Section 9: "Optimal" estimators

ARE 210

November 2, 2017

The section notes are available on the section Github at github.com/johnloeser/are210 in the "section9" folder.

1 Definitions

- **Statistics**
  - A function of the data $T(X)$ is a **statistic**
  - $X \sim P_{\theta_0}$, $\theta_0 \in \Theta$, $T(X)$ is **sufficient** for $\mathbf{P} = \{P_\theta : \theta \in \Theta\}$ (or sufficient for θ) if $P_\theta[X|T] = P[X|T]$
  - **Factorization theorem**: $\{f_\theta : \theta \in \Theta\}$, $T(X)$ statistic, T is sufficient $\Leftrightarrow$ $f(X, \theta) = g(T(X), \theta)h(X)$
  - **Exponential family of distributions**:
    $p(x, \theta) = h(x)\exp(\sum_{j=1}^{k} \eta_j(\theta)T_j(x) - \beta(\theta))$
  - S is **ancillary** for θ if the distribution of S does not depend on θ
  - T is **complete** for θ if $\forall$ g such that $\mathbf{E}_\theta[g(T)] = 0$ $\forall$ $\theta \in \Theta$, $g(T) = 0$ almost everywhere.
  - X_i iid, $p(X, \theta)$ from the exponential family, then
    $T(X_1, \ldots, X_n) = (\sum_{i=1}^{n} T_1(X_i), \ldots, \sum_{i=1}^{n} T_k(X_i))$ is complete for θ if
    $\{\eta_1(\theta), \ldots, \eta_k(\theta)\}$ contains an open set in $\mathbf{R}^k$. By the factorization theorem, $T(X_1, \ldots, X_n)$ is always sufficient for θ.
  - **Basu**: If T is complete and sufficient, then $T \perp$ every ancillary statistic.
- **"Optimal" statistics**
  - For estimators, unbiased = good, consistency = good, efficiency = good
  - $\phi(X)$ is **UMVUE** of $g(\theta)$ if 1) $\phi(X)$ is unbiased, and 2) $\forall$ $\delta(X)$ unbiased for $g(\theta)$, $V_\theta[\phi(X)] \leq V_\theta[\delta(X)]$ $\forall$ $\theta \in \Theta$
    * If $\phi(X)$ is UMVUE of $g(\theta)$, then it is the unique UMVUE of $g(\theta)$
  - **Rao-Blackwell**: Let $h(X)$ be any unbiased estimator of $g(\theta)$, $T(X)$ sufficient statistic for θ. Let $\phi(T) = \mathbf{E}_\theta[h(X)|T] = \mathbf{E}[h(X)|T]$, then $\phi(T)$ is also unbiased and has lower variance.
    * $\phi(T)$ only a statistic because T is sufficient

- $\phi(S(X))$ unbiased for $g(\theta)$, S is a complete statistic, then ϕ equals any other unbiased statistic that's a function of $S(X)$ almost everywhere
- **Lehmann-Scheffe**: $T(X)$ sufficient and complete for θ, $\phi(T(X))$ unbiased for $g(\theta)$, then ϕ is UMVUE for $g(\theta)$
- **Hausman principle**: Let $\mathbf{E}_\theta[W(X)] = \tau(\theta)$, then W is UMVUE of $\tau(\theta)$ if and only if W is uncorrelated with all unbiased estimators of $0 \; \forall \; \theta \in \Theta$
- Variance bounds
    * Below, assume $X \sim P_\theta$, $p(X, \theta)$ density, conditions (1) the support of p does not depend on θ, (2) $\frac{\partial p(x,\theta)}{\partial \theta}$ exists $\forall \; \theta$, for almost all x, and is finite, (3) if $T(X)$ is any statistic such that $\mathbf{E}_\theta[|T|] < \infty \; \forall \; \theta$, then $\frac{\partial}{\partial \theta} \int T(x)p(x,\theta)dx = \int T(x)\frac{\partial p(x,\theta)}{\partial \theta}dx$
    * Define $I(\theta) \equiv \mathbf{E}_\theta\left[\left(\frac{\partial}{\partial \theta}\log p(x,\theta)\right)\left(\frac{\partial}{\partial \theta}\log p(x,\theta)\right)'\right]$
    * **Information matrix equality**: $\mathbf{E}_{\theta^*}[\frac{\partial}{\partial \theta}\log p(x,\theta^*)] = 0$, $I(\theta^*) = V_{\theta^*}[\frac{\partial}{\partial \theta}\log p(x,\theta^*)]$, $I(\theta^*) = -\mathbf{E}_{\theta^*}[\frac{\partial^2}{\partial \theta \partial \theta'}\log p(x,\theta^*)]$
    * **Cramer-Rao inequality**: $T(X)$ real valued statistic, $\mathbf{E}_\theta[T(X)] = \phi(\theta)$, $I(\theta)$ non-singular. Then, $\forall \; \theta$, $V_\theta[T(X)] \geq \frac{d\phi(\theta)}{d\theta} I(\theta)^{-1}\frac{d\phi(\theta)}{d\theta}$
        · (CRI Lemma): Let Y be a scalar random variable, Z random vector, $\mathbf{E}[ZZ']^{-1}$ and $\mathbf{E}[ZY]$ exist, and either $\mathbf{E}[Z] = 0$ or Z contains a constant, then let $\beta(P) = \arg\min_c \mathbf{E}[(Y - Z'c)^2]$ and let $\mu_L(Z) = Z'\beta(P)$, then 1) $\beta(P) = \mathbf{E}[ZZ']^{-1}\mathbf{E}[ZY]$, and 2) $V[Y] \geq V[\mu_L(Z)] = \beta(P)'V[Z]\beta(P)$
        · Proof: (CRI Lemma) letting $Z = \frac{\partial \log p(x,\theta)}{\partial \theta}$ and $Y = T(X)$

2 Some useful bits

1. Unbiased estimators need not exist, when they do exist UMVUE need not exist
2. Main lessons here - OLS is great, IV with one endogenous variable and one instrument is great, finding minimally sufficient statistics = complete sufficient statistics for your model is great, writing down models with minimally sufficient statistics is great, MLE is great (when you're correctly specified-ish)
3. Exponential families have nice results
4. IME $\Rightarrow$ MLE equivalent to setting the score equal to 0, and the variance of the score is the negative Hessian of the log likelihood; we'll use this basically to get simpler expressions for the variance of estimators

3 Practice questions

1) Prove (CRI Lemma).

2) (**Identification in Exponential Families)** Consider the (canonical) exponential family $\left\{p\left(y,\theta\right):\theta\in\Theta\subset\mathbb{R}^{d}\right\}$ where

$$p\left(y,\theta\right)=h\left(y\right)\exp\left\{\sum_{k=1}^{d}\theta_{k}T_{k}\left(y\right)-A\left(\theta\right)\right\}$$

and suppose that $A\left(\theta\right)$ is continuously differentiable for all $\theta\in\Theta$. Show that θ is identified if the $d\times d$ matrix

$$I\left(\theta^{*}\right)=\mathbb{E}_{\theta^{*}}\left(\frac{d\log p\left(y,\theta^{*}\right)}{d\theta}\frac{d\log p\left(y,\theta^{*}\right)}{d\theta}'\right)$$

is non-singular for every $\theta^{*}\in\Theta$ (where $\frac{d\log p(y,\theta^{*})}{d\theta}=\left.\frac{d\log p(y,\theta)}{d\theta}\right|_{\theta=\theta^{*}}$)

3) Prove the Hausman principle.