บทที่ 10
สถิติเชิงพรรณนา

ในการวิจัยใด ๆ ข้อมูลที่ได้จากการเก็บรวบรวมจะได้ข้อมูลทั้งในลักษณะของข้อมูลเชิงปริมาณที่เป็นตัวเลข และข้อมูลเชิงคุณภาพ ที่เป็นการบรรยายผลของระดับพฤติกรรม/คุณลักษณะที่เกิดขึ้น โดยที่ในส่วนของข้อมูลเชิงปริมาณนั้นผู้วิจัยจะต้องนำข้อมูลมาดำเนินการด้วยวิธีการทางสถิติเพื่อวิเคราะห์ข้อมูล ที่จะได้สรุปผลตามวัตถุประสงค์และสมมุติฐานของการวิจัยที่กำหนดไว้

สถิติ

1. ความหมายของสถิติ

สถิติ(Statistics) เป็นคำที่มาจากภาษาเยอรมันว่า “Statistik”ที่มีรากศัพท์ว่า “State” ที่หมายถึง ข้อมูลที่เป็นตัวเลขและข้อเท็จจริงที่มีประโยชน์ต่อการบริหารงานของรัฐ(Fact of State)

นักวิชาการได้ให้ความหมายของสถิติ จำแนกได้ ดังนี้ (ศิริชัย กาญจนวาสี,ทวีวัฒน์ ปิตยานนท์ และ ดิเรก ศรีสุโข.2537 : 37; นงลักษณ์ วิรัชชัย,2543 : 212)

1.1 สถิติ ในความหมายของ “ข้อมูลสถิติ”หมายถึง ตัวเลขที่เป็นข้อเท็จจริงของประเด็นที่สนใจ อาทิ สถิติจำนวนครอบครัวในชุมชนแห่งหนึ่ง สถิติปริมาณน้ำฝนในจังหวัดอุดรธานี สถิติการประสบอุบัติเหตุทางบกของประเทศไทย ฯลฯ

1.2. สถิติ ในความหมายของ “สถิติศาสตร์”หมายถึง ศาสตร์ที่เกี่ยวกับวิธีการที่ใช้ในการศึกษาข้อมูลที่เรียกว่า “ระเบียบวิธีการทางสถิติ” ที่ประกอบด้วย 4 วิธีการ คือ การเก็บรวบรวมข้อมูล การนำเสนอข้อมูล การวิเคราะห์ข้อมูล และการแปลความหมายข้อมูล

1.3 สถิติ ในความหมายของ “ค่าสถิติ”หมายถึง ค่าตัวเลขที่คำนวณได้จากข้อมูลของกลุ่มตัวอย่าง โดยใช้สัญลักษณ์ $\overline{X}$ แทนค่าเฉลี่ย(Mean), S.D.แทนส่วนเบี่ยงเบนมาตรฐาน และ r_{xy}แทนค่าสัมประสิทธิ์สหสัมพันธ์ เป็นต้น

1.4 สถิติ ในความหมายของ “วิชาสถิติ”หมายถึง วิชาวิทยาศาสตร์แขนงหนึ่งที่มีเนื้อหาสาระและรากฐานจากวิชาคณิตศาสตร์ และตรรกศาสตร์ ที่จะนำไปใช้ในการตัดสินใจภายใต้สถานการณ์ที่ไม่แน่นอน

ศิริชัย กาญจนวาสี,ทวีวัฒน์ ปิตยานนท์ และ ดิเรก ศรีสุโข(2537 : 38) ได้นำเสนอผลการเปรียบเทียบความแตกต่างบางประการของความหมายสถิติในฐานะข้อมูลกับวิธีการ ดังแสดงใน**ตารางที่ 10.1**(ศิริชัย กาญจนวาสี,ทวีวัฒน์ ปิตยานนท์ และ ดิเรก ศรีสุโข,2537 : 38)

ตารางที่ 10.1 ความแตกต่างของความหมายสถิติในฐานะข้อมูลกับวิธีการ

สถิติในฐานะข้อมูล	สถิติในฐานะวิธีการ
1. เป็นปริมาณ 2. อยู่ในลักษณะของข้อมูลดิบ 3. เป็นการบรรยาย 4. เป็นวัตถุดิบที่ป้อนเข้าสู่กระบวนการ เมื่อวิเคราะห์แล้วนำมาช่วยพิจารณาตัดสินใจ 5. ไม่มีความหมายถ้าขาดเครื่องมือวิเคราะห์ 6. ธรรมชาติของข้อมูลเป็นตัวกำหนดวิธีการ	1. เป็นวิธีการปฏิบัติ 2. กระบวนการจัดกระทำข้อมูลดิบ 3. เป็นเครื่องมือในการวิเคราะห์ 4. เป็นระเบียบวิธีทางวิทยาศาสตร์ของการวิเคราะห์และตีความหมายข้อมูล 5. เป็นเครื่องมือที่ไร้ความหมายถ้าขาดข้อมูล 6. เครื่องมือจะเป็นตัวกำหนดวิธีการเก็บรวบรวมข้อมูล

2. ประเภทของสถิติ

ประเภทของสถิติที่ใช้ในการวิจัย สามารถจำแนกตามเกณฑ์ที่ใช้ในการพิจารณา ดังนี้ (ศิริชัย กาญจนวาสี,ทวีวัฒน์ ปิตยานนท์ และ ดิเรก ศรีสุโข.2537 : 39-40 ; นงลักษณ์ วิรัชชัย,2543 : 213)

2.1จำแนกตามบทบาทและหน้าที่

2.1.1 สถิติเชิงบรรยาย(Descriptive Statistics) เป็นสถิติที่มุ่งนำเสนอสารสนเทศเพื่อใช้บรรยายสรุปลักษณะของตัวแปรในกลุ่มตัวอย่างหรือประชากรว่าเป็นอย่างไร มีสถิติที่ใช้ดังนี้

2.1.1.1 การแจกแจงความถี่ และการนำเสนอด้วยตาราง กราฟ และแผนภูมิ

2.1.1.2 การจัดตำแหน่งและเปรียบเทียบ อาทิ สัดส่วน ร้อยละ หรือเปอร์เซนไทล์

2.1.1.3 การวัดแนวโน้มเข้าสู่ส่วนกลาง

2.1.1.4 การวัดการกระจายของข้อมูล

2.1.1.5 การวัดความสัมพันธ์ระหว่างตัวแปร

2.1.2 สถิติเชิงอ้างอิง (Inferential Statistics) เป็นเทคนิคทางสถิติที่ศึกษาข้อมูลจากกลุ่มตัวอย่างหรือค่าสถิติเพื่อใช้สรุปอ้างอิงข้อมูลไปสู่ประชากรหรือค่าพารามิเตอร์ แต่จะต้องมีวิธีการได้กลุ่มตัวอย่างที่เป็นตัวแทนที่ดีของประชากร ที่มีความสอดคล้องกับหลักการอ้างอิงที่มีประสิทธิภาพจากกลุ่มตัวอย่างสู่ประชากร มีวิธีการทางสถิติเชิงอ้างอิง ดังนี้

2.1.2.1 การประมาณค่าพารามิเตอร์ (Estimation) เป็นเทคนิคทางสถิติในการคำนวณค่าสถิติของกลุ่มตัวอย่างไปคาดคะเนค่าพารามิเตอร์ของประชากรที่สามารถดำเนินการได้ 2 ลักษณะ คือ การประมาณค่าเป็นจุด(Point Estimation) และการประมาณค่าเป็นช่วง(Interval Estimation)

2.1.2.2 การทดสอบสมมุติฐาน(Hypothesis testing)เป็นเทคนิคทางสถิติที่นำค่าสถิติของกลุ่มตัวอย่างไปทดสอบสมมุติฐานทางสถิติเกี่ยวกับค่าพารามิเตอร์ของประชากร

2.2 จำแนกตามลักษณะของข้อตกลงเบื้องต้นและวิธีการทดสอบทางสถิติ มีดังนี้

2.2.1 สถิติพาราเมตริก(Parametric Statistics) เป็นเทคนิคทางสถิติที่อ้างอิงจากค่าสถิติของกลุ่มตัวอย่างไปสู่ค่าพารามิเตอร์ของประชากร โดยมีข้อตกลงเบื้องต้นและวิธีการทดสอบ ดังนี้

2.2.1.1 วิธีการทดสอบทางสถิติ มีดังนี้

1) ใช้การทดสอบค่าที(t-test) สำหรับการทดสอบความแตกต่างระหว่างค่าเฉลี่ยของประชากร 2 กลุ่ม

2) ใช้การทดสอบค่าเอฟ(F-test) สำหรับการทดสอบความแปรปรวนระหว่างประชากรตั้งแต่ 3 กลุ่มขึ้นไป

3) ใช้การทดสอบไคร์สแควร์(Chi-square) สำหรับการทดสอบความแตกต่างระหว่างสัดส่วนของประชากร

2.2.1.2 ข้อตกลงเบื้องต้นในการใช้สถิติพาราเมตริก มีดังนี้

1) ค่าพารามิเตอร์ของประชากรทั่วไปจะต้องมีลักษณะเฉพาะและกำหนดเป็นค่าในประชากรได้

2) ลักษณะการแจกแจงตัวแปรของประชากร จะต้องมีลักษณะการแจกแจงแบบใดแบบหนึ่ง อาทิ ประชากรจะต้องมีจำนวนมากและมีการแจกแจงแบบปกติ เป็นต้น

2.2.2 สถิตินอนพาราเมตริก(Non-Parametric Statistics) เป็นเทคนิคทางสถิติที่ไม่มีการระบุค่าพารามิเตอร์ของประชากร และไม่จำเป็นต้องมีข้อตกลงเบื้องต้นเกี่ยวกับลักษณะของการแจกแจงของประชากร อาทิ การทดสอบไคร์สแควร์ในการทดสอบGoodness of Fit หรือ การทดสอบวิลคอกวอล-แมน-วิทนีย์(Wilcoxon-Mann-Whitney)สำหรับการทดสอบตำแหน่งเฉลี่ยของประชากรที่เป็นอิสระจากกัน 2 กลุ่ม เป็นต้น

2.3 จำแนกตามจำนวนของตัวแปร มีดังนี้

2.3.1 สถิติเอกนาม(Univariate Statistics) เป็นสถิติที่เกี่ยวกับตัวแปรเพียงตัวเดียว

2.3.2 สถิติทวินาม(Bivariate Statistics) เป็นสถิติที่เกี่ยวกับตัวแปร 2 ตัว

2.3.3 สถิติพหุคูณ(Multivariate Statistics) เป็นสถิติที่เกี่ยวกับตัวแปรมากกว่า 2 ตัว

3. สิ่งที่ควรคำนึงในการใช้สถิติ

ล้วน สายยศ และอังคณา สายยศ(2538 : 222) ได้ระบุสิ่งที่ควรคำนึงในการใช้สถิติมีดังนี้

3.1 สถิติเป็นเพียงวิธีการของการวิจัย ไม่ใช่จุดมุ่งหมายปลายทางของการวิจัย

3.2 สถิติยังไม่สามารถจะใช้ได้ ถ้าไม่มีการกำหนดสมมุติฐานและกำหนดขอบเขตของข้อมูลให้ชัดเจนก่อน

3.3 ข้อสรุปที่ได้จากการสุ่มตัวอย่างไม่ใช่สิ่งที่ยืนยันแน่นอน เพราะไม่แน่ว่าเป็นตัวอย่างที่แท้จริงหรือไม่ ดังนั้นจะต้องคำนึงถึงขอบเขตของความเชื่อมั่นที่กำหนด

3.4 การกระทำเกี่ยวกับข้อมูลจะต้องใช้การตรวจสอบอย่างละเอียด และแน่ใจว่าไม่ผิดพลาด อย่างน้อยต้องตรวจสอบ 2 ครั้ง

3.5 ความผิดพลาดอย่างหนึ่งมักจะเกิดขึ้นเสมอ เนื่องจากความพลั้งเผลอของมนุษย์

3.6 นักสถิติและนักโกหกมีอะไรที่ใกล้เคียงกันอยู่มาก ดังนั้นการกระทำทางสถิติต้องระมัดระวังตลอดเวลา การสรุปหรือแปลผลผิดพลาดอาจจะได้จากการกำหนดสมมุติฐานที่คลาดเคลื่อนหรือได้กลุ่มตัวอย่างจากการสุ่มที่ลำเอียง

บอร์กและกอล(Brog and Gall,1971 : 145) ได้ระบุสิ่งที่ควรคำนึงในการใช้สถิติ ดังนี้

1) เลือกใช้สถิติที่เหมาะสม หรือถูกต้องในการวิเคราะห์ข้อมูล

2) พิจารณาข้อมูลให้ชัดเจน แล้วจึงนำมาพิจารณาเลือกใช้สถิติเพื่อวิเคราะห์ข้อมูล

3) ใช้สถิติเพียงวิธีการเดียวทั้ง ๆ ที่ข้อมูลนั้นอาจใช้ได้หลายวิธีการ

4) ระมัดระวังการใช้สถิติที่ฝ่าฝืนข้อตกลงเบื้องต้นของสถิติแต่ละประเภท

5) นำเสนอความสำคัญของข้อมูลที่มีความแตกต่างกันน้อยมากที่เกิดความเป็นจริง เพื่อให้มีนัยสำคัญทาสถิติ

6) ไม่จำเป็นต้องวิเคราะห์ความสัมพันธ์ ถ้าความสัมพันธ์นั้นไม่อยู่ในรูปความสัมพันธ์แบบโมเมนต์ของผลคูณ

7) หาสัมประสิทธิ์สหสัมพันธ์ที่ไม่สอดคล้องกับลักษณะของข้อมูล

8) ใช้ตารางทดสอบนัยสำคัญของสหสัมพันธ์ผิดประเภท

โดยที่เนื้อหาสาระที่กล่าวถึงในบทนี้ จะเป็นสถิติเชิงพรรณนาที่ใช้ผลการวิเคราะห์ข้อมูลเพื่อบรรยายลักษณะของข้อมูลว่าเป็นอย่างไร และในส่วนสถิติเชิงอ้างอิงจะนำเสนอในบทต่อไป

ข้อมูล

1. ความหมายของข้อมูล

ข้อมูล (Data) เป็นข้อเท็จจริงที่ได้จากการเก็บรวบรวมคุณลักษณะที่ต้องการจากกลุ่มเป้าหมายตามจุดประสงค์ที่กำหนดไว้ ทั้งในลักษณะข้อมูลเชิงปริมาณ(คะแนนที่ได้)และเชิงคุณภาพ(พฤติกรรมหรือปรากฏการณ์ที่เกิดขึ้น)

ข้อมูล(Datum) ในความหมายที่เป็นเอกพจน์ มาจากคำว่า “Datus”ในภาษาลาติน หมายถึงสิ่งที่ใช้เป็นฐานในการให้เหตุผลหรือการสรุปอ้างอิง แต่ข้อมูล(Data)ในความหมายที่เป็นพหูพจน์ที่หมายถึง ผลของการรวบรวมสารสนเทศที่เป็นข้อเท็จจริงหรือวัตถุที่ใช้เป็นฐานสำหรับการให้เหตุผล การอภิปราย หรือการคำนวณ(Webster’s Ninth New Collegiate Dictionary,1991)

ข้อมูล หมายถึง ข้อเท็จจริง สารสนเทศความรู้ ความคิด ความรู้สึก พฤติกรรมการแสดงออก การกระทำ หลักฐานหรือคุณสมบัติของบุคคลหรือสิ่งของที่เก็บรวบรวมมาใช้ประโยชน์ในการตอบปัญหาการวิจัย(Kerlinger,1982 อ้างอิงใน นงลักษณ์ วิรัชชัย,2543 : 208)

2.ระดับของข้อมูล

ในการวัดผลได้กำหนดข้อมูลเชิงปริมาณที่เป็นตัวเลขในการนำมาวิเคราะห์เพื่อสรุปผลการวิจัยที่จำแนกได้ 4 ระดับตามแนวคิดของสตีเวนส์ มีดังนี้(Babbie,1992 : 140-144; Nachmias and Nachmias,1993 : 151)

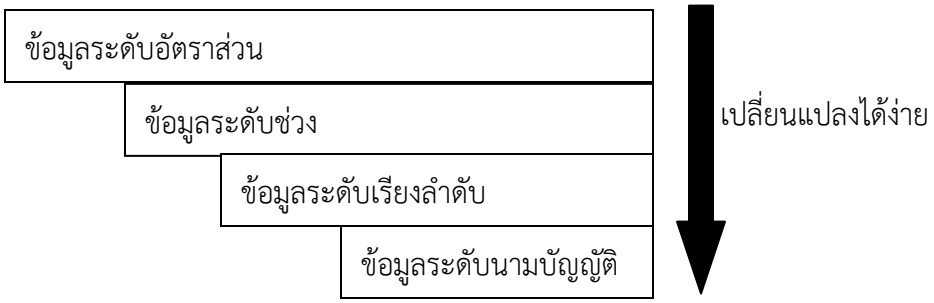
2.1 ข้อมูลระดับนามบัญญัติ(Nominal Scale) เป็นข้อมูลที่ใช้กำหนดชื่อ คุณลักษณะ/พฤติกรรมที่ได้จากการวัด/เก็บรวบรวมข้อมูล อาทิ ชื่อ-นามสกุล หมายเลขประจำตัว ศาสนา ฯลฯ หรือตัวเลข ที่ใช้จำแนกหรือจัดกลุ่มของสิ่งของ/บุคคลเท่านั้น แต่จะไม่มีความหมายในเชิงปริมาณที่จะนำไปดำเนินการด้วยวิธีการทางคณิตศาสตร์ได้ (บวก ลบ คูณ และหาร)หรือไม่สามารถนำไปเปรียบเทียบกันได้ อาทิ ให้เพศชาย แทนด้วย “1” และเพศหญิงแทนด้วย “2” ซึ่งในกรณีนี้ “2” ไม่ได้มีค่ามากกว่า “1” ในการเปรียบเทียบเชิงปริมาณ เป็นต้น ซึ่งบุญธรรม กิจปรีดาบริสุทธิ์(2534 :4-5) ได้สรุปลักษณะของข้อมูลในระดับนามบัญญัติ มีดังนี้ 1) อยู่อย่างไม่เป็นระเบียบ หรือแยกกันอยู่ 2) ข้อมูลแต่ละกลุ่มไม่มีความสัมพันธ์กัน 3) จะใช้สัญลักษณ์/ตัวเลขแทนกลุ่มไม่ได้มีความหมายในเชิงปริมาณ และ 4) สมาชิกในกลุ่มเดียวกันจะมีคุณลักษณะเหมือนกัน โดยที่ค่าสถิติที่นำมาใช้กับข้อมูลประเภทนี้ได้แก่ ความถี่ ร้อยละ หรือฐานนิยม เป็นต้น

2.2 ข้อมูลระดับเรียงลำดับ(Ordinal Scale) เป็นข้อมูลที่สามารถนำมาเรียงลำดับจากมากไปหาน้อยหรือจากน้อยไปหามากตามเกณฑ์(ความเข้มข้น ความหนาแน่น ความแข็งแรง ฯลฯ) แต่จะไม่สามารถระบุได้ว่าแต่ละลำดับแตกต่างกันเท่าไร และความแตกต่างนั้นเท่ากันหรือไม่ และยังไม่สามารถนำมาดำเนินการด้วยวิธีการทางคณิตศาสตร์ได้(บวก ลบ คูณและหาร) อาทิ การตรวจผลงานของผู้เรียนที่ครูผู้สอนสามารถเรียงลำดับของผลงานตามเกณฑ์ที่กำหนดได้ตั้งแต่ลำดับที่ 1,2,3 ถึง n แต่จะไม่สามารถจะระบุช่วงความแตกต่างระหว่างลำดับที่ 1 กับ 2 หรือ 2 กับ 3 และความแตกต่างระหว่างลำดับที่เท่ากันหรือไม่ หรือประสิทธิภาพของการปฏิบัติที่จำแนกเป็น ดี-ปานกลาง-เลว เป็นต้น ซึ่งบุญธรรม กิจปรีดาบริสุทธิ์(2534 :5-6) ได้สรุปลักษณะของข้อมูลในระดับเรียงลำดับ มีดังนี้ 1) ใช้กฎเกณฑ์ที่เป็นลักษณะเฉพาะของตัวแปร 2) ไม่สามารถเปรียบเทียบความแตกต่างของปริมาณระหว่างลำดับที่(เนื่องจากไม่ทราบความแตกต่าง) และ3) ใช้สัญลักษณ์/ตัวเลขที่เรียงลำดับกันแทนลำดับที่ที่ต้องการเรียงลำดับ โดยที่ค่าสถิติที่ใช้กับข้อมูลประเภทนี้ ได้แก่ ความถี่ ร้อยละ ฐานนิยม มัธยฐาน และเปอร์เซนไทล์ เป็นต้น

2.3 ข้อมูลระดับช่วง หรืออันตรภาค(Interval Scale) เป็นข้อมูลที่สามารถนำมาเรียงลำดับได้ และระบุความแตกต่างระหว่างลำดับที่ได้ และความแตกต่างในแต่ละช่วง(หน่วยการวัด)มีค่าเท่ากันทุก ๆ ช่วงจึงสามารถเปรียบเทียบความแตกต่างกันได้ และเป็นข้อมูลที่สามารถนำมาบวกและลบกันได้ แต่จะยังไม่สามารถนำมาคูณและหาร หรือเปรียบเทียบเป็นจำนวนเท่า เนื่องจากไม่มีระดับศูนย์แท้ (Absolute Zero) (ศูนย์แท้ หมายถึง การไม่มี/ปราศจาก/ว่างเปล่า) หรือไม่มีจุดเริ่มต้นที่แน่นอน อาทิ การที่ผู้เรียนสอบได้ 0 คะแนน ไม่ได้หมายความว่าผู้เรียนไม่มีความรู้ในวิชานั้น แต่หมายถึงผู้เรียนทำแบบทดสอบฉบับนั้นไม่ได้ และผู้เรียนที่ได้คะแนน 50 คะแนน ก็ไม่ได้หมายความว่าผู้เรียนคนนั้นมีความรู้เป็นสองเท่าของผู้เรียนที่ได้คะแนน 25 คะแนน หรือ วันนี้มีอุณหภูมิ 0 องศาเซนเซียส ไม่ได้หมายความว่าวันนี้ไม่มีอุณหภูมิ แต่หมายถึง อุณหภูมิถึงระดับ 0 องศาเซนเซียส ตามระดับมาตรฐานที่กำหนดเท่านั้น และอุณหภูมิ 60 องศาเซนเซียสก็ไม่ได้หมายความว่าจะมีความร้อนเป็นสองเท่าของอุณหภูมิ 30 องศาเซนเซียส โดยค่าสถิติที่ใช้กับข้อมูลระดับนี้ ได้แก่ การแจกแจงความถี่ ร้อยละ ฐานนิยม มัธยฐาน เปอร์เซนไทล์ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน เป็นต้น

2.4 ข้อมูลระดับอัตราส่วน(Ratio Scale) เป็นข้อมูลที่มีคุณสมบัติทางกายภาพสมบูรณ์ที่สุด โดยที่ข้อมูลสามารถระบุความแตกต่างระหว่างกลุ่ม/ลำดับที่ หรือนำมาเปรียบเทียบความมากหรือน้อย มีค่าและช่วงเท่ากัน และมีศูนย์แท้ แล้วสามารถนำข้อมูลมาดำเนินการด้วยวิธีการทางคณิตศาสตร์ บวก ลบ คูณ และหาร ได้ทุกวิธี โดยค่าสถิติที่ใช้กับข้อมูลระดับนี้จะใช้เหมือนกับข้อมูลระดับช่วงที่เป็นข้อมูลระดับที่ 3

โดยที่ข้อมูลทั้ง 4 ระดับนี้ จะสามารถที่เปลี่ยนแปลงระดับข้อมูลได้ โดยเฉพาะในระดับที่สูงกว่าไปสู่ระดับที่ต่ำกว่า คือจากระดับอัตราส่วนไปสู่ระดับนามบัญญัติ แต่จะมีความยุ่งยากในกรณีที่จะปรับเปลี่ยนข้อมูลจากระดับต่ำกว่าไปสู่ระดับที่สูงกว่า โดยแสดงความสัมพันธ์ระหว่างข้อมูลในระดับต่าง ๆ ว่าข้อมูลในระดับที่สูงกว่าจะมีคุณสมบัติของข้อมูลในระดับที่ต่ำกว่าด้วย ดังแสดงใน**ภาพที่ 10.1**



ภาพที่ 10.1 ความสัมพันธ์ของระดับข้อมูล

3. ประเภทของข้อมูล

ในการพิจารณาจำแนกประเภทของข้อมูลตามเกณฑ์ที่ใช้ มีดังนี้(นงลักษณ์ วิรัชชัย,2543 : 165-167)

3.1 จำแนกตามศาสตร์ มีดังนี้

3.1.1 ข้อมูลทางประวัติศาสตร์และมานุษยวิทยา เป็นข้อมูล/หลักฐานที่เกี่ยวกับผลผลิตหรือผลงานของมนุษย์ที่สามารถตรวจสอบได้ อาทิ สิ่งพิมพ์ หรือวัตถุโบราณ เป็นต้น

3.1.2 ข้อมูลทางวิทยาศาสตร์ จำแนกได้ดังนี้

3.1.2.1 ข้อมูลทางวิทยาศาสตร์กายภาพ เป็นข้อมูลที่ได้จากลักษณะทางกายภาพของวัตถุ และปรากฏการณ์ทางธรรมชาติ อาทิ ก้อนหิน แร่ธาตุ ปริมาณน้ำฝน และการระเบิดของภูเขาไฟ เป็นต้น

3.1.2.2 ข้อมูลทางวิทยาศาสตร์ชีวภาพ เป็นข้อมูลที่ได้จากลักษณะทางกายภาพและการทำหน้าที่ของอวัยวะต่าง ๆ ของสิ่งมีชีวิต อาทิ วัฏจักรชีวิตของแมลง การสูบฉีดโลหิตของร่างกายมนุษย์ เป็นต้น

3.1.3 ข้อมูลทางสังคมศาสตร์และพฤติกรรมศาสตร์ เป็นข้อมูลที่เป็นลักษณะทางจิตและการกระทำของสิ่งมีชีวิต โดยเฉพาะมนุษย์ที่มีพฤติกรรมที่จำแนกได้ดังนี้

3.1.3.1 พฤติกรรมภายใน เป็นพฤติกรรมที่เกิดภายในมนุษย์ที่ไม่สามารถวัดหรือสังเกตได้โดยตรง อาทิ สติปัญญา ความรู้สึก การรับรู้และการตัดสินใจ เป็นต้น

3.1.3.2 พฤติกรรมภายนอก เป็นพฤติกรรมที่มนุษย์แสดงออกแล้วสามารถวัดหรือสังเกตได้โดยใช้ประสาทสัมผัส ที่จำแนกได้ดังนี้

1) พฤติกรรมทางวาจา ที่แสดงโดยการบอกเล่าเป็นคำพูด สามารถสังเกตได้ด้วยการดู และการฟัง

2) พฤติกรรมทางกาย ที่แสดงในรูปอากัปกิริยา ท่าทางและการเปลี่ยนแปลงทางร่างกาย จำแนกได้ดังนี้

(1) พฤติกรรมโมลาร์ ที่เป็นพฤติกรรมทางกายที่สามารถสังเกตได้ง่าย อาทิ การร้องไห้ การวิ่ง การเดิน เป็นต้น

(2) พฤติกรรมโมเลคูลาร์ ที่เป็นพฤติกรรมทางกายที่ไม่สามารถสังเกตได้โดยตรงจะต้องใช้เครื่องมือในการช่วยสังเกต อาทิ การเต้นของหัวใจ ความดันโลหิต เป็นต้น

3.2 จำแนกตามแหล่งที่มาของข้อมูล มีดังนี้(ธีระวุฒิ เอกะกุล. 2544 :193-194)

3.2.1 ข้อมูลปฐมภูมิ(Primary Data)เป็นข้อมูลที่ได้จากแหล่งกำเนิดของข้อมูลหรือปรากฏการณ์โดยตรง ที่จะได้ข้อมูลที่มีรายละเอียดที่ชัดเจน มีความเที่ยงตรงและมีความเชื่อมั่นสูงที่สอดคล้องกับวัตถุประสงค์ของการนำข้อมูลมาใช้ แต่จะต้องใช้เวลา แรงงานและงบประมาณที่ค่อนข้างมาก ที่มีข้อดี-ข้อจำกัดของข้อมูลปฐมภูมิ ดังนี้

3.2.1.1 ข้อดีของข้อมูลปฐมภูมิ มีดังนี้

1) เป็นข้อมูลที่เก็บรวบรวมที่สอดคล้องกับความต้องการใช้ข้อมูลอย่างแท้จริง

2) มีความสอดคล้องกับวัตถุประสงค์ และเทคนิคการวิเคราะห์ข้อมูล

3) สามารถควบคุมคุณภาพในการเก็บรวบรวมข้อมูล ได้สอดคล้องกับความรู้ความสามารถของผู้วิจัย

3.2.1.2 ข้อจำกัดของข้อมูลปฐมภูมิ มีดังนี้

1) ใช้เวลาค่อนข้างนานในการเก็บรวบรวมข้อมูล

2) ใช้งบประมาณในการเก็บรวบรวมข้อมูลมากกว่าการใช้ข้อมูลทุติยภูมิที่มีการเก็บรวบรวมไว้แล้ว

3) ถ้าผู้วิจัยขาดความชำนาญหรือประสบการณ์ในการเก็บรวบรวมข้อมูลจะทำให้ได้ข้อมูลไม่ครบถ้วน มีความคลาดเคลื่อนจนขาดความน่าเชื่อถือ และไม่สามารถตอบสนองวัตถุประสงค์ของการวิจัยได้อย่างแท้จริง

3.2.2 ข้อมูลทุติยภูมิ(Secondary Data) เป็นข้อมูลที่ได้จากแหล่งที่ไม่ใช่ต้นกำเนิดของข้อมูลหรือปรากฏการณ์โดยตรง แต่จะได้จากแหล่งข้อมูลที่ได้มีการเก็บรวบรวมไว้แล้วเพื่อนำมาวิเคราะห์ใหม่ตามวัตถุประสงค์ ที่อาจจะได้ข้อมูลไม่ครบถ้วนตามที่ต้องการ หรือมีความเชื่อมั่นและความเที่ยงตรงที่น้อยกว่าข้อมูลปฐมภูมิ แต่จะใช้เวลา แรงงานและงบประมาณค่อนข้างน้อย ที่มีข้อดี-ข้อจำกัดของข้อมูลทุติยภูมิ มีดังนี้

3.2.2.1 ข้อดีของข้อมูลทุติยภูมิ มีดังนี้

1) ไม่ต้องเก็บรวบรวมข้อมูลที่จะต้องใช้เวลา งบประมาณและแรงงานค่อนข้างมาก

2) สามารถศึกษาเหตุการณ์ในประวัติศาสตร์ ทำให้ทราบการเปลี่ยนแปลงหรือแนวโน้มการเปลี่ยนแปลงของเหตุการณ์ตั้งแต่อดีตจนกระทั่งถึงปัจจุบัน

3.2.2.2 ข้อเสียของข้อมูลทุติยภูมิ มีดังนี้

1) ได้ข้อมูลที่อาจไม่สอดคล้องกับวัตถุประสงค์ของการวิจัยที่ต้องการอย่างแท้จริง

2) ความน่าเชื่อถือของข้อมูลทุติยภูมิมีมากหรือน้อยเพียงใด โดยพิจารณาจากวัตถุประสงค์ในการเก็บรวบรวมข้อมูลนั้น ๆ ความครบถ้วน และจะต้องดำเนินการอย่างไรเพื่อให้การนำข้อมูลมาใช้นั้นสามารถนำมาใช้ได้อย่างมีประสิทธิภาพ

3) เป็นข้อมูลที่อาจจะไม่ทันสมัย หรือสอดคล้องกับการนำมาใช้ในสถานการณ์ปัจจุบันที่มีการเปลี่ยนแปลงค่อนข้างรวดเร็ว อาทิ รายงานประจำปี เป็นต้น

4) เป็นข้อมูลที่จะต้องใช้ความรอบคอบในการพิจารณาในการให้ความหมายของประเด็นหรือข้อมูลที่ศึกษาว่าเป็นอย่างไร และแต่ละรายงานเหมือนกันและแตกต่างกันอย่างไร

3.3 จำแนกตามสภาพที่เป็นจริง

3.3.1 ข้อเท็จจริง(Fact) เป็นข้อมูลที่เป็นข้อเท็จจริงที่มีความคงที่ไม่เปลี่ยนแปลงตามเวลา และสภาพแวดล้อม ที่ได้จากบุคคล อาทิ วัน-เดือน-ปีเกิด หรือลักษณะทางกายภาพของสิ่งของ อาทิ น้ำหนัก ความกว้าง ความยาว เป็นต้น

3.3.2 เจตคติ(Attitude) เป็นข้อมูลเกี่ยวกับพฤติกรรมทางจิตใจในการตอบสนองเมื่อได้รับสิ่งเร้า และเป็นข้อมูลที่มีความคงที่ แต่อาจเปลี่ยนแปลงตามสถานการณ์หรือสิ่งเร้า อาทิ ความพึงพอใจ หรือแรงจูงใจใฝ่สัมฤทธิ์ เป็นต้น

3.3.3 ความคิดเห็น หรือความรู้สึก(Opinion or Feeling) เป็นข้อมูลของผลการรับรู้หรือปฏิกิริยาตอบสนองของบุคคลที่มีต่อสิ่งเร้า หรือสถานการณ์ และเป็นข้อมูลที่มีความคงที่น้อยที่สุดที่จะเปลี่ยนแปลงได้ตามเวลาที่เปลี่ยนแปลงแม้ว่าสิ่งเร้าหรือสถานการณ์ยังคงที่ก็ตาม

3.4 จำแนกตามคุณสมบัติของข้อมูล มีดังนี้

3.4.1 ข้อมูลปรนัย(Objective Data) เป็นข้อมูลที่ได้จากการเก็บรวบรวมโดยตรงที่ไม่ผ่านกระบวนการที่ทำให้ข้อมูลเปลี่ยนแปลงคุณสมบัติ ที่มีความเที่ยงตรงและความเชื่อมั่นสูง อาทิ การวัดลักษณะกายภาพของวัตถุ การวัดความสูงของมนุษย์ เป็นต้น

3.4.2 ข้อมูลอัตนัย(Subjective Data) เป็นข้อมูลที่ได้ผ่านกระบวนการตีความหมายของผู้เก็บรวบรวมข้อมูลแล้ว เนื่องจากกระบวนการในการเก็บรวบรวมข้อมูลขาดความชัดเจนของการใช้เครื่องมือและวิธีการ อาทิ ความพึงพอใจที่แตกต่างกันระหว่างผู้ให้ข้อมูล 2 คนที่มีระดับความพึงพอใจระดับเดียวกัน เป็นต้น

3.5 จำแนกตามลักษณะของข้อมูล จำแนกได้ดังนี้

3.5.1 ข้อมูลเชิงปริมาณ (Quantitative Data) เป็นข้อมูลที่ผู้วิจัยได้จากการสังเกตหรือวัดเป็นตัวเลข อาทิ ผลสัมฤทธิ์ทางการเรียน, ความถนัดทางการเรียนหรือคุณลักษณะด้านจิตพิสัยที่วัดผลเป็นคะแนน เป็นต้น

3.5.2 ข้อมูลเชิงคุณลักษณะ (Qualitative Data) เป็นข้อมูลที่แสดงคุณลักษณะของประเด็นที่ต้องการ อาทิ ข้อความที่เป็นความคิดเห็น หรือผลจากการสังเกตที่บรรยายรายละเอียด เป็นต้น

3.6 จำแนกตามสภาพของข้อมูลที่เกี่ยวข้องกับกลุ่มตัวอย่าง แบ่งได้ 3 ประเภท ดังนี้

3.6.1 ข้อมูลส่วนบุคคล (Personal Data) เป็นข้อมูลที่เกี่ยวกับข้อเท็จจริงส่วนบุคคลของกลุ่มตัวอย่าง อาทิ อายุ เพศ อาชีพ ระดับการศึกษา หรือการนับถือศาสนา เป็นต้น

3.6.2 ข้อมูลสิ่งแวดล้อม (Environmental Data) เป็นข้อมูลที่เกี่ยวกับข้อเท็จจริงของสิ่งแวดล้อมของกลุ่มตัวอย่าง อาทิ ลักษณะของภูมิอากาศ/ภูมิประเทศในท้องถิ่นของกลุ่มตัวอย่าง เป็นต้น

3.6.3 ข้อมูลพฤติกรรม (Behavioral Data) เป็นข้อมูลที่เกี่ยวกับข้อเท็จจริงของพฤติกรรมที่มีอยู่ในกลุ่มตัวอย่าง อาจจำแนกได้เป็น 3 ประเภท ดังนี้

3.6.3.1 ข้อมูลพุทธิพิสัย (Cognitive Domain) เป็นข้อมูลที่เกี่ยวกับพฤติกรรมทางด้านความสามารถของสมอง อาทิ ผลสัมฤทธิ์ทางการเรียนรู้ หรือความถนัดหรือสติปัญญา เป็นต้น

3.6.3.2 ข้อมูลจิตพิสัย (Affective Domain) เป็นข้อมูลที่เกี่ยวกับพฤติกรรมทางด้านจิตใจ อาทิ ความสนใจ ความวิตกกังวล แรงจูงใจใฝ่สัมฤทธิ์ หรือเจตคติ เป็นต้น

3.6.3.3 ข้อมูลทักษะพิสัย (Psychomotor Domain) เป็นข้อมูลที่เกี่ยวกับพฤติกรรมด้านทักษะทางกาย อาทิ การเคลื่อนไหว หรือการปฏิบัติงาน เป็นต้น

การวิเคราะห์ข้อมูล

1. ความหมายของการวิเคราะห์ข้อมูล

การวิเคราะห์(Analysis) มาจาก “Ana+Lyein” หรือ Analyeinในภาษากรีก หมายถึง การจำแนกสิ่งที่ต้องการศึกษาออกเป็นส่วนย่อย ๆ เพื่อศึกษาลักษณะธรรมชาติที่แท้จริงของแต่ละส่วนย่อย และศึกษาความสัมพันธ์ระหว่างส่วนย่อย ๆ เหล่านั้น(Webster’s Ninth New Collegiate Dictionary,1991)

การวิเคราะห์ข้อมูล เป็นวิธีการหาความหมายจากข้อมูล ประกอบด้วยการจัดประเภท การจัดระเบียบ การจัดกระทำ และการสรุปย่อ เพื่อให้ได้คำตอบของปัญหาการวิจัย โดยใช้หรือไม่ใช้สถิติก็ได้ หรือ การวิเคราะห์ข้อมูล เป็นวิธีการตรวจสอบทฤษฎีโดยการศึกษาความสัมพันธ์ระหว่างตัวแปรที่เป็นข้อมูลเชิงประจักษ์ หรือเป็นการอธิบายความแปรปรวนว่าตัวแปรที่ศึกษามีความแปรปรวนหรือลักษณะที่แตกต่างกันเนื่องมาจากตัวแปรใดบ้าง(Kerlinger,1982 อ้างอิงใน นงลักษณ์ วิรัชชัย,2543 : 209-210)

2. ประเภทของการวิเคราะห์ข้อมูล

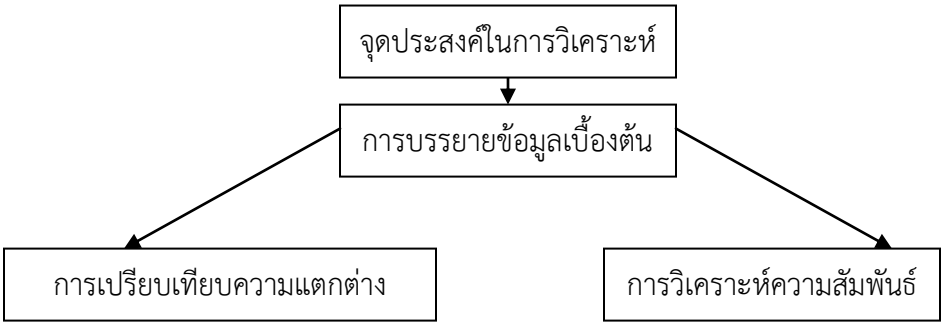
ประเภทของการวิเคราะห์วิเคราะห์ข้อมูล มีดังนี้

2.1 การวิเคราะห์เบื้องต้น(Preliminary Analysis) เป็นการวิเคราะห์เพื่อใช้บรรยายลักษณะของประชากรหรือกลุ่มตัวอย่าง คุณภาพของเครื่องมือ และลักษณะของการแจกแจงของตัวแปร

2.2 การวิเคราะห์เพื่อตรวจสอบข้อตกลงเบื้องต้นของการใช้สถิติ(Analysis to Examine Statistical Assumption) เพื่อช่วยให้เลือกใช้สถิติได้อย่างเหมาะสมอันจะส่งผลให้ผลการวิเคราะห์ข้อมูลมีประสิทธิภาพสูง

2.3 การวิเคราะห์เพื่อตอบปัญหาการวิจัย หรือเพื่อทดสอบสมมุติฐานการวิจัย

ผ่องพรรณ ตรัยมงคลกูล(2543 : 91)ได้จำแนกประเภทของการวิเคราะห์ตามจุดประสงค์ในการวิเคราะห์ ดังแสดงใน**ภาพที่ 10.2**(ผ่องพรรณ ตรัยมงคลกูล,2543 : 91)



ภาพที่ 10.2 การจำแนกการวิเคราะห์ตามจุดประสงค์

3. การวางแผนการวิเคราะห์ข้อมูล

ในการวิเคราะห์ข้อมูล ผู้วิจัยต้องมีการวางแผนในการวิเคราะห์ ดังนี้(อาธง สุทธาศาสน์,2527 : 108)

3.1 จะวิเคราะห์ตัวแปรแต่ละตัวที่ศึกษาอย่างไร และนำเสนอผลการวิเคราะห์อย่างไร

3.2 จำแนกตัวแปรอิสระ/ตัวแปรตาม และต้องการนำเสนอความสัมพันธ์ระหว่างตัวแปรทั้งสองอย่างไร ใช้สถิติอะไร

3.3 ถ้ามีสมมุติฐานการวิจัย จะนำเสนอผลการวิเคราะห์ข้อมูลเพื่อทดสอบสมมุติฐานอย่างไร ใช้สถิติอะไร

3.4 ความสัมพันธ์ของตัวแปรที่ต้องการนำเสนอในลักษณะใด เพื่อการอธิบาย ตีความ หรือการกำหนด ดังนี้

3.4.1 ความสัมพันธ์ระหว่างตัวแปรเป็นความสัมพันธ์ที่แท้จริงหรือไม่ โดยการนำตัวแปรที่คาดว่าจะมีส่วนทำให้ความสัมพันธ์ระหว่างตัวแปรลดลงมาดำเนินการควบคุมแล้วพิจารณาว่าความสัมพันธ์เดิมนั้นลดลงหรือหมดไปก็แสดงว่าความสัมพันธ์ของตัวแปรทั้งสองเป็นเพราะตัวแปรดังกล่าวไม่ใช่ความสัมพันธ์ของตัวแปรทั้งสอง

3.4.2 ความสัมพันธ์ของตัวแปรที่ต้องการนำเสนอในลักษณะตีความ เป็นการพิจารณาความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามมีสาเหตุมาจากอะไร ที่อาจเป็นเพราะมีตัวแปรแทรกระหว่างตัวแปรทั้งสองที่อาจเป็นเหตุของตัวแปรอิสระ แต่เป็นผลของตัวแปรตามดังนั้นจึงควรนำตัวแปรดังกล่าวมาทดสอบ ถ้าพบว่าความสัมพันธ์ลดลงหรือหมดไป แสดงว่าตัวแปรอิสระและตัวแปรตามมีความสัมพันธ์ผ่านตัวแปรแทรก

3.4.3 ความสัมพันธ์ของตัวแปรที่ต้องการนำเสนอในลักษณะเพื่อการกำหนดเงื่อนไข เป็นการพิจารณาว่าเมื่อนำตัวแปรมาทดลองแล้วความสัมพันธ์เดิมเป็นอย่างไรจะเปลี่ยนแปลงหรือไม่ เปลี่ยนแปลงตามเงื่อนไขหรือตัวแปรอะไรบ้าง โดยเฉพาะในกรณีที่ความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามที่มีไม่มากพอที่จะสรุปได้ ดังนั้นผู้วิจัยจะต้องคุมตัวแปรแบบกำหนดเงื่อนไขเพื่อที่จะสามารถพิจารณาความสัมพันธ์ที่มีอยู่นั้นชัดเจนมากขึ้น

3.5 นักวิจัยควรรู้ล่วงหน้าว่าจะแปลความหมายของข้อมูลที่วิเคราะห์ได้ในหลายลักษณะได้อย่างไร และจะต้องปฏิบัติอย่างไรต่อไป อาทิ ใช้การวิเคราะห์ความแปรปรวนเมื่อพบความแตกต่างอย่างมีนัยสำคัญทางสถิติแล้วควรนำข้อมูลแปแสดงการเปรียบเทียบรายคู่เชิงซ้อนต่อไป เป็นต้น

4. ข้อเสนอแนะในการวิเคราะห์ข้อมูล

กูด(Good อ้างอิงใน ล้วน สายยศ และอังคณา สายยศ,2538 : 191) ได้ให้ข้อเสนอแนะสำหรับการวิเคราะห์ข้อมูล ดังนี้

4.1พิจารณาวัตถุประสงค์ของการวิจัยที่เป็นปัญหาให้มีความเข้าใจที่ชัดเจน

4.2 ให้พิจารณาว่าแต่ละประเด็นปัญหาต้องการข้อมูลประเภทใด และจะใช้วิธีการทางสถิติอะไรบ้าง

4.3 สถิตินั้นหาได้หรือไม่กับข้อมูล เพื่อนำไปตอบปัญหาการวิจัยในแต่ละประเด็น

4.4 เลือกข้อมูลที่ได้นำมาจัดหมวดหมู่ จำแนกตามเนื้อหาของประเด็นปัญหาแต่ละประเด็น

4.5 คำนวณค่าสถิติให้ตรงกับประเด็นปัญหาที่จะตอบ

4.6 ให้แปลความหมายของข้อมูลเป็นระยะ(ระหว่างเก็บรวบรวมข้อมูล)

4.7 ออกแบบตารางการวิเคราะห์ข้อมูลที่จะนำเสนอที่ถูกต้อง ชัดเจน กะทัดรัดแต่อธิบายข้อมูลได้มาก

4.8 ถ้านำเสนอข้อมูลในลักษณะของกราฟให้พิจารณาเลือกกราฟที่ทำความเข้าใจได้ง่าย ๆ ไม่ซับซ้อน

การแจกแจงความถี่

1.ความหมายของการแจกแจงความถี่

การแจกแจงความถี่ (Frequency Distribution) เป็นการนำข้อมูลมาจัดหมวดหมู่ ประเภท หรือสร้างตารางแจกแจงความถี่ เพื่อแสดงจำนวนของข้อมูลว่าแต่ละข้อมูลนั้นมีกี่จำนวน ที่เป็นระเบียบง่าย สะดวกในการนำไปใช้ และน่าสนใจในการศึกษา (ล้วน สายยศ และ อังคณา สายยศ,2540 :21-28 ; สุวิมล ตริกานันท์,2546 : 19-24)

2. ประเภทและวิธีการวิเคราะห์การแจกแจงความถี่

2.1 การแจกแจงความถี่ที่ไม่ได้จัดหมวดหมู่(Ungroup Data) เป็นการแจกแจงความถี่ที่มีขั้นตอนในการดำเนินการอย่างง่าย ๆ ดังนี้

2.1.1 เรียงลำดับข้อมูลทุกตัวจากข้อมูลที่มีค่าสูงสุดไปต่ำสุด

2.1.2 พิจารณาข้อมูลแล้วขีดรอยขีด(Tally)แทนข้อมูลแต่ละตัวในข้อมูลนั้น ๆ แล้วให้นับจำนวนรอยขีดใส่จำนวนในช่องของความถี่

ดังแสดงการแจกแจงความถี่ที่ไม่ได้จัดหมวดหมู่ใน**ตัวอย่างที่ 10.1**

ตัวอย่างที่ 10.1 ในการทดสอบวิชาคณิตศาสตร์ของนักเรียนจำนวน 20 คน ได้คะแนนดังนี้

18 15 17 16 15 14 18 17 16 14
15 17 14 18 13 17 19 20 18 15

ให้สร้างตารางแจกแจงความถี่ของคะแนนวิชาคณิตศาสตร์ของนักเรียน

วิธีทำ สามารถแสดงได้คะแนนดังตารางแจกแจงความถี่

❶ ❷ ❸

คะแนนที่ได้	รอยขีด	จำนวนนักเรียน
13	/	1
14	///	3
15	////	4
16	//	2
17	////	4
18	////	4
19	/	1
20	/	1
	รวม	20

2.2 การแจกแจงความถี่ที่จัดหมวดหมู่(Grouped Data) เป็นการแจกแจงความถี่ของข้อมูลที่มีจำนวนมาก ๆ โดยที่นิยมจัดแบ่งข้อมูลเป็นช่วงของคะแนน มีขั้นตอนในการดำเนินการ ดังนี้

2.2.1 คำนวณหาพิสัย(Range)ของข้อมูลที่เป็นความแตกต่างระหว่างข้อมูลที่มีค่าสูงสุดกับข้อมูลที่มีค่าต่ำสุด จากสูตรคำนวณ

พิสัยของข้อมูล = ค่าของข้อมูลสูงสุด-ค่าของข้อมูลต่ำสุด

2.2.2 กำหนดจำนวนชั้นของคะแนนที่ต้องการสร้าง

2.2.3 หาอันตรภาคชั้น หรือช่วงกว้างของแต่ละชั้น จากสูตรคำนวณ

$$\text{อันตรภาคชั้น} = \frac{\text{พิสัย}}{\text{จำนวนชั้น}}$$

2.2.4 นำอันตรภาคชั้นมากำหนดช่วงข้อมูลของแต่ละชั้นให้มีความครอบคลุมข้อมูลจากมากไปน้อยหรือจากน้อยไปมาก

2.2.5 ขีดรอยขีดแทนข้อมูลแต่ละตัวในข้อมูลนั้น ๆ แล้วให้นับจำนวนรอยขีดใส่จำนวนในช่องของความถี่

สามารถแสดงการแจกแจงความถี่ที่จัดหมวดหมู่ใน**ตัวอย่างที่ 10.2**

ตัวอย่างที่ 10.2 จากข้อมูลการทดสอบคะแนนความสามารถทั่วไปของนักเรียน มีผลดังนี้

39	38	44	45	36	43	47	49	36	43	50
41	44	43	36	42	41	44	43	45	49	35
39	48	41	40	40	45	48	43	44	46	45

ให้สร้างตารางแจกแจงความถี่แบบจัดหมวดหมู่ของข้อมูล

วิธีทำ มีวิธีการสร้างตารางแจกแจงความถี่ มีดังนี้

1) คำนวณหาพิสัยของข้อมูล

พิสัยของข้อมูล = ค่าของข้อมูลสูงสุด-ค่าของข้อมูลต่ำสุด

= 50 -35 = 15

2) กำหนดจำนวนชั้นของคะแนน เป็น 6 ชั้น

3) หาอันตรภาคชั้น หรือช่วงกว้างของอันตรภาคแต่ละชั้น จากสูตรคำนวณ

$$\text{อันตรภาคชั้น} = \frac{\text{พิสัย}}{\text{จำนวนชั้น}}$$

$$= \frac{15}{6} = 2.5 \approx 3$$

4) กำหนดช่วงข้อมูลของแต่ละชั้นให้มีความครอบคลุมข้อมูล จากมากไปน้อยหรือจากน้อยไปมาก โดยเริ่มต้นที่ข้อมูลที่มากที่สุดแล้วลดช่วงละ 3 ได้แก่ ช่วงข้อมูล 50-48,47-45, 44-42,41-39,38-36,35-33

5) พิจารณาข้อมูลแล้วขีดรอยขีดแทนข้อมูลแต่ละตัวในข้อมูลนั้น ๆ แล้วให้นับจำนวนรอยขีดใส่จำนวนในช่องของความถี่

ดังแสดงการแจกแจงความถี่ในตาราง

คะแนน	รอยขีด	จำนวนนักเรียน
33-35	/	1
36-38	////	4
39-41	~~////~~ //	7
42-44	~~/////~~~~////~~	10
45-47	~~////~~ /	6
48-50	~~////~~	5
	รวม	33

การวัดแนวโน้มเข้าสู่ส่วนกลาง

1. ความหมายของการวัดแนวโน้มเข้าสู่ส่วนกลาง

การวัดแนวโน้มเข้าสู่ส่วนกลาง(Measure of Central Tendency) เป็นการหาค่ากลางที่เป็นตัวแทนของข้อมูลแต่ละชุด เพื่อใช้สรุปผลของข้อมูลในภาพรวม โดยค่ากลางที่นิยมใช้ มี 3 ค่า ได้แก่ ค่าเฉลี่ยเลขคณิต ค่ามัธยฐานและฐานนิยม (ล้วน สายยศ และ อังคณา สายยศ,2540 :52-73 ; สุวิมล ตริกานันท์,2546 : 51-55 ; พิศิษฐ ตัณฑวณิช,2543 : 36-48)

2. ประเภทและวิธีการวิเคราะห์การวัดแนวโน้มเข้าสู่ส่วนกลาง

2.1 ค่าเฉลี่ยเลขคณิต (Arithematic Mean หรือ Mean : μ(ค่าเฉลี่ยของประชากร), $\bar{x}$ (ค่าเฉลี่ยของกลุ่มตัวอย่าง)) เป็นค่าที่ได้จากการนำข้อมูลทุกจำนวนมารวมกันแล้วหารเฉลี่ยด้วยจำนวนข้อมูลทั้งหมด ที่จำแนกรายละเอียด ดังนี้

2.1.1 ค่าเฉลี่ยเลขคณิตที่ไม่แจกแจงความถี่(Ungroup) เป็นการหาค่าเฉลี่ยจากข้อมูลดิบของแต่ละบุคคล โดยใช้สูตรคำนวณ

$$\overline{X} = \frac{\sum X}{n} \text{ หรือ } \mu = \frac{\sum X}{N}$$

เมื่อ $\overline{X}$ เป็นค่าเฉลี่ยจากกลุ่มตัวอย่าง

μ เป็นค่าเฉลี่ยจากประชากร

$\sum X$ เป็นผลรวมของข้อมูล

n, N เป็นจำนวนของข้อมูลจากกลุ่มตัวอย่างและประชากร

ดังแสดงการคำนวณค่าเฉลี่ยเลขคณิตที่ไม่แจกแจงความถี่ใน**ตัวอย่างที่ 10.3**

ตัวอย่างที่ 10.3 ให้หาคะแนนเฉลี่ยของการสอบซ่อมเสริมวิชาคณิตศาสตร์ของผู้เรียนกลุ่มหนึ่งมีคะแนนดังนี้

25 26 27 29 24 26 25

วิธีทำ

จากสูตรคำนวณหาค่าเฉลี่ย $\overline{X} = \frac{\sum X}{n}$

แทนค่า $= \frac{25+26+27+29+24+26+25}{7}$

$$\overline{X} = \frac{182}{7} = 26$$

ดังนั้นคะแนนเฉลี่ยของคะแนนการสอบซ่อมเสริมครั้งนี้เท่ากับ 26 คะแนน

2.1.2 ค่าเฉลี่ยเลขคณิตที่แจกแจงความถี่(Group) จำแนกเป็น 2 ลักษณะ ดังนี้

2.1.2.1 การหาค่าเฉลี่ยจากข้อมูลที่ไม่มีการจัดหมวดหมู่ จากสูตรคำนวณ

$$\overline{X} = \frac{\sum fx}{n} \text{ หรือ } \mu = \frac{\sum fx}{N}$$

เมื่อ $\overline{X}$ เป็นค่าเฉลี่ยจากกลุ่มตัวอย่าง

μ เป็นค่าเฉลี่ยจากประชากร

$\sum fx$ เป็นผลรวมของผลคูณระหว่างความถี่กับข้อมูล

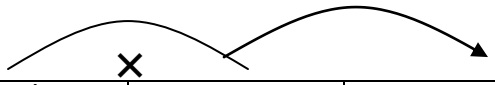
f เป็นความถี่ของข้อมูล

n เป็นจำนวนของข้อมูลของกลุ่มตัวอย่าง

N เป็นจำนวนของข้อมูลของประชากร

ดังแสดงการคำนวณค่าเฉลี่ยเลขคณิตที่แจกแจงความถี่ไม่มีการจัดหมวดหมู่ใน**ตัวอย่างที่ 10.4**

ตัวอย่างที่ 10.4 ให้หาคะแนนเฉลี่ยของวิชาคณิตศาสตร์จากตารางคะแนนของนักเรียนที่กำหนดให้



คะแนนที่ได้(X)	จำนวนนักเรียน (f)	fx
13	1	13
14	3	42
15	4	60
16	2	32
17	4	68
18	4	72
19	1	19
20	1	20
	n=20	$\sum fx = 326$

วิธีทำ จากสูตรการหาค่าเฉลี่ย $\overline{X} = \frac{\sum fx}{n}$

แทนค่า $\overline{X} = \frac{326}{20} = 16.3$

จะได้คะแนนเฉลี่ยของวิชาคณิตศาสตร์เท่ากับ 16.3 คะแนน

2.1.2 การหาค่าเฉลี่ยจากข้อมูลที่มีการจัดหมวดหมู่ ใช้สูตรคำนวณเดียวกับการหาค่าเฉลี่ยที่ไม่มีการจัดหมวดหมู่ แต่ค่าข้อมูลแต่ละจำนวนเป็นค่าที่ได้จากจุดกึ่งกลางชั้นของแต่ละช่วงชั้น โดยหาจุดกึ่งกลางชั้น(Midpoint)จากสูตรคำนวณ

$$\text{จุดกึ่งกลางชั้น } \frac{\text{ขีดจำกัดบน+ขีดจำกัดล่าง}}{2}$$

โดยที่ ขีดจำกัดบน หมายถึง ค่าเฉลี่ยระหว่างผลบวกของค่าสูงสุดของช่วงข้อมูลนั้นกับค่าต่ำสุดของช่วงข้อมูลที่มีค่าสูงกว่า

ขีดจำกัดล่าง หมายถึง ค่าเฉลี่ยระหว่างผลบวกของค่าต่ำสุดของช่วงข้อมูลนั้นกับค่าสูงสุดของช่วงข้อมูลที่มีค่าต่ำกว่า

ดังแสดงการคำนวณค่าเฉลี่ยเลขคณิตที่แจกแจงความถี่และมีการจัดหมวดหมู่ใน**ตัวอย่างที่ 10.5**

ตัวอย่างที่ 10.5 จงหาคะแนนเฉลี่ยของการทดสอบความสามารถทั่วไปของนักเรียนจากตารางข้อมูลที่กำหนดให้

วิธีทำ

คะแนนที่ได้	**จุดกึ่งกลางชั้น (x)**	**จำนวนนักเรียน(f)**	**fx**
33-35	$\frac{33+35}{2}$ =34	1	49×5=245
36-38	37	4	276
39-41	40	7	430
42-44	43	10	280
45-47	46	6	148
48-50	49	5	34
			$\sum fx = 1,413$

วิธีทำ

จากสูตรการหาค่าเฉลี่ยเลขคณิต $\overline{X} = \frac{\sum fx}{n}$

$$\overline{X} = \frac{1,413}{33} = 42.82$$

จะได้คะแนนเฉลี่ยความสามารถทั่วไปของนักเรียนกลุ่มนี้เท่ากับ 42.82 คะแนน

2.1.3 คุณสมบัติที่สำคัญของค่าเฉลี่ยเลขคณิต

ค่าเฉลี่ยเลขคณิตมีคุณสมบัติที่สำคัญ ดังนี้(ชูศรี วงศ์รัตนะ,2546 :41)

2.1.3.1 ผลรวมระหว่างความแตกต่างระหว่างข้อมูลแต่ละจำนวนจากค่าเฉลี่ยเลขคณิตของข้อมูลชุดนั้นเท่ากับศูนย์ ดังแสดงในสมการ

$$\sum_{i=1}^{N}(X_i - \overline{X}) = 0$$

2.1.3.2 ผลรวมของความแตกต่างกำลังสองของข้อมูลแต่ละตัวจากจำนวนM ใด ๆ จะมีค่าน้อยที่สุดเมื่อจำนวนนั้น ๆ(M) เป็นค่าเฉลี่ยเลขคณิตของข้อมูลทุกตัว ดังแสดงในสมการ

$$\sum_{i=1}^{N}(X_i - M)^2 \text{ เมื่อ } M = \overline{X}$$

2.1.3.3 ค่าเฉลี่ยเลขคณิตของข้อมูลจะต้องมีค่าอยู่ระหว่างข้อมูลที่มีค่าต่ำสุดและข้อมูลที่มีค่าสูงสุด ดังแสดงในสมการ

$$\overline{X}_{min} \langle \overline{X} \langle \overline{X}_{max}$$

2.1.3.4 ในกรณีที่ค่าของข้อมูลเพิ่มขึ้นหรือลดลงเท่า ๆ กันค่าเฉลี่ยเลขคณิตของข้อมูลชุดนั้นจะเปลี่ยนแปลงในปริมาณที่เท่ากัน ดังสมการ

$$\overline{X}_{x\pm c} = \overline{X} + c \text{ เมื่อ } c \text{ เป็นค่าคงที่}$$

2.1.3.5 ถ้าข้อมูลทุกตัวนำมาคูณหรือหารด้วยค่าคงที่(ตัวหารต้องไม่เป็นศูนย์)ค่าเฉลี่ยเลขคณิตของข้อมูลชุดนั้นจะมีค่าเฉลี่ยเลขคณิตคูณกับค่าคงที่นั้น ดังแสดงในสมการ

$$\overline{X}_{cx} = c\overline{X}$$
$$\overline{X}_{x/c} = \frac{\overline{X}}{c}$$

2.2 ค่ามัธยฐาน (Median : Me) เป็นค่าของข้อมูลที่มีตำแหน่งอยู่กึ่งกลางของข้อมูลเมื่อได้เรียงลำดับข้อมูลแล้ว โดยส่วนมากจะใช้กับข้อมูลในระดับมาตราเรียงลำดับ หรือข้อมูลที่มีการแจกแจงในลักษณะเบ้ ที่มีรายละเอียด ดังนี้

2.2.1 กรณีข้อมูลไม่ได้แจกแจงความถี่ มีขั้นตอนในการดำเนินการหาค่ามัธยฐาน ดังนี้

2.2.1.1 เรียงลำดับข้อมูลที่มีค่าจากมากไปหาน้อยหรือจากน้อยไปหามาก

2.2.1.2 หาตำแหน่งของข้อมูลจากสูตรคำนวณ

$$\text{ตำแหน่งของมัธยฐาน} = \frac{N+1}{2}$$

เมื่อ N เป็นจำนวนของข้อมูล

2.2.1.3 หาค่าของข้อมูลที่อยู่ในตำแหน่งที่คำนวณได้ แต่ถ้าค่าของข้อมูลอยู่กึ่งกลางของข้อมูล 2 ตัวจะต้องนำค่าของข้อมูลทั้งสองมาเฉลี่ยเป็นค่ามัธยฐาน

ดังแสดงการหามัธยฐานของข้อมูลไม่ได้แจกแจงความถี่ใน**ตัวอย่างที่ 10.6 และ 10.7**

ตัวอย่างที่ 10.6 คะแนนสอบซ่อมเสริมของนักเรียนห้องที่ 1

6 5 8 7 10 9

ให้หาค่ามัธยฐานของคะแนนสอบซ่อมเสริมของนักเรียน

วิธีทำ 1) เรียงลำดับข้อมูล

5 6 7 8 9 10

2) หาตำแหน่งของค่ามัธยฐาน จากสูตร

$$= \frac{N+1}{2}$$

แทนค่า $= \frac{6+1}{2}$

ดังนั้นตำแหน่งของค่ามัธยฐาน เท่ากับ 3.5

3) ตำแหน่งที่ 3.5 อยู่ระหว่างข้อมูลตำแหน่งที่ 3 เท่ากับ 7 และตำแหน่งที่ 4 เท่ากับ 8 ดังนั้นค่ามัธยฐานของคะแนนสอบซ่อมเสริม $(Me) = \frac{7+8}{2} = 7.5$

ตัวอย่างที่ 10.7 คะแนนสอบซ่อมเสริมของนักเรียนห้องที่ 2

6 5 8 7 10 9 11

จงหาค่ามัธยฐานของคะแนนสอบซ่อมเสริม

วิธีทำ 1) เรียงลำดับข้อมูล

5 6 7 8 9 10 11

2) หาตำแหน่งของค่ามัธยฐาน จากสูตรคำนวณ $= \frac{N+1}{2}$

แทนค่า $= \frac{7+1}{2} = 4$

3) ตำแหน่งที่ 4 ของข้อมูลมีค่าเท่ากับ 8 ดังนั้นค่ามัธยฐาน $(Me) = 8$

2.2.2 กรณีข้อมูลแจกแจงความถี่ มีรายละเอียด ดังนี้

2.2.2.1 กรณีข้อมูลแจกแจงความถี่ที่ไม่ได้จัดหมวดหมู่ มีขั้นตอนการดำเนินการหาค่ามัธยฐาน ดังนี้

1) หาความถี่สะสมของแต่ละชั้นจากข้อมูลที่มีค่าน้อยไปมาก

2) หาตำแหน่งของข้อมูลจากสูตรคำนวณ

$$\text{ตำแหน่งของมัธยฐาน} = \frac{N+1}{2}$$

เมื่อ N เป็นจำนวนของข้อมูล

3) หาค่าของข้อมูลที่ตำแหน่งมัธยฐาน

ดังแสดงการหาค่ามัธยฐานของข้อมูลที่แจกแจงความถี่ไม่ได้จัดหมวดหมู่ใน**ตัวอย่างที่ 10.8**

ตัวอย่างที่ 10.8 จงหาค่ามัธยฐานของคะแนนของวิชาคณิตศาสตร์จากตารางข้อมูลที่กำหนดให้

วิธีทำ 1) หาความถี่สะสมของข้อมูลจากข้อมูลที่มีค่าน้อยไปหาค่ามาก

คะแนนที่ได้(x)	**จำนวนนักเรียน (f)**	**ความถี่สะสม**	
13	1	1	
14	3	4	
15	4	8	
16	2	10	
17	4	14	← ตำแหน่งที่ค่ามัธยฐานอยู่
18	4	18	
19	1	19	
20	1	20	

2) หาตำแหน่งของค่ามัธยฐาน $= \frac{20+1}{2} = 10.5$

3) หาค่าของมัธยฐานจะอยู่ระหว่างตำแหน่งที่ 10 และ11 (พิจารณาจากความถี่สะสมว่าเกิน 9 แต่ไม่เกิน 14 ที่มีคะแนนเท่ากับ 17)

จะได้ค่ามัธยฐานของคะแนนวิชาคณิตศาสตร์ เท่ากับ 17 คะแนน

2.2.2.2 กรณีข้อมูลแจกแจงความถี่ที่จัดหมวดหมู่ มีขั้นตอนการดำเนินการหาค่ามัธยฐาน ดังนี้

1) คำนวณหาตำแหน่งของค่ามัธยฐาน จากสูตร $\frac{N}{2}$

2) คำนวณหาค่ามัธยฐานโดยใช้สูตรคำนวณ

$$Me = L + i\left[\frac{\frac{N}{2} - F}{f}\right]$$

เมื่อ Me เป็นค่ามัธยฐานของข้อมูล
L เป็นขีดจำกัดล่างของชั้นที่มีค่ามัธยฐานอยู่
i เป็นความก้วางของอันตรภาคชั้น
F เป็นความถี่สะสมก่อนถึงชั้นที่มัธยฐานอยู่
f เป็นความถี่ของคะแนนในชั้นที่มัธยฐานอยู่
N เป็นจำนวนข้อมูลทั้งหมด

ดังแสดงการหามัธยฐานของข้อมูลที่แจกแจงความถี่ไม่ได้จัดหมวดหมู่ใน**ตัวอย่างที่ 10.9**

ตัวอย่างที่ 10.9 จงหาค่ามัธยฐานของคะแนนการทดสอบความสามารถทั่วไปของนักเรียน

คะแนนที่ได้	**จำนวนนักเรียน (f)**	**ความถี่สะสม**
33-35	1	1
36-38	4	5
39-41	7	12 (F)
42-44	10 (f)	22 ← ตำแหน่งที่ค่ามัธยฐานอยู่
45-47	6	28
48-50	5	33
	N=33	

วิธีทำ

1) คำนวณหาตำแหน่งของค่ามัธยฐาน จากสูตร $\frac{N}{2} = \frac{33}{2} = 16.5$

2) คำนวณหาค่ามัธยฐานโดยใช้สูตรคำนวณ

$$Me = L + i\left[\frac{\frac{N}{2} - F}{f}\right]$$

แทนค่า $= 41.5 + 3\left[\frac{16.5 - 12}{10}\right] = 42.85$

จะได้ค่ามัธยฐานของคะแนนความสามารถทั่วไป เท่ากับ 42.85

2.3 ค่าฐานนิยม (Mode : Mo) เป็นค่าของข้อมูลที่มีค่าซ้ำกันมากที่สุด หรือข้อมูลที่มีความถี่สูงสุดในข้อมูลชุดนั้น ๆ จะใช้กับข้อมูลมาตรานามบัญญัติขึ้นไป มีรายละเอียด ดังนี้

2.3.1 กรณีที่ข้อมูลไม่ได้แจกแจงความถี่ จะใช้วิธีการพิจารณาข้อมูลที่ซ้ำกันมากที่สุด เป็นค่าฐานนิยม 1 ค่า หรือมีข้อมูลซ้ำมากที่สุด 2 ข้อมูลจะมีค่าฐานนิยม 2 ค่า แต่ถ้าข้อมูลทุกตัวไม่ซ้ำกันเลยแสดงว่าข้อมูลชุดนั้นไม่มีค่าฐานนิยม

ดังแสดงการหาฐานนิยมของข้อมูลที่ไม่ได้แจกแจงความถี่ใน**ตัวอย่างที่ 10.10**

ตัวอย่างที่ 10.10 วิธีการการหาฐานนิยมของข้อมูลแต่ละชุด

วิธีทำ

ข้อมูลคะแนนชุดที่ 1 มีดังนี้ 5 6 7 8 4 5 6 5 7 | มี 5 ซ้ำกัน 3

ดังนั้น ค่าฐานนิยม คือ 5 (ข้อมูลซ้ำกันมากที่สุด)

ข้อมูลคะแนนชุดที่ 2 มีดังนี้ 5 6 7 8 7 5 6 5 7 | มี 5และ 7 ซ้ำกัน 3

ดังนั้น ค่าฐานนิยม คือ 5 และ 7(จำนวนข้อมูลซ้ำกันที่เท่ากัน)

ข้อมูลคะแนนชุดที่ 3 มีดังนี้ 5 6 9 8 4 3 1 2 7

ดังนั้นข้อมูลชุดนี้ ไม่มีค่าฐานนิยม (ไม่มีข้อมูลที่ซ้ำกัน)

2.3.2 กรณีที่ข้อมูลแจกแจงความถี่ การหาฐานนิยม จำแนก เป็น 2 ลักษณะ ดังนี้

2.3.2.1 กรณีที่ข้อมูลแจกแจงความถี่ที่ไม่ได้จัดหมวดหมู่ จะใช้วิธีการพิจารณาข้อมูลที่มีความถี่สูงสุด เป็นค่าฐานนิยม

ดังแสดงการหาฐานนิยมของข้อมูลที่แจกแจงความถี่ไม่ได้จัดหมวดหมู่ใน**ตัวอย่างที่ 10.11**

ตัวอย่างที่ 10.11 จงหาค่าฐานนิยมของคะแนนสอบของวิชาคณิตศาสตร์ในตารางที่กำหนดให้

วิธีทำ

คะแนนที่ได้(x)	**จำนวนนักเรียน(f)**	
13	1	
14	3	
15	4	
16	2	
17	5	← ความถี่สูงสุด
18	4	
19	1	
20	1	

ดังนั้น ค่าฐานนิยม(Mo)ของคะแนนวิชาคณิตศาสตร์ของนักเรียน เท่ากับ 17

2.3.2.2 กรณีที่ข้อมูลแจกแจงความถี่ที่จัดหมวดหมู่แล้ว มีขั้นตอนการดำเนินการดังนี้

1) พิจารณาเลือกช่วงข้อมูลที่มีความถี่สูงสุด

2) คำนวณค่าฐานนิยมโดยใช้สูตรคำนวณ

$$Mo = L + i\left[\frac{d_1}{d_1 + d_2}\right]$$

หรือ $Mo = 3Me + 2\overline{X}$

เมื่อ Mo เป็นค่าฐานนิยมของข้อมูล

L เป็นขีดจำกัดล่างของชั้นที่มีค่าฐานนิยมอยู่

i เป็นความกว้างของอันตรภาคชั้น

d_1 เป็นผลต่างของความถี่ของชั้นที่ค่าฐานนิยมอยู่กับความถี่ของชั้นคะแนนที่ต่ำกว่า

d_2 เป็นผลต่างของความถี่ของชั้นที่ค่าฐานนิยมอยู่กับความถี่ของชั้นคะแนนที่สูงกว่า

ดังแสดงการหาฐานนิยมของข้อมูลแจกแจงความถี่ที่จัดหมวดหมู่ใน**ตัวอย่างที่ 10.12**

ตัวอย่างที่ 10.12 จงหาค่าฐานนิยมของคะแนนการทดสอบความสามารถทั่วไปของนักเรียน

วิธีทำ

คะแนนที่ได้	**จำนวนนักเรียน (f)**
33-35	1
36-38	4
39-41	7
2-44	10
45-47	6
48-50	5
	N=33

$d_1 = 10 - 7$

← ค่าฐานนิยมอยู่ชั้นนี้

$d_2 = 10 - 6$

วิธีทำ 1)ช่วงของข้อมูลที่มีความถี่สูงสุด คือ 42-44

2) คำนวณค่าฐานนิยมจากสูตร

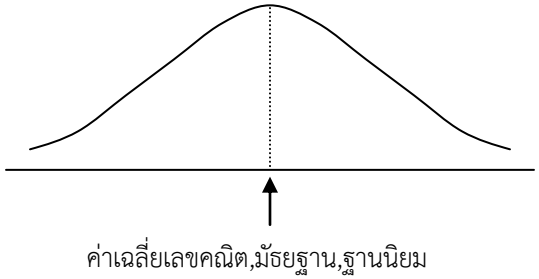
$$Mo = L + i\left[\frac{d_1}{d_1 + d_2}\right]$$

แทนค่า $= 41.5 + 3\left[\frac{3}{3+4}\right]$

ดังนั้นฐานนิยมของคะแนนชุดนี้ $(Mo) = 42.79$

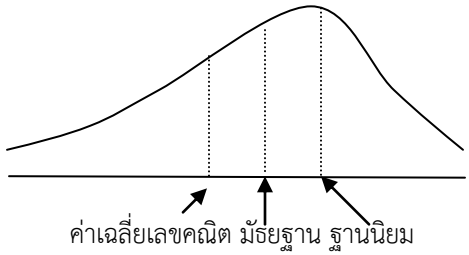
3. ข้อสังเกตในการวิเคราะห์ข้อมูลโดยใช้ค่าเฉลี่ย มัธยฐาน และฐานนิยม(ชูศรี วงศ์รัตนะ, 2546 : 45)

3.1 ในกรณีที่ข้อมูลชุดเดียวกัน มีค่าเฉลี่ยเลขคณิต มัธยฐาน และฐานนิยม เท่ากัน แสดงว่าข้อมูลมีการแจกแจงเป็นโค้งปกติ ดังแสดงใน**ภาพที่ 10.3**



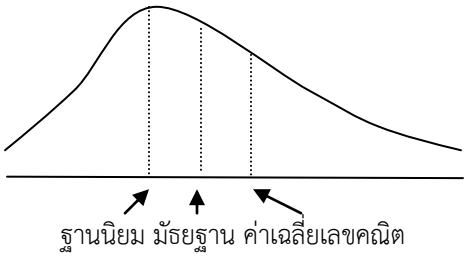
ภาพที่ 10.3 การแจกแจงเป็นโค้งปกติ กรณีที่ข้อมูลชุดเดียวกัน
มีค่าเฉลี่ยเลขคณิต มัธยฐาน และฐานนิยม เท่ากัน

3.2 ในกรณีที่ข้อมูลชุดเดียวกัน มีค่าเฉลี่ยเลขคณิต〈 มัธยฐาน 〈ฐานนิยม แสดงว่าข้อมูลมีการแจกแจงแบบเบ้ซ้าย(Negatively Skewness) ดังแสดงใน**ภาพที่ 10.4**



ภาพที่ 10.4 การแจกแจงเป็นเบ้ซ้าย กรณีที่ข้อมูลชุดเดียวกัน
มีค่าเฉลี่ยเลขคณิต〈 มัธยฐาน 〈ฐานนิยม

3.3 ในกรณีที่ข้อมูลชุดเดียวกัน มีค่าเฉลี่ยเลขคณิต 〉มัธยฐาน〉 ฐานนิยม แสดงว่าข้อมูลมีการแจกแจงแบบเบ้ขวา(Positively Skewness) ดังแสดงใน**ภาพที่ 10.5**



ภาพที่ 10.5 การแจกแจงเป็นเบ้ขวา กรณีที่ข้อมูลชุดเดียวกัน
มีค่าเฉลี่ยเลขคณิต 〉มัธยฐาน〉 ฐานนิยม

3.4 ในกรณีที่มีข้อมูลมาก ๆ ต้องการหาค่าวัดแนวโน้มเข้าสู่ส่วนกลางชุดนั้นมีความแตกต่างกันมากหรือเบ้ไปทางใดทางหนึ่งควรเลือกใช้มัธยฐาน

3.5 ในกรณีที่ข้อมูลมีการแจกแจงแบบเบ้ ควรนำเสนอค่าวัดแนวโน้มเข้าสู่ส่วนกลางทุกค่าให้พิจารณา แต่ถ้าข้อมูลมีการแจกแจงใกล้เคียงแบบโค้งปกติ ควรจะเลือกใช้ค่าเฉลี่ยเพราะเป็นค่าที่มีเสถียรภาพมากที่สุด

3.6 ในกรณีจำแนกข้อมูลตามมาตราวัด มีดังนี้

3.6.1 ข้อมูลในระดับอันตรภาค หรืออัตราส่วนจะใช้ค่าเฉลี่ยเลขคณิต มัธยฐานหรือฐานนิยมก็ได้โดยพิจารณาจากการแจกแจง

3.6.2 ข้อมูลในระดับเรียงอันดับ ให้เลือกใช้มัธยฐานหรือฐานนิยม

3.6.3 ข้อมูลในระดับนามบัญญัติให้เลือกใช้ฐานนิยม

การวัดการกระจาย

1. ความหมายของการวัดการกระจาย

การวัดการกระจาย(Measure of Variability) เป็นวิธีการหาค่าที่ใช้แสดงความแตกต่างของข้อมูลแต่ละตัวในแต่ละชุด ที่ทำให้การเปรียบเทียบผลค่าวัดแนวโน้มเข้าสู่ส่วนกลางของข้อมูลมีความชัดเจนมากยิ่งขึ้นโดยเฉพาะในกรณีที่ค่าวัดแนวโน้มเข้าสู่ส่วนกลางมีค่าเท่ากัน ผู้วิจัยจะต้องนำค่าจากการวัดการกระจายมาเปรียบเทียบกันเพื่อให้พิจารณาความแตกต่างในระดับมากหรือน้อย โดยวิธีการวัดการกระจายที่นิยมใช้ ได้แก่ พิสัย ส่วนเบี่ยงเบนควอไทล์ ส่วนเบี่ยงเบนเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ความแปรปรวน และสัมประสิทธิ์ความผันแปร เป็นต้น(ล้วน สายยศ และอังคณา สายยศ,2540 :96-113 ;สุวิมล ตริกานันท์,2546: 56-59 ; พิศิษฐ ตัณฑวณิช,2543 : 54-64)

2. ประเภทและวิธีการวิเคราะห์การวัดการกระจาย

2.1 พิสัย(Range) เป็นค่าที่แสดงความแตกต่างระหว่างค่าสูงสุดและค่าต่ำสุดของข้อมูลชุดนั้น ๆ ดังสูตรคำนวณ(Huntsberger,1967: 43)

พิสัยของข้อมูล = ค่าสูงสุดของข้อมูล-ค่าต่ำสุดของข้อมูล หรือ

$$X = X_{max} - X_{min}$$

พิสัย เป็นค่าที่วัดความแตกต่างของข้อมูลอย่างหยาบ ๆ ที่ใช้ข้อมูลเพียง 2 ค่าของข้อมูลในชุด นั้น ๆ ทำให้มีการเปลี่ยนแปลงได้ง่าย ดังนั้นควรจะใช้พิสัยระบุความแตกต่างของข้อมูลที่ต้องการอย่างรวดเร็ว หรือใช้ในการสร้างตารางแจกแจงความถี่ของข้อมูลเท่านั้น

ดังแสดงตัวอย่างการคำนวณหาค่าพิสัยของข้อมูลใน**ตัวอย่างที่ 10.13**

ตัวอย่างที่ 10.13 จงหาพิสัยของข้อมูลชุดหนึ่ง มีดังนี้ 5 6 7 8 4 5 6 5 7

วิธีทำ พิสัยของข้อมูล = ค่าของข้อมูลสูงสุด-ค่าของข้อมูลต่ำสุด

= 8 – 4 = 4

2.2 ส่วนเบี่ยงเบนควอไทล์(Quartile deviation : Q.D.) เป็นค่าที่ได้จากค่าเฉลี่ยของความแตกต่างระหว่างควอไทล์ที่ 3 และควอไทล์ที่ 1 ที่จำแนก ได้ดังนี้

2.2.1 กรณีข้อมูลไม่ได้แจกแจงความถี่ มีขั้นตอนการดำเนินการ ดังนี้

2.2.1.1 เรียงลำดับข้อมูลจากน้อยไปหามาก แล้วหาค่า Q_1 และ Q_3

2.2.1.2 นำค่า Q_1 และ Q_3 มาแทนค่าในสูตรคำนวณ

$$Q.D. = \frac{Q_3 - Q_1}{2}$$

เมื่อ $Q.D.$ เป็นส่วนเบี่ยงเบนควอไทล์ของข้อมูล

Q_1 เป็นค่าที่แบ่งข้อมูลชุดนั้น เป็น $\frac{1}{4}$ หรือ 25 %

Q_3 เป็นค่าที่แบ่งข้อมูลชุดนั้น เป็น $\frac{3}{4}$ หรือ 75 %

ดังแสดงตัวอย่างการคำนวณหาค่าส่วนเบี่ยงเบนควอไทล์ของข้อมูลที่ไม่ได้แจกแจงความถี่ใน**ตัวอย่างที่ 10.14**

ตัวอย่างที่ 10.14 ผลการทดสอบในรายวิชาคณิตศาสตร์ของนักเรียนจำนวน 8 คน มีคะแนน ดังนี้

18 20 24 46 48 58 60 70

ให้หาส่วนเบี่ยงเบนควอไทล์ของคะแนนชุดนี้

วิธีทำ 1) จากข้อมูล 18 20 24 46 48 58 60 70

แบ่งข้อมูลออกเป็น 4 ส่วน ↓ Q_2,25% ↓ Q_2,50% ↓ Q_3,75%

หา $Q_1 = \frac{20+24}{2} = 22$ และ $Q_3 = \frac{58+60}{2} = 59$

2) หาค่าส่วนเบี่ยงเบนควอไทล์จากสูตรคำนวณ

$$Q.D. = \frac{Q_3 - Q_1}{2}$$

แทนค่า $= \frac{59 - 22}{2} = 18.5$

จะได้ส่วนเบี่ยงเบนควอไทล์ เท่ากับ 18.5

2.2.2 กรณีข้อมูลแจกแจงความถี่แล้ว มีขั้นตอนการดำเนินการ ดังนี้

2.2.2.1 ทำตารางความถี่สะสมจากข้อมูลที่มีค่าต่ำไปข้อมูลที่มีค่าสูงกว่า

2.2.2.2 หาตำแหน่งของควอไทล์ที่ 1 จากสูตร $\frac{N}{4}$

ตำแหน่งของควอไทล์ที่ 3 จากสูตร $\frac{3N}{4}$

เพื่อหาช่วงข้อมูลที่ควอไทล์ที่ 1 และ 3

2.2.2.3 หาค่า Q_1 และ Q_3 จากสูตรคำนวณ

$$Q_1 = L + i\left[\frac{\frac{1}{4}N - F}{f}\right] \text{ และ } Q_3 = L + i\left[\frac{\frac{3}{4}N - F}{f}\right]$$

เมื่อ Q_1 และ Q_3 เป็นค่าควอไทล์ที่ 1 และ 3 ของข้อมูล

L เป็นขีดจำกัดล่างของชั้นที่มีค่าควอไทล์อยู่

i เป็นความกว้างของอันตรภาคชั้น

F เป็นความถี่สะสมก่อนถึงชั้นที่ค่าควอไทล์อยู่

f เป็นความถี่ของคะแนนในชั้นที่ค่าควอไทล์อยู่

N เป็นจำนวนข้อมูลทั้งหมด

ดังแสดงตัวอย่างการคำนวณหาค่าส่วนเบี่ยงเบนควอไทล์ของข้อมูลที่แจกแจงความถี่ ใน**ตัวอย่างที่ 10.15**

ตัวอย่างที่ 10.15 จงหาค่าส่วนเบี่ยงเบนควอไทล์ของคะแนนของวิชาคณิตศาสตร์

วิธีทำ 1) หาตำแหน่งของควอไทล์ที่ 1 $= \frac{N}{4} = \frac{33}{4} = 8.25$

ตำแหน่งของควอไทล์ที่ 3 $= \frac{3N}{4} = \frac{3\times 33}{4} = 24.75$

คะแนนที่ได้	จำนวนนักเรียน (f)	ความถี่สะสม	
33-35	1	1	
36-38	4	5	ตำแหน่งที่
39-41	7	12	← ค่าควอไทล์ที่ 1อยู่
42-44	10	22	ตำแหน่งที่
45-47	6	28	← ค่าควอไทล์ที่ 3 อยู่
48-50	5	33	
	N=33		

2) หาค่า Q_1 และ Q_3 จากสูตร

$$Q_1 = 38.5 + 3\left[\frac{8.25-5}{7}\right] \text{และ } Q_3 = 44.5 + 3\left[\frac{24.75-22}{6}\right]$$

$$Q_1 = 39.89 \quad \text{และ } Q_3 = 45.88$$

3) ค่าส่วนเบี่ยงเบนควอไทล์จากสูตร

$$Q.D. = \frac{Q_3 - Q_1}{2}$$

แทนค่า $Q.D. = \frac{45.88 - 39.89}{2} = 2.995$

จะได้ส่วนเบี่ยงเบนควอไทล์ของคะแนนวิชาคณิตศาสตร์ เท่ากับ 2.995

2.3 ค่าส่วนเบี่ยงเบนเฉลี่ย(Average Deviation : A.D.) เป็นค่าเฉลี่ยของความเบี่ยงเบนของคะแนนแต่ละตัวจากค่าเฉลี่ยเลขคณิตของข้อมูลชุดเดียวกัน ที่จำแนก ดังนี้

2.3.1 กรณีข้อมูลไม่ได้แจกแจงความถี่ มีขั้นตอนการดำเนินการ ดังนี้

2.3.1.1 หาค่าเฉลี่ยเลขคณิตของข้อมูล

2.3.1.2 คำนวณค่าส่วนเบี่ยงเบนเฉลี่ยจากสูตรคำนวณ

$$A.D. = \frac{\sum|X - \overline{X}|}{N}$$

เมื่อ $A.D.$ เป็นค่าส่วนเบี่ยงเบนเฉลี่ยของข้อมูล

X เป็นข้อมูลแต่ละตัว

$\overline{X}$ เป็นค่าเฉลี่ยเลขคณิตของข้อมูล

N เป็นจำนวนของข้อมูล

ดังแสดงตัวอย่างการคำนวณหาค่าส่วนเบี่ยงเบนเฉลี่ยของข้อมูลที่ไม่ได้แจกแจงความถี่ ใน**ตัวอย่างที่ 10.16**

ตัวอย่างที่ 10.16 ผลการทดสอบในรายวิชาคณิตศาสตร์(คะแนนเต็ม10)ของนักเรียนจำนวน 5 คน ได้คะแนน ดังนี้ 2 4 6 6 7
จงหาค่าส่วนเบี่ยงเบนเฉลี่ยของคะแนนชุดนี้

วิธีทำ 1) หาค่าเฉลี่ยเลขคณิตของคะแนนชุดนี้ $\overline{X} = \frac{\sum X}{n} = \frac{2+4+6+6+7}{5} = 5$

2) คำนวณหาค่าส่วนเบี่ยงเบนเฉลี่ยจากสูตรคำนวณ

$$A.D. = \frac{\sum|X - \overline{X}|}{n}$$

แทนค่า $A.D. = \frac{|2-5|+|4-5|+|6-5|+|6-5|+|7-5|}{5}$

ดังนั้น $A.D. = \frac{3+1+1+1+2}{5} = 1.60$

จะได้ส่วนเบี่ยงเบนเฉลี่ยของข้อมูลชุดนี้ เท่ากับ 1.60

2.3.2 กรณีข้อมูลแจกแจงความถี่แล้ว มีขั้นตอนดังนี้

2.3.2.1 หาค่าเฉลี่ยเลขคณิตของข้อมูล,เพิ่มช่องตารางจุดกึ่งกลางชั้น,ความถี่คูณจุดกึ่งกลางชั้น (fX) ,จุดกึ่งกลางชั้นแต่ละชั้นลบค่าเฉลี่ยเลขคณิต $(X - \overline{X})$ และ ความถี่คูณค่าสมบูรณ์ของจุดกึ่งกลางชั้นลบค่าเฉลี่ยเลขคณิต $(f|X - \overline{X}|)$

2.3.2.2 คำนวณค่าส่วนเบี่ยงเบนเฉลี่ยจากสูตรคำนวณ

$$A.D. = \frac{\sum f\left|X - \overline{X}\right|}{n}$$

เมื่อ A.D.เป็นค่าส่วนเบี่ยงเบนเฉลี่ยของข้อมูล

X เป็นข้อมูลแต่ละตัว

$\overline{X}$ เป็นค่าเฉลี่ยเลขคณิตของข้อมูล

n เป็นจำนวนของข้อมูล

f เป็นความถี่ของข้อมูล

ดังแสดงตัวอย่างการคำนวณหาค่าส่วนเบี่ยงเบนเฉลี่ยของข้อมูลที่แจกแจงความถี่แล้วใน **ตัวอย่างที่ 10.17**

ตัวอย่างที่ 10.17 จงหาค่าส่วนเบี่ยงเบนเฉลี่ยของคะแนนของวิชาคณิตศาสตร์ในการทดสอบครั้งหนึ่ง

วิธีทำ

คะแนนที่ได้	**จำนวนนักเรียน(f)**	**จุดกึ่งกลาง (x)**	fx	$X - \overline{X}$	$f\left\|X - \overline{X}\right\|$
33-35	1	34	34	-8.82	8.82
36-38	4	37	148	-5.82	23.28
39-41	7	40	280	-2.82	19.74
42-44	10	43	430	0.18	1.80
45-47	6	46	276	3.18	19.08
48-50	5	49	245	6.18	30.9
	n=33		$\sum fX = 1413$		$\sum f\left\|X - \overline{X}\right\| = 103.62$

1) หาค่าเฉลี่ยเลขคณิตของคะแนนวิชาคณิตศาสตร์ $\overline{X} = \frac{\sum fX}{n} = \frac{1413}{33} = 42.82$

2) หาค่าส่วนเบี่ยงเบนเฉลี่ยของข้อมูล จากสูตร $A.D. = \frac{\sum f\left|X - \overline{X}\right|}{n}$

แทนค่า $= \frac{103.62}{33} = 3.14$

จะได้ค่าส่วนเบี่ยงเบนเฉลี่ยของคะแนนวิชาคณิตศาสตร์ เท่ากับ 3.14

2.4 ส่วนเบี่ยงเบนมาตรฐาน และความแปรปรวน(Variance)

2.4.1 ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation : S.D.) เป็นค่าของการวัดการกระจายที่ได้หาได้จากรากที่สองของคะแนนทุกค่าที่เบี่ยงเบนไปจากค่าเฉลี่ยของข้อมูลยกกำลังสอง มีรายละเอียด จำแนกตามลักษณะของข้อมูล ดังนี้

2.4.1.1 กรณีที่ข้อมูลไม่มีการแจกแจงความถี่ มีสูตรการคำนวณดังนี้

1) กรณีที่ข้อมูลมาจากประชากร มีสูตรคำนวณ

$$\sigma = \sqrt{\frac{\sum (X-\mu)^2}{N}} \text{ หรือ } \sigma = \sqrt{\frac{N\sum X^2 - (\sum X)^2}{N^2}}$$

2) กรณีที่ข้อมูลมาจากกลุ่มตัวอย่าง มีสูตรคำนวณ

$$S.D. = \sqrt{\frac{\sum (X-\overline{X})^2}{n-1}} \text{ หรือ } S.D. = \sqrt{\frac{n\sum X^2 - (\sum X)^2}{n(n-1)}}$$

เมื่อ $S.D.$ เป็นส่วนเบี่ยงเบนมาตรฐานของข้อมูลจากกลุ่มตัวอย่าง

σ เป็นส่วนเบี่ยงเบนมาตรฐานของข้อมูลจากประชากร

$\overline{X}$ เป็นค่าเฉลี่ยเลขคณิตของข้อมูลจากกลุ่มตัวอย่าง

μ เป็นค่าเฉลี่ยเลขคณิตของข้อมูลจากประชากร

X เป็นข้อมูลแต่ละตัว

n เป็นจำนวนข้อมูลของกลุ่มตัวอย่าง

N เป็นจำนวนข้อมูลของประชากร

ดังแสดงตัวอย่างการคำนวณหาค่าส่วนเบี่ยงเบนมาตรฐานของข้อมูลที่ไม่ได้แจกแจงความถี่ ใน**ตัวอย่างที่ 10.18**

ตัวอย่างที่ 10.18 ผลการทดสอบในรายวิชาคณิตศาสตร์(คะแนนเต็ม10)ของนักเรียนจำนวน 5 คน ได้คะแนน ดังนี้ 2 4 6 6 7 ให้หาค่าส่วนเบี่ยงเบนมาตรฐานของคะแนนชุดนี้

วิธีทำ 1) หาค่าเฉลี่ยของคะแนนชุดนี้ $\overline{X} = \frac{\sum X}{n} = \frac{2+4+6+6+7}{5} = 5$

2) คำนวณหาค่าส่วนเบี่ยงเบนมาตรฐานจากสูตรคำนวณ

$$S.D. = \sqrt{\frac{\sum (X-\overline{X})^2}{n-1}}$$

$$S.D. = \sqrt{\frac{(2-5)^2+(4-5)^2+(6-5)^2+(6-5)^2+(7-5)^2}{5-1}} = 2.00$$

จะได้ส่วนเบี่ยงเบนมาตรฐานของวิชาคณิตศาสตร์ เท่ากับ 2.00

หรือ คำนวณค่าส่วนเบี่ยงเบนมาตรฐานจากสูตรคำนวณ

จากข้อมูล(X)	2	4	6	6	7	$\sum X = 25$
X^2	4	16	36	36	49	$\sum X^2 = 141$

$$S.D. = \sqrt{\frac{n\sum X^2 - (\sum X)^2}{n(n-1)}}$$

$$= \sqrt{\frac{5(141)-(25)^2}{5(5-1)}}$$

$$S.D. = 2.00$$

2.4.1.2 กรณีข้อมูลที่แจกแจงความถี่แล้ว มีสูตรการคำนวณ ดังนี้

1) กรณีที่ข้อมูลมาจากประชากร มีสูตรคำนวณ

$$\sigma = \sqrt{\frac{\sum f(X-\mu)^2}{N}} \text{ หรือ } \sigma = \sqrt{\frac{N\sum fX^2 - (\sum fX)^2}{N^2}}$$

2) กรณีที่ข้อมูลมาจากกลุ่มตัวอย่าง มีสูตรคำนวณ

$$S.D. = \sqrt{\frac{\sum f(X - \overline{X})^2}{n-1}} \text{ หรือ } S.D. = \sqrt{\frac{n\sum fX^2 - (\sum fX)^2}{n(n-1)}}$$

เมื่อ $S.D.$ เป็นส่วนเบี่ยงเบนมาตรฐานของข้อมูลจากกลุ่มตัวอย่าง

σ เป็นส่วนเบี่ยงเบนมาตรฐานของข้อมูลจากประชากร

$\overline{X}$ เป็นค่าเฉลี่ยเลขคณิตของข้อมูลจากกลุ่มตัวอย่าง

μ เป็นค่าเฉลี่ยเลขคณิตของข้อมูลจากประชากร

X เป็นข้อมูลแต่ละตัว

f เป็นความถี่ของข้อมูล

n เป็นจำนวนข้อมูลของกลุ่มตัวอย่าง

N เป็นจำนวนข้อมูลของประชากร

ดังแสดงตัวอย่างการคำนวณหาค่าส่วนเบี่ยงเบนมาตรฐานของข้อมูลที่แจกแจงความถี่แล้วใน**ตัวอย่างที่ 10.19**

ตัวอย่างที่ 10.19 จงหาค่าส่วนเบี่ยงเบนมาตรฐานของคะแนนของวิชาคณิตศาสตร์

คะแนนที่ได้	**จำนวนนักเรียน(f)**	**จุดกึ่งกลาง (x)**	fX	X^2	fX^2
33-35	1	34	34	1,156	1,156
36-38	4	37	148	1,369	5,479
39-41	7	40	280	1,600	11,200
42-44	10	43	430	1,849	18,490
45-47	6	46	276	2,116	12,696
48-50	5	49	245	2,401	12,005
	n=33		$\sum fX = 1,413$		$\sum fX^2 = 61,023$

วิธีทำ คำนวณหาค่าส่วนเบี่ยงเบนมาตรฐาน จากสูตรคำนวณ

$$S.D. = \sqrt{\frac{n\sum fX^2 - (\sum fX)^2}{n(n-1)}}$$

แทนค่า $$= \sqrt{\frac{33(61023) - (1413*1413)}{33(33-1)}} = 4.03$$

จะได้ส่วนเบี่ยงเบนมาตรฐานของคะแนนของวิชาคณิตศาสตร์เท่ากับ 4.03

2.4.1.3 การแปลความหมายของค่าส่วนเบี่ยงเบนมาตรฐาน

1) ค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับศูนย์ (σหรือS.D.=0) แสดงว่าข้อมูลชุดนั้นไม่มีการกระจาย หรือข้อมูลมีค่าเท่ากันทุกตัว

2) ค่าส่วนเบี่ยงเบนมาตรฐานเข้าใกล้ศูนย์ แสดงว่า ข้อมูลชุดนั้นมีค่าที่ใกล้เคียงกัน

3) ค่าเบี่ยงเบนมาตรฐานเท่ากับ 1 แสดงว่าการแจกแจงของข้อมูลมีลักษณะเป็นโค้งปกติ(Normal Curve) ที่จะใช้ค่าเฉลี่ยเป็นตัวแทนของข้อมูล แต่ถ้าค่าส่วนเบี่ยงเบนมาตรฐานมากกว่าค่าเฉลี่ยควรใช้ฐานนิยมหรือมัธยฐานเป็นตัวแทนของข้อมูล

4) ในกรณีการเปรียบเทียบค่าส่วนเบี่ยงเบนมาตรฐานระหว่างกลุ่ม จะได้ว่า กลุ่มที่มีค่าส่วนเบี่ยงเบนมาตรฐานต่ำกว่า แสดงว่ามีการกระจายของข้อมูลต่ำกว่าหรือข้อมูลมีค่าใกล้เคียงกัน มากกว่ากลุ่มที่มีค่าส่วนเบี่ยงเบนมาตรฐานสูงกว่า

2.4.2 ความแปรปรวน(Variance) เป็นค่าเฉลี่ยของผลรวมของข้อมูลทุกตัวที่เบี่ยงเบนไปจากค่าเฉลี่ยของข้อมูลยกกำลังสอง หรือ เป็นค่าส่วนเบี่ยงเบนมาตรฐานยกกำลังสอง ใช้สัญลักษณ์ $(S.D)^2$ แทนความแปรปรวนของข้อมูลที่ได้จากกลุ่มตัวอย่าง และ σ^2 แทนความแปรปรวนของข้อมูลที่ได้จากประชากร แต่ในการวัดการกระจายของข้อมูลโดยทั่วไปจะนิยมใช้ส่วนเบี่ยงเบนมาตรฐานมากกว่าใช้ความแปรปรวน เนื่องจากค่าส่วนเบี่ยงมาตรฐานแสดงการกระจายของข้อมูลที่อยู่ในหน่วยเดียวกับข้อมูลเดิมที่นำมาวิเคราะห์ข้อมูล

2.5 สัมประสิทธิ์ความผันแปร(CoefficientVariance : C.V.) เป็นการวัดการกระจายของข้อมูลที่คำนวณหาได้จากอัตราส่วนที่คิดเป็นร้อยละระหว่างส่วนเบี่ยงเบนมาตรฐานต่อค่าเฉลี่ยของข้อมูล ดังสูตรคำนวณ

$$C.V. = \frac{S.D.}{\overline{X}} \times 100$$

เมื่อ $C.V.$ เป็นสัมประสิทธิ์ความแปรปรวน
$S.D.$ เป็นค่าส่วนเบี่ยงเบนมาตรฐาน
$\overline{X}$ เป็นค่าเฉลี่ยเลขคณิตของข้อมูล

โดยที่สัมประสิทธิ์ความแปรปรวนจะใช้ในกรณีที่ต้องการเปรียบเทียบความแตกต่างของข้อมูล 2 ชุดที่มีหน่วยการวัดไม่เหมือนกัน

การวัดตำแหน่งและการเปรียบเทียบ

1. ความหมายของการวัดตำแหน่งและการเปรียบเทียบ

การวัดตำแหน่งและการเปรียบเทียบ เป็นวิธีการหาค่าที่ระบุตำแหน่งของข้อมูล หรือแสดงการเปรียบเทียบระหว่างข้อมูล อาทิ อัตราส่วน ร้อยละ และเปอร์เซนไทล์ เป็นต้น

2. ประเภทและวิธีการวิเคราะห์การวัดตำแหน่งและการเปรียบเทียบ

2.1 อัตราส่วน

อัตราส่วน(Ratio) เป็นการนำข้อมูลเชิงปริมาณ 2 ตัว/กลุ่ม มาเปรียบเทียบในลักษณะของข้อมูลหนึ่งเป็นกี่เท่าของข้อมูลอีกสิ่งหนึ่ง หรือ ข้อมูลที่สนใจต่อข้อมูลทั้งหมด ดังสูตรคำนวณ (สุวิมล ตริกานันท์,2546 : 30)

$$\text{อัตราส่วน} = \frac{\text{ข้อมูลที่ 1}}{\text{ข้อมูลที่ 2}} \quad \text{หรือ} \quad \frac{\text{ข้อมูลที่สนใจ}}{\text{ข้อมูลทั้งหมด}}$$

ดังแสดงตัวอย่างการคำนวณหาอัตราส่วนของข้อมูลใน**ตัวอย่างที่ 10.20**

ตัวอย่างที่ 10.20 จำนวนผู้เรียนในห้องทั้งหมด 30 คน เป็นเพศชาย 12 คน เป็นเพศหญิง 18 คน ให้หาอัตราส่วนของเพศหญิงต่อเพศชาย หรืออัตราส่วนของเพศหญิงหรือเพศชายต่อผู้เรียนทั้งหมด

วิธีทำ อัตราส่วนของเพศหญิง ต่อเพศชาย เท่ากับ $\frac{18}{12}=\frac{3}{2}$

อัตราส่วนของเพศหญิง ต่อผู้เรียนทั้งหมด เท่ากับ $\frac{12}{30}=\frac{2}{5}$

อัตราส่วนของเพศชาย ต่อผู้เรียนทั้งหมด เท่ากับ $\frac{18}{30}=\frac{3}{5}$

2.2 ร้อยละ

ร้อยละ (Percentage) เป็นอัตราส่วนที่เปรียบเทียบจำนวนที่สนใจกับกับจำนวนทั้งหมดที่เป็นจำนวนเต็มร้อย ดังสูตรคำนวณ(วิเชียร เกตุสิงห์,2541 : 64)

$$\text{ร้อยละ} = \frac{\text{จำนวนที่สนใจ}}{\text{จำนวนทั้งหมด}} \times 100$$

โดยมีข้อควรระวังในการใช้ร้อยละของผู้วิจัย มีดังนี้

1) จำนวนเต็มที่ใช้เทียบเป็นส่วนเท่ากับหนึ่งร้อย เพื่อให้การแสดงข้อมูลมีความหมาย แต่พฤติกรรมของมนุษย์ที่ไม่ควรเทียบเป็นหนึ่งร้อยเนื่องจากพฤติกรรมของมนุษย์ในแต่ละช่วงมีอัตราการเปลี่ยนแปลงที่ไม่เท่ากัน

2) ร้อยละของจำนวนที่มีฐานต่างกัน(ก่อนที่จะเปลี่ยนเป็นร้อยละ)จะนำมาบวก ลบ และหาค่าเฉลี่ยไม่ได้

3) ไม่นิยมใช้ร้อยละที่มีค่าเกิน 100 หรือกรณีที่มีฐานเป็นจำนวนน้อย ๆ แต่จะใช้การเปรียบเทียบในลักษณะของสัดส่วนจะเหมาะสมกว่า

4) การใช้ร้อยละจะทำให้ความคลาดเคลื่อนที่มีอยู่แล้วเพิ่มมากยิ่งขึ้น อาทิ ผู้เรียนได้คะแนน 15 ± 2 แสดงว่าได้คะแนนจริงระหว่าง 13 ถึง 17 คะแนน แต่ถ้าขยายเป็นร้อยละจะทำให้ความคลาดเคลื่อน±2 เพิ่มขึ้นด้วย

ดังแสดงตัวอย่างการคำนวณหาร้อยละของข้อมูลในตัวอย่างที่ 10.21

ตัวอย่างที่ 10.21 จำนวนผู้เรียนในห้องทั้งหมด 30 คน เป็นเพศชาย 12 คน เป็นเพศหญิง 18 คน ให้หาร้อยละของผู้เรียนเพศหญิงหรือเพศชายต่อผู้เรียนทั้งหมด

วิธีทำ ร้อยละของเพศชาย ต่อผู้เรียนทั้งหมด เท่ากับ $\frac{12}{30}\times100=40\%$ หรือ ร้อยละ 40

ร้อยละของเพศหญิง ต่อผู้เรียนทั้งหมด เท่ากับ $\frac{18}{30}\times100=60\%$ หรือ ร้อยละ 60

2.3 เปอร์เซนไทล์

เปอร์เซนไทล์(Percentile) เป็นตำแหน่งของข้อมูลที่ระบุให้ทราบว่ามีข้อมูลที่ต่ำกว่าข้อมูลนั้น ๆ คิดเป็นร้อยละเท่าไร อาทิ สมศักดิ์ สอบวิชาคณิตศาสตร์ได้คะแนน 75 คะแนนคิดเป็นเปอร์เซนไทล์ที่ 80 หมายความว่า ในจำนวนผู้เรียน 100 คนมีผู้เรียนที่ได้ต่ำกว่าสมศักดิ์ 80 คน มีขั้นตอนวิธีการคำนวณหาเปอร์เซนไทล์ ดังนี้(ล้วน สายยศ และอังคณา สายยศ, 2540: 82-86)

2.3.1 เรียงลำดับคะแนนคะแนนจากมากไปหาน้อย

2.3.2 แจกแจงความถี่ของคะแนนแต่ละคะแนน

2.3.3 หาความถี่สะสม((cf) ของคะแนนแต่ละคะแนนจากคะแนนต่ำสุดขึ้นสู่คะแนนสูงสุด (ดังตัวอย่าง)โดยความถี่สะสมทั้งหมดจะเท่ากับจำนวนทั้งหมด

2.3.4 คำนวณหา $cf + \frac{1}{2}f$ โดยเริ่มต้นที่ 0 จากช่องความถี่สะสม แล้วนำไปบวกครึ่งหนึ่งของความถี่ ดังตัวอย่าง

2.3.5 หาค่าเปอร์เซ็นต์ไทล์โดยนำ $cf + \frac{1}{2}f$ คูณด้วย $\frac{100}{N}$ (N เป็นจำนวนคนทั้งหมด)

หรือกรณีที่ข้อมูลมีการแจกแจงความถี่เป็นชั้นของข้อมูลสามารถคำนวณหาค่าเปอร์เซนไทล์จากสูตรคำนวณ ที่มีขั้นตอน ดังนี้

1) หาตำแหน่งของเปอร์เซนไทล์จากสูตร $\frac{NX}{100}$ แล้วนำไปพิจารณาว่าอยู่ช่วงใดของข้อมูล

2) นำค่าที่ได้จากตารางมาแทนค่าในสูตรเพื่อหาค่า ดังสูตร(สุวิมล ตริกานันท์,2546 : 39)

$$P_x = L + \left[\frac{\frac{NX}{100} - F}{f} \right] i$$

โดยที่ P_x เป็นเปอร์เซนไทล์ที่ X

L เป็นขีดจำกัดล่างของชั้นที่เปอร์เซนไทล์อยู่

N เป็นจำนวนข้อมูลทั้งหมด

F เป็นความถี่สะสมก่อนถึงชั้นที่เปอร์เซนไทล์อยู่

f เป็นความถี่สะสมของชั้นที่เปอร์เซนไทล์อยู่

i เป็นความกว้างของอันตรภาคชั้น

ดังแสดงตัวอย่างการวิเคราะห์หาเปอร์เซนไทล์ ใน**ตัวอย่างที่ 10.22**

ตัวอย่างที่ 10.22 จากตารางข้อมูลที่กำหนดให้ ให้วิเคราะห์หาเปอร์เซนไทล์

คะแนน	**จำนวน** (f)	**ความถี่สะสม** (cf)	$cf + \frac{1}{2}f$	**เปอร์เซนไทล์** $(cf + \frac{1}{2}f) \times \frac{100}{N}$
41	3	3	1.5	3.75
42	3	6	4.5	12.25
43	4	10	8	20.00
44	4	14	12	30.00
45	8	22	18	40.00
46	5	27	24.5	61.25
47	3	30	28.5	71.25
48	4	34	32	80.00
49	4	38	36	90.00
50	2	40	39	97.50
		N = 40		

ความหมายของเปอร์เซ็นต์ไทล์จากตาราง อาทิ

คะแนน 45 คะแนน ตรงกับเปอร์เซ็นต์ไทล์ที่ 40.0 หมายความว่า ในจำนวน 100 คนมีบุคคลที่ได้คะแนนต่ำกว่า 45 คะแนนเท่ากับ 40 คน

คะแนน 49 คะแนน ตรงกับเปอร์เซ็นต์ไทล์ที่ 90.0 หมายความว่า ในจำนวน 100 คนมีบุคคลที่ได้คะแนนต่ำกว่า 49 คะแนนเท่ากับ 90 คน

สหสัมพันธ์

1. ความหมายของสหสัมพันธ์

สหสัมพันธ์ เป็นการศึกษาที่สามารถระบุว่าข้อมูลจากตัวแปรที่ศึกษามีความสัมพันธ์กันหรือไม่ เท่านั้นที่หาได้จากสิ่งเดียวกันเป็นคู่ ๆ แต่จะไม่สามารถใช้พยากรณ์ตัวแปรหนึ่ง ในกรณีที่ทราบค่าของอีกตัวแปรหนึ่งได้ ดังลักษณะของข้อมูลที่จะนำมาหาสหสัมพันธ์ใน**ตารางที่ 10.2**(สุวิมล ตริกานันท์, 2546 : 39)

ตารางที่ 10.2 ลักษณะของข้อมูลที่จะหาค่าสหสัมพันธ์

ข้อมูลที่	ตัวแปรที่ 1	ตัวแปรที่ 2
1	x_1	y_1
2	x_2	y_2
3	x_3	y_3
4	x_4	y_4
5	x_5	y_5

2. ลักษณะที่สำคัญของสหสัมพันธ์

ในการหาสหสัมพันธ์ของข้อมูลมีลักษณะที่ควรพิจารณา ดังนี้(Runyon,1980 : 120)

2.1 ข้อมูลทั้ง 2 ชุดจะต้องมาจากกลุ่มตัวอย่างเดียวกัน หรือมาจากกลุ่มตัวอย่าง 2 กลุ่มที่มีลักษณะจับคู่กันโดยใช้หลักการที่เฉพาะเจาะจง อาทิ สามี-ภรรยา หรือ พ่อ-ลูก เป็นต้น

2.2 ค่าที่ได้จากการคำนวณเรียกว่าสัมประสิทธิ์สหสัมพันธ์ ที่มีค่าอยู่ระหว่าง -1 ถึง 1

2.3 ข้อมูลมีความสัมพันธ์กันทางบวก หมายถึง ถ้าข้อมูลชุดหนึ่งมีค่าสูงขึ้นข้อมูลอีกชุดหนึ่งก็มีค่าสูงขึ้นด้วย หรือถ้าข้อมูลชุดหนึ่งมีค่าต่ำลงแล้วข้อมูลอีกชุดหนึ่งก็จะมีค่าต่ำลงด้วย(ตามกัน)

2.4 ข้อมูลมีความสัมพันธ์กันทางลบ หมายถึง ถ้าข้อมูลชุดหนึ่งมีค่าสูงขึ้นแล้วข้อมูลอีกชุดหนึ่งจะมีค่าต่ำลง(ขัดแย้งกัน)

2.5 ข้อมูลมีความสัมพันธ์เป็นศูนย์ หมายถึง ข้อมูลชุดหนึ่งจะสูงขึ้นหรือต่ำลงจะไม่เกี่ยวข้องกับการสูงขึ้นหรือต่ำลงของข้อมูลอีกชุดหนึ่ง ดังแสดงด้วยกราฟในภาพที่ 10.8

หมายเหตุ ลักษณะสหสัมพันธ์ในข้อ 2.3 และ 2.4 จะเป็นจริง เมื่อข้อมูลมีความสัมพันธ์เชิงเส้นตรง

3. ประเภทสถิติในการหาค่าสหสัมพันธ์

ในการศึกษาความสัมพันธ์ระหว่างตัวแปรในการวิจัยใด ๆ จำแนกได้ดังนี้(Runyon and Harber,1980 : 119)

3.1 การศึกษาความสัมพันธ์ระหว่างตัวแปร 2 ตัว ที่มีวัตถุประสงค์ของการวิจัยว่าตัวแปรทั้งสองมีความสัมพันธ์กันหรือไม่ จำแนกตามระดับของข้อมูล ดังนี้

3.1.1 ถ้าเป็นข้อมูลในระดับอันตรภาคหรืออัตราส่วน ให้ใช้การหาสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน(Pearson Product Moment Correlation Coefficient)

3.1.2 ถ้าเป็นข้อมูลในระดับเรียงลำดับ ให้ใช้การหาสัมประสิทธิ์สหสัมพันธ์ของสเปียร์แมน(The Spearman Rank Order Correlation Coefficient)

3.1.3 ถ้าเป็นข้อมูลในระดับนามบัญญัติ ให้ใช้การหาค่าไคร์สแควร์(χ^2-test)

3.1.4 ถ้าข้อมูลชุดหนึ่งเป็นอันตรภาค ส่วนข้อมูลอีกชุดหนึ่งเป็นนามบัญญัติที่กำหนดเพียง 2 ค่า ให้ใช้การหาสัมประสิทธิ์แบบพอยท์ไบซีเรียล

3.1.5 ถ้าศึกษาความสัมพันธ์ของตัวแปร 2 ตัว แต่มีการควบคุมตัวแปรที่จะมาเกี่ยวข้องควรใช้การหาค่าสัมประสิทธิ์สหสัมพันธ์แบบส่วนย่อย(Partial Correlation Coefficient)

3.2 การศึกษาความสัมพันธ์ของตัวแปรตั้งแต่ 2 ตัวขึ้นไป ในการวิเคราะห์ข้อมูลควรเลือกใช้การหาค่าสัมประสิทธิ์พหุคูณ(Multiple Correlation Coefficient)

4. เกณฑ์พิจารณาค่าสัมประสิทธิ์สหสัมพันธ์

การพิจารณาว่า ตัวแปร 2 ตัวแปร มีความสัมพันธ์กันหรือไม่ และมีความสัมพันธ์กันในระดับใด พิจารณาได้จากค่าสัมประสิทธิ์สหสัมพันธ์ซึ่งมีเกณฑ์กว้างๆ ดังนี้

0.8 ขึ้นไปหรือต่ำกว่า-0.80 มีความสัมพันธ์กันในระดับสูงหรือสูงมาก

0.6 ถึง 0.8 หรือ -0.8 ถึง -0.6 มีความสัมพันธ์กันในระดับค่อนข้างสูง

0.4 ถึง 0.6 หรือ -0.6 ถึง -0.4 มีความสัมพันธ์กันในระดับปานกลาง

0.2 ถึง 0.4 หรือ -0.4 ถึง -0.2 มีความสัมพันธ์กันในระดับค่อนข้างต่ำ

ต่ำกว่า 0.2 หรือมากกว่า-0.2 มีความสัมพันธ์กันในระดับต่ำ

5. คุณสมบัติของค่าสัมประสิทธิ์สหสัมพันธ์

5.1 ค่าสัมประสิทธิ์สหสัมพันธ์ไม่ขึ้นอยู่กับหน่วยการวัดของตัวแปร กล่าวคือการเปลี่ยนแปลงของหน่วยการวัดจะไม่ทำให้ค่าสัมประสิทธิ์สหสัมพันธ์เปลี่ยนแปลง

5.2 ค่าสัมประสิทธิ์สหสัมพันธ์จะมีค่าอยู่ระหว่าง -1.00 ถึง 1.00 เท่านั้น

สาระสำคัญบทที่ 10 สถิติเชิงพรรณนา

ในการเรียนรู้บทนี้มีสาระสำคัญ ดังนี้

1. ความหมายของสถิติ 1) “ข้อมูลสถิติ”หมายถึง ตัวเลขที่เป็นข้อเท็จจริงของประเด็นที่สนใจ 2) “สถิติศาสตร์”หมายถึง ศาสตร์ที่เกี่ยวกับวิธีการที่ใช้ในการศึกษาข้อมูลที่เรียกว่า “ระเบียบวิธีการทางสถิติ” 3) “ค่าสถิติ”หมายถึง ค่าตัวเลขที่คำนวณได้จากข้อมูลของกลุ่มตัวอย่าง และ 4) “วิชาสถิติ” หมายถึง วิชาวิทยาศาสตร์แขนงหนึ่งที่มีเนื้อหาสาระและรากฐานจากวิชาคณิตศาสตร์ และตรรกศาสตร์ที่จะนำไปใช้ในการตัดสินใจภายใต้สถานการณ์ที่ไม่แน่นอน

2. สถิติเชิงบรรยาย เป็นสถิติที่มุ่งนำเสนอสารสนเทศเพื่อใช้บรรยายสรุปลักษณะของตัวแปรในกลุ่มตัวอย่างหรือประชากรว่าเป็นอย่างไร มีสถิติที่ใช้ดังนี้ 1) การแจกแจงความถี่ และการนำเสนอด้วยตาราง กราฟ และแผนภูมิ 2) การวัดแนวโน้มเข้าสู่ส่วนกลาง 3) การวัดการกระจายของข้อมูล 4) การวัดความสัมพันธ์ระหว่างตัวแปร และ5) การจัดตำแหน่งและเปรียบเทียบ อาทิ สัดส่วน ร้อยละ หรือเปอร์เซนไทล์

3. ข้อมูล หมายถึง ข้อเท็จจริง สารสนเทศความรู้ ความคิด ความรู้สึก พฤติกรรมการแสดงออก การกระทำ หลักฐานหรือคุณสมบัติของบุคคลหรือสิ่งของที่เก็บรวบรวมมาใช้ประโยชน์ในการตอบปัญหาการวิจัยทั้งในลักษณะข้อมูลเชิงปริมาณ(คะแนนที่ได้)และเชิงคุณภาพ(พฤติกรรมหรือปรากฏการณ์ที่เกิดขึ้น)

4. ข้อมูลเชิงปริมาณที่เป็นตัวเลขในการนำมาวิเคราะห์เพื่อสรุปผลการวิจัยที่จำแนกได้ 4 ระดับ ดังนี้ 1) ข้อมูลระดับนามบัญญัติ เป็นข้อมูลที่ใช้กำหนดชื่อ คุณลักษณะ/พฤติกรรมที่ได้จากการวัด/เก็บรวบรวมข้อมูล 2) ข้อมูลระดับเรียงลำดับ เป็นข้อมูลที่สามารถนำมาเรียงลำดับจากมากไปหาน้อยหรือจากน้อยไปหามากตามเกณฑ์ไม่สามารถระบุได้ว่าแต่ละลำดับแตกต่างกันเท่าไร และความแตกต่างนั้นเท่ากันหรือไม่ 3) ข้อมูลระดับช่วง หรืออันตรภาค เป็นข้อมูลที่สามารถนำมาเรียงลำดับได้ และระบุความแตกต่างระหว่างลำดับที่ได้ และความแตกต่างในแต่ละลำดับที่มีค่าเท่ากัน และเป็นข้อมูลที่ไม่มีระดับศูนย์แท้ และ4) ข้อมูลระดับอัตราส่วน เป็นข้อมูลที่มีคุณสมบัติทางกายภาพสมบูรณ์ที่สุด โดยที่ข้อมูลสามารถระบุความแตกต่างระหว่างกลุ่ม/ลำดับที่ หรือนำมาเปรียบเทียบความมากหรือน้อยมีค่าและช่วงเท่ากัน และมีศูนย์แท้

5. การวิเคราะห์ข้อมูล เป็นวิธีการหาความหมายจากข้อมูล ที่ประกอบด้วยการจัดหมวดหมู่ จัดระเบียบ จัดกระทำ และสรุปย่อ เพื่อให้ได้คำตอบของปัญหาการวิจัย โดยใช้หรือไม่ใช้สถิติก็ได้ หรือ การวิเคราะห์ข้อมูล เป็นวิธีการตรวจสอบทฤษฎีโดยการศึกษาความสัมพันธ์ระหว่างตัวแปรที่เป็นข้อมูลเชิงประจักษ์ หรือเป็นการอธิบายความแปรปรวนว่าตัวแปรที่ศึกษามีความแปรปรวนหรือลักษณะที่แตกต่างกันเนื่องมาจากตัวแปรใดบ้าง

6. การแจกแจงความถี่ เป็นการนำข้อมูลมาจัดหมวดหมู่ ประเภท หรือนำเสนอเป็นตาราง แจกแจงความถี่ เพื่อแสดงจำนวนของข้อมูลว่าแต่ละข้อมูลนั้นมีกี่จำนวน ที่เป็นระเบียบ ง่าย ในการศึกษา และการนำไปใช้ จำแนกเป็น การแจกแจงความถี่ที่ไม่ได้จัดหมวดหมู่และการแจกแจงความถี่ที่แจกแจงข้อมูลแล้ว

7. การวัดแนวโน้มเข้าสู่ส่วนกลาง เป็นการหาค่ากลางที่เป็นตัวแทนของข้อมูลแต่ละชุด เพื่อใช้สรุปผลของข้อมูลในภาพรวม โดยค่ากลางที่นิยมใช้ มี 3 ค่า ได้แก่ 1)ค่าเฉลี่ยเลขคณิต เป็นค่าที่ได้จากผลบวกของข้อมูลทุกจำนวนแล้วเฉลี่ยด้วยจำนวนข้อมูลทั้งหมด 2)ค่ามัธยฐานที่ เป็นค่าของข้อมูลที่มีตำแหน่งอยู่กึ่งกลางของข้อมูล เมื่อได้เรียงลำดับข้อมูลแล้ว โดยจะใช้กับข้อมูลในระดับมาตราเรียงลำดับ หรือข้อมูลที่มีการแจกแจงในลักษณะเบ้ และ 3) ฐานนิยม เป็นค่าของข้อมูลที่มีค่าซ้ำกันมากที่สุด หรือข้อมูลที่มีความถี่สูงสุดในข้อมูลชุดนั้น ๆ

8. การวัดการกระจาย เป็นวิธีการหาค่าที่ใช้แสดงความแตกต่างของข้อมูลในแต่ละชุด โดยวิธีการวัดการกระจายที่นิยมใช้ ได้แก่ 1) พิสัย ค่าที่แสดงความแตกต่างระหว่างค่าสูงสุดและค่าต่ำสุดของข้อมูลชุดนั้น ๆ 2) ส่วนเบี่ยงเบนควอไทล์ เป็นค่าที่ได้จากค่าเฉลี่ยของความแตกต่างระหว่างควอไทล์ที่ 3 และควอไทล์ที่ 1 3) ส่วนเบี่ยงเบนเฉลี่ย เป็นค่าเฉลี่ยของความเบี่ยงเบนของคะแนนแต่ละตัวจากค่าเฉลี่ยเลขคณิตของข้อมูลชุดเดียวกัน 4) ส่วนเบี่ยงเบนมาตรฐาน เป็นค่าของการวัดการกระจายที่ได้หาได้จากรากที่สองของคะแนนทุกค่าที่เบี่ยงเบนไปจากค่าเฉลี่ยของข้อมูลยกกำลังสอง 5) ความแปรปรวน เป็นค่าเฉลี่ยของผลรวมของข้อมูลทุกตัวที่เบี่ยงเบนไปจากค่าเฉลี่ยของข้อมูลยกกำลังสอง หรือ เป็นค่าส่วนเบี่ยงเบนมาตรฐานยกกำลังสองและ 6)สัมประสิทธิ์ความผันแปรที่เป็นการวัดการกระจายของข้อมูลที่คำนวณหาได้จากอัตราส่วนที่คิดเป็นร้อยละระหว่างส่วนเบี่ยงเบนมาตรฐานต่อค่าเฉลี่ยของข้อมูล

9. การวัดตำแหน่งและการเปรียบเทียบ เป็นวิธีการหาค่าที่ระบุตำแหน่งของข้อมูล หรือแสดงการเปรียบเทียบระหว่างข้อมูล มีดังนี้ 1)อัตราส่วน เป็นการนำข้อมูลเชิงปริมาณ 2 ตัว/กลุ่มมาเปรียบเทียบในลักษณะของข้อมูลหนึ่งเป็นกี่เท่าของข้อมูลอีกสิ่งหนึ่ง หรือ ข้อมูลที่สนใจต่อข้อมูลทั้งหมด 2)ร้อยละ เป็นอัตราส่วนที่เปรียบเทียบจำนวนที่สนใจกับกับจำนวนทั้งหมดที่เป็นจำนวนเต็มร้อย 3) เปอร์เซนไทล์ เป็นตำแหน่งของข้อมูลที่ระบุให้ทราบว่ามีข้อมูลที่ต่ำกว่าข้อมูลนั้น ๆ คิดเป็นร้อยละเท่าไร

10. สหสัมพันธ์ เป็นการศึกษาที่สามารถระบุว่าข้อมูลจากตัวแปรที่ศึกษามีความสัมพันธ์กันหรือไม่เท่านั้นที่หาได้จากสิ่งเดียวกันเป็นคู่ ๆ แต่จะไม่สามารถใช้ทำนายตัวแปรหนึ่ง ในกรณีที่ทราบค่าของอีกตัวแปรหนึ่งได้ ที่จำแนกความสัมพันธ์ออกเป็น 3 ลักษณะดังนี้ 1)ความสัมพันธ์กันทางบวก หมายถึง ถ้าข้อมูลชุดหนึ่งมีค่าสูงขึ้นข้อมูลอีกชุดหนึ่งก็มีค่าสูงขึ้นด้วย หรือถ้าข้อมูลชุดหนึ่งมีค่าต่ำลงแล้วข้อมูลอีกชุดหนึ่งก็จะมีค่าต่ำลงด้วย(ตามกัน) 2)ความสัมพันธ์กันทางลบ หมายถึง ถ้าข้อมูลชุดหนึ่งมีค่าสูงขึ้นแล้วข้อมูลอีกชุดหนึ่งจะมีค่าต่ำลง(ขัดแย้งกัน) และ3) ความสัมพันธ์เป็นศูนย์ หมายถึง ข้อมูลชุดหนึ่งจะสูงขึ้นหรือต่ำลงจะไม่เกี่ยวข้องกับการสูงขึ้นหรือต่ำลงของข้อมูลอีกชุดหนึ่ง

11. ในการศึกษาความสัมพันธ์ระหว่างตัวแปรในการวิจัยใด ๆ จำแนกได้ดังนี้ 1) ถ้าเป็นข้อมูลในระดับอันตรภาคหรืออัตราส่วน ให้ใช้การหาสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน 2) ถ้าเป็นข้อมูลในระดับเรียงลำดับ ให้ใช้การหาสัมประสิทธิ์สหสัมพันธ์ของสเปียร์แมน 3) ถ้าเป็นข้อมูลในระดับนามบัญญัติ ให้ใช้การหาค่าไคร์สแควร์(χ^2-test) 4) ถ้าข้อมูลชุดหนึ่งเป็นอันตรภาค ส่วนข้อมูลอีกชุดหนึ่งเป็นนามบัญญัติที่กำหนดเพียง 2 ค่า ให้ใช้การหาสัมประสิทธิ์แบบพอยท์ไบซีเรียล 5) ถ้าศึกษาความสัมพันธ์ของตัวแปร 2 ตัว แต่มีการควบคุมตัวแปรที่จะมาเกี่ยวข้องควรใช้การหาค่าสัมประสิทธิ์สหสัมพันธ์แบบส่วนย่อย และ 6)การศึกษาความสัมพันธ์ของตัวแปรตั้งแต่ 2 ตัวขึ้นไป ควรเลือกใช้การหาค่าสัมประสิทธิ์พหุคูณ

คำถามเชิงปฏิบัติการบทที่ 10 สถิติเชิงพรรณนา

คำชี้แจง ให้ตอบคำถามจากประเด็นคำถามที่กำหนดให้อย่างถูกต้องและชัดเจน

1. ในการวิจัยจะใช้สถิติเชิงบรรยาย โดยมีวัตถุประสงค์ใด
2. ให้พิจารณาข้อมูลที่กำหนดให้ ว่าเป็นข้อมูลในระดับใด
   2.1 ทักษะการคิดคำนวณของผู้เรียน
   2.2 จำเป็นได้รับคะแนนสูงสุดในห้องเรียน
   2.3 หมายเลขโทรศัพท์ของจำนอง คือ 01-511894
   2.4 ในปีนี้อุณภูมิต่ำสุดเท่ากับ 10 องศาเซนเซียส
   2.5 จำเนียรมีความพึงพอใจในผลการปฏิบัติงานของตนเอง
3. จากข้อมูลที่กำหนดให้

4	5	4	6	5	7	8	9	5	6	8
7	9	9	8	7	5	6	4	8	7	6

ให้หาความถี่ ร้อยละ ค่าเฉลี่ยเลขคณิต มัธยฐาน ฐานนิยม พิสัย ส่วนเบี่ยงเบนควอไทล์ ส่วนเบี่ยงเบนเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และความแปรปรวนของข้อมูล

4. จากตารางข้อมูลที่กำหนดให้

คะแนน	จำนวนคน
1-5	3
6-10	8
11-15	12
16-20	10
21-25	4
26-30	3

ให้หาค่าเฉลี่ยเลขคณิต มัธยฐาน ฐานนิยม พิสัย ส่วนเบี่ยงเบนควอไทล์ ส่วนเบี่ยงเบนเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และความแปรปรวนของข้อมูล

4. จากการสรุปผลเบื้องต้นของข้อมูล 2 ชุด มีดังนี้

ข้อมูลชุดที่ 1 มีค่าเฉลี่ยเลขคณิต เท่ากับ 100 และมีความแปรปรวนเท่ากับ 17.64

ข้อมูลชุดที่ 2 มีค่าเฉลี่ยเลขคณิต เท่ากับ 75 และมีความแปรปรวนเท่ากับ 12.25

ให้ท่านได้สรุปว่าในข้อมูลชุดใดมีความแตกต่างภายในกลุ่มมากกว่ากัน

5. จากตารางข้อมูลที่กำหนดให้จำนวน 2 ชุด

คนที่	ข้อมูลชุดที่ 1	ข้อมูลชุดที่ 2
1	6	8
2	7	7
3	8	6
4	6	6
5	7	7
6	8	8
7	9	8
8	8	7

ให้หาสัมประสิทธิ์สหสัมพันธ์ของข้อมูลโดยวิธีการของเพียร์สันและสเปียร์แมน