



Rendiconti
Accademia Nazionale delle Scienze detta dei XL
Memorie di Matematica e Applicazioni
118° (2000), Vol. XXIV, fasc. 1, pagg. 111-120

RITA GIULIANO ANTONINI (*)

On the Rosenblatt coefficient for normalized sums of real random variables (**)

ABSTRACT. — For a given sequence $(X_n)_{n \geq 1}$ of independent identically distributed real random variables, we consider the normalized sums U_n defined by $U_n = (X_1 + \dots + X_n)/\sqrt{n}$, and we give some results on $\mathbf{Cov}(I_A(U_p), I_B(U_q))$ with p, q integers and A, B Borel sets in \mathbb{R} .

Sul coefficiente di Rosenblatt per somme normalizzate di variabili aleatorie reali

SUNTO. — Per un'assegnata successione $(X_n)_{n \geq 1}$ di variabili aleatorie reali, indipendenti e identicamente distribuite, si considerano le somme normalizzate U_n , definite da $U_n = (X_1 + \dots + X_n)/\sqrt{n}$, e si dimostrano alcuni risultati riguardanti le covarianze del tipo $\mathbf{Cov}(I_A(U_p), I_B(U_q))$, con (p, q) coppia d'interi e (A, B) coppia d'insiemi boreliani di \mathbb{R} .

0. - INTRODUCTION

Let $(a_{k, n})_{k \geq 1, n \geq 1}$ be a matrix of real numbers. A classical problem of Probability Theory is the asymptotic behaviour, as n go to infinity, of weighted partial sums such as

$$Z_n = \sum_{k=1}^n a_{k, n} Y_{k, n},$$

where $(Y_{k, n})$ is a sequence of real random variables, satisfying suitable assumptions (see for instance Stout [1], chap. 4). In the last years many authors (see for instance

(*) Indirizzo dell'Autrice: Dipartimento di Matematica, via F. Buonarroti 2, I-56127 Pisa.
E-mail: giuliano@dm.unipi.it

(**) Memoria presentata il 20 aprile 2000 da Giorgio Letta, uno dei XL.

Brosamler [2], Schatte [3], Lacey-Philipp [4]) have considered the case

$$Y_{k,n} = I_{A_{k,n}}(U_k);$$

here $A_{k,n}$ are Borel sets in \mathbb{R} , and U_n denotes the random variable

$$(X_1 + \dots + X_n)/\sqrt{n},$$

where $(X_n)_{n \geq 1}$ is a sequence of independent identically distributed real random variables, with $E[X_1^2] = 1$ and $E[X_1] = 0$, defined on a probability space $(\Omega, \mathfrak{A}, P)$. This leads in a natural way to the problem of evaluating

$$\mathbf{Cov}(I_A(U_p), I_B(U_q)),$$

with p, q integers and A, B Borel sets in \mathbb{R} . In the present paper we give some bounds for the *Rosenblatt coefficient*

$$\sup_{A,x} |\mathbf{Cov}(I_A(U_p), I_{]-\infty, x]}(U_q))|,$$

(where A varies among all Borel sets in \mathbb{R} and x in \mathbb{R}), and for

$$\sup_{A,B} |\mathbf{Cov}(I_A(U_p), I_B(U_q))|,$$

(where A, B vary among all Borel sets in \mathbb{R} , with B included in a fixed set of finite measure). We obtain some results (Theorems 1, 2 and 3) which can be set into the framework of the so-called *almost-orthogonal* random variables; for this kind of variables some interesting laws of large numbers exist (see for ex. Lacey-Philipp [4], pag. 203, Atlagh-Weber [5], pag. 52, and, for a detailed study, Weber [6], sect. 7.4).

2. - NOTATIONS AND FIRST RESULTS

We start by recalling a result on the concentration function of a sum of random variables. Let Q_n be the concentration function of U_n , namely the function defined on \mathbb{R}_+ by

$$Q_n(\lambda) = \sup_x P\{x \leq U_n \leq x + \lambda\}.$$

If the law μ of X_1 is not degenerate, then there is a constant C , depending on μ only, such that, for every real number $\varepsilon > 0$, the inequality

$$Q_n(\varepsilon) \leq C(\varepsilon + 1/\sqrt{n})$$

holds good. See Petrov [7], pag. 49 for a proof. We recall also a result on the difference of the distribution functions of two random variables Y, Z . Let ϕ_Y and ϕ_Z be

the characteristic functions of Y and Z respectively; then we have the inequality

$$|P\{Y \leq x\} - P\{Z \leq x\}| \leq \frac{1}{\pi} \int_{\mathbb{R}} \frac{|\phi_Y(t) - \phi_Z(t)|}{|t|} dt.$$

Without loss of generality, we assume in the sequel $p \leq q$.

LEMMA 1: For every Borel set A in \mathbb{R} and for every Lipschitz function f , with Lipschitz constant L , we have

$$|\mathbf{Cov}(I_A(U_p), f(U_q))| \leq L\sqrt{2p/q}.$$

PROOF: We can assume that the event $H = \{U_p \in A\}$ is not negligible. Denote by $E_H[\cdot]$ the expectation with respect to the conditional probability measure $P(\cdot|H)$, and let $(X'_n)_n$ be an independent copy of the sequence $(X_n)_n$. Put

$$(1) \quad V_q = (X'_1 + \dots + X'_p + X_{p+1} + \dots + X_q)/\sqrt{q}.$$

Then we have

$$E_H[f(V_q)] = E[f(U_q)],$$

hence

$$\begin{aligned} |\mathbf{Cov}(I_A(U_p), f(U_q))| &= \left| \int_H f(U_q) dP - P(H) \int f(U_q) dP \right| \\ &= P(H) |E_H[f(U_q)] - E[f(U_q)]| \\ &= P(H) |E_H[f(U_q)] - E_H[f(V_q)]| \\ &\leq P(H) LE_H[|U_q - V_q|] \\ &\leq LE[|U_q - V_q|]. \end{aligned}$$

By using the second moment, we get

$$\begin{aligned} |\mathbf{Cov}(I_A(U_p), f(U_q))| &\leq L(\mathbf{Var}[U_q - V_q])^{1/2} \\ &= L\left(q^{-1} \sum_{k=1}^p \mathbf{Var}[X_k - X'_k]\right)^{1/2} \\ &= L\sqrt{2p/q}. \end{aligned}$$

The lemma is thus proved.

LEMMA 2: Let ε be a strictly positive real number. For every Borel set A in \mathbb{R} we have

$$|\mathbf{Cov}(I_A(U_p), I_{]-\infty, x]}(U_q))| \leq \frac{2}{\varepsilon} \sqrt{\frac{p}{q}} + 2Q_q(\varepsilon).$$

PROOF: Let the real numbers ε and x be fixed, and denote by f_ε the Lipschitz function defined as

$$f_\varepsilon(t) = I_{]-\infty, x]}(t) + g_\varepsilon(t) = I_{]-\infty, x]}(t) + \left(1 + \frac{x-t}{\varepsilon}\right) I_{]x, x+\varepsilon]}(t).$$

One verifies immediately that f_ε has Lipschitz constant $1/\varepsilon$. Let H be the event $\{U_p \in A\}$; we can assume again that H is not negligible. If Q denotes the conditional probability law $P(\cdot|H)$, we have

$$|\mathbf{Cov}(I_A(U_p), I_{]-\infty, x]}(U_q))| = P(H) |Q\{U_q \leq x\} - P\{U_q \leq x\}|.$$

Moreover

$$\begin{aligned} |Q\{U_q \leq x\} - P\{U_q \leq x\}| &= |\mathbf{E}^Q[(f_\varepsilon - g_\varepsilon)(U_q)] - \mathbf{E}^P[(f_\varepsilon - g_\varepsilon)(U_q)]| \\ &= |\mathbf{E}^Q[(f_\varepsilon - g_\varepsilon)(U_q)] - \mathbf{E}^Q[(f_\varepsilon - g_\varepsilon)(V_q)]| \\ &= |\mathbf{E}^Q[f_\varepsilon(U_q) - f_\varepsilon(V_q)] - \mathbf{E}^Q[g_\varepsilon(U_q) - g_\varepsilon(V_q)]|, \end{aligned}$$

where V_q are the random variables defined in (1). By arguing as in Lemma 1, we get

$$(2) \quad |\mathbf{E}^Q[(f_\varepsilon(U_q) - f_\varepsilon(V_q))]| \leq \frac{2}{\varepsilon P(H)} \sqrt{\frac{p}{q}}.$$

Since we have trivially

$$(3) \quad |\mathbf{E}^Q[(g_\varepsilon(U_q) - g_\varepsilon(V_q))]| \leq \frac{2Q_q(\varepsilon)}{P(H)},$$

from relations (2) and (3) it follows that

$$|Q\{U_q \leq x\} - P\{U_q \leq x\}| \leq \frac{2}{\varepsilon P(H)} \sqrt{\frac{p}{q}} + \frac{2Q_q(\varepsilon)}{P(H)},$$

hence the statement of the lemma.

3. - THE MAIN RESULTS.

THEOREM 1: *There is a constant K , depending on the law of X_1 only, such that, for every pair p, q of integers, we have*

$$(4) \quad \sup_{A, x} |\mathbf{Cov}(I_A(U_p), I_{] - \infty, x]}(U_q))| \leq K \sqrt[4]{\frac{p}{q}}.$$

PROOF: The statement being trivial if X_1 is degenerate, we can assume that X_1 is not degenerate; then, by Lemma 2 and Petrov's theorem (stated in the preceding section), we have for $\varepsilon > 0$,

$$\sup_{A, x} |\mathbf{Cov}(I_A(U_p), I_{] - \infty, x]}(U_q))| \leq \frac{2}{\varepsilon} \sqrt{\frac{p}{q}} + 2C \left(\varepsilon + \frac{1}{\sqrt{q}} \right).$$

The conclusion then follows since the minimum of the function

$$\varepsilon \mapsto (2/\varepsilon) \sqrt{p/q} + 2C(\varepsilon + 1/\sqrt{q})$$

is given by $4\sqrt{C} \sqrt[4]{\frac{p}{q}} + \frac{2C}{\sqrt{q}}$, which is obviously less than $(4\sqrt{C} + 2C) \sqrt[4]{\frac{p}{q}}$. This proves the statement of the theorem.

REMARK When the X_n are gaussian random variables, we have

$$(5) \quad \sup_{A, x} |\mathbf{Cov}(I_A(U_p), I_{] - \infty, x]}(U_q))| = K_0 \sqrt{\frac{p}{q}}.$$

Hence we are faced with the question of what conditions on the law of X_1 can guarantee an «optimal» relation as (5); in other words we are wondering when we are allowed to put the square root of p/q , in place of the fourth one, in relation (4). In what follows we are going to give two sufficient conditions for this to happen. We need two lemmas.

LEMMA 3: *Let ϕ be the characteristic function of X_1 and l a member of $\{0, 1\}$. Assume that there exists an integer r such that the function $t \mapsto |t^l \phi^r(t)|$ is integrable. Then the relation*

$$\sup_{p/q \leq 1/2} \int |t|^l \left| \phi \left(\frac{t}{\sqrt{q}} \right) \right|^{q-p} dt < \infty$$

holds good.

PROOF: Let L be the real function defined by

$$L(t) = E \left[X_1^2 \left(1 \wedge |t| \frac{|X_1|}{3} \right) \right].$$

It is easily seen that L is symmetric, increasing on $[0, \infty[$, bounded by 1 and has limit 0 in $t = 0$, so that there exists a real number \bar{t} on $]0, 1[$ such that $L(\bar{t}) < 1/4$. By the inequality

$$\left| e^{itx} - 1 - itx + \frac{1}{2} t^2 x^2 \right| \leq t^2 x^2 \left(1 \wedge |t| \frac{|x|}{3} \right),$$

(Kallenberg [8], pag. 69) we get $\left| \phi \left(\frac{t}{\sqrt{n}} \right) - 1 + \frac{1}{2} \frac{t^2}{n} \right| \leq \frac{t^2}{n} L \left(\frac{t}{\sqrt{n}} \right)$, for every real number t ; it follows

$$\left| \phi \left(\frac{t}{\sqrt{n}} \right) \right|^{n/2} \leq \left[\left| 1 - \frac{1}{2} \frac{t^2}{n} \right| + \frac{t^2}{n} L \left(\frac{t}{\sqrt{n}} \right) \right]^{n/2}.$$

Hence on the interval $J_n = [-\bar{t}\sqrt{n}, \bar{t}\sqrt{n}]$ we obtain

$$|t^l| \left| \phi \left(\frac{t}{\sqrt{n}} \right) \right|^{n/2} \leq |t^l| \left[1 - \frac{1}{2} \frac{t^2}{n} + \frac{1}{4} \frac{t^2}{n} \right]^{n/2} \leq |t^l| e^{-t^2/8},$$

while, on J_n^c , we have $\left| \phi \left(\frac{t}{\sqrt{n}} \right) \right| \leq \sup_{|u| \geq \bar{t}} |\phi(u)| = d < 1$, where, since $|\phi|$ is integrable, the last inequality follows from a well known result on characteristic functions (see Feller [9], pag. 501). Hence, for every pair of integers p, q , with $p \leq q/2$, one gets the inequalities

$$\begin{aligned} \int |t|^l \left| \phi \left(\frac{t}{\sqrt{q}} \right) \right|^{q-p} dt &\leq \sup_n \int |t|^l \left| \phi \left(\frac{t}{\sqrt{n}} \right) \right|^{n/2} dt \\ &\leq \int |t|^l e^{-t^2/8} dt + \sup_n \int_{J_n^c} |t|^l \left| \phi \left(\frac{t}{\sqrt{n}} \right) \right|^{n/2} dt \\ &\leq C_1 + \sup_n n^{\frac{l+1}{2}} d^{\frac{n}{2}-r} \int |u|^l \phi(u) |^r du, \end{aligned}$$

where C_1 is an absolute constant. The lemma is proved.

LEMMA 4: Let p, q be two integers, with $p \leq q$, and assume that the event $\{U_p \in A\}$ is not negligible. Denote by Φ_q and $\tilde{\Phi}_q$ the characteristic functions of U_q with respect to

P and $P_{\{U_p \in A\}}$; then we have

$$|\mathbf{Cov}(I_A(U_p), I_{] - \infty, x]}(U_q))| \leq \frac{P\{U_p \in A\}}{\pi} \int_{\mathbb{R}} \frac{|\Phi_q(t) - \tilde{\Phi}_q(t)|}{|t|} dt.$$

If in addition the function $t \mapsto |t| |\phi^r(t)|$ is integrable, then for every bounded Borel set B and for every q greater than r , we have

$$|\mathbf{Cov}(I_A(U_p), I_B(U_q))| \leq \frac{P\{U_p \in A\}}{2\pi} \text{meas}(B) \int_{\mathbb{R}} |\Phi_q(t) - \tilde{\Phi}_q(t)| dt.$$

PROOF: The first statement follows from the relation

$$\begin{aligned} |\mathbf{Cov}(I_A(U_p), I_{] - \infty, x]}(U_q))| &= P\{U_p \in A\} |P\{U_q \leq x\} - P_{\{U_p \in A\}}\{U_q \leq x\}| \\ &\leq \frac{P\{U_p \in A\}}{\pi} \int_{\mathbb{R}} \frac{|\Phi_q(t) - \tilde{\Phi}_q(t)|}{|t|} dt, \end{aligned}$$

by the relation on the difference of two distribution function (see section 2). As to the second statement, just note that, for q greater than r , we have

$$\begin{aligned} |\mathbf{Cov}(I_A(U_p), I_{] - \infty, x]}(U_q))| &= P\{U_p \in A\} |P\{U_q \in B\} - P_{\{U_p \in A\}}\{U_q \in B\}| \\ &= \frac{P\{U_p \in A\}}{2\pi} \left| \int_B dx \int_{\mathbb{R}} e^{-itx} [\Phi_q(t) - \tilde{\Phi}_q(t)] dt \right|. \end{aligned}$$

This proves the lemma.

THEOREM 2: Assume that there exists an integer r such that the function $t \mapsto |\phi^r(t)|$ is integrable. Then, for every pair of integers p, q , we have the relation

$$\sup_{A, x} |\mathbf{Cov}(I_A(U_p), I_{] - \infty, x]}(U_q))| \leq K \sqrt{\frac{p}{q}},$$

where K is a constant depending on the law of X_1 only. Moreover, if the function $t \mapsto |t| |\phi^r(t)|$ is integrable too, then for every bounded Borel set B we have

$$\mathbf{Cov}(I_A(U_p), I_B(U_q)) \leq K_1 \text{meas}(B) \sqrt{\frac{p}{q}},$$

where K_1 is a constant depending on the law of X_1 only.

PROOF: Without loss of generality we can assume that $p \leq q/2$, since for $p \geq q/2$ we

have trivially

$$|\mathbf{Cov}(I_A(U_p), I_B(U_q))| = \sqrt{\frac{p}{q}} \sqrt{\frac{q}{p}} |\mathbf{Cov}(I_A(U_p), I_B(U_q))| \leq 4 \sqrt{\frac{p}{q}}.$$

As in the proof of Lemma 1, we assume also that the event $H = \{U_p \in A\}$ is not negligible. Denote by Q the conditional probability measure $P(\cdot|H)$ and put

$$V_q = \frac{X'_1 + \dots + X'_p + X_{p+1} + \dots + X_q}{\sqrt{q}},$$

where (X'_n) is an independent copy of the sequence (X_n) ; then we have

$$\begin{aligned} |\Phi_q(t) - \tilde{\Phi}_q(t)| &= |E^Q[e^{itU_q}] - E^Q[e^{itV_q}]| \\ &= \left| \phi\left(\frac{t}{\sqrt{q}}\right) \right|^{q-p} |E^Q[e^{it\frac{X_1+\dots+X_p}{\sqrt{q}}}] - E^Q[e^{it\frac{X'_1+\dots+X'_p}{\sqrt{q}}}]| \\ &\leq \left| \phi\left(\frac{t}{\sqrt{q}}\right) \right|^{q-p} \frac{2|t|}{P(H)} \sqrt{\frac{p}{q}}. \end{aligned}$$

Lemmas 3 and 4 achieve the conclusion.

Finally, we have a «Berry-Esseen type» result. In detail:

THEOREM 3: *Assume that X_1 has finite absolute third moment. Then we have*

$$\sup_{A, x} |\mathbf{Cov}(I_A(U_p), I_{]-\infty, x]}(U_q))| \leq K_2 \sqrt{\frac{p}{q}},$$

where the constant K_2 depends on the law of X_1 only.

PROOF: As in the proof of the preceding theorem, we can assume $p \leq q/2$. Let (Y_n) be a sequence of random variables and assume that the (Y_n) are independent $\mathcal{N}(0, 1)$ and independent on (X_n) . Put

$$U'_q = \frac{X_1 + \dots + X_p + Y_{p+1} + \dots + Y_q}{\sqrt{q}}.$$

Since $\mathbf{Cov}(I_A(U_p), I_{]-\infty, x]}(U_q))$ is equal to

$$\mathbf{Cov}(I_A(U_p), I_{]-\infty, x]}(U_q) - I_{]-\infty, x]}(U'_q)) + \mathbf{Cov}(I_A(U_p), I_{]-\infty, x]}(U'_q)),$$

it will be enough to prove that the two terms in the above sum are bounded by a number of the form $K \sqrt{\frac{p}{q}}$. As in the above theorem, denote by H the (non negligible)

event $\{U_p \in A\}$ and by Q the conditional probability measure $P(\cdot|H)$; then we have

$$\begin{aligned} |\mathbf{Cov}(I_A(U_p), I_{] - \infty, x]}(U'_q))| &= P(H) |(Q\{U'_q \leq x\} - P\{U'_q \leq x\})| \\ &\leq \frac{P(H)}{\pi} \int \frac{1}{|t|} \left| \int e^{itU'_q} dQ - \int e^{itV'_q} dQ \right| dt. \end{aligned}$$

Here V'_q denotes $\frac{X'_1 + \dots + X'_p + Y_{p+1} + \dots + Y_q}{\sqrt{q}}$, where (X'_n) is a copy of (X_n) , independent on each X_n, Y_n . It follows

$$\begin{aligned} |\mathbf{Cov}(I_A(U_p), I_{] - \infty, x]}(U'_q))| &\leq \frac{P(H)}{\pi} \int e^{-\frac{t^2(q-p)}{2q}} \mathbf{E}^Q[|U'_q - V'_q|] dt \\ &\leq \frac{1}{\pi} \int e^{-\frac{t^2(q-p)}{2q}} \mathbf{E}^P[|U'_q - V'_q|^2]^{1/2} dt \\ &\leq \frac{1}{\pi} \int e^{-\frac{t^2}{2}} \mathbf{E}^P[|U'_q - V'_q|^2]^{1/2} dt \\ &\leq \mathbf{E}^P[|U'_q - V'_q|^2]^{1/2} \\ &\leq 2 \sqrt{\frac{p}{q}}. \end{aligned}$$

Let $a_{p,q}$ be the second term $\mathbf{Cov}(I_A(U_p), I_{] - \infty, x]}(U_q) - I_{] - \infty, x]}(U'_q))$; then

$$|a_{p,q}| = |b_{p,q} - P\{U_p \in A\}(P\{U_q \leq x\} - P\{U'_q \leq x\})|,$$

where $b_{p,q}$ denotes $P\{U_p \in A, U_q \leq x\} - P\{U_p \in A, U'_q \leq x\}$. By the Berry-Esseen inequality, we get

$$\begin{aligned} |b_{p,q}| &= \int_A |F_{p,q}(g(x_1, \dots, x_p)) - N(g(x_1, \dots, x_p))| d\mu(x_1, \dots, x_p) \\ &\leq \frac{E[|X_1|^3]}{\sqrt{q-p}}, \end{aligned}$$

where μ is the law of U_p under P , N the distribution function of the standard gaussian law and $g(x_1, \dots, x_p)$ the real number $\frac{x\sqrt{q} - (x_1 + \dots + x_p)}{\sqrt{q-p}}$. By arguing analogously, we get also

$$|P\{U_p \in A\}(P\{U_q \leq x\} - P\{U'_q \leq x\})| \leq \frac{E[|X_1|^3]}{\sqrt{q-p}}.$$

From the two above relations and the inequality $p \leq q/2$ it follows that

$$|a_{p,q}| \leq 2 \frac{E[|X_1|^3]}{\sqrt{q-p}} \leq 4E[|X_1|^3] \sqrt{\frac{p}{q}}.$$

This concludes the proof.

REMARK: It is for the moment an open problem whether the relation in theorems 2 e 3 holds good even without any assumption on the third moment or the characteristic function of X_1 .

Acknowledgements. I am greatly indebted to Prof. M. Weber for stimulating my interest on this topic; I am also grateful to Prof. Luca Pratelli: I would have not brought this paper to an end without his ideas and suggestions.

REFERENCES

- [1] W. F. STOUT, *Almost sure Convergence*, Academic Press, (1974).
- [2] G. A. BROSAMLER, *An almost everywhere central limit theorem*, Math. Proc. Cambridge Philos. Soc., 104 (1988), 561-574.
- [3] P. SCHATTE, *On strong versions of the almost sure central limit theorem*, Math. Nachr., 137 (1988), 249-256.
- [4] M. T. LACEY - W. PHILIPP, *A note on the Almost Sure Central Limit Theorem*, Stat. and Prob. Letters, 9 (1990), 201-205.
- [5] M. ATLAGH - M. WEBER, *Une nouvelle loi forte des grands nombres*, in *Convergence in Ergodic Theory and Probability* (Bergelson, March, Rosenblatt Eds.), de Gruyter (1996), 41-62.
- [6] M. WEBER, *Entropie métrique et convergence presque partout*, Collection Travaux en cours, 58, Hermann (1998).
- [7] V. V. PETROV, *Sums of Independent Random Variables*, Springer (1975).
- [8] O. KALLENBERG, *Foundations of Modern Probability*, Springer (1998).
- [9] W. FELLER, *An Introduction to Probability Theory and Its Applications*, vol. 2, Wiley (1971).