The Annals of Mathematical Statistics
1969, Vol. 40, No. 1, 144–151

A CENTRAL LIMIT THEOREM FOR THE NUMBER OF EDGES IN THE RANDOM INTERSECTION OF TWO GRAPHS

By O. Abe

University of Ibadan, Nigeria

0. Summary. The distribution of the statistic X which is the number of edges in the intersection graph $G_1 \cap G_2(V, E_1 \cap E_2)$ of $G_1(V, E_1)$ and $G_2(V, E_2)$ is investigated through its moments. An expression is obtained for the rth central moment and the moment ratios of X are, under a set of sufficient conditions, shown to approximate to those of a normal variable with the standardised variable.

$$Z = \{X - \epsilon(X)\}/(\text{var }(X))^{\frac{1}{2}}$$

having an asymptotically unit normal distribution.

1. Introduction. David and Barton [1] (1965), gave a set of conditions under which the number of edges in the random intersection of two graphs has an asymptotic Poisson distribution. The stantistic, X say, which corresponds to the number of such edges was first discussed by Knox [2] and [3] (1963) and (1964) respectively for 'epidemicity' in the field of epidemiological statistics. It has been recognised as providing, for the first time, a valid test, e.g. Doll [4].

So far, no conditions for asymptotic normality have been discovered (except in so far as David and Barton's Poisson limit may be used to provide a first stage in a double limit using the central limit theorem for a Poisson variate). We show here that X has an asymptotically normal distribution under fairly wide conditions.

Two graphs $G_1(V_1, E_1)$ and $G_2(V_2, E_2)$ are supposed to consist of sets of vertices V_1 and V_2 and sets of edges E_1 and E_2 respectively and the vertices of one are randomly mapped on to those of the other. Without loss of generality we may consider them as consisting of the same set of vertices (that is $V_1 \equiv V_2$) but not necessarily having the same set of edges (that is E_1 and E_2 are not necessarily the same). A pair of vertices v_1 and v_2 which have an edge between them in say, G_1, are said to be adjacent in G_1. The graph G defined by

$$G(V, E) \equiv G(V, E_1 \cap E_2);$$

$V \equiv V_1 \equiv V_2$ is called the intersection graph and X is the number of edges in it. We study the distribution of X where the vertices of G_1 are mapped at random on to those of G_2. The asymptotic behaviour is as n, the number of vertices in G, tends to infinity. Clearly we have to envisage a sequence of pairs of graphs G_1 and G_2 and the conditions describe the asymptotic behaviour of each pair as n tends to infinity.

Received 9 November 1967.

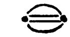
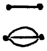
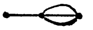
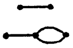
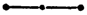
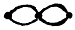
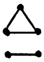
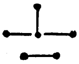
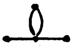
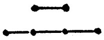
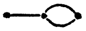
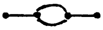
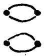
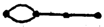
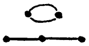
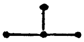
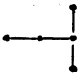
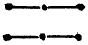
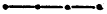
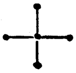
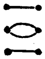
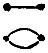
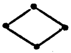
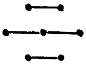
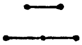
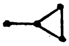
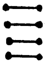
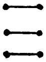
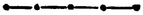
Subgraph	Notation	Subgraph	Notation	Subgraph	Notation
	ilj		il^{4}j		ilj,kl^{3}l
	il^{2}j		iljl^{3}k		ilj,klll^{2}m
	iljlk		il^{2}jl^{2}k		ilj,klllmlk
	ilj,kll		il^{2}jlkli		ilj,kllklmkln
	il^{3}j		il^{2}jilkill		ilj,klllmln
	iljl^{2}k		iljl^{2}kll		il^{2}j,kl^{2}l
	iljkli		il^{2}jkll		il^{2}j,klllm
	iljilkill		iljklllklm		iljlk,llmln
	iljkll		iljilkillilm		ilj,kll,ml^{2}n
	ilj,kl^{2}l		iljkllli		ilj,kll,mlnlp
	ilj,klllm		iljkllllj		ilj,kll,mln,plq
	ilj,kll,mln		iljkllllm		

FIG. 1

2. Notation. We require some shorthand notation for the subgraphs of both graphs G_1 and G_2. The expressions for the moments of X are in terms of the graph symmetric functions of subgraphs with 1, 2, $\cdots$, r (for the rth moment) edges. We require in particular notation for subgraphs with 1, 2, 3, or 4 edges and these are set out in Figure 1. We also let ρ_i denote the degree (sometimes called valency) at the ith vertex and $\bar{\rho}_1$, $\bar{\rho}_2$ the average degrees in G_1 and G_2, respectively. Both $\bar{\rho}_1$ and $\bar{\rho}_2$ depend on the number of vertices in and the degree of connectedness of G_1 and G_2 respectively. $\nu G(V, E)$ denotes the number of non-isomorphic ways of labelling the vertices of $G(V, E)$ with $v_1, \cdots, v_V$.

3. Moments of X. Let $\alpha_{ij} = 1(0)$ if the vertices i and j are (are not) adjacent in G_1. β_{ij} is similarly defined for G_2. Then

$$X = \sum_{i<j} \alpha_{ij}\beta_{ij} = \sum_{i<j} x_{ij}$$

where

$$x_{ij} = \alpha_{ij}\beta_{ij}, \qquad \mu_r' = \epsilon(X^r) = \epsilon(\sum_{i<j} x_{ij})^r.$$

Let W denote the set of all pairs (i, j) $i < j$ in $(1, n)$; let

$$\omega_i; \qquad i = (1, 2), (1, 3), \cdots, (1, n), (2, 3), \cdots, (n - 1, n)$$

denote the elements of W.

$$\begin{aligned}(\sum_{i<j} x_{ij})^r &= (\sum_w x_{\omega_i})^r = (\sum_w \alpha_{\omega_i}\beta_{\omega_i})^r \\ &= \sum_w \alpha_{\omega i_1}^r \cdots \alpha_{\omega i_e}\beta_{\omega i_1}^r \cdots \beta_{\omega i_1}^{r_e} r!/\prod r_i\,!\end{aligned}$$

where $\sum_{i=1}^e r_i = r$, and the summation is over all values $1 \leqq e \leqq r$ and ω_i's in W. Hence

$$\begin{aligned}\text{(3.1)} \qquad \mu_r' &= \epsilon\{\sum_W \alpha_{\omega i_1}^{r_1} \cdots \alpha_{\omega i_e}^{r_e}\beta_{\omega i_e}^{r_e} \cdots \beta_{\omega i_e}^{r_e} r!/(r_1! \cdots r_e!)\} \\ &= \sum_W \epsilon(\alpha_{\omega i_1}^{r_1} \cdots \alpha_{\omega i_e}^{r_e})\epsilon(\beta_{\omega i_1}^{r_1} \cdots \beta_{\omega i_e}^{r_e}) r!/(r_1! \cdots r_e!)\end{aligned}$$

since by the null hyopthesis G_1 is independent of G_2.

Let $\nu G(V, E)$ denote the number of non-isomorphic ways the labels $v_1, \cdots, v_V$ may be given to the vertices of $G(V, E)$ and suppose a subgraph $G(V, E)$ consists of π connected components, the ith one occurring π_i times. Then

$$\nu G(V, E) = v!\ \prod_{i=1}^c (\nu G(v_{i_1} e_i))^{\pi i}/\prod_{i=1}^c (v_i!)^{\pi i}$$

where $\sum_{i=1}^c v_i = v$; $\sum_i \sum_{k=1}^{\pi k} r_{i_k} = r$, $\sum \pi_i = \pi$; k is the number of edges in the ith component, not counting multiplicities, and there are c different connected subgraphs in $G(V, E)$. David and Barton [1] showed that the sums $\sum_w x_{\omega i_e} \cdots x_{\omega i_e}^{r_e}$ are graph symmetric functions which are fixed characteristics of G_1 and G_2. These, when divided by the number of terms they contain, are equivalent to finite population moments. That is

$$\epsilon(G(v, r)) = \mu[G(v, r)] = [G(v, r)]/NG(v, r),$$

where $NG(v, r)$, the number of terms in $G(v, r)$, is given by

$$NG(v, r) = \nu G(v, r) n_{C_v}.$$

When the symmetric functions are augmented we use square brackets and then $NG(v, r) = \nu G(v, r) AG(v, r) n_{C_v}$ where $AG(v, r)$ is the augmentation factor.

$$\mu_r' = \sum{}^* NG(v, r)\mu[G(v, r)]_1\mu[G(v, r)]_2/AG(v, r)$$

with summation over all subgraphs with r edges (including multiple ones).

$$\begin{aligned}\mu_r = \textstyle\sum_{s=0} \{ {}^{r}c_s(-1)^s[NG(2, 1)\mu[G(2, 1)]_1\mu[G(2, 1)]_2/AG(2, 1)]^s \\ \cdot \textstyle\sum^* NG(v, r-s)\mu[G(v, r-s)]_1\mu[G(v, r-s)]_2 \\ \cdot \varphi G(v, r-s)/AG(v, r-s)\end{aligned}$$

(φ being the multinomial coefficient $r!/\prod r_i!$).

$$\begin{aligned}\mu_r = {} & \textstyle\sum^{*1} NG(v, r)\mu[G(v, r)]_1\mu[G(v, r)]_2\varphi G(v, r)/AG(v, r) + \sum^{*2} N\{kG(2, 1) \\ & + G(v, r-k)\}\{\mu[kG(2, 1) + G(v, r-k)]_1\mu[kG(2, 1) \\ & + G(v, r-k)]_2\varphi\{kG(2, 1) + G(v, r-k)\}/A\{kG(2, 1) + G(v, r-k)\} \\ & + (\textstyle\sum^{*3})_{s=1}^{r} {}^{r}c_s(-1)^s\{NG(2, 1)\mu[G(2, 1)]_1\mu[G(2, 1)]_2AG(2, 1)\}^s \\ & \cdot \textstyle\sum^* NG(v, r-s)\mu[G(v, r-s)]_1\mu G(v, r-s)]_2\varphi G(v, r-s)/AG(v, r-s) \\ & + \textstyle\sum^{*4} N\{k_1G(2, e \geqq 2) + k_2G(3, e \geqq 2)\}\mu[k_1G(2, e \geqq 2) \\ & + k_2G(3, e \geqq 2)]_1\mu[k_1G(2, e \geqq 2) + k_2G(3, e \geqq 2]_2\varphi\{k_1G(2, e \geqq 2) \\ & + k_2G(3, e \geqq 2)\}(A\{k_1G(2, e \geqq 2) + k_2G(3, e \geqq 2)\})\end{aligned}$$

where

$\sum^{*1}$ is over all terms for which $s = 0$ and the corresponding subgraphs have any number of connected components of which at least one has $v \geqq 3$ and $e_i > 2$ and none is of the form $i1j$;

$\sum^{*2}$ contains terms (when $s = 0$) for which there is at least 1 ($k \geqq 1$) connected subgraph of the form $i1j$;

$\sum^{*3}$ is over all terms in μ_r for which $s \neq 0$ and

$\sum^{*4}$ is over all terms whose corresponding subgraphs are either $i1^sj$ with $2 \leqq s$ or $i1^{s_1}j1^{s_2}k$, $s_1, s_2 \geqq 1$.

4. The normal limit. The numerical values of the graph symmetric functions evidently depend on the degree of connectedness of G_1 and G_2 and these numerical values, can largely be expressed in terms of the local and average degrees of these graphs.

THEOREM 4.1. *If $n^{\frac{1}{2}} < \bar{p}_1\bar{p}_2 \to \infty$ but $\{\bar{p}_1\bar{p}_2/n\}^r \to 0$ for $r > 2$ as $n \to \infty$ the standardised form of X, that is $Z = \{X - \epsilon(X)\}/(\mathrm{var}\,(X))^{\frac{1}{2}}$, has a unit normal distribution.*

First we prove two Lemmas

LEMMA 1. *The contribution to μ_r from a term corresponding to a subgraph with k connected components, increase with k.*

PROOF. Now $1 \leqq k \leqq 2r$. When $k = 2r$, each of the connected components in the subgraph has the form i1j and

$$\begin{aligned}[G(v, r)] &= [i1j]^r + k_1[i1j]^{r-2}[i1j1k] + k_2[i1j]^{r-1}[i1^2j] + \cdots \\ &= \{\tfrac{1}{2}(n\bar{p})\}^r + k_1(n\bar{p}/2)^{r-2}n/2\textstyle\sum \bar{p}_i{}^{(2)} + \cdots \\ &= (n\bar{p}/2)^r + O(n^{r-1}\bar{p}^r) \quad k_1, k_2 \text{ are finite constants.}\end{aligned}$$

When $k = 2r - 1$ that is one of the subgraphs has the form $i1j1k$ while each of the others have the from $i1j$

$$[G(v, r)] = [i1j]^{r-2}[i1j1k] + k_3[i1j]^{r-3}[i1j1^2k] + k_4[i1j]^{r-4}[i1j1k]^2 + \cdots$$
$$= (n\bar{p}/2)^{r-2}n\sum \rho_i^{(2)}/2 + (n\bar{\rho}/2)^{r-3}n\sum \rho_i^{(2)} + \cdots$$
$$\cong (n\bar{p}/2)^{r-1} + O(n\bar{\rho})^{r-2}.$$

It can similarly be shown that as k decreases so does the contribution to $\mu[G(v, r)]$, when $G(v, r)$ has k connected components.

Lemma 2. $S_{2r} = \sum_{s=0}^{2r} \binom{2r}{s}(-1)^s(n^{(2)})^{-s}(n^{(4r-2s)})^{-1} = O(n^{-5r})$.

Proof.

$$S_{2r} = (n^{(2)})^{-2r}n!\sum_{t=0}^{2r} \binom{2r}{t}(-1)^{2r-t}(n - 2t)!(n^{(2)})^t$$
$$= (n^{(2)})^{-2r}n!\int_0^\infty \sum_{t=0}^{2r} \binom{2r}{t}x^{n-2t}(n^{(2)})^t(-1)^{2r-t}e^{-x}\, dx$$
$$= (n^{(2)})^{-2r}n!\int_0^\infty x^{n-4r}(x^2 - n^{(2)})^{2r}e^{-x}\, dx$$
$$= (n^{(2)})^{-2r}n!(2nn^{\frac{1}{2}})^{2r}e^{-x}n^{-4r}n^{\frac{1}{2}}\int_{-n^{\frac{1}{2}}}^\infty (u + O(n^{-\frac{1}{2}}))^{2r} \exp\,[-\tfrac{1}{2}u^2 + O(n^{-\frac{1}{2}})]\, du$$
$$\cong \{n^{-5r}\int_{-\infty}^\infty u^{2r}e^{-u^2/2}/(2\pi)^{\frac{1}{2}}\}\ (1 + O(n^{\frac{1}{2}}))\, du \quad \text{as} \quad n \to \infty$$
$$= n^{-5r}(2r - 1)(2r - 3) \cdots 3.1$$
$$= O(n^{-5r}) \quad \text{for } r \text{ finite as} \quad n \to \infty.$$

Proof of Theorem 4.1. We shall prove Theorem 4.1 by considering the sums $\sum^{*1}$, $\sum^{*2} + \sum^{*3}$, and $\sum^{*4}$ separately in the limit as $n \to \infty$, in the expression for μ_{2r} and μ_{2r+1} writing $\mu_{2r}(z) = \sum_{2r}^{*1} + \sum_{2r}^{*2} + \sum_{2r}^{*3} + \sum_{2r}^{*4}$.

$\sum^{*1}$: The subgraph in this sum for μ_{2r}, written $\sum_{\mu_{2r}}^{*1}$, with the largest number of connected components has $r - 2$, all except one of which have the form $i1j1k$. This remaining one may have the form $i1^2j1k1l$, $i1j1^2k1l$, $i1^2ji1ki1l$, $i1^2j1k1i$, $i1j1k1lk1m$, $i1ji1ki1li1m$, $i1j1k1l1i$, $i1j1k1l1i$ or $i1j1k1l1m$. The contribution from this subgraph with the largest number of connected components to $\mu_{2r}(z)$ is given by

$$(2r)!(n^2\bar{\rho}_1^2\bar{\rho}_2^2)^{r-2}n^2\bar{\rho}_1^{c-1}\bar{\rho}_2^{c-1}/kn^{(3r-6+c)}(r - 2)!\mu_2^r(x);$$

where k is finite and c is the number of vertices in the only connected component of the subgraph which does not have the form $i1j1k$. This sum is of the order of $\bar{\rho}_1^{2r+c-s}\bar{\rho}_2^{2r+c-s}/n^{r+c-s}\mu_2^r(x)$. There are a finite number of such subgraphs, so that

$$\sum_{2r}^{*1} = O((\bar{\rho}_1\bar{\rho}_2)^{2r+c-s}/n^{r+c-2}\mu_2^r(x), \qquad 3 \leqq c \leqq 5;$$

and
$$\mu_2(x) = 2[i1^2j]_1[i1^2j]_2/n^{(2)} + 4[i1j1k]_1[i1j1k)_2/n^{(3)}$$
$$+ 4[i1j, k1l]_1[i1j, k1l]_2/n^{(4)} + 4[i1j]_1^2[i1j]_2^2/(n^{(2)})^2$$
$$= O(n^2\bar{\rho}_1^2\bar{\rho}_2^2/n^{(2)} = O((\bar{\rho}_1\bar{\rho}_2)^2).$$

It follows then that

$$\sum\nolimits_{2r}^{*1} \leqq (\bar{\rho}_1\bar{\rho}_2)^r/n^{r+3} \tag{4.2}$$

which converges to 0 by the condition of the theorem. By Lemma 1, contributions from other subgraphs also tend to zero.

$\sum_{2r}^{*2} + \sum_{2r}^{*3}$: The leading term in this sum for $\mu_{2r}(z)$ may be expressed as

$$\begin{aligned} S'_{2r} &= 2^{2r} \sum\nolimits_{s=0}^{2r} \tbinom{2r}{s}(-1)^s[i1j]_1^r[i1j]_2^r(n^{(2)})^{-s}(n^{(4r-2s)})^{-1}\mu_2^{-r}(x) \\ &= 2^{2r}[i1j]_1^{2r}[i2j]_2^{2r}S_{2r}\mu_2^{-r}(x) \\ &= O((16n^{-1})^r) \qquad \text{by Lemma 2} \end{aligned}$$

and other subgraphs contribute terms less than this by Lemma 1.

$$\sum\nolimits_{2r}^{*2} + \sum\nolimits_{2r}^{*3} = O\{(n^{-1}\bar{\rho}_1\bar{\rho}_2)^r\} \to 0 \tag{4.3}$$

since we have a finite number of such subgraphs.

$\sum_{2r}^{*4}$: The graph symmetric functions in $\sum_{2r}^{*4}$ have $v = 2$, $e \geqq 2$ or $v = 3$, $e \geqq 2$ (with multiplicities). The subgraph with the largest number of connected components has j of the form $i1^2j$ and $r - j$ of the form $i1j1k$, $0 \leqq j \leqq r$ and the contribution from this to μ_{2r} is given by

$$\sum\nolimits_{2r}^{**4} = \sum\nolimits_{j=0}^{r} k_j[i1^2j]_1[i1^2j]_2^2[i1j1k]_1^{r-j}[i1jk]_2^{r-j}\cdot\mu_2^{-r}(x),$$

$$\begin{aligned} k_j = \varphi(jG(2, e = 2) + (r - j)G(3, 2))/\{\nu(jG(2, 2) \\ + (r - j)G(3, 2))A^2(jG(3, 2) + (r - j)G(3, 2))\tbinom{n}{3r-j}\} \\ = (2r)!\,2^{r-j}/(r - j)!\,n^{(3r-j)}. \end{aligned}$$

$$\begin{aligned} \sum\nolimits_{2r}^{*4} &= \{(2r)!/2^r r!\} \sum\nolimits_{s=0}^{r} {}_rc_s 2^s[i1^2j]_1^s[i1^2j]_2^s 4^{r-s}[i1j1k]_1^{r-s}[i1j1]_2^{r-s}\cdot\mu_2^{-r}(x) \\ &\quad + O(\mu_2^{-1}(x)) \\ &= (2r)!/2^r r!(2[i1^2j]_1[i1^2j]_2/n^{(2)}4[i1j1k]_1[i1j1k]_2)^r\mu_2^r(x)n^{(3)} \\ &= (2r)!/2^r r! + O(\mu_2^{-1})(x) \end{aligned} \tag{4.4}$$

which tends to $(2r)!/2^r r!$ as n tends to ∞. For odd values of r consider μ_{2r+1} :

$$\sum\nolimits_{\mu_{2r+1}}^{*1} = O\{(\bar{\rho}_1\bar{\rho}_2)^{r+\frac{1}{2}}/n^{r+1}\} \tag{4.5}$$

and

$$\sum\nolimits_{\mu_{2r+1}}^{*2} + \sum\nolimits_{\mu_{2r+1}}^{*3} = O\{(\bar{\rho}_1\bar{\rho}_2)^{r+\frac{1}{2}}/n^{r+\frac{1}{2}}\} \tag{4.6}$$

can be proved as in corresponding sums in μ_{2r} . $\sum^{*4}$ consists of terms corresponding to subgraphs of the form $i1^rj$ and $i1^{r_1}j1^{r_2}k$; $r \geqq 2r_1$, $r_2 \geqq 1$. The leading term which is the one having the largest number of connected subgraphs, have r, all except one of which have the form $i1^2j$ or $i1j1k$, the one having the form $i1^{r_1}j$ of $i1^{r_2}j1^{r_3}k$, $r_3 \geqq 1$. Its contribution to μ_{2r+1} is given therefore to be of the order of $\mu_2^r(x)/\mu_2^{r+\frac{1}{2}}(x)$ which tends to zero as $n \longrightarrow \infty$. That is

$$\textstyle\sum^{*4}_{\mu_{2r+1}} = O(\mu_2^{-\frac{1}{2}}). \tag{4.7}$$

From equations (4.2) to (4.4) we have

$$\lim_{n\to\infty} \mu_{2r}(x)/\{\mu_2(x)\}^r = (2r)!/2^r\cdot r! \tag{4.8}$$

while from equations (4.5) to (4.7) we have

$$\lim_{n\to\infty} \mu_{2r+1}(x)/\{\mu_2(x)\}^{r+\frac{1}{2}} = 0. \tag{4.9}$$

It follows from equations (4.8) and (4.9) using the Frechet-Shohat limit theorem that $Z = \{X - \epsilon(X)\}/(\text{var }(X))^{\frac{1}{2}}$ has an asymptotic unit normal distribution.

5. Discussion. The conceptual model of a graph whose number of vertices tends to infinity, needs to be evaluated in relation to the particular statistical application. In epidemiological applications, the graphs G_1 and G_2 have the same set of vertices V corresponding to a set of patients suffering from a particular disease; the coordinates of whose domicile and the time of onset of disease are recorded; but their sets of edges E_1 and E_2 are not necessarily the same. If two cases are 'adjacent' in space the corresponding vertices are joined in G_1 and if they are 'adjacent' in time they are joined in G_2. Adjacency is defined as being less than a distance d km apart in space or separated by an interval less than t days in time. In the case of Knox's X for patients with Leukaemia in Northumberland and Durham (1964) [2] if G_1 denotes the space graph, G_1 is then essentially a population map of Northumberland and Durham with locations of the different cases labelled 1, 2, 3, $\cdots$, n in the order in which they occurred. A single edge is drawn between every adjacent pair in space if they are less than 1 km apart. Defining adjacency in time as being less than 60 days apart G_2 is then the time graph and it is 'linear.' The choice of the critical values d and t ($d = 1$ and $t = 60$ in Knox's case) is arbitrary, though they have to behave so that the limiting conditions in Theorem 4.1 hold. For example, if we consider $n \longrightarrow \infty$ due to sampling being extended over an indefinite period of time, but the region from which patients are taken remains fixed, d has to decrease to enable conditions in Theorem 4.1 to hold; t may remain fixed. Conversely, if n increases due to inclusion of wider geographical area, but cases are all drawn from the same period, then t has to decrease with d held fixed.

For values of $\bar{\rho}_1\bar{\rho}_2 < n^{\frac{1}{2}}$, X has an asymptotic Poisson distribution (David and Barton (1965)[1]) and since the variance does not increase with n, there is no valid ground, under the conditions they gave, for the application of the central limit theorem for a Poisson variable.

The choice of d and t, to give the most powerful test ought to have regard to the nature of the disease (particularly the length of its incubation period and its method of spreading) and the density of the population exposed to the disease; this has been discussed by David and Barton (1966) [5].

REFERENCES

[1] Barton, D. E. and David, F. N. (1965). The random intersection of two graphs. *Research Papers in Statistics, Festschrift for Neyman.* Wiley, New York.

[2] Knox, G. (1964). Epidemiology of childhood leukaemia in Northumberland and Durham. *Brit. J. Prev. Soc. Med.* **18** 17–24.

[3] Knox, G. (1963). Detection of low intensity epidemicity. *Brit. J. Prev. Soc. Med.* **17** 121–27.

[4] Doll, R. (1965). Leukaemia. The Epidemiological picture. *Symposium on Current Research in Leukaemia.* Cambridge University Press.

[5] David, F. N. and Barton, D. E. (1966). Two space-time tests for epidemicity. *Brit. J. Prev. Soc. Med.* **20** 44–48.