Stat 543  1-21-05

Recall:

Thm (FNHS Factorization)

\[ T(X) \text{ sufficient } \iff f(x|\theta) = g(T(x), \theta) h(x) \]

"the shape of the loglikelihood depends on x only through \( T(x) \)"
$T(x)$ sufficient $\iff L(\theta) = g(T(x), \theta) h(x) \ln L(\theta) = \ln g(T(x), \theta) + \ln h(x)$

$\ln L(\theta)$

\[ \text{suppose } x, x', x'' \]
\[ T(x) = T(x') = T(x'') \]

for $x$

for $x'$

for $x''$
Example \( X_1, \ldots, X_n \) iid Poisson (\( \lambda \))

\( \lambda \geq 0 \)

\[
f(x | \lambda) = \begin{cases} \frac{e^{-\lambda} \lambda^{x_i}}{x_i!} & \text{all } x_i \text{ are nonnegative integers} \\ 0 & \text{otherwise} \end{cases}
\]

\[
= \frac{e^{-\lambda} \sum x_i}{\prod x_i!} \quad I\left[ \text{all } x_i \text{ are nonnegative integers} \right]
\]
Let \( g(t, \lambda) = \lambda^t \ \cdot e^{-\lambda t} \)

\[
h(x) = \frac{1}{\prod x_i!} \left[ \text{all } x_i \text{ are nonnegative integers} \right]
\]

\[
T(x) = \sum x_i
\]

We can then write

\[
f(x|\lambda) = g(T(x), \lambda) \cdot h(x)
\]

So \( T(X) = \sum X_i \) is sufficient for \( \lambda \).
Example \( X_1, X_2, \ldots, X_n \) iid Beta \((\alpha, \beta)\)

\[
f(x \mid \alpha, \beta) = \left( \frac{1}{\text{B}(\alpha, \beta)} \right)^n \prod_{i=1}^{n} \frac{x_i^{\alpha-1} (1-x_i)^{\beta-1}}{\text{B}(\alpha, \beta)} I[\text{all } x_i \text{ are in } (0,1)]
\]

Let \( g(t_1, t_2, \alpha, \beta) = \left( \frac{1}{\text{B}(\alpha, \beta)} \right)^n t_1^{\alpha-1} t_2^{\beta-1} \)

\( h(x) = I[\text{all } x_i \text{ are in } (0,1)] \)

\( T(x) = (\prod x_i, \prod (1-x_i)) \)
you can check that
\[ f(x | \alpha, \beta) = g(T(x), \alpha, \beta) h(x) \]
so that by FNHS
\[ T(x) = (\prod X_i, \prod (1-X_i)) \]
is sufficient for \((\alpha, \beta)\)

Proof of FNHS in discrete case:

Suppose that \(T(X)\) is sufficient for \(\Theta\) and let \(\Theta_0 \in \Theta\)
For any \(t\) and \(\Theta\) with
\[ P_\theta \left[ T(X) = t \right] > 0 \]

Sufficiency says that
\[ P_\theta \left[ X = x \mid T(X) = t \right] = P_\theta \left[ X = x \mid T(X) = t \right] \]

So
\[ f(x \mid \theta) = P_\theta \left[ X = x \right] = P_\theta \left[ X = x \text{ and } T(X) = T(x) \right] \]
\[ = P_\theta \left[ X = x \mid T(X) = T(x) \right] P_\theta \left[ T(X) = T(x) \right] \]

where
\[ g(t, \theta) = P_\theta \left[ T(X) = t \right] \]
\[ \frac{1}{n(x)} \]
On the other hand, suppose $g(t, \theta)$ and $h(x)$ so that

$$f(x|\theta) = g(T(x), \theta) \cdot h(x)$$

Then

$$P_0 \left[ T(X) = t \right] = \sum_{x \text{ s.t. } T(x) = t} g(T(x), \theta) \cdot h(x)$$

$$= g(t, \theta) \sum_{x \text{ s.t. } T(x) = t} h(x)$$
So for any $t$ with $P_\theta[T(x) = t] > 0$

$$P_\theta \left[ X = x \mid T(x) = t \right] = \frac{P_\theta \left[ X = x \text{ and } T(x) = t \right]}{P_\theta \left[ T(x) = t \right]}$$

$$= \mathbb{I} \left[ T(x) = t \right] \frac{g(t, \theta) h(x)}{g(t, \theta) \sum_{x \leq t, \frac{T(x)}{t} = t} h(x)}$$

$$= \mathbb{I} \left[ T(x) = t \right] \frac{h(x)}{\sum_{x \leq t, \frac{T(x)}{t} = t} h(x)}$$
and this doesn't involve $\theta$, i.e. $T(X)$ is sufficient for $\theta$.

Example: $X_1, \ldots, X_n$ iid $N(\mu, \sigma^2)$

$\mu \in \mathbb{R}, \sigma^2 > 0$

$T(X) = (\sum X_i, \sum X_i^2) = (T_1(X), T_2(X))$

"equivalent to" $\bar{X}, S^2$
\[ f(x \mid \mu, \sigma^2) = \left( \frac{1}{2\pi \sigma^2} \right)^{N/2} \exp \left[ -\frac{1}{2\sigma^2} \left( \sum (x_i - \mu)^2 \right) \right] \]

\[ = \left( \frac{1}{2\pi \sigma^2} \right)^{N/2} \exp \left[ -\frac{1}{2\sigma^2} \left( \sum x_i^2 - 2\mu \sum x_i + n\mu^2 \right) \right] \]

So with

\[ g(t, \mu, \sigma^2) = \left( \frac{1}{2\pi \sigma^2} \right)^{N/2} \exp \left[ -\frac{1}{2\sigma^2} \left( t_i - 2\mu t_2 + \sigma^2 \right) \right] \]

\[ h(x) = 1 \]

I apply Factorization Thm and \( T(X) \) is sufficient for \((\mu, \sigma^2)\).