Maximum Likelihood (ML) Estimation

Consider a parametric family of pdf's: $\{p(\,\cdot\,;\theta) : \theta \in \Theta\}$. For fixed data y, $p(y;\,\cdot\,)$ is called the *likelihood function*

Definition: The ML estimator (MLE) of θ is

$$\boxed{\hat{\theta}_{ML}(y) := \arg\max_{\theta\in\Theta} p(y;\theta)}$$

In general, $\hat{\theta}_{ML}$ is biased, but it has optimal asymptotic properties

PROPERTIES OF ML

1. Principle of (Functional) Invariance

If $\hat{\theta}$ is the MLE of $\theta \in \Theta \subseteq \mathbb{R}^p$, then $f(\hat{\theta})$ is the MLE of $f(\theta)$, where $f : \Theta \to Z \subseteq \mathbb{R}^n$, with $n \le p$

2. Relation to Cramér-Rao (C-R) Bound

If there is an unbiased estimator that achieves the C-R bound, it is also the MLE

Let $\hat{\theta}_N$ be the MLE of θ based on N i.i.d. samples $y = \{x_1, \ldots, x_N\}$

3. Consistency

$$\boxed{\hat{\theta}_N \xrightarrow[N\to\infty]{a.s.} \theta}$$

4. Asymptotic normality

$$\boxed{\sqrt{N}(\hat{\theta}_N - \theta) \xrightarrow[N\to\infty]{d} \mathcal{N}(0, \bar{M}_\theta^{-1})}, \qquad \bar{M}_\theta := \frac{1}{N} E\left\{ \frac{\partial \ln p(y;\theta)}{\partial \theta} \frac{\partial \ln p(y;\theta)}{\partial \theta^T} \right\}$$