Ratio estimator for the Sightability Model

Carl James Schwarz

2020-04-05

Contents

1 Introduction

This vignette documents the development of the ratio estimator in the *SightabilityModel* package (Fieberg, 2012).

The theory for this extension is found in Wong (1996). A portion of her thesis is appended to this document

2 Implementation

Equation 2.4.2 is the ratio estimator formed as the ratio of estimates of the total for variables a and b (e.g. bulls:cow ratio) tat are adjusted using the inverse of sightability (Θ_{ij}) for group j in primary sampling unit i.

The delta-method is used to find the variance of the ratio as a function of the variances of the estimates of the numerator and denominator and the covariance between the two estimates (Equation 2.4.4). The variance of the numerator and denominator can be obtained using the *Sight.Est()* function in the *SightabilityModel* package.l So that all remains is to estimate the covariance between the estimate of the numerator and denominator ($Cov(\widehat{\tau}_a, \widehat{\tau}_b)$)).

The first Equation 2.4.5 gives the formula for the covariance of the numerator and denominator which is estimated as noted just below Equation 2.4.5 as

$$\widehat{Cov}(\widehat{\tau}_a, \widehat{\tau}_b) = \widehat{\tau}_a\widehat{\tau}_b - \widehat{\tau_a\tau_b}$$

NOTICE THE MINUS SIGN between the two terms.

The population value for the second term above ($\tau_a\tau_b$) is given by the equation between the two Equations 2.4.5 which consists of three terms.

This is combined in the second Equation 2.4.5 and Equation 2.4.6. However both equations have two errors.

First, the minus sign above has been dropped, so the proper form of the two estimating equations is:

$$\widehat{Cov}(\widehat{\tau}_a, \widehat{\tau}_b) = \widehat{\tau}_a\widehat{\tau}_b - [term2 + term3 + term4]$$

rather than simply adding the 4 terms together.

Second, the estimator of $\sum_i^N \sum_j^{M_i} a_{ij}b_{ij}$ (term2 above) is

$$\widehat{term2} = \sum_i^N \frac{I_i}{\pi_i} \sum_j^{M_i} Z_{ij} a_{ij} b_{ij} \Theta_{ij}^2$$

i.e. BOTH a_{ij} and b_{ij} must be expanded by the estimate of the inverse of the sightability factor $\widehat{\Theta}_{ij}$ leading to the square of $\widehat{\Theta}$. This is similar to what happens in the rest of the terms.

These corrections were validated using simulation.

3 References

Fieberg, J. (2012). Estimating Population Abundance Using Sightability Models: R SightabilityModel Package. Journal of Statistical Software, 51(9), 1-20. URL http://www.jstatsoft.org/v51/i09/.

Wong, C. (1996). Population size estimation using the modified Horvitz-Thompson estimator with estimated sighting probabilities. Dissertation, Colorado State University, Fort Collins, USA. Available from ProQuest Dissertations & Theses A&I. (304301226). Retrieved from http://proxy.lib.sfu.ca/login?url=https://search-proquest-com.proxy.lib.sfu.ca/docview/304301226?accountid=13800

2.4 Ratio estimator

Samuel et. al. (1992) developed a modified Horvitz-Thompson estimator for a population age or sex ratio. Given the same assumptions as section 2.1, define

τ_a = population size of a type animals,

τ_b = population size of b type animals,

$\hat{\tau}_{a\pi}$ = population size estimator of a type animals when sighting probabilities are known,

$\hat{\tau}_{b\pi}$ = population size estimator of b type animals when sighting probabilities are known,

$\hat{\tau}_{aLR}$ = population size estimator of a type animals when sighting probabilities are unknown and are estimated via a logistic model,

$\hat{\tau}_{bLR}$ = population size estimator of b type animals when sighting probabilities are unknown and are estimated via a logistic model,

a_{ij} = the number of a type animals in the jth group of ith primary unit,

b_{ij} = the number of b type animals in the jth group of ith primary unit,

and with the notation defined in the last section, an estimator for the ratio $R = \tau_a/\tau_b$, using (2.1.2) when sighting probabilities g_{ij} are specified, is

$$\widehat{R}_\pi = \frac{\hat{\tau}_{a\pi}}{\hat{\tau}_{b\pi}} = \frac{\sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j=1}^{M_i} Z_{ij} a_{ij} \Theta_{ij}}{\sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j=1}^{M_i} Z_{ij} b_{ij} \Theta_{ij}}. \tag{2.4.1}$$

Using (2.1.4) when the $\Theta_{ij} = 1/g_{ij}$'s are unknown and are estimated via a logistic model, the ratio estimator is

$$\widehat{R}_{LR} = \frac{\hat{\tau}_{aLR}}{\hat{\tau}_{bLR}} = \frac{\sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j=1}^{M_i} Z_{ij} a_{ij} \widehat{\Theta}_{ij}}{\sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j=1}^{M_i} Z_{ij} b_{ij} \widehat{\Theta}_{ij}}. \tag{2.4.2}$$

By applying the delta method,

$$\frac{\hat{\tau}_a}{\hat{\tau}_b} \doteq \frac{\tau_a}{\tau_b} + \frac{1}{\tau_b}(\hat{\tau}_a - \tau_a) - \frac{1}{{\tau_b}^2}(\hat{\tau}_b - \tau_b), \tag{2.4.3}$$

In (2.4.3), $\hat{\tau}_a$, $\hat{\tau}_b$ are $\hat{\tau}_{a\pi}$, $\hat{\tau}_{b\pi}$ (see (2.1.2)) if sighting probabilities are known and they are $\hat{\tau}_{aLR}$, $\hat{\tau}_{bLR}$ (see (2.1.4)) if sighting probabilities are unknown. Thus, we obtain an approximate variance of a ratio estimator as

$$\mathrm{var}(\widehat{R}) \doteq R^2\left[\frac{\mathrm{var}(\hat{\tau}_a)}{{\tau_a}^2} + \frac{\mathrm{var}(\hat{\tau}_b)}{{\tau_b}^2} - \frac{2\mathrm{cov}(\hat{\tau}_a, \hat{\tau}_b)}{\tau_a\tau_b}\right]. \tag{2.4.4}$$

To obtain an (approximate) estimator of $\mathrm{var}(\widehat{R})$ in (2.4.4), we replace the ratio, the population sizes, variances and covariances of the population size estimator by their respective unbiased estimators. We have already derived (approximate and/or asymptotic) unbiased estimators for the ratio, the population sizes (see (2.1.2) and (2.1.4)) and variances of the population size estimator (see (2.2.2) and (2.2.13)). Therefore, to extend the results of Samuel et. al. (1992), we only need to find an unbiased estimator for the covariance term. To do so, we write

$$\mathrm{cov}(\hat{\tau}_a, \hat{\tau}_b) = E[\hat{\tau}_a\hat{\tau}_b] - E[\hat{\tau}_a]E[\hat{\tau}_b] = E[\hat{\tau}_a\hat{\tau}_b] - \tau_a\tau_b. \tag{2.4.5}$$

Then, the expression $\hat{\tau}_a\hat{\tau}_b - \hat{\tau}_{ab}$, where $\hat{\tau}_{ab}$ is an unbiased estimator of $\tau_a\tau_b$, is an unbiased estimator of the covariance term in (2.4.5). To find an unbiased estimator of $\tau_a\tau_b$, we note that

$$\tau_a\tau_b = \left(\sum_{i=1}^{N}\sum_{j=1}^{M_i} a_{ij}\right)\left(\sum_{i=1}^{N}\sum_{j=1}^{M_i} b_{ij}\right) = \sum_{i=1}^{N}\sum_{j=1}^{M_i} a_{ij}b_{ij} + \sum_{i=1}^{N}\sum_{j=1}^{M_i}\sum_{j'=1}^{M_{i'}} a_{ij}b_{ij'} + \sum_{i\neq i'}^{N}\sum_{j=1}^{M_i}\sum_{j'=1}^{M_{i'}} a_{ij}b_{i'j'}.$$

Hence, an unbiased estimator for the covariance term in (2.4.5) is

$$\widehat{\mathrm{cov}}(\hat{\tau}_a, \hat{\tau}_b) = \hat{\tau}_{a\pi}\hat{\tau}_{b\pi} + \sum_{i=1}^{N}\frac{I_i}{\pi_i}\sum_{j=1}^{M_i}\frac{Z_{ij}a_{ij}b_{ij}}{g_{ij}} + \sum_{i=1}^{N}\frac{I_i}{\pi_i}\sum_{j\neq j'}^{M_i}\frac{Z_{ij}Z_{ij'}a_{ij}b_{ij'}}{g_{ij}g_{ij'}} + \sum_{i\neq i'}^{N}\frac{I_iI_{i'}}{\pi_{ii'}}\sum_{j=1}^{M_i}\sum_{j'=1}^{M_{i'}}\frac{Z_{ij}Z_{i'j'}a_{ij}b_{i'j'}}{g_{ij}g_{i'j'}}. \tag{2.4.5}$$

This needs a square exponent

for the case of known g_{ij}'s.

Change this + sign to a minus sign

For the case of unknown Θ_{ij}'s estimated via a logistic model. an unbiased estimator for the covariance term in (2.4.5) is

This needs a square exponent

$$\widehat{\text{cov}}(\hat{\tau}_a, \hat{\tau}_b) = \hat{\tau}_{aLR}\hat{\tau}_{bLR} + \sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j=1}^{M_i} Z_{ij} a_{ij} b_{ij} \widehat{\Theta}_{ij}$$

$$+ \sum_{i=1}^{N} \frac{I_i}{\pi_i} \sum_{j \neq j'}^{M_i} Z_{ij} Z_{ij'} a_{ij} b_{ij'} [-\widehat{\text{cov}}(\widehat{\Theta}_{ij}, \widehat{\Theta}_{ij'}) + \widehat{\Theta}_{ij}\widehat{\Theta}_{ij'}]$$

$$+ \sum_{i \neq i'}^{N} \frac{I_i I_{i'}}{\pi_{ii'}} \sum_{j=1}^{M_i} \sum_{j'=1}^{M_{i'}} Z_{ij} Z_{i'j'} a_{ij} b_{i'j'} [-\widehat{\text{cov}}(\widehat{\Theta}_{ij}, \widehat{\Theta}_{i'j'}) + \widehat{\Theta}_{ij}\widehat{\Theta}_{i'j'}]. \quad (2.4.6)$$

Change this` + sign to a minus sign

Notice the brackets