

ใบงานที่ 3.4 เรื่อง ค่าวัดทางสถิติ : ค่าวัดการกระจาย

รหัสวิชา ค33202 วิชาเสริมทักษะคณิตศาสตร์ 6

ชั้นมัธยมศึกษาปีที่ 6

ค่าวัดทางสถิติ

2. ค่าวัดการกระจาย

การวัดการกระจายของข้อมูลแบ่งได้เป็น 2 วิธี คือ

1. การกระจายสัมบูรณ์ (absolute variation) คือ _____

ในที่นี้จะศึกษาค่าวัดการกระจายสัมบูรณ์ 4 ชนิด คือ

1) _____ 2) _____

3) _____ 4) _____

2. การกระจายสัมพัทธ์ (relative variation) คือ _____

ในที่นี้จะศึกษาค่าวัดการกระจายสัมพัทธ์เพียงชนิดเดียวคือสัมประสิทธิ์ของการแปรผัน

การวัดการกระจายสัมบูรณ์

1) พิสัย (range) คือ _____

ตัวอย่างที่ 14

ผลผลิตน้ำตาลใน พ.ศ. 2561/62 ของจีน สหรัฐอเมริกา ไทย อินเดีย ออสเตรเลีย และบราซิล แสดงได้ดังนี้

ประเทศ	จีน	สหรัฐอเมริกา	ไทย	อินเดีย	ออสเตรเลีย	บราซิล
ผลผลิต (ล้านตัน)	10.60	8.12	14.19	33.07	4.90	29.50

จงหาพิสัยของข้อมูลชุดนี้

--

จากตัวอย่างข้างต้น สามารถนำพิสัยมาใช้ในการอธิบายอย่างคร่าว ๆ ว่าผลผลิตน้ำตาลใน พ.ศ. 2561/62 ของแต่ละประเทศจะต่างกันไม่เกิน 28.17 ล้านตัน

2) พิสัยระหว่างควอร์ไทล์ (interquartile range) คือ _____

เขียนแทนพิสัยระหว่างควอร์ไทล์ด้วย IQR

$$IQR = \underline{\hspace{2cm}}$$

การวัดการกระจายสัมบูรณ์โดยใช้พิสัยระหว่างควอร์ไทล์มีข้อดีในกรณีที่ชุดข้อมูลมีข้อมูลที่แตกต่างจากข้อมูลตัวอื่นมาก เนื่องจากการคำนวณหาพิสัยระหว่างควอร์ไทล์จะใช้เพียง Q_1 และ Q_3 เท่านั้น ส่วนข้อมูลที่แตกต่างจากข้อมูลตัวอื่นมากจะมีค่าน้อยกว่า Q_1 หรือมากกว่า Q_3

นอกจากนี้ IQR สามารถนำไปใช้ในการตรวจสอบว่าข้อมูลใดเป็นค่านอกเกณฑ์ โดยค่านอกเกณฑ์คือข้อมูลที่มีค่าน้อยกว่า _____ หรือมากกว่า _____

ตัวอย่างที่ 15

ปริมาณพลังงาน (กิโลแคลอรี) ของอาหารจานเดียว 7 รายการ ที่จำหน่ายในโรงอาหารของโรงเรียนแห่งหนึ่ง แสดงได้ดังนี้

อาหารจานเดียว	ปริมาณพลังงาน (กิโลแคลอรี)
ข้าวราดแกงเขียวหวานไก่	338
ข้าวราดแกงไตปลา	319
ข้าวราดแกงส้มผักรวม	255
ข้าวราดผัดเผ็ดหอยลาย	424
ข้าวราดแกงพะแนงหมู	40+
ข้าวราดแกงฉู่ฉี่ปลาทุ	365
ข้าวราดผัดผักรวม	353

จงหาพิสัยระหว่างควอร์ไทล์ของข้อมูลชุดนี้

เมื่อเปรียบระหว่างพิสัยและพิสัยระหว่างควอร์ไทล์ จะเห็นว่าพิสัยสามารถหาได้สะดวก แต่ไม่เหมาะสำหรับใช้วัดการกระจายของข้อมูลในกรณีที่ข้อมูลที่มีค่าสูงหรือต่ำกว่าข้อมูลอื่นมาก โดยเฉพาะอย่างยิ่งในกรณีที่ข้อมูลมีค่า

นอกเหนือไปในขณะที่พิสัยระหว่างควอร์ไทล์สามารถใช้วัดการกระจายของข้อมูลในลักษณะนี้ได้ อย่างไรก็ตาม ทั้งพิสัยและพิสัยระหว่างควอร์ไทล์ไม่ได้ใช้ข้อมูลทุกตัวในการคำนวณเพื่อวัดการกระจาย

3) ส่วนเบี่ยงเบนมาตรฐาน (standard deviation) คือ ค่าที่ใช้วัดการกระจายของข้อมูลโดยเป็นค่าที่บอกให้ทราบว่าข้อมูลแต่ละตัวอยู่ห่างจากค่าเฉลี่ยเลขคณิตโดยเฉลี่ยประมาณเท่าใด สูตรของส่วนเบี่ยงเบนมาตรฐานมีดังนี้

ให้ $x_1, x_2, x_3, \dots, x_N$ แทนข้อมูล เมื่อ N แทนขนาดประชากร และให้ μ แทนค่าเฉลี่ยเลขคณิตของข้อมูลชุดนี้

ส่วนเบี่ยงเบนมาตรฐานของประชากร เขียนแทนด้วย σ (อ่านว่า ซิกมา) หาได้จาก

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}$$

ให้ $x_1, x_2, x_3, \dots, x_n$ แทนข้อมูล เมื่อ n แทนขนาดตัวอย่าง และให้ $\bar{x}$ แทนค่าเฉลี่ยเลขคณิตของข้อมูลชุดนี้

ส่วนเบี่ยงเบนมาตรฐานของตัวอย่าง เขียนแทนด้วย s หาได้จาก

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

ตัวอย่างที่ 16

ความสูง (เซนติเมตร) ของนักวอลเลย์บอลหญิงของโรงเรียนแห่งหนึ่งจำนวนทั้งหมด 10 คน แสดงได้ดังนี้

174 171 170 184 180 179 169 178 181 160

จงหาส่วนเบี่ยงเบนมาตรฐานของข้อมูลชุดนี้

4) ความแปรปรวน (variance) คือ ค่าที่ใช้วัดการกระจายของข้อมูล โดยคำนวณจากกำลังสองของส่วนเบี่ยงเบนมาตรฐาน จะได้สูตรของความแปรปรวน ดังนี้

ให้ $x_1, x_2, x_3, \dots, x_N$ แทนข้อมูล เมื่อ N แทนขนาดประชากร และให้ μ แทนค่าเฉลี่ยเลขคณิตของข้อมูลชุดนี้

ความแปรปรวนของประชากร เขียนแทนด้วย σ^2 (อ่านว่า ซิกมา) หาได้จาก

$$\sigma^2 = \underline{\hspace{2cm}}$$

ให้ $x_1, x_2, x_3, \dots, x_n$ แทนข้อมูล เมื่อ n แทนขนาดตัวอย่าง และให้ $\bar{x}$ แทนค่าเฉลี่ยเลขคณิตของข้อมูลชุดนี้

ความแปรปรวนของตัวอย่าง เขียนแทนด้วย s^2 หาได้จาก

$$s^2 = \underline{\hspace{2cm}}$$

จากตัวอย่างที่ 16 จะได้ว่าความแปรปรวนของความสูงของนักวอลเลย์บอลหญิงจำนวน 10 คน คือ _____ เซนติเมตร²

ตัวอย่างที่ 17

ในการศึกษาอายุขัยเฉลี่ย (ปี) ของสัตว์เลี้ยงลูกด้วยน้ำนม นักวิทยาศาสตร์ได้สุ่มตัวอย่างสัตว์เลี้ยงลูกด้วยน้ำนมมา 10 ชนิด พบว่าอายุขัยเฉลี่ยของสัตว์แต่ละชนิดเป็นดังนี้

สัตว์เลี้ยงลูกด้วยน้ำนม	อายุขัยเฉลี่ย (ปี)
แมว	13
วัว	15
สุนัข	12
ลา	12
แพะ	8
หนูตะเภา	4
ม้า	20

หนู	10
กระต่าย	5
แกะ	12

จงหาส่วนเบี่ยงเบนมาตรฐานและความแปรปรวนของอายุขัยเฉลี่ยของสัตว์เลี้ยงลูกด้วยน้ำนม 10 ชนิดนี้

การวัดการกระจายสัมพัทธ์

ในการเปรียบเทียบข้อมูลตั้งแต่สองชุดขึ้นไป เพื่อพิจารณาว่าข้อมูลชุดใดมีการกระจายมาก ข้อมูลชุดใดมีการกระจายน้อย ถ้านำค่าที่ได้จากการวัดการกระจายสัมบูรณ์ของข้อมูลแต่ละชุดมาเปรียบเทียบกันโดยตรง อาจให้ข้อมูลสรุปที่คลาดเคลื่อนไปจากความเป็นจริง เช่น ข้อมูลชุดหนึ่งมีค่าตั้งแต่ 0 ถึง 10 มีส่วนเบี่ยงเบนมาตรฐาน 2.2 และข้อมูลชุดอีกชุดหนึ่งมีค่าตั้งแต่ 200 ถึง 800 มีส่วนเบี่ยงเบนมาตรฐาน 60.5 ถ้าพิจารณาเฉพาะส่วนเบี่ยงเบนมาตรฐานของข้อมูลทั้งสองชุด อาจทำให้เข้าใจว่าข้อมูลชุดที่หนึ่งมีการกระจายน้อยกว่าข้อมูลชุดที่สองซึ่งอาจไม่ถูกต้องนัก เพราะค่าของข้อมูลสองชุดนี้ต่างกันมาก ค่ากลางและค่าวัดการกระจายของข้อมูลทั้งสองชุดย่อมต่างกันมากเช่นกัน เพื่อให้การเปรียบเทียบมีความหมาย จึงนิยามหาอัตราส่วนของค่าที่ได้จากการวัดการกระจายสัมบูรณ์กับค่ากลางของ

ข้อมูลชุดนั้น ๆ แล้วจึงนำอัตราส่วนที่หาได้มาเปรียบเทียบกัน ในที่นี้จะพิจารณาเฉพาะสัมประสิทธิ์การแปรผัน (coefficient of variation) โดยมีสูตรดังนี้

สัมประสิทธิ์การแปรผันของประชากร = _____

สัมประสิทธิ์การแปรผันของตัวอย่าง = _____

สัมประสิทธิ์ของการแปรผันอาจเขียนในรูปเปอร์เซ็นต์ ได้ดังนี้

สัมประสิทธิ์การแปรผันของประชากร = _____

สัมประสิทธิ์การแปรผันของตัวอย่าง = _____

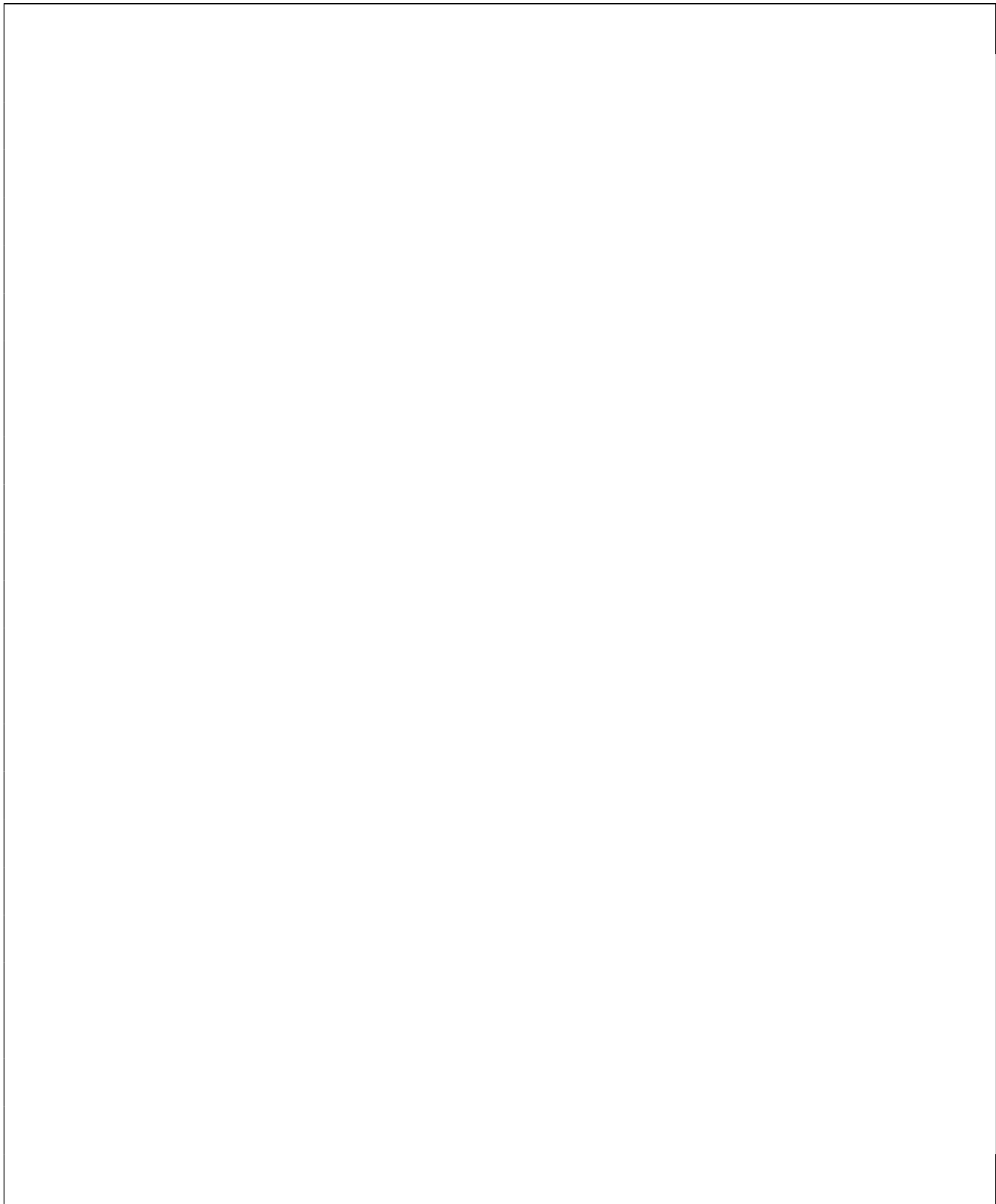
การเปรียบเทียบการกระจายของข้อมูลโดยใช้สัมประสิทธิ์การแปรผันนั้น ถ้าสัมประสิทธิ์การแปรผันของข้อมูลชุดใดมีค่ามากกว่า หมายความว่าข้อมูลชุดนั้นมีการกระจายออกจากค่าเฉลี่ยเลขคณิตมากกว่า หรือกล่าวได้ว่าข้อมูลชุดนั้นเกาะกลุ่มกันน้อยกว่าข้อมูลอีกชุดหนึ่ง

ตัวอย่างที่ 18

ในการเปรียบเทียบคะแนนสอบวิชาคณิตศาสตร์และวิชาภาษาอังกฤษของนักเรียนห้องหนึ่งซึ่งมีคะแนนเต็มวิชาละ 100 คะแนน ครูประจำชั้นได้สุ่มตัวอย่างนักเรียนห้องนี้มา 10 คน พบว่าคะแนนสอบแต่ละวิชาของนักเรียนแต่ละคนเป็นดังนี้

นักเรียนคนที่	1	2	3	4	5	6	7	8	9	10
คะแนนสอบวิชาคณิตศาสตร์	56	62	76	90	78	81	88	79	80	75
คะแนนสอบวิชาภาษาอังกฤษ	78	74	63	89	76	75	85	90	73	74

จงหาสัมประสิทธิ์การแปรผันของคะแนนสอบวิชาคณิตศาสตร์และวิชาภาษาอังกฤษของนักเรียนที่สุ่มตัวอย่างมา 10 คนนี้ พร้อมทั้งเปรียบเทียบการกระจายของคะแนนสอบทั้งสองวิชาของนักเรียนที่สุ่มตัวอย่างมา 10 คนนี้



ข้อสังเกต สัมประสิทธิ์การแปรผันไม่มีหน่วย