Biostatistics 602 - Statistical Inference
Lecture 07
Exponential Family

Hyun Min Kang

January 31st, 2013

Last Lecture

1. What are differences between complete statistic and minimal sufficient statistics?
2. What is the relationship between complete statistic and ancillary statistics?
3. What is the characteristic shared among non-constant functions of complete statistics?
4. What is the Basu's Theorem?
5. Any example where Basu's Theorem is helpful?

Exponential Family

Definition 3.4.1

The random variable X belongs to an exponential family of distributions, if its pdf/pmf can be written in the form

$$f(x|\boldsymbol{\theta}) = h(x)c(\boldsymbol{\theta})\exp\left[\sum_{j=1}^{k} w_j(\boldsymbol{\theta})t_j(x)\right]$$

where

- $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_d),\ d \leq k$.
- $w_j(\theta),\ j \in \{1, \cdots, k\}$ are functions of $\boldsymbol{\theta}$ alone.
- and $t_j(x)$ and $h(x)$ only involve data.

Example of Exponential Family

Problem

Show that a Poisson(λ) ($\lambda > 0$) belongs to the exponential family

Proof

$$f_X(x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{1}{x!}e^{-\lambda}\exp\left(\log\lambda^x\right) = \frac{1}{x!}e^{-\lambda}\exp\left(x\log\lambda\right)$$

Define $h(x) = 1/x!$, $c(\lambda) = e^{-\lambda}$, $w(\lambda) = \log\lambda$, and $t(x) = x$, then

$$f_X(x|\lambda) = h(x)c(\lambda)\exp\left[w(\lambda)t(x)\right]$$

Normal Distribution Belongs to an Exponential Family

$$f_X(x|\boldsymbol{\theta} = (\mu, \sigma^2)) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{x^2}{2\sigma^2} + \frac{2\mu x}{2\sigma^2} - \frac{\mu^2}{2\sigma^2}\right]$$

Define $h(x) = 1,\ c(\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{\mu^2}{2\sigma^2}\right]$,
$k = 2,\ w_1(\boldsymbol{\theta}) = \frac{\mu}{\sigma^2},\ t_1(x) = x,\ w_2(\boldsymbol{\theta}) = -\frac{1}{2\sigma^2},\ t_2(x) = x^2$, then

$$f_X(x|\boldsymbol{\theta}) = h(x)c(\boldsymbol{\theta}) \exp\left[\sum_{j=1}^{k} w_j(\boldsymbol{\theta})t_j(x)\right]$$

A Specialized Normal Distribution : $\mathcal{N}(\mu, \mu^2)$

$$f_X(x|\mu) = \frac{1}{\sqrt{2\pi\mu^2}} \exp\left[-\frac{(x-\mu)^2}{2\mu^2}\right] = \frac{1}{\sqrt{2\pi\mu^2}} \exp\left[-\frac{x^2}{2\mu^2} + \frac{2\mu x}{2\mu^2} - \frac{\mu^2}{2\mu^2}\right] = \frac{1}{\sqrt{2\pi\mu^2}} \exp\left(-\frac{1}{2}\right) \exp\left[-\frac{1}{2\mu^2}x^2 + \frac{1}{\mu}x\right]$$

Define $h(x) = 1,\ c(\mu) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}}$,
$k = 2,\ w_1(\mu) = \frac{1}{\mu},\ t_1(x) = x,\ w_2(\mu) = -\frac{1}{2\mu^2},\ t_2(x) = x^2$, then

$$f_X(x|\mu) = h(x)c(\mu) \exp\left[\sum_{j=1}^{k} w_j(\mu)t_j(x)\right]$$

Curved and Full Exponential Families

Definition

For an exponential family, if $d = \dim(\boldsymbol{\theta}) < k$, then this exponential family is called *curved exponential family*. if $d = \dim(\boldsymbol{\theta}) = k$, then this exponential family is called *full exponential family*.

Examples

- Poisson(λ), $\lambda > 0$ is a full exponential family
- $\mathcal{N}(\mu, \sigma^2), \mu \in \mathbb{R}, \sigma > 0$ is also a full exponential family
- $\mathcal{N}(\mu, \mu), \mu \in \mathbb{R}$ is also a curved exponential family

Alternative Parametrization of Exponential Families

An alternative parametrization of the exponential family of distributions in terms of "natural" or "canonical" parameters can be written as follows.

$$f_X(x|\boldsymbol{\eta}) = h(x)c^*(\boldsymbol{\eta}) \exp\left[\sum_{j=1}^{k} \boldsymbol{\eta} t_j(x)\right]$$

The alternative parametrization can be achieved by defining $\eta_i = w_j(\boldsymbol{\theta})$ from the following equation,

$$f_X(x|\boldsymbol{\theta}) = h(x)c(\boldsymbol{\theta}) \exp\left[\sum_{j=1}^{k} w_j(\boldsymbol{\theta})t_j(x)\right]$$

where $c^*(\boldsymbol{\eta}) = c \circ w(\boldsymbol{\theta})$. This alternative parametrization is most often used in a GLM (Generalized Linear Model) course.

Sufficient Statistic for Exponential Families

Theorem 6.2.10

- Let $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} f_X(x|\boldsymbol{\theta})$ that belongs to an exponential family given by

$$f_X(x|\boldsymbol{\theta}) = h(x)c(\boldsymbol{\theta})\exp\left[\sum_{j=1}^{k} w_j(\boldsymbol{\theta})t_j(x)\right]$$

  where $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_d), \ d \leq k$.
- Then the following $T(\mathbf{X})$ is a sufficient statistic for $\boldsymbol{\theta}$.

$$T(\mathbf{X}) = \left(\sum_{j=1}^{n} t_1(X_j), \cdots, \sum_{j=1}^{n} t_k(X_j)\right)$$

Sufficient Statistic for Normal Distribution

Problem

Let $X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$, where $\mu \in \mathbb{R}$, and σ^2 is known. Find a sufficient statistic for μ.

Sufficient Statistic for Normal Distribution - Solution

$$\begin{aligned} f_X(x|\mu) &= \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \\ &= \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{x^2}{2\sigma^2}\right)\exp\left(-\frac{\mu^2}{2\sigma^2}\right)\exp\left(\frac{\mu}{\sigma^2}x\right) \\ &= h(x)c(\mu)\exp\left[w(\mu)t(x)\right] \end{aligned}$$

where

$$\begin{cases} h(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{x^2}{2\sigma^2}\right] \\ c(\mu) = \exp\left(-\frac{\mu^2}{2\sigma^2}\right) \\ w(\mu) = \mu/\sigma^2 \\ t(x) = x \end{cases}$$

Therefore, $T(\mathbf{X}) = \sum_{i=1}^{n} t(X_i) = \sum_{i=1}^{n} X_i$ is a sufficient statistic for μ by Theorem 6.2.10.

Theorem 5.2.11

Suppose $X_1, \cdots, X_n$ is a random sample from pdf or pmf $f_X(x|\theta)$ where

$$f_X(x|\boldsymbol{\theta}) = h(x)c(\boldsymbol{\theta})\exp\left[\sum_{j=1}^{k} w_j(\boldsymbol{\theta})t_j(x)\right]$$

is a member of an exponential family. Define a statistic $T(\mathbf{X})$ by

$$\mathbf{T}(\mathbf{X}) = \left(\sum_{j=1}^{n} t_1(X_j), \cdots, \sum_{j=1}^{n} t_k(X_j)\right)$$

If the set $\{w_1(\boldsymbol{\theta}), \cdots, w_k(\boldsymbol{\theta}), \ \forall \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$ contains an open subset of $\mathbb{R}^k$, then the distribution of $\mathbf{T}(\mathbf{X})$ is an exponential family of the form

$$f_T(u_1, \cdots, u_k|\boldsymbol{\theta}) = H(u_1, \cdots, u_k)[c(\boldsymbol{\theta})]^n \exp\left[\sum_{j=1}^{k} w_j(\boldsymbol{\theta})u_i\right]$$

Theorem 6.2.25

Suppose $X_1, \cdots, X_n$ is a random sample from pdf or pmf $f_X(x|\theta)$ where

$$f_X(x|\boldsymbol{\theta}) = h(x)c(\boldsymbol{\theta})\exp\left[\sum_{j=1}^{k} w_j(\boldsymbol{\theta})t_j(x)\right]$$

is a member of an exponential family. Then the statistic $T(\mathbf{X})$

$$\mathbf{T}(\mathbf{X}) = \left(\sum_{j=1}^{n} t_1(X_j), \cdots, \sum_{j=1}^{n} t_k(X_j)\right)$$

is complete as long as the parameter space $\boldsymbol{\Theta}$ contains an open set in $\mathbb{R}^k$

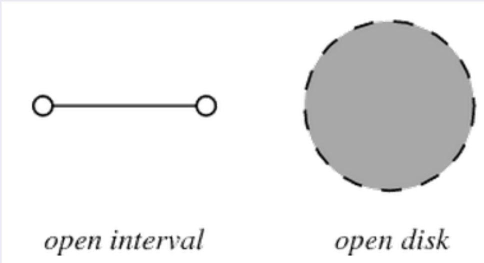
What is the "open set"?

Definition : Open Set

A set A is open in $\mathbb{R}^k$ of for every $x \in A$, there exists a ϵ-ball $B(x, \epsilon)$ around x such that $B(x, \epsilon) \subset A$.

$$B(x, \epsilon) = \{y : |y - x| < \epsilon, \ y \in \mathbb{R}^k\}$$

Examples (from Wolfram MathWorld)



Examples of open set test

- $A = (-1, 1)$: A is open in $\mathbb{R}$
- $A = [-1, 1]$: A is not open in $\mathbb{R}$
- $A = (-\infty, 0) \times \mathbb{R}$: A is open in $\mathbb{R}^2$
- $A = (-\infty, 0] \times \mathbb{R}$: A is not open in $\mathbb{R}^2$
- $A = \{(x, y) : x \in (-1, 1), y = 0\}$: A is not open in $\mathbb{R}^2$
- $A = \{(x, y) : x \in \mathbb{R}, y \in x^2\}$: A is not open in $\mathbb{R}^2$
- $A = \{(x, y) : x^2 + y^2 < 1\}$: A is open in $\mathbb{R}^2$

Exponential Family Example

Problem

$X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} f_X(x|\theta) = \theta x^{\theta-1}$ where $0 < x < 1, \theta > 0$. Is $\prod_{i=1}^n X_i$ (1) a sufficient statistic? (2) a complete statistic? (3) a minimal sufficient statistic?

How to solve it

- Show that $f_X(x|\theta)$ is a member of an exponential family.
- Apply Theorem 6.2.10 to obtain a sufficient statistic and see if it is equivalent to or related to $\prod_{i=1}^n X_i$.
- Apply Theorem 6.2.25 to show that it is complete.
- If they are both sufficient and complete, Theorem 6.2.28 will imply that it is also a minimal sufficient statistic.

$f_X(x|\theta)$ belong to an exponential family

$$\begin{aligned} f_X(x|\theta) &= \theta x^{\theta-1} I(0<x<1) \\ &= I(0<x<1)x^{-1}\theta x^{\theta} \\ &= I(0<x<1)x^{-1}\theta \exp(\log x^{\theta}) \\ &= I(0<x<1)x^{-1}\theta \exp(\theta \log x) \\ &= h(x)c(\theta)\exp(w(\theta)t(x)) \end{aligned}$$

where

$$\begin{cases} h(x) = I(0<x<1)x^{-1} \\ c(\theta) = \theta \\ w(\theta) = \theta \\ t(x) = \log x \end{cases}$$

Therefore, $f_X(x|\theta)$ belongs to an exponential family.

Apply Theorem 6.2.10

$$f_X(x|\theta) = h(x)c(\theta)\exp(w(\theta)t(x))$$

$$\begin{cases} h(x) = I(0<x<1)x^{-1} \\ c(\theta) = \theta \\ w(\theta) = \theta \\ t(x) = \log x \end{cases}$$

By Theorem 6.2.10, $T(\mathbf{X}) = \sum_{i=1}^n t(X_i) = \sum_{i=1}^n \log X_i$ is a sufficient statistic for θ.

$$\prod_{i=1}^n X_i = \exp\left(\log \prod_{i=1}^n X_i\right) = \exp\left(\sum_{i=1}^n \log X_i\right) = e^{T(\mathbf{X})}$$

Because $\prod_{i=1}^n X_i$ is an one-to-one function of $T(\mathbf{X})$, it is also a sufficient statistic.

Apply Theorem 6.2.25 and 6.2.28

$T(\mathbf{X})$ is a complete statistic

Let $A = \{w(\theta) : \theta \in \Omega\} = \{\theta : \theta > 0\}$. A contains an open subset in $\mathbb{R}$. By Theorem 6.2.25, $T(\mathbf{X}) = \sum_{i=1}^n \log X_i$ is a complete statistic for θ.

$T(\mathbf{X})$ is a minimal sufficient statistic

By Theorem 6.2.28, because $T(\mathbf{X})$ is both sufficient and complete, it is also minimal sufficient.

$\prod_{i=1}^n X_i = e^{T(\mathbf{X})}$ is also minimal sufficient and complete

Because $\prod_{i=1}^n X_i = e^{T(\mathbf{X})}$ is an one-to-one function of $T(\mathbf{X})$, $\prod_{i=1}^n X_i$ is sufficient, complete, and minimal sufficient.

Exponential Family

Today

- Curved and full exponential families
- Sufficient statistics for exponential families
- Complete statistics for exponential families

Next Lecture

- Review of Chapter 6