การวิเคราะห์ความแปรปรวนแบบทางเดียว

5

ในบทที่แล้วเราได้กล่าวถึงการเปรียบเทียบความแตกต่างระหว่างค่าเฉลี่ย 2 ค่าโดยใช้ t-test ที่แตกต่างกันในแต่ละแบบ แต่ถ้าหากมีค่าเฉลี่ยมากกว่า 2 ค่า การวิเคราะห์ด้วย t-test ทำให้เสียเวลาและเกิดความคลาดเคลื่อนได้มาก ดังนั้นจึงมีสถิติอีกตัวหนึ่งที่สามารถแก้ปัญหานี้ได้ ชื่อว่า การวิเคราะห์ความแปรปรวนแบบทางเดียว (One-way ANOVA)

ใช้คำสั่ง oneway ในการวิเคราะห์ความแปรปรวนแบบทางเดียว เป็นสถิติใช้สำหรับวิเคราะห์ความแตกต่างระหว่างค่าเฉลี่ยตั้งแต่ 2 ค่าขึ้นไป โดยจะต้องมีตัวแปรตามมีระดับการวัดอยู่ในระดับ Interval Scale และตัวแปรอิสระมีเพียงตัวเดียวอยู่ในระดับ Nominal Scale แบ่งออกเป็น k ระดับ โดยมีข้อตกลงเบื้องต้นในการทดสอบ One-way ANOVA ดังนี้

1. กลุ่มตัวอย่างแต่ละกลุ่มจะต้องสุ่มมาจากประชากรที่มีการแจกแจงเป็นปกติ
2. กลุ่มตัวอย่างแต่ละกลุ่มจะต้องสุ่มมาจากประชากรที่ความแปรปรวนไม่แตกต่างกัน
3. หน่วยสมาชิกในกลุ่มตัวอย่างแต่ละหน่วยจะต้องสุ่มมาอย่างอิสระ
4. กลุ่มตัวอย่างแต่ละกลุ่มจะต้องเป็นอิสระจากกัน

จากข้อตกลงเบื้องต้นกลุ่มตัวอย่างจะต้องสุ่มมาจากประชากรที่มีความแปรปรวนไม่แตกต่างกัน เราสามารถทดสอบได้ด้วยสถิติต่อไปนี้

1. Hartley F_{max} Test สำหรับทดสอบความแปรปรวนของกลุ่มตัวอย่างเพียง 2 กลุ่ม ซึ่งแต่ละกลุ่มจะมีจำนวนตัวอย่างเท่ากัน

สมมติฐาน

$$H_0 : \sigma_1^2 = \sigma_2^2$$
$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

สูตรคำนวณ

$$F_{max} = \frac{S^2_{max}}{S^2_{min}} \quad ; \quad df = n - 1$$

ในสูตรนี้ ความแปรปรวนของกลุ่มตัวอย่าง 2 กลุ่ม กลุ่มใดมีค่ามากที่สุดให้เป็นตัวตั้ง และหารด้วยอีกกลุ่มหนึ่ง หากค่าที่คำนวณ มากกว่าค่าจากตาราง จะปฏิเสธ H_0

2. Cochran Test สำหรับทดสอบความแปรปรวนของกลุ่มตัวอย่างหลายกลุ่ม ซึ่งกลุ่มตัวอย่างแต่ละกลุ่มจะมีจำนวนตัวอย่างเท่ากันหรือไม่เท่ากันก็ได้

สมมติฐาน

$H_0 : \sigma_1^2 = \sigma_2^2 = ... = \sigma_k^2$

H_1 : มี σ^2 อย่างน้อย 1 คู่ที่แตกต่างกัน

สูตรคำนวณ

$$C = \frac{S_{max}^2}{\Sigma S_i^2} \quad ; \quad df = n - 1$$

ถ้าค่า C ที่คำนวณมากกว่า C ตาราง จะปฏิเสธ H_0

3. Bartlett Test สำหรับทดสอบความแปรปรวนของกลุ่มตัวอย่างหลายกลุ่ม เหมาะสำหรับกลุ่มตัวอย่างที่มีขนาดใหญ่ และจะมีการแจกแจงคล้ายการแจกแจงของไคสแคร์

สมมติฐาน

$H_0 : \sigma_1^2 = \sigma_2^2 = ... = \sigma_k^2$

H_1 : มี σ^2 อย่างน้อย 1 คู่ที่แตกต่างกัน

สูตรคำนวณ

$$B = \frac{1}{C}[(N-k)\ln(\frac{SS_W}{N-k}) - \sum_{j=1}^{k}(n_j - 1)\ln S_j^2]$$

เมื่อ $N = \sum_{i=1}^{k} n_j$

$$C = 1 + \frac{1}{3(k-1)}[\sum_{j=1}^{k}(\frac{1}{n_j - 1}) - \frac{1}{N-k}]$$

$$SS_W = \Sigma(n_j - 1)S_j^2$$

หากค่า B ที่คำนวณมากกว่าค่าจากตารางไคสแควร์ ที่ df = n – 1 จะปฏิเสธ H_0

ในการวิเคราะห์ความแปรปรวนแบบทางเดียวมีสมมติฐานในการทดสอบดังนี้

สมมติฐาน

$H_0 : m_1 = m_2 ... m_N$ (เมื่อ n คือจำนวนกลุ่มตัวอย่าง)

H_1 : มีค่าเฉลี่ยอย่างน้อย 1 คู่ที่แตกต่างกัน

สูตรทดสอบ

$$F = \frac{MS_{BG}}{MS_{WG}}$$

df = k – 1 และ N – k

ค่าของ MS_{BG} และ MS_{WG} สามารถคำนวณได้จากสูตร

$$MS_{BG} = \frac{SS_{BG}}{p-1} \qquad MS_{WG} = \frac{SS_{WG}}{p(n-1)}$$

บรรดาค่า Sum of Square (SS) ต่าง ๆ สามารถคำนวณได้ด้วยสูตรดังนี้

$$SS_{TO} = \sum_{j=1}^{p}\sum_{i=1}^{n} Y_{ij}^2 - \frac{\left(\sum_{j=1}^{p}\sum_{i=1}^{n} Y_{ij}\right)^2}{np}$$

$$SS_{BG} = \sum_{j=1}^{p}\frac{\left(\sum_{i=1}^{n} Y_{ij}\right)^2}{n} - \frac{\left(\sum_{j=1}^{p}\sum_{i=1}^{n} Y_{ij}\right)^2}{np}$$

$$SS_{WG} = \sum_{j=1}^{p}\sum_{i=1}^{n} Y_{ij}^2 - \sum_{j=1}^{p}\frac{\left(\sum_{i=1}^{n} Y_{ij}\right)^2}{n}$$

จากนั้นนำค่าที่คำนวณได้ใส่ตารางวิเคราะห์ความแปรปรวน

แหล่งความแปรปรวน	SS	df	MS	F
1. Between Groups	SS_{BG}	p - 1	MS_{BG}	$\frac{MS_{BG}}{MS_{WG}}$
2. Within Group	SS_{WG}	p(n - 1)	MS_{WG}	
3. Total	SS_{TO}	np - 1		

หากค่าสถิติ F ที่คำนวณได้มีค่ามากกว่า F จากตาราง จะปฏิเสธ H_0 ยอมรับ H_1 แสดงว่ามีค่าเฉลี่ยอย่างน้อย 1 คู่ที่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

หากต้องการทราบว่ามีคู่ใดบ้างที่แตกต่างกันให้ดำเนินการเปรียบเทียบพหุคูณต่อไป ซึ่งมีวิธีการหลายวิธีดังนี้

1. วิธี Least significance difference
2. วิธี Duncan's multiple-range test
3. วิธี Student-Newman-Keuls test
4. วิธี Turkey's alternate test
5. วิธี Scheffe's test

ฯลฯ

วิธีการคำนวณการเปรียบเทียบพหุคูณเหล่านี้ สามารถศึกษาได้จากหนังสือสถิติทั่วไป

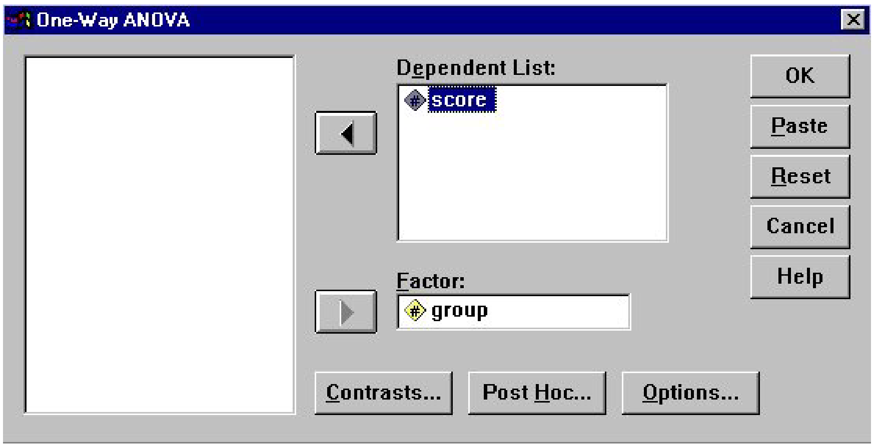
ตัวอย่าง 5.1

สุ่มนักเรียนมา 30 คน แล้วแบ่งออกเป็น 3 กลุ่ม แต่ละกลุ่มได้รับตัวแปรทดลองต่างกัน ผลของการทดลอง มีดังนี้

กลุ่ม 1	4	6	8	6	7	8	10	8	6	11
กลุ่ม 2	14	13	10	14	7	10	13	16	10	18
กลุ่ม 3	7	8	10	7	8	10	11	12	14	15

ทำการทดสอบความแตกต่างของกลุ่มตัวอย่างทั้ง 3 กลุ่ม

ใช้เมนูหลัก “Analyze” เมนูรอง “Compare Means” และเมนูย่อย “One-Way ANOVA” จะเกิดหน้าต่าง



ภาพประกอบ 5.1

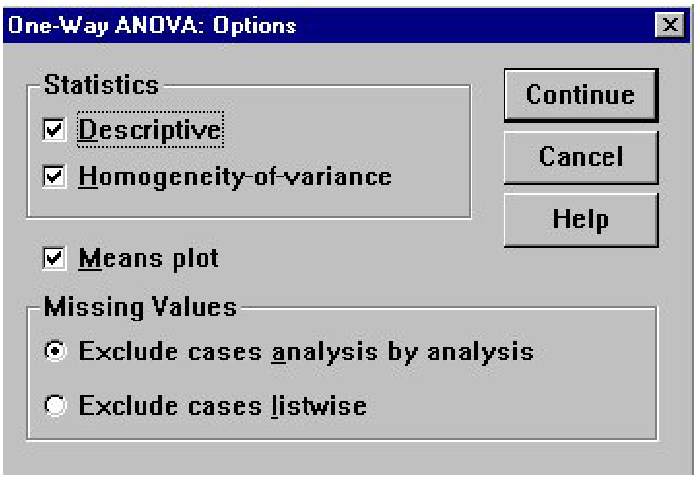
ให้เลือกตัวแปรตามใส่ช่อง “Dependent List :” และตัวแปรอิสระใส่ช่อง “Factor :” ในที่นี้ตัวแปรตามคือตัวแปร “score” และตัวแปรอิสระคือ “group”

หากต้องการเปรียบเทียบพหุคูณคลิกที่ปุ่ม “Post Hoc...” แล้วเลือกวิธีการเปรียบเทียบดังภาพประกอบ

ภาพประกอบ 5.2

ในที่นี้เลือกการเปรียบเทียบพหุคูณด้วยวิธี “Scheff” จากนั้นคลิกปุ่ม “Continue” และคลิกปุ่ม “OK” โปรแกรมแสดงผลลัพธ์ที่คำนวณได้ในหน้าต่าง “Output”

หากต้องการค่าสถิติพื้นฐานของตัวแปรและทดสอบความเป็นเอกพันธ์ของความแปรปรวนให้คลิกที่ปุ่ม “Options...” หากต้องการให้โปรแกรมสร้างกราฟแสดงค่าเฉลี่ยของแต่ละกลุ่มก็คลิกที่ข้อความ “Means plot” ดังภาพประกอบ



ภาพประกอบ 5.3

สำหรับการทดสอบความเป็นเอกพันธ์ของความแปรปรวนนั้น จะแสดงการทดสอบความเป็นเอกพันธ์ด้วยสูตรของ Levene ตัวอย่างผลลัพธ์มีดังนี้

Descriptives

SCORE

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
1.00	10	7.4000	2.0656	.6532	5.9224	8.8776	4.00	11.00
2.00	10	12.5000	3.2745	1.0355	10.1576	14.8424	7.00	18.00
3.00	10	10.2000	2.8206	.8919	8.1823	12.2177	7.00	15.00
Total	30	10.0333	3.4087	.6223	8.7605	11.3062	4.00	18.00

Test of Homogeneity of Variances

SCORE

Levene Statistic	df1	df2	Sig.
1.097	2	27	.348

ANOVA

SCORE

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	130.467	2	65.233	8.529	.001
Within Groups	206.500	27	7.648		
Total	336.967	29			

Multiple Comparisons

Dependent Variable: SCORE

Scheffe

(I) GROUP	(J) GROUP	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-5.1000*	1.2368	.001	-8.3033	-1.8967
	3.00	-2.8000	1.2368	.096	-6.0033	.4033
2.00	1.00	5.1000*	1.2368	.001	1.8967	8.3033
	3.00	2.3000	1.2368	.197	-.9033	5.5033
3.00	1.00	2.8000	1.2368	.096	-.4033	6.0033
	2.00	-2.3000	1.2368	.197	-5.5033	.9033

*. The mean difference is significant at the .05 level.

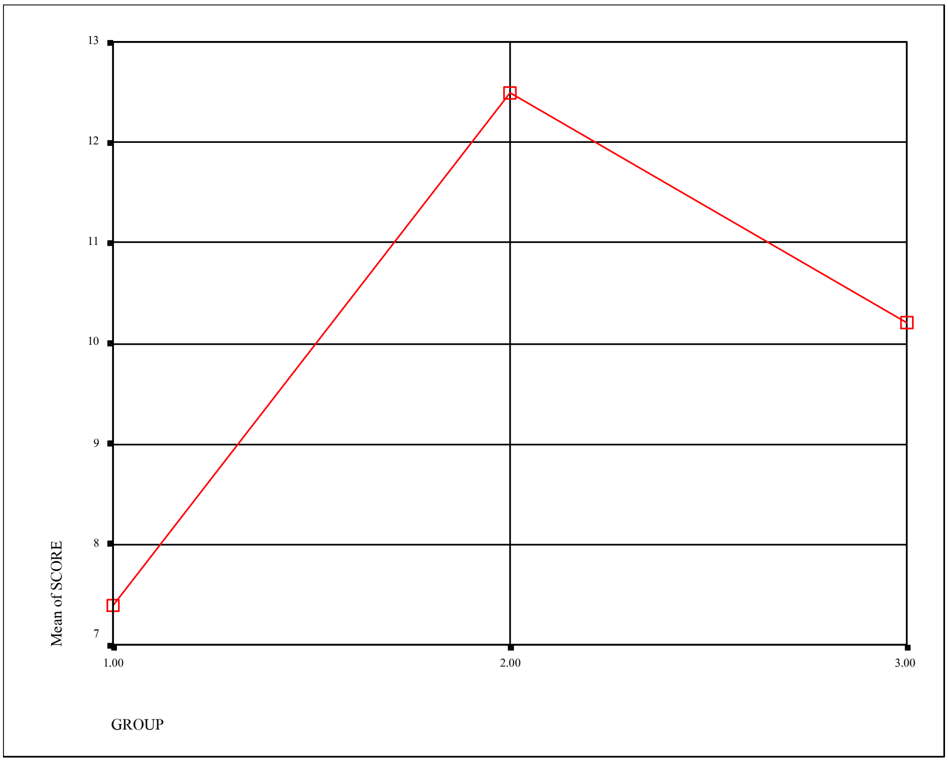
SCORE

Scheffe[a]

GROUP	N	Subset for alpha = .05	
		1	2
1.00	10	7.4000	
3.00	10	10.2000	10.2000
2.00	10		12.5000
Sig.		.096	.197

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 10.000.



ภาพประกอบ 5.4

ผลการวิเคราะห์ความแปรปรวนแบบทางเดียวปรากฏว่าค่า F-test คำนวณได้ 8.5293 มีนัยสำคัญทางสถิติที่ .001 ซึ่งน้อยกว่า .01 แสดงว่าทั้ง 3 กลุ่มมีค่าเฉลี่ยแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 จึงทำการเปรียบเทียบพหุคูณด้วยวิธีเชฟเฟ่ ผลปรากฏว่า กลุ่มที่ 1 และกลุ่มที่ 2 แตกต่างกันอย่างมีนัยสำคัญทางสถิติ โดยกลุ่มที่ 2 มีคะแนนเฉลี่ยสูงกว่ากลุ่มที่ 1

สำหรับการทดสอบความเป็นเอกพันธ์ของความแปรปรวนนั้นพบว่าไม่มีนัยสำคัญทางสถิตินั่น คือความแปรปรวนของกลุ่มทั้ง 3 มีความเป็นเอกพันธ์กัน ซึ่งเป็นไปตามข้อตกลงเบื้องต้นของการ วิเคราะห์ความแปรปรวน