

## การสำรวจตัวอย่าง: เครื่องมือในการวิจัย

ประชุม สุวัฒน์\*

### 1. คำนำ

การวิจัยโดยการสำรวจ เป็นวิธีการวิจัยที่ใช้กันแพร่หลายมากที่สุด โดยเฉพาะอย่างยิ่งในวงการวิจัยสาขาสังคมศาสตร์ ปัจจุบันการวิจัยทางสังคมวิทยา รัฐประศาสนศาสตร์ จิตวิทยา การบริหารธุรกิจ สาธารณสุขศาสตร์ ภูมิศาสตร์ ประชากรศาสตร์ ฯลฯ ในประเทศไทยอาศัยการสำรวจเป็นเครื่องมือที่สำคัญในการเก็บรวบรวมข้อมูล ดังจะเห็นได้ว่าการสำรวจประชามติ การวิจัยตลาด การสำรวจความต้องการของชาวนาหรือของโรงงาน ส่วนใหญ่เป็นการสำรวจด้วยตัวอย่าง แทนที่จะเป็นการสำมะโน

ผลงานของ Jacob Bernoulli ใน ค.ศ. 1713 เป็นรากฐานสำคัญของทฤษฎีความน่าจะเป็น และเป็นเอกสารฉบับแรกที่เกี่ยวข้องกับทฤษฎีของการสุ่มตัวอย่าง<sup>1</sup> ส่วนเอกสารอื่นๆ ที่เกี่ยวข้องกับทฤษฎีการสุ่มตัวอย่างและที่ตีพิมพ์ก่อนปี ค.ศ. 1900 มีน้อยมาก นับแต่ต้นคริสต์ศตวรรษที่ 20 เป็นต้นมา มีผู้ใช้การสุ่มตัวอย่างกันอย่างกว้างขวาง และมีนักสถิติศึกษาทฤษฎีและวิธีการสุ่มตัวอย่างกันมากมาย เช่น ในปี ค.ศ. 1934 Jersey Neyman<sup>2</sup> ได้เสนอวิธีการสุ่ม (randomization) อันดีอ ได้ว่าเป็นจุดเริ่มต้นของการสุ่มตัวอย่าง โดยอาศัยความน่าจะเป็น ในปี ค.ศ. 1943 Hansen and Hurwitz<sup>3</sup> ได้เสนอแนวความคิดเกี่ยวกับการสุ่มตัวอย่าง โดยกำหนดความน่าจะเป็นไม่เท่ากัน พร้อมทั้งการสุ่มหลายชั้น ฯลฯ

ในบทความนี้ ผู้เขียนมุ่งเสนอลักษณะวิธีการสุ่มตัวอย่างเป็นระบบตามแบบวิธีที่ใช้ความน่าจะเป็น พร้อมทั้งจะชี้ให้เห็นความสำคัญของการสุ่มตัวอย่างต่อ

\* ศาสตราจารย์ รองอธิการบดีฝ่ายบริหาร สถาบันบัณฑิตพัฒนบริหารศาสตร์

การวิจัย แต่ก่อนที่จะเสนอวิธีการสุ่มตัวอย่างแบบต่าง ๆ จะได้เสนอานิยามของคำที่เกี่ยวข้อง ลำดับขั้นการดำเนินงานสำรวจด้วยตัวอย่าง โครงการการสำรวจที่ดำเนินอยู่ทั้งในและนอกประเทศบางโครงการ ตลอดจนปัญหาบางประการที่เกิดขึ้นในการดำเนินการสำรวจด้วยตัวอย่าง

1.1 บทบาทของการสำรวจด้วยตัวอย่างในการวิจัย ในปัจจุบันการสำรวจด้วยตัวอย่างถือกันโดยทั่วไปทั่วโลกว่าเป็นวิธีที่ใช้ในการศึกษาคุณลักษณะและความสัมพันธ์ของหน่วยของประชากร (ในความหมายทางสถิติ) โดยการเลือกตัวอย่างเป็นตัวแทนของทั้งหมด เก็บรวบรวมข้อมูลจากหน่วยในตัวอย่างมาศึกษาและใช้ข้อมูลเหล่านั้นให้เกิดประโยชน์ คือประมาณค่าหรือทดสอบความจริงเกี่ยวกับคุณสมบัติของประชากรที่ทำการศึกษา การดำเนินการเช่นกล่าวนี้เป็นไปในแนวเดียวกับการวิจัยโดยทั่วไป แม้ว่าการวิจัยอาจจะใช้วิธีอื่นในการเก็บรวบรวมข้อมูลก็ได้ เช่น อาจจะทำการศึกษาเฉพาะกรณี (Case study) การทำการทดลองเพื่อให้เกิดข้อมูลเพื่อนำมาศึกษา (Controlled experiment) การสังเกตโดยตรง การวิเคราะห์ข้อมูลที่มีอยู่แล้ว ฯลฯ แต่การสำรวจด้วยตัวอย่างเป็นวิธีที่ใช้กันแพร่หลายมากที่สุด หากดำเนินการไปโดยรอบคอบควบคุมให้ดีจะได้ข้อมูลที่มีคุณภาพสูง คือวัดสิ่งที่ต้องการได้ถูกต้องเหมาะสมโดยรวดเร็ว ซ้ำยังสามารถใช้จ่ายเงินให้อยู่ในวงจำกัดได้อีกด้วย ถ้าเราจะพิจารณาประเภทของข้อมูลโดยแบ่งตามวิธีการรวบรวมข้อมูล เราอาจจะแบ่งได้เป็น 2 พวกใหญ่ ๆ คือ

(1) ข้อมูลที่ได้จากการทดลอง (Experimental data) หมายถึงข้อมูลที่ได้โดยปกติมิได้มีอยู่ แต่หากสนใจจะต้องทำให้เกิดขึ้น เช่น น้ำหนักของพืชอันเป็นผลจากการใช้ปุ๋ย ระยะทางที่รถยนต์ชนิดหนึ่งวิ่งได้ต่อน้ำมัน 1 ลิตร ปริมาณการขายของบริษัทอันเป็นผลจากการโฆษณาวิธีต่าง ๆ ฯลฯ

(2) ข้อมูลที่ได้จากการสำรวจ (Survey data) หมายถึงข้อมูลที่ตามปกติมีอยู่แล้ว หากต้องการก็ไปรวบรวมมา เช่น จำนวนคนในครัวเรือน รายได้ของคนงานในโรงงาน ความเป็นอยู่ของคนในหมู่บ้าน ฯลฯ การรวบรวมข้อมูลอาจทำได้โดยการสังเกตหรือวัดโดยตรง หรือโดยการสอบถามก็ได้ ในบางกรณีอาจจะมีผู้เก็บรวบรวมข้อมูลจากรายงานหรือสมุดหนังสือที่มีข้อมูลสถิติ โดยการคัดลอกหรือเลือกมาใช้ ในกรณีที่มีการเลือกใช้ข้อมูลบางส่วน การสุ่มตัวอย่างย่อมมีบทบาทสำคัญเช่นเดียวกับการสุ่มตัวอย่างจากประชากรอื่น ๆ

การวิจัยทางสังคมศาสตร์มักมีการรวบรวมข้อมูลประเภทหลัง การสำรวจด้วยตัวอย่างจึงมีบทบาทสำคัญยิ่งในการเก็บรวบรวมข้อมูลให้ถูกต้อง เหมาะสมกับสภาพการวิจัย โดยเสียเวลาน้อยและประหยัดค่าใช้จ่าย ในบางกรณีอาจมีการรวบรวมข้อมูลประเภทแรก คือ ทำการทดลองเพื่อให้เกิดข้อมูลที่สนใจ ในการนี้ การสุ่มตัวอย่างมีบทบาทสำคัญในการกำหนด หน่วยตัวอย่างที่ใช้ในการทดลองนั้น ๆ โดยมักถือว่าเป็นตัวอย่างสุ่ม (Random sample) จากประชากรที่ประกอบไปด้วยหน่วยตัวอย่างที่สนใจทั้งหมด การสำรวจด้วยตัวอย่างจึงเป็นเครื่องมือที่ใช้ในการรวบรวมข้อมูลที่สำคัญยิ่งในการดำเนินการวิจัย การที่ข้อยุติจากการวิจัยจะถูกต้อง ย่อมต้องอาศัยข้อมูลที่ถูกต้อง และเป็นตัวแทนของสิ่งที่จะศึกษาได้อย่างถูกต้อง ต้องการตัวอย่างที่ดีเหมาะสม วิธีการสุ่มตัวอย่างและการสำรวจด้วยตัวอย่าง จึงเป็นสิ่งที่มีความสำคัญยิ่งต่อการวิจัยชนิดต่าง ๆ

**1.2 การสำมะโนและการสำรวจด้วยตัวอย่าง** การเก็บรวบรวมข้อมูลที่สนใจอาจจะเก็บมาจากหน่วยทุกหน่วย (ซึ่งอาจจะเป็นคน ครัวเรือน ร้านค้า หมู่บ้าน โรงงาน ฯลฯ) หรืออาจจะรวบรวมมาจากบางหน่วย ในกรณีแรกเป็นการสำรวจมาจากทุกหน่วยในประชากร เราเรียกว่า การสำมะโน (Census or

complete enumeration survey) ในกรณีหลังเป็นการสำรวจด้วยตัวอย่าง (Sample survey) โดยทั่วไปการเก็บรวบรวมข้อมูลโดยวิธีสำมะโน ย่อมต้องการแรงงาน เวลา และค่าใช้จ่าย มากกว่าวิธีสำรวจด้วยตัวอย่าง การจะเลือกใช้โดยวิธีใดนั้น ผู้ดำเนินการวิจัยจะต้องชั่งใจดูให้ดี ควรพิจารณาว่าต้องการข้อมูลจากทุกหน่วยหรือไม่ จำเป็นให้ได้ข้อมูลที่ครอบคลุมร้อยละเก้าสิบหรือเปล่า การเก็บข้อมูลจากทุกหน่วยจะให้ข้อมูลที่ถูกต้องตามต้องการหรือไม่ และควรพิจารณาด้วยว่าจะใช้การสำรวจด้วยตัวอย่างแทนการสำมะโนที่จะทำนั้นดีกว่าหรือไม่

การที่เราจะตัดสินใจเลือกใช้การสำรวจชนิดใดก็ตาม ขึ้นอยู่กับจุดประสงค์ของการรวบรวมข้อมูลนั้น ในปัญหาที่ว่าต้องการข้อมูลจากทุกหน่วยหรือไม่นั้น เป็นปัญหาที่เกี่ยวข้องกับจุดประสงค์ของการรวบรวมข้อมูลโดยตรง หากไม่ต้องการข้อมูลแสดงรายละเอียดเป็นข้อมูลของแต่ละหน่วย การรวบรวมข้อมูลโดยวิธีอื่นต่างจากการสำมะโนอาจจะเป็นสิ่งที่ควรทำ อาจจะประหยัดและทำงานได้สะดวกและรวดเร็วกว่า ในกรณีที่ต้องพิจารณาจากแต่ละหน่วยแยกจากกัน เช่น ข้อมูลเพื่อประโยชน์ในการเก็บภาษี รายชื่อผู้มีสิทธิออกเสียงเลือกตั้ง รายชื่อผู้จะเข้ารับการศึกษา ฯลฯ การสำมะโนเป็นสิ่งจำเป็น แต่หากต้องการข้อเท็จจริงที่แสดงคุณสมบัติของประชากรโดยไม่สนใจข้อมูลแยกเป็นหน่วย ๆ ก็ไม่จำเป็นต้องใช้การสำมะโน แต่อาจใช้การสำรวจด้วยตัวอย่างก็สามารถหาข้อเท็จจริงที่ต้องการได้

ในหลายกรณีเราไม่ต้องการข้อเท็จจริงที่ถูกต้องหมดร้อยเปอร์เซ็นต์ ภายใต้งานที่เรามีอยู่รวมทั้งหมดอันจำกัด เราต้องการข้อเท็จจริงที่มีผิดพลาดในขอบเขตที่จะยอมรับได้ ขนาดของข้อผิดพลาดที่ยอมรับได้เราเรียกว่า permissible error เช่นในการกะประมาณจำนวนประชากรในประเทศไทย ในขณะนี้เรามักจะ

กล่าวเป็นตัวเลขลงตัวเป็นล้านแสนหมื่นพันร้อยคน ในทำนอง 43,100,000 คน  
 ทั้ง ๆ ที่จำนวนที่แท้จริงอาจจะเป็น 43,110,017 คน เป็นต้น ข้อผิดพลาดใน  
 กรณีนี้คือ 0.02 % อยู่ในขอบเขตที่จะยอมรับได้ แต่หากจะใช้ตัวเลข 35,000,000  
 คน แทนจำนวนจริง เช่นนี้มีข้อผิดพลาดถึง 18.81 % ซึ่งมีขนาดใหญ่เกิน  
 กว่าที่จะยอมรับได้

ในปัญหาเกี่ยวกับความถูกต้องของข้อมูลนี้ควรระวังไว้ด้วยว่า การสำรวจ  
 โดยวิธีสำมะโนไม่จำเป็นต้องให้ข้อเท็จจริงถูกต้องเสมอไป ในทางปฏิบัติ หากการ  
 ควบคุมการทำงานของพนักงานสนาม การให้คำแนะนำไม่ดีพอแล้ว การสำมะโน  
 จะก่อให้เกิดข้อผิดพลาดมากกว่าการสำรวจด้วยตัวอย่างด้วยซ้ำ ทั้งนี้เพราะว่าการ  
 การสำรวจด้วยตัวอย่างเป็นงานขนาดเล็กกว่าการสำมะโน การควบคุมคุณภาพย่อม  
 ทำได้ง่ายกว่า คุณภาพของข้อมูลจะดีกว่า และอาจใช้ทฤษฎีและวิธีการทางสถิติ  
 ช่วยให้เราสามารถประมาณค่าพารามิเตอร์ ที่สนใจได้ถูกต้องกว่าข้อมูลที่ ได้จาก การ  
 สำมะโน ในการทำสำมะโนอาจมีข้อผิดพลาดอันเนื่องมาจากนิยามของคำที่ใช้  
 การที่คลุมเครือครอบคลุมทุกหน่วย การวัดคลาดเคลื่อน และการประมวลผลคลาด-  
 เคลื่อนเนื่องมาจากปริมาณข้อมูลที่เกี่ยวข้องมีมากมายเหลือเกิน ส่วนในการสำรวจ  
 ด้วยตัวอย่างปริมาณงานเล็กกว่าในการสำมะโน จึงสามารถควบคุมคุณภาพของ  
 ข้อมูลที่ได้ดีกว่า

ในการพิจารณาใช้การสุ่มตัวอย่างแทนการสำมะโนนั้น บางที่เราจำเป็นต้องใช้  
 การสุ่มตัวอย่าง โดยเฉพาะอย่างยิ่งเมื่อการสำรวจนั้นจะทำลายหน่วยตัวอย่าง เช่น  
 การทดสอบเพื่อหาอายุการใช้งานของหลอดไฟ เหตุผลสำคัญที่เราใช้การสุ่มตัว-  
 อย่างแทนการสำมะโนก็คือ การประหยัดทรัพยากรทั้งทางด้านแรงงาน เวลาและ  
 ค่าใช้จ่าย

1.3 ปัญหาเกี่ยวกับตัวอย่าง ในการสำรวจด้วยตัวอย่าง เรามีการเลือกตัวอย่าง (ส่วนหนึ่ง) จากประชากร แล้วใช้ข้อเท็จจริงจากตัวอย่างที่สำรวจมาให้เกิดประโยชน์ คือ ประมาณค่าของคุณลักษณะบางประการที่สนใจของประชากรที่ศึกษา ถ้าถามที่มีผู้ถามนักสถิติท่านน้อยๆ ก็คือ แน่ใจได้อย่างไรว่าตัวอย่างที่เลือกมานั้นเป็น “ตัวแทน” ที่ดีของประชากร หากจะตอบคำถามนี้ให้เข้าใจได้ดี ผู้ฟังก็ต้องรู้หลักวิชาสถิติบ้าง แต่หากจะพูดกันด้วยภาษาชาวบ้านก็บอกได้ง่าย ๆ ว่า ตัวอย่างที่ได้มาตามวิธีที่ใช้เลือกโดยไม่มีควมลำเอียง ถือได้ว่าเป็นตัวอย่างที่ดี เป็นตัวแทนที่ดีของประชากรทั้งนั้น หากจะใช้ค่าทางคำนวณวิชาการสักเล็กน้อย จะต้องพูดกันถึงการสุ่มตัวอย่างแบบใช้ความน่าจะเป็น (Probability sampling) กับแบบไม่ใช้ความน่าจะเป็น (Non-probability sampling)

ในทางสถิติเมื่อพูดถึง “ตัวอย่าง” เราหมายถึง “ตัวอย่างสุ่ม” (Random sample) ซึ่งได้แก่ตัวอย่าง (ส่วนหนึ่งของประชากร) ที่เลือกโดยอาศัยหลักเกณฑ์เกี่ยวกับความน่าจะเป็น กล่าวคือ เป็นตัวอย่างที่เลือกมาโดยหน่วยแต่ละหน่วยในประชากรมีโอกาสหรือความน่าจะเป็นที่จะถูกเลือกมาใช้เป็นตัวอย่างตามที่กำหนดล่วงหน้า เช่น หน่วยแต่ละหน่วยมีโอกาสที่จะถูกเลือกใช้เท่า ๆ กัน เป็นต้น

การสุ่มตัวอย่างแบบใช้ความน่าจะเป็นได้แก่ การสุ่มตัวอย่างชนิดที่หน่วยทุกหน่วยในประชากรมีโอกาสที่จะถูกเลือกใช้เป็นตัวอย่าง ตัวอย่างนั้นจะเลือกโดยกระบวนการสุ่ม (Randomization process) เพื่อให้ความน่าจะเป็น เป็นไปตามที่กำหนดไว้ และในการวิเคราะห์ข้อมูลที่รวบรวมจากตัวอย่างมีการใช้นาหนักเหมาะสมกับความน่าจะเป็นที่กำหนดด้วย

ตัวอย่างที่เลือกมาโดยวิธีอื่น เรามักถือว่าไม่เป็นตัวอย่างสุ่ม ตัวอย่างชนิดนี้อาจเลือกมาโดยใช้วิจารณญาณของผู้เลือก ในฐานะที่คุ้นเคยกับประชากรที่ หรือ คาดว่าจะได้ “ตัวอย่างที่ดี” ตามความคิดของผู้เลือก ผู้เลือกอาจใช้ข้อเท็จจริง บางประการที่ทราบล่วงหน้า เจาะจงเลือกบางหน่วยที่คิดว่าจะเป็นตัวแทนที่ดีของ ประชากรได้ วิธีการสุ่มตัวอย่างแบบนี้มีชื่อเรียกกันว่า *Judgement sampling* หรือ *Purposive sampling* วิธีการแบบนี้ไม่เป็นที่ยอมรับ และใช้น้อยในวง การสำรวจขนาดใหญ่ เพราะที่ไม่สามารถใช้ข้อเท็จจริงในการประมาณค่าที่สนใจ ได้ถูกต้องตามหลักวิชาการ และนอกจากนั้นยังไม่สามารถประมาณค่าความคลาด-เคลื่อน อันเนื่องจากการสุ่มตัวอย่าง (*Sampling error*) การสุ่มแบบกำหนด โควตา (*Quota sampling*) คือกำหนดคุณสมบัติของหน่วยตัวอย่างที่ต้องการ ไว้ล่วงหน้า เป็นตัวอย่างของตัวอย่างที่มีใช้ตัวอย่างสุ่ม วิธีการสุ่มตัวอย่างแบบนี้ มักใช้กันทั่วไปในการวิจัยตลาด และการสำรวจความคิดเห็น แต่เริ่มใช้น้อยลง ทุกที่ ๆ ที่กล่าวมานี้มิได้หมายความว่าใช้ได้ตัวอย่างดังกล่าวไม่ได้ เราอาจจะใช้ ตัวอย่างไม่สุ่มในการทดสอบแบบสอบถาม หรือทดสอบวิธีการที่จะใช้ในการสำ-รวจ อย่างไรก็ตาม หากต้องการให้การสำรวจมีผลถูกต้องตามหลักวิทยาศาสตร์ และใช้ผลนั้นในการวางแผนหรือสรุปผลการวิจัยแล้วควรจะใช้การสุ่มตัวอย่างแบบ ใช้ความน่าจะเป็น

ที่เราจะพิจารณาต่อไปนี้เป็นวิธีการสุ่มตัวอย่างแบบใช้ความน่าจะเป็นเท่านั้น

## 2. แนวความคิดเบื้องต้น

2.1 นิยามของคำที่เกี่ยวข้อง เพื่อให้การพิจารณาวิธีการสำรวจด้วยตัว อย่างเป็นไปโดยสะดวกและถูกต้อง เราควรพิจารณาความหมายของคำต่าง ๆ ที่ เกี่ยวข้องก่อน

คำว่า “หน่วยตัวอย่าง” หรือ “หน่วย” (Sampling unit, elementary unit or unit) ได้แก่สมาชิก หรือกลุ่มสมาชิกของเซตๆ หนึ่งที่จะให้ข้อมูล สถิติที่เราสนใจได้ เช่น บุคคล ครอบครัว ครัวเรือน ร้านค้า โรงงาน ตลาด ฯลฯ หน่วยตัวอย่างอาจจะแตกต่างจากหน่วยที่ใช้เก็บข้อมูล (Recording unit) เช่น เมื่อหน่วยตัวอย่างเป็นครัวเรือน หน่วยที่เก็บข้อมูลอาจจะได้แก่คนแต่ละคนใน ครัวเรือนตัวอย่างก็ได้ ส่วนหน่วยที่ใช้ในการวิเคราะห์ (Unit of analysis) อาจ จะแตกต่างกันออกไปอีก หรือเหมือนกับหน่วยตัวอย่างก็ได้

เซตของหน่วยตัวอย่างทั้งหมดที่ต้องการจะศึกษา เราเรียกกันว่า “ประชากร” (Population or universe) ควรสังเกตว่าประชากรในที่นี้มิได้แปลว่า กลุ่มคน หากจะเป็นกลุ่มคนก็เป็นกรณีพิเศษเท่านั้น เช่น ประชากรอาจจะได้แก่ ครัวเรือนทั้งหมดในเขตกรุงเทพมหานคร หรืออาจจะได้แก่โรงงานทั้งหมดใน ประเทศ หรือหมู่บ้านทั้งหมดในประเทศ หรือสัตว์เลี้ยงทั้งหมดในภาคหนึ่ง ของประเทศ ฯลฯ ในบางกรณีในการทำกรวิจัยหรือสำรวจ เราอาจจะสนใจ เฉพาะบางส่วนของประชากรก็ได้ ในกรณีเช่นนั้นเราใช้คำว่า “ประชากรที่ ศึกษา” (Domain of study or sub-population)

ในทางทฤษฎีการสุ่มตัวอย่างเราถือว่า “ขนาด” ของประชากร ได้แก่ จำนวนหน่วยตัวอย่างของประชากรนั้น มิได้หมายความว่าถึงขนาดความใหญ่โต ใน ลักษณะของความกว้าง ความยาว ความสูง ฯลฯ ของประชากรเลย ในทำนอง เดียวกัน “ขนาด” ของตัวอย่างได้แก่ จำนวนหน่วยตัวอย่างในตัวอย่างที่สุ่มมา จากประชากรที่ศึกษา

“กรอบตัวอย่าง” (Sampling frame) ได้แก่รายการหน่วยตัวอย่างทั้งหมด ของประชากรที่ศึกษารวมทั้งตำบลที่อยู่ หรือสิ่งที่ใช้แสดงหน่วยตัวอย่างนั้นๆ



หรืออาจจะเป็นแผนที่ที่แสดงอาณาเขตของหน่วยตัวอย่าง (อันได้แก่พื้นที่) กรอบตัวอย่างชนิดแรกเรียกว่า “กรอบรายการ” (List frame) และประเภทหลังเรียกว่า “กรอบแผนที่” (Map frame or area frame) เนื่องจากกรอบตัวอย่างเป็นสิ่งสำคัญใช้ในการเลือกตัวอย่าง จึงควรให้ความสนใจว่ากลุ่มหน่วยตัวอย่างทั้งหมดในประชากรที่ศึกษา และไม่รวมหน่วยตัวอย่างของประชากรอื่น กรอบตัวอย่างควรสะท้อนสมัย ถูกต้อง ไม่มีหน่วยตัวอย่างใดอยู่ซ้ำ ในบางกรณีกรอบตัวอย่างยังประกอบด้วยข้อมูลอื่นที่จะช่วยให้การสุ่มตัวอย่างมีประสิทธิภาพสูงขึ้น ในการวางแผนการสำรวจด้วยตัวอย่าง ผู้วางแผนควรให้ความสนใจต่อการสร้างกรอบตัวอย่างให้มาก เพราะถือว่าเป็นรากฐานสำคัญของความถูกต้อง ความสมบูรณ์ของข้อมูลที่ใช้ในการศึกษาวิจัย หรือการรวบรวมข้อมูลให้ผู้อื่นเป็นผู้ใช้ ปัญหาเกี่ยวกับกรอบตัวอย่างเป็นปัญหาใหญ่ในการดำเนินการสำรวจด้วยตัวอย่างทั่วไป

ในการรวบรวมข้อมูลจากหน่วยตัวอย่างในตัวอย่างสุ่มเราอาจทำได้หลายวิธี แต่เมื่อได้ข้อมูลมาแล้วเราจะนำเอาข้อมูลนั้นมาวิเคราะห์ด้วยจุดประสงค์ที่จะเรียนรู้เกี่ยวกับ “พารามิเตอร์” (Parameter) อันได้แก่สิ่งที่แสดงคุณสมบัติของประชากรที่ศึกษา เราใช้ข้อมูลนั้นมาคำนวณค่าสถิติที่ใช้เป็น “ตัวประมาณค่า” (Estimate) ของพารามิเตอร์ ซึ่งโดยทั่วไปก็ย่อมผิดพลาดได้ แต่หากใช้วิธีการที่ถูกต้องเราเชื่อได้ว่าข้อผิดพลาดจะน้อย การสุ่มตัวอย่างจะทำให้เราสามารถกะประมาณได้ว่า มีความคลาดเคลื่อนสักเท่าใดก็กด้วย

2.2 ขั้นตอนการดำเนินงานสำรวจด้วยตัวอย่าง การดำเนินการสำรวจด้วยตัวอย่าง มีลักษณะคล้ายกับการดำเนินการวิจัยโดยทั่วไป กล่าวคือมีขั้นตอนเตรียมงานหรือวางแผนงานขั้นเก็บรวบรวมข้อมูล ขั้นประมวลผลข้อมูลซึ่งอาจรวมการวิเคราะห์ข้อมูลไว้ด้วย และขั้นการเตรียมรายงาน

2.2.1 การวางแผนงานสำรวจด้วยตัวอย่าง ประกอบไปด้วยการตั้งจุดประสงค์การดำเนินการทางกฎหมาย (หากมี), การประชาสัมพันธ์ การกะงบประมาณค่าใช้จ่าย การสร้างปฏิทินการสำรวจ การเตรียมค้ำบริหารงานสำรวจ การร่วมมือกับหน่วยงานที่เกี่ยวข้อง การเตรียมแผนที่และเตรียมรายการหน่วยตัวอย่าง วางแผนเกี่ยวกับการเก็บรวบรวมข้อมูล การประมวลผล การสร้างแบบสอบถาม การกำหนดวิธีการสุ่มตัวอย่าง การทดลองดำเนินงาน และทดสอบแบบสอบถาม การเตรียมงานบุคลากรรวมทั้งการฝึกพนักงานด้วย ในที่นี้เราไม่สามารถพิจารณารายละเอียดที่เกี่ยวข้องได้

2.2.2 การเก็บรวบรวมข้อมูล อาจทำได้โดยการสังเกตหรือวัดโดยตรง การสัมภาษณ์โดยตรง การสอบถามทางไปรษณีย์ ทางโทรศัพท์ การลงทะเบียน และการคัดลอกจากรายงานสถิติ วิธีต่าง ๆ ย่อมมีข้อดีข้อเสียของตน ควรเลือกใช้ให้เหมาะสมกับจุดประสงค์ เวลา งบประมาณ และกำลังคนที่มี

หากมีการใช้แบบสอบถามต้องระมัดระวังให้ดี เพราะแบบสอบถามมีผลสำคัญต่อคุณภาพของข้อมูลที่จะได้ แบบสอบถามต้องให้ข้อมูลที่ต้องการ ได้ครบแต่จะต้องกระชับรัดกุม เพราะหากแบบสอบถามยาวเกินความจำเป็น คุณภาพของคำตอบที่ได้ในตอนท้ายของการสัมภาษณ์จะค่อยลง และอาจจะถึงขนาดเชื่อไม่ได้เลยก็เป็นได้ การตั้งคำถามก็เป็นสิ่งสำคัญที่ควรระมัดระวังให้มาก เมื่อสร้างแบบสอบถามเสร็จแล้ว ควรทดสอบและแก้ไขให้เหมาะสม เมื่อจะใช้ต้องมีการอบรมฝึกพนักงานเจ้านับให้ที่ดีด้วย

ข้อมูลที่รวบรวมจะต้องสัมพันธ์กับเวลาที่ต้องการ (Reference period) ซึ่งอาจจะต่างจากเวลาที่สำรวจ (Survey period) และเวลาที่รายงาน (Reporting period) และที่สำคัญยิ่งอีกประการหนึ่งก็คือ การควบคุมและการให้คำแนะนำ

(Supervision) หากเป็นการสำรวจขนาดใหญ่จะต้องเตรียมเรื่องนี้ให้ดี แต่ในงานสำรวจขนาดย่อมการควบคุมและแนะนำก็มีความสำคัญยิ่ง คุณภาพของข้อมูลจะดีขึ้นขึ้นอยู่กับ การควบคุมและแนะนำมาก ในการเก็บรวบรวมข้อมูลนี้ ต้องพยายามขจัดความลำเอียงของพนักงานเฝ้าให้มากที่สุดเท่าที่จะทำได้ การเตรียมงานอย่างดี การเตรียมแบบสอบถามดี การอบรมฝึกฝนพนักงาน และการควบคุมแนะนำการทำงาน ล้วนแต่ทำให้คุณภาพของข้อมูลที่ดีขึ้นทั้งสิ้น

2.2.3 การประมวลผลข้อมูล รวมถึงการสร้างตารางจากข้อมูลที่รวบรวมได้ การวิเคราะห์ด้านวิชาการหรือสิ่งที่สนใจ และการวิเคราะห์ด้านสถิติ

การสร้างตารางเป็นการสรุปข้อมูลที่รวบรวมได้ไว้ในรูปตาราง การเสนอตารางนั้นควรจะมีข้อเท็จจริงเกี่ยวกับภูมิหลัง เช่น จุดประสงค์ของการสำรวจ วิธีการสุ่มตัวอย่างที่ใช้ วิธีเก็บรวบรวมข้อมูล และการสร้างตาราง รวมทั้งความคลาดเคลื่อนของข้อมูล ทั้งนี้โดยมุ่งหวังให้ผู้ใช้อข้อมูลและสถิติที่เสนอสามารถแปลความหมายของข้อมูลได้ถูกต้องและเที่ยงตรง

การวิเคราะห์เนื้อหาวิชาที่เกี่ยวข้อง กระทำหลังจากได้สร้างตารางสถิติแล้ว อาจรวมถึงการพิจารณาตารางสองทาง หรือสามทาง แยกตามเกณฑ์ต่างๆ มากกว่าหนึ่งเกณฑ์ เป็นการวิเคราะห์ทางด้านเทคนิค ซึ่งอาจกินเวลา จึงควรกระทำแยกจากส่วนแรก

การวิเคราะห์ทางด้านสถิติที่กล่าวถึงในที่นี้ หมายถึงการวิเคราะห์เพื่อประโยชน์ในการปรับปรุงเทคนิคการสำรวจทางสถิติ เช่น การวิเคราะห์เกี่ยวกับความคลาดเคลื่อนเกี่ยวกับการสุ่มตัวอย่าง (Sampling error) ความแปรปรวน ค่าใช้จ่าย และความคลาดเคลื่อนอันมิได้เนื่องจากการสุ่มตัวอย่าง (Non-sampling error) ที่เกิดขึ้นในการสำรวจชั้นต่าง ๆ การวิเคราะห์เช่นนี้อาจกินเวลาเช่นเดียว

กัน จึงควรพิจารณาแยกจากผลการสำรวจ และผู้ทำการสำรวจส่วนใหญ่มักจะละเลย ไม่พิจารณากันเสียเลย

การประมวลผลข้อมูลจากการสำรวขนาดใหญ่นั้น จะต้องมีการวางแผนและกำหนดงานเป็นขั้นตอนจึงจะทำให้สามารถดำเนินการประมวลผลไปโดยมีประสิทธิภาพและลดข้อผิดพลาดต่าง ๆ ต้องคำนึงถึงคุณภาพของข้อมูลอยู่เสมอ อาจต้องมีการตรวจสอบผลด้วย งานขั้นแรกของผู้รับผิดชอบการประมวลผลต้องทำก็คือ ตรวจสอบผลการรวบรวมข้อมูล เช่น แบบสอบถาม ให้แน่ใจว่าข้อมูลที่ได้รับมามีความถูกต้อง ไม่มีข้อบกพร่อง หากพบว่าข้อมูลผิดพลาดควรปรึกษาผู้ดำเนินการสำรวจควรมีแบบฟอร์มตารางที่ต้องการสร้างให้พร้อมครบทุกตาราง เพื่ออาจทำให้สามารถสร้างตารางที่คล้ายคลึงกันในเวลาเดียวกันได้ จะทำให้ประหยัดเวลาขึ้น งานขนาดใหญ่ที่ต้องใช้เครื่องจักรคำนวณในการประมวลผล ต้องมีการถ่ายข้อมูลเป็นรหัสแล้วนำไปเจาะบัตร บัตรที่ได้้อาจนำไปจำแนกประเภทโดยใช้เครื่องหรือใช้กับเครื่องจักรคำนวณ ควรพยายามหาวิธีตรวจสอบคุณภาพ เช่น ใช้การตรวจสอบเป็นจุด (Spot check) ควรมีการบันทึกเวลาทำงานเพื่อช่วยในการควบคุมการทำงาน คัดค่าใช้จ่ายและเป็นรากฐานในการสำรวจภายหลัง

2.2.4 การเตรียมรายงาน ในการเตรียมรายงานการสำรวจด้วยตัวอย่าง United Nations Sub-Commission on Statistical Sampling<sup>4</sup> ได้เสนอให้ประกอบด้วยสิ่งต่าง ๆ แยกสรุปได้ดังนี้

2.2.4.1 รายงานทั่วไป ควรประกอบไปด้วยคำอธิบายเกี่ยวกับการสำรวจ เพื่อประโยชน์แก่ผู้ใช้ข้อมูลที่รวบรวมได้ ควรสรุปผลของการสำรวจ การออกแบบตัวอย่าง การดำเนินการและการวิเคราะห์ คือ ควรประกอบด้วย

- (1) จุดประสงค์ของการสำรวจ และการใช้ข้อมูลที่จะสำรวจได้

(2) ขอบข่าย (Scope) และค้ำมรวม (Coverage) ให้ทราบขอบเขตของประชากรที่จะสำรวจและสิ่งที่ต้องการ

(3) วิธีการรวบรวมข้อมูล รวมปัญหาที่เกิดขึ้นในการรวบรวมข้อมูลด้วย

(4) ระยะเวลาที่เกี่ยวข้อง บอกระยะเวลาของการสำรวจ รวมทั้งระยะเวลาอ้างอิง (Reference period) และระยะเวลาที่รายงาน (Reporting period) ด้วย

(5) ชนิดของการสำรวจ เป็นการสำรวจครั้งเดียว หรือเป็นการสำรวจซ้ำ (Ad hoc or repetitive survey)

(6) วิธีการสุ่มตัวอย่าง และวิธีการประมาณค่า บอกให้ทราบเกี่ยวกับหน่วยตัวอย่าง กรอบตัวอย่าง และวิธีการสุ่ม บอกขนาดของตัวอย่าง วิธีการหาขนาดของตัวอย่าง บอกสูตรที่ใช้ในการประมาณค่า

(7) วิธีการประมวลผล ทำด้วยมือหรือเครื่องจักร ขั้นตอนการดำเนินการพร้อมทั้งปัญหาและวิธีแก้ไข

(8) เสนอผลเป็นรูปตาราง พร้อมกราฟ แผนภูมิภาพ และคำอธิบายประกอบ

(9) ความถูกต้องแม่นยำ ควรบอกได้ว่าการสำรวจนั้นให้ผลถูกต้องเพียงไร บอกให้ทราบถึงอัตราการไม่ตอบ การแก้ไข้ปัญหา บอกทั้งความคลาดเคลื่อนอันเนื่องจากการสุ่มตัวอย่าง และความคลาดเคลื่อนอันเนื่องจากสิ่งอื่น ๆ

(10) ค่าใช้จ่าย บอกให้ทราบว่าค่าใช้จ่ายทั้งหมดเป็นเท่าไร แยกรายละเอียดได้อย่างไรบ้าง เช่น แยกเป็นค่าเดินทาง ค่ารวบรวมข้อมูล ค่าลงรหัส ค่าเจาะบัตร ค่าเครื่องคำนวณ ฯลฯ รวมทั้งจำนวนคนและเวลาที่ใช้

(11) ความรับผิดชอบ บอกให้ทราบว่าหน่วยงานใดรับผิดชอบ และให้ความสนับสนุน

(12) บรรณานุกรม บอกให้ทราบว่ารายงานหรือเอกสารใดที่เกี่ยวข้องกับการสำรวจนั้นบ้าง

2.2.4.2 รายงานทางเทคนิค ควรจัดทำขึ้น หากการสำรวจนั้นมีความสำคัญมาก หรือใช้เทคนิควิธีใหม่ที่จะเป็นที่สนใจทั่วไป ซึ่งควรประกอบด้วย

(1) การออกแบบการสำรวจ ใช้วิธีการสุ่มตัวอย่างแบบใด อธิบายถึงกรอบตัวอย่าง หน่วยตัวอย่าง เศษส่วนการสุ่ม (Sampling fraction)<sup>6</sup> หลักเกณฑ์ในการแบ่งเป็นชั้นภูมิ วิธีการสุ่มในชั้นต่าง ๆ

(2) บุคลากรและเครื่องมือ จำนวนคนที่ใช้ในการวางแผน การรวบรวมข้อมูล การประมวลผล รวมทั้งประสบการณ์ของบุคคลเหล่านั้น วิธีการจัดหาบุคคล การอบรม การตรวจสอบ การควบคุมแนะนำ รวมทั้งเครื่องมือที่ใช้

(3) วิธีการวิเคราะห์ทางสถิติและวิธีคำนวณ บอกสูตรที่ใช้ วิธีการประมาณค่าพร้อมทั้งเหตุผลที่ใช้

(4) ความถูกต้องแม่นยำของผลการสำรวจ บอกความคลาดเคลื่อนมาตรฐาน (Standard error) การตรวจสอบคุณภาพ

(5) การวิเคราะห์ทางเทคนิค และการแปรความหมายผลของการสำรวจ อาจมีการทดสอบสมมติฐาน หรือการตอบคำถามบางประการ

(6) การเปรียบเทียบกับข้อมูลจากที่อื่น

(7) การวิเคราะห์ค่าใช้จ่าย บอกให้ทราบค่าใช้จ่ายของการสำรวจ แยกตามขั้นการดำเนินงาน เช่น การวางแผน การเดินทาง การแจกจ่าย การแนะนำ การประมวลผล การวิเคราะห์ และค่าใช้จ่ายอื่น ๆ

(8) ประสิทธิภาพ เปรียบเทียบประสิทธิภาพของวิธีการที่ใช้กับวิธีอื่น ๆ

(9) ข้อสังเกตอื่น ๆ โดยเฉพาะข้อสังเกตของนักสถิติ และนักวิชาการที่เกี่ยวข้อง

### 8. ระบบการเลือกตัวอย่าง

ปัจจุบันมีผู้ใช้คำว่าแบบสำรวจ (Survey design) แบบตัวอย่าง (Sample design) และแบบการสุ่มตัวอย่าง (Sampling plan or sampling procedure or selection process) สับสนกันอยู่มาก

คำว่า “แบบสำรวจ” หมายถึง การดำเนินการสำรวจรวมทั้งการวางแผนการดำเนินการ การเก็บรวบรวมข้อมูล การเลือกและอบรมพนักงานแจงนับ การประมวลผลข้อมูล และการเสนอรายงานข้อมูล มิใช่เฉพาะการเลือกตัวอย่างและการประมาณค่าพารามิเตอร์เท่านั้น

“แบบการสุ่มตัวอย่าง” หมายถึง กฎเกณฑ์ และการดำเนินการเลือกตัวอย่าง โดยอาศัยกรอบตัวอย่าง กล่าวคือ เป็นแบบการเลือกหน่วยตัวอย่างจากกรอบตัวอย่างที่เตรียมไว้ล่วงหน้า การเลือกตัวอย่างหลายวิธีหากปฏิบัติตามกฎเกณฑ์ หรือนิยามของวิธีการสุ่มตัวอย่างแบบนั้น จะทำได้โดยยาก จึงอาจมีวิธีการเลือกหน่วยตัวอย่างเพื่อให้ได้ตัวอย่างที่มีคุณสมบัติตามกฎเกณฑ์ของวิธีการนั้น ๆ

“แบบตัวอย่าง” หมายความว่าวิธีการเลือกตัวอย่าง รวมทั้งวิธีการประมาณค่าเพื่อหาค่าสถิติที่ใช้ประมาณค่าพารามิเตอร์

การพิจารณาแบบการสุ่มตัวอย่างในที่นี้ จะไม่กล่าวถึงวิธีการประมาณค่า หากแต่จะมุ่งพิจารณาแบบการสุ่มตัวอย่างโดยอาศัยหลักเกณฑ์บางประการ ดังนี้

3.1 แบบการสุ่มตัวอย่างแบ่งตามความน่าจะเป็น ในที่นี้จะพิจารณาเฉพาะวิธีการสุ่มตัวอย่างชนิดใช้ความน่าจะเป็นเท่านั้น จะไม่พิจารณาวิธีการสุ่ม

ตัวอย่างแบบใช้วิจารณ์แผนของผู้ชำนาญพิเศษ ไม่มีการใช้โควต้า ไม่มีการใช้ความลำเอียงใด ๆ หากแบ่งวิธีการสุ่มตัวอย่างโดยอาศัยความน่าจะเป็น จะแบ่งได้เป็น 2 พวก คือ

3.1.1 แบบการสุ่มตัวอย่างที่ทุกหน่วยตัวอย่างมีความน่าจะเป็นที่ถูกเลือกเท่ากัน (Equal probability sampling) เป็นวิธีที่หน่วยตัวอย่างทุกหน่วยในประชากร ซึ่งปรากฏอยู่ในกรอบตัวอย่างมีโอกาสถูกเลือกเท่า ๆ กัน ได้แก่การสุ่มตัวอย่างอย่างง่าย (Simple random sampling) และการสุ่มตัวอย่างแบบมีระบบ (Systematic sampling)

“วิธีการสุ่มตัวอย่างอย่างง่าย” เป็นวิธีการสุ่มตัวอย่างที่ตัวอย่างที่เป็นไปได้ทุกตัวอย่างมีโอกาสถูกเลือกมาใช้เท่า ๆ กัน เช่น ถ้า  $N$  เป็นขนาดของประชากร และ  $n$  เป็นขนาดของตัวอย่าง จำนวนตัวอย่างที่ประกอบไปด้วยหน่วยตัวอย่างที่ต่างกันหมดคือ  $\binom{N}{n}$  และแต่ละตัวอย่างมีโอกาสที่จะถูกเลือกมาใช้เท่ากับ  $1/\binom{N}{n}$  แต่ในทางปฏิบัติเราเลือกตัวอย่างแบบนี้ โดยเลือกหน่วยตัวอย่างมาทีละหน่วย ในแต่ละครั้งให้แต่ละหน่วยมีโอกาสถูกเลือกเท่า ๆ กัน ซึ่งอาจจะทำได้โดยการจับสลากหรือใช้ตารางเลขสุ่ม

“วิธีการเลือกตัวอย่างแบบมีระบบ” เป็นการเลือกตัวอย่างจากรายการหน่วยตัวอย่างที่เรียงกันอยู่ จากช่วงหนึ่ง ๆ เลือกหน่วยตัวอย่างมา 1 หน่วย เช่น เลือกมา 1 หน่วยจากทุก ๆ 20 หน่วยเป็นต้น ในแต่ละช่วงเราเลือกหน่วยตัวอย่างมา 1 หน่วย โดยใช้ความน่าจะเป็นที่เท่ากัน ในทางปฏิบัติเราเลือก “จุดเริ่มต้น” (Random start) เป็นตัวเลขจำนวนเต็มระหว่าง 1 กับ  $k$  เมื่อ  $k$  เป็นขนาดของช่วง สมมติว่าเป็น  $r$  แล้ว หน่วยตัวอย่างที่ถูกเลือกเป็นตัวอย่างได้แก่ หน่วยที่  $r, r + k, r + 2k, \dots, r + (n - 1)k$  เมื่อ  $n$  เป็นขนาดของตัวอย่าง ตามที่



กล่าวข้างบนนี้ใช้ได้เมื่อ  $n$  ทหาร  $N$  ได้ลงตัว และเท่ากับ  $k$  ในที่นี้  $N$  เป็นขนาดของประชากร ในกรณีที่  $n$  ทหาร  $N$  ไม่ลงตัวบางทีก็เลือก  $k$  เป็นจำนวนเต็มใกล้เคียงกับ  $\frac{N}{n}$  และดำเนินการเช่นที่กล่าว แต่ก็ยังมีปัญหาต่าง ๆ ตามมา จึงมีผู้คิดค้นแปลงวิธีการดังกล่าวต่อไปอีก เช่น “วิธีสุ่มแบบมีระบบค้ำแปลง” (Modified systematic sampling) ซึ่งเลือกจุดเริ่มต้น  $r$  ระหว่าง 1 กับ  $N$  และได้หน่วยตัวอย่างที่  $r, r \pm k, r \pm 2k, \dots$  ซึ่งอยู่ระหว่าง 1 กับ  $N$  เป็นตัวอย่าง “วิธีสุ่มแบบมีระบบวงกลม” (Circular systematic sampling) เลือกจุดเริ่มต้น  $r$  ระหว่าง 1 กับ  $N$  และได้หน่วยตัวอย่างที่  $r, r + k, r + 2k, \dots, r + (n-1)k$  (ใช้ modulus  $N$ ) เป็นตัวอย่าง

3.1.2 แบบการสุ่มตัวอย่างที่ใช้ความน่าจะเป็นไม่เท่ากัน (Unequal probability sampling) ในกรณีที่เราทราบธรรมชาติของหน่วยตัวอย่างในประชากร เราอาจใช้ข้อมูลที่มีอยู่ให้เกิดประโยชน์ในการเลือกตัวอย่าง และในการประมาณค่าได้ การทำเช่นนี้จะทำให้ความน่าจะเป็นที่หน่วยตัวอย่างจะถูกเลือกแต่ละครั้งไม่เท่ากัน ตัวอย่างประมาณค่าที่ใช้ก็มักจะมีประสิทธิภาพ (Efficiency) สูงกว่าตัวอย่างประมาณค่าในวิธีการอื่น วิธีการเลือกตัวอย่างประเภทที่สำคัญคือ การสุ่มตัวอย่างโดยใช้ความน่าจะเป็นได้สัดส่วนกับขนาด (Sampling with probability proportional to size)

“วิธีการสุ่มตัวอย่างโดยใช้ความน่าจะเป็นได้สัดส่วนกับขนาด” เป็นการเลือกตัวอย่างที่ให้ความสำคัญแก่ “ขนาด” ของหน่วยตัวอย่าง คำว่า “ขนาด” ในที่นี้เราหมายถึงความถี่ของตัวแปรที่เกี่ยวข้องกับตัวแปรที่ต้องการศึกษา ในการเลือกตัวอย่างโดยใช้ความน่าจะเป็นได้สัดส่วนกับขนาดจากกรอบตัวอย่างที่เป็นรายการอาจจะทำได้โดย “วิธีสะสมขนาด” (Cumulative total method)  $z_1$  ,

$z_1 + z_2 + \dots + z_1 + z_2 + \dots + z_N = Z$  ของหน่วยตัวอย่างที่ 1, 2, 3, ..., N แล้วเลือกตัวเลขสุ่มระหว่าง 1 กับ  $Z$  เพื่อเทียบกับหน่วยตัวอย่าง ทำซ้ำๆ จนได้ตัวอย่างตามต้องการ หรืออาจจะใช้ "วิธีของ Lahiri" เลือกตัวเลขสุ่ม 1 คู่ ตัวหนึ่งระหว่าง 1 กับ  $N$  (ขนาดของประชากร) และอีกตัวหนึ่งระหว่าง 1 กับค่าสูงสุดของขนาดของหน่วยตัวอย่าง ถ้าเลือกได้  $i$  และจำนวนที่น้อยกว่า  $z_i$  ถือว่าเลือกได้หน่วยที่  $i$  มิฉะนั้นจะต้องเลือกตัวเลขสุ่มซ้ำใหม่ ทำซ้ำๆ จนได้ตัวอย่างที่มีขนาดตามต้องการ

หากต้องการเลือกพื้นที่ จากแผนที่ โดยใช้ความน่าจะเป็นได้สัดส่วนกับขนาดของพื้นที่ เลือกตัวเลขสุ่ม 2 ตัว ตัวหนึ่งจะแสดงความกว้าง อีกตัวหนึ่งแสดงความยาว หากจุดที่มีโคออร์ดิเนตเท่าตัวเลขที่สุ่มได้ตกอยู่ในหน่วยพื้นที่ใด ก็ถือว่าเลือกได้หน่วยนั้น หากจุดนั้นตกอยู่นอกเขตที่สนใจก็เลือกซ้ำอีก จนกระทั่งได้ตัวอย่างที่ต้องการ

ในบางกรณีเราก็ไม่ทราบค่าของตัวแปรที่เกี่ยวข้องกับค่าของตัวแปรที่สนใจ แต่อาจจะประมาณค่าได้ ในกรณีเช่นนั้นเราใช้การสุ่มตัวอย่างที่ใช้ความน่าจะเป็นได้สัดส่วนกับค่าประมาณของขนาด

การสุ่มตัวอย่างแบบใช้ความน่าจะเป็นไม่เท่ากัน อาจจะเป็นไปในรูปแบบอื่นๆ อีก อาจกำหนดความน่าจะเป็นในลักษณะอื่นๆ เช่น ให้  $p_1, p_2, \dots, p_N$

$$\sum_{i=1}^N p_i = 1 \quad \text{เป็นความน่าจะเป็นในการสุ่มได้หน่วยตัวอย่างที่ } 1, 2, \dots, N$$

ตามลำดับ ซึ่งอาจจะเป็นเท่าไรก็ได้เพียงแต่ให้  $\sum_{i=1}^N p_i = 1$  และ  $0 < p_i < 1$  เท่านั้น

**3.2 แบบการสุ่มตัวอย่างแบ่งตามการแทนที่** ในการเลือกตัวอย่าง โดยทั่วไปเรามักจะเลือกทีละหน่วยจนครบจำนวนที่ต้องการ จึงมีปัญหว่าในการเลือกแต่ละครั้งหลังครั้งแรก จะมีการคืนหน่วยตัวอย่างที่เลือกได้แล้วก่อนการเลือกครั้งต่อไปหรือไม่ นั่นคือเราแบ่งแบบการสุ่มตัวอย่างตามเกณฑ์ได้ 2 แบบ คือ

**3.2.1 แบบการสุ่มตัวอย่างที่มีการคืนหน่วยตัวอย่างก่อนเลือกครั้งต่อไป (Sampling with replacement)** เป็นแบบที่จะพิจารณาหน่วยตัวอย่างที่เลือกได้ในแต่ละครั้ง แล้วคืนไว้ในกลุ่มเดิมก่อนเลือกครั้งต่อไป การทำเช่นนั้นทำให้สามารถเลือกตัวอย่างได้ง่าย ไม่ต้องคิดคำนวณยุ่งยาก เช่น ในการเลือกตัวอย่างแบบใช้ความน่าจะเป็นได้สัดส่วนกับขนาด หากเลือกทีละหน่วยตัวอย่างแล้วคืนไว้ในกลุ่มก่อนที่จะเลือกครั้งต่อไป ทำให้ไม่ต้องคิดค่าความน่าจะเป็นที่กำหนดให้แก่แต่ละหน่วยใหม่

**3.2.2 แบบการสุ่มตัวอย่างที่ไม่มีการคืนหน่วยตัวอย่างก่อนเลือกครั้งต่อไป (Sampling without replacement)** เป็นแบบที่เมื่อเลือกได้หน่วยตัวอย่างใดแล้ว ในการเลือกครั้งต่อไปจะไม่พิจารณาหน่วยตัวอย่างนั้นอีก การทำเช่นนี้อาจทำให้การเลือกตัวอย่างยุ่งยากขึ้น เช่น อาจต้องคำนวณค่าความน่าจะเป็นที่ใช้ในการเลือกใหม่อยู่เสมอๆ แต่มีข้อดีที่สำคัญอยู่ว่า ข้อมูลที่ได้จะใช้ประโยชน์ได้ดีกว่าแบบแรก กล่าวคือความคลาดเคลื่อนของตัวประมาณค่าของพารามิเตอร์เดียวกัน รูปเดียวกัน ที่คิดจากตัวอย่างที่เลือกโดยวิธีนี้ จะน้อยกว่าวิธีแรก<sup>๕</sup> ด้วยเหตุนี้ในทางปฏิบัติเราจึงมักจะใช้วิธี การสุ่มตัวอย่าง แบบนี้แทนที่จะใช้แบบแรก โดยเฉพาะวิธีการสุ่มอย่างง่ายมักเป็นแบบไม่แทนที่

**3.3 แบบการสุ่มตัวอย่างแบ่งตามการแบ่งประชากรเป็นกลุ่ม** เราอาจจะพิจารณาวិธีการเลือกตัวอย่างโดยพิจารณาว่าประชากรที่ต้องการศึกษา ประ-

กอบด้วยหน่วยตัวอย่างที่แบ่งเป็นพวกหรือไม่ ในกรณีที่ไม่มีการแบ่งประชากรเป็นส่วน ๆ (ที่จะเรียกว่าชั้นภูมิ) เราอาจจะเลือกตัวอย่างตามแบบใดที่กล่าวมาแล้ว เช่น วิธีสุ่มอย่างง่าย หรือมีระบบ หรือใช้ความน่าจะเป็น ได้สัดส่วนกับขนาด ในกรณีที่มีการแบ่งประชากรเป็นส่วน ๆ เราอาจจะพิจารณาวิธีการสุ่มแยกได้เป็น 2 พวก คือ

3.3.1 แบบการสุ่มตัวอย่างที่มี การ แบ่ง ประชากร เป็น ชั้น ภูมิ (Stratified sampling) ในกรณีนี้เราแบ่งประชากรเป็นส่วน ๆ เรียกว่า “ชั้นภูมิ” แล้วสุ่มตัวอย่างมาจากแต่ละชั้นภูมิทุกชั้นภูมิโดยวิธีใดวิธีหนึ่ง เช่น ใช้การสุ่มอย่างง่ายหรืออย่างมีระบบ หรือโดยใช้ความน่าจะเป็น ได้สัดส่วนกับขนาด ฯลฯ วิธีการสุ่มแบบนี้ใช้กันแพร่หลายมาก อาจจะเป็นที่นิยมมากที่สุดก็เป็นได้

3.3.2 แบบการสุ่มตัวอย่างโดยการแบ่งกลุ่ม (Cluster sampling) ได้แก่การสุ่มตัวอย่างชนิดที่ถือว่าหน่วยตัวอย่าง ประกอบไปด้วยหน่วยย่อยหลายหน่วย เมื่อสุ่มได้กลุ่มหน่วยตัวอย่างใดแล้วจะเก็บรวบรวมข้อมูลมาจากทุกหน่วยย่อย การเลือกกลุ่ม (Clusters) อาจจะทำโดยวิธีสุ่มอย่างง่ายหรือแบบมีระบบ หรือใช้ความน่าจะเป็นไม่เท่ากันก็ได้ และนอกจากนี้ยังอาจจะแยกกลุ่มออกเป็นส่วน ๆ ก่อนจะเลือกจากแต่ละส่วนหรือชั้นภูมิก็ได้

คำว่า “กลุ่ม” ในที่นี้อาจจะได้แก่ กลุ่มของหน่วยตัวอย่างที่เกิดขึ้นตามธรรมชาติ เช่น บุคคลในครัวเรือน หรือครัวเรือนในหมู่บ้าน สัตว์ในฝูง ฯลฯ หรืออาจจะเป็นกลุ่มที่สร้าง ขึ้นเองโดย รวมหน่วย ที่อยู่ใกล้เคียงกัน ที่อาจจะสำรวจร่วมกันได้ง่ายก็ได้ ในกรณีที่มีการแบ่งพื้นที่ของประชากรออกเป็น ส่วน ๆ และเลือกพื้นที่เป็นตัวอย่าง เราเรียกรูปแบบการสุ่มตัวอย่างชนิดนี้ว่า “การสุ่มพื้นที่” (Area sampling)

3.4 **แบบการสุ่มตัวอย่างแบ่งตามขั้นการสุ่ม** เราอาจจะพิจารณาวีธีการสุ่มตัวอย่างตามจำนวนขั้นการสุ่มก็ได้ หากมีการสุ่มตัวอย่างครั้งเดียวแล้วเก็บข้อมูลจากหน่วยตัวอย่าง เราเรียกการสุ่มนี้ว่า “การสุ่มขั้นเดียว” (Single-stage sampling) ในกรณีที่มีการสุ่มตัวอย่างมากกว่าหนึ่งขั้น มีการสุ่มย่อยจากตัวอย่างที่เลือกได้ในขั้นก่อน เราเรียกวีธีการสุ่มตัวอย่างนี้ว่า “วีธีการสุ่มหลายขั้น” (Multi-stage sampling) การสุ่มในแต่ละขั้นอาจจะทำได้โดยวิธีใดก็ได้ เช่น อาจจะใช้วิธีแบ่งเป็นชั้นภูมิ หรือแบบสุ่มอย่างง่าย ฯลฯ วีธีการสุ่มแบบหลายขั้นเป็นแบบที่นิยมใช้กันมาก โดยเฉพาะถ้ามีการสำรวจขนาดใหญ่

3.5 **แบบการสุ่มตัวอย่างชนิดอื่น** นอกจากที่กล่าวมาแล้วยังมีแบบการสุ่มตัวอย่างแบบอื่น ๆ อีก ดังจะได้กล่าวพอเป็นสังเขปต่อไป

3.5.1 **การสุ่มหลายตอน (Multiphase sampling)** หมายถึงการสุ่มตัวอย่างชนิดที่มีการหาข้อเท็จจริงจากหน่วยตัวอย่างที่สุ่มได้ และใช้ประโยชน์จากข้อเท็จจริงที่รวบรวมได้ให้เกิดประโยชน์ในการสุ่มขั้นต่อไป อาจจะใช้ประโยชน์ในการประมาณค่าพารามิเตอร์ด้วย

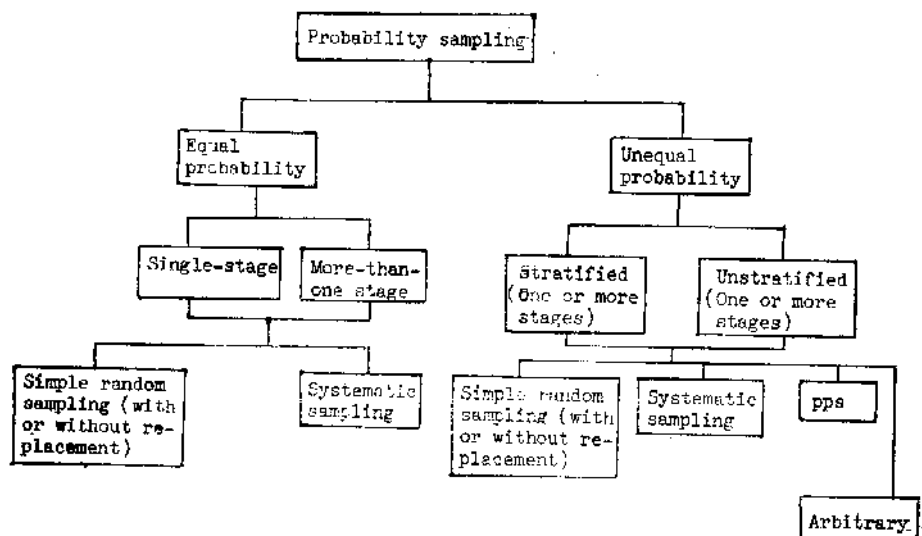
3.5.2 **การสุ่มหลายครั้ง (Sampling on successive occasions)** การสำรวจด้วยตัวอย่างที่กระทำซ้ำ โดยคลุ่มคุ่มรวมเดียวกันในเวลาต่างกัน เช่น การสำรวจประชากร การสำรวจแรงงาน ฯลฯ อาจจะใช้วิธีต่าง ๆ ได้หลายวิธี เช่น โดยการเลือกตัวอย่างใหม่ทุกครั้ง หรือโดยการแทนตัวอย่างบางส่วนคงไว้บางส่วน หรือโดยการสุ่มจากตัวอย่างที่สุ่มไว้ก่อน วีธีการสุ่มแบบนี้จะมีจุดประสงค์ที่จะวัดการเปลี่ยนแปลง หรือเพื่อใช้ข้อมูลผสมกัน

3.5.3 **การสุ่มกลับ (Inverse sampling)** ได้แก่การสุ่มชนิดที่เลือกหน่วยตัวอย่างไปเรื่อย ๆ จนกระทั่งได้สิ่งที่ต้องการครบ เช่น ได้หน่วยตัวอย่างที่ต่างกัน มีจำนวนเท่าที่ต้องการ

3.5.4 การสุ่มแบบควบคุม (Controlled sampling) ได้แก่การสุ่มที่ความน่าจะเป็นที่กำหนดให้แก่หน่วยตัวอย่างบางหน่วยใหญ่กว่าหน่วยอื่น ๆ มาก จึงมีโอกาสที่จะถูกเลือกมากกว่า มักใช้กับกรณีที่มีหน่วยตัวอย่างชั้นแรกขนาดใหญ่ และมีไม่มากนัก

3.5.5 การสุ่มแบบลูกหิมะ (Snowball sampling) หมายถึงวิธีการสุ่มตัวอย่างชนิดที่ใช้หน่วยตัวอย่างที่สุ่มได้ เป็นผู้ส่งชววนำไปสู่หน่วยตัวอย่างอื่น เช่น ตามคนที่สุ่มได้ว่าใครทราบเรื่องหนึ่งจากใครแล้ว ผู้รวบรวมข้อมูลก็ไปตามผู้ที่ถูกเอ่ยชื่อ เช่นนี้ต่อไปเรื่อย ๆ เป็นต้น

เมื่อพิจารณาแบบการสุ่มตัวอย่างดังกล่าวมาแล้ว เราอาจจะแสดงระบบการสุ่มตัวอย่างด้วยแผนภูมิต่อไปนี้



#### 4. ตัวอย่างการสำรวจที่ใช้การสุ่มตัวอย่าง

ได้มีผู้ดำเนินการสำรวจด้วยตัวอย่าง ทั้งในประเทศไทยและต่างประเทศกัน อย่างแพร่หลาย แต่ที่น่าสนใจและใคร่สรุปเสนอ ณ ที่นี้พอเป็นสังเขป คือ :-

4.1 การสำรวจค่าใช้จ่ายของครัวเรือน<sup>7</sup> (Household expenditure survey) ประเทศออสเตรเลียได้ทำการสำรวจค่าใช้จ่ายของครัวเรือนในปี ค.ศ. 1974 โดยใช้การสุ่มแบบ 3 ชั้น ในชั้นแรกเลือกหน่วยที่สร้างขึ้นสำหรับสามะโน (ประมาณ 250 ครัวเรือน) โดยใช้การสุ่มแบบมีระบบ ในชั้นที่สองภายในแต่ละหน่วยเลือกบล็อกมา 2 บล็อก โดยใช้ความน่าจะเป็นได้สัดส่วนกับขนาด ในชั้นที่สามเลือกครัวเรือนภายในบล็อกที่สุ่มได้ โดยใช้วิธีแบ่งเป็นชั้นภูมิ และใช้วิธีสุ่มอย่างง่ายภายในแต่ละชั้นภูมิ

4.2 การวิเคราะห์ทางเศรษฐศาสตร์ สังคม การเมือง และการบริหาร ที่มีผลต่อการพัฒนาชนบท<sup>8</sup> (Analysis of the socio-economic, political and administrative factors affecting the uneven progress of rural areas) เป็นการสำรวจโดยแผนกประชากรศาสตร์แห่ง Gokhale, Institute of Politics and Economics, Poona, India เพื่อพิจารณาการวางแผนครอบครัว โดยพยายามหาสิ่งที่ทำให้การวางแผนครอบครัวไม่ได้ผลเท่าที่ควร แบบการสุ่มที่ใช้คือ แบบแบ่งประชากรเป็นชั้นภูมิ โดยแบ่งสถานื่อนามัยมาจากแต่ละเขต เก็บข้อมูลเกี่ยวกับการวางแผนครอบครัวจากสถานื่อนามัยที่เลือกได้

4.3 การสำรวจค่าใช้จ่ายในการศึกษา ของเด็กระดับมัธยมศึกษาในเขตชุมชน<sup>9</sup> (A survey of the per-student family cost of education in urban secondary schools) เป็นการสำรวจที่ดำเนินการโดย Child and Youth Research Center, Philippines เพื่อหาค่าใช้จ่ายที่ครอบครัวจะต้องจ่ายในการศึกษาชั้นมัธยม

ศึกษาของเด็กคนหนึ่งๆ การแจกแจงของค่าใช้จ่าย และพิจารณาความแตกต่างของค่าใช้จ่าย แยกตามขนาดของครัวเรือนที่มีเด็กไปโรงเรียน ภาค อาริฟ และระดับการศึกษาของบิดามารดา วิธีการสุ่มตัวอย่างที่ใช้เป็นแบบสองชั้น ในชั้นที่ 1 ใช้วิธีการสุ่มแบบแบ่งประชากรเป็นชั้นภูมิ ซึ่งได้แก่ภาคปกครองทั้ง 11 ของประเทศ เลือกโรงเรียนชั้นมัธยมจากแต่ละภาค โดยวิธีสุ่มอย่างง่าย ในชั้นที่ 2 เลือกนักเรียนจากแต่ละโรงเรียนที่เลือกได้ในชั้นแรก โดยใช้วิธีสุ่มอย่างง่าย

4.4 การสำรวจวิทยุและโทรทัศน์ที่วราชอาณาจักร พ.ศ. 2517<sup>10</sup> เป็นการสำรวจผู้รับฟังวิทยุ และโทรทัศน์ เพื่อหาข้อมูลเกี่ยวกับจำนวนเครื่องรับและลักษณะนิสัยในการฟังวิทยุและโทรทัศน์ ไปใช้ในการวางแผนจัดรายการให้ตรงกับความต้องการของประชาชน การสุ่มตัวอย่างที่ใช้แยกเป็น 2 พวก ข้อมูลเกี่ยวกับเครื่องรับ ใช้การสุ่มแบบ 2 ชั้น ชั้นแรก เลือกจังหวัดจากแต่ละภาคปกครอง โดยใช้ความน่าจะเป็นได้สัดส่วนกับจำนวนครัวเรือนในจังหวัด ในชั้นที่ 2 แบ่งจังหวัดที่เลือกได้เป็น 3 ชั้นภูมิ คือในเขตเทศบาล เขตสุขาภิบาล และนอกเขตสำหรับในเขตเทศบาล เลือกชุมชนอาคารโดยใช้ความน่าจะเป็นได้สัดส่วนกับจำนวนครัวเรือนในชุมชนอาคาร ส่วนในเขตสุขาภิบาล สุ่มเขตสุขาภิบาลมา 1 สุขาภิบาลจากกลุ่มสุขาภิบาล 4 กลุ่ม โดยใช้ความน่าจะเป็นได้สัดส่วนกับจำนวนครัวเรือนในสุขาภิบาล แล้วเลือกหมู่บ้านในสุขาภิบาลอีกทีหนึ่ง สำหรับบริเวณนอกเขตเทศบาลและสุขาภิบาลเลือกหมู่บ้านมาจำนวนหนึ่ง โดยใช้ความน่าจะเป็นได้สัดส่วนกับจำนวนครัวเรือนในหมู่บ้าน

ในส่วนที่เกี่ยวข้องกับธรรมเนียมของการรับฟังหรือชมรายการ เลือกครัวเรือนมาจากตัวอย่างประเภทแรก โดยใช้การสุ่มแบบมีระบบ



### เชิงอรรถ

- 1 ผู้สนใจประวัติย่อของทฤษฎีการสุ่มตัวอย่าง อาจจะได้จากเอกสารหลายแห่ง เช่นใน Suwattec, Prachoom. Some Estimator, Variances and Variance Estimators for Point-Cluster Sampling of Digraphs, Ph.D, Thesis, Department of Statistics, North Carolina State University at Ralaigh, N.C. 1970.
- 2 Neyman, J. On two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. J. Roy. Stat. Soc. 97 : 558 - 606, 1934.
- 3 Hansen, M.H. and W.N. Hurwity 1943. On the theory of sampling from finite populations. Ann. Math. Stat. 14 : 333 - 862.
- 4 United Nations. Recommendations for the Preparation of Sample Survey Reports. Sale No: 64, XVII. 7, New York, 1964.
- 5 Sampling fraction ไท้แก่ อัตราส่วนของขนาดของตัวอย่าง ต่อ ขนาดของประชากร.
- 6 ผู้สนใจต้องการศึกษาเพิ่มเติมโปรดดู เช่น Basu, D. On sampling with and without replacement. Sankhya. 20 : 287 - 294, 1958 หรือ Raj, D., and S.H. Khamis. Some remarks on sampling with replacement. Annals of Mathematical Statistics 29 : 550 - 557, 1958 เป็นต้น.

- 7 UN, ECAFE, Sample Surveys in the ESCAP Region, twelfth report  
(Survey no. 1.3) 1975.
- 8 Ibid. (Survey no. 7.21).
- 9 Ibid. (Survey no. 15.7).
- 10 คู่มือรายงานเบื้องต้นการสำรวจวิทยุ และโทรทัศน์ ทวีธาพร อานาจจักร พ.ศ.  
2517 สำนักงานสถิติแห่งชาติ สำนักงานนายกรัฐมนตรี.