Exponential Family Examples

P.12. Bernoulli(ϕ)

$$
\begin{aligned}
p(y;\phi) &= \phi^{y}(1-\phi)^{1-y} \\
&= e^{\log \phi^{y}(1-\phi)^{1-y}} \\
&= e^{y\log\phi + (1-y)\log(1-\phi)} \\
&= e^{y\log\phi + \log(1-\phi) - y\log(1-\phi)} \\
&= e^{y\log\frac{\phi}{1-\phi} + \log(1-\phi)} \\
&= \underbrace{1}_{b(y)} \cdot e^{\underbrace{y}_{T(y)}\left(\underbrace{\log\frac{\phi}{1-\phi}}_{\eta}\right) - \underbrace{(-\log(1-\phi))}_{\text{log-partition function } a(\eta)}}
\end{aligned}
$$

Write $-\log(1-\phi)$ as a function of η:

First, solve for ϕ in $\eta = \log\frac{\phi}{1-\phi}$:

$$
\begin{aligned}
\eta &= \log\frac{\phi}{1-\phi} \\
e^{\eta} &= \frac{\phi}{1-\phi} \\
e^{\eta} - e^{\eta}\phi &= \phi \\
e^{\eta} &= \phi + e^{\eta}\phi = (1+e^{\eta})\phi \\
\phi &= \frac{e^{\eta}}{1+e^{\eta}} \boxed{= \frac{1}{e^{-\eta}+1}} \text{ sigmoid function}
\end{aligned}
$$

Substitute ϕ in $\cancel{\log\frac{\phi}{1-\phi}}$ $-\log(1-\phi)$

$$
\begin{aligned}
a(\eta) &= -\log\left(1 - \frac{e^{\eta}}{1+e^{\eta}}\right) \\
&= -\log\left(\frac{1}{1+e^{\eta}}\right) \\
&= -\left(-\log(1+e^{\eta})\right) \\
&= \boxed{\log(1+e^{\eta})} \leftarrow a(\eta)
\end{aligned}
$$

P.13. Normal distribution $N(\mu, 1)$

$$
\begin{aligned}
p(y;\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y-\mu)^2} &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y^2+\mu^2-2y\mu)} \\
&= \left[\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}\right] e^{-\frac{1}{2}(\mu^2-2y\mu)} \\
&= \left[\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}\right] e^{\underbrace{\mu}_{\eta}\underbrace{y}_{T(y)} - \frac{\mu^2}{2}} \quad a(\eta) = \frac{\mu^2}{2} = \frac{\eta^2}{2}
\end{aligned}
$$

inverse: $\mu = \eta$

P16. Poisson(λ)

$$p(y;\lambda) = \frac{\lambda^y e^{-\lambda}}{y!} = \frac{1}{y!}\lambda^y e^{-\lambda} = \frac{1}{y!} e^{\log \lambda^y} e^{-\lambda}$$

$$= \frac{1}{y!} e^{y\log\lambda - \lambda}$$

$$= \underbrace{\frac{1}{y!}}_{b(y)} e^{\underbrace{\log\lambda}_{\eta}\cdot \underbrace{y}_{T(\eta)} - \underbrace{\lambda}_{\text{partition function } a(\eta)}}$$

since $\eta = \log\lambda$,

we have $\lambda = e^\eta$

Then, $a(\eta) = \lambda = \boxed{e^\eta} \leftarrow a(\eta)$

P26. Multinomial($\phi_1, \dots, \phi_k$)

Since $\sum_{i=1}^{k}\phi_i = 1$, it suffices to ~~have~~ use $k-1$ parameters to define a multinomial distribution of k random variables.

Let $T(y) \in \mathbb{R}^{k-1}$ be defined as $T(y) = \begin{bmatrix} 1\{y=1\} \\ \vdots \\ 1\{y=k-1\} \end{bmatrix}$

where $1\{y=i\}$ is the indicator

function $1\{y=i\} = \begin{cases} 1 & y=i \\ 0. & y\neq i. \end{cases}$

Denote the i th element of $T(y)$ as $T(y)_i = 1\{y=i\}$

$$p(y;\phi) = \prod_{i=1}^{k} \phi_i^{1\{y=i\}}$$

$$= \left(\prod_{i=1}^{k-1} \phi_i^{T(y)_i}\right) \phi_k^{1\{y=k\}} = \left(\prod_{i=1}^{k-1} \phi_i^{T(y)_i}\right) \phi_k^{1-\sum_{i=1}^{k-1} T(y)_i}$$

$$= e^{T(y)_i \log \prod_{i=1}^{k-1}\phi_i + (1-\sum_{i=1}^{k-1} T(y)_i)\log\phi_k}$$

$$= e^{T(y)_i \sum_{i=1}^{k-1}\log\phi_i + \log\phi_k - \sum_{i=1}^{k-1} T(y)_i \log\phi_k}.$$

$$= e^{\sum_{i=1}^{k-1}(T(y)_i \log\phi_i - T(y)_i\log\phi_k) + \log\phi_k}$$

$$= e^{\sum_{i=1}^{k-1}(T(y_i)\log(\frac{\phi_i}{\phi_k})) + \log\phi_k}.$$

$$= \underbrace{1}_{b(y)=1}\cdot e^{\underbrace{\begin{bmatrix}\log\frac{\phi_1}{\phi_k} \\ \vdots \\ \log\frac{\phi_{k-1}}{\phi_k}\end{bmatrix}^T}_{\eta} \underbrace{T(y)}_{T(y)=\begin{bmatrix}1\{y=1\} \\ \vdots \\ 1\{y=k-1\}\end{bmatrix}} + (\log\phi_k)}$$

$(\log\phi_k) \to$ How to write it as a function of η?

$\to a(\eta) = -\log\phi_k.$

Since $\eta = \begin{bmatrix} \log\frac{\phi_1}{\phi_k} \\ \vdots \\ \log\frac{\phi_{k-1}}{\phi_k} \end{bmatrix}$, we generalize the notation η_i to include $i=k$.

i.e. $\eta_i = \log\frac{\phi_i}{\phi_k}$ $(1 \le i \le k)$

and $\eta_k = \log\frac{\phi_k}{\phi_k} = 0$

Then we have $\eta_i = \log\frac{\phi_i}{\phi_k}$ ← canonical link function

$$e^{\eta_i} = \frac{\phi_i}{\phi_k}$$

$$\phi_i = \phi_k \cdot e^{\eta_i} \qquad (1)$$

since $\sum_{i=1}^{k}\phi_i = 1$, $1 = \sum_{i=1}^{k}\phi_i = \sum_{i=1}^{k}\phi_k \cdot e^{\eta_i}$

Then we can solve for ϕ_k:

$$1 = \phi_k \sum_{i=1}^{k} e^{\eta_i}$$

$$\phi_k = \frac{1}{\sum_{i=1}^{k} e^{\eta_i}} \qquad (2)$$

Then $a(\eta) = -\log\phi_k = -\log\frac{1}{\sum_{i=1}^{k}e^{\eta_i}} = \underline{\log\sum_{i=1}^{k}e^{\eta_i}}$ ← $a(\eta)$

Further, plug (2) into (1),

$$\phi_i = \frac{e^{\eta_i}}{\sum_{i=1}^{k}e^{\eta_i}}$$ ← canonical response function