

Mathematico-Statistical Papers  
on Sampling Design

I

---

This volume contains several papers on some  
sorts of sampling design.

Content

- 1 Sur un Exemple de la Méthode d'échantillon  
Systématique  
Hiroshi Midzuno
- 2 Sur un Exemple de la Méthode d'échantillon à  
Certaine Cité  
Nisihira Sigeki
- 3 Sampling Design in a Small City  
Masatsugu D. Ishida
- 4 Sampling Design in the Socio-Psychological  
Survey of Abnormal Juveniles  
Chikio Hayashi
- 5 Sampling Design of the Survey Measuring the  
Ability for Arithmetic of the Primary Schools'  
Boys at Yokohama City  
Chikio Hayashi
- 6 Sampling Design for the Survey of Opinions of  
Housewives on Bread  
Hitoshi Kimura
- 7 Sampling Design for Opinion Survey on School  
Recommendation  
Nisihira Sigeki
- 8 Sampling Design in Literacy Survey  
Chikio Hayashi
- 9 Sampling Design in the Social Survey of  
Language at the City of Shirakawa  
Chikio Hayashi and  
Masatsugu D. Ishida
- 10 Sampling Design in the Social Survey of Language  
at the city of Tsuruoka  
Chikio Hayashi

## ABSTRACTS

### 1 Sur un Exemple de la Méthode d'Échantillon Systématique

Hiroshi Midzuno

Dans une méthode d'échantillon systématique, nous devons utiliser la période des nombres des lignes du registre.

Nous deûmes choisir 1/60 dans le registre. Il y avait 20 lignes à chaque page du registre. Alors, nous choisîmes certaine ligne des toutes les trois pages, et nous eûmes 1681. Mais nous eûmes besoin de 1200 environs.

$$1200 = 1681 \times 5/7$$

$$5/7 = 2/3 + 1/3 \times 1/7$$

Alors, une fois nous eûmes jeté 1/3, et nous ressuscitâmes 1/7.  
*ressuscité mis*

### 2 Sur un Exemple de la Méthode d'Échantillon à certaine cité

Nisihira sigeki

Quand nous avons seulement population à chaque quartier dans la cité, nous proportionnons ses échantillons à sa population. Ensuite nous devons décider le point où nous faisons des enquêtes dans chaque quartier. Cependant nous n'avons pas ces mathéiaux frequentment. Dans ce cas, nous parta-gons la quartier en petits carrés sur la carte, et nous choisissons un carré.

Ce n'est pas la méthode d'échantillon par region, car des nombres d'échantillon à chaque quartier proportionnent à sa population et le carré choisi n'est que le point de départ.

3

### Sampling Design in a Small City

Masatsugu D. Ishida

A survey of sanitary status was held in the city of Mito, with population about 60,000. Whole city was divided into 25 blocks, containing about the same number of households.

Sample households were chosen at random on the map.

### 4 Sampling Design in the Socio-Psychological

#### Survey of Abnormal Juveniles

Chikio Hayashi

This survey will be performed as a pre-test. The aim is not only to estimate the total numbers of several social and psychological types of abnormal juveniles and to analyse the multi-dimensional factors influencing the ratio of them to the total but also to find out their psychological characteristics by using the projective tests.

Thus survey is divided into 2 parts

- (1) Estimation of the numbers
- (2) Finding the characteristics

In (1) simple sampling method, and in (2) sub-sampling method are used.

Primary sampling units are schools, and Secondary sampling units are abnormal juveniles, the numbers of which are previously unknown in schools.

The population is constructed by giving the equal sampling probabilities to the primary sampling units and giving the equal sampling probabilities to the secondary sampling units in the selected primary units. The stratification is performed with respect to the primary units. We construct the strata to make as homogeneous in the total member of students and in the ecological state of the site as possible, because the homogeneous strata will be made by doing so from the statistical point of view.

12 strata are made and only one primary sampling unit is sampled from every stratum. 200 or 300 abnormal juveniles will be tested. The coefficient of variation of the estimated value of the aiming characteristics in this sampling (in (1) or (2)) is about 0.05, (where the several estimated values are quantified from the statistical and psychological stand point).

5 Sampling Design of the Survey Measuring the Ability for Arithmetic of the Primary Schools' Boys at Yokohama City

Chikio Hayashi

The universe of our survey consists of all the boys in the primary schools at Yokohama City, whose labels are test scores of arithmetic by some tests.

Sub-sampling method was adopted. Primary sampling unit is a school, secondary sampling unit is a boy. Stratification was made.

Fortunately the results of achievement tests at Yokohama City that were obtained by random sampling method at 1949, are used for stratification. The data of all schools concerning some labels can be used.

The results of sampling survey are analysed and the relations between the ability for arithmetic and the labels mentioned above have been obtained in the sense of multiple regression theory.

$$\bar{y} = \sum_{i=1}^{10} a_i x_i + a_0$$

Where  $\bar{y}$  is an estimated value of ability scores,  $x_i$  is a factor influencing on the ability score.

Thus the averages of test scores of all schools are estimated. Using these data, all schools are stratified into 10 strata. Next, boys are drawn from the schools sampled. Boys are stratified by using the intelligence test scores, and sample boys of size  $n'$  (previously determined) are drawn. This gain is due to the correlation coefficient,  $\rho$ , between ability for arithmetic and intelligence being high.

Suppose that population mean  $\bar{X}$  of test scores of

arithmetic will be estimated

For this sample mean  $\bar{x}$  is used

$$\bar{x} = \sum_{i=1}^R p_i \bar{x}_i$$

Where  $\bar{x}_i$  is the sample mean of  $i$  stratum,  
 $N_i$  is the population size of  $i$  stratum,

$$p_i = \frac{N_i}{\sum_{i=1}^R N_i}$$

$R$  is the number of strata.

Let  $\sigma_{\bar{x}}$  be the variance of  $\bar{x}$ .

The allocation of sample to each stratum is decided to be proportionate to the population size of each stratum.

Only one school is drawn from each stratum with probability proportionate to the size of school,

boys of size  $n \cdot \frac{N_i}{\sum_{i=1}^R N_i}$  are drawn from the school sampled of  $i$  stratum with equal probability.

Under some proper hypotheses,  $\sigma_{\bar{x}}^2$  is approximately as follows.

$$\sigma_{\bar{x}}^2 \doteq a \cdot \frac{\sigma^2(1-p^2)}{n} + \sum_{i=1}^R p_i^2 \sigma_{b_i}^2$$

Where  $a$  is sampling ratio in the sense of mean,  
 $n$  is sample size,

$\rho$  is the correlation coefficient between ability scores for arithmetic and intelligence test scores,

$\sigma_{b_i}^2$  is the variance of ability scores of boys  
 $\sigma_{b_i}^2$  is the between variance of  $i$  stratum.

Let  $\bar{X}$  be the population mean.

The coefficient of variation of  $\bar{x}$  is numerically obtained, where sample size is 700 that is decided by some statistical considerations.

$$\frac{\sigma_{\bar{x}}}{\bar{X}} = 0.015$$

Under the level of confidence 95%, the confidence interval is 0.03 multiplied by the population mean.

From this, the precision of this sampling design is considered to be high.

6      Sampling Design for the Survey of Opinions  
         of Housewives on Bread

Hitoshi Kimura

This survey aims to secure opinions of housewives on bread. Results should be obtained in 2 days. And area sampling was employed in Metropolitan area of Tokyo. According to the view that opinions are controlled by circumstances, ku(ward) were stratified by industrial character. Sample spots, cho(quarter), were chosen systematically from list of cho in ku.

Sample housewives were chosen as follows :  
Every 5-th house from the house, at which interviewer at first recognized being in his sample spot, were sampled. At a corner, the direction to advance was selected at random.

7      Sampling Design for Opinion Survey on  
         Schoup Recommendation

Hitoshi Kimura

Opinion survey on Schoup Recommendation was held in Kanto area. In this survey, stratified sub-sampling method was employed. Strata of wards, cities, town and villages were constructed by combining "atoms" of strata which had been constructed with respect to industrial characters from the data of 1947. Census.

In Opinion survey to peasants, strata were constructed by combining the strata above regarding industrial character.

Assuming that the ratios of peasants to the population in spots were constant in every stratum, method of common sample spots to two survey, expressed in preceding volume, was employed.

8      Sampling Design in Literacy Survey

Chikio Hayashi

The Literacy Survey was performed under the direction of the Committee organized for this purpose. This short report is that of sampling design in this survey, a part of works done by the statistical committees. Under some restrictions, for example, time, expenses, or testers required, this design was made.

(i) The form of this Survey

The test materials having been constructed, the testees are to be made answer by filling in each entry of the test paper, instructions given in the same way as in a general intelligence test. In order to carry out the test as strictly as possible, the testees should be tested under the same condition. To do this, they are to be called to a certain place, where the test with its instructions is given within certain limits of time.

(ii) Sampling

Next, making use of the results obtained from the pretests, we established a sampling design and a concrete plan for the survey. It is rather difficult to draw each individual with equal probability in such a large scale sampling. For this, it is very useful to divide the universe (the Japanese) into homogeneous groups. Then we have to consider the factors affecting the abilities. The factors which are taken into consideration, for the stratification are naturally assumed and known by the pretest.

These are followings.

(A) Japanese Area, (B) Urban and Rural

Area : This reflects such characteristics as population size and density, economic-ecological structure and degree of urbanization.

Thus the numbers of strata are 87 in urban area and 86 in rural. The sub-sampling method from these strata was adopted. The next step is to determine the number of testees to be sampled and to allocate to each test spot. In order to obtain the required precision, necessary sample size is 17,100. Including nonresponses which may not give any bias to the whole, the sample size was determined to be 21,008. Having fixed the sample size in this manner, we proceeded to allocate them. This sampling design, however, is intended for estimating the distribution of population in terms of scores. Allocation such as is optimum for a single label is considered to be improper in the survey consideration which is so

complicated that the reading and writing abilities can be measured only by many criteria combined. Therefore, in view of the conveniences of the counting, the samples were allocated in proportion to the population in each stratum.

9            Sampling Design in the Social Survey of

Language at the City of Shirakawa

Chikio Hayashi and Masatsugu D. Ishida

The purpose of the survey is to make clear the linguistic factors which disturb rationalization of national social life. In this, interview method is adopted.

Under the real conditions the stratified and double sampling method is used at the sample size 500. And several devices are made concerning with this sampling survey (design and performance)

10           Sampling Design in the Social Survey of

Language at the City of Tsuruoka

Chikio Hayashi

The aim of this survey is same as that of the City of Shirakawa. In this survey double sampling method is used. And some devices concerning with the allocation of sample have been made.

Considering the several restrictions, sample size is decided to be about 500. In this case sample size is a random variable, because of some devices used.

For the convenience of analyses of the results, sample size of high age is designed to become large, and sample size of low age is to become small. But the operation of counting must be of course simple and total sample size must be not more than 500. This is the important point of this survey.

In this survey, Rorschach test is used in order to make clear the relation between the degree of using common language and personality.

This is a number of a projected series of reports entitled "The Research Report of the I. S. M." "The Research Report of the I. S. M." publishes the reports of researches done in the application of Statistical Mathematics such as initial preparation, study designs, practical procedures and handling of data.

The series aims to be beneficial for research workers who are engaged in the practical problems of Surveying, Analysis and so on, rather than for the theoretical workers.

The series will be published monthly or bi-monthly and will contain reports on a variety of problems.

<b>Editor</b>	Hiroshi Midzuno
<b>Published by</b>	The Institute of Statistical Mathematics 10, Sangenjaya-cho, Setagaya-ku, Tokyo
<b>Printed by</b>	Sobunsha Co. 13, Takata toyokawa-cho, Bunkyo-ku, Tokyo