

สถิติ (Statistics)

สถิติ (Statistics - ในระดับอุดมศึกษา เรามักเรียกสั้นๆ ว่า Stat) เป็นวิชาที่มีประโยชน์ และสำคัญมากในวิชาชีพและ
ข้อมูลและเราสามารถนำข้อมูลเหล่านั้น มาเป็นเครื่องมือ ซึ่งช่วยในการตัดสินใจทางเศรษฐกิจ ได้

ซึ่งหัวข้อที่ต้องรู้เกี่ยวกับเรื่องนี้

1. ข้อมูล (Data) คืออะไร หมายถึง สิ่งที่มีขอบเขตชัดเจนหรือที่แท้จริงของเรื่องที่เราสนใจ

ข้อมูล อาจมีหลายแบบ เช่น

▷ ข้อมูลเชิงคุณภาพ (Qualitative Data) = ความสูง, ความยาว, ความถี่, ความเร็ว, ความดัน
(หน่วย cm) = 172, 162, 167, 165

▷ ข้อมูลเชิงปริมาณ (Quantitative Data) = นม, นม, นม, นม, นม

↳ ข้อมูลที่เป็นตัวเลข

2. แหล่งข้อมูล (Source) ได้มาจาก

- 2.1 ไปคัดลอกข้อมูลที่มีอยู่รวบรวมไว้แล้ว
- 2.2 รวบรวมโดยตรงจากสิ่งที่เราสนใจศึกษา

การนำเสนอข้อมูล

- นำเสนอด้วยตาราง
- นำเสนอด้วยแผนภูมิแท่ง
- นำเสนอด้วยแผนภูมิวงกลม
- นำเสนอด้วยภาพเคลื่อนไหว

P. 88

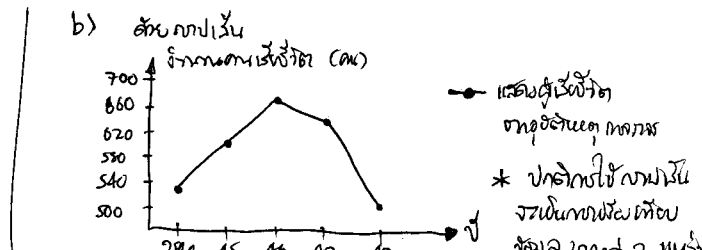
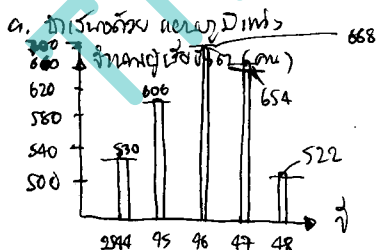
พื้นที่เกษตร

ตารางที่ 1 จำนวนผู้เลี้ยงสัตว์จากข้อมูลของกรมการเกษตร พ.ศ. 2544-2548

ปี	2544	2545	2546	2547	2548
ผู้เลี้ยงสัตว์	530	606	668	654	522

1. นำข้อมูลมาแสดงด้วยกราฟแท่ง

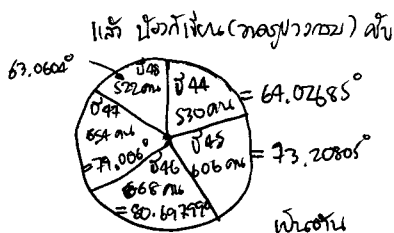
↳ แสดงด้วย 2-3 แกนก็ได้ หรือ อาจจะแสดงเป็นกราฟที่ 0. ส่วนในของกราฟก็ได้



c) นำเสนอด้วยแผนภูมิวงกลม

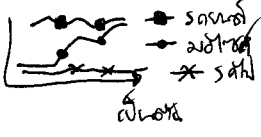
ปีแรกผู้เลี้ยงสัตว์ทั้งหมดรวมกัน = 360° ส่วนแล้ว

$$\begin{aligned} \text{ปี 44} &= \frac{530}{2980} \times 360^\circ \\ \text{ปี 45} &= \frac{606}{2980} \times 360^\circ \\ \text{ปี 46} &= \frac{668}{2980} \times 360^\circ \\ \text{ปี 47} &= \frac{654}{2980} \times 360^\circ \\ \text{ปี 48} &= \frac{522}{2980} \times 360^\circ \end{aligned}$$



$$\begin{aligned} 530 &\rightarrow \frac{530 \times 360}{2980} = 64.02685^\circ \\ 606 &\rightarrow \frac{606 \times 360}{2980} = 73.20805^\circ \\ 668 &\rightarrow \frac{668 \times 360}{2980} = 80.67785^\circ \\ 654 &\rightarrow \frac{654 \times 360}{2980} = 79.00671^\circ \\ 522 &\rightarrow \frac{522 \times 360}{2980} = 63.0604^\circ \end{aligned}$$

เมื่อ นำข้อมูลมาแสดงด้วยกราฟแท่งจาก
ตารางแล้ว ก็แสดงด้วยกราฟวงกลม
ก็จะได้ออกมา 2 แบบ ก็คือแบบที่ 1
ก็แสดงเป็น 3 เส้น แทน



ใช้ MS. Excel ทำก็ได้
ถ้าไม่แน่ใจในกราฟที่มีได้แล้ว (มีเลข 080-571-4623)

2) จากภาพที่ 3 แบบที่ มีพื้นที่ใช้สอยพื้นที่แคบ

จากถนนขนาด 3 เมตร เพื่อเลือกแบบที่พอ เพื่อแบบภูมิอากาศได้สบาย

มีแต่ใช้บรรทัดได้ใช้ใช้เส้นนี้

ถ้ามีโปร + 1 เมตร ใช้เส้นนี้

ใช้เส้นนี้ ใช้เส้นนี้ใช้บรรทัดมาตาม

ตารางที่ 2 จำนวนผู้เสียชีวิต 10 คนตามแผนผัง ช่วงสัปดาห์

2. มีตัวเลข 7 ระดับ กรณีนี้ เขาอธิบายจาก

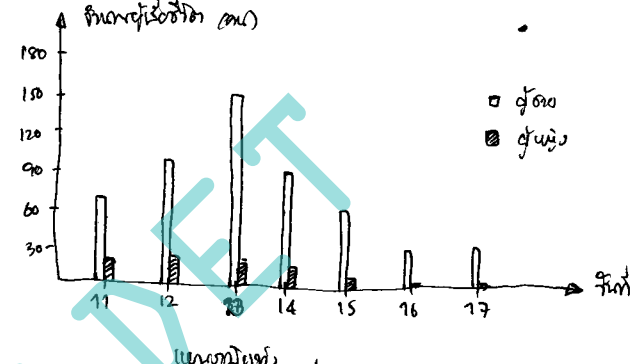
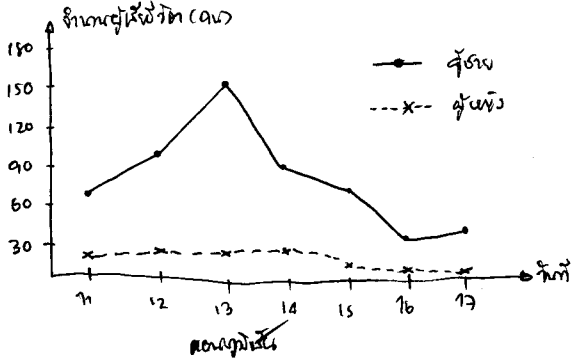
▷ วันที่ 11 ผู้ป่วยมีอาการอ่อน ผู้ที่เสียชีวิต

▷ วันที่ 12 _____

▷ วันที่ 13 _____

กรณีนี้ แบบภูมิอากาศไม่เหมาะสม: หากไม่มีการปรับพื้นที่ ในกรณีนี้ ผู้ป่วยมีอาการอ่อน ผู้ที่เสียชีวิต

ตัวแปรที่จะเป็นพื้นที่ได้ 30 คน เส้น และแบบภูมิอากาศ ดังนี้



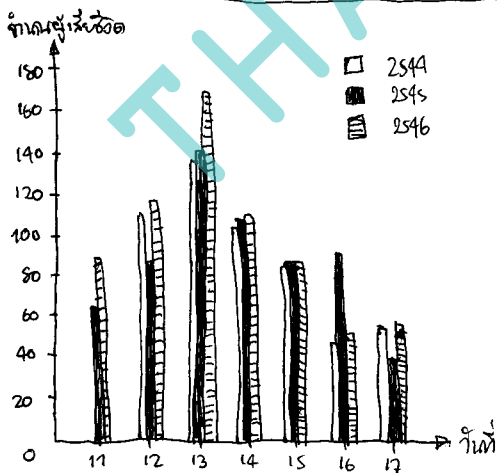
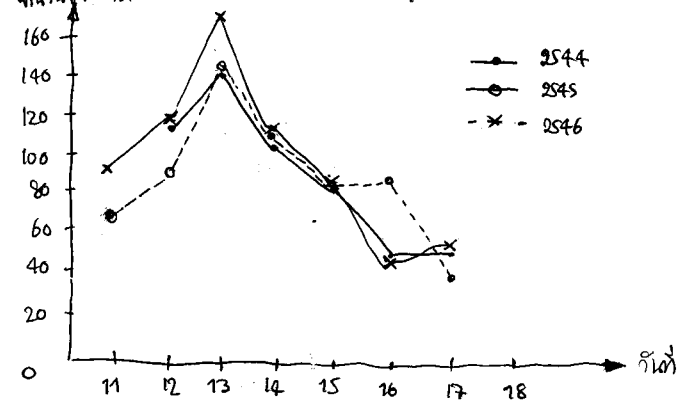
กรณีนี้หรือแบบภูมิอากาศไม่เหมาะสม และ (แบบภูมิอากาศ) ได้ข้อสังเกต

ตารางที่ 3

จำนวนผู้เสียชีวิต จากอุบัติเหตุทางรถ ช่วงสัปดาห์

วันที่	จำนวนผู้เสียชีวิต		
	2544	2545	2546
11	-	63	88
12	111	85	117
13	50	39	54

มีจุดสังเกตของพื้นที่เดิม คือแบบภูมิอากาศ และแบบภูมิอากาศ



ตอนนี้อาจไปทำเอง ใช้ scale ที่ถูกต้องตามระดับ แล้วใช้ใช้บรรทัด และบรรทัดมาตาม เพื่อตัวผู้เขียน *

P. 92 การนำเสนอข้อมูลในรูปแบบของแผนภูมิ

ข้อสอบ P. 95 ตารางแจกแจงความถี่ของอายุที่เริ่มต้นสอบ Math

ความถี่สอบ Math ของเด็ก 80 คน มีค่าเฉลี่ย = 100 คะแนน

จากข้อมูลในข้อ
P. 96

65	41	72	34	70
⋮				⋮
83	78	98	...	98
				60

1. ค่าเฉลี่ยที่ต่ำสุด = 34
ค่าเฉลี่ยที่สูงสุด = 98
2. พิสัย (Range) ได้มาจก ค่าเฉลี่ยที่สูงสุด - ค่าเฉลี่ยที่ต่ำสุด
= 98 - 34 = 64
3. เขตร้อยเปอร์เซ็นต์ = $\frac{\text{พิสัย}}{\text{ความถี่รวมทั้งหมด}}$

ซึ่งนี่เราคำนวณเองตามความหมาย
นั่นคือ ถ้าเราคำนวณโดยใช้วิธีนี้ค่าที่น้อย (ช่วงที่น้อยไป) จะได้อัตราเปอร์เซ็นต์ที่
เกิน 100 เปอร์เซ็นต์
→ ข้อสอบ

ถ้าเราคำนวณโดยใช้ความถี่ที่มีค่ามาก เราจะได้จำนวนเปอร์เซ็นต์ที่น้อย
(ช่วงที่มากกว่าๆ)
∴ เราต้องกำหนดให้พอดี ตามข้อที่ถามไว้ก่อน, กำหนดให้ = 10

∴ จำนวนเปอร์เซ็นต์ = $\frac{64}{10} = 6.4$

- แต่จำนวนเปอร์เซ็นต์จะเป็นทศนิยมไว้ได้ ต้องเป็นจำนวนทศนิยม ดังนั้น 6.4 ที่ได้อาจใช้ปัดเป็นจำนวนเต็ม *
* และถ้าตัวเลขเป็นจำนวนเต็ม ก็ใช้บอกเพิ่มไปอีก 1

- เรียงตามแนวนอน อยู่ในตารางชั้นแรก
∴ ค่าเฉลี่ยที่ต่ำสุด อยู่ในตารางชั้นแรก

ที่ข้อที่ 5. และ 6. จะได้ตารางแจกแจงความถี่ ดังนี้

ค่าเฉลี่ยที่ต่ำสุดคือ 34 อยู่ระหว่าง 31 และ 40
(31)

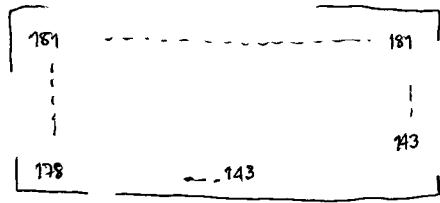
ความถี่สอบ	ช่วงที่ใดก็ได้ที่ทราบแล้ว	ความถี่
31 - 40		6
41 - 50		14
⋮	⋮	⋮
91 - 100		5
	รวม	80

ค่าเฉลี่ยที่สูงสุดคือ 98
อยู่ในชั้นนี้

มีข้อที่ 1 และข้อที่ 2 ต้องทำด้วย
จำนวนของนักเรียนที่แก้ข้อ (กลุ่มตัวอย่างทั้งหมดที่สอบ)

note: ความถี่รวมของเปอร์เซ็นต์อาจใช้ไม่ได้ทั้งที่ข้อที่ 5
ที่ขึ้นอยู่กับข้อที่ 6 หรือข้อที่ 7 ใช้ ข้อที่ 7 ของการแก้ไขข้อที่ 5 หรือข้อที่ 6 และคำนวณค่าของข้อที่ 5 และข้อที่ 6

ข้อ 4 จาก (76) จำนวนไม้ * มีขนาดที่ สามารถรวมตัวเท่ากับ 76 สามารถแบ่งออกเป็น



ค่าที่จุด = 139
ค่าที่จุด = 181

ตามคำถาม

1. ไม้ = ค่าที่จุด - ค่าที่จุด
= 181 - 139 = 42
2. กำหนดจำนวนชั้นตามต้น = 6
แล้ว จำนวนชั้นตามต้น = $\frac{46}{6} = 7$
ตัวเลข 7 เป็นเลขจำนวนเต็ม ให้บวก 1 เป็น 7+1 = 8
 \therefore จำนวนชั้นตามต้น = 8
3. ที่เหลือของไม้

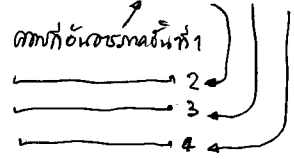
เลข 136 \rightarrow 141
เป็นเลขที่เรากำหนด
* 0 ตัวที่ก่อน ค่าที่จุด (คือ 139)
ตั้งอยู่ในตอนแรก

ท่อนไม้ (cm)	รอยต่อ	จำนวน
136 - 141		11
142 - 147		47
148 - 153		8
154 - 159		1
160 - 165		2
166 - 171		0
172 - 177		0
178 - 183		7

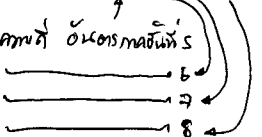
ค่าที่จุดคือ 181
อยู่ในตอนแรกด้วย

(76) \rightarrow ค่าที่จุดคือ 76
= ไม้รวม 76 จำนวน

4. ตามคำถาม
- 1) จำนวนไม้ที่เกินกำหนดคือ ท่อนไม้ที่อยู่ในช่วง 142 - 147 cm มีจำนวนที่รวมตามต้นที่จุด (คือ 47)
- 2) มีไม้เกินกำหนดที่ท่อนไม้ที่มีขนาด 160 cm อยู่ $= (11 + 47 + 8 + 1) = 67$ จำนวน



- 3) มีไม้เกินกำหนดที่ท่อนไม้ที่มีขนาด 160 cm นี้ อยู่ $2 + 0 + 0 + 7 = 9$ จำนวน



เลข 175 อยู่ระหว่าง 172 ถึง 177

- 4) จำนวนไม้ที่เกินกำหนด = 175 ไม้ที่เกิน กำหนดตามต้นที่จุดตามต้น 172 - 177 ไม้ #

การกระจายของคนที่ไปชมคอนเสิร์ต

1. อัตราคนขึ้นที่ 2 ถึง 9 มีความกว้างเท่ากันคือ 5 $\left[\begin{array}{l} \text{ใน } 15 \text{ ถึง } 19 \\ \text{เลข } 40 \text{ ถึง } 44 \end{array} \right]$

พออัตราคนขึ้นแรก มีความกว้าง = 15 (เลข 0 ถึงเลข 14)

อัตราคนขึ้นสูงที่ 9 ไม่สามารถหาความกว้างได้ เพราะไม่รู้ค่าช่วงบนที่มีอยู่สูงๆจนเกินไป \rightarrow อาจเป็น 120 ก็ได้!

2. หาค่าของ f ในข้อ 1.

3. หาค่าของ f ในข้อ 1.

4. ช่วงคน 10 คน ที่มีจำนวนคนที่ดู = คนที่ขึ้นรถที่ 5 อยู่ในช่วงอายุ 30-34 ปี ช่วงนี้มีจำนวน 439 คน (จากที่มีคนขึ้นรถได้)

5. จากตาราง เหนือ, อายุ 30-34 = 16^2

35-39 = 134

40-44 = 66

45-49 = 34

รวม 403 คน

6. ช่วงคนแต่ละช่วง อายุต่ำกว่า 50 ปี มี $20 + 2 + 56 + 157 + 292 + 254 + 144 + 87$

รวมที่รวมกับอัตราคนขึ้นแรกคืออัตราคนขึ้นที่ 8 (45-49)

= 992 คน

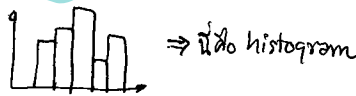
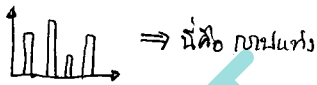
7. หาค่าของ $1070 - 651 = 419$ คน

8. บอกไม่ได้ เพราะไม่รู้จำนวนในตารางแจกแจงความถี่

P.101

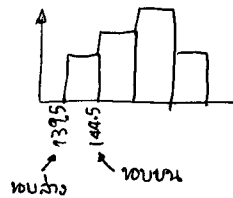
รอบล่าง - รอบบน

แบบ รอบล่าง - รอบบน เมื่อเขียน histogram (ฮิสโตแกรม) = แบ่งช่วงความถี่ \rightarrow ดูค่าเฉลี่ยของแต่ละช่วงความถี่



โดย แทนเลข เลขความถี่ของแต่ละช่วงอัตราคนขึ้น (ส่วนนี้ไม่ใช่ค่าที่ขึ้น) และตัวแปรตามที่จะใช้

* จุดปลายของด้านกว้างของรูปสี่เหลี่ยมมุมฉากแต่ละรูป คือ รอบล่าง และ รอบบน ของอัตราคนขึ้นที่เรากำลังดู



นี่ตรงกลางของแต่ละอัตราคนขึ้น มีรอบล่าง และรอบบน
ดังนั้น ความกว้างของแต่ละอัตราคนขึ้น = รอบบน - รอบล่าง
และ จุดกึ่งกลางที่ขึ้น = $\frac{\text{รอบบน} + \text{รอบล่าง}}{2}$ หรือ $\frac{\text{รอบล่าง} + \text{รอบบน}}{2}$

* ที่นำค่าที่ขึ้นนี้คือ เมื่อเราหาจุดกึ่งกลางของแต่ละอัตราคนขึ้นได้แล้ว

แล้วเราก็หาจุดกึ่งกลางของแต่ละจุดนั้น และต่อให้มีจุดบนแกน

ทวนซ้ำ 1 จุด และทวนซ้ำอีก 1 จุด ซึ่งเป็นจุดกึ่งกลางของแต่ละอัตราคนขึ้นก่อน อัตราคนขึ้นที่จุด

11. จุดกึ่งกลางของอัตราคนขึ้น ที่อยู่ถัดอัตราคนขึ้นสูงสุด เมื่อเราลากเส้นเชื่อมกันแล้ว

เราก็จะได้รูปคล้ายๆ รูปสามเหลี่ยม \rightarrow เรียกรูปนี้ว่า " รูปเขตรวมแล้วของตาราง "

ดูตัวอย่างหน้า 105 ครับ และ ~~***~~ เขาก็จะตรงกับค่าของช่วงล่างแล้ว

1. อ่านตารางที่ 1 รอบ ความถี่หรือเรื่อง
2. ให้หาช่วงของรอบล่าง ที่ 20, 35 หรืออื่นๆ
3. ดูตัวอย่างหน้าถัดไป ซึ่งจะเห็นในข้อถัดไป

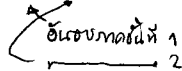
4. โทรหาผม

086-571-4623 \rightarrow มีข้อสงสัย

1. มีนักเรียนจากโรงเรียนต่าง ๆ 55 คน ใช้รถจักรยาน

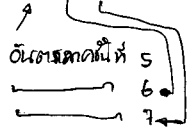
เวลา (นาที)	จำนวน (คน)
น้อยกว่า 15 นาที	3
15 - 20	7
21 - 24	12
25 - 30	15
31 - 35	8
36 - 40	3
มากกว่า 40 นาที	2
	50

1. time ≤ 20 นาที มี $3+7 = 10$ คน



2. ส่วนใหญ่ใช้รถจักรยานประมาณ 25-30 นาที (max. frequency = 15 คน)

3. time > 30 นาที มี $8+3+2 = 13$ คน



4. เวลาที่ใช้ในรถจักรยาน 55. มากกว่า 40 นาที มี 2 คน

5. นอนได้เมื่อไม่ใช้รถจักรยานใช้รถจักรยาน 25 นาที มี 15 คน * อาจไม่ได้มีนอน 5 นาที ใช้รถจักรยาน 15 คน นอนได้เมื่อใช้รถจักรยานที่ (4)

6. 33 นาที มี 1 คน * 8 คน (5)

* จากข้อ 5. และ 6. จักรยานที่นอนได้ใช้รถจักรยาน 25 นาที มี 1 คน, 2 คน หรือ 12 คน (ที่นอนได้) นอนได้เมื่อใช้รถจักรยาน 25-30 นาที มี 15 คน นอนได้เมื่อใช้รถจักรยาน

และนอนได้ 6 คน มี 1 คนนอนได้เมื่อใช้รถจักรยาน

* ถ้าใช้รถจักรยาน 10 นาทีขึ้นไปรถจักรยานที่นอนได้

7. รถจักรยานที่นอนได้ อาจมีคนที่ใช้รถจักรยาน 55. มากกว่า 1 ชั่วโมง จำนวน 2 คน

เพราะ คนที่ใช้รถจักรยาน 40 นาที มี 2 คน

(มากกว่า 40 นาที อาจเป็น 41, 42, 45, 50, 60, 70, 80, ... น นาที
ซึ่งมากกว่า 40 ดังนั้น everything is possible.)

2. เวลาว่างของนักเรียนที่สอบได้คะแนน จำนวน 80 คน ดังนี้

16	25	30
50		
72	14	49

คะแนนที่ได้ที่ ค่าต่ำสุด = 12

ค่าสูงสุด = 78

เฉลี่ย = ค่าสูงสุด - ค่าต่ำสุด

$$= 78 - 12 = 66$$

ค่าเบี่ยงเบนมาตรฐานของคะแนน = 10 แล้ว

$$\therefore \text{จำนวนนักเรียน} = \frac{66}{10} = 6.6 \text{ แล้ว } \frac{1}{10} \text{ ของ } 78 \text{ คือ } 7.8$$

มีข้อสังเกตด้วย ช่วง มีช่วงอายุ 70-79 คนนี้ เป็นคนในแถวแรกแถวสุดท้ายก็ได้ ตัวนี้ ถูกใส่แทนไป

ช่วงอายุ	รอยขีด	คน
10 - 19	HHH HHH	11
20 - 29	HHH HHH	10
30 - 39	HHH HHH	11
40 - 49	HHH HHH HHH	18
50 - 59	HHH HHH	13
60 - 69	HHH HHH	12
70 - 79	HHH	5
รวม		80

- จำนวนคนที่มอดอกสีม่วงมากที่สุดอยู่ในช่วงอายุ 40-49 ปี (มี 18 คน)
- จำนวนคนที่มอดอกสีม่วง ซึ่งมียุคที่มากกว่า 40 ปี มี $11 + 10 + 11 = 32$ คน
- ผู้มียุคที่มีน้อยกว่า 60 ปีมีไป มอดอกสีม่วง = $12 + 5 = 17$ คน
- ผู้มียุคที่มีน้อยกว่า 50 ปี มอดอกสีม่วง $11 + 10 + 11 + 18 = 50$ คน

3. สร. กำนันดำทอกา และครอบครัวในครอบครัวอื่น 1 พันคนในครอบครัว 100 คนในครอบครัวอื่น 300 คน ด้วย histogram ดังนี้

[ขอบเขตคะแนน]

1. ความกว้างของแต่ละชั้นคะแนน = ขอบบน - ขอบล่าง (จะเอาชั้นไหนก็ได้)

$$= 39.5 - 29.5 = 10$$

2. จุดกึ่งกลางชั้นคะแนน $60 - 69 = \frac{69 + 60}{2} = \frac{129}{2} = 64.5$

3. ส่วนใหญ่ ร้อยได้คะแนนในชั้นคะแนนที่ 3 [ขอบล่าง 49.5] [ขอบบน 59.5]

4. คนที่สอบได้คะแนน > 80 % มีจำนวน $15 + 5 = 20$ คน

↑
คะแนนในชั้นคะแนนที่ 6
← 7 →

5. คนที่สอบได้ไม่ถึง 50% มีจำนวนกี่คน → คนนี้ มีจัวไม่นับ 49.5 นาดับ

$$\approx 35 + 50 = 85 \text{ คน}$$

↑
คะแนนในชั้นคะแนนที่ 1
← 2 →

6. กำแพงที่รองรับมีคะแนน คือ คะแนนสอบ > 70 คะแนน 7. รับจิวเกิน = $35 + 15 + 5 = 55$ คน

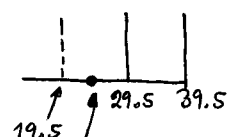
↑ ↑ ↑
ชั้นที่ 5 ชั้นที่ 6 ชั้นที่ 7

*** 7. เขียน histogram * ข้อต้องนิยามชัด มีนมีแถวความถี่อยู่ 7 แถว

- แต่ละแถวต้องมีจุดกึ่งกลางชั้นของตัวเอง นั่นคือ แถวต้องมาจุดกึ่งกลางนั้น
- แถวต้องมาจุด ขอบซ้ายสุด และ ขอบขวาสุด ใช้ histogram ด้วย โดย

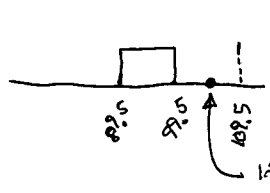
เส้นขอบซ้ายสุด ใช้ระดับความถี่ ในชั้นคะแนนแรก มี ขอบล่าง = 29.5 และ มีนมี ความกว้างชั้นคะแนนนั้น = $39.5 - 29.5 = 10$ ดังนั้น ก็ จะมีย่อตรงกลาง ชั้นนั้นแล้ว มีน มี ขอบขวาสุดนั้น

และ จุดกึ่งกลางชั้นของ 19.5 และ 29.5 คือ $\frac{19.5 + 29.5}{2} = 24.5$



*** เลข 24.5 นี้แปลแล้ว จะเป็นจุดกึ่งกลาง เส้นนับความถี่ของบรรทัดแรกและรวมของกรรมที่ ***

สี่เหลี่ยมผืนผ้าปกติ เส้นตั้ง เวลาตัวแปรสุ่ม อินเทอร์เน็ต สมมุติ ตัวนี้

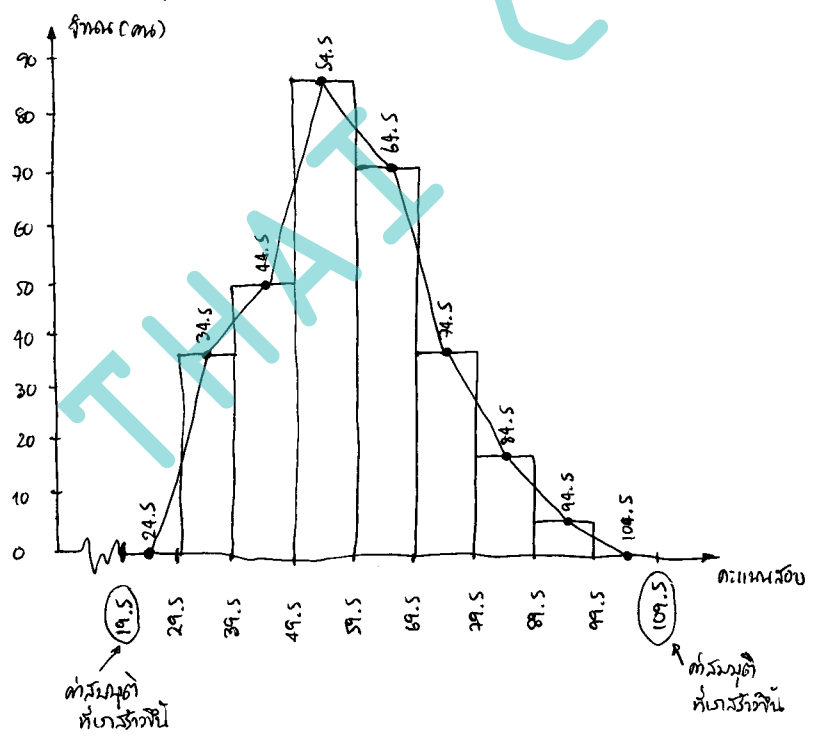


$$\begin{aligned} \text{แล้วหาจุดกึ่งกลางอินเทอร์เน็ตสมมุติ} &= \frac{109.5 + 99.5}{2} \\ &= \frac{209}{2} = 104.5 \end{aligned}$$

ที่นี้ เราก็หาจุดกึ่งกลางของแต่ละอินเทอร์เน็ตใน histogram ของเรากันด้วย

อินเทอร์เน็ตที่	จุดต่ำ	จุดบน	จุดกึ่งกลาง
สมมุติข้อแรก	19.5	29.5	$\frac{19.5 + 29.5}{2} = 24.5$
1	29.5	39.5	$\frac{29.5 + 39.5}{2} = 34.5$
2	39.5	49.5	$\frac{39.5 + 49.5}{2} = 44.5$
3	49.5	59.5	$\vdots = 54.5$
4	59.5	69.5	$\vdots = 64.5$
5	69.5	79.5	$\vdots = 74.5$
6	79.5	89.5	$\vdots = 84.5$
7	89.5	99.5	$\vdots = 94.5$
สมมุติข้อสุดท้าย	99.5	109.5	$\frac{99.5 + 109.5}{2} = 104.5$

ที่นี้ ก็เขียน รูปบางเส้นของกราฟก็ได้แล้วด้วย ตัวนี้



จำนี่ ไว้ด้วย !!

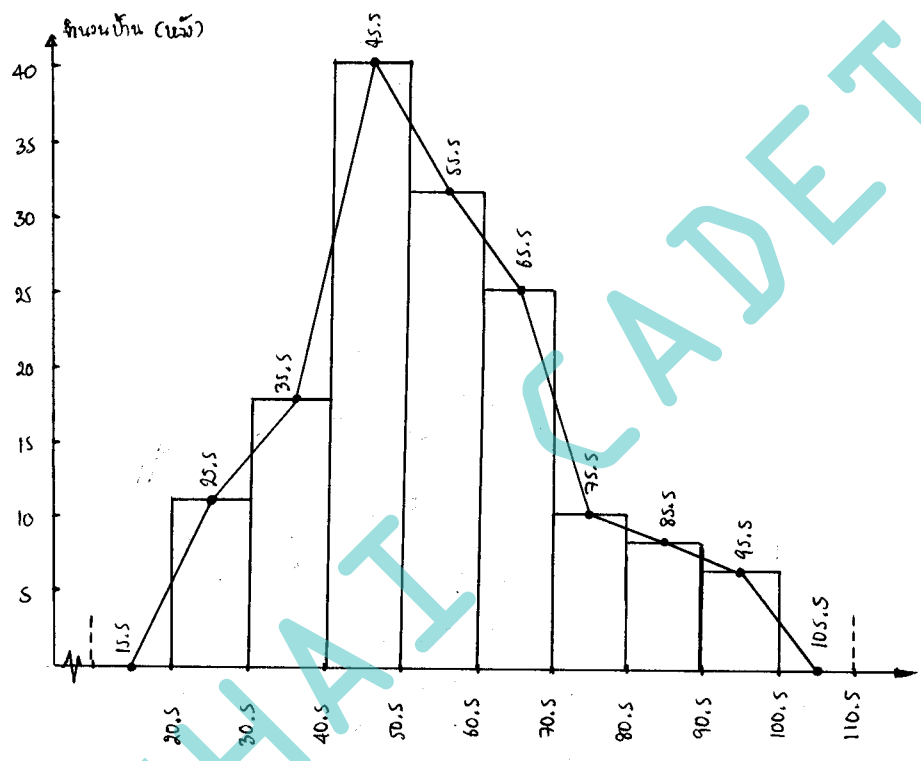
4. สืบหา histogram ไล่ตัวนี้

ปีนที (ขนาด)	จำนวน	ขอบล่าง	ขอบบน	จุดกึ่งกลางชั้น
21 - 30	11	20.5	30.5	$\frac{20.5 + 30.5}{2} = 25.5$
31 - 40	18	30.5	40.5	$\frac{30.5 + 40.5}{2} = 35.5$
41 - 50	40	40.5	50.5	$\vdots = 45.5$
51 - 60	32	50.5	60.5	$\vdots = 55.5$
61 - 70	25	60.5	70.5	$\vdots = 65.5$
71 - 80	10	70.5	80.5	$\vdots = 75.5$
81 - 90	8	80.5	90.5	$\vdots = 85.5$
91 - 100	6	90.5	100.5	$\frac{90.5 + 100.5}{2} = 95.5$

* ค่าเฉลี่ย จุดกึ่งกลางชั้น สมมุติตัวตั้งเป็น 15.5
ซึ่ง ค่าเป็น 15.5

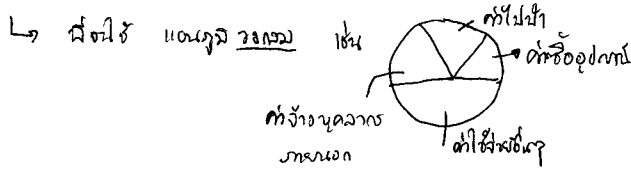
* และ จุดกึ่งกลางชั้น สมมุติตัวตั้งเป็น 105.5
ซึ่ง ค่าเป็น 105.5

รวม 150

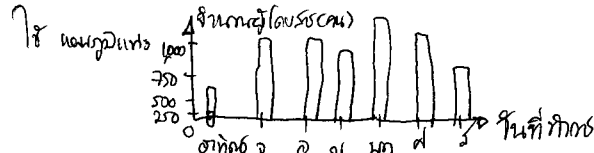


จงวิเคราะห์ข้อมูลต่อไปนี้ที่ ตราสินค้าของขนมปังโดยอิงตามกรรมวิธี

1. ข้อมูลรวบรวมค่าใช้จ่ายในครัวเรือน ผลการวัดของ 2 เดือน ตามปกติกฎระเบียบ



2. สรุปจำนวนผู้โดยสาร BTS ใน 1 วัน (หน่วย: คน)



เพื่อ Histogram ปี

3. ข้อมูลอัตราเงินเดือน ของข้าราชการ และเจ้าหน้าที่ใน อ. แนนนึ่ง
 (ตาราง) ตัวอย่างคนที่ ระดับข้าราชการ เป็นอย่างไร และ อัตรา: ระดับมีจำนวนเงินเดือนเท่าไร

ระดับข้าราชการ	อัตราเงินเดือน
ระดับ 1	7,500
ระดับ 2	9,500
ระดับ 3	10,500
ระดับ 4	12,000
ระดับ 5	15,000

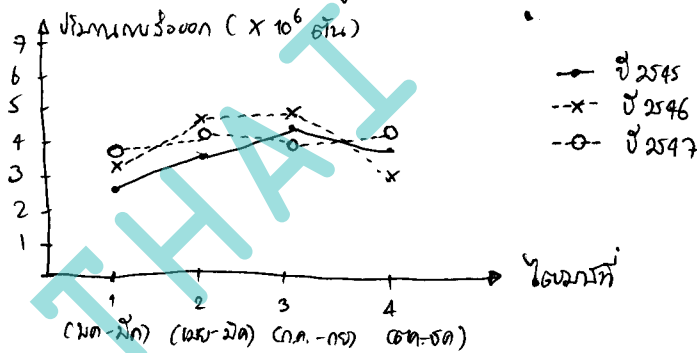
อัตราค่าจ้างนอกตัว แยกเป็นสองส่วน

4. ข้อมูลผลผลิตของข้าวโพด ของไรชาวนา ช่วง ปี 2545 - 2547

ปี 2545 1 ปี 2546 4 ไร่ชาวนา

และ ไร่ชาวนา มี 3 ปี คือ 2545, 2546 และ 2547

∴ ต้องการนำเสนอด้วยแผนภูมิเส้น



หมายเหตุ: ไร่ชาวนาที่ปลูกข้าวโพด

- ไร่ชาวนาของ ปี 45-47 รวม 3 ปี

ถ้าไร่ชาวนา มีอัตราการผลิตนี้ - ผลของไร่

แต่ไม่มีปัจจัยอื่น ที่ทำให้มีการเพิ่ม/ลด

ของอัตราการผลิตของไร่ชาวนาในปี 7 ไร่

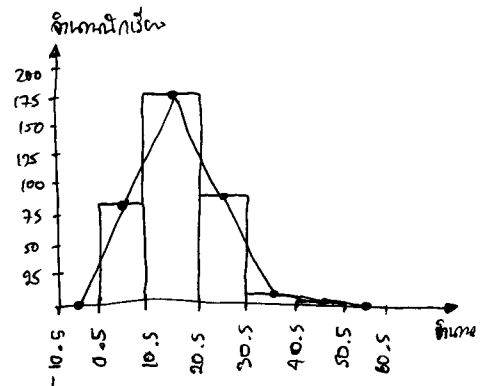
5. จำนวนเงินที่ 1 เดือนได้เป็นค่าจ้างพนักงาน 3 เดือน 350 คน.

จำนวนเงินที่ได้อีก / วัน (บาท)	จำนวนเงิน
1 - 10	82
11 - 20	175
21 - 30	84
31 - 40	5
41 - 50	4

รวม 350

นี่คือการนำเสนอข้อมูลในแผนภูมิ Histogram ได้โดย

$$\left\{ \begin{array}{l} \text{ขอบบน} \\ \text{ขอบล่าง} \\ \text{จุดกึ่งกลางชั้น} \end{array} \right. = \frac{\text{ขอบบน} + \text{ขอบล่าง}}{2}$$



3.2 ค่ากลางของข้อมูล

ค่าเฉลี่ยเลขคณิต (\bar{X}) = จำนวนที่ได้จากการผลบวกของข้อมูลทั้งหมดหารด้วยจำนวนข้อมูล

$$= \frac{\sum \text{ข้อมูลทั้งหมด}}{\text{จำนวนข้อมูล}}$$

เช่น บ้าน 3 หลัง มีรถจอด 10, 15, 20 คัน

น $\bar{X} = \frac{10+15+20}{3} = \frac{45}{3} = 15$
 3 → มี 3 ข้อมูล

∴ เฉลี่ยแล้วปลูกบ้านละ 15 คัน
 ↓
 รถจอดคัน

ข้อ 7 ข้อที่ 1, 2 และ 3 จงหาค่ากลางของ \bar{X} ลงท้ายด้วย

โดยเลข: ข้อ 2 และ 3 หาค่ากลางที่มาจากข้อมูลใช้ \bar{X} เป็นตัวตั้ง

และดูข้อ 3.2 ก

1. จงหาค่า \bar{X}

1) $\bar{X} = \frac{3+2+5+8+14+14+5+3+17}{9} = 7.888$
 9 → ข้อมูลมี 9 ตัว

2) $\bar{X} = \frac{2.8+2.7+5.7+2.1+3.3+2.8+2.8+3.2+2.1+5.1}{10} = \frac{32}{10} = 3.2$
 10 → ข้อมูลมี 10 ตัว

3) $\bar{X} = \frac{72+88+90+65+72+68}{6} = \frac{453}{6} = 75.5$

4) $\bar{X} = \frac{150+86+225+345+410+330+176}{7} = 246$

2. หาค่าข้อมูลทีไรทีในวงนั้น

1) เครื่องจักรเส้นทอขนาด 15 เครื่อง (มีข้อมูล 15 ตัว ตั้งแต่ 2, 5, 1, 2, ..., 38, 40)

2) จำนวนชั่วโมงน้อยที่สุด = 1 ชั่วโมง (ข้อที่ 3)

3) จำนวนชั่วโมงมากที่สุด = 40 ชั่วโมง (ข้อที่ 15)

4) $\bar{X} = \frac{2+5+1+2+14+10+11+18+14+28+26+23+31+38+40}{15} = \frac{263}{15} = 17.533$

** 3. ค่าเฉลี่ยคะแนนสอบ ของนักเรียนชาย 6 คน และนักเรียนหญิง 4 คน เป็น 51 คะแนน

จาก $\bar{X} = \frac{\text{คะแนนจาก นร.ชาย } 6 \text{ คน} + \text{คะแนนจาก นร.หญิง } 4 \text{ คน}}{10} = 51$

∴ 1) คะแนนรวมของคะแนนสอบในข้อนี้ = $\bar{X} \times 10$
 = $51 \times 10 = 510$ คะแนน

2) ถ้า \bar{X} ของนักเรียนชาย เป็น 49 จงหา \bar{X} ของนักเรียนหญิง

คือค่าที่ \bar{X} รวม = 51

∴ เป็นไปได้ที่ นร.ชายทุกคนได้ 51 คะแนน และ นร.หญิงทุกคนได้ 51 คะแนน

ที่ในนี้ $\bar{X} = \frac{(\bar{X}_{\text{นร.ชาย}})(6) + (\bar{X}_{\text{นร.หญิง}})(4)}{10}$

↖ เป็นเลข 10 เพราะมีคนทั้งหมด 10 คน

= $\frac{(51)(6) + (51)(4)}{10} = \frac{51(10)}{10} = 51$

∴ $\bar{X} = \frac{\bar{X}_{\text{ชาย}}(6) + \bar{X}_{\text{หญิง}}(4)}{10} = 51$

$(49)(6) + \bar{X}_{\text{หญิง}}(4) = 51(10) = 510$

∴ $\bar{X}_{\text{หญิง}} = \frac{510 - 294}{4} = \frac{216}{4} = 54$

∴ $\bar{X}_{\text{นักเรียนหญิง}} = 54 \quad \#$

** 4. ใน 6 วัน ส้มลงได้เงินได้รวม 120 บาท ในวันที่ 7 ส้มลงต้องได้เงินให้ได้เท่ากับ \bar{x} ซึ่งเท่ากับวันอื่น 2 บาท

$$\bar{x}_{6 \text{ วัน}} = \frac{120}{6} = 20$$

แสดงว่า หากส้มลงได้เงินวันละ 20 บาท ก็จะได้ $\bar{x}_{6 \text{ วัน}} = 20$ บาท

$$\text{ถ้าส้มลงได้ } \bar{x}_{7 \text{ วัน}} = 22$$

$$\therefore \bar{x}_{7 \text{ วัน}} = \frac{(\bar{x}_{6 \text{ วัน}})(6) + (\bar{x}_{\text{วันที่ 7}})(1)}{7} = 22$$

$$\therefore \bar{x}_{\text{วันที่ 7}} = \frac{(22)(7) - \bar{x}_{6 \text{ วัน}}(6)}{1}$$

$$= \frac{154 - 20(6)}{1} = \frac{154 - 120}{1} = 34$$

\therefore ในวันที่ 7 นี้ ส้มลงต้องได้เงิน 34 บาท

5. ผลการเรียน 5 ภาคเรียนของดิฉัน คือ

		ระดับผลการเรียน	จำนวนหน่วยการเรียน
ม.1	ภาค 1	3.20	15
	ภาค 2	4.00	15
ม.2	ภาค 1	3.50	14.5
	ภาค 2	3.00	16.5
ม.3	ภาค 1	3.65	16

$$\begin{aligned} \text{จงหา ระดับผลการเรียนเฉลี่ย ทั้ง 5 ภาคการเรียน} &= \frac{(3.2)(15) + (4.00)(15) + (3.5)(14.5) + (3.00)(16.5) + (3.65)(16)}{15 + 15 + 14.5 + 16.5 + 16} \\ &= \frac{266.65}{97} = 3.46298 \end{aligned}$$

* ทราบว่า ผู้รู้ได้ตัวอย่างไร ที่ต้องคิดแบบนี้ ที่ไม่ใช้ผลบวกกันของ grade แล้วหาจำนวนตอนมาหาร

$$\text{กล่าวคือ } \bar{x} = \frac{3.2 + 4 + 3.5 + 3.0 + 3.65}{5} = 3.47$$

คือ 3.47 ก็ใกล้เคียงกันได้เยอะมากกับ 3.46298

* แต่เวลาสอบคิดดูด้วยว่า

$$\text{เกรดเฉลี่ยแต่ละตอน} = \frac{\text{กำหนดอันดับของทั้งหมด} \times \text{ผลลัพธ์} \Delta}{\text{หน่วยการเรียนรวม}}$$

$$\text{เกรดเฉลี่ยแต่ละตอน} = \frac{\Delta}{\text{หน่วยการเรียน}} \quad \therefore \Delta = (\text{เกรดเฉลี่ยแต่ละตอน}) \times (\text{หน่วยการเรียน})$$

$$\text{และที่นี้ } \Delta_1 = (3.2)(15) \quad \Delta_4 = (3.0)(16.5)$$

$$\Delta_2 = (4.0)(15) \quad \Delta_5 = (3.65)(16)$$

$$\Delta_3 = (3.5)(14.5)$$

$$\text{และ } \bar{\Delta} = \frac{\Delta_1 + \Delta_2 + \Delta_3 + \Delta_4 + \Delta_5}{15 + 15 + 14.5 + 16.5 + 16} = \frac{266.65}{97} = 3.46298$$

(11คองดู)

กรณีที่ไม่ได้ข้อมูล จำนวนครั้งที่ขึ้นแต่ละขั้นในกราฟของ
รูปที่ 2 → 10

∴ เรายก \bar{x} ไม่ได้ เพราะไม่มีข้อมูลในครั้งที่ 2 → 10

P. 117 มัชฐาน (Median)

ขั้นแรกจัดเรียงข้อมูลจาก น้อยไปมาก
หรือจาก มากไปน้อย

แล้ว ข้อมูลที่อยู่ตรงกลาง คือค่า มัชฐาน นั้นเอง

▷ ถ้า กลุ่มข้อมูลเป็นเลขคู่ เช่น 1, 1, 2, 1, 5, 1, 2, 3, 7

จัดเรียงข้อมูลจากน้อย → มาก ได้ 1, 1, 1, 1, 2, 2, 3, 5, 7

↑
เลขที่อยู่ตรงกลาง

อย่างนี้ เราจะพิจารณาไม่ได้ ไม่ชากัน

▷ แต่ ถ้ากลุ่มข้อมูลเป็นเลขคู่ เช่น 1, 1, 2, 5, 1, 3, 3, 7

เรียงข้อมูลจากน้อย → มาก ได้ 1, 1, 1, 2, 3, 3, 5, 7

จะเห็นว่า ค่ากลางคือเลข 2 และเลข 3

แล้วจะทำได้ ? เทำไป ถ้าค่ากลางคือ 2 ค่ากลางอีกอัน แล้วหาค่าเฉลี่ย 2

$$\text{กล่าวคือ มัชฐาน} = \frac{2+3}{2} = \frac{5}{2} = 2.5$$

▷ สรุปว่า มัชฐาน คือ ค่ากลางของข้อมูล จากกลุ่มข้อมูลที่ถูกเรียงจากน้อยไปมาก

หรือ ถูกเรียงจาก มากไปน้อย แล้ว จำนวนข้อมูลที่น้อยกว่าค่านั้น = จำนวนข้อมูลที่มากกว่าค่านั้น

ถ้าชุดข้อมูลเป็นเลขคู่ มัชฐานจะเป็นเลขเต็มตัวเดียว อยู่ตรงกลางพอดี

ถ้าชุดข้อมูลเป็นเลขคี่ มัชฐานจะเกิดจากเลข 2 ตัวตรงกึ่งกลาง โดยที่ $\text{มัชฐาน} = \frac{\text{เลขตัวแรก} + \text{เลขตัวที่สอง}}{2}$

P. 119 จักรนิยม (mode) คือ ข้อมูลที่มีค่ามากที่สุดหรือค่าสูงสุดในข้อมูลชุดนั้น

Ex1 ชุดข้อมูล 4, 3, 3, 7, 5, 8, 7, 9, 7, 3, 2, 2, 7

เรียงข้อมูลจากน้อยไปมาก ได้ 2, 2, 3, 3, 3, 4, 5, 7, 7, 7, 7, 8, 9
↑ เลข 5 คือ มัชฐาน
↑ จำนวนนิยมคือ 7

$$\text{และมี } \bar{x} = \frac{2+2+3+\dots+8+9}{13} = \frac{67}{13} = 5.15$$

Ex2 ชุดข้อมูล 1, 1, 2, 2, 3, 2, 1

เรียงข้อมูลจากน้อยไปมาก ได้ 1, 1, 1, 2, 2, 2, 3
มี 3 ตัว มี 3 ตัว
∴ จักรนิยม คือเลข 1 และเลข 2

Ex3 ชุดข้อมูล 2, 1, 2, 2, 1, 1

เรียงข้อมูลจากน้อยไปมาก ได้ 1, 1, 1, 2, 2, 2
มี 3 ตัว มี 3 ตัว
∴ ข้อมูลชุดนี้ ไม่มีจกรนิยม → เพราะค่าของข้อมูลมีค่ากันหมด
ยกเว้น เลข 1 และเลข 2

1. หากมีข้อมูล 100 ชุด ข้อมูลต่อไปนี้ (ห้อยเรียงจากน้อย → มาก โดยค้น)
 น้อยต่อไปห้อยเรียงเอาเองนะค้น

1) 15, 17, 17, 18, 25, 29, 37, 49 และ 62

↑ มีศูนย์กลางที่ตัวเลข 25

2) 0.8, 0.8, 4.3, 5.1, 6.5, 7.2, 10.2, 11.3

$$\text{มีศูนย์กลาง} = \frac{5.1 + 6.5}{2} = \frac{11.6}{2} = 5.8$$

3) 48, 56, 58, 72, 72, 72, 90

↑ มีศูนย์กลางที่ตัวเลข 72

4) 10, 11, 12, 12, 12, 12, 15, 16, 20, 20

$$\text{มีศูนย์กลาง} = \frac{12 + 12}{2} = \frac{24}{2} = 12$$

2. จงหาฐานนิยมของข้อมูลแต่ละชุดต่อไปนี้

1) 4, 4, 5, 7, 7, 7, 8, 8, 10, 11

ฐานนิยม คือ เลข 7

2) 35.2, 37.3, 38.5, 38.7, 39.3, 40.1, 41.4, 43.9

ไม่มีเลขใดซ้ำกัน ∴ ไม่มีฐานนิยม

3) 11, 11, 11, 11, 15, 15, 15, 17, 18, 18, 21

มี 4 จำนวน, มี 3 จำนวน, มี 2 จำนวน

∴ ฐานนิยม คือ เลข 11

3. จากข้อมูลต่อไปนี้

1) คนส่วนใหญ่ซื้อเสื้อไซส์ 12 ประเภทได้คือ 10 ตัว

(ในคนที่ซื้อไซส์ 15 ขายได้แค่ 3 ตัว)

2) เฉลี่ยแล้วแต่ละวันคนซื้อได้กี่ตัว $= \frac{7+5+6+10+8+9+3}{9} = 6.57142$ ตัว

≈ 6 ตัวครึ่ง → เมื่อพิจารณาจากที่พบเมื่อได้ 2 วัน 13 ตัว #

4.
$$\bar{x} = \frac{2(4) + 3(6) + 4(6) + 5(12) + 6(13) + 7(20) + 8(16) + 9(10) + 10(6) + 11(4) + 12(3)}{4 + 6 + 6 + 12 + 13 + 20 + 16 + 10 + 6 + 4 + 3} = \frac{686}{100} = 6.86$$

มีศูนย์กลาง ให้พิจารณาคนที่ซื้อเสื้อรวม 100 ตัว หากเรียงข้อมูลจากน้อยไปมากแล้ว จำนวนลำดับที่ 50 คือเลข 7

7 หรือ 8 คือเลข 7 เช่นกัน

∴ มีศูนย์กลาง = $\frac{7+7}{2} = \frac{14}{2} = 7$

ฐานนิยม คือคนที่ซื้อเสื้อที่สูงสุด = 11 ตัว # มีคนซื้อเสื้อเข้ารวม 20 ตัว #

5. จงหาตัวกลางโดยที่แบบผสมกับข้อมูลในชุดต่อไปนี้ (ห้อยต่อเรียงจากน้อย → มาก ค้นนะค้น)

1) 65, 70, 72, 73, 75, 76, 77, 79, 81 และ 8923

ใช้ \bar{x} ไม่ได้ เพราะเลข 8923 แยกออกจากเลขกลุ่ม 65 → 81 มากๆ

ใช้ \bar{x} มีศูนย์กลาง → ใช้เลข $\frac{75+76}{2} = 75.5$

2) 2, 2, 2, 2, 2, 2, 16, 18, 21 และ 28

ใช้ฐานนิยม = เลข 2

3) 2, 3, 4, 5, 6, 7, 18, 18, 18 และ 200

ใช้ \bar{x} ไม่ได้ เพราะเลข 200 แยกจากเลขหลักอื่นๆ และหลักสิบ จึงมีค่าไม่เกิน 20

∴ ค่าใช้ มีศูนย์กลาง คือ $\frac{6+7}{2} = 6.5$

ทำให้ฐานนิยม คือเลข 18 (มี 3 จำนวน)

4) 1, 2, 3, 4, 5, 16, 19, 28, 29 และ 30

ฐานนิยม = ไม่มี

มีศูนย์กลาง = $\frac{5+16}{2} = \frac{21}{2} = 10.5$

และ $\bar{x} = \frac{1+2+3+4+5+16+19+28+29+30}{10}$

= 13.7

ห้อยต่อ \bar{x} เมื่อ มีศูนย์กลางก็ได้ ตามที่คำนวณแล้ว

ภยได้/เดือน	120,000	50,000	15,000	12,000	10,000	8,500
จำนวนนักงาน (คน)	1	3	10	12	16	8

$$1) \text{ นก } \bar{x} = \frac{120,000(1) + 50,000(3) + 15,000(10) + 12,000(12) + 10,000(16) + 8,500(8)}{1 + 3 + 10 + 12 + 16 + 8}$$

$$= \frac{792,000}{50} = 15,840$$

มีฐาน = 12,000 เมฆ; จากนักงาน 50 คน ถือว่าเป็นกลุ่ม 50 คน
 เมื่อเริ่มข้อมูลจากน้อย → มาก ; ข้อมูลที่ 25 คือเลขเงินเดือน 12,000
 → 26 ก็คือเลขเงินเดือน 12,000 เช่นกัน
 \therefore มีฐาน = $\frac{12,000 + 12,000}{2} = 12,000$
 ฐานใหม่ = จำนวนที่มีค่ามากที่สุด = เงินเดือน 10,000 บาท

2) กำมการจาเมื่อปรับเงินเดือนขึ้นให้นักงาน โดยมีผู้เกี่ยวข้อง 3 ฝ่าย คือ

- เจ้าของบริษัท
- ต้อแทนนักงาน และ
- คนกลาง ผู้ไกล่เกลี่ย

ผู้คิดว่ เตต้องเลือก \bar{x} เมฆะ สำปรับเงินเดือนน้อยๆ ก็เลือก มีฐาน บางได้ แต่เลข 12,000 บาท
 ถ้าเลือก ฐานใหม่ บางได้แต่เลข 10,000 บาท

แต่ถ้าเลือก \bar{x} บางได้ 15,840

และไม่ต้องห่วงผู้บริหร หรือ ผู้ได้เงินเดือนเยอะ เมฆะ เขาไม่ถูกลดเงินเดือนหรอกครับ
 ยกเว้น บริษัทเกิดสวทះ ยี่สิบสอง ทางกรมมือ

หน้า 123

ตัดสินอย่างไร?

งานที่นำส้ มีทงข้อง 100 เมตร ถือเป็น ทรเชิงเมื่อเน้นเขาเจ้ากรรมวิ โดยท้งขี้

แต่ละคนส้ มีนักว้ 3 คน

$$\therefore \text{ ส้คนส้ } = 3 \times 4 = 12 \text{ คน}$$

- ผลการท้งวันคือ
- | | | | |
|-------|-------------|------|----------------------------|
| ที่ 1 | : ส้เข้ขวมี | 1 คน | ใช้ทล 13 วันท้ |
| ที่ 2 | : ส้ม่วง | 2 คน | |
| | : ส้เน้ง | 2 คน | รวม 4 คน
ใช้ทล 14 วันท้ |
| ที่ 3 | : ส้เข้ช | 1 คน | |
| | : ส้ม่วง | 1 คน | รวม 4 คน
ใช้ทล 15 วันท้ |
| | : ส้เน้ง | 2 คน | |
| ที่ 4 | : ส้เน้ง | 1 คน | รวม 3 คน
ใช้ทล 16 วันท้ |
| | : ส้เข้ช | 1 คน | |
| | : ส้เน้ง | 1 คน | |

ถ้าตัดสินท้ง-ชนะ คือเป็นทีม โดยดูท้ ท้มใดใช้ทลว้เร็วเดส้ น้สท้

ใ้พิจารณาด้วยส้ ; ส้เน้ง = $\frac{2(15) + 1(16)}{3} = \frac{30+16}{3} = 15.33$ วันท้

ส้เข้ช = $\frac{1(13) + 1(15) + 1(16)}{3} = \frac{44}{3} = 14.66$ วันท้

ส้เน้ง = $\frac{2(14) + 1(16)}{3} = \frac{44}{3} = 14.66$ วันท้

* ส้ม่วง = $\frac{2(14) + 1(15)}{3} = \frac{43}{3} = 14.33$ วันท้

\therefore ถ้าตัดสินท้งคน นักว้สีเข้ช ถ้องได้แม้งทล (เมฆคนส้ช)

* แต่ ถ้าตัดสินท้ง นั้น ทีมส้ม่วง ชนะด้วย เมฆ: ใช้ทลเร็วเดส้ น้สท้เร็ว ท้กับ 14.33 วันท้

ตอบ

ตัวอย่างหน้า 124 เป็นตัวอย่างที่ดี ของกลุ่มข้อมูล 2 กลุ่ม ซึ่ง

ข้อมูลทั้งสองกลุ่ม มี \bar{x} เท่ากัน
 มีมัธยฐาน เท่ากัน
 มีฐานนิยม เท่ากัน

แต่ สิ่งแตกต่าง กลุ่มข้อมูล A นั้นมีค่ากลุ่มกัน

กลุ่มข้อมูล B มีการกระจายตัวของข้อมูลมากกว่าของกลุ่ม A

* องค์ประกอบที่ทางสถิติ มักจะใช้กัน คือ เฉลี่ย และ ส่วนเบี่ยงเบนมาตรฐาน

เมื่อเฉลี่ย = คะแนนสูงสุด - คะแนนต่ำสุด

** \therefore เฉลี่ย จะเป็น ตัวบอก มรกระจายตัว ของ ข้อมูล นั้นเอง *

และเราจะหา ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)

1. หา \bar{x}
2. หา ส่วนเบี่ยงเบน เมื่อผลต่างระหว่างแต่ละคะแนน กับ \bar{x} โดยนำแต่ละคะแนน - ค่าเฉลี่ยเลขคณิต
3. หาค่าเฉลี่ยของ ส่วนเบี่ยงเบนแต่ละค่า ที่ได้จากข้อ 2.
4. หา ค่าเฉลี่ยเลขคณิต ของค่าเฉลี่ยของส่วนเบี่ยงเบนที่ได้จากข้อ 3.
5. หาากที่ล่อง ซึ่งเป็นบวก ของค่าเฉลี่ยเลขคณิต ที่ได้ในข้อ 4.

ซึ่ง ผลลัพธ์ที่ได้ในข้อ 5. จะเป็นส่วนเบี่ยงเบนมาตรฐาน ที่มีหน่วยเดียวกับหน่วยของข้อมูล

หน้า 125 หาส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) ของนักเรียนกลุ่ม A

1. $\bar{x} = \frac{28 + 29 + 30 + 32 + 34 + 36 + 37 + 38 + 38 + 38}{10} = \frac{340}{10} = 34$

2. หาส่วนเบี่ยงเบน

คะแนน	\bar{x}	ส่วนเบี่ยงเบน	(ส่วนเบี่ยงเบน) ²
28	34	28 - 34 = -6	(-6) ² = 36
29	34	29 - 34 = -5	(-5) ² = 25
30	34	30 - 34 = -4	(-4) ² = 16
32	34	32 - 34 = -2	(-2) ² = 4
34	34	34 - 34 = 0	0 ² = 0
36	34	36 - 34 = 2	2 ² = 4
37	34	37 - 34 = 3	3 ² = 9
38	34	38 - 34 = 4	4 ² = 16
38	34	38 - 34 = 4	4 ² = 16
38	34	38 - 34 = 4	4 ² = 16

3. หาค่าเฉลี่ยของ ส่วนเบี่ยงเบน

4. หา ค่าเฉลี่ยเลขคณิต จากข้อ 3. ; $\frac{36 + 25 + 16 + 4 + 0 + 4 + 9 + 16 + 16 + 16}{10} = 14.2$

5. จากที่ล่องที่เป็นบวก ของค่าเฉลี่ยเลขคณิตที่ได้จากข้อ 4.

คือ $\sqrt{14.2} \approx 3.8$

ดังนั้น ส่วนเบี่ยงเบนมาตรฐานของคะแนนกลุ่ม A มีค่าประมาณ 3.8 คะแนน

ภาคน	ส่วนเบี่ยงเบน	(ส่วนเบี่ยงเบน) ²
8	8-34 = -26	(-26) ² = 676
16	16-34 = -18	(-18) ² = 324
27	27-34 = -7	(-7) ² = 49
33	33-34 = -1	(-1) ² = 1
34	34-34 = 0	0 ² = 0
36	36-34 = 2	2 ² = 4
38	38-34 = 4	4 ² = 16
38	38-34 = 4	4 ² = 16
50	50-34 = 16	16 ² = 256
60	60-34 = 26	26 ² = 676

$$\text{ได้ } \bar{x} = \frac{676+324+49+1+0+4+16+16+256+676}{10} = 201.8$$

$$\text{แล้ว } \sqrt{201.8} \sim 14.2056$$

$$\sim 14.21$$

เมื่อได้ค่าเฉลี่ยแล้ว นี้อ่านค่าอธิบาย จากหน้า 127 จะพบว่า "การกระจายจากค่าเฉลี่ยเลขคณิตของกลุ่ม B มีมากกว่าค่าเบี่ยงเบนค่าเฉลี่ยของกลุ่ม A"

และจะเห็นว่า แล้วกลุ่ม B จะมีนิสัย = 60-8 = 52 ซึ่งเป็นช่วงที่กว้าง

แต่จะเห็นว่า ส่วนเบี่ยงเบนมาตรฐานได้ ~ 14.21 ซึ่งเป็นค่าที่มากกว่า เพราะได้ใช้ทุกคะแนนของกลุ่ม B มาเกี่ยวข้องกับความสัมพันธ์

หน้า 128) ตัวนี้ มาที่แบบฝึกข้อต่อไป

1. เขาค้นหาค่า (Range) และ ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) ของข้อมูลต่อไปนี้

1) 0, 3, 5, 6, 11 มีนิสัย = 11-0 = 11 ; มี $\bar{x} = \frac{0+3+5+6+11}{5} = \frac{25}{5} = 4.166$

เรียงข้อมูล	ส่วนเบี่ยงเบน	(ส่วนเบี่ยงเบน) ²
0	0-4.166 = -4.166	17.3555
3	3-4.166 = -1.166	1.3595
5	5-4.166 = 0.834	0.6955
6	6-4.166 = 1.834	3.3635
11	11-4.166 = 6.834	46.7015

รวม 187.4855

$$\bar{x} = \frac{187.4855}{5} = 37.4971$$

$$\text{แล้ว } \sqrt{\bar{x}} = \sqrt{37.4971} \approx 6.1234875$$

$$\approx 6.1235$$

สังเกตว่า นิสัย = 4.166

ส่วนเบี่ยงเบนมาตรฐาน ≈ 6.1235

2) 6, 6, 8, 8, 12, 14 มี $\bar{x} = \frac{6+6+8+8+12+14}{6} = 9$ และ นิสัย = 14-6 = 8

ข้อมูล	ส่วนเบี่ยงเบน	(ส่วนเบี่ยงเบน) ²
6	6-9 = -3	(-3) ² = 9
6	6-9 = -3	(-3) ² = 9
8	8-9 = -1	(-1) ² = 1
8	8-9 = -1	(-1) ² = 1
12	12-9 = 3	3 ² = 9
14	14-9 = 5	5 ² = 25

รวม 54

$$\bar{x} = \frac{54}{6} = 9$$

$$\text{แล้ว } \sqrt{\bar{x}} = \sqrt{9} = 3$$

ตัวนี้ ส่วนเบี่ยงเบนมาตรฐาน = 3

นิสัย = 8

(ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)

มีน้อย แสดงว่า การกระจายกลุ่มข้อมูลก็มีน้อยเช่นกัน)

3) กลุ่มข้อมูล 9, 9, 9, 9, 9, 9

จะเห็นว่า $\bar{x} = 9$ แน่นอน

กล่าวคือ $\bar{x} = \frac{9+9+9+9+9+9}{6} = 9$

และ นิสัย = $9 - 9 = 0$

* แสดงว่า ไม่มีสมการกระจายตัวข้อมูลเลย

ข้อมูล	ส่วนเบี่ยงเบน	(ส่วนเบี่ยงเบน) ²
9	$9 - 9 = 0$	$0^2 = 0$
9	$9 - 9 = 0$	$0^2 = 0$
9	$9 - 9 = 0$	$0^2 = 0$
9	$9 - 9 = 0$	$0^2 = 0$
9	$9 - 9 = 0$	$0^2 = 0$
9	$9 - 9 = 0$	$0^2 = 0$

รวม 0

$\therefore \bar{x} = \frac{0}{6} = 0$

แล้ว $\sqrt{\bar{x}} = \sqrt{0} = 0$

แสดงให้เห็นว่า นิสัย = 0

ส่วนเบี่ยงเบนมาตรฐาน คือ = 0 เช่นกัน

4) กลุ่มข้อมูล 0, 2, 3, 3, 4, 4, 6, 7, 9, 12

มี $\bar{x} = \frac{0+2+3+3+4+4+6+7+9+12}{10} = \frac{50}{10} = 5$

นิสัย = $12 - 0 = 12 \rightarrow$ คือเป็นช่วงที่กว้างพอสมควร

ข้อมูล	ส่วนเบี่ยงเบน	(ส่วนเบี่ยงเบน) ²
0	$0 - 5 = -5$	$(-5)^2 = 25$
2	$2 - 5 = -3$	$(-3)^2 = 9$
3	$3 - 5 = -2$	$(-2)^2 = 4$
3	$3 - 5 = -2$	$(-2)^2 = 4$
4	$4 - 5 = -1$	$(-1)^2 = 1$
4	$4 - 5 = -1$	$(-1)^2 = 1$
6	$6 - 5 = 1$	$1^2 = 1$
7	$7 - 5 = 2$	$2^2 = 4$
9	$9 - 5 = 4$	$4^2 = 16$
12	$12 - 5 = 7$	$7^2 = 49$

รวม 114

$\bar{x} = \frac{114}{10} = 11.4$

แล้ว $\sqrt{\bar{x}} = \sqrt{11.4} \approx 3.3764$

นิสัย (Range) = 5

ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) = 3.3764

6. กลุ่มข้อมูล 2, 3, 3, 4, 4, 4, 5, 7, 9, 9

มี $\bar{x} = \frac{2+3+3+4+4+4+5+7+9+9}{10} = \frac{50}{10} = 5$

นิสัย = $9 - 2 = 7$

ข้อมูล	ส่วนเบี่ยงเบน	(ส่วนเบี่ยงเบน) ²
2	$2 - 5 = -3$	$(-3)^2 = 9$
3	$3 - 5 = -2$	$(-2)^2 = 4$
3	$3 - 5 = -2$	$(-2)^2 = 4$
4	$4 - 5 = -1$	$(-1)^2 = 1$
4	$4 - 5 = -1$	$(-1)^2 = 1$
4	$4 - 5 = -1$	$(-1)^2 = 1$
5	$5 - 5 = 0$	$0^2 = 0$
7	$7 - 5 = 2$	$2^2 = 4$
9	$9 - 5 = 4$	$4^2 = 16$
9	$9 - 5 = 4$	$4^2 = 16$

รวม 56

$\bar{x} = \frac{56}{10} = 5.6$

แล้ว $\sqrt{\bar{x}} = \sqrt{5.6} = 2.3664$

\therefore นิสัย = 7

ส่วนเบี่ยงเบนมาตรฐาน = 2.3664

2. จานวน \bar{x} , มัชฌิม (median), ฐานที่สาม (mode), นิสัย (Range) และส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) ของข้อมูลต่อไปนี้

คะแนน	ความถี่	ส่วนเบี่ยงเบน	(ส่วนเบี่ยงเบน) ²
70	2	70 - 85.5 = -15.5	(-15.5) ² = 240.25
75	1	75 - 85.5 = -10.5	(-10.5) ² = 110.25
80	4	80 - 85.5 = -5.5	(-5.5) ² = 30.25
85	3	85 - 85.5 = -0.5	(-0.5) ² = 0.25
90	6	90 - 85.5 = 4.5	(4.5) ² = 20.25
95	4	95 - 85.5 = 9.5	(9.5) ² = 90.25
		รวม	491.5

$$\sum \bar{x} = \frac{491.5}{20} = 24.575$$

$$\sqrt{\bar{x}} = \sqrt{24.575} = 4.9575$$

$$\bar{x} = \frac{70(2) + 75(1) + 80(4) + 85(3) + 90(6) + 95(4)}{20} = \frac{1,710}{20} = 85.5$$

มัชฌิม คือ เมื่อเรียงลำดับ คะแนนจากน้อยไปมากแล้ว

จะได้ 70, 70, 75, 80, 80, 80, 80, 85, 85, 85, 90, 90, 90, 90, 90, 90, 95, 95, 95, 95

$$\text{มัชฌิม} = \frac{85 + 90}{2} = \frac{175}{2} = 87.5$$

ฐานที่สาม = เลขคะแนน (90)

นิสัย = คะแนนมากที่สุด - คะแนนน้อยสุด

$$= 95 - 70 = 25$$

และ ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) = 4.9575

ตอบ

หน้า 129

ข้อแล้วสั้นๆ (ขี้ขี้เนี่ยะ?)

1. ค่า Standard Deviation (ส่วนเบี่ยงเบนมาตรฐาน) เป็นจำนวนลบได้หรือไม่?

คำตอบคือ ไม่ได้ เพราะ Standard Deviation คือ รากที่สอง ที่บวกของ \bar{x} ที่คิดได้

2. Standard Deviation เป็น 0 ได้มั๊ย. ตอบ: ค่า $\sqrt{\bar{x}} = 0$

หรือว่า $\bar{x} = 0$

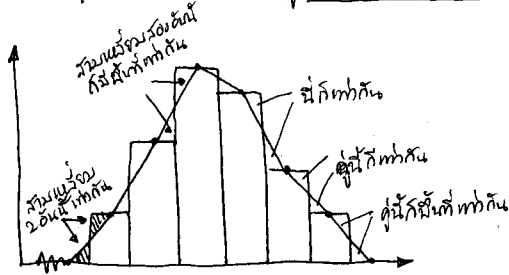
นั่นหมายความว่า

$$\frac{\sum (\text{ส่วนเบี่ยงเบน})^2}{\text{ความถี่ทั้งหมด}} = 0$$

หรือ ผลบวกของ (ส่วนเบี่ยงเบน)² ทั้งหมด = 0

แสดงว่า ข้อมูลไม่มีค่าเบี่ยงเบนเลย ***

สังเกตว่า เมื่อเราหา histogram จากได้ เราจะเห็น "รูปแปลงแล้วของความถี่" ได้ด้วย



สังเกตว่า พื้นที่รูปแปลงแล้วของความถี่ = ความสูงของพื้นที่ของรูปสี่เหลี่ยมมุมฉากทุกรูปใน histogram

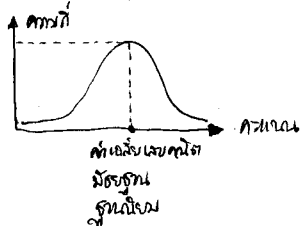
ถ้า จำนวนข้อมูลมากพอ แล้วถ้าจำนวนตัวอย่างน้อยลงมากพอ แล้ว รูปแปลงแล้วของความถี่ จะกลายเป็นลักษณะของเส้นโค้ง ดังนั้น ทฤษฎีที่จะมีพบบน รูปแปลงแล้วของความถี่ ได้เช่นเส้นโค้งปกติ โดย จุดที่ได้ได้เส้นโค้งจะมีค่าเท่ากับ รูปแปลงแล้วของความถี่เป็นเซต

ข้อ เส้นโค้งแบบนี้ ถูกเรียกว่า "เส้นโค้งของความถี่" (frequency Curve)

แบ่งเป็น 3 ลักษณะ คือ

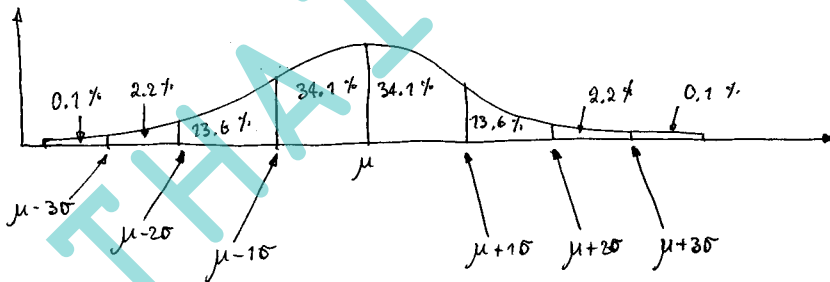
1. โค้งลาดเอียงซ้าย
 2. โค้งลาดเอียงขวา
 3. โค้งปกติ รูปประจักษ์
- ดูรายละเอียด หน้า 132-133

โดยทั่วไป เส้นโค้งปกติที่ดูประจักษ์ เป็นข้อมูลที่ไม่มีค่าแจกแจงปกติ จะมี \bar{X} , มัชฌิม และฐานนิยม เป็น ค่าเดียวกัน



Keywords ในเรื่องนี้คือ เส้นโค้งปกติ จะเกี่ยวข้องกับค่าเฉลี่ยเลขคณิต (\bar{X}) และ ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) ด้วย

ถ้าเราใช้ μ แทนค่า \bar{X}
 σ แทนค่า ส่วนเบี่ยงเบนมาตรฐาน

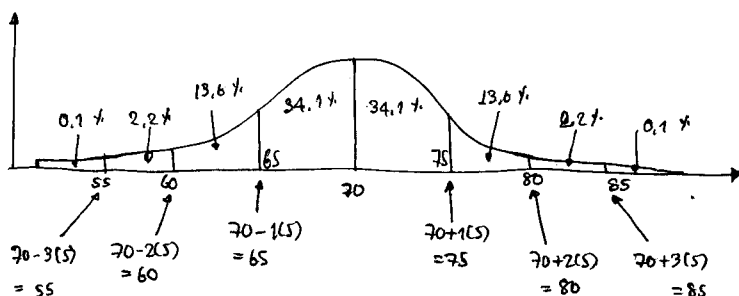


การสรุป พื้นที่ใต้เส้นโค้งปกติ ระหว่าง $\mu-1\sigma$ และ $\mu+1\sigma$ เท่ากับ $2 \times 34.1 = 68.2\%$
 พื้นที่ใต้เส้นโค้งปกติ ระหว่าง $\mu-2\sigma$ และ $\mu+2\sigma$ เท่ากับ $2 \times (34.1 + 13.6) = 95.4\%$
 พื้นที่ใต้เส้นโค้งปกติ ระหว่าง $\mu-3\sigma$ และ $\mu+3\sigma$ เท่ากับ $2 \times (34.1 + 13.6 + 2.2) = 99.8\%$

Ex หน้า 135 ; $\bar{X} = 70$ คะแนน

ส่วนเบี่ยงเบนมาตรฐาน = 5 คะแนน

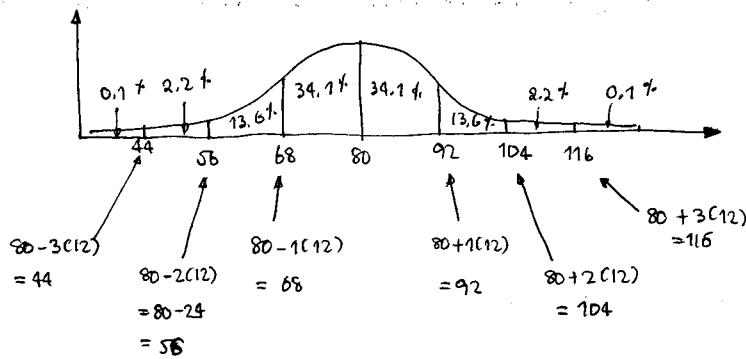
เราใช้ส่วนเบี่ยงเบนมาตรฐาน = 5 แบ่งออกเป็น เริ่มจากคะแนน 70 ไปทางขวา และ ซ้าย จะได้พื้นที่ใต้เส้นโค้งปกติ ดังนี้



ค่าต่างๆ ส่วนมากจะได้อัตโนมัติกับเส้นโค้งปกตินี้
 เช่น พื้นที่ระหว่างคะแนน 65 และ 80
 เท่ากับ $34.1 + 34.1 + 13.6 = 81.8\%$ พื้นที่

1. $\bar{x} = 80$ hours

Standard Deviation = 12 hours.



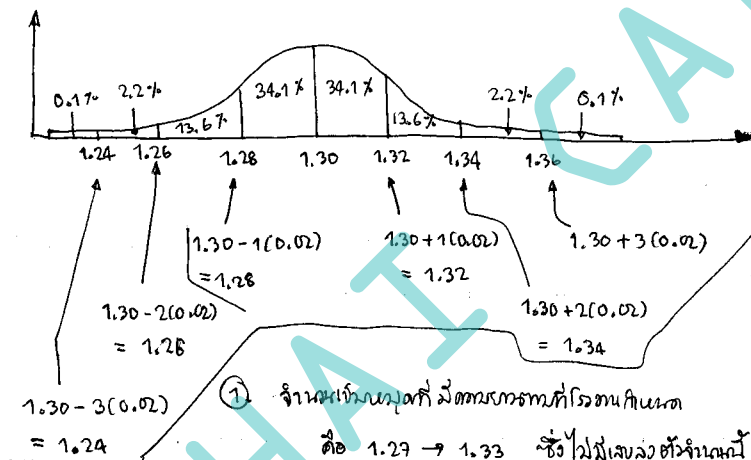
1) พื้นที่ใต้กราฟ ช่วง 68 ชม. ถึง 92 ชม. = = 34.1 + 34.1 = 68.2 %

2) พื้นที่ใต้กราฟของข้อมูลที่มีอยู่ 56 % = = 0.1 + 2.2 = 2.3 %

2. การควบคุมคุณภาพคือข้อบกพร่อง คือ 1.27" ถึง 1.33" จำนวนการผลิต 5,000 ชิ้น

มี $\bar{x} = 1.30$ " (อยู่ในเกณฑ์ที่รับใช้ได้) $\Rightarrow 1.27 \leq 1.30 \leq 1.33$

และ Standard Deviation = 0.02"



① จำนวนชิ้นงานที่มีคุณภาพคือมีข้อบกพร่องที่รับใช้ได้
คือ 1.27 \rightarrow 1.33 ซึ่งไม่มีตัวเลขตัวจำนวนนี้ (มีตัวเลข 1.26 ถึง 1.28 และเลข 1.32 ถึง 1.34)

\therefore เลขที่ถูกต้องมีจำนวน คือเลข 1.28 \rightarrow 1.32

ซึ่งมีจำนวน 34.1 + 34.1 = 68.2 %

คิดเป็นจำนวน $\frac{68.2}{100} \times 5000 = 3410$ ชิ้น

\therefore มีจำนวนชิ้นที่ได้มาตรฐาน ทั้งหมด 3,410 ชิ้น (มากกว่านี้ไม่ผ่าน แต่ไม่ผ่านก็รับใช้ได้)

② จากข้อแรก ① งานที่ทำเสร็จแล้วแต่ไม่ได้มาตรฐาน = 100 - 68.2 = 31.8 %

ซึ่ง 31.8 % คิดเป็น $\frac{31.8}{100} \times 5000 = 1590$ ชิ้น

\therefore ชิ้นงานที่ไม่ได้มาตรฐาน มีอยู่รวมกัน 1,590 ชิ้น

(ซึ่งข้อบกพร่องนี้ ถ้าไม่พบข้อบกพร่องที่แน่นอน อาจจะไปถึงจุดเลข 1.27 และ 1.33

จะไปตกอยู่ที่ด้านแบ่ง ที่เลข 1.30)

ข้อ 7 ลงรถที่จุดที่ 1 และจะจอดที่จุดที่ 2 ซึ่งมีคนขึ้นรถที่ท่าที่จอดที่ 1

หน้า 141 เรื่องข้อได้เป็นข้อได้

สมมติข้อที่ 1 : ประมวลผล หากขึ้นรถที่นั่นก่อนขึ้นรถที่จุดที่ 2 ต้องออกสองคนก่อนขึ้นรถที่จุดที่ 2

- ใครขึ้นก่อนขึ้นรถ
- เมื่อคนที่ขึ้นรถ อยู่แล้ว อยู่ตำแหน่งใดบ้าง

แล้ว เมื่อใช้ใช้รถที่นั่นไปสักพักแล้ว

- คนที่ขึ้นรถก่อนขึ้นรถ มีคนขึ้นรถหรือไม่
- คนที่ขึ้นรถอยู่แล้ว ~~มี~~ ไม่มีคนขึ้นรถที่นั่นใช้หรือไม่?

สมมติข้อที่ 2 : ประเด็นคือ รถที่ใช้มีคนขึ้นรถที่นั่น มีกี่คน เมื่อรถขึ้นรถที่จุดที่ 2 "คนขึ้นรถ" ขึ้นรถที่นั่นหรือไม่ (คนขึ้นรถที่นั่นมาขึ้นรถที่นั่นหรือไม่) ขึ้นรถที่นั่นหรือไม่

ดังนั้น ข้อมูลที่บอกนี้ เป็นข้อมูลรวม รถที่จอดที่นั่นไม่ได้

สมมติข้อที่ 3 : ถ้าจอด OK เพราะ ข้อมูลที่ได้ รวมรถที่จอดที่นั่น เช่นคนขึ้นรถที่นั่นจอดที่นั่น (จอดที่นั่น = 5 ในกรณีที่ ไม่จอดที่นั่น = 0 เป็นต้น)

สมมติข้อที่ 4 : รถที่จอดที่นั่น OK

สมมติข้อที่ 5 : ถ้าจอด OK แต่จอดที่นั่น

- รถที่จอดที่นั่น มีกี่คนขึ้นรถที่นั่น
- แล้วรถที่จอดที่นั่น (ที่ km/hr ; miles/hr) เพราะคนที่ขึ้นรถ, คนที่ขึ้นรถที่นั่น เป็นที่ขึ้นรถที่นั่นใช้หรือไม่ (คนที่ขึ้นรถที่นั่นใช้หรือไม่) เป็นที่ขึ้นรถที่นั่น
- ใครขึ้นรถที่นั่น และรถที่นั่น
- อุณหภูมิ และคนที่ขึ้นรถที่นั่น ขึ้นรถที่นั่น เป็นที่ขึ้นรถที่นั่น และจะมีรถที่นั่น

หากเป็นรถที่ขึ้นรถที่นั่น โดยที่รถที่ขึ้นรถที่นั่น ไปขึ้นรถที่นั่นกับ รถที่ขึ้นรถที่นั่น (จอดที่นั่น) ในจีนและรถที่นั่น และรถที่นั่น มีข้อมูลขึ้นรถที่นั่น เมื่อขึ้นรถที่นั่น จริง จึงมีรถที่นั่น รถที่ขึ้นรถที่นั่น จริง เมื่อขึ้นรถที่นั่น (จอดที่นั่น) เป็นที่ขึ้นรถที่นั่น

หน้า 145 รถที่ขึ้นรถที่นั่น และใช้เงินที่นั่น รถที่ขึ้นรถที่นั่น รถที่ขึ้นรถที่นั่น 5 คนขึ้นรถที่นั่น 2547

ส่วนที่ขึ้นรถที่นั่น

- คนที่ขึ้นรถที่นั่น รถที่ขึ้นรถที่นั่น 11 ล้าน หรือรถที่นั่น 13.5 ล้านคนขึ้นรถที่นั่น ใช้มีรถที่นั่น รถที่ขึ้นรถที่นั่น
- รถที่ขึ้นรถที่นั่น รถที่ขึ้นรถที่นั่น
 - รถที่ขึ้นรถที่นั่น รถที่ขึ้นรถที่นั่น 5 %
 - รถที่ขึ้นรถที่นั่น รถที่ขึ้นรถที่นั่น รถที่ขึ้นรถที่นั่น รถที่ขึ้นรถที่นั่น

66 จะ จบที่ 3 จบ

เจอกันที่ 4 "ทุกเรื่องจบที่นั่น"

เดือน ๓.๕๒ จบ
นี่แหละ