Mathematical Social Sciences 80 (2016) 108–114

Contents lists available at ScienceDirect

Mathematical Social Sciences

journal homepage: www.elsevier.com/locate/econbase

Box model of migration channels

Nikolay K. Vitanov[a,b,*], Kaloyan N. Vitanov[a]

[a] Institute of Mechanics, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Bl. 4, 1113 Sofia, Bulgaria

[b] Max-Planck Institute for the Physics of Complex Systems, Nöthnitzerstr. 36, 01187, Dresden, Germany

HIGHLIGHTS

- Model of migration channels is proposed.
- The model is based on the truncated Waring distribution.
- Large number of these migrants remain in the first country of the channel.
- Large part of migrants concentrate in an attractive final destination country.

ARTICLE INFO

Article history:
Received 20 October 2015
Received in revised form
11 December 2015
Accepted 6 February 2016
Available online 23 February 2016

ABSTRACT

We discuss a mathematical model of migration channel based on the truncated Waring distribution. The truncated Waring distribution is obtained for a more general model of motion of substance through a channel containing finite number of boxes. The model is applied then for case of migrants moving through a channel consisting of finite number of countries or cities. The number of migrants in the channel strongly depends on the number of migrants that enter the channel through the country of entrance. It is shown that if the final destination country is very popular then large percentage of migrants may concentrate there.

1. Introduction

Population migration involves the relocation of individuals, households or moving groups between geographical locations. Much efforts are directed to the study of internal migration in order to understand this migration and to make projection of the internal migration flows that may be very important for taking decisions about economic development of regions of a country (Armitage, 1986; Bracken and Bates, 1983; Champion et al., 2002). In addition the study of migration becomes very actual after the large migration flows directed to Europe in September 2015. From the point of view of migrating units migration models may be classified as macromodels or micromodels (Stillwell and Congdon, 1991; Cadwallader, 1989, 1992). Micromodels are based on the individual migrating unit (person, group or household) and on the processes underlying the decision of the migrant to remain in the current location or to move somewhere else (Maier and Weiss, 1991). Macromodels are for the aggregate migration flows. Many macromodels are explanatory and they use non-demographic information. An important class of these models are the gravity models having their roots in the ideas of Ravenstein about the importance of distance for migration and their first applications in the study of intercity migration by Zipf (Ravenstein, 1885; Zipf, 1946). Further development of these models was made by the concept of spatial interactions (Wilson, 1970; Stillwell, 1978; Fotheringham et al., 2001) or by introduction of statistical spatial interaction models (Congdon, 1991; Flowerdew, 1991; Nakaya, 2001; Fotheringham et al., 2002).

Other kinds of macromodels use demographic information from the field of multi-state demography in order to generate projections of migration flows. In the first models of this kind one used a cohort component model which involved the estimation of the population in the studied region at the beginning of a projection period. Then a projection of the number of births during the future time period and the survival of those in existence or being born during the period was made (Bowley, 1924; Whelpton, 1936). In the course of the time these models become multi-regional stochastic models and the requirement to the used information was changed: the first models required little information about migration and the more sophisticated models require maximum information about migration, e.g. a migration

* Corresponding author at: Institute of Mechanics, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Bl. 4, 1113 Sofia, Bulgaria.
E-mail address: vitanov@imbm.bas.bg (N.K. Vitanov).

http://dx.doi.org/10.1016/j.mathsocsci.2016.02.001

flow information disaggregated by single year of age and sex (Rees and Wilson, 1977; Rogers, 1990, 1995).

Migration models may be classified as probability models or deterministic models with respect to their mathematical features. Many probability models of migration (Willekens, 1999, 2008) are focused on the change of address in the process of migration involving the crossing of an administrative boundary. The migration data can be connected to the events of migration (event data or movement data) or to the place of residence (migrant status) at a given point of time (status data). Interesting kind of data is the duration data where the time is measured as a duration of since a reference event (event-origin that may be associated with the start of the process of migration). The observation of place of residence of an individual at two points in time leads to collection of transition data (Ledent, 1980; Willekens, 1999). Data obtained by recording the migrant status at several points of time are called panel data.

Different kinds of data are connected to different probability models. The *duration data* may be used in the *exponential model of migration* (Willekens, 1999). This model considers two states—origin state and a destination state. It is assumed that at the onset of migration all individuals are in the origin state. Individuals may leave the origin state for different reason but only the leaving connected to migration is considered. And the event of migration is assumed to be experienced by an individual only once (non-repeatable event). Let the size of population (of identical individuals) be m and let T be the time at which an individual migrates ($T = 0$ denotes the birth of the individual). One can define probability distribution $F(t) = P(T < t)$ ($T < t$ means that the migration event happens between $T = 0$ and $T = t$). $f(t)$ is the probability density function connected to $F(t)$ and $S(t) = 1 - F(t)$ is called survival function. It is connected to the size $mS(t)$ of the risk set of individuals (those who have not migrated but are exposed to the risk of migration in the future). One can introduce the conditional density function $\mu(t) = f(t)/S(t)$ which represents probability that the migration event occurs in a small interval following t provided that it was not occurred before t. The survival function and the probability density functions then are

$$S(t) = \exp\left[-\int_0^t dt\mu(t)\right]; \qquad f(t) = \mu(t)\exp\left[-\int_0^t dt\mu(t)\right]. \tag{1}$$

If $\mu(t)$ is constant (equal to μ) it is referred to as a migration rate. Migration rate μ can be determined on the basis of an appropriate likelihood function (Willekens, 1999). The exponential model of migration can be extended in different directions, e.g., to the case of non-identical individuals; the life span can be split into age intervals (of one or several years) (Blossfeld and Rohwer, 2002; Blossfeld et al., 2007); the number of possible destination states may become larger than 1 (Hachen, 1988).

The exponential model of migration is obtained on the basis of assumption that migration is a non-repeatable event. Let us now assume that the migration is a repeatable event and there is no upper limit on the number of migrations in a time interval. If the event data are available the *Poisson model* of migration can be used. The Poisson model describes the number of migrations during an interval of unit length (e.g. year, month, etc.). In this model the probability of observing n migrations during the unit interval is given by the Poisson distribution

$$P(N = n) = \frac{\lambda^n}{n!}\exp(-\lambda) \tag{2}$$

where λ is the expected number of migrations during the unit interval.

As we have mentioned above the Poisson model is applicable when event data are available. When status data for migration are available, i.e., migration is measured by comparing the places of residence at two consecutive points of time, then the probability of observing n migrants among a sample population of m individuals is given by the binomial distribution (*binomial model of migration*)

$$P(N = n) = \frac{m!}{n!(m-n)!}p^n(1-p)^{m-n} \tag{3}$$

of index m and parameter p representing the probability of being migrant. If the number of possible destinations for migration is larger than 1 (let us assume that the number of possible destinations be K) then the probability distribution connected to migration is given by the multinomial distribution (*multinomial model of migration*)

$$P(N_1 = n_1, \dots, N_K = n_K) = \frac{m!\prod_{i=1}^{K} p_i^{n_i}}{\prod_{i=1}^{K} n_i!}; \qquad \sum_{i=1}^{K} p_i = 1; \qquad \sum_{i=1}^{K} n_i = m. \tag{4}$$

The transition between the states (e.g. addresses occupied in different ages) is given by transitional probabilities $p_{ij}(x)$ which correspond to the probability that an individual who resides in state i at x resides in state j at $t + 1$. p_{ij} may be represented as (Willekens, 2008)

$$p_{ij}(x) = \frac{\exp[\beta_{j0}(x) + \beta_{j1}(x)Y_i(x)]}{\sum_{r=1}^{K}\exp[\beta_{j0}(x) + \beta_{j1}(x)Y_r(x)]} \tag{5}$$

for the case when only the most recent state occupancy is relevant and the person has single relevant attribute (covariate) Y_i that is equal to 1 if the state is occupied and is equal to 0 otherwise. Then the state probability π_i that an individual occupies state i is given by

$$\text{logit}(\pi_j) = \ln\left(\frac{\pi_j}{1-\pi_j}\right) = \beta_{j0}(x) + \beta_{j1}(x)Y_i(x). \tag{6}$$

From the transition probabilities for discrete time processes one may turn to transition rates and this way leads to the Markov chain models (Singer and Spilerman, 1979). The parameters of these models may be estimated even if some data are missing (McLachlan and Krishnan, 1997). Markov chain models are useful to demographers concerned with problems of movement of people (Collins, 1972, 1975) as these models are appropriate for describing and analysing the nature of changes generated by the movement from one state of a system to another possible state and in some cases Markov models may be useful for forecasting future changes. Bayesian modelling framework may be used for generating estimates of place-to-place migration flows too (Raymer, 2007; Bierley et al., 2008).

From the deterministic models we shall give brief additional information about the gravity models and about models connected to the replicator dynamics. Finally we shall mention some urn models of interest for our study.

The gravity model of migration (already mentioned above in the text) is a place-to-place migration model that assumes that interregional migration is directly related to the population of the origin and of the destination regions and inversely related to the distance between them (Greenwood, 2005). The classic gravity model can be written as

$$\ln(M_{ij}) = \ln(G) + \beta_1\ln(P_i) + \beta_2\ln(P_j) - \alpha\ln(D_{ij}) \tag{7}$$