The Rank-Size Rule of a Set of Cities with No or Weak Interaction: A Case Study in Japan

Masayoshi KIKUCHI*

Abstract  S.M. Guseyn-Zade (1977) presents a new model for the relation of rank and size of cities with noninteractions between the cities within a region. It is different from the Zipf formula. According to S.M. Guseyn-Zade (1977), the expected population of the i-th rank city is theoretically shown as follows:

\[ E(N_{i}) = \frac{N}{n} \sum_{k} \frac{1}{k} \]

where \( N \) and \( n \) are the total population and the number of cities respectively, and where \( N_{i} \) is the total population of the i-th rank city. The purpose of this paper is to examine to what extent this model fits the size distribution of cities in seven selected urban areas: Tokyo, Osaka, Nagoya, Sapporo, Hakodate, Hachinohe and Morioka areas, excluding the top rank city in each area. The result shows S.M. Guseyn-Zade's model is appropriate for Tokyo, Osaka and Nagoya areas, and Zipf formula for the other areas.

1. Introduction

Much research has been carried out on the rank-size problem of cities within an urban system.1) This research has progressed in two directions. One direction is the empirical study of size distribution of cities within an urban system, especially within a national urban system. These have focused on the fit of the generalized rank-size rule (that is, the Pareto distribution) to the size distribution of cities (e.g., Suzuki, 1972; Rosen and Resnick, 1980).

The other direction of research consists of theoretical studies concerning the reasons why the generalized rank-size rule fits so well with the size distribution of cities. The stochastic model, as one of theoretical models, is frequently used to explain this. Curry (1964) attempts to demonstrate how the generalized rank-size rule is derived from the maximum entropy principle. One of the other popular stochastic models is the allogemetric growth model proposed by Beckmann (1958). He indicates that the size distribution of cities derived from this model coincides with the generalized rank-size rule.

Many theoretical models, including the stochastic models which were devised to explain the size distribution of cities, are built without explicitly taking into account whether there are interactions between cities or not. For the case where there are no interactions, an interesting model on the relation of rank and size of cities is proposed by S.M. Guseyn-Zade (1977). He uses the term “interaction” between cities in the statistical sense. This means that, from the point of view of

* Faculty of Economics, Toyo University
settlement patterns, the probability of a person’s settling in one of the cities is directly proportional to its size. He asserts that this model shows a good fit for the relations between rank and size of cities in Moscow Oblast and Leningrad Oblast, except for Moscow and Leningrad, respectively.

The purpose of this paper is to examine how far this model fits the size distribution of cities in seven urban areas in Japan. The model is explained in section 2. A series of empirical results are given in section 3, and they are compared with the results obtained from the application of generalized rank-size rule to the same urban areas. The conclusions are shown in section 4.

2. Model

Suppose a region which consists of \( n \) cities, \( C_1, C_2, \ldots, C_n \). Then, assume that the population of \( C_i \) is \( N_i \) and the total population is \( N = \sum_{i=1}^{n} N_i \), where \( N \) is a sufficiently large number. We represent \( N_i/N \) by \( x_i \):

\[
\sum_{i=1}^{n} x_i = 1
\]  

(1)

Therefore \((x_1, x_2, \ldots, x_n)\) consists of \( n \) non-negative numbers and is a point in \( n \)-dimensional Euclidian space \( R^n \) subject to the formula (1). If \( N_i (i = 1, 2, \ldots, n) \) is able to move from zero to \( N \), \((x_1, x_2, \ldots, x_n)\) forms a true \((n-1)\)-dimensional tetrahedron in \( R^n \). S.M. Guseyn-Zade calls it a phase space. He considers that the cities \( C_i \) do not interact if the distribution of probabilities in the phase space is uniform.

Let \( x_{(i)} \) be the largest within these \( x_1, x_2, \ldots \), \( x_{(i)} \) be the next \( x_i \) in order of magnitude and so on. Then the two constraints comes into existence:

\[
\sum_{i=1}^{n} x_{(i)} = 1
\]

\[
1 \geq x_{(1)} \geq x_{(2)} \geq \cdots \geq x_{(n)} \geq 0
\]

(2) (3)

Moreover let us denote a city corresponding to \( x_{(i)} \) and its populations by \( N_{(i)} \).

Under the two constraints, the moving range of \((x_{(1)}, x_{(2)}, \ldots, x_{(n)})\) is restricted in a \((n-1)\)-dimensional hyperplane with \( n \) apexes \((1, 0, 0, \ldots, 0), (1/2, 1/2, 0, \ldots, 0), (1/3, 1/3, 1/3, 0, \ldots, 0), \ldots, (1/n, 1/n, \ldots, 1/n)\).

Since the population \( N_{(i)} \) of the \( i \)-th rank city is equal to \( N x_{(i)} \), its mathematical expectation \( E(N_{(i)}) \) is equal to \( N \) times \( E(x_{(i)}) \):

\[
E(N_{(i)}) = E(N x_{(i)}) = NE(x_{(i)})
\]

\[
E(x_{(1)}), E(x_{(2)}), \ldots, E(x_{(n)})
\]

(4)

mathematically corresponds to the center of gravity in the above stated hyperplane, namely

\[
E(x_{(i)}), E(x_{(2)}), \ldots, E(x_{(n)})
\]

\[
= \frac{1}{n}[(1, 0, \ldots, 0) + (1/2, 1/2, 0, \ldots, 0) + \ldots + (1/n, 1/n, \ldots, 1/n)]
\]

\[
= \frac{1}{n}[(1+1/2+\ldots+1/n, 1/2+1/3+\ldots+1/n, \ldots, 1/(n-1)+1/n, 1/n)]
\]

Therefore, \( E(x_{(i)}) \) is expressed as follows:

\[
E(x_{(i)}) = \frac{1}{n} \sum_{k=1}^{n} \frac{1}{k}
\]

(5)

Hence,

\[
E(N_{(i)}) = \frac{N}{n} \sum_{k=1}^{n} \frac{1}{k}
\]

(6)

If \( n \) is sufficiently large, the following formula is asymptotically established:

\[
1 + 1/2 + \cdots + 1/n \approx \log(1 + n) + K
\]

where \( K \) is Euler’s constant shown as follows:

\[
\lim_{n \to \infty} (1 + 1/2 + \cdots + 1/n - \log n) = K
\]

The efore, \( E(N_{(i)}) \) can be approximately expressed as follows:
The formula (8) suggests that the relation between $E(N(i))$ and the logarithms of its rank of $C(i)$ is approximately expressed as linear.

The most distinctive aspect of Guseyn-Zade model is that if the total population $N$ and the number of cities $n$ in a set of urban places are given, Eq. (8) determines the expected populations of these cities uniquely. Therefore, we do not need to obtain the regression line by the least squares method, which is different from obtaining the regression line of the generalized rank size rule in which we must estimate the two parameters, Pareto exponent and constant term, by the least squares method.

3. Empirical results

In order to examine the validity of the model (8) for the relation between rank and size of a set of noninteracting urban places, it has been applied in seven urban areas in Japan: Tokyo, Osaka, Nagoya, Sapporo, Hakodate, Hachinohe and Morioka. The results are then compared with the results of the generalized rank size rule. The data used for analysis are population by each county (shi-cho-son) taken from the population census of Japan in 1980.

If metropolises are excluded, Tokyo, Osaka and Nagoya metropolitan areas can be regarded as a set of urban places with no interactions because there are very strong interactions only between metropolitan centers and other cities and weak interactions among the latter cities. Therefore, Tokyo, Osaka and Nagoya areas are selected as a set of urban places to which the model is applicable. In addition, to establish how well the equation (8) fits other urban areas excluding the top rank city, Sapporo, Hakodate, Hachinohe and Morioka areas have been selected. In the following discussion we call the Sapporo area the quasi-metropolitan area and Hakodate, Hachinohe and Morioka the non-metropolitan area.

Definition of the boundary of each urban area is a very difficult problem. The boundaries of the metropolitan areas are set up along prefectural boundaries. The quasi- or non-metropolitan areas are done in a few neighboring counties.

Eq. (8) has been used in the data analysis, though it should be noted that the discrepancy between (6) and (8) depends on $n$ and $i$. To compare the results of Guseyn-Zade model, the generalized rank-size rule $N(i)=A+a \log i$ (where $A$ and $a$ are constant) has been applied to the same areas.

The results of this analysis are shown in Fig. 1-(a) ~ (g) and Fig. 2-(a) ~ (g), respectively. The former depict rank on a log scale and population on a normal scale and the latter on a double log scale.

In Fig. 1-(a) ~ (g), the straight line of each figure is easily derived from Eq. (8). The relation between rank and size of the cities in Tokyo and Osaka areas coincides well with the equation. Only the first rank city, Sakai city, in the Osaka area, does not fit well. The fit is also good for the Nagoya area, excluding the upper rank cities.

On the other hand, in the case of the quasi- and non-metropolitan areas, the fit of the observations to the straight line is relatively bad. The results are shown in Fig. 1-(d) ~ (g).

In Fig. 2-(a) ~ (g), the straight line in each figure is that of the generalized rank-size rule which was derived by the least squares method. In Fig. 2-(a) ~ (c), it is seen that the data do not fit the
KIKUCHI: The Rank-Size Rule of a Set of Cities with no or weak Interaction

Fig. 1 Rank-Size Relation derived from Guseyn-Zade Model

$n$: the number of cities within each area

---

183
straight line well especially in the upper and lower rank cities in comparison with that of Fig. 1-(a)−(c). Conversely Fig. 2-(d)−(g) indicate that the data follow the generalized rank-size rule in quasi- and non-metropolitan areas.

In Fig. 1-(a)−(g) and Fig. 2-(a)−(g), interactions
between metropolis and other cities in the metropolitan areas are extremely strong in comparison with interactions among the latter cities. We can regard the set of cities excluding the metropolis in the metropolitan area as a set of non-interacting ones. So, Guseyn-Zade model fits to these areas well. On the other hand, in the quasi- and non-metropolitan areas the generalized rank-size rule fits better than Guseyn-Zade’s model because interactions between the largest city and other cities are not so strong in comparison with interactions among the latter cities.

4. Conclusions

The study of the size distribution of cities in the past focussed only on the generalized rank-size rule in both theoretical and empirical aspects. The results in section 3 show that Eq. (8) is applicable to metropolitan areas like Tokyo, Osaka and Nagoya. Although selected case are limited and the selection of boundaries in each urban area is arbitrary, this study clarified that the Guseyn-Zade model fits better than the generalized rank-size rule to the set of cities with no or weak interactions.

Acknowledgement

The author wishes to express his thanks to the referee for his helpful suggestions. The author is also indebted to Professor Joseph B.R. Whitney of University of Toronto for reading of the manuscript. (Accepted July 15, 1986)

Footnotes


References


相互作用の弱い都市群の順位規模分布に対する
Guseyn–Zade モデルの適合性について

菊 池 正 佳

都市システム内の各都市間に相互作用がない場合の都市の順位と規模の関係について、S.M. Guseyn-Zade (1977) は、Zipf 法則とは異なる関係式が成り立つことを示している。彼によると、上位から i 番目の都市の理論的期待人口は、

\[ E(N_{x(i)}) = \frac{N}{n} \sum_{k=i}^{n} \frac{1}{k} \]

であり、漸近的（n が十分大の時）には

\[ E(N_{x(i)}) \approx \frac{N}{n} (\log(1+n) - \log i) \]

によって求まる。ここで、N は都市システム内の総人口、n は都市の総数、x_{i} は i 目の都市の人口 N_{i}の総人口 N に対する比 N_{i}/N である。

ところで、日本の三大都市圏域の都市群の場合を考えると、各都市圏域内において、最上位の都市とその他の都市との相互作用が非常に強い反面、後者の都市間の相互作用は弱く、まさに Guseyn-Zade モデルの想定する状態に近似的に対応する。

そこで本稿では、相互作用を持たない又は弱いと想定される都市群の例として、首都地域を除く三大都市地域の都市群を取りあげ、Guseyn-Zade のモデルと Zipf の法則のいずれがよくあてはまるかを検討した。また、これらの都市群との比較のため、札幌、函館、八戸、盛岡の各都市周辺の首位都市を除く都市群に対しても、両モデルの適用を試みた。

結論として、大都市圏域の都市群については、Guseyn-Zade のモデルが、その他の都市群については、Zipf の法則がよくあてはまることが示された。

* 東洋大学経済学部