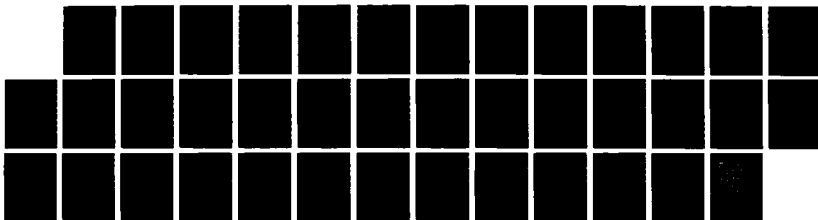
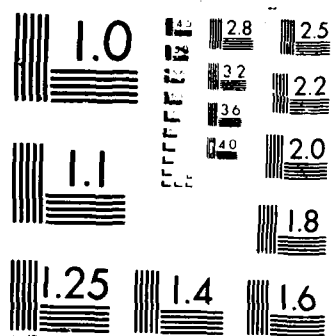


AD-A189 341 CONVOLUTION METRICS AND RATES OF CONVERGENCE IN THE CLT 1/1
(CENTRAL LIMIT TH. (U) NORTH CAROLINA UNIV AT CHAPEL
HILL CENTER FOR STOCHASTIC PROC. S T RACHEV ET AL.
UNCLASSIFIED OCT 87 TR-210 AFOSR-TR-87-1915 F/G 12/3 NL





2

DTIC FILE COPY

CENTER FOR STOCHASTIC PROCESSES

AD-A189 341

AFOSR-TR. 87-1315

Department of Statistics
University of North Carolina
Chapel Hill, North Carolina



Convolution Metrics

and

Rates of Convergence in the CLT

by

S.T. Rachev

and

J.E. Yukich

DTIC
ELECTE
JAN 14 1988
S H

Technical Report No. 210

October 1987

DISTRIBUTION STATEMENT

Approved for

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified			1b. RESTRICTIVE MARKINGS		
2a. SECURITY CLASSIFICATION AUTHORITY			3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.		
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE					
4. PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Report No. 210			5 MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 87-1915		
6a. NAME OF PERFORMING ORGANIZATION University of North Carolina	6b OFFICE SYMBOL (If applicable)	7a NAME OF MONITORING ORGANIZATION AFOSR/NM			
6c. ADDRESS (City, State, and ZIP Code) Statistics Dept. 321-A Phillips Hall 039-A Chapel Hill, NC 27514		7b ADDRESS (City, State, and ZIP Code) AFOSR/NM Bldg 410 Bolling AFB DC 20332-6448			
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR	8b OFFICE SYMBOL (If applicable) NM	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-NO. F49620 85C 0144.			
8c. ADDRESS (City, State, and ZIP Code) AFOSR/NM Bldg 410 Bolling AFB DC 20332-6448		10 SOURCE OF FUNDING NUMBERS			
		PROGRAM ELEMENT NO 61102F	PROJECT NO 2304	TASK NO. A5	WORK UNIT ACCESSION NO
11 TITLE (Include Security Classification) Convolution metrics and rates of convergence in the CLT (Central Limit Theorem)					
12. PERSONAL AUTHOR(S) Rachey, S.T. and Yukich, J.E.					
13a. TYPE OF REPORT October 1987	13b. TIME COVERED FROM 9/87 TO 8/88	14. DATE OF REPORT (Year, Month, Day) October 1987		15 PAGE COUNT 36	
16 SUPPLEMENTARY NOTATION					
17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB-GROUP	Key Words & Phrases: Ideal probability metrics, convolution metrics, rate of convergence, stable random variables.		
19 ABSTRACT (Continue on reverse if necessary and identify by block number)					
Abstract Let B be a complete separable Banach space and let $\mathbb{E} = \mathbb{E}(B)$ be the vector space of all random variables defined on a probability space $\Omega, \mathcal{G}, P$ and taking values in B. It is known that metrics on $\mathbb{E}$ of convolution type enjoy a variety of interesting properties. In this article it is shown that convolution metrics may also be used to obtain rates of convergence in CLT's involving a stable limit law. The rates are expressed in terms of a variety of uniform metrics on $\mathbb{E}$ and include the total variation metric and the uniform metrics between density and characteristic functions. The results represent both an improvement and an extension of existing results. Weak convergence properties of convolution metrics are also explored.					
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS			21 ABSTRACT SECURITY CLASSIFICATION Unclassified/unlimited		
22a NAME OF RESPONSIBLE INDIVIDUAL Maj. Brian Woodruff			22b TELEPHONE (Include Area Code) (202) 767-5027	22c OFFICE SYMBOL NM	

Convolution Metrics
and
Rates of Convergence in the CLT

by

S. T. Rachev^{* 1}
Center for Stochastic Processes
Department of Statistics
University of North Carolina

and

J.E. Yukich
Institute of Mathematics
Sofia, Bulgaria

* On leave from Institute of Mathematics, Sofia, Bulgaria.

¹ Research supported by the Air Force Office of Scientific Research Grant
No. F49620 85C 0144.

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	



By	
Dissemination	
A 81 0115 0000	
A-1	

Abstract. Let $(B, \|\cdot\|)$ be a complete separable Banach space and let $\mathfrak{X} = \mathfrak{X}(B)$ be the vector space of all random variables defined on a probability space $(\Omega, \mathcal{Q}, P)$ and taking values in B . It is known that metrics on $\mathfrak{X}$ of convolution type enjoy a variety of interesting properties. In this article it is shown that convolution metrics may also be used to obtain rates of convergence in CLT's involving a stable limit law. The rates are expressed in terms of a variety of uniform metrics on $\mathfrak{X}$ and include the total variation metric and the uniform metrics between density and characteristic functions. The results represent both an improvement and an extension of existing results. Weak convergence properties of convolution metrics are also explored.

AMS 1980 Subject Classifications

Primary: 60F05, 60G50, 60B10

Secondary: 60B12, 60E07

Key words and phrases: ideal probability metrics, convolution metrics, rate of convergence, stable random variables.

§0. Introduction

In this article we shall be concerned with various convolution pseudo-metrics on $P(\mathbb{R}^k)$, or more generally $P(B)$, B a Banach space, where $P(B)$ denotes the collection of all Borel probability measures on B . By convolution pseudo-metrics on $P(\mathbb{R}^k)$ we mean distances of the form

$$d_g(P, Q) := \sup_{x \in \mathbb{R}^k} \left| \int g(x-y)(dP-dQ)(y) \right| \quad P, Q \in P(\mathbb{R}^k),$$

where g is a kernel belonging to $L^1(\mathbb{R}^k) \cap C_0(\mathbb{R}^k)$ and where $C_0(\mathbb{R}^k)$ is the collection of bounded continuous functions on $\mathbb{R}^k$ vanishing at infinity. Every such kernel generates an associated distance d_g on $P(\mathbb{R}^k)$. When the set $\{t: \hat{g}(t) = 0\}$ has empty interior, d_g is actually a metric (here $\hat{g}$ denotes the Fourier transform of g); we refer the reader to Theorem A below for more on this.

We recall that these convolution pseudo-metrics enjoy a variety of properties and that they are actually quite useful from a statistical point of view. For example, as the following shows, one may characterize those convolution pseudo-metrics metrizing weak convergence and one can also obtain a CLT for $d_g(P_n, P)$, where here and henceforth P_n denotes the n^{th} empirical probability measure for P .

Theorem A [Yukich (1985)] Let $g \in L^1(\mathbb{R}^k) \cap C_0(\mathbb{R}^k)$. The following are equivalent statements for the pseudo-metric d_g :

- (i) d_g metrizes the topology of weak convergence in $P(\mathbb{R}^k)$,
- (ii) the set $\{t: \hat{g}(t) = 0\}$ has empty interior, and
- (iii) $d_g(P, Q) \neq 0$ if $P \neq Q$.

Theorem B [Yukich (1985)] Let $g \in L^1(\mathbb{R}^k) \cap C_0(\mathbb{R}^k)$ be a decreasing function of $\|\cdot\|$. Then for all $P \in \mathcal{P}(\mathbb{R}^k)$ we have

$$(0.1) \quad n^{1/2} d_g(P_n, P) \xrightarrow{D} \sup_{x \in \mathbb{R}^k} |G_P(g(x-\cdot))|,$$

where G_P is a mean zero Gaussian process indexed by the translates of g .

Of course, by drawing on the results of Dudley and Philipp (1983), one could easily deduce a bounded LIL result from (0.1), but we will not pursue this here. Likewise, from the theory of the function indexed empirical process, one can also obtain refined exponential bounds for $n^{1/2} d_g(P_n, P)$. Also, the above results can be easily extended to the group setting; see Yukich (1987).

Keeping in mind that we are free to choose from among a variety of kernels g , it is sometimes relatively easy to calculate $d_g(P, Q)$. From a statistical viewpoint, this is important. For example, if P is a uniform probability measure then $d_g(P_n, P)$ may be easily evaluated if g is the bilateral exponential kernel. For more on this, see Yukich (1987). In this context we should mention that other well known metrics, e.g. the Prokhorov and dual bounded Lipschitz metrics, are relatively difficult to calculate and moreover, do not always satisfy the general CLT result (0.1); see the recent results of Giné and Zinn (1987).

This article, which may be regarded as an extension of previous work surrounding convolution metrics, shows that certain modified convolution metrics are useful in another context: they are extremely appropriate for providing rates of convergence (with respect to a variety of metrics) in the general CLT. In fact, some of our results will hold in the Banach space setting. We will also examine the close relationship between convolution metrics and the Kantorovich - Wasserstein distance. Thus, convolution metrics enjoy a variety of useful properties, making their possible statistical use seem

especially attractive.

Let us be more precise about the contents of this article. Let $(B, \|\cdot\|)$ be a complete separable Banach space equipped with the usual Borel sets $\mathcal{B}$ and let $\mathcal{X} := \mathcal{X}(B)$ be the vector space of all random variable defined on a probability space $(\Omega, \mathcal{Q}, P)$ and taking values in B . We will choose to work with metrics on the space $\mathcal{X}$ instead of the space $\mathcal{P}(B)$. We will show that certain convolution metrics on $\mathcal{X}$ may be used to provide improved rates of convergence of normalized sums to a stable limit law. The rates of convergence, which hold uniformly in n , are expressed in terms of a variety of uniform metrics on $\mathcal{X}$. In the Banach space setting the convergence rates hold with respect to the total variation metric and represent both an improvement and an extension of existing results. Even in the classical Euclidean space setting our approach provides improved rates of convergence and it also allows a determination of convergence rates with respect to uniform metrics between density and characteristic functions.

§1. Definitions, notation and terminology

A mapping $\mu: \mathcal{X} \times \mathcal{X} \rightarrow [0, \infty]$ is called an ideal probability metric of order $r \in \mathbb{R}$ if for any random variables $X_1, X_2, Z \in \mathcal{X}$ and any non-zero constant c the following two properties are satisfied [cf. Zolotarev(1976)]:

(i) Regularity: $\mu(X_1+Z, X_2+Z) \leq \mu(X_1, X_2)$, and

(ii) Homogeneity of order r :

$$\mu(cX_1, cX_2) = |c|^r \mu(X_1, X_2).$$

When μ is a simple metric, i.e., its values are determined by the marginal distributions of the random variables being compared, then it is assumed in addition that the random variable Z is independent of X_1 and X_2 in condition (i). All metrics μ in this article are simple.

Zolotarev (1976) showed the existence of an ideal metric of a given

order $r \geq 0$ and he defined the ideal metric

$$(1.1) \quad \zeta_r(X_1, X_2) := \sup\{|\mathbb{E}(f(X_1) - f(X_2))| : |f^{(m)}(x) - f^{(m)}(y)| \leq \|x - y\|^\beta\},$$

where $m \in \mathbb{N}^+$ and $\beta \in (0, 1]$ satisfy $m + \beta = r$ and $f^{(m)}$ denotes the m^{th} Fréchet derivative of f for $m \geq 0$ and $f^{(0)}(x) = f(x)$. He also

obtained an upper bound for ζ_k , $k \in \mathbb{N}^+$, in terms of the so-called difference pseudomoment κ_k , where for $r > 0$

$$\kappa_r(X_1, X_2) := \sup\{|\mathbb{E}(f(X_1) - f(X_2))| : |f(x) - f(y)| \leq \|x\|^{r-1} \|y\|^{r-1}\}.$$

If $B = \mathbb{R}$, $\|x\| = |x|$, then

$$(1.2) \quad \kappa_r(X_1, X_2) := r \int_{-\infty}^{\infty} |x|^{r-1} |F_{X_1} - F_{X_2}| dx, \quad r > 0,$$

and where F_X denotes the distribution function for X .

In this article we introduce and study two ideal metrics of convolution type on the space $\mathfrak{X}$. In addition to their ideality and convolution structure, these metrics have the following useful and special properties:

(P1) they have upper bounds which can be explicitly calculated in terms of the so-called difference pseudomoments, and

(P2) they have a weaker uniform structure than most other ideal metrics and thus yield better rates of convergence.

These ideal metrics will be used to provide improved convergence rates for convergence to an α -stable random variable in the Banach space setting. Moreover, the rates will hold with respect to a variety of uniform metrics on $\mathfrak{X}$.

More precisely, letting $X, X_1, X_2, \dots$ denote i.i.d. random variables and Y_α denote an α -stable random variable we use ideal metrics to describe the rate of convergence

$$(1.3) \quad \frac{X_1 + \dots + X_n}{n^{1/\alpha}} \xrightarrow{D} Y_\alpha$$

with respect to the following uniform metrics on $\mathfrak{X}$.

Distance in variation metric:

$$(1.4) \quad \begin{aligned} \sigma(X_1, X_2) &:= \sup_{A \in \mathfrak{B}} |P\{X_1 \in A\} - P\{X_2 \in A\}|, \quad X_1, X_2 \in \mathfrak{X}(B), \\ &:= \sup\{|Ef(X_1) - Ef(X_2)| : f: B \rightarrow \mathbb{R} \text{ is measurable} \\ &\quad \text{and } \forall x, y \in B \quad |f(x) - f(y)| \leq I(x, y) \text{ where} \\ &\quad I(x, y) = 1 \text{ if } x \neq y \text{ and } 0 \text{ otherwise}\}, \end{aligned}$$

Total variation metric:

$$(1.5) \quad \begin{aligned} \text{Var}(X_1, X_2) &:= \ell_1(X_1, X_2) := \sup\{|Ef(X_1) - Ef(X_2)| : f: B \rightarrow \mathbb{R} \text{ is} \\ &\quad \text{measurable and } \|f\|_\infty := \sup_{x \in B} |f(x)| \leq 1\} \\ &= 2\sigma(X_1, X_2), \quad X_1, X_2 \in \mathfrak{X}(B). \end{aligned}$$

In $\mathfrak{X}(\mathbb{R}^n)$ we have $\text{Var}(X_1, X_2) := \int |d(F_{X_1} - F_{X_2})|$.

Uniform metric between densities (p_X denotes the density for $X \in \mathfrak{X}(\mathbb{R}^k)$):

$$(1.6) \quad \ell(X_1, X_2) := \text{esssup}_x |p_{X_1}(x) - p_{X_2}(x)|.$$

Uniform metric between characteristic functions:

$$(1.7) \quad \chi(X_1, X_2) := \sup_t |\varphi_{X_1}(t) - \varphi_{X_2}(t)|,$$

where φ_X denotes the characteristic function of X . The metric χ is topologically weaker than Var , which is itself topologically weaker than ℓ by Scheffé's Theorem; see Billingsley (1968), p. 224.

The convergence rates with respect to σ will hold in the Banach space setting, thereby extending and generalizing results of Zolotarev (1976, 1977), Senatov (1980) and Paulauskas (1973, 1976) who consider uniform rates in terms of ζ_r and κ_r as well as stronger metrics.

Even in the special case $B = \mathbb{R}$, our results improve upon those of Senatov (1980) and Paulauskas (1973, 1976), primarily because of the special property (P2). Our results with respect to the χ and ℓ metrics, which seem to be the first of their kind, will follow from the easy applicability of the methods used for the Var metric. We note that not only does our method enjoy wide applicability to a variety of situations, but it is also remarkably simple.

Following this brief introductory remark, let us describe the contents of this article. This section concludes with notational remarks and section two discusses ideal convolution metrics and their properties, especially (P1) and (P2).

In section three we use an ideal convolution metric to obtain the rate of convergence in (1.3) in terms of the metric Var in the Banach space setting. Sections four and five illustrate the wide applicability of our method and describe rates in terms of the uniform metrics χ and ℓ , respectively. Here, the unifying theme is that ideal metrics, especially those of convolution type, provide improved convergence rates in (1.3) with respect to various uniform metrics, e.g., Var , χ , or ℓ .

Section six shows that the new ideal convolution metrics metrize the convergence in distribution of random variables and section seven contains concluding remarks.

Notation and Terminology

For each $X_1, X_2 \in \mathfrak{X}$ we write $X_1 + X_2$ to mean the sum of

independent random variables with laws P_{X_1} and P_{X_2} , respectively. For any $X \in \mathcal{X}$, p_X denotes the density of X if it exists. We reserve the letter Y_α (or Y) to denote a strictly stable symmetrical random variable with parameter $\alpha \in (0, 2]$, i.e., $Y_\alpha \stackrel{D}{=} -Y_\alpha$ and for any $n = 1, 2, \dots$, $X'_1 + \dots + X'_n \stackrel{D}{=} n^{1/2} Y_\alpha$, where $X'_1, X'_2, \dots, X'_n$ are i.i.d. random variables with the same distribution as Y_α . If $Y_\alpha \in \mathcal{X}(\mathbb{R})$ we assume that Y_α has characteristic function

$$\varphi_Y(t) = \exp(-|t|^\alpha), \quad t \in \mathbb{R}.$$

For any $f : B \rightarrow \mathbb{R}$, $\|f\|_L := \sup_{x \neq y} \frac{|f(x) - f(y)|}{\|x - y\|}$ denotes the Lipschitz norm of f , $\|f\|_\infty$ the essential supremum of f , and when $B = \mathbb{R}^k$, $\|f\|_p$ denotes the L^p norm.

We will use the following metrics on $\mathcal{X}$.
Kolmogorov metric:

$$(1.8) \quad \rho(X_1, X_2) := \sup_{x \in \mathbb{R}} |F_{X_1}(x) - F_{X_2}(x)|, \quad X_1, X_2 \in \mathcal{X}(\mathbb{R}).$$

Weighted χ metric:

$$(1.9) \quad \chi_r(X_1, X_2) := \sup_{t \in \mathbb{R}} |t|^{-r} |\varphi_{X_1}(t) - \varphi_{X_2}(t)|,$$

L^p version of ζ_m :

$$(1.10) \quad \zeta_{m,p}(X_1, X_2) := \sup\{|\mathbb{E}(f(X_1) - f(X_2))| : \|f^{(m+1)}\|_q \leq 1\}, \quad \frac{1}{p} + \frac{1}{q} = 1, \\ m = 0, 1, 2, \dots$$

(If $\zeta_{m,p}(X_1, X_2) < \infty$ then $\zeta_{m,p}(X_1, X_2) = \left\| \int_{-\infty}^{\bullet} \frac{(\bullet - t)^m}{m!} d(F_{X_1}(t) - F_{X_2}(t)) \right\|_p$).

Generalized Kantorovich-Wasserstein metric:

$$(1.11) \quad W_p^p(X_1, X_2) := \sup \left\{ \left| \int f dF_{X_1} + \int g dF_{X_2} \right| : \|f\|_{\infty} + \|g\|_L < \infty, \right.$$

$$\left. \|g\|_{\infty} + \|g\|_L < \infty \text{ and } f(x) + g(y) \leq \|x-y\|^p \quad \forall x, y \in B \right\}, \quad p \geq 1.$$

§2. Ideal convolution metrics and their properties

Let $\theta \in \mathfrak{X}(\mathbb{R}^k)$ and define for every $r > 0$ the convolution metric

$$(2.1) \quad \mu_{\theta,r}(X_1, X_2) := \sup_{h \in \mathbb{R}} |h|^r \ell(X_1 + h\theta, X_2 + h\theta) \quad X_1, X_2 \in \mathfrak{X}(\mathbb{R}^k).$$

Thus, each random variable θ generates a metric $\mu_{\theta,r}$, $r > 0$. When $\theta \in \mathfrak{X}(B)$ we will also consider convolution metrics of the form [cf. Rachev and Ignatov (1984)]:

$$\nu_{\theta,r}(X_1, X_2) := \sup_{h \in \mathbb{R}} |h|^r \ell_1(X_1 + h\theta, X_2 + h\theta) \quad X_1, X_2 \in \mathfrak{X}(B).$$

Lemmas 2.1 and 2.2 below show that $\mu_{\theta,r}$ and $\nu_{\theta,r}$ are ideal of order $r-1$ and r , respectively. In general, $\mu_{\theta,r}$ and $\nu_{\theta,r}$ are actually only pseudo-metrics (cf. Theorem A), but this distinction is not of importance in what follows and so we omit it.

When θ is an α -stable random variable we will write $\mu_{\alpha,r}$ and $\nu_{\alpha,r}$ (or simply μ_r and ν_r when it is understood) in place of $\mu_{\theta,r}$ and $\nu_{\theta,r}$. Also, if θ has a density g , then the metric $\mu_{\theta,r}$ represents a generalization of the convolution metric

$$d_g(X_1, X_2) := \ell(X_1 + h\theta, X_2 + h\theta)$$

described in the introduction.

One of the central themes of this article is that every ideal metric on $\mathcal{X}$ can be used to provide convergence rates in (1.3) in terms of some uniform metric (e.g. Var , χ or ℓ) corresponding to the proposed ideal metric.

For example, we show that μ_r can be used to describe the convergence rate in (1.3) with respect to the uniform metric Var . The method for the ideal metric μ_r is strikingly simple and yet general enough to handle other ideal metrics of non-convolution type: for example, the ideal metric χ_r describes the rate of convergence in (1.3) with respect to the uniform metric χ ; also μ_r and ν_r , when taken together, describe the convergence rate in terms of the uniform metric ℓ . There are few published results concerning convergence rates in terms of ℓ and χ ; as for the latter metric, Banys (1976) has found convergence rates which do not hold uniformly over $\mathbb{R}$, but only over intervals increasing with the n , the sample size.

We note that Zolotarev (1976, 1977) and Senatov (1980) use the metric ζ_r to develop the convergence rate in (1.3) with $\alpha = 2$ in terms of Var and ρ . Even in this special case our results are sharper and more refined. Zolotarev and Rachev (1984) and Omev and Rachev (1987) use the ideal weighted Kolmogorov metric

$$\rho_r(X_1, X_2) := \sup_{x \in \mathbb{R}} |x|^r |F_{X_1}(x) - F_{X_2}(x)|, \quad r > 1$$

to obtain the rate of convergence for the normalized maxima

$$n^{-1} \max(X_1, \dots, X_n) \xrightarrow{D} U,$$

where $P\{U < x\} = \exp\{-x^{-1}\}$, $x \geq 0$.

The remainder of this section describes the special properties of the ideal convolution (or smoothing) metrics $\mu_{\theta,r}$ and $\nu_{\theta,r}$. We first verify ideality.

Lemma 2.1. For all $\theta \in \mathfrak{X}$ and $r > 0$, $\mu_{\theta,r}$ is an ideal metric of order $r-1$.

Proof. If Z does not depend upon X_1 and X_2 then

$$\ell(X_1+Z, X_2+Z) \leq \ell(X_1, X_2) ,$$

and hence $\mu_{\theta,r}(X_1+Z, X_2+Z) \leq \mu_{\theta,r}(X_1, X_2)$. Additionally, for any $c \neq 0$

$$\begin{aligned} \mu_{\theta,r}(cX_1, cX_2) &= \sup_{h \in \mathbb{R}} |h|^r \ell(cX_1+h\theta, cX_2+h\theta) \\ &= \sup_{h \in \mathbb{R}} |ch|^r \ell(cX_1+ch\theta, cX_2+ch\theta) \\ &= |c|^{r-1} \mu_{\theta,r}(X_1, X_2) . \end{aligned}$$

Q.E.D.

The proof of the next lemma is analogous to the one above.

Lemma 2.2. For all $\theta \in \mathfrak{X}$ and $r > 0$ $\nu_{\theta,r}$ is an ideal metric of order r .

We now turn to special property (P1) and show that both $\mu_{\theta,r}$ and $\nu_{\theta,r}$ are bounded above by the difference pseudomoment whenever θ has a density which is smooth enough.

Lemma 2.3. Let $k \in \mathbb{N}^+$ and suppose that $X, Y \in \mathfrak{X}(\mathbb{R})$ satisfy $EX^j = EY^j$, $j = 1, \dots, k-2$. Then for every $\theta \in \mathfrak{X}(\mathbb{R})$ with a density g which

is $k-1$ times differentiable,

$$(2.3) \quad \mu_{\theta,k}(X_1, X_2) \leq \frac{\|g^{(k-1)}\|_{\infty}}{(k-1)!} \kappa_{k-1}(X_1, X_2),$$

where $\|\bullet\|_{\infty} := \text{ess sup}_{x \in \mathbb{R}} (\bullet)$.

Proof. In view of the inequality [Zolotarev (1979)]

$$(2.4) \quad \varsigma_{k-1}(X_1, X_2) \leq \frac{1}{(k-1)!} \kappa_{k-1}(X_1, X_2),$$

it suffices to show that

$$(2.5) \quad \mu_{\theta,k}(X_1, X_2) \leq \varsigma_{k-1}(X_1, X_2).$$

But

$$\begin{aligned} \mu_{\theta,k}(X_1, X_2) &= \sup_{h \in \mathbb{R}} |h|^k \sup_{x \in \mathbb{R}} \frac{1}{|h|} \left| \int_{-\infty}^{\infty} g\left(\frac{x-y}{h}\right) d(F_{X_1}(y) - F_{X_2}(y)) \right| \\ &= \sup_{h \in \mathbb{R}} |h|^{k-1} \sup_{x \in \mathbb{R}} \left| \int_{-\infty}^{\infty} (F_{X_1}(y) - F_{X_2}(y)) g^{(1)}\left(\frac{x-y}{h}\right) \frac{1}{h} dy \right| \\ &= \sup_{h \in \mathbb{R}} |h|^{k-2} \sup_{x \in \mathbb{R}} \left| \int_{-\infty}^{\infty} g^{(1)}\left(\frac{x-y}{h}\right) dF_{X_1}^{(-1)}(y) - F_{X_2}^{(-1)}(y) \right| \\ &\vdots \\ &= \sup_{h \in \mathbb{R}} |h| \sup_{x \in \mathbb{R}} \left| \int_{-\infty}^{\infty} g^{(k-1)}\left(\frac{x-y}{h}\right) \frac{1}{h} (F_{X_1}^{(-k+2)}(y) - F_{X_2}^{(-k+2)}(y)) dy \right|. \end{aligned}$$

where

$$(2.6) \quad F_X^{(-k+1)}(x) := \int_{-\infty}^x \frac{(x-t)^{k-1}}{(k-1)!} dF_X(t) .$$

Therefore, by (1.10) and $\zeta_{k-1} = \zeta_{k-2,1}$, we have

$$\begin{aligned} \mu_{\theta,k}(X_1, X_2) &\leq \|g^{(k-1)}\|_{\infty} \int_{-\infty}^{\infty} |F_{X_1}^{(2-k)}(y) - F_{X_2}^{(2-k)}(y)| dy \\ &= \|g^{(k-1)}\|_{\infty} \zeta_{k-1}(X_1, X_2) . \end{aligned}$$

Q.E.D.

Analogously, Lemmas 2.5 and 2.6 below show that $\nu_{\theta,k}$ is bounded by the difference pseudomoments.

Under similar hypotheses we may also show that the smoothing metrics $\mu_{\theta,k}$ and $\nu_{\theta,k}$ are weaker than $\zeta_{k,p}$. This, together with inequality (2.5), helps illustrate property (P2).

Lemma 2.4. For every $\theta \in \mathcal{X}(\mathbb{R})$ with a density g which is m times differentiable and for all $X_1, X_2 \in \mathcal{X}(\mathbb{R})$

$$\mu_{\theta,r}(X_1, X_2) \leq C(m, p, g) \zeta_{m-1,p}(X_1, X_2) ,$$

where $r = m + \frac{1}{p}$, $m \in \mathbb{N}^+$, and

$$(2.7) \quad C(m, p, g) := \|g^{(m)}\|_q , \quad \frac{1}{p} + \frac{1}{q} = 1 .$$

Lemma 2.5. Under the hypotheses of Lemma 2.4 we have

$$(2.8) \quad \nu_{\theta,r}(X_1, X_2) \leq C(m, p, g) \zeta_{m-1,p}(X_1, X_2) ,$$

where $r = m - 1 + \frac{1}{p}$, $m \in \mathbb{N}^+$ and $C(m,p,g)$ is as in (2.7).

Lemma 2.6. [cf. Theorem 2 of Maejima and Rachev (1987)]. Let $m \in \mathbb{N}^+$ and suppose $E(X_1^j - X_2^j) = 0$ $j = 0, 1, \dots, m-1$. Then for $p \in [1, \infty)$

$$(2.9) \quad \zeta_{m,p}(X_1, X_2) \leq \begin{cases} \kappa_1^{1/p}(X_1, X_2) & m = 0, \\ \frac{\Gamma(1+\frac{1}{p})}{\Gamma(r)} \kappa_r(X_1, X_2) & r = m + \frac{1}{p}, m = 1, 2, \dots \end{cases}$$

The proofs of the above three lemmas follow from straightforward modifications of the techniques used in Ignatov and Rachev (1983), Maejima and Rachev (1987) and Zolotarev (1979). The details are left to the reader.

§3. Rates of convergence in the total variation metric

In this section we develop rates of convergence with respect to the Var metric.

Throughout we suppose that $X, X_1, X_2, \dots$ denotes a sequence of i.i.d. random variables in $\mathfrak{X}(B)$, where B is a separable Banach space. $Y \in \mathfrak{X}(B)$ denotes a strictly α -stable random variable. The ideal convolution metric $\nu_r := \nu_{\alpha,r}$ (i.e., $\theta = Y$) will play a central role.

Our main theorem is

Theorem 3.1. Let Y be an α -stable random variable. Let

$$r = s + \frac{1}{p} > \alpha \text{ for some integer } s \text{ and } p \in [1, \infty), a = \frac{1}{2^{r/\alpha} A},$$

and $A := 2(2^{r/\alpha-1} + 3^{r/\alpha})$. If $X \in \mathfrak{X}(B)$ satisfies

$$(3.1) \quad \tau_0 := \tau_0(X, Y) := \max(\text{Var}(X, Y), \nu_{\alpha, r}(X, Y)) \leq a ,$$

then $\forall n \geq 1$

$$\text{Var} \left(\frac{X_1 + \dots + X_n}{n^{1/\alpha}}, Y \right) \leq A(a) \tau_0 n^{1-r/\alpha} \leq 2^{-r/\alpha} n^{1-r/\alpha} .$$

Remarks

- (i) A result of this type was proved by Senatov (1980) for the special case $B = \mathbb{R}^k$, $s = 3$, and $\alpha = 2$ via the ζ_r metric. We will follow Senatov's method with some refinements.
- (ii) Theorem 3.1 is optimal in the sense that the power of n is the smallest possible; i.e., the exponent $1-r/\alpha$ cannot be decreased. This follows, for example, from Theorem 3.4.1 of Ibragimov and Linnik (1971) and the inequality $\rho \leq \sigma = \frac{1}{2} \text{Var}$ in $\mathfrak{X}(\mathbb{R})$.

Before proving Theorem 3.1 we need a few auxiliary results.

Lemma 3.2. For any $X_1, X_2 \in \mathfrak{X}(B)$ and $\sigma > 0$

$$\text{Var}(X_1 + \sigma Y, X_2 + \sigma Y) \leq \sigma^{-r} \nu_r(X_1, X_2) .$$

Proof. Since Y and $(-Y)$ have the same distribution

$$\nu_r(X_1, X_2) = \sup_{h>0} h^r \ell_1(X_1 + hY, X_2 + hY)$$

and thus

$$\ell_1(X_1 + hY, X_2 + hY) \leq h^{-r} \sup_{h>0} h^r \ell_1(X_1 + hY, X_2 + hY)$$

$$= h^{-r} \nu_r(X_1, X_2) .$$

Q.E.D.

Lemma 3.2 closely resembles Lemma 1 of Senatov (1980) for the metric ζ_r . The next result resembles Lemma 2 of Senatov (1980) proved for $B = \mathbb{R}^k$. We note that estimates of this sort have been used by Sazonov (1972) and Sazonov and Ul'yanov (1979).

Lemma 3.3. For any $X_1, X_2, U, V \in \mathfrak{X}(B)$ the following inequality holds:

$$\text{Var}(X_1+U, X_2+U) \leq \text{Var}(X_1, X_2)\text{Var}(U, V) + \text{Var}(X_1+V, X_2+V) .$$

Proof. By the definition (1.5) and the triangle inequality

$$\begin{aligned} \text{Var}(X_1+U, X_2+U) &= \sup\{|\text{Ef}(X_1+U) - \text{Ef}(X_2+U)| : \|f\|_{\infty} \leq 1\} \\ &= \sup\{|\int_B f(u)(P_{X_1+U} - P_{X_2+U}) du| : \|f\|_{\infty} \leq 1\} \\ &\leq \sup\{|\int_B \bar{f}(x)(P_{X_1} - P_{X_2}) dx| : \|f\|_{\infty} \leq 1\} + \\ &\quad + \text{Var}(X_1+U, X_2+V), \end{aligned}$$

where

$$\bar{f}(x) := \int_B f(u)(P_U - P_V)(du - x) = \int_B f(u+x)(P_U - P_V) du,$$

and where P_X denotes the law of the random variable X . Since

$\|f\|_{\infty} \leq 1$ then

$$\begin{aligned} \|\bar{f}\|_{\infty} &= \sup_{x \in B} \left| \int_B f(u+x)(P_U - P_V) du \right| \\ &\leq \left| \int_B f(u)(P_U - P_V) du \right| \end{aligned}$$

$$\leq \text{Var}(U, V), \text{ by (1.5),}$$

and thus $\sup\left\{\left|\int_B \bar{f}(x)(P_{X_1} - P_{X_2})dx\right| : \|f\|_\infty \leq 1\right\}$ is bounded by

$$\begin{aligned} &\leq \sup\left\{\left|\int_B g(x)(P_{X_1} - P_{X_2})dx\right| : \|g\|_\infty \leq \text{Var}(U, V)\right\} \\ &= \text{Var}(X_1, X_2) \text{Var}(U, V). \end{aligned}$$

Q.E.D.

We now prove Theorem 3.1; throughout, $Y_1, Y_2, \dots$ denote i.i.d. copies of Y .

Proof. We proceed by induction; for $n = 1$ the assertion of the theorem is trivial. For $n = 2$, the assertion follows from the inequality

$$\begin{aligned} \text{Var}\left(\frac{X_1 + X_2}{2^{1/\alpha}}, Y\right) &= \text{Var}\left(\frac{X_1 + X_2}{2^{1/\alpha}}, \frac{Y_1 + Y_2}{2^{1/\alpha}}\right) \\ &= \text{Var}(X_1 + X_2, Y_1 + Y_2) \\ &\leq 2\text{Var}(X_1, Y_1) \\ &\leq A(a)\tau_0 2^{1-r/\alpha} \end{aligned}$$

since $A(a) \geq 2^{r/\alpha}$. A similar calculation holds for $n = 3$. Suppose now that the estimate

$$(3.2) \quad \text{Var}\left(\frac{X_1 + \dots + X_j}{j^{1/\alpha}}, Y\right) \leq A(a)\tau_0 j^{1-r/\alpha}$$

holds for all $j < n$. To complete the induction we only need to show that (3.2) holds for $j = n$.

Thus, assuming (3.2) we have by (3.1)

$$(3.3) \quad \text{Var}\left(\frac{X_1 + \dots + X_j}{j^{1/\alpha}}, Y\right) \leq A(a)a = 2^{-r/\alpha}.$$

For any integer $n \geq 4$ and $m = [\frac{n}{2}]$, where $[\cdot]$ denotes integer part, the triangle inequality gives

$$\begin{aligned} V &:= \text{Var}\left(\frac{X_1 + \dots + X_n}{n^{1/\alpha}}, Y\right) \\ &= \text{Var}\left(\frac{X_1 + \dots + X_m}{n^{1/\alpha}}, \frac{Y_1 + \dots + Y_n}{n^{1/\alpha}}\right) \\ &\leq \text{Var}\left(\frac{X_1 + \dots + X_m}{n^{1/\alpha}} + \frac{X_{m+1} + \dots + X_n}{n^{1/\alpha}}, \frac{Y_1 + \dots + Y_m}{n^{1/\alpha}} + \frac{X_{m+1} + \dots + X_n}{n^{1/\alpha}}\right) \\ &\quad + \text{Var}\left(\frac{Y_1 + \dots + Y_m}{n^{1/\alpha}} + \frac{X_{m+1} + \dots + X_n}{n^{1/\alpha}}, \frac{Y_1 + \dots + Y_m}{n^{1/\alpha}} + \frac{Y_{m+1} + \dots + Y_n}{n^{1/\alpha}}\right) \end{aligned}$$

Hence, by Lemma 3.3

$$(3.4) \quad V \leq I_1 + I_2 + I_3,$$

where

$$\begin{aligned} I_1 &:= \text{Var}\left(\frac{X_1 + \dots + X_m}{n^{1/\alpha}}, \frac{Y_1 + \dots + Y_m}{n^{1/\alpha}}\right) \text{Var}\left(\frac{X_{m+1} + \dots + X_n}{n^{1/\alpha}}, \frac{Y_{m+1} + \dots + Y_n}{n^{1/\alpha}}\right) \\ I_2 &:= \text{Var}\left(\frac{X_1 + \dots + X_m}{n^{1/\alpha}} + \frac{Y_{m+1} + \dots + Y_n}{n^{1/\alpha}}, \frac{Y_1 + \dots + Y_m}{n^{1/\alpha}} + \frac{Y_{m+1} + \dots + Y_n}{n^{1/\alpha}}\right), \text{ and} \end{aligned}$$

$$I_3 := \text{Var}\left(\frac{Y_1 + \dots + Y_m}{n^{1/\alpha}} + \frac{X_{m+1} + \dots + X_n}{n^{1/\alpha}}, \frac{Y_1 + \dots + Y_m}{n^{1/\alpha}} + \frac{Y_{m+1} + \dots + Y_n}{n^{1/\alpha}}\right).$$

We first estimate I_1 . By (3.3) and (3.2)

$$\begin{aligned} I_1 &\leq 2^{-r/\alpha} A(a) \tau_0 (n-m)^{1-r/\alpha} \\ (3.5) \quad &\leq \frac{1}{2} A(a) \tau_0 n^{1-r/\alpha}. \end{aligned}$$

In order to estimate I_2 and I_3 we will use Lemma 3.2 and the relation

$$(3.6) \quad \frac{Y_1 + \dots + Y_n}{n^{1/\alpha}} \stackrel{D}{=} Y_1.$$

Thus, by (3.6), Lemma 3.2, and the fact that ν_r is ideal of order r we deduce

$$\begin{aligned} I_2 &= \text{Var}\left(\frac{X_1 + \dots + X_m}{n^{1/\alpha}} + \left(\frac{n-m}{n}\right)^{1/\alpha} Y, \frac{Y_1 + \dots + Y_m}{n^{1/\alpha}} + \left(\frac{n-m}{n}\right)^{1/\alpha} Y\right) \\ &\leq \left(\frac{n-m}{n}\right)^{-r/\alpha} \nu_r\left(\frac{X_1 + \dots + X_m}{n^{1/\alpha}}, \frac{Y_1 + \dots + Y_m}{n^{1/\alpha}}\right) \\ &\leq 2^{r/\alpha} m \nu_r\left(\frac{X_1}{n^{1/\alpha}}, \frac{Y_1}{n^{1/\alpha}}\right) \\ (3.7) \quad &\leq 2^{r/\alpha-1} n^{1-r/\alpha} \nu_r(X_1, Y_1). \end{aligned}$$

Analogously, we estimate I_3 by

$$\begin{aligned}
 I_3 &= \text{Var}\left(\frac{X_1 + \dots + X_{n-m}}{n^{1/\alpha}} + \left(\frac{m}{n}\right)^{1/\alpha} Y, \frac{Y_1 + \dots + Y_{n-m}}{n^{1/\alpha}} + \left(\frac{m}{n}\right)^{1/\alpha} Y\right) \\
 &\leq \left(\frac{m}{n}\right)^{-r/\alpha} \nu_r\left(\frac{X_1 + \dots + X_{n-m}}{n^{1/\alpha}}, \frac{Y_1 + \dots + Y_{n-m}}{n^{1/\alpha}}\right) \\
 (3.8) \quad &\leq 3^{r/\alpha} n^{1-r/\alpha} \nu_r(X_1, Y_1).
 \end{aligned}$$

Taking (3.4), (3.5), (3.7) and (3.8) into account we obtain

$$\begin{aligned}
 V &\leq \left(\frac{1}{2} A(a) + 2^{r/\alpha-1} + 3^{r/\alpha}\right) \tau_0 n^{1-r/\alpha} \\
 &\leq A(a) \tau_0 n^{1-r/\alpha},
 \end{aligned}$$

since $A(a)/2 = 2^{r/\alpha-1} + 3^{r/\alpha}$.

Q.E.D.

§4. Rates of convergence in the χ metric

In this section we develop rates of convergence in (1.3) with respect to the χ metric; our purpose here is to show that the methods of proof for Theorem 3.1 can be easily extended to deduce analogous results with respect to χ . The metric χ_r will play a role analogous to that played by ν_r in Section 3.

Throughout, $X, X_1, X_2, \dots$ denotes a sequence of i.i.d. copies of an α -stable random variable.

Our main theorem is

Theorem 4.1. Let Y be an α -stable random variable in $\mathfrak{X}(\mathbb{R})$.

Let $r > \alpha$, $b := \frac{1}{2^{r/\alpha} B}$, and $B := \max(3^{r/\alpha}, 2C_r(2^{r/\alpha-1} + 3^{r/\alpha}))$ where

$C_r := (\frac{r}{\alpha e})^{r/\alpha}$. If $X \in \mathcal{X}(\mathbb{R})$ satisfies

$$(4.1) \quad \tau_r := \tau_r(X, Y) := \max\{\chi(X, Y), \chi_r(X, Y)\} \leq b,$$

then for all $n \geq 1$

$$(4.2) \quad \chi\left(\frac{X_1 + \dots + X_n}{n^{1/\alpha}}, Y\right) \leq B \tau_r n^{1-r/\alpha} = 2^{-r/\alpha} n^{1-r/\alpha}.$$

Remarks

- (i) In comparing conditions (3.1) and (4.1) it is useful to note that the metric χ is topologically weaker than Var , i.e., $\text{Var}(X_n, Y) \rightarrow 0$ implies $\chi(X_n, Y) \rightarrow 0$ but not conversely. Also, it is easy to show that if $r = m + \beta$, $m = 0, 1, \dots$, $\beta \in (0, 1)$ then

$$(4.3) \quad \chi_r \leq C_\beta \zeta_r \quad \text{where } C_\beta := \sup_t |t^{-\beta}(1 - e^{it})|;$$

if $r = m$, $m = 1, 2, \dots$ then $\chi_r \leq \zeta_r$.

- (ii) Actually, one may show that for $r \in \mathbb{N}^+$ the metric χ_r has a convolution type structure. In fact, with a slight abuse of notation,

$$\chi_r(F_{X_1}, F_{X_2}) = \chi(F_{X_1} * p_r, F_{X_2} * p_r),$$

where $p_r(t) = t^r/r! I_{\{t>0\}}$ is the density of an unbounded positive measure on the half line $[0, \infty)$.

- (iii) As in Theorem 3.1, the exponent $1-r/\alpha$ cannot be reduced; this follows from Theorem 3.4.1 of Ibragimov and Linnik (1971) and

the fact that χ convergence implies ρ convergence.

- (iv) As noted earlier, Banys (1976) has obtained a result similar to Theorem 4.1: his result is weaker since it only considers the sup norm difference between characteristic functions over finite intervals depending on n . Additionally, his result is expressed in terms of the so-called r^{th} absolute pseudomoment

$$\bar{\nu}_r(X, Y) := \int |x|^r |p_X(x) - p_Y(x)| dx ;$$

since χ_r is topologically weaker than $\bar{\nu}_r$ (i.e., $\chi_r \leq \text{const } \zeta_r \leq \text{const } \bar{\nu}_r$ and χ_r -convergence does not, in general, imply $\bar{\nu}_r$ -convergence), our estimate (4.2) is clearly more refined, even over finite intervals.

The proof of Theorem 4.1 is very similar to that of Theorem 3.1 and uses the following auxiliary results, which are completely analogous to Lemmas 3.2 and 3.3. We leave the details to the reader.

Lemma 4.2. For any $X_1, X_2 \in \mathcal{X}(\mathbb{R})$, $\sigma > 0$, and $r > \alpha$

$$\chi(X_1 + \sigma Y, X_2 + \sigma Y) \leq C_r \sigma^{-r} \chi_r(X_1, X_2) ,$$

where $C_r := \left(\frac{r}{\alpha e}\right)^{r/\alpha}$.

Proof. We have

$$\chi(X_1 + \sigma Y, X_2 + \sigma Y) := \sup_{t \in \mathbb{R}} |\varphi_{X_1}(t) - \varphi_{X_2}(t)| \varphi_{\sigma Y}(t)$$

$$\begin{aligned}
 &= \sup_{t \in \mathbb{R}} |\varphi_{X_1}(t) - \varphi_{X_2}(t)| \exp\{-|\sigma t|^\alpha\} \\
 &\leq \sup_{t \in \mathbb{R}} |\sigma t|^{-r} |\varphi_{X_1}(t) - \varphi_{X_2}(t)| \sup_{u>0} u^r e^{-u^\alpha} \\
 &= C_r \sigma^{-r} \chi_r(X, Y),
 \end{aligned}$$

since $C_r = \sup_{u>0} u^r e^{-u^\alpha}$ by a simple computation.

Q.E.D.

Lemma 4.3. For any $X_1, X_2, Z, W \in \mathfrak{X}(\mathbb{R})$ the following inequality holds:

$$\chi(X_1+Z, X_2+W) \leq \chi(X_1, X_2) \chi(Z, W) + \chi(X_1+W, X_2+W).$$

Proof. From the inequality

$$\begin{aligned}
 |\varphi_{X_1+Z}(t) - \varphi_{X_2+W}(t)| &\leq |\varphi_{X_1}(t) - \varphi_{X_2}(t)| |\varphi_Z(t) - \varphi_W(t)| + \\
 &\quad + |\varphi_{X_1}(t) - \varphi_{X_2}(t)| |\varphi_W(t)|
 \end{aligned}$$

we obtain the desired result.

Q.E.D.

§5. Rates of convergence in the ℓ metric

In this section we develop convergence rates with respect to the ℓ -metric and thus we naturally restrict attention to the subset $\mathfrak{X}^*$ of $\mathfrak{X}(\mathbb{R}^k)$ of random variables with densities. Throughout $X, X_1, X_2, \dots$ denotes a sequence of i.i.d. random variables in $\mathfrak{X}^*$ and $Y = Y_\alpha$

denotes an α -stable random variable. The ideal convolution metrics $\mu_r := \mu_{\alpha,r}$ and $\nu_r := \nu_{\alpha,r}$ (i.e., $\theta=Y$) will play a central role.

Our main result is

Theorem 5.1. Let Y be an α -stable random variable in $\mathfrak{X}(\mathbb{R}^k)$. Let

$$r = m + \frac{1}{p} > \alpha \text{ for some integer } m \text{ and } p \in [1, \infty), \quad a := \frac{1}{2^{r/\alpha_A}},$$

$$A := 2(2^{r/\alpha-1} + 3^{(r+1)/\alpha}) \quad \text{and} \quad D := 3^{1/\alpha} 2^{r/\alpha}. \quad \text{If } X \in \mathfrak{X}^* \text{ satisfies}$$

$$(i) \quad \tau(X, Y) := \max(\ell(X, Y), \mu_{\alpha,r}(X, Y)) \leq a \quad \text{and} \quad (5.1)$$

$$(ii) \quad \tau_0(X, Y) := \max(\text{Var}(X, Y), \nu_{\alpha,r}(X, Y)) \leq \frac{1}{A(a)D} \leq a,$$

then

$$(5.2) \quad \ell\left(\frac{X_1 + \dots + X_n}{n^{1/\alpha}}, Y\right) \leq A(a) \tau(X, Y) n^{1-r/\alpha}.$$

Remarks.

(i) Conditions (i) and (ii) describe the domain of attraction of a stable Y_α random variable; in fact they guarantee ℓ -closeness (of order $n^{1-r/\alpha}$) between Y and the normalized sums

$$(5.3) \quad n^{-1/\alpha}(X_1 + \dots + X_n).$$

(ii) From property (P2) and especially Lemmas 2.3, 2.5 and 2.6 we know that $\mu_{r+1}(X, Y)$ and $\nu_r(X, Y)$, $r = m-1 + \frac{1}{p}$, $m = 1, 2, \dots$ can be approximated from above by the h difference pseudomoment κ_r whenever X and Y share the same first $(m-1)$ moments. Thus conditions (i)

and (ii) could be expressed in terms of difference pseudomoments, which of course amounts to conditions on the tails of X .

(iii) That (5.2) is of the right order of magnitude may be seen from Theorem 4.5.1 of Ibragimov and Linnik (1971).

To prove Theorem 5.1 we need a few auxiliary results similar in spirit to Lemmas 3.2 and 3.3.

Lemma 5.2. Let $X_1, X_2 \in \mathfrak{X}(\mathbb{R}^k)$. Then

$$\ell(X_1 + \sigma Y, X_2 + \sigma Y) \leq \sigma^{-r} \mu_r(X_1, X_2) .$$

Proof: $\ell(X_1 + \sigma Y, X_2 + \sigma Y) \leq \sigma^{-r} \sigma^r \ell(X_1 + \sigma Y, X_2 + \sigma Y) \leq \sigma^{-r} \mu_r(X_1, X_2) .$

Q.E.D.

Lemma 5.3. For any $X, Y, U, V \in \mathfrak{X}^*(\mathbb{R}^k)$ the following inequality holds:

$$(5.4) \quad \ell(X+U, Y+U) \leq \ell(X, Y) \text{Var}(U, V) + \ell(X+V, Y+V)$$

Proof: Using the triangle inequality we obtain

$$\begin{aligned} \ell(X+U, Y+U) &= \sup_{x \in \mathbb{R}^k} \left| \int (p_X(x-y) - p_Y(x-y)) \Pr(U \in dy) \right| \\ &\leq \sup_{x \in \mathbb{R}^k} \left| \int (p_X(x-y) - p_Y(x-y)) (\Pr\{U \in dy\} - \Pr\{V \in dy\}) \right| + \\ &\quad + \sup_{x \in \mathbb{R}^k} \left| \int (p_X(x-y) - p_Y(x-y)) \Pr\{V \in dy\} \right| \\ &\leq \ell(X, Y) \text{Var}(U, V) + \ell(X+V, Y+V) . \end{aligned}$$

Q.E.D.

To prove Theorem 5.1 one only needs to use the method of proof for Theorem 3.1 combined with the above two auxiliary results. The complete details are left to the reader.

§6. The ideal smoothing metrics and weak convergence

We conclude our discussion of the ideal metrics μ_r and ν_r by showing that they satisfy the same weak convergence properties as do the Kantorovich-Wasserstein distance W_r and the pseudomoments κ_r .

Theorem 6.1 Let $k \in \mathbb{N}^+$, $0 < \alpha \leq 2$, and $X_n, U \in \mathcal{X}(\mathbb{R})$ with $EX_n^j = EU^j \quad \forall j = 1, \dots, k-2$. If k is odd then the following are equivalent as $n \rightarrow \infty$:

- (i) $\mu_{\alpha,k}(X_n, U) \rightarrow 0$,
- (ii) (a) $X_n \xrightarrow{D} U$ and (b) $E|X_n|^{k-1} \rightarrow E|U|^{k-1}$,
- (iii) $W_{k-1}(X_n, U) \rightarrow 0$,
- (iv) $\kappa_{k-1}(X_n, U) \rightarrow 0$, and
- (v) $\nu_{\alpha,k-1}(X_n, U) \rightarrow 0$.

Before proving this we note that (ii) $\Leftrightarrow$ (iii) follows immediately from Theorem 4.1 of Rachev (1984 c), Theorem 1 of Rachev (1984 b), Theorem 2 of Rachev (1984 a) and the identity $EX^{k-1} = E|X|^{k-1}$ for k odd. Also, (ii) $\Leftrightarrow$ (iv) follows from Rachev (1982); (iv) $\Rightarrow$ (i) follows from Lemma 2.3, and (iv) $\Rightarrow$ (v) from Lemmas 2.5 and 2.6; thus the only new result here are the implications (i) $\Rightarrow$ (ii) and (v) $\Rightarrow$ (ii).

Now (i) $\Rightarrow$ (ii)(a) follows easily from Fourier transform arguments since the Fourier transform of g never vanishes. Similarly if (v)

holds then $X_n + Y \xrightarrow{D} U + Y$ and thus (ii)(a) follows. To prove (i) $\Rightarrow$ (ii)(b) we need a lemma.

Lemma 6.2. Let $0 < \alpha \leq 2$ and consider the associated metric $\mu_r := \mu_{r,\alpha}$. For all k there is a constant $\beta := \beta(\alpha, k) < \infty$ such that for all $X, U \in \mathcal{X}(\mathbb{R})$

$$(6.1) \quad \mu_k(X, U) \geq \beta \left| \int_{-\infty}^{\infty} F_X^{(2-k)}(z) - F_U^{(2-k)}(z) dz \right|.$$

Here $F^{(2-k)}$ is as in (2.6).

Proof of Theorem 6.1. Using equality of the first $k-2$ moments and applying (6.1) to X_n and U yields

$$\begin{aligned} \beta^{-1} \mu_k(X_n, U) &\geq \left| \int_{-\infty}^{\infty} \int_{-\infty}^z \frac{(z-t)^{k-2}}{(k-2)!} (dF_{X_n} - dF_U)(t) dz \right| \\ &= \left| \int_{-\infty}^0 (\cdot) dt + \int_0^{\infty} (\cdot) dt \right| \\ (6.2) \quad &:= \left| I_1 + I_2 \right|. \end{aligned}$$

To estimate I_1 and I_2 we first note that since

$$\int_{-\infty}^{\infty} \frac{(z-t)^{k-2}}{(k-2)!} (dF_{X_n} - dF_U)(t) = E(z - X_n)^{k-2} - E(z - U)^{k-2} = 0,$$

we obtain

$$\begin{aligned}
 \int_{-\infty}^{\infty} \frac{(z-t)^{k-2}}{(k-2)!} (dF_{X_n} - dF_U)(t) &= - \int_z^{\infty} \frac{(z-t)^{k-2}}{(k-2)!} (dF_{X_n} - dF_U)(t) \\
 (6.3) \qquad \qquad \qquad &= (-1)^{k-1} \int_z^{\infty} \frac{(t-z)^{k-2}}{(k-2)!} (dF_{X_n} - dF_U)(t).
 \end{aligned}$$

Thus by (6.3) and Fubini's theorem we get

$$\begin{aligned}
 I_2 &= (-1)^{k-1} \int_0^{\infty} \int_z^{\infty} \frac{(t-z)^{k-2}}{(k-2)!} (dF_{X_n} - dF_U)(t) dz \\
 &= (-1)^{k-1} \int_0^{\infty} \int_0^t \frac{(t-z)^{k-2}}{(k-2)!} dz d(F_{X_n} - F_U)(t) \\
 (6.4) \qquad \qquad \qquad &= \int_0^{\infty} \frac{(-t)^{k-1}}{(k-1)!} (dF_{X_n} - dF_U)(t).
 \end{aligned}$$

Another application of Fubini's theorem gives

$$\begin{aligned}
 I_1 &= \int_{-\infty}^0 \int_t^0 \frac{(z-t)^{k-2}}{(k-2)!} dz (dF_{X_n} - dF_U)(t) \\
 (6.5) \qquad \qquad \qquad &= \int_{-\infty}^0 \frac{(-t)^{k-1}}{(k-1)!} (dF_{X_n} - dF_U)(t).
 \end{aligned}$$

Combining (6.3), (6.4) and (6.5) gives

$$\beta^{-1} \mu_k(X_n, U) \geq \left| \int_{-\infty}^{\infty} \frac{(-t)^{k-1}}{(k-1)!} (dF_{X_n} - dF_U)(t) \right|$$

$$= \frac{1}{(k-1)!} \left| E(X_n^{k-1} - U^{k-1}) \right| ,$$

which gives the desired implication (i) $\Rightarrow$ (ii)(b).

To prove (v) $\Rightarrow$ (ii)(b) we integrate by parts to obtain

$$\begin{aligned} \nu_k(X_n, U) &\geq \int_{-\infty}^{\infty} |p_{X_n+Y}(x) - p_{U+Y}(x)| dx \\ &= \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} p_Y^{(k)}(x-z) \left(\int_{-\infty}^z \frac{(z-t)^{k-1}}{(k-1)!} d(F_{X_n}(t) - F_U(t)) dz \right) dx \right| \\ &\geq \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_Y^{(k)}(x-z) dx \left(\int_{-\infty}^z \frac{(z-t)^{k-1}}{(k-1)!} d(F_{X_n}(t) - F_U(t)) dz \right) \right| \\ &= \left| \int_{-\infty}^{\infty} \int_{-\infty}^z \frac{(z-t)^{k-1}}{(k-1)!} d(F_{X_n}(t) - F_Y(t)) dz \right| \cdot \left| \int_{-\infty}^{\infty} p_Y^{(k)}(x) dx \right| . \end{aligned}$$

By (6.2) - (6.5) we obtain

$$\nu_k(X_n, U) \geq \left| \int_{-\infty}^{\infty} p_Y^{(k)}(x) dx \right| \left| E(X_n^k - U^k) \right| ,$$

showing (v) $\Rightarrow$ (ii)(b) and completing Theorem 6.1.

Q E D.

It only remains to give the

Proof of Lemma 6.2. Integration by parts yields

$$\begin{aligned}
 \mu_k(X, U) &= \sup_{h \in \mathbb{R}} |h|^k \sup_{x \in \mathbb{R}} |p_{X+hY}(x) - p_{U+hY}(x)| \\
 &= \sup_{h \in \mathbb{R}} |h|^k \sup_{x \in \mathbb{R}} \left| \int_{-\infty}^{\infty} p_{hY}(z) d(F_{X_n}(x-z) - F_U(x-z)) \right| \\
 (6.6) \quad &= \sup_{h \in \mathbb{R}} |h|^k \sup_{x \in \mathbb{R}} \left| \int_{-\infty}^{\infty} F_X^{(2-k)}(x-z) - F_U^{(2-k)}(x-z) p_{hY}^{(k-1)}(z) dz \right|.
 \end{aligned}$$

Now $2\pi p_{hY}(z) = \int_{-\infty}^{\infty} e^{-itz} e^{-|ht|^\alpha} dt$ and differentiating

$p_{hY}^{(k-1)}$ times gives (setting $\tilde{t} = ht$):

$$\begin{aligned}
 2\pi |h^k p_{hY}^{(k-1)}(z)| &= \left| h^k \int_{-\infty}^{\infty} (it)^{k-1} e^{itz} e^{-|ht|^\alpha} dt \right| \\
 &= \left| h^k \int_{-\infty}^{\infty} \left(i \frac{\tilde{t}}{h}\right)^{k-1} e^{i\tilde{t}z/h} e^{-|\tilde{t}|^\alpha} d\left(\frac{\tilde{t}}{h}\right) \right| \\
 &= \left| \int_{-\infty}^{\infty} (i\tilde{t})^{k-1} e^{i\tilde{t}z/h} e^{-|\tilde{t}|^\alpha} d\tilde{t} \right|.
 \end{aligned}$$

Since $\beta := \beta(\alpha, k) := \frac{1}{2\pi} \int_{-\infty}^{\infty} |it|^{k-1} e^{-|t|^\alpha} dt < \infty$ we obtain

$$\lim_{h \rightarrow \infty} |h^k p_{hY}^{(k-1)}(z)| = \frac{1}{2\pi} \left| \int_{-\infty}^{\infty} \lim_{h \rightarrow \infty} (it)^{k-1} e^{itz/h} |t|^\alpha dt \right| = \beta.$$

Now multiply both sides of (6.6) by β^{-1} . Since

$$\int_{-\infty}^{\infty} |t|^{k-1} e^{-|t|^\alpha} dt \quad \text{and} \quad \int |F_X^{(2-k)}(x-z) - F_U^{(2-k)}(x-z)| dt = \zeta_{k-1}(X, U) \quad \text{are}$$

both finite, $\beta^{-1} \mu_k(X, U)$ is

$$\begin{aligned} & \geq \beta^{-1} \sup_{x \in \mathbb{R}} \left| \int_{-\infty}^{\infty} (F_X^{(2-k)}(x-z) - F_U^{(2-k)}(x-z)) \lim_{h \rightarrow \infty} (h^k p_{hY}^{(k-1)}(z)) dz \right| \\ & = \sup_{x \in \mathbb{R}} \left| \int_{-\infty}^{\infty} F_X^{(2-k)}(x-z) - F_U^{(2-k)}(x-z) dz \right| \\ & = \sup_{x \in \mathbb{R}} \left| \int_{-\infty}^{\infty} F_X^{(2-k)}(z) - F_U^{(2-k)}(z) dz \right|. \end{aligned}$$

Q.E.D.

§7. Concluding remarks

The results above show that the "ideal" structure of the convolution metrics μ_r and ν_r may be used to determine optimal rates of convergence in the general central limit theorem problem. The rates are expressed in terms of the uniform metrics Var , χ and ℓ and hold uniformly in n under the sufficient conditions (3.1), (4.1), and (5.1), respectively. We have not explored the possible weakening of these conditions or even their possible necessity. This would be an interesting line of future research.

The ideal convolution metrics μ_r and ν_r are not limited to the

context of Theorems 3.1, 4.1 and 5.1, but they can also be successfully employed to study other questions of interest. For example, we only mention here that ν_r can be used to prove a Berry-Esseen type of estimate for the Kolmogorov metric ρ (1.8).

More precisely, if $X, X_1, X_2, \dots$ denotes a sequence of i.i.d. random variables in $\mathcal{X}(\mathbb{R})$ and $Y \in \mathcal{X}(\mathbb{R})$ an α -stable random variable, then for all $r > \alpha$ and $n \geq 1$

$$(7.1) \quad \rho\left(\frac{X_1 + \dots + X_n}{n^{1/\alpha}}, Y\right) \leq C \nu_{\alpha, r}(X, Y) n^{1-r/\alpha} + \\ + C \max\{\rho(X, Y), \nu_{\alpha, 1}(X, Y), \nu_{\alpha, r}^{1/(r-\alpha)}(X, Y)\} n^{-1/\alpha},$$

where C is an absolute constant.

Clearly, whenever $\nu_{\alpha, 1}(X, Y) < \infty$ and $\nu_{\alpha, r}(X, Y) < \infty$ we obtain the right order estimate in the Berry-Esseen theorem in terms of the metric $\nu_{\alpha, r}$. Inequalities of this type have been proved by Paulauskas (1982), who uses the ρ metric, and by Senatov (1981), who uses σ instead of ρ and who only considers the normal case $\alpha = 2$ together with the $\zeta_{2,1}$ metric. The estimate (7.1), which proceeds by induction and which will be detailed in a forthcoming article, represents an improvement over earlier estimates since the $\nu_{\alpha, r}$ distance is weaker than the $\zeta_{m,p}$ distance ($r = m - 1 + \frac{1}{p}$); see e.g. Lemma 2.5.

Thus metrics of the convolution type, especially those with the "ideal" structure, are extremely appropriate when investigating sums of independent random variables converging to a stable limit law. We can only conjecture that there are other ideal convolution metrics, other than those explored in this article, which may furnish additional results in related limit theorem problems.

References

Banys, J. (1976) On the convergence rate in the multidimensional local

- limit theorem with limiting stable law, Litovskii Mat. Sbornik, 16, 13-20 (in Russian); English translation: Lithuanian Math. J. 16, (1976) pp.320-325.
- Billingsley, P. (1968) Convergence of Probability Measures, J. Wiley and Sons.
- Dudley, R.M. and Walter Philipp (1983) Invariance principles for sums of Banach space valued random elements and empirical processes, Z. Wahr. V. Geb., 62, pp. 509-552.
- Giné, E. and J. Zinn (1987) Empirical processes indexed by Lipschitz functions, Ann. Prob., 14, 4, pp. 1329-1338.
- Ibramigov, I.A. and Yu. V. Linnik (1971) Independent and Stationary Sequences of Random Variables, Wolters-Noordhoff, Groningen.
- Ignatov, Ts.G. and Rachev, S.T. (1983) Minimality of ideal probabilistic metrics (Stability Problems of Stochastic Models - Seminar Proceedings, (VNIISI, Moscow) pp. 36-48 (in Russian); English translation: J. Soviet Math 32, 6 (1986) pp.595-608.
- Maejima, M. and Rachev, S.T. (1987) An ideal metric and the rate of convergence to a self-similar process, Ann. of Prob., 15, no. 2, pp. 708-727.
- Omey, E. and Rachev, S.T. (1987) On the rate of convergence in extreme value theory, Theor. Prob. Appl. (to appear).
- Paulauskas, V.I. (1973) Some nonuniform estimates in limit theorems of probability theory, Soviet Math. Dokl., vol. 14, 4, pp. 1125-1127.
- Paulauskas, V.I. (1976) On the rate of convergence in the central limit theorem in certain Banach spaces, Theor. Prob. Appl. 21, pp. 754-769.

- Paulauskas, V.I. (1982) Convergence to stable laws and their simulation, *Litovskii Mat. Sbornik*, 22, 3, pp. 146-156.
- Rachev, S.T. (1982) Minimal metrics in the random variables space, *Pub. Inst. Statist. Univ. Paris*, XXVII, pp. 27-47.
- Rachev, S.T. (1984a) The Monge-Kantorovich mass transference problem and its stochastic applications, *Theor. Prob. Appl.*, 29, pp. 647-676.
- Rachev, S.T. (1984b) On a class of minimal functionals on a space of probability measures, *Theor. Prob. Appl.*, 29, pp. 41-49.
- Rachev, S.T. (1984c) Extreme functionals in the space of probability measures, *Lect. Notes in Math*, v. 1155, pp. 320-348.
- Rachev, S.T. and Ignatov, Ts.G. (1984) Ideal quadratic metrics (Stability Problems of Stochastic Models - Seminar Proceedings, VNIISI, Moscow) pp. 119-128 (Russian) = *J. Soviet Math*, 35, 2, (1986), pp. 2386-2394 (English).
- Sazonov, V.V. (1972) On a bound for the rate of convergence in the multi-dimensional central limit theorem, *Proc. 6th Berkeley Symposium on Mathematical Statistics and Probability*, 2, pp. 563-581.
- Sazonov, V.V. and Ul'yanov, V.V. (1979) On the speed of convergence in the central limit theorem, *Adv. Appl. Prob.*, 11, pp. 269-270.
- Senatov, V.V. (1980) Uniform estimates of the rate of convergence in the multi-dimensional central limit theorem, *Theor. Prob. Appl.*, 25, 4, pp. 745-759.
- Yukich, J.E. (1985) The convolution metric d_g , *Math. Proc. Camb. Phil. Soc.*, 98, pp. 533-540.
- Yukich, J.E. (1987) Smoothing metrics for measures on groups, preprint.

- Zolotarev, V.M. (1976) Approximation of distributions of sums of independent random variables with values in infinite dimensional spaces, Theor. Probab. Appl., 21, 4, pp. 721-737.
- Zolotarev, V.M. (1977) Ideal metrics in the problems of approximating distributions of sums of independent random variables, Theor. Prob. Appl., 22, pp. 433-449.
- Zolotarev, V.M. (1979) Ideal metrics in the problems of probability theory, Austral. J. Statist., 21, pp. 193-208.
- Zolotarev, V.M. and Rachev, S.T. (1984) Rate of convergence in limit theorems for the max scheme, Stability Problems of Stochastic Models, Lecture Notes in Mathematics, v. 1155, pp. 415-442.

S. T. Rachev
Department of Applied Mathematics
and Statistics
State University of New York
at Stony Brook
Stony Brook, New York 11794-3651

J. E. Yukich
Department of Mathematics
Lehigh University
Bethlehem, PA 18015

END

DATE

FILMED

MARCH

1988

DTIC