Exponential Families

Exponential families and completeness

Exponential family:

$$f(x|\theta) = h(x)\, c(\theta) \exp\left\{ \sum_{i=1}^{k} w_i(\theta)\, t_i(x) \right\}$$

Suppose $X_1, \dots, X_n$ iid from an exponential family

Sufficiency: $T(X_1, \dots, X_n) = \left(\sum_{j=1}^{n} t_1(X_j), \dots, \sum_{j=1}^{n} t_k(X_j) \right)$

is sufficient for θ

Completeness: If, in addition, $\{(w_1(\theta), \dots, w_k(\theta))\}$ contain an open set in $\mathbb{R}^k$, then

$T(X_1, \dots, X_n) = \left(\sum_{j=1}^{n} t_1(X_j), \dots, \sum_{j=1}^{n} t_k(X_j) \right)$

is complete.

For Bernoulli: $w_1(p) = \log\left(\frac{p}{1-p}\right)$

Poisson: $w_1(p) = \log(p)$

Example

Suppose that $X_1, \ldots, X_n \overset{iid}{\sim} Bernoulli(p)$.

- Show the the Bernoulli is an exponential family distribution
- Find a complete, sufficient statistic
- Use the complete, sufficient statistic to find a best unbiased estimator of p

$$f(x|p) = p^x (1-p)^{1-x} = (1-p)\left(\frac{p}{1-p}\right)^x$$

$$= \underbrace{(1-p)}_{c(p)} \exp\Big\{ \underbrace{x}_{t_1(x)} \underbrace{\log\left(\tfrac{p}{1-p}\right)}_{w_1(p)} \Big\} \qquad h(x) = 1$$

$T(X_1, \ldots, X_n) = \sum_{i=1}^{n} t_1(X_i) = \sum_i X_i$ is sufficient

and $\left\{ \log\left(\frac{p}{1-p}\right) : p \in [0,1] \right\}$ contains an open set in $\mathbb{R}$

so $\sum_i X_i$ is also complete $\Rightarrow$ $\hat{p} = \frac{1}{n}\sum_i X_i$ is the best unbiased est.

Example

Suppose that $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$.

- Show the the normal is an exponential family distribution
- Find a complete, sufficient statistic
- Use the complete, sufficient statistic to find a best unbiased estimator of σ^2

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ - \frac{(x-\mu)^2}{2\sigma^2} \right\}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ - \frac{\mu^2}{2\sigma^2} \right\} \exp\left\{ - \frac{x^2}{2\sigma^2} + \frac{\mu x}{\sigma^2} \right\}$$

$h(x) = 1$ $\qquad c(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{\mu^2}{2\sigma^2} \right\}$

$w_1(\mu, \sigma^2) = -\frac{1}{2\sigma^2}$ $\qquad t_1(x) = x^2$

$w_2(\mu, \sigma^2) = \frac{\mu}{\sigma^2}$ $\qquad t_2(x) = x$

complete, sufficient statistic: $\left(\sum_i X_i^2 , \sum_i X_i \right)$

$$s^2 = \frac{1}{n-1} \sum_i (X_i - \bar{X})^2 = \frac{1}{n-1} \left(\sum_i X_i^2 - n\left(\frac{1}{n} \sum_i X_i \right)^2 \right)$$

s^2 is a function of complete, sufficient statistic ✓

$\mathbb{E}[s^2] = \sigma^2$ (unbiased for σ^2) ✓

$\Rightarrow$ if μ is unknown, s^2 is the best unbiased estimator of σ^2