NOTE

Comparing vertical plankton profiles with replication

Andrew Beet[1,*], Andrew R. Solow[1], Stephen M. Bollens[2]

[1]**Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA**
[2]**Department of Biology and Romberg Tiburon Center for Environmental Studies, San Francisco State University, 3152 Paradise Drive, Tiburon, California 94920, USA**

ABSTRACT: Traditional statistical methods for testing for differences between vertical plankton distributions are invalid if the distributions are patchy. This Note describes a statistical test for differences in vertical plankton distributions in the presence of patchiness when replicate samples are available. An illustration involving a test for day/night differences in the vertical distribution of the copepod *Calanus pacificus* is presented.

KEY WORDS: *Calanus pacificus* · Negative binomial distribution · Patchiness

INTRODUCTION

Variation in the vertical distribution of zooplankton has long been recognized as an important factor controlling the structure and dynamics of marine and aquatic food webs (Russell 1927, Cushing 1951, Banse 1964, Longhurst 1976, Lampert 1989). In analyzing data from field and laboratory studies, it is commonly of interest to test for a difference in the vertical distributions of zooplankton under different conditions (e.g. day/night, presence/absence of predators). Venrick (1986) pointed out that patchiness in zooplankton distributions invalidates standard statistical methods, such as the Kolmogorov-Smirnov test, that have been used to address this issue. Specifically, statistical tests that do not account for patchiness will reject the null hypothesis of no difference in vertical distribution at a rate that can far exceed the nominal significance level of the test. Solow et al. (2000) proposed a modification of the Kolmogorov-Smirnov test that avoids this problem (see also Smith et al. 1998). This modified test was developed for the case in which there is no replication, i.e. for each condition only a single sample at each depth is available. The purpose of this Note is to describe and illustrate a test that can be used when replicate samples are available.

METHOD

Interest centers on testing for differences in the vertical distributions under a total of T different conditions. To do so, n replicate samples are collected at each of a fixed set of D depths under each of these conditions. The assumption that the number of replicate samples is the same at each depth under each condition can be relaxed. Let the random variable X_{ijk} be the number of individuals collected under condition i at depth j in replicate k. We will assume that X_{ijk} has a negative binomial distribution with unknown mean μ_{ij} and unknown dispersion coefficient c. The negative binomial distribution is commonly used to describe count data with a variance:mean ratio exceeding 1 (e.g. Hayek & Buzas 1997). Under this model, the variance of X_{ijk} is $\mu_{ij} + c\mu_{ij}^2$. As $c \to 0$, the negative binomial distribution approaches the Poisson distribution and the variance:mean ratio approaches 1.

A general multiplicative model for the mean is:

$$\mu_{ij} = \alpha_i \beta_j (\tau + \phi_{ij}) \qquad (1)$$

with side conditions $\alpha_T = \beta_D = 1$ and $\phi_{iD} = \phi_{Tj}$ for all i and j (Thall 1992). Briefly, α_i represents the main effect of condition i, β_j is the main effect of depth j, and ϕ_{ij} is the interaction of condition i and depth j. The side condi-

*Email: abeet@whoi.edu

tions enable the parameters of the model to be identified. Let ϕ be the vector of interaction terms. Under this model, the problem can be formalized as testing the null hypothesis H_0: $\phi = 0$. Specifically, under this hypothesis, the mean abundance at each depth under one condition is the same constant multiple of the mean abundance at the same depth under another condition, so that the shapes of the depth profiles of mean abundance are the same under all conditions.

Paul & Banerjee (1998) described a test of H_0. The test statistic is:

$$B = n\sum_{i=1}^{T}\sum_{j=1}^{D}\frac{(\overline{Y}_{ij}-\hat{\mu}_{ij})^2}{\hat{\mu}_{ij}(1+\hat{c}\hat{\mu}_{ij})} \qquad (2)$$

where $\overline{Y}_{ij}$ is the average of the replicate counts for condition i at depth j and $\hat{\mu}_{ij}$ and $\hat{c}$ are the maximum likelihood (ML) estimates under H_0 of μ_{ij} and c, respectively. The log likelihood function is given by:

$$L = \sum_{i=1}^{T}\sum_{j=1}^{D}Y_{ijk}\log\mu_{ij}-(Y_{ijk}+c^{-1})\log(1+c\mu_{ij})$$
$$+ \sum_{l=1}^{Y_{ijk}}\log[1+c(l-1)] \qquad (3)$$

To find the ML estimates under H_0, substitute Eq. (1) for μ_{ij} in Eq. (3) with $\phi_{ij} = 0$ and maximize the resulting expression over the remaining parameters α_i, β_j, τ, and c. Let $\hat{\alpha}_i$, $\hat{\beta}_j$, and $\hat{\tau}$ be the ML estimates of the first 3 of these parameters. The ML estimate of μ_{ij} is given by the product $\hat{\alpha}_i\hat{\beta}_j\hat{\tau}$. The maximization of Eq. (3) must be done numerically and a MATLAB program can be provided upon request. Under H_0, the test statistic B has an approximate chi-squared distribution with $(T-1)(D-1)$ degrees of freedom.

Paul & Banerjee (1998) presented some simulation results showing that this test performs well. That is, the test falsely rejects H_0 at a rate that is reasonably close to its nominal significance level and has good power against a range of departures from H_0. Our own simulation results confirm these results.

ILLUSTRATION

This section presents the results of applying the test outlined in the previous section to some data pertaining to day/night differences in the vertical distribution of the planktonic copepod *Calanus pacificus* in Dabob Bay, Washington, USA. The data, which were extracted from a larger study by Bollens & Frost (1989) and Bollens et al. (1993), are given in Table 1. These data comprise 2 sets (August 20–22, 1985, and August 5–6, 1986) collected as paired day and night samples taken at $D = 6$ depths with $n = 4$ replicates at each depth. Here, we will test for day/night differences in both sets of samples, so that each test involves a comparison of $T = 2$ conditions.

In Table 2, we report the value of the test statistic B and the corresponding significance level (or p-value) based on the chi-squared distribution with 5 degrees of freedom. The results for the 2 day/night comparisons are quite similar and, in both cases, H_0 is decisively rejected. As a further illustration, we also compared the 2 daytime samples and the 2 nighttime samples. These results are also reported in Table 2. In contrast to the day/night comparisons, in these cases, it is not possible to reject H_0. Under H_0, the ML estimate $\hat{c}$ of the dispersion parameter is 0.55 for comparing Day–August 1985 to Day–August 1986 and 0.31 for comparing Night–August 1985 to Night–August 1986. This level of over-dispersion is sufficient to invalidate a test based on the assumption of Poisson counts. Briefly, under the Poisson assumption, the standard test would amount to testing for independence in a 6-by-2 contingency table (e.g. Everitt 1977), where the rows in the table correspond to depth, the columns correspond to condition, and the entries in the table are the sums of the 4 replicate counts. If this test is applied to that Day–August 1985 versus Day–August 1986 com-

Table 1. Replicate (n = 4) counts of *Calanus pacificus* in 6 depth bins for 2 sets of paired day and night samples collected in Dabob Bay, Washington, USA (see Bollens & Frost 1989 and Bollens et al. 1993 for details)

Depth bin	Depth (m)	1	2	3	4	1	2	3	4
		Day–August 1985				**Day–August 1986**			
1	0–10	1	0	1	0	0	0	0	1
2	10–25	1	1	2	1	0	1	0	1
3	25–50	1	1	3	1	1	1	1	1
4	50–75	2	1	44	3	1	4	4	1
5	75–125	67	14	50	79	86	16	14	47
6	125–175	4	1	3	4	0	0	1	37
		Night–August 1985				**Night–August 1986**			
1	0–10	460	208	163	234	65	121	120	120
2	10–25	1	28	66	123	21	24	14	10
3	25–50	48	13	16	63	18	16	2	2
4	50–75	25	19	17	14	9	7	1	1
5	75–125	21	1	8	16	10	9	2	2
6	125–175	3	11	1	4	1	0	2	1

Table 2. Values of the test statistic B and the corresponding p-value for selected comparisons

Comparison	B	p
Day–August 1985 vs Night–August 1985	19.68	0.001
Day–August 1986 vs Night–August 1986	19.47	0.002
Day–August 1985 vs Day–August 1986	6.51	0.259
Night–August 1985 vs Night–August 1986	1.43	0.921

parison, the result is highly significant ($p < 0.001$), while the analogous p-value is around 0.05 for comparing Night–August 1985 to Night–August 1986. These results are in sharp contrast to those in Table 2 and underline the importance of accounting for over-dispersion.

DISCUSSION

The purpose of this Note has been to describe and illustrate a test for differences in vertical plankton distributions in the presence of patchiness when replicate samples are available. This test is a complement to the test of Solow et al. (2000), which is used in situations when replicate samples are not available. The test of Solow et al. (2000) could also be applied when there is replication by pooling replicates, and the question therefore arises as to which procedure is better in that case. Although we have not undertaken a direct comparison, for theoretical reasons, we prefer to use the present test when there is replication. Briefly, Solow et al.'s (2000) test involved a modification of the 2-sample Kolmogorov-Smirnov statistic to make its sampling distribution under the null hypothesis robust to patchiness. In contrast, the test described in this Note exploits the replicate samples to characterize the degree of patchiness (as measured in the over-dispersion parameter c) and to use this directly in the test. This information is lost when replicates are pooled, which would be necessary when using the Solow et al. test, and it is for this reason that we prefer the test described herein.

LITERATURE CITED

Banse K (1964) On the vertical distribution of zooplankton in the Sea. Prog Oceanogr 2:53–125

Bollens SM, Frost BW (1989) Zooplanktivorous fish and variable diel vertical migration in the marine planktonic copepod *Calanus pacificus.* Limnol Oceanogr 34: 1072–1083

Bollens SM, Frost BW, Osgood K, Watts SD (1993) Vertical distributions and susceptibilities to vertebrate predation of the marine copepods *Metridia lucens* and *Calanus pacificus.* Limnol Oceanogr 38:1841–1851

Cushing DH (1951) The vertical migration of planktonic Crustacea. Biol Rev 26:158–192

Everitt B (1977) The analysis of contingency tables. Chapman & Hall, London

Hayek LA, Buzas M (1997) Surveying natural populations. Columbia University Press, New York

Lampert WH (1989) The adaptive significance of diel vertical migration of zooplankton. Funct Ecol 3:21–27

Longhurst AL (1976) Vertical migration. In: Cushing DH, Walsh JJ (eds) The ecology of the seas. Blackwell, Oxford, p 116–137

Paul SR, Banerjee T (1998) Analysis of two-way layout of count data involving multiple counts in each cell. J Am Stat Assoc 93(444):1419–1429

Russell FS (1927) The vertical distribution of plankton in the sea. Biol Rev 2:213–262

Smith W, Beet A, Solow AR (1998) Testing for shifts in the vertical distribution of plankton using a robust Kolmogorov-Smirnov like statistic. J Agric Biol Environ Stat 3(4): 421–429

Solow AR, Bollens SM, Beet A (2000) Comparing two vertical plankton distributions. Limnol Oceanogr 45(2): 506–509

Thall PF (1992) Score tests in two-way layout of counts. Commun Stat Theory Methods 21:3017–3036

Venrick EL (1986) The Smirnov statistic: an incorrect test for vertical distribution patterns. Deep-Sea Res 33: 1275–1277