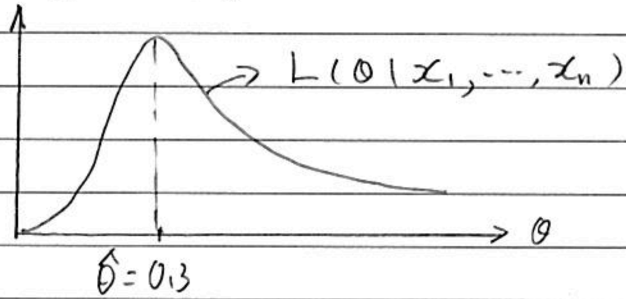
According to graph 0.3 is more likely $\hat{\theta} = 0.3 \rightarrow ML\theta$



$\Rightarrow$ Chapter 5. special Model

Exponential families

$\Rightarrow$ Defination: $f(x \mid \theta)$ is said to be a model in exponential family if

$$f(x \mid \theta) = C(\theta)\, h(x) \cdot e^{\sum_{i=1}^{K} \pi_i(\theta) \cdot T_i(x)}$$

Note: $\theta = (\theta_1, \theta_2, \cdots \theta_d)$ d may be $< K$

$\Rightarrow$ Example:

1) $X_1, \cdots, X_n \mid \theta \overset{iid}{\sim} \exp(\theta)$ θ is scale

$X = (X_1, \cdots, X_n)$

$$f(x \mid \theta) = \prod_{i=1}^{n} \frac{1}{\theta} e^{-\frac{x_i}{\theta}} = \theta^{-n} e^{-\frac{1}{\theta} \sum_{i=1}^{n} x_i}$$

$\pi_i(\theta) = -\frac{1}{\theta}$ $T_i(x) = \sum_{i=1}^{n} x_i$ $C(\theta) = \theta^{-n}$ $h(x) = 1$

2) $X_1, \cdots x_n \mid \theta \overset{iid}{\sim} \text{Gamma}(\alpha, \theta)$

$$f(x \mid \theta) = \prod_{i=1}^{n} \frac{x_i^{\alpha-1}}{\Gamma(\alpha) \cdot \theta^{\alpha}} e^{-\frac{x_i}{\theta}}$$

$$= \underbrace{\left(\prod_{i=1}^{n} x_i\right)^{\alpha-1} \left(\Gamma(\alpha)\right)^{-n}}_{h(x)} \underbrace{\theta^{-n\alpha}}_{C(\theta)} \cdot e^{\underbrace{-\frac{1}{\theta}}_{\Pi_i(\theta)} \underbrace{\sum_{i=1}^{n} x_i}_{T_i(x)}}$$

③ $X_1, \cdots, X_n \mid \theta \overset{iid}{\sim} \text{Beta}(a, b)$

$$\theta = (a, b), \quad f(x\mid\theta) = \prod_{i=1}^{n} \frac{1}{\text{Beta}(a,b)} x_i^{a-1} \cdot (1-x_i)^{b-1}$$

$$f(x\mid\theta) = \{\text{Beta}(a,b)\}^{-n} \cdot \left(\prod_{i=1}^{n} x_i\right)^{a-1} \cdot \left(\prod_{i=1}^{n} (1-x_i)\right)^{b-1}$$

$$= \underbrace{\{\text{Beta}(a,b)\}^{-n}}_{C(\theta)} \cdot e^{\underbrace{(a-1)}_{\Pi_1(\theta)} \underbrace{\sum \log x_i}_{T_1(x)} + \underbrace{(b-1)}_{\Pi_2(\theta)} \underbrace{\sum \log(1-x_i)}_{T_2(x)}}$$

$k = d$:
how many terms in the exponential

④ $X_1, \cdots X_n \mid \theta \overset{iid}{\sim} N(\mu, \sigma^2) \qquad \theta = (\mu, \sigma^2)$

$$f(x\mid\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\sum \frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{\sum(x_i^2 - 2\mu x_i + \mu^2)}{2\sigma^2}}$$

$$= \underbrace{(2\pi\sigma^2)^{-\frac{n}{2}} \cdot e^{-\frac{n\mu^2}{2\sigma^2}}}_{C(\theta)} \cdot e^{\underbrace{\left(-\frac{1}{2\sigma^2}\right)}_{\Pi_1(\theta)} \underbrace{\sum x_i^2}_{T_1(x)} + \underbrace{\frac{\mu}{\sigma^2}}_{\Pi_2(\theta)} \underbrace{\sum x_i}_{T_2(x)}}$$

⑤ $X_1, \cdots, X_n \mid \theta \sim \text{Bern}(\theta)$

$$f(x\mid\theta) = \prod_{i=1}^{n} \theta^{x_i} \cdot (1-\theta)^{1-x_i}$$

$$= e^{\log\left(\prod_{i=1}^{n} \theta^{x_i} \cdot (1-\theta)^{1-x_i}\right)}$$

$$= e^{\underbrace{\log\theta}_{\Pi_1(\theta)} \cdot \underbrace{\Sigma x_i}_{T_1(x)} + \underbrace{\log(1-\theta)}_{\Pi_2(\theta)} \cdot \underbrace{(n-\Sigma x_i)}_{T_2(x)}}$$

$d=1, \quad k=2$

A model is not in exponential family:

$X \mid \theta \sim \text{Cauchy}(\theta)$

$$f(x \mid \theta) = \frac{1}{\pi} \frac{1}{1+(x-\theta)^2}$$

$$= \frac{1}{\pi} \cdot e^{-\log(1+(x-\theta)^2)}$$

cannot be written as $\Sigma \Pi_i(\theta) T_i(x)$

$\Rightarrow \int f(x \mid \theta)\, dx = 1$

Suppose we can change $\frac{\partial}{\partial\theta}$ and $\int$

[Sometimes we cannot

$X \mid \theta \sim \text{unif}(0, \theta)$

$$\int_0^\theta \frac{1}{\theta}\, dx = 1 \Rightarrow \frac{\partial}{\partial\theta} \int_0^\theta \frac{1}{\theta}\, dx \neq \int_0^\theta \frac{\partial}{\partial\theta}\left(\frac{1}{\theta}\right) dx \;]$$

$$E_{x|\theta}\left(\frac{\partial}{\partial\theta_j} \log f(x \mid \theta) \mid \theta\right) = 0 \quad \text{for } j = 1, 2, \dots d.$$

$$= E_{x|\theta}\left\{\frac{\partial}{\partial\theta_i \partial\theta_j} \log f(x \mid \theta) \mid \theta\right\} = -E_{x|\theta}\left(\frac{\partial \log f}{\partial\theta_i} \cdot \frac{\partial \log f}{\partial\theta_j}\right)$$

proof: Right hand side

$$= E\left(\frac{\partial \log f}{\partial\theta_j} \mid \theta\right) = \int \frac{\partial \log f}{\partial\theta_j} \cdot f(x \mid \theta)\, dx$$

$$= \int \frac{1}{f(x|\theta)} \cdot \frac{\partial f(x|\theta)}{\partial \theta_j} \cdot f(x|\theta)\, dx$$

$$= \int \frac{\partial f(x|\theta)}{\partial \theta_j}\, dx \quad = \frac{\partial}{\partial \theta_j} \int f(x|\theta)\, dx$$

$$= \frac{\partial}{\partial \theta_j}(1) \quad = 0$$

$\Rightarrow$ If $f(x|\theta) = C(\theta) \cdot h(x) \cdot \exp^{\sum \Pi_i(\theta) T_i(x)}$

$$\log f(x|\theta) = \log C(\theta) + \log h(x) + \sum \Pi_i(\theta) \cdot T_i(x)$$

$$\frac{\partial \log f(x|\theta)}{\partial \theta_j} = \frac{1}{C(\theta)} \cdot \frac{\partial C(\theta)}{\partial \theta_j} + \sum \frac{\partial \Pi_i(\theta)}{\partial \theta_j} \cdot T_i(x)$$

So $E\left(\underset{0}{\underset{\shortparallel}{\frac{\partial \log f(x|\theta)}{\partial \theta_j}}}\right) = \frac{1}{C(\theta)} \cdot \frac{\partial C(\theta)}{\partial \theta_j} + \sum \frac{\partial \Pi_i(\theta)}{\partial \theta_j} \cdot E(T_i(x)) = 0$

for $i = 1, \cdots, K$

we can find $E[T_i(x)]$ by solving these equations.

Example : $f(x|\mu, \sigma^2) = e^{-\frac{1}{2\sigma^2}\sum x_i^2 - \frac{\mu}{\sigma^2}\sum x_i}$

$$E(\sum x_i^2) = 0 \quad , \quad E(\sum x_i) = 0$$

$$E(T_i(x)) = 0$$

<< Mar 06. 2015 >> STAT 846

⇒ if $f(x|\theta)$ is in the exponential family; then

$$\frac{\partial}{\partial\theta_j}\log C(\theta) + \sum_{i=1}^{k}\frac{\partial\pi_i}{\partial\theta_j}E(T_i(x)) = 0 \text{ for } j=1,2,\dots d$$

⇒ Def: Natural parameterization

Suppose $f(x|\theta) = C(\theta)h(x)e^{\sum_{i=1}^{k}\pi_i(\theta)T_i(x)}$

Natural parameter

Let $\boxed{\eta_i = \pi_i(\theta)}$, $\eta = (\eta_1, \cdots \eta_n)$

$f(x|\eta) = C'(\eta)h(x)e^{\sum \eta_i T_i(x)}$

in such case, we say $f(x|\eta)$ is in its natural parameterization.

⇒ Example: $f(x|\theta) = (2\pi)^{-\frac{n}{2}}\cdot(\sigma^2)^{-\frac{n}{2}}e^{-\frac{n\mu^2}{2\sigma^2}}$

$$\cdot e^{-\frac{\sum x_i^2}{2\sigma^2} + \frac{\mu}{\sigma^2}\sum x_i} \qquad \theta = (\mu, \sigma^2)$$

Let $\eta_1 = -\frac{1}{2\sigma^2}$, $\eta_2 = \frac{\mu}{\sigma^2}$ $\quad (\mu, \sigma^2) \Leftrightarrow (\eta_1, \eta_2)$

$f(x|\eta) = (2\pi)^{-\frac{n}{2}}\hat{C}(\eta)e^{\eta\sum x_i^2 + \eta_2\sum x_i}$

⇒ Def: Natural Parameterization space:

$$\Pi = \{\eta = (\pi_1(\theta), \cdots \pi_k(\theta)) \mid \theta \text{ such that } C(\theta) \text{ is finite}\}$$

K is dimention of Π (Natural parameterization space)

$\theta = (\theta_1, \theta_2, \cdots \theta_d)$ $\quad \Pi = (\Pi_1(\theta), \cdots \Pi_K(\theta))$

$\underbrace{\theta_d}_{d}$ $\quad \underbrace{\Pi_K}_{K}$

if $d = K$ $\boxed{\text{dimention}(\theta) = \text{dimention}(\Pi)}$, then we say $f(x|\theta)$ is in full exponential family

If $d < K$, then we say $f(x|\theta)$ is in curved exponential family. (Reduce)

=> example:

1) $X_1, \cdots X_n | \mu, \sigma^2 \overset{iid}{\sim} N(\mu, \sigma^2)$, it is in full exponential family

2) $X_1, \cdots, X_n | \theta \sim$ Bernuli (θ)

$$f(x|\theta) = \prod_{i=1}^{n} \theta^{x_i} (1-\theta)^{1-x_i}$$

$$= e^{\log\theta \sum x_i + \log(1-\theta) \sum (1-x_i)}$$

here, $\Pi_1 = \log(\theta)$ $\quad \Pi_2 = \log(1-\theta)$

(Π_1, Π_2) $\quad K = 2$ $\quad d = 1$ (only θ)

therefore, $f(x|\theta)$ is in a curved exponential family

=> A fact: {E, and Cov. for natural exponential family}

If $f(x|\theta)$ is in not parameterization, i.e.

$$f(x|\theta) = C(\theta) h(x) \cdot e^{\sum_{i=1}^{K} \theta_i T_i(x)}, \text{ where, } \Pi_i(\theta) = \theta_i$$

$$\frac{\partial}{\partial \theta_j} \log(C(\theta)) + \sum_{i=1}^{K} \frac{\partial \Pi_i}{\partial \theta_j} E(T_i(x)) = 0$$

$$\frac{\partial}{\partial \theta_j} \log C(\theta) + E(T_j(x)) = 0$$

$$E(T_j(x)) = -\frac{\partial}{\partial \theta_j} \log C(\theta)$$

Also, we can show

$$\operatorname{cov}(T_i(x), T_j(x) \mid \theta) = -\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log C(\theta)$$

$\Rightarrow$ Lemma 5.1

Suppose that, $f(x|\theta) = C(\theta) h(x) e^{\sum_{i=1}^{k} \pi_i(\theta) T_i(x)}$

the the joint distribution of $(T_1, \cdots, T_k)$ is

in exponential family. natural parameter $(\pi_1, \cdots \pi_k)$

proof: Let x be discrete

$$\Pr(T_1 = y_1, \cdots, T_k = y_k \mid \theta) = \sum_{\left\{x \,\middle|\, \begin{array}{c} T_1(x) = y_1 \\ \vdots \\ T_k(x) = y_k \end{array}\right\}} \Pr(X = x \mid \theta)$$

$$= \sum_{\{x \mid T_i(x) = y_i\}} C^n(\theta) h(x) \underbrace{e^{\sum_{i=1}^{k} \pi_i(\theta) T_i(x)}}_{\text{constant over } \Sigma}$$

$$= C^n(\theta) \left(\sum_{\{x \mid T_i(x) = y_i} h(x) \right) e^{\sum \pi_i(\theta) \cdot y_i}$$

$$= C^n(\theta) \cdot \tilde{h}(y) e^{\sum \pi_i(\theta) \cdot y_i}$$

$\Rightarrow$ Lemma 5.2

Suppose that $f(x|\theta) = C(\theta) \cdot h(x) \cdot e^{\sum \pi_i(\theta) T_i(x)}$

Let S be a subset of $\{1, 2, \cdots, k\}$, then the

discrete $T_i \quad i \in S \mid T_j \quad j \notin S$

is of exponential family with natural parameterization

$\{ \pi_i(\theta) \mid i \in S \}$

$\Rightarrow$ proof: Let x be discrete, Let $S = \{1, 2, \dots \ell\}$

$Pr(T_1 = y_1, \dots T_k = y_k \mid \theta)$

$= C(\theta) \tilde{h}(y) e^{\sum \pi_i(\theta) y_i}$

$Pr(T_1 = y_1, \dots T_\ell = y_\ell \mid T_{\ell+1} = y_{\ell+1}, \dots T_k = y_k, \theta)$

$= Pr(T_1 = y_1, \dots T_k = y_k) \Big/ \sum_{y_1 \dots y_\ell} [Pr(T_1 = y_1 \dots T_k = y_k \mid \theta)]$

$= C(\theta) \tilde{h}(y) \prod e^{\pi_i(\theta) y_i} \Big/ \sum_{y_1 \dots y_\ell} C(\theta) \tilde{h}(y) \prod e^{\pi_i(\theta) y_i}$

$$= \frac{\tilde{h}(y)}{\left(\sum_{y_1 \dots y_\ell} \tilde{h}(y) \cdot e^{\sum_{i=1}^{\ell} \pi_i(\theta) y_i} \right)} \cdot e^{\sum_{i=1}^{\ell} \pi_i(\theta) y_i}$$

$\uparrow$

free of $y_1, \dots y_\ell$, related to θ i.e. can be written as $\tilde{C}(\theta)^{-1}$

<<Chapter 6>>: Sufficiency and Completness

<< Mar 09. 2015 >> STAT 846

⇒ Sufficient statistics

likelihood function: $L(\theta; x) = f(x|\theta)$

θ is variable, x is fixed.

⇒ Example: $X_1, \cdots X_n | \mu, \sigma^2 \sim N(\mu, \sigma^2)$

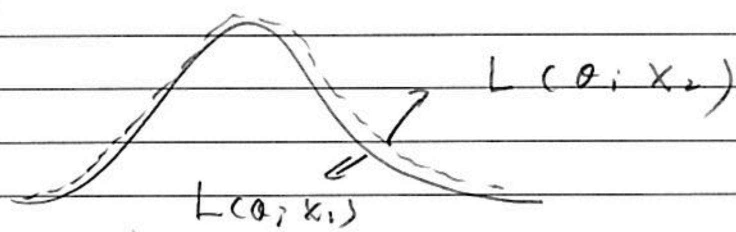
$x = (X_1, \cdots, X_n)$, $\theta = (\mu, \sigma^2)$

$$L(\theta, x) = f(x|\theta) = (2\pi)^{-\frac{n}{2}} \cdot (\sigma^2)^{-\frac{n}{2}} \cdot e^{\frac{-n\mu}{2\sigma^2}} e^{-\frac{1}{2\sigma^2}\Sigma x_i^2 + \frac{\mu}{\sigma^2}\Sigma x_i}$$

$$= (2\pi)^{-\frac{n}{2}} (\sigma^2)^{-\frac{n}{2}} \cdot e^{-\frac{1}{2\sigma^2}[n(\mu-\bar{x})^2 + (n-1)v^2]}$$

⇒ informally: $L(\theta; x)$ is determined by $\bar{x}$ and v^2, it means that for two data sets X_1 and X_2 of given, if $\begin{cases} \bar{X}_1 = \bar{X}_2 \\ V_1^2 = V_2^2 \end{cases}$, then

$L(\theta; x_1) = L(\theta; x_2)$



⇒ Definition:
A stististics $T = T(x)$ is sufficient for θ, if one of the following conditions are true:

1) Factorization:

$$f(x; \theta) = \underbrace{h(x)}_{\text{irrelavant to } \theta} g(t(x); \theta)$$

2) likelihood ratio theorem.

For any pair of data sets x and x'

such that $t(x) = t(x')$, we have

$$\Lambda_x(\theta_1, \theta_2) = \Lambda_{x'}(\theta_1, \theta_2)$$

where, $\Lambda_x(\theta_1, \theta_2) = \frac{f(\theta_2, x)}{f(\theta_1, x)}$

3) $f(x \mid t(x), \theta)$ is independent of θ

i.e $f(x \mid t(x), \theta) = f(x \mid t(x))$

$\Rightarrow$ Lamma 6.1

Condition 1 $\Leftrightarrow$ Condition 2

pf: For 1) $\Rightarrow$ 2) suppose $t(x) = t(x')$

$$f(x; \theta) = h(x)\, g(t(x), \theta)$$

$$\Lambda_x(\theta_1, \theta_2) = h(x) \cdot g(t(x), \theta_2) \,/\, h(x) \cdot g(t(x), \theta_1)$$

$$= g(t(x), \theta_2) \,/\, g(t(x), \theta_1)$$

$$\Lambda_{x'}(\theta_1, \theta_2) = g(t(x'), \theta_2) \,/\, g(t(x'), \theta_1)$$

$$\Rightarrow \Lambda_x(\theta_1, \theta_2) = \Lambda_{x'}(\theta_1, \theta_2)$$

For 2) $\Rightarrow$ 1)

pf: Fixed θ_0, for any θ

$\dfrac{f(x;\theta)}{f(x,\theta_0)}$ as a function of θ; is the same

as $\dfrac{f(x',\theta)}{f(x',\theta_0)}$ as long as $t(x)=t(x')$

it implies $\dfrac{f(x,\theta)}{f(x,\theta_0)} = g^*(t(x);\theta,\theta_0)$

$$f(x,\theta) = f(x,\theta_0)\cdot g^*(t(x);\theta,\theta_0)$$

$$= h(x)\cdot g(t(x);\theta)$$

$\Rightarrow$ Condition 3) $\Leftrightarrow$ Condition 1)

from 1) $\Rightarrow$ 3)

$f(x|\theta) = h(x)\cdot g(t(x);\theta)$

suppose that x is discrete for all x, t such that $t=t(x)$

$$f(x|t(x),\theta) = f(x, t=t(x)|\theta) \,/\, f(t(x)=t|\theta)$$

$$= f(x|\theta) \,/\, \sum_{t(x)=t} f(x|\theta)$$

$$= h(x)\cdot \underline{g(\overset{=t}{t(x)};\theta)} \,/\, \sum_{t(x)=t} h(x)\cdot \underbrace{g(t(x)=t,\theta)}_{\text{constant}}$$

$$= \frac{h(x)}{\sum h(x)} \quad \text{free of } \theta$$

example for above pf:

$$P(x_1=1, x_2=2 \mid x_1+x_2=3,\theta) = \frac{P(x_1=1, x_2=2, x_1+x_3=3|\theta)}{P(x_1+x_2=3|\theta)}$$

Form 3) $\Rightarrow$ 1)

$f(x|t(x),\theta)$ is free at θ

$\exists\, h(x)$ such that $f(x|t(x),\theta) = h(x)$

For $\cancel{X}$, t such that $t(x) = t$, then $f(x|\theta) = f(x, t(x) | \theta)$

$= f(t(x) = t | \theta) \cdot f(x | t(x) = t, \theta)$

by product rule

$= h(x) \cdot g(t(x) | \theta)$

« Mar 11, 2015 » STAT 846.

⇒ Definition: T is a sufficient statistic if for any pair x, x' with $T(x) = T(x')$, then.

$\Lambda_x(\theta_1, \theta_2) = \Lambda_{x'}(\theta_1, \theta_2)$ for all θ_1, θ_2

in other words: $L(\theta, x) = L(x, x') L(\theta, x')$

⇒ Definition: Minimal sufficient statistics (MSS)

A sufficient statistics $T(x)$ is minimal sufficient statistics if $T(x)$ is a function of any other sufficient statistics, i.e. For any sufficient statistic S, where exist a g such that $T(x) = g(S)$

⇒ Remark:

(1) Minimal sufficient statistic is not unique

example: $X_1, X_2, \cdots X_n | \theta \sim \text{poisson}(\theta)$

$T_1(x) = \Sigma X_i$ is minimal sufficient

$S(x) = 2\Sigma X_i$ is minimal sufficient

i.e $\exists$ one-to one g such that $T_1 = g(T_2)$

⇒ Lemma 6.2:

if T and S are two minimal sufficient statistics, then $\exists$ injective function g_1 and g_2

such that $T = g_1(S)$ $S = g_2(T)$

$\Rightarrow$ theorem 6.1:

$T(x)$ is minimal sufficient statistics

$\Leftrightarrow$ For any pair X and X', $T(x) = T(x')$

$\Leftrightarrow \Lambda_x(\theta_1, \theta_2) = \Lambda_{x'}(\theta_1, \theta_2)$ for any θ_1 and θ_2

$\Rightarrow$ example 6.1.

$X_1, \cdots X_n \mid \mu, \sigma^2 \sim N(\mu, \sigma^2)$

$$f(x; \mu, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{\Sigma x_i^2}{2\sigma^2} + \frac{\mu}{\sigma^2}\Sigma x_i - \frac{n\mu^2}{2\sigma^2}}$$

$T(x) = (\Sigma x_i^2, \Sigma x_i)$

$$\Lambda_x(\theta_1, \theta_2) = \frac{C(\theta_1)}{C(\theta_2)} \cdot e^{-\frac{\Sigma x_i^2}{2}\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right) + \Sigma x_i\left(\frac{\mu_1}{\sigma_1^2} - \frac{\mu_2}{\sigma_2^2}\right)}$$

For any pair of data set x and x'

if $T(x) = T(x')$

the $\Lambda_x(\theta_1, \theta_2) = \Lambda_{x'}(\theta_1, \theta_2)$ for all θ_1 and θ_2

$\Rightarrow$ T is sufficient.

$\Rightarrow$ proof theorem 6.1

For any pair of data sets X and X'

If $\Lambda_x(\theta_1, \theta_2) = \Lambda_{x'}(\theta_1, \theta_2)$ for some θ_1, θ_2

then we want to say $T(x) = T(x')$

This is equivelent to prove that

if $T(x) \neq T(x')$ then $\lambda_x(\theta_1, \theta_2) \neq \lambda_{x'}(\theta_1, \theta_2)$

for some θ_1 and θ_2.

Minimal : if $L(\theta, x') = c\,L(\theta, x)$, then
$T(x) = T(x')$

$\Leftrightarrow$ if $T(x) \neq T(x')$ then $L(\theta, x') \neq c\,L(\theta, x)$

$\Rightarrow$ Example 6.2:

If $f(x_i \mid \theta)$ is in exp family i.e

$f(x \mid \theta) = c(\theta)\, h(x) \cdot e^{\sum \pi_i(\theta) T_i(x)}$

and Θ contains an open rectangle.

Then $T(x) = (T_1(x), \dots, T_k(x))$ is
minimal suffi statistic Indeed it is

complete too.

$\Rightarrow$ HW:

$X_1 \cdots X_n \mid \theta \sim$ poisson (θ)

to show. $T\left(\sum_{i=1}^{5} x_i, \sum_{i=6}^{n} x_i\right)$ is not

minimal sufficient by using the therom 6.1

3) $\tau^2 \mid \theta, \mu_1, \cdots, \mu_p, x_1, \cdots, x_p$

$$f(\tau^2 \mid \cdots\cdots) \propto \prod_{i=1}^{p} f(\mu_i \mid \tau^2, \theta) \cdot \pi_{\tau^2}(\tau^2)$$

$$= e^{\sum_{i=1}^{p} -\frac{(\mu_i-\theta)^2}{2\tau^2}} \cdot (\tau^2)^{-\frac{p}{2}} \cdot (\tau^2)^{-(\alpha^*+1)} \cdot e^{-\frac{\beta^*}{\tau^2}}$$

$$= e^{-\frac{1}{\tau^2}\left(\underline{\sum_{i=1}^{p} \frac{(\mu_i-\theta)^2}{2} + \beta^*}\right)} \cdot (\tau^2)^{-(\underline{\alpha^* + \frac{p}{2}} + 1)}$$

$$\Rightarrow \tau^2 \mid \cdots \sim \text{inv-Gamma}\left(\frac{2\alpha^*+p}{2}, \frac{\sum_{i=1}^{p}(\mu_i-\theta)^2 + 2\beta^*}{2}\right)$$

Let $\alpha^* = \frac{K^*}{2}$, $\beta^* = \frac{\lambda^*}{2}$

$$\Rightarrow \tau^2 \mid \cdots \sim \text{inv-Gamma}\left(\frac{K^*+p}{2}, \frac{\lambda^* + \sum_{i=1}^{p}(\mu_i-\theta)^2}{2}\right)$$

$$E(\tau^2 \mid \cdots) = \frac{\lambda^* + \sum_{i=1}^{p}(\mu_i-\theta)^2}{K^* + p}$$

$\Rightarrow$ Gibbs sampling

$(\theta_1, \cdots, \theta_p) \sim \pi(\theta_1, \cdots, \theta_p)$

starting from $\theta^{(0)} = (\theta_1^{(0)}, \cdots, \theta_p^{(0)})$

reapeat N times of :

(1). Draw $\theta_1^{(1)} \sim \pi(\theta_1 \mid \theta_2^{(0)}, \cdots \theta_p^{(0)})$

(2). Draw $\theta_2^{(1)} \sim \pi(\theta_2 \mid \theta_1^{(1)}, \theta_3^{(0)}, \cdots, \theta_p^{(0)})$

$\vdots$

(K). Draw $\theta_K^{(1)} \sim \pi(\theta_K \mid \theta_1^{(1)}, \theta_2^{(1)} \cdots \theta_{K-1}^{(1)}\ \theta_{K+1}^{(0)} \cdots \theta_p^{(0)})$

$\vdots$

(p). Draw $\theta_p^{(1)} \sim \pi(\theta_p \mid \theta_1^{(1)}, \theta_2^{(1)} \cdots\cdots \theta_{p-1}^{(1)})$

set $\theta^{(0)} = \theta^{(1)}$

informally, Repeat N times for $i = 1, \cdots, p$

$\theta_i \sim \pi(\theta_i \mid \theta_{-i})$

<< Mar 13, 2015 >> STAT 846

=> Mini sufficient

A statistic T is said to be MSS $\overset{iff}{\Leftrightarrow}$

For any pair of data sets x, x'

(1) $T(x) = T(x') \Rightarrow \Lambda_x(\theta_1, \theta_2) = \Lambda_{x'}(\theta_1, \theta_2)$
for all θ_1, θ_2.

(2) $T(x) \neq T'(x) \Rightarrow \Lambda_x(\theta_1, \theta_2) \neq \Lambda_x(\theta_1, \theta_2)$
for some θ_1, θ_2.

=> Complete statistic.

A Statistic T is said to be complete

for any real-value function g, if

$E(g(T) \mid \theta) = 0$ for all θ, then $Pr(g(T) = 0 \mid \theta) = 1$

=> what completeness implies?

if T is complete, then there exists only

one unbiased estmator for θ, Let

$E(g_1(t) \mid \theta) = 0, \quad \forall \theta.$

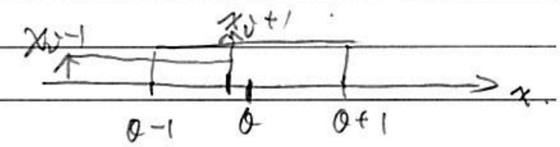
$E(g_2(t) \mid \theta) = 0, \quad \forall \theta$

$E(g_1(T) - g_2(T) \mid \theta) = 0$

if T is not complete for θ unbiased estimator for θ is not unique.

If T is complete, then unbiased estimator is unique.

=> example: (not complete)



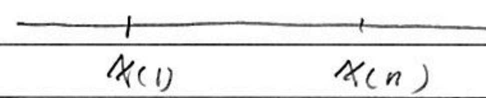
$X_1, \dots X_n | \theta \overset{iid}{\sim} unif(\theta-1, \theta+1)$

$f(x) = \prod_{i=1}^{n} I(\theta-1 < x_i < \theta+1)$

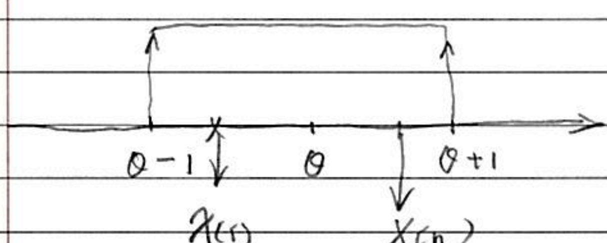
$= \prod_{i=1}^{n} I(x_i - 1 < \theta < x_i + 1)$

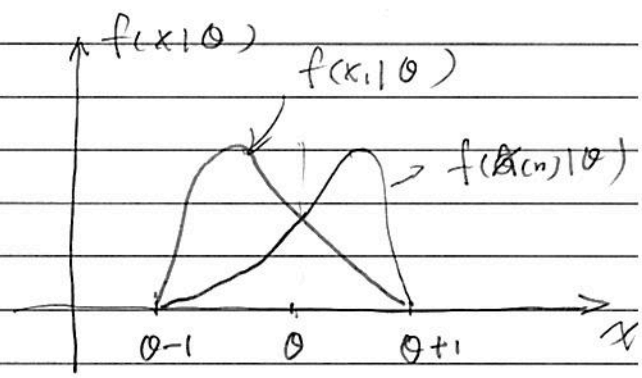
$= I(\theta \in (x_i - 1, x_i + 1))$ for all $i=1, 2, \dots n$)

$= I(\theta \in (X_{(n)} - 1, x_{(1)} + 1))$



$T(X_{(1)}, T_{(n)})$
is minimal sufficient statistic.
but we will see
T is not complete



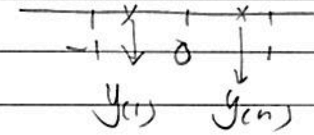


$f(x_{(1)} | \theta)$ is symetric about θ with $f(x_{(n)}, \theta)$

$$E\left(\frac{X_{(1)} + X_{(n)}}{2}\right) = \theta$$

The distribution of $X_{(n)} - X_{(1)}$ is independent of θ,

Let $Y_i = X_i - \theta \sim unif(-1, 1)$



$$Y_{(n)} = X_{(n)} - \theta$$
$$Y_{(1)} = X_{(1)} - \theta$$

$X_{(n)} - X_{(1)} = Y_{(n)} - Y_{(1)}$ the distribution of $Y_{(n)} - Y_{(1)}$ doesn't depend on θ so the distribution of $X_{(n)} - X_{(1)}$ doesn't depend on θ too!

$E(X_{(n)} - X_{(1)} | \theta) = c$ for all θ, c doesn't depend on θ.

Now we have:

$$E\left(\frac{X_{(1)} + X_{(n)}}{2} \Big| \theta\right) = \theta \text{ for all } \theta$$

$$E\left(\frac{X_{(1)} + X_{(n)}}{2} - ((X_{(n)} - X_{(1)}) - c) \Big| \theta\right) = \theta$$

for all

There are two unbiased estimator which are function $f(X_{(1)}, X_{(n)})$, for θ.

⇒ Lemma 6.3

If $T = (T_1, \cdots, T_K)$ is the natural statistical of an exp family also Θ contains an open rectangle, then, T is complete.

proof: the pdf / pmf of $(T_1, \cdots, T_K)$ has the form as follows

$$f(t_1, \cdots, t_K | \theta) = C(\theta)\, h(t) \cdot e^{\sum_{i=1}^{K} \theta_i t_i}$$

suppose for a real-value function g such that

$$E(g(T_1, \cdots, T_K) | \theta) = 0 \quad \text{for all } \theta$$

we will show $g(t_1, \cdots, t_K) = 0$ a.s.

Let $t = (t_1, \cdots, t_K)$

$$E(g(T_1, \cdots, T_K) | \theta)$$

$$= \int g(t)\, C(\theta)\, h(t)\, e^{\sum_{i=1}^{K} \theta_i t_i}\, dt$$

$$= C(\theta) \int g(t) \cdot e^{\sum_{i=1}^{K} \theta_i t_i}\, dt$$

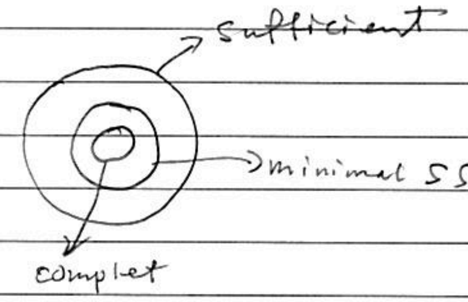
This is the Laplace transformal of $g(t)\, h(t)$, so integral $= 0 \Rightarrow g(t)\, h(t) = 0$

//

<< Mar 16, 2015 >> STAT 846

⇒ Lehman - Scheffe

theorem: if T is sufficient and complete, then T is minimal sufficient statistic (MSS)



⇒ UMVUE: we says a statistic $T(x)$ is an uniformly minimum variance unbiased estimator for θ if

1) $E(T(x) \mid \theta) = \theta$ for all θ

2) $V(T(x) \mid \theta) \le V(d(x) \mid \theta)$, for all θ, ; for all $d(x)$

with $E(d(x) \mid \theta) = \theta$

$V(T(x) \mid \theta) = E((T(x) - \theta)^2 \mid \theta)$

because $E(T \mid \theta) = \theta$

⇒ theorem 6.3 (Rao - Blackwell Theorem)

Given

$E(d_1(x) \mid \theta) = \theta$

T is a sufficient statistics

$g(T) = E(d_1(x) \mid T, \theta)$

then,

1) $g(T)$ is a statistic i.e free of θ

2) $E(g(T)|\theta) = \theta$

3) $Var(g(T)|\theta) \leq Var(d(x)|\theta)$

proof:

1) T is sufficient for θ

$X|T$ is independent of θ

so $E(d_1(x)|T, \theta)$ is free of θ

2) $E(g(T)|\theta) = E_T(E_X(d_1(x)|T,\theta))$

$$= E_X(d_1(x)|T)$$
$$= \theta$$

3)

$$E_X[(d_1(x)-\theta)^2]$$

$$= E_T E_X((d_1(x)-\theta)^2|T)$$

Note that:

$$E_X((d_1(x)-\theta)^2|T) = (g(T)-\theta)^2$$

$$\geq (E(d_1(x)|T)-\theta)^2 \text{ by Jensen Inequality}$$

Then, we see that

$$E_X[(d_1(x)-\theta)^2] \geq E_T[(g(T)-\theta)^2]$$

Remarks:

1) if T is complete; $g(T)$ is unique

So, the $g(T)$ is UMVUE

T is complete

$d_1(x)$, $d_2(x)$, ..., $d_n(x)$

$\Downarrow$ $\Downarrow$ $\Downarrow$

$g_1(T) = E(d_1(x)|T)$ $g_2(T) = E(d_2(x)|T)$ .... $g_n(T) = E(d_n(x)|T)$

★ ⇒ Methods for finding UMVUE for θ.

1) Find the complete T for θ

2) Find an estimator statistic $d_1(x)$ such that

$E(d_1(x)|\theta) = \theta$

3) Find $E(d_1(x)|T) \equiv g(T)$

$g(T)$ is the UMVUE.

⇒ example: $X_1, \dots X_n \overset{iid}{\sim}$ poisson(λ)

Find the UMVUE

Review: $X \sim$ poisson(λ), $Y \sim$ poisson(μ)

$X \perp Y$

$X+Y \sim$ poisson$(\lambda+\mu)$

For $k \geq x$

$$P(X=x \mid X+Y=k) = \frac{P(X=\cancel{k}x,\ Y=k-x)}{P(\cancel{X+Y=k})}$$

$$= \frac{P(X=\cancel{k}x)\cdot P(Y=k-x)}{P(X+Y=k)}$$

$$= \frac{e^{-\lambda}\lambda^x}{x!}\cdot\frac{e^{-\mu}\cdot\mu^{k-x}}{(k-x)!} \Big/ \frac{e^{-(\lambda+\mu)}\cdot(\lambda+\mu)}{k!}$$

$$= \binom{k}{x}\left(\frac{\lambda}{\lambda+\mu}\right)^x\left(1-\frac{\lambda}{\lambda+\mu}\right)^{k-x}$$

$\Rightarrow \ x \mid x+y = k \sim Bin\left(k,\ \frac{\lambda}{\lambda+\mu}\right)$

Solution:

(1) $T = \Sigma X_i$ is complete statistic, because poisson (λ) is in exp family

$E(X_1) = \lambda$, Let $d_1(x) = X_1$, the UMVUE for λ is $g(T) = E(X_1 \mid T)$

$X_1 \sim$ poisso(λ) ; $X_2 + \cdots + X_n \sim$ poisso($(n-1)\lambda$)

$X_1 \perp X_2 + \cdots + X_n$

$X_1 \mid X_1 + \cdots + X_n = K \sim Bin\left(K\ \ \frac{\lambda}{n\lambda}\right)$

$X_1 \mid T \sim Bin($

$= Bin\left(K,\ \frac{1}{n}\right)$

in other words

$$X_1 \mid T \sim Bin\left(T,\ \frac{1}{n}\right)$$

$$E(X_1 \mid T) = \frac{T}{n} = \bar{X}$$

(2) Find an UMVUE for λ^2

$T = \Sigma X_i$ is complete for λ^2

$Var(X_1) = \lambda = E(X_1^2) - (E(X_1))^2$

$\Rightarrow E(X_1^2 - X_1) = \lambda^2$

We will find $g(T) = E(X_1^2 - X_1 \mid T)$

$V(X_1 \mid T) + (E(X_1 \mid T))^2 = E(X_1^2 \mid T) - E(X_1 \mid T)$

$= \left[T \cdot \frac{1}{n}\left(1 - \frac{1}{n}\right) + \left(T \cdot \frac{1}{n}\right)^2\right] - \frac{T}{n}$