# DENSITY ESTIMATION BY VARIABLE AREA TRANSECT

KEITH R. PARKER, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Southwest Fisheries Center, La Jolla, CA 92038

*Abstract:* The frequency distribution of the distance from a randomly selected point to the $r$th nearest organism contained in a variable area transect is described. Estimators of density and variance are derived for both randomly distributed and aggregated populations. For randomly distributed populations exact confidence intervals are available. On an effort-precision criterion the method is more useful than other spatial distribution based density estimators. Under the condition where effort is a linear function of distance covered and numbers of organisms counted, the effort expended in achieving a desired level of precision is the same as that for quadrats; the variance of effort depends on the difficulty in counting and in covering the distance. An example using pismo clams (*Tivela stultorum*) is given: the density of clams was found to be 1 per 2.87 m$^2$. The method was found to be superior to that of other distance-measures and quadrats.

*J. WILDL. MANAGE. 43(2):484–492*

Two methods are commonly employed when measuring the density of plant communities and closed animal populations where the animals are relatively stationary: quadrats and distance-measures. Quadrats are simple to apply. The fact that means of plot samples provide an unbiased estimate of true mean density enhances the usefulness of quadrats. Unfortunately, problems arise in quadrat sampling. Quadrats of different size yield different estimates of density (Greig-Smith 1964). Also, the degree of nonrandomness detected depends on plot size (Pielou 1957). Distance-measure methods, where the density is based on the spatial distribution of organisms, avoid these difficulties. The density estimate is based on the distance between some randomly chosen point and some $r$th closest organism and not on a fixed plot.

Unfortunately, there are also problems in using distance-measures. The precision of the estimate is inverse to the product of the number of samples taken, $n$, and the number in each sample, $r$ (Pollard 1971). In the field, an $r$ beyond 3 or 4 is often impractical. This difficulty arises not from the measurement to the $r$th organism but in determining which is the $r$th. Increased precision is afforded only at the expense of increasing the number of sample points. This has discouraged many investigators from using distance-measures.

One way to avoid the difficulties of quadrats and distance-measures is to sample in such a way that the search for the $r$th organism is a consequence of traversing a transect. In this way searching and measuring the distance are accomplished at the same time. A large sample size is easily obtained and the estimate is still free of the restrictions imposed by a fixed plot size.

The author would like to thank L. L. Eberhardt for his suggestions in the original formulation of the problem and D. G. Chapman for his suggestions, especially those on the comparison of methods. Personnel from the California Department of Fish and Game participated in data collection.

## ESTIMATION

### Maximum Likelihood Estimation of Density

Consider the following: A resource management biologist proceeds along a transect of constant and fixed width, $w$, until he reaches the $r$th organism. If the

organisms are Poisson randomly distributed with constant density, the probability of $r$ or more organisms being contained in a plot of area, $wx$, where $x$ is the measure of distance to the $r$th organism, is

$$F_r(x) = \sum_{i=r}^{\infty} (\lambda w x)^i \exp(-\lambda w x)/i!. \quad (1)$$

The density function, upon differentiating with respect to $x$, is

$$F_r(x) = (\lambda w)^r x^{r-1} \exp(-\lambda w x)/(r-1)!. \quad (2)$$

If we have $x_1, x_2, \ldots, x_n$ sample distances from $n$ sample point then the modified (unbiased) maximum likelihood estimator of the density is (see Appendix, Part A)

$$\hat{\lambda} = \frac{nr-1}{w \sum_{i=1}^{n} x_i}, \quad (3)$$

where the variance is

$$\operatorname{var} \hat{\lambda} = \frac{\lambda^2}{nr-2}. \quad (4)$$

For the variance to be finite the product $nr$ must be greater than 2.

### Confidence Intervals

Often an investigator is interested in knowing not only an estimate of the density but also an interval of confidence in which it is expected to occur. Equation (2) is used to derive the characteristic function of $x$ which is used to construct an exact confidence interval for the density. The characteristic function is unique for any distribution. Thus, for a given density function the characteristic function can be used to index the underlying distribution. Once the underlying distribution is known the confidence interval can be readily identified as will be presently shown.

By definition, the characteristic function of $x$ the distance to the $r$th organism, is the expected value of $\exp^{itx}$, i.e.,

$$\phi_x(t) = E[\exp(itx)]$$

which from equation (2) is

$$= \int_0^{\infty} \frac{(\lambda w)^r}{(r-1)!} x^{r-1} \cdot \exp[-x(\lambda w - it)]\, dx$$

$$= (1 - it/\lambda w)^{-r}. \quad (5)$$

The characteristic function for the chi-square distribution with $m$ degrees of freedom (df) is

$$\phi(t) = (1 - i2t)^{-m/2}.$$

Thus from equation (5), $2\lambda wx$ is chi-square distributed with $2r$ df and since the sum of a chi-square is a chi-square with df added $2\lambda wn\bar{x}$ is chi-square with $2nr$ df. This statistic can be used to construct a confidence interval about $\lambda$.

For a $1 - \alpha$ percent confidence interval

$$\Pr(C_1 \leq 2\lambda wn\bar{x} \leq C_2) = 1 - \alpha$$

or

$$\Pr(C_1/2wn\bar{x} \leq \lambda \leq C_2/2\lambda wn\bar{x}) = 1 - \alpha, \quad (6)$$

where $C_1$ and $C_2$ are the lower and upper $\alpha/2$ percentage points from a chi-square distribution with $2nr$ df.

## COMPARISON OF METHODS

### Variable Area Transect and Closest-Individual

The variance of the modified maximum likelihood estimator is the same for the variable area transect (equation 2) and the closest-individual method (Pollard 1971):

$$\operatorname{Var}(\lambda) = \lambda^2(nr-2). \quad (7)$$

Under the same sampling conditions an

equal amount of effort is expended in locating $n$ random sample points for both methods. So, in deciding which method requires the least effort to attain a given sample size, and hence, variance, the relative difficulty in measuring $r$ should be examined.

For the closest-individual method, an $r$ beyond 3 or 4 is often impractical; the difficulty arises from determining which organism is the $r$th. If the investigator searches either in increasing circles or works back and forth across the sample area, a great deal of time and effort is expended. Only when the $r$th organism can be determined from the sample point does the closest-individual method require little searching effort.

For the variable area transect the determination of the $r$th organism is a consequence of traversing the transect. Less effort is likely to be expended in determining the order of organisms. Consequently, in most field applications searching effort should be less.

Difficulties might arise in determining transect boundaries. In some situations these boundaries could be laid out as with quadrats, 3 of the sides being fixed with the 4th being variable. Practical alternatives would depend on the terrain. In some situations the investigator could carry a rod and allow the ends of the rod to determine the side boundaries. In any event, the transect should be narrow enough so that it is easy to determine whether or not an organism lies inside the transect boundaries.

## Variable Area Transect and Quadrats

In terms of the amount of effort expended in achieving a desired level of precision, the variable area transect and quadrats can be compared under 2 conditions: (1) the major portion of effort is expended in the examination of the area for specimens and not in either determining boundaries or locating sample points, and (2) the effort invested is assumed to be a linear function of the distance covered and the number of organisms counted.

Under the first condition, Holgate (1964) shows a natural way to compare quadrats and the closest-individual method. This reasoning can be extended to variable area transects as well.

The variance of the density estimator, say $\mathring{\lambda}$, for quadrats is

$$\mathrm{Var}(\mathring{\lambda}) = \lambda/a \qquad (8)$$

where $a$ is the total area sampled.

For the variable area transect the total expected area for $n$ samples can be derived from equation (2):

$$\begin{aligned} E(a) &= nwE(x) \\ &= nw\int_0^\infty (\lambda w)^r x^r \exp(-\lambda w x)/(r-1)! \\ &= nr/\lambda. \end{aligned}$$

Substituting $E(a)$ for $a$ in equation (8),

$$\mathrm{Var}(\mathring{\lambda}) = \lambda^2/nr, \qquad (9)$$

and since from equation (4)

$$\mathrm{Var}(\hat{\lambda}) = \lambda^2/(nr-2)$$

the variance of a quadrat estimator is smaller. However, for large sample sizes the variances can be assumed to be nearly equal.

Under the second set of conditions we weigh whether less effort is required either to count a random number of organisms over a fixed area or to determine the area for a fixed number of organisms.

Let the following linear relationship describe effort:

$$\begin{aligned} \text{effort} = {} & d\cdot(\text{distance covered}) \\ & + b\cdot(\text{number counted}), \end{aligned}$$

where $d$ and $b$ are arbitrary constants reflecting the difficulty in covering the distance and in counting the organisms, respectively.

If $E(T_1)$ is the expected effort for the variable area transect then:

$$E(T_1) = dnE(x) + b(nr) = drn/(\lambda w) + b(nr). \quad (10)$$

If $E(T_2)$ is the expected effort for quadrats then

$$E(T_2) = da/w + bE(k) = da/w + b\lambda a, \quad (11)$$

where $k$ is the number counted.

For convenience, letting the width, $w$, equal unity and the area, $a$, equal $nr/\lambda$, the expected efforts are equal:

$$E(T_1) = E(T_2) = nr(d/\lambda + b). \quad (12)$$

Since effort as defined here is subject to statistical error, we should also investigate its variance.

The variance of $T_1$ is

$$\mathrm{Var}(T_1) = d^2\mathrm{var}(x) + 0. \quad (13)$$

The variance of $x$ is

$$\mathrm{Var}(x) = [E(x)]^2 - [E(x^2)] = nr/(\lambda w)^2, \quad (14)$$

and therefore

$$\mathrm{Var}\, T_1 = d^2nr/(\lambda w)^2. \quad (15)$$

The variance of $T_2$ is

$$\mathrm{Var}(T_2) = 0 + b^2\mathrm{var}(k) = b^2\lambda a \quad (16)$$

since $k$ is a Poisson random variable.

Again, letting $w$ equal unity and $a = nr/\lambda$,

$$\mathrm{Var}(T_1) = d^2nr/\lambda^2, \quad (17)$$

$$\mathrm{Var}(T_2) = b^2nr. \quad (18)$$

The variances are equal when

$$\lambda = d/b. \quad (19)$$

Without prior knowledge of the area to be sampled it is difficult to assign relative values to $b$ and $d$. However, we can consider 2 cases where a comparison is easily made.

If counting is difficult, relative to that in covering the distance ($b > d$) and if the population is dense, then the variable area transect should have the least variance of effort. So, in this situation, the variance of effort would be more consistent using the variable area transect.

In the opposite situation, where the distance is difficult to cover relative to the difficulty in counting the organisms, $d > b$, and the density is low, quadrats offer the effort of least variance.

## AN EXAMPLE

Pismo clams (*Tivela stultorum* Mawe) inhabit open sandy beaches from the intertidal zone to beyond the surf line. The density of clams on the portion of the beach utilized by recreational clam diggers is valuable management information.

The following example is from data collected at Seal Beach, California, during the midday low tide of 16 February 1977. Transect lines were laid from randomly selected points on the exposed beach. Along the lines, the beach was probed with a 6-prong clamming fork until the 3rd clam was reached. Because of heavy surf an $r$ greater than 3 was not practical.

The width of the fork, 0.23 m, is the transect width. The measured distances to the 3rd clam are listed in Table 1. Fourteen samples were taken.

The sum of distances to the 3rd clam is

$$\sum_{i=1}^{14} x_i = 510.8.$$

Since $nr = 42$, the density estimate from

Table 1. The distance from a randomly chosen point to the third organism, $X_i$, in meters. The $Z$ statistic and the empirical distribution function, $S_n$, used for testing the hypothesis of randomness.

| $X_i$ | 5.4 | 15.3 | 33.2 | 7.8 | 43.0 | 60.5 | 49.8 | 63.4 | 42.3 | 35.8 | 65.8 | 15.3 | 13.9 | 59.3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $Z_i$ | 0.07 | 0.14 | 0.21 | 0.28 | 0.36 | 0.42 | 0.50 | 0.57 | 0.64 | 0.71 | 0.78 | 0.85 | 0.92 | 1.00 |
| $S_n$ | 0.01 | 0.04 | 0.11 | 0.12 | 0.21 | 0.32 | 0.42 | 0.55 | 0.63 | 0.72 | 0.83 | 0.86 | 0.88 | 1.00 |

equation (3) is $\lambda = 0.35/m^2$ or 1 clam per 2.87 $m^2$.

The lower and upper 0.025 percentage points from a chi-square distribution with $2nr = 84$ df are $C_1 = 60.5$, $C_2 = 111.2$. Hence, from equation (7) the 95% confidence limits are [0.26, 0.47] clams per square meter. The confidence interval is skewed because it is based on the chi-square distribution.

The statistic $Z_j$, where

$$Z_j = \sum_{i=1}^{j} x_i \Big/ \sum_{i=1}^{14} x_i,$$

is an ordered statistic from the uniform distribution on the interval (0, 1) (Seshadri, Csorgo, and Stevens 1969) and can be used to test the hypothesis of randomness. The $Z_j$ and the empirical distribution $S_n$ $(=j/n)$ are listed in Table 1. The supremum of 0.15 is well below the critical value of 0.349 (from Owen 1962) for the 1 sample Kolmogorov-Smirnov test and the hypothesis of randomness cannot be rejected at the 0.05 level.

In this example it probably would not have been as practical to use the closest-individual method. Because waves constantly washed over the sampling area it would have been difficult to determine exactly what areas had been sampled. Resampling, and consequently wasted effort, would have been difficult to avoid.

There was some difficulty in counting because the clams had to be dug to be identified. The distance was relatively easy to cover since there were no obstacles on the exposed beach. Therefore, the variable area transect offered advantages over using quadrats. However, had the surf been heavier and more troublesome to covering the distance, quadrats would probably have been more advantageous.

This example helps to illustrate several points about the theoretical framework of the model and its applications. The precision of the method is not enhanced by inclusion of the distance measures to the 1st and 2nd organisms. Because the organisms are modeled as being randomly distributed the information content of the measurements to the 3rd organism contains all the information expressed by those distances to the 2nd and 1st. Note that while the number of measurements is $n$, the variance, equation (4) is inverse to the total number encountered, the product $nr$.

If the density is zero the variable area transect method will be of no value because the density estimate, equation (3), is dependent on at least 1 measure of distance. For the density to be zero, the distance, the denominator of equation (3), must be infinity. In this case quadrats would be superior to the variable area transect because for quadrats the variable being measured, the number of organisms, is in the numerator of the assembly equation and can be zero. Investigating the area beforehand for organisms can eliminate this problem.

The clams were assumed to be of uniform density over the entire area sampled. Because a significant deviation from randomness could be used as evidence to the contrary and because no such deviation was found, the assump-

tion is probably true. This is most likely because the exposed area of the beach is relatively flat, making the environment homogeneous to life-supporting and life-damaging influences. However, it is not difficult to imagine that beyond the surf line, where inundation and protection from clam harvesting is greater, the density is higher than that of the exposed beach. The density could also change along the beach. Substrate, angle of wave action, and utilization by clam diggers cannot be thought of as being invariant over the entire stretch of the beach. For the situation where distinct areas of different densities can be defined, an estimator is described in the following section. When distinct areas are not easily defined difficulties can arise. It would probably be best for the investigator to define the areas as well as possible and then proceed. If the investigator does not have some prior knowledge about the interaction of organisms and terrain then some presampling may be appropriate.

## DENSITY ESTIMATOR FOR RANDOM AND CLUSTERED DISTRIBUTIONS

An alternative to randomness often encountered in biological problems is that of aggregation, or heterogeneity. A special case of aggregation occurs when a population is divided into distinct regions of different densities, where the organisms are randomly distributed within each region. For the closest individual method Morisita (1957) proposes an unbiased estimator of density for this situation. For the variable area transect the analog to Morisita's estimator is (see Appendix, Part B)

$$\hat{\lambda} = \frac{(r-1)\sum_{i=1}^{n} x_i^{-1}}{nw} \qquad (20)$$

or

$$\hat{\lambda} = \frac{\sum_{i=1}^{n} \hat{\lambda}_i}{n},$$

where $\hat{\lambda}_i$ is the density defined for each sample point. It can be shown that $\hat{\lambda}_i$ is unbiased for the random distribution described in equations (1) and (2) with the following variance (see Appendix, Part B)

$$\mathrm{Var}(\hat{\lambda}) = \frac{\lambda^2}{n(r-2)}, \qquad (21)$$

which has an unbiased replicated sample estimate

$$\mathrm{Var}(\hat{\lambda}) = \frac{\sum(\hat{\lambda}_i - \hat{\lambda})^2}{n(n-1)}. \qquad (22)$$

Thus when the population is uniformly random over the area sampled there are 2 density estimators: $\hat{\lambda}$, equation (3) and $\hat{\lambda}$, equation (20). Since the variance is always less for $\hat{\lambda}$, $(nr-2)$ always being greater than $n(r-2)$, it would be preferable to use $\hat{\lambda}$ over $\hat{\lambda}$ in this situation.

For the situation where the densities are variant for different regions the sampling must be done from random points within each region. Also, a separate test of randomness should be carried out for each region.

The negative binomial distribution is often used to model aggregated populations. Parker (1976) shows that $\hat{\lambda}$ is unbiased for the negative-binomial distribution with the following variance:

$$\mathrm{Var}(\hat{\lambda}) = \frac{\lambda^2}{n(r-2)}\left[1 + \frac{n(r-1)}{k}\right]. \qquad (23)$$

Two parameters must be estimated: $\lambda$, the density and $k$, the heterogeneity parameter. In discussing Morisita's estimator, Eberhardt (1967) noted that $k$ will tend to vary with $r$. This is also the case for equation (23); for a given variance $k$ must change with $r$. Consequently, the negative binomial model may not be en-

tirely realistic for distance-measure methods. If an investigator suspects aggregation, in the sense that it cannot be divided into distinct regions of randomness, he is advised to use the quadrat methods as outlined by Bliss and Fisher (1953).

## DISCUSSION

For density estimation an investigator is usually interested in a method which is accurate, precise, and requires minimum field effort. When organisms are distributed at random (Poisson) the variable area transect offers a useful alternative to other distance-measures methods, and quadrats. For small sample sizes the variable area transect should be especially useful because the confidence intervals are exact.

Both the variable area transect and closest-individual methods provide unbiased estimates of density with equal variances. The relative effort expended in achieving a desired sample size should be less for most sampling situations with the variable area transect because determining the order of organisms could likely require less searching effort than with the closest-individual method.

Both quadrats and variable area transects can be used for unbiased density estimation under the condition of randomness. For an equal sampling area the variance for quadrats is slightly less. When effort is a linear function of distance covered and number counted, both methods require an equal sampling effort. When counting is difficult and the population is dense, the effort is more consistent for the variable area transect. When the density is low and the distance is difficult to cover, the effort is more consistent for quadrats.

A density estimator can be derived that is useful for both uniformly random populations and aggregated populations, where the variant density for different regions is randomly distributed. The estimator is unbiased for the negative binomial distribution as well, the implication being that the estimator is robust for a wide set of conditions as yet undefined.

In conclusion, the variable area transect provides a useful sampling alternative. Because of ease of operation and because the resulting estimate is free of quadrat size, the variable area transect should be useful in a wide range of biological sampling problems.

## LITERATURE CITED

BLISS, C. I., AND R. A. FISHER. 1953. Fitting the negative binomial distribution to biological data; note on the efficient fitting of the negative-binomial. Biometrics 9:176–200.

EBERHARDT, L. L. 1967. Some developments in 'distance sampling.' Biometrics 23:207–216.

GREIG-SMITH, P. 1964. Quantitative plant ecology. Butterworth and Co., Ltd., London, England. 485pp.

HOLGATE, P. 1964. The efficiency of nearest neighbor estimators. Biometrics 20:647–649.

MORISITA, M. 1957. A new method for the estimation of density by the spacing method applicable to non-randomly distributed populations. Physiol. and Ecol. 7:134–144 (Japanese). U.S. Dep. Agric. Transl. 11116.

OWEN, D. B. 1962. Handbook of statistical tables. Addison-Wesley, Inc., Reading, Mass. 580pp.

PARKER, K. R. 1976. Variable area transect estimation of population density. M.S. Thesis. Univ. Washington, Seattle, 38pp.

PIELOU, E. C. 1957. The effect of quadrat size on the estimation of the parameters of Neyman's and Thomas' distributions. J. Ecol. 45:31–47.

POLLARD, J. H. 1971. On distance estimators of density in randomly distributed forests. Biometrics 27:991–1002.

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

SESHADRI, V., M. CSORGO, AND M. A. STEVENS. 1969. Tests for the exponential distribution using Kolmogorov-type statistics. J. R. Stat. Soc. 31:499–509.

*Received 17 January 1978.*
*Accepted 20 October 1978.*

## APPENDIX A. MAXIMUM LIKELIHOOD ESTIMATION OF DENSITY

The maximum likelihood estimator is that which maximizes the likelihood that a set of distance measures $x_1, x_2, \ldots, x_n$ assume a particular value given the density function of the population density. Maximum likelihood estimators are appealing to use because they are asymptotically efficient, normal, and unbiased.

From the density function, equation (2), the likelihood function is

$$L(\lambda) = \left[(\lambda w)^{nr}(x_1{}^{r-1}\cdot x_2{}^{r-1}\cdot \cdots x_n{}^{r-1}) \cdot \exp\left(-\lambda w \sum_{i=1}^{n} x_i\right)\right] \Big/ [(r-1)!]^n.$$

Taking the partial derivative with respect to $\lambda$, equating to 0 and solving for $\lambda$ the maximum likelihood estimator is

$$\tilde{\lambda} = nr \Big/ \left(w \sum_{i=1}^{n} x_i\right).$$

As a check for bias, the expected value of $\tilde{\lambda}$ is

$$E(\tilde{\lambda}) = \int_0^\infty \cdots \int_0^\infty \frac{nr}{w\sum_{i=1}^{n} x_i} \prod_{i=1}^{n} \frac{(\lambda w)^r x_i{}^{r-1}\exp(-\lambda w x_i)}{(r-1)!}\, dx_i. \quad \text{(A1)}$$

By making the following transformation integration of equation (A1) is simplified:

$$U_j = \sum_{i=1}^{j} x_i.$$

$U_n$ is integrated from 0 to $\infty$, and each $U_j$ is integrated from $U_j$ to $U_{j+1}$. The first $n-1$ integrals are beta, the last is a gamma:

$$E(\tilde{\lambda}) = \frac{nr}{nr-1}\lambda.$$

Although asymptotically unbiased, $\tilde{\lambda}$ is biased for small sample sizes. From the invariance principle of maximum likelihood estimators, the modified (unbiased) maximum likelihood estimator is

$$\hat{\lambda} = \frac{(nr-1)}{nr} \cdot \frac{nr}{w\sum_{i=1}^{n} x_i} = \frac{nr-1}{w\sum_{i=1}^{n} x_i}.$$

To estimate the variance of $\hat{\lambda}$ the second moment of $\hat{\lambda}$ is computed with the same transformations as above, where

$$E(\hat{\lambda}^2) = \frac{(nr-1)}{(nr-2)}\lambda^2,$$

and then the variance of $\hat{\lambda}$ is, from definition,

$$\mathrm{Var}(\hat{\lambda}) = E(\hat{\lambda}^2) - [E(\hat{\lambda})]^2 = \frac{\lambda^2}{(nr-2)}.$$

## APPENDIX B. EXPECTATION AND VARIANCE OF GENERAL ESTIMATOR

Eberhardt (1967) showed that the expected value of the reciprocal of the measure of distance to the $r$th organism is equivalent to Morisita's (1957) estimator. The situation is the same with the variable area transect. From equation (2)

$$E(x^{-1}) = \frac{(\lambda w)^r}{(r-1)!}\int_0^\infty x^{r-2}\exp(-\lambda w x)\, dx = \frac{\lambda w}{r-1},$$

and then an estimate of the density is

$$\hat{\lambda} = \frac{(r-1)E(x^{-1})}{w}.$$

As a test for bias

$$E(\hat{\lambda}) = \int_0^\infty \cdots \int_0^\infty \frac{(r-1)\sum_{i=1}^{n} x_i{}^{-1}}{nw}$$

$$\cdot \prod_{i=1}^{n} \frac{(\lambda w) x_i^{r-1} \exp(-\lambda w x_i)}{(r-1)!} dx_i,$$

which can be solved directly without transformations:

$$E(\hat{\lambda}) = \lambda.$$

The 2nd moment is derived similarly where

$$E(\hat{\lambda}^2) = \frac{\lambda^2 + \lambda^2 n(r-2)}{n(r-2)}$$

and the variance is

$$\begin{aligned} \mathrm{Var}(\hat{\lambda}) &= E(\lambda^2) - [E(\lambda)]^2 \\ &= \frac{\lambda^2}{n(r-2)}. \end{aligned}$$