LÝ THUYẾT ƯỚC LƯỢNG VÀ THỐNG KÊ ε — ĐỦ

Lê Ngọc Bường

Bài báo này trình bày một số kết quả về lý thuyết ước lượng dựa vào khái iệm thống kê ε — đủ và khái niệm (G, ε, δ) — đầy đủ. Khái niệm (G, ε, δ) — ầy đủ đã được trình bày ở [1]. Để dễ dàng theo dõi các kết quả, ở đây nhắc lại ịnh nghĩa và một vài tính chất quan trọng của thống kê ε — đủ.

Chúng tôi luôn luôn xét các vấn đề trên cấu trúc thống kê $(æ, \mathcal{A}, \mathcal{P})$ trong đó là không gian các đối tượng quan sát, A là σ — trường Boren trên æ còn **P** là ọ các độ đo xác suất cho trên A được các xác định bởi tham số θ.

$$P = [P_\theta, \theta \in \Theta]$$

Θ là không gian các tham số. Họ P luôn luôn được giả thiết là bị làm trội ởi độ đo dương, σ — hữu hạn μ và khi đó ta ký hiệu các mật độ Radom — Nyôdym của P_θ đối với μ là $f(x, \theta)$. Khoảng cách ρ giữa hai độ đo xác suất p và được xác định như sau:

$$\rho(p, q) = \sup_{A \in \mathcal{A}^*} \frac{|P(A) - q(A)|}{P(A) + q(A)}$$

rong đó $$A^* = [A \in A : P(A) + q(A) \neq 0]$$

òn khoảng cách ρ^* giữa hai hàm không âm f và g được cho bởi

$$\rho^*(f, g) = \sup_{x \in æ^*} \frac{|f(x) - g(x)|}{f(x) + g(x)}$$

rong đó

$$æ^* = [x \in æ : f(x) + g(x) \neq 0$$

iả sử T là thống kê ánh xạ từ (æ, A, P) vào $(Y, \mathcal{B}, P^T)$

Định nghĩa 1: Thống kê T được gọi là thống kê ε — đủ cho họ P nếu có òn tại họ các độ đo xác suất $Q = [q_\theta, \theta \in \Theta]$ sao cho T là thống kê đủ cho ọ Q và

$$\sup_{\theta \in \Theta} \rho(P_\theta, q_\theta) \leqslant \varepsilon$$

Định lý 1: Giả sử rằng T là thống kê ε — đủ cho họ **P**. Khi đó tồn tại họ ác độ đo xác suất $Q = [q_\theta]$ sao cho T là thống kê đủ cho họ Q và

$$\sup_{\theta \in \Theta} \rho(p_\theta, q_\theta) \leqslant \frac{2\varepsilon}{1 + \varepsilon^2}$$

à $\forall A \in T^{-1}(\mathcal{B})$ ta có $P_\theta(A) = q_\theta(A)$. Đồng thời với mọi hàm X là P — khả tích hì cũng là Q — khả tích và

$$|E_{q_\theta}(X/T) - E_{p_\theta}(X/T)| \leqslant \frac{2\varepsilon}{1 + \varepsilon^2} [E_{q_\theta}(|X|/T) + E_{p_\theta}(|X|/T)] \; (P_\theta.\text{h. k. n})$$

Định lý 2: Giả thiết rằng tồn tại các hàm $h(x)$ và $g(x, \theta)$ sao cho

i) $h(x) \geqslant 0$, $g(x, \theta) \geqslant 0 \quad \forall (x, \theta)$

ii) $h(x)$ là A — đo được còn $g(x, \theta)$ là $T^{-1}(B)$ đo được

iii) h bằng không trên mọi tập $\mathcal{P}$ — bỏ qua.

iV)
$$\sup_{\theta \in \Theta} \rho^*(f, h.g) \leqslant \varepsilon$$

Khi đó T là thống kê $\dfrac{2\varepsilon}{1-\varepsilon}$ — đủ cho họ $\mathcal{P}$.

Định nghĩa 2: Giả sử G là một lớp con các hàm A — đo được và P — kh tích, Khi đó cấu trúc thống kê $(æ, A, P)$ được gọi là (G, ε, δ) — đầy đủ nếu từ b đẳng thức

$$|E_\theta \varphi| \leqslant \varepsilon \quad \forall \theta \in \Theta$$

trong đó $\varphi \in G$ thì suy ra

$$|\varphi(x)| \leqslant \delta \quad (P.h.k.n)$$

Nếu G bao gồm tất cả các hàm A — đo được và P — khả tích thì ta nói rằng cấ trúc $(æ, A, P)$ là (ε, δ) — đầy đủ. Nếu G bao gồm tất cả các hàm A — đo đượ P — khả tích và giới nội thì ta nói cấu trúc thống kê $(æ, A, P)$ là (ε, δ) — đầy đ giới nội.

Thống kê T được gọi là (G, ε, δ) — đầy đủ nếu cấu trúc $(æ, T^{-1}(B), P)$ là ($\varepsilon, \delta)$ — đầy đủ.

Định nghĩa 3: Giả sử $g(\theta)$ là hàm của tham số θ ánh xạ đo được từ Θ vào R Thống kê $\varphi(x)$ ánh xạ đo được từ $(æ, A)$ vào (Y, B^*) được gọi là ước lượng ε chệch cho $g(\theta)$ nếu

$$|E_\theta \varphi - g(\theta)| \leqslant \varepsilon \quad \forall \theta \in \Theta$$

Định nghĩa 4: Giả sử $g(\theta)$ là hàm của tham số θ còn S là một ước lượn không chệch cho $g(\theta)$ tức là

$$E_\theta s = g(\theta) \, \forall \theta \in \Theta$$

Khi đó thống kê $U(x)$ được gọi là ước lượng (ε, δ) — tốt hơn $S(x)$ cho $g(\theta)$ nế

i) $U(x)$ là ước lượng ε — chệch cho $g(\theta)$

ii) $D_\theta U \leqslant \delta + D_\theta S \quad \forall \theta \in \Theta$

trong đó D_θ là toán tử lấy phương sai theo P_θ.

Định lý 3: Giả sử $g(\theta)$ là hàm của tham số θ có ước lượng không chệch $S($ thỏa mãn:

$$E_\theta S^2 \leqslant M^2 < +\infty \quad \forall \theta \in \Theta$$

Giả sử $T(x)$ là thống kê ε — đủ cho họ P. Khi đó có tồn tại một ước lượn $\left(\dfrac{4M\varepsilon}{(1-\varepsilon)^2}, \dfrac{40M^2\varepsilon}{(1-\varepsilon)^4}\right)$ — tốt hơn $S(x)$ cho $g(\theta)$. Ước lượng đó là hàm của $T(x)$.

Chứng minh: theo định lý 1 có tồn tại họ các độ đo xác suất $Q = [q_\theta]$ sa cho T là thống kê đủ cho họ Q.

Đặt
$$U(x) = E_{q_\theta}[S(.)/T(x)]$$

Rõ ràng $U(x)$ là một thống kê và cũng theo định lý 1 thì

$$|U(x) - E_{p_\theta}(S/T(x))| \leqslant \frac{2\varepsilon}{1+\varepsilon^2}[\widetilde{U}(x) + E_{p_\theta}(|S|/T(x))] \quad (P_\theta.h.k.n) \quad (1)$$

trong đó
$$\widetilde{U}(x) = E_{p_\theta}(|S|/T(x))$$

Ta cũng có:

$$[\widetilde{U}(x) - E_{p_\theta}(|S|/T(x))] \leqslant \frac{2\varepsilon}{1+\varepsilon^2}[\widetilde{U}(x) + E_{p_\theta}(|S|/T(x))] \, (P_\theta.h.k.n)$$

ừ đó ta có:

$$\widetilde{U}(x) \leqslant \left(\frac{1+\varepsilon}{1-\varepsilon}\right)^2 E_{p_\theta}(|S|/T(x)) \quad (p_\theta.\text{h.k.n}) \tag{2}$$

hế (2) vào (1) ta có:

$$|U(x) - E_{p_\theta}(S/T(x))| \leqslant \frac{4\varepsilon}{(1-\varepsilon)^2} E_{p_\theta}(|s|/T(x)) \quad (p_\theta \text{ h.k.n.}) \tag{3}$$

ừ (3) ta có:

$$|E_{p_\theta} U(x) - g(\theta)| = |E_{p_\theta} E_{p_\theta}(U/T(x)) - E_{p_\theta}(S/T(x))| \leqslant E_{p_\theta}(|U(x) -$$

$$E_{p_\theta}(S/T(x)|/T(x)) \leqslant \frac{4\varepsilon}{(1-\varepsilon)^2} E_{p_\theta}|S| \leqslant \frac{4M\varepsilon}{(1-\varepsilon)^2} \tag{4}$$

ậy (4) cho thấy kết luận U(x) là ước lượng $\frac{4M\varepsilon}{(1-\varepsilon)^2}$ — chệch cho $g(\theta)$

iếp theo ta có:

$$D_{p_\theta} S(x) = E_{p_\theta}(S(x) - g(\theta))^2 = E_{p_\theta}(S(x) - U(x) + E_{p_\theta} U(x) - g(\theta))^2 +$$

$$+ E_{p_\theta}(U(x) - E_{p_\theta} U(x))^2 + 2E_{p_\theta}(U(x)) - E_{p_\theta} U(x))(S(x) - U(x) + E_{p_\theta} U(x) -$$

$$- g(\theta)) \geqslant D_{p_\theta} U(x) + 2E_{p_\theta}(U(x) - E_{p_\theta} U(x))(S(x) - U(x) + E_{p_\theta} U(x) - g(\theta))$$

ặt

$$U^*(x) = U(x) - E_{p_\theta} U(x)$$

$$S^*(x) = S(x) - E_{p_\theta} S(x)$$

a có

$$D_{p_\theta} S(x) \geqslant D_{p_\theta} U(x) + 2E_{p_\theta} U^*(x)(S^*(x) - U^*(x)) \tag{5}$$

ét biểu thức:

$$= E_{p_\theta} U^*(x)(S^*(x) - U^*(x)) = E_{p_\theta} U^*(x)(E_{p_\theta}(S^*/T) - U^*(x)) \tag{6}$$

i

$$E_{p_\theta}(S^*/T(x) = E_{p_\theta}(S/T(x)) - g(\theta)$$

ên

$$E_{p_\theta}(S^*/T(x) - U^*(x) = E_{p_\theta}(S/T(x) - U^*(x) - g(\theta)) = E_{p_\theta}(S/T(x) -$$

$$- E_{q_\theta}(S/T(x)) + E_{p_\theta}(U(x) - g(\theta))$$

ừ đó ta có:

$$|E_{p_\theta}(S/T(x)) - U^*(x)| \leqslant |E_{p_\theta}(S/T(x)) - E_{q_\theta}(S/T(x))| + |E_{p_\theta}(U(x) -$$

$$- g(\theta))| \leqslant \frac{4\varepsilon}{(1-\varepsilon)^2} E_{p_\theta}(|S|/T(x)) + \frac{4M\varepsilon}{(1-\varepsilon)^2}.$$

ặt khác ta cũng có

$$|U^*(x)| = |E_{q_\theta}(S/T(x)) - E_{p_\theta}(U(x))| \leqslant E_{q_\theta}(|S|/T(x)) + |g(\theta)| +$$

$$+ |E_{p_\theta}(U(x) - g(\theta))| \leqslant \left(\frac{1+\varepsilon}{1-\varepsilon}\right)^2 E_{p_\theta}(S/T(x)) + |E_{p_\theta}(S)| +$$

$$+ \frac{4M\varepsilon}{(1-\varepsilon)^2} \leqslant \left(\frac{1+\varepsilon}{1-\varepsilon}\right)^2 E_{p_\theta}(|S|/T(x)) + M + \frac{4M\varepsilon}{(1-\varepsilon)^2}$$

hay các bất đẳng thức trên vào (6) ta có

$$|A| \leqslant E_{p_\theta}\left\{\left(\frac{1+\varepsilon}{1-\varepsilon}\right)^2 E_{p_\theta}(|S|/T(x)) + M + \frac{4M\varepsilon}{(1-\varepsilon)^2}\right\} \times$$

$$\times\left\{\frac{4\varepsilon}{(1-\varepsilon)^2}E_{P_\theta}(|S|/T(x))+\frac{4M\varepsilon}{(1-\varepsilon)^2}\right\}=\frac{4\varepsilon(1+\varepsilon)^2}{(1-\varepsilon)^4}E_{P_\theta}(|S|/T(x))^2+$$
$$+\frac{4\varepsilon(1-\varepsilon)^2+4\varepsilon(1+\varepsilon)^2}{(1-\varepsilon)^4}ME_{P_\theta}(E_{P_\theta}(|S|/T(x))+\frac{4\varepsilon(1-\varepsilon)^2+16\varepsilon^2}{(1-\varepsilon)^4}M^2\leqslant$$
$$\leqslant\frac{40M^2\varepsilon}{(1-\varepsilon)^4}$$

Thay (7) vào (5) ta có $D_\theta S(x) \geqslant D_\theta U(x) - \dfrac{40M^2\varepsilon}{(1-\varepsilon)^4}$

***Định nghĩa 5*:** Hai thống kê T_1 và T_2 được gọi là Δ – tương đương nếu

$$|T_1(x) - T_2(x)| \leqslant \Delta\ (\mathcal{P}.h.k.n)$$

***Định lý 4*:** Nếu T_1 và T_2 là hai thống kê Δ – tương đương thì

$$|E_\theta T_1 - E_\theta T_2| \leqslant \Delta$$
$$|\sqrt{D_\theta T_1} - \sqrt{D_\theta T_2}| \leqslant 2\Delta$$

***Chứng minh*:** Từ định nghĩa ta có

$$T_2(x) - \Delta \leqslant T_1(x) \leqslant T_2(x) + \Delta\ (P.h.k.n)$$

Lấy kỳ vọng hai vế ta có

$$E_\theta T_2(x) - \Delta \leqslant E_\theta T_1(x) \leqslant E_\theta T_2(x) + \Delta$$

Từ đó

$$|E_\theta T_1 - E_\theta T_2| \leqslant \Delta$$

Ta cũng có:

$$|(E_\theta T_1 - T_1) - (E_\theta T_2 - T_2)| \leqslant 2\Delta$$

Nên $|E_\theta T_1 - T_1| \leqslant |E_\theta T_2 - T_2| + 2\Delta$

$$\Rightarrow E_\theta(T_1 - E_\theta T_1)^2 \leqslant E_\theta(T_2 - E_\theta T_2)^2 + 4\Delta E_\theta|T_2 - E_\theta T_2| + 4\Delta^2 \leqslant$$
$$\leqslant E_\theta(T_2 - E_\theta T_2)^2 + 4\Delta\sqrt{E_\theta(T_2 - E_\theta T_2)^2} + 4\Delta^2 = (\sqrt{E_\theta(T_2 - E_\theta T_2)^2} + 2\Delta)^2$$

Vậy $$\sqrt{D_\theta T_1} \leqslant \sqrt{D_\theta T_2} + 2\Delta$$

Hoàn toàn tương tự ta có:

$$\sqrt{D_\theta T_2} \leqslant \sqrt{D_\theta T_1} + 2\Delta$$

Vậy $$|\sqrt{D_\theta T_2} - \sqrt{D_\theta T_1}| \leqslant 2\Delta$$

***Định lý 5*:** Giả sử $T(x)$ là thống kê ε – đủ và (η, ξ) – đầy đủ cho họ $\mathcal{P}$. Giả s $g(\theta)$ là hàm của tham số θ có tồn tại ước lượng không chệch $S(x)$ thỏa mãn

$$E_\theta S^2 \leqslant M^2 < +\infty \qquad \forall\theta \in \Theta$$

Khi đó có tồn tại ước lượng $\left(\dfrac{4M\varepsilon}{(1-\varepsilon)^2}, \dfrac{40M^2\varepsilon}{(1-\varepsilon)^4}\right)$ – tốt hơn $S(x)$ cho $g(\theta)$. Ướ lượng đó là hàm của T và là ước lượng duy nhất theo nghĩa $\dfrac{8M\varepsilon\xi}{(1-\varepsilon)^2\eta}$ – tươn trong lớp tất cả các ước lượng $\dfrac{4M\varepsilon}{(1-\varepsilon)^2}$ – chệch cho $g(\theta)$ và là hàm của $T(x)$

***Chứng minh*:** Sự tồn tại ước lượng $U(x)$ là hàm của $T(x)$ và là $\left(\dfrac{4M\varepsilon}{(1-\varepsilon)^2}, \dfrac{40M^2\varepsilon}{(1-\varepsilon)^4}\right)$ tốt hơn $S(x)$ cho $g(\theta)$ đã được chỉ ra ở định lý 3. Vậy giờ ta chứng minh tin duy nhất theo nghĩa $\dfrac{8M\varepsilon\xi}{(1-\varepsilon)^2\eta}$ – tương đương trong lớp tất cả các ước lượn là $\dfrac{4M\varepsilon}{(1-\varepsilon)^2}$ – chệch cho $g(\theta)$ và là hàm của $T(x)$. Muốn vậy ta chỉ còn chứng min rằng nếu $U_1(x)$ là hàm của $T(x)$ và là ước lượng $\dfrac{4M\varepsilon}{(1-\varepsilon)^2}$ – chệch bất kỳ cho g(θ thì

$$| U(x) - U_1(x) | \leqslant \frac{8M\varepsilon}{(1-\varepsilon)^2}.$$

Ta có :

$$| E_\theta U(x) - g(\theta) | \leqslant \frac{4M\varepsilon}{(1-\varepsilon)^2}$$

$$| E_\theta U_1(x) \; | \; g(\theta) | \leqslant \frac{4M\varepsilon}{(1-\varepsilon)^2}$$

$$\Rightarrow | E_\theta (U_1(x) - U_1(x)) | \leqslant \frac{8M\varepsilon}{(1-\varepsilon)^2}$$

Do tính $T^{-1}(B)$ – đo được của $U-U_1$ và tính (η, ζ) – đầy đủ của $T(x)$ nên ta có :

$$| U(x) - U_1(x) | \leqslant \frac{8M\varepsilon\zeta}{(1-\varepsilon)^2\eta} (\mathcal{P}.\text{h.k.n})$$

Định nghĩa 6: Họ các độ đo xác suất $\mathcal{P} = [P_\theta, \theta \in \Theta]$ được gọi là thỏa mãn điều kiện chính qui nếu họ các mật độ $f(x, \theta)$ đối với μ của nó thỏa mãn các điều kiện sau :

i) Θ là toàn bộ đường thẳng thực hoặc một khoảng trong đường thẳng thực.

ii) $\frac{d}{d\theta} f(x,\theta)$ tồn tại và hữu hạn P_θ – h.k.n với mỗi θ

iii) $\int \left| \frac{d^i}{d\theta^i} f(x,\theta) \right| d\mu < +\infty$ đối với bất kỳ $\theta \in \Theta$
$i = 1, 2$

iV) $E_\theta \left\{ \frac{d}{d\theta} \log f(x,\theta) \right\}^2 < +\infty \; \forall \theta \in \Theta$

Đặt $S(x,\theta) = \frac{d}{d\theta} \log f(x,\theta)$

Định nghĩa 7: Hàm

$$I_x(x) = E_\theta S^2(x, \theta)$$

Được gọi là lượng thông tin Fisher.

Vì giới hạn của bài báo nên dưới đây chúng tôi chỉ phát biểu một định lý mà không chứng minh.

Định lý 6: Giả sử họ $\mathcal{P} = [P_\theta, \theta \in \Theta]$ là họ các độ đo xác suất thỏa mãn điều kiện chính qui. Giả thiết thêm rằng

V) $\sup_{\theta \in \Theta} \left| \frac{d^2}{d\theta^2} f(x,\theta) \right| \leqslant m(x) < +\infty$ và

$$\int m(x)\, d\mu \leqslant m < +\infty$$

Vi) $I_x(\theta) \leqslant M \; \forall \theta \in \Theta$

Giả sử rằng $T(x)$ là thống kê ε – đủ cho họ $\mathcal{P}$ và cảm sinh ra họ các độ đo P_θ^T chính qui. Khi đó ta có

$$I_x(\theta) \geqslant I_t(\theta) \geqslant I_x(\theta) - 1000M\sqrt{\varepsilon}.$$

TÀI LIỆU THAM KHẢO

1. LÊ NGỌC BƯỜNG. Khái niệm (G, ε, δ) đầy đủ, Tạp chí Khoa học trường Đại học Tổng hợp Hà Nội, Số 3—1987, 1—6.

LÊ NGỌC BƯỜNG

ESTIMATION THEORY AND ε — SUFFICIENT STATISTICS

In this paper we consider some properties of estimations of parameter functions for the family of probability measures which have ε—sufficient statistic. As an example we prove existence of an estimation function of ε—sufficient statistic which is$\left(\frac{4M\varepsilon}{(1-\varepsilon)^2}, \frac{4M^2\varepsilon}{(1-\varepsilon)^4}\right)$ — better than an unbiased estimation of a paramater function. Further we prove the information inequality in form

$$I_t(\theta) \leqslant I_x(\theta) \leqslant I_t(\theta) + 1000M\sqrt{\varepsilon}$$

where T is a ε — sufficient statistic of θ

Khoa Toán—Cơ—Tin học
Trường Đại học Tổng hợp Hà Nội

Nhận bài ngày 28—10—198