

## 栽培漁業を事例とした累積データによる相関分析の問題点

北田 修一\*<sup>1</sup>・松宮 義晴\*<sup>2</sup>

\*<sup>1</sup>日本栽培漁業協会, \*<sup>2</sup>三重大学生物資源学部

### Influence of Data Cumulation on Correlation applied in Sea Ranching Programs

Shuichi KITADA\*<sup>1</sup> and Yoshiharu MATSUMIYA\*<sup>2</sup>

\*<sup>1</sup>Japan Sea-Farming Association

\*<sup>2</sup>Faculty of Bioresources, Mie University

#### Abstract

Examination was made of the effect of cumulated dependent and independent variables on correlation coefficients. Correlations between annual number of hatchery-release and annual catch in weight, and prediction of the day when crab larvae metamorphose to crab-stage fingerlings were determined. Through simulation studies on cumulating variables, the distribution patterns of sample correlation coefficients were clarified and precision of prediction was assessed.

**Key words:** data cumulation・correlation coefficient・regression analysis・sea ranching

種苗放流効果を定量的に評価するためには、基本的には市場の水揚げの標本調査によって回収率や回収量を推定することが望ましい。2段サンプリングの場合の推定方法が提案され、誤差や調査計画についても検討されている<sup>1)</sup>。この調査では、放流魚は何らかの標識によって天然魚と識別できることが要求される。放流尾数とそれに対応する漁獲量の相関分析は、放流効果の簡便な把握方法としてよく用いられている。一般に寿命の長い栽培対象種の放流効果は数年にわたって発現するため、単純にある年の放流尾数と漁獲量の関係を検討することは適当ではない。このため放流尾数や漁獲量のデータを累積して両者の相関を検討し、得られた回帰直線によって放流効果を試算している<sup>2,3)</sup>。種苗生産においても飼育水温の累積データによる相関分析に基づいて、ガザミ幼生

の稚ガニへの変態時期が予測されている<sup>4)</sup>。

変数  $X$  に関する情報に基づいて変数  $Y$  の値を予測したい場合、説明変数  $X$  や目的変数  $Y$  の累積が相関分析に及ぼす影響については十分検討されていない。本論文の目的は累積データに基づく相関分析の統計学的な問題点について、栽培漁業における3タイプのデータを事例として精査検討することである。

#### 資料および方法

説明変数  $X_i$  と目的変数  $Y_i$  の両方を累積する事例1<sup>2)</sup>として、アワビの年々の累積放流個数  $X$  と累積漁獲量  $Y$  をとりあげる。 $X_i$  のみを累積する事例2<sup>3)</sup>として、マガイの5年間の累積放流尾数  $X$  とその翌年すなわち6年目の年間漁獲量  $Y$  を検討する。さらに  $X_i$  が  $Y_i$  の一部として含まれる事例3<sup>4)</sup>として、ガザミ種苗生産水

槽内での飼育積算水温  $X$  と稚ガニ変態までの総積算水温  $Y$  を吟味する。

計算機シミュレーションによるブートストラップ的方法に従って、累積データの相関分析に及ぼす影響を検討した。計算機から2組の一様乱数を  $J$  個 ( $j=1, \dots, J$ ) 取り出し  $A_j, B_j$  とする。事例1では  $X_i = \sum_{j=1}^J A_j$ ,  $Y_i = \sum_{j=1}^J B_j$ , 事例2では  $X_i = \sum_{j=1}^J A_j$ ,  $Y_i = B_{i+5}$  である。ここで、データ数は  $n(i=1, \dots, n)$  であり、 $X_i$  は放流尾数、 $Y_i$  は漁獲重量に対応する。事例3では  $X_i = \sum_{j=1}^k A_j$ ,  $Y_i = \sum_{j=1}^k A_j$  であり、 $A_j$  はガザミ幼生がふ化後  $c+1$  日目に稚ガニに変態すると考えた場合の毎日の計測水温に対応する。 $X_i$  は  $k$  日までの積算水温であり、 $Y_i$  は  $c$  日まですなわち稚ガニ変態までの総積算水温である。ここでは  $k=3, 5, 7, 10$  と  $c=13$  の場合について検討する。

母相関係数  $\rho=0$  の下での標本相関係数  $r$  の確率分布は、 $r=0$  のまわりに対称形をなしており、 $\rho$  が1に近づくにつれて左に歪んだものとなる<sup>5)</sup>。 $\rho=0$  の母集団から  $X_i$  と  $Y_i$  をランダムに抽出して上記の事例で見られる累積操作を行えば、 $r$  の分布はデータを累積する影響に反応するはずである。一方、標本相関係数の確率密度関数<sup>5)</sup>は複雑な形をしており、累積データの影響を解析的に精査することは困難である。

$X_i$  と  $Y_i$  から何回も繰り返して  $r$  を計算しヒストグラムを描き、標本相関係数の密度関数を近似する。このような計算機シミュレーションによってデータの累積操作の影響を検討することが可能である。具体的には、事例1および事例2では  $0 \leq A_j \leq 100$ ,  $0 \leq B_j \leq 100$ ,  $n=10$ , 事例3では  $20 \leq A_j \leq 30$ ,  $n$  は実際の例<sup>4)</sup>に合わせて18とした。また、 $r$  の計算は2,000回実施した。生成する2組の乱数の数  $J$  は、事例1で  $J=n=10$ , 事例2で  $J=n+5=15$ , 事例3で  $J=c=13$  である。

## 結 果

事例1のシミュレーションにおける標本相関係数  $r$  の分布は極端に1に偏ったものとなる (Fig. 1 上)。全く関係のない母相関係数  $\rho=0$  の2つの現象でも、2変数のデータを累積することにより標本相関係数の分布は1に偏ったものとなる。このことは事例1のような解析方法によって放流効果を議論するのは適当でないことを

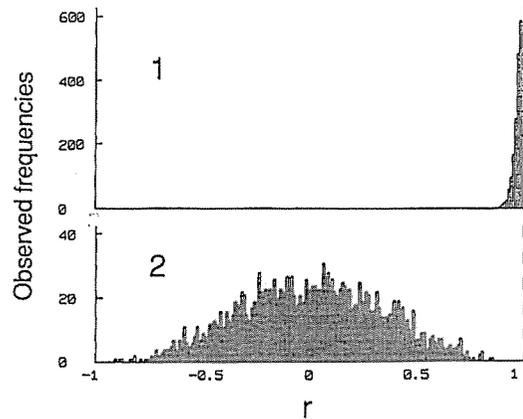


Fig. 1 Distributions of sample correlation coefficients obtained from 2,000 simulations.  $r$ : sample correlation coefficient, 1: Case 1, 2: Case 2.

意味している。

事例2のシミュレーションにおける標本相関係数  $r$  は  $\rho=0$  のまわりに対称形に分布しており (Fig. 1 下)、 $\rho=0$  からとられた分布と考えると良いであろう。このことは事例2のような累積操作は標本相関係数の分布に影響しないことを意味している。

事例3のシミュレーションでは飼育水温を積算する日数  $k$  が大きくなるにつれて、標本相関係数  $r$  も大きくなり左に歪んだ分布となる (Fig. 2)。 $r$  の期待値は  $k$  の増大につれて大きくなり、予測の精度が高くなることを示唆している。

## 考 察

事例1のシミュレーションについて標本相関係数  $r$  の分布が1に偏ったものになる理由は、累積した  $X_i$  と  $Y_i$  のそれぞれが互いに独立でないことによる。標本数  $n$  が大きくなるほど分散は小さくなるので、 $r$  の分布は益々1に偏っていくと想定される。

事例2の場合は  $X_i$  については互いに独立でないが、 $X_i$  と  $Y_i$  については独立であるため、もとの  $\rho=0$  のときの標本相関係数の分布が保存されている。従って、事例2の場合の累積操作は統計学的には問題ない。但し、種々の原因によりこの方法によって放流効果を抽出することは困難である。条件が整った場合を除いて、多くの

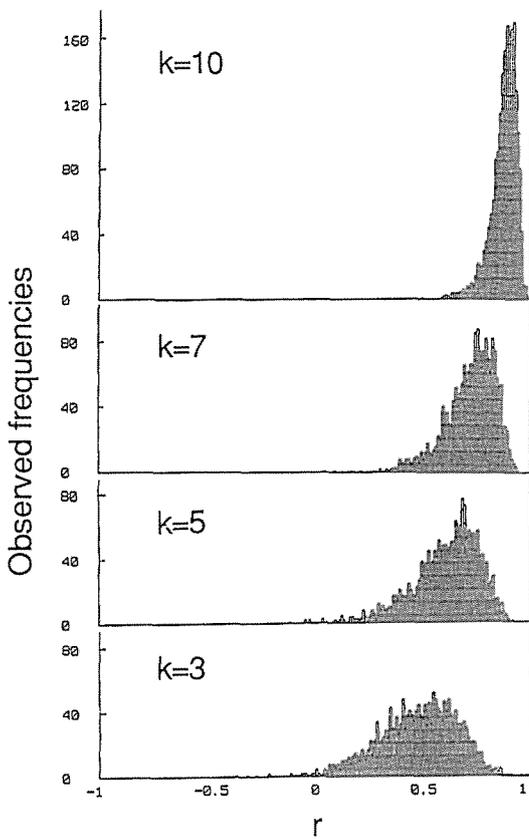


Fig. 2 Distributions of sample correlation coefficients for the Case 3 obtained from 2,000 simulations.  $r$ : sample correlation coefficient,  $k$ : the number of cumulative days

場合放流効果が漁獲変動の中に埋没してしまうことが最大の原因である。

標本数との関連を精査するため、事例2について同様の方法で標本数  $n=10$  のほか、 $n=20$  と  $50$  に基づくシミュレーションを実施した (Fig. 3)。久保・吉原<sup>6)</sup>は「放流尾数と漁獲量の相関係数が大きければ放流効果が大きいことを物語る」と述べたが、標本数  $n$  が小さい場合は実際に無相関であってもみかけ上大きな値の相関係数が得られる確率が高いことに留意する必要がある。

事例3については、 $X_i = \sum_{j=1}^i A_j$ ,  $Y_i = \sum_{j=1}^i A_j$  ( $k < 13$ ) であるから、積算する日数  $k \rightarrow 13$  のとき  $r \rightarrow 1$  と当然の結果となる。ここでは明らかに  $\rho > 0$  であるので、ガザミの例<sup>4)</sup>で行っているような  $\rho = 0$  の仮説検定をわ

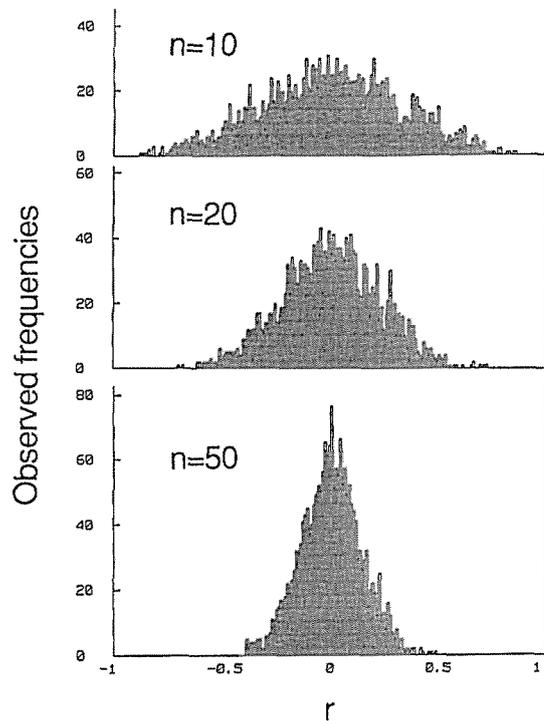


Fig. 3 Distributions of sample correlation coefficients for the Case 2 with different sample sizes obtained from 2,000 simulations.  $r$ : sample correlation coefficient,  $n$ : sample size

ざわざおこなう必要はない。ガザミの例<sup>4)</sup>では飼育日数の経過に伴って、稚ガニ変態までの総積算水温の4つの予測式を得たが、予測の精度については未検討である。

予測の精度を明かにするため、事例3と同様の方法で変数をつくり、回帰係数および回帰式による予測値の平均と分散を2,000回のシミュレーションから計算した。回帰係数の分散は  $k$  が大きくなるほど小さくなり、 $k=10$  では  $k=3$  の場合の約1/10となる (Table 1)。 $X_i$  の最大値および  $X_i$  の平均値に対する予測値の分散は、 $k=10$  では  $k=3$  の場合に比べてかなり小さい (Table 2)。予測の精度は平均値から離れるに従って悪くなるのが実際に確かめられた。

$k$  が大きくなると予測の精度も良くなるので、稚ガニに近い段階の予測が飼育初期段階に比べて精度において優れている。事例3では  $20 \leq A_j \leq 30$  すなわち毎日の飼育水温を  $20^\circ\text{C}$  から  $30^\circ\text{C}$  の間の乱数で与えたが、実際の飼育水温の変動はもっと小さいので予測の精度はさら

**Table 1.** Means and variances of the partial regression coefficients obtained from 2,000 simulations for the Case 3

$k$	Mean	Variance
3	0.9915	0.1984
5	1.0019	0.0962
7	0.9916	0.0593
10	0.9957	0.0189

$k$ : the number of cumulative days

**Table 2.** Means and variances of the regressed values to the maximum and the mean values of  $X$ , obtained from 2,000 simulations for the Case 3

$k$	Maximum value of $X$		Mean value of $X$	
	Estimate	Variance	Estimate	variance
3	329.93	35.14	317.02	8.14
5	333.48	8.64	318.47	3.73

$k$ : the number of cumulative days

に良くなることが予想される。ここでは飼育の途中での累積積算水温と稚ガニに変態する時点での総積算水温の関係について検討したが、飼育水温はコントロールされており変動は小さいことを考慮すると、幼生収容時の飼育水温と稚ガニに変態するまでの日数の関係を検討しておけば十分実用に耐え得る予測が可能である。

## 要 約

本論文では、栽培漁業を事例とした相関分析に関し、説明変数  $X$  と目的変数  $Y$  の累積の影響を計算機シミュ

レーションにより統計学的に検討した。 $X$  を放流尾数、 $Y$  をそれに対応する漁獲量とした相関分析では、 $X$  だけの累積は問題ないが  $Y$  も同時に累積した場合、 $X$  と  $Y$  が無関係であっても標本数が大きくなるほど相関係数が 1 に近づいていく。このような方法による放流効果の評価は適当ではない。また、 $X$  を累積飼育水温、 $Y$  をガザミが稚ガニに変態するまでの総積算水温とした分析では、累積する量が増えるほど予測の精度が良くなる。

## 引用文献

- 1) Kitada, S., Taga, Y., and Kishino, H.: Effectiveness of a stock enhancement program evaluated by a two-stage sampling survey of commercial landings. *Can. J. Fish. Aquat. Sci.*, 49: 1573-1582 (1992).
- 2) Saito, K.: Studies on propagation of Ezo abalone, *Haliotis discus hannai* Ino-I, Analysis of the relationship between transplantation and catch in Funka Bay coast. *Nippon Suisan Gakkaishi*, 45: 695-704 (1979).
- 3) 高間 浩：神奈川県沿岸における放流の成果と問題点。マダイの資源培養技術（田中 克・松宮義晴編），恒星社厚生閣，東京，pp. 127-143 (1986)。
- 4) 金丸彦一郎：有効積算水温を用いたガザミ種苗生産における稚ガニへの変態時期の推定。佐賀県栽培センター研報，1: 83-88 (1987)。
- 5) 芝 祐順：行動科学における相関分析法（第2版），東大出版会，東京，pp. 202-204 (1975)。
- 6) 久保伊津男・吉原友吉：水産資源学（改訂版），共立出版，東京，pp. 357-367 (1969)。
- 7) 浜崎活幸・福永恭平・丸山敬悟：ガザミのふ化までの所要日数と飼育水温の関係およびふ化日の同調化について。栽培技研：20, 73-85 (1992)。