

บทที่ 3

วิธีการดำเนินการศึกษา

ในการศึกษาแนวคิดและกระบวนการของแผนแบบการสุ่มตัวอย่างแบบอันดับ (Ranked Set Sampling) ได้กำหนดสัญลักษณ์ต่างๆและความหมายของสัญลักษณ์ ดังนี้

N	หมายถึง	ขนาดประชากร
n	หมายถึง	ขนาดตัวอย่างทั้งหมด
k	หมายถึง	ขนาดตัวอย่างที่ได้มาจากแผนแบบการสุ่มตัวอย่างแบบอันดับในแต่ละวัฏจักร
m	หมายถึง	จำนวนวัฏจักรของการเลือกตัวอย่างจากแผนแบบการสุ่มตัวอย่างแบบอันดับ
r	หมายถึง	อันดับของหน่วยตัวอย่างซึ่งได้มาจากจากแผนแบบการสุ่มตัวอย่างแบบอันดับ
\bar{x}	หมายถึง	ค่าเฉลี่ยของตัวอย่างที่ได้มาจากแผนแบบการสุ่มตัวอย่างแบบง่าย
\bar{x}_{rss}	หมายถึง	ค่าเฉลี่ยของตัวอย่างที่ได้มาจากแผนแบบการสุ่มตัวอย่างแบบอันดับ
\bar{X}	หมายถึง	ค่าเฉลี่ยของประชากร
$Var(\bar{x})$	หมายถึง	ความแปรปรวนของตัวประมาณ \bar{y} ที่ได้มาจากแผนแบบการสุ่มตัวอย่างแบบง่าย
$Var(\bar{x}_{rss})$	หมายถึง	ความแปรปรวนของตัวประมาณ \bar{y} ที่ได้มาจากแผนแบบการสุ่มตัวอย่างแบบอันดับ
$h(X)$	หมายถึง	ฟังก์ชันของตัวแปร X
$h(X_{[r]})$	หมายถึง	ฟังก์ชันของตัวแปร X ที่มีลำดับเท่ากับ r
\bar{X}_h	หมายถึง	ค่าเฉลี่ยของฟังก์ชันตัวแปร X
$\bar{X}_{h[r]}$	หมายถึง	ค่าเฉลี่ยของฟังก์ชันตัวแปร X ที่มีลำดับเท่ากับ r
S_h^2	หมายถึง	ความแปรปรวนของฟังก์ชันตัวแปร X

$S_{h[r]}^2$	หมายถึง	ความแปรปรวนของฟังก์ชันตัวแปร X ที่มีลำดับเท่ากับ r
$F_{r:k}(X)$	หมายถึง	ฟังก์ชันการแจกแจงสะสมของ X ที่มีลำดับเท่ากับ r จากตัวอย่างขนาดเท่ากับ k
$f_{r:k}(X)$	หมายถึง	ฟังก์ชันความหนาแน่นความน่าจะเป็นของ X ที่มีลำดับเท่ากับ r จากตัวอย่างขนาดเท่ากับ k

การอธิบายแนวคิดและกระบวนการของแผนแบบการสุ่มตัวอย่างแบบอันดับ (Ranked Set Sampling) รวมทั้งการประมาณค่าเฉลี่ยของประชากรได้ทำการศึกษาค้นคว้าจากตำราและวารสารที่เกี่ยวข้องทั้งภาษาไทยและภาษาต่างประเทศ จากการศึกษาค้นคว้าพบว่าแผนแบบการสุ่มตัวอย่างแบบอันดับเป็นแผนแบบการสุ่มตัวอย่างที่ยังไม่เป็นที่รู้จักมากนักและเป็นแผนแบบการสุ่มตัวอย่างที่พัฒนาขึ้นมาใหม่ แผนแบบการสุ่มตัวอย่างแบบอันดับนั้นโดยทั่วไปจะมี 2 ประเภทคือ แผนแบบการสุ่มตัวอย่างแบบอันดับชนิดสมดุล (Balanced Ranked Set Sampling) และแผนแบบการสุ่มตัวอย่างแบบอันดับชนิดไม่สมดุล (Unbalanced Ranked Set Sampling) กำหนดให้ n_r เป็นจำนวนหน่วยตัวอย่างลำดับที่ r เมื่อ $r = 1, 2, 3, \dots, k$ โดยที่ $\sum_{r=1}^k n_r = N$ โดยให้ $X_{[r]i}$ แทนหน่วยตัวอย่างกลุ่มที่ i ลำดับที่ r กลุ่มตัวอย่างที่จัดอันดับแล้วจะแสดงได้ดังต่อไปนี้

$X_{[1]1}$	$X_{[1]2}$	\dots	$X_{[1]n_1}$
$X_{[2]1}$	$X_{[2]2}$	\dots	$X_{[2]n_2}$
$X_{[3]1}$	$X_{[3]2}$	\dots	$X_{[3]n_3}$
\dots	\dots	\dots	\dots
$X_{[k]1}$	$X_{[k]2}$	\dots	$X_{[k]n_k}$

ถ้า $n_1 = n_2 = n_3 = \dots = n_k$ แสดงว่าแผนแบบการสุ่มตัวอย่างแบบอันดับมีความสมดุล (Balanced Ranked Set Sampling) นอกเหนือจากกรณีนี้จะถือว่าเป็นแผนแบบการสุ่มตัวอย่างแบบอันดับที่มีความไม่สมดุล (Unbalanced Ranked Set Sampling) การศึกษาครั้งนี้ทำการศึกษาเฉพาะแผนแบบการสุ่มตัวอย่างแบบอันดับชนิดสมดุลเท่านั้น เนื่องจากแผนแบบการสุ่มตัวอย่างแบบอันดับชนิดสมดุลเป็นพื้นฐานที่สำคัญในการศึกษาเรื่องแผนแบบการสุ่มตัวอย่างแบบอันดับ อีกทั้ง

ในงานวิจัยทั่วไปก็จะนำหลักการของแผนแบบการสุ่มตัวอย่างแบบอันดับชนิดสมมูลไปประยุกต์ใช้ เป็นส่วนใหญ่

ในการศึกษาแผนแบบการสุ่มตัวอย่างแบบอันดับจะอธิบายในสองส่วน คือ ในส่วนของการสุ่มตัวอย่างและในส่วนของการประมาณค่า โดยในส่วนของการสุ่มตัวอย่างจะเป็นการเลือกตัวอย่างชนิดไม่ใส่คืน(Sampling without Replacement) เนื่องจากในการเลือกตัวอย่างแบบใส่คืนเป็นการเลือกตัวอย่างที่เมื่อเลือกตัวอย่างได้หน่วยตัวอย่างใดแล้วจะนำตัวอย่างที่ได้มานั้นใส่กลับคืนประชากรก่อนที่จะทำการเลือกตัวอย่างต่อไป ดังนั้น ตัวอย่างที่ได้ อาจมีหน่วยซ้ำปรากฏอยู่ในตัวอย่าง บางครั้งอาจทำให้การประมาณค่าหรือการอธิบายลักษณะประชากรมีความคลาดเคลื่อน เพราะบางหน่วยของประชากรที่ต้องการศึกษาไม่ได้ถูกเลือกมาใช้เป็นตัวอย่าง สำหรับการสุ่มตัวอย่างชนิดไม่ใส่คืนเป็นการเลือกตัวอย่างที่กำหนดให้หน่วยตัวอย่างที่ถูกเลือกแล้วไม่สามารถถูกเลือกใหม่ได้อีก ดังนั้น ประชากรในขณะทีเลือกตัวอย่างทีละหน่วยมีลักษณะแตกต่างกันคือมีขนาดลดลงเรื่อยๆซึ่งวิธีการนี้มีประโยชน์จากการที่ตัวอย่างขนาดหนึ่งที่เลือกมาได้ให้ข้อมูลที่ครบถ้วนตามขนาดที่กำหนด คือไม่มีหน่วยซ้ำนั่นเอง และในส่วนของการประมาณค่าจะทำการพิจารณาตัวประมาณที่เป็นตัวประมาณที่ขึ้นอยู่กับค่าเฉลี่ยประชากร จากทฤษฎีสถิติ พบว่า ตัวอย่างที่นำมาพิจารณาเป็นตัวประมาณสุ่มจากการแจกแจงที่หลากหลาย ตัวประมาณค่าเฉลี่ยประชากรเป็นตัวประมาณที่มีประสิทธิภาพ กล่าวคือ เป็นตัวประมาณที่ไม่มีความเอนเอียง มีความคงเส้นคงวาหรือแบบนัย และมีค่าความแปรปรวนใกล้เคียงกับค่าที่กำหนดในทฤษฎีสถิติ สำหรับตัวประมาณของยอดรวมประชากร(Population Total) สัดส่วนประชากร (Population Proportion) และอัตราส่วนประชากร (Population Ratio) นั้น คุณสมบัติของตัวประมาณสามารถหาได้จากการปรับเปลี่ยนจากกรณีของตัวประมาณค่าเฉลี่ยประชากรทั้งสิ้น

ในการวัดคุณภาพของตัวประมาณ พิจารณาจากคุณสมบัติความไม่เอนเอียงของตัวประมาณ ซึ่งความไม่เอนเอียงของตัวประมาณก็คือ การที่ตัวประมาณนั้นมีค่าคาดหวังหรือค่าเฉลี่ยของการแจกแจงตัวอย่างของตัวประมาณมีค่าเท่ากับค่าจริงพอดี การทราบว่าตัวประมาณใดเป็นตัวประมาณที่ไม่มีความไม่เอนเอียง กล่าวคือ เมื่อเลือกตัวอย่างมาขนาดหนึ่งจากประชากรขนาดหนึ่ง ด้วยวิธีการเลือกตัวอย่างและใช้ตัวประมาณตัวนี้ ตัวอย่างที่อาจเกิดขึ้นได้แต่ละตัวอย่างจะให้ค่าประมาณค่าหนึ่ง เมื่อพิจารณาค่าประมาณทั้งหลายที่อาจเกิดขึ้นได้พร้อมกับความน่าจะเป็นที่จะเกิดขึ้นแล้ว ค่าเฉลี่ยของตัวประมาณเหล่านี้จะเท่ากับค่าเฉลี่ยประชากรที่ต้องการประมาณพอดีหรือค่ากลางของการแจกแจงของตัวประมาณตกอยู่ ณ จุดเดียวกับประชากร นอกจากนี้ ความไม่เอนเอียงส่งผลให้สามารถใช้ความแปรปรวนในการวัดความถูกต้องของตัวประมาณด้วย ดังนั้นผู้ศึกษาจึงได้พิจารณาความแปรปรวนของตัวประมาณ โดยความแปรปรวนของตัวประมาณเป็นค่าที่ใช้วัดความเที่ยงตรงของตัวประมาณ ซึ่งในกลุ่มของตัวประมาณที่ไม่มีความเอนเอียง การเลือกตัวประมาณควรจะเลือก

ตัวประมาณที่มีความแปรปรวนต่ำที่สุด ค่าความแปรปรวนจะมีค่ามากหรือน้อยนั้นขึ้นอยู่กับแผนแบบวิธีการเลือกตัวอย่างที่นำมาใช้ว่าเหมาะสมกับลักษณะประชากรที่สนใจศึกษาหรือไม่และความแปรปรวนขึ้นอยู่กับขนาดตัวอย่างด้วยกล่าวคือ ถ้าขนาดตัวอย่างใหญ่ขึ้น ค่าความแปรปรวนของตัวประมาณจะลดลง ดังนั้นในการศึกษาคุณภาพของตัวประมาณจะทำการพิจารณา 2 กรณีคือ

- ตัวประมาณค่าเฉลี่ยประชากรที่ได้มาจากแผนแบบการสุ่มตัวอย่างแบบอันดับเป็นตัวประมาณค่าที่ไม่เอนเอียงของค่าเฉลี่ยประชากร
- ความแปรปรวนของตัวประมาณค่าเฉลี่ยประชากรที่ได้มาจากแผนแบบการสุ่มตัวอย่างแบบอันดับมีค่าน้อยกว่าความแปรปรวนของตัวประมาณค่าเฉลี่ยประชากรที่ได้มาจากแผนแบบการสุ่มตัวอย่างแบบง่าย

ในการศึกษาครั้งนี้ทำการเปรียบเทียบประสิทธิภาพของแผนแบบการสุ่มตัวอย่างแบบอันดับกับแผนแบบการสุ่มตัวอย่างแบบง่าย โดยจะเปรียบเทียบจากการพิจารณาค่า ประสิทธิภาพสัมพัทธ์ (Relative Efficiency) ซึ่ง ค่าประสิทธิภาพสัมพัทธ์นี้มีค่าเท่ากับ ความแปรปรวนของตัวประมาณค่าเฉลี่ยประชากรที่ได้มาจากแผนแบบการสุ่มตัวอย่างแบบง่ายหารด้วยความแปรปรวนของตัวประมาณค่าเฉลี่ยประชากรที่ได้มาจากแผนแบบการสุ่มตัวอย่างแบบอันดับ กล่าวคือ

$$\text{ค่าประสิทธิภาพสัมพัทธ์} = \frac{\text{ความแปรปรวนของตัวประมาณค่าเฉลี่ยประชากรที่ได้มาจากแผนแบบการสุ่มตัวอย่างแบบง่าย}}{\text{ความแปรปรวนของตัวประมาณค่าเฉลี่ยประชากรที่ได้มาจากแผนแบบการสุ่มตัวอย่างแบบอันดับ}}$$

ดังนั้นถ้าค่าประสิทธิภาพสัมพัทธ์มีค่ามากกว่า 1 แสดงว่าแผนแบบการสุ่มตัวอย่างแบบอันดับมีประสิทธิภาพดีกว่าแผนแบบการสุ่มตัวอย่างแบบง่าย และถ้าค่าประสิทธิภาพสัมพัทธ์มีค่าน้อยกว่า 1 แสดงว่าแผนแบบการสุ่มตัวอย่างแบบง่ายมีประสิทธิภาพดีกว่าแผนแบบการสุ่มตัวอย่างแบบอันดับ เพราะการมีค่าสัมประสิทธิ์สัมพัทธ์ที่น้อยกว่า 1 หมายถึง ความแปรปรวนในการใช้แผนแบบการสุ่มตัวอย่างแบบง่ายมีน้อยกว่า เมื่อเทียบกับความแปรปรวนที่เกิดจากแผนแบบการสุ่มตัวอย่างแบบอันดับ