

บทเรียนที่ 8
การแจกแจง
ความถี่ของข้อมูล



สำหรับการเรียนรู้

- การแจกแจงความถี่ของข้อมูล
- การแจกแจงความถี่โดยใช้กราฟ
- การแจกแจงความถี่สะสมโดยใช้กราฟ



การแจกแจงความถี่ของข้อมูล

ข้อมูลที่ได้จากการเก็บรวบรวมข้อมูลโดยวิธีต่าง ๆ อาจจะมีอยู่ในลักษณะที่ไม่เป็นระเบียบ จำเป็นต้องจัดข้อมูลดังกล่าวให้มีระเบียบเป็นหมวดเป็นหมู่ เพื่อให้ง่ายต่อการคำนวณและสะดวกต่อการจัดทำข้อมูลเพื่อการวิเคราะห์ข้อมูลขั้นต่อไป เราเรียกวิธีทางสถิติว่า **การแจกแจงความถี่ของข้อมูล**

ความถี่ของค่าจากการสังเกตหรือข้อมูล คือ จำนวนครั้งของค่าของข้อมูลแต่ละค่าที่ได้มาจากการเก็บรวบรวมข้อมูล การหาความถี่นิยมใช้วิธีทำรอยขีด (tally) แล้วจึงสร้างตารางแจกแจงความถี่ (Frequency table)



การแจกแจงความถี่ของข้อมูล

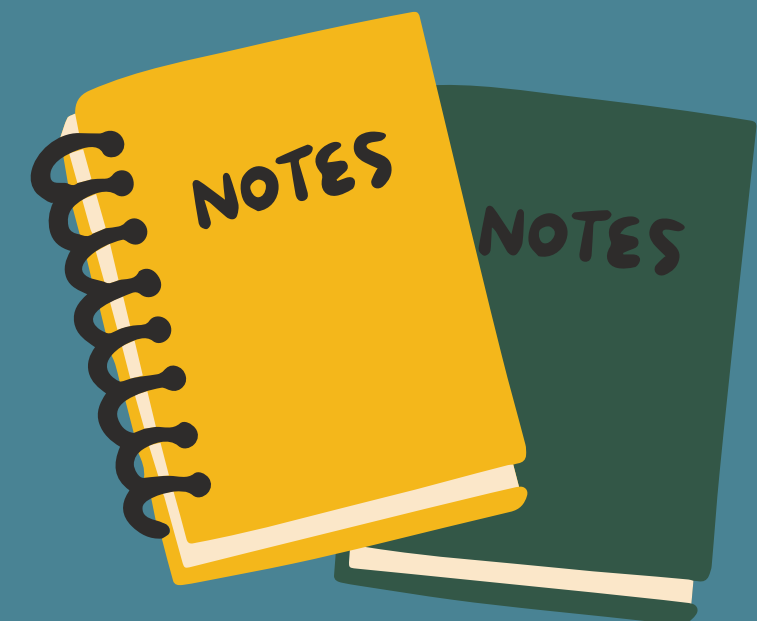
การแจกแจงความถี่ สามารถจัดเป็นลักษณะดังต่อไปนี้

1. การแจกแจงความถี่ (Frequency Distribution)
2. การแจกแจงความถี่สัมพัทธ์ (Relative Frequency Distribution)
3. การแจกแจงความถี่สะสม (Cumulative Frequency Distribution)
4. การแจกแจงความถี่สะสมสัมพัทธ์ (Relative Cumulative Frequency Distribution)

1. การแจกแจงความถี่โดยวิธีจัดเรียงค่าของข้อมูลตามลำดับ

2. การแจกแจงความถี่โดยวิธีจัดข้อมูลเป็นช่วงหรืออันตรภาคชั้น

การแจกแจงความถี่ด้วยตารางทำได้ 2 วิธี คือ



การแจกแจงความถี่ของข้อมูล

1. การแจกแจงความถี่โดยวิธีจัดเรียง ค่าของข้อมูลตามลำดับ

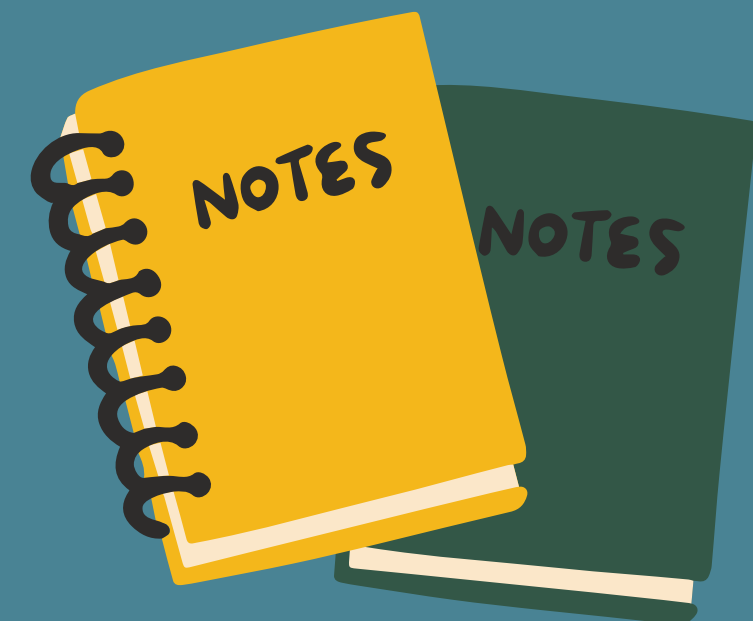
เป็นการนำข้อมูลทั้งหมดมาเรียง
ตามลำดับจากค่ามากไปน้อยหรือค่าน้อย
ไปมาก เหมาะสำหรับข้อมูลที่มีค่าแตกต่างกัน
กันน้อยและมีจำนวนข้อมูลไม่มากนัก

ตัวอย่างที่ 1 ในการสอบวิชาคณิตศาสตร์
พื้นฐานอาชีพของนักเรียน 20 คน ทำ
คะแนนได้ดังนี้

17 15 21 19 14 20 15 18 15 16
17 20 21 15 18 19 19 20 20 15

คะแนน	14	15	16	17	18	19	20	21
รอบขีด	I		I	II	II	III		II
ความถี่	1	5	1	2	2	3	4	2

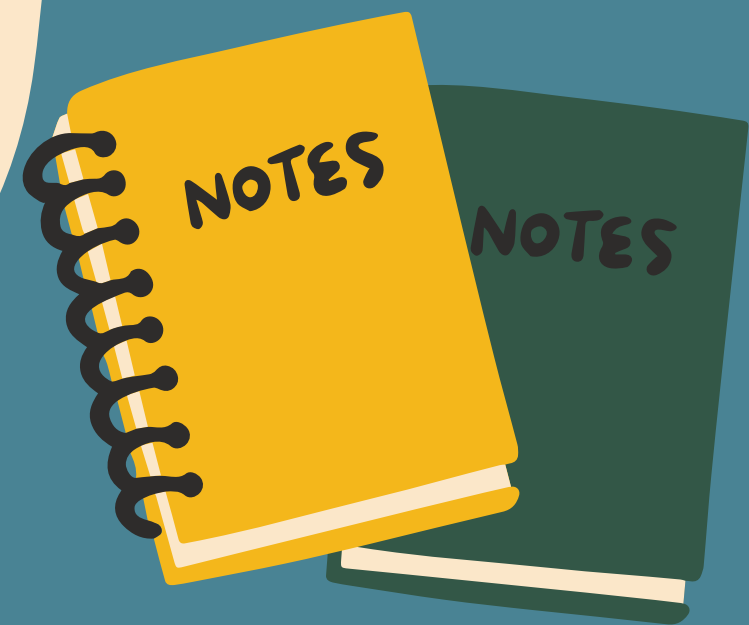
สร้าง
ตาราง
แจกแจง
ความถี่ได้
ดังนี้



การแจกแจงความถี่ของข้อมูล

2. การแจกแจงความถี่โดยวิธีจัดข้อมูลเป็นช่วงหรือ อันตรภาคชั้น

ในกรณีข้อมูลที่มีเป็นจำนวนมาก ถ้านำข้อมูลทำการแจก
แจงความถี่โดยวิธีจัดเรียงค่าของข้อมูลตามลำดับ อาจจะทำให้
ตารางมีขนาดใหญ่ คงจะไม่สะดวกและไม่มีประโยชน์มาก
นัก ดังนั้นจึงแบ่งข้อมูลออกเป็นช่วง ซึ่งเรียกว่า อันตรภาค
ชั้น (Class Interval) แล้วหารอยขีด จากนั้นนับรอยขีดรวม
เป็นความถี่



การแจกแจงความถี่ของข้อมูล

1.1 คำสำคัญที่ เกี่ยวกับตาราง แจกแจงความถี่

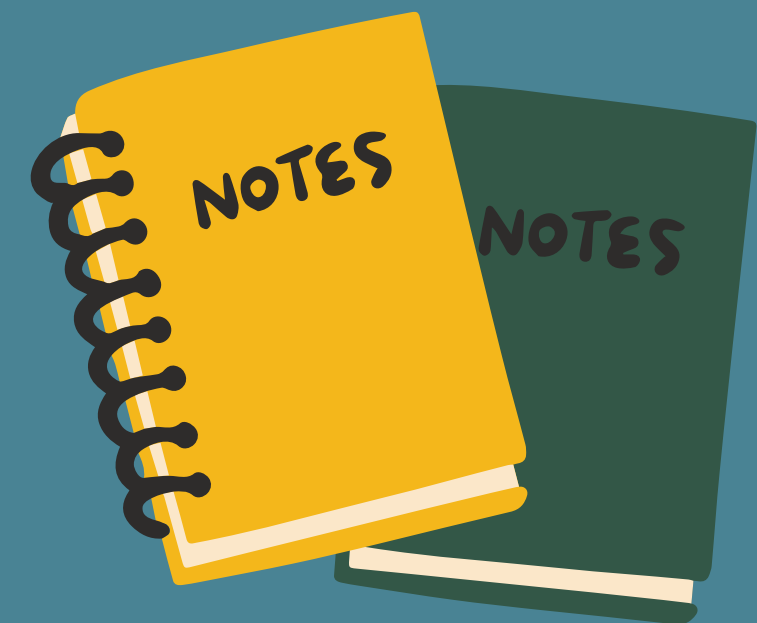
ตารางที่ 1 ความถี่ของคะแนนสอบวิชา
คณิตศาสตร์ของนักเรียนห้องหนึ่ง

คะแนน	ความถี่
30 - 39	1
40 - 49	19
50 - 59	10
รวม	30

➔ **1.1.1 อันตรภาคชั้น (Class Interval)** เป็นช่วง
คะแนนแต่ละช่วง คือ 30 - 39, 40 - 49, 50 - 59

การแจกแจงความถี่ของข้อมูล

1.1.2 จำนวนอันดับชั้น หมายถึง
จำนวนอันดับชั้นทั้งหมดในตารางแจก
แจงความถี่
ดังตารางที่ 1 นี้มีจำนวนอันดับชั้น 3 ชั้น



การแจกแจงความถี่ของข้อมูล

1.1.3 ขีดจำกัดของ อันดับชั้น (Class limit)

- 1) ขีดจำกัดล่างของอันดับชั้น (Lower class limit)
คือ ค่าน้อยที่สุดของแต่ละอันดับชั้นนั้น
- 2) ขีดจำกัดบนของอันดับชั้น (Upper class limit)
คือ ค่ามากที่สุดของแต่ละอันดับชั้นนั้น

จากตารางที่ 1

อันดับชั้นที่ 1 มีขีดจำกัดล่างคือ 30 ขีดจำกัดบนคือ 39
อันดับชั้นที่ 2 มีขีดจำกัดล่างคือ 40 ขีดจำกัดบนคือ 49
อันดับชั้นที่ 3 มีขีดจำกัดล่างคือ 50 ขีดจำกัดบนคือ 59

หมายถึง ตัวเลขที่แสดงถึงช่วงหรือบอกให้ทราบ
ว่ามีตัวเลขอะไรที่อยู่ในชั้นนั้นบ้าง ซึ่งขีดจำกัด
อันดับชั้นประกอบด้วย



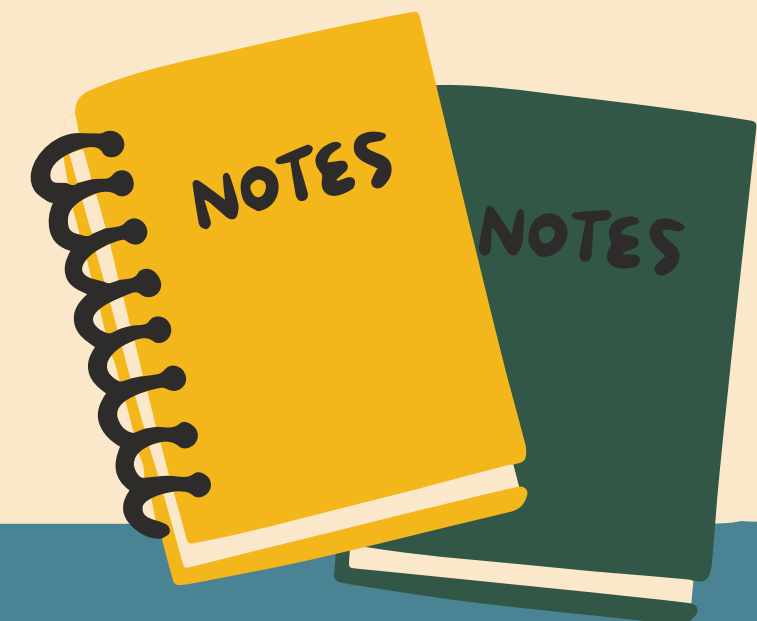
การแจกแจงความถี่ของข้อมูล

1.1.4 ขอบเขตอันตรภาคชั้น (Class boundary) หมายถึง ขีดจำกัดชั้นที่แท้จริงของข้อมูล เป็นค่าที่ขยายหรือครอบคลุมขีดจำกัดชั้น เพื่อให้ขอบเขตในแต่ละชั้นต่อเนื่องกัน

1) ขอบล่างของอันตรภาคชั้น (Lower boundary) คือ ค่ากึ่งกลางระหว่างขีดจำกัดล่างของอันตรภาคชั้นกับขีดจำกัดบนของอันตรภาคชั้นต่ำกว่าที่อยู่ติดกัน

$$\text{ขอบล่างของอันตรภาคชั้น} = \frac{(\text{ขีดจำกัดบนของชั้นก่อน} + \text{ขีดจำกัดล่างของชั้นนั้น})}{2}$$

2) ขอบบนของอันตรภาคชั้น (Upper boundary) คือ ค่ากึ่งกลางระหว่างขีดจำกัดบนของอันตรภาคชั้นกับขีดจำกัดล่างของอันตรภาคชั้นสูงกว่าที่อยู่ติดกัน

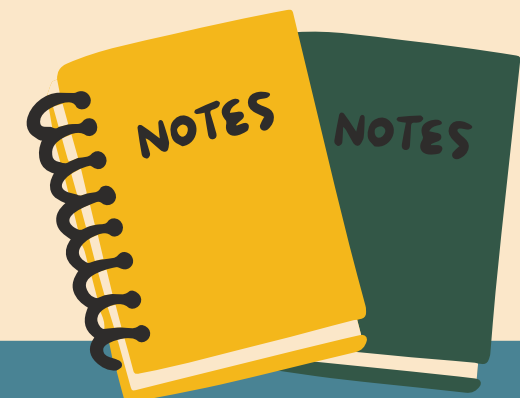
$$\text{ขอบบนของอันตรภาคชั้น} = \frac{(\text{ขีดจำกัดบนของชั้นนั้น} + \text{ขีดจำกัดล่างของชั้นถัดไป})}{2}$$


การแจกแจงความถี่ของข้อมูล

1.1.5 ความกว้างของอันตรภาคชั้น (Class Interval) คือ ค่าที่เป็นไปได้ทั้งหมดในอันตรภาคชั้นนั้น เขียนแทนด้วย I หรือ C เช่น อันตรภาคชั้นที่ 1 มีช่วงคะแนน 30 - 39 ประกอบด้วยค่าที่เป็นไปได้ 10 ค่า คือ 30, 31, ... , 39 หรือหาได้จากผลต่างระหว่างขอบบนและขอบล่างของแต่ละอันตรภาคชั้น

1.1.6 จุดกึ่งกลางของอันตรภาคชั้น (Mid point) คือ ค่ากึ่งกลางระหว่างขอบล่างกับขอบบนของอันตรภาคชั้น หรือค่ากึ่งกลางระหว่างขีดจำกัดล่างกับขีดจำกัดบนของอันตรภาคชั้นนั้น นั่นคือ

$$\begin{aligned}\text{จุดกึ่งกลางของอันตรภาคชั้น} &= \frac{(\text{ขอบล่าง} + \text{ขอบบน})}{2} \\ &= \frac{(\text{ขีดจำกัดล่าง} + \text{ขีดจำกัดบน})}{2}\end{aligned}$$



การแจกแจงความถี่ของข้อมูล

1.2 การสร้างตารางแจกแจงความถี่



โดยจัดข้อมูลเป็นช่วงหรือ
อันดับภาคชั้น มีขั้นตอนดังนี้
1.2.1 หาพิสัย (Range)
ของข้อมูล คือ หาผลต่างของ
ค่าสูงสุดกับค่าต่ำสุดของ
ข้อมูลชุดนั้น

นั่นคือ
$$\text{พิสัย(R)} = \text{ค่าสูงสุด} - \text{ค่าต่ำสุด}$$

1.2.2 กำหนดจำนวนอันดับภาคชั้น
พอสมควรที่นิยมใช้กัน ประมาณ 5 - 15
ชั้น

1.2.3 กำหนดความกว้างของ
อันดับภาคชั้นโดยประมาณได้จาก
สูตร

$$\text{ความกว้างของอันดับภาคชั้น (I)} = \frac{\text{ค่าสูงสุด} - \text{ค่าต่ำสุด}}{\text{จำนวนอันดับภาคชั้น}}$$

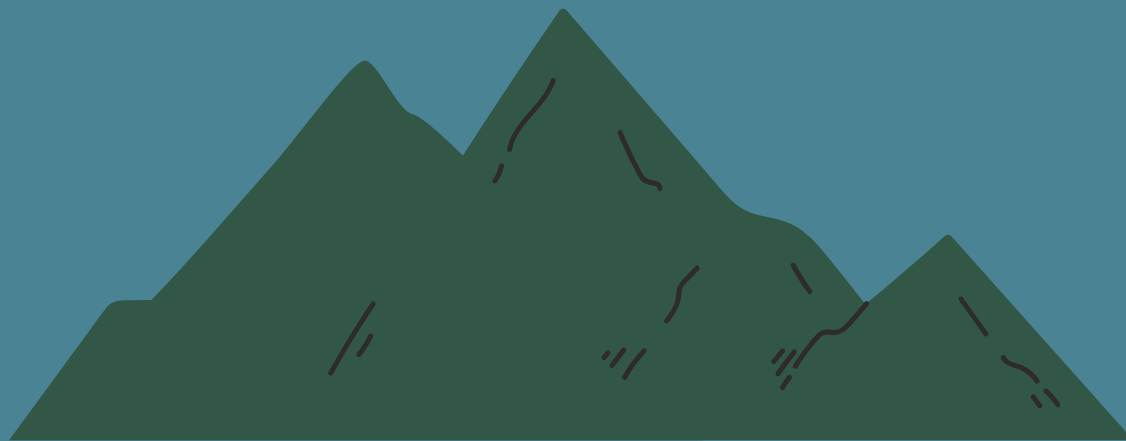


การแจกแจงความถี่ของข้อมูล

1.2 การสร้างตารางแจกแจงความถี่

1.2.4 เขียนอันตรภาคชั้นแต่ละอันตรภาคชั้น เรียงตามลำดับโดยยึดหลักว่าค่าต่ำสุดต้องอยู่ในอันตรภาคชั้นต่ำสุด และค่าสูงสุดต้องอยู่ในอันตรภาคชั้นสูงสุด

1.2.5 นับรอยขีดและหาความถี่



การแจกแจงความถี่ของข้อมูล

1.3 การแจกแจงความถี่สัมพัทธ์

ความถี่สัมพัทธ์ (Relative frequency) ของแต่ละอันตรภาคชั้น คือ อัตราส่วนระหว่างความถี่ของอันตรภาคชั้นนั้นกับผลรวมของความถี่ทั้งหมด

ความถี่สัมพัทธ์อาจแสดงอยู่ในรูปของเศษส่วน ทศนิยม หรือบางครั้งอาจแสดงอยู่ในรูปร้อยละ เรียกว่า “ร้อยละของความถี่สัมพัทธ์” นั่นคือ

$$\begin{aligned}\text{ความถี่สัมพัทธ์} &= \frac{\text{ความถี่ของอันตรภาคชั้นนั้น}}{\text{ผลรวมของความถี่ทั้งหมด}} \\ \text{ร้อยละของความถี่สัมพัทธ์} &= \frac{\text{ความถี่ของอันตรภาคชั้นนั้น}}{\text{ผลรวมของความถี่ทั้งหมด}} \times 100\%\end{aligned}$$

ข้อสังเกต

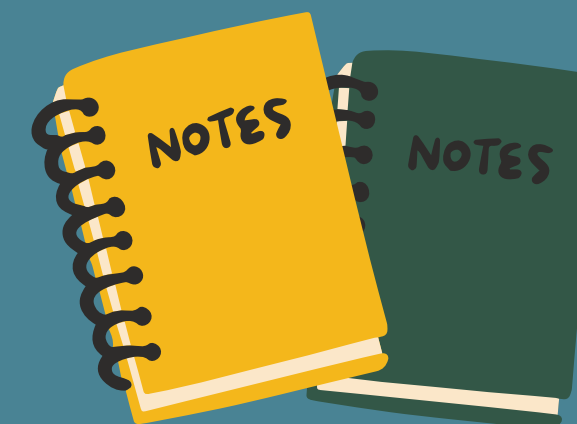
1. ผลรวมของความถี่สัมพัทธ์ของทุกอันตรภาคชั้นจะมีค่าเท่ากับ 1 เสมอ

2. ผลรวมของร้อยละความถี่สัมพัทธ์ของทุกอันตรภาคชั้นจะมีค่าเท่ากับ 100 เสมอ

1.4 การแจกแจงความถี่สะสม

ความถี่สะสม (Cumulative frequency) ของแต่ละอันตรภาคชั้น คือ ผลรวมของความถี่ของอันตรภาคชั้นนั้นกับความถี่ของอันตรภาคชั้นที่ต่ำกว่าทั้งหมดหรือสูงกว่าทั้งหมดอย่างใดอย่างหนึ่ง

ข้อสังเกต
โดยทั่วไปการหาความถี่สะสมนิยมใช้ความถี่สะสมจากน้อยไปมาก



การแจกแจงความถี่ของข้อมูล

1.5 การแจกแจงความถี่สะสมสัมพัทธ์

ความถี่สะสมสัมพัทธ์ (Relative cumulative frequency) ของแต่ละอันตรภาคชั้น คือ อัตราส่วนระหว่างความถี่สะสมของอันตรภาคชั้นนั้นกับผลรวมของความถี่ทั้งหมด

ความถี่สะสมสัมพัทธ์อาจแสดงในรูปเศษส่วน ทศนิยม หรือบางครั้งอาจจะแสดงอยู่ในรูปร้อยละ เรียกว่า “ร้อยละของความถี่สะสมสัมพัทธ์” นั่นคือ

$$\begin{aligned}\text{ความถี่สะสมสัมพัทธ์} &= \frac{\text{ความถี่สะสมของอันตรภาคชั้นนั้น}}{\text{ผลรวมของความถี่ทั้งหมด}} \\ \text{ร้อยละของความถี่สะสมสัมพัทธ์} &= \frac{\text{ความถี่สะสมของอันตรภาคชั้นนั้น}}{\text{ผลรวมของความถี่ทั้งหมด}} \times 100\%\end{aligned}$$

1.6 ข้อสังเกตเกี่ยวกับการแจกแจงความถี่

1) ความกว้างของแต่ละอันตรภาคชั้นไม่จำเป็นต้องเท่ากันทุกอันตรภาคชั้น ทั้งนี้ขึ้นอยู่กับวัตถุประสงค์ของการใช้ข้อมูล และอาจขึ้นอยู่กับความแตกต่างระหว่างค่าจากการสังเกตของข้อมูลชุดนั้น

2) ถ้าค่าจากการสังเกตบางค่าซึ่งมีค่าต่างไปจากค่าอื่นๆ ในข้อมูลชุดนั้นมาก ควรกำหนดอันตรภาคชั้นแรกหรือชั้นสุดท้ายเป็นอันตรภาคชั้นเปิด (Open - end class Interval) กล่าวคือเป็นอันตรภาคชั้นที่ไม่มีขอบล่างหรือขอบบน

3) การสร้างตารางแจกแจงความถี่บางอันตรภาคชั้นอาจมีความถี่เป็นศูนย์ได้



การแจกแจงความถี่ โดยใช้กราฟ

การแจกแจงความถี่โดยใช้กราฟ จะทำให้
เห็นการกระจายของข้อมูลได้ชัดเจนกว่าการดู
จากตารางแจกแจงความถี่ กราฟที่แสดง
การแจกแจงความถี่มีดังนี้

- ฮิสโทแกรม (Histogram)
- รูปหลายเหลี่ยมของความถี่ (Frequency polygon)
- เส้นโค้งความถี่ (Frequency curve)
- แผนภาพต้น - ใบ (Stem and leaf display)





ฮิสโทแกรม (Histogram)

- เป็นกราฟแท่งรูปสี่เหลี่ยมมุมฉากวางติดกัน โดยความกว้างของรูปสี่เหลี่ยมมุมฉากแทนความกว้างของแต่ละอันดับภาคขึ้นตามแกนแนวนอน และความยาวของรูปสี่เหลี่ยมมุมฉากแต่ละรูปแทนความถี่ของแต่ละอันดับภาคขึ้นตามแกนตั้ง
- วิธีการเขียนรูปฮิสโทแกรมอาจใช้ขอบล่าง ขอบบน หรือจุดกึ่งกลางชั้นของแต่ละอันดับภาคขึ้นในตารางแจกแจงความถี่ก็ได้



รูปหลายเหลี่ยม ของความถี่ (Frequency polygon)

- คือ รูปเหลี่ยมที่เกิดจากการต่อจุดกึ่งกลางของยอดของแท่งสี่เหลี่ยมมุมฉากของฮิสโทแกรม รวมทั้งจุดกึ่งกลางของอันตรภาคชั้นที่ต่ำกว่าและสูงกว่าอย่างละ 1 ชั้น



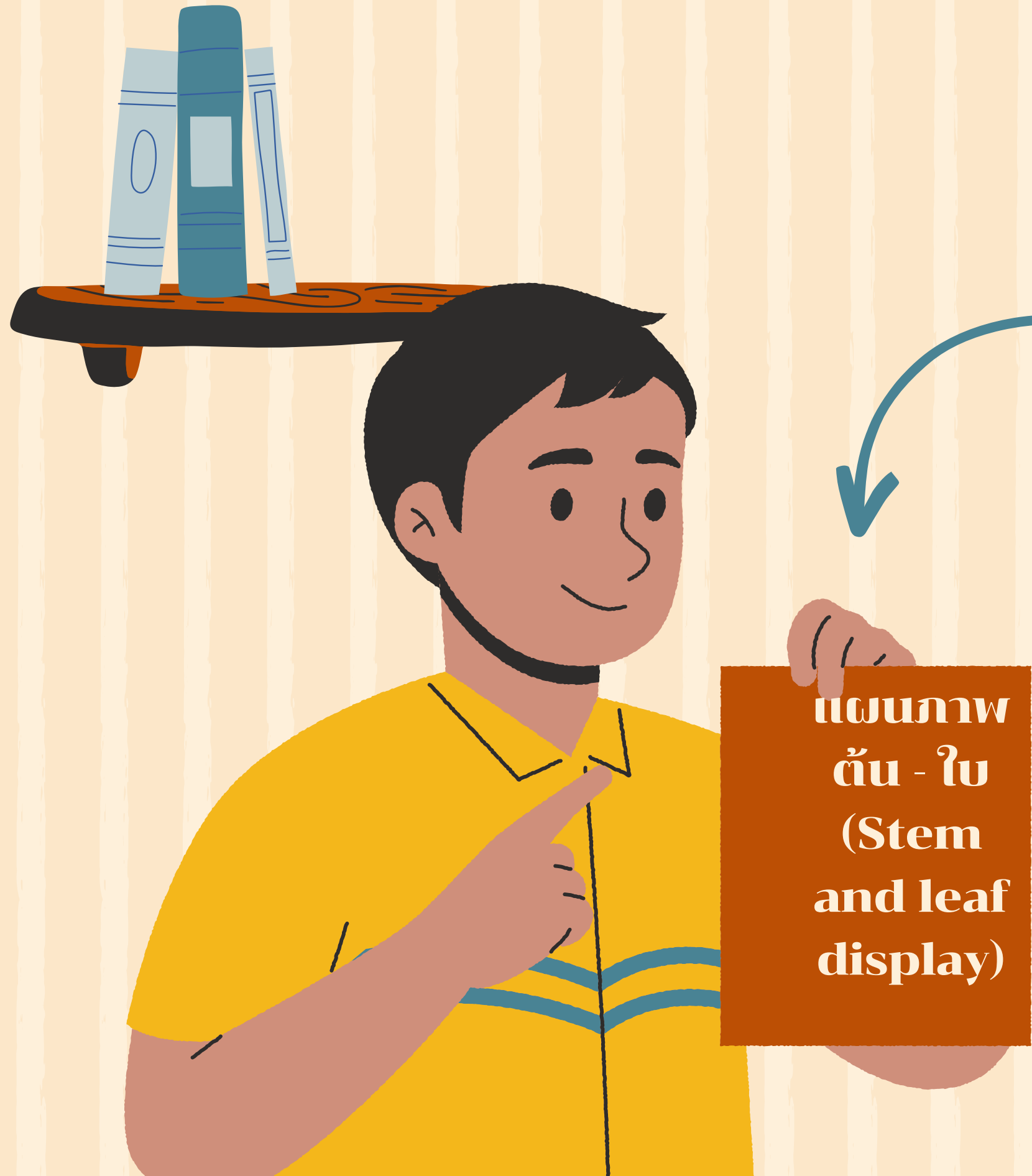
เส้นโค้งความถี่ (Frequency curve)

- คือ เส้นโค้งที่ได้จากการปรับด้านของรูปหลายเหลี่ยมของความถี่ให้เรียบขึ้น โดยการปรับพื้นที่ใต้เส้นโค้ง ให้มีพื้นที่ใกล้เคียงกับพื้นที่รูปหลายเหลี่ยมของความถี่



แผนภาพต้น - ใบ (Stem and leaf display)

- ในการจัดข้อมูลเป็นกลุ่ม ๆ นอกจากจะใช้ตารางแจกแจงความถี่หรือการแจกแจงความถี่โดยใช้กราฟแล้ว อาจจะใช้วิธีการสร้างแผนภาพเพื่อแจกแจงความถี่และสามารถวิเคราะห์ข้อมูลเบื้องต้นไปพร้อมๆ กันที่เรียกว่า แผนภาพต้น-ใบ (stem and leaf display) โดยมีขั้นตอนดังต่อไปนี้



แผนภาพต้น - ใบ (Stem and leaf display)

1. จัดแบ่งข้อมูลออกเป็นกลุ่ม ๆ โดยเรียงลำดับจากน้อยไปหามาก หรือมากไปน้อย
 - ถ้ากลุ่มข้อมูลเป็นเลขสองหลักให้ ใช้เลขโดดของหลักแรก (หลักสิบ)
 - ถ้ากลุ่มข้อมูลเป็นเลขสามหลักให้ ใช้เลขโดดของสองหลักแรก คือ หลักร้อยและหลักสิบ
2. นำกลุ่มข้อมูลในข้อ 1 มาใส่ในหลัก (แนวตั้งในหลักแรก) เพื่อสร้างเป็นลำต้น (stem)



แผนภาพต้น - ใบ (Stem and leaf display)

3. เมื่อได้ลำดับ จังหวะและความถี่
โดยพิจารณาจากข้อมูลที่มีอยู่ในแต่ละแถว
แนวนอน นำมาต่อเป็นใบ (leaf) ของแต่ละ
กลุ่ม

4. เรียงลำดับข้อมูลจากน้อยไปมากของแต่ละ
กลุ่ม (แต่ละแถวแนวนอน) ก็จะได้แผนภาพ
ต้น-ใบ (stem and leaf display)

แผนภาพ
ต้น - ใบ
(Stem
and leaf
display)

การแจกแจง ความถี่สะสมโดย ใช้กราฟ

กราฟแสดงความถี่สะสม เรียกว่า “เส้นโค้ง
ความถี่สะสม” หรือเส้นโค้งโอจีฟ (Ogive
curve) เส้นโค้งความถี่สะสม เป็นกราฟแสดง
ความสัมพันธ์ระหว่างขอบบนของอันตรภาค
ชั้นกับความถี่สะสม แล้วโยงจุดนั้นด้วย
เส้นตรง จากนั้นปรับให้เป็นเส้นโค้งเรียบ
กราฟแสดงความถี่สะสมสามารถนำไป
ใช้ประโยชน์ในการประมาณความถี่ของค่าของ
ข้อมูลที่ต่ำกว่าหรือสูงกว่าคะแนนที่กำหนดให้

(Ogive
curve)

