Ann. I. H. Poincaré – PR 38, 6 (2002) 1071–1086

S0246-0203(02)01140-8/FLA

ESTIMATES OF THE RATE OF APPROXIMATION IN A DE-POISSONIZATION LEMMA

ESTIMATION DE LA VITESSE D'APPROXIMATION DANS UN LEMME DE DÉPOISSONISATION

Andrei Yu. ZAITSEV [1]
Laboratory of Statistical Methods, St. Petersburg Branch of the Steklov Mathematical Institute, 27 Fontanka, St. Petersburg 191011, Russia
Received 4 April 2001, revised 15 May 2002

ABSTRACT. – Estimates of the rate of approximation in a de-Poissonization lemma of Beirlant and Mason [1] is obtained for the case where the distribution satisfy a quantitative condition of existence of exponential moments introduced in [6].

MSC: 60F05; 60F15; 60F17

Keywords: Strong approximation; Prokhorov distance; Central limit theorem; De-Poissonization

RÉSUMÉ. – Des estimées de la vitesse d'approximation dans un lemme de dépoissonisation dû à Beirlant et Mason sont obtenues lorsque la distribution vérifie une condition quantitative d'existence de moments exponentiels introduite dans [6].

1. Introduction

Beirlant and Mason [1] introduced a general method for deriving the asymptotic normality of the L_p-norm of empirical functionals. They proved and essentially used the following "de-Poissonization" Lemma A.

LEMMA A (Beirlant and Mason [1]). – *Let* (*for each* $n \in \mathbf{N}$) $\eta_{1,n}$ *and* $\eta_{2,n}$ *be independent Poisson random variables with* $\eta_{1,n}$ *being Poisson* $(n(1-\alpha_n))$ *and* $\eta_{2,n}$

E-mail address: zaitsev@pdmi.ras.ru (A.Yu. Zaitsev).

[1] Research partially supported by an NSF Grant while visiting the University of Delaware, by Russian Foundation of Basic Research (RFBR) Grant 99-01-00112, by INTAS Grant 99-01317 and by RFBR-DFG Grant 99-01-04027.

being Poisson $(n\alpha_n)$ *where* $\alpha_n \in (0,1)$ *and* $\alpha_n \to \alpha \in (0,1)$ *as* $n \to \infty$. *Denote* $\eta_n = \eta_{1,n} + \eta_{2,n}$ *and set* $U_n = n^{-1/2}(\eta_{1,n} - n(1-\alpha_n))$ *and* $V_n = n^{-1/2}(\eta_{2,n} - n\alpha_n)$. *Let* $\{S_n\}_{n=1}^{\infty}$ *be a sequence of random variables such that the random vector* (S_n, U_n) *is independent of* V_n *and for some* $\beta^2 < \infty$, *and* γ *such that* $(1-\alpha)\beta^2 - \gamma^2 > 0$, $(S_n, U_n) \to_d (\beta Z_1, \sqrt{1-\alpha}\, Z_2)$ *as* $n \to \infty$, *where* Z_1 *and* Z_2 *are standard normal random variables with* $\mathrm{cov}(\beta Z_1, \sqrt{1-\alpha}\, Z_2) = \gamma$. *Then, for all* x, $\mathbf{P}\{S_n \leqslant x \mid \eta_n = n\} \to \mathbf{P}\{\sqrt{\beta^2 - \gamma^2}\, Z_1 \leqslant x\}$.

This lemma was used to prove the Central Limit Theorem (CLT) for L_p-norms of some kernel estimates of densities by Beirlant and Mason [1] and for L_1-norm by Giné, Mason and Zaitsev [5]. Using independence properties of random samples of Poissonized size, one can establish the CLT for some vectors (S_n, U_n) by means of known CLT for sums of 1-dependent random vectors. Lemma A provides a possibility to transfer the CLT to the case when we have samples of fixed size n. We are going to prove estimates of the rate of approximation in Lemma A, assuming that the distributions of (S_n, U_n) belong to some classes of distributions with finite exponential moments which are close to Gaussian ones. Our results could be useful to derive the estimates of the rate of convergence in the CLT's of Beirlant and Mason [1] and Giné, Mason and Zaitsev [5].

To simplify the notation, we shall omit the subscript n considering $S, U, V, \eta, \eta_1, \eta_2, \alpha$ instead of $S_n, U_n, \ldots$. Denote

$$\chi = \mathrm{cov}(S, U), \qquad \Pi = U + V = n^{-1/2}(\eta - n), \qquad \Xi = (S, \Pi) \in \mathbf{R}^2. \tag{1.1}$$

Assume that

$$\mathbf{E}S = 0 \quad \text{and} \quad |\chi|\big(\mathrm{Var}(S)\big)^{-1/2} \leqslant c_1, \quad \text{for some } c_1 < 1. \tag{1.2}$$

We shall treat c_1 as an absolute constant so that any constant depending on c_1 only is considered as well as absolute one. Such constants will be denoted by $c_2, c_3, \ldots$ or c. The same symbol c may be used for different constants even in the same formulas when we do not need to fix their values. Condition (1.2) means that the distribution $\mathcal{L}(\Xi)$ of the vector Ξ is non-degenerated. Let Z_0 be a standard normal random variable independent of $\{S, U, V\}$ and $b > 0$. Set

$$S^* = S + bZ_0, \qquad \Psi = (S^*, \Pi) \in \mathbf{R}^2. \tag{1.3}$$

The conditional density of S^* given $\eta = n$ will be denoted $p(x)$, $x \in \mathbf{R}$. We assume that

$$Q \stackrel{\mathrm{def}}{=} \mathcal{L}(\Xi) = \mathcal{L}\big((S, U) + (0, V)\big) \in \mathcal{A}_2(\tau), \tag{1.4}$$

where $\mathcal{A}_d(\tau)$, $\tau \geqslant 0$, $d \in \mathbf{N}$, denote classes of d-dimensional distributions, introduced in Zaitsev [6], see as well Zaitsev [7–9]. The class $\mathcal{A}_d(\tau)$ (with a fixed $\tau \geqslant 0$) consists of d-dimensional distributions F for which the function $\varphi(z) = \varphi(F, z) = \log \int_{\mathbf{R}^d} \mathrm{e}^{\langle z, x\rangle} F\{dx\}$ $(\varphi(0) = 0)$ is defined and analytic for $\|z\|\tau < 1$, $z \in \mathbf{C}^d$, and $|d_u d_v^2 \varphi(z)| \leqslant \|u\|\tau \langle \mathbb{D}v, v\rangle$ for all $u, v \in \mathbf{R}^d$ and $\|z\|\tau < 1$, where $\mathbb{D} = \mathrm{cov}\, F$, the covariance operator corresponding to F, and $d_u\varphi$ is the derivative of the function φ in direction u.

It is easy to see that $\tau_1 < \tau_2$ implies $\mathcal{A}_d(\tau_1) \subset \mathcal{A}_d(\tau_2)$. Moreover, if $F_1, F_2 \in \mathcal{A}_d(\tau)$, then $F_1 F_2 \stackrel{\text{def}}{=} F_1 * F_2 \in \mathcal{A}_d(\tau)$. The class $\mathcal{A}_d(0)$ coincides with the class of all Gaussian distributions in $\mathbf{R}^d$. See Zaitsev [6–9] for further properties of classes $\mathcal{A}_d(\tau)$. Thus, obviously, $\Phi \stackrel{\text{def}}{=} \mathcal{L}((bZ_0, 0)) \in \mathcal{A}_2(0) \subset \mathcal{A}_2(\tau)$, for any $b \in \mathbf{R}$ and $\tau \geqslant 0$. Using the closeness of $\mathcal{A}_2(\tau)$ with respect to convolution, (1.1), (1.3) and (1.4), we conclude that

$$F \stackrel{\text{def}}{=} \mathcal{L}(\Psi) = \mathcal{L}\big((S, U) + (0, V) + (bZ_0, 0)\big) \in \mathcal{A}_2(\tau), \tag{1.5}$$

for any $b \in \mathbf{R}$. The summand $(bZ_0, 0)$ will play a smoothing role ensuring the existence of the conditional density $p(x)$ in (2.13) below. The value of $b \geqslant 0$ will be optimized later.

Throughout the following $\theta, \theta_1, \theta_2, \ldots$ symbolize quantities *depending* on variables involved in corresponding formulas and not exceeding one in absolute value. The same symbol θ may be used for different quantities even in the same formulas.

THEOREM 1.1. – *There exist absolute positive constants $c_2, \ldots, c_7$ such that if*

$$c_2 n^{-1/2} \leqslant \tau \leqslant c_3 b, \quad b \leqslant 1, \quad |x| < c_4/\tau, \quad \operatorname{Var}(S) = 1, \tag{1.6}$$

$$5\alpha^{-1} \exp\big(-5\alpha/432\tau^2\big) \leqslant \tau, \tag{1.7}$$

then

$$p(x) = (2\pi)^{-1/2} B^{-1} \exp\big(-x^2/2B^2\big) \times \exp\big(c_5\big(\theta\tau\big(|x|^3 + 1\big) + \theta \exp\big(-b^2/72\tau^2\big)\big)\big), \tag{1.8}$$

$$2^{-1}(2\pi)^{-1/2} B^{-1} \exp\big(-x^2/B^2\big) \leqslant p(x) \leqslant 2(2\pi)^{-1/2} B^{-1} \exp\big(-x^2/4B^2\big), \tag{1.9}$$

where $B^2 = 1 + b^2 - \chi^2$, $B > 0$, $|\theta| \leqslant 1$ and, moreover, for any $x \in \mathbf{R}$,

$$p(x) \leqslant c_6 B^{-1} \exp\big(-\min\big\{x^2/4B^2, c_7|x|/\tau\big\}\big), \tag{1.10}$$

provided that c_2, c_6 are sufficiently large and c_3, c_4, c_7 sufficiently small absolute constants.

The rather cumbersome condition (1.7) is obviously satisfied for sufficiently large $n \geqslant n_0$ when we consider a scheme of series with $\tau = \tau_n \to 0$ as $n \to \infty$ and fixed $\alpha > 0$.

THEOREM 1.2. – *Let the conditions of Theorem 1.1 be satisfied. Then there exist absolute positive constants c_8 and c_9 such that*

$$\mathbf{P}\big(BZ \leqslant z - \gamma(z)\big) \leqslant \mathbf{P}(S^* \leqslant z \mid \eta = n) \leqslant \mathbf{P}\big(BZ \leqslant z + \gamma(z)\big), \tag{1.11}$$

for $|z| \leqslant c_8\tau^{-1}$, where $\gamma(z) = c_9(\tau(z^2 + 1) + \exp(-b^2/72\tau^2))$ and Z is a standard normal random variable.

Theorem 1.1 will be proved in Section 2. The proof of Theorem 1.2 is sufficiently long and complicated but it is standard. It repeats almost literally the derivation of Lemma 2.1 from Lemma 1.6 in Zaitsev [8]. Theorem 1.2 may be deduced from Theorem 1.1 in a similar way. Therefore we omit the proof of Theorem 1.2.

COROLLARY 1.1. – *Let the conditions of Theorem* 1.2 *be satisfied. Then there exist absolute constants* c_{10}, c_{11}, c_{12} *such that, for any fixed* b *satisfying* (1.6), *one can construct on a probability space random variables* ξ *and* Z *so that the distribution of* ξ *is* F_1, *the conditional distribution of* S^* *given* $\eta = n$, Z *is a standard normal random variable and*

$$|BZ - \xi| \leqslant \gamma(\xi), \quad \textit{for } |\xi| \leqslant c_{10}\tau^{-1}. \tag{1.12}$$

Moreover,

$$|BZ| \geqslant c_{11}\tau^{-1} \quad \textit{if } |\xi| \geqslant c_{10}\tau^{-1} \textit{ and } \tau \leqslant c_{12}b. \tag{1.13}$$

Proof. – Denote $F_1(x) = F_1\{(-\infty, x]\} = \mathbf{P}(S^* \leqslant x|\eta = n)$, $F_2(x) = \mathbf{P}(BZ \leqslant x)$. We assume that a random variable ξ with $\mathbf{P}(\xi \leqslant x) = F_1(x)$ is already constructed and define Z as the unique solution of the equation $F_1(\xi) = F_2(BZ)$. Now inequality (1.12) is an easy consequence of (1.11) if $c_{10} \leqslant c_8$. In order to prove (1.13) it suffices to use (1.12) for $|\xi| = c_{10}\tau^{-1}$ with sufficiently small $c_{10} \leqslant c_8$ and take into account (1.6) and the fact that Z is an increasing function of ξ. □

THEOREM 1.3. – *Let the conditions of Theorem* 1.1 *be satisfied. Then there exist absolute constants* c_{13}, c_{14}, c_{15} *such that, for* τ *satisfying*

$$c_2 n^{-1/2} \leqslant \tau \leqslant c_{13}, \tag{1.14}$$

and for any fixed $\lambda > 0$, *one can construct on a probability space random variables* ζ *and* Z *so that the distribution of* ζ *is the conditional distribution of* S *given* $\eta = n$, Z *is a standard normal random variable and*

$$\mathbf{P}\Big(\Big|\sqrt{1-\chi^2}Z - \zeta\Big| \geqslant \lambda\Big) \leqslant c_{14}\exp(-c_{15}\lambda/\tau). \tag{1.15}$$

The Prokhorov distance is defined by $\pi(F, G) = \inf\{\lambda\colon\ \pi(F, G, \lambda) \leqslant \lambda\}$, where $\pi(F, G, \lambda) = \sup_X \max\{F\{X\} - G\{X^\lambda\}, G\{X\} - F\{X^\lambda\}\}$, $\lambda > 0$, and X^λ is the λ-neighborhood of the Borel set X. Inequality (1.15) implies the following statement.

COROLLARY 1.2. – *Let the conditions of Theorem* 1.3 *be satisfied. Then the bounds* $\pi(F_3, F_4, \lambda) \leqslant c_{14}\exp(-c_{15}\lambda/\tau)$ *and* $\pi(F_3, F_4) \leqslant c\tau(|\log\tau| + 1)$ *hold, where* F_3 *is the conditional distribution of* S *given* $\eta = n$, *and* F_4 *is the centered normal law with variance* $1 - \chi^2$.

Zaitsev [6] has shown that the same bounds are valid for the normal approximation of two-dimensional distributions $Q, F \in \mathcal{A}_2(\tau)$. These bounds do not imply however the inequalities for conditional distributions considered in Corollary 1.2.

One can show that if we consider a scheme of series with $\tau = \tau_n \to 0$ as $n \to \infty$ and $\chi = \chi_n \leqslant c_1 < 1$ for all $n \in \mathbf{N}$, then $(1 - F_3(x))/(1 - F_4(x))$ and $F_3(-x)/F_4(-x)$ tend to 1 as $n \to \infty$, if $0 < x = x_n = \mathrm{o}(\tau_n^{-1/3})$. It suffices to apply the inequality which follows from Corollary 1.2 to the sets $(-\infty, x]$ and $\lambda = \tau_n^{1/3}$. If, in addition, $\chi = \chi_n \to 0$ as $n \to \infty$, then the same limit relation with $F_4(\cdot)$ replaced by the standard normal distribution function is valid for $0 < x = x_n = \mathrm{o}(\min\{\tau_n^{-1/3}, \chi_n^{-1}\})$.

THEOREM 1.4. – *Let the conditions of Theorem* 1.3 *be satisfied. Then there exists absolute constants* c_{16}, c_{17} *such that, for any fixed* b *satisfying* (1.6) *and* $\tau \leqslant c_{12} b$, *one can construct on a probability space random variables* ζ *and* Z *with distributions described in Theorem* 1.3 *so that*

$$\mathbf{P}\Big(\Big|\sqrt{1-\chi^2}\,Z-\zeta\Big| \geqslant c_9 \exp\big(-b^2/72\tau^2\big)+\lambda\Big) \leqslant c_{16}\exp(-c_{17}\lambda/\tau)+2\mathbf{P}\big(|\omega|>\lambda/6\big), \tag{1.16}$$

for any $\lambda > 0$, *where* ω *have the centered normal distribution with variance* b^2.

Theorems 1.3 and 1.4 are proved in Section 3. Comparing Theorems 1.3 and 1.4, we observe that in Theorem 1.3 the probability space depends essentially on λ, while in Theorem 1.4 we proved (1.16) on the same probability space (depending on b) for any $\lambda > 0$. However, (1.16) is weaker than (1.15) for some values of λ. The same rate of approximation (as in (1.15)) is contained in (1.16) if $b^2 \geqslant 72\tau^2\log(1/\tau)$ and $\lambda \geqslant b^2/\tau$ only.

2. Proof of Theorem 1.1

Let $\tau \geqslant 0$, $F = \mathcal{L}(\xi) \in \mathcal{A}_d(\tau)$, $\|h\|\tau < 1$, $h \in \mathbf{R}^d$. Then the Cramér transform $\overline{F} = \overline{F}(h)$ is defined by $\overline{F}(h)\{dx\} = (\mathbf{E}\mathrm{e}^{\langle h,\xi\rangle})^{-1}\mathrm{e}^{\langle h,x\rangle}F\{dx\}$. Also we shall use below the notation $F_h = \overline{F}(h)$ so that $\widehat{F}_h(t) = \int \mathrm{e}^{i\langle t,x\rangle}\overline{F}(h)\{dx\}$. Denote by $\bar{\xi}(h)$ a random vector with $\mathcal{L}(\bar{\xi}(h)) = \overline{F}(h)$. It is clear that $\overline{F}(0) = F$ and the convolution of Cramér transforms is the Cramér transform for convolution with the same h. Below we shall need the following facts.

LEMMA 2.1 (Zaitsev [6, Lemmas 2.1, 3.1]). – *Suppose that* $\tau \geqslant 0$, $F \in \mathcal{A}_d(\tau)$, $h \in \mathbf{R}^d$, $\|h\|\tau \leqslant 1/2$, $F = \mathcal{L}(\xi)$, $\mathbb{D} = \operatorname{cov}\xi$, $\mathbb{D}(h) = \operatorname{cov}\bar{\xi}(h)$ *and* $\mathbf{E}\xi = 0$. *Then* $\overline{F}(h) \in \mathcal{A}_d(2\tau)$,

$$\big\langle \mathbb{D}(h)u,u\big\rangle = \langle \mathbb{D}u,u\rangle\big(1+\theta_1\|h\|\tau\big), \quad \textit{for all } u \in \mathbf{R}^d, \tag{2.1}$$

$$\log \mathbf{E}\mathrm{e}^{\langle h,\xi\rangle} = 2^{-1}\langle \mathbb{D}h,h\rangle\big(1+\theta_2\|h\|\tau/3\big), \tag{2.2}$$

$$\log \mathbf{E}\mathrm{e}^{\mathrm{i}\langle h,\xi\rangle} = -2^{-1}\langle \mathbb{D}h,h\rangle\big(1+\theta_3\|h\|\tau/3\big), \tag{2.3}$$

$$\big(\det \mathbb{D}(h)\big)^{1/2} = (\det \mathbb{D})^{1/2}\exp\big(c\theta_4 d\|h\|\tau\big), \tag{2.4}$$

where c *is an absolute positive constant and* θ_j *satisfy* $|\theta_j| \leqslant 1$.

LEMMA 2.2 (Zaitsev [6, Lemma 3.2]). – *Let* $\Omega = \{\bar{x} \in \mathbf{R}^d\colon\ 4.8\tau\sigma^{-1}\|\mathbb{D}^{-1/2}\bar{x}\| \leqslant 1\}$. *Then, in the conditions of Lemma* 2.1, *for any* $\bar{x} \in \Omega$, *there exists an* $\tilde{h} = \tilde{h}(\bar{x}) \in \mathbf{R}^d$ *such that* $\mathbf{E}\bar{\xi}(\tilde{h}) = \bar{x}$,

$$\sigma\|\tilde{h}\| \leqslant \big\|\mathbb{D}^{1/2}\tilde{h}\big\| \leqslant 2.4\big\|\mathbb{D}^{-1/2}\bar{x}\big\|, \tag{2.5}$$

$$\mathbf{E}\exp\big(\langle\tilde{h},\xi\rangle - \langle\tilde{h},\bar{x}\rangle\big) = \exp\big(-2^{-1}\big\|\mathbb{D}^{-1/2}\bar{x}\big\|^2 + 10.08\theta\tau\sigma^{-1}\big\|\mathbb{D}^{-1/2}\bar{x}\big\|^3\big), \tag{2.6}$$

where σ^2 *is the minimal eigenvalue of the operator* $\mathbb{D}$ *and* $|\theta| \leqslant 1$.

Now we shall operate similarly to the proof of Lemma 4.1 of Zaitsev [6]. Let $r > 0$, $b \geqslant r^{-1}$ and $\widehat{P}_0(u) = \exp(-u^2b^2/2)$, $u \in \mathbf{R}$. For $m = 0, 1, 2, \ldots$, introduce the functions

$\psi_m(u)$ and $\alpha_m(u)$ by putting $\psi_m(u) = \widehat{P}_0(u + mr)$, for $u \geqslant 0$, and $\psi_m(u) = \widehat{P}_0(u - mr)$, for $u \leqslant 0$, and

$$\alpha_m(u) = \psi_m(u) - \psi_{m+1}(u + r) - \psi_{m+1}(u - r) + \psi_{m+2}(u). \tag{2.7}$$

Notice that $\psi_0(u) = \widehat{P}_0(u)$. Clearly, $0 \leqslant \alpha_m(u) \leqslant \psi_m(u)$, for $u \in \mathbf{R}$, and $\alpha_m(u) = 0$, for $|u| \geqslant r$. Further,

$$a_m = \max_{u\in\mathbf{R}} \alpha_m(u) \leqslant \max_{u\in\mathbf{R}} \psi_m(u) = \widehat{P}_0(m) = \exp\bigl(-r^2m^2b^2/2\bigr), \quad m = 0, 1, 2, \ldots. \tag{2.8}$$

It is especially easy to check these properties of functions $\psi_m(u)$ and $\alpha_m(u)$ looking on their graphs.

LEMMA 2.3. – *For any characteristic function $\widehat{W}(u)$ of a one-dimensional distribution W, any $\delta \in \mathbf{R}$ and any $m = 0, 1, 2, \ldots,$*

$$\left|\int_{\mathbf{R}} \widehat{W}(u)\psi_m(u - \delta)\, du\right| \leqslant \int_{\mathbf{R}} \widehat{W}(u)\psi_m(u)\, du. \tag{2.9}$$

Proof. – For $b \geqslant r^{-1}$, it may be shown that $\widehat{P}_0'(u) < 0$ and $\widehat{P}_0''(u) > 0$, for $u > r$. Therefore, by Polya's criterion (see Feller [4]), the function $\widehat{P}_m(u) = \psi_m(u)/a_m$ is the characteristic function of a probability distribution P_m, for each $m = 1, 2, \ldots$. Let $d_m(y)$ be the density corresponding to the distribution P_m, $m = 0, 1, 2, \ldots$. By Parseval's equality, we get (2.9):

$$\left|\int_{\mathbf{R}} \widehat{W}(u)\widehat{P}_m(u - \delta)\, du\right| = \left|2\pi \int_{\mathbf{R}} \mathrm{e}^{\mathrm{i}y\delta} d_m(y)\, W(dy)\right| \leqslant \int_{\mathbf{R}} \widehat{W}(u)\widehat{P}_m(u)\, du. \qquad \square$$

Proof of Theorem 1.1. – Restriction (1.2) turns now into

$$|\chi| \leqslant c_1 < 1 \tag{2.10}$$

since we assumed $\mathrm{Var}(S) = 1$. Note that (1.6) and (2.10) imply that

$$B^2 \asymp \mathrm{e}^{c\theta}. \tag{2.11}$$

Consider the characteristic function $\phi(t_1, t_2) = \mathbf{E}\exp(\mathrm{i}t_1S^* + \mathrm{i}t_2\Pi)$. Clearly,

$$\phi(t_1, t_2) = \sum_{k=0}^{\infty} \exp\left(\frac{\mathrm{i}t_2(k - n)}{\sqrt{n}}\right)\mathbf{E}\left(\exp(\mathrm{i}t_1S^*) \,\middle|\, \Pi = \frac{k - n}{\sqrt{n}}\right)\mathbf{P}\left(\Pi = \frac{k - n}{\sqrt{n}}\right).$$

From this we see by Fourier's inversion that the conditional characteristic function of S^*, given $\eta = n$ or, equivalently, $\Pi = 0$, is

$$\varphi(t_1) = \frac{1}{2\pi\mathbf{P}(\Pi = 0)\sqrt{n}} \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} \exp\bigl(-t_1^2b^2/2\bigr)\mathbf{E}\exp(\mathrm{i}t_1S + \mathrm{i}t_2U)\mathbf{E}\exp(\mathrm{i}t_2V)\, dt_2. \tag{2.12}$$

We shall use the following inversion formula expressing $p(x)$ via characteristic functions:

$$p(x)=\frac{1}{(2\pi)^2\mathbf{P}(\Pi=0)\sqrt{n}}\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}t_1x}\left(\int_{-\pi\sqrt{n}}^{\pi\sqrt{n}}\widehat{F}(t)\,dt_2\right)dt_1, \tag{2.13}$$

where $\widehat{F}(t)=\widehat{\Phi}(t)\widehat{G}(t)\widehat{H}(t)$, $t=(t_1,t_2)\in\mathbf{R}^2$, denote below the characteristic function of F, $\widehat{\Phi}(t)=\exp(-t_1^2b^2/2)$, $\widehat{G}(t)=\mathbf{E}\exp(\mathrm{i}t_1S+\mathrm{i}t_2U)$, $\widehat{H}(t)=\mathbf{E}\exp(\mathrm{i}t_2V)$.

Let $h=(h_1,h_2)\in\mathbf{R}^2$ satisfy $\|h\|\tau<1/2$. For the conditional density of $\overline{S^*}(h)$, given $\langle\overline{\Psi}(h),e_2\rangle=0$, we shall use the notation $p_h(x)$, $x\in\mathbf{R}$. Arguing similarly as deriving (2.13), we may express $p_h(x)$ via characteristic functions:

$$p_h(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{i}t_1x}\varphi_h(t_1)\,dt_1,\quad x\in\mathbf{R}, \tag{2.14}$$

where

$$\varphi_h(t_1)=\frac{1}{2\pi\mathbf{P}(\langle\overline{\Psi}(h),e_2\rangle=0)\sqrt{n}}\int_{-\pi\sqrt{n}}^{\pi\sqrt{n}}\widehat{F}_h(t)\,dt_2 \tag{2.15}$$

is the characteristic function corresponding to the density $p_h(x)$ and

$$\widehat{F}_h(t)=\widehat{\Phi}_h(t)\widehat{G}_h(t)\widehat{H}_h(t),\quad t\in\mathbf{R}^2,$$

is the characteristic function of $\overline{F}(h)$. For $h=0$, (2.14) turns into (2.13) since $\langle\Psi,e_2\rangle=\Pi$ (see (1.3)). It is easy to see that $\widehat{\Phi}_h(t)=\exp(\mathrm{i}t_1bh_1-t_1^2b^2/2)$ and $\varphi_h(t_1)=\beta_h(t_1)\exp(\mathrm{i}t_1bh_1-t_1^2b^2/2)$, where

$$\beta_h(t_1)=\left(2\pi\mathbf{P}\left(\left\langle\overline{\Psi}(h),e_2\right\rangle=0\right)\sqrt{n}\,\right)^{-1}\int_{-\pi\sqrt{n}}^{\pi\sqrt{n}}\widehat{Q}_h(t)\,dt_2.$$

Note that in a similar way one can establish that $\beta_h(t_1)$ is the characteristic function of the conditional distribution of $\overline{S}(h)$, given $\langle\overline{\Psi}(h),e_2\rangle=0$. Thus, by (2.14),

$$p_h(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\beta_h(t_1)\exp\left(\mathrm{i}t_1bh_1-\mathrm{i}t_1x-t_1^2b^2/2\right)dt_1,\quad x\in\mathbf{R}. \tag{2.16}$$

By (2.13), (2.14) and (2.15), we have

$$p_h(x)\mathbf{P}\left(\left\langle\overline{\Psi}(h),e_2\right\rangle=0\right)=p(x)\mathbf{P}(\Pi=0)\left(\mathbf{E}\mathrm{e}^{\langle h,\Psi\rangle}\right)^{-1}\mathrm{e}^{h_1x},\quad x\in\mathbf{R}. \tag{2.17}$$

Collecting (2.13), (2.14) and (2.17), we get

$$p(x) = \frac{\mathbf{E}\mathrm{e}^{\langle h,\Psi\rangle - h_1 x}}{(2\pi)^2 \mathbf{P}(\Pi = 0)\sqrt{n}} \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}t_1 x} \left(\int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} \widehat{F}_h(t)\, dt_2 \right) dt_1, \quad x \in \mathbf{R}. \tag{2.18}$$

By the definition of $\psi_0(u)$ and (2.14)–(2.16),

$$\begin{aligned} 2\pi p_h(x) &= \int_{-\infty}^{\infty} \widehat{W}(u)\psi_0(u)\, du \\ &= \frac{1}{2\pi \mathbf{P}(\langle \overline{\Psi}(h), e_2\rangle = 0)\sqrt{n}} \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}t_1 x} \left(\int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} \widehat{F}_h(t)\, dt_2 \right), \end{aligned} \tag{2.19}$$

where $\widehat{W}(u) = \beta_h(u)\exp(\mathrm{i}ubh_1 - \mathrm{i}ux)$ is the characteristic function of a one-dimensional probability distribution. Now, using (2.7) with $m = 0$, $r = (6\tau)^{-1}$, $b \geqslant 6\tau$, we expand the first integral in (2.19) into a sum of integrals and, applying (2.9), estimate

$$\left| \int_{-\infty}^{\infty} \widehat{W}(u)\psi_0(u)\, du - I_0 \right| \leqslant 2|J_1| + |J_2| \stackrel{\text{def}}{=} \sum\nolimits_1 |J_m|, \tag{2.20}$$

where we denote $I_m = \int \widehat{W}(u)\alpha_m(u)\, du$, $J_m = \int \widehat{W}(u)\psi_m(u)\, du$, for $m = 0, 1, \ldots$. The sum $\sum_1 |J_m|$ will be estimated with the help of a sequential procedure based on identity (2.7) and inequality (2.9). In each of the integrals J_m we again replace $\psi_m(u)$ using (2.7) and apply (2.9). As a result we obtain the inequality $\sum_1 |J_m| \leqslant \sum_1 |I_m| + \sum_2 |J_m|$. Each of the terms in the sum $\sum_1 |J_m|$ generates one (corresponding) term in the sum $\sum_1 |I_m|$ and three terms in the sum $\sum_2 |J_m|$. The index m of each term generated in $\sum_2$ is by at least one greater then the corresponding index of the generating term in $\sum_1$. Continuing to operate in the same fashion, at the sth step we obtain the inequality

$$\sum\nolimits_1 |J_m| \leqslant \sum_{k=1}^{s-1} \sum\nolimits_k |I_m| + \sum\nolimits_s |J_m|, \tag{2.21}$$

in which indices m occurring in $\sum_k$, $k = 1, \ldots, s$, are at least k and the number of terms is 3^k. It is easy to show that $\sum_s |J_m| \leqslant 3^s c_1(b,\tau)\exp(-c_2(b,\tau)s^2) \to 0$ as $s \to \infty$, where $c_j(b,\tau)$, $j = 1, 2$, are positive quantities depending on b and τ only. By (2.20) and (2.21),

$$\left| \int_{-\infty}^{\infty} \widehat{W}(u)\psi_0(u)\, du - I_0 \right| \leqslant \sum_{k=1}^{\infty} \sum\nolimits_k |I_m|. \tag{2.22}$$

By (2.8), the right-hand side of (2.22) may be estimated by

$$\sum_{k=1}^{\infty} 3^k \exp\bigl(-r^2k^2b^2/2\bigr) K_0 \leqslant c \exp\bigl(-r^2b^2/2\bigr) K_0, \tag{2.23}$$

where

$$K_0 = \int_{-r}^{r} \bigl|\widehat{W}(u)\bigr|\,du = \int_{-r}^{r} \bigl|\beta_h(u)\bigr|\,du$$
$$= \frac{1}{2\pi\,\mathbf{P}(\langle\overline{\Psi}(h), e_2\rangle = 0)\sqrt{n}}(K_1 + K_2), \tag{2.24}$$

and $K_j = \int_{T_j} |\widehat{Q}_h(t)|\,dt,\ \ j = 1, 2$, where

$$T_1 = \bigl\{t = (t_1, t_2) \in \mathbf{R}^2\colon\ |t_1| \leqslant r,\ |t_2| \leqslant r\bigr\}, \tag{2.25}$$
$$T_2 = \bigl\{t = (t_1, t_2) \in \mathbf{R}^2\colon\ |t_1| \leqslant r,\ r \leqslant |t_2| \leqslant \pi\sqrt{n}\bigr\}. \tag{2.26}$$

By (2.7), (2.8) and (2.24), $|\int_{-r}^{r} \widehat{W}(u)\psi_0(u)\,du - I_0| \leqslant 3\exp(-r^2b^2/2)K_0$. Together with (2.22) and (2.23), this inequality implies that

$$\left|\int_{-r}^{r} \widehat{W}(u)\psi_0(u)\,du - \int_{-\infty}^{\infty} \widehat{W}(u)\psi_0(u)\,du\right| \leqslant c\exp\bigl(-r^2b^2/2\bigr)K_0. \tag{2.27}$$

Taking into account (1.5), we may apply Lemma 2.1 which implies that $\mathcal{L}(\overline{\Psi}(h)) \in \mathcal{A}_2(2\tau)$. The characteristic function of $\overline{\Psi}(h) - \mathbf{E}\overline{\Psi}(h)$ is $\widehat{F}_h(t)\exp(-\mathrm{i}\langle t, \mathbf{E}\overline{\Psi}(h)\rangle)$. Using relation (2.3) of Lemma 2.1 with doubled parameter τ, we obtain

$$\log\bigl(\widehat{F}_h(t)\exp\bigl(-\mathrm{i}\bigl\langle t, \mathbf{E}\overline{\Psi}(h)\bigr\rangle\bigr)\bigr)$$
$$= -2^{-1}\bigl\langle \mathbb{D}(h)t, t\bigr\rangle\bigl(1 + 2\theta\|t\|\tau/3\bigr), \quad \text{for } \|t\|\tau \leqslant 1/4, \tag{2.28}$$

where $\mathbb{D}(h) = \operatorname{cov}\overline{\Psi}(h)$ (we denote as well $\mathbb{D} = \operatorname{cov}\Psi$). According to (1.1) and (1.5), we have $\det\mathbb{D} = 1 + b^2 - \chi^2 = B^2$. Moreover, for $u = (u_1, u_2) \in \mathbf{R}^2$,

$$\langle \mathbb{D}u, u\rangle = \bigl\|\mathbb{D}^{1/2}u\bigr\|^2 = \mathbf{E}\langle\Psi, u\rangle^2 = \bigl(1 + b^2\bigr)u_1^2 + u_2^2 + 2u_1u_2\chi. \tag{2.29}$$

Furthermore, one may calculate that

$$\bigl\langle \mathbb{D}^{-1}u, u\bigr\rangle = \bigl\|\mathbb{D}^{-1/2}u\bigr\|^2 = \bigl((1 + b^2)u_2^2 + u_1^2 - 2u_1u_2\chi\bigr)(\det\mathbb{D})^{-1}. \tag{2.30}$$

Applying relation (2.1) of Lemma 2.1, (2.10) and (2.29), we see that

$$\bigl\langle \mathbb{D}(h)u, u\bigr\rangle \geqslant c\|u\|^2. \tag{2.31}$$

Using the inequality $|\mathrm{e}^{z_1} - \mathrm{e}^{z_2}| \leqslant |z_1 - z_2|\max\{|\mathrm{e}^{z_1}|, |\mathrm{e}^{z_1}|\}$, $z_1, z_2 \in \mathbf{C}$, and relation (2.28) and (2.31), we find, for $t \in T_1$ and for sufficiently small c_4, that

$$\bigl|\widehat{F}_h(t) - \exp\bigl(-2^{-1}\bigl\langle \mathbb{D}(h)t, t\bigr\rangle + \mathrm{i}\bigl\langle t, \mathbf{E}\overline{\Psi}(h)\bigr\rangle\bigr)\bigr| \leqslant c\tau\exp\bigl(-c\|t\|^2\bigr). \tag{2.32}$$

It is easy to see that

$$\int_{-r}^{r} \widehat{W}(u)\psi_0(u)\,du = \frac{1}{2\pi\,\mathbf{P}(\langle\overline{\Psi}(h), e_2\rangle = 0)\sqrt{n}} \int_{T_1\cup T_2} \mathrm{e}^{-\mathrm{i}t_1x}\,\widehat{F}_h(t)\,dt. \tag{2.33}$$

Now we expand the integral in (2.33) into a sum of integrals and estimate

$$\Bigl|\int_{T_1\cup T_2} \mathrm{e}^{-\mathrm{i}t_1 x}\widehat{F}_h(t)\,dt - L_0\Bigr| \leqslant L_1 + L_2 + L_3, \tag{2.34}$$

where

$$L_0 = \int_{\mathbf{R}^2} \exp\Bigl(-\frac{1}{2}\langle \mathbb{D}(h)t, t\rangle\Bigr)\exp\bigl(-\mathrm{i}t_1 x + \mathrm{i}\langle t, \mathbf{E}\overline{\Psi}(h)\rangle\bigr)\,dt, \tag{2.35}$$

$$L_1 = \Bigl|\int_{T_1} \mathrm{e}^{-\mathrm{i}t_1 x}\Bigl(\widehat{F}_h(t) - \exp\Bigl(-\frac{1}{2}\langle \mathbb{D}(h)t, t\rangle\Bigr)\exp\bigl(\mathrm{i}\langle t, \mathbf{E}\overline{\Psi}(h)\rangle\bigr)\Bigr)\,dt\Bigr|, \tag{2.36}$$

$$L_2 = \Bigl|\int_{T_2} \mathrm{e}^{-\mathrm{i}t_1 x}\widehat{F}_h(t)\,dt\Bigr|, \tag{2.37}$$

$$L_3 = \Bigl|\int_{\mathbf{R}^2\setminus T_1} \exp\Bigl(-\frac{1}{2}\langle \mathbb{D}(h)t, t\rangle\Bigr)\exp\bigl(-\mathrm{i}t_1 x + \mathrm{i}\langle t, \mathbf{E}\overline{\Psi}(h)\rangle\bigr)\,dt\Bigr|. \tag{2.38}$$

By (2.4),

$$|L_0| \leqslant 2\pi\bigl(\det \mathbb{D}(h)\bigr)^{-1/2} = 2\pi(\det\mathbb{D})^{-1/2}\exp\bigl(c\theta\|h\|\tau\bigr). \tag{2.39}$$

Coupled with (2.36), inequality (2.32) implies that $L_1 \leqslant c\tau$. Estimating L_2, we first note that $|\widehat{F}_h(t)| \leqslant |\widehat{Q}_h(t)| \leqslant |\widehat{H}_h(t)| = |\mathbf{E}\exp((h_2+\mathrm{i}t_2)V)/\mathbf{E}\exp(h_2 V)|$. By Example 1.2 in Zaitsev [6], we have $\mathcal{L}(V) \in \mathcal{A}_1(c/\sqrt{n})$. Clearly, $\mathbf{E}V = 0$, $\mathrm{Var}(V) = \alpha$. The function $s(t_2) = \mathbf{E}\exp((h_2+\mathrm{i}t_2)V)/\mathbf{E}\exp(h_2 V)$ may be considered as the characteristic function of the one-dimensional distribution $\overline{\mathcal{L}(V)}(h_2)$. Applying a one-dimensional version of (2.1) and (2.3) of Lemma 2.1, we get

$$\log s(t_2) = -2^{-1}\alpha t_2^2\bigl(1 + c\theta|t_2|/\sqrt{n}\,\bigr)\bigl(1 + c\theta|h_2|/\sqrt{n}\,\bigr), \quad \text{for } |t_2| \leqslant c_{10}\sqrt{n}, \tag{2.40}$$

with a sufficiently small absolute constant c_{10}. Thus, for sufficiently large c_2, (2.40) gives

$$\bigl|s(t_2)\bigr| \leqslant \exp\bigl(-5\alpha t_2^2/12\bigr), \quad \text{for } |t_2| \leqslant c_{10}\sqrt{n} \quad \text{and} \quad (6\tau)^{-1} \leqslant c_{10}\sqrt{n} \leqslant \pi\sqrt{n}, \tag{2.41}$$

if c_{10} is small enough. The function $|s(t_2)|$ may be easily calculated:

$$\bigl|s(t_2)\bigr| = g(t_2)/g(0), \quad \text{where } g(t_2) = \exp\bigl(\alpha\bigl(\mathrm{e}^{h_2}\cos\bigl(t_2/\sqrt{n}\,\bigr) - 1\bigr)\bigr). \tag{2.42}$$

The function g is even and decreasing for $0 \leqslant t_2 \leqslant \pi\sqrt{n}$. Therefore, using (1.7), (2.26), (2.37), (2.41) and (2.42), we obtain (if c_2 is sufficiently large and c_3 sufficiently small)

$$\begin{aligned}\max\{L_2, K_2\} &\leqslant \frac{1}{3\tau}\Bigl(\int_{|u|\geqslant\tau^{-1}/6} \exp\Bigl(-\frac{5}{12}\alpha u^2\Bigr)\,du + 2\pi\sqrt{n}\exp\Bigl(-\frac{5}{12}\alpha c_{10}^2 n\Bigr)\Bigr)\\ &\leqslant (3\tau)^{-1}\bigl(14.4\tau\alpha^{-1}\exp\bigl(-5\alpha/432\tau^2\bigr) + 2\pi\sqrt{n}\exp\bigl(-5\alpha c_{10}^2 n/12\bigr)\bigr)\\ &\leqslant 5\alpha^{-1}\exp\bigl(-5\alpha/432\tau^2\bigr) \leqslant \tau. \end{aligned}\tag{2.43}$$

Let us estimate L_3. Using (2.25), (2.38) and (2.31), we obtain $L_3 \leqslant 4\int_{\tau^{-1}/6}^{\infty} \mathrm{e}^{-cy^2}\,dy \leqslant c\tau$. Collecting (2.34) and bounds for L_j, we get (if c_2 is sufficiently large and c_3, c_4 sufficiently small)

$$\left| \int\limits_{T_1\cup T_2} \mathrm{e}^{-\mathrm{i}t_1 x}\widehat{F}_h(t)\,dt - L_0 \right| \leqslant c\tau. \tag{2.44}$$

Arguing similarly to the proof of (2.28) and (2.31), we may show that

$$\log\bigl(\widehat{Q}_h(t)\exp\bigl(-\mathrm{i}\langle t, \mathbf{E}\overline{\Xi}(h)\rangle\bigr)\bigr) = -2^{-1}\langle \mathbb{B}(h)t, t\rangle\bigl(1+2\theta\|t\|\tau/3\bigr), \quad \text{for } \|t\|\tau \leqslant 1/4,$$

where $\mathbb{B}(h) = \operatorname{cov}\overline{\Xi}(h)$ and $\langle \mathbb{B}(h)t, t\rangle \geqslant c\|t\|^2$. Hence, $K_1 \leqslant c$.

Let us fix an $x \in \mathbf{R}$ satisfying $|x| < c_4/\tau$. Let us apply Lemma 2.2 with $\xi = \Psi$, $\bar{x} = (x, 0) \in \mathbf{R}^2$. Relation (2.10) coupled with independence of (S, Π) and $(bZ_0, 0)$ implies that in this case

$$\sigma^2 = \min_{\|t\|=1} \operatorname{Var}\bigl(\langle \Psi, t\rangle\bigr) = \min_{\|t\|=1}\bigl(t_1^2(1+b^2) + t_2^2 + 2t_1t_2\chi\bigr) \geqslant c. \tag{2.45}$$

Moreover, $4.8\tau\sigma^{-1}\|\mathbb{D}^{-1/2}\bar{x}\| \leqslant 5\tau\sigma^{-2}|x| \leqslant c\tau|x| < 1$, where $\mathbb{D} = \operatorname{cov}\Psi$ if c_4 is sufficiently small. Applying now Lemma 2.2, we get an $\tilde{h} = \tilde{h}(\bar{x}) = (\tilde{h}_1, \tilde{h}_2) \in \mathbf{R}^2$, satisfying $\mathbf{E}\overline{\Psi}(\tilde{h}) = \bar{x}$, (2.5) and (2.6) with $\xi = \Psi$, $\bar{x} = (x, 0)$. We shall first estimate $p(x)$ using (2.18) with $h = \tilde{h}$. Note that (2.5) yields $\|\tilde{h}\|\tau < 1/2$. So we can apply for $h = \tilde{h}$ all the relations derived above. In particular, using (2.11) and (2.30), we get $\|\mathbb{D}^{-1/2}\bar{x}\|^2 = x^2/B^2 \leqslant cx^2$. Substituting this bound into (2.6) and using (2.45), we obtain

$$\mathbf{E}\exp\bigl(\langle \tilde{h}, \Psi\rangle - \tilde{h}_1 x\bigr) = \exp\bigl(-x^2/2B^2 + c\theta\tau|x|^3\bigr). \tag{2.46}$$

Relations $\mathbf{E}\overline{\Psi}(\tilde{h}) = \bar{x}$, $\|\tilde{h}\|\tau < 1/2$ and $\det\mathbb{D} = B^2$ together with (2.11), (2.35) and (2.39) imply $L_0 = 2\pi B^{-1}\exp(c\theta\|\tilde{h}\|\tau) = \mathrm{e}^{c\theta}$. By (2.5), $\|\tilde{h}\| \leqslant c|x|$, if c_4 is sufficiently small. Hence, $L_0 = 2\pi B^{-1}\exp(c\theta|x|\tau)$. Inequalities (1.6) and (2.44) give now the relation

$$\int\limits_{T_1\cup T_2} \mathrm{e}^{-\mathrm{i}t_1 x}\widehat{F}_{\tilde{h}}(t)\,dt = L_0 + c\theta\tau = 2\pi B^{-1}\exp\bigl(c\theta\tau\bigl(|x|+1\bigr)\bigr), \tag{2.47}$$

if c_2 is sufficiently large and c_4 is sufficiently small. Recall that $r = (6\tau)^{-1}$. Using now (2.11), (2.19), (2.24), (2.27), (2.33), (2.43), (2.47) and $K_1 \leqslant c$, we get

$$\int\limits_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}t_1 x}\left(\int\limits_{-\pi\sqrt{n}}^{\pi\sqrt{n}} \widehat{F}_{\tilde{h}}(t)\,dt_2\right)dt_1 = 2\pi B^{-1}\exp\bigl(c\theta\tau\bigl(|x|+1\bigr) + c\theta\exp\bigl(-b^2/72\tau^2\bigr)\bigr). \tag{2.48}$$

Applying Stirling's formula, $n! = (n/\mathrm{e})^n\sqrt{2\pi n}\,\mathrm{e}^{\theta_n/n}$ for some $0 < \theta_n < 1/12$, we obtain

$$2\pi\mathbf{P}(\eta = n) = 2\pi\mathrm{e}^{-n}n^n/n! = (2\pi/n)^{1/2}\mathrm{e}^{\theta/n}. \tag{2.49}$$

Collecting bounds (2.18) with $h = \tilde{h}$, (2.46), (2.47) and (2.49) and using (1.6), we complete the proof of (1.8). Inequalities (1.9) follows from (1.8) and (2.11) if c_4 is small enough.

To prove inequality (1.10), we define $h^* = (h_1^*, h_2^*) \in \mathbf{R}^2$ by $h_2^* = 0$ and $h_1^* = c_4\tau^{-1}/2$. Below we choose c_4 so small as it is necessary. Taking $c_4 < 1$, we ensure the validity of $\|h^*\|\tau < 1/2$. So we can apply for $h = h^*$ all the relations derived above. By (2.11) and (2.39), $|L_0| \leqslant c$. Coupled with inequality (2.44), this implies that $|\int_{T_1\cup T_2} \mathrm{e}^{-\mathrm{i}t_1x}\widehat{F}_{h^*}(t)\,dt| \leqslant |L_0| + c\tau \leqslant c$. Using (2.19), (2.24), (2.27), (2.33), (2.43) and (2.47) together with $K_1 \leqslant c$, we get

$$\left|\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}t_1x}\left(\int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} \widehat{F}_{h^*}(t)\,dt_2\right)dt_1\right| \leqslant c. \tag{2.50}$$

Using relation (2.2) of Lemma 2.1, we obtain, for $x \geqslant c_4\tau^{-1}$, $\log \mathbf{E}\mathrm{e}^{\langle\Psi,h^*\rangle - h_1^*x} \leqslant c_4^2/4\tau^2 - c_4x/2\tau \leqslant -c_4x/4\tau$. Now (1.10) for $x \geqslant c_4\tau^{-1}$ follows from (2.18), (2.49) and (2.50). For $x \leqslant -c_4\tau^{-1}$, it may be verified in a similar way. For $|x| \leqslant c_4\tau^{-1}$, it follows from (1.9). □

3. Proof of Theorems 1.3 and 1.4

LEMMA 3.1. – *Let the conditions of Theorem* 1.3 *be satisfied. Let positive* ε, δ *and* y *satisfy*

$$c_2n^{-1/2} \leqslant \tau \leqslant c_{18}b, \quad b \leqslant c_{19}, \quad \varepsilon = 2\mathrm{e}^2\delta, \quad \delta \leqslant c_{20}/\tau, \quad y = c_{21}\sqrt{\delta/\tau}, \quad \tau \leqslant c_{22}\delta. \tag{3.1}$$

Absolute positive constants $c_{18}, c_{19}, c_{20}, c_{21}, c_{22}$ *and* c_{23} *may be chosen so small that, for any closed set* $X \subset [-y, y]$,

$$F_1\{X\} \leqslant F_2\{X^\varepsilon\} + 2\exp(-c_{23}\delta/\tau) + 3\Delta, \quad \textit{where } \Delta = c_5\exp(-b^2/72\tau^2) \tag{3.2}$$

and distributions F_1 *and* F_2 *are defined by* $F_1(x) = \mathbf{P}(S^* \leqslant x \mid \eta = n)$, $F_2(x) = \mathbf{P}(BZ \leqslant x)$.

Proof. – Choosing $c_{18} \leqslant c_3$ and $c_{19} \leqslant 1$, we are under the hypotheses of Theorem 1.1. Denote by $w(x) = \frac{1}{\sqrt{2\pi}B}\exp(-x^2/2B^2)$ the density of F_2. Then, by virtue of (1.8), (2.11) and (3.1),

$$p(x+\delta) \leqslant w(x)\mathrm{e}^\Delta, \qquad w(x+\delta) \leqslant p(x)\mathrm{e}^\Delta, \tag{3.3}$$

for $x \geqslant 0$, and $2B \leqslant x + \delta \leqslant 2c_{21}\sqrt{\delta/\tau} = 2y$, while

$$p(x-\delta) \leqslant w(x)\mathrm{e}^\Delta, \qquad w(x-\delta) \leqslant p(x)\mathrm{e}^\Delta, \tag{3.4}$$

for $x \leqslant 0$, and $-2y \leqslant x - \delta \leqslant -2B$, if c_{19}, c_{20} and c_{21} are small enough. In similar fashion, choosing c_{18} and c_{19} to be sufficiently small, we can show with the help of

(1.8), (2.11), (3.1) and (3.2) that, for $|x| \leqslant 2B$,

$$p(x) \leqslant w(x)\exp\bigl(c_5(c\tau+\Delta)\bigr) \leqslant 6w(x)/5 \quad \text{and} \quad p(x) \geqslant 18w(x)/19. \tag{3.5}$$

By choosing $c_{20} \leqslant \sqrt{c_{21}}/4\mathrm{e}^4$ to be sufficiently small, we can assure the inequality

$$\varepsilon = 2\mathrm{e}^2\delta \leqslant c_{21}\sqrt{\delta/\tau} = y. \tag{3.6}$$

Let X be an arbitrary closed subset of $[-y, y]$. Consider the collection $\{\Pi_\gamma\}_{\gamma\in\Gamma}$ of open intervals $\Pi_\gamma \subset \mathbf{R}\backslash X$ of lengths at least 2ε. Write $Y = \mathbf{R}\backslash \bigcup_{\gamma\in\Gamma} \Pi_\gamma$. Then

$$X \subset Y \subset [-y, y], \qquad X^\varepsilon = Y^\varepsilon. \tag{3.7}$$

The set Y may be represented as a union of disjoint closed intervals $M_j \subset [-y, y]$, $j = 1, \ldots, l$, separated by intervals whose lengths are at least 2ε. Therefore,

$$F_1\{Y\} = \sum_{j=1}^{l} F_1\{M_j\}, \qquad F_2\bigl\{Y^\varepsilon\bigr\} = \sum_{j=1}^{l} F_2\bigl\{M_j^\varepsilon\bigr\}. \tag{3.8}$$

Observe that, by (3.6),

$$M_j \subset M_j^\varepsilon \subset [-2y, 2y], \quad j = 1, \ldots, l. \tag{3.9}$$

Let us fix j and compare $F_1\{M_j\}$ with $F_2\{M_j^\varepsilon\}$. Let $M_j = [\alpha_j, \beta_j]$. Then $M_j^\varepsilon = (\alpha_j - \varepsilon, \beta_j + \varepsilon)$. Consider separately the four possible cases:

(a) $(\alpha_j, \beta_j) \cap [-2B, 2B] = \emptyset$ and $0 \notin (\alpha_j - \varepsilon, \beta_j + \varepsilon)$;

(b) $0 \in (\alpha_j, \beta_j)$ and $[-2B, 2B] \subset (\alpha_j - \varepsilon, \beta_j + \varepsilon)$;

(c) at least one of the intervals $(\alpha_j - \varepsilon, \alpha_j)$ or $(\beta_j, \beta_j + \varepsilon)$ lies in the interval $[-2B, 2B]$;

(d) one of the intervals $(\alpha_j - \varepsilon, \alpha_j)$ or $(\beta_j, \beta_j + \varepsilon)$ contains at least one of the intervals $[0, 2B]$ or $[-2B, 0]$.

In case (a) we assume for definiteness that $0 \leqslant \alpha_j - \varepsilon$ and $\alpha_j \geqslant 2B$. Then we have, in view of (3.3) and (3.9), $F_1\{M_j\} \leqslant \mathrm{e}^\Delta F_2\{M_j^\varepsilon\}$. If $\beta_j + \varepsilon \leqslant 0$ and $\beta_j \leqslant -2B$, then the same inequality follows from (3.4) and (3.9).

Consider case (b). By Bernstein's inequality and (2.11),

$$F_2\bigl\{(y, \infty)\bigr\} = F_2\bigl\{(-\infty, -y)\bigr\} \leqslant \exp\bigl(-y^2/4B^2\bigr) \leqslant \exp(-c\delta/\tau). \tag{3.10}$$

So that if $\beta_j + \varepsilon > y$, then $F_2\{(\beta_j + \varepsilon, \infty)\} \leqslant \exp(-c\delta/\tau)$. But if $\beta_j + \varepsilon \leqslant y$, then since $\beta_j > 0$, $\beta_j + \varepsilon \geqslant 2B$ and $\varepsilon > \delta$, we have, in view of (3.3) and (3.10),

$$\begin{aligned} F_2\bigl\{(\beta_j + \varepsilon, \infty)\bigr\} &\leqslant \exp(-c\delta/\tau) + \int_{\beta_j+\varepsilon-\delta}^{y-\delta} w(x+\delta)\,dx \\ &\leqslant \exp(-c\delta/\tau) + \mathrm{e}^\Delta F_1\bigl\{(\beta_j, \infty)\bigr\}. \end{aligned}$$

Thus, irrespective of the mutual disposition of the numbers $\beta_j + \varepsilon$ and y, we have the inequality

$$F_2\{(\beta_j + \varepsilon, \infty)\} \leqslant \exp(-c\delta/\tau) + \mathrm{e}^{\Delta} F_1\{(\beta_j, \infty)\}. \tag{3.11}$$

By means of (3.4), it can be shown in similar fashion that

$$F_2\{(-\infty, \alpha_j - \varepsilon)\} \leqslant \exp(-c\delta/\tau) + \mathrm{e}^{\Delta} F_1\{(-\infty, \alpha_j)\}. \tag{3.12}$$

Adding the left-hand and right-hand sides of (3.11) and (3.12), we obtain $1 - F_2\{M_j^\varepsilon\} \leqslant 2\exp(-c\delta/\tau) + \mathrm{e}^{\Delta}(1 - F_1\{M_j\})$ which yields the inequality

$$F_1\{M_j\} \leqslant 2\exp(-c\delta/\tau) + \Delta + F_2\{M_j^\varepsilon\}. \tag{3.13}$$

To consider case (c), we introduce the sets $N_j = M_j^\varepsilon \cap [-2B, 2B]$, $P_j = M_j^\varepsilon \backslash (N_j \cup R_j)$ and $R_j = (M_j^\varepsilon \backslash N_j) \cap ([-2B - \delta, -2B] \cup [2B, 2B + \delta])$. From (3.1) and (3.5), it follows that $F_1\{N_j\} \leqslant \mathrm{e}^{\Delta} F_2\{N_j\} + \delta/\sqrt{2\pi}B$, if c_{18}, c_{19} and c_{22} are sufficiently small. Further, by condition (c), the definition of R_j, (1.9) and (3.9) $F_1\{R_j\} \leqslant \delta/\sqrt{2\pi}B$. If the set P_j is non-empty, then it is concentrated entirely either on the positive or on the negative real axis. For definiteness, let $P_j \subset \{x{:}\ x \geqslant 2B + \delta\}$. Then $P_j - \delta \subset P_j \cup R_j$ and so, by (3.3) and (3.9), $F_1\{P_j\} \leqslant \mathrm{e}^{\Delta} F_2\{P_j \cup R_j\}$. Similarly, we can establish this bound also in the case where $P_j \subset \{x{:}\ x \leqslant -2B - \delta\}$. It is also clear that in case (c) $F_2\{M_j^\varepsilon \backslash M_j\} \geqslant \varepsilon \mathrm{e}^{-2}/\sqrt{2\pi}B = 2\delta/\sqrt{2\pi}B$. Now from (3.9) and the above inequalities it follows that $F_1\{M_j\} = F_1\{N_j\} + F_1\{R_j\} + F_1\{P_j\} \leqslant \mathrm{e}^{\Delta} F_2\{M_j^\varepsilon\}$.

In case (d), $F_2\{M_j^\varepsilon \backslash M_j\} > 0.475$, $F_2\{M_j \cap [-2B, 2B]\} < 1/2$ and $F_1\{[-2B, 2B]\} > 0.9$ (see (3.5)). Similarly, we obtain $F_1\{M_j\} < 0.1 + 1.2F_2\{M_j \cap [-2B, 2B]\} < F_2\{M_j^\varepsilon\}$.

Thus, we have proved that $F_1\{M_j\} \leqslant \mathrm{e}^{\Delta} F_2\{M_j^\varepsilon\}$ for the cases (a), (c) and (d). Only inequality (3.13) has been established for case (b). But there cannot be more than one of the closed intervals M_j containing zero. Therefore, choosing c_{18} to be sufficiently small and using (3.1), (3.7) and (3.8), we obtain $F_1\{X\} \leqslant F_1\{Y\} \leqslant \mathrm{e}^{\Delta} F_2\{Y^\varepsilon\} + 2\exp(-c\delta/\tau) + \Delta \leqslant F_2\{X^\varepsilon\} + 2\exp(-c\delta/\tau) + 3\Delta$, proving (3.2). □

Proof of Theorem 1.3. – Observe now that, by (1.10) and (2.11),

$$F_1\{u{:}\ |u| \geqslant x\} \leqslant c \max\{\exp(-x^2/8B^2), \exp(-cx/\tau)\}, \quad \text{for any } x \geqslant 0. \tag{3.14}$$

Inequalities (3.2) and (3.14) imply that under conditions (3.1), for any closed set $X_0 \subset \mathbf{R}$,

$$\begin{aligned} F_1\{X_0\} &\leqslant F_1\{X_0 \cap [-y, y]\} + F_1\{u{:}\ |u| \geqslant y\} \\ &\leqslant F_2\{X_0^\varepsilon\} + 3\exp(-c_{24}\delta/\tau) + 3\Delta. \end{aligned} \tag{3.15}$$

The same inequality is valid also for arbitrary Borel set X_0 since $X_0^\varepsilon = (\overline{X}_0)^\varepsilon$, where $\overline{X}_0$ is the closure of X_0. Moreover, since $(\mathbf{R} \backslash X_0^\varepsilon)^\varepsilon \subset \mathbf{R} \backslash X_0$, (3.15) implies that, for arbitrary Borel set X_0, $F_2\{X_0\} \leqslant F_1\{X_0^\varepsilon\} + 3\exp(-c_{24}\delta/\tau) + 3\Delta$. Set now $b^2 = \min\{\delta\tau, c_{19}^2\}$. Assume $\tau \leqslant \delta \min\{c_{18}^2, c_{22}\}$, $\delta \leqslant \min\{c_{19}^2, c_{20}\}/\tau$ and $c_{13} \leqslant c_{18}^2 c_{19}^2$. Then $b^2 = \delta\tau$. By the Strassen–Dudley theorem (see Dudley [3]), one can construct on the same probability space the random variables μ and ν having distributions F_1 and F_2 respectively so that

$$\mathbf{P}(|\mu - \nu| > \varepsilon) \leqslant 3\exp(-c_{24}\delta/\tau) + 3\Delta \leqslant 6\exp(-c\delta/\tau). \tag{3.16}$$

By Lemma A of Berkes and Philipp [2], we can assume that $\mu = \zeta + \omega$, $\nu = \sqrt{1-\chi^2}Z + \phi$, where random variables ζ and Z have the needed distributions and random variables ω and ϕ are independent of ζ and Z respectively and have centered normal distributions with variance b^2. Then, using (3.1), we get

$$\begin{aligned}\mathbf{P}\big(\big|\zeta - \sqrt{1-\chi^2}Z\big| > 3\varepsilon\big) &\leqslant \mathbf{P}\big(|\mu-\nu| > \varepsilon\big) + \mathbf{P}\big(|\omega| > \varepsilon\big) + \mathbf{P}\big(|\phi| > \varepsilon\big)\\ &\leqslant 6\exp(-c\delta/\tau) + 2\exp\big(-c\varepsilon^2/b^2\big)\\ &\leqslant 8\exp(-c\delta/\tau). \qquad (3.17)\end{aligned}$$

Let now $\delta \geqslant \min\{c_{19}^2, c_{20}\}/\tau$ and X_0 be an arbitrary Borel set. Then $b^2 = e^{c\theta}$. If $X_0 \cap [-\delta,\delta] = \emptyset$, then, by (3.14), we have

$$F_1\{X_0\} \leqslant F_1\big\{u\colon |u| \geqslant \delta\big\} \leqslant c\exp(-c\delta/\tau). \qquad (3.18)$$

If $X_0 \cap [-\delta,\delta] \neq \emptyset$, then, by Bernstein's inequality and (2.11),

$$F_1\{X_0\} - F_2\big\{X_0^{2\delta}\big\} \leqslant F_2\big\{u\colon |u| \geqslant \delta\big\} \leqslant 2\exp\big(-\delta^2/4B^2\big) \leqslant 2\exp(-c\delta/\tau). \qquad (3.19)$$

Applying again Lemma A of Berkes and Philipp [2], we construct $\mu, \nu, \zeta, Z, \omega$ and ϕ so that $\mathbf{P}(|\mu-\nu| > 2\delta) \leqslant c\exp(-c\delta/\tau)$ and

$$\begin{aligned}\mathbf{P}\big(\big|\zeta - \sqrt{1-\chi^2}Z\big| > 4\delta\big) &\leqslant \mathbf{P}\big(|\mu-\nu| > 2\delta\big) + \mathbf{P}\big(|\omega| > \delta\big) + \mathbf{P}\big(|\phi| > \delta\big)\\ &\leqslant c\exp(-c\delta/\tau) + 2\exp\big(-c\delta^2/b^2\big)\\ &\leqslant c\exp(-c\delta/\tau). \qquad (3.20)\end{aligned}$$

If $\tau \geqslant \delta\min\{c_{18}^2, c_{22}\} > 0$, then, evidently, for any ζ and Z with needed distributions

$$\begin{aligned}\mathbf{P}\big(\big|\zeta - \sqrt{1-\chi^2}Z\big| > 4\delta\big) &\leqslant 1 \leqslant \exp\big(1 - \min\big\{c_{18}^2, c_{22}\big\}\delta/\tau\big)\\ &\leqslant 3\exp(-c\delta/\tau). \qquad (3.21)\end{aligned}$$

Collecting bounds (3.17), (3.20) and (3.21), we obtain (1.15) with $\lambda = 3\varepsilon = 6e^2\delta$. □

Proof of Theorem 1.4. – By Corollary 1.1, one can construct on the same probability space the random variables μ and ν having distributions F_1 and F_2 respectively so that $|\mu-\nu| \leqslant \gamma(\mu)$, if $|\mu| \leqslant c_{10}\tau^{-1}$, where $\gamma(z)$ is defined in Theorem 1.2. Moreover, $|\nu| \geqslant c_{11}\tau^{-1}$, if $|\mu| \geqslant c_{10}\tau^{-1}$. By Lemma A of Berkes and Philipp [2], we can assume that $\mu = \zeta + \omega$, $\nu = \sqrt{1-\chi^2}Z + \phi$, where ζ and Z have the needed distributions and ω and ϕ are independent of ζ and Z respectively and have centered normal distributions with variance b^2. Then

$$\begin{aligned}&\mathbf{P}\big(\big|\zeta - \sqrt{1-\chi^2}Z\big| > c_9\exp\big(-b^2/72\tau^2\big) + \lambda\big)\\ &\quad\leqslant \mathbf{P}\big(|\mu-\nu|\mathbf{1}\big(|\mu| \leqslant c_{10}\tau^{-1}\big) > c_9\exp\big(-b^2/72\tau^2\big) + \lambda/3\big)\\ &\qquad + \mathbf{P}\big(|\mu|\mathbf{1}\big(|\mu| \geqslant c_{10}\tau^{-1}\big) > \lambda/6\big) + \mathbf{P}\big(|\nu|\mathbf{1}\big(|\nu| \geqslant c_{11}\tau^{-1}\big) > \lambda/6\big)\\ &\qquad + 2\mathbf{P}\big(|\omega| > \lambda/6\big). \qquad (3.22)\end{aligned}$$

Without loss of generality we assume that $c_9\tau \leqslant \lambda/6$. Therefore, by (2.11) and (3.14),

$$\mathbf{P}\big(|\mu-\nu|\mathbf{1}\big(|\mu|\leqslant c_{10}\tau^{-1}\big)>c_9\exp\big(-b^2/72\tau^2\big)+\lambda/3\big)$$
$$\leqslant \mathbf{P}\big(c_9\mu^2\mathbf{1}\big(|\mu|\leqslant c_{10}\tau^{-1}\big)>\lambda/6\tau\big)\leqslant c\exp(-c\lambda/\tau), \tag{3.23}$$
$$\mathbf{P}\big(|\mu|\mathbf{1}\big(|\mu|\geqslant c_{10}\tau^{-1}\big)>\lambda/6\big)\leqslant c\exp(-c\lambda/\tau) \tag{3.24}$$

and

$$\mathbf{P}\big(|\nu|\mathbf{1}\big(|\nu|\geqslant c_{11}\tau^{-1}\big)>\lambda/6\big)\leqslant c\exp(-c\lambda/\tau). \tag{3.25}$$

Inequality (1.16) follows now from (3.22)–(3.25). □

Acknowledgements

The main part of the results of this paper was obtained by the author while visiting the Delaware University. He would like to thank Professor David M. Mason for his hospitality and for the statement of the problem. The author thanks the anonymous referee for careful reading the manuscript and useful suggestions.

REFERENCES

[1] J. Beirlant, D.M. Mason, On the asymptotic normality of L_p-norms of empirical functionals, Math. Methods Statist. 4 (1995) 1–19.

[2] I. Berkes, W. Philipp, Approximation theorems for independent and weakly dependent random vectors, Ann. Probab. 7 (1979) 29–54.

[3] R.M. Dudley, Probability and Metrics, in: Lectures Notes Aarhus Univ., 1976.

[4] W. Feller, An Introduction to Probability Theory and its Applications, Vol. II, Wiley, New York, 1966.

[5] E. Giné, D.M. Mason, A.Yu. Zaitsev, The L_1-norm density estimator process, Ann. Probab., 2001, Accepted for publication.

[6] A.Yu. Zaitsev, Estimates of the Lévy–Prokhorov distance in the multivariate central limit theorem for random variables with finite exponential moments, Theor. Probab. Appl. 31 (1986) 203–220.

[7] A.Yu. Zaitsev, Estimates for quantiles of smooth conditional distributions and multidimensional invariance principle, Siberian Math. J. 37 (1996) 807–831 (in Russian).

[8] A.Yu. Zaitsev, Multidimensional version of the results of Komlós, Major and Tusnády for vectors with finite exponential moments, ESAIM: Probability and Statistics 2 (1998) 41–108.

[9] A.Yu. Zaitsev, Multidimensional version of the results of Sakhanenko in the invariance principle for vectors with finite exponential moments. I; II; III, Theor. Probab. Appl. 45 (2000) 718–738; 46 (2001) 535–561; 744–769.