IN 542003 Introduction to Statistic for Economists

บทที่ 2 การกำหนดประชากรและกลุ่มตัวอย่าง
(Population and Sample Determination)

สอนโดย
อ.ปรัชญา หมื่นอภัย

วัตถุประสงค์การเรียนรู้

1. **อธิบาย**ความหมายของประชากรและกลุ่มตัวอย่างได้
2. **จำแนก**ประเภทของประชากรและกลุ่มตัวอย่างได้
3. **อธิบาย**เหตุผลของการใช้กลุ่มตัวอย่างแทนการศึกษาประชากรทั้งหมดได้
4. **อธิบาย**แนวคิดของพารามิเตอร์และค่าสถิติ
5. **เลือกวิธี**การสุ่มตัวอย่างที่เหมาะสมกับปัญหาทางเศรษฐศาสตร์ได้
6. **อธิบาย**ความคลาดเคลื่อนจากการสุ่มตัวอย่างและอคติในการเลือกตัวอย่างได้

ประชากรและกลุ่มตัวอย่าง

ประชากร

ประชากร (Population) หมายถึง กลุ่มของหน่วย **"ทุกหน่วย"** ที่ผู้วิจัยสนใจศึกษา ซึ่งมีลักษณะหรือคุณสมบัติบางอย่างร่วมกัน

ประเภทของประชากร

1. **ประชากรที่มีจำนวนจำกัด (Finite population)** คือ ประชากรที่**สามารถนับ**ได้เป็นตัวเลขอย่างชัดเจน เช่น ครัวเรือนในตำบลแห่งหนึ่ง จำนวน 3,480 ครัวเรือน
2. **ประชากรที่มีจำนวนไม่จำกัด (Infinite population)** คือ ประชากรที่**ไม่สามารถนับ**ได้เป็นตัวเลขอย่างชัดเจน เช่น ผู้โดยสารที่ใช้บริการรถโดยสารสาธารณะในแต่ละวัน ผู้บริโภคที่ซื้อสินค้าออนไลน์ในประเทศไทย

การเก็บข้อมูลจากประชากรทุกหน่วย เรียกว่า **สำมะโนประชากร (Census)

กลุ่มตัวอย่าง

กลุ่มตัวอย่าง (Sample) หมายถึง **"ส่วนหนึ่ง"** ของประชากรที่ถูกเลือกขึ้นมาเพื่อใช้ในศึกษา

ในทางปฏิบัติ หาก**ประชากรมีขนาดใหญ่**มากๆ จะเก็บเพียง **"กลุ่มตัวอย่าง"** เนื่องจาก

1) ใช้เวลานานในการเก็บรวบรวมข้อมูลจึงจะครบตามจำนวนประชากรทั้งหมด
2) ใช้ทรัพยากร เช่น ทรัพยากรเงิน ทรัพยากรบุคลากร จำนวนมากเก็บรวบรวมข้อมูล
3) ข้อมูลอาจมีความคลาดเคลื่อนเกิดขึ้น เนื่องจากต้องเก็บข้อมูลจำนวนมาก
4) การดำเนินการวิจัยอาจไม่ชัดเจนเนื่องจากนักวิจัยเวลาจำกัด

หัวข้อที่สนใจศึกษา	ประชากร	กลุ่มตัวอย่าง	ตัวอย่างข้อมูลดิบ
1. เวลาที่นักศึกษาใช้สื่อสังคมออนไลน์ต่อวัน	นักศึกษาระดับปริญญาตรีทั้งหมดในมหาวิทยาลัยแห่งหนึ่ง	นักศึกษาระดับปริญญาตรี จำนวน 50 คน	2, 3, 5 (ชั่วโมง/วัน)
2. กรุ๊ปเลือดของนักเรียนชั้นมัธยมศึกษาปีที่ 6	นักเรียนชั้น ม.6 ทั้งหมดในโรงเรียนแห่งหนึ่ง	นักเรียนชั้น ม.6 จำนวน 30 คน	A, O, B, AB, O
3. ค่าใช้จ่ายในการเดินทางไปทำงานต่อวันของประชากรวัยทำงาน	ประชากรวัยทำงานในจังหวัดขอนแก่น	ประชากรวัยทำงานจำนวน 50 คน	30, 50, 40, 60, 35, 80, 45, 50 (บาท/วัน)
4. ระดับความเครียดของนักศึกษาวิชาสถิติเศรษฐศาสตร์	นักศึกษาที่ลงทะเบียนเรียนวิชาสถิติเศรษฐศาสตร์ทั้งหมด	นักศึกษา 35 คน	มากที่สุด, มาก, ปานกลาง, น้อย, น้อยที่สุด

การกำหนดขนาดกลุ่มตัวอย่าง
(Sample Size Determination)

เราควรเก็บตัวอย่างจำนวนเท่าไรดี ?

- **การกำหนดขนาดกลุ่มตัวอย่าง (Sample Size Determination)** คือ กระบวนการกำหนดจำนวนหน่วยตัวอย่างที่เหมาะสมสำหรับการศึกษา เพื่อให้ข้อมูลที่เก็บรวบรวมสามารถเป็นตัวแทนของประชากรได้อย่างน่าเชื่อถือ และมีความคลาดเคลื่อนอยู่ในระดับที่ยอมรับได้

ปัจจัยการกำหนดขนาดกลุ่มตัวอย่าง

1. ขนาดของประชากร (Population Size; N)
2. ระดับความเชื่อมั่น (Confidence Level)
3. ค่าความคลาดเคลื่อนที่ยอมรับได้ (Margin of Error; e)

วิธีการกำหนดขนาดกลุ่มตัวอย่าง

การกำหนดขนาดกลุ่มตัวอย่างสามารถทำได้หลายวิธี เช่น

1. การใช้สูตรคำนวณทางสถิติ
2. การใช้ตารางสำเร็จรูป เช่น ตารางของ Krejcie และ Morgan
3. การใช้โปรแกรมคอมพิวเตอร์หรือซอฟต์แวร์ทางสถิติ
4. การอ้างอิงจากงานวิจัยที่มีลักษณะใกล้เคียงกัน

การกำหนดขนาดกลุ่มตัวอย่างเมื่อ ทราบจำนวนประชากร

Yamane (1973) เสนอสูตรคำนวณจำนวนตัวอย่าง เมื่อทราบจำนวนประชากรแน่นอน

$$n = \frac{N}{1 + Ne^2}$$

เมื่อ e แทน ค่าความคลาดเคลื่อนที่ยอมรับได้ เช่น 0.05

ตัวอย่างการคำนวณ

หากประชากรมีจำนวน 2,000 คน และกำหนดค่าความคลาดเคลื่อนที่ยอมรับได้เท่ากับ 0.05

$$n = \frac{N}{1+Ne^2} = \frac{2,000}{1+2,000(0.05)^2} =$$

ดังนั้น ควรใช้กลุ่มตัวอย่างอย่างน้อย **..... คน**

การกำหนดขนาดกลุ่มตัวอย่างเมื่อ ไม่ทราบจำนวนประชากร

Cochran (1977) เสนอสูตรคำนวณจำนวนตัวอย่าง เมื่อไม่ทราบจำนวนประชากรแน่นอน หรือประชากรมีขนาดใหญ่มาก (Infinite Population)

$$n = \frac{Z^2 pq}{e^2}$$

เมื่อ e แทน ค่าความคลาดเคลื่อนที่ยอมรับได้ เช่น 0.05

p แทน สัดส่วนของประชากรที่มีลักษณะตรงกับที่สนใจ

q = 1 – p

** โดยทั่วไป ใช้ p = 0.5

Z แทน ระดับความเชื่อมั่น

ระดับความเชื่อมั่น	ค่า Z
90%	1.645
95%	1.96
99%	2.576

ตัวอย่างการคำนวณ

ผู้วิจัยต้องการศึกษาพฤติกรรมการใช้บริการแอปพลิเคชันด้านการเงินของประชาชน แต่ไม่ทราบจำนวนประชากรทั้งหมด กำหนดให้ p = 0.5 , ความคลาดเคลื่อน 5% และระดับความเชื่อมั่น 95%

$$n = \frac{Z^2 pq}{e^2} = \frac{(1.96)^2(0.5)(0.5)}{(0.05)^2} =$$

ดังนั้น ควรใช้กลุ่มตัวอย่างอย่างน้อย …… **คน**

ตารางของ Krejcie และ Morgan

ขนาดของประชากร	ขนาดของกลุ่มตัวอย่าง	ขนาดของประชากร	ขนาดของกลุ่มตัวอย่าง	ขนาดของประชากร	ขนาดของกลุ่มตัวอย่าง
10	10	220	140	1,200	291
15	14	230	144	1,300	297
20	19	240	148	1,400	302
25	24	250	152	1,500	306
30	28	260	155	1,600	310
35	32	270	159	1,700	313
40	36	280	162	1,800	317
45	40	290	165	1,900	320
50	44	300	169	2,000	322

ที่มา: https://krootor.com/2062

กรอบการสุ่มตัวอย่าง (Sampling Frame)

กรอบการสุ่มตัวอย่าง (Sampling Frame)

- **กรอบการสุ่มตัวอย่าง (Sampling Frame)** คือ รายชื่อหรือรายการของหน่วยทั้งหมดในประชากรเป้าหมายที่สามารถใช้เป็นพื้นฐานในการคัดเลือกกลุ่มตัวอย่าง

- **ตัวอย่าง:** มหาวิทยาลัยมีรายชื่อนักศึกษาทั้งหมด 12,350 คน พร้อมรหัสนักศึกษาและข้อมูลการติดต่อ → มีกรอบการสุ่มตัวอย่าง

- **ตัวอย่าง:** ทราบว่าประชาชนในจังหวัดหนึ่งมีจำนวน 1,500,000 คน แต่ไม่มีรายชื่อประชาชนแต่ละคน ✗ มีกรอบการสุ่มตัวอย่าง

ประเภทของการสุ่มตัวอย่าง

ประเภทของการสุ่มตัวอย่าง

การสุ่มแบบอาศัยความน่าจะเป็น (Probability Sampling)

- การสุ่มอย่างง่าย (Simple Random Sampling)
- การสุ่มแบบเป็นระบบ (Systematic Sampling)
- การสุ่มแบบแบ่งชั้น (Stratified Sampling)
- การสุ่มแบบกลุ่ม (Cluster Sampling)

การสุ่มแบบไม่อาศัยความน่าจะเป็น (Non-probability Sampling)

- การเลือกตามความสะดวก (Convenience Sampling)
- การเลือกแบบเจาะจง (Purposive Sampling)
- การเลือกแบบโควตา (Quota Sampling)
- การเลือกแบบลูกโซ่ (Snowball Sampling)

การสุ่มแบบอาศัยความน่าจะเป็น
(Probability Sampling)

การสุ่มแบบอาศัยความน่าจะเป็น (Probability Sampling)

- วิธีการเลือกกลุ่มตัวอย่างที่สมาชิกทุกหน่วยของประชากร**มีโอกาสถูกเลือก**เป็นกลุ่มตัวอย่าง
- วิธีการสุ่มประเภทนี้ช่วย**ลดอคติ**ในการคัดเลือก (Selection Bias)

1. การสุ่มอย่างง่าย (Simple Random Sampling)

การสุ่มอย่างง่าย คือ การเลือกกลุ่มตัวอย่างโดยให้ทุกหน่วยของประชากรมีโอกาสได้รับการคัดเลือกเท่ากัน และเป็นอิสระจากกัน

ตัวอย่าง: ผู้วิจัยต้องการศึกษาพฤติกรรมการออมของนักศึกษาเศรษฐศาสตร์ จึงหมายเลขประจำตัวและใช้โปรแกรมคอมพิวเตอร์สุ่มตัวเลข

ข้อดี

- ยุติธรรมและลดอคติในการคัดเลือก
- เหมาะสำหรับประชากรที่มีลักษณะคล้ายกัน
- หลักการและขั้นตอนการสุ่มเข้าใจง่าย

ข้อเสีย

- ต้องมีรายชื่อประชากรครบถ้วน
- ไม่เหมาะกับประชากรที่มีขนาดใหญ่มาก
- ไม่เหมาะกับประชากรที่กระจายอยู่หลายพื้นที่

2. การสุ่มแบบเป็นระบบ (Systematic Sampling)

- การสุ่มแบบเป็นระบบ คือ การเลือกกลุ่มตัวอย่างโดย**กำหนดช่วงการสุ่ม (Sampling Interval)** จากนั้นสุ่มเลือกหน่วยแรก แล้วเลือกหน่วยถัดไปตามช่วงที่กำหนดอย่างสม่ำเสมอ

$$k = \frac{N}{n}$$

ข้อดี

- ดำเนินการได้ง่ายและรวดเร็ว
- เหมาะสำหรับประชากรที่มีรายชื่อเรียงลำดับ
- ประหยัดเวลาและค่าใช้จ่าย

ข้อเสีย

- หากลำดับของประชากรมีรูปแบบซ้ำ (Pattern) อาจทำให้ผลการสุ่มเกิดความเอนเอียงได้

2. การสุ่มแบบเป็นระบบ (Systematic Sampling)

การสุ่มแบบเป็นระบบ คือ การเลือกกลุ่มตัวอย่างโดย**กำหนดช่วงการสุ่ม (Sampling Interval)** จากนั้นสุ่มเลือกหน่วยแรก แล้วเลือกหน่วยถัดไปตามช่วงที่กำหนดอย่างสม่ำเสมอ

$$k = \frac{N}{n}$$

ตัวอย่าง: ต้องการกลุ่มตัวอย่าง 100 คนจากประชากร 1,000 คน ดังนั้น คำนวณ k = 10

หากเริ่มต้นเก็บข้อมูลตัวอย่างเลขที่ 7 ลำดับถัดไปจะเป็น 17, 27, 37, ... จนครบ 100 คน

ข้อดี

- ดำเนินการได้ง่ายและรวดเร็ว
- เหมาะสำหรับประชากรที่มีรายชื่อเรียงลำดับ
- ประหยัดเวลาและค่าใช้จ่าย

ข้อเสีย

- หากลำดับของประชากรมีรูปแบบซ้ำ (Pattern) อาจทำให้ผลการสุ่มเกิดความผิดพลาดได้

3. การสุ่มแบบแบ่งชั้น (Stratified Sampling)

การสุ่มแบบแบ่งชั้น คือ การแบ่งประชากรออกเป็นกลุ่มย่อย (Strata) ที่มีลักษณะคล้ายคลึงกันภายในกลุ่ม เช่น เพศ อายุ คณะ หรือระดับรายได้ จากนั้นสุ่มตัวอย่างจากแต่ละชั้นตามสัดส่วนหรือจำนวนที่กำหนด

ตัวอย่าง: มหาวิทยาลัยมีนักศึกษาชาย 40% และนักศึกษาหญิง 60%

หากต้องการกลุ่มตัวอย่าง 500 คน จะสุ่มนักศึกษาชาย 200 คน และนักศึกษาหญิง 300 คน

ข้อดี

- เหมาะสำหรับประชากรที่มีความแตกต่างกันอย่างชัดเจน

ข้อเสีย

- ต้องทราบข้อมูลของประชากรเพื่อใช้แบ่งชั้น

4. การสุ่มแบบกลุ่ม (Cluster Sampling)

การสุ่มแบบกลุ่ม คือ การแบ่งประชากรออกเป็นกลุ่มย่อย (Strata) ที่มีลักษณะคล้ายคลึงกันภายในกลุ่ม เช่น เพศ อายุ คณะ หรือระดับรายได้ จากนั้นสุ่มตัวอย่างจากแต่ละชั้นตามสัดส่วนหรือจำนวนที่กำหนด

ตัวอย่าง: หากต้องการศึกษาพฤติกรรมการใช้จ่ายของนักเรียนระดับมัธยมศึกษาในจังหวัดขอนแก่นผู้วิจัยสุ่มเลือกโรงเรียนจำนวน 10 แห่งจากทั้งหมด 80 แห่ง แล้วเก็บข้อมูลจากนักเรียน

ข้อดี

- เหมาะกับประชากรที่กระจายอยู่หลายพื้นที่
- ลดค่าใช้จ่ายและเวลาในการเก็บข้อมูล

ข้อเสีย

- ความแม่นยำต่ำลง หากสมาชิกภายในแต่ละกลุ่มมีลักษณะคล้ายกันมาก หรือกระจายกลุ่มตัวอย่างไม่ดี

การสุ่มแบบไม่อาศัยความน่าจะเป็น
(Non-probability Sampling)

1. การเลือกตามความสะดวก (Convenience Sampling)

- การเลือกตามความสะดวก คือ การคัดเลือกกลุ่มตัวอย่างจากบุคคลที่สามารถเข้าถึงได้ง่าย สะดวก และยินดีให้ข้อมูล โดยไม่ใช้กระบวนการสุ่ม

- **ตัวอย่าง:** ผู้วิจัยต้องการศึกษาพฤติกรรมการซื้อเครื่องดื่มของนักศึกษา จึงเลือกสอบถามนักศึกษาที่กำลังเดินผ่านบริเวณโรงอาหารในช่วงเวลาพักกลางวัน

ข้อดี

- ดำเนินการได้ง่าย รวดเร็ว
- ประหยัดค่าใช้จ่าย
- ใช้เวลาน้อยในการเก็บข้อมูล

ข้อเสีย

- กลุ่มตัวอย่างอาจไม่เป็นตัวแทนของประชากร
- มีโอกาสเกิด Bias Selection สูง

2. การเลือกแบบเจาะจง (Purposive Sampling)

- **การเลือกแบบเจาะจง** คือ การคัดเลือกกลุ่มตัวอย่างโดยอาศัยดุลยพินิจของผู้วิจัย เพื่อเลือกผู้ที่มีคุณสมบัติเฉพาะหรือมีความรู้และประสบการณ์ตรงกับประเด็นที่ศึกษา เหมาะกับการวิจัยเชิงคุณภาพ
- **ตัวอย่าง:** ผู้วิจัยศึกษาปัจจัยความสำเร็จของผู้ประกอบการวิสาหกิจชุมชน จึงเลือกสัมภาษณ์เฉพาะผู้ประกอบการที่ดำเนินกิจการต่อเนื่องมาแล้วไม่น้อยกว่า 5 ปี

ข้อดี

- เหมาะสำหรับการวิจัยที่ต้องการข้อมูลเชิงลึกจากผู้เชี่ยวชาญ
- ได้ข้อมูลที่มีคุณภาพและตรงกับวัตถุประสงค์

ข้อเสีย

- อาจเกิดความเอนเอียงจากดุลยพินิจของผู้วิจัย
- สามารถอ้างอิงผลไปยังประชากรทั้งหมดได้

3. การเลือกแบบโควตา (Quota Sampling)

- **การเลือกแบบโควตา** คือ การกำหนดสัดส่วนกลุ่มตัวอย่างของแต่ละกลุ่มแยกย่อยตามลักษณะที่ต้องการศึกษา เช่น เพศ หรืออาชีพ จากนั้นผู้วิจัยเลือกผู้ให้ข้อมูลตามจำนวนที่กำหนด โดยไม่ใช้การสุ่ม
- **ตัวอย่าง:**การสำรวจความคิดเห็นของประชาชนจำนวน 400 คน โดยกำหนดให้มีเพศชาย 200 คน และเพศหญิง 200 คน จากนั้นผู้วิจัยเลือกผู้ตอบแบบสอบถามให้ครบตามจำนวนที่กำหนด

ข้อดี

- ควบคุมสัดส่วนของกลุ่มตัวอย่างได้

ข้อเสีย

- อาจเกิด Bias Selection
- ไม่สามารถคำนวณความคลาดเคลื่อนจากการสุ่มได้

4. การเลือกแบบลูกโซ่ (Snowball Sampling)

- **การเลือกแบบลูกโซ่** คือ การคัดเลือกกลุ่มตัวอย่างโดยเริ่มจากผู้ให้ข้อมูลกลุ่มแรก แล้วให้ผู้ให้ข้อมูลแนะนำบุคคลอื่นที่มีคุณสมบัติตรงตามเกณฑ์การวิจัยต่อไปเป็นลำดับ จนได้ครบจำนวนกลุ่มตัวอย่างตามต้องการ
- **ตัวอย่าง:**ผู้วิจัยศึกษาพฤติกรรมการลงทุนในสินทรัพย์ดิจิทัลของนักลงทุนรายย่อย โดยเริ่มสัมภาษณ์นักลงทุนจำนวนหนึ่ง และขอให้ผู้ให้ข้อมูลแนะนำผู้ลงทุนรายอื่นที่มีคุณสมบัติตรงตามเกณฑ์ของการวิจัย

ข้อดี

- ควบคุมสัดส่วนของกลุ่มตัวอย่างได้

ข้อเสีย

- อาจเกิด Bias Selection
- ไม่สามารถคำนวณความคลาดเคลื่อนจากการสุ่มได้