Stat 543 1-28-05

Correction to last example:

$$R(x,1,1) = \exp\left(\sum x_i - \frac{n}{2}\right)$$

$$R(x,0,2) = \left(\frac{1}{\sqrt{2}}\right)^n \exp\left(\frac{1}{4}\sum x_i^2\right)$$

A type of model in which it is easy to identify a minimal sufficient statistic (and which has lots of other convenient mathematical properties) is an "exponential family" model

Def If $T_1, T_2, \ldots, T_k$ are linearly independent real-valued functions on $\mathbb{R}^q$ or on some discrete space, $h(x) \geq 0$ and for parameters $\eta \in \mathbb{R}^k$ pdfs or pmfs

$$f(x|\eta) = A(\eta)\, h(x) \exp \sum_{j=1}^{k} \eta_j T_j(x)$$

we'll call the family of dsns an exponential family

(with "natural parameter" η)

Fact: $\Xi = \left\{ \eta \in \mathbb{R}^k \;\middle|\; \int h(x) \exp \sum_{j=1}^{k} \eta_j T_j(x)\, dx < \infty \right\}$

is a convex subset of $\mathbb{R}^k$ — for a given observation space and set of functions $T_1, \ldots, T_k$ this is the largest possible parameter space for the exponential family

$$\left(A(\eta) = \left(\int \qquad \right)^{-1} \right)$$

— Ξ is called the "natural parameter space"

Note: If $X_1, X_2, \ldots, X_n$ are iid $f(x|\eta)$

the joint pdf or pmf is of the form

$$f(x|\eta) = A(\eta)^n \prod_{i=1}^{n} h(x_i) \exp \sum_{j=1}^{k} \eta_j \left(\sum_{i=1}^{n} T_j(x_i) \right)$$

and the factorization Thm says that

$$T(X) = \left(\sum T_1(X_i), \sum T_2(X_i), \ldots, \sum T_k(X_i) \right)$$

is sufficient for η in any $\mathcal{E}^* \subset \mathcal{E}$ —

This is called the "natural sufficient statistic" for the parameter η (of for the family $\mathcal{P} = \{ P_\eta \}_{\eta \in \mathcal{E}^*}$)

not only is the natural sufficient statistic sufficient but provided $\mathcal{E}^*$ is "big enough" it is also minimal sufficient — that is

Thm If $\mathcal{E}^*$ $(\subset \mathcal{E})$ contains an open rectangle in $\mathbb{R}^k$ then $T(X)$ above is minimal sufficient

— Argument for this is on a handout .... it involves showing
  i) $T(X)$ has a property called "completeness"
  ii) Bahadur's Thm to be true

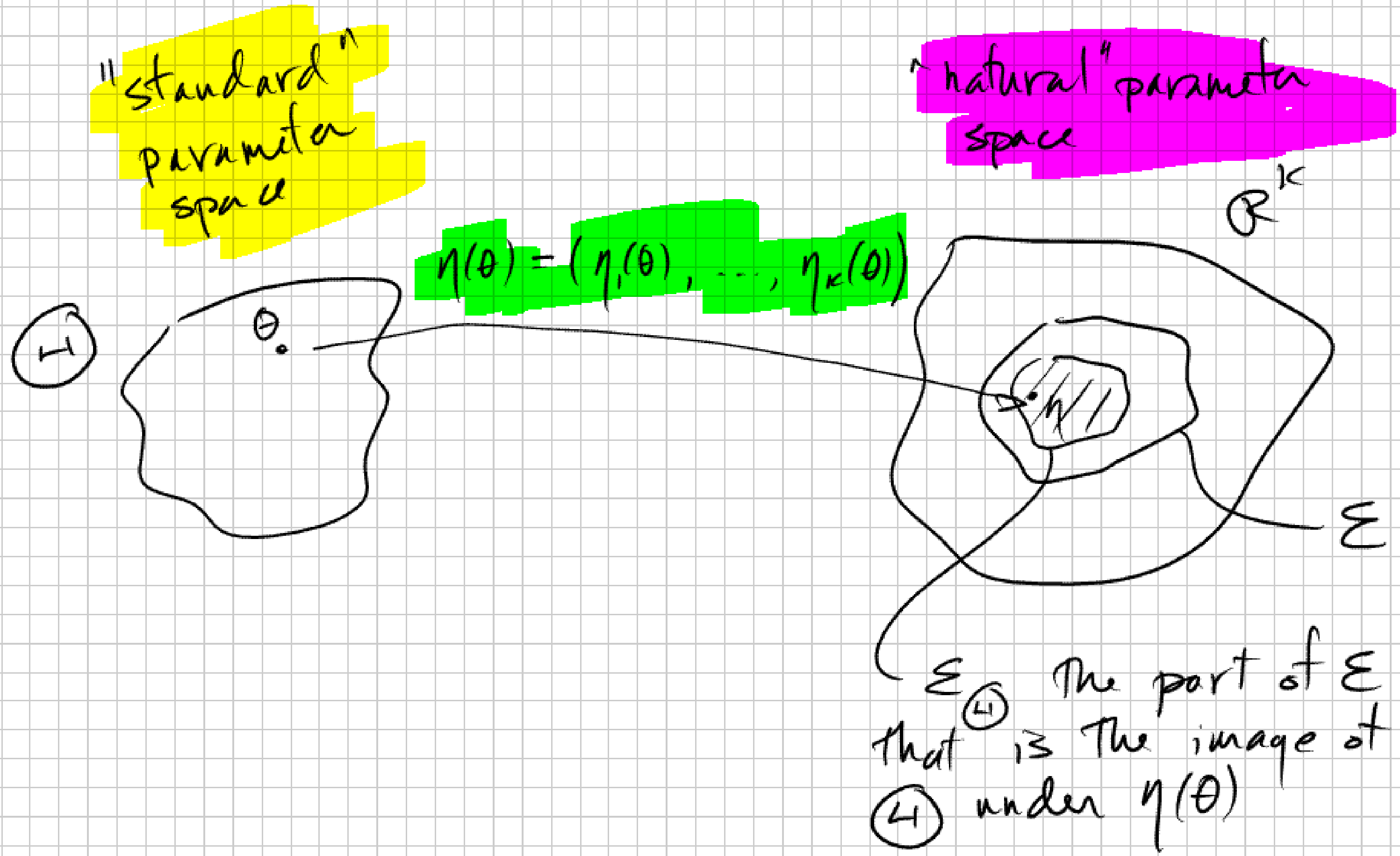
This is not really the form in which one likes to apply this result

Example   Poisson (λ) — marginal pmf here on $\{0,1,2,\ldots\}$ is

$$f(x \mid \lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$= e^{-\lambda} \underbrace{\frac{1}{x!}}_{h(x)} e^{\overbrace{x}^{T_1(x)} \underbrace{\log \lambda}_{\eta_1}}$$

That is, the "standard" parameterization (involving λ) is not the "natural" parameterization (involving η)


"standard" parameter space
"natural" parameter space
$\eta(\theta) = (\eta_1(\theta), \ldots, \eta_k(\theta))$
Θ
θ
$\mathbb{R}^k$
η
$\mathcal{E}$
$\mathcal{E}_\Theta$ The part of $\mathcal{E}$ that is the image of Θ under $\eta(\theta)$

$T(X)$ sufficient for $\eta \in \mathcal{E}$

$\Rightarrow T(X)$ sufficient of $\eta \in \mathcal{E}_{\Theta}$

and if $\mathcal{E}_{\Theta}$ contains an open rectangle we apply the theorem to conclude that $T(X)$ is minimal sufficient $\underbrace{\text{for } \eta \in \mathcal{E}_{\Theta}}_{\text{for } \theta \in \Theta}$

Back to Poisson Example

$$f(x|\lambda) = e^{-\lambda} \frac{1}{x!} e^{x \log \lambda}$$

$$\eta(\lambda) = \log \lambda \qquad \lambda \in (0, \infty) = \Theta$$

$\mathcal{E}_{\Theta}$ = the image of Θ under the log function

$= \mathbb{R}^1$) contains an open interval

So in an iid problem, since we're working in an exponential family

$$T(X) = \sum_{i=1}^{n} T_1(X_i) = \sum_{i=1}^{n} X_i$$

is not only sufficient but is minimal sufficient

Example Normal (μ, σ^2) marginal

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2} + \frac{x\mu}{\sigma} - \frac{\mu^2}{2\sigma^2}\right)$$

use $T_1(x) = x^2$ $T_2(x) = x$

$\eta_1 = -\frac{1}{2\sigma^2}$ $\eta_2 = \frac{\mu}{\sigma^2}$