PRML 演習問題2.7, 2.8 解答

演習問題2.7

μの事前分布はベータ分布であるため、(2.15) からμの事前平均は

$$\frac{a}{a+b}$$

同様に m,l に対する事後平均は (2.20) となるため

$$\frac{m+a}{m+a+l+b}$$

また、最尤推定量は次で与えられる。

$$\frac{m}{m+l}$$

問題文から、λと 1-λを用いて、事前平均と最尤推定量から事後平均を表すと

$$\lambda\frac{a}{a+b}+(1-\lambda)\frac{m}{m+l}=\frac{m+a}{m+a+l+b}$$

この時、λが

$$0\leq\lambda\leq 1 \qquad \cdots\cdots\cdots (*)$$

となる事を確認できれば、事後平均が事前平均と最尤推定解との間の値になる事が示される事になる。上の式をλについて計算すると

$$\begin{aligned} &\lambda\left(\frac{a}{a+b}-\frac{m}{m+l}\right)=\frac{m+a}{m+a+l+b}-\frac{m}{m+l} \\ \Leftrightarrow\quad &\lambda\frac{al-bm}{(a+b)(m+l)}=\frac{al-bm}{(m+a+l+b)(m+l)} \end{aligned}$$

となる。ここで左辺のλの係数が 0 の時は

$$\frac{a}{a+b}-\frac{m}{m+l}=\frac{al-bm}{(a+b)(m+l)}$$

ここで、λの係数について考えると、λの係数が 0 になる場合は al=bm の時であり、この場合は右辺も 0 となるため、λの値に関わらず両辺は等しくなる。
一方、al=bm でない場合は、両辺をλの係数で割ることができ

$$\lambda=\frac{a+b}{m+a+l+b}$$

となる。a,b,m,l はいずれも非負である事から、(*) を満たす事がわかる。

演習問題 2.8

一つ目の式はまず、

$$E_y[E_x[x|y]] = \int_y \left(\int_x xp(x|y)dx \right) p(y)dy$$

ベイズの定理から $p(x|y) = \frac{p(x,y)}{p(y)}$ となり、これを代入すると

$$\begin{aligned} E_y[E_x[x|y]] &= \int_y \left(\int_x x\frac{p(x,y)}{p(y)}dx \right) p(y)dy \\ &= \int_y \int_x xp(x,y)dxdy \\ &= \int_x x \left(\int_y p(x,y)dy \right) dx \\ &= \int_x xp(x)dx \\ &= E[x] \end{aligned}$$

と導ける。ただし

$$\int_y p(x,y)dy = p(x)$$

を利用している。
二つ目の式は、まず右辺第二項から見ると、(1.39) と上の式から

$$\begin{aligned} var_y[E_x[x|y]] &= E_y\left[E_x[x|y]^2\right] - E_y\left[E_x[x|y]\right]^2 \\ &= E_y\left[E_x[x|y]^2\right] + E[x]^2 \end{aligned} \tag{1}$$

続けて、右辺第一項は、(1.34) と P46 のページ下部訳注から

$$\begin{aligned} E_y[var_x[x|y]] &= \int_y \left(\int_x \{x - E_x[x|y]\}^2 p(x|y)dx \right) p(y)dy \\ &= \int_y \left(\int_x \{x^2 - 2xE_x[x|y] + E_x[x|y]^2\} p(x|y)dx \right) p(y)dy \end{aligned}$$

ここで $E_x[x|y]$ は x の関数ではないので x の積分から外す事ができ

$$E_y[var_x[x|y]] = \int_y \left(\int_x x^2 p(x|y)dx - 2E_x[x|y]\int_x xp(x|y)dx + E_x[x|y]^2 \int_x p(x|y)dx \right) p(y)dy$$

すると $\int_x xp(x|y)dx = E_x[x|y]$ と $\int_x p(x)dx = 1, \int_x p(x|y)dx = 1$ から

$$\begin{aligned} E_y[var_x[x|y]] &= \int_y \left(\int_x x^2 p(x|y)dx - 2E_x[x|y]^2 + E_x[x|y]^2\right) p(y)dy \\ &= \int_y \left(\int_x x^2 p(x|y)dx - E_x[x|y]^2\right) p(y)dy \\ &= \int_y \left(\int_x x^2 p(x|y)dx\right) p(y)dy - \int_y E_x[x|y]^2 p(y)dy \\ &= \int_y \left(\int_x x^2 p(x|y)dx\right) p(y)dy - E_y[E_x[x|y]^2] \end{aligned}$$

また、ここでもやはりベイズの定理より $p(x|y) = \frac{p(x,y)}{p(y)}$ を利用して

$$\begin{aligned} E_y[var_x[x|y]] &= \int_y \left(\int_x x^2 p(x,y)dx\right) dy - E_y[E_x[x|y]^2] \\ &= \int_x \int_y x^2 p(x,y)dydx - E_y[E_x[x|y]^2] \\ &= \int_x x^2 p(x)dx - E_y[E_x[x|y]^2] \\ &= E[x^2] - E_y[E_x[x|y]^2] \end{aligned} \tag{2}$$

すると (1) と (2) から

$$\begin{aligned} E_y[var_x[x|y]] + var_y[E_x[x|y]] &= \{E[x^2] - E_y[E_x[x|y]^2]\} + \{E_y\left[E_x[x|y]^2\right] + E[x]^2\} \\ &= E[x^2] - E[x]^2 \\ &= var[x] \end{aligned}$$