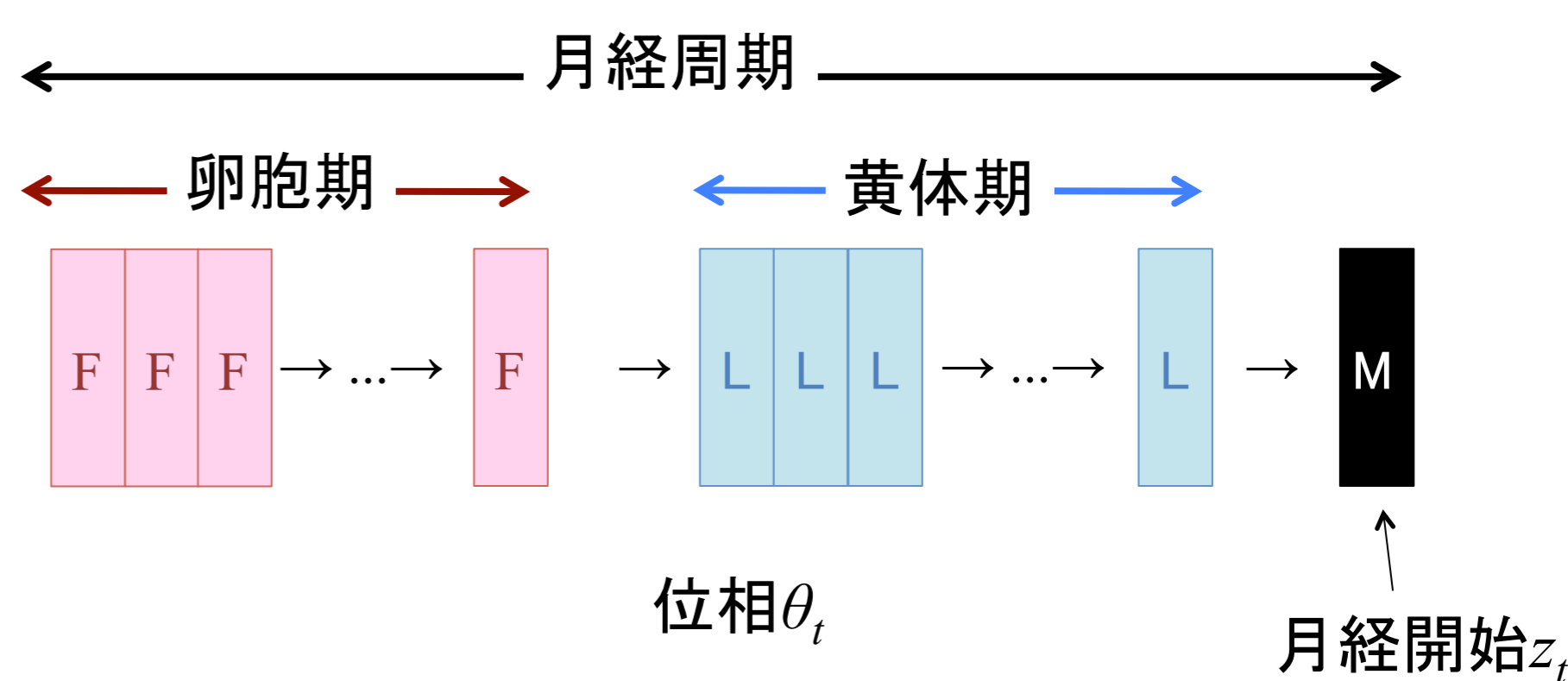


自己励起型状態空間モデルを用いた月経周期解析

川森 愛 データ科学研究系 特任研究員

[月経周期]

- ・1周期は約28日
- ・排卵を境に卵胞期[F]と黄体期[L]に分かれる
- ・卵胞期の日数長はばらつきが大きい
- ・黄体期の日数長はばらつきが小さい(約14日)
- ・平均体温は卵胞期で低く、黄体期には高い(0.3~0.5°C差)



[目標]

体温データから次の月経開始日を予測する

[状態空間モデル]

1. 過程モデル

$$\theta_t = \theta_{t-1} + \varepsilon_t \text{ mod } 1$$

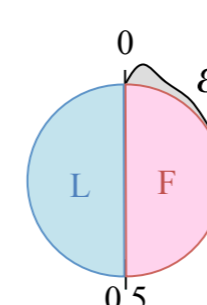
$$\varepsilon_t \sim \text{Gamma}(\alpha_t, \beta_t)$$

$$(\alpha_t, \beta_t) = (\alpha_1, \beta_1) \text{ when } \theta_t < 0.5$$

$$= (\alpha_2, \beta_2) \text{ when } \theta_t \geq 0.5$$

自己励起モデル

位相 θ_t

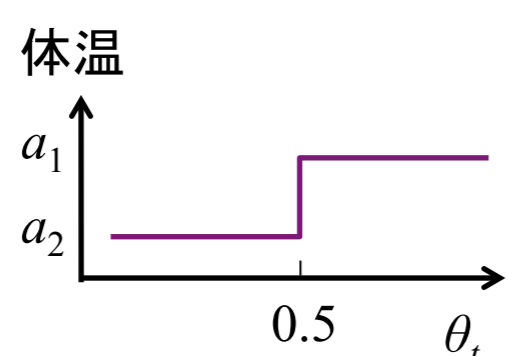


2. 観測モデル (体温)

$$y_t \sim \text{Norm}(a_t, \sigma_t)$$

$$(a_t, \sigma_t) = (a_1, \sigma_1) \text{ when } \theta_t < 0.5$$

$$= (a_2, \sigma_2) \text{ when } \theta_t \geq 0.5$$



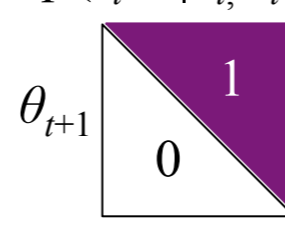
(月経開始)

$$z_t = 0 \text{ when } \theta_t > \theta_{t-1}$$

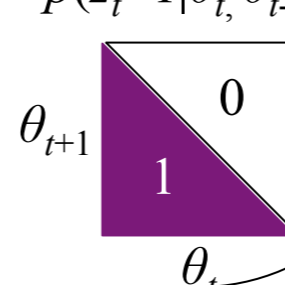
$$z_t = 1 \text{ when } \theta_t \leq \theta_{t-1}$$

$$p(z_t | \theta_t, \theta_{t-1}) = (1-z_t)I(\theta_t > \theta_{t-1}) + z_t I(\theta_t \leq \theta_{t-1})$$

$$p(z_t=0 | \theta_t, \theta_{t-1})$$

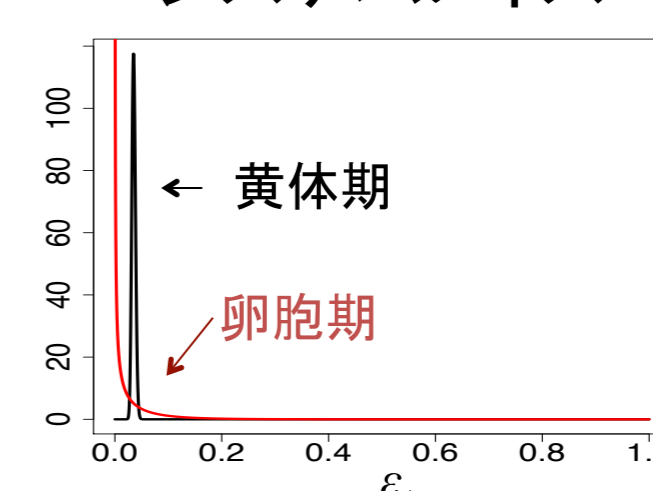


$$p(z_t=1 | \theta_t, \theta_{t-1})$$



[パラメータ推定値]

システムノイズ



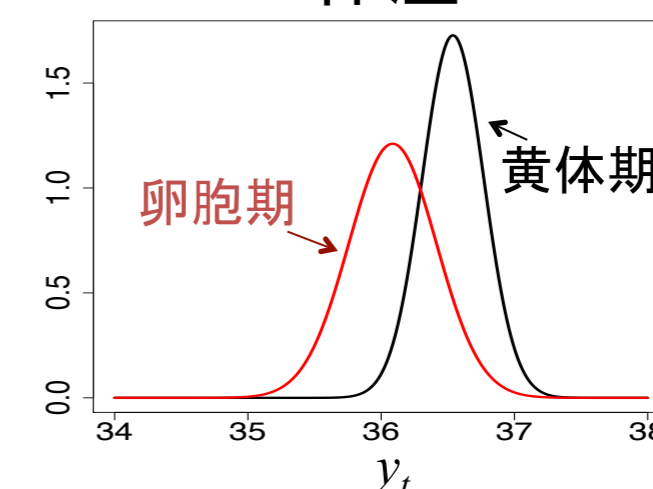
$$\alpha_1 = 0.3255499$$

$$\beta_1 = 12.4047963$$

$$\alpha_2 = 106.9998992$$

$$\beta_2 = 3036.2378939$$

体温



$$\alpha_1 = 36.0861182$$

$$\sigma_1 = 0.3295040$$

$$\alpha_2 = 36.5387787$$

$$\sigma_2 = 0.2309690$$

[位相の推移]

逐次ベイズフィルタ

予測分布

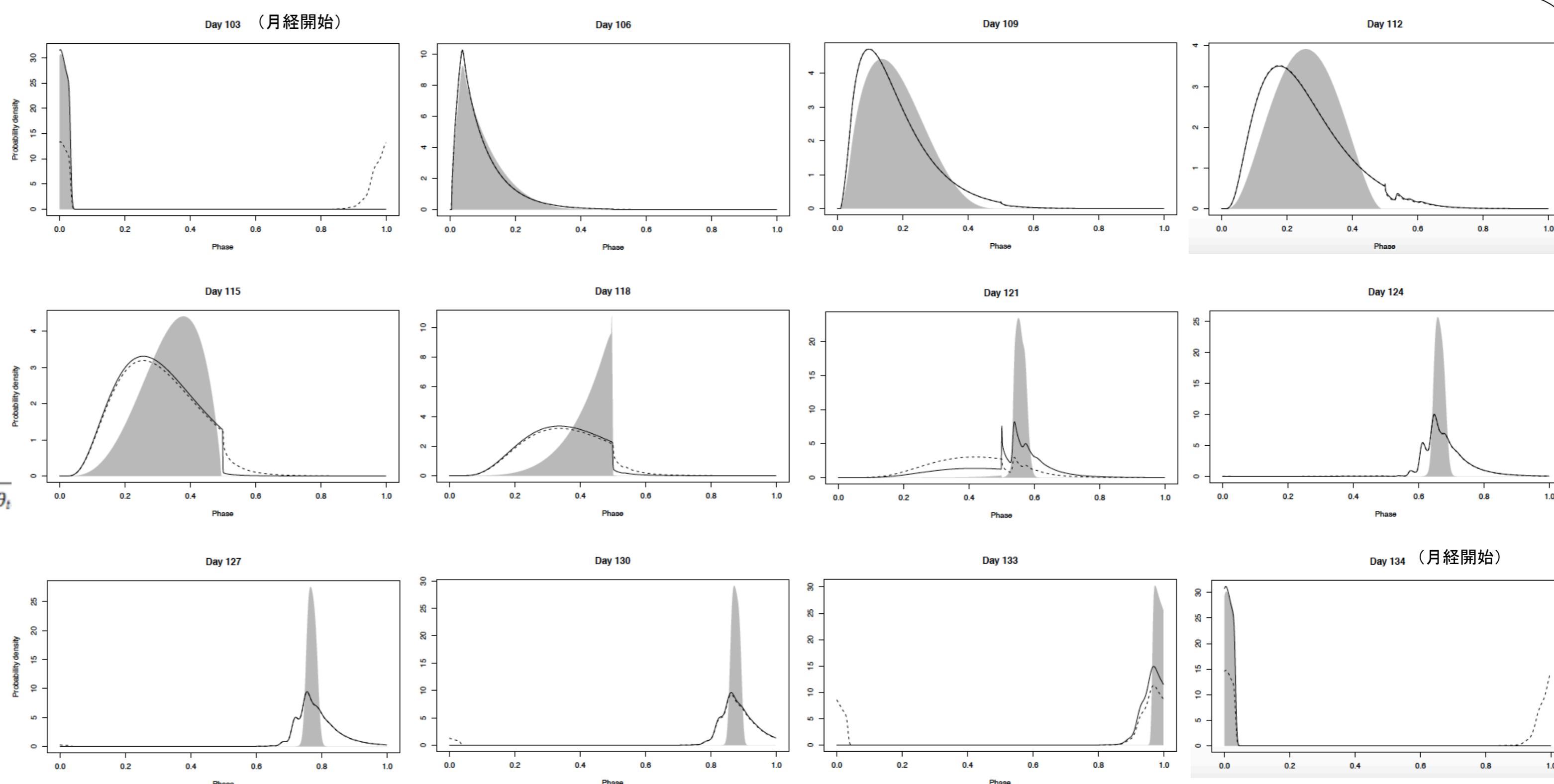
$$p(\theta_{t+1}, \theta_t | Y_t, Z_t) = p(\theta_{t+1} | \theta_t) p(\theta_t | Y_t, Z_t)$$

フィルタ分布

$$p(\theta_t, \theta_{t-1} | Y_t, Z_t) = \frac{p(y_t | \theta_t) p(z_t | \theta_t, \theta_{t-1}) p(\theta_t, \theta_{t-1} | Y_{t-1}, Z_{t-1})}{\int p(y_t | \theta) p(z_t | \theta, \theta_{t-1}) p(\theta, \theta_{t-1} | Y_{t-1}, Z_{t-1}) d\theta d\theta_{t-1}}$$

平滑化分布

$$p(\theta_t, \theta_{t-1} | Y_T, Z_T) = \frac{p(\theta_t, \theta_{t-1} | Y_t, Z_t) p(\theta_t | Y_T, Z_T)}{p(\theta_t | Y_t, Z_t)}$$



[月経開始日予測]

STEP1. s日後にフェイズ移行し,

k日後に月経が起こる確率 $f_s(k|\theta_t)$ を求める

STEP2. k日後に月経が起こる確率 $f(k|\theta_t)$ を求める

$$\sum_{s=1}^k f_s(k|\theta_t) = f(k|\theta_t)$$

STEP3. 月経予測分布 $h(k|Y_t, Z_t)$ を求める

$$h(k|Y_t, Z_t) = \int_0^1 f(k|\theta_t) p(\theta_t | Y_t, Z_t) d\theta_t$$

$f_s(k|\theta_t)$ の求め方

$$F \rightarrow F \quad P[\theta_{t+1}^F | \theta_t^F] = P[\theta_t + \varepsilon_{t+1}^F < 0.5] = P[\varepsilon_{t+1}^F < 0.5 - \theta_t]$$

$$F \rightarrow L \quad P[\theta_{t+1}^L | \theta_t^F] = P[\theta_t + \varepsilon_{t+1}^F \geq 0.5] - P[\theta_t + \varepsilon_{t+1}^F < 1]$$

$$= P[\varepsilon_{t+1}^F \geq 0.5 - \theta_t] - P[\varepsilon_{t+1}^F < 1 - \theta_t]$$

$$L \rightarrow L \quad P[\theta_{t+1}^L | \theta_t^L] = P[\theta_t + \varepsilon_{t+1}^L < 1] = P[\varepsilon_{t+1}^L < 1 - \theta_t]$$

$$L \rightarrow M \quad P[\theta_{t+1}^M | \theta_t^L] = P[\varepsilon_{t+1}^L \geq 1 - \theta_t]$$

$$f_s(k|\theta_t) = \prod_{i=1}^s [F \rightarrow F] \times [F \rightarrow L] \times \prod_{i=1}^{k-s} [L \rightarrow L] \rightarrow [L \rightarrow L] \rightarrow [L \rightarrow M]$$

$P[\varepsilon < x]$ 及び $P[\varepsilon \geq x]$ の求め方

$\varepsilon^F \sim f(t), \varepsilon^L \sim g(t)$ ただし、 $f(t), g(t)$ はガンマ分布の確率密度関数

畳み込みにより、 $m(t) = f * g(t)$

$$= \int_0^t f(\tau) g(t-\tau) d\tau$$

累積分布関数 $M(t)$ $M(t) = \int_0^t m(u) du$

$$= \int_0^t \int_0^u f(\tau) g(u-\tau) d\tau du$$

$$= 1 - \int_t^\infty \int_0^u f(\tau) g(u-\tau) d\tau du$$

$$P[\varepsilon < x] = M(x), \quad P[\varepsilon \geq x] = 1 - M(x)$$