日水研報告，(39): 19—35, 1989.
Bull. Japan Sea Reg. Fish. Res. Lab., (39): 19—35, 1989.

An Interval Estimation for the PETERSEN Method using Bayesian Statistics

Tatsuro AKAMINE[1)]

Abstract

The statistical model for the PETERSEN method is a hypergeometric distribution. Approximation to a binomial distribution has been used, and the usual method for this binomial model is based on approximation to a normal distribution. The Bayesian statistical model for a binomial distribution, which assumes that the prior distribution of parameter is uniform, corresponds well with the conventional method. However, the Bayesian statistical method for a hypergeometric distribution which assumes the uniform prior distribution is not feasible. The prior distribution according to the inverse squared parameter is natural for this model. Beta function and zeta function are important to understand these methods. This model is simpler to understand and easier to calculate by micro-computer than the conventional method.

Key words Bayesian statistics, PETERSEN method, hypergeometric distribution, binomial distribution, beta function, zeta function

Introduction

The PETERSEN method is the simplest method to employ in mark-recapture experiments. Let N: total number of individuals to estimate, M: number of marked individuals in N, n: total number of sampled individuals, m: number of mark-recaptured individuals in n (Fig. 1). Mean and variance of estimater N are well known, as follows:

$$E(N)=\frac{Mn}{m}, \tag{1.1}$$

$$V(N)=\frac{M^2n(n-m)}{m^3} \quad \text{or} \tag{1.2}$$

$$V(N)=\frac{M(M-m)n(n-m)}{m^3}. \tag{1.3}$$

But, these are not efficient for interval estimation.

In this paper, interval estimation based on Bayesian statistics is developed.

Accepted: October 20, 1988. Contribution A *No.* 454 from the Japan Sea Regional Fisheries Research Laboratory.

1) Japan Sea Regional Fisheries Research Laboratory, 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)

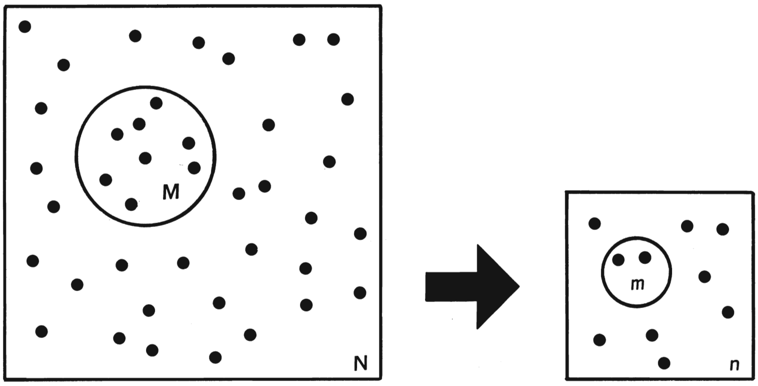

Fig. 1. The image of the PETERSEN method.

First, the method for a binomial distribution model is shown and compared with the conventional method. Next, the method for a hypergeometric distribution model is shown. These methods are logical extensions of the conventional method.

The method for a binomial distribution model

1. Hypergeometric distribution

The statistical model for the PETERSEN method is a hypergeometric distribution. It is expressed as follows:

$$P(N, m)=\frac{\binom{M}{m}\binom{N-M}{n-m}}{\binom{N}{n}}. \tag{2.1}$$

Where, the binomial coefficient (number of combinations):

$$\binom{a}{b}=\frac{a!}{b!(a-b)!}=\frac{a(a-1)\cdots\cdots(a-b+1)}{b(b-1)\cdots\cdots 1}. \tag{2.2}$$

This expression gives probability of m for fixed N, M and n. Then,

$$\sum_{m=0}^{n} P(N, m)=1. \tag{2.3}$$

This equation is proved easily by the binomial theorem:

$$(p+q)^a=\sum_{b=0}^{a}\binom{a}{b}p^b q^{a-b}. \tag{2.4}$$

Expanding both sides of

$$(1+x)^M(1+x)^{N-M}=(1+x)^N \tag{2.5}$$

by (2.4) and then equating coefficients of x to the nth power gives (2.3).

On the other hand,

$$P(N, m)=\frac{\binom{M}{m}\binom{N-M}{n-m}}{\binom{N}{n}}$$

$$=\frac{M!}{m!(M-m)!}\cdot\frac{(N-M)!}{(n-m)!(N-M-n+m)!}\cdot\frac{n!(N-n)!}{N!}$$

$$=\frac{\binom{n}{m}\binom{N-n}{M-m}}{\binom{N}{M}}. \tag{2.6}$$

Then M and n are interchangeable. There are two expressions for estimator N as follows:

$$N=\frac{M}{m/n}=\frac{M}{p},\qquad N=\frac{n}{m/M}=\frac{n}{s}. \tag{2.7}$$

In this paper, let $M \geqq n$. Then, the next approximation is useful for the former expression of (2.7). If $M<n$, then, let interchange M and n.

2. Approximation to a binomial distribution

A hypergeometric distribution is approximated to a binomial distribution as follows:

$$P(N, m)=\binom{n}{m}\frac{M\cdots\cdots(M-m+1)(N-M)\cdots\cdots(N-M-n+m+1)}{N\cdots\cdots(N-n+1)}$$

$$=\binom{n}{m}\frac{M}{N}\frac{M-1}{N-1}\cdots\cdots\frac{M-m+1}{N-m+1}\frac{N-M}{N-m}\cdots\cdots\frac{N-M-n+m+1}{N-n+1}$$

When $N \gg n$ and $M \gg m$, this is approximated to

$$P(N, m)\doteqdot\binom{n}{m}\left(\frac{M}{N}\right)^m\left(\frac{N-M}{N}\right)^{n-m}=\binom{n}{m}p^m q^{n-m}$$

$$=P(p, m). \tag{2.8}$$

Where, $p=M/N$, $\quad p+q=1$.

3. Bayesian statistical method

The following equation is essential for this method.

$$S=\int_0^1\binom{n}{m}p^m q^{n-m}\mathrm{d}p=\frac{1}{n+1}. \tag{2.9}$$

The proof of this equation is shown in the next chapter.

The Bayesian statistical method is demonstrated as follows: Let θ: parameter to estimate, t: data, $P(\theta, t)$: probability of data for each θ, $P^\circ(\theta)$: prior distribution of θ. Where,

$$\sum P^\circ(\theta)=1. \tag{2.10}$$

Let $P^*(\theta)$: posterior distribution of θ. Then,

$$P^*(\theta)=\frac{P^\circ(\theta)P(\theta, t)}{\sum P^\circ(\theta)P(\theta, t)}. \tag{2.11}$$

Where,

$$\sum P^*(\theta)=1. \tag{2.12}$$

In particular, the prior distribution of θ is a uniform distribution:

$$P^\circ(\theta)\equiv\varepsilon=\text{const}. \tag{2.13}$$

The posterior distribution becomes as follows:

$$P^*(\theta)=\frac{P(\theta,t)}{\sum P(\theta,t)}. \tag{2.14}$$

In this case, θ is p. It is natural to set the prior distribution of p to a uniform distribution from 0 to 1. Although (2.11) and (2.14) are equations for discrete distributions, the distribution of p is a continuous distribution. Let $\sum\to\int d\theta$ for (2.11) and (2.14) to obtain equations for a continuous distribution. Then, prior and posterior distributions of p become as follows by (2.9).

$$P^\circ(p)\equiv 1=\text{const}. \tag{2.15}$$

$$P^*(p)=(n+1)P(p,m). \tag{2.16}$$

MATSUBARA (1985) stated that this model is the original Bayesian theorem by THOMAS BAYES.

4. Beta function

The beta function is defined as follows:

$$B(a,b)=\int_0^1 x^{a-1}(1-x)^{b-1}dx. \tag{2.17}$$

On the other hand, the gamma function is defined as follows:

$$\Gamma(a)=\int_0^\infty e^{-x}x^{a-1}dx. \tag{2.18}$$

The following formulas are well known.

$$B(a,b)=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \tag{2.19}$$

$$\Gamma(a+1)=a!. \tag{2.20}$$

Then, the following equation is easily obtained.

$$B(a+1,b+1)=\frac{a!\,b!}{(a+b+1)!}=\frac{1}{a+b+1}\,\frac{1}{\binom{a+b}{a}}. \tag{2.21}$$

The gamma function is regarded as a generalized factorial function and the beta function is considered to be a generalized (reciprocal number of) combination. These relations give (2.9) easily, as follows:

$$S=\binom{n}{m}B(m+1,n-m+1)=\frac{1}{n+1}. \tag{2.9'}$$

Alternatively, the 'integration by parts' gives the same result as follows:

$$B(a+1,b+1)=\int_0^1 x^a(1-x)^b\,dx$$

$$=\left[\frac{1}{a+1}x^{a+1}(1-x)^b\right]_0^1+\frac{b}{a+1}\int_0^1 x^{a+1}(1-x)^{b-1}\,\mathrm{d}x$$

$$=\frac{b}{a+1}B(a+2,\,b)\,. \tag{2.22}$$

Recursion of (2. 22) gives,

$$B(a+1,\,b+1)=\frac{b}{a+1}\cdot\frac{b-1}{a+2}\cdots\cdots\frac{1}{a+b}B(a+b+1,\,1)\,.$$

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

Then, we get (2. 9′).

5. A BASIC program

Although calculations of $a!$ are not easy because $a!$ is too large, $P(p,\,m)$ is not large. In this paper, $P(p,\,m)$ is calculated as follows: The binomial coefficient is calculated by the following equation:

$$\binom{a}{b}=\prod_{i=0}^{b-1}\frac{a-i}{b-i}\,. \tag{2.2'}$$

Then, multiplying by p and q we get $P(p,\,m)$.

Let $f(p)$ be as follows:

$$f(p)=p^m q^{n-m}=p^m(1-p)^{n-m}, \tag{2.23}$$

$$\frac{\mathrm{d}f}{\mathrm{d}p}=\frac{m-np}{p(1-p)}f(p)\,. \tag{2.24}$$

Therefore, when $p=m/n$, $P(p,\,m)$ is max.

An example of a BASIC program is shown in 'Program 1'. This program calculates only the sum of $P(p,\,m)$. It is regarded as the 'middle point formula' or 'trapezoidal formula' of numerical integrations.

[Example 1] Estimate N when $M=2000$, $n=100$ and $m=20$.

Point estimation is $N=Mn/m=10000$.

Interval estimation is as follows: The result of Program 1 shown in Table 1 and Fig. 2. The 95% interval is $p=0.13\sim0.28$. Then, $N=7100\sim15400$. Although this interval of p is minimum, that of N is not.

6. The conventional method

The non-Bayesian statistical method is as follows: When $n\to\infty$, a binomial distribution approaches a normal distribution:

$$N(\mu,\,\sigma)=\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}\,. \tag{2.25}$$

Where $\mu=np$, $\sigma=\sqrt{npq}$.

Let z be as follows:

$$z=\frac{np-m}{\sqrt{npq}},\qquad p+q=1\,. \tag{2.26}$$

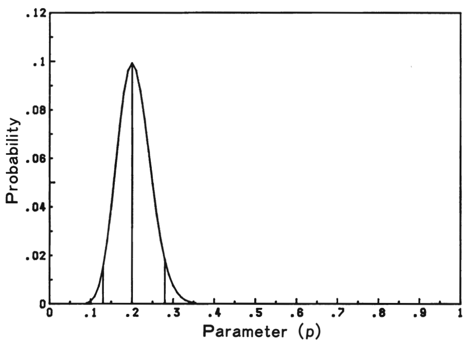

Fig. 2. The graph of $P(p, m)$ when $n=100$ and $m=20$.

Table 1. Values of binomial distributions: $P(p, m)$ when $n=100$ and $m=20$.

p	P	$101\times\sum P$
.00	.00000000	.00000000
.01	2.398650×10^{-20}	2.422637×10^{-18}
.12	.00743687	1.23904116
.13	.01477606	2.73142296
.14	.02579812	5.33703350
.19	.09616673	41.02511873
.20	.09930021	51.05444043
.21	.09631735	60.78249251
.27	.02643963	95.12671528
.28	.01815146	96.96001313
.29	.01196132	98.16810662
.40	.00001053	99.99923869
.60	2.864017×10^{-19}	100.00000000
1.00	.00000000	100.00000000

Then z distributes according to $N(0, 1)$. The confidence interval for z is easily obtained (ex. 95% confidence interval is $-1.96\leqq z\leqq 1.96$). From (2.26)

$$p=\frac{m}{n}+z\sqrt{\frac{pq}{n}}\,. \qquad (2.27)$$

Let n of the last term be fixed. Then, we get the rough estimator as variance of p:

$$V(p)=\sigma^2(p)=\frac{pq}{n}=\frac{1}{n}\,\frac{m}{n}\left(1-\frac{m}{n}\right)=\frac{m(n-m)}{n^3}\,. \qquad (2.28)$$

The rough 95% confidence interval is given by $m/n\pm2\sigma(p)$.

On the other hand, from $N=M/p$

$$\mathrm{d}N=-\frac{M}{p^2}\,\mathrm{d}p\,. \qquad (2.29)$$

This is approximated to

$$\Delta N^2=\frac{M^2}{p^4}\,\Delta p^2\,. \qquad (2.30)$$

This is regarded as the 'error propagation rule'. From (2.28) and (2.30)

$$V(N)=\frac{M^2}{p^4}\,V(p)=\frac{M^2n^4}{m^4}\,\frac{m(n-m)}{n^3}=\frac{M^2n(n-m)}{m^3}\,.$$

Then, we get (1.2).

The variance of a hypergeometric distribution is as follows:

$$V(m)=\frac{N-n}{N-1}\,npq\,. \qquad (2.31)$$

On the other hand, the variance of a binomial distribution is npq. Multiply,

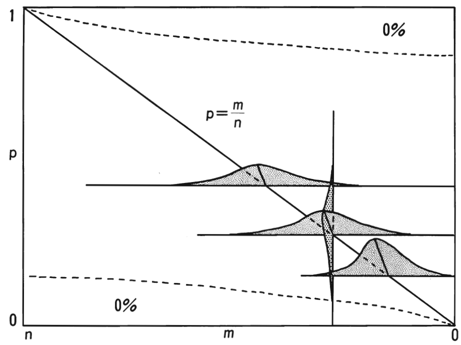

Fig. 3. The image of (m, p) coordinates.

$$\frac{N-n}{N-1} \fallingdotseq \frac{N-n}{N} = \frac{M-m}{M}$$

into (1. 2), then we get (1. 3).

7. Relation between the Bayesian statistical method and the conventional method

Fig. 3 shows an image of (m, p) coordinates. From (2. 26)

$$dz = \frac{n\{(n-2m)p+m\}}{2\{np(1-p)\}^{3/2}} dp . \tag{2. 32}$$

And (2. 25) shows that the ratio of the height of $N(\mu, \sigma)$ to that of $N(0, 1)$ is $1/\sqrt{\sigma}$. Then, from Fig. 3

$$P(p) = P(z)/\sqrt{npq} . \tag{2. 33}$$

Substituting (2. 32) and (2. 33) for the following equation:

$$P(p)dp = T(z)P(z)dz . \tag{2. 34}$$

Then, we get

$$T(z) = \frac{2p(1-p)}{(n-2m)p+m} . \tag{2. 35}$$

Therefore, the following equations hold:

$$\int_{-\infty}^{\infty} P(z)dz = 1 \tag{2. 36}$$

$$\int_0^1 P(p)dp = \int_{-\infty}^{\infty} T(z)P(z)dz \fallingdotseq \frac{1}{n+1} \tag{2. 37}$$

The graph of $(z, T(z))$ for Example 1 is shown in Fig. 4–a.

AKAMINE (1989) demonstrates the case of extraction. In this case, n is required when p and m is fixed. From (2. 26)

$$dz = \frac{np+m}{2n\sqrt{npq}} dn . \tag{2. 38}$$

Similar to (2. 35), we obtain

$$T(z) = \frac{2n}{np+m} . \tag{2. 39}$$

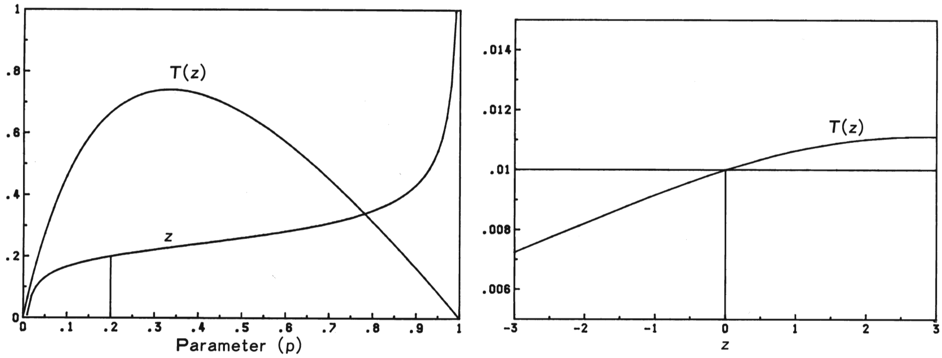

Fig. 4-a. The graph of $T(z)$ for the PETERSEN method.

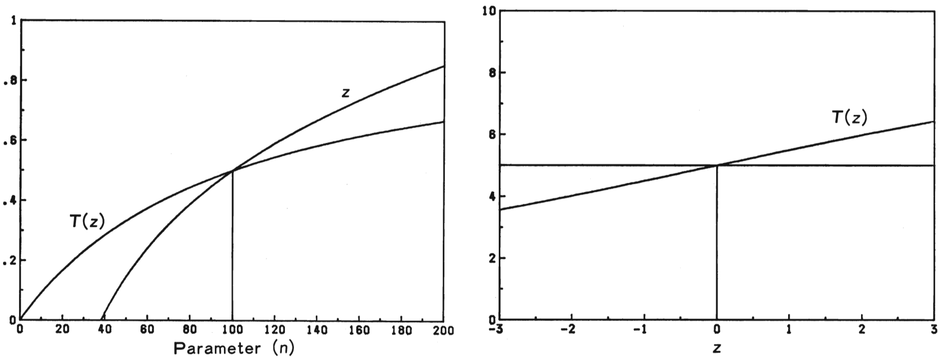

Fig. 4-b. The graph of $T(z)$ for extraction.

The graph of $(z, T(z))$ for Example 1 of AKAMINE (1989) is shown in Fig. 4-b. Comparing Fig. 4-a and b suggests that the bias of each model is almost equal. In AKAMINE (1989), 'Theorem 2' proves that precent points for the ordinate correspond to that for the abscissa in Fig. 5 (which is equivalent to Fig. 3 in this paper). Then, in this model, the same feature is expected. Fig. 5 suggests this relation.

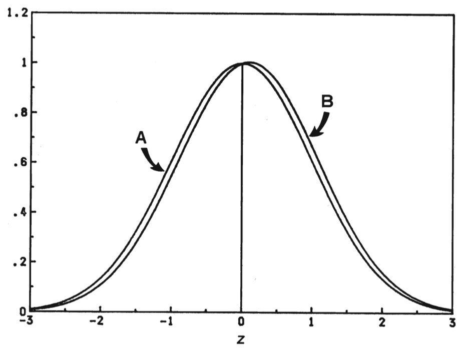

Fig. 5. Comparison of a normal distribution and $T(z)$.
A: $y=\exp(-z^2/2)$,
B: $y=(1+0.1z)\exp(-z^2/2)$.

On the right side of Fig. 4-b, the line $T(z)$ is as follows: When $z=0$, $n=m/p$,

$$T(z)=\frac{2n}{np+m} \to \frac{2n}{2m}=\frac{1}{p}.$$

On the other hand,

$$\frac{\mathrm{d}T}{\mathrm{d}z}=\frac{\mathrm{d}T}{\mathrm{d}n}\frac{\mathrm{d}n}{\mathrm{d}z}=\frac{4nm\sqrt{npq}}{(np+m)^3} \to \frac{1}{2}\sqrt{\frac{1-p}{mp^2}}. \tag{2.40}$$

Therefore, this line is

$$T(z)=\frac{1}{p}\left(1+\frac{1}{2}\sqrt{\frac{1-p}{m}}\,z\right). \tag{2.41}$$

In Fig. 5, the following equation holds.

$$\int_{-\infty}^{\infty}(a+bx)\exp(-x^2/2)\,\mathrm{d}x=a\int_{-\infty}^{\infty}\exp(-x^2/2)\,\mathrm{d}x \tag{2.42}$$

[Example 1′] Estimate N by the non-Bayesian method when $M=2000$, $n=100$ and $m=20$.

95% confidence interval becomes as follows:

Let $z=\pm 1.96 \doteqdot \pm 2$. Then, (2.26) becomes

$4=(100p-20)^2/100p(1-p)$

$p-p^2=(5p-1)^2$

$26p^2-11p+1=0$

$p=0.21\pm 0.08=0.13,\ 0.29$

Then, $N=6900\sim 15400$.

On the other hand, from (1.2),

$V(N)=4000000$, then $\sigma(N)=2000$,

$N=Mn/m\pm 2\sigma(N)=10000\pm 4000$.

From (1.3),

$V(N)=3960000$, then $\sigma(N)=1990$,

$N=10000\pm 3980$.

The result of (2.26) almost corresponds with the Bayesian statistical method (Example 1). This is logical from Fig. 5.

The model for a hypergeometric distribution

1. RIEMANN's zeta function

when $N\to\infty$, $P(N, m)$ becomes as follows:

$$P(N, m)=a\frac{(N-M)\cdots\cdots(N-M-n+m+1)}{N\cdots\cdots(N-n+1)}$$

$$=\frac{a\left(1-\frac{M}{N}\right)\cdots\cdots\left(1-\frac{M+n-m-1}{N}\right)}{N^m\left(1-\frac{1}{N}\right)\cdots\cdots\left(1-\frac{n-1}{N}\right)}\xrightarrow[(N=\infty)]{}\frac{a}{N^m}. \tag{3.1}$$

Where, $a=\binom{n}{m}M\cdots\cdots(M-m+1)=\text{const.}$

Fig. 6 shows $P(N, m)$ when the prior distribution of N is uniform.

This is according to RIEMANN's zeta function:

$$\zeta(s)=\sum_{k=1}^{\infty}\frac{1}{k^s}=\frac{1}{1^s}+\frac{1}{2^s}+\frac{1}{3^s}+\cdots\cdots. \tag{3.2}$$

The convergence values of the zeta function are well known as follows:

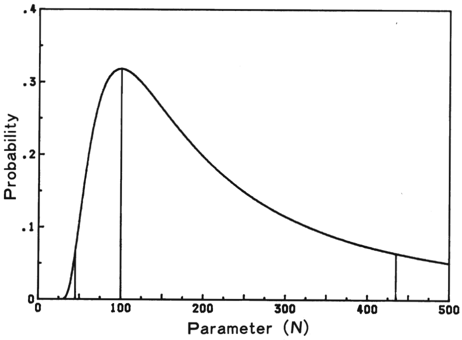

Fig. 6. The graph of $P(N, m)$ when $M=20$, $n=10$ and $m=2$.

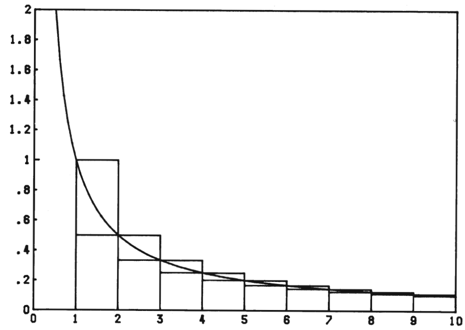

Fig. 7. The graph of $y=1/x^s$ and zeta function.

$$\zeta(s)=\begin{cases}\infty & (s\leqq 1)\\ \alpha<\infty & (s>1)\end{cases} \tag{3.3}$$

It is easy to understand (3. 3) by comparison with integration of

$$F(s)=\int_1^{\infty}\frac{1}{x^s}\,\mathrm{d}x \tag{3.4}$$

(See Fig. 7).

From (3. 3) when $m=0$, 1 we get

$$\sum_{N=M+n-m}^{\infty} P(N, m)=\infty. \tag{3.5}$$

Therefore, the prior distribution of N as a uniform distribution to ∞ is not adequate.

2. Bayesian statistical method

It is natural to let the prior distribution of N accord to the binomial distribution model in the previous chapter. The prior distribution of $p(=M/N)$ is a uniform distribution. From the equation:

$$\mathrm{d}p=-(M/N^2)\,\mathrm{d}N, \tag{3.6}$$

we get the following approximation:

$$\Delta p=-(M/N^2)\,\Delta N. \tag{3.6'}$$

Then, let the prior distribution of N be as follows:

$$P^{\circ}(N)=\frac{M+1}{(N+2)(N+1)}. \tag{3.7}$$

Where the following equations hold.

$$\sum_{N=M}^{\infty} P^{\circ}(N)=1, \tag{3.8}$$

$$\sum_{N=M+n-m}^{\infty} P^{\circ}(N)P(N, m)=\frac{1}{n+1}. \tag{3.9}$$

These equations are proved in the next chapter. Then the posterior distribution of N becomes as follows:

$$P^*(N)=\frac{P^\circ(N)P(N,m)}{\sum P^\circ(N)P(N,m)}=\frac{(M+1)(n+1)}{(N+2)(N+1)}P(N,m)\,. \tag{3.10}$$

For geometric images, let it be as follows:

$$S(N)=P^\circ(N)P(N,m). \tag{3.11}$$

Where S: area, P°: width and P: height. Although P is a discrete distribution, S is a continuous distribution. It is natural to use point estimation by P and interval estimation by S, because interval estimation is an original image for a continuous distribution.

3. Difference and summation

The factorial function is defined as follows:

$$x^{(r)}=x(x-1)(x-2)\cdots\cdots(x-r+1)\,. \tag{3.12}$$

In addition, the difference is defined as follows:

$$\Delta f(x)=f(x+1)-f(x)\,. \tag{3.13}$$

Then, we get

$$\begin{aligned}\Delta x^{(r)}&=(x+1)^{(r)}-x^{(r)}\\&=(x+1)x(x-1)\cdots\cdots(x-r+2)-x(x-1)(x-2)\cdots\cdots(x-r+1)\\&=rx(x-1)(x-2)\cdots\cdots(x-r+2)\\&=rx^{(r-1)}\,.\end{aligned} \tag{3.14}$$

On the other hand, The negative factorial function is defined as follows:

$$x^{(-r)}=\frac{1}{(x+r)^{(r)}}=\frac{1}{(x+r)(x+r-1)\cdots\cdots(x+1)}\,. \tag{3.15}$$

Then, we get

$$\begin{aligned}\Delta x^{(-r)}&=(x+1)^{(-r)}-x^{(-r)}\\&=\frac{1}{(x+r+1)^{(r)}}-\frac{1}{(x+r)^{(r)}}\\&=\frac{1}{(x+r+1)(x+r)\cdots\cdots(x+2)}-\frac{1}{(x+r)(x+r-1)\cdots\cdots(x+1)}\\&=\frac{-r}{(x+r+1)(x+r)\cdots\cdots(x+2)(x+1)}\\&=-rx^{(-r-1)}\,.\end{aligned} \tag{3.14$'$}$$

Therefore, (3.14) holds when r is negative.

The difference of the product becomes as follows:

$$\begin{aligned}\Delta\{f(t)g(t)\}&=f(t+1)g(t+1)-f(t)g(t)\\&=f(t+1)g(t+1)-f(t)g(t+1)+f(t)g(t+1)-f(t)g(t)\end{aligned}$$

$$=\{f(t+1)-f(t)\}g(t+1)+f(t)\{g(t+1)-g(t)\}$$

$$=\Delta f(t)g(t+1)+f(t)\Delta g(t). \tag{3.16}$$

On the other hand, from (3. 13)

$$\Delta f(a)=f(a+1)-f(a),$$

$$\Delta f(a+1)=f(a+2)-f(a+1),$$

$$\cdots\cdots\cdots\cdots\cdots$$

$$\cdots\cdots\cdots\cdots\cdots$$

$$\Delta f(b)=f(b+1)-f(b).$$

Summarizing both sides gives the following result:

$$\sum_{x=a}^{b}\Delta f(x)=f(b+1)-f(a).$$

Let $h(x)=\Delta f(x)$ and $\Delta^{-1}h(x)=f(x)$. Where, Δ^{-1}: summation. Then, we get the summation formula:

$$\sum_{x=a}^{b}h(x)=\Delta^{-1}h(b+1)-\Delta^{-1}h(a)$$

$$=\Big[\ \Delta^{-1}h(x)\ \Big]_a^{b+1}. \tag{3.17}$$

And, from (3. 14)

$$\Delta^{-1}x^{(r)}=\frac{1}{r+1}x^{(r+1)}. \tag{3.14$''$}$$

From the difference of the product

$$\Delta f(t)g(t)=f(t)g(t-1)-f(t)\Delta g(t-1),$$

$$h(t)g(t)=\Delta^{-1}h(t)g(t-1)-\Delta^{-1}f(t)\Delta g(t-1).$$

Then, we get the 'summation by parts' formula:

$$\sum_{a}^{b}h(t)g(t)=\Big[\Delta^{-1}h(t)g(t-1)\Big]_a^{b+1}-\sum_{a}^{b}\Delta^{-1}f(t)\Delta g(t-1). \tag{3.18}$$

Therefore, (3.8) is proved as follows:

$$\sum_{N=M}^{\infty}P^{\circ}(N)=\sum_{N=M}^{\infty}\frac{M+1}{(N+2)(N+1)}=(M+1)\sum_{N=M}^{\infty}N^{(-2)}$$

$$=(M+1)\Big[\ -N^{(-1)}\ \Big]_M^{\infty}=1.$$

And (3. 9) is proved as follows:

$$\sum_{N=M+n-m}^{\infty}P^{\circ}(N)P(N,m)=\binom{n}{m}(M+1)^{(m+1)}\sum_{N=M+n-m}^{\infty}\frac{(N-M)^{(n-m)}}{(N+2)^{(n+2)}}$$

Where,

$$\sum_{N=M+n-m}^{\infty}\frac{(N-M)^{(n-m)}}{(N+2)^{(n+2)}}=\sum_{N=M+n-m}^{\infty}(N-M)^{(n-m)}(N-n)^{(-n-2)}$$

$$=\Big[-\frac{1}{n+1}(N-M-1)^{(n-m)}(N-n)^{(-n-1)}\Big]_{M+n-m}^{\infty}$$

$$+\frac{n-m}{n+1}\sum_{N=M+n-m}^{\infty}(N-M-1)^{(n-m-1)}(N-n)^{(-n-1)}$$

The first term of the right side is equal to 0. The result from recursion of this equation is

$$\frac{n-m}{n+1}\ \frac{n-m-1}{n}\cdots\cdots\frac{1}{m+2}\ \sum_{N=M+n-m}^{\infty}(N-n)^{(-m-2)}.$$

Where, the last term becomes

$$\left[-\frac{1}{m+1}(N-n)^{(-m-1)}\right]_{M+n-m}^{\infty}=\frac{1}{m+1}(M-m)^{(-m-1)}$$

$$=\frac{1}{m+1}\ \frac{1}{(M+1)^{(m+1)}}.$$

Then,

$$\sum_{N=M+n-m}^{\infty}\frac{(N-M)^{(n-m)}}{(N+2)^{(n+2)}}=\frac{1}{n+1}\ \frac{1}{(M+1)^{(m+1)}}\ \frac{m!(n-m)!}{n!}$$

Therefore, (3.9) is proved.

4. A BASIC program

From (2.1) we get the following recurrence formula:

$$P(N+1,\ m)=\frac{(N-M+1)(N-n+1)}{(N+1)(N-M-n+m+1)}P(N,\ m). \qquad (3.19)$$

Let $f(N)$ be as follows:

$$f(N)=\frac{(N-M+1)(N-n+1)}{(N+1)(N-M-n+m+1)}$$

$$=1+\frac{Mn-(N+1)m}{(N+1)(N-M-n+m+1)}. \qquad (3.20)$$

Then, $f(N)$ is monotone decreasing. In addition,

$$f\left(\frac{Mn}{m}-1\right)=1. \qquad (3.21)$$

Let $N_0=\max\ (N|N\leqq Mn/m)$ where N is natural number. The result from (3.21) is:

$$\max P(N)=P(N_0). \qquad (3.22)$$

In addition, when $N_0=Mn/m$,

$$\max P(N)=P(N_0)=P(N_0-1). \qquad (3.22')$$

Therefore, point estimation of N is N_0.

On the other hand, from (3.19)

$$P(N-1,\ m)=\frac{N(N-M-n+m)}{(N-M)(N-n)}P(N,\ m). \qquad (3.19')$$

A BASIC program using (3.19) and (3.19′) is shown in 'Program 2'.

[Example 2] Estimate N when $M=20$, $n=10$, $m=2$.

Point estimation is $N=Mn/m=100$.

Table 2. Values of hypergeometric distributions: $P(N, m)$ when $M=20$, $n=10$ and $m=2$.

N	P	$\sum S$
28	1.447828×10^{-5}	3.844232×10^{-6}
42	.04129029	.02568505
43	.04858810	.03135366
44	.05631803	.03763843
45	.06441605	.04452099
46	.07281814	.05197711
88	.31116199	.50152099
98	.31808012	.58258717
99	.31817063	.58986414
100	.31817063	.59699843
101	.31808432	.60399229
434	.06455245	.98758007
435	.06430816	.98765804
436	.06406522	.98773536
500	.05094430	.99159015

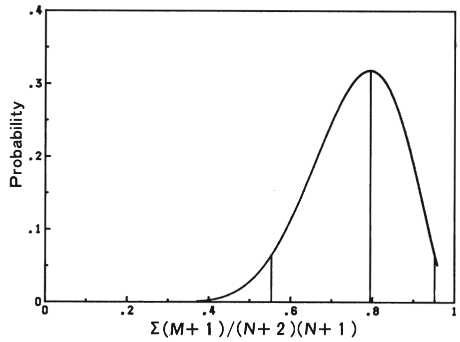

Fig. 8. The posterior distribution of Example 1. Each height shows $P(N, m)$, width shows prior probability and area shows posterior probability.

Interval estimation is as follows: The result of Program 2 is shown in Table 2 and Fig. 8. The 95% interval is $N=45\sim435$. The interval of N is not minimum, because that of p is minimum.

[Example 2′] Estimate 95% confidence interval for the above problem using the conventional method.

(a) From (1.2), $V(n)=4000$, $\sigma(n)=63$, $N=Mn/m\pm2\sigma=100\pm126$, then $N=-26\sim226$.

(b) From (1.3), $V(n)=3600$, then $N=-20\sim220$.

(c) Let $z=\pm1.96\doteqdot\pm2$. Then (2.26) becomes

$$4=(10p-2)^2/10p(1-p)$$

$$35p^2-20p+1=0$$

$$p=0.286\pm0.230=0.06,\ 0.52$$

Then $N=38\sim333$.

(d) Results of 'Program 1' are as follows:

(d–1) Both side 2.5% points are $p=0.06$, 0.52, then $N=38\sim333$. This is in correspondence with (c).

(d–2) The minimum interval (highest density credibility interval) of p is $p=0.04\sim0.48$, then $N=42\sim500$.

Example 1 and 2 leads to the following conclusion:

(1) For small samples, Program 2 is the best.

(2) For large samples, Program 1 is better, but the conventional method is sufficiently useful.

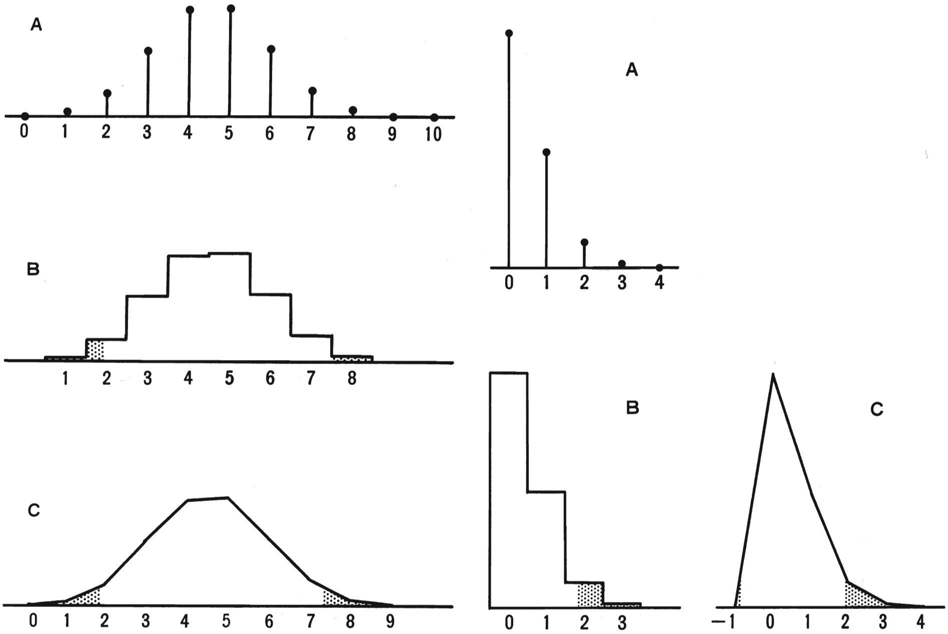

Fig. 9. Confidence intervals for $P(N, m)$ when $N=44$ (left), 436 (right), $M=20$, $n=10$ and $m=2$. A: $P(N, m)$. B: the case for the middle point rule. C: the case for the trapezoidal rule.

5. Discussion

In Example 2, $N=436$ is cut off even at $P(436,2)=6.4\%$. It is difficult to estimate the confidence interval with a discrete distribution. In Fig. 9, two types of confidence intervals are shown for $N=44$, 436. These appear to suppose the Bayesian statistical method.

Acknowledgements

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

References

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

MATSUBARA, N. (1985) *Ishi kettei no kiso* (*Base for determination*). *Shinban*, Asakurashoten, Tokyo, 169–195. (in Japanese)

ベイズ統計によるピーターセン法の区間推定

赤 嶺 達 郎

ピーターセン法は超幾何分布と一致する．従来は二項分布に近似し，さらに正規分布に近似することによって区間推定を行ってきた．母数の事前分布を一様分布と仮定するベイズ統計モデルは従来の手法とよく一致する．しかし，超幾何分布においては一様分布は事前分布として不合理である．このモデルには母数の逆二乗に従う事前分布が自然である．これらの手法を理解するためにはベータ関数とゼータ関数が重要である．このモデルは従来の手法と比較して単純で理解しやすく，小型計算機による計算も容易である．

Appendix

Program 1. An example of BASIC program to culculate a confidence interval for the PETERSEN method using Bayesian statistics based on binomial distributions.

```
100  '----------------------------------------------------------
110  '        Interval estimation for Petersen method
120  '                 (Binomial distribution)
130  '                              by Tatsuro Akamine
140  '                                         1988-08-31
150  '----------------------------------------------------------
1000 DEFINT I-N
1010 DEFDBL A-H,O-Z
1020 N2=100 : M2=20
1040 AREA=0#
1050 FOR I=0 TO 100
1060    P1=I/100# : Q1=1#-P1 : C1=1#
1070    FOR J=0 TO M2-1
1080       C1=C1*(N2-J)/(M2-J)*P1
1090    NEXT J
2000    FOR K=1 TO N2-M2
2010       C1=C1*Q1
2020    NEXT K
2030    AREA=AREA+C1*(N2+1)
2040    PRINT P1, C1, AREA
3000 NEXT I
3010 END
```

Program 2. An example of BASIC program to calculate a confidence interval for the PETERSEN method using Bayesian statistics based on hypergeometric distributions.

```
100  '-----------------------------------------------------
110  '        Interval estimation for Petersen method
120  '             (Hypergeometric distribution)
130  '                                 by Tatsuro Akamine
140  '                                              1988-08-31
150  '-----------------------------------------------------
1000 DEFINT I-N
1010 DEFDBL A-H,O-Z
1020 M1=20 : N2=10 : M2=2
1030 N1=INT(M1*N2/M2)
1040 N9=M1  : M9=M2 : GOSUB *COMBI : C5=C1
1050 N9=N1-M1 : M9=N2-M2 : GOSUB *COMBI : C6=C1
1060 N9=N1  : M9=N2 : GOSUB *COMBI : C7=C1
1070 PROB=C5*C6/C7
1080 AREA=PROB*(M1+1)*(N2+1)/(N1+2)/(N1+1)
1090 PRINT N1,PROB,AREA
2000 BPRO=PROB*(N1-M1+1)*(N1-N2+1)/(N1+1)/(N1-M1-N2+M2+1)
2010 SPRO=PROB*N1*(N1-M1-N2+M2)/(N1-M1)/(N1-N2)
2020 N1S=N1  : N1B=N1
2030 *REPEAT
2040  IF BPRO>SPRO GOTO *RIGHT
3000 *LEFT
3010  N1S=N1S-1
3020  N1S9=N1S : SPRO9=SPRO
3030  AREAS=SPRO*(M1+1)*(N2+1)/(N1S+2)/(N1S+1)
3040  AREA=AREA+AREAS
3050  SPRO=SPRO*N1S*(N1S-M1-N2+M2)/(N1S-M1)/(N1S-N2)
3060  GOTO *CHECK
4000 *RIGHT
4010  N1B=N1B+1
4020  N1B9=N1B : BPRO9=BPRO
4030  AREAB=BPRO*(M1+1)*(N2+1)/(N1B+2)/(N1B+1)
4040  AREA=AREA+AREAB
4050  BPRO=BPRO*(N1B-M1+1)*(N1B-N2+1)/(N1B+1)/(N1B-M1-N2+M2+1)
5000 *CHECK
5010  IF AREA<.95# GOTO *REPEAT
5020  PRINT N1S9,SPRO9,AREAS
5030  PRINT N1B9,BPRO9,AREAB
5040  PRINT AREA
5050 END
9000 *COMBI
9010  C1=1#
9020  FOR I=0 TO M9-1
9030     C1=C1*(N9-I)/(M9-I)
9040  NEXT I
9050  RETURN
```