SANKHYĀ

THE INDIAN JOURNAL OF STATISTICS

Edited By : P. C. MAHALANOBIS

VOL. 14 PART 3 1954

ON THE ERROR IN CROP CUTTING EXPERIMENT DUE TO THE BIAS ON THE BORDER OF GRID

[APPLICATION OF INTEGRAL GEOMETRY TO AREAL SAMPLING PROBLEMS—PART IV]

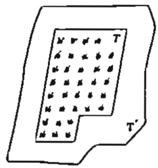
By MOTOSABURO MASUYAMA

Tokyo University, Tokyo and Indian Statistical Institute, Calcutta

1. INTRODUCTION

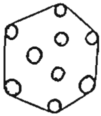
The aim of this note is to estimate the order of bias in the circular cutting method of estimating crop yield due to over- or under-estimation of numbers of plants on the boundary of the circle.

In what follows, we assume that the condition of unbiasedness with regard to the size T of the field holds. That is to say, every movable oval F of area F which has at least one point in common with the α-th small oval F_a of area F_a and of perimeter L_a in T is completely contained in T. Otherwise we must assume



an extended area T'' instead of T [Masuyama, 1953 (Part I)].

Consider the cross-section of plant at the ground level and replace a set of cross sections of tillers in a bunch by its convex closure. A mechanical analogy



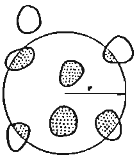
of the convex closure is the area surrounded by an elastic belt on wheels, where wheels correspond to the cross sections of tillers. For the sake of simplicity we call this convex closure the 'plant' or the 'small oval'.

2. THE FIRST MODEL

We assume that the yield of a bunch is proportional to its small oval. Then the kinematic expectation of the meet M between the set of fixed small ovals $\{F_a\}$ and a movable oval F is given by

$$KE\{M\} = \phi\left(\frac{F}{T}\right)^{\dagger} \qquad (= m, \text{ say}), \qquad \dots \quad (2.1)$$

under the condition of unbiasedness, where ϕ is the sum of areas $F_a (\alpha = 1, 2, \dots, \nu)$.



In our case we have $F = \pi r^2$, where r is the radius of circular cut. Thus we have

$$\Delta m/m = 2\Delta r/r \qquad \dots \quad (2.2)$$

which gives the relative error due to over- or under-estimation of r.

Let us assume that $r = 48$ inches and $\Delta r = 1$ inch, then we have

$$\Delta m/m = 0.042. \qquad \dots \quad (2.3)$$

If we use a movable square, as is used in Japan*, we have

$$\Delta m/m = 2\Delta a/a, \qquad \dots \quad (2.4)$$

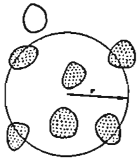
where a^2 is the size of the movable square. If we assume the areas of the same size, $\sqrt{\pi} r = a$. Thus $\Delta a = \Delta r$ leads to $\sqrt{\pi}\Delta a/a = \Delta r/r$.

† In (2.1) we have taken the kinematic density. However, it has been proved later [Masuyama, 1954] that the rotation of the movable oval is not necessary to estimate ϕ.

* It is called "Tubogari". "Tubo" is approximately equal to 36 square feet. "Gari" is the corresponding *mediae palatale* to the *tenues palatale* "Kari", which means "out".

3. The second model

We assume here that the yield per movable oval is proportional to the number of plants counted with regard to the movable oval.



We count here the number of plants which are partially or completely included in the movable oval. Then under the condition of unbiasedness the average number of plants in the circular cut is given by

$$p = (\phi+\lambda r+\nu\pi r^2)/T \quad \dots \quad (3.1)$$

[cf. Masuyama, 1953 (Part I)]

or,
$$pt = f+lr+\pi r^2, \quad \dots \quad (3.2)$$

where we put $T/\nu = t, \phi/\nu = f$ and $\lambda/\nu = l$.

We shall discuss three different models in this class.

i) *The bias due to the over- or under-estimation of the number of plants on the boundary of the movable oval.* As was shown in a previous paper [Masuyama, 1953 (Part II)], the number of plants on the boundary is equal to

$$y = 2\lambda r/T = 2lr/t \quad \dots \quad (3.3)$$

on the average. Thus the relative error of p will be

$$\begin{aligned}\Delta p/p &= \Delta(\lambda r)/(\phi+\lambda r+\nu\pi r^2)\\ &= t\Delta y/\{2(f+lr+\pi r^2)\}.\end{aligned} \quad \dots \quad (3.4)$$

Let us assume that $r = 4$ ft., $l = 0.78$ ft., $f = 0.054$ sq.ft. and $t = 1.0$ sq.ft.,* then we have

$$\begin{aligned}p &= 0.054+0.78r+3.14r^2\\ &= 0.05+3.12+50.27=53.44,\end{aligned} \quad \dots \quad (3.5)$$

and accordingly
$$\Delta p/p = \Delta y/106.88. \quad \dots \quad (3.6)$$

* These eye-estimates were obtained by the author at the Indian Statistical Institute's (ISI) experimental field of rice in Giridih. According to the direct counting by ISI trainees, $2lr/t$ is nearly 6 and p is nearly 50 in this field [cf. Masuyama & Sengupta, 1954 (Part V)].

If we exaggerate the actual problem, assuming that all plants on the boundary are erroneously excluded, then we have

$$\Delta y = 2lr/t = 6.24, \qquad \dots \quad (3.7)$$

and accordingly $$\Delta p/p = 0.058. \qquad \dots \quad (3.8)$$

In the case of a movable square of the same size we have

$$p = 0.054+0.50a+a^2$$
$$= 0.05+3.54+50.27 = 53.86, \qquad \dots \quad (3.9)$$

and $$\Delta y = 4la/(\pi t) = 7.08. \qquad \dots \quad (3.10)$$

Thus we obtain $$\Delta p/p = 0.066. \qquad \dots \quad (3.11)$$

That is to say, the square cut is a little bit worse than the circular cut. This is a natural conclusion because of the isoperimetric inequality.

We note that in our formulae (3.5) and (3.9), f is negligibly small compared with the area of movable oval and the second term which depends upon the mean perimeter is less than one tenth of the third term.

(ii) *The bias due to the over- or under-estimation of r.* In this case we have

$$t\Delta p = (l+2\pi r)\Delta r, \qquad \dots \quad (3.12)$$

and accordingly $$\Delta p/p = g(r)\,\Delta r, \qquad \dots \quad (3.13)$$

where we put $$g(r) = (l+2\pi r)/(f+lr+\pi r^2). \qquad \dots \quad (3.14)$$

In the case of our numerical example,

$$g(r) = (0.78+6.28r)/(0.054+0.78r+3.14r^2)$$
$$= 25.01/53.44 = 0.48, \qquad \dots \quad (3.15)$$

and accordingly $$\Delta p/p = 0.040, \qquad \dots \quad (3.16)$$

where we put $\Delta r = 1/12$ ft.

Differentiating $g(r)$ with regard to r, we get

$$p^2t^2g'(r) = (2\pi f-l^2/2)-2\pi^2(r+l/2\pi)^2. \qquad \dots \quad (3.17)$$

Thus we have a positive root of $g'(r)$, say r_0, only when $2\pi f > l^2$, but $g''(r_0) < 0$ in this case. That is to say, there is no positive r which minimizes $g(r)$. To make $g(r)$ small, we should make at least one of f, l and r large. We note that in our numerical example $g(r)$ is approximately equal to $2/r$ [cf. (2.2)].

If we assume further that each fixed oval is circular, we have

$$g(r) = (m+r)/\{\sigma^2+(m+r)^2\}, \qquad \dots \quad (3.18)$$

where m and σ^2 are the mean and the variance of the radii of fixed ovals. Thus $g(r)$ is a decreasing function of r, m and σ.

In the case of the movable square we have

$$g(a) = (0.50+2a)/p = 0.263, \qquad \dots \quad (3.19)$$

and accordingly $$\Delta p/p = 0.022, \qquad \dots \quad (3.20)$$

where we put $\Delta a = 1/12$ ft.

(iii) *The bias due to the under- or over-estimation of perimeters.* We have seen in (i) that the average number of plants on the boundary is equal to $2lr/t$. In this case we have

$$\Delta p/p = h(r)\Delta l \qquad \dots \quad (3.21)$$

where we assume the circular cut and

$$h(r) = r/(f+lr+\pi r^2). \qquad \dots \quad (3.22)$$

In the case of our numerical example we obtain, putting $\Delta l = 1/12$ ft.,

$$\Delta p/p = 0.0062. \qquad \dots \quad (3.23)$$

Differentiating $h(r)$ with regard to r, we have

$$p^2t^2h'(r) = f-\pi r^2. \qquad \dots \quad (3.24)$$

As $$h''(r_0) < 0, \quad \text{where } r_0 = \sqrt{f/\pi}, \qquad \dots \quad (3.25)$$

there is no minimum, except the trivial case where $r = 0$. In our numerical example $r_0 = 0.13$ ft.

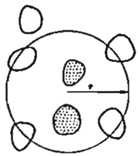
In the case of our movable square we have

$$\Delta p/p = 0.014. \qquad \dots \quad (3.26)$$

4. The third model

We put the same assumption as in § 3, except that we count only the number of plants which are completely included in the movable oval. Then the average number of plants in the circular cut is given by

$$p = (f - lr + \pi r^2)/t. \quad \ldots \quad (4.1)$$



Thus there is not much difference between this model and the previous model. In our numerical example, $p = 47.20$.

References

Mahalanobis, P. C. and Sengupta, J. M. (1951) : On the size of sample cuts in crop-cutting experiments in the Indian Statistical Institute : 1939–1950. *Bull. Inter. Stat. Inst.*, 33, 359–404.

Masuyama, M. (1953) : A rapid method of estimating basal area in timber survey (Part I). *Sankhyā*, 12, 291–302.

——— (1953) : Analysis of the 1939 model sample survey results from the viewpoint of integral geometry (Part II). *Sankhyā*, 13, 229–234.

——— (1953) : Rapid methods of estimating the sum of specified areas in a field of given size (Part III). *Rep. Stat. Appl. Res.*, 2, 113–119.

——— (1954) : A note on "Rapid methods of estimating the sum of specified areas in a field of given size". *Rep. Stat. Appl. Res.*, 3, No. 2, 32.

Masuyama, M. and Sengupta, J. M. (1954) : On a bias in crop-cutting experiments (Part V).* *Forthcoming.*

* Read at the meeting of the Indian Science Congress in January, 1954.