

# ベイズ生存時間分析をSTANでやってみる

伊庭 幸人

モデリング研究系 教授

(データ同化研究開発センター, 統計思考院 兼任)

状態空間モデルにもとづくベイズ生存時間分析は古くから提案されている (Gamerman(1991), Fahrmeir(1994)) が, その割にあまり普及していないように思われる. 平滑化事前分布(1階・2階)の利用により, 柔軟な枠組みが得られるのが特徴である.

ここでは, STANなどのMCMCツールを用いれば手軽に実装でき, 興味深い拡張が可能であることを示す.

(共同研究者 立森久照・加藤直広 国立精神神経医療研究センター精神保健研究所)

例1. 比例ハザードモデルの拡張 .. 「比例」の条件を緩和 (共変量が2値, すなわち処置群と対照群の比較の場合) 先行研究にすでにあるが, STAN上に簡単に実装できる

「ベースとなるハザードが滑らかに変化」及び「対数ハザード比が滑らかに変化」と仮定(2階差分)

Y イベント数, Z リスクセット 10%/年脱落  
模擬データのサンプルサイズ5000 追跡期間10年

観測過程(尤度)

$$Y_t^{(0)} \sim \text{Poisson}(\lambda_t^{(0)} Z_t^{(0)})$$

$$Y_t^{(1)} \sim \text{Poisson}(\lambda_t^{(1)} Z_t^{(1)})$$

```
transformed parameters {
  real prob[T,2];
  for(t in 1:T)
    prob[t,1] <- exp(beta[t]);
  for(t in 1:T)
    prob[t,2] <- exp(beta[t]+a[t]); }
```

リンク関数

$$\log \lambda_t^{(0)} = \beta_t$$

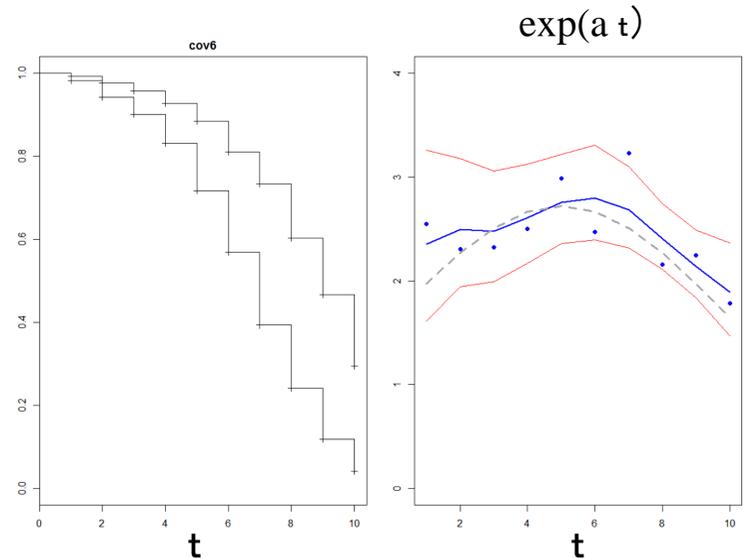
$$\log \lambda_t^{(1)} = \beta_t + a_t$$

```
model {
  for(t in 3:T)
    beta[t] ~ normal(2*beta[t-1]-beta[t-2], s_mu);
  for(t in 3:T)
    a[t] ~ normal(2*a[t-1]-a[t-2], s_a);
  for(t in 1:T)
    for(n in 1:2)
      Y[t,n] ~ poisson(Z[t,n]*prob[t,n]); }
```

システム過程(事前分布の一部)

$$\beta_t = 2\beta_{t-1} - \beta_{t-2} + \epsilon_t$$

$$a_t = 2a_{t-1} - a_{t-2} + \eta_t$$



右の図で, 青:事後平均, 赤:95%確信区間, 青点:割り算推定値 点線:正解

例2. 「観察時間の成分」 $\beta(t)$ と「年齢の成分」 $\gamma(s)$  に分解 (暴露or前兆→発症)の時間経過t と年齢sの両方に依存

Y イベント数, Z リスクセット (観察開始からの経過時間 t と イベント発生時の年齢 s ごとに)

模擬データのサンプルサイズ1000 追跡期間10年  
追跡開始時の年齢は(X~X+19)に一様分布  
10%/年脱落

「どちらの成分も徐々に変化」と仮定(1階差分)

観測過程(尤度)

$$Y_{ts} \sim \text{Poisson}(\lambda_{ts} Z_{ts})$$

リンク関数

$$\log \lambda_{ts} = \alpha + \beta_t + \gamma_s$$

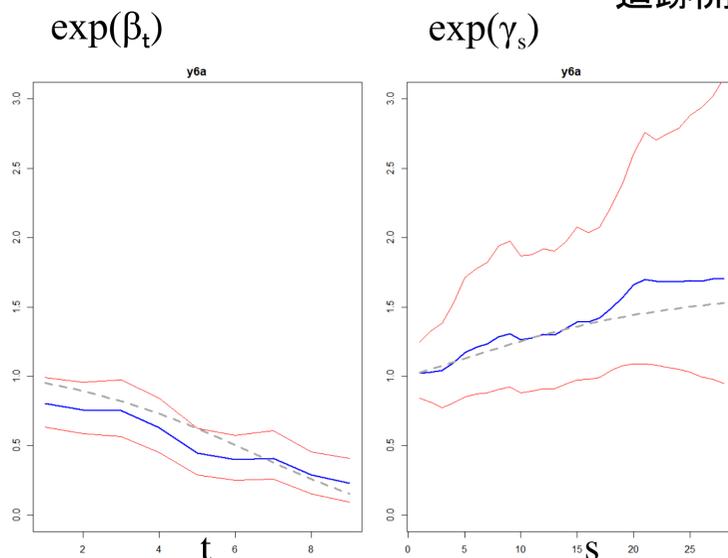
システム過程(事前分布の一部)

$$\beta_1 = 0$$

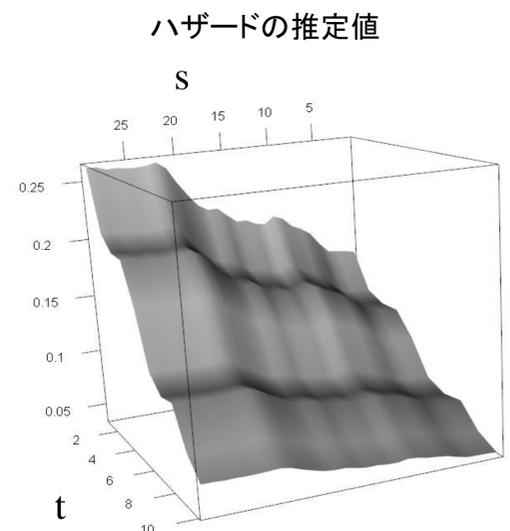
$$\beta_t = \beta_{t-1} + \epsilon_t$$

$$\gamma_1 = 0$$

$$\gamma_s = \gamma_{s-1} + \eta_s$$



左右とも, 青:事後平均, 赤:95%確信区間, 点線:正解



総研大大学院生募集 他大学の院生も在籍のまま受け入れ可能(1~2年:特別共同利用研究員制度)  
— 具体的な問題を扱いながら, 統計モデリングやモンテカルロ法の使い方が学べます —

医学・疫学データの時空間階層ベイズ解析  
外部の保健・医学研究者と組んで具体的なデータを解析

レアイベントサンプリングで解ける問題各種

組み合わせの数の計算や時間逆転シミュレーション,  
リスク解析などさまざまな手法や応用が考えられます

政党支持率データの状態空間モデル解析  
メディアによる違いや時間変動を統一的に扱う手法を開発