

การวัดค่ากลางของข้อมูล

การวัดค่ากลางของข้อมูล ก็คือการคำนวณค่าใดค่าหนึ่งให้เป็นตัวแทนของข้อมูลทั้งหมด ในการวิเคราะห์ข้อมูลเบื้องต้นนั้น นอกจากจะทำได้โดยการแจกแจงความถี่ด้วยตารางและด้วยกราฟ แล้ว ยังสามารถวิเคราะห์ข้อมูลด้วยการหาค่ากลางมาเป็นตัวแทนของข้อมูลทั้งหมด การวิเคราะห์ข้อมูลด้วยค่ากลางนี้ จะทำให้ทราบลักษณะโดยรวมของข้อมูลทั้งหมดได้ดีขึ้น และสะดวก ในการจดจำหรือสรุปเรื่องราวที่เกี่ยวข้องกับข้อมูลนั้น ๆ ได้มากขึ้น เช่น ผู้อำนวยการโรงเรียน ต้องการทราบว่าผลการเรียนคณิตศาสตร์ของนักเรียนตั้งแต่มัธยมศึกษาปีที่ 1 ถึงชั้นมัธยมศึกษาปีที่ 3 ในปีที่ผ่านมาเป็นอย่างไร แทนที่ผู้อำนวยการจะต้องทราบคะแนนวิชาคณิตศาสตร์ของนักเรียนแต่ละคนแต่ละระดับชั้น ผู้อำนวยการอาจจะทราบเพียงคะแนนเฉลี่ยของวิชาคณิตศาสตร์ของนักเรียนในแต่ละระดับชั้น ก็เพียงพอที่ตัดสินใจได้โดยกว้าง ๆ ว่า ผลการเรียนวิชาคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 ถึงชั้นมัธยมศึกษาปีที่ 3 ในปีที่ผ่านมาเป็นอย่างไร ควรจะปรับปรุงและพัฒนาการเรียนการสอนวิชาคณิตศาสตร์มากน้อยแค่ไหน เป็นต้น การกำหนดค่ากลางเพื่อเป็นตัวแทนของข้อมูลทั้งหมด มีวิธีกำหนดหลายวิธี เช่น

- 1) การกำหนดค่ากลางในรูปของค่าเฉลี่ยเลขคณิต (arithmetic mean)
- 2) การกำหนดค่ากลางในรูปของค่าเฉลี่ยเรขาคณิต (geometric mean)
- 3) การกำหนดค่ากลางในรูปของค่าเฉลี่ยฮาร์มอนิก (harmonic mean)
- 4) การกำหนดค่ากลางในรูปของค่ามัธยฐาน (median)
- 5) การกำหนดค่ากลางในรูปของค่าฐานนิยม (mode)
- 6) การกำหนดค่ากลางในรูปของค่ากึ่งกลางพิสัย (midrange)

ค่ากลางของข้อมูลที่นิยมใช้กันทั่วไปมี 3 ชนิดคือ ค่าเฉลี่ยเลขคณิต ค่ามัธยฐานและค่าฐานนิยม

การคำนวณค่ากลางทั้งสามชนิดนี้ แบ่งเป็น 2 แบบใหญ่ คือ

แบบที่ 1 การหาค่ากลางของข้อมูลที่ไม่ได้แจกแจงความถี่ (ungroup data)

แบบที่ 2 การหาค่ากลางของข้อมูลที่ได้แจกแจงความถี่แล้ว (group data) ซึ่งจะได้

กล่าวถึงการหาค่ากลางชนิดต่าง ๆ ในหัวข้อต่อไป

1. การหาค่าเฉลี่ยเลขคณิต (mean)

ค่าเฉลี่ยเลขคณิต เหมาะที่จะนำมาใช้เป็นค่ากลางของข้อมูลในกรณีที่ข้อมูลชุดนั้นไม่มีค่าใดค่าหนึ่งน้อยกว่าหรือมากกว่าค่าอื่น ๆ ในข้อมูลชุดนั้นอย่างผิดปกติ เช่น จากการชั่งน้ำหนัก (กิโลกรัม) ของนักเรียน 5 คน ได้ผลดังนี้

48 , 45 , 50 , 52 , 105

ในตัวอย่างนี้ น้ำหนัก 105 กิโลกรัม ของข้อมูลชุดนี้ถือว่ามากผิดปกติ

การหาค่าเฉลี่ยน้ำหนักของนักเรียน 5 คนนี้ จะได้คำตอบที่ไม่สอดคล้องกับความเป็นจริง และถือได้ว่าคำตอบที่ได้นี้ไม่ใช่ตัวแทนที่ดีของข้อมูลชุดนี้

การหาค่าเฉลี่ยเลขคณิตของข้อมูลที่ไม่ได้แจกแจงความถี่

การหาค่าเฉลี่ยเลขคณิตของข้อมูลที่ไม่ได้แจกแจงความถี่ จะสามารถหาได้ง่าย ๆ โดยการนำข้อมูลทุกค่ารวมกัน แล้วหารด้วยจำนวนข้อมูลทั้งหมด เช่น พี่น้องสามคนมีอายุ 13 , 15 , 17 ปี

ถามลำดับ จะได้ว่า อายุเฉลี่ยของพี่น้องสามคนนี้ คือ $\frac{13+15+17}{3} = \frac{45}{3} = 15$ ปี

ถ้าให้ $X_1, X_2, X_3, \dots, X_N$ เป็นข้อมูล N ค่าและให้ \bar{X} แทนค่าเฉลี่ยเลขคณิตของข้อมูลชุดนี้ จะได้ว่า

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_N}{N}$$

$$\bar{X} = \frac{\sum_{i=1}^N X_i}{N}$$

ตัวอย่างที่ 1 ในการซ้อมวิ่งผลัด 4 x 100 เมตร ของทีมไทย นักวิ่ง 4 คน ทำเวลาได้ 9.9 , 10.1 , 10.3 และ 10.5 วินาที ตามลำดับ จงหาว่าในระยะทาง 100 เมตร ทีมไทยทำเวลาเฉลี่ยได้เท่าไร

วิธีทำ ทีมไทยทำเวลาเฉลี่ยได้ คือ $\frac{9.9+10.1+10.3+10.5}{4} = \frac{40.8}{4} = 10.2$

ตอบ 10.2 วินาที

ตัวอย่างที่ 2 ธนาคารแห่งหนึ่งมีเจ้าหน้าที่ 12 คน มีอายุเรียงตามลำดับดังนี้ 22, 24, 25, 28, 30, 32, 35, 37, 38, 40, 44, 49 ปี จงหาอายุเฉลี่ยของเจ้าหน้าที่ของธนาคารแห่งนี้

วิธีทำ อายุเฉลี่ยของเจ้าหน้าที่ธนาคารแห่งนี้ $\frac{22+24+25+28+30+32+35+37+38+40+44+49}{12} = \frac{404}{12} = 33.67$

ตอบ 33.67 ปี

ข้อสังเกต

1) จาก $\bar{X} = \frac{\sum_{i=1}^N X_i}{N}$ จะทำให้ได้ว่า $N\bar{X} = \sum_{i=1}^N X_i$

นั่นคือ $N\bar{X}$ คือ ผลรวมของข้อมูลทั้งหมด

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

เหตุผล เพราะว่า $\sum_{i=1}^N (X_i - \bar{X}) = \sum_{i=1}^N X_i - \sum_{i=1}^N \bar{X}$
 $= \sum_{i=1}^N X_i - N\bar{X} \quad ; \quad \bar{X} \text{ เป็นค่าคงตัว}$
 $= \sum_{i=1}^N X_i - N \frac{\sum_{i=1}^N X_i}{N}$
 $= \sum_{i=1}^N X_i - \sum_{i=1}^N X_i$
 $= 0$

3) ถ้า C เป็นค่าคงตัวใด ๆ และ \bar{X} เป็นค่าเฉลี่ยเลขคณิตของข้อมูล
ถ้าให้ $X_1, X_2, X_3, \dots, X_N$ แล้ว จะได้ว่า $\sum_{i=1}^N (X_i - \bar{X})^2 \leq \sum_{i=1}^N (X_i - C)^2$ เสมอ

เช่น จากข้อมูล 4, 6, 2, 8 จะได้ $\bar{X} = \frac{4+6+2+8}{4} = 5$

$$\begin{aligned} \sum_{i=1}^N (X_i - \bar{X})^2 &= (4-5)^2 + (6-5)^2 + (2-5)^2 + (8-5)^2 \\ &= 20 \end{aligned}$$

(1) ถ้า $C = 3$ จะได้ $\sum_{i=1}^4 (X_i - 3)^2 = 36$

(2) ถ้า $C = 4$ จะได้ $\sum_{i=1}^4 (X_i - 4)^2 = 24$

(3) ถ้า $C = 7$ จะได้ $\sum_{i=1}^4 (X_i - 7)^2 = 36$