Theoretical Statistics, STOR 655

Asymptotic Efficiency

Andrew Nobel

March 2023

Asymptotic Efficiency

Setting

- Family $\mathcal{P} = \{f(x|\theta) : \theta \in \Theta\}$ of densities on $(\mathcal{X}, \mathcal{A})$
- Observations $X_1, X_2, \ldots \in \mathcal{X}$ iid with $X_i \sim f(x|\theta_0)$

Definition: A sequence of estimators $\tilde{\theta}_n$, $n \geq 1$, is *asymptotically efficient* if for every $\theta_0 \in \Theta$

$$n^{1/2}(\tilde{\theta}_n(X_1^n) - \theta_0) \Rightarrow \mathcal{N}_p(0, I(\theta_0)^{-1})$$

Example: MLEs are asymptotically efficient under suitable conditions

Asymptotic Efficiency

Note: If $\tilde{\theta}_1, \tilde{\theta}_2, \ldots$ is asymptotically efficient then $\tilde{\theta}_n \approx \mathcal{N}_p(\theta_0, (nI(\theta_0))^{-1})$

- $\tilde{\theta}_n$ is approximately unbiased
- $\mathrm{Var}_{\theta_0}(\tilde{\theta}_n)$ is approximately the CR lower bound

Can one do better, i.e., achieve smaller asymptotic variance?

Example: (Hodges) Consider a one-dimensional problem for which MLEs $\hat{\theta}_n$ are asymptotically efficient. Fix $\theta^* \in \Theta$ and define

$$\tilde{\theta}_n = \begin{cases} \theta^* & \text{if } n^{1/4}|\hat{\theta}_n - \theta^*| \leq 1 \\ \hat{\theta}_n & \text{otherwise} \end{cases}$$

Can show $n^{1/2}(\tilde{\theta}_n - \theta_0) \Rightarrow \mathcal{N}(0, \sigma^2(\theta))$ where $\sigma^2(\theta) = I(\theta_0)^{-1}\mathbb{I}(\theta \neq \theta^*)$.
But one can only improve $\hat{\theta}_n$ this way on a set of Lebesgue measure zero

Newton's Method

Idea: Iterative procedure to find a local maxima of a smooth function h

Setting

- Function $h : A \to \mathbb{R}$ of interest
- $A \subseteq \mathbb{R}^d$ open
- h is twice continuously differentiable
- $\ddot{h}(x) = \nabla^2 h(x) \in \mathbb{R}^{d \times d}$ non-singular for every $x \in A$

Goal: Find a local maximum of h

Newton's Method

0. (Initialize) Set $k = 0$ and fix $x_0 \in A$

1. Form a quadratic approximation to h at the point x_k

$$q_k(x) \;:=\; h(x_k) + (x - x_k)^t \, \dot{h}(x_k)^t + \frac{1}{2}(x - x_k)^t \, \ddot{h}(x_k)(x - x_k)$$

2. Maximize $q_k(x)$ by solving the equation

$$\dot{q}_k(x) = \dot{h}(x_k)^t + \ddot{h}(x_k)(x - x_k) \;=\; 0$$

   Solution $x_{k+1} := x_k - \ddot{h}(x_k)^{-1} \, \dot{h}(x_k)^t$ updates previous guess

3. If $||x_k - x_{k+1}|| \leq \tau$ then stop. Otherwise, increment $k := k + 1$ and return to step 1

Application: Improving Sub-Efficient Estimates

Setting: $X_1, X_2, \ldots \in \mathcal{X}$ iid with $X_i \sim f(x|\theta_0) \in \mathcal{P} = \{f(x|\theta) : \theta \in \Theta\}$

- Assume $\dot{\ell}_n(\theta)$ and $\ddot{\ell}_n(\theta)$ exist and are continuous on Θ
- Assume $\ddot{\ell}_n(\theta) < 0$ for every θ

Given: Sequence of estimators $\tilde{\theta}_n : \mathcal{X}^n \to \Theta$, $n \geq 1$

Idea: For each $n \geq 1$ apply one step of Newton's method to the estimate $\tilde{\theta}_n = \tilde{\theta}_n(X_1^n)$ in order to produce an improved estimate

$$\tilde{\theta}_n^{(1)} = \tilde{\theta}_n - \ddot{\ell}_n(\tilde{\theta}_n)^{-1} \dot{\ell}_n(\tilde{\theta}_n)^t$$

Application: Improving Sub-Efficient Estimates

Theorem: Suppose that regularity conditions for asymptotic normality of MLE hold at θ_0 and that

1. $\tilde{\theta}_n = \tilde{\theta}_n(X_1^n)$ is strongly consistent for θ_0
2. $n^{1/2}(\tilde{\theta}_n - \theta_0) = O_p(1)$

Then the modified estimates

$$\tilde{\theta}_n^{(1)} = \tilde{\theta}_n - \ddot{\ell}_n(\tilde{\theta}_n)^{-1}\dot{\ell}_n(\tilde{\theta}_n)^t$$

are asymptotically equivalent to any sequence $\hat{\theta}_n$ such that $\dot{\ell}_n(\hat{\theta}_n) = 0$ and $\hat{\theta}_n \to \theta_0$ wp1, in the sense that $n^{1/2}(\tilde{\theta}_n^{(1)} - \hat{\theta}_n) = o_p(1)$

Application: Improving Sub-Efficient Estimates

Cor: The modified estimates $\tilde{\theta}_n^{(1)}$ are asymptotically efficient

$$n^{1/2}(\tilde{\theta}_n^{(1)} - \theta_0) \Rightarrow \mathcal{N}_p(0, I(\theta_0)^{-1})$$