

A Likelihood Ratio Test for Equality of Deviations from Randomness

D. L. JOHNSON AND A. J. PETKAU*

Agriculture and Agri-Food Canada Research Centre, P.O. Box 3000, Lethbridge, AB, Canada T1J 4B1

*Department of Statistics, The University of British Columbia, Vancouver, BC, Canada V6T 1Z2

Abstract. The use of the sample variance-to-mean ratio as a measure of deviation from randomness in spatial pattern is reviewed. The likelihood ratio method of constructing a statistical test for the equality of several population variance-to-mean ratios is described, and details are provided for the special case where counts are modelled as arising from a negative binomial distribution. This test is illustrated by application to example data sets in ecology. Likelihood ratio tests represent a general methodology whereby relationships among several indices of aggregation can be systematically investigated, provided one is able to specify a suitable parametric form for the underlying distributions.

Key words: variance-to-mean ratios, deviations from randomness, spatial pattern and counts, Poisson, negative binomial.

Introduction

The examination of spatial pattern is fundamental to many ecological investigations, including studies of mechanisms as well as determination of sampling schemes. Various methods of analyzing observational data and detecting deviations from a random distribution have been developed.

Spatial pattern is usually represented by the distribution of counts per unit area, volume, time, individual host or site. The simplest model of dispersion is a hypothesis of randomness, described by the Poisson probability distribution:

$$P[X=x] = \frac{e^{-\lambda} \lambda^x}{x!}, \quad \text{for } x=0, 1, \dots \quad (\lambda > 0).$$

Both the mean and the variance of this probability distribution are equal to the parameter λ .

It was noted early in the century, however, that the Poisson rarely provides an adequate description of ecological data (e.g., Student 1919; Blackman 1935; Clapham 1936). The sample variance-to-mean ratio provides a measure of the degree of departure from Poisson expectation. Ratios significantly greater than 1 are taken to indicate clumping or "contagion". Such deviations can be easily tested since if the underlying distribution is Poisson, then for reasonably large sample size n , the sampling distribution of $(n-1)s^2/\bar{x}$ can be reasonably approximated as χ^2 with $n-1$ degrees of freedom (Fisher et al. 1922; see

also Hoel 1943). Fisher (1950) provided an example in which the variance-to-mean ratio was more sensitive than the usual χ^2 tests in detecting deviations from the Poisson. Morisita (1962) suggests that because s^2/\bar{x} typically varies with density (i.e., \bar{x}), it should not be recommended as an index of dispersion representing degree of aggregation, but nevertheless it provides a useful test of significance of departure from randomness. Perry and Mead (1979) found the sample variance-to-mean ratio was quite powerful as a test of randomness and they recommended its use in a wide variety of situations.

The relative degree of departure from randomness exhibited by independent samples or treatments is often of interest. Such comparisons may provide insight into ecological processes or consequences, such as behavioral tendencies to aggregate or disperse. The variance-to-mean ratio has been chosen as the basis for such comparisons in a wide variety of studies. For example, Boiteau et al. (1979) compared s^2/\bar{x} ratios calculated from sweepnet samples of the bean leaf beetle among maturity groups and sampling dates. Erwin (1977) used s^2/\bar{x} ratios from samples of counts of foraging birds per 10-minute period to determine which species foraged more randomly in time. Mason (1970) computed s^2/\bar{x} for counts of five snail species from monthly collections, and drew ecological inferences from seasonal trends. Myers and Harris (1980) compared the s^2/\bar{x} of distributions of two gall flies, *Urophora affinis* (Frld.) and *U. quadrifasciata* (Meig.), on diffuse and spotted knapweed (*Centaurea*

diffusa Lam. and *C. maculosa* Lam.) over time and between species. In a classic study of spatial distribution of soil insects, Salt and Hollick (1946) compared the s^2/\bar{x} of counts per soil sample for three age classes of click beetles. In a study of winter aggregation of the northern red-backed vole (*Clethrionomys rutilus*), West (1977) compared the indices of clumping ($s^2/\bar{x}-1$) calculated from monthly catches of voles in traps.

Suggestions for assessing the equality of indices of aggregation have been made previously. David and Moore (1954) proposed a z-score, based on asymptotic normality of $\log(s^2/\bar{x})$, for testing the equality of variance-to-mean ratios in two populations. Their method relies on an approximation, due to Bateman (1950), to the sampling distribution of s^2/\bar{x} from a Neyman Type A contagious distribution. Reed (1983) investigated the use of the jackknife as a distribution-free method of forming confidence intervals for such indices and suggested specific transformations intended to improve small-sample performance. This method also lends itself to assessing the equality of indices of aggregation from two populations. Hutchison and Lyons (1989) suggested the use of pairwise z-scores, based on asymptotic normality of s^2/\bar{x} , for testing the equality of variance-to-mean ratios for binomial, Poisson and negative binomial populations. Their z-scores depend upon which parametric form is used as the model for the underlying distribution. Hutchison and Lyons (1989) noted that in their simulation studies for negative binomial populations, the jackknife underestimated the variance more often than their moment-based estimators. None of these authors addressed the issue of testing the equality of indices of aggregation from several populations.

In this paper, we consider the problem of inference for variance-to-mean ratios. We demonstrate construction of the likelihood ratio test of the hypothesis of equality of variance-to-mean ratios for several populations based on independent samples of counts. The methods presented can be applied to statistical distributions such as the Neyman Type A, Poisson-lognormal, negative binomial and Polya distributions which model data showing aggregation ($s^2 > \bar{x}$). The negative binomial is detailed because of its versatility and wide application.

The negative binomial is usually defined as:

$$P[X=x] = \binom{k+x-1}{x} \left(\frac{m}{m+k}\right)^x \left(\frac{k}{m+k}\right)^k, \\ \text{for } x=0, 1, \dots \quad (k, m > 0)$$

(e.g., Bishop et al. 1980). The parameter m is the mean, and the variance is $m(1+m/k)$. The geometric and Poisson distributions are special cases with $k=1$ and $k=\infty$, respectively. Defining $\theta=m/k$, we obtain the reparametrization of the negative binomial distribution first employed by Fisher (1941):

$$P[X=x] = \binom{k+x-1}{x} \left(\frac{\theta}{\theta+1}\right)^x \left(\frac{1}{\theta+1}\right)^k, \\ \text{for } x=0, 1, \dots \quad (k, \theta > 0).$$

For the case of p populations with parameters m_i and k_i , the hypothesis of equality of ratios can then be expressed as:

$$H_0: \theta_1 = \theta_2 = \dots = \theta_p.$$

The likelihood ratio test

Suppose interest focuses on the null hypothesis $H_0: \theta \in \Omega_0$, that is, the parameter vector θ lies in a specified subset Ω_0 of the parameter space Ω . The likelihood ratio is given by:

$$\lambda = L(\hat{\theta}^0)/L(\hat{\theta}),$$

where $L(\theta)$ is the likelihood function and $\hat{\theta}$ is the maximum likelihood estimate (MLE) of θ , while $\hat{\theta}^0$ is the MLE of θ when H_0 is assumed to be true (θ is restricted to lie in Ω_0). Note that $0 \leq \lambda \leq 1$; values of λ close to 1 correspond to outcomes in good agreement with H_0 and values of λ close to 0 indicate outcomes in poor agreement.

In special cases, the sampling distribution of the likelihood ratio λ can be explicitly determined. In general, however, appeal must be made to a result concerning its large sample distribution. Suppose the parameter space Ω is of dimension r (there are r unknown parameters), and the hypothesis H_0 restricts θ to a subspace of dimension s (there are s unknown parameters under H_0). Under rather general conditions (for a detailed discussion, see Rao 1965), if the sample size n is reasonably large, and H_0 is true:

$$-2 \log \lambda = 2[l(\hat{\theta}) - l(\hat{\theta}^0)] \approx \chi^2_{(r-s)},$$

where $l(\theta) = \log L(\theta)$ is the log-likelihood function.

For a sample of n counts from a negative binomial parametrized by θ and k , let f_j be the observed frequency of a count of j ($\sum_{j=0}^{\infty} f_j = n$). The log-likelihood function is:

$$l(\theta, k) = \sum_{j=0}^{\infty} f_j \log \left[\binom{k+j-1}{j} \left(\frac{\theta}{\theta+1}\right)^j \left(\frac{1}{\theta+1}\right)^k \right] \\ = \sum_{j=0}^{\infty} f_j \left[\sum_{i=0}^{j-1} \log(k+i) + j \log \theta - \right. \\ \left. (k+j) \log(\theta+1) - \log(j!) \right].$$

Levin and Reeds (1977) established that the MLE ($\hat{\theta}, \hat{k}$) is unique and corresponds to a finite k in the case of interest ($s^2 > \bar{x}$). The likelihood equations can be written as:

$$\frac{\partial l}{\partial \theta} = \sum_{j=0}^{\infty} f_j \left[\frac{j}{\theta} - \frac{k+j}{\theta+1} \right] = \frac{n}{\theta(\theta+1)} (\bar{x} - k\theta) = 0,$$

$$\frac{\partial l}{\partial k} = \sum_{j=0}^{\infty} f_j \left[\sum_{i=0}^{j-1} \frac{1}{(k+l)} - \log(\theta+1) \right]$$

$$= \sum_{i=0}^{\infty} f_i^+ / (k+l) - n \log(\theta+1) = 0,$$

where

$$f_i^+ = \sum_{j=i+1}^{\infty} f_j$$

is the number of counts exceeding l . The choice:

$$\theta = \bar{x}/k,$$

therefore maximizes the likelihood for variation in θ . Substitution of this choice into the second equation leads to the equation:

$$\sum_{i=0}^{\infty} f_i^+ / (k+l) - n \log(1 + \bar{x}/k) = 0,$$

which must be solved in an iterative fashion. Computations leading to the MLE's \hat{k} and $\hat{\theta} = \bar{x}/\hat{k}$ were first described by Fisher (1941) and Haldane (1941).

In the case of samples of n_i counts from p independent negative binomials with parameters θ_i and k_i ($i=1, 2, \dots, p$), the log-likelihood function is:

$$l = l(\theta_1, \theta_2, \dots, \theta_p, k_1, k_2, \dots, k_p) = l(\theta, k) = \sum_{i=1}^p l_i(\theta_i, k_i),$$

where the log-likelihood for the i th sample is:

$$l_i(\theta_i, k_i) = \sum_{j=0}^{\infty} f_{ij} \left[\sum_{i=0}^{j-1} \log(k_i+l) + j \log \theta_i - (k_i+j) \log(\theta_i+1) - \log(j!) \right];$$

here f_{ij} is the observed frequency of a count of j in the i th sample ($\sum_{j=0}^{\infty} f_{ij} = n$). Computations of maximum likelihood estimates can be carried out for each sample separately.

The null hypothesis of interest,

$$H_0: \theta_1 = \theta_2 = \dots = \theta_p = \theta \text{ say,}$$

leaves the common value θ unspecified. The log-likelihood function is:

$$l = l(\theta, k_1, k_2, \dots, k_p) = l(\theta, k) = \sum_{i=1}^p l_i(\theta, k_i),$$

leading to the likelihood equations

$$\frac{\partial l}{\partial \theta} = \frac{1}{\theta(\theta+1)} \sum_{i=1}^p n_i(\bar{x}_i - k_i\theta) = 0,$$

$$\frac{\partial l}{\partial k_i} = \sum_{i=0}^{\infty} f_{i+}^+ / (k_i+l) - n_i \log(\theta+1) = 0$$

for $i=1, 2, \dots, p,$

where

$$f_{i+}^+ = \sum_{j=l+1}^{\infty} f_{ij}$$

is the number of counts exceeding l in the i th sample. The choice:

$$\theta = \frac{\sum_{j=1}^p n_j \bar{x}_j / \sum_{j=1}^p n_j k_j,$$

therefore maximizes the likelihood for variation in θ . Substitution leads to the remaining equations:

$$\sum_{i=0}^{\infty} f_{i+}^+ / (k_i+l) - n_i \log \left(1 + \frac{\sum_{j=1}^p n_j \bar{x}_j / \sum_{j=1}^p n_j k_j}{k_i} \right) = 0$$

for $i=1, 2, \dots, p.$

Each of these equations involves all of k_1, k_2, \dots, k_p , so the system of p equations must be solved simultaneously in an iterative fashion. This leads to the MLE's under H_0 :

$$\hat{k}_1^0, \hat{k}_2^0, \dots, \hat{k}_p^0 \text{ and } \hat{\theta}^0 = \frac{\sum_{j=1}^p n_j \bar{x}_j / \sum_{j=1}^p n_j \hat{k}_j^0,$$

and

$$-2 \log \lambda = 2[l(\hat{\theta}, \hat{k}) - l(\hat{\theta}^0, \hat{k}^0)] \approx \chi_{(p-1)}^2$$

yields an approximate P -value for the likelihood ratio test of

$$H_0: \theta_1 = \theta_2 = \dots = \theta_p.$$

In the event the data are in good agreement with H_0 , one might wish to proceed to inferences concerning the parameters under the simpler statistical model specified by H_0 (e.g., construction of a confidence interval for the common θ). Estimates of variances and covariances of the MLE's to facilitate such inferences for both the general model and the simpler model specified by H_0 are provided in the Appendix.

Remark 1:

Note that s^2/\bar{x} , the method of moments estimate of the population variance-to-mean ratio is not equal to $\hat{\theta}-1$, the MLE under the negative binomial model. Since the method of moments is not fully efficient, the MLE is a better estimate of the population variance-to-mean ratio.

Remark 2:

The corresponding computations for the null hypothesis:

$$H_0: k_1 = k_2 = \dots = k_p = k,$$

are considerably simpler because even under H_0 , maximum likelihood estimation involves the iterative solution of only a single equation rather than a system of equations. For details, see Bliss and Fisher (1953), Bliss (1958), and Bliss and Owen (1958). This problem has attracted considerable attention and application in the ecological literature because the parameter k has also been

Table 1. Statistics for cankers on apple trees (Johnson et al. 1982).

Cultivar	\bar{x}	s^2/\bar{x}	n	Unrestricted MLE		MLE under H_0	
				\hat{k}	$\hat{\theta}$	\hat{k}^0	$\hat{\theta}^0$
Golden Delicious	6.60	6.73	48	0.85	7.74	1.15	5.18
Spartan	6.91	8.83	45	1.22	5.67	1.31	5.18
McIntosh	10.73	5.74	33	2.45	4.38	2.13	5.18
Delicious	12.67	4.55	36	4.18	3.03	2.61	5.18

used to indicate the degree of aggregation.

Examples

Apple cankers

Nectria galligena Bres. is a common fungal parasite of apple trees. On infected trees, localized areas of bark and underlying cambium die and tissue sloughs from the face of the canker, leaving a blackened ellipsoid series of concentric callus ridges. Johnson et al. (1982) counted the cankers per tree in an orchard to assess whether cultivar might affect canker distribution, as measured by the respective variance-to-mean ratios. The Poisson distribution did not provide an acceptable fit for any of the four apple cultivars examined (Table 1), but the negative binomial distribution provided a reasonable fit for all four cultivars ($P > 0.05$ for GOF χ^2). The likelihood ratio test for common θ led to $-2 \log \lambda = 5.07$ and indicates that the differences in aggregation of the cultivars may be due to chance ($P \approx 0.17$ from $\chi^2_{(3)}$). Hence, there appears to be no

compelling reason for entertaining research hypotheses which depend upon or seek to explain differences among the variance-to-mean ratios of counts of cankers on the four apple varieties.

Mountain goats

Foster (1982) counted the number of mountain goats in 129 quadrats (each 400 m²) in the Nass Ranges, B.C., Canada, each month for 16 months. He calculated the variance-to-mean ratio for each month as an index of deviation from randomness. If the samples for different months are assumed to be independent, our test can be applied to examine differences in distribution over time as a possible aid to understanding the timing of mountain goat aggregation.

The negative binomial provides a reasonable fit ($P > 0.05$ for GOF χ^2) to the data of all months except May and November, 1977 ($P < 0.01$). The data lead to $-2 \log \lambda = 32.31$, suggesting differences among months in deviations from a random pattern are not due to chance ($P \approx 0.006$ from $\chi^2_{(15)}$). Removing the two samples for which the

Table 2. Statistics for Foster's (1982) mountain goats. $n=129$ for each sample (month).

Month		\bar{x}	s^2/\bar{x}	Unrestricted MLE		MLE under H_0	
				\hat{k}	$\hat{\theta}$	\hat{k}^0	$\hat{\theta}^0$
May	1977	1.38	5.18	0.183	7.56	0.180	7.76
June	1977	2.05	9.15	0.178	11.55	0.215	7.76
July	1977	1.03	8.17	0.080	12.86	0.100	7.76
August	1977	1.97	12.18	0.125	15.80	0.171	7.76
September	1977	1.76	8.73	0.250	7.04	0.238	7.76
October	1977	1.06	4.79	0.171	6.20	0.153	7.76
November	1977	1.59	5.30	0.187	8.51	0.196	7.76
December	1977	1.03	16.53	0.066	15.59	0.088	7.76
January	1978	1.17	5.43	0.209	5.61	0.177	7.76
February	1978	1.78	5.52	0.299	5.93	0.259	7.76
March	1978	1.74	4.28	0.370	4.71	0.279	7.76
April	1978	1.46	3.37	0.520	2.80	0.284	7.76
May	1978	1.73	6.68	0.248	6.96	0.235	7.76
June	1978	1.48	6.74	0.171	8.63	0.181	7.76
July	1978	0.75	6.48	0.075	10.04	0.084	7.76
August	1978	1.02	10.63	0.119	8.53	0.124	7.76

Table 3. Statistics for the prairie vegetation example (Steiger 1930). $n=40$ quadrats for each sample.

Species	Site	\bar{x}	s^2/\bar{x}	Unrestricted MLE		MLE under H_0	
				\hat{k}	$\hat{\theta}$	\hat{k}^0	$\hat{\theta}^0$
<i>A. multiflorus</i>	low prairie	2.53	6.66	0.328	7.69	0.372	6.18
	high prairie	1.35	4.35	0.313	4.31	0.255	6.18
$-2 \log \lambda = 0.749$ ($P \approx 0.39$ from $\chi^2_{(1)}$)							
<i>S. glaberrima</i>	low prairie	6.45	10.50	0.504	12.79	0.498	13.08
	high prairie	3.23	19.52	0.238	13.55	0.242	13.08
$-2 \log \lambda = 0.010$ ($P \approx 0.92$ from $\chi^2_{(1)}$)							

negative binomial does not provide a good fit results in $-2 \log \lambda = 32.20$ ($P \approx 0.002$ from $\chi^2_{(13)}$), so, in this case, lack of fit in some of the samples does not affect the conclusions.

Prairie vegetation

Steiger (1930), in a detailed study of prairie grasses and forbs, published tables of the distributions of counts of individual plants found on "high prairie" and "low prairie" sites in Nebraska. Clapham (1936) calculated variance-to-mean ratios for Steiger's samples but found the Poisson was not a reasonable model for most of the samples. He attributed the degree of over-dispersion (aggregation) of plants to the mode of reproduction with vegetative reproduction resulting in highly aggregated spatial patterns and reproduction by seed resulting in less aggregation especially as the dispersal distance increased.

The negative binomial provides a good fit to the distribution of counts for *Aster multiflorus* Ait and *Solidago glaberrima* Martens in 40 quadrats of high and low prairie. We applied the likelihood ratio test to determine whether either species has different variance-to-mean ratios in high and low prairie (Table 3). There is no strong evidence for rejection of the hypothesis of equality in either case.

Concluding remarks

We do not suggest that the variance-to-mean ratio be adopted as a general index of spatial pattern. We present the details of the likelihood ratio test for the equality of several such population ratios, and expressions for variances and covariances of maximum likelihood estimates, so that researchers who choose to use this index as a measure of departure from Poisson expectation can properly make inference. We do not suggest the negative binomial as a general model of contagion; our development demonstrates a method to produce the details necessary for other models which describe contagion.

Maximum likelihood estimation and likelihood ratio tests provide a general statistical methodology whereby relationships among several indices of aggregation, or other population parameters, can be investigated systematically. If, for example, one wished to investigate the suitability of a power law relationship between the variances and means of several related populations ($\sigma^2 = \alpha \mu^\beta$, Taylor 1961), maximum likelihood estimation of the parameters α and β would lead directly to a likelihood ratio test to assess the adequacy of the postulated power law as a representation of the relationship. Note that equality of variance-to-mean ratios corresponds to the special case $\beta = 1$. Of course, these methods only apply if one is able to specify a suitable parametric form for the underlying distribution of the data.

A copy of the Fortran program used to perform the Newton-Raphson iteration and the calculations for the examples presented is available from D. L. Johnson.

Acknowledgements: This research was supported in part by the Natural Sciences and Engineering Research Council (A. J. Petkau and D. L. Johnson), by a Killam fellowship at the University of British Columbia (D. L. Johnson), and by Agriculture and Agri-Food Canada, Research Centre, Lethbridge (D. L. Johnson). This paper is Lethbridge Research Centre Contribution No. 3879443.

References

- Bateman, G. I. (1950) The power of the χ^2 index of dispersion test when Neyman's contagious distribution is the alternative hypothesis. *Biometrika* 37: 59-63.
- Bishop, Y. M. M., S. E. Fienberg and P. W. Holland (1980) *Discrete multivariate analysis: theory and practice*. MIT press, Cambridge, Mass.
- Blackman, G. E. (1935) A study by statistical methods of the distribution of species in grassland associations. *Ann. Bot.* 49: 749-777.
- Bliss, C. I. (1958) The analysis of insect counts as negative binomial distributions. *Proceedings of the Tenth International Congress of Entomology* 2: 1015-1032.
- Bliss, C. I. and A. R. G. Owen (1958) Negative binomial distributions with a common k . *Biometrika* 45: 37-58.
- Bliss, C. I. and R. A. Fisher (1953) Fitting the negative binomial

- distribution to biological data. *Biometrics* **9**: 176–200.
- Boiteau, G., J. R. Bradley Jr., J. W. Van Duyn and R. E. Stinner (1979) Bean leaf beetle micro-spatial patterns and sequential sampling of field populations. *Environ. Entomol.* **8**: 1139–1144.
- Clapham, A. R. (1936) Over-dispersion in grassland communities and the use of statistical methods in plant ecology. *J. Ecol.* **24**: 232–251.
- David, F. N. and P. G. Moore (1954) Notes on contagious distributions in plant populations. *Ann. Bot.* **18**: 47–53.
- Erwin, R. M. (1977) Foraging and breeding adaptations to different food regimes in three seabirds: the Common Tern, *Sterna hirundo*, Royal Tern, *Sterna maxima*, and Black Skimmer, *Rynchops nigra*. *Ecology* **58**: 389–397.
- Fisher, R. A. (1941) The negative binomial distribution. *Ann. Eugenics* **11**: 182–187.
- Fisher, R. A. (1950) The significance of deviations from expectation in a Poisson series. *Biometrics* **6**: 17–24.
- Fisher, R. A., H. G. Thornton and W. A. Mackenzie (1922) The accuracy of the plating method of estimating bacterial populations. *Ann. Appl. Biol.* **9**: 325–359.
- Foster, B. F. (1982) *Observability and habitat characteristics of the mountain goat (Oreamnos americanus Blainville, 1816) in west-central British Columbia*. M.Sc. Thesis. University of British Columbia, Vancouver, B.C., Canada.
- Haldane, J. B. S. (1941) The fitting of binomial distributions. *Ann. Eugenics* **11**: 179–181.
- Hoel, P. G. (1943) On indices of dispersion. *Ann. Math. Stat.* **14**: 155–162.
- Hutcheson, K. and N. I. Lyons (1989) A significance test for Morisita's index of dispersion and the moments when the population is negative binomial and Poisson. pp. 335–346 *In* L. McDonald, B. Manly, J. Lockwood and J. Logan (eds.) *Estimation and analysis of insect populations*. Lecture notes in Statistics, Vol. 55. Springer-Verlag, New York.
- Johnson, D. L., J. Lovett Doust and G. W. Eaton (1982) The effect of European canker and its spatial pattern on four apple cultivars in British Columbia. *J. Appl. Ecol.* **19**: 603–609.
- Levin, B. and J. Reeds (1977) Compound multinomial likelihood functions are unimodal: proof of a conjecture by I. J. Good. *Ann. Stat.* **5**: 79–87.
- Mason, C. F. (1970) Snail populations, beech litter production, and the role of snails in litter decomposition. *Oecologia* **5**: 215–239.
- Morisita, M. (1962) I_p -index, a measure of dispersion of individuals. *Res. Popul. Ecol.* **4**: 1–7.
- Myers, J. H. and P. Harris (1980) Distribution of *Urophora* galls in flower heads of diffuse and spotted knapweed in British Columbia. *J. Appl. Ecol.* **17**: 359–367.
- Perry, J. N. and R. Mead (1979) On the power of the index of dispersion test to detect spatial pattern. *Biometrics* **35**: 613–622.
- Rao, C. R. (1965) *Linear statistical inference and its applications*. John Wiley and Sons, New York.
- Reed, W. J. (1983) Confidence estimation of ecological aggregation indices based on counts—a robust procedure. *Biometrics* **39**: 987–998.
- Salt, G. and F. S. J. Hollick (1946) Studies of wireworm populations. II. Spatial distribution. *J. Exp. Biol.* **23**: 1–46.
- Steiger, T. L. (1930) Structure of prairie vegetation. *Ecology* **11**: 170–217.
- Student (1919) An explanation of deviations from Poisson's law in practice. *Biometrika* **12**: 211–215.
- Taylor, L. R. (1961) Aggregation, variance and the mean. *Nature* **189**: 732–735.
- West, S. D. (1977) Midwinter aggregation in the northern red-backed vole, *Clethrionomys rutilus*. *Can. J. Zool.* **55**: 1404–1409.

Received 23 January 1995; Accepted 11 September 1995

Appendix. Variances and covariances of maximum likelihood estimates

In making inferences about population parameters (e.g., construction of confidence intervals), estimated variances and covariances of maximum likelihood estimates are required. For a log-likelihood function $l(\theta)$ depending on a vector of r parameters θ , typically the MLE can be found by solving the likelihood equations. If $\hat{\theta}$ must be determined iteratively, the Newton-Raphson method is most commonly employed:

$$\theta^{(k+1)} = \theta^{(k)} - [l''(\theta^{(k)})]^{-1} l'(\theta^{(k)});$$

here $\theta^{(k)}$ denotes the approximation to the solution $\hat{\theta}$ at the k th iteration, $l'(\theta)$ denotes the $r \times 1$ vector of first partial derivatives and $[l''(\theta)]$ denotes the $r \times r$ matrix of second partial derivatives. Under rather general conditions (for further discussion, see Rao 1965), the variance-covariance matrix of $\hat{\theta}$ can be approximated by $[-l''(\hat{\theta})]^{-1}$; note that the latter matrix is computed in the course of the Newton-Raphson iteration.

For a sample of n counts from a single negative binomial population.

$$[-l''(\hat{\theta}, \hat{k})] = \begin{bmatrix} \frac{n\hat{k}}{\hat{\theta}(\hat{\theta}+1)} & \frac{n}{(\hat{\theta}+1)} \\ \frac{n}{(\hat{\theta}+1)} & S \end{bmatrix}$$

where

$$S = S(\hat{k}) = \sum_{l=0}^{\infty} f_l^+ / (\hat{k} + l)^2.$$

Inversion of this matrix leads to the estimates:

$$\begin{aligned} \text{v\hat{a}r}(\hat{\theta}) &= (\hat{\theta} + 1)/D, \\ \text{cov\hat{a}r}(\hat{\theta}, \hat{k}) &= -(n/S)/D, \\ \text{v\hat{a}r}(\hat{k}) &= (n/S)(\hat{k}/\hat{\theta})/D, \end{aligned}$$

where

$$D = D(\hat{\theta}, \hat{k}) = n[\hat{k}/\hat{\theta} - (n/S)/(\hat{\theta} + 1)];$$

the estimate $\text{v\hat{a}r}(\hat{k})$ can also be obtained directly in the course of the Newton-Raphson iteration for \hat{k} . In the case of samples of n_i counts from p independent negative binomials with parameters θ_i and k_i ($i=1, 2, \dots, p$) and with no restrictions (the general model), the above expressions apply for each population separately.

If the data are in good agreement with $H_0: \theta_1 = \theta_2 = \dots = \theta_p = \theta$ say, one might wish to proceed to inferences concerning the parameters under the simpler model specified by H_0 . In this case,

$$[-l''(\hat{\theta}^{\circ}, \hat{k}^{\circ})] = \begin{bmatrix} a & b \\ b & S \end{bmatrix},$$

where

$$a = \sum_{i=1}^p n_i k_i^{\circ} / [\hat{\theta}^{\circ}(\hat{\theta}^{\circ} + 1)],$$

the $p \times 1$ vector b has i th element b_i given by

$$b_i = n_i / (\hat{\theta}^{\circ} + 1),$$

and the $p \times p$ diagonal matrix S has i th diagonal element S_i given by

$$S_i = (S)_{ii} = \sum_{l=0}^{\infty} f_{il}^+ / (\hat{k}_i^{\circ} + l)^2.$$

Inversion of this matrix leads to the estimates:

$$\begin{aligned} \text{v\hat{a}r}(\hat{\theta}^{\circ}) &= (\hat{\theta}^{\circ} + 1)/D, \\ \text{cov\hat{a}r}(\hat{\theta}^{\circ}, \hat{k}_i^{\circ}) &= -(n_i/S_i)/D, \\ \text{v\hat{a}r}(\hat{k}_i^{\circ}) &= 1/S_i + (n_i/S_i)^2 / [(\hat{\theta}^{\circ} + 1)D], \\ \text{cov\hat{a}r}(\hat{k}_i^{\circ}, \hat{k}_j^{\circ}) &= (n_i/S_i)(n_j/S_j) / [(\hat{\theta}^{\circ} + 1)D], \end{aligned}$$

where

$$D = D(\hat{\theta}^{\circ}, \hat{k}^{\circ}) = \sum_{j=1}^p n_j [\hat{k}_j^{\circ} / \hat{\theta}^{\circ} - (n_j/S_j) / (\hat{\theta}^{\circ} + 1)].$$

The estimates $\text{v\hat{a}r}(\hat{k}_i^{\circ})$ and $\text{cov\hat{a}r}(\hat{k}_i^{\circ}, \hat{k}_j^{\circ})$ can also be obtained directly in the course of the Newton-Raphson iteration for \hat{k}° .