Sampling Techniques Applicable to Studies on Rice-Stem Borer Egg-Mass Populations*

Akio ÔTAKE**

(Laboratory of Applied Entomology)

ニカメイガ卵塊個体群の研究に適用しうるサンプリング法

大 竹 昭 郎 (応用昆虫学研究室)

During the three years 1957—1959, the writer made observations on the egg mortality of the rice-stem borer, *Chilo suppressalis* WALKER, in its first brood, in a paddy field at Nogi, Matsue City. Every observation day in the first two years, a new map was drawn showing the location of rice-bunches which bore, on that day, egg masses not yet hatched or just after hatching in a plot established in the field. In 1960, a map of the same kind was made only once.

The writer intends, in this paper, to investigate whether some of these maps answer the following two questions:

1) Whether the sequential sampling technique is applicable to classifying given paddy fields according to population levels of egg masses existing in them.

2) Whether a kind of the spacing method as proposed by MORISITA (1957) is adequate for estimating the population density of egg masses in a given paddy field.

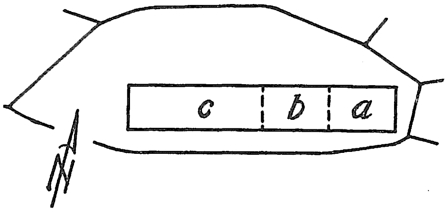
Making of Maps Showing the Distribution of Egg Masses in the Plot

In the paddy field where the plot was established in 1957, rice-bunches had regularly been planted at distances of 24 cm. from each other. The plot was large enough to contain 600 bunches at first and in the next year it was enlarged to double size (14.4 m. ×4.8 m.) towards the west (cf. Fig. 1). In each of these two years, egg masses deposited in the plot were searched at intervals of one or two days through their first occurrence period (from the end of May to the beginning of July). In 1960, the plot was again enlarged to double size (28.8 m. ×4.8 m.) and such a search was made only once on the 20th of June.

In 1957 and 1958, every egg mass discovered in the plot was marked by a stick and wire set beside the bunch bearing it, and at the same time the location of the bunch was recorded on a map where each of all the bunches planted in the plot was expressed as a point. Such a map was newly made every observation day to show the existence of egg masses. It was not until the ascertainment of their hatching that they were taken away from the bunches and ceased to

* Contributions from Laboratory of Applied Entomology, Shimane Agricultural College, No. 31.

** Now at: Hokuriku National Agricultural Experiment Station, Takada, Niigata Prefecture.



be recorded on the map.

Among the maps thus made, those of the following three dates are taken up in this paper: the 10th and 19th of June, 1958 and the 20th of June, 1960 (Figs. 1−3). Of the two of 1958, the former was chosen because it came in the early emergence period of

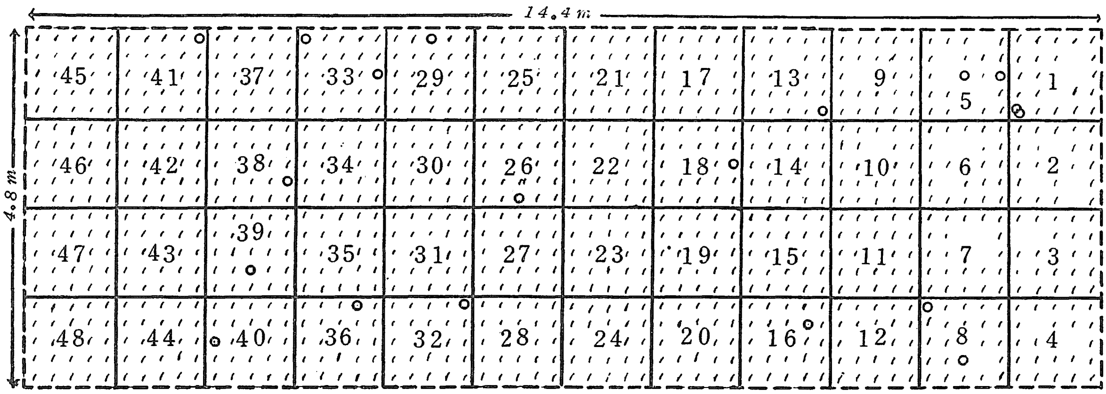


Fig. 1. Distribution of rice-stem borer egg masses in the plot, observed on the 10th of June, 1958. On the left upside is shown the situation of the plot in the paddy field; *a*, $a+b$ and $a+b+c$ are the areas observed in 1957, 1958 and 1960, respectively. A dot is a rice-bunch bearing no egg mass. A hollow circle represents a mass deposited on a bunch. Th plot is divided into 48 units by solid lines, each of them being numbered for making a sequential analysis.

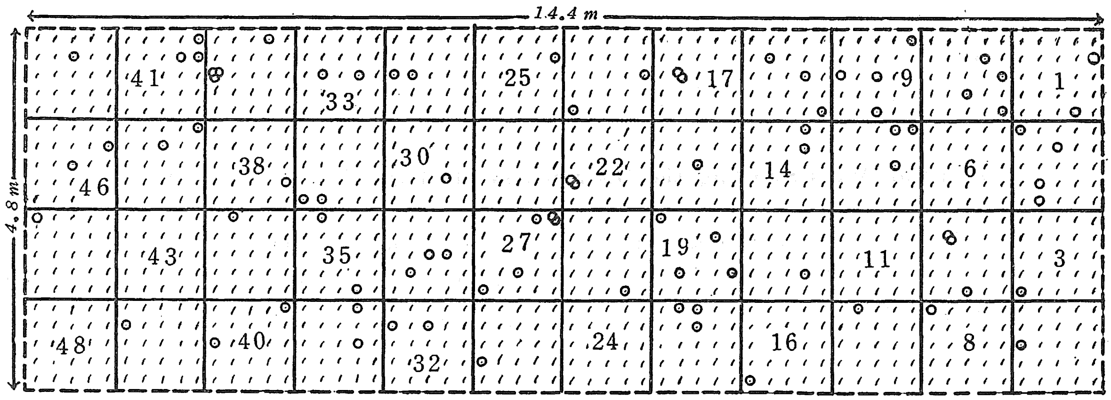


Fig. 2. Distribution of rice-stem borer egg masses in the plot, observed on the 19th of June, 1958. The symbols are the same as in Fig. 1.

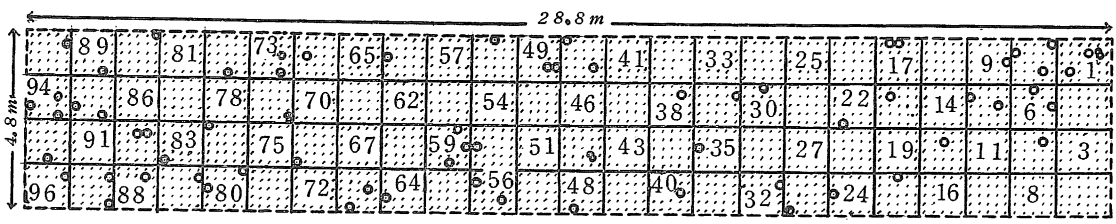


Fig. 3. Distribution of rice-stem borer egg masses in the plot, observed on the 20th of June, 1960. The plot is divided into 96 units. The symbols are the same as in Fig. 1.

moths of the borer in its first brood and the latter in the period when egg masses were most abundantly existing in the plot. The data obtained in 1957 are not used here, the reason for which is that area observed in that year was so small that the data seemed to be insufficient for giving any statistical analysis.

Application of Sequential Analysis

The sequential sampling technique has often been used, mainly in forest entomology, as a rapid method for predicting the infestation which may be caused by a given insect pest, for determing the degree of parasitism by a given parasite and for other purposes (MORRIS, 1960). Involving no fixed number of samples is a distinctive feature of this method. In this paper, the writer will explore the possibility of applying it to classification of paddy fields by population levels of rice-stem borer egg masses.

(1) *Determination of the type of frequency distribution.* ···The first step in the application of the sequential technique is to determine the type of frequency distribution. We devide the plot into squares containing 25 rice-bunches (1.2 m.×1.2 m.) as shown in Figs. 1−3 and assume

Table 1. Frequency distributions of rice-stem borer egg masses obtained from all the units in Figs. 1−3. Agreement of observed counts with those expected by Poisson series is tested by chi-square.

No. of egg masses in an unit	June 10, 1958 (Fig. 1)			June 19, 1958 (Fig. 2)			June 20, 1960 (Fig. 3)		
	Observed frequency	Expected frequency	χ^2	Observed frequency	Expected frequency	χ^2	Observed frequency	Expected frequency	χ^2
0	33	32.30	0.015	6	7.86	0.438	47	45.35	0.041
1	11 } 15	12.79 } 15.70		16	14.22	0.223	31	34.01	0.266
2	4 }	2.53 }	0.031	14	12.87	0.100	14 } 18	12.75 } 16.64	
3		⋮ }		6 } 12	7.76 } 13.05		3 }	3.19 }	0.111
4				5 }	3.51 }	0.086	1 }	0.60 }	
5				1 }	1.27 }			⋮ }	
					⋮ }				
Total	48	48.00	0.046	48	48.00	0.847	96	96.00	0.418
Average	0.396		d.f.=1 P> 0.80	1.813		d.f.=3 P> 0.80	0.750		d.f.=2 P> 0.80

each of them as an unit. Then we can get, in each of the figures, a frequency distribution exhibited by the egg mass count in the whole units. As shown in Table 1, all the distributions conform very well to the Poisson, $m^r e^{-m} / r\,!$

(2) *Calculation of acceptance and rejection lines.*···The drawing of acceptance and rejection lines is the point in the sequential plan; the accumulated total of egg masses counted in sampling units which are chosen at random in a paddy field is plotted as a series of points on a graph in which the lines have been drawn, and when these points cross either of the lines, the population level of egg masses is determined and sampling can be discontinued, as far as the field is concerned.

The following formulae are used in obtaining the lines for a Poisson distribution (MILLER, 1955):

$$d_1 = -h_1 + sn$$

and

$$d_2 = h_2 + sn$$

where

$$h_1 = \frac{b}{\log_e m_2 - \log_e m_1} \quad \text{and } b = \log_e \frac{1-\alpha}{\beta},$$

$$h_2 = \frac{a}{\log_e m_2 - \log_e m_1} \quad \text{and } a = \log_e \frac{1-\beta}{\alpha},$$

$$s = \frac{m_2 - m_1}{\log_e m_2 - \log_e m_1},$$

d_1 and d_2 = cumulative number of egg masses,

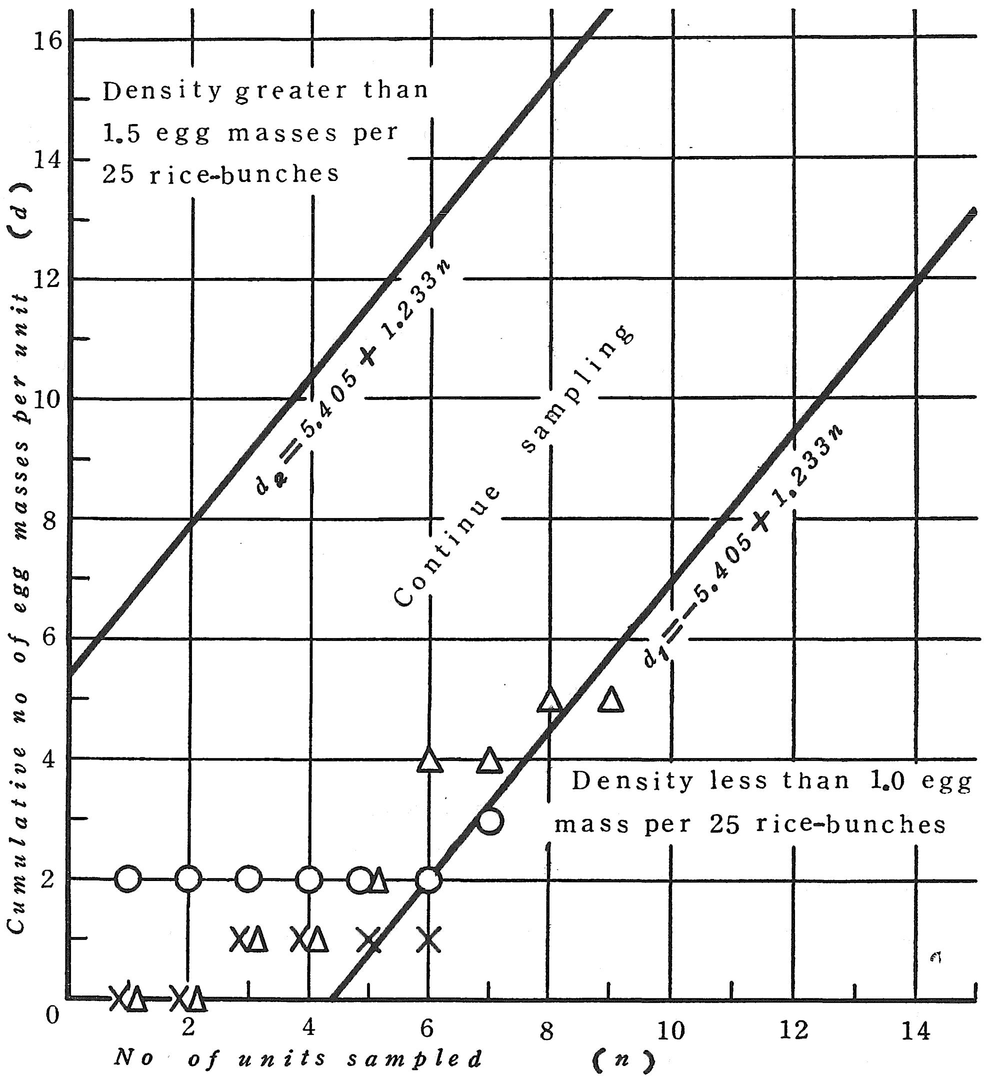


Fig. 4. Graph of sequential sampling plan. Accumulated number of rice-stem borer egg masses counted in units randomly chosen from Fig. 1 (see text). ○ : Series a_1; △ : Series a_2; × : Series a_3.

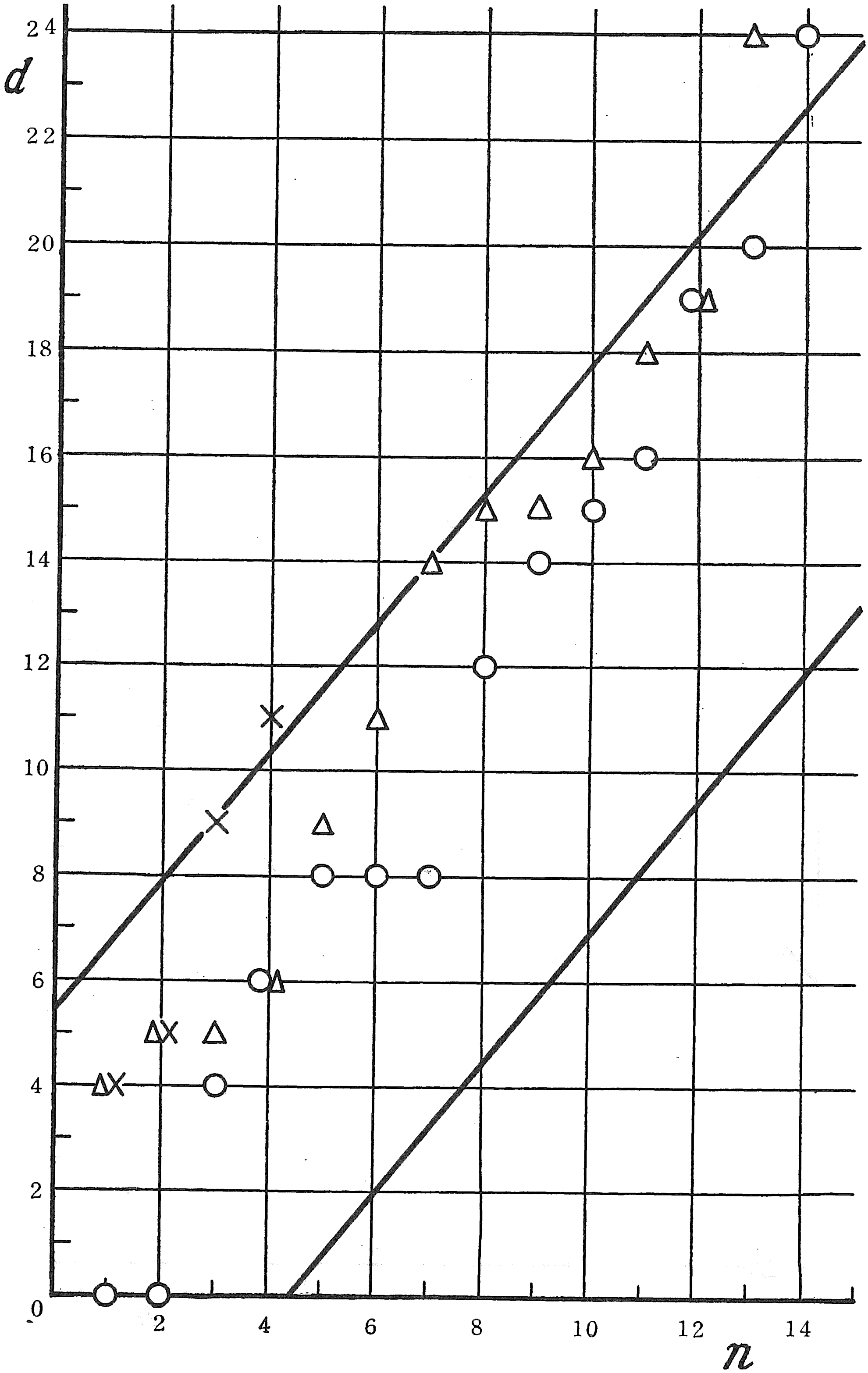


Fig. 5. Graph of sequential sampling plan. Accumulated number of rice-stem borer egg masses counted in units randomly chosen from Fig. 2. ○ : Series b_1 ; △ : Series b_2 ; × : Series b_3.

n=number of sampling units containing 25 rice-bunches,

h_1 and h_2=intercepts of one set of lines,

s=slope of the line,

m_1=the smaller mean,

m_2=the larger mean,

α=the probability of accepting $m=m_2$ when $m=m_1$,

β=the probability of accepting $m=m_1$ when $m=m_2$.

The set of lines shown in each of Figs. 4—6 are got by substituting the following values in the above formulae:

$m_1=1.0$, i.e., the density of one egg mass per 25 rice-bunches,

$m_2=1.5$.

This means that an egg-mass population is determined to be on either of the two levels : densities less than 1.0 ($m<1.0$) and greater than 1.5 ($m>1.5$).

α and $\beta=0.10$, i.e., 10 chances in 100 that a given paddy field will be incorrectly classified as either of the two population levels.

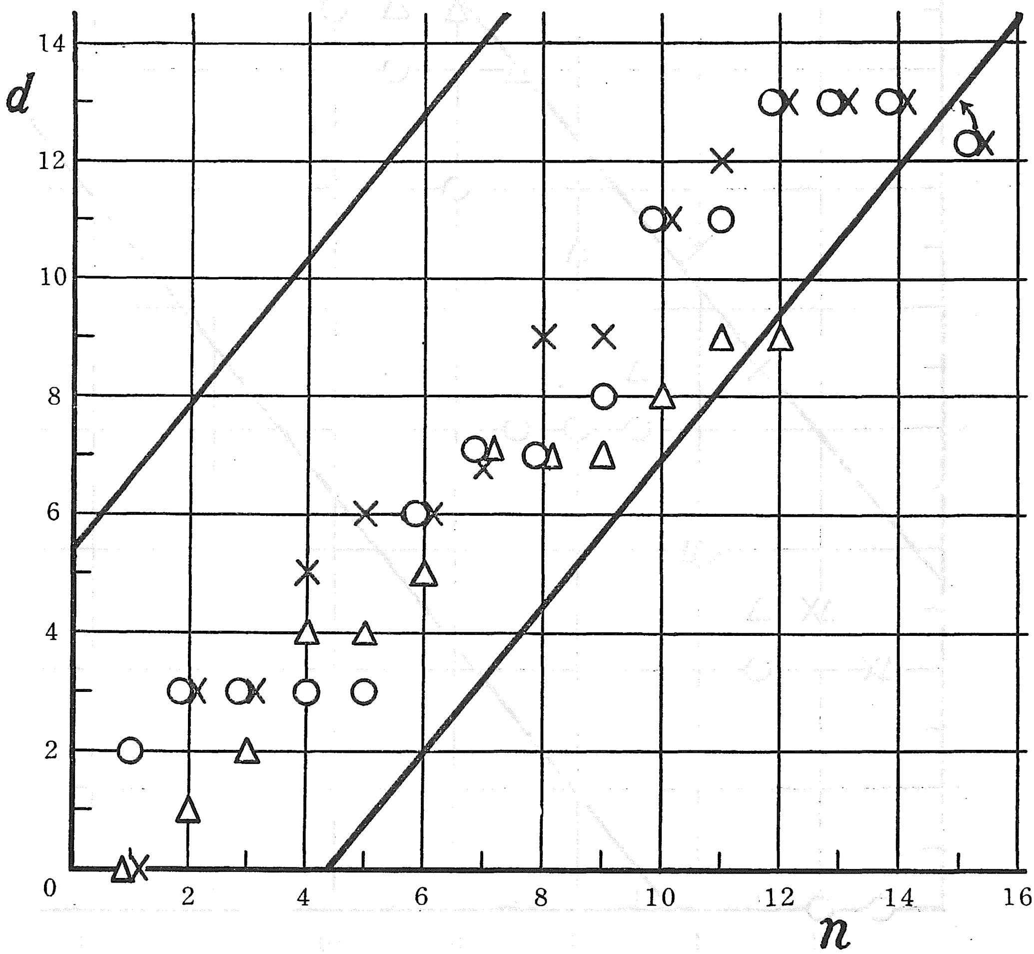


Fig. 6. Graph of sequential sampling plan. Accumulated number of rice-stem borer egg masses counted in units randomly chosen from Fig. 3. ○ : Series c_1 ; △ : Series c_2 ; × : Series c_3.

(3) *Sampling made on the maps.* ··· As shown in each of Figs. 1−3, the whole units are numbered consecutively. Then, using a table of random numbers, we draw out the following three series of unit numbers from each of the figures:

Fig. 1:	a_1	33	6	3	44	45	20	18	17	32	8	41	11	37	23	22	......
	a_2	22	28	16	10	41	33	9	32	20	34	21	23	35	45	12	......
	a_3	46	30	26	31	27	42	18	48	20	38	10	35	46	3	11	......
Fig. 2:	b_1	48	6	19	29	22	43	26	5	42	15	38	10	4	2	45	......
	b_2	2	3	43	39	7	21	20	12	48	18	34	47	27	11	17	......
	b_3	37	44	2	34	17	26	31	28	3	11	25	40	21	20	35	......
Fig. 3:	c_1	74	69	60	72	11	5	71	62	53	94	65	17	36	31	23	......
	c_2	81	64	69	6	86	15	49	11	78	28	95	33	12	91	50	......
	c_3	39	94	4	10	34	60	28	49	76	73	18	38	2	75	62	......

Now, assuming that the units arranged above are to be sampled successively, we proceed on Figs. 4−6 to plot the cumulative numbers of egg masses counted in them.

In Series b_3, for example, the first unit sampled, No. 37, contains 4 egg masses and then a mark is made on Fig. 5 at [$n=1$, $d=4$]; the second unit, No. 44, has one mass, so a mark of $4+1=5$ masses is plotted over $n=2$, leaving us still in the band, "Continue sampling." A mark of the cumulated 11 masses plotted over $n=4$ puts us outside the band and into the area of "Density greater than 1.5 egg masses" and sampling is discontinued. Here, only by taking up the first four units we can state (in 90 per cent confidence) that the average number of egg masses per 25 rice-bunches in the plot was above 1.5 on the 19th of June, 1958. This statement is not contradictory to the result shown in Table 1 ($\overline{m}=1.813$).

In other series, generally speaking, much more units are necessary to determine the population level of egg masses as seen in Figs. 4−6, the maximum being 15 units as in Series b_1, c_1 and c_3. In any case, however, we can recognize that the determination of population level in this way is regarded as appropriate, judging from the actual mean density (cf. Table 1).

(4) *Calculation of the average sample number.*···While the sample size is not fixed in a sequential plan, an average sample number for examining densities at a specified level may be given. In the Poisson case the formulae which give this number are as follows (MILLER, 1955)*:

$$\text{For } m=m_1,\ \bar{n}_{m1} = \frac{(1-\alpha)h_1-\alpha h_2}{s-m_1}\ ;$$

$$m=m_2,\ \bar{n}_{m2} = \frac{(1-\beta)h_2-\beta h_2}{m_2-s}\ .$$

Here we have

$$\bar{n}_{m1} = \frac{(1-0.1)\times 5.405-0.1\times 5.405}{1.233-1.0}=18.56$$

$$\bar{n}_{m2} = \frac{(1-0.1)\times 5.405-0.1\times 5.405}{1.5-1.233}=16.19.$$

As m approaches the slope of the line, $s(=1.233)$, the average sample number increases, and in plotting the cumulative number of egg masses on a graph of sequential plan, we can

*See also GOULDEN (1952, p. 434)

see that since neither line is crossed, density of masses per 25 rice-bunches is greater than 1.0 but less than 1.5.

Application of MORISITA's Spacing Method

(1) *Application procedures.*······Differing from the sequential sampling technique, the spacing method proposed by MORISITA (1957) is used for estimating the population density itself. This method has an outstanding advantage which may be applicable to any population, distributed either randomly or non-randomly, without determining the type of the frequency distribution beforehand.

The application procedures of the method are as follows:

1) Place N sampling points randomly or regularly*, on an observation area, A.
2) Divide the circle of infinite radius surrounding each sampling point into k sectors ($k \geqq 4$) and measure the distance, r, from the n-th nearest individual in each sector ($n \geqq 3$).
3) By putting here

$$\hat{u}_1 = \frac{n-1}{N} \sum_{i=1}^{N} \sum_{j=1}^{K} \frac{1}{r^2_{ij}}$$

and

$$\hat{u}_2 = \frac{nk-1}{N} \sum_{i=1}^{N} \frac{K}{\sum_{j=1}^{K} \sum r^2_{ij}}$$

$\hat{u}_1$ and $\hat{u}_2$ become the unbiased estimates of u for the random distribution in the case of random sampling,

$$u = \pi \frac{\mathrm{T}}{\mathrm{A}},$$

T=total number of individuals in A,

T/A=population density per unit area.

In the case of regular sampling, too, u can be estimated without a strong bias by calculating the values of $\hat{u}_1$ and $\hat{u}_2$.

As for the non-random distribution,

$$u_0 = \frac{\hat{u}_1 + \hat{u}_2}{2} \quad (\text{when } \hat{u}_1 < \hat{u}_2)$$

or

$$\hat{u}_1 \quad (\text{when } \hat{u}_1 > \hat{u}_2)$$

may be used as an estimate of u. When N is small, however, u_0 calculated from $\sqrt{\hat{u}_1 \cdot \hat{u}_2}$ will be more appropriate to the estimation of u than that from $(\hat{u}_1 + \hat{u}_2)/2$.

---

*In its practical use in the field, regular sampling will usually be less laborious than random sampling although it holds a potential danger that it may bring about a statistically meaningless result if the organisms to be dealt with take a spatial distribution with high regularity. As pointed out by MILNE (1959), however, it may be said that in a field population, there scarcely occurs such a kind of regularity as makes it impossible to expect it in advance and take adequate procedures for it.

4) The estimated density per unit area is thus obtained from dividing the estimate of u by π, the total of individuals being estimated by multiplying the quotient by A.

In the course of practical application, the following should be noted:

1) For the convenience of ready working it will be proper to take $k=4$ and $n=3$ (cf. Fig.7).

2) When there are two or more individuals with the same distance from the sampling point in a sector, they are regarded as in the same order. By numbering them as one the third nearest individual is determined (cf. the second quadrant in Fig. 7).

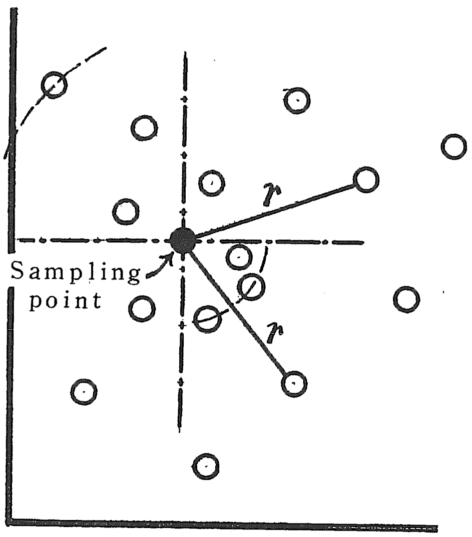


Fig. 7. Schema drawn in explanation of practical application of MORISITA's spacing method (see text).

3) We cannot use a sector in which r is greater than the shortest distance from the sampling point to the border line of the area (cf. the third quadrant in Fig. 7) or in which the third nearest individual itself does not exist (cf. the fourth quadrant in Fig. 7). When this happens, $\hat{u}_1$ and $\hat{u}_2$ are calculated from the following equations:

$$\hat{u}_1 = \frac{k(n-1)}{K} \sum_{i=1}^{N} \sum_{j=1}^{k'} \frac{1}{r_{ij}^2}$$

and

$$\hat{u}_2 = \frac{\sum_{i=1}^{N} k' u_{2.i}}{K}$$

where K=the total of used sectors,

k'=the number of sectors used at each sampling point.

$$\hat{u}_2.i = k\,(k'\,n-1)\frac{1}{\sum_{j=1}^{k'} r_{ij}^2}$$

(2) *Application of the method on maps.*……Now, we attempt to apply the method to Fig. 2 representing the egg-mass population on the 19th of June, 1958, the density of which is highest

Table 2. Estimation of the total egg masses existing in the plot on the 19th of June, 1958, by the use of the spacing method in which the sampling points are arranged as shown in Figs. 8A, B and C. Eighty-seven is the actual number of egg masses existing in Fig. 8.

Arrangement of sampling points	No. of sampling points	$\hat{u}_1$	$\hat{u}_2$	Estimated value: $\hat{u}_0 = \sqrt{\hat{u}_1 \cdot \hat{u}_2}$ (when $\hat{u}_1 < \hat{u}_2$) or $\hat{u}_1$ (when $\hat{u}_1 > \hat{u}_2$)	Estimated value: Total no. of egg masses
Fig. 8A	7	3.63493	3.81997	3.72631	82.0
Fig. 8B	12	4.29481	4.20488	4.29481	94 5
Fig. 8C	8	3.51361	3.93601	3.71882	81.8

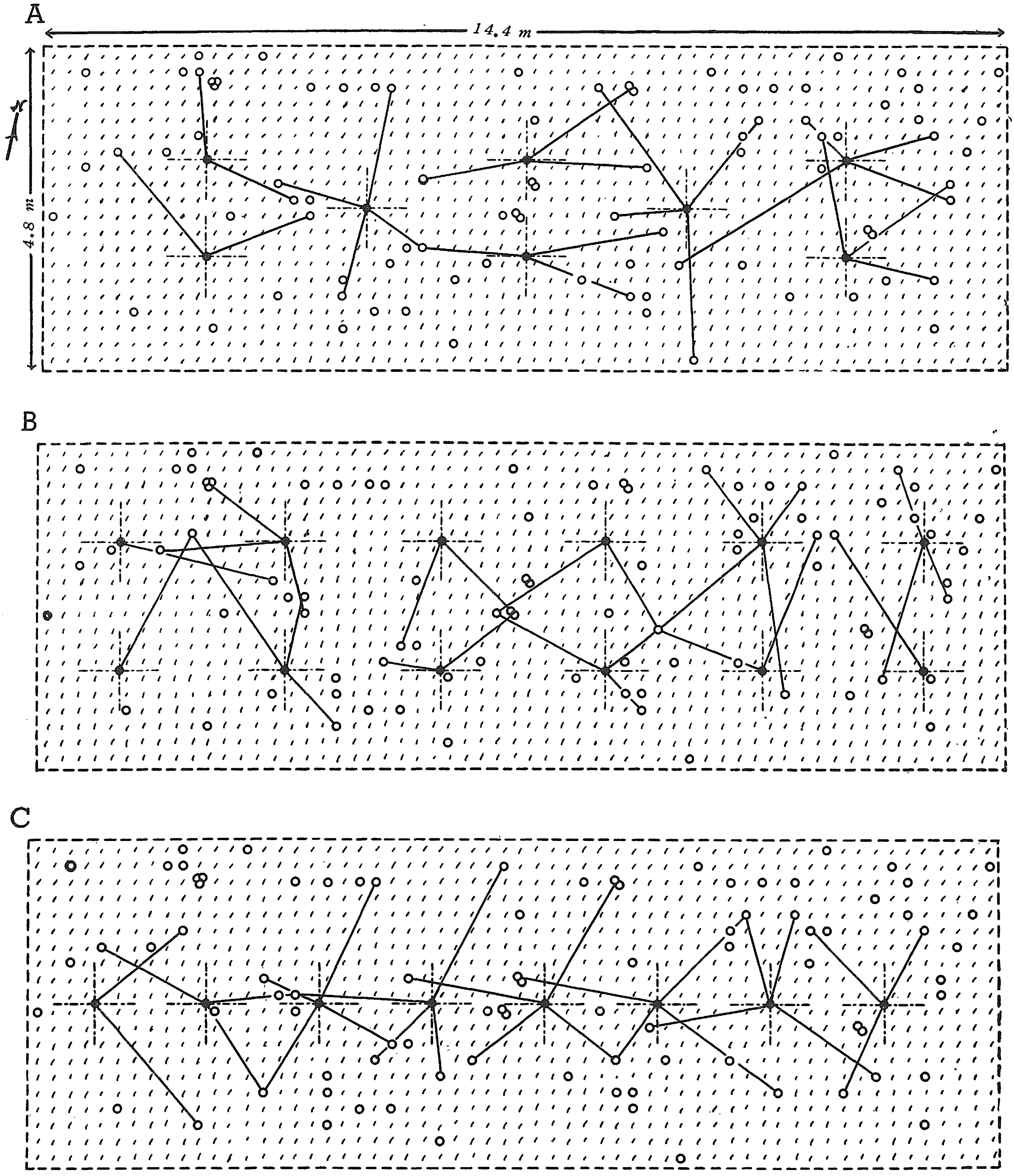


Fig. 8. Application of the spacing method on the map showing the distribution of egg masses observed on the 19th of June, 1958. Dots and hollow circles are the same as in Fig. 2. A solid circle represents a sampling point, which gives four sectors. A hollow circle connected with a sampling point by a line is the third nearest egg mass in the sector concerned. For particulars, see text.

among the three populations shown in Figs. 1—3. Three sets of sampling points are individually put on Figs. 8A, B and C, in each of which the points are differently arranged among rice-bunches. The number of sectors is determined as four ($k=4$) and the distance in metre is measured from the third nearest individual ($n=3$). Since every set contains points with unused sectors; $\hat{u}_1$ and $\hat{u}_2$ are calculated from the equations given at the end of the preceding subsection. Estimated totals of egg masses on the plot are shown in Table 2, together with calculated values of $\hat{u}_1$ and $\hat{u}_2$ and estimates of u.

Though there is rather a small number of sampling points taken, each of the above three estimated totals agrees pretty well with the actual total, 87. That obtained from Set B is remotest from the actual value. This may be attributable to the fact that in this set there are a considerable number of unused sectors; MORISITA (1957 and personal communication) points out that the existence of a lot of unused sectors often makes the precision of the estimation reduce, when the organisms are more or less regularly distributed on an area. In our plot the distribution of egg masses may have been somewhat affected by regularity in the arrangement of the rice-bunches*.

Discussion and Conclusion

In the sequential analysis, it is necessary, as mentioned already, to determine the type of frequency distribution in advance; as for Figs. 1–3, the Poisson was very well applied to the egg-mass populations shown in them, so that it seems to be regarded as generally applicable to any population in a paddy field where the egg-mass density remains at least below the highest of values shown in Table 1 (about two egg masses per 25 rice-bunches). In a case where a much higher density has been attained in a field, we have no clue as to what type of frequency distribution the egg-mass population has in that field. The writer's experience, however, tells that in the Matsue district from late spring to early summer even if there exist paddy fields bearing egg masses in such a high density, the frequency of their appearance will be perhaps very low. The formulae given in this paper for calculating the acceptance and rejection lines will therefore be brought into general use in our district at least, for classifying paddy fields quickly according to population levels of rice-stem borer egg masses existing in them.

The sequential analysis proves its intrinsic merit when it is employed to predict the infestation of a given insect pest by the abundance of the species at any of its earlier developmental stages, as shown, for example, by MORRIS (1954) on the spruce budworm and by CONNOLA *et al.* (1957) on the forest tent caterpillar. In the case of the rice-stem borer, if it is possible to obtain the data which clearly show, for example, the relationship between the maximum density of egg masses recorded during their occurrence period in a paddy field and the damage caused by the borers originating in the masses, we may successfully predict their infestation from the cumulative number of egg masses found in units sampled successively. Through his observations made in the plot during 1957–1959, however, the writer could not gain any information about the existence of such a kind of relationship, so we could, in this paper, get no further than appyling the sequential sampling method to the classification of egg-mass populations by their density.

Through its application on the map (Fig. 8) showing the distribution of egg masses on the 19th of June, 1958, the spacing method as proposed by MORISITA (1957) seems to prove its general usefulness in estimating the density of egg masses (per 1 m^2.) in a paddy field. As

*In spite of that apprehension, we dared to make regular sampling because the regularity which might exist in the distribution of egg masses seemed to be very low.

guessed from Fig. 8, however, in a field where egg masses are not so abundant, a "third nearest mass" (cf. p. 216) obtained by this method will be generally distant from the sampling point thus compelling us to take a rather hard labor when searching it and measuring the distance. This method will therefore be most favorably applicable in the case of a high density of egg masses: the more densely egg masses exist in a field, the shorter the distance to be measured and accordingly the less laborious the method.

In this paper, the distance to a "third nearest mass" was measured on the maps in which rice-bunches were reduced to points. This means that the measurement was made between the sampling point and the centre of a rice-bunch bearing the mass, regardless of its situation in the bunch. Such an easier method was permissible because at the time of observation the rice plants had not been developed so large yet. It is needless to say that we must measure the distance up to the actual place on a plant where a "third nearest mass" is deposited, in later season when plants have been developed so large that they come to cover almost completely the surface of the field.

Acknowledgement: The writer is deeply indebted to Dr. G. B. OAKLAND, Dominion Bureau of Statistics, Canada Department of Trade and Commerce, Ottawa, and Dr. C. A. MILLER, Green River Laboratory, St. Jacques, N. B., Canada, for advice on the application of the sequential analysis, and to Dr. M. MORISITA, Department of Biology, Faculty of Science, Kyushu University, Fukuoka, Japan, for replying his inquiries concerning the spacing method which he proposed, and for reading the manuscript. He also wishes to thank Mr. T. FUNAKI, who kindly allowed him to establish an observation plot in one of his paddy fields.

Summary

By examining the distribution of rice-stem borer egg masses in an observation plot established in a paddy field at Nogi, Matsue City, the writer made an attempt to answer the following two questions:

1) Whether the sequential sampling technique is applicable to classifying paddy fields according to population levels of egg masses existing in them.
2) Whether the spacing method proposed by MORISITA (1957) is adequate for estimating the population density of egg masses (or the total number of masses) in a given paddy field.

As a result of their application on the distribution maps, the above two techniques proved to be useful in principle at least in the case of rice-stem borer egg-mass populations.

要 約

松江市乃木の1筆の水田の中に設けた調査区からえられたニカメイガ卵塊の分布のデータを基にして，

1) ニカメイガ卵塊の密度の違いによって，いくつかの水田を類別するのに，sequential sampling technique が使えるかどうか？

2) あるひとつの水田の中のニカメイガ卵塊の密度（あるいは卵塊の総数）を推定するのに，森下（1957）によって提案された一種の間隔法が使えるかどうか？

の2つの疑問に答えようとした．

上の2つのサンプリング法を，卵塊の分布図にあてはめてみたところ，少くとも原則的には，これらの方法はニカメイガ卵塊個体群に使えることが判った．

Literature Cited

1 CONNOLA, D. P., W. E. WATERS, & W. E. SMITH : Bull. N. Y. St. Mus. No. 366, 22 pp., 1957.

2 GOULDEN, C. H. : *Methods of statistical analysis*. 2nd ed., New York, John Wiley, 467 pp., 1952.

3 MILLER, C. A. : Can. J. Zool. **33**, 5–17, 1955.

4 MILNE, A. : Biometrics **15**, 270–297, 1959.

5 MORISITA, M. : Physiol. and Ecol. **7**, 134–144, 1957. [In Japanese with a summary in English.]

6 MORRIS, R. F. : Can. J. Zool. **32**, 302–313, 1954.

7 MORRIS, R. F. : Ann. Rev. Ent. **5**, 243–264, 1960.