2. สถิติไร้พารามิเตอร์

1. ความรู้เบื้องต้นเกี่ยวกับสถิติไร้พารามิเตอร์
2. การทดสอบโดยใช้เครื่องหมาย
   (Sign Test)
3. การทดสอบแบบจัดอันดับเครื่องหมาย
   (Signed Rank Test)
4. แบบทดสอบผลรวมของอันดับ
   (Rank Sum Test)
5. แบบทดสอบการสุ่มโดย run
6. การทดสอบสหสัมพันธ์แบบจัดอันดับ
7. แบบฝึกหัด

1. ความรู้เบื้องต้นเกี่ยวกับสถิติไร้พารามิเตอร์

ในการทดสอบสมมติฐานเกี่ยวกับค่าพารามิเตอร์ของประชากรในบทที่ 9 (ตำรา) จะต้องมีเงื่อนไขหรือข้อสมมติเกี่ยวกับประชากรที่เราสุ่มตัวอย่างมาว่า เป็นประชากรที่มีการแจกแจงแบบปกติ หรือเป็นประชากรที่มีขนาดโตมากจนมีการแจกแจงโดยประมาณแบบโค้งปกติ อย่างไรก็ตามไม่ใช่ทุกประชากรจะมีการแจกแจงแบบปกติเสมอไป และแม้ในบทที่ 10 จะได้กล่าวถึงการทดสอบการแจกแจงของประชากร ซึ่งถ้าผลการทดสอบสรุปว่ามีการแจกแจงแบบปกติก็ตาม เราก็ยังเชื่อไม่ได้ 100% เพราะจะต้องเผื่อไว้สำหรับความผิดพลาดด้วย ดังนั้น ในบางโอกาสที่เราไม่ทราบว่ามีการแจกแจงแบบปกติหรือไม่ เราจะใช้การทดสอบในบทที่ 9 ซึ่งเป็นการทดสอบแบบมีพารามิเตอร์ หรือแบบทราบการแจกแจงไม่ได้ จะต้องใช้วิธีการทดสอบแบบไร้พารามิเตอร์ หรือแบบไม่คำนึงถึงรูปร่างการแจกแจงของประชากร (nonparametric or distribution-free tests)

ข้อดีของสถิติไร้พารามิเตอร์

1. ไม่ต้องใช้ข้อสมมุติว่าประชากรมีรูปร่างการแจกแจงแบบโค้งปกติหรือโค้งอื่นใดตามข้อกำหนด

2. เข้าใจง่ายเพราะการคำนวณง่ายกว่าแบบมีพารามิเตอร์ เช่น ไม่ต้องหาค่าเบี่ยงเบนมาตรฐาน แม้ข้อมูลจะมีค่าสูงก็ไม่ทำให้การคำนวณยาก เพราะจะจัดข้อมูลให้อยู่ในรูปอันดับ 1,2,3,...

3. ข้อมูลบางชุดอาจไม่ต้องจัดอันดับ เช่น การจำแนกผลได้เพียงแค่จำแนกว่า “ดีกว่า” อีกตัวหนึ่ง ซึ่งใช้กับผลได้ที่ไม่สามารถวัดค่าแน่นอนได้ จึงควรใช้วิธีการของสถิติไร้พารามิเตอร์

ข้อเสียของสถิติไร้พารามิเตอร์

สถิติไร้พารามิเตอร์มีข้อเสีย 2 ข้อ คือ

1. ขาดรายละเอียดข่าวสารจากข้อมูล เช่น ถ้ามีข้อมูล 5 จำนวน คือ 13.33, 76.50, 102, 113, 190 สถิติไร้พารามิเตอร์จะแทนข้อมูลแต่ละตัวด้วยอันดับ 1,2,3,4,5 ดังนั้นถ้าตัวสุดท้ายมีค่าเป็น 1900 เมื่อใช้สถิติไร้พารามิเตอร์ ก็จะถูกแทนค่าด้วย 5 เหมือนกับ 190 เพราะเป็นค่าที่โตที่สุดใน 5 จำนวน นั้น

2. บ่อยครั้งที่การทดสอบแบบไร้พารามิเตอร์จะไม่ "เพียงพอ" หรือไม่ "แน่ชัด" (efficient or sharp) เหมือนการทดสอบแบบมีพารามิเตอร์ เช่น ช่วงเชื่อมั่นแบบไร้พารามิเตอร์อาจกว้างเป็น 2 เท่าของแบบมีพารามิเตอร์ เมื่อใช้ระดับความเชื่อมั่นเดียวกัน

การทดสอบแบบไร้พารามิเตอร์แบ่งกว้าง ๆ เป็น 3 แบบคือ

1. แบบใช้เครื่องหมาย (sign test)
2. แบบจัดอันดับเครื่องหมาย (signed rank test)
3. แบบผลรวมของอันดับ (rank sum test)

2. แบบทดสอบโดยใช้เครื่องหมาย (sign test)

2.1 การทดสอบใน 1 ประชากร

บ่อยครั้งที่เราต้องการทดสอบว่า π อยู่ในระดับ π_0 หรือไม่ โดยเฉพาะในการควบคุมคุณภาพของการผลิตสินค้า (quality control) ซึ่งเราจะกำหนดระดับ π_0 ไว้ล่วงหน้า นั่นคือเรามีสมมติฐาน ดังนี้

$$H_0 : \pi = \pi_0, \quad H_a : \pi > \pi_0 \tag{1}$$

$$H_0 : \pi = \pi_0, \quad H_a : \pi < \pi_0 \tag{2}$$

$$H_a : \pi = \pi_0, \quad H_a : \pi \neq \pi_0 \tag{3}$$

โดยที่ $X_1, X_2, \ldots, X_n$ คือ ตัวอย่างสุ่มจากการแจกแจงแบบ X ซึ่งเป็นการแจกแจงแบบต่อเนื่อง ซึ่งมีฟังก์ชันความหนาแน่น $f(x)$ และ x คือค่าที่สังเกตได้จากตัวอย่างของ X ให้ L คือระดับที่กำหนดให้ และเมื่อ $x < L$ เราจะกล่าวว่าเกิดความสำเร็จ (S) แต่ถ้า $x > L$ เราจะกล่าวว่าเกิดความล้มเหลว (F) (ในการควบคุมคุณภาพสินค้า S และ F คือ จำนวนสินค้าชำรุดและสินค้าไม่ชำรุด ตามลำดับ) ดังนั้นเมื่อเราได้ข้อมูลจากตัวอย่างคือ $x_1, x_2, \ldots, x_n$ แล้ว เราจะหาค่า y ซึ่งเป็นค่าที่สังเกตได้ของตัวแปรเชิงสุ่ม Y = จำนวนความสำเร็จ ดังนั้น Y จะมีการแจกแจงแบบทวินามด้วยพารามิเตอร์ n และ π (และถ้า H_0 เป็นจริง, π จะมีค่าเท่ากับ π_0)

ดังนั้น เราจึงต้องหาเขตวิกฤตภายใต้การแจกแจงแบบทวินาม (n, π_0) ให้สอดคล้องกับสมมติฐาน ข้อ (1) - (3) ดังนี้

เขตวิกฤติของ (1) คือจะปฏิเสธ H_0 เมื่อมีเหตุผลสนับสนุนเพียงพอว่า y มีค่าโตเกินไป

เขตวิกฤตของ (2) คือจะปฏิเสธ H_0 เมื่อมีเหตุผลสนับสนุนเพียงพอว่า y มีค่าเล็กเกินไป

และเขตวิกฤตของ (3) คือจะปฏิเสธ H_0 เมื่อมีหลักฐานสนับสนุนเพียงพอว่า y มีค่าโตเกินไป หรือเล็กเกินไป

ทั้งนี้ใช้ระดับนัยสำคัญ α_0

ตัวอย่างที่ 1 หลอดทรานซิสเตอร์ชนิดหนึ่งจะถือว่ามีคุณภาพดีถ้าสามารถใช้การได้ในระดับอุณหภูมิสูงสุด 200°C ขึ้นไป ถ้าระดับอุณหภูมิสูงสุดที่ใช้การได้ของหลอดใดต่ำกว่า 200°C ถือว่าหลอดชำรุด ถ้าการตรวจสอบหลอดที่สุ่มมา 20 หลอด มีระดับอุณหภูมิสูงสุดที่ใช้งานได้ดังนี้

221,	210,	213,	$\boxed{190,}$	$\boxed{187,}$	245,	205,	$\boxed{183,}$	206,	231,
227,	$\boxed{179,}$	209,	217,	216,	235,	208,	$\boxed{195,}$	202,	220

ถ้าโรงงานพอใจเมื่อระดับหลอดชำรุดไม่เกิน 10% จงใช้ $\alpha_0 = .05$ ทดสอบโดยใช้วิธี "ปลอดการแจกแจง" (distribution-free method)

1. $H_0 : \pi = .10,\ H_a : \pi > .10$

ให้ $L = 200°$ C ถ้า $x < L$ คือหลอดชำรุด ถ้า $x > L$ คือหลอดไม่ชำรุด จากตัวอย่าง $n = 20$ มี $x < L$ อยู่ 5 จำนวน ดังนั้น $y = 5$

2. ถ้า H_0 เป็นจริง Y จะมีการแจกแจงแบบทวินาม ด้วย $n = 10, \pi = .10$ จากตารางทวินาม

$P(Y \geq 5) = 1 - P(Y \leq 4) = 1 - .95683 = .04317 < .05$

ดังนั้นเขตวิกฤตคือ $y \geq 5$

$P(Y \geq 4) = 1 - P(Y \leq 3) = 1 - .86705 = .13295 > .05$

3. จากตัวอย่าง ค่าที่สังเกตได้ของ Y คือ $y = 5$ และอยู่ในเขตวิกฤต จึงปฏิเสธ H_0 และสรุปว่าอัตราสินค้าชำรุดเกิน 10%

ตัวอย่างที่ 2 บริษัทหนึ่งใช้หลอดไฟเป็นจำนวนมากทุกปี ปัจจุบันใช้หลอดไฟที่มีอายุใช้งานโดยเฉลี่ย 1000 ชั่วโมง และมีค่าเบี่ยงเบนมาตรฐาน 100 ชั่วโมง โรงผลิตหลอดไฟอีกแห่งหนึ่งได้เสนอขายหลอดไฟซึ่งมีราคาถูกกว่า บริษัทจึงตั้งเกณฑ์การตัดสินใจว่าจะเปลี่ยนใช้ยี่ห้อใหม่ ถ้ามีหลักฐานสนับสนุนเพียงพอที่ระดับนัยสำคัญ 5% ว่ามีอายุการใช้งานโดยเฉลี่ยไม่น้อยกว่าชนิดเดิมที่ใช้อยู่ จากการสุ่มตัวอย่างหลอดชนิดใหม่มาทดลองใช้ 100 หลอด พบว่ามีอายุการใช้งานโดยเฉลี่ย 985 ชั่วโมง และสมมุติให้มีค่าเบี่ยงเบนมาตรฐานเท่ายี่ห้อเดิม บริษัทจะตัดสินใจใช้ชนิดใด ?

สำหรับข้อนี้ถ้าสมมุติว่าอายุการใช้งานของหลอดไฟมีการแจกแจงแบบปกติ เราจะทดสอบแบบมีพามิเตอร์หรือมีการแจกแจง ดังนี้

$H_0 : \mu = 1{,}000, \; H_a : \mu < 1{,}000$

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{985 - 1000}{100/\sqrt{100}} = -\frac{15}{10} = -1.5$$

ค่าเปิดตารางคือ $Z_{.05} = -1.645$

$Z_c = -1.5$ ไม่อยู่ในเขตวิกฤต จึงยอมรับ H_0 และตัดสินใจเปลี่ยนใช้ชนิดใหม่ซึ่งราคาต่ำกว่า

แต่ถ้าเราไม่สมมุติว่าอายุการใช้งานของหลอดไฟมีการแจกแจงแบบปกติ เราจะใช้การทดสอบแบบไร้พารามิเตอร์ หรือแบบไม่สนใจการแจกแจง ดังนั้นบริษัทจะตั้งค่ามัธยฐานของอายุใช้งานคือ L = 1000 ชั่วโมง ถ้าหลอดใดมีอายุการใช้งานต่ำกว่า 1000 ชั่วโมง ให้เป็นเครื่องหมายลบ ถ้าหลอดใดมีอายุการใช้งานสูงกว่า 1,000 ชั่วโมงให้เป็นเครื่องหมายบวก ถ้ามีเครื่องหมายลบ 60 หลอด เราจะใช้การทดสอบแบบใช้เครื่องหมาย (sign test) ดังนี้

1. $H_0 : \pi = .5, \; H_a : \pi > .50$ ในเมื่อ π คือสัดส่วนหลอดไฟที่มีอายุการใช้งานต่ำกว่า 1000 ชั่วโมง เราจะปฏิเสธ H_0 เมื่อ p มีค่าโตเกินไป ซึ่งแสดงว่ามัธยฐานของอายุใช้งานของหลอดชนิดใหม่ต่ำกว่ากว่า 1000 ชั่วโมง

2. จากตัวอย่าง $n = 100$, ให้ Y คือจำนวนหลอดไฟที่มีอายุการใช้งานต่ำกว่า 1000 ชั่วโมง ค่าสังเกตของ Y จากตัวอย่าง 100 หลอดคือ $y = 60$ หรือมี $p = .60$ ที่อายุการใช้งานต่ำกว่า 1000 ชั่วโมง

ถ้า H_0 เป็นจริง Y จะมีการแจกแจงแบบทวินาม ด้วยพารามิเตอร์ $n = 100$, $\pi = .5$ และเราจะปฏิเสธ H_0 ถ้ามีจำนวนเครื่องหมายลบมากเกินไป

3. จากตารางทวินาม $n = 100$, $\pi = .5$

$P(Y \geqslant 58) = 1 - P(Y \leqslant 57) = 1 - .93339 = .06661 > .05$

$P(Y \geqslant 59) = 1 - P(Y \leqslant 58) = 1 - .97156 = .02844 = \alpha_0$

เราจะปฏิเสธ H_0 ถ้า $y \geqslant 59$

4. จากตัวอย่าง $y = 60$ จึงอยู่ในเขตวิกฤต จึงปฏิเสธ H_0 นั่นคือชนิดใหม่มีคุณภาพด้อยกว่า บริษัทจึงยังคงใช้ชนิดเดิม

จะเห็นว่าผลการทดสอบแบบมีพารามิเตอร์จะมีพลังการทดสอบสูงกว่า เพราะการทดสอบแบบไร้พารามิเตอร์ ไม่คำนึงถึงค่าที่แท้จริงของข้อมูล ดูแต่เครื่องหมายเท่านั้น จึงขาดรายละเอียดที่สำคัญ

2.2 การทดสอบใน 2 ประชากรที่ไม่เป็นอิสระกัน (ข้อมูลแบบจับคู่)

สำหรับข้อมูลแบบจับคู่ เมื่อเราหาผลต่าง D_i ของคู่ต่าง ๆ n คู่แล้ว ให้ M คือค่ามัธยฐานของ D_i ให้นำ D_i ทุกตัวเทียบกับ M_D ถ้าตัวใดน้อยกว่าให้มีเครื่องหมายลบ ถ้าตัวใดมากกว่า M_D ให้มีเครื่องหมายบวก และกำหนดให้ตัวแปรเชิงสุ่ม Y คือจำนวนเครื่องหมายลบของ D_i และจะมีสมมติฐาน ดังนี้

$$H_0 : \pi = .5, \quad H_a : \pi > .5 \tag{4}$$

$$H_0 : \pi = .5, \quad H_a : \pi < .5 \tag{5}$$

$$H_0 : \pi = .5, \quad H_a : \pi \neq .5 \tag{6}$$

π คือสัดส่วนของเครื่องหมายลบ

ตัวอย่างที่ 3 ให้พนักงานเข้าอบรมเกี่ยวกับการขาย 10 คน ภายหลังอบรมแล้วได้เทียบจำนวนขายต่อเดือนก่อนและหลังการอบรม ดังนี้

หลังอบรม	135	200	160	182	147	200	172	186	194	141
ก่อนอบรม	127	195	162	170	143	205	168	175	197	136
D_i	8	5	−2	12	4	−5	4	11	−3	5

ทดสอบว่าโปรแกรมอบรมทำให้จำนวนขายเปลี่ยนแปลงหรือไม่

ให้ Y คือจำนวนเครื่องหมายลบ ถ้าโปรแกรมไม่มีประโยชน์ เราควรได้จำนวนเครื่องหมายบวกและลบเท่า ๆ กัน นั่นคือ Y มีการแจกแจงแบบทวินามด้วย $n = 10$ และ $\pi = .5$ แต่ถ้าโปรแกรมมีประโยชน์ เราย่อมคาดว่าจะได้เครื่องหมายลบ **น้อย** กว่าเครื่องหมายบวก

1. $H_0 : \pi = .50$ (โปรแกรมไม่ช่วยเพิ่มยอดขาย)
2. $H_a : \pi < .50$ (โปรแกรมช่วยเพิ่มยอดขาย)
3. $\alpha = .025$
4. จากตารางทวินาม $n = 10,\ \pi = .50$

   $P(Y \leqslant 1) = .01074 < .025 \leftarrow$ เขตวิกฤต

   $P(Y \leqslant 2) = .05469 > .025$

   ดังนั้นเราจะปฏิเสธ H_0 เมื่อ $y \leqslant 1$

5. จากตัวอย่าง ค่าที่สังเกตได้คือ $y = 3$ ซึ่งไม่อยู่ในเขตวิกฤต จึงยอมรับ H_0 และสรุปว่าโปรแกรมฝึกอบรมไม่ช่วยเพิ่มยอดขาย

แบบฝึกหัด

2.1 โรงงานผลิตน้ำมะเขือเทศกระป๋องขนาดบรรจุ 46 ออนซ์ โรงงานไม่ต้องการให้บรรจุน้ำหนักถัวเฉลี่ยมากกว่าหรือน้อยกว่า 46 ออนซ์ ถ้าจากตัวอย่างที่สุ่มมา 25 กระป๋องมีน้ำหนักดังนี้

45.63, 45.82, 45.77, 45.80, 46.21, 46.07, 46.16, 45.91
45.87, 46.03, 46.16, 45.92, 45.76, 45.89, 45.85, 46.16
45.80, 45.79, 45.86, 46.10, 45.88, 45.91, 45.87, 45.96
45.99

จงใช้การทดสอบแบบเครื่องหมาย ทดสอบด้วย $\alpha_0 = .05$ และสรุปผลการทดสอบ
(CR:$y \leqslant 7$; $y \geqslant 18$, ปฏิเสธ H_0, สรุปว่าน้ำหนักเฉลี่ยต่ำกว่า 46, ค่าสังเกตได้ของ $y = 18$)

2.2 นักจิตวิทยาต้องการหนูที่เจริญเติบโตเร็วกว่าพันธุ์มาตรฐาน ถ้าผลการทดลองหนูพันธุ์ใหม่พบว่ามี 75% ที่ใช้เวลาเติบโตเต็มที่ไม่เกิน 240 วัน เขาจะเปลี่ยนมาใช้พันธุ์ใหม่ ถ้าผลการทดลองเลี้ยงหนูพันธุ์ใหม่ 50 ตัว มี 36 ตัวที่โตเต็มที่ในเวลาน้อยกว่า 240 วัน ส่วนที่เหลืออีก 14 ตัวใช้เวลาโตเต็มที่เกิน 240 วัน จงใช้ $\alpha_0 = .10$ ทดสอบ
($H_0 : \pi = .25$, $H_a : \pi > .25$, y คือ จำนวนหนูที่ใช้เวลาเกิน 240 วัน CR : $y \geqslant 17$ ด้วย $\alpha = .09831$, ค่าสังเกต $y = 14$ จึงไม่ปฏิเสธ H_0 สรุปว่าเปลี่ยนใช้พันธุ์ใหม่ หรือ $Z_C = 0.49$, ควรเปลี่ยนใช้หนูพันธุ์ใหม่)

2.3 อาจารย์สอนวิชาสถิติจะเพิ่มชั่วโมงสอนเป็นสัปดาห์ละ 2 ครั้ง ถ้าพิสูจน์ได้ว่าค่ามัธยฐานของคะแนนสอบสูงเกิน 60 คะแนน ถ้าทดลองสอนนักเรียนกลุ่มหนึ่งจำนวน 100 คน สัปดาห์ละ 2 ครั้ง มี 38 คน ที่ได้คะแนนต่ำกว่า 60 คะแนน และ 62 คนได้คะแนนสูงกว่า 60 คะแนน จงใช้แบบทดสอบเครื่องหมายทดสอบโดยใช้ระดับนัยสำคัญ 5%
($Z_C = 2.4$, ควรเพิ่มชั่วโมงสอน)

2.4 สำนักงานแห่งหนึ่งต้องการซื้อหลอดไฟขนาด 100 วัตต์ ซึ่งเมื่อเปิดใช้ใน 10 นาที จะกินไฟโดยเฉลี่ยน้อยกว่า 50 ยูนิต เมื่อทดลอง 50 หลอดมี 35 หลอดที่กินไฟน้อยกว่า 50 ยูนิต อีก 15 หลอดกินไฟเกิน 50 ยูนิต จงใช้ $\alpha = .05$ ทดสอบโดยใช้เครื่องหมาย

2.5 โรงงานผลิตยาทาแก้ผื่นคันอ้างว่ายาที่ผลิตใหม่มีคุณภาพช่วยทำให้หายคันเร็วกว่าชนิดเดิมที่มีขายในท้องตลาด ถ้าใช้คนที่เป็นผื่นคัน 50 คน ทดลองโดยแบ่งบริเวณที่มีผื่นคันเป็น 2 บริเวณ บริเวณหนึ่งใช้ทาด้วยยาชนิดเดิม และอีกบริเวณหนึ่งทาด้วยยาชนิดใหม่ ให้ $d_1, d_2, \ldots, d_{50}$ คือผลต่างของเวลาที่ใช้จนรู้สึกหายคัน โดยนำยาใหม่ไปลบออกจากยาเดิม ถ้ามี 14 คนที่มีเครื่องหมายลบ จะสรุปผลที่ระดับนัยสำคัญ 1% ว่าอย่างไร

2.6 ข้อมูลต่อไปนี้คืออัตราสินค้าชำรุดก่อนและหลังการปรับปรุงระบบงาน จงทดสอบว่าการปรับปรุงระบบงานช่วยลดอัตราชำรุด โดยใช้ $\alpha = .05$

ก่อน	8	7	6	9	11	10	8	6	5	8	9	10
หลัง	6	5	8	6	9	8	10	7	4	6	9	10

ตัวอย่างที่ 4 การประมาณการทดสอบโดยใช้เครื่องหมายด้วยโค้งปกติ

มีเครื่องจักรผลิตสกรูขนาดเดียวกัน 2 เครื่อง ให้ A คือจำนวนชำรุดต่อชั่วโมงของเครื่องจักรที่ I และ B คือจำนวนชำรุดต่อชั่วโมงของเครื่องที่ II จากการสังเกตการทำงานของเครื่องจักรทั้งสองเป็นเวลา 18 ชั่วโมง ได้บันทึกจำนวนชำรุดต่อชั่วโมงไว้ จงใช้การทดสอบแบบเครื่องหมายตรวจสอบความแตกต่างของการทำงานของเครื่องทั้ง 2 โดยใช้ระดับนัยสำคัญ 5%

ชั่วโมง	A จำนวนชำรุด ของเครื่องที่ I	B จำนวนชำรุด ของเครื่องที่ II	เครื่องหมาย A – B
1	12	10	+
2	8	12	
3	15	12	+
4	8	7	+
5	16	16	0
6	20	22	
7	18	20	
8	19	10	+
9	15	12	+
10	27	25	+
11	11	16	
12	24	21	+
13	17	17	0
14	19	15	+
15	13	17	–
16	9	10	
17	11	10	+
18	26	7	+

เนื่องจากมี 0 อยู่ 2 จำนวน จึงไม่มีเครื่องหมาย ขนาดตัวอย่างจึงเหลือ 16 และมีเครื่องหมายบวก 10 จำนวน เครื่องหมายลบ 6 จำนวน ให้ Y คือจำนวนเครื่องหมาย + ถ้าเครื่องทั้งสองทำงานด้วยประสิทธิภาพเท่ากันก็ควรจะได้อัตราชำรุดเท่ากัน นั่นคือ อัตราการเกิดเครื่องหมายบวก และลบเท่ากัน คือ $\pi = .5$ ดังนั้นจึงตั้งสมมติฐาน

$H_0 : \pi = .5, \quad H_a : \pi \neq .5$

การแจกแจงที่แท้จริงของ Y เมื่อ H_0 เป็นจริงคือการแจกแจงแบบทวินาม ที่มี $n = 16$, $\pi = .5$ แต่เราจะลองทำแบบใช้การแจกแจงปกติประมาณ เพราะทั้ง $n\pi > 5$ และ $n(1 - \pi) > 5$

โดยให้ $\mu = n\pi = 16(.5) = 8$

$$\sigma = \sqrt{n\pi(1 - \pi)} = \sqrt{16(.5)(.5)} = 2$$

เราอาจเปลี่ยนสมมติฐานใหม่เป็น

$H_0 : \mu = 8,\ H_a : \mu \neq 8$ ก็ได้

$$Z = \frac{X - n\pi_0}{\sqrt{n\pi_0(1 - \pi_0)}} = \frac{10 - 8}{2} = 1$$

เขตวิกฤตคือ $|Z| > 1.96$

แต่ $Z_C = 1$ ซึ่งไม่อยู่ในเขตวิกฤต เราจึงยอมรับ H_0 ว่า $\pi = .5$ นั่นคือ อัตราชำรุดของทั้ง 2 เครื่องไม่ต่างกันอย่างมีนัยสำคัญ

3. การทดสอบแบบจัดอันดับเครื่องหมาย (SIGNED RANK TEST)

3.1 แบบทดสอบของวิลค็อกซ์สัน (WILCOXON SIGNED RANK TEST) (สำหรับข้อมูลแบบจับคู่)

การทดสอบโดยใช้เครื่องหมายมีจุดอ่อนที่ไม่คำนึงถึงขนาดความแตกต่าง สนใจแต่เครื่องหมายเพียงอย่างเดียว Frank Wilcoxon ได้สร้างแบบทดสอบโดยพิจารณาขนาดความแตกต่างของแต่ละคู่ โดยการจัดอันดับค่า absolute value ของ D_i (ไม่คิดเครื่องหมาย) จากน้อยที่สุดไปหาค่ามากที่สุด โดยให้ D_i ที่มีค่าน้อยที่สุดเป็นอันดับ 1 ค่าน้อยถัดไปเป็นอันดับ 2 เรื่อย ๆ ไป ส่วนค่า D_i ที่เป็น 0 ให้ตัดทิ้งไป เนื่องจากการจัดอันดับไม่คำนึงถึงเครื่องหมาย ดังนั้นผลต่าง -1 หรือ +1 จะอยู่ในอันดับเดียวกัน, -2 และ +2 จะอยู่ในอันดับเดียวกัน เช่นนี้เรื่อยไป เมื่อจัดอันดับครบทุกจำนวนแล้ว จึงใส่เครื่องหมายเดิมให้กับอันดับ แล้วหาผลรวมของอันดับโดยแยก

เป็นผลรวมของเครื่องหมายบวก และเครื่องหมายลบ และจะใช้ผลรวมของอันดับนี้สำหรับทดสอบในขั้นต่อไป

ให้ T คือผลรวมของอันดับ เมื่อ n มีค่า 8 ขึ้นไป T จะมีการแจกแจงโดยประมาณแบบปกติ

ด้วยค่าเฉลี่ย $E(T) = \frac{n(n+1)}{4}$

และค่าเบี่ยงเบนมาตรฐาน $\sigma_T = \sqrt{\frac{n(n+1)(2n+1)}{24}}$

และ $Z = \frac{T - E(T)}{\sigma_T}$ จะมีการแจกแจงแบบปกติ

หมายเหตุ จะให้ T แทนผลรวมของอันดับของเครื่องหมายลบ หรือบวกก็ได้

ตัวอย่างที่ 5 เลือกหมูมาครอกละ 2 ตัว จำนวน 10 ครอก แบ่งเลี้ยงด้วยสูตรอาหาร A และ B จนครบเวลาการทดลอง จึงเอาน้ำหนักหมูจากครอกเดียวกัน (แต่กินอาหารต่างสูตรกัน) มาเปรียบเทียบ ให้ d_i คือผลต่างของสูตร A-B ให้ทดสอบว่าสูตร A ด้อยกว่าสูตร B ที่ $\alpha = .05$

คู่	A	B	$d_i = A - B$	อันดับ	เครื่องหมาย	
					ลบ	บวก
1	17	14	+3	4		4
2	17	21	-4	5	-5	
3	21	36	-15	10	-10	
4	18	20	-2	2.5	-2.5	
5	17	24	-7	7	-7	
6	14	12	+2	2.5		2.5
7	19	28	-9	9	-9	
8	17	16	+1	1		1
9	16	21	-5	6	-6	
10	12	20	-8	8	-8	
ผลรวมของอันดับ = T					-47.5	7.5

ให้ T = 47.5 (ไม่คิดเครื่องหมาย)

$$E(T) = \frac{n(n+1)}{4} = \frac{10(11)}{4} = 27.5$$

$$\sigma(T) = \sqrt{\frac{n(n+1)(2n+1)}{24}} = \sqrt{\frac{10(11)(21)}{24}}$$

$$= \sqrt{\frac{2310}{24}} = \sqrt{96.25} = 9.81$$

H_0 : A และ B มีคุณภาพไม่ต่างกัน

H_a : A มีคุณภาพด้อยกว่า B

ถ้า H_0 เป็นจริงเราย่อมคาดว่าจะได้ผลรวมอันดับของเครื่องหมายบวกและลบใกล้เคียงกัน แต่ถ้า H_a เป็นจริง เนื่องจาก d_i = A - B d_i จะมีค่าติดลบหลายตัว และจะทำให้ผลรวมอันดับของเครื่องหมายลบมากเกินไปจนอยู่ในเขตวิกฤต คือ $Z_C > Z_{.05} = 1.645$

$$Z = \frac{T - E(T)}{\sigma(T)} = \frac{47.5 - 27.5}{9.81}$$

$$= \frac{20}{9.81} = 2.04$$

2.04 > 1.645 จึงปฏิเสธ H_0 และสรุปว่าสูตร A ด้อยกว่า B

ในกรณีที่ $n < 8$ ยังไม่ควรใช้ normal approximation จึงต้องใช้ตารางเฉพาะซึ่ง Wilcoxon สร้างไว้

ตาราง B_{20} คือตารางของ Wilcoxon ค่าในตารางคือ $P(T \leqslant t)$ เมื่อ n = 5 ถึง n = 20 หากต้องการเขตวิกฤต 2 ด้าน ต้องใช้ความสัมพันธ์

$$P(T \geqslant t) = P[T \leqslant \frac{n(n+1)}{2} - t]$$

นั่นคือ $$t_2 = \frac{n(n+1)}{2} - t_1$$

จากตัวอย่างที่ 5 ถ้าเราต้องการทดสอบแบบ 2 ด้าน คือต้องการทดสอบว่าสูตรอาหารทั้ง A และ B มีคุณภาพแตกต่างกันหรือไม่ โดยใช้ α = .05 เราจะต้องแบ่งเขตวิกฤตไว้ 2 ด้าน ด้านละ $\alpha/2$ = .025

จากตาราง B_{20} เมื่อ $n = 10$

$P(T \leqslant 8) = .0244 < .025 \leftarrow$ เขตวิกฤตคือ $T \leqslant 8$

$P(T \leqslant 9) = .0322 > ,025$

ให้ $\quad t_1 = 8$

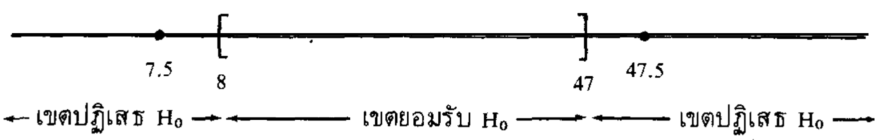
$$t_2 = \frac{n(n+1)}{2} - t_1$$

$$= \frac{10(10+1)}{2} - 8$$

$$= 47$$

นั่นคือ $P(T \geqslant 47) = .0244 \leftarrow$ เขตวิกฤตคือ $t \geqslant 47$

ดังนั้นเขตวิกฤตคือ $t \leqslant 8$ และ $t \geqslant 47$



ผลการคำนวณได้ $T = 47.5$ และ $T = 7.5$

จะเห็นว่าทั้ง 2 จำนวนอยู่นอกเขตวิกฤต จึงปฏิเสธ H_0 ด้วยระดับนัยสำคัญ

$.0244 + .0244 = .0488$

3.2 การใช้แบบทดสอบของ Wilcoxon ใน 1 ประชากร

ดังได้กล่าวแล้วว่าการทดสอบโดยใช้เครื่องหมายไม่พิจารณาค่าแท้จริงของข้อมูลเดิม จึงขาดข่าวสารที่สำคัญ ดังเช่นในแบบฝึกหัดที่ 2.1 เรื่องการทดสอบคุณภาพของน้ำมะเขือเทศกระป๋องขนาด 46 ออนซ์ ด้วย $\alpha = .05$ สมมุติสุ่มมา 10 กระป๋อง

ให้ x_i คือ น้ำหนักของ 10 กระป๋องนั้น

d_i คือ $x_i - 46$

ตัวอย่างที่ 6

x_i	d_i	อันดับ	เครื่องหมาย	
			ลบ	บวก
45.63	−.37	10	−10	
45.82	−.18	6	-6	
45.77	−.23	9	-9	
45.80	−.20	7	-7	
46.21	+.21	8		8
46.07	+.07	2		2
46.16	+.16	5		5
45.91	−.09	3		
45.87	−.13	4	-3	
46.03	+.03	.1	-4	
				1
		T	-39	16

เขตวิกฤตได้จากตาราง B_{20} โดยแบ่ง 2 ด้าน ๆ ละ .025

$P(T \leqslant 8) = .0244 < .025 \leftarrow$ เขตวิกฤต

$P(T \leqslant 9) = .0322 > .025$

ให้ $t_1 = 8$

$$t_2 = \frac{n(n+1)}{2} - t_1$$

$$= \frac{10(11)}{2} - 8$$

$$= 47$$

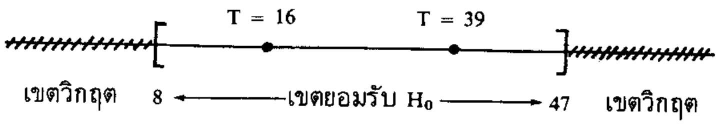
นั่นคือ $P(T \geqslant 47) = .0244$

ดังนั้นเขตวิกฤตคือ $t \leqslant 8$ และ $t \geqslant 47$

ด้วยระดับนัยสำคัญ $.0244 + .0244 = .0488$

ค่าที่สังเกตจากตัวอย่างคือ $T = 16$ (ให้ใช้ค่าเล็กของ T) ไม่อยู่ในเขตวิกฤต จึงไม่ปฏิเสธ H_0 แสดงว่าน้ำหนักบรรจุได้มาตรฐานตามที่ได้กำหนดไว้

หมายเหตุ กรณีที่ใช้เขตวิกฤต 2 ด้าน จะใช้ค่า T ตัวใดก็ได้ เพราะทั้งคู่จะให้ข้อสรุปเหมือนกัน เช่น จากตัวอย่างจะเห็นว่าทั้ง $T = 39$ และ $T = 16$ อยู่นอกเขตวิกฤต



สมมุติว่าเป็นการทดสอบด้านเดียวโดยใช้ $\alpha = .025$

เขตวิกฤตคือ $t \leqslant 8$ ด้วย $\alpha = .0244$

ค่าสังเกตของ T คือ 39 และ 16 ให้เลือกค่าน้อยสำหรับเปรียบเทียบ คือ $T = 16$ ซึ่งอยู่นอกเขตวิกฤต

4. การทดสอบแบบผลรวมของอันดับ (Rank Sum Tests)

4.1 การทดสอบของ Wilcoxon ใน 2 ประชากรที่เป็นอิสระกัน (Wilcoxon Two-Sample Test)

เมื่อต้องการทดสอบ $H_0 : \mu_1 = \mu_2$ เราจะใช้ Z-test เมื่อทราบว่าประชากรมีการแจกแจงแบบปกติ หรือเมื่อขนาดตัวอย่างโตมาก เราจะใช้ t-test สำหรับตัวอย่างขนาดเล็กที่มีการแจกแจงแบบปกติ เรียกว่าเป็นการทดสอบโดยมีพารามิเตอร์ และเงื่อนไขการแจกแจง เราจะทำอย่างไร เมื่อเรามีตัวอย่างขนาดเล็กซึ่งไม่ทราบการแจกแจง หรือทราบว่ามีลักษณะการแจกแจงแบบไม่สมมาตรหรือการแจกแจงที่ไม่ใช่แบบปกติ เราจึงต้องใช้การทดสอบแบบไร้พารามิเตอร์ หรือแบบไม่คำนึงการแจกแจง วิธีการของ Wilcoxon คือ การนำข้อมูลทั้ง 2 กลุ่มมารวมกันเพื่อจัด

อันดับ ให้ n_1 คือขนาดตัวอย่างของกลุ่มที่มีจำนวนข้อมูลน้อย n_2 คือขนาดตัวอย่างของกลุ่มที่มีข้อมูลมากกว่า ให้ W_1, W_2 คือผลรวมของอันดับของกลุ่มขนาดตัวอย่าง n_1, n_2 ตามลำดับ ถ้า $n_1 = n_2$ ให้ W_1 คือผลรวมอันดับของกลุ่มที่มีค่าเฉลี่ยต่ำ ในการจัดอันดับข้อมูลต้องพิจารณาสมมติฐานรอง ดังนี้

$H_0 : \mu_1 = \mu_2$, $H_a : \mu_1 < \mu_2$ เราจะปฏิเสธ H_0 เมื่อ $\overline{X}_1$ เป็นค่าเล็กเกินไป เราจึง **จัดอันดับจากน้อยไปหามาก** คือค่าต่ำสุดเป็นอันดับ 1 ค่าสูงสุดเป็นอันดับสุดท้าย เราจะปฏิเสธเมื่อ W_1 เป็นค่าเล็กเกินไป

$H_0 : \mu_1 = \mu_2$, $H_a : \mu_1 > \mu_2$ เราจะปฏิเสธ H_0 เมื่อ $\overline{X}_1$ เป็นค่าโตเกินไป จึง **จัดอันดับจากมากไปหาน้อย** คือให้ค่าสูงสุดเป็นอันดับ 1 ค่าต่ำสุดเป็นอันดับสุดท้าย และเราจะปฏิเสธเมื่อ W_1 เป็นค่าน้อยเกินไป

$H_0 : \mu_1 = \mu_2$, $H_a : \mu_1 \neq \mu_2$ เราจะปฏิเสธ H_0 เมื่อ $\overline{X}_1$ มีค่าที่เล็กเกินไป หรือโตเกินไป แต่เขตวิกฤตในตารางให้แต่ด้านน้อยด้านเดียว จึงต้องเปรียบเทียบ $\overline{X}_1$ และ $\overline{X}_2$ ถ้า $\overline{X}_1 > \overline{X}_2$ ให้ **จัดอันดับจากค่าสูงสุดเป็นอันดับ 1** และให้ค่าต่ำสุดเป็นอันดับสุดท้าย เพื่อว่าจะปฏิเสธ H_0 เมื่อ W_1 เป็นค่าที่เล็กเกินไป และถ้า $\overline{X}_1 < \overline{X}_2$ **ก็จัดอันดับจากน้อยไปหามาก** และปฏิเสธ H_0 เมื่อ W_1 เป็นค่าน้อยเกินไป

สรุปคือไม่ว่าจะมีสมมติฐานรองแบบใด เราจะปฏิเสธเมื่อ W_1 เป็นค่าเล็กเกินไป โดยคำนวณค่าสถิติ U

$$U = W_1 - \frac{n_1(n_1+1)}{2}$$

ให้นำค่า U เทียบกับค่าในตาราง Wilcoxon คือตารางที่ 13 และ 14 และเมื่อ $n_2 > 20$ ให้ ประมาณโดยการแจกแจงแบบปกติ ด้วย

$$\mu_U = \frac{n_1 n_2}{2}$$

$$\sigma_U = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

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

ตัวอย่างที่ 7 ต้องการทราบว่าวัคซีนชนิดใหม่ช่วยรักษาโรคลิวคีเมียหรือไม่ โดยนำหนูที่เป็นโรคลิวคีเมียระดับสูงมา 9 ตัว ฉีดวัคซีนให้กับหนู 5 ตัว อีก 4 ตัวไม่ฉีด แล้วนับอายุรอดเป็นปี ได้ผลดังนี้

ฉีดยา	2.1 5.3 1.4 4.6 0.9	$\overline{X}_2 = 2.86$
ไม่ฉีดยา	1.9 0.5 2.8 3.1	$\overline{X}_1 = 2.075$

ให้ทดสอบว่าวัคซีนช่วยเพิ่มอายุหนูที่เป็นโรคหรือไม่ ด้วย $\alpha = .05$

$n_1 = 4, n_2 = 5, \overline{X}_1 = 2.075 \ \overline{X}_2 = 2.86$

$H_0 : \mu_1 = \mu_2$

$H_a : \mu_1 < \mu_2$

นำข้อมูลทั้งหมดมาเรียงจากน้อยไปหามากเพื่อจัดอันดับ

ข้อมูลเดิม	0.5	0.9	1.4	1.9	2.1	2.8	3.1	4.6	5.3
อันดับ	$\underline{1}$	2	3	$\underline{4}$	5	$\underline{6}$	$\underline{7}$	8	9

ที่ขีดเส้นใต้คืออันดับของข้อมูลในกลุ่มที่ 1

$$\begin{aligned} W_1 &= 1+4+6+7 = 18 \\ U &= W_1 - n_1(n_1+1)/2 \\ &= 18 - 4(5)/2 \\ &= 8 \end{aligned}$$

จากตารางที่ 13 เมื่อ $n_2 = 5, n_1 = 4$

$P(U \leqslant 2) = .032 < .05 \leftarrow$ เขตวิกฤต

$P(U \leqslant 3) = .056 > .05$

ดังนั้นจะปฏิเสธ H_0 เมื่อ $U \leqslant 2$

แต่ค่าคำนวณ $U = 8$ จึงไม่อยู่ในเขตวิกฤต

จึงยังไม่ปฏิเสธ H_0 และสรุปว่า วัคซีนไม่ช่วยทำให้หนูมีอายุยืนขึ้นกว่าไม่ฉีด

ตัวอย่างที่ 8 ต้องการเปรียบเทียบปริมาณสารนิโคตินในบุหรี่ 2 ชนิด ซึ่งสุ่มมาแล้ววัดได้เป็น มิลลิกรัม ดังนี้

A	22.1	24.0	26.3	25.4	24.8	23.7	26.1	23.3			$\overline{X}_1$ = 24.5
B	24.1	20.6	23.1	22.5	24.0	26.2	21.6	22.2	21.9	25.4	$\overline{X}_2$ = 23.2

จงทดสอบว่าสารนิโคตินโดยเฉลี่ยมีความแตกต่างกันที่ α = .05 หรือไม่ ?

$n_1 = 8$, $n_2 = 10$, $\overline{X}_1 = 24.5$, $\overline{X}_2 = 23.2$

$H_0 : \mu_1 = \mu_2$

$H_a : \mu_1 \neq \mu_2$

เนื่องจาก $\overline{X}_1 > \overline{X}_2$ จึงต้องเรียงข้อมูลจากมากมาหาน้อย ดังนี้

ข้อมูลเดิม	26.3	26.2	26.1	25.4	25.4	24.8	24.1	24.0	24.0
อันดับ	<u>1</u>	2	<u>3</u>	<u>4.5</u>	4.5	<u>6</u>	7	8.5	<u>8.5</u>

ข้อมูลเดิม	23.7	23.3	23.1	22.5	22.2	22.1	21.9	21.6	20.6
อันดับ	<u>10</u>	<u>11</u>	12	13	14	<u>15</u>	16	17	18

ค่าที่ขีดเส้นใต้คืออันดับของกลุ่มที่มีตัวอย่างน้อย คือ $n_1 = 8$

$$W_1 = 1+3+4.5+6+8.5+10+11+15 = 59$$

$$U = 59 - (8)(9)/2 = 23$$

จากตารางที่ 14 เมื่อ $n_2 = 10$, $n_1 = 8$, $\alpha = .05$ (ทดสอบ 2 ด้าน)

ค่าวิกฤตของ U คือ 17 นั่นคือจะปฏิเสธ H_0 เมื่อ $U \leqslant 17$

แต่ค่าคำนวณจากตัวอย่างของ U = 23 จึงไม่อยู่ในเขตวิกฤต

จึงไม่ปฏิเสธ H_0 และสรุปว่าปริมาณสารนิโคตินโดยเฉลี่ยในบุหรี่ทั้ง 2 ชนิดไม่ต่างกันอย่างมีนัยสำคัญ

4.2 แบบทดสอบของ Mann-Whitney (Mann Whitney U Test)

แบบทดสอบของวิทนี่มีวิธีการคล้ายคลึงกับแบบทดสอบของ Wilcoxon คือนำข้อมูลทั้งหมดมาเรียงจากน้อยไปหามากเพื่อจัดอันดับ **ให้อันดับ 1 กับข้อมูลที่น้อยที่สุด** เมื่อจัดเสร็จแล้วก็หาผลรวมของอันดับของกลุ่มที่ 1 คือ R_1 แล้วหาเขตวิกฤตของ R_1 โดยหาขีดจำกัดบนและล่าง ดังนี้

$$a = \text{ขีดจำกัดล่าง} = \frac{n_1(n_1+1)}{2}$$

$$b = \text{ขีดจำกัดบน} = n_1n_2 + \frac{n_1(n_1+1)}{2}$$

ถ้า R_1 อยู่ในช่วง a และ b จะยอมรับ H_0 ว่าประชากรทั้ง 2 มีค่าเฉลี่ยไม่ต่างกัน แต่ถ้า R_1 มีค่าน้อยเกินไปคือน้อยกว่า a ต้องปฏิเสธ แสดงว่า เรายอมรับ $H_a : \mu_1 < \mu_2$ และถ้า R_1 มีค่ามากเกินไปจน $R_1 > b$ เราจะยอมรับ $H_a : \mu_1 > \mu_2$

ตัวอย่างที่ 9

$n_1 = 3$: 4 17 12

$n_2 = 2$: 14 10

จัดเรียงลำดับ	:	4 10 12 14 17
อันดับ	:	$\underline{1}$ 2 $\underline{3}$ 4 $\underline{5}$

ที่ขีดเส้นใต้คือกลุ่มที่ 1 ซึ่งมี $n_1 = 3$

$$R_1 = 1+3+5 = 9$$

เขต **ยอมรับ** H_0 คือ

$$a = \frac{n_1(n_1+1)}{2} = \frac{3(4)}{2} = 6$$

$$b = n_1n_2 + \frac{n_1(n_1+1)}{2} = 3(2)+6 = 12$$

นั่นคือจะยอมรับ $H_0 : \mu_1 = \mu_2$ ถ้า R_1 อยู่ระหว่าง 6 ถึง 12

แต่ถ้า $R_1 < 6$ จะยอมรับ $H_a : \mu_1 < \mu_2$

ถ้า $R_1 > 12$ จะยอมรับ $H_a : \mu_1 > \mu_2$

แต่จากตัวอย่าง คำนวณได้ $R_1 = 9$ อยู่ในเขตยอมรับ H_0

จึงสรุปว่า $\mu_1 = \mu_2$

แต่ถ้าเราต้องการทดสอบนัยสำคัญ โดยระบุค่า α เมื่อ n_1, n_2 มีค่าเป็น 10 ขึ้นไป เราจะคำนวณค่าสถิติ U และใช้ normal approximation

$$U = n_1n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$\mu_U = \frac{n_1n_2}{2}$$

$$\sigma_U = \sqrt{\frac{n_1n_2(n_1+n_2+1)}{12}}$$

ตัวอย่าง

$$n_1 = 15,\ n_2 = 15$$

$$R_1 = 247,\ R_2 = 218$$

$$\begin{aligned} U &= 15(15) + \frac{15(16)}{2} - 247 \\ &= 225 + 120 - 247 \\ &= 98 \leftarrow \text{ค่าสถิติ U} \end{aligned}$$

$$\mu_U = n_1n_2/2 = 15(15)/2 = 112.5$$

$$\begin{aligned} \sigma_U &= \sqrt{\frac{15(15)(15+15+1)}{12}} \\ &= \sqrt{\frac{6975}{12}} = \sqrt{581.25} = 24.1 \end{aligned}$$

$H_0 : \mu_1 = \mu_2,\ H_a : \mu_1 \neq \mu_2$

ให้ $\alpha = .15$

จากตาราง $Z_{.075} = 1.44$

นั่นคือจะปฏิเสธ H_0 เมื่อ $Z_C > 1.44$ หรือ $Z < -1.44$

$$\begin{aligned} Z &= \frac{U - \mu_U}{\sigma_U} = \frac{98 - 112.5}{24.1} \\ &= -0.60 \end{aligned}$$

$Z_C = -0.60$ ไม่อยู่ในเขตวิกฤต จึงสรุปว่า $\mu_1 = \mu_2$

หมายเหตุ เราจะใช้ R_1 หรือ R_2 ก็ได้ในการคำนวณค่าสถิติ U
ถ้าใช้ R_2 ต้องเปลี่ยน n_1 เป็น n_2 คือ

$$
\begin{aligned}
U &= n_1n_2 + \frac{n_2(n_2+1)}{2} - R_2 \\
&= 15(15) + \frac{15(16)}{2} - 218 \\
&= 225 + 120 - 218 \\
&= 127 \\
Z &= (U - \mu_U)/\sigma_U \\
&= (127 - 112.5)/24.1 \\
&= 0.60
\end{aligned}
$$

Z_C = 0.60 ไม่อยู่ในเขตวิกฤต จึงสรุปว่า $\mu_1 = \mu_2$ เช่นกัน

แบบฝึกหัด

2.7 A และ B เป็นระบบการผลิตขดลวดไฟฟ้า ให้ X คือจำนวนขดลวดชำรุดต่อชั่วโมงของระบบ A และ Y คือจำนวนชำรุดต่อชั่วโมงของระบบ B ในเมื่อ A และ B ผลิตในชั่วโมงเดียวกัน จงประมาณด้วยการทดสอบแบบโค้งปกติว่าระบบทั้ง 2 ทำงานแตกต่างกันหรือไม่ เมื่อใช้ $\alpha = .01$

ชั่วโมง	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
X	20	10	61	18	25	28	22	20	18	15	23	25	18	5	30	31
Y	21	25	64	25	21	22	30	29	16	10	18	35	34	25	19	13

2.8 การทดลองเช่นเดียวกับตัวอย่างที่ 5 แต่ใช้หมู 12 คู่ แบ่งเลี้ยงด้วยสูตร A และ B ได้ผลต่างของน้ำหนัก (A – B) ของแต่ละคู่ ดังนี้

คู่	1	2	3	4	5	6	7	8	9	10	11	12
(A – B)	+3	–4	–15	–2	–7	+2	0	–9	0	+1	–5	–8

ก) ให้ทดสอบว่าสูตร A มีคุณภาพด้อยกว่าสูตร B อย่างมีนัยสำคัญที่ $\alpha = 0.05469$ (ถ้าใช้การทดสอบแบบเครื่องหมาย)

ข) ใช้ normal approximation ใช้ $\alpha = .05$

ค) ใช้การทดสอบของวิลค็อกซ์สัน โดยใช้ $\alpha = .01$

2.9 โรงงานผลิตยางรถยนต์กำลังพิจารณาว่า ควรจะเปลี่ยนใช้ระบบการผลิตใหม่หรือไม่ โรงงานจึงตั้งเกณฑ์ว่า ถ้ามีหลักฐานสนับสนุนว่าระบบใหม่ทำให้ยางที่ผลิตได้มีอายุการใช้งานเพิ่มขึ้นจากเดิมซึ่งมีค่ามัธยฐาน 20,000 ไมล์ โรงงานจึงจะเปลี่ยนใช้ระบบใหม่ โดยให้ระดับนัยสำคัญไม่เกิน 1% จากการทดลองยางซึ่งผลิตโดยระบบใหม่ 16 เส้น มีอายุการใช้งานเป็นไมล์ ดังนี้

20,600	19,900	23,000	21,800
20,800	22,200	18,900	21,300
21,200	21,700	20,500	18,400
19,800	20,700	21,500	20,900

จงใช้การทดสอบแบบจัดอันดับเครื่องหมายตรวจสอบว่าควรใช้ระบบการผลิตแบบใด

2.10 สำนักงานแห่งหนึ่งตั้งมาตรฐานหลอดไฟขนาด 100 วัตต์ ที่จะซื้อมาใช้ในสำนักงานจำนวนมากว่า ต้องมีหลักฐานสนับสนุนว่า เมื่อใช้ 10 นาทีแล้วกินไฟน้อยกว่า 50 ยูนิต ถ้าทดลองใช้ 12 หลอด จำนวนยูนิตที่ใช้ใน 10 นาที มีดังนี้

48.7	49.6	49.4	49.0	51.5	49.7
48.8	49.1	49.2	51.8	49.3	48.9

จงใช้การทดสอบแบบจัดอันดับเครื่องหมาย ทดสอบว่าสำนักงานจะตกลงใช้หลอดไฟชนิดนี้หรือไม่ ถ้าใช้ $\alpha = .01$

2.11 ข้อมูลต่อไปนี้คือน้ำหนักของกระเป๋าเดินทางและสัมภาระของการเดินทางโดยเรือบินของนักฟุตบอล 2 สมาคม คือ A และ B

A	34	39	41	28	33	
B	36	40	35	31	39	36

จงใช้การทดสอบของ Wilcoxon ด้วย $\alpha = .05$ ทดสอบว่า น้ำหนักเฉลี่ยของกระเป๋าเดินทางของนักฟุตบอลสมาคม B สูงกว่าสมาคม A หรือไม่ ?

2.12 มีนักเรียนคณิตศาสตร์ 12 คน ซึ่งมีความสามารถเท่า ๆ กัน และเรียนคณิตศาสตร์โดยมีโปรแกรมวัสดุ เลือกมา 5 คนแบบสุ่มและครูสอนเพิ่มเติมจากเดิม และให้สอบข้อสอบชุดเดียวกัน ได้คะแนนสอบดังนี้

ครูสอนเพิ่มเติม	87	69	78	91	80		
โปรแกรมเดิม	75	88	64	82	93	79	67

จงใช้การทดสอบของ Wilcoxon และ $\alpha = .05$ ตรวจสอบว่าการสอนเพิ่มเติม ช่วยให้นักเรียนได้คะแนนสูงกว่าเดิมหรือไม่ ?

2.13 น้ำหนักของคน 5 คนซึ่งชั่งก่อนและหลังหยุดสูบบุหรี่ 5 สัปดาห์ มีดังนี้

คน	1	2	3	4
ก่อนเลิกสูบ	148	176	153	116
ภายหลังเลิกสูบ	154	179	151	121

จงใช้การทดสอบของ Wilcoxon ทดสอบว่าการเลิกสูบบุหรี่ทำให้น้ำหนักเพิ่มขึ้นหรือไม่ ? $\alpha = .05$

2.14 รัฐบาลให้ทุนอุดหนุนวิทยาลัยเกษตรกรรม 9 แห่ง เพื่อเปรียบเทียบข้าวพันธุ์ใหม่ 2 พันธุ์ โดยกำหนดให้ทุกวิทยาลัยใช้พื้นที่ทดลองเท่ากัน ข้อมูลต่อไปนี้คือผลผลิตเป็นบุชเชลต่อ 1 เอเคอร์

	วิทยาลัย								
	1	2	3	4	5	6	7	8	9
พันธุ์ 1	38	23	35	41	44	29	37	31	38
พันธุ์ 2	45	25	31	38	50	33	36	40	43

จงใช้ $\alpha = .05$ และการทดสอบของ Wilcoxon ตรวจสอบว่าผลผลิตโดยเฉลี่ยของทั้ง 2 พันธุ์ มีความแตกต่างอย่างมีนัยสำคัญ หรือไม่ ?

2.15 ในการเปรียบเทียบคุณภาพของยาง 2 ชนิดคือ A และ B ได้นำยางทั้ง 2 ไปใช้กับรถแท็กซี่ 8 คัน โดยสุ่มใช้กับล้อหลังด้านซ้ายหรือขวาด้านละ 1 ชนิด ได้ข้อมูลดังนี้

รถ	**ชนิด A**	**ชนิด B**
1	21,400	22,800
2	28,300	29,100
3	22,800	23,400
4	19,900	19,300
5	30,100	29,700
6	20,400	22,600
7	23,700	24,200
8	18,700	19,600

จงใช้การทดสอบของ Wilcoxon ทดสอบ $H_0 : \mu_A = \mu_B$ และ $H_a : \mu_A < \mu_B$, $\alpha = .01$

2.16 ผู้ผลิตอาหารลดน้ำหนักอ้างว่าจะทำให้ลดได้ 10 ปอนด์ใน 2 สัปดาห์ ถ้าน้ำหนักก่อนและหลังการทดลองของหญิง 7 คน มีดัง

	หญิง						
	1	2	3	4	5	6	7
น้ำหนักก่อน	129	133	136	152	141	138	125
น้ำหนักหลัง	130	121	128	137	129	132	120

จงใช้การทดสอบของ Wilcoxon ตรวจสอบว่าอาหารดังกล่าวสามารถลดน้ำหนักได้โดยเฉลี่ย 10 ปอนด์ หรือต่ำกว่า 10 ปอนด์ โดยใช้ $\alpha = .05$

2.17 จงทดสอบว่าไม่มีความแตกต่างของอายุระหว่างพนักงานชายและหญิงของโรงงานหนึ่ง โดยใช้ แบบทดสอบ U ของแมน-วิทนี่ และ $\alpha = .05$

ชาย	26	25	38	33	42	40	44	26	43	35
หญิง	44	30	34	47	35	46	35	47	48	34

2.18 ราคาขายปลีกของสินค้า A และ B จากร้านที่สุ่มมาจากเมืองต่าง ๆ มีดังนี้

A	89	90	92	81	76	88	85	
B	78	93	81	87	89	90	96	82

ราคาขายปลีกของพิมพ์ดีด A และ B แตกต่างกันหรือไม่ ถ้าใช้ $\alpha = .01$ ยกเว้นข้อจำกัด $n_1, n_2 < 10$ และใช้แบบทดสอบของแมน-วิทนี่

2.19 ข้อมูลต่อไปนี้คือจำนวนหน่วยของสินค้าที่ผลิตได้ต่อวันของพนักงานในกลุ่มอายุที่ต่างกัน 2 กลุ่มอายุ จงใช้แบบทดสอบแมน-วิทนี่ ตรวจสอบว่ามีความแตกต่างระหว่างผลผลิตของกลุ่มอายุต่ำ และอายุสูงหรือไม่ โดยใช้ $\alpha = .05$ และยกเว้น ข้อจำกัด $n_1, n_2 < 10$

เกิน 40	24	28	15	47	23	25	53	20		
ต่ำกว่า 40	22	12	30	16	26	14	18	21	16	18

2.20 บริษัทอีกแห่งหนึ่งต้องการทราบว่าอัตราผลผลิตของคนงานชายและหญิงมีความแตกต่างกันหรือไม่ ผู้จัดการฝ่ายผลิตได้เก็บผลงานของพนักงานหญิงและชายเป็นเวลา 1 สัปดาห์ ภายหลังเลขาของผู้จัดการได้เก็บข้อมูลเดิมผิดแฟ้ม จนค้นข้อมูลเดิมไม่ได้ ทราบเพียงว่า $\sigma_U = 64.26$, $\mu_U = 420$ และ $R_1 = 830$

ผู้จัดการจำไม่ได้ว่าใช้ขนาดตัวอย่างเท่าใด แต่จำได้ว่าขนาดตัวอย่างของกลุ่มชายคือ n_2 มากกว่า n_1 อยู่ 2 คน

จงคำนวณค่า Z และทดสอบนัยสำคัญที่ $\alpha = .05$ และจงหาค่า n_1, n_2 และ R_2

($n_1 = 28$, $n_2 = 30$, $R_2 = 881$, ยอมรับ H_0)

4.3 แบบทดสอบของ Kruskal - Wallis

แบบทดสอบนี้ใช้แทนการวิเคราะห์ความแปรปรวน (ANOVA) **ของการทดสอบแบบมีพารามิเตอร์** เนื่องจากการใช้ F-test ต้องมีข้อสมมุติว่าประชากรที่เรานำค่า μ_j มาเปรียบเทียบกันนั้น ต้องเป็นประชากรแบบปกติ และมีความแปรปรวนเป็นเอกภาพ นั่นคือมีข้อสมมุติว่า $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2 = \sigma^2$ และสมมติฐาน $H_0: \mu_1 = \mu_2 = \ldots = \mu_k = \mu$ ถ้าข้อสมมติไม่สมบูรณ์ เราจะใช้วิธีการของสถิติไร้พารามิเตอร์ โดยใช้การทดสอบแบบผลรวมของอันดับ เรียกว่าแบบทดสอบของ Kruskal-Wallis ซึ่งให้ชื่อตามผู้คิดค้น 2 ท่าน ในปี 1952 บางครั้งจะเรียกสั้น ๆ ว่า แบบทดสอบ H(H test) เพราะใช้ตัวสถิติ H ทดสอบ แบบทดสอบนี้ไม่ต้องใช้ข้อสมมุติเกี่ยวกับประชากร นอกจากต้องเป็นประชากรแบบต่อเนื่อง ดังนั้นข้อมูลที่เก็บมาจึงต้องเป็นตัวแปรเชิงสุ่มแบบต่อเนื่องและมีสมมุติฐานดังนี้

H_0 : ประชากรทั้งหลายมีการแจกแจงเหมือนกัน (identical distributions)

H_a : ประชากรทั้งหลายมีการแจกแจงไม่เหมือนกันทั้งหมด

โดยมีทั้งหมด k ประชากร n_j คือขนาดตัวอย่างที่สุ่มจากแต่ละประชากรซึ่งไม่จำเป็นต้องเท่ากันทั้งหมด วิธีการทดสอบคล้าย ๆ กับการทดสอบ $H_0: \mu_1 = \mu_2$ จากตัวอย่าง 2 กลุ่มที่เป็นอิสระด้วยวิธีการของ Wilcoxon คือต้องเอาข้อมูลทั้งหมดมาเรียงกันจากค่าน้อยไปมาก เพื่อจัดอันดับค่าน้อยที่สุดจะเป็นอันดับ 1 ค่าสูงจุดจะเป็นอันดับสุดท้ายคือ อันดับ N ในเมื่อ $N = \Sigma n_j$ แล้วหาผลรวมของอันดับคือ R_j ของทุก ๆ กลุ่ม แล้วคำนวณค่าสถิติ H

$$H = \frac{12}{N(N+1)}\sum_{j=1}^{k}\frac{R_j^2}{n_j} - 3(N+1)$$

H จะมีการแจกแจงแบบ χ^2 ที่มี df = (k-1), k = จำนวนกลุ่ม และจะปฏิเสธ H_0 เมื่อ $H > \chi^2_{\alpha\,(k-1)}$

ตัวอย่างที่ 10 ผลสอบของวิธีสอนภาษาอังกฤษ 3 วิธี โดยมี $n_1 = 9, n_2 = 10$ และ $n_3 = 10$ ได้คะแนนสอบ และจัดอันดับดังนี้

กลุ่มที่ 1		กลุ่มที่ 2		กลุ่มที่ 3	
คะแนน	อันดับ	คะแนน	อันดับ	คะแนน	อันดับ
93	23.5	89	19	78	9
77	8	90	20	80	11
93	23.5	85	15	75	6
79	10	76	7	81	12
92	22	84	14	91	21
99	28	95	26	88	18
98	27	82	13	86	16
71	3	72	4	94	25
87	17	73	5	69	2
		68	1	100	29
$R_1 = 162$		$R_2 = 124$		$R_3 = 149$	

ให้ทดสอบว่าตัวอย่างทั้ง 3 กลุ่มนี้มาจากประชากรที่มีการแจกแจงเหมือนกันด้วย $\alpha = .05$

$n_1 = 9,\quad n_2 = 10,\quad n_3 = 10,\quad k = 3,\quad N = \Sigma n_j = 29$

$R_1 = 162,\quad R_2 = 124,\quad R_3 = 149$

คำนวณค่าสถิติ H

$$H = \frac{12}{N(N+1)} \Sigma \frac{R_j^2}{n_j} - 3(N+1)$$

$$= \frac{12}{29(30)} \left[\frac{(162)^2}{9} + \frac{(124)^2}{10} + \frac{(149)^2}{10} \right] - 3(29+1)$$

$$= \frac{12}{870} (2916 + 1537.6 + 2220.1) - 90$$

$$= 2.05$$

H จะมีการแจกแจงแบบ $\chi^2_{(k-1),\ \alpha}$

$\chi^2_{2,\ .05} = 5.991$

เราจะปฏิเสธ H_0 ถ้า $H > 5.991$

แต่ค่าคำนวณของ $H = 2.05 < 5.991$

จึงไม่ปฏิเสธ H_0 และสรุปว่าวิธีการสอนทั้ง 3 วิธีไม่ต่างกันอย่างมีนัยสำคัญ

หมายเหตุ การทดสอบของ Kruskal-Wallis จะเป็นการทดสอบแบบหางเดียวเสมอ และเราจะปฏิเสธ H_0 ได้เมื่อ H เป็นค่าโตเกินไปจนตกอยู่ในพื้นที่ α ที่ปลายหางของโค้ง $\chi^2_{(k-1)}$

แบบฝึกหัด

2.21 แบบทดสอบผลรวมของอันดับและการวิเคราะห์ความแปรปรวนมีส่วนที่เหมือนกัน และแตกต่างกันอย่างไรบ้าง ?

2.22 สมมติฐานของการทดสอบแบบผลรวมของอันดับว่าอย่างไร ?

2.23 การทดสอบแบบผลรวมของอันดับเป็นการทดสอบแบบด้านเดียวหรือแบบ 2 ด้าน

2.24 ถ้าใช้การทดสอบผลรวมของอันดับแทนการวิเคราะห์แปรปรวน ส่วนค่าวิกฤตเราจะยังคงใช้ตาราง f หรือไม่ ? และหาค่า df อย่างไร ?

2.25 ปลูกพืช 3 พันธุ์ ในแปลงทดลองที่มีความสมบูรณ์ของดินเท่ากันทั้ง 20 แปลง ได้ผลผลิตต่อไร่ ดังนี้

พันธุ์ 1	20	21	22	23	24	26	28
พันธุ์ 2	19	20	23	18	25	13	12
พันธุ์ 3	15	20	16	17	14	27	

จงทดสอบว่าพันธุ์ทั้ง 3 ให้ผลผลิตต่างกันหรือไม่ โดยใช้ $\alpha = .05$ ($H = 5.692$, ไม่ปฏิเสธ H_0)

2.26 ต้องการทราบว่ายางรถยนต์ 4 ยี่ห้อมีอายุการใช้งานโดยเฉลี่ยต่างกันหรือไม่ จึงสุ่มจากยี่ห้อต่าง ๆ มาชนิดละ 5 เส้นและได้ทดลองใช้ มีอายุการใช้งานเป็นไมล์ดังนี้ ให้ทดสอบว่ายางทั้ง 4 ชนิดมีคุณภาพไม่ต่างกัน $\alpha = .05$

ชนิดที่ 1	ชนิดที่ 2	ชนิดที่ 3	ชนิดที่ 4
35,000	34,500	34,000	34,600
37,000	33,000	33,500	38,000
37,500	34,300	36,000	38,400
32,000	32,500	37,400	40,000
38,500	31,000	36,700	39,100

2.27 สำนักงานแห่งหนึ่งได้จำแนกพนักงานเป็น 3 ประเภท ฝ่ายจัดการต้องการทราบว่าพนักงาน 3 กลุ่มนี้มีคุณธรรมแตกต่างกันหรือไม่ จึงสุ่มมากลุ่มละ 10 คน และให้คะแนนคุณธรรมจากคะแนนเต็ม 40 คะแนน ดังนี้

กลุ่มที่ 1	21	23	25	27	31	33	30	35	13	39
กลุ่มที่ 2	20	22	26	28	30	32	19	34	36	38
กลุ่มที่ 3	15	16	17	24	18	29	30	14	37	24

จงทดสอบโดยใช้ $\alpha = .01$

(H = 3.76, ไม่ปฏิเสธ H_0)

2.28 คะแนนสอบไล่วิชาเศรษฐศาสตร์ สถิติ และจิตวิทยา ของนักเรียน 5 คน มีดังนี้

เศรษฐศาสตร์	สถิติ	จิตวิทยา
98	90	99
86	91	89
78	95	80
84	97	88
87	93	96

จงทดสอบว่านักเรียนทั้ง 5 คน มีความสัมฤทธิผลใน 3 วิชา ไม่ต่างกันโดยใช้ $\alpha = .01$

2.29 ใช้วิธีการต่าง ๆ กัน 3 วิธีเพื่อทำลายเชื้อแบคทีเรียซึ่งแฝงอยู่ในผลิตภัณฑ์ชนิดหนึ่ง จำนวนแบคทีเรียที่ยังเหลือตกค้างอยู่ภายหลังจากใช้วิธีการทั้ง 3 มีดังนี้

วิธีการที่ 1	110	108	105	98	102	111	131	106
วิธีการที่ 2	120	109	115	104	122	119	123	117
วิธีการที่ 3	130	125	135	133	128	103	100	127

วิธีการทั้ง 3 มีประสิทธิภาพแตกต่างกันหรือไม่ ถ้าใช้ α = .05 ?

(H = 4.965, ยังไม่ปฏิเสธ H_0)

5. แบบทดสอบรันสำหรับ 1 กลุ่มตัวอย่าง (One-Sample Runs Tests)

ในการศึกษาวิชาสถิตินั้น จำเป็นต้องมีข้อจำกัดว่า ตัวอย่างที่ได้มาต้องเป็นตัวอย่างแบบสุ่มเสมอ มิฉะนั้นจะนำผลจากตัวอย่างมาวิเคราะห์ไม่ได้ เพราะถือว่าไม่ใช่ตัวแทนที่ดีของประชากร เช่นในการรับสมัครพนักงานในสำนักงานแห่งหนึ่ง โดยตั้งเกณฑ์ไว้ล่วงหน้าว่าจะไม่คำนึงถึงเพศของผู้สมัคร ให้ ญ = หญิง และ ช = ชาย ถ้าพนักงานที่เข้ามาสัมภาษณ์ 16 คน อยู่ในลำดับ ดังนี้

ญ,ญ,ญ,ญ, ช,ช,ช,ช, ญ,ญ,ญ,ญ ช,ช,ช,ช

จะเห็นว่ามี ญ = 8 = ช แต่ลำดับการเข้าสัมภาษณ์ ไม่เป็นแบบสุ่ม เพราะมีโอกาสน้อยมากที่จะได้อนุกรมของ ญ และ ช เป็นชุดละ 4 คน โดยสลับกันเช่นนี้

สมมุติว่า ลำดับการเข้าสัมภาษณ์ เป็นแบบใหม่ ดังนี้

ญ, ช, ญ, ช, ญ, ช, ญ, ช, ญ, ช, ญ, ช, ญ, ช, ญ, ช

เราก็เชื่อว่าลำดับการเข้าสัมภาษณ์ไม่เป็นแบบสุ่มอีกเช่นกัน

ดังนั้นนักสถิติจึงได้พัฒนา ทฤษฎีของ **รัน** (runs)

ตัวอย่างของ รัน คือ ถ้าผู้เข้าสัมภาษณ์ 6 คนแรก อยู่ในลำดับต่อไปนี้ จะมีทั้งหมด 3 รัน

$\underbrace{\text{ญ,}}\ \ \underbrace{\text{ช,ช,ช,ช,}}\ \ \underbrace{\text{ญ}}$

รัน → 1 2 3

และอนุกรมต่อไปนี้มี 7 รัน

$\underbrace{\text{ช,}}\ \ \underbrace{\text{ญ,ญ,}}\ \ \underbrace{\text{ช,ช,}}\ \ \underbrace{\text{ญ,}}\ \ \underbrace{\text{ช,ช,ช,ช,}}\ \ \underbrace{\text{ญ,ญ,ญ,ญ,}}\ \underbrace{\text{ช}}$

รัน → 1 2 3 4 5 6 7

ดังนั้น รัน คืออนุกรมของสัญลักษณ์ที่เหมือนกัน ซึ่งอาจจะตามหรือนำสัญลักษณ์อื่นๆ หรือไม่มีสัญลักษณ์อื่นตามหรือนำเลยก็ได้

เราใช้ผลรวมของรัน (R) ในการจัดเรียงสัญลักษณ์ 2 ชนิดขึ้นไป เพื่อทดสอบการสุ่มของการจัดเรียง (randomness)

5.1 การทดสอบการสุ่มใน 1 กลุ่มตัวอย่าง

ตัวอย่างที่ 11 โรงงานผลิตบะหมี่สำเร็จรูปใช้เครื่องจักรบรรจุของแถมซึ่งเป็นเครื่องเล่นเด็กในซองบะหมี่แบบสุ่ม ของเล่นที่แถมมี 2 ชนิด และบริษัทไม่ต้องการให้เด็กได้ของเล่นซ้ำกัน จึงลองหยิบบะหมี่ที่บรรจุในลำดับติดต่อกันมา 60 ห่อ ให้ A และ B แทนของเล่นแต่ละชนิด ได้ผลดังนี้

B, A, B, B, B A, A, A, B, B, A, B, B, B, B, A, A, A, A, B,

A, B, A, A, B, B, B, A, A, B, A, A, A, A, B, B, A, B, B, A

A, A, A, B, B, A, B, B, B, B, A, A, B, B, A, B, A, A, B, B

n_1 = จำนวนเครื่องหมาย A = 29

n_2 = จำนวนเครื่องหมาย B = 31

R = 29 = จำนวนรัน

$$\mu_R = \frac{2n_1n_2}{n_1+n_2}+1$$

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

และ $Z = \dfrac{R-\mu_R}{\sigma_R}$ จะมีการแจกแจงแบบปกติ

ดังนั้น
$$\mu_R = \frac{2(29)(31)}{29+31}+1$$

$$= 29.97+1$$

$$= 30.97$$

$$\sigma_R = \sqrt{\frac{2(29)(31)[2(29)(31 - 29 - 31)]}{(29+31)^2(29+31-1)}}$$

$$= \sqrt{\frac{(1798)(1738)}{(60)^2(59)}}$$

$$= \sqrt{\frac{3{,}124{,}924}{212{,}400}} = \sqrt{14.71} = 3.84$$

ดังนั้น $Z = \frac{29 - 30.97}{3.84}$

$= -0.51$

H_0 : การบรรจุของเล่นเป็นแบบสุ่ม

H_a : การบรรจุของเล่นไม่เป็นแบบสุ่ม

ถ้าใช้ $\alpha = .20$ เขตวิกฤตคือ $Z_C > 1.28$ และ $Z_C < -1.28$

แต่ $Z_C = -0.51$ ไม่อยู่ในเขตวิกฤต จึงสรุปว่าการบรรจุของเล่นเป็นแบบสุ่ม

แบบฝึกหัด

2.30 จงทดสอบว่าข้อมูลต่อไปนี้เป็นแบบสุ่มหรือไม่ ด้วย $\alpha = .05$

A, B, B, B, B, A, B, A, B, B, A, A, B, A, A, A, B, B, B, A,

B, B, A, A, A, A, B, A, B, B, A, A, A, B, A, A, B, A, B, B

2.31 ผลิตภัณฑ์เครื่องแก้วซึ่งขนส่งทางเรือ เมื่อตรวจสอบพบลำดับของดี (A) และชำรุด (D) ดังนี้ A, A, A, A, D, A, D, D, D, A, A, D, D, A, A, A, A, A, A, A, D, D, D, D, D ให้ทดสอบว่าสินค้าชำรุดเป็นแบบสุ่มหรือไม่ โดยใช้ $\alpha = .10$

(ปฏิเสธ H_0, ไม่เป็นแบบสุ่ม)

2.32 ข้อมูลต่อไปนี้คือเปอร์เซ็นต์สินค้าชำรุดของเครื่องจักรเครื่องหนึ่งในวันทำการ 25 วันติดต่อกัน จงทดสอบว่าอนุกรมเป็นแบบสุ่มหรือไม่โดยใช้ $\alpha = .05$ โดยเปลี่ยนค่าเหล่านี้เป็นสูงกว่าและต่ำกว่าค่ามัธยฐาน

8.2	9.4	11.1	10.4	8.6
10.3	12.3	12.0	9.3	9.7
8.9	10.0	11.8	9.9	10.9
9.4	8.4	10.1	12.2	11.9
10.3	11.4	8.8	7.4	11.2

2.33 เจ้าของภัตตาคารแห่งหนึ่งสังเกตว่า ลูกค้าสูงอายุ มักมาทานอาหารเย็นเร็วกว่าคู่หนุ่มสาว ซึ่งอาจเป็นผู้สูงอายุไม่ต้องรับภาระเรื่องลูกที่ยังเล็ก จึงปลีกเวลามาได้ก่อน เขาจึงจดบันทึก ลูกค้าที่เริ่มมาทานอาหารเย็นตั้งแต่ 17.30 น. ตามลำดับของลูกค้าที่มา โดยให้ A แทน ลูกค้าที่มีอายุ 30 ปีขึ้นไป และ B แทน ลูกค้าที่มีอายุต่ำกว่า 30 ปี ดังนี้

(17.30 น.) A, A, A, A, A, A, B, A, A, A, A, A, A, B, B

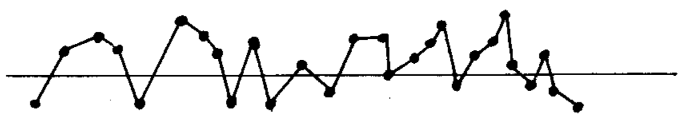
B, A, B, B, B, B, B, B, A, B, B, B, B, B, B, A (4 ทุ่ม)

จงใช้ $\alpha = .05$ ทดสอบว่าอายุของลูกค้าในเวลาต่าง ๆ เป็นแบบสุ่มหรือไม่ ?

5.2 การทดสอบการเปลี่ยนแปลงแบบขึ้น-ลง (Tests of Runs Up and Down)

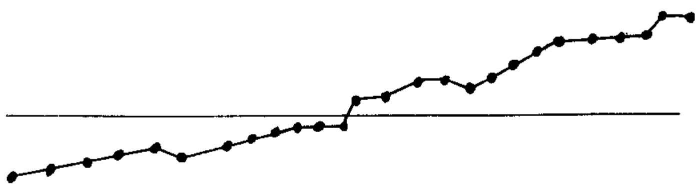
รูปที่ 1.1 แสดงการเคลื่อนไหวของอนุกรม เมื่อเป็นแบบสุ่มและแบบไม่สุ่ม รูป (ก) แสดงการเคลื่อนไหวแบบสุ่ม ส่วนรูป (ข - จ) แสดงการเคลื่อนไหวแบบไม่สุ่ม อาจมีลักษณะแนวโน้ม หรือมีลักษณะเคลื่อนไหว หรืออาจมีลักษณะเปลี่ยนแปลงแบบไม่คงที่

รูปที่ 1.1 *แสดงอนุกรมแบบสุ่มและไม่สุ่ม*



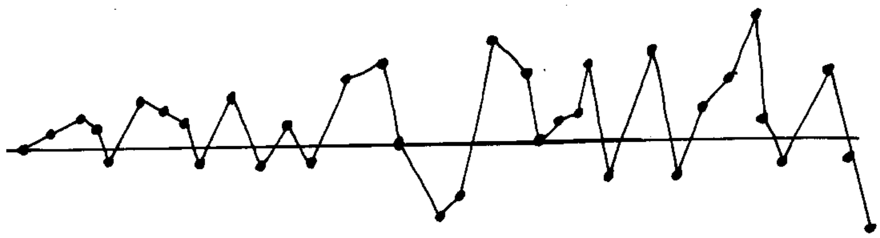
(ก) อนุกรมแบบสุ่มโดยปกติ

(Random Normal Process)



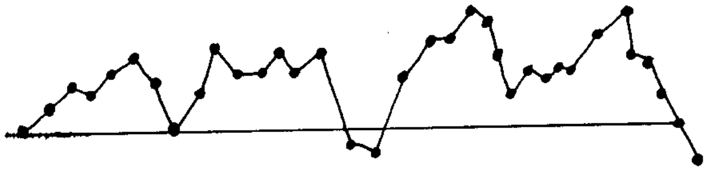
(ข) อนุกรมมีแนวโน้มสูงขึ้น

(Process Involving An Upward Trend)



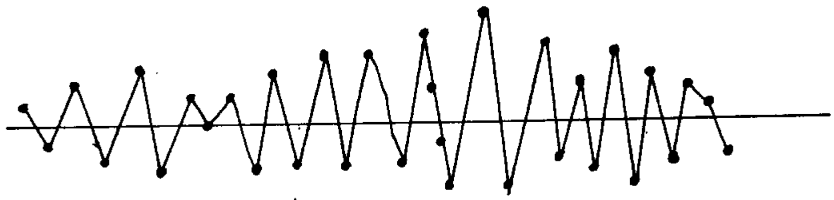
(ค) อนุกรมมีความผันแปรสูงขึ้น

(Process Involving Increasing Variability)



(ง) อนุกรมมีการเปลี่ยนแปลงแบบวัฏจักร

(Process Involving Cyclical Behavior)



(ค) อนุกรมมีการเปลี่ยนแปลงทิศทางอย่างปกติ

(Process Involving Regular and Frequent Changes in Direction)

สมมติฐานของการทดสอบคือ

H_0 : การเปลี่ยนแปลงเป็นแบบสุ่ม

H_a : การเปลี่ยนแปลงไม่เป็นแบบสุ่ม

วิธีการทดสอบคือต้องใช้สัญลักษณ์ + และ - แทนการเปลี่ยนแปลง ถ้าเปลี่ยนแปลงในทางสูงขึ้นใช้เครื่องหมาย + ถ้าเปลี่ยนแปลงแบบลดลงใช้เครื่องหมาย - แล้วหาจำนวน run ให้ R คือจำนวน run ทั้งหมด และเมื่อขนาดตัวอย่าง n ใหญ่พอควร R จะมีการแจกแจงโดยประมาณแบบโค้งปกติ ซึ่งมี

$$\mu_R = \frac{2n - 1}{3}$$

และ $$\sigma_R = \sqrt{\frac{16n - 29}{90}}$$

ตัวอย่างที่ 12 ระดับน้ำสูงสุดของแม่น้ำแห่งหนึ่งในเวลา 30 ปีติดต่อกัน มีดังนี้

ปี	ระดับน้ำ	การเปลี่ยนแปลง	รัน
1	6.63		
2	6.59	-	1
3	6.46	-	
4	6.49	+	2
5	6.45	-	3
6	6.41	-	
7	6.38	-	
8	6.26	-	
9	6.09	-	
10	5.99	-	
11	5.92	-	
12	5.93	+	4
13	5.83	-	5
14	5.82	-	
15	5.95	+	6
16	5.91	-	7
17	5.81	-	
18	5.64	-	
19	5.51	-	
20	5.31	-	
21	5.36	+	8
22	5.17	-	9
23	5.07	-	
24	4.97	-	
25	5.00	+	10
26	5.01	+	
27	4.85	-	11
28	4.79	-	
29	4.73	-	
30	4.76	+	12

$$R = 12,\ n = 30$$

$$\mu_R = \frac{2n-1}{3} = \frac{2(30)-1}{3} = 19.67$$

$$\sigma_R = \sqrt{\frac{16n-29}{90}} = \sqrt{\frac{16(30)-29}{90}}$$

$$= \sqrt{5.01} = 2.24$$

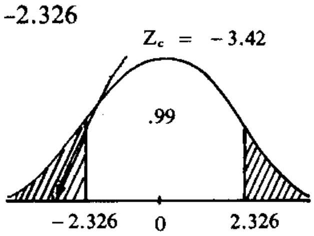
เมื่อใช้ $\alpha = .01$ เขตปฏิเสธ H_0 คือ $Z_C > 2.326$

หรือ $Z < -2.326$

$$Z = \frac{R-\mu_R}{\sigma_R} = \frac{12-19.67}{2.24}$$

$$= -3.42$$

$Z_C = -3.42$ อยู่ในเขตวิกฤต



จึงสรุปว่าการเปลี่ยนแปลงของระดับสูงสุดของน้ำในระยะติดต่อกัน 30 ปีไม่เป็นแบบสุ่ม และจากข้อมูลมีแนวโน้มที่ลดลง

6. แบบทดสอบสหสัมพันธ์แบบจัดอันดับ (The Rank Correlation Test)

เมื่อเราศึกษาเรื่องความถดถอย เราทราบวิธีการหาสัมประสิทธิ์สหสัมพันธ์ ซึ่งต้องมีข้อสมมุติว่าตัวแปรทั้งคู่มีการแจกแจงแบบปกติ ดังนั้น เมื่อเราไม่แน่ใจว่าตัวแปรคู่หนึ่งมีการแจกแจงแบบปกติหรือไม่ เราจึงควรใช้วิธีการของสถิติไร้พารามิเตอร์ ซึ่งมีอยู่หลายวิธีการ แต่วิธีการที่ดีที่สุดคือแบบทดสอบที่สเปียร์แมนสร้างไว้ในปี ค.ศ. 1904 ซึ่งใช้วิธีการง่าย ๆ คือ จัดอันดับข้อมูลในแต่ละกลุ่มตัวอย่าง แล้วคำนวณสหสัมพันธ์ระหว่างอันดับของข้อมูล 2 กลุ่มนั้น วิธีการหาสหสัมพันธ์แบบจัดอันดับจึงเป็นที่นิยม เพราะบางกรณีข้อมูลที่เก็บมาไม่เป็นตัวเลขอยู่ในรูปนามบัญญัติ เราจึงต้องวัดดีกรีของความพอใจ แล้วจัดอันดับความพอใจ ข้อได้เปรียบอีกประการหนึ่งคือเมื่อขนาดตัวอย่างโตมาก การคำนวณค่า r แบบมีพารามิเตอร์จะยุ่งยาก ถ้าไม่มีเครื่องทุ่นแรงช่วย แต่ถ้าเราเปลี่ยนเลขเหล่านั้นให้อยู่ในรูปอันดับ การคำนวณจะง่ายขึ้น เช่น ถ้าเราต้องการวัดความเกี่ยวพันระหว่างผลการเรียนและผลการทำงาน เรามีพนักงานที่ใช้เป็นตัวอย่าง 5 คน เราอาจไม่ทราบคะแนนสอบที่แท้จริงของพนักงานทั้ง 5 แต่เราพอทราบว่าพนักงานคนไหนเรียนเก่งที่สุดหรือเก่งกว่าใครบ้าง ถ้าให้อันดับที่ 1 แทนคะแนนต่ำสุด อันดับ

ที่ 5 แทนผู้ได้คะแนนสูงสุด เราก็สามารถให้อันดับกับพนักงานทั้ง 5 ได้ ส่วนอีกคอลัมน์หนึ่งคือความสัมฤทธิ์ผลในการงานภายหลังจบการศึกษา 10 ปี เราก็สามารถจัดอันดับโดยให้ผู้ที่สัมฤทธิ์ผลน้อยที่สุดเป็นอันดับ 1 ผู้ประสบความสำเร็จในการงานสูงสุดเป็นอันดับ 5 ถ้าได้ข้อมูลดังนี้

ตัวอย่างที่ 13

พนักงาน	อันดับ ผลการเรียน	อันดับ ผลงาน
เมธี	4	4
เมธาวี	3	3
สมศรี	1	1
สมบัติ	2	2
ระวิวรรณ	5	5

จะเห็นว่าผลการเรียนและผลการทำงานสอดคล้องกันโดยสมบูรณ์ สูตรหาค่าสัมประสิทธิ์สหสัมพันธ์ คือ

$$r_s = 1 - \frac{6\Sigma d_i^2}{n(n^2 - 1)}$$

ในเมื่อ r_s = สัมประสิทธิ์สหสัมพันธ์แบบจัดอันดับซึ่งคิดค้นโดย Spearman

n = จำนวนคู่ของข้อมูล

d = ผลต่างของอันดับของแต่ละคู่

จากตัวอย่างข้างบน ทุกคู่มี d_i เป็น 0 ดังนั้น $\Sigma d_i = 0$

ดังนั้น
$$r_s = 1 - \frac{6(0)}{5(5^2 - 1)}$$
$$= 1 - \frac{0}{120}$$
$$= 1 - 0$$
$$= 1$$

$r_S = 1$ แสดงว่าการศึกษาและผลงานมีสหสัมพันธ์แบบสมบูรณ์ในทิศทางเดียวกัน ลองสมมุติใหม่ว่า ผลงานตรงข้ามกับผลการศึกษา ในตารางต่อไปนี้

ตัวอย่างที่ 14

พนักงาน	อันดับของ ผลการเรียน	อันดับของ ผลงาน	d_i	d_i^2
ชาลี	5	1	4	16
ทิวา	1	5	–4	16
ราตรี	3	3	0	0
ชาลินี	2	4	–2	4
กัทลี	4	2	2	4
			Σd_i^2 =	40

$$r_S = 1 - \frac{6(40)}{5(25 - 1)}$$

$$= 1 - \frac{240}{120}$$

$$= 1 - 2$$

$$= -1$$

ค่า $r_S = -1$ แสดงว่าผลการเรียนและผลงานมีสหสัมพันธ์แบบสมบูรณ์ในทิศทางตรงข้าม เช่น ถ้าทราบว่า $r_S = -1$ และเราทราบว่าชาลีเรียนเก่งที่สุด เราจะพยากรณ์ล่วงหน้าอย่างถูกต้องว่าอีก 10 ปีข้างหน้า ชาลีจะได้เงินเดือนน้อยที่สุด (ในระหว่าง 5 คนนี้)

ตัวอย่างที่ 13 - 14 แสดงความสัมพันธ์โดยสมบูรณ์ของ 2 ตัวแปร แต่ในทางปฏิบัติมักไม่ค่อยเกิดกรณีทั้ง 2 นี้ เนื่องจากความสำเร็จในผลงานมิได้เกี่ยวข้องกับพื้นฐานการศึกษาอย่างเดียว ยังมีปัจจัยอื่นอีก แต่ก็เป็นแนวทางให้เราทราบว่า ขีดจำกัดของ r_S ก็เหมือนกับ r คือมีค่าสูงสุดเป็น +1 หรือ -1 เครื่องหมาย + หรือ - เป็นเพียงบอกทิศทางของความสัมพันธ์ และค่าต่ำสุดคือ 0 เมื่อ $r_S = 0$ แสดงว่าไม่มีความสัมพันธ์ระหว่างตัวแปรคู่นั้น

ตัวอย่างที่ 15 ต้องการทราบความสัมพันธ์ของคุณภาพอากาศ และโรคปอด องค์การสุขภาพจึงทำการศึกษาปัญหานี้จากเมือง 11 เมือง ดังนี้

เมือง	อันดับ คุณภาพอากาศ	อันดับ การเป็นโรคปอด	d_i	d_i^2
A	4	5	-1	1
B	7	4	3	9
C	9	7	2	4
D	1	3	-2	4
E	2	1	1	1
F	10	11	-1	1
G	3	2	1	1
H	5	10	-5	25
I	6	8	-2	4
J	8	6	2	4
K	11	9	2	4
			$\Sigma d_i^2 =$	58

1 = อันดับเลวที่สุด

11 = อันดับดีที่สุด

$$r_s = 1 - \frac{6\Sigma d_i^2}{n(n^2-1)}$$

$$r_s = 1 - \frac{6(58)}{11(121-1)}$$

$$= 1 - \frac{348}{1,320}$$

$$= 1 - .264$$

$$= .736$$

ค่า r_S = .736 ชี้ให้เห็นว่าคุณภาพของอากาศ และโรคปอดมีความเกี่ยวพันกันแบบเชิงบวก

การทดสอบนัยสำคัญของ r_S

จากตัวอย่างที่ 15 จากเมืองตัวอย่าง 11 เมือง ได้ค่า r_S = .736 เราจะกล่าวถึงประชากรว่าอย่างไร ? ดังนั้นสมมติฐานคือ

$H_0 : \rho_S = 0$ (ไม่มีสหสัมพันธ์ระหว่างอันดับของข้อมูลในประชากร)
$H_a : \rho_S \neq 0$ (มีสหสัมพันธ์ระหว่างอันดับของข้อมูลในประชากร)

ต้องนำค่า r_S ที่คำนวณได้เปรียบเทียบกับค่าวิกฤตในตารางที่ 11 ท้ายเล่ม ซึ่งใช้สำหรับการทดสอบ 2 ด้าน โดยดูที่ $n = 11$ ถ้าใช้ $\alpha = .05$ ค่าวิกฤตจากตารางคือ $\pm .6091$ นั่นคือ ถ้าคำนวณค่าสถิติ r_S ได้ โตกว่า .6091 จึงจะปฏิเสธ H_0 ได้

จากตัวอย่าง ค่าคำนวณ $r_S = .736 > .6091$ จึงปฏิเสธ H_0 และสรุปว่ามีสหสัมพันธ์ระหว่างคุณภาพของอากาศ และโรคปอดในเมืองต่าง ๆ ทั่วโลก

และเมื่อ $n > 30$ จะไม่มีค่าวิกฤตในตาราง จึงต้องประมาณด้วยการแจกแจงแบบปกติด้วย $\mu_r = 0$ และ $\sigma_r = \frac{1}{\sqrt{n-1}}$

$$\text{ดังนั้นตัวสถิติ } Z = \frac{r_S - 0}{1/\sqrt{n-1}} = r_S\sqrt{n-1}$$

จะมีการแจกแจงแบบปกติ

หมายเหตุ นักสถิติบางคนจะใช้ตัวสถิติ $T = \frac{r_S\sqrt{n-2}}{\sqrt{1-r_S^2}}$ เป็นตัวทดสอบเหมือนกับการทดสอบ $H_0 : \rho = 0$ แบบมีพารามิเตอร์และตัวสถิติ T จะมีการแจกแจงแบบ t ด้วย df = n - 2

ตัวอย่าง 16

จากตัวอย่างที่ 15 $n = 11$, $r_S = .736$

ทดสอบแบบ T ดังนี้

$$T = \frac{.736\sqrt{11-2}}{\sqrt{1-(.736)^2}} = \frac{.736(3)}{.677}$$

$$= 3.26$$

จากตาราง $t_{.025,\ 7} = 2.365$

จึงปฏิเสธ H_0 ได้เช่นเดียวกับการเปรียบเทียบ r_S กับตารางที่ 11

ตัวอย่างที่ 17 นักสังคมวิทยาผู้หนึ่งสังเกตว่าผู้มีสติปัญญาสูง มักเลือกคู่ครองที่มีสติปัญญาสูงด้วย เขาได้เลือกคู่สมรสแบบสุ่มมา 32 คู่ แล้ววัดระดับ IQ ของทั้งสามีและภริยา แล้วได้เปลี่ยนค่า IQ ให้เป็นอันดับ โดยอันดับ 1 คือระดับ IQ ต่ำสุด และอันดับ 32 คือผู้มี IQ สูงสุด และได้หาสัมประสิทธิ์สหสัมพันธ์ระหว่างอันดับของสามีและภริยา ได้ข้อมูลดังนี้

$$n = 32,\ \Sigma d_i^2 = 1043.50$$

$$r_S = 1 - \frac{6(1043.50)}{32(1024-1)}$$

$$= 1 - \frac{6,261}{32,736}$$

$$= 1 - .191$$

$$= .809$$

$H_0 : \rho = 0$ (ไม่มีสหสัมพันธ์ระหว่างอันดับในประชากร นั่นคือสติปัญญาของสามีและภริยาเป็นแบบสุ่ม)

$H_a : \rho > 0$ (มีสหสัมพันธ์แบบเชิงบวกในประชากร นั่นคือ ผู้ฉลาดจะเลือกคู่สมรสที่ฉลาดด้วย)

$\alpha = .01$ นั่นคือจะปฏิเสธ H_0 เมื่อ $Z > 2.33$

$$\sigma_r = 1/\sqrt{n-1} = 1/\sqrt{31} = 1/5.568$$

$$= 0.1796$$

$$Z = r_S/\sigma_r = .809/.1796$$

$$= 4.5$$

ค่า $Z_C = 4.5 > 2.33$ จึงปฏิเสธ H_0 และสรุปว่าผู้มีสติปัญญาสูงจะเลือกคู่สมรสที่มีสติปัญญาสูงด้วย

คุณสมบัติพิเศษของสหสัมพันธ์แบบจัดอันดับ

สหสัมพันธ์แบบจัดอันดับมีข้อดีที่ไม่มีผลกระทบสูงจากข้อมูลที่มีค่าแตกต่างอย่างมากจากตัวอื่น ๆ ตัวอย่างเช่น

X :	10	13	16	19	25
Y :	34	40	45	51	117

จะเห็นว่าค่าสุดท้ายของ Y แตกต่างจากอีก 4 ค่าที่เหลือเป็นอย่างมาก ดังนั้นถ้าคำนวณค่า r แบบมีพารามิเตอร์ อาจจะได้ค่า r ที่สูงมากจนมีนัยสำคัญ ซึ่งเกิดจากผลกระทบของค่าสูงสุดนี้ แต่ถ้าใช้สหสัมพันธ์แบบจัดอันดับ ค่า 117 จะอยู่ในอันดับที่ 5 จะไม่มีผลกระทบสูงต่อค่า r_S

แบบฝึกหัด

2.34 ให้ X คืออัตราความก้าวร้าว

Y คือจำนวนขายในปีที่ผ่านมา

ข้อมูลได้จากพนักงานขาย 8 คน จงทดสอบว่าสหสัมพันธ์อันดับมีนัยสำคัญที่ $\alpha = .05$ หรือไม่ ?

X	30	17	35	28	42	25	19	34
Y	35	31	40	46	50	32	33	42

(r_S มีนัยสำคัญ)

2.35 ผู้เชี่ยวชาญเกษตรได้จัดอันดับคนงาน 8 คน ตามจำนวนชั่วโมงที่ทำงานล่วงเวลาและระยะเวลาการเข้าทำงาน (ปี) ดังนี้

ช.ม.ล่วงเวลา	5.0	8.0	4.0	3.0	7.0	1.0	6.0	2.0
จำนวนปีทำงาน	1.0	6.0	2.0	7.0	8.0	4.5	3.0	4.5

จงใช้ $\alpha = .01$ ทดสอบนัยสำคัญของ r_S

2.36 ในการคัดเลือกผู้จัดการสาขาเปิดใหม่ของบริษัทหนึ่ง ได้มีผู้สมัครทั้งสิ้น 14 คน และใช้กรรมการสัมภาษณ์ 2 คน กรรมการทุกคนจะสัมภาษณ์ผู้สมัครและให้คะแนนอย่างเป็นอิสระกัน จงทดสอบว่าคะแนนของกรรมการทั้ง 2 คนมีสหสัมพันธ์แบบเชิงบวกหรือไม่ ด้วยระดับนัยสำคัญ 1%

ผู้สมัคร	1	2	3	4	5	6	7	8	9	10	11	12	13	14
กรรมการ 1	1	11	12	2	13	10	3	4	14	5	6	9	7	8
กรรมการ 2	4	12	11	2	13	10	1	3	14	8	6	5	9	7

2.37 การทดสอบสหสัมพันธ์แบบจัดอันดับเป็นการทดสอบแบบ 2 ด้าน เสมอไปหรือไม่ ?

2.38 กรรมการ 2 คน ให้คะแนนผู้เข้าประกวดความงาม 11 คน ดังนี้

ผู้เข้าประกวด	A	B	C	D	E	F	G	H	I	J	K
กรรมการ 1	1	2	3	4	5	6	7	8	9	10	11
กรรมการ 2	2	3	1	6	4	5	8	7	10	11	9

จงทดสอบนัยสำคัญของ r_S โดยใช้ $\alpha = .01$

2.39 มีตำแหน่งทางบริหารว่าง 1 ตำแหน่ง มีผู้สมัครทั้งสิ้น 6 คน กรรมการมี 2 คน คือรอง-ประธานบริษัท และผู้จัดการฝ่ายบุคคล ผลการให้คะแนนโดยจัดอันดับความพอใจของ

กรรมการทั้ง 2 คน มีดังนี้

ผู้สมัคร	A	B	C	D	E	F
รองประธาน	1	2	3	4	5	6
ผู้จัดการฝ่ายบุคคล	2	1	3	5	6	4

สหสัมพันธ์ระหว่างอันดับของกรรมการทั้ง 2 คน มีนัยสำคัญที่ α = .01 หรือไม่ ?

2.40 สายการบินได้รับการร้องเรียนว่าเสียงดังรบกวนมาก จึงได้ติดตั้งเครื่องกำจัดเสียงที่ท่อไอเสียของเรือบินเจ็ท ข้อมูลข้างล่างคือจดหมายร้องเรียนก่อนและหลังการติดเครื่องกำจัดเสียง จงทดสอบด้วย α = .05 ว่า เครื่องที่ติดตั้งช่วยลดเสียงรบกวน เพราะถ้าไม่มีหลักฐานสนับสนุนเพียงพอ สายการบินจะถอดออก เพราะในการติดตั้งเครื่องดังกล่าวทำให้เครื่องบินใช้เชื้อเพลิงเพิ่มจากเดิม

จำนวนจดหมายร้องทุกข์ ก่อนและหลังการติดเครื่องกำจัดเสียง

กอน	15	20	24	18	30	46	15	29	17	21	18
หลัง	23	19	12	9	16	12	28	20	16	14	11

2.41 นักอุตุนิยมวิทยาคนหนึ่งมีความเชื่อว่า ถ้าฝนตกหนักติดต่อกัน 6 - 7 ปี จะตามด้วยปีที่แห้งแล้งกว่าปกติจำนวนหนึ่ง แต่นักอุตุนิยมอีกคนหนึ่ง เชื่อว่าปริมาณฝนตกแต่ละปีมีลักษณะเป็นแบบสุ่ม เขาได้ตรวจดูปริมาณฝนย้อนหลังไปหลายปีโดยนำปริมาณฝนแต่ละปีเปรียบ-เทียบกับค่ามัธยฐาน ถ้าสูงกว่าใช้สัญลักษณ์ A ถ้าต่ำกว่าใช้สัญลักษณ์ B ได้ข้อมูลสรุปดังนี้

A, A, B, B, B, B, A, B, A, A, A, B, A, B, A, B, A, A, B, B, B, A, A, B, A,
B, A, A, B, B, B, A, B, B, B, A, B, A, A, A, B, A, A, A, B, A, B, B, A

จงใช้ α = .05 ตรวจดูลักษณะฝนว่ามีรูปแบบหรือเป็นแบบสุ่ม

2.42 ข้อมูลข้างล่างคือคะแนนความพอใจจากผู้ทดลองใช้ผงซักฟอก 2 ชนิด เป็นเวลาติดต่อกัน 3 สัปดาห์ (คนเดียวใช้ 2 ชนิด) จงใช้ $\alpha = .05$ ทดสอบว่าไม่มีความแตกต่างระหว่างผงซักฟอก 2 ชนิดนี้

A	4	4	5	5	3	2	5	3	1	2	5	3	4	2	5	5
B	2	3	3	3	3	3	3	4	3	2	3	2	2	3	3	4

2.43 จำนวนผลผลิตต่อสัปดาห์ของพนักงาน 10 คน ก่อนและหลังการหยุดพักผ่อนประจำปีมีดังนี้

คนงาน	ก่อน	หลัง	คนงาน	ก่อน	หลัง
1	61	69	6	36	38
2	74	77	7	39	33
3	95	93	8	18	20
4	47	48	9	70	73
5	95	99	10	85	89

1. จงทดสอบว่าผลผลิตเปลี่ยนแปลงหรือไม่ โดยใช้ $\alpha = .05469$
2. จงใช้การทดสอบแบบใช้เครื่องหมายและจัดอันดับตรวจสอบว่ามีความแตกต่างของระดับการผลิตที่ $\alpha = .05$ หรือไม่ ?

2.44 ในการตรวจดูประสิทธิภาพของระบบรักษาความปลอดภัยในโรงงานแห่งหนึ่ง โดยเก็บสถิติอุบัติเหตุ ข้อมูลคือจำนวนชั่วโมงแรงงานโดยเฉลี่ยต่อเดือนที่เสียไป เนื่องจากผลของอุบัติเหตุ จากโรงงาน 10 แห่ง ก่อนและหลังการใช้ระบบรักษาความปลอดภัย มีดังนี้

โรงงาน	1	2	4	5	6	7	8	9	10	
ก่อน	28	37	8	65	43	14	15	6	28	115
หลัง	19	38	7	53	31	19	13	4	26	100

ก. จงใช้ $\alpha = .01074$ ตรวจดูว่าโปรแกรมมีอิทธิพลช่วยรักษาความปลอดภัยหรือไม่
(จำนวนเครื่องหมายบวกมี 8 อัน, ยังไม่ปฏิเสธ H_0)

ข. จงใช้การทดสอบแบบจัดอันดับเครื่องหมาย ด้วย $\alpha = .05$

($Z = 2.039$, ปฏิเสธ H_0)

2.45 A B C เป็นพันธุ์พืชที่ใช้ปลูกในแปลงทดลองขนาดและความสมบูรณ์เท่ากัน โดยปลูก A ใน 4 แปลง ปลูก B ใน 5 แปลง และปลูก C ใน 3 แปลง ได้ผลผลิต ดังนี้

A	5	6	11	4	
B	12	7	10	8	9
C	3	2	1		

จงทดสอบว่าผลผลิตเฉลี่ยของพืช 3 พันธุ์นี้ไม่ต่างกันที่ $\alpha = .05$

2.46 เลี้ยงไก่ด้วยอาหาร 3 สูตร แต่ละสูตรใช้เลี้ยงไก่ 5 ตัว มีน้ำหนักเพิ่มขึ้น ดังนี้

กลุ่ม 1	กลุ่ม 2	กลุ่ม 3
3	16	6
14	5	9
6	4	10
7	12	11
2	13	15

จงใช้ระดับนัยสำคัญ 5% ทดสอบว่าสูตรทั้ง 3 มีอิทธิพลต่อการเพิ่มน้ำหนักแตกต่างกันหรือไม่ ?

2.47 เลี้ยงหมูด้วยสูตรอาหาร 3 สูตร มีน้ำหนักเพิ่มขึ้นดังนี้

กลุ่มที่ 1	กลุ่มที่ 2	กลุ่มที่ 3
23	12	20
24	13	19
25	17	14
26	18	15
27		

จงเปรียบเทียบคุณภาพของสูตรอาหารทั้ง 3 ชนิดโดยใช้ $\alpha = .05$

2.48 เมือง X, Y, Z มีประชากรเท่า ๆ กัน สถิติอาชญากรรมภายใน 6 เดือนของ 3 เมืองต่าง ๆ มีดังนี้

เดือน	X	Y	Z
1	3	7	10
2	19	8	17
3	4	11	12
4	5	9	2
5	15	14	18
6	6	16	13

(คำตอบ : H = 1.2 ไม่ปฏิเสธ H_0)

จงทดสอบว่าจำนวนอาชญากรรมโดยเฉลี่ยแตกต่างกันที่ $\alpha = .05$ หรือไม่ ?

2.49 ทางด่วน A, B, C, D มีความยาวเท่ากัน และมีจำนวนอุบัติเหตุจาก 5 วันที่สุ่มมา ดังนี้

วัน	A	B	C	D
1	1	4	6	9
2	2	5	7	10
3	3	16	8	13
4	17	19	11	14
5	18	20	12	15

มีเหตุผลเพียงพอที่จะกล่าวว่าอุบัติเหตุบนทางด่วนเหล่านี้ไม่เท่ากันที่ระดับนัยสำคัญ 5% ไหม ?

2.50 ใช้แบบทดสอบอันเดียวกันกับคน 3 กลุ่ม ๆ ละ 6 คน (ผู้สมัครเข้างาน) ได้คะแนนทดสอบดังนี้

กลุ่ม 1	กลุ่ม 2	กลุ่ม 3
64	73	67
68	77	84
74	80	86
76	81	88
78	82	94
79	83	96

จงตรวจดูว่าคะแนนทดสอบของ 3 กลุ่มนี้เหมือนกันที่ระดับนัยสำคัญ 5% ไหม ?

2.51 ให้คนงาน 3 คน ทำงานชนิดหนึ่ง แล้วจดผลผลิตต่อวัน เป็นเวลา 5 วัน ดังนี้

A	B	C
44	50	45
46	56	48
55	60	50
58	64	52
59	65	61

การทำงานของคนงานทั้ง 3 แตกต่างกันไหม ? $\alpha = .01$ ($H = 3.615$, ไม่ปฏิเสธ H_0)

2.52 บริษัทโฆษณาแห่งหนึ่งต้องการทราบว่า จำนวนครั้งของการโฆษณาทางโทรทัศน์คือ X มีความสัมพันธ์กับจำนวนขาย (Y) ของสินค้าชนิดหนึ่งหรือไม่ เขามีข้อมูลดังนี้

เมือง	X	Y	เมือง	X	Y
1	9	29	7	5	9
2	11	67	8	16	58
3	14	49	9	8	28
4	4	12	10	1	10
5	7	11	11	13	77
6	6	24	12	15	94

1. จงหาสัมประสิทธิ์สหสัมพันธ์ r_S
2. จงทดสอบนัยสำคัญของ r_S, $\alpha = .05$

2.53 ให้ X คือจำนวนผักกาดหอมปลีที่เก็บเกี่ยว และ Y คือราคาในช่วงเวลาต่าง ๆ 10 ช่วงดังนี้

X	1	2	3	4	5	6	7	8	9	10
Y	4	7	6	8	9	5	11	10	12	15

ก) จงหาค่า r_S (.85)

ข) จงทดสอบ $H_0 : \rho_S = 0$, $\alpha = .05$

($T = 4.65$; ปฏิเสธ H_0)

2.54 ห้างเฟอร์นิเจอร์ต้องการทราบความสัมพันธ์ระหว่างจำนวนปีการทำงานเป็นพนักงานและจำนวนขาย จากการสุ่มพนักงานขายมา 9 คน มีจำนวนปีการทำงาน (X) และจำนวนขายต่อปี (Y) ดังนี้

X	1	2	3	4	5	6	7	8	9
Y	2	1	3	8	4	5	6	9	7

ก) จงหาค่า r_S

ข) จงทดสอบ $H_0 : \rho_S = 0$, $\alpha = .05$

2.55 จำนวนชั่วโมงที่ใช้ทำการบ้านวิชาสถิติ (X) และคะแนนวิชาสถิติ (Y) ของนักเรียนที่สุ่มมา 6 คน มีดังนี้

X	10	11	14	7	8	9
Y	38	67	49	12	10	24

ก) จงหาค่า r_S (.89)

ข) จงทดสอบ $H_0 : \rho_S = 0, \alpha = .05$

($T = 3.9$, ปฏิเสธ H_0)

2.56 ทดลองปลูกถั่วเหลือง 6 แปลงเพื่อศึกษาความสัมพันธ์ระหว่างผลผลิต (Y) กับจำนวนน้ำ (X) ได้ผลดังนี้

X	1	2	3	4	5	6
Y	20	28	32	25	30	31

ก. จงหาค่า r_S

ข. จงทดสอบ $H_0 : \rho = 0, H_a : \varrho > 0; \alpha = .01$

2.57 ให้ X คือความสูงของพ่อ, Y คือความสูงของบุตรชาย เป็นเซนติเมตร ดังนี้

X	162	158	168	177	155	175	157	150	178
Y	165	159	160	180	158	170	161	155	175

ก. จงหาค่า r_S

ข. จงทดสอบ $H_0 : \rho_S = 0., \alpha = .05$