Correction of

A simply variance formula for population size estimators by conditioning, Statistical Methodology 5 (2008), 410-423

by Dankmar Böhning

March 5, 2018

I am indebted to Julia Tranninger (Technical University of Graz, Austria) for pointing out the following error in Böhning (2008).

The situation is as follows. A zero-truncated sample of counts of size n is given with frequencies $f_1, f_2, ..., f_m$ where f_x is the frequency of count x. The aim is to provide an estimate of the population size N. The error has occurred in the computation of the variance of the estimator of Chao (1987) given as $n + f_1^2/(2f_2)$.

The error arose when computing $\mathrm{Var}_n\{E_{\hat\lambda_0|n}(n+\hat\lambda_0)\}$ in equation (25) of Böhning (2008). It was falsely assumed that $E_{\hat\lambda_0|n}(n+\hat\lambda_0)$ would be $n+\lambda_0$ and that then λ_0 would disappear when computing $\mathrm{Var}_n(n+\lambda_0)$. But this is *not* so. In fact, for $\hat\lambda_0 = \hat f_0 = f_1^2/(2f_2)$ we have that

$$E_{\hat\lambda_0|n}(n+\hat\lambda_0) = n + E_{\hat\lambda_0|n}(\hat f_0)$$

$$= n + n\frac{p_0}{1-p_0} = n/(1-p_0),$$

where p_0 is the probability for a zero-count. Hence

$$\mathrm{Var}_n(n/(1-p_0)) = \frac{1}{(1-p_0)^2} N p_0(1-p_0) = N\frac{p_0}{1-p_0}.$$

This can be estimated by

$$\widehat{\mathrm{Var}}_n(n/(1-p_0)) = \frac{\hat f_0}{1-\frac{\hat f_0}{n+\hat f_0}} = \hat f_0 + \hat f_0^2/n = \frac{f_1^2}{2f_2} + \frac{f_1^4}{4f_2^2 n}.$$

As

$$\widehat{\mathrm{Var}}_{\hat\lambda_0|n}(n + f_1^2/(2f_2)) = \frac{f_1^3}{f_2^2}\left(1 + \frac{1}{4}\frac{f_1}{f_2}(1-f_2/n)\right),$$

we find the total variance as

$$\widehat{\mathrm{Var}}_{\hat{\lambda}_0|n}(n + f_1^2/(2f_2)) + \widehat{\mathrm{Var}}_n(n/(1-p_0))$$

$$= \frac{f_1^3}{f_2^2} + \frac{1}{4}\frac{f_1^4}{f_2^3} - \frac{1}{4}\frac{f_1^4}{f_2^2 n} + \frac{f_1^2}{2f_2} + \frac{f_1^4}{4f_2^2 n}$$

$$= \frac{f_1^3}{f_2^2} + \frac{1}{4}\frac{f_1^4}{f_2^3} + \frac{f_1^2}{2f_2},$$

which corresponds exactly to the variance estimate given in Chao (1987, p. 786).

References

Böhning, D. (2008). A simply variance formula for population size estimators by conditioning, *Statistical Methodology* **5**, 410–423.

Chao, A. (1987). Estimating the population size for capture-recapture data with unequal catchability. *Biometrics* **43**, 783–791.