FISHER INFORMATION AND CONVERGENCE TO STABLE LAWS

S. G. BOBKOV[1,4], G. P. CHISTYAKOV[2], AND F. GÖTZE[3]

Abstract. The convergence to stable laws is studied in relative Fisher information for sums of i.i.d. random variables.

1. Introduction

Let $(X_n)_{n\geq 1}$ be independent identically distributed random variables, and define the normalized sums

$$Z_n = \frac{X_1 + \cdots + X_n}{b_n} - a_n$$

for given (non-random) normalizing sequences $a_n \in \mathbf{R}$ and $b_n > 0$. Assuming that Z_n converges weakly in distribution to a random variable Z with a non-degenerate stable law, we consider the Fisher information distance

$$I(Z_n||Z) = \int_{-\infty}^{+\infty} \Big(\frac{p_n'(x)}{p_n(x)} - \frac{\psi'(x)}{\psi(x)}\Big)^2 p_n(x)\,dx,$$

where p_n respectively ψ denote the densities of Z_n respectively Z. The definition makes sense, if p_n is absolutely continuous and has a Radon-Nikodym derivative p_n'. Otherwise one puts $I(Z_n||Z) = +\infty$.

If X_1 has finite second moment with mean zero and variance one, the classical central limit theorem holds, with $a_n = 0$, $b_n = \sqrt{n}$, and Z being standard normal. In this case a striking result of Barron and Johnson [B-J] indicates that $I(Z_n||Z) \to 0$, as $n \to \infty$, as soon as $I(Z_n||Z) < +\infty$, for some n, that is, if for some n, Z_n has finite Fisher information

$$I(Z_n) = \int_{-\infty}^{+\infty} \frac{p_n'(x)^2}{p_n(x)}\,dx.$$

This observation considerably strengthens a number of results on the central limit theorem for strong distances involving the total variation and the relative entropy. It raises at the same time the question about possible extensions to non-normal limit stable laws (as mentioned e.g. in [J], p.104). The question turns out to be rather tricky, and

1991 *Mathematics Subject Classification.* Primary 60E.
Key words and phrases. Fisher information, limit theorems, stable laws.
1) School of Mathematics, University of Minnesota, USA; Email: bobkov@math.umn.edu.
2) Faculty of Mathematics, University of Bielefeld, Germany; Email: chistyak@math.uni-bielefeld.de.
3) Faculty of Mathematics, University of Bielefeld, Germany; Email: goetze@math.uni-bielefeld.de.
4) Research partially supported by NSF grant DMS-1106530 and SFB 701.

it is not evident that $I(Z_n)$ needs to be even bounded for large n (a property which is guaranteed by Stam's inequality in case of a finite second moment).

The present note gives an affirmative solution of the problem in case of the so-called non-extremal stable laws (see Definition 1.2 below). In the sequel, we shall consider non-degenerate distributions only.

Theorem 1.1. *Assume that the sequence of normalized sums Z_n defined above converges weakly to a random variable Z with a non-extremal stable limit law. Then $I(Z_n||Z) \to 0$, as $n \to \infty$, if and only if $I(Z_n||Z) < +\infty$, for some n.*

The normal case is included in this assertion. Note, however, that if X_1 has an infinite second moment, but still belongs to the domain of normal attraction, we have $I(Z_n||Z) = +\infty$ for all n. Hence, in this special case there is no convergence in relative Fisher information.

In the remaining cases Z has a stable distribution with some parameters $0 < \alpha < 2$, $-1 \leq \beta \leq 1$, with characteristic function $f(t) = \mathbf{E}\, e^{itZ}$ described by

$$\log f(t) = \exp\big\{iat - c|t|^\alpha\big(1 + i\beta\, \mathrm{sign}(t)\, \omega(t,\alpha)\big)\big\}, \tag{1.1}$$

where $a \in \mathbf{R}$, $c > 0$, and $\omega(t,\alpha) = \tan(\frac{\pi\alpha}{2})$ in case $\alpha \neq 1$, and $\omega(t,\alpha) = \frac{2}{\pi}\log|t|$ for $\alpha = 1$. In particular, $|f(t)| = e^{-c|t|^\alpha}$ which implies that Z has a smooth density $\psi(x)$.

Definition 1.2. *A stable distribution is called non-extremal, if it is normal or, if $0 < \alpha < 2$ and $-1 < \beta < 1$. In the latter case, the density ψ of Z is positive on the whole real line and satisfies asymptotic relations*

$$\psi(x) \sim c_0\, |x|^{-(1+\alpha)} \quad (x \to -\infty), \qquad \psi(x) \sim c_1\, x^{-(1+\alpha)} \quad (x \to +\infty) \tag{1.2}$$

with some constants $c_0, c_1 > 0$.

The property that X_1 belongs to the domain of attraction of a stable law of index $0 < \alpha < 2$ may be expressed explicitly in terms of the distribution function $F_1(x) = \mathbf{P}\{X_1 \leq x\}$. Namely, we have $Z_n \Rightarrow Z$ with some $b_n > 0$ and $a_n \in \mathbf{R}$, if and only if

$$F_1(x) = (c_0 + o(1))\, |x|^{-\alpha} B(|x|) \qquad (x \to -\infty), \tag{1.3}$$

$$1 - F_1(x) = (c_1 + o(1))\, x^{-\alpha}\, B(x) \qquad (x \to +\infty), \tag{1.4}$$

for some constants $c_0, c_1 \geq 0$ that are not both zero, and where $B(x)$ is a slowly varying function in the sense of Karamata. This description reflects a certain behavior of the characteristic function $f_1(t) = \mathbf{E}\, e^{itX_1}$ near the origin (cf. [I-L], [Z]).

In connection with Theorem 1.1, note that a similar assertion has recently been proved in [B-C-G1] for the relative entropy

$$D(Z_n||Z) = \int_{-\infty}^{+\infty} p_n(x) \log \frac{p_n(x)}{\psi(x)}\, dx.$$

Namely, it is shown that $D(Z_n||Z) \to 0$, if and only if $Z_n \Rightarrow Z$ and $D(Z_n||Z) < +\infty$ for some n (thus extending Barron's entropic central limit theorem, [B]). In the normal case, it is known that, if $\mathbf{E}X_1 = \mathbf{E}Z$ and $\mathrm{Var}(X_1) = \mathrm{Var}(Z) = \sigma^2$, then

$$\frac{\sigma^2}{2}\, I(Z_n||Z) \geq D(Z_n||Z). \tag{1.5}$$

Hence, the convergence in Fisher information distance is a stronger property than in relative entropy. The question how these two distances are related to each other with respect to other stable laws does not seem to have been addressed in the literature. Apparently it is a question about the existence of certain weak logarithmic Sobolev inequalities for probability distributions with heavy tails, and we do not touch it here. However, it is natural to conjecture that the situation is similar as in the normal case via a suitable analogue of (1.5).

Another obvious question concerns the description of distributions satisfying the conditions of Theorem 1.1. In the non-normal case, the property $I(Z_n||Z) < +\infty$ may be simplified to $I(Z_n) < +\infty$. Taking, for example, $n = 1$, we obtain $I(X_1) < +\infty$ as a sufficient condition. This is however a rather strong condition, which may be considerably weakened by choosing larger values of n. One may wonder therefore what assumptions need to be added to (1.3)-(1.4) in terms of F_1 or f_1 to obtain the convergence of Z_n to Z in relative Fisher information. A direct characterization may be given in terms of the behavior of f_1 at infinity, at least in one particular case. As shown in [B-C-G1], if X_1 has a finite first absolute moment, the property $I(Z_n) < +\infty$ for some n is equivalent to any of the following two conditions:

a) For some $\varepsilon > 0$, $|f_1(t)| = O(t^{-\varepsilon})$, as $t \to +\infty$;
b) For some $\nu > 0$,

$$\int_{-\infty}^{+\infty} |f_1(t)|^\nu\, t^2 dt < +\infty. \tag{1.6}$$

This characterization may be used in Theorem 1.1 in case $1 < \alpha \leq 2$, since then, by (1.3)-(1.4), we have $\mathbf{E}\,|X_1|^\delta < +\infty$, for all $0 < \delta < \alpha$.

Note that, when X_1 has finite second moment (which corresponds to the case $\alpha = 2$), (1.6) may be replaced by a weaker condition

$$\int_{-\infty}^{+\infty} |f_1(t)|^\nu\, |t| dt < +\infty, \quad \text{for some } \ \nu > 0. \tag{1.7}$$

Removing the weight $|t|$ from the integral (1.7), we obtain a further weaker condition. It will be equivalent to the property that Z_n has an absolutely continuous distribution with a bounded continuous density p_n, for some n (and consequently for any sufficiently large n). In that and only that case, the following uniform local limit theorem holds: $\sup_x |p_n(x) - \psi(x)| \to 0$, as $n \to \infty$ (provided that $Z_n \Rightarrow Z$ with $0 < \alpha \leq 2$, cf. [I-L]).

The paper is organized as follows. First we state some general bounds on Fisher information and some properties of densities which can be represented as convolutions of densities with finite Fisher information (Sections 2-4). A main result used here has been already proved in recent work [B-C-G3] by the authors. In Section 5 we turn to the stable case and discuss a number of auxiliary results such as local limit theorems, as well as questions about the behavior of characteristic functions of Z_n near zero. In Section 6 we reduce Theorem 1.1 to showing that the Fisher information $I(Z_n)$ is bounded in n. The subsequent sections are therefore focused on this boundedness problem. Section 7 introduces a special decomposition of convolutions, and the final steps of the proof of Theorem 1.1 can be found in Section 8. We will shall complement the proofs by comments explaining why condition (1.6) is sufficient for the validity of Theorem 1.1.

2. **General Results about Fisher Information**

If a random variable X has an absolutely continuous density p with Radon-Nikodym derivative p', its Fisher information is defined by

$$I(X) = I(p) = \int_{-\infty}^{+\infty} \frac{p'(x)^2}{p(x)}\, dx, \tag{2.1}$$

where the integration may be restricted to the set $\{p(x) > 0\}$. In any other case, $I(X) = +\infty$.

If $I(X)$ is finite then necessarily the distribution of X has to be absolutely continuous with density $p(x)$ such that the derivative $p'(x)$ exists and is finite on a set of full Lebesgue measure. Furthermore, one can show that, if $I(X) < +\infty$, then $p'(x) = 0$ at any point, where $p(x) = 0$.

It follows immediately from the definition that the I-functional is translation invariant and homogeneous of order -2, that is, $I(a + bX) = \frac{1}{b^2}\, I(X)$, for all $a \in \mathbf{R}$ and $b \neq 0$.

Since the function u^2/v is convex in the upper half-plane $u \in \mathbf{R}$, $v > 0$, this functional is convex. That is, for all densities $p_1, \dots, p_n$, we have Jensen's inequality

$$I(\alpha_1 p_1 + \dots + \alpha_n p_n) \ \leq \ \sum_{k=1}^n \alpha_k I(p_k) \qquad \Big(\alpha_k > 0, \ \sum_{k=1}^n \alpha_k = 1\Big).$$

The inequality may be generalized to arbitrary "continuous" mixtures of densities. In particular, for the convolution

$$p * q(x) = \int_{-\infty}^{+\infty} p(x-y) q(y)\, dx$$

of any two densities p and q , we have

$$I(p * q) \leq \min\{I(p), I(q)\}. \tag{2.2}$$

In other words, if X and Y are independent random variables with these densities, then

$$I(X + Y) \leq \min\{I(X), I(Y)\}.$$

This property may be viewed as monotonicity of the Fisher information: This functional decreases when adding an independent summand. In fact, a much stronger inequality is available.

Proposition 2.1. (Stam [St]) *If X and Y are independent random variables, then*

$$\frac{1}{I(X+Y)} \geq \frac{1}{I(X)} + \frac{1}{I(Y)}. \tag{2.3}$$

Let us also introduce the Fisher information distance

$$I(X||Z) = \int_{-\infty}^{+\infty} \Big(\frac{p'(x)}{p(x)} - \frac{\psi'(x)}{\psi(x)}\Big)^2\, p(x)\, dx,$$

defined with respect to a random variable Z having a stable law. We need the following elementary observation, which shows that the question of boundedness of the Fisher information $I(Z_n)$ and of the Fisher information distance $I(Z_n||Z)$ for the normalized sums Z_n as introduced in Theorem 1.1 are in fact equivalent.

Proposition 2.2. *If Z has a non-extremal stable law of some index $0 < \alpha < 2$, then, for any random variable X,*

$$I(X||Z) \ \leq\ 2I(X) + c(Z), \tag{2.4}$$

$$I(X) \ \leq\ 2I(X||Z) + c(Z), \tag{2.5}$$

where $c(Z)$ depends on the distribution of Z, only. In particular, $I(X||Z) < +\infty$, if and only if $I(X) < +\infty$.

The assertion is based on the fact that any non-extremal non-normal stable distribution has a smooth positive density ψ such that, for all $k = 1, 2, \dots$,

$$\big|(\log \psi(x))^{(k)}\big| \sim \frac{(k-1)!}{|x|^k} \qquad (|x| \to +\infty),$$

(cf. [I-L], [Z]). In particular, it yields

$$\frac{c_1}{1+|x|} \leq \frac{|\psi'(x)|}{\psi(x)} \leq \frac{c_2}{1+|x|} \qquad (x \in \mathbf{R}) \tag{2.6}$$

with some positive constants c_j. Hence, assuming that $I(X) < +\infty$, then writing

$$\Big(\frac{p'(x)}{p(x)} - \frac{\psi'(x)}{\psi(x)}\Big)^2 \ \leq\ 2\Big(\frac{p'(x)}{p(x)}\Big)^2 + 2\Big(\frac{\psi'(x)}{\psi(x)}\Big)^2 \ \leq\ 2\Big(\frac{p'(x)}{p(x)}\Big)^2 + 2c_2^2$$

and integrating this inequality with weight $p(x)$, we obtain (2.4). Similarly,

$$\Big(\frac{p'(x)}{p(x)}\Big)^2 \ \leq\ 2\Big(\frac{p'(x)}{p(x)} - \frac{\psi'(x)}{\psi(x)}\Big)^2 + 2c_2^2,$$

which leads to (2.5).

Similar arguments for normal case ($\alpha = 2$), however lead to a different conclusion.

Proposition 2.3. *If Z is normal, then $I(X||Z) < +\infty$, if and only if $I(X) < +\infty$ and $\mathbf{E}X^2 < +\infty$.*

Note that in case where X and Z have equal means and variances, we have $I(X||Z) = I(X) - I(Z)$.

3. Connection with Functions of Bounded Variation

Applying Cauchy's inequality, one immediately obtains from the definition (2.1) the following elementary lower bound on the Fisher information.

Proposition 3.1. *If X has an absolutely continuous density p with Radon-Nikodym derivative p', then*

$$\sqrt{I(X)} \geq \int_{-\infty}^{+\infty} |p'(x)|\,dx. \tag{3.1}$$

Here, the integral represents the total variation norm of the function p as used in Real Analysis,

$$\|p\|_{\rm TV} \, = \, \sup \sum_{k=1}^{n} |p(x_k) - p(x_{k-1})|,$$

where the supremum runs over all finite collections $x_0 < x_1 < \dots < x_n$.

The densities p with finite total variation are vanishing at infinity and are uniformly bounded by $\|p\|_{\rm TV}$. Moreover, their characteristic functions

$$f(t) = \int_{-\infty}^{+\infty} e^{itx} p(x)\,dx \qquad (t \in \mathbf{R}),$$

admit, by integration by parts, a simple bound on the decay at infinity

$$|f(t)| \leq \frac{\|p\|_{\rm TV}}{|t|} \qquad (t \neq 0). \tag{3.2}$$

Hence, by Proposition 3.1, if a random variable X has finite Fisher information, its density p and characteristic function $f(t) = \mathbf{E}\,e^{itX}$ satisfy similar bounds

$$\max_x p(x) \leq \sqrt{I(X)}, \qquad |f(t)| \leq \frac{\sqrt{I(X)}}{|t|} \quad (t \neq 0). \tag{3.3}$$

In general, the inequality (3.1) cannot be reversed, though this is possible for convolutions of three densities of bounded variation. The following statement may be found in [B-C-G3].

Proposition 3.2. *If independent random variables X_j $(j = 1,2,3)$ have densities p_j of bounded variation, then $S = X_1 + X_2 + X_3$ has finite Fisher information, and moreover,*

$$I(S) \, \leq \, \frac{1}{2}\Big[\|p_1\|_{\rm TV}\,\|p_2\|_{\rm TV} + \|p_1\|_{\rm TV}\,\|p_3\|_{\rm TV} + \|p_2\|_{\rm TV}\,\|p_3\|_{\rm TV}\Big]. \tag{3.4}$$

Note that the convolution of two densities of bounded variation (e.g., corresponding to the uniform distribution on finite intervals) may have an infinite Fisher information.

Remark 3.3. A similar bound on the Fisher information may also be given in terms of characteristic functions. In view of (3.4), it suffices to bound the total variation norm, which can be done by applying the inverse Fourier formula, at least in case of finite first absolute moment. Namely, one can easily show that, if a random variable X has a continuously differentiable characteristic function $f(t)$ for $t > 0$, and

$$\int_{-\infty}^{+\infty} t^2 \left(|f(t)|^2 + |f'(t)|^2\right) dt < +\infty, \tag{3.5}$$

then X must have an absolutely continuous distribution with density p of bounded total variation satisfying

$$\|p\|_{\rm TV} \leq \left(\int_{-\infty}^{+\infty} |tf(t)|^2\, dt \int_{-\infty}^{+\infty} |(tf(t))'|^2\, dt\right)^{1/4}. \tag{3.6}$$

4. Classes of Densities Representable as Convolutions

General bounds like (3.3) may considerably be sharpened in the case where p is representable as convolution of several densities with finite Fisher information. Here, we consider the collection $\mathfrak{P}_2(I)$ of all functions on the real line which can be represented as convolution of two probability densities with Fisher information at most I. Correspondingly, let $\mathfrak{P}_2 = \cup_I \mathfrak{P}_2(I)$ denote the collection of all functions representable as convolution of two probability densities with finite Fisher information. Note that, by (2.3), $I(p) \leq \frac{1}{2}\, I$, for any $p \in \mathfrak{P}_2(I)$.

Thus, a random variable $X = X_1 + X_2$ has density p in $\mathfrak{P}_2$, if the density may be written as

$$p(x) = \int_{-\infty}^{+\infty} p_1(x-y)p_2(y)\, dx \tag{4.1}$$

in terms of absolutely continuous densities p_1, p_2 of the independent summands X_1, X_2 having finite Fisher information. Differentiating under the integral sign, we obtain a Radon-Nikodym derivative of the function p,

$$p'(x) = \int_{-\infty}^{+\infty} p_1'(x-y)p_2(y)\, dy = \int_{-\infty}^{+\infty} p_1'(y)p_2(x-y)\, dy. \tag{4.2}$$

The latter expression shows that p' is an absolutely continuous function and has the Radon-Nikodym derivative

$$p''(x) = \int_{-\infty}^{+\infty} p_1'(y)p_2'(x-y)\, dy. \tag{4.3}$$

In other words, p'' appears as the convolution of the functions p_1' and p_2' which are integrable, according to Proposition 3.1.

Note that (4.3) defines $p''(x)$ at every individual point x, not just almost everywhere (which is typical for a Radon-Nikodym derivative). Using the property $p_i(x) = 0 \Rightarrow p_i'(x) = 0$ in case of finite Fisher information, we obtain a similar implication $p(x) = 0 \Rightarrow p''(x) = 0$, which holds for any x.

Moreover, a direct application of the inequality (3.1) in (4.3) shows that p' has finite total variation

$$\|p'\|_{\rm TV} = \int_{-\infty}^{+\infty} |p''(x)|\,dx \leq I.$$

These formulas may be used to derive various point-wise and integral relations within the class $\mathfrak{P}_2$ such as the following.

Proposition 4.1. *Given a density p in $\mathfrak{P}_2(I)$, for all $x \in \mathbf{R}$,*

$$|p'(x)| \ \leq \ I^{3/4} \sqrt{p(x)} \ \leq \ I. \tag{4.4}$$

In addition,

$$\int_{-\infty}^{+\infty} \frac{p''(x)^2}{p(x)}\,dx \leq I^2. \tag{4.5}$$

To be more precise, integration in (4.5) is restricted to the set $\{p(x) > 0\}$. This proposition can be found in [B-C-G3]; since the proof is short, we shall include it here for completeness. Starting with (4.1), where $I(p_i) \leq I$, define the functions $u_i(x) = \frac{p_i'(x)}{\sqrt{p_i(x)}}\,1_{\{p_i(x)>0\}}$ $(i = 1,2)$. Applying Cauchy's inequality in (4.2), we get

$$\begin{aligned} p'(x)^2 \ &= \ \Big(\int_{-\infty}^{+\infty} u_1(x) \cdot \sqrt{p_1(x-y)}\,p_2(y)\,dy\Big)^2 \\ &\leq \ I(X_1) \int_{-\infty}^{+\infty} p_1(x-y)\,p_2(y)^2\,dy \ \leq \ I(X_1)\sqrt{I(X_2)}\ p(x), \end{aligned}$$

where we used $p_2(y) \leq \sqrt{I(X_2)}$ on the last step. Hence, we obtain the first inequality in (4.4), and the second follows from $p(x) \leq \sqrt{I}$. Similarly, rewrite (4.3) as

$$p''(x) = \int_{-\infty}^{+\infty} \big(u_1(x-y)u_2(y)\big)\,\sqrt{p_1(x-y)p_2(y)}\,dx,$$

to get

$$p''(x)^2 \ \leq \ \int_{-\infty}^{+\infty} u_1(x-y)^2\,u_2(y)^2\,dy \int_{-\infty}^{+\infty} p_1(x-y)p_2(y)\,dx \ = \ u(x)^2 p(x),$$

where we define $u \geq 0$ by

$$u(x)^2 = \int_{-\infty}^{+\infty} u_1(x-y)^2\, u_2(y)^2\, dy.$$

It also follows that

$$\int_{-\infty}^{+\infty} u(x)^2\, dx = I(X_1)I(X_2) \leq I^2,$$

which implies (4.5) and thus proves Proposition 4.1.

The analytic properties of densities in $\mathfrak{P}_2$ allow to make use of different formulas for the Fisher information (by using integration by parts). For example,

$$I(X) = -\int_{-\infty}^{+\infty} p''(x)\, \log p(x)\, dx,$$

provided that the integrand is Lebesgue integrable.

We will need the following "tail-type" estimate for the Fisher information.

Corollary 4.2. *If p is in $\mathfrak{P}_2(I)$, then for any T real,*

$$\int_T^{+\infty} \frac{p'(x)^2}{p(x)}\, dx \ \leq\ I^{3/4} \sqrt{p(T)}\, |\log p(T)| + I \left(\int_T^{+\infty} p(x) \log^2 p(x)\, dx \right)^{1/2}. \tag{4.6}$$

Proof. Let us decompose the open set $G = \{x > T : p(x) > 0\}$ into the union of disjoint intervals (a_n, b_n), irrespective whether they are bounded or not. In particular, $p(a_n) = p(b_n) = 0$ (if $a_n > T$), so, $p'(x) \log p(x) \to 0$, as $x \downarrow a_n$, by Proposition 4.1, and similarly for b_n.

Let $a_n < T_1 < T_2 < b_n$. Since p' is an absolutely continuous function of bounded variation, integration of by parts is justified and yields

$$\int_{T_1}^{T_2} \frac{p'(x)^2}{p(x)}\, dx \ =\ \int_{T_1}^{T_2} p'(x)\, d\log p(x) = p'(x) \log p(x)\Big|_{x=T_1}^{T_2} - \int_{T_1}^{T_2} p''(x)\, \log p(x)\, dx.$$

Letting $T_1 \to a_n$ and $T_2 \to b_n$, we get in case $a_n > T$

$$\int_{a_n}^{b_n} \frac{p'(x)^2}{p(x)}\, dx = -\int_{a_n}^{b_n} p''(x)\, \log p(x)\, dx$$

and

$$\int_{a_n}^{b_n} \frac{p'(x)^2}{p(x)}\, dx = -p'(T) \log p(T) - \int_{a_n}^{b_n} p''(x)\, \log p(x)\, dx$$

in case $a_n = T$ (if such n exists). Anyhow, the summation over n gives

$$\int_{(T,+\infty)\cap G} \frac{p'(x)^2}{p(x)}\, dx \ \leq\ |p'(T) \log p(T)| + \int_{(T,+\infty)\cap G} |p''(x)\, \log p(x)|\, dx. \tag{4.7}$$

Here the first term on the right-hand side can be estimated by virtue of (4.4), which leads to the first term on the right-hand side of (4.6). Using (4.5) together with Cauchy's inequality, for the the last integral we also have

$$\bigg(\int_{(T,+\infty)\cap G} \frac{|p''(x)|}{\sqrt{p(x)}}\,\sqrt{p(x)}\,|\log p(x)|\,dx\bigg)^2 \,\leq\, I^2 \int_T^{+\infty} p(x)\log^2 p(x)\,dx,$$

thus proving Corollary 4.2.

5. **Stable Laws and Uniform Local Limit Theorems**

Let us return to the normalized sums

$$Z_n = \frac{1}{b_n}\,(X_1 + \dots + X_n) - a_n, \qquad (a_n \in \mathbf{R},\ b_n > 0), \tag{5.1}$$

associated with independent identically distributed random variables $(X_n)_{n\geq 1}$. In this section we discuss uniform limit theorems for densities p_n of Z_n and behavior of their characteristic functions near the origin. As before, if $Z_n \Rightarrow Z$, the density of the stable limit Z is denoted by ψ.

Introduce the characteristic functions of X_1 and Z_n,

$$f_1(t) = \mathbf{E}\,e^{itX_1}, \qquad f_n(t) = \mathbf{E}\,e^{itZ_n} = e^{-ita_n} f_1(t/b_n)^n \qquad (t \in \mathbf{R}).$$

To avoid confusion, we make the convention that $Z_1 = X_1$, i.e., $a_1 = 0$ and $b_1 = 1$.

Proposition 5.1. *Assume that $Z_n \Rightarrow Z$ weakly in distribution. If*

$$\int_{-\infty}^{+\infty} |f_1(t)|^\nu\,dt < +\infty, \quad \textit{for some } \nu > 0, \tag{5.2}$$

then for all n large enough, Z_n have bounded continuous densities p_n such that

$$\lim_{n\to\infty}\,\sup_x\,|p_n(x) - \psi(x)| \,=\, 0. \tag{5.3}$$

Proposition 5.2. *Assume that $Z_n \Rightarrow Z$ weakly in distribution. If*

$$\int_{-\infty}^{+\infty} |f_1(t)|^\nu\,|t| dt < +\infty, \quad \textit{for some } \nu > 0, \tag{5.4}$$

then for all n large enough, Z_n have continuously differentiable densities p_n with bounded derivatives, and moreover

$$\lim_{n\to\infty}\,\sup_x\,|p_n'(x) - \psi'(x)| \,=\, 0. \tag{5.5}$$

The first assertion is well-known, cf. [I-L], p.126. The condition (5.2) is actually equivalent to the property that for all sufficiently large n, say $n \geq n_0$, Z_n have bounded continuous densities p_n. In that case, the characteristic functions f_n are integrable whenever $n \geq 2n_0$. Conversely, under (5.2), these densities for $n \geq \nu$ are given by the inversion formula

$$p_n(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx}\, f_n(t)\, dt. \tag{5.6}$$

Under the stronger assumption (5.4), the above equality may be differentiated, and we get a similar representation for the derivative.

$$p_n'(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} (-it)\, e^{-itx}\, f_n(t)\, dt. \tag{5.7}$$

Although Proposition 5.2 is not stated in [I-L], its proof is identical to the proof of Proposition 5.1. An important ingredient in the argument is the fact that the weak convergence $Z_n \Rightarrow Z$ forces f_1 to be regularly behaving near the origin. This fact can also be used in the study of the boundedness of the Fisher information distance $I(Z_n||Z)$, so let us state it separately.

Proposition 5.3. *If $Z_n \Rightarrow Z$ weakly in distribution, where Z has a stable law of index $\alpha \in (0,2]$, then*

$$|f_1(t)| = \exp\{-c\,|t|^\alpha h(1/|t|)\}, \tag{5.8}$$

where $c > 0$ and $h = h(x)$ is a slowly varying function for $x \to +\infty$ such that

$$\lim_{n \to \infty} \frac{nh(b_n)}{b_n^\alpha} = 1. \tag{5.9}$$

One of the consequences of this relation is that, given $0 < \delta < \alpha$, the characteristic functions f_n admits on a relatively large interval the bound

$$|f_n(t)| \leq e^{-c(\delta)|t|^\delta} \qquad (|t| \leq \varepsilon b_n) \tag{5.10}$$

with some positive constants ε and $c(\delta)$, which are independent of n, cf. [I-L], p.123. This bound shows that the parts of the integrals (5.6)-(5.7) taken over the region $T \leq |t| \leq \varepsilon b_n$ with fixed $T > 0$ are indeed small, while the assumptions (5.2) and (5.4) guarantee smallness of these integrals taken over the remaining region $|t| \geq \varepsilon b_n$.

In comparison with (5.8) a more precise statement is obtained in [I-L], cf. Theorem 2.6.5, p.85. Namely, if $Z_n \Rightarrow Z$, then for all t small enough,

$$\log f_1(t) = \exp\big\{i\gamma t - c|t|^\alpha h(1/|t|)\big(1 + i\beta\, \mathrm{sign}(t)\, \omega(t,\alpha)\big)\big\},$$

where γ is real, $c > 0$, and the parameters $\alpha \in (0,2]$, $\beta \in [-1,1]$, and the function $\omega(t,\alpha)$ are the same as in the representation (1.1) for the logarithm of the characteristic function $f(t)$ of Z. By lengthy computations in the proof of Theorem 2.6.5 in [I-L], it was shown that the function $B(x)$ appearing in the asymptotic relations (1.3)-(1.4) and the function $h(x)$ are connected via

$$h(x) = (1 + o(1))B(x), \quad \text{as } \ x \to +\infty.$$

Taking into account (5.9), this implies that in the case $0 < \alpha < 2$ there is a constant $c > 0$ such that, as $n \to \infty$,

$$\mathbf{P}\{|X_1| > b_n\} \sim \frac{c}{n}. \tag{5.11}$$

Let us return to local limit theorems. From (5.3) and (5.5) we immediately obtain the convergence of a "truncated" Fisher information distance.

Corollary 5.5. *Assume that $Z_n \Rightarrow Z$ weakly in distribution, where Z has a non-extremal stable law. If $I(Z_{n_0}) < +\infty$, for some n_0, then for all n large enough, the random variables Z_n have continuously differentiable densities p_n, and for every fixed $T > 0$,*

$$\int_{-T}^{T} \Big(\frac{p_n'(x)}{p_n(x)} - \frac{\psi'(x)}{\psi(x)}\Big)^2 p_n(x)\, dx \,=\, o(1), \qquad n \to \infty. \tag{5.12}$$

The only property of non-extremal stable laws which is used to show (5.12) is the fact that ψ is everywhere positive.

By the assumption, we have $I(Z_n) < +\infty$, for all $n \geq n_0$, and by (3.3),

$$|f_{n_0}(t)| \leq \frac{c}{|t|} \quad (t \neq 0)$$

with $c = \sqrt{I(Z_{n_0})}$. Hence, the condition (5.4) is fulfilled with $\nu = 3n_0$. Therefore, we get both (5.3) and (5.5), and in particular, $p_n(x) \geq \varepsilon > 0$ in $|x| \leq T$, for all n large enough. As a result, the integrand in (5.12) is uniformly bounded by a sequence tending to zero.

6. Moderate Deviations

As before, for independent identically distributed random variables $(X_n)_{n\geq 1}$ put

$$Z_n = \frac{S_n}{b_n} - a_n, \quad \text{where} \quad S_n = X_1 + \dots + X_n \quad (a_n \in \mathbf{R},\ b_n > 0). \tag{6.1}$$

It is well known that if $Z_n \Rightarrow Z$, where Z has a stable law of some index $0 < \alpha \leq 2$, then necessarily

$$b_n = n^{1/\alpha} h(n), \tag{6.2}$$

where h is a slowly varying function in the sense of Karamata.

To study the behaviour of $I(Z_n||Z)$ in the non-extremal non-normal case, it is worthwhile noting that this Fisher information distance is finite, if and only $I(Z_n)$ is finite (Proposition 2.2). In the normal case, $I(Z_n||Z) < +\infty$, if and only if $I(Z_n) < +\infty$ and $\mathbf{E}Z_n^2 < +\infty$ (Proposition 2.3). The last inequality is equivalent to $\mathbf{E}X_1^2 < +\infty$ and then for the weak convergence $Z_n \Rightarrow Z$ with a standard normal limit one may take $b_n = \sqrt{n\,\mathrm{Var}X_1}$ and $a_n = \mathbf{E}X_1\sqrt{n}/\sqrt{\mathrm{Var}X_1}$.

In any case, the requirement that $I(Z_{n_0}) < +\infty$ implies that for all $n \geq n_0$, Z_n have absolutely continuous bounded densities which we denote in the sequel by p_n. Moreover, $p_n \in \mathfrak{P}_2$ whenever $n \geq 2n_0$, and then p_n have continuous derivatives p_n' of bounded variation (as discussed in the previous section).

As the next step towards Theorem 1.1, we prove:

Lemma 6.1. *Assume that $Z_n \Rightarrow Z$ weakly in distribution, where Z has a non-extremal stable law. If $\limsup_{n\to\infty} I(Z_n) < +\infty$, then*

$$\lim_{n\to\infty} I(Z_n||Z) \,=\, 0. \tag{6.3}$$

Proof. As before, denote by ψ the density of Z.

By the assumptions, we have for some n_0 and with some constant I'

$$I(S_n) \leq I' b_n^2, \qquad n \geq n_0.$$

If $n \geq 2n_0$, write $n = n_1 + n_2$ with $n_1 = [\frac{n}{2}]$, $n_2 = n - n_1$. Then $n_1 \geq n_0$ and $n_2 \geq n_0$, and hence

$$I(S_{n_1}) \leq I' b_{n_1}^2 \leq I b_n^2, \qquad I(S_n - S_{n_1}) \leq I' b_{n_2}^2 \leq I b_n^2$$

with some constant I in view of the almost polynomial behavior of b_n as described in (6.2). Thus,

$$Z_n = \Big(\frac{S_{n_1}}{b_n} - a_n\Big) + \frac{S_n - S_{n_1}}{b_n}$$

represents the sum of two independent random variables with Fisher information at most I. Therefore, $p_n \in \mathfrak{P}_2(I)$, for all $n \geq 2n_0$, and we may invoke Corollary 4.2.

In view of Corollary 5.5 we only need to show that, given $\varepsilon > 0$, one may choose $T > 0$ such that the integral

$$J = \int_{|x|>T} \Big(\frac{p_n'(x)}{p_n(x)} - \frac{\psi'(x)}{\psi(x)}\Big)^2\, dx$$

is smaller than ε, for all n large enough.

Clearly, $J \leq 2J_1 + 2J_2$, where

$$J_1 = \int_{|x|>T} \frac{p_n'(x)^2}{p_n(x)}\, dx, \qquad J_2 = \int_{|x|>T} \Big(\frac{\psi'(x)}{\psi(x)}\Big)^2 p_n(x)\, dx.$$

Recall that $\frac{|\psi'(x)|}{\psi(x)} \leq \frac{c}{1+|x|}$ with a constant c depending on ψ, only (cf. (2.6)). Hence,

$$J_2 \,\leq\, \Big(\frac{c}{1+T}\Big)^2,$$

which thus can be made as small, as we wish.

It remains to estimate J_1. We now apply (4.6) giving

$$J_1 \le I^{3/4} \Big(\sqrt{p_n(T)}\,|\log p_n(T)| + \sqrt{p_n(-T)}\,|\log p_n(-T)|\Big) + 2I \left(\int_{|x|\ge T} p_n(x)\log^2 p_n(x)\,dx\right)^{1/2}. \tag{6.4}$$

Using the uniform local limit theorem in the form (5.3) together with the asymptotic relation (1.1) for $\psi(x)$ at infinity, we easily get

$$\sqrt{p_n(\pm T)}\,|\log p_n(\pm T)| \le c\,\frac{\log T}{T^{(1+\alpha)/2}} + \varepsilon_n, \tag{6.5}$$

which holds for all sufficiently large n and all $T \ge T_0$ with $\varepsilon_n \to 0$ (as $n \to \infty$) and with constants $c > 0$ and $T_0 \ge 10$ depending on ψ, only.

To bound the integral in (6.4), we partition $\{x : |x| \ge T\}$ into the set

$$A = \{x : |x| \ge T,\, p_n(x) \le |x|^{-4}\}$$

and its complement B. By the definition,

$$\int_A p_n(x)\log^2 p_n(x)\,dx \le 16 \int_{|x|\ge T} |x|^{-4}\log^2|x|\,dx \le \frac{32}{T}. \tag{6.6}$$

On the other hand, p_n are uniformly bounded, namely, $\sup p_n(x) \le \sqrt{I}$, for all $n \ge 2n_0$ (cf. (3.3)). Hence, on the set B,

$$|\log p_n(x)| \le \log\frac{\sqrt{I}}{p_n(x)} + |\log\sqrt{I}\,| \le 4\log|x| + |\log I|,$$

and therefore

$$\int_B p_n(x)\log^2 p_n(x)\,dx \le c \int_{|x|\ge T} p_n(x)\log^2|x|\,dx, \tag{6.7}$$

where the constant depends on I.

Finally, we use the property that the moments $\mathbf{E}\,|Z_n|^\delta$ are uniformly bounded in n, whenever $0 < \delta < \alpha$ (cf. [I-L], p.142). Choosing $\delta = \alpha/2$ and using $c_\alpha \log^2|x| \le |x|^{\alpha/2}$ (where $|x| \ge T_0 \ge 10$), we obtain with some constant K that

$$\begin{aligned} K \ge \mathbf{E}\,|Z_n|^{\alpha/2} &\ge T^{\alpha/4}\,\mathbf{E}\,|Z_n|^{\alpha/4}\,1_{\{|Z_n|\ge T\}} \\ &= T^{\alpha/4}\int_{|x|\ge T}|x|^{\alpha/4}p_n(x)\,dx \ge c_\alpha T^{\alpha/4}\int_{|x|\ge T}p_n(x)\log^2|x|\,dx. \end{aligned}$$

Thus, the second integral in (6.7) may be bounded by $cT^{-\alpha/4}$ with some constant c independent of n. Combining this with (6.6), we obtain a similar bound for the integral in (6.4), and taking into account (6.5), we get $J_1 \le cT^{-\alpha/8} + \varepsilon_n$. This completes the proof of Lemma 6.1.

7. Binomial decomposition of convolutions

To show that the assumption in Lemma 6.1 holds as long as $I(Z_{n_0}) < +\infty$, for some n_0, we introduce a special decomposition of densities p_n of Z_n. To simplify the argument, assume $n_0 = 1$, so that $I(p) = I(X_1) < +\infty$, where $p = p_1$ denotes the density of X_1.

Keeping the same notations as in the previous sections, we use a suitable truncation (which is actually not needed in case $\alpha > 1$). Introduce the probability densities

$$\tilde{p}_n(x) = \frac{b_n}{1-\delta_n}\, p(b_n x)\, 1_{\{|x|\leq 1\}}, \qquad \tilde{q}_n(x) = \frac{b_n}{\delta_n}\, p(b_n x)\, 1_{\{|x|>1\}}$$

together with their characteristic functions

$$\tilde{f}_n(t) = \frac{1}{1-\delta_n} \int_{-b_n}^{b_n} e^{itx/b_n} p(x)\, dx, \qquad \tilde{g}_n(t) = \frac{1}{\delta_n} \int_{|x|>b_n} e^{itx/b_n} p(x)\, dx,$$

where $\delta_n = \int_{|x|>b_n} p(x)\, dx \sim \frac{c}{n}$ with some constant $c > 0$, as emphasized in (5.11).

Then we have a binomial decomposition for convolutions

$$p_n \,=\, ((1-\delta_n)\tilde{p}_n + \delta_n \tilde{q}_n)^{n*} \,=\, \sum_{k=0}^n \binom{n}{k} (1-\delta_n)^k\, \delta_n^{n-k}\, \tilde{p}_n^{k*} * \tilde{q}_n^{(n-k)*}. \tag{7.1}$$

Note that each convolution $\tilde{p}_n^{k*} * \tilde{q}_n^{(n-k)*}$ appearing in this weighted sum represents a probability density with characteristic function $\tilde{f}_n(t)^k\, \tilde{g}_n(t)^{n-k}$.

In this section we establish some properties of $\tilde{f}_n$, which will be needed in the proof of Theorem 1.1. The corresponding density $\tilde{p}_n$ is supported on $[-1,1]$, and without loss of generality one may assume that it has mean zero. This can always be achieved by shifting the distribution of X_1, that is, the density p. Thus, let $\tilde{f}_n'(0) = 0$.

In the sequel we assume that $Z_n \Rightarrow Z$ weakly in distribution, where Z has a non-extremal stable law with index $0 < \alpha < 2$. The next two lemmas do not use the assumption $I(p) < +\infty$ and may be stated for general distributions from the domain of attraction of these stable laws.

Lemma 7.1. *For all real t, with some constant C depending on p, only,*

$$|\tilde{f}_n'(t)| \leq \frac{C}{n}\, |t|. \tag{7.2}$$

Proof. Using the assumption $\tilde{f}_n'(0) = 0$, we deduce the upper bound

$$|\tilde{f}_n'(t)| \,\leq\, \frac{1}{b_n(1-\delta_n)} \int_{-b_n}^{b_n} |e^{itx/b_n} - 1||x|\, dF_1(x) \,\leq\, \frac{|t|}{b_n^2(1-\delta_n)} \int_{-b_n}^{b_n} x^2\, dF_1(x),$$

where F_1 is the distribution function of X_1. Integrating by parts we have

$$\int_{-b_n}^{b_n} x^2\,dF_1(x) = -b_n^2(1 - F_1(b_n) + F_1(-b_n)) + 2\int_0^{b_n} x(1 - F_1(x) + F_1(-x))\,dx \le 2\int_0^{b_n} x(1 - F_1(x) + F_1(-x))\,dx.$$

Since $1 - \delta_n \to 1$, we obtain that

$$|\tilde f_n'(t)| \le \frac{C|t|}{b_n^2}\int_0^{b_n} x(1 - F_1(x) + F_1(-x))\,dx \tag{7.3}$$

with some constant C depending on p.

Recall that in the asymptotical formulas (1.3)-(1.4) for F_1, the function B is equivalent to the slowly varying function h associated with the characteristic function of X_1. Thus, with some $c_0 \ge 0$, $c_1 \ge 0$ ($c_0 + c_1 > 0$), we have

$$F(x) = \frac{c_0 + o(1)}{(-x)^\alpha}\,h(-x), \quad x < 0; \qquad F(x) = 1 - \frac{c_1 + o(1)}{x^\alpha}\,h(x), \quad x > 0,$$

Hence, up to a constant, the integral in (7.3) does not exceed

$$\int_0^{b_n} \frac{h(x)}{x^{\alpha-1}}\,dx = b_n^{2-\alpha}h(b_n)\int_0^1 \frac{h(sb_n)}{h(b_n)}\,\frac{ds}{s^{\alpha-1}}.$$

But by the well-known result on slowly varying functions ([Se], pp. 66–67),

$$\int_0^1 \frac{h(sb_n)}{h(b_n)}\,\frac{ds}{s^{\alpha-1}} \to \int_0^1 \frac{ds}{s^{\alpha-1}} = \frac{1}{2-\alpha}, \quad \text{as } n \to \infty.$$

Therefore, returning to (7.3), we get with some other constant

$$|\tilde f_n'(t)| \le Cb_n^{-\alpha}h(b_n)\,|t|.$$

It remains to simplify the right-hand side by applying equation (5.9) of Proposition 5.3, telling us that $h(b_n) \sim b_n^\alpha/n$. Lemma 7.1 is proved.

Lemma 7.2. *Let $\delta \in (0,\alpha)$ and $\eta \in (0,1)$ be fixed. There exist positive constants ε, c, C, depending on p, δ, η, with the following property: If $k \ge \eta n$, then*

$$|\tilde f_n(t)|^k \le Ce^{-c|t|^\delta}, \qquad \text{for } |t| \le \varepsilon b_n. \tag{7.4}$$

Proof. This is an analogue of the bound (5.10) for the characteristic functions of Z_n. In order to prove this upper bound, assume $|t| \ge 1$ and note that

$$\tilde f_n(t) = \frac{1}{1-\delta_n}\big(f_1(t/b_n) - \delta_n \tilde g_n(t)\big), \qquad t \in \mathbf{R}. \tag{7.5}$$

To proceed, we apply Proposition 5.3. First recall that, according to Karamata's theorem, any positive slowly varying function $h(x)$ defined in $x \geq 0$ has a representation

$$h(x) = c(x) \exp\Big\{\int_{x_0}^x \frac{w(y)}{y}\,dy\Big\}, \qquad x \geq x_0,$$

where $x_0 > 0$, $c(x) \rightarrow c \neq 0$ and $w(x) \rightarrow 0$, as $x \rightarrow +\infty$. For $x_0 = \min_{n \geq 1} b_n$, $1 \leq |t| \leq \varepsilon b_n$, where $0 < \varepsilon \leq 1$ is fixed, this representation implies

$$\frac{h(b_n/|t|)}{h(b_n)} \geq |t|^{-\gamma} \qquad \text{with} \quad \gamma = \gamma(\varepsilon) = \sup_{y \geq 1/\varepsilon} |w(y)|.$$

Hence, from (5.8)-(5.9)

$$|f_1(t/b_n)| \ = \ \exp\{-c|t|^\alpha\, b_n^{-\alpha} h(b_n/|t|)\} \ \leq \ \exp\{-c'|t|^{\alpha-\gamma}/n\}$$

with some constant $c' > 0$.

We choose $\varepsilon > 0$ to be small enough so that $\gamma < \alpha - \delta$. Now, applying the above estimate in (7.5), we get in the region $1 \leq |t| \leq \varepsilon b_n$

$$\begin{aligned} |\tilde f_n(t)| \ &\leq \ \frac{1}{1-\delta_n}\,\big(|f_1(t/b_n)| + \delta_n\big) \\ &\leq \ \frac{1}{1-\delta_n}\,\big(\exp\{-c'|t|^{\alpha-\gamma}/n\} + \delta_n\big). \end{aligned}$$

One can simplify the right-hand side by noting that $\frac{c'|t|^{\alpha-\gamma}}{n} \leq \frac{c' b_n^{\alpha-\gamma}}{n} < K$ with some constant K. Using $\log x \leq x - 1$ (valid for all x) and $e^{-x} \leq 1 - (1-e^{-K})x$ for $0 \leq x \leq K$, we then have

$$\begin{aligned} \log\big(\exp\{-c'|t|^{\alpha-\gamma}/n\} + \delta_n\big) \ &\leq \ \exp\{-c'|t|^{\alpha-\gamma}/n\} + \delta_n - 1 \\ &\leq \ -\big(1 - e^{-K}\big)\,\frac{c'|t|^{\alpha-\gamma}}{n} + \delta_n \ \leq \ \frac{c_1}{n} - \frac{c_2|t|^{\alpha-\gamma}}{n} \end{aligned}$$

with positive constants c_j. As a result,

$$|\tilde f_n(t)| \ \leq \ \exp\Big\{\frac{1}{n}\,\big(c_1 - c_2\,|t|^{\alpha-\gamma}\big)\Big\}$$

with some other positive constants c_1 and c_2 (independent of n). It remains to raise this inequality to the power k, and (7.4) follows.

We will now develop a first application of Lemmas 7.1-7.2 using the assumption $I(p) < +\infty$. The latter forces p to have bounded variation and vanish at infinity. Hence,

$$\|\tilde p_n\|_{\rm TV} = b_n(1-\delta_n)^{-1} \|p\, 1_{\{|x| \leq b_n\}}\|_{\rm TV} \leq b_n (1-\delta_n)^{-1} \sqrt{I(p)}. \tag{7.6}$$

Using the inequality (3.2), we see that the characteristic function of $\tilde p_n$ satisfies

$$|\tilde f_n(t)| \leq \frac{cb_n}{|t|} \qquad (t \neq 0) \tag{7.7}$$

with some constant $c = c(p)$, depending on p, only.

Corollary 7.3. *If $I(p) < +\infty$, then under the assumptions of Lemma 7.2 with $k \geq 4$, with some constant C depending on p, δ, η, only,*

$$\int_{-\infty}^{+\infty} (1 + |t|)\, \big|\tilde f_n^k(t)\big|\, dt \,\leq\, C, \tag{7.8}$$

$$\int_{-\infty}^{+\infty} t^2 \big|(\tilde f_n^k)'(t)\big|^2\, dt \,\leq\, C. \tag{7.9}$$

Proof. We have $(\tilde f_n^k)'(t) = k \tilde f_n'(t)\, \tilde f_n(t)^{k-1}$, while by (7.2),

$$\int_{-\infty}^{+\infty} t^2\, |\tilde f_n'(t)|^2\, |\tilde f_n(t)|^{2(k-1)}\, dt \,\leq\, \frac{C^2}{n^2} \int_{-\infty}^{+\infty} t^4\, |\tilde f_n(t)|^{2(k-1)}\, dt.$$

To estimate the last integral, first we use (7.4) which gives

$$\int_{|t| \leq \varepsilon b_n} t^4\, |\tilde f_n(t)|^{2(k-1)}\, dt \leq C.$$

For the complementary region $|t| > \varepsilon b_n$, note that

$$\tilde f_n(b_n t) = \frac{1}{1 - \delta_n} \int_{-b_n}^{b_n} e^{itx} p(x)\, dx,$$

which shows that these functions are separated from 1 uniformly in n in $|t| \geq \varepsilon$. (This can easily be seen by using general separation bounds for characteristic functions which are discussed in [B-C-G2]).

Thus, $\sup_{|t| \geq \varepsilon} |\tilde f_n(b_n t)| \leq e^{-c}$, for some constant $c > 0$ independent of n. In addition, by (7.7),

$$t^4\, |\tilde f_n(b_n t)|^6 \leq \frac{c}{t^2}$$

with some other constant. Hence,

$$\int_{|t| \geq \varepsilon b_n} t^4\, |\tilde f_n(t)|^{2(k-1)}\, dt \,\leq\, b_n^5\, e^{-2c(k-4)} \int_{|t| \geq \varepsilon} t^4\, |\tilde f(b_n t)|^6\, dt \,\leq\, C b_n^5\, e^{-2ck}.$$

The last expression is exponentially small with respect to n by the constraint on k, and we arrive at (7.9). The first inequality (7.8), which is simpler, is proved similarly.

8. Boundedness of Fisher Information. Proof of Theorem 1.1

In this section we complete the last step in the proof of Theorem 1.1. Keeping the same notations as in the previous sections and recalling Lemma 6.1, we only need:

Lemma 8.1. *Assume that $Z_n \Rightarrow Z$ weakly in distribution, where Z has a non-extremal stable law. If $I(Z_{n_0}) < +\infty$, for some n_0, then $\sup_{n \geq n_0} I(Z_n) < +\infty$.*

In the normal case, when X_1 has a finite second moment, the assertion immediately follows from Stam's inequality (2.3). In view of Lemma 6.1, we therefore obtain Barron-Johnson theorem, i.e., $I(Z_n||Z) \to 0$. Thus, we may focus on the case $0 < \alpha < 2$.

To simplify the argument and the notations, we assume $n_0 = 1$ (otherwise, mild modifications connected with the binomial decomposition are needed). Thus, let $I(p) < +\infty$, where p is the density of X_1. As before, we denote by p_n the density of Z_n, and assume that $Z_n \Rightarrow Z$ weakly in distribution, where Z has a non-extremal stable law.

By Stam's inequality (2.3),

$$I(Z_n) \le \frac{b_n^2}{n}\, I(p).$$

Although the right-hand side tends to infinity, this inequality may be used for small values of n, and here it will be sufficient to show that $\sup_{n \ge n_0} I(Z_n) < +\infty$, for some n_0.

Our basic tool is the binomial decomposition (7.1) of the previous section. Note that, by the convexity of the I-functional,

$$I(p_n) \ \le \ \sum_{k=0}^{n} \binom{n}{k} (1-\delta_n)^k \delta_n^{n-k}\, I\big(\tilde p_n^{k*} * \tilde q_n^{(n-k)*}\big), \tag{8.1}$$

and it suffices to properly estimate the terms in this sum. To this aim, we fix a number $\eta \in (0,1)$ and distinguish two cases.

Lemma 8.2. *If $k \le n-3$, then*

$$I\big(\tilde p_n^{k*} * \tilde q_n^{(n-k)*}\big) \ \le \ C (n b_n)^2\, I(p) \tag{8.2}$$

with some constant C depending on p, only.

Proof. By the monotonicity property (2.2), $I(\tilde p_n^{k*} * \tilde q_n^{(n-k)*}) \le I(\tilde q_n^{(n-k)*})$. On the other hand, by Proposition 3.2, if $n-k \ge 3$,

$$I(\tilde q_n^{(n-k)*}) \ \le \ \frac{1}{2} \Big(\|\tilde q_n^{[(n-k)/3]*}\|_{\rm TV}^2 + 2\, \|\tilde q_n^{[(n-k)/3]*}\|_{\rm TV} \cdot ||\tilde q_n^{n-k-2[(n-k)/3]*}\|_{\rm TV} \Big).$$

But the total variation norm decreases when taking convolutions, so that $\|\tilde q_n^{s*}\|_{\rm TV} \le ||\tilde q_n||_{\rm TV}$ $(s = 1,2,\dots)$. Hence,

$$I(\tilde q_n^{(n-k)*}) \ \le \ \frac{3}{2}\, \|\tilde q_n\|_{\rm TV}^2.$$

In turn, by means of the inequality $\|p\|_{\rm TV} \le \sqrt{I(p)}$ (Proposition 3.1), we have

$$\|\tilde q_n\|_{\rm TV} = b_n \delta_n^{-1} \|p\, 1_{\{|x| > b_n\}}\|_{\rm TV} \le b_n \delta_n^{-1} \|p\|_{\rm TV} \le b_n \delta_n^{-1} \sqrt{I(p)},$$

where we used the property $p(-\infty) = p(+\infty) = 0$ for the first inequality. Thus

$$I\big(\tilde p_n^{k*} * \tilde q_n^{(n-k)*}\big) \ \le \ \frac{3}{2} \big(\sqrt{I(p)}\ b_n \delta_n^{-1}\big)^2.$$

Recalling that $\delta_n \sim \frac{c}{n}$, Lemma 8.2 is proved.

Lemma 8.3. *If* $15 \le \eta n \le k \le n$, *then*

$$I\big(\tilde p_n^{k*} * \tilde q_n^{(n-k)*}\big) \ \le \ C \tag{8.3}$$

with some constant C *depending on* p *and* η, *only.*

Proof. Again appealing to the monotonicity of the Fisher information, we will use the bound

$$I(\tilde p_n^{k*} * \tilde q_n^{(n-k)*}) \le I(\tilde p_n^{k*}).$$

Thus, it suffices to show that

$$I(\tilde p_n^{k*}) \le C. \tag{8.4}$$

Assume first that the following weaker condition holds: $\eta_0 n \le k \le n$, where $0 < \eta_0 < \eta$. Since $\|\tilde p_n\|_{\rm TV} \le C b_n \sqrt{I(p)} < +\infty$ (see (7.6) and Proposition 3.2), the convolution powers $\tilde p_n^{k*}$ have finite Fisher information, whenever $k \ge 3$. In view of the bound (7.7) on the characteristic functions, we may invoke inversion formulas like in (5.6)-(5.7) to write, for any $x \in \mathbf{R}$,

$$\tilde p_n^{k*}(x) \ = \ \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \tilde f_n(t)^k\,dt, \tag{8.5}$$

$$(\tilde p_n^{k*})'(x) \ = \ \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} (-it) \tilde f_n(t)^k\,dt, \tag{8.6}$$

$$\tilde p_n^{k*}(x) + x(\tilde p_n^{k*})'(x) \ = \ -\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx}\, tk \tilde f_n(t)^{k-1} \tilde f_n'(t)\,dt, \tag{8.7}$$

where for reasons of integrability it is safe to assume that $k \ge 5$.

Corollary 7.3 tells us that the Fourier transforms in (8.5) and (8.7) are well-defined for square integrable functions whose L^2-norms are bounded by a constant independent of k and n. Hence, the same is true for

$$x\,(\tilde p_n^{k*})'(x) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \left(\tilde f_n(t)^k + tk \tilde f_n(t)^{k-1} \tilde f_n'(t)\right) dt,$$

and we may write

$$|(\tilde p_n^{k*})'(x)| \le \frac{\psi_{nk}(x)}{|x|} \tag{8.8}$$

with

$$\|\psi_{nk}\|_2^2 \ = \ \int_{-\infty}^{+\infty} \psi_{nk}(x)^2\,dt \le C. \tag{8.9}$$

Moreover, according to (7.8), L^1-norms of the functions $(-it)\tilde f_n(t)^k$ in (8.6) are also bounded by a constant independent of k and n. Hence,

$$\sup_x\, |(\tilde p_n^{k*})'(x)| \le C$$

for all n and $\eta_0 n \le k \le n$. As a result, (8.8) may be sharpened to

$$|(\tilde p_n^{k*})'(x)| \le \frac{\psi_{nk}(x)}{1+|x|}$$

with some functions ψ_{nk} satisfying (8.9). By applying Cauchy's inequality, the latter immediately implies that

$$\|\tilde p_n^{k*}\|_{\rm TV} \ = \ \int_{-\infty}^{+\infty} |(\tilde p_n^{k*})'(x)|\,dx \le C'\,\|\psi_{nk}\|_2 \ \le \ C, \tag{8.10}$$

where the resulting constant C may depend on p and η_0 (by choosing, for example, $\delta = \alpha/2$ in the previous auxiliary lemmas of the previous section).

We now apply Proposition 3.2 to convolutions of any three densities $\tilde p_n^{k*}$, as above. That is, if $\eta_0 n \le k_j \le n$ and $k_j \ge 5$ $(j = 1,2,3)$, we obtain by (3.4) and (8.10) that

$$I(\tilde p_n^{(k_1+k_2+k_3)*}) \le \frac{3}{2}\,C^2. \tag{8.11}$$

Starting with $k \ge 15$, put $k_1 = k_2 = [\frac{k}{3}]$, $k_3 = k - (k_1 + k_2)$, so that $k_j \ge 5$. Also, if $k \ge \eta n$, we have $k_j \ge [\frac{\eta n}{3}] \ge \frac{\eta n}{6}$. Hence, we may choose $\eta_0 = \frac{\eta}{6}$, and thus (8.11) implies (8.3)-(8.4).

Proof of Lemma 8.1. In the case $15 \le \eta n \le n-3$, we may combine Lemmas 8.2 and 8.3 to get from (8.1) with some constant $C = C(p,\eta)$

$$\begin{aligned} I(p_n) \ &\le \ C(nb_n)^2 \sum_{0 \le k < \eta n} \binom{n}{k} (1-\delta_n)^k \delta_n^{n-k} + C \sum_{\eta n \le k \le n} \binom{n}{k} (1-\delta_n)^k \delta_n^{n-k} \\ &\le \ C(nb_n)^2 \cdot 2^n\, \delta_n^{(1-\eta)n} + C \ \le \ C', \end{aligned}$$

where the last inequality holds for all sufficiently large n (by using $\delta_n \sim \frac{c}{n}$) with, for example, $\eta = \frac{1}{2}$. Lemma 8.1 and therefore Theorem 1.1 are now proved.

Remark 8.4. Finally, let us comment on the conditions $a)-b)$ from the introductory section. In view of the general bound (3.3), $a)$ is always necessary for the finiteness of $I(Z_n)$ with some n. Also, $b)$ is weaker than $a)$, so we need explain the opposite direction.

If $1 < \alpha \le 2$, then X_1 has finite first absolute moment $C = \mathbf{E}\,|X_1|$. Hence, under (1.6), the condition (3.5) is fulfilled and thus the bound (3.6) is applicable to all Z_n with $n \ge (\nu+2)/2$. More precisely, denoting by $g_n(t) = f_1(t)^n$ the characteristic function of $S_n = X_1 + \dots + X_n$, we have

$$|(tg_n)'(t)| \le |g_n(t)| + |t||g_n'(t)| \le (1 + Cn|t|)\,|f_1(t)|^{n-1},$$

so S_n has a density $r_n(x)$ whose total variation norm satisfies

$$\|r_n\|_{\rm TV}^4 \ \le \ \int_{-\infty}^{+\infty} t^2\,|f_1(t)|^{2n}\,dt \int_{-\infty}^{+\infty} (1 + Cn|t|)^2\,|f_1(t)|^{2(n-1)}\,dt \ < \ +\infty.$$

By Proposition 3.2, we get $I(S_{3n}) < +\infty$.

References

[B] A. R. Barron. Entropy and the central limit theorem. Ann. Probab. 14 (1986), no. 1, 336-342.

[B-J] A. R. Barron and O. Johnson. Fisher information inequalities and the central limit theorem. Probab. Theory Related Fields, 129 (2004), no. 3, 391-409.

[B-C-G1] S. G. Bobkov, G. P. Chistyakov and F. Götze. Convergence to stable laws in relative entropy. J. Theor. Probab., September 2011 (online version).

[B-C-G2] S. G. Bobkov, G. P. Chistyakov and F. Götze. Bounds for characteristic functions in terms of quantiles and entropy. Bielefeld, Preprint (2011). To appear in: Electron. Comm. Probab.

[B-C-G3] S. G. Bobkov, G. P. Chistyakov and F. Götze. Fisher information and the central limit theorem. arXiv:1204.6650 [math.PR], (2012)

[I-L] I. A. Ibragimov and Ju. V. Linnik. Independent and stationary sequences of random variables. Izdat. "Nauka", Moscow, 1965, 524 pp.

[J] O. Johnson. Information theory and the central limit theorem. Imperial College Press, London, 2004, xiv+209 pp.

[Se] E. Seneta. Regularly varying functions. Lecture Notes in Math. 508, Springer-Verlag, Berlin-Haideberg-New York, 1976.

[St] A. J. Stam. Some inequalities satisfied by the quantities of information of Fisher and Shannon. Information and Control 2 (1959), 101-112.

[Z] V. M. Zolotarev. One-dimensional stable distributions. Probability Theory and Mathematical Statistics, "Nauka", Moscow, 1983, 304 pp.

Sergey G. Bobkov
School of Mathematics, University of Minnesota
127 Vincent Hall, 206 Church St. S.E., Minneapolis, MN 55455 USA
E-mail address: bobkov@math.umn.edu

Gennadiy P. Chistyakov
Fakultät für Mathematik, Universität Bielefeld
Postfach 100131, 33501 Bielefeld, Germany
E-mail address: chistyak@math.uni-bielefeld.de

Friedrich Götze
Fakultät für Mathematik, Universität Bielefeld
Postfach 100131, 33501 Bielefeld, Germany
E-mail address: goetze@mathematik.uni-bielefeld.de