

# การสุ่มตัวอย่างแบบแบ่งกลุ่ม

## ความหมาย

การสุ่มตัวอย่างแบบแบ่งกลุ่ม (Cluster Sampling) หมายถึง การสุ่มสมาชิกจากประชากรที่มีหน่วยการสุ่มประกอบด้วยกลุ่มของสมาชิกที่มีลักษณะคล้ายกัน โดยวิธีการสุ่มตัวอย่างแบบง่าย

## หลักการและเหตุผล

ก. โดยทั่วไปปัญหาวิจัยชื่อเกี่ยวกับประชากร ครัวเรือนและที่นาถือครองของชาวนา ฯลฯ มักจะมีการเปลี่ยนแปลงอยู่เสมอ ไม่เป็นปัจจุบัน จึงทำให้หน่วยการสุ่มที่เป็นสมาชิกแต่ละสมาชิกไม่เหมาะสม จึงควรใช้หน่วยการสุ่มที่ใหญ่กว่า ซึ่งเรียกว่ากลุ่มหรือ cluster จะเหมาะสมกว่า

ข. แม้ว่าจะมีปัญหาวิจัยชื่อที่จะทำการสุ่มครบถ้วนสมบูรณ์ และเป็นปัจจุบันแล้วก็ตาม ซึ่งอาจจะทำการสุ่มตัวอย่างแบบง่ายได้ แต่ก็ไม่ควรทำทั้งนี้เพราะสมาชิกที่ถูกคัดเลือกเป็นตัวอย่าง อาจจะอยู่อย่างกระจัดกระจายทั่วทั้งประชากรทำให้ต้องเสียเวลา และค่าใช้จ่ายในการเก็บข้อมูลมาก แต่การสุ่มตัวอย่างแบบแบ่งกลุ่ม สามารถช่วยลดค่าใช้จ่ายเกี่ยวกับการเดินทางเก็บข้อมูล เพราะสมาชิกภายในกลุ่มมีลักษณะทางสภาพภูมิศาสตร์คล้ายกันมาก ทำให้ไม่จำเป็นต้องทำการวัดทั้งหมด และการเดินทางเก็บข้อมูลภายในกลุ่มจะสั้นเปลืองน้อยกว่า การเดินทางเก็บข้อมูลระหว่างสมาชิกภายในประชากร เมื่อทำการสุ่มตัวอย่างแบบง่าย

## วิธีการสุ่มตัวอย่างแบบแบ่งกลุ่ม

ก่อนอื่นจำเป็นจะต้องระบุกลุ่ม ที่มีความเหมาะสม และชัดเจนเสียก่อน คือสมาชิกภายในกลุ่มจะต้องมีลักษณะทางด้านกายภาพคล้ายคลึงกัน หรือกล่าวอีกนัยหนึ่ง คือค่าสหสัมพันธ์ระหว่างสมาชิกภายในกลุ่มค่อนข้างสูง ดังนั้นการที่จะทำการวัดสมาชิกภายในกลุ่มจำนวนมากๆ ก็มิได้ทำให้ค่าที่ต้องการจะกะประมาณเปลี่ยนแปลงไปมากแต่อย่างใด เพราะสมาชิกมีลักษณะคล้ายกัน

มากและถ้าหากกำหนดให้กลุ่ม (cluster) ใหญ่เกินไป ก็ทำให้เสียค่าใช้จ่ายในการเก็บข้อมูลมาก วิธีการสุ่มตัวอย่างแบบแบ่งกลุ่มนี้พยายามให้สมาชิกภายในกลุ่มมีจำนวนน้อยๆ เมื่อเปรียบเทียบกับขนาดของประชากร และให้มีจำนวนกลุ่มหลายๆ

ตัวอย่าง สมมติว่าเอาโรงเรียนในหมู่บ้านเป็นกลุ่ม (cluster) ของการสุ่มครัวเรือนในท้องที่แห่งหนึ่ง ภายในแต่ละกลุ่มจะมีหลายครัวเรือน ในท้องที่นี้มีโรงเรียนเพียง ๒-๓ แห่ง ก็จะแบ่งได้ ๒-๓ กลุ่มเท่านั้น กรณีเช่นนี้ ตัวอย่างที่ได้อาจจะไม่เป็นตัวแทนของประชากร เพราะว่าครัวเรือนภายในกลุ่มที่ยัดเอาโรงเรียนเป็นหลัก อาจจะมีความเป็นเอกพันธ์กันมาก ในคุณลักษณะหรือตัวแปรที่ต้องการจะวัด แต่ถ้าแบ่งให้มีจำนวนกลุ่มหลายๆ และภายในมีสมาชิกจำนวนน้อยๆ ก็จะช่วยทำให้ตัวอย่างที่ได้ เป็นตัวแทนของประชากรดีขึ้น การประมาณค่าเฉลี่ยและผลรวมของประชากร

เนื่องจากการสุ่มตัวอย่างแบบแบ่งกลุ่ม เปรียบเสมือนการสุ่มตัวอย่างแบบง่าย ที่มีหน่วยการสุ่มประกอบด้วยสมาชิกจำนวนหนึ่ง ดังนั้นการกะประมาณค่าเฉลี่ยของประชากร ( $\mu$ ) และผลรวมของประชากร ( $T$ ) จึงอาศัยหลักการเดียวกันกับการสุ่มตัวอย่างแบบง่าย โดยมีวิธีการดังนี้

ให้  $\bar{Y}$  เป็นการประมาณค่าเฉลี่ยของประชากร ( $\mu$ )  
 $N$  เป็นจำนวนกลุ่มในประชากร  
 $n$  เป็นจำนวนกลุ่มในตัวอย่าง  
 $m_i$  เป็นจำนวนสมาชิกในกลุ่มที่  $i$  ( $i=1, 2, \dots, n$ )

$$\bar{m} = \frac{1}{n} \sum_{i=1}^n m_i = \text{ขนาดเฉลี่ยของกลุ่มในตัวอย่าง}$$

$$M = \sum_{i=1}^N m_i = \text{จำนวนสมาชิกทั้งหมดในประชากร}$$

$\bar{M} = \frac{M}{N}$  = ขนาดเฉลี่ยของกลุ่มในประชากร

$Y_i$  = ค่าที่วัดได้ทั้งหมดในกลุ่มที่  $i$

ดังนั้นค่าเฉลี่ยของตัวอย่าง  $\bar{Y}$  คำนวณได้จาก

$$\bar{Y} = \frac{\sum_{i=1}^n Y_i}{n \sum_{i=1}^n m_i}$$

ซึ่งเป็นการประมาณค่าเฉลี่ยของประชากร (๔)

และความแปรปรวนของ  $\bar{Y}$  คือ

$$\hat{V}(\bar{Y}) = \left( \frac{N-n}{Nn\bar{M}^2} \right) \frac{\sum_{i=1}^n (Y_i - \bar{Y}m_i)^2}{n-1}$$

ค่าประมาณผลรวม (population total)  $T$   
หรือ  $M$  หาได้จาก  $M\bar{Y}$  ซึ่ง

$$M\bar{Y} = M \frac{\sum_{i=1}^n Y_i}{n \sum_{i=1}^n m_i}$$

ความแปรปรวนของ  $M\bar{Y}$  คือ

$$\hat{V}(M\bar{Y}) = M^2 \hat{V}(\bar{Y})$$

$$= N^2 \left( \frac{N-n}{Nn} \right) \frac{\sum_{i=1}^n (Y_i - \bar{Y}m_i)^2}{n-1}$$

อนึ่ง ถ้าหากไม่ทราบค่า  $\bar{M}$  อาจใช้  $\bar{m}$  แทนค่า  $\bar{M}$  ได้

โปรดสังเกตว่า  $\hat{V}(\bar{Y})$  ซึ่งเป็นความแปรปรวนของ  $\bar{Y}$  จะไม่ลำเอียง เมื่อ  $n$  มีจำนวนมาก กล่าวคือ  $n \geq 20$  นอกจากนั้นถ้าขนาดของกลุ่ม  $m_1, m_2, \dots, m_N$  มีขนาดเท่ากัน ก็จะทำให้การลำเอียงมีค่าน้อยลงอีกด้วย

อนันต์ ศรีโสภา

## บรรณานุกรม

อนันต์ ศรีโสภา เทคนิคการสุ่มตัวอย่าง คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร ๒๕๒๔.

Cochran, W.G. Sampling Techniques. 2nd ed. New York, John Wiley and Sons, Inc., 1963.

Kish, L. Survey Sampling. New York, John Wiley and Sons, Inc., 1965.