Population characterizes collective behaviors of human beings. In this paper we investigate cumulative population distribution of municipalities (cities, towns and villages) of Japan to show that each of them can be represented by a characteristic distribution function. We also introduce a numerical model for population dynamics to explain the origin of the distributions. From our analysis it has become clear that each distribution originates from the lower and upper bounds of population in the definitions of cities and towns.

KEYWORDS: lognormal distribution, power law distribution, population, boundary condition

1. Introduction

Population is an important indicator characterizing collective human behaviors. The dynamics of population is a complicated phenomenon because its property is strongly affected by many factors such as politics, economy and religion.

Sometimes the dynamics of population is investigated on the basis of the population distribution within a given nation and region. For example, Auerbach found that the population distribution of an urban area in Russia obeys a power law distribution,\(^1\) which was later formulated as Zipf’s law.\(^2\) Zipf’s law has the form as

\[
F(x) = ax^{-b},
\]

where \(F(x)\) and \(x\) are the rank of the cities and the population, respectively. Especially, classic version of Zipf’s law is the case of \(b = 1\). It can be seen that Zipf’s law is the discrete case of Pareto distribution.\(^3\) On the other hand, Eeckhout revealed that the population distribution of all municipalities (cities, towns, boroughs and villages) in U.S. follows a lognormal distribution.\(^4\) Recently, Kobayashi et al. also revealed that the population distribution of prefectures in Japan obeys a lognormal distribution.\(^5\)

Apart from the population dynamics, lognormal distribution and power law distribution are often found in various complicated phenomena in nature. Some examples of power law distributions are size distribution of lunar craters,\(^6\) the Gutenberg-Richter’s law for earthquakes,\(^7\) and size distribution of islands.\(^8\) Some examples of lognormal distributions are the productivity of scientists publishing research papers,\(^9\) the fragmentation of glass rods,\(^10\) income distribution of families and single individuals in U.S.,\(^11\) life span of animals,\(^5\) food fragmentation by chewing\(^12\) and the life span of aged people.\(^13\) The lognormal distribution \(n(x)\) is written as

\[
n(x) = \frac{1}{\sqrt{2\pi} \sigma x} \exp \left( -\frac{[\ln(x/T)]^2}{2\sigma^2} \right),
\]

where \(\sigma\) and \(T\) are dispersion and average, respectively. The cumulative form of the lognormal distribution (CLND) is given by

\[
N(x) = \frac{N_T}{2} \left( 1 - \text{erf} \left( \frac{\ln(x/T)}{\sqrt{2\sigma}} \right) \right),
\]

where \(N_T\) is the total number of samples, and \(\text{erf}(z)\) is the error function defined as

\[
\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \exp(-y^2)dy.
\]

A stochastic process which generates lognormal distribution is random multiplicative stochastic process. If an event is composed of the succession of \(m\) independent events with the probabilities \(p_i\) (\(1 \leq i \leq m\)), the probability of the event, \(P_r^{(m)}\), is given by the product of those independent random variables as \(P_r^{(m)} = \prod_{i=1}^{m} p_i\), so that the logarithm of \(P_r^{(m)}\) is described as \(\log P_r^{(m)} = \sum_{i=1}^{m} \log p_i\). If \(m\) is sufficiently large, the distribution of \(\log P_r^{(m)}\) becomes a normal distribution due to the central limit theorem, so that the distribution of \(P_r^{(m)}\) becomes a lognormal distribution. If \(p_i\) is replaced by the growth rate of the quantity at \(i\)-th step, that is called Gibrat’s law.\(^14\)\(^15\)

The aims of this paper are investigating the cumulative population distribution of municipalities in Japan and explaining the origin of those distributions by a simple model which hopefully captures the essence of population migration. The organization of this paper is as follows. In the next section, we show the cumulative population distribution of all municipalities in Japan and those of each municipality. In §3, we construct a simple model of population migration to show that the model can reproduce those characteristic distributions. In §4, we summarize our results. In our analysis we obtained the population data from Statistics Bureau, Ministry of International Affairs and Communications, Japan.\(^16\)

2. Typical Population Structure of Each Municipality

Kobayashi et al. reported that the population distribution of the prefectures of Japan obeys a lognormal distribution and is divided into two parts due to the high economic growth after the World War II.\(^5\) Thus, in this section, we firstly investigate the population distribution
of all municipalities which are smaller than the prefectures.

Figure 1 is the cumulative population distribution of 2217 municipalities of Japan, except for Miyake village, in 2005. Most part of the distribution is fitted by eq.(3) with $N_T = 2217$, $T = 17200$ and $\sigma = 1.5$ while the tail end is fitted by eq.(1) with $b = 1.2$. It is known that the consumer income distribution of U.S. in 1935/1936 is similar to this distribution.

The tail end is composed of 14 government-designated cities such as Yokohama and Osaka. Therefore, the population distribution of large cities differs from that of others.

Here we explain why a large part of the cumulative population distribution is approximated by CLND. As previously mentioned, a growth process of a quantity generates a lognormal distribution when the growth rate of the physical quantity is independent of the quantity, which is called as Gibrat’s law. Figure 2 shows the relation between the population of municipalities $x$ and the population growth rate for each municipality from 2000 to 2005. The population growth rate is defined as

$$\frac{100 \times (N_d^{2005} - N_d^{2000})}{N_d^{2000}}$$

where $N_d^{2000}$ and $N_d^{2005}$ are the populations in 2000 and 2005, respectively. As seen in Fig.2, this relation indicates that the population growth rate of each municipality is almost constant and has no correlation with the population. Thus, it can be seen that most part of the cumulative population distribution is approximated by CLND due to the noncorrelation between the population growth rate and the populations of municipalities.

Secondly, we show the cumulative population distributions of cities, towns and villages, respectively. Figure 3(a) shows that the cumulative population distribution of cities exhibits the power law with $b = 1.2$. Thus, it can be seen that the tail end of the cumulative population distribution of all municipalities of Japan corresponds to that of cities. In the case of towns, the discrepancy between the cumulative population distribution and CLND becomes large in the large population region (Fig.3(b)). In the case of villages, however, the cumulative population distribution is well fitted by eq.(3) with $N_T = 288$, $T = 3.2 \times 10^3$ and $\sigma = 0.86$. Those characteristics are seen also before the great Heisei merger in 2000.

Here we discuss why such characteristic distributions appear for each municipality. We focus on the promotion condition that a town or a village is promoted to a city or a town, respectively. For the promotion of a town and a village to a city, they must fulfill a population condition, an urban functional condition, and so on. The population condition is that the number of residents in the municipality is more than 50,000 persons. The urban functional condition is that the percentage of the people who works for urban-business such as commerce and industry and their family is larger than 60 percent of all population. Therefore, the lower bound for population of cities is 50,000 persons. Several previous works showed that a multiplicative process with a lower boundary con-
dution generates a power law distribution. Thus, it can be seen that the lower bound generates the power law distribution in the cumulative population distribution of cities.

For the promotion of villages to towns, each prefecture has its own population conditions and urban functional conditions. Thus, the towns have lower bound on the number of residents. For the promotion of towns to cities, towns have already satisfied the urban functional condition in the promotion from villages to towns. If a town develops normally, an urban functional condition is not important in the promotion from a town to a city. Thus, the towns have also upper bound on the number of residents. In the case of villages, it can be seen that the villages do not have an urban function. Therefore, we can regard that the villages do not have a population condition.

To confirm our conjectures, we present all the distributions in Fig.4. Figure 4 shows that the cumulative population distribution of cities is hardly overlapped to that of towns or villages. On the other hand, the cumulative population distribution of towns is comparatively overlapped to that of villages. Additionally, we draw the line of 50,000 persons as a reference. The line shows that the population condition is so strong restriction in the cumulative population distribution of towns. Moreover, in the region of more than 50,000 persons, the cumulative population distribution of cities is well fitted by the power law distribution. Therefore, it can be seen that the cumulative population distribution of villages is not affected by the population condition, while the cities and towns are affected. From these facts, we expect that the characteristic cumulative distributions for population of each municipality are attributed to the population condition in the promotion of villages to towns to cities or towns.

3. Model for Population Migration

In this section we construct a toy model for the population migration to investigate the effect of the population condition on the characteristic population distributions. With the aid of the model, we separately simulate the population migration of each municipality by introducing upper or lower bound of the population which corresponds to the population condition.

Let us introduce our model. Our model consists of 1000 sites, each of which has the initial population 1.0. In the simulation for each municipality, all the sites belong to the municipality. At each simulation step, a portion of population migrate between two chosen sites. The procedure of one simulation step is summarized as follows:

1. We randomly choose a source site $i$ with the population $n_i$.
2. We choose a group of sites, $G_{n<n_i}$ or $G_{n>n_i}$. Here $G_{n<n_i}$ and $G_{n>n_i}$ are the groups of the sites whose population are less and more than $n_i$, respectively. The probability to choose $G_{n<n_i}$ is $\alpha$ (migration parameter) while that to choose $G_{n>n_i}$ is $1-\alpha$.
3. Among the group of sites chosen in the previous step, we randomly choose the destination site $j$ for migration.
4. $P_{ij}$ percent of $n_j$ are transferred to the site $j$, and thus the populations of sites $i$ and $j$ vary in quantity as $n_i - P_{ij}n_i$ and $n_j + P_{ij}n_i$, respectively.

In the last step, $P_{ij}$ is randomly chosen in the range from 0 to 90. We iterate this procedure 10000 times in our simulation.

Here we explain how to introduce the lower and upper bounds into the model. We choose the lower and upper bounds from normal random numbers with average $T_n$ and dispersion $\sigma_n$ at each simulation step. In our simulation process, when the number of residents of a site exceeds the upper bound, the number of residents of the site is replaced by a random value smaller than the upper bound. This procedure corresponds to the change that a village is promoted to a town followed by the promotion from a village or a town to a city. If, for instance, a town is promoted to a city, the town is removed from the group of towns. On the other hand, cities are regarded as having a clear lower bound owing to the strong population constraint. When a lower bound is introduced, we iterate the basic procedure 10000 steps so that the population of a site is not less than the lower bound. For sample average, we averaged results of 100 samples in our simulation.

We show the cumulative population distributions obtained in our simulation in Fig.5. To obtain these results, the value of $\alpha$ is set to 0.05. The cumulative population distribution in Fig.5(a) is obtained when the lower bound is introduced with $T_n = 0.3 \times 10^{-2}$ and $\sigma_n = 0.5 \times 10^{-2}$. In this case, the cumulative population distribution exhibits the power law distribution with $b = 0.55$ and is similar to that of actual cities in Fig.3(a), although the slopes are different.

The cumulative population distribution in Fig.5(b) is obtained when the upper bound is introduced with $T_n = 2.0$ and $\sigma_n = 0.2$ and the lower bound with $T_n = 0.7 \times 10^{-4}$ and $\sigma_n = 0.3 \times 10^{-4}$. In this case, a large part of the cumulative population distribution is well fitted by CLND with $N_T = 1000$, $T = 2.2 \times 10^{-3}$ and $\sigma = 2.9$, although the discrepancy between the data and the fitting curve becomes larger in the region of large population. This is due to the cut off coming from the upper bound. The cumulative population distribution is similar to that of actual towns in Fig.3(b).

The cumulative population distribution in Fig.5(c) is obtained when we introduce neither the lower nor the
upper bound. In this case, the cumulative population distribution is well fitted by CLND with $N_T = 1000$, $T = 5.1 \times 10^{-3}$ and $\sigma = 3.3$ and is similar to that of actual villages in Fig.3(c).

Here we investigate how $\alpha$ affects the cumulative population distribution when we do not introduce the population condition into the model. In Fig.6 we show the cumulative population distributions obtained by introducing neither the lower nor the upper bound. Figure 6(a), (b) and (c) are the cumulative population distributions obtained in the cases $\alpha = 0.0$, 0.5 and 1.0, respectively. For $\alpha = 0.0$ the cumulative population distribution is well fitted by CLND while, for $\alpha$ larger than 0, the discrepancy from CLND becomes larger with the increase of $\alpha$.

To characterize this tendency, we investigate the population growth rate $N_s^{10000}/N_s^{8000}$ of each site, where $N_s^{8000}$ and $N_s^{10000}$ are the number of residents at 8000 and 10000 simulation steps, respectively. For example, Fig.7 shows the relation between $N_s^{8000}$ and the growth rate in the case of $\alpha = 1.0$. The solid line is drawn by the Gauss-Newton method and has the slope $-0.9$. The relation between $\alpha$ and the average of the slope is shown in Fig.8, where the average of the slope is calculated over 100 samples. For small $\alpha$, the average value of the slopes approaches 0, which indicates that Gibrat’s law is satisfied. Thus, it can be seen that the distribution approaches CLND with the decrease of $\alpha$, as shown in Fig.6.

Let us discuss the difference of the power exponents of the population distributions of cities. In the simulation of the population migration for cities, we have found that the power exponent, $b = 0.55$, is different from the actual value, $b = 0.12$. This discrepancy is due to the fact that the promotion of smaller municipalities to cities is not introduced into our model. In realistic systems, historically, many towns and villages have been promoted to small cities, which is the origin of the large power exponent. Thus, we believe that a more realistic model can be constructed by taking into account the promotion of villages and towns to cities, which will generate the larger power exponent $b$.

Finally, we would like to comment on how both the migration rate $P_{ij}$ and the initial distribution of the number of residents depend on the cumulative population distributions. We introduce the migration rate $P_{ij}$ as a fixed value provided that both the maximum of simulation steps and $\alpha$ are fixed. In this case, the cumulative population distribution approaches not a lognormal distribution but a power law distribution with the decrease of $P_{ij}$. In addition, we randomly give initial numbers of residents to all the sites by using uniform random numbers ranging from 0 to 1 provided that $P_{ij}$, the maximum of simulation steps, and $\alpha$ are fixed. Also in this case, this model reproduces approximately the same distributions as those obtained when the initial numbers of residents are uniform.

4. Concluding Remarks

In summary we have investigated the cumulative population distribution of all municipalities in Japan. A large part of the cumulative population distribution of all the municipalities exhibits a lognormal distribution while the tail end of the distribution exhibits a power law. Investigating the relation between the population growth rate and the population, we have concluded that the population data of all municipalities satisfy the Gibrat’s law, which is the reason why a large part of the cumulative population distribution exhibits a lognormal distribution.

We have also investigated the cumulative population distributions of cities, towns and villages, respectively. The cumulative population distributions of cities and villages are well fitted by power law distribution and CLND, respectively, while the cumulative population distribution of towns shows a large discrepancy from CLND in the large population region. This discrepancy is clearly due to the promotion condition for a town to become a city. On the assumption that these characteristic distributions of cities, towns and villages are generated by the promotion condition of each municipality, we have constructed a population migration model which can reproduce the actual cumulative population distributions of each municipality in Japan. Thus, it can be seen that the characteristic distributions of cities, towns and villages are strongly affected by the promotion conditions that are unique to each municipality.
Fig. 6. The cumulative population distributions generated by our model with (a) $\alpha = 0.0$, (b) $\alpha = 0.5$ and (c) $\alpha = 1.0$, respectively. The parameters in (a) are $N_T = 1000$, $T = 0.23 \times 10^{-2}$ and $\sigma = 3.7$, while the parameters in (b) are $N_T = 1000$, $T = 0.22$ and $\sigma = 1.7$. And the parameters in (c) are $N_T = 1000$, $T = 0.79$ and $\sigma = 0.85$.

For future work, we should (i) investigate the effect of $\alpha$ in the actual municipalities, (ii) carry out simulations using the model with large number of sites, and (iii) refine our model so as to generate larger power exponents $b$ in the simulation for cities. We hope that the discussions of this study are useful for researches of the human migration property.

Fig. 7. Log-log plot for relation of $N_{8000}$ and growth rate in the case of $\alpha = 1.0$. The line is drawn by the Gauss-Newton method.

Fig. 8. Relation between $\alpha$ and the average of power law exponents. Here the power law exponent is given by the relation between values of each site and the growth rates. Each circle and the error bar show an average of 100 samples and the standard deviation, respectively.

Acknowledgement

We would like to thank Prof. M. Katori and Prof. T. Nakano for their valuable comments. This study is supported by a Grant-in-Aid of the Ministry of Education, Culture, Sports, Science, and Technology (MEXT), Japan (Grant No.18340115).

1) F. Auerbach: Petermanns Geogr Mitt (1913) 73.
3) V. Pareto: Cours d’Economic Politique. (1897).
16) http://www.stat.go.jp/.