Transpn Res. Vol. 1, pp. 253–269. Pergamon Press 1967. Printed in Great Britain

A STATISTICAL THEORY OF SPATIAL DISTRIBUTION MODELS

A. G. WILSON

Economic Planning Group, Ministry of Transport, London, England

(*Received* 9 *May* 1967)

1. INTRODUCTION

THIS paper offers a theoretical basis for some spatial distribution models which are in common use in locational analysis, and uses the new methodology to generate new models. Such analysis is likely to be concerned with the spatial location of activities among the zones of a region, and measures of interaction between the zones. For example, there may be an interest in zonal activity levels in the form of numbers of workers in residential zones, and the numbers of jobs in employment zones; the interaction between these activities is the journey to work. To fix ideas, this example will be used to illustrate the methods proposed in this paper.

Let T_{ij} be the number of (work) trips, and d_{ij} the distance, between zones i and j, let O_i be the total number of work trip origins in i, and let D_j be the total number of work trip destinations in j. A spatial distribution model then estimates T_{ij} as a function of the O_i's, the D_j's and the d_{ij}'s. These variables could, of course, themselves be functions of other independent variables.

The simplest such model is the so-called gravity model developed by analogy with Newton's law of the gravitational force F_{ij} between two masses m_i and m_j separated by a distance d_{ij}.

$$F_{ij} = \gamma \frac{m_i m_j}{d_{ij}^2} \tag{1}$$

where γ is a constant. The analogous transport gravity model is then

$$T_{ij} = k \frac{O_i D_j}{d_{ij}^2} \tag{2}$$

using the variables defined above, and where k is a constant. This model has some sensible properties: T_{ij} is proportional to each of O_i and D_j and inversely proportional to the square of the distance between them. But the equation has at least one obvious deficiency: if a particular O_i and a particular D_j are each doubled, then the number of trips between these zones would quadruple according to the equation (2), when it would be expected that they would double also. To put this criticism of (2) more precisely, the following constraint equations on T_{ij} should always be satisfied, and they are not satisfied by (2):

$$\sum_j T_{ij} = O_i \tag{3}$$

$$\sum_i T_{ij} = D_j \tag{4}$$

That is, the row and column sums of the trip matrix should be the numbers of trips generated in each zone, and the number of trips attracted, respectively. These constraint equations can be satisfied if sets of constants A_i and B_j associated with production zones and attraction zones respectively are introduced. They are sometimes called balancing factors. Also, there is no reason to think that distance plays its part in the transport equation (2) as it does in the world of Newtonian physics, and so a general function of distance is introduced. The modified gravity model is then

$$T_{ij} = A_i B_j O_i D_j f(d_{ij}) \tag{5}$$

where

$$A_i = \left[\sum_j B_j D_j f(d_{ij})\right]^{-1} \tag{6}$$

and

$$B_j = \left[\sum_i A_i O_i f(d_{ij})\right]^{-1} \tag{7}$$

The equations for A_i and B_j are solved iteratively, and it can easily be checked that they ensure that the T_{ij} given in equation (5) satisfies the constraint equations (3) and (4). Note also that d_{ij} in such a model should be interpreted as a general measure of impedance between i and j, which may be measured as actual distance, as travel time, as cost, or as some weighted combination of such factors sometimes referred to as a "generalized cost". With this proviso, equations (5)–(7) describe a gravity model which has been extensively used, and the discussion above has shown its heuristic derivation by analogy with Newton's gravitational law.

A second approach to trip distribution uses the intervening opportunities model. Interzonal impedance does not appear explicitly in this model, but possible destination zones away from an origin zone i have to be ranked in order of increasing impedance from i. A notation is needed to describe this. Let $j_\mu(i)$ be the μth destination zone in this rank order away from i; $j_\mu(i)$ will be referred to simply as j_μ in cases where it is clear to which i it refers. The intervening opportunities model was first developed by Stouffler (1940) in a simple form, assuming that the number of trips from an origin zone to a destination zone is proportional to the number of opportunities at the destination zone, and inversely proportional to the number of intervening opportunities. The underlying assumption of the model is that the tripper considers each opportunity, as reached, in turn, and has a definite probability that his needs will be satisfied. The model will be derived here in the form developed by Schneider, originally for use in the Chicago Area Transportation Study (1960). To see how the basic assumption operates, consider a situation in which destination zones are rank ordered away from an origin zone as defined earlier. Let U_{ij_μ} be the probability that one tripper will continue beyond the μth zone away from i. Suppose there is a chance L that an opportunity will satisfy this single tripper when it is offered. Then, to the first order in L,

$$U_{ij_1} = 1 - LD_{j_1}$$

where D_{j_1} is the number of opportunities in the zone j_1, nearest to i. Then combining successive probabilities multiplicatively,

$$U_{ij_2} = U_{ij_1}(1 - LD_{j_2})$$

$$U_{ij_3} = U_{ij_2}(1 - LD_{j_3})$$

and so on. In general

$$U_{ij_\mu} = U_{ij_{\mu-1}}(1 - LD_{j_\mu}) \tag{8}$$

This equation can be written

$$\frac{U_{ij_\mu} - U_{ij_{\mu-1}}}{U_{ij_{\mu-1}}} = -LD_{j_\mu}. \tag{9}$$

Let A_{j_μ} be the number of opportunities passed up to and including zone j_μ. Then

$$D_{j_\mu} = A_{j_\mu} - A_{j_{\mu-1}} \tag{10}$$

and (9) can be written

$$\frac{U_{ij_\mu} - U_{ij_{\mu-1}}}{U_{ij_{\mu-1}}} = -L(A_{j_\mu} - A_{j_{\mu-1}}) \tag{11}$$

This equation can be written, assuming continuous variation, as

$$\mathrm{d}U/U = -L\,\mathrm{d}A \tag{12}$$

which integrates to

$$\log U = -LA + \text{constant}$$

so that

$$U_{ij_\mu} = k_i \exp(-LA_{j_\mu}) \tag{13}$$

where k_i is a constant. But

$$T_{ij_\mu} = O_i(U_{ij_{\mu-1}} - U_{ij_\mu}) \tag{14}$$

where T_{ij_μ} is the number of trips from i to the μth destination away from i, for a total of O_i trips originating at i. Substitution from (13) to (14) gives

$$T_{ij_\mu} = k_i O_i[\exp(-LA_{j_{\mu-1}}) - \exp(-LA_{j_\mu})] \tag{15}$$

and this is the usual statement of the intervening opportunities model.

Note that k_i can be chosen so that the resulting matrix T_{ij} satisfies the constraint equation (3):

$$\sum_j T_{ij} = k_i O_i[1 - \exp(-A_{j_N})] = O_i$$

where N is the total number of zones. Since $\exp(-LA_{j_N})$ should be very small, k_i will be very nearly one for each i. The constraint equation (4) on the total number of trip attractions cannot be satisfied, however, within the model structure itself, but if actual D_j's are known, the matrix can be adjusted by the same balancing process as that implied by equations (6) and (7).

(Thus, in general, if there is a matrix T_{ij}^* and it is required to transform it to a matrix T_{ij} whose columns and rows sum to O_i and D_j, then this can be done by the transformation

$$T_{ij} = A_i B_j T_{ij}^* \tag{16}$$

where

$$A_i = O_i\left(\sum_j B_j T_{ij}^*\right)^{-1} \tag{17}$$

$$B_j = D_j\left(\sum_i A_i T_{ij}^*\right)^{-1} \tag{18}$$

and these equations can be seen to reduce to (6) and (7) for the gravity model. This process is accomplished in practice by factoring rows and columns successively by D_j/D_j^* and O_i/O_i^* where the D_j^* are the column sums, and O_i^* the row sums, of the matrix reached after the immediately preceding operation. This is the balancing process which can be applied, if required, to the intervening opportunities model trip matrix.)

It is often argued that the intervening opportunities model is "better" than the gravity model because the theoretical derivation (as outlined above) is sound, whereas the derivation of the gravity model is at best heuristic and based on an analogy with Newton's gravitational law in the physical sciences. The statistical theory of spatial distribution models proposed in this paper is based on an analogy with a different branch of physics, statistical mechanics, and does offer a sound theoretical base for the gravity model. It is also possible, by changing the assumptions, to derive the intervening opportunities model. In fact, the proposed method provides a reasonably general rule for deriving appropriate spatial distribution models for a variety of purposes and situations, and for comparing different models which are supposed to apply to the same situation.

The key to the method is to define a set of variables which completely specify the system, and to enumerate any constraints on these variables. It will usually be possible to do this using the variables T_{ij} defined previously, the number of trips between i and j (still assuming that this is for one purpose and a reasonably homogeneous set of travellers). A set of T_{ij}'s, defined by $\{T_{ij}\}$ then defines a *distribution* of trips. It is also possible to define a *state* of the system as one of the ways in which a distribution is brought about at the micro-level: thus, if the system is made up of individual trippers, a state of the system, for example for the journey to work, is just one way in which these trippers decide on their journey to work in a consistent way (that is, subject to the usual constraints). Note that, at the end of the analysis, the only interest is in distributions—total numbers of trips between points—and not in states—that is, the individual trippers who are making up these trip bundles. The crucial assumption of the new method can now be stated: that the probability of a distribution $\{T_{ij}\}$ occurring is proportional to the number of states of the system which give rise to the distribution $\{T_{ij}\}$. Thus, if $w(T_{ij})$ is the number of ways in which individuals can arrange themselves to produce the overall distribution $\{T_{ij}\}$, then the probability of $\{T_{ij}\}$ occurring is proportional to $w(T_{ij})$. The total number of such arrangements is

$$\sum w(T_{ij}) \tag{19}$$

where the summation is over the distributions which satisfy the constraints of the problem. It will be shown that in most cases there is one distribution $\{T_{ij}\}$ for which $w(T_{ij})$ dominates all other terms of the sum (19) overwhelmingly, and so forms the most probable distribution.

Section 2 of this paper states and discusses this theory for the usual gravity model; Section 3 then applies the theory to a variety of new situations, including an application to multimodal distribution, and also including a derivation of the intervening opportunities model. The final section summarizes the conclusions.

2. A STATISTICAL THEORY OF THE GRAVITY MODEL

2.1. *The "conventional" gravity model*

The new method is most easily illustrated by a single-purpose example, such as the journey to work example discussed above, and so the previously defined example and notation are used. The trip matrix, T_{ij}, must satisfy the two constraint equations (3) and (4), which, for convenience, are stated again here:

$$\sum_j T_{ij} = O_i \tag{20}$$

$$\sum_i T_{ij} = D_j \tag{21}$$

It will also be assumed that another constraint equation is satisfied:

$$\sum_i \sum_j T_{ij} c_{ij} = C \tag{22}$$

where c_{ij} is the impedance, or generalized cost, of travelling between i and j, and so replaces the d_{ij} of the introduction to emphasize that the measure of impedance need not be distance.

This constraint, then, implies that the total amount spent on trips in the region at this point in time, is a fixed amount C. The use of, and need for, this constraint will be made clear as the method develops.

The basic assumption of the method is that the probability of the distribution $\{T_{ij}\}$ occurring is proportional to the number of states of the system which give rise to this distribution, and which satisfy the constraints. Suppose

$$T = \sum_i O_i = \sum_j D_j \tag{23}$$

is the total number of trips. Then the number of distinct arrangements of individuals which give rise to the distribution $\{T_{ij}\}$ is

$$w(T_{ij}) = (T!) \Big/ \Big(\prod_{ij} T_{ij}!\Big) \tag{24}$$

since there is no interest in arrangements within a particular trip bundle. The total number of possible states is then

$$W = \sum w(T_{ij}) \tag{25}$$

where the summation is over T_{ij} satisfying (20)–(22). However, the maximum value of $w(T_{ij})$ turns out to dominate the other terms of the sum to such an extent that the distribution $\{T_{ij}\}$, which gives rise to this maximum is overwhelmingly the most probable distribution. This maximum will now be obtained, and its sharpness, and the validity of the method in general, will then be discussed in Section 2.3 below, following a section which discusses the interpretation of particular terms.

To obtain the set of T_{ij}'s which maximizes $w(T_{ij})$ as defined in (24) subject to the constraints (20)–(22), the function M has to be maximized, where

$$M = \log w + \sum_i \lambda_i^{(1)} \Big(O_i - \sum_j T_{ij}\Big) + \sum_j \lambda_j^{(2)} \Big(D_j - \sum_i T_{ij}\Big) + \beta \Big(C - \sum_i \sum_j T_{ij} c_{ij}\Big) \tag{26}$$

and where $\lambda_i^{(1)}$, $\lambda_j^{(2)}$ and β are Lagrangian multipliers. Note that it is more convenient to maximize $\log w$ rather than w, and then it is possible to use Stirling's approximation

$$\log N! = N \log N - N \tag{27}$$

to estimate the factorial terms. The T_{ij}'s which maximize M, and which therefore constitute the most probable distribution of trips, are the solutions of

$$\frac{\partial M}{\partial T_{ij}} = 0 \tag{28}$$

and the constraint equations (20)–(22). Using Stirling's approximation, (27), note that

$$\frac{\partial \log N!}{\partial N} = \log N \tag{29}$$

and so

$$\frac{\partial M}{\partial T_{ij}} = -\log T_{ij} - \lambda_i^{(1)} - \lambda_j^{(2)} - \beta c_{ij} \tag{30}$$

and this vanishes when

$$T_{ij} = \exp(-\lambda_i^{(1)} - \lambda_j^{(2)} - \beta c_{ij}) \tag{31}$$

Substitute in (20) and (21) to obtain $\lambda_i^{(1)}$ and $\lambda_j^{(2)}$:

$$\exp[-\lambda_i^{(1)}] = O_i \Big/ \left[\sum_j \exp(-\lambda_j^{(2)} - \beta c_{ij})\right] \tag{32}$$

$$\exp[-\lambda_j^{(2)}] = D_j \Big/ \left[\sum_i \exp(-\lambda_i^{(1)} - \beta c_{ij})\right] \tag{33}$$

To obtain the final result in more familiar form, write

$$A_i = \exp(-\lambda_i^{(1)})/O_i \tag{34}$$

and

$$B_j = \exp(-\lambda_j^{(2)})/D_j \tag{35}$$

and then

$$T_{ij} = A_i B_j O_i D_j \exp(-\beta c_{ij}) \tag{36}$$

where, using equations (32)–(35),

$$A_i = \left[\sum_j B_j D_j \exp(-\beta c_{ij})\right]^{-1} \tag{37}$$

$$B_j = \left[\sum_i A_i O_i \exp(-\beta c_{ij})\right]^{-1} \tag{38}$$

Thus the most probable distribution of trips is the same as the gravity model distribution discussed earlier, and defined in equations (5)–(7), and so this statistical derivation constitutes a new theoretical base for the gravity model. Note that C in the cost constraint equation (22) need not actually be known, as this equation is not in practice solved for β. This parameter would be found by the normal calibration methods. However, if C was known, then (22) could be solved numerically for β.

This statistical theory is effectively saying that, given total numbers of trip origins and destinations for each zone for a homogeneous person–trip purpose category, given the costs of travelling between each zone, and given that there is some fixed total expenditure on transport in the region, then there is a most probable distribution of trips between zones, and this distribution is the same as the one normally described as the gravity model distribution. Students of statistical mechanics will recognize the method as a variation of the micro-canonical ensemble method for analysing systems of particles, for example, the molecules of a gas.

2.2. *Interpretation of terms*

It has always been a feature of statistical mechanics that the terms which occur in the equation giving the most probable distribution are then seen to have physical significance. This is true here also. O_i, D_j and c_{ij} were defined previously. The expression $\exp(-\beta c_{ij})$ appears in this formulation as the preferred form of distance deterrence function, and the parameter β is determined in theory by the cost constraint equation (22). It does, however, have its usual interpretation: it is closely related to the average distance travelled. The greater β, the less is the average distance travelled. This is then obviously related to C of equation (22). If C is increased, then more is spent on travelling and distances will increase, but an examination of the left-hand side of (22) shows that β would decrease. The remaining task is to interpret A_i and B_j.

Suppose one of the D_j's changes, say D_1. Then

$$T_{i1} = A_i B_1 O_i D_1 \exp(-\beta c_{i1}) \tag{39}$$

and if D_1 changes substantially, the trips from each i to zone 1 would change in proportion.

The next largest change will be in the A_i's as defined by (37), but the change will not be large as the expression involving D_1 in each A_i is only one of a number of terms. The B_j's will probably be affected even less, as any change is brought about through changes in the A_i's.

Suppose, then, that D_1 is substantially increased. Then the T_{i1}'s will increase more or less in proportion. The A_i's will decrease by a lesser amount relatively, and the B_j's will increase even more slightly. The role of the A_i's, then, will be to reduce all trips slightly to compensate for the increase in trips to zone 1. A_i can thus be seen as a competition term which reduces most trips due to the increased attractiveness of one zone. The denominator of A_i is also commonly used as a measure of accessibility, and the increase in D_1 could be said to increase the accessibility of everyone to opportunities at 1, though more usually such an interpretation would be reserved for changes in the c_{ij}'s. Thus, this analysis establishes a competition–accessibility interpretation of the A_i's. The B_j's play a similar role, and would be responsible for the main adjustments if the major change was in an O_i rather than a D_j. A change in the c_{ij}'s, or several O_i's and D_j's simultaneously, would bring about complex readjustments through the A_i's and the B_j's.

One consequence of this interpretation, and of the use of the new method which gives a fundamental role to the A_i's and B_j's, is that it shows that the interpretation of the A_i's and B_j's suggested by Dieter (1962) is wrong. Dieter suggested that the A_i's and B_j's should be associated with terminal costs, say a_i and b_j in origin and destination zones i and j respectively. This can be checked by replacing c_{ij} by $a_i+b_j+c_{ij}$ in the preceding analysis, and this gives for T_{ij}:

$$T_{ij} = A_i B_j O_i D_j \exp(-\beta a_i - \beta b_j - \beta c_{ij}) \tag{40}$$

Thus, new terms $\exp(-\beta a_i)$ and $\exp(-\beta b_j)$ are introduced, but the A_i's and B_j's are still present independently of the existence of terminal costs.

2.3. *Validity of the method*

There are two possible points of weakness in the method. Firstly, is Stirling's approximation, in equation (27), valid for the sort of T_{ij}'s that occur in practice? Secondly, is the maximum value of

$$\frac{T!}{\prod_{ij} T_{ij}!}$$

a very sharp maximum?

The first of the doubts can be answered by analogy. The use of Stirling's approximation underlies one particular approach in statistical mechanics and is used, as here, to produce most probable distributions. There is, however, a second method, the Darwin–Fowler method, which actually calculates the individual terms of the sum in equation (19) by using a generating function and complex integration. These terms are then used as weights to calculate the means of all the distributions, and these mean values have been shown to be the same as the most probable values obtained by using Stirling's approximation, even in the cases where the numbers involved are obviously so small that Stirling's theorem is not valid. It is a safe conjecture that the same result applies here: that theoretically valid results can be obtained, using the method above, at all times.

The second question can be answered explicitly if small changes in $\log w(T_{ij})$ are examined near the maximum. At, or very near, the maximum the terms of $\mathrm{d}[\log w(T_{ij})]$ which are linear in $\mathrm{d}T_{ij}$ vanish, and

$$\mathrm{d}[\log w(T_{ij})] = \tfrac{1}{2}\sum_i \sum_j \frac{\partial^2 \log w}{\partial T_{ij}^2} (\mathrm{d}T_{ij})^2 \tag{41}$$

It will be recalled that

$$\frac{\partial \log w}{\partial T_{ij}} = -\log T_{ij}$$

and so

$$\frac{\partial^2 \log w}{\partial T_{ij}^2} = -T_{ij}^{-1} \tag{42}$$

Substituting in equation (41),

$$\mathrm{d}\,[\log w(T_{ij})] = -\tfrac{1}{2}\sum_i\sum_j \frac{(\mathrm{d}T_{ij})^2}{T_{ij}} = -\tfrac{1}{2}\sum_i\sum_j \left(\frac{\mathrm{d}T_{ij}}{T_{ij}}\right)^2 T_{ij} \tag{43}$$

Thus (43) can be written

$$\mathrm{d}\,[\log w(T_{ij})] = -\tfrac{1}{2}\sum_i\sum_j p^2 T_{ij} \tag{44}$$

where p is the percentage change in each T_{ij} away from the most probable distribution.

To evaluate this expression, the size distribution of elements of the trip matrix is needed. Suppose there are N size groups, and that the nth group has T_{ij}'s with a mean value T_n, and there are S_n such trip matrix elements in this group. Then, (44) can be written

$$\mathrm{d}\,[\log w(T_{ij})] = -\tfrac{1}{2}p^2 \sum_n S_n T_n \tag{45}$$

Now consider a typical example: take a large urban area with, say 1000 zones. Suppose 1000 trip interchanges have each got 10^4 trips, 10,000 have 10^3 and 100,000 have 10^2. Let $p = 10^{-3}$. Then

$$\mathrm{d}\log w \approx -\tfrac{1}{2}10^{-6}(10^7 + 10^7 + 10^7) \approx -15$$

Thus, $\log w$ changes by -15 for a change of one part in a thousand of each element of the trip matrix away from the most probable distribution. Thus w drops by the enormous factor of e^{-15}, which gives an indication of just how sharp the maximum can be. Such an estimate of $\mathrm{d}(\log w)$ can be calculated as a check in any particular case. One of the advantages of this new approach is that it gives the possibility of doing this check, and ruling out certain situations as being unsuitable for the gravity model approach should the maximum turn out not to be a sharp one.

A third result of interest can also be stated by analogy with the corresponding result in statistical mechanics (cf. Tolman, 1938). That is:

$$\frac{\overline{(T_{ij} - \bar{T}_{ij})^2}}{\bar{T}_{ij}^2} = T_{ij}^{-1} - T^{-1} \tag{46}$$

This gives the dispersion of T_{ij} and indicates, as is well known in practice, that estimates are better for large flows than for small ones.

This analysis has shown that the gravity model has a sound base. However, it should be recalled that the whole analysis has been for a single trip purpose, and for a homogeneous group of travellers. People are not identical in the way that particles in physics are identical, and so no theory of this form (indeed, no theory period) can be expected to apply exactly. This analysis has shown, in effect, that good results can be expected if trips can be classified by purpose and by person type in a reasonably uniform way.

3. NEW APPLICATIONS OF THE STATISTICAL THEORY

3.1. *Rules for constructing distribution models*

The previous sections have shown that a statistical theoretical base can be given to the conventional gravity model. The principle on which this derivation is based is, however, quite general. The only assumption is that the probability of a distribution occurring is proportional to the number of states of the system which give rise to that distribution, subject to a number of constraints. It can easily be seen that if there were no constraints at all, then all the T_{ij}'s would have an equal share of the total number of trips. In other words, it is the constraints which have the effect of giving a distribution of trips other than a trivial one. One of the remarkable features of this statistical theory is that it produces the conventional gravity model using constraints which really say relatively little, that is, are relatively unrestrictive. The new theory reveals that an effective way of developing better models is to refine the constraints which are applied to behaviourable variables to make them more restrictive. This makes more precise for this problem what is simply the normal method of scientific research. (See, for example, Popper, 1959.) This new theory is a powerful general method for constructing spatial distribution models: the main task to produce a model in this framework is to discover the constraints on the variables which describe the spatial distribution problem, and then to maximize, as a function of the distribution variables, the number of states which can give rise to the distribution subject to the constraints.

The derivation of the conventional gravity model above followed this pattern. The following sections of the paper apply the general theory to a number of situations, ranging from new ones being tackled for the first time to a derivation of the intervening opportunities model.

3.2. *The single competition term gravity model*

The so-called conventional gravity model was described by equations (5)–(7) above. A much-used variant of this simply takes all the B_j's as one at the expense of failing to satisfy the constraints (4). So this model is

$$T_{ij} = A_i O_i D_j f(c_{ij}) \tag{47}$$

where

$$A_i = \left[\sum_j D_j f(c_{ij})\right]^{-1} \tag{48}$$

to ensure that

$$\sum_j T_{ij} = O_i \tag{49}$$

but where, if

$$D_j^* = \sum_i T_{ij},$$

D_j^* is not necessarily equal to D_j. This model may be used in a variety of special circumstances: for example, when D_j is simply some measure of attraction, and D_j^* is then *defined* to be the resulting number of trips produced in the model. That is, no constraint of the form (4) is assumed to hold. The new method can be applied to this situation easily: maximize

$$\frac{T!}{\prod_{ij} T_{ij}!}$$

subject to the constraints (49), and subject to a generalized cost constraint (22). The resulting distribution, which is analogous to (31) as derived from (30), is

$$T_{ij} = \exp(-\lambda_i^{(1)} - \beta c_{ij}) \tag{50}$$

where the $\lambda_i^{(1)}$'s and β are Lagrangian multipliers. The $\lambda_i^{(1)}$ can be found by substituting in (49) with the result

$$T_{ij} = \frac{O_i \exp(-\beta c_{ij})}{\sum_k \exp(-\beta c_{ik})} \tag{51}$$

Now this equation resembles (47) and (48), but the D_j term of the latter equation is now missing. However, it will be recalled that in this case D_j is likely to be a measure of attraction, and not a number of trip ends. Such a term can be introduced into equation (51) by the following interesting device. Assume that the traveller to j receives some benefit W_j which can be set against the transport cost c_{ij} over and above the benefit which can be obtained from going to other zones. (So W_j may be a measure of the scale economies available to the shopper in large shopping centres.) Equation (51) can then be rewritten with c_{ij} being replaced by $c_{ij}-W_j$ as

$$T_{ij} = \frac{O_i \exp(\beta W_j - \beta c_{ij})}{\sum\limits_k \exp(\beta W_k - \beta c_{ik})} \tag{52}$$

so that we can now identify the new model with that of equations (47)–(48) by taking $\exp(\beta W_j)$ as the attractive measure D_j. Since D_j in such applications is usually taken as a zonal size variable which is a proxy for scale benefits of the form W_j, it seems intuitively reasonable to expect W_j to vary as $\log D_j$, as implied here, rather than D_j.

3.3. *Distribution of trips when there are several transport modes*

This section discusses a more important problem. Consider the basic situation illustrated earlier by the journey to work, and described by variables T_{ij}, O_i, D_j and c_{ij}. Suppose now, however, that there are several possible modes of transport between i and j, and the cost of travelling by the kth mode is $c_{ij}{}^k$. Let T_{ij}, O_i and D_j be defined as before, and let $T_{ij}{}^k$, $O_i{}^k$ and $D_j{}^k$ be the proportions of these trip totals carried by mode k.

The constraint equations which describe this situation are

$$\sum_i \sum_k T_{ij}{}^k = D_j \tag{53}$$

$$\sum_j \sum_k T_{ij}{}^k = O_i \tag{54}$$

$$\sum_i \sum_j \sum_k T_{ij}{}^k c_{ij}{}^k = C \tag{55}$$

and the maximand, subject to these constraints, is

$$\frac{T!}{\prod\limits_{ijk} T_{ij}{}^k!}$$

Note that the constraint equations (53) and (54) do not use any knowledge of trip end modal split, and because of this, there are the same number of Lagrangian multipliers as for the single-mode case. The most probable distribution, obtained in the usual way, is

$$T_{ij}{}^k = \exp(-\lambda_i{}^{(1)} - \lambda_j{}^{(2)} - \beta c_{ij}{}^k) \tag{56}$$

and, substituting in (53) and (54),

$$\exp(-\lambda_j{}^{(2)}) = \frac{D_j}{\sum\limits_i \sum\limits_k \exp(-\lambda_i{}^{(1)} - \beta c_{ij}{}^k)} \tag{57}$$

$$\exp(-\lambda_i{}^{(1)}) = \frac{O_i}{\sum\limits_j \sum\limits_k \exp(-\lambda_j{}^{(2)} - \beta c_{ij}{}^k)} \tag{58}$$

So, putting

$$A_i = \exp(-\lambda_i{}^{(1)})/O_i, \quad B_j = \exp(-\lambda_j{}^{(2)})/D_j$$

as before,

$$T_{ij}{}^{k} = A_i B_j O_i D_j \exp(-\beta c_{ij}{}^{k}) \tag{59}$$

where

$$A_i = \left[\sum_j \sum_k B_j D_j \exp(-\beta c_{ij}{}^{k})\right]^{-1} = \left[\sum_j B_j D_j \sum_k \exp(-\beta c_{ij}{}^{k})\right]^{-1} \tag{60}$$

$$B_j = \left[\sum_i \sum_k A_i O_i \exp(-\beta c_{ij}{}^{k})\right]^{-1} = \left[\sum_i A_i O_i \sum_k \exp(-\beta c_{ij}{}^{k})\right]^{-1} \tag{61}$$

Note also that

$$T_{ij} = \sum_k T_{ij}{}^{k} = A_i B_j O_i D_j \sum_k \exp(-\beta c_{ij}{}^{k}) \tag{62}$$

Thus, equations (59)–(61) define a multi-mode distribution model. $T_{ij}{}^{k}$ and T_{ij} can be divided, using equations (59) and (62), to give the modal split as

$$\frac{T_{ij}{}^{k}}{T_{ij}} = \frac{\exp(-\beta c_{ij}{}^{k})}{\sum_k \exp(-\beta c_{ij}{}^{k})} \tag{63}$$

as the proportion travelling by mode k between i and j. Note that in the two-mode case, say with modes 1 and 2 representing public and private transport, a plot of, say, $T_{ij}{}^{1}/T_{ij}$ would give a curve which has the shape of the usual diversion curve.

Note further that the T_{ij} derived in (62) above can be wholly identified with the T_{ij} derived in the conventional gravity model in equation (36) provided that

$$\exp(-\beta c_{ij}) = \sum_k \exp(-\beta c_{ij}{}^{k}) \tag{64}$$

This equation is of the greatest importance because it shows how a composite measure of impedance, $\exp(-\beta c_{ij})$, or average generalized cost, c_{ij}, can be derived from the modal impedances $\exp(-\beta c_{ij}{}^{k})$ where these are known individually. Such composite impedances are valuable in a variety of planning models, but past practice has been to use one of a number of arbitrary averaging procedures.

It should also be remarked that the modal split formula (63) is identical in form to that derived from a statistical approach to modal split using discriminant analysis (Quarmby, 1967). There could be complete identification if the generalized cost $c_{ij}{}^{k}$ could be identified with the discriminant function used by the statisticians. If such an identification can be made, then discriminant analysis would provide a method for determining the generalized costs.

Finally, let us examine the case where there is independent information on trip end estimation by mode. Suppose it is known that not only are there O_i trips in total from zone i, but that $O_i{}^{k}$ are by mode k, and that similarly there are $D_j{}^{k}$ trips by mode k into zone j. The constraints analogous to (53) and (54) can be written as

$$\sum_i T_{ij}{}^{k} = D_j{}^{k} \tag{65}$$

$$\sum_j T_{ij}{}^{k} = O_i{}^{k} \tag{66}$$

Sets of Lagrangian multipliers $\lambda_j{}^{(2)k}$, $\lambda_i{}^{(1)k}$ are now defined to be associated with these constraints. The maximization process is carried through in the usual way giving

$$T_{ij}{}^{k} = A_i{}^{k} B_j{}^{k} O_i{}^{k} D_j{}^{k} \exp(-\beta c_{ij}{}^{k}) \tag{67}$$

where

$$A_i^k = \left[\sum_j B_j^k D_j^k \exp(-\beta c_{ij}^k)\right]^{-1} \tag{68}$$

$$B_j^k = \left[\sum_i A_i^k O_i^k \exp(-\beta c_{ij}^k)\right]^{-1} \tag{69}$$

Notice that in this case, T_{ij}^k/T_{ij} is not as simple a ratio as with the previous one and no longer agrees with the results of the discriminant analysis approach.

An alternative result is obtained if the constraint (53) is used with (66) together with the usual cost constraint (55). (This represents the situation where trip generations are known by mode but trip attractions only in total for each zone.) This can be worked out in the usual way and the result is

$$T_{ij}^k = A_i^k B_j O_i^k D_j \exp(-\beta c_{ij}^k) \tag{70}$$

where

$$A_i^k = \left[\sum_j B_j D_j \exp(-\beta c_{ij}^k)\right]^{-1} \tag{71}$$

and

$$B_j = \left[\sum_i \sum_k A_i^k O_i^k \exp(-\beta c_{ij}^k)\right]^{-1} \tag{72}$$

The objective of obtaining these results and comparing them is to assess their importance for modal split applications in transportation studies, since in some cases it is assumed that trip ends by mode can be estimated, for example by regression analysis.

The model represented by equations (67)–(69), that is assuming that trip end modal split is known, is commonly used. Examination of the structure of these equations shows that they are completely separated for each mode: the trip ends are estimated separately, and then each mode is distributed separately. The weakness of this method arises in the need for forecasting as well as to explain the present, or in any situation where some of the parameters change: any aggregate switching of mode can only be brought about by changes in the O_i^k's and D_j^k's. By contrast, with the first multi-modal distribution model derived above, described by equations (53)–(55), the aggregate levels of modal choice are determined, given only total O_i's and D_j's by the relative costs, the c_{ij}^k's. This seems a much more fundamental mechanism than the separate mode methods, and is directly related to the sort of inter-modal comparison a traveller may be expected to make. Its connection with the discriminant analysis approach also adds weight to the view that it is a preferable method.

The intermediate method, that is the use of the model described by equations (70)–(72), falls between the two stools, and is probably not in current use anyway. However, it will be seen in the next section that a model of this structure will become important in the case where two types of traveller are considered: car owners and non-car owners.

3.4. *Extension of the multi-mode model to the case where some users have access to only a subset of all modes*

The multi-mode distribution models produced in Sections 3.3 assume implicitly that all travellers have access to all modes. There is at least one obvious case in real life where this is not so: non-car owners do not have the possibility of travel by car. This is important for current transportation study models, where usual practice is often to assume that trip ends for car owners, for example, can be separately estimated, and the trips separately distributed. This suffers from the same forecasting deficiency as the separate modal distributions of the previous section. It is, in fact, an exactly analogous assumption. However, once again the situation can be described by the appropriate constraints and a model can be produced which seems more appropriate to the situation. The basis of the constraint equations will be that, if any set of travellers should have only a subset of the modes available, then the trip end productions should be similarly categorized, but not trip attractions, so that all travellers compete for the same attractions, but non-car owners, for example, cannot generate trips by car. Thus trip productions would be generated separately for car owners and non-car owners, but only total trip attractions would be estimated.

The first step, however, is to get a general formulation of the problem and to develop an appropriate notation. Let n represent a class of travellers, and let $\gamma(n)$ be the set of modes available to travellers in category n. k will denote mode as usual, and

$$\sum_{k\epsilon\gamma(n)}$$

denotes summation over the subset of modes k available to persons of the type n. The new constraints, embodying the principles discussed above, are then

$$\sum_{j}\sum_{k\epsilon\gamma(n)} T_{ij}{}^{kn} = O_i{}^n \tag{73}$$

$$\sum_{i}\sum_{n}\sum_{k\epsilon\gamma(n)} T_{ij}{}^{kn} = D_j \tag{74}$$

$$\sum_{i}\sum_{j}\sum_{n}\sum_{k\epsilon\gamma(n)} T_{ij}{}^{kn} = C \tag{75}$$

where, in an obvious notation, $T_{ij}{}^{kn}$ is the number of trips from i to j by mode k by traveller type n, $O_i{}^n$ is the number of trip generations at i by travellers of type n, and other variables have been defined before. The maximand is now

$$\frac{T!}{\prod\limits_{ijkn} T_{ij}{}^{kn}!} \tag{76}$$

subject to the constraints (73)–(75). Introduce Lagrangian multipliers $\lambda_i{}^{(1)n}$, $\lambda_j{}^{(2)}$ and β in the usual way and the maximizing condition is

$$-\log T_{ij}{}^{kn} - \lambda_i{}^{(1)n} - \lambda_j{}^{(2)} - \beta c_{ij}{}^k = 0 \tag{77}$$

so

$$T_{ij}{}^{kn} = \exp(-\lambda_i{}^{(1)n} - \lambda_j{}^{(2)} - \beta c_{ij}{}^k) \tag{78}$$

and writing

$$A_i{}^n = \exp(-\lambda_i{}^{(1)n})/O_i{}^n \tag{79}$$

and

$$B_j = \exp(-\lambda_j{}^{(2)})/D_j \tag{80}$$

the usual manipulation gives

$$T_{ij}{}^{kn} = A_i{}^n B_j O_i{}^n D_j \exp(-\beta c_{ij}{}^k) \tag{81}$$

where

$$A_i{}^n = \left[\sum_{j}\sum_{k\epsilon\gamma(n)} B_j D_j \exp(-\beta c_{ij}{}^k)\right]^{-1} \tag{82}$$

and

$$B_j = \left[\sum_{j}\sum_{n}\sum_{k\epsilon\gamma(n)} A_i{}^n O_i{}^n \exp(-\beta c_{ij}{}^k)\right]^{-1} \tag{83}$$

Note that we can now get the total inter-zonal trips by mode (by summing over n, denoted by T_{ij}^k), by traveller type (by summing over $k \in \gamma(n)$, denoted by T_{ij}^n), and in total (by summing over $k \in \gamma(n)$ and n, denoted by T_{ij}). Thus

$$T_{ij}^k = B_j D_j \left(\sum_n A_i^n O_i^n \right) \exp(-\beta c_{ij}^k) \tag{84}$$

$$T_{ij}^n = A_i^n B_j O_i^n D_j \sum_{k \in \gamma(n)} \exp(-\beta c_{ij}^k) \tag{85}$$

$$T_{ij} = B_j D_j \sum_n \sum_{k \in \gamma(n)} A_i^n O_i^n \exp(-\beta c_{ij}^k) \tag{86}$$

Note that in equation (85), the results for one person category n is linked to the other person type variables through the B_j's defined in (83). This arises because different person categories are competing for the same attractions.

A result which is suitable for a car owner/non-car owner split can now be obtained easily from these general equations.

3.5. *Derivation of the intervening opportunities model*

The intervening opportunities model was derived in the traditional way in the introduction to this paper and its main equation was derived as equation (15). It is of some interest to attempt to derive this using the new methodology, since, if this is possible, the gravity and opportunities models are related by this common base and can be compared in a new light.

Using the variables defined in the introduction, it is also possible to define in addition

$$S_{ij_\mu} = O_i U_{ij_\mu} \tag{87}$$

as the number of trips from i continuing beyond the μth ranked zone away from i. Note that, since

$$T_{ij_\mu} = S_{ij_{\mu-1}} - S_{ij_\mu} \tag{88}$$

the variables S_{ij_μ} define the new system as a possible alternative to T_{ij_μ}. To derive the opportunities model, the new method is applied to the variables S_{ij_μ}. Thus if S is the total numbers of states for a given distribution $\{S_{ij_\mu}\}$, then the maximand will be

$$\frac{S}{\prod_{ij_\mu} S_{ij_\mu}!}$$

It is now necessary to establish appropriate constraints. As seen earlier, the opportunities model does not have a constraint on trip attractions of the form of equation (4), but does need a constraint on trip generations of the form (3). For the variables S_{ij_μ}, the strictly analogous constraint to (3) is the inequality

$$S_{ij_\mu} \leqslant O_i \tag{89}$$

as there cannot be more trips continuing beyond a point from i than originally set out from i. If these are summed over j_μ to get a constraint of the form (3), the resulting equation is

$$\sum_{j_\mu} S_{ij_\mu} = k_i' O_i \tag{90}$$

where k_i' is some constant, and $1 \leqslant k_i' \leqslant N$, where N is the total number of zones. Finally, a constraint analogous to the gravity model cost constraint, equation (22), is needed.

The main assumption of the intervening opportunities model, as commonly stated, is that the number of trips between i and j is determined by the number of opportunities at j, and varies inversely as the number of intervening opportunities. This gives the clue for the cost constraint: to use intervening opportunities as a proxy for cost. Thus, if S_{ij_μ} trips are to be made beyond j_μ, then these will incur costs greater than those for trips which have been made to nearer zones. Suppose, then, for trips from i, we take the number of opportunities passed as a measure of the cost of getting so far. Thus, the minimum cost for the remaining trips beyond j_μ is $A_{j_\mu} S_{ij_\mu}$. If this is summed over j_μ and then over all origin zones i, this gives a function which behaves in some ways like a total cost function, and the corresponding constraint is

$$\sum_i \sum_{j_\mu} A_{ij_\mu} S_{ij_\mu} = C \tag{91}$$

Since, as can be derived from the definitional equation (10),

$$A_{j_\mu} = \sum_{n=1}^{\mu} D_{j_n} \tag{92}$$

and, as can be derived from the definition of S_{ij_μ},

$$S_{ij_\mu} = \sum_{n=\mu+1}^{N} T_{ij_n} \tag{93}$$

where N is the total number of zones, it can easily be seen that the coefficient of T_{ij_μ} in the summation in equation (91) is [substituting for S_{ij_μ} from (93)]

$$(\mu-1)\, D_{j_1} + (\mu-2)\, D_{j_2} + \ldots + D_{j_{\mu-1}} \tag{94}$$

and so the opportunities passed contribute to the cost associated with a particular element of the trip matrix weighted by the number of times they have been "passed" or "have intervened".

Now, maximizing

$$\frac{S}{\prod_{ij_\mu} S_{ij_\mu}!}$$

subject to the constraints (90) and (91), introducing Lagragian multipliers $\lambda_i^{(1)}$ and L for these constraints, the most probable distribution occurs when

$$-\log S_{ij_\mu} - LA_{j_\mu} - \lambda_i^{(1)} = 0$$

so

$$S_{ij} = \exp(-LA_{j_\mu} - \lambda_i^{(1)}). \tag{95}$$

$\lambda_i^{(1)}$ can be obtained in the usual way by substituting from (95) into (90):

$$\exp(-\lambda_i^{(1)}) = \frac{k_i' O_i}{\sum_{j_\mu} \exp(-LA_{j_\mu})} \tag{96}$$

and so, writing

$$k_i = \frac{k_i'}{\sum_{j_\mu} \exp(-LA_{j_\mu})} \tag{97}$$

$$S_{ij_\mu} = k_i O_i \exp(-LAj_\mu) \tag{98}$$

and, using equation (88),

$$T_{ij_\mu} = k_i O_i [\exp(-LA_{j_{\mu-1}}) - \exp(-LA_{j_\mu})] \tag{99}$$

which is identical to equation (15). Thus, the main equation of the intervening opportunities model has been obtained using the new method.

This derivation has been made at the expense of using a rather strange cost constraint equation (92), and assuming a cost of getting from i to j_μ implied by equation (94). Perhaps this is an argument in itself for preferring the gravity model to the intervening opportunities model.

3.6. *A distribution model of gravity type which uses intervening opportunities as a measure of cost*

As a final example of the application of the new method, consider the distribution model which is obtained if intervening opportunities are used as a measure of cost, but not weighted in the form (94). This is perhaps a more plausible assumption.

Note that j_μ, as originally defined, is properly a function of i and should be written as $j_\mu(i)$. The assumption now proposed is

$$c_{ij} = A_{j_\mu(i)} \tag{100}$$

This cost can now be substituted in either of the two gravity models derived above, the so-called conventional model described by equations (36)–(38), or the so-called single competition term model described by equations (47)–(48) with $f(c_{ij})$ as $\exp(-\beta c_{ij})$. Thus, substituting for c_{ij} from (100), the double competition term "gravity/opportunity" model is

$$T_{ij_\mu} = a_i b_{j_\mu} O_i D_{j_\mu} \exp(-\beta A_{j_\mu(i)}) \tag{101}$$

$$a_i = \left[\sum_{j_\mu} b_{j_\mu} D_{j_\mu} \exp(-\beta A_{j_\mu(i)})\right]^{-1} \tag{102}$$

$$b_{j_\mu} = \left[\sum_{i} a_i O_i \exp(-\beta A_{j_\mu(i)})\right]^{-1} \tag{103}$$

Small a's and b's are used for the balancing factors to avoid confusion with the A_{j_μ}'s.

The single competition term "gravity/opportunity" model is

$$T_{ij_\mu} = a_i O_i D_{j_\mu} \exp(-\beta A_{j_\mu(i)}) \tag{104}$$

$$a_i = \left[\sum_{j_\mu} D_{j_\mu} \exp(-\beta A_{j_\mu(i)})\right]^{-1} \tag{105}$$

If it could be argued that (100) represents a better account of cost than (94), then the models represented by equations (101)–(103) and (104)–(105) may give better answers than the conventional intervening opportunities model. A test of these new models would be welcomed.

4. CONCLUSIONS

A new statistical theory of spatial distribution models has been demonstrated. This gives a new method for constructing such models to meet a wide variety of circumstances. It appears to show that the gravity model has a more plausible theoretical base than the intervening opportunities model. But above all, the new method offers a technique for extending conventional spatial distribution models to cover new situations which now often present serious problems, especially in transportation study models: for example, modal split, especially in the multi-mode case, and the necessity of separating car owners and non-car owners, but allowing them to compete for the same attractions. Finally, a comparison of the gravity models and intervening opportunities model leads to a suggestion of a completely different type of model for quite simple and conventional situations.

REFERENCES

Chicago Area Transportation Study (1960). Final report, Vol. II.

Dieter K. H. (1962). Distribution of work trips in Toronto. *Proc. Am. Soc. civ. Engrs* **88**, 9–28.

Popper K. R. (1959). *The Logic of Scientific Discovery*. Hutchinson, London.

Quarmby D. A. (1967). Choice of travel mode for the journey to work: some findings. *J. Transp. Economics and Policy*. To be published.

Stouffler S. A. (1940). Intervening opportunities: A theory relating mobility and distance. *Am. Soc. Rev.* **5**, No. 6.

Tolman R. C. (1938). *Principles of Statistical Mechanics*. Clarendon Press, Oxford.