# The Exponential Family

$$p(y;\eta) = b(y)\exp(\eta^T T(y) - a(\eta))$$

$\eta$ : natural parameter (canonical parameter)

$T(y)$ : sufficient statistic (usually $T(y)=y$)

$a(\eta)$ : log partition function

$\exp(-a(\eta))$ : usually normalization constant

$\sum p(y;\eta) = 1$ or $\int p(y;\eta) = 1$

$T, a, b ; \eta$ → defines the family of distribution

<Bernoulli distribution>

$\phi$ : mean

$p(y=1;\phi) = \phi$

$p(y=0;\phi) = 1-\phi$

$$\begin{aligned}
p(y;\phi) &= \phi^y(1-\phi)^{1-y} \\
&= \exp\{y\log\phi + (1-y)\log(1-\phi)\} \\
&= \exp\{y\log\phi + \log(1-\phi) - y\log(1-\phi)\} \\
&= \exp\{y(\log\phi - \log(1-\phi)) + \log(1-\phi)\} \\
&= \exp\{y \times \log\frac{\phi}{1-\phi} + \log(1-\phi)\} \\
&= \exp\{\log\left(\frac{\phi}{1-\phi}\right)y + \log(1-\phi)\}
\end{aligned}$$

$T(y) = y$ ✓

$\eta = \log\left(\frac{\phi}{1-\phi}\right)$ ✓

$e^\eta = \frac{\phi}{1-\phi}$

$(1-\phi)e^\eta = \phi$

$e^\eta = \phi(1+e^\eta)$

$\phi = \frac{e^\eta}{1+e^\eta}$

$\phi = \frac{1}{e^{-\eta}+1}$ ← sigmoid

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

$b(y) = 1$ ✓