

Stat 543 1-24-05

Recall

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

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

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

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

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

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

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

Thm If $T(X)$ and $S(X)$ are equivalent, $T(X)$ is sufficient iff $S(X)$ is sufficient for Θ

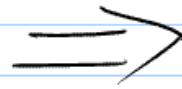
There is a Bayesian notion of sufficiency - for a Bayesian it is the posterior that is of interest, so presumably we have

Def For a prior π on Θ , a statistic $T(X)$ is Bayes sufficient for Θ if the posterior π for $\Theta | X=x$ 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 θ



$T(X)$ Bayes
sufficient for
any prior G

Why?

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

$T(X)$ Bayes
sufficient for all G



$T(X)$ is
sufficient for θ

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

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 θ and for any other sufficient statistic $S(X)$ \exists a function $g(s)$ on the range of $S(X)$ such that

$$T(x) = g(S(x))$$

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 θ . If

the existence of a number $k(x,y) > 0$ such that

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

$\Rightarrow T(y) = T(x)$

Then $T(X)$ is minimal sufficient

Note:

()

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

\approx
 \approx

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

$$\ln L_y(\theta) = \ln L_x(\theta) + \ln k(x,y)$$

the loglikelihoods for
 y and x have the same
shape

So roughly The Theorem says

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

$\Rightarrow T(X)$ is minimal sufficient

Example X_1, X_2, \dots, X_n iid with marginal pmf

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

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

iid truncated Poisson r.v.'s

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

$$f(x|\theta) = c^n(\theta) \frac{\theta^{\sum x_i}}{\prod x_i!} \mathbb{I}[\max x_i \leq \theta]$$

The factorization thm then says that

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

is sufficient - use

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

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

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

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

$f(x|\theta)$ 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 $\max x_i = \max y_i$

$$C^n(\theta) \frac{\theta^{\sum x_i}}{\pi x_i!} = k(y, x) C^n(\theta) \frac{\theta^{\sum y_i}}{\pi y_i!}$$

as a function of θ requires

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

This is a polynomial in θ that is 0 on an interval ... that requires

$$\left(\text{and } \frac{1}{\prod x_i!} = k(y, x) \frac{1}{\prod y_i!} \right)$$

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