Accid. Anal. & Prev. Vol. 22, No. 5, pp. 487–498, 1990
Printed in Great Britain.

0001-4575/90 $3.00 + .00

A BIVARIATE NEGATIVE BINOMIAL MODEL TO EXPLAIN TRAFFIC ACCIDENT MIGRATION

M. J. MAHER
Traffic Safety Division, Transport and Road Research Laboratory, Crowthorne, Berkshire, RG11 6AU United Kingdom

(*Received 26 January* 1989; *in revised form* 23 *January* 1990)

Abstract—The phenomenon of "regression to the mean" is now widely known in the study of the effectiveness of remedial treatment of traffic accident blackspots. What happens is that the criterion used for selection of sites at which treatment is to be applied gives rise to bias in the estimate of the effectiveness: the conditional expectation of the after frequency is less than the true mean, even if the treatment is totally ineffective. It has been reported in some previous studies that accident "migration" has been observed. This is the phenomenon whereby the accident rate apparently rises at sites that are untreated but that are neighbours to treated sites. If this were a genuine effect, it would have serious implications for the assessment of remedial treatments. This paper aims to explain this migration effect in purely probabilistic terms, without recourse to the concept of physical migration. The model used is a new bivariate negative binomial distribution, incorporating spatial correlation between the true mean site accident rates. As with the regression to mean effect, the migration effect can then be explained in terms of the conditioning implicit in the selection process.

INTRODUCTION

In assessing the effectiveness of some remedial treatment in road safety studies, it is natural to compare the before and after accident frequencies, B_T and A_T, at the treated sites, and estimate the effectiveness θ by the statistic $(A_T - B_T)/B_T$. If there is the possibility of a shift in accident rate over time, unconnected with the treatment, control sites might well be used, with corresponding accident frequencies B_C and A_C. In this case θ would be estimated by the statistic $(A_TB_C - B_TA_C)/B_TA_C$. (It should be made clear at this point that the term "accident rate" will, in this paper, refer to the mean number of accidents *per unit time*). However, although such an approach is well founded in subjects in which experiments can be designed, in road accident studies this is not generally possible. The treated sites are not a random sample from the population of all sites, but have usually been selected because of their poor accident record in the before period. (In practice, however, the criterion for selection is often far from precise). As Hauer (1980a; 1980b) has pointed out so clearly, this bias in the selection procedure leads to a bias in the estimate used to assess the effectiveness of the treatment, a phenomenon known as "regression to the mean." Since *un*treated sites, too, may have been selected nonrandomly, even if only in a negative fashion, the statistic $(A_U - B_U)/B_U$ (which might naively be expected to be zero since no treatment has been applied) could also be biased—in the opposite direction to that for the treated sites. McGuigan (1985a) has referred to this as a "reverse regression to mean" effect. If the untreated sites were used as controls, the bias would, therefore, be even further increased.

Boyle and Wright (1984) have described work carried out on accident data from London boroughs. In addition to looking at the changes in accident rate at treated sites, they also looked at untreated sites that were next to treated sites, and referred to as "neighbours." They found that at these neighbouring sites, the accident rates had apparently *increased*. This was a disturbing finding, because, if this was a genuine effect, it would mean that the treatment was having the effect, not just of reducing accidents at the treated sites, but of spreading them to the neighbouring sites. Other work, too, has reported data which apparently display the same "migration" effect. Persaud (1987) has given some results from a study in which two-way STOP signs in Philadelphia were replaced by four-way STOP signs. Adams (1985) reports data on fatalities on the Baden-Wurttemburg autobahn; here, also, an apparent migration effect is displayed.

Boyle and Wright proposed a hypothesis to explain this migration of accidents from treated to neighbouring sites. This hypothesis was a behavioural mechanism, based on the idea of risk compensation. They hypothesised that "successful treatment of the blackspot will reduce the proportion of drivers leaving the blackspot who are behaving cautiously so that the number of accidents in the surrounding area will tend to increase towards its 'natural' level." McGuigan (1985a) argued that this migration effect could be explained purely in terms of reverse regression to mean, since the neighbours were untreated. However, Boyle and Wright (1985) pointed out that the size of this effect depended on P_T, the proportion of sites that are treated, and that for realistic values of P_T, the magnitude of the reverse regression to mean was considerably less than the 10% observed migration effect in the London data. McGuigan (1985b) was forced to concede that the support for his argument "became less cogent."

Maher (1987) proposed a new, purely probabilistic explanation for accident migration. This was based on the spatial correlation between the true accident rates at nearby, or adjacent sites in the road network. Since it has been well established by, for example, Maycock and Hall (1984), Hall (1987), and Turner and Thomas (1986) amongst others, that the accident rate at a site is highly dependent on flow at the site, and because the flow through the network is essentially continuous, it follows that accident rates at neighbouring sites will tend to be positively correlated. Maher used simulation to illustrate how this spatial correlation could produce a migration effect appreciably larger in magnitude than the simple reverse regression to mean of McGuigan. This gave strength, therefore, to the belief that accident migration could be explained purely in probabilistic terms, without necessarily any resort to a behavioural mechanism. The simulation approach, however, has its limitations, in that it is largely illustrative and does not produce any formulae for predicting the size of the migration effect. The purpose of this present paper is to give an analytical treatment of the problem, confirming the earlier simulation tests and resulting in formulae and a computer programme from which the size of the expected migration effect can be calculated for any values of the relevant parameters. To aid understanding of the process, and to summarise the results, graphs are presented.

REGRESSION TO THE MEAN

It has been widely assumed in previous work that the number of accidents Y at a site is Poisson distributed with a mean X, and that X varies from site to site, dependent on flow through the site and site characteristics including geometry and other explanatory variables. It has become standard to take the distribution of X to be gamma, denoted by $G(\alpha + \beta, \lambda)$ where the mean μ is $(\alpha + \beta)/\lambda$. $(\alpha + \beta)$ is the shape parameter (the reason for this style of parameterisation will appear later). The unconditional distribution of Y is then negative binomial with a probability function (pf) of, say, $Q(y)$:

$$Q(y) = \frac{\Gamma(\alpha + \beta + y)}{\Gamma(\alpha + \beta)y!}\left[\frac{\lambda}{\lambda + 1}\right]^{\alpha+\beta}\left[\frac{1}{\lambda + 1}\right]^{y} \quad (y = 0,1,2, \ldots) \tag{1}$$

Suppose that a site is selected for treatment if its observed accident frequency in the before period is k or more. The probability of selection is

$$P_T = \sum_{k}^{\infty} Q(y)$$

and the conditional observed before frequency is

$$B_T = E(Y|Y \geq k) = \sum_{k}^{\infty} y\, Q(y)/P_T. \tag{2}$$

It can be shown (see, for example, Robbins (1980)) that

$$E(X|Y = y) = (y + 1)\ Q(y + 1)/Q(y) \tag{3}$$

from which it follows that, if the treatment is ineffective (i.e. $\theta = 1$), the expected after frequency is

$$A_T = E(X|Y \geq k) = \sum_{k+1}^{\infty} y\ Q(y)/P_T, \tag{4}$$

and the regression to mean effect is found from (2) and (4) to be:

$$P_{reg} = (A_T - B_T)/B_T = -kQ(k) \Big/ \left(\mu - \sum_{0}^{k-1} y\ Q(y)\right). \tag{5}$$

Similarly, for untreated sites, the "reverse" regression to mean effect is

$$P_{unt} = (A_U - B_U)/B_U = k\ Q(k) \Big/ \sum_{0}^{k-1} y\ Q(y). \tag{6}$$

If μ and $(\alpha + \beta)$ were both 2 and k were 10, the proportion of sites treated P_T, would be .59%, the apparent effect of the treatment, P_{reg}, would be -41% and the apparent increase in accident rate at the untreated sites, P_{unt}, would be 1.39%, considerably less than the 10% value observed by Boyle and Wright for the migration effect in the London data. For reasonable values of the parameters, Boyle and Wright's (1985) claim that the reverse regression to mean argument was not sufficient to explain the observed level of accident migration, is confirmed.

Migration

Consider, now, two sites in the network. Eventually, we shall be considering these as a pair of neighbouring sites: one selected for treatment ($Y \geq k$) and the other not selected ($Y < k$), but for the moment they are simply any two sites. The assumptions are identical to those made above:

$$\left.\begin{array}{l} Y_i \text{ is Poisson with mean } X_i \\ \text{and} \\ X_i \text{ is gamma } G(\alpha + \beta, \lambda) \text{ with mean } \mu\ (= (\alpha + \beta)/\lambda). \end{array}\right\} \quad (i = 1, 2)$$

Therefore, the marginal distribution of Y_1 (and, by symmetry, Y_2) is again negative binomial, with pf $Q(y)$ defined as in (1). However, we now take X_1 and X_2 to be constructed as the sum of other, mutually independent gamma distributed random variables,

$$X_1 = U + V \quad \text{and} \quad X_2 = U + W,$$

in which U is $G(\alpha, \lambda)$ and both V and W are $G(\beta, \lambda)$. Because U is common to both, X_1 and X_2 are positively correlated with a coefficient of correlation $\rho = \alpha/(\alpha + \beta)$. Y_1 and Y_2 are also positively correlated (with coefficient of correlation $\alpha/((\alpha + \beta)(\lambda + 1))$—smaller by a factor $(\lambda + 1)$ than that of the X_i), and it is their bivariate pf, $P(y_1, y_2)$, which we shall be developing later and which is the key to all the results we obtain. For the present, we note that, since the marginal distributions of the Y_i are negative binomial it seems natural to refer to the distribution of (Y_1, Y_2) as a *bivariate* negative binomial. However, it should be pointed out that this is a different distribution from that first described by Bates and Neyman (1952) and recently used by Senn and Collie (1988). There, the X_i were *perfectly* correlated, with $X_2 = aX_1$. Here the degree of correlation can be varied between 0 and 1 by varying the relative sizes of α and β; if α is zero, Y_1 and Y_2 are independent and $P(y_1, y_2)$ is $Q(y_1)Q(y_2)$, whereas if β is zero the distribution becomes virtually that of Bates and Neyman (apart from the growth factor a). In fact, with a minor modification this new model can be generalised further, so that the Bates and Neyman model is a special case of it. However, for the present purpose, the model will suffice in the form defined above.

From the assumptions about the Y_i and X_i, it follows that the bivariate pf for the Y_i can be written:

$$P(y_1, y_2) = \iiint f_u(u)f_v(v)f_w(w)\, e^{-(u+v)} \frac{(u + v)^{y_1}}{y_1!}\, e^{-(u+w)} \frac{(u + w)^{y_2}}{y_2!}\, dudvdw$$

where $f_U(u)$, $f_V(v)$ and $f_W(w)$ are the probability density functions of the three gamma variables U, V, and W. It follows quite simply from this that the expected value of X_2 $(= U + W)$ given that $Y_1 = y_1$ and $Y_2 = y_2$ is:

$$\iiint f_u(u)f_v(v)f_w(w)\, e^{-(u+v)} \frac{(u + v)^{y_1}}{y_1!}\, e^{-(u+w)} \frac{(u + w)^{y_2+1}}{y_2!}\, dudvdw/P(y_1, y_2)$$
$$= (y_2 + 1)\, P(y_1, y_2 + 1)/P(y_1, y_2), \quad (7)$$

which is a similar result to that obtained in (3) for the univariate case. Now suppose that site 1 is a treated site and site 2 is an untreated neighbour (i.e. $y_1 \geq k$ and $y_2 < k$). Denoting these conditions by C, the probability of them is:

$$P_C(k) = \sum_{k}^{\infty} \sum_{0}^{k-1} P(y_1, y_2).$$

Again assuming that any treatment is ineffective (or simply that no treatment was actually applied), the conditional expected frequencies at the neighbour, before and after, are:

$$B_N = E(Y_2|C) = \sum_{k}^{\infty} \sum_{0}^{k-1} y_2\, P(y_1, y_2)/P_C(k) \quad (8)$$

and

$$A_N = E(X_2|C) = \sum_{k}^{\infty} \sum_{0}^{k-1} (y_2 + 1)\, P(y_1, y_2 + 1)/P_C(k)$$
$$= \sum_{k}^{\infty} \sum_{1}^{k} y_2\, P(y_1, y_2)/P_C(k) \quad (9)$$

The size of the apparent migration effect is then $(A_N - B_N)/B_N$ which, after a little rearrangement of the limits of the summations involved (so as to use only finite sums) is:

$$P_{mig} = \frac{k\, Q(k) - \sum_{0}^{k-1} k\, P(y_1, k)}{\sum_{0}^{k-1} y\, Q(y) - \sum_{0}^{k-1} \sum_{0}^{k-1} y_1\, P(y_1, y_2)} \quad (10)$$

Therefore, once $P(y_1, y_2)$ has been found, the finite sums in the above expression can be evaluated (along with those involving $Q(y)$) for any desired values of the parameters μ, $(\alpha + \beta)$, ρ and k, and thence P_{mig}, P_{reg} and P_{unt} calculated.

It is clear that as $\alpha \to 0$ (i.e. as $\rho \to 0$) $P_{mig} \to P_{unt}$. That is, with *no* spatial correlation, the migration effect reduces simply to the reverse regression to mean effect.

Numerical evaluation

The derivation of the bivariate pf is given in the appendix, together with simple recursive formulae which enable the values of the $P(y_1, y_2)$ to be determined in an accurate and straightforward manner. A FORTRAN computer program has been written to implement these relations and hence to calculate, for any chosen values of μ, $(\alpha + \beta)$ and ρ, the bivariate pf $P(y_1, y_2)$ and, using (1), (5), (6), and (10) the values of P_T, P_{reg}, P_{unt} and (of most relevance here) P_{mig} for a range of values of k. For example, if $\mu = 2$, $(\alpha + \beta) = 2$ and $\rho = 0.5$, a section of the output from the programme would look like:

k	P_T (%)	P_{reg} (%)	P_{unt} (%)	P_{mig} (%)
6	6.25	−36.2	10.61	20.60
7	3.52	−37.8	6.39	15.22
8	1.95	−39.1	3.86	11.57
9	1.07	−40.2	2.32	9.00
10	0.586	−41.0	1.39	7.11
11	0.318	−41.8	0.822	5.69
12	0.171	−42.4	0.481	4.60

From this table we can see that as k increases, the proportion of sites which would be selected for treatment, P_T, decreases as do both P_{unt} and P_{mig}. P_{unt}, it will be remembered, is the reverse regression to mean effect, proposed initially by McGuigan (1985a) as an explanation of the apparent migration of accidents noted by Boyle and Wright (1984) in the London data. The mechanism proposed here contains a spatial correlation component in addition to that of reverse regression to mean, and so the value of P_{mig} is inevitably greater than P_{unt} for any given k (if $\rho > 0$). More importantly, however, the *relative* magnitude of P_{mig} to P_{unt} *increases* as P_T decreases, from a ratio of around 2 when k is 6 to one of almost 10 when k is 12. Therefore, at realistically low values of P_T, the spatial correlation mechanism proposed here is a far stronger contender than is the reverse regression to mean mechanism alone, in attempting to provide a plausible statistical explanation of the accident migration phenomenon. The next section will explore more fully the dependence of P_{mig} on μ, $(\alpha + \beta)$ and ρ.

Results

It is clear from what has been said already that P_{mig} decreases as ρ decreases, reaching P_{unt} as ρ reaches zero (accident rates at different sites then being independent). The manner in which it reduces is exemplified in Fig. 1. There, the lower of the two curves is a plot of P_{mig} against ρ for the case $\mu = 2$, $\alpha + \beta = 2$ and $k = 10$: it can be seen that the relationship is virtually linear. So too is the upper curve, in which instead of k being kept constant at a value of 10, linear interpolation is used to estimate, for each ρ value, the value of P_{mig} which would correspond to $P_T = 1\%$.

The same interpolation method is used to draw Figs. 2 and 3. In Fig. 2, P_{mig} is plotted against μ for a number of values of the shape parameter $\alpha + \beta$ for $P_T = 1\%$ and $\rho = 1$; Fig. 3 shows the same but for $P_T = .1\%$. It is evident from these plots that the shape parameter has an appreciable influence on P_{mig}. In figures 4 and 5 are similar plots but of P_{unt} instead of P_{mig}, and here the shape parameter has much less influence, as can be seen from the close clustering of the curves. The circumstances which give rise to a large value for the apparent migration of accidents from a treated site to a neighbour, even when the treatment is ineffective, are: high ρ (spatial correlation) and low $\alpha + \beta$ (high between site variation in true accident rate). The effect of μ is not uniform, as can be seen from figures 2 and 3. The apparent migration effect P_{mig} is still large for small values of P_T, whereas P_{unt} declines markedly as P_T diminishes. For example, when

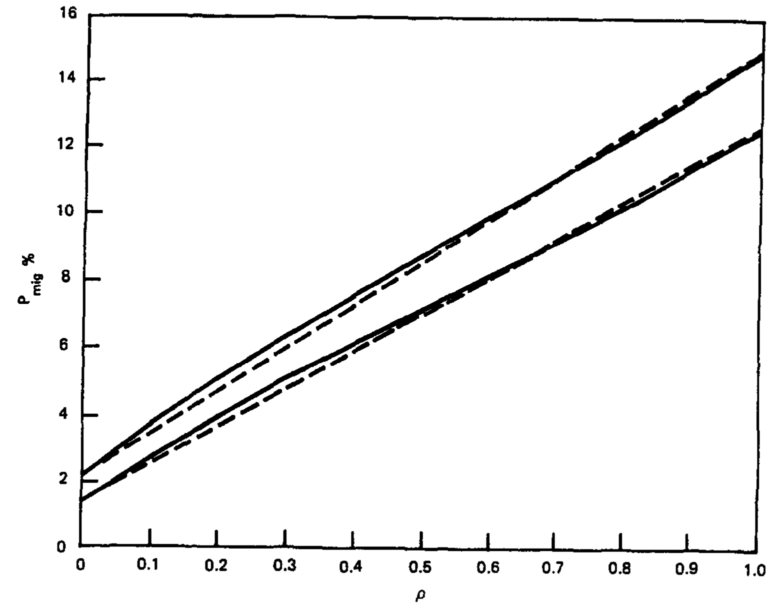

Fig. 1. P_{mig} (%) against ρ, when $\mu = 2$ and $\alpha + \beta = 2$, for (i) $k = 10$ (lower curve) and (ii) k chosen such that $P_T = 1\%$ (upper curve).

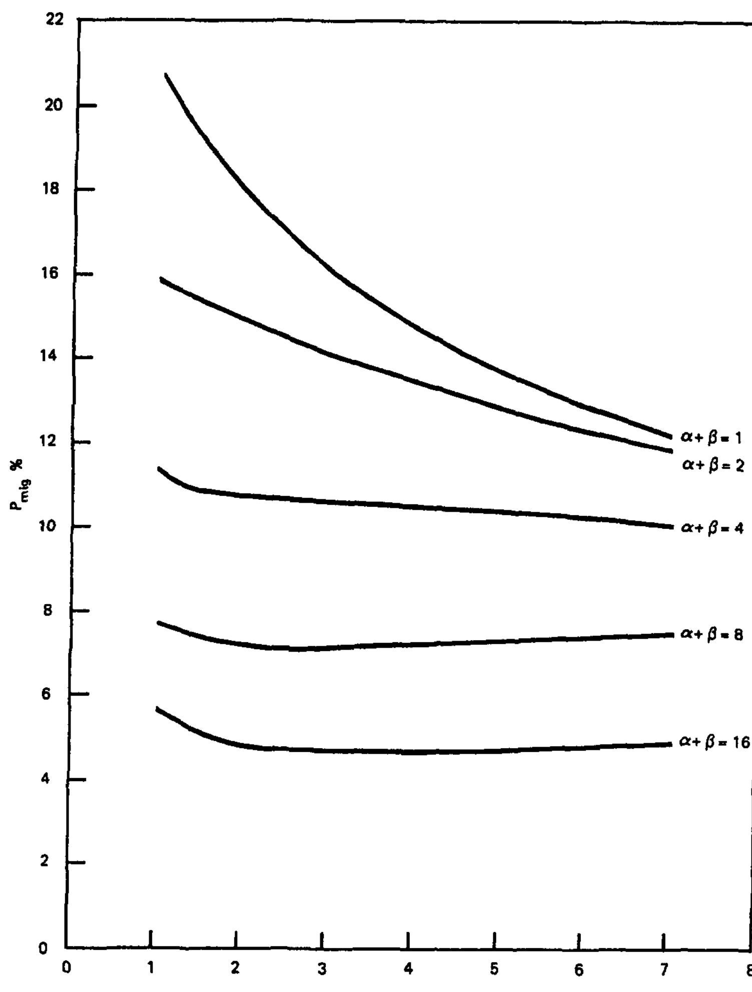

Fig. 2. P_{mig} (%) against μ, when $\rho = 1$ and $P_T = 1\%$, for (from top to bottom) $\alpha + \beta = 1, 2, 4, 8,$ and 16.

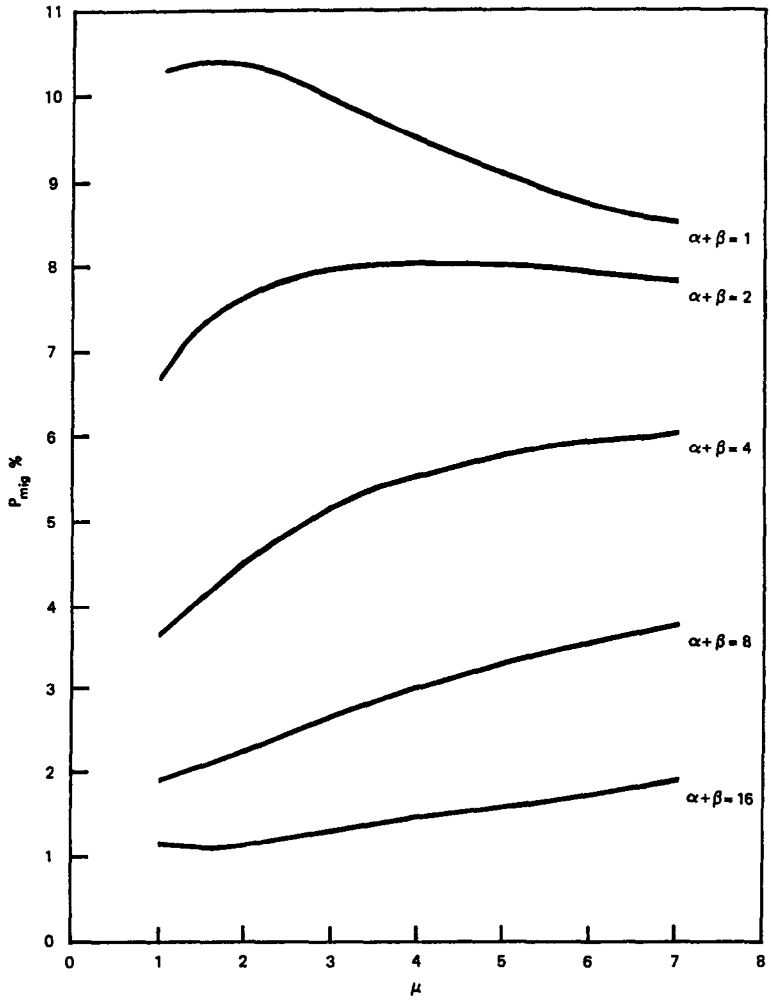

Fig. 3. P_{mig} (%) against μ, when $\rho = 1$ and $P_T = .1\%$, for (from top to bottom) $\alpha + \beta = 1, 2, 4, 8,$ and 16.

$\rho = 0.6$, $\mu = 3$, $\alpha + \beta = 1$ and $k = 16$, the proportion of sites treated is 1%, $P_{mig} = 9.95\%$, $P_{unt} = 1.4\%$ and $P_{reg} = -21.1\%$, this last being the size of the regression to mean effect observed by Boyle and Wright. Whilst this does certainly not, in any sense, *prove* that accident migration is a statistical artifact, it does, at least, demonstrate that with reasonable parameter values the probabilistic model *can* produce results which match those observed.

Model extensions

Any mathematical model is, of course, a simplification of reality, and the one considered here is no exception. An attempt has been made to build into the model certain features that were lacking in other models and that could contribute to an understanding of the phenomenon of apparent migration. The most important of these is the spatial correlation between the true site accident rates. This correlation is expected to occur chiefly through the mechanism of accident-flow relationships coupled with the continuity of flow through the network. It could also be contributed to, however, by the incorrect coding of the location of accidents in the network. For example, if at each of two neighbouring sites just 10% of accidents were wrongly allocated to the other site, the correlation between the two mean accident rates would be 0.22 even if the two true accident rates *were independent*.

The criterion for selection of a site for treatment is, in practice, not always a clearly defined one. It is influenced, but not completely determined, by the number of accidents

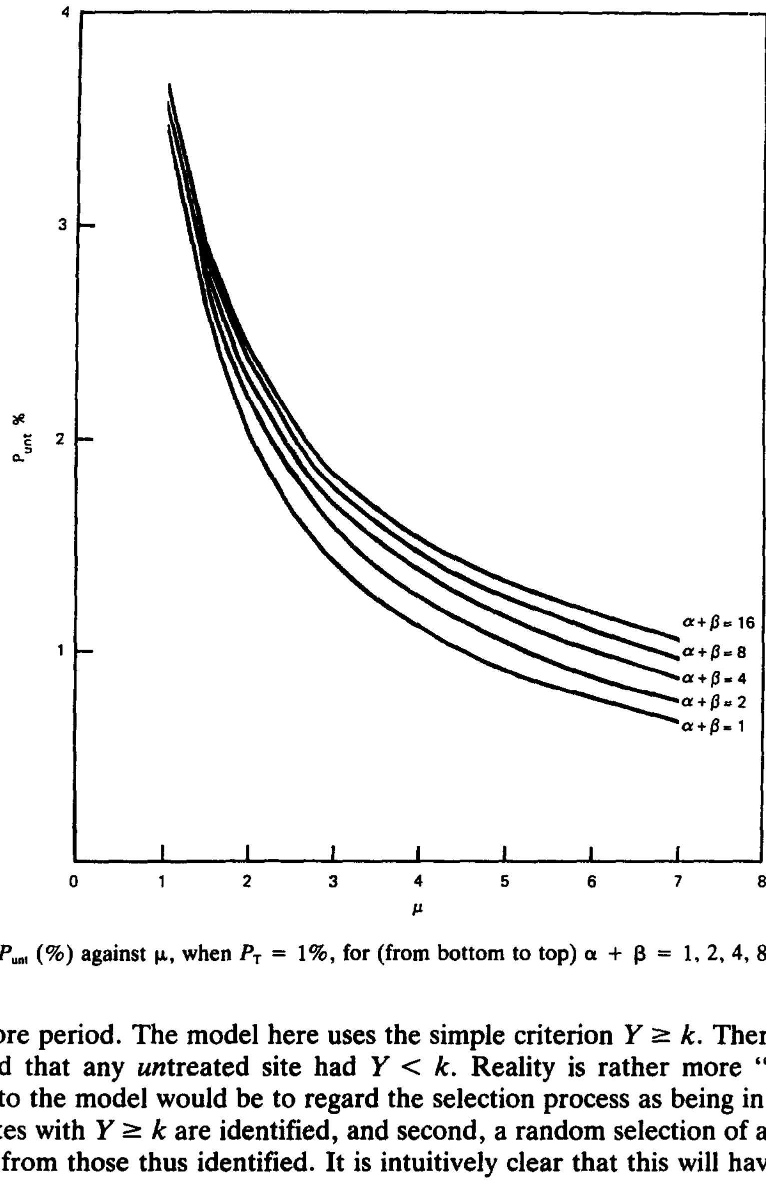

Fig. 4. P_{unt} (%) against μ, when $P_T = 1\%$, for (from bottom to top) $\alpha + \beta = 1, 2, 4, 8$, and 16.

in the before period. The model here uses the simple criterion $Y \geq k$. Therefore it can be inferred that any *un*treated site had $Y < k$. Reality is rather more "fuzzy". An extension to the model would be to regard the selection process as being in two stages: first, all sites with $Y \geq k$ are identified, and second, a random selection of a proportion p is made from those thus identified. It is intuitively clear that this will have the effect of diluting the apparent migration effect; that is, of diminishing P_{mig}. It will also diminish P_{unt}, but P_{reg} will remain unaffected for any given k. The smaller p becomes, the more the process will approach that of a random selection of sites for treatment.

The main effect this modification has on the development of the model is that it can no longer be taken that an untreated neighbour has $Y < k$. It could also be a site for which $Y \geq k$ but which was not selected in the second stage. It can be shown that the effects on the expressions for P_{unt} and P_{mig} in equations (6) and (10) are the addition in each case of an extra term in the denominator. For P_{unt} this term is simply $\mu(1 - p)/p$. For P_{mig} it is:

$$\frac{1-p}{p}\left[\mu\left(1 - \sum_{0}^{k-1} Q(y)\right) + \frac{\rho}{\lambda + 1}\left(\sum_{0}^{k-1} Q(y) - \sum_{0}^{k-1} y\, Q(y)\right)\right]. \tag{11}$$

As before, when $\rho = 0$, $P_{mig} = P_{unt}$. This extension to the model has been incorporated into the FORTRAN program so that the effect of this two stage selection process can be quantified. To give an example: suppose that $\rho = 1$, $\mu = 2$, $\alpha + \beta = 2$ and $p = 1$. Then if k is such that 1% of sites are treated, $P_{mig} = 15\%$. However, if the value taken by p is altered to .75, .5 and .25 in turn (whilst adjusting the value of k as necessary to maintain P_T at 1%), the values of P_{mig} are respectively 10.7%, 7% and 3.6%. Not

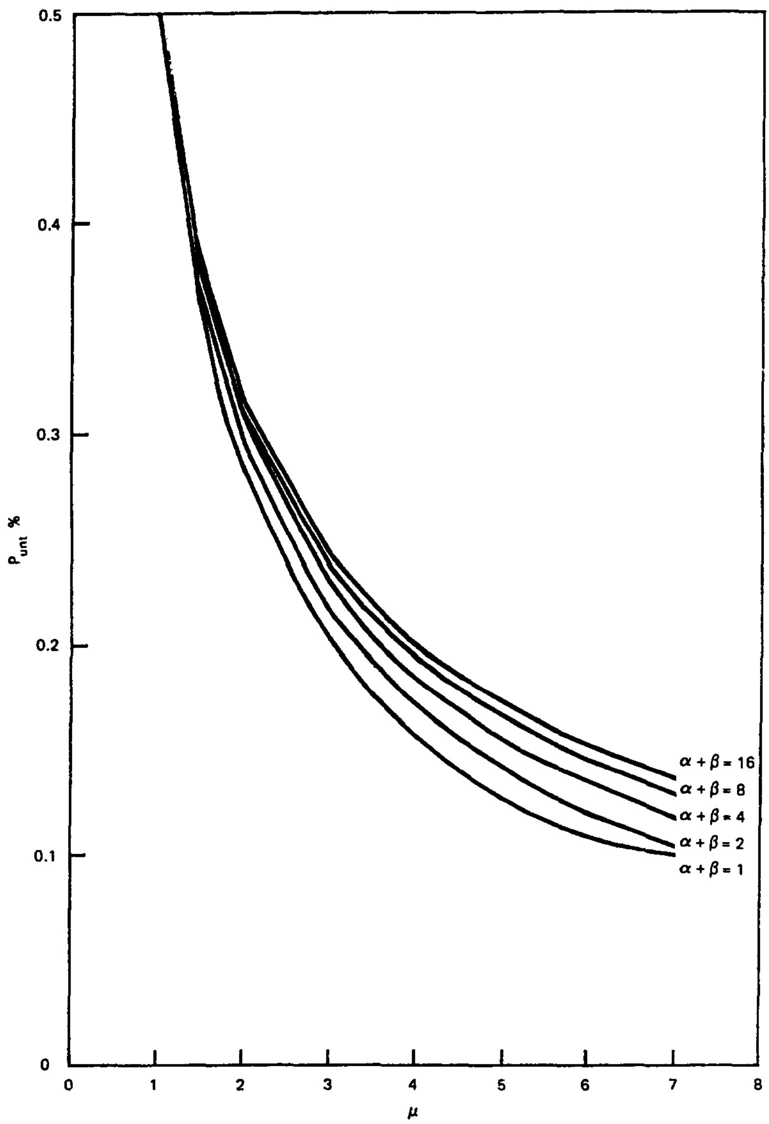

Fig. 5. P_{unt} (%) against μ, when P_T = .1%, for (from bottom to top) $\alpha + \beta$ = 1, 2, 4, 8, and 16.

surprisingly, then, the nature of the selection process has a profound influence on the size of the apparent migration effect.

DISCUSSION

The method of selection of sites at which remedial treatment is to be applied gives rise to the now widely known regression to mean effect. This is due to the conditioning, such as $Y \geq k$, used in the selection process. This bias can work in reverse for the unselected sites (where $Y < k$), even when the X_i, the true site mean frequencies, are independent. This reverse regression to mean effect gives an apparent migration of accidents from treated to untreated sites, even when the treatment is ineffective. However, the magnitude of this effect is not sufficient in practice to provide a satisfactory explanation for the migration observed by Boyle and Wright.

The introduction of spatial correlation between the X_i into the model increases the expected size of the migration effect considerably from that predicted by the reverse regression to mean argument alone. The neighbour is now not only an untreated site but also adjacent to a treated site, and so its true mean will tend to be higher than would be expected under the independence model. A computer program has been written to implement the probabilistic model developed here and to output results for P_T, P_{mig}, P_{unt} and P_{reg} for any combination of values of ρ, $\alpha + \beta$, and μ. An extension to the model allows for the case in which the selection of sites for treatment can be thought of as a two stage process: of those sites for which $Y \geq k$, a proportion of p are randomly selected.

The established relationships between accident frequency and traffic flow and the continuity of flows within the road network provide a compelling argument for the existence of the spatial correlations assumed here. The extent of such correlations is as yet undetermined; there will inevitably be some characteristic structure within the road network, because such networks are by nature hierarchical, and flow continuity will have most impact on the correlation between accident frequencies at the *main junctions* and less between main and minor junctions (e.g. minor road accesses carrying low flows and with small accident frequencies). Some more work is needed to estimate the size of the correlations. However, it is clear that the numerical properties of the model developed here show that it could provide a satisfactory and feasible alternative to the behavioural explanation for accident migration hypothesised by Boyle and Wright.

The bivariate negative binomial model developed in this paper could have other useful applications. Bates and Neyman described *their* bivariate model as an "optimistic" one, because it assumed that the after accident rate (μ_2 in their model) was directly proportional to (and therefore perfectly correlated with) the before rate (i.e. $\mu_2 = a\mu_1$). The new model presented here requires more calculation, but with modern computing power (vastly increased since 1952 when Bates and Neyman's work was carried out), this is of no great significance, and the benefits to be obtained from this more general form of bivariate negative binomial distribution could well justify its use. Its application need not be limited to before and after data, but could be extended to, for example, data on two types of accident (slight and severe or nighttime and daytime) or data on accidents and conflicts counted at sites.

Acknowledgements—The work described in this paper forms part of the programme of the Transport and Road Research Laboratory and the paper is published by permission of the Director.

REFERENCES

Adams, J. Risk and freedom: The record of road safety regulation. Cardiff, U.K.: Transport Publishing Projects: 1985.
Bates, G. E.; Neyman, J. Contributions to the theory of accident proneness: I. An optimistic model of the correlation between light and severe accidents. Univ. Calif. Publ. Statistics 1(9): 215–254; 1952.
Boyle, A. J.; Wright, C. C. Accident "migration" after remedial treatment at accident blackspots. Traffic Eng. Control 25:260–267; 1984.
Boyle, A. J.; Wright, C. C. Letter to the editor. Traffic Eng. Control 26:389; 1985.
Feller, W. An Introduction to probability theory and its application. Volume 1 (2nd ed). New York: Wiley International; 1957.
Hall, R. D. Accidents at four-arm single carriageway urban traffic signals. Contractor report 65. Crowthorne, Berkshire, U.K.: Transport and Road Research Laboratory; 1987.
Hauer, E. Selection for treatment as a source of bias in before-and-after studies. Traffic Eng. and Control, 21(8/9):419–421; 1980a.
Hauer, E. Bias by selection: Overestimation of the effectiveness of safety countermeasures caused by the process of selection for treatment. Accid. Anal. Prev. 15:323–328; 1980b.
McGuigan, D. R. D. Accident "migration"—or a flight of fancy? Traffic Eng. Control 26:229–233; 1985a.
McGuigan, D. R. D. Letter to the editor. Traffic Eng. Control 26:390–391; 1985b.
Maher, M. J. Accident migration—a statistical explanation? Traffic Eng. Control 28:480–483; 1987.
Maycock, G.; Hall, R. D. Accidents at four-arm roundabouts. Laboratory report 1120. Crowthorne, Berkshire, U.K. Transport and Road Research Laboratory; 1984.
Persaud, B. "Migration" of accident risk after remedial blackspot treatment. Traffic Eng. Control 28:23–26; 1987.
Robbins, H. Some estimation problems for the compound Poisson distribution. In: asymptotic theory of statistical tests and estimation. New York: Academic Press; 251–257; 1980.
Senn, S. J.; Collie, G. S. Accident blackspots and the bivariate negative binomial. Traffic Eng. Control 29:168–169; 1988.
Turner, D. J.; Thomas, R. Motorway accidents: An examination of accident totals, rates and severity and their relationship with traffic flow. Traffic Eng. Control 26:377–383; 1986.

APPENDIX

Derivation of bivariate pf

If Y is Poisson distributed with mean μ, its probability generating function (pgf) is (see, for example, Feller (1957)):

$$G(s) = E(s^Y) = \exp(\mu(s - 1))$$

If X is gamma $G(\alpha, \lambda)$ distributed, its moment generating function (mgf) is:

$$M(s) = E(\exp(sX)) = \left[1 - \frac{s}{\lambda}\right]^{-\alpha} \tag{A1}$$

The bivariate pgf of Y_1 and Y_2 is:

$$\begin{aligned} G(s_1, s_2) &= E(s_1^{Y_1} s_2^{Y_2}) \\ &= E_X[\exp(X_1(s_1 - 1)) \exp(X_2(s_2 - 1))] \\ &= E_{U,V,W}[\exp((U + V)(s_1 - 1)) \exp((U + W)(s_2 - 1))] \\ &= E_U[\exp((s_1 + s_2 - 2)U)]\, E_V[\exp((s_1 - 1)V)]\, E_W[\exp((s_2 - 1)W)] \\ &= M_U(s_1 + s_2 - 2)\, M_V(s_1 - 1)\, M_W(s_2 - 1) \end{aligned}$$

where M_U, M_V, and M_W are the mgfs of U, V, and W. Using (A1), therefore, it follows that $G(s_1, s_2)$ can be written as:

$$\left[1 - \frac{s_1 + s_2 - 2}{\lambda}\right]^{-\alpha} \left[1 - \frac{s_1 - 1}{\lambda}\right]^{-\beta} \left[1 - \frac{s_2 - 1}{\lambda}\right]^{-\beta} \tag{A2}$$

Since the pgf is defined as $G(s_1, s_2) = \Sigma_{y_1} \Sigma_{y_2} s_1^{y_1} s_2^{y_2} P(y_1, y_2)$, expanding (A2) and identifying coefficients of $s_1^{y_1} s_2^{y_2}$ with $P(y_1, y_2)$ enables the bivariate pf to be obtained. After some algebra, we obtain the result:

$$P(y_1, y_2) = S(y_1, y_2) \sum_0^{y_1} \sum_0^{y_2} T(i, j) \tag{A3}$$

where

$$S(y_1, y_2) = \frac{\Gamma(\beta + y_1)\, \Gamma(\beta + y_2)}{\Gamma(\beta) y_1!\, \Gamma(\beta) y_2!} \left[\frac{\lambda}{\lambda + 2}\right]^{\alpha} \left[\frac{\lambda}{\lambda + 1}\right]^{2\beta} \left[\frac{1}{\lambda + 1}\right]^{y_1 + y_2}, \tag{A4}$$

and

$$T(i, j) = \binom{y_1}{i} \binom{y_2}{j} \frac{\Gamma(\alpha + i + j)\, \Gamma(\beta + y_1 - i)\, \Gamma(\beta + y_2 - j)}{\Gamma(\alpha)\;\; \Gamma(\beta + y_1)\;\; \Gamma(\beta + y_2)} \left[\frac{\lambda + 1}{\lambda + 2}\right]^{i+j}. \tag{A5}$$

Although these expressions appear rather unwieldy, quite simple recursive formulae can be obtained from them:

$$S(y_1, y_2) = S(y_1 - 1, y_2) \left[\frac{\beta + y_1 - 1}{y_1}\right] \left[\frac{1}{\lambda + 1}\right] \quad (y_1 \geq 1), \tag{A6}$$

and

$$T(i, j) = T(i - 1, j) \left[\frac{\lambda + 1}{\lambda + 2}\right] \left[\frac{y_1 - i + 1}{i}\right] \left[\frac{\alpha + i + j - 1}{\beta + y_1 - i}\right] (i \geq 1). \tag{A7}$$

Similar relations hold between $S(y_1, y_2)$ and $Sy_1, y_2 - 1)$ and between $T(i, j)$ and $T(i, j - 1)$, and are obtained by reversing the roles of y_1 and y_2 and of i and j. The $S(y_1, y_2)$ and $T(i, j)$ can then be calculated recursively in the order (0,0), (1,0), (2,0), . . . (1,1), (2,1), . . . etc. To start the recursion, we have:

$$S(0, 0) = \left[\frac{\lambda}{\lambda + 2}\right]^{\alpha} \left[\frac{\lambda}{\lambda + 1}\right]^{2\beta} \text{ and } T(0, 0) = 1$$

Regression of Y_2 on Y_1

To find the conditional expectation of Y_2 given that $Y_1 = k$, note first that in the univariate case, from (1) and (3):

$$E(X|Y = k) = (k + 1)Q(k + 1)/Q(k) = (\alpha + \beta + k)/(\lambda + 1). \tag{A8}$$

It follows, therefore, in the bivariate case, that:

$$E(X_1|Y_1 = k) = E(U + V|Y_1 = k) = (\alpha + \beta + k)/(\lambda + 1). \tag{A9}$$

Now, to find the regression of Y_2 on Y_1, we can write:

$$E(Y_2|Y_1 = k) = E(X_2|Y_1 = k) = E(U|Y_1 = k) + E(W|Y_1 = k) \tag{A10}$$

Since U and V are independent gammas, $U + V$ is gamma and $U/(U + V)$ is beta, and the two are independent. Therefore,

$$E(U|Y_1 = k) = E\left(\frac{U}{U + V}\right) E(U + V|Y_1 = k) = \frac{\alpha}{\alpha + \beta} E(U + V|Y_1 = k). \tag{A11}$$

Since $Y_1 = k$ sheds no light on W, $E(W|Y_1 = k) = \beta/\lambda$ and so, from (A9), (A10) and (A11):

$$E(Y_2|Y_1 = k) = \frac{\alpha}{\alpha + \beta}\frac{\alpha + \beta + k}{\lambda + 1} + \frac{\beta}{\lambda} = (1 - \rho)\mu + \rho\left[\mu\frac{\lambda}{\lambda + 1} + \frac{k}{\lambda + 1}\right]. \tag{A12}$$