Econ 240A Econometrics **Fall 2018**

Problem Set 3 Solutions

Due date: Sept. 26, 2018 *Fengshi Niu*

1. **Sufficient statistic of Gamma random sample**
   Write the joint density in the canonical form of the exponential family

$$
\begin{aligned}
f_{\alpha,\beta}(x) &= \prod_{i=1}^{n}\left\{\frac{1}{\Gamma(\alpha)\beta^{\alpha}}x_i^{\alpha-1}\exp\left(-\frac{x_i}{\beta}\right)\mathbf{1}\{x_i \geq 0\}\right\} \\
&= \exp\left(-\frac{1}{\beta}\sum_{i=1}^{n}x_i + (\alpha-1)\sum_{i=1}^{n}\log(x_i) - n\log\left(\Gamma(\alpha)\beta^{\alpha}\right)\right)\prod_{i=1}^{n}\mathbf{1}\{x_i \geq 0\}.
\end{aligned}
$$

   Since $\sum_{i=1}^n X_i$ and $\sum_{i=1}^n \log(X_i)$ are linearly independent and the natural parameter space contains an open set $\{(-\frac{1}{\beta}, \alpha-1) : \alpha, \beta > 0\} \supset (-\infty, 0)\times(-1,\infty)$, $T(X) = (\sum_{i=1}^n X_i, \sum_{i=1}^n \log(X_i))$ is a complete sufficient statistic.

2. **Poisson random sample**

   (a) The joint pmf of X at $x \in \mathbb{N}^n$ is

$$
f_\lambda(x) = \prod_{i=1}^{n}\frac{\lambda^{x_i}}{x_i!}e^{-\lambda} = \underbrace{\exp\left(\log\lambda\sum_{i=1}^{n}x_i - n\lambda\right)}_{g(\lambda,T(x))}\underbrace{\frac{1}{\prod_{i=1}^{n}(x_i!)}}_{h(x)}.
$$

   By factorization theorem, $T(X) = \sum_{i=1}^n X_i$ is sufficient.

   (b) $\mathbb{E}_\lambda X_1 = \lambda$, so $\hat{\lambda}_{MM} = \bar{X}$. The log likelihood is

$$
l_X(\lambda) = \log f_\lambda(X) = \log\lambda T(X) - n\lambda + c(X)
$$

   which is concave and maximized over $\lambda \geq 0$ at $\hat{\lambda}_{ML} = \bar{X}$.

   (c) $\mathbb{E}_\lambda\bar{X} = \frac{1}{n}\sum_{i=1}^n \mathbb{E}_\lambda(X_i) = \lambda$.

   (d) The CRLB_λ is

$$
\begin{aligned}
\frac{1^2}{\mathbb{E}_\lambda\left(\left(\frac{\partial}{\partial\lambda}\log f_\lambda(X)\right)^2\right)} &= \frac{1}{-\mathbb{E}_\lambda\left(\frac{\partial^2}{\partial\lambda^2}\log f_\lambda(X)\right)} \\
&= \frac{1}{\mathbb{E}_\lambda\left(\frac{T(X)}{\lambda^2}\right)} \\
&= \frac{\lambda}{n}.
\end{aligned}
$$

   (e) $\hat{\lambda}_{MM}$ is UMVU, because it is unbiased and it achieves the CRLB_λ for all $\lambda > 0$

$$
\text{Var}_\lambda(\hat{\lambda}_{MM}) = \frac{\lambda}{n} = \text{CRLB}_\lambda.
$$

Remark. *Another way to prove $\hat{\lambda}_{MM}$ is UMVU is to show*

1. *It's unbiased.*
2. *It's a function of a complete sufficient statistic.*

3. Uniform location family

(a) Simplify the joint density

$$f_\theta(x) = \prod_{i=1}^n f_\theta(x_i) = \prod_{i=1}^n \frac{1}{\theta}\mathbf{1}(\theta - \frac{1}{2} \le x_i \le \theta + \frac{1}{2}) = \underbrace{\frac{1}{\theta^n}\mathbf{1}(x_{(1)} \ge \theta - \frac{1}{2})\mathbf{1}(x_{(n)} \le \theta + \frac{1}{2})}_{g(\theta, T(x))}.$$

By factorization theorem, $T(x) = \big(X_{(1)}, X_{(n)}\big)$ is sufficient. To show it's minimal, we just need to show for any x, y

$$f_\theta(x) \propto_\theta f_\theta(y) \Longrightarrow T(x) = T(y).$$

Here, $f_\theta(x) \propto_\theta f_\theta(y)$ means $\exists\ c(x,y)$ s.t. $f_\theta(x) = c(x,y) f_\theta(y),\ \forall \theta \in \Theta$. (I will use this notation throughout to show minimal sufficiency).

$$\begin{aligned}
&f_\theta(x) \propto_\theta f_\theta(y) \\
\Longleftrightarrow\ &\frac{1}{\theta^n}\mathbf{1}(x_{(1)} \ge \theta - \frac{1}{2})\mathbf{1}(x_{(n)} \le \theta + \frac{1}{2}) \propto_\theta \frac{1}{\theta^n}\mathbf{1}(y_{(1)} \ge \theta - \frac{1}{2})\mathbf{1}(y_{(n)} \le \theta + \frac{1}{2}) \\
\Longleftrightarrow\ &\mathbf{1}(x_{(n)} - \frac{1}{2} \le \theta \le x_{(1)} + \frac{1}{2}) \propto_\theta \mathbf{1}(y_{(n)} - \frac{1}{2} \le \theta \le y_{(1)} + \frac{1}{2}) \\
\Longrightarrow\ &x_{(1)} = y_{(1)},\ x_{(n)} = y_{(n)}.
\end{aligned}$$

Hence, T is minimal sufficient.

(b) To show T is not complete, we will find a nonzero function of T with expectation 0 for all $\theta \in \Theta = \mathbb{R}$. Consider a function h of $T(X)$,

$$h(T(X)) = X_{(n)} - X_{(1)} - \frac{n-1}{n+1}.$$

For any $\theta \in \mathbb{R}$, the expectation

$$\begin{aligned}
\mathbb{E}_\theta\Big(h(T(X))\Big) &= \mathbb{E}_\theta\big(X_{(n)} - X_{(1)} - \frac{n-1}{n+1}\big) = \mathbb{E}_\theta\big(X_{(n)} - \theta - (X_{(1)} - \theta) - \frac{n-1}{n+1}\big) \\
&= \mathbb{E}_0\big(X_{(n)} - X_{(1)} - \frac{n-1}{n+1}\big) = \frac{n}{n+1} - \frac{1}{n+1} - \frac{n-1}{n+1} \\
&= 0.
\end{aligned}$$

Obviously, $\mathbb{P}_\theta(h(T(X)) = 0) < 1$. Hence, T is not complete.

(c) The new estimator is

$$\delta(X) = \mathbb{E}(\bar{X} | T(X)).$$

The conditional distribution of X given $T(X)$ is

$$\begin{aligned}
f_{X|T}(x|t) &= \frac{f_{X,T}(x,t)}{f_T(t)} \\
&= \frac{f_X(x)\mathbf{1}(x_{(1)} = t_1)\mathbf{1}(x_{(n)} = t_n)}{f_T(t)} \\
&= \frac{1^n\mathbf{1}(x_{(1)} \geq \theta - \frac{1}{2})\mathbf{1}(x_{(n)} \leq \theta + \frac{1}{2})}{n(n-1)\cdot 1^2(t_2 - t_1)^{n-2}}\mathbf{1}(x_{(1)} = t_1)\mathbf{1}(x_{(n)} = t_n) \\
&= \frac{1}{n(n-1)(t_2 - t_1)^{n-2}}\mathbf{1}(x_{(1)} = t_1)\mathbf{1}(x_{(n)} = t_2)\mathbf{1}(t_1 \geq \theta - \frac{1}{2})\mathbf{1}(t_n \leq \theta + \frac{1}{2}) \\
&= \frac{1}{n(n-1)(t_2 - t_1)^{n-2}}\mathbf{1}(x_{(1)} = t_1)\mathbf{1}(x_{(n)} = t_2)
\end{aligned}$$

The last step follows because we are considering only those t in the support of T. Now we are ready to compute $\delta(X)$.

$$\delta(X) = \mathbb{E}(\bar{X}|T(X)) = \frac{X_{(1)} + X_{(n)}}{2}.$$

(d) *

(e) *

4. Minimal sufficiency of the likelihood ratio

(a) The likelihood function is $T(X) = (p_\theta(X))_{\theta\in\Theta} = (p_{\theta_0}(X), p_{\theta_1}(X), \ldots, p_{\theta_m}(X))$. Notice for any $x \in \mathcal{X}$,

$$p_\theta(x) = \underbrace{\Pi_\theta(T(x))}_{g(\theta,T(x))},$$

where $\Pi_\theta(t)$ is the projection of a vector t to its θ coordinate. One way to write it is $\Pi_\theta(t) = te_\theta,\ e_\theta := (\mathbf{1}(\theta_0 = \theta), \mathbf{1}(\theta_1 = \theta), \ldots, \mathbf{1}(\theta_m = \theta))'$. By factorization theorem, $T(X) = (p_\theta(X))_{\theta\in\Theta}$ is sufficient.
I will give an example illustrating the likelihood function is not minimal sufficient in general in (c).

(b) The likelihood ratio function is $T(x) = \left(\frac{p_\theta(x)}{p_{\theta_0}(x)}\right)_{\theta\in\Theta} = \left(\frac{p_{\theta_0}(x)}{p_{\theta_0}(x)} = 1, \frac{p_{\theta_1}(x)}{p_{\theta_0}(x)}, \ldots, \frac{p_{\theta_m}(x)}{p_{\theta_0}(x)}\right)$. Factor the density function as the product

$$p_\theta(x) = \frac{p_\theta(x)}{p_{\theta_0}(x)} \cdot p_{\theta_0}(x) = \underbrace{\Pi_\theta(T(x))}_{g(\theta,T(x))} \cdot \underbrace{p_{\theta_0}(x)}_{h(x)}.$$

By factorization theorem, $T(X)$ is sufficient. To show T is minimal sufficient, we just need to show for any $x, y \in \mathcal{X}$,

$$p_\theta(x) \propto_\theta p_\theta(y) \Rightarrow T(x) = T(y).$$

Here, as usual, $p_\theta(x) \propto_\theta p_\theta(y)$ means $\exists\ c(x,y)$ s.t. $p_\theta(x) = c(x,y)p_\theta(y),\ \forall \theta \in \Theta$. Notice

$$\begin{aligned}
&p_\theta(x) = c(x,y)p_\theta(y),\ \forall\theta \in \Theta \\
\implies &\frac{p_\theta(x)}{p_{\theta_0}(x)} = \frac{p_\theta(y)c(x,y)}{p_{\theta_0}(y)c(x,y)} = \frac{p_\theta(y)}{p_{\theta_0}(y)},\ \forall\theta\in\Theta \\
\implies &T(x) = \left(\frac{p_{\theta_0}(x)}{p_{\theta_0}(x)} = 1, \frac{p_{\theta_1}(x)}{p_{\theta_0}(x)}, \dots, \frac{p_{\theta_m}(x)}{p_{\theta_0}(x)}\right) = \left(\frac{p_{\theta_0}(y)}{p_{\theta_0}(y)} = 1, \frac{p_{\theta_1}(y)}{p_{\theta_0}(y)}, \dots, \frac{p_{\theta_m}(y)}{p_{\theta_0}(y)}\right) = T(y).
\end{aligned}$$

Hence, $T(X)$ is minimal sufficient.

(c) Here is an example showing likelihood function is not minimal sufficient in general:

$$\begin{aligned}
\mathcal{X} &= \{0,1,2\} && \mathcal{X} \text{ is a finite sample space} \\
p_\theta(0) = 1-3\theta,\ p_\theta(1) &= \theta,\ p_\theta(2) = 2\theta && p_\theta \text{ is a generic distribution} \\
\mathcal{P} &= \left\{p_\theta :\ \theta \in \left\{\frac{1}{4}, \frac{1}{3}\right\}\right\} && \mathcal{P} \text{ is the model.}
\end{aligned}$$

We can represent the distributions using the table:

	$x=0$	$x=1$	$x=2$
$p_{\frac{1}{4}}(x)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
$p_{\frac{1}{3}}(x)$	0	$\frac{1}{3}$	$\frac{2}{3}$.

In this case the likelihood function takes value

$$L(x) = \left(p_{\frac{1}{4}}(x), p_{\frac{1}{3}}(x)\right) = \begin{cases} (\frac{1}{4}, 0) & x = 0 \\ (\frac{1}{4}, \frac{1}{3}) & x = 1 \\ (\frac{1}{2}, \frac{2}{3}) & x = 2 \end{cases}.$$

The likelihood ratio function takes value

$$LR(x) = \left(\frac{p_{\frac{1}{4}}(x)}{p_{\frac{1}{4}}(x)} = 1, \frac{p_{\frac{1}{3}}(x)}{p_{\frac{1}{4}}(x)}\right) = \begin{cases} (1, 0) & x = 0 \\ (1, \frac{4}{3}) & x = 1 \\ (1, \frac{4}{3}) & x = 2 \end{cases}.$$

The likelihood function $L(X)$ is not a function of the likelihood ratio function $LR(X)$ in this case. In particular, $LR(1) = LR(2)$ but $L(1) \neq L(2)$ (also $P_\theta(X=1),\ P_\theta(X = 2) > 0$). Hence, the likelihood function is not minimal sufficient by definition. The idea is any two points whose probability (density) are the same up to a constant (for all θ) will necessarily take the same value in any minimal sufficient statistic. This doesn't hold for likelihood function in general, but does hold for likelihood ratio function.

Remark. *The sufficient statistics $L(X)$ and minimal sufficient statistic $LR(X)$ are random variables whose values are functions of θ (represented as "vectors" with Θ being the index set). In this sense, $L(X)$ and $LR(X)$ are* **random functions** *and we call them likelihood function and likelihood ratio function respectively. In general, when the parameter space Θ is* *either finite or infinite, we have random functions*

$$L(X) = g(\cdot, X) : \theta \mapsto p_\theta(X)$$
$$LR(X) = h(\cdot, X) : \theta \mapsto \frac{p_\theta(X)}{p_{\theta_0}(X)}$$

being sufficient and minimal sufficient respectively (under some regularity conditions).