สถิติไร้พารามิเตอร์

Fete, Time, Occasion, Chance and Change?
. . . ..to these all things are subject

Percy Bysshe Shelly

การอนุมานเชิงสถิติที่เกี่ยวกับประชากรซึ่งได้กล่าวมาเกือบทั้งหมดตั้งอยู่บนข้อสมมติที่ว่าตัวอย่างสุ่มนั้นเลือกมาจากประชากรที่เราทราบว่ามีการแจกแจงเป็นอย่างไร เช่นแจกแจงเป็นแบบปกติทวินามเป็นต้น ดังนั้นฟังก์ชันของการแจกแจงจึงขึ้นอยู่กับพารามิเตอร์ของประชากรนั้น เราจึงเรียกการอนุมานเกี่ยวกับประชากรนั้นว่าสถิติในเชิงพารามิเตอร์ (Parametric Statistics) ซึ่งเป็นการกะประมาณหรือทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ของประชากรที่ไม่ทราบค่านั้นด้วยตัวสถิติ (Statistic) จากตัวอย่างสุ่มที่เลือกมาจากประชากร

ในการใช้สถิติเชิงพารามิเตอร์เพื่อกะประมาณและทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์หรือค่าคงที่ของประชากรนั้น ถ้าจะให้มีประสิทธิภาพแล้วต้องสอดคล้องกับเงื่อนไข 2 ประการคือ

(1) การแจกแจงของประชากรไม่ว่าจะเป็นแบบใดจะต้องมีลักษณะปกติ (Normality of Populations)

(2) ความแปรปรวนจะต้องคงที่ (Stability of Variance) ในแบบทดสอบ t และ F ที่กล่าวมานั้นเราได้ยึดข้อสมมติเกี่ยวกับประชากรทั้ง 2 ประการ ในการอนุมานเกี่ยวกับวิเคราะห์ถดถอยและสหสัมพันธ์เราก็ได้อาศัยข้อสมมติทั้งสองเช่นกัน

ในแบบทดสอบเกี่ยวกับการปรับที่ดี (Goodness -of-Fit) หรือแบบทดสอบเกี่ยวกับการแจกแจงของประชากรและแบบทดสอบความเป็นอิสระเชิงสถิตินั้นเราก็มิได้กำหนดแน่ชัดลงไปว่าประชากรที่เราสุ่มตัวอย่างมานั้นมีการแจกแจงแบบไหน กรณีเช่นนี้อาจถือได้ว่าเรามิได้กำหนดข้อสมมติเกี่ยวกับการแจกแจงของประชากรเลย

ในสภาพการณ์บางอย่างเราไม่สามารถจะระบุไปอย่างชัดแจ้งว่าประชากรมีการแจกแจงแบบใด ทั้งยังไม่อาจตั้งข้อสมมติได้ว่าการแจกแจงของประชากรเป็นไปตามเงื่อนไข 2 ประการนั้น กรณีเช่นนี้เราจึงใช้สถิติในเชิงพารามิเตอร์มิได้ นั่นคือเราจะใช้สถิติไร้พารามิเตอร์นั่นเอง อย่างไรก็ตามสถิติแบบนี้ยังต้องอาศัยข้อสมมติที่ว่าประชากรมีการแจกแจงแบบต่อเนื่อง

ในการทดสอบสมมติฐานเกี่ยวกับประชากรโดยไม่ต้องอาศัยข้อสมมติเกี่ยวกับการแจกแจงของประชากรไว้อย่างเคร่งครัดและไม่จำเป็นต้องทำการทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์

ของประชากรนั้น ๆ วิธีการทางสถิติหรือแบบทดสอบที่ใช้สรุปผลจากค่าที่ได้จากตัวอย่างโดยตรงนั้นได้ชื่อว่าแบบทดสอบที่เป็นอิสระจากรูปการแจกแจง (Distribution-Free Statistical Test) หรือแบบทดสอบที่ไร้พารามิเตอร์ แต่ชื่อเรียกทั้งสองนี้แตกต่างกันเล็กน้อยดังนี้

แบบทดสอบไร้พารามิเตอร์เป็นแบบทดสอบที่ไม่ได้ระบุเงื่อนไขเกี่ยวกับพารามิเตอร์ของประชากรที่ตัวอย่างสุ่มได้เลือกมาหรือเป็นแบบทดสอบสมมติฐานที่ไม่เป็นคำกล่าวเกี่ยวกับพารามิเตอร์แบบทดสอบนี้ใช้เมื่อเงื่อนไขหรือข้อสมมติเกี่ยวกับแบบทดสอบมาตรฐาน (Standard Test) ขาดไป และส่วนมากใช้เป็นแบบทดสอบคู่ขนานกับแบบทดสอบเกี่ยวกับพารามิเตอร์ที่เรากล่าวมาแล้ว

แบบทดสอบที่เป็นอิสระจากรูปการแจกแจงเป็นแบบทดสอบสมมติฐานที่ไม่มีข้อสมมติเกี่ยวกับธรรมชาติ รูปร่าง หรือรูปแบบ (Nature, Shape or Form) ของประชากรที่เลือกสุ่มตัวอย่างมา แต่อย่างไรก็ตามแบบทดสอบนี้ไม่ได้เป็นอิสระในเชิงข้อสมมติเกี่ยวกับรูปแบบของการแจกแจงตัวอย่าง (Sampling Distribution) นั่นคือยังขึ้นกับการแจกแจงตัวอย่างของตัวสถิติเหมือนกับแบบทดสอบที่เกี่ยวกับพารามิเตอร์

เราจะเห็นได้ว่าทั้งเทอมนี้ไม่เหมือนกันนักแต่เราก็ใช้แทนกันได้ เพราะวิธีการในการทดสอบจะเหมือนกันตามปกติเรามักใช้คำว่าแบบทดสอบไร้พารามิเตอร์นี้ ให้รวมถึงแบบทดสอบทั้งอย่างนั้นในระยะหลังมีผู้คิดค้นแบบทดสอบในสถิติพารามิเตอร์กันมากและเป็นที่ยอมรับกันทั่วไป

มีข้อที่น่าสังเกตไว้คือในกรณีที่สามารถทำการทดสอบได้ทั้งสองแบบคือแบบที่เกี่ยวกับพารามิเตอร์และไร้พารามิเตอร์นั้น เราก็ไม่ควรนำแบบทดสอบไร้พารามิเตอร์มาใช้ทั้งนี้ก็เพราะแบบทดสอบไร้พารามิเตอร์มีอำนาจทดสอบหรือความถูกต้องเชื่อถือน้อยกว่าแบบทดสอบที่เกี่ยวกับพารามิเตอร์มาก อย่างไรก็ตามแบบทดสอบไร้พารามิเตอร์จะใช้ได้ผลก็ต่อเมื่อได้ใช้ตัวอย่างสุ่มขนาดโตพอควร

การทดสอบโดยใช้แบบทดสอบไร้พารามิเตอร์นี้ก็มีข้อดีหลายประการคือ นำไปประยุกต์ได้ง่าย การคำนวณก็ไม่ยุ่งยาก ง่ายต่อการอธิบายและทำความเข้าใจ การพัฒนาทฤษฎีก็ใช้คณิตศาสตร์ไม่ยากนัก

8.1 การทดสอบโดยอาศัยเครื่องหมาย (Sign Test)

การทดสอบประเภทนี้จะใช้กับค่าสังเกตตัวอย่างที่มี 2 ลักษณะ หรือใช้กับข้อมูลตัวอย่างที่สามารถแปลงให้เป็น 2 ลักษณะได้ และลักษณะทั้งสองจะใช้เครื่องหมาย + และ - แทน นั่นคือการทดสอบประเภทนี้จะถือว่าตัวอย่างสุ่มนั้นได้มาจากประชากรชนิดสองค่า แบบทดสอบที่ใช้เครื่องหมายมีที่น่าสนใจดังนี้

8.1.1 แบบทดสอบเครื่องหมายชนิดตัวอย่างเดียว (One-Sample Sign Test)

แบบทดสอบนี้ใช้ทดสอบมัธยฐานหรือค่าเฉลี่ยที่ระบุไว้ของประชากรชนิดสมมาตร นั่นคือสมมติฐานหลักจะเป็น

$$H_0 : \mu = \mu_0$$

ในเมื่อ μ_0 เป็นค่าที่ระบุไว้ของมัธยฐาน

สำหรับตัวอย่างขนาด n ถ้าค่าสังเกตนั้นเป็นเครื่องหมาย + หรือ - เมื่อให้ x เป็นจำนวนเครื่องหมาย + แล้ว x จะเป็นตัวสถิติทดสอบซึ่งมีการแจกแจงทวินาม $(n, \pi = 1/2)$ สำหรับ H_0 เป็นจริง

เมื่อ ตัวอย่างขนาดโตเราก็จะได้ตัวสถิติทดสอบ Z ซึ่งมีการแจกแจงปกติมาตร ฐานดังนี้

$$Z = (x - n/2) / \sqrt{n/4}$$

ตัวอย่าง สมมติว่าอายุเฉลี่ย (Median life) ของหลอดไฟที่ผลิตโดยกระบวนการแบบเก่าเป็น 1000 ชั่วโมง บริษัทผู้ผลิตจะเปลี่ยนกระบวนการผลิตใหม่ ถ้าพบว่าระดับนัยสำคัญไม่มากกว่า 0.05 ที่กระบวนการผลิตใหม่ จะให้อายุเฉลี่ยน้อยกว่ากระบวนการแบบเก่า

สุ่มตัวอย่างหลอดไฟที่ผลิตด้วยกระบวนการใหม่มา 100 หลอด และทดสอบอายุ ปรากฏว่ามี 60 หลอด มีอายุน้อยว่า 1000 ชั่วโมง จะสรุปผลอย่างไร

ก. $H_o : \mu = 1000$, $H_a : \mu < 1000$

ข. $\alpha = .05$, $n = 100$

ค. ตัวสถิติทดสอบ X ที่เป็นเครื่องหมาย + และ X มีการแจกแจงทวินาม ที่มี $n = 100$, $\pi = .5$ ดังนี้

$$f(x) = \binom{100}{x} (.5)^x (.5)^{100-x}, \quad x = 0, 1, 2, \ldots, 100$$

ปฏิเสธ Ho ถ้าจำนวนเครื่องหมาย + ที่สังเกตได้น้อยเกินไป หรือถ้า

$$p = P(X \leq x) = \sum_{S=0}^{x} \binom{n_o}{s} P^s Q^{n-s} < \alpha$$

ง. $x = 40$, $$p \doteq \sum_{s=0}^{40} \binom{100}{s} (.5)^s (.5)^{100-s} = 0.0176$$

จ. ปฏิเสธ Ho เพราะ $p < .05$ แสดงว่าบริษัทจะต้องผลิตหลอดไฟโดยกระบวนการแบบเก่าต่อไป

8.1.2 แบบทดสอบคอกซ์และสจวร์ทเกี่ยวกับแนวโน้ม (Cox and Stuart Test for Trend)

แบบทดสอบนี้พัฒนาจากแบบทดสอบเครื่องหมายเพื่อใช้ทดสอบแนวโน้มของอนุกรมค่าสังเกต นั่นคือสมมติฐานหลักจะเป็น

Ho ; อนุกรมค่าสังเกตไม่มีแนวโน้ม

อนุกรมค่าสังเกตของตัวอย่างขนาด n นั้นเมื่อจับคู่ค่าสังเกตครึ่งแรกกับครึ่งหลังเป็น (X_1, X_{1+c}), (X_2, X_{2+c}),..., (X_{n-c}, X_n) ในเมื่อ c เท่ากับ n/2 ถ้า n เป็นเลขคู่ และเท่ากับ (n+ 1)/2 ถ้า n เป็นเลขคี่แล้วแทนแต่ละคู่ด้วยเครื่องหมาย + เมื่อ $X_i < X_{i+c}$ หรือ - เมื่อ $X_i > X_{i+c}$ และตัดคู่ที่เท่ากันทิ้งไป

จำนวนเครื่องหมาย + จะเป็นตัวสถิติทดสอบ ซึ่งจะให้เป็น T แล้วการแจกแจงของ T จะเป็นทวินามที่บารามิเตอร์ n และ π = 0.50 กฎตัดสินใจจะเหมือนกับแบบทดสอบเครื่องหมายที่กล่าวมาแล้วนั่นเอง

ตัวอย่าง ฝ่ายผลิตสินค้าได้บันทึกต้นทุนเฉลี่ยต่อหน่วยของสินค้าชนิดหนึ่งใน 19 เดือนได้ข้อมูลซึ่งเป็นต้นทุนดังนี้

	45.25	45.83	41.77	36.26	45.37	52.25	35.37	57.16	35.37
58.32	41.05	33.72	45.73	37.90	41.72	36.07	49.83	36.24	39.90

ต้นทุนเฉลี่ย ต่อหน่วยของสินค้ามีแนวโน้มหรือไม่?

Ho : ไม่มีแนวโน้มในต้นทุนเฉลี่ยต่อหน่วย

Ha : มีแนวโน้มในต้นทุนเฉลี่ยต่อหน่วย

เมื่อจับคู่ค่าสังเกตจะได้เป็น

(45.25,41.05),(45.83,33.72),...,(35.37,39.90)

เราได้ T= 4 สำหรับ α= .039 เราได้ค่าวิกฤตของ T เป็น 1 และ 8 ดังนั้นจึงปฏิเสธ Ho ไม่ได้ นั่นคือต้นทุนเฉลี่ยต่อหน่วยของสินค้าไม่มีแนวโน้ม

แบบทดสอบนี้ยังดัดแปลงใช้ทดสอบเกี่ยวกับสหสัมพันธ์ และทดสอบแบบแผนชนิดไม่เป็นเชิงสุ่มได้อีก

8.1.3 แบบทดสอบเครื่องหมายชนิดสองตัวอย่างที่เป็นอิสระกัน (Two Independent Sample Sign Test)

แบบทดสอบนี้ใช้ทดสอบสมมติฐานหลักที่ว่าสองตัวอย่างสุ่มที่เป็นอิสระกันมาจากประชากรแบบเดียวกัน การทดสอบก็อาศัยการจับคู่กันแบบสุ่ม (Randomly Paired) ของค่าสังเกตจากตัวอย่างทั้งสอง และแทนแต่ละคู่ด้วยเครื่อง + และ – นั่นคือค่าของตัวอย่างแรกน้อยกว่าตัวอย่างที่สอง ก็ให้เป็น + แต่ถ้ามากกว่าก็ให้เป็น – ในกรณีที่ขนาดตัวอย่างทั้งสองไม่เท่ากันก็ต้อง

ทิ้งค่าของตัวอย่างที่ใหญ่กว่าเพื่อให้มีขนาดตัวอย่างเท่ากัน

สำหรับกระบวนการทดสอบก็ทำได้เช่นเดียวกับแบบทดสอบเครื่องหมายชนิดตัวอย่างเดียวนั่นเอง

ในการแปลงค่าสังเกตให้เป็นเครื่องหมายนั้นจะอาศัยมัธยฐานร่วม (Common Median) ก็ได้ นั่นคือจากค่าสังเกตของตัวอย่างทั้งสอง เราหามัธยฐานร่วม แล้วแปลงค่าในตัวอย่างทั้งสองให้เป็นเครื่องหมาย + และ – โดยแทนค่าสังเกตที่มากกว่ามัธยฐานร่วมเป็น + และน้อยกว่าเป็น – แล้วเราจะได้ตารางดังนี้

	1	2	
+	X_{11}	X_{12}	X_1
–	X_{21}	X_{22}	X_2
	n_1	n_2	n

ในเมื่อ X_{11} และ X_{12} เป็นจำนวนเครื่องหมาย + ในตัวอย่างที่ 1 และ 2 ตามลำดับ ส่วน X_{21} และ X_{22} ก็เป็นจำนวนเครื่องหมาย –

จากตารางนั้นก็ใช้แบบทดสอบสัดส่วนประยุกต์ นั่นคือใช้ตัวสถิติทดสอบ

$$Z = (P_1 - P_2)/\sqrt{PQ(1/n_1 + 1/n_2)}$$

ในเมื่อ $P_1 = X_{11}/n_1$, $P_2 = X_{12}/n_2$, $P = X_1/n$ และ $Q = 1 - P = X_2/n$

ตัวสถิติทดสอบ Z นี้จะมีการแจกแจงปกติมาตรฐานถ้า n_1 และ n_2 โตพอ

ตัวอย่าง ในการศึกษาวิธีการผลิต 2 วิธี ว่าให้ผลผลิตเฉลี่ยต่อชั่วโมงสำหรับการผลิตสินค้าชนิดหนึ่ง ปรากฏว่าได้ผลผลิตจากการศึกษาดังนี้

วิธีผลิต ก.

25	13	19	46	25
30	17	20	17	20
37	25	26	23	20
17	18	26	11	36
30	12	32	48	24
20	16	18	21	37
31	26			

วิธีผลิต ข.

31	43	21	42	38
30	19	20	36	29
13	50	32	41	28
30				

วิธีการผลิตทั้งสองวิธีให้ผลผลิตเฉลี่ยต่อชั่วโมงแตกต่างกันหรือไม่?

Ho : วิธีผลิตทั้งสองวิธีให้ผลผลิตเฉลี่ยไม่แตกต่างกัน

Ha : วิธีการผลิตทั้งสองวิธีให้ผลผลิตเฉลี่ยแตกต่างกัน

จากผลผลิต 48 ครั้ง เราหามัธยฐานร่วมได้เป็น (25+26)/2 =25.50 แล้วเราแทนผลผลิตในแต่ละวิธีที่มากกว่าหรือน้อยกว่ามัธยฐานร่วมด้วยเครื่องหมาย + และ − ซึ่งจะได้ผลดังตารางต่อไปนี้

วิธีการผลิต	ก	ข	รวม
+	12	12	24
−	20	4	24
ขนาดตัวอย่าง	32	16	48

เราได้ P = 24/48 = 0.50 Q = 0.50

$$Z = \frac{12/32 - 12/16}{\sqrt{0.50(0.50)(1/32+1/16)}} = -2.45$$

สำหรับ α = .05 ค่าวิกฤตของตัวสถิติทดสอบ Z จะเป็น ± 1.96 เนื่องจากค่าของตัวสถิติ Z เป็น −2.45 ซึ่งน้อยกว่า −1.96 และอยู่ในเขตวิกฤต จึงปฏิเสธ Ho นั่นคือผลผลิตเฉลี่ยในวิธีการผลิตทั้งสองจะแตกต่างกัน (วิธีการผลิต ข. จะให้ผลผลิตเฉลี่ยมากกว่าวิธีการผลิต ก.)

เนื่องจากแบบทดสอบนี้ได้ใช้มัธยฐานร่วม จึงได้ชื่อว่าแบบทดสอบมัธยฐาน (Median Test)

8.14 แบบทดสอบเครื่องหมายชนิดจับคู่ตัวอย่าง (Paired-Sample Sign Test)

แบบทดสอบนี้เหมือนกับ แบบทดสอบเครื่องหมายชนิดตัวอย่างเดียว ยกเว้นแต่ว่าข้อมูลจับคู่กันเป็นคู่ ๆ (Matched Pairs) การวิเคราะห์จึงอาศัยวิธีการดังกล่าวนั้น

ตัวอย่าง บริษัทผู้ผลิตอาหารสำเร็จรูป ต้องการทดสอบผงชูรสสองชนิดที่ใช้ในการปรุงอาหารว่าแตกต่างกันหรือไม่ โดยอาศัยผู้ชิมที่สุ่มมา 20 คน ให้แต่ละคนชิมอาหารสองที่ซึ่งเหมือนกัน แต่ว่าใช้ผงชูรสต่างกัน เมื่อผู้ชิมชอบชนิดไหนมากกว่าก็ให้เป็น + ชอบน้อยกว่าก็เป็น −

จากผลทดลองนี้ปรากฏว่า ผงชูรสชนิดแรกมีผู้ชอบมากกว่า 14 ราย จะสรุปผลทดลองว่าอย่างไร?

Ho : ผงชูรสทั้งสองชนิดไม่แตกต่างกัน

Ha : ผงชูรสทั้งสองชนิดแตกต่างกัน

เมื่อ $n = 20; X = 14$ เราได้

$$Z = \frac{14-20/2}{\sqrt{20/4}} = 1.79$$

สำหรับ α = .05 จะปฏิเสธ Ho ไม่ได้ นั่นคือผงชูรสทั้งสองชนิดไม่แตกต่างกัน

8.1.5 แบบทดสอบเครื่องหมายชนิดหลายตัวอย่าง (Mult sample Sign Test)

แบบทดสอบนี้ทำหน้าที่เช่นเดียวกับแบบทดสอบเครื่องหมายชนิดสองตัวอย่างที่เป็นอิสระกัน แต่ใช้ประยุกต์กับตัวอย่างสุ่มอย่างน้อยสองตัวอย่าง นั่นคือแบบทดสอบจะใช้ตรวจสอบว่าหลายตัวอย่างมาจากประชากรที่มีมัธยฐานสลับกันหรือไม่ดังนั้นสมมติฐานหลักจึงเป็น

Ho : ประชากรทั้งหมดมีมัธยฐานเท่ากัน

หรือ Ho : ตัวอย่างต่าง ๆ มาจากประชากรที่มีมัธยฐานเหมือนกัน

ในการทดสอบมัธยฐานก็อาศัยตัวอย่างสุ่มจากประชากรต่าง ๆ เมื่อหามัธยฐานร่วมจากตัวอย่างเหล่านั้น แล้วแปลงข้อมูลในตัวอย่างต่าง ๆ นั้น ให้เป็น + และ – แล้วแต่ว่าข้อมูลนั้นมากกว่าหรือน้อยกว่ามัธยฐานร่วม เมื่อสรุปผลก็จะได้ตาราง 2 x k ดังนี้

ตัวอย่าง	1	2	k	รวม
+	O_{11}	O_{12}	O_{1k}	a
–	O_{21}	O_{22}	O_{2k}	b
รวม	n_1	n_2	n_k	n

ในเมื่อ $n_1, n_2, \ldots, n_k$ เป็นขนาดตัวอย่างจากประชากร 1,2,....,k ตามลำดับ O_{1j} และ O_{2j} เป็นจำนวนค่าสังเกตหรือข้อมูลที่แปลงเป็น + และ – และ a,b เป็นจำนวนข้อมูลทั้งหมดที่แปลงเป็น + และ –

ตัวสถิติทดสอบ จะกำหนดไว้ดังนี้

$$X^2 = \frac{n^2}{ab}\sum (O_{1j} - n_j p)^2/n_j \qquad = \frac{n^2}{ab}\sum O_{1j}^2/n_j - na/b$$

ซึ่งมีการแจกแจงไคสแควร์ด้วยองศาความเป็นอิสระ k-1

ตัวอย่าง นักเกษตรต้องการเปรียบเทียบวิธีการปลูกข้าวโพด 4 วิธี ว่าให้ผลผลิตต่อไร่แตกต่างกันหรือไม่ จากการทดลองได้ผลผลิตมาดังนี้

วิธีปลูก 1	83	91	94	89	89	96	91	92	90	
2	91	90	81	83	84	83	88	91	89	84
3	101	100	91	93	96	95	94			
4	78	82	77	79	81	80	81	81		

Ho : วิธีปลูกทั้ง 4 วิธีให้ผลผลิตไม่แตกต่างกัน

จากข้อมูลทั้งหมด 34 จำนวน ปรากฏว่าได้มัธยฐานร่วมเป็น (89+89)/2 = 89 ดังนั้น จำนวนเครื่องหมาย + และ – ซึ่งสรุปในตาราง 2x k ดังนี้

วิธีปลูก	1	2	3	4	รวม
+	6	3	7	0	16
–	3	7	0	8	18
n_j	9	10	7	8	34

$$X^2 = \frac{34^2}{16(18)}\left\{6^2/9+3^2/10+7^2/7+0^2/4\right\}-34(16)/18$$
$$= 17.60$$

สำหรับ α = .05 เราได้ $X^{2(3)}_{.05} = 7.825$ ยังปฏิเสธ Ho ซึ่งแสดงว่าวิธีปลูกทั้ง 4 วิธีให้ผลผลิตต่อไร่ไม่เท่ากันหมด

8.1.6 แบบทดสอบเครื่องหมายชนิดหลายตัวอย่างที่สัมพันธ์กัน (Sign Test for Several Related Samples)

แบบทดสอบนี้ใช้เปรียบเทียบมัธยฐานของประชากรต่าง ๆ โดยที่ข้อมูลหรือค่าสังเกตนั้นได้จากการวางแผนทดลองชนิดแบ่งบลอคสมบูรณ์ และแปลงค่าสังเกตในแต่ละบลอคให้เป็นเครื่องหมาย + และ – ถ้าค่าสังเกตนั้นมากกว่าหรือน้อยกว่ามัธยฐานในแต่ละบลอคนั้น เมื่อรวมเครื่องหมาย + และ – ในแต่ละกรรมวิธีหรือตัวอย่างแล้วเราจะได้ตาราง 2xk ดังนี้

ตัวอย่าง	1	2	....	k	รวม
+	O_{11}	O_{12}	....	O_{1k}	a
–	O_{21}	O_{22}	....	O_{2k}	b
รวม	n	n		n	kn

ในเมื่อ O_{1j} และ O_{2j} เป็นจำนวนเครื่องหมาย + ในแต่ละตัวอย่าง a และ b เป็นผลรวมของเครื่องหมาย + และ – ของทุกตัวอย่าง k เป็นจำนวนตัวอย่าง และ n เป็นจำนวนบลอค

ตัวสถิติทดสอบก็จะเป็นเช่นเดียวกับแบบทดสอบเครื่องหมายชนิดหลายตัวอย่างที่เป็นอิสระกัน นั่นคือ

$$X^2 = \frac{nk^2}{ab}\sum O_{1j}^2 - nka/b$$

ตัวสถิตินี้มีการแจกแจงไคสแควร์ ด้วยองศาความเป็นอิสระ k–1 ถ้ามัธยฐานของทุกประชากรเท่ากันหมด

ตัวอย่าง ในการศึกษาปุ๋ย 4 ชนิด ว่าให้ผลผลิตเฉลี่ยต่อไร่ของข้าวพันธุ์หนึ่งแตกต่างกันหรือไม่ เมื่อทำการทดลองโดยใช้แปลงทดลอง 72 แปลง ที่มีขนาดเท่ากัน แปลงทดลอง 72 แปลงนี้แบ่งเป็น 18 บลอค ปรากฏว่าได้ผลทดลองที่แปลงเป็นเครื่องหมาย + และ — แล้วดังนี้

ปุ๋ย	1	2	3	4	รวม
+	3	14	8	10	35
—	15	4	10	a	37
รวม	18	18	18	18	72

Ho: ปุ๋ยทั้ง 4 ชนิดให้ผลผลิตเฉลี่ยต่อไร่ไม่แตกต่างกัน

$$X^2 = \frac{18(4^2)}{35(37)}\{3^2+14^2+8^2+\ 10^2\} - 18(4)(35)/37 = 13.955$$

สำหรับ $X^2_{.05}(3) = 7.815$ เราจึงสรุปได้ว่าปุ๋ยทั้ง 4 ชนิดให้ผลผลิตเฉลี่ยต่อไร่ไม่เท่ากันหมด

8.2 การทดสอบโดยใช้อันดับ (Rank Test)

การทดสอบสมมติฐานเกี่ยวกับมัธยฐานหรือค่าเฉลี่ยส่วนมากจะอาศัยอันดับ นั่นคือค่าสังเกตที่ได้จากตัวอย่างนั้นอย่างน้อยต้องเป็นแบบอันดับ แบบทดสอบที่ใช้อันดับมีต่าง ๆ กันดังนี้

8.2.1 แบบทดสอบอันดับที่เครื่องหมายของวิลคอกซัน (Wilcoxon Signed Rank Test)

แบบทดสอบนี้ใช้ทดสอบสมมติฐานเกี่ยวกับมัธยฐานประชากรที่ระบุไว้ (μ_o) โดยที่ประชากรนั้นเป็นแบบสมมาตรเนื่องจากแบบทดสอบนี้ทำหน้าที่เช่นเดียวกับแบบทดสอบเครื่องหมายแต่ใช้ข้อมูลข่าวสารจากตัวอย่างมากกว่า จึงมีอำนาจทดสอบมากกว่า **สมมติฐานหลักที่จะ**ทดสอบจึงเป็น

$$Ho: \mu = \mu_o$$

ตัวสถิติทดสอบกำหนดจากตัวอย่างสุ่ม $X_1, X_2, \ldots, X_n$ โดยหาผลต่าง Z_i

$$Z_i = X_i - \mu_o;\ T=1,2,\ldots,n$$

แล้วเรียงลำดับค่าสมบูรณ์ $|Z_i|$ จากน้อยไปมาก เมื่อให้ R_i แทนอันดับที่ของ $|Z_i|$ แล้วตัวสถิติทดสอบจะเป็น

$$T = \min(T+\ T-)$$

โดย T_+ และ T_- เป็นผลรวมของอันดับที่ในเมื่อ Z_i เป็นบวกและลบตามลำดับ

ตัวสถิติ T นี้เป็นค่าที่น้อยสุดระหว่างผลรวมของอันดับที่เครื่องหมายบวกและลบนั่นเอง

และตัวสถิตินี้มีค่าเฉลี่ยและความแปรปรวนเป็น

$$E(T) = n(n+1)/4$$

$$V(T) = n(n+1)(2n+1)/24$$

เมื่อตัวอย่างขนาดโต เราก็อาศัยตัวสถิติ Z ซึ่งมีการแจกแจงปกติมาตรฐาน ดังนี้

$$Z = \frac{T-E(T)}{\sqrt{V(T)}}$$

สำหรับตัวอย่างขนาดเล็ก เราอาศัยตารางพิเศษช่วยในการตัดสินใจเกี่ยวกับสมมติฐานหลักที่ทดสอบ

ตัวอย่าง ในการศึกษาเพื่อทดสอบคำกล่าวที่ว่า ค่าใช้จ่ายเฉลี่ยต่อวันของนักศึกษา ม.ร. จะประมาณ 30 บาท นั้นได้ใช้ตัวอย่างของนักศึกษา ม.ร. 10 คน ได้ค่าใช้จ่ายต่อวันเป็นดังนี้

28 29 35 43 36 39 38 26 44 37 จะสรุปผลว่าอย่างไร?

$$Ho: \mu = 30 ; Ha: \mu \neq 30$$

จากข้อมูลเราได้ $Z_i = X_i - 30$ และอันดับที่ของ $|z_i|$ เป็นดังนี้

Z_i	-2	-1	5	13	6	9	8	-4	14	7
R_i	2	1	4	9	5	8	7	3	10	6

$T_+ = 4+9+5+8+7+10+6 = 49$

$T_- = 2+1+3 = 6$

ดังนั้น $T = \min (49.6) = 6$

สำหรับ $\alpha = .05$ เราได้ค่าวิกฤตเป็น 8 ดังนั้นจึงปฏิเสธ Ho แสดงว่าค่าใช้จ่ายต่อวันของนักศึกษา ม.ร.มากกว่า 30 บาท

8.2.2 แบบทดสอบอันดับเครื่องหมายวิลคอกซัน ชนิดสองตัวอย่างจับคู่กัน (Wilcoxon Matched-Pairs Signed-Ranks Test)

แบบทดสอบวิลคอกซันนี้ใช้ทดสอบผลต่างของค่าเฉลี่ยเช่นเดียวกับแบบทดสอบที่ชนิดข้อมูลจับคู่นั่นคือสมมติฐานหลักที่ทดลองจะเป็น

$$Ho: \mu_1 = \mu_2 \quad \text{หรือ} \quad \mu_D = 0$$

ข้อมูลที่ใช้วิเคราะห์จะเป็น $(X_i, Y_i); i=1,2,\ldots n$ ซึ่งมีผลต่างเป็น $Z_i = X_i - Y_i$ เมื่อเรียงค่าสมบูรณ์ของ z_i จากน้อยไปมากแล้วกำหนดอันดับที่ให้แก่ค่า $|z_i|$ เหล่านั้นแล้วเราจะได้ตัวสถิติทดสอบ T

$$T = \min (T_+, T_-)$$

ในเมื่อ T_+ และ T_- เป็นผลรวมของอันดับที่ซึ่งมี z_i เป็น + และ - ตามลำดับ

ตัวสถิติ T นี้มีค่าคาดหวัง และความแปรปรวนเป็น

$$E(T) = n(n+1)/4$$

$$V(T) = n(n+1)(2n+1)/24$$

ดังนั้นในเมื่อตัวอย่างขนาดโต เราจึงใช้ตัวสถิติทดสอบ ซึ่งมีการแจกปกติมาตรฐาน

$$Z = \frac{T-E(T)}{\sqrt{V(T)}}$$

ตัวอย่าง ในการศึกษาวิธีการผลิต 2 วิธี โดยใช้ผู้ควบคุมการผลิต 11 คน ควบคุมการผลิตทั้งสองวิธี ปรากฏว่าได้ผลผลิตเฉลี่ยต่อชั่วโมงดังนี้

วิธี	1	88	77	76	64	96	65	90	65	80	81	72
	2	86	71	77	68	91	77	91	70	71	88	87

วิธีการผลิต 1 จะให้ผลผลิตเฉลี่ยต่อชั่วโมงน้อยกว่าวิธี 2 หรือไม่?

$$Ho: \mu_1 = \mu_2 \quad ; \quad Ha: \mu_1 < \mu_2$$

เมื่อหาผลต่าง z_i และกำหนดอันดับที่ให้แก่ $|z_i|$ เราจะได้เป็นค่าของตัวสถิติ T_+ และ T_- เป็น

$$T_+ = 3+7+5.5+9 = 24.5$$

$$T_- = 1.5+4+10+1.5+5.5+8+11 = 41.5$$

$$T = \min(24.5, 41.5) = 24.5$$

เมื่อ n= 11 และ α = .05 เราได้ค่าวิกฤตเป็น ซึ่งแสดงว่าวิธีการผลิตทั้งสองวิธีให้ผลผลิตเฉลี่ยไม่แตกต่าง

8.2.3 แบบทดสอบแมนน์-วิทนีย์ (Mann-Whitney U Test)

แบบทดสอบแมนน์-วิทนีย์ เป็นแบบทดสอบแบบผลรวมอันดับที่ใช้ทดสอบสมมติฐานที่ว่าตัวอย่างที่เป็นอิสระสองตัวอย่างนั้นสุ่มมาจากประชากรเดียวกันหรือเหมือน ๆ กัน นั่นคือ

$$Ho: \mu_1 = \mu_2$$

แบบทดสอบนี้ใช้แทนการทดสอบ ที่ชนิดสองตัวอย่างเป็นอิสระกันนั่นเอง

ค่าสังเกตจากตัวอย่างทั้งสองเมื่อกำหนดอันดับที่รวม ๆ กันจาก 1 ถึง n_1+n_2 ของค่าสังเกตที่เรียงลำดับจากน้อยไปมากแล้วเราจะได้ตัวสถิติทดสอบเป็น

$$U = \min(U_1, U_2)$$

ในเมื่อ

$$U_1 = n_1 n_2 + n_1(n_1+1)/2 - R_1$$

$$U_2 = n_1 n_2 + n_2(n_2+1)/2 - R_2$$

โดยที่ R_1 และ R_2 เป็นผลรวมอันดับของค่าสังเกตในตัวอย่างที่ 1 และ 2 ตามลำดับ

ตัวสถิติ U จะมีค่าเฉลี่ย และความแปรปรวนเป็น

$$E(U) = n_1 n_2/2$$

$$V(U) = n_1 n_2 \ (n_1+n_2+)/12$$

เมื่อตัวอย่างขนาดโต เราจึงต้องใช้ตัวสถิติซึ่งมีการแจกแจงแบบปกติมาตรฐาน ดังนี้

$$Z = \frac{V+E(U)}{\sqrt{V(U)}}$$

ถ้าให้ R เป็นผลรวมอันดับที่ซึ่งกำหนดให้แก่ตัวอย่างที่มีขนาดเล็กกว่า ($n_1 \leq n_2$) แล้ว แล้วตัวสถิติทดสอบจะเป็น

$$Z = \frac{R-E(R)}{\sqrt{V(r)}}$$

ในเมื่อ $E(r) = n_1(n_1+n_2+1)/2$ และ $V(r) = n_1 n_2(n_1+n_2+1)/12$

ตัวอย่าง ในการศึกษาอายุใช้งานของหลอดไฟ 2 ยี่ห้อโดยอาศัยตัวอย่างของหลอดไฟ ได้ผลของอายุใช้งานดังนี้

ยี่ห้อ	ก	981	952	1342	1051	1005	974	1216		
	ข	1380	1004	1032	1263	1040	990	1102	1170	1205

หลอดไฟยี่ห้อ ข. มีอายุใช้งานเฉลี่ยน้อยกว่ายี่ห้อ ก.หรือไม่ ?

$$H_o: \mu_1 = \mu_2; \ H_a: \mu_1 < \mu_2$$

เมื่อเรียงตัวอย่างทั้งสองจากน้อยไปมากและกำหนดอันดับที่ แล้วเราจะได้ค่าของตัวสถิติทดสอบ R เป็น

$$R = 1+2+3+6+9+13+15 = 49$$

สำหรับ $n_1 = 7, n_2 = 9$ และ $\alpha = .05$ เราจะได้ค่าวิกฤตของ R เป็น 43

ดังนั้นเราจึงปฏิเสธ Ho ไม่ได้ แสดงว่าหลอดไฟทั้งสองยี่ห้อมีอายุใช้งานเฉลี่ยไม่แตกต่างกัน

8.2.4 แบบทดสอบซีเกล-ทูกี้ (Siegel-Tukey Test for Equal Variability)

แบบทดสอบนี้ใช้ทดสอบสมมติฐานที่ว่าสองประชากรมีความผันแปรเท่ากัน (ถ้าเชื่อว่าประชากรทั้งสองมีค่ากลางหรือมัธยฐานเท่ากัน) วิธีการจะคล้ายกับแบบทดสอบแมนน์–วิทนีย์ แต่การจัดอันดับที่ไม่เหมือนกัน นั่นคือการกำหนดอันดับที่นั้นจะกำหนดอันดับ 1 ให้แก่ค่าน้อยที่สุดอันดับ 2 ให้แก่ค่ามากที่สุด อันดับ 3 ให้แก่ค่ามากรองลงมา อันดับ 4 ให้แก่ค่าน้อยอันดับสอง อันดับ 5 ให้แก่ค่าน้อยอันดับสาม และต่อ ๆ ไป ดังนั้นอันดับจะเป็น

$$1\ 4\ 5\ 8\ 9 \ldots (n_1+n_2) \ldots 7\ 6\ 3\ 2$$

ตัวสถิติทดสอบจะเป็นเช่นเดียวกับแบบทดสอบแมนน์วิทนีย์

8.2.5 แบบทดสอบของมูดเกี่ยวกับการกระจาย (MoodTest for Dispersion)

แบบทดสอบของมูดใช้เปรียบเทียบการกระจายของสองประชากรแบบต่อเนื่องที่เป็นอิสระกัน และประชากรทั้งสองจะมีมัธยฐานไม่แตกต่างกัน แต่การกระจายอาจจะแตกต่างกันได้

ตัวสถิติทดสอบจะเป็น

$$M = \sum_{i=1}^{n_1} (R_i - \bar{R})^2$$

ในเมื่อ n_1 เป็นขนาดตัวอย่างของกลุ่มตัวอย่าง 1 และ $n_1 \leq n_2$ ส่วน R_i เป็นอันดับที่ของค่าสังเกตในตัวอย่าง i ของการเรียงอันดับที่ร่วมกันของค่าสังเกตจากสองตัวอย่าง และ $\bar{R} = (n+1)/2$ เป็นอันดับที่เฉลี่ยของค่าสังเกต $(n = n_1+n_2)$

ตัวสถิติ M จะมีค่าเฉลี่ย และความแปรปรวนเป็น

$$E(M) = n_1 (n^2 - 1)/12$$

$$V(M) = n_1 n_2 (n - 1)(n^2 - 4)/180$$

เมื่อขนาดตัวอย่างโต $(n > 30)$ เราจึงใช้ตัวสถิติ Z ซึ่งมีการแจกแจงปกติมาตรฐาน

$$Z = [M - E(M)]/\sqrt{V(M)}$$

ถ้าขนาดตัวอย่างน้อยกว่า 30 เราก็ใช้ตัวสถิติ Z ได้แต่ต้องแก้ไขความต่อเนื่อง นั่นคือตัวสถิติจะเป็น Z

$$Z = [M - E(M)]/\sqrt{V(M)} \quad + \quad 1/2\sqrt{V(M)}$$

ตัวอย่าง ในการเปรียบการกระจายของประชากรโดยอาศัยตัวอย่าง ได้ข้อมูลมาดังนี้

ตัวอย่าง 1 3.84 2.60 1.19 2.00

2 3.97 2.50 2.70 3.36 2.30

$$Ho: \sigma_1^2 = \sigma_2^2 \quad ; \quad Ha: \sigma_1^2 \neq \sigma_2^2$$

เรียงค่าสังเกต และกำหนดอันดับที่ จะได้เป็น

ค่าสังเกต	1.19	2.00	2.30	2.50	2.60	2.70	3.36	3.84	3.97
กลุ่มตัวอย่าง	1	1	2	2	1	2	2	1	2
อันดับที่	1	2	3	4	5	6	7	8	9

$$\bar{R} = (4+5+1)/2 = 5$$

$$M = (1-5)^2 + (2-5)^2 + (5-5)^2 + (8-5)^2 = 34$$

$$E(M) = 4(9^2 - 1)/12 = 26.67$$

$$V(M)\ 4(5)\ (9+1)\ (9^2 - 4)/180 = 85.56$$

$$Z = (34 - 26.67)/\sqrt{85.56} + 1/2\sqrt{85.56}$$

$$= 0.846$$

เมื่อ $\alpha = .05$ เราได้ $Z_{.025} = 1.96$ ดังนั้นจึงปฏิเสธ Ho ไม่ได้ แสดงว่าตัวอย่างทั้งสองมีการกระจายไม่แตกต่างกัน

8.2.6 แบบทดสอบครัสคัล-วอลลิส (Kruskal-Wallis Test)

แบบทดสอบนี้เป็นรูปทั่วไปของแบบทดสอบแมนน์-วิทนีย์ ซึ่งใช้ทดสอบการเท่ากันของหลาย ๆ ค่าเฉลี่ยที่เป็นอิสระกัน หรือใช้วิเคราะห์ข้อมูลที่แจกแจงทางเดียวโดยอาศัยอันดับที่ ดังนั้นแบบทดสอบครัสคัล-วอลลิส จึงได้ชื่อว่าการวิเคราะห์ความแปรปรวนทางเดียวโดยอาศัยอันดับที่ (One-way ANOVA by Ranks) สมมติฐานที่จะทดสอบจึงเป็น

$$Ho: \mu_j = \mu \quad \forall j \quad (j = 1, 2, \ldots, k)$$
$$Ha: \mu_j \neq \mu \quad \exists j$$

ตัวสถิติทดสอบ จะอาศัยอันดับที่ซึ่งเป็นผลรวมของอันดับที่ของค่าสังเกตในตัวอย่างนั้นการกำหนดอันดับที่นั้นให้กำหนดร่วมกันจาก 1 ถึง n ในเมื่อ $n = n_1+n_2+ \ldots + n_k$ ดังนั้นตัวสถิติทดสอบจึงเป็น

$$H = 12/n(n+1) \sum n_j (\bar{R}_j - \bar{R})^2$$
$$= \frac{12}{n(n+1)} \sum R_j^2/n_j = 3(n+1)$$

ซึ่งมีการแจกแจงไคสแควร์ ด้วยองศาความเป็นอิสระ k - 1 ถ้าตัวอย่างขนาดโต และ Ho เป็นจริง ในเมื่อ R_j เป็นผลรวมของอันดับที่ในตัวอย่างที่ j $\bar{R}_j = R_j/n_j$ และ $\bar{R} = \sum R_j/n = (n+1)/2$

ตัวอย่าง ในการศึกษาวิธีปลูกข้าวโพด 4 วิธี ว่าจะให้ผลผลิตต่อไร่ แตกต่างกันหรือไม่ ได้ผลทดลองดังนี้

วิธีปลูก 1	2	3	4
83(11)	91(23)	101(34)	78(2)
91(23)	90(19.5)	100(33)	82(9)
94(28.5)	81(6.5)	91(23)	81(6.5)
89(17)	83(11)	93(27)	77(1)
89(17)	84(13.5)	96(31.5)	79(3)
96(31.5)	83(11)	95(30)	81(6.5)
91(23)	88(15)	94(28.5)	80(4)
90(19.5)	89(17)		81(6.5)
	84(13.5)		
R_j 196.5	153.0	207.0	38.5

Ho : วิธีปลูกข้าวโพดทั้ง 4 วิธีนี้จะให้ผลผลิตต่อไร่ไม่แตกต่างกัน

$$H = \frac{12}{34(34-1)}\left(196.5^2/9 + 153.0^2/10 + 207.0^2/7 + 38.5^2/8\right) - 3(34+1)$$

$$= 25.46$$

เมื่อ α = .05 เราได้ $\chi^2(4-1) = 7.815$ จึงปฏิเสธ Ho แสดงว่ามีวิธีปลูกข้าวโพดบางวิธีให้ผลผลิตต่อไร่ต่างจากวิธีอื่น

เมื่อปฏิเสธ Ho และต้องการทราบว่าสองตัวอย่างใด ๆ มีค่าเฉลี่ยต่างกัน ก็ทำได้โดยอาศัยช่วงเชื่อมั่นของ $\mu_i - \mu_j$ $(i < j)$ ดังนี้

$$\mu_i - \mu_j = (\bar{R}_i - \bar{R}_j) \pm \sqrt{\chi^2_{\alpha}(k-1)}\sqrt{\frac{n(n+1)}{12}(1/n_i + 1/n_j)}$$

ถ้าช่วงคาบเกี่ยว O ไว้ด้วย แสดงว่าค่าเฉลี่ยทั้งสองไม่แตกต่างกัน

บางครั้งถ้าสนใจความแผกกัน ψ เพื่อดูความแตกต่างของกลุ่มตัวอย่างต่าง ๆ ได้ จะทำได้โดยการสร้างช่วงเชื่อมั่นพหุคูณของ ψ ดังนี้

$$\Psi = \hat{\Psi} \pm \sqrt{\chi^2_\alpha(k-1)} \sqrt{(n(n+1)/12)\sum a_j^2/n_j}$$

โดยที่ $\Psi = \sum a_j M_j$: $\hat{\Psi} = \sum a_j R_j$ ในเมื่อ $\sum a_j = 0$

8.2.7 **แบบทดสอบฟรีดแมน** (Friedman Test)

แบบทดสดบฟรีดแมนเป็นรูปทั่วไปของแบบทดสอบวิลคอกซันชนิดจับคู่เพื่อใช้ทดสอบค่าเฉลี่ยของกลุ่มตัวอย่างที่สัมพันธ์กัน หรือใช้วิเคราะห์ข้อมูล 2 ทาง ซึ่งทำให้แบบทดสอบนี้ได้ชื่อว่า การวิเคราะห์ความแปรปรวนสองทางแบบอันดับที่ นั่นคือสมมติฐานหลักจะเป็น

$$H_0: \mu_j = \mu \; ; \; \forall j \; (j = 1, 2, \ldots, k)$$

ตัวสถิติทดสอบสำหรับสมมติฐานนี้จะกำหนดได้จากอันดับที่ซึ่งกำหนดในแต่ละบลอดจาก 1 ถึง k นั่นคือตัวสถิติทดสอบเป็น

$$S = \frac{12b}{k(k+1)} \sum (\bar{R}_j - \bar{R})^2$$

$$= \frac{12}{bk(k+1)} \sum R_j^2 - 3b(k+1)$$

ในเมื่อ R_j เป็นผลรวมของอันดับที่ในตัวอย่างที่ j $\bar{R}_j = R_j/b$ เป็นค่าเฉลี่ย และ $\bar{R} = \sum R_j/bk = (k+1)/2$ เป็นค่าเฉลี่ยรวมของอันดับ

ในเมื่อตัวอย่างขนาดโต ($b \Rightarrow \infty$) ตัวสถิติทดสอบ S จะมีการแจกแจงไคสแควร์ด้วยองศาความเป็นอิสระ k-1 ดังนั้นเกณฑ์ตัดสินใจจึงกำหนดไว้ว่า "ปฏิเสธ Ho ณ ระดับนัยสำคัญ α ถ้า

$$S \geqslant \chi^2_\alpha(k-1)$$

ตัวอย่าง ชาวสวนที่ปลูกหญ้าขายคนหนึ่งต้องการทราบว่าหญ้าชนิดใดเป็นที่นิยมของประชาชน จากการทดลองกับเจ้าของบ้าน 12 ราย โดยให้จัดอันดับที่ของหญ้า 4 ชนิดได้ผลดังนี้ (1 ชอบน้อยที่สุด 4 ชอบมากที่สุด)

ชนิดของหญ้า	ก	ข	ค	ง
เจ้าบ้าน 1	4	3	2	1
2	4	2	3	1
3	3	1	2	4
4	3	1	2	4
5	4	2	1	3

6	3	1	2	4
7	1	3	2	4
8	2	4	1	3
9	3	1	2	4
10	4	1	3	2
11	4	2	3	1
12	3	1	2	4
R_j	38	22	25	35

สมมติฐาน Ho : หญ้าทั้งสี่ชนิดได้รับความนิยมพอ ๆ กัน

Ha : มีหญ้าบางชนิดได้รับความนิยมมากกว่า

$$S = \frac{12}{12(4)(5)}(38^2+22^2+25^2+35^2) - 3(12)(5)$$
$$= 8.9$$

เนื่องจากค่า S โตกว่าค่าวิกฤต $X^2_{.05}(4-1) = 7.815$ จึงปฏิเสธ Ho นั่นคือมีหญ้าบางชนิดได้รับความนิยมมากกว่า

เมื่อปฏิเสธ Ho แล้วต้องการเปรียบเทียบระหว่างกรรมวิธี หรือระหว่างตัวอย่างก็ทำได้โดยการเปรียบเทียบเชิงพหุคูณ นั่นคือสร้างช่วงเชื่อมั่น $100(1-\alpha)\%$ สำหรับ $\mu_i - \mu_j \; (i<j)$ ได้เป็น

$$\mu_i - \mu_j = (\bar{R}_i - \bar{R}_j) \pm \sqrt{X^2_{\alpha}(k-1)}\sqrt{k(k+1)/6b}$$

ถ้าช่วงใดคาบเกี่ยว 0 ไว้ด้วยก็แสดงว่าค่าเฉลี่ยไม่แตกต่างกัน

ถ้าต้องการเปรียบเทียบโดยอาศัยความแผกกัน (ψ) หรือผลรวมเชิงเส้นของค่าเฉลี่ยก็ทำได้ดังนี้

$$\psi = \hat{\psi} \pm \sqrt{X^2_{\alpha}(k-1)}\;\sqrt{V(\hat{\psi})}$$

ในเมื่อ $\hat{\psi} = a_1\bar{R}_1 + a_2\bar{R}_2 + \dots + a_k\bar{R}_k \; ; a_1 + a_2 \dots + a_k = 0$ และ $V(\hat{\psi})\; \frac{k(k+1)}{12}\sum a_j^2/b$

8.3 การทดสอบโดยอาศัยรัน (Run Test)

รัน (Run) เป็นอนุกรมของสัญญลักษณ์ที่เหมือนกันซึ่งอาจจะตามหรือนำสัญญลักษณ์อื่น ๆ หรือไม่มีสัญญลักษณ์อื่นตามหรือนำเลยก็ได้

ผลรวมของรัน (R) ในการจัดเรียงสัญญลักษณ์สองชนิดหรือมากกว่า จะนำไปใช้ตรวจสอบการสุ่ม (Randomness) ของการจัดเรียง และใช้ทดสอบสองตัวอย่างสุ่มที่เป็นอิสระกันว่ามาจากประชากรที่เป็นแบบเดียวกันหรือไม่

8.3.1 **ทดสอบการสุ่ม** (Test for Randomness)

ในการตรวจสอบการสุ่มโดยอาศัยวันนั้นมีวิธีการต่าง ๆ กัน คือใช้ผลรวมของรัน ใช้วันที่อยู่เหนือและใต้มัธยฐาน และใช้วันที่อยู่บนและล่างดังนี้

(1) ผลรวมของรัน (Total number of Runs) ผลรวมของรันใช้ทดสอบการสุ่มของอนุกรมสัญญลักษณ์สองชนิด ค่าเฉลี่ยและความแปรปรวนของผลรวมของรัน R จะเป็น

$$E(R) = \frac{2n_1n_2}{n_1+n_2} + 1$$

$$V(R) = \frac{2n_1n_2(2n_1n_2-n_1-n_2)}{(n_1+n_2)^2(n_1+n_2-1)}$$

ในเมื่อ n_1 เป็นจำนวนสัญญลักษณ์ประเภทหนึ่ง และ n_2 เป็นจำนวนสัญญลักษณ์อีกประเภทหนึ่ง

เมื่อตัวอย่างขนาดโต (n_1 และ n_2 โตกว่า 20) การทดสอบการสุ่มจึงใช้ตัวสถิติ Z ซึ่งมีการแจกแจงปกติมาตรฐาน

$$Z = [R - E(R)]/\sqrt{V(R)}$$

ถ้าตัวอย่างขนาดเล็กจะอาศัยตารางพิเศษในการกำหนดค่าวิกฤต

ตัวอย่าง ในการศึกษาเกี่ยวกับอันดับเพศของลูกค้าที่เข้ามาติดต่อธนาคาร 50 ราย ปรากฏผลดังนี้

M F M F M M M F F M F M F M F

M M M M F M F M F M M F F F M

F M F M F M M F M M F M M M M

F M F M M

อันดับเพศเป็นแบบ สุ่มหรือไม่ ? (M:ชาย F:หญิง)

Ho : อันดับเพศของลูกค้าที่เข้ามาติดต่อธนาคารเป็นแบบสุ่ม

Ha : อันดับเพศไม่เป็นแบบสุ่ม

จากข้อมูลเราได้ $n_1 = 30$ $n_2 = 20$ $R = 35$

$$E(R) = \frac{2(30)(20)}{30+20} + 1 = 25$$

$$V(R) = \frac{2(30)(20)(2(30)(20)-30-20)}{(30+20)^2(30+20-1)} = 11.265$$

$$Z = (35 - 25)/\sqrt{11.265} = 2.98$$

เนื่องจาก $Z_{.025}$ 1.96 จึงปฏิเสธ Ho แสดงว่าการเข้ามาติดต่อของลูกค้าตามเพศไม่เป็นไปแบบสุ่ม

(2) รันที่อยู่เหนือและใต้มัธยฐาน (Runs above and below the median) รันชนิดนี้ใช้ทดสอบการสุ่มของตัวอย่างโดยอาศัยอันดับที่ของค่าจากตัวอย่างที่เลือกสุ่มมา ค่าแต่ละค่าจะแทนด้วยอักษร a หรือ b ตามแต่ว่ามันมีค่ามากกว่าหรือน้อยกว่ามัธยฐานของตัวอย่างแล้วใช้การทดสอบของรันในแบบ (1) ประยุกต์เข้ากับอนุกรมของ a และ b

ถ้าจำนวนตัวอย่างเป็นจำนวนคูและมากกว่า 25 โดยที่ประชากรเป็นแบบต่อเนื่อง แล้วการแจกแจงของจำนวนรัน R จะเป็นแบบปกติที่มีค่าเฉลี่ย และ ความแปรปรวนดังนี้

$$E(R) = n/2 + 1$$

$$V(R) = n(n-2)/4(n-1)$$

นั่นคือตัวสถิติ Z จะมีการแจกแจงปกติมาตรฐาน

$$Z = (R - E(R)) / \sqrt{V(R)}$$

ตัวอย่าง จากการศึกษาผลการสอบสถิติ 203 ของนักศึกษาโดยอาศัยตัวอย่างขนาด 26 ราย ปรากฏว่าได้คะแนนดังนี้

97 89 25 81 11 83 16 96 44 32 98 19 68

33 25 54 74 82 17 49 33 22 62 20 92 80

มัธยฐานจะเป็น (54+49)2 = 51.5

จากข้อมูล เราแทนคะแนนด้วย a หรือ b ดังนี้

a a b a b a b a b b a b a b b a a a b b b a b a a

$$n = 26 \quad n_1 = 9 \quad n_2 = 8 \quad r = 17$$

$$E(r) = 26/2+1 = 14$$

$$V(r) = 26\ (26-2)/4(26-1) = 6.24$$

$$Z = 17-14/\sqrt{6.24} = 1.21$$

เนื่องจาก Z_{025} = 1.96 จึงสรุปได้ว่าตัวอย่างนี้เป็นตัวอย่างสุ่ม

(3) รันที่อยู่บนและล่าง (Runs up and down) รันชนิดนี้ใช้ทดสอบการสุ่มเช่นเดียวกัน แต่พิจารณาค่าสังเกตทีละคู่ที่ติดต่อกันแล้วแทนด้วยเครื่องหมาย + หรือ - ตามแต่ว่าค่าหลังมากกว่าหรือน้อยกว่าค่าแรก แล้วเราจะได้รัน เมื่อใช้การทดสอบรันในแบบ (1) เข้าประยุกต์กับอนุกรมของ + และ - ก็จะตัดสินใจได้ว่าตัวอย่างเป็นแบบสุ่มหรือไม่

จำนวนรันของเครื่องหมาย + และ - นี้จะมีค่าเฉลี่ยและความแปรปรวนดังนี้

$$E(R) = (1/3)(2n - 1)$$

$$V(R) = (1/90)(16n - 2g)$$

เมื่อตัวอย่างขนาดโต $(n \geq 20)$ แล้วจะได้ตัวสถิติ Z ซึ่งมีการแจกแจงปกติมาตรฐาน

$$Z = \frac{R-E(R)}{\sqrt{V(R)}}$$

ตัวอย่าง จากตัวอย่างที่ผ่านมาเราได้เครื่องหมาย + และ - เป็น

-- + - + - + -- + - + -- + + + - + -- + - + --

ดังนั้น

$$R=19 \quad n_1 = 15 \quad n_2 = 11$$

$$E(R) = (1/3)\{2(26-1)\} = 17$$

$$V(R) = (1/90)\{16(26)-29\} = 4.3$$

$$Z = \frac{19-17}{\sqrt{4.3}} = 1.4$$

เมือ $\alpha = .05$ เราได้ $Z_{.025} = 1.96$ จึงสรุปได้ว่าตัวอย่างเป็นตัวอย่างสุ่ม

8.3.2 แบบทดสอบของวอลด์-วูลโฟวิทซ์(Wald-Wolfowitz Run Test)

แบบทดสอบนี้ใช้ทดสอบว่าตัวอย่างสุ่มสองตัวอย่างมาจากประชากรที่เหมือนกันหรือไม่ วิธีการของแบบทดสอบก็คือเรียงข้อมูลของสองตัวอย่างตามขนาด แล้วแทนแต่ละค่าด้วย 1 หรือ 2 แล้วแต่ว่าข้อมูลนั้นมาจากตัวอย่าง 1 หรือตัวอย่าง 2 นั่นคือ เราจะได้รันของ 1 และ 2 เมื่อใช้การทดสอบรันประยุกต์ก็จะตัดสินใจเกี่ยวกับสมมติฐานที่ว่า "สองตัวอย่างมาจากประชากรที่เหมือนกัน" ได้

ตัวอย่าง ในการศึกษาความก้าวร้าวของเด็ก 4 ขวบ โดยใช้ตัวอย่างของเด็กชาย 12 คน และเด็กหญิง 12 คน ได้คะแนนความก้าวร้าวดังนี้

ช	66	69	72	65	113	65	118	45	141	104	41	50
ญ	55	40	22	58	16	7	9	16	26	36	20	15

Ho : ความก้าวร้าวของเด็กทั้งสองเพศไม่แตกต่างกัน

Ha : ความก้าวร้าวของเด็กทั้งสองเพศแตกต่างกัน

เมื่อเรียงตามขนาดของคะแนนความก้าวร้าวและแทนด้วย 1 หรือ 2 เราจะได้เป็น

2 2 2 2 2 2 2 2 2 2 1 1 1 2 2 1 1 1 1 1 1 1 1 1

$$n_1 = 12 \quad n_2 = 12 \qquad R = 4$$

$$E(R) = \frac{2(12)(12)}{12+12} + 1 = 13$$

$$V(R) = \frac{2(12)(12)\,[2(12)(12)-12-12]}{(12+12)^2\,(12+12-1)} = 5.74$$

$$Z = \frac{4-13}{\sqrt{5.74}} = -3.76$$

เนื่องจาก $-Z_{.025} = -1.96$ เราจึงปฏิเสธ Ho ซึ่งแสดงว่าความก้าวร้าวของเด็กทั้งสองเพศแตกต่างกัน

8.4 แบบทดสอบอื่น (Other Tests)

แบบทดสอบไร้พารามิเตอร์ที่น่าสนใจยังมีอีกมากมาย แต่จะขอกล่าวไว้บางแบบทดสอบดังต่อไปนี้

8.4.1 แบบทดสอบอัตราส่วนของ ฟอน นิวแมนน์ (Von Neumann Ratio Test for Independence)

แบบทดสอบนี้ใช้ทดสอบอนุกรมของค่าสังเกตว่าเป็นแบบสุ่มหรือไม่โดยอาศัยผลต่างที่ตัดกัน ดังนั้น จึงได้ชื่ออีกอย่างว่า แบบทดสอบกำลังสองเฉลี่ยของผลต่างที่ติดกัน

ให้ $X_1, X_2, \ldots X_n$ เป็นอนุกรมค่าสังเกต ซึ่งเรียงตามเกณฑ์อย่างหนึ่ง (ตามปฏิทิน)ถ้า X_i ขึ้นอยู่กับ X_{i-1} $(i = 1, 2, \ldots n)$ นั่นคือ อนุกรม X_i ไม่เป็นเชิงสุ่ม แล้วอนุกรมค่าสังเกตจะให้รูปแบบ (Pattern) บางอย่าง ในประชากรเราเรียกความไม่เป็นเชิงสุ่มหรือความพึ่งพิงของอนุกรมค่าสังเกตว่า "สหสัมพันธ์ในตัวเอง (Autocorrelation) แต่ในตัวอย่างเรียกว่า "สหสัมพันธ์เชิงอนุกรม (Serial Correlation)" แต่สองคำนี้มักจะใช้แทนกัน เมื่อ X_i ขึ้นอยู่กับ X_{i-1} แล้วผลต่างที่ติดกัน หรือ $(X_i - X_{i-1})$ จะมีค่าน้อย เพื่อที่จะขจัดเครื่องหมายของผลต่างนี้ จึงใช้กำลังสองเฉลี่ยของผลต่างที่ตัดกัน นั่นคือ

$$\delta^2 = \frac{\Sigma(X_i - X_{i-1})^2}{n-1}$$

ถ้าอนุกรมของ X_i เป็นแบบสุ่ม (ซึ่งเป็นสมมติฐานหลัก Ho) แล้ว จะได้ว่า

$$E(\delta^2) = 2\sigma^2$$

แต่ไม่ทราบความแปรปรวนประชากร σ^2 จึงประมาณด้วยความแปรปรวนตัวอย่าง S_0^2

$$S_0^2 = \frac{\sum(X_i - \bar{X})^2}{n}$$

ซึ่งเป็นค่าประมาณที่เอียงเฉ

ดังนั้น ถ้ากำหนดอัตราส่วน ฟอน นิวแมนน์ (VNR) เป็น

$$VNR = \delta^2/S_0^2$$

แล้วค่าคาดหวังของ VNR จะประมาณ 2

ถ้า X แจกแจงแบบปกติ และถ้า n > 60 แล้ว VNR จะแจกแจงใกล้กับการแจกแจงปกติด้วยค่าเฉลี่ยและความแปรปรวนเป็น

$$E(VNR) = 2n/(n-1)$$

$$V(VNR) = 4n^2(n-2)/(n+1)(n-1)^3$$

เมื่อ n โตมาก $E(VNR) \simeq 2$ และ $V(VNR) \simeq 4/n$

สำหรับตัวอย่างขนาดเล็ก (n < 60) ค่าวิกฤตพิจารณาได้จากตารางพิเศษ แต่ถ้าตัวอย่างขนาดโต (n > 60) เราก็ใช้ตัวสถิติ Z ซึ่งมีการแจกแจงปกติมาตรฐาน

$$Z = \frac{VNR - E(VNR)}{\sqrt{V(VNR)}}$$

ตัวอย่าง ในการศึกษารายได้จากการขายสินค้าชนิดหนึ่งใน 50 วัน ติดต่อกัน ปรากฏว่าได้รายได้สรุปเป็นดังนี้

$$\sum x = 4864 \quad \sum x^2 = 598628 \quad \text{และ} \quad \sum(x_i - x_{i-1})^2 = 273993$$

รายได้เป็นแบบสุ่มหรือไม่ ?

Ho : รายได้ต่อวันเป็นแบบสุ่ม

Ha: รายได้ต่อวันไม่เป็นแบบสุ่ม

$$\delta^2 = \frac{\sum(x_i - \bar{x}_{i-1})^2}{n-1} = 273993/49 = 5591.69$$

$$S_o^2 = \frac{\sum(x_i - \bar{x})^2}{n} = \frac{\sum x^2 - (\sum x)^2/n}{n}$$

$$= \frac{598628 - (4864)^2/50}{50} = 2509.16$$

$$VNR = \delta^2/S_o^2 = 5591.69/2509.16$$

$$= 2.229$$

ดังนั้น ถ้าใช้ตัวสถิติ Z จะได้ค่าเป็น

$$Z = \frac{VNR - E(VNR)}{\sqrt{V(VNR)}} = \frac{2.2-29-2}{\sqrt{4150}} = 0896$$

เมื่อ α = 05 จึงปฏิเสธ Ho ไม่ได้ แสดงว่ารายได้ต่อวันเป็นแบบสุ่ม

8.4.2 แบบทดสอบโคลโมโกรอฟ (Kolmogorov Test)

แบบทดสอบนี้ใช้ทดสอบการปรับที่ดี (Goodness-of-fit) เช่นเดียวกับแบบทดสอบ**ไคสแควร์** แต่แบบทดสอบไคสแควร์นั้นใช้ประยุกต์กับข้อมูลนามบัญญัติ ส่วนแบบทดสอบโคลโมโกรอฟใช้กับข้อมูลแบบอันดับหรือสูงกว่า

แบบทดสอบการปรับที่ดีของโคลโมโกรอฟนี้จะสนใจที่ฟังก์ชันแจกแจงสะสม (Cdf, Cumulative distribution function) ที่กล่าวไว้ในสมมติฐานและที่สังเกตได้ เราจะให้ $S(n)$ และ $F_0(X)$ แทนฟังก์ชันแจกแจงสะสมที่สังเกตได้และที่กล่าวไว้ในสมมติฐาน ดังนั้น สมมติฐานเกี่ยวกับฟังก์ชันแจกแจงสะสมของประชากรจะเป็นอย่างใดอย่างหนึ่งดังนี้

(1) $H_0 : F(X) = F_0(X) , \forall_x$

$H_a : F(X) \neq F_0(X) ; \exists_x$

(2) $H_o : F(X) \geq F_0(X) , \forall_x$

$H_a : F(X) < F_0(X) , \exists_x$

(3) $H_o : F(X) \leq F_0(X) ; \forall_x$

$H_a : F(X) > F_0(X) , \exists_x$

ในการทดสอบสมมติฐานหลักนั้นก็อาศัยตัวอย่างสุ่มขนาด n จากประชากรที่ไม่ทราบฟังก์ชันแจกแจง F(x) และให้ S(x) เป็นฟังก์ชันแจกแจงตัวอย่างที่สังเกตได้ โดยกำหนดได้ดังนี้

S(x) = สัดส่วนของค่าสังเกตตัวอย่างที่น้อยกว่าหรือเท่ากับ x

ดังนั้น ตัวสถิติทดสอบจึงกำหนดไว้ดังนี้

$$D = \max_{x} | S(x) - F_0(x)|$$

นั่นคือ D เป็นระยะทางแนวตั้งที่มากสุดระหว่าง S(x) กับ $F_0(x)$

เกณฑ์ตัดสินใจสำหรับสมมติฐานหลักนั้นได้ใช้ตารางพิเศษช่วย

ตัวอย่าง ในการศึกษาคะแนนความถนัดเชิงวิชาการเพื่อทดสอบสมมติฐานที่ว่าคะแนนความถนัดนั้นมีการแจกแจงปกติซึ่งมีค่าเฉลี่ย 85 และส่วนเบี่ยงเบนมาตรฐาน 15 ได้คำแนนความถนัดจากตัวอย่างมาดังนี้

58	78	84	90	97	70	90	86	82
59	90	70	74	83	90	76	88	84
68	93	70	94	70	110	67	68	75
80	68	82	104	92	112	84	98	80

$H_o : F(X) = F_o(X) ; \forall_x$

$H_a : F(X) \neq F_0(X) ; \exists_x$

ในเมื่อ $F_o(X)$ เป็นการแจกแจงสะสมของประชากรปกติที่มีค่าเฉลี่ย 85 และส่วนเบี่ยงเบนมาตราฐาน 15 เราได้ S(x), $F_o(X)$ และ d= $|S(X)-F_o(X)|$ ดังตารางต่อไปนี้

x	58	59	67	68	74	75
$F_o(x)$	.036	.042	.115	.129	.233	.251
S(x)	.028	.056	.083	.167	.306	.333
d	.008	.014	.032	.119	.073	.082

x	76	78	80	82	83	84	86
$F_o(x)$	.274	.320	.371	.421	.448	.472	.528
S(x)	.361	.389	.444	.500	.528	.611	.639
d	.087	.069	,073	.079	.080	.139	.111

x	88	90	92	93	94	97	98
$F_o(x)$	.579	.629	.681	.702	.726	.788	,808
S(x)	.667	.778	.806	,833	.861	.889	.917
d	.088	$\boxed{.149}$	.125	.131	.135	.101	.109

x	104	110	112
$F_o(x)$	.898	.953	.964
S(x)	.944	.972	1.000
d	.046	.019	,036

$$D = \max_x |S(X)-F_o(X)| = 0.149$$

จากตารางเมื่อ n = 36 และ α = .05 เราได้ค่าวิกฤติเป็น .221 ดังนั้นจึงปฏิเสธ Ho ไม่ได้ แสดงว่าคะแนนความถนัดนี้มีการแจกแจงปกติที่มีค่าเฉลี่ย 85 คะแนนและส่วนเบี่ยงเบนมาตรฐาน 15 คะแนน

8.4.3 แบบทดสอบสเมอร์นอฟ (Smirnov Test)

แบบทดสอบสเมอร์นอฟนี้มีวิธีการเหมือนกับแบบทดสอบโคลไมโกรอฟ แต่ใช้ทดสอบว่าสองตัวอย่างมาจากประชากรมีเหมือนกันทั้งค่ากล่าวและการกระจายหรือไม่ และแบบทดสอบนี้ฉับไว (Sensitive) ต่อความแตกต่างทุกแบบที่อาจจะมีอยู่ระหว่างสองการแจกแจง โดยปกติแบบสอบความเป็นเอกภาพก็ใช้ทดสอบได้เช่นเดียวกับแบบทดสอบนี้ แต่มักใช้กับข้อมูลนามบัญญัติ

ให้ F(x) และ G(x) เป็นฟังก์ชั่นแจกแจงสะสมที่ไม่ทราบค่าของตัวแปร x และ y แล้วสมมติฐานเกี่ยวกับฟังก์ชั่นแจกแจงทั้งสองจะเป็น

(1) $H_0: F(X) = G(X) \; ; \; \forall_x$

$H_a: F(X) \neq G(X) \; ; \; \exists_x$

(2) $H_0: F(X) \leq G(X) \; ; \; \forall_x$

$H_a: F(X) > G(X) \; ; \; \exists_x$

(3) $H_0: F(X) \geq G(X) \; ; \; \forall_x$

$H_a: F(X) < G(X) \; ; \; \exists_x$

ในการทดสอบสมมติฐานหลักต้องอาศัยตัวอย่างที่เป็นอิสระขนาด m และ n จากสองประชากรที่สนใจโดยที่ค่าสังเกตมีสเกลการวัดอย่างน้อยเป็นแบบอันดับ เมื่อได้ $S_1(x)$ และ $S_2(x)$ แทนฟังก์ชั่นแจกแจงที่สังเกตได้ของ X และ Y ตามลำดับ โดยที่

$S_1(x)$ = สัดส่วนที่ค่าสังเกต x น้อยกว่าหรือเท่ากับ x

$S_2(x)$ = สัดส่วนที่ค่าสังเกต x น้อยกว่าหรือเท่ากับ x ตัวสถิติทดสอบจึงกำหนดไว้เป็น

$$D = \max_x | S_1(x) - S_2(x) |$$

เกณฑ์ตัดสินใจสำหรับสมมติฐาน Ho จึงกำหนดไว้ว่าปฏิเสธ Ho ถ้า D มากกว่าค่าวิกฤ จากตารางพิเศษ

ถ้าทั้งสองการแจกแจงประชากรเป็นแบบต่อเนื่อง แล้วตัวสถิติทดสอบสเมอร์นอฟจะให้ค่าจริง แต่ถ้าเป็นแบบไม่ต่อเนื่องจะให้ค่าประมาณ

เมื่อตัวอย่างขนาดโต จะใช้ค่าประมาณของค่าวิกฤตในท้ายตารางพิเศษนั้น หรือจะใช้ตัวสถิติทดสอบดังนี้

$$X^2 = 4D^2\left(\frac{n_1 n_2}{n_1+n_2}\right)$$

ซึ่งมีการแจกแจงไคสแควร์ ด้วยองศาความเป็นอิสระ 2 ตัวสถิติทดสอบนี้ยังใช้ได้ดีกับตัวอย่างขนาดเล็กด้วย

ตัวอย่าง ในการศึกษาวิธีการสอนภาษาไทย 2 แบบ โดยอาศัยตัวอย่างได้ข้อมูลซึ่งเป็นคะแนนสอบภาษาไทยดังนี้

วิธี 1	206	238	224	257	230	
2	236	209	278	276	252	251

พอจะสรุปได้ไหมว่า สองตัวอย่างนี้มาจากประชากรเดียวกัน

$$H_o: F(x) = G(x); \forall x$$

$$H_a: F(x) = G(x); \exists x$$

x	y	$S_1(x)-S_2(x)$	x	y	$S_1(x)-S_2(x)$
206		$1/5-0 = 6/30$		251	$4/5-3/6 = 9/30$
	209	$1/5-1/6 = 1/30$		252	$4/5-4/6 = 4/30$
224		$2/5-1/6 = 7/30$	257		$5/5-4/6 = 10/30$
230		$3/5-1/6 = 13/30$		276	$1-5/6 = 5/30$
	236	$3/5-2/6 = 8/30$		276	$1-1 = 0$
236		$4/5-2/6 = 14/30$			

$$D = \max_x | S_1(x)-S_2(x) | = 14/30$$

จากตารางเมื่อ $m=5$ และ $n=6$ ณ ระดับนัยสำคัญ .05 ได้ค่าวิกฤต D เป็น $2/3 = 20/30$ จึงปฏิเสธ Ho ไม่ได้ แสดงว่าสองตัวอย่างนั้นมาจากประชากรที่มีฟังก์ชันไม่แตกต่างกัน