3. มาตราวัดแนวโน้มเข้าสู่ส่วนกลาง

1. การวัดแนวโน้มเข้าสู่ส่วนกลาง
ด้วยค่าเฉลี่ยเลขคณิต

2. การวัดแนวโน้มเข้าสู่ส่วนกลาง
ด้วยค่าเฉลี่ยแบบถ่วงน้ำหนัก

3. การวัดแนวโน้มเข้าสู่ส่วนกลาง
ด้วยค่าเฉลี่ยเรขาคณิต

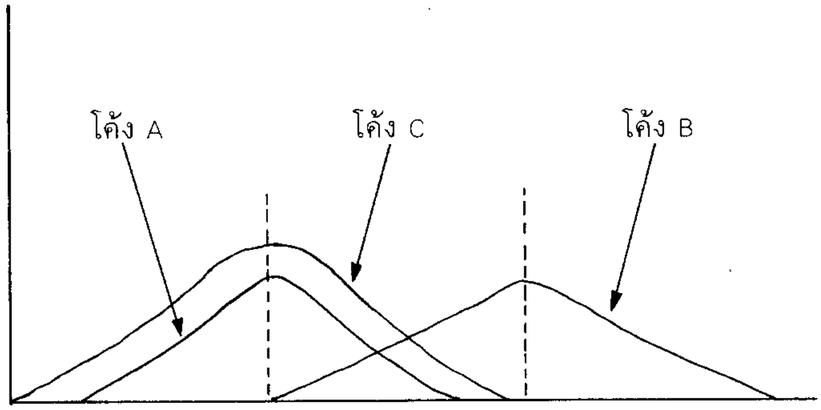
4. การวัดแนวโน้มเข้าสู่ส่วนกลาง
ด้วยค่ามัธยฐาน

5. การวัดแนวโน้มเข้าสู่ส่วนกลาง
ด้วยฐานนิยม

6. เปรียบเทียบระหว่างค่าเฉลี่ย
ค่ามัธยฐาน และฐานนิยม

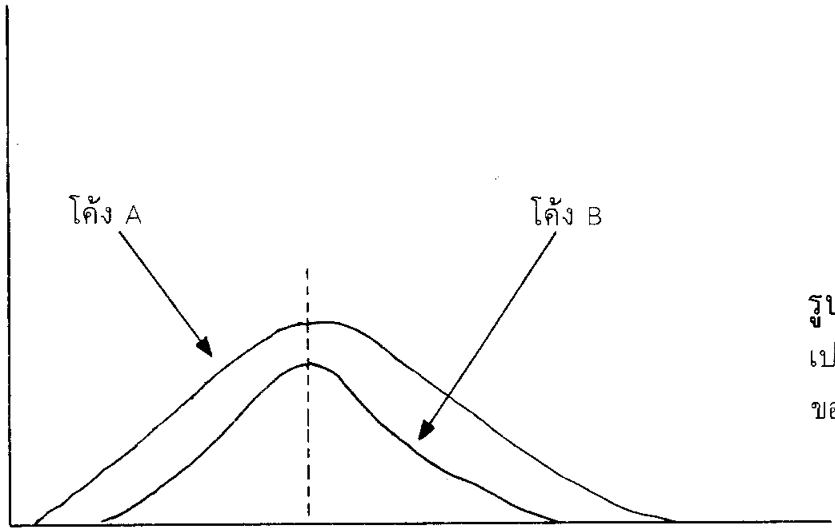
ค่าสถิติเชิงสรุปที่บรรยายลักษณะของโค้งความถี่ มี 4 อย่างคือ

1. **วิธีวัดแนวโน้มเข้าสู่ส่วนกลาง** เช่น**ค่าเฉลี่ย** จะบอกให้ทราบที่ตั้งของศูนย์กลางแนวโน้มเช่นจากรูปที่ 3.1 จุดกึ่งกลางของ B อยู่ทางขวามือของโค้ง A และ C พึงสังเกตว่า โค้ง A และ C มีจุดกึ่งกลางอยู่ ณ ตำแหน่งเดียวกัน



รูปที่ 3.1
เปรียบเทียบจุดกึ่งกลางของ 3 โค้ง

2. **วิธีวัดการกระจาย** การกระจายหมายถึงการแผ่ขยายของข้อมูล range เป็นวิธีวัดการกระจายอย่างหนึ่ง จากรูปที่ 3.2 โค้ง A จะมีการกระจายมากกว่าโค้ง B

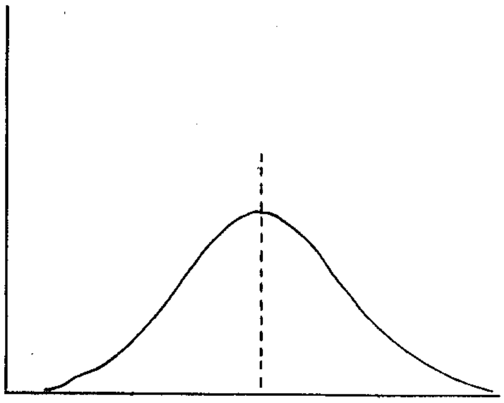


รูปที่ 3.2
เปรียบเทียบการกระจายของ 2 โค้ง

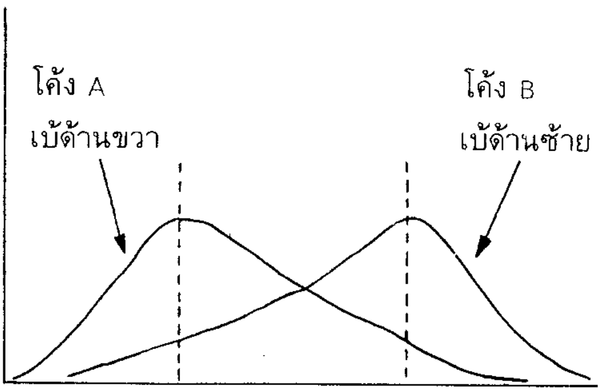
3. การวัดความเบ้

โค้งที่สมมาตรคือ โค้งในรูป 3.3 นั่นคือ เมื่อลากเส้นดิ่งจากจุดยอดของโค้งมายังฐาน เส้นดิ่งนั้นจะแบ่งโค้งออกเป็น 2 ส่วนเท่า ๆ กัน โดยที่แต่ละด้านคือภาพในกระจกของอีกด้านหนึ่ง

แต่โค้งในรูป 3.4 เป็นโค้งที่เบ้ เนื่องจากข้อมูลส่วนใหญ่จะไปรวมกันอยู่ที่ปลายหางด้านใดด้านหนึ่งแทนที่จะอยู่ตรงกลาง เช่น รูป A แสดงโค้งที่เบ้ด้านขวามือ หรือเบ้เชิงบวก เพราะหางด้านขวามือยาวมาก ส่วนโค้ง B เป็นโค้งที่เบ้ด้านซ้ายมือหรือเบ้เชิงลบ เพราะหางด้านซ้ายมือยาวมาก



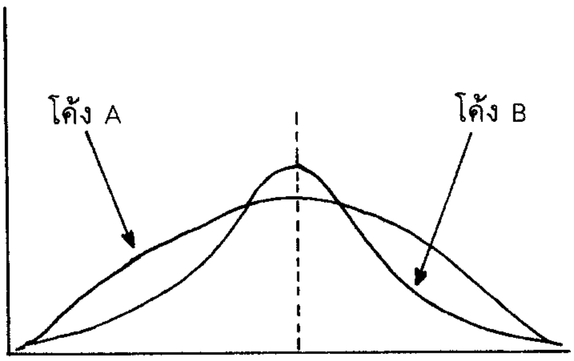
รูป 3.3
แสดงโค้งที่สมมาตร



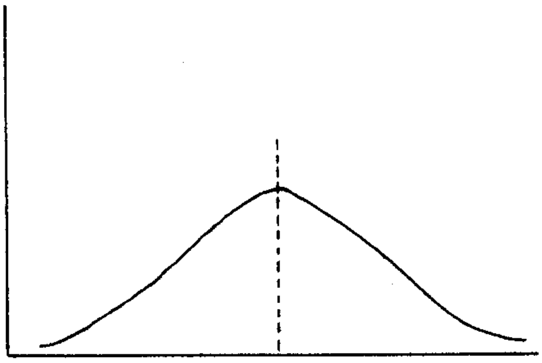
รูป 3.4
แสดงโค้งที่เบ้

4. การวัดความชัน (Kurtosis)

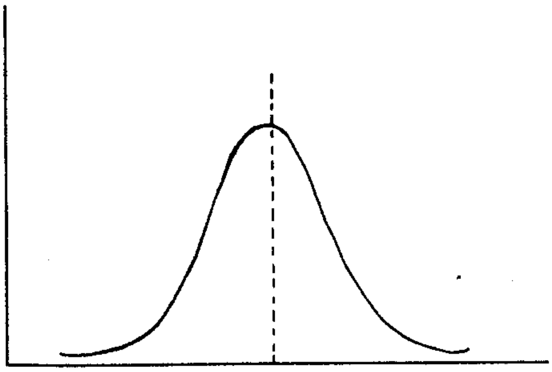
ค่า Kurtosis จะแสดงความสูงของยอดโค้ง ดังรูปที่ 3.5 โค้ง A และ B มีความสูงต่างกัน แม้ว่าจะมีจุดศูนย์กลาง และการกระจายเท่ากัน นักสถิติจะบอกว่า โค้งทั้ง 2 มีดีกรีของ "kurtosis" ต่างกัน ดีกรีของ kurtosis มี 3 ระดับ ดังรูปที่ 3.6 เรียกว่า "mesokurtic" คือโค้งที่มีความชันหรือความสูงแต่เพียงปานกลาง หรือปกติ เพราะ "meso" แปลว่า "Intermediate" ส่วนรูป 3.7 เรียกว่า "leptokurtic" จะมีความชันมากขึ้น โดยที่ "lepto" หมายถึง "slender" และรูป 3.8 แสดงโค้งแบบ "Platykurtic" คือไม่ค่อยชัน และค่อนข้างแบนราบเนื่องจาก "platy" หมายถึง-"Broad" หรือ "flat"



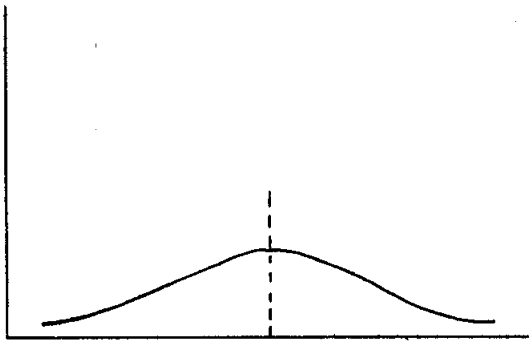
รูป 3.5
แสดงโค้งที่มีดีกรีความชันต่างกัน



รูป 3.6
แสดงโค้งที่มีความชันปานกลาง
หรือ Mesokurtic



รูป 3.7
แสดงโค้งที่มีความชันสูง
เรียกว่า "Leptokurtic"



รูป 3.8
แสดงโค้งที่มีความชันต่ำ
เรียกว่า "Platykurtic"

การวัดแนวโน้มเข้าสู่ส่วนกลาง

1. **ค่าเฉลี่ยเลขคณิต** มี 2 อย่างคือ

ผลรวมของข้อมูลทั้งหมดในประชากร

ค่าเฉลี่ยของประชากร
เป็นค่าพารามิเตอร์ $\mu = \frac{\Sigma X}{N}$

จำนวนหน่วยทั้งหมดของประชากร

ผลรวมของข้อมูลทั้งหมดในตัวอย่าง

ค่าเฉลี่ยของตัวอย่าง
เป็นค่าสถิติ

$$\bar{x} = \frac{\Sigma x}{n}$$

ขนาดของตัวอย่าง

ค่าพารามิเตอร์ คือค่าที่แสดงลักษณะของประชากร และใช้อักษรกรีกเป็นสัญญลักษณ์ เช่น μ (อ่านว่า mu) ใช้แทนค่าเฉลี่ยของประชากร ส่วน**ค่าสถิติ** คือ ค่าที่แสดงลักษณะของตัวอย่าง ใช้แทนด้วยอักษรโรมัน เช่น $\bar{x}$ แทนค่าเฉลี่ยจากตัวอย่าง

ถ้าข้อมูลอยู่รวมกันเป็นกลุ่ม

$$\bar{x} = \frac{\Sigma (f \times x)}{n}$$

f = จำนวนความถี่
x = จุดกึ่งกลางของชั้น

การแปลงข้อมูล (Coding)

การแปลงค่าช่วยให้การคำนวณง่ายขึ้น สูตรในการหาค่าเฉลี่ยคือ

$$\bar{x} = x_o + W \frac{\Sigma (U \times f)}{n}$$

x_o = จุดกึ่งกลางชั้นที่แปลงค่าให้เป็น 0
W = อันตรภาคชั้น , U = ค่าที่แปลงแล้ว
f = ความถี่ , n = ขนาดตัวอย่าง

ตัวอย่าง

ชั้น	จุดกึ่งกลาง (x)	แปลงค่า (U)		ความถี่ (f)		U × f	
0-7	3.5	-2	×	2	=	-4	
8-15	11.5	-1	×	6	=	-6	
16-23	19.5	0	×	3	=	0	← x_0
24-31	27.5	1	×	5	=	5	
32-39	35.5	2	×	2	=	4	
40-47	43.5	3	×	2	=	6	
		Σf =	n =	20		5	← $\Sigma(U \times f)$

$$\bar{x} = 19.5 + (8)\left(\frac{5}{20}\right)$$
$$= 21.5$$

ค่าเฉลี่ยเลขคณิตในฐานะที่ใช้เป็นตัวแทนของข้อมูลทั้งกลุ่ม จะมีทั้งข้อดี และเสีย ดังนี้

ข้อดี

1. เป็นที่ยอมรับและรู้จักใช้กันโดยแพร่หลาย
2. ข้อมูลทุกกลุ่มต้องมีค่าเฉลี่ย 1 จำนวนเสมอ
3. ใช้ในกระบวนการสถิติ เช่น การเปรียบเทียบระหว่างค่าเฉลี่ยของข้อมูลหลาย ๆ กลุ่ม

ส่วน**ข้อเสีย** คือ

1. เนื่องจากค่าเฉลี่ยได้จากผลรวมของทุกจำนวน ดังนั้น ถ้ามีจำนวนหนึ่งที่มีค่าแตกต่างจากข้อมูลส่วนใหญ่ในกลุ่ม ค่านั้นจะมีผลกระทบต่อค่าเฉลี่ยมาก เช่น ข้อมูลกลุ่มหนึ่งมี 7 ตัว คือ 4.2, 4.3, 4.7, 4.8, 5.0, 5.1 และ 9.0 จะเห็นว่า ค่าสุดท้ายต่างจากค่าอื่น ๆ มากเมื่อนำมารวมกันหาค่าเฉลี่ยประชากร

$$\mu = \frac{\Sigma X}{N} = \frac{(4.2 + 4.3 + \ldots + 9.0)}{7}$$
$$= \frac{37.1}{7} = 5.3$$

แต่ถ้าไม่นำตัวสุดท้ายมารวมจะได้ค่าเฉลี่ยเพียง 4.7 ซึ่งดูจะเป็นตัวแทนของประชากรดีกว่า 5.3

2. ถ้ามีข้อมูลหลายจำนวน จะทำให้เสียเวลา เพราะต้องนำทุก ๆ ค่ามาพิจารณา

3. ถ้าข้อมูลซึ่งจัดรวมเป็นกลุ่ม ไม่มีขีดจำกัดบน หรือล่าง จะหาค่าเฉลี่ยไม่ได้

2. ค่าเฉลี่ยแบบถ่วงน้ำหนัก

จงหาต้นทุนถัวเฉลี่ยต่อชั่วโมงของสินค้า 2 ชนิดนี้

ประเภท คนงาน	ค่าจ้างต่อ ชั่วโมง (x)	ชั่วโมงแรงงานที่ใช้ต่อสินค้า 1 หน่วย สินค้า A	สินค้า B
1	฿4.00	1	4
2	6.00	2	3
3	8.00	5	3

ถ้าคิดโดยไม่ถ่วงน้ำหนัก

$\bar{x}$ = ค่าแรงถัวเฉลี่ย = $\frac{4 + 6 + 8}{3}$ = ฿6/ชั่วโมง

ดังนั้น

ต้นทุนต่อหน่วยของสินค้า A

= 6 (1 + 2 + 5) = ฿48

ต้นทุนต่อหน่วยของสินค้า B

= 6 (4 + 3 + 3) = ฿60

แต่วิธีนี้ไม่ถูกต้อง เนื่องจากสินค้า 2 ชนิดใช้ชั่วโมงแรงงานต่างกัน ดังนั้น

ต้นทุนต่อหน่วยของ A

= $(4 \times 1) + (6 \times 2) + (8 \times 5)$ = ฿56

ต้นทุนต่อหน่วยของ B

= $(4 \times 4) + (6 \times 5) + (8 \times 3)$ = ฿58

ส่วนต้นทุนถั่วเฉลี่ยต่อหน่วยต่อชั่วโมงของ A

= 56/8 = ฿7

ของ B = 58/10 = ฿5.8

มีวิธีการคิดอีกวิธีโดยใช้ ค่าเฉลี่ยแบบถ่วงน้ำหนัก ดังนี้

A ใช้เวลาผลิตทั้งหมด 8 ชั่วโมง โดยมีแรงงานประเภทต่าง ๆ คือ $\frac{1}{8}$, $\frac{2}{8}$ และ $\frac{5}{8}$ ของเวลาตามลำดับ ให้ใช้สัดส่วนเหล่านี้เป็นน้ำหนัก จะได้ต้นทุน แรงงาน ถัวเฉลี่ยต่อชั่วโมงของ A ดังนี้

$(\frac{1}{8} \times ฿4) + (\frac{2}{8} \times ฿6) + (\frac{5}{8} \times ฿8) = ฿7$/ชั่วโมง

และของ B ดังนี้

$(\frac{4}{10} \times ฿4) + (\frac{3}{10} \times ฿6) + (\frac{3}{10} \times ฿8) = ฿\ 5.8$/ชั่วโมง

3. ค่าเฉลี่ยเรขาคณิต (Geometric Mean)

G.M. = $\sqrt[n]{\text{ผลคูณของค่า } x \text{ ทั้งหมด}}$

ค่าเฉลี่ยเรขาคณิตเหมาะสำหรับหาค่าเฉลี่ยของอัตราการเปลี่ยนแปลง เช่น อัตราดอกเบี้ยซึ่งเปลี่ยนแปลงเสมอ ดังในตัวอย่าง สมมุติว่า ฝากเงิน 100 บาท เป็นเวลา 5 ปี โดยไม่ถอนเลย ดอกเบี้ยจึงทบต้น และอัตราดอกเบี้ยยังเปลี่ยนแปลงทุกปี ดังนี้

ปี	อัตราดอกเบี้ย	ตัวคูณ (ดัชนีเติบโต)	เงินออมปลายปี
1	7%	1.07	107.00
2	8	1.08	115.56
3	10	1.10	127.12
4	12	1.12	142.37
5	18	1.18	168.00

จะเห็นว่า ค่าเฉลี่ยของดัชนีเติบโต = (1.07 + 1.08 1.10 1.12 - 1.18)/5 = 1.11 นั่นคือ โดยถัวเฉลี่ยใน 5 ปี จะได้รับดอกเบี้ยปีละ 11% ดังนั้น ถ้าฝาก 5 ปี จากเงินต้น 100 บาท ในปลายปีที่ 5 จะได้

$$100 \times 1.11 \times 1.11 \times 1.11 \times 1.11 \times 1.11 = 168.51$$

ซึ่งต่างจาก 168.00 เล็กน้อย ทั้งนี้เพราะวิธีนี้ใช้ค่าเฉลี่ย เลขคณิตไม่ได้ ต้องใช้ค่าเฉลี่ยเรขาคณิต

$$GM = \sqrt[5]{1.07 \times 1.08 \times 1.10 \times 1.12 \times 1.18}$$

$$= 1.1093 \qquad = 11.093\%$$

ดังนั้นเงินรวมของเงินต้น 100 บาท ในปลายปีที่ 5

$$= 100 \times 1.1093 \times 1.1093 \times 1.1093 \times 1.1093 \times 1.1093 = 167.975$$

$$\approx 168 \text{ บาท}$$

4. ค่ามัธยฐาน (Median)

ค่ามัธยฐานคือ ค่าที่อยู่กึ่งกลางของข้อมูลที่เรียงตาม**ลำดับ** ถ้า n เป็นเลขคี่ ค่ามัธยฐานจะอยู่กึ่งกลางพอดี แต่ถ้า n เป็นเลขคู่

ค่ามัธยฐาน = ข้อมูลลำดับที่ $(\frac{n+1}{2})$

ถ้าเป็นข้อมูลแบบจับกลุ่ม ต้องหาว่าค่ามัธยฐานคือ $(\frac{n+1}{2})$ อยู่ในผลรวมของความถี่ ชั้นใด แล้วใช้วิธีเทียบบัญญัติไตรยางค์หาค่ามัธยฐาน ดังตัวอย่าง

จำนวนเงิน	ความถี่	ความถี่สะสม
0 — 49.99	78	78
50.00 — 99.99	123	201
100.00 — 149.99	187	388
150.00 — 199.99	82	470
200.00 — 249.99	51	521
250.00 — 299.99	47	568
300.00 — 349.99	13	581
350.00 — 399.99	9	590
400.00 — 449.99	6	596
450.00 — 499.99	4	600
	600	

มัธยฐาน = $\frac{600+1}{2}$ = 300.5 จะอยู่รวมในชั้นที่มีความถี่สะสม 388 ซ้ำคือจำนวนเงิน 100.00 - 149.99 บาท ในชั้นนี้ มีความถี่ 187, จำนวนเงินต่อ 1 ความถี่ = $\frac{49.99}{187}$ = 0.2673 บาท ค่า 300.5 คือความถี่ลำดับที่ = 300.5 — 201 = 99.5 ในชั้นนี้ (ซึ่งมีความถี่ทั้งหมด 187) คิดเป็นจำนวนเงิน = 99.5 × 0.2673 = 26.5963 แต่ชั้นนี้มีค่าเริ่มต้นที่ 100.00 บาท ดังนั้น ค่ามัธยฐานคือจำนวนเงิน

100 + 26.5963 = 126.5963 บาท

ข้อดีและเสียของค่ามัธยฐาน

เมื่อเปรียบเทียบค่ามัธยฐานกับค่าเฉลี่ยเลขคณิต จะมีข้อได้เปรียบที่ค่ามัธยฐานจะไม่ถูกผลกระทบจากตัวเลขที่แตกต่างไปจากกลุ่มมาก ๆ เหมือนกับค่าเฉลี่ย นอกจากนั้น วิธีหาค่ามัธยฐานยังง่ายกว่า หรือแม้แต่ข้อมูลแบบจับกลุ่มซึ่งไม่มีขีดจำกัดบนหรือล่างก็หาค่ามัธยฐานได้ แต่หาค่าเฉลี่ยไม่ได้ นอกจากนี้ ยังมีข้อมูลบางชนิดซึ่งจำแนกตามคุณภาพ เช่น ความเข้มของสีก็หาค่ามัธยฐานได้แต่หาค่าเฉลี่ยเลขคณิตไม่ได้ แต่มีข้อเสียที่เมื่อนำไปใช้กับกระบวนการสถิติจะต้องใช้วิธีการซับซ้อนกว่าการใช้ค่าเฉลี่ยเลขคณิต และมีข้อเสียที่ต้องเสียเวลาจัดเรียงข้อมูลตามลำดับก่อนหาค่ามัธยฐาน โดยเฉพาะถ้าข้อมูลมีจำนวนมาก จะเสียเวลามากกว่าใช้ค่าเฉลี่ยเลขคณิต

5. ฐานนิยม (Mode)

ค่าฐานนิยมคือ ข้อมูลที่มีการซ้ำมากที่สุด (มีความถี่สูงสุด) มักใช้กับข้อมูลแบบจับกลุ่ม ฐานนิยมมีข้อดีเหมือนค่ามัธยฐาน คือ ไม่ถูกกระทบจากค่าผิดปกติ, ใช้กับข้อมูลเชิงคุณภาพได้ แต่ไม่เป็นที่นิยมใช้เหมือนกับค่าเฉลี่ยและมัธยฐาน เนื่องจากข้อมูลบางชุดอาจไม่มีฐานนิยม หรือบางชุดอาจมีฐานนิยมมากกว่า 1 อัน ซึ่งทำให้การเปรียบเทียบยุ่งยากขึ้น