Ecology, 77(1), 1996, pp. 75–80
© 1996 by the Ecological Society of America

A STATISTICAL TEST FOR A DIFFERENCE BETWEEN THE SPATIAL DISTRIBUTIONS OF TWO POPULATIONS[1]

STEPHEN E. SYRJALA
Alaska Fisheries Science Center, National Marine Fisheries Service,
NOAA, 7600 Sand Point Way Northeast, Bin C15700, Seattle, Washington 98115-0070 USA

Abstract. A statistical hypothesis test for a difference between the spatial distributions of two populations is presented. The test is based upon a generalization of the two-sample Cramér-von Mises test for a difference between two univariate probability distributions. It is designed to be sensitive to differences in the way the populations are distributed across the study area but insensitive to a difference in abundance between the two populations. The procedure is nonparametric and uses the methodology of a randomization test to determine the level of significance of the test statistic.

Key words: Cramér-von Mises; differences between distributions; distributions of populations; geographic distributions; hypothesis test; nonparametric test; randomization test; spatial distributions; statistical test.

INTRODUCTION

In the study of spatial or geographic distributions of populations, questions arise about whether the distribution of one population differs from the distribution of another population. For example, within a species, are males and females distributed differently or are juveniles and adults distributed differently? Or do the populations of two species have the same distribution? Questions may also arise as to whether the spatial distribution of a population has changed over time. In the following sections, I develop a statistical test to test the null hypothesis that there is no difference in the spatial distributions of two populations against the alternative hypothesis that some unspecified difference exists between the two distributions. This test is specifically designed to be insensitive to differences in the total abundances in the study area but sensitive to differences in the distributions given the respective sizes of the two populations.

The proposed test is based on a bivariate generalization of the Cramér-von Mises nonparametric test for a difference between two univariate probability distribution functions (Conover 1980). The test is constructed using empirical distribution functions only; no underlying theoretical distributions are considered. For the Cramér-von Mises test, individuals are sampled at random from a population, the value of a random variable is determined for each individual sampled, and an empirical cumulative distribution function is constructed across the range of that variable. For a true bivariate generalization of the test to compare two spatial distributions, the random variable of interest would be the locations of the sampled individuals. In a wild population, however, individuals cannot be sampled at random, independent of their location. Thus, for the present purpose the generalization of the Cramér-von Mises test must be modified.

In most field studies, sampling locations are selected and then the individuals who happen to be at the selected locations at the time of sampling constitute the sample. If the sampling at each location provides a measure of population density, a spatial distribution function can still be constructed across the study area. The random variable in this case is the observed density at the sampling location, not the location itself. The magnitude of the increment of the cumulative distribution function at each sampling location is proportional to the population density at the location; the sum of the increments in the cumulative distribution function is 1. In examining the difference between the distributions of two populations, the test statistic is the square of the difference between the cumulative distribution functions, summed over all sample locations. In comparing two distributions, this approach requires that the density sampling be conducted for each population at the same set of sampling locations.

Upton and Fingleton (1985) have previously examined the associations between the distributions of species. Rather than compare two distributions directly, their approach was to identify whether there is an attraction, a repulsion, or independence between the two species under study. While their discussion is primarily based on using a complete enumeration of individual locations and the distances between individuals, two of their methods—those in which the data are reduced to the numbers of individuals per quadrat—can be adapted to the type of data discussed herein, where one has sample density observations at various locations. The first approach is to use a 2 × 2 contingency table. Each quadrat is categorized by the presence or absence of each of two species. The second approach involves calculating the Pearson correlation coefficient between

[1] Manuscript received 27 June 1994; revised 17 January 1995; accepted 28 March 1995; final version received 22 May 1995.

the two sets of species density measurements, one set for each species.

The Hypothesis Test

Within a study area, consider the distributions of two populations—or, alternatively, the distributions of two disjoint groups or sub-populations within a single population. The null hypothesis is that across the study area, the normalized distributions of the two populations—the distributions conditional on the respective population sizes—are the same. The alternative hypothesis is that there is some unspecified difference in the underlying normalized distributions. The two distributions are normalized in order to remove the effect of differing population sizes.

Population density data are collected at K sampling locations on two populations; observations must be made for each population at each location. Although the study area may be of any shape, when calculating the cumulative distribution functions, it is useful to think in terms of a rectangle A fully enclosing the actual study area. Assume a Cartesian coordinate system superimposed on the rectangle A with the origin at an arbitrarily selected corner of A and with the axes defined so that A lies in the first quadrant of the Cartesian plane (i.e., so that all sampling locations have positive coordinates). Let (x_k, y_k) denote the coordinates of the k^{th} sampling location, $\{k = 1, \ldots, K\}$; let $d_\iota(x_k, y_k)$ denote the sample density at the k^{th} sampling location of the i^{th} population. The question being evaluated is whether the data $\{d_\iota(x_k, y_k)$: $k = 1, \ldots, K$; $i = 1,2\}$ indicate a difference in the underlying normalized distributions of the two populations or whether the data may reasonably simply represent random variations drawn from a single, common underlying distribution.

To construct a test that is independent of the population sizes, first normalize the observed density data: divide each density observation by the sum of all density observations for that species; that is, let

$$\gamma_\iota(x_k, y_k) = \frac{d_\iota(x_k, y_k)}{D_\iota} \tag{1}$$

define the normalized density observations where

$$D_\iota = \sum_{k=1}^{K} d_\iota(x_k, y_k).$$

The value of the cumulative distribution function at the location (x_k, y_k) for the i^{th} population, denoted $\Gamma_\iota(x_k, y_k)$, is the sum of all normalized density observations, $\gamma_\iota(x, y)$, whose location (x, y) is such that $x \leq x_k$ and $y \leq y_k$. All of these sample locations lie within or on a rectangle within A whose diagonal runs between the origin of the Cartesian coordinates, (0, 0), and the point (x_k, y_k). Thus, the cumulative distribution function for the i^{th} population at the k^{th} sampling location can be defined as

$$\Gamma_\iota(x_k, y_k) = \sum_{\forall x \leq x_k, \forall y \leq y_k} \gamma_\iota(x, y). \tag{2}$$

Following the Cramér-von Mises analog, a statistic to test the null hypothesis is the square of the difference between the two cumulative distribution functions, summed over all sampling locations; that is

$$\Psi = \sum_{k=1}^{K} [\Gamma_1(x_k, y_k) - \Gamma_2(x_k, y_k)]^2. \tag{3}$$

Unfortunately, the statistic ψ is not invariant with respect to the corner of the rectangle A that is chosen as the origin of the coordinate system. In a univariate model with observations distributed across a single dimension, the Cramér-von Mises test statistic is invariant with respect to which extreme of the distribution is selected as the starting point for calculating the cumulative distribution functions. The set of observations included in the cumulative distribution function at a specific location when that distribution is calculated from one extreme is the complement of the set of observations included in the cumulative distribution function at the same location when the cumulative distribution function is calculated starting at the other extreme of the distribution. In contrast, when calculating a cumulative distribution function across a two-dimensional sample space, the set of observations included in the cumulative distribution function at a specific location is not the complement of the set of points included when calculating the cumulative distribution function for that location but starting at a different extreme of the distribution (i.e., at a different corner of A). Thus the invariance property of the test statistic is lost.

Likewise, a rotation of the circumscribing rectangle through an angle other than a multiple of 90° will change the value of the test statistic. But such a change is unlikely to substantially affect the level of significance of the test statistic, especially in the situation where an approximate randomization test is conducted and the level of significance of the test statistic will already vary somewhat depending on the choice of the seed for the randomization process (i.e., the random number chosen to initialize the pseudo-random number generator) used to select a subset of all possible permutations.

In general, there is no one corner of A that one would naturally define as the origin of the coordinate system. A solution to this dilemma was proposed by Zimmerman (1993), in a paper that presents a hypothesis test of whether a spatial distribution is a random distribution: calculate ψ four times, once with each corner of A defined as the origin, and average the four values to obtain a test statistic. Eqs. 2 and 3 can be rewritten as

$$\Gamma_\iota(x_{c,k}, y_{c,k}) = \sum_{\forall x \leq x_{c,k}, \forall y \leq y_{c,k}} \gamma_\iota(x, y) \tag{4}$$

and

$$\psi_c = \sum_{k=1}^{K} [\Gamma_1(x_{c,k}, y_{c,k}) - \Gamma_2(x_{c,k}, y_{c,k})]^2, \tag{5}$$

$\{c = 1,2,3,4\}$, where each value of c identifies a different corner of the rectangle A as the origin of the coordinate axes and where $(x_{c,k},y_{c,k})$ denotes the location of the k^{th} sampling location relative to the origin at corner c. $\Gamma_i(x_{c,k}, y_{c,k})$ is the i^{th} cumulative distribution function calculated with the sampling location coordinates defined relative to corner c. Finally, define the test statistic as the average of the four statistics,

$$\Psi = \frac{1}{4}\sum_{c=1}^{4} \psi_c. \tag{6}$$

The level of significance of the test statistic Ψ can be determined using the methodology of a randomization test (Edgington 1980). Under the null hypothesis, at a given sampling location (x_k, y_k), either density observation $\gamma_i(x_k, y_k)$, $i = 1,2$, is equally likely for each population. Thus, for a given data set, the distribution of the test statistic can be constructed by calculating the value of the test statistic for all 2^K pairwise permutations of the data set. The need to permute the paired observations is why each population must be sampled at each sample location. The level of significance of a specific realization of the test statistic Ψ is determined from its position in the ordered set of test statistic values from all 2^K permutations.

For most studies, the number of sampling locations, K, is large enough that it becomes computationally impractical—if not infeasible—to calculate the value of the test statistic for all 2^K permutations of the data. For example, with 15 sampling stations, there are more than 32 000 possible permutations. In that case, an approximate randomization test may be used; a large (but not exhaustive) number of randomly selected permutations is used to approximate the distribution of the test statistic Ψ and hence the level of significance of the observed value.

As part of this work, a *QuickBASIC* program was written to calculate the test statistic and the associated level of significance.[2]

An Example

The Resource Assessment and Conservation Engineering Division of the Alaska Fisheries Science Center, National Marine Fisheries Service (NMFS), conducts an annual bottom trawl survey of the eastern Bering Sea. The survey includes 329 sampling stations regularly spaced on a ≈37 × 37 km (20 × 20 nautical mile) grid. Among the species sampled in the survey is Pacific cod (*Gadus macrocephalus*). Questions of ecological interest arise about the distribution of the species: whether the sexes are similarly distributed, whether the distribution of juveniles differs from that of adults, whether the species distribution has changed from one year to the next.

Fig. 1 shows the distribution of the normalized (per Eq. 1) catch-per-unit-effort (CPUE) data by sex from the 1990 NMFS survey (Armistead and Nichol 1993). While the survey area is not rectangular, a rectangle can be drawn to encompass the study area (e.g., the frame surrounding each map in Fig. 1) and then the above-described statistical test can be used to test whether the observed difference between the two distributions is statistically significant. Because the significance test is nonparametric, there is no need to make any assumptions about the underlying distributions of the fish.

Consider one corner of the rectangular frame (say, 180° W, 54° N) as the origin of a set of coordinate axes. Calculate the two cumulative distribution functions (Eq. 4) using the normalized CPUE data as the measure of density and the sampling locations defined relative to the origin. Sum the square of the differences between the cumulative distribution functions over all sampling locations (Eq. 5) to obtain $\psi_1 = 0.244$. Repeat this process with each of the other three corners of the frame defined as the origin of the coordinate axes, yielding $\psi_2 = 0.229$, $\psi_3 = 0.094$, and $\psi_4 = 0.327$. The mean of the four statistics (Eq. 6) is the test statistic, $\Psi = 0.224$.

To calculate the level of significance of the test statistic (the P value), 1000 permutations of the data were examined, the observed permutation plus 999 pseudorandom permutations. The data set consists of 327 pairs of observations (unusable tows occurred at two stations). For each pseudo-random permutation, one normalized observation from each pair was randomly assigned to females and the other observation to males. The data were again normalized and a test statistic value calculated. The P value is the proportion of the 1000 test-statistic values that were greater than or equal to the observed test statistic. In this case, 212 of the pseudo-random test statistic values were greater than or equal to the observed test statistic $\Psi = 0.224$, yielding a value of $P = 0.213$, and indicating that the observed difference between the distributions of females and males was not statistically significant (Fig. 1).

In contrast, Fig. 2 shows the distribution of the normalized CPUE data by age class from the 1990 NMFS survey (Armistead and Nichol 1993). The test statistic for these data is $\Psi = 8.04$. Of 999 pseudo-random permutations examined, none had a test statistic value as large as 8.04 (the largest one was 5.42), yielding $P = 0.001$, and leading one to conclude that in the eastern Bering Sea, juvenile Pacific cod are distributed differently than are adults. A look at Figs. 1 and 2 shows that while there is an observable difference in the distribution of Pacific cod by sex, there is a much greater difference in the distribution by age class. This is reflected in the levels of significance of the respective test statistics.

[2] A copy of the program, either in *QuickBASIC* code or as an executable program, is available on a 3½″ diskette as ESA Supplementary Publication Service Document No. 9503. For a copy of this program, contact the author or order from The Ecological Society of America, 328 East State Street, Ithaca, NY 14850-4318 USA. There is a small fee for this service.

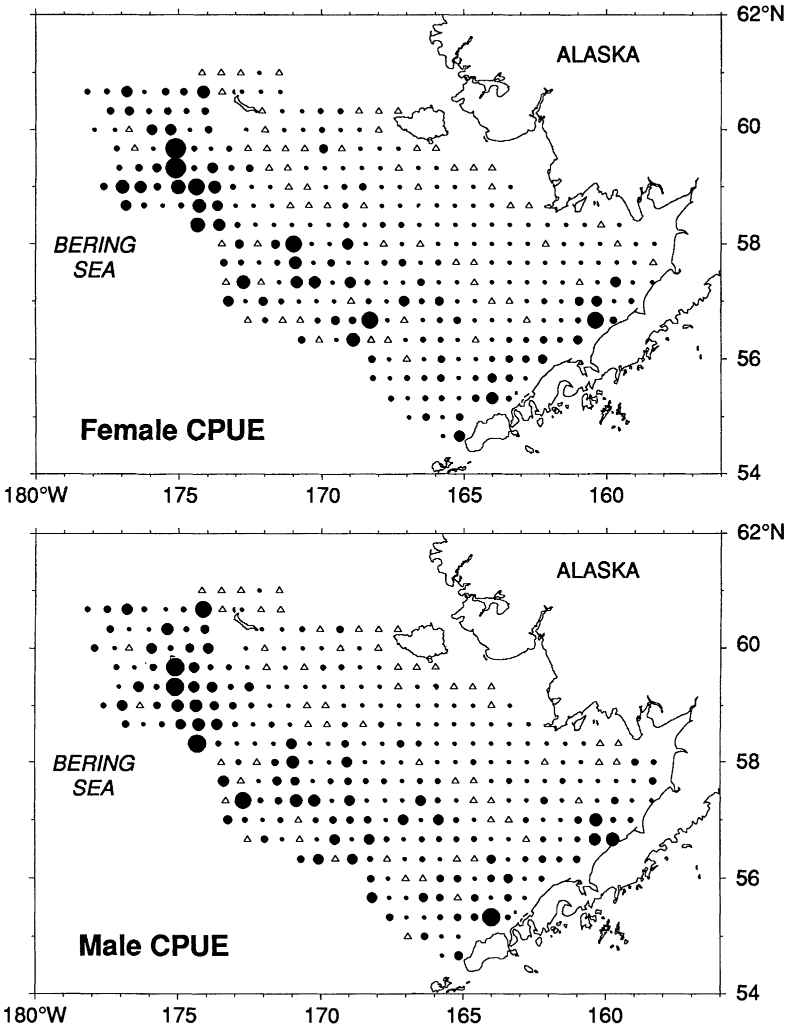

FIG. 1. Distribution of Pacific cod catch-per-unit-effort (CPUE) by sex from the National Marine Fisheries Service 1990 eastern Bering Sea bottom trawl survey (adapted from Armistead and Nichol 1993). The size of each dot represents the relative CPUE at that location; a "Δ" indicates a zero catch. The difference between the two spatial distributions is not statistically significant ($\Psi = 0.224$, $P = 0.213$).

DISCUSSION

The statistical hypothesis test presented in this paper is designed to test for a difference between the spatial distributions of two populations—or, alternatively, for a change in the distribution of a population over time. The test is in contrast with the tests offered by Zimmerman (1993) and Perry and Smith (1994). Zimmerman's test is a test of spatial randomness, of whether the spatial distribution of a single population is a random distribution. Perry and Smith's test is a test for an association between an environmental factor and the spatial distribution of a population.

The hypothesis test is nonparametric; no assumptions are required about the distributions of the two populations. The null hypothesis is that the distributions of the two populations are the same. The test statistic is based on the difference between two cumulative distribution functions. Specifically, the test statistic is the squared difference between the two cumulative distribution functions summed over all sampling locations, a Cramér-von Mises type of statistic.

A Kolmogorov-Smirnov type of test statistic (Conover 1980) was also considered while developing this test. The Kolmogorov-Smirnov statistic is the greatest difference between the same two cumulative distribution functions. When testing for a difference between two univariate probability distributions, the Kolmogorov-Smirnov test is more commonly used than is the Cramér-von Mises test. Analyses of several unpublished data sets indicated that the Kolmogorov-Smirnov approach is more sensitive to a small number of inordinately large density observations. If one population has a small number of density observations, even just one or two, that are much larger than the rest of the

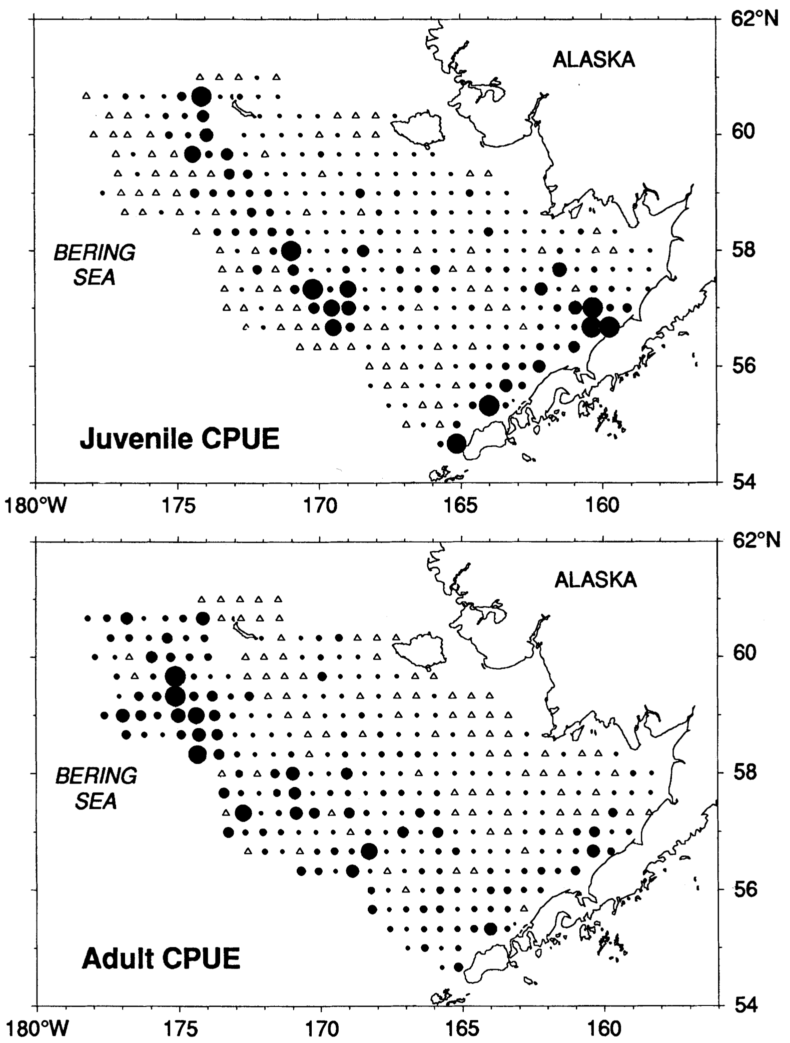

FIG. 2. Distribution of Pacific cod catch-per-unit-effort (CPUE) by age class from the National Marine Fisheries Service 1990 eastern Bering Sea bottom trawl survey (adapted from Armistead and Nichol 1993). The size of each dot represents the relative CPUE at that location; a "Δ" indicates a zero catch. The difference between the two spatial distributions is statistically significant ($\Psi = 8.04$, $P = 0.001$).

observations, the Kolmogorov-Smirnov approach is more likely to yield a statistically significant test statistic than is the Cramér-von Mises approach. In sampling the distributions of fish and other species that tend to school or otherwise exhibit contagious behavior, a small number of very large catches is not uncommon (McConnaughey and Conquest 1993). But statistical significance, in this case due to a few such large observations, does not necessarily indicate a fundamental difference in the underlying population distributions being examined. Perry and Smith (1994) developed their test in conjunction with a study of whether changes in water temperature, salinity, and depth affected the distribution of fish species. Their test uses a Kolmogorov-Smirnov type test statistic. R. I. Perry (*personal communication*) confirmed finding that their procedure is sensitive to a few large density observations, particularly if those observations occurred in an area where the environmental factor being examined was at an extreme of its range. A small number of such observations could well be due to the aggregative behavior of the species being studied rather than due to a change in distribution in response to an environmental factor. Thus, the use of a Cramér-von Mises type test statistic appears to be preferable since it appears to be more robust in the presence of a few extreme observations.

If a statistically significant difference between two distributions is found, a natural next step is to try to characterize that difference. Some differences between distributions are easily characterized—for example, two similar distributions but with one shifted in a more northerly direction or two distributions with the same central location but where one population is more widely spread across the study area than is the other—while others are not. Unfortunately, the differences between the distributions of animal populations are usually not easily described. The observed differences may reflect differing responses to irregularly distributed habitats or environmental conditions. When differences are apparent, at present the best follow-up option may be to offer a descriptive view of the differences based on the data (either with text or with graphics such as those in Figs. 1 and 2) combined, if available, with information on the habitat and environmental patterns across the study area. Perhaps an integrated approach based on this method and that of Perry and Smith (1994) would address the issue in a more comprehensive and qualitative fashion.

Acknowledgments

Thanks to Mark Zimmermann who initially raised the question of whether two species of flounder were distributed differently and especially to Bob McConnaughey for the numerous discussions and for his constructive comments after reading a draft of this manuscript. And thanks to two anonymous referees whose comments improved the quality of this manuscript.

Please note that reference to a trade name does not imply endorsement by the National Marine Fisheries Service, NOAA.

Literature Cited

Armistead, C. E., and D. G. Nichol. 1993. 1990 Bottom trawl survey of the eastern Bering Sea continental shelf. United States Department of Commerce, NOAA Technical Memorandum **NMFS-AFSC-7**.

Conover, W. J. 1980. Practical nonparametric statistics. Second edition. John Wiley & Sons, New York, New York, USA.

Edgington, E. S. 1980. Randomization tests. Second edition. Marcel Dekker, New York, New York, USA.

McConnaughey, R., and L. Conquest. 1993. Trawl survey estimation using a comparative approach based on lognormal theory. United States National Marine Fisheries Service Fishery Bulletin **91**:107–118.

Perry, R. I., and S. J. Smith. 1994. Identifying habitat associations of marine fishes using survey data: an application to the northwest Atlantic. Canadian Journal of Fisheries and Aquatic Sciences **51**:589–602.

Upton, G. J. G., and B. Fingleton. 1985. Spatial data analysis by example. Volume 1. Point pattern and quantitative data. John Wiley & Sons, New York, New York, USA.

Zimmerman, D. L. 1993. A bivariate Cramér-von Mises type of test for spatial randomness. Applied Statistics **42**:43–54.