Chapter 1 & 2: Basics and Sufficiency *

1 Set-up for Statistics

- $(\Omega, \mathcal{F}, \mathcal{P}) \xrightarrow{X} (\mathcal{X}, \mathcal{B})$
- X is $\langle \mathcal{F}, \mathcal{B} \rangle$- measurable if $X^{-1}(B) \in \mathcal{F}$ for any $B \in \mathcal{B}$.
- Dominance: $\mu_1 \ll \mu_2$. Equivalently, $\mu_2(B) = 0 \implies \mu_1(B) = 0$
- σ-finite measure μ: $\exists A_1, \ldots \in \Omega, \Omega = \cup_{i=1}^{\infty} A_i, \mu(A_i) < \infty$.
- Radon-Nikodym Theorem: $P_\theta \ll \mu, \mu$ is σ-finite, then for each $\theta \in \Theta, \exists f_\theta : \mathcal{X} \to \mathbb{R}$ such that

$$P_\theta(B) = \int_B f_\theta d\mu, \forall B \in \mathcal{B}.$$

- $f_\theta(X)$ is called the likelihood function, with respect to measure μ. When X is completely discrete, μ is counting measure; when X is completely continuous, μ is Lebesgue measure; otherwise, μ can be a mixture of these two.

2 Sufficiency and Related Concepts

- Statistic $T(X) : (\mathcal{X}, \mathcal{B}) \to (\mathcal{T}, \mathcal{F})$ is a measurable map.
- σ-algebra generated by T is $\mathcal{B}_T := \sigma\langle T \rangle = \{T^{-11}(A) : A \in \mathcal{F}\} \subset \mathcal{B}$
- Sufficiency: $[X|T(X)] \perp \theta$. Equivalently, $T(x) = T(x') \implies f_\theta(x) = c(x, x')f_\theta(x')$. Equivalently, $\forall B \in \mathcal{B}, \exists \mathcal{B}_T$-measurable random variable $Y_B : \mathcal{X} \to \mathbb{R}$ such that

$$Y_B \equiv E_\theta(I_B|\mathcal{B}_T) = P_\theta(B|\mathcal{B}_T)$$

  a.s. P_θ for all $\theta \in \Theta$, i.e. the conditional proabablity of X given $T(X)$ does not depend on θ.

- Factorization Theorem (Halmos-Savage): Suppose $\mathcal{P} \ll \mu, \mu$ is a σ-finite measure on $(\mathcal{X}, B)$. Then $T(X)$ is sufficient for $\mathcal{P} \iff \exists$ nonnegative $\mathcal{B}$-measurable function $h : \mathcal{X} \to \mathbb{R}$ and a $\mathcal{F}$- measurable function $g_\theta : \mathcal{T} \to \mathbb{R}$ such that

$$\frac{dP_\theta}{d\mu}(x) = f_\theta(x) = g_\theta(T(X))h(x)$$

  a.s. μ for all $\theta \in \Theta$.

  - Lehmann's Theorem: (Lemma 1.2 / Page 37 of Shao): Let $T : (\mathcal{X}, \mathcal{B}) \to (\mathcal{T}, \mathcal{F})$ be measurable and let $\phi : (\mathcal{X}, \mathcal{B}) \to (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ be $\mathcal{B}_T \equiv \sigma\langle T \rangle$ -measurable. Then, there exists an $\mathcal{F}$ -measurable function $\psi : (\mathcal{T}, \mathcal{F}) \to (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ such that

$$\phi(x) = \psi(T(x))$$

  - Lemma 03 (Lemma 2.1/Page 104 of Shao): $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ is dominated by a σ -finite measure μ if and only if $\mathcal{P}$ is dominated by a probability measure λ of the form

$$\lambda = \sum_{i=1}^{\infty} c_i P_{\theta_i}$$

    for some countable subset $\{\theta_i\} \subset \Theta$ and a countable $\{c_i\}$ with $c_i \geq 0$ and $\sum_{i=1}^{\infty} c_i = 1$

*STA643: Advanced Theory of Statistical Inference. Instructed by Dr. Daniel Nordman. Arranged by Zhiling Gu

– Lemma 04: Suppose $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ is dominated by a σ -finite measure μ and $T : (\mathcal{X}, \mathcal{B}) \to (\mathcal{T}, \mathcal{F})$. Then, T is sufficient for $\mathcal{P}$ if and only if there exists a nonnegative $\mathcal{F}$ -measurable function $g_\theta : (\mathcal{T}, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that

$$\frac{dP_\theta}{d\lambda}(x) = g_\theta(T(x)) \quad \text{a.s. } \lambda$$

using a fixed probability measure λ in the form of Lemma 03.

- Minimal Sufficiency: A sufficient statistic $T : (\mathcal{X}, \mathcal{B}) \to (\mathcal{T}, \mathcal{F})$ is minimal sufficient for $\mathcal{P}$(or θ) provided for every sufficient statistic $S : (\mathcal{X}, \mathcal{B}) \to (\mathcal{S}, \mathcal{G})$, there is a function $U : \mathcal{S} \to \mathcal{T}$ such that

$$T = U \circ S \quad \text{a.s. } \mathcal{P}$$

(that is, the set $A = \{x \in \mathcal{X} : T(x) \neq U(S(x))\}$ satisfies $P_\theta(A) = 0$ for any θ)

– Theorem 06: Suppose $\mathcal{P}$ is dominated by a σ-finite measure μ and $T : (\mathcal{X}, \mathcal{B}) \to (\mathcal{T}, \mathcal{F})$ is sufficient. Suppose further that, if given versions of densities $\frac{dP_\theta}{d\lambda} = f_\theta$ and some $\mathcal{P}$ -null set N_0, it turns out that, for two datasets $x, y \in \mathcal{X} \backslash N_0$, the existence of a constant $k(x, y) > 0$ such that

$$f_\theta(x) = f_\theta(y) k(x, y) \quad \forall \theta \in \Theta \quad (*)$$

in turn implies that $T(x) = T(y)$. Then, T is minimal sufficient. (Null set $N_0 \in \mathcal{B}$ and $P_\theta(N_0) = 0$ for all $P_\theta \in \mathcal{P}$.) Proof: only need to show T is a function of $S \iff$ wherever $S(X) = S(Y), T(X) = T(Y)$.

– Theorem 07: For finite dimension measure $\mathcal{P} = \{P_i\}_{i=1}^k$, $T(X) = \left(\frac{f_1(X)}{f_0(X)} \cdots \frac{f_k(X)}{f_0(X)}\right)$ is minimal sufficient for $\mathcal{P}$.

– Theorem 08: $\mathcal{P} \ll \mathcal{P}_0\, (i.e. P(B) = 0, \forall P \in \mathcal{P}_0 \implies P(B) = 0, \forall P \in \mathcal{P}), \mathcal{P}_0 \subset \mathcal{P}$, T sufficient for $\mathcal{P}, \mathcal{P}_0$, T minimal sufficient for $\mathcal{P}_0 \implies T$ minimal sufficient for $\mathcal{P}$.