Math. Proc. Camb. Phil. Soc. (1979), **86**, 313
With 4 text figures
Printed in Great Britain

Multi-indexing and multiple clustering

By Z. A. MELZAK
University of British Columbia, Vancouver

(*Received* 25 *August* 1978)

1. *Simple linear clustering.* One of the simplest clustering problems is the following: if n points are taken independently and uniformly at random on the interval $I = [0, L]$, what is the probability $Q = Q(n, a, L)$ that no two points are closer than a? The well-known answer

$$Q(n, a, L) = [1-(n-1)\,a/L]^n \tag{1}$$

if $(n-1)\,a \leqslant L$, and $Q = 0$ otherwise, can be obtained in a variety of ways. Perhaps the simplest one is the original method of E. C. Molina (5). If $x_1, \ldots, x_n$ are the abscissas of the n points, then there are $n!$ equiprobable orderings of the x_i's consistent with the condition that no two points are closer than a. The probability of each one is

$$L^{-n}\,\mathrm{Prob}\{0 \leqslant x_1 \leqslant x_2 - a \leqslant x_3 - 2a \leqslant \ldots \leqslant x_n - (n-1)\,a \leqslant L - (n-1)\,a\}.$$

Let $y_i = x_i - (i-1)\,a$ for $i = 1, \ldots, n$, then the above becomes

$$L^{-n}\,\mathrm{Prob}\{0 \leqslant y_1 \leqslant y_2 \leqslant \ldots \leqslant y_n \leqslant L - (n-1)\,a\},$$

which is

$$L^{-n}\int_0^{L-(n-1)a}\int_0^{y_n}\ldots\int_0^{y_2} dy_1 \ldots dy_{n-1}\,dy_n.$$

Performing the integrations and multiplying by the number $n!$ of orderings, we get (1).

Similar simple reasoning yields a slight generalization of (1). Let n points labelled $1, \ldots, n$ be taken as before independently and uniformly at random on $[0, L]$. Let $A = (a_{ij})$ be a symmetric $n \times n$ matrix of real non-negative numbers such that $a_{ii} = 0$ and the triangle condition $a_{ij} + a_{jk} \geqslant a_{ik}$ holds for all i, j, k. Then the probability $Q(n, A, L)$ that the ith and the jth point are no closer than a_{ij} for $1 \leqslant i < j \leqslant n$ is

$$Q(n, A, L) = \frac{1}{n!}\sum_{\sigma}\left[\max\left(0, L - \sum_{k=1}^{n-1} a_{\sigma(k)\,\sigma(k+1)}\right)\Big/ L\right]^n.$$

The first summation extends over the $n!$ permutations of the symmetric group. To prove this, it suffices to consider the contribution arising from a particular ordering, say $0 \leqslant x_1 \leqslant \ldots \leqslant x_n$, and then to sum over the permutations. For this ordering we must have

$$0 \leqslant x_1 \leqslant x_2 - A_2 \leqslant \ldots \leqslant x_n - A_n \leqslant L - A_n,$$

where

$$A_k = \max\,(a_{i_1 i_2} + a_{i_2 i_3} + \ldots + a_{i_{s-1} i_s}),$$

the maximum being taken over all sets $i_1, \ldots, i_s$ of integers, satisfying

$$1 \leqslant i_1 < i_2 < \ldots < i_s \leqslant k.$$

0305-0041/79/0000-6540 $03.50 © 1979 Cambridge Philosophical Society

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

The triangle condition implies at once that

$$A_k = \sum_{i=1}^{k-1} a_{ii+1}$$

and now $Q(n, A, L)$ is obtained as (1) before.

Formula (1) gives us the probability of non-occurrence of a simple clustering for a uniform one-dimensional distribution on a segment. Here 'simple' refers to pairs of points: a simple cluster occurs if some two points are too close. The situation is quite different if the segment I is replaced by a region R in two or more dimensions, or if simple clustering is replaced by multiple clustering, or if the distribution of the n points is non-uniform: no closed-form explicit solution is known then. This simultaneous triple generalization is the problem attacked in the present paper. The methods used are a development and a correction of those initiated in (3) and (4). We start by describing our problem in more detail.

2. *Multiple clustering.* Let R be a sufficiently regular region in the d-dimensional Euclidean space and let n points be taken in R, independently and identically distributed according to a sufficiently well-behaved density f on R. In what follows the points will be often identified with the corresponding vectors from the origin. An integer k, $2 \leqslant k \leqslant n$, and a positive number a are given. We wish now to set up conditions which guarantee that some k of our n points form a spherical shape k-tuple cluster of size $\leqslant a$. The following four conditions are equivalent:

(a) a ball of diameter a contains more than $k-1$ points,
(b) the smallest ball covering k points has diameter $\leqslant a$,
(c) of the n balls of diameter a about the n points as centres, some k have non-empty intersection,
(d) among the n points there are some k, of which no two are further apart than $((m+1)/2m)^{\frac{1}{2}} a$, where $m = \min(d, k-1)$.

The last condition hinges on the theorem of Jung (2): the smallest sphere containing a set of diameter D in the Euclidean d-space has radius $\leqslant (d/(2d+2))^{\frac{1}{2}} D$. Another way to define k-tuple clusters would be by means of the order statistics for the successive nearest neighbours. Here one has the following condition for a k-cluster of size $\leqslant a$:

(e) among the n points $p_1, \ldots, p_n$ the kth successive nearest neighbour to some point p_i, starting with that p_i as its own first nearest neighbour, is at the distance $\leqslant b_k a$ from p_i, where $b_2 = 1$ and $b_k = \frac{1}{2}$ for $k > 2$.

However, (e) is not equivalent to the previous conditions (a)–(d). Spherical shape clusters enter in all cases (a)–(e) but in (a)–(d) the centre of the cluster is arbitrary whereas in (e) the cluster centres on one of the n points if $k > 2$ and midway between two such points if $k = 2$.

We agree now that a k-cluster is defined by any one of the conditions (a)–(d) and our problem is to find the probability $Q = Q(n, k, d, a, R, f)$ that no k-cluster occurs. When $k = 2$, $d = 1$, $R = I$ and f is uniform on I, then Q becomes the probability $Q(n, a, L)$ of (1).

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

The procedure will depend on a suitable geometrical dissection of the sample space of all configurations, and the desired probability Q will be obtained by applying the inclusion–exclusion principle, (6), not in its standard form but extended to multi-indexed sets. We shall thus obtain Q as a finite series

$$Q = 1 - K_1 + K_2 - \ldots + (-1)^N K_N \quad \left(N = \binom{n}{k}\right), \tag{2}$$

where the terms K_r are finite sums of certain basic multiple integrals which will be called clusters. Both the number of clusters in K_r and their complexity grow rapidly as r increases. Hence (2) might appear to be a formal rather than an actual solution. However, when a is small enough the first few terms in (2) give a good approximation to Q because the series (2) is then alternating, and rapidly decreasing, so that the truncation error can be estimated.

Especial interest attaches to the asymptotic case when n is large and a is small while $n^{k/(k-1)d}a = u$ stays constant. As $n \to \infty$ it will turn out that each term K_r contributes in the limit only one cluster, and we prove that

$$\lim_{n\to\infty} Q(n, k, d, un^{-k/(k-1)d}, R, f) = \exp\left[-c(2^{1-k}k)^d u^{d(k-1)}\right], \tag{3}$$

where

$$c = \frac{1}{k!}(\pi^{\frac{1}{2}d}/\Gamma(1+d/2))^{k-1}\int_R f^k(x)\,dx. \tag{4}$$

This generalizes a result of Efron (1) valid for simple clusters only, that is for $k = 2$. Finally, for comparison with the asymptotics as well as for their own interest, some calculations will be made relative to clustering in the unit square.

3. *A special case.* Our rather lengthy method may be introduced by deriving with it (1) again. The n points have abscissas $x_1, \ldots, x_n$ and the sample space of n-tuples $(x_1, \ldots, x_n)$ is the n-cube

$$H = \{(x_1, \ldots, x_n)\colon 0 \leqslant x_m \leqslant L, m = 1, \ldots, n\}.$$

For $1 \leqslant i < j \leqslant n$, let $H(i,j)$ be the subset of H given by

$$H(i,j) = \{(x_1, \ldots, x_n)\colon 0 \leqslant x_m \leqslant L, m = 1, \ldots, n, |x_i - x_j| \leqslant a\}.$$

$H(i,j)$ is that part of the sample space H for which the ith and the jth points are in the forbidden configuration, that is, too close. On account of the independence and the uniform distribution

$$Q(n, a, L) = \mathrm{Vol}\left(H - \bigcup_{1\leqslant i<j\leqslant n} H(i,j)\right)/\mathrm{Vol}\,H. \tag{5}$$

When the $\binom{n}{2}$ sets $H(i,j)$ are removed from the cube H the remainder of H consists of certain $n!$ simplexes corresponding to the $n!$ orderings of the abscissas $x_1, \ldots, x_n$. As it happens, those $n!$ simplexes can be translated so as to fit together and to make up, exactly and without overlap, a smaller n-cube of side $L-(n-1)a$. The required translations are simply described: each simplex is moved so that its vertex closest to the origin is shifted to the origin itself. Therefore, as (5) states, Q is the ratio of the volumes of the two n-cubes, which is (1) again.

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

Such convenient fitting together of the sample space of allowable configurations does not appear to hold for any case other than that of simple linear clustering. It may therefore be expected that the general probability Q of no k-tuple clustering will be something considerably more complicated than (1). However, even without knowing of the lucky fluke of the exact geometrical fit into a smaller n-cube, we could still start with the obvious formula (5). It is necessary then to obtain either of the quantities

$$W = \mathrm{Vol}\Big(\bigcup_{1\leqslant i<j\leqslant n} H(i,j)\Big), \quad V = \mathrm{Vol}\Big(H - \bigcup_{1\leqslant i<j\leqslant n} H(i,j)\Big). \tag{6}$$

These are respectively the contents of the set of forbidden and of allowed configurations. If the units are normalized so that $\mathrm{Vol}\, H = 1$, then W and V are the complementary probabilities of occurrence and of non-occurrence of simple clustering.

4. *Multi-indexing for pairs.* An obvious way to obtain W in (6) would be to re-enumerate the $N = \binom{n}{2}$ sets $H(i,j)$ with a single index as $H_1, \ldots, H_N$. Then an application of the inclusion–exclusion principle would give us

$$W = \sum_{1\leqslant i\leqslant N} \mathrm{Vol}\, H_i - \sum\sum_{1\leqslant i<j\leqslant N} \mathrm{Vol}\,(H_i \cap H_j) + \ldots + (-1)^N \mathrm{Vol}\,(H_1 \cap \ldots \cap H_N).$$

However, such disastrous merging of pairs of indices into the anonymity of a single index destroys the useful pair-structure which, on the contrary, must be preserved and used. But we still want to use the powerful counting method represented by the inclusion–exclusion. Now, the ordinary inclusion–exclusion is symbolized conveniently by

$$\prod_{i=1}^{n}(1-X_i) = 1 - \sum_{1\leqslant i\leqslant n} X_i + \sum\sum_{1\leqslant i<j\leqslant n} X_i X_j - \ldots. \tag{7}$$

We replace it with its multi-indexed (here: pair-indexed) variant:

$$\prod_{1\leqslant i<j\leqslant n}\prod [1-X(i,j)] = 1 - \Sigma\, X(i_1,j_1) + \Sigma\Sigma\, X(i_1,j_1)\, X(i_2,j_2) - \ldots. \tag{8}$$

This will be written as

$$\prod_{1\leqslant i<j\leqslant n}\prod [1-X(i,j)] = 1 - \Sigma_1 + \Sigma_2 - \ldots + (-1)^N \Sigma_N, N = \binom{n}{2}, \tag{9}$$

where

$$\Sigma_r = \sum_{i_1,j_1} \cdots \sum_{i_r,j_r} X(i_1,j_1) \ldots X(i_r,j_r) \tag{10}$$

and the summation in (10) extends over all the distinct $\binom{N}{r}$ r-tuples of pairs of indices.

The equations (9) and (10) will be referred to as inclusion–exclusion for pair-indexing. By assigning meaning to the symbols $X(i,j)$ we now connect (9) with the formula (6) for V.

There are n points $p_1, \ldots, p_n$ on the interval I or in the region R and the indices i and j refer to them. $X(i,j)$ is the elementary coincidence operator referring to those configurations of the n-tuple $(p_1, \ldots, p_n)$ in which the ith and the jth point are within distance a. The complementary operator $1 - X(i,j)$ represents elementary exclusion. Multiplication in the L.H.S. of (9) is in the logical or Boolean sense as the 'and' operation and the whole product synthesizes the desired exclusion: no two points are closer

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

than a. We note that the result of applying $X(i,j)$ to H is $H(i,j)$; we recall also that the units are normalized so as to make Vol $H = 1$. Therefore, applying the formal identity (9) to H we get the formula for V in (6):

$$\mathrm{Vol}\,(H - \bigcup_{1 \leqslant i < j \leqslant n} H(i,j)) = 1 - \Sigma_1 + \Sigma_2 - \ldots + (-1)^N \Sigma_N, \tag{11}$$

where

$$\Sigma_r = \sum_{i_1, j_1} \ldots \sum_{i_r, j_r} \mathrm{Vol}\,[H(i_1,j_1) \cap \ldots \cap H(i_r,j_r)]; \tag{12}$$

in the above the summation extends over the $\binom{N}{r}$ distinct r-tuples of pairs.

5. *Simple clustering in R.* As an application of the preceding, we obtain Q in the special case when $k = 2$, R is arbitrary, and f is uniform over R. As usual, we normalize units so that R has content 1; f is therefore the characteristic function $\chi(R)$ of R. Hence $Q = Q(n, 2, d, a, R, \chi(R))$ is the probability that, when n points are taken independently and uniformly at random in the d-dimensional region R, then no two are closer than a. The sample space is now the Cartesian product

$$H = \{(p_1, \ldots, p_n) : p_m \in R, m = 1, \ldots, n\}$$

and $H(i,j)$ has the same meaning as before:

$$H(i,j) = \{(p_1, \ldots, p_n) : p_m \in R, m = 1, \ldots, n, |p_i - p_j| \leqslant a\}. \tag{13}$$

Therefore Q is given by (11) and (12). This is, at best, only a formal solution of our problem. However, we shall introduce the device of indexing by incidence graphs, which will enable us to control the complicated summations in (12). In addition to this simple though important book-keeping function, graph-indexing will enable us to distinguish between terms of different orders of magnitude. When the various clusters contributing to Σ_r are considered, it will turn out that, the more disconnected the graph of a cluster, the bigger its contribution to Σ_r. This will lead eventually to an explicit asymptotic form of the probability Q.

6. *Incidence graphs and cluster integrals.* The probability $Q = Q(n, 2, d, a, R, \chi(R))$ is

$$Q = 1 - \Sigma_1 + \Sigma_2 - \ldots + (-1)^N \Sigma_N, \quad N = \binom{n}{2}, \tag{14}$$

where Σ_r is given by (12) and (13) as the sum of the volumes of the intersections of the sets $H(i,j)$, taken r at a time. Let us examine the first few terms in (14). The term Σ_1 is just the sum of the volumes

$$\Sigma_1 = \sum_{1 \leqslant i < j \leqslant n} \mathrm{Vol}\, H(i,j).$$

The $N = \binom{n}{2}$ sets $H(i,j)$ are all congruent so that $\Sigma_1 = N \,\mathrm{Vol}\, H(1,2)$. Also, each $H(i,j)$ is a cylinder set: a Cartesian product of a subset of $R_i \times R_j$ with all the remaining R_m's. This leads to the first cluster integral

$$\mathrm{Vol}\, H(i,j) = \mathrm{Vol}\, H(1,2) = I_{11} = \int\int_{\substack{p_1 \in R, p_2 \in R \\ |p_1 - p_2| \leqslant a}} dp_1 dp_2. \tag{15}$$

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

For uniformity of notation with other terms, which are yet to come, we write $N_{11} = N$ so that

$$\Sigma_1 = N_{11} I_{11}, \tag{16}$$

where I_{11} is given in (15). The next term Σ_2 is the sum of volumes of the intersections of all pairs of sets $H(i,j)$. While such intersections are many, their types are only two:

N_{21} congruent sets of the first type, $H(1,2) \cap H(3,4)$, with all indices distinct, and

N_{22} congruent sets of the second type, $H(1,2) \cap H(1,3)$, with one index shared.

Simple combinatorial calculation shows that

$$N_{21} = n(n-1)(n-2)(n-3)/8, \quad N_{22} = n(n-1)(n-2)/2. \tag{17}$$

In either expression the numerators are obvious; the denominators represent factoring out of symmetries, necessary for correct counting. It is verified that

$$N_{21} + N_{22} = \left(\binom{n}{2} \right)$$
$$2$$

as it should be, since all the pairs of pairs of indices are either of type 1 or of type 2. As to the volumes of the two types of intersections, it follows from the cylindrical structure of $H(i,j)$ that

$$I_{21} = \mathrm{Vol}\,(H(1,2) \cap H(3,4)) = I_{11}^2. \tag{18}$$

The second type of intersection leads to a new cluster integral

$$I_{22} = \mathrm{Vol}\,(H(1,2) \cap H(1,3)).$$

In terms analogous to (15), we have

$$I_{22} = \int\int\int_{\substack{p_i \in R \\ |p_3 - p_1| \leqslant a,\ |p_3 - p_1| \leqslant a}} dp_1\, dp_2\, dp_3. \tag{19}$$

Therefore Σ_2 may be written as

$$\Sigma_2 = N_{21} I_{21} + N_{22} I_{22}. \tag{20}$$

The pattern of formation of the terms Σ_r is now beginning to emerge. To write down the next term Σ_3 we begin, as in (16) or (20), with

$$\Sigma_3 = \sum_{i=1}^{5} N_{3i} I_{3i}. \tag{21}$$

Here we have five terms rather than one as in (16) or two as in (20) because there are exactly five types of index sharing among triples of pairs. These are given in Fig. 1 together with the corresponding incidence graphs. Let N_{3i} be the number of intersections of type i. Simple combinatorial computation, together with correct symmetry removal, shows that

$$N_{31} = (n)_6/48, \quad N_{32} = (n)_5/4, \quad N_{33} = (n)_4/2, \quad N_{34} = (n)_4/6, \quad N_{35} = (n)_3/6, \tag{22}$$

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

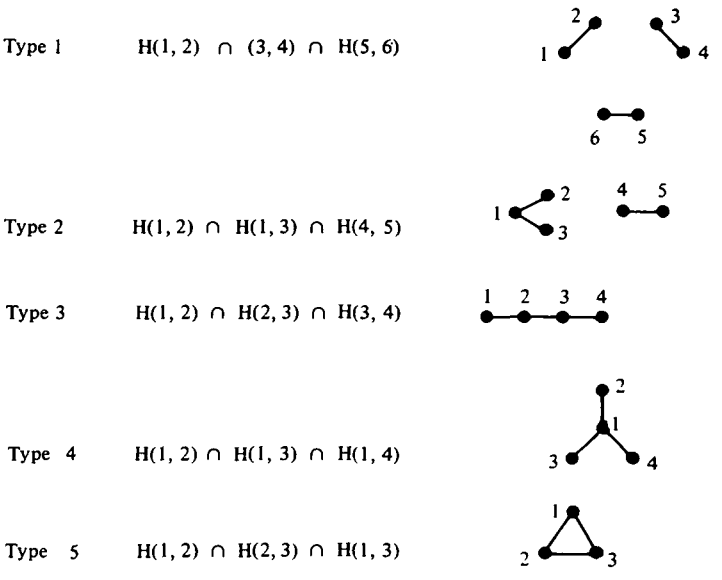


Fig. 1. Incidence graphs for $r = 3$.

where $(n)_s = n(n-1)(n-2)\dots(n-s+1)$. Again it is verified that

$$\sum_{i=1}^{5} N_{3i} = \left(\binom{n}{2}\right)$$

since the five types comprise all possible triples of pairs. The first two intersection-volumes, that is, the cluster integrals I_{31} and I_{32}, correspond to disconnected incidence graphs. Therefore they split up into products of previous ones on account of cylindricity:

$$I_{31} = I_{11}^3, \quad I_{32} = I_{11} I_{22}. \tag{23}$$

The other three correspond to connected incidence graphs and so they lead to new clusters:

$$I_{33} = \int\int\int\int_{\substack{p_i \in R,\ |p_1-p_3|\leqslant a \\ |p_2-p_3|\leqslant a,\ |p_3-p_4|\leqslant a}} dp_1 dp_2 dp_3 dp_4, \quad I_{34} = \int\int\int\int_{p_i \in R,\ |p_1-p_j|\leqslant a} dp_1 dp_2 dp_3 dp_4,$$

$$I_{35} = \int\int\int_{p_i \in R,\ |p_i-p_j|\leqslant a} dp_1 dp_2 dp_3. \tag{24}$$

Consider now the general term Σ_r. Let there be $t(r)$ distinct types of index sharing among the r-tuples of pairs, and let the corresponding incidence graphs be

$$G_{r1}, G_{r2}, \dots, G_{rt(r)}.$$

Here G_{ri} is an unlabelled graph with the following properties:

(*a*) no vertex is isolated,

(*b*) two vertices are connected by at most one edge,

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

(*c*) no edge connects a vertex to itself,

(*d*) there are exactly r edges,

(*e*) the graphs G_{ri}, $i = 1, \ldots, t(r)$, are all the distinct graphs with the properties (*a*)–(*d*).

Let $v = v(G)$ be the number of vertices of the graph G. For the graph G_{ri} let $N_{ri}(n)$ be the number of distinct ways of labelling the $v(G_{ri})$ vertices of G_{ri} with the integers $1, \ldots, n$. That is, $N_{ri}(n)$ is the number of distinct r-tuples of pairs corresponding to the type i of index sharing:

$$N_{ri}(n) = (n)_{v(G_{ri})}/D_{ri}, \tag{25}$$

where D_{ri} is the number of symmetries of G_{ri}. Finally, together with G_{ri} goes the cluster integral I_{ri}, which is the volume of that intersection of the r-tuple of sets $H(i,j)$ of the required type:

$$I_{ri} = \int \cdots \int_{L(G_{ri})} dp_1 \ldots dp_{v(G_{ri})}. \tag{26}$$

Here the specification $L(G_{ri})$ describes the domain of integration as follows. Let the $v = v(G_{ri})$ vertices of G_{ri} be labelled with integers $1, \ldots, v$ in any way whatever. $L(G_{ri})$ consists of the v relations $p_1 \in R, \ldots, p_v \in R$ which state that the points (or vectors) $p_1, \ldots, p_v$ vary over R, and of the r distance-constraints $|p_i - p_j| \leqslant a$ for every pair i, j such that v_i and v_j are joined in G_{ri} by an edge. The cylindricity of the sets $H(i,j)$ shows that if G_{ri} is disconnected with m components then I_{ri} splits into a product of m cluster integrals belonging to lower values of r. We have therefore the following expression for the probability Q:

$$Q = 1 + \sum_{r=1}^{N} (-1)^r \sum_{i=1}^{t(r)} N_{ri}(n)\, I_{ri} \quad \left(N = \binom{n}{2}\right). \tag{27}$$

The initial assumption that the points $p_1, \ldots, p_n$ were uniformly distributed in R may be modified with only slight changes. Let the n points be independently distributed with the density f over R. Then Q is still given by (27) except that the clusters, instead of being given by (26), now have the form

$$I_{ri} = I_{ri}(f) = \int \cdots \int_{L(G_{ri})} \prod_{j=1}^{v(G_{ri})} f(p_j)\, dp_1 \ldots dp_{v(G_{ri})}. \tag{28}$$

7. *Extension to multiple clusters.* The originally proposed problem of finding the probability $Q(n, k, d, a, R, f)$ will now be solved, at least in the same formal sense as before. It is first necessary to extend the inclusion–exclusion formulae (9) and (10) from pair indexing to indexing by k-tuples. The new pair of formulae is

$$\prod_{1 \leqslant i < j < \ldots < q \leqslant n} \cdots \prod [1 - X(i,j,\ldots,q)] = 1 - \Sigma_1 + \Sigma_2 - \ldots + (-1)^M \Sigma_M, \tag{29}$$

where $M = \binom{n}{k}$ and there are k indices $i, j, \ldots, q$, and

$$\Sigma_r = \sum_{i_1, j_1, \ldots, q_1} \cdots \sum_{i_r, j_r, \ldots, q_r} X(i_1, j_1, \ldots, q_1) \ldots X(i_r, j_r, \ldots, q_r); \tag{30}$$

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

here the summations extend over all the $\binom{M}{r}$ distinct r-tuples of k-tuples of indices.

Thereafter the procedure continues as before, with r-tuples of pairs replaced by r-tuples of k-tuples, till we reach the matter of incidence graphs. These are replaced by the entirely analogous incidence hypergraphs or, more precisely, incidence k-hypergraphs. The idea is that a k-hypergraph shows the index sharing of k-tuples just as a graph shows it for pairs.

Let S be a simplex, for instance the regular simplex, with v vertices and let G be an incidence graph with r edges and v vertices. On account of the defining properties (a)–(e) in the last section G is simply a collection of some r of the $\binom{v}{2}$ edges, that is, one-dimensional faces of S. Similarly, a k-hypergraph with r faces is merely a collection of some r of the $\binom{v}{k}$ $(k-1)$-dimensional faces of S. We use here the usual absorption rule from the elementary algebraic topology: if two $(k-1)$-dimensional faces of S share the vertices of a j-dimensional face of S then they share this whole face as well.

With $k = 3$, the faces of 3-hypergraphs are triangles; let $t(r)$ be the number of distinct ones with r faces. Then $t(1) = 1$, and $t(2) = 3$ because there are three types of index sharing for pairs of triples: 0, 1, or 2 indices are shared. The three corresponding 3-hypergraphs consist of two separate triangles for the type $H(1,2,3) \cap H(4,5,6)$, two triangles sharing a vertex for the type $H(1,2,3) \cap H(1,4,5)$, and two triangles sharing an edge for the type $H(1,2,3) \cap H(1,2,4)$. For $r = 3$, $t(3)$ is found to be 12 and the 12 3-hypergraphs with three faces are shown in Fig. 2. The multiplicities $N_{ri}(n)$ are defined for incidence hypergraphs in the same way as for incidence graphs and we find that $N_{ri}(n) = (n)_v/D_{ri}$, where v is the number of vertices and D_{ri} is the number of symmetries. For instance, for the first three 3-hypergraphs of Fig. 2 the multiplicities are

$$(n)_9/(3!)^4, \quad (n)_8/2^3 3!, \quad (n)_7/2^2 3!.$$

Of course, for the sum of all twelve multiplicities we have

$$\sum_{i=1}^{12} N_{3i}(n) = \left(\binom{n}{3}\right).$$

The cluster integrals I_{ri} are also as before, substituting k-hypergraphs and k-tuples for graphs and pairs. We define H as before to be the Cartesian product of n identical replicas of R, and we use the abbreviation $CC_k(1, \ldots, k; a)$ for: the k-tuple clustering condition holds for the points $p_1, \ldots, p_k$. In place of the sets $H(i,j)$ of elementary pair-coincidence we have now sets of elementary k-tuple coincidence: for $1 \leqslant i_1 < \ldots < i_k \leqslant n$

$$H(i_1, \ldots, i_k) = \{(x_1, \ldots, x_n): x_m \in R, m = 1, \ldots, n, CC_k(i_1, \ldots, i_k; a)\}.$$

The cluster integrals are defined as volumes of intersections of multiples of the sets $H(i_1, \ldots, i_k)$ if the density f is uniform over R, and as integrals of the type (28) otherwise. The specification $L(G_{ri})$ describes the domain of integration much as before, in (28). Let $v(G_{ri}) = v$ be the number of vertices of the k-hypergraph G_{ri} of the ith type and

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

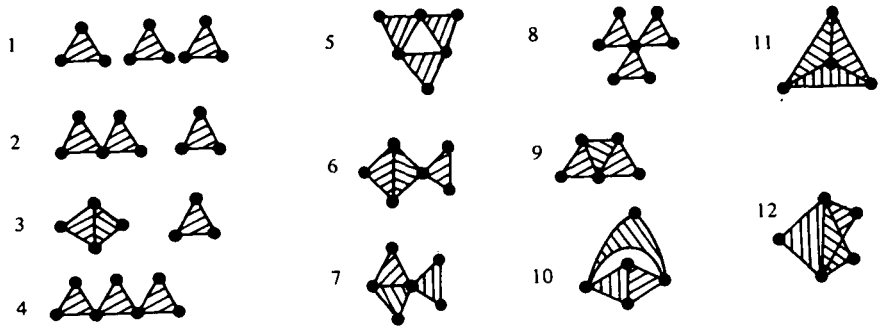


Fig. 2. Incidence 3-hypergraphs for $r = 3$.

with r faces, and let the vertices be enumerated in any way. Then $L(G_{ri})$ consists of the v relations $p_i \in R, i = 1, \ldots, v$, and of the r relations $CC_k(i_1, \ldots, i_k; a)$, one for every k-tuple $p_{i_1}, \ldots, p_{i_k}$, such that the vertices $v_{i_1}, \ldots, v_{i_k}$ are connected in G_{ri} by a $(k-1)$-dimensional face. We find now that Q is given by the formula (27) but with $N = \binom{n}{k}$ rather than with $N = \binom{n}{2}$.

8. *First two clusters for cubes.* If the formal expressions for Q are to be of any use it is necessary to evaluate at any rate the first few cluster integrals I_{ri}. We consider now the first cluster I_{11} for $k = 2$; the density f is uniform and R is the unit cube in p dimensions. To show the dependence on p and a we write $I_{11} = I_{11}(p, a)$. This is the probability that two points taken independently and uniformly at random in the unit p-cube are at the distance $\leqslant a$. As might be expected, $I_{11}(p, a)$ has different forms depending on the size of a, relative to the various diagonals of the cube, that is, on the ranges $0 \leqslant a \leqslant 1, 1 \leqslant a \leqslant 2^{\frac{1}{2}}, \ldots, (p-1)^{\frac{1}{2}} \leqslant a \leqslant p^{\frac{1}{2}}$. On the first range $0 \leqslant a \leqslant 1$ $I_{11}(p, a)$ is a polynomial in a.

A simple geometrical argument based on fitting two isosceles right-angled triangles into a square shows that

$$I_{11}(1, a) = 1 - (1-a)^2 = 2a - a^2 \tag{31}$$

for $a \leqslant 1$, and $I_{11}(1, a) = 1$ if $a \geqslant 1$. Alternatively, $I_{11}(1, a)$ is given as a double integral

$$I_{11}(1, a) = \int_0^1 \int_m^M dx_1 \, dx_2 \tag{32}$$

with $$m = \max(0, x_2 - a), \quad M = \min(1, x_2 + a).$$

For $p = 2$, we take two points (x_1, y_1), (x_2, y_2) in the unit square and $I_{22}(2, a)$ is a quadruple integral of the type (32):

$$I_{11}(2, a) = \int_0^1 \int_{m_1}^{M_1} \int_0^1 \int_{m_2}^{M_2} dx_1 \, dx_2 \, dy_1 \, dy_2 \tag{33}$$

with

$$m_1 = \max(0, y_2 - a), \quad M_1 = \min(1, y_2 + a), \quad m_2 = \max(0, x_2 - (a^2 - (y_1 - y_2)^2)^{\frac{1}{2}}),$$
$$M_2 = \min(1, x_2 + (a^2 - (y_1 - y_2)^2)^{\frac{1}{2}}).$$

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

This arrangement of limits corresponds to letting the point (x_1, y_1) move freely over the square while the other one lies within the allowed distance. There are two cases to consider depending on whether $a \leqslant 1$ or $a \geqslant 1$. If $a \leqslant 1$ we put, for brevity,

$$D = (a^2-(y_1-y_2)^2)^{\frac{1}{2}}$$

and we start on the inner double integral of (33)

$$\int_0^1 \int_m^M dx_1 dx_2, \tag{34}$$

where $m = \max(0, x_2-D)$ and $M = \min(1, x_2+D)$. In view of (32) this is $I_{11}(1, D)$. With new variables $y_1-y_2 = u, y_2 = y$

$$I_{11}(2, a) = \int_0^1 \int_m^M I_{11}(1, (a^2-u^2)^{\frac{1}{2}})\, du\, dy, \tag{35}$$

where $m = \max(-y, -a)$ and $M = \min(1-y, a)$. To evaluate this, we change the order of integrations getting

$$\begin{aligned} I_{11}(2, a) &= \int_{-a}^0 (1+u)\, I_{11}(1, (a^2-u^2)^{\frac{1}{2}})\, du + \int_0^a (1-u)\, I_{11}(1, (a^2-u^2)^{\frac{1}{2}})\, du \\ &= 2\int_0^a (1-u)\, I_{11}(1, (a^2-u^2)^{\frac{1}{2}})\, du = \pi a^2 - \tfrac{8}{3}a^3 + \tfrac{1}{2}a^4 \quad (a \leqslant 1). \end{aligned} \tag{36}$$

The same procedure is used when $a \geqslant 1$ except that the inner double integral (34) is now

$$F(u) = \begin{cases} 1 & \text{if} \quad -(a^2-1)^{\frac{1}{2}} \leqslant u \leqslant (a^2-1)^{\frac{1}{2}} \\ 2(a^2-u^2)^{\frac{1}{2}} - (a^2-u^2) & \text{otherwise.} \end{cases}$$

In place of (36) we have

$$I_{11}(2, a) = \int_{-1}^0 \int_{-u}^1 F(u)\, dy\, du + \int_0^1 \int_0^{1-u} F(u)\, dy\, du = 2\int_0^1 (1-u)\, F(u)\, du.$$

The range $[0, 1]$ is split into two parts, $[0, (a^2-1)^{\frac{1}{2}}]$ and $[(a^2-1)^{\frac{1}{2}}, 1]$ and we get

$$\begin{aligned} I_{11}(2, a) = {} & \tfrac{1}{3} - 2a^2 - \tfrac{1}{2}a^4 + \tfrac{4}{3}(a^2-1)^{\frac{1}{2}} + \tfrac{8}{3}a^2(a^2-1)^{\frac{1}{2}} \\ & + 2a^2 \arcsin((2-a^2)/a^2) \quad (1 \leqslant a \leqslant 2^{\frac{1}{2}}). \end{aligned} \tag{37}$$

Finally,
$$I_{11}(2, a) = 1 \quad (2^{\frac{1}{2}} \leqslant a). \tag{38}$$

Now (36)–(38) give us $I_{11}(2, a)$ for all values of a. It may be verified that $I_{11}(2, a)$ is continuous and continuously differentiable in a, but there are discontinuities in the second derivative at the critical values $a = 1$ and $a = 2^{\frac{1}{2}}$. The dimensional recursion represented by (35) holds generally: $I_{11}(p, a)$ is reduced to $I_{11}(p-1, a)$ and we have

$$I_{11}(p, a) = \int_0^1 \int_m^M I_{11}(p-1, (a^2-(z_1-z_2)^2)^{\frac{1}{2}})\, dz_1\, dz_2, \tag{39}$$

where $m = \max(0, z_2-a)$ and $M = \min(1, z_2+a)$. Introducing new variables $z_1-z_2 = u$ and $z_2 = y$, and interchanging the order of integrations, we obtain

$$I_{11}(p, a) = 2\int_0^M (1-u)\, I_{11}(p-1, (a^2-u^2)^{\frac{1}{2}})\, du \tag{40}$$

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

with $M = \min(1, a)$. By suitable splitting of the range of integration this allows us to obtain $I_{11}(p, a)$ for all values of a.

When $a \leqslant 1$ the procedure simplifies: we put $u = a\sin\theta$ in (40) so that

$$I_{11}(p,a) = 2a\int_0^{\frac{1}{2}\pi} (1 - a\sin\theta)\, I_{11}(p-1, a\cos\theta)\cos\theta\, d\theta. \tag{41}$$

An easy induction on p shows that $I_{11}(p, a)$ is a polynomial of the form

$$I_{11}(p,a) = \sum_{j=p}^{2p} A_{pj} a^j.$$

Let $$c_j = \int_0^{\frac{1}{2}\pi} \cos^j\theta\, d\theta, \quad A_{pj} = 0 \quad \text{if} \quad j < p \quad \text{or} \quad j > 2p,$$

then (41) leads to the recursion

$$A_{p+1\,j} = 2c_j A_{p\,j-1} - (2/j) A_{p\,j-2}.$$

This recursion holds even down to $p = 1$ provided that we define $I_{11}(0, a)$ to be 1. For the first few values of p, it is found that

$$\left.\begin{aligned}
&I_{11}(0,a) = 1, \quad I_{11}(1,a) = 2a - a^2, \quad I_{11}(2,a) = \pi a^2 - (8/3)a^3 + (1/2)a^4,\\
&I_{11}(3,a) = (4\pi/3)a^3 - (3\pi/2)a^4 + (8/5)a^5 - (1/6)a^6,\\
&I_{11}(4,a) = (\pi^2/2)a^4 - (32\pi/15)a^5 + \pi a^6 - (64/105)a^7 + (1/24)a^8,\\
&I_{11}(5,a) = (8\pi^2/15)a^5 - (5\pi^2/6)a^6 + (32\pi/21)a^7 - (\pi/12)a^8 + (32/189)a^9 - (1/120)a^{10}.
\end{aligned}\right\} \tag{42}$$

The leading coefficient is the volume of the p-dimensional unit ball and the last coefficient is $(-1)^p/p!$.

The next cluster integral splits on account of cylindricity and we have $I_{21} = I_{11}^2$. The higher clusters get rapidly more involved but we shall sketch the procedure for the cluster $I_{22}(p, a)$ when $a \leqslant \frac{1}{2}$. Roughly speaking, this limitation on a means that the incidence graph, which consists of two edges of length a meeting at a common vertex, can be 'stretched out' in full inside the unit cube (or square). Dimensional induction on p will be applied again but, as it happens, there is not enough structure in the thing for a full inductive hold with just one parameter a. Therefore an additional parameter b is brought in and the generalized cluster integral $I_{22}(p, a, b)$ is considered. Here

$$I_{22}(1,a,b) = \int_0^1 \int_{m_1}^{M_1} \int_{m_2}^{M_2} dx_3\, dx_2\, dx_1 \tag{43}$$

with

$$m_1 = \max(0, x_1 - b), \quad M_1 = \min(1, x_1 + b), \quad m_2 = \max(0, x_1 - a)$$

and

$$M_2 = \min(1, x_1 + a).$$

Generally, $I_{22}(p, a, b)$ is the probability that when three points p_1, p_2, p_3 are chosen independently and uniformly at random in the unit p-cube, then p_3 is within distance a, and p_2 within distance b, of p_1. The cluster we want is $I_{22}(p, a)$ and it is obtained as $I_{22}(p, a, a)$. First, for $p = 2$

$$I_{22}(2,a,b) = \int_{m_6}^{M_6} \cdots \int_{m_1}^{M_1} dx_3\, dx_2\, dx_1\, dy_3\, dy_2\, dy_1,$$

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

where

$$m_1 = \max(0, x_1 - (a^2-(y_3-y_1)^2)^{\frac{1}{2}}), \quad M_1 = \min(1, x_1 + (a^2-(y_3-y_1)^2)^{\frac{1}{2}}),$$
$$m_2 = \max(0, x_1 - (b^2-(y_2-y_1)^2)^{\frac{1}{2}}), \quad M_2 = \min(1, x_1 + (b^2-(y_2-y_1)^2)^{\frac{1}{2}}),$$
$$m_4 = \max(0, y_1 - a), \quad M_4 = \min(1, y_1 + a),$$
$$m_5 = \max(0, y_1 - b), \quad M_5 = \min(1, y_1 + b),$$
$$m_3 = m_6 = 0, \quad M_3 = M_6 = 1.$$

Therefore we get the start of our dimensional induction:

$$I_{22}(2,a,b) = \int_0^1 \int_{m_1}^{M_1} \int_{m_2}^{M_2} I_{22}(1, (a^2-(y_3-y_1)^2)^{\frac{1}{2}}, (b^2-(y_2-y_1)^2)^{\frac{1}{2}})\, dy_3\, dy_2\, dy_1, \tag{44}$$

where

$$m_1 = \max(0, y_1 - b), \quad M_1 = \min(1, y_1 + b), \quad m_2 = \max(0, y_1 - a)$$

and

$$M_2 = \min(1, y_1 + a).$$

Similarly, for general p,

$$I_{22}(p,a,b) = \int_0^1 \int_{m_1}^{M_1} \int_{m_2}^{M_2} I_{22}(p-1, (a^2-(z_3-z_1)^2)^{\frac{1}{2}}, (b^2-(z_2-z_1)^2)^{\frac{1}{2}})\, dz_3\, dz_2\, dz_1 \tag{45}$$

with

$$m_1 = \max(0, z_1 - b), \quad M_1 = \min(1, z_1 + b), \quad m_2 = \max(0, z_1 - a)$$

and

$$M_2 = \min(1, z_1 + a).$$

From (43) $I_{22}(1,a,b)$ is obtained as

$$I_{22}(1,a,b) = \int_0^1 [\min(1, x+b) - \max(0, x-b)][\min(1, x+a) - \max(0, x-a)]\, dx. \tag{46}$$

The range $[0,1]$ is split into five parts: $[0,m]$, $[m,M]$, $[M,1-M]$, $[1-M,1-m]$, $[1-m,1]$ where

$$m = \min(a,b), \quad M = \max(a,b). \tag{47}$$

On each of the five sub-ranges the integrand of (46) is a polynomial; carrying out the integrations, we get

$$I_{22}(1,a,b) = 4mM - \tfrac{1}{3}m^3 - m^2M - 2mM^2 \tag{48}$$

with m and M given by (47). In particular,

$$I_{22}(1,a) = I_{22}(1,a,a) = 4a^2 - \tfrac{10}{3}a^3. \tag{49}$$

To get $I_{22}(2,a)$ we use (45) with $p = 2$. New variables u, v, z are introduced with $z_3 - z_1 = u$, $z_2 - z_1 = v$, $z_1 = z$ and

$$I_{22}(2,a,b) = \int_0^1 \int_{m_1}^{M_1} \int_{m_2}^{M_2} I_{22}(1, (a^2-u^2)^{\frac{1}{2}}, (b^2-v^2)^{\frac{1}{2}})\, du\, dv\, dz,$$

where

$$m_1 = \max(-z, -b), \quad M_1 = \min(1-z, b), \quad m_2 = \max(-z, -a)$$

and

$$M_2 = \min(1-z, a).$$

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

Putting $a = b$ and interchanging the order of integrations we can execute the z-integration, obtaining

$$I_{22}(2,a) = \int_0^a\int_0^a [1-\max(u,v)]\, I_{22}\, du\, dv + \int_{-a}^0\int_{-a}^0 [1+\min(u,v)]\, I_{22}\, du\, dv$$
$$+\int_{-a}^0\int_0^a (1+v-u)\, I_{22}\, du\, dv + \int_0^a\int_{-a}^0 (1-v+u)\, I_{22}\, du\, dv$$

where I_{22} in the integrands stands for

$$I_{22} = I_{22}(1, (a^2-u^2)^{\frac{1}{2}}, (a^2-v^2)^{\frac{1}{2}}).$$

Next, splitting various squares in the uv-plane into pairs of isosceles right-angled triangles, and observing various symmetries, we get

$$I_{22}(2,a) = \int_0^a\int_0^u (8-8u-4v)\, I_{22}(1, (a^2-u^2)^{\frac{1}{2}}, (a^2-v^2)^{\frac{1}{2}})\, dv\, du.$$

The trigonometric substitution $u = a\sin\theta$, $v = a\sin\phi$ puts it in the form

$$I_{22}(2,a) = 4a^2\int_0^{\frac{1}{2}\pi}\int_0^{\theta} (2-2a\sin\theta-a\sin\phi)\, I_{22}(1, a\cos\theta, a\cos\phi)\cos\theta\cos\phi\, d\phi\, d\theta,$$

where $I_{22}(1, a\cos\theta, a\cos\phi)$ is given by (48). Here the maximum M and the minimum m are given by $m = a\cos\theta$, $M = a\cos\phi$ because $\cos\theta \leqslant \cos\phi$ if $0 \leqslant \phi \leqslant \theta \leqslant \frac{1}{2}\pi$. Now the integrations can be carried out and we get finally

$$I_{22}(2,a) = \pi^2a^4 - [(8\pi/3)+(256/45)]\, a^5 + [(59\pi/48)+(8/9)]\, a^6. \tag{50}$$

9. *First hypercluster integral.* In this section we consider the first cluster integral $I_{11} = I_{11}(k,d,a,R)$. The values of k and d are general, R is any sufficiently regular region in the d-dimensional Euclidean space, normalized to unit content, and the density f is uniform over R. This I_{11} could be called the simplest k-hypercluster integral or, briefly, the first or principal hypercluster. It is the probability that k points, independently and uniformly at random in R, form a spherical k-tuple cluster of size $\leqslant a$. Our aim is to obtain the principal term

$$I_{11} = Ca^{(k-1)d} + O(a^{(k-1)d+1}), \quad a \text{ small}, \tag{51}$$

and the constant C is needed. The reason for this need is the asymptotic value of the probability Q when n is large and a is small, while $n^{k/(k-1)d}a$ is fixed. It will develop that in this case all the contributions to the general term Σ_r are negligible except for one of them, and that one is precisely the principal hypercluster I_{11}. It corresponds to the most disconnected incidence k-hypergraph consisting of r separate $(k-1)$-dimensional simplexes. Further, even in that hypercluster I_{11} it is necessary only to know the leading coefficient C of (51).

We begin with the special case: $d = 1$ and R is the interval $I = [0,1]$. Here, as it happens, it will be possible to obtain the exact value of the first hypercluster $I_{11} = I_{11}(k,1,a,I)$. In the Euclidean k-space E^k, let H be the entire sample space, which is the unit k-cube

$$H = \{(x_1,\ldots,x_k): 0 \leqslant x_m \leqslant 1, m = 1,\ldots,k\}.$$

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

Since the k-tuple clustering is the event whose probability I_{11} is wanted, any one of the defining conditions (a)–(d) of section 2 may be used. We choose the condition (a): a ball of diameter a contains k points, that is, all the points $x_1, \ldots, x_k$. I_{11} is the volume of a certain subset $S = S(a)$ of H, and with a view to later cases, which will be more complicated, we examine S in some detail. When the size a of the hypercluster is 0, then $S = S(0)$ is simply the diagonal

$$X = \{(x, \ldots, x): 0 \leqslant x \leqslant 1\}$$

of H because the points $x_1, \ldots, x_k$ must then be lumped together. Since $d = 1$ and the one-dimensional ball of diameter a, centred at x, is the interval $[x - \frac{1}{2}a, x + \frac{1}{2}a]$, we define for $0 \leqslant x \leqslant 1$

$$H_a(x) = \{(x_1, \ldots, x_k): |x_m - x| \leqslant a/2, m = 1, \ldots, k\}.$$

Then $H \cap H_a(x)$ is the sample space of all k-tuples $(x_1, \ldots, x_k)$ which form a k-cluster of size $\leqslant a$ centred at x, and so

$$S(a) = H \cap \bigcup_{0 \leqslant x \leqslant 1} H_a(x)$$

is the sample space of all k-clusters that we want. For small a, this $S(a)$ consists of the diagonal X 'padded out' into a long thin para-diagonal spindle.

As was stated, we shall be able to find the exact value of $S(a)$. This is a lucky geometrical accident, like the earlier fitting together of simplexes into a cube, which enabled us to find $Q(n, a, L)$ by simple geometry. We shall therefore develop a technique which will work for more complicated cases, at least as far as finding the principal term is concerned. The idea is somewhat analogous to a change of basis in a vector space, or to an orthogonal transformation to principal axes. Namely, the para-diagonal set $S(a)$ will be approximated by a cylinder set $Z_a(X)$ which is the Cartesian product of the diagonal X and a set in the hyperplane complementary to X. For a simple illustration we take $k = 2$ and we refer to Fig. 3. Here H is the unit square and its hexagonal subset, hugging the diagonal as shown in Fig. 3a, is the exact set $S(a)$, that is, H intersected with the union of all sets $H_a(x)$, $0 \leqslant x \leqslant 1$; one such set $H_a(x)$ is shown. The cylinder $Z_a(X)$ is shown in Fig. 3 (b); it is a diagonally situated rectangle. We have then

$$I_{11}(2, 1, a, I) = 2a - a^2 = \text{Area } S(a),$$
$$\text{Area } Z_a(X) = 2^{\frac{1}{2}} \cdot 2^{\frac{1}{2}} a$$

and so the principal term, which is the lowest power of a, is the same in $S(a)$ and in $Z_a(X)$.

We return now to our first special case, that of finding $I_{11}(k, 1, a, I)$. Consider the set $H_a(a/2)$ which is the same as the cube

$$\{(x_1, \ldots, x_k): 0 \leqslant x_m \leqslant a, m = 1, \ldots, k\}.$$

Let $G_1, \ldots, G_k$ be its k $(k-1)$-dimensional faces meeting at the origin. Let $P_1, \ldots, P_k$ be the generalized k-dimensional rectangles lying in the positive hyperoctant, based on $G_1, \ldots, G_k$, and of unit heights. Thus P_i reaches from a $(k-1)$-dimensional face of H to

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

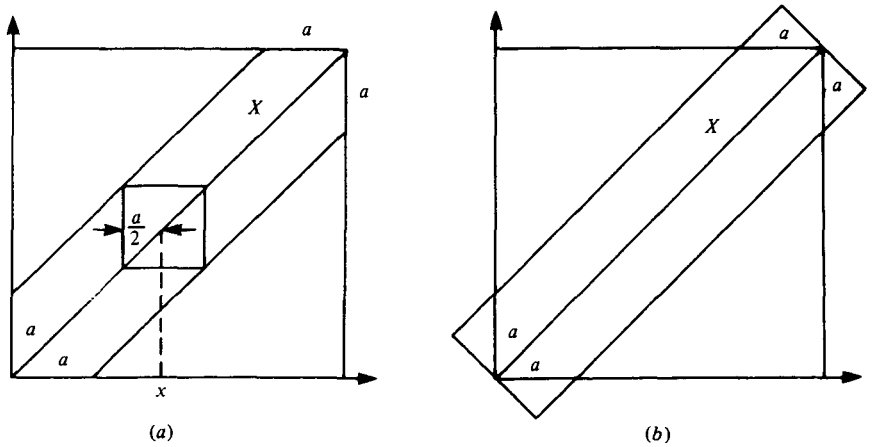


Fig. 3. Para-diagonal subsets.

the face parallel and opposite to it. An application of the Cavalieri principle, aided perhaps by a check of the special cases $k = 2$ and $k = 3$, shows that

$$S(a) = \bigcup_{m=1}^{k} P_m$$

so that
$$\operatorname{Vol} S(a) = \sum_{m=1}^{k} \operatorname{Vol} P_m - (k-1) \operatorname{Vol} \bigcap_{m=1}^{k} P_m.$$

Therefore
$$I_{11}(x, 1, a, I) = ka^{k-1} - (k-1)a^k. \tag{52}$$

We check that, for $k = 2$, this gives us the correct value (31).

Next, we take up a second special case which is at the other extreme: here $k = 2$ but d and R are arbitrary. The sample space is now the generalized square

$$H = \{(p_1, p_2): p_1 \in R, p_2 \in R\}$$

with the diagonal

$$X = \{(p, p): p \in R\}.$$

It is necessary to find the content $\operatorname{Cont}(X)$ of X. To do this we observe that

$$\operatorname{Cont}(X) = C_1 \operatorname{Cont}(R), \tag{53}$$

where the constant C_1 does not depend on R but only on the dimension d. Therefore, for the purpose of determining C_1, R could be any d-dimensional set, and we take it to be the d-dimensional unit cube. Then

$$R = \{(x_1, \ldots, x_d): 0 \leqslant x_m \leqslant 1, m = 1, \ldots, d\}$$

and
$$X = \{(x_1, \ldots, x_d, x_1, \ldots, x_d): 0 \leqslant x_m \leqslant 1, m = 1, \ldots, d\}.$$

Since the content is invariant under symmetries, we have

$$\operatorname{Cont}(X) = \operatorname{Cont}(Y),$$

where
$$Y = \{(x_1, x_1, \ldots, x_d, x_d): 0 \leqslant x_m \leqslant 1, m = 1, \ldots, d\}.$$

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

But Y is clearly a d-dimensional cube of side $2^{\frac{1}{2}}$ so that $\mathrm{Cont}\,(Y) = 2^{\frac{1}{2}d}$ and therefore (53) gives us

$$\mathrm{Cont}\,(X) = 2^{\frac{1}{2}d}\,\mathrm{Cont}\,(R). \tag{54}$$

Actually rather more is true: the 'diagonal' X is similar to the 'side' R and the similarity ratio is $2^{\frac{1}{2}d}:1$. Entirely analogous reasoning shows that for the generalized k-cube

$$H = \{(p_1, \ldots, p_k): p_m \in R, m = 1, \ldots, k\},$$

with the diagonal

$$X = \{(p, \ldots, p): p \in R\},$$

we have

$$\mathrm{Cont}\,(X) = k^{\frac{1}{2}d}\,\mathrm{Cont}\,(R); \tag{55}$$

here again X is similar to R with the similarity ratio $k^{\frac{1}{2}d}:1$.

It may be worth noting that we are dealing here with a certain dimensional generalization of the Pythagorean theorem. Consider k sets $A_1, \ldots, A_k$, each a subset of the d-dimensional Euclidean space. Let these sets be placed in k pairwise orthogonal d-flats of the Euclidean kd-dimensional space, one set in each such flat. Consider the Cartesian product H of all the A_i's and suppose that a diagonal X of H can be unambiguously defined. This is certainly the case when all A_i's are copies of the same set A, and we have then

$$X = \{(p, \ldots, p): p \in A\}.$$

Somewhat more generally, for $1 \leqslant i \leqslant k$, A_i might be similar to the fixed set A, with the similarity ratio $\lambda_i:1$. Now the diagonal is given by

$$X = \{(\lambda_1 p, \ldots, \lambda_k p): p \in A\}.$$

Under these conditions we have the following generalization of the Pythagorean theorem:

$$\sum_{i=1}^{k} [\mathrm{Cont}\,(A_i)]^{\frac{1}{2}d} = [\mathrm{Cont}\,(X)]^{\frac{1}{2}d}.$$

The usual Pythagorean theorem results when we take $d = 1$; A_i is then a segment and $\mathrm{Cont}\,(A_i)$ is just the length a_i of A_i. The similarities of every segment A_i to the unit segment A are automatically assured, $\mathrm{Cont}\,(X)$ is the length x of the 'long' diagonal of the k-dimensional rectangular box of sides $a_1, \ldots, a_k$, so that

$$\sum_{i=1}^{k} a_i^2 = x^2.$$

We return now to the problem of finding the value of $I_{11} = I_{11}(2, d, a, R)$. In analogy to the first special case, let

$$H_a(p) = \{(p_1, p_2): |p - p_1| \leqslant \tfrac{1}{2}a,\ |p - p_2| \leqslant \tfrac{1}{2}a\} \quad (p \in R).$$

Then $H \cap H_a(p)$ is the sample space of 2-clusters (p_1, p_2) centred at p. The set $H_a(p)$ is a generalized square: the Cartesian product of two balls of diameter a. As before,

$$S(a) = H \cap \bigcup_{p \in R} H_a(p)$$

is the sample space of all 2-clusters and so

$$I_{11} = I_{11}(2, d, a, R) = \mathrm{Vol}\,S(a).$$

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

Here we are unable to evaluate I_{11} exactly and so we use the previously introduced technique based on comparing the volumes of the sets

$$S(a) = H \cap \bigcup_{p \in R} H_a(p), Z_a(X). \tag{56}$$

Here the cylinder $Z_a(X)$ is a generalized rectangle: a Cartesian product of its 'long' side X which is the diagonal of the generalized square $H = R \times R$, and its 'short' side which is the diagonal of the generalized square $H_a(p)$. We recall that the latter is the Cartesian product $B \times B$ of two balls of diameter a. Actually, the long first diagonal is the principal diagonal of the Cartesian type $\{(p, p): \ldots\}$ while the short other diagonal is the secondary diagonal of the Cartesian type $\{(p, -p): \ldots\}$. Since $H_a(p) = B \times B$, we define $-p$ quite simply: $-p$ is the reflexion of p in the centre of B. The same would apply if Minkowski distances were used instead of Euclidean ones, since the Minkowski ball is a convex body with a centre of symmetry. Let

$$\gamma(d) = \pi^{\frac{1}{2}d}/\Gamma(1 + \tfrac{1}{2}d)$$

so that

$$\text{Cont}\,(B) = \gamma(d)\,(\tfrac{1}{2}a)^d.$$

Now (54) allows us to determine the contents of both the long diagonal and the short one; these are

$$2^{\frac{1}{2}d} \quad \text{and} \quad 2^{\frac{1}{2}d}\gamma(d)\,(\tfrac{1}{2}a)^d. \tag{57}$$

Since the volumes of the sets in (56) agree in the lowest power of a, the product of the two factors in (57) is the principal term of I_{11}:

$$I_{11}(2, d, a, R) = \gamma(d)\,a^d + O(a^{d+1}). \tag{58}$$

As a check, we verify that this agrees with the values of $I_{11}(p, a)$ found in (42).

Finally, we consider the general first hypercluster $I_{11} = I_{11}(k, d, a, R)$. The kd-dimensional sample space is the generalized cube

$$H = \{(p_1, \ldots, p_k): p_m \in R,\ m = 1, \ldots, k\}$$

in the Euclidean space E^{kd}. The diagonal is

$$X = \{(p, \ldots, p): p \in R\}$$

and we define

$$H_a(p) = \{(p_1, \ldots, p_k): |p - p_m| \leqslant \tfrac{1}{2}a,\ m = 1, \ldots, k\} \quad (p \in R). \tag{59}$$

This is a Cartesian product of k replicas of the d-dimensional ball of diameter a, and $H \cap H_a(p)$ is the sample space of k-clusters of size $\leqslant a$ centred at p. As before, we set up the two sets (56). Then $I_{11} = \text{Vol}\,S(a)$ and we approximate to $\text{Vol}\,S(a)$ by $\text{Vol}\,Z_a(X)$. The cylinder set $Z_a(X)$ is a Cartesian product of X and the projection P of $H_a(p)$ onto the $[(k-1)\,d]$-dimensional flat complementary in E^{kd} to X. For the volume of $Z_a(X)$, we need $\text{Cont}\,(P)$; this is

$$\text{Cont}\,(P) = C_2[\gamma(d)\,(\tfrac{1}{2}a)^d]^{k-1}. \tag{60}$$

The projection constant C_2 in (60) depends only on k and d, and we find that $C_2 = k^{\frac{1}{2}d}$. The factor $\text{Cont}\,(X)$ is given by (55) and now

$$\text{Vol}\,Z_a(X) = \text{Cont}\,(X)\,\text{Cont}\,(P).$$

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

Therefore $$I_{11}(k,d,a,R) = (2^{1-k}k)^d\,[\gamma(d)\,a^d]^{k-1} + O(a^{d(k-1)+1}). \tag{61}$$

One final necessary generalization is easily accomplished: we suppose that the density f is no longer uniform. Then the only change is that I_{11}, instead of being (approximately) the volume of the cylinder $Z_a(X)$, is now the integral over $Z_a(X)$ of a product of densities f, just as in (28). Now, $Z_a(X)$ is the Cartesian product of the long diagonal X and the projection P of $H_a(p)$. When a is small, then so is the diameter of P, and so the only variation in the integrand, essential to the principal term, occurs along the long diagonal X. The changes at right angles to X, that is across P, can be neglected to first approximation. Therefore by using the mean value theorem we find that, for sufficiently smooth f (e.g. f in the class C^1), (61) is replaced by

$$I_{11}(k,d,a,R,f) = (2^{1-k}k)^d\,[\gamma(d)\,a^d]^{k-1}\int_R f^k(x)\,dx + O(a^{d(k-1)+1}). \tag{62}$$

10. *Asymptotics for simple linear clustering.* Let us return to the probability (1) for simple linear uniform clustering. With the normalization $L = 1$ this becomes

$$Q(n,a) = [1-(n-1)\,a]^n \tag{63}$$

and our interest now is in the asymptotics: the limiting behaviour of Q when

$$n \to \infty, \quad a \to 0 \quad (n^2 a = x).$$

Here simple analysis of (63) shows that

$$\lim_{n\to\infty} Q(n, x/n^2) = e^{-x} + e^{-x}(x - x^2/2)\,n^{-1} + O(n^{-2}). \tag{64}$$

This is, of course, quite trivial to derive since we know the explicit form (63) of Q but the important question is: can we derive asymptotic results of the type (64) *without* knowing Q explicitly? This question will be answered affirmatively, and this constitutes perhaps the best justification of our method since there is now the possibility of deriving asymptotic results for the general probability $Q(n,k,d,a,R,f)$.

To begin with, (64) will be derived again, starting with the inclusion–exclusion expansion of Q in the cluster-series

$$Q(n,a) = 1 - N_{11}I_{11} + (N_{21}I_{21} + N_{22}I_{22}) - (N_{31}I_{31} + \ldots + N_{35}I_{35}) + \ldots. \tag{65}$$

It will turn out that not all terms are needed nor even many terms; roughly speaking, only the multiplicities N_{ri} and clusters I_{ri} corresponding to the most disconnected incidence graphs will contribute significantly. For instance, toward the leading term e^{-x} of (64) we need consider only the principal term $N_{r1}I_{r1}$ from each bracket of (65). This term arises from the principal incidence graph consisting of exactly r components each of which is one edge. Accordingly,

$$N_{r1}(n) = (n)_{2r}/2^r r! \sim n^{2r}/2^r r!, \quad n \text{ large},$$
$$I_{r1}(a) = (2a-a^2)^r \sim 2^r a^r, \qquad a \text{ small},$$

and so the leading asymptotic term is

$$\sum_{r=0}^{\infty} (-1)^r n^{2r} a^r / r! \quad \text{or} \quad e^{-x}$$

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

in the limit, as it should be. To get the next term of (64), the one involving $1/n$, it is necessary to develop some crude estimates for the multiplicity $N_{ri}(n)$ when n is large, and for the cluster integral $I_{ri}(a)$ when a is small. For convenience, we write

$$N_{ri}(n) = N(G_{ri}), \quad I_{ri}(a) = I(G_{ri})$$

to bring out the dependence on the incidence graph G_{ri}. The number of vertices of G_{ri} is $v = v(G_{ri})$, the number of connected components of G_{ri} is $c = c(G_{ri})$ and the symmetry-reduction factor is $D(G_{ri})$. Then

$$N(G_{ri}) = n^v/D(G_{ri}) - v(v-1)\,n^{v-1}/2D(G_{ri}) + O(n^{v-2}). \tag{66}$$

For the cluster integral $I(G_{ri})$, we have

$$I(G_{ri}) = \prod_{j=1}^{c} I(G_j), \tag{67}$$

where $G_1, \ldots, G_c$ are the c connected components of G_{ri}. Further,

$$I(G_j) = O(1)\,a^{d[v(G_j)-1]} + O(a^{d[v(G_j)-1]+1})$$

because any one of the $v(G_j)$ vertices of G_j, say P_1, is a point which is free to lie anywhere on the basic interval $[0, 1]$ or, more generally, in the d-dimensional set R, and the remaining $v(G_j) - 1$ vertices must lie somewhere within a d-dimensional ball of radius ba, centred at P_1. So, by (67),

$$I(G_{ri}) = O(1)\,a^{dv(G_{ri})-dc(G_{ri})} + O(a^{dv(G_{ri})-dc(G_{ri})+1}). \tag{68}$$

To compute the $1/n$ term in (64) we shall need three incidence graphs from the general rth bracket in (65); these are shown in Fig. 4. The first one is the principal incidence graph. In the corresponding multiplicity $N_{r1}(n)$ we do not take the highest power of n, as given in (66); that has already been used for the main asymptotic term e^{-x}. Instead, we use the next-highest power from (66), which is

$$-2r(2r-1)\,n^{2r-1}/2^{r+1}r!.$$

The principal power of a, from I_{r1}, is as before $2^r a^r$, and so the contribution is

$$-2r(2r-1)\,n^{2r-1}\,2^r a^r/2^{r+1}r!$$

or in the limit,

$$-n^{-1}r(2r-1)\,x^r/r!.$$

Therefore the total contribution from such terms is

$$-n^{-1}\sum_{r=1}^{\infty}(-1)^r r(2r-1)\,x^r/r! = -n^{-1}e^{-x}(2x^2-x). \tag{69}$$

Next we consider the terms associated with the second incidence graph G_{r2}; here

$$v = 2r-1, \quad D(G_{r2}) = 2^{r-1}(r-2)!, \quad I_{r2}(a) = (2a-a^2)^{r-2}\,I_{22}(a).$$

The value of $I_{22}(a)$ is given by (49). Therefore, neglecting certain powers of n and of a, we find the total contribution here to be

$$\sum_{r=2}^{\infty}(-1)^r n^{2r-1}(2a)^{r-2}\,4a^2/2^{r-1}(r-2)!$$

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

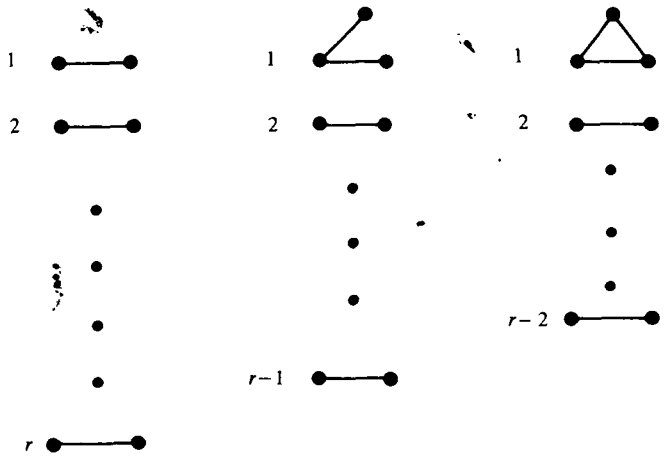


Fig. 4. Highly disconnected incidence graphs.

or in the limit,

$$2n^{-1}\sum_{r=2}^{\infty}(-1)^r x^r/(r-2)! = 2n^{-1}x^2 e^{-x}. \tag{70}$$

Finally there are the contributions associated with G_{r3}. Here

$$v = 2r-3, \quad D(G_{r3}) = 2^{r-3}3!\,(r-3)!, \quad I_{r3}(a) = (2a-a^2)^r I_{35}(a).$$

The value of $I_{35}(a)$ is obtained from (52) on putting $k=3$: $I_{35}(a) = 3a^2 - 2a^3$. Hence the total contribution here is

$$\sum_{r=3}^{\infty}(-1)^r n^{2r-3}(2a)^{r-3}.3a^2/6.(r-3)!\,2^{r-3}$$

or in the limit,

$$(2n)^{-1}x^2\sum_{r=3}^{\infty}(-1)^r x^{r-3}/(r-3)! = -(2n)^{-1}x^2 e^{-x}. \tag{71}$$

Therefore the complete $1/n$ term is the sum of (69), (70) and (71), or

$$n^{-1}e^{-x}(x-x^2/2),$$

in agreement with (64). Of course, the question arises whether there are indeed no other contributions to the $1/n$ term in (64).

As an example, consider two further incidence graphs with r edges, of the highly disconnected type, where all the components are single edges except for the first components; these are respectively the third and the fourth graphs of Fig. 1. Here the estimate (68) and a simple computation of multiplicities show that the contributions are of the order n^4a^3. Therefore in the limit we get

$$n^{-2}(n^6a^3) \cong n^{-2}x^3$$

and so there is indeed no contribution to the $1/n$ term. Similar computation shows that there is no contribution to that term, other than the three considered.

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

11. *General asymptotics.* We derive now the leading term in the asymptotic expansion of the general probability $Q = Q(n, k, d, a, R, f)$ for the limiting case when

$$n \to \infty, \quad a \to 0, \quad n^{k|(k-1)\,d} a = u. \tag{72}$$

The procedure is formally almost the same as in the last section. We start with the inclusion–exclusion expansion of Q in cluster series. Instead of incidence graphs we have now incidence k-hypergraphs G_{ri}. The principal incidence hypergraph G_{r1} consists of r separate copies of a regular simplex with k vertices; it is the maximally disconnected one. For the multiplicity $N_{r1}(n)$ we have

$$N_{r1}(n) = (n)_{rk}/(k!)^r r!.$$

Here the rk vertices of G_{r1} can be labelled in $(n)_{rk}$ ways, one factor $k!$ is obtained by symmetry reduction for each component simplex, and $r!$ arises from the symmetry of r identical components. Therefore

$$N_{r1}(n) = n^{rk}/(k!)^r r! + O(n^{rk-1}).$$

We do not have the value of the principal hypercluster $I_{11}(k, d, a, R, f)$ but its principal term J is given by (62):

$$J = (2^{1-k}k)^d [\gamma(d)\, a^d]^{k-1} \int_R f^k(x)\, dx. \tag{73}$$

Therefore

$$\lim_{n\to\infty} Q(n, k, d, un^{-k|(k-1)d}, R, f) = \lim_{n\to\infty} \sum_{r=0}^{\infty} (-1)^r n^{rk} J^r/(k!)^r r!,$$

which in view of (73) gives us (3).

It is now necessary to justify our procedure, and such justification involves two different things. First, it must be shown that there are indeed no other contributions to the main asymptotic term except those arising from the principal k-hypergraph G_{r1} for different values of r. This presents no new difficulties and we need only some crude estimates of the multiplicities and hyperclusters, of the same type as (68). Such estimates, though perhaps somewhat messy, are obtained in a straightforward manner, much as (68) itself.

Then, it is necessary to justify the passages to the limit. We have started with the inclusion–exclusion series for Q

$$Q = \Sigma_0 - \Sigma_1 + \Sigma_2 - \ldots + (-1)^M \Sigma_M, \quad M = \binom{n}{k}, \tag{74}$$

where $\Sigma_0 = 1$ and the rth term is

$$\Sigma_r = \sum_{j=1}^{t(r)} N_{rj}(n)\, I_{rj}(a) \quad (r = 1, 2, \ldots). \tag{75}$$

Here $t(r)$ is the number of types of hypergraphs with r faces, $N_{rj}(n)$ is the corresponding multiplicity and $I_{rj}(a)$ the corresponding cluster integral. Since $t(r)$ depends neither on n nor on a, there is no problem in justifying the passage to the limit in (75) and

$$\lim \Sigma_r = \lim N_{r1}(n)\, I_{r1}(a), \tag{76}$$

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

because all contributions from (75) with $j > 1$ drop out in the limit. In (76) as in other limits to follow, lim is taken in the sense of (72). The value of the R.H.S. in (76) is the rth term in the power series for the exponential in (3). Since the number $M = M(n)$ of terms in (74) depends on n, it is necessary to justify the interchange of limits and summation in

$$\lim \sum_{r=0}^{M(n)} [(-1)^r \Sigma_r] = \sum_{r=0}^{\infty} (-1)^r \lim \Sigma_r.$$

We get round this difficulty by using the enveloping property, possessed by the series (74). From the very nature of the inclusion–exclusion principle, and from the positivity of probabilities, the successive partial sums of (74) overestimate Q and underestimate it, in turn. Letting $s_m(n, a)$ be the mth partial sum, $m = 0, 1, \ldots$, we express this enveloping property as

$$s_{2m}(n, a) \geqslant Q \geqslant s_{2m+1}(n, a).$$

Keeping m fixed, we let n approach infinity and a approach 0 in the sense of (72). It follows that lim Q is bracketed between the $(2m)$th and the $(2m+1)$th partial sums for the exponential series of the exponential in (3), and we are finished.

12. *Numerical results for the unit square.* As is shown in Section 8, the evaluation of the cluster integrals I_{rj} gets rapidly complicated when r increases, even for a region with the geometry as simple as the square. Since those clusters I_{rj} are necessary for the numerical computation of the probability Q, we have attempted to calculate some of them numerically by the Monte-Carlo methods. The results are reported in this section; all calculations refer to the unit square and to uniform simple clustering. The first two important clusters $I_{11}(a)$ and $I_{22}(a)$ are given by (36) and (50); to four decimal places they are

$$I_{11}(a) = 3{\cdot}1416a^2 - 2{\cdot}6667a^3 + 0{\cdot}5000a^4 \quad (0 \leqslant a \leqslant 1), \tag{77}$$

$$I_{22}(a) = 9{\cdot}8696a^4 - 14{\cdot}0665a^5 + 4{\cdot}7504a^6 \quad (0 \leqslant a \leqslant \tfrac{1}{2}). \tag{78}$$

As a check and a preliminary feasibility test, these were first determined empirically by the Monte-Carlo method. For $I_{11}(a)$ $1{\cdot}5 \times 10^7$ pairs of random points were taken in the unit square, and the empirical frequencies were found for $I_{11}(a)$ at values $a = i/400$, $i = 1, \ldots, 400$. Then a polynomial of the form $Aa^2 - Ba^3 + Ca^4$ was fitted to these data by the least squares method; it turned out to be

$$I_{11}^{\mathrm{Emp}}(a) = 3{\cdot}1405a^2 - 2{\cdot}6649a^3 + 0{\cdot}4993a^4, \tag{79}$$

the total running time (on the Amdahl V6-11 machine) was 282 seconds. The difference $|I_{11}(a) - I_{11}^{\mathrm{Emp}}(a)|$ attains its maximum value at $a = 0{\cdot}585$; dividing this maximum value by $I_{11}(0{\cdot}585)$ gives us the relative error $1{\cdot}6 \times 10^{-4}$.

For $I_{22}(a)$ $1{\cdot}5 \times 10^7$ triples of random points were taken in the unit square and the empirical frequencies for $I_{22}(a)$ were found at $a = i/800$ for $i = 1, \ldots, 400$. The least-squares fitted polynomial was

$$I_{22}^{\mathrm{Emp}}(a) = 9{\cdot}8297a^4 - 13{\cdot}8624a^5 + 4{\cdot}5035a^6 \tag{80}$$

and the total machine time used was 300 seconds. The difference $|I_{22}(a) - I_{22}^{\mathrm{Emp}}(a)|$ attains its maximum on the range $0 \leqslant a \leqslant \frac{1}{2}$ at $a = 0{\cdot}449$, giving the relative error $4{\cdot}4 \times 10^{-4}$.

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

From the above calculations it appears that the first few clusters can be numerically determined in reasonable time with an accuracy of between 3 and 4 significant places. These preliminary results for the clusters we know were considered sufficiently encouraging, and the next three clusters, which we do not know, were similarly computed. These are the clusters $I_{33}(a), I_{34}(a), I_{35}(a)$ corresponding to the three connected incidence graphs for $r = 3$ shown in Fig. 1. The least-squares fits were

$$I_{33}^{\text{Emp}}(a) = 30{\cdot}8642a^6 - 57{\cdot}3222a^7 + 24{\cdot}1126a^8 \quad (0 \leqslant a \leqslant \tfrac{1}{3}), \tag{81}$$

$$I_{34}^{\text{Emp}}(a) = 31{\cdot}0063a^6 - 56{\cdot}7718a^7 + 25{\cdot}2248a^8 \quad (0 \leqslant a \leqslant \tfrac{1}{2}), \tag{82}$$

$$I_{35}^{\text{Emp}}(a) = 5{\cdot}8017a^4 - 6{\cdot}7786a^5 + 1{\cdot}9093a^6 \quad (0 \leqslant a \leqslant \tfrac{1}{2}), \tag{83}$$

In computing these 5×10^7 random quadruples of points were taken for (82), as many random triples for (83), and 10^8 random quadruples for (81). The known leading coefficient in (82) is π^3. The total running time for the three fits was 3461 seconds.

With the known values (77) and (78) for I_{11} and I_{22}, and the empirically fitted values (81)–(83) for I_{33}, I_{34} and I_{35}, the probability $Q = Q(n, a, S)$ was computed as follows. This Q is the probability that when n points are taken independently and uniformly at random in the unit square S, then no two are closer than a. The exact value of Q is given by (27); we use the approximation

$$Q \cong 1 - \Sigma_1 + \Sigma_2 - \Sigma_3 + \Sigma_4' - \Sigma_5' + \Sigma_6'. \tag{84}$$

Here Σ_r is the exact term referring to summation over all incidence graphs with r edges. Σ_r' is the truncated term in which the summation extends only over those incidence graphs for which the clusters can be computed. In our case, this summation is over all the disconnected graphs of r edges for which no component has more than three edges. There are six such graphs for $r = 4$, nine for $r = 5$, and sixteen for $r = 6$.

The numerical results look reasonably good when $na \leqslant 1$. The following rows give the seven partial sums from (84) for $Q(n, 1/n, S)$ at $n = 10, 20, 50, 100, 200$:

1	−0·2960	0·5329	0·1253	0·2431	0·2215	0·2248
1	−0·4295	0·5895	0·0681	0·2718	0·2102	0·2254
1	−0·5133	0·6313	0·0366	0·2777	0·1968	0·2200
1	−0·5419	0·6467	0·0266	0·2754	0·1932	0·2165
1	−0·5563	0·6547	0·0218	0·2733	0·1919	0·2143

The numbers in the last column may be compared with the limiting value from (3)

$$\lim_{n\to\infty} Q(n, 1/n, S) = e^{-\frac{1}{2}\pi} = 0{\cdot}2079.$$

The convergence looks even better when $na < 1$; the following rows show the partial sums for $n = 10$, $a = 0{\cdot}08$, for $n = 20$, $a = 0{\cdot}03$, and for $n = 50$, $a = 0{\cdot}015$:

1	0·1557	0·5069	0·3859	0·4104	0·4073	0·4077
1	0·4764	0·6130	0·5842	0·5890	0·5884	0·5885
1	0·1451	0·5103	0·4007	0·4271	0·4218	0·4227

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press

However, the approximation (84) deteriorates very rapidly if na exceeds 1 even by a little, and it degenerates completely when na reaches the values 1·5–2. This is shown in the following rows of partial sums for $n = 20$, $a = 0{\cdot}06$, for $n = 50$, $a = 0{\cdot}025$, for $n = 50$, $a = 0{\cdot}03$, and for $n = 50$, $a = 0{\cdot}04$:

1	−1·0406	1·0368	−0·4484	0·3506	0·0195	0·1297
1	−1·3545	1·4165	−0·7999	0·5638	−0·1224	0·1702
1	−2·3759	3·3214	−3·1758	2·4775	−1·5150	0·8585
1	−4·9500	12·7518	−22·6340	30·8870	−34·2612	31·9633

These calculations suggest that in matters of convergence, monotonicity, possible unimodality of even and of odd partial sums, etc. the inclusion–exclusion cluster series mimics, in an exaggerated fashion perhaps, the behaviour of the exponential series for e^{-bx}. In view of (1), (3) and (64) this is not surprising.

The following acknowledgements are gratefully made. Prof. D. W. Boyd (University of British Columbia), Dr J. M. Hammersley (Trinity College, Oxford), Prof. I. Olkin (Stanford University) and Prof. G. C. Rota (M.I.T.) have kindly supplied various relevant information and suggestions for improvements. Mr A. Jones (University of British Columbia) has done the programming. The Department of Computer Science (University of British Columbia) was generous with allowances of computing time. The National Research Council of Canada supported the work by a research grant for 1976–9.

REFERENCES

(1) Efron, B. *Ann. Math. Stat.* **38** (1967), 298.
(2) Jung, H. W. E. *J. reine angew. Math.* **123** (1901), 241.
(3) Melzak, Z. A. *Quart. Appl. Math.* **20** (1962), 151.
(4) Melzak, Z. A. *Bell System Tech. J.* **47** (1968), 1105.
(5) Molina, E. C. *Trans. A.I.E.E.* **44** (1929), 294.
(6) Riordan, J. *An introduction to combinatorial analysis* (New York, John Wiley, 1958).

https://doi.org/10.1017/S0305004100056140 Published online by Cambridge University Press