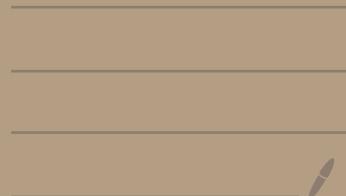


ベイズ推論

3.3



3.3. 128元がウス分布の確率.

$$\mathcal{N}(x | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma^2} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

… $\sim N(\mu, \sigma^2)$

$\leadsto \begin{cases} \text{① 平均値 } \mu \text{ が } \mathbb{R} \text{ の確率} \\ \text{② 分散 } \sigma^2 \text{ が } \mathbb{R} \text{ の確率} \\ \text{③ 両方 } (\mu, \sigma^2) \text{ の確率} \end{cases}$

二二で、精度 (precision) $\sim N(\mu, \sigma^2)$ を導入可.

3.3.1. 平均が未知の場合.

仮定

$\sim N(\mu, \sigma^2)$: 正規分布. N 個

… 精度 $\sim N(\mu, \sigma^2)$ $\lambda \in \mathbb{R}_+$: 固定.

$\leadsto \mu$ が事前分布を設定.

観測値 $x \in \mathbb{R}$ の分布

$$p(x | \mu) = \mathcal{N}(x | \mu, \lambda^{-1}) \left(= \sqrt{\frac{\lambda}{2\pi}} \exp\left(-\frac{\lambda(x-\mu)^2}{2}\right) \right) \quad (3.47)$$

μ が事前分布

$$p(\mu) = \mathcal{N}(\mu | m, \lambda_m^{-1}) \left(= \sqrt{\frac{\lambda_m}{2\pi}} \exp\left(-\frac{\lambda_m(\mu-m)^2}{2}\right) \right) \quad (3.48)$$

where

$m \in \mathbb{R}$, $\lambda_m \in \mathbb{R}_+$ are hyper parameters.

Observed Data

$$X = \{x_1, \dots, x_N\}$$

$$\begin{aligned}
 p(\mu | \mathbf{x}) &\propto p(\mathbf{x} | \mu) p(\mu) \quad (\because \text{Bayes}) \\
 &= \left(\prod_{i=1}^N p(x_i | \mu) \right) p(\mu) \\
 &= \prod_{i=1}^N N(x_i | \mu, \lambda^{-1}) N(\mu | m, \lambda_m^{-1}) \quad (3.47)
 \end{aligned}$$

对数似然

$$\begin{aligned}
 \ln p(\mu | \mathbf{x}) &= \sum_{i=1}^N \ln N(x_i | \mu, \lambda^{-1}) \\
 &\quad + \ln N(\mu | m, \lambda_m^{-1}) + \text{const.} \\
 &= \sum_{i=1}^N \ln \sqrt{\frac{\lambda}{2\pi}} \exp\left(-\frac{(x_i - \mu)^2}{2\lambda}\right) \\
 &\quad + \ln \sqrt{\frac{\lambda_m}{2\pi}} \exp\left(-\frac{\lambda_m(\mu - m)^2}{2}\right) + \text{const.} \\
 &= \sum_{i=1}^N \left[\frac{1}{2} (\ln \lambda - \ln 2\pi) - \frac{\lambda (x_i - \mu)^2}{2} \right] \\
 &\quad + \frac{1}{2} (\ln \lambda_m - \ln 2\pi) - \frac{\lambda_m (\mu - m)^2}{2} + \text{const.} \\
 &= -\frac{\lambda}{2} \sum_{i=1}^N (x_i^2 - 2x_i \mu + \mu^2) \\
 &\quad - \frac{\lambda_m}{2} (\mu^2 - 2\mu m + m^2) + \text{const.} \\
 &= -\frac{1}{2} \left[\underbrace{(\lambda N + \lambda_m)}_{+ \text{const.}} \mu^2 - 2 \left(\underbrace{\sum_{i=1}^N x_i}_{+ \text{const.}} \lambda + m \lambda_m \right) \mu \right]
 \end{aligned}$$

(3.50)

$\Rightarrow \mu$ は関数上に凸な2次関数。

事後分布が2次元ガウス分布となる。

$$p(\mu | \mathbf{x}) = N(\mu | \hat{m}, \hat{\lambda}_\mu^{-1}) \left(= \sqrt{\frac{\hat{\lambda}_\mu}{2\pi}} \exp\left(-\frac{\hat{\lambda}_\mu(\mu - \hat{m})^2}{2}\right) \right) \quad (3.51)$$

対数をとる。

$$\begin{aligned} \ln p(\mu | \mathbf{x}) &= \frac{1}{2} (\ln \hat{\lambda}_\mu - \ln 2\pi) - \frac{\hat{\lambda}_\mu}{2} (\mu - \hat{m})^2 \\ &= -\frac{1}{2} \left(\hat{\lambda}_\mu \mu^2 - 2\hat{m} \hat{\lambda}_\mu \mu \right) + \text{const.} \end{aligned} \quad (3.52)$$

(3.50), (3.52) より。

$$\hat{\lambda}_\mu = N\lambda + \lambda_\mu \quad (3.53)$$

$$\hat{m} = \frac{\lambda \sum_{i=1}^N x_i + \lambda_\mu m}{\hat{\lambda}_\mu} \quad (3.54)$$

が得られる。

② 平均 μ の 事後分布の精度 $\hat{\lambda}_\mu$

… 事前分布の精度は $N\lambda$ だけ加えたも。

∴ データ N が大きくなると事後分布の精度が大きくなる。

③ 平均 μ の 事後分布の平均 \hat{m}

… 事前分布によると知識 m と観測データの和 $\sum_{i=1}^N x_i$ の重み合計。

⇒ データを観測すると、事前分布の平均 μ の影響はうすく、
データによると計算された平均値 $\frac{1}{N} \sum_{i=1}^N x_i$ が支配的。

λ_μ が小さくてもデータが支配的。

④ 未観測データ x_* は次の予測分布,

$$p(x_*) = \int p(x_* | \mu) p(\mu) d\mu$$
$$= \int N(x_* | \mu, \lambda^{-1}) N(\mu | m, \lambda_m^{-1}) d\mu \quad (3.55)$$

⇒ 条件付き確率の計算算子.

⑤ 予測分布 $p(x_*)$ と事前分布 $p(\mu)$ の間に因縁.

$$p(\mu | x_*) = \frac{p(x_* | \mu) p(\mu)}{p(x_*)} \quad (\because \text{Bayes}) \quad (3.56)$$

対応式 \propto , $p(x_*)$ は内側に左へ.

$$\ln p(\mu | x_*) = \ln p(x_* | \mu) + \underbrace{\ln p(\mu)}_{\text{const.}} - \ln p(x_*)$$
$$\therefore \ln p(x_*) = \ln p(x_* | \mu) - \ln p(\mu | x_*) + \underbrace{\text{const.}}_{\text{const.}} \quad (3.57)$$

(3.57) はおいた. $p(\mu | x_*)$ は, (3.51) が $N \leftarrow 1$

かつ λ が λ_0 に等しい.

$$\therefore p(\mu | x_*) = N(\mu | \underline{m(x_*)}, \bar{\lambda}) \quad (3.58)$$

where

平均値 $m(x_*)$ の対応式 (3.47), (3.51).

$$\bar{\lambda} = \lambda + \lambda_m$$

$$m(x_*) = \frac{\lambda x_* + \lambda_m m}{\lambda + \lambda_m} \quad (3.59)$$

つまり, $p(x_* | \mu)$ は内側に左へ, (3.47) や.

5, 7.

$$\ln p(x_*) = \ln p(x_* | \mu) - \ln(\mu | x_*) + \text{const.}$$

$$= \ln N(x_* | \mu, \lambda^{-1}) - \ln N(\mu | m(x_*), \lambda + \lambda_m)$$

$$= \ln \left[\sqrt{\frac{\lambda}{2\pi}} \exp \left(-\frac{\lambda(x_* - \mu)^2}{2} \right) \right] + \text{const.}$$

$$- \ln \left[\sqrt{\frac{\lambda + \lambda_m}{2\pi}} \exp \left(-\frac{(\lambda + \lambda_m)(\mu - m(x_*))^2}{2} \right) \right] + \text{const.}$$

$$= \frac{1}{2} (\ln \lambda - \ln 2\pi) - \frac{\lambda}{2} (x_* - \mu)^2$$

$$- \frac{1}{2} (\ln(\lambda + \lambda_m) - \ln 2\pi) - \frac{1}{2} (\lambda + \lambda_m) (\mu - m(x_*))^2$$

$$= -\frac{1}{2} \left[\lambda (x_* - \mu)^2 - (\lambda + \lambda_m) (\mu - m(x_*))^2 \right] + \text{const.}$$

$$= -\frac{1}{2} \left[\lambda (x_*^2 - 2x_* \mu + \mu^2) - (\lambda + \lambda_m) \left(\mu - \frac{\lambda x_* + \lambda_m \mu}{\lambda + \lambda_m} \right)^2 \right] + \text{const.}$$

$$= -\frac{1}{2} \left[\lambda x_*^2 - 2\cancel{\lambda M x_*} + \cancel{\lambda \mu^2} \right. + \left. \frac{\lambda^2 x_*^2 + 2\lambda \lambda_m M x_* + \lambda^2 \mu^2}{\lambda + \lambda_m} \right] - \frac{(\lambda x_* + \lambda_m \mu)^2}{\lambda + \lambda_m}$$

$$= -\frac{1}{2} \left[\left(\lambda - \frac{\lambda^2}{\lambda + \lambda_m} \right) x_*^2 - \frac{2m\lambda\lambda_m}{\lambda + \lambda_m} x_* \right] + \text{const.}$$

$$= -\frac{1}{2} \left[\frac{\lambda\lambda_m}{\lambda + \lambda_m} x_*^2 - \frac{2m\lambda\lambda_m}{\lambda + \lambda_m} x_* \right] + \text{const.}, \quad (3.60)$$

よ、 λ 、予測分布 $p(\lambda_*)$ は。

$$p(\lambda_*) = N(\lambda_* | \mu_*, \lambda_*^{-1}) \quad (3.61)$$

where

$$\begin{cases} \lambda_* = \frac{\lambda \lambda_m}{\lambda + \lambda_m} \\ \mu_* = m \end{cases} \quad (3.62)$$

==

$$\frac{1}{\lambda_*} = \frac{\lambda + \lambda_m}{\lambda \lambda_m} = \frac{1}{\lambda_m} + \frac{1}{\lambda} \quad (3.63)$$

\Rightarrow 予測分布の平均 (\bar{x}) : 事前分布の平均
" 不確実性 : 觀測分布と事前分布の不確実性

★ $\bar{x} \rightarrow$ 觀測後、予測分布 $p(\lambda_* | \bar{x})$
 $\cdots m \leftarrow \hat{m}, \lambda_m \leftarrow \hat{\lambda}_m$

\therefore 予測分布の分布 $\lambda_*^{-1}, \lambda_*^{-1} + \lambda_m^{-1}$.

① $\bar{x} \rightarrow$ m の精度 λ_m は

$\bar{x} \rightarrow$ が土算入るほど大きくなる

but.

$\bar{x} \rightarrow$ 生成自体に精度入る、変わらぬ。

② λ の分布も考慮する場合、かくこと事前分布 $p(\lambda)$ を追加して推論。

$$\ln N(\lambda_* | \mu_*, \lambda_*^{-1}) \\ = -\frac{1}{2} [\lambda_* \lambda_*^2 - 2\mu_* \lambda_* + \text{const.}]$$

よし。

$$\lambda_* = \frac{\lambda \lambda_m}{\lambda + \lambda_m}, \mu_* \lambda_* = \frac{m \lambda \lambda_m}{\lambda + \lambda_m}$$

3.3.2. 精度が未知

仮定

• μ : 已知

• λ : 未知 \Rightarrow データの尤度.

観測) モデル

$$p(x | \lambda) = N(x | \mu, \lambda^{-1}) \left(= \sqrt{\frac{\lambda}{2\pi}} \exp\left(-\frac{\lambda(x-\mu)^2}{2}\right) \right)$$

支承式で明示。

(3.64)

λ の事前分布: ガンマ分布

$$p(\lambda) = G_{am}(\lambda | a, b) \left(= \frac{b^a}{\Gamma(a)} \lambda^{a-1} e^{-b\lambda} \right) \quad (3.65)$$

~ ガンマ分布は出入り分布の精度パラメータの共役事前分布。

2次元、事後分布を求めることが示す。

$$\begin{aligned} p(\lambda | x) &\propto p(x | \lambda) p(\lambda) \\ &= \prod_{i=1}^N p(x_i | \lambda) p(\lambda) \\ &= \prod_{i=1}^N N(x_i | \mu, \lambda^{-1}) G_{am}(\lambda | a, b) \quad (3.66) \end{aligned}$$

対数式となる。

$$\begin{aligned}
 \ln p(\lambda | \mathbf{x}) &= \sum_{i=1}^N \ln \left(\sqrt{\frac{\lambda}{2\pi}} \exp \left(-\frac{\lambda(x_i - \mu)^2}{2} \right) \right) \\
 &\quad + \ln(C_\alpha(a, b) \lambda^{a-1} e^{-b\lambda}) + \text{const.} \\
 &= \frac{N}{2} (\ln \lambda - \ln 2\pi) - \sum_{i=1}^N \frac{\lambda(x_i - \mu)^2}{2} \\
 &\quad + \underline{\ln C_\alpha(a, b)} + (a-1) \ln \lambda - b\lambda + \underline{\text{const.}} \\
 &= \left(\frac{N}{2} + a - 1 \right) \ln \lambda - \left(\frac{1}{2} (x_i - \mu)^2 + b \right) \lambda + \text{const.} \quad \downarrow \\
 &\quad (3.67)
 \end{aligned}$$

∴ $\ln p(\lambda | \hat{a}, \hat{b})$ の対数凸.

$$\ln G_m(\lambda | \hat{a}, \hat{b}) = (\hat{a} - 1) \ln \lambda - \hat{b} \lambda + \text{const.}$$

∴ λ の対数凸を考慮する. 事後分布 $p(\lambda | \mathbf{x})$ は,

$$p(\lambda | \mathbf{x}) = G_m(\lambda | \hat{a}, \hat{b}) \quad (3.68)$$

where

$$\begin{cases} \hat{a} = \frac{N}{2} + a \\ \hat{b} = \frac{1}{2} (x_i - \mu)^2 + b \end{cases} \quad (3.69)$$

ただし、予測分布を考慮し、事前分布を利用した予測分布.

$$p(x_* | \mathbf{x}) = \int p(x_* | \lambda) p(\lambda) d\lambda. \quad (3.70)$$

対数凸と二乗法の近似で算出する方法を用いる.

ベイズの定理.

$$p(\lambda | x_*) = \frac{p(x_* | \lambda) p(\lambda)}{p(x_*)} \quad (3.71)$$

対数凸と二乗法,

$$\begin{aligned}
 \ln p(\lambda | x_*) &= \ln p(x_* | \lambda) + \ln p(\lambda) - \ln p(x_*) \\
 p(x_*) \text{ は既に考慮.}
 \end{aligned}$$

$$\ln p(x_*) = \ln p(x_* | \lambda) - \ln p(\lambda | x_*) + \text{const.} \quad (3.72)$$

∴ $p(\lambda | \lambda_*) \propto 1^{\frac{1}{2}} \lambda_*$ と観測した後 λ の分布を
考えよ。

$$p(\lambda | \lambda_*) = \text{Gam} \left(\lambda \mid \frac{1}{2} + a, b(\lambda_*) \right) \quad (3.73)$$

where $(\because (3.69))$

$$b(\lambda_*) = \frac{1}{2} (\lambda_* - \mu)^2 + b \quad (3.74)$$

∴

$$\ln p(\lambda_*) = \ln p(\lambda_* | \lambda) - \ln p(\lambda | \lambda_*) + \text{const.}$$

$$= \ln \mathcal{N}(\lambda_* | \mu, \lambda^{-1}) - \ln \text{Gam} \left(\lambda \mid \frac{1}{2} + a, b(\lambda_*) \right) + \text{const.}$$

$$= \underbrace{\frac{1}{2} (\ln \lambda - \ln 2\pi) - \frac{\lambda}{2} (\lambda_* - \mu)^2}_{-\left[\ln C_G \left(\frac{1}{2} + a, b(\lambda_*) \right) + \underbrace{\left(\frac{1}{2} + a - 1 \right) \ln \lambda - b(\lambda_*) \lambda} \right] + \text{const.}}$$

$$= -\cancel{\frac{\lambda}{2} (\lambda_* - \mu)^2} - \left[\cancel{\frac{1}{2} (\lambda_* - \mu)^2} + b \right] \lambda - \left(\ln \left(b + \frac{1}{2} (\lambda_* - \mu)^2 \right)^{\frac{1}{2} + a} - \ln T \left(\frac{1}{2} + a \right) \right) \downarrow$$

$$= -\frac{2a+1}{2} \ln b \left(1 + \frac{1}{2b} (\lambda_* - \mu)^2 \right) + \text{const.} \quad + \text{const.}$$

$$= -\frac{2a+1}{2} \ln \left(1 + \frac{1}{2b} (\lambda_* - \mu)^2 \right) + \text{const.}, \quad (3.75)$$

$\sim \lambda \hat{T}_2 - \hat{T} \sim \lambda$ が λ の分布を表す。

$$S_{\lambda}(\lambda | \mu_s, \lambda_s, v_s) = \frac{T(\frac{v_s+1}{2})}{T(\frac{v_s}{2})} \left(\frac{\lambda_s}{\pi v_s} \right)^{\frac{1}{2}} \left[1 + \frac{\lambda_s}{v_s} (\lambda - \mu_s)^2 \right]^{-\frac{v_s+1}{2}} \quad (3.76)$$

(3.74), 対数式とし.

$$\ln S(x | \mu_s, \lambda_s, v_s)$$

$$\begin{aligned} &= \ln T\left(\frac{v_s+1}{2}\right) - \ln T\left(\frac{v_s}{2}\right) + \frac{1}{2} \ln\left(\frac{\lambda_s}{\pi v_s}\right) \\ &\quad - \frac{v_s+1}{2} \ln\left[1 + \frac{\lambda_s}{v_s} (x - \mu_s)^2\right] \\ &= -\frac{v_s+1}{2} \ln\left[1 + \frac{\lambda_s}{v_s} (x - \mu_s)^2\right] + \text{const.} \end{aligned} \quad (3.77)$$

(3.75), (3.77) を比較可.

$$p(x_s) = S(x_s | \mu_s, \lambda_s, v_s) \quad (3.78)$$

where

$$\begin{cases} v_s = 2a \\ \lambda_s/v_s = 1/2b \\ \mu_s = M \end{cases} \therefore \begin{cases} \mu_s = M \\ \lambda_s = a/b \\ v_s = 2a \end{cases} \quad (3.79)$$

3.3.3, 平均・精度が未知の場合.

観測モード

$$p(x | \mu, \lambda) = N(x | \mu, \lambda^{-1}) \left(= \sqrt{\frac{\lambda}{2\pi}} \exp\left(-\frac{\lambda(x-\mu)^2}{2}\right) \right) \quad (3.80)$$

~ 前分布は ガウス、ガニス分布

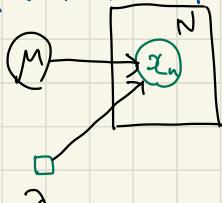
$$\begin{aligned} p(\mu, \lambda) &= NG(\mu, \lambda | m, \beta, a, b) \\ &= N(\mu | m, (\beta \lambda)^{-1}) G_{am}(\lambda | a, b) \end{aligned} \quad (3.81)$$

モードは.

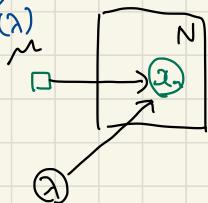
$$p(X, \mu, \lambda) = p(X | \mu, \lambda) p(\mu, \lambda)$$

12次元ガウス分布の導出モデル

$$p(x, \lambda, \mu) = p(x|\mu, \lambda) \\ p(\mu)p(\lambda)$$

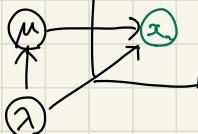


。 平均未知



。 精度未知

$$p(x, \mu, \lambda) \\ = p(x|\mu, \lambda) p(\mu|\lambda) p(\lambda)$$



$$p(x|\mu, \lambda) \\ p(\mu|\lambda)$$

。 平均・精度未知

平均パラメータも

精度パラメータも未知

事後分布を求める。

① μ の事後分布を求める。

μ の事前分布が

$$p(\mu) = \mathcal{N}(\mu | m, (\beta\lambda)^{-1})$$

で与えられる。実際には正確か?

$$p(\mu | \lambda, x) \propto p(x | \mu, \lambda) p(\mu) \quad \text{fix } \lambda, x$$

② $p(\mu, \lambda | x)$

$$= \frac{p(x, \mu, \lambda)}{p(x)} = \frac{p(x, \mu, \lambda)}{p(\lambda|x)} \cdot \frac{p(x)}{p(\lambda|x)}$$

$$= \underbrace{p(\mu | \lambda, x)}_{\text{fix } \lambda, x} \underbrace{p(\lambda | x)}_{\text{fix } x}$$

$$= \prod_{i=1}^N \mathcal{N}(x_i | \mu, \lambda^{-1}) \mathcal{N}(\mu | m, (\beta\lambda)^{-1})$$

(3.49) はよいか $\lambda \leftarrow (\beta\lambda)$ とLT2形。

よって、事後分布は、(3.51), (3.53), (3.54) が

$$\textcircled{2} \quad p(\mu | \lambda, x) = \mathcal{N}(\mu | \hat{m}, \hat{\lambda}^{-1}) = \mathcal{N}(\mu | \hat{m}, (\hat{\beta}\lambda)^{-1}) \quad (3.82)$$

where

$$\hat{\lambda} = N\lambda + \beta\lambda = (N + \beta)\lambda, \quad \hat{\beta} = N + \beta$$

$$\hat{m} = \frac{\lambda \sum_{i=1}^N x_i + m\beta\lambda}{(N + \beta)\lambda} = \frac{1}{\hat{\beta}} \left(\sum_{i=1}^N x_i + \beta m \right) \quad (3.83)$$

である。

② λ の事後分布

同時分布と条件付分布の積とで書き下す。

$$p(X, \lambda, M) = p(M|\lambda, X) p(\lambda|X) p(X) \quad (3.84)$$

これを、 λ の事後分布は、

$$\begin{aligned} p(\lambda|X) &\propto \frac{p(X, \lambda, M)}{p(M|\lambda, X)p(X)} \\ &\propto \frac{p(X, \lambda, M)}{p(M|\lambda, X)} \end{aligned} \quad (3.85)$$

ここで、同時分布は

$$p(X, M, \lambda) = p(X|M, \lambda) p(M, \lambda)$$

で、左辺。

左辺、(3.85) の左辺は λ の

$$\ln p(\lambda|X)$$

$$= \underbrace{\ln p(X|M, \lambda)}_{\text{blue}} + \underbrace{\ln p(M, \lambda)}_{\text{green}} - \underbrace{\ln p(M|\lambda, X)}_{\text{orange}} + \text{const.}$$

$$\begin{aligned} &= \sum_{i=1}^N \underbrace{\ln \mathcal{N}(x_i | M, \lambda^{-1})}_{(3.80)} + \underbrace{\ln \mathcal{N}(M | m, (\beta \lambda)^{-1})}_{(3.81)} + \underbrace{\ln \text{Gam}(\lambda | a, b)}_{\text{green}} \\ &\quad - \underbrace{\ln \mathcal{N}(M | \hat{m}, (\hat{\beta} \lambda)^{-1})}_{(3.82)} + \text{const.} \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^N \left[\frac{1}{2} (\ln \lambda - \ln 2\pi) - \frac{1}{2} \lambda (x_i - \mu)^2 \right] \\
&\quad + \frac{1}{2} (\ln(\beta \lambda) - \ln 2\pi) - \frac{1}{2} \beta \lambda (\mu - m)^2 + (a-1) \ln \lambda - b \lambda + \ln C(a, b) \\
&\quad - \left[\frac{1}{2} (\ln(\hat{\beta} \lambda) - \ln 2\pi) - \frac{1}{2} \hat{\beta} \lambda (\mu - \hat{m})^2 \right] \stackrel{+ \text{const.}}{=} \\
&= \left(\frac{N}{2} + a - 1 \right) \ln \lambda \\
&\quad + \left[-\frac{1}{2} \sum_{i=1}^N (x_i - \mu)^2 - \frac{1}{2} \beta (\mu - m)^2 - b + \frac{1}{2} \hat{\beta} (\mu - \hat{m})^2 \right] \lambda \stackrel{+ \text{const.}}{=}
\end{aligned}$$

$$\begin{aligned}
&\stackrel{=: \text{7'.'}}{=} \\
&- \frac{1}{2} \sum_{i=1}^N (x_i - \mu)^2 \\
&= -\frac{1}{2} \sum_{i=1}^N (x_i^2 - 2\mu x_i + \mu^2) = -\frac{1}{2} \sum_{i=1}^N x_i^2 + \cancel{\mu \sum_{i=1}^N x_i} - \frac{1}{2} N \mu^2 \\
&\quad - \frac{1}{2} \beta (\mu - m)^2 - b \\
&= -\frac{1}{2} \beta (\mu^2 - 2\mu m + m^2) - b \\
&= -\frac{1}{2} \beta \mu^2 + \cancel{\beta \mu m} - \frac{1}{2} \beta m^2 - b
\end{aligned}$$

$$\begin{aligned}
&\frac{1}{2} \hat{\beta} (\mu - \hat{m})^2 \\
&= \frac{1}{2} \hat{\beta} (\mu^2 - 2\mu \hat{m} + \hat{m}^2) \\
&= \frac{1}{2} \hat{\beta} \mu^2 - \hat{\beta} \mu \cdot \frac{1}{\hat{\beta}} \left(\sum_{i=1}^N x_i + \beta m \right) + \frac{1}{2} \hat{\beta} \hat{m}^2 \\
&= \frac{1}{2} \hat{\beta} \mu^2 - \mu \sum_{i=1}^N x_i - \cancel{\beta \mu m} + \frac{1}{2} \hat{\beta} \hat{m}^2 \\
&(\because \hat{\beta} = N + \beta)
\end{aligned}$$

$$\lambda \geq 0, \lambda_0 > 0,$$

$$\left(-\frac{1}{2} \sum_{i=1}^N x_i^2 - \frac{1}{2} \beta_m^2 + \frac{1}{2} \hat{\beta}_m^2 - b \right) \lambda$$

$$= - \left[\frac{1}{2} \left(\sum_{i=1}^N x_i^2 + \beta_m^2 - \hat{\beta}_m^2 \right) + b \right] \lambda$$

2次上式

$$\ln p(\lambda | \mathbf{x}) = \left(\frac{N}{2} + a - 1 \right) \ln \lambda$$

$$- \left[\frac{1}{2} \left(\sum_{i=1}^N x_i^2 + \beta_m^2 - \hat{\beta}_m^2 \right) + b \right] \lambda$$

+ const.

(3.86)

これは、 λ の二点分布の対数と同一形

$$\left(\ln \text{Gam}(\lambda | \bar{a}, \bar{b}) = (\bar{a} - 1) \ln \lambda - \bar{b} \lambda + \text{const.} \right)$$

よし、事後分布は

$$\textcircled{1} p(\lambda | \mathbf{x}) = \text{Gam}(\lambda | \hat{a}, \hat{b}) \quad (3.87)$$

where

$$\hat{a} = \frac{N}{2} + a \quad (3.88)$$

$$\hat{b} = \frac{1}{2} \left(\sum_{i=1}^N x_i^2 + \beta_m^2 - \hat{\beta}_m^2 \right) + b$$

$$p(\mu | \mathbf{x}, \lambda), p(\lambda | \mathbf{x}) \quad (3.82), \quad (3.86) \quad \text{事後分布が同一形式。}$$

$\sim (M 1^\circ - 1^\circ \rightarrow -\gamma) m, \beta, a, b$ が都道府県で異なる
土地形

② 予測分布を計算.

$$p(x_*) = \iint p(x_* | \mu, \lambda) p(\mu, \lambda) d\mu d\lambda \quad (3.89)$$

～～ ベイズの定理を用いて、 $p(x_*)$ を求めよ。かくして？

$$p(\mu, \lambda | x_*) = \frac{p(x_* | \mu, \lambda) p(\mu, \lambda)}{p(x_*)}$$

よし、対数を取る。

$$\ln p(\mu, \lambda | x_*) = \ln p(x_* | \mu, \lambda) + \ln p(\mu, \lambda) - \ln p(x_*)$$

$$\therefore \ln p(x_*) = \ln p(x_* | \mu, \lambda) - \ln p(\mu, \lambda | x_*) + \text{const.} \quad (3.90)$$

ここで、 $p(\mu, \lambda | x_*)$ は、つまり x_* の観測後、事後分布を表す。

$$\begin{aligned} p(\mu, \lambda | x_*) &= p(\mu | x_*, \lambda) p(\lambda | x_*) \\ &\stackrel{(3.82) \text{ 式で}}{\times} \stackrel{(3.87) \text{ 式で}}{\times} \\ &\quad \text{X} \leftarrow x_* \sim (T_2 \pm \sigma) \quad \text{X} \leftarrow x_* \sim (T_2 \pm \sigma) \\ &= N(\mu | m(x_*), \{(1+\beta)\lambda\}^{-1}) \text{Gam}(\lambda | \frac{1}{2} + a, b(x_*)) \end{aligned} \quad (3.91)$$

$T_2 = T_1 - L$.

$$\hat{m} = \frac{x_* + \beta m}{1 + \beta} \doteq m(x_*)$$

$$\begin{aligned} \hat{b} &= \frac{1}{2} \left(x_*^2 + \beta m^2 - (1+\beta) \cdot \frac{(x_* + \beta m)^2}{(1+\beta)^2} \right) + b \\ &= \frac{1}{2} \left[x_*^2 + \beta m^2 - \frac{x_*^2 + 2\beta m x_* + \beta^2 m^2}{1+\beta} \right] + b \\ &= \frac{1}{2} \left[\frac{(1+\beta-1)x_*^2 - 2\beta m x_* + (\beta((1+\beta)-\beta)m^2)}{1+\beta} \right] + b \\ &= \frac{\beta}{2(1+\beta)} (x_* - m)^2 + b \doteq b(x_*) \end{aligned} \quad (3.92)$$

(3.90) (2) (3.80), (3.91) 代入.

$$\ln p(x_*)$$

$$= \ln p(x_* | M, \lambda) - \ln (M, \lambda | x_*) + \text{const.}$$

$$= \ln \mathcal{N}(x_* | M, \lambda)$$

$$- \ln \mathcal{N}(\mu | m(x_*), \{(1+\beta)\lambda\}^{-1})$$

$$- \ln \text{Gam}(\lambda | \frac{1}{2} + a, b(x_*)) + \text{const.}$$

$$= \frac{1}{2} (\ln \lambda - \ln 2\pi) - \frac{1}{2} \lambda (x_* - M)^2$$

$$- \left[\frac{1}{2} (\ln ((1+\beta)\lambda) - \ln 2\pi) - \frac{1}{2} (1+\beta)\lambda (M - m(x_*))^2 \right]$$

$$- \left((\frac{1}{2} + a - 1) \ln \lambda - b(x_*) \lambda + \ln \text{Ga}(\frac{1}{2} + a, b(x_*)) + \text{const.} \right)$$

$$= - \frac{1}{2} \lambda (x_*^2 - 2Mx_* + M^2) + \frac{1}{2} (1+\beta) \lambda \left(M - \frac{x_* + \beta M}{1+\beta} \right)^2$$

$$+ \frac{\beta}{2(1+\beta)} (x_* - M)^2 \lambda + \left(\frac{1}{2} + a \right) \ln b(x_*) - \ln \Gamma(\frac{1}{2} + a) + \text{const.}$$

$$\geq \frac{1}{2} (1+\beta) \lambda \left(M - \frac{x_* + \beta M}{1+\beta} \right)^2$$

$$= \frac{1}{2} (1+\beta) \lambda \left(M^2 - \frac{2(x_* + \beta M)}{1+\beta} M + \frac{(x_* + \beta M)^2}{(1+\beta)^2} \right)$$

$$= \frac{1}{2} (1+\beta) \lambda M^2 - 2\lambda M(x_* + \beta M) + \frac{\lambda(x_* + \beta M)^2}{2(1+\beta)}$$

$$= \frac{1}{2} (1+\beta) \lambda M^2 - 2\lambda Mx_* - 2\lambda \beta M M + \frac{\lambda}{2(1+\beta)} (x_*^2 + 2\beta M x_* + \beta^2 M^2)$$

$$= \frac{1}{2} (1+\beta) \lambda M^2 - 2\lambda Mx_* - 2\lambda \beta M + \frac{\lambda}{2(1+\beta)} x_*^2 + \frac{\beta M \lambda}{1+\beta} x_* + \frac{\beta^2 \lambda M^2}{2(1+\beta)}$$

$$\frac{\beta}{2(1+\beta)} (x_* - m)^2$$

$$= \frac{\beta \lambda}{2(1+\beta)} (\cancel{x_*^2} - 2x_* m + m^2)$$

$$\left(\frac{1}{2} + \alpha\right) \ln b(x_*)$$

$$= \left(\frac{1}{2} + \alpha\right) \ln \left[\frac{\beta}{2(1+\beta)} (x_* - m)^2 + b \right]$$

よし、 x_* は角 B の観測値.

$$\ln p(x_*)$$

$$= -\frac{1}{2} \cancel{\lambda x_*^2} + M \cancel{\lambda x_*}$$

$$\frac{\lambda}{2(1+\beta)} + \frac{\beta \lambda}{2(1+\beta)} = \frac{1}{2}$$

$$-\cancel{\lambda M x_*} + \frac{\lambda}{2(1+\beta)} x_*^2 + \frac{\beta m \lambda}{1+\beta} x_*$$

$$+ \frac{\beta \lambda}{2(1+\beta)} x_*^2 - \frac{\beta m \lambda}{1+\beta} x_*$$

$$+ \left(\frac{1}{2} + \alpha\right) \ln \left[1 + \frac{\beta}{2(1+\beta)b} (x_* - m)^2 \right] + \text{const.}$$

$$= \left(\frac{1}{2} + \alpha\right) \ln \left[\frac{\beta}{2(1+\beta)b} (x_* - m)^2 \right] + \text{const.}$$

(3.13)

これは、 λ -分布の対応式である.

$$p(x_*) = S_t (x_* | \mu_s, \lambda_s, v_s) \quad (3.14)$$

$$\text{where } \mu_s = m$$

$$\lambda_s = \frac{\beta \alpha}{(1+\beta)b} \quad (3.15)$$

$$v_s = 2\alpha$$

\rightsquigarrow これは確率密度 $p(x_* | X)$ は、事前分布 $\pi(X)$ と事後分布 $\pi(X | x_*)$ を入力する.