# CHAPTER 7

## A CLT WHEN SUMMANDS COME RANDOMLY FROM r POPULATIONS

### 7.1 INTRODUCTION AND STATEMENT OF THE MAIN RESULT:

Let $\{X_n\}$ be a sequence of mutually independent r.v.s with respective d.f.s $\{G_n\}$, all of which belong to the domain of normal attraction of a symmetric stable law G with index $\alpha$, $0< \alpha\leq 2$. Suppose further that at most r of the d.f.s. $\{G_n\}$ are distinct i.e. $G_n \in \{F_1, F_2, \ldots , F_r\}$.

Without loss of generality assume that $EX_1 = 0$, $i = 1, 2, \ldots$ whenever it exists.

For each n, let $\tau_j(n)$ be the number of r.v.s among $X_1, X_2, \ldots , X_n$ which have $F_j$ as their d.f.

Suppose that $\tau_j(n)$, for fixed j, is a r.v. possibly depending upon $\{X_n\}$.

If $\tau_j(n)$, $j = 1, 2, \ldots , r$ are constants dependent on n, Sreehari (1970) and Mason (1970) have proved that

$$P\{S_n\leq xB_n\} \to G(x), \text{ as } n \to \infty,$$

with $B_n = \left[\sum_{j=1}^{r} C_j^{\alpha}(\tau_j(n))\right]^{1/\alpha}$. Here $C_j(n)$ is proportional to $n^{1/\alpha}$, the constant of proportionality changing with j.

In this chapter, we permit $\tau_j(n)$, $j = 1, 2, \ldots , r$, to be r.v.s satisfying the condition:

For $j = 1, 2, \ldots, r$, $\tau_j(n)/n \to N_j$ in probability, where $P(N_j > 0) = 1$.



We prove the following theorem.

**Theorem 7.1.1:** *Under the assumptions of this section,* $P\{S_n \le x\nu(n)\} \to G(x)$, *as* $n \to \infty$, *where* $\nu(n)$ *is given by*

$$\nu^{\alpha}(n) = \sum_{j=1}^{r} C_j^{\alpha}(\tau_j(n)).$$

The proof of the theorem is given in Section 7.3. Section 7.2 is devoted to some preliminary results required for the proof of the theorems.

*Remark 7.1.1:* We are not assuming $\tau_j$'s to be independent of $\{X_n\}$.

**7.2 PRELIMINARY RESULTS:**

For $A, B \in \mathbb{F}$, denote the conditional probability of A given B by $P(A|B)$. If $P(B) = 0$, then we use the convention $P(A|B) = P(A)$.

**Definition 7.2.1:** A sequence $\{A_n\}$ of events is said to be **P-mixing** if $\lim_{n \to \infty}[P(A_n|A) - P(A_n)] = 0$ for every $A \in \mathbb{F}$.

**Definition 7.2.2:** Let $\xi(t)$ be an independent separable homogeneous process with independent increments defined on [0,1] such that $\xi(0) = 0$ and $\xi(t)$

(i) is stochastically continuous on the right,

(ii) has at most a denumerable number of discontinuities, all of the first kind, and

(iii) is defined by $E[\exp\{iu\xi(t)\}] = \exp\{-t\theta|u|^{\alpha}\}$, $0< \alpha\leq 2$, $\theta> 0$.

Then the process $\{\xi(t)\}$ is called a **symmetric stable process** with exponent $\alpha$.

**Lemma 7.2.1** (Barndorff-Neilson Lemma):

Let $\{k_n\}$ and $\{m_n\}$ with $k_n< m_n$ be two increasing sequences of positive integers with $k_n \to \infty$ and let $\{A_n\}$ be a sequence of events of $\mathbb{F}$ such that $A_n$ depends only on $X_{k_n}$, ... , $X_{m_n}$. Then $A_n$ is P-mixing.

Whenever the observations do not come randomly from r populations (i.e. $\tau_j(n)$ are not r.v.s ) but positive integer valued functions of n such that $\sum_{j=1}^{r}\tau_j(n) = n$, then $\tau_j(n)$ will be, in the remaining part of this section, denoted by $t_j(n)$, $j = 1, 2, \ldots, r$.

Let us define $\psi^{\alpha}(n) = \sum_{j=1}^{r} C_j^{\alpha}(t_j(n))$, $\sum_{j=1}^{r} t_j(n) = n$.

**Lemma 7.2.2:** The sequence $\{A_n\}$ defined by $A_n = \{S_n\leq x\psi(n)\}$ is P-mixing.

**Proof**: Let $\varepsilon> 0$ be an arbitrary constant and let A be any event of $\mathbb{F}$.

Define $E_n = \{|S_{[\log(n)]}|> \varepsilon\psi(n)\}$, where $[x]$ is n if $n\leq x< n+1$.

Denote $\delta_n = P(E_n|A)$.

Writing $P[S_n \le x\psi(n)|A]$

$$= P[S_n - S_{[\log(n)]} + S_{[\log(n)]} \le x\psi(n)|A]$$

and intersecting with the event $E_n$ we get after usual manipulations

$$P[S_n - S_{[\log(n)]} \le (x-\varepsilon)\psi(n)|A] - \delta_n$$

$$\le P[S_n \le x\psi(n)|A]$$

$$\le P[S_n - S_{[\log(n)]} \le (x+\varepsilon)\psi(n)|A] + \delta_n. \qquad \ldots(7.2.1)$$

Note that $\delta_n = P(E_n|A)$

$$= P[|S_{[\log(n)]}| > \varepsilon(\psi(n)/\psi(\log(n)))\ \psi(\log(n))]/P(A) \to 0,$$

because $\psi(n)/\psi(\log n) \to \infty$ as $n \to \infty$.

Similarly, it can be proved that

$$P[S_n - S_{[\log(n)]} \le (x-\varepsilon)\psi(n)] - \delta_n^*$$

$$\le P[S_n \le x\psi(n)]$$

$$\le P[S_n - S_{[\log(n)]} \le (x+\varepsilon)\psi(n)] + \delta_n^*, \qquad \ldots(7.2.2)$$

where $\delta_n^* = P(E_n)$.

Inequalities (7.2.1) and (7.2.2) imply that

$$P[S_n - S_{[\log(n)]} \le (x-\varepsilon)\psi(n)|A]$$

$$- P[S_n - S_{[\log(n)]} \le (x+\varepsilon)\psi(n)] - \delta_n - \delta_n^*$$

$$\le P[S_n \le x\psi(n)|A] - P[S_n \le x\psi(n)]$$

$\leq P[S_n - S_{[\log(n)]} \leq (x+\varepsilon)\psi(n) | A]$

$- P[S_n - S_{[\log(n)]} \leq (x-\varepsilon)\psi(n)] + \delta_n + \delta_n^*$. ...(7.2.3)

Note that in view of Lemma 7.2.1, the limit of the first term of the inequality (7.2.3), after adding and subtracting the term $P[S_n - S_{[\log(n)]} \leq (x-\varepsilon)\psi(n)]$, will be bounded below by $G(x-2\varepsilon) - G(x+2\varepsilon)$, whereas the limit of the third term of the inequality (7.2.3), after adding and subtracting the term $P[S_n - S_{[\log(n)]} \leq (x+\varepsilon)\psi(n)]$, will be bounded above by $G(x+2\varepsilon) - G(x-2\varepsilon)$, and hence as $n \to \infty$, we get

$-[G(x+2\varepsilon) - G(x-2\varepsilon)]$

$\leq \lim_{n \to \infty}^{*} \{P[S_n \leq x\psi(n) | A] - P[S_n \leq x\psi(n)]\}$

$\leq [G(x+2\varepsilon) - G(x-2\varepsilon)]$.

Now, allowing $\varepsilon \to 0$, we have the result that

$\lim_{n \to \infty} \{P[S_n \leq x\psi(n) | A] - P[S_n \leq x\psi(n)]\} = 0$

which proves the lemma.□

**Lemma 7.2.3**: Let $H(x) = P\{\sup_{0 \leq t \leq 1} |\xi(t^{\alpha})| \leq x\}$. Then $\lim_{n \to \infty} P[\max_{0 \leq j \leq n} |S_j| \leq xB_n] = H(x)$ for an appropriately chosen normalizing sequence of positive constants $B_n$.

*Remark 7.2.1*: This lemma is due to Sreehari (1970, Theorem 5.2) and hence the proof is omitted.

## 7.3 PROOF OF THE THEOREM:

We shall prove the theorem for $r = 2$; for $r > 2$, the proof is analogous.

Let $k$ be a positive integer to be chosen later appropriately. For convenience we shall denote $N_1$ by $N$ so that $N_2 = 1-N$.

Let us denote for a positive integer $k$,

$$B_i = \{(i/k) < N \leq (i+1)/k\},\ i = 0, 1, 2, \ldots, k-1;$$

$$D_{n,k} = \{|(\tau_1(n)/n) - N| < (1/k)\},$$

$$\upsilon^{\alpha}(n) = C_1^{\alpha}(\tau_1(n)) + C_2^{\alpha}(\tau_2(n)),$$

$$J_i(n) = P[\{S_n \leq x\upsilon(n)\} \cap D_{n,k} \cap B_i],\ i = 1, 2, \ldots, k-2;$$

$$\eta_n = P[\{S_n \leq x\upsilon(n)\} \cap D_{n,k} \cap \{N > (k-1)/k\}],$$

$$\gamma_n = P[\{S_n \leq x\upsilon(n)\} \cap D_{n,k} \cap \{N \leq 1/k\}],$$

$$\xi_n = P[\{S_n \leq x\upsilon(n)\} \cap D'_{n,k}],$$

where $D'_{n,k}$ is the complement of $D_{n,k}$ in $\Omega$.

We have

$$P[S_n \leq x\nu(n)] = \xi_n + \gamma_n + \eta_n + \sum_{i=1}^{k-1} J_i(n). \qquad \ldots(7.3.1)$$

For fixed $i$, define

$$\alpha_{ni} = [n(i-1)/k],$$

$$\beta_{ni} = [n(i+2)/k],$$

$\theta_{ni} = n-\alpha_{ni},$

$\delta_{ni} = n-\beta_{ni},$

$\upsilon_1^{\alpha}(n,i) = C_1^{\alpha}(\alpha_{ni}) + C_2^{\alpha}(\delta_{ni}),$

$\upsilon_2^{\alpha}(n,i) = C_1^{\alpha}(\beta_{ni}) + C_2^{\alpha}(\theta_{ni}).$

On the event $\{B_i \cap D_{n,k}\}$, $\nu_1(n,i) \le \nu(n) \le \nu_2(n,i)$ a.s.

We first prove the theorem in the case $x \ge 0$; when $x < 0$ the steps will be exactly similar.

Let

$C_{1,i}(n) = P[\{S_n \le x\upsilon_1(n,i)\} \cap D_{n,k} \cap B_i],$

$C_{2,i}(n) = P[\{S_n \le x\upsilon_2(n,i)\} \cap D_{n,k} \cap B_i].$

Then, note that

$C_{1,i}(n) \le P[\{S_n \le x\upsilon(n)\} \cap D_{n,k} \cap B_i] \le C_{2,i}(n).$

...(7.3.2)

Further,

$C_{2,i}(n)$

$= P[\{S_n \le x\upsilon_2(n,i)\} \cap D_{n,k} \cap B_i]$

$\le P[\{S_{1,\beta_{ni}'}+S_{2,\theta_{ni}} - \max_{\alpha_{ni}\le j\le\beta_{ni}} |S_{1,j}-S_{1,\beta_{ni}}|$

$- \max_{\delta_{ni}\le j\le\theta_{ni}} |S_{2,j}-S_{2,\theta_{ni}}| \le x\nu_2(n,i)\} \cap D_{n,k} \cap B_i].$ ...(7.3.3)

Let

$$V_{2,n,i} = \{\max_{\alpha_{ni} \le j \le \beta_{ni}} |S_{1,j}-S_{1,\beta_{ni}}| > \varepsilon\nu_2(n,i)\}$$

$$W_{2,n,i} = \{\max_{\delta_{ni} \le j \le \theta_{ni}} |S_{2,j}-S_{2,\theta_{ni}}| > \varepsilon\nu_2(n,i)\}.$$

Then, using elementary results in probability, we get from (7.3.3),

$C_{2,i}(n)$

$\le P[\{S_{1,\beta_{ni}}+S_{2,\theta_{ni}} \le (x+2\varepsilon)\nu_2(n,i)\} \cap B_i]$

$+ P(V_{2,n,i} \cap B_i) + P(W_{2,n,i} \cap B_i).$

Now using Lemmas 7.2.1, 7.2.2 and 7.2.3, we get from (7.3.1),

$\limsup_{n \to \infty} P[S_n \le x\nu(n)]$

$\le \sum_{i=1}^{k-2} P(B_i)\{\limsup P[\{S_{1,\beta_{ni}}+S_{2,\theta_{ni}} \le (x+2\varepsilon)\nu_2(n,i)\}]$

$+ \limsup P(V_{2,n,i}) + \limsup P(W_{2,n,i})\}$

$+ P\{N \le (1/k)\} + P\{N > (k-1)/k\}$

$\le \sum_{i=1}^{k-2} P(B_i)\{G(x+2\varepsilon)+2\{1-H((\varepsilon(k+3)/3)^{1/\alpha})\}\}$

$+ P\{N \le (1/k)\} + P\{N > (k-1)/k\}$

$\le G(x+2\varepsilon),$ ...(7.3.4).

by letting $k \to \infty$ because H is a proper d.f.

Now consider,

$$\sum_{i=1}^{k-2} C_{1,i}(n) = \sum_{i=1}^{k-2} P[\{S_n \le x\nu_1(n,i)\} \cap D_{n,k} \cap B_i]$$

$$\ge \sum_{i=1}^{k-2} P[\{S_{1,\alpha_{ni}} + S_{2,\delta_{ni}} + \max_{\alpha_{ni}\le j\le\beta_{ni}} |S_{1,j} - S_{1,\alpha_{ni}}|$$

$$+ \max_{\delta_{ni}\le j\le\theta_{ni}} |S_{2,j} - S_{2,\delta_{ni}}| \le x\nu_1(n,i)\} \cap B_i] - P(D'_{n,k})$$

$$\ge \sum_{i=1}^{k-2} P(B_i)\{P[\{S_{1,\alpha_{ni}} + S_{2,\delta_{ni}} \le (x-2\varepsilon)\nu_1(n,i)\} | B_i]$$

$$- P(V_{2,n,i}|B_i) - P(W_{2,n,i}|B_i)\} - P(D'_{n,k}),$$

where

$$V_{1,n,i} = \{\max_{\alpha_{ni}\le j\le\beta_{ni}} |S_{1,j} - S_{1,\alpha_{ni}}| > \varepsilon\nu_1(n,i)\},$$

$$W_{1,n,i} = \{\max_{\delta_{ni}\le j\le\theta_{ni}} |S_{2,j} - S_{2,\delta_{ni}}| > \varepsilon\nu_1(n,i)\}.$$

By using Lemmas 7.2.1, 7.2.2 and 7.2.3, we get from (7.3.1),

$$\liminf_{n\to\infty} P[S_n \le x\nu(n)]$$

$$\ge \sum_{i=1}^{k-2} P(B_i)\{\liminf P[\{S_{1,\alpha_{ni}} + S_{2,\delta_{ni}} \le (x-2\varepsilon)\nu_1(n,i)\}]$$

$$- \limsup P(V_{1,n,i}) - \limsup P(W_{1,n,i})\} - \limsup P(D'_{n,k})$$

$$- P\{N \le (1/k)\} - P\{N > (k-1)/k\}$$

$$\ge \sum_{i=1}^{k-2} P(B_i)\,[G(x-2\varepsilon) - 2\{1 - H((\varepsilon(k-3)/3)^{1/\alpha})\}]$$

$$- P\{N \le (1/k)\} - P\{N > (k-1)/k\}$$

$$\ge G(x-2\varepsilon), \text{ by letting } k \to \infty. \qquad \ldots(7.3.5)$$

The required result now follows from (7.3.4) and (7.3.5) on allowing $\varepsilon \to 0$.□

### 7.4 CONCLUDING REMARKS:

*Remark 7.4.1*: It is well known that the limit distribution of normalized sums $S_n$ of independent r.v.s $X_1, X_2, \ldots, X_n$ with $X_n \sim G_n \in \{F_1, F_2\}$, $F_i$ belonging to the domain of normal attraction of a stable law with index $\alpha_i$, $i = 1, 2$ with $\alpha_1 < \alpha_2$, need not always exists. Also it is known from the Theorem 4.1 of Sreehari (1970) that this limit distribution exists iff $\lim_{n \to \infty}\{C_2(\tau_2)/C_1(\tau_1)\} = \lambda > 0$ exists and is finite.

If we assume that the limit distribution of normalized sums $S_n$ exists, then an analog of Theorem 7.1 in this set up becomes:

**Theorem 7.4.1**: *Under the assumption that* $\{C_2(\tau_2)/C_1(\tau_1)\} \to^{P} \lambda>0$, *exists and is finite, as* $n \to \infty$, *we have*

$$P\{S_n \le x\nu(n)\} \to G^*(x),$$

*where* $G^*(x)$ *denotes composition of the stable laws with indices* $\alpha_1$ *and* $\alpha_2$.

The proof of this theorem is on the lines of the proof of Theorem 7.1 with slight modifications and hence it is omitted.

*Remark 7.4.2*: In both the theorems, restriction on the d.f.s to be in the domain of normal attraction of stable law(s) is due to the fact that Lemma 7.2.3 holds only for distributions in the domain of normal attraction of stable laws.

*Remark 7.4.3*: In the next chapter we shall obtain a local limit theorem version of the result proved in this chapter.