

MỤC LỤC

Lời nói đầu	1
1 Một số kiến thức cần dùng	3
1.1 Hồi quy hàng số	3
1.2 Hàm tổn thất	8
1.3 Thống kê đủ và đầy đủ	9
1.4 Ước lượng không chệch phương sai bé đều nhất (UMVU) . .	11
1.5 Ước lượng chấp nhận được và ước lượng tối ưu	11
1.6 Một số kiến thức khác	12
2 Tính tối ưu trong lớp các ước lượng không chệch và lớp các ước lượng đều	15
2.1 Tính chất tối ưu trong lớp các ước lượng không chệch	15
2.2 Tham số tịnh tiến và lớp các ước lượng đều	20
3 Đặc trưng của họ phân phối chuẩn	29
3.1 Đặc trưng của phân phối chuẩn thông qua tính chấp nhận được của các ước lượng tuyến tính tối ưu của tham số tịnh tiến đối với hàm tổn thất dạng bình phương	29
3.2 Điều kiện để trung bình mẫu chấp nhận được trong một lớp con các ước lượng của tham số tịnh tiến	38

3.2.1	Các ước lượng không chệch chỉ phụ thuộc trung bình mẫu và một thống kê tuyến tính khác	38
3.2.2	Lớp các ước lượng đa thức	41
3.3	Đặc trưng của luật chuẩn bởi tính tối ưu của hàm của ước lượng $\hat{L} = \sum c_j^0 X_j$	49
3.4	Đặc trưng luật chuẩn bởi tính chấp nhận được và tính tối ưu của ước lượng bình phương bé nhất trong mô hình Gauss - Markov	52
	Kết luận	59
	Tài liệu tham khảo	61

LỜI CẢM ƠN

Trước hết tác giả xin bày tỏ lòng biết ơn sâu sắc đến thầy giáo, PGS.TS Đào Hữu Hồ, chính thầy là người đã trực tiếp giảng dạy 3 chuyên đề rất khó của Xác suất - Thống kê và đã tận tình chỉ bảo để tác giả hoàn thành Luận văn này.

Tác giả cũng xin bày tỏ lòng biết ơn chân thành đến các thầy giáo tham gia giảng dạy các chuyên đề cao học toán Khóa 2008 - 2010 nói chung và lớp cao học Xác suất - Thống kê 2008 - 2010 nói riêng. Xin cảm ơn Khoa Toán - Cơ - Tin học đã luôn tạo điều kiện thuận lợi để học viên yên tâm học tập, nghiên cứu.

Cuối cùng, cho tôi được gửi lời cảm ơn đến tất cả bạn bè, người thân đã luôn cổ vũ, động viên tôi trong suốt khóa học và quá trình làm Luận văn.

Hà Nội, ngày 13 tháng 11 năm 2010

Tác giả

Trần Thị Thu Lan

BẢNG KÝ HIỆU

$\mathcal{X} :=$ Không gian mẫu.

$\mathcal{A} :=$ σ -đại số các tập con của $\mathcal{X}$.

$\mathbb{P}_\theta :=$ Độ đo xác suất trên $\mathcal{A}$, $\theta \in \Theta$.

$\Theta :=$ Không gian tham số.

$\mathcal{L}_\theta^2 :=$ Không gian Hilbert các hàm đo được trên $(\mathcal{X}, \mathcal{A})$ bình phương khả tích theo độ đo $\mathbb{P}_\theta$ đối với tích vô hướng thông thường.

$$\mathcal{L}^2 := \bigcap_{\theta \in \Theta} \mathcal{L}_\theta^2.$$

$\mathcal{H}_\theta :=$ Tập hợp tất cả các ước lượng không chệch của không trong $\mathcal{L}_\theta^2$.

$$\mathcal{H} := \bigcap_{\theta \in \Theta} \mathcal{H}_\theta.$$

$\mathcal{T} :=$ Lớp các ước lượng đều.

$\mathcal{L}_\theta^2(T) :=$ Bao đóng theo chuẩn trong $\mathcal{L}_\theta^2$ của tập tất cả các đa thức theo T .

$$\mathcal{L}^2(T) := \bigcap_{\theta \in \Theta} \mathcal{L}_\theta^2(T).$$

LỜI NÓI ĐẦU

Nghiên cứu tính chất đặc trưng của các phân phối xác suất có một lịch sử lâu đời. Thực chất, Gauss (1809) đã chứng tỏ (với một số điều kiện) rằng ước lượng hợp lý cực đại của tham số tịnh tiến của một phân phối là trung bình mẫu nếu và chỉ nếu đó là phân phối chuẩn.

Nói một cách khái quát, một bài toán đặc trưng có dạng: Giả sử rằng với một vectơ ngẫu nhiên X , có một họ các phân phối $\mathcal{F}$ sao cho nếu $\mathcal{L}(X) \in \mathcal{F}$ thì X phải có tính chất $\mathcal{P}$ nào đó; và ngược lại, tức là, chỉ ra rằng nếu vectơ ngẫu nhiên X có tính chất $\mathcal{P}$ thì $\mathcal{L}(X) \in \mathcal{F}$.

Có hai thành phần của bài toán đặc trưng: Họ các phân phối $\mathcal{F}$ và tính chất $\mathcal{P}$. Các họ phân phối có thể là phân phối chuẩn, phân phối đều, phân phối mũ, Beta, Gamma, phân phối hình học, phân phối Poisson, Cauchy hay Wishart, trong đó phân phối chuẩn được quan tâm nhiều nhất. Các tính chất rất đa dạng, khó có thể đề cập cụ thể ở đây. Tuy nhiên, ta có thể chia thành một số loại tính chất như là: một số hàm cụ thể của X có phân phối nào đó, một số hàm của X là độc lập với nhau, một số hàm của X có hồi quy đặc biệt đối với một số hàm khác của X , một số hàm của X là ước lượng hợp lý cực đại đối với một tham số chưa biết nào đó trong phân phối của hàm X hoặc một số hàm của X là ước lượng tối ưu (hoặc chấp nhận được) trong một lớp các ước lượng tham số nào đó...

Như vậy, có thể thấy rằng bài toán đặc trưng rất đa dạng, phong phú. Trong luận văn này, tác giả chỉ tập trung nghiên cứu tính chất đặc trưng thông qua tính chấp nhận được và tính tối ưu của các ước lượng. Chúng ta có thể lấy một ví dụ: Chúng ta biết rằng trung bình mẫu $\bar{X}$ được xem là ước lượng của trung bình chuẩn và nó có một số tính chất tối ưu. Một câu hỏi nảy sinh tự nhiên là liệu phân phối chuẩn có phải là phân phối duy nhất để trung bình mẫu $\bar{X}$ có một trong các tính chất tối ưu đó hay không? Đây chính là bài toán đặc trưng thông qua tính chất tối ưu (hoặc chấp nhận được) của các ước

lượng. Trong luận văn tác giả sẽ trình bày những vấn đề khái quát của ví dụ nêu trên.

Luận văn bao gồm ba chương:

Chương 1: Trình bày một số kiến thức chuẩn bị liên quan đến nội dung chính của luận văn. Bao gồm hồi quy hằng số, hàm tổn thất, thống kê đủ và đầy đủ, ước lượng không chệch phương sai bé đều nhất (UMVU), ước lượng chấp nhận được, ước lượng tối ưu và một số kiến thức khác.

Chương 2: Nghiên cứu tính tối ưu trong lớp các ước lượng không chệch, tham số tịnh tiến và lớp các ước lượng đều, ước lượng Pitman.

Chương 3: Đây là nội dung chính của luận văn. Trong chương này, tác giả sẽ nghiên cứu bốn vấn đề chính. Một là đặc trưng của phân phối chuẩn thông qua tính chấp nhận được của ước lượng tuyến tính tối ưu của tham số tịnh tiến với hàm tổn thất dạng bình phương. Hai là điều kiện để trung bình mẫu là ước lượng chấp nhận được trong một số lớp con của các ước lượng cho tham số tịnh tiến. Ba là đặc trưng của phân phối chuẩn thông qua tính tối ưu của các hàm của ước lượng không chệch tuyến tính tối ưu cho tham số tịnh tiến. Cuối cùng, tác giả nghiên cứu đặc trưng của phân phối chuẩn thông qua tính chấp nhận được và tính tối ưu của các ước lượng bình phương bé nhất trong mô hình Gauss-Markov.

Chương 1

Một số kiến thức cần dùng

1.1 Hồi quy hằng số

Định nghĩa 1.1.1. Cho hai biến ngẫu nhiên X và Y , kỳ vọng $\mathbb{E}Y$ tồn tại. Khi đó Y được gọi là có hồi quy hằng số đối với X nếu $\mathbb{E}(Y|X) = \mathbb{E}Y$ hầu khắp nơi.

Người ta chứng minh được rằng Y có hồi quy hằng số đối với X khi và chỉ khi $\mathbb{E}(Ye^{itX}) = \mathbb{E}Y \cdot \mathbb{E}e^{itX}$. Những định lý sau đây chỉ ra đặc trưng của phân phối thông qua tính hồi quy hằng số của một thống kê tuyến tính trên một thống kê tuyến tính khác.

Lấy $X_1, \dots, X_n$ là các biến ngẫu nhiên độc lập không nhất thiết cùng phân phối. Xem xét hàm tuyến tính $a_1X_1 + \dots + a_nX_n$ với các hệ số khác không. Bằng phép co giãn thích hợp, ta đưa về dạng

$$L = X_1 + \dots + X_n. \quad (1.1.1)$$

Hơn nữa, lấy

$$b_{i1}X_1 + \dots + b_{in}X_n, \quad i = 1, \dots, p \geq 2$$

là p hàm tuyến tính độc lập tuyến tính, mà sau một phép biến đổi thích hợp có thể đưa về dạng chuẩn tắc

$$\begin{aligned}
M_1 &= X_1 + c_{11}X_{p+1} + \cdots + c_{1n}X_n \\
&\dots \\
M_p &= X_p + c_{p1}X_{p+1} + \cdots + c_{pn}X_n
\end{aligned} \tag{1.1.2}$$

Chúng ta ký hiệu ma trận các hệ số c_{ij} bởi $\mathbf{C}$ có cấp là $p \times (n - p)$ và cột thứ i được ký hiệu là C_i .

Định lý 1.1.1. Cho $p \geq 2$ và giả thiết các vector cột của $\mathbf{C}$ không trực giao với nhau và cũng không trực giao với bất cứ cột nào của ma trận đơn vị $\mathbf{I}_p$ cỡ $p \times p$. Nếu $\mathbb{E}X_i = 0$, khi đó điều kiện

$$\mathbb{E}(L|M_1, \dots, M_p) = 0 \tag{1.1.3}$$

suy ra rằng các biến X_i có phân phối chuẩn.

Hệ quả 1.1.1. Cho $L_1, \dots, L_n$ là các hàm tuyến tính độc lập tuyến tính của các biến ngẫu nhiên $X_1, \dots, X_n$ với các hệ số khác không. Giả sử rằng $\mathbb{E}X_i = 0, \forall i$. Khi đó điều kiện

$$\mathbb{E}(L_1|L_2, \dots, L_n) = 0 \tag{1.1.4}$$

đảm bảo rằng các biến X_i có phân phối chuẩn, nếu $n \geq 3$.

Ta xét hai tập gồm q và $p \geq 2$ các hàm tuyến tính được giả thiết là độc lập tuyến tính. Chúng ta viết lại dưới dạng chuẩn tắc

$$\begin{aligned}
L_j &= \sum_{u=1}^p a_{ju}X_u + \sum_{i=p+1}^n b_{j,i-p}X_i, & j &= 1, \dots, q, \\
M_i &= X_i + \sum_{u=p+1}^n c_{i,u-p}X_u, & i &= 1, \dots, p \geq 2.
\end{aligned} \tag{1.1.5}$$

Định lý 1.1.2. Cho $X_1, \dots, X_n$ là các biến ngẫu nhiên độc lập với $\mathbb{E}X_i = 0, \forall i$. Xét các hàm tuyến tính $L_1, \dots, L_q$ và $M_1, \dots, M_p$ được định nghĩa bởi (1.1.5). Xét các ma trận $\mathbf{A} = (a_{ij})$, $\mathbf{B} = (b_{ij})$, $\mathbf{C} = (c_{ij})$ với các cấp tương ứng là $q \times p$, $q \times (n - p)$ và $p \times (n - p)$. Nếu mỗi cột của các

ma trận **A** và **B** có ít nhất một phần tử khác không và không có cột nào của **C** tỉ lệ với bất cứ một cột nào của ma trận đơn vị $\mathbf{I}_p$ thì đẳng thức $\mathbb{E}(L_i|M_1, \dots, M_p) = 0, i = 1, \dots, q$ suy ra các X_i có phân phối chuẩn.

Ta xét các hàm độc lập tuyến tính

$$L_i = X_i + \sum_{u=q+1}^n w_{i,u-q} X_u, \quad i = 1, \dots, q, \quad (1.1.6)$$

$$L_j = X_j + \sum_{u=1}^q w_{j-q,i} X_i, \quad j = q+1, \dots, n. \quad (1.1.7)$$

Các biến X_{q+i} và X_{q+j} được gọi là liên kết được trong (1.1.6) nếu chúng đồng thời xuất hiện với các hệ số khác 0 trong ít nhất một trong những phương trình (1.1.6). Ta có hệ quả

Hệ quả 1.1.2. Nếu trong (1.1.6) các biến X_{q+i} đều liên kết với ít nhất một biến trong số X_{q+j} và hơn nữa, trong biểu diễn (1.1.7) không có cột hệ số w_{ji} nào là vectơ không thì đẳng thức $\mathbb{E}(L_i|L_{q+1}, \dots, L_n) = 0, i = 1, \dots, q$ sẽ dẫn đến khẳng định là các biến X_j có phân phối chuẩn nếu $n - q \geq 2$.

Bây giờ ta nhắc lại khái niệm phân phối chia vô hạn.

Định nghĩa 1.1.2. Hàm đặc trưng f được gọi là chia vô hạn nếu với mỗi số nguyên, dương n , tồn tại hàm đặc trưng $f_n(t)$ thỏa mãn:

$$f(t) = (f_n(t))^n, \quad \forall t \in \mathbb{R}^1.$$

Khai triển Levy của hàm đặc trưng chia vô hạn:

$$\ln f(t) = i\beta t - \frac{\sigma^2}{2}t^2 + \int_{(0,\infty)} h(t,u)dN(u) + \int_{(-\infty,0)} h(t,u)dM(u)$$

trong đó $\beta \in \mathbb{R}^1$, $\sigma \geq 0$, $h(t,u) = e^{itu} - 1 - itu/(1+u^2)$, và các hàm M , N thỏa mãn các điều kiện sau đây:

(a) M và N là không giảm và liên tục phải trên các khoảng $(-\infty, 0)$ và $(0, \infty)$ tương ứng;

(b) $M(-\infty) = N(\infty) = 0$; và

(c) Các tích phân $\int_{(-a,0)} u^2 dM(u)$ và $\int_{(0,a)} u^2 dN(u)$ hữu hạn với bất kỳ $a > 0$.

Ta xét trường hợp các biến $X_1, \dots, X_n$ là các biến ngẫu nhiên độc lập, cùng phân phối với $\mathbb{E}X_i = 0$. Giả sử có hai dạng tuyến tính $a_1X_1 + \dots + a_nX_n$ và $b_1X_1 + \dots + b_nX_n$ sao cho

$$\mathbb{E}(a_1X_1 + \dots + a_nX_n | b_1X_1 + \dots + b_nX_n) = 0. \quad (1.1.8)$$

Không mất tổng quát, ta giả sử $|b_n| = \max_i |b_i| \neq 0$ và $|a_n| = \max_i |a_i| \neq 0$. Khi đó, với phép co giãn thích hợp, ta đưa (1.1.8) về dạng

$$\mathbb{E}(-\alpha_iX_i - \dots - \alpha_{n-1}X_{n-1} + X_n | \beta_1X_1 + \dots + \beta_{n-1}X_{n-1} + X_n) = 0. \quad (1.1.9)$$

trong đó $|\beta_j| \leq 1$ với $j = 1, \dots, n-1$.

Định lý 1.1.3. Cho $X_1, \dots, X_n$ là các biến ngẫu nhiên độc lập, cùng phân phối thỏa mãn (1.1.8). Giả sử rằng có chính xác k số β_i là khác nhau về độ lớn (chúng ta có thể chọn là $\beta_1, \dots, \beta_k$) và khác ± 1 . Cho $j = 1, \dots, k$, ta đặt

$$\epsilon_j = \sum \{\sigma_i : b_i = b_j\} \quad , \quad \epsilon_0 = \sum \{\sigma_i : \beta_i = 1\}$$

và

$$\epsilon'_0 = \sum \{\sigma_i : \beta_i = -1\}.$$

Nếu $\epsilon'_0 = 0$, $\gamma_i = \frac{\sigma_i}{1 - \epsilon_0} > 0$ với $i = 1, \dots, k$ và λ là nghiệm thực duy nhất của phương trình $\sum_1^k \gamma_i |\beta_i|^\lambda = 1$, khi đó ta có các khẳng định sau

(i) các X_i là suy biến nếu $\lambda \leq 1$ hoặc $\lambda > 2$

(ii) các X_i là chuẩn (có thể suy biến) nếu $\lambda = 2$; và

(iii) với $1 < \lambda < 2$, các X_i có hàm đặc trưng f mà trong biểu diễn Levy $L(\mu, \sigma^2, M, N)$ đối với $\ln f$, ta có $\sigma = 0$ và tùy vào bản chất của vector $\mathbf{B} = (B_1, \dots, B_n)$ trong đó $B_j = -\ln \beta_j$, các hàm M và N có những dạng như sau

(a) ít nhất hai thành phần của $\mathbf{B}$ là không thông ước với nhau, khi đó $M(u) = \xi|u|^{-\lambda}$ và $N(u) = -\eta u^{-\lambda}$, trong đó ξ và η là hai số thực không âm thỏa mãn $\xi + \eta > 0$; hơn nữa $\xi = \eta$ nếu $p < n$ hoặc nếu $\lambda = 1$, vì vậy, với một số thực c nào đó, $f(t)e^{ict}$ là hàm đặc trưng của một luật đối xứng ổn định với số mũ λ ; nếu $p = n$ và $\lambda \neq 1$, khi đó f là hàm đặc trưng của một phân phối ổn định với số mũ λ .

(b) các thành phần của $\mathbf{B}$ là thông ước, tức tồn tại $\rho > 0$ sao cho $m_j = B_j/\rho$ là các số tự nhiên có ước chung lớn nhất bằng 1. Hơn nữa trong số các $m_1, \dots, m_p$ có số lẻ, trong các số $m_{p+1}, \dots, m_n$ có số lẻ. Khi đó

$$M(u) = \frac{\xi(\ln |u|)}{|u|^\lambda} \quad , \quad N(u) = \frac{-\eta(\ln u)}{|u|^\lambda}$$

trong đó ξ và η là các hàm không âm liên tục phải trên đường thẳng thực với chu kỳ ρ , hơn nữa nếu $p < n$ thì $\xi = \eta$.

(c) các thành phần của $\mathbf{B}$ là thông ước. Hơn nữa $m_1, \dots, m_p$ là các số chẵn và các số $m_{p+1}, \dots, m_n$ là các số lẻ và $p < n$. Khi đó

$$M(u) = \frac{\xi(\ln |u|) + \eta(\ln |u|)}{|u|^\lambda} \quad , \quad N(u) = \frac{-\xi(\ln u) + \eta(\ln u)}{|u|^\lambda}$$

trong đó ξ và η là các hàm liên tục phải, $\xi(x+p) = \xi(x)$ và $\eta(x+p) = -\eta(x)$ với mọi x .

Bây giờ ta xét (1.1.8) với $X_1, \dots, X_n$ là các biến ngẫu nhiên độc lập cùng phân phối với $\mathbb{E}X_1 = 0$ và f là hàm đặc trưng của X_1 . Giả sử $a_i \neq 0$, $b_i \neq 0$, và đặt $c_i = \frac{a_i}{b_i}$, khi đó (1.1.8) dẫn đến phương trình

$$\prod_{i=1}^n [f(b_i t)]^{c_i} = 1 \quad (1.1.10)$$

với t nằm trong lân cận nào đó của 0. Không mất tổng quát, ta giả sử rằng $c_1 > 0, \dots, c_p > 0$ và $c_{p+1} < 0, \dots, c_n < 0$. Khi đó (1.1.10) có thể viết lại dưới dạng

$$\prod_{i=1}^p [f(b_i t)]^{c_i} = \prod_{i=p+1}^n [f(b_i t)]^{|c_i|}. \quad (1.1.11)$$

Ta giả sử thêm là

$$\max\{|b_1|, \dots, |b_p|\} \neq \max\{|b_{p+1}|, \dots, |b_n|\}. \quad (1.1.12)$$

Khi đó ta có định lý

Định lý 1.1.4. Cho $X_1, \dots, X_n$ là các biến ngẫu nhiên độc lập cùng phân phối với $\mathbb{E}X_1 = 0$, thỏa mãn các điều kiện (1.1.8) và (1.1.12). Lấy γ là không điểm lớn nhất của hàm $G(\lambda) = \sum c_i |b_i|^\lambda$. γ chắc chắn tồn tại và dương. Nếu X_i có mô men hữu hạn cấp $2m$, trong đó $m = [(\gamma + 2)/2]$. Khi đó X_1 có phân phối chuẩn (có thể suy biến). Đặc biệt, khẳng định trên đúng nếu X_1 có mô men mọi cấp.

1.2 Hàm tổn thất

Xét không gian xác suất $(\mathcal{X}, \mathcal{A}, \mathbb{P}_\theta)$, trong đó $\mathcal{X}$ là không gian mẫu, $\mathcal{A}$ là một σ -đại số và $\{\mathbb{P}_\theta\}$ là họ các độ đo xác suất được đánh chỉ số bởi tham số $\theta \in \Theta$. Giả sử thống kê $T(X)$ là một ước lượng cho hàm tham số $t(\theta)$, và $r(T, t)$ là một hàm tổn thất không âm. Kỳ vọng toán của hàm tổn thất

$$R(T, \theta) = \mathbb{E}_\theta r(T, t)$$

được gọi là hiểm của T khi giá trị đúng của tham số là θ .

Một số ví dụ về hàm tổn thất:

1. Hàm tổn thất dạng bình phương: $r(T, t) = (T - t)^2$
2. Hàm tổn thất dạng trị tuyệt đối: $r(T, t) = |T - t|$
3. Hàm tổn thất dạng 0 - 1: $r(T, t) = \begin{cases} 0, & |T - t| \leq b \\ 1, & |T - t| > b \end{cases}$

với $b > 0$ nào đó.

Định nghĩa 1.2.1 (Hàm lồi). Hàm nhận giá trị thực Φ xác định trên một khoảng mở (a, b) (hữu hạn hoặc vô hạn) được gọi là lồi nếu

$$\forall x, y \in (a, b) : a < x < y < b, \quad \forall \gamma : 0 < \gamma < 1$$

thì

$$\Phi(\gamma x + (1 - \gamma)y) \leq \gamma \Phi(x) + (1 - \gamma)\Phi(y).$$

Nếu dấu “ $<$ ” xảy ra thì Φ được gọi là lồi chặt (strictly convex) hay lồi thực sự.

Chúng ta chỉ xét những hàm tổn thất lồi theo đối số thứ nhất (tức là lồi theo T)

Định lý 1.2.1 (Giá trị cực tiểu của hàm lồi). Giả sử ρ là hàm lồi xác định trên $(-\infty, +\infty)$, X là biến ngẫu nhiên sao cho

$$\Phi(a) = \mathbb{E}[\rho(X - a)] \text{ hữu hạn.}$$

Nếu ρ là hàm không đơn điệu thì hàm $\Phi(a)$ sẽ nhận giá trị nhỏ nhất, và $\{a : \Phi(a) \text{ min}\}$ là khoảng đóng.

Nếu ρ lồi chặt thì giá trị nhỏ nhất là duy nhất.

Chứng minh. Xem trong [14, tr. 50].

Ví dụ 1.2.1.

1. Nếu hàm tổn thất có dạng bình phương thì

$$\mathbb{E}(X - a)^2 \text{ min} \Leftrightarrow a = \mu = \mathbb{E}X.$$

2. Nếu hàm tổn thất có dạng trị tuyệt đối thì $\mathbb{E}|X - a|$ đạt min trên khoảng đóng tại các median của X .

1.3 Thống kê đủ và đầy đủ

Định nghĩa 1.3.1. Thống kê $T(X)$ được gọi là thống kê đủ đối với họ độ đo xác suất $\{\mathbb{P}_\theta\}$ hay thống kê đủ đối với tham ẩn θ nếu $\mathbb{P}_\theta(X \in A | T(X) = t)$ không phụ thuộc $\theta, \forall t, \forall \theta \in \Theta$.

Định lý 1.3.1 (Tiêu chuẩn tách Neyman). Điều kiện cần và đủ để thống kê $T : (\mathcal{X}, \mathcal{A}, \mathbb{P}_\theta) \longrightarrow (\mathcal{T}, \mathcal{B}, \mathbb{P}_\theta^T)$ là thống kê đủ đối với $\theta \in \Theta$ là tồn tại $g_\theta(t) \geq 0$ $\mathcal{B}$ - đo được và $h(x) \geq 0$ $\mathcal{A}$ - đo được sao cho mật độ của biến ngẫu nhiên X có dạng

$$p_\theta(x) = \frac{d\mathbb{P}_\theta(x)}{d\mu(x)} = g_\theta(T(X)) \cdot h(x).$$

Định nghĩa 1.3.2. Xét mô hình thống kê $(\mathcal{X}, \mathcal{A}, P_\theta)$, $\theta \in \Theta$. Họ $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$ được gọi là họ đầy đủ nếu với mọi hàm $\varphi(x)$ $\mathcal{A}$ - đo được, từ hệ thức

$$\mathbb{E}_\theta \varphi(x) = 0, \quad \forall \theta \in \Theta$$

suy ra $\varphi(x) = 0$ ($\mathcal{P}$ - h.k.n).

Định lý 1.3.2 (Cải tiến ước lượng dựa vào thống kê đủ). Xét mô hình thống kê $(\mathcal{X}, \mathcal{A}, \mathbb{P}_\theta)$, $X = (X_1, \dots, X_n)$. Giả sử $T(X)$ là thống kê đủ đối với họ $\{\mathbb{P}_\theta\}$ và $\varphi(x)$ là một ước lượng không chệch nào đó của $g(\theta)$. Khi đó

$$\hat{g}(T) = \mathbb{E}_\theta(\varphi(X) | T(X))$$

cũng sẽ là một ước lượng không chệch của $g(\theta)$ và hàm tổn thất tương ứng của nó thỏa mãn

$$\mathbb{E}_\theta r(\hat{g}, g) \leq \mathbb{E}_\theta r(\varphi, g), \quad \forall \theta \in \Theta.$$

Dấu bằng xảy ra khi và chỉ khi $\varphi(X)$ là đo được đối với σ - đại số sinh bởi thống kê $T(X)$.

Chứng minh. Xem [2, Chương 3, tr. 37-38].

Định lý 1.3.3 (Blackwell - Rao - Lehmann - Fisher). Vẫn giả thiết như Định lý 1.3.2 và thêm điều kiện thống kê T đầy đủ thì $\hat{g}(T)$ sẽ là ước lượng không chệch của $g(\theta)$ và hàm tổn thất dạng bình phương là bé nhất đồng thời $\hat{g}(T)$ là ước lượng duy nhất có tính chất đó.

Chứng minh. Xem [2, Chương 3, tr. 39].

1.4 Ước lượng không chệch phương sai bé đều nhất (UMVU)

Định nghĩa 1.4.1. Xét mô hình thống kê $(\mathcal{X}, \mathcal{A}, \mathbb{P}_\theta)$. Thống kê $\delta(X)$ được gọi là ước lượng không chệch cho $g(\theta)$ nếu

$$\mathbb{E}_\theta \delta(X) = g(\theta), \quad \forall \theta \in \Theta.$$

Nếu $g \in \mathbb{R}^d$ thì điều kiện trên được thay bằng

$$\mathbb{E}_\theta \delta_i(X) = g_i(\theta), \quad \forall \theta \in \Theta, \quad i = 1, \dots, d.$$

Định nghĩa 1.4.2. Ước lượng không chệch $\delta(X)$ được gọi là ước lượng không chệch phương sai bé đều nhất (UMVU estimators) nếu với mọi ước lượng không chệch $\delta'(X)$ khác ta có

$$\text{Var}_\theta \delta(X) \leq \text{Var}_\theta \delta'(X), \quad \forall \theta \in \Theta.$$

1.5 Ước lượng chấp nhận được và ước lượng tối ưu

Định nghĩa 1.5.1. Giả sử ước lượng $T'(x)$ nằm trong lớp K các ước lượng của hàm tham số $t(\theta)$. $T'(x)$ được gọi là chấp nhận được trong K với hàm tổn thất $r(T, t)$ nếu không tồn tại ước lượng T thuộc K sao cho

$$R(T, \theta) \leq R(T', \theta), \quad \forall \theta \in \Theta \quad (1.5.1)$$

và dấu “<” xảy ra với ít nhất một giá trị của θ .

Nếu ước lượng T' là chấp nhận được trong lớp tất cả các ước lượng thì T' được gọi là chấp nhận được tuyệt đối.

Định nghĩa 1.5.2. Ước lượng $T^0 \in K$ là tối ưu trong K (với K là lớp bất kỳ các ước lượng của hàm tham số $t(\theta)$) với hàm tổn thất $r(T, t)$ nếu với mọi $T \in K$ ta có

$$R(T^0, \theta) \leq R(T, \theta), \quad \forall \theta \in \Theta. \quad (1.5.2)$$

Định nghĩa 1.5.3. Ước lượng $T_0 \in K$ được gọi là tối ưu tại điểm θ_0 (hay còn gọi là tối ưu địa phương) trong K với hàm tổn thất $r(T, t)$ nếu với mọi $T \in K$ ta có

$$R(T_0, \theta_0) \leq R(T, \theta_0). \quad (1.5.3)$$

1.6 Một số kiến thức khác

Định nghĩa 1.6.1 (Điều kiện chính quy). Giả sử họ phân phối xác suất $\mathbb{P}_\theta$, $\theta \in \Theta$ liên tục tuyệt đối với độ đo μ , có đạo hàm Radon-Nikodym đối với μ là $f(x, \theta) = \frac{d\mathbb{P}_\theta}{d\mu}(x)$. Họ các hàm mật độ $\{f(x, \theta), \theta \in \Theta\}$ được gọi là chính quy nếu

- Θ hoặc là đường thẳng thực hoặc là một khoảng của đường thẳng thực và $\{x : f(x, \theta) > 0\}$ không phụ thuộc vào θ .
- $\frac{\partial f(x, \theta)}{\partial \theta}$ tồn tại và hữu hạn $\mathbb{P}_\theta$ - hầu chắc chắn với mọi $\theta \in \Theta$.
- $\int \left| \frac{\partial f(x, \theta)}{\partial \theta} \right| \mu(dx) < \infty$ với mọi $\theta \in \Theta$.
- $0 < \mathbb{E}_\theta \left(\frac{\partial \ln f(x, \theta)}{\partial \theta} \right)^2 < \infty$ với mọi $\theta \in \Theta$.

Định lý 1.6.1 (Bất đẳng thức Cramer-Rao). Giả sử không gian tham số Θ là một khoảng nào đó của đường thẳng thực và tất cả các phân phối $\mathbb{P}_\theta$ là liên tục tuyệt đối theo độ đo μ với hàm mật độ $p(x, \theta) = d\mathbb{P}_\theta/d\mu$ thỏa mãn điều kiện chính quy. Nếu T là một thống kê sao cho $\mathbb{E}_\theta T = t(\theta) + b(\theta)$, T có phương sai hữu hạn và $t(\theta), b(\theta)$ khả vi thì

$$\mathbb{E}_\theta [T - t(\theta)]^2 \geq [b(\theta)]^2 + \frac{[t'(\theta) + b'(\theta)]^2}{I(\theta)}, \quad (1.6.1)$$

trong đó $I(\theta) = \mathbb{E}_\theta \left[\frac{\partial \log p}{\partial \theta} \right]^2$ là lượng thông tin Fisher.

Chứng minh. Xem [18, tr. 323-325].

Định lý 1.6.2. *Giả sử thống kê T là một ước lượng cho $t(\theta)$ và hàm tổn thất $r(T, t)$ là hàm lồi theo T . Nếu họ phân bố $\{\mathbb{P}_\theta\}$ có tồn tại thống kê đủ S thì sẽ tồn tại ước lượng $\hat{T}$ chỉ phụ thuộc vào S thỏa mãn*

$$\mathbb{E}_\theta r(\hat{T}, t(\theta)) \leq \mathbb{E}_\theta r(T, t(\theta)), \quad \forall \theta.$$

Chứng minh. Xem [4], [12], [15] và [18].

Từ Định lý 1.6.2 ta suy ra rằng những ước lượng mà không phải là hàm của thống kê đủ thì sẽ là không chấp nhận được đối với hàm tổn thất lồi theo đối số thứ nhất. Phương pháp tìm $\hat{T}$ từ T được gọi là cải tiến ước lượng dựa vào thống kê đủ (Định lý “Blackwell - Rao”).

Nếu hàm tổn thất có dạng bình phương thì chúng ta hướng sự quan tâm đặc biệt đến các ước lượng không chệch của hàm tham ẩn.

Có thể chỉ ra rằng trong Định lý 1.6.2 nếu T là một ước lượng không chệch của $t(\theta)$ thì $\hat{T}$ cũng là ước lượng không chệch của $t(\theta)$.

Đối với các họ đầy đủ, mọi hàm tham ẩn $t(\theta)$ rõ ràng là có tối đa một ước lượng không chệch phụ thuộc vào thống kê đủ (nếu các ước lượng trùng nhau hầu khắp nơi đối với từng độ đo $\mathbb{P}_\theta$ thì ta coi là như nhau). Thật vậy, giả sử $\psi(S)$ và $\phi(S)$ là hai ước lượng không chệch phụ thuộc vào thống kê đủ S cho $t(\theta)$. Khi đó rõ ràng $\mathbb{E}_\theta \psi(S) = \mathbb{E}_\theta \phi(S) = t(\theta), \forall \theta$, suy ra $\mathbb{E}_\theta [\psi(S) - \phi(S)] = 0, \forall \theta$. Theo tính chất của họ đầy đủ, ta suy ra ngay $\psi(S) - \phi(S) = 0, \mathbb{P}_\theta$ hầu khắp nơi, $\forall \theta$. Từ đó $\psi(S) = \phi(S), \mathbb{P}_\theta$ hầu khắp nơi, $\forall \theta$.

Từ Định lý 1.6.2 ta suy ra rằng đối với các họ đầy đủ, mọi thống kê $\hat{T} = \hat{T}(S)$ là tối ưu trong lớp các ước lượng không chệch của hàm tham ẩn $t(\theta) = \mathbb{E}_\theta(\hat{T})$ (xem [16]). Tính chất này sẽ được sử dụng thường xuyên trong các nội dung về sau.

Để nhận biết tính đầy đủ của họ $\{\mathbb{P}_\theta\}$, ta có tiêu chuẩn sau đây:

Định lý 1.6.3 (Lehman). *Giả sử phân phối của vector thống kê đủ $S =$*

$(S_1, \dots, S_m) \in \mathbb{R}^m$ có hàm mật độ (theo độ đo Lebesgue) dạng

$$\pi(s, \theta) = c(\theta) q(s) \exp \left(\sum_1^m c_i(\theta) s_i \right),$$

trong đó $s = (s_1, \dots, s_m)$ và vector $(c_1(\theta), \dots, c_m(\theta))$ phủ một miền nào đó trong $\mathbb{R}^m$ khi $\theta \in \Theta$. Khi đó họ phân bố $\{\mathbb{P}_\theta\}$ là đầy đủ đối với thống kê đủ S .

Chứng minh. Xem trong [2, Chương 3] và [13].

Chương 2

Tính tối ưu trong lớp các ước lượng không chệch và lớp các ước lượng đều

2.1 Tính chất tối ưu trong lớp các ước lượng không chệch

Trong mục này ta nghiên cứu tính chất tối ưu trong trường hợp hàm tổn thất có dạng bình phương.

Ký hiệu $\mathcal{L}_\theta^2$ là không gian Hilbert các hàm đo được trên $(\mathcal{X}, \mathcal{A})$ bình phương khả tích theo độ đo $\mathbb{P}_\theta$ đối với tích vô hướng thông thường. Đặt $\mathcal{L}^2 = \bigcap_{\theta \in \Theta} \mathcal{L}_\theta^2$.

Định nghĩa 2.1.1. Thống kê $h = h(x)$ được gọi là ước lượng không chệch của không nếu $\mathbb{E}_\theta h = 0$ với mọi $\theta \in \Theta$.

Ký hiệu $\mathcal{H}_\theta$ là tập hợp tất cả các ước lượng không chệch của không trong $\mathcal{L}_\theta^2$ và đặt $\mathcal{H} = \bigcap_{\theta \in \Theta} \mathcal{H}_\theta$.

Xét Bổ đề sau:

Bổ đề 2.1.1. Giả sử $T \in \mathcal{L}^2$. T là một ước lượng không chệch tối ưu của $t(\theta) = \mathbb{E}_\theta T$ khi và chỉ khi

$$\mathbb{E}_\theta (Th) \equiv 0, \quad \forall h \in \mathcal{H}, \quad \forall \theta \in \Theta \quad (2.1.1)$$

Tương tự: Giả sử $T \in \mathcal{L}_\theta^2$. T là một ước lượng không chệch tối ưu địa

phương tại điểm θ của $t(\theta) = \mathbb{E}_\theta T$ khi và chỉ khi

$$\mathbb{E}_\theta (Th) \equiv 0, \quad \text{với } \theta \text{ đã cho và } \forall h \in \mathcal{H}_\theta. \quad (2.1.2)$$

Bổ đề trên đây bắt nguồn từ một nghiên cứu của Fisher, mặc dù nó có vẻ đơn giản nhưng lại là một công cụ hết sức hữu dụng trong việc nghiên cứu các bài toán về đặc trưng.

Chứng minh.

Điều kiện cần: Giả sử $T \in \mathcal{L}_\theta^2$ là ước lượng không chệch của $t(\theta)$ tối ưu địa phương tại θ và $h \in \mathcal{H}_\theta$. Với λ là hằng số tùy ý, $T + \lambda h \in \mathcal{L}_\theta^2$ và là một ước lượng không chệch cho $t(\theta)$. Thật vậy,

$$\mathbb{E}_\theta (T + \lambda h) = \mathbb{E}_\theta T + \lambda \mathbb{E}_\theta h = t(\theta)$$

vì $\mathbb{E}_\theta h = 0$ (do h là ước lượng không chệch của không) và $t(\theta) = \mathbb{E}_\theta T$ theo giả thiết. Ngoài ra, $T \in \mathcal{L}_\theta^2$, $h \in \mathcal{H}_\theta \subset L_\theta^2$ nên $T + \lambda h \in \mathcal{L}_\theta^2$.

Xét

$$\begin{aligned} \mathbb{E}_\theta (T + \lambda h - t)^2 &= \mathbb{E}_\theta (T - t)^2 + \mathbb{E}_\theta \lambda^2 h^2 + 2\mathbb{E}_\theta \lambda h (T - t) \\ &= \mathbb{E}_\theta (T - t)^2 + \lambda^2 \mathbb{E}_\theta h^2 + 2\lambda \mathbb{E}_\theta (Th) - 2\lambda \mathbb{E}_\theta (th) \\ &= \mathbb{E}_\theta (T - t)^2 + \lambda^2 \mathbb{E}_\theta h^2 + 2\lambda \mathbb{E}_\theta (Th) - 2\lambda t(\mathbb{E}_\theta h) \\ &= \mathbb{E}_\theta (T - t)^2 + \lambda^2 \mathbb{E}_\theta h^2 + 2\lambda \mathbb{E}_\theta (Th). \end{aligned}$$

Có các biến đổi phía trên là do $t(\theta)$ là hàm tất định đối với mỗi giá trị của θ , không phải biến ngẫu nhiên nên $2\lambda \mathbb{E}_\theta h t = 2\lambda t(\mathbb{E}_\theta h)$. Ngoài ra vì h là ước lượng không chệch của không nên $\mathbb{E}_\theta h = 0$.

Nếu $\mathbb{E}_\theta (Th) \neq 0$ thì ta có thể chọn giá trị thích hợp của λ để cho $f(\lambda) = \lambda^2 \mathbb{E}_\theta h^2 + 2\lambda \mathbb{E}_\theta (Th) < 0$. Thật vậy, xét hai trường hợp:

Trường hợp 1. $\mathbb{E}_\theta (Th) > 0$

Nếu $\mathbb{E}_\theta h^2 = 0$ thì ta chọn $\lambda < 0$, khi đó $f(\lambda) = 2\lambda \mathbb{E}_\theta (Th) < 0$.

Nếu $\mathbb{E}_\theta h^2 > 0$ thì $f(\lambda)$ là một tam thức bậc hai đối với λ có các nghiệm là 0 và $\frac{-2\mathbb{E}_\theta (Th)}{\mathbb{E}_\theta h^2}$, $f(\lambda) < 0$ nếu $\frac{-2\mathbb{E}_\theta (Th)}{\mathbb{E}_\theta h^2} < \lambda < 0$.

Trường hợp 2. $\mathbb{E}_\theta (Th) < 0$

Nếu $\mathbb{E}_\theta h^2 = 0$ thì ta chọn $\lambda > 0$, khi đó $f(\lambda) = 2\lambda\mathbb{E}_\theta (Th) < 0$.

Nếu $\mathbb{E}_\theta h^2 > 0$ thì $f(\lambda) < 0$ nếu $0 < \lambda < \frac{-2\mathbb{E}_\theta (Th)}{\mathbb{E}_\theta h^2}$.

Tóm lại, ta luôn chọn được λ để cho $f(\lambda) < 0$. Khi đó rõ ràng từ

$$\mathbb{E}_\theta (T + \lambda h - t)^2 = \mathbb{E}_\theta (T - t)^2 + \lambda^2 \mathbb{E}_\theta h^2 + 2\lambda \mathbb{E}_\theta (Th)$$

ta suy ra $\mathbb{E}_\theta (T + \lambda h - t)^2 < \mathbb{E}_\theta (T - t)^2$ với λ thích hợp. Điều này mâu thuẫn với giả thiết tối ưu của T tại θ . Vậy $\mathbb{E}_\theta (Th) = 0, \forall h \in \mathcal{H}_\theta$.

Điều kiện đủ: Giả sử $\mathbb{E}_\theta (Th) = 0, \forall h \in \mathcal{H}_\theta$ và $T' \in \mathcal{L}_\theta^2$ là một ước lượng không chệch bất kỳ khác của $t(\theta)$. Khi đó $\mathbb{E}_\theta T = \mathbb{E}_\theta T' = t(\theta)$, suy ra $\mathbb{E}_\theta (T - T') = 0$. Đặt $T' - T = h$ thì h là một ước lượng không chệch của không nên $h \in \mathcal{H}_\theta$. Hơn nữa,

$$\begin{aligned} \mathbb{E}_\theta (T' - t)^2 &= \mathbb{E}_\theta (T' - T + T - t)^2 = \mathbb{E}_\theta (h + T - t)^2 \\ &= \mathbb{E}_\theta (T - t)^2 + 2\mathbb{E}_\theta (T - t)h + \mathbb{E}_\theta h^2 \\ &= \mathbb{E}_\theta (T - t)^2 + 2\mathbb{E}_\theta (Th) - 2\mathbb{E}_\theta (th) + \mathbb{E}_\theta h^2 \\ &= \mathbb{E}_\theta (T - t)^2 + 2\mathbb{E}_\theta (Th) - 2t(\mathbb{E}_\theta h) + \mathbb{E}_\theta h^2 \\ &= \mathbb{E}_\theta (T - t)^2 + 2\mathbb{E}_\theta (Th) + \mathbb{E}_\theta h^2 \\ &\geq \mathbb{E}_\theta (T - t)^2. \end{aligned}$$

Bất đẳng thức cuối cùng suy ra từ điều kiện $\mathbb{E}_\theta (Th) = 0$. Vậy T tối ưu tại θ . $\square$

Khẳng định của Bổ đề 2.1.1 trong trường hợp $T \in \mathcal{L}^2$ được chứng minh hoàn toàn tương tự.

Bổ đề nói trên chỉ tỏ ra hữu ích trong những trường hợp mà họ phân bố đang khảo sát có khá nhiều sự hỗ trợ từ phía các ước lượng không chệch trong một lớp thích hợp. Đặc biệt là những phân bố được sinh bởi các mẫu ngẫu nhiên rút ra từ tổng thể với các tham số vị trí và tỉ lệ. Điều kiện (2.1.2) là điều kiện cần không chỉ đối với tính tối ưu mà còn là điều kiện cần đối với tính chấp nhận được của ước lượng T trong những trường hợp mà $\mathbb{E}_\theta (T - t)^2$

là hằng số và nếu ta sử dụng ước lượng không chệch của không h thì $\mathbb{E}_\theta(Th)$ và $\mathbb{E}_\theta h^2$ là các hằng số.

Nếu ta sử dụng hàm tổn thất không có dạng bình phương thì ta không có được khẳng định như Bổ đề 2.1.1 (và thậm chí ta không thể biết được lớp các ước lượng nào có thể đặt ra điều kiện cho tính chất tối ưu).

Bây giờ ta áp dụng Bổ đề 2.1.1 để chứng minh tính không chệch tối ưu được bảo toàn đối với các phép biến đổi hàm tương đối tổng quát. Giả sử rằng đối với họ phân bố $\{\mathbb{P}_\theta\}$ đang xét, tập hợp $\mathcal{H}'$ các ước lượng không chệch của không h , thỏa mãn $\mathbb{E}_\theta h^4 < \infty$ với mọi $\theta \in \Theta$, trù mật trong $\mathcal{H}$ theo chuẩn trong $\mathcal{L}_\theta^2$. Ta gọi điều kiện này là điều kiện (C).

Nếu

$$\mathbb{E}_\theta(|T|^k) < \infty, \quad k = 1, 2, \dots \quad (2.1.3)$$

thì ta sẽ ký hiệu $\mathcal{L}_\theta^2(T)$ là bao đóng theo chuẩn trong $\mathcal{L}_\theta^2$ của tập tất cả các đa thức theo T , và ký hiệu $\mathcal{L}^2(T) = \bigcap_{\theta \in \Theta} \mathcal{L}_\theta^2$.

Định lý 2.1.1 (xem [17]). *Giả sử họ phân bố $\{\mathbb{P}_\theta\}$ thỏa mãn điều kiện (C). Nếu $T \in \mathcal{L}^2$ là ước lượng không chệch tối ưu trong $\mathcal{L}^2$ của $t(\theta) = \mathbb{E}_\theta(T)$, thỏa mãn (2.1.3) với mọi θ thì mọi ước lượng $\beta \in \mathcal{L}^2(T)$ cũng sẽ là ước lượng không chệch tối ưu trong $\mathcal{L}^2$ của kỳ vọng toán của nó $b(\theta) = \mathbb{E}_\theta \beta$.*

Chứng minh. Nếu $T \in \mathcal{L}^2$ là ước lượng không chệch tối ưu của $t(\theta) = \mathbb{E}_\theta(T)$ thì theo Bổ đề 2.1.1, với mọi h' thuộc $\mathcal{H}'$ và mọi θ ta phải có

$$\mathbb{E}_\theta(Th') = 0. \quad (2.1.4)$$

Theo bất đẳng thức Cauchy-Bunyakovski:

$$\mathbb{E}_\theta(T^2 h'^2) \leq \mathbb{E}_\theta(T^4)^{1/2} \cdot \mathbb{E}_\theta(h'^4)^{1/2} < \infty$$

bởi vì $\mathbb{E}_\theta h'^4 < \infty$ và $\mathbb{E}_\theta(T^4) < \infty$. Vậy $Th' \in \mathcal{L}^2$. Nhưng vì Th' là ước lượng không chệch của không (theo (2.1.4)) nên $Th' \in \mathcal{H}, \forall h' \in \mathcal{H}'$. Từ đó suy ra $\mathbb{E}_\theta(T^2 h') = \mathbb{E}_\theta(T.Th') = 0, \forall \theta$. Mặt khác theo giả thiết $\mathcal{H}'$ trù mật trong $\mathcal{H}$ nên $\forall h \in \mathcal{H}$, tồn tại dãy $h'_n \in \mathcal{H}'$ sao cho $h'_n \rightarrow h$ theo chuẩn trong

$\mathcal{L}_\theta^2$ khi $n \rightarrow \infty$, tức là

$$\mathbb{E}_\theta (h'_n - h)^2 \rightarrow 0 \text{ khi } n \rightarrow \infty.$$

Ta có

$$\begin{aligned} \left| \mathbb{E}_\theta T^2 h - \mathbb{E}_\theta T^2 h'_n \right| &= \left| \mathbb{E}_\theta T^2 (h - h'_n) \right| \\ &\leq \left(\mathbb{E}_\theta T^4 \right)^{1/2} \cdot \left(\mathbb{E}_\theta (h - h'_n)^2 \right)^{1/2} \rightarrow 0 \text{ khi } n \rightarrow \infty, \end{aligned}$$

vì $\mathbb{E}_\theta (T^4) < \infty$ và $\mathbb{E}_\theta (h'_n - h)^2 \rightarrow 0$ khi $n \rightarrow \infty$. Nhưng $\mathbb{E}_\theta (T^2 h'_n) = 0$ do $h'_n \in \mathcal{H}'$ và chứng minh ở trên, suy ra $\mathbb{E}_\theta (T^2 h) = 0$, với mọi $h \in \mathcal{H}$. Tương tự ta có thể chứng minh được rằng

$$\mathbb{E}_\theta (T^k h) = 0, \forall \theta, \forall h \in \mathcal{H}, \quad k = 0, 1, \dots$$

Xét $f(T)$ là một đa thức theo biến T , khi đó $f(T)$ có dạng

$$f(T) = \sum_i a_i T^i.$$

Ta có $\mathbb{E}_\theta (f(T) h) = \sum_i a_i \mathbb{E}_\theta (T^i h) = 0$. Vì tập hợp các đa thức theo biến T là trù mật trong $\mathcal{L}_\theta^2(T)$ nên $\forall \beta \in \mathcal{L}^2(T)$, tồn tại dãy đa thức $f_n(T)$ sao cho $f_n(T) \rightarrow \beta$ theo chuẩn trong $\mathcal{L}_\theta^2$, tức là

$$\mathbb{E}_\theta (f_n(T) - \beta)^2 \rightarrow 0 \text{ khi } n \rightarrow \infty.$$

Lại có

$$\begin{aligned} \left| \mathbb{E}_\theta (f_n(T) h) - \mathbb{E}_\theta (\beta h) \right| &= \left| \mathbb{E}_\theta h (f_n(T) - \beta) \right| \\ &\leq \left(\mathbb{E}_\theta h^2 \right)^{1/2} \cdot \left(\mathbb{E}_\theta (f_n(T) - \beta)^2 \right)^{1/2} \rightarrow 0 \end{aligned}$$

vì $\mathbb{E}_\theta h^2 < \infty$ ($h \in \mathcal{H} \subset \mathcal{L}^2$) và $\mathbb{E}_\theta (f_n(T) - \beta)^2 \rightarrow 0$. Theo chứng minh ở trên, $\mathbb{E}_\theta (f_n(T) h) = 0, \forall h \in \mathcal{H}$, suy ra $\mathbb{E}_\theta (\beta h) = 0, \forall h \in \mathcal{H}, \forall \theta \in \Theta$.

Theo Bổ đề 2.1.1, β là ước lượng không chệch tối ưu trong $\mathcal{L}^2$ của $b(\theta) = \mathbb{E}_\theta \beta$. $\square$

Ta nhận thấy rằng ước lượng không chệch tối ưu sẽ là hiệu quả theo nghĩa xác định bởi bất đẳng thức Cramer-Rao (vế phải của (1.6.1) bằng phương sai của ước lượng này). Tính chất này sẽ không được bảo toàn với các phép biến đổi hàm. Tuy nhiên với các điều kiện của Định lý 2.1.1, các hàm của ước lượng không chệch hiệu quả cũng sẽ là ước lượng không chệch tối ưu của kỳ vọng toán học của nó.

Nhận xét 2.1.1. Điều kiện (C) có lẽ không cần thiết đối với Định lý 2.1.1. Trong mọi trường hợp nó có thể được lược bỏ nếu tồn tại một hằng số C thỏa mãn $|T(x)| < C$ với xác suất bằng 1 theo $\mathbb{P}_\theta, \forall \theta$.

Nhận xét 2.1.2. Tính chất của ước lượng $T \in \mathcal{L}_\theta^2$ là không chệch tối ưu địa phương của $t(\theta)$, nói chung không bảo toàn đối với các phép biến đổi hàm.

2.2 Tham số tịnh tiến và lớp các ước lượng đều

Giả sử rằng không gian các quan sát $\mathcal{X}$ là $\mathbb{R}^n$ và không gian tham số Θ là $\mathbb{R}^1$. Nếu phân phối $\mathbb{P}_\theta$ phụ thuộc vào tham số θ theo nghĩa dưới đây

$$\mathbb{P}_\theta(A) = \int \cdots \int_A dF(x_1 - \theta, \dots, x_n - \theta) \quad (2.2.1)$$

thì ta nói rằng θ là tham số tịnh tiến. Phân phối có dạng (2.2.1) nảy sinh từ việc xem xét các lược đồ biến đổi tuyến tính, khi mà các quan sát có dạng

$$x_i = \theta + \varepsilon_i, \quad i = 1, \dots, n$$

trong đó sai số ε_i (nói chung phụ thuộc) có phân bố đồng thời $F(x_1, \dots, x_n)$.

Để ước lượng cho tham số θ lẽ tự nhiên ta sẽ chọn ra một lớp các ước lượng $\tilde{\theta} = \tilde{\theta}(X_1, \dots, X_n)$ thỏa mãn điều kiện sau đây với mọi số thực c :

$$\tilde{\theta}(x_1 + c, \dots, x_n + c) = \tilde{\theta}(x_1, \dots, x_n) + c. \quad (2.2.2)$$

Ước lượng nói trên được đưa ra đầu tiên bởi Pitman (xem [8]). Chúng ta sẽ gọi những ước lượng này là ước lượng đều (regular estimators) và ký hiệu lớp các ước lượng đều bởi $\mathcal{J}$.

Nếu hàm tổn thất chỉ phụ thuộc vào hiệu các đối số, tức là $r(\tilde{\theta}, \theta) = r(\tilde{\theta} - \theta)$ thì ta dễ dàng chỉ ra rằng với $\tilde{\theta} \in \mathcal{T}$, hàm hiểm

$$R(\tilde{\theta}, \theta) = \mathbb{E}_{\theta} r(\tilde{\theta} - \theta) = \text{hằng số (không phụ thuộc vào } \theta \in \mathbb{R}^1).$$

Thật vậy, theo (2.2.2) ta có

$$\tilde{\theta}(x_1, \dots, x_n) - \theta = \tilde{\theta}(x_1 - \theta, \dots, x_n - \theta).$$

Đặt $y_i = x_i - \theta$. Khi đó

$$\begin{aligned} & \mathbb{E}_{\theta} r[\tilde{\theta}(x_1, \dots, x_n) - \theta] \\ &= \int \cdots \int_{\mathbb{R}^n} r[\tilde{\theta}(x_1, \dots, x_n) - \theta] dF(x_1 - \theta, \dots, x_n - \theta) \\ &= \int \cdots \int_{\mathbb{R}^n} r[\tilde{\theta}(x_1 - \theta, \dots, x_n - \theta)] dF(x_1 - \theta, \dots, x_n - \theta) \\ &= \int \cdots \int_{\mathbb{R}^n} r[\tilde{\theta}(y_1, \dots, y_n)] dF(y_1, \dots, y_n) \\ &= \mathbb{E}_0 r(\tilde{\theta}). \end{aligned}$$

Từ kết quả trên đây ta có thể chỉ ra được rằng, đối với các hàm tổn thất có dạng hiệu, một ước lượng $\tilde{\theta} \in \mathcal{T}$ hoặc là tối ưu trong $\mathcal{T}$ hoặc không chấp nhận được trong $\mathcal{T}$. Thật vậy, giả sử $T \in \mathcal{T}$ là một ước lượng khác cho θ . Khi đó theo kết quả trên

$$\mathbb{E}_{\theta} r(\tilde{\theta} - \theta) = c_{\tilde{\theta}}, \forall \theta$$

$$\mathbb{E}_{\theta} r(T - \theta) = c_T, \forall \theta$$

Nếu $c_{\tilde{\theta}} \leq c_T, \forall T$ thì $\mathbb{E}_{\theta} r(\tilde{\theta} - \theta) \leq \mathbb{E}_{\theta} r(T - \theta), \forall T, \forall \theta$. Vậy $\tilde{\theta}$ là ước lượng tối ưu.

Nếu tồn tại T_0 sao cho $c_{T_0} < c_{\tilde{\theta}}$ thì $\mathbb{E}_{\theta} r(T_0 - \theta) < \mathbb{E}_{\theta} r(\tilde{\theta} - \theta), \forall \theta$. Vậy $\tilde{\theta}$ không chấp nhận được.

Định nghĩa 2.2.1. *Ước lượng $\hat{\theta}$ tối ưu trong lớp $\mathcal{T}$ theo nghĩa hàm tổn thất $R(\hat{\theta}, \theta) = \min_{\tilde{\theta} \in \mathcal{T}} R(\tilde{\theta}, \theta)$ được gọi là ước lượng Pitman của tham số tịnh tiến θ tương ứng với hàm tổn thất $r(\tilde{\theta} - \theta)$.*

Nếu $\min_{\tilde{\theta} \in \mathcal{T}} R(\tilde{\theta}, \theta)$ đạt được duy nhất tại thống kê $\hat{\theta}$ (Chúng ta đồng nhất các thống kê trùng nhau hầu khắp nơi theo từng độ đo $\mathbb{P}_\theta$) thì rõ ràng điều kiện cần để ước lượng $\tilde{\theta} \in \mathcal{T}$ chấp nhận được là

$$\tilde{\theta} = \hat{\theta} \text{ hầu khắp nơi theo } \mathbb{P}_\theta, \quad \forall \theta \in \mathbb{R}^1, \quad (2.2.3)$$

hay tương đương với

$$\tilde{\theta} = \hat{\theta} \text{ hầu khắp nơi theo } \mathbb{P}_0. \quad (2.2.4)$$

Thật vậy, giả sử $\tilde{\theta}$ là chấp nhận được trong $\mathcal{T}$. Vì đối với hàm tổn thất dạng hiệu, như đã chứng minh ở trên, mọi ước lượng trong $\mathcal{T}$ hoặc không chấp nhận được, hoặc tối ưu. Vì thế $\tilde{\theta}$ là một ước lượng tối ưu trong $\mathcal{T}$. Vì $\min_{\tilde{\theta} \in \mathcal{T}} R(\tilde{\theta}, \theta)$ đạt được duy nhất tại thống kê $\hat{\theta}$ nên

$$\tilde{\theta} = \hat{\theta} \text{ hầu khắp nơi theo } \mathbb{P}_\theta, \quad \forall \theta \in \mathbb{R}^1.$$

Khi đó ta có

$$\mathbb{E}_\theta r(\tilde{\theta} - \theta) = \mathbb{E}_\theta r(\hat{\theta} - \theta).$$

Đẳng thức trên tương đương với

$$\mathbb{E}_0 r(\tilde{\theta}) = \mathbb{E}_0 r(\hat{\theta}).$$

Vậy $\tilde{\theta} = \hat{\theta}$ hầu khắp nơi theo $\mathbb{P}_0$.

Đối với một vài ước lượng và hàm tổn thất đặc biệt, điều kiện (2.2.4) khá phù hợp để đặc trưng cho hàm phân phối $F(x)$ thông qua tính chấp nhận được của $\tilde{\theta}$.

Giả sử rằng

$$\int_{\mathbb{R}^n} x_j^2 dF(x_1, \dots, x_n) < \infty, \quad j = 1, \dots, n \quad (2.2.5)$$

và cho $L = \sum_1^n c_j x_j$ với $\sum_1^n c_j = 1$ là một thống kê tuyến tính nào đó.

Bổ đề 2.2.1.

(i) Dưới điều kiện (2.2.5), ước lượng cho bởi

$$\begin{aligned}\hat{\theta} &= L - \mathbb{E}_0 \left(L | X_2 - X_1, \dots, X_n - X_1 \right) \\ &= L - \mathbb{E}_0 \left(L | Y \right), Y = (X_2 - X_1, \dots, X_n - X_1)\end{aligned}\quad (2.2.6)$$

là ước lượng Pitman ứng với hàm tổn thất dạng bình phương, đồng thời $\mathbb{E}_\theta \left(\tilde{\theta} - \theta \right)^2 > \mathbb{E}_\theta \left(\hat{\theta} - \theta \right)^2$ với $\tilde{\theta} \in \mathcal{T}$, $\theta \in \mathbb{R}^1$ trừ trường hợp $\tilde{\theta} = \hat{\theta}$ hầu khắp nơi đối với $\mathbb{P}_0$ (hay $\tilde{\theta} = \hat{\theta}$ hầu khắp nơi đối với $\mathbb{P}_\theta, \forall \theta$).

(ii) Nếu $F(x)$ là liên tục tuyệt đối với $f(x)$ là hàm mật độ (theo độ đo Lebesgue), thì

$$\hat{\theta} = \frac{\int_{-\infty}^{+\infty} u f(x_1 - u, \dots, x_n - u) du}{\int_{-\infty}^{+\infty} f(x_1 - u, \dots, x_n - u) du}.$$

Chứng minh.

(i) Lấy $\tilde{\theta}(x_1, \dots, x_n) \in \mathcal{T}$. Đặt

$$\begin{aligned}y_1 &= x_1 \\ y_n &= x_n - x_1, \quad n \geq 2.\end{aligned}$$

Khi đó theo (2.2.2) ta có

$$\begin{aligned}\tilde{\theta}(x_1, x_2, \dots, x_n) &= \tilde{\theta}(y_1, y_2 + y_1, \dots, y_n + y_1) \\ &= \tilde{\theta}(0, y_2, \dots, y_n) + y_1.\end{aligned}$$

Tương tự,

$$\hat{\theta}(x_1, x_2, \dots, x_n) = \hat{\theta}(0, y_2, \dots, y_n) + y_1.$$

Lúc đó

$$\begin{aligned}\tilde{\theta} - \hat{\theta} &= \tilde{\theta}(0, y_2, \dots, y_n) - \hat{\theta}(0, y_2, \dots, y_n) \\ &= \left(\tilde{\theta} - \hat{\theta} \right)(0, y_2, \dots, y_n) := \psi(Y)\end{aligned}$$

Vậy

$$\tilde{\theta} = \hat{\theta} + \psi(Y).$$

Nếu $\mathbb{E}_0 \tilde{\theta}^2 = \infty$ thì hiển nhiên là $\mathbb{E}_\theta (\hat{\theta} - \theta)^2 \leq \mathbb{E}_\theta (\tilde{\theta} - \theta)^2$, $\theta \in \mathbb{R}^1$.

Nếu $\mathbb{E}_0 \tilde{\theta}^2 < \infty$ thì ta có

$$\mathbb{E}_\theta (\tilde{\theta} - \theta)^2 = \mathbb{E}_\theta (\tilde{\theta} - \hat{\theta})^2 + 2\mathbb{E}_\theta (\tilde{\theta} - \hat{\theta}) (\hat{\theta} - \theta) + \mathbb{E}_\theta (\hat{\theta} - \theta)^2.$$

Nhưng

$$\mathbb{E}_\theta (\tilde{\theta} - \hat{\theta}) (\hat{\theta} - \theta) = \mathbb{E}_\theta \hat{\theta} (\tilde{\theta} - \hat{\theta}) - \theta \mathbb{E}_\theta (\tilde{\theta} - \hat{\theta}).$$

Với phép biến đổi $t_n = x_n - \theta$, trong đó $x = (x_1, \dots, x_n)$ và $t = (t_1, \dots, t_n)$ ta có:

$$\begin{aligned} & \mathbb{E}_\theta \hat{\theta} (\tilde{\theta} - \hat{\theta}) \\ &= \int \hat{\theta}(x_1, \dots, x_n) (\tilde{\theta}(x_1, \dots, x_n) - \hat{\theta}(x_1, \dots, x_n)) dF(x - \theta) \\ &= \int \hat{\theta}(t_1 + \theta, \dots, t_n + \theta) (\tilde{\theta}(t_1 + \theta, \dots, t_n + \theta) - \hat{\theta}(t_1 + \theta, \dots, t_n + \theta)) dF(t) \\ &= \int (\hat{\theta}(t_1, \dots, t_n) + \theta) \left((\tilde{\theta}(t_1, \dots, t_n) + \theta) - (\hat{\theta}(t_1, \dots, t_n) + \theta) \right) dF(t) \\ &= \int (\hat{\theta}(t_1, \dots, t_n) + \theta) (\tilde{\theta}(t_1, \dots, t_n) + \theta - \hat{\theta}(t_1, \dots, t_n) - \theta) dF(t) \\ &= \int (\hat{\theta}(t_1, \dots, t_n) + \theta) (\tilde{\theta}(t_1, \dots, t_n) - \hat{\theta}(t_1, \dots, t_n)) dF(t) \\ &= \int \hat{\theta} (\tilde{\theta} - \hat{\theta}) dF(t) + \theta \int (\tilde{\theta} - \hat{\theta}) dF(t) \\ &= \mathbb{E}_0 \hat{\theta} (\tilde{\theta} - \hat{\theta}) + \theta \mathbb{E}_0 (\tilde{\theta} - \hat{\theta}). \end{aligned} \tag{2.2.7}$$

$$\begin{aligned} & \mathbb{E}_\theta (\tilde{\theta} - \hat{\theta}) \\ &= \int (\tilde{\theta}(x_1, \dots, x_n) - \hat{\theta}(x_1, \dots, x_n)) dF(x - \theta) \\ &= \int (\tilde{\theta}(t_1 + \theta, \dots, t_n + \theta) - \hat{\theta}(t_1 + \theta, \dots, t_n + \theta)) dF(t) \\ &= \int \left((\tilde{\theta}(t_1, \dots, t_n) + \theta) - (\hat{\theta}(t_1, \dots, t_n) + \theta) \right) dF(t) \end{aligned}$$

$$\begin{aligned}
&= \int \left(\tilde{\theta}(t_1, \dots, t_n) - \hat{\theta}(t_1, \dots, t_n) \right) dF(t) \\
&= \mathbb{E}_0 \left(\tilde{\theta} - \hat{\theta} \right).
\end{aligned} \tag{2.2.8}$$

Kết hợp (2.2.7) và (2.2.8), ta có

$$\mathbb{E}_\theta \left(\tilde{\theta} - \hat{\theta} \right) \left(\hat{\theta} - \theta \right) = \mathbb{E}_0 \hat{\theta} \left(\tilde{\theta} - \hat{\theta} \right).$$

$$\begin{aligned}
\mathbb{E}_\theta \left(\tilde{\theta} - \hat{\theta} \right) \left(\hat{\theta} - \theta \right) &= \mathbb{E}_0 \hat{\theta} \left(\tilde{\theta} - \hat{\theta} \right) = \mathbb{E}_0 \left(\mathbb{E}_0 \left(\hat{\theta} (\tilde{\theta} - \hat{\theta}) \middle| Y \right) \right) \\
&= \mathbb{E}_0 \left(\mathbb{E}_0 \left(\hat{\theta} \psi(Y) \middle| Y \right) \right) = \mathbb{E}_0 \left(\psi(Y) \mathbb{E}_0 \left(\hat{\theta} \middle| Y \right) \right) \\
\mathbb{E}_0 \left(\hat{\theta} \middle| Y \right) &= \mathbb{E}_0 \left(L - \mathbb{E}_0(L|Y) \middle| Y \right) \\
&= \mathbb{E}_0(L|Y) - \mathbb{E}_0(L|Y) = 0.
\end{aligned}$$

Suy ra $\mathbb{E}_\theta \left(\tilde{\theta} - \hat{\theta} \right) \left(\hat{\theta} - \theta \right) = 0$. Lúc đó $\mathbb{E}_\theta \left(\tilde{\theta} - \theta \right)^2 = \mathbb{E}_\theta \left(\tilde{\theta} - \hat{\theta} \right)^2 + \mathbb{E}_\theta \left(\hat{\theta} - \theta \right)^2$. Vậy $\mathbb{E}_\theta \left(\hat{\theta} - \theta \right)^2 \leq \mathbb{E}_\theta \left(\tilde{\theta} - \theta \right)^2$. Nếu $\mathbb{E}_\theta \left(\hat{\theta} - \theta \right)^2 = \mathbb{E}_\theta \left(\tilde{\theta} - \theta \right)^2$ thì rõ ràng $\mathbb{E}_0 \left(\hat{\theta} \right)^2 = \mathbb{E}_0 \left(\tilde{\theta} \right)^2$, tức là $\tilde{\theta} = \hat{\theta}$ hầu khắp nơi theo $\mathbb{P}_0$, tương đương với $\tilde{\theta} = \hat{\theta}$ hầu khắp nơi theo $\mathbb{P}_\theta, \forall \theta$.

(ii) Do $\sum_1^n c_j = 1$, ta có

$$\begin{aligned}
L &= \sum_1^n c_j X_j = c_1 X_1 + \sum_2^n c_j (X_j - X_1) + \sum_2^n c_j X_1 \\
&= \sum_1^n c_j X_1 + \sum_2^n c_j (X_j - X_1) \\
&= \left(\sum_1^n c_j \right) X_1 + \sum_2^n c_j (X_j - X_1) \\
&= X_1 + \sum_2^n c_j (X_j - X_1).
\end{aligned}$$

Kéo theo

$$\begin{aligned}
\hat{\theta} &= L - \mathbb{E}_0(L|Y) \\
&= X_1 + \sum_2^n c_j(X_j - X_1) - \mathbb{E}_0\left(X_1 + \sum_2^n c_j(X_j - X_1) \middle| Y\right) \\
&= X_1 - \mathbb{E}_0(X_1|Y) + \sum_2^n c_j(X_j - X_1) - \mathbb{E}_0\left(\sum_2^n c_j(X_j - X_1) \middle| Y\right) \\
&= X_1 - \mathbb{E}_0(X_1|Y).
\end{aligned}$$

Phép biến đổi

$$\begin{aligned}
y_1 &= x_1 \\
y_n &= x_n - x_1, \quad n \geq 2
\end{aligned}$$

có $|J| = 1$ với J là ma trận Jacobian.

Ta có

$$F(y_1, \dots, y_n) = \int_{\left\{ \begin{array}{l} x_1 < y_1 \\ x_2 - x_1 < y_2 \\ \dots \\ x_{n-1} - x_1 < y_n \end{array} \right\}} f(x_1, \dots, x_n) dx_1 \cdots dx_n.$$

Suy ra

$$f_Y(y_1, \dots, y_n) = f(y_1, y_2 + y_1, \dots, y_n + y_1).$$

Mật độ có điều kiện của X_1 với điều kiện $y = (y_2, \dots, y_n)$ là

$$\frac{f_Y(y_1, \dots, y_n)}{\int_{-\infty}^{\infty} f_Y(y_1, \dots, y_n) dy_1} = \frac{f(y_1, y_2 + y_1, \dots, y_n + y_1)}{\int_{-\infty}^{\infty} f(t, y_2 + t, \dots, y_n + t) dt}.$$

Kỳ vọng có điều kiện của $\mathbb{E}_0(X_1|Y = y)$ là

$$\begin{aligned}
\mathbb{E}_0(X_1|Y = y) &= \int_{-\infty}^{\infty} y_1 \frac{f_Y(y_1, \dots, y_n)}{\int_{-\infty}^{\infty} f_Y(t, y_2, \dots, y_n) dt} dy_1 \\
&= \frac{\int_{-\infty}^{\infty} t f(t, y_2 + t, \dots, y_n + t) dt}{\int_{-\infty}^{\infty} f(t, y_2 + t, \dots, y_n + t) dt}.
\end{aligned}$$

Do vậy

$$\begin{aligned}\hat{\theta} &= X_1 - \mathbb{E}_0(X_1|Y) = y_1 - \frac{\int_{-\infty}^{\infty} t f(t, y_2 + t, \dots, y_n + t) dt}{\int_{-\infty}^{\infty} f(t, y_2 + t, \dots, y_n + t) dt} \\ &= \frac{\int_{-\infty}^{\infty} (y_1 - t) f(t, y_2 + t, \dots, y_n + t) dt}{\int_{-\infty}^{\infty} f(t, y_2 + t, \dots, y_n + t) dt}.\end{aligned}$$

Đặt $u = y_1 - t$ suy ra $t = y_1 - u$, ta sẽ có

$$\begin{aligned}\hat{\theta} &= \frac{-\int_{-\infty}^{\infty} u f(y_1 - u, y_2 + y_1 - u, \dots, y_n + y_1 - u) du}{-\int_{-\infty}^{\infty} f(y_1 - u, y_2 + y_1 - u, \dots, y_n + y_1 - u) du} \\ &= \frac{\int_{-\infty}^{\infty} u f(x_1 - u, x_2 - u, \dots, x_n - u) du}{\int_{-\infty}^{\infty} f(x_1 - u, x_2 - u, \dots, x_n - u) du}.\end{aligned}$$

□

Bổ đề trên đây là một trường hợp riêng của Bổ đề 2.1.1. Thật vậy, tính chất tối ưu của ước lượng $\hat{\theta}$ trong lớp $\mathcal{T}$ suy ra tính trực giao của $\hat{\theta}$ đối với các ước lượng không chệch của không có dạng $h(X_2 - X_1, \dots, X_n - X_1)$. Chúng ta đưa ra một cách chứng minh của Bổ đề 2.2.1 bởi vì cách chứng minh này có liên hệ với tính tối ưu của các ước lượng trong một lớp đặc biệt.

Bổ đề 2.2.2. Với điều kiện

$$\int |x_j| dF(x_1, \dots, x_n) < \infty, \quad j = 1, \dots, n$$

thì ước lượng

$$\bar{\theta} = L - \text{med}_0(L|Y) \tag{2.2.9}$$

là ước lượng Pitman đối với hàm tổn thất dạng trị tuyệt đối

$$r(\tilde{\theta} - \theta) = |\tilde{\theta} - \theta|.$$

Lưu ý rằng trong biểu thức phía trên $\text{med}_0(L|Y)$ được hiểu là median của phân phối có điều kiện tương ứng với phân phối $\mathbb{P}_0$ của thống kê L với điều kiện Y đã cho. Khi đó $\text{med}_0(L|Y)$ cũng là một thống kê.

Chứng minh. Lấy $\tilde{\theta} \in \mathcal{T}$. Khi đó tương tự như cách chứng minh Bổ đề 2.2.1, dễ thấy $\tilde{\theta} = L + \psi(Y)$ và theo giả thiết hàm tổn thất có dạng hiệu nên ta có

$$\begin{aligned} \mathbb{E}_\theta |\tilde{\theta} - \theta| &= \mathbb{E}_0 |\tilde{\theta}| = \mathbb{E}_0 |L + \psi(Y)| \\ &= \mathbb{E}_0 \left(\left(\mathbb{E}_0 |L + \psi(Y)| \middle| Y \right) \right) \\ &\geq \mathbb{E}_0 \left(\left(\mathbb{E}_0 |L - \text{med}_0(L|Y)| \middle| Y \right) \right) = \mathbb{E}_0 |\bar{\theta}| = \mathbb{E}_\theta |\bar{\theta} - \theta|. \end{aligned}$$

Bởi vì với mỗi y , $\min_{\psi(y)} \mathbb{E}_0 \left(|L + \psi(Y)| \middle| Y = y \right)$ đạt được khi $\psi(y) = -\text{med}_0(L|Y = y)$. $\square$

Đối với trường hợp tổng quát của hàm tổn thất, không có một sự biểu diễn tường minh cho ước lượng Pitman.

Chương 3

Đặc trưng của họ phân phối chuẩn

3.1 Đặc trưng của phân phối chuẩn thông qua tính chấp nhận được của các ước lượng tuyến tính tối ưu của tham số tịnh tiến đối với hàm tổn thất dạng bình phương

Chúng ta giả sử rằng các biến ngẫu nhiên $X_1, \dots, X_n$ có phân phối độc lập và $F_1(x - \theta), \dots, F_n(x - \theta)$ là các hàm phân phối tương ứng phụ thuộc vào tham số tịnh tiến $\theta \in \mathbb{R}^1$. Ngoài ra,

$$0 < \int x^2 dF_j(x) = \sigma_j^2, \quad \forall j. \quad (3.1.1)$$

(trường hợp $\sigma_j = 0$ với j nào đó là tầm thường theo quan điểm về ước lượng cho θ). Để cho đơn giản, ta giả sử rằng

$$\int x dF_j(x) = 0, \quad \forall j \quad (3.1.2)$$

khi đó θ là giá trị trung bình của tổng thể với hàm phân phối $F_j(x - \theta)$. Thật vậy,

$$\mathbb{E}_\theta X_i = \int x dF_i(x - \theta) = \int (y + \theta) dF_i(y)$$

$$= \int y dF_i(y) + \theta \int dF_i(y) = \theta.$$

Cho trước mẫu ngẫu nhiên $X_1, \dots, X_n$. Khi đó đối với hàm tổn thất dạng bình phương, ước lượng không chệch tuyến tính tối ưu cho θ có dạng

$$\hat{L} = \sum_1^n c_j^0 X_j, \text{ trong đó } c_j^0 = \frac{1/\sigma_j^2}{\sum_1^n 1/\sigma_i^2}. \quad (3.1.3)$$

Thật vậy, xét một ước lượng tuyến tính cho θ có dạng $L = \sum_1^n c_j X_j$. Để L là ước lượng không chệch cho θ thì ta cần phải có

$$\mathbb{E}_\theta L = \theta \Leftrightarrow \sum_1^n c_j \mathbb{E}_\theta X_j = \theta \Leftrightarrow \sum_1^n c_j \theta = \theta \Leftrightarrow \theta \sum_1^n c_j = \theta \Leftrightarrow \sum_1^n c_j = 1.$$

Ta thấy

$$\begin{aligned} \mathbb{E}_0 L^2 &= \mathbb{E}_0 \left(\sum_1^n c_j X_j \right)^2 = \mathbb{E}_0 \sum_1^n \sum_1^n c_i c_j X_i X_j \\ &= \sum_1^n \sum_1^n \mathbb{E}_0 c_i c_j X_i X_j = \sum_1^n c_j^2 \mathbb{E}_0 X_j^2 = \sum_1^n c_j^2 \sigma_j^2 \end{aligned}$$

vì nếu $i \neq j$ thì $\mathbb{E}_0 c_i c_j X_i X_j = c_i c_j (\mathbb{E}_0 X_i) (\mathbb{E}_0 X_j) = 0$.

Mặt khác, áp dụng bất đẳng thức Bunyakovski ta có

$$\begin{aligned} \left(\sum_1^n c_j \right)^2 &= \left(\sum_1^n (c_j \sigma_j) \cdot \frac{1}{\sigma_j} \right)^2 \leq \sum_1^n (c_j \sigma_j)^2 \cdot \sum_1^n \frac{1}{\sigma_j^2} \\ &\Leftrightarrow \frac{1}{\sum_1^n 1/\sigma_j^2} \leq \sum_1^n (c_j \sigma_j)^2. \end{aligned}$$

Dấu bằng xảy ra khi và chỉ khi

$$\frac{c_j \sigma_j}{\frac{1}{\sigma_j}} = \text{hằng số} \Leftrightarrow c_j \sigma_j^2 = \text{hằng số} \Leftrightarrow c_j = \frac{\text{hằng số}}{\sigma_j^2}, \quad \forall j.$$

Vì

$$\begin{aligned}\sum_1^n c_j = 1 &\Leftrightarrow \sum_1^n \frac{\text{hàng số}}{\sigma_j} = 1 \Leftrightarrow (\text{hàng số}) \cdot \sum_1^n \frac{1}{\sigma_j} = 1 \\ &\Leftrightarrow \text{hàng số} = \frac{1}{\sum_1^n 1/\sigma_j^2}\end{aligned}$$

nên dấu bằng đạt được khi

$$c_j = \frac{\text{hàng số}}{\sigma_j^2} = \frac{1/\sigma_j^2}{\sum_1^n 1/\sigma_j^2}.$$

(Nếu điều kiện (3.1.2) không thỏa mãn, và $\int x dF_j(x) = b_j$ thì (3.1.3) được thay bởi $\hat{L} = \sum_1^n c_j^0 (X_j - b_j)$ và mọi lập luận vẫn giữ nguyên).

Xét định lý sau

Định lý 3.1.1 (xem [9]). Cho $n \geq 3$ và các biến ngẫu nhiên X_j độc lập có các hàm phân phối lần lượt là $F_j(x - \theta)$ với $j = 1, \dots, n$. Giả sử các điều kiện (3.1.1) và (3.1.2) được thỏa mãn. Khi đó, đối với hàm tổn thất dạng bình phương ước lượng tuyến tính tối ưu $\hat{L}$ cho bởi (3.1.3) là chấp nhận được trong lớp các ước lượng không chệch của θ khi và chỉ khi F_j là phân phối chuẩn, $\forall j = 1, \dots, n$.

Chứng minh.

Điều kiện cần: Giả sử $\hat{L}$ là chấp nhận được trong lớp các ước lượng không chệch của θ . Theo Bổ đề 2.2.1, ước lượng cho bởi

$$\hat{\theta} = \hat{L} - \mathbb{E}_0(\hat{L} | Y), \text{ trong đó } Y = (X_2 - X_1, \dots, X_n - X_1)$$

là ước lượng Pitman.

Vì hàm tổn thất có dạng hiệu nên

$$\begin{aligned}\mathbb{E}_\theta(\hat{\theta} - \theta) &= \mathbb{E}_0 \hat{\theta} = \mathbb{E}_0(\hat{L} - \mathbb{E}_0(\hat{L} | Y)) \\ &= \mathbb{E}_0 \hat{L} - \mathbb{E}_0(\mathbb{E}_0(\hat{L} | Y)) \\ &= \mathbb{E}_0 \hat{L} - \mathbb{E}_0 \hat{L} = 0.\end{aligned}$$

Suy ra $\mathbb{E}_\theta \hat{\theta} = \theta$, tức $\hat{\theta}$ là một ước lượng không chệch của θ . Nếu $\hat{L} \neq \hat{\theta}$ thì theo Bổ đề 2.2.1, ta phải có

$$\mathbb{E}_\theta \left(\hat{L} - \theta \right)^2 > \mathbb{E}_\theta \left(\hat{\theta} - \theta \right)^2, \quad \forall \theta \in \mathbb{R}^1.$$

Bất đẳng thức trên chứng tỏ $\hat{L}$ không chấp nhận được. Điều này mâu thuẫn với giả thiết $\hat{L}$ chấp nhận được. Vậy ta phải có $\hat{L} = \hat{\theta}$ hầu khắp nơi theo $[\mathbb{P}_0]$, tức là

$$\mathbb{E}_0 \left(\hat{L} | Y \right) = 0. \quad (3.1.4)$$

Đặt

$$\begin{aligned} L_1 &= \hat{L} = \sum_1^n c_j^0 X_j \\ L_2 &= X_2 - X_1 \\ &\dots \\ L_n &= X_n - X_1. \end{aligned}$$

Khi đó $L_1, \dots, L_n$ độc lập tuyến tính, vì

$$\begin{vmatrix} c_1^0 & -1 & -1 & \dots & -1 \\ c_2^0 & 1 & 0 & \dots & 0 \\ c_3^0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_n^0 & 0 & 0 & \dots & 1 \end{vmatrix} = \begin{vmatrix} \sum_1^n c_j^0 & 0 & 0 & \dots & 0 \\ c_2^0 & 1 & 0 & \dots & 0 \\ c_3^0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_n^0 & 0 & 0 & \dots & 1 \end{vmatrix} = \sum_1^n c_j^0 = 1 \neq 0.$$

Rõ ràng tất cả các hệ số $c_j^0 \neq 0$ theo cách xác định c_j^0 . Hơn nữa $\mathbb{E}_0 X_j = 0, \forall j$. Vậy theo Hệ quả 1.1.1 của Định lý 1.1.1 ta suy ra X_j có phân phối chuẩn nếu $j \geq 3$.

Điều kiện đủ: Giả sử F_j là các hàm phân phối chuẩn $N(0, \sigma_j^2)$. Mật độ đồng thời của $X_1, \dots, X_n$ là

$$\begin{aligned}
L(x, \theta) &= \left\{ \prod_1^n (2\pi\sigma_j^2)^{-1/2} \right\} \exp \left[-\frac{1}{2} \sum_1^n \frac{x_j^2}{\sigma_j^2} + \theta \sum_1^n \frac{x_j}{\sigma_j^2} - \frac{1}{2} \theta^2 \sum_1^n \frac{1}{\sigma_j^2} \right] \\
&= \left\{ \prod_1^n (2\pi\sigma_j^2)^{-1/2} \exp \left[-\frac{1}{2} \sum_1^n \frac{x_j^2}{\sigma_j^2} \right] \right\} \cdot \exp \left[\theta \sum_1^n \frac{x_j}{\sigma_j^2} - \frac{1}{2} \theta^2 \sum_1^n \frac{1}{\sigma_j^2} \right] \\
&= H(x) \cdot G(T(x), \theta).
\end{aligned}
\tag{3.1.5}$$

Theo định lý tách của Neyman, dễ thấy thống kê $T = \sum_1^n \frac{x_j}{\sigma_j^2}$ (hoặc thống kê tương đương $S = \sum_1^n c_j^0 X_j$) là thống kê đủ đối với họ mật độ (3.1.5), với $\theta \in \mathbb{R}^1$. Phân phối của S là chuẩn (vì S là tổ hợp tuyến tính của các biến ngẫu nhiên có phân phối chuẩn), trong đó

$$\begin{aligned}
\mathbb{E}_\theta S &= \sum_1^n c_j^0 \mathbb{E}_\theta X_j = \theta \sum_1^n c_j^0 = \theta, \\
\text{Var}_\theta S &= \sum_1^n c_j^0 \text{Var}_\theta X_j = \sum_1^n c_j^0 \sigma_j^2 := \sigma^2.
\end{aligned}$$

Khi đó S có mật độ

$$\begin{aligned}
p(s, \theta) &= \frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{(s - \theta)^2}{2\sigma^2} \right] \\
&= \left\{ \frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{\theta^2}{2\sigma^2} \right] \right\} \cdot \exp \left[-\frac{s^2}{2\sigma^2} \right] \cdot \exp \left[\frac{s\theta}{\sigma^2} \right].
\end{aligned}$$

Theo Định lý 1.6.3 dễ thấy S là thống kê đầy đủ. S đủ, đầy đủ và S là ước lượng không chệch đối với θ nên nó là ước lượng UMVU cho θ . Vì hiểm đối với hàm tổn thất dạng bình phương chính là phương sai, mà phương sai của S bé nhất nên S là ước lượng không chệch tối ưu cho θ . $\square$

Hệ quả 3.1.1. Nếu $F = F_j$ với mọi $j = 1, \dots, n$ ($n \geq 3$), và F thỏa mãn (3.1.1) và (3.1.2) thì tính chấp nhận được của trung bình mẫu $\bar{X}$ trong lớp các ước lượng không chệch của tham số tịnh tiến θ là một tính chất đặc trưng của phân phối chuẩn.

Chứng minh. Thật vậy, từ $F = F_j$ suy ra $\sigma_j = \sigma, \forall j, c_j^0 = \frac{1}{n}, \forall j$. Khi đó $\widehat{L} = \frac{1}{n} \sum_1^n X_j = \overline{X}$. Theo Định lý 3.1.1, $\overline{X}$ chấp nhận được trong lớp các ước lượng không chệch của θ là tính chất đặc trưng của phân phối chuẩn. $\square$

Hệ quả 3.1.2. *Giả sử $X_1, \dots, X_n$ và $Y_1, \dots, Y_n$ là các biến ngẫu nhiên độc lập, $n \geq 3$, trong đó $X_j \sim F_j(x - \theta - \Delta), Y_j \sim G_j(y - \theta)$, F_j và G_j thỏa mãn (3.1.1) và (3.1.2) với $\theta \in \mathbb{R}^1, \Delta \in \mathbb{R}^1$. Để một ước lượng có dạng $\sum_1^n a_j (X_j - Y_j)$ là chấp nhận được trong lớp các ước lượng không chệch của tham số Δ đối với hàm tổn thất dạng bình phương, điều kiện cần và đủ là tất cả F_j và G_j phải là phân phối chuẩn, ngoài ra $\sigma_j^2 = \int x^2 dF_j(x)$ và $\rho_j^2 = \int y^2 dG_j(y)$ liên hệ với nhau bởi hệ thức $\sigma_j^2 = \lambda \rho_j^2, \forall j$.*

Chứng minh. Trước hết, từ tính chấp nhận được của $\sum_1^n a_j (X_j - Y_j)$ trong lớp các ước lượng có dạng $\sum_1^n \alpha_j (X_j - Y_j)$ ta suy ra $a_j \neq 0, \forall j$. Bởi lẽ tất cả các ước lượng dạng trên với $\sum_1^n a_j = 1$ đều là ước lượng không chệch của Δ . Hiếm của những ước lượng này là hằng số và bằng $\sum_1^n a_j^2 (\sigma_j^2 + \rho_j^2)$. Vì vậy tính chấp nhận được đồng nghĩa với tính tối ưu (do đối với hàm tổn thất dạng hiệu, một ước lượng tuyến tính hoặc tối ưu, hoặc không chấp nhận được), tức là các a_j phải có dạng giống như c_j^0 trong (3.1.3), suy ra $a_j \neq 0, \forall j$. Đặt $Z_j = X_j - Y_j, x' = x - \theta - \Delta, y' = y - \theta$, ta có

$$\begin{aligned} P_{\theta, \Delta}(Z_j < u) &= \iint_{x-y < u} dF_j(x - \theta - \Delta) dG_j(y - \theta) \\ &= \iint_{x'-y' < u-\Delta} dF_j(x') dG_j(y') = H_j(u - \Delta), \end{aligned}$$

tức là, hàm phân phối của Z_j chỉ phụ thuộc vào tham số tịnh tiến Δ . Từ tính chấp nhận được của ước lượng $\sum_1^n a_j Z_j$ trong lớp các ước lượng không chệch của Δ và Định lý 3.1.1, ta suy ra Z_j có phân phối chuẩn. Nhưng theo Định lý Cramer, X_j và Y_j cũng có phân phối chuẩn: $X_j \sim N(0, \sigma_j^2)$ và $Y_j \sim N(0, \rho_j^2)$. Hơn thế nữa, dễ dàng chỉ ra rằng với họ phân phối chuẩn

$$F_1(x_1 - \theta - \Delta) \cdots F_n(x_n - \theta - \Delta) G_1(y_1 - \theta) \cdots G_n(y_n - \theta) \quad (3.1.6)$$

trong đó $\theta \in \mathbb{R}^1, \Delta \in \mathbb{R}^1$, các tổ hợp tuyến tính $\sum_1^n c_j^0 X_j$ và $\sum_1^n d_j^0 Y_j$ trong đó

$$c_j^0 = \frac{1}{\sigma_j^2} \left(\sum_1^n \frac{1}{\sigma_i^2} \right)^{-1}, \quad d_j^0 = \frac{1}{\rho_j^2} \left(\sum_1^n \frac{1}{\rho_i^2} \right)^{-1}$$

là các thống kê đủ và đầy đủ. Theo Định lý 1.6.2, từ tính chấp nhận được của $\sum_1^n a_j (X_j - Y_j)$ trong lớp các ước lượng không chệch của Δ , suy ra

$$\begin{aligned} \sum_1^n a_j (X_j - Y_j) &= \beta \sum_1^n c_j^0 X_j + \gamma \sum_1^n d_j^0 Y_j \\ &= \sum_1^n \beta c_j^0 X_j + \sum_1^n \gamma d_j^0 Y_j. \end{aligned}$$

Kéo theo $a_j = \beta c_j^0, -a_j = \gamma d_j^0$. Vì các $a_j \neq 0, \forall j$ nên $\beta \neq 0$ và $\gamma \neq 0$, khi đó

$$\beta c_j^0 = -\gamma d_j^0 \Leftrightarrow c_j^0 = -\frac{\gamma}{\beta} d_j^0 := \frac{1}{\xi} d_j^0.$$

Dễ thấy $\sum_1^n c_j^0 = \sum_1^n d_j^0 = 1$, suy ra

$$\sum_1^n c_j^0 = \frac{1}{\xi} \sum_1^n d_j^0 = 1 \Rightarrow \xi = 1.$$

Dẫn tới $c_j^0 = d_j^0$, từ đó suy ra $\sigma_j^2 = \lambda \rho_j^2, \quad \forall j = 1, \dots, n$.

Ước lượng $\sum_1^n a_j (X_j - Y_j)$ là hàm của thống kê đủ, đầy đủ và là ước lượng không chệch của Δ nên là ước lượng không chệch tối ưu duy nhất cho Δ .

Ngược lại, Nếu F_j và G_j là phân phối chuẩn, đồng thời có hệ thức $\sigma_j^2 = \lambda \rho_j^2, \forall j = 1, \dots, n$ trong đó $\sigma_j^2 = \int x^2 dF_j(x)$ và $\rho_j^2 = \int y^2 dG_j(y)$ thì ta đặt $Z_j = X_j - Y_j$. Vì X_j và Y_j có phân phối chuẩn nên Z_j có phân phối chuẩn với phương sai $\omega_j^2 = \sigma_j^2 + \rho_j^2$. Xét $a_j = \frac{1/\omega_j^2}{\sum_1^n 1/\omega_i^2}$. Khi đó, theo Định lý 3.1.1, $L = \sum_1^n a_j Z_j$ là một ước lượng chấp nhận được trong lớp các ước lượng không chệch cho tham số Δ .

□

Nhận xét 3.1.1. Kết quả của Định lý 3.1.1 có thể được mở rộng trong trường hợp $n = \infty$ nếu kèm thêm điều kiện $\sum_1^n \frac{1}{\sigma_j^2} < \infty$ (nếu tổng này phân kỳ thì luôn tồn tại một ước lượng tuyến tính $\sum_1^\infty c_j^0 X_j$ với phương sai nhỏ tùy ý). Nếu tổng này hội tụ thì theo Định lý 3.1.1, ước lượng $\hat{L} = \sum_1^\infty c_j X_j$, trong đó $c_j^0 = \frac{1/\sigma_j^2}{\sum_1^\infty 1/\sigma_j^2}$ là chấp nhận được trong lớp các ước lượng không chệch khi và chỉ khi các F_j là phân phối chuẩn. Để chứng minh tính chuẩn của các hàm phân phối $F_1, \dots, F_n$ ta lưu ý, từ $\mathbb{E}_0 \left(\hat{L} | X_2 - X_1, \dots \right) = 0$ ta suy ra

$$\mathbb{E}_0 \left(\hat{L} | X_2 - X_1, \dots, X_n - X_1 \right) = 0. \quad (3.1.7)$$

Thật vậy, chúng ta biết kết quả sau đây từ xác suất cơ sở: Nếu $\mathcal{F}_1$ và $\mathcal{F}_2$ là hai σ - đại số sao cho $\mathcal{F}_1 \subset \mathcal{F}_2$ thì $\mathbb{E}(X | \mathcal{F}_1) = \mathbb{E} \left(\mathbb{E}(X | \mathcal{F}_2) | \mathcal{F}_1 \right) = \mathbb{E} \left(\mathbb{E}(X | \mathcal{F}_1) | \mathcal{F}_2 \right)$, và dễ thấy

$$\sigma(X_2 - X_1, \dots, X_n - X_1) \subset \sigma(X_2 - X_1, \dots).$$

Khi đó, áp dụng kết quả trên ta được

$$\begin{aligned} 0 &= \mathbb{E}_0 \left(\mathbb{E}_0 \left(\hat{L} | X_2 - X_1, \dots \right) | X_2 - X_1, \dots, X_n - X_1 \right) \\ &= \mathbb{E}_0 \left(\mathbb{E}_0 \left(\hat{L} | X_2 - X_1, \dots, X_n - X_1 \right) | X_2 - X_1, \dots \right) \\ &= \mathbb{E}_0 \left(\hat{L} | X_2 - X_1, \dots, X_n - X_1 \right). \end{aligned}$$

Vì $\hat{L} = \sum_1^n c_j X_j + \sum_{n+1}^\infty c_j X_j$ và $\sum_{n+1}^\infty c_j X_j$ là biến ngẫu nhiên độc lập với $Y = (X_2 - X_1, \dots, X_n - X_1)$ nên từ (3.1.7) ta có

$$\begin{aligned} \mathbb{E}_0 \left(\hat{L} | Y \right) &= \mathbb{E}_0 \left(\sum_1^n c_j X_j | Y \right) + \mathbb{E}_0 \left(\sum_{n+1}^\infty c_j X_j | Y \right) \\ &= \mathbb{E}_0 \left(\sum_1^n c_j X_j | Y \right) + \mathbb{E}_0 \left(\sum_{n+1}^\infty c_j X_j \right) \\ &= \mathbb{E}_0 \left(\sum_1^n c_j X_j | Y \right) + \sum_{n+1}^\infty c_j \mathbb{E}_0 X_j \end{aligned}$$

$$= \mathbb{E}_0 \left(\sum_1^n c_j X_j \middle| Y \right).$$

Do đó

$$\mathbb{E}_0 \left(\sum_1^n c_j X_j \middle| Y \right) = 0.$$

Đẳng thức trên chính là (3.1.4) trong chứng minh của Định lý 3.1.1, suy ra $F_1, \dots, F_n$ có phân phối chuẩn.

Ngược lại, Nếu các F_j có phân phối chuẩn $N(0, \sigma_j^2)$, trong đó $\sum_1^n \frac{1}{\sigma_j^2} < \infty$ thì dễ dàng kiểm tra rằng $\sum_1^\infty c_j^0 X_j$ là thống kê đủ, đầy đủ đối với họ phân bố xác suất $\prod_1^\infty F_j(x - \theta)$, $\theta \in \mathbb{R}^1$ (sử dụng Định lý 8.2.1 của [10] và các kết quả trong [5] hoặc [11]). Do đó ước lượng này là ước lượng tối ưu trong lớp các ước lượng không chệch của tham số θ .

Nhận xét 3.1.2. Khi $n = 2$, tính chấp nhận được của $\widehat{L}$ trong lớp các ước lượng không chệch của θ không phải là một đặc trưng của phân phối chuẩn. Để lấy ví dụ, ta có thể xét trường hợp $F_1 = F_2 = F$, trong đó F là một phân phối đối xứng nào đó. Khi đó $c_1^0 = c_2^0 = \frac{1}{2}$, (3.1.4) luôn được thỏa mãn. Vì thế ước lượng Pitman trong trường hợp này là $\overline{X}$. Theo Định lý Stein (xem [19]), điều này đảm bảo cho tính chấp nhận được (thậm chí là chấp nhận được tuyệt đối) của $\overline{X}$ đối với phân phối F với giả thiết là $\int |x|^3 dF(x) < \infty$.

Định lý 3.1.2. Với giả thiết của Định lý 3.1.1, tính chấp nhận được tuyệt đối của $\widehat{L} = \sum_1^n c_j^0 X_j$ trong lớp các ước lượng cho θ là một tính chất đặc trưng cho tính chuẩn của F_j .

Chứng minh. Hiển nhiên ta chỉ cần chứng minh rằng nếu F_j là phân phối chuẩn thì $\widehat{L}$ là chấp nhận được tuyệt đối. Theo Định lý 1.6.2, chúng ta chỉ cần giới hạn trong lớp các ước lượng là hàm của thống kê đủ $\widehat{L}$. Dễ thấy, đối với một ước lượng bất kỳ $\tilde{\theta}(\widehat{L})$ với $\mathbb{E}_\theta \tilde{\theta}^2 < \infty$, $\theta \in \mathbb{R}^1$, điều kiện chính quy đòi hỏi khi rút ra bất đẳng thức (1.6.1), được thỏa mãn. Đặt $\mathbb{E}_\theta \tilde{\theta} = \theta + b(\theta)$, $\sigma^2 = \sum_1^n (c_j^0)^2 \sigma_j^2$, khi đó áp dụng (1.6.1) với $t(\theta) = \theta$,

$$I(\theta) = \mathbb{E}_\theta \left(\frac{\partial \log p}{\partial \theta} \right)^2 = \frac{1}{\sigma^2} \text{ ta có}$$

$$\mathbb{E}_\theta (\tilde{\theta} - \theta)^2 \geq [b(\theta)]^2 + [1 + b'(\theta)]^2 \sigma^2.$$

Nếu ước lượng $\hat{L}$ là không chấp nhận được (trong lớp các ước lượng là hàm của thống kê đủ $\hat{L}$), thì với $\tilde{\theta}$ nào đó ta phải có

$$[b(\theta)]^2 + [1 + b'(\theta)]^2 \sigma^2 \leq \mathbb{E}_\theta (\tilde{\theta} - \theta)^2 \leq \mathbb{E}_\theta (\hat{L} - \theta)^2 = \sigma^2 \quad (3.1.8)$$

với $\theta \in \mathbb{R}^1$ và dấu “<” xảy ra tại ít nhất một điểm $\theta \in \mathbb{R}^1$ nào đó. Từ (3.1.8) ta suy ra rằng $b'(\theta) \leq 0$ và $b(\theta)$ bị chặn. Đó đó luôn tồn tại $\lim b(\theta)$ khi $\theta \rightarrow \pm\infty$, suy ra $\overline{\lim} b'(\theta)$ tồn tại và $= 0$ khi $\theta \rightarrow \pm\infty$. Nhưng cũng từ (3.1.8) ta suy ra rằng $\lim b(\theta) = 0$ khi $\theta \rightarrow \pm\infty$. $b(\theta)$ là một hàm đơn điệu giảm, bị chặn, lại có $b(-\infty) = b(+\infty) = 0$ nên $b(\theta) = 0, \forall \theta$. Vì thế, nếu (3.1.8) được thỏa mãn thì $\tilde{\theta}(\hat{L})$ là một ước lượng không chệch của θ . Nhưng theo phần chứng minh điều kiện đủ của Định lý 3.1.1, $\hat{L}$ là một ước lượng không chệch tối ưu cho θ . Vì đối với họ đầy đủ, tham số θ có tối đa một ước lượng không chệch phụ thuộc thống kê đủ cho nên ta phải có $\tilde{\theta}(\hat{L}) = \hat{L}$ hầu chắc chắn. Điều này mâu thuẫn với (3.1.8), suy ra $\hat{L}$ chấp nhận được tuyệt đối. $\square$

3.2 Điều kiện để trung bình mẫu chấp nhận được trong một lớp con các ước lượng của tham số tịnh tiến

3.2.1 Các ước lượng không chệch chỉ phụ thuộc trung bình mẫu và một thống kê tuyến tính khác

Giả sử rằng $X_1, \dots, X_n$ là các biến ngẫu nhiên độc lập cùng phân phối với hàm phân phối $F(x - \theta)$, $\theta \in \mathbb{R}^1$. F thỏa mãn điều kiện

$$\int x dF(x) = 0, \quad \int x^2 dF(x) < \infty. \quad (3.2.1)$$

Giả sử rằng, ngoài trung bình mẫu $\bar{X}$ ra thì giá trị của thống kê tuyến tính khác $L = \sum_1^n a_j X_j$ được biết theo kinh nghiệm. Một câu hỏi đặt ra là: Đối với những loại hàm phân phối F nào thì θ tồn tại ước lượng không chệch chỉ phụ thuộc L và $\bar{X}$ và ước lượng này sẽ tốt hơn $\bar{X}$ theo nghĩa khi ta sử dụng hàm tổn thất dạng bình phương? Để trả lời cho câu hỏi này, ta xét thống kê sau

$$V = L - (n\bar{a})\bar{X} = \sum_1^n b_j X_j, \quad b_j = a_j - \bar{a}, \quad \bar{a} = \frac{1}{n} \sum_1^n a_j.$$

Dễ thấy rằng $\sum_1^n b_j = 0$, $\frac{1}{n} \sum_1^n (X_j - \theta) = \bar{X} - \theta$, còn $\sum_1^n b_j (X_j - \theta) = V - \theta \sum_1^n b_j = V$. Mặt khác, từ tính chất của tham số tịnh tiến θ , ta luôn có

$$\begin{aligned} \mathbb{E}_\theta \left(\frac{1}{n} \sum_1^n (X_j - \theta) \middle| \sum_1^n b_j (X_j - \theta) \right) &= \mathbb{E}_0 \left(\frac{1}{n} \sum_1^n X_j \middle| \sum_1^n b_j X_j \right) \\ &= \mathbb{E}_0(\bar{X} | V). \end{aligned}$$

Dẫn đến

$$\mathbb{E}_\theta(\bar{X} - \theta | V) = \mathbb{E}_0(\bar{X} | V). \quad (3.2.2)$$

Ta suy ra

$$\mathbb{E}_\theta(\mathbb{E}_0(\bar{X} | V)) = \mathbb{E}_\theta(\mathbb{E}_\theta(\bar{X} | V) - \theta) = \mathbb{E}_\theta \bar{X} - \theta = 0.$$

Ngoài ra ta cũng có

$$\begin{aligned} \mathbb{E}_\theta \left(\mathbb{E}_\theta(\bar{X} - \theta | V) \right)^2 &= \mathbb{E}_\theta \left(\mathbb{E}_\theta \left(\frac{1}{n} \sum_1^n (X_j - \theta) \middle| \sum_1^n b_j (X_j - \theta) \right) \right)^2 \\ &= \mathbb{E}_0 \left(\mathbb{E}_0 \left(\frac{1}{n} \sum_1^n X_j \middle| \sum_1^n b_j X_j \right) \right)^2 \\ &= \mathbb{E}_0(\mathbb{E}_0(\bar{X} | V))^2. \end{aligned} \quad (3.2.3)$$

Kết hợp (3.2.2) với (3.2.3) ta có

$$\mathbb{E}_\theta \left(\mathbb{E}_0(\bar{X} | V) \right)^2 = \mathbb{E}_\theta \left(\mathbb{E}_\theta(\bar{X} - \theta | V) \right)^2 = \mathbb{E}_0 \left(\mathbb{E}_0(\bar{X} | V) \right)^2. \quad (3.2.4)$$

Vì thế, ước lượng $\tilde{\theta} = \overline{X} - \mathbb{E}_0(\overline{X}|V)$ thỏa mãn

$$\begin{aligned}
\mathbb{E}_\theta(\tilde{\theta} - \theta)^2 &= \mathbb{E}_\theta(\overline{X} - \theta)^2 - 2\mathbb{E}_\theta(\overline{X} - \theta)\mathbb{E}_0(\overline{X}|V) + \mathbb{E}_\theta(\mathbb{E}_0(\overline{X}|V))^2 \\
&= \mathbb{E}_\theta(\overline{X} - \theta)^2 - 2\mathbb{E}_\theta(\overline{X} - \theta)\mathbb{E}_\theta(\overline{X} - \theta|V) + \mathbb{E}_\theta(\mathbb{E}_0(\overline{X}|V))^2 \\
&= \mathbb{E}_\theta(\overline{X} - \theta)^2 - 2\mathbb{E}_\theta\left(\mathbb{E}_\theta\left((\overline{X} - \theta)\mathbb{E}_\theta(\overline{X} - \theta|V)\middle|V\right)\right) \\
&\quad + \mathbb{E}_\theta(\mathbb{E}_0(\overline{X}|V))^2 \\
&= \mathbb{E}_\theta(\overline{X} - \theta)^2 - 2\mathbb{E}_\theta\left(\mathbb{E}_\theta(\overline{X} - \theta|V)\right)^2 + \mathbb{E}_\theta(\mathbb{E}_0(\overline{X}|V))^2.
\end{aligned}$$

Thay (3.2.3) và (3.2.4) vào đẳng thức trên ta được

$$\mathbb{E}_\theta(\tilde{\theta} - \theta)^2 = \mathbb{E}_\theta(\overline{X} - \theta)^2 - \mathbb{E}_0(\mathbb{E}_0(\overline{X}|V))^2.$$

Vì thế, điều kiện cần để $\overline{X}$ là một ước lượng tuyến tính trong lớp các ước lượng không chệch chỉ phụ thuộc $\overline{X}$ và L không thể cải thiện được như trên đó là

$$\mathbb{E}_0(\overline{X}|V) = 0. \quad (3.2.5)$$

Trong các kết quả về hồi quy hàng số ta đã chứng minh được (3.2.5) không đảm bảo cho tính chuẩn của X_j . Tuy nhiên, (3.2.5) cùng với một vài điều kiện ràng buộc cho b_j và F_j thì lại đảm bảo cho tính chuẩn này. Bây giờ ta minh họa một kết quả như vậy, kết quả này suy trực tiếp từ Định lý 1.1.4.

Ta xác định hàm

$$G(\lambda) = \sum_{b_i > 0} b_j^{\lambda-1} - \sum_{b_i < 0} |b_j|^{\lambda-1}$$

và giả sử rằng $G(\lambda) \not\equiv 0$. Giả σ là cận trên đúng của phần thực của các không điểm của $G(\lambda)$.

Định lý 3.2.1. Cho $X_1, \dots, X_n$ là các mẫu ngẫu nhiên rút từ tổng thể với hàm phân phối $F(x - \theta)$, $\theta \in \Theta$, thỏa mãn điều kiện (3.2.1) và điều kiện bổ sung

$$\int x^{2m} dF(x) < \infty, \text{ trong đó } m = [1 + \sigma/2].$$

Khi đó tính chấp nhận được của $\overline{X}$ đối với hàm tổn thất dạng bình phương trong lớp các ước lượng không chệch cho θ có dạng $\tilde{\theta} \left(\overline{X}, \sum_1^n a_j X_j \right)$ là một tính chất đặc trưng của phân phối chuẩn.

3.2.2 Lớp các ước lượng đa thức

Chúng ta giữ lại các giả thiết ở các phần trước, xét các biến ngẫu nhiên độc lập cùng phân phối $X_1, \dots, X_n, X_j \sim F(x - \theta), \theta \in \mathbb{R}^1$. Giả thiết rằng F thỏa mãn (3.2.1) và với số nguyên $k \geq 1$,

$$\mu_{2k} = \int x^{2k} dF(x) < \infty. \quad (3.2.6)$$

Trong trường hợp này, tập tất cả các đa thức $Q(X_1, \dots, X_n)$ có bậc $\leq k$ tạo thành không gian Hilbert M_k với tích vô hướng thông thường: $(Q_1, Q_2) = \mathbb{E}_0(Q_1 Q_2)$. M_k chứa đủ nhiều các ước lượng không chệch cho θ (tất cả ước lượng này đều là một phần tử của M_k , có phương sai hữu hạn); chẳng hạn các ước lượng có dạng $\overline{X} + Q(X_2 - X_1, \dots, X_n - X_1)$, trong đó đa thức Q thỏa mãn điều kiện $\mathbb{E}_0 Q = 0$. Trong trường hợp này $\mathbb{E}_\theta Q = 0$. Ký hiệu $P_k, P_k \subset M_k$ là lớp tất cả các phần tử của M_k là ước lượng không chệch cho θ .

Định lý 3.2.2 (xem [6], [7]). Giả sử $n \geq 3$ và $X_1, \dots, X_n$ là mẫu ngẫu nhiên rút từ tổng thể với hàm phân phối $F(x - \theta), \theta \in \mathbb{R}^1$, và thỏa mãn điều kiện (3.2.1) và (3.2.5). Khi đó tính chấp nhận được của $\overline{X}$ với tư cách là một ước lượng cho θ trong lớp P_k đối với hàm tổn thất dạng bình phương là tính chất đặc trưng của phân phối F với $k+1$ mô men đầu tiên trùng với trùng với các mô men tương ứng của phân phối chuẩn nào đó (với kỳ vọng bằng 0), tức là

$$\mu_m = \begin{cases} 0, & \text{nếu } m \text{ lẻ,} \\ (m-1)!!\sigma^m, & \text{nếu } m \text{ chẵn,} \end{cases} \quad 1 \leq m \leq k+1$$

trong đó $\sigma \geq 0$ và $n!! = 1 \cdot 3 \cdot 5 \cdots n$ với n lẻ.

Để chứng minh định lý này, trước hết ta chứng minh Bổ đề 3.2.1 và Bổ đề 3.2.2 dưới đây.

Ký hiệu Λ_k là không gian con các đa thức có dạng

$$Q(X_2 - X_1, \dots, X_n - X_1).$$

Dễ thấy $\Lambda_k \subset M_k$. Xét ước lượng

$$\hat{\theta}_k = \overline{X} - \hat{\mathbb{E}}_0(\overline{X} | \Lambda_k) \quad (3.2.7)$$

trong đó $\hat{\mathbb{E}}_0(\cdot | \Lambda_k)$ là toán tử hình chiếu trên không gian con Λ_k . Ước lượng (3.2.7) là dạng ước lượng Pitman.

Bổ đề 3.2.1. Với mọi $\theta \in \mathbb{R}^1$,

$$\mathbb{E}_\theta \hat{\theta}_k = \theta \quad (3.2.8)$$

và

$$\mathbb{E}_\theta (\hat{\theta}_k - \theta)^2 \leq \mathbb{E}_\theta (\overline{X} - \theta)^2 \quad (3.2.9)$$

trong đó đẳng thức hoặc bất đẳng thức trong (3.2.9) xảy ra hầu đều đối với mọi $\theta \in \mathbb{R}^1$, và đẳng thức xảy ra khi và chỉ khi

$$\hat{\mathbb{E}}_0(\overline{X} | \Lambda_k) = 0. \quad (3.2.10)$$

Chứng minh. Vì $\mathbf{1} \in \Lambda_k$, trong khi $\hat{\mathbb{E}}_0(\overline{X} | \Lambda_k)$ là hình chiếu của $\overline{X}$ lên Λ_k nên $\mathbf{1}$ trực giao với $\overline{X} - \hat{\mathbb{E}}_0(\overline{X} | \Lambda_k)$, tức là $\mathbb{E}_0[\overline{X} - \hat{\mathbb{E}}_0(\overline{X} | \Lambda_k)] = 0$, hay $\mathbb{E}_0 \overline{X} = \mathbb{E}_0[\hat{\mathbb{E}}_0(\overline{X} | \Lambda_k)] = 0$, nên

$$\mathbb{E}_\theta \hat{\theta}_k = \mathbb{E}_\theta \overline{X} - \mathbb{E}_\theta[\hat{\mathbb{E}}_0(\overline{X} | \Lambda_k)] = \theta - \mathbb{E}_\theta[\hat{\mathbb{E}}_0(\overline{X} | \Lambda_k)].$$

Lại chú ý rằng, do $\hat{\mathbb{E}}_0(\overline{X} | \Lambda_k) \in \Lambda_k$, nên

$$\mathbb{E}_\theta[\hat{\mathbb{E}}_0(\overline{X} | \Lambda_k)] = \mathbb{E}_0[\hat{\mathbb{E}}_0(\overline{X} | \Lambda_k)] = 0.$$

Ta nhận được (3.2.8) và do đó $\hat{\theta}_k \in P_k$. Hơn nữa

$$\begin{aligned} \mathbb{E}_\theta (\overline{X} - \theta)^2 &= \mathbb{E}_0 (\overline{X})^2 = \mathbb{E}_0 [\hat{\theta}_k + \hat{\mathbb{E}}_0(\overline{X} | \Lambda_k)]^2 \\ &= \mathbb{E}_0 (\hat{\theta}_k)^2 + \mathbb{E}_0 [\hat{\mathbb{E}}_0(\overline{X} | \Lambda_k)]^2 \end{aligned}$$

$$\begin{aligned}
&= \mathbb{E}_\theta \left(\hat{\theta}_k - \theta \right)^2 + \mathbb{E}_0 \left[\hat{\mathbb{E}}_0 \left(\overline{X} | \Lambda_k \right) \right]^2 \\
&\geq \mathbb{E}_\theta \left(\hat{\theta}_k - \theta \right)^2.
\end{aligned}$$

Lưu ý rằng $\hat{\theta}_k$ trực giao với không gian Λ_k nên $\mathbb{E}_0 \left[\hat{\theta}_k \hat{\mathbb{E}}_0 \left(\overline{X} | \Lambda_k \right) \right] = 0$.

Đẳng thức xảy ra khi và chỉ khi

$$\mathbb{E}_0 \left[\hat{\mathbb{E}}_0 \left(\overline{X} | \Lambda_k \right) \right]^2 = 0 \Leftrightarrow \hat{\mathbb{E}}_0 \left(\overline{X} | \Lambda_k \right) = 0.$$

□

Bổ đề 3.2.2. Nếu $n \geq 3$, khi đó từ đẳng thức $\hat{\mathbb{E}}_0(\overline{X}|\Lambda_k) = 0$ suy ra rằng $(k+1)$ mô men đầu tiên của F trùng với những mô men cấp tương ứng của một phân phối chuẩn (có thể suy biến) nào đó.

Chứng minh. Chúng ta sẽ chứng minh theo phương pháp quy nạp. Với $k = 1$, Bổ đề hiển nhiên đúng. Vì $\Lambda_k \supset \Lambda_{k-1}$, nên từ $\hat{\mathbb{E}}_0(\overline{X}|\Lambda_k) = 0$ ta suy ra rằng $\hat{\mathbb{E}}_0(\overline{X}|\Lambda_{k-1}) = 0$. Do đó, chúng ta có thể giả sử rằng k mô men đầu tiên của F trùng với những mô men cấp tương ứng của phân phối chuẩn nào đó, và chỉ phải chứng minh rằng mô men μ_{k+1} của F trùng với mô men thứ $(k+1)$ của phân phối chuẩn nói trên. Điều kiện (3.2.10) tương đương với

$$\hat{\mathbb{E}}_0(\overline{X}|\Lambda_{k-1}) = 0 \tag{3.2.11}$$

và

$$\mathbb{E}_0 \left[\overline{X} (X_{j_1} - X_1) \cdots (X_{j_k} - X_1) \right] = 0 \tag{3.2.12}$$

trong đó (3.2.12) phải đúng với mọi phép chọn $j_1, \dots, j_k$ từ các số $2, \dots, n$. Sự tương đương của đẳng thức (3.2.10) và (3.2.11) – (3.2.12) là bởi vì Λ_{k-1} và các hàm $(X_{j_1} - X_1) \cdots (X_{j_k} - X_1)$ sinh ra Λ_k . Ta viết

$$(X_{j_1} - X_1) \cdots (X_{j_k} - X_1) = (X_{i_1} - X_1)^{\alpha_1} \cdots (X_{i_s} - X_1)^{\alpha_s} \tag{3.2.13}$$

trong đó $i_1, \dots, i_s$ là những số phân biệt và $\sum_{r=1}^s \alpha_r = k$.

Với ký hiệu tổ hợp C_k^r , ta có

$$\begin{aligned} (X_{i_1} - X_1)^{\alpha_1} \cdots (X_{i_s} - X_1)^{\alpha_s} &= \sum_{p_1=0}^{\alpha_1} \cdots \sum_{p_s=0}^{\alpha_s} (-1)^p C_{\alpha_1}^{p_1} \cdots C_{\alpha_s}^{p_s} \\ &\quad \times X_{i_1}^{p_1} \cdots X_{i_s}^{p_s} X_1^{k-p}, \quad p = \sum_{i=1}^s p_i. \end{aligned} \quad (3.2.14)$$

Và vì vậy, ta có

$$\begin{aligned} &\mathbb{E}_0 \left[\overline{X} (X_{i_1} - X_1)^{\alpha_1} \cdots (X_{i_s} - X_1)^{\alpha_s} \right] \\ &= \frac{1}{n} \sum_{p_1=0}^{\alpha_1} \cdots \sum_{p_s=0}^{\alpha_s} (-1)^p C_{\alpha_1}^{p_1} \cdots C_{\alpha_s}^{p_s} \mu_{p_1} \cdots \mu_{p_s} \mu_{k+1-p} \\ &\quad + \frac{1}{n} \sum_{q=1}^s \sum_{p_1=0}^{\alpha_1} \cdots \sum_{p_s=0}^{\alpha_s} (-1)^p C_{\alpha_1}^{p_1} \cdots C_{\alpha_s}^{p_s} \mu_{p_1} \cdots \mu_{p_{q-1}} \mu_{p_q+1} \mu_{p_{q+1}} \cdots \mu_{p_s} \mu_{k-p}. \end{aligned} \quad (3.2.15)$$

Chúng ta xem xét riêng biệt hai trường hợp k chẵn và k lẻ.

Trường hợp k chẵn. Từ giả thiết quy nạp, tất cả các mô men cấp lẻ từ 1 tới $(k-1)$ (bao gồm cả $k-1$) là 0 (vì mô men cấp lẻ của phân phối chuẩn bằng 0). Do đó, từ (3.2.15) ta suy ra

$$\begin{aligned} &n \mathbb{E}_0 \left[\overline{X} (X_{i_1} - X_1)^{\alpha_1} \cdots (X_{i_s} - X_1)^{\alpha_s} \right] \\ &= \sum^* (-1)^p C_{\alpha_1}^{p_1} \cdots C_{\alpha_s}^{p_s} \mu_{p_1} \cdots \mu_{p_s} \mu_{k+1-p} \\ &\quad + \sum_{q=1}^s \sum_q^* (-1)^p C_{\alpha_1}^{p_1} \cdots C_{\alpha_s}^{p_s} \mu_{p_1} \cdots \mu_{p_{q-1}} \mu_{p_q+1} \mu_{p_{q+1}} \cdots \mu_{p_s} \mu_{k-p}. \end{aligned} \quad (3.2.16)$$

trong đó, $\sum^*$ là tổng lấy hạn chế trên những số chẵn trong các số $p_1, \dots, p_s$ và $\sum_q^*$ là tổng lấy trên những số chẵn $p_1, \dots, p_{q-1}, p_{q+1}, \dots, p_s$ còn số p_q là số lẻ. Trong tổng $\sum^*$, p luôn là số chẵn và do đó $k+1-p$ là số lẻ. Nếu như $k+1-p \leq k-1$ thì $\mu_{k+1-p} = 0$ theo giả thiết quy nạp. Còn trong tổng $\sum_q^*$, p luôn là số lẻ, vì vậy $k-p$ cũng là số lẻ. Cũng vì $k-p \leq k-1$ nên

$\mu_{k-p} = 0$. Vì vậy

$$n\mathbb{E}_0 \left[\overline{X} (X_{i_1} - X_1)^{\alpha_1} \cdots (X_{i_s} - X_1)^{\alpha_s} \right] = \mu_{k+1}$$

và điều kiện (3.2.12) tương đương với

$$\mu_{k+1} = 0. \quad (3.2.17)$$

Trường hợp k lẻ. Chúng ta quay lại đẳng thức (3.2.16). Ta chia 2 tổng ở đó ra như sau

$$\sum^* = S_0 + S_2 + \cdots + S_{k-1}$$

trong đó S_{2m} là phần của tổng $\sum^*$ tương ứng với tất cả $p_1, \dots, p_s$, $0 \leq p_1 \leq \alpha_1, \dots, 0 \leq p_s \leq \alpha_s$ mà thỏa mãn $p_1 + \cdots + p_s = 2m$, và

$$\sum_{q=1}^s \sum_q^* = S_1 + S_3 + \cdots + S_k$$

trong đó S_{2m+1} là phần tổng bội tương ứng với những giá trị của $p_1, \dots, p_s$ thỏa mãn $p_1 + \cdots + p_s = 2m + 1$. Chúng ta chú ý rằng, với $n \geq 3$, tồn tại những đa thức dạng $(X_{i_1} - X_1)^{\alpha_1} \cdots (X_{i_s} - X_1)^{\alpha_s}$ với $s \geq 2$ và các giá trị α_i dương.

Chúng ta chỉ cần xem xét những đa thức có dạng như vậy. Ta sẽ chỉ ra rằng, với $0 < 2m \leq k - 1$, $S_{2m} + S_{2m+1} = 0$. Chú ý rằng, với điều kiện $\alpha_1 > 0, \dots, \alpha_s > 0$, trong tổng $\sum_{q=1}^s \sum_q^*$, ta luôn luôn có $p_q + 1 \leq k$ và vì vậy, ta suy ra từ giả thiết quy nạp rằng

$$\mu_{p_q+1} = p_q \mu_{p_q-1} \sigma^2 \quad (3.2.18)$$

với $\sigma^2 \geq 0$ nào đó. Khi đó,

$$\begin{aligned} S_{2m+1} &= - \sum_{q=1}^s \sum_q^* C_{\alpha_1}^{p_1} \cdots C_{\alpha_s}^{p_s} \mu_{p_1} \cdots \mu_{p_{q-1}} \mu_{p_q+1} \mu_{p_{q+1}} \cdots \mu_{p_s} \mu_{k-1-2m} \\ &= - \sum_{p_1 + \cdots + p_s = 2m}^* C_{\alpha_1}^{p_1} \cdots C_{\alpha_s}^{p_s} \mu_{p_1} \cdots \mu_{p_s} \mu_{k-1-2m} \\ &\quad \times \sigma^2 \left(\frac{C_{\alpha_1}^{p_1+1}}{C_{\alpha_1}^{p_1}} (p_1 + 1) + \cdots + \frac{C_{\alpha_s}^{p_s+1}}{C_{\alpha_s}^{p_s}} (p_s + 1) \right). \end{aligned} \quad (3.2.19)$$

Lại dễ thấy rằng

$$\frac{C_{\alpha_1}^{p_1+1}}{C_{\alpha_1}^{p_1}}(p_1 + 1) + \cdots + \frac{C_{\alpha_s}^{p_s+1}}{C_{\alpha_s}^{p_s}}(p_s + 1) = \sum_{i=1}^s (\alpha_i - p_i) = k - 2m.$$

Do vậy,

$$\mu_{k-1-2m}\sigma^2(k - 2m) = \mu_{k+1-2m}$$

với $m > 0$. Thay đẳng thức trên vào (3.2.19) ta được

$$S_{2m} + S_{2m+1} = 0. \quad (3.2.20)$$

Từ (3.2.20), ta thấy rằng (3.2.12) tương đương với

$$S_0 + S_1 = 0. \quad (3.2.21)$$

Tuy nhiên, ta lại có

$$S_0 = \mu_{k+1}$$

và

$$S_1 = -(\alpha_1 + \cdots + \alpha_s)\mu_2\mu_{k-1} = -k\sigma^2\mu_{k-1}.$$

Hệ quả là (3.2.21) sẽ tương đương với

$$\mu_{k+1} = k\sigma^2\mu_{k-1}. \quad (3.2.22)$$

Cùng với giả thiết quy nạp rằng mô men của k cấp đầu tiên trùng với mô men tương ứng của một phân phối chuẩn có kỳ vọng 0 và phương sai σ^2 nào đó, công thức (3.2.22) kéo theo $\mu_{k+1} = 1 \cdot 3 \cdots k\sigma^{k+1}$. Bổ đề được chứng minh. $\square$

Chúng ta trở lại việc chứng minh Định lý 3.2.2. Tính chuẩn của phân phối F ở đây sẽ được suy ra ngay từ tính chấp nhận được của $\overline{X}$ trong lớp P_k . Thật vậy, theo Bổ đề 3.2.1, ước lượng $\hat{\theta}_k = \overline{X} - \hat{\mathbb{E}}_0(\overline{X}|\Lambda_k)$ cho tham số θ là một ước lượng không chệch và với mọi $\theta \in \mathbb{R}^1$, có phương sai nhỏ hơn phương sai của $\overline{X}$ (trừ trường hợp $\hat{\mathbb{E}}_0(\overline{X}|\Lambda_k) = 0$). Theo Bổ đề 3.2.2, điều kiện này đảm bảo tính chuẩn của phân phối F .

Công việc còn lại là cần chỉ ra rằng nếu $(k + 1)$ mô men của F trùng với những mô men tương ứng của phân phối chuẩn thì ta có tính chấp nhận được của $\overline{X}$ nếu coi như là một ước lượng của θ trong lớp P_k . Đầu tiên, ta chú ý rằng $\hat{\theta}_k$ là một ước lượng tối ưu của θ trong lớp các đa thức cấp không quá k có dạng $\overline{X} + Q(X_2 - X_1, \dots, X_n - X_1)$. Thật vậy, vì

$$\mathbb{E}_0 \left[\hat{\theta}_k \cdot \left(Q(X_2 - X_1, \dots, X_n - X_1) + \hat{\mathbb{E}}_0(\overline{X} | \Lambda_k) \right) \right] = 0$$

nên

$$\begin{aligned} & \mathbb{E}_\theta \left(\overline{X} + Q(X_2 - X_1, \dots, X_n - X_1) - \theta \right)^2 \\ &= \mathbb{E}_0 \left(\overline{X} + Q(X_2 - X_1, \dots, X_n - X_1) \right)^2 \\ &= \mathbb{E}_0 \left[\hat{\theta}_k + Q(X_2 - X_1, \dots, X_n - X_1) + \hat{\mathbb{E}}_0(\overline{X} | \Lambda_k) \right]^2 \\ &= \mathbb{E}_0 (\hat{\theta}_k)^2 + \mathbb{E}_0 \left[Q(X_2 - X_1, \dots, X_n - X_1) + \hat{\mathbb{E}}_0(\overline{X} | \Lambda_k) \right]^2 \\ &\geq \mathbb{E}_0 (\hat{\theta}_k)^2 = \mathbb{E}_\theta (\hat{\theta}_k - \theta)^2. \end{aligned}$$

Hơn nữa, nếu $(k + 1)$ mô men ban đầu của F trùng với những mô men tương ứng của một phân phối chuẩn nào đó, khi đó $\hat{\theta}_k = \overline{X}$. Vì điều kiện $\hat{\mathbb{E}}_0(\overline{X} | \Lambda_k) = 0$ được thỏa mãn với mọi F có các mô men $\mu_1, \dots, \mu_k$ như phân phối chuẩn nói trên hoặc không thỏa mãn với mọi F có các mô men như vậy. Tuy nhiên, nếu F có phân phối chuẩn, khi đó $\hat{\mathbb{E}}_0(\overline{X} | \Lambda_k) = 0$, vì trong trường hợp này, $\overline{X}$ độc lập với các vectơ $(X_2 - X_1, \dots, X_n - X_1)$. Điều cần chứng minh được suy ra ngay từ bổ đề sau:

Bổ đề 3.2.3. *Nếu μ_{2k} tồn tại thì $\hat{\theta}_k$ là chấp nhận được trong lớp M_k tất cả các đa thức có bậc $\leq k$ (và hiển nhiên cũng chấp nhận được trong lớp P_k).*

Chứng minh. Mọi đa thức có bậc không quá k có thể viết dưới dạng

$$\begin{aligned} Q(X) &= \overline{X}^m q_0(X_2 - X_1, \dots, X_n - X_1) \\ &\quad + \overline{X}^{m-1} q_1(X_2 - X_1, \dots, X_n - X_1) + \dots \\ &\quad + q_m(X_2 - X_1, \dots, X_n - X_1) \end{aligned}$$

trong đó $m \leq k$ và ta giả sử rằng $\mathbb{P}_0[q_0 \neq 0] > 0$ (trong trường hợp $q_0 = 0$ hầu chắc chắn đối với $[\mathbb{P}_0]$ thì ta thay Q bởi $Q - \overline{X}^m q_0$). Các q_j với $0 \leq j \leq m$ là các đa thức theo các biến của chúng. Giả sử rằng đối với đa thức Q nào đó điều kiện sau thỏa mãn với mọi θ

$$\mathbb{E}_\theta(Q - \theta)^2 \leq \mathbb{E}_\theta(\hat{\theta}_k - \theta)^2 = \mathbb{E}_0(\hat{\theta}_k)^2. \quad (3.2.23)$$

Ta có

$$\begin{aligned} \mathbb{E}_\theta(Q - \theta)^2 &= \mathbb{E}_0 \left[(\overline{X} + \theta)^m q_0(X_2 - X_1, \dots, X_n - X_1) + \dots \right. \\ &\quad \left. + q_m(X_2 - X_1, \dots, X_n - X_1) - \theta \right]^2 \\ &= \theta^{2m} \mathbb{E}_0 q_0^2 + \pi_{2m-1}(\theta) \end{aligned}$$

trong đó bậc của $\pi_{2m-1}(\theta) \leq \max(2, 2m-1)$. Khi đó rõ ràng là bất đẳng thức (3.2.23) được thỏa mãn với mọi θ chỉ khi $m \leq 1$, tức là chỉ các đa thức có dạng $Q = \overline{X}q_0 + q_1$ mới có thể tốt hơn ước lượng $\hat{\theta}_k$; đối với những đa thức ước lượng này bởi vì nếu $m \leq 1$, $\mathbb{E}_\theta(Q - \theta)^2$ là một đa thức bậc $2m$ của θ , và hiểm của nó tiến đến ∞ khi $\theta \rightarrow \infty$. Trong khi $\mathbb{E}_\theta(\hat{\theta} - \theta)^2 < \mathbb{E}_0(\overline{X})^2 =$ hằng số, tức là sẽ tồn tại θ để hiểm của $\hat{\theta}$ nhỏ hơn hiểm của Q tại θ . Với $m = 1$ ta có

$$\begin{aligned} \mathbb{E}_\theta(Q - \theta)^2 &= \theta^2 \mathbb{E}_0(q_0 - 1)^2 + 2\theta \mathbb{E}_0[(q_0 - 1)\overline{X}q_1] \\ &\quad + \mathbb{E}_0(\overline{X}q_0 + q_1)^2. \end{aligned}$$

Từ đó suy ra rằng nếu (3.2.23) đúng thì $q_0(X_2 - X_1, \dots, X_n - X_1) = 1$ với xác suất bằng 1 theo $[\mathbb{P}_0]$ (và vì thế đúng với xác suất bằng 1 theo $[\mathbb{P}_\theta]$, với mọi θ) vì nếu trái lại $\mathbb{E}_\theta(Q - \theta)^2 \rightarrow \infty$ khi $\theta \rightarrow \infty$. Nhưng trong số các đa thức ước lượng có dạng $\overline{X} + q_0(X_2 - X_1, \dots, X_n - X_1)$ thì $\hat{\theta}_k$ không những là ước lượng chấp nhận được mà còn là ước lượng tối ưu. Do đó (3.2.23) không thể thỏa mãn với mọi θ , điều này chứng tỏ $\hat{\theta}_k$ chấp nhận được trong lớp M_k . Bổ đề 3.2.3 và Định lý 3.2.2 vì thế được chứng minh. $\square$

Hệ quả 3.2.1. *Giả sử các điều kiện của Định lý 3.2.2 được thỏa mãn và phân phối F có mô men tuyệt đối mọi cấp. Khi đó tính chấp nhận được của trung bình mẫu $\bar{X}$ trong lớp $P_\infty = \bigcup_1^\infty P_k$ là tính chất đặc trưng cho của chuẩn.*

Chứng minh. Thật vậy, theo Định lý 3.2.2 tất cả các mô men của F trùng với các mô men tương ứng của luật phân phối chuẩn nào đó, vì thế F là phân phối chuẩn. $\square$

Lưu ý rằng lớp P_∞ thực sự bé hơn lớp các ước lượng không chệch cho θ vì thế hệ quả trên không mạnh hơn Hệ quả 3.1.1 của Định lý 3.1.1 bởi vì Hệ quả 3.1.1 không cần đến giả thiết về sự tồn tại mô men cấp cao bất kỳ của F .

Nhận xét 3.2.1. Việc kiểm tra kỹ lưỡng phần chứng minh của Định lý 3.2.2 nảy sinh ra kết quả sau đây: Cho (X_1, X_2) là mẫu ngẫu nhiên rút từ tổng thể với hàm phân phối $F(x - \theta)$, thỏa mãn điều kiện (3.2.1) và (3.2.5). Tính chấp nhận được của $\frac{1}{2}(X_1 + X_2)$ cho tham số θ trong lớp P_k là tính chất đặc trưng của hàm phân phối F với F là phân phối có các mô men cấp lẻ (kể cả đến cấp $k + 1$) đều bằng 0, tức là

$$\begin{aligned} \mu_1 = \mu_3 = \cdots = \mu_k & \quad \text{nếu } k \text{ lẻ} \\ \mu_1 = \mu_3 = \cdots = \mu_{k+1} & \quad \text{nếu } k \text{ chẵn} \end{aligned}$$

Tính chấp nhận được của $\frac{1}{2}(X_1 + X_2)$ trong lớp P_k không áp đặt cho các mô men cấp chẵn.

3.3 Đặc trưng của luật chuẩn bởi tính tối ưu của hàm của ước lượng $\hat{L} = \sum c_j^0 X_j$

Chúng ta trở lại với những giả thiết trong phần 3.1. Cho $X_1, \dots, X_n$ là những biến ngẫu nhiên độc lập có phân phối $X_j \sim F_j(x - \theta)$, $\theta \in \mathbb{R}^1$ và F_j

thỏa mãn điều kiện (3.1.1) và (3.1.2). Với điều kiện

$$\int x^{2k} dF_j(x) < \infty, \quad j = 1, \dots, n \quad (3.3.1)$$

đa thức $q(\widehat{L}) = a_0 \widehat{L}^k + \dots + a_k$ là một ước lượng không chệch của đa thức tham số $g(\theta) = \mathbb{E}_\theta q = c_0 \theta^k + \dots + c_k$ với phương sai hữu hạn với mọi θ . Tính tối ưu của $q(\widehat{L})$ trong lớp các ước lượng không chệch của $g(\theta)$ với hàm tổn thất dạng bình phương, hóa ra lại đặc trưng cho phân phối chuẩn. Thật vậy, ta có định lý sau

Định lý 3.3.1. Cho $X_1, \dots, X_n$, $n \geq 3$, là những biến ngẫu nhiên độc lập với $X_j \sim F_j(x - \theta)$, trong đó $\theta \in \mathbb{I}$ với $\mathbb{I}$ một khoảng không suy biến nào đó, và F_j thỏa mãn điều kiện (3.3.1). Đa thức cấp $k \geq 1$: $q(\widehat{L}) = a_0 \widehat{L}^k + \dots + a_k$, $a_0 \neq 0$ là tối ưu trong lớp các ước lượng không chệch trong $\mathcal{L}^2$ của hàm tham số $g(\theta) = \mathbb{E}_\theta(q)$ với mọi $\theta \in \mathbb{I}$ và hàm tổn thất bình phương nếu và chỉ nếu mọi phân phối F_j là phân phối chuẩn.

Chứng minh.

Điều kiện cần: Theo Bổ đề 2.1.1, từ tính tối ưu của $q(\widehat{L})$ trong lớp các ước lượng không chệch trong $\mathcal{L}^2$ của $g(\theta)$, kéo theo rằng, với mọi hàm

$$h(X_2 - X_1, \dots, X_n - X_1) = h \in \mathcal{H}$$

ta có

$$\mathbb{E}_\theta [q(\widehat{L}) \cdot h] = 0, \quad \theta \in \mathbb{I}. \quad (3.3.2)$$

Nhưng

$$\begin{aligned} \mathbb{E}_\theta(qh) &= \mathbb{E}_0 [q(\widehat{L} + \theta)h] \\ &= \mathbb{E}_0 \left([a_0(\widehat{L} + \theta)^k + \dots + a_k] h \right). \end{aligned} \quad (3.3.3)$$

Vế phải của (3.3.3) là một đa thức đối với θ . Từ (3.3.2) ta suy ra rằng mọi hệ số của đa thức này phải bằng không. Hệ số của θ^{k-1} là

$$\mathbb{E}_0 [(a_0 k \widehat{L} + a_1)h] = k a_0 \mathbb{E}_0 (\widehat{L}h).$$

Vì $a_0 \neq 0$ nên $\mathbb{E}_0(\widehat{L}h) = 0$ với mọi $h \in \mathcal{H}$. Nhưng điều này không khác gì điều kiện (3.1.5), vì thế F_j phải là phân phối chuẩn.

Điều kiện đủ: Nếu F_j là phân phối chuẩn (với kỳ vọng bằng 0) thì $\widehat{L}$ là một thống kê đủ và đầy đủ đối với họ $F_1(x_1 - \theta) \cdots F_n(x_n - \theta)$. Do đó mọi hàm của thống kê $\widehat{L}$ với phương sai hữu hạn là ước lượng không chệch tối ưu của kỳ vọng toán tương ứng của chúng trong $\mathcal{L}^2$. Vì thế, Định lý 3.3.1 được chứng minh. $\square$

Đối với các tập Borel A , đặt

$$\nu_\theta(A) = \int_{L \in A} dF_1(x_1 - \theta) \cdots dF_n(x_n - \theta).$$

Đối với một hàm bị chặn $\psi(u)$, áp dụng như mục 1.2 với $d\mathbb{P}_\theta = d\nu_\theta(u)$, ta xây dựng $\mathcal{L}_\theta^2(\psi)$ và $\mathcal{L}^2(\psi) = \bigcap_{\theta \in \mathbb{I}} \mathcal{L}_\theta^2(\psi)$ trong đó $\mathbb{I}$ là một khoảng không suy biến nào đó. Ta đưa ra điều kiện

$$“\mathcal{L}^2(\psi) \text{ chứa một đa thức } q \text{ không phải là hằng số}”. \quad (3.3.4)$$

Khi đó ta có hệ quả sau đây của Định lý 3.3.1

Hệ quả 3.3.1. Cho $X_1, \dots, X_n$ với $n \geq 3$ là các biến ngẫu nhiên độc lập, $X_j \sim F_j(x - \theta)$, $\theta \in \mathbb{I}$. Nếu các điều kiện (3.1.1) và (3.3.4) được thỏa mãn thì tính tối ưu của $\psi(\widehat{L})$ trong $\mathcal{L}^2$ với ψ bị chặn trong lớp các ước lượng không chệch của $\mathbb{E}_\theta \psi(\widehat{L})$ là tính chất đặc trưng cho luật chuẩn của hàm phân phối F_j .

Chứng minh. Ta sẽ chỉ ra rằng tính tối ưu trong $\mathcal{L}^2$ của $\psi(\widehat{L})$ sẽ kéo theo tính chuẩn của F_j (điều ngược lại được chứng minh từ thực tế rằng $\widehat{L}$ là một thống kê đủ và đầy đủ). Theo Định lý 2.1.1, tính chuẩn trong $\mathcal{L}^2$ của $\psi(\widehat{L})$ trong lớp các ước lượng không chệch của $\mathbb{E}_\theta \psi$ và điều kiện (3.3.4) đảm bảo rằng đa thức $q(\widehat{L})$ trong lớp các ước lượng không chệch cho $\mathbb{E}_\theta q$ là tối ưu trong $\mathcal{L}^2$, $\theta \in \mathbb{I}$. Nhưng khi đó theo Định lý 3.3.1, các F_j là chuẩn. $\square$

Để kết luận, ta chú ý rằng nếu thống kê tuyến tính $L^* = \sum_1^n c_j X_j$ khác với $\widehat{L}$ thì không tồn tại một đa thức $q_0(L^*)$ có bậc ≥ 1 và một hàm $\psi(L^*)$ thỏa mãn các điều kiện của Hệ quả 3.3.1, mà có thể là ước lượng tối ưu (cho $\theta \in \mathbb{I}$) với tư cách là ước lượng không chệch của kỳ vọng toán học của đa thức đó.

3.4 Đặc trưng luật chuẩn bởi tính chấp nhận được và tính tối ưu của ước lượng bình phương bé nhất trong mô hình Gauss - Markov

Cho vectơ ngẫu nhiên $\mathbf{Y}' = (Y_1, \dots, Y_n)$ ($\mathbf{Y}'$ là vectơ chuyển vị của vectơ $\mathbf{Y}$) trong đó Y_i là các biến ngẫu nhiên độc lập, $i = 1, \dots, n$. $\mathbf{Y}'$ có vectơ trung bình $\boldsymbol{\theta}' = (\theta_1, \dots, \theta_n)$. Trong mô hình Gauss - Markov, vectơ trung bình $\boldsymbol{\theta}$ có dạng $\boldsymbol{\theta} = \mathbf{C}\boldsymbol{\beta}$, với $\mathbf{C}$ là ma trận cấp $n \times m$ chưa biết và vectơ tham số $\boldsymbol{\beta}' = (\beta_1, \dots, \beta_m)$ đã biết và nhận giá trị trong miền $U \subset \mathbb{R}^m$. Nếu hạng của ma trận $\mathbf{C}$ bằng r thì không mất tổng quát, ta giả sử $\mathbf{C}$ có dạng chính tắc

$$\mathbf{A} = \begin{pmatrix} \mathbf{I}_r \\ \mathbf{B} \end{pmatrix} \quad (3.4.1)$$

trong đó $\mathbf{I}_r$ là ma trận đơn vị cấp $r \times r$ và $\mathbf{B}$ là ma trận cấp $n - r \times r$ thỏa mãn, nếu

$$\mathbf{Z} = \begin{pmatrix} -\mathbf{B}' \\ \mathbf{I}_{n-r} \end{pmatrix} \quad (3.4.2)$$

thì $\mathbf{A}'\mathbf{Z} = \mathbf{0}$.

Cuối cùng ta giả sử rằng các hàm phân phối F_j của các sai số ε_j trong biểu diễn $Y_j = \theta_j + \varepsilon_j$, $j = 1, 2, \dots, n$, có kỳ vọng bằng 0 và phương sai bằng nhau chưa biết.

Chúng ta biết rằng ước lượng bình phương bé nhất (LSE) của mọi hàm tham số ước lượng được $\mathbf{p}'\boldsymbol{\beta}$ (tức là tồn tại ước lượng tuyến tính $\mathbf{q}'\mathbf{Y}$ thỏa

mã $\mathbb{E}_\beta(\mathbf{q}'\mathbf{Y}) = \mathbf{p}'\beta$, $\beta \in U$) có dạng $\boldsymbol{\lambda}'\mathbf{A}'\mathbf{Y}$ với cách chọn vector $\boldsymbol{\lambda}'$ thích hợp, và mọi thống kê tuyến tính có dạng $\boldsymbol{\lambda}'\mathbf{A}'\mathbf{Y}$ là ước lượng LSE của kỳ vọng toán học của nó (xem [18]). Các ước lượng LSE là tối ưu trong lớp các ước lượng không chệch tuyến tính đối với hàm tổn thất dạng bình phương với hàm phân phối F_j tùy ý. Ta cũng biết rằng nếu F_j là phân phối chuẩn với mọi j thì các ước lượng LSE sẽ tối ưu trong lớp ước lượng không chệch của các hàm tham ẩn tương ứng (xem [18]).

Định lý 3.4.1. *Giả sử rằng hai hàng bất kỳ của ma trận $\mathbf{B}$ cho bởi (3.4.1) có các phần tử khác không trên ít nhất một cột chung và không cột nào của ma trận $\mathbf{B}$ chứa toàn phần tử không. Khi đó với vector $\boldsymbol{\lambda}$ tùy ý, nếu $\boldsymbol{\lambda}'\mathbf{A}'\mathbf{Y}$ là chấp nhận được trong lớp các ước lượng không chệch của hàm $\mathbb{E}_\beta(\boldsymbol{\lambda}'\mathbf{A}'\mathbf{Y}) = \mathbf{p}'\beta$ thì F_j là phân phối chuẩn.*

Nói cách khác, với ma trận $\mathbf{B}$ xác định như trên, nếu ít nhất một trong số các F_j không phải là phân phối chuẩn và hoàn toàn biết được dạng phân phối của nó thì đối với vector $\boldsymbol{\lambda}$ nào đó, ước lượng $\boldsymbol{\lambda}'\mathbf{A}'\mathbf{Y}$ không chấp nhận được trong lớp các ước lượng không chệch của $\mathbf{p}'\beta$ đối với hàm tổn thất dạng bình phương.

Chứng minh. Xét ước lượng sau đây cho $\mathbf{p}'\beta$:

$$T = \boldsymbol{\lambda}' \left[\mathbf{A}'\mathbf{Y} - \mathbb{E}_0(\mathbf{A}'\mathbf{Y}|\mathbf{Z}'\mathbf{Y}) \right]. \quad (3.4.3)$$

Dễ thấy ước lượng trên có dạng như (2.2.6). Ngoài ra

$$\mathbb{E}_\beta T = \mathbb{E}_\beta(\boldsymbol{\lambda}'\mathbf{A}'\mathbf{Y}) = \mathbf{p}'\beta,$$

và

$$\begin{aligned} \mathbb{E}_\beta(T - \mathbf{p}'\beta)^2 &= \mathbb{E}_\beta(\boldsymbol{\lambda}'\mathbf{A}'\mathbf{Y} - \mathbf{p}'\beta)^2 - \mathbb{E}_0 \left[\mathbb{E}_0(\boldsymbol{\lambda}'\mathbf{A}'\mathbf{Y}|\mathbf{Z}'\mathbf{Y}) \right]^2 \\ &\leq \mathbb{E}_\beta(\boldsymbol{\lambda}'\mathbf{A}'\mathbf{Y} - \mathbf{p}'\beta)^2 \end{aligned} \quad (3.4.4)$$

trong đó dấu đẳng thức xảy ra với mọi $\beta \in U$ khi và chỉ khi

$$\mathbb{E}_0(\boldsymbol{\lambda}'\mathbf{A}'\mathbf{Y}|\mathbf{Z}'\mathbf{Y}) = \boldsymbol{\lambda}'\mathbb{E}_0(\mathbf{A}'\mathbf{Y}|\mathbf{Z}'\mathbf{Y}) = 0. \quad (3.4.5)$$

Nếu với λ bất kỳ, ước lượng $\lambda'A'Y$ chấp nhận được trong lớp các ước lượng không chệch của $p'\beta$ thì từ (3.4.5) ta suy ra rằng $\mathbb{E}_0(A'Y|Z'Y) = 0$. Gọi $L_i, i = 1, \dots, r$ là các vectơ cột của $A'Y$ và $L_i, i = r + 1, \dots, n$ là các vectơ cột của $Z'Y$. Gọi b_{ij} là phần tử hàng i cột j của ma trận B , ta có

$$L_1 = X_i + \sum_1^{n-r} b_{j-r,i} Y_j \text{ với } i = 1, \dots, r$$

và

$$L_1 = X_i - \sum_1^r b_{i,j} Y_j \text{ với } i = r + 1, \dots, n.$$

Với giả thiết hai hàng bất kỳ của B đều có phần tử khác không tại ít nhất một cột ta suy ra rằng các X_i liên kết được với nhau với $i = r + 1, \dots, n$. Cùng với giả thiết không cột nào của B chứa toàn phần tử 0, áp dụng hệ quả của Định lý 1.1.2, ta kết luận rằng, đối với ma trận A xác định như trong điều kiện của định lý thì các F_j là phân phối chuẩn. $\square$

Ta lưu ý rằng điều kiện của Định lý 3.4.1 về tính chấp nhận được của ước lượng $\lambda'A'Y$ với mọi λ tương đương với tính chấp nhận được của ước lượng LSE của r hàm tham ẩn độc lập tuyến tính $p'_1\beta, \dots, p'_r\beta$. Bởi vì mỗi vectơ $p'_i\beta$, đều tồn tại λ_i sao cho $p'_i\beta = \lambda'_i A'Y$ và ngược lại mọi $\lambda'A'Y$ đều là tổ hợp tuyến tính của các $p'_1\beta, \dots, p'_r\beta$.

Nếu các F_j trùng nhau thì tồn tại một cột của ma trận B có ít nhất hai phần tử khác không.

Ta tiếp tục nghiên cứu các điều kiện mà qua đó tính chuẩn của F_j được suy ra từ tính chấp nhận được của ước lượng LSE với k hàm tham ẩn độc lập tuyến tính nào đó, $k < r$. Trước hết ta xét trường hợp hàm một ẩn.

Định lý 3.4.2. Cho $L = \sum_1^n a_j Y_j$ là ước lượng LSE của $p'\beta$. Giả sử rằng mọi a_j đều khác không và B là ma trận mà mọi hàng của nó đều khác vectơ không hoặc không tỉ lệ với một hàng nào đó của I_r hoặc không tỉ lệ với một hàng khác của chính ma trận B . Khi đó tính chấp nhận được của L trong lớp các ước lượng không chệch của $p'\beta$ suy ra tính chuẩn của F_j .

Chứng minh. Tương tự cách chứng minh của Định lý 3.4.1, từ tính chấp nhận được của L ta suy ra

$$\mathbb{E}_0(L|\mathbf{Z}'\mathbf{Y}) = 0.$$

Sử dụng Định lý 1.1.1 ta thu được điều cần chứng minh. $\square$

Cho $\mathbf{D}$ là ma trận cấp $n \times k$ và $\mathbf{D}'\mathbf{Y}$ là ước lượng LSE cho hàm vector $\mathbb{E}_\beta(\mathbf{D}'\mathbf{Y}) = \mathbf{D}'\mathbf{A}\beta = \mathbf{G}'\beta$ với $\mathbf{G}$ là ma trận cấp $m \times k$. Chúng ta sẽ khẳng định rằng $\mathbf{D}'\mathbf{Y}$ là chấp nhận được trong lớp các ước lượng không chệch của $\mathbf{G}'\beta$ nếu không tồn tại ước lượng không chệch $\mathbf{T} = \mathbf{T}(\mathbf{Y})$ thỏa mãn

$$\mathbb{E}_\beta(\mathbf{T} - \mathbf{G}'\beta)(\mathbf{T} - \mathbf{G}'\beta)' \leq \mathbb{E}_\beta(\mathbf{D}'\mathbf{Y} - \mathbf{G}'\beta)(\mathbf{D}'\mathbf{Y} - \mathbf{G}'\beta)', \beta \in U. \quad (3.4.6)$$

Hơn thế nữa, với ít nhất một $\beta \in U$, các ma trận bên trái và bên phải hoàn toàn phân biệt. Ta viết $\mathbf{A}_1 \leq \mathbf{A}_2$ nếu $\mathbf{A}_2 - \mathbf{A}_1$ là ma trận xác định không âm.

Định lý 3.4.3. *Giả sử $\mathbf{D}$ là ma trận mà không hàng nào của nó chứa toàn phần tử không và $\mathbf{B}$ là ma trận thỏa mãn các điều kiện xác định trong Định lý 3.4.2. Khi đó, tính chấp nhận được của $\mathbf{D}'\mathbf{Y}$ trong lớp các ước lượng không chệch của hàm tham ẩn vector $\mathbf{G}'\beta = \mathbb{E}_\beta(\mathbf{D}'\mathbf{Y})$ suy ra tính chuẩn của hàm phân phối F_j .*

Chứng minh. Tương tự các định lý ở trên, từ tính chấp nhận được của $\mathbf{D}'\mathbf{Y}$ ta suy ra $\mathbb{E}_0(\mathbf{D}'\mathbf{Y}|\mathbf{Z}'\mathbf{Y}) = 0$. Sử dụng Định lý 1.1.2 ta suy ra điều phải chứng minh. $\square$

Trong Định lý 3.4.1 - 3.4.3, ta đã nghiên cứu tính chấp nhận được của ước lượng LSE với giả thiết đã biết trước phân phối F_j của sai số ε_j (không phụ thuộc vào các tham ẩn chưa biết). Bây giờ ta sẽ bỏ giả thiết này và thu được các điều kiện trong những tính hướng khác mà ở đó tính chuẩn của hàm phân phối của sai số có thể được bảo đảm.

Như ở phần trước, chúng ta sẽ xét mô hình Gauss - Markov: $\mathbf{Y} = \mathbf{C}\beta + \varepsilon$, trong đó $\mathbb{E}_\beta \mathbf{Y} = \mathbf{C}\beta$ và ε là vectơ sai số. Cho các biến ngẫu nhiên

$Y_j - \mathbb{E}_\beta Y_j$ có cùng phân phối F_θ phụ thuộc vào tham số chưa biết $\theta \in \Theta$ và θ có thể phụ thuộc vào β . Các ước lượng LSE luôn tối ưu trong lớp các ước lượng không chệch tuyến tính theo biến $\mathbf{Y}$ của kỳ vọng toán học của chúng bất kể độ phức tạp của các tham ẩn. Bây giờ ta giả sử các ước lượng LSE là tối ưu, tức là nó có phương sai nhỏ nhất với mọi $\theta \in \Theta$ trong một lớp bất kỳ các ước lượng không chệch của tham ẩn. Trong trường hợp đó, ta kết luận được gì về phân phối F_θ ? Để trả lời câu hỏi này, ta xét một lớp đặc biệt các ước lượng không chệch được cho bởi Định lý 3.4.4 và 3.4.5, từ đó suy ra rằng trong trường hợp đặc biệt đó thì tính chất tối ưu của các ước lượng LSE trong lớp tất cả các ước lượng không chệch cùng với các điều kiện nào đó cho F_θ sẽ kéo theo tính chuẩn của F_θ .

Trước hết ta xét trường hợp hạng của ma trận $\mathbf{C}$ bằng 1.

Định lý 3.4.4. Cho hạng $\mathbf{C} = 1$ và $L = b_1 Y_1 + \cdots + b_n Y_n$ là ước lượng LSE của hàm tham ẩn $\mathbb{E}_\beta(\mathbf{b}'\mathbf{Y}) = \mathbf{b}'\mathbf{C}\beta = \mathbf{p}'\beta$. Giả sử thêm rằng $\int x^{2s} dF_\theta(x)$ tồn tại với mọi θ , s nguyên dương nào đó và $b_j \neq 0, \forall j$. Nếu L tối ưu trong lớp các ước lượng không chệch đa thức có bậc $\leq s$ của $\mathbf{p}'\beta$, và vector $\mathbf{b}' = (b_1, \dots, b_n)$ không tỉ lệ với vector có các thành phần là ± 1 , thì $(s+1)$ mô men đầu tiên của F_θ trùng với các mô men tương ứng của phân phối chuẩn nào đó.

Chứng minh. Vì hạng $\mathbf{C} = 1$ nên tồn tại $n-1$ hàm số

$$Z_j = c_{j1}Y_1 + \cdots + c_{jn}Y_n, \quad j = 1, \dots, n-1$$

thỏa mãn $\mathbb{E}_\beta(Z_j) = 0$, $\mathbb{E}_{\beta,\theta}(Z_i Z_j) = 0$ với $i \neq j$. Thực chất ta có thể lấy $C_j = (c_{j1}, \dots, c_{jn})'$ là $n-1$ vector đôi một trực giao và trực giao đối với không gian vector một chiều sinh bởi các vector cột của $\mathbf{C}$. Thống kê Z_j có thể được xem là ước lượng không chệch của không, hơn nữa $\mathbb{E}_{\beta,\theta} Z_j^2 < \infty$. Theo Bổ đề 2.1.1 ta kết luận rằng $\mathbb{E}_{\beta,\theta}(Z_j L) = 0$, khi đó $Z_j L$ là ước lượng không chệch của không với phương sai hữu hạn nếu $s \geq 2$ (trường hợp $s = 1$ tầm thường). Lại sử dụng Bổ đề 2.1.1 ta thu được

$$\mathbb{E}_{\beta,\theta}(Z_j L^r) = 0, \quad r = 1, \dots, s; \quad j = 1, \dots, n-1.$$

Xét hàm

$$\begin{aligned}\phi_j(t) &= \mathbb{E}\left[iZ_j \exp it(L - \mathbb{E}_\beta L)\right] \\ &= \left[c_{j1}\psi(b_1t) + \cdots + c_{jn}\psi(b_nt)\right]f(b_1t) \cdots f(b_nt)\end{aligned}\quad (3.4.7)$$

trong đó f là hàm đặc trưng tương ứng với phân phối F_θ (ta sẽ không chỉ ra sự phụ thuộc của f vào θ), $\psi = f'/f$ và (3.4.7) đúng khi $|t| < \delta$ với $\delta > 0$ nào đó. Vì F_θ có mô men hữu hạn cấp $2s$ nên ψ khả vi cấp s . Chú ý rằng đạo hàm cấp r của $\phi_j(t)$ tại điểm không chính là $-\mathbb{E}[Z_j(L - \mathbb{E}_\beta L)^r]$. Sử dụng các điều kiện $\mathbb{E}_\beta(Z_j) = 0$ và $\mathbb{E}_{\beta,\theta}(Z_j L^r) = 0, r = 1, \dots, s$, ta suy ra rằng $\mathbb{E}[Z_j(L - \mathbb{E}_\beta L)^r] = 0, r = 1, \dots, s$, tức là các đạo hàm cấp r của $\phi_j(t)$ tại 0 phải bằng 0. Lấy đạo hàm lần lượt từ bậc 1 đến bậc s của ϕ đồng thời sử dụng lần lượt tính chất $\phi_j^{(r)}(0) = 0, r = 1, \dots, s$, ta thu được

$$\left(\sum_1^n c_{ji}b_i^r\right)\kappa_{r+1}(\theta) = 0, \quad r \leq s, \quad j = 1, \dots, n-1 \quad (3.4.8)$$

trong đó $\kappa_{r+1}(\theta)$ là đạo hàm cấp r của hàm ψ tại 0, tức là đạo hàm cấp $r+1$ của hàm $\ln f$ tại 0. Giá trị này được gọi là bán bất biến thứ $r+1$ của F_θ . Từ (3.4.8) ta kết luận rằng có ít nhất một trong hai nhân tử bên vế trái phải bằng không. Vì $\mathbb{E}_{\beta,\theta}(Z_j L) = 0$ suy ra $\sum_1^n c_{ji}b_i = 0$ với $j = 1, \dots, n-1$. Nếu $\kappa_{r+1}(\theta) \neq 0$ thì $\sum_1^n c_{ji}b_i^r = 0, j = 1, \dots, n-1$. Như vậy các vectơ $(b_1, \dots, b_n)'$ và $(b_1^r, \dots, b_n^r)'$ cùng trực giao với hệ $n-1$ vectơ đôi một trực giao trong $\mathbb{R}^n$ nên hai vectơ này tỉ lệ với nhau, tức là $b_i^r = \lambda b_i, i = 1, \dots, n$, với λ nào đó. Với $r > 1$, đẳng thức này mâu thuẫn với giả thiết $\mathbf{b}' = (b_1, \dots, b_n)$ không tỉ lệ với vectơ mà các thành phần của nó là ± 1 , suy ra $\kappa_{r+1}(\theta) = 0, r = 2, \dots, s$. Như vậy $(s+1)$ bán bất biến đầu tiên của F_θ trùng với các bán bất biến tương ứng của phân phối chuẩn nào đó. Mà các bất biến có thể biểu diễn qua các mô men của F_θ , nên dễ dàng suy ra rằng $(s+1)$ mô men đầu tiên của F_θ trùng với các mô men tương ứng của phân phối chuẩn nào đó. Định lý được chứng minh. $\square$

Chúng ta lưu ý rằng nếu F_θ có mô men mọi cấp thì theo các điều kiện

trong Định lý 3.4.4 về các hệ số của ước lượng LSE: $L = b_1 Y_1 + \cdots + b_n Y_n$, tính chuẩn của F_θ suy ra từ tính tối ưu của L trong lớp tất cả các ước lượng không chệch của $\mathbf{p}'\beta$. Thật vậy, nếu F_θ có mô men mọi cấp, theo chứng minh của định lý trên ta suy ra rằng các bán bất biến của F_θ sẽ trùng với các bán bất biến của một phân phối chuẩn tức là khai triển Mac Lauren của hàm $\ln f$ sẽ trùng với khai triển hàm logarit của hàm đặc trưng của một phân phối chuẩn. Tức là hàm đặc trưng f sẽ trùng với hàm đặc trưng của một phân phối chuẩn tại lân cận của 0. Suy ra F_θ có phân phối chuẩn.

Bây giờ ta xét trường hợp tổng quát: hạng $\mathbf{C} = r > 1$. Cho $L_1, \dots, L_r$ là các ước lượng LSE của r hàm tham ẩn độc lập tuyến tính $\mathbf{p}'_1\beta, \dots, \mathbf{p}'_r\beta$. Tồn tại $n - r$ hàm tuyến tính dạng

$$Z_j = c_{j1}Y_1 + \cdots + c_{jn}Y_n, \quad j = 1, \dots, n - r$$

thỏa mãn $\mathbb{E}_\beta(Z_j) = 0, \mathbb{E}_{\beta, \theta}(Z_i Z_j) = 0$ với $i \neq j$. Xét dạng chính tắc của ma trận $\mathbf{C}$: $\begin{pmatrix} \mathbf{I}_r \\ \mathbf{B} \end{pmatrix}$. Ta gọi $\mathbf{C}$ là ma trận ngoại lệ nếu trong dạng chính tắc của nó mỗi hàng của ma trận $\mathbf{B}$ chứa không nhiều hơn một phần tử khác không và phần tử này (nếu tồn tại) phải bằng ± 1 . Chúng ta phát biểu một định lý trong trường hợp tổng quát:

Định lý 3.4.5. Cho hạng $\mathbf{C} = r$, ma trận $\mathbf{C}$ không ngoại lệ, và đối với số nguyên dương s nào đó, $\int x^{2s} dF_\theta < \infty$, với mọi θ . Nếu các ước lượng LSE $L_1, \dots, L_r$ tối ưu trong lớp biệt các ước lượng không chệch đa thức có bậc $\leq s$ là của $\mathbf{p}'_1\beta, \dots, \mathbf{p}'_r\beta$ tương ứng, thì $s + 1$ mô men đầu tiên của F_θ trùng với các mô men tương ứng của phân phối chuẩn nào đó.

Chứng minh. Tương tự Định lý 3.4.4. □

KẾT LUẬN

Người ta nói rằng phân phối chuẩn là sợi chỉ đỏ xuyên suốt các giáo trình Xác suất - Thống kê. Nói như vậy là chính xác, bởi lẽ rất nhiều kết quả trong Xác suất, Thống kê chỉ đúng cho phân phối chuẩn mà không đúng cho phân phối khác, và bởi lẽ trong thực tế có rất nhiều đại lượng ngẫu nhiên tuân theo luật chuẩn hoặc gần chuẩn.

Do vậy trong các bài toán đặc trưng phân phối xác suất, đặc trưng cho phân phối chuẩn chiếm một vị trí đặc biệt.

Ta có thể đặc trưng cho họ chuẩn bởi nhiều tính chất đặc trưng khác nhau dựa trên tính độc lập, tính cùng phân phối, tính hồi quy hằng số, ... giữa một thống kê đối với một thống kê khác.

Luận văn này tập trung trình bày các kết quả đặc trưng phân phối chuẩn dựa trên tính chấp nhận được và tính tối ưu của các ước lượng với hiểm dạng bình phương. Cụ thể, luận văn đã giải quyết những vấn đề sau:

- Đặc trưng của luật chuẩn thông qua tính chấp nhận được của các ước lượng tuyến tính tối ưu của tham số tịnh tiến đối với hàm tổn thất dạng bình phương.
- Đưa ra điều kiện để trung bình mẫu chấp nhận được trong một lớp con các ước lượng của tham số tịnh tiến. Ta xét hai lớp con đặc biệt. Một là lớp các ước lượng không chệch chỉ phụ thuộc vào trung bình mẫu và một thống kê tuyến tính khác. Hai là lớp các ước lượng đa thức không chệch.
- Đặc trưng của luật chuẩn bởi tính tối ưu của hàm của ước lượng tuyến tính tối ưu $\hat{L} = \sum c_j^0 X_j$.
- Đặc trưng của luật chuẩn bởi tính chấp nhận được và tính tối ưu của ước lượng bình phương bé nhất trong mô hình Gauss-Markov.

Trong khuôn khổ của Luận văn, tác giả chỉ trình bày những kết quả về đặc trưng của phân phối chuẩn trong trường hợp các quan trắc là độc lập và hàm tổn thất dạng bình phương. Một số vấn đề lý thú và đáng quan tâm khác là trường hợp các quan trắc phụ thuộc và trường hợp hàm tổn thất có một dạng khác chẳng hạn như dạng trị tuyệt đối. Hơn nữa, chúng ta có thể quan tâm đến đặc trưng của một số phân phối khác. Đó có thể là vấn đề nghiên cứu của chúng ta trong thời gian tới.

TÀI LIỆU THAM KHẢO

- [1] Đào Hữu Hồ (2010), *Xác suất - Thống kê*, NXB ĐHQGHN, Tái bản lần thứ XII.
- [2] Đào Hữu Hồ, Hoàng Hữu Như, Nguyễn Văn Hữu (2004), *Thống kê Toán học*, NXB ĐHQGHN.
- [3] Nguyễn Hữu Dư, Nguyễn Văn Hữu (2003), *Phân tích Thống kê và Dự báo*, NXB ĐHQGHN.
- [4] Blackwell D. (1947), Conditional Expectation and Unbiased Sequential Estimation, *Ann. Math. Statist.* 18, pp. 105-110.
- [5] Hajek J. (1958), On a Property of Normal Distribution of an Arbitrary Stochastic Process, *Czechoslovak. Math. Journal* 8, pp. 610-618.
- [6] Kagan A. M. (1966), On the Estimation Theory of Location Parameters, *Sankhya, Ser. A* 28, pp. 335-352.
- [7] Kagan A. M. (1968), Theory of Estimation for Families with Shift-, Scale- and Exponential Parameters, *Trudy Matem. Inst. Steklov. AN SSSR* 104, pp. 19-87.
- [8] Kagan A. M., Klebanov L. B., Zinger A. A. (1969), The Sample Mean as an Estimator for a Shift Parameter under Loss Functions other than Quadratic, *DAN SSSR* 189, pp. 29-30.

- [9] Kagan A. M., Linnik Y. V., Rao C. R. (1965), *On a Characterization of the Normal Law Based on a Property of the Sample Average*, *Sankhya*, Ser. A 27, pp. 405-406.
- [10] Kagan A. M., Linnik Y. V., Rao C. R. (1973), *Characterization Problems in Mathematical Statistics*, by John Willey & Sons, New York.
- [11] Kakutani S. (1950), On the Equivalence of Infinite Product Measures, *Ann. Math.* 49, pp. 303-326.
- [12] Kolmogorov A. N. (1950), Unbiased Estimators, *Izvestiia AN SSSR Ser. Matem.* 14, pp. 303-326.
- [13] Lehmann E. L. (1959), *Testing of Statistical Hypotheses*, John Willey, New York.
- [14] Lehmann E. L., George Casella (1998), *Theory of Point Estimation*, Second Edition.
- [15] Rao C. R. (1945), Information and Accuracy Attainable in the Estimation of Statistical Parameters, *Bull. Calcutta Math. Soc.* 37, pp. 81-91.
- [16] Rao C. R. (1948), Sufficient Statistics and Minimum Variance Estimates, *Proc. Camb. Phil. Soc.* 45, pp. 215-218.
- [17] Rao C. R. (1952), Some Theorems on Minimum Variance Estimation, *Sankhya*, Ser. A 12, pp. 27-42.
- [18] Rao C. R. (1973), *Linear Statistical Inference and Its Applications*, John Wiley, New York, Second Edition.
- [19] Stein C. (1959), The Admissibility of Pitman's Estimator for a Single Location Parameter, *Ann. Math. Statist.* 30, pp. 970-978.