志田祐一郎＠野生生物総合研究所

8.隠れデータモデル

ベイズ解析の魅力的な特性は階層モデルが支障なく扱われることである。

階層モデルは深く構造化されており、複数のレベルの確率性を有している。確率的なモデリングは、それらの分布に影響するパラメータを記述するために、固定分布からの抽出によってデータの記述を越えて広がる。従来のモデリングは time-indexed なパラメータ θ_1, θ_2,･･･,θ_T のセットを推定するためにデータを使用するだろう。一方階層モデリングは θ_T 間の確率的な構造を許容する。$\theta_t = a + bt$ のような、決定論的なパラメータ間の結合を強要されることはない。まるでそれらが無関係な量であるかのように θ_t を推定する必要はなく、またパラメータ間の変動や科学的に意味のないサンプリングによる変動の識別に苦心しながら推定量 $\hat{\theta}$ 間のパターンを明らかにしようと努力する必要はない。階層モデルはパラメータ間の確率的な構造を仮定しーたとえば $\theta_t = a + bt + \varepsilon_t$ ここで ε_t は互いに独立で同一の分布 $N(0, \sigma^2)$ に従うランダム変数としてー、その構造を観察データから直接調べることを許容する。

階層モデルは様々な状況から発生し多くの使用法がある。回帰モデルに多くの邪魔なパラメータが存在するときの節約法の導入－分割デザインーや、応答変数が完全に観察されていないときに、潜在的な変数を記述する際に使用されうる。

この章では、理想的なデータセット（“持ちたいデータ”から構成される；このフレーズは Draper,1995 からの借用）の記述から始まる隠れデータモデル―階層モデル―を調べる。持ちたいデータが実際に観察されたデータに縮退するプロセスの記述を始める。焦点は構造の特定に際しての計算論的及び推論的利点である。完全データ尤度（CDL）と頻度論的解析でよく使われる観察データ尤度（ODL）との比較から始める。CDL は持ちたいデータに基づくもので、ODL よりも特定しやすく伝統的な ODL と同じ推論が得られる。データ拡大は、特異な ODL と一致する、充実したコンピュータ的利点を有する CDL の特定のプロセスである。これらは 8.1 節で導入される。

次に、犯罪調査の無作為化回答データセットを考える（8.2 節）。これらのデータは生態学データに共通するいくつかの特性を持っている。ODL よりも CDL を使用したほうが説明が簡単になる。8.3～8.5 節では、占有モデル、距離サンプリングモデル、有限母集団サンプリングモデルについて述べる。これらは通常は完全に区別できるが、類似した構造を共有しているように見えるだろう。それぞれのモデルは縮退したあるいは隠れたデータを含んでおり、モデリングは持ちたいデータの記述から始まる。この場合、CDL に基づくベイズ解析は簡単なことである。

8.1 完全データ尤度

一般的に言って、統計的推測は、未知のパラメータ θ に関するデータ Y の確率を記述して、$[Y|\theta]$ のデータ分布を特定することから始まる。データ分布とみなされる、$[Y|\theta]$ は変数 Y の関数で θ は固定量である。データが与えられたとき、我々は $[Y|\theta]$ を θ が固定された Y に伴って変化する関数として扱うことを望むかもしれない。我々はそれぞれの θ に対してのデータによって与えられる相対的なサポートを評価するために、異なった θ に対して $[Y|\theta]$ の値を比較する。この内容は、以下のような式で書ける。

$$L(\theta|Y) = k[Y|\theta]$$

ここで k は任意の定数、$L(\theta|\mathrm{Y})$ は尤度関数と呼ぶ（3.1 節参照）

目に見えるデータは持ちたいデータの一部というのは良くあるケースである。ここで、Y を $\mathrm{Y}=\{\mathrm{Y}^{obs}, \mathrm{Y}^{mis}\}$ －obs は観察される、mis は観察されないという意味－に分割できる。多くの古典的な方法では、ODL の基本的な原理として、データの観察された一部による完全データ Y のモデルの特定から始める。この過程は、観察されたデータの分布を得ることで観察されないデータの可能値の統合あるいは要約を含む（**式 8.1**）。結果的に得られる尤度 $L(\theta|\mathrm{Y}^{obs})=k[\mathrm{Y}^{obs}|\theta]$ は ODL である。

別のアプローチは、パラメータ θ を扱うように $[\mathrm{Y}^{obs}, \mathrm{Y}^{mis}|\theta]$ の中で観察されない Y^{mis} を正確に扱うことである。ここで、Y^{obs} を固定値として扱い様々な Y^{mis} と θ の値に対する対応を比較できる。ゆえに CDL を $L(\theta, \mathrm{Y}^{mis}|\mathrm{Y}^{obs}) \propto [\mathrm{Y}^{obs}, \mathrm{Y}^{mis}|\theta]$ と定義できる。ODL ではなく CDL を使用するのに様々な利点がある。第 1 に、ODL は CDL より非常に複雑なパラメータの関数となりがちである。第 2 に、CDL の使用は観察されない量の予測のための自然な枠組みを提供する；CDL の観点での欠測値の記載もまた予測のために必要とされる仮定をしばしば明らかにする。CDL は時には頻度論者の分析にも使用される（例えば EM アルゴリズム、Dempster et al., 1977）。しかしベイズの理論的枠組みの元でのほうがより自然で楽に扱える、すべての観察されない量—パラメータ、予測、非観察値—が同等に扱われる。

例示の前に、“データ拡大”—ベイズ法の使用者が直面するであろう、CDL に密接に関係する概念—について言及する。データ拡大は、観察されたデータによるモデルを観察されない構造という観点で拡張することから成り、通常はコンピュータ計算による効果を生み出すという目的を備える。拡大されたモデルは観察値からのデータ分布と観察されない構造を含み、式 8.1 のように、拡大された変数で積分することで ODL を生成する。データ論証の例は 8.4 節にある。

8.2 無作為化された回答データ

ODL と CDL の違いのよい説明のために、無作為化回答調査からのデータを用いる。これは、調査された個人に関する特別な情報を要求することなく不名誉な行動の母比率を推定することに使用されてきた。

このような調査は Warner(1965)によって最初に記述された。彼はこのように始めた。

「謙遜、頑固に思われることへの恐れ、ただ単に他人に秘密を打ち明ける不本意さなどの理由で、多くの個人はインタビュアーから与えられた確かな質問にまともに答えないようにする。・・・調査されることを完全に拒絶したり、調査は承認するが意図的に誤った回答を提供したり・・・人々が回避しがちな質問は過度の暴露を要求する質問である。」

Warner は賢い解決手法を提供している。彼の技術はアルコールの使用や中毒、喫煙、中絶、学術的不正行為、常習的犯罪、性癖、規制の遵守（税の支払いから釣りの規制まで）に適用されてきている。Brewer(1981)は、オーストラリア統計局によるキャンベラにおけるマリファナ使用調査の際に、この手法を使用したことを報告している。

我々は、仲間たちが“行動 X[1]”に関与したことがあるかどうか、データを集めた。

我々は関心ある疑問を直接聞くことはできない、なぜなら仲間たちは行動 X を告白するのは渋るかもしれないからだ。代わりに参加者は、結果を我々に知らせることなしに、最初にコインを投げることを求められた。彼らが行動 X をと

[1] 意味は著者の想像に任せる

ったらあるいはコインの表がでたらYesと応答することが求められた。コインの裏が出てかつ行動Xをとらなかった人はNoと応答することが求められた。

この調査デザインでは、Yesと答えた人の理由が行動Xをとったためかコインをはじいて表が得られたためかわからないことを保証するために、データは意図的にコイントスの結果と交絡されている。この交絡にもかかわらず、交絡を二項分布でモデリングすることで、我々は集団の中及び研究グループの中の行動Xのとった人の割合を推測できる。

可能な結果の2×2元表を考えることで行動Xの確率がどのように推測されるかを見ることはたやすい。pは行動Xの確率、πは無作為化機構と関連する確率を示し、これらの事象は独立としたとき、4つの結果の確率はTABLE 8.1のように計算される[2]。無作為化回答データのなかで、4つの結果のうちの3つ（表がでて行動Xをとる、裏がでて行動Xをとる、表がでて行動Xをとらない）はすべてYesという回答を包含する（影つきセル）。Yesと答える人の確率は$\theta=\pi+(1-\pi)p$;Noと答える人の確率は$\theta=(1-\pi)(1-p)$である（影なしセル）。θをpについて解くことができ、$p=(\theta-\pi)(1-\pi)$が得られる。xをn個体のひとつの標本でYesと応答する確率とすると、θの最尤推定量(MLE)は$\hat{\theta}=x/n$である。ゆえに$\hat{\theta}\geq\pi$が与えられたら、MLEの不偏特性からpのMLEは以下の式となる。

$$\theta \geq \pi のとき \hat{p}=\frac{\hat{\theta}-\pi}{1-\pi}=\frac{x-n\pi}{n(1-\pi)}$$

もし$\hat{\theta}\leq\pi$なら、すなわち、無作為化事象によるものより小さい正の応答が存在すれば、pのMLEは0となる[3]。

モデル化の視点から、どのようにデータをシミュレートするのかを考えることは考えることは有益だ。第一に、個人 i に指示変数D_iを使って、それぞれの対象に行動 X の状態を割り当てる必要があるだろう。次の段階では、表という結果に指示変数C_iを使って、個々のコイン投げの結果がシミュレートされるだろう。無作為化回答のためD_iとC_iは共に潜在変数であり、潜在的に観測できないことを意味する。しかし我々が観測できるものは、人 i がコインは裏が出て行動Xをとらないならば0と等しくなる変数Y_iである。そうでなければ、Y_iは1となる。Y_iはD_iとC_iの関数（式8.2;$Y_i=1-(1-D_i)(1-C_i)$）と書ける。

n個体が大集団からのランダムサンプルとみなせれば、D_iの値をシミュレートする合理的な方法は、成功確率pのベルヌーイランダム変数を発生させることである。シミュレートされたD_iの値が得られたら、次は成功確率πのベルヌーイランダム変数からC_iの値を取り出す。観察された値Y_iはD_iとC_iから式8.2のように構成される。

この過程を、有向非巡回グラフ（DAG）を用いてFIGURE 8.1に表現した。DAGは、階層モデルの中での量間の関係を表現する、直感的に理解できる手法である。1本線の矢印は、確率論的な従属（e.g. 成功確率pのベルヌーイ試行D）を意味し、2本線の矢印は決定論的な関係（e.g. Yは直接DとCから計算される）を示す。DとCの独立は分割された確率的従属性の欠如によって示される。我々のモデルでは、pとDとの関係はその逆のケースよりも$[D|p]$という観点から記述されるだろうから、DAGは有向;$p\rightarrow D$と描かれる。DAGは非環式である; X→Y→Z→Xという形のそれ自身で回り戻るモデルの特定を避け、意味のないモデルの記述を終わりにしうる[4]。

この特殊な問題では、変数CとDはYを通して部分的に観察される。$Y_i=0$の場合、$C_i=D_i=0$である。上付きの

---

2 コインは公平に投げられるものとし、ゆえに$\pi=1/2$。無作為化は別のやり方でも使用でき、πは別の値をとりうる。重要な性質はπが既知だということ。

3 MLEは切り捨てのために偏っている。Warner(1965)は推定量を不偏として記述しているが、結果は負の推定量の容認を必要としている。

4 例えば$[X|Y]=N(Y,1)$ & $[Y|X]=N(X,1)$は合理的なモデル特定かもしれない。しかし$[X,Y]$の同時分布はなく、これらの条件付分布が与えられる。

mis と obs を使用して、CDL は次の式のように書ける。

$$L(p, C^{obs}, D^{mis}|Y) \propto [\,Y, C, D|p] = \prod_{i=1}^{n} I(Y_i, C_i, D_i)\times[C][D|p] = \prod_{i=1}^{n} I(Y_i, C_i, D_i)\times\pi_i^{Ci}(1-\pi)^{1-Ci}\times p_i^{Di}(1-p)^{1-Di}$$

ここで $I(Y_i, C_i, D_i)$は式 8.2 で与えられた関係を定義するための指標である。ゆえに、もし(1)$Y_i=1$ で$\{C_i, D_i\}\neq\{0,0\}$あるいは(2) $Y_i=0$、$C_i=0$、 $D_i=0$ であるなら $I(Y_i, C_i, D_i)=1$ であり、そうでないなら $I(Y_i, C_i, D_i)=0$ である。この指示関数の役割は、データが C_i と D_i の許容値に課す制約を実行することだ。この制約は CDL において観察された値 Y が p に関する情報を規定する役割を定義する。

　　　ベイズ推測の問題として、すべての不明量の事後分布が要求される。これは不明なパラメータの同時事前分布を乗じた CDL に比例する。p の事前分布としてベータ分布を使用すると、**式 8.3** のようになる。

8.2.1 事後分布の計算

p の推測のために、多くの方向に進むことができる。ひとつのアプローチは、密度[p|Y]の明確な記述を発見しようと努めることである。形式上、式 8.3 中の潜在変数 C_i と D_i の可能値を統合する。TABLE 8.1 に基づく初めの ODL の計算で得られるものと結果が同じで、次いで p の事前分布と乗算されることを示す。

より簡単な道を進もう。値 Yi は変換可能（条件付独立）なベルヌーイ試行である。TABLE 8.1 で分析したように $\theta\equiv Pr(Yi=1)=\pi+(1-\pi)p$ であり、ゆえに $X=\Sigma_{i=1}^{n} Y_i$ は指数 n で成功率 θ の二項ランダム変数である。ゆえに ODL は $L(p|Y)\propto\theta^X(1-\theta)^{(n-X)}$である。$p$ の事前分布と乗算されたのが**式 8.4** である。

p の事後分布密度は任意の α と β による閉じた形では存在しない。R などのソフトウェアを用いて、要求された規格化定数を発見することは可能だが、より簡単な解法が望ましい。これまで 2 つの事前分布から得られた密度を FIGURE 8.2 に示す。行動 X のデータ分析は MCMC を用いてより楽に行うことができる。3 つの異なった方法を提示する。

・BUGS を用いた CDL に基づく MCMC

事後分布の要約のために、PANEL 8.1 のコードを使って BUGS のモデルに適合させることは容易だ。このコードの興味深い特性は、Y_i をパラメータ $\alpha[i]=1-(1-D[i])(1-C[i])$をもつベルヌーイランダム変数としてモデル化したことだ。もし Y_i が FIGURE 8.1 に示されているような C_i と D_i の決定論的関数として与えられるならば、このコーディングは多分直観に反したものだろう。いつも 0 か 1 となる $\alpha[i]$が与えられたとき、$Y\equiv\alpha[i]$だから、コーディングは不必要に複雑に見えるかもしれない。しかし、BUG コードによる確率論的な Y_i の特定は、尤度が指示変数 $I(Y_i, C_i, D_i)$を通して観察された変数 Y_i の関数であることを保証する[5]。

BUGS プログラムと PANEL 8.1 のコードを用いて、サイズ 100 万のマルコフチェイン事後分布サンプルを発生させた。初めの 1 万の値は焼き捨て、p が 0.58～0.98 の間に入る 95％の機会があると結論付けた。この結論は Be(1,1)の事前分布を使用して得られた。ジェフリーの事前分布 Be(1/2,1/2)を使用した場合、その範囲は 0.62～0.99 である。これらの

---

5 Y[i] <- 1-(1-D[i]) * (1-C[i])と書くと、Y[i]が 2 回定義されるため、エラーメッセージが出るだろう。1 つがデータの記述、もう 1 つが独立的に生成された D[i]と C[i]の関数になっている。

結論は、FIGURE 8.2 に要約された解析的計算と一致する。

FIGURE 8.2 の結果のとおり、事前分布の選択に幾分敏感だ。ここで観察値 X＝22 は p 値が 1 に近いことを支持すると解釈される。許容範囲の境界付近に存在するパラメータの推定は難しい統計的な問題だ。頻度論的な推定では、漸近的な結果を適用するため、境界付近のパラメータ推定には非常に大きなサンプルサイズが必要とされる。

二項分布の成功率 $\theta=\pi+(1-\pi)p$ の推定に対して、標準正規分布に基づく漸近的な結果は、$n\theta=\mathrm{E}[X]$と $n(1-\theta)=\mathrm{E}[n-X]$がともに 5 を超えていれば信頼できる。しかし $n-X=1$ ということは $n(1-\theta)$が十分に 5 よりも小さいことを示唆する。ゆえに漸近的な結果は信頼できない。実際、もし通常の漸近的な CI を計算するなら、1 の境界を越えて拡張されるような満足できない(0.87, 1.04)という 95％信用区間を得る。代替の頻度論者の手続きでは、(0.78, 1.00)という p の(0.56, 1.00)につながる“正確な”95％CI が得られる。しかしこの CI は、そのカバーする比率は真の p の値に依存して 95％を超えうるという意味で、十分に保守的ではない[6]。

これらのデータのベイズ解析は漸近的な近似を当てにしない。しかし、ベイズ推測を行うためには、選択により敏感に応答する小さなサンプルサイズのために、p の事前分布を特定しなければならない。

CDL を使用したモデリングの有用な特性は事後予測が直接的であることだ。例えば、Yes と答えた人が行動 X をとったかを、対応する値 D_i を監視することで予測できる。同様に、特別な関心がなくてもコイントスの結果を予測できる。

p に Be(1,1)を使用したとき、D_i の 100 万の値のうち 909348 が 1 であった。個人 1 が行動 X をとる事後確率は、事前分布のもとで 0.91 である。共変量がないため、事後確率は無作為化実験の中では Yes と応答するどんな人にとっても同じである。同様に、C_i の 100 万の値のうち 545293 が 1 であり、Yes と応答して個々が表に返るとなる事後確率は 0.55 である。ジェフリーの事前分布のときは、これらの確率は 0.93 と 0.54 にわずかに変化する。

パラメータ p は、23 人がサンプリングされた仮想的な有限母集団を表す。あるいは、行動 X の事前分布として解釈してもよく、データが集められたとき、その値は不在であった 24 番目の人を表すために使用されるだろう。この値自体は、行動 X をとった 23 人の割合とは区別される。後者の量 $\Sigma_{i=1}^{23} D_i/23$ は p の推定量として使用されるが、それ自身は p ではない。しかし、23 人中行動 X をとった人の数に我々は興味を持つだろう。CDL を使用することで、$\Sigma_{i=1}^{23} D_i$ の推測が自明で簡単にできるだろう。PANEL 8.1 の BUGS コードでは、合計から生成されたパラメータを単純に定義する。我々は更なる努力なしに単純な事後分布を得る。事後分布は FIGURE8.3 に図示され、特に、98％の信頼度で少なくとも 23 のうち 15 が行動 X をとった。しかし、23 の事後分布は 0 である；我々の実験デザインによって 1 人が確実に行動 X をとっていないと結論づけられた[7]。

・BUGS を用いた ODL に基づく MCMC

前述の解析は CDL に基づくものであった。特別な問題については、ODL に基づいた解析も採用されうる。PANEL 8.2 のように BUGS でコードされる。gamma[1]と gamma[2]が、Yes と答えたときに行動 X をとった確率(γ_1)そして Yes と

---

6 正確な頻度論者の CI の詳細は 3.2 節参照。

7 それは、No と報告する個人が正直であるという条件での話である。行動 X のオッズは無作為化実験の正の応答によっては倍増する；参加者は、このオッズの変化を直感あるいは計算し、不正直な No という回答に傾くかもしれない。あるいは、サンプリングした集団を考えると、負の応答と関係する兆候があるかもしれない。両方の応答は参加者にコインを 2 度投げるよう依頼し、表が 2 回でたら Yes、裏が 2 階出たら No、そうでなければ正直に回答してもらうことで無作為化されるだろう。

答えたときにコインの表が出た確率(γ_2)の事後分布を計算するために使用される。TABLE 8.1 の検査により、$\gamma_1 = p/\theta$ でかつ $\gamma_2 = \pi/\theta$ であることが明らかである。これらは p の関数で、生成されたパラメータである。ゆえに事後分布の標本は容易に p の事後標本から得られる。

我々の ODL に基づく解析は個々の行動 X の性質を示す D_i を含んでいないことに注意せよ。しかし我々は D_i の事後予測分布を計算することで発生する事後標本を活用できる。ゆえに次の式のように $\Pr(D_i = 1|Y_i = 1)$を計算する。

$$\Pr(D_i = d|\, Y_i = 1) = \int_{\gamma 1} \Pr(D_i = d|\, Y_i = 1) \times [\gamma_1|\mathrm{Data}]d\gamma_1 = \int_{\gamma 1} \gamma_1 \times [\gamma_1|\mathrm{Data}]d\gamma_1 = E[\gamma_1|\mathrm{Data}]$$

同様に、$\Pr(C_i = d|\, Y_i = 1) = E[\gamma_2|\mathrm{Data}]$

これらの確率は γ_1 と γ_2 のために生成されたマルコフチェインの平均として計算される。

・BUGS を用いた部分的に統合された CDL に基づく MCMC

別のアプローチとして、モデルにとって邪魔な性状である潜在変数(ここでは C_i)を統合し、部分的に統合した CDL という観点でモデル化することがある。BUGS コードは PANEL 8.3。ここで、行動 X をとったら 1、とらなければ $\pi = 0.5$ を α[i]がとることで、Y_i は確率論的にモデル化される。形式的には、コイントスの起こりうる結果を通して合計することで Pr (Report"Yes"|D)を得ている。

$$\Pr(Y_i = 1|D_i) = \Pr(Y_i = 1|D_i, C_i = 0)\Pr(C_i = 0) + \Pr(Y_i = 1|D_i, C_i = 1)\Pr(C_i = 1)$$

ここで D＝0 ならば π、D＝1 ならば 1 で、ゆえに一般式は $\pi(1 - D) + D$ である。

・ギブスサンプラーのコーディング

これまで行動 X に関する 4 つの分析方法を述べてきた。まず、数値積分により推測に必要な規格化定数を計算することで、p の事後分布を ODL L(p|Y)と事前分布[p]の積と比例するように記述できた。次に、PANEL 1 の BUGS コードを使用して CDL に基づく MCMC の使用を提案した。3 番目と 4 番目の分析は ODL(PANEL 8.2)及び部分的に統合した CDL(PANEL 8.3)に基づく MCMC を使用した。

この項では、CDL を使用したギブスサンプリング構想の定式化に基づくデータ解析の最後のアプローチを提示する。この章に、潜在変数とパラメータの同等性を強調すること、そして読者にギブスサンプラーを理解することの勧めを含める。BUGS の使用はベイズモデルへの適合を容易にするが、MCMC のための洞察力を養うためには自分自身のコードを書くしかない。ここでの解析は良い単純な例で、R のようなプログラムを使って容易に実行できる。

ギブスサンプリングは全条件付分布の循環的サンプリングであることを思い出してほしい。未知の $\theta = (\theta_1, \theta_2 \cdots \theta_k)$ とデータ X が与えられたとき、θ_i の全条件付分布は、すべての j が i ではないときの θ_j の固定値を伴う[X|θ][θ]と比例する。ギブスサンプリングは、全条件付分布からのサンプリングによる連続的な θ_j の値の更新からなる。

行動 X の完全データは Y、C、D という名前の長さ 23 のベクトルからなる。ベクトルの要素は、表が出る指標 C_i、行動をとる指標 D_i、無作為化回答調査で Yes と回答する指標 $Y_i = 1 - (1 - C_i)/(1 - D_i)$ である。Y_i のすべての値は観察され、C_i と D_i は直接的には観察されない。$Y_{23} = 0$ だから $C_{23} = D_{23} = 0$ と演繹できるが、i が 23 でないとき $Y_i = 1$ だから

対応する C_i と D_i の値はわからない。このような C_i と D_i をギブスサンプラーに含めなければならない。

必要な C_i と D_i の全条件付分布は TABLE 8.1 の調査により容易に得られる。C_i 以外の Y、C、D が[$C_i|\cdot$]の計算の必要条件となる。c＝0,1 である次の式を得る。

$$\Pr(C_i = c|D_i = 1, Y_i = 1, \pi) = \pi^c(1-\pi)^{(1-c)} \quad \text{and} \quad \Pr(C_i = c|D_i = 0, Y_i = 1, \pi) = c$$

C_i の全条件付分布は、比率パラメータ $D_i \cdot \pi + (1-D_i)$ として、ベルヌーイ分布とすることで要約される、同様に、全条件付分布[Di|･]は Ci、Yi、p に従属する。ここで、d＝0,1 である次の式を得る。

$$\Pr(Di = d|Ci = 1, Y_i = 1, p) = p^d(1-p)^{(1-d)} \quad \text{and} \quad \Pr(D_i = d|C_i = 0, Y_i = 1, p) = d$$

C_i の全条件付分布は、比率パラメータ $C_i p + (1-C_i)$ として、ベルヌーイ分布とすることで要約される。

最後に、全条件付分布[$p|\cdot$]は D_i の値の集合で決定され、p の事前分布は $Be(\alpha, \beta)$。値 D_i はベルヌーイ試行だから、その合計は二項ランダム変数でベータ事前分布が共役となる。ゆえに全条件付分布[$p|\cdot$]は $Be(\alpha + \Sigma D_i, \beta + n - \Sigma D_i)$ である。ギブスサンプラーは以下のように記述される。

Step1：i＝1～22 において $C_i^{(1)} = 0$、$D_i^{(1)} = 0$ と初期化する。これらの 22 は未知の値で、$Y_i = 1$。初期値の選択は任意だが、$Y_i = 1$ と一致しなければならない。$p^{(1)} \in (0, 1)$を初期化する。上付きの数字は、生成された値のマルコフチェインの中での連続する番号を意味する。

Step2：i＝1～22 について、成功パラメータが $D_i^{(j-1)} + (1-D^{(j-1)})$ のベルヌーイ試行としての $C_i^{(j)}$ を発生させる。

Step3：i＝1～22 について、成功パラメータが $C_i^{(j-1)} + (1-C^{(j-1)})$ のベルヌーイ試行としての $D_i^{(j)}$ を発生させる。

Step4：$p^{(j)} \sim Be(\alpha + \Sigma_{i=1}^{n} D_i^{(j)}, \beta + n - \Sigma_{i=1}^{n} D_i^{(j)})$を発生させる。

Step5：Step2～4 を j＝1,2,…繰り返す。

読者には R か他の言語でのギブスサンプラーのコーディングを薦める。その代わりに、BUGS はシミュレート装置としても使える、分析は PANEL 8.4 のように全条件付分布を使用して実行できる。全条件付分布の特定の後、データの記述はない；観察されたデータ Y は、p の事前分布としての Be（1,1）を組み合わせて全条件付分布の定義に使用した、非明示的なものである。

8.2.2 意見

この例では、応答の交絡は意図的な、あるデータ収集の仕掛けである。コイントスを通して、興味のデータである D_i は Y_i に縮退した。ゆえに、D_i は今潜在的だ。データ縮退機構を思慮深く選択することで、D_i を観察しないにもかかわらず確率 p が推測できる。いろいろな意味で、無作為化回答調査は、階層モデルとして問題を表現しているにもかかわらず推測が促進される、縮退データ問題の基準的な例としての役割を果たす。

生物学的データはしばしば交絡についての類似の要素をもつ。しかし違いは、交絡機構が調査者よりもむしろ自然によって選ばれるフィールドデータを備えていることだ。結果として、データを生みだすプロセスに注意深い考察が必要とされるはずだ。データの縮退につながるプロセスから興味ある潜在変数を発生させる過程を明示的に分離することで、階層モデルはこの考察を促進させる。データ縮退モデルを特定したとき、調査者は潜在変数のモデリングに自由に専念できる。このアプローチが“持ちたいデータをモデリングすること”である。そこでは、興味ある潜在変数という

観点でモデル化を行う。

例えば、共変量ベクトル Z_iー行動 X に関与したかどうかを決定することと関係があるだろうと我々が信じる各人の記録ーがあったとしよう。共変量は年齢、性、音楽の好み、など。もし我々がそのような情報を持っていたら、D_i を Z_i のロジスティック回帰を使用してモデル化するのが自然だろう。D_i は潜在変数ではあるが、D_i がパラメータ p をもつベルヌーイランダム変数とする代わりに、個々の特定パラメータ p_i をもつベルヌーイ変数とするようにモデルを拡張することは容易だ。ここでロジット(p_i)＝$Z'_i\beta$ であり、この中のパラメータベクトル β は回帰係数を含む(PANEL 8.4)。加法モデル構造では PANEL 8.1 の BUGS コードの若干の微修正が要求されるだろう。