PH 712 Probability and Statistical Inference

Part V: Point Estimation I

Zhiyang Zhou (zhou67@uwm.edu, zhiyanggeezhou.github.io)

2024/Oct/08 00:29:58

Recall types of the stochastic model

- Stochastic model: the distribution of RVs of interest
  - Parametric model
  - Non-parametric model
  - Semi-parametric model

Parametric model

- A set of pdfs/pmfs indexed by p-dimensional unknown $\boldsymbol{\theta}$ (constrained in $\boldsymbol{\Theta}$) with small or moderate dimension p, i.e., $\{f(\cdot \mid \boldsymbol{\theta}) : \boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^p\}$ with
  - $f(\cdot \mid \boldsymbol{\theta})$: either a pdf or a pmf
  - $\boldsymbol{\Theta}$: the set of allowed values of $\boldsymbol{\theta}$
- True parameters, say $\boldsymbol{\theta}_0$, believed to be fixed (frequentist statistics)
  - Rather than randomizing $\boldsymbol{\theta}_0$ (Bayesian statistics)
- Estimator: a statistic (i.e., a function of the sample); a guess about $\boldsymbol{\theta}_0$
- Estimate: plugging the realization into the estimator
- $p = 1$ hereafter, i.e., considering only one unknown parameter

Method of moments (MM, CB Sec 7.2.1)

- Procedure
  1. Equate RAW moments ($\mathrm{E}(X_i^k)$) to their empirical counterparts ($n^{-1} \sum_{i=1}^n X_i^k$).
  2. Solve the resulting simultaneous equations for θ.
- Pros and cons
  - Easy implementation
  - Start point for more complex methods
  - No constraint
  - Not uniquely defined

Example Lec5.1

- Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} f(x \mid \theta) = \theta x^{\theta-1} \mathbf{1}_{[0,1]}(x)$, $\theta > 0$.
  a. Find an MM estimator of θ.
  b. Can we employ the second (raw) moment instead of the first one?

Maximum Likelihood (ML) Estimator (MLE, CB Sec 7.2.2)

- Likelihood:
  - a real-valued function of unknown θ

$$L(\theta) = L(\theta; x_1, \ldots, x_n) = f_{X_1,\ldots,X_n}(x_1, \ldots, x_n \mid \theta), \quad \theta \in \Theta$$

  - $f_{X_1,\ldots,X_n}$: the joint pdf or pmf of $X_1, \ldots, X_n$
- Log-likelihood: the natural logarithm of likelihood

$$\ell(\theta) = \ell(\theta; x_1, \ldots, x_n) = \ln L(\theta; x_1, \ldots, x_n), \quad \theta \in \Theta$$

- If $\hat{\theta}(x_1, \ldots, x_n)$ is the maximizer of $L(\theta)$ (equiv. the maximizer of $\ell(\theta)$) with respect to θ constrained in Θ, i.e.,

$$\hat{\theta}(x_1, \ldots, x_n) = \arg\max_{\theta \in \Theta} L(\theta) = \arg\max_{\theta \in \Theta} \ell(\theta),$$

  then the statistic $\hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$ is the MLE for θ.
- Invariance (CB Thm 7.2.10): if $\hat{\theta}$ is the MLE of θ, then $g(\hat{\theta})$ is the MLE of $g(\theta)$ for any given function $g(\cdot)$.

How to locate the MLE constrained in Θ?

- If $\ell(\theta)$ is monotonic with respect to $\theta \in \Theta$, then the MLE lies at one boundary point of Θ
- If $\ell(\theta)$ is non-monotonic but differentiable with respect to $\theta \in \Theta$, then
  1. Collect all the candidates:
     - Stationary points, i.e., solutions to the equation $S(\theta) = 0$ subject to $\theta \in \Theta$
       * $S(\theta) = S(\theta; x_1, \ldots, x_n) = \ell'(\theta; x_1, \ldots, x_n)$: the score
     - Boundary points of Θ
  2. Compare the values of log-likelihood or likelihood evaluated at all the above candidates

Example Lec5.2

- Suppose $X_1, \ldots, X_n$ are iid as the following distributions. Find the MLEs for corresponding parameters.
  a. $\mathcal{N}(\mu, \sigma^2)$, with unknown $\mu \in \mathbb{R}$ and known $\sigma \in \mathbb{R}^+$.
  b. $\mathcal{N}(\mu, \sigma^2)$, with known $\mu \in \mathbb{R}$ and unknown $\sigma \in \mathbb{R}^+$.
  c. $Bern(\theta)$: $p(x \mid \theta) = \theta^x (1-\theta)^{1-x} \mathbf{1}_{\{0,1\}}(x)$, $\theta \in [0, 1/2]$.
  d. Exponential: $f(x \mid \beta) = \beta^{-1} \exp(-x/\beta) \mathbf{1}_{(0,\infty)}(x)$, $\beta \in \mathbb{R}^+$.