

Ecological Archives E092-070-A2

Kathryn A. Williams, Peter C. Frederick, and James D. Nichols. 2011. Use of the superpopulation approach to estimate breeding population size: an example in asynchronously breeding birds. *Ecology* 92:821-828.

APPENDIX B. Calculation of standard errors, superpopulation estimates, and confidence intervals for Alley North samples.

We conducted aerial surveys and photo analysis of sample areas from the large (2000 m) Alley North wading bird colony in 2005 (White Ibises) and 2006 (White Ibises and Great Egrets). We calculated a gross superpopulation estimate (N^*) for each sample for however long it was surveyed (Schwarz and Arnason 1996), using Program MARK's POPAN data function (White and Burnham 1999, Arnason and Schwarz 1995). We took the number of nests present in each sample area on the survey date closest to the colony peak count date as the raw count (the equivalent of the peak count for the individual sample). The superpopulation estimates were then compared to the raw counts for all samples in order to obtain a ratio expression of the difference between the two estimates. We divided the superpopulation estimate by the raw count for each sample, to obtain each sample ratio, then averaged these ratios across all samples. We averaged individual ratios rather than finding a summed ratio across all samples (as in Rao 2005), as it is unlikely that there is one underlying level of nest turnover within the colony; different nesting cohorts within the colony are likely to have different levels of nest turnover at different times within the breeding season.

The average ratio of this proportional error across all samples was our estimated colony-wide proportional difference between the peak count and extrapolated superpopulation count, as follows:

$$\frac{\sum_i \left(\hat{N}_i^* / R_i \right)}{n} = \frac{\hat{N}_{colony}^*}{P_{colony}}$$

where \hat{N}_i^* is the superpopulation estimate for sample i ; R_i is the raw count for sample i ; n is the

number of samples in the colony for that year and species. $\hat{\mu} = \frac{\sum_i \left(\hat{N}_i^* / R_i \right)}{n}$ is the average

proportion of the total number of nest starts that were seen in the raw counts. P_{colony} is the peak

count for the colony, and we solved for “ \hat{N}_{colony}^* ,” the entire colony’s extrapolated

superpopulation estimate:

$$\hat{N}_{colony}^* = \frac{(P_{colony}) \sum_i \left(\hat{N}_i^* / R_i \right)}{n}$$

The raw count can be treated as a binomial random variable, and we can thus estimate the approximate variance of the average ratio, $\text{vâr}(\hat{\mu})$, as (e.g., Mood et al. 1970):

$$\text{vâr}(\hat{\mu}) \approx \frac{1}{n^2} \sum_i \left[\frac{\text{vâr}(\hat{N}_i^*)}{R_i^2} \right] = \frac{1}{n^2} \sum_i \left[\left(\frac{(\hat{N}_i^*)^2}{R_i^2} \right) \left(\frac{\text{vâr}(\hat{N}_i^*)}{(\hat{N}_i^*)^2} + \frac{\text{vâr}(R_i)}{R_i^2} \right) \right] = \frac{1}{n^2} \sum_i \left[\frac{\text{vâr}(\hat{N}_i^*)}{R_i^2} + \frac{\hat{N}_i^* (\hat{N}_i^* - R_i)}{R_i^3} \right]$$

This calculation assumes no covariance between the raw count and superpopulation estimate.

Based on a comparison of peak counts to superpopulation estimates for the small Great Egret colonies examined in this study (Table 1), we would suggest that this covariance is probably fairly small, and may be safely excluded from our calculations.

Finally, we compute the conditional (on the peak counts and sample size) variance of the superpopulation estimate as (e.g., Goodman 1960):

$$\text{var}(N_{colony}^*) \approx P_{colony}^2 \text{var}(\hat{\mu}) + \hat{\mu}^2 \text{var}(P_{colony}) - \text{var}(\hat{\mu}) \text{var}(P_{colony}) \approx P_{colony}^2 \text{var}(\hat{\mu}) + \hat{\mu}^2 (P_{colony})(1 - 1/\hat{\mu}) - \text{var}(\hat{\mu})(P_{colony})(1 - 1/\hat{\mu})$$

Thus the average ratio of raw counts to superpopulation estimates for all samples ($\hat{\mu}$), with its associated variance, may be used with the peak count for the entire colony in order to estimate the approximate superpopulation variance for the entire colony. We assumed that the average ratio between the superpopulation estimates of samples and the raw counts of samples is equivalent to the ratio between the superpopulation estimate for the entire colony and the peak count for the colony.

We manually calculated approximate 95% confidence intervals for the superpopulation estimates according to the following estimation procedure, which focuses on f_0 , the number of animals not captured (Chao 1989, Rexstad and Burnham 1992, Williams et al. 2002):

Lower and upper confidence interval bounds for \hat{N}^* : $M_{k+1} + \hat{f}_0 / C$, $M_{k+1} + \hat{f}_0 \cdot C$

where M_{k+1} is the number of individual animals captured in k samples (approximated for large colonies by the peak count number),

$$\hat{f}_0 = \hat{N}_{colony}^* - M_{k+1}$$

$$\text{and } C = \exp \left\{ 1.96 \left[\ln \left(1 + \text{var}(\hat{N}_{colony}^*) / \hat{f}_0^2 \right) \right]^{1/2} \right\}$$

This estimation method assumes that f_0 is log-normally distributed (rather than requiring an assumption of normality for superpopulation estimates, which is often not the case for situations with small numbers of captures and recaptures; Williams et al. 2002).

Example: The 2005 White Ibis samples. The sample raw counts and superpopulation estimates for White Ibis sample areas from the 2005 Alley North colony are shown in Table B1.

Substituting these values into the above equations, we find an average proportional difference between superpopulation estimate and raw count of 3.00. The peak count for this colony in 2005 was 12,750 nests, so we can simply multiply this peak count by $\hat{\mu}$ to find the extrapolated gross superpopulation estimate (38,275 nests). We can estimate the approximate variance of the average ratio, $\text{var}(\hat{\mu})$, as 0.54, and $\text{var}(N_{colony}^*) \approx 8262266$.

95% confidence intervals for the superpopulation estimate may then be calculated as follows:

$$M_{k+1} + \hat{f}_0 / C, M_{k+1} + \hat{f}_0 \cdot C$$

where $M_{k+1} \approx 12750$, $\hat{f}_0 = 25525$, and $C \approx 2.01$, and approximate CIs for the gross superpopulation estimate are thus 25443, 64080.

TABLE B1. Raw Counts and gross superpopulation estimates for White Ibis sample area from the Alley North colony, 2005.

Sample	Raw count	N^*
IT 1	28	51
IT 3	36	67
GST 1	3	19
GST 2	14	30

LITERATURE CITED

- Arnason, A. N., and C. J. Schwarz. 1995. POPAN-4: enhancements to a system for the analysis of mark-recapture data from open populations. *Journal of Applied Statistics* 22:785–800.
- Chao, A. 1989. Estimating population size for sparse data in capture-recapture experiments. *Biometrics* 45: 427–438.
- Goodman, L. A. 1960. On the exact variance of products. *Journal of the American Statistical Association* 55:708–713.
- Mood, A. M., F. A. Graybill, and D. C. Boes. 1974. *Introduction to the theory of statistics*, Third edition. McGraw-Hill, New York, New York, USA.
- Rao, P. S. R. S. 2005. Ratio Estimators—II. Pages 6978–6983 *in* S. Kotz, C. B. Read, N. Balakrishnan, and B. Vidakovic, editors. *Encyclopedia of Statistical Sciences*, 10.
- Rexstad, E. A., and K. P. Burnham. 1992. User's guide for interactive program CAPTURE. Abundance estimation of closed animal populations. Colorado State University, Fort Collins, Colorado, USA. Available at: <http://www.mbr-pwrc.usgs.gov/software/doc/capturemanual.pdf>.
- Schwarz, C. J., and A. N. Arnason. 1996. A general methodology for the analysis of capture-recapture experiments in open populations. *Biometrics* 52:860–873.
- White, G. C., and K. P. Burnham. 1999. Program MARK: Survival estimation from populations of marked animals. *Bird Study* 46S:120–138.
- Williams, B. K., J. D. Nichols, and M. J. Conroy. 2002. *Analysis and management of animal populations*. Academic Press, New York, New York, USA.