Stat 543 1-24-05

Recall

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

$T(X) = \left( \sum X_i, \sum X_i^2 \right)$ is sufficient for

of course we're more used to thinking in terms $\bar{X}, S$ as appropriate in this problem – clearly

$\bar{X} = \frac{1}{n}\left(\sum X_i\right) \quad S = \sqrt{\frac{1}{n-1}\left(\sum X_i^2 - \frac{(\sum X_i)^2}{n}\right)}$

and $\sum x_i = n\overline{X}$ $\quad \sum x_i^2 = (n-1)s^2 + \frac{(n\overline{X})^2}{n}$

<u>Def</u> Two statistics $T(X)$ and $S(X)$ are equivalent if $\exists$ functions $p(t)$ and $q(s)$ s.t.

$$T(x) = q(S(x)) \text{ and } S(x) = p(T(x))$$

<u>Thm</u> If $T(X)$ and $S(X)$ are equivalent, $T(X)$ is sufficient iff $S(X)$ is sufficient for $\theta$

There is a Bayesian notion of sufficiency - for a Bayesian it is the posterior that is of interest, so presumably we have

Def For a prior dsn $G$, a statistic $T(X)$ is Bayes sufficient for $\theta$ if the posterior dsn for $\theta \mid X = x$ depends on $x$ only through $T(x)$

(2 $x$'s with the same value of $T$ have the same posterior)

Pretty generally

$T(X)$ sufficient for $\theta$ $\Longrightarrow$ $T(X)$ Bayes sufficient for any prior $G$

Why? $$g(\theta|x) \propto \underbrace{L(\theta)}_{g(T(x),\theta)h(x)} g(\theta)$$

---

$T(X)$ Bayes sufficient for all $G$

or

$T(X)$ Bayes sufficient for $G$ with $g(\theta) > 0 \ \forall \theta$

$\Longrightarrow$ $T(X)$ is sufficient for $\theta$

**Minimal Sufficiency** Idea is to reduce $X$ as far as one can w/o losing sufficiency

**Def** A statistic $T(X)$ is called minimal sufficient provided it is sufficient for $\theta$ and for any other sufficient statistic $S(X)$ $\exists$ a function $q(s)$ on the range of $S(X)$ such that

$$T(x) = q(S(x))$$

How to recognize a minimal sufficient statistic?

Thm Suppose $f(x|\theta)$ $\forall$ $\theta \in \Theta$ is either a pdf for $X$ on $\mathbb{R}^k$ or a pmf for $X$ and $T(X)$ is sufficient for $\theta$. If

$$\left( \begin{array}{l} \text{the existence of} \\ \text{a number } k(x,y) > 0 \\ \text{such that} \\ f(y|\theta) = f(x|\theta)\, k(x,y) \\ \qquad \forall\, \theta \end{array} \right) \Longrightarrow T(y) = T(x)$$

Then $T(X)$ is minimal sufficient

Note:

$$\Big( \qquad \Big) \approx$$ the existence of a number $k(x,y) > 0$ s.t.

$$\underbrace{\ln L_y(\theta) = \ln L_x(\theta) + \ln k(x,y)}_{\text{the loglikelihoods for } y \text{ and } x \text{ have the same shape}}$$

roughly: $\frac{f(y|\theta)}{f(x|\theta)}$ is free of $\theta$

So roughly the Theorem says

$T(X)$ sufficient and $\left(\begin{array}{l}\text{loglikelihoods for } x, y \\ \text{with same shape} \Rightarrow T(x) = T(y)\end{array}\right)$

$\Longrightarrow T(X)$ is minimal sufficient

Example $X_1, X_2, \ldots, X_n$ iid with marginal pmf

$$f(x|\theta) = \begin{cases} c(\theta)\dfrac{\theta^x}{x!} & x = 0, 1, 2, \ldots, \lfloor\theta\rfloor \\ 0 & \text{otherwise} \end{cases}$$

$$c(\theta) = \left(\sum_{x=0}^{\lfloor\theta\rfloor} \frac{\theta^x}{x!}\right)^{-1}$$

iid truncated Poisson r.v.'s

Joint pmf on $\{0,1,2,\ldots\}^n$

$$f(x|\theta) = c^n(\theta)\,\frac{\theta^{\sum x_i}}{\pi x_i!}\, I[\max x_i \le \theta]$$

The factorization thm then says that

$$T(X) = (\max X_i, \Sigma X_i)$$

is sufficient - use

$$g(t,\theta) = c^n(\theta)\,\theta^{t_2}\, I[t_1 \le \theta]$$

$$h(x) = \frac{1}{\pi x_i!}$$

Then suppose $\exists$ $k(x,y) > 0$ s.t. as functions of $\theta$

$$f(y|\theta) = f(x|\theta)\, k(x,y)$$

$f(x|\theta)$ is 0 for $\theta < \max x_i$ and positive for $\theta \geq \max x_i$ (same for rhs) — something similar for $f(y|\theta)$ — equality then forces $\max x_i = \max y_i$

$$\cancel{c^n(\theta)} \frac{\theta^{\sum x_i}}{\Pi x_i!} = k(y,x)\, \cancel{c^n(\theta)} \frac{\theta^{\sum y_i}}{\Pi y_i!}$$

as a function of $\theta$ requires

$$\frac{\theta^{\Sigma x_i}}{\Pi x_i!} - k(y,x)\frac{\theta^{\Sigma y_i}}{\Pi y_i!} \equiv 0$$

This is a polynomial in $\theta$ that is $0$ on an interval ... that requires

$$\Sigma x_i = \Sigma y_i$$

(and $\frac{1}{\Pi x_i!} = k(y,x)\frac{1}{\Pi y_i!}$ )

And thus requires $T(x) = T(y)$ and $T(X)$ is minimal sufficient