

## หน่วยที่ 8 การวิเคราะห์ข้อมูลและการนำเสนอข้อมูล

### สถิติพรรณนา

**การแจกแจงความถี่ (frequency distribution)** คือ การจัดข้อมูลออกเป็นหมู่หรือเป็นชั้น ๆ แล้วแจกแจงดูว่าในแต่ละชั้นมีข้อมูลหรือตัวเลขที่มีค่าอยู่ในชั้นนั้นเป็นจำนวนกี่ตัว จำนวนข้อมูลที่ได้เรียกว่าความถี่ วิธีนับหาความถี่นิยมใช้เครื่องหมายรอยคะแนน (tally mark) นำมาสร้างเป็นตารางแจกแจงความถี่ ประกอบด้วย ช่องแสดงชั้น (class) หรือชั้นของข้อมูล ช่องแสดงความถี่ (frequency) หรือจำนวนของข้อมูลในชั้นนั้น มีการสรุปรายได้ 2 วิธี

1. ตารางแจกแจงความถี่แบบธรรมดา (simple frequency table) เช่น

ระดับคอเลสเตรอล (มก/ดล)	ความถี่
106	2
114	1
116	2
118	2
รวม	5

วิธีนี้ไม่เป็นที่นิยม เพราะไม่กะทัดรัด มักใช้กับข้อมูลที่มีค่าใกล้กัน

2. ตารางแจกแจงความถี่แบบมีอันตภาคชั้น (frequency distribution in class intervals) คือการจัดการข้อมูลออกเป็นชั้น ๆ ให้มีอันตภาคชั้น ซึ่งอาจเท่ากันทุกชั้นหรือไม่เท่ากันก็ได้ เท่ากันจะสะดวกในการคำนวณ การคำนวณจะผิดจากวิธีแรกไปบ้างแต่ไม่มาก ถ้ามีข้อมูลมากๆ ค่าความคลาดเคลื่อนจะน้อยลง ตารางแจกแจงความถี่ประกอบด้วย

**ค่าต่ำสุดของข้อมูล (minimum value)** คือ ตัวเลขที่มีค่าต่ำสุดของข้อมูลชุดนั้น

**ค่าสูงสุดของข้อมูล (maximum value)** คือ ตัวเลขที่มีค่าสูงสุดของข้อมูลชุดนั้น

**พิสัย (range)** คือค่าแตกต่างระหว่างค่าสูงสุดกับค่าต่ำสุดของข้อมูล  $\text{พิสัย} = \text{ค่าสูงสุด} - \text{ค่าต่ำสุดของข้อมูล}$

**ชั้น (class)** คือ ข้อมูลที่จัดเป็นชั้นๆ เรียงจากแถบบนลงมาล่าง ต้องครอบคลุม ค่าต่ำสุดและค่าสูงสุดของข้อมูล

$$\text{สูตร } k = 1 + 3.322(\log n) \quad k = \text{จำนวนชั้น} \quad n = \text{จำนวนข้อมูล}$$

**อันตภาคชั้น (class interval)** หมายถึงช่วงห่างของชั้น ค่าอาจเป็นตัวเลขจำนวนเต็มหรือมีจุดทศนิยมก็ได้ สูตรเท่ากับ  $\frac{\text{พิสัย}}{\text{จำนวนชั้น}}$

**ขีดจำกัดชั้น (class limit)** ประกอบด้วยขีดจำกัดต่ำสุดของชั้นกับขีดจำกัดสูงสุดของชั้น แบ่งออกได้เป็น 2 อย่างคือ

- เขตของชั้น (state class limit) คือเขตของชั้นที่บอกคร่าว ๆ ว่าเป็นเท่าใด เช่น

10 – 29

30 – 49

- ขอบของชั้นที่แท้จริง (true class limit) คือขอบเขตของชั้นที่กำหนดลงไปว่าจะมีขอบเขตที่แน่นอนจริง ๆ เป็นเท่าใด

ขอบบนของชั้น = ขีดจำกัดสูงสุดของชั้น + 0.5

ขอบล่างของชั้น = ขีดจำกัดต่ำสุดของชั้น - 0.5 ตัวอย่างเช่น

ขีดจำกัด

ขอบจำกัด

10 – 29

9.5 – 29.5

30 – 49

29.5 – 49.5

จุดกึ่งกลางชั้น (midpoint) คือค่าที่อยู่ตรงกลางระหว่างขอบล่างกับขอบบนของชั้น

$$\text{จุดกึ่งกลางชั้น} = \frac{\text{ขอบล่างของชั้น} + \text{ขอบบนของชั้น}}{2}$$

2

**ความถี่** (frequency) คือจำนวนข้อมูลที่แจกแจงหรือนับได้ในแต่ละชั้น

**ความถี่สะสม** (cumulative frequency) ความถี่สะสมของชั้นใด คือผลรวมของความถี่ตั้งแต่ชั้นแรกจนถึงชั้นนั้น ดังนั้น

ความถี่สะสมชั้นสุดท้ายจะเท่ากับความถี่ทั้งหมด

หลักในการสร้างตารางความถี่แบบมีอันตภาคชั้น

1. หาค่าสูงสุดของข้อมูล
2. หาค่าต่ำสุดของข้อมูล
3. หาพิสัย พิสัย = ค่าสูงสุด - ค่าต่ำสุด
4. กำหนดจำนวนชั้น
5. หาอันตภาคชั้น = พิสัย / จำนวนชั้น
6. การจัดเรียงชั้น
7. แจกแจงความถี่

ตัวอย่าง ในหนังสือ หน้า 11 (เล่ม 2) คู่มือกิจกรรมหน้า 12 ประกอบ

### การวัดแนวโน้มเข้าสู่ส่วนกลาง

เป็นการใช้สถิติพรรณนาเพื่อหาตัวแทนของกลุ่ม ได้แก่ การบอกว่าข้อมูลชุดนั้น ข้อมูลอะไรเป็นตัวแทนกลุ่ม หรือมีค่าเฉลี่ยเท่าใดนั่นเอง นิยมใช้กันอยู่ทั่วไปมี 3 ชนิด คือ มัชฌิมเลขคณิต มัชยฐาน ฐานนิยม

**มัชฌิมเลขคณิต** (arithmetic mean) เป็นค่าเฉลี่ยที่ได้จากวิธีเลขคณิตธรรมดา คือ การรวมข้อมูลแต่ละตัวเข้าด้วยกันหารด้วยจำนวนข้อมูลทั้งหมด สัญลักษณ์ที่ใช้แทนมี 2 ตัวคือ กรณีเป็นค่าเฉลี่ยประชากรใช้  $\mu$  (มิว) และกรณีเป็นค่าเฉลี่ยของตัวอย่างใช้  $\bar{X}$  (เอ็กซ์บาร์)

1. การคำนวณหามัชฌิมเลขคณิตจากข้อมูลที่ไม่ได้แจกแจงความถี่ หมายถึง การหาค่าเฉลี่ยของข้อมูลที่ยังไม่ได้มีการแจกแจงความถี่ ยังอยู่ในสภาพกระจัดกระจาย หรือเป็นข้อมูลดิบ มักใช้กับข้อมูลที่มีจำนวนไม่มาก

สูตร  $\bar{X} = \frac{\sum x}{n}$  เช่น

ตัวอย่าง สมมติว่ามีนักเรียนอนุบาลจำนวน 10 คน ได้รับเงินค่าขนมสำหรับไปโรงเรียนจากผู้ปกครอง ดังนี้ 3,3,4,5,5,6,7,7 และ 8 บาท จงคำนวณหามัชฌิมเลขคณิตของข้อมูลชุดนี้

วิธีทำ

$$\bar{X} = \frac{3+3+4+5+5+6+7+7+8}{10} = \frac{53}{10} = 5.3$$

นั่นคือ ค่าเฉลี่ยของค่าขนมของเด็กนักเรียนอนุบาลเท่ากับ 5.3 บาท

2. การคำนวณหามัชฌิมเลขคณิตจากข้อมูลที่มีการแจกแจงความถี่ (ตัวอย่างหน้า 15)

### ข้อดีของมัชฌิมเลขคณิต

1. เข้าใจง่ายและคำนวณง่าย
2. เป็นค่าที่ได้มาจากค่าของข้อมูลทุกตัว

3. เป็นค่าที่แน่นอน
4. เป็นค่าตัวกลางที่เหมาะสมกับการกระจายปกติ หรือเกือบปกติ
5. เหมาะสำหรับการคำนวณขั้นต่อไป

### ข้อเสีย

1. มีค่าตรงกับความเป็นจริงของข้อมูลไม่กี่ราย หรืออาจจะไม่มีเลย
2. เป็นค่าที่ถูกกระทบกระเทือน โดยค่าของข้อมูลที่สูงหรือต่ำมากอย่างผิดปกติ หรือข้อมูลที่ผิดปกติ

**มัธยฐาน** (median) คือ ค่าของข้อมูลที่มีตำแหน่งตรงกลางของข้อมูลชุดหนึ่ง ซึ่งได้จัดเรียงลำดับค่า หรือจัดหมู่ไว้เรียบร้อยแล้ว นั่นคือการเรียงข้อมูลจากน้อยไปหามาก การคำนวณมัธยฐานมี 2 วิธี

1. การคำนวณมัธยฐานจากข้อมูลที่ไม่ได้แจกแจงความถี่
2. การคำนวณมัธยฐานจากข้อมูลที่มีการแจกแจงความถี่ ( สูตรคำนวณ หน้า 17 )

### ข้อดีของมัธยฐาน

1. เข้าใจง่าย
2. ค่ามัธยฐานที่หาได้จากข้อมูลที่เรียงลำดับแล้ว จะเป็นค่าที่ตรงกับค่าจริงของข้อมูลนั้น
3. เป็นค่าที่ไม่ถูกกระทบกระเทือน โดยค่าของข้อมูลที่มีค่าสูงหรือต่ำมากผิดปกติ หรือข้อมูลที่ผิดปกติ
4. ถ้าทราบค่าข้อมูลในตำแหน่งกลาง ๆ ก็พอจะคำนวณหาค่ามัธยฐานได้
5. ไม่ได้ใช้ทุกค่ามาคำนวณ เพราะในการแจกแจงความถี่ที่อยู่ในรูปของตารางแจกแจงความถี่ที่มีช่วงเปิด เช่น น้อยกว่า หรือมากกว่าก็สามารถคำนวณได้

### ข้อเสียของมัธยฐาน

1. ค่ามัธยฐานไม่แน่นอน ถ้าการแจกแจงข้อมูลไม่สมมาตร
2. ค่ามัธยฐานที่ได้จากข้อมูลที่จัดหมวดหมู่ จะเป็นค่าที่ไม่ถูกต้องตามความเป็นจริง
3. จะหาค่ามัธยฐานไม่ได้ หากไม่ได้จัดเรียงลำดับค่าของข้อมูลให้เรียบร้อยก่อน

**ฐานนิยม** (mode) ฐานนิยมคือค่าที่มีความถี่มากที่สุด หรือค่าที่ซ้ำกันมากที่สุด ( ตัวอย่างคำนวณ หน้า 20 )

### ข้อดีของฐานนิยม

1. เข้าใจง่าย
2. หาได้ง่ายจากข้อมูลที่เรียงลำดับค่าแล้ว
3. เป็นค่าที่ไม่ถูกกระทบกระเทือน โดยค่าของข้อมูลที่มีค่าสูง หรือค่าต่ำมากผิดปกติ หรือข้อมูลที่ผิดปกติ
4. อาศัยเสียงข้างมากเป็นเกณฑ์

### ข้อเสียของฐานนิยม

1. ฐานนิยมอาจมีหลายค่า
2. ฐานนิยมบางค่าอยู่ก่อนไปทางค่าสูงมาก บางค่าก่อนไปทางค่าที่ต่ำมาก
3. หาค่าที่แน่นอนไม่ได้จากข้อมูลที่จัดเป็นหมู่
4. ไม่เหมาะที่จะคำนวณขั้นต่อไป

**เปอร์เซ็นต์ไทล์** (Percentiles) คือค่าของตำแหน่งข้อมูล (99 ค่า) ที่จะแบ่งข้อมูลชุดหนึ่งออกเป็น 100 ส่วนเท่า ๆ กันมักใช้ในงานสาธารณสุข โดยเฉพาะทางด้านโภชนาการ เช่น อยากทราบว่าน้ำหนักของเด็กในวัยทารกที่ไปสำรวจนั้นในเปอร์เซ็นต์ไทล์ที่ 90 มีค่าเท่าใด ( สูตร ตัวอย่างคำนวณ หน้า 22 )

**ควอร์ไทล์** (quartiles or Q) คือค่าของตำแหน่งของข้อมูล (3 ค่า) ที่แบ่งข้อมูลชุดหนึ่งออกเป็น 4 ส่วนเท่า ๆ กัน โดยที่ข้อมูลนั้นมีการเรียงค่าไว้แล้ว จากน้อยไปหามาก (ตัวอย่างหน้า 23)

**มัธยฐานเลขคณิต** คือการถอนกรณีที่  $n$  (Root  $n$  th) ของผลคูณของข้อมูลทั้งหมดมักใช้กับข้อมูลทางด้านห้องปฏิบัติการ เช่น ข้อมูลจากการทดลองทางด้านจุลชีววิทยา ปรสตีวิทยา (สูตร คำนวณหน้า 25)

การวัดการกระจาย (measures of dispersion) เป็นวิธีวิเคราะห์ข้อมูลทางสถิติเพื่อให้ทราบว่า ค่าต่าง ๆ ของข้อมูลชุดหนึ่ง ๆ นั้นมีความแตกต่าง สูง ต่ำ กว่ากันมากน้อยเพียงใด หรือมีการกระจายมากหรือน้อย ถ้าข้อมูลชุดใดมีการกระจายมาก แสดงว่าค่าต่าง ๆ ของข้อมูลชุดนั้นมีค่าสูงต่ำผิดกันมาก หรือถ้าหากมีการกระจายน้อย แสดงว่าค่าต่าง ๆ ของข้อมูลชุดนั้นใกล้เคียงกัน การวัดการกระจายที่นิยมมี 3 ชนิด คือ 1. พิสัย 2. ส่วนเบี่ยงเบนควอร์ไทล์ 3. ส่วนเบี่ยงเบนมาตรฐาน

1. **พิสัย** คือ ผลต่างระหว่างค่าของข้อมูลที่สูงที่สุด และค่าที่ต่ำสุดของข้อมูลชุดหนึ่ง ๆ = ค่าสูงสุด - ค่าต่ำสุด

2. **ส่วนเบี่ยงเบนควอร์ไทล์** (quartile deviation) คือ ค่า ครึ่งหนึ่งของผลต่างระหว่างควอร์ไทล์ที่ 3 กับควอร์ไทล์ที่ 1 ของข้อมูลชุดหนึ่ง ๆ  $Q.D. = \frac{Q3 - Q1}{2}$

2 (ตัวอย่าง คำนวณหน้า 28)

3. **ส่วนเบี่ยงเบนมาตรฐาน** (standard deviation) ใช้สัญลักษณ์  $\sigma$  (sigma อ่านว่าซิกมา) แทนส่วนเบี่ยงเบนมาตรฐานของประชากร และ  $s$  แทนส่วนเบี่ยงเบนมาตรฐานของตัวอย่าง เป็นค่าที่ได้จากการถอนกรณีที่สอง (square root) ของค่าเฉลี่ยของกำลังสองของผลต่างระหว่างค่าแต่ละค่าของข้อมูลกับมัธยฐานเลขคณิต เป็นการวัดการกระจายที่ดีที่สุดในงานวิจัย (ตัวอย่าง หน้า 29)

ถ้าส่วนเบี่ยงเบนมาตรฐานมีค่ามาก หมายความว่าค่าต่าง ๆ ของข้อมูลชุดนั้นมีการกระจายมากรูปโค้งจะกว้างแบน ถ้าส่วนเบี่ยงเบนมาตรฐานมีค่าน้อย แสดงว่าค่าต่าง ๆ ของข้อมูลชุดนั้นมีการกระจายน้อยรูปโค้งจะแคบและสูง

**ความแปรปรวน** (variance) คือ ค่าของส่วนเฉลี่ยของกำลังสองของผลต่างระหว่างข้อมูลแต่ละตัวจากค่ามัธยฐานเลขคณิต หรือค่ากำลังสองของส่วนเบี่ยงเบนมาตรฐาน (ตัวอย่างคำนวณหน้า 34) คูณกรรม 8.13 ประกอบ ประชากร 1 กลุ่ม ทราบค่าความแปรปรวนของประชากรสถิติที่ใช้คือ **Z-test** (หน้า 39) ประชากร 2 กลุ่มเป็นอิสระต่อกัน

- ทราบค่าความแปรปรวนของประชากรใช้ Z-test
- ไม่ทราบค่าความแปรปรวนของประชากรใช้ t-test

ประชากร 2 กลุ่ม ไม่เป็นอิสระต่อกัน ใช้ **paired t-test**

ประชากร 2 กลุ่มที่เป็นอิสระต่อกันมีการวัดที่แยกจากกัน มีการสังเกตค่าของข้อมูลครั้งเดียว ใช้ **Z-test** (หน้า 42)

ประชากรมากกว่า 2 กลุ่มที่ไม่เป็นอิสระต่อกัน ใช้ สถิติ **Cochran Q test**

สูตรที่ใช้หาค่าสัมประสิทธิ์สหสัมพันธ์คือ **Pearson Product Moment Correlation** (หน้า 47)

ถ้ามีตัวแปรอิสระเพียง 1 ตัว ถือว่าเป็นการวิเคราะห์ถดถอยแบบง่าย (หน้า 48)

ถ้ามีตัวแปรอิสระมากกว่า 1 ตัว เรียกว่า การถดถอยพหุ

การวิเคราะห์จำแนกพหุ หมายถึง การจำแนกที่หลาย ๆ ตัวพร้อมกัน ไม่ใช่การจำแนกทีละตัว (หน้า 49)

เทคนิคการวิเคราะห์จำแนกประเภทมีวัตถุประสงค์ที่สำคัญคือ การทดสอบว่าตัวแปรที่ได้เก็บรวบรวมมาสามารถที่จะจำแนกวิเคราะห์ออกเป็นกลุ่ม ๆ ได้อย่างถูกต้องหรือไม่ (หน้า 50)

**การอนุมานเชิงสถิติ** เป็นการประเมินผลหรือสรุปผลของสมมุติฐานที่ตั้งขึ้นเกี่ยวข้องกับประชากร โดยใช้ค่าหรือผลที่ได้จากตัวอย่างที่สุ่มมาจากประชากรกลุ่มนี้ ไปอ้างอิงประชากรที่ศึกษาการใช้สถิติที่ถูกต้องและเหมาะสมจะทำให้ผลการอ้างอิงนั้นแม่นยำถูกต้องและเชื่อถือได้

การอนุมานเชิงสถิติที่ใช้ในงานวิจัยด้านวิทยาศาสตร์สุขภาพ ที่นิยมและใช้กันมากได้แก่การเปรียบเทียบค่าเฉลี่ย ค่าสัดส่วน การหาความสัมพันธ์ระหว่างตัวแปร และการหาค่าตัวทำนายตัวแปรตาม

การอนุมานเชิงสถิติ จะต้องมีความรู้เบื้องต้นเกี่ยวกับระดับของข้อมูลหรือตัวแปรเป็นอย่างดีและวัตถุประสงค์ หรือสมมุติฐานในงานวิจัยนั้นๆ

Z-test เป็นการทดสอบค่าเฉลี่ยของประชากร 1 กลุ่ม ที่ทราบค่าความแปรปรวน

ใช้สถิติ paired t-test ประชากร 2 กลุ่ม ไม่เป็นอิสระต่อกัน เช่น การชั่งน้ำหนักเครื่องที่ 1 เป็น การวัดครั้งที่ 1 ชั่งน้ำหนักเครื่องที่ 2 เป็นการวัดครั้งที่ 2

การหาค่าตัวทำนายในการอนุมานเชิงสถิติ หมายถึง การนำเอาตัวแปรอิสระ (X) ที่ศึกษา หรือได้จากการอ่านบททวนวรรณกรรมที่เกี่ยวข้องกับตัวแปรตาม (y) มาทำนายการเกิดตัวแปรตาม สถิติที่ใช้คือ การวิเคราะห์ถดถอย การวิเคราะห์จำแนกหมู่ การวิเคราะห์จำแนกประเภท

## การนำเสนอข้อมูล

**การนำเสนอข้อมูลในรูปบทความ** เป็นการนำเสนอข้อมูลที่มีลักษณะเป็นบทความสั้น ๆ ปนไป กับตัวเลขที่มีรายการน้อย นับว่าเป็นข้อมูลธรรมดาและง่ายที่สุด พบได้ทั่วไปในหน้าหนังสือพิมพ์ รายการวิทยุ โทรทัศน์ หรือสรุปรายงานต่าง ๆ

**การเสนอข้อมูลในรูปบทความกิ่งตาราง** เป็นการเสนอด้วยบทความกับตัวเลขที่จัดเป็นหมวดหมู่ซึ่งมีจำนวนไม่มากและไม่จำเป็นต้องทำเป็นรูปตาราง

**การนำเสนอข้อมูลในรูปตาราง** เป็นการนำข้อมูลที่มีหลายรายการหรือซ้ำ ๆ กัน มาแสดงไว้ในตารางให้เป็นแถวตามแนวตั้ง และแถวตามแนวนอนเพื่อความเป็นระเบียบเรียบร้อย และสะดวกในการเปรียบเทียบ

## การนำเสนอข้อมูลในรูปกราฟและแผนภูมิ

### 1. การนำเสนอแบบรูปกราฟเส้น ( line graph) (หน้า 59)

- **กราฟสเกลธรรมดา** คือกราฟที่มีมาตราส่วนธรรมดา ทั้งแกนนอนและแกนตั้ง (หน้า 60)

- **กราฟการเปลี่ยนแปลงเป็นสัดส่วน** คือกราฟเส้นที่ใช้ในการเปรียบเทียบกันระหว่างข้อมูล 2 ชุดขึ้นไปว่า มีการเปลี่ยนแปลงที่สัมพันธ์กันอย่างไร (หน้า 62)

- **กราฟสเกลกึ่งลอการิทึม** คือ กราฟที่มีมาตราส่วนธรรมดาและของอีกแกนหนึ่งซึ่งมักได้แก่แกนตั้งเป็นมาตราส่วนลอการิทึม (หน้า 62)

### 2. การนำเสนอแบบแผนภูมิแท่ง (bar chart)

- **การนำเสนอแผนภูมิแท่งแบบธรรมดา** (หน้า 64)

- **การนำเสนอแบบกราฟแท่งรวม** (multiple bar chart หรือ compound bar chart) หมายถึง การแสดงรายละเอียดย่อยภายในแท่งเดียวกัน เพื่อการเปรียบเทียบตั้งแต่ 2 รายละเอียดย่อย โดยเอาแท่งแต่ละชุดมาเรียงติดกันหรือใกล้กัน ทำเครื่องหมายระบายสีให้แตกต่างกันออกไป (หน้า 65)

- **การนำเสนอแบบกราฟแท่งแยกส่วนประกอบ** (component bar chart) มีลักษณะเดียวกันนำมาเปรียบเทียบกัน โดยใช้ความยาวของแท่งแสดงค่าของข้อมูลประเภทนั้น ๆ แล้วนำมาบรรจุในแท่งเดียวกัน แต่งสีหรือทำเครื่องหมายของแต่ละส่วนให้แตกต่างกันไป (หน้า 65)

3. **การนำเสนอแบบแผนภูมิวงกลม (pie diagram)** คือ ภาพรูปวงกลม แสดงรายการในภาพวงกลม (หน้า 67)

4. **การนำเสนอแบบแผนภูมิรูปภาพ (pictogram)** เป็นการนำเสนอโดยใช้รูปภาพเป็นเครื่องหมายแสดงปริมาณหรือความถี่ของข้อมูลที่ต้องการเสนอ เช่น ภาพคน ภาพสิ่งของ (หน้า 68)

5. **การนำเสนอแบบแผนภูมิแผนที่ (map diagram)** คือ การนำเสนอข้อมูลที่เกี่ยวข้องกับสภาพภูมิศาสตร์หรือสถานที่ เช่น อัตราป่วยด้วยมาลาเรีย ภาวะเจริญพันธุ์ของประเทศต่าง ๆ ในโลก อัตราการเกิด ตาย ป่วย การอพยพ (หน้า 69)

6. **การนำเสนอแบบ แผนภาพกระจาย หรือ แผนภาพสหสัมพันธ์ (scatter or correlation diagram)** เหมาะสำหรับใช้ดูความสัมพันธ์ระหว่างข้อมูลชนิดปริมาณค่าต่อเนื่องกัน 2 ชนิด แขนงนอนและแกนตั้ง สบเป็นคู่ ๆ แสดงไว้ด้วย . (จุด)

- จุดกระจัดกระจายไปทั่ว หมายถึงไม่มีความสัมพันธ์กัน
- ถ้าจุดจับกลุ่มอยู่ตามแนวเส้นทแยง โดยเกิดจากค่าน้อยของแกนนอน สบกับค่าน้อยของแกนตั้ง และค่ามากของแกนนอน สบกับค่ามากของแกนตั้ง หมายถึงความสัมพันธ์กันไปทางบวก
- ถ้าจุดจับกลุ่มอยู่ตามแนวเส้นทแยง โดยเกิดจาก ค่าน้อยของแกนนอน สบกับค่ามากของแกนตั้ง และค่ามากของแกนนอน สบกับค่าน้อยของแกนตั้ง หมายถึงว่า มีความสัมพันธ์กัน ไปทางลบ
- ถ้าจุดจับกลุ่มอยู่เป็นแนวโค้ง เรียกว่ามีความสัมพันธ์กันแบบเชิงเส้นโค้ง (หน้า 70)

7. **การนำเสนอแบบฮิสโตแกรม (histogram)** คือ รูปกราฟแท่งที่เรียงติดกันและฐานของแท่งเป็นตัวแปรในสเกลช่วงมาตรา หรืออัตราส่วนมาตรา มีหลักดังนี้


- ใช้กับข้อมูลที่มีการแจกแจงความถี่
- ใช้แสดงปริมาณด้วยพื้นที่ ๆ สร้างขึ้นเป็นแท่ง ตั้งอยู่บนแกนนอน
- แท่งเหล่านี้นวางชิดติดกัน
- เพื่อความสะดวกในการสร้างควรทำให้ทุกอันตรภาคชั้น (class interval) มีค่าเท่า ๆ กัน ถ้าไม่เท่ากันต้องทำให้เป็นความหนาแน่นของความถี่เสียก่อน (กราฟ หน้า 71)

รูปฮิสโตแกรม คือ รูปแท่ง ส่วนภาพเส้นประที่จะสร้างต่อที่หลัง คือ รูปหลายเหลี่ยมแห่งความถี่

8. **การนำเสนอรูปหลายเหลี่ยมแห่งความถี่ (frequency polygon)** เปรียบเทียบข้อมูลแจกแจงความถี่ตั้งแต่ 2 ชุดขึ้นไป ดัดแปลงมาจากรูปฮิสโตแกรมให้ดูง่ายซึ่งมีเนื้อหาอยู่ครบตามเดิม (รูป กราฟ หน้า 72 )

 - รูปสี่เหลี่ยมผืนผ้า หมายถึงว่ามีความถี่เท่ากันหมด

 - รูปประฆังสมมาตร หมายถึงว่าส่วนใหญ่มีคะแนนปานกลาง คะแนนน้อย ๆ และคะแนนมาก ๆ มีเป็นจำนวนน้อย

 - รูปประฆังเบ้ขวา หมายถึงว่าคะแนนน้อย ๆ มีเป็นจำนวนมาก คะแนนมาก ๆ มีเป็นจำนวนน้อย

 - รูปประฆังเบ้ซ้าย หมายถึงว่าคะแนนน้อย ๆ มีเป็นจำนวนน้อย คะแนนมากมีเป็นจำนวนมาก

- รูปตัว U หมายถึงว่า คะแนนน้อยและคะแนนมากมีจำนวนมากกว่ามีจำนวนมากกว่าคะแนนกลาง ๆ เช่นการแจกแจงของปริมาณรถในช่วงเช้าก่อนทำงานและ ช่วงเย็นหลังเลิกงาน

- รูปตัว J หมายถึงว่าคะแนนน้อยมีจำนวนมากกว่าคะแนนกลาง ๆ แต่น้อยกว่าคะแนนมาก เช่นการแจกแจงอายุประชากรตามอัตราตายตามกลุ่มอายุ

- รูปตัว J กลับ หมายถึงว่าคะแนนน้อยมีจำนวนมากกว่าคะแนนกลาง ๆ และคะแนนมากมีน้อยกว่าคะแนนน้อย

**9. การนำเสนอแบบแผนภูมิพีระมิด (pyramidal chart) หรือพีระมิดประชากร (population pyramid) แสดงถึงโครงสร้างประชากรตามกลุ่มอายุและเพศ เป็นภาพสี่โตแกรม 2 ภาพติดกัน แกน x แทนด้วยประชากรแยกตามเพศชาย เพศหญิง แกน Y แทนกลุ่มอายุ นิยมจัดกลุ่มอายุให้มีอันตรภาคชั้นเท่ากันเท่ากับ 5 สามารถแสดงได้ 3 รูปแบบ**

- **พีระมิดประชากร แสดงเป็นค่าจำนวนประชากร** ทำให้ผู้อ่านสามารถทราบได้ว่า จำนวนประชากรแยกตามเพศและกลุ่มอายุ มีจำนวนกี่คน
- **พีระมิดประชากร แสดงเป็นค่าร้อยละของประชากรทั้งหมด** ทำให้ผู้อ่านทราบได้ว่า มีประชากรในแต่ละกลุ่มอายุและเพศ มีกี่เปอร์เซ็นต์ของประชากรทั้งหมด ประชากรวัยต่าง ๆ เช่นเด็กต่ำกว่า 5 ปี มีเป็นร้อยละเท่าไรของประชากรทั้งหมด
- **พีระมิดประชากร แสดงเป็นค่าร้อยละของประชากรแต่ละเพศ** ทำให้ผู้อ่านทราบว่าประชากรแต่ละกลุ่ม อายุ และเพศ มีเป็นร้อยละเท่าไรของจำนวนประชากรรวมแต่ละเพศนั้น ๆ เป็นการเปรียบเทียบกันเอง

**10. การนำเสนอข้อมูลแบบอื่น ๆ เช่น รูปสามเหลี่ยมมิติ แต่ไม่ค่อยนิยมใช้ในงานวิทยาศาสตร์สุขภาพ**