

Estimating the Number of Animals:

Consistency of Population Density Estimates by the Plot-Removal Method

D. S. Robson

Abstract

The 8 November, 1968 issue of Science contains a report by W. R. Hanson on a new sampling design and estimation formulas for estimating animal population density in a study area. Two alternative estimators are given, one of which is intended for use only in situations where density is "essentially uniform" throughout the area. This estimator is shown here to be consistent if and only if the spatial distribution of animals is Poisson; the other and supposedly more robust estimator is consistent if and only if the probability distribution of number of animals per plot is a distribution on the two integers 0 and 2.

Submitted to Science for publication
November, 1968

Estimating the Number of Animals:

Consistency of Population Density Estimates by the Plot-Removal Method

Hanson (1) has presented two alternative formulas for estimating animal density (k) in a study area sampled by his proposed "plot-removal method". One formula is intended to apply only when the density of animals is "essentially uniform" throughout the area in question while the other is claimed to be more generally applicable. From a statistical viewpoint the criterion most widely adopted as a necessary condition for applicability is the property of consistency; as sample size increases a consistent estimator of k will come arbitrarily close to k with probability approaching 1. Conditions for consistency of Hanson's estimators are determined here under the simplifying assumptions that the animals do not move and that during any one survey of the study area each animal has an equal and independent probability p of being observed.

The study area is assumed to be partitioned into L equal plots, and the plot-removal method consists of a sequence of cursory surveys of those plots on which no animals have yet been seen. Initially all L plots are surveyed and a total of x_1 animals are seen on l of these plots; x_2 animals are then seen in the second count which surveys only the $L-l$ plots on which no animals were previously seen, and so on. The unknown mean density k is to be estimated from data collected in a sequence of $n \geq 2$ such surveys on successively fewer plots. For the case $n = 2$ Hanson gives two alternative estimators of k :

$$\hat{k}_1 = \frac{(x_1/L)^2[1 - (l/L)]}{(x_1/L)[1 - (l/L)] - (x_2/L)} \quad \hat{k}_2 = \frac{x_1/L}{1 - \left(\frac{x_2/L}{x_1/L}\right)}$$

If the relative frequency of plots containing z animals approaches the probability distributions $f_k(z)$ as the number of plots (L) is increased then,

by the law of large numbers, the random variables x_1/L , x_2/L and t/L converge (in probability) to

$$x_1/L \rightarrow kp \quad x_2/L \rightarrow pq \frac{d}{dq} Q_k(q) \quad t/L \rightarrow 1 - Q_k(q)$$

where $q = 1 - p$ and $Q_k(s)$ is the probability generating function

$$Q_k(s) = \sum_z s^z f_k(z)$$

with $Q'_k(1) = k$. Thus, \hat{k}_1 and \hat{k}_2 converge in probability to

$$\hat{k}_1 \rightarrow \frac{k^2 p}{k - q \frac{d}{dq} \log Q_k(q)} \quad \hat{k}_2 \rightarrow \frac{kp}{1 - \left[\frac{q}{k} \frac{d}{dq} Q_k(q) \right]}$$

and so $\hat{k}_1 \rightarrow k$ if and only if

$$Q_k(q) = e^{-k(1-q)}$$

that is, if and only if the number of animals per plot follows a Poisson distribution with mean k . The estimator \hat{k}_2 converges on k if and only if

$$Q_k(q) = 1 - (1 - q^2)(k/2)$$

that is, if and only if the number of animals on a plot is either 0 or 2 with probabilities $1 - (k/2)$ and $k/2$, respectively.

This quaint result disqualifies \hat{k}_2 as an estimator of k and apparently reflects the unsoundness of the basic relation which Hanson used in deriving \hat{k}_2 ; namely:

$$\frac{\text{Number unobserved in } l \text{ plots}}{\text{Number unobserved in } L \text{ plots}} = \frac{\text{number observed in } l \text{ plots}}{\text{number present in } L \text{ plots}}.$$

Note that in the Poisson case \hat{k}_2 is an estimate of $kp / [1 - \sqrt{qe^{-kp}}]$, which is approximately $2k/(k+1)$ when p and kp are both small.

With appropriate modifications to eliminate negative estimates, \hat{k}_1 is both consistent and efficient in the Poisson case, since it is then the maximum likelihood estimator -- but only in the Poisson case. If, for example, the frequency distribution of number of animals per plot is negative binomial with mean k and variance σ^2 ($\sigma^2 > k$) then \hat{k}_1 estimates $k[p + q(k/\sigma^2)]$, which is less than k . The behavior of this estimator is also influenced by any intra-plot correlation among the observations; in the extreme case where either all animals in a plot are observed (with probability p) or all are in hiding (with probability $q = 1 - p$) then under the Poisson model \hat{k}_1 estimates $k[p + qe^k]$, which is greater than k . Likewise, movement of animals between plots can influence the behavior of \hat{k}_1 ; if there is complete reshuffling between the counts x_1 and x_2 then \hat{k}_1 becomes arbitrarily large with probability approaching 1 as L increases. If none of these complications arise but the investigator is more attentive in surveying the $L-l$ plots first observed to be empty then in the Poisson case \hat{k}_1 estimates

$k/[1 - \epsilon(q_1/p_1)]$ where ϵ is the relative change in p between the first and second surveys, $\epsilon = (p_2 - p_1)/p_1$. Thus it appears that \hat{k}_1 , as well as \hat{k}_2 , could be subject to serious bias in practice.

Douglas S. Robson
Biometrics Unit, Cornell University
Ithaca, N.Y.

References

1. W. R. Hanson, Science 162, 675 (1968).