STAT6201: Theoretical Statistics I

Satvik Saha

Homework 2

1. Let $\mathcal{P}_0 \subseteq \mathcal{P}$, such that $P(A) = 0$ for all $P \in \mathcal{P}_0$ implies that $P(A) = 0$ for all $P \in \mathcal{P}$. Furthermore, let T be sufficient for $\mathcal{P}$, and complete for $\mathcal{P}_0$. We claim that T is complete sufficient for $\mathcal{P}$. Indeed, let f be a measurable function such that $\mathbb{E}_P[f(T)] = 0$ for all $P \in \mathcal{P}$. By completeness of T for $\mathcal{P}_0$, we have $f(T) = 0$, P almost surely for all $P \in \mathcal{P}_0$. Set $A = \{x \in \mathcal{X} \colon f(T(x)) \neq 0\}$, so we have $P(A) = 0$ for all $P \in \mathcal{P}_0$; by assumption, we have $P(A) = 0$ for all $P \in \mathcal{P}$, i.e. $f(T) = 0$, P almost surely for all $P \in \mathcal{P}$. It follows that T is complete sufficient for $\mathcal{P}$.

2. (a) Let U be minimal sufficient, and T be complete sufficient for some family of distributions $\mathcal{P}$. Minimality of U means that $U = g(T)$ for some measurable g. Define

$$h \colon T(\mathcal{X}) \to g \circ T(\mathcal{X}), \qquad t \mapsto \mathbb{E}_P\left[T \mid U = g(t)\right] - t.$$

Note that this is well defined; sufficiency of U guarantees that the conditional expectation $\mathbb{E}_P[T \mid U = u]$ is a function of u independent of $P \in \mathcal{P}$! With this, observe that

$$\mathbb{E}_P\left[h(T)\right] = \mathbb{E}_P\left[\mathbb{E}_P[T \mid U]\right] - \mathbb{E}_P[T] = \mathbb{E}_P\left[T\right] - \mathbb{E}_P[T] = 0.$$

Completeness of T forces $h(T) = 0$, P almost surely for all $P \in \mathcal{P}$. This immediately gives

$$T = \mathbb{E}_P[T \mid U] := \tilde{g}(U),$$

whence T is a function of U. It follows that T is minimal sufficient for $\mathcal{P}$, being in bijection with the minimal sufficient statistic U. Alternatively, note that for any sufficient statistic V, we have $U = g_V(V)$ for some measurable g_V, so $T = \tilde{g} \circ g_V(V)$, whence T is minimal sufficient for $\mathcal{P}$.

(b) First, observe that (measurable) functions of complete sufficient statistics are complete. Indeed, if T is complete sufficient and g is a measurable function, then for measurable f, we have $\mathbb{E}_P[f(g(T)) = 0]$ for all $P \in \mathcal{P}$, which gives $f \circ g(T) = 0$ for all $P \in \mathcal{P}$ via completeness of T, whence $g(T)$ is complete for $\mathcal{P}$.

With this, let $X_1, \dots, X_n \overset{iid}{\sim} \text{Uniform}(\theta - 1/2, \theta + 1/2)$. Then,

$$\begin{aligned} f(\boldsymbol{x}; \theta) &= \prod_{i=1}^{n} \mathbf{1}(\theta - 1/2 < x_i < \theta + 1/2) \\ &= \mathbf{1}(\theta - 1/2 < x_{(1)} \leq x_{(n)} < \theta + 1/2), \end{aligned}$$

so

$$\frac{f(\boldsymbol{x}; \theta)}{f(\boldsymbol{y}; \theta)} = \frac{\mathbf{1}(\theta - 1/2 < x_{(1)} \leq x_{(n)} < \theta + 1/2)}{\mathbf{1}(\theta - 1/2 < y_{(1)} \leq y_{(n)} < \theta + 1/2)}$$

is independent of θ if and only if the indicator functions describe precisely the same intervals, i.e. $(x_{(1)}, x_{(n)}) = (y_{(1)}, y_{(n)})$. Thus, $T(\boldsymbol{X}) = (X_{(1)}, X_{(n)})$ is minimal sufficient for $\theta \in \mathbb{R}$.

Now, suppose that a complete sufficient statistic U for $\theta \in \mathbb{R}$ exists; then, $T = g(U)$ for some measurable function g by minimality of T, whence T is complete minimal sufficient for $\theta \in \mathbb{R}$. By setting $h(\boldsymbol{t}) = t_2 - t_1$, we have

$$\begin{aligned}
\mathbb{E}_\theta[h(T)] &= \mathbb{E}_\theta[X_{(n)} - X_{(1)}] \\
&= \mathbb{E}_\theta[(X_{(n)} - \theta) - (X_{(1)} - \theta)] \\
&= \mathbb{E}_0[X_{(n)} - X_{(1)}] \\
&= C, \qquad\qquad (= (n-1)/(n+1))
\end{aligned}$$

using the fact that $\{X_i - \theta\}_{i=1}^n$ are iid Uniform$(-1/2, 1/2)$ random variables. But we certainly do not have $h(T) = X_{(n)} - X_{(1)} = C$, P_θ almost surely! This is a contradiction, whence no complete sufficient statistic for $\theta \in \mathbb{R}$ exists.

3. We have $X_1, \ldots, X_n \overset{iid}{\sim} f_\theta$, with density $f(x;\theta) = e^{-(x-\theta)}\mathbf{1}(\theta \leq x)$. It follows that

$$f(\boldsymbol{x};\theta) = e^{-\left(\sum_{i=1}^n (x_i - \theta)\right)} \prod_{i=1}^n \mathbf{1}(\theta \leq x_i) = e^{-n(\bar{x}-\theta)}\mathbf{1}(\theta \leq x_{(1)}).$$

(a) We immediately have $T = X_{(1)}$ sufficient for $\theta \in \mathbb{R}$ via Neyman-Fisher factorization. Also,

$$\frac{f(\boldsymbol{x};\theta)}{f(\boldsymbol{y};\theta)} = \frac{e^{-n\bar{x}}\mathbf{1}(\theta \leq x_{(1)})}{e^{-n\bar{y}}\mathbf{1}(\theta \leq y_{(1)})}$$

becomes independent of θ if and only if $x_{(1)} = y_{(1)}$, whence $T = X_{(1)}$ is minimal sufficient for $\theta \in \mathbb{R}$.

Now, compute

$$\begin{aligned}
P_\theta(T \geq t) &= P_\theta(X_{(1)} \geq t) \\
&= P_\theta(X_{(1)} - \theta \geq t - \theta) \\
&= P_0(X_{(1)} \geq t - \theta) \\
&= \prod_{i=1}^n P_0(X_i \geq t - \theta) \\
&= \prod_{i=1}^n e^{-(t-\theta)} \\
&= e^{-n(t-\theta)}
\end{aligned}$$

for $t \geq \theta$. It follows that the density of $T = X_{(1)}$ is given by

$$f_T(t) = ne^{-n(t-\theta)}\mathbf{1}(\theta \leq t).$$

Suppose that g is measurable, and that $\mathbb{E}_\theta[g(T)] = 0$ for all $\theta \in \mathbb{R}$. This means that

$$\int_\theta^\infty g(t)\, e^{-nt} e^{n\theta}\, dt = 0, \qquad \text{for all } \theta \in \mathbb{R}.$$

Thus, for any $\alpha, \beta \in \mathbb{R}$, we have

$$\int_{(\alpha,\beta]} g(t)\, e^{-nt}\, dt = \left(\int_\alpha^\infty - \int_\beta^\infty\right) g(t)\, e^{-nt}\, dt = 0.$$

Indeed, the Monotone Convergence Theorem gives

$$\int_{\mathbb{R}} g(t)\, e^{-nt}\, dt = 0.$$

This is sufficient to show that $g = 0$ almost everywhere; the collection

$$\mathcal{E} = \left\{ E \in \mathcal{B}(\mathbb{R}) \colon \int_E g(t)\, e^{-nt}\, dt = 0 \right\}$$

is a λ-system containing the π-system of intervals of the form $(-\alpha, \beta]$ which generates the Borel σ-algebra $\mathcal{B}(\mathbb{R})$, hence must be equal to $\mathcal{B}(\mathbb{R})$ by Dynkin's Lemma. It follows that

$$\int_E g(t)\, e^{-nt}\, dt = 0$$

for all Borel sets E. From this, we may conclude that $g(t)\, e^{-nt} = 0$, hence $g = 0$ almost everywhere.

Remark: If f is $(\Omega, \mathcal{M}, \mu)$ measurable and

$$\int_E f\, d\mu = 0, \qquad \text{for all } E \in \mathcal{M},$$

then setting $E_k = \{\omega \in \Omega \colon f(\omega) > 1/k\}$ gives

$$0 = \int_{E_k} f\, d\mu \geq \int_{E_k} \frac{1}{k}\, d\mu = \frac{1}{k}\mu(E_k),$$

whence all $\mu(E_k) = 0$, so $\mu(\{\omega \colon f(\omega) > 0\} = 0$ via continuity from below. Similarly, $\mu(\{\omega \colon f(\omega) < 0\}) = 0$, whence $f = 0$, μ-almost everywhere.

Remark: We have shown that the signed Borel measure ν described by $d\nu(t) = g(t)\, e^{-nt}\, dt$ agrees with the zero measure on all intervals (θ, ∞), hence must be zero.

(b) Now suppose that our family is restricted to $\theta \in (-\infty, 0)$. Set

$$c_1 = \int_0^1 e^{-nt}\, dt > 0, \qquad c_2 = \int_1^\infty e^{-nt}\, dt > 0,$$

and

$$g(t) = c_2 \mathbf{1}_{(0,1]}(t) - c_1 \mathbf{1}_{(1,\infty)}(t).$$

Then, for $\theta < 0$,

$$\begin{aligned}
\mathbb{E}_\theta[g(T)] &= \int_\theta^\infty g(t)\, e^{-nt} e^{n\theta}\, dt \\
&= e^{n\theta} \int_0^1 c_2 e^{-nt}\, dt - e^{n\theta} \int_1^\infty c_1 e^{-nt}\, dt \\
&= e^{n\theta}\,(c_2 c_1 - c_1 c_2) \\
&= 0.
\end{aligned}$$

However, we do not have $g(T) = 0$, P_{-1}-everywhere! Thus, T cannot be complete for $\theta \in (-\infty, 0)$.

4. Consider an exponential family $\{P_\theta\}_{\theta \in \Theta}$ on $\mathcal{X}$, with densities

$$f(x; \theta) = \frac{dP_\theta}{d\nu} = e^{\eta(\theta) T(x) - B(\theta)} h(x),$$

where ν is σ-finite. Furthermore, let Θ be an open interval in $\mathbb{R}$, on which η, B are infinitely differentiable, $\eta' \neq 0$, and we have sufficient regularity to interchange differentiation with respect to θ and integration with respect to ν.

(a) Set $\tau(\theta) = \mathbb{E}_\theta[T(X)]$, and observe that

$$\int_{\mathcal{X}} e^{\eta(\theta)T(x)-B(\theta)} h(x)\, d\nu(x) = \int_{\mathcal{X}} f(x;\theta)\, d\nu(x) = 1$$

for all $\theta \in \Theta$ gives

$$\int_{\mathcal{X}} (\eta'(\theta)T(x) - B'(\theta)) e^{\eta(\theta)T(x)-B(\theta)} h(x)\, d\nu(x) = \frac{\partial}{\partial\theta} \int_{\mathcal{X}} f(x;\theta)\, d\nu(x) = 0.$$

Rearranging, we have

$$\eta'(\theta) \int_{\mathcal{X}} T(x)\, f(x;\theta)\, d\nu(x) = B'(\theta) \int_{\mathcal{X}} f(x;\theta)\, d\nu(x), \tag{$\star$}$$

which is precisely $\eta'(\theta)\mathbb{E}_\theta[T(X)] = B'(\theta)$, whence $\tau(\theta) = B'(\theta)/\eta'(\theta)$.

(b) Differentiating ($\star$) after dividing by $\eta'(\theta)$, we have

$$\int_{\mathcal{X}} T(x)(\eta'(\theta)T(x) - B'(\theta))\, f(x;\theta)\, d\nu(x) = \tau'(\theta),$$

which means that

$$\eta'(\theta)\mathbb{E}_\theta[(T(X))^2] - B'(\theta)\mathbb{E}_\theta[T(X)] = \tau'(\theta).$$

This simplifies to

$$\mathbb{E}_\theta[(T(X))^2] = \frac{\tau'(\theta)}{\eta'(\theta)} + (\tau(\theta))^2.$$

It follows that

$$\mathrm{Var}_\theta(T(X)) = \mathbb{E}_\theta[(T(X))^2] - (\mathbb{E}_\theta[T(X)])^2 = \frac{\tau'(\theta)}{(\eta'(\theta)}.$$

5. We will prove the Lehmann-Scheffé Theorem. Let U be unbiased for $g(\theta)$ in the family $\{P_\theta\}_{\theta\in\Theta}$, and let T be a complete sufficient statistic for $\theta \in \Theta$. Set

$$h(t) = \mathbb{E}_\theta\left[U \mid T = t\right],$$

whence $\delta = h(T)$ is a well-defined statistic; this is free of θ by sufficiency of T! Now, δ is unbiased for $g(\theta)$ since $\mathbb{E}_\theta\left[\mathbb{E}_\theta[U \mid T]\right] = \mathbb{E}_\theta[U] = g(\theta)$. Furthermore, the Blackwell-Rao Theorem guarantees that for any competing unbiased estimator S of $g(\theta)$, we have

$$\mathrm{Var}_\theta(\delta) \le \mathrm{Var}_\theta(S), \qquad \text{for all } \theta \in \Theta,$$

since these variances evaluate to precisely the risks under the convex squared error loss $\ell(\,\cdot\,,\theta) = (\,\cdot\, - g(\theta))^2$. Thus, δ is Uniformly Minimum Variance Unbiased (UMVU) for $g(\theta)$.

It remains to show that δ is unique. Let $\tilde{\delta} = \tilde{h}(T)$ also be unbiased for $g(\theta)$, whence

$$\mathbb{E}_\theta\left[(h - \tilde{h})(T)\right] = \mathbb{E}_\theta[\delta] - \mathbb{E}_\theta[\tilde{\delta}] = 0$$

via unbiasedness. Then, $(h - \tilde{h})(T) = 0$, i.e. $\delta = \tilde{\delta}$, P_θ almost surely, via completeness of T.

Remark: We have shown that any unbiased function of a complete sufficient statistic is UMVU.

6. Let $X_1, \ldots, X_n \overset{iid}{\sim} \text{Exp}(\theta)$ for $\theta > 0$, so $\boldsymbol{X}$ has density

$$f(\boldsymbol{x}; \theta) = \prod_{i=1}^{n} \theta e^{-\theta x_i} \mathbf{1}_{(0,\infty)}(x_i) = \theta^n e^{-n\theta\bar{x}} \mathbf{1}_{(0,\infty)}(x_{(1)}).$$

(a) Observe that we have a 1-parameter exponential family, with natural parameter $\eta(\theta) = -\theta$, and $T(\boldsymbol{X}) = n\bar{X} = \sum_{i=1}^n X_i$. Since our natural parameter space $(-\infty, 0)$ has non-empty interior, T is complete minimal sufficient for $\theta \in (0, \infty)$.

(b) Let $a > 0$, and let $\phi(\theta) = e^{-\theta a}$. Define

$$\hat{\phi}(T) = \left(1 - \frac{a}{T}\right)^{n-1} \mathbf{1}(T \geq a).$$

Since this is a function of the complete sufficient statistic T, we need only check that $\hat{\phi}$ is unbiased for ϕ for us to conclude that $\hat{\phi}$ is UMVU, via uniqueness in the Lehmann-Scheffé Theorem!

Note that $T \sim \Gamma(n, \theta)$; the characteristic function of an exponential random variable is simply

$$\varphi_{X_1}(t) = \mathbb{E}_\theta \left[e^{itX_1}\right] = \int_0^\infty \theta e^{-(\theta - it)x} \, dx = \frac{\theta}{\theta - it} = \left(1 - \frac{it}{\theta}\right)^{-1}.$$

Thus, the characteristic function of T satisfies

$$\varphi_T(t) = \varphi_{\sum_{i=1}^n X_i}(t) = \prod_{i=1}^{n} \varphi_{X_i}(t) = \left(1 - \frac{it}{\theta}\right)^{-n},$$

from which the claim follows. This means that T has density

$$f_T(t; \theta) = \frac{\theta^n}{\Gamma(n)} t^{n-1} e^{-\theta t} \mathbf{1}_{(0,\infty)}(t).$$

Finally,

$$\begin{aligned}
\mathbb{E}_\theta[\hat{\phi}] &= \frac{\theta^n}{\Gamma(n)} \int_a^\infty \left(1 - \frac{a}{t}\right)^{n-1} t^{n-1} e^{-\theta t} \, dt \\
&= \frac{\theta^n}{\Gamma(n)} \int_a^\infty (t - a)^{n-1} e^{-\theta t} \, dt \\
&= \frac{\theta^n}{\Gamma(n)} \int_0^\infty t^{n-1} e^{-\theta(t+a)} \, dt \\
&= \frac{\theta^n}{\Gamma(n)} \int_0^\infty t^{n-1} e^{-\theta t} \, dt \cdot e^{-\theta a} \\
&= e^{-\theta a} \\
&= \phi(\theta).
\end{aligned}$$