PRML 演習問題2.3,2.4

@americiumian

平成 24 年 10 月 15 日

2.3

この演習問題では，二項分布 (2.9) が正規化されていることを証明する．まず，全部で N 個ある対象から m 個の同じものを選ぶ組み合わせの数の定義 (2.10) を用いて，

$$\begin{pmatrix} N \\ m \end{pmatrix} + \begin{pmatrix} N \\ m-1 \end{pmatrix} = \begin{pmatrix} N+1 \\ m \end{pmatrix} \tag{1}$$

を示せ．この結果を用い，帰納法で次の結果を証明せよ．

$$(1+x)^N = \sum_{m=0}^{N} \begin{pmatrix} N \\ m \end{pmatrix} x^m \tag{2}$$

これは**二項定理** (binomial theorem) として知られ，全ての実数値について成り立つ．最後に，二項分布が次のように正規化されていることを，$(1-\mu)^N$ を和の外に出してから，二項定理を使うことで示せ．

$$\sum_{m=0}^{N} \begin{pmatrix} N \\ m \end{pmatrix} \mu^m (1-\mu)^{N-m} = 1 \tag{3}$$

2.3.1

式 (1) を証明する．

$$\begin{aligned}
(\text{左辺}) &= \begin{pmatrix} N \\ m \end{pmatrix} + \begin{pmatrix} N \\ m-1 \end{pmatrix} \\
&= \frac{N!}{m!(N-m)!} + \frac{N!}{(m-1)!(N-m+1)!} \\
&= N! \left(\frac{N-m+1}{m!(N-m+1)!} + \frac{m}{m!(N-m+1)!} \right) \\
&= \frac{(N+1)!}{m!(N+1-m)!} \\
&= \begin{pmatrix} N+1 \\ m \end{pmatrix} \\
&= (\text{右辺})
\end{aligned}$$

より，成立．

2.3.2

式 (2) を，数学的帰納法を用いて証明する．

1. $N = 0$ の時

$$\begin{aligned}(\text{左辺}) &= (1+x)^0 = 1 \\ (\text{右辺}) &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} x^0 = 1\end{aligned}$$

より，成立

2. $N = k$ の時成立すると仮定し，$N = k+1$ の時に成立することを示す．
$N = k$ の時に成立すると仮定したため，

$$(1+x)^k = \sum_{m=0}^{k} \begin{pmatrix} k \\ m \end{pmatrix} x^m \tag{4}$$

式 (4) の両辺に $(1+x)$ をかけて，

$$\begin{aligned}(1+x)^{k+1} &= \sum_{m=0}^{k} \left\{ \begin{pmatrix} k \\ m \end{pmatrix} x^m + \begin{pmatrix} k \\ m \end{pmatrix} x^{m+1} \right\} \\ &= \sum_{m=0}^{k} \begin{pmatrix} k \\ m \end{pmatrix} x^m + \sum_{m=0}^{k} \begin{pmatrix} k \\ m \end{pmatrix} x^{m+1} \\ &= \left(1 + \sum_{m=1}^{k} \begin{pmatrix} k \\ m \end{pmatrix} x^m \right) + \left(\sum_{m=1}^{k} \begin{pmatrix} k \\ m-1 \end{pmatrix} x^m + x^{k+1} \right) \\ &= 1 + x^{k+1} + \sum_{m=1}^{k} \left\{ \begin{pmatrix} k \\ m \end{pmatrix} + \begin{pmatrix} k \\ m-1 \end{pmatrix} \right\} x^m \\ &= 1 + x^{k+1} + \sum_{m=1}^{k} \begin{pmatrix} k+1 \\ m \end{pmatrix} x^m \\ &= \sum_{m=0}^{k+1} \begin{pmatrix} k+1 \\ m \end{pmatrix} x^m\end{aligned}$$

より成立．

従って，任意の自然数 N について，

$$(1+x)^N = \sum_{m=0}^{N} \begin{pmatrix} N \\ m \end{pmatrix} x^m \tag{5}$$

が成立する．

2.3.3

式 (3) を証明する．

$$
\begin{aligned}
(\text{左辺}) &= \sum_{m=0}^{N} \begin{pmatrix} N \\ m \end{pmatrix} \mu^m (1-\mu)^{N-m} \\
&= (1-\mu)^N \sum_{m=0}^{N} \begin{pmatrix} N \\ m \end{pmatrix} \left(\frac{\mu}{1-\mu}\right)^m \\
&= (1-\mu)^N \left(1 + \frac{\mu}{1-\mu}\right)^N \\
&= (1-\mu)^N \left(\frac{1}{1-\mu}\right)^N \\
&= 1
\end{aligned}
$$

より，成立.

2.4

二項分布の平均が (2.11) であることを示せ．これには，正規化条件 (2.64) の両辺を μ で微分し，変形して n の平均を求めよ．同様に，(2.264) の両辺を μ について 2 階微分し，二項分布の平均 (2.11) も用いて，二項分布の分散の結果 (2.12) を証明せよ.

2.4.1

二項分布の平均が式 (2.11) で表されることを示す.
式 (3) の両辺を μ で偏微分して，

$$
\begin{aligned}
& \sum_{m=0}^{N} \begin{pmatrix} N \\ m \end{pmatrix} \left\{ m\mu^{m-1}(1-\mu)^{N-m} - (N-m)\mu^N (1-\mu)^{N-m-1} \right\} = 0 \qquad (6) \\
\Leftrightarrow \quad & \frac{1}{\mu} \sum_{m=0}^{N} \begin{pmatrix} N \\ m \end{pmatrix} m\mu^{m-1}(1-\mu)^{N-m} = \frac{1}{1-\mu} \sum_{m=0}^{N} \begin{pmatrix} N \\ m \end{pmatrix} (N-m)\mu^N (1-\mu)^{N-m}.
\end{aligned}
$$

ここで，

$$
\begin{aligned}
\sum_{m=0}^{N} \begin{pmatrix} N \\ m \end{pmatrix} m\mu^{m-1}(1-\mu)^{N-m} &= E[m] \\
\sum_{m=0}^{N} \begin{pmatrix} N \\ m \end{pmatrix} \mu^{m-1}(1-\mu)^{N-m} &= 1
\end{aligned}
$$

より，

$$
\begin{aligned}
\frac{1}{\mu} E[m] &= \frac{1}{1-\mu}(N - E[m]) \\
\frac{1}{\mu(1-\mu)} E[m] &= \frac{1}{1-\mu} N \\
E[m] &= N\mu
\end{aligned}
$$

2.4.2

二項分布の分散が式 (2.12) で表されることを示す.
式 (6) の両辺を μ で偏微分して,

$$
\begin{aligned}
&\sum_{m=0}^{N}\begin{pmatrix} N \\ m \end{pmatrix}\{m(m-1)\mu^{m-2}(1-\mu)^{N-m} - 2m(N-m)\mu^{m-1}(1-\mu)^{N-m-1} \\
&+(N-m)(N-m-1)\mu^{m}(1-\mu)^{N-m-2}\} = 0 \\
\Leftrightarrow\ &\frac{1}{\mu^2}(E[m^2]-E[m]) - \frac{2}{\mu(1-\mu)}(NE[m]-E[m^2]) + \frac{1}{(1-\mu)^2}\{N(N-1)+E[m^2]-N(2N-1)\} = 0
\end{aligned}
$$

式 (2.11) より, $E[m] = N\mu$ なので,

$$
\begin{aligned}
\Leftrightarrow\ &(1-\mu)^2E[m^2] - N\mu(1-\mu)^2 - 2N^2\mu^2(1-\mu) + 2\mu(1-\mu)E[m^2] \\
&+N(N-1)\mu^2 + \mu^2E[m^2] - N(2N-1)\mu^3 = 0 \\
\Leftrightarrow\ &(1-2\mu+\mu^2+2\mu-2\mu^2+\mu^2)E[m^2] = N\mu(1-\mu)^2 + 2N^2\mu^2(1-\mu) - N(N-1)\mu^2 + N(2N-1)\mu^3 \\
\Leftrightarrow\ &E[m^2] = \mu N(-\mu+\mu N+1) \\
\Leftrightarrow\ &E[m^2] = N\mu(1-\mu) + N^2\mu^2
\end{aligned}
$$

したがって,

$$
\begin{aligned}
\sum_{m=0}^{N}(x-E[x])^2Bin(x|m,N) &= \sum_{m=0}^{N}x^2Bin(x|m,N) - 2E[x]\sum_{m=0}^{N}xBin(x|m,N) + E[x]^2\sum_{m=0}^{N}Bin(x|m,N) \\
&= E[x^2] - 2E[x]^2 + E[x]^2 \\
&= E[x^2] - E[x]^2 \\
&= \{N\mu(\mu-1) + N^2\mu^2\} - N^2\mu^2 \\
&= N\mu(1-\mu)
\end{aligned}
$$