

บทที่ 1

ความรู้ทั่วไปเกี่ยวกับสถิติ

สถิติได้เข้ามามีบทบาทในชีวิตประจำวันของคนเรามากขึ้น ส่วนใหญ่เรามักจะได้ยินหรือได้ฟังสถิติอยู่เสมอ ไม่ว่าจะเป็นจากการอภิปราย จากการสนทนา จากสื่อทางวิทยุ จากโทรทัศน์ และสื่อสิ่งพิมพ์ ต่างก็มีตัวเลขสถิติแสดงอยู่บ่อย ๆ เช่น การจัดบันทึกค่าใช้จ่ายรายเดือน การจัดบันทึกการขายของบริษัท เป็นต้น เมื่อได้ยินคำว่าสถิติหลายคนจะนึกถึงตัวเลข ทั้งนี้เพราะสถิติที่คนทั่วไปรู้จักและเกี่ยวข้องจะอยู่ในรูปของตัวเลข ตัวเลขมักแสดงในรูปของยอดรวม ซึ่งประมวลจากข้อมูลเบื้องต้นเป็นจำนวนมาก ๆ หรืออาจอยู่ในรูปใด ๆ ซึ่งได้มาด้วยการดำเนินการคำนวณอย่างหนึ่งอย่างใดตามวิธีการทางสถิติ เพื่อทราบความสัมพันธ์หรือลักษณะโดยรวมของข้อมูลเบื้องต้นเหล่านั้น ดังนั้นสถิติเป็นวิชาการที่เกี่ยวข้องโดยตรงกับทุกคน เพราะการดำรงชีวิตของเราอยู่กับการเปรียบเทียบ การวัด และการประมาณค่า ตลอดจนการนำตัวเลขมาเป็นเกณฑ์มาตรฐานต่าง ๆ และยังเป็นเครื่องมือที่ใช้ในทางวิทยาศาสตร์เพื่อใช้สรุปผลการทดลองได้อย่างมากมาย โดยสถิติเป็นวิชาการที่เกี่ยวข้องกับข้อมูลและการประมวลผลของข้อมูล ยังรวมถึงระเบียบวิธีการทางสถิติด้วย

ความหมายของสถิติ

คำว่า สถิติ ตรงกับคำศัพท์ภาษาอังกฤษว่า Statistics เป็นคำที่แปลมาจากคำศัพท์ Statistik ในภาษาเยอรมันว่า ซึ่งคิดค้นขึ้นโดย Gottfried Achenwall ในปีค.ศ. 1749 เป็นคำที่มีรากศัพท์เดียวกับคำว่า State ที่หมายถึง ข้อมูลที่เกี่ยวข้องและจะเป็นประโยชน์ต่อการบริหารงานของรัฐ และขณะเดียวกันก็ถือว่าเป็นรัฐศาสตร์แขนงหนึ่งว่าด้วยการใช้ข้อมูลมาบริหารกิจการของรัฐ (เกรียง กิจบำรุงรัตน์. 2562 : 10 และปริตากรณ์ กาญจนสำราญวงศ์. 2560 : 1)

ดังนั้น สมัยก่อน สถิติ จึงหมายถึง ข้อมูลหรือข่าวสารที่เป็นประโยชน์ต่อการบริหารจัดการในภาครัฐ เช่น ข้อมูลการสำรวจกำลังคนเพื่อประโยชน์ในด้านการทหาร ข้อมูลด้านการเก็บภาษีอากร เพื่อเป็นรายได้ของรัฐ ข้อมูลการสำรวจที่ดินทำกิน การทำสำมะโนประชากร เป็นต้น ต่อมาสถิติได้มี

ความหมายที่กว้างขวางขึ้นกว่าเดิม เนื่องจากนักคณิตศาสตร์และนักสถิติได้คิดค้นและพัฒนาทฤษฎีทางสถิติไว้มากมาย จึงสามารถสรุปความหมายของสถิติได้ 2 ความหมาย ดังนี้ (ปริดาภรณ์ กาญจนสำราญวงศ์. 2560 : 1)

ความหมายที่ 1 สถิติ หมายถึง ตัวเลขที่แทนจำนวนหรือข้อเท็จจริงเกี่ยวกับเรื่องต่าง ๆ ที่เราสนใจ เช่น สถิติปริมาณน้ำฝนที่ตกในจังหวัดบุรีรัมย์ สถิติการนำเข้าและส่งออกน้ำมันปาล์ม สถิติการเกิดอุบัติเหตุบนท้องถนน เป็นต้น

ความหมายที่ 2 สถิติ หมายถึง ศาสตร์ที่ว่าด้วยวิธีการที่ใช้ในการศึกษาข้อมูล หรือระเบียบวิธีการทางสถิติ ซึ่งประกอบด้วย การเก็บรวบรวมข้อมูล การนำเสนอข้อมูล การวิเคราะห์ข้อมูล และการแปลความหมายข้อมูล ตามสถิติในแง่ที่เป็นศาสตร์ แบ่งออกเป็น 2 ประเภท ดังต่อไปนี้ (Mann. 2011 : 2)

1. สถิติพรรณนา (Descriptive Statistics) เป็นวิธีการทางสถิติที่เกี่ยวกับการเก็บรวบรวมข้อมูล โดยอาจนำเสนอข้อมูลด้วยการบรรยาย ตาราง กราฟ แผนภูมิ แผนภาพ หรืออาจนำเสนอค่าวัดคุณลักษณะของสิ่งที่สนใจ เช่น การวัดแนวโน้มเข้าสู่ส่วนกลาง การวัดการกระจาย เป็นต้น ทั้งนี้จะไม่นำผลที่ได้จากการวิเคราะห์ไปอ้างสรุปถึงข้อมูลกลุ่มอื่น (พิศมัย หาญมงคลพิพัฒน์. 2560 : 1)

2. สถิติเชิงอนุมาน (Inferential Statistics) เป็นการเก็บรวบรวมข้อมูลจากกลุ่มตัวอย่างแล้ววิเคราะห์ข้อมูลโดยใช้ทฤษฎีความน่าจะเป็นเข้ามาประยุกต์ใช้ เพื่อนำไปสรุปลักษณะบางประการของประชากร เช่น การประมาณค่า (Estimation) และการทดสอบสมมติฐาน (Testing Hypothesis) เป็นต้น (ปริดาภรณ์ กาญจนสำราญวงศ์. 2560 : 2 และยุทธ ใภยวรรณ. 2555 : 4)

ความหมายของข้อมูล

ข้อมูล (Data) หมายถึง ข้อเท็จจริงเกี่ยวกับเรื่องใดเรื่องหนึ่งที่สนใจศึกษา ซึ่งอาจเกี่ยวข้องกับคน สัตว์ หรือสิ่งของ ข้อมูลที่ยังไม่ผ่านการประมวลผลทางสถิติจะเรียกว่า ข้อมูลดิบ (Raw Data) โดยข้อมูลอาจได้มาจากการนับ การสังเกต หรือการจดบันทึกในรูปของข้อความหรือตัวเลขที่ไขบอกปริมาณหรือลักษณะของสิ่งที่สนใจ สำหรับข้อมูลที่เก็บรวบรวมในรูปของตัวเลขเพื่อนำมาประมวลผลหาความหมายที่แน่นอนจะเรียกว่า ข้อมูลสถิติ (Statistical Data) (พิศมัย หาญมงคลพิพัฒน์. 2560 : 1 และนพพร ณะชัยจันทร์. 2557 : 9) เช่น สถิติปริมาณน้ำฝนในจังหวัดบุรีรัมย์ สถิติ

ประชากร สถิติข้าวที่ประเทศไทยผลิตได้ สถิติการเกิดอุบัติเหตุ สถิติสินค้าส่งออกของประเทศไทย ระหว่างปีพ.ศ. 2561-2564 เป็นต้น

ประเภทของข้อมูล

ปรีดาภรณ์ กาญจนสำราญวงศ์ (2560 : 3-6) ; ยุธีย์ วรวิชัยยันต์ (2551 : 4-5) ; พิศมัย หาญมงคลพิพัฒน์ (2560 : 2-3) และ Doane & Seward (2009 : 25-27) ได้กล่าวถึง ประเภทของข้อมูล โดยข้อมูลที่เก็บรวบรวมเพื่อวิเคราะห์จะประกอบด้วยข้อมูลหลาย ๆ ประเภทรวมกัน การแบ่งประเภทของข้อมูลจึงมีความสำคัญ จึงสามารถแบ่งข้อมูลตามลักษณะต่าง ๆ ดังนี้

1. การแบ่งตามแหล่งที่มาของข้อมูล

ถ้าแบ่งตามแหล่งที่มาของข้อมูลจะแบ่งข้อมูลออกได้ 2 ประเภท คือ ข้อมูลปฐมภูมิ (Primary Data) และข้อมูลทุติยภูมิ (Secondary Data) ดังนี้

1.1 ข้อมูลปฐมภูมิ เป็นข้อมูลที่ผู้ใช้หรือหน่วยงานที่จะใช้ข้อมูลเป็นผู้เก็บรวบรวมข้อมูลเองตามวัตถุประสงค์ที่ต้องการศึกษา อาจมาจากการสัมภาษณ์ สังเกตการณ์ หรือแบบสอบถาม ข้อมูลปฐมภูมิจึงมีรายละเอียดของข้อมูลตรงตามที่ใช้ต้องการ แต่จะเสียเวลาและค่าใช้จ่ายจำนวนมาก และข้อมูลที่ไต่ยังเป็นข้อมูลดิบ

1.2 ข้อมูลทุติยภูมิ เป็นข้อมูลที่ผู้ใช้หรือหน่วยงานที่จะใช้ข้อมูลไม่ได้เก็บรวบรวมข้อมูลเอง แต่มีหน่วยงานหรือผู้อื่นเก็บรวบรวมไว้แล้ว ผู้ใช้สามารถนำมาใช้ได้เลยจึงเป็นการประหยัดเวลาและค่าใช้จ่าย แต่มีข้อเสียคือ ข้อมูลทุติยภูมิจะไม่ตรงกับความต้องการของผู้ใช้หรืออาจมีรายละเอียดไม่เพียงพอ ผู้ที่นำข้อมูลมาใช้จึงต้องใช้ด้วยความระมัดระวัง

2. การแบ่งตามมาตรการวัด

การแบ่งตามมาตรการวัดจะแบ่งข้อมูลออกได้ 4 ประเภท คือ มาตรการวัดนามบัญญัติ (Nominal Scale) มาตรการวัดอันดับ (Ordinal Scale) มาตรการวัดแบบช่วงหรือมาตรการวัดอันตรภาค (Interval Scale) และมาตรการวัดอัตราส่วน (Ratio Scale) ดังนี้

1. มาตรฐานนามบัญญัติ เป็นการวัดค่าที่ง่ายที่สุดหรือสะดวกต่อการใช้อย่างมากที่สุด เพราะเป็นการแบ่งกลุ่มของข้อมูล โดยถือว่าแต่ละกลุ่มมีความเสมอภาคหรือเท่าเทียมกัน ค่าตัวเลขที่กำหนดให้แต่ละกลุ่มจะไม่มี ความหมายและไม่สามารถนำมาคำนวณได้ เช่น อาชีพ เพศ สถานภาพ เชื้อชาติ บ้านเลขที่ เบอร์โทรศัพท์ รหัสไปรษณีย์ เป็นต้น

2. มาตรฐานอันดับ เป็นการวัดที่แสดงว่าข้อมูลที่อยู่ในแต่ละกลุ่มมีความแตกต่างกัน โดยพิจารณาจากลำดับ นั่นคือ สามารถบอกได้ว่ากลุ่มใดที่มากกว่าหรือน้อยกว่ากลุ่มอื่น ๆ แต่ไม่สามารถบอกถึงปริมาณของความมากกว่าหรือน้อยกว่าได้ว่าเป็นเท่าใด และค่าที่กำหนดให้แต่ละกลุ่มไม่สามารถนำมาคำนวณได้ เช่น ระดับความคิดเห็น ระดับการศึกษา ระดับความพึงพอใจ อันดับเพลง ยอดนิยม อันดับหนังยอดเยี่ยม การแข่งขันต่าง ๆ การประกวดต่าง ๆ เป็นต้น

3. มาตรฐานช่วงหรือมาตรฐานอัตราภาค เป็นการวัดที่แบ่งสิ่งที่ศึกษาออกเป็นระดับหรือเป็นช่วง ๆ โดยแต่ละช่วงมีขนาดหรือระยะห่างเท่า ๆ กัน สามารถบอกระยะห่างของช่วงได้อีกทั้งบอกได้ว่ามากหรือน้อยกว่ากันเท่าไร เช่น อุณหภูมิ คะแนนสอบ ซึ่งสามารถนำตัวเลขเหล่านี้มาบวก ลบได้ แต่คุณ หารไม่ได้ โดยค่าศูนย์ของข้อมูลชนิดนี้เป็นศูนย์สมมติ ไม่ใช่ศูนย์แท้ เช่น อุณหภูมิ 0 องศาเซลเซียส (0°C) ไม่ได้หมายความว่า ณ จุดนั้นไม่มีความร้อนอยู่เลย หรือนักศึกษาได้คะแนนสอบ 0 คะแนน ก็ไม่ได้หมายความว่า นักศึกษาคนนั้นไม่มีความรู้เลย แต่เป็นเพียงตัวเลขที่บอกว่า นักศึกษาคนนั้นทำข้อสอบไม่ได้ ความแตกต่างระหว่าง 50 องศาเซลเซียส (50°C) กับ 25 องศาเซลเซียส (25°C) นั้นบอกได้เพียงว่าห่างกัน 25 องศาเซลเซียส (25°C) แต่ไม่ได้หมายความว่า อุณหภูมิต่างกันเป็น 2 เท่า หรือถ้าได้คะแนนสอบ 70 คะแนน ก็ไม่ได้หมายความว่ามีความรู้เป็น 2 เท่าของผู้ที่สอบได้คะแนนสอบ 35 คะแนน

4. มาตรฐานอัตราส่วน เป็นการวัดที่ละเอียดและสมบูรณ์ที่สุดที่สามารถบอกความแตกต่างในเชิงปริมาณได้ โดยแบ่งสิ่งที่ศึกษาออกเป็นช่วง ๆ เหมือนมาตรฐานอัตราภาค ที่แต่ละช่วงมีระยะห่างกันเท่ากัน และค่าศูนย์ของข้อมูลชนิดนี้เป็นศูนย์แท้ ซึ่งหมายถึงไม่มีอะไรเลย หรือมีจุดที่เริ่มต้นที่แท้จริง และสามารถนำตัวเลขนี้มาบวก ลบ คูณ หรือหารได้ เช่น เวลา น้ำหนัก ความสูง อายุ ระยะทาง ความเร็ว เป็นต้น

ตาราง 1.1 คุณสมบัติของมาตรการวัดในระดับต่าง ๆ

มาตรการของ การวัด	คุณสมบัติ					
	แบ่งเป็น กลุ่มได้	สามารถ บอกลำดับได้	แบ่งเป็นช่วง เท่า ๆ กันได้	บวกลบ ได้	คูณหาร ได้	มีศูนย์ แท้
นามบัญญัติ						
อันดับ	✓	-	-	-	-	-
อันดับภาค	✓	✓	-	-	-	-
อัตราส่วน	✓	✓	✓	✓	-	-
	✓	✓	✓	✓	✓	✓

ที่มา : ปริดาภรณ์ กาญจนสำราญวงศ์ (2560 : 4)

3. การแบ่งตามลักษณะของข้อมูล

ถ้าแบ่งตามลักษณะของข้อมูลจะแบ่งข้อมูลออกได้ 2 ประเภท คือ ข้อมูลเชิงปริมาณ (Quantitative Data) และข้อมูลเชิงคุณภาพ (Qualitative Data) ดังนี้

3.1 ข้อมูลเชิงปริมาณ หมายถึง ข้อมูลที่เป็นตัวเลขและสื่อความหมายตามค่าตัวเลขได้โดยตรง เช่น ยอดขายสินค้า รายได้ รายจ่าย น้ำหนัก เป็นต้น ซึ่งแบ่งเป็น 2 ประเภท คือ

1. ข้อมูลแบบไม่ต่อเนื่อง (Discrete Data) หมายถึง ข้อมูลที่ได้จากการนับซึ่งมีค่าเป็นจำนวนเต็ม เช่น จำนวนนักศึกษา จำนวนสมาชิกในครัวเรือน จำนวนสินค้า จำนวนคอมพิวเตอร์ เป็นต้น (อนุวัติ คุณแก้ว. 2560 : 27)

2. ข้อมูลแบบต่อเนื่อง (Continuous Data) หมายถึง ข้อมูลที่ได้จากการวัดซึ่งมีค่าได้ทุกช่วงในช่วงที่กำหนด เช่น ความสูง อายุ ระยะทาง ความยาว รายได้ ความลึก เป็นต้น (อนุวัติ คุณแก้ว. 2560 : 27)

3.2 ข้อมูลเชิงคุณภาพ หมายถึง ข้อมูลที่เป็นข้อความหรือตัวเลขที่แสดงลักษณะประเภท หรือกลุ่ม ซึ่งไม่สามารถวัดค่าออกมาเป็นตัวเลขและสื่อความหมายตามค่าตัวเลขได้โดยตรง เช่น เพศ สัญชาติ ระดับการศึกษา ทักษะ ทักษะ อาชีพ ความคิดเห็น เป็นต้น

ตาราง 1.2 ความสัมพันธ์ระหว่างประเภทของข้อมูลที่แบ่งตามมาตรฐานของการวัดและแบ่งตามลักษณะของข้อมูล

มาตรฐานของการวัด	ลักษณะของข้อมูล
มาตรวัดนามบัญญัติ มาตรวัดอันดับ มาตรวัดอันตรภาค มาตรวัดอัตราส่วน	ข้อมูลเชิงคุณภาพ ข้อมูลเชิงปริมาณ

ที่มา : ปรีดาภรณ์ กาญจนสำราญวงศ์ (2560 : 6)

นิยามศัพท์ที่เกี่ยวข้อง

นิยามศัพท์ที่เกี่ยวข้องกับสถิติที่ควรทราบ เพื่อใช้ในบทถัดไป ดังต่อไปนี้ (อำนาจ วังจิ้น และพรรณี บุญสุยา. 2553 : 1-2 และปรีดาภรณ์ กาญจนสำราญวงศ์. 2560 : 10-11)

1. ประชากร (Population) หมายถึง กลุ่มของสิ่งต่าง ๆ ที่เราสนใจศึกษาทั้งหมด ซึ่งอาจจะเป็นบุคคล สัตว์ สิ่งของ หรือเหตุการณ์ที่เราสนใจ เช่น การสำรวจความพึงพอใจของนักศึกษาในรายวิชาหลักสถิติ ดังนั้นประชากร คือ นักศึกษาทุกคนที่เรียนในรายวิชาหลักสถิติ เป็นต้น ประชากรแบ่งได้ 2 ประเภท คือ

1.1 ประชากรที่มีจำนวนจำกัด หมายถึง ประชากรที่มีจำนวนสมาชิกจำกัดและสามารถนับจำนวนได้แน่นอน เช่น จำนวนอุปกรณ์การเรียน จำนวนพนักงานในบริษัท จำนวนคนในจังหวัดบุรีรัมย์ เป็นต้น

1.2 ประชากรที่มีจำนวนอนันต์ หมายถึง ประชากรที่มีสมาชิกไม่สามารถนับจำนวนที่แน่นอนได้ เช่น จำนวนฝุ่นในอากาศ จำนวนเม็ดทรายในหาดทรายทอง เป็นต้น

2. ตัวอย่าง (Sample) หมายถึง ประชากรจำนวนหนึ่งที่ได้ถูกเลือกออกมาเป็นตัวแทนของประชากรทั้งหมดที่เราสนใจศึกษา เช่น ในการสำรวจค่าใช้จ่ายเฉลี่ยของพนักงานที่ทำงานอยู่ในจังหวัดบุรีรัมย์ ประชากร คือ พนักงานทุกคนที่ทำงานอยู่ในจังหวัดบุรีรัมย์ แต่เนื่องจากเราไม่สามารถ

สอบถามคนทั้งจังหวัดได้ ดังนั้น จึงเลือกคนเพียงบางส่วนมาเป็นผู้ตอบ ซึ่งในที่นี้ตัวอย่าง คือ พนักงาน
บางคนที่ทำงานอยู่ในจังหวัดบุรีรัมย์ที่ถูกเลือกมาเป็นตัวอย่าง เป็นต้น

3. ค่าพารามิเตอร์ (Parameter) หมายถึง ค่าคงที่ที่คำนวณมาจากทุก ๆ หน่วยของ
ประชากรที่ต้องการศึกษา เป็นตัวที่บ่งชี้ถึงคุณลักษณะของประชากร

4. ค่าสถิติ (Statistics) หมายถึง ค่าที่คำนวณได้จากเป็นตัวอย่างที่ศึกษา เป็นตัวที่บ่งชี้ถึง
คุณลักษณะของตัวอย่าง

จากความหมายของค่าพารามิเตอร์และค่าสถิติ สามารถสรุปค่าพารามิเตอร์และค่าสถิติ
ต่าง ๆ ดังตาราง 1.3

ตาราง 1.3 สัญลักษณ์ค่าพารามิเตอร์และค่าสถิติ

ความหมาย	สัญลักษณ์	
	ค่าพารามิเตอร์	ค่าสถิติ
ค่าเฉลี่ย	μ (อ่านว่า มิว)	\bar{X} (อ่านว่า เอกซ์บาร์)
ส่วนเบี่ยงเบนมาตรฐาน	σ (อ่านว่า ซิกมา)	S (อ่านว่า เอส)
ความแปรปรวน	σ^2 (อ่านว่า ซิกมาสแควร์)	S^2 (อ่านว่า เอสสแควร์)
สัดส่วน	p (อ่านว่า พี)	\hat{p} (อ่านว่า พีแคป)
สัมประสิทธิ์สหสัมพันธ์	ρ (อ่านว่า โรล)	r (อ่านว่า อาร์)
สัมประสิทธิ์การถดถอย	β (อ่านว่า เบต้า)	b (อ่านว่า บี)
จำนวน	N	n

ที่มา : ปรีดาภรณ์ กาญจนสำราญวงศ์ (2560 : 11)

การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลเป็นขั้นตอนหนึ่งของกระบวนการทางสถิติที่มีความสำคัญ เพื่อให้ได้มาซึ่งข้อมูลที่ตอบสนองตามวัตถุประสงค์ ควรทำความเข้าใจทั้งลักษณะและประเภทของข้อมูล ตลอดจนวิธีการเก็บข้อมูล ซึ่งวิธีการเก็บรวบรวมข้อมูลนั้นสามารถทำได้หลายวิธี ขึ้นอยู่กับความต้องการว่าเราต้องการข้อมูลลักษณะเช่นไร โดยทั่วไปวิธีเก็บข้อมูลแบ่งได้ 3 วิธี คือ (เกรียง กิจบำรุงรัตน์. 2562 : 20-21 และปรีดาภรณ์ กาญจนสำราญวงศ์. 2560 : 10-11)

1. การเก็บรวบรวมข้อมูลจากงานทะเบียนหรือการบันทึก (Registration)

เป็นการเก็บข้อมูลโดยการบันทึกข้อมูลอย่างต่อเนื่องกันไป อาจมีการแก้ไขปรับปรุงเปลี่ยนแปลงรายการหรือระบบวิธีการบันทึกเก็บข้อมูลให้ทันสมัยขึ้นก็ได้ เช่น ทะเบียนบ้าน ทะเบียนผู้ป่วยนอก เวชระเบียน ทะเบียนประวัตินักศึกษา เป็นต้น การเก็บข้อมูลโดยวิธีนี้เป็นการคัดลอกจากทะเบียนที่มีผู้บันทึกไว้แล้วเมื่อมีเหตุการณ์เกิดขึ้น ผู้ศึกษาเพียงแต่ไปทำการคัดลอกแล้วจัดให้เป็นหมวดหมู่ตามต้องการ ซึ่งมีข้อจำกัดในเรื่องของข้อมูลที่ได้อาจไม่ตรงกับความต้องการ ดังนั้นผู้ศึกษาควรทำความเข้าใจถึงแหล่งที่เหมาะสมที่จะไปคัดลอกข้อมูลมา แต่ทั้งนี้วิธีการเก็บข้อมูลจากงานทะเบียนหรือการบันทึกก็มีข้อดีในแง่ประหยัดเวลา และค่าใช้จ่ายในการไปเก็บข้อมูลเอง

2. การเก็บรวบรวมข้อมูลจากการสำรวจ (Survey) เป็นการเก็บรวบรวมข้อมูลจาก

หน่วยที่เราสนใจโดยตรง โดยอาศัยวิธีการสัมภาษณ์ สังเกต การวัด การทดสอบ แบบสอบถาม เป็นต้น การเก็บรวบรวมข้อมูลจากการสำรวจแบ่งเป็น 2 ประเภท คือ

2.1 การสำมะโน (Census) เป็นการเก็บรวบรวมข้อมูลจากทุกหน่วยในขอบเขตของประชากรที่จะศึกษา ซึ่งคำว่าสำมะโนไม่ได้หมายถึง สำมะโนประชากรเพียงอย่างเดียว แต่หมายถึงเรื่องใดก็ได้ที่มีการศึกษาทุกหน่วยในเรื่องที่สนใจ เช่น การสำมะโนน้ำในเขตพื้นที่จังหวัดบุรีรัมย์ ก็คือ การเข้าไปเก็บรวบรวมข้อมูลจากน้ำทั้งหมดที่อยู่ในเขตพื้นที่จังหวัดบุรีรัมย์ เป็นต้น ข้อดีของการทำสำมะโนคือ ได้ข้อมูลที่แท้จริงจากทุกหน่วยในเรื่องที่เราสนใจศึกษา แต่ก็มีข้อเสียคือ เสียเวลาและค่าใช้จ่ายจำนวนมาก รวมทั้งกว่าจะได้ข้อมูลครบถ้วน ข้อมูลอาจจะล้าสมัยไปแล้ว วิธีนี้จึงไม่นิยม แต่ก็ยังคงมีการทำสำมะโนประชากรทุก ๆ 10 ปี

2.2 การสำรวจด้วยตัวอย่าง (Sample Survey) เป็นการเก็บรวบรวมข้อมูลจากตัวอย่างหรือเก็บรวบรวมข้อมูลจากเพียงบางส่วนของประชากร วิธีนี้จึงประหยัดและใช้ระยะเวลาสั้นกว่า การสำรวจด้วยตัวอย่างสามารถทำได้หลายวิธี แต่ไม่ว่าจะเลือกใช้วิธีไหน สิ่งหนึ่งที่ต้องคำนึงถึงคือ ตัวอย่างที่ได้มานั้นควรจะเป็นตัวแทนที่ดีของประชากร หมายถึง มีคุณลักษณะต่าง ๆ ของประชากรครบถ้วน

3. การเก็บรวบรวมข้อมูลจากการทดลอง (Experimentation) เป็นการเก็บรวบรวม โดยจะต้องอาศัยเนื้อหาการวางแผนการทดลองมาช่วย ซึ่งใช้กับการทดลองทางด้านเกษตร วิทยาศาสตร์ การแพทย์ เช่น การศึกษาวิธีการปลูกพืชที่แตกต่างกัน 3 วิธี การเปรียบเทียบผลผลิตของข้าว 4 พันธุ์ และการเปรียบเทียบคุณภาพของยารักษาโรคผิวหนัง 4 ชนิด ทดสอบวัคซีนที่ใช้รักษาโรคติดเชื้อไวรัสโคโรนา 2019 เป็นต้น

ประโยชน์ของสถิติกับงานด้านต่าง ๆ

บทบาทของสถิติมีประโยชน์ในหลายด้าน ดังต่อไปนี้ (เมฆิยา แยมเจริญกิจ. 2559 : 5-6 ; สรชัย พิศาลบุตร. 2551 : 15-16 และมนตรี สังข์ทอง. 2557 : 14-16)

1. **ด้านการศึกษา** สถิติมีประโยชน์ต่อการศึกษาเนื่องจากการศึกษาต้องมีการพัฒนาและปรับปรุงคุณภาพตลอดเวลา ดังนั้น สถิติจึงเป็นเครื่องมือที่ทำให้ผู้สอน ผู้เรียน และบุคลากรที่เกี่ยวข้องเห็นการพัฒนาหรือเห็นการเปลี่ยนแปลงของการศึกษา ยกตัวอย่างเช่น ใช้สถิติในการเปรียบเทียบผลสัมฤทธิ์ทางการเรียนจากวิธีการสอนวิธีต่าง ๆ ใช้สถิติในการเปรียบเทียบคะแนนก่อนเรียนและคะแนนหลังเรียนของผู้เรียน ใช้สถิติในการเปรียบเทียบคะแนนผู้เรียนแต่ละห้องหรือแต่ละสาขาวิชา และใช้สถิติในการทดสอบว่าผู้เรียนมีคะแนนผ่านเกณฑ์ที่กำหนดไว้หรือไม่ เป็นต้น

2. **ด้านธุรกิจ** ในการประกอบธุรกิจนั้นธุรกิจจำเป็นต้องมีการวางแผนในการดำเนินงาน และต้องตัดสินใจที่จะดำเนินงานในด้านต่าง ๆ เช่น การวางแผนการผลิต การจัดจำหน่าย วิธีการผลิต จำนวนการผลิต เครื่องจักรที่ใช้ในการผลิต การบริหารจัดการ เป็นต้น นอกจากนั้นในบางครั้งอาจเกิดปัญหาเกี่ยวกับธุรกิจ เช่น ปัญหาด้านการผลิต ปัญหาด้านการตลาด ปัญหาเกี่ยวกับพนักงาน ปัญหาการควบคุมคุณภาพสินค้า เป็นต้น นักธุรกิจจึงต้องทำการตัดสินใจเพื่อแก้ปัญหาหรือดำเนินการตามแผนการธุรกิจ ซึ่งการวางแผนที่ดีและการตัดสินใจที่ถูกต้องย่อมส่งผลให้การดำเนินทางธุรกิจนั้น

ราบรื่น และมีโอกาสประสบความสำเร็จสูง ดังนั้น ในการวางแผนและการตัดสินใจทางธุรกิจ จำเป็นต้องอาศัยข้อมูลในอดีตที่ผ่านมาประกอบการตัดสินใจ โดยใช้วิธีทางสถิติในการวิเคราะห์ข้อมูล นอกจากนี้วิธีการทางสถิติยังสามารถช่วยพยากรณ์อนาคตของธุรกิจนั้น ๆ ได้อีกด้วย

3. ด้านการเกษตรกรรม ผู้ที่เกี่ยวข้องกับการเกษตรสามารถนำสถิติไปใช้ประโยชน์ได้มากมาย ยกตัวอย่างเช่น เกษตรกรต้องการเปรียบเทียบวิธีการปลูกพืชสองวิธีว่าวิธีการใดดีกว่ากัน ต้องการเปรียบเทียบชนิดของปุ๋ยว่าปุ๋ยชนิดใดทำให้พืชเจริญเติบโตดีกว่า ต้องการประมาณค่าผลผลิตที่จะได้ต่อไร่ ต้องการทดสอบว่าผลผลิตเป็นไปตามที่คาดไว้หรือไม่ การทดลองเกี่ยวกับคุณภาพของพันธุ์ข้าวที่เหมาะสมกับแต่ละพื้นที่การเกษตร และการทดลองเกี่ยวกับคุณภาพของปุ๋ยที่ส่งผลต่อปริมาณผลผลิต เป็นต้น นอกจากนี้สถิติยังมีประโยชน์ต่อการเกษตรเนื่องจากสถิติถูกนำมาใช้ในการกำหนดกำหนดนโยบายและวางแผนพัฒนาทางการเกษตรของประเทศอีกด้วย ซึ่งข้อมูลที่ต้องการใช้ได้แก่ คริวเรือนที่ทำการเกษตร เนื้อที่การเพาะปลูก ผลิตผลทางการเกษตร จำนวนปศุสัตว์ ราคาสินค้าเกษตรกรรม เครื่องมือเครื่องใช้ทางการเกษตร ภาวะเศรษฐกิจและสังคมของครัวเรือนเกษตร การประมง การป่าไม้ ข้อมูลเกี่ยวกับแหล่งน้ำและการชลประทาน เป็นต้น

4. ด้านอุตสาหกรรม ผู้ที่เกี่ยวข้องกับงานด้านอุตสาหกรรมสามารถนำสถิติไปใช้ประโยชน์ได้ ยกตัวอย่างเช่น ข้อมูลเกี่ยวกับความสามารถในการผลิตของเครื่องจักร ความต้องการของการตลาด ต้นทุนการผลิต จำนวนแรงงาน ค่าใช้จ่ายของสถานประกอบการ ประสิทธิภาพของแรงงาน เป็นต้น นอกจากนี้ยังสามารถใช้สถิติในการจัดทำแผนงานหรือกำหนดนโยบายและส่งเสริมอุตสาหกรรม ส่งเสริมการลงทุน และพัฒนาเทคโนโลยีทางด้านอุตสาหกรรมได้อีกด้วย

5. ด้านการพัฒนาประเทศ ในการพัฒนาประเทศไม่ว่าจะเป็นทางด้านเศรษฐกิจหรือด้านสังคม รัฐบาลและผู้ที่เกี่ยวข้องจำเป็นต้องทราบสถิติข้อมูลเกี่ยวกับประชากร และปัจจัยอื่น ๆ ที่เกี่ยวข้อง เช่น สถิติจำนวนประชากร สถิติการเกิด สถิติการตาย สถิติการย้ายถิ่นฐาน สถิติการศึกษา สถิติแรงงาน สถิติการเกษตร สถิติการส่งสินค้าออกหรือการนำเข้าสินค้าเข้า เป็นต้น สถิติเหล่านี้จะช่วยให้รัฐบาลและผู้ที่เกี่ยวข้องสามารถวางแผนหรือกำหนดนโยบายในการพัฒนาคนและพัฒนาประเทศให้ดียิ่งขึ้น

6. ด้านสาธารณสุข ผู้ที่เกี่ยวข้องกับงานด้านสาธารณสุขสามารถนำสถิติไปใช้ประโยชน์ในการปรับปรุง การจัดทำแผนพัฒนาด้านสาธารณสุข การพัฒนางานวิชาการทางการแพทย์และสาธารณสุข เพื่อให้ประชาชนมีสุขภาพอนามัยที่ดี โดยการพัฒนาจำเป็นต้องใช้ข้อมูลทางสถิติเกี่ยวกับ

อัตราการเกิด อัตราการตาย น้ำหนักเด็กแรกเกิด การเจ็บป่วย การรักษาพยาบาล ความเป็นอยู่และสภาพทางสังคมของประชากร อัตราการป่วย การอนามัยและสุขภาพ พฤติกรรมด้านการบริโภค การสูบบุหรี่และดื่มสุรา เป็นต้น

7. **ด้านคมนาคมและขนส่ง** ผู้ที่เกี่ยวข้องกับงานด้านคมนาคมและขนส่งสามารถนำสถิติไปใช้ประโยชน์ในการปรับปรุงบริการ พัฒนางานทางด้านการคมนาคมและขนส่ง และการสื่อสารของประเทศ เพื่อสนับสนุนการพัฒนาในด้านต่าง ๆ และกระจายความเจริญไปสู่ภูมิภาค ซึ่งข้อมูลที่ใช้ ได้แก่ รายรับ-รายจ่ายของการประกอบการขนส่ง ปริมาณผู้ใช้บริการในแต่ละเส้นทาง ปริมาณการขนส่งทางถนน ทางน้ำ และทางอากาศ รายละเอียดเส้นทางคมนาคม ข้อมูลเกี่ยวกับการจัดสรรความถี่วิทยุ จำนวนคร่าว์เรือนที่มีเครื่องรับวิทยุ-โทรทัศน์ เป็นต้น

เครื่องหมายผลรวม

พิศมัย หาญมงคลพิพัฒน์ (2560 : 5-7) ได้กล่าวถึง การหาผลรวมของข้อมูลหลาย ๆ ค่า เขียนแทนด้วยสัญลักษณ์ \sum (อ่านว่า ซิกม่า ; Sigma) แทนผลรวมได้

$$\sum_{i=1}^N X_i \quad \text{อ่านว่า ซิกม่า เอ็กซ์-ไอ โอมิค่าตั้งแต่ 1 ถึง เอ็น หมายถึง ผลรวมของข้อมูลทุก}$$

หน่วย จากหน่วยที่ 1 จนถึงหน่วยที่ N ดังนั้น

$$\sum_{i=1}^N X_i = X_1 + X_2 + \dots + X_N$$

โดยที่ X_1 แทน ค่าของข้อมูลในหน่วยที่ 1

X_2 แทน ค่าของข้อมูลในหน่วยที่ 2

.

.

.

X_N แทน ค่าของข้อมูลในหน่วยที่ N

N แทน จำนวนข้อมูลทั้งหมด

ในบางครั้งตัวแปรที่จะใช้ฝนการหาผลรวม อาจจะอยู่ในรูปของเลขยกกำลัง หรืออยู่ในรูปผลคูณก็ได้ เช่น

$$\sum_{i=1}^N X_i^2 = X_1^2 + X_2^2 + \dots + X_N^2$$

$$\left(\sum_{i=1}^N X_i \right)^2 = X_1 + X_2 + \dots + X_N^2$$

$$\sum_{i=1}^N X_i Y_i = X_1 Y_1 + X_2 Y_2 + \dots + X_N Y_N$$

$$\left(\sum_{i=1}^N X_i \right) \left(\sum_{i=1}^N Y_i \right) = X_1 + X_2 + \dots + X_N \quad Y_1 + Y_2 + \dots + Y_N$$

ตัวอย่าง 1.1 กำหนด $X_1 = 3, X_2 = 4, X_3 = 5, Y_1 = 1, Y_2 = 2, Y_3 = 3$ จงคำนวณค่าต่อไปนี้

$$1. \sum_{i=1}^3 X_i$$

$$2. \sum_{i=1}^3 Y_i$$

$$3. \sum_{i=1}^3 X_i^2$$

$$4. \sum_{i=1}^3 Y_i^2$$

$$5. \sum_{i=1}^3 X_i Y_i$$

$$6. \left(\sum_{i=1}^3 X_i \right) \left(\sum_{i=1}^3 Y_i \right)$$

วิธีทำ

$$\begin{aligned} 1. \sum_{i=1}^3 X_i &= X_1 + X_2 + X_3 \\ &= 3 + 4 + 5 \\ &= 12 \end{aligned}$$

$$\begin{aligned} 2. \sum_{i=1}^3 Y_i &= Y_1 + Y_2 + Y_3 \\ &= 1 + 2 + 3 \\ &= 6 \end{aligned}$$

$$3. \sum_{i=1}^3 X_i^2 = X_1^2 + X_2^2 + X_3^2$$

$$\begin{aligned}
 &= 3^2 + 4^2 + 5^2 \\
 &= 9 + 16 + 25 \\
 &= 50
 \end{aligned}$$

$$\begin{aligned}
 4. \sum_{i=1}^3 Y_i^2 &= Y_1^2 + Y_2^2 + Y_3^2 \\
 &= 1^2 + 2^2 + 3^2 \\
 &= 1 + 4 + 9 \\
 &= 14
 \end{aligned}$$

$$\begin{aligned}
 5. \sum_{i=1}^3 X_i Y_i &= X_1 Y_1 + X_2 Y_2 + X_3 Y_3 \\
 &= 3 \cdot 1 + 4 \cdot 2 + 5 \cdot 3 \\
 &= 3 + 8 + 15 \\
 &= 26
 \end{aligned}$$

$$\begin{aligned}
 6. \left(\sum_{i=1}^3 X_i \right) \left(\sum_{i=1}^3 Y_i \right) &= (X_1 + X_2 + X_3) (Y_1 + Y_2 + Y_3) \\
 &= (3 + 4 + 5) (1 + 2 + 3) \\
 &= 12 \cdot 6 \\
 &= 72
 \end{aligned}$$

ตัวอย่าง 1.2 กำหนด $X_1 = 1$, $X_2 = -2$, $Y_1 = -1$, $Y_2 = 3$ จงคำนวณค่าต่อไปนี้

$$1. \sum_{i=1}^2 X_i^2$$

$$2. \sum_{i=1}^2 X_i Y_i$$

$$3. \left(\sum_{i=1}^2 X_i \right) \left(\sum_{i=1}^2 Y_i \right)$$

$$4. \left(\sum_{i=1}^2 X_i \right) \left(\sum_{i=1}^2 Y_i \right)^2$$

วิธีทำ

$$1. \sum_{i=1}^2 X_i^2 = X_1^2 + X_2^2$$

$$\begin{aligned}
 &= 1^2 + (-2)^2 \\
 &= 1 + 4 \\
 &= 5
 \end{aligned}$$

$$\begin{aligned}
 2. \sum_{i=1}^2 X_i Y_i &= X_1 Y_1 + X_2 Y_2 \\
 &= 1(-1) + (-2)(3) \\
 &= -1 + (-6) \\
 &= -7
 \end{aligned}$$

$$\begin{aligned}
 3. \left(\sum_{i=1}^2 X_i \right) \left(\sum_{i=1}^2 Y_i \right) &= (X_1 + X_2)(Y_1 + Y_2) \\
 &= (1 + (-2))(-1 + 3) \\
 &= (-1)(2) \\
 &= -2
 \end{aligned}$$

$$\begin{aligned}
 4. \left(\sum_{i=1}^2 X_i \right) \left(\sum_{i=1}^2 Y_i \right)^2 &= (X_1 + X_2)(Y_1 + Y_2)^2 \\
 &= (1 + (-2))(-1 + 3)^2 \\
 &= (-1)(2)^2 \\
 &= (-1)(4) \\
 &= -4
 \end{aligned}$$

กฎการใช้เครื่องหมายผลรวม

กฎการใช้เครื่องหมายผลรวมนั้นจะได้กล่าวเป็นสมบัติดังต่อไปนี้ (พิศมัย หาญมงคล พิพัฒน์. 2560 : 8-9)

1. ผลรวมของผลบวกของตัวแปรตั้งแต่ 2 ตัวขึ้นไป จะเท่ากับผลรวมของแต่ละตัวแปรบวกกัน นั่นคือ

$$\sum_{i=1}^N X_i + Y_i + Z_i = \sum_{i=1}^N X_i + \sum_{i=1}^N Y_i + \sum_{i=1}^N Z_i$$

2. ผลรวมของค่าคงที่ จะเท่ากับจำนวนของค่าคงที่คูณกับค่าคงที่ นั่นคือ ถ้า C เป็นค่าคงที่ จะได้

$$\sum_{i=1}^N C = NC$$

3. ผลรวมของค่าคงที่คูณกับตัวแปร จะมีค่าเท่ากับค่าคงที่คูณกับผลรวมของตัวแปร นั่นคือ ถ้า C เป็นค่าคงที่ จะได้

$$\sum_{i=1}^N CX_i = C \sum_{i=1}^N X_i$$

4. ผลรวมของตัวแปรบวกกับค่าคงที่ จะมีค่าเท่ากับผลรวมของตัวแปรบวกกับผลคูณระหว่างจำนวนของค่าคงที่คูณกับค่าคงที่ นั่นคือ ถ้า C เป็นค่าคงที่ จะได้

$$\sum_{i=1}^N X_i + C = \sum_{i=1}^N X_i + NC$$

5. ผลรวมของตัวแปรลบกับค่าคงที่ จะมีค่าเท่ากับผลรวมของตัวแปรลบกับผลคูณระหว่างจำนวนของค่าคงที่คูณกับค่าคงที่ นั่นคือ ถ้า C เป็นค่าคงที่ จะได้

$$\sum_{i=1}^N X_i - C = \sum_{i=1}^N X_i - NC$$

6. ผลรวมของค่าคงที่คูณกับตัวแปร 2 ตัวใด ๆ จะมีค่าเท่ากับค่าคงที่คูณกับผลรวมของผลคูณระหว่างตัวแปร 2 ตัวนั้น นั่นคือ ถ้า C เป็นค่าคงที่ จะได้

$$\sum_{i=1}^N CX_i Y_i = C \sum_{i=1}^N X_i Y_i$$

ตัวอย่าง 1.3 กำหนด $X_1 = 2, X_2 = -3, Y_1 = -1, Y_2 = 4$ จงคำนวณค่าต่อไปนี้

$$1. \sum_{i=1}^2 X_i + 2 + \sum_{i=1}^2 Y_i - 3$$

$$2. \sum_{i=1}^2 2X_i - 3Y_i + 5$$

$$3. \sum_{i=1}^2 X_i^2 - 3X_i Y_i$$

วิธีทำ

$$1. \sum_{i=1}^2 X_i + 2 + \sum_{i=1}^2 Y_i - 3 = \sum_{i=1}^2 X_i + \sum_{i=1}^2 2 + \sum_{i=1}^2 Y_i - \sum_{i=1}^2 3$$

$$\begin{aligned}
&= 2 + -3 + 2 \cdot 2 + -1 + 4 - 3 \cdot 2 \\
&= -1 + 4 + 3 - 6 \\
&= 0
\end{aligned}$$

$$\begin{aligned}
2. \sum_{i=1}^2 2X_i - 3Y_i + 5 &= \sum_{i=1}^2 2X_i - \sum_{i=1}^2 3Y_i + \sum_{i=1}^2 5 \\
&= 2 \sum_{i=1}^2 X_i - 3 \sum_{i=1}^2 Y_i + 5 \cdot 2 \\
&= 2 \cdot 2 + -3 \cdot -3 - 1 + 4 + 5 \cdot 2 \\
&= 2 - 1 - 3 \cdot 3 + 10 \\
&= -2 - 6 + 10 \\
&= 2
\end{aligned}$$

$$\begin{aligned}
3. \sum_{i=1}^2 X_i^2 - 3X_i Y_i &= \sum_{i=1}^2 X_i^2 - \sum_{i=1}^2 3X_i Y_i \\
&= \sum_{i=1}^2 X_i^2 - 3 \sum_{i=1}^2 X_i Y_i \\
&= 2^2 + -3 \cdot 2^2 - 3[2 \cdot -1 + -3 \cdot 4] \\
&= 4 + 9 - 3 \cdot -2 - 12 \\
&= 13 - 3 - 14 \\
&= 13 + 42 \\
&= 55
\end{aligned}$$

สรุป

สถิติสามารถสรุปได้ 2 ความหมาย ความหมายแรก หมายถึง ตัวเลขที่แทนจำนวนหรือข้อเท็จจริงเกี่ยวกับเรื่องต่าง ๆ ที่เราสนใจ เช่น สถิติปริมาณน้ำฝนที่ตกในจังหวัดบุรีรัมย์ สถิติการนำเข้าและส่งออกน้ำมันปาล์ม เป็นต้น ส่วนความหมายที่สองหมายถึง ศาสตร์ที่ว่าด้วยวิธีการที่ใช้ในการศึกษาข้อมูล หรือระเบียบวิธีการทางสถิติ ซึ่งประกอบด้วย การเก็บรวบรวมข้อมูล การนำเสนอข้อมูล การวิเคราะห์ข้อมูล และการแปลความหมายข้อมูล ตามสถิติในแง่ที่เป็นศาสตร์จำแนกสถิติออกเป็น 2 ประเภท คือ สถิติพรรณนาและสถิติเชิงอนุมาน นอกจากนี้ในบทนี้ยังกล่าวถึงข้อมูลและประเภทของข้อมูล โดยจำแนกประเภทของข้อมูลตามแหล่งที่มาของข้อมูล ตามมาตรการวัด และตามตามลักษณะของข้อมูล แล้วในการเก็บรวบรวมข้อมูลจะเป็นขั้นตอนหนึ่งของกระบวนการทางสถิติที่มีความสำคัญ ซึ่งในการเก็บรวบรวมข้อมูลสามารถทำได้ 3 วิธีคือ การเก็บรวบรวมข้อมูลจากงานระบบทะเบียนหรือการบันทึก การเก็บรวบรวมข้อมูลจากการสำรวจ และการเก็บรวบรวมข้อมูลจากการทดลอง