

Completeness

Outline

- 1) Completeness
- 2) Ancillarity
- 3) Basu's Theorem

Completeness

Def $T(X)$ is complete for $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$

$$\text{if } \mathbb{E}_\theta f(T(X)) = 0 \quad \forall \theta$$

$$\Rightarrow f(T) \stackrel{a.s.}{=} 0 \quad \forall \theta$$

[Name comes from a prior notion that

$\mathcal{P}^T = \{P_\theta^T : \theta \in \Theta\}$ is "complete basis"

wrt inner product $\langle f, P_\theta^T \rangle = \int f(t) dP_\theta^T(t)$

(see HW 3)

Ex. (Cont'd) Laplace location family has
minimal suff. stat. $S = (X_{(i)})_{i=1}^n$. Complete?

No: Let $M(s) = \text{median}(X)$

$$\bar{X}(s) = \frac{1}{n} \sum X_i$$

$$\mathbb{E}_\theta \bar{X} = \mathbb{E}_\theta M = \theta \quad (\text{by symmetry})$$

$$\mathbb{E}_\theta [\bar{X}(s) - M(s)] = 0 \quad \forall \theta$$

$S(X)$ still has "a lot of extra fluff"

Ex $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} U[0, \theta] \quad \theta \in (0, \infty)$

Can show $T(X) = X_{(n)}$ min. suff. Complete?

Find density of $T(X)$:

$$P_{\theta}(T \leq t) = \left(\frac{t}{\theta} \wedge 1\right)^n = \left(\frac{t}{\theta}\right)^n \wedge 1$$

$$\Rightarrow p_{\theta}(t) = \frac{d}{dt} P_{\theta}(T \leq t)$$

$$= n \frac{t^{n-1}}{\theta^n} \mathbb{1}_{\{t \leq \theta\}}$$

Suppose $0 = E_{\theta} f(T) \quad \forall \theta > 0$

$$= \frac{n}{\theta^n} \int_0^{\theta} f(t) t^{n-1} dt \quad \forall \theta > 0$$

$$\Rightarrow \int_0^{\theta} f(t) t^{n-1} dt = 0 \quad \forall \theta > 0$$

$$\Rightarrow f(t) t^{n-1} = 0 \quad \text{a.e. } t > 0$$

Def Assume $\mathcal{P} = \{P_\eta : \eta \in \Xi\}$ has densities

$$p_\eta(x) = e^{\eta' T(x) - A(\eta)} h(x)$$

If $T(x)$ satisfies no linear constraint $\left(\begin{matrix} \exists \beta \neq 0, \alpha: \\ \beta' T(x) \stackrel{a.s.}{=} \alpha \end{matrix} \right)$
and Ξ contains an open set, we say
 \mathcal{P} is full-rank

If \mathcal{P} is not full-rank we say it is curved

[Note: If $T(x)$ satisfies linear constraint, then
 \mathcal{P} might still be full-rank for a lower-dim.
sufficient statistic]

Theorem If \mathcal{P} is full rank then
 $T(x)$ is complete sufficient

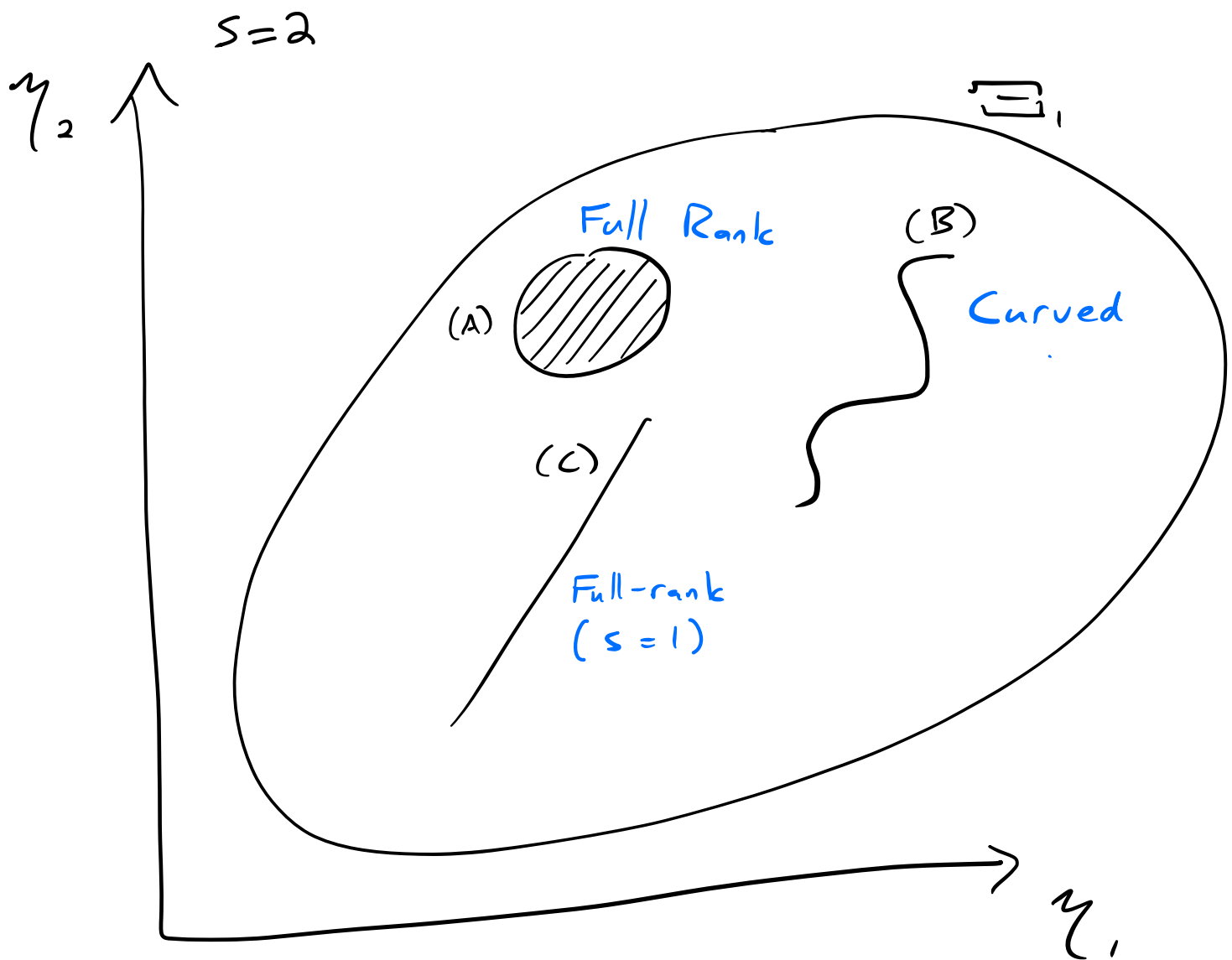
Proof in Lehmann & Romano, Thm. 4.3.1

Proof idea wlog $T(x) = x$, $p_\eta(x) = e^{\eta' x - A(\eta)}$, $0 \in \Xi^{\text{interior}}$

Write $f(x) = f^+(x) - f^-(x)$, for $f^+, f^- \geq 0$

$$\int e^{\eta' x} \underbrace{f^+(x)}_{d\nu^+(x)} d\mu(x) = \int e^{\eta' x} \underbrace{f^-(x)}_{d\nu^-(x)} d\mu(x) \quad \text{Both MGFs}$$

Uniqueness of MGFs $\Rightarrow f^+ = f^-$



Theorem If $T(X)$ complete sufficient
for \mathcal{P} then $T(X)$ is minimal

Game plan for completeness proofs: show two things are
a.s. equal by showing they have = expectation.

Proof Assume $S(X)$ is minimal suff

$$\text{Let } \bar{T}(S(X)) = \mathbb{E}_{\theta} [T(X) \mid S(X)]$$

~~θ~~ $\leftarrow S \text{ suff.}$

$$\text{Claim: } \bar{T}(S(X)) \stackrel{\text{a.s.}}{=} T(X)$$

$$\text{We have } S(X) \stackrel{\text{a.s.}}{=} f(T(X)) \quad (S \text{ minimal suff})$$

$$\text{Let } g(t) = t - \bar{T}(f(t))$$

$$\begin{aligned} \mathbb{E}_{\theta} [g(T(X))] &= \mathbb{E}_{\theta} T(X) - \mathbb{E}_{\theta} [\bar{T}(S(X))] \\ &= \mathbb{E}_{\theta} T(X) - \mathbb{E}_{\theta} [\mathbb{E}[T \mid S]] \\ &= 0 \end{aligned}$$

$$\Rightarrow g(T(X)) \stackrel{\text{a.s.}}{=} 0 \quad (\text{completeness}) \quad \square$$

Ancillarity

Two reasons to care about completeness:

1) Uniqueness of unbiased estimators using T

$$\text{If } \mathbb{E}_\theta \delta_1(T) = \mathbb{E}_\theta \delta_2(T) = g(\theta), \forall \theta \in \Theta$$

$$\text{Then } \mathbb{E}_\theta[\delta_1 - \delta_2] = 0 \Rightarrow \delta_1 \stackrel{a.s.}{=} \delta_2$$

[We will explore this further next time]

2) Basu's theorem: neat way to show independence

Def $V(X)$ is ancillary for $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$
if its distribution does not depend
on θ . (V carries no info. about θ)

(Aside:) Conditionality Principle

If $V(X)$ is ancillary then all inference
should be conditional on $V(X)$

[will return to this in testing & CI unit]

Basu's Theorem

Theorem (Basu)

If $T(X)$ is complete sufficient and $V(X)$ is ancillary for \mathcal{J} , then

$$V(X) \perp\!\!\!\perp T(X) \quad \text{for all } \theta \in \Theta$$

Proof

Want $P_\theta(V \in A, T \in B) = P_\theta(V \in A) P_\theta(T \in B)$ all A, B, θ

$$\text{Let } q_A(T(X)) = P_\theta(V \in A \mid T) \quad \leftarrow T \text{ suff.}$$

$$\rho_A = P_\theta(V \in A) \quad \leftarrow V \text{ ancillary}$$

$$E_\theta[q_A(T) - \rho_A] = \rho_A - \rho_A = 0, \quad \forall \theta$$

$$\Rightarrow q_A(T) \stackrel{\text{a.s.}}{=} \rho_A \quad \forall \theta$$

$$\begin{aligned} P_\theta(V \in A, T \in B) &= \int q_A(t) 1\{t \in B\} dP_\theta^T(t) \\ &= \rho_A \int 1\{t \in B\} dP_\theta^T(t) \\ &= P(V \in A) P_\theta(T \in B) \quad \square \end{aligned}$$

Using Basu's Theorem

Ancillarity, Completeness, Sufficiency are all properties wrt a family \mathcal{P}

Independence is a property of a distribution

If you can't verify the thm's hypotheses for one family, try a different family!

Ex. $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2) \quad \mu \in \mathbb{R}, \sigma^2 > 0$

Sample mean $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$

Sample variance $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

Want to show $\bar{X} \perp\!\!\!\perp S^2$

But neither stat. is ancillary or sufficient in the full family with μ, σ^2 unknown

To apply Basu, use family with σ^2 known:

$$\mathcal{P} = \{ N(\mu, \sigma^2)^n : \mu \in \mathbb{R} \}$$

In \mathcal{P} , \bar{X} is complete sufficient

and S^2 is ancillary since

$$S^2 = \sum (z_i - \bar{z})^2 \text{ for } z_i = X_i - \mu \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$$

↖ not statistics
but doesn't matter

Therefore $\bar{X} \perp\!\!\!\perp S^2$

[Conclusion has nothing to do with "known"
or "unknown" parameters]