



ตะลุยโจทย์คณิตศาสตร์

สถิติ และการวิเคราะห์ ข้อมูลเบื้องต้น



ระดับชั้นมัธยมศึกษาตอนปลาย

ประกอบด้วย

- ♦ สรุปเนื้อหาตามบทเรียน พร้อมตัวอย่างและแนวคิด
- ♦ เหมาะสำหรับใช้เป็นคู่มือเตรียมสอบ
และสอบเข้ามหาวิทยาลัยทุกแห่ง
ทั้งระบบรับตรงและระบบกลาง (ADMISSIONS)

เรียบเรียงโดย กวียา เนาวประทีป



โครงการตำราเรียน สำนักพิมพ์ฟิสิกส์เซ็นเตอร์

ชื่อหนังสือ	ตะลุยโจทย์คณิตศาสตร์ : สถิติ และการวิเคราะห์ข้อมูลเบื้องต้น
เรียบเรียงโดย	กวียา เนาวประทีป
ราคา	79 บาท
ISBN	978-616-306-048-8
จัดพิมพ์จำหน่ายเมื่อ	มีนาคม 2556

ติดต่อสั่งซื้อได้ที่ สำนักพิมพ์ฟิสิกส์เซ็นเตอร์

45, 47 ซอยจรัญสนิทวงศ์ 40 ถนนจรัญสนิทวงศ์ แขวงบางยี่ขัน
เขตบางพลัด กรุงเทพฯ 10700
โทร. 0-2433-7755-7, 0-2433-7704-6 แฟกซ์. 0-2433-7703

ธนาคารที่ส่งจ่าย ปณ. พระปิ่นเกล้า ในนาม หจก. สำนักพิมพ์ฟิสิกส์เซ็นเตอร์
บรรณาธิการ พรทิพย์ แพงสุด
ปก/รูปเล่ม ฝ่ายศิลปกรรม

สงวนลิขสิทธิ์ตามพระราชบัญญัติลิขสิทธิ์

เนื้อหา รูปเล่ม และภาพประกอบในหนังสือเล่มนี้เป็นลิขสิทธิ์ของสำนักพิมพ์ฟิสิกส์เซ็นเตอร์
ห้ามคัดลอก ยกเว้นได้รับอนุญาต

ข้อมูลทางบรรณานุกรมของหอสมุดแห่งชาติ

กวียา เนาวประทีป

ตะลุยโจทย์คณิตศาสตร์ : สถิติ และการวิเคราะห์ข้อมูลเบื้องต้น. - -

กรุงเทพฯ : ฟิสิกส์เซ็นเตอร์, 2556.

139 หน้า.

1. คณิตศาสตร์ (สถิติ และการวิเคราะห์ข้อมูลเบื้องต้น). I. ชื่อเรื่อง.

510

ISBN 978-616-306-048-8

website : www.physicscenter.com

e-mail : sales@physicscenter.com

คำนำ

หนังสือ **ตะลุยโจทย์คณิตศาสตร์ : สถิติ และการวิเคราะห์ข้อมูลเบื้องต้น** เล่มนี้ เป็นฉบับปรับปรุงใหม่ล่าสุดที่ผู้เขียนได้ทำการรวบรวมและปรับปรุงเนื้อหาให้ตรงตามมาตรฐานการเรียนรู้กลุ่มสาระการเรียนรู้คณิตศาสตร์ สาระการเรียนรู้เพิ่มเติม โดยมีจุดมุ่งหมายให้นักเรียนและผู้สนใจที่จะสมัครสอบเข้าศึกษาต่อในสถาบันอุดมศึกษาในระบบกลาง (แอดมิสชันส์ : ADMISSIONS) ได้ใช้ศึกษาทบทวนความรู้ความเข้าใจในเนื้อหาที่สรุปย่อด้วยภาษาที่อ่านเข้าใจง่าย รวมทั้งนำเสนอเทคนิคการวิเคราะห์โจทย์ปัญหา และตัวอย่างข้อสอบอย่างเป็นระบบ ซึ่งจะช่วยให้ผู้อ่านสามารถทำความเข้าใจได้ในเวลาที่ไม่มากนัก

ผู้เขียนหวังเป็นอย่างยิ่งว่า หนังสือเล่มนี้จะอำนวยประโยชน์ให้แก่ผู้ใช้ได้เป็นอย่างดีอีกเล่มหนึ่ง

ด้วยความปรารถนาดี

กวียา เนาวประทีป

คำนำสำนักพิมพ์

การสอบเข้ามหาวิทยาลัยเป็นจุดหมายของนักเรียน นักศึกษา ระดับชั้นมัธยมศึกษาตอนปลายแทบทุกคน และปลายทางของความมุ่งมั่นนั้น คือ การสอบได้ แต่หนทางที่จะก้าวเดินไปนั้นมีการสอบหลายวิชาเพื่อเป็นเครื่องทดสอบความรู้ความสามารถของนักเรียน ซึ่งในปัจจุบันมีการสอบในระบบกลาง (แอดมิสชันส์ : ADMISSIONS) คณิตศาสตร์ ก็เป็นอีกวิชาหนึ่งที่มีความสำคัญในการตัดสินความไฝ่ฝันของนักเรียน นักศึกษาได้ สำนักพิมพ์ฟิลิกส์เซ็นเตอร์ได้เล็งเห็นความสำคัญของการสอบในระบบนี้จึงได้จัดทำหนังสือ **ตะลุยโจทย์คณิตศาสตร์** ระดับชั้นมัธยมศึกษาตอนปลาย กลุ่มสาระการเรียนรู้คณิตศาสตร์ สถิติ และการวิเคราะห์ข้อมูลเบื้องต้น เล่มนี้ขึ้นมา เพื่อให้นักเรียน นักศึกษา ได้นำไปใช้เป็นแนวทางในการทำข้อสอบเข้าศึกษาต่อในมหาวิทยาลัยได้สะดวกยิ่งขึ้น

สำนักพิมพ์ฯ หวังเป็นอย่างยิ่งว่า หนังสือเล่มนี้คงจะอำนวยประโยชน์อย่างมากแก่นักเรียน นักศึกษาทุกคนที่มีโอกาสได้ใช้เพื่อการสอบเข้าศึกษาต่อในมหาวิทยาลัยต่อไป

ฝ่ายวิชาการ
สำนักพิมพ์ฟิลิกส์เซ็นเตอร์

สารบัญ

บทที่ 1	การวิเคราะห์ข้อมูลเบื้องต้น	6
1.1	การเก็บรวบรวมข้อมูล	6
1.2	แหล่งข้อมูล	6
1.3	การจำแนกข้อมูล	6
1.4	การนำเสนอข้อมูล	6
1.5	การแจกแจงความถี่	7
1.6	ความหมายของสัญลักษณ์ Σ	7
1.7	การวัดค่ากลางของข้อมูล	8
1.8	การวัดตำแหน่งของข้อมูล	19
1.9	การวัดการกระจายของข้อมูล	21
1.10	การวัดการกระจายสัมพัทธ์	22
1.11	ความสัมพันธ์ระหว่างตารางแจกแจงความถี่ ค่ากลาง และการกระจายของข้อมูล	23
	ตะลุยโจทย์ สถิติ และการวิเคราะห์ข้อมูลเบื้องต้น	24
บทที่ 2	การแจกแจงปกติ	95
2.1	ค่ามาตรฐาน	95
2.2	การแจกแจงปกติ และเส้นโค้งปกติ	97
	ตะลุยโจทย์ การแจกแจงปกติ	100
บทที่ 3	ความสัมพันธ์เชิงฟังก์ชันระหว่างข้อมูล	124
3.1	ความสัมพันธ์เชิงฟังก์ชันที่มีกราฟเป็นเส้นตรง	124
3.2	ความสัมพันธ์เชิงฟังก์ชันที่มีกราฟเป็นรูปพาราโบลา	124
3.3	ความสัมพันธ์เชิงฟังก์ชันที่มีกราฟเป็นรูปเอกซ์โพเนนเชียล	125
	ตะลุยโจทย์ ความสัมพันธ์เชิงฟังก์ชันระหว่างข้อมูล	125

บทที่ 1

การวิเคราะห์ข้อมูลเบื้องต้น

1.1 การเก็บรวบรวมข้อมูล

วิธีการเก็บรวบรวมข้อมูลสถิติแบ่งได้เป็น 4 วิธีใหญ่ๆ ดังนี้

1. การเก็บรวบรวมข้อมูลจากทะเบียนประวัติ
2. การเก็บรวบรวมข้อมูลจากการสำรวจ
3. การเก็บรวบรวมข้อมูลจากการทดลอง
4. การเก็บรวบรวมข้อมูลจากการสังเกต

1.2 แหล่งข้อมูล

1. แหล่งปฐมภูมิ (Primary Source) คือ ข้อมูลที่ได้จากแหล่งกำเนิดของข้อมูล
2. แหล่งทุติยภูมิ (Secondary Source) คือ ข้อมูลที่ได้จากแหล่งที่เก็บรวบรวมไว้แล้ว

1.3 การจำแนกข้อมูล

1. จำแนกตามคุณภาพ (Qualitative Classification)
2. จำแนกตามปริมาณ (Quantitative Classification)
3. จำแนกตามเวลา (Chronological Classification)
4. จำแนกตามภูมิศาสตร์ (Geographical Classification)

1.4 การนำเสนอข้อมูล

การนำเสนอข้อมูล คือการแสดงข้อมูลที่รวบรวมได้ให้ผู้สนใจได้ศึกษา เป็นการเผยแพร่ข้อมูล เพื่อนำไปวิเคราะห์ต่อไป วิธีนำเสนอข้อมูลมีวิธีการใหญ่ๆ 2 วิธี คือ

1. การนำเสนอข้อมูลอย่างไม่เป็นแบบแผน ซึ่งประกอบด้วย
 1. การนำเสนอข้อมูลในรูปของบทความ
 2. การนำเสนอข้อมูลในรูปบทความกึ่งตาราง

2. การนำเสนอข้อมูลอย่างเป็นแบบแผน ประกอบด้วย

1. การนำเสนอในรูปแบบตาราง
2. การนำเสนอในรูปแบบภูมิแท่ง
3. การนำเสนอในรูปแบบภูมิวง
4. การนำเสนอในรูปแบบภูมิรูปภาพ
5. การนำเสนอในรูปแบบที่สถิติ
6. การนำเสนอในรูปแบบของกราฟ

1.5 การแจกแจงความถี่ (Frequency Distribution)

การแจกแจงความถี่ คือ การจัดข้อมูลที่ยังไม่เป็นระเบียบ หรือข้อมูลดิบ (raw data) ให้เป็นระเบียบเพื่อความสะดวกในการวิเคราะห์ข้อมูล ซึ่งการจัดระเบียบของข้อมูลทำได้ 2 วิธี คือ

1. จัดเรียงลำดับข้อมูลหรือคะแนนทั้งหมดจากน้อยไปมาก หรือมากไปน้อยก็ได้
2. จัดเรียงลำดับข้อมูลหรือคะแนนเป็นกลุ่มและจำนวนคะแนนในแต่ละกลุ่มเรียกว่า “ความถี่”

(frequency) ช่วงของคะแนนในแต่ละกลุ่ม เรียกว่า “อันตรภาคชั้น” (class interval)

การแจกแจงความถี่เป็นแผนภูมิ

1. ฮิสโทแกรม (Histogram)
2. รูปหลายเหลี่ยมของความถี่ (Polygon Frequency)
3. โค้งของการแจกแจงความถี่ (Curve of Frequency Distribution)

1.6 ความหมายของสัญลักษณ์ Σ (Summation Notation)

1.
$$\sum_{i=1}^N X_i = X_1 + X_2 + X_3 + \dots + X_N$$
2.
$$\sum_{i=1}^N kX_i = k \sum_{i=1}^N X_i$$
3.
$$\sum_{i=1}^N (X_i \pm Y_i) = \sum_{i=1}^N X_i \pm \sum_{i=1}^N Y_i$$
4.
$$\sum_{i=1}^N k = Nk$$
5.
$$\sum_{x=1}^N x = \frac{N}{2} (N + 1)$$
6.
$$\sum_{x=1}^N x^2 = \frac{N}{6} (2N + 1) (N + 1)$$
7.
$$\sum_{x=1}^N x^3 = \left[\frac{N}{2} (N + 1) \right]^2$$

1.7 การวัดค่ากลางของข้อมูล (Measures of Central Tendency)

1. ค่าเฉลี่ยเลขคณิต (Arithmetic Mean) เขียนแทนด้วย $\bar{X}$ ถ้า $X_1, X_2, X_3, \dots, X_N$ เป็นข้อมูล หรือคะแนน N จำนวน

$$\begin{aligned}\bar{X} &= \frac{X_1 + X_2 + X_3 + \dots + X_N}{N} \\ &= \frac{\sum_{i=1}^N X_i}{N} = \frac{\Sigma X}{N}\end{aligned}$$

และถ้า $f_1, f_2, f_3, \dots, f_n$ เป็นความถี่ของข้อมูลหรือคะแนนแต่ละอันตรภาคชั้น

$$\begin{aligned}\bar{X} &= \frac{f_1 X_1 + f_2 X_2 + f_3 X_3 + \dots + f_n X_n}{N} \\ &= \frac{\sum_{i=1}^n f_i X_i}{N} = \frac{\Sigma fX}{N}\end{aligned}$$

หรือหาโดยวิธีหาค่า หรือวิธีลัด

$$\bar{X} = a + \frac{\sum_{i=1}^n f_i D_i}{N} \times I$$

เมื่อ a คือ ค่าเฉลี่ยสมมติ

I คือ ขนาดของอันตรภาคชั้น

$$D_i = \frac{X_i - a}{I}$$

ตัวอย่างที่ 1 จงหาค่าเฉลี่ยเลขคณิตของข้อมูล 10, 12, 8, 6, 14

วิธีทำ

$$\begin{aligned}\therefore \bar{X} &= \frac{10 + 12 + 8 + 6 + 14}{5} \\ &= \frac{50}{5} = 10\end{aligned}$$

ตัวอย่างที่ 2 จงหาค่าเฉลี่ยเลขคณิตจากตารางแจกแจงความถี่ต่อไปนี้

X_i	2	4	6	8	10	
f_i	1	4	10	3	2	N = 20

วิธีทำ

X_i	f_i	$f_i X_i$
2	1	2
4	4	16
6	10	60
8	3	24
10	2	20
$N = 20$		$\Sigma f_i X_i = 122$

$$\begin{aligned}\therefore \bar{X} &= \frac{\sum_{i=1}^n f_i X_i}{N} \\ &= \frac{122}{20} \\ &= 6.1\end{aligned}$$

ตัวอย่างที่ 3

จงหาค่าเฉลี่ยเลขคณิตจากตารางแจกแจงความถี่ต่อไปนี้

อันตรภาคชั้น	5-9	10-14	15-19	20-24	25-29	
ความถี่	2	3	7	5	3	$N = 20$

วิธีทำ

อันตรภาคชั้น	X_i	f_i	$f_i X_i$
5 - 9	7	2	14
10 - 14	12	3	36
15 - 19	17	7	119
20 - 24	22	5	110
25 - 29	27	3	81
		$N = 20$	$\Sigma f_i X_i = 360$

$$\begin{aligned}\therefore \bar{X} &= \frac{\sum_{i=1}^n f_i X_i}{N} = \frac{360}{20} \\ &= 18\end{aligned}$$

หรือหาโดยวิธีทอนค่า (วิธีลัด)

วิธีทำ

$$a = 17$$

$$I = 5$$

X_i	f_i	$D = \frac{X_i - a}{I}$	$f_i D_i$
7	2	-2	-4
12	3	-1	-3
17	7	0	0
22	5	1	5
27	3	2	6
$N = 20$			$\Sigma f_i D_i = 4$

$$\begin{aligned} \therefore \bar{X} &= a + \left(\frac{\Sigma f_i D_i}{N} \right) I \\ &= 17 + \left(\frac{4}{20} \right) 5 \\ &= 17 + 1 = 18 \end{aligned}$$

1. ค่าเฉลี่ยเลขคณิตรวม (Pooled mean)

เช่น ข้อมูล k ชุด จำนวนข้อมูลในแต่ละชุดเท่ากับ $N_1, N_2, N_3, \dots, N_k$ และมีค่าเฉลี่ยเลขคณิตของแต่ละชุดเท่ากับ $\bar{X}_1, \bar{X}_2, \bar{X}_3, \dots, \bar{X}_k$

$$\bar{X}_{\text{รวม}} = \frac{N_1 \bar{X}_1 + N_2 \bar{X}_2 + N_3 \bar{X}_3 + \dots + N_k \bar{X}_k}{N_1 + N_2 + N_3 + \dots + N_k}$$

ตัวอย่าง

คะแนนสอบวิชาคณิตศาสตร์ของนักเรียนหญิงห้อง 5/1 จำนวน 20 คน ได้ค่าเฉลี่ยเลขคณิต 70 คะแนน และคะแนนสอบวิชาเดียวกันของนักเรียนชายห้อง 5/1 จำนวน 30 คน ได้ค่าเฉลี่ยเลขคณิต 80 คะแนน จงหาคะแนนเฉลี่ยของการสอบวิชาคณิตศาสตร์ของนักเรียนห้อง 5/1 ทั้งหมด

วิธีทำ

$$\begin{aligned} \therefore N_{\text{ญ}} &= 20, \quad \bar{X}_{\text{ญ}} = 70 \\ N_{\text{ช}} &= 30, \quad \bar{X}_{\text{ช}} = 80 \end{aligned}$$

$$\begin{aligned} \therefore \bar{X}_{\text{รวม}} &= \frac{N_{\text{ญ}} \bar{X}_{\text{ญ}} + N_{\text{ช}} \bar{X}_{\text{ช}}}{N_{\text{ญ}} + N_{\text{ช}}} \\ &= \frac{(20)(70) + (30)(80)}{20 + 30} \end{aligned}$$

$$\begin{aligned}
 &= \frac{1,400 + 2,400}{50} \\
 &= \frac{3,800}{50} \\
 &= 76 \text{ คะแนน}
 \end{aligned}$$

∴ คะแนนเฉลี่ยของการสอบวิชาคณิตศาสตร์ของนักเรียนห้อง 5/1 = 76 คะแนน

สมบัติเฉพาะของ $\bar{X}$

1. ถ้าค่าของข้อมูลแต่ละตัวของข้อมูลชุดใดมีค่าเท่ากับ $\bar{X}$ จะมีค่าเท่ากับค่าของข้อมูลแต่ละตัวนั่นเอง

เช่น ถ้าข้อมูลทุกตัวของข้อมูลชุดหนึ่ง = c

∴ $\bar{X}$ ของข้อมูลชุดนั้น = c

2. ถ้าข้อมูล $X_1, X_2, X_3, \dots, X_n$ มีค่าเฉลี่ยเลขคณิตเท่ากับ M แล้ว ข้อมูล $aX_1 + b, aX_2 + b, aX_3 + b, \dots, aX_n + b$ มีค่าเฉลี่ยเลขคณิตเท่ากับ $aM + b$ เมื่อ a, b เป็นจำนวนจริง

3. ผลบวกทางพีชคณิตของข้อมูลแต่ละตัวที่แตกต่างไปจากค่าเฉลี่ยเลขคณิต จะมีค่าเท่ากับศูนย์เสมอ กล่าวคือ

$$\sum_{i=1}^N (x_i - \bar{X}) = 0$$

4. ผลบวกของกำลังสองของข้อมูลแต่ละตัวลบด้วยค่าคงที่ที่มีค่าน้อยที่สุด เมื่อค่าคงที่นั้นเป็นค่าเฉลี่ยเลขคณิต กล่าวคือ

$$\sum_{i=1}^N (x_i - k)^2 \text{ จะมีค่าน้อยที่สุด เมื่อ } k = \bar{X}$$

2. มัธยฐาน (Median)

มัธยฐาน คือค่าของข้อมูลที่มีตำแหน่งอยู่กึ่งกลางของข้อมูลทั้งหมด เมื่อเรียงข้อมูลนั้นตามลำดับ

ตัวอย่างที่ 4

มัธยฐานของข้อมูล 10, 11, 13, 15, 17 คือ 13 และมัธยฐานของ 50, 52, 55, 56, 57, 60

$$\text{คือ } \frac{55 + 56}{2} = 55.5$$

ถ้าเป็นข้อมูลที่มีการแจกแจงความถี่ เราหามัธยฐานโดยใช้สูตร

$$\text{มัธยฐาน} = L + \left(\frac{\frac{N}{2} - \Sigma f_L}{f_M} \right) I$$

L เป็นขอบชั้นล่างของอันตรภาคชั้นมัธยฐาน

$\frac{N}{2}$ คือตำแหน่งมัธยฐานของข้อมูล

Σf_L เป็นผลรวมของความถี่ของทุกอันตรภาคชั้นที่เป็นช่วงคะแนนต่ำกว่าชั้นที่มีมัธยฐานอยู่

I ความกว้างของอันตรภาคชั้นที่เป็นมัธยฐาน

f_M เป็นความถี่ของชั้นที่มีมัธยฐานอยู่

ตัวอย่างที่ 5 จงหามัธยฐานจากตารางแจกแจงความถี่ต่อไปนี้

คะแนน	5	8	9	10	12	
ความถี่	3	4	6	5	2	N = 20

วิธีทำ

คะแนน	ความถี่	ความถี่สะสม
5	3	3
8	4	7
9	6	13
10	5	18
12	2	20
	N = 20	

$$\text{ตำแหน่งมัธยฐาน คือตำแหน่งที่ } \frac{N + 1}{2} = \frac{20 + 1}{2} = 10.5$$

$$\therefore \text{มัธยฐาน} = 9$$

ตัวอย่างที่ 6 จงหามัธยฐานจากตารางแจกแจงความถี่ต่อไปนี้

อันตรภาคชั้น	5-9	10-14	15-19	20-24	25-29	
ความถี่	3	4	7	3	3	N = 20

วิธีทำ

ขอบชั้น	ความถี่	ความถี่สะสม
4.5 - 9.5	3	3
9.5 - 14.5	4	7
14.5 - 19.5	7	14
19.5 - 24.5	3	17
24.5 - 29.5	3	20
	N = 20	

ตำแหน่งมัธยฐานคือ $\frac{N}{2} = \frac{20}{2} = 10$ ซึ่งอยู่ในขอบชั้น 14.5 - 19.5

$$\therefore L = 14.5, \Sigma f_L = 7, f_M = 7, I = 5$$

$$\begin{aligned} \therefore \text{มัธยฐาน} &= L + \left(\frac{\frac{N}{2} - \Sigma f_L}{f_M} \right) I \\ &= 14.5 + \left(\frac{10 - 7}{7} \right) \times 5 \\ &= 14.5 + \frac{15}{7} \\ &= 16.6 \end{aligned}$$

สมบัติของมัธยฐาน

สมบัติที่สำคัญข้อหนึ่งของมัธยฐานคือผลรวมของค่าสัมบูรณ์ของผลต่างระหว่างข้อมูลแต่ละค่า

กับมัธยฐานของข้อมูลชุดนั้นจะมีค่าน้อยที่สุด กล่าวคือ $\sum_{i=1}^N |X_i - \text{มัธยฐาน}|$ มีค่าน้อยที่สุด

นั่นคือ $\sum_{i=1}^N |X_i - a|$ จะมีค่าน้อยที่สุดเมื่อ a คือ ค่ามัธยฐานของข้อมูลชุดนั้น

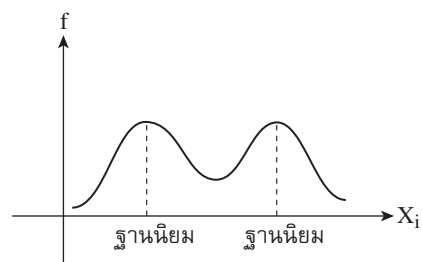
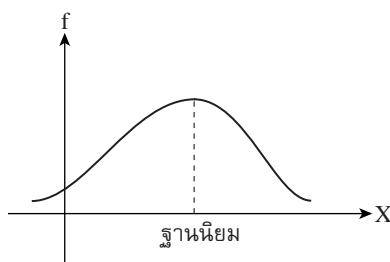
3. ฐานนิยม (Mode)

ฐานนิยม คือค่าของข้อมูลที่มีความถี่สูงสุด

ตัวอย่างที่ 7

ฐานนิยมของข้อมูล 5, 6, 6, 6, 7, 8, 7, 9, 6, 6 คือ 6 เพราะว่าข้อมูล 10 ตัวนี้ เป็น 6 เสีย 5 ตัว ซึ่งถือว่ามีความถี่สูงสุด

สำหรับข้อมูลที่มีการแจกแจงความถี่แล้ว เมื่อเขียนเป็นเส้นโค้งของความถี่ของข้อมูลที่แจกแจงความถี่แล้ว ฐานนิยม คือ ค่าของ X ณ จุดสูงสุดบนเส้นโค้งความถี่ แต่ถ้าเส้นโค้งความถี่มีจุดสูงสุดสองจุด ข้อมูลชุดนั้นจะมีฐานนิยม 2 ค่า



สำหรับการคำนวณหาฐานนิยมของข้อมูลที่มีการแจกแจงความถี่แล้วนั้นคำนวณได้โดยการหาจุดกึ่งกลางของอันตรภาคชั้นที่มีฐานนิยมอยู่ ค่าที่ได้จะเป็นค่าของฐานนิยมโดยประมาณ