

บทที่ 2

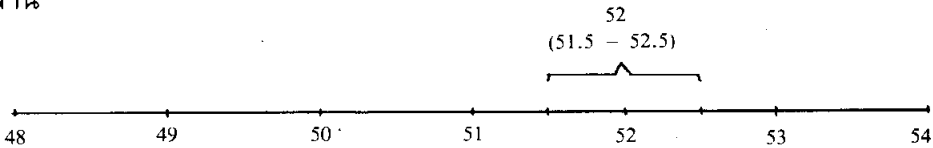
การแจกแจงความถี่

2.1 ลักษณะของคะแนน

ให้เราพิจารณาคะแนนผลการทดสอบวิชาสังคมสามค่าที่เรียงกันดังนี้ 51, 52, 53 เหตุผลที่เชื่อว่าคะแนน 52 ใช้เป็นตัวแทนได้ดีกว่าคะแนน 51 หรือ 53 คะแนน 52 อาจมีความหมายขยายออกไปอีกจาก 51.5 ถึง 52.5 นี้จะแปลความหมายของคะแนนได้จากรูปที่ 2.1 ก โดยทั่วไป การพิจารณาขอบเขตของคะแนนจะขยายออกได้จากการลดหรือเพิ่มขึ้นของหน่วยที่เล็กที่สุดของคะแนนที่วัดได้เศษหนึ่งส่วนสอง สมมติว่าเราวัดอะไรสักอย่างหนึ่งได้ 2.3 นิ้ว พิสัยที่ใช้แทนคะแนน 2.3 นิ้ว คือ 2.3 ± 0.05 นิ้ว หรือ 2.25 ถึง 2.35 นิ้ว สมมติว่าเราชั่งถ่านหินได้ 780 ปอนด์ น้ำหนักของถ่านหิน 780 ปอนด์แทนได้ด้วย 780 ± 5 ปอนด์ หรือจาก 775 ปอนด์ถึง 785 ปอนด์

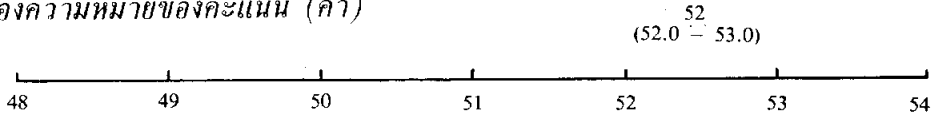
โอกาสการเกิดขึ้นของคะแนนที่มีความหมายแตกต่างไปจากที่กล่าวมาข้างต้นยากมาก อายุเป็นตัวอย่างที่มีธรรมชาติที่สุด เมื่อใครมีใครคนหนึ่งพูดว่าเขามีอายุ 52 ปี โดยธรรมชาติแล้ว เขาหมายความว่า เขาผ่านวันเกิดของเขาแล้วครั้งที่ 52 แต่เขายังไม่ถึงวันเกิดของเขาครั้งที่ 53 ในความหมายเหล่านี้ คะแนน 52 ใช้แทนช่วงจาก 52.0 ถึง 53.0 ดังรูป 2.1 ข

กลุ่มของคะแนนบางกลุ่มอาจมีค่าเป็นลบ ดังในบันทึกเกี่ยวกับอุณหภูมิ คะแนนที่มีค่าเป็นลบสามารถเกิดขึ้นได้ในการใช้ทดสอบคะแนนด้วยการลบหมายเลขของคำตอบที่ผิดออกจากหมายเลขที่ถูกได้ด้วย ในการพิจารณาการวัดระยะทาง การบันทึกค่าอาจจะสะดวกในรูปของค่าคลาดเคลื่อนที่ประมาณขึ้นมากเกินไปหรือน้อยเกินไป การใช้ค่าลบและบวกแสดงถึงค่าความแตกต่างเหล่านี้



ก. คะแนนทดสอบวิชาสังคม

รูปที่ 2.1 สองความหมายของคะแนน (ค่า)



ข. อายุของวันเกิดครั้งล่าสุด

2.2 ข้อมูลดิบ

ข้อมูลดิบหมายถึงข้อมูลที่ได้ถูกเก็บรวบรวมขึ้น แต่ข้อมูลเหล่านี้ยังมิได้มีการจัดเข้าไว้เป็นระเบียบหรือแก้ไขสิ่งบกพร่องบางอย่าง จะนำไปวิเคราะห์ทางสถิติไม่ได้ ดังตัวอย่าง เช่น ผลการสอบของนักศึกษา 50 คน ในวิชาเศรษฐศาสตร์ของแต่ละคนได้คะแนนดังตารางที่ 2.1

ตารางที่ 2.1

60, 33, 85, 52, 65, 77, 84, 65, 57, 74,
 71, 81, 35, 50, 35, 64, 74, 47, 68, 54,
 80, 41, 61, 91, 55, 73, 59, 53, 45, 77,
 41, 78, 55, 48, 69, 85, 67, 39, 76, 60,
 94, 66, 98, 66, 73, 42, 65, 94, 89, 88

จากตัวเลขข้างต้นทำให้นักศึกษาเกิดความยุ่งยากหรือลำบาก เมื่อต้องการจะเปรียบเทียบคนที่ได้คะแนนสูงสุดกับต่ำสุด ว่าแตกต่างกันเท่าไรหรือต้องการทราบว่าคะแนนส่วนใหญ่จะมีนักศึกษาสักกี่คน วิธีที่จะทำให้ตัวเลขเหล่านี้อยู่ในลักษณะที่อ่านง่ายขึ้น ถ้าได้รับการจัดเรียงตามขนาดมากน้อยเสียใหม่ คะแนนสูงสุดคือ 98 และต่ำสุด 33 การบันทึกลำดับของคะแนนได้เป็น 98,94, 91, 33 ให้เรากลับมาเรียงคะแนนที่ยังไม่ได้ลำดับและนำกลับมาจัดลำดับพร้อมด้วยความถี่ที่เกิดขึ้นแต่ละค่าดังตารางที่ 2.2 เรียกว่า การแจกแจงความถี่

ตารางที่ 2.2 คะแนนจากตารางที่ 2.1 มาจัดลำดับขนาด

คะแนน	ความถี่	คะแนน	ความถี่	คะแนน	ความถี่	คะแนน	ความถี่	คะแนน	ความถี่
98	1	84	1	70	0	56	0	42	1
97	0	83	0	69	1	55	2	41	2
96	0	82	0	68	1	54	1	40	0
95	0	81	1	67	1	53	1	39	1
94	2	80	1	66	2	52	1	38	0
93	0	79	0	65	3	51	0	37	0
92	0	78	1	64	1	50	1	36	0
91	1	77	2	63	0	49	0	35	2
90	0	76	1	62	0	48	1	34	0
89	1	75	0	61	1	47	1	33	1
88	1	74	2	60	2	46	0		
87	0	73	2	59	1	45	1		
86	0	72	0	58	0	44	0		
85	2	71	1	57	1	43	0		

การจัดระเบียบนี้มีหลาย ๆ ลักษณะที่สนใจมองเห็นได้ง่าย เราจะพบว่า คะแนน 65 จะอยู่ล่างศูนย์กลางของการแจกแจงเล็กน้อย ถึงแม้ว่าพิสัยของคะแนนจาก 33 ถึง 98 แต่ขนาดของการแจกแจงจะอยู่ระหว่าง 39 และ 94 จะมีนักศึกษาอยู่คนเดียวที่อยู่เหนือและสามคนอยู่ล่าง หกสิบห้าจึงไม่ได้อยู่ห่างไกลจากการดำเนินงานที่ดีในชั้นนี้

ในกรณีเช่นนี้ การจัดลำดับคะแนนก็เป็นการเพียงพอสำหรับวัตถุประสงค์ แต่ถ้าเมื่อไรการกระจายของคะแนนมีการกระจายมากก็จำเป็นจะต้องจัดคะแนนออกเป็นกลุ่มแทนการจัดลำดับคะแนน

2.3 การจัดคะแนนเป็นกลุ่ม

ในกรณีที่มีข้อมูลจำนวนมาก ๆ การจัดข้อมูลให้เรียงตามลำดับขนาด ก็มีข้อเสียอยู่หลายอย่าง เช่นเสียเวลาและเนื้อที่ ในบางครั้งการคำนวณหาค่าสถิติบางค่าก็ไม่จำเป็นจะต้องใช้ข้อมูลหรือตัวเลขจำนวนมาก (หรือทั้งหมด) แต่เราก็สามารถที่จะย่อ ข้อมูลทั้งหมดหรือจำนวนมาก ๆ ลงได้โดยมิให้เสียเนื้อความที่สำคัญกลับได้ประโยชน์ทั้งในด้านคำนวณทางสถิติและเข้าใจวิธีย่อข้อมูลหรือตัวเลขในเรื่องใดเรื่องหนึ่ง โดยทำข้อมูลให้เป็นกลุ่ม เป็นจำนวนหลาย ๆ ชั้น จัดเป็นรูปตารางที่เรียกว่า ตารางแจกแจงความถี่

ตารางที่ 2.3

การแจกแจงความถี่ของคะแนนนักศึกษา 50 คน

ชั้นของคะแนน	ใส่คะแนน	ความถี่
30-39	III	4
40-49	IIII	6
50-59	IIIIII	8
60-69	IIIIIIII	12
70-79	IIIIIIII	9
80-89	IIIIII	7
90-99	IIII	4

จากตารางที่ 2.3 ได้จัดข้อมูลเข้าเป็นกลุ่มหรือชั้น โดยให้ชั้นหนึ่ง ๆ มีช่วงกว้างเท่ากันหมด การแจกแจงจำนวนข้อมูลของแต่ละชั้นจะสังเกตได้จากช่องใส่คะแนน (tally) เพื่อบอกให้ทราบว่า ข้อมูลของแต่ละชั้นมีจำนวนหรือความถี่มากน้อยเท่าใด ตารางแจกแจงความถี่นี้อยู่ในลักษณะที่ รัดกุมและอ่านเข้าใจง่าย มีข้อดีและข้อเสียดังต่อไปนี้

ข้อดี

1. เมื่อรหัสของคะแนนกระจายกว้าง การเข้าใจความหมายอาจจะง่ายขึ้นเมื่อลดข้อมูล ให้เป็นจำนวนกลุ่มเล็ก ๆ
2. เมื่อไรที่จะต้องมีการคำนวณค่าอื่น ๆ โดยไม่ใช่เครื่องคำนวณการทำคะแนนเป็นกลุ่ม ทำให้งานคำนวณง่ายยิ่งขึ้น
3. การจัดคะแนนเป็นกลุ่มช่วยทำให้การแจกแจงความถี่ที่ไม่เป็นระเบียบให้เรียบได้ ถ้าหากว่าการแจกแจงมีแนวโน้มไปในรูปลักษณะเฉพาะทำให้สามารถมองเห็นได้ง่ายยิ่งขึ้น เมื่อ ทำคะแนนให้เป็นกลุ่ม
4. เมื่อไรที่ต้องทำการแจกแจงความถี่ไปเสนอด้วยกราฟ กราฟจะมีการสับสนน้อยกว่า เมื่อทำคะแนนเป็นกลุ่ม

ข้อเสีย

1. เมื่อไรที่ทำคะแนนให้เป็นกลุ่ม ข้อมูลข่าวสารบางอย่างจะสูญหายไป อย่างเช่น มี นักศึกษาได้คะแนนอยู่ในชั้น 60-69 12 คน แต่เราจะไม่ทราบว่านักศึกษา 12 คนนี้ได้คะแนนจริง ๆ เท่าไร อาจจะได้คะแนนเท่ากันหมดคือ 60 หรือ 61 หรือ 62 หรือ 69 คะแนน หรืออาจจะมีอยู่เพียง 2 คน ที่ได้คะแนน 60 อีก 10 คนได้คะแนน 61 หรือ 62 หรือ หรือ 69 คะแนนได้ ทั้งนั้น
2. กลุ่มของคะแนนดิบ (รูปของกลุ่มคะแนนรูปเดิม) ไม่มีผลต่อแบบของคะแนนที่จัดเป็นกลุ่ม ได้เพียงแบบเดียวเท่านั้น สามารถจัดได้หลายรูปแบบต่าง ๆ กัน

2.4 คุณลักษณะของอันตรภาคชั้น

ในการเปลี่ยนคะแนนดิบให้เป็นข้อมูลหรือคะแนนที่จัดเป็นกลุ่มนั้น มีอยู่หลายหลักการด้วยกัน ที่ควรจะต้องเก็บรักษาและพิจารณาเพื่อช่วยในการอ้างถึงตารางที่ 2.3

1. อัตราภาคชั้นควรจะไม่มีการกระจายหรือข้อมูลร่วมกัน นั่นคือ คะแนนหนึ่งหรือข้อมูลหนึ่งไม่สามารถอยู่ได้เกินหนึ่งอัตราภาคชั้น
2. ทุก ๆ อัตราภาคชั้นควรจะถูกกำหนด เมื่อไรที่อัตราภาคชั้นไม่เท่ากันจะเป็นเหตุให้งานสถิติอื่น ๆ เกิดความยุ่งยาก
3. ทุก ๆ อัตราภาคชั้นควรจะต้องเนื่องตลอดการแจกแจง
4. ขอบเขตของแต่ละอัตราภาคชั้นจะบอกถึงขอบเขตของคะแนนอย่างเช่น อัตราภาคชั้น 30-39 จะรวมถึงค่า 29.5-39.5 จำนวนค่าเหล่านี้เรียกว่าขอบเขตของชั้น ค่า 29.5 เรียกว่าขอบเขตของชั้นล่าง 39.5 เรียกว่าขอบเขตของชั้นบน สำหรับค่า 30 เรียกว่า ขีดจำกัดล่างปรากฏ ค่า 39 เรียกว่าขีดจำกัดบนปรากฏ การคำนวณค่าขอบเขตของชั้นได้ด้วยการเอาขีดจำกัดบนปรากฏของชั้นบวกกับขีดจำกัดล่างปรากฏของชั้นคะแนนมากกว่าแล้วหารด้วยสอง อย่างเช่น เราต้องการหาขอบเขตของชั้น 30-39 ขอบเขตของชั้นล่างคำนวณหาได้จาก $(29 + 30)/2 = 29.5$ สำหรับขอบเขตของชั้นบนคำนวณหาได้จาก $(39+40)/2 = 39.5$

2.5 วิธีดำเนินการแจกแจงความถี่

วิธีดำเนินการแจกแจงความถี่มีวิธีการที่จะเป็นหลักปฏิบัติได้ ดังนี้

1. หาพิสัย (R) ค่าของพิสัยหาได้จากการเอาข้อมูลหรือคะแนนที่มีค่ามากที่สุดลบด้วยข้อมูลหรือคะแนนที่มีค่าน้อยที่สุดนั่นคือ

$$\text{พิสัย (R)} = X_N - X_1$$

ในเมื่อ X_N = ค่าของข้อมูลหรือคะแนนที่มีค่ามากที่สุด

X_1 = ค่าของข้อมูลหรือคะแนนที่มีค่าน้อยที่สุด

2. หาความกว้างของชั้น ค่าความกว้างของชั้นหาได้จากเอาค่าพิสัยซึ่งหาได้จากข้อ 1 หารด้วยจำนวนชั้นซึ่งเรากำหนดเอาเองว่าจะใช้สักกี่ชั้น

$$I = \frac{R}{C}$$

I = ค่าความกว้างของชั้น (ปัดเป็นเลขจำนวนเต็มไม่ว่าเศษนั้นจะมีค่าน้อยกว่า 0.5)

C = จำนวนชั้นที่เรากำหนดขึ้น

3. หาขีดจำกัดล่างปรากฏของชั้นแรก โดยเอาค่า $(I \times C - R)/2$ นำไปลบออกจากคะแนนต่ำสุดของข้อมูล (ใช้ผลลัพธ์จำนวนเต็ม) หรืออาจใช้คะแนนต่ำสุดของข้อมูลก็ได้

$$\text{ขีดจำกัดล่างปรากฏของชั้นแรก} = X_1 - (I \times C - R)/2$$

เมื่อได้ค่าขีดจำกัดล่างของชั้นแรก เราก็สามารถสร้างตารางการแจกแจงความถี่ได้ดังตัวอย่าง 2.3 หากค่า $R = X_n - X_1 = 98 - 33 = 65$ ค่า C เรากำหนดให้มี 7 ชั้น หากค่า $I = \frac{R}{C} = \frac{65}{7} = 9.3$ เราปัดเป็นเลขจำนวนเต็มได้ 10 ต่อไปคำนวณหาขีดจำกัดล่างของชั้นแรกได้

$$\begin{aligned} L &= X_1 - (I \times C - R)/2 \\ &= 33 - (10 \times 7 - 65)/2 \\ &= 33 - 5/2 = 30.5 \end{aligned}$$

เราปัดเป็นเลขจำนวนเต็มได้เท่ากับ 30 จากนั้นเราสามารถสร้างตารางแจกแจงความถี่ได้

จุดกลางของชั้น (Class mark)

จุดกลางของชั้นเป็นจุดกึ่งกลางของความกว้างของแต่ละชั้น คำนวณหาได้จากการบวกของขีดจำกัดล่างปรากฏกับขีดจำกัดบนปรากฏของชั้นหารด้วยสอง จุดกลางของความกว้างของชั้น 30-39 คือ $(30+39)/2 = 34.5$ ในการวิเคราะห์ทางสถิติ จุดกึ่งกลางของชั้นใช้เป็นตัวแทนของจำนวนข้อมูลหรือคะแนนของชั้น ดังนั้น คะแนนที่จะเป็นตัวแทนของชั้น 30-39 ก็คือ 34.5

2.6 การแจกแจงความถี่สัมพัทธ์

ในกรณีที่ต้องการหาความถี่ของแต่ละชั้นให้เป็นความถี่ในรูปของสัดส่วนหรือเปอร์เซ็นต์ จึงจำเป็นต้องมีการแจกแจงความถี่สัมพัทธ์ ความถี่สัมพัทธ์ของชั้นใดชั้นหนึ่งหมายถึงความถี่ของชั้นนั้นหารด้วยความถี่รวมของชั้นทั้งหมด อย่างเช่น ความถี่สัมพัทธ์ของชั้น 30-39 ก็จะมีค่าเท่ากับ $4/50 = 0.08$ หรือ 8 เปอร์เซ็นต์ ผลบวกของความถี่สัมพัทธ์ของชั้นทั้งหมดมีค่าเท่ากับหนึ่งหรือ 100%

ประโยชน์ของความถี่สัมพัทธ์ช่วยในการเปรียบเทียบสองการแจกแจงความถี่ หรือมากกว่าในที่ซึ่งจำนวนความถี่ของแต่ละการแจกแจงไม่เท่ากัน แต่สำหรับการเปรียบเทียบความถี่สองกลุ่มไม่ใช่ของง่าย ถ้าหากว่าเปลี่ยนเป็นความถี่สัมพัทธ์ทำให้การแจกแจงทั้งสองอยู่บนฐานเดียวกันวิธีการเปรียบเทียบก็ง่ายขึ้น นอกจากนั้นความถี่สัมพัทธ์ยังสามารถนำไปคำนวณหาค่ามัธยฐานเลขคณิต ส่วนเบี่ยงเบนมาตรฐานและค่าฐานนิยม

การนำเสนอด้วยกราฟของการแจกแจงความถี่สัมพัทธ์สามารถหาค่าได้จาก แผนภูมิหรือหลายเหลี่ยมความถี่ โดยการเปลี่ยนสเกลในแนวตั้งจากความถี่เป็นความถี่สัมพัทธ์ รูปที่ได้จะอยู่ในลักษณะเดิมเรียกว่า แผนภูมิความถี่สัมพัทธ์หรือแผนภูมิเปอร์เซ็นต์และหลายเหลี่ยมความถี่สัมพัทธ์หรือหลายเหลี่ยมเปอร์เซ็นต์ตามลำดับ

ตารางที่ 2.4

คะแนน	ความถี่ (f)	ความถี่สัมพัทธ์ (r.f)	r.f. %
30-39	4	.08	8
40-49	6	.12	12
50-59	8	.16	16
60-69	12	.24	24
70-79	9	.18	18
80-89	7	.14	14
90-99	<u>4</u>	<u>.08</u>	<u>8</u>
	<u>50</u>	<u>1.00</u>	<u>100</u>

2.7 การแจกแจงความถี่สะสม

สมมติว่ามีใครคนหนึ่งถามว่า มีนักศึกษาได้เกรดสูงกว่าหรือเท่ากับ 60 อยู่กี่คน หรือมีนักศึกษาได้เกรดต่ำกว่าหรือเท่ากับ 40 อยู่กี่คน คำถามเช่นนั้นเกิดขึ้นบ่อย ๆ ที่มีความประสงค์จะรู้ความถี่สูงกว่าหรือเท่ากับ และความถี่ต่ำกว่าหรือเท่ากับค่าที่กำหนดให้ของตัวแปรค่า X ตัวอย่างเช่นใครคนหนึ่งมีความประสงค์หรืออยากจะทราบว่า มีพนักงานขายสักกี่คนสามารถขายได้มากกว่าจำนวนที่กำหนดของแผนกต่าง ๆ ในร้านขายสรรพสินค้า มีอยู่ที่แผนกขายได้จำนวนน้อยกว่าจำนวนที่กำหนด ปัญหาที่จะตอบนี้เราจะต้องจัดรวบรวมข้อมูลให้เป็นตารางความถี่สะสมและเส้นโค้งความถี่สะสมดังแสดงพร้อมด้วยการดำเนินงานอย่างง่าย

ตารางที่ 2.5
ตารางความถี่สะสม

	ชั้น	ความถี่	ความถี่สะสม (ต่ำกว่า)	ชั้น	ความถี่	ความถี่สะสม (สูงกว่า)
น้อยกว่า	39	4	4	มากกว่า	30	4
"	49	6	10	"	40	6
"	59	8	18	"	50	8
"	69	12	30	"	60	12
"	79	9	39	"	70	9
"	89	7	46	"	80	7
"	99	<u>4</u>	50	"	90	<u>4</u>
		50			50	

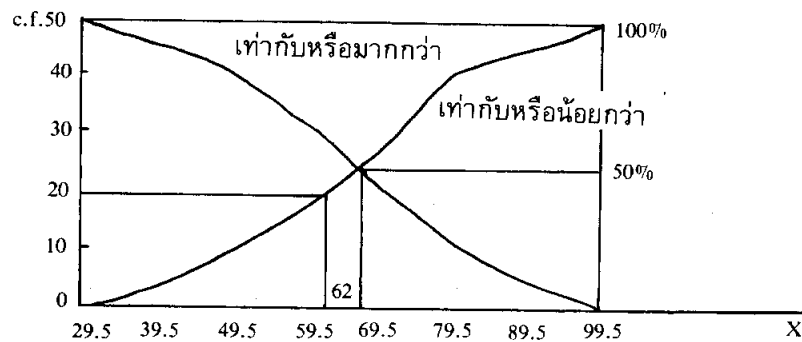
การสร้างการแจกแจงความถี่สะสม เราเริ่มด้วยการแจกแจงความถี่เสมอ ดังแสดงในคอลัมน์ที่สี่ (ในกรณีน้อยกว่า) คอลัมน์ที่แปด (ในกรณีมากกว่า) ของตารางที่ 2.5 เริ่มต้นที่ชั้นล่าง (กรณีมากกว่า) เรียบันทึกแต่ละอันตรภาคชั้นของความถี่รวมกรณีที่เกิดอยู่ต่ำกว่าขีดจำกัดบนของแต่ละอันตรภาคชั้น ตัวเลขเหล่านี้แสดงในคอลัมน์ที่แปดซึ่งคำนวณหาได้โดยการบวกความถี่ของอันตรภาคชั้นที่กำหนดให้เข้ากับความถี่สะสมที่ได้บันทึกสำหรับอันตรภาคชั้นที่ต่ำกว่าต่อไป การตรวจสอบจากการคำนวณ ความถี่สะสมสำหรับอันตรภาคชั้นสูงสุดควรเท่ากับ N ในกรณีน้อยกว่า เราเริ่มจากอันตรภาคชั้นบนสุดและใช้วิธีการแบบเดียวกันกับกรณีมากกว่า

ใช้ตาราง 2.3 มาสร้างตารางความถี่สะสมได้ดังตารางที่ 2.5 ข้างต้นในบางครั้งเราต้องการจะศึกษาถึงจำนวนของข้อมูลที่มีค่าต่ำกว่าหรือเท่ากับ และสูงกว่าหรือเท่ากับของขีดจำกัดแต่ละชั้น โดยเฉพาะ ตัวอย่างเช่นนักศึกษาต่ำกว่าหรือเท่ากับ 39 หมายถึงจำนวนนักศึกษาได้เกรดตั้งแต่ 39 ลงมามีอยู่ 4 คน จำนวนนักศึกษาได้เกรดต่ำกว่า 49 คะแนนมีอยู่ 10 คน เป็นต้น ในกรณีที่มีมากกว่าหรือเท่ากับ 30 คะแนน หมายความว่าจำนวนนักศึกษาซึ่งสอบได้คะแนนมากกว่าหรือเท่ากับ 30 อยู่ 50 คน หรือจำนวนนักศึกษามีคะแนนมากกว่าหรือเท่ากับ 40 หมายความว่า จำนวนนักศึกษาซึ่งสอบได้คะแนนมากกว่าหรือเท่ากับ 40 อยู่ 46 คน ค่าเหล่านี้ดูได้จากตารางที่ 2.5 ความถี่สะสมนี้สามารถนำไปหาค่ามัธยฐาน มาตรการวัดตำแหน่งต่าง ๆ ได้

เส้นโค้งความถี่สะสม

เราเอาค่าจากตารางความถี่สะสมมาเขียนกราฟ โดยให้แกนตั้งแสดงความถี่สะสม แกนนอนแสดงเกรดหรือคะแนน เราพลอตระหว่างคะแนนหรือเกรดกับความถี่สะสม (ต่ำกว่า) หรือระหว่างคะแนนหรือเกรดกับความถี่สะสม (สูงกว่า)

วิธีพลอต กราฟ ต่ำกว่าหรือเท่ากับของเส้นโค้งความถี่สะสม เราพลอตโดยใช้คะแนนของขอบเขตของชั้น 29.5, 39.5, ... 99.5 กับความถี่สะสม (ต่ำกว่า) 4, 10, ... 46 และ 50 ลากเส้นเชื่อมโยงระหว่าง Coordinates ก็จะได้เส้นโค้งความถี่สะสม (ต่ำกว่า) ดังรูปที่ 2.2 ในทำนองเดียวกัน ถ้าเราพลอต เส้นโค้งความถี่ระหว่างค่า ขอบเขตของชั้นกับความถี่สะสม (สูงกว่า) ก็จะได้เส้นโค้งความถี่สะสมตามต้องการ ดูรูปที่ 2.2



รูปที่ 2.2

2.8 การนำเสนอข้อมูล

ข้อมูลที่เก็บรวบรวมได้ เมื่อได้ตรวจสอบความถูกต้อง หรือข้อผิดพลาดอื่น ๆ ซึ่งเกิดจากความเลินเล่อของเจ้าหน้าที่สำรวจและผู้ตอบ ข้อมูลส่วนมากมักจะเป็นประโยชน์ต่อรัฐบาล เพราะฉะนั้น รัฐบาลจึงเป็นผู้เก็บรวบรวมและนำเสนอต่อประชาชน การนำเสนอมีอยู่หลายวิธี คือ

1. การนำเสนอด้วยตาราง แบบฟอร์มตารางที่จะนำเสนอประกอบด้วย
 - ก. ชื่อเรื่อง จะต้องแสดงว่าเป็นเรื่องเกี่ยวกับอะไร ที่ไหน เมื่อไร
 - ข. หัวเรื่อง เป็นคำอธิบายเกี่ยวกับตัวเลข ในแนวตั้งอาจจะมีได้หลายหัวข้อตามจำนวนที่นำเสนอ
 - ค. หัวข้อ เป็นคำอธิบายเกี่ยวกับตัวเลขในแนวนอน อาจจะมีได้หลายหัวข้อเช่นเดียวกับหัวเรื่อง
 - ง. ตัวเรื่อง ประกอบด้วยตัวเลขที่นำเสนอ
 - จ. หมายเหตุแหล่งที่มาของตัวเลข ว่าตัวเลขในตารางได้มาจากที่ใด
 - ฉ. หน่วย ตัวเลขในตารางควรกำกับหน่วยไว้ทุกครั้ง

ตารางที่ 2.6

ประชากรของประเทศสหรัฐอเมริกา } ชื่อเรื่อง

(หน่วย : ล้านคน) } หน่วย

ปี {	ปี ค.ศ.	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	หัวเรื่อง
	ประชากร	17.1	23.2	31.4	39.8	50.2	62.9	76.0	92.0	105.7	122.8	131.7	151.1	179.3	ตัวเรื่อง

ที่มา : Bureau of the Census

การนำเสนอด้วยตารางจะใช้แบบไม่ตีเส้นก็มีคนนิยมใช้กันมาก ขึ้นอยู่กับความนิยมหรือลักษณะงานที่จะนำเสนอ ดังเช่น ตารางที่ 2.7

ตารางที่ 2.7
รายชื่อยานพาหนะอย่างย่อ
(หน่วย 1,000 คัน)

รายชื่อ	1950	1955	1958
รถพ่วงชนิดโรงงาน	6666	7920	4258
รถแทรกและรถโดยสารโรงงาน	1337	1249	877
ยานพาหนะที่ใช้ทำความสะอาด	3234	4392	3173

ที่มา: กรมการค้าของประเทศสหรัฐ

ตารางที่ 2.8
จำนวนผลิตผลของข้าวสาลีและข้าวโพดจากฟาร์ม PQR
ในปี 1950 - 1960 คิดเป็นจำนวนบุชเชล

ปี	จำนวนบุชเชล ของข้าวสาลี	จำนวนบุชเชล ของข้าวโพด
1950	200	75
1951	185	90
1952	225	100
1953	250	85
1954	240	80
1955	195	100
1956	210	110
1957	225	105
1958	250	95
1959	230	110
1960	235	100

ที่มา: ฟาร์ม PQR

ตารางที่ 2.9

แสดงถึงสถานะของเพศชายและหญิงในสหรัฐของ ปี 1958 คิดเป็นเปอร์เซ็นต์

สถานะ	เพศชาย %	เพศหญิง %
โสด	24.5	18.8
สมรส	69.8	66.0
เป็นหม้าย	3.9	12.8
หย่าร้าง	1.8	2.3

ที่มา: สำนักงานสำมะโนประชากร (สหรัฐ)

2. การนำเสนอด้วยกราฟ การนำเสนอด้วยวิธีนี้เป็นกรนำเสนอที่ทำให้ผู้อ่านเข้าใจความหมายของข้อมูลได้ดีและรวดเร็วกว่าการนำเสนอวิธีอื่น ๆ เพราะผู้อ่านสามารถมองเห็นได้ทันที การเสนอด้วยกราฟที่นิยมใช้กันมาก มีแผนภูมิแท่ง, แผนภาพวง, รูปภาพ, แผนที่และเส้นกราฟ การนำเสนอดังกล่าว อาจถือได้ว่าเป็นการเปรียบเทียบอย่างใดอย่างหนึ่ง 1 ใน 3 อย่าง ดังเช่น

1. เปรียบเทียบระหว่างขนาดเพียงอย่างเดียว (รูปที่ 2.15)
2. เปรียบเทียบขนาดในคาบเวลาต่างกัน (รูปที่ 2.4, 2.12 - 2.14)
3. เปรียบเทียบอัตราการเปลี่ยนแปลงของขนาด การนำเสนอแบบนี้มีข้อเสียอยู่บ้าง การให้รายละเอียดได้น้อยกว่าตารางและการสร้างกราฟเพื่อนำเสนอกีใช้เวลามาก ข้อมูลที่จะนำมาเปรียบเทียบโดยใช้กราฟก็มีจำนวนน้อยชุดด้วยกัน

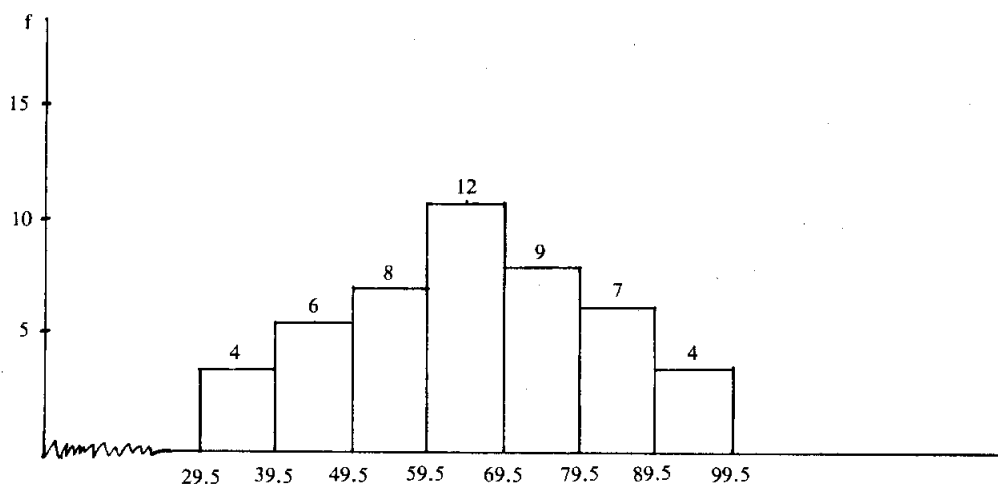
2.9 หลักการนำเสนอ

การนำเสนอด้วยกราฟ มีหลักสำคัญที่ควรยึดถือได้คือ

1. มาตรฐาน ควรเลือกให้อยู่ในขนาดที่เหมาะสมเพื่อได้ภาพที่ถูกต้องตามข้อเท็จจริง
2. มาตรฐานแกนตั้ง (แกน y) ควรเริ่มต้นจากค่าศูนย์ ในบางกรณีที่ต้องเริ่มจากค่าอื่น จะต้องทำเครื่องหมายบางอย่างให้ผู้อ่านเข้าใจ
3. ภาพนำเสนอ ไม่ควรมีเส้นนำมากเกินความจำเป็น
4. แผนภูมิหรือเส้นกราฟที่นำเสนอจะต้องมีชื่อเรื่องบอกกำกับไว้

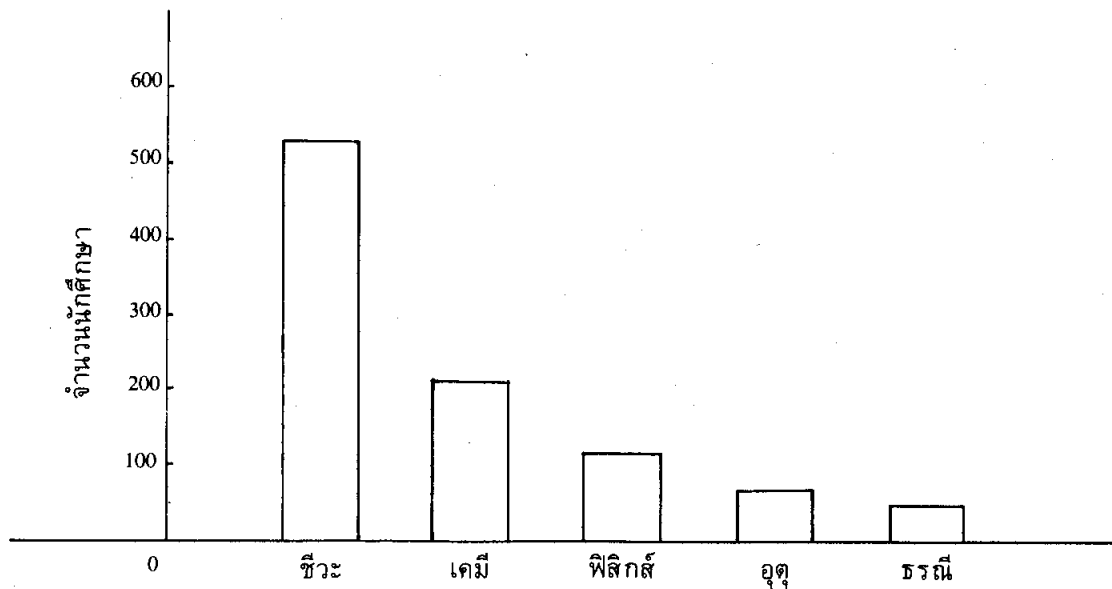
5. การนำเสนอด้วยเส้นกราฟ ควรลากเส้นกราฟที่แสดงข้อมูลให้หนักกว่าเส้นอื่น
6. เส้นกราฟแสดงข้อมูลมีจำนวนมากกว่า 1 เส้น ควรเขียนชื่อกำกับเส้นกราฟทุกเส้น เพื่อบอกให้รู้ว่าแทนข้อมูลอะไร
7. ที่มาของตัวเลขข้อมูล ควรทำหมายเหตุไว้ทุกครั้ง

ก. แผนภูมิ กราฟแบบนี้ใช้นำเสนอการแจกแจงความถี่อันประกอบด้วยอนุกรมของแท่งสี่เหลี่ยมผืนผ้า แต่ละแท่งใช้แทนความถี่ของคะแนนในชั้นหนึ่งของตารางการแจกแจง อย่างเช่น ตารางที่ 2.3 ใช้แทนการแจกแจงความถี่ของคะแนนนักศึกษา 50 คน ข้อมูลเหล่านี้นำมาเขียนเป็นกราฟในรูปของแผนภูมิ ดังรูปที่ 2.3



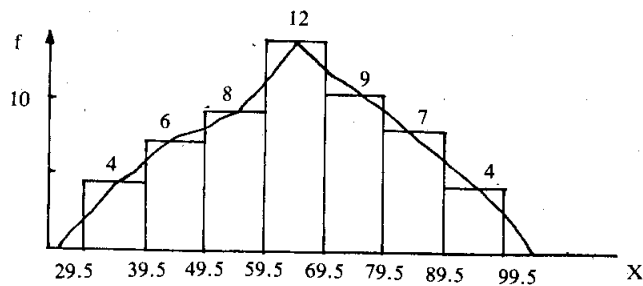
รูปที่ 2.3

ข. แผนภาพแท่ง ในกรณีข้อมูลแบ่งออกเป็นประเภท แผนภูมิจึงใช้เสนอได้ แต่แผนภาพแท่งใช้ได้ดีกว่า แผนภาพแท่งมีลักษณะคล้ายคลึงกับแผนภูมิมาก เว้นแต่ว่ามีช่องว่างอยู่ระหว่างแท่งสี่เหลี่ยมผืนผ้า แผนภาพแท่งมีข้อดีในการเสนอข้อมูลที่ไม่ต่อเนื่อง หรือข้อมูลชนิดลำดับความกว้างของแต่ละแท่งเท่ากันหมด



รูปที่ 2.4 แผนภาพแท่ง เปรียบเทียบความถี่ของนักศึกษาวิชาเอก คณะวิทยาศาสตร์

ค. หลายเหลี่ยมความถี่ ใช้ข้อมูลชุดเดียวกันกับการเสนอแบบแผนภูมิในรูปที่ 2.3 มาเสนอแบบหลายเหลี่ยมความถี่ กราฟชนิดนี้เสนอได้โดยพลอตจุดบนจุดกึ่งกลางของความกว้างของแต่ละชั้น ลากเส้นเชื่อมโยงจุดกลางด้วยเส้นตรง เราก็จะได้รูปหลายเหลี่ยมความถี่ดังรูปที่ 2.5 ดังตัวอย่าง จุดกลางของความกว้างของชั้น 40-49 คือ $(40 + 49)/2 = 44.5$



รูปที่ 2.5

หมายเหตุ ในการลากเส้นเชื่อมจุดกึ่งกลางของความกว้างของชั้น 30-39 และ 20-29 แม้ว่าไม่มีค่าความถี่ในความกว้างของชั้น 20-29 ก็ได้ลากไปยังจุดกึ่งกลางของความกว้างของชั้น 20-29 ในทำนองเดียวกันความกว้างของชั้นสุดท้ายก็ทำเช่นเดียวกัน

ง. เส้นโค้งความถี่ (Frequency curve)

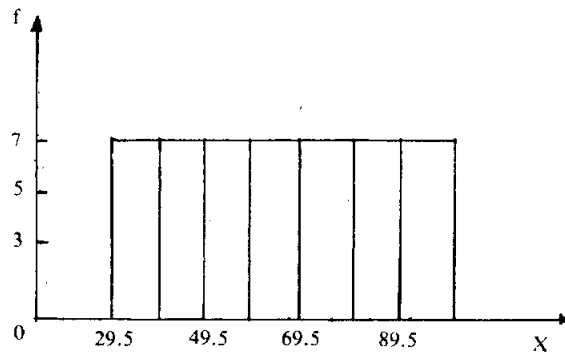
เส้นโค้งความถี่ เป็นการนำเสนอด้วยกราฟในทางทฤษฎีเท่านั้น ทางปฏิบัติไม่ค่อยเกิดขึ้น เช่น สมมุติว่าข้อมูลที่มีการแจกแจงความถี่จำนวนมากและจัดเข้าเป็นชั้นเล็ก ๆ มากชั้น หลายเหลี่ยมความถี่ที่สร้างจากข้อมูลจะให้เส้นโค้งเรียบขึ้นแต่เรียบเหมือนเส้นโค้งความถี่ไม่ได้ นอกจากเราจะต้องเกลกลายเหลี่ยมความถี่ให้เป็นเส้นโค้งความถี่โดยไม่ให้เนื้อที่ภายใต้เหลี่ยมเหลี่ยมความถี่เสียไป

เส้นโค้งความถี่ มีชื่อต่าง ๆ กันตามชนิดของการแจกแจงความถี่

การแจกแจงแบบสี่เหลี่ยมผืนผ้า ถ้าหากเกรดของนักศึกษาเป็นดังตารางที่ 2.10 และการแจกแจงก็จะเป็นอย่างรูปที่ 2.6 ลักษณะของการแจกแจงเป็นแบบสี่เหลี่ยมผืนผ้าส่วนใหญ่ใช้ในทางทฤษฎี

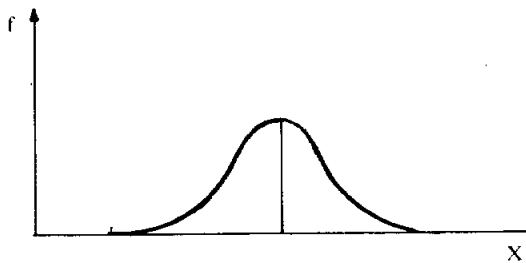
ตารางที่ 2.10

ชั้น	f
30 - 39	7
40 - 49	7
50 - 59	7
60 - 69	7
70 - 79	7
80 - 89	7
90 - 99	7
	<u>49</u>



รูปที่ 2.6

การแจกแจงแบบปกติ การแจกแจงแบบปกติเป็นการแจกแจง แบบสมมาตรรอบมัธยฐาน เลขคณิตหรือแบบรูปประฆังตามรูปที่ 2.7 การแจกแจงแบบนี้มีความสำคัญมากที่จะนำไปใช้วิเคราะห์ในขั้นต่อไป

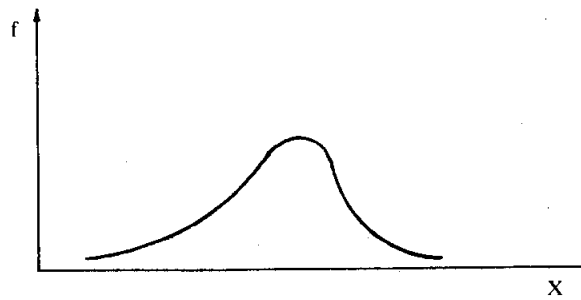


รูปที่ 2.7

การแจกแจงแบบรูปประฆังเบ้ ถ้าเกรตถูกแจกแจงอย่างไม่สมมาตรรอบมัชฌิมเลขคณิต การแจกแจงตารางที่ 2.11 จะให้การแจกแจงแบบรูปประฆังเบ้ดังรูป 2.8

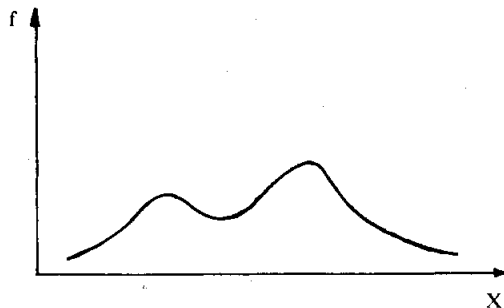
ตารางที่ 2.11

	f
30 - 39	1
40 - 49	3
50 - 59	6
60 - 69	10
70 - 79	12
80 - 89	15
90 - 99	3
	<u>50</u>



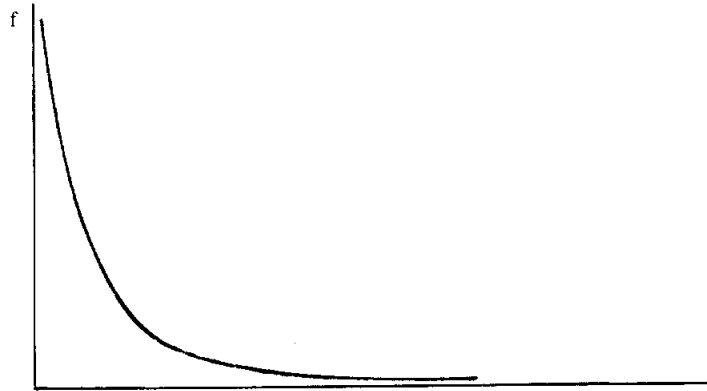
รูปที่ 2.8

การแจกแจงจะเบ้ไปทางซ้ายหรือขวาก็ขึ้นอยู่กับข้อมูลที่เกี่ยวข้องกับเรื่องอะไรเป็นสำคัญ เช่น จำนวนของฟาร์มจัดตามขนาดในสหรัฐอเมริกาจะมีการแจกแจงที่เบ้ไปทางขวา ปริมาณของจำนวนร้านขายปลีกย่อยในสหรัฐก็มีการแจกแจงที่เบ้ไปทางขวา การแจกแจงเกรดในโรงเรียนกฎหมายมีลักษณะเบ้ไปทางซ้าย



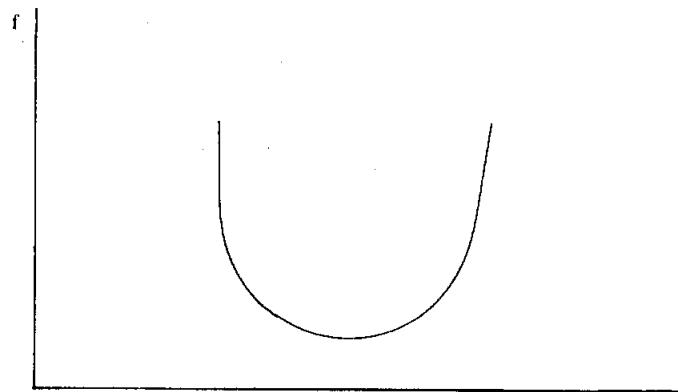
รูปที่ 2.9

การแจกแจงแบบสองฐานนิยม ในบางกรณีที่ค่าของข้อมูลที่เกิดบ่อยครั้งมีมากกว่าหนึ่งค่า เช่น รูปที่ 2.9 การแจกแจงแบบนี้โดยมากเกิดจากข้อมูลที่ไม่ใช่ชนิดเดียวกันหรือเกิดจากการนำเอาการแจกแจงที่ต่างกัน 2 อัน มารวมกันเข้า หรือข้อมูลที่ใช้มีอยู่น้อยจำนวนเกินไป ตัวอย่าง เช่น การแจกแจงของเกรดในวิชาสถิติโดยจัดเป็นกลุ่มที่เรียงดีกับเรียงไม่ดี จะได้รับการแจกแจงแบบสองฐานนิยม



รูปที่ 2.10

การแจกแจงแบบรูปตัว J กลับ ดังรูป 2.10

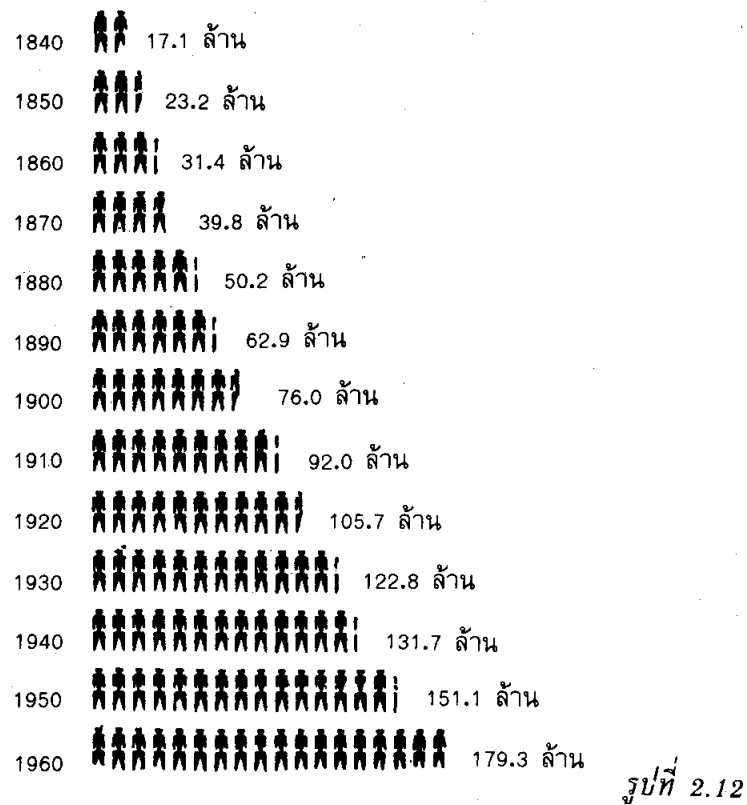


รูปที่ 2.11

การแจกแจงแบบรูปตัว U ดังรูป 2.11

การแจกแจงความถี่ทั้ง 8 รูป ในทางสถิติการแจกแจงแบบปกติ ดังรูป 2.7 จัดว่ามีความสำคัญมากที่สุด

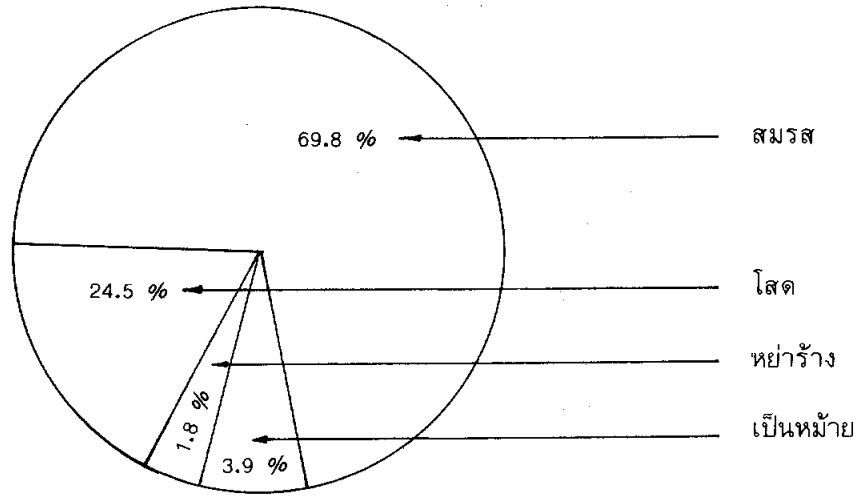
จ. กราฟแบบรูปภาพ เป็นการนำเสนอข้อมูลออกมาเป็นรูปภาพของข้อมูลที่ได้มาและเป็นกราฟเรียกร้องความสนใจจากผู้อ่านได้มาก ผู้อ่านเข้าใจความหมายได้ทันที



ที่มา: Bureau of the Census

รูปที่ 2.12 เป็นแผนภาพ ที่เรียกว่า กราฟรูปภาพหรือแผนภาพรูปภาพ เป็นกราฟเรียกร้องความสนใจจากผู้อ่านได้มากและผู้อ่านเข้าใจความหมายได้ทันที เราใช้มาตราส่วน 10,000,000 คนต่อรูปหนึ่ง

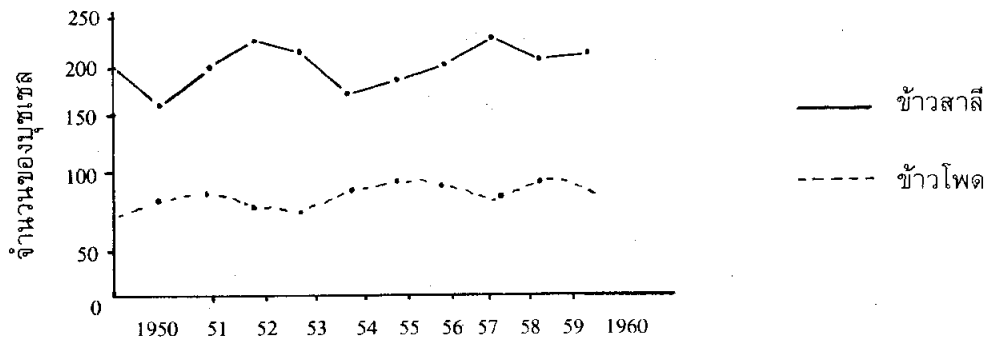
ฉ. แผนภาพวง เป็นการเสนอข้อมูลที่แบ่งออกเป็นประเภทต่าง ๆ ในรูปวงกลม ซึ่งแสดงการเปรียบเทียบอัตราเป็นเปอร์เซ็นต์



ที่มา: สำนักงานสำมะโนประชากร (สหรัฐ)

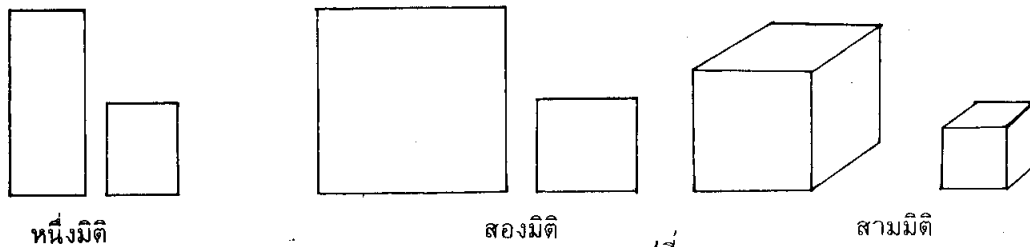
รูปที่ 2.13

ข. กราฟเส้น เป็นการนำเสนอข้อมูลชนิดเดียวกันออกมาเป็นเส้น



รูปที่ 2.14

การเปรียบเทียบขนาด



รูปที่ 2.15

แบบฝึกหัด

1. นำนักคิดเป็นปอนด์ของนักศึกษาชาย 40 คน ที่มหาวิทยาลัยแห่งหนึ่งในสหรัฐมีดังนี้

138	164	150	132	144	125	149	157
146	158	140	147	136	148	152	144
168	126	138	176	163	119	154	165
146	173	142	147	135	153	140	135
161	145	135	142	150	156	145	128

จงสร้างตารางการแจกแจงความถี่

2. จงสร้าง (ก) แผนภูมิ (ข) หลายเหลี่ยมความถี่ จากตารางการแจกแจงน้ำหนักในข้อหนึ่ง
3. จงสร้างเส้นโค้งความถี่สะสมทั้งชนิดน้อยกว่าหรือเท่ากับและชนิดมากกว่าหรือเท่ากับจากข้อหนึ่ง
4. อายุของคนงานแบ่งออกเป็นกลุ่มหรือแจกแจงตามขีดจำกัดของชั้น 20-29, 30-39, 40-49, 50-59 และ 60-69 ปี จงหาค่าของ (ก) ขอบเขตของชั้น (ข) จุดกลาง และ (ค) ความกว้างของชั้น
5. จุดกลางของการแจกแจงราคาอาหารย่อย ๆ เป็น 15, 22, 29, 36 และ 43 เซ็นต์ จงหาขีดจำกัดของชั้นแต่ละชั้น