A SIMPLE FORMULA FOR CALCULATING THE VARIANCE OF PRODUCTS AND DIVIDENDS

DOUGLAS C. HEARD

DEPARTMENT OF RENEWABLE RESOURCES

GOVERNMENT OF THE NWT

YELLOWKNIFE, NWT

1987

Northwest Territories Renewable Resources

Manuscript Report 20

ABSTRACT

This report presents a simple general formula for determining the variance of products or dividends of uncorrelated random variables.

TABLE OF CONTENTS

INTRODUCTION

Ecologists frequently need to estimate the product of two or more other estimates. For example, to estimate the number of pregnant caribou cows on a calving ground (w), I multiply the estimate of the calving ground population (x) by the proportion of those animals estimated to be pregnant females (y). Similarly, it may be of interest to estimate survival of gyrfalcons over their first year of life (w) by multiplying the estimate of egg and nestling survival (x) by the estimate of survival from fledging to their first birthday (y). In both cases, $w = xy$.

Because x and y are estimates, "random variables" to a statistician, they each have an associated variance, V(x) and V(y). Estimating the variance of a random variable depends on the statistic and many texts cover those methods (e.g., Sokal and Rohlf 1969, Seber 1973, Ricker 1975). The problem addressed in this paper is how to estimate the variance of the product or dividend of two or more independent random variables (e.g., V(w)).

RESULTS

The general formula for dealing with functions of random variables was given to me by Dr. Conrad Wehrhahn (University of British Columbia, Vancouver, B.C.), but see also Goodman (1960).

If

$w = f(x,y)$,

then

$$V(w) = [(df/dx)^2]\ V(x) + [(df/dy)^2]\ V(y) + (df/dx)(df/dy)Cov(x,y)$$

(equation 1)

For example

if $w = cx$, $V(w) = c^2\ V(x)$ (equation 1a)

if $w = x + y$, $V(w) = V(x) = V(y)$ (equation 1b)

if $w = xy$, $V(w) = x^2V(y) + y^2V(x)$ (equation 1c)

if $w = x/y$, $V(w) = (1/y)^2V(x) + (x^2/y^4)V(y)$ (equation 1d)

where c is a constant and x and y are independent (uncorrelated) estimates (i.e.) $Cov(x,y) = 0$).

If the estimates are correlated then the variance will be over or underestimated (see equation 1). This bias will probably be insignificant if either of the coefficients of variation are small (see Goodman 1960).

I make a judgement as to the probability of the variables being correlated. When the covariance term cannot be ignored other methods, e.g., Monte Carlo simulations can be used to empirically describe the distribution of the product (Rubinstein 1981).

Equation 1 gets messy when dealing with dividends (e.g., equation 1d and Heard 1981) and when dealing with functions of more than two variables (e.g., $w = xyz$ or $w = x/yz$ or $w = (xy/ab)$. The calculations can be simplified by dealing with coefficients of variation, where $CV(x) = \sqrt{V(x)}/x$.

The general formula for the variance of products and dividends involving any number of variables is:

$$[CV(w)]^2 = [CV(x)]^2 + [CV(y)]^2 + [CV(z)]^2 + \ldots \qquad \text{(equation 2a)}$$

or

$$V(w) = (w)^2 ([CV(x)]^2 + [CV(y)]^2 + [CV(z)]^2 + \ldots) \qquad \text{(equation 2b)}$$

where

$$w = [(x)(y)]/(z) \ldots$$

or

$$w = (x)/[(y)(z) \ldots]$$

That is, w can be a function of any number of uncorrelated variables multiplied or divided in any way.

Proof of the two variable case:

If

$$w = x/y$$

Then from equation 1c

$$V(w) = (1/y^2)\ V(x) + (x^2/y^4)\ V(y) \qquad \text{(equation 3a)}$$

Because

$CV(x) = \sqrt{V(x)}/(x)$ and $CV(y) = \sqrt{V(y)}/(y)$,

It is possible to substitute for V(x) and V(y), in equation 3a where

$V(x) = (x^2)([CV(x)]^2)$ and $V(y) = (x^2)([CV(y)]^2)$

producing

$$V(w) = (1/y^2)(x^2)([CV(x)]^2) + (x^2/y^4)(y^2)([CV(y)]^2)$$

$$= (x^2/y^2)([CV(x)]^2 + [CV(y)]^2) \qquad \text{(equation 3b)}$$

$$= (w^2)([CV(x)]^2 + [CV(y)]^2) \qquad \text{(equation 2b)}$$

NOTE: equation 2b cannot be used for addition and subtraction functions.

Example (from Heard 1981: 36-37)

Variable	Estimate	Variance	Coefficient of variance
a. calving ground estimate	16503	3747000	0.1173
b. proportion of breeding females on the calving ground	0.896	0.00145	0.0425
c. proportion of females in the population	0.604	0.00215	0.0768
d. proportion of females that breed	0.694	0.00120	0.0499

population estimate

$= ab/cd$

$= 35276$

population variance

$= (35276)^2 (0.1173^2 + 0.0425^2 + 0.0768^2 + 0.0499^2)$

$= 29.81 \times 10^6$

Heard (1981:37) reported a population variance of 29.78×10^6. The difference is due to rounding errors.

ACKNOWLEDGEMENT

I thank Mitch Taylor for his constructive comments on this paper.

LITERATURE CITED

Goodman, L.A. 1960. On the exact variance of products. J. Amer. Statistical Assoc. 55: 708-713.

Heard, D.C. 1981. An estimate of the size and structure of the Kaminuriak caribou herd in 1977. NWT Wild. Serv. File Rep. No. 17. 37 pp.

Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Canada Dept. of the Environment, Fisheries and Marine Service, Ottawa, Bull. 191. 382 pp.

Rubinstein, R.Y. 1981. Simulation and the Monte Carlo method. John Wiley and Sons. New York. 279 pp.

Seber, G.A.F. 1973. The estimation of animal abundance. Hafner Press, New York. 506 pp.

Sokal, R.R. and R.L. Rohlf. 1969. Biometry. W.H. Freeman Co., San Francisco. 776 pp.