1 A binary child given a binary parent

Let X and Y be two binary variables. The conditional probability of the child-node X given the parent-node Y is expressed as follows:

$$p(X|Y) = p_{x|y}^{I(Y=0)I(X=0)} p_{\bar{x}|y}^{I(Y=0)I(X=1)} p_{x|\bar{y}}^{I(Y=1)I(X=0)} p_{\bar{x}|\bar{y}}^{I(Y=1)I(X=1)}$$

This conditional probability can be expressed in different exponential forms as follows:

- **First form**:

$$\begin{aligned}
\ln p(X|Y) &= \theta^T s(X,Y) - A(\theta) \\
&= \begin{pmatrix} \ln p_{x|y} \\ \ln p_{\bar{x}|y} \\ \ln p_{x|\bar{y}} \\ \ln p_{\bar{x}|\bar{y}} \end{pmatrix}^T \begin{pmatrix} I(X=0)I(Y=0) \\ I(X=1)I(Y=0) \\ I(X=0)I(Y=1) \\ I(X=1)I(Y=1) \end{pmatrix} - 0
\end{aligned}$$

- **Second form**:

$$\begin{aligned}
\ln p(X|Y) &= \theta(Y)^T s(X) - A(Y) \\
&= \begin{pmatrix} I(Y=0)\ln p_{x|y} + I(Y=1)\ln p_{x|\bar{y}} \\ I(Y=0)\ln p_{\bar{x}|y} + I(Y=1)\ln p_{\bar{x}|\bar{y}} \end{pmatrix}^T \begin{pmatrix} I(X=0) \\ I(X=1) \end{pmatrix} - 0 \\
&= \begin{pmatrix} m_0^Y \cdot \theta_0 + m_1^Y \cdot \theta_2 \\ m_0^Y \cdot \theta_1 + m_1^Y \cdot \theta_3 \end{pmatrix}^T \begin{pmatrix} I(X=0) \\ I(X=1) \end{pmatrix} - 0
\end{aligned}$$

- **Third form**:

$$\begin{aligned}
\ln p(X|Y) &= \theta(X)^T s(Y) - A(X) \\
&= \begin{pmatrix} I(X=0)\ln p_{x|y} + I(X=1)\ln p_{\bar{x}|y} \\ I(X=0)\ln p_{x|\bar{y}} + I(X=1)\ln p_{\bar{x}|\bar{y}} \end{pmatrix}^T \begin{pmatrix} I(Y=0) \\ I(Y=1) \end{pmatrix} - 0 \\
&= \begin{pmatrix} m_0^X \theta_0 + m_1^X \theta_1 \\ m_0^X \theta_2 + m_1^X \theta_3 \end{pmatrix}^T \begin{pmatrix} I(Y=0) \\ I(Y=1) \end{pmatrix} - 0
\end{aligned}$$

2 A multinomial child given a set of multinomial parents

Let X be a multinomial variable whose state space is $\{x^1, ..., x^k\}$, and let $\mathbf{Y} = Y_1, \ldots, Y_n$ denote the set of parents of X, such that all of them are multinomial. Each parent Y_i has r_i states $\{{y_1}^{r_1}, {y_2}^{r_2}, \ldots, {y_n}^{r_n}\}$ with $r_i \geq 2$. The log-conditional probability of the child-node X given the parent-nodes $\mathbf{Y}$ can be expressed as follows:

$$\ln p(X|Y_1, \ldots, Y_n) =$$

Similarly the above conditional probability can be expressed in the following exponential forms:

- **First form**:

$$\ln p(X|Y_1, \ldots, Y_n) \quad = \quad \theta^T s(X, Y_1, \ldots, Y_n) - A(\theta)$$

- **Second form**:

$$\ln p(X|Y_1, \ldots, Y_n) \quad = \quad \theta(Y_1, \ldots, Y_n)^T s(X) - A(Y_1, \ldots, Y_n)$$

- **Third form**:

$$\ln p(X|Y_1, \ldots, Y_n) \quad = \quad \theta(X, Y_1, \ldots, Y_{l-1}, Y_l, \ldots, Y_n)^T s(Y_l) - A(X, Y_1, \ldots, Y_{l-1}, Y_l, \ldots, Y_n)$$

3 A normal child given a set of normal parents

Let X be a normal variable and $\mathbf{Y} = \{Y_1, \ldots, Y_n\}$ denote the set of parents of X, such that all of them are normal. The log-conditional probability of X given $\mathbf{Y}$ can be expressed as follows:

$$\ln p(X|Y_1, \ldots, Y_n) = \ln \left(\frac{1}{\sigma\sqrt{2(\beta_0 + \sum_i^n \beta_i Y_i)}} \mathrm{e}^{-\frac{(y-(\beta_0+\sum_i^n \beta_i Y_i))^2}{2\sigma^2}} \right)$$

Similarly the above conditional probability can be expressed in the following exponential forms:

- **First form**:

$$\begin{aligned} \ln p(X|Y) &= \theta^T s(X,Y) - A(\theta) + h(Y) \\ &= \begin{pmatrix} \frac{-1}{2\sigma^2} & = & \theta_{\text{-}1} \\ \frac{-\beta_1^2}{2\sigma^2} & = & \theta_{1^2} \\ \cdots & & \\ \frac{-\beta_n^2}{2\sigma^2} & = & \theta_{n^2} \\ \frac{\beta_0}{\sigma^2} & = & \theta_0 \\ \frac{\beta_1}{\sigma^2} & = & \theta_1 \\ \cdots & & \\ \frac{\beta_n}{\sigma^2} & = & \theta_n \\ \frac{-\beta_0\beta_1}{\sigma^2} & = & \theta_{01} \\ \cdots & & \\ \frac{-\beta_0\beta_n}{\sigma^2} & = & \theta_{0n} \\ \frac{-\beta_1\beta_2}{\sigma^2} & = & \theta_{12} \\ \cdots & & \\ \frac{-\beta_1\beta_n}{\sigma^2} & = & \theta_{1n} \\ \cdots & & \\ \frac{-\beta_{n-1}\beta_n}{\sigma^2} & = & \theta_{n\text{-}1n} \end{pmatrix}^T \begin{pmatrix} X^2 \\ Y^2 \\ \cdots \\ Y_n^2 \\ X \\ XY_1 \\ \cdots \\ XY_n \\ Y_1 \\ \cdots \\ Y_n \\ Y_1Y_2 \\ \cdots \\ Y_1Y_n \\ \cdots \\ Y_{n-1}Y_n \end{pmatrix} - \left(\frac{\beta_0^2}{2\sigma^2} + \ln\sigma \right) + \frac{1}{\ln\sqrt{2\mu_{X|Y}}} \end{aligned}$$

- **Second form**:

$$\begin{aligned} \ln p(X|Y) &= \theta(Y)^T s(X) - A(\theta(Y)) + h(Y) \\ &= \begin{pmatrix} \frac{\mu_{X|Y}}{\sigma^2} \\ \frac{-1}{2\sigma^2} \end{pmatrix}^T \begin{pmatrix} X \\ X^2 \end{pmatrix} - \left(\frac{\mu_{X|Y}^2}{2\sigma^2} + \ln\sigma \right) + \ln \frac{1}{\sqrt{2\mu_{X|Y}}} \\ &= \begin{pmatrix} \theta_0 + \theta_i m_0^{Y_i} \\ \theta_{\text{-}1} \end{pmatrix}^T \begin{pmatrix} X \\ X^2 \end{pmatrix} - \left(\frac{\mu_{X|Y}^2}{2\sigma^2} + \ln\sigma \right) + \ln \frac{1}{\sqrt{2\mu_{X|Y}}} \end{aligned}$$

where $\mu_{X|Y} = \beta_0 + \sum_i^n \beta_i Y_i$

- **Third form**:

$$
\begin{aligned}
\ln p(X|Y) &= \theta(X)^T s(Y) - A(\theta(X)) + h(Y) \\
&= \begin{pmatrix} -\frac{\beta_1^2}{2\sigma^2} \\ \cdots \\ -\frac{\beta_n^2}{2\sigma^2} \\ \frac{\beta_1(X-\beta_0)}{\sigma^2} \\ \cdots \\ \frac{\beta_n(X-\beta_0)}{\sigma^2} \\ -\frac{\beta_1\beta_2}{\sigma^2} \\ \cdots \\ -\frac{\beta_1\beta_n}{\sigma^2} \\ \cdots \\ -\frac{\beta_{n-1}\beta_n}{\sigma^2} \end{pmatrix}^T \begin{pmatrix} Y_1^2 \\ \cdots \\ Y_n^2 \\ Y_1 \\ \cdots \\ Y_n \\ Y_1 Y_2 \\ \cdots \\ Y_1 Y_n \\ \cdots \\ Y_{n-1} Y_n \end{pmatrix} - \left(\frac{(X-\beta_0)^2}{\sigma^2} + \ln \sigma \right) + \frac{1}{\ln \sqrt{2\mu_{X|Y}}} \\
&= \begin{pmatrix} \theta_{1^2} \\ \cdots \\ \theta_{n^2} \\ \theta_1 m_0^X + \theta_{01} \\ \cdots \\ \theta_n m_0^X + \theta_{0n} \\ \theta_{12} \\ \cdots \\ \theta_{1n} \\ \cdots \\ \theta_{n\text{-}1n} \end{pmatrix}^T \begin{pmatrix} Y_1^2 \\ \cdots \\ Y_n^2 \\ Y_1 \\ \cdots \\ Y_n \\ Y_1 Y_2 \\ \cdots \\ Y_1 Y_n \\ \cdots \\ Y_{n-1} Y_n \end{pmatrix} - \left(\frac{(X-\beta_0)^2}{\sigma^2} + \ln \sigma \right) + \frac{1}{\ln \sqrt{2\mu_{X|Y}}}
\end{aligned}
$$

Notations

The list below presents a summary of the used notations:

X	Child variable
k	Range of possible values of a multinomial variable X
Y	Parent variable
$\mathbf{Y} = \{Y_1, \ldots, Y_n\}$	Set of parent variables
n	Number of parent variables
r_i	Range of possible values of a multinomial variable Y_i
i	Parent variable index
p	Probability distribution