บทที่ 11
สถิติไม่มีพารามิเตอร์

เนื่องจากการวิเคราะห์ความแปรปรวน การทดสอบแบบเอฟ หรือแบบที จะต้องมีข้อสมมุติเกี่ยวกับประชากรว่าต้องเป็นประชากรแบบปกติ ต้องมีความแปรปรวนไม่ต่างกัน และการตั้งสมมติฐานจะเกี่ยวกับพารามิเตอร์ จึงเรียกว่าสถิติแบบมีพารามิเตอร์ แต่ในบางครั้งอาจมีปัญหาไม่ทราบลักษณะของประชากรว่ามีการแจกแจงแบบใด หรือทราบเพียงว่ามีการแจกแจงแบบต่อเนื่อง ดังนั้น วิธีการทางสถิติที่ไม่คำนึงถึงการแจกแจงของประชากรและพารามิเตอร์ จึงเรียกว่าสถิติไม่มีพารามิเตอร์ (Nonparametric or distribution-free) ซึ่งมีวิธีการเพียงนำการแจกแจงของประชากรมาเปรียบเทียบกันโดยไม่คำนึงถึงพารามิเตอร์ สำหรับวิธีการต่าง ๆ ที่จะกล่าวถึงในบทนี้จะกล่าวถึงเฉพาะเมื่อมีประชากร 2 ประชากรขึ้นไป

1. แบบทดสอบของครัสคัล–วอลลิส สำหรับการวิเคราะห์ความแปรปรวนแบบจำแนกทางเดียวโดยการจัดอันดับ (Kruskal-Wallis one-way analysis of variance by ranks)

ในปี 1952 ครัสคัลและวอลลิส ได้สร้างแบบทดสอบโดยใช้การจัดอันดับเพื่อใช้ทดสอบสมมติฐานว่างเปล่าว่า ตัวอย่างที่เป็นอิสระกัน k กลุ่มมาจากประชากร k ประชากรที่มีการแจกแจงแบบเดียวกัน โดยมีข้อจำกัดว่าข้อมูลที่เก็บมาอย่างน้อยที่สุดต้องเป็นแบบ ordinal และมีการแจกแจงแบบต่อเนื่อง การทดสอบแบบนี้ทำหน้าที่เหมือนตารางวิเคราะห์ความแปรปรวนของการวางแผนแบบสุ่มสมบูรณ์นั่นเอง

ตัวอย่าง สมมุติต้องการทดสอบว่าการเรียนรู้ของเด็กในชั้นเดียวกันแต่มาจากครอบครัวที่มีฐานะทางเศรษฐกิจและสังคมต่างกัน จะมีการเรียนรู้ต่างกันหรือไม่ จึงสุ่มตัวอย่างจากเด็กประถมหนึ่ง จากระดับเศรษฐกิจสังคมที่ต่าง ๆ กัน 4 ระดับ ๆ ละ 8 คน การให้คะแนนใช้ครู 3 คนต่างคนต่างให้คะแนนจากคะแนนเต็ม 12 คะแนน แล้วนำค่ามัธยฐานของครู 3 คน ดังข้อมูลในตารางข้างล่าง แล้วนำคะแนนของเด็กทั้ง 32 คน มาจัดอันดับ คะแนนต่ำสุดเป็นอันดับที่ 1 คะแนนสูงสุดเป็นอันดับที่ 32 ในกรณีที่มีคะแนนซ้ำกัน ให้ใช้ค่าเฉลี่ยจากผลรวมของอันดับ แล้วคำนวณค่าสถิติ H ในกรณีที่มีค่าซ้ำกันต้องใช้ค่าสถิติ H′ ตัวสถิติ H หรือ H′ จะมีการแจกแจงแบบไคสแควร์ ด้วย df. = k-1 จากตัวอย่างค่าสถิติที่คำนวณได้มากกว่าค่าเปิดตาราง จึงปฏิเสธ Ho และสรุปว่าตัวอย่าง 4 กลุ่มไม่ได้มาจากประชากรที่มีการแจกแจงแบบเดียวกัน

b_1	R_1	b_2	R_2	b_3	R_3	b_4	R_4
3	8.5	4	13.0	7	23.0	7	23.0
6	19.5	5	16.0	8	26.0	8	2b.0
3	X.5	4	13.0	7	23.0	9	28.5
3	8.5	3	8.5	6	19.5	8	2b.0
1	1.0	2	3.5	5	16.0	10	30.5
2	3.5	3	8.5	6	19.5	10	30.5
2	3.5	4	13.5	5	16.0	Y	2X.5
2	3.5	3	8.5	6	19.5	11	32.0
	Σ=56.5		Σ=84.0		Σ=162.5		Σ=225.0

$$\sum_j \sum_i r_{ij} = 56.5 + 84.0 + 162.5 + 225.0 = 528.0$$

check : $N(N+1)/2 = 32(32+1)/2 = 528.0$

$$H = \frac{12}{N(N+1)}\left[\sum_j \frac{(\sum_i R_{ij})^2}{n_j}\right] - 3(N+1)$$

$$= \frac{12}{32(32+1)}\left[\frac{(56.5)^2}{8} + \frac{(84.0)^2}{8} + \frac{(162.5)^2}{8} + \frac{(225.0)^2}{8}\right]$$

$$- 3(32+1) = 25.31$$

$$C = 1 - \left[\frac{t^3 - t)}{N^3 - N}\right]$$

$$= 1 - \left[\frac{2(2^3-2) + 4(3^3-3) + 2(4^3-4) + (6^3-6)}{32^3 - 32}\right] = .99$$

$$H' = H/C = 25.37/.99 = 25.63^*$$

$$\chi^2_{.05,3} = 7.82$$

หมายเหตุ t หมายถึงจำนวนค่าสังเกตที่ซ้ำค่ากันในกลุ่มตัวเลข นั่นคือ มีจำนวน 2 ชุด ที่ค่าสังเกตที่ซ้ำกันมี 2 จำนวน มี 4 ชุดที่ค่าสังเกตซ้ำกันมี 3 จำนวน เป็นต้น

การเปรียบเทียบระหว่างกลุ่ม

จะใช้วิธีการของ Mann-Whitney U test ซึ่งใช้ค่าสถิติ U ซึ่งมีการแจกแจงตามค่าในตาราง D.15 เมื่อ n_1, n_2 คือขนาดตัวอย่างของ 2 กลุ่มที่นำมาเปรียบเทียบกันไม่เกิน 8 จำนวน สมมุติฐานว่างเปล่าคือ " ตัวอย่างทั้งสองมาจากประชากรที่มีการแจกแจงแบบเดียวกัน" วิธีการคือต้องคำนวณค่าสถิติ U และระดับนัยสำคัญแบบ experimentalwise error rate (α) ค่าในตารางต้องคำนวณค่า α' โดยที่

$$\alpha' = \frac{2\alpha}{k(r-1)}$$

α คือ experimentalwise error rate

k คือจำนวนวิธีการ

r **คือจำนวนกรรมวิธีที่เรียงลำดับแล้วที่อยู่ในช่วงเปรียบเทียบ**

จากตัวอย่าง เมื่อเรียงลำดับของอันดับแล้ว b_1 มีผลรวมของอันดับต่ำสุด และ b_4 มีผลรวมของอันดับสูงที่สุด เมื่อต้องการทดสอบว่า b_1 และ b_4 มาจากประชากรแบบเดียวกันหรือไม่ มีวิธีการดังนี้

b_1	R_1	b_4	R_4
3	6.0	7	9.0
6	8.0	8	10.5
3	6.0	9	12.5
3	6.0	8	10.5
1	1.0	10	14.5
2	3.0	10	14.5
2	3.0	9	12.5
	$\Sigma = 36.0$		$\Sigma=100.0$

$$\text{Check : } \sum_{1}^{k}\sum_{1}^{n} R_{ij} = N(N+1)/2 = 136$$

$$U = n_1n_2 + \{n_1(n_1+1)\}/2 - \sum_{i}^{n} R_{i1}$$

$$= (8)(8) + 8(8+1)/2 - 36$$

$$U' = n_1n_2 - U = (8)(8) - 64 = 0$$

จะใช้ค่าสถิติ U หรือ U′ ตัวใดก็ตามที่มีค่าน้อยที่สุด

ถ้ากำหนดระดับนัยสำคัญ $\alpha = .05$

$$\alpha' = \frac{2(.05)}{4(3-1)} = .013$$

จากตาราง D.15 จะปฏิเสธ Ho ได้เมื่อความน่าจะเป็นจากตารางต้องน้อยกว่า $\alpha' = .013$ จากตาราง เมื่อ n = 8 และ U = 0 ได้ความน่าจะเป็น .000 ซึ่งน้อยกว่า .013 จึงปฏิเสธ Ho และสรุปว่า b_1 และ b_4 มาจากประชากรที่มีการแจกแจงต่างกัน

ต่อไป เมื่อจะทดสอบความแตกต่างของคู่อื่น ต้องคำนวณค่า α' และ U หรือ U' และเทียบกับค่าจากตารางเป็นคู่ๆไป เช่นระหว่าง b_1 กับ b_2

$$\alpha' = \frac{2(.05)}{4(2-1)} = .025$$

และเทียบกับค่าจากตาราง D.15 ค่าตารางมากกว่าจึงยังปฏิเสธ Ho ไม่ได้ ผลการทดสอบความแตกต่าง จากที่จับคู่ได้ 6 คู่ สรุปได้ในตาราง ดังนี้

comparison	r	α	U or U'	obtained probability	Decision
b_1 **versus** b_4	4	.009	**0**	.001	Reject Ho
b_1 **versus** b_3	3	.013	3.5	.001	Reject Ho
b_4 **versus** b_2	3	.013	0	.001	Reject Ho
b_1 **versus** b_2	2	.025	17.0	.065	**do** not Reject Ho
b_4 versus b_3	2	.025	3.0	.001	Reject Ho
b_2 versus b_3	2	.025	1.0	.001	Reject Ho

เมื่อขนาดตัวอย่างมากกว่า 8 ค่าตารางใช้ไม่ได้ ต้องประมาณด้วยการแจกแจงแบบปกติ โดยคำนวณค่าเฉลี่ย และความแปรปรวนของ U ดังนี้

$$\mu_U = \frac{n_1 n_2}{2} \quad ; \quad \sigma^2_U = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$$

$$\text{และ} \quad Z = \frac{U - \mu_U}{\sigma_U}$$

จะมีการแจกแจงแบบปกติมาตรฐาน (N (0,1)

ในกรณีที่มีค่าซ้ำ ความแปรปรวนของ U คือ

$$\sigma^2_U = \frac{n_1 n_2}{12} \left[n_1 + n_2 + 1 - \frac{\Sigma (t^3 - t)}{(n_1 + n_2)(n_1 + n_2 - 1)} \right]$$

ในเมื่อ t คือจำนวนค่าสังเกตที่ซ้ำกันในแต่ละเซ็ท

2. การวิเคราะห์ความแปรปรวนแบบจำแนก 2 ทางตามวิธีการของฟริดแมน

(Friedman Two-Way Analysis Of Variance By Ranks)[4]

ในปี 1937 ฟริดแมนได้สร้างแบบทดสอบแบบไม่มีพารามิเตอร์โดยใช้ผลรวมของอันดับ สำหรับข้อมูลที่เก็บจากตัวอย่างที่ไม่เป็นอิสระกัน ซึ่งใช้แทนการทดสอบของการวางแผน

แบบบล็อคสมบูรณ์ วิธีการ คือ ต้องจัดอันดับค่าสังเกตในแต่ละบล็อค ถ้ามีจำนวนซ้ำให้ใช้ค่าเฉลี่ยของอันดับ แล้วจึงคำนวณค่าสถิติซึ่งจะมีการแจกแจงแบบไคสแควร์ด้วย df = (k − 1)

ตัวอย่าง

	b_1	R_1	b_2	R_2	b_3	R_3	b_4	R_4
S_1	3	1	4	2	7	3.5	7	3.5
S_2	6	2	5	1	8	3.5	8	3.5
S_3	3	1	4	2	7	3	9	4
S_4	3	1.5	3	1.5	6	3	8	4
S_5	1	1	2	2	5	3	10	4
S_6	2	1	3	2	6	3	10	4
S_7	2	1	4	2	5	3	9	4
S_8	2	1	3	2	6	3	11	4
		9.5		14.5		25.0		31.0

$$\chi^2 = \frac{12}{nk(k+1)} [\Sigma(\Sigma R_{ij})^2] - 3n(k+1)$$

$$= \frac{12}{8(4)(4+1)} [(9.5^2 + 14.5^2 + 25.0^2 + 31.0^2)] - 3(8)(4+1)$$

$$= .075(1.886.50) - 120 = 21.49^*$$

$$\chi^2_{.05,3df} = 7.82$$

การวัดความเกี่ยวพันกันของตัวอย่างที่ไม่เป็นอิสระกัน k กลุ่ม

เมื่อต้องการทราบว่าการจัดอันดับให้กับ k รายการ จำนวน n ครั้งนั้นมีความสอดคล้องกันเพียงไร ต้องใช้ตัวสถิติ W ซึ่งเคนดอลล์ได้สร้างขึ้นในปี 1948 เรียกว่า "Kendall's coefficient of concordance" W คืออัตราส่วนของความแปรปรวนของผลรวมของอันดับ k รายการเทียบกับความแปรปรวนที่สูงสุดที่จะพึงมีของผลรวมของอันดับ k รายการนี้ ถ้าการจัดอันดับสอดคล้องกันทั้ง n ครั้ง W จะมีค่าเป็น 1 แต่ถ้ามีความขัดแย้งกันมาก (lack of concordance) ค่าของ W จะเท่ากับ 0

โดยที่

$$W = \frac{12[\Sigma(\Sigma R_{ij})^2]}{n^2k(k^2 - 1)} \quad \frac{3(k+1)}{k-1}$$

จากตัวอย่าง

$$W = \frac{12(9.5^2 + 14.5^2 + \ldots + 31.0^2)}{8^2.4(4^2 - 1)} - \frac{3(4+1)}{4-1} = .90$$

หมายความว่าความแปรปรวนของ k ผลรวมของอันดับเป็น 90% ของค่าสูงสุดที่เป็นไปได้ อนึ่งมีวิธีอธิบายอีกวิธีหนึ่งโดยการคำนวณค่าเฉลี่ยของสหสัมพันธ์อันดับ $\bar{r}_s$ ซึ่งคือค่าเฉลี่ยสหสัมพันธ์แบบจัดอันดับของการจับคู่ทั้งหมด

$$\bar{r}_s = \frac{nW - 1}{n - 1} \qquad \frac{8(.90) - 1}{8 - 1} = .89$$

คือค่าเฉลี่ยของสหสัมพันธ์แบบจัดอันดับของ $n(n-1)/2 = 28$ คู่ ของการจัดอันดับของ 8 เซ็ท ขีดจำกัดของ r_s คือ $-1/(n-1)$ และ 1 จากตัวอย่างสรุปได้ว่าการจัดอันดับสอดคล้องกัน มีวิธีสังเกตอีกวิธีคือ ถ้าค่าสถิติ X^2 ที่คำนวณได้มีนัยสำคัญ ค่าของ $W \neq 0$

การเปรียบเทียบระหว่างกลุ่มของข้อมูลแบบจับคู่

จะใช้วิธีจัดอันดับของการจับคู่และทดสอบด้วยวิธีการของ วิลคอกซ์สัน มีวิธีการคือ ให้จับคู่และหาผลต่างโดยเก็บเครื่องหมายไว้ แล้วนำผลต่างมาจัดอันดับ แล้วรวมผลรวมของอันดับโดยแยกเครื่องหมาย (แต่เวลาจัดอันดับไม่คำนึงถึงเครื่องหมาย) ผลรวมอันดับของเครื่องหมายที่มีค่าน้อยที่สุดคือค่าสถิติที่นำไปใช้คำนวณในขั้นต่อไป

(1)	(2)	(3)	(4)	(5)	(6)
			rank of		
b_1	b_2	d	$\lvert d \rvert$	−	+
3	4	−1	3.5	−3.5	−
6	5	1	3.5	−	3.5
3	4	−1	3.5	−3.5	
3	3	0	−	−	−
1	2	−1	3.5	−3.5	−
2	3	−1	3.5	−3.5	−
2	4	−2	7.0	−7.0	−
2	3	−1	3.5	−3.5	−
					T = 3.5

หมายเหตุ ค่าสถิติ T คือผลรวมอันดับของเครื่องหมายที่ให้ผลรวมน้อยที่สุด

$$Z = \frac{T - \mu_T}{\sigma_T} = \frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

$$= \frac{3.5 - \frac{7(7+1)}{4}}{\sqrt{\frac{7(7+1)(2)(7)+1)}{24}}}$$

$$= \frac{3.5 - 14.0}{\sqrt{35.0}} = \frac{-10.5}{5.9} = -1.78$$

$z_{.025} = 1.96$

จึงสรุปว่าความแตกต่างไม่มีนัยสำคัญ

เปรียบเทียบข้อดี-เสียของการใช้สถิติไม่มีพารามิเตอร์

ข้อดี

- มีข้อสมมุติเกี่ยวกับประชากรน้อยกว่าแบบมีพารามิเตอร์
- การคำนวณไม่ยุ่งยากมาก
- สามารถใช้กับข้อมูลบางชุดซึ่งไม่เหมาะกับการทดสอบแบบมีพารามิเตอร์

ข้อเสีย

- มีพลังการทดสอบน้อยกว่าแบบมีพารามิเตอร์สำหรับกรณีที่มีข้อสมมุติที่สมบูรณ์ของแบบมีพารามิเตอร์
- ให้ข่าวสารจากข้อมูลที่เก็บมาน้อยกว่าแบบมีพารามิเตอร์ที่มีข้อสมมุติสมบูรณ์