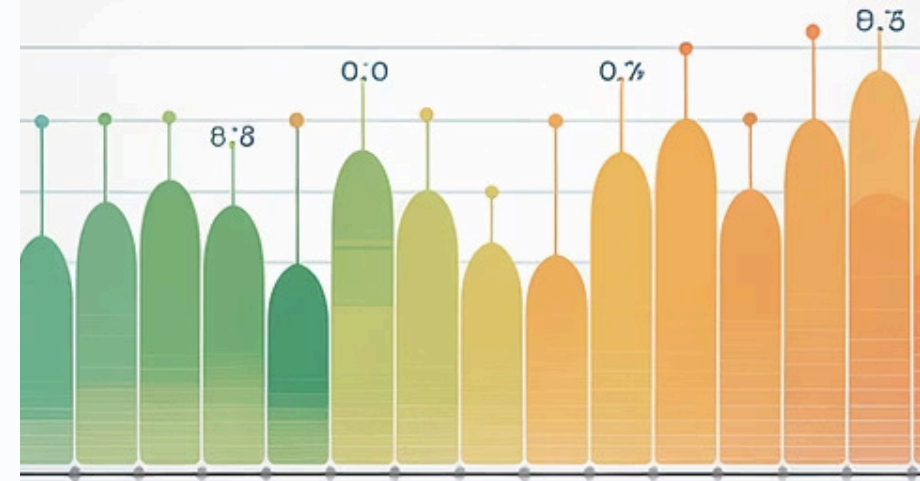


ตำแหน่งเปอร์เซ็นต์ไทล์และการวัดตำแหน่งข้อมูล

BUA3127 สถิติธุรกิจ

Dr. Khwanchat Wongjantip

Understanding Statistical Distribution



ประเภทของการวัดตำแหน่งข้อมูล

ควอไทล์ (Quartile: Qr)

แบ่งข้อมูลออกเป็น 4 ส่วนเท่าๆ กัน

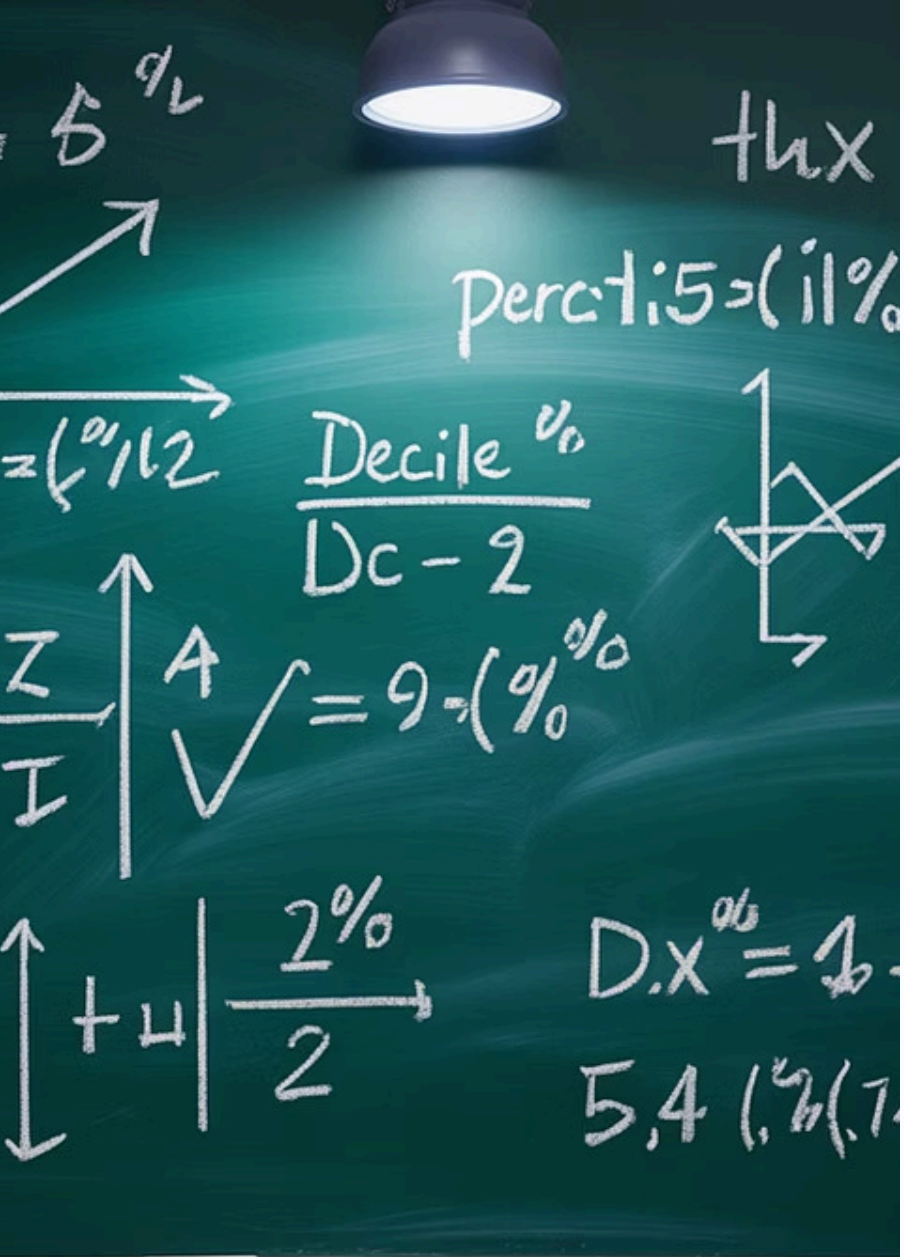
เดไซล์ (Decile: Dr)

แบ่งข้อมูลออกเป็น 10 ส่วนเท่าๆ กัน

เปอร์เซ็นต์ไทล์ (Percentile: Pr)

แบ่งข้อมูลออกเป็น 100 ส่วนเท่าๆ กัน

การวัดตำแหน่งเหล่านี้ช่วยให้เราเข้าใจการกระจายตัวของข้อมูลและสามารถเปรียบเทียบค่าต่างๆ ในชุดข้อมูลได้อย่างมีประสิทธิภาพ



สูตรการคำนวณตำแหน่งข้อมูล

$$Q_r = \frac{r}{4}(N + 1)$$

$$D_r = \frac{r}{10}(N + 1)$$

$$P_r = \frac{r}{100}(N + 1)$$

โดยที่ r คือลำดับของตำแหน่งที่ต้องการหา และ N คือจำนวนข้อมูลทั้งหมด

ความสัมพันธ์ระหว่างการวัดตำแหน่ง



ควอไทล์ (Quartile)

$Q1 = P25 = D2.5$

$Q2 = P50 = D5$

$Q3 = P75 = D7.5$



เดไซล์ (Decile)

$D1 = P10$

$D2 = P20$

$D3 = P30$

และอื่นๆ



เปอร์เซ็นต์ไทล์ (Percentile)

ละเอียดที่สุด แบ่งเป็น 100 ส่วน

$P1, P2, P3, \dots, P99$

ตัวอย่างที่ 1: การหาค่า P30

ข้อมูล: 34, 18, 16, 22, 38, 12, 40, 18, 30

จัดเรียงข้อมูล: 12, 16, 18, 18, 22, 30, 34, 38, 40

$$P_r = \frac{r}{100}(N + 1)$$

$$P_{30} = \frac{30}{100}(9 + 1) = 3$$

ดังนั้น P30 คือข้อมูลลำดับที่ 3 ซึ่งมีค่าเท่ากับ 18

ตัวอย่างที่ 2: การหาค่า P50

ข้อมูล: 54, 44, 59, 55, 46, 41, 35, 37

จัดเรียงข้อมูล: 35, 37, 41, 44, 46, 54, 55, 59

$$P_{50} = \frac{50}{100}(8 + 1) = 4.50$$

ตำแหน่ง 4.5 อยู่ระหว่างข้อมูลลำดับที่ 4 (44) และลำดับที่ 5 (46)

$$P_{50} = 44 + (46-44) \times (4.5-4) \div (5-4) = 44 + 2 \times 0.5 = 45$$

สูตรการคำนวณสำหรับข้อมูลแจกแจงความถี่

$$Q_r = L + I \left[\frac{\frac{Nr}{4} - \sum F_l}{f_m} \right]$$

$$D_r = L + I \left[\frac{\frac{Nr}{10} - \sum F_l}{f_m} \right]$$

$$P_r = L + I \left[\frac{\frac{Nr}{100} - \sum F_l}{f_m} \right]$$

เมื่อ L = ขีดจำกัดล่างที่แท้จริงของชั้นที่ Q_r , D_r , P_r อยู่

i = ความกว้างของอันตรภาคชั้น

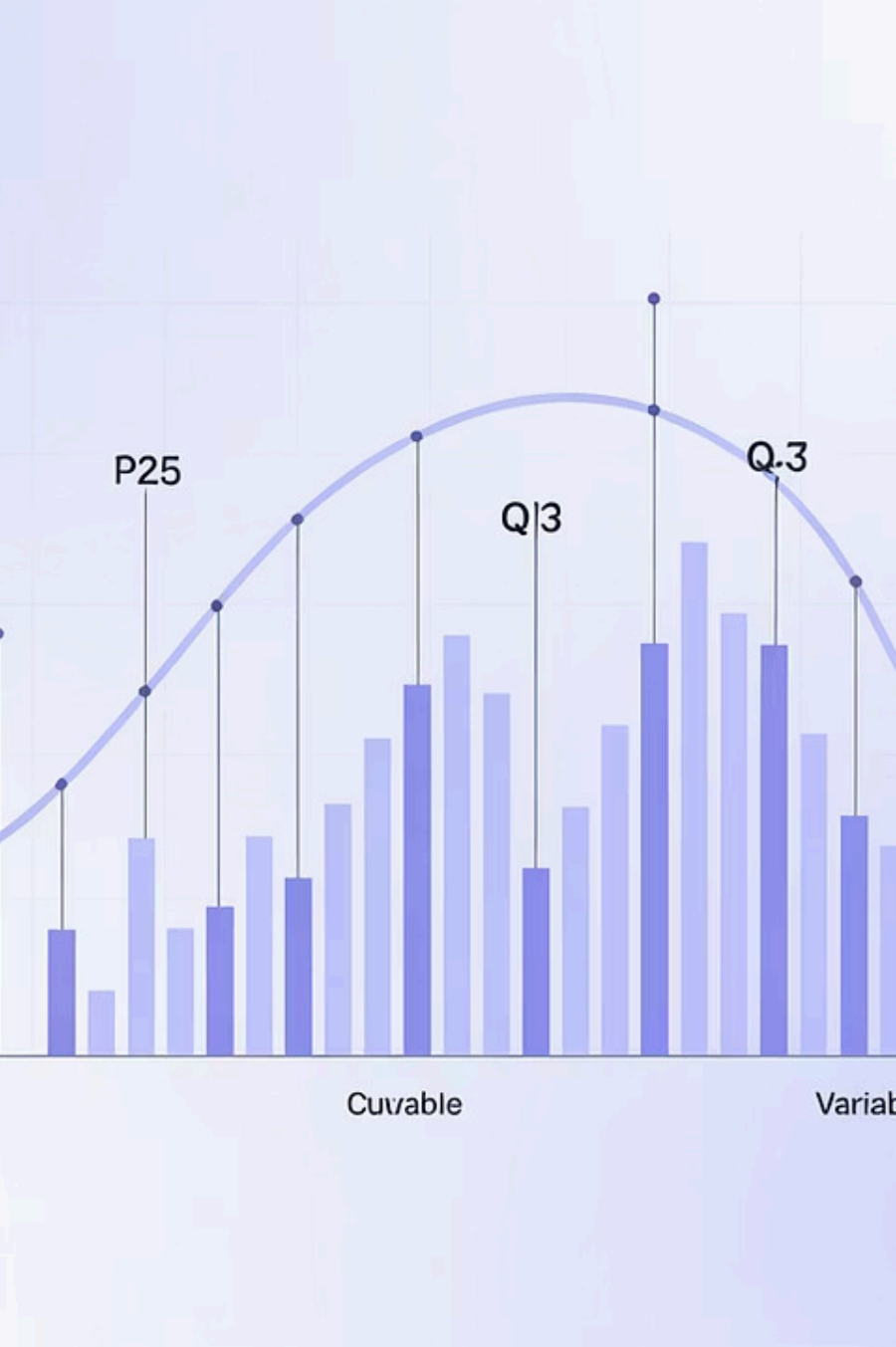
$\sum F_l$ = ความถี่สะสมชั้นที่อยู่ก่อนชั้นที่มีมัธยฐานไปทางค่าน้อย

f_m = ความถี่ของค่ามัธยฐานในชั้นที่มีมัธยฐาน

ตัวอย่างที่ 3: ข้อมูลแจกแจงความถี่

ชั้น	ความถี่	ความถี่สะสม
5-9	13	13
10-14	14	27
15-19	13	40
20-24	17	57
25-29	16	73
30-34	14	87
35-39	12	99
40-44	13	112

จำนวนข้อมูลทั้งหมด n = 112



การหาค่า P25 และ Q3 จากข้อมูล แจกแจงความถี่

การหาตำแหน่ง

P25 อยู่ที่ตำแหน่ง: $(112 \times 25) \div 100 = 28$

จากตารางความถี่สะสม P25 อยู่ในชั้น 15-19

Q3 อยู่ที่ตำแหน่ง: $(112 \times 3) \div 4 = 84$

จากตารางความถี่สะสม Q3 อยู่ในชั้น 30-34

การคำนวณค่า

$$P_{25} = 14.5 + 5 \left[\frac{28 - 27}{13} \right] = 14.5 + 0.38 = 14.88$$

$$Q_3 = 29.5 + 5 \left[\frac{84 - 73}{14} \right] = 29.5 + 3.93 = 33.43$$

ตัวอย่างเพิ่มเติม: การหาค่า P30 และ D5

ข้อมูล: 34, 8, 6, 22, 38, 2, 40, 18, 30

จัดเรียงข้อมูล: 2, 6, 8, 18, 22, 30, 34, 38, 40

การหาค่า P30

$$P_{30} = \frac{30}{100}(9 + 1) = 3$$

ดังนั้น P30 คือข้อมูลลำดับที่ 3 ซึ่งมีค่าเท่ากับ 8

การหาค่า D5

$$D_5 = \frac{5}{10}(9 + 1) = 5$$

ดังนั้น D5 คือข้อมูลลำดับที่ 5 ซึ่งมีค่าเท่ากับ 22

ข้อมูลชุดใหญ่: 32 ค่า

ข้อมูล: 71, 70, 69, 69, 69, 64, 64, 63, 61, 60, 59, 58, 58, 57, 56, 55

54, 54, 54, 54, 53, 52, 52, 51, 50, 50, 49, 47, 40, 39, 34, 30

Q1 (P25)

$$Q_1 = \frac{1}{4}(32 + 1) = 8.25$$

ตำแหน่ง 8.25 อยู่ระหว่างข้อมูลลำดับที่ 8 (63) และลำดับที่ 9 (61)

$$Q_1 = 63 - (63-61) \times 0.25 = 63 - 0.5 = 62.5$$

Q2 (P50)

$$Q_2 = \frac{2}{4}(32 + 1) = 16.5$$

ตำแหน่ง 16.5 อยู่ระหว่างข้อมูลลำดับที่ 16 (55) และลำดับที่ 17 (54)

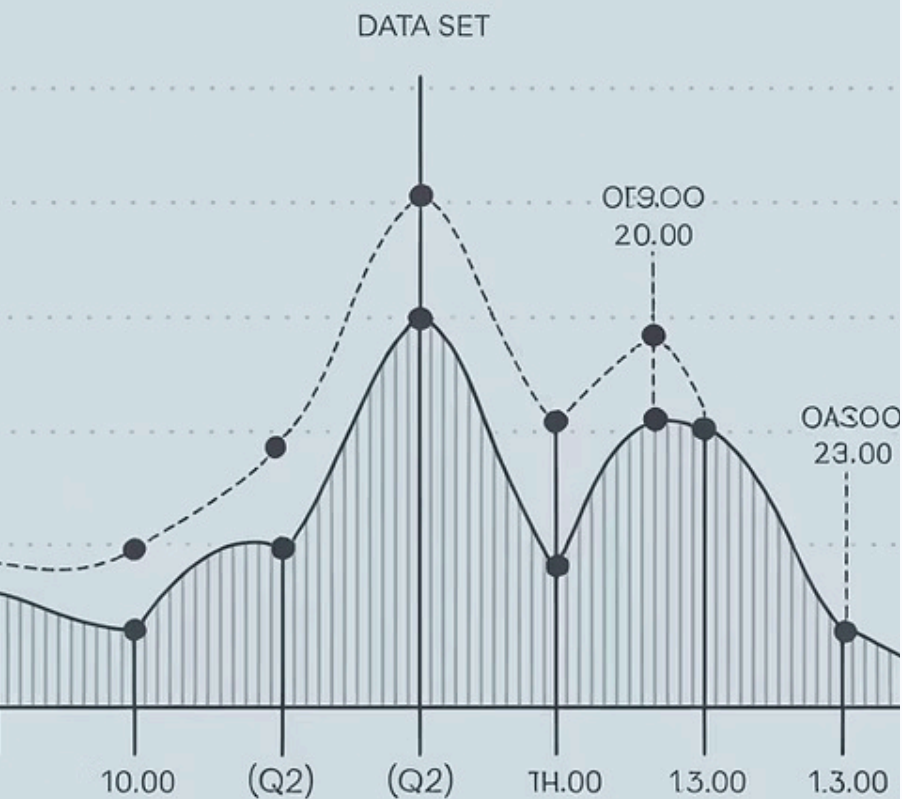
$$Q_2 = 55 - (55-54) \times 0.5 = 55 - 0.5 = 54.5$$

Q3 (P75)

$$Q_3 = \frac{3}{4}(32 + 1) = 24.75$$

ตำแหน่ง 24.75 อยู่ระหว่างข้อมูลลำดับที่ 24 (51) และลำดับที่ 25 (50)

$$Q_3 = 51 - (51-50) \times 0.75 = 51 - 0.75 = 50.25$$

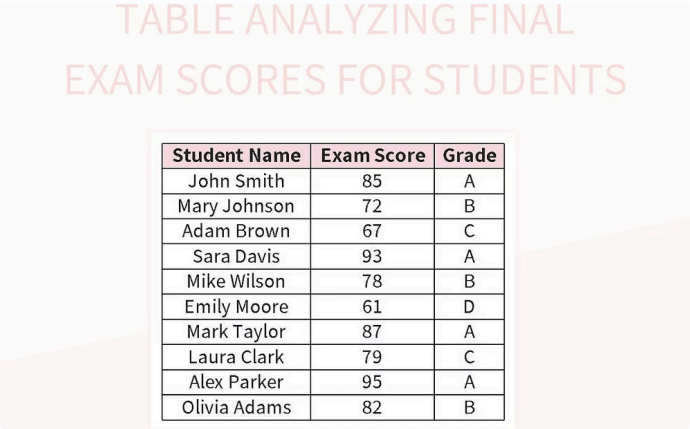


การวิเคราะห์ข้อมูลด้วยตำแหน่งเปอร์เซ็นต์ไทล์

การวัดตำแหน่งข้อมูลช่วยให้เราสามารถ:

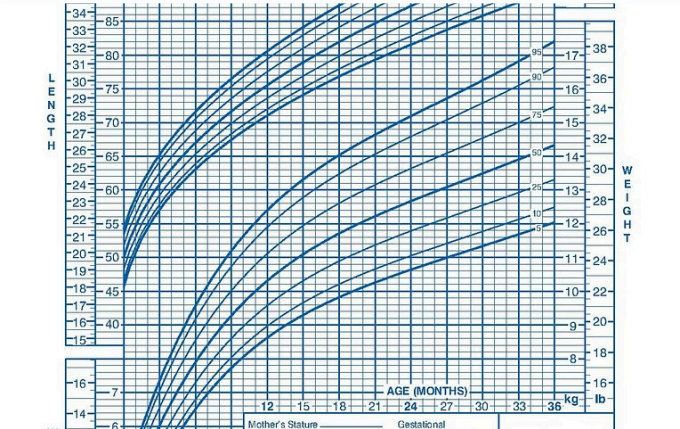
- เปรียบเทียบตำแหน่งของข้อมูลในชุดข้อมูลเดียวกัน
- ระบุค่าผิดปกติ (Outliers) ในชุดข้อมูล
- สร้างแผนภาพกล่อง (Box Plot) เพื่อแสดงการกระจายตัวของข้อมูล
- วิเคราะห์การกระจายตัวของข้อมูลโดยไม่ต้องใช้ค่าเฉลี่ย

การประยุกต์ใช้ในชีวิตประจำวัน



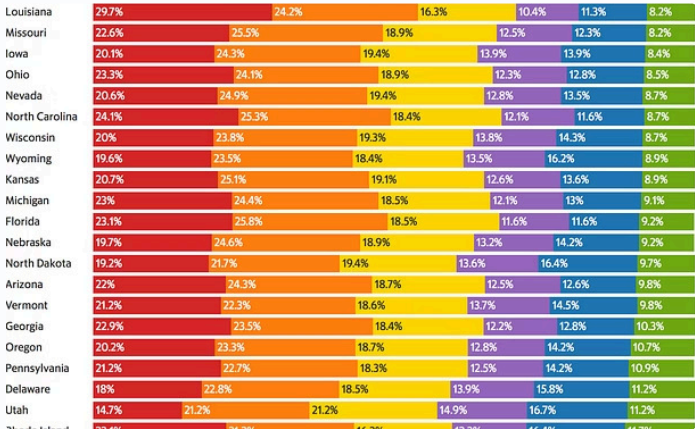
การศึกษา

ใช้เปอร์เซ็นต์ไทล์ในการวัดผลการเรียน เช่น คะแนนสอบมาตรฐาน เพื่อเปรียบเทียบผลการเรียนของนักเรียนกับกลุ่ม



การแพทย์

ใช้เปอร์เซ็นต์ไทล์ในการติดตามการเจริญเติบโตของเด็ก เช่น น้ำหนักและส่วนสูงตามอายุ



เศรษฐศาสตร์

ใช้เปอร์เซ็นต์ไทล์ในการวิเคราะห์การกระจายรายได้ในประชากร เช่น รายได้มัธยฐาน



ข้อควรระวังในการใช้ตำแหน่งเปอร์เซ็นต์

⚠️ ข้อจำกัดของการวัดตำแหน่ง

- ไม่ได้ให้ข้อมูลเกี่ยวกับค่าเฉลี่ยหรือการกระจายตัวของข้อมูลทั้งหมด
- อาจให้ผลลัพธ์ที่คลาดเคลื่อนในกรณีที่มีข้อมูลน้อยเกินไป
- มีหลายวิธีในการคำนวณซึ่งอาจให้ผลลัพธ์ที่แตกต่างกัน
- ต้องระมัดระวังในการตีความผลลัพธ์ โดยเฉพาะเมื่อเปรียบเทียบระหว่างกลุ่มข้อมูลที่แตกต่างกัน

สรุป: การวัดตำแหน่งข้อมูล

เข้าใจประเภทของการวัดตำแหน่ง

ควอไทล์ (4 ส่วน), เดไซล์ (10 ส่วน), เปอร์เซ็นไทล์ (100 ส่วน)

คำนวณและตีความผลลัพธ์

ใช้ผลลัพธ์เพื่อเปรียบเทียบตำแหน่งของข้อมูลและวิเคราะห์การกระจายตัว

เลือกสูตรที่เหมาะสม

ข้อมูลดิบใช้สูตร $Pr = r(N+1)/100$ ส่วนข้อมูลแจกแจงความถี่ใช้สูตรที่ซับซ้อนขึ้น

ประยุกต์ใช้ในสถานการณ์จริง

นำไปใช้ในการศึกษา การแพทย์ เศรษฐศาสตร์ และอื่นๆ

การวัดตำแหน่งข้อมูลเป็นเครื่องมือสำคัญในการวิเคราะห์ข้อมูลทางสถิติ ช่วยให้เราเข้าใจการกระจายตัวของข้อมูลได้ดียิ่งขึ้น