

1 Derivation of dispersion approximation

Buckland et al. (2015) describes how to approximate the number of sampling point transects to achieve desired precision given information from a pilot study.

$$K = \frac{K_0 \{cv(\hat{D})\}^2}{\{cv_t(\hat{D})\}^2} \quad (1)$$

where

- K_0 number of points in pilot study,
- $cv(\hat{D})$ coefficient of variation from pilot study and
- $cv_t(\hat{D})$ desired (target) coefficient of variation

If the pilot study was too small to produce an estimate of \hat{D} so there is no coefficient of variation from the pilot study, an alternative formulation could be employed:

$$K = \frac{K_0}{n_0} \frac{b}{\{cv_t(\hat{D})\}^2} \quad (2)$$

where

- n_0 number of detections in pilot study.

The two expressions can be equated and solved for b

$$\begin{aligned} \frac{K_0 \{cv(\hat{D})\}^2}{\{cv_t(\hat{D})\}^2} &= \frac{K_0}{n_0} \frac{b}{\{cv_t(\hat{D})\}^2} \\ b &= n_0 \cdot \frac{cv(\hat{D})^2}{\{cv_t(\hat{D})\}^2} \cdot cv_t(\hat{D})^2 \\ &= n_0 \cdot cv(\hat{D})^2 \end{aligned} \quad (3)$$

Applying this formula to the daytime duiker data produces a \hat{b} of $10284 \cdot 0.27^2 = 750$.

2 Alternative formulation

Buckland et al. (2001) presents an alternative formulation to approximate b Eqn. 7.2:

$$b \doteq \left\{ \frac{\text{var}(n)}{n} + \frac{n \cdot \text{var}\{\hat{f}(0)\}}{\{f(0)\}^2} \right\} \quad (4)$$

Applying this formula to Howe et al. (2017) duiker data, in which $\text{var}(n)$ is computed as the variance in detections across camera locations. Fig. 1 of Howe et al. (2017) shows 23 locations, but Fig. 3 of the same paper shows 21 camera locations. The data from the file `DaytimeDistances.txt` holds data from 21 locations and seems consistent with Fig. 3, with stations B1 and E4 having fewer than 5 detections.

Components of Eqn. 4 from the duiker data are

- $\text{var}(n) = 411911$
- $n = 10284$
- $\frac{\text{var}\{\hat{f}(0)\}}{\{f(0)\}^2} = cv(\hat{f}(0))^2 = 0.0394^2 = 0.00155236$ (for UNI:cos3 model)

This produces an estimate of the dispersion parameter of 56. Allowing for possible miscalculation of the number of stations and not accounting for loss of detections arising from truncation during analysis would not account for the discrepancy between \hat{b} computed using Eqn. 4 (56) and the estimate (750) produced using Eqn. 3.

The Burnham et al. (1980) monograph discusses the derivation of \hat{b}

The 2 equations needed are

$$E(n) = 2LD/f(0)$$

and

$$(cv(\hat{D}))^2 = \frac{\text{var}(n)}{(E(n))^2} + \frac{\text{var}(\hat{f}(0))}{(f(0))^2}$$

where $cv(\hat{D})$ is the coefficient of variation of \hat{D} and $\text{var}(\cdot)$ denotes a sampling variance. To a first approximation, $\text{var}(n) = a_1 n$ and $\text{var}(\hat{f}(0)) = (f(0))^2 a_2 / n$ can be written. Those relations are justified in many models and have been confirmed by experience. The constants a_1 and a_2 are unknown parameters, and it is important to note that in a given study they are independent (or almost so) of n, L and $f(0)$. Thus, replacing $E(n)$ by just n ,

$$(cv(\hat{D}))^2 = \frac{1}{n} [a_1 + a_2] = \frac{b}{n},$$

I let the issue rest with their justification that

$$b = a_1 + a_2$$

and this relationship has been "confirmed by experience."