PH 712 Probability and Statistical Inference

Part IV: Point Estimation I

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

2025/Oct/20 11:40:51

Recall the framework of statistical inference/learning

- Goal: infer/learn the distribution of RV X, say f_X, from a random sample $X_1, \dots, X_n$
- Assumption: $f_X \approx \hat{f}_X$ (statistical model)
  - E.g., $\hat{f}_X = \mathcal{N}(\mu, \sigma^2)$, reducing the task to estimating (μ, σ)
- Point estimation: make the "best" guess about unknown parameter(s)
  - E.g., estimate (μ, σ) by $(\hat{\mu}, \hat{\sigma})$
- Hypothesis testing
  - E.g., confirm whether $\mu = 0$ by testing $H_0 : \mu = 0$ vs. $H_1 : \mu \neq 0$
- Interval estimation: construct an interval likely to cover the unknown parameter
  - E.g., construct an interval, say (c_1, c_2), such that $c_1 < \mu < c_2$ with a high probability

Point estimation

- θ: the unknown parameter
  - A unknown scalar (i.e., we only consider cases with one unknown parameter)
- The generation of a guess on the value of θ based on the random sample $X_1, \dots, X_n$
- Estimator: the generated guess, say $\hat{\theta}$
  - A statistic (why?) and hence an RV
  - E.g., sometimes, $\hat{\theta} = \bar{X} = n^{-1} \sum_{i=1}^n X_i$ (sample mean) is an estimator of certain parameter θ
- Estimate: plugging the realization of the random sample, say $x_1, \dots, x_n$, into the estimator
  - A number (why?) and NOT randomized
  - E.g., $\hat{\theta} = n^{-1} \sum_{i=1}^n x_i$ is an estimate of certain parameter θ

Maximum Likelihood (ML) Estimator (MLE)

- Θ: the set of allowed values of θ
- Likelihood function: an alias of joint pdf/pmf

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

  - $f_{X_1,\dots,X_n}$: the joint pdf/pmf of $X_1, \dots, X_n$

- Log-likelihood function: the natural logarithm of likelihood function

$$\ell(\theta) = \ln L(\theta), \quad \theta \in \Theta$$

- $\hat{\theta}_{\mathrm{ML}}$ is the MLE for θ if $\hat{\theta}_{\mathrm{ML}}$ is the maximizer of $L(\theta)$ (equiv. the maximizer of $\ell(\theta)$) with respect to θ constrained in Θ
  - In the math notation,
$$\hat{\theta}_{\mathrm{ML}} = \arg\max_{\theta\in\Theta} L(\theta) = \arg\max_{\theta\in\Theta} \ell(\theta)$$
  - That is to say, $L(\hat{\theta}_{\mathrm{ML}}) \geq L(\theta)$ and $\ell(\hat{\theta}_{\mathrm{ML}}) \geq \ell(\theta)$, for all $\theta \in \Theta$.
- Invariance property of MLE: if $\hat{\theta}_{\mathrm{ML}}$ is the MLE of θ, then $g(\hat{\theta}_{\mathrm{ML}})$ is the MLE of $g(\theta)$ for any given function $g(\cdot)$.

How to locate the MLE constrained in Θ?

- Theoretical way:
  - If $L(\theta)$ (or equiv. $\ell(\theta)$) 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 including:
       * Stationary points, i.e., solutions to the equation $S(\theta) = 0$ subject to $\theta \in \Theta$
         · Where $S(\theta) = \ell'(\theta)$ is called the score/gradient
       * Boundary points of Θ
    2. Compare the values of log-likelihood or likelihood evaluated at all the above candidates
- Numerical way:
  - Merely working when realizations of $X_1, \ldots, X_n$ are given
  - `R` function `optim()`

Example Lec5.1

- Suppose $X_1, \ldots, X_n$ is an iid sample following $\mathcal{N}(\mu, \sigma^2)$, i.e., $f_{X_i}(x \mid \theta) = (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$, $x \in \mathbb{R}$, with unknown μ and known $\sigma > 0$. Find the MLE of μ.

- Suppose $X_1, \ldots, X_n$ is an iid sample following $\mathcal{N}(\mu, \sigma^2)$, i.e., $f_{X_i}(x \mid \theta) = (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$, $x \in \mathbb{R}$, with known μ and unknown $\sigma > 0$. Find the MLE of σ.

- Suppose $X_1, \ldots, X_n$ is an iid sample following $p_{X_i}(x \mid \theta) = \theta^x(1-\theta)^{1-x}\mathbf{1}_{\{0,1\}}(x)$, $\theta \in [0, 1/2]$. Find the MLE of θ.

- Suppose $X_1, \ldots, X_n$ is an iid sample following an exponential distribution, i.e., $f_X(x \mid \beta) = \beta^{-1}\exp(-x/\beta)\mathbf{1}_{(0,\infty)}(x)$, $\beta > 0$. Find the MLE of θ.

- Suppose $X_1, \ldots, X_n$ is an iid sample following a beta distribution, i.e., $f_X(x \mid \theta) = \theta x^{\theta-1}\mathbf{1}_{[0,1]}(x)$, $\theta > 0$. Find the MLE of θ.