Bull. Jpn. Soc. Fish. Oceanogr. 80(2) 115–123, 2016

A quick method for estimating egg density in sample bottle from fishery stock survey

Masatoshi HASEGAWA[1†], Toshihide KITAKADO[2], Shigehide IWATA[2] and Eiji TANAKA[2]

A rapid assessment of surveys of egg production of a fishery stock, such as the Japanese anchovies *Engraulis japonicus* or Japanese sardines *Sardinops melanostictus*, is one of the information necessary for predicting of the daily catch in the local whitebait fishery. We proposed a sequential combined method that consists the method of Most Probable Number (MPN) and a direct-count in order to reduce the time required to obtain unbiased estimate of the egg densities of the samples. In the method, the mean of egg density is estimated by the MPN if the density is low, but is done by a combination of the MPN and the direct-count if it is high. Effectiveness of the method was evaluated by simulation using the survey data of the Japanese anchovy in Suruga Bay. The results showed that the method provides good estimates of the true mean saving the time. We discussed about the application of this method.

Key words: egg production survey, most probable number, direct-count, maximum likelihood estimation, quick assessment

Introduction

In Japan, surveys of egg production in a fishery stock, such as the Japanese anchovies *Engraulis japonicus* or Japanese sardines *Sardinops melanostictus*, are carried out by towing a plankton net at various sampling points; this is done in order to estimate the abundance of spawning stock (Kuroda, 1991; Oozeki, 2010; Watanabe et al., 1996; Zenitani and Kimura, 1997), to forecast the number of recruits (Kuroda, 1991), and to predict the catch for the whitebait fishery (Yasue et al., 2006). To predict the catch amount, a rapid assessment of egg production is one of the necessary information as well as the environmental factor because daily information on the abundance of whitebait is used to make a decision of operation in fishing ground and to prepare the amount of man power for processing the landed whitebait. However, a "direct-count" of the whole number of eggs and larvae (below paragraphs, referred as "eggs") in raw sample consumes a great deal of time, and it is essential for quick estimation of the mean of the egg density to develop a rapid measurement of the egg densities in many samples.

One way to quickly estimate the mean of egg density is to apply the method of the Most Probable Number (MPN) (McCrady, 1915; Oblinger and Koburger, 1975; Salama et al., 1978), because this method requires the only information on the presence or absence of eggs in each sub sample that is a part of raw sample, and so is faster than performing a direct-count. In the MPN, the frequency of number of eggs in a sample is expressed by the Poisson distribution with mean, say μ, and the value of μ is estimated by the Maximum Likelihood Estimation using the binomial distribution with the probability of the presence $(1-e^{-\mu})$.

However, the MPN has the two problems. First, the inverse transformation of expectation of a variable transformation of random variable is not necessarily consistent with the expectation of the random variable. In the MPN, the parameter μ is transformed ($f(\mu)=e^{-\mu}$) from the assumption of the Poisson distribution and the likelihood is derived using the binominal distribution. Since the parameter μ is estimated by inverse transformation to the end, then the same problem occurs. The reason is because the variance that remains in the second-order terms in the Taylor expansion at $x=\mu$ of $f(x)$ is the bias. In the MPN, because size of sample is small, the bias becomes remarkable (Haas, 1989). In order to correct for this bias, Salama et al. (1978) proposed an effective bias-corrected expression, derived from the Taylor expansion of the logarithm of the likelihood function. After the proposal of Salama et al. (1978), Salama correction expression have been adopted in the MPN table by Haas (1989), and was incorporated into the computer program of the MPN by Klee (1993). The other study of bias correction has the following cases. In the case of a larger variance than the Poisson distribution, Haas (1989) considered the application of the negative binomial distribution and the ED_{50} estimation method. To reduce bias, Beliaeff and Mary (1993) have applied the binomial distribution, instead of the Poisson distribution to the distribution of the eggs. Garthright (1993) corrected the bias in logarithm of MPN estimate by using the idea of Salama et al. (1978).

Received June 11, 2015; Accepted January 20, 2016
[1] Shizuoka Prefectural Institute of Fishery Izu Branch, 251–1 Shirahama, Shimoda, Shizuoka 415–0012, Japan
[2] Tokyo University of Marine Science and Technology, 4–5–7 Konan, Minato, Tokyo 108–8477, Japan
[†] E-mail: masatoshi1_hasegawa@pref.shizuoka.lg.jp

Second, the MPN will fail to obtain a significant estimate if the sample densities are so high; for example, if all of the samples include one or more eggs, then the expectation of the density is infinite. In such high densities, therefore, it is not appropriate to depend directly on the MPN, and some modification is necessary.

In this paper, we propose a sequential method that combines the MPN with a direct-count. In the method, the mean of egg density is estimated by the MPN if the density is low, but is done by a combination of the MPN and the direct-count if it is high. The bias is corrected by the Salama expression (Salama et al., 1978) for the MPN and the new expressions for the others. Effectiveness of the method was evaluated by simulation using survey data of the Japanese anchovy in Suruga Bay in order to examine performances of the method, such as the extent of bias reduction and time saved. We will also discussed applications and suggest further modifications.

Materials and Methods

The Method of the Most Probable Number

We consider estimation of egg density in a "raw sample" bottle filled on board. We made some "sub-samples" of different volume levels using the raw sample, and counted the number of eggs in the sub-samples. The volume levels consist of unit one, say v_1 (m*l*), and multiple units ($v_2<v_3<\cdots<v_m$, i is ascending order about v_i). For each level a certain sub-samples are made, for example three sample bottles for each level.

We expect that the number of eggs per unit volume (eggs/m*l*) in the raw sample is a random variable of the Poisson distribution with expectation μ. Let m and w_i ($w_1=v_1/v_1=1<w_2=v_2/v_1<\cdots<w_m=v_m/v_1$, i is ascending order about w_i) be the number of volume levels and the ratio of the volume in i(i=1, 2, $\cdots$, m)th volume level to the volume in the unit one sample (i=1), respectively. Let N_i and n_i be the number of observations in the ith volume level and the number of samples that contain one or more eggs in the ith volume level, respectively. Note that the value of w_i can be set to an arbitrary level. We showed the above-mentioned procedure in Fig. 1 as a schematic diagram.

The probability that a sub-sample includes one or more eggs is $1-e^{-w_i\mu}$ (Salama et al., 1978), and n_i follows the binomial distribution. The likelihood to be maximized is

$$L\left(\mu \middle| N,n,w\right)=\prod_{i=1}^{m} {}_{N_i}\mathrm{C}_{n_i}\left(1-e^{-w_i\mu}\right)^{n_i} e^{-(N_i-n_i)w_i\mu} \tag{1}$$

And the Maximum Likelihood Estimate (MLE) of μ, denoted by $\hat{\mu}$, satisfies (Salama et al., 1978)

$$\sum_{i=1}^{m}\frac{w_i n_i}{1-e^{-w_i\hat{\mu}}}=\sum_{i=1}^{m}N_i w_i \tag{2}$$

The bias-corrected MLE (Seber, 1982; Haas, 1989), say $\hat{\mu}_c$, is

$$\hat{\mu}_c=\hat{\mu}-b \tag{3}$$

where the bias (Salama et al., 1978), say b, is

$$b=\frac{1}{2}\sum_{i=1}^{m}\left.\frac{\partial^2\hat{\mu}}{\partial\theta_i^{\,2}}\right|_{\theta_i=n_i} n_i\left(1-e^{-w_i\hat{\mu}}\right)e^{-w_i\hat{\mu}} \tag{4a}$$

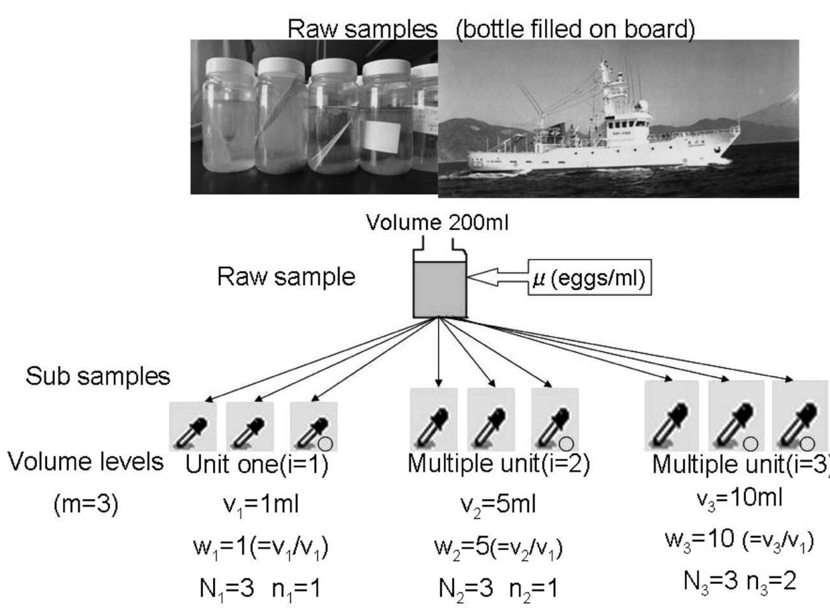

Figure 1. Schematic diagram of the sample example in the MPN in a case of m=3 and N_i=3. See text for notation. Open circle near the pipette expresses an egg.

where

$$\left.\frac{\partial^2\hat{\mu}}{\partial\theta_i^2}\right|_{\theta_i=n_i} = -\frac{F_i}{G^2}\times\frac{w_i^2}{\cosh\left(w_i\hat{\mu}\right)-1} + \frac{1}{2}\frac{F_i^2}{G^3}\times\sum_{j=1}^{m}\frac{n_j w_j^2\sinh\left(w_j\hat{\mu}\right)}{\left(\cosh\left(w_j\hat{\mu}\right)-1\right)^2} \tag{4b}$$

$$F_i = \frac{w_i}{1-e^{-w_i\hat{\mu}}} \tag{4c}$$

$$G = \sum_{j=1}^{m}\frac{n_j w_j^2}{2\left(\cosh\left(w_j\hat{\mu}\right)-1\right)} \tag{4d}$$

Value of $\hat{\mu}$ is numerically computed by such optimization methods as the Newton one.

Sequential Combined Method

The flowchart for our Sequential Combined Method (SCM) is presented in Fig. 2, as an example for m=3 and N_i=3(i=1, 2, 3), that is adopted in the paper.

In the flowchart, either (state of presence/absence of egg) or (number of eggs) in the all sub sample are measured in the order from the minimum volume level to the maximum. To reduce the bias in estimate and save the time consumption, for the first two sub samples from the same volume (state of presence/absence) are measured. If the both are in (presence), (number of eggs) in the third sample are measured. If the third contain egg and more, the rest of sequential sub samples (number of eggs) are measured because densities in the sub samples will be high. When the states of the first two sub samples are different, for the third sample (state of presence/absence) is observed. In the next volume the same procedure is applied (See Fig. 2 for details).

As the results there are the four cases of measurement combination in the Fig. 2. Whereas the Case 4 uses the MPN for the all sub samples, in Case 1 the direct-count is applied to seven in the nine samples (excepting the first two samples). For Cases 1–4 the data sets of (state of presence/absence) or (number of eggs) is tabulated in Table 1.

Based on the data sets for Cases 1–4, the following likelihood to be maximized

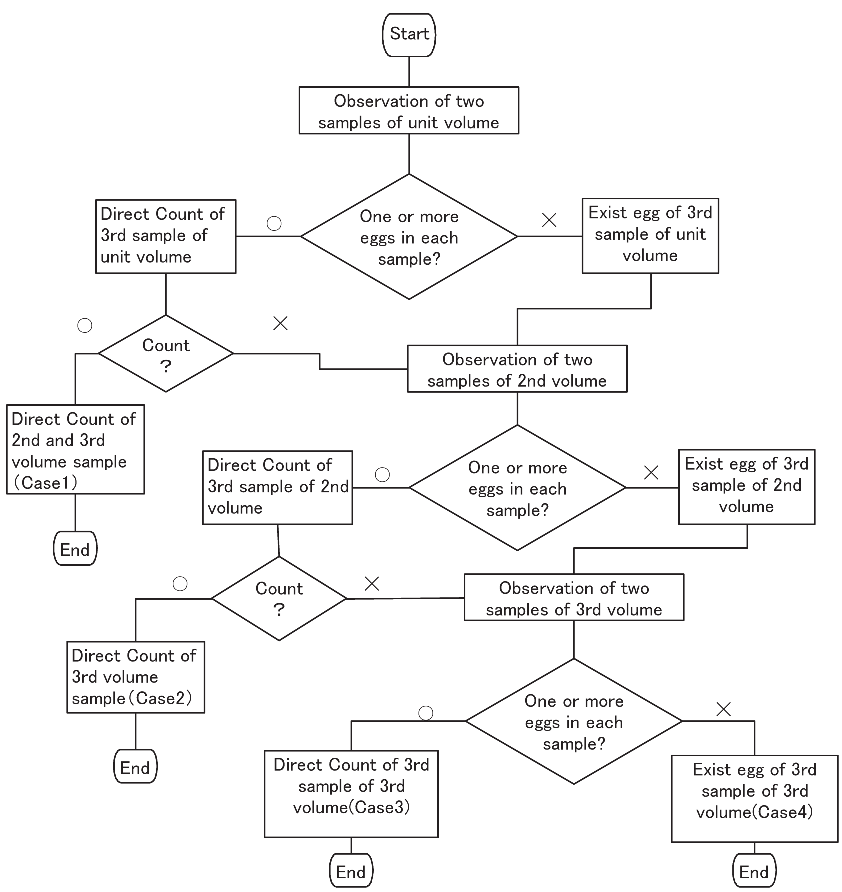

Figure 2. Flowchart of a Sequential Combined Method in a case of m=3 and N_i=3. Push forward a process from a start terminal towards a bottom. Rectangle shows the process. By the judgment of the rhombus, you choose in yes (○) or no (×). You finally choose case 1–4.

Table 1. The data sets of (state of presence/absence) or (number of eggs) for each cases.

Case	Unit volume			2nd volume			3rd volume		
	Sample1	2	3	1	2	3	1	2	3
1	P	P	C	C	C	C	C	C	C
	n_1		$x_{1,3}$	$x_{2,1}$	$x_{2,2}$	$x_{2,3}$	$x_{3,1}$	$x_{3,2}$	$x_{3,3}$
2	P/A	P/A	P/A	P	P	C	C	C	C
		n_1		n_2		$x_{2,3}$	$x_{3,1}$	$x_{3,2}$	$x_{3,3}$
3	P/A	P/A	P/A	P/A	P/A	P/A	P	P	C
		n_1			n_2		n_3		$x_{3,3}$
4	P/A	P/A	P/A	P/A	P/A	P/A	P/A	P/A	P/A
		n_1			n_2			n_3	

P/A: state of presence/absence.
P: state of presence.
C: number of eggs.
n_i: the number of samples that contain one or more eggs.
$x_{i,j}$: the number of eggs counted in j-th sample of i-th volume level.

$$L\left(\mu\middle|N,n,w,x\right) = \begin{cases} B\left(2\middle|2,w_1\mu\right)\times P\left(x_{1.3}\middle|w_1\mu\right) \times \prod_{i=2}^{3}\prod_{j=1}^{3} P\left(x_{i,j}\middle|w_i\mu\right) & \text{(Case1) (5a)} \\ B\left(n_1\middle|N_1,w_1\mu\right)\times B\left(2\middle|2,w_2\mu\right) \times P\left(x_{2,3}\middle|w_3\mu\right)\times \prod_{j=1}^{3} P\left(x_{3,j}\middle|w_3\mu\right) & \text{(Case2) (5b)} \\ \prod_{i=1}^{2} B\left(n_i\middle|N_i,w_i\mu\right)\times B\left(2\middle|2,w_3\mu\right) \times P\left(x_{3,3}\middle|w_3\mu\right) & \text{(Case3) (5c)} \\ \prod_{i=1}^{3} B\left(n_i\middle|N_i,w_i\mu\right) & \text{(Case4) (5d)} \end{cases}$$

$$\mathrm{B}(n|N,\mu)={}_N\mathrm{C}_n(1-\mathrm{e}^{-\mu})^n e^{-(N-n)\mu} \quad (6)$$

$$\mathrm{P}\left(x\middle|\mu\right)=\frac{\mu^x}{x!}e^{-\mu} \quad (7)$$

where $x_{i,j}$ is the number of eggs counted in j-th sample of i-th volume level.

The MLE of μ, denoted by $\hat{\mu}$, satisfies (See expression (2) for Case 4)

$$\frac{w_1 n_1}{1-e^{-w_1\hat{\mu}}}+\frac{x_{1,3}+\sum_{i=2}^{3}\sum_{j=1}^{3}x_{i,j}}{\hat{\mu}}=3\sum_{i=1}^{3}w_i \quad \text{(Case1) (8a)}$$

$$\sum_{i=1}^{2}\left(\frac{w_i n_i}{1-e^{-w_i\hat{\mu}}}\right)+\frac{x_{2,3}+\sum_{i=1}^{3}x_{3,j}}{\hat{\mu}}=3\sum_{i=1}^{3}w_i \quad \text{(Case2) (8b)}$$

$$\sum_{i=1}^{3}\left(\frac{w_i n_i}{1-e^{-w_i\hat{\mu}}}\right)+\frac{x_{3,3}}{\hat{\mu}}=3\sum_{i=1}^{3}w_i \quad \text{(Case3) (8c)}$$

The expressions for the bias-corrected estimate for each case are presented in Table 2.

Simulation for survey of Japanese anchovy egg production in Suruga Bay

We evaluated estimates from several methods by conducting a simulation (using the R language). The bias-corrected SCM (bc-SCM) was compared with the MPN, the MPN with the Salama correction, a direct-count, Case 3, and the SCM without the bias correction.

In the anchovy eggs production survey in Suruga Bay (Washiyama et al., 2014), a 200 m*l* bottle was usually adopted and the whole number of eggs was counted in the past surveys. The most of the mean density (98% of the whole) ranged from 0 (eggs/m*l*) to 1.0. We change a parameter μ as follows (μ=0.01 and 0.05–1.00 at intervals of 0.05).

The performance statistics for comparing efficiency are 1) point estimate, 2) mean square error (MSE) of the estimate, and 3) the mean of observation volume for an index of time consumption. The procedure of the simulation of the SCM is summarized in the below paragraphs, as an example of m=3 and N_i=3.

One set of computation consists of the following four steps:

(1) Generation of data of the number of eggs

Using the value of μ, artificial data of the nine Poisson random number ($x_{i,j}$) are generated.

(2) Data conversion

In each volume level, the number of samples with one or more eggs is counted.

Table 2. The expressions for the bias-corrected estimate for each case.

$$b=\frac{1}{2}\sum_{i=1}^{k}\left.\frac{\partial^2\hat{\mu}}{\partial\theta_i^2}\right|_{\theta_i=n_i,x_{i,j}} V\left(\hat{n}_i\ or\ \hat{x}_{i,j}\right)$$

$V(\hat{n}_i)$: the variance of the binominal distribution (the MPN)
$V(\hat{x}_{i,j})$: the variance of the Poisson distribution (the sirect-count)

Case		F_i	G	$\left.\frac{\partial^2\hat{\mu}}{\partial\theta_i^2}\right\|_{\theta=n_i,x_{i,j}}$
Case1	i=1	$\frac{w_i}{1-e^{-w_i\hat{\mu}}}$	$\frac{n_1w_1^2}{2\,(\cosh\,(w_1\hat{\mu})-1)}+\frac{n_{1,3}+\sum_{i=2}^{3}\sum_{j=1}^{3}x_{i,j}}{\hat{\mu}^2}$	$-\frac{F_i}{G^2}\times\frac{w_i^2}{\cosh(w_i\hat{\mu})-1}+\frac{1}{2}\frac{F_i^2}{G^3}\times\frac{n_1w_i^2\sinh(w_1\hat{\mu})}{\left(\cosh(w_1\hat{\mu})-1\right)^2}+\frac{4\left(x_{1,3}+\sum_{i=2}^{3}\sum_{j=1}^{3}x_{i,j}\right)}{\hat{\mu}^3}$
	i=2–8	$\frac{1}{\hat{\mu}}$		$-\frac{F_i}{G^2}\times\frac{2}{\hat{\mu}^2}+\frac{1}{2}\frac{F_i^2}{G^3}\times\frac{n_1w_1^2\sinh(w_1\hat{\mu})}{\left(\cosh(w_1\hat{\mu})-1\right)^2}+\frac{4\left(x_{1,3}+\sum_{i=2}^{3}\sum_{j=1}^{3}x_{i,j}\right)}{\hat{\mu}^3}$
Case2	i=1,2	$\frac{w_i}{1-e^{-w_i\hat{\mu}}}$	$\sum_{j=1}^{2}\frac{n_jw_j^2}{2\,(\cosh\,(w_j\hat{\mu})-1)}+\frac{x_{2,3}+\sum_{j=1}^{3}x_{3,j}}{\hat{\mu}^2}$	$-\frac{F_i}{G^2}\times\frac{w_i^2}{\cosh(w_i\hat{\mu})-1}+\frac{1}{2}\frac{F_i^2}{G^3}\times\sum_{j=1}^{2}\frac{n_jw_j^2\sinh(w_j\hat{\mu})}{\left(\cosh(w_j\hat{\mu})-1\right)^2}+\frac{4\left(x_{2,3}+\sum_{j=1}^{3}x_{3,j}\right)}{\hat{\mu}^3}$
	i=3–6	$\frac{1}{\hat{\mu}}$		$-\frac{F_i}{G^2}\times\frac{2}{\hat{\mu}^2}+\frac{1}{2}\frac{F_i^2}{G^3}\times\sum_{j=1}^{2}\frac{n_jw_j^2\sinh(w_j\hat{\mu})}{\left(\cosh(w_j\hat{\mu})-1\right)^2}+\frac{4\left(x_{2,3}+\sum_{j=1}^{3}x_{3,j}\right)}{\hat{\mu}^3}$
Case3	i=1,2,3	$\frac{w_i}{1-e^{-w_i\hat{\mu}}}$	$\sum_{j=1}^{3}\frac{n_jw_j^2}{2\,(\cosh\,(w_j\hat{\mu})-1)}+\frac{x_{3,3}}{\hat{\mu}^2}$	$-\frac{F_i}{G^2}\times\frac{w_i^2}{\cosh(w_i\hat{\mu})-1}+\frac{1}{2}\frac{F_i^2}{G^3}\times\sum_{j=1}^{3}\frac{n_jw_j^2\sinh(w_j\hat{\mu})}{\left(\cosh(w_j\hat{\mu})-1\right)^2}+\frac{4x_{3,3}}{\hat{\mu}^3}$
	i=4	$\frac{1}{\hat{\mu}}$		$-\frac{F_i}{G^2}\times\frac{2}{\hat{\mu}^2}+\frac{1}{2}\frac{F_i^2}{G^3}\times\sum_{j=1}^{3}\frac{n_jw_j^2\sinh(w_j\hat{\mu})}{\left(\cosh(w_j\hat{\mu})-1\right)^2}+\frac{4x_{3,3}}{\hat{\mu}^3}$
Case4	i=1,2,3	$\frac{w_i}{1-e^{-w_i\hat{\mu}}}$	$\sum_{j=1}^{3}\frac{n_jw_j^2}{2\,(\cosh\,(w_j\hat{\mu})-1)}$	$-\frac{F_i}{G^2}\times\frac{w_i^2}{\cosh(w_i\hat{\mu})-1}+\frac{1}{2}\frac{F_i^2}{G^3}\times\sum_{j=1}^{3}\frac{n_jw_j^2\sinh(w_j\hat{\mu})}{\left(\cosh(w_j\hat{\mu})-1\right)^2}$

$$n_i = \sum_{j=1}^{3} \mathrm{sgn}\left(x_{i,j}\right) \tag{9}$$

Here sgn (x) is the sign function expressed by

$$\mathrm{sgn}\left(x\right) = \begin{cases} 1 & (x > 0) \\ 0 & (x = 0) \end{cases} \tag{10}$$

(3) Case identification
Using n_i, the Cases are identified due to the flowchart in Fig. 2:

Case 1 (n_i=3),
Case 2 ($n_1 \neq 3$ and n_2=3),
Case 3 ($n_1 \neq 3$, $n_2 \neq 3$ and n_3=3),
Case 4 (elsewhere)

(4) Estimation of μ
The value of $\hat{\mu}$ is estimated using expressions (2) and (8a)–(8c).
The set from (1) to (4) was iterated 100,000 times.

The other simulations were conducted to estimate the mean of the observation volume. Due to the mathematical relation between the Poisson distribution and the waiting time (Yamada and Kitada, 1999), the measured volume in observation of presence/absence, say $y_{i,j}$, was calculated by the following exponential distribution:

$$f(y_{i,j}) = \mu \exp\left(-\mu \times y_{i,j}\right) \quad (0 < y_{i,j} < \infty) \tag{11}$$

Generating nine exponential random number ($y_{i,j}$), the Cases were identified by the number of sample with one or more eggs, calculated by the following expression:

$$n_i = \sum_{j=1}^{3} \mathrm{sgn}\left(w_i - y_{i,j}\right) \tag{12}$$

The volume, say $V_{i,j}$, is expressed by

$$V_{i,j} = \begin{cases} y_{i,j} & \text{(presence/absence)} \quad (13a) \\ w_i v & \text{(direct-count)} \quad (13b) \end{cases}$$

The performance statistics were calculated from 100,000 estimates of μ and $V_{i,j}$, as follows:

$$\bar{\hat{\mu}} = \frac{1}{100,000} \sum_{k=1}^{100,000} \hat{\mu}_k \tag{14}$$

$$\mathrm{MSE} = \frac{1}{100,000} \sum_{k=1}^{100,000} \left(\hat{\mu}_k - \mu\right)^2 \tag{15}$$

$$\bar{V} = \frac{1}{100,000} \sum_{k=1}^{100,000} \sum_{i=1}^{3} \sum_{j=1}^{3} V_{i,j,k} \tag{16}$$

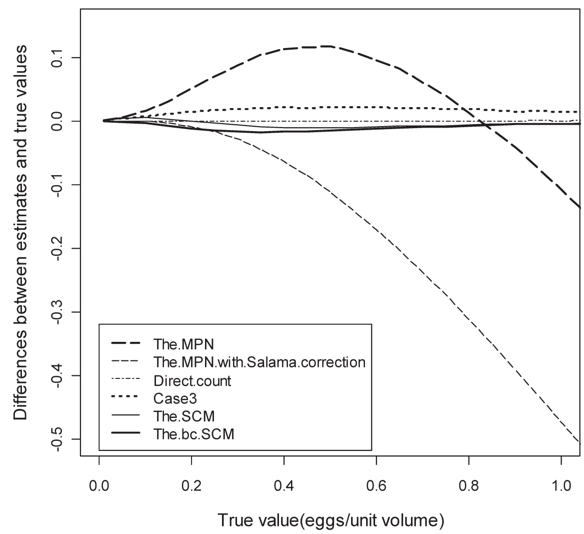

Figure 3. Relationship between the true value (μ) of egg density and the differences ($\bar{\hat{\mu}} - \mu$) between the estimate and true value of egg density by simulation. The method were the MPN, the MPN with the Salama correction, a direct-count, Case 3, the SCM, and the bc-SCM. The conditions of simulation are as follows. The number of simulations was 100,000. μ were 0.01 and 0.05–1.00 at interval of 0.05. The number of volume levels (m) was 3. The number of observations in each volume level (N_i) was 3. Volume of unit one was 1 ml (v_1). Volumes of Multiple units were 5 ml (v_2) and 10 ml (v_3).

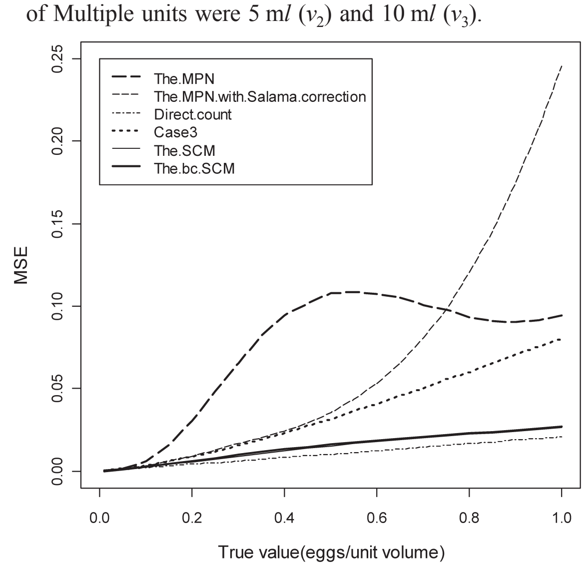

Figure 4. Relationship between the true value (μ) of egg density and the mean square error (MSE) by simulation. The methods were the MPN, the MPN with the Salama correction, a direct-count, Case 3, the SCM, and the bc-SCM. The conditions of simulation are as follows. The conditions of simulation are as follows. The number of simulations was 100,000. μ were 0.01 and 0.05–1.00 at interval of 0.05. The number of volume levels (m) was 3. The number of observations in each volume level (N_i) was 3. Volume of unit one was 1 ml (v_1). Volumes of Multiple units were 5 ml (v_2) and 10 ml (v_3).

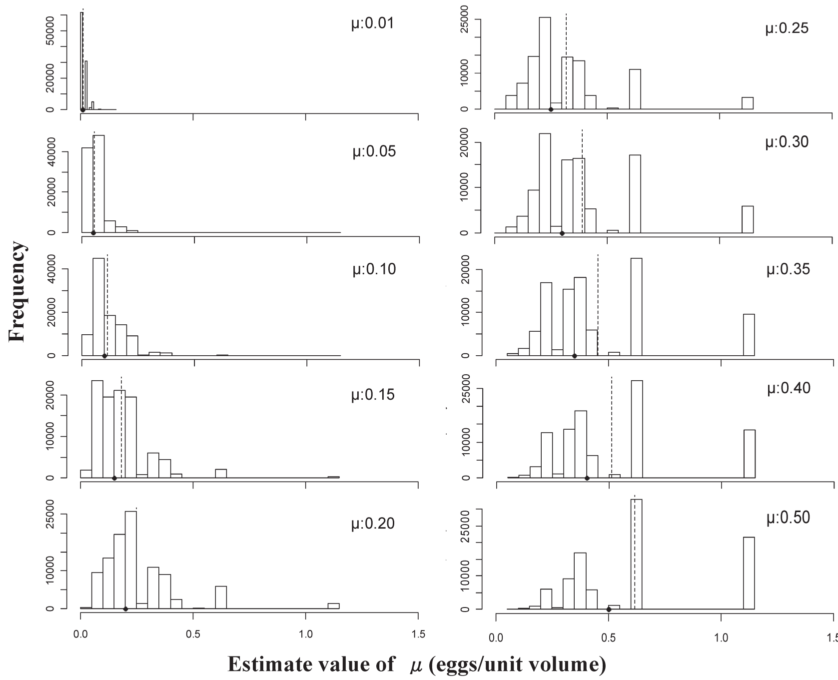

Figure 5. Histogram of the MPN estimate value. True values (μ) were 0.01 and 0.05–0.50 at intervals of 0.05. The frequency distribution of the estimate value of μ were provided as a result of 100,000 times of simulation against a truth value. Solid circles represent the true value. Dotted lines represent the average of estimate value of μ. The conditions of simulation are as follows. The number of volume levels (m) was 3. The number of observations in each volume level (N_i) was 3. Volume of unit one was 1 m*l* (v_1). Volumes of Multiple units were 5 m*l* (v_2) and 10 m*l* (v_3).

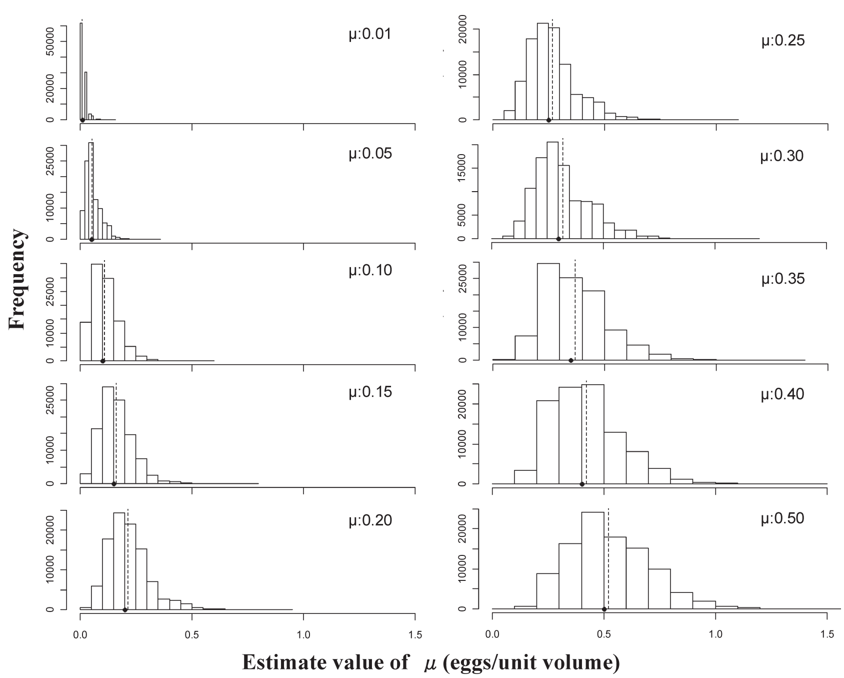

Figure 6. Histogram of the estimate value for Case3. True values (μ) were 0.01 and 0.05–0.50 at intervals of 0.05. The frequency distribution of the estimate value of μ were provided as a result of 100,000 times of simulation against a truth value. Solid circles represent the true value. Dotted lines represent the average of estimate value of μ. The conditions of simulation are as follows. The number of volume levels (m) was 3. The number of observations in each volume level (N_i) was 3. Volume of unit one was 1 m*l* (v_1). Volumes of Multiple units were 5 m*l* (v_2) and 10 m*l* (v_3).

Results

Simulation for survey of Japanese anchovy egg production in Suruga Bay

Fig. 3 compares $\bar{\hat{\mu}}-\mu$ with μ from the results of the simulations. The figure showed that 1) the curve of the MPN estimate against the true is upwardly convex and the absolute extent of positive bias at true=0.45 are 0.12, the negative bias at true 1.0 is 0.11, 2) the absolute extent of negative bias in the MPN with the Salama correction is over 40% at 1.0 although the extent is almost zero within the range that true is less than 0.25, 3) with case 3, the positive bias slightly exists, and 4) the extent of bias in the other methods is negligible.

Fig. 4 illustrates the mean square error for the methods. In the figure, 1) the MSE of the MPN is high to true=0.7, and in more than true=0.75 the MPN with the Salama correction is the highest, 2) the MSE of Case3 is high next to the MPN and the MPN with the Salama correction, and 3) the MSE in the other methods are low with an about the same value.

Fig. 5 and 6 show the frequency of the estimates of the MPN and Case 3 as an example in the SCM, respectively. Comparing the two figures, both shapes of the frequencies within $\mu \leq 0.2$ are similar to each other and seemed to be continuous. In $\mu \geq 0.2$, however, the distribution of the estimates in the MPN is discontinuous although the distribution in Case 3 is smooth one around μ.

Fig. 7 shows the expectation of the total observation volume for the MPN, the bc-SCM, and a direct-count. The averaged volume in the bc-SCM over $0.01 \leq \mu \leq 1$ was less than 80% of the volume in the direct-count and means that 20% and more of the expected time of observation is saved.

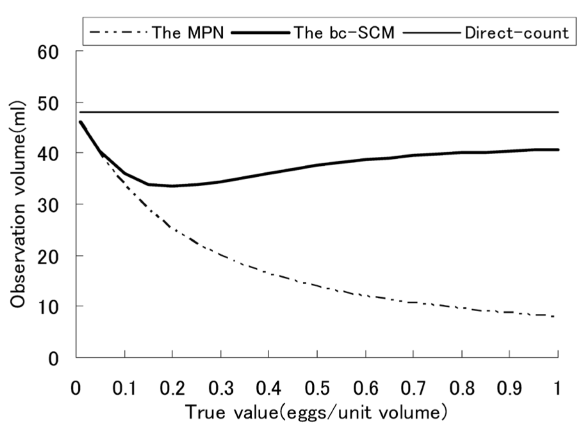

Figure 7. Relationship between the true value (μ) of egg density and the observation volume by simulation. Thick solid lines represent the bc-SCM. Solid lines represent a direct-count. Two-dot chain lines represent the MPN. The conditions of simulation are as follows. The number of simulations was 100,000. μ were 0.01 and 0.05–1.00 at interval of 0.05. The number of volume levels (m) was 3. The number of observations in each volume level (N_i) was 3. Volume of unit one was 1 ml (v_1). Volumes of Multiple units were 5 ml (v_2) and 10 ml (v_3).

Discussion

The combination of measurement (P/A or direct-count) in the SCM and bc-SCM was classified into Cases 1–4 (Fig. 2). Although some bias existed in the point estimate of the MLE in Cases 3 and 4 in the results of simulations, the bias was corrected by the new and past expressions for bias correction. Therefore the bias in the bc-SCM will be negligible. Although Fig. 3 showed that the extent of bias in the SCM and bc-SCM is equal, bc-SCM should be used for reasons of the bias existing in Cases 3 and 4.

The bc-SCM is expected to quickly give the unbiased estimate. Fig. 3 showed good performance of expectation of the estimates from the bc-SCM. The volume necessary for the bc-SCM and the SCM was the only 17–24% of the total volume (200 ml) counted by a direct-count. The Shizuoka Prefectural Research Institute of Fisheries carries out egg production surveys for 26 stations each month. The time necessary for counting the eggs and larvae of the Japanese anchovy of the 26 samples by the bc-SCM is empirically less than nineteen hour-person and this will sufficiently enable us to realize an effective catch prediction of the whitebait fishery based on the quick assessment.

Although the simulations in the paper were based on the only information of the distribution of density of anchovy eggs in Suruga Bay, the results of the simulation can be applied to various densities by adjusting the size of the minimum volume of the unit sub-sample. For an example, if the expectation of egg density is greater than the range covered in the simulation, the results will be effective by reducing the size of the minimum volume to the level of the expectation covered by the simulations, or by diluting the sample.

In addition the simulations will suggest an appropriate size of the minimum volume of the sub-sample. Fig. 7 showed that the observation volume in the bc-SCM was small when the true value is in the range from 0.15 to 0.3 and therefore the minimum volume having the expectation of 0.15–0.3 is recommended as an appropriate volume.

The paper gives a general idea for the quick assessment. The simulation study was conducted for a candidate in the minimum sets of the sub-samples, where (m, N_i)=(3, 3) is adopted in the past studies of the MPN (Oblinger and Koburger, 1975). However, the case (m, N_i) can be set to an arbitrary level. Based on the given condition of survey scale and number of staffs, someone may develop more appropriate combination of the volume and the number of the sub-sample or the decision rule of shift from (P/A) to direct-count using the simulation studies.

In the actual survey the samples of egg density were collected from various sampling stations for estimating the average of egg density in the whole survey areas, and the efficient estimate may be obtained changing the (m, N_i) in station. In such cases the optimum (m, N_i) in station will be determined using the prior information on egg density.

The bc-SCM will be applied to density estimation of other fields, of which the conditions is the sequential measurement, the variable quadrat (ex. volume, quadrat) in the measurement and the countable number.

References

Beliaeff, B. and J. Y. Mary (1993) The "most probable number" estimate and its confidence limits. Wat. Res., **27**, 799–805.

Garthright, W. E. (1993) Bias in the logarithm of microbial density estimates from serial dilutions. Biom. J., **35**, 299–314.

Haas, C. N. (1989) Estimation of microbial densities from dilution count experiments. Appl. Environ. Microbio., **55**, 1934–1942.

Klee, A. J. (1993) A computer program for the determination of most probable number and its confidence limits. J. Microbiol. Method, **18**, 91–98.

Kuroda, K. (1991) Studies on the recruitment process focusing on the early life history of the Japanese sardine, *Sardinops melanostictus* (Schlegel). Bull. Natl. Res. Inst. Fish. Sci., **3**, 25–278 (in Japanese with English abstract).

McCrady, M. H. (1915) The numerical interpretation of fermentation-tube results. J. Infec. Dis., **17**, 183–212.

Oblinger, J. L. and J. A. Koburger (1975) Understanding and teaching the most probable number technique. J. Milk Food Technol., **38**, 540–545.

Oozeki, Y. (2010) Temporal and spatial variation pattern of reproductive traits of small pelagic fish. Bull. Jpn. Soc. Fish. Oceanogr., **74**, 94–98 (in Japanese with English abstract).

Salama, I. A., G. G. Koch and H. D. Tolley (1978) On the estimation of the most probable number in a serial dilution technique. Commun. Stat. Theor. Methodol., A7: 1267–1281.

Seber, G. A. F. (1982) Some statistical methods. In: The estimation of animal abundance and related parameters second edition, The Blackburn Press, Caldwell, New Jersey, USA, 1.3.1–1.3.8.

Washiyama, H., M. Hasegawa and H. Hiyoshi (2014) Research on egg/larvae of important fishing species and plankton. Shizuoka Prefectural Institute of Fishery Annual Report 2012, 26–27 (in Japanese).

Watanabe, Y., H. Zenitani and R. Kimura (1996) Offshore expansion of spawning of the Japanese sardine, *Sardinops melanostictus*, and its implication for egg and larval survaival. Can. J. Fish. Aquqt. Sci., **53**, 55–61.

Yamada, S. and S. Kitada (1999) Chapter 3 Various Probability Distribution, Seibutu-shigen-toukeigaku (Biostatistics), Suisan-Kaiyou Library 4, Seizando-syoten Publishing, Tokyo, 28–52 (in Japanese).

Yasue N., R. Utsumi and Y. Takeda (2006) Relationship between the larval catch and egg production of Japanese anchovy, *Engraulis japonicus*, in the western inshore waters off Kii Peninsula. Bull. Jpn. Soc. Fish. Oceanogr., **70**, 163–169 (in Japanese with English abstract).

Zenitani, H. and R. Kimura (1997) Increase in late winter egg production of the Japanese anchovy as related to recovery of the stock size along the Pacific coast of Japan. Nippon Suisan Gakkaishi, **63**, 665–671 (in Japanese with English abstract).

漁業資源卵稚仔調査サンプルの卵稚仔数推定のための迅速法

長谷川雅俊[1†]・北門利英[2]・岩田繁英[2]・田中栄次[2]

カタクチイワシ資源等の卵稚仔量の迅速な評価は，ある地域のシラス漁業の漁況予測のために必要な情報の一つである．迅速な評価のために卵稚仔サンプルの測定時間を短縮し，かつ平均密度の推定値の偏りを小さくする方法として最確数法と直接計数を結合させた逐次結合法を開発した．この方法は密度が低いときは最確数法だが，測定が進むにつれ，高密度が明らかになったなら，偏りのない直接計数を採用し推定する方法である．この方法を駿河湾のカタクチイワシの調査データを使いシミュレーションで評価した．その結果，平均値の推定値は真値とよく一致し，推定時間も短くなることがわかった．この方法の利活用について議論した．

キーワード：卵稚仔調査，最確数法，直接計数，最尤推定，迅速評価

2015年6月11日受付，2016年1月20日受理

[1] 静岡県水産技術研究所伊豆分場，〒415–0012　静岡県下田市白浜251–1

[2] 東京海洋大学，〒108–8477　東京都港区港南4–5–7

[†] E-mail: masatoshi1_hasegawa@pref.shizuoka.lg.jp