

第8章 MCMC法とベイズ統計モデル

2015/02/03 高倉担当

前章の内容：観測できない“個体差”をGLMMで記述

- ランダム効果 r_i を含めて、めでたしめでたし
 - 尤度を求めるときに r_i で積分していた
- 個体差が複数になってもGLMM
 - 複数の r_i について**多重**積分
 - 計算が難しいか、そもそもできないことも

この章：積分なしに**尤度**を計算するMCMC法

Markov Chain Monte Carlo

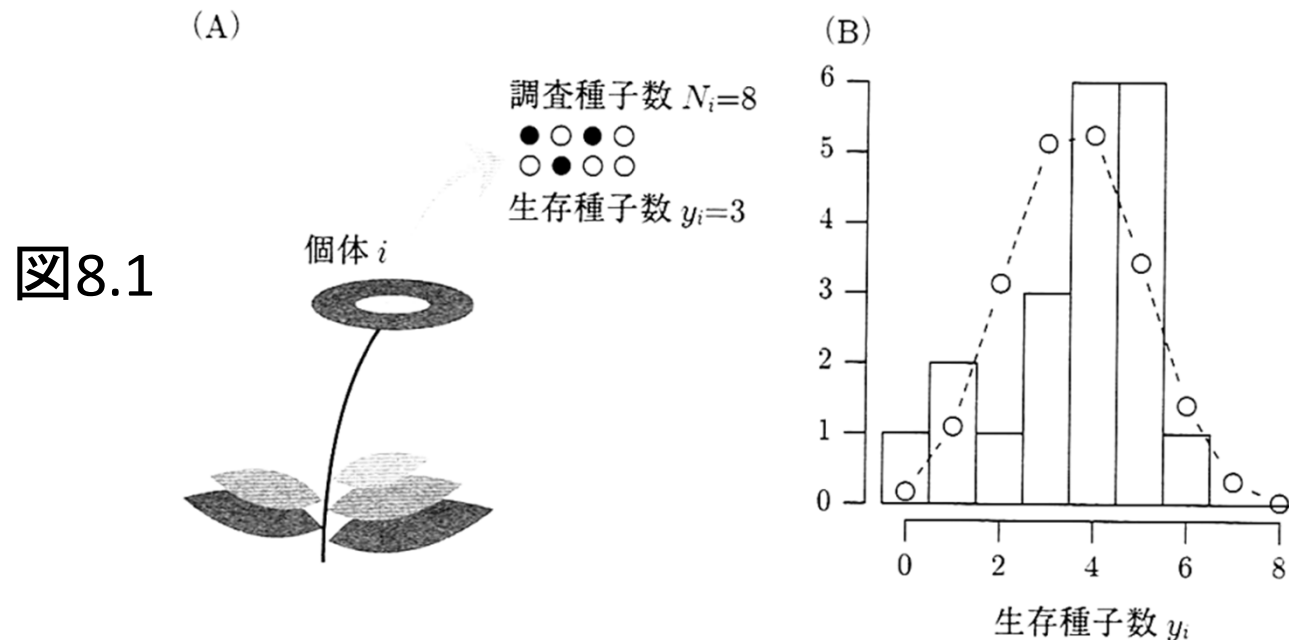
+ ベイズ統計の考え方

例題8.1 種子の生存確率（個体差なし）

ごく簡単な（GLMで事足りる）例で考える

- 8種子/個体の生存率を20個体分調べた

$$Y = \{y_1, y_2, \dots, y_{20}\} = \{4, 3, 4, 5, 5, 2, \dots, 4\}$$



ある個体 i の生存種子が y_i 個である確率

$$p(y_i|q) = \binom{8}{y_i} q^{y_i} (1 - q)^{8-y_i}$$

尤度 $L(q)$ は20個体分の「データが得られる確率の積だから」

$$L(q) = p(\mathbf{Y}|q) = \prod_i p(y_i|q)$$

これの対数をとると

対数尤度

$$\log L(q) = \sum_i \{y_i \log q + (8 - y_i) \log(1 - q)\} + (\text{定数})$$

これを最大化する q が最尤推定値 $\hat{q}$
微分してゼロとおいて計算すると次に

$$\hat{q} = \frac{\text{合計生存種子数}}{\text{合計調査種子数}} = \frac{73}{8 \times 20} = 0.45625$$

生存確率は0.46ぐらい

8.2 ふらふら試行錯誤＝数値計算

先程やったのは解析的な解法

同じことを試行錯誤による最尤推定で

でたらめな初期値から出発、尤度が高まるようにずらす

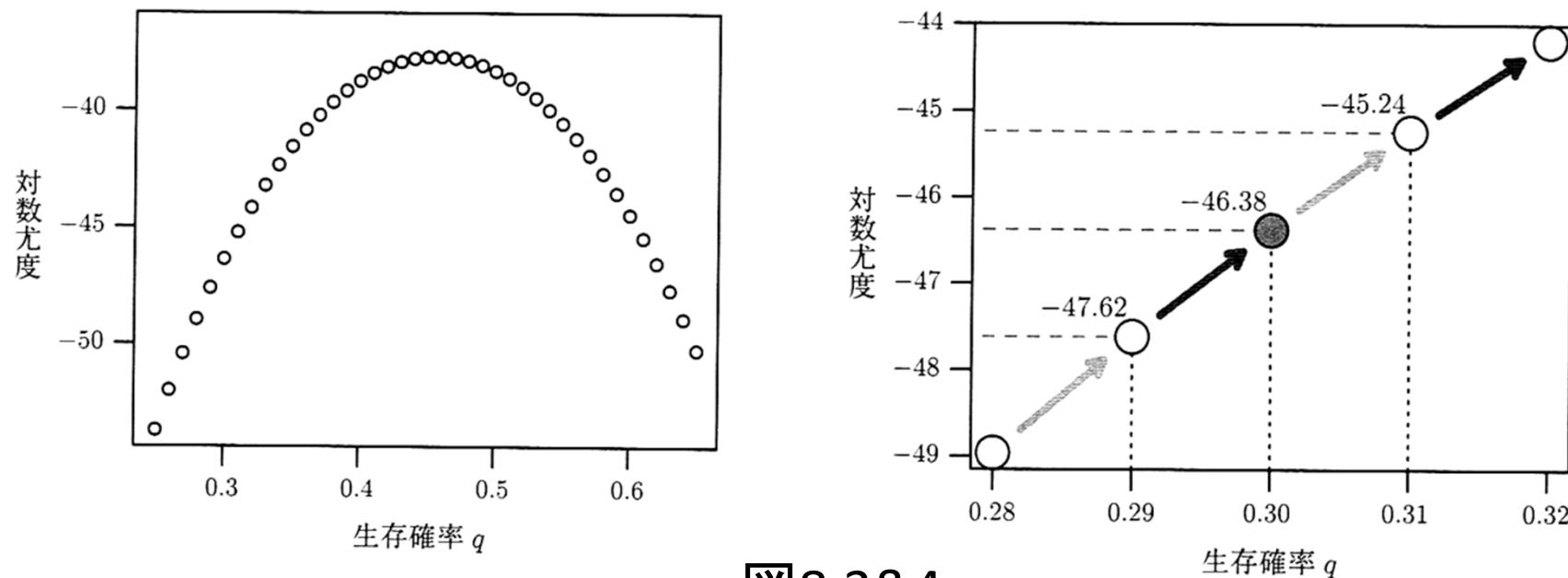


図8.3&4

8.2 ふらふら試行錯誤＝数値計算

先程やったのは解析的な解法

同じことを試行錯誤による最尤推定で

でたらめな初期値から出発、尤度が高まるようにずらす

尤度が最大化する地点へ
＝最大尤度・最尤推定値
が 求められた！

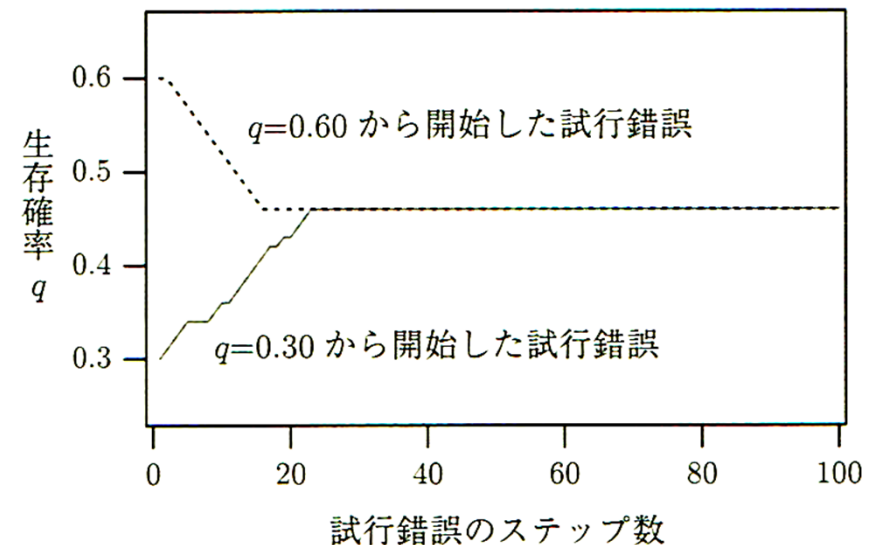


図8.5

8.2 ふらふら試行錯誤 = 数値計算

1. パラメータ q の初期値を選ぶ
2. ランダムにずらして $q^{\text{新}}$ にする
3. 尤度が大きくなったら、 q を $q^{\text{新}}$ に更新
 当てはまりが良くなったら
 2に戻る

8.3 メトロポリス法(MCMC法の一つ)

1. パラメータ q の初期値を選ぶ
2. ランダムにずらして $q^{\text{新}}$ にする
3. 尤度が大きくなったら、 q を $q^{\text{新}}$ に更新
当てはまりが良くなったら
4. 尤度が小さくなっても、確率 r で $q^{\text{新}}$ に更新

当てはまりが悪くなっても

ただし、悪すぎると更新されにくい

2に戻る

$$r = \frac{L(q^{\text{新}})}{L(q)}$$

8.3.1 メトロポリス法でサンプリング

この辺りからだんだん騙された気になってくる

ふらふら試行錯誤と同じようにする
尤度の山を登りながらも、行きつ戻りつ
尤度比=確率 r で戻る

でたらめなんだけど、前の状態に依存：これがMarkov chainと言われる所以

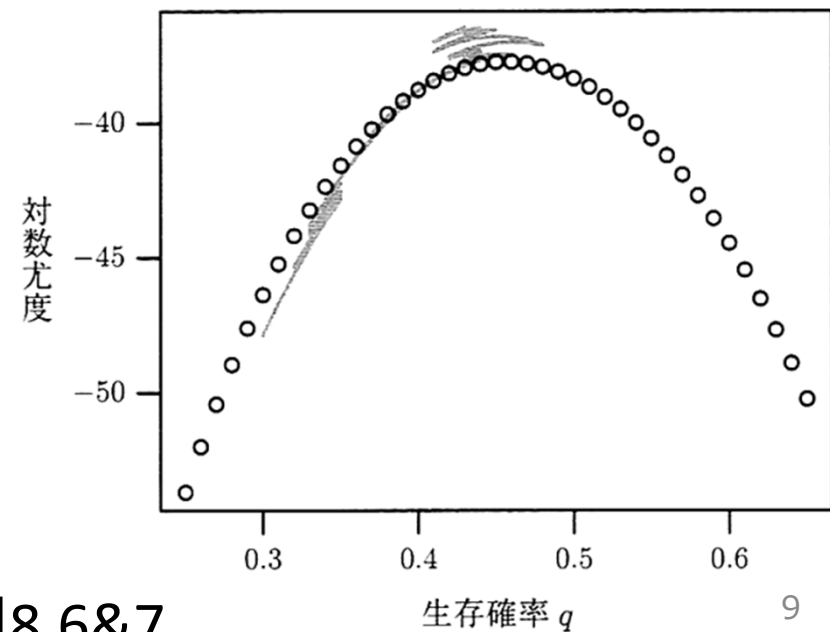
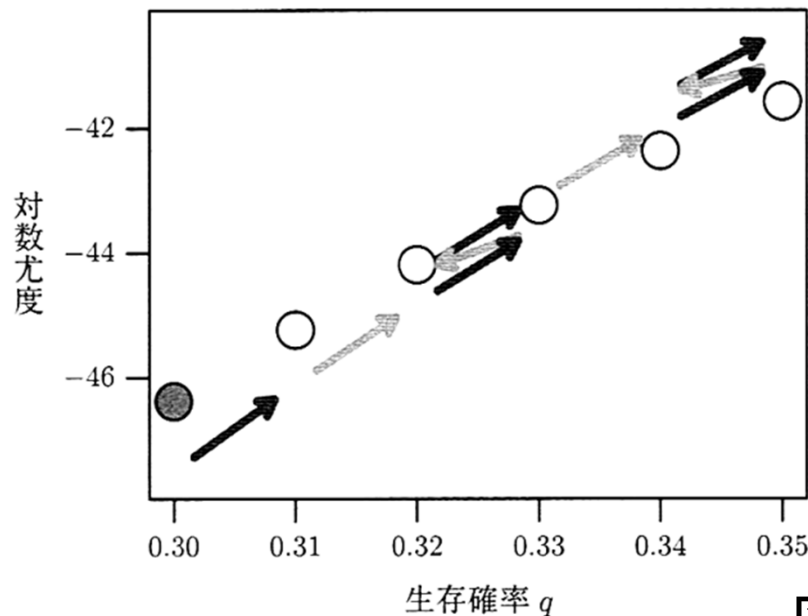


図8.6&7

もっとステップ数を増やしてみたらどうなるか？

徐々に定常状態に落ち着く

- ・ 最初の頃
- ・ 少ステップ数 ダメ

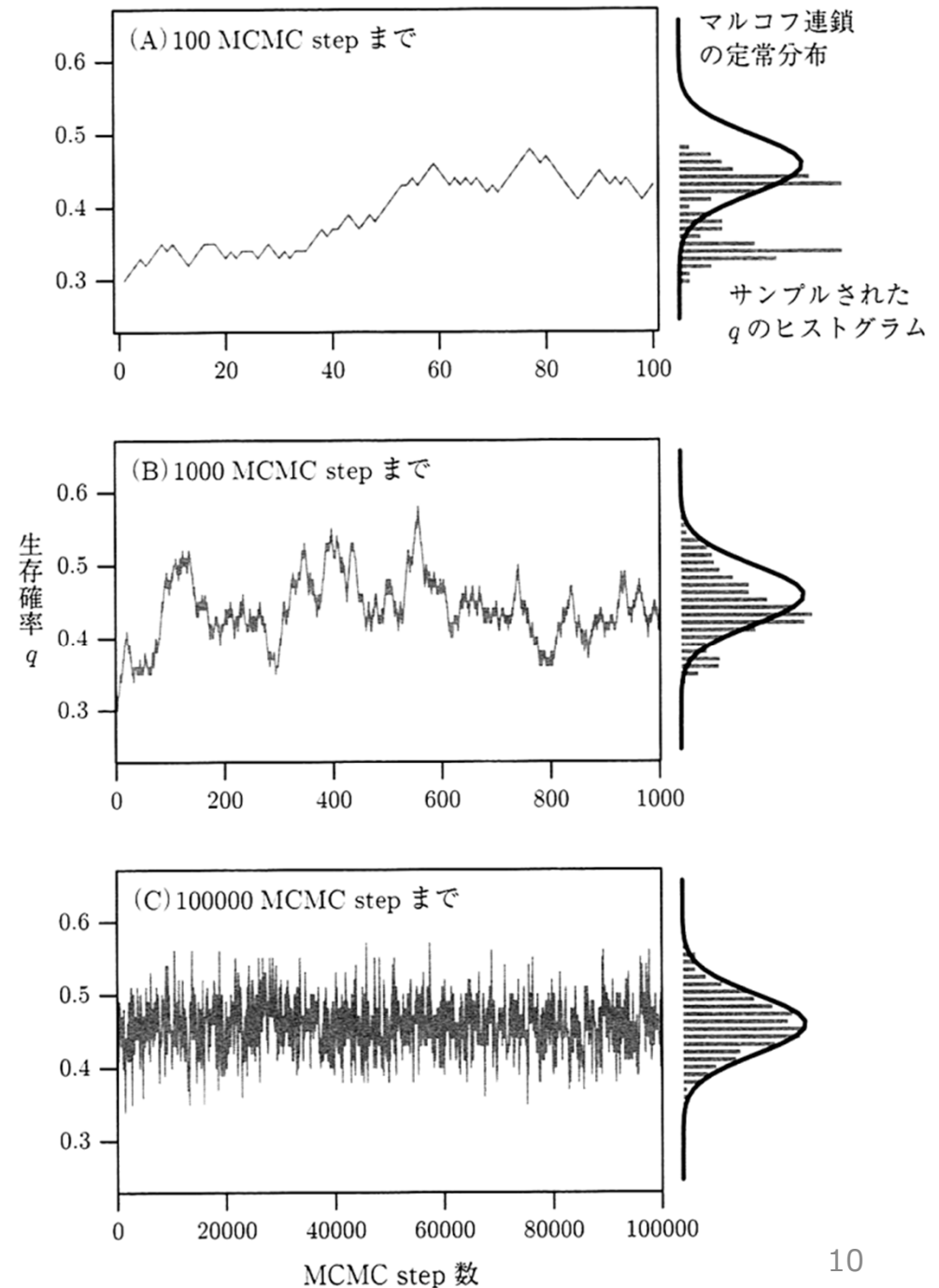


図8.8