Outline

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

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

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

\( \Rightarrow f(T) \) a.s. \( = 0 \quad \forall \theta \)

[Name comes from a prior notion that \( P^T = \{P^T_\theta : \theta \in \Theta \} \) is "complete basis"

wrt inner product \( \langle f, P^T_\theta \rangle = \int f(x) dP^T_\theta(x) \) (see HW 3)]

Ex. (Cont’d) Laplace location family has minimal suff stat. \( S = (X_{(i)})_{i=1} \). Complete?

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

\( \overline{X}(S) = \frac{1}{n} \sum X_i \)

\( \mathbb{E}_\theta \overline{X} = \mathbb{E}_\theta M = \theta \) (by symmetry)

\( \mathbb{E}_\theta [\overline{X}(S) - M(S)] = 0 \quad \forall \theta \)

\( S(x) \) still has "a lot of extra fluff"
\[
\text{Ex} \quad X_1, \ldots, X_n \overset{i.d.}{\sim} U[0, \theta] \quad \theta \in (0, \infty)
\]

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

Find density of \( T(X) \):
\[
\begin{align*}
\Pr_\theta(T \leq t) &= \left( \frac{t}{\theta} \wedge 1 \right)^n = \left( \frac{t}{\theta} \right)^n \wedge 1 \\
\Rightarrow \rho_\theta(t) &= \frac{d}{dt} \Pr_\theta(T \leq t) \\
&= nt^{n-1}\theta^n \quad 1 \{ t \leq \theta \}
\end{align*}
\]

Suppose \( O = \mathbb{E}_\theta f(t) \quad \forall \theta > 0 \)
\[
= \frac{n}{\theta^n} \int_0^\theta f(t) t^{n-1} dt \quad \forall \theta > 0
\]
\[
\Rightarrow \int_0^\theta f(t) t^{n-1} dt = 0 \quad \forall \theta > 0
\]
\[
\Rightarrow f(t) t^{n-1} = 0 \quad \text{a.e.} \quad t > 0
\]
**Def** Assume \( P = \{ P_\theta : \theta \in \Theta \} \) has densities
\[
\rho_\theta(x) = e^{\theta' T(x) - A(\theta)} h(x) \quad (\forall \theta \neq 0, \alpha: \theta' T(x) = \alpha)
\]

If \( T(x) \) satisfies no linear constraint and \( \Theta \) contains an open set, we say \( P \) is **full-rank**

If \( P \) is not full-rank we say it is **curved**

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

**Theorem** If \( P \) is full rank then \( T(x) \) is complete sufficient

**Proof** in Lehmann & Romano, Thm. 4.3.1

Proof idea wlog \( T(x) = x \), \( \rho_\theta(x) = e^{\theta' x - A(\theta)} \), \( \theta \in \Theta^\circ \)

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

\[
\int e^{\theta' x} f^+(x) dm(x) = \int e^{\theta' x} f^-(x) dm(x) \quad \text{Both MGFs}
\]

Uniqueness of MGFs \( \Rightarrow f^+ = f^- \)
Full Rank

Full-rank (s = 1)

Curved
Theorem: If \( T(x) \) complete sufficient for \( P \) 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

Let \( \bar{T}(S(x)) = \mathbb{E}_\Theta \left[ T(x) \mid S(x) \right] \)

Claim: \( \bar{T}(S(x)) \) a.s. = \( T(x) \)

We have \( S(x) \) a.s. \( f(T(x)) \) (\( S \) minimal suff)

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

\[
\mathbb{E}_\Theta [g(T(x))] = \mathbb{E}_\Theta T(x) - \mathbb{E}_\Theta [\bar{T}(S(x))]
\]

\[
= \mathbb{E}_\Theta T(x) - \mathbb{E}_\Theta [\mathbb{E}[T \mid S]]
\]

\[
= 0
\]

\[
g(T(x)) \text{ a.s.} = 0 \text{ (completeness)}
\]

\( \Box \)
Ancillarity

Two reasons to care about completeness:

1) Uniqueness of unbiased estimators using $T$:
   If $E_\theta \delta_1(T) = E_\theta \delta_2(T) = g(\theta), \forall \theta \in \Theta$
   Then $E_\theta[\delta_1 - \delta_2] = 0 \implies \delta_1 \equiv \delta_2$  
   [We will explore this further next time]

2) Basu’s theorem: neater way to show independence

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

(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 $\Theta$, then $V(X) \perp T(X)$ 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 $A, B, \Theta$

Let $q_A(T(X)) = P_{\Theta}(V \in A \mid T)$

$$P_{\Theta}(V \in A)$$

$$1 E_{\Theta}[q_A(T) - P_{\Theta}(V \in A)] = P_{\Theta}(V \in A) - P_{\Theta}(V \in A) = 0 \quad \forall \Theta$$

$$\Rightarrow q_A(T) = P_{\Theta}(V \in A) \quad \forall \Theta$$

$$P_{\Theta}(V \in A, T \in B) = \int q_A(t) 1_{T \in B} \, dP_{\Theta}(t)$$

$$= P_{\Theta}(V \in A) P_{\Theta}(T \in B)$$

$$= P_{\Theta}(V \in A, T \in B)$$
Using Basu's Theorem

Ancillarity, Completeness, Sufficiency are all properties wrt a family \( \mathcal{D} \)

Independence is a property of a distribution

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

Ex. \( X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2) \) \( \mu \in \mathbb{R}, \sigma^2 > 0 \)

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 \( \bar{X} \parallel S^2 \)

But neither stat. is ancillary or sufficient in the full family with \( \mu, \sigma^2 \) unknown

To apply Basu, use family with \( \sigma^2 \) known:

\[
\mathcal{D} = \{ N(\mu, \sigma^2)^n : \mu \in \mathbb{R} \}
\]
In $\mathcal{P}$, $\bar{X}$ is complete sufficient and $S^2$ is ancillary since

$$S^2 = \sum (Z_i - \bar{Z})^2 \quad \text{for} \quad Z_i = X_i - \mu \sim \mathcal{N}(0, \sigma^2)$$

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

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