

ベースball場

4.3



4.3. ポアソン混合モデルにおける推論。
• λ に関するデータに対してポアソン混合モデルを導入し、事後分布を推論するためのアルゴリズムを導出。

4.3.1. ポアソン混合モデル

① 各峰性の離散非負データを掌握する際に利用できる。

パラメータ: $\lambda = \{\lambda_1, \dots, \lambda_K\}$

観測モデル (グラスラーに付く)

$$p(x_n | \lambda_k) = \text{Poi}(x_n | \lambda_k) \left(= \frac{\lambda_k^{x_n}}{x_n!} e^{-\lambda_k} \right) \quad (4.27)$$

~) 混合分布における条件分布は。

$$p(\lambda_n | S_n, \lambda) = \prod_{k=1}^K \text{Poi}(x_n | \lambda_k)^{S_{n,k}} \quad (4.28)$$

… K 個あるポアソン分布のうちで S_n が S_n に沿って指定される。
(c.f. (4.3))

ポアソン分布 λ のパラメータ $\lambda = \{\lambda_1, \dots, \lambda_K\}$ は次の事前分布。

… ガンマ分布 (共役事前分布)

$$p(\lambda_k) = \text{Gam}(\lambda_k | a, b) \quad (4.29)$$

where a, b : hyper parameter (fixed)

仮定: a, b はグラスラー内で共通。

… 通常予想なども応用して考慮できる場合は (はじめからグラスラー) の hyper parameter を明記してアルゴリズムを導出するが良い。

S は実可算事前分布

$$p(S_n | \pi) = \text{Cat}(S_n | \pi) \quad (\text{c.f. (4.2)})$$

π は実可算事前分布

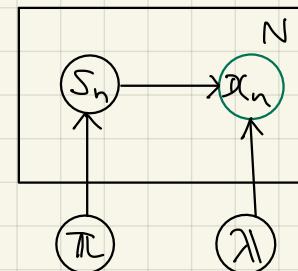
$$p(\pi) = \text{Dir}(\pi | \alpha) \quad (\text{c.f. (4.4)})$$

~) 同時分布は。 (4.5)

$$\begin{aligned} p(X, S, \lambda, \pi) &= p(X | S, \lambda) p(S | \pi) p(\lambda) p(\pi) \\ &= \left\{ \prod_{n=1}^N p(x_n | S_n, \lambda) p(S_n | \pi) \right\} \left\{ \prod_{k=1}^K p(\lambda_k) \right\} p(\pi) \end{aligned}$$

Remark

構築した同時分布自体がデータの発生過程 (= 実可算) の仮説を明確に表現している。



4.3.2. ギブスサニ=7°ルニ=7°

ホトトノ混合モテル、事後分布が $\text{Pois}(\lambda)$ 、 π と潜在変数 S をサニ=7°ルする
アルゴリズムを導出可。

$$p(S, \lambda, \pi | X) \quad (4.30)$$

X が観測された後、条件付分布。

混合分布では、潜在変数 S を $\text{Pois}(\lambda)$ で分けてサニ=7°ルする。十分に簡単な確率分布が得られることが知られていく。

$$\sim S \sim p(S | X, \lambda, \pi) \quad (4.31)$$

$$\lambda, \pi \sim p(\lambda, \pi | X, S) \quad (4.32)$$

1. (4.31) が $\text{S} = \{S_1, \dots, S_N\}$ をサニ=7°ルする条件付分布を求める。

$$\frac{p(S | X, \lambda, \pi)}{p(X, \lambda, \pi)} = \frac{p(X | S, \lambda, \pi)}{p(X, \lambda, \pi)} \quad (\because \text{Bayes})$$

$$\begin{aligned} & \text{テーク } X \\ & \text{Pois}(\lambda, \pi) = p(X | S, \lambda, \pi) p(S | \pi) p(\lambda) p(\pi) \quad (\because (4.5)) \\ & \text{がでる} \rightarrow \text{下の} \end{aligned}$$

$$\begin{aligned} & \text{S の分布} \propto p(X | S, \lambda) p(S | \pi) \\ & = \prod_{n=1}^N p(x_n | S_n, \lambda) p(S_n | \pi) \quad (4.33) \end{aligned}$$

S_1, \dots, S_N の独立性の程

② テーク $\text{Pois}(\lambda)$ が $\text{S}_1, \dots, \text{S}_N$ の独立性を保つ。つまり S_n が独立にサニ=7°ルできること。(条件付独立性)

具体的に計算可。

$$\frac{\lambda^{x_n}}{x_n!} e^{-\lambda}$$

$$\begin{aligned} \ln p(x_n | S_n, \lambda) &= \sum_{k=1}^K S_{n,k} \ln \underbrace{p_{\text{Pois}}(x_n | \lambda_k)}_{\lambda_k^{x_n} / x_n!} \quad (\because (4.28)) \\ &= \sum_{k=1}^K S_{n,k} \left(x_n \ln \lambda_k - \ln x_n! - \lambda_k \right) \quad (4.34) \end{aligned}$$

$$\ln p(S_n | \pi) = \ln \underbrace{C_{\text{Cat}}(S_n | \pi)}_{\prod_{k=1}^K \pi_k^{S_{n,k}}} \quad (\because (4.2))$$

$$= \sum_{k=1}^K S_{n,k} \ln \pi_k \quad (4.35)$$

5.7.

$$\begin{aligned} \ln p(x_n | S_n, \pi) p(S_n | \pi) &= \sum_{k=1}^K S_{n,k} (x_n \ln \lambda_k - \ln x_n! - \lambda_k + \ln \pi_k) \\ &= \sum_{k=1}^K S_{n,k} (x_n \ln \lambda_k - \lambda_k + \ln \pi_k) + \underbrace{\text{const.}}_{(\because \sum_{k=1}^K S_{n,k} = 1)} \quad (4.36) \end{aligned}$$

補足) $\sum_{k=1}^K S_{n,k} = 1$ が、これは S_n の制約条件であるべき。

$$\mathcal{S}_n \sim \text{Cat}(\mathcal{S}_n | \pi_n) \quad \left(= \prod_{k=1}^K \pi_{n,k}^{s_{n,k}} \right) \quad (4.37)$$

where

$$\pi_{n,k} \propto \exp(x_n \ln \lambda_k - \lambda_k + \ln \pi_k)$$

s.t. $\sum_{k=1}^K \pi_{n,k} = 1 \quad (4.38)$

\sim \mathcal{S}_n は確率で $\pi_{n,k}$ を計算する二乗。力でこの分布が \mathcal{S}_n が \mathcal{X} の

2. (4.32) が \mathcal{X} の \mathcal{S} を規則化によりに及ぼす。

$$\begin{aligned} p(\lambda, \pi | \mathcal{X}, \mathcal{S}) &= \frac{p(\mathcal{X}, \mathcal{S}, \lambda, \pi)}{p(\mathcal{X}, \mathcal{S})} \quad (\because \text{Bayes}) \\ &= \frac{p(\mathcal{X} | \mathcal{S}, \lambda) p(\mathcal{S} | \pi) p(\lambda) p(\pi)}{p(\mathcal{X}, \mathcal{S})} \quad (\because (4.5)) \\ &\propto \underbrace{p(\mathcal{X} | \mathcal{S}, \lambda)}_{\lambda \mid \mathcal{X} \text{の確率}} \underbrace{p(\mathcal{S} | \pi)}_{\pi \mid \mathcal{S} \text{の確率}} \underbrace{p(\lambda)}_{\lambda \mid \mathcal{X} \text{の確率}} \underbrace{p(\pi)}_{\pi \mid \mathcal{S} \text{の確率}} \quad (4.39) \end{aligned}$$

$$\lambda \mid \mathcal{X} \text{の確率} = \prod_{n=1}^N \prod_{k=1}^K S_{n,k} \ln \text{Poi}(x_n | \lambda_k)$$

$$= \sum_{n=1}^N \sum_{k=1}^K S_{n,k} \ln \underbrace{\frac{\lambda_k^{x_n}}{x_n!} e^{-\lambda_k}}_{\text{Poi}(x_n | \lambda_k)} + \sum_{k=1}^K \ln \underbrace{\text{Gam}(\lambda_k | a, b)}_{\frac{b^a}{\Gamma(a)} \lambda_k^{a-1} e^{-b\lambda_k}} \quad (\because (4.28), (4.29))$$

$$\sum_{n=1}^N \sum_{k=1}^K S_{n,k} \ln \text{Poi}(x_n | \lambda_k)$$

$$= \sum_{n=1}^N \sum_{k=1}^K S_{n,k} (x_n \ln \lambda_k - \ln x_n! - \lambda_k)$$

$$= \sum_{n=1}^N x_n \sum_{k=1}^K S_{n,k} \ln \lambda_k - \underbrace{\sum_{n=1}^N \ln x_n!}_{\sum_{n=1}^N S_{n,k}} - \sum_{n=1}^N \sum_{k=1}^K S_{n,k} \lambda_k$$

$$\sum_{k=1}^K \ln \text{Gam}(\lambda_k | a, b)$$

$$= \sum_{k=1}^K \ln \frac{b^a}{\Gamma(a)} + (a-1) \ln \lambda_k - b \lambda_k$$

$$\lambda \mid \mathcal{X} \text{の確率} = p(\mathcal{X} | \mathcal{S}, \lambda) p(\lambda)$$

$$= \sum_{k=1}^K \left\{ \left(\sum_{n=1}^N x_n S_{n,k} + a-1 \right) \ln \lambda_k - \left(\sum_{n=1}^N S_{n,k} + b \right) \lambda_k \right\} + \text{const.} \quad (4.40)$$

\sim 独立して λ_k は、力で \mathcal{S} 分布が \mathcal{X} の

$$\lambda_k \sim \text{Gam}(\lambda_k | a_k, b_k) \quad (4.41)$$

where

$$\begin{aligned} a_k &= \sum_{n=1}^N x_n S_{n,k} + a \\ b_k &= \sum_{n=1}^N S_{n,k} + b \end{aligned} \quad (4.42)$$

π の分布は π

$$\ln p(S | \pi) p(\pi)$$

$$= \sum_{n=1}^N \ln \underbrace{C_0}_{= C_0(\alpha)} \underbrace{(S_n | \pi)}_{= \prod_{k=1}^K \pi_k^{s_{n,k}}} + \ln \underbrace{\text{Dir}(\pi | \alpha)}_{= C_0(\alpha) \prod_{k=1}^K \pi_k^{\alpha_k - 1}} \quad (\because (4.4), (4.35))$$

$$= \sum_{n=1}^N \sum_{k=1}^K s_{n,k} \ln \pi_k + \sum_{k=1}^K (\alpha_k - 1) \ln \pi_k + \text{const.}$$

$$= \sum_{k=1}^K \left(\sum_{n=1}^N s_{n,k} + \alpha_k - 1 \right) \ln \pi_k + \text{const.} \quad (4.43)$$

よし.

$$\pi \sim \text{Dir}(\pi | \hat{\alpha}) \quad (4.44)$$

where

$$\hat{\alpha}_k = \sum_{n=1}^N s_{n,k} + \alpha_k \quad (4.45)$$

以上より、 π の混合モデルの事後分布は $\pi \sim \text{Dir}(\pi | \hat{\alpha})$ である.

Algorithm 4.2 : π の混合モデルの π のギブスサンプリング
初期値 π を用いて π を $\pi = \pi$ とする.

For $i = 1$ to MAXITER do

 For $n = 1$ to N do
 (4.37) を用いて s_n を $\pi = \pi$ する

 End For

 For $k = 1$ to K do
 (4.41) を用いて λ_k を $\pi = \pi$ する

 End For

 (4.44) を用いて π を $\pi = \pi$ する

End For

Remark

ここで S を最初に $\pi = \pi$ してから、 π を初期値として与えれば π が近づく.

4.3.3 変分推論

更新式を得るには、事後分布に好適な分解近似の仮定をおく必要がある。
 ~> 潜在変数 λ と π を分けて事後分布を近似

$$p(\mathbf{x}, \lambda, \pi | \mathbf{X}) \approx g(\mathbf{x}) g(\lambda, \pi) \quad (4.46)$$

… 変分ベイズ EM アルゴリズム
 (潜在変数 λ と π の分布を分けて近似する手続き)

1. $g(\mathbf{x})$ に対する変分推論、公式 (4.25) を当てはめよ。

$$\begin{aligned} \ln g(\mathbf{x}) &= \langle \ln p(\mathbf{X}, \mathbf{x}, \lambda, \pi) \rangle_{g(\lambda, \pi)} + \text{const.} \quad (\because (4.25)) \\ &= \langle \ln p(\mathbf{X} | \mathbf{x}, \lambda) + \ln p(\mathbf{x} | \pi) + \underbrace{\ln p(\lambda) + \ln p(\pi)}_{\mathbf{x} \text{ は固く定取}} \rangle_{g(\lambda, \pi)} \\ &+ \text{const.} \quad (\because (4.39)) \\ &= \langle \ln p(\mathbf{x} | \mathbf{x}, \lambda) \rangle_{g(\lambda, \pi)} + \langle \ln p(\mathbf{x} | \pi) \rangle_{g(\lambda, \pi)} + \text{const.} \\ &= \sum_{n=1}^N \left\{ \underbrace{\langle \ln p(x_n | \mathbf{x}_n, \lambda) \rangle_{g(\lambda)}}_{\pi \text{ は積分が消え}} + \underbrace{\langle \ln p(x_n | \pi) \rangle_{g(\pi)}}_{\lambda \text{ は積分が消え}} \right\} + \text{const.} \quad (4.47) \end{aligned}$$

… x_1, x_2, \dots, x_N 独立な分布 $g(x_1), \dots, g(x_N)$ に分解。

ここで、 λ は半正、期待値 $\mathbb{E}[\lambda]$ は。

$$\begin{aligned} &\langle \ln p(x_n | \mathbf{x}_n, \lambda) \rangle_{g(\lambda)} \\ &= \int \cdots \int \ln \prod_{k=1}^K P_{0i}(x_n | \lambda_k)^{S_{n,k}} g(\lambda_k) d\lambda_1 \cdots g(\lambda_K) d\lambda_K \quad (\because (4.28)) \\ &= \sum_{k=1}^K \int_R S_{n,k} \ln P_{0i}(x_n | \lambda_k) g(\lambda_k) d\lambda_k \quad (\because \text{幾何形}) \\ &= \sum_{k=1}^K \int_R S_{n,k} \left(x_n \ln \lambda_k - \frac{\lambda_k^{x_n} e^{-\lambda_k}}{x_n!} \right) g(\lambda_k) d\lambda_k \\ &= \sum_{k=1}^K S_{n,k} \left(x_n \langle \ln \lambda_k \rangle - \langle \lambda_k \rangle \right) + \text{const.} \quad (4.48) \end{aligned}$$

よって、

$$\begin{aligned} &\langle \ln p(\mathbf{x}_n | \pi) \rangle_{g(\pi)} \\ &= \int \cdots \int \ln \frac{C_{0i}(\mathbf{x}_n | \pi)}{\prod_{k=1}^K \pi_k^{S_{n,k}}} g(\pi) d\pi \quad (\because (4.2)) \end{aligned}$$

$$\begin{aligned} &= \frac{1}{\prod_{k=1}^K} \int \ln \prod_{k=1}^K \pi_k^{S_{n,k}} g(\pi_k) d\pi_k \\ &= \sum_{k=1}^K S_{n,k} \langle \ln \pi_k \rangle \quad (4.49) \end{aligned}$$

である。

(4.48), (4.49) より、 \mathcal{S}_n の分布はカテゴリ分布.

$$g(\mathcal{S}_n) = C_{\alpha} \cdot (\mathcal{S}_n | \eta_n) \quad (4.50)$$

where

$$\eta_n \propto \exp \left\{ \lambda_n \langle \ln \lambda_k \rangle - \langle \lambda_k \rangle + \langle \ln \lambda_k \rangle \right\}$$

$$\text{s.t. } \sum_{k=1}^K \eta_{n,k} = 1$$

Remark

- 更新式の完成までの流れやπに対する期待値計算が必要.
- … $g(\lambda)$ や $g(\pi)$ の形式が明確 (=?) でないことが必要,
- … 一回後回し

2. 1° マルコフ(=対角線)分布の更新式

$$\begin{aligned} \ln g(\lambda, \pi) &= \langle \ln p(X, \mathcal{S}, \lambda, \pi) \rangle_{g(\mathcal{S})} + \text{const.} \quad (\because (4.25)) \\ &= \langle \ln p(X|\mathcal{S}, \lambda) + \ln p(\mathcal{S}|\pi) + \underbrace{\ln p(\lambda) + \ln p(\pi)}_{g(\mathcal{S}) \text{ が独立}} \rangle_{g(\mathcal{S})} + \text{const.} \quad (\because (4.39)) \\ &= \underbrace{\langle \ln p(X|\mathcal{S}, \lambda) \rangle_{g(\mathcal{S})} + \ln p(\lambda)}_{\lambda \text{ が } \mathcal{S}} + \underbrace{\langle \ln p(\mathcal{S}|\pi) \rangle_{g(\mathcal{S})} + \ln p(\pi)}_{\pi \text{ が } \mathcal{S}} + \text{const.} \quad (4.52) \end{aligned}$$

$$\rightsquigarrow g(\lambda, \pi) = g(\lambda) g(\pi) \quad (\text{独立})$$

λ の独立性.

$$\ln g(\lambda)$$

$$\begin{aligned} &= \sum_{n=1}^N \langle \ln p(x_n | \mathcal{S}_n, \lambda) \rangle_{g(\mathcal{S}_n)} + \sum_{k=1}^K \ln \text{Gam}(\lambda_k | a, b) + \text{const.} \\ &\approx \dots \end{aligned}$$

$$\langle \ln p(x_n | \mathcal{S}_n, \lambda) \rangle_{g(\mathcal{S}_n)}$$

$$= \left\langle \ln \prod_{k=1}^K \text{Poi}(x_n | \lambda_k)^{s_{n,k}} \right\rangle_{g(\mathcal{S}_n)} \quad (\because (4.28))$$

$$= \int \ln \prod_{k=1}^K \text{Poi}(x_n | \lambda_k)^{s_{n,k}} g(\mathcal{S}_n) d\mathcal{S}_n$$

$$= \sum_{k=1}^K \int \ln \frac{\text{Poi}(x_n | \lambda_k)^{s_{n,k}}}{\frac{\lambda_k^{x_n}}{x_n!} e^{-\lambda_k}} g(\mathcal{S}_n) d\mathcal{S}_n \quad (\because \text{独立性})$$

$$= \sum_{k=1}^K \langle s_{n,k} (\ln \lambda_k - \ln x_n! - \lambda_k) \rangle_{g(\mathcal{S}_n)}$$

$$= x_n \sum_{k=1}^K \ln \lambda_k \langle s_{n,k} \rangle - \ln x_n! \sum_{k=1}^K \langle s_{n,k} \rangle - \sum_{k=1}^K \lambda_k \langle s_{n,k} \rangle$$

よし

$$\ln \frac{\text{Gam}(\lambda_k | a, b)}{\frac{b^a}{\Gamma(a)} \lambda_k^{a-1} e^{-b\lambda_k}} = \underbrace{\ln b^a - \ln \Gamma(a) + (a-1) \ln \lambda_k - b\lambda_k}_{\text{const.}}$$

∴

$$\ln g(\lambda)$$

$$= \sum_{n=1}^N \left(x_n \sum_{k=1}^K \ln \lambda_k \langle S_{n,k} \rangle - \sum_{k=1}^K \lambda_k \langle S_{n,k} \rangle \right) + \sum_{k=1}^K (a-1) \ln \lambda_k - b \lambda_k + \text{const.}$$

$$= \sum_{k=1}^K \left(\left(\sum_{n=1}^N x_n \langle S_{n,k} \rangle + a-1 \right) \ln \lambda_k - \left(\sum_{n=1}^N \langle S_{n,k} \rangle + b \right) \lambda_k \right) + \text{const.} \quad (4.53)$$

$\Rightarrow \lambda$ が 26 件事後分布は K 個の独立なガウス分布。

$$g(\lambda_k) = \text{Gam}(\lambda_k | \hat{a}, \hat{b}) \quad (4.54)$$

$$\begin{aligned} \hat{a}_k &= \sum_{n=1}^N x_n \langle S_{n,k} \rangle + a \\ \hat{b}_k &= \sum_{n=1}^N \langle S_{n,k} \rangle + b \end{aligned} \quad (4.55)$$

π も λ と同様。

$$\ln g(\pi) = \sum_{n=1}^N \langle \ln (\delta_n | \pi) \rangle_{g(\delta_n)} + \ln p(\pi) + \text{const.}$$

$$= \sum_{n=1}^N \left\langle \ln \text{Cat}(S_n | \pi) \right\rangle_{g(\delta_n)} + \ln \frac{\text{Dir}(\pi | \alpha)}{C_0(\alpha) \prod_{k=1}^K \pi_k^{\alpha_{k-1}}} + \text{const.} \quad (\because (4.2), (4.4))$$

∴

$$\begin{aligned} &\left\langle \ln \text{Cat}(S_n | \pi) \right\rangle_{g(\delta_n)} \\ &= \int \ln \prod_{k=1}^K \pi_k^{S_{n,k}} g(\delta_n) d\delta_n \\ &= \sum_{k=1}^K \int S_{n,k} \ln \pi_k g(\delta_n) d\delta_n. \quad (\because \text{統計的}) \\ &= \sum_{k=1}^K \ln \pi_k \langle S_{n,k} \rangle_{g(\delta_n)} \end{aligned}$$

よって

$$\ln \text{Dir}(\pi | \alpha) = \underbrace{\ln C_0(\alpha)}_{\text{const.}} + \sum_{k=1}^K (\alpha_{k-1}) \ln \pi_k \quad \therefore$$

$$\begin{aligned} \langle \ln g(\pi) \rangle &= \sum_{n=1}^N \left(\sum_{k=1}^K \ln \pi_k \langle S_{n,k} \rangle \right) + \sum_{k=1}^K (\alpha_{k-1}) \ln \pi_k + \text{const.} \\ &= \sum_{k=1}^K \left(\sum_{n=1}^N \langle S_{n,k} \rangle + \alpha_{k-1} \right) \ln \pi_k + \text{const.} \quad (4.56) \end{aligned}$$

∴ π は π の分布を有する。

$$g(\pi) = \text{Dir}(\pi | \hat{\alpha}) \quad (4.57)$$

$$\begin{aligned} \hat{\alpha}_k &= \sum_{n=1}^N \langle S_{n,k} \rangle + \alpha_k \\ &= \sum_{n=1}^N x_n \langle S_{n,k} \rangle + \alpha_k \end{aligned} \quad (4.58)$$

~ サーチング分布と事前分布と同じ種類。

$\gamma(\lambda)$, $\gamma(\pi)$ の更新 (\approx 12. 期待値 $\langle \gamma_{n,k} \rangle = \langle S_{n,k} \rangle_{\gamma(\beta_n)}$ が必須。
いま $\gamma(\beta_n) = \text{Cat}(\beta_n | \eta_n)$ (\because 4.50)

$$\text{左} \cdot \langle S_{n,k} \rangle_{\gamma(\beta_n)} = \eta_{n,k} \quad (4.59)$$

で次。

2段階、期待値 $\langle \lambda \rangle$, $\langle \ln \lambda \rangle$, $\langle \ln \pi \rangle$ を考え。

$\gamma(\lambda) = \text{ガム分布}$. $\gamma(\pi) = \text{ディリクレ分布}$ が。

$$\langle \lambda_k \rangle = \frac{\hat{a}_k}{\hat{b}_k} \quad (\because (2.59)) \quad (4.60)$$

$$\langle \ln \lambda_k \rangle = \psi(\hat{a}_k) - \ln \hat{b}_k \quad (\because (2.60)) \quad (4.61)$$

$$\langle \ln \pi_k \rangle = \psi(\hat{\alpha}_k) - \psi\left(\sum_{i=1}^K \hat{\alpha}_i\right) \quad (\because (2.52)) \quad (4.62)$$

where $\psi(\cdot)$ is digamma function.

Remark

期待値の対数 ≠ 対数の期待値

ホーリーは完全モードの下で γ の実効性論をアーティスティックな示す。

$\gamma(\lambda), \gamma(\pi)$ の初期化。

For $i = 1$ to MAXITER do

For $n = 1$ to N do

Update $\gamma(\beta_n)$ using (4.50).

End For

For $k = 1$ to K do

Update $\gamma(\lambda_k)$ using (4.54).

End For

Update $\gamma(\pi)$ using (4.57).

End For

Remark

• 每回の更新が π の下で完全モードの ELBO で実行される。
進行度を追跡する方法も。

• 更新の際の期待値計算は $\pi = \gamma$ に依存するが、 $\pi = \gamma$ は $\pi = \gamma$,

4.3.4. 前工場型 ギガスサニゴリ=7°

通常、前工場型 ギガスサニゴリ=7°

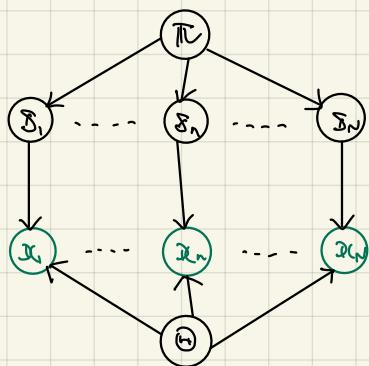
… 同時分布から1度×→を周辺化除去。

$$p(X, S) = \iint p(X, S, \lambda, \pi) d\lambda d\pi \quad (4.63)$$

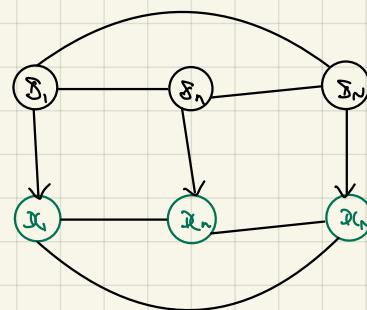
→ S を $p(S|X)$ が S ニゴリで互換性よ。

1度×→の事後分布は、必要に応じてサニゴリエキス X から計算でいる。

混合モデル



周辺化エキスモデル



… 各 $\delta_1, \dots, \delta_N$ は互いに依存関係を持つ。(完全グラフ)

→ 面積同時 $\delta_1, \dots, \delta_N$ は $\pi = 7^{\circ}$ するため $\delta_1, \dots, \delta_N$ は互いに独立、組み合わせをすべて計算して正規化が必要がある、
… 非現実

周辺化エキスモデルの事後分布 \Leftarrow
 $p(S|X)$ は $\delta_1, \dots, \delta_N$ が
互換性。(各 $\delta_1, \dots, \delta_N$ が
1度×→で互換性か?)

→ $\pi = 7^{\circ}$ (たとえ δ_n が δ_{n+1} すべて
潜在変数の集合
 $\delta_n = \{\delta_1, \dots, \delta_{n-1}, \delta_{n+1}, \dots, \delta_N\}$
 $=$ δ_n の条件分布
 $p(\delta_n | X, \delta_{-n})$ が十分に簡単な確率分布とし得るればよ)

$$p(\delta_n | X, \delta_{-n})$$

$$= \frac{p(x_n, \delta_n, \delta_{-n})}{p(x_n, \delta_{-n})} \quad (4.64)$$

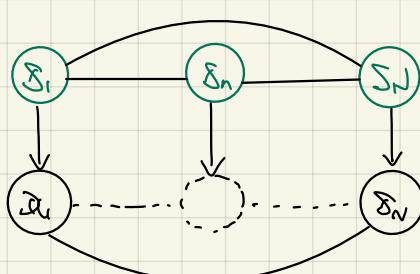
δ_n に直接関係づく。

$$p(X_{-n}, \delta_n, \delta_{-n})$$

$$\propto p(x_n | X_{-n}, \delta_n, \delta_{-n}) p(x_{-n} | \delta_n, \delta_{-n}) p(\delta_n | \delta_{-n}) \quad (4.65)$$

$$\propto p(x_n | X_{-n}, \delta_n, \delta_{-n}) p(\delta_n | \delta_{-n})$$

$$(4.66)$$



δ_n は X_{-n} と共同観の関係 (δ_n と X_{-n} が観察される)。

$$\Rightarrow p(X_{-n} | \delta_n, \delta_{-n}) = p(X_{-n} | \delta_n) \quad (4.67)$$

よ、7. $p(\delta_n | \mathcal{X}_{-n}, \delta_{-n})$ (2つめの2つめ)を分解せよ.

$$p(\delta_n | \delta_{-n}) \quad (4.68)$$

$$p(x_n | \mathcal{X}_{-n}, \delta_n, \delta_{-n}) \quad (4.69)$$

①いま計算(たまご)で、 δ_n は角子離散分布.

(4.68)、(4.69)に対して、各 k で $S_{n,k} = 1$ となる場合の値を計算し、最後に足し算(たなみ)る(近似化する)。 δ_n をサニフーリするための確率(カテゴリ分布)が得られる。

1. (4.68)について、下記のように δ_n の予測分布でみて解説です。

$$\begin{aligned} p(\delta_n | \delta_{-n}) &= \int p(\delta_n, \pi | \delta_{-n}) d\pi \\ &= \int \frac{p(\delta_n, \pi, \delta_{-n})}{p(\delta_{-n})} d\pi \\ &= \int \frac{p(\delta_n, \delta_{-n} | \pi) p(\pi)}{p(\delta_{-n})} d\pi \quad p(\delta_{-n}, \pi) \\ &= \int \frac{p(\delta_n | \pi) \frac{p(\delta_{-n} | \pi) p(\pi)}{p(\delta_{-n})}}{p(\delta_{-n})} d\pi \quad (\because \text{独立}) \\ &= \int p(\delta_n | \pi) p(\pi | \delta_{-n}) d\pi \quad (4.70) \end{aligned}$$

Cat($\delta_n | \pi$)

N-1個のπによる δ_{-n}
を観測したときの
板、T₂までの「πの事後分布」

ベイズ的定理よ。

$$p(\pi | \delta_{-n}) \propto p(\delta_{-n} | \pi) p(\pi) \quad (4.71)$$

~3章、カテゴリ分布とディリクレ分布を用いた二乗法による予測と同じ。
(3.27) を用いよと解説的におぼえよ。

$$p(\pi | \delta_{-n}) = \text{Dir}(\pi | \hat{\alpha}_{-n}) \quad (4.72)$$

where

$$\hat{\alpha}_{-n,k} = \sum_{n' \neq n} S_{n',k} + \alpha_k \quad (4.73)$$

$p(\pi | \delta_{-n})$ の分布が分かって、(4.70) のように π を近似化除去すれば $p(\delta_n | \delta_{-n})$ が求まる。

~3章、結果を用いよ。

$$\begin{aligned} p(\delta_n | \delta_{-n}) &= \int \text{Cat}(\delta_n | \pi) \text{Dir}(\pi | \hat{\alpha}_{-n}) d\pi \\ &= \text{Cat}(\delta_n | \eta_{-n}) \quad (\because (3.32)) \quad (4.74) \end{aligned}$$

where $\eta_{-n,k} \propto \hat{\alpha}_{-n,k}$ (4.75)

2. (4.69) (左) 分布と右分布の関係論.

$$\begin{aligned}
 & p(x_n | X_{-n}, S_n, S_{-n}) \\
 &= \int p(x_n, \lambda | X_{-n}, S_n, S_{-n}, \lambda) d\lambda \\
 &= \int \frac{p(x_n, X_{-n}, S_n, S_{-n}, \lambda)}{p(X_{-n}, S_n, S_{-n})} d\lambda \\
 &= \int \frac{p(x_n, X_{-n}, S_n, S_{-n} | \lambda) p(\lambda)}{p(X_{-n}, S_n, S_{-n})} d\lambda \\
 &= \int \frac{p(x_n, S_n | \lambda) p(X_{-n}, S_{-n} | \lambda) p(\lambda)}{p(X_{-n}, S_n, S_{-n})} d\lambda \quad (\because \text{独立性}) \\
 &= \int p(x_n | S_n, \lambda) p(S_n) \frac{p(X_{-n}, S_{-n}, \lambda)}{p(S_n) p(X_{-n}, S_{-n})} d\lambda \\
 &= \int p(x_n | S_n, \lambda) p(\lambda | X_{-n}, S_{-n}) d\lambda \quad (4.76)
 \end{aligned}$$

$\prod_{k=1}^K \text{Poi}(x_n | \lambda_k)^{S_{n,k}}$ 各後分布(左より右)

$$\prod_{n' \neq n} \prod_{k=1}^K \text{Poi}(x_{n'} | \lambda_k)^{S_{n',k}}$$

ベイズの定理より、

$$p(\lambda | X_{-n}, S_{-n}) \propto p(X_{-n} | S_{-n}, \lambda) p(\lambda) \quad (4.77)$$

→ ポisson 分布とガーベル分布の関係論

$$\begin{aligned}
 & \ln p(\lambda | X_{-n}, S_{-n}) \frac{\lambda^{x_n}}{x_n!} e^{-\lambda} \\
 &= \sum_{n' \neq n} \sum_{k=1}^K S_{n',k} \ln \text{Poi}(x_{n'} | \lambda_k) + \sum_{k=1}^K \ln \text{Gam}(\lambda_k | a, b) \\
 &= \sum_{n' \neq n} \sum_{k=1}^K S_{n',k} (x_{n'} \ln \lambda_k - \ln x_{n'}! - \lambda_k) + \sum_{k=1}^K \left(\ln \frac{b^a}{T(a)} + (a-1) \ln \lambda_k - b \lambda_k \right) \\
 &= \sum_{k=1}^K \left\{ \left(\sum_{n' \neq n} S_{n',k} x_{n'} + a-1 \right) \ln \lambda_k - \left(\sum_{n' \neq n} S_{n',k} + b \right) \lambda_k \right\} + \text{const.} \quad (4.78)
 \end{aligned}$$

左, 右 = 2 分布の関係.

$$p(\lambda | X_{-n}, S_{-n}) = \prod_{k=1}^K \text{Gam}(\lambda_k | \hat{a}_{-n,k}, \hat{b}_{-n,k}) \quad (4.79)$$

where

$$\begin{aligned}
 \hat{a}_{-n,k} &= \sum_{n' \neq n} S_{n',k} x_{n'} + a \\
 \hat{b}_{-n,k} &= \sum_{n' \neq n} S_{n',k} + b
 \end{aligned} \quad (4.80)$$

27(2). この分布を用いて $\pi = \gamma^*$ の積分消去. x_n に対する種の予測分布を計算する.

$S_{n,k}$ / of k .

式 k は $S_{n,k} = 1$ の場合のみを考える. (4.76) より,

$$\begin{aligned} p(x_n | X_{-n}, S_{n,k} = 1, \delta_{-n}) &= \int p(x_n | S_{n,k} = 1, \lambda) p(\lambda | X_{-n}, \delta_{-n}) d\lambda \\ &= \int p(x_n | \lambda_k) p(\lambda_k | X_{-n}, \delta_{-n}) d\lambda_k \\ &= \int \text{Poi}(x_n | \lambda_k) \text{Gam}(\lambda_k | \hat{a}_{-n,k}, \hat{b}_{-n,k}) d\lambda_k \quad (\because (4.27), (4.79)) \\ &= NB(x_n | \hat{a}_{-n,k}, \frac{1}{\hat{b}_{-n,k} + 1}) \quad (\because (3.43)) \quad (4.81) \end{aligned}$$

次上式.

(4.68), (4.69) の確率分布は,

$$p(S_n | \delta_{-n}) = \text{Cat}(S_n | \eta_{-n}) \quad (\because (4.14), (4.15))$$

where $\eta_{-n} \propto \hat{\alpha}_{n,k}$

$$p(x_n | X_{-n}, S_n, \delta_{-n}) = \prod_{k=1}^K NB(x_n | \hat{a}_{-n,k}, \frac{1}{\hat{b}_{-n,k} + 1})^{S_{n,k}} \quad (\because (4.81))$$

でよし.

~ 実際にはこの確率分布が S_n で $\pi = \gamma^*$ で $\gamma = \gamma^*$ と T_2 の δ_{-n} , S_n の実現値 ($S_n = (1, 0, \dots, 0)^T, \dots, (0, \dots, 0, 1)^T$) まで (4.74), (4.81) に用いられる. 得られた T_2 を個々の $\pi = \gamma^*$ で正規化.

→ S_n で $\pi = \gamma^*$ とするために $\pi = \gamma^*$ が $\pi = \gamma^*$ 分布が得られる.

$$\begin{aligned} p(S_n | X, \delta_{-n}) &\propto p(x_n | X_{-n}, S_n, \delta_{-n}) p(S_n | \delta_{-n}) \\ &= \left\{ \prod_{k=1}^K NB(x_n | \hat{a}_{-n,k}, \frac{1}{\hat{b}_{-n,k} + 1})^{S_{n,k}} \right\} \text{Cat}(S_n | \eta_{-n}) \\ &= \left\{ \prod_{k=1}^K NB(x_n | \hat{a}_{-n,k}, \frac{1}{\hat{b}_{-n,k} + 1}) \right\} \eta_{-n,k}^{S_{n,k}} \quad \text{二本を $\pi = \gamma^*$ で $\pi = \gamma^*$ 分布} \end{aligned}$$

③ 前回型モデルで $\pi = \gamma^* = \gamma^*$ では、各 n 番目 $\pi = \gamma^*$ 点で $\pi = \gamma^* = \gamma^*$. 每回 $\pi = \gamma^* = \gamma^*$ は $N - 1$ 回 $\pi = \gamma^*$ や潜在変数 $\pi = \gamma^*$ に属する足し合わせを実行してみよう(必要はなし).

e.g., 3番目 $\pi = \gamma^*$ を得た後で 4番目 $\pi = \gamma^*$ を得る.

事後分布の計算 (4.73) 12.

$$\hat{\alpha}_{-j,k} = \hat{\alpha}_{-i,k} + \underline{S_{i,k}} - \underline{S_{j,k}} \quad (4.82)$$

$\pi = \gamma^* = \gamma^*$
12倍, 12重

実際 $\pi = \gamma^* = \gamma^*$
 $\pi = \gamma^* = \gamma^*$
 $S_{i,k}$

過去 $\pi = \gamma^*$
 $\pi = \gamma^*$

現在古 $S_{j,k}$

(4.80) についても同様.

$$\begin{aligned}\hat{a}_{-j,k} &= \hat{a}_{-i,k} + s_{i,k} x_i - s_{j,k} x_j \\ \hat{b}_{-j,k} &= \hat{b}_{-i,k} + s_{i,k} - s_{j,k}\end{aligned}\quad (4.83)$$

⑨ 実現上は、N個分の $\hat{\alpha}_k$ は計算上 $\hat{\alpha}_k$, \hat{a}_k , \hat{b}_k を複数上へ保存。

→ 新しい s_n が $=70\%$ で直前の -2 入前からの値 $\hat{\alpha}_k$ は更新量を $\hat{\alpha}_k$, \hat{a}_k , \hat{b}_k から一時的に引く。

→ s_n が $=70\%$

→ 新たに得た t_2 量を加える。

Algorithm : ホーリン混合モデルの t_2 の崩壊型ガウス $\sigma = 70\% = 9^\circ$

潜在変数 x_i が $=70\%$ で x_1, \dots, x_N 初期値を設定。

$\hat{\alpha}$, \hat{a} , \hat{b} を計算

For $i = 1$ to MAXITER do

For $n = 1$ to N do

(4.82), (4.83) を用いて x_n の肉可了係数を計算を除く。

For $k = 1$ to K do

(4.81) を用いて $p(x_n | x_{-n}, s_{n,k} = 1, \hat{s}_n)$ を計算。

EndFor

(4.74) やよし $p(x_n | x_{-n}, \hat{s}_n, \hat{s}_{-n})$ を用いて \hat{s}_n を $=70\%$

(4.82) やよし (4.83) を用いて x_n の肉可了係数を追加。

EndFor

EndFor