

บทที่ 13

การวิเคราะห์ความแปรปรวนแบบทางเดียว (ONE WAY ANALYSIS OF VARIANCE)

วัตถุประสงค์

เมื่อท่านศึกษาเนื้อหาบทที่ 13 โดยละเอียดแล้ว ควรมีความสามารถ ดังนี้

1. บอกสาเหตุของการนำการวิเคราะห์ความแปรปรวนมาใช้ในการทดสอบเกี่ยวกับค่าเฉลี่ยของประชากรได้
2. วิเคราะห์ข้อมูลทางการศึกษาโดยใช้วิธีวิเคราะห์ความแปรปรวนแบบทางเดียวได้

เนื้อหา

13.1 บทนำ

13.2 การวิเคราะห์ความแปรปรวนแบบทางเดียว

13.2.1 ข้อตกลงเบื้องต้น

13.2.2 สัญลักษณ์

13.2.3 การตั้งสมมุติฐาน

13.2.4 สถิติที่ใช้ในการทดสอบสมมุติฐาน

13.2.5 ขอบเขตวิกฤต

13.2.6 ตารางแสดงผลการวิเคราะห์ความแปรปรวน

เนื้อหาที่ 13.1 บทนำ

การวิเคราะห์ความแปรปรวน (ANOVA) หรือ F-test เป็นสถิติที่ใช้ในการทดสอบความแตกต่างของค่าเฉลี่ยของกลุ่มตัวอย่างหลาย ๆ กลุ่ม เราเรียกการทดสอบแบบนี้ว่า F-test ในบทนี้ผู้เขียนจะอธิบายถึงหลักการเบื้องต้นของการวิเคราะห์ความแปรปรวนแบบทางเดียวพอเป็นสังเขปเพื่อเป็นพื้นความรู้ที่จะเรียนต่อไปเท่านั้น นักศึกษาที่ประสงค์จะเรียนรู้เกี่ยวกับการวิเคราะห์ความแปรปรวนให้ลึกซึ้งอาจจะศึกษารายละเอียดได้จากหนังสืออ้างอิง ในบรรณานุกรมท้ายเล่มของตำราเล่มนี้ หรือศึกษาจากกระบวนวิชาต่อจากกระบวนวิชานี้

ในบทก่อนได้อธิบายถึงการใช้ Z-test หรือ t-test สำหรับทดสอบความแตกต่างระหว่างค่าเฉลี่ยของผลงานของกลุ่มตัวอย่าง 2 กลุ่ม อย่างไรก็ตามถ้าเราต้องการเปรียบเทียบค่าเฉลี่ยของผลงานของกลุ่มตัวอย่างมากกว่า 2 กลุ่มขึ้นไป เราไม่นิยมใช้ Z - test หรือ t - test แต่จะมาใช้ F - test แทน ทั้งนี้เนื่องจากสาเหตุสำคัญดังนี้

1) ถ้าใช้ t - test (หรือ Z - test) ทดสอบ แล้วค่าของระดับนัยสำคัญทางสถิติ (∞ - error) จะเพิ่มขึ้นจากที่กำหนดไว้ ซึ่งจะมีผลทำให้ความเชื่อถือได้ของผลการวิจัยจะลดน้อยลง

2) ปัญหาเชิงปฏิบัติอย่างหนึ่งในการทดสอบนัยสำคัญของความแตกต่างคือ ต้องเสียเวลามาก นั่นคือถ้ามีกลุ่มตัวอย่าง 3 กลุ่ม ก็ต้องทดสอบ 3 คู่ กลุ่มตัวอย่าง 6 กลุ่ม ต้องทดสอบ 10 คู่ เป็นต้น ทั้งนี้อาจเป็นไปได้ว่า ไม่มีความแตกต่างคู่ใดเลยที่ต่างกันอย่างมีนัยสำคัญทางสถิติ ในสถานการณ์เช่นนี้ถ้าจะมีวิธีทดสอบกลุ่มตัวอย่างทั้งหมดพร้อมกันว่าจะมีคู่ใดแตกต่างกันอย่างมีนัยสำคัญทางสถิติก็จะเป็นการดีถ้าผลการทดสอบระบุว่าไม่มีความแตกต่าง การทดสอบก็จบเพียงเท่านั้นไม่ต้องเสียเวลามาก แต่ถ้าผลการทดสอบระบุว่ามีความแตกต่าง เราจึงจะทำการทดสอบว่ามีคู่ใดบ้างที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ดังนั้นการวิเคราะห์ความแปรปรวน หรือ F-test จึงถูกนำมาใช้ในการวิเคราะห์ข้อมูลด้วยสาเหตุดังกล่าวข้างต้น

เนื้อหาที่ 13.2 การวิเคราะห์ ความแปรปรวนแบบทางเดียว

ในกรณีที่มีกลุ่มตัวอย่าง 3 กลุ่มขึ้นไป และนักวิจัยต้องการดูว่ามีความแตกต่างของค่าเฉลี่ยระหว่างกลุ่มหรือไม่ สถิติที่นิยมใช้คือการวิเคราะห์ความแปรปรวนแบบทางเดียวซึ่งมีชื่อเรียกอย่างย่อ ๆ ว่า ANOVA (อ่านว่า แอน-โน-วา) หรือ F-test

13.2.1 ข้อตกลงเบื้องต้น

การใช้ F-test มีข้อตกลงเบื้องต้น ดังนี้

1. กลุ่มตัวอย่างภายในแต่ละกลุ่ม และระหว่างกลุ่มต้องเป็นอิสระจากกัน
2. การแจกแจงของประชากรเป็นแบบโค้งปกติ
3. ความแปรปรวนของประชากรแต่ละกลุ่มต้องเท่ากัน

13.2.2 สัญลักษณ์

ให้ j ใช้แทนกลุ่มของการทดลอง
 i แทนคนในกลุ่ม

การวิเคราะห์ความแปรปรวนแบบทางเดียว

n_j แทนจำนวนกลุ่มตัวอย่างในกลุ่ม j

X_{ij} แทนคะแนนหรือผลที่เกิดจากการกระทำของคน i ในกลุ่ม j

$\sum_{i=1}^n X_{ij}$ แทนผลบวก (Σ) ของผลที่ได้จากการทดลอง (X_{ij}) จากตัวแรก ($i = 1$) ถึงตัวสุดท้าย (n_j) ในกลุ่ม j ตัวอย่างเช่น ถ้าผลที่ได้จากการทดลองในกลุ่ม j มี 3 ตัว คือ 4, 2, 5 แล้ว

$$\sum_{i=1}^n X_{ij} \quad \text{จะเท่ากับ } 4 + 2 + 5 = 11 \quad \text{และ } n_j = 3$$

N แทนจำนวนกลุ่มตัวอย่างทั้งหมด

S_j^2 แทนความแปรปรวนของกลุ่ม j

จุด (.) แทนการรวมข้อมูล สัญลักษณ์จุด (.) จะถูกใช้แทนที่ i หรือ j เพื่อเป็นการบ่งบอกว่าให้รวม i และ/หรือ j เข้าด้วยกัน

$\bar{X}_{.j}$ แทนค่าเฉลี่ยของกลุ่มที่ j

$\bar{X}_{..}$ แทนค่าเฉลี่ยของกลุ่มตัวอย่างทั้งหมด (Grand mean)

13.2.3 การตั้งสมมติฐาน

สมมติฐานเพื่อใช้ในการทดสอบโดยการวิเคราะห์ความแปรปรวนแบบทางเดียวเขียนได้เป็นรูปแบบเดียวกันดังนี้

H_0 : $\mu_1 = \mu_2 = \mu_3 = \dots = \mu_j$

H_1 : มีค่า μ อย่างน้อย 1 ตัวที่แตกต่างจาก μ ตัวอื่น ๆ

เป็นที่น่าสังเกตว่า ถ้าเราใช้ t -test หรือ Z -test ในการทดสอบแล้ว สมมติฐานกลางที่ตั้งจะมีลักษณะดังนี้

$$H_0 : \mu_1 = \mu_2$$

$$H_0 : \mu_1 = \mu_3$$

$$H_0 : \mu_1 = \mu_4$$

$$H_0 : \mu_{j-1} = \mu_j$$

13.2.4 สถิติที่ใช้ในการทดสอบสมมุติฐาน

สถิติที่ใช้ในการทดสอบสมมุติฐานเมื่อใช้วิธีวิเคราะห์ความแปรปรวนแบบทางเดียว คือ

$$F = \frac{MS_B}{MS_W}, df = J-1, N-J \quad \dots\dots(13.1)$$

เมื่อ $MS_B = SS_B/df_B$

และ $MS_W = SS_W/df_W$

เมื่อชั้นแห่งความอิสระระหว่างกลุ่ม (df_B) = J-1 และชั้นแห่งความอิสระภายในกลุ่ม (df_W) = N-J

ดังนั้น $MS_B = SS_B/J-1$

และ $MS_W = SS_W/N-J$

ในกรณีนี้ $SS_B =$ Sum Square Between Group
 $=$ ผลบวกของกำลังสองของผลต่างระหว่างค่าเฉลี่ย
 ของแต่ละกลุ่มกับค่าเฉลี่ยของกลุ่มตัวอย่างทั้งหมด
 $= \sum_i n_i (\bar{X}_i - \bar{X})^2$

การวิเคราะห์ความแปรปรวนแบบทางเดียว

$$= \sum_j \frac{(\sum_i X_{ij})^2}{n_j} - \frac{(\sum_j \sum_i X_{ij})^2}{N}$$

ส่วน

$SS_w =$ Sum Square Withim Group

$=$ ผลบวกของกำลังสองของผลต่างระหว่างคะแนนแต่ละตัวกับค่าเฉลี่ยของแต่ละกลุ่ม

$$= \sum_j \sum_i (X_{ij} - \bar{X}_{.j})^2$$

$$= \sum_j \sum_i X_{ij}^2 - \sum_j \frac{(\sum_i X_{ij})^2}{n_j}$$

และเนื่องจาก $SS_B + SS_w = SS_T$

ดังนั้นถ้าต้องการตรวจสอบความถูกต้องในการคำนวณหาค่า SS_B กับ SS_w ก็อาจคำนวณหาค่า SS_T (Sum Square Total) มาเปรียบเทียบกับ โดยคำนวณได้จากสูตร

$SS_T =$ ผลบวกกำลังสองของผลต่างระหว่างคะแนนแต่ละตัวกับค่าเฉลี่ยของกลุ่มตัวอย่างทั้งหมด

$$= \sum_j \sum_i (X_{ij} - \bar{X}_{..})^2$$

$$= \sum_j \sum_i X_{ij}^2 - \frac{(\sum_j \sum_i X_{ij})^2}{N}$$

13.2.5 ขอบเขตวิกฤต

เราจะปฏิเสธสมมุติฐานกลาง (H_0) ถ้าค่า F ที่คำนวณได้จากสูตร 13.1 ($F = MS_B/MS_w$) มีค่ามากกว่า $F_{1-\alpha, df = J-1, N-J}$

จากตารางในภาคผนวก J-1 คือ V_1 และ $N-J$ คือ V_2

13.2.6 ตารางแสดงผลการวิเคราะห์ความแปรปรวน

เมื่อกำหนดค่า MS_B และ MS_W ได้แล้วให้นำรายละเอียดทั้งหมดมาแสดงไว้ในตารางวิเคราะห์ความแปรปรวนตามแบบแผนที่กำหนดไว้ดังตาราง 13.1 ข้างล่างนี้

ตาราง 13.1 ตารางแสดงผลการวิเคราะห์ความแปรปรวน

แหล่งของความแปรปรวน	SS	df	MS	F
ระหว่างกลุ่ม		J-1	$SS/J-1$	$\frac{MS_B}{MS_W}$
ภายในกลุ่ม		N-J	$SS_W/N-J$	
รวม		N-1		

ตัวอย่าง 13.1 ในการทดลองใช้วิธีสอน 4 แบบ สอนวิชาคณิตศาสตร์กับนักเรียนชั้นป. 3 กลุ่มละ 1 แบบในเนื้อหาเรื่องการคูณ เมื่อเสร็จสิ้นการสอนได้ทำการทดสอบความรู้ของนักเรียนเรื่องการคูณทั้ง 4 กลุ่ม ผลการทดสอบแสดงไว้ในตารางข้างล่าง อยากทราบว่า การสอนทั้ง 4 แบบให้ผลการเรียนที่แตกต่างกันหรือไม่ ผลการทดสอบมีดังนี้

กลุ่มที่ 1	2	3	4
7	6	4	3
3	10	2	4
4	8	2	5
3	5	1	4
6		2	
		1	

การวิเคราะห์ความแปรปรวนแบบทางเดียว

n_j	5	4	6	4
$\Sigma_i X_{ij}$	23	29	12	16
$(\Sigma_i X_{ij})^2$	529	841	144	256
$\frac{(\Sigma_i X_{ij})^2}{n_j}$	105.80	210.25	24.00	64.00
$\bar{X}_{.j}$	$\frac{23}{5}$ (= 4.6)	$\frac{29}{4}$ (= 7.25)	$\frac{12}{6}$ (= 2.0)	$\frac{16}{4}$ (= 4.0)

$$\Sigma_j \Sigma_i X_{ij}^2 = 7^2 + 3^2 + 4^2 + 3^2 + \dots + 4^2 + 5^2 + 4^2 = 440.00$$

$$\begin{aligned} \sum_j \frac{(\Sigma_i X_{ij})^2}{n_j} &= \frac{(23)^2}{5} + \frac{(29)^2}{4} + \frac{(12)^2}{6} + \frac{(16)^2}{4} \\ &= 105.80 + 210.25 + 24.0 + 64.0 = 404.05 \end{aligned}$$

$$\frac{(\Sigma_j \Sigma_i X_{ij})^2}{N} = \frac{80^2}{19} = \frac{6400}{19} = 336.84$$

$$\bar{X}_{..} = \frac{\Sigma_j \Sigma_i X_{ij}}{N} = \frac{23 + 29 + 12 + 16}{19} = \frac{80}{19} = 4.21$$

การคำนวณหาค่า SS_B

ค่า SS_B คำนวณได้จากสูตรใดสูตรหนึ่งต่อไปนี้

$$1) \quad SS_B = \Sigma_j \frac{(\Sigma_i X_{ij})^2}{n_j} - \frac{(\Sigma_j \Sigma_i X_{ij})^2}{N}$$

$$2) \quad SS_B = \Sigma_j n_j (\bar{X}_{.j} - \bar{X}_{..})^2$$

ในที่นี้จะแสดงวิธีหาค่า SS_B จากทั้ง 2 สูตร เพื่อเปรียบเทียบให้เห็นจริง ดังนี้

$$1) \quad SS_B = \Sigma_j \frac{(\Sigma_i X_{ij})^2}{n_j} - \frac{(\Sigma_j \Sigma_i X_{ij})^2}{N}$$

$$= 404.05 - 336.84$$

$$= 67.21$$

$$2) \quad SS_B = \sum_j n_j (\bar{X}_j - \bar{X})^2$$

$$= 5(4.6 - 4.21)^2 + 4(7.25 - 4.21)^2 + 6(2.0 - 4.21)^2 + 4(4.0 - 4.21)^2$$

$$= 5(0.39)^2 + 4(3.04)^2 + 6(-2.21)^2 + 4(-0.21)^2$$

$$= 5(0.15) + 4(9.24) + 6(4.88)$$

$$= 0.75 + 36.96 + 29.28 + 0.16$$

$$= 67.15$$

การหาค่า SS_B โดยวิธีที่ 1 และวิธีที่ 2 จะได้ค่าใกล้เคียงกัน แต่เนื่องจากวิธีที่ 2 มีความยุ่งยากในการคำนวณ เพราะมีจุดทศนิยมเข้ามาเกี่ยวข้อง และถ้ามีข้อมูลมาก ๆ การปัดค่าของทศนิยม อาจจะทำให้ค่าที่คำนวณได้คลาดเคลื่อนไป จึงมีผู้นิยมหา SS_B โดยวิธีที่ 1 มากกว่าวิธีที่ 2

การคำนวณหาค่า SS_W

ค่า SS_W คำนวณได้จากสูตรใดสูตรหนึ่งต่อไปนี้

$$1) \quad SS_W = \sum_j \sum_i X_{ij}^2 - \sum_j \frac{(\sum_i X_{ij})^2}{n_j}$$

$$2) \quad SS_W = \sum_j \sum_i (\bar{X}_{ij} - \bar{X}_j)^2$$

ในที่นี้จะแสดงวิธีหาค่า SS_W จากทั้ง 2 สูตร ให้ดูดังนี้

$$1) \quad SS_W = \sum_j \sum_i X_{ij}^2 - \sum_j \frac{(\sum_i X_{ij})^2}{n_j}$$

การวิเคราะห์ความแปรปรวนแบบทางเดียว

$$= 440 - 404.05$$

$$= 35.95$$

$$2) \quad SS_w = \sum_j \sum_i (X_{ij} - \bar{X}_{.j})^2$$

$$= (7-4.6)^2 + (3-4.6)^2 + (4-4.6)^2 + (3-4.6)^2 + (6-4.6)^2 +$$

$$(6-7.25)^2 + (10-7.25)^2 + (8-10.25)^2 + (5-7.25)^2 + (4-2)^2 +$$

$$(2-2)^2 + (2-2)^2 + (1-2)^2 + (2-2)^2 + (1-2)^2 + (3-4)^2 +$$

$$(4-4)^2 + (5-4)^2 + (4-4)^2$$

$$= (2.4)^2 + (-1.6)^2 + (0.6)^2 + (-1.6)^2 + (1.4)^2 + (-1.25)^2 + (2.75)^2$$

$$+ (2.25)^2 + (2.25)^2 + 2^2 + 0^2 + 0^2 + (-1)^2 + (0)^2 + (-1)^2 +$$

$$(-1)^2 + 0^2 + 1^2 + 0^2$$

$$= 5.76 + 2.56 + 0.36 + 2.56 + 1.96 + 1.56 + 7.56 + 5.06 + 5.06 +$$

$$4 + 0 + 0 + 1 + 0 + 1 + 1 + 0 + 1 + 0$$

$$= 40.44$$

จากตัวอย่างที่ 13.1 แสดงให้เห็นว่าในการคำนวณหาค่า SS_w โดยวิธีที่ 2 จะได้ค่าที่คลาดเคลื่อนไปจากการหาค่า SS_w โดยวิธีที่ 1 พอประมาณ ทั้งนี้เนื่องจากการปัดทศนิยม ดังนั้นในกรณีที่มีกลุ่มตัวอย่างมาก ๆ จึงควรใช้การคำนวณโดยวิธีที่ 1 เพราะจะมีความคลาดเคลื่อนน้อยกว่า

อนึ่งจากตัวอย่างนี้ จะยึดตัวเลขที่ได้จากการหาค่า SS_B และ SS_w โดยวิธีที่ 1 ในการวิเคราะห์ข้อมูล

การหาค่า MS_B และ MS_w

$$MS_B = \frac{SS_B}{J-1} = \frac{67.21}{3} = 22.40$$

$$MS_w = \frac{SS_w}{N-J} = \frac{35.95}{15} = 2.40$$

แสดงค่าตัวเลขในตารางการวิเคราะห์ความแปรปรวนได้ดังนี้

ตาราง 13.1 แสดงผลการวิเคราะห์ความแปรปรวน

แหล่งของความแปรปรวน	SS	df	MS	F
ระหว่างกลุ่ม (B)	67.21	3	22.40	9.33
ภายในกลุ่ม (W)	35.95	15	2.40	
ทั้งหมด (T)	103.16	18		

$$F_{1-\alpha, df = J-1, N-J} = F_{.95, df = 3, 15} = 3.29$$

ดังนั้นจึงปฏิเสธสมมุติฐานกลาง ($H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$) เพราะค่า F ที่คำนวณได้มากกว่า $F_{.95, df = 3, 15}$

แสดงว่าค่าเฉลี่ยของบางกลุ่มแตกต่างไปจากกลุ่มอื่น ๆ นั่นคือมีวิธีสอนอย่างน้อย 1 วิธีให้ผลสัมฤทธิ์ทางการเรียน แตกต่างจากวิธีอื่น

สรุปเนื้อหาบทที่ 13

1. ในการทดสอบเพื่อหาความแตกต่างของค่าเฉลี่ยของประชากรเมื่อมีกลุ่มตัวอย่างตั้งแต่ 3 กลุ่มขึ้นไป สถิติที่นิยมใช้คือ การวิเคราะห์ความแปรปรวนหรือการทดสอบแบบเอฟ (F - test)
2. ข้อตกลงเบื้องต้นของการใช้ F - test คือ
 - 1) กลุ่มตัวอย่างภายในแต่ละกลุ่ม และระหว่างกลุ่มต้องเป็นอิสระจากกัน
 - 2) การแจกแจงของประชากรเป็นแบบโค้งปกติ
 - 3) ความแปรปรวนของประชากรแต่ละกลุ่มต้องเท่ากัน

$$3. F = \frac{MS_B}{MS_W}, df = J-1, N-J$$

$$\text{เมื่อ } MS_B = SS_B/J-1$$

$$\text{และ } MS_W = SS_W/N-J$$

$$4. SS_B = \sum_j \frac{(\sum_i X_{ij})^2}{n_j} - \frac{(\sum_j \sum_i X_{ij})^2}{N}$$

$$5. SS_W = \sum_j \sum_i X_{ij}^2 - \sum_j \frac{(\sum_i X_{ij})^2}{n_j}$$

คำถามท้ายบทที่ 13

1. ในการทดลองเกี่ยวกับวิธีสอน แบบต่าง ๆ 4 แบบกับนักเรียนกลุ่มละ 6 คนให้ผลการเรียนดังตารางข้างล่างอยากทราบว่าวิธีสอน 4 แบบให้ผลการเรียนเหมือนกันหรือไม่ที่ระดับนัยสำคัญทางสถิติ .05

ผลการทดลองเป็นดังนี้

กลุ่มที่ 1	กลุ่ม 2	กลุ่ม 3	กลุ่ม 4
64	73	77	78
72	61	83	91
68	90	97	97
77	80	69	82
56	97	79	85
95	67	87	77

การวิเคราะห์ความแปรปรวนแบบทางเดียว

2. ในการทดลองเกี่ยวกับการเรียนรู้ กลุ่มตัวอย่าง 10 คน ถูกกำหนดให้อยู่ในแต่ละกลุ่ม การทดลอง 3 กลุ่ม อย่างสุ่ม แต่ละกลุ่มจะต้องปฏิบัติงานแบบเดียวกัน แต่กำหนดให้ สภาพการทดลองต่างกัน อยากทราบว่าค่าเฉลี่ยของผลการทดลองของกลุ่มแตกต่างกันหรือไม่

กลุ่มที่	1	2	3
	41	40	36
	40	36	33
	39	40	29
	41	34	30
	39	34	45
	41	39	39
	36	36	33
	35	3 4	32
	35	41	34
	37	37	34