

STA732

Statistical Inference

Lecture 04: Completeness and Ancillarity

Yuansi Chen

Spring 2023

Duke University

<https://www2.stat.duke.edu/courses/Spring23/sta732.01/>



- **Sufficient statistic T :** when T is known, we have enough information to estimate θ . Check the conditional $P_\theta(X = x \mid T = t)$
- **Factorization theorem:** if we can factor the likelihood, then we prove T is sufficient

$$p_\theta(x) = g_\theta(T(x))h(x)$$

- **Minimal sufficiency:** we can order the sufficient statistics and define the minimal (intuitively it is the most compressed version of sufficient statistics)

Goal of Lecture 04

1. Define ancillarity and completeness
2. Complete statistics in exponential family
3. Basu's theorem

Chap. 3.5 in Keener or Chap. 1.6 in Lehmann and Casella

Ancillarity and Completeness

Definition. Ancillarity

Suppose X has distribution from a family $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$. A statistics $V(X)$ is called **ancillary** if its distribution does not depend on θ . So $V(X)$ by itself provides no information about θ .

- In the Laplace location family example of Lecture 3, the joint density is

$$p_{\theta}(x) = \frac{1}{2^n} \exp \left\{ - \sum_{i=1}^n |x_i - \theta| \right\}$$

We showed that the order statistics is minimal sufficient. But $X_{(1)} - X_{(3)}$ is ancillary. The distribution of $(X_{(1)} - \theta) - (X_{(3)} - \theta)$ does not depend on θ

- Similarly for $X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\theta, 1)$. $X_1 - X_2$ is ancillary

Definition. First-order ancillary

A statistics $V(X)$ is called **first-order ancillary** if $\mathbb{E}_\theta[V(X)]$ does not depend on θ .

Example:

$X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2)$, X_1 is first-order ancillary but it is not ancillary.

Definition. Completeness

Suppose X has distribution from a family $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$. A statistics $T(X)$ is **complete** if no non-constant function of T is first-order ancillary. In other words, for any non-constant f ,

$$\mathbb{E}_\theta[f(T(X))] = 0, \forall \theta \Rightarrow f(T(X)) = 0, \text{ a.s.}, \forall \theta$$

Definition. Completeness

Suppose X has distribution from a family $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$. A statistics $T(X)$ is **complete** if no non-constant function of T is first-order ancillary. In other words, for any non-constant f ,

$$\mathbb{E}_\theta[f(T(X))] = 0, \forall \theta \Rightarrow f(T(X)) = 0, \text{ a.s.}, \forall \theta$$

Remark:

- In some sense, the complete + sufficient formalizes our **ideal** notion of “optimal data compression”
- The minimal sufficiency is our **achievable** notion of “optimal data compression”

Example 1: minimal sufficient statistic is not necessarily complete

In the Laplace location family example with joint density

$$p_{\theta}(x) = \frac{1}{2^n} \exp \left\{ - \sum_{i=1}^n |x_i - \theta| \right\}$$

$S = (X_{(1)}, \dots, X_{(n)})$ is minimal sufficient statistics. But it is not complete! Consider

$$\text{median}(X_{(i)}) - \text{mean}(X_{(i)})$$

why is its expectation 0? symmetry

Example 2: a minimal sufficient complete statistic

$X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} \text{Unif}[0, \theta], \theta \in (0, \infty)$. Show that $T(X) = X_{(n)}$ is minimal sufficient and complete.

proof:

Example 3: a complete statistic that is not sufficient

$X_1, \dots, X_{10} \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\theta, 1)$.

- X_1 is not sufficient
- X_1 is complete

proof:

Theorem 3.17 in Keener

If T is complete and sufficient, then T is minimal sufficient.

proof:

Complete statistics in exponential family

Definition. Full-rank exponential family

An exponential family \mathcal{P} with densities $p_\theta(x) = \exp\{\eta(\theta)^\top T(x) - B(\theta)\} h(x)$, $\theta \in \Omega$ is said to be **full rank** if the interior of $\eta(\Omega)$ is not empty and if T_1, \dots, T_s do not satisfy a linear constraint of the form $v^\top T \stackrel{\text{a.s.}}{=} c$.

Otherwise, we say the family \mathcal{P} is **curved**

Which of the following exponential family is full rank?

$$\bullet p_{\theta}(x) = \exp \left(\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}^{\top} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - B(\theta) \right) h(x)$$

$$\bullet p_{\theta}(x) = \exp \left(\begin{pmatrix} \theta_1 \\ \theta_1^2 \end{pmatrix}^{\top} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - B(\theta) \right) h(x)$$

$$\bullet p_{\theta}(x) = \exp \left(\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}^{\top} \begin{pmatrix} x_1 \\ -2x_1 \end{pmatrix} - B(\theta) \right) h(x)$$

$T(X)$ in full-rank exponential family is complete sufficient

Theorem 3.19 in Keener

In an exponential family of full rank, $T(X)$ is complete sufficient.

proof in Lehmann & Romano, Thm 4.3.1

Key proof ideas:

- W.L.O.G., assume the parameter space contain the rectangle $I = \{\theta \mid |\theta_j| \leq a\}$
- For $f = f^+ - f^-$ satisfying $\mathbb{E}_\theta f(T) = 0$, we have for $\theta \in I$

$$\int e^{\theta^\top t} f^+(t) d\nu(t) = \int e^{\theta^\top t} f^-(t) d\nu(t)$$

- We can normalize both sides to make appear moment generating functions on both sides. Equality by moment generating functions implies $f^+ = f^-$

Basu's theorem

Theorem 3.21 in Keener

If T is complete and sufficient for $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$, and if V is ancillary, then T and V are independent under P_θ for any $\theta \in \Omega$.

Theorem 3.21 in Keener

If T is complete and sufficient for $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$, and if V is ancillary, then T and V are independent under P_θ for any $\theta \in \Omega$.

Remark

- Ancillarity, completeness, sufficiency are all properties of a statistic with respect to a family \mathcal{P}
- Independence is a property with respect to a particular distribution P_θ

proof of Basu's theorem:

To show independence, we have to show $\forall \theta, A, B$ that

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

Define

$$q_A(t) = P_{\theta}(V \in A \mid T = t)$$

$$\rho_A = P_{\theta}(V \in A)$$

- First, show that they do not depend on θ .
- Second, show that they are equal by completeness

$X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$. Consider the subfamily $\mathcal{P}_{\sigma^2} = \{\mathcal{N}(\mu, \sigma^2)^n : \mu \in \mathbb{R}\}$.

- $\bar{X} = \frac{1}{n} \sum X_i$ is complete sufficient for \mathcal{P}_{σ^2} .
- $S^2 = \frac{1}{n-1} \sum (X_i - \bar{X})^2$ is ancillary **why?**

Apply Basu's theorem

Review basics of conditional expectations

Let X, Z be random variables with joint density $p(x, z) = p(z | x)p(x)$. Let h be a measurable function with $\int \int |h(x, z)| p(x, z) dx dz < \infty$. The conditional expectation of h given X is

$$\mathbb{E}[h(X, Z) | X = x] = \int h(x, z)p(z | x)dz$$

Basic properties of conditional expectation

- Pull-out property: if $h(x, z) = h_1(x)h_2(z)$, then

$$\mathbb{E}[h(X, Z) \mid X = x] = h_1(x)\mathbb{E}[h_2(Z) \mid X = x]$$

- Tower property:

$$\mathbb{E}[\mathbb{E}[h(X, Z) \mid X]] = \mathbb{E}[h(X, Z)]$$

- Conditional expectation under independence: if $p(x, z) = p(x)p(z)$, then

$$\mathbb{E}[h(Z) \mid X = x] = \mathbb{E}[h(Z)]$$

- V is **ancillary** if its distribution does not depend on θ
- **Completeness + sufficiency** as the ideal notion of optimal data compression
- **Basu's theorem** is useful to prove independence

- Rao-Blackwell Theorem
- UMVU

Thank you

