(1) Todian Mishtaku Sept 2022

III

$$\Theta = \log\left(\frac{\mu}{1-\mu}\right) ; \mu = ?$$

$$\frac{1-\mu+\cancel{2\mu}}{1-\mu} = -1+2x$$

$$e^{\Theta} = \frac{\mu}{1-\mu} \Rightarrow \mu = (1-\mu) e^{\Theta}$$

$$\mu = e^{\Theta} - \mu e^{\Theta} \Rightarrow \mu(1+e^{\Theta}) = e^{\Theta} \Rightarrow$$

$$\boxed{\mu = \frac{e^{\Theta}}{1+e^{\Theta}}}$$

$$p(X|\mu) = \mu^{x}(1-\mu)^{1-x} = e^{\ln[\mu^{x}(1-\mu)^{1-x}]} =$$

$$= e^{\ln \mu^{x} + \ln(1-\mu)^{1-x}} = e^{x\ln\mu + (1-x)\ln(1-\mu)}$$

$$\phi(x) = [\mathbb{I}(x=1), \mathbb{I}(x=0)]$$

$$\Theta = [\ln(\mu), \ln(1-\mu)]$$

Exponential family:

$$\boxed{p(X|\Theta) = \frac{1}{Z(\Theta)} h(x) e^{\Theta^T\phi(x)}} = \boxed{h(x) e^{[\Theta^T\phi(x) - A(\Theta)]}}$$

$$\boxed{Z(\Theta) = \int_{x^m} h(x) e^{[\Theta^T\phi(x)]} dx} \quad A(\Theta) = \log(Z(\Theta))$$

(2)

Exponential family if

$$P(X|\theta) = \frac{1}{Z(\theta)} \cdot h(x) \cdot e^{\theta^T \phi(x)} = \boxed{\frac{1}{Z(\theta)} h(x) \cdot \exp(\theta^T \phi(x))} \quad (1)$$

$$\boxed{P(X|\theta) = h(x)\, e^{[\theta^T \phi(x) - A(\theta)]} = \boxed{h(x) \exp(\theta^T \phi(x) - A(\theta))}} \quad (2)$$

for $x = (x_1, \ldots, x_m) \in \mathcal{X}^m$, $\theta \in \Theta \subseteq R^d$

where

$$Z(\theta) = \int_{\mathcal{X}^m} h(x) \exp(\theta^T \cdot \phi(x))\, dx \quad (3)$$

$$A(\theta) = \log(Z) \quad (4) \qquad \log(u) = \log_e(u) = \ln u$$

$$= \boxed{\log \int_{\mathcal{X}^m} h(x) \exp(\theta^T \cdot \phi(x))\, dx} \quad (5) \qquad \log \overset{\text{adnot}}{=} \ln$$

$$\frac{dA}{d\theta} = \frac{d\left(\log \int h(x) \cdot \exp(\theta^T \phi(x))\right) dx}{d\theta} = \frac{d \log u(\theta)}{d(\theta)} =$$

where $u(\theta) = \boxed{\int h(x) \cdot \exp[\theta^T \phi(x)]\, dx}$

$$= \frac{u'(\theta)}{u(\theta)} = \frac{\frac{du(\theta)}{d\theta}}{u(\theta)} = \frac{d \int h(x) \exp[\theta^T \phi(x)]\, dx}{d\theta} \cdot \frac{1}{\int h(x) \exp[\theta^T \phi(x)]\, dx} =$$

$$= \frac{\frac{d}{d\theta} \int h(x) \exp[\theta^T \cdot \phi(x)]\, dx}{\int h(x) \exp[\theta^T \cdot \phi(x)]\, dx}$$

$$A(\theta) = \log \int h(x) \exp(\theta^T \phi(x))\, dx$$

$$e^{A(\theta)} = \int h(x) e^{\theta^T \phi(x)}\, dx$$

$$\boxed{\exp(A(\theta)) = \int h(x) \exp[\theta^T \phi(x)]\, dx}$$

$$= \frac{\frac{d}{d\theta} \int h(x) \exp[\theta^T \phi(x)]\, dx}{\exp(A(\theta))} =$$

$$\boxed{\frac{d}{d\theta} \int g(x) \cdot e^{\theta^T \phi(x)} = \int g(x) \cdot \phi(x)\, e^{\theta^T \phi(x)}} \quad g(x) = h(x)\, dx$$

$$= \frac{\int h(x) \phi(x) \exp[\theta^T \phi(x)]\, dx}{\exp(A(\theta))} = \frac{1}{e^{A(\theta)}} \int h(x) \cdot \phi(x) \cdot e^{\theta^T \phi(x)}\, dx =$$

$$= \int h(x) \phi(x)\, e^{[\theta^T \phi(x) - A(\theta)]} \cdot dx$$

③

$$= \int \phi(x) \cdot \underbrace{h(x) \exp[\theta^T \phi(x) - A(\theta)]}_{\substack{p(x|\theta) \\ p(x)}} dx =$$

$$= \int \phi(x)\, p(x) \cdot dx = E[\phi(x)]$$

i.e. $\frac{d A(\theta)}{d\theta} = E[\phi(x)]$ $\boxed{E[\phi(x)] = \frac{d A(\theta)}{d\theta}}$

Briefly:

$$\frac{dA}{d\theta} = \frac{d(\log \int h(x) \exp[\theta^T \phi(x)] dx)}{d\theta} =$$

$\frac{d \log u(\theta)}{d\theta} = \frac{u'(\theta)}{u(\theta)}$

$$\frac{\frac{d}{d\theta} \int h(x) \exp[\theta^T \phi(x)] dx}{\int h(x) \exp[\theta^T \phi(x)] dx} = \frac{\int \phi(x) h(x) \exp[\ ] dx}{\exp(A(\theta))}$$

$$= \int \frac{1}{\exp A(\theta)} \cdot \phi(x) h(x) \exp[\theta^T \phi(x)] = \int \phi(x)\, h(x) \frac{e^{\theta^T \phi(x)}}{e^{A(\theta)}} dx =$$

$$= \boxed{\int \phi(x) \underbrace{h(x) \exp[\theta^T \phi(x) - A(\theta)]}_{p(x)} \cdot dx} = \int \phi(x) p(x) dx$$

(6) (7)

$$= E[\phi(x)]$$

(4) $$\frac{d^2A}{d\theta^2} = \frac{d}{d\theta}\left(\frac{dA}{d\theta}\right) = \frac{d}{d\theta}\int \phi(x) h(x) \exp[\theta^T\phi(x) - A(\theta)] \cdot dx$$

$$V = \theta^T \cdot \phi(x) - A(\theta) \quad \Big| \quad = \frac{d}{d\theta}\int \phi(x) h(x) e^{V(\theta)} dx =$$

$$= \cancel{V'(\theta) \cdot \int \phi(x) h(x) e^{V(\theta)} dx} =$$

$$= \int \phi(x) \cdot h(x) \cdot V'(\theta) e^{V(\theta)} dx = \left\{ \theta^T \cdot \underbrace{\phi(x)}_{\text{const}} - A(\theta) \Big|_{\theta}' = \phi(x) - A'(\theta) \right\}$$

$$= \int \phi(x) h(x) \cdot [\phi(x) - A'(\theta)] e^{[\theta^T\phi(x) - A(\theta)]} dx =$$

$$= \int \phi^2(x) \underbrace{h(x) \exp[\theta^T\phi(x) - A(\theta)]}_{p(x)} dx - A'(\theta) \int \phi(x) \underbrace{h(x) e^{\theta^T\phi(x) - A(\theta)}}_{p(x)} dx$$

$$\frac{dA}{d\theta} = E[\phi(x)]$$

$$= \int \phi^2(x) p(x)\, dx - A'(\theta) \underbrace{\int \phi(x) p(x)\, dx}_{E[\phi(x)]} =$$

$$= E[\phi^2(x)] - E^2[\phi(x)] = \operatorname{var}[\phi(x)]$$

$$E[t^2] - E^2[t] = \operatorname{var}(t)$$

$$\boxed{\frac{dA}{d\theta} = E[\phi(x)]}$$

$$\boxed{\frac{d^2A}{d\theta^2} = \operatorname{var}[\phi(x)]}$$

Used for calculations!

(5)

$$p(x|\theta) = \boxed{h(x)\exp[\theta^T\cdot\phi(x) - A(\theta)]} \quad (2)$$

$$= \boxed{\frac{1}{Z(\theta)} h(x)\exp[\theta^T\cdot\phi(x)]} \quad (1)$$

$$Z(\theta) = \int_{X^m} h(x)\exp[\theta^T\cdot\phi(x)]\,dx$$

$$A(\theta) = \log Z(\theta)$$

$$e^{\log(u)} = u$$

$$\frac{1}{Z(\theta)} h(x)\exp[\theta^T\cdot\phi(x)] = h(x)\exp[\theta^T\phi(x)]\exp[\log Z(\theta)]$$

$$= h(x)\exp[\theta^T\cdot\phi(x) - A(\theta)]$$

$$\frac{d^2 A}{\partial\theta_i\,\partial\theta_j} = E[\phi_i(x)\phi_j(x)] - E[\phi_i(x)]\cdot E[\phi_j(x)]$$

$$= \mathrm{cov}[\phi_i, \phi_j]$$

$$\boxed{\nabla^2 A(\theta) = \mathrm{cov}[\phi(x)]}$$

secunda cov > 0 (def)

$\Rightarrow A(\theta)$ is convex

$$p(x|\theta) = \boxed{h(x)\exp[\theta^T \phi(x) - A(\theta)]}\ (\ast)$$

$$\frac{1}{Z(\theta)} h(x) \exp[\theta^T \phi(x)]\ (\ast)$$

$$Z(\theta) = \int_{x} h(x) \exp[\theta^T \phi(x)]\, dx$$

$$A(\theta) = \log Z(\theta)$$

$$\frac{1}{Z(\theta)} h(x) \exp[\theta^T \phi(x)] = h(x) \exp[\theta^T \phi(x)] \exp[-\log Z(\theta)]$$

$$= h(x) \exp[\theta^T \phi(x) - A(\theta)]$$

$$\frac{\partial^2 A}{\partial \theta_i \partial \theta_j} = E[\phi_i(x)\phi_j(x)] - E[\phi_i(x)] \cdot E[\phi_j(x)]$$

$$= \operatorname{cov}[\phi_i, \phi_j]$$

$$\boxed{\nabla^2 A(\theta) = \operatorname{cov}[\phi(x)]}$$

covariance matrix

$\Rightarrow A(\theta)$ convex