TESTS FOR SPACE-TIME INTERACTION AND A POWER FUNCTION

D. E. BARTON*
UNIVERSITY COLLEGE, LONDON
F. N. DAVID**
UNIVERSITY COLLEGE, LONDON
and
UNIVERSITY OF CALIFORNIA, BERKELEY
EVELYN FIX**
UNIVERSITY OF CALIFORNIA, BERKELEY
MAXINE MERRINGTON
UNIVERSITY COLLEGE, LONDON
and PIERO MUSTACCHI***
UNIVERSITY OF CALIFORNIA MEDICAL CENTER, SAN FRANCISCO

1. Introduction

The problem of the detection of low order epidemicity is an old one and has been discussed moderately frequently in epidemiological literature. The first workable test may be that due to Knox [5], [6], [7], who, considering a swarm of points in the three dimensions of time and space, defined somewhat arbitrarily criteria for deciding whether points were adjacent in time and adjacent in space. He conjectured that the small number of points adjacent in both time and space was distributed Poissonwise. It was subsequently shown by Barton and David [2] that his conjecture was substantially correct. In this present paper we show that his test is very sensitive to departures from randomness in time and space commonly associated with epidemic conditions.

Various other test criteria have been proposed. It is clear that they all may be divided into three broad types, and that with appropriate modifications the same ideas regarding the power function for Knox's test can be applied to obtain power functions for any of the others.

2. Test criteria

N points $\{x_i, y_i, t_i\}$ are supposed where $\{x_i, y_i\}$ are the space points and $\{t_i\}$ those of time. It will sometimes be convenient to write

* Now at the Institute of Computer Sciences and Queen Mary College, University of London.
** With the partial support of the National Institutes of Health, USPHS Grant GM-10525.
*** With the support of National Institute of Health Research Grant 9751.

$$d_{ij}^2 = (x_i - x_j)^2 + (y_i - y_j)^2, \tag{2.1}$$

$$t_{ij} = |t_i - t_j|, \tag{2.2}$$

where $i \neq j$. Two standards $\mathfrak{D}$ and $\mathfrak{J}$ are chosen. The coordinates (x_i, y_i, t_i) and (x_j, y_j, t_j) are defined as being adjacent in space if $d_{ij} < \mathfrak{D}$ and as adjacent in time if $t_{ij} < \mathfrak{J}$.

3. Criteria depending on both $\mathfrak{J}$ and $\mathfrak{D}$

3.1. *Knox's method of counting.* Knox counts $d_{ij} = 1$ if it is less than $\mathfrak{D}$ and zero if it is greater than $\mathfrak{D}$, with a similar counting procedure for t_{ij}. Let there be N_{1S} adjacencies in space when all possible $N(N-1)/2$ comparisons are made and N_{1T} adjacencies in time. If n_{ST} denotes the number of adjacencies in both time *and* space then under the hypothesis of randomness

$$E(n_{ST}) = \frac{2N_{1S}\,N_{1T}}{N(N-1)} = \lambda, \tag{3.1}$$

$$E(n_{ST})^2 = \frac{2N_{1S}\,N_{1T}}{N^{(2)}} + \frac{4N_{2S}\,N_{2T}}{N^{(3)}} + \frac{4}{N^{(4)}}[N_{1S}^{(2)} - 2N_{2S}][N_{1T}^{(2)} - 2N_{2T}], \tag{3.2}$$

where N_2 denotes the number of pairs of adjacencies with a common point. Barton and David [2] showed that for mathematical conditions corresponding to low level epidemicity n_{ST} may be distributed in a Poisson distribution as suggested by Knox (to a good approximation).

3.2. *Extensions of Knox's method.* It is clear that Knox's method of counting is a special case of a much more general procedure. Thus, if we confine our attention to those pairs of points judged adjacent in time and judged adjacent in space, we may begin by ignoring the nonadjacencies—which is equivalent to counting them all zero—but what we do then is a matter of choice. We can take some function of the d_{ij} and of the t_{ij}, either the quantities themselves, or their squares or their reciprocals for example, and are then left with sums such as

$$R_f \equiv \sum_{N_T}\sum_{N_S} f(d_{ij})\,f(t_{ij}) \equiv \sum_{n_{ST}} f(d_{ij})\,f(t_{ij}). \tag{3.3}$$

The distribution under randomization of any of such criteria can be obtained and the choice of the most sensitive can only be made by a calculation of the power of the criteria to detect departures from randomness such as that described below for Knox's case. Given modern computing methods, this does not present any intrinsic difficulty. Barton and David [2] provide tables which enable the first four moments of R_f under randomization to be computed. From these, percentage points are easily obtained by standard graduation methods.

4. Criteria depending on $\mathfrak{J}$ or on $\mathfrak{D}$

4.1. *Group according to* $\mathfrak{J}$. Barton and David [2] suggested making a division of the N observations into groups according to $\mathfrak{J}$ and then calculating the within groups variation of the space distances. Such a criterion has been shown to be sensitive for many diseases where the epidemicity is reasonably widespread and the contagion known to be present.

4.2. *Group according to* $\mathfrak{D}$. Although the procedure under subsection 4.1 is simple because the $\{t_{ij}\}$ are arranged in natural order (over j for a given i), there is no reason why the data should not be divided into groups according to $\mathfrak{D}$ and then within groups variation of the time distances computed.

5. Criteria depending on neither $\mathfrak{I}$ nor $\mathfrak{D}$

We may choose a function of the space distance and a function of the time distance between any two points and sum their product over all possible points. Thus, we would take

$$\sum_{i>j} f(d_{ij})\, f(t_{ij}). \tag{5.1}$$

Mantel has put forward criteria based on $f(d_{ij}) = d_{ij}$, $f(t_{ij}) = t_{ij}$ or their reciprocals; David and Fix have suggested, for algebraic convenience, taking the easily computed statistic

$$f(d_{ij}) = d_{ij}^2, \quad f(t_{ij}) = t_{ij}^2. \tag{5.2}$$

Many other similar criteria can be suggested.

6. Criteria and data inadequacies

In studying the set of N observations numerically it will often be observed that there are very large intervals in both space and time distances. Such large intervals are commonly regarded with suspicion, and workers in this field seem to prefer to choose criteria which will minimize their effect. Knox's criterion and those allied to it (as mentioned in section 1) get over this difficulty of the large intervals simply by ignoring them. With any criterion, however, based on a division of the observations into groups, one is faced with having to decide what values to take for $\mathfrak{I}$ and/or $\mathfrak{D}$. While this may be done, using the simple rule set out in Barton, David, and Merrington [3], the effect of having to make such a choice introduces a certain arbitrary element into the analysis.

To refuse to make a division of the observations and to take a criterion R_f as in section 3 means that one accepts the fact that there are these long intervals—which may be due to observations not having been recorded—and that one's test will be correspondingly less sensitive in picking out the space-time interaction. The device of taking

$$f(d_{ij}) = \frac{1}{d_{ij}}, \qquad f(t_{ij}) = \frac{1}{t_{ij}} \tag{6.1}$$

is not to be recommended. The data are commonly such that t_{ij} or d_{ij} can be zero—that is, two or more cases can be recorded as occurring simultaneously—and to assign an arbitrary value to t_{ij} or d_{ij} in such cases is to beg the question.

It is worth noting that the Knox statistic is of the R_f form, f being a zero-one (indicator) function.

7. A model for power function

Barton and David pointed out that for Knox's method of analysis the data could be represented in terms of a space graph, a time graph, and a space-time graph. Points which are adjacent (according to the given definition) are joined, otherwise no join is made. Under the null hypothesis, the time and space graph is a random intersection of the time graph with the space graph. Under the alternative hypothesis proposed here a small number out of the N^* cases have arisen by contagion. We suppose the observed time and space graphs to be built up in two stages.

(i) A set of N primary cases are allocated to the N time and N space points independently. There will be N_T joins in the time graph and N_S joins in the space graph, and $X = n_{ST}$ joins in the combined space-time graph.

(ii) Assume that some of the N primary cases give rise to m secondary cases within the adjacency space-time vicinity of a primary. Consequently we observe not X but $X^* \geqq X + m$ actual cases adjacent in time and space, out of the $N^* = N + m$ cases which are actually recorded. The m secondaries will provide m extra joins with their own primaries in the space-time graph and possibly others either within themselves or with existing primaries other than those which gave rise to them. The total time adjacencies will also have increased at least by m, say to N_{1T}^*, and similarly for space adjacencies, say N_{1S}^*.

(iii) It is further assumed that each secondary is equally likely to arise from any primary and that the secondaries themselves arise independently. If each secondary is assumed to be so close to its primary as to share all its adjacency properties this will be equivalent for working purposes to assuming the primary and the secondary are coincident. We should emphasize that this is for working purposes only. In practice all that will be observed will be two cases, very close in time and space, which share each other's adjacencies.

8. Asymptotic power function using Poisson limit

It is intuitively clear that if

$$\lambda = \frac{2N_{1T}\,N_{1S}}{N(N-1)}, \qquad \lambda^* = \frac{2N_{1T}^*\,N_{1S}^*}{(N+m)(N+m-1)}, \tag{8.1}$$

then under the "incoherent" limit condition [2], $\lambda^* \to \lambda$ and $(X^* - m)$ tends to be distributed as a Poisson distribution with mean λ. Thus, if we take Knox's data, where $\lambda = 5/6$, and write the *asymptotic* power function as $\beta(m)$, then using an approximate five per cent significance level we have $\beta(0) = 0.052$, $\beta(1) = 0.203$, $\beta(2) = 0.565$, $\beta(3) = 1.000$. To study the power function more accurately we need to investigate

$$E(\lambda^*) = \Lambda, \qquad E(X^*) = \mu \tag{8.2}$$

(say), since the first approximation to the power function will be to calculate the effect of using the percentage points of a Poisson distribution with mean Λ when X^* is actually distributed with (noncentral) mean μ.

9. Use of mean Λ with Knox's data

Let the m secondaries be divided into m_1 (singles), m_2 (doubles), m_3 (triplets), and so on, so that

$$m = \sum_{r=1} rm_r. \tag{9.1}$$

If ρ_{jT}^* equals the observed local degree of the jth point of the time graph given $(m_1, m_2, \cdots)$, then

$$\begin{aligned} E(N_{1T}^*) &= \frac{1}{2}\sum_{j=1}^{N} E(\rho_{jT}^*) \\ &= \frac{1}{2}\sum_{j=1}^{N} \rho_{jT} + \sum_{j=1}^{N} (\rho_{jT} + 1)\frac{m}{N} + \frac{1}{2}\sum_{j=1}^{N} \rho_{jT} \left[\frac{m^2 - \sum r^2 m_r}{N^{(2)}}\right], \end{aligned} \tag{9.2}$$

where the sums on the right side refer to primary-primary joins, primary-secondary joins, and secondary-secondary joins (in that order).

Consequently, conditional on $(m_1, m_2, \cdots)$, we have

$$E(N_{1T}^*) = m + N_{1T}\left[1 + \frac{2m}{N} + \frac{m^2 - \sum r^2 m_r}{N^{(2)}}\right]. \tag{9.3}$$

Now

$$E\left(\sum_r m_r r^2\right) = N\left[\frac{m^{(2)}}{N^2} + \frac{m}{N}\right], \qquad \mathrm{Var}\left(\sum_r m_r r^2\right) = \frac{2m^{(2)}}{N^3 N^{(2)}}, \tag{9.4}$$

and it follows that

$$\begin{aligned} E(\lambda^*) &= \frac{2}{(N+m)^{(2)}} E(N_{1S}^* N_{1T}^*) \\ &= \frac{1}{{}^{N+m}C_2}\left\{{}^{N}C_2\lambda\left[\left(1 + \frac{2m}{N} + \frac{m^{(2)}}{N^{(2)}}\right)^2 + \frac{2m^{(2)}}{N^3 N^{(2)}}\right]\right. \\ &\qquad \left. + m\left(1 + \frac{2m}{N} + \frac{m^{(2)}}{N^2}\right)(N_{1T} + N_{1S}) + m^2\right\}. \end{aligned} \tag{9.5}$$

If we consider Knox's data, $N = 96$, $\lambda = 5/6$, $N_{1S} = 152$, $N_{1T} = 25$ and we let $m = 3$, the greatest value which was required for the $\beta(m)$ of the power function above, we have, on substitution

$$E(\lambda^*) = 1.004 = \Lambda. \tag{9.6}$$

This is greater than the value for λ, but it is interesting to note that the change from λ to Λ is not large. In particular the critical value which gives an approximate five per cent level of significance is the same using 1.004 as it is using 5/6. This question is discussed further in the next section.

10. Validity of asymptotic power functions

To study the mean of X^*, we use the fact that the m secondaries are independent and consider the distribution of secondaries for a fixed set of primary cases. Let α_j denote the number of joins of the jth secondary to primaries and β_j denote the joins of the jth secondary to other secondaries. Consequently,

$$X^* = X + \sum_{j=1}^{m} \alpha_j + \frac{1}{2}\sum_{j=1}^{m} \beta_j, \tag{10.1}$$

$$E(\sum \alpha_j) = m\left(1 + \frac{2X}{N}\right), \tag{10.2}$$

$$E(\sum \beta_j) = \frac{m^{(2)}}{N^2}(N + 2X) \tag{10.3}$$

giving

$$E(X^*) = X\left[1 + \frac{2m}{N} + \frac{m^{(2)}}{N^2}\right] + m + \frac{m^{(2)}}{2N}. \tag{10.4}$$

Averaging now over the primary cases, since

$$E(X) = \frac{N_{1T} N_{1S}}{N}, \tag{10.5}$$

we have

$$\begin{aligned} E(X^*) &= \frac{N_{1T} N_{1S}}{N}\left[1 + \frac{2m}{N} + \frac{m^{(2)}}{N^{(2)}}\right] + m + \frac{m^{(2)}}{2N} \\ &\doteqdot \lambda + m + O\left(\frac{1}{N}\right). \end{aligned} \tag{10.6}$$

For general values of m it is seen that $X^* - m$ has expected value of the form $\lambda + \epsilon_m$ where, in the case $\lambda = 5/6$, $N = 96$ (as for Knox's data), $\epsilon_m = 0.017$, 0.045, 0.084 for $m = 1, 2, 3$, respectively. The chance that a Poisson variable (of mean Λ) is 3 or more we have seen to be 0.052 when Λ is 5/6. When $\Lambda = 1.004$ it is 0.081, while the chance the variable exceeds 3 is 0.019 in this case. Thus, since Λ takes intermediate values for $m = 1$ and 2, the approximate five per cent point has the value 3 for each of $m = 0, 1, 2, 3$. That is to say, we should, to this degree of approximation, use the same critical values for the whole effective range of m. Using this value and approximating to the distribution of $X^* - m$ by that of a Poisson variable with mean $\lambda + \epsilon_m$, we find the power of the test to have the values 0.052, 0.210, 0.605, 1.000 for $m = 0, 1, 2, 3$, respectively. This indicates that the "asymptotic" value of the power is an adequate approximation for all practical purposes.

In the foregoing, the use of m as the parameter for the power function for a fixed number of primaries has meant that as m increases so does the number of cases (analogous to "sample number" in classical statistical theory of tests). It is arguable that a better, if less realistic, idea of the power would be given if the "sample number" were held constant. On the other hand, the fundamental consideration which necessitates the randomization test (such as the X test) is the fact that the time and the space graphs both have such a complicated irregularity that it is not possible to specify them by a deterministic pattern nor yet by any acceptable random distribution of time or space coordinates. We have therefore preferred to keep their specification fixed as actually observed, at the cost of allowing the "sample number" of cases to vary. In the event, we have seen that this variation is negligible and does not arise in the "asymptotic" power, which our results suggest gives an adequate description. It needs to be noted, however, that there is a further factor which arises whether the sample number is kept fixed or not and which is due to the discrete nature of the variation of X and hence of the critical value for it as a test function. That is to say, in terms of the particular case considered, while the nominal level of significance is held at 5 per cent, this is only approximate. In the case considered the actual level varies from 5.2 per cent to 8.1 per cent over the range of m concerned. This variation is only marginal in this case and does not arise in the asymptotic form, but clearly has to be considered on its merits in respect of actual time or space graphs when the number of cases is small.

The general form for the variance of $X^* - m$ is expressible, after some algebraic labor by

$$\text{Var}\,(X^* - m) = \sigma_X^2(1 + A) + \lambda B + \lambda^2 C + DM_2 + EM_3 + F, \tag{10.7}$$

where A, B, C, D, E, F are purely functions of m and N (as specified below), λ and σ_X^2 are the null hypothesis mean and variance of X, and M_r is the mean rth factorial moment of the local degree of the time and space graph. In terms of the corresponding moments of the time graph and of the space graph (denoted by suffixes S and T) we have $M_r = M_{rT}M_{rS}/(N - 1)^{(r)}$. For the particular time and space graphs of Knox's data, $M_{rT} = 19/6, 71/16, 85/4$; $M_{rS} = 25/48, 5/24, 0$, for $r = 1, 2, 3$ respectively. Thus, $M_2 = 71/685824$ and $M_3 = 0$. The $\{M_r\}$ will generally be as negligible as this from the incoherence of the parent graphs. The numbers A, B, C, D, E, and F are all small, when m is small compared with N as here. Thus, $X^* - m$ has substantially the same variance as X and substantially the same mean which provides additional confirmation of the validity of the approximation by the asymptotic results. Explicitly,

$$\begin{aligned}
E &= \frac{m^{(3)}}{4N^2}, \qquad D = m + \frac{m^{(3)}\,(m-1)}{4N^{(2)}} - \frac{5m^{(3)}}{4NN^{(2)}},\\
A &= \frac{4m}{N} - \frac{2m^{(2)}}{N^2} + \frac{4m^{(3)}}{N^3\,(N-1)} + \frac{m^{(4)}}{N^4},\\
B &= \frac{2m}{N} - \frac{2(m-1)m^2}{N^2} - \frac{(m^{(2)})^2}{N^3} - \frac{m^{(3)}}{N^2\,N^{(2)}} + \frac{(3N+1)m^{(4)}}{2N^2\,N^{(2)}},\\
C &= -\frac{2m}{N} + \frac{3m^{(2)}}{N^2} + \frac{4m^{(3)} + m^{(4)}}{N^2\,N^{(2)}},\\
F &= \frac{m^{(2)}}{4N} + \frac{m^{(4)} - (m^{(2)})^2}{4N^2}.
\end{aligned} \tag{10.8}$$

Since $\sigma_X^2 = 0.802$ for Knox's data we see that Var $(X^* - m) = 0.802, 0.839, 0.880, 0.892$, for $m = 0, 1, 2, 3$. These may be compared with the mean for $X^* - m$ which takes the values 0.8333, 0.850, 0.878, 0.917. So far as agreement of mean and variance are concerned, it is clear that the Poisson approximation is actually better over this range of m than it is under the null hypothesis. Since this is clearly due to a relative lengthening of the upper tail of the distribution, the improvement will be particularly in just that region of the distribution which affects the power function.

The formulae for any m are complicated but since the power for $m = 1$ is of reasonable size and since for this value the formulae are comparatively simple, they are perhaps worth noting. Let E_1 denote the operation of averaging over all possible positions of the single secondary point. We have

$$(10.9) \qquad E_1(N_1^*) = 1 + \frac{N+2}{N}N_1, \qquad E_1(N_2^*) = N_2 + \frac{3(N_2 + 2N_1)}{N}.$$

For Knox's data where $N_{2T} = 426$, $N_{2S} = 6$ and the other values are as previously given, we have

$$(10.10) \qquad E(X^* - 1) = 0.85, \qquad \text{Var}\ (X^* - 1) = 0.84,$$

indicating that the Poisson approximation is probably still adequate. For m large and N small the Poisson approximation will not be a good one. The model set up for the alternative hypothesis corresponds almost exactly with that set up by Neyman in devising his "contagious" distributions, and we may expect the variance in our case to become greater than the mean as m increases. Given, however, that numerical exploration shows that this is indeed the case, there would appear to be no difficulty in assuming a compound Poisson distribution for X^* in the alternative.

The variants on Knox's criterion may be investigated along similar lines. Let t_{ij} be the time distance between the ith and jth points, always counted positive, whether $i < j$ or $i > j$, and d_{ij} the corresponding space distance. Let t_{ij} equal zero if it is greater than $\mathfrak{T}$ and similar for $d_{ij} > \mathfrak{D}$ and write

$$(10.11) \qquad \varphi = \sum_{i \neq j} t_{ij}\, d_{ij}.$$

Under the null hypothesis

$$(10.12) \qquad E(\varphi) = N^{(2)}\, E(t_{ij})\, E(d_{ij}).$$

The expected value of t_{ij} will just be the average over the whole set of those distances which are counted. If we write these as $T_{\ell k}(\ell = 1, \cdots, N, k = 1, \cdots, N, \ell \neq k)$, then

$$(10.13) \qquad E(t_{ij}) = \frac{1}{N^{(2)}} \sum_{\ell \neq k} T_{\ell k} = \overline{T}_1,$$

say. Consequently,

$$(10.14) \qquad E(\varphi) = N^{(2)}\, \overline{T}_1 \overline{D}_1.$$

Adding a secondary point to any point of the graph and averaging over all points we have

(10.15)
$$E_1(t_{ij}) = \frac{(N+2)(N-1)}{(N+1)^{(2)}} \overline{T}_1,$$
$$E_1(\varphi) = \frac{(N+2)^2(N-1)^2}{(N+1)^{(2)}} \overline{T}_1\overline{D}_1,$$

so that the expectation of φ is increased. Again, we have

(10.16)
$$\varphi^2 = 2\sum_{i\neq j} t_{ij}^2 d_{ij}^2 + 4\sum_{i\neq j\neq k} t_{ij}t_{ik}d_{ij}d_{ik} + \sum_{i\neq j\neq k\neq \ell} t_{ij}t_{k\ell}d_{ij}d_{k\ell}.$$

Under the null hypothesis

(10.17)
$$E(t_{ij}^2) = \frac{1}{N^{(2)}}\sum_{u\neq v} T_{uv}^2 = \overline{T}_2, \qquad \text{say,}$$
$$E(t_{ij}t_{ik}) = \frac{1}{N^{(3)}}\sum_{u\neq v\neq w} T_{uv}T_{uw},$$
$$E(t_{ij}t_{k\ell}) = \frac{1}{N^{(4)}}\left[(N^{(2)})^2\overline{T}_1^2 - 2N^{(2)}\overline{T}_2 - 4\sum_{u\neq v\neq w} T_{uv}T_{uw}\right],$$

(10.18)
$$E(\varphi^2) = 2N^{(2)}\overline{T}_2\overline{D}_2 + \frac{4}{N^{(3)}}\sum_{u\neq v\neq w} T_{uv}T_{uw}\sum_{u\neq v\neq w} D_{uv}D_{uw}$$
$$+ \frac{1}{N^{(4)}}\left[(N^{(2)})^2\overline{T}_1^2 - 2N^{(2)}\overline{T}_2 - 4\sum_{u\neq v\neq w} T_{uv}T_{uw}\right]$$
$$\left[(N^{(2)})^2\overline{D}_1^2 - 2N^{(2)}\overline{D}_2 - 4\sum_{u\neq v\neq w} D_{uv}D_{uw}\right].$$

Under the alternative hypothesis

(10.19)
$$E(t_{ij}^2) = \frac{(N+2)(N-1)}{(N+1)^{(2)}}\overline{T}_2,$$
$$E(t_{ij}t_{ik}) = \frac{1}{(N+1)^{(3)}}\left[(N-1)\overline{T}_2 + \frac{N+5}{2N}\sum_{i\neq j\neq k} T_{ij}T_{ik}\right],$$
$$E(t_{ij}t_{\ell k}) = \frac{1}{(N+1)^{(4)}}\left[\frac{N+1}{N}\right]\left[(N^{(2)})^2\overline{T}_1^2 - 2N^{(2)}\overline{T}_2 - 4\sum_{u\neq v\neq w} T_{uv}T_{uw}\right],$$

with the second moment of φ following in the usual way. An assumption of normality for the distribution of φ would appear to be adequate.

A criterion containing unit powers of t_{ij} and d_{ij} is perhaps preferable in that its randomization distribution is apt to be closer to the normal approximation than those involving higher powers. It was this, possibly, which led Mantel to choose for his criterion (section 5)

(10.20)
$$\sum_{i<j} t_{ij}d_{ij}.$$

On the other hand, the square root sign implicit in d_{ij} makes it difficult to express the mean, variance and higher moments of the criterion in simple form and while the task of computing all possible space and time distances may be carried out fairly easily by modern high speed machines, there is no doubt that it is preferable to be able to express moments in terms of the original observations. To illustrate this, let us consider

$$\theta = \frac{1}{N^{(2)}} \sum_{i \neq j} t_{ij}^2 d_{ij}^2, \tag{10.21}$$

a criterion suggested by David and Fix. For illustrative purposes suppose we assume the conditions of section 3 which means that we consider the whole set of data. Let

$$\begin{aligned} m_a &= \frac{1}{N} \sum_{i=1}^{N} (t_i - \bar{t})^a, \\ M_{ab} &= \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^a (y_i - \bar{y})^b. \end{aligned} \tag{10.22}$$

Then,

$$E(\theta) = \frac{4N^2}{(N-1)^2} m_2 (M_{20} + M_{02}) = \frac{4N^2}{(N-1)^2} m_2 A, \tag{10.23}$$

say, and after some manipulation, using the augmented monomial symmetric functions of David and Kendall and the augmented binomial symmetric functions of David and Fix [4], we have

$$\begin{aligned} \operatorname{Var} \theta = {} & \frac{4N}{(N-1)^{(3)}(N-1)^2} \\ & \Big\{ m_4 [B(N^3 - N^2 - 2N + 4) - A^2(N^3 - N^2 - 2N - 4) \\ & - 8C(N-1)] + m_2^2 \Big[-B(N^3 - N^2 + 6N - 12) \\ & + A^2 \left(\frac{N^4 - 2N^3 - N^2 + 6N - 12}{N-1} \right) \\ & + 8C(N^2 - 3N + 3) \Big] \Big\}, \end{aligned} \tag{10.24}$$

where A, B and C are the invariants previously defined by Barton and David [1], namely,

$$\begin{aligned} A &= M_{20} + M_{02}, \\ B &= M_{40} + 2M_{22} + M_{04}, \\ C &= M_{20}^2 + 2M_{11}^2 + M_{02}^2. \end{aligned} \tag{10.25}$$

The higher moments are not difficult to obtain, given the necessary symmetric function tables, although the algebraic manipulation is fairly heavy. The fact that they may all be expressed as simple functions of the moments of the coordinates of the set $\{x_i, y_i, t_i\}$, with $i = 1, 2, \cdots, N$ makes the numerical calculations quite straightforward.

The moments of θ under the alternate hypothesis that there is one point which is really a doublet follow the lines previously indicated. For example,

$$E(t_{ij}^2)_{H_1} = \frac{(N+2)(N-1)}{(N+1)^{(2)}} E(t_{ij}^2)_{H_0}, \tag{10.26}$$

the other quantities required following in similar fashion. The fact that in calculating the moments of this criterion it is possible to consider the points themselves rather than the distances makes the entire algebraic manipulation much easier to set out.

It will be recognized that the scheme for the alternative hypothesis to randomness in space-time which we have set out here—primary points with secondary points attached to them—is almost exactly that envisaged by Neyman [8] when he derived his contagious distributions. It does, moreover, seem particularly applicable to the space-time interaction problem, since the addition of very few secondary points indeed seems to raise the power of detection of the departure from randomness nearly to unity. The difficulties with the calculation of such a power function are introduced purely through the technique of randomization which, it is generally agreed, is the appropriate method for this particular type of problem. The possibility of other alternate hypotheses, or of reframing the present alternative so that the combinatorial processes involved are not so intricate, is one which is at present being explored.

REFERENCES

[1] D. E. Barton and F. N. David, "Randomization bases for multivariate tests. I. Bivariate case, randomness on N points in a plane," *Bull. Inst. Statist. Internat.*, Vol. 39 (1962), pp. 158–159; pp. 455–467.

[2] — —, "The random intersection of two graphs," *Research Papers in Statistics* (edited by F. N. David), New York, Wiley, 1966.

[3] D. E. Barton, F. N. David, and M. Merrington, "A criterion for testing contagion in time and space," *Ann. Hum. Genet.*, Vol. 29 (1965), pp. 97–102.

[4] F. N. David, M. G. Kendall, and D. E. Barton, *Symmetric Functions and Allied Tables*, Cambridge, Cambride University Press, 1966.

[5] G. Knox, "Secular pattern of congenital oesophageal atresia," *Brit. J. Prev. Soc. Med.*, Vol. 13 (1959), pp. 222–226.

[6] ———, "Detection of low intensity epidemicity. Application to cleft lip and palate," *Brit. J. Prev. Soc. Med.*, Vol. 17 (1963), pp. 121–127.

[7] ———, "Epidemiology of childhood leukemia in Northumberland and Durham," *Brit. J. Prev. Soc. Med.*, Vol. 18 (1964), pp. 17–24.

[8] J. Neyman, "On a new class of 'contagious' distributions applicable in entomology and bacteriology," *Ann. Math. Statist.*, Vol. 10 (1939), pp. 35–57.