ตัวประมาณไม่เอนเอียงโดยการใช้สารสนเทศของตัวแปรช่วยสำหรับค่าเฉลี่ยประชากรจำกัด ภายใต้การเลือกตัวอย่างแบบสองเฟส

Unbiased Estimators Using Auxiliary Information for Finite Population Under Two-Phase Sampling

สังเคราะห์และถอดบทเรียนจากงานวิจัยเรื่อง ตัวประมาณไม่เอนเอียงโดยการใช้สารสนเทศของตัวแปรช่วยสำหรับค่าเฉลี่ยประชากรจำกัด ภายใต้การเลือกตัวอย่างแบบสองเฟส

ชื่อทุนส่งเสริมการตีพิมพ์สำหรับนิสิตระดับบัณฑิตศึกษา (ปริญญาโท)

ปีงบประมาณ.....2567.....

ชื่อ *รองศาสตราจารย์ ดร.นิภาพร ชุติมันต์ (ผู้เรียบเรียง)*

นายอธิปกรณ์ นาถมทอง (ผู้เรียบเรียง)

ภาควิชา คณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยมหาสารคาม

ในการศึกษาคุณลักษณะของประชากรจะถูกเรียกว่า พารามิเตอร์ (Parameter) ซึ่งมีข้อจำกัดในการได้มาซึ่งข้อมูลไม่ว่าจะเป็นกำลังคน เวลา หรืองบประมาณ ส่งผลให้ไม่สามารถศึกษาข้อมูลทั้งหมดประชากรได้ จึงต้องอาศัยการสุ่มตัวอย่างหรือวิธีการเลือกตัวอย่าง (Sampling Technique) ขึ้นมาโดยกำหนดวิธีการเลือกตัวอย่างที่เหมาะสมกับลักษณะของประชากรที่ทำการศึกษาและเก็บข้อมูลมาเพียงบางหน่วยของประชากรซึ่งจะเรียกว่าเป็นการสำรวจ (Survey) การเลือกตัวอย่างเชิงสำรวจในการประมาณค่าประชากร จะทำการเลือกตัวอย่างจากประชากร มาวิเคราะห์และอธิบายถึงคุณลักษณะต่าง ๆ ของประชากร ด้วยวิธีการเลือกตัวอย่างแล้วนำข้อมูลไปประมาณค่าด้วยตัวประมาณ (Estimator) ตัวประมาณอีกรูปแบบหนึ่งซึ่งนำเอาสารสนเทศของตัวแปรช่วย X มาช่วยเพิ่มประสิทธิภาพในการประมาณค่าเฉลี่ยของประชากรเพื่อให้ได้ค่าประมาณที่มีความแม่นยำมากขึ้นให้ได้ค่าที่ใกล้เคียงกับค่าพารามิเตอร์มากที่สุด และตัวแปรช่วย (Auxiliary variable) X ต้องเป็นตัวแปรที่ทราบค่าผลรวมทั้งหมดหรือค่าเฉลี่ยของประชากร โดยตัวแปรช่วย X นั้นอาจมีได้มากกว่า 1 ตัว ถ้าตัวแปร X มีความสัมพันธ์ในทิศทางเดียวกันกับตัวแปรที่ต้องการศึกษา Y ในระดับสูงจะเรียกตัวประมาณ ดังกล่าวว่า ตัวประมาณอัตราส่วน (Ratio Estimator) ซึ่งมีรูปแบบดังนี้

$$\bar{y}_R = \bar{y}\left(\frac{\bar{X}}{\bar{x}}\right) \quad \text{เมื่อ ทราบค่า } \bar{X}$$

โดย $\bar{X}$ คือ ค่าเฉลี่ยประชากรของตัวแปรช่วย

ในทางกลับกันถ้าตัวแปรช่วย X มีความสัมพันธ์ในทิศทางตรงกันข้ามกับตัวแปรที่ต้องการศึกษา Y ในระดับสูงจะเรียกตัวประมาณนี้ว่า ตัวประมาณผลคูณ (Product Estimator) มีรูปแบบดังนี้

$$\bar{y}_P = \bar{y}\left(\frac{\bar{x}}{\bar{X}}\right) \quad \text{เมื่อ ทราบค่า } \bar{X}$$

โดย $\bar{X}$ คือ ค่าเฉลี่ยประชากรของตัวแปรช่วย

ในการวิจัยครั้งนี้ผู้วิจัยสนใจศึกษาทฤษฎีเกี่ยวกับตัวประมาณอัตราส่วนไม่เอนเอียง และตัวประมาณผลคูณไม่เอนเอียง ที่เสนอโดย Tin (1965) Robson (1957) และ Mahanty & Mishra (2020) ภายใต้การเลือกตัวอย่างสองเฟส และเปรียบเทียบประสิทธิภาพตัวประมาณที่ศึกษากับตัวประมาณอัตราส่วนและตัวประมาณผลคูณแบบทั่วไป โดยผลการวิจัยพบว่า ตัวประมาณไม่เอนเอียง ภายใต้การเลือกตัวอย่างสองเฟส (Proposed Unbiased Estimator under Two-Phase Sampling) คือ

ตัวประมาณอัตราส่วนที่เสนอโดย Tin (1965)

$$\bar{y}_{Tin}^{(2)} = \bar{y}^{(2)} \frac{\bar{x}^{(1)}}{\bar{x}^{(2)}} \left[1 + f_2 \left(\frac{s_{yx}^{(2)}}{\bar{x}^{(2)}\bar{y}^{(2)}} - \frac{\left(s_x^{(2)}\right)^2}{\left(\bar{x}^{(2)}\right)^2} \right) \right]$$

โดยมีค่าความแปรปรวนดังนี้

$$V\left(\bar{y}_{Tin}^{(2)}\right) \approx \bar{Y}^2 \left(f_1 C_Y^2 + f_2 \left(C_Y^2 + C_X^2 - 2\rho_{YX} C_Y C_X \right) \right)$$

ตัวประมาณผลคูณที่เสนอโดย Robson (1957)

$$\bar{y}_{Rob}^{(2)} = \frac{\bar{y}^{(2)}\bar{x}^{(2)}}{\bar{x}^{(1)}} - f_2 \frac{s_{yx}^{(2)}}{\bar{x}^{(1)}}$$

โดยมีความแปรปรวนดังนี้

$$V\left(\bar{y}_{Rob}^{(2)}\right) \approx \bar{Y}^2 \left(f_1 C_Y^2 + f_2 \left(C_Y^2 + C_X^2 + 2\rho_{YX} C_Y C_X \right) \right)$$

ตัวประมาณที่เสนอโดย Mahanty & Mishra (2020)

$$\bar{y}_{MU}^{(2)} = \lambda_0^{(2)} \left(\bar{y}^{(2)} \frac{\bar{x}^{(2)}}{\bar{x}^{(1)}} - f_2 \frac{s_{yx}^{(2)}}{\bar{x}^{(1)}} \right) + \lambda_1^{(2)} \bar{x}^{(2)}$$

โดยมีความแปรปรวนดังนี้

$$V\left(\bar{y}_{MU}^{(2)}\right) \approx \bar{Y}^2 \left(f_1 C_Y^2 + f_2 C_X^2 \left(1 - \rho_{YX}^2 \right) \right)$$

และเปรียบเทียบประสิทธิภาพของตัวประมาณในเชิงทฤษฎีผ่านการการอธิบายเชิงตัวเลข แสดงดังตาราง

Efficiency Comparisons

$n^{(1)}$	$n^{(2)}$	$\hat{V}(\bar{y}_1) - \hat{V}(\bar{y}_2)$	$\rho_{YX} = 0.8$	$\rho_{YX} = 0.6$	$\rho_{YX} = 0.3$	$\rho_{YX} = -0.8$	$\rho_{YX} = -0.6$	$\rho_{YX} = -0.3$
50	10	$\hat{V}\left(\bar{y}_{Tin}^{(2)}\right) - \hat{V}\left(\bar{y}_{MU}^{(2)}\right)$	24.6549	25.5443	27.2677	*	*	*
		$\hat{V}\left(\bar{y}_{Rob}^{(2)}\right) - \hat{V}\left(\bar{y}_{MU}^{(2)}\right)$	*	*	*	24.6112	25.5535	27.4772
100	20	$\hat{V}\left(\bar{y}_{Tin}^{(2)}\right) - \hat{V}\left(\bar{y}_{MU}^{(2)}\right)$	12.2381	12.8821	13.8350	*	*	*
		$\hat{V}\left(\bar{y}_{Rob}^{(2)}\right) - \hat{V}\left(\bar{y}_{MU}^{(2)}\right)$	*	*	*	12.3443	12.8778	13.7319
100	50	$\hat{V}\left(\bar{y}_{Tin}^{(2)}\right) - \hat{V}\left(\bar{y}_{MU}^{(2)}\right)$	3.0918	3.2289	3.4260	*	*	*
		$\hat{V}\left(\bar{y}_{Rob}^{(2)}\right) - \hat{V}\left(\bar{y}_{MU}^{(2)}\right)$	*	*	*	3.0760	3.2115	3.42799
200	50	$\hat{V}\left(\bar{y}_{Tin}^{(2)}\right) - \hat{V}\left(\bar{y}_{MU}^{(2)}\right)$	4.6267	4.8060	5.1138	*	*	*
		$\hat{V}\left(\bar{y}_{Rob}^{(2)}\right) - \hat{V}\left(\bar{y}_{MU}^{(2)}\right)$	*	*	*	4.5899	4.8192	5.1050

Efficiency Comparisons								
$n^{(1)}$	$n^{(2)}$	$\hat{V}(\bar{y}_1)-\hat{V}(\bar{y}_2)$	$\rho_{YX}=0.8$	$\rho_{YX}=0.6$	$\rho_{YX}=0.3$	$\rho_{YX}=-0.8$	$\rho_{YX}=-0.6$	$\rho_{YX}=-0.3$
200	100	$\hat{V}(\bar{y}^{(2)}_{Tin})-\hat{V}(\bar{y}^{(2)}_{MU})$	1.5398	1.6087	1.7122	*	*	*
		$\hat{V}(\bar{y}^{(2)}_{Rob})-\hat{V}(\bar{y}^{(2)}_{MU})$	*	*	*	1.5363	1.6027	1.7073
500	100	$\hat{V}(\bar{y}^{(2)}_{Tin})-\hat{V}(\bar{y}^{(2)}_{MU})$	2.4605	2.5757	2.7478	*	*	*
		$\hat{V}(\bar{y}^{(2)}_{Rob})-\hat{V}(\bar{y}^{(2)}_{MU})$	*	*	*	2.4624	2.5666	2.7421
500	200	$\hat{V}(\bar{y}^{(2)}_{Tin})-\hat{V}(\bar{y}^{(2)}_{MU})$	0.9247	0.9660	1.0308	*	*	*
		$\hat{V}(\bar{y}^{(2)}_{Rob})-\hat{V}(\bar{y}^{(2)}_{MU})$	*	*	*	0.9215	0.9663	1.0249
500	300	$\hat{V}(\bar{y}^{(2)}_{Tin})-\hat{V}(\bar{y}^{(2)}_{MU})$	0.4105	0.4282	0.4575	*	*	*
		$\hat{V}(\bar{y}^{(2)}_{Rob})-\hat{V}(\bar{y}^{(2)}_{MU})$	*	*	*	0.4101	0.4277	0.4557

** เป็นตัวประมาณที่ประชากรมีความสัมพันธ์แตกต่างกัน จึงไม่ทำการเปรียบเทียบ*

เมื่อเปรียบเทียบประสิทธิภาพของตัวประมาณในเชิงทฤษฎีผ่านการการอธิบายเชิงตัวเลขของตัวประมาณภายใต้การเลือกตัวอย่างสองเฟส พบว่า ตัวประมาณที่ผู้วิจัยนำเสนอ $\bar{y}^{(2)}_{MU}$ มีประสิทธิภาพสูงกว่าตัวประมาณอื่นที่ศึกษา ทั้งในประชากรมีความสัมพันธ์กันในทิศทางเดียวกันและตรงกันข้าม ในทุกขนาดตัวอย่าง

เอกสารอ้างอิง

Irfan, M., Javed, M., Bhatti, S. H., Raza, M. A., & Ahmad, T. (2020). Almost unbiased optimum estimators for population mean using dual auxiliary information. Journal of King Saud University – Science, 32, 2835-2844.

Mahanty, B., & Mishra, G. (2020). An Unbiased Estimator of Finite Population Mean Using Auxiliary Information. Journal of Statistical Theory and Applications, 19(4), 534-539.

Robson, D. S. (1957). Applications of Multivariate Polykays to the Theory of Unbiased Ratio-Type Estimation. Journal of the American Statistical Association, 52, 511-522.

Tin, M. (1965). Comparison of Some Ratio Estimators. Journal of the American Statistical Association, 60, 294-307.

Vishwakarma, G. K., & Zeeshan, S. M. (2021). Generalized Ratio-cum-Product Estimator for Finite Population Mean under Two-Phase Sampling Scheme. Journal of Modern Applied Statistical Methods, 19(1), 1-16.

ชื่อผู้เรียบเรียง รองศาสตราจารย์ ดร.นิภาพร ชุติมันต์ และ นายอธิปกรณ์ นาถมทอง

ที่อยู่ ภาควิชาคณิตศาสตร์ มหาวิทยาลัยมหาสารคาม

เบอร์โทรศัพท์ 043-754244 E-mail : nipaporn.c@msu.ac.th