CHAPTER III

A SOLUTION FOR SHAPIRO'S PROBLEM, WHEN THE POWER IS NOT GREATER THAN $\frac{1}{2}$

In the previous chapter we stated Shapiro's problem and his results in both the identically distributed case and the case where the random variables were not necessarily identically distributed. In the latter case, when the power r is greater than $\frac{1}{2}$, the limit distribution is the same as it is in the former case; but if the power r is not greater than $\frac{1}{2}$, then the determination of the limit distribution is an open problem.

In this chapter we present a solution to the problem when r is not greater than $\frac{1}{2}$.

1. The case $r = \frac{1}{2}$

Theorem 8. Let $X_1, X_2, \ldots$ be a sequence of independent random variables and for each k, let f_k be the density function of X_k, $k = 1,2,\ldots$, Assume that the sequence $\{f_k\}$ satisfies the following properties:

1. $f_k(0)$ is nonzero, for every k,
2. the family $\{f_k\}$ is equicontinuous at 0,
3. the Cesàro's limit, L, of the sequence $\{f_k(0)\}$ exists and is positive, that is,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f_k(0) = L > 0 ,$$

4) there exists a real number M such that

$$\frac{f_k(0)}{\log \log k} \leq M ,$$

for all $k = 3, 4, \ldots$.

Then

$$(1.1) \qquad \lim_{n\to\infty} P\Big(\frac{1}{B_n} \sum_{k=1}^{n} \frac{1}{|X_k|^{\frac{1}{2}}} - A_n \leq x\Big) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt ,$$

the normal distribution, and the constants $B_n > 0$ and A_n can be chosen as follows:

$$B_n^2 = \sum_{k=1}^{n} \Big\{ \int_{|y| > \frac{1}{n \log \log n}} f_k(y) \cdot \frac{1}{|y|}\, dy - \Big(\int_{|y| > \frac{1}{n \log \log n}} f_k(y) \cdot \frac{1}{\sqrt{|y|}}\, dy \Big)^2 \Big\},$$

(1.2)

$$A_n = \frac{1}{B_n} \sum_{k=1}^{n} \int_{|y| > \frac{1}{n \log \log n}} f_k(y) \cdot \frac{1}{\sqrt{|y|}}\, dy .$$

Proof. We first note that from property 2), for the constant L (the L given in property 3)) there is a positive real number δ such that

$$(1.3) \qquad |f_k(x) - f_k(0)| < \frac{L}{2}$$

for all x with $|x| < \delta$, and for all k, $k = 1, 2, \ldots$.

To prove the theorem, we let, for each k

$$y_k = \frac{1}{|X_k|^{\frac{1}{2}}}$$

and G_k be the distribution function of Y_k.

From theorem 3, it is possible to find real constants A_n and $B_n > 0$ having the property that the distribution of the sums

$$\frac{Y_1 + Y_2 + \dots + Y_n}{B_n} - A_n$$

converges to the normal law

$$\emptyset(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt ,$$

and the variables $\frac{Y_k}{B_n}$, $k = 1,2,\dots,n$ and $n = 1,2,\dots$ are infinitesimal if and only if there exists a sequence of real constants $\{c_n\}$, $c_n \to \infty$, such that as $n \to \infty$

$$(1.4) \qquad \sum_{k=1}^{n} \int_{|x|>c_n} dG_k(x) \to 0 \qquad \text{and}$$

$$\frac{1}{c_n^2} \sum_{k=1}^{n} \{ \int_{|x|<c_n} x^2 dG_k(x) - (\int_{|x|<c_n} x dG_k(x))^2 \} \to + \infty .$$

Let F_k be the distribution function of X_k. Then, for each k, we note that

$$(1.5) \qquad G_k(x) = P(Y_k \le x) = P(\frac{1}{|X_k|^{\frac{1}{2}}} < x)$$

$$= \begin{cases} F_k(-\frac{1}{x^2}) + 1 - F_k(\frac{1}{x^2}) & \text{if } x > 0, \\ 0 & \text{if } x \le 0. \end{cases}$$

To show that there exists a sequence $\{c_n\}$ satisfying (1.4), we let $c_1 = c_2 = 1$ and for $n \geq 3$

$$c_n = \sqrt{n \log \log n} .$$

Let n be so large that $\frac{1}{n \log \log n} < \delta$ (the δ given in 1.3). We note that

$$0 \leq \int_{|x|>c_n} dG_k(x) = \int_{x>\sqrt{n \log \log n}} d\left(F_k\left(-\frac{1}{x^2}\right) - F_k\left(\frac{1}{x^2}\right)\right)$$

$$= \int_0^{1/n \log \log n} (f_k(-y) + f_k(y))dy, \qquad \left(y = \frac{1}{x^2}\right).$$

And hence

$$0 \leq \sum_{k=1}^{n} \int_{|x|>c_n} dG_k(x) = \sum_{k=1}^{n} \int_0^{1/n \log \log n} (f_k(-y) + f_k(y))dy$$

$$\leq \sum_{k=1}^{n} (2f_k(0)+L) \int_0^{1/n \log \log n} dy \quad , \quad \text{(by (1.3))}$$

$$= \sum_{k=1}^{n} (2f_k(0)+L) . \frac{1}{n \log \log n}$$

$$= \left(2 . \frac{1}{n} . \sum_{k=1}^{n} f_k(0) + L\right) . \frac{1}{\log \log n} .$$

Since $\frac{1}{\log \log n}$ tends to zero and the term in the parenthesis is positive and finite, taking the limit as $n \to \infty$ we have the first assertion of (1.4). To show the second condition of (1.4), we observe that

$$0 \le \int_{|x|<c_n} x\,dG_k(x) = \int_{0<x<c_n} x\,d(F_k(-\frac{1}{x^2}) - F_k(\frac{1}{x^2}))$$

$$= \int_{1/n \log\log n}^{\infty} \frac{1}{\sqrt{y}} (f_k(-y) + f_k(y))dy \;, \qquad (y = \frac{1}{x^2})$$

$$\le \int_0^{\delta} \frac{1}{\sqrt{y}}(f_k(-y)+f_k(y))dy + \int_{\delta}^{\infty} \frac{1}{\sqrt{y}}(f_k(-y)+f_k(y))dy$$

$$\le 2(2f_k(0)+L)\sqrt{\delta} + \frac{1}{\sqrt{\delta}} \;.$$

That is,

$$(1.6) \qquad 0 \le \int_{|x|<c_n} x\,dG_k(x) \le (4f_k(0)+2L)\sqrt{\delta} + \frac{1}{\sqrt{\delta}} \;.$$

And $\int_{|x|<c_n} x^2 dG_k(x) = \int_{0<x<c_n} x^2 d(F_k(-\frac{1}{x^2}) - F_k(\frac{1}{x^2}))$

$$= \int_{y>1/(n \log\log n)} \frac{1}{y} \cdot (f_k(-y) + f_k(y))dy, \qquad (y = \frac{1}{x^2}),$$

$$\ge \int_{1/(n \log\log n)}^{\delta} \frac{1}{y} (f_k(-y) + f_k(y))dy$$

$$\ge (2f_k(0)-L) \int_{1/(n \log\log n)}^{\delta} \frac{1}{y}\, dy$$

$$= (2f_k(0)-L) \log(\delta n \log\log n) \;.$$

Thus $\frac{1}{c_n^2} \sum_{k=1}^{n} \{ \int_{|x|<c_n} x^2 dG_k(x) - (\int_{|x|<c_n} x\,dG_k(x))^2 \}$

$$\geq \frac{1}{n \log\log n} \sum_{k=1}^{n} \{(2f_k(0)-L)\log(\delta n \log\log n)-[(4f_k(0)+2L)\sqrt{\delta}+\frac{1}{\sqrt{\delta}}]^2\}$$

$$= \{\frac{2\sum_{k=1}^{n} f_k(0)}{n} - L\} \frac{\log(\delta n \log\log n)}{\log\log n}$$

$$- \frac{1}{n \log\log n} \{\sum_{k=1}^{n} [16f_k^2(0)\delta+(16L\delta+8)f_k(0)+4L^2\delta+\frac{1}{\delta}+4L]\}$$

$$\geq \{\frac{2\sum_{k=1}^{n} f_k(0)}{n} - L\} \frac{\log(\delta n \log\log n)}{\log\log n} -$$

$$\{\frac{16M\delta \sum_{k=1}^{n} f_k(0)}{n} + [(16L\delta+8)\frac{\sum_{k=1}^{n} f_k(0)}{n} + 4L^2\delta + \frac{1}{\delta} + 4L].\frac{1}{\log\log n}\}.$$

Note that $\{\frac{2\sum_{k=1}^{n} f_k(0)}{n} - L\} \to L > 0$, $\frac{\log(\delta n.\log\log n)}{\log\log n} \to \infty$,

and the last term tends to $16ML\delta$ which is finite, as $n \to \infty$. So we have the second condition of (1.4).

The assertion concerning $\{A_n\}$ and $\{B_n\}$ is an immediate consequence of (4.12) in Ch.I.

The theorem is now proved.

2. The case $r < \frac{1}{2}$.

Let $X_1, X_2, \ldots$ be a sequence of independent random variables, f_k the probability density function of X_k and suppose that the following conditions hold:

1) $f_k(0)$ is nonzero, for every k,

2) the family $\{f_k\}$ is equicontinuous at 0 and

3) $\lim\limits_{n\to\infty} \frac{\sum\limits_{k=1}^{n} f_k(0)}{n}$, the Cesàro limit of $f_k(0)$, exists and is positive, say equal to L.

Note that, by the equicontinuity of $\{f_k\}$ at 0, there is a positive real number δ such that

$$(2.1) \qquad |f_k(x) - f_k(0)| < 1$$

for all x with $|x| < \delta$ and for all k. Moreover, using the continuity of f_k at 0 and condition 1), there exists, for each k, a positive real number δ_k such that

$$(2.2) \qquad |f_k(x) - f_k(0)| < \frac{f_k(0)}{2}$$

for all x with $|x| < \delta_k$. For each k, we denote by F_k, the distribution function of X_k, $k = 1,2,\ldots$.

We then get that

$$0 \leq \int_{\mathbb{R}} \frac{1}{|x|^{2r}} \, dF_k(x) = \int_{|x|<\delta_k} \frac{1}{|x|^{2r}} \, dF_k(x) + \int_{|x|>\delta_k} \frac{1}{|x|^{2r}} \, dF_k(x)$$

$$\leq \int_{|x|<\delta_k} \frac{1}{|x|^{2r}} f_k(x)dx + \frac{1}{\delta_k^{2r}}$$

$$\leq \frac{3f_k(0)}{1-2r} \cdot \delta_k^{(1-2r)} + \frac{1}{\delta_k^{2r}} < +\infty.$$

Hence $\int_{\mathbb{R}} \frac{1}{|x|^{2r}} dF_k(x)$ is finite for all k.

By the Cauchy-Bunyakovskii inequality,

$$0 \leq \left(\int_{\mathbb{R}} \frac{1}{|x|^r} dF_k(x)\right)^2 \leq \int_{\mathbb{R}} \frac{1}{|x|^{2r}} dF_k(x) < +\infty \quad , \quad \text{for all k.}$$

Thus

$$0 \leq \int_{\mathbb{R}} \frac{1}{|x|^{2r}} dF_k(x) - \left(\int_{\mathbb{R}} \frac{1}{|x|^r} dF_k(x)\right)^2 < +\infty \quad , \quad \text{for all k.}$$

We assume further that the system satisfies the property that the Cesàro sums of $\int_{\mathbb{R}} \frac{1}{|x|^{2r}} dF_k(x) - \left(\int_{\mathbb{R}} \frac{1}{|x|^r} dF_k(x)\right)^2$ converge to a positive number σ^2, that is to say

4) $$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \left\{\int_{R} \frac{1}{|x|^{2r}} dF_k(x) - \left(\int_{R} \frac{1}{|x|^r} dF_k(x)\right)^2\right\} = \sigma^2 > 0.$$

With the above four conditions we have the following theorem.

<u>Theorem 9</u>. Let $X_1, X_2, \ldots$ be a sequence of independent random variables and for each k, let f_k be the density function of X_k which we assume satisfy properties 1)-4). Then for some suitably chosen constant $A_n(r)$

$$\lim_{n\to\infty} P\left(\frac{1}{\sigma\sqrt{n}} \sum_{k=1}^{n} \frac{1}{|X_k|^r} - A_n(r) \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt, \tag{2.3}$$

the normal distribution, where $r < \frac{1}{2}$ and σ is given in 4).

<u>Proof</u> Let F_k be the distribution function of X_k, $k = 1,2,\dots$.

Let $\xi_{nk} = \dfrac{1}{\sigma\sqrt{n}\,|X_k|^r}$ and G_{nk} be the distribution function of ξ_{nk}, for $k = 1,2,\dots,n$ and $n = 1,2,\dots$,

To obtain the conclusion of the theorem it suffices (according to theorem 2) to show that for all $\varepsilon > 0$

$$(2.4), \qquad \lim_{n\to\infty} \sum_{k=1}^{n} \int_{|x|\geq\varepsilon} dG_{nk}(x) = 0 \qquad \text{and}$$

$$\lim_{n\to\infty} \sum_{k=1}^{n} \{ \int_{|x|<\varepsilon} x^2 dG_{nk}(x) - (\int_{|x|<\varepsilon} x dG_{nk}(x))^2 \} = 1 .$$

To show that (2.4) holds for every $\varepsilon > 0$, we first note that

$$(2.5), \quad G_{nk}(x) = P(\xi_{nk} \leq x) = P(\frac{1}{\sigma\sqrt{n}\ |X_k|^r} \leq x)$$

$$= \begin{cases} F_k(\dfrac{-1}{(\sigma\sqrt{n}x)^{\frac{1}{r}}}) + 1 - F_k(\dfrac{1}{(\sigma\sqrt{n}x)^{\frac{1}{r}}}) & , \ x > 0, \\ 0 & , \ x \leq 0. \end{cases}$$

Let ε be any positive real number. Let n be so large that

$$\frac{1}{(\sigma\sqrt{n}\varepsilon)^{\frac{1}{2}}} < \delta .$$

$$\sum_{k=1}^{n} \int_{|x|\geq\varepsilon} dG_{nk}(x) = \sum_{k=1}^{n} \int_{\varepsilon\leq x<+\infty} d(F_k(\frac{-1}{(\sigma\sqrt{n}x)^{\frac{1}{r}}}) - F_k(\frac{1}{(\sigma\sqrt{n}x)^{\frac{1}{r}}}))$$

$$= \sum_{k=1}^{n} \{\int_0^{\frac{1}{(\sigma\sqrt{n}\varepsilon)^{\frac{1}{r}}}} (f_k(-y)+f_k(y))dy, \ (y = \frac{1}{(\sigma\sqrt{n}x)^{\frac{1}{r}}}),$$

$$\leq \sum_{k=1}^{n} (2f_k(0)+2) \int_0^{\frac{1}{(\sigma\sqrt{n}\varepsilon)^{\frac{1}{r}}}} dy$$

$$= \sum_{k=1}^{n} (2f_k(0)+2)\left(\frac{1}{(\sigma\sqrt{n}\varepsilon)^{\frac{1}{r}}}\right)$$

$$= \sum_{k=1}^{n} \left(\frac{2f_k(0)}{n} + \frac{2}{n}\right)\left(\frac{1}{(\sigma\varepsilon)^{\frac{1}{r}}}\right) \cdot \frac{1}{n^{\frac{1}{2r}-1}} \ .$$

Since $(\frac{1}{2r})-1$ is positive, the last term tends to zero as $n\to\infty$ and we have the first assertion of (2.4).

Let $S_n = \sum_{k=1}^{n} \{ \int_{|x|<\varepsilon} x^2 dG_{nk}(x) - (\int_{|x|<\varepsilon} x dG_{nk}(x))^2\}$. Then, for n so large that that $\frac{1}{(\sigma\sqrt{n}\varepsilon)^{\frac{1}{r}}} < \delta$, we get that

$$S_n = \sum_{k=1}^{n} \{ \int_{0<x<\varepsilon} x^2 d\left(F_k\left(\frac{-1}{(\sigma\sqrt{n}x)^{\frac{1}{r}}}\right) - F_k\left(\frac{1}{(\sigma\sqrt{n}x)^{\frac{1}{r}}}\right)\right) -$$

$$\left(\int_{0<x<\varepsilon} x d\left(F_k\left(\frac{-1}{(\sigma\sqrt{n}x)^{\frac{1}{r}}}\right) - F_k\left(\frac{1}{(\sigma\sqrt{n}x)^{\frac{1}{r}}}\right)\right)\right)^2\}.$$

From now on, for each positive integer n we represent the constant $\frac{1}{(\sigma\sqrt{n}\varepsilon)^{\frac{1}{r}}}$ by θ_n. Then by substituting $y = \frac{1}{(\sigma\sqrt{n}x)^{\frac{1}{r}}}$ in the above expression we get that

$$S_n = \sum_{k=1}^{n} \int_{\theta_n}^{\infty} \left(\frac{1}{(\sigma^2 n y^{2r})}\right)(f_k(-y)+f_k(y))\, dy \, -$$

$$\left(\int_{\theta_n}^{\infty} \left(\frac{1}{(\sigma\sqrt{n}y^r)}\right)(f_k(-y)+f_k(y))dy\right)^2\}$$

$$= \frac{1}{\sigma^2 n} \sum_{k=1}^{n} \{\int_{\theta_n}^{\infty} (1/y^{2r})(f_k(-y)+f_k(y))dy - (\int_{\theta_n}^{\infty} (1/y^r)(f_k(-y)+f_k(y))dy)^2\}.$$

Since

$$\int_{\theta_n}^{\infty} (1/y^{2r})(f_k(-y)+f_k(y))dy = \int_{\theta_n}^{\infty} (1/y^{2r})f_k(-y)dy + \int_{\theta_n}^{\infty} (1/y^{2r})f_k(y)dy$$

and

$$\int_{\theta_n}^{\infty} (1/y^{2r})f_k(-y)dy \quad = \int_{\theta_n}^{\infty} (1/|y|^{2r})\, f_k(-y)dy$$

$$= \int_{-\infty}^{-\theta_n} (1/|z|^{2r})f_k(z)dz \; , \quad (z = -y),$$

we get that

$$\int_{\theta_n}^{\infty} (1/y^{2r})(f_k(-y)+f_k(y))dy = \int_{-\infty}^{-\theta_n} (1/|y|^{2r})f_k(y)dy + \int_{\theta_n}^{\infty} (1/|y|^{2r})f_k(y)dy$$

$$= \int_{|y|>\theta_n} (1/|y|^{2r})f_k(y)dy.$$

So we have that

$$(2.6) \quad S_n = \frac{1}{\sigma^2 n} \sum_{k=1}^{n} \{ \int_{|y|>\theta_n} (1/|y|^{2r}) f_k(y)dy - (\int_{|y|>\theta_n} (1/|y|^{r}) f_k(y)dy)^2 \}.$$

For n sufficiently large we get that

$$0 \le \frac{1}{n} \sum_{k=1}^{n} \{\int_R (1/|y|^{2r}) f_k(y)dy - (\int_R (1/|y|^{r}) f_k(y)dy)^2\} -$$

$$\frac{1}{n} \sum_{k=1}^{n} \int_{|y|>\theta_n} (1/|y|^{2r}) f_k(y)dy - (\int_{|y|>\theta_n} (1/|y|^{r}) f_k(y)dy)^2\}$$

$$= \frac{1}{n} \sum_{k=1}^{n} \int_{|y|<\theta_n} (1/|y|^{2r}) f_k(y)dy -$$

$$[(\int_R (1/|y|^{r}) f_k(y)dy)^2 - (\int_{|y|>\theta_n} (1/|y|^{r}) f_k(y)dy)^2]\}$$

$$\le \frac{1}{n} \sum_{k=1}^{n} \int_{|y|<\theta_n} (1/|y|^{2r}) f_k(y)dy$$

$$\le \frac{1}{n} \sum_{k=1}^{n} (f_k(0)+1) \int_{|y|\le\theta_n} 1/|y|^{2r} \, dy$$

$$= \frac{1}{n} \sum_{k=1}^{n} (f_k(0)+1) \cdot \frac{2}{1-2r} \cdot \frac{1}{(\sigma\sqrt{n}\varepsilon)^{\frac{1}{r}(1-2r)}}$$

Since $\frac{1}{r}^{(1-2r)} = \frac{1}{r} - 2 > 0$, the last term tends to zero as $n \to \infty$.

So we get that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \int_{|y|>\theta_n} (1/|y|^{2r}) f_k(y)dy - (\int_{|y|>\theta_n} (1/|y|^{r}) f_k(y)dy)^2\}$$

$$= \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \{\int_{\mathbb{R}} (1/|y|^{2r}) f_k(y)dy - (\int_{\mathbb{R}} (1/|y|^{r}) f_k(y)dy^2\}$$

$$= \sigma^2 .$$

That is,

$$\lim_{n\to\infty} \frac{1}{\sigma^2 n} \sum_{k=1}^{n} \{ \int_{|y|>\theta_n} (1/|y|^{2r}) f_k(y)dy - (\int_{|y|>\theta_n} (1/|y|^{r} f_k(y)dy)^2\} = 1.$$

Thus we have the second assertion of (2.4) and this proves the theorem.