

第11章

RでWinBUGSを使うインターフェイス

11.1 WinBUGS入門

11.2 WinBUGSとRを繋つなぐインターフェイス

11.3 codaパッケージを使ったMCMC診断

茨城大学工学部情報工学科
菊池裕紀

11.1 WinBUGS入門

- はじめに

WinBUGS

- マルコフ連鎖モンテカルロ法 (MCMC) において代表的な手法であるギブスサンプリングを行うことのできるフリーソフト
(メトロポリスヘイスティングスは不可)
- MCMCのアルゴリズムを使って複雑な統計モデルをベイズ法で当てはめることが目的
- Windows用のソフトであるがMacでも利用可能

WinBUGSの簡単な例

- 二項分布 $\text{binomial}(n,p)$ に従う y を観測する
 p : ベータ事前分布 $\text{beta}(\alpha,\beta)$ ※($\alpha=0.5$, $\beta=0.5$)
 $n=50$ (標本数)
 - $y=7$ 回の成功を観測し、 p の90%区間推定を行う。

```
model{
  y ~ dbin(p , n)
  p ~ dbeta(alpha , beta)
}
data
list(y=7,n=50,alpha=0.5,beta=0.5)
inits
list(p=0.1)
```

出力結果

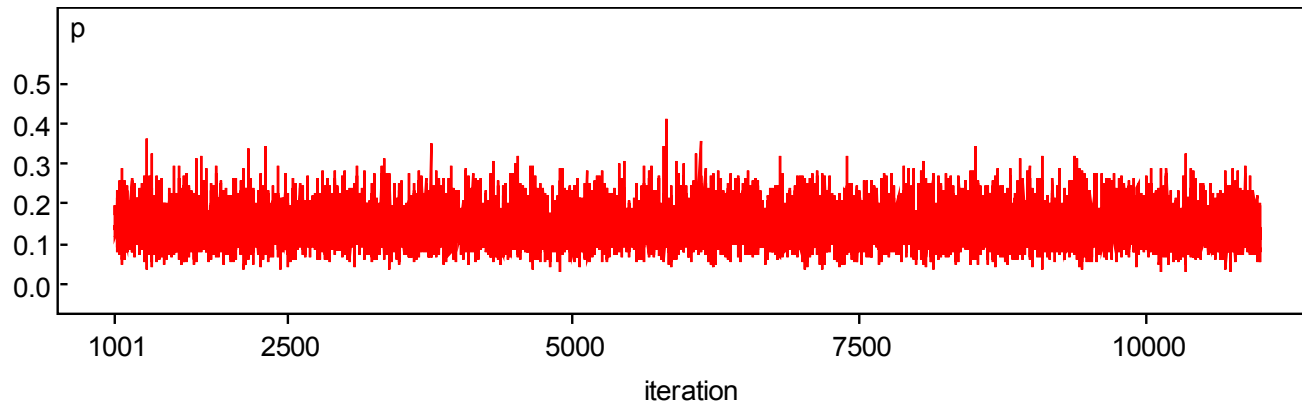


図1: history

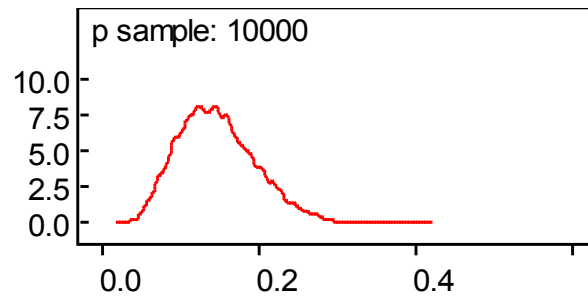


図2: density

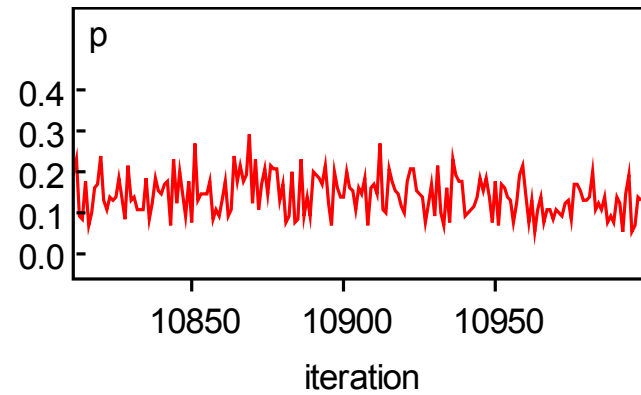


図3: trace

11.2 WinBUGSとRをつなぐインターフェイス

事前準備

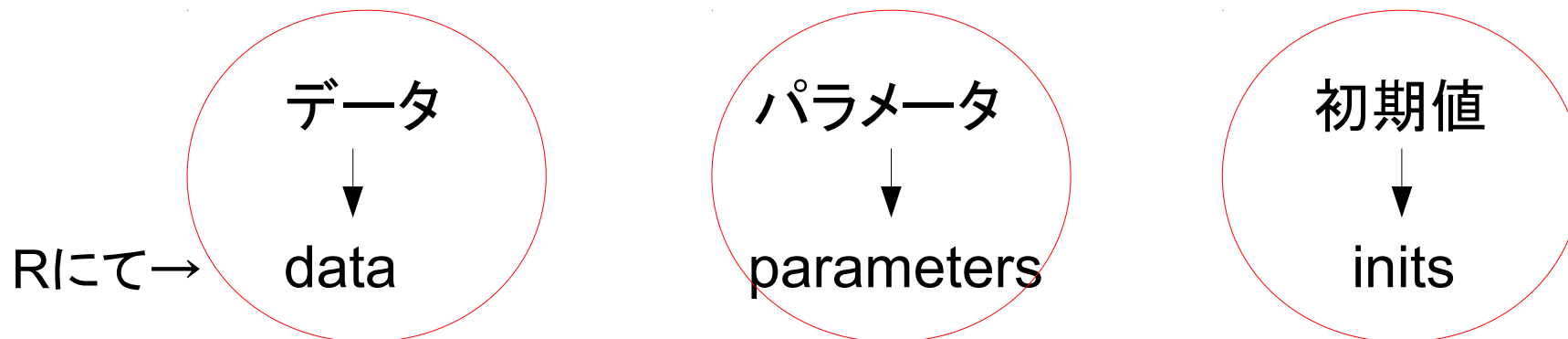
- WinBUGS、OpenBUGSのインストール
(<http://www.mrc-bsu.cam.ac.uk/bugs/>)
- Rのパッケージインストール
(R2WinBUGS、BRugs など)



RからOpenBUGSを呼び出すためのパッケージ

RでWinBUGSを実行する

- Rで実行する際も与える引数は”4”つ
 - **モデル** : 統計モデルを記述
 - **データ** : Rで直接定数、ベクトル、行列などで入力
 - **パラメータ** : Rでモニタするパラメータを指定
 - **初期値** : Rのコンソールパラメータの初期値を指定



ベイズモデルからのシミュレーション

```
model.sim <- bugs(data, inits, parameters, "model.bug")
```

上記のコードをRで入力することで、WinBUGSがバックグラウンドで実行され、モデルからパラメータをシミュレートする。

bugs() : WinBUGSからの出力を保存した構造

model.sim: パラメータのシミュレーション標本にアクセスできる

bugs()に渡す引数

- n.cahins : マルコフ連鎖の数を指定
- n.iter : 各連鎖での繰り返し数の指定
- n.burnin : 破棄する初期の繰り返しを指定
- n.thin : 間引き率を指定
- bin : 結果を保存するまでの繰り返しの数を指定
- bugs.directory : WinBUGSをインストールしたフォルダ
- working.directory : 作業ディレクトリ
- debug : デバッグ指定(=TRUE)

11.3 codaパッケージを使ったMCMC診断

- ユーザ側で行うべきこと
 - シミュレーション結果を診断し、関心のある事後分布を近似的に表現しているかを確認すること

- シミュレーションでいくつ連鎖を実行すべきか。連鎖の初期値の選択で違いが生じるおうか。
- シミュレーションによる標本が事後分布からの標本を近似的に表現するまで、どれくらいのバーイン期間が必要か。
- 事後分布の要約の正確な近似値を得るまでにシミュレーション標本を何回収集すべきか。
- 事後分布のある要約量のシミュレーション標準誤差はいくつか。
- 連続するシミュレーション標本に相関があるか。

codaパッケージ

そんな面倒な処理を簡単に…

- codaパッケージ

MCMCの出力を診断するためのパッケージ

- バッチ平均に基づき、相関する出力の平均や標準偏差、最高確立区間、分位点、シミュレーション標準誤差など多様な統計量を求める
- 異なるパラメータからシミュレーションした標本の自己相関や相互相関を比較
- 収束に関するさまざまな診断量を計算
- ラグ相関、密度推定、移動平均などの多様なプロットを作成