STAT 5010 Tutorial 2*

Oct. 2022

Definitions

1. A parametric family $\{P_\theta : \theta \in \Theta\}$ dominated by a σ-finite measure ν on $(\Omega, \mathcal{F})$ is called an *exponential family* if and only if
$$\frac{dP_\theta}{d\nu}(\omega) = \exp\left\{[\eta(\theta)]^\tau T(\omega) - \xi(\theta)\right\} h(\omega), \quad \omega \in \Omega,$$
where $\exp\{x\} = e^x, T$ is a random p-vector with a fixed positive integer p, η is a function from Θ to $\mathcal{R}^p, h$ is a nonnegative Borel function on $(\Omega, \mathcal{F})$, and $\xi(\theta) = \log\left\{\int_\Omega \exp\left\{[\eta(\theta)]^\tau T(\omega)\right\} h(\omega) d\nu(\omega)\right\}$

2. In an exponential family, consider the reparameterization $\eta = \eta(\theta)$ and
$$f_\eta(\omega) = \exp\left\{\eta^\tau T(\omega) - \zeta(\eta)\right\} h(\omega), \quad \omega \in \Omega,$$
where $\zeta(\eta) = \log\left\{\int_\Omega \exp\left\{\eta^\tau T(\omega)\right\} h(\omega) d\nu(\omega)\right\}$. This is the *canonical form* for the family, which is not unique. The new parameter η is called the *natural parameter*. The new parameter space $\Xi = \{\eta(\theta) : \theta \in \Theta\}$, a subset of $\mathcal{R}^p$, is called the *natural parameter space*. An exponential family in canonical form is called a *natural exponential family*. If there is an open set contained in the natural parameter space of an exponential family, then the family is said to be of *full rank*.

3. A measurable function of $X, T(X)$, is called a *statistic* if $T(X)$ is a known value whenever X is known, i.e., the function T is a known function.

4. Let X be a sample from an unknown population $P \in \mathcal{P}$, where $\mathcal{P}$ is a family of populations. A statistic $T(X)$ is said to be *sufficient* for $P \in \mathcal{P}$ (or for $\theta \in \Theta$ when $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ is a parametric family) if and only if the conditional distribution of X given T is known (does not depend on P or θ).

5. Let T be a sufficient statistic for $P \in \mathcal{P}$. T is called a *minimal sufficient* statistic if and only if, for any other statistic S sufficient for $P \in \mathcal{P}$, there is a measurable function ψ such that $T = \psi(S)$ a.s. $\mathcal{P}$

6. A statistic $V(X)$ is said to be *ancillary* if its distribution does not depend on the population P and *first-order ancillary* if $E[V(X)]$ is independent of P.

7. A statistic $T(X)$ is said to be *complete* for $P \in \mathcal{P}$ if and only if, for any Borel $f, E[f(T)] = 0$ for all $P \in \mathcal{P}$ implies $f(T) = 0$ a.s. $\mathcal{P}$. T is said to be *boundedly complete* if and only if the previous statement holds for any bounded Borel f.

Propositions and Theorems

1. If η_0 is an interior point of the natural parameter space, then the m.g.f. ψ_{η_0} of $P_{\eta_0} \circ T^{-1}$ is finite in a neighborhood of 0 and is given by
$$\psi_{\eta_0}(t) = \exp\left\{\zeta\left(\eta_0 + t\right) - \zeta\left(\eta_0\right)\right\}$$

2. (The factorization theorem) Suppose that X is a sample from $P \in \mathcal{P}$ and $\mathcal{P}$ is a family of probability measures on $(\mathcal{R}^n, \mathcal{B}^n)$ dominated by a σ-finite measure ν. Then $T(X)$ is sufficient for $P \in \mathcal{P}$ if and only if there are nonnegative Borel functions h (which does not depend on P) on $(\mathcal{R}^n, \mathcal{B}^n)$ and g_P (which depends on P) on the range of T such that
$$\frac{dP}{d\nu}(x) = g_P(T(x))h(x).$$

3. Suppose that $\mathcal{P}$ contains p.d.f.'s f_P w.r.t. a σ-finite measure and that there exists a sufficient statistic $T(X)$ such that, for any possible values x and y of $X, f_P(x) = f_P(y)\phi(x, y)$ for all P implies $T(x) = T(y)$, where ϕ is a measurable function. Then $T(X)$ is minimal sufficient for $P \in \mathcal{P}$.

4. If P is in an exponential family of full rank with p.d.f.'s given as in Definition. 2, then $T(X)$ is complete and sufficient for $\eta \in \Xi$.

*Dept. of Stat., CUHK. TA: YX

5. (Basu's theorem) Let V and T be two statistics of X from a population $P \in \mathcal{P}$. If V is ancillary and T is boundedly complete and sufficient for $P \in \mathcal{P}$, then V and T are independent w.r.t. any $P \in \mathcal{P}$.
6. A complete and sufficient statistic is also minimal sufficient. However, a minimal sufficient statistic is not necessarily complete.

Question 1

1. Let X and Y be two random variables such that Y has the binomial distribution with size N and probability π and, given $Y = y$, X has the binomial distribution with size y and probability p. Suppose that $p \in (0,1)$ and $\pi \in (0,1)$ are unknown and N is known. Show that (X, Y) is minimal sufficient for (p, π).

Solution:
Let $A = \{(x,y) : x = 0,1,\ldots,y, y = 0,1,\ldots,N\}$. The joint probability density of (X,Y) with respect to the counting measure is

$$\begin{aligned} &\binom{N}{y} \pi^y (1-\pi)^{N-y} \binom{y}{x} p^x (1-p)^{y-x} I_A \\ = &\exp\left\{ x \log \frac{p}{1-p} + y \log \frac{\pi(1-p)}{1-\pi} + N \log(1-\pi) \right\} \binom{N}{y} \binom{y}{x} I_A . \end{aligned}$$

Hence, (X,Y) has a distribution from an exponential family of full rank ($0 < p < 1$ and $0 < \pi < 1$). This implies that (X,Y) is minimal sufficient for (p, π)

Question 2

1. Let $X_1, \ldots, X_n$ be i.i.d. random variables from P_θ, the uniform distribution $U(\theta, \theta+1), \theta \in \mathcal{R}$. Prove that $T = \left(X_{(1)}, X_{(n)}\right)$ is minimal sufficient.

Solution: Suppose that $n > 1$. The joint Lebesgue p.d.f. of $(X_1, \ldots, X_n)$ is

$$f_\theta(x) = \prod_{i=1}^n I_{(\theta, \theta+1)}(x_i) = I_{\underline{(x_{(n)}-1, x_{(1)})}}(\theta), \quad x = (x_1, \ldots, x_n) \in \mathcal{R}^n,$$

where $x_{(i)}$ denotes the i th smallest value of $x_1, \ldots, x_n$. By the factorization theorem, $T = \left(X_{(1)}, X_{(n)}\right)$ is sufficient for θ. Note that

$$x_{(1)} = \sup\{\theta : f_\theta(x) > 0\} \quad \text{and} \quad x_{(n)} = 1 + \inf\{\theta : f_\theta(x) > 0\}.$$

If $S(X)$ is a statistic sufficient for θ, then by the factorization theorem, there are Borel functions h and g_θ such that $f_\theta(x) = g_\theta(S(x))h(x)$. <u>For x with $h(x) > 0$, $x_{(1)} = \sup\{\theta : g_\theta(S(x)) > 0\}$ and $x_{(n)} = 1 + \inf\{\theta : g_\theta(S(x)) > 0\}$</u>. Hence, there is a measurable function ψ such that $T(x) = \psi(S(x))$ when $h(x) > 0$. Since $h > 0$ a.s. $\mathcal{P}$, we conclude that T is minimal sufficient.

Proofs for Some propositions

- Prop. 3.

  Proof. From Bahadur (1957), there exists a minimal sufficient statistic $S(X)$. The result follows if we can show that $T(X) = \psi(S(X))$ a.s. $\mathcal{P}$ for a measurable function ψ. By the factorization theorem, there are Borel functions g_P and h such that $f_P(x) = g_P(S(x))h(x)$ for all P. Let $A = \{x : h(x) = 0\}$. Then $P(A) = 0$ for all P. For x and y such that $S(x) = S(y), x \notin A$ and $y \notin A$,

$$\begin{aligned} f_P(x) &= g_P(S(x))h(x) \\ &= g_P(S(y))h(x)h(y)/h(y) \\ &= f_P(y)h(x)/h(y) \end{aligned}$$

  for all P. Hence $T(x) = T(y)$. This shows that there is a function ψ such that $T(x) = \psi(S(x))$ except for $x \in A$. It remains to show that ψ is measurable. Since S is minimal sufficient, $g(T(X)) = S(X)$ a.s. $\mathcal{P}$ for a measurable function g. Hence g is one-to-one and $\psi = g^{-1}$. The measurability of ψ follows from Theorem 3.9 in Parthasarathy (1967). ∎

- Prop. 4.

  Proof. Obviously, T is sufficient. Suppose that there is a function f such that $E[f(T)] = 0$ for all $\eta \in \Xi$. Then,

$$\int f(t) \exp\{\eta^\tau t - \zeta(\eta)\}\, d\lambda = 0 \qquad \text{for all } \eta \in \Xi,$$

  where λ is a measure on $(\mathcal{R}^p, \mathcal{B}^p)$. Let η_0 be an interior point of Ξ. Then

$$\int f_+(t) e^{\eta^\tau t} d\lambda = \int f_-(t) e^{\eta^\tau t} d\lambda \qquad \text{for all } \eta \in N(\eta_0),$$

  where $N(\eta_0) = \{\eta \in \mathcal{R}^p : \|\eta - \eta_0\| < \epsilon\}$ for some $\epsilon > 0$. In particular,

$$\int f_+(t) e^{\eta_0^\tau t} d\lambda = \int f_-(t) e^{\eta_0^\tau t} d\lambda = c.$$

  If $c = 0$, then $f = 0$ a.e. λ. If $c > 0$, then $c^{-1} f_+(t) e^{\eta_0^\tau t}$ and $c^{-1} f_-(t) e^{\eta_0^\tau t}$ are p.d.f.'s w.r.t. λ and this implies that their m.g.f.'s are the same in a neighborhood of 0. Thus, $c^{-1} f_+(t) e^{\eta_0^\tau t} = c^{-1} f_-(t) e^{\eta_0^\tau t}$, i.e., $f = f_+ - f_- = 0$ a.e. λ. Hence T is complete. ∎

Prop 1. $\eta_0 \longleftrightarrow P_{\eta_0}$. $\boxed{T = T(X)}$ pdf.

$$\psi_{\eta_0}(t) = E\exp\{t^T T\} = \int \exp\{t^T T\}\exp\{\eta_0^T T(\omega) - \zeta(\eta_0)\} h(\omega)d\omega.$$

$$= \int \exp\{(t+\eta_0)^T T(\omega) - \zeta(\eta_0)\} h(\omega)d\omega.$$

$$= \frac{\int \exp\{(t+\eta_0)^T T(\omega)\} h(\omega)d\omega.}{\exp(\zeta(\eta_0))}$$

$$\exp(\zeta(\eta)) = \int_\Omega \exp\{\eta^T T(\omega)\} h(\omega)d\omega.$$

$$\exp(\zeta(\eta+t)) = \int_\Omega \exp\{(\eta+t)^T T(\omega)\} h(\omega)d\omega.$$

$$= \frac{\exp\{\zeta(\eta_0+t)\}}{\exp\{\zeta(\eta_0)\}}.$$

$$= \exp\{\zeta(\eta_0+t) - \zeta(\eta_0)\}.$$

Prop 6. Suppose T c.s. $\boxed{S}$ minimal suff. $S = g(T)$.

Complete & sufficient stat. may not exist.

$$E[X|Y] = f(Y). \quad E[X|Y=y] = f(y).$$

consider: $T - E[T|S] = T - E[T|g(T)]$.

$$= T - \psi(g(T)).$$

$$E[T - E[T|S]] = ET - E[E[T|S]]$$

$$= ET - ET = 0.$$

$\Rightarrow P(T - E[T|S] = 0) = 1.$ by completeness of T.

$\Rightarrow T = E[T|S] = f(S). \Rightarrow T$ is min. suff.

$S = g(T)$. $E[f(S)] = E[f(g(T))] = 0$. $f(g(T)) = 0 \Leftrightarrow f(S) = 0$.