APMA 1655 Notes: Chapter 9

MVUE: Minimal Variance Unbiased Estimator

- Definition: MVUE is the unbiased estimator among all unbiased estimators of population parameter θ with the minimal variance
- Statistic: Any function of samples $\{X_1, ..., X_n\}$
- Likelihood $L_\theta(x_1, ..., x_n)$: defined to be joint probability distribution of iid samples $\{X_1, ..., X_n\}$
  - Discrete samples: assume the probability mass function of X_i is $p_\theta(x)$, then:
  $$L_\theta(x_1, ..., x_n) = \Pi_{i=1}^n p_\theta(x_i)$$
  - Continuous samples: assume the probability density function of X_i is $f_\theta(x)$

Sufficient Statistic

- Definition: Let $t = (t_1, ..., t_m)$ be a collection of statistics, that is, $t_i = t_i(X_1, ..., X_n)$ for each i
- t is sufficient for estimating θ if $L_\theta(x_1, ..., x_n) = g_\theta(t)h(x_1, ..., x_n)$ where $g_\theta(t)$ is a function of θ and t only, and h doesn't depend on θ
  - Bernoulli Distribution: Let $\{X_1, ..., X_n\}$ be iid Bernoulli random variables with $P(X_i = 1) = p = 1 - P(X_i = 0)$; then
  $$t = \Sigma_{i=1}^n X_i$$
  is sufficient for estimating p
  - Poisson Distribution: Let $\{X_1, ..., X_n\}$ be iid Poisson random variables with parameter λ; then
  $$t = \Sigma_{i=1}^n X_i$$
  is sufficient for estimating λ
  - Uniform Distribution: Let $\{X_1, ..., X_n\}$ be iid random variables uniformly distribution on $[0, \theta]$; then
  $$t = max\{X_1, ..., X_n\}$$
  is sufficient for estimating θ
  - Normal Distribution: Let $\{X_1, ..., X_n\}$ be iid samples from $N(\mu, \sigma^2)$ where μ and σ are both unknown; then:
  $$t = (\Sigma_{i=1}^n X_i, \Sigma_{i=1}^n X_i^2)$$
  is sufficient for estimating μ and σ^2

Meaning of Sufficiency

- The conditional distribution of $\{X_1, ..., X_n\}$ given a sufficient statistic $t = t(X_1, ..., X_n)$ does not depend on θ. In other words, a sufficient statistic contains all the information of θ from samples

Variance Decomposition

- Let X and Y be any random variables. Define $\hat{X} = E[X|Y]$. Then
$$E[X] = E[\hat{X}]$$
$$Var[\hat{X}] \leq Var[X]$$

Rao-Blackwell Theorem

- Let $\hat{\theta} = \hat{\theta}(X_1, ..., X_n)$ be an unbiased estimator for θ, and t any sufficient statistic; Define
$$\hat{\theta}^* = E[\hat{\theta}(X_1, ..., X_n)|t]$$
Then $\hat{\theta}^*$ is an unbiased estimator for θ and Var$[\hat{\theta}^*] \leq$ Var$[\hat{\theta}]$

Completeness

- Definition: A statistic t is said to be complete if $E[h(t)] = 0$ for every θ implies $h(t) = 0$
- given a complete statistic t, there is at most one function of t that is an unbiased estimator for θ. If t is also sufficient, then this unbiased estimator has to be MVUE

MVUE via Sufficient and Complete Statistics

- Classic Approach to MVUE
  1. Find a sufficient statistic, say t
  2. Argue that this statistic is complete (most mathematically difficult)
  3. Find a function of t, say $h(t)$ such that $h(t)$ is an unbiased estimator for θ: to find functin h, theoretically can pick any unbiased estimator $\hat{\theta}$ and define $h(t) = E[\hat{\theta}|t]$
  4. This estimator h(t) is MVUE

Criterion for Sufficient and Complete Statistics

- Let $\{X_1, ..., X_n\}$ be iid sample
- Assume that the likelihood function $L_\theta(X)$ with $x = (x_1, ..., x_n)$ can be expressed as
$$L_\theta(x) = f(\theta)h(x)exp\{\Sigma_{i=1}^k w_i(\theta)t_i(x)\}$$
- Sufficient: define $t(x) = (t_1(x), ..., t_k(x))$, then t is sufficient
- Complete: if $[w_1(\theta), ..., w_k(\theta)]$: all possible$\theta$ contains an open set in $\mathbb{R}^k$, then t is also complete

MLE: Maximum Likelihood Estimate

- Definition: Let $\{X_1, ..., X_n\}$ be iid samples with likelihood function $L_\theta(x_1, ..., x_n)$; Let $\hat{\theta} = \theta thatmaximizes L_\theta(X_1, ..., X_n)$; then $\hat{\theta}$ is said to be the MLE of θ
- often more convenient to maximize log-likelihood log $L_\theta(x)$ instead of just L
- MLE is always a function of any sufficient statistic

MLE for Poisson

- Let $\{X_1, ..., X_n\}$ be iid samples from Poisson distribution with parameter λ, find the MLE for λ
- Likelihood:

$$L_\theta(x_1, ..., x_n) = \Sigma_{i=1}^n e^{-\lambda} \frac{\lambda^{x_i}}{x_i!}$$

- Maximizing likelihood amounts to maximizing:

$$-n\lambda + log\lambda\Sigma_{i=1}^n x_i$$

- maximizer is $\hat{x}$ and thus $\hat{\lambda} = \hat{X}$ is MLE for λ

MLE for Uniform

- Let $\{X_1, ..., X_n\}$ be iid samples from uniform distribution on $[0, \theta]$, find the MLE for θ
- the Likelihood is:

$$L_\theta(x_1, ..., x_n) =$$

$$\begin{cases} 1/\theta^n & \text{if } \theta \geq \max\{x_1, ..., x_n\} \\ 0 & \text{otherwise} \end{cases}$$

- maximizing likelihood amounts to finding the minimum of θ under the constraint $\theta \geq \max\{x_1, ..., x_n\}$
- $\hat{\theta} = \max\{X_1, ..., X_n\}$ is MLE for θ

MLE for Normal

- Suppose $\{X_1, ..., X_n\}$ are iid samples from $N(\mu, \sigma^2)$; want the MLE for μ and σ^2
- Likelihood is:

$$L_\theta(x_1, ..., x_n) = (\frac{1}{\sqrt{2\pi}\sigma})^n exp\{-\Sigma_{i=1}^n \frac{(x_i - \mu)^2}{2\sigma^2}\}$$

- maximizing the likelihood amounts to maximizing:

$$-n\log\sigma - \Sigma_{i=1}^n \frac{(x_i - \mu)^2}{2\sigma^2}$$

- take derivatives of μ and σ and set them to zero, the MLEs are then:

$$\hat{\mu} = \bar{X}, \hat{\sigma}^2 = \frac{1}{n}\Sigma_{i=1}^n (X_i - \bar{X})^2$$

Properties of MLE

- MLE is consistent: Let θ^* be the true value, then $\hat{\theta} \to \theta^*$ as sample size n increases
- Assume $\{X_1, ..., X_n\}$ are iid samples from density $f_\theta^*(x)$; MLE maximizes the log-likelihood:

$$L_n(\theta) = \frac{1}{n}\Sigma_{i=1}^n \log f_\theta(X_i)$$

- MLE is asymptotically normal distributed with mean θ^* and variance $[nI(\theta^*)]^{-1}$, where

$$I(\theta^*) = E[-\frac{\partial^2}{\partial\theta^2}\log f_\theta(x)|_{\theta=\theta^*}]$$

- $I(\theta^*)$ is called Fisher information