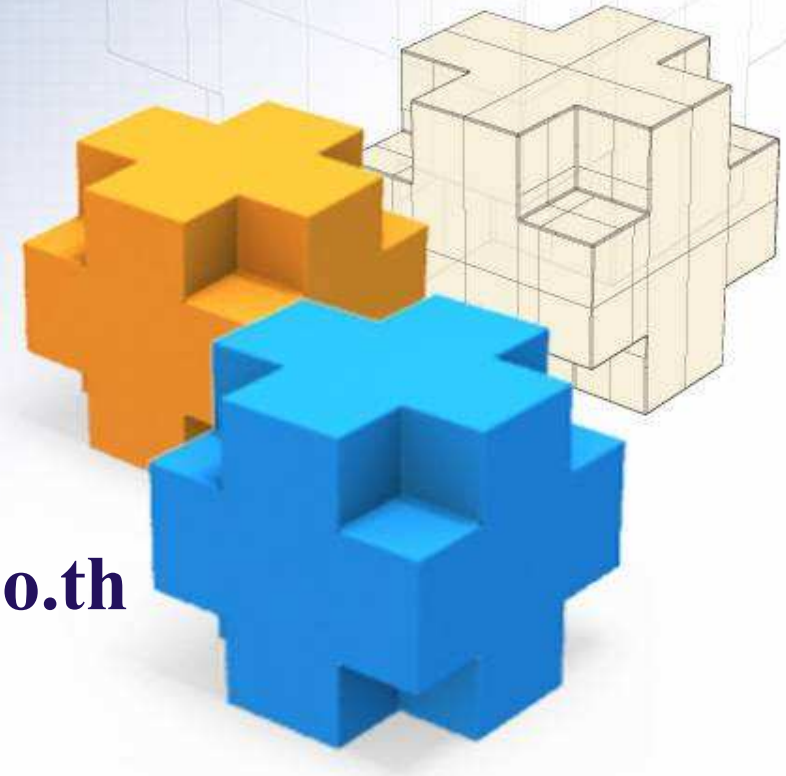
การแจกแจงความถี่

(Frequency)

อ.จันทร์เพ็ญ ไชยนุ้ย

Keawwan_02@hotmail.co.th



รายละเอียดของเนื้อหา

1. การแจกแจงความถี่ (Frequency)
2. สัดส่วน (Proportion)
3. ร้อยละ (Percentage)

การแจกแจงความถี่ (Frequency)

การแจกแจงความถี่ หมายถึง การนำข้อมูลดิบ (Raw Data) ซึ่งเป็นข้อมูลที่เก็บรวบรวมได้มาเรียงลำดับจากค่าน้อยที่สุดไปหาค่ามากที่สุด หรือเรียงลำดับจากค่ามากที่สุดไปหาน้อยที่สุดก็ได้ เพื่อแสดงให้ทราบว่าข้อมูลแต่ละค่าหรือแต่ละกลุ่มเกิดขึ้นซ้ำ ๆ กันกี่ครั้ง จำนวนข้อมูลที่ซ้ำ ๆ กันนี้เรียกว่า **"ความถี่ (Frequency)"**

การแจกแจงความถี่มี 2 วิธี คือ

1. การแจกแจงความถี่ของข้อมูลแบบไม่จัดเป็นกลุ่ม (Ungrouped Data)

2. การแจกแจงความถี่ของข้อมูลแบบจัดเป็นกลุ่ม (Grouped Data)

1. การแจกแจงความถี่ของข้อมูลแบบไม่จัดเป็นกลุ่ม (Ungrouped Data)

เป็นการนำข้อมูลทั้งหมดที่เก็บรวบรวมได้มาเรียงลำดับจากค่าน้อยที่สุดไปหาค่ามากที่สุด หรือเรียงลำดับจากค่ามากที่สุดไปหาค่าน้อยที่สุด แล้วพิจารณาข้อมูลแต่ละค่านั้นมีค่าซ้ำกันมากน้อยเพียงใด ซึ่งวิธีนี้เหมาะกับข้อมูลที่ค่าสูงสุดและค่าต่ำสุดของข้อมูลไม่แตกต่างกันมาก

วิธีการสร้างตารางแจกแจงความถี่ของข้อมูลแบบไม่จัดเป็นกลุ่ม มีขั้นตอนดังนี้

1. หาค่าของข้อมูลที่มีค่าต่ำสุดและค่าสูงสุด
2. เรียงลำดับข้อมูลใหม่จากค่าน้อยที่สุดไปหาค่ามากที่สุด หรือจากค่ามากที่สุดไปหาค่าน้อยที่สุด ค่าใดที่ไม่ปรากฏในข้อมูลก็ไม่ต้องบันทึก สำหรับข้อมูลที่ซ้ำกันให้บันทึกเพียงครั้งเดียว
3. พิจารณาดูข้อมูลค่าหนึ่ง ๆ เกิดขึ้นกี่ครั้งให้ขีดในช่องรอยขีด (Tally)
4. นับจำนวนรอยขีดแล้วรวมใส่ในช่องความถี่ของข้อมูล และเมื่อรวมความถี่ของข้อมูลแล้วต้องเท่ากับจำนวนข้อมูลทั้งหมด

ตัวอย่างที่ 1 ในการทดสอบวิชาหลักสถิติกับนักศึกษาวิทยาลัยประมงปัตตานี 20 คน คะแนนเต็ม 20 คะแนน ปรากฏผลการทดสอบดังนี้

8	9	8	6
6	5	9	7
5	8	7	8
6	9	8	6
8	9	10	7

จงสร้างตารางแจกแจงความถี่ของข้อมูลแบบไม่จัดเป็นกลุ่ม

วิธีทำ

1. ข้อมูลที่มีค่าต่ำสุด คือ 5 และข้อมูลที่มีค่าสูงสุดคือ 10

2. เรียงลำดับข้อมูลใหม่ ในที่นี้เรียงลำดับข้อมูลจากค่าน้อยที่สุดไปหาค่ามากที่สุด ได้ดังนี้ 5, 6, 7, 8, 9, 10

3. พิจารณาดูว่าข้อมูลค่าหนึ่ง ๆ เกิดขึ้นกี่ครั้งให้ขีดในช่องรอยขีด (tally)

4. นับจำนวนรอยขีดแล้วใส่ในช่องความถี่ (f)

ตารางที่ 1 แสดงการแจกแจงความถี่ของคะแนนการทดสอบวิชาหลักสถิติของนศ.

คะแนน	รอยขีด	ความถี่ (f)
5	//	2
6	////	4
7	///	3
8	~~////~~ /	6
9	////	4
10	/	1
	รวม	20

2. การแจกแจงความถี่ของข้อมูลแบบจัดเป็นกลุ่ม (Grouped Data)

เป็นการนำข้อมูลดิบมาจัดเป็นหมวดหมู่ ในแต่ละชั้นจะมีข้อมูลหลายค่าแทนที่จะเป็นค่าเดียว วิธีการสร้างตารางแจกแจงความถี่ของข้อมูลแบบจัดเป็นกลุ่ม มีขั้นตอนดังนี้

1. หาค่าพิสัยของข้อมูล (Rang; R) โดยที่ พิสัยคือ ผลต่างของข้อมูลที่มีค่าสูงสุดกับข้อมูลที่มีค่าต่ำสุด

พิสัย (R) = ข้อมูลที่มีค่าสูงสุด – ข้อมูลที่มีค่าต่ำสุด

2. จำนวนชั้นของข้อมูล (k) หมายถึง จำนวนชั้นหรือจำนวนกลุ่มของข้อมูลที่จะถูกจัดในตาราง โดยทั่วไปจะกำหนด 5-20 ชั้น ทั้งนี้จำนวนชั้นขึ้นอยู่กับช่วงกว้างของข้อมูล

ในกรณีที่ไม่ทราบว่าควรจะกำหนดชั้นกี่ชั้น อาจคำนวณจำนวนชั้น (k) ได้จากสูตร

$$k = 1 + 3.3 \log N$$

โดยที่

k = จำนวนชั้น

N = จำนวนข้อมูลทั้งหมด

3. ความกว้างของอันตรภาคชั้น (Class Interval ; I) คือ ช่วงห่างของข้อมูลในแต่ละชั้น โดยใช้สูตรดังนี้

$$\text{ความกว้างของอันตรภาคชั้น (I)} = \frac{\text{พิสัย (R)}}{\text{จำนวนชั้น (I)}}$$

โดยที่ ค่าความกว้างของอันตรภาคชั้น <u>*จะต้องเป็นจำนวนเต็มเสมอ*</u>

3.1 ถ้าคำนวณความกว้างของอันตรภาคชั้นเป็นทศนิยมให้ปัดเศษขึ้นเป็นจำนวนเต็ม ไม่ว่า*เศษจะมีค่าต่ำกว่าหรือมากกว่า 0.5*

เช่น พิสัย = 22

จำนวนชั้น = 5

ความกว้างของอันตรภาคชั้น $= \frac{22}{5} = 4.4$

≈ 5

3.2 ถ้าคำนวณความกว้างของอันตรภาคชั้นเป็นเลขจำนวนเต็มจะบวกด้วยหนึ่ง

เช่น พิสัย = 20

จำนวนชั้น = 5

ความกว้างของอันตรภาคชั้น $= \frac{20}{5} = 4+1$

$= 5$

4. เขียนขีดจำกัดชั้น (Class Limit) เป็นค่าที่บอกว่าแต่ละชั้นจะมีตัวเลขอยู่ในชั้นนั้น ๆ บ้าง โดยเริ่มจากชั้นของข้อมูลต่ำสุดไปหาข้อมูลสูงสุด หรือจากชั้นของข้อมูลสูงสุดไปหาชั้นของข้อมูลต่ำสุดก็ได้ ซึ่งแบ่งเป็น

- ขีดจำกัดบน (Upper Limits) หมายถึง ค่าสูงสุดของแต่ละชั้น
- ขีดจำกัดล่าง (Lower Limits) หมายถึง ค่าต่ำสุดของแต่ละชั้น

เช่น ข้อมูลที่มีค่าต่ำสุด คือ 40 เป็นขีดจำกัดล่างของชั้นแรก และความกว้างของอันตรภาคชั้น = 5 ดังนั้นขีดจำกัดบนของชั้นแรกคือ 40 + 5 – 1 = 44 นั่นคือ ช่วงของข้อมูลชั้นแรก คือ 40 – 44

5. พิจารณาว่าข้อมูลแต่ละค่าอยู่ในช่วงข้อมูลใด ให้ขีดในช่องรอยขีด

6. นับจำนวนรอยขีดแล้วใส่ในช่องความถี่ (f)

ตัวอย่างที่ 2 คนงานช่างหล่อของโรงหล่อเหล็กแห่งหนึ่งมีจำนวน 60 คน ได้ ค่าจ้างรายวัน ดังนี้

165	153	148	157	142	168	174	151	156	158
151	148	156	154	164	161	158	159	143	149
158	145	149	166	157	154	162	159	153	164
161	165	152	157	148	145	170	165	165	168
154	148	153	142	167	162	151	172	140	157
163	168	147	150	151	157	153	167	164	147

จงสร้างตารางแจกแจงความถี่ของข้อมูลแบบจัดเป็นกลุ่ม

วิธีทำ

1. หาค่าพิสัย (Rang; R) = ค่าสูงสุด – ค่าต่ำสุด

=

=

2. จำนวนชั้น (k) = 1 + 3.3 log N

=

=

=

=

วิธีทำ

3. หาความกว้างของอันตรภาคชั้น (Class Interval; I) $= \frac{\text{พิสัย (R)}}{\text{จำนวนชั้น (k)}}$

=

=

=

4. เขียนขีดจำกัดชั้น (Class Limit)

- ขีดจำกัดล่างของชั้นแรก =

- ขีดจำกัดบนของชั้นแรก =

จากนั้นพิจารณาว่าข้อมูลแต่ละค่าอยู่ในช่วงข้อมูลใด ให้ขีดในช่องรอยขีด นับจำนวนรอยขีดแล้วใส่ในช่องความถี่ (f) ดังตารางที่ 2

ตารางที่ 2 แสดงการแจกแจงความถี่ของค่าจ้างรายวันของคนงานช่างหล่อจำนวน 60 คน

ค่าจ้าง (บาท)	รอยขีด	ความถี่ (f)
รวม		

5. ขอบเขตจำกัดของชั้น (Class Boundaries) เป็นค่าของขีดจำกัดชั้นที่แท้จริง หมายถึงค่าแต่ละชั้นที่มีค่าต่อเนื่องกัน คำนวณได้จากสูตร

$$\text{ขอบเขตจำกัดชั้นบน} = \frac{\text{ขีดจำกัดบนของชั้นนั้น} + \text{ขีดจำกัดล่างของชั้นถัดไป}}{2}$$

$$\text{ขอบเขตจำกัดชั้นล่าง} = \frac{\text{ขีดจำกัดล่างของชั้นนั้น} + \text{ขีดจำกัดบนของชั้นถัดไป}}{2}$$

ค่าของขอบเขตจำกัดชั้น จะมีจำนวนทศนิยมมากกว่าจำนวนทศนิยมของข้อมูลเดิมหนึ่งตำแหน่งเสมอ

6. จุดกึ่งกลาง (Mid point) หมายถึง ค่าที่อยู่กึ่งกลางระหว่างขีดจำกัดบนและขีดจำกัดล่าง ซึ่งคำนวณได้จากสูตร

$$\text{จุดกึ่งกลาง} = \frac{\text{ขีดจำกัดล่าง} + \text{ขีดจำกัดบน}}{2}$$

เช่น ช่วงค่าของข้อมูล = 140 - 144

$$\text{จุดกึ่งกลางชั้น} = \frac{140 + 144}{2}$$

$$= 142$$

ตารางที่ 3 แสดงค่าจ้างรายวันของคนงานช่างหล่อจำนวน 60 คน

ค่าจ้าง (บาท)	ขอบเขตของชั้น	จุดกึ่งกลาง	ความถี่ (f)
140 -144			
145-149			
150- 154			
155-159			
160-164			
165-169			
170-174			
รวม			

7. การแจกแจงความถี่สะสม (Cumulative Frequency Distribution)

การหาความถี่สะสม (Cumulative Frequency ; Cf) คือ การหาผลรวมของความถี่ของแต่ละชั้นเข้าด้วยกัน จากชั้นแรกถึงชั้นที่ต้องการ ซึ่งสามารถหาได้ 2 ลักษณะ คือ

1. การหาความถี่สะสมของข้อมูลจากค่าน้อยที่สุดไปหาค่ามากที่สุด (Less than Cumulative Frequency)

2. การหาความถี่สะสมของข้อมูลจากค่ามากที่สุดไปหาค่าน้อยที่สุด (More than Cumulative Frequency)

ตารางที่ 4 แสดงการแจกแจงความถี่สะสมของข้อมูลจากค่าน้อยไปหาค่ามาก

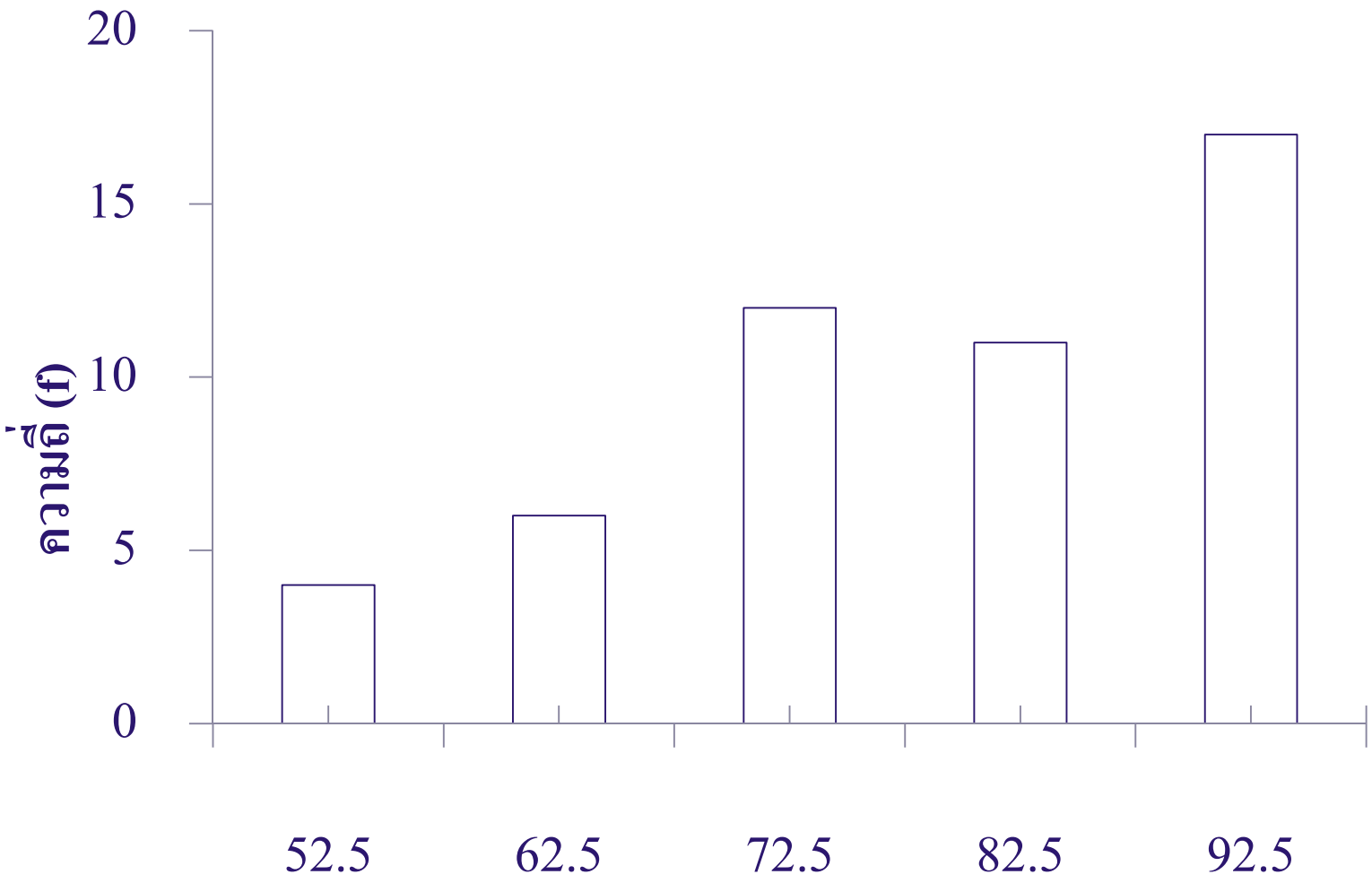
ค่าจ้าง (บาท)	ความถี่ (f)	ความถี่สะสม (cf)
140 -144		
145-149		
150- 154		
155-159		
160-164		
165-169		
170-174		

การแจกแจงความถี่ด้วยแผนภูมิ

เป็นการแจกแจงความถี่ของข้อมูลด้วยกราฟแท่งแบ่งออกได้หลายรูปแบบ คือ

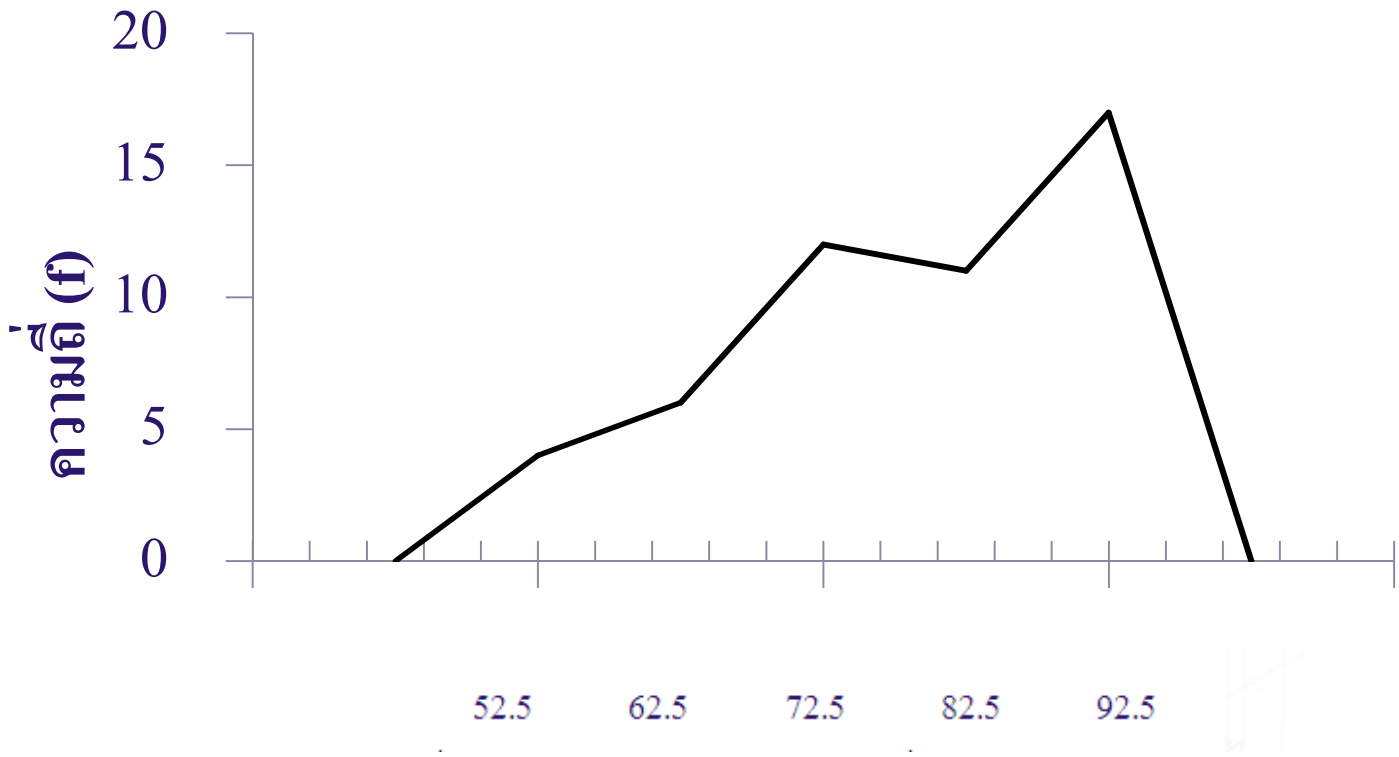
1. ฮิสโตแกรม (Histogram)
2. รูปหลายเหลี่ยมของความถี่ (Frequency Polygon)
3. โค้งแห่งความถี่ (Curve of Frequency Distribution)

1) ฮิสโตแกรม (Histogram) คือ เป็นการนำข้อมูลที่ได้แจกแจงความถี่มาเสนอในรูปกราฟแท่งสี่เหลี่ยม โดยใช้ความสูงของสี่เหลี่ยมแปรตามความถี่ของคะแนนแต่ละชั้น และความกว้างแต่ละชั้นจะมีความกว้างเท่า ๆ กัน และให้จุดกึ่งกลางของแต่ละชั้นเป็นจุดกึ่งกลางของแท่งสี่เหลี่ยม



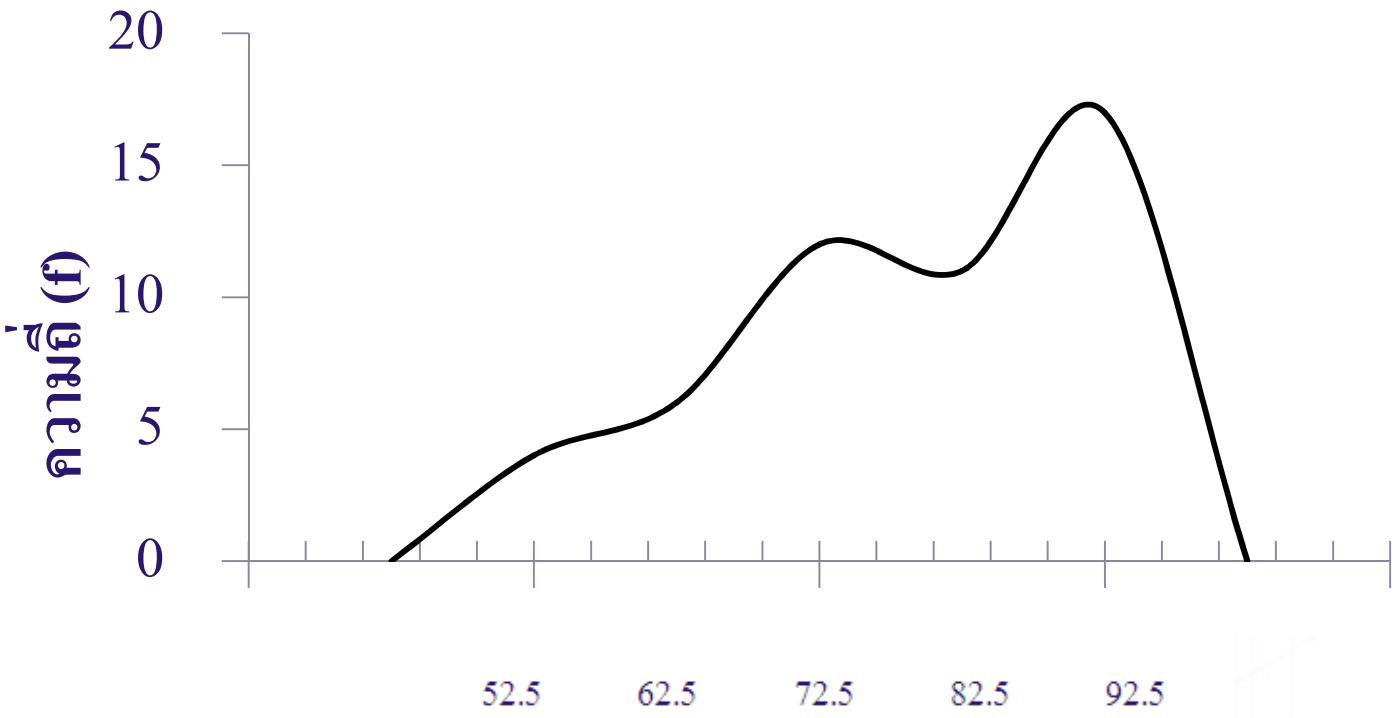
ภาพที่ 1 ฮิสโตแกรมแสดงการแจกแจงความถี่ของคะแนนสอบ

2) รูปหลายเหลี่ยมแห่งความถี่ (Frequency Polygon) เป็นการนำเสนอข้อมูล โดยใช้รูปของฮิสโตแกรม (Histogram) มาพิจารณาด้วยกัน โยงเส้นจุดกึ่งกลางแต่ละชั้นของฮิสโตแกรม และเพิ่มจุดกึ่งกลางต่ำสุด และสูงสุดอีก 2 ชั้นเข้าไปด้วย



ภาพที่ 2 โพลิกอนของคะแนนสอบ

3) รูปโค้งความถี่ (Frequency Curve) เป็นรูปภาพที่ได้จากการนำรูปหลายเหลี่ยม แห่งความถี่มาปรับให้เป็นเส้นโค้ง



ภาพที่ 3 โค้งความถี่ของคะแนนสอบ

การหาสัดส่วน (Proportion)

ค่าสัดส่วน (proportion) คือ ค่าที่เปรียบเทียบระหว่างข้อมูลที่สนใจศึกษากับข้อมูลทั้งหมด โดยการนำข้อมูลแต่ละกลุ่มหารด้วยผลรวมของข้อมูลทั้งหมด

สาขางาน	จำนวนนักศึกษา	ค่าสัดส่วน
การประมง	15	15/60 =
เกษตรทั่วไป	25	25/60 =
อุตสาหกรรม	20	
รวม	60	

ร้อยละ (Percentage)

ร้อยละ (percentage) คือ การนำจำนวนข้อมูลแต่ละระดับหารด้วยจำนวนข้อมูลทั้งหมด แล้วนำไปคูณด้วย 100 หรือ สัดส่วน x 100 นั่นเอง

สาขางาน	จำนวนนักศึกษา	ร้อยละ
การประมง	15	(15/60)x 100 =
เกษตรทั่วไป	25	(25/60)x100 =
อุตสาหกรรม	20	
รวม	60	

ลองฝึกหัดเพิ่มเติม

ใบงานกิจกรรมที่ 2