

430.714

Estimation Theory

Prof. Songhwai Oh
ECE, SNU

Exponential Family

Exponential Family

- Represents a broad class of distributions
- Includes Gaussian, exponential, gamma, chi-squared, beta, Dirichlet, Bernoulli, binomial, multinomial, Poisson, geometric, ...

$$p(\mathbf{x}|\boldsymbol{\eta}) = h(\mathbf{x})g(\boldsymbol{\eta}) \exp \{ \boldsymbol{\eta}^T \mathbf{u}(\mathbf{x}) \}$$

$\boldsymbol{\eta}$: natural parameters

$$g(\boldsymbol{\eta}) \int h(\mathbf{x}) \exp \{ \boldsymbol{\eta}^T \mathbf{u}(\mathbf{x}) \} d\mathbf{x} = 1$$

Example

$$p(x|\mu) = \text{Bern}(x|\mu) = \mu^x (1 - \mu)^{1-x}$$

$$\eta = \ln \left(\frac{\mu}{1 - \mu} \right)$$

$$\begin{aligned} p(x|\mu) &= \exp \{ x \ln \mu + (1 - x) \ln(1 - \mu) \} \\ &= (1 - \mu) \exp \left\{ \ln \left(\frac{\mu}{1 - \mu} \right) x \right\}. \end{aligned}$$

$$\mu = \sigma(\eta) \quad 1 - \sigma(\eta) = \sigma(-\eta)$$

$$\sigma(\eta) = \frac{1}{1 + \exp(-\eta)}$$

$$p(x|\eta) = \sigma(-\eta) \exp(\eta x)$$

$$u(x) = x$$

$$h(x) = 1$$

$$g(\eta) = \sigma(-\eta)$$

$$\eta = \ln \left(\frac{\mu}{1 - \mu} \right)$$

Maximum Likelihood Estimators and Sufficient Statistics

$$g(\boldsymbol{\eta}) \int h(\mathbf{x}) \exp \{ \boldsymbol{\eta}^T \mathbf{u}(\mathbf{x}) \} d\mathbf{x} = 1 \quad \nabla g(\boldsymbol{\eta}) \int h(\mathbf{x}) \exp \{ \boldsymbol{\eta}^T \mathbf{u}(\mathbf{x}) \} d\mathbf{x} \\ + g(\boldsymbol{\eta}) \int h(\mathbf{x}) \exp \{ \boldsymbol{\eta}^T \mathbf{u}(\mathbf{x}) \} \mathbf{u}(\mathbf{x}) d\mathbf{x} = 0$$

$$-\frac{1}{g(\boldsymbol{\eta})} \nabla g(\boldsymbol{\eta}) = \int h(\mathbf{x}) \exp \{ \boldsymbol{\eta}^T \mathbf{u}(\mathbf{x}) \} \mathbf{u}(\mathbf{x}) d\mathbf{x} = \mathbb{E}[\mathbf{u}(\mathbf{x})]$$

$$-\nabla \ln g(\boldsymbol{\eta}) = \mathbb{E}[\mathbf{u}(\mathbf{x})]$$

$$\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$$

$$p(\mathbf{X}|\boldsymbol{\eta}) = \left(\prod_{n=1}^N h(\mathbf{x}_n) \right) g(\boldsymbol{\eta})^N \exp \left\{ \boldsymbol{\eta}^T \sum_{n=1}^N \mathbf{u}(\mathbf{x}_n) \right\}$$

MLE:

$$-\nabla \ln g(\boldsymbol{\eta}_{\text{ML}}) = \frac{1}{N} \sum_{n=1}^N \mathbf{u}(\mathbf{x}_n)$$

Sufficient Statistic:

$$\sum_{n=1}^N \mathbf{u}(\mathbf{x}_n)$$

As $N \rightarrow \infty$, RHS $\rightarrow \mathbb{E}[\mathbf{u}(\mathbf{x})]$ by the law of large numbers.

Hence, $\boldsymbol{\eta}_{\text{ML}} \rightarrow \boldsymbol{\eta}$ as $N \rightarrow \infty$ as desired.

Exponential Family: a Gaussian random variable

$$\begin{aligned} p(x|\mu, \sigma^2) &= \frac{1}{(2\pi\sigma^2)^{1/2}} \exp \left\{ -\frac{1}{2\sigma^2} (x - \mu)^2 \right\} \\ &= \frac{1}{(2\pi\sigma^2)^{1/2}} \exp \left\{ -\frac{1}{2\sigma^2} x^2 + \frac{\mu}{\sigma^2} x - \frac{1}{2\sigma^2} \mu^2 \right\} \end{aligned}$$

$$p(\mathbf{x}|\boldsymbol{\eta}) = h(\mathbf{x})g(\boldsymbol{\eta}) \exp \{ \boldsymbol{\eta}^T \mathbf{u}(\mathbf{x}) \}$$

$$-\nabla \ln g(\boldsymbol{\eta}_{ML}) = \frac{1}{N} \sum_{n=1}^N \mathbf{u}(\mathbf{x}_n)$$

$$-\frac{\partial \log g(\boldsymbol{\eta})}{\partial \eta_1} = -\frac{\eta_1}{2\eta_2} = \frac{\mu/\sigma^2}{2 \cdot 1/(2\sigma^2)} = \mu$$

$$-\frac{\partial \log g(\boldsymbol{\eta})}{\partial \eta_2} = \sigma^2 + \mu^2$$

$$\begin{aligned} \boldsymbol{\eta} &= \begin{pmatrix} \mu/\sigma^2 \\ -1/2\sigma^2 \end{pmatrix} \\ \mathbf{u}(x) &= \begin{pmatrix} x \\ x^2 \end{pmatrix} \\ h(\mathbf{x}) &= (2\pi)^{-1/2} \\ g(\boldsymbol{\eta}) &= (-2\eta_2)^{1/2} \exp \left(\frac{\eta_1^2}{4\eta_2} \right) \end{aligned}$$

$$\hat{\mu}_{ML} = \frac{1}{N} \sum_{n=1}^N x_n$$