ST102 Week 12

I. Introduction

II. Course structure

(i) attendance

(ii) assignment (check your submission)

(iii) office hours   12:00 – 13:00 Thu
Workspace 3, LSE Life

Assumption. Quite often, I.I.D. is assumed for a random sample $\{X_i\}_{i=1}^n$, with each $X_i \sim f(x; \theta)$.

Under such assumption, a sample has joint pf/p.d.f.

$$f(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} f(x_i; \theta)$$

Notation. Random sample $\{X_i\}_1^n$ v.s. realized sample $\{x_i\}_1^n$

Def. A statistic is a known function of r.v.'s $\{X_i\}_1^n$ in a random sample.

Def. A sampling distribution is a probability distribution in the value of a statistic.

The case of i.i.d. r.v.'s

Recall, for any seq. of independent r.v.'s,

1) $E(\sum_i a_i X_i) = \sum_i a_i E(X_i)$ Linearity

2) $Var(\sum_i a_i X_i) = \sum_i a_i^2 Var(X_i)$

Then for any i.i.d. sample $\{X_i\}_1^n$ with sample mean $\bar{X} = \frac{1}{n}\sum_i X_i$, we immediately know:

$$E(\bar{X}) = E\left(\sum_{i=1}^n X_i / n\right) = \sum_i \frac{1}{n} E(X_i) = E(X_1)$$

$$Var(\bar{X}) = Var\left(\sum_i \frac{1}{n} X_i\right) = \sum_i \frac{1}{n^2} Var(X_i) = \frac{1}{n} Var(X_1)$$

The case of i.i.d. Normal

Assume i.i.d. $X_i \sim N(\mu, \sigma^2)$:

It can be shown (like using the uniqueness of mgf) that

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Thm. (Central Limit Thm (CLT))

Given i.i.d. r.v.'s $\{X_i\}_1^n$, $E(X_i) = \mu < \infty$, and $Var(X_i) = \sigma^2 < \infty$, then

$$\lim_{n \to \infty} \mathbb{P}\left(\frac{\overline{X} - \mu}{\sigma / \sqrt{n}} \le z \right) = \Phi(z) \quad , \forall z \in \mathbb{R},$$

where Φ is the cdf. of $N(0, 1)$.

Rk. For any $n \in \mathbb{Z}_+$, the above is always an approximation. Need large n for good approximation.