

บทที่ 1

บทนำ



1.1 ความนำ

ในการสำรวจตัวอย่างข้อมูลที่ได้จากการสำรวจจะมีความคลาดเคลื่อน (error) ที่สำคัญอยู่ 2 ส่วน ส่วนแรกเรียกว่าความคลาดเคลื่อนที่เกิดจากการเลือกตัวอย่าง (Sampling errors) หมายถึง ความคลาดเคลื่อนซึ่งเกิดจากการใช้ข้อมูลเพียงบางส่วนจากประชากรโดยไม่ใช้ข้อมูลทั้งหมด และส่วนที่สอง เรียกว่า ความคลาดเคลื่อนที่ไม่ได้เกิดจากการเลือกตัวอย่าง (Nonsampling errors) ความคลาดเคลื่อนประเภทนี้เกิดขึ้นในขั้นตอนอื่น ๆ ของการสำรวจตัวอย่าง

ปัญหาที่มักเกิดขึ้น เสมอในการเก็บข้อมูลปฐมภูมิคือการไม่ได้รับคำตอบ (Nonresponse) ซึ่งเป็นสาเหตุหนึ่งของความคลาดเคลื่อนที่ไม่ได้เกิดจากการเลือกตัวอย่าง และเป็นกรณีที่จะเกิดจากการไม่ยอมให้ความร่วมมือในการตอบคำถาม เกี่ยวกับ เรื่องบาง เรื่องที่ผู้ตอบคิดว่าอาจก่อปัญหาให้หรือต้อง เสียผลประโยชน์บางอย่าง เช่น การถามเกี่ยวกับรายได้ผู้ให้สัมภาษณ์อาจคิดว่า ถ้า เขาตอบไปแล้วอาจมีปัญหาในเรื่องของการเสียภาษี หรือการถามเกี่ยวกับเรื่องราคาสินค้า บางอย่าง เจ้าของห้างร้านอาจจะปฏิเสธในการให้คำตอบโดยสิ้นเชิง เนื่องจากเขาคิดว่าพนักงาน สัมภาษณ์จะนำข้อมูล เกี่ยวกับราคาสินค้าไปให้ร้านค้าคู่แข่ง เพื่อตัดราคาสินค้าทำให้ยอดขายของร้านเขาลดลง เป็นต้น สำหรับผลที่เกิดจากการไม่ได้รับคำตอบในข้อถามบางข้อคือทำให้ข้อมูลของตัวแปรที่สนใจศึกษานั้นขาดหายไป Haitosky (1968 :67) ได้กล่าวถึงลักษณะความแตกต่างของข้อมูลของตัวแปรที่ขาดหายไปเป็น 3 กรณี คือ

กรณีที่ 1 คำสั่ง เกตบางคำของตัวแปรหายไปอย่างสุ่ม

กรณีที่ 2 รายละ เอียดบางอย่างของตัวแปรขาดหายไป หมายความว่าผู้ที่ทำการวิจัย ไม่สามารถได้รับคำตอบ เกี่ยวกับรายละ เอียดบางอย่างที่ต้องการจากหน่วยที่ตก เป็นตัวอย่าง เช่น ในการสำรวจราคาผลไม้ที่ปลูกได้ในประเทศไทยสำหรับช่วง เดือนใด เดือนหนึ่ง ผู้วิจัยจะไม่สามารถทราบราคาของผลไม้บางชนิดได้ เนื่องจาก เดือนที่สำรวจไม่ใช่ฤดูของผลไม้ชนิดนั้น

กรณีที่ 3 คำสั่ง เกตบางค่าของตัวแปรไม่ได้หายไปอย่างสมบูรณ์ เป็นกรณีที่ผู้วิจัยเชื่อว่าการที่หน่วยตัวอย่างไม่ยอมตอบคำถามบางคำถามนั้น เนื่องจากมีเหตุผลส่วนตัวสำหรับการไม่ตอบคำถามนั้น

จากเหตุผลดังกล่าวข้างต้น ข้อถาม เกี่ยวกับ เรื่องรายได้ที่ไม่ได้รับคำตอบมักจะเป็นไปตามกรณีที่ 3 เพราะผู้ทำการสำรวจได้คาดการณ์ไว้ล่วงหน้าว่าผู้ตอบต้องมีรายได้จำนวนหนึ่ง ทำให้ไม่เป็นไปตามกรณีที่ 2 และข้อมูลที่ขาดหายไป อาจมีผลมาจากการปฏิเสธในการให้คำตอบโดยจงใจซึ่ง เป็นไปตามกรณีที่ 3 มากกว่าคำสั่ง เกตของข้อมูลขาดหายไปอย่างไม่ตั้งใจตามกรณีที่ 1

ในการวิจัยทางสังคมศาสตร์ ตัวแปรที่นักวิจัยมักคำนึงถึง เป็นลำดับแรกคือ "รายได้" เนื่องจากเป็นตัวแปรที่สามารถใช้วิเคราะห์ปัญหาต่าง ๆ ได้หลายกรณี เช่น

1. การศึกษาภาวะการว่างงาน ผู้วิจัยอาจต้องศึกษาว่ามีสาขาใดที่นักศึกษาจบออกมาแล้วส่วนใหญ่ไม่มีงานทำ เพราะผลกระทบที่ตามมาก็คือ นักศึกษากลุ่มนี้อาจจะต้องทำงานที่มีรายได้ต่ำกว่าวุฒิการศึกษาที่จบออกมา หรือในบางสาขาที่ยังไม่มีสถาบันใดเปิดทำการสอน แต่ตลาดมีความต้องการแรงงานทางด้านนี้ ผู้วิจัยอาจจะต้องทำการศึกษาดูว่าควรจะเปิดสาขานี้หรือไม่โดยประมาณจากรายได้ที่คาดว่าจะได้รับจากสาขานี้และความต้องการสาขานี้ในอนาคต

2. การร้องเรียนขอขึ้นค่าจ้างแรงงาน ผู้บริหารก็ต้องศึกษาถึงปัจจัยต่าง ๆ โดยอาศัยตัวแปร รายได้ เป็นหลัก

3. การจัดสรรงบประมาณสำหรับการพัฒนาประเทศ ต้องคำนึง ถึงการกระจายรายได้ของประชากรภายในประเทศ ผลผลิตมวลรวมทั้งประเทศหรือกล่าวอีกนัยหนึ่งคือรายได้ต่อคนของประชากร

จะเห็นได้ว่าตัวแปรรายได้ มีบทบาทสำคัญทั้งในการวิจัย การศึกษา การพัฒนาประเทศ และสวัสดิการด้านต่าง ๆ เป็นต้น ผลกระทบที่จะเกิดขึ้นอันเนื่องมาจากตัวแปรรายได้ คือ ถ้าผู้วิจัยไม่ได้รับคำตอบในเรื่องรายได้ ในอัตราค่อนข้างสูงแล้ว ในการประมาณค่าพารามิเตอร์เพื่อวิเคราะห์ปัญหาหรือพยากรณ์ค่าต่าง ๆ จะให้ผลการประมาณที่ไม่ถูกต้อง เนื่องจาก จำนวนตัวอย่างลดลงและหน่วยที่ไม่ตอบในเรื่องรายได้มักจะเป็นหน่วยที่เป็นปัญหา เช่น บุคคลนั้นมีรายได้น้อยเกินไป หรือมีรายได้สูงเกินไป ด้วยสาเหตุเหล่านี้ นักสถิติได้เริ่มให้ความสนใจศึกษาเกี่ยวกับตัวแปรรายได้ที่มีค่าบางค่าขาดหายไปมากขึ้น

Greenlees, Reece and Zieschang (1982 :251) ได้จำแนกวิธีการประมาณค่าที่ขาดหายไปของตัวแปรรายได้ เป็น 2 ประเภทคือ

(1) Ignorable Response Mechanism เป็นการประมาณค่ารายได้ โดยสมมติว่าการแจกแจงของการตอบ (response distribution) ไม่ขึ้นอยู่กับระดับของรายได้ของผู้ตอบ หรือกล่าวอีกนัยหนึ่งว่า ทำการประมาณค่าของรายได้โดยไม่สนใจสาเหตุของการไม่ตอบคำถามในเรื่องรายได้

(2) Nonignorable Response Mechanism เป็นการประมาณค่ารายได้ โดยสมมติว่าความน่าจะเป็นของการตอบเรื่องรายได้ขึ้นอยู่กับระดับรายได้ของผู้ตอบ

1.2 วัตถุประสงค์ของการวิจัย

1. ศึกษาวิธีการหาค่ารายได้ที่ขาดหายไปในการสำรวจตัวอย่าง
2. เปรียบเทียบวิธีการหาค่ารายได้ที่ขาดหายไปในการสำรวจตัวอย่างด้วยวิธีต่าง ๆ ว่าวิธีใดควรจะเป็นวิธีที่เหมาะสมและใกล้เคียงกับความเป็นจริงมากที่สุด

1.3 ขอบเขตของการวิจัย

1. ในการศึกษาวิธีการหาค่ารายได้ที่ขาดหายไปในการสำรวจตัวอย่างจะศึกษาในกรณี ที่ ความน่าจะเป็นที่ผู้ตอบให้คำตอบในเรื่องรายได้ไม่ขึ้นอยู่กับระดับรายได้ของผู้ตอบ
2. ข้อมูลในเรื่องรายได้ที่นำมาศึกษาครั้งนี้ใช้ข้อมูลเกี่ยวกับรายได้หลักของหัวหน้าครัวเรือนในกรุงเทพมหานคร เท่านั้น
3. หน่วยตัวอย่างที่นำมาศึกษา เลือกเฉพาะหน่วยตัวอย่างที่หัวหน้าครัวเรือนมีรายได้หลัก และทราบข้อมูลเกี่ยวกับตัวแปรอื่น ๆ ที่เกี่ยวข้อง

1.4 ข้อมูลและแหล่งที่มาของข้อมูล

การวิจัยครั้งนี้ได้ใช้ข้อมูลทุติยภูมิ (Secondary data) ในเรื่อง รายได้หลักของหัวหน้าครัวเรือน และตัวแปรอื่น ๆ ที่เกี่ยวข้อง จากโครงการวิจัยเรื่อง สภาวะทางเศรษฐกิจสังคมและประชากรของกรุงเทพมหานคร โดย ร่องศาสตราจารย์ ดร.เทียนฉาย กีระนันทน์ และคณะ สำหรับขนาดตัวอย่างและวิธีการเลือกตัวอย่าง เป็นดังนี้

1.4.1 ตัวอย่าง ขนาด และการเลือกตัวอย่าง¹

ในขั้นต้นก่อนการเลือกตัวอย่างนั้นได้มีการกำหนดกลุ่มประชากรเป้าหมาย (target population) ให้ชัดเจนเพื่อให้เลือกตัวอย่างได้และเป็นไปตามวัตถุประสงค์ของการวิจัย โดยได้กำหนดประชากรที่อยู่ในขอบข่ายของการสำรวจไว้ว่า หมายถึง ครว เรือนส่วนบุคคลที่อาศัยอยู่ในกรุงเทพมหานคร แต่ไม่รวมถึงครว เรือนสถาบัน ครว เรือนพิเศษ ครว เรือน-ทูตและชาวต่างประเทศที่อาศัยอยู่ชั่วคราว ทั้งนี้โดยอาศัยนิยามของสำนักงานสถิติแห่งชาติ ได้นิยามครว เรือนส่วนบุคคลตามความหมายที่อยู่ในขอบข่ายของการวิจัยนี้ และครว เรือนสถาบันกับครว เรือนพิเศษ ซึ่งอยู่นอกขอบข่ายของการวิจัยนี้ไว้ว่า

ครว เรือนส่วนบุคคล หมายถึง ครว เรือนที่ประกอบขึ้นด้วยบุคคลที่กินอยู่ร่วมกัน โดยมีความสัมพันธ์เกี่ยวข้องกันฉันญาติพี่น้อง หรือบุคคลไม่มีความสัมพันธ์ฉันญาติพี่น้องกันเลย แต่มาอยู่ร่วมกันด้วยความสมัครใจและมีจำนวนไม่เกิน 5 คน ครว เรือนส่วนบุคคลนี้จะอาศัยอยู่ในบ้านประเภทใดก็ได้ เช่น โรงเรือน ห้องแถว เรือนแพ เป็นต้น ครว เรือนส่วนบุคคลแบ่งออกเป็น 2 ประเภทคือ

ก. ครว เรือนคนเดียว ได้แก่ ครว เรือนที่ประกอบด้วย บุคคลคนเดียว อาจมีฐานะเป็นเจ้าของผู้เช่า ผู้อาศัย หรือผู้รับจ้างเฝ้าบ้าน โดยไม่เกี่ยวข้องเป็นสมาชิกของครว เรือนอื่นใดที่อยู่ในบ้านเดียวกันนั้น

ข. ครว เรือนหลายคน ได้แก่ ครว เรือนที่ประกอบด้วย บุคคลตั้งแต่สองคนขึ้นไปอยู่ร่วมกันในบ้านเดียวกัน และใช้สิ่งอุปโภค บริโภค อันจำเป็นแก่การครองชีพร่วมกัน บุคคลเหล่านี้จะเกี่ยวข้อง เป็นญาติหรือไม่เป็นญาติกัน เลยกก็ได้

ครว เรือนสถาบัน หมายถึง ครว เรือนที่ประกอบขึ้นด้วยบุคคลที่กินอยู่ร่วมกัน เนื่องจากเป็นสมาชิกของสถาบันเดียวกัน โดยไม่แยกที่อยู่เป็นส่วนเฉพาะคน เช่น สถานกักกัน วัด กรมทหาร นักเรียนที่อยู่ประจำที่โรงเรียน หรือในหอพักนักเรียน เป็นต้น

ครว เรือนพิเศษ หมายถึง ครว เรือนซึ่งประกอบด้วยบุคคลหลายคนอยู่กินร่วมกันเกินกว่า 5 คน ได้แก่ ที่พักคนงาน หอพักที่ไม่ใช่หอพักนักเรียนโดยเฉพาะ

¹เทียนฉาย กิระนันท์และคณะ, สภาวะทางเศรษฐกิจ สังคม และประชากรของกรุงเทพมหานคร; รายงานการวิจัย เล่ม 1 (จุฬาลงกรณ์มหาวิทยาลัย, 2525), หน้า 16-27.

ส่วนกรอบการสัมภาษณ์นั้นได้กำหนดไว้ดังนี้

ก. การสัมภาษณ์จะสัมภาษณ์หัวหน้าครัวเรือนที่ตกเป็นตัวอย่าง

ข. หัวหน้าครัวเรือนคือบุคคลที่เป็นหลักสำคัญในเชิง เศรษฐกิจของครัวเรือนมากที่สุด เช่น เป็นบุคคลที่หารายได้ให้ครัวเรือนมากที่สุด เป็นต้น

ค. ในกรณีที่มีบุคคลที่เป็นหลักสำคัญมากที่สุดดังกล่าวมากกว่า 1 คน หรือไม่อาจบอกได้เด็ดขาดว่าบุคคลใดเป็นหลักสำคัญมากที่สุด ให้เลือกสัมภาษณ์บุคคลที่อายุมากกว่า

ง. ในกรณีที่หัวหน้าครัวเรือนตามข้อ ข ไม่อยู่หรือสัมภาษณ์ไม่ได้ ให้สัมภาษณ์ภรรยาหรือบุคคลที่เป็นหลักสำคัญรองลงมาแทน

สำหรับขั้นตอนที่สำคัญในการเลือกตัวอย่างครัวเรือนมีดังนี้

ก. แบ่งบริเวณกรุงเทพมหานคร ซึ่งประกอบด้วย 24 เขตการปกครองออกเป็น 2 ส่วน คือ เขตเทศบาล และนอกเขตเทศบาล (รวมเขตสุขาภิบาล)¹

ข. แบ่งเขตชั้นเทศบาลออกเป็นชุมนุมอาคารย่อย ๆ โดยให้แต่ละชุมนุมอาคารประกอบด้วย จำนวนครัวเรือนประมาณ 75 ถึง 200 ครัวเรือน การแบ่งชุมนุมอาคารได้ยึดแนวถนน ลำคลอง ฯลฯ เป็นหลัก ได้จำนวนชุมนุมอาคารทั้งสิ้น 3,116 ชุมชนอาคาร ใช้วิธีสุ่มอย่างมีระบบ (systematic sampling) ที่สัดส่วนตัวอย่างเป็นหนึ่งในแปด ได้จำนวนชุมนุมอาคารทั้งสิ้น 403 ชุมชนอาคารกระจายกันอยู่ใน 24 เขต และ 82 แขวง

ค. เลือกหมู่บ้านตัวอย่างโดยวิธีสุ่มอย่างมีระบบ ที่สัดส่วนตัวอย่างเป็นหนึ่งในสิบจากหมู่บ้านทั้งหมดที่อยู่นอกเขตเทศบาลของเขตต่าง ๆ ในกรุงเทพมหานคร ได้จำนวนหมู่บ้านตัวอย่างทั้งสิ้น 72 หมู่บ้าน กระจายกันอยู่ใน 53 ตำบลของ 11 เขต คือ เขตบางกะปิ บางเขน พระโขนง มีนบุรี ลาดกระบัง หนองจอก ดลิ่งชัน บางขุนเทียน ภาษีเจริญ หนองแขม และราชบุรีบูรณะ

ง. เลือกครัวเรือนส่วนบุคคลขึ้นมาเป็นตัวอย่าง 7 ถึง 8 ครัวเรือนจากแต่ละชุมนุมอาคารตัวอย่าง และเลือกครัวเรือนส่วนบุคคลขึ้นมาเป็นตัวอย่าง 5 ครัวเรือนจากแต่ละหมู่บ้านตัวอย่าง โดยวิธีสุ่มอย่างมีระบบ ได้จำนวนครัวเรือนตัวอย่างรวมทั้งสิ้น 3,306 ครัวเรือน

¹ นิยามตามนัยแห่งกฎหมายเดิม คือ เขตเทศบาล และนอกเขตเทศบาล

รายละเอียดเกี่ยวกับจำนวนครัวเรือนตัวอย่างจำแนกตาม เขตได้แสดงไว้ในตารางต่อไปนี้

<u>เขต</u>	<u>จำนวนชมรมอาคาร/หมู่บ้านตัวอย่าง</u>	<u>ขนาดตัวอย่าง</u>
พญาไท	45	332
ดุสิต	33	237
ปทุมวัน	18	135
บ่อมปราบศตรุพ่าย	18	120
พระนคร	18	125
บางรัก	15	86
สัมพันธวงศ์	11	75
ห้วยขวาง	11	83
ธนบุรี	31	231
คลองสาน	20	144
บางกอกใหญ่	13	97
ยานนาวา	43	317
พระโขนง	56	418
บางเขน	23	156
บางกะปิ	16	115
บางกอกน้อย	27	198
ภาษีเจริญ	15	88
ราชบุรีบูรณะ	4	30
หนองจอก	10	50
มีนบุรี	8	50
ลาดกระบัง	5	35
บางขุนเทียน	12	78
ตลิ่งชัน	12	76
หนองแขม	5	30
รวม	469	3,306

1.5 วิธีดำเนินการวิจัย

1. ศึกษาวิธีการหาค่ารายได้ที่ขาดหายไป ตามวิธีต่าง ๆ จากวารสารวิชาการสถิติ
2. ศึกษาลักษณะของหน่วยตัวอย่างที่มีค่ารายได้ขาดหายไป โดยศึกษาถึงตัวแปรต่าง ๆ ที่มีความสัมพันธ์กับรายได้

3. ตัวแปรที่ใช้ในการวิจัยมีดังนี้

- สำหรับหัวหน้าครัวเรือน

- (1) เขตที่อยู่
- (2) อายุ (อายุเต็ม)
- (3) ระดับการศึกษา (เรียนจบชั้นอะไรโดยใช้ประเด็นของจำนวนปีการศึกษา)
- (4) สถานภาพการทำงานอาชีพหลัก
- (5) อาชีพหลัก
- (6) ชั่วโมงทำงานต่อสัปดาห์ของอาชีพหลัก
- (7) การได้รับค่าตอบแทนจากอาชีพหลัก
- (8) รายได้ต่อเดือนจากอาชีพหลักโดยประมาณ
- (9) รายจ่ายทั้งหมดต่อเดือนสำหรับครัวเรือน (บาท)
- (10) ระดับฐานะทางเศรษฐกิจของผู้ให้สัมภาษณ์

4. จากตัวอย่างทั้งหมด 3,306 ครัวเรือนของโครงการวิจัยเรื่อง สภาวะทาง

เศรษฐกิจ สังคม และประชากรของกรุงเทพมหานคร ดังได้กล่าวมาแล้ว เมื่อคัดหน่วยตัวอย่างที่ให้ค่าตัวแปรทุกตัวข้างต้น ยกเว้นตัวแปรรายได้ที่อาจได้รับค่าตอบหรือไม่ได้รับค่าตอบก็ได้ และคัดเลือกเฉพาะหัวหน้าครัวเรือนที่ได้รับค่าตอบแทนจากอาชีพหลักเท่านั้น เหลือจำนวนตัวอย่าง 1,513 ครัวเรือน ในจำนวนนี้มีพวกที่ไม่ตอบในเรื่องรายได้ 77 ครัวเรือน ดังนั้นจะเหลือครัวเรือนที่ให้ค่าตอบในเรื่องรายได้ 1,436 ครัวเรือน

5. นำครัวเรือนที่ให้ค่าตอบในเรื่องรายได้ 1,436 ครัวเรือน มาเลือกตัวแทนหน่วยที่ไม่ตอบ เรื่องรายได้ โดยที่ตัวแทนเหล่านี้มีลักษณะคล้ายกับหน่วยตัวอย่างทั้ง 77 ครัวเรือนที่มีรายได้ขาดหายไปดังกล่าวมาแล้วข้างต้น และสมมติว่า รายได้ของตัวแทน ที่เลือกขึ้นมาไม่ได้ขาดหายไปอย่างสุ่ม เพื่อศึกษาเกี่ยวกับวิธีการประมาณค่ารายได้ที่เหมาะสมสำหรับการวิจัยครั้งนี้โดยเปรียบเทียบ ผลการประมาณค่ารายได้ที่ได้จากวิธีต่าง ๆ กับรายได้ที่มีอยู่จริงของตัวแทนชุดที่เลือกขึ้นมาในลักษณะข้างต้นสำหรับวิธีการ เลือกตัวแทนในลักษณะดังกล่าวได้อาศัยตัวแปรพื้นฐานที่มีความ

สัมพันธ์กับรายได้ เป็นหลักในการคัดเลือก

6. นำครัวเรือนที่ให้คำตอบในเรื่องรายได้ 1,436 ครัวเรือน มาเลือกตัวแทนอย่างสุ่มเพื่อแทนหน่วยที่ไม่ตอบเรื่องรายได้ กล่าวคือไม่สนใจลักษณะของหน่วยตัวอย่างทั้ง 77 ครัวเรือนที่มีรายได้ขาดหายไป แล้วศึกษาวิธีประมาณค่ารายได้ เช่นเดียวกับวิธีที่จะนำมาใช้ในข้อ 5

7. ทดสอบวิธีต่าง ๆ ที่ใช้ในการประมาณค่าที่ขาดหายไป เพื่อหาวิธีที่เหมาะสมที่สุดในการหาค่ารายได้ที่ไม่ได้รับคำตอบ

1.6 ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย

1. ช่วยให้เข้าใจลักษณะของผู้ที่ไม่ตอบในเรื่องรายได้ เพื่อเป็นประโยชน์ในการสำรวจตัวอย่างครั้งต่อไป
2. เป็นแนวทางในการตัดสินใจว่าจะใช้วิธีใดในการวิเคราะห์ข้อมูลเพื่อหารายได้ของผู้ที่ไม่ตอบ
3. นำผลที่ได้ไปประยุกต์ใช้ในงานวิจัยอื่นที่เกี่ยวข้องกับรายได้ในกรณีที่กำลัง เกิดขาดหายไป

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย