \le l-j,$$

$$l_j = \min(j, k), \quad 1 \le j \le l,$$

$$\mu_i = \begin{cases} 1 & \text{if } i = 0, \\ S_1 \cdots S_i & \text{if } 1 \le i \le l-1. \end{cases}$$

According to the assumptions of the band recovery model $\{f_t, S_h\}$, the random variables of the recovery R_{ij} have a binomial distribution; that is,

$$R_{ij} \sim \text{bin}\left(N_i, \mu_{j-i} f_j\right), \quad 1 \le i \le k, \quad i \le j \le l.$$

Define the following sums of differences:

$$U_j = \sum_{i=1}^{\tau_j} \left(\frac{R_{i,j+1}}{N_i} - \frac{R_{i+1,j+1}}{N_{i+1}} S_j \right), \quad 1 \le j \le l-1,$$

$$W_j = \sum_{i=1}^{l_j} \left(\frac{R_{ij}}{N_i} - \mu_{j-i} f_j \right), \quad 1 \le j \le l.$$

It is easy to see that the sums U_j and W_j have zero mean. This observation leads directly to establishing the following set of unbiased estimating equations to estimate survival rates S_j and recovery rates f_j:

$$U_j = 0\,,\ \ 1 \le j \le l-1\,, \tag{1}$$
$$W_j = 0\,,\ \ 1 \le j \le l\,. \tag{2}$$

Solving the above set of unbiased estimating equations for survival rates S_j and recovery rates f_j, we obtain the following estimators:

$$\tilde{S}_j = \sum_{i=1}^{\tau_j} \frac{R_{i,j+1}}{N_i} \Big/ \sum_{i=1}^{\tau_j} \frac{R_{i+1,j+1}}{N_{i+1}}\,,\ \ 1 \le j \le l-1\,, \tag{3}$$
$$\tilde{f}_j = \sum_{i=1}^{l_j} \frac{R_{ij}}{N_i} \Big/ \sum_{i=1}^{l_j} \tilde{\mu}_{j-i}\,,\ \ 1 \le j \le l\,, \tag{4}$$

where

$$\tilde{\mu}_{j-i} = \begin{cases} 1 & \text{if } i = j, \\ \tilde{S}_1 \cdots \tilde{S}_{j-i} & \text{if } 1 \le i \le j. \end{cases}$$

The set of U_j and W_j determine unbiased estimating functions for S_j and f_j, denoted $\tilde{S}_j$ and $\tilde{f}_j$. For details on the estimating function approach, see Godambe and Heyde (1987).

To explore asymptotic properties of $\tilde{S}_j$ and $\tilde{f}_j$, let $x_{ij} = R_{ij} / N_j$. Then $E(x_{ij}) = \mu_{j-i} f_j$ and $Var(x_{ij}) = \mu_{j-i} f_j (1 - \mu_{j-i} f_j) / N_j$. By the strong law of large numbers (see Chung, 1974, section 5.4), we have

$$x_{ij} \xrightarrow{a.e.} \mu_{j-i} f_j\,,$$

and by the central limit theorem (see Chung, 1974, section 7.1), we obtain

$$\left(x_{ij} - \mu_{j-i} f_j\right) \Big/ \sqrt{Var(x_{ij})} \xrightarrow{D} N(0,1).$$

Hence,

$$\tilde{S}_j = \sum_{i=1}^{\tau_j} x_{i,j+i} \Big/ \sum_{i=1}^{\tau_j} x_{i+1,j+i} \xrightarrow{a.e.} \sum_{i=1}^{\tau_j} \mu_j f_{j+i} \Big/ \sum_{i=1}^{\tau_j} \mu_{j-1} f_{j+i} = S_j$$

Consequently $\tilde{S}_j$ is a strongly consistent estimator of S_j and $\hat{V}ar(x_{ij}) = x_{ij}(1 - x_{ij}) / N_i$ is a strongly consistent estimator of $Var(x_{ij})$ since

$$\sum_{i=1}^{\tau_j} x_{i+1,j+i} \xrightarrow{a.e.} \sum_{i=1}^{\tau_j} \mu_{j-1} f_{j+i}\,,$$

and hence, with probability 1,

$$Var\left(\tilde{S}_j\right) \Big/ Var\left(\sum_{i=1}^{\tau_j} x_{i,j+i} \Big/ \sum_{i=1}^{\tau_j} \mu_{j-1} f_{j+i}\right) \to 1 .$$

Because $\left\{x_{i,j+1}; 1 \le i \le \tau_j\right\}$ are mutually independent random variables,

$$\begin{aligned} Var\left(\sum_{i=1}^{\tau_j} x_{i,j+i} \Big/ \sum_{i=1}^{\tau_j} \mu_{j-1} f_{j+i}\right) &= \left(\sum_{i=1}^{\tau_j} \mu_{j-1} f_{j+i}\right)^{-2} Var\left(\sum_{i=1}^{\tau_j} x_{i,j+i}\right) \\ &= \left(\sum_{i=1}^{\tau_j} \mu_{j-1} f_{j+i}\right)^{-2} \sum_{i=1}^{\tau_j} Var\left(x_{i,j+i}\right). \end{aligned}$$

Therefore,

$$Var\left(\tilde{S}_j\right) = \left(\sum_{i=1}^{\tau_j} x_{i+1,j+i}\right)^{-2} \sum_{i=1}^{\tau_j} \hat{V}ar\left(x_{i,j+i}\right) \tag{5}$$

is a strongly consistent estimator of $Var(\tilde{S}_j)$. Applying the central limit theorem, we deduce that

$$\left(\tilde{S}_j - S_j\right) \Big/ \sqrt{\hat{V}ar(\tilde{S}_j)} \xrightarrow{D} N(0,1).$$

Similarly, we obtain the following properties for the estimated recovery rate $\tilde{f}_j$:

$$\tilde{f}_j \xrightarrow{a.e.} f_j ,$$

$$\hat{V}ar\left(\tilde{f}_j\right) = \left(\sum_{i=1}^{l_j} \tilde{\mu}_{j-i}\right)^{-2} \sum_{i=1}^{l_j} Var\left(x_{ij}\right) \xrightarrow{a.e.} Var\left(f_j\right),$$

and

$$\left(\tilde{f}_j - f_j\right) \Big/ \sqrt{\hat{V}ar(\tilde{f}_j)} \xrightarrow{D} N(0,1).$$

We summarize these results about the survival rates S_j and recovery rates f_j in the following theorem.

Theorem 1**:** Both $\tilde{S}_j$ and $\tilde{f}_j$ are strongly consistent estimators of the survival rate S_j and recovery rate f_j, respectively. In addition, both $\tilde{S}_j$ and $\tilde{f}_j$ are asymptotically normally distributed.

Note that Burnham and Rexstad (1993), assuming a beta distribution on survival rates, presented the estimator

$$\hat{S}_j = \frac{\hat{S}_1 + a(j-1)^b}{1 + a(j-1)^b}, \quad j = 1, 2, \dots, \tag{6}$$

and its standard error

$$\hat{\sigma}_j^2 = \hat{S}_j \left(\hat{S}_{j+1} - \hat{S}_j\right), \quad j = 1, 2, \dots, \tag{7}$$

where a and b need to be solved using a nonlinear program. In contrast, both estimators $\tilde{S}_j$ and $\tilde{f}_j$ derived under the estimating function approach can be easily computed without the aid of computer programs or statistical software.

3. Real Data Examples

In this section, we present two examples to illustrate the estimating function approach. The first example considers real data on banded birds. The second example uses survey data describing the operating cycle of small stalls in a local night market.

***Example 1*:** Brownie et al. (1985, p. 146) gave a set of banded data, reproduced in Table 2a, on adult male mallards. The mallards were banded during winters in Illinois, USA, from 1963 to 1970. Estimated survival rates are reported in Table 2b. Note that $k = l = 8$. The average survival rate after seven years is 52.5% with 95% confidence interval from 34.7% to 70.2%. The mean survival rate during one year following release is 67.4% with standard error 3.3%. The survival rates estimated by Burnham and Rexstad (1993) are also shown in Table 2a in which the mean survival rate during one year is about 67.6% with standard error 14.2%. Clearly, the standard error of our approach is much smaller.

Table 2a. Adult Male Mallards Banded in Winters in Illinois, 1963–1970.

Year	Number	Recovery							
	Banded	1	2	3	4	5	6	7	8
1963	2583	91	89	24	18	16	11	8	7
1964	3075	—	141	45	52	50	17	30	21
1965	1195	—	—	27	31	21	8	19	7
1966	3418	—	—	—	156	92	44	50	49
1967	3100	—	—	—	—	113	68	57	65
1968	2400	—	—	—	—	—	63	52	59
1969	2601	—	—	—	—	—	—	91	80
1970	4433	—	—	—	—	—	—	—	222

Table 2b. Comparison of Survival Estimates of Adult Male Mallards

Parameter	Proposed Survival Rate Estimator			Burnham and Rexstad (1993) Estimator	
	Estimated $\tilde{S}_i$	Standard Error of $\tilde{S}_i$	95% CI	Estimated $\tilde{S}_i$	Standard Error of $\tilde{S}_i$
S_1	0.6740	0.0327	(0.6099, 0.7381)	0.6758	0.1415
S_2	0.6774	0.0396	(0.5998, 0.7550)	0.7055	0.1233
S_3	0.7308	0.0452	(0.6422, 0.8194)	0.7270	0.1105
S_4	0.7613	0.0466	(0.6700, 0.8526)	0.7438	0.1009
S_5	0.5915	0.0542	(0.4853, 0.6977)	0.7575	0.0933
S_6	0.7222	0.0893	(0.5472, 0.8972)	0.7690	0.0870
S_7	0.5248	0.0905	(0.3474, 0.7022)	0.7788	0.0818

***Example 2*:** The Feng Chia University business district (FCBD), located near the university campus in Taichung, Taiwan, is a thriving night market with approximately four hundred small stalls. Due to the highly competitive market environment, stall turnover is considerable.

Stalls were surveyed every six months from December 2001 to June 2003. In the first survey, all stalls were banded as new stalls. Each stall was recorded in the subsequent three surveys as either still operating or vanished. In addition, all new stalls observed were banded and contributed to the total release count for that survey. Stalls that vanish by the next survey comprise the number of banded recoveries. The banded and recovered stall numbers are presented in Table 3.

Table 3. Banded and Recovered Stalls in FCBD

Survey	Number Banded	Recovery* 1	2	3	Estimated $\tilde{S}_i$	Standard Error	95% Confidence Interval
Dec 2001	433	63	74	94	0.5065	0.0523	(0.4040, 0.6090)
Jun 2002	91	—	34	31	0.6373	0.0582	(0.5232, 0.7514)
Dec 2002	44	—	—	28			

*The recovery surveys of vanished stalls were at the beginning of June 2002, December 2002, and June 2003, respectively.

The estimated survival rates S_j by the next survey, as well as standard errors and 95% confidence intervals, are given in Table 3. The mean survival rate during the six months following the initial survey, for example, is 50.7%. The average survival rate after one year is 63.7% with 95% confidence interval 52.3% to 75.1%.

4. Conclusion

In this study, we introduce an unbiased estimating function approach to the band recovery model to derive explicit estimators of population survival and recovery rates. We obtain explicit closed-form expressions for the two estimators and show that these estimators and their variance estimators are strongly consistent and asymptotically normal. The present computation method has no boundary problems and is much less laborious than that of the estimation procedure proposed by Burnham and Rexstad (1993). Moreover, the standard error with the presented estimating function approach is much smaller than their parametric method.

The proposed unbiased estimating function approach is used to estimate survival rates for local stalls in a nearby business district. The results show that approximately two out of three new stalls in Feng Chia University's night market survived at least one year over the experiment period. This indicates that it is a very competitive night market.

References

Brownie, C., D. R. Anderson, K. P. Burnham, and D. S. Robson, (1985), *Statistical Inference from Band Recovery Data: A Handbook*, 2nd edition, U.S. Fish and Wildlife Service Resource Publishing.

Burnham, K. P. and E. A. Rexstad, (1993), "Modeling Heterogeneity in Survival Rates of Banded Waterfowl," *Biometrics*, 49, 1194-1208.

Chung, K. L., (1974), *A Course in Probability Theory*, 2nd edition, New York, NY: Academic Press.

Godambe, V. P. and C. C. Heyde, (1987), "Quasi-Likelihood and Optimal Estimation," *International Statistical Review*, 55, 231-244.

Kristek, M. C., (2002), "Use of Capture/Recapture Estimation in Property Audits," *The Journal of Government Financial Management*, 51, 33-37.

Mason, M. C. and R. T. Harrison, (2000), "The Size of the Informal Venture Capital Market in the United Kingdom," *Small Business Economics*, 15, 137-148.

Seber, G. A. F., (1970), "Estimating Time-Specific Survival and Reporting Rates for Adult Birds from Band Returns," *Biometrika*, 57, 313-318.

Voas, J. M. and G. McGraw, (1997), *Software Fault Injection: Inoculating Program against Errors*, New York, NY: Wiley.

Yip, P., (1995), "Estimating the Number of Errors in a System Using a Martingale Approach," *IEEE Transactions on Reliability*, 44, 322-326.