บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการวิจัยครั้งนี้ได้ทำการสุ่มตัวอย่างอย่างง่าย และ การสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง เพื่อเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าเฉลี่ย 10 วิธี ภายใต้แผนการสุ่มตัวอย่างดังกล่าวนี้ ซึ่งผู้วิจัยได้ศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้อง ดังนี้

1. วิธีการสุ่มตัวอย่าง

    1.1 การสุ่มตัวอย่างอย่างง่าย

    1.2 การสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง

2. โครงสร้างของตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง

    2.1 สถิติลำดับ

    2.2 การแจกแจงของสถิติลำดับที่ 1 และที่ n

3. การแจกแจงของประชากรที่ใช้ในงานวิจัย

    3.1 การแจกแจงแบบปกติ

    3.2 การแจกแจงแบบยูนิฟอร์ม

    3.3 การแจกแจงแบบแกมมา

4. วิธีการประมาณค่า

    4.1 การประมาณค่าแบบง่าย (Simple Estimation)

        4.1.1 การประมาณค่าเฉลี่ยแบบง่าย ภายใต้แผนการสุ่มตัวอย่างอย่างง่าย

        4.1.2 การประมาณค่าเฉลี่ยแบบง่าย ภายใต้แผนการสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง

    4.2 การประมาณค่าด้วยตัวแปรที่เกี่ยวข้อง (Composite Estimation)

        4.2.1 การประมาณพารามิเตอร์ โดยวิธีการประมาณค่าเฉลี่ยแบบอัตราส่วน ภายใต้แผนการสุ่มตัวอย่างอย่างง่าย

        4.2.2 การประมาณพารามิเตอร์ด้วย โดยวิธีการประมาณค่าเฉลี่ยแบบถดถอย ภายใต้แผนการสุ่มตัวอย่างอย่างง่าย

        4.2.3 การประมาณพารามิเตอร์ โดยวิธีการประมาณค่าเฉลี่ยแบบอัตราส่วน ของ Samawi and Muttlak ภายใต้แผนการสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง

4.2.4 การประมาณพารามิเตอร์ โดยวิธีการประมาณค่าเฉลี่ยแบบอัตราส่วนของ Kadilar ภายใต้แผนการสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง

4.2.5 การประมาณพารามิเตอร์ โดยวิธีการประมาณค่าเฉลี่ยแบบถดถอยของ Yu and Lam ภายใต้แผนการสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง

5. งานวิจัยที่เกี่ยวข้อง

วิธีการสุ่มตัวอย่าง

1. การสุ่มตัวอย่างอย่างง่าย

แผนการสุ่มตัวอย่างอย่างง่ายเป็นแผนการเลือกหน่วยตัวอย่างขนาด n หน่วย จากประชากรจำกัดขนาด N หน่วยที่ทุก ๆ หนึ่งในชุดตัวอย่างที่เป็นไปได้จากชุดตัวอย่างที่แตกต่างกัน ${}^{N}C_{n}$ ชุดตัวอย่างมีโอกาสที่จะได้รับการคัดเลือกมาเป็นตัวอย่างในการศึกษาเท่า ๆ กัน ด้วยการกำหนดหมายเลขของหน่วยประชากร และจะเลือกหน่วยตัวอย่างทีละหน่วยจนครบตามขนาดตัวอย่างที่ต้องการศึกษา โดยวิธีการจับฉลาก ตารางเลขสุ่ม หรือการใช้คอมพิวเตอร์สุ่มตัวเลข เป็นต้น ซึ่งมีโอกาสถูกเลือกเท่า ๆ กัน ในกรณีที่เลือกหน่วยตัวอย่างด้วยความน่าจะเป็นเท่ากัน มี 2 กรณี คือ การสุ่มตัวอย่างแบบไม่แทนที่ และการสุ่มตัวอย่างแบบแทนที่ ซึ่งงานวิจัยนี้ทำการสุ่มตัวอย่างแบบไม่แทนที่เท่านั้น

กรณีสุ่มตัวอย่างแบบไม่แทนที่ (Sampling without Replacement: SRSWOR)

สำหรับการคัดเลือกหน่วยตัวอย่างแรกด้วยความน่าจะเป็นเท่ากับ n/N และการคัดเลือกหน่วยตัวอย่างครั้งที่สองด้วยความน่าเป็นเท่ากับ $(n-1)/(N-1)$ และใน 1 ชุดตัวอย่างที่คัดเลือกหน่วยตัวอย่างที่แตกต่างกัน n หน่วยตัวอย่าง จะเกิดขึ้นด้วยความน่าจะเป็นเท่ากับ

$$\frac{n}{N}\cdot\frac{(n-1)}{(N-1)}\cdot\frac{(n-2)}{(N-2)}\cdots\cdots\frac{(1)}{(N-n+1)}=\frac{n!(N-n)!}{N!}=\frac{1}{{}^{N}C_{n}}$$

2. การสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง

การสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่างเหมาะกับสำหรับสถานการณ์ที่วัดค่าสังเกตของหน่วยตัวอย่างอาจมีความยุ่งยาก ใช้แรงงาน เวลามาก และเสียค่าใช้จ่ายสูง ซึ่งเป็นวิธีการสุ่มเลือกตัวอย่างโดยการเรียงลำดับที่ให้กับหน่วยตัวอย่างในแต่ละชุดตัวอย่าง และเลือกหน่วยตัวอย่างมาเพียงหนึ่งหน่วยของแต่ละชุดตัวอย่างมีการดำเนินการสุ่มตัวอย่างทั้งสิ้นทั้งหมด r รอบ โดยแต่ละรอบ มีขั้นตอนการสุ่มเลือกตัวอย่างดังนี้

กรณีการสุ่มเลือกตัวอย่างรอบที่ 1 $(j=1)$

ขั้นตอนที่ 1 สุ่มตัวอย่างของชุดตัวอย่างที่ 1 $(m=1)$ ในรอบที่ 1 โดยสุ่มตัวอย่างแบบอย่างง่ายมา k หน่วย แล้วเรียงลำดับค่าสังเกตจากน้อยไปมาก ซึ่งกำหนดให้อยู่ในลำดับที่ l โดยที่ $l=1,2,\ldots,k$ จะได้ค่าสังเกตที่เรียงลำดับแล้วเป็น $x_{(1)11}, x_{(2)11}, \cdots, x_{(k)11}$ หรือกำหนดเป็นสัญลักษณ์ คือ $x_{(l)mj}$ เป็นค่าสังเกตจากการสุ่มหน่วยตัวอย่างลำดับที่ l ของชุดตัวอย่างที่ m ในรอบที่ j โดยที่ $l=1,2,\ldots,k;\ m=1$ และ $j=1$ ในการสุ่มตัวอย่างของชุดตัวอย่างที่ 1 จะเลือกค่าสังเกตที่ใช้ในการศึกษาเป็นค่าที่ลำดับต่ำสุดอันดับที่ 1 คือ $x_{(1)11}$ ซึ่งกำหนดให้เป็น $x_{(1)1}$ หรือกำหนดเป็นสัญลักษณ์ คือ $x_{(i)j}$ เป็นตัวแทนค่าสังเกตของหน่วยตัวอย่างที่ใช้ในการศึกษาจากลำดับที่ l ของชุดตัวอย่างที่ m ซึ่งตำแหน่งของลำดับที่กับชุดตัวอย่างที่ต้องเท่ากันกำหนดเป็นสัญลักษณ์ i ในรอบที่ j โดยที่ $i=1,2,\ldots,k$ และ $j=1$

ขั้นตอนที่ 2 สุ่มตัวอย่างของชุดตัวอย่างที่ 2 ถึงชุดตัวอย่างที่ k $(m=2,\ldots,k)$ ในรอบที่ 1 โดยแต่ละชุดตัวอย่างทำการสุ่มตัวอย่างแบบอย่างง่ายมา k หน่วย และทำการสุ่มตัวอย่างซ้ำอีก $k-2$ ครั้ง และในแต่ละครั้งจะเรียงลำดับค่าสังเกตจากน้อยไปมากได้ค่าสังเกต ดังนี้

$$\begin{array}{c}
x_{(1)21}, x_{(2)21}, x_{(3)21}, \cdots, x_{(k)21} \\
x_{(1)31}, x_{(2)31}, x_{(3)31}, \cdots, x_{(k)31} \\
\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
x_{(1)k1}, x_{(2)k1}, x_{(3)k1}, \cdots, x_{(k)k1}
\end{array}$$

และในการเลือกค่าสังเกตที่ใช้ในการศึกษาของชุดตัวอย่างที่ 2 ถึงชุดตัวอย่างที่ k ในรอบที่ 1 จะเลือกค่าสังเกตที่ใช้ในการศึกษาเป็นค่าที่ลำดับต่ำสุดเป็นอันดับที่ l ของชุดตัวอย่างที่ m โดยที่ $l=2,3,\ldots,k$ และ $m=2,3,\ldots,k$ ซึ่งตำแหน่งของลำดับที่กับชุดตัวอย่างที่ต้องเท่ากัน จะได้ $x_{(2)21}, x_{(3)31}, \cdots, x_{(k)k1}$ ซึ่งกำหนดให้เป็น $x_{(2)1}, x_{(3)1}, \cdots, x_{(k)1}$ ตามลำดับ และการสุ่มตัวอย่างรอบที่ 1 จะได้ตัวแทนค่าสังเกต ดังนี้

$$x_{(1)1}, x_{(2)1}, x_{(3)1}, \cdots, x_{(k)1}$$

โดยเลือกค่าสังเกตที่ใช้ในการศึกษาในรอบที่ 1 มีจำนวน k หน่วย หรือแสดงไว้ในตาราง 2 ของรอบที่ 1

กรณีการสุ่มเลือกตัวอย่างรอบที่ 2 ถึงรอบที่ r $(j = 2, 3, \ldots, r)$

สุ่มตัวอย่างในลักษณะเดียวกันกับกรณีการสุ่มเลือกตัวอย่างรอบที่ 1 จะได้ค่าสังเกตที่ใช้ในการศึกษา $x_{(i)j}$ โดยที่ $i = 1, 2, \ldots, k$ และ $j = 2, 3, \ldots, r$ คือ

$$\begin{array}{c}
x_{(1)2}, x_{(2)2}, x_{(3)2}, \ldots, x_{(k)2} \\
x_{(1)3}, x_{(2)3}, x_{(3)3}, \ldots, x_{(k)3} \\
\vdots \quad \vdots \quad \vdots \\
x_{(1)r}, x_{(2)r}, x_{(3)r}, \ldots, x_{(k)r}
\end{array}$$

โดยค่าสังเกตที่ใช้ในการศึกษาจะมีจำนวนทั้งสิ้น kr หน่วย $(n = kr)$ มาจากการ**เลือกหน่วยตัวอย่าง**ทั้งหมด k^2r หน่วยของประชากร โดยที่ $k^2r < N$ หรือแสดงในตาราง 2 ของ**รอบที่ 2 ถึงรอบที่** r

ตาราง 2 แสดงค่าสังเกตโดยกระบวนการสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง

รอบที่	ชุดตัวอย่างที่	ลำดับที่					ตัวแทนค่าสังเกต
		1	2	3	…	k	
1	1	$x_{(1)11}$	$x_{(2)11}$	$x_{(3)11}$	…	$x_{(k)11}$	$x_{(1)1}$
	2	$x_{(1)21}$	$x_{(2)21}$	$x_{(3)21}$	…	$x_{(k)21}$	$x_{(2)1}$
	3	$x_{(1)31}$	$x_{(2)31}$	$x_{(3)31}$	…	$x_{(k)31}$	$x_{(3)1}$
	⋮				⋮		⋮
	k	$x_{(1)k1}$	$x_{(2)k1}$	$x_{(2)k1}$	…	$x_{(k)k1}$	$x_{(k)1}$
2	1	$x_{(1)12}$	$x_{(2)12}$	$x_{(3)12}$	…	$x_{(k)12}$	$x_{(1)2}$
	2	$x_{(1)22}$	$x_{(2)22}$	$x_{(3)22}$	…	$x_{(k)22}$	$x_{(2)2}$
	3	$x_{(1)32}$	$x_{(2)32}$	$x_{(3)32}$	…	$x_{(k)32}$	$x_{(3)2}$
	⋮				⋮		⋮
	k	$x_{(1)k2}$	$x_{(2)k2}$	$x_{(3)k2}$	…	$x_{(k)k2}$	$x_{(k)2}$

ตาราง 2 (ต่อ)

รอบที่	ชุดตัวอย่างที่	ลำดับที่ 1	ลำดับที่ 2	ลำดับที่ 3	ลำดับที่ …	ลำดับที่ k	ตัวแทนค่าสังเกต
3	1	$x_{(1)13}$	$x_{(2)13}$	$x_{(3)13}$	…	$x_{(k)13}$	$x_{(1)3}$
	2	$x_{(1)23}$	$x_{(2)23}$	$x_{(3)23}$	…	$x_{(k)23}$	$x_{(2)3}$
	3	$x_{(1)33}$	$x_{(2)33}$	$x_{(3)33}$	…	$x_{(k)33}$	$x_{(3)3}$
	⋮				⋮		⋮
	k	$x_{(1)k3}$	$x_{(2)k3}$	$x_{(3)k3}$	…	$x_{(k)k3}$	$x_{(k)3}$
⋮				…			⋮
r	1	$x_{(1)1r}$	$x_{(2)1r}$	$x_{(3)1r}$	…	$x_{(k)1r}$	$x_{(1)r}$
	2	$x_{(1)2r}$	$x_{(2)2r}$	$x_{(3)2r}$	…	$x_{(k)2r}$	$x_{(2)r}$
	3	$x_{(1)3r}$	$x_{(2)3r}$	$x_{(3)3r}$	…	$x_{(k)3r}$	$x_{(3)r}$
	⋮				⋮		⋮
	k	$x_{(1)kr}$	$x_{(2)kr}$	$x_{(3)kr}$	…	$x_{(k)kr}$	$x_{(k)r}$

และสามารถสรุปตัวแทนค่าสังเกตของหน่วยตัวอย่างที่นำมาศึกษาไว้ในตาราง 3

ตาราง 3 แสดงลักษณะตัวแทนค่าสังเกตของหน่วยตัวอย่างที่ได้จากการสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง

	รอบที่ 1	รอบที่ 2	รอบที่ 3	…	รอบที่ r
ชุดตัวอย่างที่/ลำดับที่ 1	$x_{(1)1}$	$x_{(1)2}$	$x_{(1)3}$	…	$x_{(1)r}$
ชุดตัวอย่างที่/ลำดับที่ 2	$x_{(2)1}$	$x_{(2)2}$	$x_{(2)3}$	…	$x_{(2)r}$
ชุดตัวอย่างที่/ลำดับที่ 3	$x_{(3)1}$	$x_{(3)2}$	$x_{(3)3}$	…	$x_{(3)r}$
⋮	⋮	⋮	⋮	⋮	⋮
ชุดตัวอย่างที่/ลำดับที่ k	$x_{(k)1}$	$x_{(k)2}$	$x_{(1)3}$	…	$x_{(k)r}$

โครงสร้างของตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง

โครงสร้างของตัวอย่างที่ได้จากการสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่างจะแตกต่างจากตัวอย่างที่ได้จากการสุ่มตัวอย่างอย่างง่าย กล่าวคือ ทำการสุ่มตัวอย่างมา 1 รอบ ขนาด k หน่วย โดยพิจารณาโครงสร้างตัวอย่างแบบ RSS ของสถิติอันดับ โดยกำหนดให้ $x_1, x_2, \ldots, x_k$ เป็นตัวอย่างขนาด k หน่วยที่ได้จากการสุ่มตัวอย่างแบบ SRS จากประชากรที่มีฟังก์ชันการแจกแจงความน่าจะเป็นแบบต่อเนื่อง $f(x)$ และฟังก์ชันการแจกแจงความน่าจะเป็นสะสม $F(x)$ และให้ $x_1^*, x_2^*, \ldots, x_k^*$ เป็นตัวอย่างขนาด k ตัวอย่าง ที่ได้จากการสุ่มตัวอย่างแบบ RSS ในการสุ่มตัวอย่างแบบ SRS ค่าสังเกต k ค่าเป็นอิสระกัน ซึ่งค่าสังเกตนั้นไม่มีโครงสร้างที่แสดงถึงความสัมพันธ์ของค่าสังเกตนั้นกับค่าสังเกตอื่น และกำหนดให้ $x_{(1)} \le x_{(2)} \le \ldots \le x_{(k)}$ เป็นสถิติลำดับจากตัวอย่างที่ได้จากการสุ่มตัวอย่างแบบ SRS (ณัฐสุวัชร์ ถาวรธิรา, 2552, หน้า 9)

สำหรับตัวอย่างที่ได้จากการสุ่มแบบ RSS ข้อมูลและโครงสร้างความสัมพันธ์ของค่าสังเกตนั้นกับค่าสังเกตอื่นได้ถูกเพิ่มจากกระบวนการในการเลือกตัวอย่างมาเป็นตัวแทน k ตัวอย่างจากตัวอย่างทั้งหมด k^2 ตัวอย่าง มาจากการสุ่มตัวอย่างประชากร โดยกำหนด $x_{(1)}^* \le x_{(2)}^* \le \ldots \le x_{(k)}^*$ เป็นสถิติลำดับจากการสุ่มตัวอย่างแบบ RSS แต่ละค่าสังเกตอิสระกัน ซึ่งเป็นโครงสร้างที่เพิ่มขึ้นนี้ทำให้การสุ่มตัวอย่างแบบ RSS มีประสิทธิภาพมากกว่าการสุ่มตัวอย่างแบบ SRS ที่ขนาดตัวอย่างเท่ากัน ซึ่งจะกล่าวในหน้า 22 (ณัฐสุวัชร์ ถาวรธิรา, 2552, หน้า 10)

1. สถิติลำดับ

นิยาม สถิติลำดับ (Order Statistic) ของตัวอย่างสุ่ม $x_1, \ldots, x_n$ คือ ค่าของตัวอย่างที่เรียงลำดับจากน้อยไปมาก มักแทนด้วย $x_{(1)}, \ldots, x_{(n)}$ เรียกว่า สถิติลำดับที่ $1, \ldots, n$ ตามลำดับ สถิติลำดับเป็นตัวแปรสุ่ม และ $x_{(1)} \le x_{(2)} \le \ldots \le x_{(n)}$

โดยที่

$x_{(1)} = \min(x_1, \ldots, x_n)$

$x_{(2)}$ = ตัวแปรสุ่มที่มีค่าต่ำเป็นลำดับที่ 2

$\vdots \qquad \vdots \qquad \vdots$

$x_{(n)} = \max(x_1, \ldots, x_n)$

(ประชุม สุวัตถี, 2553, หน้า 30)

2. การแจกแจงของสถิติลำดับที่ 1 และที่ n

ทฤษฎีบท ถ้า $x_1, \ldots, x_n$ เป็นตัวอย่างสุ่มมาจากประชากรที่มีฟังก์ชันความหนาแน่น $f(x)$ และมีฟังก์ชันการแจกแจง $F(x)$ แล้วฟังก์ชันความหนาแน่นของ $x_{(r)}, 1 \le r \le n$ ได้แก่

$$g_r(x_r) = \frac{n!}{(r-1)!(n-r)!}\left[F(x_r)\right]^{r-1}\left[1-F(x_r)\right]^{n-r} f(x_r)$$

ทฤษฎีบท ถ้า $x_1,\ldots,x_n$ เป็นตัวอย่างสุ่มมาจากประชากรที่มีฟังก์ชันความหนาแน่น $f(x)$ และมีฟังก์ชันการแจกแจง $F(x)$ แล้ว ฟังก์ชันความหนาแน่นของ $x_{(1)}$ ได้แก่

$$g_1(x_1) = n\left[1-F(x_1)\right]^{n-1} f(x_1)$$

และฟังก์ชันความหนาแน่นของ $x_{(n)}$ ได้แก่

$$g_n(x_n) = n\left[F(x_n)\right]^{n-1} f(x_n)$$

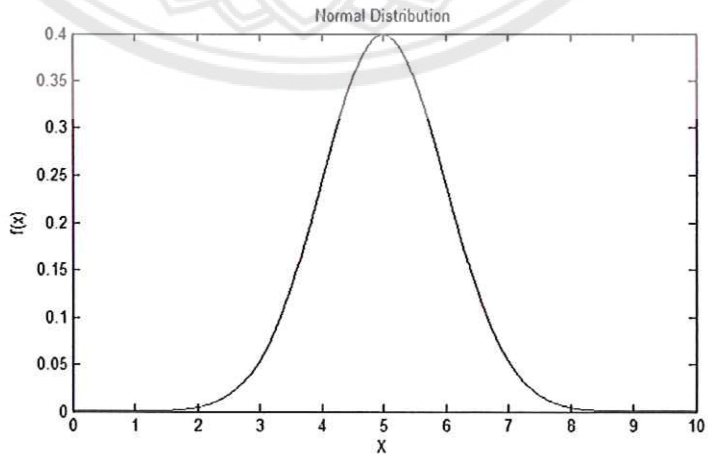
(ประชุม สุวัตถี, 2553, หน้า 34)

การแจกแจงของประชากรที่ใช้ในงานวิจัย

1. **การแจกแจงแบบปกติ** มีฟังก์ชันความหนาแน่น $f(x)$ คือ

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \quad \text{เมื่อ} \quad -\infty < x < \infty, -\infty < \mu < \infty \text{ และ } \sigma^2 > 0$$

มีค่าเฉลี่ยเท่ากับ μ และค่าความแปรปรวนเท่ากับ σ^2

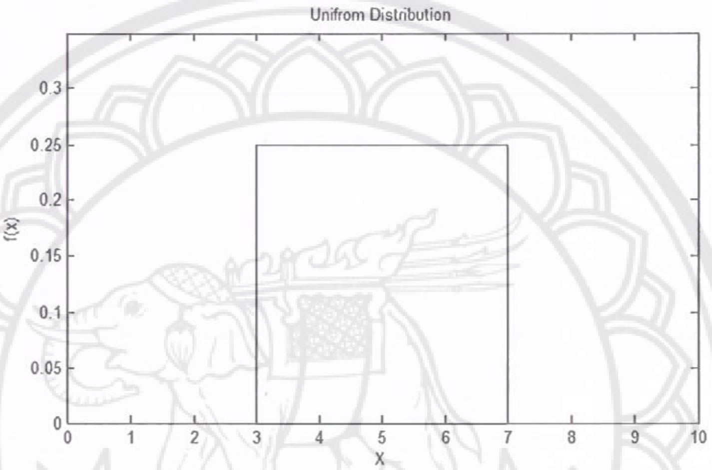


ภาพ 1 แสดงการแจกแจงแบบปกติ ที่มีค่าเฉลี่ย 5 และค่าความแปรปรวน 1

2. การแจกแจงแบบยูนิฟอร์ม มีฟังก์ชันความหนาแน่น $f(x)$ คือ

$$f(x;a,b)=\frac{1}{b-a} \quad \text{เมื่อ} \quad a\le x\le b$$

มีค่าเฉลี่ยเท่ากับ $\frac{a+b}{2}$ และค่าความแปรปรวนเท่ากับ $\frac{(b-a)^2}{12}$

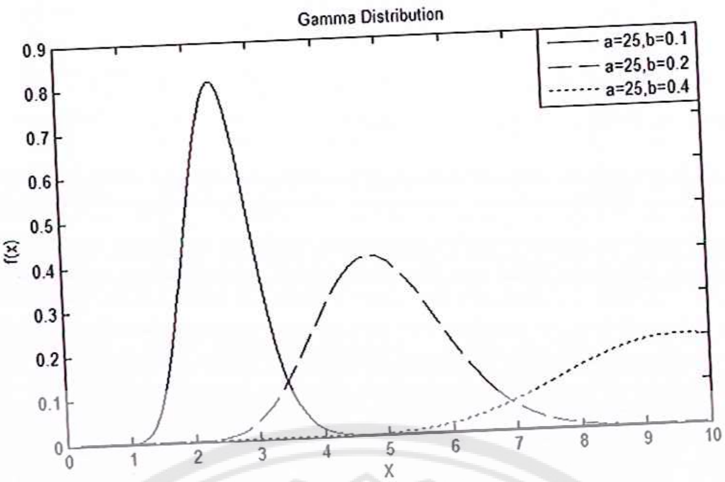


ภาพ 2 แสดงการแจกแจงแบบยูนิฟอร์ม ที่มีค่าพารามิเตอร์ $a=3$ และ $b=7$

3. การแจกแจงแบบแกมมา มีฟังก์ชันความหนาแน่น $f(x)$ คือ

$$f(x;a,b)=\frac{1}{\Gamma(a)b^a}x^{a-1}e^{-\frac{x}{b}} \quad \text{เมื่อ} \quad x>0, a>0 \text{ และ } b>0$$

มีค่าเฉลี่ยเท่ากับ ab และค่าความแปรปรวนเท่ากับ ab^2



ภาพ 3 แสดงการแจกแจงแบบแกมมา ที่มีค่าพารามิเตอร์ $a = 25$ และ $b = 0.1, 0.2, 0.4$

วิธีการประมาณค่า

สมมติว่าผู้ศึกษาจะใช้ตัวแปรสุ่ม y ในการแสดงคุณสมบัติของประชากร y อาจจะเป็นตัวแปรสุ่มมิติเดียว หรือหลายมิติ (เวกเตอร์) ก็ได้ ซึ่งโดยส่วนใหญ่แล้วค่าพารามิเตอร์มักจะไม่ทราบค่า ดังนั้นวิธีการประมาณค่าพารามิเตอร์จึงเป็นสิ่งที่ต้องการศึกษา เมื่อประมาณค่าพารามิเตอร์ได้แล้วสามารถประมาณค่าพารามิเตอร์อื่น ๆ ที่สนใจได้อีก ซึ่งงานวิจัยนี้ผู้วิจัยสนใจเฉพาะการประมาณค่าเฉลี่ยเท่านั้น โดยปกติการประมาณค่าจะใช้วิธีการประมาณค่า 2 วิธี ดังนี้

1. การประมาณค่าแบบง่าย (Simple Estimation)

การประมาณค่าแบบง่ายเป็นการประมาณค่าเฉลี่ยโดยอาศัยข้อมูลตัวแปรที่สนใจศึกษาเท่านั้น

1.1 **การประมาณค่าเฉลี่ยแบบง่าย ภายใต้แผนการสุ่มตัวอย่างอย่างง่าย**

การสุ่มตัวอย่างอย่างง่ายแบบไม่แทนที่ (SRSWOR) ขนาด n หน่วย จากประชากรขนาด N หน่วย (ทวีศักดิ์ ศิริพรไพบูลย์, 2549, หน้า 47)

1) ค่าประมาณของค่าเฉลี่ยตัวอย่าง คือ

$$\hat{\mu}_y = \frac{1}{n}\sum_{i=1}^{n} y_i$$

และ $\hat{\mu}_y$ เป็นค่าประมาณที่ไม่เอนเอียงสำหรับค่าเฉลี่ยประชากร μ_y

2) ค่าความแปรปรวนของค่าประมาณค่าเฉลี่ยตัวอย่าง คือ

$$V\left(\hat{\mu}_y\right) = \frac{N-n}{N}\frac{S_y^2}{n}$$

โดยที่

$$S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \mu_y)^2 \text{ และ } \mu_y = \frac{1}{N}\sum_{i=1}^{N} y_i$$

3) ค่าประมาณความแปรปรวนของค่าประมาณค่าเฉลี่ยตัวอย่าง คือ

$$v(\hat{\mu}_y) = \frac{N-n}{N}\frac{s_y^2}{n}$$

โดยที่

$$s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \hat{\mu}_y)^2$$

1.2 การประมาณค่าเฉลี่ยแบบง่าย ภายใต้แผนการสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง

การสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่างเป็นการสุ่มตัวอย่างอย่างง่ายมาจำนวน k^2r หน่วยตัวอย่างจากประชากร N โดยเราจะเลือกหน่วยตัวอย่างที่นำมาศึกษาขนาด kr หน่วย $\left(n = kr\right)$ โดยที่ $k^2r < N$

1. ค่าประมาณของค่าเฉลี่ยตัวอย่าง คือ

$$\hat{\mu}_{RSS} = \frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r} y_{(i)j}$$

และ $\hat{\mu}_{RSS}$ เป็นค่าประมาณที่ไม่เอนเอียงสำหรับค่าเฉลี่ยประชากร μ_y

2. ค่าความแปรปรวนของค่าประมาณค่าเฉลี่ยตัวอย่าง คือ

$$V\left(\hat{\mu}_{RSS}\right) = \frac{1}{k^2r}\sum_{i=1}^{k}\sigma_{(i)}^2$$

โดยที่

$$\sigma_{(i)}^{2}=\frac{1}{r}\sum_{j=1}^{r}\left(y_{(i)j}-\mu_{y[i]}\right)^{2}$$

เมื่อ $\sigma_{(i)}^{2}$ คือ ความแปรปรวนของหน่วยตัวอย่างสถิติลำดับที่ i

$\mu_{y[i]}$ คือ ค่าเฉลี่ยของหน่วยตัวอย่างที่สนใจศึกษาของสถิติลำดับที่ i

หรือ

$$V\left(\hat{\mu}_{RSS}\right)=\frac{1}{kr}\left[\sigma_{y}^{2}-\frac{1}{k}\sum_{i=1}^{k}\left(\mu_{y[i]}-\mu_{y}\right)^{2}\right]$$

โดยที่

$$\sigma_{y}^{2}=\frac{1}{k}\sum_{i=1}^{k}\sigma_{(i)}^{2}+\frac{1}{k}\sum_{i=1}^{k}\left(\mu_{y[i]}-\mu_{y}\right)^{2}$$

เมื่อ σ_{y}^{2} คือ ความแปรปรวนของประชากร

3. ค่าประมาณความแปรปรวนของค่าประมาณค่าเฉลี่ยตัวอย่าง คือ

$$v\left(\hat{\mu}_{RSS}\right)=\frac{1}{k^{2}r}\sum_{i=1}^{k}\hat{\sigma}_{(i)}^{2}$$

โดยที่

$$\hat{\sigma}_{(i)}^{2}=\frac{1}{r-1}\sum_{j=1}^{r}\left[y_{(i)j}-\hat{\mu}_{y[i]}\right]^{2} \text{ และ } \hat{\mu}_{y[i]}=\frac{1}{r}\sum_{j=1}^{r}y_{(i)j}$$

เมื่อ $\hat{\mu}_{y[i]}$ คือ ค่าประมาณของ $\mu_{y[i]}$

$\hat{\sigma}_{(i)}^{2}$ คือ ค่าประมาณของ $\sigma_{(i)}^{2}$

หรือ

$$v\left(\hat{\mu}_{RSS}\right)=\frac{1}{k^{2}r\left(r-1\right)}\sum_{i=1}^{k}\sum_{j=1}^{r}\left(y_{(i)j}-\hat{\mu}_{y[i]}\right)^{2}$$

พิสูจน์

1. กำหนดให้ $E\left(y_{(i)j}\right)=E\left(y_{(i)1}\right)=E\left(y_{(i)}\right)$ เมื่อ $i=1,2,\ldots,k;\ j=1,2,\ldots,r$

สามารถอธิบายได้ว่า ค่าคาดหวังของ $y_{(i)j}$ โดยที่ j เป็นค่าใด ๆ ก็ตามจะมีค่าเท่ากับค่าคาดหวังของ $y_{(i)1}$ สำหรับ $j=1$

$$E\left(\hat{\mu}_{RSS}\right)=\frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r}E\left(y_{(i)j}\right) \tag{1}$$

$$=\frac{1}{k}\sum_{i=1}^{k}E\left(y_{(i)1}\right)=\frac{1}{k}\sum_{i=1}^{k}E\left(y_{(i)}\right)$$

$$=\frac{1}{k}\sum_{i=1}^{k}\int_{-\infty}^{\infty}y\frac{k!}{(i-1)!(k-i)!}\left[F(y)\right]^{i-1}\left[1-F(y)\right]^{k-i}f(y)\,dy$$

$$=\frac{1}{k}\sum_{i=1}^{k}\left\{\int_{-\infty}^{\infty}ky\binom{k-1}{i-1}\left[F(y)\right]^{i-1}\left[1-F(y)\right]^{k-i}f(y)\,dy\right\}$$

$$=\int_{-\infty}^{\infty}yf(y)\left\{\sum_{i=1}^{k}\binom{k-1}{i-1}\left[F(y)\right]^{i-1}\left[1-F(y)\right]^{k-i}\right\}dy \tag{2}$$

กำหนดให้ $z=i-1$ แทนค่าลงในสมการ (2) จะได้ว่า

$$E\left(\hat{\mu}_{RSS}\right)=\int_{-\infty}^{\infty}yf(y)\left\{\sum_{i=1}^{k}\binom{k-1}{z}\left[F(y)\right]^{z}\left[1-F(y)\right]^{(k-1)-z}\right\}dy$$

เนื่องจาก $$\sum_{i=1}^{k}\binom{k-1}{z}\left[F(y)\right]^{z}\left[1-F(y)\right]^{(k-1)-z}=1 \tag{3}$$

แทนค่าสมการ (3) จะได้ว่า

$$E\left(\hat{\mu}_{RSS}\right)=\int_{-\infty}^{\infty}yf(y)\,dy=\mu_y$$

ดังนั้น $\hat{\mu}_{RSS}$ เป็นค่าประมาณที่ไม่เอนเอียงสำหรับค่าเฉลี่ยประชากร μ_y

2. การพิสูจน์หา $V\left(\hat{\mu}_{RSS}\right)$ ได้ดังนี้

$$V\left(\hat{\mu}_{RSS}\right)=V\left(\frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r}y_{(i)j}\right)$$

$$=\left(\frac{1}{kr}\right)^2\sum_{i=1}^{k}\sum_{j=1}^{r}V\left(y_{(i)j}\right)$$

$$=\left(\frac{1}{kr}\right)^2\sum_{i=1}^{k}\sum_{j=1}^{r}\sigma_{(i)}^2$$

$$=\frac{1}{k^2r}\sum_{i=1}^{k}\sigma_{(i)}^2$$

จะได้ว่า $V\left(\hat{\mu}_{RSS}\right)=\frac{1}{k^2r}\sum_{i=1}^{k}\sigma_{(i)}^2$ เมื่อ $\sigma_{(i)}^2=\frac{1}{r}\sum_{j=1}^{r}\left(y_{(i)j}-\mu_{y[i]}\right)^2$ (4)

กำหนดให้ $\mu_{y[i]}=E\left(y_{(i)}\right)$ เมื่อ $i=1,2,\ldots,k$ จะได้ว่า

$$E\left(y_{(i)}-\mu_y\right)^2=E\left[\left(y_{(i)}-\mu_{y[i]}+\mu_{y[i]}-\mu_y\right)^2\right]$$

$$=E\left[\left(y_{(i)}-\mu_{y[i]}\right)^2\right]+\left(\mu_{y[i]}-\mu_y\right)^2$$

ดังนั้น $E\left[\left(y_{(i)}-\mu_{y[i]}\right)^2\right]=E\left(y_{(i)}-\mu_y\right)^2-\left(\mu_{y[i]}-\mu_y\right)^2$ (5)

แทนค่าสมการ (5) ลงในสมการ (4)

$$V\left(\hat{\mu}_{RSS}\right)=\frac{1}{k^2r}\sum_{i=1}^{k}E\left[\left(y_{(i)}-\mu_y\right)^2\right]-\frac{1}{k^2r}\sum_{i=1}^{k}\left(\mu_{y[i]}-\mu_y\right)^2 \quad (6)$$

พิจารณา $\sum_{i=1}^{k}E\left[\left(y_{(i)}-\mu_y\right)^2\right]$ จะได้ว่า

$$\sum_{i=1}^{k} E\left[\left(y_{(i)}-\mu_y\right)^2\right]=\sum_{i=1}^{k}\int_{-\infty}^{\infty} k\left(y-\mu_y\right)^2\binom{k-1}{i-1}\left[F(y)\right]^{i-1}\left[1-F(y)\right]^{k-i} f(y)\,dy$$

$$=k\int_{-\infty}^{\infty}\left(y-\mu_y\right)^2 f(y)\left\{\sum_{i=1}^{k}\binom{k-1}{i-1}\left[F(y)\right]^{i-1}\left[1-F(y)\right]^{k-i}\right\}dy \quad (7)$$

กำหนดให้ $z=i-1$ แทนค่าลงในสมการ (7) จะได้ว่า

$$\sum_{i=1}^{k} E\left[\left(y_{(i)}-\mu_y\right)^2\right]=k\int_{-\infty}^{\infty}\left(y-\mu_y\right)^2 f(y)\left\{\sum_{i=1}^{k}\binom{k-1}{z}\left[F(y)\right]^{z}\left[1-F(y)\right]^{(k-1)-z}\right\}dy$$

แทนค่าสมการ (3) จะได้ว่า

$$\sum_{i=1}^{k} E\left[\left(y_{(i)}-\mu_y\right)^2\right]=k\int_{-\infty}^{\infty}\left(y-\mu_y\right)^2 f(y)dy=k\sigma_y^2 \quad (8)$$

แทนค่าสมการ (8) ลงในสมการ (6)

$$V\left(\hat{\mu}_{RSS}\right)=\frac{1}{k^2 r}\left\{k\sigma_y^2-\sum_{i=1}^{k}\left(\mu_{y[i]}-\mu_y\right)^2\right\}$$

$$=\frac{1}{kr}\left\{\sigma_y^2-\frac{1}{k}\sum_{i=1}^{k}\left(\mu_{y[i]}-\mu_y\right)^2\right\}$$

ดังนั้น $V\left(\hat{\mu}_{RSS}\right)=\frac{1}{kr}\left\{\sigma_y^2-\frac{1}{k}\sum_{i=1}^{k}\left(\mu_{y[i]}-\mu_y\right)^2\right\}$ (9)

จากสมการ (9) เมื่อพิจารณาจะเห็นว่า

$$V\left(\hat{\mu}_{RSS}\right)=\frac{1}{kr}\left\{\sigma_y^2-\frac{1}{k}\sum_{i=1}^{k}\left(\mu_{y[i]}-\mu_y\right)^2\right\}$$

$$=\frac{\sigma_y^2}{kr}-\frac{1}{k^2 r}\sum_{i=1}^{k}\left(\mu_{y[i]}-\mu_y\right)^2$$

$$= Var\left(\hat{\mu}_y\right) - \frac{1}{k^2 r}\sum_{i=1}^{k}\left(\mu_{y[i]} - \mu_y\right)^2$$

เนื่องจาก $\sum_{i=1}^{k}\left(\mu_{y[i]} - \mu_y\right)^2 \geq 0$ ดังนั้น $Var\left(\hat{\mu}_{RSS}\right) \leq Var\left(\hat{\mu}_y\right)$ เสมอ

3) เนื่องจาก $\hat{\sigma}_{(i)}^2 = \frac{1}{r-1}\sum_{j=1}^{r}\left[y_{(i)j} - \hat{\mu}_{y[i]}\right]^2$ เป็นตัวประมาณโมเมนต์ของ $\sigma_{(i)}^2$ จะได้

$$v\left(\hat{\mu}_{RSS}\right) = \frac{1}{k^2 r}\sum_{i=1}^{k}\hat{\sigma}_{(i)}^2$$

โดยที่

$$\hat{\sigma}_{(i)}^2 = \frac{1}{r-1}\sum_{j=1}^{r}\left[y_{(i)j} - \hat{\mu}_{y[i]}\right]^2 \text{ และ } \hat{\mu}_{y[i]} = \frac{1}{r}\sum_{j=1}^{r}y_{(i)j}$$

ดังนั้น
$$v\left(\hat{\mu}_{RSS}\right) = \frac{1}{k^2 r\left(r-1\right)}\sum_{i=1}^{k}\sum_{j=1}^{r}\left(y_{(i)j} - \hat{\mu}_{y[i]}\right)^2$$

2. การประมาณค่าด้วยตัวแปรที่เกี่ยวข้อง (Composite Estimation) ภายใต้แผนการสุ่มตัวอย่างอย่างง่าย และแผนการสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง

การประมาณค่าด้วยตัวแปรที่เกี่ยวข้องเป็นการประมาณค่าข้อมูลที่เราสนใจศึกษา และใช้ตัวแปรที่คาดว่ามีความสัมพันธ์กับข้อมูลเข้ามาช่วยในการประมาณ เช่น การประมาณค่าเฉลี่ยแบบอัตราส่วน และการประมาณค่าเฉลี่ยแบบถดถอย เป็นต้น

2.1 การประมาณพารามิเตอร์ โดยวิธีการประมาณค่าเฉลี่ยแบบอัตราส่วนภายใต้แผนการสุ่มตัวอย่างอย่างง่าย

การสุ่มตัวอย่างแบบไม่แทนที่อย่างง่ายขนาด n หน่วย จากประชากรขนาด N หน่วย เมื่อตัวแปรที่ต้องการศึกษากับตัวแปรช่วยมีความสัมพันธ์กัน (ทวีศักดิ์ ศิริพรไพบูลย์, 2549, หน้า 74-78)

1. ค่าประมาณของค่าเฉลี่ยตัวอย่าง คือ

$$\hat{\mu}_{RSRS} = \frac{\hat{\mu}_y}{\hat{\mu}_x}\mu_x = \hat{R}\mu_x$$

โดยที่

$$\hat{\mu}_y = \frac{1}{n}\sum_{i=1}^{n} y_i,\ \hat{\mu}_x = \frac{1}{n}\sum_{i=1}^{n} x_i \text{ และ } \mu_x = \frac{1}{N}\sum_{i=1}^{N} x_i$$

2. ค่าความแปรปรวนของค่าประมาณค่าเฉลี่ยตัวอย่าง คือ

$$V\left(\hat{\mu}_{RSRS}\right) = \left(1-\frac{n}{N}\right)\frac{S_d^2}{n}$$

โดยที่

$$S_d^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(y_i - Rx_i\right)^2 \text{ และ } R = \frac{\mu_y}{\mu_x}$$

3. ค่าประมาณความแปรปรวนของค่าประมาณค่าเฉลี่ยตัวอย่าง คือ

$$v\left(\hat{\mu}_{RSRS}\right) = \left(1-\frac{n}{N}\right)\frac{s_d^2}{n}$$

โดยที่

$$s_d^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(y_i - \hat{R}x_i\right)^2$$

2.2 การประมาณพารามิเตอร์ โดยวิธีการประมาณค่าเฉลี่ยแบบถดถอย ภายใต้แผนการสุ่มตัวอย่างอย่างง่าย

การสุ่มตัวอย่างแบบไม่แทนที่อย่างง่ายขนาด n หน่วย จากประชากรขนาด N หน่วย เมื่อตัวแปรที่ต้องการศึกษากับตัวแปรช่วยมีความสัมพันธ์กันเชิงเส้นตรง (ทวีศักดิ์ ศิริพรไพบูลย์, 2549, หน้า 81-84)

1. ค่าประมาณของค่าเฉลี่ยตัวอย่าง คือ

$$\hat{\mu}_{LSRS} = \hat{\mu}_y + \hat{\beta}\left(\mu_x - \hat{\mu}_x\right)$$

โดยที่

$$\hat{\beta} = \frac{\sum_{i=1}^{n}\left(x_i - \hat{\mu}_x\right)\left(y_i - \hat{\mu}_y\right)}{\sum_{i=1}^{n}\left(x_i - \hat{\mu}_x\right)^2}$$

เมื่อ $\hat{\beta}$ คือ ค่าประมาณสัมประสิทธิ์ถดถอยด้วยวิธีกำลังสองน้อยที่สุด

2. ค่าความแปรปรวนของค่าประมาณค่าเฉลี่ยตัวอย่าง คือ

$$V\left(\hat{\mu}_{LSRS}\right)=\left(1-\frac{n}{N}\right)\frac{S_z^2}{n}$$

โดยที่

$$S_z^2=\frac{1}{N-1}\sum_{i=1}^{N}\left[\left(y_i-\mu_y\right)-\beta\left(x_i-\mu_x\right)\right]^2$$

และ

$$\beta=\frac{\sum_{i=1}^{N}\left(x_i-\mu_x\right)\left(y_i-\mu_y\right)}{\sum_{i=1}^{N}\left(x_i-\mu_x\right)^2}$$

3. ค่าประมาณความแปรปรวนของค่าประมาณค่าเฉลี่ยตัวอย่าง คือ

$$v\left(\hat{\mu}_{LSRS}\right)=\left(1-\frac{n}{N}\right)\frac{s_z^2}{n}$$

โดยที่

$$s_z^2=\frac{1}{n-2}\left[\sum_{i=1}^{n}\left(y_i-\hat{\mu}_y\right)^2-\hat{\beta}^2\sum_{i=1}^{n}\left(x_i-\hat{\mu}_x\right)^2\right]$$

2.3 การประมาณพารามิเตอร์ โดยวิธีการประมาณค่าเฉลี่ยแบบอัตราส่วนของ Samawi and Muttlak ภายใต้แผนการสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง

ในปี ค.ศ.1996 Samawi and Muttlak ได้เสนอตัวประมาณแบบอัตราส่วนค่าเฉลี่ยของประชากร ดังนี้

1. ค่าประมาณค่าเฉลี่ยอัตราส่วน คือ

$$\hat{\mu}_{SRSS}=\frac{\hat{\mu}_{RSS}}{\hat{\mu}_{xRSS}}\mu_x=\hat{R}_{SRSS}\mu_x$$

โดยที่

$$\hat{\mu}_{RSS} = \frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r} y_{[i]j}\,,\ \hat{\mu}_{xRSS} = \frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r} x_{(i)j} \text{ และ } \hat{R}_{SRSS} = \frac{\hat{\mu}_{RSS}}{\hat{\mu}_{xRSS}}$$

เมื่อ $\hat{\mu}_{RSS}$ คือ ค่าเฉลี่ยของตัวแปรที่สนใจศึกษา

$\hat{\mu}_{xRSS}$ คือ ค่าเฉลี่ยของตัวแปรช่วย

โดยมีค่าเอนเอียงของตัวประมาณของค่าเฉลี่ยอัตราส่วน คือ

$$B\left(\hat{\mu}_{SRSS}\right) \cong \mu_y\left[\gamma\left(C_x^2 - C_{yx}\right) - \left(W_{x(i)}^2 - W_{yx(i)}\right)\right]$$

โดยที่

$$\gamma = \frac{1}{kr},\ C_x = \frac{\sigma_x}{\mu_x},\ C_{yx} = \rho C_y C_x,\ W_{yx(i)} = \frac{1}{k^2 r \mu_x \mu_y}\sum_{i=1}^{k}\tau_{yx(i)},\ W_{x(i)}^2 = \frac{1}{k^2 r \mu_x^2}\sum_{i=1}^{k}\tau_{x(i)}^2,$$

และ $W_{y[i]}^2 = \dfrac{1}{k^2 r \mu_y^2}\sum_{i=1}^{k}\tau_{y[i]}^2$

2. ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของค่าเฉลี่ยอัตราส่วน คือ

$$MSE\left(\hat{\mu}_{SRSS}\right) \cong V\left(\hat{\mu}_{RSS}\right) + R^2 V\left(\hat{\mu}_{xRSS}\right) - 2R\operatorname{cov}\left(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS}\right)$$

โดยที่

$$V\left(\hat{\mu}_{RSS}\right) = \frac{1}{kr}\left(\sigma_y^2 - \frac{1}{k}\sum_{i=1}^{k}\tau_{y[i]}^2\right),\ \tau_{y[i]} = \mu_{y[i]} - \mu_y$$

$$V\left(\hat{\mu}_{xRSS}\right) = \frac{1}{kr}\left(\sigma_x^2 - \frac{1}{k}\sum_{i=1}^{k}\tau_{x(i)}^2\right),\ \tau_{x(i)} = \mu_{x(i)} - \mu_x$$

และ $\operatorname{cov}\left(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS}\right) = \dfrac{1}{kr}\left(\sigma_{yx} - \dfrac{1}{k}\sum_{i=1}^{k}\tau_{yx(i)}\right),\ \tau_{yx(i)} = \left(\mu_{y[i]} - \mu_y\right)\left(\mu_{x(i)} - \mu_x\right)$

หรือ

$$MSE\left(\hat{\mu}_{SRSS}\right) \cong \frac{1}{kr}\left(\sigma_y^2 - 2R\sigma_{yx} + R^2\sigma_x^2\right) - \frac{1}{k^2 r}\left(\sum_{i=1}^{k}\tau_{y[i]}^2 - 2R\sum_{i=1}^{k}\tau_{yx(i)} + R^2\sum_{i=1}^{k}\tau_{x(i)}^2\right)$$

3. ค่าประมาณความคลาดเคลื่อนกำลังสองเฉลี่ยของค่าเฉลี่ยอัตราส่วน คือ

$$\widehat{MSE}(\hat{\mu}_{SRSS}) \cong v(\hat{\mu}_{RSS}) + \hat{R}^2 v(\hat{\mu}_{xRSS}) - 2\hat{R}\widehat{\mathrm{cov}}(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS})$$

โดยที่

$$v(\hat{\mu}_{RSS}) = \frac{1}{kr}\left(\hat{\sigma}_y^2 - \frac{1}{k}\sum_{i=1}^{k}\hat{\tau}_{y[i]}^2\right), \hat{\tau}_{y[i]} = \hat{\mu}_{y[i]} - \hat{\mu}_{RSS}$$

$$v(\hat{\mu}_{xRSS}) = \frac{1}{kr}\left(\hat{\sigma}_x^2 - \frac{1}{k}\sum_{i=1}^{k}\hat{\tau}_{x(i)}^2\right), \hat{\tau}_{x(i)} = \hat{\mu}_{x(i)} - \hat{\mu}_{xRSS}$$

และ $\widehat{\mathrm{cov}}(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS}) = \frac{1}{kr}\left(\hat{\sigma}_{yx} - \frac{1}{k}\sum_{i=1}^{k}\hat{\tau}_{yx(i)}\right), \hat{\tau}_{yx(i)} = \left(\hat{\mu}_{y[i]} - \hat{\mu}_{RSS}\right)\left(\hat{\mu}_{x(i)} - \hat{\mu}_{xRSS}\right)$

หรือ

$$\widehat{MSE}(\hat{\mu}_{SRSS}) \cong \frac{1}{kr}\left(\hat{\sigma}_y^2 - 2\hat{R}\hat{\sigma}_{yx} + \hat{R}^2\hat{\sigma}_x^2\right) - \frac{1}{k^2 r}\left(\sum_{i=1}^{k}\hat{\tau}_{y[i]}^2 - 2\hat{R}\sum_{i=1}^{k}\hat{\tau}_{yx(i)} + \hat{R}^2\sum_{i=1}^{k}\hat{\tau}_{x(i)}^2\right)$$

พิสูจน์

1. กำหนดให้ $e_0 = \frac{\hat{\mu}_{RSS} - \mu_y}{\mu_y}$ จะได้ว่า $\hat{\mu}_{RSS} = \mu_y(1+e_0)$

และ $e_1 = \frac{\hat{\mu}_{xRSS} - \mu_x}{\mu_x}$ จะได้ว่า $\hat{\mu}_{xRSS} = \mu_x(1+e_1)$

เนื่องจาก $E(e_0) = 0$ และ $E(e_1) = 0$

พิจารณาค่า

$$E(e_0^2) = \frac{1}{\mu_y^2}V(\hat{\mu}_{RSS})$$

$$= \frac{1}{\mu_y^2 kr}\left(\sigma_y^2 - \frac{1}{k}\sum_{i=1}^{k}\tau_{y[i]}^2\right) \qquad ; \tau_{y[i]} = \mu_{y[i]} - \mu_y$$

$$= \gamma C_y^2 - \frac{1}{k^2 r \mu_y^2}\sum_{i=1}^{k}\tau_{y[i]}^2 \qquad ; \gamma = \frac{1}{kr}, C_y = \frac{\sigma_y}{\mu_y}$$

$$= \gamma C_y^2 - W_{y[i]}^2 \qquad ; W_{y[i]}^2 = \frac{1}{k^2 r \mu_y^2} \sum_{i=1}^{k} \tau_{y[i]}^2$$

ในทำนองเดียวกันจะได้ $E\left(e_1^2\right)$ ดังนี้

$$E\left(e_1^2\right) = \gamma C_x^2 - W_{x[i]}^2$$

พิจารณาค่า

$$E\left(e_0 e_1\right) = \frac{1}{\mu_y \mu_x} \operatorname{cov}\left(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS}\right)$$

$$= \frac{1}{\mu_y \mu_x kr} \left(\sigma_{yx} - \frac{1}{k} \sum_{i=1}^{k} \tau_{yx(i)} \right)$$

$$= \frac{1}{kr} C_{yx} - \frac{1}{k^2 r \mu_x \mu_y} \sum_{i=1}^{k} \tau_{yx(i)} \qquad ; C_{yx} = \frac{\sigma_{yx}}{\mu_y \mu_x}$$

$$= \gamma C_{yx} - W_{yx(i)} \qquad ; \gamma = \frac{1}{kr}, W_{yx(i)} = \frac{1}{k^2 r \mu_x \mu_y} \sum_{i=1}^{k} \tau_{yx(i)}$$

และ $$E\left(e_0 e_1^2\right) = E\left[\left(\frac{\hat{\mu}_{RSS} - \mu_y}{\mu_y} \right) \left(\frac{\hat{\mu}_{xRSS} - \mu_x}{\mu_x} \right)^2 \right]$$

$$= \frac{1}{\mu_y \mu_x^2} E\left[\left(\hat{\mu}_{RSS} - \mu_y \right) \left(\hat{\mu}_{xRSS} - \mu_x \right)^2 \right]$$

และ เนื่องจากตัวหารมีค่ามากขึ้นแล้วทำให้ $E\left(e_0 e_1^2\right)$ มีค่าเข้าใกล้ศูนย์ ดังนั้น $E\left(e_0 e_1^2\right) \cong 0$

เนื่องจาก $\hat{\mu}_{SRSS} = \frac{\hat{\mu}_{RSS}}{\hat{\mu}_{xRSS}} \mu_x = \frac{\mu_y \left(1 + e_0\right)}{\mu_x \left(1 + e_1\right)} \mu_x = \mu_y \left(1 + e_0\right) \left(1 + e_1\right)^{-1}$

จะได้ว่า $\hat{\mu}_{SRSS} \cong \mu_y \left(1 + e_0\right) \left(1 - e_1 + e_1^2\right)$

$$= \mu_y\left(1 - e_1 + e_1^2 + e_0 - e_0 e_1 + e_0 e_1^2\right)$$

$$\hat{\mu}_{SRSS} - \mu_y = \mu_y\left(-e_1 + e_1^2 + e_0 - e_0 e_1 + e_0 e_1^2\right)$$

ดังนั้น ค่าเอนเอียงของตัวประมาณของค่าเฉลี่ยอัตราส่วน คือ

$$B\left(\hat{\mu}_{SRSS}\right) \cong \mu_y\left(\gamma C_x^2 - W_{x(i)}^2 - \gamma C_{yx} + W_{yx(i)}\right)$$

$$= \mu_y\left[\gamma\left(C_x^2 - C_{yx}\right) - \left(W_{x(i)}^2 - W_{yx(i)}\right)\right]$$

2. การพิสูจน์หาค่า $MSE\left(\hat{\mu}_{SRSS}\right)$ ได้ดังนี้

$$MSE\left(\hat{\mu}_{SRSS}\right) = E\left(\hat{R}_{SRSS}\mu_x - R\mu_x\right)^2$$

$$= E\left(\frac{\hat{\mu}_{RSS}}{\hat{\mu}_{xRSS}}\mu_x - R\mu_x\right)^2$$

$$= \left(\frac{\mu_x}{\hat{\mu}_{xRSS}}\right)^2 E\left(\hat{\mu}_{RSS} - R\hat{\mu}_{xRSS}\right)^2$$

เนื่องจาก $\frac{1}{\hat{\mu}_{xRSS}} = \frac{1}{\mu_x + \hat{\mu}_{xRSS} - \mu_x} = \frac{1}{\mu_x}\left(1 + \frac{\hat{\mu}_{xRSS} - \mu_x}{\mu_x}\right)^{-1} \cong \frac{1}{\mu_x}$ (โดยใช้ Taylor series)

จะได้ว่า $MSE\left(\hat{\mu}_{SRSS}\right) \cong E\left(\hat{\mu}_{RSS} - R\hat{\mu}_{xRSS}\right)^2$

$$= E\left(\hat{\mu}_{RSS} - \mu_y + \mu_y - R\hat{\mu}_{xRSS}\right)^2$$

$$= E\left[\left(\hat{\mu}_{RSS} - \mu_y\right) - R\left(\hat{\mu}_{xRSS} - \mu_x\right)\right]^2$$

$$= E\left[\left(\hat{\mu}_{RSS} - \mu_y\right)^2 + \left(R\left(\hat{\mu}_{xRSS} - \mu_x\right)\right)^2 - 2R\left(\hat{\mu}_{RSS} - \mu_y\right)\left(\hat{\mu}_{xRSS} - \mu_x\right)\right]$$

$$= E\left(\hat{\mu}_{RSS} - \mu_y\right)^2 + R^2 E\left(\hat{\mu}_{xRSS} - \mu_x\right)^2 - 2RE\left[\left(\hat{\mu}_{RSS} - \mu_y\right)\left(\hat{\mu}_{xRSS} - \mu_x\right)\right]$$

$$= E\left(\hat{\mu}_{RSS}^2\right) + \mu_y^2 - 2\mu_y E\left(\hat{\mu}_{RSS}\right) + R^2 V\left(\hat{\mu}_{xRSS}\right) - 2R\operatorname{cov}\left(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS}\right)$$

$$= \left(V\left(\hat{\mu}_{RSS}\right) + \mu_y^2\right) + \mu_y^2 - 2\mu_y^2 + R^2 V\left(\hat{\mu}_{xRSS}\right) - 2R\operatorname{cov}\left(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS}\right)$$

$$= V\left(\hat{\mu}_{RSS}\right) + R^2 V\left(\hat{\mu}_{xRSS}\right) - 2R\operatorname{cov}\left(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS}\right)$$

ดังนั้น $MSE\left(\hat{\mu}_{SRSS}\right) \cong V\left(\hat{\mu}_{RSS}\right) + R^2 V\left(\hat{\mu}_{xRSS}\right) - 2R\operatorname{cov}\left(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS}\right)$

เนื่องจาก

$$V\left(\hat{\mu}_{RSS}\right) = \frac{1}{kr}\left(\sigma_y^2 - \frac{1}{k}\sum_{i=1}^{k}\tau_{y[i]}^2\right), \tau_{y[i]} = \mu_{y[i]} - \mu_y$$

$$V\left(\hat{\mu}_{xRSS}\right) = \frac{1}{kr}\left(\sigma_x^2 - \frac{1}{k}\sum_{i=1}^{k}\tau_{x(i)}^2\right), \tau_{x(i)} = \mu_{x(i)} - \mu_x$$

และ $\operatorname{cov}\left(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS}\right) = \frac{1}{kr}\left(\sigma_{yx} - \frac{1}{k}\sum_{i=1}^{k}\tau_{yx(i)}\right), \tau_{yx(i)} = \left(\mu_{y[i]} - \mu_y\right)\left(\mu_{x(i)} - \mu_x\right)$

ดังนั้น สามารถเขียนสมการ $MSE\left(\hat{\mu}_{SRSS}\right)$ ได้ดังนี้

$$MSE\left(\hat{\mu}_{SRSS}\right) \cong V\left(\hat{\mu}_{RSS}\right) + R^2 V\left(\hat{\mu}_{xRSS}\right) - 2R\operatorname{cov}\left(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS}\right)$$

$$= \frac{1}{kr}\left(\sigma_y^2 - \frac{1}{k}\sum_{i=1}^{k}\tau_{y[i]}^2\right) + R^2\left[\frac{1}{kr}\left(\sigma_x^2 - \frac{1}{k}\sum_{i=1}^{k}\tau_{x(i)}^2\right)\right]$$

$$-2R\left[\frac{1}{kr}\left(\sigma_{yx} - \frac{1}{k}\sum_{i=1}^{k}\tau_{yx(i)}\right)\right]$$

$$=\frac{1}{kr}\sigma_y^2-\frac{1}{k^2r}\sum_{i=1}^{k}\tau_{y[i]}^2+\frac{1}{kr}R^2\sigma_x^2-\frac{1}{k^2r}R^2\sum_{i=1}^{k}\tau_{x(i)}^2$$

$$-\frac{1}{kr}2R\sigma_{yx}+\frac{1}{k^2r}2R\sum_{i=1}^{k}\tau_{yx(i)}$$

$$=\frac{1}{kr}\left(\sigma_y^2-2R\sigma_{yx}+R^2\sigma_x^2\right)-\frac{1}{k^2r}\left(\sum_{i=1}^{k}\tau_{y[i]}^2-2R\sum_{i=1}^{k}\tau_{yx(i)}+R^2\sum_{i=1}^{k}\tau_{x(i)}^2\right)$$

ดังนั้น $MSE\left(\hat{\mu}_{SRSS}\right)\cong\frac{1}{kr}\left(\sigma_y^2-2R\sigma_{yx}+R^2\sigma_x^2\right)-\frac{1}{k^2r}\left(\sum_{i=1}^{k}\tau_{y[i]}^2-2R\sum_{i=1}^{k}\tau_{yx(i)}+R^2\sum_{i=1}^{k}\tau_{x(i)}^2\right)$

3. เมื่อพิจารณาค่าประมาณความคลาดเคลื่อนกำลังสองเฉลี่ยของค่าเฉลี่ยอัตราส่วน $\left(\widehat{MSE}\left(\hat{\mu}_{SRSS}\right)\right)$ ซึ่งการประมาณค่าเป็นการประมาณโดยอาศัยข้อมูลตัวอย่างที่เราสนใจ ดังนั้นค่าของ $\widehat{MSE}\left(\hat{\mu}_{SRSS}\right)$ ได้ดังนี้

$$\widehat{MSE}\left(\hat{\mu}_{SRSS}\right)\cong v\left(\hat{\mu}_{RSS}\right)+\hat{R}^2v\left(\hat{\mu}_{xRSS}\right)-2\hat{R}\widehat{\operatorname{cov}}\left(\hat{\mu}_{RSS},\hat{\mu}_{xRSS}\right)$$

เนื่องจาก

$$v\left(\hat{\mu}_{RSS}\right)=\frac{1}{kr}\left(\hat{\sigma}_y^2-\frac{1}{k}\sum_{i=1}^{k}\hat{\tau}_{y[i]}^2\right),\ \hat{\tau}_{y[i]}=\hat{\mu}_{y[i]}-\hat{\mu}_{RSS}$$

$$v\left(\hat{\mu}_{xRSS}\right)=\frac{1}{kr}\left(\hat{\sigma}_x^2-\frac{1}{k}\sum_{i=1}^{k}\hat{\tau}_{x(i)}^2\right),\ \hat{\tau}_{x(i)}=\hat{\mu}_{x(i)}-\hat{\mu}_{xRSS}$$

และ $\widehat{\operatorname{cov}}\left(\hat{\mu}_{RSS},\hat{\mu}_{xRSS}\right)=\frac{1}{kr}\left(\hat{\sigma}_{yx}-\frac{1}{k}\sum_{i=1}^{k}\hat{\tau}_{yx(i)}\right),\ \hat{\tau}_{yx(i)}=\left(\hat{\mu}_{y[i]}-\hat{\mu}_{RSS}\right)\left(\hat{\mu}_{x(i)}-\hat{\mu}_{xRSS}\right)$

ดังนั้น สามารถเขียนสมการ $\widehat{MSE}\left(\hat{\mu}_{SRSS}\right)$ ได้ดังนี้

$$\widehat{MSE}\left(\hat{\mu}_{SRSS}\right)\cong\frac{1}{kr}\left(\hat{\sigma}_y^2-2\hat{R}\hat{\sigma}_{yx}+\hat{R}^2\hat{\sigma}_x^2\right)-\frac{1}{k^2r}\left(\sum_{i=1}^{k}\hat{\tau}_{y[i]}^2-2\hat{R}\sum_{i=1}^{k}\hat{\tau}_{yx(i)}+\hat{R}^2\sum_{i=1}^{k}\hat{\tau}_{x(i)}^2\right)$$

2.4 การประมาณพารามิเตอร์ โดยวิธีการประมาณค่าเฉลี่ยแบบอัตราส่วนของ Kadilar ภายใต้แผนการสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง

ในปี ค.ศ.2007 Kadilar, et al. ได้เสนอตัวประมาณแบบอัตราส่วนค่าเฉลี่ยของประชากร ดังนี้

1. ค่าประมาณค่าเฉลี่ยอัตราส่วน คือ

$$\hat{\mu}_{KRSS} = \frac{K\hat{\mu}_{RSS}}{\hat{\mu}_{xRSS}}\mu_x = \hat{R}_{KRSS}\mu_x$$

โดยที่

$$K = \frac{1+\gamma\rho C_y C_x - W_{yx(i)}}{1+\gamma C_y^2 - W_{y[i]}^2}, \gamma = \frac{1}{kr}, C_y = \frac{\sigma_y}{\mu_y}, C_x = \frac{\sigma_x}{\mu_x}, W_{yx(i)} = \frac{1}{k^2 r\mu_x\mu_y}\sum_{i=1}^{k}\tau_{yx(i)},$$

$$W_{y[i]}^2 = \frac{1}{k^2 r\mu_y^2}\sum_{i=1}^{k}\tau_{y[i]}^2, \hat{\mu}_{RSS} = \frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r} y_{[i]j}, \hat{\mu}_{xRSS} = \frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r} x_{(i)j}, \text{ และ } \hat{R}_{KRSS} = \frac{K\hat{\mu}_{RSS}}{\hat{\mu}_{xRSS}}$$

เมื่อ K คือ ค่าคงที่ ซึ่งทำให้ค่า $MSE(\hat{\mu}_{KRSS})$ มีค่าน้อยที่สุด

C_y คือ ค่าสัมประสิทธิความผันแปรของตัวแปรสุ่มที่สนใจศึกษา

C_x คือ ค่าสัมประสิทธิความผันแปรของตัวแปรช่วย

โดยมีค่าเอนเอียงของตัวประมาณของค่าเฉลี่ยอัตราส่วน คือ

$$B(\hat{\mu}_{KRSS}) \cong (K-1)\mu_y - K\mu_y\left[\gamma\left(C_x^2 - C_{yx}\right) - \left(W_{x(i)}^2 - W_{yx(i)}\right)\right]$$

โดยที่

$$C_{yx} = \rho C_y C_x \text{ และ } W_{x(i)}^2 = \frac{1}{k^2 r\mu_x^2}\sum_{i=1}^{k}\tau_{x(i)}^2$$

2. ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของค่าเฉลี่ยอัตราส่วน คือ

$$MSE(\hat{\mu}_{KRSS}) \cong K^2 V(\hat{\mu}_{RSS}) + \mu_y^2 (K-1)^2 + R^2 V(\hat{\mu}_{xRSS}) - 2RK\operatorname{cov}(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS})$$

โดยที่

$$V\left(\hat{\mu}_{RSS}\right)=\frac{1}{kr}\left(\sigma_y^2-\frac{1}{k}\sum_{i=1}^{k}\tau_{y[i]}^2\right), \tau_{y[i]}=\mu_{y[i]}-\mu_y$$

$$V\left(\hat{\mu}_{xRSS}\right)=\frac{1}{kr}\left(\sigma_x^2-\frac{1}{k}\sum_{i=1}^{k}\tau_{x(i)}^2\right), \tau_{x(i)}=\mu_{x(i)}-\mu_x$$

และ $\operatorname{cov}\left(\hat{\mu}_{RSS},\hat{\mu}_{xRSS}\right)=\frac{1}{kr}\left(\sigma_{yx}-\frac{1}{k}\sum_{i=1}^{k}\tau_{yx(i)}\right), \tau_{yx(i)}=\left(\mu_{y[i]}-\mu_y\right)\left(\mu_{x(i)}-\mu_x\right)$

หรือ

$$MSE\left(\hat{\mu}_{KRSS}\right)\cong\frac{1}{kr}\left(K^2\sigma_y^2-2RK\sigma_{yx}+R^2\sigma_x^2\right)+\mu_y^2\left(K-1\right)^2-\frac{1}{k^2r}$$

$$\times\left(K^2\sum_{i=1}^{k}\tau_{y[i]}^2-2RK\sum_{i=1}^{k}\tau_{yx(i)}+R^2\sum_{i=1}^{k}\tau_{x(i)}^2\right)$$

3. ค่าประมาณความคลาดเคลื่อนกำลังสองเฉลี่ยของค่าเฉลี่ยอัตราส่วน คือ

$$\widehat{MSE}\left(\hat{\mu}_{KRSS}\right)\cong\hat{K}^2v\left(\hat{\mu}_{RSS}\right)+\hat{\mu}_y^2\left(\hat{K}-1\right)^2+\hat{R}^2v\left(\hat{\mu}_{xRSS}\right)-2\hat{R}\hat{K}\widehat{\operatorname{cov}}\left(\hat{\mu}_{RSS},\hat{\mu}_{xRSS}\right)$$

โดยที่

$$\hat{K}=\frac{1+\gamma\hat{\rho}\hat{C}_y\hat{C}_x-\hat{W}_{yx(i)}}{1+\gamma\hat{C}_y^2-\hat{W}_{y[i]}^2}$$

เมื่อ $\hat{\rho}=\frac{\hat{\sigma}_{yx}}{\hat{\sigma}_x\hat{\sigma}_y}$, $\hat{C}_y=\frac{\hat{\sigma}_y}{\hat{\mu}_y}$, $\hat{C}_x=\frac{\hat{\sigma}_x}{\hat{\mu}_x}$, $\hat{W}_{yx(i)}=\frac{1}{k^2r\hat{\mu}_x\hat{\mu}_y}\sum_{i=1}^{k}\left(\hat{\mu}_{y[i]}-\hat{\mu}_y\right)\left(\hat{\mu}_{x(i)}-\hat{\mu}_x\right)$,
และ $\hat{W}_{y[i]}^2=\frac{1}{k^2r\hat{\mu}_y^2}\sum_{i=1}^{k}\left(\hat{\mu}_{y[i]}-\hat{\mu}_y\right)^2$

$$v\left(\hat{\mu}_{RSS}\right)=\frac{1}{kr}\left(\hat{\sigma}_y^2-\frac{1}{k}\sum_{i=1}^{k}\hat{\tau}_{y[i]}^2\right), \hat{\tau}_{y[i]}=\hat{\mu}_{y[i]}-\hat{\mu}_{RSS}$$

$$v\left(\hat{\mu}_{xRSS}\right)=\frac{1}{kr}\left(\hat{\sigma}_x^2-\frac{1}{k}\sum_{i=1}^{k}\hat{\tau}_{x(i)}^2\right),\ \hat{\tau}_{x(i)}=\hat{\mu}_{x(i)}-\hat{\mu}_{xRSS}$$

และ $\widehat{\operatorname{cov}}\left(\hat{\mu}_{RSS},\hat{\mu}_{xRSS}\right)=\frac{1}{kr}\left(\hat{\sigma}_{yx}-\frac{1}{k}\sum_{i=1}^{k}\hat{\tau}_{yx(i)}\right),\ \hat{\tau}_{yx(i)}=\left(\hat{\mu}_{y[i]}-\hat{\mu}_{RSS}\right)\left(\hat{\mu}_{x(i)}-\hat{\mu}_{xRSS}\right)$

หรือ

$$\widehat{MSE}\left(\hat{\mu}_{KRSS}\right)\cong\frac{1}{kr}\left(\hat{K}^2\hat{\sigma}_y^2-2\hat{R}\hat{K}\hat{\sigma}_{yx}+\hat{R}^2\hat{\sigma}_x^2\right)+\hat{\mu}_y^2\left(\hat{K}-1\right)^2-\frac{1}{k^2r}$$

$$\times\left(\hat{K}^2\sum_{i=1}^{k}\hat{\tau}_{y[i]}^2-2\hat{R}\hat{K}\sum_{i=1}^{k}\hat{\tau}_{yx(i)}+\hat{R}^2\sum_{i=1}^{k}\hat{\tau}_{x(i)}^2\right)$$

พิสูจน์

1. การพิสูจน์หาค่าเอนเอียงของตัวประมาณของค่าเฉลี่ยอัตราส่วน ได้ดังนี้

เนื่องจาก $e_0=\frac{\hat{\mu}_{RSS}-\mu_y}{\mu_y},\ \hat{\mu}_{RSS}=\mu_y\left(1+e_0\right),\ e_1=\frac{\hat{\mu}_{xRSS}-\mu_x}{\mu_x},\ \hat{\mu}_{xRSS}=\mu_x\left(1+e_1\right),$

$$E\left(e_0\right)=0,\ \ E\left(e_1\right)=0,\ \ E\left(e_0^2\right)=\gamma C_y^2-W_{y[i]}^2,\ \ E\left(e_1^2\right)=\gamma C_x^2-W_{x[i]}^2,\ \ E\left(e_0e_1^2\right)\cong 0,$$

และ $E\left(e_0e_1\right)=\gamma C_{yx}-W_{yx(i)}$

พิจารณา $\hat{\mu}_{KRSS}=\frac{K\hat{\mu}_{RSS}}{\hat{\mu}_{xRSS}}\mu_x=K\frac{\mu_y\left(1+e_0\right)}{\mu_x\left(1+e_1\right)}\mu_x=K\mu_y\left(1+e_0\right)\left(1+e_1\right)^{-1}$

จะได้ว่า $\hat{\mu}_{KRSS}\cong K\mu_y\left(1+e_0\right)\left(1-e_1+e_1^2\right)$

$$=K\mu_y\left(1-e_1+e_1^2+e_0-e_0e_1+e_0e_1^2\right)$$

$$\hat{\mu}_{KRSS}-\mu_y\cong K\mu_y-\mu_y+K\mu_y\left(-e_1+e_1^2+e_0-e_0e_1+e_0e_1^2\right)$$

ดังนั้น ค่าเอนเอียงของตัวประมาณของค่าเฉลี่ยอัตราส่วน คือ

$$B(\hat{\mu}_{KRSS}) \cong (K-1)\mu_y - K\mu_y\left(\gamma C_x^2 - W_{x(i)}^2 - \gamma C_{yx} + W_{yx(i)}\right)$$

$$= (K-1)\mu_y - K\mu_y\left[\gamma\left(C_x^2 - C_{yx}\right) - \left(W_{x(i)}^2 - W_{yx(i)}\right)\right]$$

2. ในทำนองเดียวกันกับการพิสูจน์ $MSE(\hat{\mu}_{SRSS})$ จะสามารถพิสูจน์ $MSE(\hat{\mu}_{KRSS})$ ได้ดังนี้

$$MSE(\hat{\mu}_{KRSS}) = E\left(\hat{R}_{KRSS}\mu_x - R\mu_x\right)^2$$

$$= E\left(\frac{K\hat{\mu}_{RSS}}{\hat{\mu}_{xRSS}}\mu_x - R\mu_x\right)^2$$

$$= \left(\frac{\mu_x}{\hat{\mu}_{xRSS}}\right)^2 E\left(K\hat{\mu}_{RSS} - R\hat{\mu}_{xRSS}\right)^2$$

เนื่องจาก $\frac{1}{\hat{\mu}_{xRSS}} = \frac{1}{\mu_x + \hat{\mu}_{xRSS} - \mu_x} = \frac{1}{\mu_x}\left(1 + \frac{\hat{\mu}_{xRSS} - \mu_x}{\mu_x}\right)^{-1} \cong \frac{1}{\mu_x}$ (โดยใช้ Taylor series)

จะได้ว่า $MSE(\hat{\mu}_{KRSS}) \cong E\left(K\hat{\mu}_{RSS} - R\hat{\mu}_{xRSS}\right)^2$

$$= E\left(K\hat{\mu}_{RSS} - \mu_y + \mu_y - R\hat{\mu}_{xRSS}\right)^2$$

$$= E\left\{\left(K\hat{\mu}_{RSS} - \mu_y\right) - R\left(\hat{\mu}_{xRSS} - \mu_x\right)\right\}^2$$

$$= E\left\{\left(K\hat{\mu}_{RSS} - \mu_y\right)^2 + \left(R\left(\hat{\mu}_{xRSS} - \mu_x\right)\right)^2 - 2R\left(K\hat{\mu}_{RSS} - \mu_y\right)\left(\hat{\mu}_{xRSS} - \mu_x\right)\right\}$$

$$= E\left(K\hat{\mu}_{RSS} - \mu_y\right)^2 + R^2E\left(\hat{\mu}_{xRSS} - \mu_x\right)^2 - 2RE\left(K\hat{\mu}_{RSS} - \mu_y\right)\left(\hat{\mu}_{xRSS} - \mu_x\right)$$

$$= K^2E\left(\hat{\mu}_{RSS}^2\right) + \mu_y^2 - 2K\mu_y E\left(\hat{\mu}_{RSS}\right) + R^2V\left(\hat{\mu}_{xRSS}\right) - 2RK\operatorname{cov}\left(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS}\right)$$

$$= K^2\left(V\left(\hat{\mu}_{RSS}\right) + \mu_y^2\right) + \mu_y^2 - 2K\mu_y^2 + R^2V\left(\hat{\mu}_{xRSS}\right) - 2RK\operatorname{cov}\left(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS}\right)$$

$$= K^2V\left(\hat{\mu}_{RSS}\right) + K^2\mu_y^2 + \mu_y^2 - 2K\mu_y^2 + R^2V\left(\hat{\mu}_{xRSS}\right) - 2RK\operatorname{cov}\left(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS}\right)$$

$$= K^2V\left(\hat{\mu}_{RSS}\right) + \mu_y^2\left(K^2 - 2K + 1\right) + R^2V\left(\hat{\mu}_{xRSS}\right) - 2RK\operatorname{cov}\left(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS}\right)$$

$$= K^2V\left(\hat{\mu}_{RSS}\right) + \mu_y^2\left(K - 1\right)^2 + R^2V\left(\hat{\mu}_{xRSS}\right) - 2RK\operatorname{cov}\left(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS}\right)$$

ดังนั้น $MSE\left(\hat{\mu}_{KRSS}\right) \cong K^2V\left(\hat{\mu}_{RSS}\right) + \mu_y^2\left(K - 1\right)^2 + R^2V\left(\hat{\mu}_{xRSS}\right) - 2RK\operatorname{cov}\left(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS}\right)$
เนื่องจาก

$$V\left(\hat{\mu}_{RSS}\right) = \frac{1}{kr}\left(\sigma_y^2 - \frac{1}{k}\sum_{i=1}^{k}\tau_{y[i]}^2\right), \tau_{y[i]} = \mu_{y[i]} - \mu_y$$

$$V\left(\hat{\mu}_{xRSS}\right) = \frac{1}{kr}\left(\sigma_x^2 - \frac{1}{k}\sum_{i=1}^{k}\tau_{x(i)}^2\right), \tau_{x(i)} = \mu_{x(i)} - \mu_x$$

และ $\operatorname{cov}\left(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS}\right) = \frac{1}{kr}\left(\sigma_{yx} - \frac{1}{k}\sum_{i=1}^{k}\tau_{yx(i)}\right), \tau_{yx(i)} = \left(\mu_{y[i]} - \mu_y\right)\left(\mu_{x(i)} - \mu_x\right)$

ดังนั้น สามารถเขียนสมการ $MSE\left(\hat{\mu}_{KRSS}\right)$ ได้ดังนี้

$$MSE\left(\hat{\mu}_{KRSS}\right) \cong K^2V\left(\hat{\mu}_{RSS}\right) + \mu_y^2\left(K - 1\right)^2 + R^2V\left(\hat{\mu}_{xRSS}\right) - 2RK\operatorname{cov}\left(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS}\right)$$

$$= K^2\left[\frac{1}{kr}\left(\sigma_y^2 - \frac{1}{k}\sum_{i=1}^{k}\tau_{y[i]}^2\right)\right] + \mu_y^2\left(K - 1\right)^2 + R^2\left[\frac{1}{kr}\left(\sigma_x^2 - \frac{1}{k}\sum_{i=1}^{k}\tau_{x(i)}^2\right)\right]$$

$$-2RK\left[\frac{1}{kr}\left(\sigma_{yx} - \frac{1}{k}\sum_{i=1}^{k}\tau_{yx(i)}\right)\right]$$

$$= \frac{1}{kr} K^2 \sigma_y^2 - \frac{1}{k^2 r} K^2 \sum_{i=1}^{k} \tau_{y[i]}^2 + \mu_y^2 (K-1)^2 + \frac{1}{kr} R^2 \sigma_x^2 - \frac{1}{k^2 r} R^2 \sum_{i=1}^{k} \tau_{x(i)}^2$$

$$- \frac{1}{kr} 2RK\sigma_{yx} + \frac{1}{k^2 r} 2RK \sum_{i=1}^{k} \tau_{yx(i)}$$

$$= \frac{1}{kr}\left(K^2\sigma_y^2 - 2RK\sigma_{yx} + R^2\sigma_x^2\right) + \mu_y^2 (K-1)^2 - \frac{1}{k^2 r}$$

$$\times \left(K^2 \sum_{i=1}^{k} \tau_{y[i]}^2 - 2RK \sum_{i=1}^{k} \tau_{yx(i)} + R^2 \sum_{i=1}^{k} \tau_{x(i)}^2 \right)$$

ดังนั้น $MSE\left(\hat{\mu}_{KRSS}\right) \cong \frac{1}{kr}\left(K^2\sigma_y^2 - 2RK\sigma_{yx} + R^2\sigma_x^2\right) + \mu_y^2 (K-1)^2 - \frac{1}{k^2 r}$

$$\times \left(K^2 \sum_{i=1}^{k} \tau_{y[i]}^2 - 2RK \sum_{i=1}^{k} \tau_{yx(i)} + R^2 \sum_{i=1}^{k} \tau_{x(i)}^2 \right)$$

การคำนวณหาค่า K ที่ทำให้ MSE มีค่าน้อยที่สุด สามารถกระทำได้โดยการอนุพันธ์ค่าของ $MSE\left(\hat{\mu}_{KRSS}\right)$ เทียบกับ K มีค่าเท่ากับ 0 หรือ $\frac{\partial MSE\left(\hat{\mu}_{KRSS}\right)}{\partial K} = 0$ **และอนุพันธ์อันดับที่ สองของ** $MSE\left(\hat{\mu}_{KRSS}\right)$ เทียบกับ K มีค่ามากกว่า 0 หรือ $\frac{\partial^2 MSE\left(\hat{\mu}_{KRSS}\right)}{\partial K^2} > 0$ ซึ่งผลการจากอนุพันธ์แสดงได้ดังนี้

พิสูจน์ว่า $\frac{\partial MSE\left(\hat{\mu}_{KRSS}\right)}{\partial K} = 0$

จากสูตร

$$MSE\left(\hat{\mu}_{KRSS}\right) \cong \frac{1}{kr}\left(K^2\sigma_y^2 - 2RK\sigma_{yx} + R^2\sigma_x^2\right) + \mu_y^2 (K-1)^2 - \frac{1}{k^2 r}$$

$$\times \left(K^2 \sum_{i=1}^{k} \tau_{y[i]}^2 - 2RK \sum_{i=1}^{k} \tau_{yx(i)} + R^2 \sum_{i=1}^{k} \tau_{x(i)}^2 \right)$$

เนื่องจาก

$$\sigma_y = C_y\mu_y,\ R = \frac{\mu_y}{\mu_x},\ \sigma_{yx} = \rho\sigma_x\sigma_y,\ \sigma_x = C_x\mu_x,\ W_{y[i]}^2 = \frac{1}{k^2 r \mu_y^2} \sum_{i=1}^{k} \tau_{y[i]}^2,$$

$$W_{yx(i)} = \frac{1}{k^2 r \mu_x \mu_y} \sum_{i=1}^{k} \tau_{yx(i)} \text{ และ } W_{x(i)}^2 = \frac{1}{k^2 r \mu_x^2} \sum_{i=1}^{k} \tau_{x(i)}^2$$

แทนค่า $\sigma_y, R, \sigma_{xy}, \sigma_x, W_{y[i]}^2, W_{yx(i)}$ และ $W_{x(i)}^2$ ลงใน $MSE(\hat{\mu}_{KRSS})$ ได้ดังนี้

$$MSE(\hat{\mu}_{KRSS}) \cong \frac{1}{kr}\left(K^2\sigma_y^2 - 2RK\sigma_{yx} + R^2\sigma_x^2\right) + \mu_y^2(K-1)^2 - \frac{1}{k^2 r}$$

$$\times\left(K^2\sum_{i=1}^{k}\tau_{y[i]}^2 - 2RK\sum_{i=1}^{k}\tau_{yx(i)} + R^2\sum_{i=1}^{k}\tau_{x(i)}^2\right)$$

$$= \frac{1}{kr}\left[K^2\left(C_y\mu_y\right)^2 - 2K\left(\frac{\mu_y}{\mu_x}\right)\left(\rho C_x\mu_x C_y\mu_y\right) + \left(\frac{\mu_y}{\mu_x}\right)^2\left(C_x\mu_x\right)^2\right]$$

$$+\mu_y^2(K-1)^2 - \frac{1}{k^2 r}K^2\sum_{i=1}^{k}\tau_{y[i]}^2 + \frac{1}{k^2 r}2K\left(\frac{\mu_y}{\mu_x}\right)\sum_{i=1}^{k}\tau_{yx(i)}$$

$$-\frac{1}{k^2 r}\left(\frac{\mu_y}{\mu_x}\right)^2\sum_{i=1}^{k}\tau_{x(i)}^2$$

$$= \mu_y^2\left[\frac{1}{kr}\left(K^2C_y^2 - 2K\rho C_y C_x + C_x^2\right) + (K-1)^2\right.$$

$$\left.-\left(K^2W_{y[i]}^2 - 2KW_{yx(i)} + W_{x(i)}^2\right)\right]$$

ดังนั้น $MSE(\hat{\mu}_{KRSS}) \cong \mu_y^2\left[\frac{1}{kr}\left(K^2C_y^2 - 2K\rho C_y C_x + C_x^2\right) + (K-1)^2\right.$

$$\left.-\left(K^2W_{y[i]}^2 - 2KW_{yx(i)} + W_{x(i)}^2\right)\right]$$

พิจารณาหา K ได้ดังนี้

$$\frac{\partial MSE(\hat{\mu}_{KRSS})}{\partial K} = 0$$

$$\frac{1}{kr}\left(2KC_y^2-2\rho C_yC_x\right)+2K-2-2KW_{y[i]}^2+2W_{yx(i)}=0$$

$$2\gamma KC_y^2+2K-2KW_{y[i]}^2=2\gamma\rho C_yC_x-2W_{yx(i)}+2$$

$$K\left(1+\gamma C_y^2-W_{y[i]}^2\right)=1+\gamma\rho C_yC_x-W_{yx(i)}$$

$$K=\frac{1+\gamma\rho C_yC_x-W_{yx(i)}}{1+\gamma C_y^2-W_{y[i]}^2}$$

ดังนั้น $K=\dfrac{1+\gamma\rho C_yC_x-W_{yx(i)}}{1+\gamma C_y^2-W_{y[i]}^2}$

พิสูจน์ว่า $\dfrac{\partial^2 MSE\left(\hat{\mu}_{KRSS}\right)}{\partial K^2}>0$

เนื่องจาก $\dfrac{\partial MSE\left(\hat{\mu}_{KRSS}\right)}{\partial K}=\dfrac{1}{kr}\left(2KC_y^2-2\rho C_yC_x\right)+2K-2-2KW_{y[i]}^2+2W_{yx(i)}$

$$=2\left(K+\frac{KC_y^2}{kr}-KW_{y[i]}^2\right)-\frac{2\rho C_yC_x}{kr}-2+2W_{yx(i)}$$

ดังนั้น $\dfrac{\partial^2 MSE\left(\hat{\mu}_{KRSS}\right)}{\partial K^2}=2\left(1+\dfrac{C_y^2}{kr}-W_{y[i]}^2\right)$

$$=2\left[1+\frac{1}{kr}\left(\frac{\sigma_y}{\mu_y}\right)^2-\frac{1}{k^2r\mu_y^2}\sum_{i=1}^{k}\left(\mu_{y[i]}-\mu_y\right)^2\right]$$

$$=2\left\{1+\frac{1}{\mu_y^2}\left\{\frac{1}{kr}\left[\sigma_y^2-\frac{1}{k}\sum_{i=1}^{k}\left(\mu_{y[i]}-\mu_y\right)^2\right]\right\}\right\}$$

$$= 2\left[1 + \frac{1}{\mu_y^2}\left(V\left(\hat{\mu}_{RSS}\right)\right)\right]$$

จะเห็นได้ว่า $\frac{\partial^2 MSE\left(\hat{\mu}_{KRSS}\right)}{\partial K^2} = 2\left[1 + \frac{1}{\mu_y^2}\left(V\left(\hat{\mu}_{RSS}\right)\right)\right] > 0$ ดังนั้นสามารถสรุปได้ว่าค่า $MSE\left(\hat{\mu}_{KRSS}\right)$ มีค่าน้อยที่สุด

3. เมื่อพิจารณาค่าประมาณความคลาดเคลื่อนกำลังสองเฉลี่ยของค่าเฉลี่ยอัตราส่วน $\left(\widehat{MSE}\left(\hat{\mu}_{KRSS}\right)\right)$ ซึ่งการประมาณค่าเป็นการประมาณโดยอาศัยข้อมูลตัวอย่างที่เราสนใจ ดังนั้นค่าของ $\widehat{MSE}\left(\hat{\mu}_{KRSS}\right)$ ได้ดังนี้

$$\widehat{MSE}\left(\hat{\mu}_{KRSS}\right) \cong \hat{K}^2 v\left(\hat{\mu}_{RSS}\right) + \hat{\mu}_y^2\left(\hat{K}-1\right)^2 + \hat{R}^2 v\left(\hat{\mu}_{xRSS}\right) - 2\hat{R}\hat{K}\widehat{\mathrm{cov}}\left(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS}\right)$$

เนื่องจาก

$$\hat{K} = \frac{1 + \gamma\hat{\rho}\hat{C}_y\hat{C}_x - \hat{W}_{yx(i)}}{1 + \gamma\hat{C}_y^2 - \hat{W}_{y[i]}^2}$$

เมื่อ $\hat{\rho} = \frac{\hat{\sigma}_{yx}}{\hat{\sigma}_x\hat{\sigma}_y}$, $\hat{C}_y = \frac{\hat{\sigma}_y}{\hat{\mu}_y}$, $\hat{C}_x = \frac{\hat{\sigma}_x}{\hat{\mu}_x}$, $\hat{W}_{yx(i)} = \frac{1}{k^2 r\hat{\mu}_x\hat{\mu}_y}\sum_{i=1}^{k}\left(\hat{\mu}_{y[i]} - \hat{\mu}_y\right)\left(\hat{\mu}_{x(i)} - \hat{\mu}_x\right)$,

และ $\hat{W}_{y[i]}^2 = \frac{1}{k^2 r\hat{\mu}_y^2}\sum_{i=1}^{k}\left(\hat{\mu}_{y[i]} - \hat{\mu}_y\right)^2$

$$v\left(\hat{\mu}_{RSS}\right) = \frac{1}{kr}\left(\hat{\sigma}_y^2 - \frac{1}{k}\sum_{i=1}^{k}\hat{\tau}_{y[i]}^2\right), \hat{\tau}_{y[i]} = \hat{\mu}_{y[i]} - \hat{\mu}_{RSS}$$

$$v\left(\hat{\mu}_{xRSS}\right) = \frac{1}{kr}\left(\hat{\sigma}_x^2 - \frac{1}{k}\sum_{i=1}^{k}\hat{\tau}_{x(i)}^2\right), \hat{\tau}_{x(i)} = \hat{\mu}_{x(i)} - \hat{\mu}_{xRSS}$$

และ $\widehat{\mathrm{cov}}\left(\hat{\mu}_{RSS}, \hat{\mu}_{xRSS}\right) = \frac{1}{kr}\left(\hat{\sigma}_{yx} - \frac{1}{k}\sum_{i=1}^{k}\hat{\tau}_{yx(i)}\right), \hat{\tau}_{yx(i)} = \left(\hat{\mu}_{y[i]} - \hat{\mu}_{RSS}\right)\left(\hat{\mu}_{x(i)} - \hat{\mu}_{xRSS}\right)$

ดังนั้น สามารถเขียนสมการ $\widehat{MSE}(\hat{\mu}_{KRSS})$ ได้ดังนี้

$$\widehat{MSE}(\hat{\mu}_{KRSS}) \cong \frac{1}{kr}\left(\hat{K}^2\hat{\sigma}_y^2 - 2\hat{R}\hat{K}\hat{\sigma}_{yx} + \hat{R}^2\hat{\sigma}_x^2\right) + \hat{\mu}_y^2\left(\hat{K}-1\right)^2 - \frac{1}{k^2r}$$

$$\times\left(\hat{K}^2\sum_{i=1}^{k}\hat{\tau}_{y[i]}^2 - 2\hat{R}\hat{K}\sum_{i=1}^{k}\hat{\tau}_{yx(i)} + \hat{R}^2\sum_{i=1}^{k}\hat{\tau}_{x(i)}^2\right)$$

2.5 การประมาณพารามิเตอร์ โดยวิธีการประมาณค่าเฉลี่ยแบบถดถอยของ Yu and Lam ภายใต้แผนการสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง

ในปี ค.ศ. 1997 Yu and Lam ได้เสนอสมการการถดถอยของ y บน x เป็นเชิงเส้น เมื่อทราบค่า μ_x ดังนี้

$$y = \mu_y + \rho\frac{\sigma_y}{\sigma_x}(x-\mu_x) + \varepsilon \tag{10}$$

เมื่อ x และ ε เป็นอิสระกัน และ ε มีค่าเฉลี่ยเป็น 0 และค่าความแปรปรวนเป็น $\sigma_y^2\left(1-\rho^2\right)$

กำหนดให้ $x_{(i)j}$ คือ ค่าของ x ที่น้อยที่สุดของลำดับที่ i ในรอบที่ j

$y_{[i]j}$ คือ ค่าที่สอดคล้องของ y จากตัวอย่างที่ i ในรอบที่ j

จากสมการ (10) สามารถเขียนเป็นสมการได้ดังนี้

$$y_{[i]j} = \mu_y + \rho\frac{\sigma_y}{\sigma_x}\left(x_{(i)j}-\mu_x\right) + \varepsilon_{[i]j} \qquad ; i = 1,2,\ldots,k,\ j = 1,2,\ldots,r \tag{11}$$

เมื่อทราบค่า μ_x เราพิจารณาตัวประมาณค่าความแตกต่าง

$$\overline{y}_d = \hat{\mu}_{RSS} + \beta\left(\mu_x - \hat{\mu}_{xRSS}\right)$$

โดยที่ $\hat{\mu}_{RSS} = \frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r}y_{[i]j}$, $\hat{\mu}_{xRSS} = \frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r}x_{(i)j}$ และ β เป็นค่าคงที่

เพราะว่า $\hat{\mu}_{RSS}$ และ $\hat{\mu}_{xRSS}$ เป็นตัวประมาณค่าที่ไม่เอนเอียง ดังนั้น $\overline{y}_d$ ก็เป็นตัวประมาณค่าที่ไม่เอนเอียงด้วย จะได้ว่า $V\left(\overline{y}_d\right)$ คือ

$$V\left(\overline{y}_d\right) = E\left(\overline{y}_d - \mu_y\right)^2$$

$$= E\left[\hat{\mu}_{RSS} - \beta\left(\hat{\mu}_{xRSS} - \mu_x\right) - \mu_y\right]^2$$

$$= E\left[\left(\hat{\mu}_{RSS} - \mu_y\right) - \beta\left(\hat{\mu}_{xRSS} - \mu_x\right)\right]^2$$

$$= \beta^2 V\left(\hat{\mu}_{xRSS}\right) - 2\beta\rho\frac{\sigma_y}{\sigma_x}V\left(\hat{\mu}_{xRSS}\right) + V\left(\hat{\mu}_{RSS}\right)$$

ดังนั้น $$V\left(\overline{y}_d\right) = \beta^2 V\left(\hat{\mu}_{xRSS}\right) - 2\beta\rho\frac{\sigma_y}{\sigma_x}V\left(\hat{\mu}_{xRSS}\right) + V\left(\hat{\mu}_{RSS}\right)$$

โดยที่

$$V\left(\hat{\mu}_{xRSS}\right) = \frac{\sigma_x^2}{kr}\left[1 - \frac{1}{k}\sum_{i=1}^{k}\left(\frac{E\left(x_{(i)j}\right) - \mu_x}{\sigma_x}\right)^2\right]$$

และ $$V\left(\hat{\mu}_{RSS}\right) = \frac{\sigma_y^2}{kr}\left[1 - \frac{\rho^2}{k}\sum_{i=1}^{k}\left(\frac{E\left(x_{(i)j}\right) - \mu_x}{\sigma_x}\right)^2\right]$$

ที่มา: Yu and Lam (1997, p.1071)

กรณีไม่ทราบ β จะประมาณด้วย $\hat{\beta}$

$$\hat{\beta} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{r}\left(x_{(i)j} - \hat{\mu}_{xRSS}\right)\left(y_{[i]j} - \hat{\mu}_{RSS}\right)}{\sum_{i=1}^{k}\sum_{j=1}^{r}\left(x_{(i)j} - \hat{\mu}_{xRSS}\right)^2}$$

1. เมื่อทราบค่า μ_x ตัวประมาณแบบถดถอยในการสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่างของ μ_y คือ

$$\hat{\mu}_{LRSS} = \hat{\mu}_{RSS} + \hat{\beta}\left(\mu_x - \hat{\mu}_{xRSS}\right)$$

โดยที่

$$\hat{\mu}_{RSS} = \frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r} y_{[i]j},\ \hat{\mu}_{xRSS} = \frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r} x_{(i)j}$$

และ

$$\hat{\beta} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{r}\left(x_{(i)j} - \hat{\mu}_{xRSS}\right)\left(y_{[i]j} - \hat{\mu}_{RSS}\right)}{\sum_{i=1}^{k}\sum_{j=1}^{r}\left(x_{(i)j} - \hat{\mu}_{xRSS}\right)^2}$$

2. ค่าความแปรปรวนของค่าประมาณค่าเฉลี่ยแบบถดถอย คือ

$$V\left(\hat{\mu}_{LRSS}\right) = \frac{\sigma_y^2}{kr}\left(1-\rho^2\right)\left[1 + E_1\left(\frac{\bar{z}_{RSS}^2}{\sigma_z^2}\right)\right]$$

โดยที่

$$z_{(i)j} = \frac{x_{(i)j} - \mu_x}{\sigma_x},\ \bar{z}_{RSS} = \frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r} z_{(i)j} \text{ และ } \sigma_z^2 = \frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r}\left(z_{(i)j} - \bar{z}_{RSS}\right)^2$$

3. ค่าประมาณความแปรปรวนของค่าประมาณค่าเฉลี่ยแบบถดถอย คือ

$$v\left(\hat{\mu}_{LRSS}\right) = \frac{\hat{\sigma}_y^2}{kr}\left(1-\hat{\rho}^2\right)\left(1 + \frac{\hat{\bar{z}}_{RSS}^2}{\hat{\sigma}_z^2}\right)$$

โดยที่

$$\hat{z}_{(i)j} = \frac{x_{(i)j} - \hat{\mu}_x}{\hat{\sigma}_x},\ \hat{\bar{z}}_{RSS} = \frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r} \hat{z}_{(i)j} \text{ และ } \hat{\sigma}_z^2 = \frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r}\left(\hat{z}_{(i)j} - \hat{\bar{z}}_{RSS}\right)^2$$

พิสูจน์

จากสมการของ $\hat{\mu}_{LRSS}$ ที่กำหนดไว้ มีคุณสมบัติดังต่อไปนี้

1. การพิสูจน์หา $E\left(\hat{\mu}_{LRSS}\right)$ ได้ดังนี้

จะได้ว่า

$$E(\hat{\mu}_{LRSS}) = E_1E_2\left[\hat{\mu}_{RSS} + \hat{\beta}\left(\mu_x - \hat{\mu}_{xRSS}\right) \mid \underset{\sim}{x}, \underset{\sim}{y}\right]$$

$$= E_1\left[E_2(\hat{\mu}_{RSS} \mid \underset{\sim}{x}, \underset{\sim}{y}) + (\mu_x - \hat{\mu}_{xRSS}) E_2\left(\hat{\beta} \mid \underset{\sim}{x}, \underset{\sim}{y}\right)\right]$$

$$= E_1\left[\mu_y + \beta(\mu_x - \hat{\mu}_{xRSS})\right]$$

$$= E_1(\mu_y) + \beta E_1(\mu_x - \hat{\mu}_{xRSS})$$

$$= \mu_y + \beta(\mu_x - \mu_x)$$

$$= \mu_y$$

ดังนั้น $\hat{\mu}_{LRSS}$ เป็นค่าประมาณที่ไม่เอนเอียงสำหรับค่าเฉลี่ยประชากรของ μ_y

2. การพิสูจน์หา $V(\hat{\mu}_{LRSS})$ ได้ดังนี้

จะได้ว่า

$$V(\hat{\mu}_{LRSS}) = E_1\left[V_2(\hat{\mu}_{LRSS}) \mid \underset{\sim}{x}, \underset{\sim}{y}\right] + V_1\left[E_2(\hat{\mu}_{LRSS}) \mid \underset{\sim}{x}, \underset{\sim}{y}\right]$$

$$= E_1\left[E_2\left(\hat{\mu}_{LRSS} - \mu_y\right)^2 \mid \underset{\sim}{x}, \underset{\sim}{y}\right] + V_1(\mu_y)$$

$$= E_1\left[E_2\left(\hat{\mu}_{RSS} + \hat{\beta}(\mu_x - \hat{\mu}_{xRSS}) - \mu_y\right)^2\right]$$

$$= E_1\left[E_2\left((\hat{\mu}_{RSS} - \mu_y) - \hat{\beta}(\hat{\mu}_{xRSS} - \mu_x)\right)^2\right]$$

$$= E_1\left[\frac{\sigma_y^2}{kr} + \hat{\beta}^2\left(\frac{\sigma_x^2}{kr}\right) - 2\hat{\beta}\rho\frac{\sigma_y\sigma_x}{kr}\right]$$

$$= \frac{\sigma_y^2}{kr} + \left(\frac{\sigma_x^2}{kr}\right)E_1\left(\hat{\beta}^2\right) - 2\rho\frac{\sigma_y\sigma_x}{kr}E_1\left(\hat{\beta}\right) \qquad ; E_1\left(\hat{\beta}\right) = \beta$$

$$= \frac{\sigma_y^2}{kr}\left[1+\frac{\sigma_x^2}{\sigma_y^2}\left(\frac{\sigma_y^2\left(1-\rho^2\right)}{\sigma_x^2}E_1\left(\frac{\overline{z}_{RSS}}{\sigma_z^2}\right)+\beta^2\right)-2\rho\frac{\sigma_y\sigma_x}{\sigma_y^2}\beta\right]$$

$$= \frac{\sigma_y^2}{kr}\left[1+\left(1-\rho^2\right)E_1\left(\frac{\overline{z}_{RSS}}{\sigma_z^2}\right)+\frac{\sigma_x^2}{\sigma_y^2}\left(\rho^2\frac{\sigma_y^2}{\sigma_x^2}\right)-2\rho\frac{\sigma_y\sigma_x}{\sigma_y^2}\left(\rho\frac{\sigma_y}{\sigma_x}\right)\right]$$

$$= \frac{\sigma_y^2}{kr}\left[1+\left(1-\rho^2\right)E_1\left(\frac{\overline{z}_{RSS}}{\sigma_z^2}\right)+\rho^2-2\rho^2\right]$$

$$= \frac{\sigma_y^2}{kr}\left[1-\rho^2+\left(1-\rho^2\right)E_1\left(\frac{\overline{z}_{RSS}}{\sigma_z^2}\right)\right]$$

$$= \frac{\sigma_y^2}{kr}\left(1-\rho^2\right)\left[1+E_1\left(\frac{\overline{z}_{RSS}^2}{\sigma_z^2}\right)\right]$$

เมื่อ $V_1\left(\hat{\beta}\right)=\frac{\sigma_y^2}{\sigma_x^2}\left(1-\rho^2\right)E_1\left(\frac{\overline{z}_{RSS}}{\sigma_z^2}\right)$

โดยที่

$z_{(i)j}=\frac{x_{(i)j}-\mu_x}{\sigma_x}$, $\overline{z}_{RSS}=\frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r}z_{(i)j}$ และ $\sigma_z^2=\frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r}\left(z_{(i)j}-\overline{z}_{RSS}\right)^2$

ดังนั้น $V\left(\hat{\mu}_{LRSS}\right)=\frac{\sigma_y^2}{kr}\left(1-\rho^2\right)\left[1+E_1\left(\frac{\overline{z}_{RSS}^2}{\sigma_z^2}\right)\right]$

3. เมื่อพิจารณาค่าประมาณความแปรปรวนของค่าประมาณค่าเฉลี่ย $\left(v\left(\hat{\mu}_{LRSS}\right)\right)$ จะได้ว่า

$$v\left(\hat{\mu}_{LRSS}\right)=\frac{\hat{\sigma}_y^2}{kr}\left(1-\hat{\rho}^2\right)\left(1+\frac{\hat{\overline{z}}_{RSS}^2}{\hat{\sigma}_z^2}\right)$$

โดยที่

$$\hat{z}_{(i)j}=\frac{x_{(i)j}-\hat{\mu}_x}{\hat{\sigma}_x},\ \hat{\overline{z}}_{RSS}=\frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r}\hat{z}_{(i)j}\ \text{และ}\ \hat{\sigma}_z^2=\frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r}\left(\hat{z}_{(i)j}-\hat{\overline{z}}_{RSS}\right)^2$$

งานวิจัยที่เกี่ยวข้อง

ผู้วิจัยทำการศึกษางานวิจัยที่เกี่ยวข้องกับวิธีการประมาณค่าเฉลี่ย โดยสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง ดังนี้

จิรชัย แสงไพศาล (2547) ทำการศึกษาความแม่นยำของตัวประมาณค่าเฉลี่ยของประชากรด้วยการสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง ศึกษาภายใต้ข้อกำหนดว่าไม่มีความคลาดเคลื่อนในการให้ลำดับที่กับหน่วยตัวอย่าง โดยทำการเปรียบเทียบกับการสุ่มตัวอย่างอย่างง่าย จากประชากรที่ศึกษามีการแจกแจงแบบปกติมาตรฐาน การแจกแจงแบบยูนิฟอร์มที่มีพารามิเตอร์ คือ $\beta = 2, 3, 4, 5$ และการแจกแจงแบบเอกซ์โพเนนเชียล ที่มีพารามิเตอร์ คือ $\lambda = 2, 3, 4, 5$ เมื่อกำหนดขนาดของชุดตัวอย่าง (m) ตั้งแต่ 5 ถึง 10 และกำหนดจำนวนรอบที่ทำการสุ่มตัวอย่างซ้ำ (r) ตั้งแต่ 2 ถึง 5 ในแต่ละสถานการณ์ที่ศึกษาทำการทดลองซ้ำทั้งหมด 500 รอบ เกณฑ์การเปรียบเทียบการประมาณค่าเฉลี่ยพิจารณาจากค่าความแม่นยำสัมพัทธ์ (Relative Precision: RP) ในการประมาณค่าเฉลี่ยประชากรเมื่อเทียบกับวิธีการสุ่มตัวอย่างอย่างง่าย และค่าประสิทธิภาพสัมพัทธ์ (Relative Efficiency: RE) เมื่อสุ่มตัวอย่างจากการแจกแจงของประชากรทั้ง 3 แบบ ผลปรากฏว่า ในทุกการแจกแจงเกิดทำนองเดียวกัน คือ ขนาดของชุดตัวอย่าง (m) มีอิทธิพลต่อความแม่นยำในการประมาณค่าเฉลี่ยของประชากรมากกว่าจำนวนรอบที่สุ่มตัวอย่างซ้ำ (r) เมื่อพิจารณาการศึกษาโดยการจำลองแบบ พบว่า การสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่างมีประสิทธิภาพมากกว่าการสุ่มตัวอย่างอย่างง่าย และเมื่อพิจารณาการศึกษาความแม่นยำในการประมาณค่าเฉลี่ยของประชากรโดยวิธีการสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง ในกรณีขนาดตัวอย่างเท่ากัน พบว่า ประชากรที่มีการแจกแจงแบบยูนิฟอร์ม มีความแม่นยำสูงกว่าการแจกแจงแบบปกติมาตร และการแจกแจงแบบเอกซ์โพเนนเชียล ตามลำดับ

นวภา ปลุกปลื้ม (2549) ทำการศึกษาประสิทธิภาพของวิธีการสุ่มตัวอย่างแบบทั่วไปของการเลือกลำดับที่ของชุดตัวอย่าง (GRSS) เมื่อจัดลำดับสมบูรณ์ โดยทำการเปรียบเทียบกับตัวประมาณค่าเฉลี่ยของประชากรด้วยวิธีการสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง การสุ่มตัวอย่างแบบกลุ่ม (Cluster) การสุ่มตัวอย่างแบบแบ่งชั้นภูมิ (Stratified) และการสุ่มตัวอย่างอย่างง่าย โดยประชากรที่ศึกษามีการแจกแจง 3 ลักษณะ คือ การแจกแจงแบบปกติมาตรฐาน การแจกแจงแบบสม่ำเสมอต่อเนื่องในช่วง (0,1) และการแจกแจงแบบเอ็กซ์โพเนนเชียล ที่มีพารามิเตอร์ คือ $\lambda = 0.5, 1$ และ 3 ภายใต้ข้อจำกัดว่าไม่มีความคลาดเคลื่อนในการให้ลำดับที่กับหน่วยตัวอย่าง เมื่อกำหนดขนาดของชุดตัวอย่าง (m) มีค่าเท่ากับ 2, 3, 4, 5 และ 10 จำนวนรอบที่ทำการสุ่มตัวอย่างซ้ำ (r) มีค่าเท่ากับ 1, 2, 3, 4, 5 และ 10 และจำนวนหน่วยตัวอย่างที่เลือกจากแต่ละชุดตัวอย่าง (τ) มี 2 ระดับ คือ 1 และ 2 ในแต่ละสถานการณ์ที่ศึกษาทำการทดลองซ้ำทั้งหมด 500

รอบ เกณฑ์การเปรียบเทียบประสิทธิภาพของตัวประมาณค่าเฉลี่ยพิจารณาจากค่าประสิทธิภาพสัมพัทธ์ (RE) จากการสุ่มประชากรทั้ง 3 ลักษณะ ผลปรากฏว่า ค่าเฉลี่ยที่ประมาณได้จากการสุ่มตัวอย่างทั้ง 5 วิธี เป็นตัวประมาณค่าที่ไม่เอนเอียง (Unbiased Estimator) เมื่อจำนวนชุดตัวอย่าง m เพิ่มขึ้น จำนวนรอบการสุ่มตัวอย่างซ้ำ r คงที่ หรือ จำนวนรอบการสุ่มตัวอย่างซ้ำ r เพิ่มขึ้น จำนวนชุดตัวอย่าง m คงที่ ค่าความแปรปรวนของค่าเฉลี่ยของตัวอย่างมีค่าลดลง และประชากรมีการแจกแจงแบบสมมาตร วิธีการสุ่มตัวอย่างแบบทั่วไปของการเลือกลำดับที่ของชุดตัวอย่างมีประสิทธิภาพมากกว่าวิธีการสุ่มตัวอย่างอื่น ๆ ค่อนข้างสูง แต่เมื่อประชากรมีการแจกแจงแบบเบ้ วิธีการสุ่มตัวอย่างแบบทั่วไปของการเลือกลำดับที่ของชุดตัวอย่างมีประสิทธิภาพมากกว่าวิธีการสุ่มตัวอย่างอื่น ๆ ค่อนข้างน้อย

ณัฐสวัชร์ ถาวรธิดา (2552) ทำการศึกษาเปรียบเทียบแผนภูมิควบคุมค่าเฉลี่ยที่ได้จากการสุ่มตัวอย่างอย่างง่าย (SRS) มาสร้างแผนภูมิควบคุมค่าเฉลี่ย $(\bar{x}-chart)$ การสุ่มตัวอย่างแบบกลุ่มลำดับ (RSS) มาสร้างแผนภูมิควบคุมค่าเฉลี่ย $(\bar{x}-RSS\ chart)$ และแผนภูมิควบคุมค่ามัธยฐาน (MRSS chart) และการสุ่มตัวอย่างแบบกลุ่มมัธยฐาน (Median-Set Sampling: MSS) มาสร้างแผนภูมิควบคุมค่าเฉลี่ย $(\bar{x}-MSS\ chart)$ และแผนภูมิควบคุมค่ามัธยฐาน (MMSS chart) โดยประชากรที่ศึกษามีการแจกแจงความน่าจะเป็นแบบปกติ ที่มีค่าเฉลี่ย (μ) เท่ากับ 50 และส่วนเบี่ยงเบนมาตรฐาน (σ) เท่ากับ 1, 3 และ 5 กระบวนการผลิตมีการเปลี่ยนแปลง (δ) เท่ากับ 0, 0.25, 0.5, 0.75, 1, 3 และ 5 เกณฑ์การตัดสินใจ คือ ค่าความยาววิ่งเฉลี่ย (The Average Run Length: ARL) ผลปรากฏว่าค่าความยาววิ่งเฉลี่ย (ARL) ของแผนภูมิควบคุมค่ามัธยฐานจากสถิติลำดับของการสุ่มตัวอย่างแบบกลุ่มลำดับ (MRSS chart) จะดีกว่าแผนภูมิควบคุมอื่น ๆ ในกรณีที่มีค่าเฉลี่ยของกระบวนการผลิตไม่มีการเปลี่ยนแปลง และแผนภูมิควบคุมค่ามัธยฐานจากสถิติลำดับของการสุ่มตัวอย่างแบบกลุ่มมัธยฐาน (MMSS chart) จะดีกว่าแผนภูมิควบคุมอื่น ๆ ในกรณีที่มีค่าเฉลี่ยของกระบวนการผลิตมีการเปลี่ยนแปลงเล็กน้อย ส่วนกรณีที่ค่าเฉลี่ยของกระบวนการผลิตมีการเปลี่ยนแปลงระดับสูง ในกระบวนการผลิตมีส่วนเบี่ยงเบนมาตรฐานต่ำ แผนภูมิควบคุมทั้ง 5 ชนิดมีประสิทธิภาพพอ ๆ กัน และในเกือบทุกกรณีแผนภูมิควบคุมค่าเฉลี่ยจากการสุ่มตัวอย่างแบบอย่างง่าย $(\bar{x}-chart)$ มีประสิทธิภาพด้อยกว่าแผนภูมิควบคุมแบบอื่น ๆ

Yu and Lam (1997) ทำการศึกษาเปรียบเทียบประสิทธิภาพของตัวประมาณค่าเฉลี่ยแบบถดถอย ภายใต้การสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง กำหนดข้อสมมติว่าตัวแปร X และ Y มีการแจกแจงแบบปกติ ทำการศึกษาในกรณีการเรียงลำดับของตัวแปรช่วย X โดย

เปรียบเทียบประสิทธิภาพของตัวประมาณค่าเฉลี่ยแบบถดถอย ภายใต้การสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง ตัวประมาณค่าเฉลี่ยแบบง่าย ภายใต้การสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง และตัวประมาณค่าเฉลี่ยแบบถดถอย ภายใต้การสุ่มตัวอย่างอย่างง่าย เกณฑ์การเปรียบเทียบตัวประมาณค่าเฉลี่ยพิจารณาจากค่าความแม่นยำสัมพัทธ์ (RP) พบว่า ตัวประมาณค่าเฉลี่ยแบบถดถอย ภายใต้การสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง มีประสิทธิภาพมากกว่าตัวประมาณค่าเฉลี่ยแบบง่าย ภายใต้การสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง ยกเว้นความสัมพันธ์ระหว่าง X และ Y ระดับต่ำ $(|\rho| < 0.4)$ และ ทุกระดับความสัมพันธ์ ตัวประมาณค่าเฉลี่ยแบบถดถอย ภายใต้การสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง มีประสิทธิภาพมากกว่าตัวประมาณค่าเฉลี่ยแบบถดถอย ภายใต้การสุ่มตัวอย่างอย่างง่าย

Kadilar, et al. (2007) ทำการศึกษาเปรียบเทียบประสิทธิภาพของตัวประมาณค่าเฉลี่ยแบบอัตราส่วนสำหรับค่าเฉลี่ยประชากร โดยใช้แผนการสุ่มตัวอย่างแบบเลือกลำดับที่ของชุดตัวอย่าง กำหนดประชากรที่ศึกษามีการแจกแจงแบบปกติ ทำการศึกษาในกรณีการเรียงลำดับของตัวแปรช่วย X โดยเปรียบเทียบประสิทธิภาพของตัวประมาณค่าเฉลี่ยแบบอัตราส่วนแบบดั้งเดิม ตัวประมาณค่าเฉลี่ยแบบอัตราส่วนของ Prasad (1989) ตัวประมาณค่าเฉลี่ยแบบอัตราส่วนของ Samawi and Muttlak (1996) และตัวประมาณค่าเฉลี่ยแบบอัตราส่วนของ Kadilar (2007) กำหนดให้ขนาดของชุดตัวอย่าง (m) มีค่าเท่ากับ 3 และกำหนดจำนวนรอบที่ทำการสุ่มตัวอย่างซ้ำ (r) มีค่าเท่ากับ 4 เกณฑ์การเปรียบเทียบตัวประมาณค่าเฉลี่ยพิจารณาจากค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) พบว่า ตัวประมาณค่าเฉลี่ยแบบอัตราส่วนของ Kadilar (2007) มีประสิทธิภาพมากที่สุด รองลงมา คือ ตัวประมาณค่าเฉลี่ยอัตราส่วนของ Prasad (1989) ตัวประมาณค่าเฉลี่ยแบบอัตราส่วนของ Samawi and Muttlak (1996) และตัวประมาณค่าเฉลี่ยแบบอัตราส่วนแบบดั้งเดิม ตามลำดับ