

# I. ค่ากลางของข้อมูล

## ค่าเฉลี่ยเลขคณิต

ไม่แจกแจงความถี่

$$\mu = \frac{x_1 + x_2 + x_3 + \dots + x_N}{N} = \frac{\sum_{i=1}^N x_i}{N}$$

แจกแจงความถี่

$$\mu = \frac{f_1 x_1 + f_2 x_2 + f_3 x_3 + \dots + f_k x_k}{f_1 + f_2 + f_3 + \dots + f_k} = \frac{\sum_{i=1}^k f_i x_i}{\sum_{i=1}^k f_i}$$

$x_i$  คือค่ากึ่งกลางของอันตรภาคชั้น

## ค่าเฉลี่ยเลขคณิตแบบถ่วงน้ำหนัก

$$\mu = \frac{w_1 x_1 + w_2 x_2 + w_3 x_3 + \dots + w_N x_N}{w_1 + w_2 + w_3 + \dots + w_N} = \frac{\sum_{i=1}^N w_i x_i}{\sum_{i=1}^N w_i}$$

## ค่าเฉลี่ยเลขคณิตรวม

$$\mu_{รวม} = \frac{N_1 \mu_1 + N_2 \mu_2 + N_3 \mu_3 + \dots + N_k \mu_k}{N_1 + N_2 + N_3 + \dots + N_k} = \frac{\sum_{i=1}^k N_i \mu_i}{\sum_{i=1}^k N_i}$$

## I. ค่ากลางของข้อมูล (ต่อ)

### ค่ามัธยฐาน

ค่ามัธยฐาน คือ ค่าที่แบ่งข้อมูลออกเป็นสองส่วนเท่าๆ กัน

ไม่แจกแจงความถี่ (ตัวๆ)	แจกแจงความถี่ (ช่วงๆ)
<p>หลักการหาค่า <i>Med</i></p> <ol style="list-style-type: none"> <li>เรียงข้อมูลจากน้อยไปมาก</li> <li>หาดำแหน่งของ <i>Med</i> คือ <math>\frac{N+1}{2}</math></li> <li>หาค่าที่ตำแหน่งนั้นคือ ค่า <i>Med</i></li> </ol>	<p>หลักการหาค่า <i>Med</i></p> <ol style="list-style-type: none"> <li>เรียงข้อมูลจากน้อยไปมาก</li> <li>หาดำแหน่งของ <i>Med</i> คือ <math>\frac{N}{2}</math></li> <li>ค่า <i>Med</i> = <math>L + I \left( \frac{\frac{N}{2} - \sum f_L}{f_m} \right)</math></li> </ol> <p><i>L</i> คือ ขอบล่างของชั้นที่ <i>Med</i> อยู่  <math>\sum f_L</math> คือ ความถี่สะสมชั้นก่อนหน้า  <math>f_m</math> คือ ความถี่ชั้นมัธยฐาน</p>

### ค่าฐานนิยม

ค่าฐานนิยม คือ ค่าของข้อมูลที่ซ้ำกันมากที่สุด

ไม่แจกแจงความถี่ (ตัวๆ)
<p><i>Mode</i> คือ ค่าของข้อมูลที่ซ้ำกันมากที่สุด</p> <p><i>Mode</i> ของข้อมูลชุดหนึ่ง อาจมีได้มากกว่า 1 ค่า หรือไม่มีก็ได้</p>

## I. ค่ากลางของข้อมูล (ต่อ)

### ค่ากึ่งกลางพิสัย

$$\text{ค่ากึ่งกลางพิสัย} = \frac{x_{\max} + x_{\min}}{2}$$

### สมบัติของค่ากลาง

1.  $\sum (x_i - \mu)^2$  มีค่าน้อยสุด ;  $\sum (x_i - \mu) = 0$
2.  $\sum |x_i - \text{Med}|$  มีค่าน้อยสุด
3. ถ้ามีข้อมูล 2 ชุด ซึ่งมีจำนวนเท่ากัน  
ชุดที่ 1  $x_1, x_2, \dots, x_i, \dots, x_N$   
ชุดที่ 2  $y_1, y_2, \dots, y_i, \dots, y_N$   
มีสมการความสัมพันธ์  $y_i = Ax_i + B$  จะได้ว่า  
$$\mu_y = A\mu_x + B$$
$$\text{Med}_y = A\text{Med}_x + B$$
$$\text{Mode}_y = A\text{Mode}_x + B$$

### TRICK !! “การลดทอน”

กรณี  $I$  เท่ากัน

$$d_i = \frac{x_i - a}{I}$$

$$x_i = a + d_i I$$

$$\mu = a + I\bar{d}$$

กรณี  $I$  ไม่เท่ากัน

$$d_i = x_i - a$$

$$x_i = a + d_i$$

$$\mu = a + \bar{d}$$

## II. การวัดตำแหน่งของข้อมูล

$Q_r, D_r, P_r$

**ควอร์ไทล์** คือ การแบ่งข้อมูลออกเป็น 4 ส่วนเท่าๆ กัน หลังจากเรียงคะแนนจากน้อยไปมาก มีตัวแบ่งหลักอยู่ 3 ตัว  $Q_1, Q_2$  และ  $Q_3$   $\overline{Q_1 \quad Q_2 \quad Q_3}$

**เดซิล์** คือ การแบ่งข้อมูลออกเป็น 10 ส่วนเท่าๆ กัน หลังจากเรียงคะแนนจากน้อยไปมาก มีตัวแบ่งหลักอยู่ 9 ตัว  $D_1, D_2, \dots, D_9$   $\overline{D_1 \quad D_2 \dots D_9}$

**เปอร์เซ็นต์ไทล์** คือ การแบ่งข้อมูลออกเป็น 100 ส่วนเท่าๆ กัน หลังจากเรียงคะแนนจากน้อยไปมาก มีตัวแบ่งหลักอยู่ 99 ตัว  $P_1, P_2, \dots, P_{99}$   $\overline{P_1 \quad P_2 \dots P_{99}}$

**หลักการหาค่า  $Q_r, D_r, P_r$**

เรียงข้อมูลจากน้อยไปมาก

ไม่แจกแจงความถี่ (ตัวๆ)	แจกแจงความถี่ (ช่วงๆ)
<p>หลักการหาค่า <math>Q_r, D_r, P_r</math></p> <ol style="list-style-type: none"> <li>เรียงข้อมูลจากน้อยไปมาก</li> <li>หาตำแหน่งของ <math>Q_r, D_r, P_r</math> คือ <math>\frac{r(N+1)}{4}, \frac{r(N+1)}{10}, \frac{r(N+1)}{100}</math> ตามลำดับ</li> <li>หาค่าที่ตำแหน่งนั้น คือ ค่า <math>Q_r, D_r, P_r</math></li> </ol>	<p>หลักการหาค่า <math>Q_r, D_r, P_r</math></p> <ol style="list-style-type: none"> <li>เรียงข้อมูลจากน้อยไปมาก</li> <li>หาตำแหน่งของ <math>Q_r, D_r, P_r</math> คือ <math>\frac{rN}{4}, \frac{rN}{10}, \frac{rN}{100}</math> ตามลำดับ</li> <li>ค่า <math>Q_r, D_r, P_r</math> <math>= L + I \frac{(\text{ตำแหน่ง} - \text{ความถี่สะสมขั้นก่อนหน้า})}{\text{ความถี่ชั้น } Q, D, P}</math></li> </ol>

### NOTE

- $Q_2 = P_{50} = D_5 = \text{Med}$
- ถ้าตำแหน่งของ  $P, Q, D$  เท่ากับความถี่สะสม แล้วค่า  $P, Q, D$  จะเท่ากับขอบบนของชั้นนั้น



### III. การวัดการกระจาย

#### การกระจายสัมบูรณ์ การกระจายสัมพัทธ์

การกระจายสัมบูรณ์ ใช้วัดการกระจายข้อมูลเพียงชุดเดียว

การกระจายสัมพัทธ์ ใช้เปรียบเทียบการกระจายของข้อมูลมากกว่า 1 ชุด

การกระจายสัมบูรณ์		การกระจายสัมพัทธ์	
พิสัย (R)	$x_{\max} - x_{\min}$	ส.ป.ส. การแปรผันพิสัย (C.R.)	$\frac{x_{\max} - x_{\min}}{x_{\max} + x_{\min}}$
ส่วนเบี่ยงเบน ควอร์ไทล์ (Q.D.)	$\frac{Q_3 - Q_1}{2}$	ส.ป.ส. ส่วนเบี่ยงเบน ควอร์ไทล์ (C.Q.)	$\frac{Q_3 - Q_1}{Q_3 + Q_1}$
ส่วนเบี่ยงเบน เฉลี่ย (M.D.)	$\frac{\sum  x - \mu }{N}$	ส.ป.ส. ส่วนเบี่ยงเบน เฉลี่ย (C.M.)	$\frac{M.D.}{\mu}$
ส่วนเบี่ยงเบน มาตรฐาน $\sigma$	$\sqrt{\frac{\sum (x - \mu)^2}{N}}$ $\sqrt{\frac{\sum x^2}{N} - \mu^2}$	ส.ป.ส. การแปรผัน (C.V.)	$\frac{\sigma}{ \mu }$

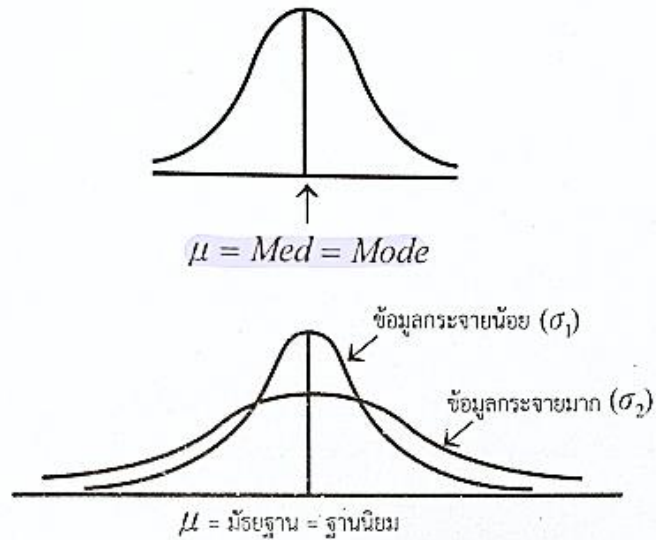
#### NOTE

$$S.D. = \text{ส่วนเบี่ยงเบนมาตรฐานกลุ่มตัวอย่าง} = \sqrt{\frac{\sum (x - \bar{x})^2}{N-1}} = \sqrt{\frac{\sum x^2 - n\bar{x}^2}{N-1}}$$

## IV. การแจกแจงปกติกับค่ามาตรฐาน

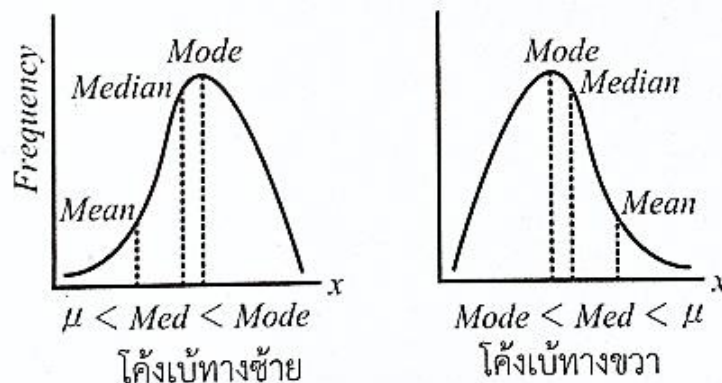
### การแจกแจงปกติ

ความสัมพันธ์ระหว่างค่ากลางกับการกระจายของข้อมูล



$$\sigma_1 < \sigma_2, \frac{\sigma_1}{\mu_1} < \frac{\sigma_2}{\mu_2} \quad \frac{\sigma}{\mu} \text{ ยิ่งน้อยยิ่งดี}$$

NOTE การแจกแจงไม่ปกติ



$$|\mu - \text{Mode}| = 3|\mu - \text{Med}|$$

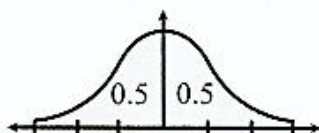
## IV. การแจกแจงปกติกับค่ามาตรฐาน (ต่อ)

### การแจกแจงค่ามาตรฐาน

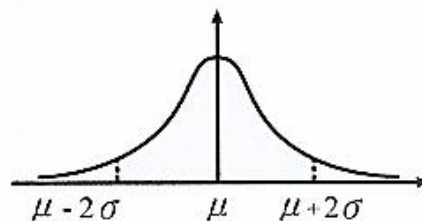
ค่ามาตรฐาน  $z = \frac{x_i - \mu}{\sigma}$

สมบัติของค่า  $Z$

1.  $\bar{Z} = 0 \rightarrow \sum Z = 0$
2.  $\sigma_z = 1 \rightarrow \sum Z^2 = N$
3. พื้นที่ใต้เส้นโค้งมีค่าเท่ากับ 1



The 95% rule



- จำนวนข้อมูลที่อยู่ในช่วง  $(\mu - 2\sigma, \mu + 2\sigma) \approx 95\%$
- $R \approx 4\sigma \rightarrow \sigma \approx \frac{R}{4}$