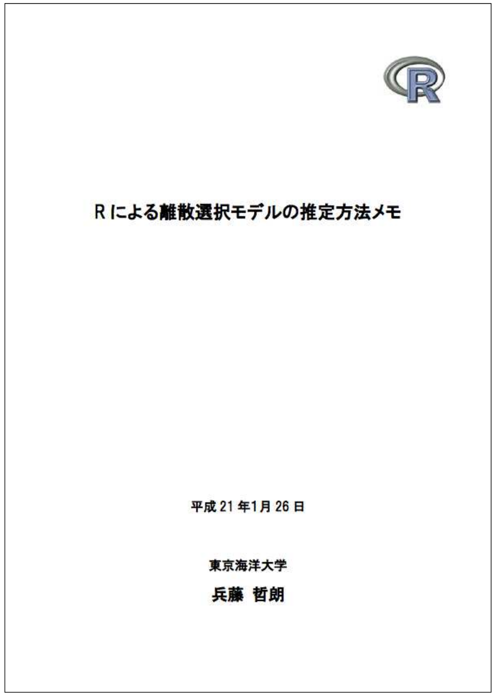
ベイズ統計データ分析＋兵藤先生のメモ
第4章 離散選択モデル


Rによる離散選択モデルの推定方法メモ

平成21年1月26日

東京海洋大学
兵藤 哲朗

http://www2.kaiyodai.ac.jp/~hyodo/Logit_by_R.pdf

2009/6/24(水)
ベイズゼミ#2
M2 北川直樹

目次

1. MCMC法のアルゴリズム
   - ギブズサンプラー
   - メトロポリスヘスティング
2. 離散選択モデル
   - 二項プロビットモデル
     - ギブズサンプラー, MHアルゴリズム
   - 多項プロビットモデル
3. 数値例
   - 最尤推定
   - ベイズ推定

MCMC法のアルゴリズム

- Markov チェイン
  - 3蓮の葉(A,B,C)の上を蛙が飛び移る.
  - 各葉の遷移確率が3*3行列の場合, t→∞の蛙の存在確率を計算する。
  - 等高線は対数尤度関数.
  - より良い解を求めて蓮の葉を飛び回る.
  - 蓮の葉間の飛び方を規定するのが, MCMC 法

QuickTimeý Ç²
êLí£ÉvÉçÉOÉâÉÄ
Ç™Ç±ÇÃÉsÉNÉ`ÉÉÇ¾å©ÇÈÇžÇ½Ç…ÇÕïKóvÇ-ÇÅB

MCMC法のアルゴリズム

- Gibbs Sampler(GS)
  - 各パラメータが与えられたとき，他のパラメータの条件付事後分布が定義できる場合に使われる.
  - 例）平均0,分散1 ,相関係数ρの2 変量正規分布

$$\theta_1|\theta_2 \approx N(\rho\theta_2, 1-\rho^2), \quad \theta_2|\theta_1 \approx N(\rho\theta_1, 1-\rho^2)$$

  - 2 つの正規乱数を入れ子で発生させれば，得るべき2 変量正規分布に従う.
  - 蛙の例で言えば，蛙は次に跳ぶべき蓮への望ましい跳躍方向を知っており，それに従いrandomness をもって飛び回る.

MCMC法のアルゴリズム

- Metropolis-Hastings アルゴリズム(MH)
  - 条件付事後分布が書き下せない場合に使われる.
  - 様々なモデルに対応できるが，計算効率が悪い.
  - アルゴリズムの手順
    - 次に飛ぶ蓮をランダムに選ぶ.
    - GS と違い，その蓮に飛んでしまったことを想定し，解の”改善度”を [0,1] の範囲内に収まる基準値で計算.
    - [0,1] の一様乱数を発生させ，両者の大小関係で，この蓮に飛ぶか否かを決断する.
    - 元の蓮に止まる場合，今の計算が無駄になる.

離散選択モデル

- 二項プロビット
  - 個人i，選択肢集合y_i={1,0}，説明変数x_i
    パラメータβ，誤差項ε_i

$$y_i = \begin{cases} 1, & if \;\; x_i\beta + \varepsilon_i > 0 \\ 0, & otherwise \end{cases}$$

  - 条件付き事後分布は，

$$P(y_i = 1 | x_i, \beta) = \Phi(x_i\beta)$$
$$P(y_i = 0 | x_i, \beta) = 1 - \Phi(x_i\beta)$$

  - Φの関数形
    - 正規分布：プロビット
    - ガンベル分布：ロジット

離散選択モデル

- 二項プロビット
  - Φの数値積分が困難
  - 潜在変数y*を用いてギブズサンプラーを適用した事後分布推定が可能

$$y_i^*|\beta \sim \begin{cases} IN_{(0,\infty)}(x_i\beta,1), & if\ \ y_i = 1 \\ IN_{(-\infty,0]}(x_i\beta,1), & if\ \ y_i = 0 \end{cases}$$

  - $IN_{[a,b]}(\mu,\sigma^2)$は，区間[a,b]で切断させた正規分布
  - βの事前分布をβ～N(b_0,B_0)とすると，y*が与えられたときの条件付き事後分布は，

$$\beta|y^* \sim N(b_1,B_1)$$

  - ただし，$b_1 = B_1(B_0^{-1}b_0 + X^T y^*),\quad B_1^{-1} = B_0^{-1} + X^T X,\quad X = (x_1,\cdots x_n)^T$

二項プロビットモデル

- ギブズ・サンプラー
  1. 繰り返し回数を $s=0,\cdots,ndraw$ とする.
  2. $s=0$ とし，初期値である事前分布 $\beta^{s=0}\approx N(b_0,B_0)$ を決定
  3. β^s から β^{s+1} を生成
     1. $y_i^{*s+1}\big|\beta^s$ を生成する

$$y_i^*\big|\beta\approx\begin{cases}N_{(0,\infty)}(x_i\beta,1), & if\ \ y_i=1\\ N_{(-\infty,0]}(x_i\beta,1), & if\ \ y_i=0\end{cases}$$

     1. $\beta^{s+1}\big|y_i^{*s+1}$ を生成する

$$\beta\big|y^*\approx N(b_1,B_1)$$

ただし，$b_1=B_1(B_0^{-1}b_0+X^Ty^*),\quad B_1^{-1}=B_0^{-1}+X^TX,\quad X=(x_1,\cdots x_n)^T$

  1. $s+1<ndraw$ のときに3.に戻る. $s=ndraw$ のとき終了

二項プロビットモデル

- MHアルゴリズム
  - $\beta^{s=0} \sim N(b_0, B_0)$ を決定した後，提案分布の密度関数を与えて $q(\beta^s, \beta|y)$ 受容確率 を計算し $\beta^s, \beta|y$ サンプリングする.
  - 例えば，最尤法によるパラメータ推定値 $\hat{\beta}$ と共分散 $\hat{\Sigma}$ を使い，提案密度分布を $\beta|y \sim N(\beta^{s-1}, \tau\hat{\Sigma})$ とすると，次の受容確率を計算する.

$$\rho(\beta^{s-1}, \beta|y) = min\left(\frac{p(\beta|y)}{p(\beta^{s-1}|y)} \frac{q(\beta, \beta^{s-1}|y)}{q(\beta^{s-1}, \beta|y)}, 1 \right)$$

二項プロビットモデル

- MHアルゴリズム
  1. 繰り返し回数を $s=1,\cdots,ndraw$ とする.
  2. $s=1$ とし，初期値である事前分布 $\beta^{s-1} \approx q(b_0, B_0)$ を決定
  3. β^{s-1} から $\beta \approx q(\beta^{s-1}, \beta | y)$ を生成
  4. 受容確率 $\rho(\beta^{s-1}, \beta | y)$ を生成する

$$\rho(\beta^{s-1}, \beta | y) = min\left(\frac{p(\beta|y)}{p(\beta^{s-1}|y)} \frac{q(\beta, \beta^{s-1}|y)}{q(\beta^{s-1}, \beta|y)}, 1 \right)$$

  1. 一様乱数 $u \approx u(0,1)$ を生成する

$$\beta^{s} = \begin{cases} \beta & if \ \ u \leq \rho(\beta^{s-1}, \beta | y) \\ \beta^{s-1} & otherwise \end{cases}$$

  2. $s+1 < ndraw$ のときに3.に戻る. $s = ndraw$ のとき終了

多項プロビットモデル

- 個人iがp個の選択肢集合から選択効用u_{ij}が最大となる選択肢jを選択する問題.
- p個の選択肢集合に対して，j=0,･･p-1とラベル付けし，p-1個の潜在効用wを考える.

$$w_i = X_i^d \beta + \varepsilon_i$$

$$w_{ij} = u_{ij} - u_{ip},\ \varepsilon_{ij} = \xi_{ij} - \xi_{ip},\ \varepsilon_i \sim N(0,\Sigma)$$

$$X_i^d = \begin{pmatrix} x_{i1}^T - x_{ip}^T \\ \vdots \\ x_{ip-1}^T - x_{ip}^T \end{pmatrix}$$

- $w_{ij}<0$のとき，個人iは選択肢pを選択する.
- w_{ij}の中で1つでも正があれば，w_{ij}が最大の選択肢jを選択

多項プロビットモデル

- 多項プロビットを推定する場合，(β,Σ)の事前分布は，

$$\beta \sim N(b_0, B_0), \quad \Sigma \sim IW(v_0, R_0)$$

- $IW(v_0, R_0)$は 逆ウィシャート分布である．
- 事後分布は，

$$\beta | \Sigma, w \sim N(b_1, B_1), \quad \Sigma | \beta, w \sim IW(v_1, R_1)$$

- ただし，

$$b_1 = B_1 (B_0^{-1} b_0 + \sum_{i=1}^{n} (X_i^d)^T \Sigma^{-1} w_i), \quad B_1^{-1} = B_0^{-1} + \sum_{i=1}^{n} (X_i^d)^T \Sigma^{-1} X_i^d$$

$$v_1 = v_0 + v, \quad R_1^{-1} = R_0^{-1} + \sum_{i=1}^{n} (w_i - X_i^d \beta)(w_i - X_i^d \beta)^T$$

- vは逆ウィシャート分布の自由度

多項プロビットモデル

- ギブズサンプラーでは，以下の手順で条件付き分布を算出

$$\beta|\Sigma,w \quad and \quad \Sigma|\beta,w \rightarrow w|\beta,\Sigma,y,X^d \rightarrow y|\beta,\Sigma,w,X^d$$

- w_iは(p-1)次元に切断された正規分布であるため，wの条件付き事後分布を直接求めるのは難しい.
- w_{ij}の事後分布を以下のような切断された正規分布とする.

$$w_{ij}|w_{i-j},y_i,\beta,\Sigma \sim$$

$$N(m_{ij},\tau_{jj}^2)\times[I(j=y_i)I(w_{ij}>\max(w_{i-j},0))+I(j\neq y_i)I(w_{ij}<\max(w_{i-j},0))]$$

$$m_{ij}=x_{jj'}^d\beta+(-\sigma_{jj}\gamma_{j,-j})^T(w_{i,-i}-X_{i,-j}^d\beta),\ \ \tau_{jj}^2=1/\sigma_{jj}$$

- $\gamma_{j,-i}$はΣ^{-1}のj番目の要素を除いたj番目の行
- σjjはΣ^{-1}の(j,j)要素

$$\Sigma^{-1}=\begin{pmatrix}\gamma_1'\\ \vdots \\ \gamma_{p-1}'\end{pmatrix}$$

離散選択モデル

- 多項プロビットのギブズ・サンプラー
  - 繰り返し回数を $s=0,\cdots,ndraw$ とする.
  - 適当な初期値 b_0,B_0,v_0,R_0 を与える.
  - $s=0$ とし，事前分布 β^s,Σ^s,w^s を生成
  - w^s から $w^{s+1}\big|\beta^s,\Sigma^s$ を生成
  - β^s から $\beta^{s+1}\big|w^{s+1},\Sigma^s$ を生成
  - Σ^s から $\Sigma^{s+1}\big|\beta^{s+1},w^{s+1}$ を生成
  - $s+1<ndraw$ のときに3.に戻る.
  - $s=ndraw$ のとき終了

パッケージと関数

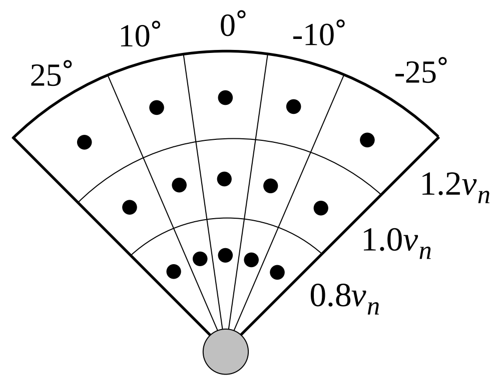
- 二項プロビット
  - ギブズサンプラー
    - MCMCprobit(M)
    - rbprobitGibbs(B)
- 多項プロビット
  - ギブズサンプラー
    - rmnpGibbs(B)

※(M):MCMCpackパッケージ
 (B):bayesmパッケージ

- 二項ロジット
  - MHアルゴリズム
    - MCMClogit(M)
- 多項ロジット
  - MHアルゴリズム
    - MCMCmnl(M)
    - rmnlIndepMetrop(B)

数値例

- 歩行者の速度-角度選択モデル
  - 選択肢集合
    - 速度3分類
    - 角度5分類
    - 合計15分類
  - 説明変数
    - 目的地を目指す
      - 距離，角度
    - 速度を保つ
      - 加速，減速
    - 定数項



$$V_{vd_n} = \beta_{ddist} ddist_{vd_n} + \beta_{ddir} ddir_{d_n} + \beta_{acc} I_{v,acc} v_n + \beta_{dec} I_{v,dec} v_n + \beta_{vd_n}$$

数値例

- 最尤推定
  - 尤度関数

$$L = \prod_{n=1}^{N} \prod_{i=1}^{j} P_n(i|x,\theta)$$

  - 誤差構造
    - 多項ロジットモデル
  - 最適化手法
    - 準ニュートン法

- ベイズ推定
  - ベイズの定理

$$P(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}$$

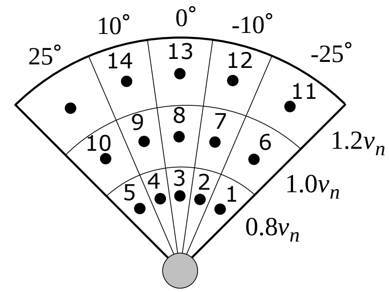
  - 誤差構造
    - 多項ロジットモデル
  - MCMC法
    - MHアルゴリズム
    - 棄却回数:1000回
    - 繰返回数:10000回
  - 初期分布
    - 平均0，分散0

推定結果

	最尤推定		ベイズ			最尤推定		ベイズ	
説明変数	推定値	T値	平均	平均/分散	説明変数	推定値	T値	平均	平均/分散
β_{ddist}	-5.821	-7.05^{**}	-6.968	-8.11	β_6	0.054	0.17	-2.622	-11.85
β_{ddir}	-0.058	-0.20	-0.292	-1.03	β_7	1.544	5.40^{**}	-1.051	-6.17
β_{acc}	0.181	1.71	0.290	2.64	β_8	2.713	9.66^{**}	0.134	0.84
β_{dec}	-1.264	-10.65^{**}	-1.998	-15.27	β_9	1.623	5.70^{**}	-0.968	-5.70
β_1	0.570	1.86	-2.139	-10.36	β_{10}	0.245	0.79	-2.368	-11.15
β_2	1.492	5.12^{**}	-1.139	-6.39	β_{11}	1.352	4.73^{**}	-0.692	-4.08
β_3	2.300	8.04^{**}	-0.315	-1.90	β_{12}	2.263	8.44^{**}	0.318	2.44
β_4	1.292	4.41^{**}	-1.362	-7.58	β_{13}	3.317	12.14^{**}	1.277	11.07
β_5	0.686	2.25^{*}	-1.967	-11.57	β_{14}	2.456	9.22	0.525	4.07
	サンプル	3085	サンプル	3085		初期尤度	-8354	受容率ρ_0	1.000
	適合度	0.132				最終尤度	-7251	受容率ρ	0.641

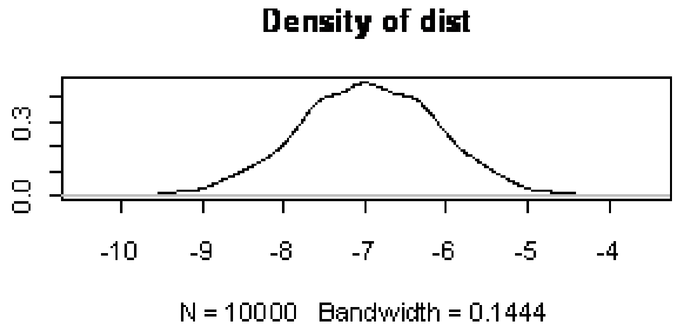
- 赤は，推定値⇔平均値，t値⇔平均/分散，ベイズの絶対値が大きい．
- 青は，推定値-2≒平均値，t値-10≒平均/分散ベイズの値が小さい．

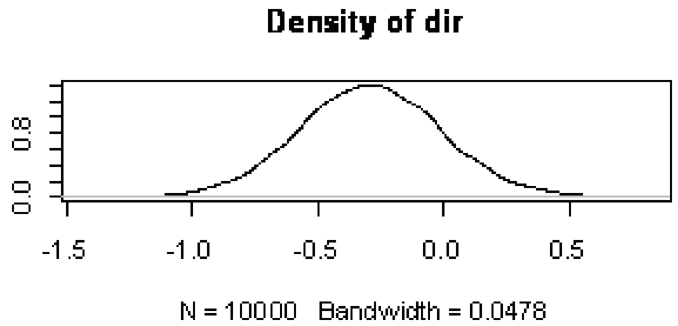
$$V_{vd_n} = \beta_{ddist} ddist_{vd_n} + \beta_{ddir} ddir_{d_n} + \beta_{acc} I_{v,acc} v_n + \beta_{dec} I_{v,dec} v_n + \beta_{vd_n}$$



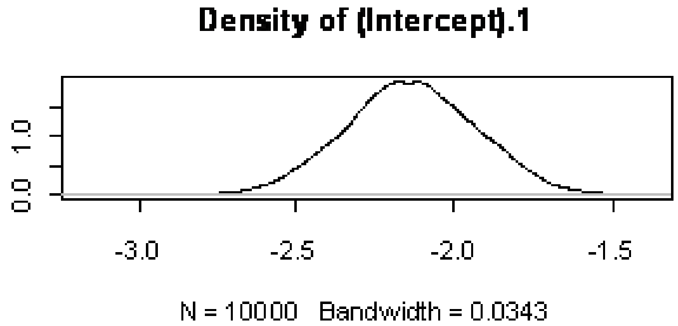
分布図

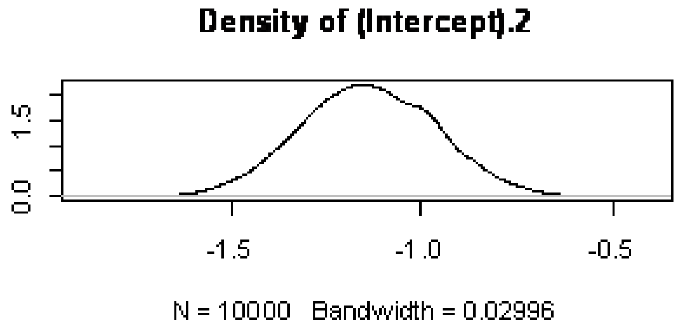
- 赤の変数





- 青の変数





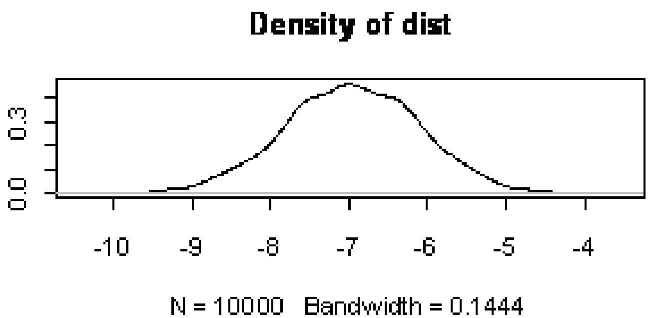
推定結果

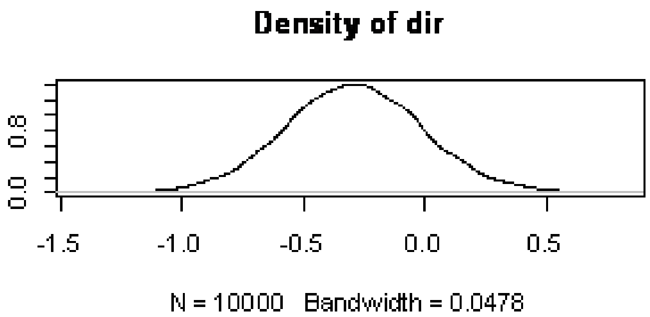
- 最尤推定法との比較
  - MCMC 法から得られたパラメータの分散共分散と，ヘッセ行列の逆行列から得られた分散共分散の一致性は証明できるのであろうか．
- 解の安定性
  - 概ね，正規分布に漸近していることが理解できるが，必ずしも単峰性が満たされているわけでもない．これは何度も乱数初期値や，繰り返し回数，MCMC法を変えたりしながら常に成立する特性か否かを判断する必要があろう．

MCMC法の比較

- independent MH
  - θがθ^(i-1)と独立

$$q(\theta^{(i-1)},\theta|y)=q(\theta|y)$$





- random walk Metropolis
  - θがθ^(i-1)に依存

$$q(\theta^{(i-1)},\theta|y)=f(\theta-\theta^{(i-1)})$$