

## ใบความรู้ที่ 5 เรื่อง การแจกแจงความถี่ของข้อมูล

การวิเคราะห์ข้อมูลเบื้องต้นเกี่ยวกับเรื่องใดๆ เป็นการวิเคราะห์เพื่อทราบลักษณะโดยรวมของข้อมูลเกี่ยวกับเรื่องนั้นๆ เช่น

1. การวิเคราะห์ข้อมูลเบื้องต้นเกี่ยวกับโรงเรียนใด โรงเรียนหนึ่ง ลักษณะโดยรวมของข้อมูลเกี่ยวกับโรงเรียนที่ควรจะทราบคือ

จำนวนนักเรียนจำแนกตามชั้นเรียนที่เปิดสอน

จำนวนนักเรียน โดยเฉลี่ยต่อห้อง

จำนวนครูจำแนกตามเพศ อายุและระดับการศึกษา

2. การวิเคราะห์ข้อมูลเบื้องต้นเกี่ยวกับความเป็นอยู่ของประชาชนในท้องที่แห่งหนึ่ง ลักษณะโดยรวมของข้อมูลเกี่ยวกับความเป็นอยู่ของประชาชนที่ควรจะทราบคือ

จำนวนสมาชิกโดยเฉลี่ยของครอบครัว

จำนวนครอบครัวจำแนกตามอาชีพ

รายได้เฉลี่ยต่อครอบครัว และการกระจายของรายได้ในแต่ละครอบครัว

จำนวนครอบครัวจำแนกตามระดับการศึกษาของหัวหน้าครอบครัว

ลักษณะโดยรวมของข้อมูลตามตัวอย่างข้างต้น จะเห็นได้ว่าการวิเคราะห์ที่เกี่ยวกับเรื่องต่างๆ ต่อไปนี้ทั้งสิ้น

การแจกแจงความถี่ของข้อมูล

การหาค่ากลางของข้อมูล

การหาค่าการกระจายของข้อมูล

กล่าวได้ว่า การสรุปหรืออธิบายและนำเสนอข้อมูลด้วยค่าสถิติชุดหนึ่ง บอกลักษณะบางประการที่เป็นประโยชน์ต่อการทำความเข้าใจข้อมูลชุดนั้นได้ รวมทั้งการนำไปใช้ในเชิงวิเคราะห์หรือช่วยการตัดสินใจ ค่าสถิติที่สำคัญมากสองค่า คือ ค่าที่บอกแนวโน้มสู่ส่วนกลางของข้อมูล กับค่าที่บอกถึงการกระจายของค่าต่างๆ จากค่ากลางในข้อมูลชุดนั้น

ก่อนกล่าวถึงรายละเอียดเกี่ยวกับวิธีวิเคราะห์ข้อมูลเบื้องต้น ควรจะต้องทราบความหมายของ "ตัวแปร" ที่จำเป็นต้องกล่าวถึงอยู่เสมอๆ ในการวิเคราะห์ข้อมูลเสียก่อน ดังนี้

"ตัวแปร" (Variable) หมายถึง ลักษณะของประชากรที่เราสนใจวิเคราะห์ โดยที่ลักษณะนั้นๆ ของประชากรสามารถเปลี่ยนแปลงค่าได้ ไม่ว่าจะเป็นเชิงปริมาณหรือเชิงคุณภาพ เช่น

1) จำนวนสมาชิกของครอบครัวที่เป็นตัวแปร เนื่องจากเปลี่ยนค่าได้ ตั้งแต่ 1, 2, 3 เรื่อยไป และเป็นตัวแปรเชิงปริมาณ

2) คะแนนสอบเป็นตัวแปร เนื่องจากเปลี่ยนค่าได้ ตั้งแต่ 1, 2, เรื่อยไปจนถึง 100 (คะแนนเต็ม 100) และเป็นตัวแปรเชิงปริมาณ

3) เพศ เป็นตัวแปร เนื่องจากเปลี่ยนค่าเป็น ชายและหญิง และเป็นตัวแปรเชิงคุณภาพ  
ตัวแปรที่นำมาวิเคราะห์ส่วนใหญ่มักเป็นตัวแปรที่เปลี่ยนค่าในเชิงปริมาณ ซึ่งสามารถวัดออกมาเป็นจำนวนได้ เช่น อายุ รายได้ ราคาสินค้า จำนวนคนงาน ฯลฯ

ถ้าให้  $X$  เป็นตัวแปรที่แสดงผลการสอบซึ่งมีคะแนนเต็ม 10 คะแนน คะแนนที่นักเรียนแต่ละคนสอบได้นี้เรียกว่า "ค่าจากการสังเกต" และเรียกคะแนนที่อาจเป็นไปได้สำหรับการสอบ ซึ่งมี 11 ค่า คือ 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, และ 10 ว่า "ค่าที่เป็นไปได้"

ตัวอย่างของตัวแปรเชิงคุณภาพและค่าที่เป็นไปได้ ได้แก่

ตัวแปร	ค่าที่เป็นไปได้
ฤดูกาล	ฤดูใบไม้ผลิ ฤดูใบไม้ร่วง ฤดูร้อน ฤดูหนาว
สถานภาพการสมรส	โสด สมรส หย่า
วันในสัปดาห์	วันจันทร์ วันอังคาร วันพุธ วันพฤหัสบดี วันศุกร์ วันเสาร์ วันอาทิตย์

### การแจกแจงความถี่ของข้อมูล

การแจกแจงความถี่ เป็นวิธีการทางสถิติอย่างหนึ่งที่ใช้ในการจัดข้อมูลที่มีอยู่ หรือที่เก็บรวบรวมมาได้ให้อยู่เป็นพวกๆ เพื่อความสะดวกในการนำเสนอข้อมูลและการวิเคราะห์ข้อมูลเหล่านั้น การแจกแจงความถี่มักจะทำเมื่อข้อมูลที่จะทำการศึกษามีเป็นจำนวนมาก หรือข้อมูลมีค่าซ้ำกันอยู่มาก เพราะจะช่วยประหยัดเวลา และสรุปผลได้ชัดเจนขึ้น และเหมาะสมที่จะนำไปใช้เป็นประโยชน์ต่อไป

1. การแจกแจงความถี่โดยไม่จัดข้อมูลเป็นกลุ่ม(อันตรภาคชั้น) เป็นวิธีที่เหมาะสมกับข้อมูลที่มีจำนวนไม่มากนัก และค่าต่ำสุดของข้อมูลมีค่าไม่ต่างกัน โดยจัดข้อมูลไว้ในตาราง เรียกว่า "ตารางแจกแจงความถี่" ซึ่งมีขั้นตอนดังนี้

ขั้นที่ 1 เรียงข้อมูลจากข้อมูลน้อยไปหาข้อมูลมากหรือจากข้อมูลมากไปหาน้อย

ขั้นที่ 2 เขียนข้อมูลทั้งหมดให้ต่อเนื่องกันจากน้อยไปหามากหรือมากไปหาน้อย

ขั้นที่ 3 เขียนรอยขีดลงในช่องรอยขีดเพื่อแสดงว่าข้อมูลนั้นมีจำนวนเท่าใด ซึ่งเรียกว่า

ความถี่ของข้อมูล

ตัวอย่างที่ 1 จากผลการสอบวิชาคณิตศาสตร์พื้นฐานของนักเรียนจำนวน 30 คน ปรากฏคะแนนดังนี้

24 23 22 26 25 28 29 23 24 27  
25 24 22 24 23 25 26 25 27 25  
28 25 26 22 24 23 26 25 28 27

วิธีทำ เรียงคะแนนจากน้อยหามากได้ดังตาราง

คะแนน	รอยขีด	จำนวนรอยขีด
22	///	3
23	////	4
24	/////	5
25	///// //	7
26	////	4
27	///	3
28	///	3
29	/	1
	รวม (N)	30

จากตาราง จำนวนรอยขีดของข้อมูลแต่ละค่า เรียกว่า "ความถี่"

จะเห็นว่า ตารางแจกแจงความถี่ข้างต้น ไม่ค่อยมีประโยชน์ต่อการวิเคราะห์ข้อมูลมากนัก ทั้งนี้เนื่องจาก ค่าที่เป็นไปได้มีน้อยและค่าที่ได้จากการสังเกตก็มีจำนวนน้อยด้วย แต่ถ้าค่าที่เป็นไปได้มีจำนวนมาก เช่น คะแนนเต็มในการสอบวิชาคณิตศาสตร์เป็น 100 แทนที่จะเป็น 10 คะแนน และจำนวนนักเรียนที่สอบก็มีจำนวนมากด้วย เช่น มีจำนวน 60 คน แทนที่จะมีเพียง 6 คน การแจกแจงความถี่โดยใช้ทุกๆค่าที่เป็นไปได้ทั้งหมด ซึ่งในกรณีนี้มี 101 ค่า ทำให้เสียเวลาในการรวบรวมข้อมูลให้อยู่เป็นพวกและยังเป็นการยากในการสรุปผลที่สำคัญๆ ให้รัดกุมและง่ายต่อการที่จะนำไปใช้อีกด้วย

**2. การแจกแจงความถี่โดยจัดเป็นกลุ่ม (อันตรภาคชั้น) เหมาะสำหรับข้อมูลที่มีจำนวนมาก และค่าสูงสุดกับค่าต่ำสุดของข้อมูลต่างกันมาก**

การสร้างตารางแจกแจงความถี่โดยไม่จัดเป็นกลุ่ม (อันตรภาคชั้น) มีขั้นตอนดังนี้

ขั้นที่ 1 หาข้อมูลสูงสุดและต่ำสุด

ขั้นที่ 2 หาพิสัยซึ่งพิสัย = ข้อมูลสูงสุด-ข้อมูลต่ำสุด

ขั้นที่ 3 กำหนดจำนวนชั้น โดยปกติจำนวนชั้นจะอยู่ระหว่าง 5-15 ชั้นซึ่งแบ่งเป็น

1. ถ้ากำหนดจำนวนชั้นให้จะต้องหาความกว้างของอันตรภาคชั้น ดังนี้

$$\text{ความกว้างของอันตรภาคชั้น} = \frac{P}{N}$$

เมื่อ  $P = \text{พิสัย} = \text{ข้อมูลสูงสุด} - \text{ข้อมูลต่ำสุด}$

$N = \text{จำนวนชั้น}$

2. ถ้ากำหนดความกว้างของอันตรภาคชั้นให้ จะต้องหาจำนวนชั้น ดังนี้

$$\text{จำนวนชั้น} = \frac{P}{R}$$

เมื่อ  $P = \text{พิสัย} = \text{ข้อมูลสูงสุด} - \text{ข้อมูลต่ำสุด}$

$R = \text{ความกว้างของอันตรภาคชั้น}$

ทั้งนี้ ผลลัพธ์ที่ได้จากการหารในข้อ 1 และข้อ 2 ถ้าเป็นจำนวนเต็มให้บวกด้วย 1 แต่ถ้าเป็นทศนิยมจะต้องปัดให้เป็นจำนวนเต็มเสมอ

ชั้นที่ 4 เขียนอันตรภาคชั้นจากชั้น ข้อมูลต่ำสุดไปหาชั้นข้อมูลสูงสุดหรือจากชั้นของ ข้อมูลสูงสุดไปหาชั้นข้อมูลต่ำสุดก็ได้

ชั้นที่ 5 พิจารณาว่าข้อมูลแต่จำนวน จำนวนใดอยู่ในช่วงข้อมูลใดแล้วจัดลงในช่องรอย ข้อมูล โดยให้หนึ่งขีดแทนข้อมูล 1 จำนวน

ชั้นที่ 6 จำนวนรอยขีดแต่ละชั้น คือ ความถี่ของข้อมูลในชั้นนั้น

ตัวอย่างที่ 2 จากผลการสอบวิชาคณิตศาสตร์พื้นฐานของนักเรียนจำนวน 40 คน เป็นดังนี้

35	100	56	49	64	85	64	65	51	84
95	84	66	72	83	89	64	66	73	69
65	87	56	78	77	69	87	56	47	95
47	79	76	55	83	68	75	76	41	72

จงสร้างตารางแจกแจงความถี่ให้มีจำนวนชั้นเป็น 7

วิธีทำ คะแนนสูงสุดเท่ากับ 100 และ คะแนนต่ำสุดเท่ากับ 35

$$\therefore \text{พิสัย} = 100 - 35 = 65$$

$$\text{ความกว้างของอันตรภาคชั้น} = \frac{P}{N}$$

$$= \frac{65}{7}$$

$$= 9.2$$

$$\approx 10$$

คะแนน	รอยขีด	จำนวนรอยขีด
35-44	//	2
45-54	////	4
55-64	///// //	7
65-74	///// ////	10
75-84	///// ////	10
85-94	////	4
95-104	///	3
	รวม (N)	40

### 1. การแจกแจงความถี่สะสม

**ความถี่สะสม** (cumulative frequency) ของค่าที่เป็นไปได้ค่าใดหรืออันตรภาคชั้นใด คือ ผลรวมของความถี่ของค่านั้นหรือของอันตรภาคชั้นนั้นหรือของอันตรภาคชั้นนั้นกับความถี่ของค่าหรือของอันตรภาคชั้นที่มีช่วงคะแนนต่ำกว่าทั้งหมด (หรือสูงกว่าทั้งหมด) อย่างใดอย่างหนึ่ง

**การแจกแจงความถี่สะสม** คือ การหาผลรวมของความถี่แต่ละชั้นต่อเนื่องกันมาตั้งแต่ต้นตามลำดับจากชั้นที่มีค่าน้อยที่สุดไปหาชั้นที่มีค่ามากที่สุด หรือจากชั้นที่มีค่ามากที่สุดไปหาชั้นที่มีค่าน้อยที่สุด

**ตัวอย่างที่ 3** จงสร้างตารางแจกแจงความถี่สะสม จากตารางแจกแจงความถี่ต่อไปนี้

คะแนน	10-19	20-29	30-39	40-49	50-59	60-69	70-79
ความถี่	7	13	18	22	26	10	4

### วิธีทำ

คะแนน	ความถี่	ความถี่สะสม
10-19	7	7
20-29	13	7+13=20
30-39	18	20+18=38
40-49	22	38+22=60
50-59	26	60+26=86
60-69	10	86+10=96
70-79	4	96+4=100

จากตารางแสดงว่า

คะแนนที่มีค่าต่ำกว่า 9.5 มี 0 จำนวน

คะแนนที่มีค่าต่ำกว่า 19.5 มี 7 จำนวน

คะแนนที่มีค่าต่ำกว่า 29.5 มี 20 จำนวน

ทั้งนี้ วัตถุประสงค์สำคัญในการสร้างตารางแจกแจงความถี่ ก็เพื่อเปรียบเทียบความถี่ของค่าจากการสังเกตทั้งหมดที่ตกอยู่ในแต่ละค่าที่เป็นไปได้หรือในแต่ละอันตรภาคชั้น

## 2. การแจกแจงความถี่สัมพัทธ์

ความถี่สัมพัทธ์ (relative frequency) ของค่าที่เป็นไปได้ค่าใดหรืออันตรภาคชั้นใด คือ อัตราส่วนระหว่างความถี่ของค่านั้นหรือของอันตรภาคชั้นนั้นกับผลรวมของความถี่ทั้งหมด ซึ่งอาจจะแสดงในรูปเศษส่วน ทศนิยม หรือร้อยละ ดังตัวอย่างที่ 4

ตัวอย่างที่ 4 จงสร้างตารางแจกแจงความถี่สัมพัทธ์ จากตารางแจกแจงความถี่ในตัวอย่างที่ 3

วิธีทำ

คะแนน	ความถี่	ความถี่สัมพัทธ์	ร้อยละของความถี่สัมพัทธ์
10-19	7	0.07 (7 ÷ 100)	7.0
20-29	13	0.13 (13 ÷ 100)	13.0
30-39	18	0.18 (18 ÷ 100)	18.0
40-49	22	0.22 (22 ÷ 100)	22.0
50-59	26	0.26 (26 ÷ 100)	26.0
60-69	10	0.10 (10 ÷ 100)	10.0
70-79	4	0.04 (4 ÷ 100)	4.0
รวม	100	1.00	100

ทั้งนี้ วัตถุประสงค์สำคัญในการสร้างตารางแจกแจงความถี่สัมพัทธ์ เพื่อหาว่าความถี่ของแต่ละค่าที่เป็นไปได้หรือแต่ละอันตรภาคชั้นมีความถี่สะสมเป็นจำนวนมากน้อยเพียงใดเมื่อเทียบกับความถี่ทั้งหมด

## 3. การแจกแจงความถี่สะสมสัมพัทธ์

ความถี่สะสมสัมพัทธ์ (relative cumulative frequency) ของค่าที่เป็นไปได้ค่าใดหรืออันตรภาคชั้นใด คือ อัตราส่วนระหว่างความถี่สะสมของค่านั้น หรือของอันตรภาคชั้นนั้นกับผลรวมของความถี่ทั้งหมด ซึ่งอาจแสดงในรูปเศษส่วน ทศนิยม หรือร้อยละ ดังตัวอย่างที่ 5

ตัวอย่างที่ 5 จงสร้างตารางแจกแจงความถี่สะสมสัมพัทธ์ จากตารางแจกแจงความถี่ในตัวอย่างที่ 3  
วิธีทำ

คะแนน	ความถี่	ความถี่สะสม	ความถี่สะสมสัมพัทธ์	ร้อยละของความถี่สัมพัทธ์
10-20	7	7	0.07	7.0
20-29	13	20	0.20	20.0
30-39	18	38	0.38	38.0
40-49	22	60	0.60	60.0
50-59	26	86	0.86	86.0
60-69	10	96	0.96	96.0
70-79	4	100	1.00	100.0

ทั้งนี้ ตารางแจกแจงความถี่สะสมสัมพัทธ์ ใช้เพื่อหาว่าแต่ละค่าที่เป็นไปได้หรือแต่ละอันตรภาคชั้น มีความถี่สะสมเป็นจำนวนมากน้อยเพียงใดเมื่อเทียบกับความถี่ทั้งหมด เช่น จำนวนนักเรียน ที่ทำคะแนนสอบได้ต่ำกว่า 40 คะแนน คิดเป็นร้อยละ 38 ของนักเรียนที่เข้าสอบทั้งหมด

---