〈寄稿論文〉

A Model of Class Identification:*

Generalization of the Fararo-Kosaka Model using Lyapounov's Central Limit Theorem

Hiroshi HAMADA**

1. Introduction: Fararo-Kosaka model of class identification

In this paper, we explore a mechanism for generating a class identification process. We use the concept of "class identification" in the following sense. In empirical studies based on social surveys, class identification is often defined operationally as an answer to the question 'If we were to divide the entire society into several levels, which level would you say you belong to?' We treat the respondents' answer to this question as their self-location in society. For example, in a Social Stratification and Mobility survey, the response categories for the question of class identification are 'upper, upper middle, lower middle, upper lower, and lower lower.' Empirical studies of class identification mainly focus on specifying its social and economic determinants. Such studies have demonstrated historical changes affecting social status such as household and individual income, education, occupation, gender, and age based on class identification (Shirahase 2010). These empirical studies provide basic assumptions for making an abstract model.

There are many empirical studies of class identification based on statistical analysis whereas few theoretical studies attempt to explain a mechanism for generating a class identification process. Formalization studies are likely to explain a mechanism by making a model rather than describing historical changes in the determinants of class identification. The model of the image of stratification formalized by Fararo and Kosaka is one of the most significant studies in the field of social stratification because the model attempts to explain how individual actors located in the structure of any given social system acquire a stable image of the structure, which varies with location, and how the actors locate themselves within their structural image. The Fararo-Kosaka model (hereafter, FK model) explicitly formalized the process whereby people compare their social status with others to build an image of social stratification where they could locate their class identity based on a subjective image of stratification. The FK model has many interesting implications and it successfully explains the magnification of middle class identification (Kosaka 2006; Fararo and Kosaka 2003). Many studies have attempted to develop the FK model further (Shirakura and Yosano 1991; Yosano 1996; Ishida 2003, 2005; Karpinski 2009; Maeda 2011).

The FK model assumes that each dimension of social status, such as education, income, and occupation, has an internal rank and that the ranges of ranks are equal among dimensions. Suppose that r_1, r_2, ..., r_n are the maximum values of numbers in the ranks of each dimension. For example r_1 is the

---

*Key Words: class identification, Fararo-Kosaka model, central limit theorem
**Associate Professor, Tohoku University

maximum value in ranks of the first dimension, r_2 is that of the second dimension, and so on. The FK model assumes that $r_1 = r_2 = \cdots = r_n$ for all dimensions. This assumption implies that if education is measured using five ranks, then other dimension of social status, for example income, should be also divided into five categories. The number of ranks in each dimension directly has an effect on the position of self-location in a society. Therefore, the assumption of rank homogeneity may be too strong. However, few studies have explored the implications when the rank homogeneity assumption is relaxed from both theoretical and empirical perspectives. In this paper, we generalize the rank homogeneity assumption and show that the empirical validity of the model will be improved by this generalization. In other words, we attempt to generalize the FK model theoretically and empirically.

The remainder of this paper is organized as follows. In section 2, we apply Lyapounov's central limit theorem to the FK model to generalize the condition of the marginal distribution of social status and we prove that the main proposition of the FK model still holds when each dimension of social status obeys a non-identical discrete uniform distribution with different ranges. In section 3, we further generalize the FK model by applying the central limit theorem and we show that the distribution of standardized class identification asymptotically obeys the normal distribution, if each distribution of the dimensions has a finite mean and variance, and the maximum value of ranks and the dimensions are independent of each other. In section 4, we test the empirical validity of the generalized FK model by checking the correlation between the theoretical predictions of class identification from the FK model and observed data of class identification in an SSM survey. In section 5, we summarize the results of the mathematical analysis and empirical test.

2. Non-identical discrete uniform distribution model

Definition (a set of profiles of social status). Social stratification consists of n dimensions and each dimension has an internal rank. The rank in each dimension is given by a natural number. For the ith dimension, $\{1, 2, ..., r_i\}$ is as a set of ranks. r_i indicates the highest rank of the ith dimension. A set of profiles of social status is expressed as direct products of the sets of ranks. We call this S and it is explicitly given by

$$S = \{1, 2, ..., r_1\} \times \{1, 2, ..., r_2\} \times \cdots \times \{1, 2, ..., r_n\}.$$

Individual social status profiles are defined as a point $(x_1, x_2, ..., x_n) \in S$. Generally, the ranges of ranks for dimensions are different from each other. We allow cases such as $r_i \neq r_j$. Conversely, if the special condition $r_1 = r_2 = \cdots = r_n$ is satisfied, the definition of a set of social status profiles corresponds to that of the FK model.

For example, suppose there are two dimensions such as education and wealth. The first dimension has three ranks (low, middle, and high) while the second dimension has two ranks (poor and rich). A set of profile of social status is then,

$$S = \{1, 2, 3\} \times \{1, 2\} = \{(1, 1), (1, 2), (2, 1), (2, 2), (3, 1), (3, 2)\}.$$

Individual social status profiles are (1, 1), (1, 2), (2, 1), (2, 2), (3, 1) and (3, 2).

The FK model assumes that actors are stratified along a number of salient dimensions and that the dimensions of stratification are themselves ordered relative to their importance. Therefore, it is possible to combine positions occupied by the actors in each dimension to arrive at a single ordering of the members of the system. In the FK model, the ordering is defined using the lexicographic order (Fararo and Kosaka 2003; Karpinski 2009). According to this rule, if person p occupies a higher position than person o in the most socially significant dimension, then p is higher than o in the overall ordering, regardless of the positions they occupy in the remaining dimensions of stratification. Thus, if p and o are status-equals with respect to the first dimension, but o is higher along the second one, then o is also higher in the overall ordering of positions. For example, suppose that there is a society with structure $S = \{1, 2, 3\} \times \{1, 2\}$ and the first dimension is more important than the second dimension. Suppose that the magnitude of numbers corresponds to the order of rank (larger number indicates high status). The lexicographic order determines the rank of profiles as follows; $(3, 2) \succ (3, 1) \succ (2, 2) \succ (2, 1) \succ (1, 2) \succ (1, 1)$.

Corollary 1 (rank of class identification). Suppose that the focal actor has profile $(x_1, x_2, ..., x_n)$. The rank of self-location in its image of stratification is

$$\sum_{i=1}^{n} x_i - (n-1).$$

This equation indicates the rank of class identification from the bottom.

Proof. The proof of this corollary is almost trivial because we can use the result of a previous study (Kosaka 2006). Let $(x_1, x_2, ..., x_n)$ be the profile of the focal actor. In the first dimension, the number of cases where the ranks of other are smaller than the focal actor's rank is $x_1 - 1$. The number of cases where the ranks of others are equal to the focal actor's rank in the first dimension and the ranks of others are smaller than the focal actor's rank in second dimension are $x_2 - 1$. Similarly, the number of cases where the ranks of others are equal to the focal actor's rank from first dimension to $n-1$th dimension and ranks of others are smaller than the focal actor's rank in the nth dimension are $x_n - 1$. Therefore, all possible patterns where the ranks of other profiles are below the focal actor's profile is $\sum_{i=1}^{n}(x_i - 1)$. The focal actor's subjective class identification from the bottom is $\sum_{i=1}^{n}(x_i - 1)$. We obtain

$$\sum_{i=1}^{n}(x_i - 1) + 1 = \sum_{i=1}^{n} x_i - n + 1 = \sum_{i=1}^{n} x_i - (n-1)$$

Proposition 1 (probability function of class identification). Suppose that the rank of each dimension is represented using non-identical and independent discrete uniform distributions. The probability that an actor's class identification becomes x is

$$\Pr\{X = x\} = \left(\prod_{i=1}^{n} r_i\right)^{-1} \beta_x$$

where

$$\beta_x = \#\{y = (y_1, y_2, ..., y_n) \,|\, y \in S,\ \sum_{i=1}^{n} y_i - (n-1) = x\}.$$

In the above equation $\#A$ indicates the number of elements in set A.

Proof. Class identification of profile $y=(y_1, y_2, ..., y_n)$ in S is $\sum_{i=1}^{n} y_i-(n-1)$. Let this value be x. There are several profiles whose class identification becomes x. The number of patterns can be represented by

$$\#\{y=(y_1, y_2, ..., y_n) \mid y \in S, \sum_{i=1}^{n} y_i-(n-1)=x\}. \tag{1}$$

Based on the assumption of the independence of random variables, for all $y=(y_1, y_2, ..., y_n)$, we have

$$\Pr\{(Y_1=y_1)\cap(Y_2=y_2)\cap\cdots\cap(Y_n=y_n)\}=\Pr\{Y_1=y_1\}\Pr\{Y_2=y_2\}\cdots\Pr\{Y_n=y_n\}=\frac{1}{r_1}\frac{1}{r_2}\cdots\frac{1}{r_n}.$$

The number of patterns where the class identification becomes x is expressed by (1), and each pattern is mutually exclusive. Thus,

$$\Pr\{X=x\}=\left(\prod_{i=1}^{n} r_i\right)^{-1}\beta_x.$$

Figure 1 shows the distributions of class identification generated by the probability function in Proposition 1. In Figure 1, 8×2 means that the number of ranks in the first dimension is 8 and the number of ranks in second dimension is 2. Similarly, $8\times2\times3$ means that 8 is the number of ranks in first dimension, 2 is that in the second dimension, and 3 is that in the third dimension.

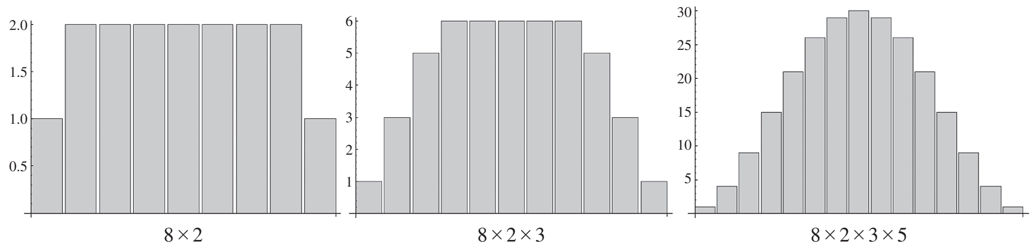

Figure 1 Distribution of class identification. Numbers at the bottom of the graph indicate the rank of each dimension. The distribution is computed given the condition that there is only one person in each profile.

In terms of probability theory, we then have to specify the distribution of class identification derived from the FK model where the range of ranks for each dimension is different from each other.

Let, X_i be a discrete random variable and 1, 2, ..., r_i be its values. We can define a new random variable S_n by composing n random variables as follows.

$$S_n=\sum_{i=1}^{n} X_i-(n-1)$$

The class identification distribution is specified by identifying a probability function $\Pr\{S_n=j\}=f(j)$.

Yosano (1996) proved the very important proposition that the class identification distribution obeys the normal distribution when a random variable in each dimension independently obeys an identical uniform distribution. The rank homogeneity assumption, i.e., the identical distribution assumption, means that education, income, and occupation are subject to identical distribution with the same range. We know that the formalization approach often uses such strong assumptions to simplify a model while the strong assumption itself has no problem. However, if we can relax the strong assumptions without any loss of generality, this will expand the universality of the theory. Moreover, we can expect to improve empirical validity of the model by this generalization.

We will show that, even when a random variable in each dimension obeys a non-identical and independent uniform distribution, the class identification also obeys a normal distribution. In order to test this proposition, we use the following general theorem of the probability theory.

Theorem 2 (Lyapounov's central limit theorem). Suppose that $\{X_j\}$ is a sequence of independent random variables and each random variable has a mean of $E(X_j) = \mu_j$ and a variance $V(X_j) = \sigma_j^2$. Suppose that $B_n = \sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2}$ and for some $\delta > 0$, condition

$$\lim_{n\to\infty} \frac{1}{B_n^{2+\delta}} \sum_{j=1}^{n} E[|X_j - \mu_j|^{2+\delta}] = 0$$

holds. Then,

$$Y_n = ((X_1 - \mu_1) + (X_2 - \mu_2) + \cdots + (X_n - \mu_n))/B_n$$

obeys a standard normal distribution $N(0, 1)$.
Proof. Details are given in Billingsley (1986).

Note that, in Theorem 2 the components of the random variable sequence $\{X_j\}$ are not necessarily identical. It is enough that the mean and variance hold the condition of $2+\delta$th order moment. We refer to this as Lyapounov's condition.

Proposition 3. Suppose that a set of social status profiles is expressed as the direct product of sets of ranks, $S = \{1, 2, ..., r_1\} \times \{1, 2, ..., r_2\} \times \cdots \times \{1, 2, ..., r_n\}$. When a random variable in each dimension obeys a non-identical and independent uniform distribution, the standardized class identification distribution obeys a standard normal distribution $N(0, 1)$ if n is sufficiently large.

Proof. Let X_j be a random variable in a discrete uniform distribution of the jth dimension. The values of X_j are 1, 2, ..., r_j. Let $K > 0$ be the maximum number of ranks in all dimension. Thus, for all j, $\max_j(r_j) = K$. For all j, an absolute third order moment of X_j satisfies

$$E[|X_j - \mu_j|^3] = \sum_{i=1}^{r_j} |i - \mu_j|^3 \frac{1}{r_j} = \sum_{i=1}^{r_j} |i - \mu_j|\,|\mathrm{i} - \mu_j|^2 \frac{1}{r_j} < \sum_{i=1}^{r_j} K |i - \mu_j|^2 \frac{1}{r_j} = K\sigma_j^2 .$$

We use the following relation $|i - \mu_j| < K$ for all i ($i = 1, 2, ..., r_j$). Based on the inequality with respect to an absolute third order moment, we have

$$\frac{1}{B_n^3} \sum_{j=1}^{n} E[|X_j - \mu_j|^3] < \frac{1}{B_n^3} \sum_{j=1}^{n} K\sigma_j^2$$

and

$$\lim_{n\to\infty}\frac{1}{B_n^3}\sum_{j=1}^{n}K\sigma_j^2=\lim_{n\to\infty}\frac{KB_n^2}{B_n^3}=\lim_{n\to\infty}\frac{K}{B_n}=0\,.$$

Since the limit of the sum of absolute values is not negative,

$$\lim_{n\to\infty}\frac{1}{B_n^3}\sum_{j=1}^{n}E\,[\,|X_j-\mu_j|^3\,]$$

should not be negative. Because $\forall n,\ a_n<b_n\Rightarrow\lim_{n\to\infty}a_n\le\lim_{n\to\infty}b_n$, we obtain

$$0\le\lim_{n\to\infty}\frac{1}{B_n^3}\sum_{j=1}^{n}E\,[\,|X_j-\mu_j|^3\,]\le\lim_{n\to\infty}\frac{1}{B_n^3}\sum_{j=1}^{n}K\sigma_j^2=0\Rightarrow\lim_{n\to\infty}\frac{1}{B_n^3}\sum_{j=1}^{n}E\,[\,|X_j-\mu_j|^3\,]=0$$

This implies that, for $\delta=1$, Lyapounov's condition is satisfied. Therefore,

$$Y_n=((X_1-\mu_1)+(X_2-\mu_2)+\cdots+(X_n-\mu_n))/B_n$$

obeys a standard normal distribution $N(0,\,1)$. As shown, the class identification is given by

$$Y=X_1+X_2+\cdots+X_n-(n-1)\,.$$

If we standardize Y as $Y_n=\{Y+(n-1)-\sum_{i=1}^{n}\mu_i\}/B_n$, Y_n obeys a standard normal distribution $N(0,\,1)$ based on the central limit theorem.

Using the following corollary, we can specify the class identification distribution.

Corollary 2. When a random variable in each dimension independently obeys a non-identical uniform distribution, the non-standardized class identification distribution obeys an approximately normal distribution with a mean $(\sum_{j=1}^{n}\mu_j)-(n-1)$ and a variance $\sum_{j=1}^{n}\frac{r_j^2}{12}$, if n is sufficiently large.

Proof. From $Y_n=((X_1-\mu_1)+(X_2-\mu_2)+\cdots+(X_n-\mu_n))/B_n$, $X_1+X_2+\cdots+X_n=B_nY_n+\sum_{j=1}^{n}\mu_j$.

Because Y_n obeys a standard normal distribution, $X_1+X_2+\cdots+X_n$ obeys a normal distribution. When X_j obeys a discrete uniform distribution with the range $[1,\,r_j]$, its variance is $V(X_j)=r_j^2/12$. Therefore, the variance of $X_1+X_2+\cdots+X_n$ will be the variance of Y_n times B_n^2.

$$1\cdot B_n^2=\sigma_1^2+\sigma_2^2+\cdots+\sigma_n^2=\sum_{j=1}^{n}\frac{r_j^2}{12}\,.$$

Therefore, $X_1+X_2+\cdots+X_n$ obeys a normal distribution,

$$N\Big((\sum_{j=1}^{n}\mu_j),\ \sum_{j=1}^{n}\frac{r_j^2}{12}\Big).$$

Subtracting constant, $X_1+X_2+\cdots+X_n-(n-1)$ obeys a normal distribution,

$$N\Big((\sum_{j=1}^{n}\mu_j)-(n-1),\ \sum_{j=1}^{n}\frac{r_j^2}{12}\Big).$$

Figure 2 shows an example of this corollary.

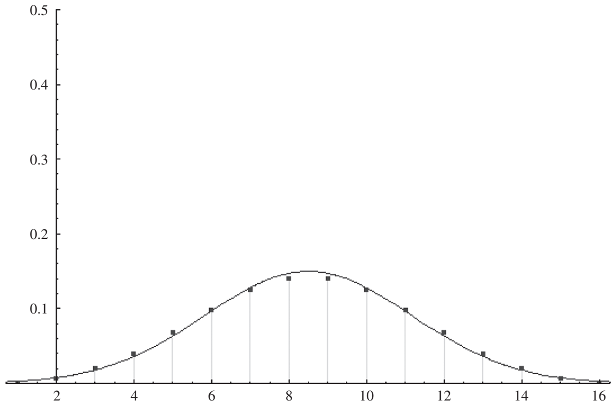

Figure 2 Plot of a discrete probability function and a continuous probability density function of a class identification distribution, which is theoretically specified where $r_1=3$, $r_2=4$, $r_3=2$, $r_4=6$, $r_5=5$. Points on the graph correspond to the values of the probability function while the curved line corresponds to a plot of the probability density function[1).]

3. General model

Note that Proposition 3 and Corollary 2 hold only when a random variable in each dimension independently obeys a uniform distribution. This condition is not usually satisfied in the real world. However, when Lyapounov's condition holds we obtain a normal distribution as a class identity distribution. Thus, we do not have to assume a uniform distribution in order to derive a normal class identity distribution from the FK model.

Proposition 4. Suppose that a set of social status profiles is expressed as a direct product of sets of ranks, $S=\{1, 2, ..., r_1\}\times\{1, 2, ..., r_2\}\times\cdots\times\{1, 2, ..., r_n\}$. When random variables from each dimension are independent of each other, they have a finite mean $E(X_j)=\mu_j$ and a variance $V(X_j)=\sigma_j^2$, respectively. If $\lim_{n\to\infty} \sigma_1^2+\sigma_2^2+\cdots+\sigma_n^2=\infty$ holds, then the standardized class identification distribution obeys a standard normal distribution $N(0, 1)$, if n is sufficiently large.

Proof. Let X_j be a random variable in the ith dimension. The values of X_j are 1, 2, ..., r_j. Let $K>0$ be a maximum number of ranks in all dimension and $f_j(x)$ be a probability function of X_j. For all j, the absolute third order moment of X_j satisfies

$$E[|X_j-\mu_j|^3]=\sum_{i=1}^{r_j}|i-\mu_j|^3 f_j(i)\leq\sum_{i=1}^{r_j}K|i-\mu_j|^2 f_j(i)=K\sigma_j^2.$$

Based on a similar logic to the proof of Proposition 3, we have

1) More accurately, the line indicates the plot of a probability density function for a normal distribution with the mean $\left(\sum_{j=1}^{5}\mu_j\right)-4$ and the variance $\left(\sum_{j=1}^{5}r_j^2/12\right)$.

$$\lim_{n\to\infty}\frac{1}{B_n^{\,3}}\sum_{j=1}^{n}E\left[\left|X_j-\mu_j\right|^3\right]=0$$

This implies that, for $\delta=1$, Lyapounov's condition is satisfied. When we standardize the class identification by $Y_n=\{Y+(n-1)-\sum_{i=1}^{n}\mu_i\}/B_n$, Y_n obeys a standard normal distribution $N(0,\,1)$ based on the central limit theorem.

Proposition 4 has a very strong and general implication for the FK model. Whenever each social status dimension has a finite mean and variance (this is a quite natural assumption), the class identification distribution obeys a normal distribution if the number of dimensions is sufficiently large. This is very important because Proposition 4 allows us to assume any empirical distribution for each dimension of social stratification. As Figure 2 shows, four or five dimensions are enough for approximation by a normal distribution, although the number of dimensions n should be infinitely large in theory. It is important for us to reconsider the assumption of the independence of random variables in dimensions. It is well known that the dimensions of social status are correlated with each other. For example, education and income have a positive correlation while occupational prestige score and income also have a positive correlation. These correlations may affect the fitness of data.

4. Empirical validity of the generalized FK model

In this section, we test the empirical validity of the FK model based on the results from the previous section. We used Social Stratification and Mobility Survey (SSM) data from 1955 to 2005. For time-series comparisons, we used only the male date set. The procedure of the empirical test was as follows.

First, we defined the range of ranks in each dimension. Dimensions such as education are easy to fix the number of ranks because it is naturally divided into finite categories. In order to compare data in time-series, we used three categories for education. The codes were: 1: elementary and junior high school level; 2: high school level; and 3: university and higher level. In contrast, it is not easy to fix the number of ranks for dimensions such as income and occupation because they do not have natural distinctive points for categorization. Therefore, we used different numbers of ranks experimentally. According to previous research, we know that people often view occupation as four or five categories (Shiotani 2010). We used 3 to 8 ranks for occupation. The ranks are computed based on the occupational prestige score of the SSM survey. Incomes were measured using various scales in each survey period. For example, a 10 point scale was used in 1955, a continuous scale in 1965, and a 30 point scale in the 1995 and 2005 surveys. Therefore, we used 5 to 30 ranks for dividing the income range.

Second, the theoretical class identification prediction was computed as follows. Three independent variables were used, i.e., family income, education level, and occupation prestige score. For example, if a subject i has a profile such as (family income 10, education level 2, prestige score 5), we assume that his theoretical class identification is $10+2+5-2=15$ based on Corollary 1. We computed all theoretical prediction of class identification for all possible patterns of combinations of ranks, i.e., sample size $\times$ 6 (patterns of occupation rank) $\times$ 26 (patterns of income rank).

Third, we computed the correlation coefficient between the theoretical class identification prediction from the generalized FK model and the observed class identification data collected by the SSM surveys. As mentioned earlier in this paper, the SSM survey asks respondents to locate their subjective class identification into five categories: Upper, Upper Middle, Lower Middle, Upper Lower, and Lower Lower.

In the following tables, the numbers in each cell indicate the correlation coefficient between the observed class identification from data and the theoretical prediction of the model. Rows correspond to number of rank divisions for occupational prestige while columns correspond to the number of rank divisions of household income.

Table 1 Correlation coefficients for the theoretical class identification predictions of the model and the observed data from the 1955 SSM (males).

Number of ranks of occupation	*Number of ranks of household income* 5	6	7	8	9	10
3	0.378721	0.378921	0.381383	0.380175	0.378539	0.378720
4	0.385464	0.387769	0.390463	0.389768	0.388521	0.389020
5	0.371599	0.378095	0.381955	0.383151	0.383669	0.384581
6	0.360643	0.369531	0.373567	0.376091	0.378210	0.379310
7	0.357685	0.368491	0.372785	0.376281	0.379526	0.381111
8	0.353845	0.364953	0.369191	0.373178	0.377194	0.378731

Table 2 Correlation coefficients for the theoretical class identification predictions of the model and the observed data from the 1965 SSM (males).

Number of ranks of household income

Number of ranks of occupation	5	6	7	8	9	10	11	12	13	14	15	16	17
3	0.347073	0.351024	0.353398	0.355487	0.355041	0.355746	0.353729	0.351272	0.351797	0.349756	0.349041	0.349078	0.350913
4	0.340050	0.346461	0.350294	0.353770	0.354005	0.355633	0.354165	0.352545	0.353258	0.351624	0.351004	0.351304	0.353183
5	0.328911	0.337287	0.342521	0.347140	0.348417	0.350959	0.350425	0.349717	0.350751	0.349743	0.349541	0.350093	0.352203
6	0.322503	0.332087	0.338296	0.343959	0.345831	0.349197	0.349325	0.349568	0.350841	0.350399	0.350478	0.351404	0.353592
7	0.318410	0.328622	0.335300	0.341517	0.343838	0.347455	0.348152	0.348873	0.350379	0.350173	0.350524	0.351618	0.353881
8	0.319416	0.329639	0.336516	0.342946	0.345401	0.349369	0.350307	0.351420	0.352948	0.353051	0.353447	0.354672	0.356951

Number of ranks of occupation	18	19	20	21	22	23	24	25	26	27	28	29	30
3	0.347560	0.347238	0.347147	0.346351	0.344636	0.343792	0.342751	0.342247	0.342837	0.342134	0.341054	0.340449	0.340167
4	0.349988	0.349775	0.349652	0.348902	0.347204	0.346381	0.345406	0.344817	0.345425	0.344740	0.343612	0.342996	0.342706
5	0.349253	0.349245	0.349285	0.348684	0.347166	0.346458	0.345537	0.345070	0.345671	0.345064	0.344038	0.343462	0.343172
6	0.350921	0.351122	0.351246	0.350754	0.349348	0.348763	0.347908	0.347436	0.348065	0.347505	0.346511	0.345950	0.345661
7	0.351365	0.351715	0.351863	0.351487	0.350159	0.349701	0.348840	0.348436	0.349033	0.348497	0.347550	0.347029	0.346699
8	0.354595	0.354985	0.355120	0.354783	0.353488	0.353021	0.352228	0.351754	0.352345	0.351821	0.350844	0.350300	0.349965

Table 3 Correlation coefficients for the theoretical class identification predictions of the model and the observed data from 1975 SSM (males).

Number of ranks of household income

Number of ranks of occupation	5	6	7	8	9	10	11	12	13	14	15	16	17
3	0.216181	0.222304	0.226590	0.223880	0.223649	0.224583	0.224967	0.224365	0.223691	0.227413	0.224773	0.222517	0.222795
4	0.202164	0.211051	0.216679	0.215976	0.217569	0.218137	0.219790	0.220253	0.220055	0.223600	0.221602	0.219941	0.220492
5	0.203985	0.212898	0.218701	0.218397	0.220176	0.220725	0.222462	0.222984	0.222835	0.226261	0.224356	0.222723	0.223262
6	0.198935	0.208440	0.214583	0.215158	0.217734	0.218039	0.220379	0.221445	0.221550	0.224801	0.223289	0.222012	0.222705
7	0.194131	0.204009	0.210414	0.211697	0.215054	0.215170	0.218078	0.219713	0.220042	0.223177	0.222030	0.221132	0.221993
8	0.192511	0.202455	0.208943	0.210780	0.214577	0.214681	0.217957	0.219927	0.220449	0.223472	0.222634	0.221981	0.222965

Number of ranks of occupation	18	19	20	21	22	23	24	25	26	27	28	29	30
3	0.220109	0.221673	0.221153	0.220678	0.220179	0.219516	0.21986	0.219097	0.218839	0.219449	0.219490	0.219122	0.217932
4	0.218223	0.219809	0.219392	0.219206	0.218934	0.218353	0.218782	0.218138	0.218014	0.218670	0.218718	0.218457	0.217334
5	0.221016	0.222509	0.222125	0.221867	0.221525	0.220924	0.221304	0.220640	0.220442	0.221059	0.221092	0.220764	0.219634
6	0.220721	0.222178	0.221870	0.221771	0.221558	0.220994	0.221417	0.220821	0.220694	0.221327	0.221350	0.221077	0.219990
7	0.220285	0.221749	0.221488	0.221584	0.221513	0.221010	0.221483	0.220959	0.220913	0.221572	0.221589	0.221380	0.220339
8	0.221449	0.222903	0.222722	0.222902	0.222906	0.222444	0.222938	0.222453	0.222434	0.223087	0.223119	0.222924	0.221911

Table 4　Correlation coefficients for the theoretical class identification predictions of the model and the observed data from the 1985 SSM (males).

Number of ranks of occupation	*Number of ranks of household income* 5	6	7	8	9	10	11	12	13	14	15	16	17
3	0.249602	0.262136	0.265811	0.275087	0.274646	0.279240	0.279301	0.278391	0.279923	0.279364	0.283050	0.284976	0.283089
4	0.243616	0.257325	0.262176	0.271820	0.272488	0.277546	0.278230	0.277922	0.279744	0.279524	0.283321	0.285356	0.283720
5	0.249107	0.261771	0.266199	0.275333	0.275937	0.280699	0.281357	0.281097	0.282811	0.282638	0.286229	0.288128	0.286533
6	0.232579	0.246834	0.252781	0.262366	0.264825	0.270369	0.272125	0.272982	0.275307	0.275938	0.279762	0.282194	0.281256
7	0.229497	0.243604	0.249875	0.259238	0.262361	0.268082	0.270275	0.271611	0.274139	0.275155	0.279000	0.281581	0.280971
8	0.215211	0.229955	0.237208	0.246793	0.251178	0.257452	0.260513	0.262809	0.265844	0.267619	0.271654	0.274765	0.274824

Number of ranks of occupation	18	19	20	21	22	23	24	25	26	27	28	29	30
3	0.281070	0.282188	0.283885	0.281900	0.281656	0.283318	0.283467	0.283977	0.282972	0.282733	0.281051	0.280985	0.283165
4	0.281893	0.283098	0.284822	0.282975	0.282793	0.284469	0.284626	0.285193	0.284208	0.283975	0.282345	0.282279	0.284457
5	0.284740	0.285880	0.287468	0.285633	0.285419	0.287021	0.287113	0.287634	0.286602	0.286332	0.284704	0.284567	0.286694
6	0.279906	0.281384	0.283262	0.281867	0.281863	0.283606	0.283938	0.284709	0.283754	0.283704	0.282306	0.282262	0.284470
7	0.279839	0.281435	0.283407	0.282231	0.282318	0.284048	0.284482	0.285339	0.284402	0.284426	0.283129	0.283104	0.285308
8	0.274142	0.276106	0.278351	0.277660	0.277963	0.279878	0.280544	0.281669	0.280840	0.281107	0.280081	0.280156	0.282473

Table 5　Correlation coefficients for the theoretical class identification predictions of the model and the observed data from the 1995 SSM (males).

Number of ranks of occupation	*Number of ranks of household income* 5	6	7	8	9	10	11	12	13	14	15	16	17
3	0.323868	0.322506	0.329275	0.325582	0.323249	0.330070	0.322198	0.327029	0.323808	0.328289	0.322230	0.325484	0.321760
4	0.322582	0.322721	0.329464	0.327075	0.325382	0.331615	0.324770	0.329425	0.326490	0.330637	0.325012	0.328275	0.324658
5	0.316764	0.319454	0.326371	0.325316	0.324780	0.330764	0.325077	0.329895	0.327299	0.331524	0.326386	0.329489	0.326227
6	0.314240	0.318414	0.325365	0.325539	0.325977	0.331383	0.326913	0.331772	0.329601	0.333706	0.329143	0.332223	0.329270
7	0.305938	0.312066	0.319110	0.320768	0.322414	0.327670	0.324670	0.329840	0.328185	0.332492	0.328720	0.331756	0.329368
8	0.303795	0.310214	0.317073	0.319450	0.321517	0.326381	0.324102	0.329334	0.327860	0.332068	0.328753	0.331849	0.329673

Number of ranks of occupation	18	19	20	21	22	23	24	25	26	27	28	29	30
3	0.318195	0.3218391	0.322225	0.320547	0.317373	0.319765	0.317400	0.318128	0.316412	0.314984	0.316946	0.314775	0.315041
4	0.321249	0.3246893	0.325023	0.323365	0.320278	0.322533	0.320170	0.320875	0.319109	0.317695	0.319530	0.317368	0.317583
5	0.322986	0.3262530	0.326710	0.325162	0.322158	0.324375	0.322064	0.322732	0.321034	0.319601	0.321406	0.319250	0.319470
6	0.326268	0.3293321	0.329798	0.328332	0.325478	0.327606	0.325336	0.325982	0.324298	0.322881	0.324606	0.322468	0.322621
7	0.326706	0.3296463	0.330255	0.329068	0.326445	0.328594	0.326497	0.327157	0.325594	0.324226	0.325968	0.323894	0.324083
8	0.327288	0.3300480	0.330716	0.329663	0.327175	0.329320	0.327310	0.328002	0.326447	0.325188	0.326824	0.324841	0.325038

Table 6　Correlation coefficients for the theoretical class identification predictions of the model and the observed data from the 2005 SSM (males).

Number of ranks of occupation	*Number of ranks of household income* 5	6	7	8	9	10	11	12	13	14	15	16	17
3	0.364370	0.373050	0.379977	0.383128	0.383615	0.387946	0.386800	0.386803	0.386142	0.387477	0.386776	0.386876	0.385396
4	0.349550	0.360114	0.368037	0.373151	0.375335	0.379707	0.379997	0.380667	0.380664	0.382445	0.382393	0.382965	0.381917
5	0.357971	0.367727	0.375000	0.380025	0.382068	0.386226	0.386665	0.387350	0.387232	0.388813	0.388747	0.389105	0.388062
6	0.337069	0.348545	0.357205	0.364092	0.367915	0.372499	0.374532	0.376280	0.376941	0.379303	0.380011	0.381109	0.380781
7	0.339446	0.350700	0.358980	0.366143	0.370132	0.374507	0.376900	0.378611	0.379418	0.381731	0.382616	0.383726	0.383474
8	0.329467	0.341179	0.349842	0.357758	0.362591	0.367026	0.370249	0.372490	0.373742	0.376456	0.377775	0.379341	0.379461

Number of ranks of occupation	18	19	20	21	22	23	24	25	26	27	28	29	30
3	0.384884	0.386452	0.385528	0.386053	0.385177	0.384683	0.383834	0.383193	0.381597	0.380903	0.382507	0.382416	0.381898
4	0.381807	0.383347	0.382851	0.383515	0.382901	0.382653	0.381920	0.381457	0.379991	0.379465	0.381074	0.381098	0.380653
5	0.387859	0.389403	0.388747	0.389270	0.388574	0.388136	0.387422	0.386786	0.385342	0.384670	0.386166	0.386078	0.385579
6	0.381186	0.382791	0.382788	0.383636	0.383304	0.383287	0.382857	0.382558	0.381258	0.380899	0.382489	0.382590	0.382187
7	0.383926	0.385476	0.385511	0.386287	0.386015	0.385956	0.385568	0.385240	0.383954	0.383591	0.385098	0.385189	0.384777
8	0.380287	0.381840	0.382275	0.383228	0.383199	0.383419	0.383154	0.383049	0.381863	0.381692	0.383250	0.383453	0.383135

For every survey year, the goodness of fit for the generalized model was improved compared with a model given the homogeneity of rank assumption. Tables show that the goodness of fit was relatively high in 1955, 1965, 1995, and 2005, but relatively low in 1975 and 1985. The results of our empirical tests correspond to the findings of statistical analyses in previous studies of social stratification. It is difficult for the FK model to explain why these historical changes occurred.

5. Conclusion

By applying Lyapounov's central limit theorem to probability theory, we relaxed the assumptions of the FK model without loss of generality. In section 2, we succeeded in proving that the assumption of identical uniform distributions is not a necessary condition for deriving a normal class identification distribution from the FK model. Furthermore, in section 3, we proved that the FK model does not have to assume a uniform distribution for each social status dimension in order to derive a normal distribution. In section 4, we showed that a generalization of the FK model can contribute to the sophistication of the model and to the improvement of the empirical validity of the model. We should also emphasize that the FK model can provide a theoretical basis for empirical research on class identification. By Corollary 1, we proved the proposition that class identification can be expressed as a linear combination of social status. In empirical research on class identification, this proposition is a common implicit assumption when using a statistical model for a linear combination of independent variables. The FK model provides a theoretical explanation for using an approximation of a linear combination of social and economic variables when describing class identification. Thus, this theoretical model can contribute to the development of empirical studies.

Acknowledgments

I thank the 2005 SSM Research Committee for allowing me to use the SSM survey data. This work was supported by a Grant-in-Aid for Young Scientists (B) 21730396 and Scientific Research (S) 23223002. The first draft of this paper was presented at the 49th Conference of the Japanese Association for Mathematical Sociology at Ritsumeikan University in Kyoto on March 7–8, 2010. I am grateful to Professor Kenji Kosaka, who guided me on mathematical sociology.

References

Billingsley, Patrick. 1986. *Probability and Measure, 2nd edition.* Wiley.

Fararo, Thomas J. and Kenji Kosaka. 2003. *Generating Images of Stratification: A Formal Theory.* Dordrecht: Kluwer Academic Publishers.

Ishida, Atsushi. 2003. "Cognitive Efficiency and Images of Stratification: FK Model Revisited under Condition of Truncation of Scanning Process." *Sociological Theory and Methods.* 18(2): 211–228.（石田淳．2003.「認識の効率性と階層イメージ－スキャニング打ち切り条件を課した FK モデル」『理論と方法』18(2): 211–28.）

———. 2005. "Mathematical Sociology as an Empirical Science: Objective Stratification System and Verification of FK Model." *Sociological Theory and Methods.* 20(1): 97–108.（「経験科学としての数理社会学－FK モデルにおける客観階層とモデル検証問題」『理論と方法』20(1) 37 号：97–108.）

Karpinski, Zbigniew. 2009. "Testing a Formal Theory of Images of Stratification: A Proposal for a Research Design." *ASK Research and Methods,* Vol.18: 97–122.

Kosaka, Kenji. 2006. *A Formal Theory in Sociology, revised edition.* Harvest Press.（高坂健次．2006.『社会学

におけるフォーマルセオリー　－階層イメージにおける FK モデル改訂版』ハーベスト社）

Maeda, Yutaka. 2011. “Mathematical Model of Class Identification Reflecting Discrimination Process: FK Model with Comparative Reference Group.” *Sociological Theory and Methods.* 26(2): 303–320.（前田豊．2011.「識別過程を考慮した階層帰属意識の数理モデル－比較準拠集団を組み入れた FK モデル」『理論と方法』26(2): 303–20.）

Shiotani, Yoshiya. 2010. “Construction Image of Occupational Status and Status Orientation”, *Sociological Theory and Methods,* Vol.25 No.1: 65–80.

Shirahase, Sawako. 2010. “Japan as a Stratified Society: With a Focus on Class Identification.” *Social Science Japan Journal,* Vol.13, No.1: 31–52.

Shirakura, Yukio and Arinori Yosano. 1991. “Image of stratification and Self-placement in a Class Society.” *Sociological Theory and Methods.* 6(2): 37–54.（白倉幸男・与謝野有紀．1991「階層認知と階層意識－階層認知の数理モデル」『理論と方法』6(2): 37–54.）

Yosano, Arinori. 1996. “Polification of Class Identifying Factors and Self Identification.” *Sociological Theory and Methods.* 11(1): 21–36.（与謝野有紀．1996.「階層評価の多様化と階層意識」『理論と方法』11(1): 21–36.）

A Model of Class Identification:

Generalization of the Fararo-Kosaka Model using Lyapounov's Central Limit Theorem

ABSTRACT

The purpose of this paper is to generalize the Fararo-Kosaka model by applying Lyapounov's central limit theorem to strengthen the linkage between theoretical and empirical studies on class identification. The Fararo-Kosaka model is one of the most significant theories explaining how images of social stratification are generated and under what conditions a distribution of class identification holds a stable form. Yosano improved the Fararo-Kosaka model and proved that class identification as a random variable obeys a normal distribution under the assumption that each social status dimension obeys an identical uniform distribution. We show that these assumptions can be generalized by the application of Lyapounov's central limit theorem and that the class identification obeys a normal distribution irrespective of the distribution of each dimension, where Lyapounov's condition holds. Moreover, we conduct an empirical test by checking the degree of goodness of fit between predictions from the model and observed data from a social survey. We show that our method improves the theoretical universality and empirical validity of the Fararo-Kosaka model.

Key Words: class identification, Fararo-Kosaka model, central limit theorem