Lecture 3. Exponential Families (cont.)

Nan Ye

School of Mathematics and Physics
University of Queensland

Exponential Families

Recall: one parameter exponential family

An exponential family distribution has a PDF (or PMF for a discrete random variable) of the form

$$f(y \mid \theta, \phi) = \exp\left(\frac{\eta(\theta)T(y) - A(\theta)}{b(\phi)} + c(y, \phi)\right)$$

- $T(y)$ is called the *natural statistic*.
- $\eta(\theta)$ is called the *natural parameter*.
- ϕ is a *nuisance parameter* treated as known.

Range of y must be independent of θ.

This Lecture

- Expectation of natural statistics
- Variance of natural statistics
- Mean determines natural parameter

Properties

Expectation of natural statistics

The expected value of T is

$$\mathbb{E}(T) = A'(\theta)/\eta'(\theta).$$

In particular, for a natural exponential family (i.e. $\theta = \eta$),

$$\mathbb{E}(T) = A'(\eta).$$

Proof. For the discrete case, $\sum_y f(y;\theta,\phi) = 1$, that is,

$$\sum_y \exp\left(\frac{\eta(\theta)T(y) - A(\theta)}{b(\phi)} + c(y,\phi)\right) = 1$$

Multiply by $\exp(A(\theta)/b(\phi))$, and differentiate w.r.t. θ, we have

$$\sum_y \exp\left(\frac{\eta(\theta)T(y)}{b(\phi)} + c(y,\phi)\right)\frac{T(y)\eta'(\theta)}{b(\phi)} = \exp\left(\frac{A(\theta)}{b(\phi)}\right)\frac{A'(\theta)}{b(\phi)}.$$

Divide both sides by $\exp(A(\theta)/b(\phi))\eta'(\theta)/b(\phi)$, we obtain

$$\sum_y f(y \mid \theta,\phi)T(y) = A'(\theta)/\eta'(\theta).$$

This completes the proof. The continuous case is similar.

Variance of natural statistic

The variance of T is

$$\mathrm{var}(T) = \left(A''(\theta) - \mathbb{E}(T)\eta''(\theta)\right) \frac{b(\phi)}{\eta'(\theta)^2}.$$

In particular, for a natural exponential family (i.e. $\theta = \eta$),

$$\mathrm{var}(T) = A''(\eta) b(\phi).$$

Example 1. Gaussian distribution $N(\mu, \sigma^2)$ (known σ^2)

- Recall: $\eta(\mu) = \mu$, $T(y) = y$, $A(\mu) = \frac{1}{2}\mu^2$, $b(\sigma) = \sigma^2$.
- This is a natural exponential family.
- The mean is given by

$$\mathbb{E}(Y) = \mathbb{E}(T) = A'(\mu) = \mu.$$

- The variance is given by

$$\text{var}(Y) = \text{var}(T) = A''(\mu)b(\sigma) = \sigma^2.$$

- Compare with evaluating

$$\mathbb{E}(Y) = \int_{-\infty}^{+\infty} y \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy$$

$$\text{var}(Y) = \int_{-\infty}^{+\infty} (y-\mu)^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy$$

Example 2. Bernoulli distribution $B(p)$

- Recall: $\eta(p) = \ln \frac{p}{1-p}$, $T(y) = y$, $A(p) = -\ln(1-p)$, $b(\phi) = 1$.
- The mean is given by

$$\mathbb{E}(Y) = A'(p)/\eta'(p) = \frac{1}{1-p}\Big/\frac{1}{p(1-p)} = p.$$

- The variance is given by

$$\begin{aligned}
\text{var}(Y) = \text{var}(T) &= \left(A''(\theta) - \mathbb{E}(T)\eta''(\theta)\right)\frac{b(\phi)}{\eta'(\theta)^2} \\
&= \left(\frac{1}{(1-p)^2} - p\frac{2p-1}{p^2(1-p)^2}\right)\Big/\frac{1}{p^2(1-p)^2} \\
&= p(1-p)
\end{aligned}$$

- Recall (natural parametrization): $A(\eta) = \ln(1 + e^{\eta})$.
- The mean is given by

$$\mathbb{E}(Y) = \mathbb{E}(T) = A'(\eta) = \frac{1}{1 + e^{-\eta}} = p.$$

- The variance is given by

$$\mathrm{var}(Y) = \mathrm{var}(T) = A''(\eta) = \frac{e^{-\eta}}{(1 + e^{-\eta})^2} = p(1 - p).$$

We get the same answer using different parametrizations.

Mean determines natural parameter

For any natural exponential family with nonzero variance, the natural parameter η is uniquely determined by the mean $\mu = \mathbb{E}(T)$.

Put it in another way, we can always parametrize a natural exponential family using its mean.

Proof. It suffices to show that $\mu = \mathbb{E}(T) = A'(\eta)$ is strictly monotonic in η.
Since $A''(\eta) = \text{var}(T)/b(\phi)$ and $\text{var}(T) \neq 0$ by assumption, we have $\text{var}(T) > 0$, and thus $A''(\eta)$ is always positive if $b(\phi) > 0$, or always negative if $b(\phi) < 0$. This implies that $A'(\eta)$ is strictly monotonic, and thus $\eta = A'^{-1}(\mu)$.

Example

- For the Bernoulli distribution $B(p)$, given the mean p, we can uniquely determine the natural parameter $\eta = \ln \frac{p}{1-p}$.

What You Need to Know

- $\mathbb{E}(T) = A'(\theta)/\eta'(\theta)$.
- $\text{var}(T) = (A''(\theta) - \mathbb{E}(T)\eta''(\theta)) \frac{b(\phi)}{\eta'(\theta)^2}$.
- Mean determines natural parameter: $\eta = A'^{-1}(\mu)$.