

Bayesian Inference: With ecological applications

W. A. Link and R. J. Barker (2010)

Chapter 11: Open Population Models (11.3~11.5)

11.3 Open population mark-recapture – The CJS model (P.247)

Cormack-Jolly-Seber (CJS) モデル (セクション 5.5 を参照) は、セクション 11.2 で登場した band-recovery model と密接に関係している。これらの両モデルが適用されたデータのタイプは異なるが、それぞれの尤度関数は代数的に等しいものであることが簡単にわかる。(注釈 5 : CJS モデルの下で得られたデータは m_{ij} 配列の形で要約することができ、この形は band-recovery matrix と同じ形である (Table 11.1)。CJS 尤度は、banding model の再パラメータ化を通して得られた。)

CJS モデルは、離散的な時系列サンプリングデザインの下でタグ付けと放流、そして再捕獲を記録した動物個体群データの解析に適している。CJS モデルの重要な特徴であり、Jolly-Seber (JS) モデル (11.4 に登場) との相違点は、放流を必要条件とし、各動物個体の最初の捕獲のモデリングを試みないことである ; つまり、再捕のみをモデリングする。CJS モデルは動物の生存に注目している。

各サンプリング時において全ての動物個体が捕獲されるわけではないので、CJS モデルでは生存と再捕獲のプロセスの部分的な交絡が生じる。ある動物個体を再捕獲できなかった場合、1) サンプリングが行われた時点ではまだ生きているが捕獲できなかった、2) すでに死亡している、のどちらが原因となっているかを知ることができない。後に行われるサンプリングで捕獲された場合はその動物がそれまで生存してきたことを示すが、死亡した個体を再捕獲することはできない。サンプリング終了まで個体を再捕獲できなかったという事実は、すでに死亡している可能性だけでなく、再捕獲を避け続けていた可能性も意味する。このように、生存と再捕獲のプロセスを解きほぐすことは難題である。

データの基本的な構造は、closed mark-recapture model と同じである。行が各動物個体、列がサンプリング時に対応した捕獲情報を表す X^{obs} の配列によって調査結果が要約される。closed モデルと CJS モデルの重要な違いは、(CJS モデルは) 各動物個体の最初の放流と、適用できる場合は捕獲時の損失 (losses on capture) を必要条件とすることである。したがって、(CJS モデルでは) セクション 9.2 で取り扱った X^{mis} 項は含まれない。

捕獲時の損失は複雑である ; 標識再捕法の研究では、罾による捕獲時の動物の死亡、タグの脱落、未標識状態での放流などが原因となる。

以下の条件で、捕獲時の損失の項を CJS モデルに組み込むと、生存率や捕獲率に対する捕獲時の損失の影響を考慮する必要がなくなる。

- i) サンプル j における捕獲時の損失は、生存以外の事前の出来事と関連しない (以前の捕獲などと独立ということ?)
- ii) 捕獲時の損失パラメータは、モデル上では他のパラメータと区別できる

An Observed Data Likelihood (P. 248)

y_i が各動物個体を明示的に表している要素とする、動物の初捕獲時を表すベクトルを y とした場合、最初の捕獲という条件付きの X^{obs} の全ての情報に基づく Observed Data Likelihood (ODL) は以下の様に示すことができる。

$$\begin{aligned} \mathcal{L}(\phi, p, v | X^{obs}, y) &\propto [X^{obs} | \phi, p, y, v] \\ &= [X^{obs} | R, r, m] \times [R | n, v] \times [r | R, \phi, p] [m | T, \phi, p] \end{aligned} \quad (11.5)$$

パラメータと統計値の定義は以下の通り (Table 11.4)。

Parameters

- ϕ_j 調査時 j から $j+1$ への生存率 ($j = 1, \dots, k-1$)
- p_j 調査時 j において再捕獲可能な個体を捕獲する確率 ($j = 2, \dots, k$)
- v_j 調査時 j において捕獲個体を再放流する確率 ($j = 1, \dots, k$)

Statistics

- n_j 調査時 j における全捕獲個体数 ($j = 1, \dots, k$)
 - R_j 調査時 j における標識放流個体数 ($j = 1, \dots, k$)
 - r_j 調査時 j で標識放流され、後に再捕獲される個体数 ($j = 1, \dots, k-1$)
 - m_j 調査時 j において捕獲された標識個体数 ($j = 2, \dots, k$)
 - T_j 調査時 j における個体群中の標識個体数 ($j = 2, \dots, k$)
-

先述の bird-banding model で用いられた生存率 S_j に代わって、 ϕ_j が生存率を表すパラメータとして用いられている。 ϕ_j は「見かけの」生存率である。調査地から永続的な移出が生じた場合、この移出は死亡と区別できない。(注釈 6: 永続的な移出あるいはランダムで一時的な移出の仮定の下では、CJS 尤度は妥当である。これらのアイデアは Burnham (1993) が詳しい。)

式 11.5 の 4 つの項の内、 $[R | n, v]$ は捕獲時の損失を扱い、最も興味が低い。この項は、捕獲個体数 n_i とあまり興味のない (nuisance) パラメータ v_i に依存するという条件の下での調査時 i における標識放流数をモデル化している。

3 番目と 4 番目の項は、必要最小限の統計値によって表現される CJS モデルの ODL に対応している。CJS モデルは 2 項分布で表現することができる。

$$\begin{aligned} [r | R, \phi, p] &= \prod_{j=1}^{k-1} B(r_j; R_j, \lambda_j) \\ [m | T, \phi, p] &= \prod_{j=2}^k B(m_j; T_j, \tau_j) \end{aligned}$$

これらのパラメータの内、 λ_i と τ_i はセクション 5.4 に対応している。 $(\lambda_i : i$ 時点で放流され、

i 以後に再び捕獲される確率、 τ_i : i 以前に放流され、 i 時点かそれ以後に再び捕獲される確率)。この CJS モデルの式についての Gibbs sampler の記述はセクション 5.4 を参照。

ODL の最初の項 $[X^{obs}|R, r, m]$ にパラメータ v, ϕ, p が含まれていないことは注目すべき点である。この項は多項超幾何分布の産物として表現され、分割表(contingency tables)に基づいて goodness of fit tests に用いることができる(例として Pollock et al., 1990 を参照)。

セクション 5.4 で議論した通り、ODL の複雑な形は、このモデルの式を用いた場合、 ϕ と p の全条件付分布 (full conditional distribution) は得られないことを意味している。しかしながら、調査によって捕獲された各動物個体を要素とする隠れベクトル d によってデータを補完して、扱いやすい式に変形できる。もし、 h 時から $h+1$ 時 ($h < k$) の間に個体 i が死亡した場合、 $d_i = h$ となる。また、もし k 時まで生きていたら $d_i = k$ となる。

もし、初捕獲を表すベクトル y と同じ次元で、捕獲時の損失を示すためにベクトル v を用いる場合、最初の放流と捕獲時の損失の条件の下での完全なデータモデルは以下の様
に書くことができる。

$$\begin{aligned} [X^{obs}, d | y, v, \phi, p] &= [X^{obs} | d, y, v, \phi, p] \times [d | y, v, \phi, p] \\ &= [X^{obs} | d, y, v, p] \times [d | y, v, \phi] \end{aligned}$$

この式は、 ϕ と p に対して共役事前分布が使われている条件の下で、 d の未知で打ち切られた要素に対する既知の全条件付分布に基づく Gibbs sampler を導く (Dupis, 1995 ; Schofield, 2007)。

11.3.1 A Gibbs Sampling Algorithm for the CJS Model (P. 250)

事前分布 $\text{Be}(\alpha_\phi, \beta_\phi)$ を ϕ_j に対して用いることで、全条件付分布もまた下記の様なベータ分布となる。

$$[\phi_j | \cdot] = \text{Be}(M_j^+ - D_j + \alpha_\phi, D_j + \beta_\phi)$$

ここで、 M_j^+ は j 時直後における個体群内の標識個体数を、 D_j は M_j^+ の内で j 時から $j+1$ 時の間に死亡した個体数を表す。捕獲時の損失が生じた動物個体は j 時から $j+1$ 時の間に死亡したとみなさないことに注意。これらの個体は統計値 M_j^+ と D_j の算出に用いられない。

事前分布 $\text{Be}(\alpha_p, \beta_p)$ を p_j に対して用いることで、全条件付き分布もまた下記の様なベータ分布になる。

$$[p_j | \cdot] = \text{Be}(m_j + \alpha_p, M_j - m_j + \beta_p)$$

ここで、 M_j は j 時直前における個体群内の標識個体数を、 m_j は j 時に捕獲された標識個体数

を表す。

$\{d_j\}$ の未知かつ打ち切られた要素に関する全条件付分布は、モデル $[d|y, \phi]$ によって得られる。個体 i に対し、 $[d_i|y_i, \phi]$ はサンプル空間 $\{y_i, y_{i+1}, \dots, k\}$ とパラメータベクトル $\{\xi_{y_i, y_i}, \xi_{y_i, y_{i+1}}, \dots, \xi_{y_i, k}\}$ を持つカテゴリカル分布となる。 ξ_{ij} は i 時に初めて捕獲され、 j 時に死亡する確率を表す。このパラメータベクトルは $i = 1, \dots, k - 1$ において下記の生存率に従うものとする。

$$\xi_{ij} = \begin{cases} 1 - \phi_i & i = j \\ \phi_{k-1} & i = k - 1; \quad j = k \\ \xi_{i+1, j} \phi_i & i = 1, \dots, k - 1; \quad j = 2, \dots, k \end{cases}$$

例として、 $k = 4$ となる調査における ξ_{ij} の値を Table 11.5 に示した。隠れベクトル d の更新 (updating) は以下のいずれかに従って行われる。

(i) もし個体 i が k 時点 (最終サンプリング時) で捕獲されたら、 d_i は k に等しくなる。なぜなら、この個体は k 時以降に死亡することが明確だから。

(ii) l_i 時点で個体 i が最後に捕獲された場合、 d_i は l_i で打ち切られる。そして、サンプル空間 $\{l_i, l_{i+1}, \dots, k\}$ とパラメータベクトル $(\xi_{l_i, l_i}, \xi_{l_i, l_{i+1}}, \dots, \xi_{l_i, k})$ を持つカテゴリカル分布から d_i の値を引き出す。

(iii) l_i 時点での捕獲によって個体 i が損失した場合、望むのであれば (ii) の条件の下で d_i の値を引き出すことができる。この値は、個体 i の予測死亡時間とみなすことができる。他のすべての未知なものに関する推定が妥当であったとしても、これらの個体 i に対する d_i の更新は本質的な事ではないことに注意。

上記で記述された全条件付分布に基づく効果的な Gibbs sampler は、R のような言語を用いて簡単にプログラミングすることができる。しかしながら、隠れ死亡時間を含む CJS モデルは、Panel 11.2 で示された BUGS (Schofield *et al.*, 2008) によって簡単に実行することもできる。

BUGS コードの特徴は以下の通り。

- $S[j]$ 、 $p[j]$ はそれぞれ j 時の生存率と捕獲率を表す。
- $first[i]$ は動物個体 i が初めて捕獲・標識放流されたサンプル時を表すデータである。
- $AvailUntil[i]$ は捕獲によって動物が生きていることが確認された最後のサンプリング時を表すデータである。 h 時に動物個体 i が取り除かれた (死亡した) 場合、 h が値として割り当てられ、最後まで取り除かれなかった (生き残った) 場合、 k が値として割り当てられる。このデータが示す時間断面以降の捕獲率は 0 が与えられる。CDL 解析では、死亡時間はこれらの動物に帰属され、推定死亡時間は動物が研究の間生き残っていることとして解釈される。
- 隠れ変数であるノード $a[i, j]$ は、動物個体 i が j 時において捕獲可能な場合は 1、そうで

ない場合は 0 となる。 $a[i, first[i]]$ は 1 が割り当てられているデータである。

・ノード $asuse[]$ 、 $sv[]$ 、 $apuse[]$ 、 $pcap[]$ は、未標識や死亡している動物個体は尤度に貢献しないことを確実にするために用いられる。

・ノード $avail[]$ は動物個体が再補され得るかどうか（つまり、標識され個体群から除外されていない）を追いつける。

Panel 11.2 で指定されている CDL は 1) ベルヌーイ分布、2) 動物が生存しているかどうかと捕獲可能かどうかに依存する確率、によって記述されている。このコードは動物の死亡時間の隠れベクトルを明示的に含んでいないが、このベクトルは行列 $a[]$ から簡単に得ることができる。（ $a[]$ は動物の「利用可能状態」を表すベクトル）

死亡の隠れ時間も含んだ CDL によって CJS モデルを記述することの主な利点は、計算上の便利さと MCMC 実行の速さである。死亡の隠れ時間は、個体群内に残っている標識個体数の予測を得るために用いられることもあるが、このこと（標識個体数の予測？）が興味の対象となることはめったにない。

CJS モデルは個体群サイズや出生率を含んでいない。これらはサンプリング時における放流個体数を条件つけることで制限されている。現存量と出生の推定を行うために、完全な捕獲-再捕獲データセットに関する尤度を用いなければならない。特に、最初の捕獲に関するモデルに含まなければならない。

11.4 Full Open Population Modeling Formalities (P. 252)

閉鎖系個体群の研究では、研究期間内に変化しない単一の個体群サイズによって個体群が記述される。開放系個体群の記述は、個体群動態の性質を反映して複雑になる。これらの違いにもかかわらず、閉鎖系と開放系の個体群モデルは同一の一般的な方法で記述することができる。この違いとは、研究期間内の個体群の変化を特徴づけるために用いることができる様々な要点 (summaries) である。

閉鎖系個体群の研究における興味のある変量は個体群サイズである。開放系個体群の研究における興味の対象は、研究期間中に生存している動物個体数である。これらの動物の総数を“個体群”と呼ぶのは魅力的であるが、困難も伴う。ある動物は隣接するサンプリング時の間（例： j 時と $j+1$ 時）に出生・死亡し、それゆえに実験において目に見えない状態となる。これらの個体は解析対象から除外することが望ましい。例えば、調査期間の賢明な選択はこれらの個体数を無視することができる。したがって、我々は、少なくとも 1 回のサンプリング時に生存しており捕獲可能な全個体を記述するために“個体群”という言葉を使う。（注釈 7：“supermodel”のように、Schwarz and Arneson (1996) で使われた“superpopulation”という言葉は、仮定上の何かを暗示していて紛らわしい。現実の個体数を表す言葉としては“population”が望ましい。）

個体群中の捕獲可能な動物個体数を表すために A を用いることで、CDL を以下の様に表すことができる。

$$CDL \propto [X^{obs}, X^{miss}, I | \theta^{obs}, \theta^{miss}, \Lambda]$$

この式は、追加された指標 (inclusion indicator) I を除けば本質的には閉鎖個体群に用いたものと同一の一般的表現である。 $A \times t$ 行列はサンプリング j 時に個体 i が捕獲可能かつ捕獲された場合 $I_{ij} = 1$ となり、そうでない場合は $I_{ij} = 0$ となる。また、個体群は開放系であるために、パラメータベクトル θ は、捕獲確率と同じように、個体群へ移入と個体群からの移出をモデリングするのに用いられる要素を含む。

もし、動物個体がサンプリング j 時以前に個体群に移入し、サンプリング j 時以降に個体群から移出したとするならば、その個体はサンプリング j 時に捕獲可能である。したがって、 $A \times 1$ ベクトル b と a を用いて追加された指標 I における情報を要約化できる。ここで、要素 b_i および d_i は、それぞれ個体 i の出生および死亡が発生した期間を表す。

この章の冒頭で議論したように、標識再補法の研究においては、全ての出生と死亡の時間は打ち切られている。研究開始時においてすでに存在する個体の出生時間 b は左側に打ち切られており、 t_1 (調査開始時) 以前に出生したということしかわからない。研究期間の最後まで生存していた個体の死亡時間は右側に打ち切れ、 t_k (調査終了時) 以降に死亡したということしかわからない。研究期間の間に出生と死亡の両方が生じた場合、区間打ち切りとなる。

開放系個体群モデルは出生時間 b_i と死亡時間 d_i 、そしてそれらの間の関係に注目する。CDL を用いることで、サンプリング j 時に捕獲可能な動物個体数 N_j と j 時から $j+1$ の間に生まれて少なくとも $j+1$ 時まで生存している個体数 B_j といった興味ある変量を調べることができる。

Figure 11.1 は、4 つの研究期間に関連した 18 個体の一生を描いたものである。16 個体は研究期間の間生存していたが、行列 X で説明できるのは 13 個体しかない。隣り合うサンプリング時の間に出生と死亡が生じている 3 個体はモデルから除外される。 $A = 13$ 個体それぞれが、(i) t_1 以前に生まれた、(ii) t_1 と t_2 の間に生まれた、(iii) t_2 と t_3 の間に生まれた、(iv) t_3 と t_4 の間に生まれた、のどれかによって記述される。 t_1 以前に生まれた個体数は B_0 で表現され、 $B_0 = 1, B_1 = 5, B_2 = 4, B_3 = 3$ を Fig 11.1 より得ることができる。また、同様に $N_1 = 1, N_2 = 6, N_3 = 7, N_4 = 8$ も得ることができる。

Jolly-Seber Model (P. 254)

最初の完全な開放系個体群モデルは Jolly-Seber (JS) model である (Jolly, 1965, Seber, 1965)。このモデルはセクション 11.3 の $[X^{obs} | \phi, p, v, y]$ に最初の捕獲に関するモデル $[y | p, U]$ を追加したものである。ここで、サンプリング j 時に個体群中に存在する未標識個体数 U_j を要素とするベクトルは $U = (U_1, \dots, U_k)'$ と定義される。ちなみに、Jolly と Seber は同じモデルを独立に提案し、Biometrika に 1965 年に並んで発表している。

JS モデルでは、「どの時点で動物が始めて捕獲されたのか」という y に含まれる情報は、サンプリング j 時に捕獲された未標識個体数 u_j を要素とするベクトル $u = (u_1, \dots, u_k)'$ によって要約される。最初の捕獲に関するモデルは以下の様になる。

$$[y|p, U] \propto \prod_{j=1}^k [u_j|p_j, U_j]$$

ここで、 $[u_j|p_j, U_j]$ は指標 U_j とパラメータ p_j の二項分布である。

U_j の値は Fig 11.1 から推測することができない。なぜなら、 U_j はサンプリング j 時に新しく生まれ、それゆえ未標識な個体 B_{t-1} だけではなく、 t_{i-1} 以前に生まれたけどまだ捕獲されていない個体も含んでいるからである。この出生と捕獲のプロセスの個体の交絡によって、パラメータ U は生物学的に特別な興味とならない。しかしながら、 U_j とモデル内の他のパラメータ推定値を得ることに価値はあり、 N_j と B_j の推定値を求めることを可能にする。 U に関する尤度をパラメータ化することの欠点は、モデルを階層的に拡張させて複雑にすることである。 U は個体群と標識放流実験の複雑な相互作用（上記の様な交絡）を反映しているため、このパラメータの事前分布として妥当なプロセスを考え出すのは困難である。より生物学的に意味のある変数である N_j と B_j をモデル化できることを我々は望んでいるため、プロセス事前分布(process priors)は興味の対象である。

11.5 CMSA Model and Extensions (P. 255)

JS モデルの重要なバリエーションの 1 つが、まず Crosbie and Manly (1985) によって開発され、後に Schwarz and Arneson (1996) によって拡張された。我々はこのモデルを JS モデルの CMSA パラメータ化と呼ぶ。Link and Barker (2005) で議論されたように、この定式化の利点は階層的な拡張が容易であることである。

CMSA モデルは A とベクトル $\beta = (\beta_0, \dots, \beta_{k-1})'$ についてパラメータ化される。この β_j は捕獲可能な動物個体がサンプリング j 時から $j+1$ 時の間に個体群に移入する確率である。出生イベントは独立し、ベクトル β は各動物個体で同値という仮定の下で、出生頻度のベクトル B は多項ランダム変数(multinomial random variables) $B|A, \beta \sim M_k(A, \beta)$ としてモデル化され得る。

これは、CJS の ODL 拡張としての CMSA の基本である。

CMSA モデルに関する ODL の詳しい記述は、Pradel (1996) で用いられたモデルとの比較の議論も合わせて Link and Barker (2005) に示されている。ここでは、隠れ出生時間 b と隠れ死亡時間 d についての CDL に関する推論に基礎をおく。このアプローチでは、 N_i あるいは B_i は b と d の単純な関数として表されるため、これらの要約に関する推論は簡単である。(CMSA が?) CDL と共に作用することはより効果的な Gibbs sampler を導くことができ、後に示すコードの様に、BUGS の実行が比較的容易である。

11.5.1 Complete Data Likelihood (p. 255)

CMSA モデルに関する CDL は以下の式に比例する。

$$[X, b, d | \Lambda, \beta, p, \phi] = [X | p, \Lambda, b, d][d | \phi, b, \Lambda][b | \beta, \Lambda]$$

ここで、 X は完全な $(A \times k)$ X 行列、 $\begin{pmatrix} X^{\text{obs}} \\ 0 \end{pmatrix}$ である。単純化のため、捕獲時の損失は無視。

$[X | p, \Lambda, b, d]$ 項は、各動物個体の出生前や死亡後は捕獲率が 0 になるという点以外は閉鎖系再捕獲モデルの $[X | p, N]$ と似ている。同様に、 $[d | \phi, b, \Lambda]$ 項は、調査地点への初侵入ではなく出生を条件付けしているという点以外は CJS モデルの平衡項と同じである。

CDL モデルにおける出生をモデリングするために、以下の式を用いた。

$$[b | \beta, \Lambda] = \prod_{i=1}^A [b_i | \beta]$$

ここで、 $[b_i | \beta]$ はサンプル空間 $\{1, \dots, k\}$ とパラメータベクトル β を持つカテゴリカル分布である。ここでは、動物がサンプリング $h-1$ 時から h 時の間に生まれたことを $b_i = h$ で示し ($h = 2, \dots, k$)、最初のサンプリング以前に出生した場合は、 $b_i = 1$ となる。今回は調査地への初侵入ではなく出生を条件としたこと以外、CJS モデルで行ったものと同じ方法で死亡時間 d_i のモデリングを行った。我々はすべての動物 A に対して死亡時間のモデリングを行った。つまり、我々はサンプル空間 $\{b_i, b_{i+1}, \dots, k\}$ とパラメータベクトル $(\xi_{b_i b_i}, \xi_{b_i, b_{i+1}}, \dots, \xi_{b_i, k})$ を持つカテゴリカル分布として $[d_i | \phi, b_i]$ を指定することによって d_i ($i = 1, \dots, A$) をモデル化した。要素 ξ_{ij} はセクション 11.3.1 で定義したものである。

CJS モデルでは、死亡が部分的に観測 (partially observed) される。つまり、最後の捕獲時以降に死亡が発生することが分かる。同様に、出生も部分的に観測される。最初の捕獲以前に出生が発生していることがわかる。

CMSA モデルに適合するために、Panel 11.2 の BUGS コードは比較的簡単に以下のようにして拡張することができる。

1. A の更新に伴う次元の変化を明らかにするために、データ増強ステップ (data-augmentation step) を追加
2. 出生時間を説明するために、打ち切り多項式項を追加
3. 調査地への初侵入時間ではなく出生時を基にした見込み死亡時間を用いて d に関するモデルを適合

我々は、一度も捕獲されていない動物に関する k にセットされた項目とともに、打ち切り出生時を含むデータに、 $(A \times 1)$ ベクトル ($birthCT$) を追加した (?)。superpopulation のメンバーがすでに死亡しているかどうかを表すために、我々は部分的に観察された $(A \times k)$ データ行列 ($notyetdead$) を含めた。これは調査開始時に全動物に 1 が値として割り当てられ、動

物 i の最初と最後の視認の間の全てのサンプルに対して要素 ij は 1 が割り当てられ、最後の捕獲の後は全てのサンプルに NA が割り当てられる。

Panel 11.3 のコードを用いることで、行列 $avail []$ と $a []$ の単純な関数として N_i と B_i の予測を容易に構築することができる (BUGS コードは Panel 11.4)。しかしながら、パラメータ ϕ と β の単純な推定値と N あるいは B の予測値は、殆どの研究者を満足させないだろう。より大きな興味は、より高いレベルの要約を手に入れるための手段として、あるいは生物学的に面白い問いに答えるために、これらの変量をモデリングすることである。

11.5.2 Example: Nonmelanic *Gonodontis* (P. 256)

Bishop et al. (1978) は、イングランドのリバプール近郊の Cressington Park で 689 の *Gonodontis bidentata* の非黒色雄を標識再放流する実験を毎日行った研究の結果を報告した。*Gonodontis* は、黒色・非黒色の 2 つの形態を出産し、工業汚染のために黒色タイプの優占度が増加した。これは自然選択を示す最も古典的な研究例の一つである。

セクション 5.5.2 では、*Gonodontis* のデータに CJS モデルを当てはめた。ここでは、Fig 11.2 で示された事後分布の要約と Panel 11.3 の BUGS コードを用いて、*Gonodontis* のデータに対して、日別の現存量・出生・生存確率に関する CMSA モデルを当てはめた。

ベイズ推定の特徴は、セクション 5.2 で議論したように、異なるパラメータ化を用いて推定を表すことの簡便さである。 N や B に関して個体群 demographics を記述することに代わって、個体あたりの出生率 f (先述の bird-banding model で用いた recovery rate パラメータと同一ではないことに注意)、あるいは個体群成長率 λ の様な変量についての事後分布密度を、Link and Barker (2005) で示された関係を用いることによって構築することができる。これらの再パラメータ化は次のように行われる。

$$f_j = \frac{\beta_j}{d_j} \quad j = 1, \dots, k-1$$

ここで、以下のような条件となる

$$d_j = \begin{cases} \beta_0 & j = 1 \\ d_{j-1} S_{j-1} + \beta_{j-1} & j = 2, \dots, k-1 \end{cases}$$

$$\lambda_j = S_j + f_j \quad j = 1, \dots, k-1$$

f と λ の事後分布密度からサンプルを生成するために、我々は上記の機能的な関係を用いて、 S と β の Markov chain を f と λ の Markov chain に単純に変形した。*Gonodontis* の解析に関する f と λ の事後分布の要約は Fig 11.3 に示されている。

日別個体群増加率に関する事後分布密度を基に、Cressington Park の *Gonodontis* 個体群の非黒色雄は、終盤ごろに個体群の明確な減少を示したが研究期間を通してたいていは増加していることが示された。蛾の出生数 (B_j) と現存量 (N_j) から同様の結論が示されているが、CMSA パラメータ化の利点は、平均増加率について個体群の挙動を要約化する手段として λ_j の様なパラメータにランダム効果を入れ、そして parsimony を導入することにモデルを簡単に適合させられることである。

我々がこの本の中で何回か強調しているさらなる利点は、階層ベイズモデリングは研究者が取り扱う問題を大きく拡張できることである。1つの例はパラメータセット間の関係をモデリングできることである。共通の環境の影響によって、生存確率と個体あたりの出生率の間には関係があると考えられることから、Link and Barker (2005)は以下の様な統計的制約を持つパラメータ ϕ と f に関する多項正規ランダム効果モデルを構築した。

$$(\text{logit}(\phi_j), \ln(f_j))' \sim MVN(\mu_j, \Sigma)$$

ここで、要素 μ_j は $\text{logit}(\phi_j) = \mu_{1j}$ の推定値と $\ln(f_j) = \mu_{2j}$ の推定値である。パラメータ Σ は、パラメータの結合分布の共分散を表す分散-共分散行列である。

この”プロセス共分散”は、demographic パラメータ間の生物学的に明確な関係を記述しており、推定のプロセスに由来したパラメータ間 (between-parameter) を「推定」する生物学的に見間違いなサンプリング共分散との混乱を生じない。

このモデルを当てはめるために、 β の代わりとして f についてのモデルの再パラメータ化と、 ϕ と f の結合分布についての多項正規事前分布の追加によって、Panel 11.3 の BUGS コードを適合することができる。Link and Baker (2005)は、 μ の要素についてはあいまいな正規事前分布(vague normal prior distribution)を用い、 $[\Sigma]$ に関しては逆 Wishart $IW_2(V, df)$ 分布を用いた。

階層的な関係から、このコードはデータの混合が遅い。やきもきしながら (Being impatient?)、Link and Barker (2005)は、 u や研究中に捕獲された動物個体数の条件付けによって得られた部分尤度 (partial likelihood) を代わりに用いた。これは、Pradel (1996)でも用いられた手法である。データである u を制限することによって、重要な情報がモデルから廃棄されるのは明らかである。しかしながら、Link and Baker (2005)は、 u は事実上補助的であり、モデルの中で同一視可能なパラメータの推定についての有用な情報を何も含んでいないことを意味することを示した。 ϕ と f に関する階層モデルを含んだ後は、Link and Barker (2005)は2つのパラメータセットの間の正の関係に関する証拠を見つけた (Fig 11.4)。

ϕ と f の間の明確な関係に関するあり得そうな説明は、両パラメータともに蛾の現存量に依存することによる密度効果である。CDL は現存量 N_j の予測に用いることができるため、生存確率や個体あたりの出生率に対する予測子として現存量を含むことも可能である。このアプローチは、同じデータセットに対して Schofield et al. (2008)で示されている。