HÀM PHÂN TÁN - MỘT SỐ TÍNH CHẤT VÀ SỰ HỘI TỤ TRONG KHÔNG GIAN L_k

Tô Anh Dũng, Mai Trăng Thanh
Trường Đại học Khoa học Tự nhiên, ĐHQG-HCM
(Bài nhận ngày 29 tháng 03 năm 2007, hòan chỉnh sửa chữa ngày 19 tháng 09 năm 2007)

TÓM TẮT: *Bài báo trình bày một số tính chất, định nghĩa và nghiên cứu sự hội tụ của hàm phân tán bậc k.*
Từ khóa: *Hàm phân tán bậc k, sự hội tụ của dãy hàm phân tán, khoảng cách giữa hai hàm phân tán.*

1.ĐẶT VẤN ĐỀ

Phương pháp L_1-chuẩn có nhiều ứng dụng trong bài toán kiểm định giả thuyết thống kê tuyến tính, xây dựng khoảng tin cậy, phân tích phương sai… Nhiều lĩnh vực của phân tích dữ liệu thống kê dựa trên cơ sở của L_1-chuẩn như ước lượng mật độ, phân tích chuỗi thời gian và phân tích phương sai nhiều chiều.

Trên cơ sở của phương pháp L_1-chuẩn, độ lệch tuyệt đối trung bình $\delta_\mu(X)$ và độ lệch tuyệt đối trung vị $\delta_{Md}(X)$ đã được xây dựng, với

$\delta_\mu(X) = E|X-\mu|$, có thể được coi như chuẩn $\|X-\mu\|_{L_1}$.

$\delta_{Md}(X) = E|X-Md|$, có thể được coi như chuẩn $\|X-Md\|_{L_1}$.

Bên cạnh những ứng dụng tốt trong xác suất thống kê, các độ lệch tuyệt đối $\delta_\mu(X)$ và $\delta_{Md}(X)$ cũng đã bộc lộ những hạn chế của chúng. Cùng với độ phức tạp trong tính toán, $\delta_\mu(X)$ và $\delta_{Md}(X)$ chỉ nói lên được độ lệch giữa biến ngẫu nhiên với trung bình μ hoặc trung vị Md.

Để khắc phục hạn chế nói trên, hàm phân tán đã được xây dựng như một thước đo tổng quát cho độ phân tán các giá trị của biến ngẫu nhiên *X*. Giả sử (Ω, A, P) là không gian xác suất, L_1 là tập hợp các biến ngẫu nhiên khả tích trên không gian xác suất (Ω, A, P). Giả sử biến ngẫu nhiên $X \in L_1$ và F_X là hàm phân phối của biến ngẫu nhiên *X*. Hàm phân tán $D_X(u)$ của biến ngẫu nhiên X được xác định bởi

$D_X(u) = E|X-u|$ với mỗi $u \in \mathbb{R}$.

Một số tính chất của hàm phân tán $D_X(u)$ đã được nghiên cứu trong các bài báo [2], [3], [4], [5], như:

1. $D_X(u)$ là hàm lồi trên $\mathbb{R}$.
2. $\forall X, Y \in L_1$, $\max\limits_{x\in\mathbb{R}} |D_X(x) - D_Y(x)| < \infty$.

3. Với $X \in L_1$ là biến ngẫu nhiên có phân phối liên tục tuyệt đối, thì

$$D_X(u) = u - \mu + 2\int_{x \geq u} (x-u)dF_X(x) = \mu - u + 2\int_{x<u} (u-x)dF_X(x)$$

4. Với $X \in L_1$ là biến ngẫu nhiên có phân phối rời rạc với dãy phân phối xác suất $p_n = P(X = x_n)$, $(n \geq 0)$, thì

$$D_X(u) = u - \mu + 2\sum_{n:x_n \geq u} (x_n - u)p_n = \mu - u + 2\sum_{n:x_n<u} (u - x_n)p_n$$

5. $\lim\limits_{u\to+\infty} (D_X(u) - u) = -EX$ và $\lim\limits_{u\to-\infty} (D_X(u) + u) = EX$

6. Trong trường hợp X là biến ngẫu nhiên có phương sai hữu hạn, thì

$$\int_{-\infty}^{+\infty} \left|D_X(u) - D_{EX}(u)\right| du = \sigma^2$$

Với $D_{EX}(u) = E\left|\mu - u\right|$ là hàm phân tán của biến ngẫu nhiên suy biến tại $EX = \mu$.

7. Với $X, Y \in L_2$, $EX = EY$, thì $\int_{-\infty}^{+\infty} \left|D_X(u) - D_Y(u)\right| du \leq VarX + VarY$

8. Với dãy biến ngẫu nhiên $X_1, X_2, ..., X_n, ...$ bất kì trong không gian xác suất, thì

$$D_{\sum_{i=1}^{n} X_i}(u) \leq \sum_{i=1}^{n} D_{X_i}\left(\frac{u}{n}\right)$$

9. Khi các biến ngẫu nhiên $X_1, X_2, ..., X_n, ...$ cùng phân phối, thì

$$D_{\sum_{i=1}^{n} X_i}(u) \leq nD_{X_1}\left(\frac{u}{n}\right)$$

10. Giả sử $X \in L_1$ và C_F là tập các điểm liên tục của F_X. Khi đó, $\forall u \in C_F$, thì

$$F_X(u) = \frac{1}{2}[D'_X(u) + 1]$$

11. $\lim\limits_{u\to+\infty} D'_X(u) = 1$ và $\lim\limits_{u\to-\infty} D'_X(u) = -1$

Trong thực tế, ngoài L_1 người ta còn gặp những biến ngẫu nhiên khả tích bậc $k > 1$, và theo chúng tôi biết thì chưa có tài liệu nào nghiên cứu các hàm phân tán bậc k của các biến đó, vì vậy chúng tôi lấy vấn đề này làm mục tiêu cho bài báo.

2. HÀM PHÂN TÁN BẬC k

Với biến ngẫu nhiên X trong không gian L_k và F_X là hàm phân phối tương ứng của X, hàm phân tán bậc k của biến ngẫu nhiên X được xác định bởi

$$D_X^k(u) = E\left|X - u\right|^k \text{ với mỗi } u \in \mathbb{R}$$

Có thể chứng minh không khó khăn các bất đẳng thức thường gặp trong xác suất sau đây đối với hàm phân tán bậc k.

a. Bất đẳng thức dạng Markov

Với mọi $\varepsilon > 0$, với mọi $u \in (-\infty, +\infty)$ và với biến ngẫu nhiên $X \in L_k$, ta có

$$P(|X-u| \geq \varepsilon) \leq \frac{1}{\varepsilon^k} D_X^k(u)$$

b. Bất đẳng thức dạng Liapunov

Với mọi $u \in (-\infty, +\infty)$, với s, t thỏa $0 < s < t$, với biến ngẫu nhiên $X \in L_t$, ta có

$$\sqrt[s]{D_X^s(u)} \leq \sqrt[t]{D_X^t(u)}$$

c. Bất đẳng thức dạng Minkowski

Với mọi $u \in (-\infty, +\infty)$, với mọi p thoả $1 \leq p < \infty$, với hai biến ngẫu nhiên $X, Y \in L_p$, ta có

$$\sqrt[p]{D_{X+Y}^p(u)} \leq \sqrt[p]{D_X^p(\frac{u}{2})} + \sqrt[p]{D_Y^p(\frac{u}{2})}$$

d. Bất đẳng thức dạng - c_r

Với mọi $u \in \mathbb{R}$, với hai biến ngẫu nhiên $X, Y \in L_r$, ta có

$$D_{X+Y}^r(u) \leq c_r D_X^r\left(\frac{u}{2}\right) + c_r D_Y^r\left(\frac{u}{2}\right)$$

Với $r \leq 1$, tương ứng với $c_r = 1$; hoặc $r \geq 1$, tương ứng với $c_r = 2^{r-1}$.

3. ĐỊNH LÍ VỀ SỰ HỘI TỤ CỦA HÀM PHÂN TÁN TRONG KHÔNG GIAN L_k

Sau đây là định lí về mối liên hệ giữa sự hội tụ của dãy hàm phân tán bậc k $D_{X_n}^k(u)$ với sự hội tụ của dãy biến ngẫu nhiên $X, X_1, X_2, ..., X_n, ...$

Định lí 1

Giả sử $X, X_1, X_2, ..., X_n, ...$ là dãy biến ngẫu nhiên thuộc không gian L_k.

Nếu tồn tại p>k>0 sao cho

$$\sup_n E|X_n|^p < \infty$$

và $X_n \xrightarrow{P} X$

Khi đó

$D_{X_n}^k(u) \to D_X^k(u)$ *khi* $n \to +\infty$, *với mọi* $u \in \mathbb{R}$

Chứng minh.

Từ điều kiện $X_n \xrightarrow{P} X$

và với $p>k$

$$\sup_n E|X_n|^p < \infty$$

ta có được

$$E|X_n - X|^k \to 0 \text{ , hay } X_n \xrightarrow{k} X.$$

a. Khi $k \le 1$

Theo bất đẳng thức c_r, ta có được

$$E|X_n - u|^k \le E|X_n - X|^k + E|X - u|^k$$

Hay

$$E|X_n - u|^k - E|X - u|^k \le E|X_n - X|^k \tag{1.1}$$

Cũng theo bất đẳng thức c_r, ta có được

$$E|X - u|^k \le E|X - X_n|^k + E|X_n - u|^k$$

Suy ra

$$E|X - u|^k \le E|X - X_n|^k + E|X_n - u|^k$$

Hay

$$E|X_n - u|^k - E|X - u|^k \ge -E|X_n - X|^k \tag{1.2}$$

Do đó, từ (1.1), (1.2) và từ kết luận $E|X_n - X|^k \to 0$, ta có được

$$\left|E|X_n - u|^k - E|X - u|^k\right| \le E|X_n - X|^k \to 0 \tag{1.3}$$

b. Khi $k > 1$

Tương tự, cũng theo bất đẳng thức Minkowski, ta có được

$$\left|E^{\frac{1}{k}}|X_n - u|^k - E^{\frac{1}{k}}|X - u|^k\right| \le E^{\frac{1}{k}}|X_n - X|^k \to 0 \tag{1.4}$$

Từ (1.3) và (1.4), ta suy ra

$$E|X_n - u|^k \to E|X - u|^k \quad \text{khi } n \to +\infty$$

Hay

$$D_{X_n}^k(u) \to D_X^k(u) \qquad \text{khi } n \to +\infty \qquad \square$$

4. HÀM PHÂN TÁN BẬC 2 CỦA CÁC BIẾN NGẪU NHIÊN TRỰC GIAO.

Trong không gian L_1, với dãy biến ngẫu nhiên $X_1, X_2, ..., X_n$ bất kì, thì

$$D_{\sum_{i=1}^n X_i}(u) \le \sum_{i=1}^n D_{X_i}\left(\frac{u}{n}\right)$$

Ta phát triển mối liên hệ này trong không gian L_2.

Định lí 2 (Một dạng định lí Pithagore).

Trước hết, với hai biến ngẫu nhiên X, Y bất kì trong không gian L_2, $X-u$ v $Y-u$ trực giao. Khi đó

$$D^2_{X+Y}(u) = D^2_X\left(\frac{u}{2}\right) + D^2_Y\left(\frac{u}{2}\right)$$

Chứng minh.

Áp dụng tính chất của biến ngẫu nhiên trực giao cho hai biến ngẫu nhiên $\left(X-\frac{u}{2}\right)$ và $\left(Y-\frac{u}{2}\right)$, ta được

$$E\left|X-\frac{u}{2}+Y-\frac{u}{2}\right|^2 = E\left|X-\frac{u}{2}\right|^2 + E\left|Y-\frac{u}{2}\right|^2$$

$$\Rightarrow E\left|X+Y-u\right|^2 = E\left|X-\frac{u}{2}\right|^2 + E\left|Y-\frac{u}{2}\right|^2$$

Hay

$$D^2_{X+Y}(u) = D^2_X\left(\frac{u}{2}\right) + D^2_Y\left(\frac{u}{2}\right) \qquad \square$$

Ta mở rộng định lí trên với dãy biến ngẫu nhiên $X_1, X_2, ..., X_n$ trong không gian L_2.

Định lí 3

Cho $X_1, X_2, ..., X_n$ là dãy biến ngẫu nhiên trực giao, và $X_i + X_j = \frac{u}{n}$ $\left(i, j \in \overline{1,n}\right)$. Khi đó

$$D^2_{\sum_{i=1}^n X_i}(u) = \sum_{i=1}^n D^2_{X_i}\left(\frac{u}{n}\right)$$

Ngoài ra, khi $n \to \infty$, thì

$$D^2_{\sum_{i=1}^n X_i}(u) \to \sum_{i=1}^{\infty} D^2_{X_i}\left(\frac{u}{n}\right).$$

5. ĐỊNH LÍ VỀ KHOẢNG CÁCH CỦA HAI HÀM PHÂN TÁN TRONG KHÔNG GIAN L_1

Định lí sau đã được chứng minh trong [4], và trong bài này, chúng tôi đưa ra một cách chứng minh đơn giản hơn.

Định lí 4

Với $D_{EX}(u) = E\left|\mu - u\right|$ là hàm phân tán của biến ngẫu nhiên suy biến tại $EX = \mu$, $X \in L_1$ và X là biến ngẫu nhiên có phương sai hữu hạn thì

$$\int_{-\infty}^{+\infty} \left|D_X(u) - D_{EX}(u)\right| du = \sigma^2 \qquad (4.1)$$

Chứng minh.

Ta viết lại vế trái

$$\int_{-\infty}^{+\infty} \left|D_X(u) - D_{EX}(u)\right| du = \int_{-\infty}^{\mu} \left|D_X(u) - D_{EX}(u)\right| du + \int_{\mu}^{+\infty} \left|D_X(u) - D_{EX}(u)\right| du$$

Với

$$\int_{-\infty}^{\mu} \left|D_X(u) - D_{EX}(u)\right| du = \int_{-\infty}^{\mu} \left|E|X-u| - E|EX-u|\right| du$$

$$= \int_{-\infty}^{\mu} \left|E(|X-u| - |\mu - u|)\right| du$$

$$= \int_{-\infty}^{\mu} \left| \int_{-\infty}^{+\infty} \left(|x-u| - |\mu-u|\right) dF(x) \right| du \qquad (4.2)$$

Ta khử dấu trị tuyệt đối ở (4.2) bằng cách chia biến x làm ba khoảng $x < u < \mu$, $u < x < \mu$ và $u < \mu < x$, khi đó (4.2) được viết thành

$$\int_{-\infty}^{\mu} \left| \int_{x<u<\mu} [(u-x)-(\mu-u)]dF(x) + \int_{u<x<\mu} [(x-u)-(\mu-u)]dF(x) + \int_{u<\mu<x} [(x-u)-(\mu-u)]dF(x) \right| du$$

$$= \int_{-\infty}^{\mu} \left| \int_{x<u<\mu} (2u - x - \mu) dF(x) + \int_{u<x<\mu} (x-\mu) dF(x) + \int_{u<\mu<x} (x-\mu) dF(x) \right| du$$

$$= \int_{-\infty}^{\mu} \left| \int_{-\infty}^{u} (2u - x - \mu) dF(x) + \int_{u}^{+\infty} (x-\mu) dF(x) \right| du$$

$$= \int_{-\infty}^{\mu} \left| \int_{-\infty}^{u} 2u dF(x) - \int_{-\infty}^{u} x dF(x) - \int_{-\infty}^{u} \mu dF(x) + \int_{u}^{+\infty} x dF(x) - \int_{u}^{+\infty} \mu dF(x) \right| du$$

$$= \int_{-\infty}^{\mu} \left| \int_{-\infty}^{u} 2u dF(x) - \int_{-\infty}^{u} x dF(x) + \int_{u}^{+\infty} x dF(x) - \mu \right| du \qquad (4.3)$$

vì

$$\int_{u}^{+\infty} x dF(x) = \mu - \int_{-\infty}^{u} x dF(x)$$

nên (4.3) có thể viết

$$\int_{-\infty}^{\mu} \left| \int_{-\infty}^{u} 2u dF(x) - \int_{-\infty}^{u} x dF(x) + \left(\mu - \int_{-\infty}^{u} x dF(x)\right) - \mu \right| du$$

$$= \int_{-\infty}^{\mu} \left| \int_{-\infty}^{u} 2u dF(x) - \int_{-\infty}^{u} 2x dF(x) \right| du$$

$$= \int_{-\infty}^{\mu} \left| \int_{-\infty}^{u} 2(u-x) dF(x) \right| du$$

$$= \int_{-\infty}^{u} \left(\int_{-\infty}^{\mu} 2(u-x) du \right) dF(x)$$

$$= \int_{-\infty}^{u} \left(\int_{x}^{\mu} 2(u-x) du \right) dF(x) \quad \text{(do lúc này } x < u < \mu \text{)}$$

$$= \int_{-\infty}^{u} [2 \frac{(u-x)^2}{2} \Big|_{x}^{\mu}] dF(x)$$

$$= \int_{-\infty}^{u} (\mu - x)^2 dF(x)$$

Thực hiện biến đổi tương tự

$$\int_{\mu}^{+\infty} \left| D_X(u) - D_{EX}(u) \right| du = \int_{\mu}^{+\infty} \left| E|X-u| - E|EX-u| \right| du$$

$$= \int_{\mu}^{+\infty} \left| E(|X-u| - |\mu - u|) \right| du$$

$$= \int_{\mu}^{+\infty} \left| \int_{-\infty}^{+\infty} (|x-u| - |\mu - u|) dF(x) \right| du \qquad (4.4)$$

Ta khử dấu trị tuyệt đối ở (4.4) bằng cách chia biến x thành ba khoảng $x < \mu < u$, $\mu < x < u$ và $\mu < u < x$, khi đó (4.4) được viết thành

$$\int_{\mu}^{+\infty} \left| \int_{x<\mu<u} [(u-x)-(u-\mu)] dF(x) + \int_{\mu<x<u} [(u-x)-(u-\mu)] dF(x) + \int_{\mu<u<x} [(x-u)-(u-\mu)] dF(x) \right| du$$

$$= \int_{\mu}^{+\infty} \left| \int_{x<\mu<u} (\mu - x) dF(x) + \int_{\mu<x<u} (\mu - x) dF(x) + \int_{\mu<u<x} (\mu + x - 2u) dF(x) \right| du$$

$$= \int_{\mu}^{+\infty} \left| \int_{-\infty}^{u} (\mu - x) dF(x) + \int_{u}^{+\infty} (\mu + x - 2u) dF(x) \right| du$$

$$= \int_{\mu}^{+\infty} \left| \int_{-\infty}^{u} \mu dF(x) - \int_{-\infty}^{u} x dF(x) + \int_{u}^{+\infty} \mu dF(x) + \int_{u}^{+\infty} x dF(x) - \int_{u}^{+\infty} 2u dF(x) \right| du$$

$$= \int_{\mu}^{+\infty} \left| \mu - \int_{-\infty}^{u} x dF(x) + \int_{u}^{+\infty} x dF(x) - \int_{u}^{+\infty} 2u dF(x) \right| du \qquad (4.5)$$

Do

$$-\int_{-\infty}^{u} xdF(x) = -\mu + \int_{u}^{+\infty} xdF(x)$$

Nên (4.5) trở thành

$$\int_{\mu}^{+\infty}\left|\mu-\mu+\int_{u}^{+\infty} xdF(x)+\int_{u}^{+\infty} xdF(x)-\int_{u}^{+\infty} 2udF(x)\right|du$$

$$=\int_{\mu}^{+\infty}\left|\int_{u}^{+\infty} 2xdF(x)-\int_{u}^{+\infty} 2udF(x)\right|du$$

$$=\int_{\mu}^{+\infty}\left|\int_{u}^{+\infty} 2(x-u)dF(x)\right|du$$

$$=\int_{u}^{+\infty}\left(\int_{\mu}^{+\infty} 2(x-u)du\right)dF(x)$$

$$=\int_{u}^{+\infty}\left(\int_{\mu}^{x} 2(x-u)du\right)dF(x)$$

$$=\int_{u}^{+\infty}\left(-\left[2\frac{(x-u)^2}{2}\Big|_{\mu}^{x}\right]\right)dF(x)$$

$$=\int_{u}^{+\infty}\left(-(x-u)^2\Big|_{\mu}^{x}\right)dF(x)$$

$$=\int_{u}^{+\infty}(x-\mu)^2 dF(x)$$

Từ đó ta có

$$\int_{-\infty}^{+\infty}\left|D_X(u)-D_{EX}(u)\right|du=\int_{-\infty}^{u}(\mu-x)^2 dF(x)+\int_{u}^{+\infty}(x-\mu)^2 dF(x)$$

$$=\int_{-\infty}^{+\infty}(x-\mu)^2 dF(x)=E(X-\mu)^2=\sigma_X^2$$

□

Nhận xét. Định lí trên không còn đúng khi $k>1$.

6.KẾT LUẬN

Hướng sắp tới nghiên cứu các biểu thức giải tích của hàm phân tán bậc $k \geq 1$ cho các lớp phân phối khả phân vô hạn và phân phối ổn định, tìm các ứng dụng trong các định lý giới hạn địa phương và đặc biệt là trong thống kê.

Định lý 2 của bài báo này, một định lý dạng Pithagore, được chứng minh cho bậc 2, tuy nhiên có thể khảo sát thêm cho các bậc $k > 2$.

THE DISPERSION FUNCTION- SOME PROPERTIES AND THE CONVERGENCE IN SPACE L_k

To Anh Dung, Mai Trang Thanh
University of Natural Sciences, VNU-HCM

ABSTRACT: *This paper presents some properties and gives the definition of k^{th} dispersion function and studies its convergence.*

Key words: *k^{th} dispersion function, convergence of sequence of dispersion functions, distance between two dispersion functions.*

TÀI LIỆU THAM KHẢO

[1]. M.Loeve, *Probability Theory*- D.Van Nostrand Company, Canada (1963).

[2]. Trần Lộc Hùng, Nguyễn Văn Sơn, *Some connections of Weak Convergence with the Convergence of the Dispersion function*, Vietnam Journal of Mathematics 31:3 , (2003).

[3]. Phạm Gia Thụ, Trần Lộc Hùng, *Bayesian estimation under estimation constraint*, Acta Mathematica Vietnamica, Volume 28, Number 2, (2003).

[4]. J.Munoz-Perez, A.Sanchez-Gomez, *A characterization of the distribution function: the dispertion function*, Statistics & Probability Letter 10 (1990).

[5]. J.Munoz-Perez, A.Sanchez-Gomez, *Dispersive ordering by dilation*, J.Appl.Prob.27 (1990).