# บทที่ 4
# การวัดการกระจาย
# (MEASURES OF VARIABILITY)

## วัตถุประสงค์

เมื่อท่านศึกษาเนื้อหาบทที่ 4 โดยละเอียดแล้ว ควรมีความสามารถดังนี้

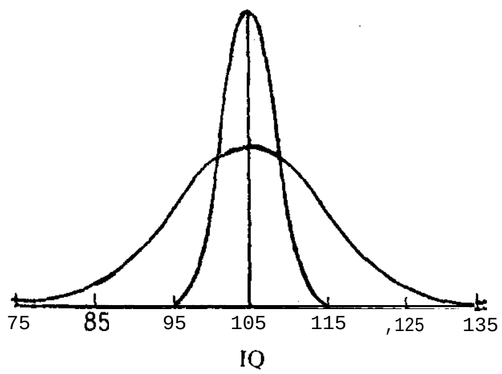
1. บอกความหมายของการวัดการกระจายได้
2. คำนวณค่าสถิติที่ใช้ในการวัดแนวโน้มเข้าสู่ส่วนกลางได้
3. เลือกสถิติที่ใช้ในการวัดแนวโน้มเข้าสู่ส่วนกลางได้เหมาะสมกับลักษณะของข้อมูล

## เนื้อหา

4.1 ความหมายของการวัดการกระจาย
4.2 พิสัย (range)
4.3 ความแปรปรวน (variance)
4.4 ส่วนเบี่ยงเบนมาตรฐาน (standard deviation)

# เนื้อหาที่ 4.1 ความหมายของการวัดการกระจาย

แม้ว่ามาตราวัดแนวโน้มเข้าสู่ส่วนกลาง จะแสดงค่าเฉลี่ยที่ตัวแทนของกลุ่มทำได้ และแสดงให้ทราบว่าข้อมูลเกาะกลุ่มกัน ณ ที่ใดแล้วก็ตาม แต่การทราบค่าตัวกลางของการวัดก็ยังไม่ให้ภาพทั้งหมดของกลุ่มตัวอย่างที่วัดได้ ตัวอย่างเช่นเด็กอายุ 6 ขวบสองกลุ่ม อาจมีค่าเฉลี่ยของ IQ เท่ากับ 105 เท่ากันทั้งสองกลุ่ม ซึ่งจากข้อมูลนี้เราอาจลงความเห็นว่าเด็กทั้งสองกลุ่มมีความฉลาดเท่าเทียมกัน ถ้าเราเชื่อว่า IQ แสดงถึงเชาวน์ปัญญา แต่ถ้าเราทราบเพิ่มเติมว่ากลุ่มที่หนึ่งมี IQ ระหว่าง 95 ถึง 115 และกลุ่มที่สองมี IQ ระหว่าง 75 ถึง 135 เราก็จะทราบได้ทันทีว่า IQ ของเด็กทั้งสองกลุ่มมีการกระจายแตกต่างกัน กลุ่มแรกจะเป็นกลุ่มที่มี IQ ใกล้กันมากกว่ากลุ่มที่สอง การแจกแจงของ IQ ของเด็ก 2 กลุ่ม แสดงให้เห็นได้ในรูป 4.1




รูป 4.1 การแจกแจงสองกลุ่มที่มีมัชฌิมเท่ากัน (IQ = 105) แต่พิสัย (การกระจาย) แตกต่างกัน

ดังนั้นการวัดการกระจายของข้อมูลจึงเป็นการระบุให้เห็นถึงความแตกต่างกันระหว่างคะแนนที่อยู่ในข้อมูลชุดเดียวกัน กล่าวคือ ถ้าคะแนนทุกรายการในข้อมูลชุดเดียวกันมีค่าเท่ากันทั้งหมด ข้อมูลชุดนั้นจะไม่มีการกระจาย หรือมีการกระจายเป็นศูนย์ แต่ถ้าคะแนนแต่ละรายการในข้อมูลชุดใดมีความแตกต่างกัน บางรายการมีค่ามากบางรายการมีค่าน้อย แสดงว่าข้อมูลชุดนั้นมีการกระจาย ซึ่งถ้าความแตกต่างของข้อมูลที่ไม่มากนัก 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 แต่ยังมีการกระจาย**น้อยกว่า**ข้อมูลชุดที่ 3

การวัดการกระจายของข้อมูล มีวิธีการวัดอยู่หลายแบบ แต่ในบทนี้จะกล่าวถึงเฉพาะวิธีวัดการกระจายที่นิยมใช้กันมากอยู่ 3 แบบ คือ พิสัย (range) ความแปรปรวน (variance) และส่วนเบี่ยงเบนมาตรฐาน (standard deviation) เท่านั้น

## เนื้อหาที่ 4.2 พิสัย

พิสัย คือระยะทางที่ห่างกันระหว่างข้อมูลที่มีค่าสูงสุดกับข้อมูลที่มีค่าต่ำสุด ดังนั้น พิสัยของคะแนนชุดใด คือ ผลต่างระหว่างคะแนนที่มีค่าสูงสุดกับคะแนนที่มีค่าต่ำสุดในกลุ่ม ดังตัวอย่าง 4.1

**ตัวอย่าง 4.1** จงหาพิสัยของข้อมูล 2 ชุด ต่อไปนี้
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 ในการหาค่าของพิสัยจึงนับว่าเป็นมาตรวัดที่หยาบ ถ้าค่าสูงสุดหรือต่ำสุดของข้อมูลชุดใดเปลี่ยนไปจะทำให้ค่าของพิสัยเปลี่ยนไปได้มาก (ดูตัวอย่าง 4.1) อย่างไรก็ตาม พิสัย เป็นมาตราวัดการกระจายของข้อมูลที่สามารถหาได้

อย่างง่าย ๆ และรวดเร็ว แต่จัดว่าเป็นมาตราวัดที่หยาบ นักสถิติจึงหาวิธีที่จะหาค่าการกระจายของข้อมูล โดยนำข้อมูลทุกรายการเข้ามาเกี่ยวข้องด้วย วิธีหนึ่งที่นิยมใช้คือการหาความแปรปรวนของข้อมูล

## เนื้อหาที่ 4.3 ความแปรปรวน

ความแปรปรวน หมายถึง ค่าเฉลี่ย (มัชฌิมเลขคณิต) ของค่ายกกำลังสองของความเบี่ยงเบนอันเกิดจากคะแนนแต่ละรายการเมื่อเทียบกับค่าเฉลี่ย (มัชฌิมเลขคณิต) ของข้อมูลชุดนั้น โดยปกติใช้สัญลักษณ์ $\sigma^2$ แทนความแปรปรวนของประชากร และใช้ $S^2$ แทนความแปรปรวนของกลุ่มตัวอย่าง ความแปรปรวนของประชากรหาได้จากสูตร 4.1

$$\sigma^2 = \frac{\Sigma(X-\mu)^2}{N} \qquad \text{.......(4.1)}$$

ส่วนความแปรปรวนของกลุ่มตัวอย่างหาได้จากสูตร 4.2

$$S^2 = \frac{\Sigma(X-\bar{X})^2}{N} \qquad \text{.......(4.2)}$$

เมื่อ $\sigma^2$ = ความแปรปรวนของประชากร
$S^2$ = ความแปรปรวนของกลุ่มตัวอย่าง
$X$ = คะแนนของนักเรียนแต่ละคน
$\mu$ = คะแนนเฉลี่ยของประชากร
$\bar{X}$ = คะแนนเฉลี่ยของกลุ่มตัวอย่าง
$N$ = จำนวนข้อมูลหรือจำนวนคะแนนทั้งหมด

สูตร 4.2 จะให้ค่าความแปรปรวนที่ปราศจากอคติก็ต่อเมื่อ จำนวนกลุ่มตัวอย่างมีขนาดใหญ่ (N มากกว่า 30) ถ้ากลุ่มตัวอย่างมีขนาดน้อยกว่าหรือเท่ากับ 30 การประมาณค่าความแปรปรวนโดยใช้สูตร 4.2 จะทำให้เกิดอคติในการประมาณค่า จึงต้องมีการปรับแก้สูตร 4.2 โดยเปลี่ยนตัวหารจาก N เป็น N−1 ดังนี้

$$S^2 = \frac{\Sigma(X-\bar{X})^2}{N-1} \qquad \text{.......(4.3)}$$

ฉะนั้นการคำนวณหาค่า $S^2$ โดยใช้สูตร 4.3 จะเป็นค่าประมาณที่ไม่ลำเอียงของ $\sigma^2$

การที่จำเป็นต้องปรับค่า N เป็น N − 1 ในการหาค่าความแปรปรวนของกลุ่มตัวอย่างที่มีขนาดน้อยกว่า 30 ดังสูตร 4.3 เนื่องจากจำนวนของค่าที่เบี่ยงเบนจากค่า

เฉลี่ยได้อย่างอิสระมีจำนวนเท่ากับ N−1 (ถ้า N คือจำนวนกลุ่มตัวอย่างเช่น สมมุติว่ามีข้อมูลอยู่ 5 จำนวน คือ 5, 7, 10, 12 และ 16 ค่าเฉลี่ยคือ 10.0 ค่าเบี่ยงเบนจากค่าเฉลี่ยคือ −5, −3, 0, +2, และ +6 ผลรวมของค่าเบี่ยงเบนจากค่าเฉลี่ยคือศูนย์ [(−5) + (−3) + (0) + (2) + (6) = 0] พิจารณาจากค่าเบี่ยงเบน 5 จำนวนนี้จะพบว่า ถ้ากำหนดค่าเบี่ยงเบนจากค่าเฉลี่ย 5 จำนวนใหม่ โดยบังคับให้ผลรวมของค่าเบี่ยงเบนเป็นศูนย์แล้ว เราจะกำหนดค่าเบี่ยงเบนได้อย่างอิสระเพียง 4 จำนวน ส่วนค่าเบี่ยงเบนตัวที่ 5 จะต้องคงที่จะแปรเปลี่ยนอย่างอิสระไม่ได้ เช่นถ้าเราสุ่มกลุ่มตัวอย่างใหม่ เราอาจทำให้ค่าเบี่ยงเบนสี่จำนวนเป็น −8, −4, +1 และ −2 และเพื่อให้ผลรวมเป็นศูนย์จำนวนที่ห้าจะต้องถูกกำหนดโดยอัตโนมัติว่าจะต้องเท่ากับ +13 ดังนั้นจึงมีเพียงสี่ (N−1) จำนวนเท่านั้นที่มีอิสระที่จะผันแปรภายใต้ข้อจำกัดที่กำหนดไว้

ในกรณีที่กลุ่มตัวอย่างมีขนาดใหญ่ (กลุ่มตัวอย่างมีจำนวนมากกว่า 30) จะใช้สูตร 4.2 หรือสูตร 4.3 ก็ได้ เพราะค่าที่คำนวณได้จะใกล้เคียงกัน

เนื่องจากสูตรการคำนวณหาความแปรปรวนของกลุ่มตัวอย่าง (สูตร 4.3) ค่อนข้างจะยุ่งยาก เพราะต้องนำคะแนนของนักเรียนแต่ละคนไปลบออกจากคะแนนเฉลี่ยแล้วนำมายกกำลังสอง จึงได้มีการดัดแปลงสูตร 4.3 ให้ง่ายต่อการคำนวณ ดังนี้

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

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

$$= \frac{\Sigma X^2 - 2\Sigma X\left(\frac{\Sigma X}{N}\right) + \Sigma\left(\frac{\Sigma X}{N}\right)^2}{N - 1}$$

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

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

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

ฉะนั้น $$S^2 = \frac{N\Sigma X^2 - (\Sigma X)^2}{N (N-1)} \qquad (4.4)$$

**ตัวอย่าง 4.2** จงหาความแปรปรวนของคะแนน IQ ของนักเรียน 2 กลุ่ม ซึ่งมี IQ เฉลี่ยเท่ากัน

คะแนน IQ ของนักเรียนกลุ่มที่ 1 มีดังนี้

109 108 107 105 105 103 102 101 101 101

99 99 97 97 96 95 95 94 93 93

คะแนน IQ ของนักเรียนกลุ่มที่ 2 มีดังนี้

185 147 121 108 106 104 103 103 102 101

99 96 91 83 82 80 74 74 71 70

**วิธีหาค่าความแปรปรวนของ IQ ของนักเรียนกลุ่มที่ 1**
เนื่องจากกลุ่มตัวอย่างมีขนาดเล็ก (N น้อยกว่า 30) จึงใช้สูตร 4.3 หรือ 4.4 ในการคำนวณหาค่าความแปรปรวนได้ดังนี้

| คะแนน IQ (X) | $X^2$ | $X-\bar{X}$ | $(X-\bar{X})^2$ |
|---|---|---|---|
| 109 | 11881 | 9 | 81 |
| 108 | 11664 | 8 | 64 |
| 107 | 11449 | 7 | 49 |
| 105 | 11025 | 5 | 25 |
| 105 | 11025 | 5 | 25 |
| 103 | 10609 | 3 | 9 |
| 102 | 10404 | 2 | 4 |
| 101 | 10201 | 1 | 1 |
| 101 | 10201 | 1 | 1 |
| 101 | 10201 | 1 | 1 |
| 99 | 9801 | -1 | 1 |
| 99 | 9801 | -1 | 1 |
| 97 | 9409 | -3 | 9 |
| 97 | 9409 | -3 | 9 |
| 96 | 9216’ | -4 | 16 |
| 95 | 9025 | -5 | 25 |
| 95 | 9025 | -5 | 25 |
| 94 | 8836 | -6 | 36 |
| 93 | 8649 | -7 | 49 |
| 93 | 8649 | -7 | 49 |
| ผลรวม ($\Sigma$) 2000 | 200480 | | 480 |

เมื่อ $N = 20, \Sigma X = 2000$ หาค่าเฉลี่ยของ IQ ได้จากสูตร

$$\bar{X} = \frac{\Sigma X}{N}$$

$$= \frac{2000}{20}$$

$$= 100$$

ดังนั้น IQ เฉลี่ยของนักเรียนกลุ่มที่ 1 เท่ากับ 100

1) คำนวณหาค่าความแปรปรวนของ IQ นักเรียนกลุ่มที่ 1 โดยใช้สูตร 4.3 ได้ดังนี้

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

$$= \frac{480}{20 - 1}$$

$$= \frac{480}{19}$$

$$= 25.26$$

2) คำนวณหาค่าความแปรปรวนของ IQ นักเรียนกลุ่มที่ 1 โดยใช้สูตร 4.4 ได้ดังนี้

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

$$= \frac{(20) (200480) - (2000)^2}{(20) (20 - 1)}$$

$$= \frac{4009600 - 4000000}{(20) (19)}$$

$$= \frac{9600}{380}$$

$$= 25.26$$

จากตัวอย่างจะเห็นว่าค่าความแปรปรวนของ IQ นักเรียนกลุ่มที่ 1 ซึ่งหาโดยใช้สูตร 4.3 เท่ากับค่าความแปรปรวนที่คำนวณโดยใช้สูตร 4.4

**วิธีหาค่าความแปรปรวนของ IQ ของนักเรียนกลุ่มที่ 2**

เนื่องจากกลุ่มตัวอย่างกลุ่มที่ 2 มีขนาดเล็ก (N น้อยกว่า 30) จึงใช้สูตร 4.3 หรือ 4.4 ในการคำนวณหาค่าความแปรปรวนได้ดังนี้

| คะแนน IQ (X) | $X^2$ | $X - \bar{X}$ | $(X - \bar{X})^2$ |
|---|---|---|---|
| 185 | 34225 | 85 | 7225 |
| 147 | 21609 | 47 | 2209 |
| 121 | 14641 | 21 | 441 |
| 108 | 11664 | 8 | 64 |
| 106 | 11236 | 6 | 36 |
| 104 | 10816 | 4 | 16 |
| 103 | 10609 | 3 | 9 |
| 103 | 10609 | 3 | 9 |
| 102 | 10404 | 2 | 4 |
| 101 | 10201 | 1 | 1 |
| 99 | 9801 | −1 | 1 |
| 96 | 9216 | −4 | 16 |
| 91 | 8281 | −9 | 81 |
| 83 | 6889 | −17 | 289 |
| 82 | 6724 | −18 | 324 |
| 80 | 6400 | −20 | 400 |
| 74 | 5476 | −26 | 676 |
| 74 | 5476 | −26 | 676 |
| 71 | 5041 | −29 | 841 |
| 70 | 4900 | −30 | 900 |
| ผลรวม (Σ) 2000 | 214218 | | 14218 |

หาค่าเฉลี่ยของ IQ ได้จากสูตร

$$\bar{X} = \frac{\Sigma X}{N}$$

$$= \frac{2000}{20}$$

$$= 100$$

1) คำนวณหาค่าความแปรปรวนของ IQ ของนักเรียนกลุ่มที่ 2 โดยใช้สูตร 4.3 ได้ดังนี้

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

$$= \frac{14218}{20-1}$$

$$= \frac{14218}{19}$$

$$= 748.32$$

2) คำนวณหาค่าความแปรปรวนของ IQ ของนักเรียนกลุ่มที่ 2 โดยใช้สูตร 4.4 ได้ดังนี้

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

$$= \frac{(20)\ (214218) - (2000)^2}{(20)\ (20 - 1)}$$

$$= \frac{4284360 - 4000000}{(20)\ (19)}$$

$$= \frac{284360}{380}$$

$$= 748.32$$

เนื่องจากค่าความแปรปรวนที่คำนวณได้จากสูตรที่ 4.3 และสูตร 4.4 มีค่าเท่ากัน ดังนั้นในกรณีกลุ่มตัวอย่างขนาดเล็ก ($N \leq 30$) ท่านจะเลือกใช้สูตร 4.3 หรือ 4.4 สูตรใดสูตรหนึ่งก็ได้ในการคำนวณหาค่าความแปรปรวน

# เนื้อหาที่ 4.4 ส่วนเบี่ยงเบนมาตรฐาน

ส่วนเบี่ยงเบนมาตรฐาน เป็นค่าเฉลี่ยอย่างหนึ่งของการเบี่ยงเบนจากมัชฌิมเลขคณิต ค่าเบี่ยงเบนมาตรฐานเป็นค่าที่ใช้กันทั่วไปในการแสดงถึงขนาดของการกระจายส่วนเบี่ยงเบนมาตรฐานหาได้จากการถอดรากที่สองของค่าความแปรปรวน สัญลักษณ์ที่ใช้แทนค่าส่วนเบี่ยงเบนมาตรฐานของประชากรคือ $\sigma$ ส่วนสัญลักษณ์ที่ใช้แทนค่าส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง คือ S หรือ SD มีสูตรสำหรับการคำนวรหาค่า ดังนี้

ส่วนเบี่ยงเบนมาตรฐานของประชากรหาได้จากสูตร 4.5

$$\sigma = \sqrt{\frac{\Sigma (X - \mu)^2}{N}} \qquad \text{.....(4.5)}$$

เมื่อ $\sigma$ = สำนวนเบี่ยงเบนมาตรฐานของประชากร

ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่างที่มีขนาดใหญ่หาได้จากสูตร 4.6

$$S = \sqrt{\frac{\Sigma (X - \bar{X})^2}{N}} \qquad \text{.....(4.6)}$$

เมื่อ S = ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง

ในกรณีที่กลุ่มตัวอย่างมีขนาดเล็กให้หาค่าส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่างโดยใช้สูตร 4.7 หรือ 4.7 ดังนี้

$$S = \sqrt{\frac{\Sigma (X - \bar{X})^2}{N - 1}} \qquad \text{.....(4.7)}$$

หรือ

$$S = \sqrt{\frac{N\Sigma X^2 - (\Sigma X)^2}{N (N-1)}} \qquad \text{.....(4.8)}$$

# สรุปเนื้อหาบทที่ 4

1. การวัดการกระจายของข้อมูลเป็นการระบุให้เห็นถึงความแตกต่างระหว่างคะแนนที่อยู่ในข้อมูลชุดเดียวกัน

2. วิธีวัดการกระจายของข้อมูลที่นิยมใช้กันมากมีอยู่ 3 แบบ คือ พิสัย ความแปรปรวน และส่วนเบี่ยงเบนมาตรฐาน

3. พิสัย คือ ระยะทางที่ห่างกันระหว่างข้อมูลที่มีค่าสูงสุดกับข้อมูลที่มีค่าต่ำสุด

4. ความแปรปรวนหมายถึงค่าเฉลี่ยของค่ายกกำลังสองของความเบี่ยงเบนอันเกิดจากคะแนนแต่ละรายการเมื่อเทียบกับค่าเฉลี่ยของข้อมูลชุดนั้น

5. ความแปรปรวนของประชากรหาได้จากสูตร

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

6. ความแปรปรวนของกลุ่มตัวอย่างขนาดใหญ่ (N มากกว่า 30) หาได้จากสูตร

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

7. ความแปรปรวนของกลุ่มตัวอย่างขนาดเล็ก (N ≤ 30) หาได้จากสูตร

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

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

8. ส่วนเบี่ยงเบนมาตรฐาน เป็นค่าเฉลี่ยอย่างหนึ่งของการเบี่ยงเบนจากมัชฌิมเลขคณิต หาได้จากการถอดรากที่สองของค่าความแปรปรวน

9. ส่วนเบี่ยงเบนมาตรฐานของประชากร หาได้จากสูตร

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

10. ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่างขนาดใหญ่หาได้จากสูตร

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

11. ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่างขนาดเล็ก หาได้จากสูตร

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

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

# คำถามท้ายบทที่ 4

1. การวัดการกระจายของข้อมูลหมายถึงอะไรและมีกี่วิธี

2. จากข้อมูลต่อไปนี้จงหาค่าพิสัย ความแปรปรวนและส่วนเบี่ยงเบนมาตรฐาน

   ก. 8, 15, 13, 6, 10, 16, 7, 12, 11, 14, 9

   ข. 12, 10, 18, 13, 4, 8, 17, 15, 6, 14

   ค. 9, 8, 9, 15, 3, 9, 11, 9, 13

   ง. 12, 28, 19, 15, 15, 35, 14, 15