捕かく一再捕かく法の一モデル

——捕かくの確率が前回の捕かく—非捕かくにより異ってくる場合——

高　橋　宏　一

(1961 年 5 月受付)

One Model for Estimating the Population Size by Capture-Recapture Method

Koiti TAKAHASI

In our model for the estimation, using the multiple-recapture method, of the total size of a population, we assume that the probability with which each individual in population is canght at the t th stage of sampling depends on whether it was canght or not in the $t-1$ th stage of sampling.

The estimate is obtained as a root of equation (6″) with respect to N.

Furthermore some asymptotic proporties of our estimate in the simplest case are shown.
Institute of Statistical Mathematics

§1. 序

一般の母集団の総数推定のための捕かく再捕かく法をもちいるモデルでは，母集団の各個体は，捕かくに対し同一確率をもつとされる．しかし，Leslie, Chitty が [1] で示唆しているように，以前における捕かく歴が，以後の捕かくの確率に影響をもつ場合があり得る．こうした場合を，もつとも一般的に取り扱おうとすれば，以前の捕かく歴全体に次の捕かくの確率が依存するとして考察することになるが，逆に簡単な場合としては，すぐ前で捕まったか否かということのみに，次の捕かくの確率が依存するというものが考えられる．後者でも，実際上起り得るものと思われる．例えば，捕かくの間隔が短かすぎ，わなが前の近くにかけられたような場合とか，[2] の例で，売春婦や麻薬中毒患者の検挙で前に検挙された人に特に注意が向けられているといった場合（この場合は以前の検挙歴全体を考慮した方がよいかもしれないが）などが考えられる．ここでは，この簡単な場合の母集団の全数推定の捕かく—再捕かく法をもちいるモデルを作り，最大尤度法的推定量を求める．

§2. 母集団の総数推定のモデルと推定量

母集団の総数を N とする．最初に各個体は同一確率 p_1 で捕かくされるとする．捕かくされたものは“1”なる記しを付けられ母集団中に戻される．次に前に捕かくされなかった各個体は p_2 の確率で，捕かくされたものは p_2' の確率で捕かくされる．ここで捕かくされたものは“2”なる記しを付けて母集団に戻される．以下同様に t 回目では，$t-1$ 回目で捕まったものは，p_t' の確率で，捕まっていないものは p_t の確率で捕かくされる．そして，“t”なる記しを付けられ母集団に戻される．こうした実験により我々の得る資料は，t_1, t_2, …, t_s 回目でのみ捕まったものの個体数，$u_{t_1\cdots t_s}$ と記す，である．

次に，この資料からの母集団総数の推定量を作らんとする．方法は [3] と全く類似である．

捕かくの回数を T 回としておく．資料 $\{u_\omega ; \omega$ は 1, 2, …, T の部分集合$\}$ が得られたとき，N の尤度は，これを $e^{L(N)}$ とおくと，

$$(1)\qquad e^{L(N)}=\frac{N!}{(N-r)!\prod\limits_{\omega} u_\omega !}\cdot(\prod_{\omega} P_\omega{}^{u_\omega})\cdot Q^{N-r}$$

(ここで r は少くも一度捕まっているものの個数，P_ω は一個体が ω (空集合でない) でのみ捕まる確率，Q は一度も捕まらない確率である.) 例えば，

$$P_{256}=(1-p_1)p_2(1-p_3{}')(1-p_4)p_5p_6{}'(1-p_7{}')(1-p_8)\cdots(1-p_T)$$

である．(1) の値を最大にするような N を求めるための尤度方程式は，(1) の対数をとり階差をつくって，0 とおくことにより得られる．すなわち，対数をとって

$$L(N)=\log N!-\log(N-r)!+(N-r)\log Q-\sum_{\omega}\log u_\omega !+\sum_{\omega} u_\omega \log P_\omega$$

階差をつくると

$$L(N)-L(N-1)=\log N-\log(N-r)+\log Q.$$

これを 0 に等しいと置いて

$$(2)\qquad \frac{N}{N-r}=\frac{1}{Q}$$

ところで

$$(3)\qquad Q=\prod_{t=1}^{T}(1-p_t)$$

また，a_t を t 回目でつかまったものの個数，$v_{t-1,t}$ は $t-1$ 回目では捕まらず t 回目で捕まっているものの個数，$w_{t-1,t}$ は T 回中少くも一度は捕まっていて，かつ $t-1$ 回目，t 回目の両方で捕まっていないものの個数とすると，p_t や $p_k{}'$ の定義と，この $v_{t-1,t}$, $w_{t-1,t}$, a_t などの意味を考え合わせれば，

$$(4)\qquad (\prod_{\omega} P_\omega{}^{u_\omega})\cdot Q^{N-r}=p_1{}^{a_1}(1-p_1)^{N-a_1}\prod_{t=2}^{T}\{p_t{}^{V_{t-1,t}}(1-p_t)^{W_{t-1,t}}(1-p_t)^{N-r}\}\cdot f(p_1{}',\cdots,p_T{}')$$

と書き表わせることがわかる．$f(p_1{}', \cdots, p_T{}')$ は $p_1, \cdots, p_T$, N には無関係な $p_1{}', \cdots, p_T{}'$ のみの函数である．そこで (1) を p_t $(t=1, 2, \cdots, T)$ の尤度と考えるとき，(1) の対数をとり，両辺を各 p_t で偏微分し，0 とおくことにより p_t $(t=1, 2, \cdots, T)$ についての尤度方程式を得る．すなわち (4) の対数をとると，

$$a_1\log p_1+(N-a_1)\log(1-p_1)+\sum_{t=2}^{T}\{v_{t-,t}\log p_t+w_{t-1,t}\log(1-p_t)+(N-r)\log(1-p_t)\}$$
$$+\log f(p_1{}', \cdots, p_T{}')$$

これを p_t で偏微分して，0 とおけば

$$\frac{a_1}{p_1}-\frac{N-a_1}{1-p_1}=0,\quad \frac{v_{t-1,t}}{p_t}-\frac{w_{t-1,t}}{1-p_t}-\frac{N-r}{1-p_t}=0\quad (t=2, \cdots, T)$$

すなわち，

$$p_1=\frac{a_1}{N},\quad p_t=\frac{v_{t-1,t}}{N-r+w_{t-1,t}+v_{t-1,t}}\quad (t=2, \cdots, T)$$

ところで，後者は，$r-v_{t-1,t}-w_{t-1,t}$ は少くも T 回中一度捕まったものから，それらの中で $t-1$ 回目と t 回目の両方で捕まらなかったものと，$t-1$ 回目で捕まらず t 回目で捕まったものとを引いたものだから，これは結局 $t-1$ 回目で捕まっているものの個数 a_{t-1} であるから，$p_t=v_{t-1,t}/N-a_{t-1}$ と書きかえられる．ここで便宜のため，$v_{0,1}=a_1$, $a_0=0$ とおけば，

$$(5)\qquad p_t=\frac{v_{t-1,t}}{N-a_{t-1}}\quad (t=1, 2, \cdots, T)$$

と，まとめて表わせる．

ここで (5) を (2) に代入すると，(Q は (3) によって $p_t(t=1, 2, \cdots, T)$ だけで表わされていることに注意して)

$$\frac{N}{N-r}=\frac{1}{\prod_{t=1}^{T}\left(1-\frac{v_{t-1,t}}{N-a_{t-1}}\right)}\quad \text{であるから}$$

(6)
$$N\prod_{t=1}^{T}(N-a_{t-1}-v_{t-1,t})=(N-r)\prod_{t=1}^{T}(N-a_{t-1})$$

$a_0=0$, $v_{0,1}=a_1$ であったことから，(6) は，

(6′)
$$\prod_{t=2}^{T}(N-a_{t-1}-v_{t-1,t})=(N-r)\prod_{t=3}^{T}(N-a_{t-1})$$

また，簡単のため $d_t=a_{t-1}+v_{t-1,t}$ $(t\geqq 2)$ とおけば (d_t の意味は $t-1$ 回目か t 回目の少くも一方で捕まったものの個数)，

(6″)
$$\prod_{t=2}^{T}(N-d_t)=(N-r)\prod_{t=2}^{T-1}(N-a_t).$$

$\varphi(N)=\prod_{t=2}^{T}(N-d_t)-(N-r)\prod_{t=2}^{T}(N-a_t)$ とおくと，$d_t<r(t=1, 2, \cdots, T)$ と仮定してかまわないから $\varphi(r)>0$，一方$\varphi(N)=(-\sum_{t=2}^{T}d_t+r+\sum_{t=2}^{T-1}a_t)N^{T-2}+O(N^{T-3})$ であり，かつ $-\sum_{t=2}^{T}d_t+r+\sum_{t=2}^{T-1}a_t=-\sum_{t=2}^{T}(a_{t-1}+v_{t-1,t})+r+\sum_{t=2}^{T-1}a_t=r-(v_{1,2}+\cdots+v_{T-1,T}+a_1+a_T)<0$ だから $\varphi(\infty)<0$. したがって母集団の総数の推定量として意味があるような r より大きい (6″) の根が少くも存在している．それを $\hat{N}$ と書き，求めている推定量とする．

§3. $T=3$ のときの $\hat{N}$ の漸近的様子

T が一般のときの考察は，式が複雑であり，かつ考え方は $T=3$ のときと同様であるから，$T=3$ に限って話をする．ここで漸近的という意味は $\{p_t\}$ $\{p_t'\}$ を固定して，N を十分大きくするということである．さて今の場合，(6″) は，$(N-d_2)(N-d_3)=(N-r)(N-a_2)$ であるから，

(7)
$$\hat{N}=\frac{d_2d_3-a_2r}{d_2+d_3-r-a_2}=r+\frac{u_1\cdot u_3}{u_{13}}.$$

$$\because\ d_2=u_1+u_2+u_{12}+u_{13}+u_{23}+u_{123},\ d_3=u_2+u_3+u_{12}+u_{13}+u_{23}+u_{123},$$
$$a_2=u_2+u_{12}+u_{23}+u_{123},\ r=u_1+u_2+u_3+u_{12}+u_{13}+u_{23}+u_{123}.$$

式を簡単に書く為，$r=X$, $u_1=Y$, $u_3=Z$, $u_{13}=W$ とおく．さて，$\{u_\omega\}$ が多項分布に従っていることから，$\left(P_r[\{u_\omega\}]=\frac{N!}{(N-r)!\prod_\omega u_\omega!}(\prod_\omega P_\omega^{u_\omega})\cdot Q^{N-r}\right)$, $E(X)=X_0$, $E(Y)=Y_0$, $E(Z)=Z_0$, $E(W)=W_0$ とおくと，

(8)
$$X_0=N(1-Q),\ Y_0=NP_1,\ Z_0=NP_3,\ W_0=NP_{13}.$$

ところで，

(9)
$$\begin{aligned}\hat{N}(X_0, Y_0, Z_0, W_0)&=N(1-Q)+N\cdot\frac{P_1\cdot P_3}{P_{13}}\\&=N\left\{1-Q+\frac{p_1(1-p_2')(1-p_3)\times(1-p_1)(1-p_2)p_3}{p_1(1-p_2')p_3}\right\}\\&=N\{1-Q+(1-p_1)(1-p_2)(1-p_3)\}=N\end{aligned}$$

簡単のため，(X_0, Y_0, Z_0, W_0) での函数の値は右下にインデックス 0 をつけて表わすことにする．(9) より，テーラーの定理によって

$$(10)\quad \hat{N}-N=(X-X_0)\left[\frac{\partial\hat{N}}{\partial X}\right]_0+(Y-Y_0)\left[\frac{\partial\hat{N}}{\omega Y}\right]_0+(Z-Z_0)\left[\frac{\partial\hat{N}}{\partial Z}\right]_0+(W-W_0)\left[\frac{\partial\hat{N}}{\partial W}\right]_0$$
$$+\frac{1}{2}(X-X_0)^2\left[\frac{\partial^2\hat{N}}{\partial X^2}\right]_\theta+\frac{1}{2}(Y-Y_0)^2\left[\frac{\partial^2\hat{N}}{\partial Y^2}\right]_\theta+\frac{1}{2}(Z-Z_0)^2\left[\frac{\partial^2\hat{N}}{\partial Z^2}\right]_\theta$$
$$+\frac{1}{2}(W-W_0)\left[\frac{\partial^2\hat{N}}{\partial W^2}\right]_\theta$$
$$+(X-X_0)(Y-Y_0)\left[\frac{\partial^2\hat{N}}{\partial X\partial Y}\right]_\theta+(X-X_0)(Z-Z_0)\left[\frac{\partial^2\hat{N}}{\partial X\partial Z}\right]_\theta$$
$$+(X-X_0)(W-W_0)\left[\frac{\partial^2\hat{N}}{\partial X\partial W}\right]_\theta$$
$$+(Y-Y_0)(Z-Z_0)\left[\frac{\partial^2\hat{N}}{\partial Y\partial Z}\right]_\theta+(Y-Y_0)(W-W_0)\left[\frac{\partial^2\hat{N}}{\partial Y\partial W}\right]_\theta$$
$$+(Z-Z_0)(W-W_0)\left[\frac{\partial^2\hat{N}}{\partial Z\partial W}\right]_\theta$$

ただし，θ は (X, Y, Z, W) と (X_0, Y_0, Z_0, W_0) を結ぶ線分上のある点での値を意味する．

ところで，

$$(11)\quad E(X-X_0)^2=NQ(1-Q),\ E(Y-Y_0)^2=NP_1(1-P_1),\ E(Z-Z_0)^2=NP_3(1-P_3),$$
$$E(W-W_0)^2=NP_{13}(1-P_{13}),\ E(X-X_0)(Y-Y_0)=NQP_1,$$
$$E(X-X_1)(Z-Z_0)=NQP_3,\ E(X-X_0)(W-W_0)=NQP_{13},$$
$$E(Y-Y_0)(Z-Z_0)=-NP_1P_3,\ E(Y-Y_0)(W-W_0)=-NP_1P_{13},$$
$$E(Z-Z_0)(W-W_0)=-NP_3\cdot P_{13}$$

また

$$(12)\quad \frac{\partial\hat{N}}{\partial X}=1,\ \frac{\partial\hat{N}}{\partial Y}=\frac{Z}{W},\ \frac{\partial\hat{N}}{\partial Z}=\frac{Y}{W},\ \frac{\partial\hat{N}}{\partial W}=-\frac{YZ}{W^2}$$
$$\frac{\partial^2\hat{N}}{\partial X^2}=\frac{\partial^2\hat{N}}{\partial Y^2}=\frac{\partial^2\hat{N}}{\partial Z^2}=0,\ \frac{\partial^2\hat{N}}{\partial W^2}=2\frac{YZ}{W^3},$$
$$\frac{\partial^2\hat{N}}{\partial X\partial Y}=\frac{\partial^2\hat{N}}{\partial X\partial Z}=\frac{\partial^2\hat{N}}{\partial X\partial W}=\frac{\partial^2\hat{N}}{\partial Y\partial Z}=0,\ \frac{\partial^2\hat{N}}{\partial Y\partial W}=-\frac{Z}{W^2},\ \frac{\partial^2\hat{N}}{\partial Z\partial W}=-\frac{Y}{W^2}$$

ここで，(10) の両辺を二乗して，平均をとれば，(11)，(12) を代入して，さらに N が十分大のとき，二次の偏微分は確率 1 で，$O(N^{-1})$ なることに注意して，

$$(13)\quad E(\hat{N}-N)^2=E(X-X_0)^2\left[\frac{\partial\hat{N}}{\partial X}\right]_0^2+E(Y-Y_0)^2\left[\frac{\partial\hat{N}}{\partial Y}\right]_0^2+E(Z-Z_1)^2\left[\frac{\partial\hat{N}}{\partial Z}\right]_0^2$$
$$+E(W-W_0)^2\left[\frac{\partial\hat{N}}{\partial W}\right]_0^2+2E(X-X_0)(Y-Y_0)\left[\frac{\partial\hat{N}}{\partial X}\right]_0\left[\frac{\partial\hat{N}}{\partial Y}\right]_0$$
$$+2E(X-X_0)(Z-Z_0)\left[\frac{\partial\hat{N}}{\partial X}\right]_0\left[\frac{\partial\hat{N}}{\partial Z}\right]_0$$
$$+2E(X-X_0)(W-W_0)\left[\frac{\partial\hat{N}}{\partial X}\right]_0\left[\frac{\partial\hat{N}}{\partial W}\right]_0$$
$$+2E(Y-Y_0)(Z-Z_0)\left[\frac{\partial\hat{N}}{\partial Y}\right]_0\left[\frac{\partial\hat{N}}{\partial Z}\right]_0$$
$$+2E(Y-Y_0)(W-W_0)\left[\frac{\partial\hat{N}}{\partial Y}\right]_0\left[\frac{\partial\hat{N}}{\partial W}\right]_0$$
$$+2E(Z-Z_0)(W-W_0)\left[\frac{\partial\hat{N}}{\partial Z}\right]_0\left[\frac{\partial\hat{N}}{\partial W}\right]_0+O(1)$$

$$=N\Big\{Q(1-Q)+P_1(1-P_1)\left(\frac{P_1}{P_{13}}\right)^2+P_3(1-P_3)\left(\frac{P_3}{P_{13}}\right)^2$$
$$+4P_{13}(1-P_{13})\left(\frac{P_1\cdot P_3}{P_{13}}\right)^2+2QP_1\left(\frac{P_3}{P_{13}}\right)+2QP_3\left(\frac{P_1}{P_{13}}\right)$$
$$+2QP_{13}\left(\frac{P_1P_3}{P_{13}^2}\right)-2P_1P_3\left(\frac{P_1\cdot P_3}{P_{13}^2}\right)+2P_1P_3\left(\frac{P_1P_3^2}{P_{13}^3}\right)$$
$$+2P_3P_{13}\left(\frac{P_1^2\cdot P_3}{P_{13}}\right)+O(N^{-1})\Big\}$$

(13) は，N の推定量 $\hat{N}$ の漸近的な平均平方誤差を表わしている．ここで同じようにして，漸近的な $\hat{N}$ の偏よりをもとめてみる．(10) のテーラー展開の式でもう一つ三次のところまで書き表わし，両辺の平均をとると，X や X, Z, W などの3次の平均の周りのモーメントは，すべて $O(N)$，また $\hat{N}$ の×次の微役数は確率 1 で，$O(N^{-2})$ であり，さらに $E(X)=X_0$, $E(Y)=Y_0$, $E(Z)=Z_0$, $E(W)=W_0$ であったことから，一次の項は 0 になり，結局2次の所の平均が問題であり，それは (10) 式の微係数をとる点 θ を，$(X_0,\ Y_0,\ Z_0,\ W_0)$ でおきかえて，平均をとれば得られるわけである．すなわち，

$$(14)\quad E(\hat{N}-N)=\frac{1}{2}E(X-X_0)^2\left[\frac{\partial^2\hat{N}}{\partial X^2}\right]_0+\frac{1}{2}E(Y-Y_0)^2\left[\frac{\partial^2\hat{N}}{\partial Y^2}\right]_0+\frac{1}{2}(Z-Z_0)^2\left[\frac{\partial^2\hat{N}}{\partial Z^2}\right]_0$$
$$+\frac{1}{2}E(W-W_0)^2\left[\frac{\partial^2\hat{N}}{\partial W}\right]_0+E(X-X_0)(Y-Y_0)\left[\frac{\partial^2\hat{N}}{\partial X\partial Y}\right]_0$$
$$+E(X-X_0)(Z-Z_0)\left[\frac{\partial^2\hat{N}}{\partial X\partial Z}\right]_0+E(X-X_0)(W-W_0)\left[\frac{\partial^2\hat{N}}{\partial X\partial W}\right]_0$$
$$+E(Y-Y_0)(Z-Z_0)\left[\frac{\partial^2\hat{N}}{\partial Y\partial Z}\right]_0+E(Y-Y_0)(W-W_0)\left[\frac{\partial^2\hat{N}}{\partial Y\partial W}\right]_0$$
$$+E(Z-Z_0)(W-W_0)\left[\frac{\partial^2\hat{N}}{\partial Z\partial W}\right]_0+O(N^{-1})$$

ここで，また (11) と (12) をもちいて，

$$=\frac{1}{2}P_{13}(1-P_{13})\cdot 2\frac{P_1\cdot P_3}{P_{13}^3}+P_1P_{13}\cdot\frac{P_3}{P_{13}^2}+P_3P_{13}\cdot\frac{P_1}{P_{13}^2}+O(N^{-1})$$
$$=\frac{P_1P_3}{P_{13}^2}+\frac{P_1\cdot P_3}{P_{13}}+O(N^{-1})$$

T が一般の時は，$\hat{N}$ を陽函数に表わせないが，(6″) 式からテーラー展開に必要な偏微係数などは求まる．

（統計数理研究所）

参考文献

[1] Leslie & Chitty: "The estimation of population parameters fiom capture-recapture data III" Biometrika (1953).

[2] 橋爪浅治："動く母集団と厚生統計" 生物統計学雑誌 6 (3) (1959).

[3] J. N. Darroch: "The multiple-recapture census" Biometicka, vol. 45 (1958).