บทที่ 13
การวิเคราะห์ในสถิติไม่อิงพารามิเตอร์ : การทดสอบเกี่ยวกับค่ากลาง

1. แนวคิดเกี่ยวกับสถิติไม่อิงพารามิเตอร์

ในสถิติเชิงอนุมาน (inferential statistics) เราจำเป็นต้องทราบเกี่ยวกับการแจกแจงความน่าจะเป็นของข้อมูล การแจกแจงความน่าจะเป็นที่รู้จักกันมากที่สุดคือ การแจกแจงแบบปกติซึ่งใช้กับข้อมูลแบบต่อเนื่อง การแจกแจงทวินามใช้กับข้อมูลที่เป็นจำนวนนับ และการแจกแจงอื่น ๆ ที่รู้จักกัน ได้แก่ ยูนิฟอร์ม ปัวซอง เอ็กซโพเนนเชียล เป็นต้น จากความรู้ทางทฤษฎีและประสบการณ์ที่ผ่านมาเราอาจมีเหตุผลพอที่จะบอกได้ว่าข้อมูลมีการแจกแจงแบบใดแบบหนึ่งได้ และจากทฤษฎีที่เรียกว่า ทฤษฎีขีดจำกัดกลาง (Central limit theorem) ก็สนับสนุนเกี่ยวกับการสรุปอ้างอิงสำหรับการวัดส่วนใหญ่ที่มีข้อตกลงเบื้องต้นว่าข้อมูลได้มาจากตัวอย่างสุ่มที่มีการแจกแจงแบบปกติ แม้พบว่าการตรวจสอบข้อตกลงเบื้องต้นเกี่ยวกับความเป็นปกติจะไม่เป็นจริงก็ตาม การสรุปอ้างอิงเกี่ยวกับพารามิเตอร์บางครั้งอาจไม่เหมาะสมหรือทำไม่ได้เนื่องจากการที่จะบอกว่ากลุ่มตัวอย่างมาจากการแจกแจงแบบใดแบบหนึ่งอาจไม่มีเหตุผลเพียงพอ ในทางปฏิบัติเราสามารถบอกได้ยากมากว่ากลุ่มตัวอย่างของเรามาจากการแจกแจงแบบใด ในกรณีเช่นนี้เหมาะสมที่จะใช้การสรุปอ้างอิงแบบไม่อิงพารามิเตอร์ แต่บางท่านชอบใช้คำว่า distribution – free มากกว่าคำว่าไม่อิงพารามิเตอร์

ในทางปฏิบัติเราใช้วิธีการง่าย ๆ ในการสรุปอ้างอิง ส่วนใหญ่เราจะสมมติว่ากลุ่มตัวอย่างของเรามาจากการแจกแจงที่สมมาตร (symmetric) นี่ไม่ได้จำกัดว่าต้องเป็นการแจกแจงแบบปกติแต่หมายความว่าเราไม่นับรวมการแจกแจงทั้งหมดที่เป็นแบบไม่สมมาตร (asymmetric) และสำหรับวิธีที่ไม่สามารถบอกได้ชัดลงไปว่าการสรุปอ้างอิงเป็นแบบอิงพารามิเตอร์หรือแบบไม่อิงพารามิเตอร์ก็มีคำกลาง ๆ ที่ใช้กันคือ semiparametic และ asymptotically distribution – free

ในกรณีที่เราไม่ทราบการแจกแจงความน่าจะเป็นของข้อมูล หรือจำนวนตัวอย่างน้อย หรือตัวอย่างสุ่มมาจากประชากรที่มีการแจกแจงไม่ใช่แบบปกติ ลักษณะข้อมูลเป็นข้อมูลเชิงคุณภาพที่มีสเกลการวัดแบบแบ่งประเภทหรือแบบอันดับ ซึ่งไม่เป็นไปตามข้อตกลงเบื้องต้นของสถิติสรุปอ้างอิงเกี่ยวกับพารามิเตอร์ ดังนั้นจึงอาจจำเป็นต้องใช้วิธีการของสถิติที่ไม่อิงพารามิเตอร์ ซึ่งได้แก่ Chi – Square test , Binomial test, Runs test , Kolmogorov – Smirnov test , Wilcoxon test , Sign test เป็นต้น สถิติทดสอบบางตัวที่ใช้สำหรับการทดสอบแบบไม่อิงพารามิเตอร์อย่างน้อยที่สุดก็มีการแจกแจงใกล้เคียงกับการแจกแจงแบบต่อเนื่องที่คุ้นเคย เช่น การแจกแจงปกติ การแจกแจง t การแจกแจง F หรือการแจกแจง χ^2

2. การทดสอบเกี่ยวกับค่ากลางของประชากร 1 กลุ่ม

2.1 การทดสอบเครื่องหมาย (Sign test)

เมื่อใช้สถิติทดสอบ t หรือ t-test ทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากรเท่ากับค่าที่กำหนดให้บางค่า ซึ่งมีข้อตกลงเบื้องต้นว่าประชากรต้องมีการแจกแจงแบบปกติ ถ้าประชากรมีการแจกแจงไม่เป็นแบบปกติ เราอาจใช้การทดสอบไม่อิงพารามิเตอร์ซึ่งมีข้อตกลงเบื้องต้นน้อยกว่า คือไม่เข้มงวดเกี่ยวกับกลุ่มตัวอย่างว่ามาจากการแจกแจงแบบหนึ่ง และวิธีการไม่อิงพารามิเตอร์ใช้ได้กับข้อมูลที่มีสเกลการวัดแบบอันดับหรือจำนวนนับของเหตุการณ์ที่สนใจศึกษาซึ่งเป็นข้อมูลแบบแบ่งประเภท

ในการศึกษาทางการแพทย์ มักดูอาการดีขึ้นของคนไข้ภายในระยะเวลาที่กำหนดภายหลังการให้ทรีทเมนต์ อาจเป็นระยะเวลาไม่กี่เดือนถึง 5 หรือ 6 ปี จากการศึกษาของ the Eastern Co-operative Oncalogy Group ในอเมริกา Dinse (1982) อ้างถึงใน P.Sprent and N.C. Smeeton (2001) ทำการศึกษาคนไข้ที่เป็นผู้ชาย 10 คน ซึ่งเป็นโรค symptomatic lympho non – Hodgkin's lymphoma ที่เข้าร่วมในการศึกษานี้ ได้ให้ข้อมูลเวลารอดคิดเป็นสัปดาห์ ทำการศึกษาเป็นเวลา 362 สัปดาห์ เราไม่รู้เวลารอดจริง ๆ ของคนไข้คนหนึ่งซึ่งยังมีชีวิตภายหลัง 362 สัปดาห์ ข้อมูลสำหรับคนไข้คนนี้เรียกว่า censored ได้ข้อมูลเวลารอดของคนไข้ทั้ง 10 คน เป็นสัปดาห์คือ 49, 58, 75, 110, 112, 132, 151, 276, 281, 362* เราใช้สัญลักษณ์ * แทนข้อมูลตัวที่เป็น censored

อยากทราบว่ามีเหตุผลสนับสนุนพอไหมที่จะบอกว่าข้อมูลเหล่านี้มีมัธยฐานของเวลารอดเท่ากับ 200 สัปดาห์ ข้อมูลตัวที่เป็น censored ทำให้เป็นปัญหาในการทดสอบแบบอิงพารามิเตอร์ แต่สามารถใช้การทดสอบแบบไม่อิงพารามิเตอร์ได้อย่างง่าย ๆ สมมติฐานที่ต้องการทดสอบคือ ค่ามัธยฐานของเวลารอดเท่ากับ 200 หรือไม่ โดยมีข้อตกลงเบื้องต้นว่าข้อมูลได้มาจากตัวอย่างซึ่งสุ่มมาจากประชากรคนไข้ที่มีเชื้อโรคนี้

ในการสรุปอ้างอิงเกี่ยวกับค่ามัธยฐานหรือการแจกแจงของเวลารอดนี้ใช้ได้กับคนไข้ทั้งหมดที่ได้รับทรีทเมนต์แบบเดียวกันนี้และที่มีคุณลักษณะอื่น ๆ ที่เหมือนกัน เช่น อายุเดียวกัน เพศเดียวกัน คนไข้ในการศึกษานี้เป็นผู้ชายทั้งหมดจึงไม่ควรอ้างอิงไปสู่เวลารอดของผู้หญิงว่ามีการแจกแจงของเวลารอดแบบเดียวกัน

ในตัวอย่างนี้เราต้องการทดสอบสมมติฐานเกี่ยวกับค่ามัธยฐานของเวลารอดของประชากรเท่ากับ 200 คู่กับสมมติฐานแย้งว่าเท่ากับค่าอื่น ๆ เขียนเป็นสัญลักษณ์ได้คือ H_0 : $\theta = 200$ คู่กับ $H_1 : \theta \neq 200$ หรือ อาจเขียนเป็นความน่าจะเป็นของจำนวนข้อมูลที่มีค่ามากกว่าหรือน้อยกว่าค่ามัธยฐานที่กำหนด โดยกำหนดให้ข้อมูลที่มีค่ามากกว่า 200 มีเครื่องหมายบวก และข้อมูลที่มีค่าน้อยกว่า 200 มีเครื่องหมายลบ จึงเขียนเป็นสมมติฐานข้างต้นใหม่ในเทอมของความน่าจะเป็นได้ว่า $H_0 : P(+) = P(-) = .5$ เพื่อทดสอบสมมติฐานที่ว่ามีจำนวนเครื่องหมายบวกและลบเท่ากัน วิธีการทดสอบต้องการเพียงจำนวนนับของข้อมูลที่มีค่ามากกว่า 200 บันทึกข้อมูลเป็น ‘ + ‘ ดังนั้นจึงเขียนสมมติฐานข้างต้นใหม่ได้คือ $H_0 : P(+) = .5$

ถ้าเรามีตัวอย่างสุ่มที่ได้จากการ แจกแจงใด ๆ ซึ่งมีมัธยฐานเท่ากับ 200 จะได้ว่าค่าของตัวอย่างแต่ละตัวจะมีค่ามากหรือน้อยกว่า 200 จำนวนพอ ๆ กัน หมายความว่า ภายใต้ H_0 จำนวนของเครื่องหมายบวกมีการแจกแจงแบบทวินาม $B(10 , .5)$ จากการสังเกต 10 ครั้ง เมื่อความน่าจะเป็นที่จะเกิดเครื่องหมายบวกคือ $p = .5$ เราสามารถคำนวณหาความน่าจะเป็นของจำนวนเครื่องหมายบวกเท่ากับ r ได้จากสูตรคือ

$$P(X = r) = \binom{10}{r} p^r (1 - p)^{10-r}$$

หรือเปิดค่าจากตารางการแจกแจงทวินามที่ $n = 10$, $p = .5$ ในตำราสถิติ (ประสิทธิ์ พยัคฆพงษ์, 2545) จะได้ค่าของความน่าจะเป็นสำหรับแต่ละค่าของ r ระหว่าง 0 ถึง 10 ดัง ตารางที่ 13.1

ตารางที่ 13.1 ค่าความน่าจะเป็นจากตารางการแจกแจงทวินามที่ $n = 10$, $p = .5$ เมื่อ $r = 0$ ถึง 10

r	0	1	2	3	4	5	6	7	8	9	10
p	.0010	.0098	.0439	.1172	.2051	.2461	.2051	.1172	.0439	.0098	.0010

กฎการตัดสินใจสำหรับสมมติฐานแย้งที่แตกต่างกัน 3 แบบคือ

1) $H_1 : P(+) > .5$ จะตัดสินใจปฏิเสธ $H_0 : P(+) = .5$ ถ้าความน่าจะเป็นที่จะเกิดเครื่องหมายบวกมากกว่าหรือเท่ากับ ระดับนัยสำคัญที่กำหนด α

2) $H_1 : P(+) < .5$ จะตัดสินใจปฏิเสธ $H_0 : P(+) = .5$ ถ้าความน่าจะเป็นที่จะเกิดเครื่องหมายบวกน้อยกว่าหรือเท่ากับระดับนัยสำคัญที่กำหนด α

3) $H_1 : P(+) \neq .5$ จะตัดสินใจปฏิเสธ $H_0 : P(+) = .5$ ถ้าความน่าจะเป็นที่จะเกิดเครื่องหมายบวกน้อยกว่าหรือเท่ากับระดับนัยสำคัญที่กำหนด α

จากข้อมูลพบว่ามีค่าสังเกต 3 ตัวที่มีค่ามากกว่าค่ามัธยฐานของเวลารอดของประชากร ดังนั้นจึงมีข้อมูลที่มีเครื่องหมายบวกจำนวนเท่ากับ 3 และจากตารางที่ 5.1 จะเห็นว่าความน่าจะเป็นที่ r น้อยกว่า หรือเท่ากับ 3 ในตัวอย่างขนาด $n = 10$ คือ $.1172 + .0439 + .0098 + .0010 = .1719$ เนื่องจากความน่าจะเป็นของจำนวนเครื่องหมายบวก ที่น้อยกว่าหรือเท่ากับ 3 และความน่าจะเป็นของจำนวนเครื่องหมายบวกที่มากกว่าหรือเท่ากับ 7 เป็นการทดสอบแบบสองทาง จะได้ค่า P เท่ากับ $2 \times .1719 = .3438$ แสดงว่าไม่มีหลักฐานเพียงพอที่จะปฏิเสธ H_0 นั่นคือมีจำนวนเครื่องหมายบวก

เท่ากับจำนวนเครื่องหมายลบ การทดสอบอย่างง่าย ๆ นี้เรียกว่าการทดสอบเครื่องหมาย (Sign test)

2.2 ตัวอย่างการทดสอบเครื่องหมายและขนาดตัวอย่าง

สำหรับหัวข้อนี้จะแสดงตัวอย่างการทดสอบเครื่องหมายของกลุ่มตัวอย่างกลุ่ม 1 ที่มีขนาดตัวอย่าง $n_1 = 10$ และตัวอย่างการทดสอบเครื่องหมายของกลุ่มตัวอย่างกลุ่ม 2 ที่มีขนาดตัวอย่าง $n_2 = 20$ เพื่อแสดงให้เห็นเกี่ยวกับอิทธิพลของขนาดตัวอย่างในการทดสอบเครื่องหมายของประชากร 1 กลุ่ม โดยอาศัยข้อมูลจากการศึกษาเรื่องเปอร์เซ็นต์น้ำในพื้นที่เกษตรกรรมของพื้นที่หลายแห่ง (จาก P. Sprent and N.C. Smeeton ; 2001) สมมติฐานที่ต้องการทดสอบคือ ค่ามัธยฐานของเปอร์เซ็นต์น้ำในพื้นที่เกษตรกรรมทั้งหลายเท่ากับ 9 หรือไม่ เราใช้ข้อมูลบางตัวจากข้อมูลชุดใหญ่ แล้วแบ่งออกเป็น 2 กลุ่ม คือ กลุ่ม 1 มีขนาดตัวอย่างเท่ากับ 10 และกลุ่ม 2 มีขนาดตัวอย่างเท่ากับ 20 เรียงลำดับค่าของข้อมูลจากน้อยไปมากของตัวอย่างทั้งสองกลุ่มได้ดัง ตารางที่ 13.2

ตารางที่ 13.2 ข้อมูลเปอร์เซ็นต์น้ำในพื้นที่เกษตรกรรมของพื้นที่หลายแห่ง

ตัวอย่างกลุ่ม 1 (n = 10)	5.5	6.0	6.5	7.6	7.6	7.7	8.0	8.2	9.1	15.1
ตัวอย่างกลุ่ม 2	5.6	6.1	6.3	6.3	6.5	6.6	7.0	7.5	7.9	8.0
(n = 20)	8.0	8.1	8.1	8.2	8.4	8.5	8.7	9.4	14.3	26.0

ในการทดสอบเครื่องหมาย ถ้ามัธยฐานของประชากรคือ 9 แล้วค่าสังเกตตัวใด ๆ ในตัวอย่างสุ่มมีค่ามากกว่า 9 เราจะให้เครื่องหมายบวกกับค่าที่มากกว่า 9

ต้องการทดสอบสมมติฐานเกี่ยวกับมัธยฐาน θ สมมติฐานที่ต้องการทดสอบคือ $H_0 : \theta = 9$ คู่กับ $H_1 : \theta \neq 9$ โดยใช้การทดสอบเครื่องหมาย หรือ $H_0 : P(+) = .5$ คู่กับ $H_1 : P(+) \neq .5$ และถ้า H_0 เป็นจริง จำนวนเครื่องหมายบวกของ

ตัวอย่างกลุ่ม 1 จะมีการแจกแจงทวินาม ซึ่งมี $p = .5$, $n = 10$ เขียนได้เป็น B(10 ,.5) และตัวอย่างกลุ่ม 2 จะมีการแจกแจงทวินาม B(20 ,.5)

ในตัวอย่างกลุ่ม 1 มีข้อมูล 2 ตัว คือ 9.1 และ 15.1 ที่มีค่าเกิน 9 ทำให้มีเครื่องหมายบวกจำนวนเท่ากับ 2 แล้วคำนวณหาความน่าจะเป็นของ B(10 , .5) หรือหาค่าได้จากตารางความน่าจะเป็นของการแจกแจงทวินาม B(n, p) ซึ่งคำนวณความน่าจะเป็นของจำนวนข้อมูลที่มีเครื่องหมายบวกเท่ากับ r จากสูตร

$$P(X = r) = \binom{n}{r} p^r (1 - p)^{n-r}$$

จะได้ความน่าจะเป็นของจำนวนเครื่องหมายบวกเท่ากับ 0, 1, และ 2 คือ .0010, .0098, และ .0439 ตามลำดับ หรือความน่าจะเป็นของจำนวนเครื่องหมายบวก เท่ากับ 8, 9, และ 10 คือ .0439, .0098, และ .0010 ตามลำดับ ดังนั้นในการทดสอบสมมติฐานแบบสองทางสามารถคำนวณหาความน่าจะเป็นของจำนวนเครื่องหมายบวกน้อยกว่าหรือเท่ากับ 2 และจำนวนเครื่องหมายบวกมากกว่าหรือเท่ากับ 8 ได้ความน่าจะเป็นคือ

$$P = (.0010 + .0098 + .0439) + (.0439 + .0098 + .0010)$$

$$= .109$$

ในตัวอย่างกลุ่ม 2 ที่มีการแจกแจงทวินาม B(20 , .5) สามารถคำนวณหาความน่าจะเป็นของจำนวนเครื่องหมายบวกเท่ากับ r หรือเปิดค่าจากตารางการแจกแจงทวินามที่ $n = 20$, $p = .5$ จะได้ค่าความน่าจะเป็นสำหรับแต่ละค่าของ r ระหว่าง 0 ถึง 20 ดังตารางที่ 13.3

ตารางที่ 13.3 ค่าความน่าจะเป็นจากตารางการแจกแจงทวินามที่ $n = 20$, $p = .5$

r	p	r	p	r	p	r	p
0	.0000	5	.0148	10	.1762	15	.0148
1	.0000	6	.0370	11	.1602	16	.0046
2	.0002	7	.0739	12	.1201	17	.0011

3	.0011	8	.1201	13	.0739	18	.0002
4	.0046	9	.1602	14	.0370	19	.0000
						20	.0000

จากข้อมูลในตัวอย่างกลุ่ม 2 พบว่ามีข้อมูล 3 ตัว คือ 9.4 , 14.3 , และ 26.0 ที่มีค่ามากกว่าค่ามัธยฐานของประชากรคือ 9 ดังนั้นจึงมีเครื่องหมายบวกจำนวนเท่ากับ 3 ในการทดสอบสมมติฐานแบบสองทางสามารถคำนวณหาความน่าจะเป็นของเครื่องหมายบวกจำนวนน้อยกว่า หรือเท่ากับ 3 และความน่าจะเป็นของเครื่องหมายบวกจำนวนมากกว่าหรือเท่ากับ 17 ได้ความน่าจะเป็นคือ

$$P = (.0000 + .0000 + .0002 + .0011) + (.0011 + .0002 + .0000 + .0000)$$
$$= .0026$$

สรุปผลสำหรับตัวอย่างกลุ่ม 1 ขนาด $n = 10$ ที่มีความน่าจะเป็นเท่ากับ .109 จึงไม่มีหลักฐานพอที่จะปฏิเสธ H_0 สำหรับตัวอย่างกลุ่ม 2 ขนาด $n = 20$ ที่มีความน่าจะเป็นเท่ากับ .0026 จึงมีหลักฐานเพียงพอที่จะปฏิเสธ H_0 ผลที่ได้นี้ยืนยันว่าการเพิ่มขนาดตัวอย่างโดยทั่วไปจะช่วยเพิ่มอำนาจของการทดสอบสมมติฐาน เราสามารถใช้โปรแกรม SPSS ช่วยในการคำนวณได้ ซึ่งคำสั่งการทดสอบเครื่องหมายของประชากร 1 กลุ่ม อยู่ในคำสั่ง 2 Related Samples ใน Nonparametric test

2.3 การใช้คำสั่ง 2 Related Samples ... สำหรับการทดสอบเครื่องหมาย

จากการศึกษาเรื่องเปอร์เซ็นต์น้ำในพื้นที่เกษตรกรรมของตัวอย่างกลุ่ม 2 ข้อมูลอยู่ในตารางที่ 13.2 ที่มีขนาดตัวอย่างเท่ากับ 20 กำหนดให้ตัวแปร median คือมัธยฐานของประชากรที่กำหนดมีค่าเท่ากับ 9 และตัวแปร group2 คือข้อมูลเปอร์เซ็นต์น้ำในพื้นที่เกษตรกรรมของตัวอย่างกลุ่ม 2 ซึ่งข้อมูลมีสเกลการวัดแบบช่วง (interval scale) สมมติฐานที่ต้องการทดสอบคือ H_0 : เปอร์เซ็นต์น้ำในพื้นที่เกษตรกรรมมีมัธยฐานเท่ากับ 9 คู่กับ H_1 : เปอร์เซ็นต์น้ำในพื้นที่เกษตรกรรมมีมัธยฐานไม่เท่ากับ 9 ทดสอบสมมติฐานโดยใช้การทดสอบเครื่องหมาย บันทึกข้อมูลในแฟ้มข้อมูล Sign1.sav มีรูปแบบข้อมูลในแฟ้มดังตารางที่ 13.4

ตารางที่ 13.4 รูปแบบการบันทึกข้อมูลเปอร์เซ็นต์น้ำในพื้นที่เกษตรกรรมของตัวอย่างกลุ่ม 2

median	group2	median	group2
9	5.6	9	8.0
9	6.1	9	8.1
9	6.3	9	8.1
9	6.3	9	8.2
9	6.5	9	8.4
9	6.6	9	8.5
9	7.0	9	8.7
9	7.5	9	9.4
9	7.9	9	14.3
9	8.0	9	26.0

ที่ใช้คำสั่ง 2 Related Samples เนื่องจากการที่ข้อมูลชุดหนึ่งถูกแบ่งออกเป็น 2 ส่วน ด้วยค่ามัธยฐาน จึงดูเสมือนกับมีตัวอย่าง 2 กลุ่ม ที่เกี่ยวข้องกัน ซึ่งใช้เป็นแนวคิดในการคำนวณของโปรแกรมเท่านั้น

จากข้อมูลในแฟ้มข้อมูล siqn1.sav มีข้อมูลที่มีเครื่องหมายบวกคือ มีค่ามากกว่า 9 จำนวนเท่ากับ 3 และมีข้อมูลที่มีเครื่องหมายลบคือ มีค่าน้อยกว่า 9 จำนวนเท่ากับ 17 การใช้โปรแกรม SPSS ช่วยในการทดสอบสมมติฐาน H_0 : เปอร์เซ็นต์น้ำในพื้นที่เกษตรกรรมมีมัธยฐานเท่ากับ 9 โดยใช้การทดสอบ Sign test มีขั้นตอน ดังนี้

1. ไปที่เมนูบาร์ คลิกที่ Analyze , Nonparametric Tests, 2 Related Samples... จะได้หน้าต่าง Two-Related-Samples Tests
2. ในหน้าต่าง Two-Related-Samples Tests

ในช่องซ้ายมือ คลิกที่ตัวแปร median และ group2 ให้ย้ายไปอยู่ในช่อง Test Pair(s) List:

ในกรอบ Test Type

เลือก ☐ Sign เพื่อทำการทดสอบ Sign test

แล้วคลิกที่ปุ่ม OK จะได้ผลลัพธ์ดังภาพที่ 13.1

NPar Tests
Sign Test

Frequencies

		N
group2 - median	Negative Differences [a]	17
	Positive Differences [b]	3
	Ties[c]	0
	Total	20

a. group2 < median

b. group2 > median

c. group2 = median

Test Statistics[b]

	group2 - median
Exact Sig. (2-tailed)	.003[a]

a. Binomial distribution used.

b. Sign Test

ภาพที่ 13.1

จากภาพผลลัพธ์ในตาราง Frequencies แสดงจำนวนข้อมูลที่มีเครื่องหมายลบและบวกในการทดสอบ Sign test คือมีเครื่องหมายลบจำนวนเท่ากับ 17 และเครื่องหมายบวกจำนวนเท่ากับ 3 รวมทั้งหมดจำนวนเท่ากับ 20

ผลการทดสอบ Sign test อยู่ในตาราง Test Statistics ได้ค่า Exact Sig. (2-tailed) เท่ากับ .003 ซึ่งมีค่าน้อยกว่าระดับนัยสำคัญที่กำหนด ($\alpha = .05$) ดังนั้นจึงปฏิเสธ H_0 นั่นคือเปอร์เซ็นต์น้ำในพื้นที่เกษตรกรรมมีมัธยฐานไม่เท่ากับ 9

2.4 การทดสอบวิลคอกซัน (Wilcoxon signed – rank test)

ทฤษฎีของการทดสอบวิลคอกซัน สมมติให้ประชากรมีการแจกแจงแบบสมมาตร ข้อตกลงเบื้องต้นเกี่ยวกับความสมมาตรนี้แสดงนัยว่าผลบวกของอันดับที่เป็นบวกและผลบวกของอันดับที่เป็นลบควรจะเกือบเท่ากัน เราสร้างข้อมูลเป็นอันดับหรือเครื่องหมายบวกหรือลบที่เทียบ

กับค่าเฉลี่ยหรือค่ามัธยฐานที่กำหนดในสมมติฐาน เมื่อ S คือ ผลบวกของอันดับที่มีเครื่องหมายบวกหรือลบ ให้ตัวที่น้อยกว่าอีกตัวหนึ่งเป็นสถิติทดสอบ

ถ้าข้อมูลมีค่าเท่ากันหลายตัว โดยไม่จำเป็นต้องมีเครื่องหมายเหมือนกัน เราจะแทนอันดับของข้อมูลเหล่านี้ด้วยค่าเฉลี่ยของอันดับของทุกจำนวนที่เท่ากัน เช่น ถ้ามีค่าสังเกต 7 ตัว คือ 1, 1, 5, 5, 8, 8, 8 และเราต้องการใช้ Wilcoxon sign – rank test ทดสอบสมมติฐานเกี่ยวกับมัธยฐานคือ $H_0 : \theta = 3$ เราจะลบข้อมูลทุกตัวด้วย 3 ได้ผลคือ -2, -2, 2, 2, 5, 5, 5 ข้อมูลที่เล็กที่สุดคือ 2 มีอยู่ 4 ตัว ดังนั้นค่าเฉลี่ยของอันดับของข้อมูล 4 ตัวนี้ คือ $(1 + 2 + 3 + 4)/4 = 2.5$ และข้อมูลอีก 3 ตัวเป็น 5 ดังนั้นค่าเฉลี่ยของอันดับของข้อมูล 3 ตัวนี้คือ $(5 + 6 + 7)/3 = 6$ สรุปได้ข้อมูลที่แปลงเป็นอันดับได้คือ -2.5, -2.5, 2.5, 2.5, 6, 6, 6 สถิติทดสอบที่เหมาะสมคือผลบวกของอันดับที่เป็นลบคือ $S_- = 5$ ซึ่งมีค่าน้อยกว่า $S_+ = 23$

ตัวอย่างการศึกษาจำนวนคำในประโยคแรกของย่อหน้าต่าง ๆ ที่สุ่มมาทั้งหมด 12 ย่อหน้า [จาก Fisher (1948) อ้างถึงใน P. Sprent and N.C. Smeeton (2001)] ได้ข้อมูลคือ 12, 18, 24, 26, 37, 40, 42, 47, 49, 49, 78, 108 ซึ่งเรียงจากน้อยไปมาก สมมติฐานที่ต้องการทดสอบเกี่ยวกับมัธยฐานคือ $H_0 : \theta \le 30$ คู่กับ $H_1 : \theta > 30$ โดยมีข้อตกลงเบื้องต้นคือ ความยาวของประโยคมีการแจกแจงแบบสมมาตร

วิธีการแปลงข้อมูลโดยลบข้อมูลทุกตัวด้วย 30 ได้ข้อมูลที่แปลงแล้วคือ

$$-18, -12, -6, -4, 7, 10, 12, 17, 19, 19, 48, 78$$

นำตัวเลขเหล่านี้มาเรียงใหม่ตามปริมาณโดยไม่ดูเครื่องหมาย แล้วคิดอันดับของข้อมูลใหม่ทุกตัวโดยไม่คำนึงถึงเครื่องหมาย ข้อมูลตัวที่เท่ากันทุกตัวให้มีอันดับเดียวกันเท่ากับค่าเฉลี่ยของอันดับ ได้ผลดังตารางที่ 13.5

ตารางที่ 13.5 การคำนวณค่าสถิติทดสอบของตัวอย่างการศึกษาจำนวนคำในประโยคแรกของย่อหน้า

จำนวนคำในประโยค	$d_i = x_i - 30$	อันดับที่ของ d_i	เครื่องหมายของอันดับที่ของ d_i
12	-18	8	-8
18	-12	5.5	-5.5
24	-6	2	-2
26	-4	1	-1
37	7	3	3
40	10	4	4
42	12	5.5	5.5
47	17	7	7
49	19	9.5	9.5
49	19	9.5	9.5
78	48	11	11
108	78	12	12

คำนวณค่าสถิติทดสอบได้ค่า $S_- = 16.5$ และ $S_+ = 61.5$ เห็นได้ชัดว่าผลบวกของอันดับที่มีเครื่องหมายลบมีค่าเล็กกว่า ดังนั้นสถิติทดสอบคือ $S_- = 16.5$

ค่าวิกฤติของสถิติทดสอบเปิดจากตาราง Probability Levels for the Wilcoxon signed-rank test (Daniel , 1995) ที่ $n = 12$, $\frac{\alpha}{2} = .026$ ซึ่งเป็นค่าที่ใกล้ที่สุดกับค่า .025 ในตารางได้ค่า $T = 14$

กฎการตัดสินใจที่สมมติฐานแย้งแตกต่างกัน 3 แบบ คือ

1) $H_1 : \theta \neq \theta_0$ จะปฏิเสธ $H_0 : \theta = \theta_0$ ที่ระดับนัยสำคัญ α ถ้าค่าสถิติทดสอบที่คำนวณได้น้อยกว่าหรือเท่ากับค่าวิกฤติ T ที่เปิดจากตาราง ที่ตัวอย่างขนาด n และที่ความน่าจะเป็น $\alpha/2$

2) $H_1 : \theta < \theta_0$ จะปฏิเสธ $H_0 : \theta \geq \theta_0$ ที่ระดับนัยสำคัญ α ถ้า S_+ น้อยกว่าหรือเท่ากับค่าวิกฤติ T ที่เปิดจากตาราง ที่ตัวอย่างขนาด n และที่ความน่าจะเป็น α

3) $H_1 : \theta > \theta_0$ จะปฏิเสธ $H_0 : \theta \leq \theta_0$ ที่ระดับนัยสำคัญ α ถ้า S_- น้อยกว่าหรือเท่ากับค่าวิกฤติ T ที่เปิดจากตาราง ที่ตัวอย่างขนาด n และที่ความน่าจะเป็น α

เนื่องจากค่าสถิติทดสอบ S_- ที่คำนวณได้มากกว่าค่าวิกฤติ T ดังนั้นจึงตัดสินใจยอมรับ H_0 นั่นคือ ประชากรของประโยคมีมัธยฐานของความยาวเท่ากับ 30 คำ

เราสามารถใช้โปรแกรม SPSS ช่วยในการคำนวณได้โดยโดยใช้คำสั่ง 2 Related Samples ใน Nonparametric test ข้อมูลอยู่ในแฟ้มข้อมูล wilcox1.sav กำหนดให้ตัวแปร median คือมัธยฐานของประชากรที่กำหนดมีค่าเท่ากับ 30 และตัวแปร words คือ จำนวนคำในประโยคแรกของย่อหน้า สมมติฐานที่ต้องการทดสอบคือ จำนวนคำในประโยคแรกของย่อหน้าต่าง ๆ มีมัธยฐานเท่ากับ 30 คำ คู่กับ H_1 : จำนวนคำในประโยคแรกของย่อหน้าต่าง ๆ มีมัธยฐานไม่เท่ากับ 30 คำ ทดสอบสมมติฐานโดยใช้การทดสอบ Wilcoxon signed – rank test มีรูปแบบข้อมูลในแฟ้มดังตารางที่ 13.6

ตารางที่ 13.6 รูปแบบการบันทึกข้อมูลจำนวนคำในประโยคแรกของย่อหน้าที่สุ่มมา 12 ย่อหน้า

median	words
30	12
30	18
30	24
30	26
30	37
30	40
30	42
30	47
30	49
30	49
30	78
30	108

ขั้นตอนการใช้คำสั่ง 2 Related Samples... สำหรับการทดสอบ Wilcoxon signed - rank test เหมือนกับขั้นตอนที่ 1 และ 2 ที่อธิบายไว้แล้วในตอนที่ 2.3 ยกเว้น ในกรอบ Test Type เลือก ☐ Wilcoxon เพื่อทำการทดสอบ Wilcoxon signed rank test จะได้ผลลัพธ์ดังภาพที่ 13.2

NPar Tests
Wilcoxon Signed Ranks Test

Ranks

		N	Mean Rank	Sum of Ranks
words - median	Negative Ranks	4[a]	4.13	16.50
	Positive Ranks	8[b]	7.69	61.50
	Ties	0[c]		
	Total	12		

a. words < median

b. words > median

c. words = median

Test Statistics[b]

	words - median
Z	-1.766[a]
Asymp. Sig. (2-tailed)	.077

a. Based on negative ranks.

b. Wilcoxon Signed Ranks Test

ภาพที่ 13.2

จากภาพผลลัพธ์ ดูในตาราง Ranks แสดงค่าสถิติ Sum of Ranks คือ ผลบวกของอันดับของข้อมูลที่มีเครื่องหมายลบ เท่ากับ 16.5 และผลบวกของอันดับของข้อมูลที่มีเครื่องหมายบวกเท่ากับ 61.5 และตัวอย่างมีขนาดเท่ากับ 12

ผลการทดสอบ Wilcoxon signed rank test อยู่ในตาราง Test Statistics ได้ค่าสถิติ Z เท่ากับ -1.766 ซึ่งคิดจากค่าสถิติทดสอบคือ ผลบวกของ

อันดับของข้อมูลที่มีเครื่องหมายลบ และ Asymp. Sig. (2-tailed) เท่ากับ .077 ซึ่งมีค่ามากกว่าระดับนัยสำคัญที่กำหนด ($\alpha = .05$) ดังนั้นจึงยอมรับ H_0 นั่นคือความยาวของประโยคแรกมีมัธยฐานเท่ากับ 30 คำ

3. การทดสอบเกี่ยวกับค่ากลางของประชากร 2 กลุ่ม ที่เป็นอิสระกัน

3.1 การทดสอบแมนวิทเนย์ (Mann – Whitney U Test)

การทดสอบที่นิยมใช้มากที่สุดในการทดสอบเกี่ยวกับความเท่ากันของค่ามัธยฐานของ 2 ประชากรที่มีการแจกแจงแบบต่อเนื่องและเป็นอิสระกัน คือ การทดสอบ Mann – Whitney U Test ตัวแปรที่สนใจศึกษามีสเกลการวัดแบบอันดับ (ordinal scale) หรือสูงกว่า โดยคำนวณค่าสถิติ U เราอาจเปรียบเทียบการทดสอบ Mann – Whitney U Test ได้กับการทดสอบ Independent Sample T – Test Group ในการวิเคราะห์ของสถิติอิงพารามิเตอร์ ซึ่งมีข้อตกลงเบื้องต้นเกี่ยวกับตัวแปรที่สนใจศึกษาที่มีสเกลการวัดแบบช่วง (interval scale) และมีการแจกแจงแบบปกติ นั่นคือเมื่อไม่ทราบการแจกแจงของข้อมูลหรือข้อมูลไม่ได้มีการแจกแจงแบบปกติ เราสามารถใช้การทดสอบ Mann – Whitney U Test แทนการทดสอบ Independent Sample T – Test ได้ และแทนที่จะทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากร 2 กลุ่ม คือ $H_0 : \mu_1 = \mu_2$ จะเปลี่ยนมาเป็นการทดสอบสมมติฐานเกี่ยวกับค่ากลางของประชากร 2 กลุ่ม คือ H_0 : ประชากรทั้ง 2 กลุ่ม มีมัธยฐานเท่ากัน คู่กับ สมมติฐานแย้งคือ H_1 : ตัวอย่างกลุ่มหนึ่งมาจากประชากรที่มีมัธยฐานมากกว่า หรือเขียนเป็นสัญลักษณ์ได้คือ $H_0 : M_A = M_B$ คู่กับ $H_1 : M_A \neq M_B$ หรือกลุ่มตัวอย่าง 2 กลุ่มมาจากประชากรเดียวกันหรือไม่ หรือมาจากประชากรที่มีการแจกแจงเหมือนกันหรือไม่

ตัวอย่างเช่น การศึกษาเปรียบเทียบตัวแบบของเครื่องคิดเลข 2 ตัวแบบคือ ตัวแบบ A ค่าของฟังก์ชันตรีโกณหาได้จากการใส่ตัวเลขก่อนกดปุ่มฟังก์ชัน และตัวแบบ B ต้องเลือกฟังก์ชันก่อนใส่ตัวเลข ครูคณิตศาสตร์อยากทราบว่าตัวแบบ A ทำการคำนวณได้เร็วกว่าตัวแบบ B หรือไม่ กลุ่มตัวอย่างคือ นักเรียน 21 คน แบ่งเป็น 2 กลุ่ม กลุ่ม A ให้ใช้เครื่องคิดเลขตัวแบบ A จำนวน 10 คน และกลุ่ม B ให้ใช้เครื่องคิดเลขตัวแบบ B จำนวน 11 คน ให้นักเรียนทุกคนคำนวณโจทย์ตรีโกณเหมือนกัน แล้ววัดเวลาในการคำนวณเป็นนาที

ได้ข้อมูลดังตารางที่ 13.7 (P. Sprent and N.C. Smeeton, 2001) อยากทราบว่าเครื่องคิดเลขทั้ง 2 แบบนี้แตกต่างกันหรือไม่

ตารางที่ 13.7 ข้อมูลเวลา(นาที) ของการคำนวณค่าของเครื่องคิดเลขที่มีตัวแบบแตกต่างกัน 2 ตัวแบบ

นักเรียนกลุ่ม A	23	18	17	25	22	19	31	26	29	33	
นักเรียนกลุ่ม B	21	28	32	30	41	24	35	34	27	39	36

สมมติฐานที่ต้องการทดสอบคือ H_0 : ประชากรเครื่องคิดเลขทั้ง 2 ตัวแบบมีมัธยฐานของเวลาในการคำนวณเท่ากัน คู่กับ H_1 : เครื่องคิดเลขตัวแบบหนึ่งประชากรมีมัธยฐานของเวลาในการคำนวณมากกว่า ถ้าประชากรทั้ง 2 กลุ่มเป็นประชากรเดียวกัน คือ มีค่ากลางเท่ากัน อันดับที่ของข้อมูลทั้ง 2 กลุ่มจะต้องเป็นไปอย่างสุ่ม

วิธีการคำนวณคือ การให้อันดับที่แก่ข้อมูลทั้ง 2 กลุ่ม โดยถือเสมือนว่าตัวอย่างทั้ง 2 กลุ่ม เป็นกลุ่มเดียวกัน แล้วเรียงลำดับจากน้อยไปมาก กรณีที่ข้อมูลมีค่าเท่ากันหลายค่าให้ใช้อันดับที่เฉลี่ยของข้อมูลที่เท่ากัน

ขนาดกลุ่มตัวอย่างของกลุ่ม A คือ $m = 10$ และกลุ่ม B คือ $n = 11$ จัดอันดับของเวลาในการคำนวณของนักเรียนกลุ่ม A และกลุ่ม B ได้ดังตารางที่ 13.8

ตารางที่ 13.8 การจัดอันดับข้อมูลเวลาของการคำนวณค่าของเครื่องคิดเลขที่มีตัวแบบแตกต่างกัน 2 ตัวแบบ

นักเรียน	อันดับของเวลาในการคำนวณ											ผลบวก
กลุ่ม A	1	2	3	5	6	8	9	12	14	16		$S_m = 76$
กลุ่ม B	4	7	10	11	13	15	17	18	19	20	21	$S_n = 155$

ฟังก์ชันของผลบวกของอันดับของกลุ่มตัวอย่างแทนด้วย S คำนวณค่าสถิติทดสอบ U ของกลุ่มตัวอย่างกลุ่ม A และกลุ่ม B ได้จากสูตร

$$U_m = S_m - \frac{1}{2}m(m+1)$$

$$U_n = S_n - \frac{1}{2}n(n+1)$$

ผลบวกของอันดับทั้งหมดจาก 1 ถึง $m + n$ เท่ากับ $\frac{1}{2}(m+n)(m+n+1) = S_m + S_n$ อาศัยความสัมพันธ์ของ U_m และ U_n ที่แต่ละตัวมีค่าน้อยที่สุดคือ 0 และ $U_m = mn - U_n$ ดังนั้นเราต้องการคำนวณเพียงตัวเดียวเท่านั้นใน 2 ตัวนี้ นั่นคือเราสามารถใช้ U_m หรือ U_n ตัวใดก็ได้ในการทดสอบสมมติฐาน ถึงแม้ว่าในตารางทั่วไปจะให้ค่าสถิติ U_m คือ ใช้สถิติทดสอบของตัวอย่างที่มีขนาดเล็ก การคำนวณ U_n และ U_m คือ

$$\begin{aligned} U_n &= S_n - \frac{1}{2}n(n+1) \\ &= 155 - \frac{1}{2}(11)(12) \\ &= 89 \\ U_m &= mn - U_n \\ &= (10)(11) - 89 \\ &= 21 \end{aligned}$$

ค่าวิกฤตเปิดจากตาราง Quantiles of the Mann – Whitney Test Statistic [จาก Daniel (1995)] ที่ $m = 10$, $n = 11$ และ $\alpha = .05$ ได้ค่าวิกฤติ $U_{.025} = 27$

กฎการตัดสินใจที่สมมติฐานแย้งแตกต่างกัน 3 แบบ คือ

1) $H_1 : M_A < M_B$ จะปฏิเสธ $H_0 : M_A \geq M_B$ ที่ระดับนัยสำคัญ α ถ้าค่า U ที่คำนวณได้น้อยกว่า U_α

2) $H_1 : M_A > M_B$ จะปฏิเสธ $H_0 : M_A \leq M_B$ ที่ระดับนัยสำคัญ α ถ้าค่า U ที่คำนวณได้มากกว่า $U_{1-\alpha}$ เมื่อ $U_{1-\alpha} = mn - U_\alpha$

3) $H_1 : M_A \neq M_B$ จะปฏิเสธ $H_0 : M_A = M_B$ ที่ระดับนัยสำคัญ α ถ้าค่า U ที่คำนวณได้น้อยกว่า $U_{\alpha/2}$ หรือมากกว่า $U_{1-\alpha/2}$ เมื่อ $U_{1-\alpha/2} = mn - U_{\alpha/2}$

เปรียบเทียบค่า U_m กับค่าวิกฤติ $U_{.025} = 27$ พบว่าค่าค่าสถิติทดสอบ U_m ที่คำนวณได้น้อยกว่าค่าวิกฤติ จึงสรุปได้ว่าปฏิเสธ H_0 นั่นคือ นักเรียนกลุ่ม B มีมัธยฐานของเวลาในการคำนวณค่าของฟังก์ชันตรีโกณมากกว่า หรือเครื่องคิดเลขที่ใช้ตัวแบบ B มีมัธยฐานของเวลาในการคำนวณค่าของฟังก์ชันตรีโกณมากกว่า

เราสามารถใช้โปรแกรม SPSS ช่วยในการคำนวณได้โดยใช้คำสั่ง **2 Independent Sample**... ใน **Nonparametric Tests**

3.2 การทดสอบวิลคอกซัน (Wilcoxon rank sum test)

วิธีการของการทดสอบนี้คือ รวมข้อมูลของกลุ่มตัวอย่างทั้ง 2 กลุ่มเข้าด้วยกัน แล้วจัดอันดับของข้อมูลทั้งหมด หลังจากนั้นแยกข้อมูลของแต่ละกลุ่มตัวอย่างออกเป็น 2 กลุ่มตามเดิม โดยมีอันดับของข้อมูลติดอยู่ด้วย สมมติฐานที่ต้องการทดสอบคือ กลุ่มตัวอย่าง 2 กลุ่มมาจากประชากรที่เหมือนกัน และสมมติฐานแย้งคือ การแจกแจงของประชากรมีความแตกต่างเกี่ยวกับค่ากลางหรือมัธยฐานเท่านั้น จากการศึกษาเปรียบเทียบตัวแบบของเครื่องคิดเลข 2 ตัวแบบ ถ้าตัวอย่างทั้ง 2 กลุ่มมาจากประชากรที่เหมือนกัน ซึ่งอาจมีรูปแบบต่อเนื่องแบบใด ๆ และไม่จำเป็นต้องมีการแจกแจงแบบสมมาตร เราคาดหวังว่าข้อมูลของกลุ่มตัวอย่างทั้ง 2 กลุ่ม จะมีข้อมูลที่มีอันดับต่ำ ปานกลาง สูง ปนเปกันไป และภายใต้สมมติฐานแย้งเราคาดหวังว่าในประชากรหนึ่งจะมีข้อมูลที่มีอันดับต่ำกว่า และในอีกประชากรหนึ่งจะมีข้อมูลที่มีอันดับสูงกว่าเด่นออกมา ซึ่งเนื่องมาจากอิทธิพลของทรีทเมนต์

การคำนวณค่าสถิติทดสอบ W คิดจากผลบวกของค่าอันดับของข้อมูลในตัวอย่างกลุ่มเล็ก (S_m) จากตารางที่ 13.8 กลุ่ม A ได้ผลบวกของค่าอันดับของข้อมูลคือ $S_m = 76$ แต่ถ้าตัวอย่างทั้ง 2 กลุ่มมีขนาดเท่ากัน ค่าสถิติทดสอบ W ได้แก่ ผลรวมของค่าอันดับของข้อมูลในกลุ่มตัวอย่างที่ถูกกำหนดให้เป็นกลุ่มแรก

ทั้งการทดสอบ Wilcoxon rank sum test และการทดสอบ Mann – Whitney U test จะให้ผลเหมือนกัน เพราะทั้ง 2 การทดสอบนี้ใช้วิธีคิดคนละแบบ แต่คำนวณได้ตัวสถิติเดียวกัน

3.3 การใช้คำสั่ง 2 Independent Sample... ทดสอบเกี่ยวกับค่ากลางของประชากร 2 กลุ่มที่เป็นอิสระกัน

จากตัวอย่างการศึกษาเปรียบเทียบตัวแบบของเครื่องคิดเลข 2 ตัวแบบในการคำนวณค่าของฟังก์ชันตรีโกณ กำหนดให้ตัวแปรกลุ่มตัวอย่างนักเรียนแทนด้วย group (1 = group A และ 2 = groupB) และตัวแปร min แทนเวลาในการคำนวณค่าฟังก์ชันตรีโกณคิดเป็นนาทีที่มีสเกลการวัดแบบอัตราส่วน (ratio scale) บันทึกข้อมูลในแฟ้มข้อมูล Nonpara1.sav มีรูปแบบของข้อมูลในแฟ้มดังตารางที่ 13.9

ตารางที่ 13.9 รูปแบบการบันทึกข้อมูลเวลาในการคำนวณค่าของเครื่องคิดเลขที่มีตัวแบบแตก ต่างกัน 2 ตัวแบบ

group	min	group	min
1	23	2	21
1	18	2	28
1	17	2	32
1	25	2	30
1	22	2	41
1	19	2	24
1	31	2	35
1	26	2	34
1	29	2	27
1	33	2	39
		2	36

การใช้โปรแกรม SPSS ช่วยในการคำนวณ สำหรับการทดสอบ Mann – Whitney U test มีขั้นตอนดังนี้

1. ไปที่เมนูบาร์ คลิกที่ Analyze, Nonparametric Tests, 2 Independent Samples... จะได้หน้าต่าง Two – Independent – Samples Tests

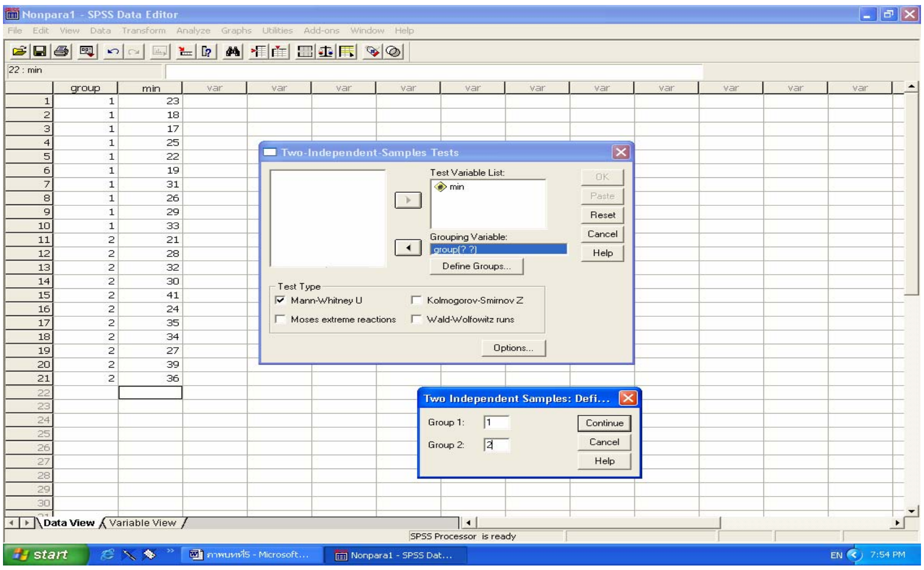
2. ในหน้าต่าง Two – Independent – Samples Tests ในช่องซ้ายมือ คลิกที่ตัวแปร min แล้วคลิกที่หัวลูกศร ➢ หน้าช่อง Test Variable List : ตัวแปร min จะย้ายเข้าไปอยู่ในช่อง Test Variable List :

คลิกที่ตัวแปร group แล้วคลิกที่หัวลูกศร ➢ หน้าช่อง Grouping Variable : ตัวแปร group จะย้ายเข้าไปอยู่ในช่องนี้ แล้วคลิกที่ปุ่ม Define Groups... จะได้หน้าต่าง Two Independent Samples : Define Groups ในช่อง Group 1 : ใส่เลข 1 หมายถึงกลุ่ม A และในช่อง Group 2 : ใส่เลข 2 หมายถึงกลุ่ม B ดังภาพที่ 13.3 แล้วคลิกที่ปุ่ม Continue หน้าต่างนี้จะถูกปิดไป

ในกรอบ Test Type

เลือก ☐ Mann – Whitney U เพื่อให้ได้ค่าสถิติทดสอบ Mann – Whitney U ในการทดสอบสมมติฐานเกี่ยวกับค่ากลางของเวลาในการคำนวณค่าฟังก์ชันตรีโกณของเครื่องคิดเลขที่มีตัวแบบแตกต่างกัน 2 ตัวแบบ

แล้วคลิกที่ปุ่ม OK จะได้ผลลัพธ์ดังภาพที่ 13.4



ภาพที่ 13.3

NPar Tests
Mann-Whitney Test

Ranks

	group	N	Mean Rank	Sum of Ranks
min	groupA	10	7.60	76.00
	groupB	11	14.09	155.00
	Total	21		

Test Statistics[b]

	min
Mann-Whitney U	21.000
Wilcoxon W	76.000
Z	-2.394
Asymp. Sig. (2-tailed)	.017
Exact Sig. [2*(1-tailed Sig.)]	.016[a]

a. Not corrected for ties.

b. Grouping Variable: group

ภาพที่ 13.4

จากภาพผลลัพธ์ในตาราง Ranks สำหรับตัวแปร min ของนักเรียนกลุ่ม A มีจำนวน N = 10 และนักเรียนกลุ่ม B มีจำนวน N = 11 ในช่อง Mean Rank แสดงค่าเฉลี่ยของอันดับที่ของเวลาการคำนวณค่าของนักเรียนกลุ่ม A เท่ากับ 7.60 และของนักเรียนกลุ่ม B เท่ากับ 14.09 ในช่อง Sum of Ranks แสดงผลบวกของอันดับที่ของเวลาการคำนวณค่าของนักเรียนกลุ่ม A เท่ากับ 76 และของนักเรียนกลุ่ม B เท่ากับ 155

ผลการทดสอบ Mann – Whitney U test สำหรับประชากร 2 กลุ่มที่เป็นอิสระกัน ทำการทดสอบสมมติฐานแบบสองทางอยู่ในตาราง Test Statistics ได้ค่าสถิติทดสอบ Mann – Whitney U เท่ากับ 21 , ค่าสถิติทดสอบ Wilcoxon W เท่ากับ 76 , และค่าสถิติทดสอบ Z เท่ากับ -2.394 พิจารณาค่า Asymp. Sig. (2

– tailed) ของการทดสอบได้เท่ากับ .017 ซึ่งน้อยกว่าระดับนัยสำคัญที่กำหนด (α = .05) จึงปฏิเสธ H_0 นั่นคือ เครื่องคิดเลขทั้ง 2 ตัวแบบ ที่ใช้คำนวณค่าฟังก์ชันตรีโกณมีค่ากลางของเวลาในการคำนวณแตกต่างกัน

สำหรับกรณีทดสอบสมมติฐานแบบทางเดียว คือ $H_1 : M_A > M_B$ หรือ $H_1 : M_A < M_B$ จะคำนวณค่า P-value = $\frac{.017}{2}$

4. การทดสอบเกี่ยวกับค่ากลางของประชากรหลายกลุ่มที่เป็นอิสระกัน

4.1 การทดสอบครัสคอล-วอลลิส (Kruskal – Wallis test)

Kruskal และ Wallis ได้ขยายวิธีการทดสอบ Mann – Whitney U Test สำหรับใช้กับกลุ่มตัวอย่าง 3 กลุ่มหรือมากกว่า การทดสอบนี้ใช้ได้กับข้อมูลที่มีสเกลการวัดแบบอันดับขึ้นไป เป็นการทดสอบสมมติฐานเกี่ยวกับความเท่ากันของค่ากลางหรือมัธยฐานของประชากร เมื่อตัวอย่างกลุ่มต่าง ๆ มาจากประชากรกลุ่มต่าง ๆ ที่เหมือนกัน และมีการแจกแจงแบบต่อเนื่อง และยังใช้ได้ดีกับการทดสอบสมมติฐานเกี่ยวกับความเหมือนกันของฟังก์ชันการแจกแจงสะสมของประชากรกลุ่มต่าง ๆ คู่กับสมมติฐานแย้งที่ว่ามีประชากรอย่างน้อย 1 กลุ่มมีการแจกแจงสะสมแตกต่างจากประชากรกลุ่มอื่น ๆ การทดสอบ Kruskal – Wallis test คล้ายกับการทดสอบ F-test ในการวิเคราะห์ความแปรปรวน (ANOVA)ในการวิเคราะห์แบบอิงพารามิเตอร์

สมมติว่ามีตัวอย่าง k กลุ่ม ตัวอย่างกลุ่มที่ i (i = 1, 2, ..., k) ประกอบด้วย ค่าสังเกตจำนวน n_i กำหนดให้ค่าสังเกตตัวที่ j เขียนแทนด้วย x_{ij} (j = 1, 2, ... , n_i) จำนวนค่าสังเกตทั้งหมดคือ $N = \sum_i n_i$ สมมติว่าไม่มีข้อมูลที่มีค่าซ้ำกัน และให้อันดับค่าสังเกตทั้ง N ตัวจากค่าที่เล็กที่สุดเป็นอันดับ 1 ถึงค่าที่ใหญ่ที่สุดเป็นอันดับ N กำหนดให้ r_{ij} เป็นอันดับของ x_{ij} และ $S_i = \sum_j r_{ij}$ คือ ผลบวกของอันดับของตัวอย่างกลุ่มที่ i คำนวณค่า $S_k = \sum_i (S_i^2/n_i)$ สถิติทดสอบคือ

$$T = \frac{12\,S_k}{N(N+1)} - 3(N+1)$$

ถ้าตัวอย่างทุกกลุ่มมาจากประชากรเดียวกัน สมมติฐานที่ต้องการทดสอบคือ H_0 : ในแต่ละกลุ่มตัวอย่างจะมีค่าสังเกตที่มีอันดับปน ๆ กัน ทั้งเล็ก กลาง และสูง คู่กับ H_1 : มีกลุ่มตัวอย่างอย่างน้อย 1 กลุ่ม จะมีค่าสังเกตที่มีอันดับสูง (หรือต่ำ) เด่นชัดกว่ากลุ่มอื่น ๆ หรือ H_0 : ค่ากลางหรือมัธยฐานของประชากรทั้ง k กลุ่ม เท่ากัน คู่กับ H_1 : มีประชากรอย่างน้อย 1 กลุ่มที่มีค่ากลางแตกต่างจากกลุ่มอื่น ๆ หรือ $H_0 : M_1 = M_2 = \ldots = M_k$ คู่กับ $H_1 : M_i \neq M_j$ อย่างน้อย 1 คู่ ที่ $i \neq j$ และ S_k จะประกอบด้วยกำลังสองของผลบวกของอันดับของอย่างน้อย 1 กลุ่มนั้น ซึ่งทำให้ค่าสถิติ T มีค่ามากขึ้น สำหรับกลุ่มตัวอย่างขนาดใหญ่ สถิติทดสอบ T มีการแจกแจงแบบไคสแควที่มี $df = k - 1$

ตัวอย่างเช่น การศึกษาเปรียบเทียบราคาขายบ้านในหมู่บ้าน A, B, และ C คิดเป็นปอนด์ (£K) มีราคาขายดังนี้ [จาก P.Sprent and N.C. Smeeton (2001)] ข้อมูลอยู่ในตารางที่ 13.10

ตารางที่ 13.10 ข้อมูลราคาขายบ้าน (£K) ในหมู่บ้าน 3 แห่ง

หมู่บ้าน A	39	45	71	
หมู่บ้าน B	51	63	88	97
หมู่บ้าน C	99	150	260	

ต้องการทดสอบสมมติฐานเกี่ยวกับราคาขายบ้านใน หมู่บ้าน 3 แห่ง ว่ามาจากประชากรเดียวกันหรือไม่ หรือ H_0 : ทั้ง 3 หมู่บ้าน มีมัธยฐานของราคาขายบ้านเท่ากัน คู่กับ H_1 : มีอย่างน้อย 1 หมู่บ้าน ที่มีมัธยฐานไม่เท่ากับหมู่บ้านอื่น

ตารางที่ 13.11 ข้อมูลราคาขายบ้าน (£K) และค่าอันดับของราคาขายของหมู่บ้าน 3 แห่ง

หมู่บ้าน A	ราคา	39	45	71	
	อันดับ	1	2	5	ผลบวกอันดับ = 8

หมู่บ้าน B	ราคา	51	63	88	97	
	อันดับ	3	4	6	7	ผลบวกอันดับ = 20
หมู่บ้าน C	ราคา	99	150	260		
	อันดับ	8	9	10		ผลบวกอันดับ = 27

คำนวณ $S_k = (8)^2/3 + (20)^2/4 + (27)^2/3 = 364.33$

คำนวณค่าสถิติทดสอบ $T = \frac{12(364.33)}{10(11)} - 3(11)$

$= 6.745$

ค่าวิกฤติเปิดจากตาราง Percentiles of the Chi-square Distribution [จาก Daniel (1995)] ที่ $df = k - 1 = 3 - 1 = 2$ และ $\alpha = .05$ ได้ค่าวิกฤติ $\chi^2_2 = 5.991$ กฎการตัดสินใจที่ระดับนัยสำคัญ α คือ ปฏิเสธ H_0 ถ้าสถิติทดสอบ T มากกว่า $\chi^2_{(k-1),1-\alpha}$

เปรียบเทียบค่าสถิติทดสอบ T กับค่าวิกฤติ $\chi^2_{2;.95}$ พบว่าค่า T มากกว่าค่าวิกฤติ จึงสรุปได้ว่าปฏิเสธ H_0 นั่นคือ ราคาขายบ้านของหมู่บ้านอย่างน้อย 1 แห่งแตกต่างจากหมู่บ้านอื่น ๆ

เราสามารถใช้โปรแกรม SPSS ช่วยในการคำนวณได้โดยใช้คำสั่ง K – Independent Samples...

4.2 การใช้คำสั่ง K – Independent Samples... สำหรับ Kruskal – Wallis test

จากตัวอย่างการศึกษาเปรียบเทียบราคาขายบ้านในหมู่บ้าน 3 แห่ง กำหนดให้หมู่บ้านแทนด้วยตัวแปร village (1 = villageA, 2 = village B และ 3 = villageC) และราคาขายบ้านแทนด้วยตัวแปร price ซึ่งข้อมูลคือราคาขายบ้านที่มีสเกล

การวัดแบบอัตราส่วน บันทึกข้อมูลในแฟ้มข้อมูล Nonpara 2.sav มีรูปแบบของข้อมูลในแฟ้มดังตารางที่ 13.12

ตารางที่ 13.12 รูปแบบการบันทึกข้อมูลราคาขายบ้านในหมู่บ้าน 3 แห่ง

village	price	village	price	village	price
1	39	2	51	2	97
1	45	2	63	3	99
1	71	2	88	3	150
				3	260

การใช้โปรแกรม SPSS ช่วยในการคำนวณ สำหรับการทดสอบ Kruskal – Wallis test มีขั้นตอนดังนี้

1. ไปที่เมนูบาร์ คลิกที่ Analyze, Nonparametric Tests, K Independent Samples... จะได้หน้าต่าง Tests for Several Independent Samples

2. ในหน้าต่าง Tests for Several Independent Samples ในช่องซ้ายมือ

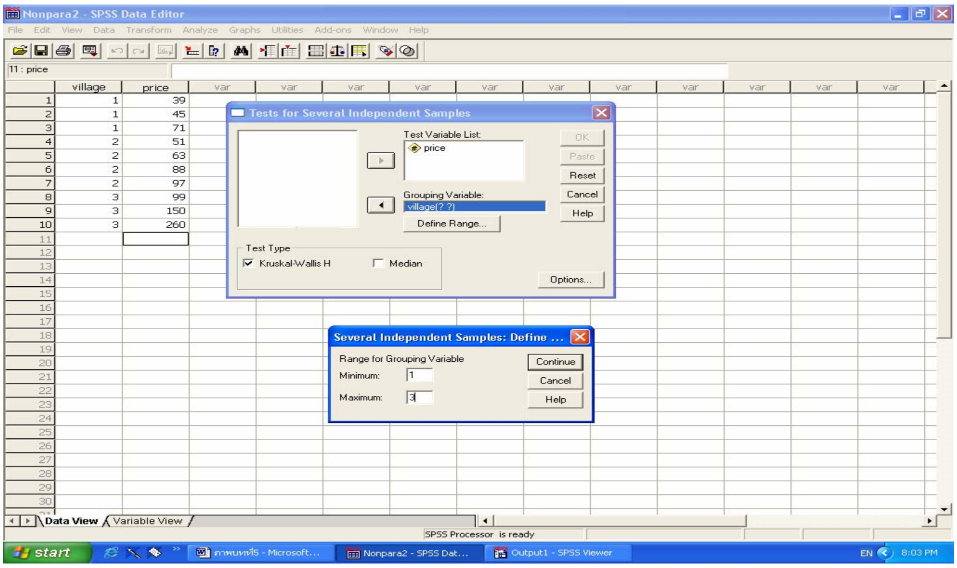
คลิกที่ตัวแปร price ให้ย้ายเข้าไปอยู่ในช่องของ Test Variable List :

คลิกที่ตัวแปร village ให้ย้ายเข้าไปอยู่ในช่องของ Grouping Variable : แล้วคลิกที่ปุ่ม Define Range... จะได้หน้าต่าง Several Independent Samples : Define Range ในช่อง Mininum : ใส่เลข 1 หมายถึง หมู่บ้าน $\boxed{A}$ และในช่อง Maximum : ใส่เลข 3 หมาย$\boxed{\text{ถึง}}$ หมู่บ้าน C ดังภาพที่ 13.5 แล้วคลิกที่ปุ่ม Continue หน้าต่างนี้จะถูกปิดไป

ในกรอบ Test Type

เลือก ☐ Kruskal – Wallis H เพื่อให้ได้ค่าสถิติทดสอบ Kruskal – Wallis H ในการทดสอบสมมติฐานเกี่ยวกับราคาบ้านในตัวอย่าง 3 หมู่บ้าน ว่ามาจากประชากรเดียวกัน

แล้วคลิกที่ปุ่ม OK จะได้ผลลัพธ์ดังภาพที่ 13.6



ภาพที่ 13.5

NPar Tests

Kruskal-Wallis Test

Ranks

	village	N	Mean Rank
price	villageA	3	2.67
	villageB	4	5.00
	villageC	3	9.00
	Total	10	

Test Statistics[a,b]

	price
Chi-Square	6.745
df	2
Asymp. Sig.	.034

a. Kruskal Wallis Test

b. Grouping Variable: village

ภาพที่ 13.6

จากภาพผลลัพธ์ในตาราง Ranks สำหรับตัวแปร price ในหมู่บ้าน A มีจำนวนข้อมูล N = 3, หมู่บ้าน B มีจำนวนข้อมูล N = 4, และหมู่บ้าน C มี

จำนวนข้อมูล N = 3 ในช่อง Mean Rank แสดงค่าเฉลี่ยของอันดับของหมู่บ้าน A , B และ C เท่ากับ 2.67 , 5.00 , และ 9.00 ตามลำดับ

ผลการทดสอบ Kruskal – Wallis test อยู่ในตาราง Test Statistics ได้ค่าสถิติทดสอบ Chi – Square เท่ากับ 6.745 ที่ df = 2 และ Asymp. Sig. เท่ากับ .034 ซึ่งน้อยกว่าระดับนัยสำคัญที่กำหนด (α = .05) จึงปฏิเสธ H_0 นั่นคือ ค่ามัธยฐานของราคาขายบ้านของหมู่บ้านอย่างน้อย 1 แห่ง แตกต่างจากหมู่บ้านอื่น ๆ

4.3 การทดสอบมัธยฐาน (Median test)

การทดสอบมัธยฐานต้องการทดสอบเพียงว่ามีเหตุผลเพียงพอที่จะบอกได้หรือไม่ว่าประชากรกลุ่มต่าง ๆ มีมัธยฐานเดียวกันซึ่งไม่ทราบค่าไม่ ถ้ากลุ่มตัวอย่างขนาด m มาจากประชากรต่าง ๆ (ไม่จำเป็นต้องมีลักษณะเหมือนกันทั้งหมด) ซึ่งแต่ละประชากรมีมัธยฐานเดียวกัน ซึ่งไม่ทราบค่า คือ θ_1 และกลุ่มตัวอย่างขนาด n มาจากประชากรต่าง ๆ ซึ่งแต่ละประชากรมีมัธยฐานเดียวกันซึ่งไม่ทราบค่าคือ θ_2 เราต้องการทดสอบสมมติฐาน $H_0 : \theta_1 = \theta_2$ คู่กับ $H_1 : \theta_1 \neq \theta_2$ หรือ $H_1 : \theta_1 > \theta_2$

ถ้าประชากรทั้งหมดมีมัธยฐานเดียวกันคือ θ แล้วมัธยฐานของตัวอย่างที่ได้มารวม ๆ กันเท่ากับ M เป็นตัวประมาณค่าของ θ การทดสอบมัธยฐานเป็นการตรวจสอบสำหรับแต่ละตัวอย่างว่ามีจำนวนกี่ค่าที่มีค่ามากกว่าและน้อยกว่าค่ามัธยฐาน M ถ้าตัวอย่างมาจากประชากรต่าง ๆ ซึ่งมีมัธยฐานเดียวกัน การแจกแจงของจำนวนในแต่ละค่าที่มีค่ามากกว่ามัธยฐาน M จะประมาณได้ด้วยการแจกแจงทวินาม $B(m , .5)$ และ $B(n , .5)$

ในกรณีที่มีกลุ่มตัวอย่าง 2 กลุ่ม ตัวอย่างแต่ละกลุ่มอาจมาจากประชากรใด ๆ ที่ไม่กำหนดคือ ประชากรทั้งหลายไม่ต้องมีการแจกแจงแบบเดียวกัน ถ้ามีตัวอย่าง k กลุ่ม ที่มีมัธยฐานของ ประชากรซึ่งไม่ทราบค่าคือ $\theta_1 , \theta_2 , \ldots , \theta_k$ สมมติฐานที่ต้องการทดสอบคือ $H_0 : \theta_1 = \theta_2 = \ldots = \theta_k$ คู่กับ $H_1 : \theta_i$ เท่ากันไม่ทุกตัว ซึ่งไม่เหมือนกับสถานการณ์ในการทดสอบ Kruskal – Wallis test

วิธีการทดสอบ ถ้า M คือ มัธยฐานของค่าสังเกตทั้งหมดของตัวอย่างที่รวม ๆ กัน ใน ตัวอย่างแต่ละกลุ่มเรานับจำนวนของค่าสังเกตที่อยู่เหนือและต่ำกว่า M เราตัดทิ้งค่า

ของตัวอย่างที่เท่ากับ M แล้ววิเคราะห์ตัวอย่างที่เหลืออยู่ สมมติว่าเอาค่าของตัวอย่างที่เท่ากับ M ออกแล้ว เหลือค่าสังเกตจำนวน n_i ในตัวอย่างที่ i มีจำนวน a_i ที่มีค่ามากกว่า M และมีจำนวน b_i ที่มีค่าน้อยกว่า M ; $i = 1, 2, \ldots , P$ ทำการบันทึกจำนวนเหล่านี้ทั้งมากและน้อยกว่า M ในตารางขนาด $k \times 2$ ที่มี k แถว และ 2 คอลัมน์ ใส่ค่า a_i และ b_i ของตัวอย่างแต่ละกลุ่ม ดังนั้น $a_i + b_i = n_i$ เมื่อไม่มีค่าสังเกตที่เท่ากับค่า M แล้ว ผลรวมทางคอลัมน์ A และ B ทั้ง 2 ตัว เท่ากับ $(.5)\ N$ ดังแสดงในตารางที่ 13.13

ตารางที่ 13.13 ตารางสำหรับการทดสอบมัธยฐาน

ค่าที่อยู่เหนือ M	ค่าที่อยู่ต่ำกว่า M	ผลรวม
a_1	b_1	n_1
a_2	b_2	n_2
.	.	.
.	.	.
.	.	.
a_k	b_k	n_k
A	B	N

ความน่าจะเป็น P^* ของค่าในช่องต่าง ๆ ในตารางข้างต้นคิดได้จากสูตร

$$P^* = \frac{A!\,B!\,\pi_i(n_i!)}{N!\,\pi_i(a_i!)\,\pi_i(b_i!)}$$

เมื่อ $\pi_i\,(x_i!)$ คือผลคูณของ $(x_1!)(x_2!)\ldots(x_k!)$

ตัวอย่าง การศึกษาเปรียบเทียบศัลยแพทย์ฟัน 6 คน ที่ถอนฟันกรามในคนไข้ที่เป็นผู้ใหญ่ ต้องการทดสอบว่าศัลยแพทย์ฟันบางคนใช้เวลาในการผ่าตัดฟันกรามออกได้เร็วกว่าคนอื่น บันทึกข้อมูลเวลาคิดเป็นนาทีที่ศัลยแพทย์ฟันทั้ง 6 คน ใช้ในการผ่าตัดเอาฟันกรามออกไปของคนไข้ทั้งหมด 28 คน ได้ข้อมูลดังตารางที่ 13.14

ตารางที่ 13.14 ข้อมูลเวลา (นาที) ในการผ่าตัดฟันกรามของศัลยแพทย์ 6 คน

ศัลยแพทย์ฟัน	เวลา (นาที)						
1	23	25	34	45			
2	7	11	14	15	17	24	40
3	5	8	16	20	26		
4	12	21	31	38			
5	30	43					
6	4	9	10	18	19	35	

จากข้อมูลหาค่ามัธยฐานของเวลา M ของคนไข้ทั้งหมด 28 คน เท่ากับ 19.5 นาที ดังนั้นไม่มีค่าของตัวอย่างที่เท่ากับมัธยฐาน ตารางที่ 13.15 แสดงจำนวนของเวลาที่มากกว่าและน้อยกว่า M แยกตามศัลยแพทย์ฟันแต่ละคน

ใช้การทดสอบมัธยฐาน ทดสอบสมมติฐาน H_0 : ตัวอย่างทั้ง 6 กลุ่ม มาจากประชากรทั้งหลายที่มีมัธยฐานของเวลาเท่ากัน คู่กับ H_1 : มัธยฐานเท่ากันไม่ทุกตัว ซึ่งเป็นการทดสอบแบบสองทาง

ตารางที่ 13.15 จำนวนค่าสังเกตที่อยู่เหนือและต่ำกว่ามัธยฐาน M ของตัวอย่างทั้งหมด

ศัลยแพทย์ฟัน	จำนวนที่อยู่เหนือ M	จำนวนที่อยู่ต่ำกว่า M	ผลรวม
1	4	0	4
2	2	5	7
3	2	3	5
4	3	1	4
5	2	0	2
6	1	5	6
ผลรวม	14	14	28

คำนวณค่า

$$P^* = \frac{(14!)(14!)(4!)(7!)(5!)(4!)(2!)(6!)}{(28!)(4!)(2!)(2!)(3!)(2!)(1!)(0!)(5!)(3!)(1!)(0!)(5!)}$$

$$= \quad 0.000126$$

สรุปได้ว่ามีหลักฐานชัดเจนในการปฏิเสธ H_0 นั่นคือ มีความแตกต่างของมัธยฐานของเวลาระหว่างศัลยแพทย์ฟันทั้ง 6 คนนี้

เนื่องจากตัวอย่างนี้มีกลุ่มตัวอย่างคนไข้ที่ศัลยแพทย์ฟันแต่ละคนทำการผ่าตัดจำนวนน้อยมาก ในทางปฏิบัติควรใช้กลุ่มตัวอย่างคนไข้มากกว่านี้ในการศึกษา สำหรับตัวอย่างนี้สามารถใช้การทดสอบ Fisher exact test สำหรับการทดสอบมัธยฐานได้

เราสามารถใช้โปรแกรม SPSS ช่วยในการคำนวณสำหรับการทดสอบ Median test โดยใช้คำสั่ง K – Independent Samples...

4.4 การใช้คำสั่ง K – Independent Samples... สำหรับ Median test

จากตัวอย่างการศึกษาเปรียบเทียบศัลยแพทย์ฟัน 6 คน ที่ถอนฟันกรามในคนไข้ที่เป็นผู้ใหญ่ กำหนดให้ศัลยแพทย์ฟันแทนด้วยตัวแปร doctor (1 = ศัลยแพทย์ฟันคนที่ 1, 2 = ศัลยแพทย์ฟันคนที่ 2, ... , 6 = ศัลยแพทย์ฟันคนที่ 6) และเวลาในการผ่าตัดฟันกรามแทนด้วยตัวแปร time ซึ่งข้อมูลคือเวลาเป็นนาทีที่มีสเกลการวัดแบบอัตราส่วน บันทึกข้อมูลในแฟ้มข้อมูล Nonpara3.sav มีรูปแบบของข้อมูลในแฟ้มดังตารางที่ 13.16

ตารางที่ 13.16 รูปแบบการบันทึกข้อมูลเวลา (นาที) ที่ใช้ในการถอนฟันของศัลยแพทย์ฟัน 6 คน

doctor	time	doctor	time	doctor	time
1	23	2	40	5	30
1	25	3	5	5	43
1	34	3	8	6	4
1	45	3	16	6	9
2	7	3	20	6	10
2	11	3	26	6	18
2	14	4	12	6	19
2	15	4	21	6	35
2	17	4	31		
2	24	4	38		

การใช้โปรแกรม SPSS ช่วยในการคำนวณสำหรับการทดสอบ Median test มีขั้นตอนดังนี้

1. ไปที่เมนูบาร์ คลิกที่ Analyze, Nonparametric Test , K independent Samples... จะได้หน้าต่าง Tests for Several Independent Samples

2. ในหน้าต่าง Test for Several Independent Samples ในช่องซ้ายมือ คลิกที่ตัวแปร time ให้ย้ายเข้าไปอยู่ในช่องของ Test Variable List:

คลิกที่ตัวแปร doctor ให้ย้ายเข้าไปอยู่ในช่องของ Grouping Variable : แล้วคลิกที่ปุ่ม Define Range... จะได้หน้าต่าง Several Independent Samples... Define Range ในช่อง Minimum: ใส่เลข 1 หมายถึง ศัลยแพทย์ฟันคนที่ 1 และในช่อง Maximum: ใส่เลข 6 หมายถึง ศัลยแพทย์ฟันคนที่ 6 แล้วคลิกที่ปุ่ม Continue หน้าต่างนี้จะถูกปิดไป

ในกรอบ Test Type

เลือก ☐ Median เพื่อให้ได้ค่าสถิติทดสอบ Median test ในการทดสอบสมมติฐานเกี่ยวกับเวลาการผ่าตัดฟันกรามของศัลยแพทย์ฟัน 6 คน มีมัธยฐานของเวลาเท่ากันหรือไม่

แล้วคลิกที่ปุ่ม OK จะได้ผลลัพธ์ดังภาพที่ 13.7

NPar Tests
Median Test

Frequencies

		doctor					
		doctor1	doctor2	doctor3	doctor4	doctor5	doctor6
time	> Median	4	2	2	3	2	1
	<= Median	0	5	3	1	0	5

Test Statistics[b]

	time
N	28
Median	19.50
Chi-Square	11.152[a]
df	5
Asymp. Sig.	.048

a. 12 cells (100.0%) have expected frequencies less than 5. The minimum expected cell frequency is 1.0.

b. Grouping Variable: doctor

ภาพที่ 13.7

จากภาพผลลัพธ์ในตาราง Frequencies สำหรับตัวแปร time แสดงค่า > median และ <= Median ของเวลาผ่าตัดฟันกรามของศัลยแพทย์ฟัน 6 คน

ผลการทดสอบ Median test อยู่ในตาราง Test Statistics มีจำนวนตัวอย่างคนไข้ทั้งหมด N = 28 คน ที่มีค่ามัธยฐานของเวลาผ่าตัดฟันกรามเท่ากับ 19.50 นาที ได้ค่าสถิติทดสอบ Chi-Square เท่ากับ 11.152 ที่ df = 5 และค่า Asymp. Sig. เท่ากับ .048 ซึ่งน้อยกว่าระดับนัยสำคัญที่กำหนด (α = .05) จึงปฏิเสธ H_0 นั่นคือ ศัลยแพทย์ฟันทั้ง 6 คน มีมัธยฐานของเวลาการผ่าตัดฟันกรามแตกต่างกัน

5. การทดสอบเกี่ยวกับค่ากลางของประชากร 2 กลุ่ม ที่ไม่เป็นอิสระกัน

5.1 การทดสอบเครื่องหมาย (Sign test)

ในการทดสอบสมมติฐานเกี่ยวกับความแตกต่างของค่ามัธยฐานเท่ากับ 0 โดยมีสมมติฐานแย้งคือ $H_1 : P(X_i > Y_i) = P(X_i < Y_i) = .5$ ในกรณีที่มีข้อมูลที่จับคู่กันของกลุ่มตัวอย่าง 2 กลุ่ม สมมติว่าคือ x และ y และไม่เป็นไปตามข้อตกลงเบื้องต้นของการทดสอบ t-test หรือข้อมูลมีสเกลการวัดแบบอันดับ เราสามารถใช้การทดสอบ Sign test

ตัวอย่าง การศึกษาเรื่องการประเมินการสอนประชาชนเกี่ยวกับการแปรงฟัน กลุ่มตัวอย่างคือ คนไข้ในคลินิกฟันจำนวน 12 คู่ จับคู่กันด้วยปัจจัยด้าน อายุ เพศ คะแนนความฉลาด และคะแนนตั้งต้นเกี่ยวกับความสะอาดของช่องปาก ตัวอย่างคนหนึ่งของแต่ละคู่จะได้รับการสอนเกี่ยวกับการแปรงฟันและการรักษาความสะอาดของช่องปาก หลังจากนั้น 6 เดือน ทำการตรวจและให้คะแนนความสะอาดของช่องปาก คนไข้ทั้ง 24 คน ความหมายของคะแนนคือคะแนนน้อย หมายความว่ามีความสะอาดในช่องปากมาก ได้คะแนนของตัวอย่างทั้ง 12 คู่ ดังตารางที่ 13.17 (Daniel, 1995)

ตารางที่ 13.17 คะแนนความสะอาดของช่องปากของตัวอย่างกลุ่มที่ได้รับการสอนและไม่ได้รับ การสอนเกี่ยวกับการแปรงฟันและการรักษาความสะอาดของช่องปาก

ตัวอย่างคู่ที่	คะแนนของกลุ่มที่ได้รับการสอน (x_i)	คะแนนของกลุ่มที่ไม่ได้รับการสอน (y_i)
1	1.5	2.0
2	2.0	2.0
3	3.5	4.0
4	3.0	2.5
5	3.5	4.0
6	2.5	3.0
7	2.0	3.5
8	1.5	3.0
9	1.5	2.5

10	2.0	2.5
11	3.0	2.5
12	2.0	2.5

จากข้อมูลคะแนนที่จับคู่กัน 1 คู่ ให้ y_i ลบออกจาก x_i ถ้า y_i น้อยกว่า x_i จะได้เครื่องหมายบวก และถ้า y_i มากกว่า x_i จะได้เครื่องหมายลบ ถ้าความแตกต่างของมัธยฐานของ x และ y เท่ากับ 0 เราคาดว่าจะมีจำนวนเครื่องหมายบวกเท่ากับจำนวนเครื่องหมายลบ

ข้อมูลคือ เครื่องหมายของความแตกต่างของคะแนนความสะอาดของตัวอย่างทั้ง 12 คู่ ดังตารางที่ 13.18

ตารางที่ 13.18 เครื่องหมายของความแตกต่าง $(x_i - y_i)$ ของคะแนนความสะอาดในช่องปาก ของตัวอย่างคนไข้ 12 คู่ โดย x_i คือคะแนนของคนที่ได้รับการสอน และ y_i คือคะแนนของคนที่ไม่ได้รับการสอน

ตัวอย่างคู่ที่	1	2	3	4	5	6	7	8	9	10	11	12
ความแตกต่างของคะแนน$(x_i - y_i)$	-	0	-	+	-	-	-	-	-	-	+	-

สมมติฐานที่ต้องการทดสอบคือ การสอนวิธีแปรงฟันและการรักษาความสะอาดในช่องปากได้ผล จะได้ความแตกต่างของคะแนน $(x_i - y_i)$ เป็นเครื่องหมายลบมากกว่า แต่ถ้าการสอนไม่ได้ผลค่ามัธยฐานของความแตกต่างของคะแนนจะเท่ากับ 0 เขียนเป็นสมมติฐานได้คือ

H_0 : มัธยฐานของความแตกต่างของคะแนนเท่ากับ 0 หรือ $P(+) = P(-) = .5$

คู่กับ H_1 : มัธยฐานของความแตกต่างของคะแนนเป็นลบ หรือ $P(+) < P(-)$

สถิติทดสอบมีการแจกแจงทวินาม $B(12\ ,\ .5)$ ถ้า H_0 เป็นจริง จากการสังเกต 12 ครั้ง เมื่อความน่าจะเป็นที่จะเกิดเครื่องหมายลบคือ $p = .5$ เราจะตัดสินใจปฏิเสธ H_0 ถ้าความน่าจะเป็นที่จะเกิดเครื่องหมายบวกน้อยกว่าหรือเท่ากับ α

วิธีคำนวณจากตารางเครื่องหมายของความแตกต่าง $(x_i - y_i)$ มีค่า 0 จำนวน 1 คู่ มีเครื่องหมายบวกจำนวน 2 คู่ และมีเครื่องหมายลบจำนวน 9 คู่ เราจะตัดค่า 0 ออกไปไม่นำมาคำนวณ และลด n ลงเหลือ 11 คู่ และ $r = 2$

เราต้องการทราบความน่าจะเป็นที่จะเกิดเครื่องหมายบวกไม่มากกว่า 2 ตัว จากตัวอย่าง 11 คู่ เมื่อ H_0 เป็นจริง คำนวณค่าความน่าจะเป็นสะสมจากการแจกแจงทวินามคือ

$$P(R \le r | n , p) = \sum_{R=0}^{r} \binom{n}{r} p^r q^{n-r}$$

หรือเปิดค่าจากตาราง Cumulative Binomial Probability Distribution (Daniel , 1995) ที่ n = 11 , p = .5 และ r = จำนวนเครื่องหมายบวก = 2 ได้ P = .0327 ซึ่งน้อยกว่าระดับนัยสำคัญที่กำหนด ($\alpha = .05$) จึงสรุปได้ว่าปฏิเสธ H_0 นั่นคือ มัธยฐานของความแตกต่างของคะแนนความสะอาดในช่องปากของกลุ่มตัวอย่าง 2 กลุ่มเป็นลบคือ การสอนการแปรงฟันและรักษาความสะอาดในช่องปากให้ประชาชนได้ผลดี

เราสามารถใช้โปรแกรม SPSS ช่วยในการคำนวณได้ โดยใช้คำสั่ง **2 Related Samples** ใน **Nonparametric test**

5.2 การใช้คำสั่ง 2 Related Samples... สำหรับ Sign test ที่มีข้อมูลแบบจับคู่

จากตัวอย่างการประเมินผลการสอนประชาชนเกี่ยวกับการแปรงฟันที่มีตัวอย่าง 12 คู่ กำหนดให้ตัวแปร instruct แทนกลุ่มที่ได้รับการสอน และตัวแปร notins แทนกลุ่มที่ไม่ได้รับการสอน บันทึกข้อมูลในแฟ้มข้อมูล Sign2.sav มีรูปแบบของข้อมูลในแฟ้มดังตารางที่ 13.19

ตารางที่ 13.19 รูปแบบการบันทึกข้อมูลคะแนนความสะอาดช่องปากของกลุ่มที่ได้รับการสอน และกลุ่มที่ไม่ได้รับการสอนวิธีแปรงฟัน

notins	instruct
2.0	1.5
2.0	2.0
4.0	3.5
2.5	3.0
4.0	3.5

3.0	2.5
3.5	2.0
3.0	1.5
2.5	1.5
2.5	2.0
2.5	3.0
2.5	2.0

การใช้โปรแกรม SPSS ช่วยในการคำนวณสำหรับการทดสอบ Sign test ที่มีข้อมูลแบบจับคู่ มีขั้นตอนดังนี้

1. ไปที่เมนูบาร์ คลิกที่ Analyze , Nonparmetric Test , 2 Related Samples... จะได้หน้าต่าง Two-Related-Samples Tests

2. ในหน้าต่าง Two-Related-Samples Tests

ในช่องซ้ายมือ คลิกที่ตัวแปร notins และ instruct ให้ย้ายไปอยู่ในช่อง Test Pair(s) List :

ในกรอบ Test Type

เลือก ☐ Sign เพื่อทำการทดสอบ Sign test

แล้วคลิกที่ปุ่ม OK จะได้ผลลัพธ์ดังภาพที่ 13.8

NPar Tests

Sign Test

Frequencies

		N
instruct - notins	Negative Differences [a]	9
	Positive Differences [b]	2
	Ties[c]	1
	Total	12

a. instruct < notins

b. instruct > notins

c. instruct = notins

Test Statistics[b]

	instruct - notins
Exact Sig. (2-tailed)	.065[a]

a. Binomial distribution used.

b. Sign Test

ภาพที่ 13.8

จากภาพผลลัพธ์ในตาราง Frequencies แสดงจำนวนตัวอย่างของเครื่องหมายลบและบวกในการทดสอบ Sign test คือ มีเครื่องหมายลบที่เกิดจากคะแนนของกลุ่ม instruct น้อยกว่าคะแนนของกลุ่ม notins จำนวน 9 ตัว และมีเครื่องหมายบวกที่เกิดจากคะแนนของกลุ่ม instruct มากกว่าคะแนนของกลุ่ม notins จำนวน 2 ตัว และ Ties คือ คะแนนของทั้ง 2 กลุ่มตัวอย่างเท่ากันมีจำนวน 1 คู่ รวมทั้งหมด 12 คู่

ผลการทดสอบ Sign test อยู่ในตาราง Test Statistics ได้ค่า Exact Sig. (2-tailed) เท่ากับ .065 แต่ตัวอย่างนี้เป็นการทดสอบแบบทางเดียว จึงได้ค่า Sig. เท่ากับ .065/2 เท่ากับ .0325 ซึ่งน้อยกว่าระดับนัยสำคัญที่กำหนด (α = .05) จึงสรุปได้ว่าปฏิเสธ H_0 นั่นคือ มัธยฐานของความแตกต่างของคะแนนเป็นลบ หรือ $P(+) < P(-)$ หรือแสดงว่าผลการสอนการแปรงฟันให้ประชาชนได้รับผลดี

5.3 การทดสอบวิลคอกซัน (Wilcoxon signed – rank test)

ใช้ในการเปรียบเทียบความแตกต่างของข้อมูลที่เป็นอันดับจากกลุ่มตัวอย่าง 2 กลุ่มที่เกี่ยวข้องกัน เป็นการทดสอบเครื่องหมายและอันดับ ใช้แทนการทดสอบ t-test ที่มีกลุ่มตัวอย่าง 2 กลุ่มที่เกี่ยวข้องกันในสถิติอิงพารามิเตอร์

การทดสอบ Wilcoxon signed-rank test เป็นการทดสอบที่พัฒนามาจากการทดสอบเครื่องหมาย (Signed test) แต่แตกต่างจากการทดสอบเครื่องหมายคือ ในกรณีที่สุ่มข้อมูลมาเป็น คู่ ๆ (paired data) ได้นำขนาดความแตกต่างของข้อมูลแต่ละคู่มาพิจารณาพร้อมกับเครื่องหมายด้วย ค่าความแตกต่างที่เท่ากันจะถูกจัดอันดับโดยใช้ค่าเฉลี่ยของอันดับที่เป็นของค่าที่เท่ากันเหล่านั้น ข้อมูลคู่ใดเท่ากันจะไม่ถูกนำเข้ามาพิจารณา และมีผลให้ขนาดตัวอย่าง n ลดลง และเนื่องจากการจัดอันดับความแตกต่างนี้ไม่คำนึงถึงเครื่องหมาย ทำ

ให้ค่าแตกต่างที่มีเครื่องหมายแตกต่างกันจัดอยู่ในอันดับเดียวกัน หลังจากการจัดอันดับค่าแตกต่างแล้วจึงใช้เครื่องหมายที่มีอยู่เดิมกำกับอันดับที่ได้ ขั้นต่อมาหาผลรวมของอันดับที่มีเครื่องหมายบวก และผลรวมของอันดับที่มีเครื่องหมายลบ แล้วเลือกค่าของผลรวมที่น้อยกว่าโดยไม่คำนึงถึงเครื่องหมายเป็นค่าสถิติทดสอบ

สมมติฐานที่ต้องการทดสอบคือ H_0 : ประชากรทั้งสองกลุ่มเหมือนกันหรือมีการแจกแจงแบบเดียวกัน คู่กับ H_1 : ประชากรทั้งสองกลุ่มแตกต่างกัน ถ้า H_0 เป็นจริง ผลรวมของอันดับที่มีเครื่องหมายบวก และผลรวมของอันดับที่มีเครื่องหมายลบควรจะเกือบเท่ากัน ถ้าผลรวมทั้งสองแตกต่างกันมากก็พอจะสรุปได้ว่าประชากรทั้งสองกลุ่มมีการแจกแจงที่ไม่เหมือนกัน นั่นคือ เราจะปฏิเสธ H_0

ถ้าเราให้ผลรวมของอันดับที่มีเครื่องหมายบวกและลบแทนด้วย $\sum R(+)$ และ $\sum R(-)$ ตามลำดับ

สมมติฐานที่ต้องการทดสอบคือ H_0 : ความแตกต่างของข้อมูลและคู่มีการแจกแจงในลักษณะสมมาตรรอบ ๆ ค่ากลางที่มีค่าเท่ากับศูนย์ ทำให้ผลรวมของอันดับที่เป็นเครื่องหมายบวกเท่ากับผลรวมของอันดับที่มีเครื่องหมายลบ เขียนสมมติฐานข้างต้นใหม่ได้คือ H_0 : $\sum R(+) = \sum R(-)$ คู่กับ H_1 : $\sum R(+) \neq \sum R(-)$

ตัวอย่างการศึกษาของ Geffen, Bradshaw and Nettleton (1973) อ้างถึงใน P.Sprent and N.C. Smeeton (2001) ที่มีวัตถุประสงค์เพื่อศึกษาว่าการแสดงตัวเลขในพื้นที่ที่มองเห็นได้จากด้านขวา (RVF) หรือจากด้านซ้าย (LVF) ทำให้มองเห็นตัวเลขได้ถูกต้องและได้เร็วกว่ากันหรือไม่แตกต่างกัน ศึกษากับกลุ่มตัวอย่าง 12 คน ให้ดูตัวเลขที่แสดงออกมาอย่างสุ่มจากด้านขวาหรือด้านซ้าย แล้ววัดเวลาเฉลี่ยของการตอบในแต่ละด้านของตัวอย่างแต่ละคน เวลาในการตอบในระหว่างตัวอย่างแต่ละคนมีความแตกต่างกันมากกว่าในระหว่างด้านของพื้นที่ที่มองเห็น ข้อมูลและความแตกต่างของ LVF – RVF ของตัวอย่างแต่ละคนแสดงในตารางที่ 13.20

ตารางที่ 13.20 เวลาเฉลี่ย (ms) ของการตอบเกี่ยวกับตัวเลขที่แสดงในพื้นที่ทางด้านซ้ายและขวา

ตัวอย่าง	1	2	3	4	5	6	7	8	9	10	11	12
LVF (1)	564	521	495	564	560	481	545	478	580	484	539	467
RVF (2)	557	505	465	562	544	448	531	458	560	485	520	445
(1) – (2)	7	16	30	2	16	33	14	20	20	-1	19	22

กำหนดให้ x_i แทนเวลาของการตอบของตัวอย่างที่ i ที่ดูตัวเลขที่แสดงออกมาจากทางด้านขวา (RVF) และ y_i แทนเวลาของการตอบของตัวอย่างที่ i ที่ดูตัวเลขที่แสดงออกมาจากทางด้านซ้าย (LVF) และหาความแตกต่าง $d_i = y_i - x_i$; $i = 1, 2, \ldots, n$ ความแตกต่าง d_i เป็นอิสระกันเนื่องจากแต่ละตัวอ้างถึงความแตกต่างตัวหนึ่งภายใต้สมมติฐานว่า H_0 : มัธยฐานของความแตกต่าง d_i เท่ากับ ศูนย์ นั่นคือ d_i มีค่าเป็นบวกหรือลบเท่า ๆ กัน และสามารถใช้การทดสอบ Sign test ถ้าเราสมมติว่าการแจกแจงของ d_i มีความสมมาตร ภายใต้ H_0 เราอาจใช้การทดสอบ Wilcoxon signed_rank test ในตัวอย่าง 2 กลุ่มนี้มีรูปแบบของการตอบหลายรูปแบบ ซึ่งสามารถทำให้ความแตกต่าง d_i มีการแจกแจงแบบสมมาตรภายใต้ H_0 โดยเฉพาะถ้าเราสมมติให้เวลาการตอบเป็นเวลาเดียวกันและเป็นอิสระกันกระจายอยู่ใน 2 กลุ่ม สำหรับหน่วยตัวอย่างใด ๆ ความแตกต่างของแต่ละหน่วยตัวอย่างจะสมมาตรอยู่รอบศูนย์ เนื่องจาก ถ้า X และ Y มีการแจกแจงแบบเดียวกันและเป็นอิสระกันแล้ว $X - Y$ และ $Y - X$ แต่ละตัวจะมีการแจกแจงแบบเดียวกันและจะมีความสมมาตรอยู่รอบศูนย์ ทำให้ความแตกต่างมีการแจกแจงแบบสมมาตรด้วย ถ้าเวลาของการตอบที่ตัวเลขแสดงออกมาจากทางด้านซ้ายและขวาของหน่วยตัวอย่างใด ๆ มีการแจกแจงอย่างสมมาตรที่แตกต่างกัน โดยที่มีค่าเฉลี่ยเท่ากัน ดังนั้นการใช้การทดสอบ Wilcoxon test มีเหตุผลสมควรเมื่อเราสมมติว่าเวลาของการตอบที่ตัวเลขแสดงออกมาจากทางด้านซ้ายและขวา มีการแจกแจงแบบเดียวกันสำหรับแต่ละหน่วยตัวอย่างใด ๆ ภายใต้ H_0 เป็นจริง สมมติฐานแย้งคือ H_1 : มัธยฐานของความแตกต่าง d_i ไม่เท่ากับศูนย์ หรือมีความแตกต่างของมัธยฐาน

วิธีการทดสอบคือ

(1) หาความแตกต่าง d_i ของ LVF–RVF ของตัวอย่างแต่ละคนดังแสดงในตารางที่ 13.21

ตารางที่ 13.21 ค่าเฉลี่ยของเวลา (ms) ในการตอบ เกี่ยวกับตัวเลขที่แสดงในพื้นที่ทางด้านซ้าย

และขวา และค่า d_i

คนที่	1	2	3	4	5	6	7	8	9	10	11	12
LVF (1)	564	521	495	564	560	481	545	478	580	484	539	467
RVF (2)	557	505	465	562	544	448	531	458	560	485	520	445
d_i = (1)–(2)	7	16	30	2	16	33	14	20	20	-1	19	22

(2) เรียงลำดับความแตกต่าง d_i ที่ได้จากตารางที่ 13.20 คือ

-1 2 7 14 16 16 19 20 20 22 30 33

(3) ให้แต่ละอันดับมีเครื่องหมายของ d_i คือ

-1 2 3 4 5.5 5.5 7 8.5 8.5 10 11 12

(4) คำนวณค่าสถิติทดสอบวิลคอกซัน คือ S_+ หรือ S_- เมื่อ S_+ คือ ผลบวกของอันดับที่มีเครื่องหมายบวกและ S_- คือ ผลบวกของอันดับที่มีเครื่องหมายลบ ตัวที่มีค่าเล็กกว่าเป็นสถิติทดสอบ

เนื่องจากค่าสถิติทดสอบวิลคอกซันคือ S_- เท่ากับ 1 เทียบกับค่าวิกฤติที่เปิดจากตาราง Probability Levels for the Wilcoxon signed – rank test ที่ $n = 12$, $\alpha/2 = .026$ ซึ่งเป็นค่าที่ใกล้ที่สุดกับค่า .025 ในตารางได้ค่า $T = 14$ สรุปได้ว่าค่า S_- น้อยกว่าค่าวิกฤติ T ดังนั้นจึงตัดสินใจปฏิเสธ H_0 นั่นคือ มัธยฐานของเวลาในการตอบเกี่ยวกับตัวเลขที่แสดงในพื้นที่ด้านขวาและ ด้านซ้ายแตกต่างกัน และเมื่อเทียบค่าสถิติทดสอบ S_- กับค่าวิกฤติ T ที่เปิดจากตาราง ที่ $n = 12, \alpha = .046$ ซึ่งเป็นค่าที่ใกล้ที่สุดกับค่า .05 ในตารางได้ค่า $T = 17$ สรุปได้ว่าเวลาในการตอบเกี่ยวกับตัวเลขที่แสดงในพื้นที่ด้านขวาเร็วกว่าพื้นที่ด้านซ้าย

เราสามารถใช้โปรแกรม SPSS ช่วยในการคำนวณได้ โดยใช้คำสั่ง **2 Related Samples** ใน **Nonparametric test**

5.4 การใช้คำสั่ง 2 Related Samples ... สำหรับ Wilcoxon signed – rank test ที่มีข้อมูล แบบจับคู่

จากตัวอย่างการทดสอบความเร็วของการตอบเกี่ยวกับตัวเลขที่แสดงอย่างสุ่มในพื้นที่ด้านขวาและด้านซ้าย กำหนดให้ตัวแปร RVF และ LVF แทนด้านที่แสดงตัวเลขทางขวาและซ้ายตามลำดับ บันทึกข้อมูลในแฟ้มข้อมูล Wilcox2.sav มีรูปแบบของข้อมูลในแฟ้มดังตารางที่ 13.22

ตารางที่ 13.22 รูปแบบการบันทึกข้อมูลเวลาในการตอบเกี่ยวกับตัวเลขที่แสดงในพื้นที่ด้านขวา และด้านซ้าย

RVF	LVF	RVF	LVF
557	564	531	545
505	521	458	478
465	495	560	580
562	564	485	484
544	560	520	539
448	481	445	467

การใช้โปรแกรม SPSS ช่วยในการคำนวณสำหรับ Wilcoxcon signed-rank test ที่มีข้อมูลแบบจับคู่ มีขั้นตอนดังนี้

1. ไปที่เมนูบาร์ คลิกที่ Analyze, Nonparametric Test, 2 Related Samples... จะได้หน้าต่าง Two – Related – Samples Tests
2. ในหน้าต่าง Two – Related Samples Tests

   ในช่องซ้ายมือ คลิกที่ตัวแปร RVF และ LVF ให้ย้ายไปอยู่ในช่อง Test Pair(s) List:

   ในกรอบ Test Type

   เลือก ☐ Wilcoxon เพื่อทำการทดสอบ Wilcoxon signed – rank test

แล้วคลิกที่ปุ่ม OK จะได้ผลลัพธ์ดังภาพที่ 13.9

NPar Tests
Wilcoxon Signed Ranks Test

Ranks

		N	Mean Rank	Sum of Ranks
LVF - RVF	Negative Ranks	1[a]	1.00	1.00
	Positive Ranks	11[b]	7.00	77.00
	Ties	0[c]		
	Total	12		

a. LVF < RVF

b. LVF > RVF

c. LVF = RVF

Test Statistics[b]

	LVF - RVF
Z	-2.983[a]
Asymp. Sig. (2-tailed)	.003

a. Based on negative ranks.

b. Wilcoxon Signed Ranks Test

ภาพที่ 13.9

จากภาพผลลัพธ์ในตาราง Ranks แสดงจำนวนตัวอย่างของอันดับที่เป็นลบมี 1 ตัว และอันดับที่เป็นบวกมี 11 ตัว รวมทั้งหมด 12 ตัวอย่าง ผลบวกของอันดับที่เป็นลบเท่ากับ 1 และผลบวกของอันดับที่เป็นบวกเท่ากับ 77

ผลการทดสอบ Wilcoxon signed – rank test อยู่ในตาราง Test Statistics ได้ค่าสถิติ Z ของอันดับที่เป็นลบเท่ากับ -2.983 และค่า Asymp. Sig. (2-tailed) เท่ากับ .003 ซึ่งมีค่าน้อยกว่าระดับนัยสำคัญที่กำหนด (α = .05) ดังนั้นจึงปฏิเสธ H_0 นั่นคือ มัธยฐานของเวลาในการตอบเกี่ยวกับตัวเลขที่แสดงในพื้นที่ด้านขวาและด้านซ้ายแตกต่างกัน

6. การทดสอบเกี่ยวกับค่ากลางของประชากรหลายกลุ่มที่ไม่เป็นอิสระกัน

6.1 การทดสอบฟรีดแมน (Friedman Test)

Friedman (1937) ได้ขยายวิธีการทดสอบ sign test มาใช้ในกรณีที่มีตัวอย่างหลายกลุ่มที่เกี่ยวข้องกัน สำหรับในการทดสอบ sign test มีวิธีการให้คะแนนเป็นแบบเครื่องหมาย - หรือ + สามารถเทียบได้เป็นอันดับที่ 1 และ 2 สำหรับในการทดสอบ Friedman test มีวิธีการให้คะแนนแต่ละหน่วยตัวอย่างเป็นอันดับที่ คล้ายกับการทดสอบ F-test ในการวิเคราะห์ความ แปรปรวนที่ออกแบบการทดลองแบบสุ่มสมบูรณ์ภายในบล็อกในการวิเคราะห์แบบอิงพารามิเตอร์ การทดสอบนี้เรียกอีกอย่างได้คือ การทดสอบ Friedman Two-Way Analysis of Variance by Ranks

ในการวิเคราะห์ความแปรปรวนสำหรับค่าสังเกตที่มีการแจกแจงแบบต่อเนื่อง เช่น การแยกความแตกต่างระหว่างผู้ป่วย (หรือบล็อก) ซึ่งเป็นแหล่งของความแปรปรวนออกก่อนทำการเปรียบเทียบทรีทเมนต์ ในการทดสอบ Friedman’s test ความแตกต่างระหว่างผู้ป่วย (หรือบล็อก) ถูกแยกออกไปแล้วโดยการแทนที่ค่าสังเกตด้วยอันดับ เมื่อเราทำอย่างนี้ภายในแต่ละบล็อกแยกกัน และถ้าข้อมูลเป็นอันดับอยู่แล้วในแต่ละบล็อกก็ใช้การทดสอบ Friedman’s test ดังนั้นการทดสอบนี้เป็นการตรวจสอบความคงเส้นคงวาของอันดับมากกว่าเป็นการทดสอบค่ากลาง

สมมติว่าในการทดลองหนึ่งมี t ทรีทเมนต์ และมี b บล็อก สมมติฐานที่ต้องการทดสอบคือ H_0 : การแจกแจงของประชากรของแต่ละทรีทเมนต์เหมือนกัน ดังนั้น ทำให้มัธยฐานของประชากรต่าง ๆ เท่ากันด้วย ถ้าสมมติฐานสูญเป็นจริง ที่ทรีทเมนต์หนึ่งจะวัดได้อันดับจาก 1 ถึง t ปน ๆ กัน แต่ถ้ามัธยฐานของประชากรต่าง ๆ แตกต่างกันแล้วที่อย่างน้อยที่สุด 1 ทรีทเมนต์จะมีอันดับส่วนใหญ่สูงหรือส่วนใหญ่ต่ำ

ตัวอย่างเช่นในการศึกษาหนึ่งมีผู้ป่วย b คน ผู้ป่วยแต่ละคนถูกวัดอัตราการเต้นของชีพจรที่ t จุดเวลา ออกแบบการทดลองแบบสุ่มสมบูรณ์ภายในบล็อกที่มี b บล็อก ให้ t ทรีทเมนต์กับหน่วยทดลอง 1 หน่วย ในแต่ละบล็อกประกอบด้วย t ทรีทเมนต์ ถ้า x_{ij} แทนค่าสังเกตที่เป็นอัตราการเต้นของชีพจรที่เวลา i ของผู้ป่วย j เมื่อ $i = 1, 2, \ldots, t$ และ $j = 1, 2, \ldots, b$ วิธีการทดสอบ Friedman’s test คือ แทน x_{ij} ของผู้ป่วยแต่ละคนด้วยอันดับที่ 1 ถึง t การให้อันดับนี้ทำแยกกันในผู้ป่วยแต่ละคน (บล็อก) สมมติว่าไม่มีอันดับซ้ำกันเลย ผลรวมของอันดับของจุดเวลา i แทนด้วย s_i เมื่อ $i = 1, 2, \ldots, t$ สถิติ Friedman คือ

$$T = \frac{12\sum_i s_i^2}{bt(t+1)} - 3b(t+1)$$

ถ้า b และ t ไม่เล็กเกินไป T จะมีการแจกแจงโดยประมาณแบบไคสแคว์ที่มีจำนวนชั้นอิสระ t – 1 ตารางที่ใช้สำหรับดูค่าวิกฤติของ T (ดูได้จาก Siegel & Castellan , 1988 , หน้า 353 อ้างถึงใน วิสาข์ เกษประทุม , 2545)

ตัวอย่างเช่นต้องการศึกษาเปรียบเทียบอัตราการเต้นของชีพจรที่ 3 จุดเวลาคือ ก่อนการออกกำลังกาย ทันทีหลังจากออกกำลังกาย และ 5 นาทีหลังการออกกำลังกาย กลุ่มตัวอย่างคือนักเรียน 7 คน ดำเนินการทดลองคือทำการวัดอัตราการเต้นของชีพจรของนักเรียนแต่ละคนที่ 3 จุดเวลาได้ข้อมูลดังตารางที่ 13.23 ทำการวิเคราะห์โดยใช้ค่าสถิติ Friedman ทดสอบความแตกต่างระหว่างอัตราการเต้นของชีพจรใน 3 จุดเวลา

วิธีการทดสอบคือ นักเรียนแต่ละคนถูกวัดอัตราการเต้นของชีพจรเป็นค่าสังเกตแต่ละตัว แล้วแทนค่าสังเกตแต่ละตัวด้วยอันดับ แล้วคำนวณค่าสถิติ Friedman ของอันดับ สมมติฐานที่ต้องการทดสอบคือ H_0 : การแจกแจงของประชากรของค่าอัตราการเต้นของชีพจร เหมือนกันทั้ง 3 จุดเวลา คู่กับ H_1 : อย่างน้อยที่สุดที่หนึ่งจุดเวลามีการแจกแจงของอัตราการเต้นของชีพจรแตกต่างไป สะท้อนให้เห็นจากค่ามัธยฐานที่เคลื่อนไป

วิธีคำนวณ ให้ข้อมูลอัตราการเต้นของชีพจรของนักเรียนแต่ละคนที่ 3 จุดเวลา แทนด้วยอันดับที่ 1 ถึง 3 ดังตารางที่ 13.24 หาค่าผลรวมของอันดับที่แต่ละจุดเวลา ยกกำลังสองผลรวมทางคอลัมน์ แล้วนำมาบวกกันคือ $10^2 + 21^2 + 11^2 = 662$ กลุ่มตัวอย่างคือนักเรียน 7 คน และทรีทเมนต์คือ 3 จุดเวลา แทนค่า $b = 7$, $t = 3$ คำนวณหาค่าสถิติ Friedman คือ

$$T = \frac{(12)(662)}{(7)(3)(3+1)} - (3)(7)(3+1) = 10.57$$

ซึ่งมีการแจกแจงโดยประมาณแบบไคสแควที่มีจำนวนชั้นอิสระ df เท่ากับ 3 - 1 = 2 เทียบกับค่าวิกฤติที่เปิดจากตารางที่ $k = 3$, $N = 7$, $\alpha \leq .05$ ได้ค่าวิกฤติเท่ากับ 7.14 สรุปได้ว่าค่า T มากกว่าค่าวิกฤติ ดังนั้นจึงตัดสินใจปฏิเสธ H_0 นั่นคือ มัธยฐานของอัตราการเต้นของชีพจรที่ 3 จุดเวลาแตกต่างกัน

ตารางที่ 13.23 อัตราการเต้นของชีพจรของนักเรียนในสถานการณ์ที่แตกต่างกัน

	เวลา

นักเรียน	ก่อนออกกำลังกาย	ทันทีหลังการออกกำลังกาย	5 นาทีหลังการออกกำลังกาย
A	72	120	76
B	96	120	95
C	88	132	104
D	92	120	96
E	74	101	84
F	76	96	72
G	82	112	76

ตารางที่ 13.24 อันดับของอัตราการเต้นของชีพจรของนักเรียนแต่ละคนในสถานการณ์ที่แตกต่างกัน

นักเรียน	เวลา		
	ก่อนออกกำลังกาย	ทันทีหลังการออกกำลังกาย	5 นาทีหลังการออกกำลังกาย
A	1	3	2
B	2	3	1
C	1	3	2
D	1	3	2
E	1	3	2
F	2	3	1
G	2	3	1
ผลรวม	10	21	11

เราสามารถใช้โปรแกรม SPSS ช่วยในการคำนวณได้โดยใช้คำสั่ง **K Related Samples** ใน **Nonparametric Test**

6.2 การใช้คำสั่ง K Related Samples…

จากตัวอย่างการเปรียบเทียบอัตราการเต้นของชีพจรใน 3 จุดเวลาคือ ก่อนการออกกำลังกาย ทันทีหลังการออกกำลังกาย และ 5 นาทีหลังการออกกำลังกาย กำหนดให้นักเรียนแทนด้วยตัวแปร student และจุดเวลาแทนด้วยตัวแปร time1 คือ ก่อนออกกำลังกาย , time2 คือ ทันทีหลังการออกกำลังกาย และ time3 คือ 5 นาทีหลังการออกกำลังกาย ข้อมูลคือ อัตราการเต้นของชีพจร บันทึกข้อมูลในแฟ้มข้อมูล Friedman1.sav มีรูปแบบของข้อมูลในแฟ้มดังตารางที่ 13.25

ตารางที่ 13.25 รูปแบบการบันทึกข้อมูลอัตราการเต้นของชีพจรใน 3 สถานการณ์

student	time1	time2	time3
1	72	120	76
2	96	120	95
3	88	132	104
4	92	120	96
5	74	101	84
6	76	96	72
7	82	112	76

การใช้โปรแกรม SPSS ช่วยในการคำนวณมีขั้นตอนดังนี้

1. ไปที่เมนูบาร์ คลิกที่ Analyze, Nonparametric Test, K Related Samples... จะได้หน้าต่าง Tests for Several Related Samples

2. ในหน้าต่าง Tests for Several Related Samples

ในช่องซ้ายมือคลิกที่ตัวแปร time1 , time2 , time3 ให้ย้ายเข้าไปอยู่ในช่องของ Test Variables :

ในกรอบ Test Type

เลือก ☐ Friedman เพื่อทำการทดสอบ Friedman test

แล้วคลิกที่ปุ่ม OK จะได้ผลลัพธ์ดังภาพที่ 13.10

NPar Tests
Friedman Test

Ranks

	Mean Rank
time1	1.43
time2	3.00
time3	1.57

Test Statistics[a]

N	7
Chi-Square	10.571
df	2
Asymp. Sig.	.005

a. Friedman Test

ภาพที่ 13.10

จากภาพผลลัพธ์ในตาราง Ranks แสดง Mean Rank ของตัวแปร time1 , time2 , และ time3

ผลการทดสอบ Friedman Test อยู่ในตาราง Test Statistics ได้ค่าสถิติทดสอบ Chi – Square เท่ากับ 10.571 ที่ df เท่ากับ 2 และค่า Asymp. Sig. เท่ากับ .005 ซึ่งมีค่าน้อยกว่าระดับนัยสำคัญที่กำหนด (α = .05) ดังนั้นจึงปฏิเสธ H_0 นั่นคือ มีความแตกต่างของอัตราการเต้นของชีพจรในระหว่างเวลาทั้ง 3 เวลานั้นอย่างน้อย 2 เวลา