lec09Claros

Andrew Claros

Monday 5th October, 2020

Relative efficiency:

- define "Relative efficiency" (RE) as the ratio of variances:\
- $RE = \frac{\mathbb{V}\text{ar}\left[\hat{\theta}^{mm}\right]}{\mathbb{V}\text{ar}\left[\hat{\theta}^{MLE}\right]} = \frac{\frac{1}{3n}}{\frac{n}{(n+1)^2(n+2)}} = \frac{(n+1)^2(n+2)}{3n^2} > 1 \Rightarrow$ MLE is better as measured by variance.
- Maybe we should be comparing the ration of MSE's? But in this case the tiny amount of bias in the MLE won't matter if n is large.

Important questions:

- Is there an theoretical minimum MSE (best) when estimating theta for a given DGP?
- if 1 is true then, for any DGP/theta is there a procedure of locating that estimator with the best MSE?
- No for both (334, C&B) because the class of all estimators is too big.

For example:

- Let $X_1, \ldots, X_n \overset{iid}{\sim} \text{Bernoulli}\,(\theta)$
- $\hat{\theta}_{Bad} = \frac{1}{2}, MSE[\hat{\theta}_{Bad}](\theta = \frac{1}{2}) = \mathbb{E}\,[[]\,(\hat{\theta}_{Bad} - \theta)^2] = \mathbb{E}\,[(]\,\frac{1}{2} - \frac{1}{2})^2 = 0$
- Means : $\hat{\hat{\hat{\theta}}}_{Bad}$ does amazing well at $\theta = \frac{1}{2}$
- There will be a "counterexample" estimator that does amazingly well for some values of θ and very badly for other values.

For only unbiased estimatorss:

- Define: minimum variance unbiased estimator (UMVUE) is the estimator is the estimator that $\hat{\theta}^*$ s.t. $\forall\hat{\theta}$ and all other unbiased estimators $\hat{\theta}$:
- $$\mathbb{V}\text{ar}\left[\hat{\theta}^*\right] \leq \mathbb{V}\text{ar}\left[\hat{\theta}\right]$$
- Is there an theoretical minimum MSE (best) when estimating theta for a given DGP? Yes. Cramer-Rao Lower Bouund (CRLB)
- Is there a procedure for locating the UMVUE? Sometimes

Camer-Rao Lower Bound proof:

Statement

- Let : $X_1, \ldots, X_n \overset{iid}{\sim} DGP[\theta]$ continuous
- $DGP \overset{iid}{\sim}$ normal: $\mathbb{V}\text{ar}\left[\bar{X}\right] = \frac{\sigma^2}{n}$
- For any unbiased estimator $\hat{\theta}, \mathbb{V}\text{ar}\left[\hat{\theta}\right] \geq \frac{I(\theta)^{-1}}{n}$
- Numerator is an irreducible core quantity based on DGP and θ
- $I(\theta) := \mathbb{E}\left[\frac{d}{dx}\ell\left(\theta : x\right)^2\right]$ "Expectation of the squared log-likelihood" called "Fisher Information" defined by Fisher in 1922

Proof

- Cauchy-Schwartz Inequality for any two r.v.'s Q and S $\mathbb{C}\text{ov}\left[Q, S\right]^2 \leq \mathbb{V}\text{ar}\left[Q\right]\mathbb{V}\text{ar}\left[S\right]$
- $\Rightarrow \mathbb{V}\text{ar}\left[Q\right] >= \frac{\mathbb{C}\text{ov}[Q,S]^2}{\mathbb{V}\text{ar}[S]}$
- $\mathbb{C}\text{ov}\left[Q, S\right] = \frac{(\mathbb{E}[RS]-\mathbb{E}[Q]\mathbb{E}[S])^2}{\mathbb{E}[S^2]-\mathbb{E}[S]^2}$
- Let $Q = \hat{\theta} \Rightarrow \mathbb{E}\left[\hat{\theta}\right] = \theta$ due to unbiasedness
- Define the "score function"
- Def 1: $S := \frac{\partial}{\partial\theta}[ln(f(X_1, \ldots, X_n : \theta))]$
- Def2 $\frac{\frac{\partial}{\partial}[f(X_1,\ldots,X_n:\theta)]}{f(X_1,\ldots,X_n:\theta)}$
- Def 3, 4, 5: $\frac{\partial}{\partial\theta}[ln(\prod f(X_i : \theta))] = \frac{\partial}{\partial}[\sum ln(f(X_i : \theta))] = \sum \frac{\partial}{\partial}[ln(ln(X_i : \theta))]$
- $\mathcal{L} = f, l := ln(\mathcal{L}) = ln(f)$
- $\frac{\partial}{\partial\theta}[l(\theta : X_1, \ldots, X_n)] = l'(\theta : X_1, \ldots, X_n) = \sum l'(\theta : X_i)$
- S and X_i's is a r.v (capital)
- Need to find: $\mathbb{E}\left[\hat{\theta}\right], \mathbb{E}\left[S^2\right], \mathbb{E}\left[S\right]$
- $\mathbb{E}\left[S\right] = \mathbb{E}\left[\frac{\frac{\partial}{\partial\theta}[f(X_1,\ldots,X_n:\theta)]}{f(X_1,\ldots,X_n:\theta)}\right] = \int \ldots^n \int \frac{\frac{\partial}{\partial}[f(X_1,\ldots,X_n:\theta)]}{f(X_1,\ldots,X_n:\theta)} f(X_1, \ldots, X_n : \theta)dx_{1n}$
- $\frac{\partial}{\partial\theta}[\int \ldots \int f(X_1, \ldots, X_n : \theta)dx_{1n}] = \frac{\partial}{\theta}[1] = 0$
- $\mathbb{E}\left[S\right] = \mathbb{E}\left[l'(\theta : X_1, \ldots, X_n)\right] = 0$
- $\mathbb{E}\left[S\right] = \mathbb{E}\left[\sum l'(\theta : X_i)\right] \overset{iid}{\sim} n\mathbb{E}\left[l'(\theta : X_i)\right] = 0 = 0 \Rightarrow l'(\theta : X_i) = 0$
- $\mathbb{V}\text{ar}\left[S\right] = \mathbb{E}\left[S^2\right] - \mathbb{E}\left[S\right]_{=0} = \mathbb{E}\left[(\sum l'(\theta : X_i))^2\right]_{=(\sum_1^n(a_i))^2=\sum_i^n a_i^2+\sum_{i\neq j} a_i a_j}$
- $\sum \mathbb{E}\left[l'(\theta : X_i)\right]^2 + \sum \mathbb{E}\left[l'(\theta : X_i)l'(\theta : X_j)\right]$