



MathStandard

เอกสารประกอบการเรียนคณิตศาสตร์

ระดับมัธยมศึกษาตอนปลาย

เรื่อง สถิติ

www.mathstandard.com

เรื่อง สถิติ

คำชี้แจง

จัดทำโดย MathStandard

ฉบับ 1.0

เอกสารประกอบการเรียน เรื่อง สถิติ ครอบคลุมข้อมูลและการนำเสนอ ค่ากลาง ตำแหน่งข้อมูล การกระจาย ค่ามาตรฐาน การแจกแจงปกติ และความสัมพันธ์ของข้อมูล

ความรู้พื้นฐานที่ใช้ ได้แก่ ร้อยละ สมการ ฟังก์ชัน ผลรวม ลำดับ การจัดหมู่ และความน่าจะเป็น

แบบฝึกหัด ประกอบ ด้วย ระดับ ง่าย กลาง ยาก ประยุกต์ และ ฝึกวิเคราะห์ พร้อม แบบ ทดสอบ ทบทวนท้ายบท เฉลยเฉพาะคำตอบและเฉลยวิธีคิดอยู่ส่วนท้ายของเอกสาร

ข้อตกลงและสัญลักษณ์

- $x_{(j)}$ คือค่าลำดับที่ j หลังเรียงข้อมูลจากน้อยไปมาก ต้องเก็บค่าซ้ำครบทุกค่า
- N, μ, σ^2, σ ใช้กับ จำนวน ค่าเฉลี่ย ความแปรปรวน และ ส่วนเบี่ยงเบนมาตรฐาน ประชากร ส่วน $n, \bar{x}, s^2, s$ ใช้กับตัวอย่าง โดยระบุขอบเขตในแต่ละข้อ
- ความกว้างอันตรภาคชั้นใช้ I ตัวพิมพ์ใหญ่ จุดกึ่งกลางชั้นใช้ m_j
- ตำแหน่งควอร์ไทล์ เดซิส์ และเปอร์เซ็นต์ไทล์รายตัวใช้ $N + 1$ และแทรกเชิงเส้นตามข้อตกลงในเนื้อหา ส่วนจัดชั้นใช้จำนวนสะสมจาก N
- ค่าจากจุดกึ่งกลางหรือการแทรกภายในชั้นเป็นค่าประมาณ เว้นแต่ทราบข้อมูลรายตัวเพียงพอ
- $P(E)$ คือความน่าจะเป็นของเหตุการณ์ E ส่วน P_r คือเปอร์เซ็นต์ไทล์ที่ r
- $\Phi(z) = P(Z \leq z)$ เป็น พื้นที่ด้านซ้าย ของโค้งปกติ มาตรฐาน ใช้ตาราง ภายใต สมมติฐาน การแจกแจงปกติ
- $\hat{y}$ คือค่าทำนาย แยกจากค่าจริง y เก็บเศษส่วนหรือทศนิยมไว้ก่อนปิดเฉพาะขั้นสุดท้ายตามที่โจทย์ กำหนด

ข้อมูลเพิ่มเติม www.mathstandard.com

เรื่อง สถิติ

สารบัญ

คำชี้แจง	1
ตอนที่ 1 ข้อมูลและการนำเสนอ	
เนื้อหาและตัวอย่าง	4
แบบฝึกหัดข้อ 1–20	11
ตอนที่ 2 ค่ากลางและการประยุกต์	
เนื้อหาและตัวอย่าง	21
แบบฝึกหัดข้อ 21–40	27
ตอนที่ 3 ตำแหน่งข้อมูลและแผนภาพกล่อง	
เนื้อหาและตัวอย่าง	37
แบบฝึกหัดข้อ 41–60	43

เรื่อง สถิติ

สารบัญ (ต่อ)

ตอนที่ 4 การกระจายและการแปลงข้อมูล

เนื้อหาและตัวอย่าง	53
แบบฝึกหัดข้อ 61–80	60

ตอนที่ 5 ค่ามาตรฐานและการแจกแจงปกติ

เนื้อหาและตัวอย่าง	70
แบบฝึกหัดข้อ 81–100	77

ตอนที่ 6 ความสัมพันธ์ของข้อมูลและแบบจำลอง

เนื้อหาและตัวอย่าง	87
แบบฝึกหัดข้อ 101–120	94
สรุปเนื้อหา เรื่อง สถิติ	104
แบบทดสอบทบทวนท้ายบท	110
เฉลยเฉพาะคำตอบ	122
เฉลยวิธีคิด	129

เรื่อง สถิติ

ตอนที่ 1 ข้อมูลและการนำเสนอ

เป้าหมาย จำแนกข้อมูล สร้างและอ่านตารางความถี่ เลือกกราฟ และพิจารณาขอบเขตจากข้อมูลจัดชั้น

1.1 ข้อมูล ประชากร และตัวอย่าง

ประชากรคือหน่วยทั้งหมดภายในขอบเขตการศึกษา ตัวอย่างคือส่วนหนึ่งที่เลือกมาศึกษา ค่าที่สรุปจากประชากรเรียกว่า **พารามิเตอร์** ส่วนค่าที่สรุปจากตัวอย่างเรียกว่า **ค่าสถิติ**

การจำแนกข้อมูล

- **เชิงคุณภาพ** บอกประเภทหรือคุณลักษณะ เช่น ชนิดสินค้า รหัสที่เขียนด้วยตัวเลขอาจเป็นเพียงชื่อประเภท
- **เชิงปริมาณไม่ต่อเนื่อง** ได้จากการนับ เช่น จำนวนผู้เข้าใช้บริการ
- **เชิงปริมาณต่อเนื่อง** ได้จากการวัด เช่น มวล เวลา และระยะทาง
- **ปฐมภูมิ** เก็บเพื่อการศึกษาโดยตรง **ทุติยภูมิ** นำข้อมูลที่มีผู้เก็บไว้แล้วมาใช้
- **ภาคตัดขวาง** เปรียบเทียบหลายหน่วยในช่วงเดียวกัน **อนุกรมเวลา** ติดตามค่าตามเวลา



สถิติเชิงพรรณนาสรุปข้อมูลที่เก็บได้ สถิติเชิงอนุมานใช้ข้อมูลตัวอย่างอ้างอิงไปยังประชากร วิธีเลือกตัวอย่างจึงมีผลต่อข้อสรุป แม้ตัวอย่างขนาดใหญ่ก็อาจมีอคติได้

ตัวอย่าง 1 โรงงานมีหลอดไฟ 20000 หลอด สุ่มทดสอบอายุการใช้งาน 100 หลอด จงระบุประชากร ตัวอย่าง และชนิดของข้อมูลอายุการใช้งาน

วิธีคิด ขอบเขตศึกษาคือหลอดไฟทั้งรุ่น ผู้ถูกทดสอบเป็นเพียงส่วนหนึ่ง และอายุการใช้งานเกิดจากการวัดเวลา

ตอบ ประชากร 20000 หลอด ตัวอย่าง 100 หลอด ข้อมูลเชิงปริมาณต่อเนื่อง

1.2 ความถี่และความถี่สะสม

ความถี่คือจำนวนครั้งที่พบค่าหรือประเภทรูปนั้น ความถี่สะสมคือจำนวนข้อมูลตั้งแต่ค่าแรกจนถึงค่าที่พิจารณา โดยต้องเรียงค่าจากน้อยไปมากก่อน กำหนด f_j เป็นความถี่ที่ค่า j และ F_j เป็นความถี่สะสมถึงค่านั้น

ผลรวมและผลต่างความถี่สะสม

$$N = \sum_{j=1}^k f_j, \quad F_j = \sum_{i=1}^j f_i, \quad f_j = F_j - F_{j-1}$$

N คือจำนวนข้อมูลทั้งหมด k คือจำนวนค่าหรือชั้น และกำหนด $F_0 = 0$

ตัวอย่าง 2 ข้อมูลจำนวนหนังสือที่ยืมคือ 0, 1, 1, 1, 2, 2, 3, 3, 3, 3 จงสร้างตารางความถี่และความถี่สะสม
วิธีคิด

ค่าข้อมูล x	ความถี่ f	ความถี่สะสม F
0	1	1
1	3	4
2	2	6
3	4	10

เช่น ค่า 1 มีสามค่า คือ 1, 1, 1 ส่วนสะสมถึง 2 นับ 0, 1, 1, 1, 2, 2 รวม 6 ค่า

ตอบ ความถี่ 1, 3, 2, 4 และสะสม 1, 4, 6, 10

1.3 ความถี่สัมพัทธ์และความถี่สะสมสัมพัทธ์

ความถี่สัมพัทธ์เปรียบเทียบความถี่กับจำนวนทั้งหมด ส่วนความถี่สะสมสัมพัทธ์เป็นสัดส่วนที่ไม่เกินค่าหรือขอบชั้นนั้น กำหนด f_j, F_j, N เป็นความถี่ ความถี่สะสม และจำนวนรวมตามลำดับ

สูตรและร้อยละ

$$r_j = \frac{f_j}{N}, \quad R_j = \frac{F_j}{N}, \quad \text{ร้อยละของชั้น} = 100r_j$$

$$\sum r_j = 1, \quad R_k = 1, \quad r_j = R_j - R_{j-1}$$

ตัวอย่าง 3 ค่าข้อมูล 1, 2, 3 มีความถี่ 2, 5, 3 จงหาความถี่สัมพัทธ์และสะสมสัมพัทธ์
วิธีคิด

ค่า	ความถี่	สัมพัทธ์	สะสมสัมพัทธ์
1	2	0.2	0.2
2	5	0.5	0.7
3	3	0.3	1

จำนวนรวม 10 จึงหารทุกความถี่ด้วย 10 เช่น สัดส่วนที่ไม่เกิน 2 คือ $\frac{2+5}{10} = 0.7$

ตอบ สัมพัทธ์ 0.2, 0.5, 0.3 สะสมสัมพัทธ์ 0.2, 0.7, 1

1.4 อันตรภาคชั้น ขอบชั้น และจุดกึ่งกลาง

ข้อมูลจำนวนมากอาจจัดเป็นช่วง โดยแต่ละค่าต้องอยู่ในชั้นเดียว กำหนด L_j, U_j เป็นขอบล่างและขอบบนของชั้นที่ j ใช้ I_j แทนความกว้างชั้น และ m_j แทนจุดกึ่งกลาง

ความกว้างและจุดกึ่งกลางชั้น

$$I_j = U_j - L_j, \quad m_j = \frac{L_j + U_j}{2}$$

ข้อมูลบันทึกละเอียด d หน่วย มีขอบชั้นห่างชิดจำกัดชั้นด้านละ $\frac{d}{2}$ เมื่อชั้นจำนวนเต็มต่อกันใช้ $d = 1$

การสร้างตารางเริ่มจากระบุค่าต่ำสุดและสูงสุด เลือกจำนวนชั้น k แล้วกำหนดความกว้างให้ครอบคลุมข้อมูลทั้งหมด การเลือก k ไม่มีคำตอบตายตัว ต้องคงรายละเอียดเพียงพอและไม่สร้างชั้นแคบเกินจำเป็น

ตัวอย่าง 4 ข้อมูลจำนวนเต็มคือ 11, 13, 14, 16, 17, 18, 19, 21, 22, 24, 25, 29 ต้องการสี่ชั้นกว้างเท่ากัน โดยเริ่มที่ 10 จงสร้างตาราง

วิธีคิด ใช้ชั้นกว้าง 5 ครอบคลุมถึง 29 แล้วนับข้อมูลแต่ละค่า

ชั้น	ขอบชั้น	จุดกึ่งกลาง	ความถี่
10–14	9.5, 14.5	12	3
15–19	14.5, 19.5	17	4
20–24	19.5, 24.5	22	3
25–29	24.5, 29.5	27	2

ตอบ ความกว้าง 5 ความถี่ 3, 4, 3, 2

1.5 แผนภูมิแท่งและฮิสโทแกรม

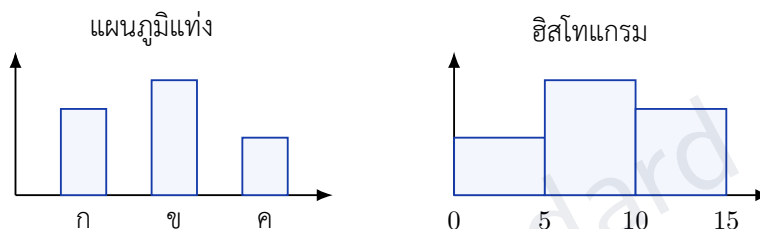
แผนภูมิแท่งเปรียบเทียบหมวดหมู่หรือค่าที่แยกจากกัน จึงเว้นช่องระหว่างแท่ง ฮิสโทแกรมแสดงข้อมูลเชิงปริมาณที่จัดเป็นช่วงต่อเนื่อง โดยฐานแท่งเป็นขอบชั้นและแท่งติดกัน

ความถี่กับพื้นที่แท่ง

เมื่อความกว้างชั้นเท่ากัน ใช้ความสูงแทนความถี่ได้ ถ้าความกว้างต่างกัน ต้องใช้ความหนาแน่นเพื่อให้พื้นที่แท่งแทนความถี่

$$\text{ความหนาแน่นความถี่} = \frac{f_j}{I_j}, \quad f_j = I_j \times \text{ความสูงแท่ง}$$

f_j คือความถี่ และ I_j คือความกว้างชั้น ถ้าใช้ความหนาแน่นสัมพัทธ์ พื้นที่แท่งแทน $\frac{f_j}{N}$



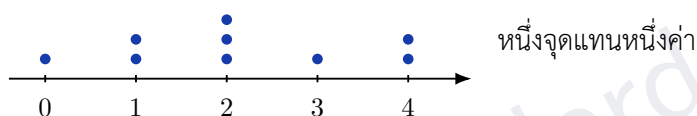
ตัวอย่าง 5 ฮิสโทแกรมมีชั้น $[0, 4)$, $[4, 10)$ ความถี่ 8, 9 จงหาความหนาแน่นและเปรียบเทียบความสูง

วิธีคิด ความกว้างคือ 4, 6 จึงได้ความสูง $\frac{8}{4} = 2$ และ $\frac{9}{6} = 1.5$ แท่งแรกสูงกว่า แม้มีความถี่น้อยกว่า

ตอบ ความหนาแน่น 2, 1.5 แท่งแรกสูงกว่า

1.6 แผนภาพจุดและแผนภาพต้นและใบ

แผนภาพจุดใช้หนึ่งจุดแทนข้อมูลหนึ่งค่า จุดที่ซ้อนเหนือค่าเดียวกันแสดงข้อมูลซ้ำ จึงรักษาค่ารายตัวไว้ ส่วนต้นและใบแยกหลักของตัวเลขและต้องกำหนดความหมายของหลักให้ชัดเจน



การอ่านข้อมูลรายตัว

จุดสามจุดเหนือเลข 2 แทนข้อมูล 2, 2, 2 ไม่ใช่ค่า 6 สำหรับต้นเป็นหลักสิบและใบเป็นหลักหน่วย 3 | 7 หมายถึง 37 ต้องเก็บใบเข้าครบทุกใบ และเรียงใบก่อนอ่านอันดับข้อมูล

ตัวอย่าง 6 ต้นเป็นหลักสิบ ใบเป็นหลักหน่วย มีข้อมูล 2 | 1 4 4 และ 3 | 0 5 จงเขียนข้อมูลจากน้อยไปมาก

วิธีคิด แฉวแรกให้ 21, 24, 24 แฉวที่สองให้ 30, 35 ใบ 4 สองใบต้องนับเป็นสองค่า

ตอบ 21, 24, 24, 30, 35

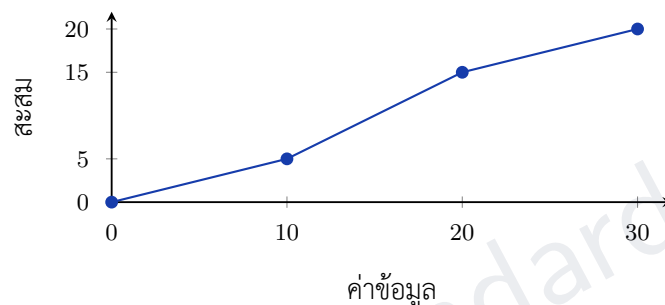
1.7 เส้นความถี่สะสมและการแทรกค่า

ใช้ขอบชั้นบนเป็น พิกัด แนวนอน และความถี่สะสมเป็น พิกัด แนวตั้ง พร้อม จุด เริ่ม ที่ ขอบ ชั้น แรก และ ความถี่สะสมศูนย์ กราฟต้องไม่ลดลง หากเชื่อมจุดเป็นเส้นตรงถือเป็นแบบจำลองการกระจายสม่ำเสมอ ภายในแต่ละชั้น

อ่านจำนวนในช่วงจากค่าที่สะสม

ให้ $F(x)$ แทนจำนวนสะสมที่ไม่เกิน x จะได้จำนวนในช่วง $(a, b]$ เท่ากับ $F(b) - F(a)$ ถ้าอ่านค่า ภายในชั้นจากเส้นตรง ผลที่ได้เป็นเพียงค่าประมาณ

ตัวอย่าง 7



กราฟสะสมผ่าน $(0, 0)$, $(10, 5)$, $(20, 15)$, $(30, 20)$ เชื่อมเป็นเส้นตรง จงประมาณจำนวนที่ไม่เกิน 16 และ จำนวนในช่วง $(10, 26]$

วิธีคิด ใช้การแทรกตามสัดส่วนระยะภายในชั้น ได้

$$F(16) = 5 + \frac{6}{10}(15 - 5) = 11, \quad F(26) = 15 + \frac{6}{10}(20 - 15) = 18$$

จำนวนในช่วง $(10, 26]$ เท่ากับ $F(26) - F(10) = 18 - 5 = 13$

ตอบ ประมาณ 11 ค่า และ 13 ค่า

1.8 เกณฑ์ที่อยู่ภายในชั้นและข้อจำกัดของข้อมูล

ตารางจัดชั้นไม่ระบุค่ารายตัว หากจุดตัดอยู่กลางชั้น จำนวนน้อยที่สุดนับเฉพาะชั้นที่ผ่านแน่นอน ส่วน จำนวนมากที่สุดรวมทั้งชั้นที่ครอบคลุมจุดตัดด้วย ต้องแยกการหาขอบเขตที่ยืนยันได้จากการประมาณด้วย เส้นตรง

หลักพิจารณาเกณฑ์

- ตรวจสอบเครื่องหมาย $>$, $\geq$, $<$, $\leq$ และความละเอียดที่บันทึก
- แปลงสะสมเป็นความถี่ก่อนแยกชั้นผ่าน ชั้นไม่ผ่าน และชั้นครอบคลุมเกณฑ์
- ไม่สมมติว่าข้อมูลกระจายเท่ากันในชั้น เว้นแต่โจทย์กำหนดแบบจำลองสำหรับการประมาณ

ตัวอย่าง 8 คะแนนจำนวนเต็ม 0–9, 10–19, 20–29, 30–39 มีความถี่ 3, 7, 8, 2 จงหาจำนวนน้อยที่สุดและมากที่สุดที่ได้มากกว่า 24

วิธีคิด ขั้นสุดท้ายผ่านแนสองคน ชั้น 20–29 มีแปดคน อาจได้ไม่เกิน 24 ทุกคนหรือเกิน 24 ทุกคน จึงได้ขอบเขต 2 ถึง 10

ตอบ น้อยที่สุด 2 คน มากที่สุด 10 คน

สรุป

เริ่มจากชนิดและขอบเขตข้อมูล แล้วเลือกการนำเสนอให้เหมาะสม ความถี่นับรายค่า สะสมเป็นผลรวมจากด้านล่าง ฮิสโทแกรมชั้นกว้างต่างกันต้องใช้พื้นที่แทนความถี่ ตารางจัดชั้นอาจให้เพียงค่าประมาณหรือขอบเขต ไม่กำหนดข้อมูลรายตัวทั้งหมด

เรื่อง สถิติ

แบบฝึกหัดตอนที่ 1

ง่าย

ข้อ 1 ต้องการศึกษเวลาดำเนินทางของนักเรียนทั้งโรงเรียน 1200 คน โดยสุ่มเก็บข้อมูล 80 คน จงระบุประชากร ตัวอย่าง และชนิดข้อมูลเวลาดำเนินทาง

ตอบ

ข้อ 2 ข้อมูลจำนวนครั้งที่ใช้ห้องสมุดคือ 0, 1, 1, 2, 2, 2, 3, 3 จงหาความถี่ของ 2 และความถี่สะสมถึง 2

ตอบ

เรื่อง สถิติ

ตอนที่ 2 ค่ากลางและการประยุกต์

เป้าหมาย คำนวณและเลือกค่ากลาง แก่ข้อมูลสรุ่ยย้อนกลับ และประยุกต์กับลำดับ อัตรา และปัญหา ค่าต่ำสุด

2.1 ค่าเฉลี่ยเลขคณิต

ค่าเฉลี่ยใช้ข้อมูลทุกค่า กำหนด $x_1, \dots, x_N$ เป็นข้อมูลทั้งชุด และ $\bar{x}$ เป็นค่าเฉลี่ย สำหรับประชากรมักใช้ สัญลักษณ์ μ ส่วนตัวอย่างใช้ $\bar{x}$

ผลรวมและค่าเฉลี่ย

$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N}, \quad \sum x_i = N\bar{x}$$

ถ้าค่า x_j มีความถี่ f_j ใช้ $N = \sum f_j$ และ $\bar{x} = \frac{\sum f_j x_j}{N}$

$$\sum (x_i - \bar{x}) = 0, \quad \overline{ax + b} = a\bar{x} + b$$

a, b เป็นค่าคงตัวเดียวกันสำหรับทุกข้อมูล ค่าเฉลี่ยไวต่อค่าที่ห่างจากกลุ่มมาก

ตัวอย่าง 9 เวลาอ่านหนังสือ 4 วันคือ 20, 30, 30, 40 นาที จงหาค่าเฉลี่ย และหาผลรวมส่วนเบี่ยงเบนจากค่าเฉลี่ย

วิธีคิด ผลรวมเวลา 120 จึงเฉลี่ย 30 ผลรวมส่วนเบี่ยงเบน $-10 + 0 + 0 + 10 = 0$

ตอบ 30 นาที และ 0

2.2 มัธยฐาน ฐานนิยม และการเลือกค่ากลาง

เรียงข้อมูลเป็น $x_{(1)} \leq \dots \leq x_{(N)}$ วงเล็บใต้สัญลักษณ์แสดงอันดับหลังเรียง มัธยฐานแบ่งตำแหน่งเป็นสองส่วน ฐานนิยมคือค่าหรือประเภทที่มีความถี่สูงสุด อาจมีหลายค่าหรือไม่มีค่าที่เด่นกว่าค่าอื่น

มัธยฐานของข้อมูลรายตัว

$$Me = \begin{cases} x_{(\frac{N+1}{2})}, & N \text{ คี่} \\ \frac{x_{(\frac{N}{2})} + x_{(\frac{N}{2}+1)}}{2}, & N \text{ คู่} \end{cases}$$

ค่าเฉลี่ยเหมาะเมื่อใช้ทุกค่าร่วมคำนวณ มัธยฐานทนต่อค่าที่สุดปลายมากกว่า ฐานนิยมใช้กับข้อมูลเชิงคุณภาพได้ ต้องเลือกตามวัตถุประสงค์และลักษณะการกระจาย

$$\begin{array}{cccccc} x_{(1)} & x_{(2)} & x_{(3)} & x_{(4)} & x_{(5)} & x_{(6)} \\ & & \underline{x_{(3)} + x_{(4)}} & & & \\ Me = & & \frac{\quad}{2} & & & \end{array}$$

ตัวอย่าง 10 ข้อมูล 2, 4, 4, 5, 7, 20 จงหาค่าเฉลี่ย มัธยฐาน และฐานนิยม พร้อมเลือกค่าที่สะท้อนตำแหน่งกึ่งกลางโดยลดอิทธิพลค่า 20

วิธีคิด ค่าเฉลี่ย $\frac{42}{6} = 7$ มัธยฐาน $\frac{4+5}{2} = 4.5$ ฐานนิยม 4 หากต้องลดอิทธิพลค่าสุดปลายให้ใช้มัธยฐาน

ตอบ 7, 4.5, 4 ตามลำดับ และเลือกมัธยฐาน

2.3 ค่าเฉลี่ยถ่วงน้ำหนักและการรวมกลุ่ม

น้ำหนักแทนความสำคัญหรือจำนวนข้อมูลที่ค่าหนึ่งเป็นตัวแทน กำหนด $w_i \geq 0$ และผลรวมน้ำหนักเป็นบวก ถ้ารวมกลุ่มให้ใช้จำนวนข้อมูลแต่ละกลุ่มเป็นน้ำหนัก

สูตรถ่วงน้ำหนัก

$$\bar{x}_w = \frac{\sum w_i x_i}{\sum w_i}, \quad \bar{x}_{รวม} = \frac{\sum n_j \bar{x}_j}{\sum n_j}$$

n_j คือจำนวนข้อมูลกลุ่มที่ j และ $\bar{x}_j$ คือค่าเฉลี่ยกลุ่มนั้น การเฉลี่ยค่าเฉลี่ยโดยตรงใช้ได้เมื่อทุกกลุ่มมีขนาดเท่ากัน

ตัวอย่าง 11 กลุ่ม ก มีผู้เรียน 12 คนเฉลี่ย 65 คะแนน กลุ่ม ข มี 8 คนเฉลี่ย 80 คะแนน จงหาค่าเฉลี่ยรวม

วิธีคิด ผลรวมคะแนน $12(65) + 8(80) = 1420$ จำนวนผู้เรียน 20 คน ค่าเฉลี่ยจึง $\frac{1420}{20}$

ตอบ 71 คะแนน

2.4 การเพิ่ม ลบ และแก้ข้อมูลผิด

เริ่มจากผลรวมเดิม ไม่แก้ค่าเฉลี่ยด้วยการบวกหรือลบค่าผิดโดยตรง ให้ S เป็นผลรวม และ N เป็นจำนวนข้อมูล

ปรับผลรวมและจำนวนข้อมูลแยกกัน

$$S_{\text{ใหม่}} = S_{\text{เดิม}} - \sum x_{\text{ออก}} + \sum x_{\text{เข้า}}, \quad \bar{x}_{\text{ใหม่}} = \frac{S_{\text{ใหม่}}}{N_{\text{ใหม่}}}$$

การแก้ค่าที่บันทึกผิดไม่เปลี่ยนจำนวนข้อมูล แต่การเพิ่มหรือนำออกต้องเปลี่ยนจำนวนด้วย หากมีข้อมูลไม่ทราบค่า ใช้ผลรวมร่วมกับลำดับ มัธยฐาน หรือความถี่ของฐานนิยม และตรวจทุกเงื่อนไข

ตัวอย่าง 12 ข้อมูล 8 ค่ามีค่าเฉลี่ยที่บันทึกไว้ 15 แต่ค่าจริง 21 ถูกบันทึกเป็น 13 จงหาค่าเฉลี่ยจริง

วิธีคิด ผลรวมที่บันทึก $8(15) = 120$ แก้เป็น $120 - 13 + 21 = 128$ จำนวนยัง 8

ตอบ ค่าเฉลี่ยจริง 16

2.5 ค่ากลางของข้อมูลจัดชั้น

ไม่ทราบค่ารายตัวจึงใช้จุดกึ่งกลางแทนค่าในแต่ละชั้นเพื่อประมาณค่าเฉลี่ย กำหนด f_j เป็นความถี่ m_j เป็นจุดกึ่งกลาง และ $N = \sum f_j$

ค่าเฉลี่ย มัธยฐาน และฐานนิยมโดยประมาณ

$$\bar{x} \approx \frac{\sum f_j m_j}{N}, \quad Me \approx L + \frac{\frac{N}{2} - F}{f} I$$

ในสูตรมัธยฐาน L คือขอบล่างของชั้นที่มีตำแหน่งสะสม $\frac{N}{2}$ ค่า F คือสะสมก่อนชั้นนั้น f คือความถี่ในชั้น และ I คือความกว้างชั้น

$$Mo \approx L_m + \frac{d_1}{d_1 + d_2} I, \quad d_1 = f_m - f_{\text{ก่อน}}, \quad d_2 = f_m - f_{\text{ถัดไป}}$$

สูตรฐานนิยมนี้ใช้กับชั้นกว้างเท่ากันและมีชั้นฐานนิยมเด่นชั้นเดียว L_m คือขอบล่างชั้นฐานนิยม f_m คือความถี่ชั้นนั้น ส่วน d_1, d_2 เป็นผลต่างความถี่กับชั้นติดกัน ไม่ใช่ความถี่สะสม

การแทรก มัธยฐาน สมมติ ข้อมูล กระจาย สม่ำเสมอใน ชั้น ส่วน สูตร ฐานนิยม อาศัย การ ประมาณ รูปทรง ฮิสโทแกรม จึงต้องระบุว่าเป็นค่าประมาณ

ตัวอย่าง 13 ข้อมูลจำนวนเต็มมีชั้น 20–24, 25–29, 30–34 ความถี่ 4, 8, 4 จงประมาณค่ากลางทั้งสามวิธีคิด

ชั้น	จุดกึ่งกลาง m	ความถี่ f	fm	สะสม
20–24	22	4	88	4
25–29	27	8	216	12
30–34	32	4	128	16

ค่าเฉลี่ย $\frac{432}{16} = 27$ ชั้นมัธยฐานมี $L = 24.5, F = 4, f = 8, I = 5$ จึง $Me \approx 24.5 + \frac{(8-4)5}{8} = 27$

ฐานนิยมมี $d_1 = d_2 = 4$ จึง $Mo \approx 24.5 + \frac{4(5)}{8} = 27$

ตอบ ประมาณ 27 ทั้งสามค่า

2.6 ค่าเฉลี่ยเรขาคณิตและค่าเฉลี่ยฮาร์โมนิก

ข้อมูลบวก $x_1, \dots, x_N$ มีค่ากลางที่เหมาะสมกับการคูณสะสมหรืออัตราส่วนต่างจากค่าเฉลี่ยเลขคณิต

ค่ากลางของข้อมูลบวก

$$AM = \frac{\sum x_i}{N}, \quad GM = \sqrt[N]{\prod x_i}, \quad HM = \frac{N}{\sum \frac{1}{x_i}}$$

$$HM \leq GM \leq AM$$

เท่ากันเมื่อข้อมูลทุกค่าเท่ากัน GM ใช้กับตัวคูณการเติบโตทบต้น ส่วนอัตราเร็วเฉลี่ยต้องใช้ระยะทางรวมหารเวลารวม กรณีระยะทางแต่ละช่วงเท่ากันจึงเป็น HM ของอัตราเร็ว

ตัวอย่าง 14 เดินทางสองช่วงระยะเท่ากัน ด้วยอัตราเร็ว 30 และ 60 กิโลเมตรต่อชั่วโมง จงหาอัตราเร็วเฉลี่ย

วิธีคิด ให้ระยะแต่ละช่วง d กิโลเมตร เวลารวม $\frac{d}{30} + \frac{d}{60} = \frac{d}{20}$ ระยะรวม $2d$ จึงอัตราเร็วเฉลี่ย 40

ตอบ 40 กิโลเมตรต่อชั่วโมง

2.7 ค่ากลางกับผลรวมความคลาดเคลื่อน

ค่าเฉลี่ยและมัธยฐานเป็นคำตอบของปัญหาต่ำสุดคนละแบบ โดย c เป็นค่าที่เลือกเปรียบเทียบกับข้อมูล

กำลังสองและค่าสัมบูรณ์

$$\sum (x_i - c)^2 = \sum (x_i - \bar{x})^2 + N(c - \bar{x})^2$$

ผลรวมกำลังสองต่ำสุดเมื่อ $c = \bar{x}$ ส่วน $\sum |x_i - c|$ ต่ำสุดที่มัธยฐาน ถ้า N คู่ ทุก c ระหว่างสองลำดับกลางให้ค่าต่ำสุดเท่ากัน

$$\sum (x_i + d)^2 = \sum x_i^2 + 2d \sum x_i + Nd^2$$

d เป็นค่าคงตัว สูตรสุดท้ายใช้เชื่อมผลรวมและผลรวมกำลังสองจากข้อมูลสรุปหลายชุด

ตัวอย่าง 15 ข้อมูล 1, 3, 7, 9 จงหาค่า c ที่ทำให้ผลรวมกำลังสองต่ำสุด และช่วง c ที่ทำให้ผลรวมค่าสัมบูรณ์ต่ำสุด

วิธีคิด ค่าเฉลี่ยคือ 5 จึงกำลังสองต่ำสุดที่ 5 ข้อมูลคู่มีสองค่ากลาง 3, 7 ค่าสัมบูรณ์จึงต่ำสุดทุกค่าระหว่างนั้น

ตอบ $c = 5$ สำหรับกำลังสอง และ $3 \leq c \leq 7$ สำหรับค่าสัมบูรณ์

2.8 ข้อมูลจากลำดับและอนุกรม

ข้อมูลที่กำหนดเป็นลำดับต้องเรียงตามค่าก่อนใช้สูตรมัธยฐาน โดยเฉพาะลำดับสลับเครื่องหมาย ส่วนค่าเฉลี่ยใช้ผลรวมพจน์ทุกพจน์ จำนวนข้อมูลไม่เปลี่ยนหลังเรียง

สูตรเชื่อมโยงที่ใช้

$$a_n = a_1 r^{n-1}, \quad \sum_{n=1}^N a_n = \frac{a_1(1-r^N)}{1-r} \quad (r \neq 1)$$

$$\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}, \quad \sum_{n=1}^N \frac{1}{n(n+1)} = 1 - \frac{1}{N+1}$$

a_1, r, N คือพจน์แรก อัตราส่วนร่วม และจำนวนพจน์ตามลำดับ ถ้าต้องย้อนหาจำนวนข้อมูลจากมัธยฐาน ให้แยกกรณีจำนวนคู่และคี่

ตัวอย่าง 16 ข้อมูล $a_n = \frac{1}{n(n+1)}$ สำหรับ $n = 1, \dots, 5$ จงหามัธยฐานและค่าเฉลี่ย

วิธีคิด ข้อมูลลดลงตาม n มัธยฐานยังเป็นพจน์ที่สามคือ $\frac{1}{12}$ ผลรวมแยกหักล้างได้ $1 - \frac{1}{6} = \frac{5}{6}$ ทหารห้าจึงได้ $\frac{1}{6}$

ตอบ $Me = \frac{1}{12}, \bar{x} = \frac{1}{6}$

สรุป

ใช้ค่าเฉลี่ยกับผลรวมและการถ่วงน้ำหนัก ใช้มัธยฐานกับอันดับหลังเรียง และใช้ฐานนิยมกับความถี่สูงสุด ข้อมูลจัดชั้นให้ค่าประมาณ การเพิ่มลบข้อมูลต้องปรับทั้งผลรวมและจำนวนข้อมูล ผลรวมกำลังสองต่ำสุดที่ค่าเฉลี่ย ส่วนผลรวมค่าสัมบูรณ์ต่ำสุดที่มัธยฐาน

เรื่อง สถิติ

แบบฝึกหัดตอนที่ 2

ง่าย

ข้อ 21 ข้อมูล 3, 5, 5, 7, 10 จงหาค่าเฉลี่ย มัธยฐาน และฐานนิยม

ตอบ

ข้อ 22 ค่าข้อมูล 2, 4, 7 มีความถี่ 2, 3, 1 จงหาค่าเฉลี่ยและมัธยฐาน

ตอบ