บทที่ 5

สรุปผลการวิจัย อภิปรายผลและข้อเสนอแนะ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อหาผลสรุปการเปรียบเทียบความแม่นยำของวิธีประมาณค่าข้อมูลที่สูญหาย 3 วิธีคือ (1) วิธีใช้ค่าเฉลี่ยจากกลุ่มตัวอย่าง (2) วิธีใช้สมการถดถอย และ (3) วิธีใช้ค่าเฉลี่ยระหว่างค่าเฉลี่ยจากกลุ่มตัวอย่างและสมการถดถอย เมื่อกำหนดจำนวนข้อมูลที่สูญหายครั้งละ 1 ค่า และครั้งละ 2 ค่า จากกลุ่มตัวอย่างขนาด 5, 10 และ 15 โดยวิเคราะห์ผลตามลักษณะข้อมูลที่สูญหายแบ่งเป็น 2 กรณี คือ 1. ไม่จำแนกเขตของข้อมูลที่สูญหาย หรือเขตของข้อมูลที่สูญหายครอบคลุมข้อมูลทั้งหมด (Cover all regions) และ 2. จำแนกข้อมูลที่สูญหายออกเป็น 3 เขต คือ มีค่ามากกว่า $\bar{X}$ + 1S.D. หรือมีค่าตั้งแต่ $\bar{X}$ - 1S.D. ถึง $\bar{X}$ + 1S.D. หรือมีค่าน้อยกว่า $\bar{X}$ - 1S.D. การใช้วิธีสมการถดถอย และวิธีใช้ค่าเฉลี่ยระหว่างค่าเฉลี่ยจากกลุ่มตัวอย่างและสมการถดถอยนั้น ได้กำหนดแผนการวิจัยให้มีสภาพคล้ายคลึงกับงานวิจัยที่ศึกษา โดยใช้ข้อมูลที่มีลักษณะการแจกแจงแบบปกติสองตัวแปร (Bivariate Normal Distribution) ที่มีค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรเท่ากับ 0.2, 0.4 และ 0.6 การวิจัยครั้งนี้ใช้เกณฑ์ในการเปรียบเทียบความแม่นยำ คือ (1) ค่าเฉลี่ยของผลต่างระหว่างค่าข้อมูลที่สูญหายกับค่าที่ประมาณได้จากวิธีประมาณค่า และ (2) ค่าเฉลี่ยกำลังสองของผลต่างระหว่างค่าข้อมูลที่สูญหายกับค่าที่ประมาณได้จากวิธีประมาณค่า

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

แผนการทดลอง

ในการทดลองเพื่อคำนวณหาค่าเฉลี่ยของผลต่างและค่าเฉลี่ยกำลังสองของผลต่างระหว่างค่าข้อมูลที่สูญหายกับค่าที่ประมาณได้จากวิธีประมาณค่า โดยกำหนดแผนการทดลองตามจำนวนข้อมูลที่สูญหาย ขนาดกลุ่มตัวอย่างและค่าสัมประสิทธิ์สหสัมพันธ์ เป็นกรณีที่จะทดลองทั้งสิ้น 18 กรณี จำแนกดังรายละเอียดต่อไปนี้

1. จำนวนข้อมูลที่สูญหายครั้งละ 1 ค่า จากกลุ่มตัวอย่างขนาดต่าง ๆ ทดลองทั้งสิ้น 9 กรณี โดยจำแนกได้ดังนี้

1.1 ขนาดกลุ่มตัวอย่าง 5 ที่ค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.2 , 0.4 และ 0.6 รวม 3 กรณี

1.2 ขนาดกลุ่มตัวอย่าง 10 ที่ค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.2 , 0.4 และ 0.6 รวม 3 กรณี

1.3 ขนาดกลุ่มตัวอย่าง 15 ที่ค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.2 , 0.4 และ 0.6 รวม 3 กรณี

2. จำนวนข้อมูลที่สูญหายครั้งละ 2 ค่าจากกลุ่มตัวอย่างขนาดต่าง ๆ ทดลองทั้งสิ้น 9 กรณี โดยจำแนกได้เช่นเดียวกับข้อ 1.1 , 1.2 และ 1.3

<u>วิธีดำเนินการทดลอง</u>

การทดลองครั้งนี้ใช้เทคนิคมอนติคาร์โลซิมูเลชั่น โดยในแต่ละกรณีกำหนดให้คอมพิวเตอร์จำลองการทดลอง 4,000 ครั้ง และในแต่ละครั้งกำหนดให้ประมาณค่าข้อมูลที่สูญหายด้วยวิธีประมาณค่า 3 วิธี คือ วิธีใช้ค่าเฉลี่ยจากกลุ่มตัวอย่าง วิธีใช้สมการถดถอย และวิธีใช้ค่าเฉลี่ยระหว่างค่าเฉลี่ยจากกลุ่มตัวอย่างและสมการถดถอย พร้อมทั้งหาค่าผลต่างและกำลังสองของผลต่างระหว่างข้อมูลที่สูญหายกับค่าที่ประมาณได้จากวิธีประมาณค่าแต่ละวิธีจำแนกตามเขตข้อมูลที่สูญหาย และแต่ละกรณีจะคำนวณหาค่าเฉลี่ยของผลต่างและของกำลังสองของผลต่างระหว่างข้อมูลที่สูญหายกับค่าที่ประมาณได้จากวิธีประมาณค่า จำแนกตามเขตข้อมูลที่สูญหาย สำหรับโปรแกรมที่ใช้ในการทดลองครั้งนี้ใช้ภาษาฟอร์แทรน 77 (FORTRAN 77) และใช้กับเครื่อง IBM 370/3031 ในระบบ OS/VS1

สรุปผลการวิจัย

ผลสรุปการเปรียบเทียบความแม่นยำของวิธีประมาณค่าข้อมูลที่สูญหาย 3 วิธี คือ (1) วิธีใช้ค่าเฉลี่ยจากกลุ่มตัวอย่าง (Mean) (2) วิธีใช้สมการถดถอย (Reg.) และ (3) วิธีใช้ค่าเฉลี่ยระหว่างค่าเฉลี่ยจากกลุ่มตัวอย่างและสมการถดถอย (Aver.) ในกรณีไม่จำแนกเขตของข้อมูลที่สูญหายเสนอผลการเปรียบเทียบเป็น 2 กรณีดังนี้

1. การเปรียบเทียบวิธีประมาณค่าโดยใช้ค่าเฉลี่ยของผลต่างระหว่างค่าของข้อมูลที่สูญหายกับค่าที่ประมาณได้จากวิธีประมาณค่า

เมื่อขนาดกลุ่มตัวอย่างเป็น 5 นั้นวิธี ใช้ค่าเฉลี่ยระหว่างค่าเฉลี่ยจากกลุ่มตัวอย่างและสมการถดถอย จะมีค่าเฉลี่ยของผลต่างระหว่าง -0.10 ถึง -0.02 จะเป็นค่าที่ดีที่สุด ซึ่งดีกว่าวิธีใช้สมการถดถอยที่มีค่าระหว่าง -0.06 ถึง 0.33 และวิธีใช้ค่าเฉลี่ยจากกลุ่มตัวอย่างที่มีค่าอยู่ระหว่าง -0.20 ถึง -0.12

เมื่อขนาดกลุ่มตัวอย่างเป็น 10 นั้นวิธีใช้สมการถดถอย จะมีค่าเฉลี่ยของผลต่างระหว่าง -0.10 ถึง 0.04 และใกล้เคียงกับวิธีใช้ค่าเฉลี่ยระหว่างค่าเฉลี่ยจากกลุ่มตัวอย่างและสมการถดถอยที่มีค่าระหว่าง -0.07 ถึง 0.05 ซึ่งทั้งสองวิธีจะดีกว่าวิธีใช้ค่าเฉลี่ยจากกลุ่มตัวอย่างที่มีค่าอยู่ระหว่าง 1.54 ถึง 1.65

เมื่อขนาดกลุ่มตัวอย่างเพิ่มขึ้นเป็น 15 นั้น วิธีใช้สมการถดถอยก็มีค่าเฉลี่ยของผลต่างระหว่าง -0.09 ถึง 0.01 และใกล้เคียงกับวิธีใช้ค่าเฉลี่ยระหว่างค่าเฉลี่ยจากกลุ่มตัวอย่างและสมการถดถอยที่มีค่าระหว่าง -0.07 ถึง 0.03 ซึ่งทั้งสองวิธีนี้ให้ค่าเฉลี่ยของผลต่างลดต่ำลงเมื่อขนาดกลุ่มตัวอย่างเพิ่มขึ้น โดยที่วิธีใช้ค่าเฉลี่ยจากกลุ่มตัวอย่างจะมีค่าอยู่ระหว่าง 1.92 ถึง 2.63

นั่นคือวิธีใช้ค่าเฉลี่ยระหว่างค่าเฉลี่ยจากกลุ่มตัวอย่างและสมการถดถอยให้ค่าเฉลี่ยของผลต่างน้อยที่สุดเมื่อขนาดกลุ่มตัวอย่างเป็น 5 ในทุกกรณี แต่เมื่อขนาดกลุ่มตัวอย่างเพิ่มขึ้นเป็น 10 หรือ 15 วิธีใช้สมการถดถอยจะมีค่าเฉลี่ยของผลต่างต่ำกว่าวิธีใช้ค่าเฉลี่ยระหว่างค่าเฉลี่ยจากกลุ่มตัวอย่างและสมการถดถอย ในทุกกรณีที่จำนวนการสูญหายครั้งละ 1 ค่า แต่เมื่อจำนวนการสูญหายเพิ่มเป็นครั้งละ 2 ค่า วิธีใช้สมการถดถอยจะให้ค่าเฉลี่ยของผลต่างต่ำกว่าวิธีใช้ค่าเฉลี่ยระหว่างค่าเฉลี่ยจากกลุ่มตัวอย่างและสมการถดถอยที่ระดับค่าสัมประสิทธิ์

สหสัมพันธ์สูง ๆ คือ $\rho = 0.4$ หรือ $\rho = 0.6$ และสำหรับวิธีใช้ค่าเฉลี่ยจากกลุ่มตัวอย่างจะประมาณได้ค่าเฉลี่ยของผลต่างสูงกว่าทั้งสองวิธีดังกล่าวในทุกขนาดกลุ่มตัวอย่าง

2. การเปรียบเทียบวิธีประมาณค่าโดยใช้ค่าเฉลี่ยกำลังสองของผลต่างระหว่างค่าของข้อมูลที่สูญหายกับค่าที่ประมาณได้จากวิธีประมาณค่า สามารถสรุปได้ดังตารางที่ 10 ดังนี้

ตารางที่ 10 วิธีประมาณค่าที่มีความแม่นยำมากที่สุดในแต่ละกรณี

n	เขตของข้อมูลที่สูญหาย	จำนวนข้อมูลที่สูญหายครั้งละ 1 ค่า			จำนวนข้อมูลที่สูญหายครั้งละ 2 ค่า		
		$\rho = 0.2$	$\rho = 0.4$	$\rho = 0.6$	$\rho = 0.2$	$\rho = 0.4$	$\rho = 0.6$
5	C.A.R.	AVER	AVER	AVER	AVER	AVER	AVER
	REGION1	MEAN	REG.	REG.	AVER	AVER	AVER
	REGION2	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN
	REGION3	REG.	REG.	REG.	AVER	AVER	AVER
10	C.A.R.	AVER	REG.	REG.	AVER	AVER	REG.
	REGION1	REG.	REG.	REG.	REG.	REG.	REG.
	REGION2	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN
	REGION3	REG.	REG.	REG.	REG.	REG.	REG.
15	C.A.R.	AVER	REG.	REG.	AVER	REG.	REG.
	REGION1	REG.	REG.	REG.	REG.	REG.	REG.
	REGION2	MEAN	MEAN	MEAN	MEAN	MEAN	MEAN
	REGION3	REG.	REG.	REG.	REG.	REG.	REG.

ผลจากตารางที่ 10 สามารถสรุปส่วนสำคัญได้ดังนี้

1. กรณีไม่จำแนกตำแหน่งของข้อมูลที่สูญหายอยู่ในช่วงใดของกลุ่มตัวอย่าง

1.1 วิธีใช้ค่าเฉลี่ยระหว่างค่าเฉลี่ยจากกลุ่มตัวอย่างและสมการถดถอย ให้ผลการประมาณค่าดีที่สุด เมื่อกลุ่มตัวอย่างเป็น 5 ในทุกกรณี และเมื่อขนาดกลุ่มตัวอย่างเพิ่มขึ้นเป็น 10 วิธีนี้ก็จะประมาณค่าได้ดีเฉพาะจำนวนข้อมูลสูญหายครั้งละ 2 ค่าและที่ค่า $\rho = 0.2$ หรือ $\rho = 0.4$ เท่านั้น

1.2 วิธีใช้สมการถดถอย ให้ผลการประมาณค่าดีที่สุดเมื่อขนาดกลุ่มตัวอย่างเป็น 10 หรือ 15 และจำนวนข้อมูลที่สูญหายครั้งละ 1 ค่า ในทุกระดับค่าสัมประสิทธิ์สหสัมพันธ์ยกเว้นที่ $\rho = 0.2$ แต่ที่จำนวนข้อมูลที่สูญหายครั้งละ 2 ค่า วิธีนี้จะประมาณค่าได้ดีเฉพาะที่มีค่าสัมประสิทธิ์สหสัมพันธ์สูง ๆ คือ ที่ $\rho = 0.6$

1.3 วิธีใช้ค่าเฉลี่ยจากกลุ่มตัวอย่าง ให้ผลการประมาณค่าไม่ดีเท่ากับวิธีใช้สมการถดถอยและวิธีใช้ค่าเฉลี่ยระหว่างค่าเฉลี่ยจากกลุ่มตัวอย่างและสมการถดถอย ในทุกกรณี

2. กรณีจำแนกตำแหน่งของข้อมูลที่สูญหายเป็น 3 เขต แต่ละวิธีให้ผลการประมาณค่าแม่นยำกว่าวิธีอื่นเฉพาะกรณีดังนี้

2.1 วิธีใช้ค่าเฉลี่ยระหว่างค่าเฉลี่ยจากกลุ่มตัวอย่างและสมการถดถอย เฉพาะขนาดกลุ่มตัวอย่างเป็น 5 และจำนวนข้อมูลที่สูญหายครั้งละ 2 ค่า ในเขตข้อมูลสูญหายมากกว่า $\bar{X}$ + 1S.D. หรือน้อยกว่า $\bar{X}$ - 1S.D. ในทุกระดับค่าสัมประสิทธิ์สหสัมพันธ์

2.2 วิธีใช้สมการถดถอย เมื่อขนาดกลุ่มตัวอย่างเป็น 5 และจำนวนข้อมูลที่สูญหายครั้งละ 1 ค่า ในเขตข้อมูลสูญหายมากกว่า $\bar{X}$ + 1S.D. ที่ $\rho = 0.4$ และ $\rho = 0.6$ และเขตข้อมูลสูญหายน้อยกว่า $\bar{X}$ - 1S.D. ในทุกระดับค่า ρ แต่เมื่อขนาดกลุ่มตัวอย่างเพิ่มขึ้นเป็น 10 หรือ 15 วิธีนี้ประมาณได้ดีกว่าวิธีอื่นทุกกรณีในเขตข้อมูลที่สูญหายมากกว่า $\bar{X}$ + 1S.D. หรือน้อยกว่า $\bar{X}$ - 1S.D.

3. วิธีใช้ค่าเฉลี่ยจากกลุ่มตัวอย่าง เมื่อข้อมูลที่สูญหายมีค่าตั้งแต่ $\bar{X}$ - 1S.D. ถึง $\bar{X}$ + 1S.D. ในทุกกรณี

อภิปรายผล

วิธีการประมาณค่าข้อมูลที่สูญหายทั้ง 3 วิธีนั้นมีเทคนิควิธีหรือลักษณะของการประมาณค่าที่แตกต่างกัน โดยวิธีใช้ค่าเฉลี่ยจากกลุ่มตัวอย่างเป็นวิธีการประมาณค่าแบบอนุรักษ์นิยม (Conservative) ในการหาค่าประมาณ เพราะค่านั้นไม่มากหรือน้อยเกินไปกว่าค่าของข้อมูลที่เหลือในกลุ่มตัวอย่าง จึงไปลดค่าความแปรปรวนของข้อมูลในกลุ่มให้ต่ำลง ซึ่งการคำนวณโดยวิธีนี้ง่ายไม่ยุ่งยากซับซ้อน แต่ให้ผลการประมาณค่าที่แม่นยำน้อยกว่าวิธีใช้สมการถดถอย และวิธีใช้ค่าเฉลี่ยระหว่างค่าเฉลี่ยจากกลุ่มตัวอย่างและสมการถดถอย ในทุกขนาดกลุ่มตัวอย่างที่ใช้ในการวิจัย

สำหรับวิธีใช้ค่าเฉลี่ยระหว่างค่าเฉลี่ยจากกลุ่มตัวอย่างและสมการถดถอย มีเทคนิควิธีการคำนวณถึง 2 ขั้นตอน คือ คำนวณหาค่าเฉลี่ยจากกลุ่มตัวอย่างก่อน แล้วคำนวณโดยใช้สมการถดถอยอีกแล้วจึงจะนำผลของทั้งสองวิธีมาเฉลี่ยกัน วิธีนี้จึงมีลักษณะอนุรักษ์นิยมค่าประมาณที่ได้มากกว่าวิธีใช้สมการถดถอย แต่ก็น้อยกว่าวิธีใช้ค่าเฉลี่ยจากกลุ่มตัวอย่าง เมื่อขนาดกลุ่มตัวอย่างเป็น 5 นั้น วิธีนี้พบว่าใช้ได้ดีที่จำนวนการสูญหายของข้อมูลเป็น 1 หรือ 2 ค่าในทุกระดับค่าสัมประสิทธิ์สหสัมพันธ์ และยังประมาณค่าได้ดีเมื่อขนาดกลุ่มตัวอย่างเพิ่มขึ้นที่จำนวนการสูญหายมากกว่า 1 ค่า และระดับค่าสัมประสิทธิ์สหสัมพันธ์ไม่มากนัก ซึ่งสอดคล้องกับอะฟิฟีและอิลาสฮอฟฟ์ (Afifi and Elashoff 1967: 16) ที่พบว่าวิธีใช้ค่าเฉลี่ยระหว่างค่าเฉลี่ยจากกลุ่มตัวอย่างและสมการถดถอยนั้น จะประมาณค่าได้ใกล้เคียงกับข้อมูลที่สูญหายมากกว่าวิธีใช้สมการถดถอย เมื่อกลุ่มตัวอย่างมีขนาดเล็ก และจำนวนการสูญหายของข้อมูลมีเพียงเล็กน้อย แม้ว่าการคำนวณจากวิธีนี้จะยุ่งยากสลับซับซ้อนเพียงไร แต่ก็คงจะให้ผลที่คุ้มค่าเช่นกัน

ในการประมาณค่าโดยใช้สมการถดถอยจากกลุ่มตัวอย่างขนาด 5 และมีข้อมูลสูญหายไปอีก 1 หรือ 2 ค่านั้น ถ้าพิจารณาค่า β ของสมการถดถอยที่คำนวณจากข้อมูลพื้นฐานเพียง 3 หรือ 4 ตัว จึงทำให้ผลในการทำนายไม่ดีกว่าวิธีใช้ค่าเฉลี่ยระหว่างค่าเฉลี่ยจากกลุ่มตัวอย่างและสมการถดถอย แต่อย่างไรก็ตามเมื่อขนาดของกลุ่มตัวอย่างมีค่ามาก ๆ แล้ว การทำนายก็จะถูกต้องแม่นยำมากยิ่งขึ้น เพราะเรามีข้อมูลพื้นฐานที่ใช้คำนวณมาก ซึ่งจะทำให้ค่า β ที่ได้มีประสิทธิภาพสูงขึ้นด้วยเช่นกัน จากผลการวิจัยที่ขนาดกลุ่มตัวอย่างเป็น 10 หรือ 15 นั้น วิธีประมาณค่าโดยใช้สมการถดถอยจะประมาณได้ดีกว่าวิธีอื่น ที่ระดับค่าสัมประสิทธิ์สหสัมพันธ์สูงขึ้นเป็น 0.4 หรือ 0.6 ผลสรุปนี้จะสอดคล้องกับการศึกษาของเคลีเจียน (Kelejean 1969: 1610)

ที่พบว่าวิธีถดถอยอันดับหนึ่ง (First-order regression) หรือวิธีใช้สมการถดถอยจะมีประสิทธิภาพเพิ่มสูงขึ้นเมื่อตัวแปรต้นและตัวแปรตามมีความสัมพันธ์กันสูง นั่นคือถ้ากำหนดให้ขนาดของกลุ่มตัวอย่างคงที่และใช้ค่าสัมประสิทธิ์สหสัมพันธ์เพิ่มสูงขึ้น วิธีใช้สมการถดถอยก็มีแนวโน้มจะประมาณค่าได้ดีขึ้นเช่นกัน

จากผลการวิจัยครั้งนี้ วิธีใช้ค่าเฉลี่ยจากกลุ่มตัวอย่างให้ผลการประมาณค่าต่ำกว่าวิธีอื่นๆ ในทุกขนาดกลุ่มตัวอย่าง เพราะวิธีนี้ใช้เพียงข้อมูลที่เหลืออยู่มาเฉลี่ยแล้วแทนค่าข้อมูลที่สูญหายเลย โดยไม่นำตัวแปรร่วมที่มีความสัมพันธ์กับข้อมูลที่สูญหายมาช่วยในการคำนวณ ซึ่งต่างจากวิธีใช้ค่าเฉลี่ยระหว่างค่าเฉลี่ยจากกลุ่มตัวอย่างและสมการถดถอย ที่นำตัวแปรควบคุมมาคิดคำนวณด้วย และก็ประมาณค่าได้ดีมากที่กลุ่มตัวอย่างขนาด 5 แต่เมื่อขนาดกลุ่มตัวอย่างเพิ่มขึ้นเป็น 10 หรือ 15 ก็ยังคงให้ผลการประมาณที่ดีเมื่อจำนวนการสูญหายมีมากกว่า 1 ค่าที่ระดับค่าสัมประสิทธิ์สหสัมพันธ์ต่ำ ๆ ซึ่งในกลุ่มตัวอย่างเป็น 10, 15 หรือที่จำนวนมากกว่านี้ เช่น 20 วิธีใช้สมการถดถอยจะประมาณค่าได้ดีที่ตัวแปรควบคุมและตัวแปรตามมีความสัมพันธ์กันสูง

การทำงานวิจัยโดยทั่ว ๆ ไปนั้น การป้องกันไม่ให้เกิดกรณีการสูญหายของข้อมูลเป็นวิธีที่ดีที่สุด เพราะไม่มีวิธีประมาณค่าวิธีใดที่สามารถประมาณได้ดีหรือเท่ากับค่าของข้อมูลที่สูญหาย แต่ถ้าผู้วิจัยไม่มีความมั่นใจว่าในการวิจัยนั้นจะมีการสูญหายของข้อมูลเกิดขึ้นหรือไม่ ผู้วิจัยก็ควรใช้กลุ่มตัวอย่างให้ใหญ่มาก ๆ เพื่อให้ข้อมูลที่สูญหายแต่ละตัวมีผลต่อการวิเคราะห์น้อยที่สุด โดยสามารถที่จะตัดข้อมูลที่สูญหายออกจากการวิเคราะห์ได้ถ้ากลุ่มตัวอย่างมีขนาดมากเพียงพอ หรือสามารถใช้วิธีประมาณค่าแทนข้อมูลที่สูญหายเพื่อให้ได้ค่าใกล้เคียงกับค่าที่สูญหาย และถ้าปฏิบัติตามกรณีข้างต้นไม่ได้คือผู้วิจัยจำเป็นต้องใช้กลุ่มตัวอย่างขนาดเล็กและคาดว่าจะมีข้อมูลสูญหายเกิดขึ้นในงานวิจัยแล้ว ก็จำเป็นอย่างยิ่งที่จะต้องเก็บรวบรวมข้อมูลในส่วนอื่นของกลุ่มตัวอย่างเพิ่มเติมโดยเฉพาะที่เป็นตัวแปรควบคุมซึ่งมีความสัมพันธ์กับข้อมูลที่สูญหายอยู่ในระดับสูง ๆ

ข้อเสนอแนะ

จากผลการวิจัยครั้งนี้ผู้วิจัยเสนอแนะเป็น 2 ส่วนดังนี้

ก. ข้อเสนอแนะเพื่อการนำไปใช้

1. ผู้วิจัยควรระมัดระวังในการทำงานวิจัย เพื่อป้องกันไม่ให้เกิดกรณีการสูญหายของข้อมูลจะเป็นวิธีที่ดีที่สุด เพราะไม่มีวิธีประมาณค่าวิธีใดที่สามารถประมาณค่าได้ดีเท่ากับค่าของข้อมูลที่สูญหาย

2. ถ้าหลีกเลี่ยงกรณีตามข้อ 1 ไม่ได้ ผู้วิจัยควรใช้ขนาดกลุ่มตัวอย่างขนาดใหญ่เพื่อให้ข้อมูลที่สูญหายแต่ละตัวมีผลต่อการวิเคราะห์น้อยที่สุด หรือสามารถที่จะตัดข้อมูลที่สูญหายออกจากการวิเคราะห์ได้ เมื่อกลุ่มตัวอย่างมีขนาดใหญ่เพียงพอ

3. ในกรณีที่ปฏิบัติตามข้อ 1 หรือ 2 ไม่ได้คือผู้วิจัยจำเป็นต้องใช้กลุ่มตัวอย่างขนาดเล็กและคาดว่าจะมีข้อมูลสูญหายเกิดขึ้นในงานวิจัย ผู้วิจัยควรเก็บรวบรวมข้อมูลที่เป็นตัวแปรควบคุมไว้ด้วย และตัวแปรที่เป็นตัวแปรควบคุมนั้นควรมีความสัมพันธ์กับข้อมูลที่สูญหายอยู่ในระดับสูง เพื่อสามารถใช้วิธีประมาณค่าแทนข้อมูลที่สูญหายได้ใกล้เคียงกับค่าที่สูญหายมากขึ้น คือ เมื่อขนาดกลุ่มตัวอย่างเป็น 5 และจำนวนข้อมูลที่สูญหาย 1 หรือ 2 ค่า ในทุกระดับค่าสัมประสิทธิ์สหสัมพันธ์ควรใช้วิธีประมาณค่าโดยใช้ค่าเฉลี่ยระหว่างค่าเฉลี่ยจากกลุ่มตัวอย่างและสมการถดถอย แต่ถ้าขนาดกลุ่มตัวอย่างเพิ่มขึ้นเป็น 10 หรือ 15 ที่มีตัวแปรควบคุมมีค่าสัมประสิทธิ์สหสัมพันธ์สูง ๆ ควรใช้วิธีประมาณค่าโดยสมการถดถอย

4. ในกรณีที่ผู้วิจัยไม่สามารถจะหาตัวแปรควบคุมใด ๆ ได้เลย ใช้เทคนิคของค่าเฉลี่ยจากกลุ่มตัวอย่างยังพอที่จะใช้ได้ ซึ่งจะให้ผลดีกว่าการคำนวณโดยไม่ได้ประมาณข้อมูลที่สูญหาย

ข. ข้อเสนอแนะเพื่อการวิจัยต่อไป

ควรทดลองหาความแม่นยำด้วยวิธีสมการถดถอย เมื่อมีตัวแปรควบคุม 2 หรือ 3 ในกรณีที่แต่ละตัวมีความสัมพันธ์กับตัวแปรเกณฑ์ขนาด 0.2 ถึง 0.4 และมีความสัมพันธ์ระหว่างกันพอสมควร เพราะอาจพบได้ทั่วไปในการปฏิบัติ