Short Paper

AIC Introduced to Parsimonious Modelling of Capture-Mark-Recapture Studies

Yoshiharu Matsumiya*

(Received October 9, 1989)

Capture-mark-recapture studies comprise repeated sampling of an animal population in a manner that allows the unique recognition of animals previously caught. The Jolly¹⁾-Seber²⁾ (J-S) model may be regarded as the cornerstone among stochastic multiple-recapture models applicable to open populations with birth, immigration, death, and emigration. Taking survival probabilities, ingress times, and capture probabilities as time-specific allows explicit estimators of parameters to be derived.

Crosbie and Manly³⁾ suggested a number of plausible alternative assumptions for survival probabilities, ingress times, and capture probabilities, and adapted the convention *IJK* to denote a particular model as *model IJK*: *I* relates to survival, *J* relates to births, and *K* relates to captures. The probability of survival can be time-specific (*I*=0), age-specific (*I*=1), or constant (*I*=2); the distribution of ingress times to the population can be time-specific mixed-uniform (*J*=0), or a beta distribution (*J*=1); and the probability of capture can be time-specific (*K*=0) or constant (*K*=1). The most general model, *model 000* is called a "modified J-S model" differing only in its ingress assumption (*J*=0). The Jolly-Seber estimators provided that animals are assumed to enter the population in batches just before sample times. Otherwise, *model 000* does not ignore animals that enter and leave the population before a sample is taken.

Crosbie and Manly³⁾ formulated the maximum likelihood estimates for models with fewer parameters than the J-S model, and estimation has to be done by a numerical optimization algorithm. The parsimonious modelling possibilities were illustrated by the analysis of five sets of real data. The most parsimonious was chosen for five sets, respectively. But, a clear standard for selection of the parsimonious model is not shown in the text.

The Akaike information criterion (AIC) was introduced by Akaike⁴⁾ for the purpose of selecting an

optimal model from within a set of proposed models. The AIC is defined as

$$AIC = -2 \log(\text{maximum likelihood}) + 2m$$

where *m* is the number of estimated parameters. The AIC provides a combined measure of good statistical fit and model parsimony, and the model with the minimum AIC value should generally be selected as the most appropriate.⁵⁾

Table 1 lists the sequence of models fitted to examples (indicates the first two of five examples after Crosbie and Manly³⁾) and AIC values. Using AIC, *model 000* (32 parameters) is chosen as the most parsimonious model in example 3, *model 210* in example 4, and *model 000* (44 parameters) in example 5. The likelihood ratio test with 5% significant level introduces another results (for instances, *model 201* in example 2 and *model 000* with 37 parameters in example 4). The introduction of AIC liberates from the vagueness of a significant level. It is possible and proper to introduce AIC to other references (for example, equal-catchability and trap-response models of Pollock *et al.*,⁶⁾ and simple and Markov models of Sandland and Kirkwood⁷⁾). AIC can be employed to select the "best" model on the parsimonious modelling passing diagnostic checks.

I would like to thank Dr. B. F. J. Manly of the University of Otago (New Zealand) for reviewing the manuscript and giving me important advice.

References

- 1) G. M. Jolly: *Biometrika*, **52**, 225-247 (1965).
- 2) G. A. F. Seber: *Biometrika*, **52**, 249-259 (1965).
- 3) S. F. Crosbie and B. F. J. Manly: *Biometrics*, **41**, 385-398 (1985).
- 4) H. Akaike: in "2nd International Symposium on Information Theory" (ed. by B. N. Petrov and F. Csaki), Akadémiai Kiadó, Budapest, 1973, pp. 267-281.
- 5) Y. Sakamoto, M. Ishiguro, and G. Kitagawa: *Akaike Information Criterion Statistics*, KTK Scientific Publishers, Tokyo, 1986, p. 290.
- 6) K. H. Pollock, J. E. Hines, and J. D. Nichols: *Biometrics*, **40**, 329-340 (1984).
- 7) R. L. Sandland and G. P. Kirkwood: *Biometrika*, **68**, 531-541 (1981).

Table 1. The sequence of models fitted to examples

Example	Model	Number of parameters	Log-likelihood	AIC	$\hat{N}$
1	J-S	12	-57.02	138.04	170
	000	12	-57.20	138.40	186
	100	11	-57.15	136.30	197
	200	10	-57.30	134.60	186
	→210	9	-57.53	133.06*	190
	211	6	-68.02	148.04	204
2	J-S	18	-70.48	176.96	2220
	000	18	-70.56	177.12	3407
	100	15	-72.51	175.02	4289
	101	10	-74.75	169.50*	4721
	→201	9	-76.52	171.04	3344
	211	6	-95.58	203.16	3580

* minimum AIC value (this model is better)

An arrow (→) indicates the most parsimonious model (by Crosbie and Manly).

$\hat{N}$ is total number of animals in the population.

* Ocean Research Institute, University of Tokyo, Tokyo 164, Japan (松宮義晴: 東京大学海洋研究所).
Present address: Faculty of Bioresources, Mie University, Tsu 514, Japan (三重大学生物資源学部).