<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 for θ iff $S(X)$ is sufficient for θ

There is a Bayesian notion of sufficiency - of course (since, for a Bayesian it is the posterior dsn that is of interest) it must be phrased in terms of the posterior

<u>Def</u> For a prior G, a statistic $T(X)$ is <u>Bayes sufficient</u> for θ if the posterior for θ (conditional of $\theta | X = x$) depends upon x only through $T(x)$ (any two x's with the same value of $T(x)$ have the same posterior)

pretty generally

$$\begin{matrix} T(X) \text{ sufficient} \\ \text{for } \theta \end{matrix} \implies \begin{matrix} T(X) \text{ Bayes} \\ \text{sufficient for} \\ \text{prior } G \end{matrix}$$

Why?

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

on the other hand

$T(X)$ Bayes sufficient for all G

or

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

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

Y&S have a nice / helpful theorem giving an equivalent condition for the factorization thm in their Section 6.1

<u>Lemma 6.1</u> For a function $T(x)$ the following are equivalent

factorization criterion $\Rightarrow$ i) $\exists$ functions $h(x)$ and $g(t,\theta)$ such that

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

the likelihood ratio criterion $\Rightarrow$ ii) For any pair x, x' with $T(x) = T(x')$

$$\frac{f(x|\theta_1)}{f(x|\theta_2)} = \frac{f(x'|\theta_1)}{f(x'|\theta_2)} \quad \forall\, \theta_1, \theta_2$$

So one immediately has

sufficiency $\Longleftrightarrow$ l.r. criterion

Example 2-class model $\Theta = \{0,1\}$ and $f(x|\theta)$

$$\Lambda(x) = \frac{f(x|1)}{f(x|0)}$$

is a sufficient statistic

simple discrete example with $f(x|\theta)$ =

θ \ x	0	1	2	3	4
1	.2	.2	.4	.1	.1
0	.1	.35	.2	.3	.05

$\Lambda(x)$ is:

x	0	1	2	3	4
$\Lambda(x)$	2	4/7	2	1/3	2

$\Lambda(X)$ is sufficient, as is anything equivalent to it, like e.g.

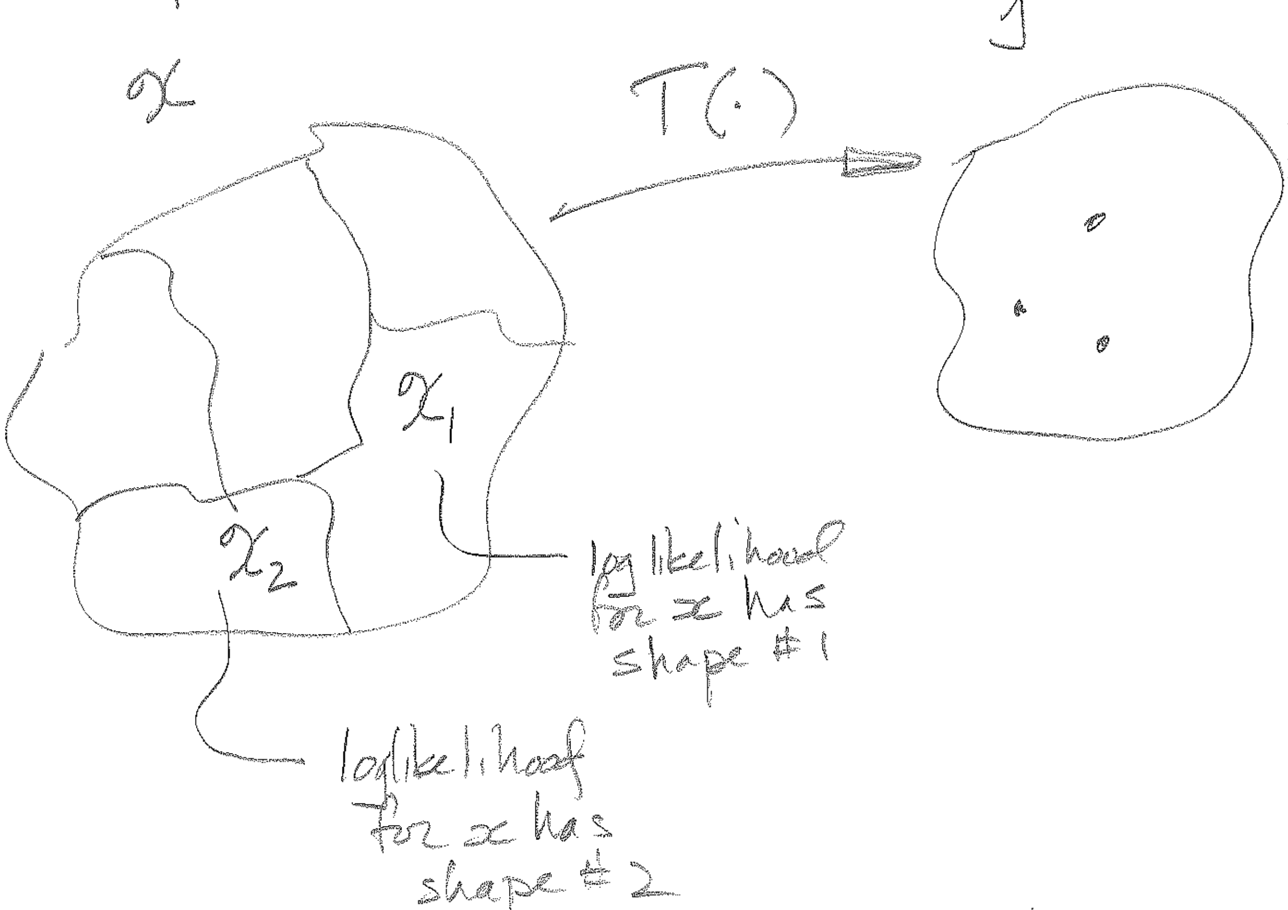
$$T(0) = T(2) = T(4) = \sqrt{\pi}$$
$$T(1) = 17 \quad \text{and} \quad T(3) = -700$$

Minimal Sufficiency

The idea here is to reduce X as far as possible 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, $\exists$ a function $q(S)$ on

The big picture regarding minimal sufficiency (that has to be quantified + qualified, but nevertheless is the fundamental story):



etc. Then T is <u>minimal sufficient iff</u>

$$T(x) = t_1 \quad \forall\ x \in \mathcal{X}_1$$
$$T(x) = t_2 \quad \forall\ x \in \mathcal{X}_2$$

etc.

where the $t_1, t_2 \cdots$ are all different

The range of S such that

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

What can be said about locating a minimal sufficient statistic and/or proving that a sufficient statistic is minimal sufficient?

day 7

<u>Thm</u> Suppose $f(x|\theta)$ $\forall\ \theta \in \Theta$ is either a pdf for X on R^k or a pmf and $T(X)$ is a sufficient statistic. <u>If</u>

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

Then $T(X)$ is minimal sufficient.

<u>Note</u>:

$$\left(\quad \right) \approx \underbrace{\begin{array}{l} \text{The existence of a} \\ \text{number } k(x,y) > 0 \text{ s.t.} \\ \ln L_y(\theta) = \ln L_x(\theta) + \ln k(x,y) \\ \quad \forall\ \theta \end{array}}$$

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

The log likelihoods for y and x have the same shape

we don't quite want to say it these ways because we don't want to worry about division by 0 or $\ln(0)$ cases — but roughly the Thm says

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

$\Longrightarrow$ $T(X)$ is minimal sufficient

Pf: See "handout" on web page ∎

Make some applications to finite Θ (K-class) models:

Application 1 $\Theta = \{0,1\}$ (2-class problem) —

We've already said that $\dfrac{f(x|1)}{f(x|0)} = \ell(x)$

sufficient since the likelihood ratio and factorization criteria are equivalent — Then if for $x, y \; \exists$ $k(x,y) > 0$ s.t. $f(x|1) = f(y|1)\, k(x,y)$ and $f(x|0) = f(y|0)\, k(x,y)$

then

$$\frac{f(x|1)}{f(x|0)} = \frac{f(y|1)}{f(y|0)} \frac{k(x,y)}{k(x,y)}$$

i.e. $\ell(x) = \ell(y)$ and $\ell(X)$ is minimal sufficient

Application 2 $\Theta = \{1, 2, \ldots, K\}$ (K-class problem) and if $f(x|1) > 0 \;\forall x$, Then with

$$\ell(x) = \left(\frac{f(x|2)}{f(x|1)}, \frac{f(x|3)}{f(x|1)}, \ldots, \frac{f(x|K)}{f(x|1)} \right)$$

the same argument will show $\ell(X)$ to be minimal sufficient

Application 3 for $\Theta = \{1, 2, \ldots, K\}$ (K-class problem) completely generally, with

$$s(x) = \sum_{\theta=1}^{K} f(x|\theta) > 0 \quad \forall x$$

(if not throw x out of the sample space, since it has 0 probability under any θ) the same argument shows that with

the posterior based on a uniform prior

$$\ell(x) = \left(\frac{f(x|1)}{s(x)}, \frac{f(x|2)}{s(x)}, \ldots, \frac{f(x|K)}{s(x)} \right)$$

The K-dimensional $\ell(X)$ is minimal sufficient

(classification problems with even high-dimensional observables $\underline{x}$ have low-dimensional minimal sufficient statistics that are sort of likelihood ratios and in fact are posteriors)

Example $K=3$ and pmf's $f(x|\theta)$ below

θ \ x	1	2	3	4	5	6
3	.1	.2	.1	.2	.2	.2
2	.2	.15	.05	.1	.1	.4
1	.1	.2	.2	.2	.2	.1

x	1	2	3	4	5	6
$\frac{f(x\|2)}{f(x\|1)}$	2	$\frac{3}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	4
$\frac{f(x\|3)}{f(x\|1)}$	1	1	$\frac{1}{2}$	1	1	2

So $T(x)$ with $T(4) = T(5) = 17$

$T(1) = 32 \quad T(2) = 63 \quad T(3) = \pi \quad T(6) = 12$

is minimal sufficient — note that altering it by

1) setting $T(1) = T(2) = 32$ destroys sufficiency

2) instead setting $T(4) = -60$ and $T(5) = 17$ maintains sufficiency (it's now equivalent to X itself, which is always sufficient) but destroys minimality

---

<u>Example</u> (standard) $X_1, X_2, \ldots, X_n$ iid Beta(α, β)

we've seen that $T(X) = (\prod X_i, \prod(1-X_i))$ is sufficient – The joint pdf on $[0,1]^n$ is

$$f(x|\alpha,\beta) = \frac{1}{(B(\alpha,\beta))^n}\left(\prod x_i\right)^{\alpha-1}\left(\prod(1-x_i)\right)^{\beta-1}$$

$f(y|\alpha,\beta) = k(y,x) f(x|\alpha,\beta)$ as a function of $(\alpha,\beta) \in (0,\infty)^2$ requires that

$$\left(\prod y_i\right)^{\alpha-1}\left(\prod(1-y_i)\right)^{\beta-1} = k(x,y)\left(\prod x_i\right)^{\alpha-1}\left(\prod(1-x_i)\right)^{\beta-1}$$

Set $\beta = 1$ and note that this requires that

$$\left(\prod y_i\right)^{\alpha} = k'\left(\prod x_i\right)^{\alpha}$$

i.e. $$\left(\frac{\prod y_i}{\prod x_i}\right)^{\alpha} = k' \text{ as a function of } \alpha$$

which requires $\prod y_i = \prod x_i$

similarly setting $\alpha = 1$ requires $\prod(1-y_i) = \prod(1-x_i)$ and T is minimal sufficient

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

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

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

iid truncated Poisson r.v.'s - joint pmf on $\{0,1,2,\ldots\}^n$ is

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

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 θ

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

$f(x|\theta)$ is 0 for $\theta < \max x_i$ and positive for $\theta \ge \max x_i$ and the same is true for $f(y|\theta)$ - That forces $\max x_i = \max y_i$

Then for $x, \underline{y} \in \{0,1,2,\ldots\}^n$ and

$\theta \ge \max x_i = \max y_i$

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

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

This is a polynomial in θ identically 0 on an interval – since the coefficients are non-zero, this can only happen if $\sum x_i = \sum y_i$ i.e. we have the same powers of θ in both $f(x|\theta)$ and $f(y|\theta)$ i.e.

$$T(x) = T(y)$$

and T is minimal sufficient