ON TROTTER METRIC AND ITS AN APPLICATION IN WEAK LAW OF LARGE NUMBERS

Tran Loc Hung

Hue University
Hue, Vietnam
Email: tlhung@hueuni.edu.vn

The main purpose of this note is to present Trotter metric and its an application in weak laws of large numbers.

Keywords: Probability metric, Trotter operator, Weak laws of large numbers.

1. INTRODUCTION

In the solution of a number of problems of probability theory the method of probability metric has attracted much attention and it has successfully been used lately as R. M. Dudley [1], H. Kirschfink [4] and V. M. Zolotarev [9], [10], [11] and [12].

The essence of this method is based on the knowledge of the properties of metrics in spaces of random variables as well as on the principle according to which in every problem of the approximating type a metric as a comparison measure much be selected in accordance with the requirements to its properties.

In recent years several results of mathematics and informatics have been established by using the probability metric approach. Results of this nature may be found in Alison L. Gibbs and Su Francis Edward [2], John E. Hutchinson and Ruschendorf Lunger [3] and Neininger Ralph and Ruschendorf Lunger [5].

The main purpose of the present note is to introduce the definition and properties of a probability metric which is based on Trotter's operator. An application in weak laws of large numbers are indicated.

The received results in the last section are extensions of that given in [6] and [7]. It should be noted that the results for depend random variables have been obtained by H. Kirschfink in [4].

2. PROBABILITY METRICS

Let us denote by Ψ the set of random variables defined on some probability space (Ω, A, P).

Definition 2.1. The mapping $d : \Psi \times \Psi \to [0,\infty)$ is called a probability metric, denoted by denoted by $d(X,Y)$, if

i. $P(X = Y) = 1$ implies $d(X,Y) = 0$

ii. $d(X,Y) = d(Y,X)$

iii. $d(X,Y) \le d(X,Z) + d(Z,Y)$

for random variables X, Y and Z in Ψ.

Definition 2.2. A probability metric d is said to be simple if its values are determined by a pair of marginal distributions P_X and P_Y. In all other cases d is called composed.

It should be noted that, for a simple probability metric the following forms are equivalent

$$d(X,Y) = d(P_X, P_Y).$$

Definition 2.3. A probability metric d is called ideal of order $s \ge 0$ on a subspace $\Psi^* \subset \Psi \times \Psi$, if for $X, Y, Z \in \Psi^*$ with X and Y independent of Z, and $c \ne 0$, the following two properties hold

i. regularity: $d(X+Z, Y+Z) \le d(X,Y)$

ii. homogeneity: $d(cX, cY) \le |c|^s d(X,Y)$.

An interesting consequence of the regularity and homogeneity properties is the semi additively of the metric d: Let $X_1, X_2, \ldots, X_n$, and $Y_1, Y_2, \ldots, Y_n$ be two collections of independent random variables, then one has for X, Y with real numbers $c_j, 1 \le j \le n, s \ge 0$

$$d\left(\sum_{j=1}^{n} X_j, \sum_{j=1}^{n} Y_j\right) \le \sum_{j=1}^{n} |c_j|^s d(X_j, Y_j).$$

We now turn to some examples for illustration of well-known probability metrics.

1. Kolmogorov metric (*Uniform* metric). Let us consider the state space, $\Omega = R^1$ and let us denote $F_X(x) = P(X < x)$ and $G_Y(x) = P(Y < x)$, then the Kolmogorov metric is denoted by

$$d_K(F,G) := \sup_{x \in R^1} |F_X(x) - G_Y(x)|. \tag{2.1}$$

This metric is a simple metric and also called the uniform metric.

2. Levy metric Let the state space $\Omega = R^1$, then the Levy metric is defined by

$$d_L(F,G) = \inf_{\delta > 0} \left\{ G(x-\delta) - \delta \le F(x) \le G(x+\delta) + \delta, \forall x \in R^1 \right\}. \tag{2.2}$$

The Levy metric does metrize weak convergence of measures on R^1 and it is a simple metric.

3. Prokhorov (or Levy-Prokhorov) metric. Let μ and ν be two Borel measures on the metric space (S, d), then the Prokhorov metric d_P is given by

$$d_P(\mu,\nu) := \inf_{\varepsilon > 0} \left\{ \mu(A) \le \nu(A^\varepsilon) + \varepsilon, \quad \text{for all Borel sets} \quad A \in (S,d) \right\}, \tag{2.3}$$

where $$A^\varepsilon := \{ y \in S; \exists x \in A : d(x,y) < \varepsilon \}.$$

This metric is theoretically important because it metrizes weak convergence on any separable metric space.

4. Zolotarev metric.The Zolotarev metric for distributions F_X and F_Y is denoted by

$$d_Z(X,Y) := \sup\left\{ |E[f(X) - f(Y)]|; f \in D_1\left(s; r+1; C\left(R^1\right)\right) \right\}, \tag{2.4}$$

where $C(R^1)$ stands for the class of all bounded, uniformly continuous real-value functions f on R^1, and

$$D_1(s;r+1;C(R^1)):=\left\{f\in C^r(R^1);\left|f^{(r)}(x)-f^{(r)}(y)\right|\le|x-y|^s\right\},$$

here $0<s\le 1;\ r\in N$.

Moreover, $C^r(R^1):=\left\{f\in C(R^1):f^{(j)}\in C(R^1),1\le j\le r\right\},r\in N$.

It should be noted that $C^r(R^1)\subset D_1(s;r+1;C(R^1))\subset C(R^1)$.

The Zolotarev metric $d_Z(X,Y)$ is ideal metric of order 3. It is easy to see that, for X_j and Y_j are pairwise independent random variables,

$$d_Z\left(\sum_{j=1}^{n}X_j,\sum_{j=1}^{n}Y_j\right)\le\sum_{j=1}^{n}d_Z(X_j,Y_j).$$

It is well – known that convergence in d_Z implies weak convergence and it plays great role in some approximation problems.

In addition, we also illustrate some relationships among probability metrics in (2.1), (2.2) and (2.3) as follows.

1. For probability measures μ, ν on R^1 with distribution functions F, G,

$$d_L(F,G)\le d_K(F,G).$$

2. If $G(x)$ is absolutely continuous (with respect to Lebesgue measures), then

$$d_K(F,G)\le(1+\sup_x|G'(x)|)\cdot d_L(F,G).$$

3. For probability measures on R^1,

$$d_L(F,G)\le d_P(F,G).$$

3. THE TROTTER METRIC

The definition and properties of Trotter metric are considered in this section.

Definition 3.1. The Trotter metric d_T of two random variables X and Y related to a function f is defined by

$$d_T(X,Y;f)=\sup_{t\in R^1}\left\{\left|E[f(X+t)]-E[f(Y+t)]\right|;f\in C^r(R^1)\right\}.$$

The most important properties of the Trotter metric are summarized in the following. The proofs easy to get from the properties of the Trotter operator (see [5], [2], [3] and [4] for the complete bibliography).

1. $d_T(X,Y;f)$ is a probability metric.

2. $d_T(X,Y;f)$ is not a simple metric because of neither regularity nor homogeneity holds.

3. If $d_T(X,Y;f)=0$ for $f\in C^r(R^1)$, then $F_X=F_Y$.

4. For all $x \in C(F)$,

$$\lim_{n \to +\infty} F_{X_n}(x) = F_X(x)$$

$$\text{if} \quad \lim_{n \to +\infty} d_T(X_n, X; f) = 0, \quad for \quad f \in C^r(R^1).$$

5. Let $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ be two collections of independent random variables, then

$$d_T\left(\sum_{j=1}^{n} X_j, \sum_{j=1}^{n} Y_j; f\right) \le \sum_{j=1}^{n} d_T\left(X_j, Y_j; f\right).$$

6. In case when $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ be two collections of i. i. d. random variables, then

$$d_T\left(\sum_{j=1}^{n} X_j, \sum_{j=1}^{n} Y_j; f\right) \le n d_T\left(X_1, Y_1; f\right).$$

7. $\sup d_T(X, Y; f) = d_Z(X, Y)$ for every $f \in D_1\left(s; r+1; C\left(R^1\right)\right)$.

4. AN APPLICATIONS IN WEAK LAWS OF LARGE NUMBERS

The following results concerning the rates of convergence in the weak laws of large numbers can illustrate the important role of the Trotter metric in theory of probability.

Let us define $S_n = n^{-1} \sum_{j=1}^{n} X_j$ and let X^0 is denoted the degenerate random variable at point 0. We are now interested in the rate of convergence of the probability metric to zero,

$$d_T(S_n; X^0; f) \to 0 \quad \text{as} \quad n \to +\infty.$$

Theorem. Let $\{X_n, n \ge 1\}$ be a sequence of i. i. d. random variables with zero means and finite r-th absolute moments $E(|X_j|^r) \le M < +\infty$ for $r \ge 1$ and for $j = 1, 2, \ldots, n$. Then, for every $f \in C^r(R^1)$ we have the following estimation

$$d_T(S_n; X^0; f) = o(n^{-(r-1)}) \quad \text{as} \quad n \to +\infty \tag{4.1}$$

Proof. Since $f \in C^r(R^1)$, we have the Taylor expansion

$$f(n^{-1}X_j + t) = \sum_{k=0}^{r} \frac{f^{(k)}(t)}{k!} n^{-k} X_j^k + (r!)^{-1}[f^{(r)}(t + \theta_1 n^{-1} X_j) - f^{(r)}(t)](n^{-1}X_j)^r,$$

where $0 < \theta_1 < 1$

Taking the expectation of both sides of last equation, we have

$$E[f(n^{-1}X_j + t)] = \sum_{k=0}^{r} \frac{f^{(k)}(t)}{k!} n^{-k} E(X_j)^k +$$

$$+(r!)^{-1}\int_{R^1}[f^{(r)}(t+\theta_1 n^{-1}x)-f^{(r)}(t)](n^{-1}x)^r\,dF_{X_j}(x), \quad \text{where} \quad 0<\theta_1<1.$$

Then
$$\left|E[f(n^{-1}X_j+t)]-f(t)\right| \le \sum_{k=1}^{r}[(k!n^k)^{-1}\|f\|_k E(X_j)^k]+$$
$$+[(r!n^r)^{-1}]\int_{R^1}\left|f^{(r)}(t+\theta_1 n^{-1}x)-f^{(r)}(t)\right|\cdot|x|^r\,dF_{X_j}(x), \tag{4.2}$$

where
$$\|f\|_k=\sup_{t\in R^1}\left|f^{(k)}(t)\right|, \quad 0<\theta_1<1.$$

Since $f\in C^r(R^1)$, it follows $\|f\|_k \le C=const$, and because of $E\left|X_j\right|^k \le M<+\infty$ for $k=1,2,\dots,r$, we get

$$\sum_{k=1}^{r}[(k!n^k)^{-1}\|f\|_k E\left|X_j\right|^k]=o(1), \quad \text{as} \quad n\to+\infty. \tag{4.3}$$

Subsequently, by estimating the integral of right side of (4.1), we have

$$[(r!n^r)^{-1}]\int_{R^1}\left|f^{(r)}(t+\theta_1 n^{-1}x)-f^{(r)}(t)\right|\cdot|x|^r\,dF_{X_j}(x)=$$
$$=[(r!n^r)^{-1}]\int_{|x|\le n\delta(\varepsilon)}\left|f^{(r)}(t+\theta_1 n^{-1}x)-f^{(r)}(t)\right|\cdot|x|^r\,dF_{X_j}(x)+$$
$$+[(r!n^r)^{-1}]\int_{|x|>n\delta(\varepsilon)}\left|f^{(r)}(t+\theta_1 n^{-1}x)-f^{(r)}(t)\right|\cdot|x|^r\,dF_{X_j}(x)=I_1+I_2.$$

Because of $f\in C^r(R^1)$, so for every $\varepsilon>0$, there is $\delta(\varepsilon)>0$, such that, for $\left|n^{-1}x\right|\le\delta(\varepsilon)$, we have

$$\left|f^{(r)}(t+\theta_1 n^{-1}x)-f^{(r)}(t)\right|<\varepsilon.$$

It follows that

$$I_1\le\varepsilon\int_{R^1}|x|^r dF_{X_j}(x)=\varepsilon E|X|^r. \tag{4.4}$$

Since $E|X|^r<+\infty$, so we get, for every $\varepsilon>0$, and for n is sufficiently large, we have

$$I_2\le 2\varepsilon\|f\|_k. \tag{4.5}$$

Combining (4.4) and (4.5) and since ε is arbitrary positive number, so we have

$$\sup_t\left|E[f(n^{-1}X_j+t)]-f(t)\right|=o(n^{-r}) \quad \text{as} \quad n\to+\infty. \tag{4.6}$$

Then we have, for $f\in C^r(R^1)$, using the properties of d_T,

$$d_T(S_n;X^0;f)\le n d_T(n^{-1}X_j;n^{-1}X_j^0;f),$$

We get the complete proof

$$d_T(S_n;X^0;f)=o(n^{-(r-1)}) \quad \text{as} \quad n\to+\infty. \ \blacksquare$$

ACKNOWLEDGMENTS

The author would like to thank Professor N. N. Troush from Belarus State University (Minsk, Belarus) for his remarks leading to the writing of this paper.

REFERENCES

1. *Dudley R. M.* Probabilities and Metrics: convergence of laws on metric spaces, with a view to statistical testing (Lect. Notes Series: N45) // Aarhus Univ, Aarhus. 1976.

2. *Alison L. Gibbs, Su Francis Edward.* On choosing and bounding probability metrics // Manuscript version. 2002. P. 1–21

3. *John E.* Hutchinson and Ruschendorf Lundger. Random fractals and probability metrics // Manuscript version. 2002. P. 1–21

4. *Kirschfnk H.* The generalized Trotter operator and weak convergence of dependent random variables in different probability metrics // Results in Mathematics. 1989. V. 15. P. 294–323.

5. *Neininger Ralph, Ruschendorf Ludger.* Rates of convergence for quicksort // Manuscript version. 2002. P. 1–10

6. *Tran Loc Hung.* Applications of the operator methods in the law of large numbers // Vietnam Journal of Mathematics. 1983. V. 11. P. 20–24.

7. *Tran Loc Hung.* The Trotter's operator method in the law of large numbers with random index // Vietnam Journal of Mathematics. 1988. V. 12. P. 4–9.

8. *Trotter H. F.* An elementary proof of the central limit theorem // Arch. Math (Basel). 1959. 10. P. 226–234.

9. *Zolotarev V. M.* Approximation of the distribution of sum of independent variables with values in infinite-dimensional spaces // Theory Prob. Appl. 1976. 21. P. 741–758.

10. *Zolotarev V. M.* Metric distances in spaces of random variables // Math. Uspekhi. 1976. V. 101 (143). N. 3 (11). P. 417–454.

11. *Zolotarev V. M.* Ideal metrics in the problems of probability theory and mathematical statistics // Austral. J. Statist. 1979. 21. (3). P. 193–208.

12. *Zolotarev V. M.* Probability metrics // Theory Prob. Appl. 1983. 28. P. 278–302.