2.4 指数型分布族

7/11(金)

鈴木直也

指数型分布族

これまで学んできた分布*[1] → 指数型分布族

x*[2]上の指数分布族はηをパラメータとして次式で定義される。

$$p(x|\eta) = h(x)g(\eta)\exp\left\{\eta^T u(x)\right\} \qquad (2.194)$$

注1：混合ガウス分布を除く

注2：スカラーでもベクトルでも、又、離散、連続を問わない

指数分布族の積分

ここで関数$g(\eta)$は、分布を正規化するための係数と解釈できるので、次式を満たす。

(2.194)式の右辺を積分すると

$$\int h(x) g(\eta) \exp\left\{\eta^T u(x)\right\} dx$$

$$g(\eta) \int h(x) \exp\left\{\eta^T u(x)\right\} dx = 1 \qquad (2.195)$$

ベルヌーイ分布、多項分布、ガウス分布が指数型分布族であることを示す。

a. ベルヌーイ分布
b. 多項分布
c. ガウス分布

a.ベルヌーイ分布

- $p(x|\mu) = Bern(x|\mu) = \mu^x(1-\mu)^{1-x}$ (2.196)

右辺の対数をとり、さらに指数をとる。

$$
\begin{aligned}
p(x|\mu) &= \exp\left[\ln\left\{\mu^x(1-\mu)^{1-x}\right\}\right] \\
&= \exp\left\{x\ln\mu + (1-x)\ln(1-\mu)\right\} \\
&= \exp\left\{\boxed{x\ln\mu} + \ln(1-\mu) - \boxed{x\ln(1-\mu)}\right\} \\
&= \exp\left\{\ln\left(1-\mu\right)\right\}\boxed{\exp\left\{x\ln\mu - x\ln(1-\mu)\right\}} \\
&= \left(1-\mu\right)\exp\left\{x\ln\left(\frac{\mu}{1-\mu}\right)\right\} \qquad (2.197)
\end{aligned}
$$

ベルヌーイ分布（続き）

ここで式(2.194)と比べてみると

$$\eta = \ln\left(\frac{\mu}{1-\mu}\right) \qquad (2.198)$$

と対応できる。これを μ について解く*[3] と

$$\mu = \sigma(\eta) = \frac{1}{\exp(-\eta)+1} \qquad (1.199)$$

が得られる。

注3：後述

式1.198から式1.199の導出

$$\exp(\eta) = \frac{\mu}{1-\mu}$$

両辺の指数をとる。

$$(1-\mu)\exp(\eta) = \mu$$

$$\exp(\eta) - \mu\exp(\eta) = \mu$$

$$\exp(\eta) = \{1+\exp(\eta)\}\mu$$

$$\mu = \frac{\exp(\eta)}{1+\exp(\eta)} = \frac{1}{1/\exp(\eta) + 1} = \frac{1}{\exp(-\eta)+1}$$

ベルヌーイ分布（続きその2）

ここで式(2.197)に式(2.199)を代入する。

$$(1-\mu)\exp\left\{x\ln\left(\frac{\mu}{1-\mu}\right)\right\} \qquad (2.197)$$

$$=\left(1-\frac{1}{1+\exp(-\eta)}\right)\exp\left\{\ln\left(\frac{1\big/1+\exp(-\eta)}{1-1\big/1+\exp(-\eta)}\right)x\right\}$$

$$=\left(\frac{1+\exp(-\eta)-1}{1+\exp(-\eta)}\right)\exp\left\{\ln\left(\frac{1}{1+\exp(-\eta)-1}\right)x\right\}$$

$$=\left(\frac{1}{1\big/\exp(-\eta)+1}\right)\exp\{\ln(\exp(\eta))x\}$$

$$=\left(\frac{1}{\exp(\eta)+1}\right)\exp(\eta x)=\sigma(-\eta)\exp(\eta x) \qquad (2.200)$$

ベルヌーイ分布（続きその3）

式(2.200)と式(2.194)を比較すると

$$p(x|\eta) = h(x)g(\eta)\exp\left\{\eta^T u(x)\right\} \qquad (2.194)$$

$$p(x|\eta) = 1 \times \sigma(-\eta)\exp(\eta x) \qquad (2.200)$$

b.多項分布

$$p(x|\mu) = mult(x|\mu) = \prod_{k=1}^{M} \mu_k^{x_k}$$

$$= \exp\left\{ \ln\left(\prod_{k=1}^{M} \mu_k^{x_k} \right) \right\} = \exp\left\{ \sum_{k=1}^{M} x_k \ln \mu_k \right\} \quad (2.204)$$

ここで $\eta_k = \ln \mu_k, \eta = (\eta_1, \eta_2, \cdots, \eta_M)^T$ と定義するとこれも指数分布の標準形で表現できる。

多項分布(2)

$$p(x|\eta) = h(x)g(\eta)\exp\left\{\eta^T u(x)\right\} \qquad (2.194)$$

$$p(x \mid \mu) = \exp\left(\eta^T \mathrm{x}\right) \qquad (2.205)$$

上記2式を比べると以下の3式が得られる。

$$\mathrm{u}(\mathrm{x}) = \mathrm{x} \qquad (2.194)$$

$$h(\mathrm{x}) = 1 \qquad (2.207)$$

$$g(\eta) = 1 \qquad (2.208)$$

多項分布(3)

また、パラメータμ_kには以下の制約がある。

$$\sum_{k=1}^{M} \mu_k = 1 \qquad (2.209)$$

このことより、このパラメータμ_kのうち、M-1個のパラメータが分かれる。

→最後のパラメータもわかる。

→(2.204)式をM-1個のパラメータで表現する。

多項分布(4)

$$\exp\left\{\sum_{k=1}^{M} x_k \ln \mu_k\right\}$$

$$= \exp\left\{\sum_{k=1}^{M-1} x_k \ln \mu_k + \left(1-\sum_{k=1}^{M-1} x_k\right)\ln\left(1-\sum_{k=1}^{M-1}\mu_k\right)\right\}$$

$$= \exp\left\{\sum_{k=1}^{M-1} x_k \ln \mu_k + \ln\left(1-\sum_{k=1}^{M-1}\mu_k\right) - \sum_{k=1}^{M-1} x_k \ln\left(1-\sum_{k=1}^{M-1}\mu_k\right)\right\}$$

$$= \exp\left\{\sum_{k=1}^{M-1} x_k \ln\left(\frac{\mu_k}{\left(1-\sum_{j=1}^{M-1}\mu_j\right)}\right) + \ln\left(1-\sum_{k=1}^{M-1}\mu_k\right)\right\} \quad (2.211)$$

多項分布(5)

- 式(2.204)と比較すること(？)によって次式が分かる。

$$\ln\left(\frac{\mu_k}{1-\sum_j \mu_j}\right)=\eta_k \qquad (2.212)$$

- これらの式より μ_k について解くことができるらしく、

$$\mu_k = \frac{\exp(\eta_k)}{1+\sum_{j=1}\exp(\eta_j)} \qquad (2.213)$$

が得られるそうです。

多項分布(6)

よって多項分布では以下の形をとり、式(2.194)と比較すると。

$$p(x|\eta) = h(x)g(\eta)\exp\left\{\eta^T u(x)\right\} \qquad (2.194)$$

$$p(x|\eta) = 1 \times \left(1 + \sum_{k=1}^{M-1} \exp(\eta_k)\right)^{-1} \exp(\eta^T \mathbf{x}) \qquad (2.214)$$

ガウス分布

$$p(x|\mu,\sigma^2)=\frac{1}{(2\pi\sigma^2)^{1/2}}\exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\} \qquad (2.218)$$

$$=\frac{1}{(2\pi\sigma^2)^{1/2}}\exp\left\{-\frac{1}{2\sigma^2}x^2+\frac{\mu}{\sigma^2}x-\frac{1}{2\sigma^2}\mu^2\right\} \qquad (2.219)$$

ここで

$$\eta_1 = \mu/\sigma^2 \qquad (2.220a)$$

$$\eta_2 = -1/2\sigma^2 \qquad (2.200b)$$

$$\eta = (\eta_{1,}\eta_2)^T \qquad (2.220)$$

と定義する。

ガウス分布(2)

(2.200b)より

$$\sigma^2 = {-1}/{2\eta_2} \qquad (2.220c)$$

この式と(2.220a)より

$$\mu = {-\eta_1}/{2\eta_2} \qquad (2.220d)$$

ガウス分布(3)

(2.220c)および(2.220d)を(2.219)に代入すると

$$\frac{1}{\left(-2\pi/2\eta_2\right)^{1/2}}\exp\left\{-\frac{1}{-2/2\eta_2}x^2+\frac{-\eta_1/2\eta_2}{-1/2\eta_2}x-\frac{1}{-2/2\eta_2}\left(-\eta_1/2\eta_2\right)^2\right\}$$

$$=\frac{\left(-2\eta_2\right)^{1/2}}{\left(2\pi\right)^{1/2}}\exp\left\{\eta_2 x^2+\eta_1 x+\left(-\eta_1^2/4\eta_2\right)\right\}$$

$$=\left(2\pi\right)^{-1/2}\left(-2\eta_2\right)^{1/2}\exp\left(-\eta_1^2/4\eta_2\right)\exp\left(\eta_1 x+\eta_2 x^2\right) \qquad (2.220f)$$

ガウス分布(4)

(2.220f)より

$$\mathrm{u}(x) = \begin{pmatrix} x \\ x^2 \end{pmatrix}$$

$$h(\mathrm{x}) = (2\pi)^{-1/2}$$

$$g(\eta) = (-2\eta_2)^{1/2} \exp\left(\frac{\eta_1^2}{4\eta_2} \right)$$

の3式が導出される。