General topic: what happens as $n \to \infty$?

Goals:

1) Desirable characteristics:

- Consistency: $\hat{\Theta}$ is consistent for Θ if as $n \to \infty$, $\hat{\Theta}$ is close to the true parameter value Θ_0 with high probability

$$\lim_{n\to\infty} P(|\hat{\Theta} - \Theta_0| < \varepsilon) = 1 \quad \text{for any } \varepsilon > 0$$

- Efficiency: $\hat{\Theta}$ is an efficient estimator of Θ if it is unbiased and has smallest variance among unbiased estimators.

→ Cramér-Rao Lower Bound: If $\hat{\Theta}$ is an unbiased estimator of Θ, then $Var(\hat{\Theta}) \geq \frac{1}{n I_1(\Theta_0)}$

the true parameter value.

→ $\hat{\Theta}^{MLE}$ is asymptotically efficient:

→ as $n \to \infty$, $\hat{\Theta}^{MLE}$ becomes unbiased and $Var(\hat{\Theta}^{MLE}) \to$ CRLB.

---

2) Set up to do confidence/credible intervals and hypothesis tests with a large sample size.

Frequentist: As $n \to \infty$, it is approximately true that

variation in $\hat{\Theta}^{MLE}$ across samples → $\hat{\Theta}^{MLE} \sim \text{Normal}\left(\Theta_0, \frac{1}{n I_1(\Theta_0)}\right)$

Bayesian: As $n \to \infty$, the posterior distribution of Θ approaches:

probabilistic statement of state of knowledge. → $\Theta | X_1, \dots, X_n \sim \text{Normal}\left(\Theta_0, \frac{1}{n I_1(\Theta_0)}\right)$

or: $\Theta | X_1, \dots, X_n \sim \text{Normal}\left(\hat{\Theta}^{MLE}, \frac{1}{n I_1(\Theta_0)}\right)$

Key ideas in proof of $\hat{\Theta}^{MLE} \sim \text{Normal}\left(\Theta_0, \frac{1}{n I_1(\Theta_0)}\right)$

- Main idea: apply Central Limit Theorem.

↳ apply to $\underbrace{\frac{1}{n} \sum_{i=1}^{n} \ell(\Theta | X_i)}_{\ell(\Theta | X_1, \dots, X_n)}$ if $X_1, \dots, X_n$ independent.

- Idea 2: Taylor Series approximation:

$$\ell(\Theta | X_1, \dots, X_n) \approx \ell(\Theta_0 | X_1, \dots, X_n) + \underbrace{\ell'(\Theta | X_1, \dots, X_n)\Big|_{\Theta=\Theta_0}}_{C} (\Theta - \Theta_0) + \frac{1}{2} \underbrace{\ell''(\Theta | X_1, \dots, X_n)\Big|_{\Theta=\Theta_0}}_{I(\Theta_0)} (\Theta - \Theta_0)^2$$

- $\ell(\Theta | \cdots)$ is maximized at $\hat{\Theta}^{MLE}$

↳ $\ell'(\Theta | \cdots) = 0$ at $\hat{\Theta}^{MLE}$

- $\ell''(\Theta | \cdots) \to$ Fisher information.