日水研報告，(40)：27-49，1990.
Bull. Japan Sea Natl. Fish. Res. Inst., (40)：27-49，1990.

An Interval Estimation of LESLIE's Method in Removal Methods

Tatsuro AKAMINE[1)]

Abstract

LESLIE's method in removal methods is modeled into the joint of the binomial distributions. An interval estimation of p (removal ratio) and n (initial population size) based on the null hypothesis for p and n is easily obtained by the approximation of each binomial distribution to the normal distribution. The confidence region on the (p, n) plane is easily obtained by micro-computers. The maximum likelihood method for a point estimation and the convenient method on the curves of partial maximum likelihood for an interval estimation are also presented.

Key words LESLIE's method, DELURY's method, removal method, binomial distribution, normal distribution, interval estimation

Introduction

LESLIE's method in removal methods is wellknown by scientists of population dynamics. SEBER (1982) said that this problem was first studied by LESLIE and DAVIS in 1939 and DELURY in 1947. This method is based on the linear regression model. This model is useful only for a point estimation, but not for an interval estimation. The probability model for this method is the joint of binomial distributions or the multinomial distribution. SEBER (1982) introduced many studies about this model, but almost all of them are not useful for interval estimations of p (removal ratio) and n (initial population size).

Although SCHNUTE (1983) presented a likelihood ratio test for an interval estimation of this model, his model was not useful and his estimating method was not adequate. The reason is that his model is a special model in which all sampling efforts are equal, and his method gives no attention to the degree of freedom of χ^2 distribution and the number of data.

The approximation of each binomial distribution to the normal distribution seems to be the best method for this problem. In this paper, we present this ap-

Accepted : October 28, 1989. Contribution A *No.* 459 from Japan Sea National Fisheries Research Institute.

1) Japan Sea National Fisheries Research Institute, Suido-cho, Niigata 951, Japan.
(〒951 新潟市水道町1丁目5939-22 日本海区水産研究所)

proximation method on the (p, n) plane and its convenient method on the curves of partial muximum likelihood for an interval estimation, the maximum likelihood method for a point estimation, and the consideration of the Bayesian statistical method based on the uniform prior distribution.

LESLIE's method

LESLIE's method is a regression model. Let

n : initial population size,

p_i : removal ratio of i th sample,

r_i : size of i th sample removed from the population.

LESLIE's method is as follows :

Generally, the following equation is supported for each removal sampling.

$$E(n_i p_i) = r_i, \quad n_i = n - R_{i-1}. \qquad (1.1)$$

Where $R_i = \sum_{k=1}^{i} r_k$.

$E(\theta)$: expected value of θ.

LESLIE's method is characterized by the next equation.

$$p_i = x_i p. \qquad (1.2)$$

x_i : units of effort expended on the i th sample.

Substituting (1. 2) into (1. 1) leads to the next equation.

$$E\left(\frac{r_i}{x_i}\right) = np - pR_{i-1}. \qquad (1.3)$$

Let $y = r_i/x_i$, $x = R_{i-1}$, then we get the regression model so called "LESLIE's method" or "DELURY's method".

This method has been used widely because point estimators of p and n can be easily obtained. However, it is difficult to obtain interval estimators of p and n. For example, when the number of data is only 2, this regression method has no interval of p and n.

When one of p and n is fixed, we can estimate the other exactly by the Bayesian statistical method (AKAMINE 1989a, 1989b), and the confidence interval of the parameter is not so wide. But in obtaining p and n simultaneously, confidence intervals of both parameter are wide. In that case the interval estimation is much more important than the point estimation. Therefore, we must use the model based on the binomial distribution mentioned in the next chapters.

The joint model of the binomial distributions

1. Model

It is widely known that LESLIE's method, the joint of the binomial distribution model, is as follows : The probability of the i th sample r_i is according to the binomial distribution :

$$P_i = \binom{n-R_{i-1}}{r_i} p_i^{r_i}(1-p_i)^{n-R_i}. \tag{2.1}$$

$$\text{Where } \binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{n^{(r)}}{r!}, \tag{2.2}$$

$$n^{(r)} = n(n-1)\cdots\cdots(n-r+1),$$
$$n! = n^{(n)}.$$

Therefore, the total probability called "likelihood" is defined as follows :

$$L = \prod_{i=1}^{m} P_i$$
$$= \frac{n^{(Rm)}}{r_1!\cdots\cdots r_m!} p_1^{r1}\cdots\cdots p_m^{rm}$$
$$\times(1-p_1)^{n-R1}\cdots\cdots(1-p_m)^{n-Rm}. \tag{2.3}$$

Where m : number of sampling (number of data).

This is equal to the multinomial distribution :

$$L = \frac{n^{(Rm)}}{r_1!\cdots\cdots r_m!} s_1^{r1}\cdots\cdots s_m^{rm} t^{n-Rm}. \tag{2.4}$$

$$\text{Where } t = 1-\sum_{i=1}^{m} s_i,$$

$$s_i = p_i \prod_{k=1}^{i-1}(1-p_k),\ t = \prod_{k=1}^{m}(1-p_k). \tag{2.5}$$

The next equation is necessary for LESLIE's method.

$$p_i = x_i p \tag{1.2}$$

For (2. 3) and (2. 4) without (1. 2), the next expression holds.

$$S_0 = \sum_{r_1=0}^{n}\left(\sum_{r_2=0}^{n-R_1}\left(\cdots\cdots\left(\sum_{rm=0}^{n-R_{m-1}} L\right)\cdots\cdots\right)\right)$$
$$= \left(\sum_{r_1=0}^{n} P_1\right)\left(\sum_{r_2=0}^{n-R_1} P_2\right)\cdots\cdots\left(\sum_{rm=0}^{n-R_{m-1}} P_m\right) \tag{2.6}$$
$$= 1\cdot 1\cdot\cdots\cdots\cdot 1$$
$$= 1$$

This equation can also be rewritten as follows :

$$S_0 = \sum_{D} L = 1,\ D : \sum_{i=1}^{m} r_i \leqq n \tag{2.7}$$

2. Maximum likelihood estimation

The estimators of $(p,\ n)$ are easily obtained by the maximum likelihood method. These are obtained by the following simultaneous equations.

$$\frac{\partial L}{\partial n} = 0, \tag{2.8}$$

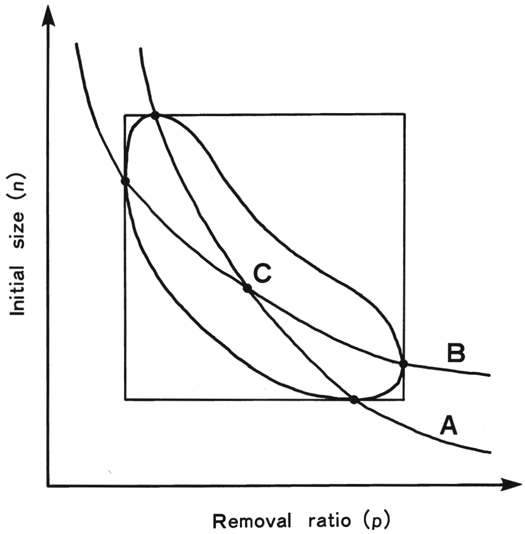

Fig. 1. The image of the confidence region on the (p, n) plane. The bent ellipse is the confidence region which is exaggerated.
A : The curve of $\partial L/\partial p=0$. B : The curve of $\partial L/\partial n=0$. C : The crossing point of A and B which means the point estimator of the maximun likelihood method.

$$\frac{\partial L}{\partial p}=0. \tag{2.9}$$

These equations are curves of partial maximum likelihood in Fig. 1. The intersection point of these curves is the maximun likelihood which is the estimator of (p, n). Fig. 1 also shows the cofidence region which is exaggerated for explanation purposes. The real confidence region is much narrower (see Fig. 2). Furthermore, these curves are so close that the precision of estimators (coordinate of intersection point) is law. Namely, the condition of this model is so bad, then we need an interval estimation and not a point estimation.

L is defined by (2. 3) and (1. 2) for LESLIE's method. From (2. 8),

$$\sum_{i=1}^{R_m}\frac{1}{n-i+1}+\sum_{i=1}^{m}\log(1-x_i p)=0. \tag{2.10}$$

From (2. 9),

$$n=\left(\frac{R_m}{p}+\sum_{i=1}^{m}\frac{R_i x_i}{1-x_i p}\right)\Big/\sum_{i=1}^{m}\frac{x_i}{1-x_i p}. \tag{2.11}$$

Substitute (2. 11) into (2. 10), then we get

$$f(p)=\sum_{i=1}^{R_m}\frac{1}{n-i+1}+\sum_{i=1}^{m}\log(1-x_i p). \tag{2.12}$$

This equation can be solved easily.

In this paper, NEWTON's method is applied. Let $n=A/B$, where A is the denominator and B is the numerator of (2. 11). Therefore, we get

$$f'=\frac{\partial f}{\partial p}=-n'\sum_{i=1}^{R_m}\frac{1}{(n-i+1)^2}-B. \tag{2.13}$$

Where $\frac{n'}{n}=\frac{A'}{A}-\frac{B'}{B}$,

$$A'=-\frac{R_m}{p^2}+\sum_{i=1}^{m}\frac{R_i x_i^2}{(1-x_i p)^2},$$

$$B'=\sum_{i=1}^{m}\frac{x_i^2}{(1-x_i p)^2}.$$

NEWTON's method corrects p step by step using the following expression :

$$\Delta p=-\frac{f}{f'}. \tag{2.14}$$

The initial value of p is searched for using the BASIC program in the appendix. An example of the BASIC program of NEWTON's method is also shown in the appendix.

In practice, calculations of Σ in (2. 12) and (2. 13) take a long time. Then we use "EULER-MACLAURIN expansion" as follows :

$$\begin{aligned} &f(a)+f(a+h)+\cdots\cdots+f(a+nh)\\ &=\frac{1}{2}\{f(a)+f(b)\}+\frac{1}{h}\int_a^b f(x)\mathrm{d}x\\ &+\frac{h}{12}\{f'(b)-f'(\mathrm{a})\}-\frac{h}{720}\{f^{(3)}(b)-f^{(3)}(a)\}\cdots\cdots.\\ &\text{Where } b=a+nh. \end{aligned} \tag{2.15}$$

Therefore,

$$\begin{aligned}\sum_a^b\frac{1}{x}&=\frac{1}{2}\left(\frac{1}{a}+\frac{1}{b}\right)+(\log b-\log a)\\ &-\frac{1}{12}\left(\frac{1}{b^2}-\frac{1}{a^2}\right)+\frac{1}{120}\left(\frac{1}{b^4}-\frac{1}{a^4}\right)\cdots\cdots\end{aligned} \tag{2.16}$$

$$\begin{aligned}\sum_a^b\frac{1}{x^2}&=\frac{1}{2}\left(\frac{1}{a^2}+\frac{1}{b^2}\right)-\left(\frac{1}{b}-\frac{1}{a}\right)\\ &-\frac{1}{6}\left(\frac{1}{b^3}-\frac{1}{a^3}\right)+\frac{1}{30}\left(\frac{1}{b^5}-\frac{1}{a^5}\right)\cdots\cdots\end{aligned} \tag{2.17}$$

[Example 1] SCHNUTE's data

$\boldsymbol{x}=(1, 1, 1)$, $\boldsymbol{r}=(90, 60, 40)$. Original (p, n) is (1/3, 270). Let initial value of p be 0. 5, and use the BASIC program in the appendix. Then we get estimators $(p, n)=$ (0. 341807, 265. 255).

Although this solution has bias, that is a general property of the maximum likelihood estimation. If the original value of n is known, the estimator of p has no bias. Let $x_i \equiv 1$ in (2. 11), then we get

$$\frac{1}{p}=1+\frac{mn-\sum R_i}{R_m}. \tag{2. 18}$$

Substituting data of example 1 into this equation we get the original value of p as follows :

$$\frac{1}{p}=1+\frac{3\times 270-(3\times 90+2\times 60+40)}{190}=3.$$

This case is the same as the next famous example : The estimators of normal ditribution $N(\mu, \sigma^2)$ by the maximum likelihood method are

$$\widehat{\mu}=\bar{x}=\frac{\sum x}{m}\cdots\cdots(a),\quad \widehat{\sigma^2}=\frac{\sum(x-\mu)^2}{m}\cdots\cdots(b).$$

Although these estimators have no bias, (b) is impossible to use because the original value of μ is unknown. Therefore, we often substitute (a) into μ of (b). Then we get (c) as an estimator by the maximum likelihood method.

$$\widehat{\sigma^2}=\frac{\sum(x-\bar{x})^2}{m}\cdots\cdots(c),\quad \widehat{\sigma^2}=\frac{\sum(x-\bar{x})^2}{m-1}\cdots\cdots(d).$$

But (c) has bias when m (number of data) $\neq \infty$. The estimator which has no bias is (d) called "unbiased variance".

This is a weak point of the maximum likelihood method. However, an interval estimation is more important than a point estimation for LESLIE's method which has a large confidence region. Therefore, unbiased estimators are not treated in this paper.

3. Existence of the solution

There is no solution of the estimator (p, n) for the maximum likelihood method in the case of the binomial distribution model :

$$P_1=\frac{n^{(r)}}{r!}p^r(1-p)^{n-r}. \tag{2. 19}$$

Where data is r. This is obvious because the number of estimators is 2 and the number of data is 1. Let's explain this by the following equations :

$\partial P_1/\partial n=0,\ \partial P_1/\partial p=0$ leads

$$\sum_{i=1}^{r}\frac{1}{n-i+1}+\log(1-p)=0, \tag{2. 20}$$

$$p=\frac{r}{n}. \tag{2. 21}$$

Substituting (2. 21) into (2. 20) we get

$$f(n)=\sum_{i=1}^{r}\frac{1}{n-i+1}+\log\left(1-\frac{r}{n}\right). \tag{2. 22}$$

But this equation has no solution because $f(n)<0$ as follows :

$$\sum_{i=1}^{r}\frac{1}{n-i+1}=\sum_{i=n-r+1}^{n}\frac{1}{i}<\int_{n-r}^{n}\frac{1}{x}\,dx=\log\frac{n}{n-r}$$
$$=-\log\left(1-\frac{r}{n}\right). \tag{2.23}$$

This is the same as the multinomial distribution model (2. 3). Where data are $r_1\sim r_m$. $\partial L/\partial n=0$, $\partial L/\partial p_i=0$ $(i=1\sim m)$ leads

$$\frac{\partial L}{\partial n}=\sum_{i=1}^{R_m}\frac{1}{n-i+1}+\sum_{i=1}^{m}\log(1-p_i), \tag{2.24}$$

$$p_i=\frac{r_i}{n-R_{i-1}}\quad (i=1\sim m). \tag{2.25}$$

Substituting (2. 25) into (2. 24) we get

$$f(n)=\sum_{i=1}^{R_m}\frac{1}{n-i+1}+\log\left(1-\frac{R_m}{n}\right). \tag{2.26}$$

This equation has no solution because $f(n)<0$ too. This is obvious because the number of estimators (p, n) is $(m+1)$ and the number of data is m. LESLIE's method decreases the number of estimators to 2 by (1. 2). Then it is possible to obtain the solution for LESLIE's method.

An approach to the Bayesian statistical method

In this chapter, we will consider the Bayesian statistical method for an interval estimation. It is natural to let the prior distribution of (p, n) be the uniform distribution $(0\leqq p\leqq 1,\ R_m\leqq n\leqq\infty)$, the same as AKAMINE (1989a, 1989b). Calculating the posterior distribution we get the following results :

1. General removal method

It is necessary for the posterior distribution to sum the likelihood on the (p, n) plane. The next expressions are essential for this calculation. For (2. 1) and (2. 3),

$$I(n,\ 1)=\int_0^1 P_1 dp_1=\frac{1}{n+1}=n^{(-1)}, \tag{3.1}$$

$$I(n,\ m)=\int_0^1\cdots\cdots\int_0^1 LJ dp_1\cdots\cdots dp_m$$
$$=\frac{1}{(n+m)\cdots\cdots(n+1)}=n^{(-m)}. \tag{3.2}$$

Where J : Jacobian.

From these we get the next sum.

$$S_1=\sum_{r=R_m}^{\infty}I(n,\ 1)=\infty, \tag{3.3}$$

$$S_m=\sum_{r=R_m}^{\infty}I(n,\ m)=\frac{R_m^{(-m+1)}}{m-1}<\infty. \tag{3.4}$$

These results are not affected by the rearrangement of the order of p and n to integrate as follows :

$$\Sigma\left(\int P\mathrm{d}p\right)=\int(\Sigma P)\mathrm{d}p. \tag{3.5}$$

(3. 3) shows that it is impossible to calculate the posterior distribution. However, (3. 4) does not show that it is possible. Because point estimators cannot exist as shown in chapter 2. (3. 1) and (3. 2) only show that the Bayesian statistical model is possible when n is known as AKAMINE (1989b).

(3. 2) is rewritten for the multinomial distribution (2. 4).

$$I(n,\ m)=\int\cdots\cdots\int_D L\,\mathrm{d}s_1\cdots\cdots\mathrm{d}s_m=n^{(-m)}. \tag{3.6}$$

Where $D: 0\leqq s_i,\ \sum s_i\leqq 1.$

This equation is equal to the definition of the multi beta function (DIRICHLET's integration). In this paper, the proof of (3. 6) by the change of variables (2. 5) is presented. First, let's calculate the Jacobian :

$$J=\frac{\partial(s_1,\cdots\cdots,s_m)}{\partial(p_1,\cdots\cdots,p_m)}=\begin{vmatrix}\dfrac{\partial s_1}{\partial p_1} & \cdots\cdots & \dfrac{\partial s_1}{\partial p_m}\\ \vdots & & \vdots\\ \dfrac{\partial s_m}{\partial p_1} & \cdots\cdots & \dfrac{\partial s_m}{\partial p_m}\end{vmatrix}$$

$$=\begin{vmatrix}1 & & & \\ -p_2 & (1-p_1) & & 0\\ \vdots & \vdots & \ddots & \\ \vdots & \vdots & \cdots\cdots & (1-p_1)\cdots\cdots(1-p_{m-1})\end{vmatrix}$$

$$=(1-p_1)^{m-1}(1-p_2)^{m-2}\cdots\cdots(1-p_{m-1})>0. \tag{3.7}$$

On the other hand, the region of integration changes as follows : From D of (3. 6),

$$\sum s_i=1-t\leqq 1.$$

This leads to

$$t=(1-p_1)\cdots\cdots(1-p_m)\geqq 0.$$

And $s_i\geqq 0$. Then D becomes

$$0\leqq p_i\leqq 1\quad(i=1\sim m). \tag{3.8}$$

Therefore, the integration of (3. 6) is obtained as follows :

$$I(n,\ m)=\int_0^1\cdots\cdots\int_0^1 LJ\mathrm{d}p_1\cdots\cdots\mathrm{d}p_m$$

$$=\frac{n^{(Rm)}}{r_1!\cdots\cdots r_m!}\int_0^1 p_1^{r_1}(1-p_1)^{n-R_1+m-1}\mathrm{d}p_1$$

$$\cdots\cdots\int_0^1 p_m^{r_m}(1-p_m)^{n-R_m}\mathrm{d}p_m$$

$$=\frac{n^{(Rm)}}{r_1!\cdots\cdots r_m!}B(r_1+1,\ n-R_1+m)$$

$$\cdots\cdots B(r_m+1,\ n-R_m+1)$$

$$=\frac{n!}{(n+m)!}=n^{(-m)}$$

Where the next formular of the beta function is used.

$$B(a,\ b)=\int_0^1 x^{a-1}(1-x)^{b-1}\mathrm{d}x=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

$$=\frac{(a-1)!(b-1)!}{(a+b-1)!} \tag{3.9}$$

In the m th dimensional space of $(p_1,\ \cdots\cdots,\ p_m)$, equation (1. 2) is a line as follows :

$$\boldsymbol{p}=\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} p \tag{3.10}$$

2. LESLIE's method

LESLIE's method is characterized by the equation (1. 2). It is impossible for this model to be a Bayesian statistical model, the same as the binomial distribution model (3. 3). First, the special case of $x_i=1$ is shown as follows : In this case, (2. 3) becomes

$$L=\frac{n^{(Rm)}}{r_1!\cdots\cdots r_m!}p^{Rm}(1-p)^{mn-\Sigma R_i}. \tag{3.11}$$

The integration of the terms for p by using (3. 9) leads to

$$\int_0^1 p^{Rm}(1-p)^{mn-\Sigma R_i}\mathrm{d}p$$

$$=\frac{R_m!\,[mn-\sum R_i]!}{[mn-\sum R_i+R_m+1]!}. \tag{3.12}$$

Then the integration of (3. 11) is

$$\int_0^1 L\mathrm{d}p=a\frac{n^{(Rm)}}{(mn-b)^{(Rm+1)}}\doteqdot\frac{a}{m^{Rm+1}}\frac{1}{n}. \tag{3.13}$$

Where a, b : const.

This equation leads to

$$S_1{*}=\sum_{r=Rm}^{\infty}\int_0^1 L\mathrm{d}p=\infty. \tag{3.14}$$

(3. 14) holds in the general case as $x_i\neq 1$. The simplest case such as the next one is enough to explain this. Let

$$I=\int_0^1 p^a(1-p)^b(1-xp)^c\mathrm{d}p. \tag{3.15}$$

Where $0<x\leqq 1$.

On the other hand,

$$1-xp=(1-x)+x(1-p). \tag{3.16}$$

Substituting (3. 16) into (3. 15) we get

$$I=(1-x)^c\int_0^1 p^a(1-p)^b\mathrm{d}p+x^c\int_0^1 p^a(1-p)^{b+c}\mathrm{d}p. \tag{3.17}$$

This equation proves that (3. 14) holds in general. Therefore, it is difficult to make Bayesian statistical models based on the hypothesis that the prior distributions of p and n are both uniform distributions.

Generally, the next hypothesis are used to let $S_1^*<\infty$.

a) $n\leqq N<\infty$.

b) The prior distribution of n is not the uniform distribution.

But these hypothesis seem to be difficult to match with AKAMINE (1989a, 1989b)'s model. In this paper, the Bayesian statistical method is not used for an interval estimation.

Interval estimation

For an interval estimation, it is important to set a null hypothesis. In this model, the next hypothesis is natural.

$$H_0: n=n_0 \text{ and } p=p_0. \tag{4.1}$$

TANAKA (1985) had already shown the confidence region of (p, n) for the similar model with LESLIE's method. Although the confidence region of the other parameter $(s=\mathrm{e}^{-M}, n)$ was a bent ellipse, that of (p, n) was a complete ellipse and not a bent ellipse in his result.

1. The approximation to the normal distribution

LESLIE's method is expressed as the joint of the binomial distributions. It is widely known by the "DE MOIVRE–LAPLACE theorem" that when $n\to\infty$, the binomial distribution will be approximately equal to the normal distribution. In practice, this approximation is useful when $n>30$, $np>5$, and $n(1-p)>5$. Therefore, for our purposes this approximation is useful because n is usually large in our sampling.

In using this theorem for (2. 3), we get

$$L=\prod_{i=1}^{m} N(n_ip_i,\ n_ip_iq_i). \tag{4.2}$$

Where $n_i=n-R_{i-1}$, $q_i=1-p_i$.

The next expression is important for LESLIE's method.

$$p_i=x_ip. \tag{1.2}$$

For each normal distribution in (4. 2), the confidence interval is easily obtained by the next inequality.

$$\frac{r_i - n_i p_i}{\sqrt{n_i p_i (1-p_i)}} \leqq z. \tag{4.3}$$

Where z is 1.96 (95%) or 2.58 (99%).

Let $p_i \to x$, $n_i \to y$ and $r_i \to r$. Then (4.3) is rewritten as follows :

$$(r - yx)^2 \leqq z^2 yx(1-x). \tag{4.4}$$

Let

$$\begin{aligned} f(x, y) &= (r - yx)^2 - z^2 yx(1-x) \\ &= x^2y^2 + x(z^2x - z^2 - 2r)y + r^2 = 0. \end{aligned} \tag{4.5}$$

The solution of this equation for y is as follows :

$$y = \frac{-b \pm \sqrt{D}}{2x}. \tag{4.6}$$

Where $b = z^2x - z^2 - 2r$,

$D = z^4(x-1)\{x - (1 + 4r/z^2)\} \geqq 0.$

These are two curves. Draw them for each normal distribution on the (p, n) plane, and we get the confidence region. The product space of all confidence regions gives the interval estimation of p and n which means that all normal distributions of (4.2) satisfy null hupothesis (4.1). The BASIC program in the appendix gives values of (4.6).

Let's consider the convenient method in this way. Either (2.8) or (2.9) is a curve which gives maximum value of likelihood L viewed from one direction. In this model, the concrete expression of (2.8) is (2.10) and that of (2.9) is (2.11). Although (2.11) is an explicit function for n, (2.10) is not. Therefore, using (2.11) it is easier to calculate n than in using (2.10). On the other hand, (2.10) gives maximun and minimum values of p in the confidence region while (2.11) gives those of n (see Fig. 1). The main purpose of LESLIE's method is to estimate n rather than p. Therefore, it is better to use (2.11) for estimation in this paper. For each normal distribution, the next function of p gives the confidence region of n on the (2.11) curve.

$$z_i = \frac{r_i - n_i p_i}{\sqrt{n_i p_i (1-p_i)}}. \tag{4.7}$$

Where n_i is given by (2.11), $p_i = x_i p$.

It is anticipated that the result of (4.7) will be almost equal to (4.6). An example of the BASIC program for this method is shown in the appendix.

2. Likelihood ratio test

(3.14) shows that it is impossible to estimate the confidence region by the superficial content under the likelihood curve which means probability. Then it is natural to use the height of likelihood curve which means probability density instead of the superficial content. The most popular approach is the likelihood ratio test. This test is based on the next theorem.

[Theorem] The likelihood ratio is defined as the next expression.

$$\lambda = \max_{\theta \in \omega} L \Big/ \max_{\theta \in \Omega} L. \tag{4.8}$$

Where $\omega \subset \Omega$, Ω : parameter space.
When m (number of data)$\to\infty$, $-2\log\lambda \sim \chi^2(k-s)$. Where k is the number of parameters of Ω, and s is the number of parameters of ω.

The condition of m (number of data) is the biggest problem when using LESLIE's method. Although this test requires m to be large, m is usually small in our sampling. Therefore, we must consider the real distribution of λ in this case. However, this problem is not treated in this paper.

The null hypothesis is (4. 1) in this paper. On the other hand, SCHNUTE (1983) used the next hypothesis.

$$H_0 : n = n_0. \tag{4.9}$$

Then he used the value of χ^2 (1). Namely, he estimated only on the (2. 11) curve. In this paper, we estimate on the whole (p, n) plane, and use the (2. 11) curve only for convenience. Therefore, we must use the value of χ^2 (2).

For calculation of L in (2. 3), the next expression called "STIRLING's formula" is useful.

$$\log(n!) = \frac{1}{2}\log(2\pi) + \left(n + \frac{1}{2}\right)\log n - n + \frac{1}{12n}. \tag{4.10}$$

This is the first part of the asymptotic expansion called "EULER–MACLAURIN expansion" which is as follows :

$$\log(n!) = \frac{1}{2}\log(2\pi) + \left(n + \frac{1}{2}\right)\log n - n + \frac{1}{12n} - \frac{1}{360n^3} + \frac{1}{1260n^5} - \frac{1}{1680n^7} + \cdots\cdots. \tag{4.11}$$

The precision of (4. 10) is enough for our calculation. An example of the BASIC program is shown in the appendix.

3. Numerical experiments

The artificial data for the experiment is shown in Table 1. Data–1 has no error, data–2 has small error, and data–3 has large error. For point estimation, program B in the appendix is used. For interval estimation, program C and D are used. For comparison, program E for the likelihood ratio test is used.

Table 1-a. The artificial data for the experiment (data-1).

i	n	p	r[1)]	R[2)]
1	10000	0.07	700	700
2	9300	0.05	465	1165
3	8835	0.10	884	2049
4	7951	0.03	636	2685
5	7315	0.04	293	2978

Table 1-b. Continued (data-2).

i	n	p	np	$\sqrt{npq}$	e	r[3)]	R[2)]
1	10000	0.07	700	25.5	0.7	718	718
2	9282	0.05	464	21.0	0.6	477	1195
3	8805	0.10	881	28.2	-0.9	856	2051
4	7949	0.08	636	24.2	0.0	636	2687
5	7313	0.04	293	16.8	-0.1	291	2978

Table 1-c. Continued (data-3).

i	n	p	np	$\sqrt{npq}$	e	r[3)]	R[2)]
1	10000	0.07	700	25.5	1.4	736	736
2	9264	0.05	463	21.0	1.2	488	1224
3	8776	0.10	878	28.1	-1.8	827	2051
4	7949	0.08	736	24.2	0.0	636	2687
5	7313	0.04	293	16.8	-0.2	290	2977

1) $r=np$. 2) $R=\sum r$. 3) $r=np+e\sqrt{npq}$.

[Example 2] Data-1

The result of the maximum likelihood method is $(p, n)=(.01003895, 9968.41)$. On the other hand, the result of the regression method is $(p, n)=(.00997327, 10026.07)$ and $r=-.999994$. The confidence regions are shown in Fig. 2a and 2b. Fig. 2a shows the interval estimation (95%) is $(p, n)=(.0042, 22026)\sim(.0156, 6974)$, and Fig. 2b shows $p=.0044\sim.0152$. Because data-1 has no error, the point estimator is natural for all samplings. On the other hand, the result of the likelihood ratio test is $(p, n)=(.0057, 16611)\sim(.0143, 7411)$.

[Example 3] Data-2

The result of the maximum likelihood method is $(p, n)=(.01107396, 9161.77)$. On the other hand, the result of the regression method is $(p, n)=(.01114463, 9151.39)$ and $r=-.989140$. The confidence regions are shown in Fig. 2c and 2d. Fig. 2c shows the interval estimation (95%) is $(p, n)=(.0053, 18034)\sim(.0164, 6701)$, and Fig. 2d shows $p=.0063\sim.0156$. The point estimator is not natural for the 3rd sampling. On the

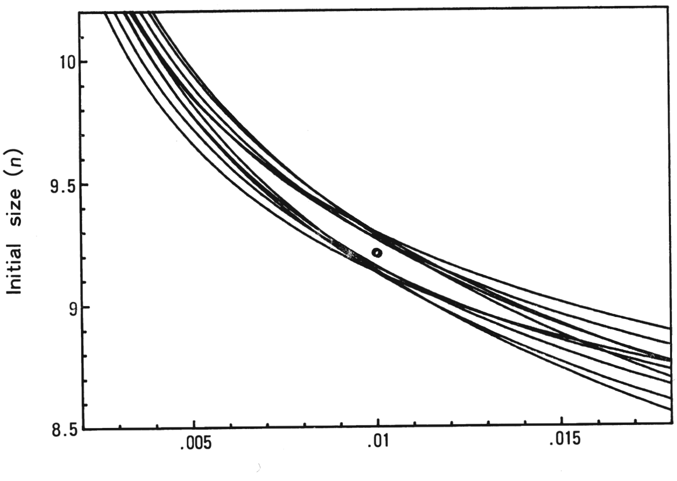

Fig. 2-a. The confidence region for data-1 on the (p, n) plane by the method based on the approximation to the normal distribution. Circles are original value (.01, 10000) and the point estimator by the maximum likelihood method.

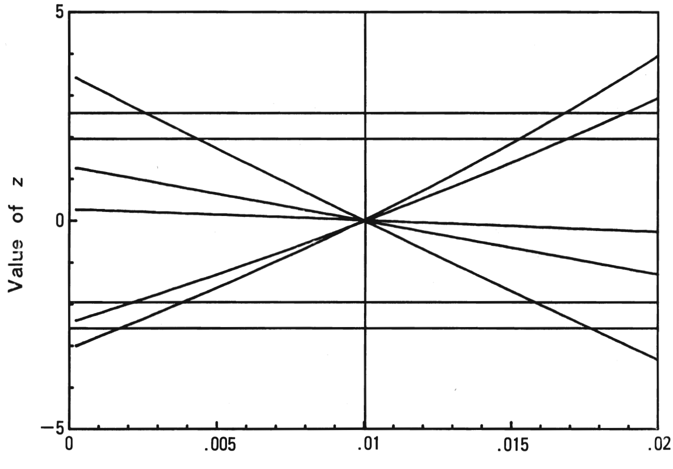

Fig. 2-b. The confidence interval of p for data-1 on the $\partial L/\partial p=0$ curve by the convenient method based on the approximation to the normal distribution.

other hand, the result of the likelihood ratio test is $(p, n)=(.0068, 14134)\sim(.0152, 7014)$.

[Example 4] Data-3

The result of the maximum likelihood method is $(p, n)=(.01202092, 8543.93)$. On the other hand, the result of the regression method is $(p, n)=(.01225645, 8477.77)$ and $r=-.964732$. The confidence regions are shown in Fig. 2e and 2f. Fig. 2e

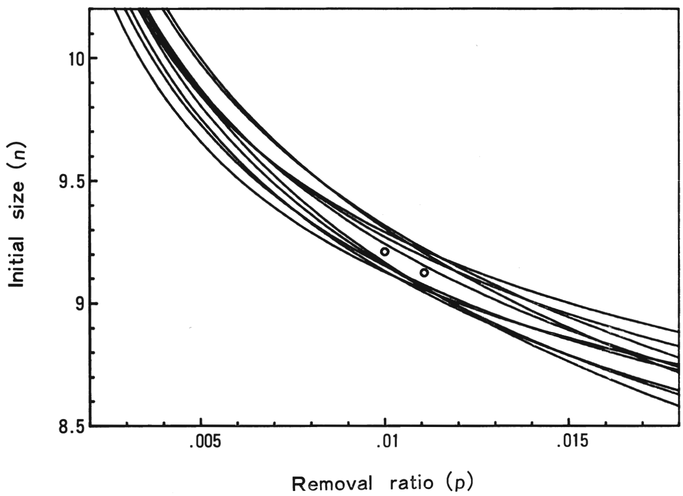

Fig. 2-c. Continued (data-2).

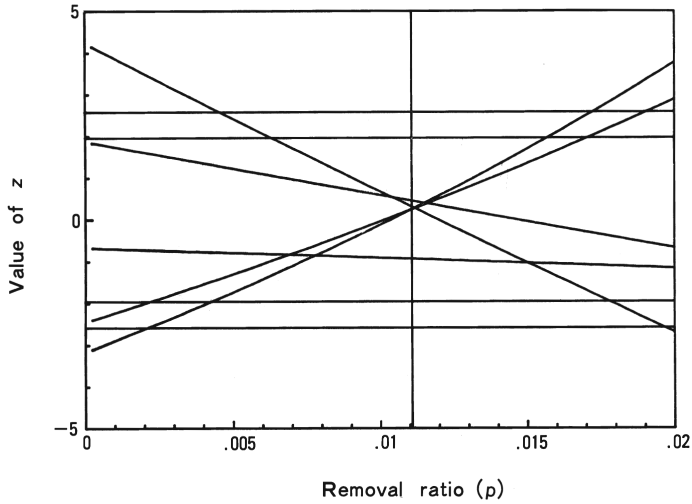

Fig. 2-d. Continued (data-2).

shows the interval estimation (95%) is $(p, n)=(.0080, 12210)\sim(.0158, 6836)$, and Fig. 2f shows $p=.0083\sim.0148$. Although the confidence interval of 95% involves the point estimator, that of 90% does not exist because the value of z_3 is too large. In practice, this case seems to occur many times because by sampling error. On the other hand, the result of the likelihood ratio test is $(p, n)=(.0077, 12632)\sim(.0161, 6697)$.

4. Consideration

For a point estimator, both the maximum likelihood method and the regression method presents a good value in practice. However, the former method has a bias. On the other hand, the latter method is difficult to estimate a confidence interval,

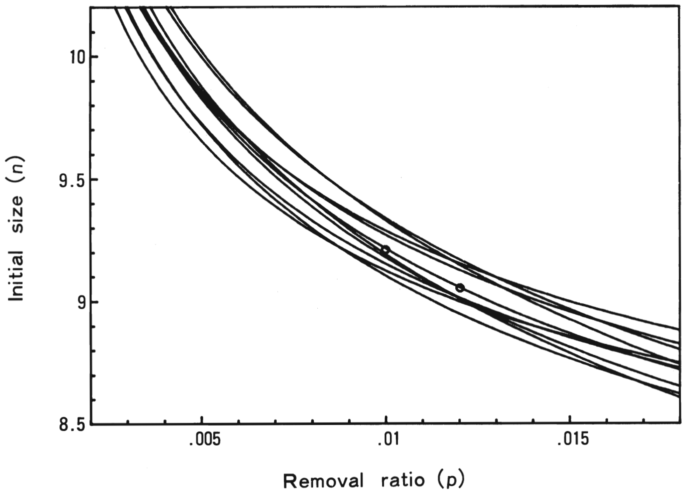

Fig. 2-e. Continued (data-3).

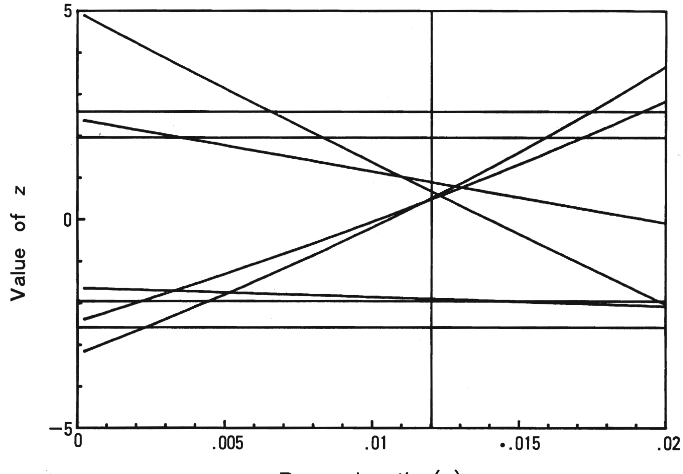

Fig. 2-f. Continued (data-3).

and correlation coefficient r does not mean any condition of the model and data.

The confidence region of (4. 3) is different from the confidence region of the likelihood ratio test given by J. NEYMAN. The former means the region in which we get $\boldsymbol{r}$ (data) in some probability. On the other hand, the latter means the region in which the true (p, n) exists in some probability. The former region of good conditional data is larger than that of bad conditional data. This is particular character of this region. If the condition of data is too bad, this region does not exist. Therefore, this region suggests the condition of the model and data. When this region does not exist in high probability, the condition of the model and data is too bad to estimate p and n.

The latter region based on the likelihood ratio test is the general confidence region. However, this region has a bias which is the same as the maximum likelihood method.

5. Conclusion

For estimations of this model, we get the results as follows :

(a) For a point estimation, the maximum likelihood method is better than the regression method because the former is able to be expanded to an interval estimation naturally. However, the former has a bias.

(b) The region based on the probability of $\boldsymbol{r}$ (data) suggests the condition of the model and data. When this region does not exist in high probability, we cannot use this model and data.

(c) The likelihood ratio test (program E) gives the general confidence region. However, this region has a bias which is the same as the maximum likelihood method. In this method, we must use $\chi^2(2)$ and not $\chi^2(1)$. The condition of m (number of data) in this method is a careful point to apply.

Acknowledgements

The author is indebted to Mr. K. HIRAMATSU of the National Research Institute of Far Seas Fisheries and to Dr. Y. MATSUMIYA of the Ocean Research Institute of the University of Tokyo for their kind advice. The author is also grateful to Mr. K. NOGAMI and Mr. S. UMEZAWA of the Japan Sea National Fisheries Research Institute for their critical reading of the manuscript.

References

AKAMINE, T. (1989a) An interval estimation for extraction using Bayesian statistics. *Bull. Japan Sea Reg. Fish. Res. Lab.*, (39), 9-17.

AKAMINE, T. (1989b) An interval estimation for the PETERSEN method using Bayesian statistics. *Bull. Japan Sea Reg. Fish. Res. Lab.*, (39), 19-33.

SCHNUTE, J. (1983) A new approach to estimating populations by the removal method. *Can. J. Fish. Aquat. Sci.*, (40), 2153-2169.

SEBER, G. A. F. (1982) The estimation of animal abundance and related parameters. 2nd edit. Griffin, London. 296-327, 500-504.

TANAKA, S. (1985) Suisan sigengaku souron (Total theory of fishery population). Kouseisha-kouseikaku, Tokyo. 231-235. (In Japanese)

除去法における LESLIE 法の区間推定

赤 嶺 達 郎

除去法における LESLIE 法は二項分布の積としてモデル化できる． p（除去率）と n（初期資源尾数）の区間推定は p と n の帰無仮説において，各二項分布を正規分布に近似することによって容易にできる．（p, n）平面上の信頼域は小型計算機で簡単に求まる． 最尤法による点推定および尤度最大の曲線による区間推定の簡便法も提供する．

Appendix

These are examples of the BASIC programs for the calculation in this paper. These have no error for the micro computer PC-9801VX (NEC). Although all variables and number jointed by "#" are defined to be double precision, the precision of results is not so high. "SQR" and "LOG" functions of N_{88}-BASIC for PC-9801VX is corrected recently, and they have no error.

Program A. This program searches the initial value for program B.

```
5  '-----------------------------------
10 '  LESLIE's method
20 '     Maximum likelihood method
30 '     Search for the initial value
40 '               by TATSURO AKAMINE
50 '                  1989/8/03
60 '-----------------------------------
1010 GOSUB *INIT1
2000 RXMAX=1#/XMAX : DELTAP0=RXMAX/20#
2020 FOR K=1 TO 20
2030    P0=K*DELTAP0
2035    B1=0 : B2=0 : E1=0
2040    FOR I=1 TO M1
2050       X2=X1(I)/(1-X1(I)*P0)
2090       B1=B1+X2
2110       B2=B2+X2*IR2(I)
2130       E1=E1+LOG(1-X1(I)*P0)
2140    NEXT I
2150    A1=IRSUM/P0+B2
2160    EN=A1/B1
2161    PRINT "P=";P0 , "N=";EN ,
2165    IF EN<IRSUM THEN END
2170    D1=0
2180    GOSUB *OYMC1
2240    F1=D1+E1
2250    PRINT "F=";F1 : PRINT
3040 NEXT K
5000 END
10000 '--------------------------
10005 ' Number of data
10010 DATA 5
10020 ' Units of effort (xi)
10030 DATA 7,5,10,8,4
10040 ' Size of i th sample (ri)
10050 DATA 700,465,884,636,293
10060 '--------------------------
```

Program B. This program requires the point estimator by the maximum likelihood method. Searching algorithm is NEWTON's method.

```
5  '-------------------------------------
10 '    LESLIE's method
20 '         Maximum likelihood method
30 '         NEWTON's method
40 '                 by TATSURO AKAMINE
50 '                    1989/8/02
60 '-------------------------------------
1010 GOSUB *INIT1
2010 READ P0
2020 FOR K=1 TO 20
2030    B1=0 : B2=0 : C1=0 : C2=0 : E1=0
2040    FOR I=1 TO M1
2050       X2=X1(I)/(1-X1(I)*P0)
2090       B1=B1+X2
2100       C1=C1+X2*X2
2110       B2=B2+X2*IR2(I)
2120       C2=C2+X2*X2*IR2(I)
2130       E1=E1+LOG(1-X1(I)*P0)
2140    NEXT I
2150    A1=IRSUM/P0+B2
2160    EN=A1/B1
2165    PRINT "P=";P0 , "N=";EN ,
2170    D1=0 : D2=0
2180    GOSUB *OYMC1
2190    GOSUB *OYMC2
2220    A2=-IRSUM/P0/P0+C2
2230    EN2=EN*(A2/A1-C1/B1)
2240    F1=D1+E1
2245    PRINT "F=";F1
2250    F2=-EN2*D2-B1
3010    DELTAP0=-F1/F2
3020    P0=P0+DELTAP0
3030    PRINT
3040 NEXT K
5000 END
10000 '---------------------------
10005 ' Number of data
10010 DATA 5
10020 ' Units of effort (xi)
10030 DATA 7,5,10,8,4
10040 ' Size of i th sample (ri)
10050 DATA 700,465,884,636,293
10060 ' Initial value of p
10070 DATA .05
10080 '---------------------------
```

Program C. This program requires the confidence region on the (p, n) plane by the method of the approximation to the normal distributions.

```
5  '---------------------------------------------
10 '  LESLIE's method
20 '     Approximation to normal distributions
30 '     The method on the (p,n) plane
40 '                        by TATSURO AKAMINE
50 '                           1989/8/07
60 '---------------------------------------------
1010 GOSUB *INIT1
1510 READ P0MIN,P0MAX,NP0
1520 Z1=1.96#
1530 P0DEL=(P0MAX-P0MIN)/NP0
2000 FOR K=1 TO M1
2001    PRINT "K=";K ,
2002    IR15=IR1(K)
2003    PRINT "R=";IR15 : PRINT
2010    FOR I=1 TO NP0
2020       P0=P0MIN+I*P0DEL
2030       P01=P0*X1(K) : X=P01
2040       D1=(X-1#)*(X-IR15*4#/Z1/Z1-1#)
2050       D2=SQR(D1)
2065       D2=Z1*Z1*D2
2070       X5=-(Z1*Z1*X-2#*IR15-Z1*Z1)
2080       EN5=(X5+D2)/X/2#
2090       EN6=(X5-D2)/X/2#
3020       ITR=IR2(K-1)
3030       END1=LOG(EN5+ITR)
3040       END2=LOG(EN6+ITR)
3050       PRINT "P=";P0,"N1=";END1,"N2=";END2
3060    NEXT I
3070    PRINT : PRINT
4000 NEXT K
5000 END
10000 '------------------------------------
10005 'Number of data
10010 DATA 5
10020 ' Units of effort (xi)
10030 DATA 7,5,10,8,4
10040 ' Size of i th sample (ri)
10050 DATA 700,465,884,636,293
10060 ' min p , max p , number of classes
10070 DATA 0#,0.02#,20
10080 '------------------------------------
```

Program D. This program requires the confidence interval of p on the $\partial L/\partial p=0$ curve by the method of the approximation to the normal distributions.

```
5  '------------------------------------------
10 '  LESLIE's method
20 '     Approximation to normal distributions
30 '     The method on the (dL/dp) curve
40 '                        by TATSURO AKAMINE
50 '                           1989/8/10
60 '------------------------------------------
1010 GOSUB *INIT1
2000 READ P0MIN,P0MAX,NP0
2010 P0DEL=(P0MAX-P0MIN)/NP0
2020 FOR K=1 TO NP0
2030    P0=P0MIN+K*P0DEL-P0DEL/2#
2035    B1=0 : B2=0 : E1=0
2040    FOR I=1 TO M1
2050       X2=X1(I)/(1-X1(I)*P0)
2090       B1=B1+X2
2110       B2=B2+X2*IR2(I)
2130       E1=E1+LOG(1-X1(I)*P0)
2140    NEXT I
2150    A1=IRSUM/P0+B2
2160    EN=A1/B1
2161    PRINT "P=";P0 , "N=";EN
3000    FOR I=1 TO M1
3005       EN2=EN-IR2(I-1) : P02=X1(I)*P0
3010       A5=IR1(I)-EN2*P02
3020       B5=EN2*P02*(1#-P02)
3030       B6=SQR(B5)
3050       Z1=A5/B6
3060       PRINT Z1,
3080    NEXT I
3085    PRINT : PRINT
3090 NEXT K
5000 END
10000 '-----------------------------------
10005 ' Number of data
10010 DATA 5
10020 ' Units of efforts (xi)
10030 DATA 7,5,10,8,4
10040 ' Size of i th sample (ri)
10050 DATA 700,465,884,636,293
10060 ' min p , max p , number of classes
10070 DATA 0#,0.02#,50
10080 '-----------------------------------
```

Program E. This program requires the value of likelihood on the $\partial L/\partial p=0$ curve for the likelihood ratio test.

```
5    '---------------------------------------
10   '    LESLIE's method
20   '         Calculation of likelihood
30   '         The method on (dL/dp) curve
40   '                  by TATSURO AKAMINE
50   '                     1989/9/04
60   '---------------------------------------
1010 GOSUB *INIT1
1600 PAI=3.141592653589793#
2000 READ P0MIN,P0MAX,NP0
2010 DELTAP0=(P0MAX-P0MIN)/NP0
2020 FOR K=1 TO NP0
2030    P0=P0MIN+K*DELTAP0
2035    B1=0 : B2=0
2040    FOR I=1 TO M1
2050       X2=X1(I)/(1-X1(I)*P0)
2090       B1=B1+X2
2110       B2=B2+X2*IR2(I)
2140    NEXT I
2150    A1=IRSUM/P0+B2
2160    EN=A1/B1
2165    IF EN<IRSUM THEN END
3010    GP=(EN+.5#)*LOG(EN)-(EN-IRSUM+.5#)*LOG(EN-IRSUM)
           -M1*.5#*LOG(PAI*2#)+(1#/EN-1#/(EN-IRSUM))/12#
3020    FOR I=1 TO M1
3025       R1=IR1(I)
3030       GP=GP-(IR1(I)+.5#)*LOG(R1)
3040       GP=GP-1#/IR1(I)/12#
3065       GP=GP+IR1(I)*LOG(X1(I)*P0)
3070       GP=GP+(EN-IR2(I))*LOG(1#-X1(I)*P0)
3080    NEXT I
3090    PRINT P0,EN,GP
4000 NEXT K
5000 END
10000 '-----------------------------------
10005 ' Number of data
10010 DATA 5
10020 ' Units of effort (xi)
10030 DATA 7,5,10,8,4
10040 ' Size of i th sample (ri)
10050 DATA 700,465,884,636,293
10060 ' min p , max p , number of classes
10070 DATA 0#,0.02#,200
10080 '-----------------------------------
```

Program F. These are subroutines for program A~E.

```
19000 '---------------------------
19010 ' Subroutines for programs
19020 '---------------------------
20000 *INIT1
20010   DEFINT I-N
20020   DEFDBL A-H,O-Z
20030   READ M1
20035   DIM X1(M1),IR1(M1),IR2(M1)
20040   XMAX=0 : IRSUM=0
20050   FOR I=1 TO M1
20060      READ X1(I)
20070      IF X1(I)>XMAX THEN XMAX=X1(I)
20080   NEXT I
20090   FOR I=1 TO M1
20100      READ IR1(I)
20110      IRSUM=IRSUM+IR1(I) : IR2(I)=IRSUM
20120   NEXT I
20130   IR2(0)=0
20200  RETURN
21000 *OYMC1
21010   S1=EN-IRSUM+1 : S2=EN
21020   S5=S1*S1 : S6=S2*S2
21030   D1=(1/S1+1/S2)/2
21040   D1=D1+LOG(S2)-LOG(S1)
21050   D1=D1-(1/S6-1/S5)/12
21060   D1=D1+(1/S6/S6-1/S5/S5)/120
21100  RETURN
22000 *OYMC2
22030   D2=(1/S5+1/S6)/2
22040   D2=D2-1/S2+1/S1
22050   D2=D2-(1/S6/S2-1/S5/S1)/6
22060   D2=D2+(1/S6/S6/S2-1/S5/S5/S1)/30
22100  RETURN
23000 *OYMC11
23010   FOR I=1 TO IRSUM
23020      D1=D1+1/(EN-I+1)
23040   NEXT I
23050  RETURN
24000 *OYMC1122
24010   FOR I=1 TO IRSUM
24020      D1=D1+1/(EN-I+1)
24030      D2=D2+1/(EN-I+1)/(EN-I+1)
24040   NEXT I
24050  RETURN
```