

# บทที่ 1

## บทนำ

### 1.1 ที่มาและความสำคัญของปัญหา

กระบวนการทางสถิติ เป็นกระบวนการของการนำข้อมูลมาผ่านวิธีการที่เหมาะสมเพื่อหาข้อสรุปตามวัตถุประสงค์ที่กำหนด สำหรับใช้ประโยชน์ในรูปแบบต่าง ๆ เช่น การทดสอบสมมติฐาน การตัดสินใจ เป็นต้น ซึ่งกระบวนการทางสถิติสามารถจำแนกตามประโยชน์การใช้งานได้เป็นสองกลุ่มคือ กระบวนการทางสถิติที่ใช้สำหรับหาข้อสรุปเกี่ยวกับข้อมูลที่มีอยู่ (Data Summarization) และกระบวนการทางสถิติที่ใช้ข้อมูลตัวอย่างอธิบายกลุ่มประชากร (Statistical Inference) โดยที่การนำข้อมูลตัวอย่างไปสรุปเกี่ยวกับประชากรนั้นต้องใช้ทฤษฎีสถิติที่เกี่ยวข้องกับการอธิบายลักษณะประชากร ได้แก่ ทฤษฎีการสำรวจตัวอย่าง (Theory of Sample Survey) และทฤษฎีการอนุมานเชิงสถิติ (Theory of Inference) สำหรับอธิบายกลุ่มประชากรที่สนใจ ทฤษฎีทั้งสองนี้มีความแตกต่างกันอย่างสิ้นเชิงทั้งในเรื่องของลักษณะและรูปแบบของประชากร คือ ทฤษฎีการอนุมานเชิงสถิติเป็นการพิจารณาประชากรอนันต์ (Infinite Population) และหน่วยแต่ละหน่วยของตัวอย่างถูกพิจารณาในลักษณะของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นกำกับ เนื่องจากประชากรที่สนใจศึกษามีขนาดใหญ่มาก และการแจกแจงความน่าจะเป็นมีบทบาทอย่างมากในการเลือกวิธีการอนุมาน เพื่อหาคุณสมบัติของการอนุมาน ดังนั้นการอนุมานเชิงสถิติจึงเป็นการอธิบายองค์รวมทั้งหมดของประชากร ในขณะที่ทฤษฎีการสำรวจตัวอย่างจะพิจารณาประชากรอันตะ (Finite Population) และโดยทั่วไปไม่สนใจการแจกแจงความน่าจะเป็นของตัวแปรสุ่ม เนื่องจากประชากรที่สนใจศึกษาเป็นประชากรที่มีการกำหนดคาบเวลาและขอบเขตของการสำรวจ ทำให้ทราบจำนวนประชากรที่สนใจศึกษาทั้งหมด ( $N$ ) หรือกล่าวได้ว่าการสำรวจตัวอย่างเป็นการอธิบายลักษณะประชากร ณ คาบเวลาใดเวลาหนึ่ง ในขอบเขตการสำรวจเท่านั้น และสภาพของการสำรวจตัวอย่างมีตัวแปรที่สนใจศึกษาเป็นจำนวนมาก จึงเป็นการยากที่จะพิจารณาการแจกแจงความน่าจะเป็นของตัวแปรต่าง ๆ พร้อมกัน

การสำรวจ (Survey) เป็นวิธีการเก็บรวบรวมข้อมูลตามสภาพที่เกิดขึ้นจริง เพื่อหาข้อสรุปเกี่ยวกับประชากร โดยการเก็บทุกหน่วยของประชากร เรียกว่าสำมะโน (Census) หรือเก็บเพียงบางหน่วยของประชากร เรียกว่าการสำรวจตัวอย่าง (Sample Survey) เนื่องจากประชากรมีขนาดใหญ่ การเก็บรวบรวมข้อมูลจากทุกหน่วยประชากรต้องใช้ทรัพยากรด้านต่าง ๆ ทั้งเวลา กำลังคน และค่าใช้จ่ายเป็นจำนวนมาก ดังนั้นภายใต้ข้อจำกัดทางด้านทรัพยากร จึงต้องทำการสุ่มตัวอย่างที่ดีขึ้นมาเพื่อเป็นตัวแทนของประชากรสำหรับอธิบายคุณลักษณะประชากร (Population Characteristic) ที่สนใจ โดยให้มีความถูกต้องและความน่าเชื่อถือ ดังนั้นจำเป็นที่จะต้องใช้ทฤษฎี

การสำรวจตัวอย่างเข้ามาช่วยในการวางแผนการเก็บรวบรวมข้อมูล เพื่อให้การดำเนินงานเป็นไปตามวัตถุประสงค์ที่ต้องการ เนื่องจากทฤษฎีการสุ่มตัวอย่างโดยทั่วไปไม่สนใจการแจกแจงความน่าจะเป็น ดังนั้นจึงเกิดความแตกต่างเชิงแนวคิด

แนวคิดในการสำรวจตัวอย่างนั้นพิจารณาได้เป็น 2 แนวคิด คือ แนวคิดของประชากรคงที่ (Fixed population Approach) และแนวคิดของอภิประชากร (Superpopulation Approach) ที่มีความแตกต่างกัน ในด้านแนวคิดของประชากรคงที่ (Fixed population Approach) จะพิจารณาหน่วยประชากร (Population Unit) เป็นค่าคงที่ที่ไม่ทราบค่าซึ่งสามารถแบ่งชนิดของการสุ่มตัวอย่างของแนวคิดนี้ออกเป็น 2 ชนิด คือ การสุ่มตัวอย่างจากประชากรที่ไม่เป็นเชิงความน่าจะเป็น (Non-Probability Sampling) และการสุ่มตัวอย่างจากประชากรเชิงความน่าจะเป็น (Probability Sampling) โดยที่การสุ่มตัวอย่างจากประชากรที่ไม่เป็นเชิงความน่าจะเป็น เป็นการสุ่มตัวอย่างที่ไม่คำนึงถึงความน่าจะเป็นที่แต่ละหน่วยในประชากรจะถูกสุ่มมาเป็นตัวอย่าง มีความจำเป็นสำหรับกรณีที่ไม่สามารถกำหนดขอบเขตของประชากรได้อย่างแน่นอน หรือไม่สามารถสร้างกรอบสำหรับเลือกตัวอย่าง (Sampling Frame) ที่ต้องการได้ แต่ตัวอย่างที่ได้ไม่เป็นตัวแทนที่ดีของประชากร และไม่สามารถอาศัยวิธีการทางสถิติในการอธิบายคุณภาพของตัวประมาณได้ ตัวอย่างของการสุ่มตัวอย่างประเภทนี้คือ การสุ่มตัวอย่างแบบบังเอิญ (Accidental Sampling) เป็นการเก็บรวบรวมข้อมูลจากตัวอย่างเท่าที่จะหาได้จนครบตามจำนวนที่ต้องการ โดยไม่มีกฎเกณฑ์ การสุ่มตัวอย่างแบบเจาะจง (Purposive Sampling or Judgement Sampling) เป็นการเลือกตัวอย่างโดยเจาะจงเอาเฉพาะตัวอย่างที่คิดว่าเป็นแบบฉบับหรือตัวแทน เพราะคิดว่าใกล้เคียงกับค่าเฉลี่ย การสุ่มตัวอย่างแบบโควตา (Quota Sampling) เป็นการเลือกตัวอย่างโดยกำหนดจำนวนตัวอย่างขึ้นเองว่าจะตัดสินใจเลือกหน่วยใดมาเป็นตัวอย่าง และการสุ่มตัวอย่างตามสะดวก (Convenience Sampling) เป็นการเลือกตัวอย่างตามความสะดวกหรือความง่ายในการเก็บรวบรวมข้อมูล ในขณะที่การสุ่มตัวอย่างจากประชากรเชิงความน่าจะเป็น เป็นการสุ่มตัวอย่างโดยคำนึงถึงความน่าจะเป็นที่แต่ละหน่วยในประชากรถูกเลือกอย่างสุ่ม แต่ละหน่วยในประชากรจะถูกเลือกเป็นตัวอย่างด้วยความน่าจะเป็นที่เท่ากันหรือไม่เท่ากันนั้น ขึ้นอยู่กับแผนการสุ่มตัวอย่างที่ใช้ ทำให้สามารถบอกคุณภาพของตัวประมาณในด้านความแม่นยำ (Precision) ได้ เนื่องจากวัตถุประสงค์หลักของการสำรวจตัวอย่าง คือ การอธิบายสภาพของประชากร วิธีการที่ใช้ในการอนุมานจากตัวอย่างไปสู่ประชากรจึงเป็นสิ่งสำคัญ ซึ่งต้องมีการวัดคุณภาพของการอนุมานด้วย ดังนั้นการสุ่มตัวอย่างต้องคำนึงถึงความสามารถในการแสดงคุณภาพของการอนุมาน จึงทำให้การสุ่มตัวอย่างเชิงความน่าจะเป็น เหมาะสำหรับการใช้งานมากกว่าการสุ่มตัวอย่างที่ไม่เป็นเชิงความน่าจะเป็น โดยการสุ่มตัวอย่างเชิงความน่าจะเป็นนี้



อาจแบ่งได้เป็น 4 แผนแบบแม่บท<sup>1</sup> คือ การสุ่มตัวอย่างแบบง่าย (Simple Random Sampling) การสุ่มตัวอย่างแบบแบ่งชั้นภูมิ (Stratified Sampling) การสุ่มตัวอย่างแบบมีระบบ (Systematic Sampling) และการสุ่มตัวอย่างแบบกลุ่ม (Cluster Sampling) โดยที่แผนการสุ่มตัวอย่างที่เหมาะสมจะขึ้นอยู่กับลักษณะความเป็นไปของประชากรที่สนใจศึกษา และสอดคล้องกับความต้องการใช้ประโยชน์เป็นสำคัญ

แนวคิดของอภิประชากร (Superpopulation Approach) คือการพิจารณาหน่วยประชากรเป็นตัวแปรสุ่ม (Random Variable) ที่ถูกกำกับด้วยโครงสร้างเชิงสโตแคสติก (Stochastic Process) นั่นคือค่าจริงของหน่วยประชากรเป็นตัวแปรสุ่มที่ต้องมีการแจกแจงความน่าจะเป็น (Probability Distribution) เข้ามาเกี่ยวข้อง ตัวประมาณถูกพิจารณาในลักษณะของค่าคาดหวังภายใต้เกณฑ์ความน่าจะเป็น ซึ่งในความเป็นจริงแล้วตัวแปรที่สนใจศึกษาของประชากรส่วนใหญ่ไม่ใช่ค่าคงที่ เช่น ตัวแปรในเรื่องรายได้ของประชาชนไทย จะเห็นได้ว่าค่าของรายได้นั้นขึ้นอยู่กับปัจจัยต่าง ๆ และมีค่าที่เป็นไปได้มากมาย จึงเหมาะสมที่จะใช้แนวคิดของอภิประชากร (Superpopulation Approach) มากกว่าแนวคิดของประชากรคงที่ (Fixed population Approach) ในการอธิบายถึงลักษณะประชากร ซึ่งถ้าพิจารณาค่าของตัวแปรที่สนใจด้วยวิธีการที่ไม่เหมาะสม จะทำให้ค่าความคลาดเคลื่อนของการประมาณที่เกิดขึ้นมีค่าสูง และการอธิบายถึงลักษณะของประชากรที่สนใจนั้นก็ผิดพลาดจากค่าที่ควรจะเป็น

ดังนั้นแนวคิดของประชากรคงที่ (Fixed population Approach) และอภิประชากร (Superpopulation Approach) มีความคล้ายคลึงกันในลักษณะของประชากรที่สนใจมีขนาดจำกัด แต่รายละเอียดต่าง ๆ ของแนวคิดทั้งสองมีความแตกต่างกัน โดยที่แนวคิดของอภิประชากร (Superpopulation Approach) นั้นได้ถูกพัฒนาขึ้นมาเนื่องจากข้อจำกัดทางด้านกรแจกแจงความน่าจะเป็นที่แนวคิดของประชากรคงที่ (Fixed population Approach) ไม่สนใจการแจกแจงความน่าจะเป็นของตัวอย่าง ถือว่าการแจกแจงเป็น Distribution Free ในขณะที่แนวคิดของอภิประชากร (Superpopulation Approach) สามารถใช้การแจกแจงความน่าจะเป็นของหน่วยประชากรในการสร้างและพัฒนาตัวแบบเชิงความน่าจะเป็น (Probabilistic Model) ทำให้สามารถวิเคราะห์ในรายละเอียดต่าง ๆ ได้ดีกว่า จะเห็นได้ว่างานวิจัยที่เกี่ยวกับการสำรวจตัวอย่างโดยทั่วไปนั้นพิจารณาลักษณะของประชากรที่สนใจด้วยแนวคิดของประชากรคงที่ (Fixed population Approach) เป็นส่วนใหญ่ เนื่องจากพิจารณาค่าของตัวแปรที่สนใจของหน่วยต่าง ๆ ในประชากร เป็นค่าคงที่ที่ไม่ทราบค่า จึงต้องทำการสำรวจตัวอย่าง แล้วทำการอนุมานไปสู่ประชากร แต่ข้อมูลบางประเภท

<sup>1</sup> สุชาติดา กิระนันท์, ทฤษฎีและวิธีการสำรวจตัวอย่าง (กรุงเทพมหานคร:โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย,2542), หน้า 25.

ไม่เหมาะสมที่จะอธิบายค่าของหน่วยประชากรเป็นค่าคงที่ ดังนั้นจึงต้องมีการพิจารณาลักษณะของประชากรด้วยแนวคิดของอภิประชากร (Superpopulation Approach) ซึ่งแนวคิดของอภิประชากร (Superpopulation Approach) ยังไม่เป็นที่แพร่หลาย และมีความซับซ้อนมาก จึงทำให้งานวิจัยต่าง ๆ นั้นพิจารณาเฉพาะแนวคิดของประชากรคงที่ (Fixed population Approach) ซึ่งทำให้ผลที่ได้คลาดเคลื่อนไปจากความเป็นจริง

ผู้วิจัยจึงสนใจที่จะศึกษาโครงสร้างเชิงลึกของแนวคิดทั้งสอง เพื่อนำเสนอแนวคิดที่เหมาะสมสำหรับงานวิจัยเชิงสำรวจ

## 1.2 วัตถุประสงค์

1. เพื่อศึกษาแนวคิดและประโยชน์ของอภิประชากร
2. เพื่อศึกษาและระบุความแตกต่างระหว่างแนวคิดของประชากรคงที่และแนวคิดของอภิประชากร
3. เพื่อเปรียบเทียบวิธีการสุ่มตัวอย่างและวิธีการประมาณค่าลักษณะประชากร ภายใต้กรอบแนวคิดของวิธีการอนุมานทางสถิติแบบคลาสสิก (Classical Statistical Inference) ของแนวคิดของอภิประชากรกับแนวคิดของประชากรคงที่

## 1.3 ขอบเขตของการวิจัย

ศึกษาเฉพาะกรณีการประมาณค่าเฉลี่ยประชากร โดยใช้แผนการสุ่มตัวอย่างแบบง่ายแบบไม่ใส่คืน ทั้งในส่วน of แนวคิดของประชากรคงที่ และแนวคิดของอภิประชากร

## 1.4 วิธีดำเนินงานวิจัย

1. ศึกษาและอธิบายแนวคิดรวมถึงประโยชน์ของแนวคิดของอภิประชากรจากตำราสถิติทั่วไป
2. ระบุความแตกต่างระหว่างแนวคิดของประชากรคงที่และแนวคิดของอภิประชากร
3. ทำการเปรียบเทียบวิธีการสุ่มตัวอย่างและวิธีการประมาณค่าลักษณะประชากร ภายใต้กรอบแนวคิดของวิธีการอนุมานทางสถิติแบบคลาสสิก (Classical Statistical Inference) ของแนวคิดของอภิประชากรกับแนวคิดของประชากรคงที่ จากคุณสมบัติความไม่เอนเอียงและความถูกต้องของตัวประมาณ

### 1.5 คำจำกัดความ

ตัวแปรสุ่ม (Random Variable) หมายถึง ฟังก์ชันที่มีค่าเป็นเลขจำนวนจริงที่นิยามได้จากสมาชิกแต่ละตัวในปริภูมิตัวอย่าง หรือหมายถึง ฟังก์ชันที่มีโดเมนเป็นกลุ่มผลการทดลองของการทดลองเชิงสุ่มและมีพิสัยเป็นเซตของเลขจำนวนจริง

ความถูกต้อง (Accuracy) หมายถึง เกณฑ์ที่ใช้ในการวัดคุณภาพของตัวประมาณโดยการพิจารณาความแตกต่างของค่าต่าง ๆ ที่เป็นไปได้ของตัวประมาณ กับค่าประชากรหรือค่าจริง ค่าที่ใช้ในการวัดความถูกต้องนั้นจะใช้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Mean Square Error)

ความไม่เอนเอียง (Unbiasedness) หมายถึง การที่ตัวประมาณมีค่าคาดหวัง (Expectation) หรือค่าเฉลี่ยของการแจกแจงตัวอย่าง (Sampling Distribution) ของตัวประมาณมีค่าเท่ากับค่าจริง

หน่วยประชากร (Population Unit) หมายถึง หน่วยที่เป็นสมาชิกของประชากรที่สามารถให้ข้อมูลที่สนใจศึกษาได้

หน่วยตัวอย่าง (Sampling Unit) หมายถึง หน่วยประชากรที่ถูกเลือกขึ้นมาเป็นตัวอย่าง

### 1.6 ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย

1. เพื่อเกิดความเข้าใจที่ชัดเจนเกี่ยวกับแนวคิดของทฤษฎีการสุ่มตัวอย่างที่แตกต่างกัน
2. ทำให้สามารถนำหลักการที่เกี่ยวกับการสุ่มตัวอย่างไปใช้ในทางปฏิบัติได้อย่างถูกต้อง
3. เพื่อเป็นแนวทางในการพัฒนาแนวคิดของอภิประชากรต่อไป

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย