Chapter 4

Central limit and Slusky's theorems

The central limit theorems (CLTs) give the asymptotic distributions of sums of independent random variables and Slutky's theorems give the asymptotic distribution of functions of random variables and of sequences that are asymptotically equivalent to other sequences.

4.1 Central limit theorems

Theorem 4.1 (Levy-Lindeberg's Theorem)
Let $\{X_n\}_{n\in I\!N}$ be a sequence of i.i.d. r.v.s with mean $E(X_n) = \mu$ and variance $V(X_n) = \sigma^2$, both finite. Then,

$$\frac{\displaystyle\sum_{i=1}^{n} X_i - E\left(\sum_{i=1}^{n} X_i\right)}{\sqrt{V\left(\displaystyle\sum_{i=1}^{n} X_i\right)}} \xrightarrow{d} N(0,1).$$

The first known CLT was the theorem below, due to De Moivre.

Corolary 4.1 (De Moivre's Theorem)
If $\{X_n\}_{n\in I\!N}$ is a sequence of i.i.d. r.v.s distributed as a Bern(p), then

$$\frac{\sum_{i=1}^{n} X_i - E\left(\sum_{i=1}^{n} X_i\right)}{\sqrt{V\left(\sum_{i=1}^{n} X_i\right)}} \xrightarrow{d} N(0,1).$$

Example 4.1 (Normal approximation of binomial distribution)
Consider $X \stackrel{d}{=} \text{Bin}(n,p)$. We know that $X = \sum_{i=1}^{n} Z_i$, where Z_i are i.i.d. with Bern(p) distribution. By the De Moivre's Theorem,

$$\frac{\sum_{i=1}^{n} Z_i - E\left(\sum_{i=1}^{n} Z_i\right)}{\sqrt{V\left(\sum_{i=1}^{n} Z_i\right)}} = \frac{X - np}{\sqrt{np(1-p)}} \xrightarrow{d} N(0,1).$$

However, the normal approximation is poor whenever either $n < 30$ or $n > 30$ but p is too small (such that $np < 5$).

Theorem 4.2 (Lindeberg-Feller's CLT)
Let $\{X_n\}_{n\in I\!N}$ be a sequence of independent r.v.s with means $E(X_n) = \mu_n$ and variances $V(X_n) = \sigma_n^2$, both finite $\forall n \in \mathcal{N}$. Let $c_n^2 = \sum_{i=1}^{n} \sigma_i^2$. If the following condition, known as the Lindeberg-Feller's condition (LFC), holds:

$$\forall \epsilon > 0, \quad \frac{1}{c_n^2} \sum_{i=1}^{n} E\left[(X_i - \mu_i)^2 \big| \, |X_i - \mu_i| \geq \epsilon\, c_n\right] \underset{n\to\infty}{\rightarrow} 0,$$

then

$$\frac{\sum_{i=1}^{n} X_i - E\left(\sum_{i=1}^{n} X_i\right)}{\sqrt{V\left(\sum_{i=1}^{n} X_i\right)}} \xrightarrow{d} N(0,1).$$

The previous theorem was also extended to sequences of triangular arrays of r.v.s of the form:

$$\begin{array}{lll} X_{11} & & \\ X_{21} & X_{22} & \\ X_{31} & X_{32} & X_{33} \\ \cdots & & \end{array},$$

where the r.v.s in each row are independent and satisfy the LFC, see the theorem below.

Theorem 4.3 (Lindeberg-Feller's CLT for triangular arrays)
Let $\{X_{ni},\ i=1,\ldots,n\}_{n\in I\!N}$ be a sequence of triangular arrays of r.v.s, where for each $n \in I\!N$, the r.v.s in n-th row $\{X_{n1},\ldots,X_{nn}\}$ are independent with finite means $E(X_{ni}) = \mu_{ni}$ and variances $V(X_{ni}) = \sigma_{ni}^2$. Let $c_n^2 = \sum_{i=1}^{n} \sigma_{ni}^2$. If the LFC holds for each row, that is, if

$$\forall \epsilon > 0, \quad \frac{1}{c_n^2} \sum_{i=1}^{n} E\left[(X_{ni} - \mu_{ni})^2 \big| \, |X_{ni} - \mu_{ni}| \geq \epsilon\, c_n\right] \underset{n\to\infty}{\longrightarrow} 0,$$

then

$$\frac{\sum_{i=1}^{n} X_{ni} - E\left(\sum_{i=1}^{n} X_{ni}\right)}{\sqrt{V\left(\sum_{i=1}^{n} X_{ni}\right)}} \xrightarrow{d} N(0,1).$$

Example 4.2 (Asymptotic normality of the LS estimator is linear regression)
Consider the sequence of r.v.s defined as

$$X_i = \alpha + \beta Z_i + e_i, \quad i = 1, 2, \ldots,$$

where $Z_1, Z_2, \ldots$ are known fixed values and $e_1, e_2, \ldots$ are i.i.d. r.v.s with $E(e_i) = 0$ and $V(e_i) = \sigma^2$, $i = 1, 2, \ldots$. Let us define $\bar{z}_n = n^{-1} \sum_{i=1}^n Z_i$ and $s_n^2 = n^{-1} \sum_{i=1}^n (Z_i - \bar{z}_n)^2$. The LS estimator of β is given by

$$\hat{\beta}_n = \frac{\sum_{i=1}^n X_i (Z_i - \bar{z}_n)}{\sum_{i=1}^n (Z_i - \bar{z}_n)^2}$$

(a) See that $\hat{\beta}_n$ can be also expressed as

$$\hat{\beta}_n = \beta + \frac{\sum_{i=1}^n e_i (Z_i - \bar{z}_n)}{\sum_{i=1}^n (Z_i - \bar{z}_n)^2}$$

(b) Using (a) and applying the Lindeberg-Feller's CLT to the sequence of random variables

$$X_{ni} = e_i (Z_i - \bar{z}_n), \quad i = 1, \ldots, n,$$

prove that if

$$\gamma_n := \max_{1 \le i \le n} \frac{(Z_i - \bar{z}_n)^2}{\sum_{j=1}^n (Z_j - \bar{z}_n)^2} \underset{n\to\infty}{\longrightarrow} 0,$$

then

$$\sqrt{n}\, s_n\, (\hat{\beta}_n - \beta) \xrightarrow{d} N(0, \sigma^2).$$

4.2 Slutsky's theorems

Theorem 4.4 (Slutsky's theorems)

Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of d-dimensional r.v.s with $X_n \xrightarrow{d} X$. Then it holds

(i) For any $f : \mathbb{R}^d \to \mathbb{R}^k$ such that $P(X \in \mathcal{C}(f)) = 1$, then

$$f(X_n) \xrightarrow{d} f(X).$$

(ii) Let $\{Y_n\}_{n\in\mathbb{N}}$ be another sequence of d-dimensional r.v.s with $X_n - Y_n \xrightarrow{P} 0$. Then,

$$Y_n \xrightarrow{d} X.$$

(iii) Let $\{Y_n\}_{n\in\mathbb{N}}$ be another sequence of d-dimensional r.v.s with $Y_n \xrightarrow{P} c \in \mathbb{R}^d$. Then,

$$\begin{pmatrix} X_n \\ Y_n \end{pmatrix} \xrightarrow{d} \begin{pmatrix} X \\ c \end{pmatrix}.$$

Example 4.3 Consider the sequence of r.v.s $X_n \xrightarrow{d} N(0,1)$ and the function $f(x) = x^2$. Since f in continuous, by Theorem 4.4 (i),

$$X_n^2 \xrightarrow{d} \mathcal{X}^2_{(1)}.$$

Example 4.4 Consider the sequence of r.v.s $X_n \xrightarrow{d} N(0,1)$ and the function $f(x) = 1/x$. Now f is not continuous at $x = 0$, but since X is an absolutely continuous r.v., $P(X \in \mathcal{C}(f)) = P(X \in \{0\}^c) = 1 - P(X = 0) = 1$. Then, by Theorem 4.4 (i),

$$1/X_n \xrightarrow{d} 1/X.$$

Example 4.5 Consider the sequence of r.v.s $X_n = 1/n$ and the function $f = \mathbf{1}_{(0,\infty)}$. It holds that $X_n \stackrel{d}{\to} 0$ but $f(X_n) = 1 \stackrel{d}{\not\to} f(X) = 0$. This happens because $\mathcal{C}(f) = \{0\}^c$ and $P(X = 0) = 1$. Then, $P(X \in \mathcal{C}(f)) = P(X \in \{0\}^c) = 0$.

Corolary 4.2 (Asymptotic distribution of functions of several sequences of random variables)
Let $\{X_n\}_{n\in I\!N}$ be a sequence of d-dimensional r.v.s with $X_n \stackrel{d}{\to} X$ and $\{Y_n\}_{n\in I\!N}$ be a sequence of k-dimensional r.v.s with $Y_n \stackrel{P}{\to} c \in I\!R^k$. Let $f : I\!R^{d+k} \to I\!R^r$ be such that

$$P\left(\begin{pmatrix} X \\ c \end{pmatrix} \in \mathcal{C}(f)\right) = 1.$$

Then it holds

$$f(X_n, Y_n) \stackrel{d}{\to} f(X, c).$$