

ทฤษฎีการใช้สถิตินอนพารามेटริก

ลาวัณย์ วิจารณ์^{1,2*}

ทฤษฎี (TIPS) หมายถึง คำแนะนำอันเป็นประโยชน์ ในที่นี้หมายความว่า คำแนะนำอันเป็นประโยชน์ ในการใช้สถิตินอนพารามेटริก (nonparametric statistics : ขอใช้ชื่อย่อว่า nparstat : เอ็นพาร์สแตท)

ผู้ทำการวิจัยที่เรียกว่า “วิทยานิพนธ์” มักจะพบความสับสนในเรื่องของการใช้สถิติทดสอบ (test statistics) ที่เรียกว่า parametric statistics (ขอใช้ชื่อย่อว่า parastat: พาร์ราสแตท) ในการทดสอบ สมมติฐานโดย parastat ที่นิยมใช้กันมาก ได้แก่ t-test, z-test, F-test (Analysis of Variance: ANOVA) และ χ^2 (เป็นทั้ง parastat และ nparstat)

ความสับสนที่สำคัญคือ ข้อตกลง 2 ประการ ในการที่จะใช้ parastat กล่าวคือ ชุดข้อมูลที่จะนำมา วิเคราะห์นั้น ต้องมีลักษณะของ normal distribution (การแจกแจงหรือการกระจายแบบปกติหรือ normal curve) และมีลักษณะของ randomness (ความสุ่ม)

ดังนั้นก่อนการตัดสินใจว่าจะใช้ parastat ต้องตรวจสอบก่อนว่า ชุดข้อมูลมีลักษณะดังกล่าวหรือไม่ เพราะจะส่งผลให้ ผลการวิเคราะห์ผิดพลาดบิดเบือนไปจากความจริงได้ ทำให้นักวิจัยนั้นไม่เกิดประโยชน์

นักสถิติได้คิดค้นวิธีการทดสอบที่ทำได้ง่าย โดยใช้ความรู้ nparstat ดังนี้ nparstat Kolmogorov-Smirnov test (K-S test) ทดสอบว่าชุดข้อมูลมีคุณสมบัติ normal distribution หรือไม่ และ nparstat Runs test ทดสอบ randomness

nparstat ได้ถูกพัฒนาขึ้น เพื่อใช้กับข้อมูลที่ได้จากการวัดในระดับ nominal scale, ordinal scale แต่ถ้า ข้อมูลเป็น interval scale ต้องแปลง (transform) เป็น nominal หรือ ordinal scale เสียก่อน

nparstat ไม่มีข้อตกลงสำคัญในเรื่อง nominal distribution และ randomness แต่อย่างใด ลักษณะการกระจายแบบ free distribution (การแจกแจงในรูปแบบใดก็ได้: ไม่อิงกับ normal distribution) จึงเหมาะกับการใช้กับชุดข้อมูลขนาดเล็ก ซึ่งสามารถใช้ทดสอบชุดข้อมูลที่มีจำนวนเท่ากับหรือน้อยกว่า 10 ได้

วิธีการของ nparstat เป็นการหา critical region (เขตวิกฤต) หรือ rejection region (เขตแห่ง การปฏิเสธ) และ level of significance: (ระดับนัยสำคัญของการทดสอบทางสถิติ) จากชุดข้อมูลโดยตรง โดยเทียบกับตาราง (Table) ของ nparstat นั้นๆ ซึ่งได้จัดทำไว้แล้ว เป็นการเฉพาะในรูปแบบต่างๆกัน

¹ ภาควิชาชีววิทยา คณะวิทยาศาสตร์ มหาวิทยาลัยรังสิต

² ที่ปรึกษาสถาบันการอาชีวศึกษาเกษตรภาคใต้

*ผู้รับผิดชอบประสานงาน (Corresponding author) e-mail: lawan.w@rsu.ac.th

Kolmogorov-Smirnov test (K-S test)

การใช้ nparstat K-S test เพื่อจะแสดงว่าชุดข้อมูลสังเกตการวัดระดับ interval scale ที่เป็นตัวเลขที่มีอยู่นั้น มาจากมวลชุดข้อมูลสังเกตทั้งหมดที่มีการกระจายแบบปกติหรือการแจกแจงเป็นโค้งปกติหรือไม่

วิธีการทำ

1. แปลง (transform) ข้อมูลดิบ (x) แต่ละตัวให้เป็นคะแนนมาตรฐาน (z) เมื่อ $\bar{x}$ คือ ค่าเฉลี่ยเลขคณิต N คือ จำนวนข้อมูลทั้งหมดของชุดนั้น s คือ ค่าเบี่ยงเบนมาตรฐาน คำนวณโดยใช้สูตร

$$z = \frac{x - \bar{x}}{s} \quad \text{สูตรที่ 1}$$

$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N} \quad \text{สูตรที่ 2}$$

$$s = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N - 1}} \quad \text{สูตรที่ 3}$$

2. เรียงลำดับค่า z จากน้อยไปมาก

3. หาค่าความน่าจะเป็นสะสมของทุกๆค่า z จากตารางซี (F_c) และสัดส่วนพื้นที่ภายใต้เส้นโค้งปกติ เรียกค่านี้ว่า $F_0(X)$ (ลัวิน สายยศ และอังคณา สายยศ, 2538)

4. หาค่าความถี่สะสมสัมพัทธ์ที่คาดหวัง (F_c)

5. หาค่าสูงสุดของความแตกต่างระหว่าง $|F_c - F_0(X)| = D_{cal.}$

หรือ $D_{cal.} = \text{maximum } |F_c - F_0(X)|$

6. เปิดตาราง E)Table E. Table of Critical Value of D in Kolmogorov-Smirnov One sample Test (Siegel,1956)) แสดงค่าวิกฤต)critical value ของ (D ที่ N , $\alpha = .05$ หรือ $\alpha = .01$

7. การตัดสินใจ

ถ้า $D_{cal.} \geq D_{tab.}$ ที่ระดับ α) level of significance) ที่กำหนด = reject H_0 (ปฏิเสธ H_0)

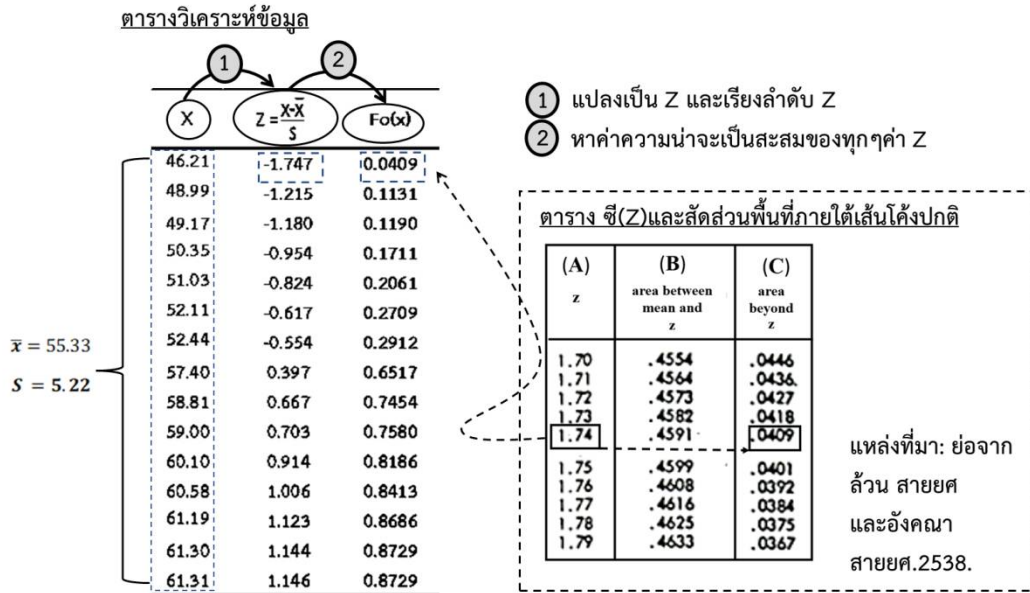
หมายความว่า ชุดข้อมูลสังเกตชุดนี้ไม่มีคุณสมบัติของการแจกแจงเป็นโค้งปกติ

ถ้า $D_{cal.} < D_{tab.}$ ที่ระดับ α) level of significance) ที่กำหนด = accept H_0 (ยอมรับ H_0)

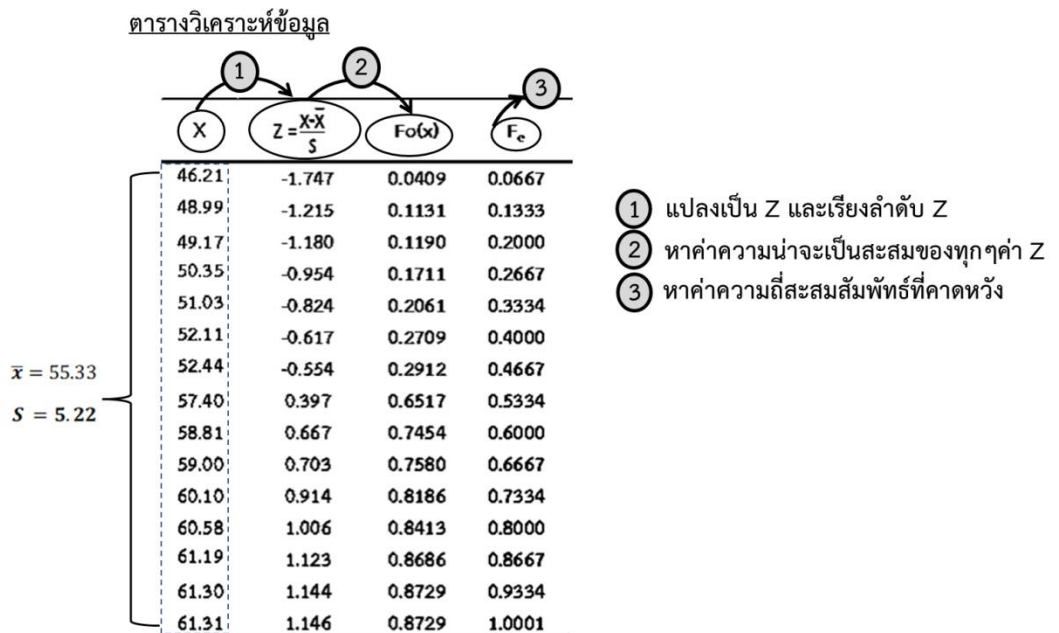
หมายความว่า ชุดข้อมูลสังเกตชุดนี้มีคุณสมบัติของการแจกแจงเป็นโค้งปกติ

ตัวอย่าง

ข้อมูลสังเกตระดับ interval scale ชุดหนึ่ง เรียงลำดับค่า x ดังแสดงในตารางวิเคราะห์ข้อมูลและตารางซี (Z)

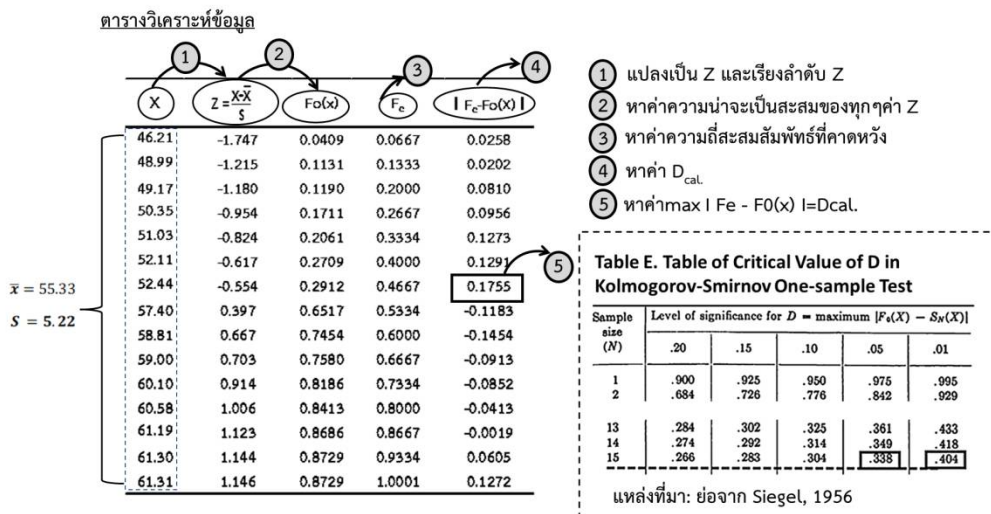


หาค่าความถี่สะสมสัมพัทธ์ที่คาดหวัง (F_e) ดังตารางวิเคราะห์ข้อมูล



หาค่าสูงสุดของความแตกต่างระหว่าง $|F_e - F_o(X)| = \max |F_e - F_o(X)| = \text{ค่า } D_{cal}.$

เปิดตาราง E (Siegel, 1956) $N=15$, $\alpha = .05$ และ $\alpha = .01$ พบว่า $D_{tab.} = 0.338$ ที่ระดับ $\alpha = .05$ และ $D_{tab.} = 0.404$ ที่ระดับ $\alpha = .01$ ตามลำดับ



จะเห็นได้ว่า

$$D_{cal.} = 0.1755 < D_{tab.} = 0.338 \text{ ที่ระดับ } \alpha = .05 = \text{accept } H_0$$

$$D_{cal.} = 0.1755 < D_{tab.} = 0.404 \text{ ที่ระดับ } \alpha = .01 = \text{accept } H_0$$

หมายความว่า ชุดข้อมูลสังเกตชุดนี้มีคุณสมบัติของการแจกแจงเป็นโค้งปกติ

Runs test

การใช้ nparstat Runs test เพื่อจะแสดงว่า ชุดข้อมูลสังเกตวัดระดับ interval scale ที่เป็นตัวเลขที่มีอยู่นั้น มีคุณสมบัติของการสุ่ม (randomness) หรือไม่

วิธีการทำ

1. พิจารณาข้อมูลสังเกตแต่ละตัวของชุดข้อมูลสังเกตนั้นว่า มีค่ามากกว่าหรือน้อยกว่าค่ามัธยฐาน (Median: Md) ของข้อมูลสังเกตชุดนั้น

2. ถ้ามีค่ามากกว่า Md ให้เครื่องหมาย (+) ถ้าน้อยกว่า Md ให้เครื่องหมายลบ (-)

3. นับจำนวนการเกิดสลับกันของเครื่องหมาย + และ - เรียกจำนวนนี้ว่า ค่า runs: $r(r_{cal})$

4. นับจำนวนเครื่องหมาย + ให้เป็น n_1 และนับจำนวนเครื่องหมาย - ให้เป็น n_2 (ถ้าเครื่องหมายเป็น 0 สุ่มให้อยู่ในกลุ่มเครื่องหมาย + หรือ -)

5. การทดสอบนัยสำคัญ แบ่งเป็น 2 กรณี

5.1 n_1 และ $n_2 < 20$

นำค่า n_1 และ n_2 ไปเปิดตาราง F (Table F. Table of Critical Values of r in the Runs Test: Siegel) 2 ตาราง คือ Table F_1 และ Table F_2 หาค่า critical value ของ r

ถ้าค่า $r_{cal.} \leq r_{tab.}$ ที่ได้จาก Table F_1 และ $r_{cal.} \leq r_{tab.}$ ที่ได้จาก Table F_2 จะ reject H_0 ที่ $\alpha = .05$ และถ้าค่า $r_{cal.}$ มีค่าอยู่ระหว่าง critical value ของ r ที่ได้จาก Table F_1 และ F_2 จะ accept H_0 ที่ $\alpha = .05$

5.2 กรณี $n_1 < 20$ หรือ $n_2 < 20$

ให้หาค่า Z (คะแนนมาตรฐาน) ดังนี้

$$Z_{cal.} = \frac{\text{จำนวน } r \text{ ที่นับได้} - \text{ค่าเฉลี่ยของ } r}{\text{ส่วนเบี่ยงเบนมาตรฐานของ } r} \quad \text{สูตรที่ 4}$$

เมื่อ ค่าเฉลี่ยของ r คือ $\bar{r}$ สามารถคำนวณได้จาก

$$\bar{r} = \frac{2n_1n_2}{n_1 + n_2} + 1 \quad \text{สูตรที่ 5}$$

และ ส่วนเบี่ยงเบนมาตรฐาน r คือ S_r สามารถคำนวณได้จาก

$$S_r = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}} \quad \text{สูตรที่ 6}$$

การตัดสินใจ

ถ้าค่า $Z_{cal.}$ > critical value ของ Z ที่เปิดจาก Table A (Siegel) หรือถ้าค่า $Z_{cal.}$ > 1.96 หรือ 2.58 จะ reject H_0 ที่ $\alpha = .05$ และ $\alpha = .01$ ตามลำดับ (Siegel, 1956)

ถ้าค่า $Z_{cal.}$ < critical value ของ Z ที่เปิดจาก Table A (Siegel) หรือถ้าค่า $Z_{cal.}$ < 1.96 หรือ 2.58 จะ accept H_0 ที่ $\alpha = .05$ และ $\alpha = .01$ ตามลำดับ (Siegel, 1956)

สรุป: ถ้า accept H_0 หมายความว่า ชุดข้อมูลสังเกตชุดนี้มีคุณสมบัติของการสุ่ม (randomness)

ตัวอย่าง

กรณี n_1 และ $n_2 < 20$ ข้อมูลสังเกตระดับ interval scale ชุดหนึ่ง คือ 49.17, 50.35, 60.58, 59.00, 52.11, 51.03, 58.81, 52.44, 61.19, 61.31, 60.10, 48.99, 46.21, 57.40 และ 61.30

วิธีการทำ

ทำตารางวิเคราะห์ข้อมูล

ข้อมูลสังเกต	เครื่องหมาย	r	ข้อมูลสังเกต	เครื่องหมาย	r
49.17	—	r_1	61.19	+	r_6
50.35	—		61.31	+	
60.58	+	r_2	60.10	+	r_7
59.00	+		48.99	—	
52.11	—	r_3	46.21	—	r_8
51.03	—		57.40	0	
58.81	+	r_4	61.30	+	
52.44	—	r_5			

$Md = 57.40$, $r_{cal.} = 8$, $n_1 = 8$, $n_2 = 7$

เปิดตาราง $F_1(8,7)$ ค่า $r_{tab.} = 4$ เปิดตาราง $F_2(8,7)$ ค่า $r_{tab.} = 13$ ดังภาพ

Table F. table of Critical Value of r in the Runs Test

Table F ₁													
$n_2 \backslash n_1$	2	3	4	5	6	7	8	9	10	11	12	13	
2													
3													
4													
5													
6													
7													
8													
9													

Table F ₂													
$n_2 \backslash n_1$	2	3	4	5	6	7	8	9	10	11	12	13	
2													
3													
4													
5													
6													
7													
8													
9													

(แหล่งที่มา: ย่อจาก Siegel, 1956)

การตัดสินใจ

ค่า $r_{cal.}(8)$ มีค่าอยู่ระหว่าง critical value ของ r ที่ได้จาก Table $F_1(4)$ และ $F_2(13) = accept H_0$ ที่ $\alpha = .05$ หมายความว่า ข้อมูลสังเกตชุดนี้มีคุณสมบัติของการสุ่ม (randomness)

ตัวอย่าง

กรณี $n_1 > 20$ หรือ $n_2 > 20$ ข้อมูลสังเกตระดับ interval scale ชุดหนึ่งดังแสดงในตาราง

วิธีการทำ

ทำตารางวิเคราะห์ข้อมูล

ข้อมูลสังเกต	เครื่องหมาย	r	ข้อมูลสังเกต	เครื่องหมาย	r	ข้อมูลสังเกต	เครื่องหมาย	r
49.17	-	r ₁	49.17	-	r ₉	49.17	-	r ₁₇
50.35	-		50.35	-		50.35	-	
60.58	+	r ₂	60.58	+	r ₁₀	60.56	+	r ₁₈
59.00	+		59.00	+		59.00	+	
52.11	-	r ₃	52.11	-	r ₁₁	52.11	-	r ₁₉
51.03	-		51.03	-		51.03	-	
58.81	+	r ₄	58.81	+	r ₁₂	58.81	+	r ₂₀
52.44	-	r ₅	52.44	-	r ₁₃	52.44	-	r ₂₁
61.19	+	r ₆	61.19	+	r ₁₄	61.19	+	r ₂₂
61.31	+		61.31	+		61.31	+	
60.10	+		60.10	+		60.10	+	
48.99	-	r ₇	48.99	-	r ₁₅	48.99	-	r ₂₃
46.21	-		46.21	-		46.21	-	
57.40	+	r ₈	57.40	+	r ₁₆	57.40	+	r ₂₄
61.30	+		61.30	+		61.30	+	

$Md=57.40$, $r = 24$, $n_1 = 24$, $n_2 = 21$

หาค่า Z (คะแนนมาตรฐาน) โดยเริ่มจากคำนวณค่าเฉลี่ยของ r ($\bar{r}$) จากสูตรที่ ดังนี้ 5

$$\bar{r} = \frac{2n_1n_2}{n_1 + n_2} + 1 = \frac{2(24)(21)}{24 + 21} + 1 = 23.4$$

คำนวณส่วนเบี่ยงเบนมาตรฐาน r คือ S_r จากสูตรที่ ดังนี้ 6

$$\begin{aligned} S_r &= \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}} \\ &= \sqrt{\frac{2(24)(21)\{2(24)(21) - 24 - 21\}}{(24 + 21)^2(24 + 21 - 1)}} \\ &= 3.3007 \end{aligned}$$

คำนวณหาค่า Z (คะแนนมาตรฐาน) โดยการแทนค่า $\bar{r} = 23.4$ $S_r = 3.3007$ และ $r = 24$ ลงในสูตรที่ 4 ดังนี้

$$Z_{cal.} = \frac{\text{จำนวน } r \text{ ที่นับได้} - \text{ค่าเฉลี่ยของ } r}{\text{ส่วนเบี่ยงเบนมาตรฐานของ } r}$$

เปิดตาราง A

Table A. Table Probabilities Associated with Values as Extreme as Observed Values of z in the Normal distribution

The body of the table gives one-tailed probabilities under H_0 of z . The left-hand marginal column gives various values of z to one decimal place. The top row gives various values to the second decimal place. Thus, for example, the one-tailed p of $z \geq .11$ or $z \leq -.11$ is $p = .4562$.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036

(แหล่งที่มา: ย่อจาก Siegel, 1956)

การตัดสินใจ

$$Z_{cal.=0.182} < 2.58 = \text{accept } H_0 \text{ ที่ } \alpha = .01$$

$$Z_{cal.=0.182} < 1.96 = \text{accept } H_0 \text{ ที่ } \alpha = .05$$

หมายความว่า ชุดข้อมูลสังเกตชุดนี้มีคุณสมบัติของการสุ่ม (randomness)

ในการวางแผนการทดลองทางการเกษตร แต่ละสิ่งทดลอง (treatment) ต้องมีการทดลองซ้ำ (replication) ของหน่วยทดลอง (experimental unit) หน่วยทดลองต้องมีความสม่ำเสมอ และจำนวนเพียงพอที่จะใช้ parastat ในการวิเคราะห์ความแตกต่างระหว่างค่าเฉลี่ยของสิ่งทดลอง

การวางแผนการทดลองจึงเป็นการทดสอบความแตกต่างของ k (จำนวน) treatments นั้นเอง มีแบบแผนการทดลองทางการเกษตรที่ใช้เป็นพื้นฐาน ในการทำวิจัยเพื่อเป็นวิทยานิพนธ์ มีชื่อว่า การวางแผนแบบสุ่มสมบูรณ์ (Completely Randomized Design: CRD) ซึ่งจะใช้ parastat One-way Analysis of Variance หรือ F-test ในการวิเคราะห์ความแตกต่างระหว่างค่าเฉลี่ยของ 3 หรือมากกว่าของสิ่งทดลอง

การใช้ F-test นี้ มีข้อที่ต้องคำนึง คือ ค่า df (degree of freedom: องศาอิสระ) ของแหล่งตัวแปรความคลาดเคลื่อน ควรมากกว่า 12 เพราะทำให้ค่า F ที่คำนวณมีค่าเพิ่มมากขึ้น ทำให้ประสิทธิภาพในการทดสอบความแตกต่างเพิ่มขึ้นด้วย

การที่ df > 12 นี้ แสดงว่าถ้าเป็นการทดลอง 3 สิ่งทดลอง ต้องใช้หน่วยทดลองอย่างน้อยเท่ากับ 16 ($N-k=16-3=13$) ซึ่งเป็นการยากที่จะหาให้ได้จำนวนเพียงพอและมีลักษณะคล้ายคลึงกัน (homogeneous) โดยเฉพาะหน่วยทดลองที่เป็นสัตว์ใหญ่และพืชที่หายาก รวมทั้งขนาดของสถานที่ทดลองที่เหมาะสมอีกด้วย ดังนั้นจำนวนหน่วยทดลอง ที่มีความสม่ำเสมอ (homogeneity) จึงเป็นข้อจำกัดสำคัญของ CRD

ถ้าผู้ทำวิจัยเพื่อเป็นวิทยานิพนธ์ประสบปัญหา ก็ควรใช้ nparstat ที่เหมาะสมกับ CRD ได้แก่ The Kruskal-Wallis 1-way Analysis of Variance by Ranks: K-W test ซึ่งจะนำเสนอเพียงสังเขปต่อไปนี้

The Kruskal-Wallis One-Way Analysis of Variance by Ranks

nparstat นี้เรียกขื่อย่อว่า K-W test ใช้ทดสอบสมมติฐานว่า ชุดข้อมูลสังเกตระดับการวัด ordinal scale k ชุด (3 ชุด) มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติหรือไม่?

วิธีการทำ

1. แปลงข้อมูลสังเกตของแต่ละชุดจาก interval scale เป็น ordinal scale (ลำดับ: rank: r)
2. หา n_i ของแต่ละชุดข้อมูลสังเกต
3. หา $N = n_1 + n_2 + \dots + n_i$
4. หา (sum of r ของแต่ละชุดข้อมูลสังเกต) $= \sum r_i$
5. หา (sum of r ของชุดข้อมูลสังเกต)²หารด้วย n ของชุดข้อมูลนั้น $= \frac{\sum r_i^2}{n_i}$
6. แทนค่าที่คำนวณได้จากข้อที่ 2 – ข้อที่ 5 ลงในสูตรที่ 7 เพื่อหา $\chi_{cal.}^2$ ดังนี้

$$\chi_{cal.}^2 = \left[\left(\frac{\sum r_i^2}{n_i} \right) \left(\frac{12}{N(N+1)} \right) - 3(N+1) \right] \quad \text{สูตรที่ 7}$$

7. เปิดตาราง $\chi^2_{tab.}$ ที่ $df = k-1$

8. การตัดสินใจ

$\chi^2_{cal.} \geq \chi^2_{tab.}$ ที่ระดับ α (level of significance) ที่กำหนด = reject H_0 (ปฏิเสธ H_0)

หมายความว่า ชุดข้อมูล 3 ชุด แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ที่ระดับ α ที่กำหนด

$\chi^2_{cal.} < \chi^2_{tab.}$ ที่ระดับ α (level of significance) ที่กำหนด = accept H_0 (ยอมรับ H_0)

หมายความว่า ชุดข้อมูล 3 ชุด ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ที่ระดับ α ที่กำหนด

ตัวอย่างที่ 1

ข้อมูลสังเกตระดับ interval scale 3 ชุดข้อมูล ดังแสดงในตารางวิเคราะห์ข้อมูล ตารางวิเคราะห์ข้อมูล

ชุดข้อมูล 1	rank (r)	ชุดข้อมูล 2	rank (r)	ชุดข้อมูล 3	rank (r)
63	19	52	9	62	18
64	20	57	13	55	11
65	21	46	3	58	14
50	7	44	1	47	4
61	17	53	10	59	15
51	8	49	6	60	16
45	2			48	5
				56	12
$n_1=7$	$r_1=94$	$n_2=6$	$r_2=42$	$n_3=8$	$r_3=95$

จากตาราง $N = n_1 + n_2 + n_3 = 21$, $\sum r_1 = 94$, $\sum r_2 = 42$, $\sum r_3 = 95$

หา (sum of r ของชุดข้อมูลสังเกต)²หารด้วย n ของชุดข้อมูลนั้น จะได้

$$\frac{\sum r_i^2}{n_i} = \frac{(94)^2}{7} + \frac{(42)^2}{6} + \frac{(95)^2}{8} = 2684.42$$

แทนค่า $N = 21$ และ $\frac{\sum r_i^2}{n_i} = 2684.42$ ลงในสูตรที่ 7 เพื่อหา $\chi^2_{cal.}$ ดังนี้

$$\begin{aligned} \chi^2_{cal.} &= \left[\left(\frac{\sum r_i^2}{n_i} \right) \left(\frac{12}{N(N+1)} \right) - 3(N+1) \right] \\ &= \left[(2684.42) \left(\frac{12}{21(21+1)} \right) - 3(21+1) \right] = 3.73 \end{aligned}$$

เปิดตาราง C (Chi Square), $df = k-1$, $= 3-1=2$, $\alpha = .05$ ได้ค่า $\chi^2_{tab.} = 5.99$ พบว่า

$\chi^2_{cal.} < \chi^2_{tab.}$ ที่ระดับ $\alpha = .05 = \text{accept } H_0$

หมายความว่า ชุดข้อมูล 3 ชุด ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ที่ระดับ $\alpha = .05$

Table C. Table of Critical Values of Chi Square

df	Probability under H_0 that $\chi^2 \geq$ chi square												
	.99	.98	.95	.90	.80	.70	.60	.50	.40	.30	.20	.10	.05
1	.00016	.00083	.0039	.016	.064	.15	.46	1.07	1.64	2.71	3.84	5.02	5.41
2	.02	.04	.10	.21	.45	.71	1.39	2.41	3.22	4.60	5.99	7.38	7.88
3	.12	.18	.35	.58	1.00	1.42	2.37	3.64	4.64	6.25	7.82	9.35	9.84
4	.30	.43	.71	1.06	1.65	2.20	3.36	4.88	5.99	7.78	9.49	11.67	12.84
5	.55	.75	1.14	1.61	2.34	3.00	4.35	6.06	7.29	9.24	11.07	13.89	15.09

(แหล่งที่มา: ย่อจากSiegel, 1956.)

การทดสอบความแตกต่างของ 3 ชุดข้อมูลสังเกต ตาม Kruskal- Wallis test อีกวิธีหนึ่ง คือ ใช้สูตร $H = \frac{12}{N(N+1)} \sum_{j=1}^k \frac{r_j^2}{n_j} - 3(N+1)$ ซึ่งศึกษาได้จาก nonparametric statistics for behavioral science (Siegel, 1956)

สุดท้ายนี้ขอสรุปว่า การใช้ nparstat ในการทดสอบสมมติฐานนี้ มีประโยชน์อย่างยิ่ง โดยเฉพาะสำหรับผู้ทำการวิจัยเพื่อเป็นวิทยานิพนธ์การทดลองทางการเกษตร ซึ่งไม่สามารถจัดหาหน่วยการทดลองได้มากเพียงพอในการใช้ parastat ในการทดสอบสมมติฐาน

เอกสารอ้างอิง

- ธีระดา ภิญโญ. (2551). สถิตินอนพาราเมตริก. คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏสวนสุนันทา: กรุงเทพฯ
- นิภา ศรีไพโรจน์. (2533). สถิตินอนพาราเมตริก. พิมพ์ครั้งที่ 2. โอเอสพริ้นติ้งเฮาส์: กรุงเทพฯ
- ล้วน สายยศ และอังคณา สายยศ. (2538). เทคนิคการวิจัยทางการศึกษา. พิมพ์ครั้งที่ 4. สุวีริยาสาส์น: กรุงเทพฯ
- Powers, Richard D. (1978). Nonparametric Statistics Handout. (Unpublished manuscript). Department of Agricultural Journalism. University of Wisconsin-Madison.
- Siegel, Sidney. (1956). Nonparametric Statistics for Behavioral Science. McGraw-Hill Book Company, INC. New York. Toronto London.