บทที่ 2
ทฤษฎีที่เกี่ยวข้องกับการวิจัย

ในบทนี้จะกล่าวถึงการประมาณค่าพารามิเตอร์แสดงตำแหน่ง และการประมาณค่าพารามิเตอร์แสดงสเกล เมื่อประชากรมีการแจกแจงแบบปกติที่มีค่าผิดปกติ โดยตัวประมาณที่นำมาเปรียบเทียบนั้นได้แก่ ตัวประมาณค่าที่มีความแกร่ง ตัวประมาณค่าที่ได้จากวิธีบูตสแตรป และตัวประมาณค่าที่ไม่เอนเอียง ซึ่งรายละเอียดมีดังต่อไปนี้

2.1 ตัวประมาณค่าพารามิเตอร์แสดงตำแหน่ง

2.1.1 ตัวประมาณค่าที่มีความแกร่งสำหรับพารามิเตอร์แสดงตำแหน่ง

ตัวประมาณค่าที่มีความแกร่งสำหรับพารามิเตอร์แสดงตำแหน่งในการวิจัยครั้งนี้ ผู้วิจัยได้พิจารณาตัวประมาณค่าที่มาจากพื้นฐานวิธี Gaussiab Skip ซึ่งมีขั้นตอนในการคำนวณดังนี้

กำหนดให้

$$U_i = \frac{(X_i - m)}{cMAD}$$

$$d_i = |X_i - m|$$

เมื่อ m แทน ค่ามัธยฐานของข้อมูล X_i

c แทน ค่าคงที่ที่เป็นบวก (c = 10)

MAD แทน ค่ามัธยฐานของข้อมูล d_i

การคำนวณหาค่า c นั้นจะพิจารณาจากค่าคงที่ที่เป็นบวกแล้วทำให้ความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด

ดังนั้นตัวประมาณค่าที่มีความแกร่งสำหรับพารามิเตอร์แสดงตำแหน่งคือ

$$M_{bi}(X) = \frac{\sum_{i=1}^{n} X_i W_{bi}(U_i)}{\sum_{i=1}^{n} W_{bi}(U_i)}$$

เมื่อ ฟังก์ชันถ่วงน้ำหนักคือ

$$W_{bi}(U_i) = (1-U_i^2)^2 \quad ; \quad |U_i| < 1$$
$$= 0 \quad ; \quad U_i \text{ เป็นค่าอื่น ๆ}$$

<u>ตัวอย่างที่ 1</u>

สมมติว่าค่าข้อมูลที่เลือกมาได้มีค่าเป็น 2.8685 , -0.3624 , -1.17 , 0.9182 , 4.3596 , 2.6195 , 1.4308 , 4.1761 , 10

จากข้อมูลของตัวอย่างพิจารณาได้ว่า ข้อมูลมีค่าผิดปกติอยู่ในระดับไม่รุนแรง (พิจารณาตามเกณฑ์ในหน้า 3) โดยที่ค่าข้อมูลเท่ากับ 10 คือค่าผิดปกติของข้อมูลชุดนี้

การคำนวณหาค่า $M_{bi}(X)$ สามารถคำนวณได้ดังนี้

จากข้อมูลที่ได้มาหาค่า m (มัธยฐานของข้อมูล) โดย

1. เรียงข้อมูลจากน้อยไปหามาก

   -1.17 , -0.3624 , 0.9182 , 1.4308 , 2.6195 , 2.8685 , 4.1761 , 4.2829 , 4.3596 , 10

2. ค่ามัธยฐาน คือ

$$m = \frac{X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}}{2}$$

เมื่อ

$X_{(\frac{n}{2})}$ แทน ค่าข้อมูลตัวที่ $\frac{n}{2}$

$X_{(\frac{n}{2}+1)}$ แทน ค่าข้อมูลตัวที่ $\frac{n}{2}+1$

$$\therefore m = \frac{2.6195 + 2.8685}{2} = 2.744$$

ข้อมูลตัวที่	X_i	d_i	U_i	$W_{bi}(U_i)$	$X_i W_{bi}(U_i)$
1	2.8685	0.1245	0.0079	0.99988	2.8682
2	-0.3624	3.1064	-0.1969	0.92396	-0.3348
3	4.2829	1.5389	0.0976	0.98104	4.2017
4	-1.17	3.914	-0.2481	0.88068	-1.0304
5	0.9182	1.8258	-0.1158	0.97336	0.8937
6	4.3596	1.6156	0.1024	0.97914	4.2687
7	2.6195	0.1245	-0.0079	0.99988	2.6192
8	1.4308	1.2679	-0.0833	0.98617	1.411
9	4.1761	1.4321	0.0908	0.98358	4.1075
10	10	7.256	0.4600	0.62157	6.2157
รวม				9.32926	25.2205

เมื่อ MAD คือ ค่ามัธยฐานของข้อมูล d_i

ซึ่ง MAD = 1.5773

$$\therefore M_{bi}(X) = \frac{\sum_{i=1}^{n} X_i W_{bi}(U_i)}{\sum_{i=1}^{n} W_{bi}(U_i)}$$

$$= \frac{25.2205}{9.32926} = 2.7034$$

โดยที่ $M_{bi}(X)$ คือ ค่าตัวประมาณพารามิเตอร์แสดงตำแหน่ง

2.1.2 ตัวประมาณค่าที่ไม่เอนเอียงสำหรับพารามิเตอร์แสดงตำแหน่ง

ตัวประมาณค่าที่ไม่เอนเอียงสำหรับพารามิเตอร์แสดงตำแหน่งนั้นมีอยู่หลายตัวด้วยกัน เช่น ค่ามัธยฐาน ซึ่งในการวิจัยครั้งนี้ได้เลือกตัวประมาณค่าที่ได้จากวิธีของพิทมัน (Method of Pitman) ซึ่งมีรายละเอียดดังนี้

ให้ $X_1, X_2, \ldots, X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีฟังก์ชันความหนาแน่น $f(x;\theta)$ โดยที่ θ เป็นพารามิเตอร์ที่เราไม่ทราบค่าแต่วัดโดยใช้มาตราเดียวกับการวัด $X_1, X_2, \ldots, X_n$

<u>บทนิยามที่ 1</u> เมื่อ $\{ f(x;\theta) , \theta \in \Omega\}$ เป็นแฟมมิลี่ของฟังก์ชันความหนาแน่นที่ขึ้นอยู่กับพารามิเตอร์ θ และปริภูมิพารามิเตอร์ Ω เป็นเส้นจำนวนจริง θ เป็นพารามิเตอร์แสดงตำแหน่ง ก็ต่อเมื่อเขียน $f(x;\theta)$ เป็นฟังก์ชันของ $x - \theta$ ได้ นั่นคือ

$$f(x;\theta) = g(x - \theta)$$

เมื่อ $g(x - \theta)$ เป็นฟังก์ชันความหนาแน่น

หาก X เป็นตัวแปรเชิงสุ่มที่มีฟังก์ชันความหนาแน่น $f(x;\theta)$ เราจะพบว่า θ เป็นพารามิเตอร์แสดงตำแหน่งของ $f(x;\theta)$ ก็ต่อเมื่อฟังก์ชันความหนาแน่น $g(x - \theta)$ ไม่ขึ้นอยู่กับ θ

<u>บทนิยามที่ 2</u> ตัวประมาณ $t(X_1, X_2, \ldots, X_n)$ ของฟังก์ชัน $\tau(\theta)$ ของพารามิเตอร์ θ เป็นตัวประมาณที่ไม่เปลี่ยนแปลงตำแหน่ง (Location invariant estimator) ก็ต่อเมื่อ

$$t(X_1+c, \ldots, X_n+c) = t(X_1, \ldots, X_n) + c$$

ทุกๆ ค่าของ $X_1, X_2, \ldots, X_n$ และทุกค่า c เช่น

$$t(X_1, X_2, \ldots, X_n) = \bar{X}$$

$$t(X_1 + c, \ldots, X_n + c) = \frac{\sum_{i=1}^{n}(X_i + c)}{n} = \bar{X} + c$$

$$= t(X_1, \ldots, X_n) + c$$

ดังนั้น $\tilde{X}$ เป็นตัวประมาณที่ไม่เปลี่ยนตำแหน่ง และ $\tilde{X}$ เป็นตัวประมาณที่ไม่เอนเอียงซึ่งนอกจาก $\bar{X}$ จะเป็นตัวประมาณที่ไม่เอนเอียงสำหรับพารามิเตอร์แสดงตำแหน่งยังมีค่ามัธยฐานของข้อมูลอีกตัวหนึ่งที่นำมาพิจารณา ซึ่งมีขั้นตอนการคำนวณดังนี้

ขั้นตอนที่ 1 เรียงลำดับข้อมูลจากน้อยไปมาก

ขั้นตอนที่ 2 ค่ามัธยฐานคือ

$$\tilde{X} = X_{\left(\frac{n+1}{2}\right)} \qquad \text{ถ้า n เป็นเลขคี่}$$

$$\tilde{X} = \frac{X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}}{2} \qquad \text{ถ้า n เป็นเลขคู่}$$

เมื่อ $X_{\left(\frac{n+1}{2}\right)}$ แทน ค่าข้อมูลตำแหน่งที่ $\left(\frac{n+1}{2}\right)$, n เป็นเลขคี่

$X_{\left(\frac{n}{2}\right)}$ แทน ค่าข้อมูลตำแหน่งที่ $\left(\frac{n}{2}\right)$, n เป็นเลขคู่

$X_{\left(\frac{n}{2}+1\right)}$ แทน ค่าข้อมูลตำแหน่งที่ $\left(\frac{n}{2}+1\right)$, n เป็นเลขคู่

2.1.3 ตัวประมาณค่าที่ได้จากวิธีบูตสแตรปสำหรับพารามิเตอร์แสดงตำแหน่ง

วิธีบูตสแตรป (Boot strap)

วิธีการหาตัวประมาณของพารามิเตอร์วิธีนี้เป็นวิธีที่เสนอขึ้นโดย แบรดเลย์ เอฟรอน (Bradley Efron) ในปี ค.ศ. 1979 โดยมีหลักเกณฑ์ดังนี้คือ จากข้อมูลที่เก็บรวบรวมมาจะทำการสุ่มตัวอย่างแบบใส่คืน (with replacement) โดยสุ่มเป็นจำนวนเท่ากับจำนวนขนาดตัวอย่างหรือข้อมูลที่มีอยู่ เพื่อสร้างข้อมูลชุดใหม่แล้วนำมาใช้ในการประมาณค่าพารามิเตอร์ที่สนใจ

สำหรับการหาตัวประมาณด้วยวิธีนี้ มีขั้นตอนการทำงานดังนี้

1. สร้างเลขสุ่มที่มีค่าอยู่ระหว่าง 0 กับ 1 เพื่อนำไปใช้ในการสุ่มตัวอย่างแบบใส่คืน

2. จากตัวอย่างที่สุ่มได้แต่ละชุดจะนำมาหาค่าประมาณของพารามิเตอร์ซึ่งรายละเอียดของวิธีการสุ่มตัวอย่างแบบใส่คืน และรายละเอียดต่าง ๆ เสนอไว้ในภาคผนวก

การหาตัวประมาณที่ได้จากวิธีบูตสแตรปมีขั้นตอนดังต่อไปนี้

(1) ทำการสร้างข้อมูลให้มีการแจกแจงตามที่ต้องการศึกษา $X_i \sim F$; $i = 1, 2, \ldots, n$ เมื่อ F เป็นการแจกแจงที่ต้องการศึกษา

(2) จากค่า $x_1, \ldots, x_n$ ทำการสุ่มตัวอย่างขนาด n แบบใส่คืนได้ $X_1^*, \ldots, X_n^*$ เมื่อ X_j^* คือตัวอย่างที่สุ่มได้ตัวที่ j จากข้อมูล $x_1, \ldots, x_n$

(3) นำค่าข้อมูล $X_1^*, \ldots, X_n^*$ ไปคำนวณหาค่าตัวประมาณคือ

(3.1) ค่าเฉลี่ยของข้อมูล $X_1^*, \ldots, X_n^*$ คือ

$$\bar{X}^* = \frac{\sum_{i=1}^{n} X_i^*}{n}$$

(3.2) ค่ามัธยฐานของข้อมูล $X_1^*, \ldots, X_n^*$ คือ

ทำการเรียงข้อมูล $X_1^*, \ldots, X_n^*$ จากน้อยไปมาก

ค่ามัธยฐานของข้อมูล $X_1^*, \ldots, X_n^*$ คือ

$$\tilde{X}^* = X^*_{\left(\frac{n+1}{2}\right)} \qquad \text{ถ้า n เป็นเลขคี่}$$

$$\tilde{X}^* = \frac{X^*_{(\frac{n}{2})} + X^*_{(\frac{n}{2}+1)}}{2} \qquad \text{ถ้า n เป็นเลขคู่}$$

เมื่อ $X^*_{(\frac{n+1}{2})}$ แทน ค่าข้อมูลตัวที่ $\frac{n+1}{2}$, n เป็นเลขคี่

$X^*_{(\frac{n}{2})}$ แทน ค่าข้อมูลตัวที่ $\frac{n}{2}$, n เป็นเลขคู่

$X^*_{(\frac{n}{2}+1)}$ แทน ค่าข้อมูลตัวที่ $\frac{n}{2}+1$, n เป็นเลขคู่

(4) กระทำตามขั้นตอนที่ (2) - (3) ซ้ำกันเท่ากับจำนวนครั้งที่ต้องการทำบูตสแตรป

(5) หาค่าเฉลี่ยของค่าประมาณที่ได้จากขั้นตอนที่ (3) ดังนั้นตัวประมาณที่ได้จากวิธีบูตสแตรป คือ

$$BMEAN = \frac{\sum_{j=1}^{B} \bar{X}_j^*}{B}$$

$$BMEDIAN = \frac{\sum_{j=1}^{B} \tilde{X}_j^*}{B}$$

เมื่อ B แทน จำนวนครั้งของการทำบูตสแตรป

การหาจำนวนครั้งของการทำบูตสแตรป

การหาจำนวนครั้งของการทำบูตสแตรปมีขั้นตอนดังต่อไปนี้

1. กำหนดจำนวนครั้งของการทำบูตสแตรปเริ่มต้นไว้ 100 ครั้ง และ 105 ครั้ง

2. ทำการประมาณค่า BMEAN, BMEDIAN ด้วยวิธีการบูตสแตรป

3. พิจารณาดูว่าค่าประมาณมีค่าคงที่แล้วหรือยัง โดยการเปรียบเทียบค่าตัวประมาณที่ได้เมื่อจำนวนครั้งของการทำบูตสแตรปเท่ากับ 100 ครั้ง กับ 105 ครั้ง ถ้าผลต่างระหว่างค่าตัวประมาณเมื่อจำนวนครั้งของการทำบูตสแตรปเท่ากับ 100 ครั้ง กับ 105 ครั้ง เกิน 1 % จะถือว่าค่าประมาณที่ได้ยังไม่คงที่ ซึ่งถ้ายังไม่คงที่ไปยังขั้นตอนที่ 4 แต่ถ้าไม่เกิน 1 % จะเลือกจำนวนครั้งสุดท้ายเป็นจำนวนครั้งของการทำบูตสแตรป

4. เพิ่มจำนวนครั้งของการทำบูตสแตรปโดยทำการเพิ่มครั้งละ 5 รอบ แล้วกลับไปยังขั้นตอนที่ 2

2.2 ตัวประมาณค่าพารามิเตอร์แสดงสเกล

2.2.1 ตัวประมาณที่มีความแกร่งสำหรับพารามิเตอร์แสดงสเกล ในการวิจัยครั้งนี้ ผู้วิจัยได้พิจารณาตัวประมาณค่าสำหรับพารามิเตอร์แสดงสเกล ซึ่งมีรายละเอียดดังต่อไปนี้

ถ้า $X_i \sim \text{iid } N(0,\sigma^2)$

และ $\bar{X} = \dfrac{\sum_{i=1}^{n} X_i}{n}$ แล้ว $n \rightarrow \infty$ จะได้ว่า

$$(nVar(\bar{X}))^{\frac{1}{2}} \longrightarrow \sigma$$

ดังนั้นความแปรปรวนของตัวประมาณค่าพารามิเตอร์แสดงตำแหน่งเพียงพอที่จะเป็นตัวประมาณค่าพารามิเตอร์แสดงสเกล จากตัวประมาณค่าพารามิเตอร์แสดงตำแหน่งทำให้ได้ตัวประมาณค่าพารามิเตอร์แสดงสเกลดังนี้

กำหนดให้

$$U_i = \frac{(X_i - m)}{cMAD}$$

$$d_i = |X_i - m|$$

เมื่อ m แทน ค่ามัธยฐานของข้อมูล X_i

c แทน ค่าคงที่ที่เป็นบวก

MAD แทน ค่ามัธยฐานของข้อมูล d_i

$$S_{mb,c} = \frac{n}{(n-1)^{1/2}} \frac{\left[\sum_{|U_i|<1}((X_i - m)^2 W_{bi}(U_i))^2\right]^{1/2}}{\left|\sum_{|U_i|<1} W_{bi}(U_i)\right|}$$

เมื่อ $S_{mb,c}$ แทน ตัวประมาณค่าพารามิเตอร์แสดงสเกล

m แทน ค่ามัธยฐานของข้อมูล

เมื่อ ฟังก์ชันถ่วงน้ำหนัก คือ

$$W_{bi}(U_i) = (1-U_i^2)^2 \quad ; \quad |U_i| < 1$$
$$= 0 \quad ; \quad U_i \text{ เป็นค่าอื่น ๆ}$$

ตัวอย่างที่ 2 จากข้อมูลในตัวอย่างที่ 1 จะได้ว่า

m = 2.744 , n = 10

MAD = 1.5773

ข้อมูลตัวที่	X_i	$(X_i - m)^2$	U_i	$W_{bi}(U_i)$	$((X_i - m)^2 W_{bi}(U_i))^2$
1	2.8685	0.0155	0.0316	0.9980	0.00024
2	-0.3624	9.6497	-0.7878	0.1439	1.9282
3	4.2829	2.3682	0.3903	0.7185	2.8954
4	-1.17	15.3194	-0.9926	0.0002	0.000012
5	0.9182	3.3335	-0.4630	0.6172	4.2329
6	4.3596	2.6102	0.4097	0.6925	3.2674
7	2.6195	0.0155	-0.0316	0.9980	0.00024
8	1.4308	1.7245	-0.3330	0.7905	1.8583
9	4.1761	2.0509	0.3632	0.7536	2.3889
10	10	52.6495	1.8401	0	0
รวม					16.5716

$$\therefore S_{mb,c} = \frac{n}{(n-1)^{\frac{1}{2}}} \frac{\left[\sum_{|U_i|<1} ((X_i - m)^2 W_{bi}(U_i))^2\right]^{\frac{1}{2}}}{\left|\sum_{|U_i|<1} W_{bi}(U_i)\right|}$$

$$= \frac{10}{(10-1)^{\frac{1}{2}}} \frac{(16.5716)^{\frac{1}{2}}}{5.7124}$$

$$= 2.624$$

ตัวประมาณค่าพารามิเตอร์แสดงสเกลเมื่อฟังก์ชันถ่วงน้ำหนักอยู่ในรูป

$$\Psi(U_i) = \sin(U_i) \quad ; \quad |U_i| \le \pi$$

$$= 0 \quad ; \quad U_i \text{ เป็นค่าอื่น ๆ}$$

ดังนั้นตัวประมาณค่าพารามิเตอร์แสดงสเกล คือ

$$S_{msi} = \frac{n * 2.1 * MAD}{(n-1)^{\frac{1}{2}}} \tan gent^{-1}\left(\frac{\left[\sum_{|U_i|\le\pi} \sin^2(U_i)\right]^{\frac{1}{2}}}{\left|\sum_{|U_i|\le\pi} \cos(U_i)\right|}\right)$$

กำหนดให้ $U_i = \dfrac{(X_i - m)}{c\,MAD}$

$d_i = |X_i - m|$

เมื่อ m แทน ค่ามัธยฐานของข้อมูล X_i

c แทน ค่าคงที่ที่เป็นบวก (c = 2.1)

MAD แทน ค่ามัธยฐานของข้อมูล d_i

<u>ตัวอย่างที่ 3</u> จากข้อมูลในตัวอย่าง จะได้ว่า

m = 2.744 , n = 10 , MAD = 1.5773

ข้อมูลตัวที่	X_i	U_i	$\sin^2(U_i)$	$\cos(U_i)$
1	2.8685	0.0376	0.0000004	0.99999
2	-0.3624	-0.9378	0.00027	0.99987
3	4.2829	0.4646	0.000066	0.99997
4	-1.17	-1.1816	0.00043	0.99979
5	0.9182	-0.5512	0.000093	0.99995
6	4.3595	0.4877	0.000072	0.99996
7	2.6195	-0.0376	0.0000004	0.99999
8	1.4308	-0.3965	0.000048	0.99998
9	4.1761	0.4324	0.000057	0.99997
10	10	2.1906	0.00146	0.99927
รวม			0.0024968	9.99874

$$S_{msi} = \frac{n * 2.1 * MAD}{(n-1)^{\frac{1}{2}}} \tan gent^{-1}\left(\frac{\left[\sum_{|U_i| \le \pi} \sin^2(U_i) \right]^{\frac{1}{2}}}{\left| \sum_{|U_i| \le \pi} \cos(U_i) \right|} \right)$$

$$= \frac{10 \quad 2.1 \times 1.5773}{(10-1)^{\frac{1}{2}}} \tan gent^{-1}\left(\frac{0.0024968^{\frac{1}{2}}}{9.99874} \right)$$

$$= 3.1613$$

2.2.2 ตัวประมาณที่ไม่เอนเอียงสำหรับพารามิเตอร์แสดงสเกล

สำหรับตัวประมาณที่ไม่เอนเอียงของพารามิเตอร์แสดงสเกล มีรายละเอียดดังนี้

กำหนดให้ $X_1, X_2, \ldots, X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีการแจกแจงแบบปกติ และ

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n}$$

ดังนั้น $\frac{nS^2}{\sigma^2} \sim \chi^2_{(n-1)}$ โดยมีฟังก์ชันความหนาแน่นอยู่ในรูป

$$f(x) = \frac{1}{\Gamma\left(\frac{n-1}{2}\right)2^{\frac{n-1}{2}}} x^{\frac{n-1}{2}-1} e^{\frac{x}{2}} \quad ; \quad x \geq 0$$

ด้วยวิธีการแปลงจาก $\frac{nS^2}{\sigma^2}$ ไปสู่ S^2 เราจะได้ ฟังก์ชันความหนาแน่นของ S^2 อยู่ในรูป

$$g(s^2) = \frac{1}{\Gamma\left(\frac{n-1}{2}\right)}\left(\frac{n}{2\sigma^2}\right)^{\frac{n-1}{2}} e^{\frac{-ns^2}{2\sigma^2}} (s^2)^{\frac{n-3}{2}} \quad ; \quad s^2 \geq 0$$

และ S มีรูปแบบฟังก์ชันความหนาแน่นอยู่ในรูป

$$h(s) = \frac{1}{\Gamma\left(\frac{n-1}{2}\right)}\left(\frac{1}{2^{\frac{n-3}{2}}}\right)\left(\frac{n}{\sigma^2}\right)^{\frac{n-1}{2}} e^{\frac{-ns^2}{2\sigma^2}} s^{n-2} \quad ; \quad s \geq 0$$

$$E(S) = \int_0^\infty s\, h(s)\, ds$$

$$= \frac{1}{\Gamma\left(\frac{n-1}{2}\right)} \frac{1}{2^{\frac{n-3}{2}}}\left(\frac{n}{\sigma^2}\right)^{\frac{n-1}{2}} \int_0^\infty e^{\frac{-ns^2}{2\sigma^2}} s^{n-1}\, ds$$

$$= \frac{1}{\Gamma\left(\frac{n-1}{2}\right)} \frac{1}{2^{\frac{n-3}{2}}}\left(\frac{n}{\sigma^2}\right)^{\frac{n-1}{2}} \int_0^\infty e^{-t} \left(\frac{2}{n}\right)^{\frac{n-1}{2}} .\sigma^{n-1}.t^{\frac{n-1}{2}}.\left(\frac{2}{n}\right)^{\frac{1}{2}} \frac{\sigma}{2}.t^{\frac{-1}{2}}\, dt$$

เมื่อ $t = \frac{ns^2}{2\sigma^2}$

$$= \frac{1}{\Gamma\left(\frac{n-1}{2}\right)} \frac{1}{2^{\frac{n-3}{2}}}\left(\frac{n}{\sigma^2}\right)^{\frac{n-1}{2}} \left(\frac{2}{n}\right)^{\frac{n}{2}} \frac{\sigma^n}{2} \int_0^\infty e^{-t}.t^{\frac{n-2}{2}}\, dt$$

$$= \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\left(\frac{2}{n}\right)^{\frac{1}{2}}\sigma$$

ดังนั้นตัวประมาณค่าที่ไม่เอนเอียงของพารามิเตอร์แสดงสเกล คือ

$$UNSD = \hat{\sigma} = \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{2}}$$

ตัวประมาณที่เอนเอียงสำหรับพารามิเตอร์แสดงสเกล

สำหรับตัวประมาณค่าที่เอนเอียงนั้นจะพิจารณาจากส่วนเบี่ยงเบนมาตรฐานของข้อมูล สามารถคำนวณได้จาก

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}}$$

2.2.3 ตัวประมาณที่ได้จากวิธีบูตสแตรปสำหรับพารามิเตอร์แสดงสเกล

สำหรับหลักการของวิธีบูตสแตรปนั้นจะใช้หลักการเดียวกันกับการประมาณค่าพารามิเตอร์แสดงตำแหน่ง

การประมาณค่าพารามิเตอร์แสดงสเกลด้วยวิธีบูตสแตรป มีขั้นตอนดังนี้

(1) ทำการสร้างข้อมูลให้มีการแจกแจงตามที่ต้องการศึกษา $X_i \sim F$; $i = 1, 2, \ldots, n$ เมื่อ F เป็นการแจกแจงที่ต้องการศึกษา

(2) จากค่า $X_1, \ldots, X_n$ ทำการสุ่มตัวอย่างขนาด n แบบใส่คืนได้ $X_1^*, \ldots, X_n^*$ เมื่อ X_j^* คือตัวอย่างที่สุ่มได้ตัวที่ j จากข้อมูล $X_1, \ldots, X_n$

(3) นำค่าข้อมูล $X_1^*, \ldots, X_n^*$ ไปคำนวณหาค่าตัวประมาณคือ

(3.1) ตัวประมาณที่เอนเอียงคือ

$$SD^* = \sqrt{\frac{\sum_{i=1}^{n}(X_i^* - \overline{X}^*)^2}{n-1}}$$

(3.2) ตัวประมาณที่ไม่เอนเอียงคือ

$$UNSD^* = \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\sqrt{\frac{\sum_{i=1}^{n}(X_i^* - \overline{X}^*)^2}{2}}$$

เมื่อ

$$\overline{X}^* = \frac{\sum_{i=1}^{n} X_i^*}{n}$$

$$\Gamma(n) = (n-1)!$$

$$\Gamma(\frac{1}{2}) = \sqrt{\pi}$$

(4) กระทำตามขั้นตอนที่ (2) - (3) ซ้ำกันเท่ากับจำนวนครั้งที่ต้องการทำบูตสแตรป

(5) หาค่าเฉลี่ยของค่าประมาณที่ได้จากขั้นตอนที่ (3) ดังนั้นตัวประมาณที่ได้จากวิธีบูตสแตรป คือ

$$BSD = \frac{\sum_{j=1}^{B} SD^*}{B}$$

$$BUNSD = \frac{\sum_{j=1}^{B} UNSD^*}{B}$$

เมื่อ B แทน จำนวนครั้งของการทำบูตสแตรป ซึ่งเท่ากับจำนวนครั้งในการบูตสแตรปสำหรับตัวประมาณค่าพารามิเตอร์แสดงตำแหน่งด้วยวิธีบูตสแตรป

ต้นฉบับ หน้าขาดหาย