7.4 ギブスサンプリングに向けて

ギブスサンプリングを考える前に

- ギブスサンプリング　for 複数のパラメーターを持つモデル

例

- 患者97名。新薬投与51人→12人が改善、偽薬投与46人→5人が改善
  患者間の相互作用は薬の効果に影響しない（＝独立）
- コインj の表のバイアス θ_j を考える（j = 1,2?）
  Nj回のコイン投げで、表が観測される数：z_j
  i番目のコイン投げの結果：y_{ji} （y=0で裏, y=1で表）
  表と裏の確率の式　…　$p(y|\theta) = \theta^y(1-\theta)^{(1-y)}$ ベルヌーイ分布

ベータ分布の共役事前分布

事前分布：ベータ関数
尤度関数：ベルヌーイ分布
事後分布：ベータ関数

- 今までとの違い…パラメータが複数存在・同時に動く
  パラメータそのものの動きを観測

どっちかっていうと
テキスト中の言葉の説明

2つのバイアスの事前分布・尤度・事後分布

- 2個のコインに、バイアスθ_1, θ_2が想定
  - バイアスはパラメータθの名前。0.5から逸脱するθの値「ではない」
- 前提：「パラメータ値の組み合わせにより事象が左右される」
  - θ_1, θ_2の組み合わせに対する確信度：p(θ_1, θ_2)　← θ_1, θ_2に支配
- 確信度…確率密度関数の形式をとる（＝尤度？）

周辺分布

- θ_1, θ_2(の信念)が独立である　→　p(θ_1, θ_2) = p(θ_1)p(θ_2)
  - （独立の場合：計算は単純）
- データセット：　D = {z_1, N_1, z_2, N_2}　とする
- $p(D|\theta_1, \theta_2) = \prod_{y_{1i} \in D_1} p(y_{1i}|\theta_1, \theta_2) \prod_{y_{2i} \in D_2} p(y2_i|\theta_1, \theta_2)$

$$= \theta_1{}^{z_1}(1-\theta_1)^{(N_1-z_1)}\theta_2{}^{z_2}(1-\theta_2)^{(N_2-z_2)}\pi r^2$$

← 尤度関数

正確に式の分布をすることによる事後分布

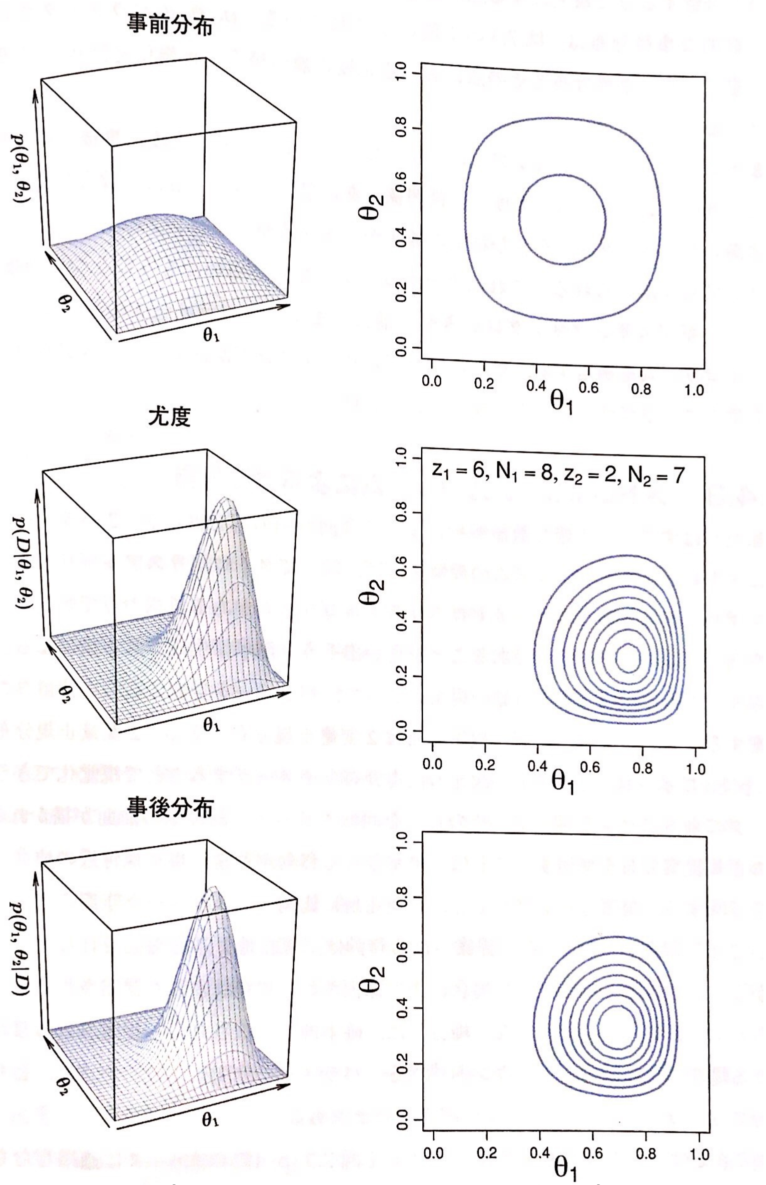
- ベータ関数：ベルヌーイ分布や二項分布の共役事前分布
  - パラメータは二つ：a,b
  - 平均はa/(a+b)
  - a=b=1のとき一様分布
  - a=bのとき1/2を中心に対称
- 事後分布　∝　尤度　×　事前分布

ベータ分布　　ベルヌーイ分布　　ベータ分布

- パラメータ $\theta_1\theta_2$の事前分布

…両方大体0.50？でも確信はない

（丘がなだらか）



遠近プロット　　等高線プロット

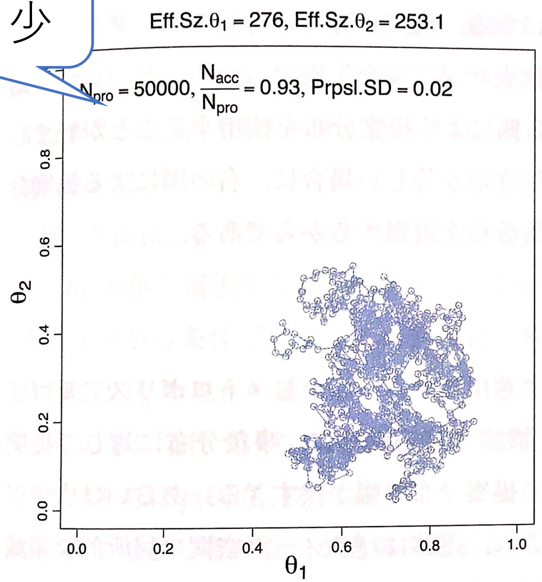
メトロポリスアルゴリズムによる事後分布

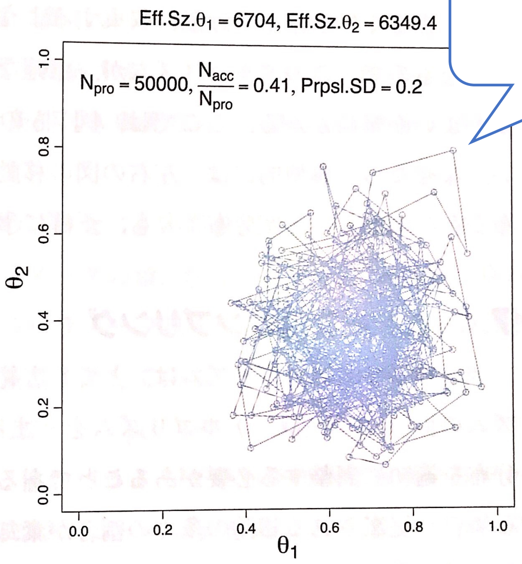
- 今までとの違い…パラメータが複数存在・同時に動く
  パラメータそのものの動きを観測

メトロポリスの場合

- 現在地の尤度＜次候補の尤度　→　必ず次候補に移動
- 次候補の尤度≦現在地　→　確率で次候補に移動/停滞

移動幅小
「停滞」少

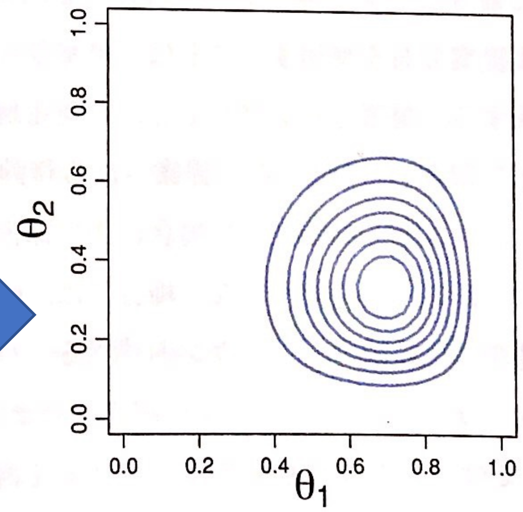




移動幅大
「停滞」多

（同じステップ数なら）
より効率的
正確に事後分布近似

似てる！



ギブスサンプリングによる事前分布

- メトロポリスアルゴリズムの問題
  - 事後分布に対して、移動先を決めるための尤度分布を調整させなきゃいけない
  - （尤度分布の幅が小さすぎor大きすぎ→停滞続き）

- ギブスサンプリング：メトロ同様ランダムウォーク
- 違い①：各ステップにおける、パラ決定の順番
  - メトロ…　$(\theta_1, \theta_2) \rightarrow (\theta_1, \theta_2) \rightarrow (\theta_1, \theta_2) \rightarrow (\theta_1, \theta_2)$
  - ギブス…　$\theta_1 \rightarrow \theta_2 \rightarrow \theta_1 \rightarrow \theta_2 \rightarrow \theta_1 \rightarrow \theta_2$

パラ数多くても、全パラに均等になるように順番こ

- 違い②：各ステップにおける、パラの移動様態
  - 次スライドにて説明

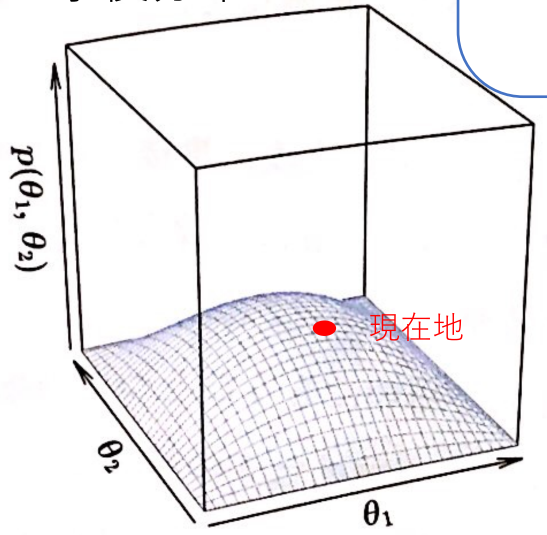
次ステップへの移動に関する、メトロとギブスの差異

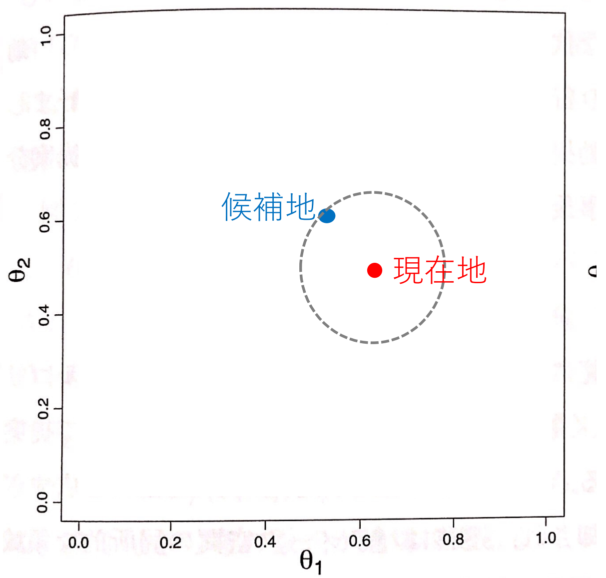
メトロ

移動しない時もある

（移動距離＝点線円の半径は既定）
①候補地の選定
②候補地の尤度計算
③移動の有無を決定
＜繰り返し＞

事後分布



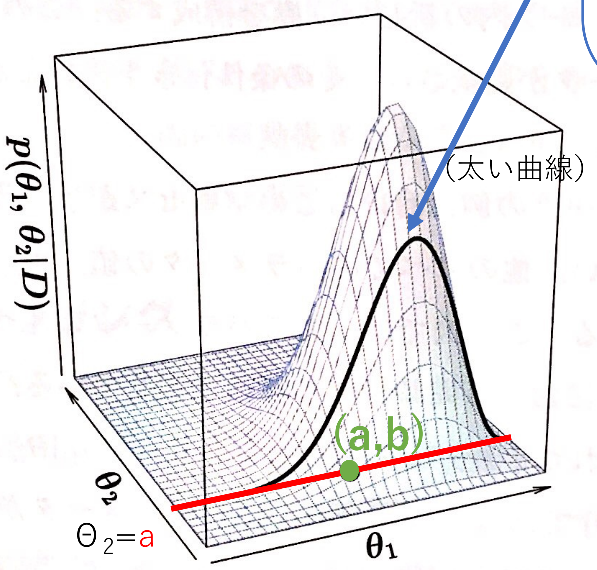


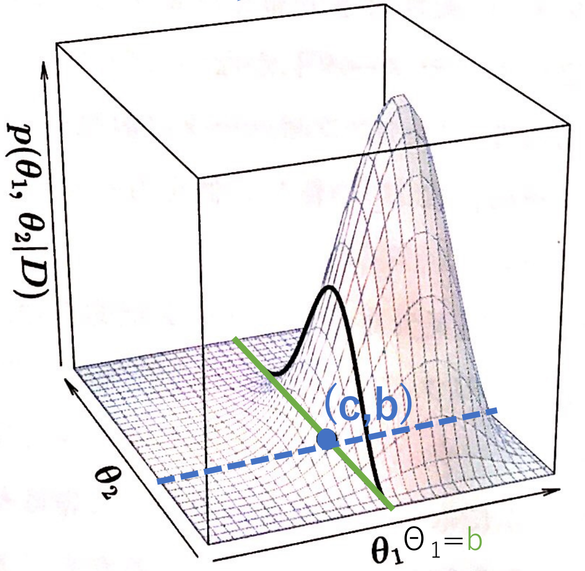
ギブス

毎回必ず移動

①暫定のパラ：θ_2=a
②θ_2=aの断面の**確率密度関数**から、確率にしたがって(a,b)を選定
③暫定のパラ：θ_1=b
④②同様、(c,b)を選定
＜繰り返し＞

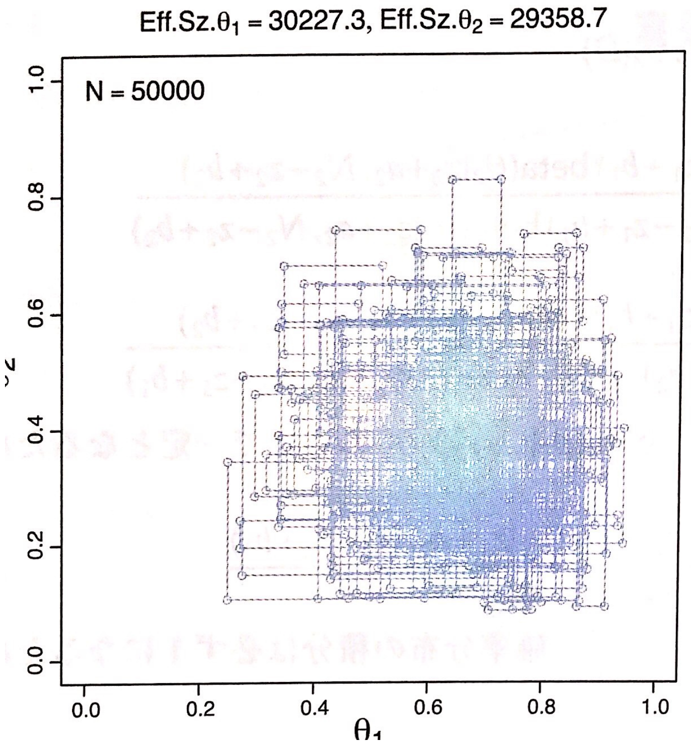
事後分布

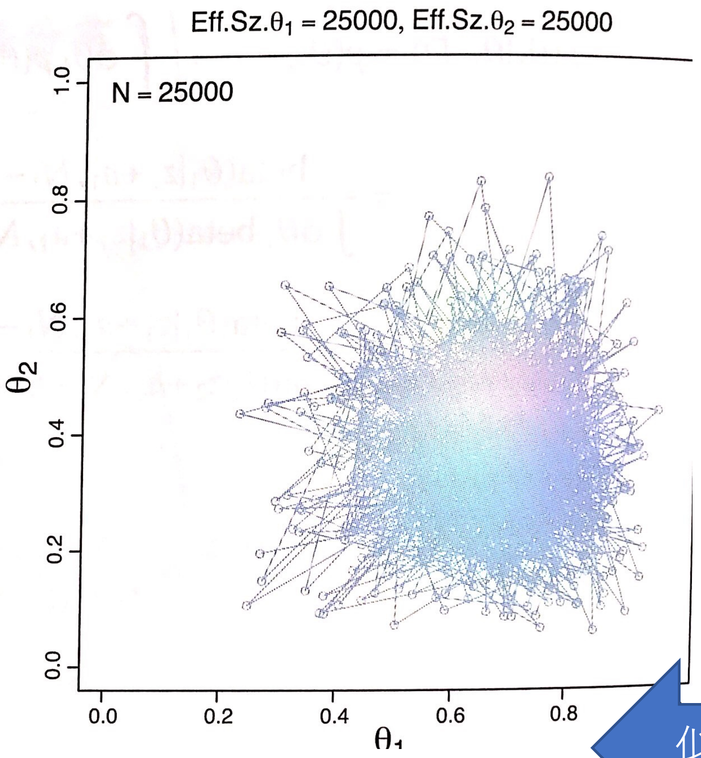




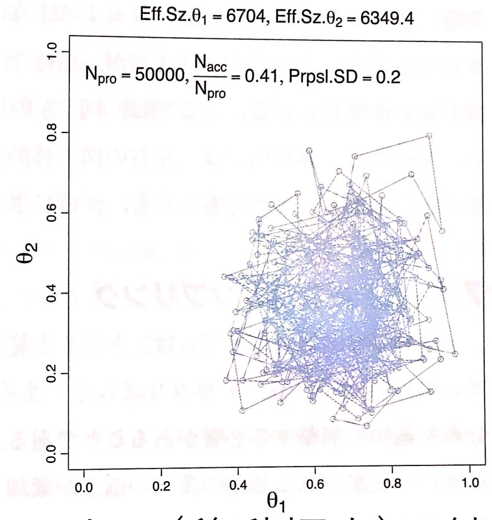
ギブスサンプリングの結果と ギブスとメトロの比較

- ギブスサンプリングの結果
  - ↓左における2ステップを1ステップにしたのが右



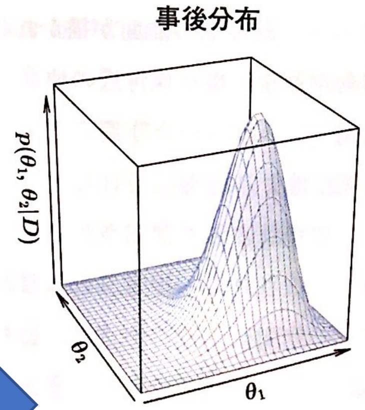


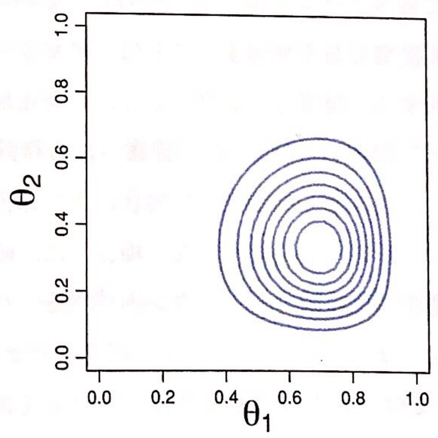
似てる



メトロ(移動幅大)の結果

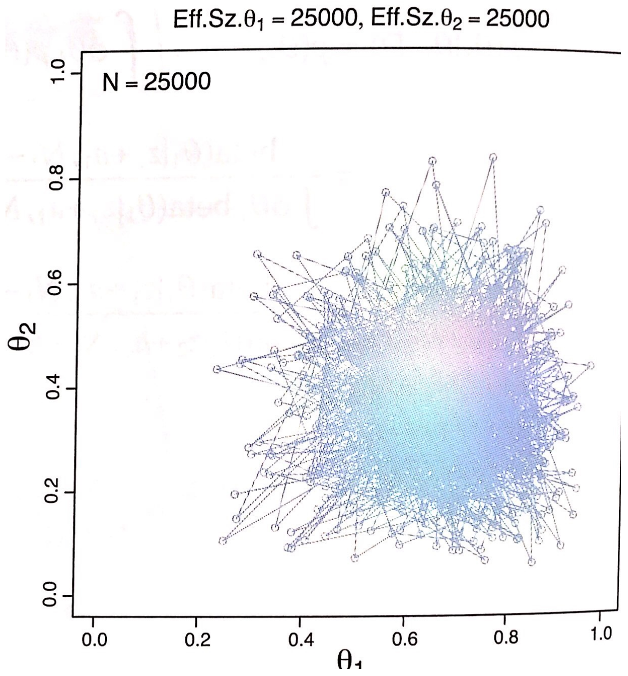
似てる



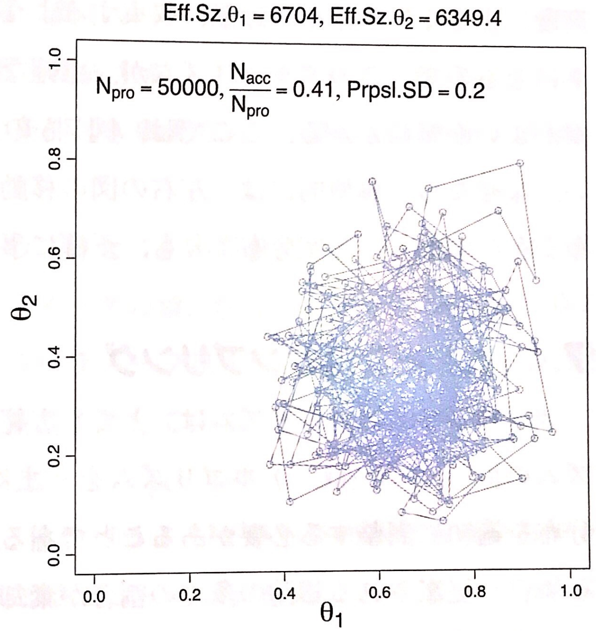


ギブスサンプリングの結果と
ギブスとメトロの比較

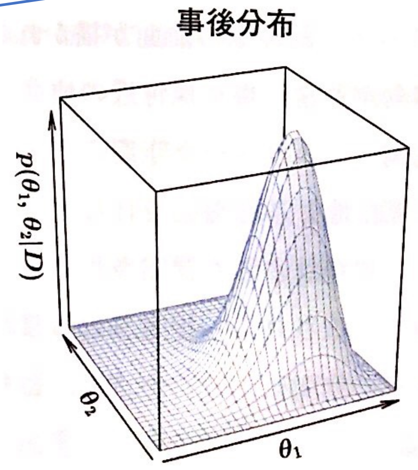
余談：このように視覚化すると…
・サンプリングの挙動がすぐにわかる
・なぜサンプリングが機能するのかを理解する助けになる？？？？

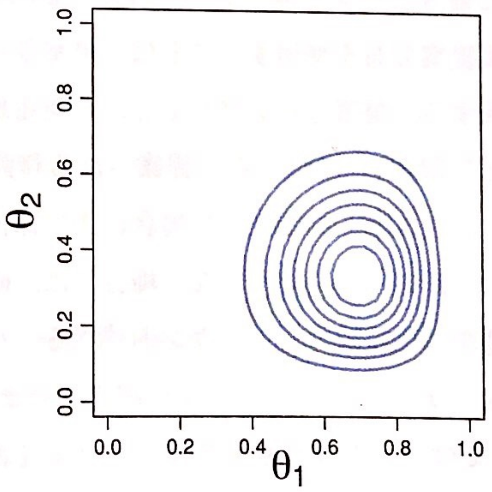


ギブスの結果



メトロ(移動幅大)の結果





- 問題：同じステップ数で、どちらが有効な結果を得られるか？
- 有効サイズ：ギブス＞メトロ（今回はたまたま）

ギブスサンプリングの長所・短所

長所

- 提案の棄却(停滞)が不要
  ＝計算省略・同ステップ数でより充実したパラメータ策定が可能
- 事後分布・尤度分布の調整が不要
  （メトロポリスアルゴリズムの欠点を克服）

短所

- 一方のパラを固定した時に、他のパラの確率分布を生成する必要がある
  （大変ってこと？）
- 相関の高いパラ同士だと計算に時間がかかる
  （今回は省略）