บทที่ 9
สถิติเบื้องต้นสำหรับการวัดและการประเมินพัฒนาการระดับปฐมวัย

เนื้อหา

เนื้อหาที่จะกล่าวถึงในบทที่ 9 มีรายละเอียด ดังนี้

9.1 ประเภทและระดับของข้อมูล

9.2 สัญญลักษณ์ผลรวมทางพีชคณิต

9.3 การวัดแนวโน้มเข้าสู่ส่วนกลาง

9.4 การวัดการกระจายของข้อมูล

9.5 การวัดความสัมพันธ์

9.1 ประเภทและระดับของข้อมูล

1. ประเภทของข้อมูล

ข้อมูล (data) หมายถึงข้อเท็จจริงเกี่ยวกับเรื่องต่าง ๆ ที่เราสนใจ ซึ่งอาจเป็น ข้อเท็จจริงที่เป็นตัวเลข เช่น จำนวนประชากร คะแนน น้ำหนัก ฯลฯ หรืออาจเป็นข้อเท็จจริงที่อยู่ในรูปคุณลักษณะ หรือคุณสมบัติ เช่น เพศ ศาสนา อาชีพ ระดับการศึกษา ฯลฯ

การแบ่งประเภทของข้อมูล อาจแบ่งได้หลายวิธี ทั้งนี้ขึ้นอยู่กับเกณฑ์ที่ใช้ในการพิจารณา ดังนี้

ก. จำแนกตามการจัดกระทำข้อมูล สามารถแบ่งออกเป็น 2 ประเภท คือ

1. ข้อมูลดิบ (raw data) เป็นข้อมูลที่ได้จากการเก็บรวบรวมข้อมูลโดยตรง โดยที่ยังไม่ได้นำมาจัดหมวดหมู่หรือจัดกระทำใด ๆ ทั้งสิ้น

2. ข้อมูลที่จัดหมวดหมู่ (group data) เป็นข้อมูลที่เกิดจากการนำข้อมูลดิบมาจัดกระทำให้เป็นหมวดหมู่อย่างมีระเบียบ เช่น นำมาแจกแจงความถี่ซึ่งจะเป็นการช่วยให้เกิดความสะดวกในการนำไปคำนวณหรือนำไปใช้

ข. จำแนกตามลักษณะข้อมูล สามารถจำแนกข้อมูลออกเป็น 2 ประเภท คือ

1. ข้อมูลเชิงปริมาณ (Quantitative data) เป็นข้อมูลที่แสดงปริมาณ หรือขนาดในลักษณะของตัวเลขโดยตรง จำแนกเป็น 2 ประเภทใหญ่ ๆ คือ

1) <u>ข้อมูลแจงนับ</u> คือข้อมูลซึ่งเกิดจากสิ่งของที่นับได้ ซึ่งจะอยู่ในรูปของความถี่ เช่น โต๊ะ 40 ตัว คน 30 คน ข้อมูลทางสังคมส่วนใหญ่จะเป็นข้อมูลแจงนับ

2) <u>ข้อมูลการวัด</u> หรือข้อมูลเมตริก คือข้อมูลที่เกิดจากการวัดสิ่งของและให้ค่าเมตริก หรือมาตรา เช่น วินัยหนัก 60 กก. สาธิตสอบได้คะแนน 70 ข้อมูลทางการศึกษาและจิตวิทยาส่วนใหญ่จะเป็นข้อมูลการวัดทั้งสิ้น

2. ข้อมูลเชิงคุณภาพ (Qualitative data) เป็นข้อมูลที่แสดงถึงคุณลักษณะไม่ได้อยู่ในรูปของตัวเลขโดยตรง เช่น เพศ ศาสนา สถานะภาพสมรส อาชีพ ระดับการศึกษา ภูมิลำเนา เชื้อชาติ เป็นต้น

2. ระดับของข้อมูลหรือมาตราการวัด

การวัดหรือการนับใด ๆ เป็นการกำหนดจำนวนหรือตัวเลขให้กับสิ่งของ หรือเหตุการณ์ จึงจำเป็นต้องมีการกำหนดหน่วยหรือมาตราให้กับตัวเลขหรือข้อมูลเหล่านั้น มาตราการวัดที่นิยมใช้กันมากที่สุดเป็นผลงานของ Steven (1946) ซึ่งแบ่งมาตราการวัดออกเป็น 4 ประเภท คือ

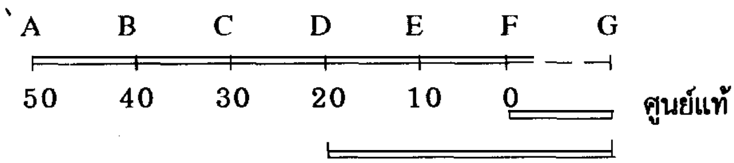
1. มาตรานามบัญญัติ (Nominal Scale)
2. มาตราอันดับ (Ordinal Scale)
3. มาตราอันตรภาค (Interval Scale)
4. มาตราอัตราส่วน (Ratio Scale)

1. มาตรานามบัญญัติ (Nominal Scale) จัดเป็นมาตราการวัดที่หยาบที่สุดเพราะเป็นเพียงการกำหนดตัวเลขหรือสัญญลักษณ์แทนคุณสมบัติอย่างใดอย่างหนึ่งของสิ่งใดสิ่งหนึ่งที่ต้องการจะวัดเพื่อจำแนกให้เห็นว่าสิ่งที่ต้องการจะวัดมีความแตกต่างกัน โดยที่ตัวเลขหรือสัญญลักษณ์เหล่านั้นไม่มีความหมายในเชิงปริมาณ เช่น เพศหญิง-ชาย นักเรียนห้อง 1,2,3,4 หรือพืช-สัตว์ ฯลฯ หรืออาจกำหนดให้เลข 1 แทนผู้ชาย และเลข 2 แทนผู้หญิง เป็นต้น เราจะเอาเลขในระบบนี้ไปบวก ลบ คูณ หาร กันไม่ได้ จะทำได้ก็เพียงแต่นับว่าแต่ละพวกมีจำนวนเท่าไรเท่านั้น

2. มาตราอันดับ (Ordinal Scale) การวัดแบบนี้มีความหมายชัดเจนกว่า การวัดในมาตรานามบัญญัติ กล่าวคือ เราสามารถเรียงคุณลักษณะของบุคคลหรือสิ่งของที่เราต้องการจะวัดจากมากไปหาน้อย หรือน้อยไปหามากได้ แต่เราไม่สามารถบอกได้ว่าความแตกต่างแต่ละอันนั้นต่างกันเท่าไร เช่น สมศักดิ์สอบได้ที่ 1 สิโรตม์สอบได้ที่ 2 สมศรีสอบได้ที่ 3 เราไม่ สามารถบอกได้ว่าสมศักดิ์เก่งเป็นกี่เท่าของสิโรตม์ หรือเราไม่สามารถบอกได้ว่าสมศักดิ์เก่งกว่าสิโรตม์เป็นปริมาณ

เท่ากับสิโรตม์เก่งกว่าสมศรี จะเห็นได้ว่า ที่ 1,2,3,...ไม่ได้แทนปริมาณและขนาดที่แท้จริง แต่จะเป็นเครื่องบอกถึงอันดับหรือตำแหน่งเท่านั้น ดังนั้น percentile rank หรืออันดับในการจัดประกวดทางด้านต่าง ๆ นั้นจัดอยู่ในมาตรานี้ทั้งสิ้น

3. มาตราอันตรภาค (Interval Scale) การวัดในระดับนี้เป็นการวัดที่สูงกว่าการวัดในมาตรานามบัญญัติและมาตราอันดับ ตรงที่แต่ละหน่วยมีขนาดเท่า ๆ กัน สามารถเปรียบเทียบได้ว่าของสิ่งหนึ่งมากกว่าหรือน้อยกว่าอีกสิ่งหนึ่งอยู่เท่าไร แต่ไม่สามารถบอกได้ว่าของสองสิ่งนั้นมากกว่ากันเป็นกี่เท่า เพราะการวัดในระดับนี้ไม่มีศูนย์แท้ (Absolute zero) คะแนนที่ได้จากการทดสอบจัดอยู่ในมาตราอันตรภาค เช่น วนิดาสอบวิทยาศาสตร์ได้ 60 คะแนน วินิจสอบได้ 50 คะแนน วินัยสอบได้ 30 คะแนน เราสามารถบอกได้ว่าวนิดาสอบได้มากกว่าวินิจ 10 คะแนน ส่วนวินิจสอบได้มากกว่าวินัย 20 คะแนน และเรายังสามารถบอกได้อีกว่าระยะห่างของคะแนนระหว่างวินิจกับวินัย (50-30) มีค่าเป็น 2 เท่า ของระยะห่างของคะแนนระหว่างวนิดากับวินิจ (60-50) แต่เราไม่สามารถบอกได้ว่าวนิดามีความรู้เป็น 2 เท่าของวินัย ทั้งนี้เนื่องจากเราไม่รู้ว่าศูนย์แท้อยู่ที่ไหน และแต่ละคนห่างจากศูนย์แท้เท่าไร ดังนั้นถ้าวดีสอบวิทยาศาสตร์ได้ 0 คะแนน มิได้หมายความว่าวดีไม่มีความรู้วิทยาศาสตร์เลย แต่มีความหมายเพียงแต่ว่าวดีทำข้อสอบวิทยาศาสตร์ไม่ได้เลยต่างหาก การใช้เทอร์โมมิเตอร์วัดอุณหภูมิก็จัดอยู่ในมาตรานี้ เพราะอุณหภูมิ 0° C มิได้หมายความว่าไม่มีความร้อนอยู่เลย เราอาจเขียนรูปแสดงให้เห็นถึงการวัดในมาตราอันตรภาคได้ดังนี้



เมื่อจุด F คือจุดที่นักเรียนสอบได้ 0 คะแนน

จุด G คือจุดที่นักเรียนไม่มีความรู้จริง ๆ

ดังนั้น ระยะระหว่าง D ถึง G จึงไม่มีค่าเป็น 2 เท่าของระยะระหว่าง F ถึง G (ดูรูปประกอบ) นั่นคือเราไม่สามารถสรุปได้ว่า D มีความรู้เป็น 20 เท่าของ F เราบอกได้แต่ เพียงว่าระยะห่างจาก C ถึง E มากเป็นสองเท่าของระยะห่างจาก D ถึง E

4. มาตราอัตราส่วน (Ratio Scale) การวัดในมาตรานี้เป็นการวัดที่สมบูรณ์กว่าการวัดใน 3 มาตราที่กล่าวมาแล้วข้างต้น ส่วนใหญ่เป็นการวัดทางวิทยาศาสตร์ เช่น การวัดความเร็ว ความยาว น้ำหนัก หรือส่วนสูง ฯลฯ การวัดในมาตรานี้มีคุณสมบัติเพิ่มเติมจากการวัด

ทั้ง 3 มาตรา ที่กล่าวมาแล้ว คือมีศูนย์แท้ เช่น ความยาว 0 นิ้ว แปลว่าไม่มีความยาวเลย หรือคนที่หนัก 40 ก.ก. ย่อมหนักเป็นสองเท่าของคนที่หนัก 20 ก.ก. เนื่องจากตัวเลขที่ได้จากการวัดในมาตรานี้ แต่ละหน่วยมีขนาดเท่ากันและมีศูนย์แท้ จึงสามารถนำตัวเลขในมาตรานี้มาเปรียบเทียบกันได้ทั้งการบวก ลบ คูณ และหาร

9.2 สัญญลักษณ์ผลรวมทางพีชคณิต

1. สัญญลักษณ์และความหมาย

ในการแสดงค่าการบวกกันของเลขหลาย ๆ จำนวนนั้น ถ้าจะเขียนแสดงตรง ๆ อาจจะดูยืดยาว โดยเฉพาะอย่างยิ่งถ้าเราต้องการแสดงสูตรซึ่งเป็นผลบวกของเทอมต่าง ๆ หลายเทอม ยิ่งจะมีความยากลำบากในการแสดงค่าการบวกกันของเลขจำนวนต่าง ๆ ในแต่ละเทอม ดังนั้นเพื่อความสะดวกจึงมีการกำหนดสัญญลักษณ์แสดงผลบวกขึ้นเพื่อให้สูตรต่าง ๆ สั้นลง สัญญลักษณ์ที่กำหนดขึ้นมีสูตรทั่ว ๆ ไปดังนี้

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + x_3 + \dots\dots + x_n$$

Σ เป็นอักษรกรีก อ่านว่า “ซิกม่า” (Sigma) นำมาใช้แทนผลรวม ดังนั้น

$\sum_{i=1}^{n} x_i$ อ่านว่า ซิกม่าของ x_i หรือผลรวมของ x_i เมื่อ i มีค่าเท่ากับ 1 ถึง n และ i เป็นตัวแปรที่จะใช้ระบุว่า x เป็นตัวที่เท่าใดและมีอยู่กี่ตัว เช่น

$$\sum_{i=1}^{3} x_i = x_1 + x_2 + x_3$$

$$\sum_{i=1}^{5} x_i = x_1 + x_2 + x_3 + x_4 + x_5$$

$$\sum_{i=3}^{5} x_i = x_3 + x_4 + x_5$$

ตัวอย่าง ถ้า $x_1 = 3, x_2 = 5, x_3 = 7, x_4 = 4$ จงหาค่าของ $\sum_{i=2}^{4} x_1$

วิธีทำ $\sum_{i=2}^{4} x_i = x_2 + x_3 + x_4$

$= 5 + 7 + 4$

$= 16$

2. ทฤษฎีเกี่ยวกับผลรวมทางพีชคณิต

$\boxed{\textit{ทฤษฎี 1}}$ $\sum_{i=1}^{n} x_i^2 = x_1^2 + x_2^2 + x_3^2 + \ldots\ldots\ldots\ldots + x_n^2$

ตัวอย่าง ถ้า $x_1 = 3, x_2 = 5, x_3 = 7, x_4 = 4, x_5 = 2$ แล้ว $\sum_{i=1}^{5} x_i^2$ มีค่าเท่าใด

วิธีทำ $\sum_{i=1}^{5} x_i^2 = x_1^2 + x_2^z + x_3^z + x_4^z + x_5^2$

$= 3^2 + 5^2 + 7^2 + 4^2 + 2^2$

$= 9 + 25 + 49 + 16 + 4$

$= 103$

$\left(\textit{ทฤ} \sum_{i=1}^{n} x_i\right)^2 = (x_1 + x_2 + x_3 + \ldots\ldots\ldots\ldots + X_n)^2$

ตัวอย่าง ถ้า $x_1 = 3, x_2 = 5, x_3 = 7, x_4 = 4, x_5 = 2$ แล้ว $\left(\sum_{i=2}^{4} x_i\right)^2$ มีค่าเท่าใด

วิธีทำ $\left(\sum_{i=2}^{4} x_i\right)^2 = (x_2 + x_3 + x_4)^2$

$= (5 + 7 + 4)^2$

$= (16)^2$

$= 256$

ทฤษฎี 3 ถ้า a เป็นค่าคงที่ จะได้ว่า

$$\sum_{i=1}^{n} a = na$$

ตัวอย่าง ถ้า a มีค่าเท่ากับ 5 ให้หา $\sum_{i=1}^{7} 5$

วิธีทำ $\sum_{i=1}^{7} 5 = 7 \times 5$

$= 35$

ทฤษฎี 4 ถ้า a เป็นค่าคงที่ และ X เป็นตัวแปร จะได้ว่า

$$\sum_{i=1}^{n} a x_i = a\sum_{i=1}^{n} x_i$$

ตัวอย่าง ถ้า a มีค่าเท่ากับ 5 และ $x_1 = 2, x_2 = 5, x_3 = 1, x_4 = 4$ ให้หา $\sum_{i=1}^{4} 5 x_i$

วิธีทำ $\sum_{i=1}^{4} 5 x_i = 5\sum_{i=1}^{4} x_i$

$= 5 (x_1 + x_2 + x_3 + x_4)$

$= 5 (2 + 5 + 1 + 4)$

$= 5 (12)$

$= 60$

ทฤษฎี 5 $$\sum_{i=1}^{n}\left(x_i + y_i\right) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i$$

(เมื่อ x และ y ต่างก็เป็นตัวแปร)

ตัวอย่าง tii $x_1 = 2,\ x_2 = 5,\ x_3 = 0$ และ $y_1 = 1,\ y_2 = 4,\ y_3 = 2$

จงหาค่าของ $\sum_{i=1}^{3}\left(x_1 + y_i\right)$

วิธีทำ $\sum_{i=1}^{3}(x_i + y_i) = \sum_{i=1}^{3} x_i + \sum_{i=1}^{3} y_i$

$= (2 + 5 + 0) + (1 + 4 + 2)$

$= 7 + 7$

$= 14$

ทฤษฎี 6 ถ้า a และ b เป็นค่าคงที่ และ x เป็นตัวแปรแล้ว จะได้ว่า

$$\sum_{i=1}^{n}(ax_i + b) = a\sum_{i=1}^{n} x_i + nb$$

ตัวอย่าง ถ้า $x_1 = 3$, $x_2 = 5$, $x_3 = 1$, $x_4 = 0$ และ a มีค่าเท่ากับ 5, b มีค่าเท่ากับ 12 แล้วจงหาค่าของ $\sum_{i=1}^{4}(5x_i + 12)$

วิธีทำ $\sum_{i=1}^{4}(5x_i + 12) = 5\sum_{i=1}^{4} x_i + (4 \times 12)$

$= 5(3 + 5 + 1 + 0) + 48$

$= 5(9) + 48$

$= 45 + 48$

$= 93$

9.3 การวัดแนวโน้มเข้าสู่ส่วนกลาง

การวัดแนวโน้มเข้าสู่ส่วนกลาง (Measure of central tendency) เป็นวิธีการหาค่าที่เป็นตัวแทนของข้อมูลชุดที่จะศึกษา เพื่อประโยชน์ในการบรรยายหรือพรรณนาลักษณะของข้อมูลชุดที่จะศึกษา หรือเพื่อประโยชน์ในการเปรียบเทียบลักษณะของข้อมูลชุดใดชุดหนึ่งกับชุดอื่น ๆ

การหาค่าตัวแทนของข้อมูล โดยใช้วิธีการวัดแนวโน้มเข้าสู่ส่วนกลางมีหลายวิธี แต่วิธีที่นิยมใช้กันแพร่หลาย มี 3 วิธี

1. มัชฌิมเลขคณิต หรือ ค่าเฉลี่ย (Mean)
2. มัธยฐาน (Median)
3. ฐานนิยม (Mode)

1. มัชฌิมเลขคณิต (หรือค่าเฉลี่ย)

มัชฌิมเลขคณิตของข้อมูลชุดใดเกิดจากการหาผลรวมของทุก ๆ รายการในข้อมูลชุดนั้นหารด้วยจำนวนรายการของข้อมูลชุดนั้น

มัชฌิมเลขคณิต นิยมเรียกสั้น ๆ ว่า "ค่าเฉลี่ย"

ค่าเฉลี่ยของกลุ่มตัวอย่าง เขียนแทนด้วย $\overline{X}$

ค่าเฉลี่ยของประชากร เขียนแทนด้วย μ

สูตรในการหาค่าเฉลี่ยมี 3 สูตร คือ

1) กรณีที่เป็นข้อมูลดิบ

สูตรที่ใช้ในการหาค่าเฉลี่ย คือ

$$\overline{X} = \frac{\Sigma X}{N} \text{ ....(1)}$$

เมื่อ $\overline{X}$ = มัชฌิมเลขคณิตของกลุ่มตัวอย่างหรือค่าเฉลี่ย

X = ข้อมูลแต่ละตัว

Σ = ผลรวมของ

N = จำนวนข้อมูล

ตัวอย่าง ข้อมูลชุดหนึ่งมีค่าคะแนนดังนี้

4, 6, 6, 7, 8, 9, 12 จงหามัชฌิมเลขคณิต (ค่าเฉลี่ย)

วิธีทำ

$$\Sigma X = 4 + 6 + 6 + 7 + 8 + 9 + 12 = 52$$

$$N = 7$$

ดังนั้นค่ามัชฌิมเลขคณิต $\overline{X} = \frac{52}{7} = 7.42$

สูตร (1) ถือเป็นสูตรพื้นฐานในการคำนวณค่ามัชฌิมเลขคณิตหรือค่าเฉลี่ย ซึ่งนิยมใช้ในกรณีที่ข้อมูลอยู่ในลักษณะเป็นข้อมูลดิบ อย่างไรก็ตาม ถ้าข้อมูลอยู่ในลักษณะของการจัดหมวดหมู่หรือจัดกลุ่ม ก็อาจดัดแปลงสูตร (1) ให้สะดวกต่อการคำนวณโดยใช้สูตร (2) ดังนี้

2) กรณีข้อมูลถูกจัดกลุ่มหรือจัดหมวดหมู่

สูตรที่ใช้ในการหาค่าเฉลี่ย คือ

$$\overline{X} = \frac{\Sigma fx}{\Sigma f} = \frac{\Sigma fx}{N} \quad \text{.............................(2)}$$

เมื่อ $\overline{X}$ = มัชฌิมเลขคณิต (หรือค่าเฉลี่ย) ของกลุ่มตัวอย่าง

f = จำนวนความถี่ของข้อมูล หรือความถี่ของคะแนนแต่ละตัว

Σf = ผลรวมของความถี่ทั้งหมด (ซึ่งเท่ากับ N)

X = ข้อมูลแต่ละตัว

Σfx = ผลรวมของค่าผลคูณระหว่างความถี่ของข้อมูลแต่ละตัว

N = จำนวนข้อมูล

ตัวอย่าง การคำนวณหาค่ามัชฌิมเลขคณิต (ค่าเฉลี่ย) ของข้อมูลที่จัดกลุ่ม

X	f	fx
57	1	57
52	1	52
47	3	141
42	4	168
37	6	222
32	7	224
27	12	324
22	6	132
17	8	136
12	2	24
รวม	Σf = 50 = N	Σfx = 1480

$$\overline{X} = \frac{\Sigma fx}{N} = \frac{1480}{50} \approx 29.60$$

3) กรณีข้อมูลถูกจัดหมวดหมู่แบบอันตรภาคชั้น

สูตรที่ใช้คำนวณหาค่าเฉลี่ยใช้สูตรที่ (2) คือ

$$\overline{X} = \frac{\Sigma fx}{\Sigma f} = \frac{\Sigma fx}{N}$$

ตัวอย่าง

อันตรภาคชั้น	f	x	fx
13-15	1	14	14
10 - 12	4	11	44
7-9	2	8	16
4-6	6	5	30
l-3	3	2	6
	$\Sigma f = 16 = N$		$\Sigma fx = 110$

$$\overline{X} = \frac{\Sigma fx}{\Sigma f} = \frac{110}{116} = 6.88$$

2. มัธยฐาน

"มัธยฐาน" คือ จุดบนมาตราการวัดซึ่งมีจำนวนครึ่งหนึ่งของข้อมูลอยู่เหนือ และอีกครึ่งหนึ่งอยู่ใต้ โดยที่ข้อมูลชุดนั้นได้มีการจัดเรียงค่าตามลำดับแล้ว

เช่น ข้อมูลชุดหนึ่งประกอบด้วย 10, 12, 15, 9, 6 เมื่อจัดเรียงค่าตามลำดับจะได้ 15, 12, 10, 9, 6 เมื่อพิจารณาแล้วข้อมูลชุดนี้มีเลข 10 อยู่กึ่งกลาง ดังนั้น "10" จึงเป็นค่ามัธยฐานของข้อมูลชุดนี้

การหาค่ามัธยฐานทำได้ 2 วิธีคือ

1. กรณีที่เป็นข้อมูลดิบ
2. กรณีที่ข้อมูลจัดเป็นหมวดหมู่

ซึ่งแต่ละกรณีมีวิธีดำเนินการ ดังนี้

1) กรณีที่เป็นข้อมูลดิบ

ถ้าข้อมูลที่ต้องการหาค่ามัธยฐานอยู่ในลักษณะตามธรรมชาติโดยมิได้นำมาจัดหมวดหมู่หรือกลุ่ม การหาค่ามัธยฐานอาจทำได้โดยการสำรวจอย่างง่าย ๆ โดยการนำข้อมูลมาจัดเรียงค่าเสียใหม่ตามลำดับจากมากไปหาน้อยหรือเรียงจากน้อยไปหามากก็ได้ แล้วสำรวจดูว่าข้อมูลใดอยู่ ณ ตำแหน่งกึ่งกลาง นั่นคือ มีข้อมูลครึ่งหนึ่งอยู่เหนือและอีกครึ่งอยู่ใต้ ข้อมูลนั้นคือมัธยฐาน (ใช้สัญลักษณ์ Mdn หรือ M.D.) การหาค่ามัธยฐานโดยการสำรวจอย่างง่าย ๆ นี้เหมาะสำหรับใช้กับข้อมูลที่มีจำนวนรายการไม่มากนัก ดังตัวอย่าง

ตัวอย่าง ข้อมูลชุดหนึ่งมีค่าคะแนนดังนี้

12, 10, 57, 50, 48, 17, 21, 42, 32, 37, 27 จงหาค่ามัธยฐานของคะแนนชุดนี้

วิธีทำ นำข้อมูลมาเรียงใหม่จากน้อยไปหามากดังนี้

10, 12, 17, 21, 27, 32, 37, 42, 48, 50, 57

ฉะนั้น ค่ามัธยฐานของคะแนนชุดนี้ คือ 32

ในกรณีที่ข้อมูลมีจำนวนเป็นเลขคู่ เช่น มีข้อมูลอยู่ 10 รายการ ค่ามัธยฐานของข้อมูลชุดนั้นคือค่าเฉลี่ยของข้อมูลตำแหน่งที่ 5 และที่ 6 ดังตัวอย่าง

ตัวอย่าง ข้อมูลชุดหนึ่งมีค่าคะแนนดังนี้

10, 12, 57, 50, 48, 17, 21, 42, 32, 37, 27, 63 จงหาค่ามัธยฐานของคะแนนชุดนี้

วิธีทำ นำข้อมูลมาเรียงใหม่จากน้อยไปหามากดังนี้

10, 12, 17, 21, 27, 32, 37, 42, 48, 50, 57, 63

เนื่องจากจำนวนข้อมูลมีอยู่ 12 รายการ ซึ่งเป็นเลขคู่ มัธยฐานของข้อมูลชุดนี้คือค่าเฉลี่ยของข้อมูลตำแหน่งที่ 6 และ 7

ฉะนั้น ค่ามัธยฐานของคะแนนชุดนี้เท่ากับ $\frac{32 + 37}{2} = 34.5$

2) กรณีที่ข้อมูลจัดเป็นหมวดหมู่

ถ้าข้อมูลถูกจัดเป็นกลุ่มหรือเป็นหมวดหมู่ การหาค่ามัธยฐานของข้อมูลที่ได้มีการจัดเรียงค่าตามลำดับแล้วสามารถทำได้โดยใช้สูตรคำนวณหาค่าดังนี้

$$\text{มัธยฐาน (Mdn)} = L + \left[\frac{N/2 - F}{f}\right] i \quad \ldots\ldots (3)$$

เมื่อ L = **ขีดจำกัดล่างแท้ของอันตรภาคชั้นที่มัธยฐานตกอยู่**

F = **ผลรวมของความถี่ที่อยู่ใต้ L (ความถี่สะสม)**

f = **ความถี่ของอันตรภาคชั้นที่มัธยฐานตกอยู่**

N = **จำนวนข้อมูล (ผลรวมของความถี่ทั้งหมด)**

i = **ขนาดของอันตรภาคชั้น**

Mdn = **ค่ามัธยฐาน**

ตัวอย่าง **จงคำนวณหาค่ามัธยฐานของคะแนนที่ได้จากการทดสอบวิชาคณิตศาสตร์ของนักเรียน 28 คน โดยคะแนนถูกจัดหมวดหมู่เป็นอันตรภาคชั้น ดังนี้**

ขนาดของชั้น	f	**ความถี่สะสม** (F)
40 - 44	1	28
35 - 39	0	27
30-34	3	27
25 - 29	5	24
20 - 24	3	19
15 - 19	10	16
10 - 14	1	6
5 - 9	1	5
o - 4	4	4
	N = 28	

เนื่องจากข้อมูลมี 28 กรณี ฉะนั้นคะแนนที่มัธยฐานตกอยู่คือ คะแนนตัวที่ 14 (หาได้จาก N/2 = 28/2 = 14) ฉะนั้นมัธยฐานจะตกอยู่ที่ชั้นของความถี่สะสม 16 ซึ่งเป็นชั้นของคะแนน 15 - 19

$L = 14.5$

$N = 28$

$F = 6$

$f = 10$

$i = 5$

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

เมื่อ $L = 14.5$

$N = 28$

$F = 1 + 1 + 4 = 6$

$f = 10$

$i = 5$

แทนค่าสูตร

$$\text{Mdn} = 14.5 + \left[\frac{(28/2 - 6)}{10}\right] \times 5$$

$$= 14.5 + \left[\frac{14 - 6}{10}\right] \times 5$$

$$= 14.5 + \left[\frac{8}{10}\right] \times 5$$

$$= 14.5 + 40$$

$$= 18.5$$

ดังนั้น คะแนนชุดนี้มีค่ามัธยฐานเท่ากับ 18.5

3. ฐานนิยม

"ฐานนิยม" คือจุดบนมาตรการวัดที่มีความถี่มากที่สุด ตามปกติฐานนิยมจะอยู่ใกล้ศูนย์กลางของการแจกแจง ในกรณีที่การแจกแจงมีลักษณะสมมาตร ฐานนิยมจะตกอยู่ที่เดียวกับมัชฌิมเลขคณิต และมัธยฐาน

ตัวอย่าง คะแนนชุดหนึ่งมีดังนี้

10, 12, 12, 15, 18, 18, 18, 25, 30 ให้หาฐานนิยม

วิธีทำ เนื่องจากคะแนน 18 มีความถี่มากที่สุด ดังนั้นฐานนิยมของคะแนนชุดนี้คือ 18

ตัวอย่าง ให้หาฐานนิยมของคะแนนที่ถูกจัดกลุ่มข้างล่างนี้

คะแนน	f
50	3
48	10
43	8
39	14
35	10
31	3
24	2

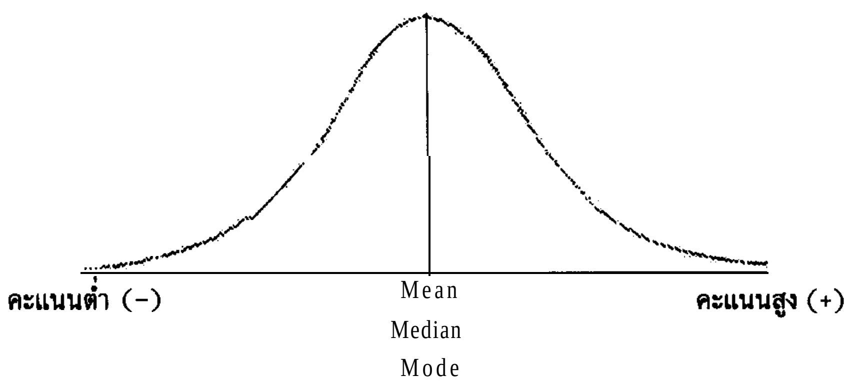
ฐานนิยม = 39

การวัดแนวโน้มเข้าสู่ส่วนกลางที่ใช้ในแต่ละมาตราการวัด

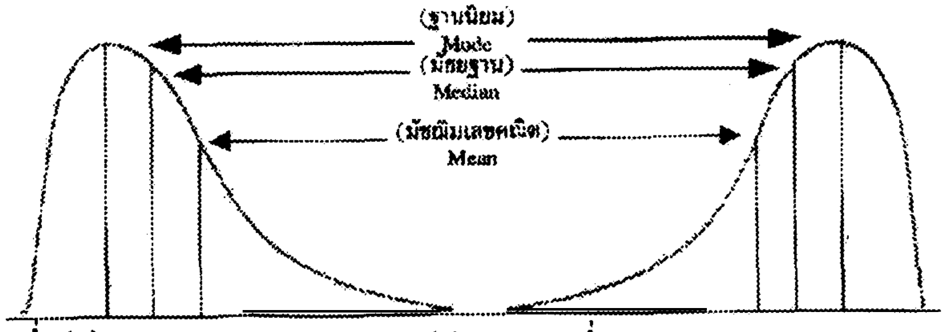
มาตราการวัด	การวัดแนวโน้มเข้าสู่ส่วนกลางที่ใช้
1. นามบัญญัติ (Nominal scale)	ฐานนิยม (Mode)
2. อันดับ (Ordinal scale)	มัธยฐาน ฐานนิยม (Median, Mode)
3. อันตรภาค (Interval scale)	มัชฌิมเลขคณิต มัธยฐาน ฐานนิยม (Mean, Median, Mode)
4. อัตราส่วน (Ratio scale)	มัชฌิมเลขคณิต มัธยฐาน ฐานนิยม (Mean, Median, Mode)

ความสัมพันธ์ระหว่างมัชฌิมเลขคณิต มัธยฐาน และฐานนิยม

ถ้าการกระจายของข้อมูลเป็นแบบปกติ นั่นคือโค้งของการกระจายมีลักษณะสมมาตร แล้ว ค่าสถิติที่ใช้ในการวัดแนวโน้มเข้าสู่ส่วนกลางทั้ง 3 ค่า คือ มัชฌิมเลขคณิต หรือค่าเฉลี่ย มัธยฐาน และฐานนิยม จะมีค่าเท่ากัน (รูป 1) แต่ถ้าการกระจายของข้อมูลมีลักษณะเบ้ ค่ามัชฌิมเลขคณิตจะอยู่ใกล้ทางปลายของโค้งที่เบ้นั้นมากที่สุด ค่ามัธยฐานจะอยู่ถัดจากค่ามัชฌิมเลขคณิต ส่วนค่าฐานนิยมจะอยู่ไกลจากปลายโค้งที่เบ้มากที่สุด (รูป 2)



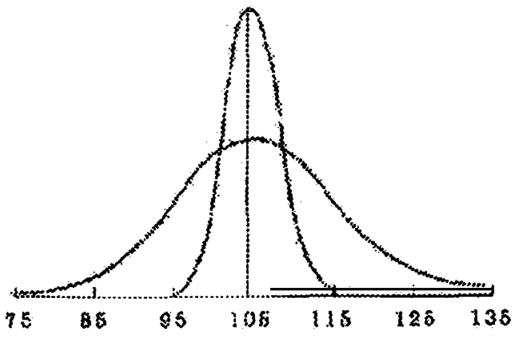
รูป 1 ความสัมพันธ์ของมัชฌิมเลขคณิต มัธยฐาน และฐานนิยม เมื่อข้อมูลมีการกระจายแบบปกติ



รูป 2 ความสัมพันธ์ของมัชฌิมเลขคณิต มัธยฐาน และฐานนิยม เมื่อข้อมูลมีการกระจายแบบเบ้ไปทางบวกและเบ้ไปทางลบ

9.4 การวัดการกระจายของข้อมูล

แม้ว่ามาตราการวัดแนวโน้มเข้าสู่ส่วนกลาง จะแสดงค่าเฉลี่ยที่ตัวแทนของกลุ่มทำได้และแสดงให้ทราบว่าข้อมูลเกาะกลุ่มกัน ณ ที่ใดแล้วก็ตาม แต่การทราบค่าตัวกลางของการวัดก็ยังไม่ให้ภาพทั้งหมดของกลุ่มตัวอย่างที่วัดได้ ตัวอย่างเช่น เด็กอายุ 6 ขวบ สองกลุ่ม อาจมีค่าเฉลี่ยของ IQ เท่ากับ 105 เท่ากันทั้งสองกลุ่ม ซึ่งจากข้อมูลนี้เราอาจลงความเห็นว่าเด็กทั้งสองกลุ่มมีความฉลาดเท่าเทียมกัน ถ้าเราเชื่อว่า IQ แสดงถึงเชาวน์ปัญญา แต่ถ้าเราทราบเพิ่มเติมว่ากลุ่มที่หนึ่งมี IQ ระหว่าง 95 ถึง 115 และกลุ่มที่สองมี IQ ระหว่าง 75 ถึง 135 เราก็จะทราบได้ทันทีว่า IQ ของเด็กทั้งสองกลุ่มมีการกระจายแตกต่างกัน กลุ่มแรกจะเป็นกลุ่มที่มีใกล้กันมากกว่ากลุ่มที่สอง การแจกแจงของ IQ ของเด็ก 2 กลุ่ม แสดงให้เห็นได้ในรูป 3 ดังนั้นการเปรียบเทียบความสามารถจึงไม่ดูเฉพาะค่าเฉลี่ย แต่จะดูการกระจายของคะแนนด้วย



รูป 3 การแจกแจงสองกลุ่มที่มีมัชฌิมเท่ากัน (IQ = 105) แต่พิสัย (การกระจาย) แตกต่างกัน

ดังนั้น การวัดการกระจายของข้อมูลจึงเป็นการระบุให้เห็นถึงความแตกต่างกันระหว่างคะแนนที่อยู่ในข้อมูลชุดเดียวกัน กล่าวคือ ถ้าคะแนนทุกรายการในข้อมูลชุดเดียวกันมีค่าเท่ากันทั้งหมด ข้อมูลนั้นจะไม่มีการกระจาย หรือมีการกระจายเป็นศูนย์ แต่ถ้าคะแนนแต่ละรายการในข้อมูลชุดใดมีความแตกต่างกัน บางรายการมีค่ามากบางรายการมีค่าน้อย แสดงว่าข้อมูลชุดนั้นมีการกระจาย ซึ่งถ้าความแตกต่างของข้อมูลมีไม่มากนัก ก็แสดงว่าข้อมูลชุดนั้นมีการกระจายน้อย แต่ถ้าข้อมูลชุดใดมีความแตกต่างกันมาก แสดงว่าข้อมูลชุดนั้นมีการกระจายมาก ดังตัวอย่างข้อมูล 3 ชุดข้างล่างนี้

ชุดที่ 1 : 10 10 10 10 10

ชุดที่ 2 : 4 4 5 7 8

ชุดที่ 3 : 1 5 7 10 15

จะเห็นได้ว่าข้อมูลชุดที่ 1 ไม่มีการกระจาย ส่วนข้อมูลชุดที่ 2 มีการกระจายบ้าง แต่ยังมีการกระจายน้อยกว่าข้อมูลชุดที่ 3

การวัดการกระจายของข้อมูลมีหลายวิธี แต่ที่นิยมใช้มี 4 วิธีคือ

1. พิสัย (Range)
2. ส่วนเบี่ยงเบนควอไทล์ (Quartile deviation)
3. ความแปรปรวน (Variance)
4. ส่วนเบี่ยงเบนมาตรฐาน (Standard deviation)

1. พิสัย (Range)

พิสัย คือ ระยะห่างระหว่างข้อมูลที่มีค่าสูงสุดกับข้อมูลที่มีค่าต่ำสุด ดังนั้นพิสัยของคะแนนชุดใด คือ ผลต่างระหว่างคะแนนที่มีค่าสูงสุดกับคะแนนที่มีค่าต่ำสุดในกลุ่ม ดังตัวอย่าง

ตัวอย่าง จงหาพิสัยของข้อมูล 2 ชุดต่อไปนี้

ชุดที่ 1 : 4 5 7 8 10 12

ชุดที่ 2 : 5 5 7 8 10 15

พิสัยของข้อมูลชุดที่ 1 $= 12-4 = 8$

พิสัยของข้อมูลชุดที่ 2 $= 15-5 = 10$

เนื่องจากในการหาค่าพิสัยของข้อมูลชุดใด เราใช้ตัวเลขเพียง 2 จำนวนริมสุดของข้อมูลชุดนั้นมาคำนวณเท่านั้น ในการหาค่าของพิสัยจึงนับว่าเป็นมาตราวัดที่หยาบ ถ้าค่าสูงสุดหรือต่ำสุดของข้อมูลชุดใดเปลี่ยนไปจะทำให้ค่าของพิสัยเปลี่ยนไปได้มาก อย่างไรก็ตามพิสัยเป็นมาตรา

วัดการกระจายของข้อมูลที่สามารถหาได้อย่างง่าย ๆ และรวดเร็ว แต่จัดว่าเป็นมาตรวัดที่หยาบ นักสถิติจึงหาวิธีที่จะหาค่าการกระจายของข้อมูล โดยนำข้อมูลรายการอื่น ๆ เข้ามาเกี่ยวข้องด้วย

2. ส่วนเบี่ยงเบนควอไทล์ (Quartile deviation)

1) "ควอไทล์" (Q) คือค่าของข้อมูลหรือจุดที่แบ่งจำนวนของข้อมูลออกเป็น 4 ส่วนเท่า ๆ กัน

Q_1 = ค่าของข้อมูลหรือจุดที่แสดงให้ทราบว่ามี 1/4 ของจำนวนข้อมูล (หรือ 25%) ที่มีค่าน้อยกว่าหรือเท่ากับค่านี้

Q_2 = ค่ามัธยฐาน

Q_3 = ค่าของข้อมูลหรือจุดที่แสดงให้ทราบว่ามี 3/4 ของจำนวนข้อมูล (หรือ 75%) ที่มีค่าน้อยกว่าหรือเท่ากับค่านี้

ตัวอย่าง นักเรียน 8 คน สอบวิชาสังคมศึกษาได้คะแนนดังนี้

12 16 19 23 24 28 30 33

ให้หา Q_1 และ Q_3

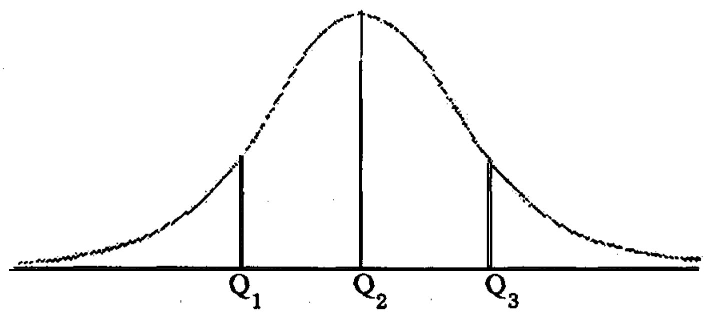
เนื่องจาก Q_1 = ค่าของข้อมูลที่แสดงให้ทราบว่ามีจำนวนข้อมูลอยู่ N/4 จำนวนที่มีค่าต่ำกว่าหรือเท่ากับค่าของ Q_1 และเนื่องจาก $N/4 = 8/4 = 2$ และ $3N/4 = 24/4 = 6$

ดังนั้น Q_1 จะอยู่ในตำแหน่งซึ่งมีข้อมูล 2 จำนวนอยู่ต่ำกว่า ซึ่งคือข้อมูลตัวที่ 3 (เมื่อข้อมูลอยู่ในลักษณะเรียงลำดับ)

$\therefore$ Q_1 มีค่าเท่ากับ 19

Q_3 มีค่าเท่ากับ 30 (มีจำนวนข้อมูลอยู่ต่ำกว่า 6 จำนวน)

ถ้าข้อมูลมีการกระจายแบบปกติ โค้งจะมีลักษณะสมมาตร ตำแหน่งของ Q_1 Q_2 และ Q_3 จะเป็นตำแหน่งที่แบ่งพื้นที่ใต้โค้งออกเป็น 4 ส่วนเท่า ๆ กัน ดังรูป 4



รูป 4 แสดงตำแหน่ง Q_1 Q_2 และ Q_3 เมื่อข้อมูลมีการกระจายแบบปกติ

2) ***"ส่วนเบี่ยงเบนควอไทล์"*** (Q.D.) เป็นการวัดการกระจายซึ่งคำนวณได้จาก ครึ่งหนึ่งของระยะห่างระหว่างควอไทล์ที่ 3 (Q_3) และควอไทล์ที่ 1 (Q_1) โดยมีสูตรดังนี้

$$Q.D. = \frac{Q_1 - Q_3}{2} \quad \text{.................................(4)}$$

จากตัวอย่างข้อมูลนักเรียน 8 คน

12 16 19 23 24 28 30 33

Q_1 = 19

Q_2 = 30

$$Q.D. = \frac{30 - 19}{2} = \frac{11}{2} = 5.5$$

สรุป 1. ค่าของส่วนเบี่ยงเบนควอไทล์มีค่ามาก ก็แสดงว่าข้อมูลมีการกระจายมาก

2. อย่างไรก็ดีการวัดการกระจายของข้อมูลโดยใช้ Quartile ไม่เป็นที่นิยมมากนัก เพราะการคำนวณหาค่า Q.D. ไม่ได้หาจากข้อมูลทั้งหมด แต่หาจากค่า Q_1 กับ Q_3 เท่านั้น

3. ความแปรปรวน (Variance)

ความแปรปรวน หมายถึง ค่าเฉลี่ยของค่ายกกำลังสองของความเบี่ยงเบนอันเกิดจากคะแนนแต่ละรายการเมื่อเทียบกับค่าเฉลี่ยของข้อมูลชุดนั้น โดยปกติใช้สัญลักษณ์ σ^2 แทนความแปรปรวนของประชากร และใช้ S^2 แทนความแปรปรวนของกลุ่มตัวอย่าง ข้อมูลชุดใดมีค่าความแปรปรวนมากกว่า จะมีการกระจายมากกว่าข้อมูลชุดที่มีค่าความแปรปรวนน้อยกว่า

1) ความแปรปรวนของประชากรหาได้จากสูตร (5)

$$\sigma^2 = \frac{\Sigma(X - \mu)^2}{N} \quad \text{.................................(5)}$$

2) ส่วนความแปรปรวนของกลุ่มตัวอย่างหาได้จากสูตร (6)

$$S^2 = \frac{\Sigma(X - \overline{X})^2}{N} \quad \text{.................................(6)}$$

เมื่อ σ^2 = ความแปรปรวนของประชากร

S^2 = ความแปรปรวนของกลุ่มตัวอย่าง

X = คะแนนของนักเรียนแต่ละคน

μ = คะแนนเฉลี่ยของประชากร

$\overline{X}$ = คะแนนเฉลี่ยของกลุ่มตัวอย่าง

N = จำนวนข้อมูลหรือจำนวนคะแนนทั้งหมด

สูตร (6) จะให้ค่าความแปรปรวนที่ปราศจากอคติก็ต่อเมื่อจำนวนกลุ่มตัวอย่างมีขนาดใหญ่ (N มากกว่า 30) ถ้ากลุ่มตัวอย่างมีขนาดน้อยกว่าหรือเท่ากับ 30 การประมาณค่าความแปรปรวนโดยใช้สูตร (6) จะทำให้เกิดอคติในการประมาณค่า จึงต้องมีการปรับแก้สูตร (6) โดยเปลี่ยนตัวหารจาก N เป็น N−1 ดังนี้

$$S^2 = \frac{\Sigma(X - \overline{X})^2}{N-1} \quad \text{..............................(7)}$$

ฉะนั้นการคำนวณหาค่า โดยใช้สูตร (7) จะเป็นค่าประมาณที่ไม่ลำเอียงของ σ^2

เนื่องจากสูตร (7) มีค่า $\overline{X}$ เข้ามาเกี่ยวข้องอาจทำให้การคำนวณเกิดความคลาดเคลื่อนได้ เนื่องจากการหาค่าเฉลี่ยอาจได้ค่าของตัวเลขที่เป็นทศนิยม จึงมีการพัฒนาสูตร (7) มาเป็นสูตร (8) เพื่อให้สะดวกในการคำนวณและไม่ต้องปัดค่าทศนิยมทิ้ง ดังนี้

$$S^2 = \frac{N\Sigma X^2 - (\Sigma X^2)}{N(N-1)} \quad \text{................................(8)}$$

แต่ถ้าข้อมูลที่นำมาหาค่าความแปรปรวน ถูกจัดหมวดหมู่ในลักษณะของความถี่ เราสามารถหาค่าความแปรปรวนของกลุ่มตัวอย่างจากสูตร (9) ดังนี้

$$S^2 = \frac{N\Sigma fX^2 - (\Sigma fX)^2}{N(N-1)} \quad \text{................................(9)}$$

เมื่อ Σ = ผลรวมของ

f = ความถี่

X = คะแนนแต่ละตัว

N = จำนวนข้อมูล

ตัวอย่าง จงหาค่าความแปรปรวนของข้อมูลข้างล่างนี้ โดยใช้สูตร $S^2 = \dfrac{N\Sigma X^2 - (\Sigma X)^2}{N(N-1)}$

เปรียบเทียบกับสูตร $S^2 = \dfrac{\Sigma(X-\overline{X})^2}{N-1}$

ข้อมูลมี 5 ตัว ดังนี้ 19 18 16 15 12

วิธีทำ (1)

X	X^2
19	361
18	324
16	256
15	225
12	144
$\Sigma X = 80$	$\Sigma X^2 = 1310$

$$S^2 = \frac{N\Sigma X^2 - (\Sigma X)^2}{N(N-1)}$$

$$= \frac{5(1310) - (80)^2}{(5)(4)}$$

$$= 7.5$$

วิธีทำ (2) หาจากสูตร $S^2 = \dfrac{\Sigma(X-\overline{X})^2}{N-1}$

เมื่อ $\overline{X} = \dfrac{\Sigma X}{N} = \dfrac{80}{5} = 16$

X	$X-\bar{X}$	$(X-\bar{X})^2$
19	3	9
18	2	4
16	0	0
15	−1	1
12	−4	16
		$\Sigma(X-\bar{X})^2 = 30$

$$S^2 = \frac{30}{4} = 7.5$$

4. ส่วนเบี่ยงเบนมาตรฐาน (Standard deviation)

ส่วนเบี่ยงเบนมาตรฐาน เป็นค่าเฉลี่ยอย่างหนึ่งของการเบี่ยงเบนจากมัชฌิมเลขคณิต ค่าเบี่ยงเบนมาตรฐานเป็นค่าที่ใช้กันทั่วไปในการแสดงถึงขนาดของการกระจาย ส่วนเบี่ยงเบนมาตรฐานหาได้จากการถอดรากที่สองของค่าความแปรปรวน สัญลักษณ์ที่ใช้แทนค่าส่วนเบี่ยงเบนมาตรฐานของประชากรคือ σ ส่วนสัญลักษณ์ที่ใช้แทนค่าส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่างคือ S หรือ S.D. มีสูตรสำหรับการคำนวณหาค่าดังนี้

ส่วนเบี่ยงเบนมาตรฐานของประชากรหาได้จากสูตร (10)

$$\sigma = \sqrt{\frac{\Sigma(x-\mu)^2}{N}} \quad \text{.................................(10)}$$

เมื่อ σ = ส่วนเบี่ยงเบนมาตรฐานของประชากร

ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่างที่มีขนาดใหญ่ ($N > 30$) หาได้จากสูตร (11)

$$S = \sqrt{\frac{\Sigma(x-\bar{x})^2}{N}} \quad \text{.................................(11)}$$

เมื่อ S = ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง

ในกรณีที่กลุ่มตัวอย่างมีขนาดเล็ก ($N \leq 30$) ให้หาส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่างโดยใช้สูตร (12) หรือ (13) ดังนี้

$$S = \sqrt{\frac{\Sigma (X - \overline{X})^2}{N - 1}} \quad \text{.................................(12)}$$

หรือ

$$S = \sqrt{\frac{N\Sigma X^2 - (\Sigma X)^2}{N(N-1)}} \quad \text{.................................(13)}$$

จากตัวอย่าง 11 ซึ่งหาค่า $S^2 = .75$

$$\therefore \ S = \sqrt{.75}$$

9.5 การวัดความสัมพันธ์ (Measures of Relationship)

ถ้าครูมีคะแนน 2 ชุด ซึ่งได้จากการทดสอบกับนักเรียนกลุ่มเดียวกัน ครูก็อยากจะรู้ว่าคะแนน 2 ชุดนั้นมีความสัมพันธ์เกี่ยวข้องกันปานใด ตัวอย่างเช่น ครูอาจจะสนใจอยากทราบความสัมพันธ์ระหว่างคะแนนวิชาคณิตศาสตร์กับวิชาภาษาไทย นั่นก็คือครูอยากจะรู้ว่าถ้านักเรียนทำคะแนนวิชาคณิตศาสตร์ได้คะแนนสูงแล้ว จะทำคะแนนในวิชาภาษาไทยได้สูงตามไปด้วยหรือไม่ สูตรการหาค่าความสัมพันธ์ของคะแนน 2 ชุดนั้น ส่วนใหญ่นิยมใช้สูตร Pearson product moment correlation coefficient (r) ซึ่งมีสูตรการคำนวณดังนี้

$$r = \frac{\Sigma[(X - \overline{X}) - (Y - \overline{Y})]}{NS_X S_Y} \quad \text{.............................(14)}$$

d a $\quad x$ = คะแนนของนักเรียนคนหนึ่งที่ได้จากการทดสอบวิชาหนึ่ง (หรือจากตัวแปรตัวหนึ่ง)

Y = คะแนนของนักเรียนคนเดียวกันที่ได้จากการทดสอบอีกวิชาหนึ่ง (หรือจากตัวแปรตัวหนึ่ง)

$\overline{X}$ = ค่าเฉลี่ยของคะแนน X

$\overline{Y}$ = ค่าเฉลี่ยของคะแนน Y

S_x = ค่าเบี่ยงเบนมาตรฐานของ X

S_y = ค่าเบี่ยงเบนมาตรฐานของ Y

N. = จำนวนของคะแนน (หรือจำนวนผู้เข้าสอบ)

ตัวอย่าง การคำนวณหาค่า r โดยใช้สูตร Pearson product moment correlation coefficie**แสดงในตาราง**

ตาราง แสดงการหาค่า r โดยใช้สูตร (14)

X	Y	$X-\bar{X}$	$(X-\bar{X})^2$	y - y	$(Y-\bar{Y})^2$	$(X-\bar{X})(Y-\bar{Y})$
50	45	20	400	16	256	320
49	50	19	361	21	441	399
30	25	0	0	-4	16	0
11	10	-19	361	-19	361	361
10	15	-20	400	-14	196	280
150	145		1522		1270	1360

$$\Sigma X = 150, \quad \Sigma Y = 145$$

$$\bar{X} = 30, \quad \bar{Y} = 29$$

$$\Sigma(X-\bar{X})^2 = 1522, \Sigma(Y-\bar{Y})^2 = 1270$$

$$S_x = \sqrt{\frac{\Sigma(X-\bar{X})^2}{N}} = \sqrt{\frac{1522}{5}} = \sqrt{304.4}$$

$$S_y = \sqrt{\frac{\Sigma(Y-\bar{Y})^2}{N}} = \sqrt{\frac{1270}{5}} = \sqrt{254}$$

$$r = \frac{\Sigma[(X-\bar{X})-(Y-\bar{Y})]}{NS_X S_Y}$$

$$= \frac{1360}{5 \times \sqrt{304.4} \times \sqrt{254}}$$

$$= .98$$

แสดงว่าคะแนน 2 ชุดนี้มีความสัมพันธ์กันสูงมาก เพราะ r มีค่าใกล้ 1.00

แต่ถ้าเราสามารถกำหนดอันดับที่ให้กับคะแนนที่จะมาหาความสัมพันธ์ เราก็สามารถหาค่าความสัมพันธ์ของคะแนน 2 ชุดนี้ได้อย่างง่าย ๆ โดยใช้สูตร Spearman rank–order correlation coefficient แต่ค่าที่ได้อาจจะคลาดเคลื่อนไปจากความเป็นจริงเล็กน้อย มีสูตรคำนวณดังนี้

$$\rho = 1 - \frac{6\sum D^2}{N(N^2-1)} \qquad \text{.................................(15)}$$

เมื่อ D คือความแตกต่างของอันดับของคะแนนแต่ละคู่

ตาราง แสดงการหาค่า ρ โดยการใช้สูตร (15)

X	Y	อันดับของ X	อันดับของ Y	D	D^2
50	45	1	2	-1	1
49	50	2	1	1	1
30	25	3	3	0	0
11	10	4	5	-1	1
10	15	5	4	1	1
					$\sum D^2 = 4$

$$\rho = 1 - \frac{6\sum D^2}{N(N^2-1)}$$

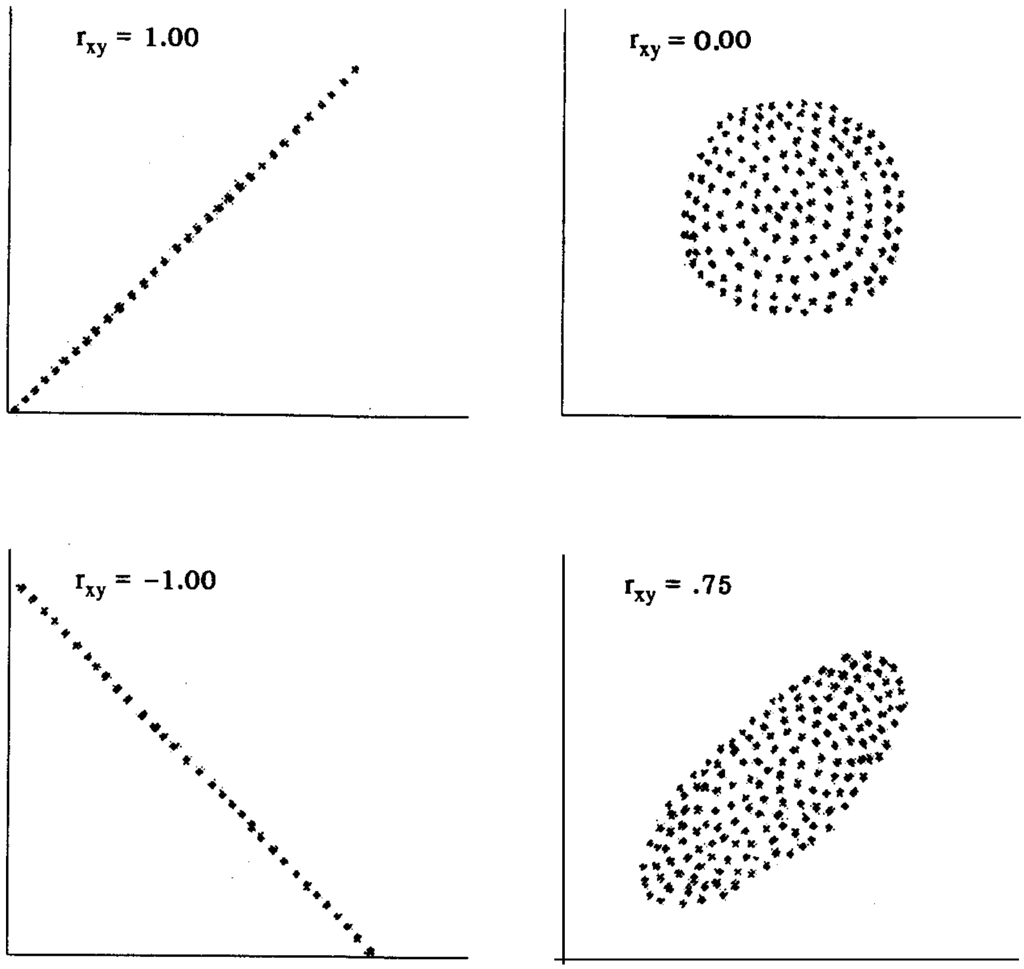
$$= 1 - \frac{6(4)}{5(5^2\quad 1)}$$

$$= I \quad - \frac{24}{120}$$

$$= .80$$

การแปลความหมายค่า r (หรือ ρ)

ค่าสหสัมพันธ์หรือ r มีค่าระหว่าง +1.00 ถึง −1.00 ถ้า r มีค่าเป็นบวก แสดงว่าตัวแปร 2 ตัวมีความสัมพันธ์ไปทางเดียวกัน นั่นคือถ้าตัวแปรตัวที่ 1 มีค่าสูง ตัวแปรตัวที่ 2 ก็จะมีค่าสูงตาม ถ้าต่ำก็จะต่ำตาม แต่ถ้า r มีค่าเป็นลบ แสดงว่าตัวแปร 2 ตัวนั้นมีความสัมพันธ์กันในทางตรงกันข้าม นั่นคือถ้าตัวแปรตัวหนึ่งสูงอีกตัวหนึ่งก็จะต่ำ ถ้า r = .00 แสดงว่าตัวแปร 2 ชุดนั้นไม่มีความสัมพันธ์ในเชิงเส้นตรง ถ้าเรานำคะแนน 2 ชุดใด ๆ ที่ได้จากการทดสอบคนกลุ่มเดียวกันมาลงจุดในกราฟจะได้ภาพตามลักษณะข้างล่างนี้



คำถามท้ายบทที่ 9

1. จากตัวอย่างต่อไปนี้ ให้ระบุระดับของข้อมูลหรือมาตราการวัดสูงสุดที่เกี่ยวข้อง
   1) จำนวนนักเรียนชายและหญิงในชั้นอนุบาล 1
   2) น้ำหนักเป็นกิโลกรัมที่เด็กคนหนึ่งยกได้
   3) ส่วนสูงของนักเรียน 10 คน
   4) ชื่อที่กำหนดให้นักเรียนกลุ่มต่าง ๆ ทำกิจกรรม
2. ถ้า $X_1=2, X_2=0, X_3=1, X_4=6, X_5=20$
   จงหาค่าของ
   1) $\sum_{i=1}^{6} x_i$
   2) $\sum_{i=1}^{4} x_i^2$
   3) $\sum_{i=1}^{4} x_i^2(x_i-3)$
   4) $\sum_{i=2}^{3} x_i^2(x_i-1)^2$
   5) $\sum_{i=2}^{5} \frac{(x_i+2)}{3}$
3. จากข้อมูลข้างล่างนี้

10	2	15	3	2	10	3	1	12	11
8	9	9	14	15	20	18	6	4	7

   จงคำนวณหาค่าต่อไปนี้
   1) ค่าเฉลี่ย
   2) ค่ามัธยฐาน
   3) ค่าฐานนิยม
   4) ค่าส่วนเบี่ยงเบนมาตรฐาน
   5) ค่าความแปรปรวน