ปัญหาเกี่ยวกับการวิเคราะห์ถดถอย

ด้วยความขยัน ด้วยความไม่ประมาท
ด้วยความสำรวมระวัง และด้วยการเอาชนะใจตนเอง
ด้วยคุณธรรมเหล่านี้แหละ ผู้มีปัญญาควรสร้างเกาะ (ที่พึ่ง)
แก่ตน ที่ห้วงน้ำ (กิเลส) ไม่สามารถท่วมได้

การวิเคราะห์ถดถอย เป็นเทคนิคทางสถิติที่ใช้กันอย่างกว้างขวางในเรื่องของความสัมพันธ์ระหว่างตัวแปรต่าง ๆ ในเชิงพารามิเตอร์จะใช้วิธีการกำลังสองต่ำสุด (Least Squares Method) สำหรับปรับเส้นถดถอยเข้ากับข้อมูลจากตัวอย่าง และในการอ้างอิงเกี่ยวกับพารามิเตอร์จะต้องยึดข้อสมมติต่าง ๆ ถ้าละเลยข้อสมมติ แล้วกระบวนการอ้างอิงอาจจะทำให้สรุปผลที่ไว้ใจไม่ได้

ในบทนี้เราจะพิจารณากระบวนการไร้พารามิเตอร์ที่ใช้ในการวิเคราะห์ถดถอย เมื่อไม่สามารถมีข้อสมมติที่จำเป็นสำหรับกระบวนการพารามิเตอร์ได้

8.1 การปรับเส้นถดถอย (Fitting Regression Line)

ให้ $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ เป็นตัวอย่างสุ่มขนาด n ของประชากรชนิดต่อเนื่องที่แต่ละหน่วยให้ข้อมูลเป็นคู่ ๆ และเราต้องการสร้างเส้นถดถอยที่มีรูปฟอร์มเป็น

$$\hat{Y} = a + bX$$

ในเมื่อ a เป็นระยะทางจากจุดกำเนิดถึงจุดตัดแกน Y และ b เป็นความชันของเส้นถดถอย ค่า a และ b เป็นค่าประมาณของพารามิเตอร์ α และ β ในสมการถดถอยประชากรที่ว่า

$$Y = \alpha + \beta X + \varepsilon$$

บราวน์ และมูด (Brown และ Mood, 1951) ได้เสนอวิธีประมาณค่าของ α และ β หรือหาค่า a และ b ไว้ วิธีการของเขานั้น จะต้องแบ่งค่าสังเกต Y เป็น 2 กลุ่ม กลุ่มที่หนึ่งเป็นค่า Y ซึ่งคู่กับค่า X ที่น้อยกว่า หรือเท่ากับมัธยฐานของ X และกลุ่มที่สองเป็นค่า Y ซึ่งคู่กับค่า X ที่มากกว่ามัธยฐานของ X ค่าของ a และ b ที่ต้องการนั้นจะให้เส้นที่มีมัธยฐานของส่วนเบี่ยงเบนที่ Y ห่างจากเส้นเป็นศูนย์ ในแต่ละกลุ่มของสองกลุ่มนั้น

ในการหาค่า a และ b เรามีวิธีการหาดังนี้

(1) สร้างแผนภาพกระจายของข้อมูลตัวอย่าง

(2) ลากเส้นดิ่งผ่านมัธยฐานของค่า X ถ้าจุดหนึ่งหรือหลายจุดอยู่บนเส้นมัธยฐานนี้ให้ย้ายเส้นไปทางขวาหรือซ้าย แล้วแต่ว่าจำนวนของจุดในแต่ละด้านของมัธยฐานเกือบเท่ากัน

(3) หามัธยฐานของ X และมัธยฐานของ Y ในกลุ่มทั้งสองนั้น ซึ่งจะได้มัธยฐาน 4

(4) ในกลุ่มแรกเรากำหนดจุดหนึ่งซึ่งแทนมัธยฐานของ X และมัธยฐาน Y และในกลุ่มที่สองเรากำหนดอีกจุดหนึ่งซึ่งแทนมัธยฐานของ X และมัธยฐาน Y

(5) ลากเส้นต่อจุดทั้งสองนั้น ซึ่งเราจะได้เส้นที่เป็นเส้นประมาณของเส้นถดถอยที่ต้องการ

(6) ถ้ามัธยฐานของส่วนเบี่ยงเบนแนวตั้งของจุดจากเส้นนี้ไม่เป็นศูนย์ในทั้งสองกลุ่ม ก็ให้เคลื่อนเส้นไปสู่ตำแหน่งใหม่ จนเห็นว่าส่วนเบี่ยงเบนในแต่ละกลุ่มมีมัธยฐานเป็นศูนย์ (ทำได้คร่าวๆ โดยใช้ไม้บรรทัดใสๆ แต่ถ้าต้องการความถูกต้องมากๆ ก็ต้องใช้วิธีการของมูด)

(7) ค่าของ a จะเป็นค่าตัดแกน Y ของเส้นถดถอยที่ได้สุดท้ายนี้ และค่าของ b หาได้จาก

$$b = (Y_1 - Y_2) / (X_1 - X_2)$$

ในเมื่อ (X_1, Y_1) และ (X_2, Y_2) เป็นจุดโคออร์ดิเนทของจุดสองจุดใดๆ ในเส้นถดถอย

ตัวอย่าง ในการศึกษาถึงความสัมพันธ์ระหว่างตัวแปร X และ Y ได้ข้อมูลจากตัวอย่างขนาด 10 เป็นดังนี้

X : 3.81 2.10 0.79 1.99 1.03 2.07 0.74 3.88 1.43 0.41

Y : 1.90 1.09 0.44 1.18 0.62 1.29 0.39 2.30 0.93 0.29

มัธยฐานของค่า X เป็น 1.71 ซึ่งแบ่งค่าสังเกตออกเป็นสองกลุ่ม กลุ่มที่อยู่ใต้มัธยฐานของ X จะมี (0.79, 0.44), (1.03, 0.62), (0.74, 0.39), (1.43, 0.93), (0.41,0.29)

และกลุ่มที่อยู่เหนือมัธยฐานจะมี

(3.81, 1.90), (2.10, 1.03), (1.99, 1.18), (2.07, 1.29), (3.88, 2.30)

กลุ่มที่อยู่ใต้มัธยฐานของค่า X จะมีมัธยฐานของค่า X และ Y เป็น 0.79 และ 0.44 ตามลำดับ ส่วนกลุ่มที่อยู่เหนือมัธยฐานของค่า X จะมีมัธยฐานของค่า X และ Y เป็น 2.10 และ 1.29 ตามลำดับ

จากข้อมูลเหล่านี้เราได้ค่าประมาณครั้งแรกของ b หรือ b' เป็น b' = (1.29 - 0.44)/(2.10 - 0.79) = 0.6489

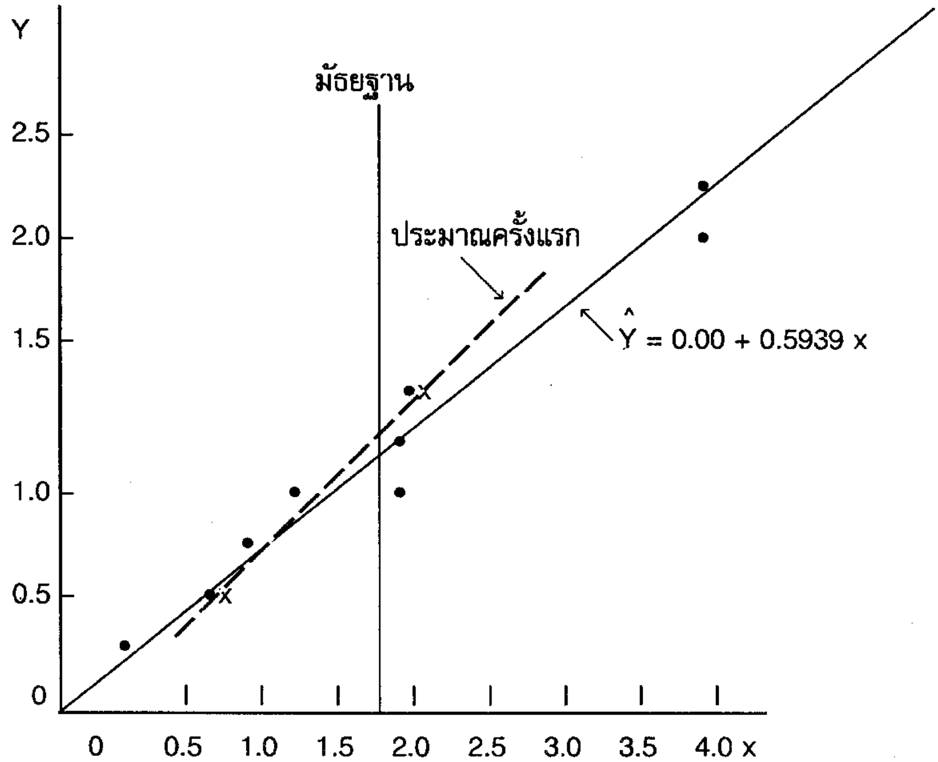
เนื่องจากมัธยฐานของค่าสังเกตจากเส้นที่ได้ครั้งแรกนี้ไม่เป็นศูนย์สำหรับสองกลุ่มนั้น จึงต้องปรับใหม่ ซึ่งจะได้ค่าของ a และ b เป็นดังนี้

$$a = 0.00 \text{ และ } b = 0.5939$$

ดังนั้นสมการหรือเส้นถดถอยที่ต้องการ คือ

$$\hat{Y} = 0.00 - 0.5939X$$

วิธีประมาณเส้นถดถอยดังที่ได้กล่าวมานี้สามารถแสดงได้ดังกราฟต่อไปนี้



ไธล์ (Theil, 1950) ได้เสนอวิธีการประมาณค่าแบบจุดของสัมประสิทธิ์ความชัน ไว้อีกวิธีหนึ่ง คือ

$$\hat{\beta} = \text{มัธยฐาน}\ \{S_{ij}\}$$

ในเมื่อ $S_{ij} = (Y_j - Y_i)/(X_j - X_i)$, $i < j$ เป็นความชันที่เป็นไปได้จากตัวอย่างขนาด n ค่าของ S_{ij} จะมี $\binom{n}{2}$ ค่า

ตัวอย่าง จากการศึกษาถึงความสัมพันธ์ระหว่างตัวแปรสองตัว ได้ข้อมูลจากตัวอย่างที่ศึกษามาดังนี้

X :	206	453	141	131	203	172	153	356	297	425
Y :	9.7	408	106	7.5	49.3	5.8	6.5	2.8	16.8	7.0

ค่าประมาณ $\hat{\beta}$ ที่เป็นไปได้จากตัวอย่างจะมี $\binom{10}{2} = 45$ ค่า เมื่อเรียงจากน้อยไปมากจะเป็นดังนี้

−588.000 −81.111 −61.333 −54.865 −47.872 −27.143 −22.000
−21.739 −13.061 −5.248 −4.214 −3.290 −2.892 −2.869
−2.083 −1.749 −0.675 −0.675 −0.121 −1.093 −0.309
−0.076 0.070 0.102 0.239 0.399 0.620 0.697 0.699
0.713 0.747 0.804 1.168 1.722 8.718 11.364 12.817
13.981 16.562 17.849 30.476 34.091 210.833 544.000

มัธยฐานของค่า S_{ij} เหล่านี้จะเป็น -0.076 ดังนั้นค่าประมาณของ β จะเป็น $\hat{\beta} = -0.076$

8.2 การทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ α และ β

บ่อยครั้งที่เราสนใจทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ α หรือ β หรือทั้ง α และ β เรามีแบบทดสอบต่างๆ ดังนี้

8.2.1 แบบทดสอบบราวน์-มูด (Brown-Mood Test)

แบบทดสอบนี้ใช้ทดสอบสมมติฐานที่ว่า

$$Ho : \alpha = \alpha_0 \text{ และ } \beta = \beta_0 \text{ หรือ } Ho : \beta = \beta_0$$

ในเมื่อ α_0 และ β_0 เป็นค่าที่ระบุไว้ของ α และ β

สำหรับสมมติฐาน Ho : $\alpha = \alpha_0$, $\beta = \beta_0$ กับ Ha : $\alpha \neq \alpha_0$ และ/หรือ $\beta \neq \beta_0$ มีวิธีการทดสอบดังนี้

สุ่มตัวอย่างขนาด n คู่ของค่าสังเกต $(X_1, Y_1), (X_2, Y_2), .., (X_n, Y_n)$ ของตัวแปรต่อเนื่อง X และ Y แล้วสร้างแผนภาพกระจาย

ลากเส้น $Y = \alpha_0 - \beta_0 X$ บนแผนภาพกระจาย และลากเส้นดิ่งผ่านมัธยฐานของค่า X ให้ n_1 เป็นจำนวนจุดหรือคู่ของข้อมูลที่อยู่เหนือเส้นถดถอยตามสมมติฐาน และอยู่ทางซ้ายของเส้นดิ่งที่ลากผ่านมัธยฐานของค่า X และให้ n_2 เป็นจำนวนจุดที่อยู่เหนือเส้นถดถอย และอยู่ทางขวามือของเส้นดิ่ง

มูด ได้แสดงว่าทั้ง n_1 และ n_2 จะมีการแจกแจงทวินาม ที่มีพารามิเตอร์ 0.50 และตัวสถิติทดสอบจะเป็น

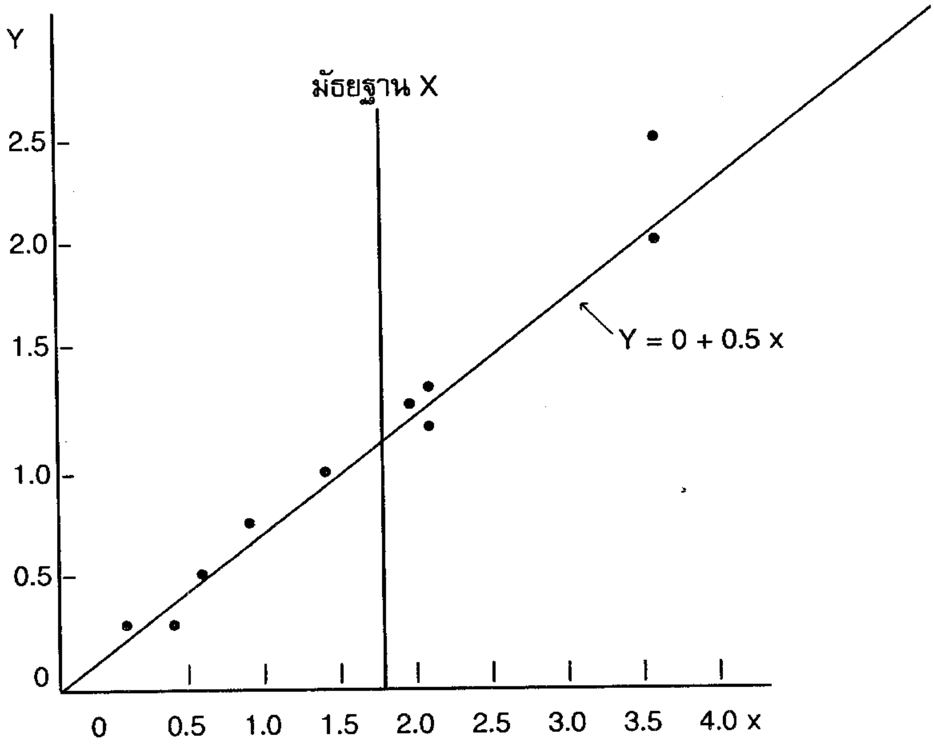
$$\chi^2 = (8/n)\{(n_1-n/4)^2+(n_2-n/4)^2\}$$

ซึ่งจะมีการแจกแจงไคสแควร์ (โดยประมาณ) และมีองศาความเป็นอิสระ 2 ถ้า Ho เป็นจริง และไม่น้อยจนเกินไป ($n \geq 10$)

ตัวอย่าง จากข้อมูลในตัวอย่าง การประมาณค่า a และ b ที่ผ่านมานั้น ถ้าเราต้องการทดสอบสมมติฐานที่ว่า

$$Ho : \alpha = 0, \beta = 0.50 \text{ และ } Ha : \alpha \neq 0, \beta \neq 0.05$$

เราจะทำได้ดังนี้



จากรูปนี้เราได้ $n_1 = 4$ และ $n_2 = 3$ ค่าของตัวสถิติทดสอบจะเป็น

$$\chi^2 = (8/10)\{(4-10/4)^2+(3-10/4)^2\} = 2.00$$

จากตารางไคสแควร์เราได้ $\chi^{2(2)}_{.05} = 5.99$ ดังนั้นเราปฏิเสธ Ho ไม่ได้ ซึ่งแสดงว่าตัวอย่างนั้นอาจจะมาจากประชากรที่เส้นถดถอยมีความชัน $\beta = 0.5$ และจุดตัดแกน $\alpha = 0$

ในการวิเคราะห์ถดถอยนั้น บางครั้งเราสนใจแต่ความชันของเส้นถดถอย β นั่นคือเราต้องการทดสอบสมมติฐานที่ว่า

$$Ho : \beta = \beta_0 \text{ และ } Ha : \beta \neq \beta_0$$

เราก็ใช้แบบทดสอบบราวน์และมูด ซึ่งมีวิธีการดัง

(1) สร้างแผนภาพกระจายจากตัวอย่างของค่าสังเกต n คู่

(2) ลากเส้นดิ่งผ่านมัธยฐานของค่า X

(3) ปรับเส้น $Y = a + \beta_0 X$ เข้ากับข้อมูล ในเมื่อ a เป็นมัธยฐานของส่วนเบี่ยงเบน $Y_i - \beta_0 X$ สำหรับทุกค่าสังเกตของ Y และ β_0 เป็นค่าของ β ที่ระบุไว้ โดยปกติเราสามารถพิจารณาเส้นตรงนี้ได้ง่าย ๆ โดยลากเส้น $Y = \beta_0 X$ และลากเส้นขนานกับเส้น $Y = \beta_0 X$ ซึ่งทำให้แบ่งจุดออกเป็นสองกลุ่มเท่า ๆ กัน

(4) นับจำนวนจุด n_1 ที่อยู่เหนือเส้น $Y = a - \beta_0 X$ และอยู่ข้างซ้ายของมัธยฐาน ของค่า X ตัวสถิติทดสอบจะเป็น

$$\chi^2_b = (16/n)(n_1 - n/4)^2$$

ถ้า Ho เป็นจริง แล้วตัวสถิติทดสอบนี้จะมีการแจกแจงไคสแควร์ ด้วยองศาความเป็นอิสระ 1 เมื่อโตพอ ($n \geq 20$)

ตัวอย่าง ในการทดสอบสมมติฐานที่ว่า Ho : $\beta = 0$ โดยอาศัยตัวอย่างขนาด 24 คู่ ได้ข้อมูลมาดังนี้

X :	23.0	18.7	17.5	21.0	20.0	19.0	15.3	14.0	14.0	13.7
Y :	9.5	9.0	9.2	9.2	9.4	9.3	9.0	8.5	9.0	8.4
	13.3	13.6	13.1	13.0	13.6	14.2	13.9	14.8	14.2	13.0
	8.8	8.9	8.5	8.7	8.6	8.7	8.5	9.1	9.1	8.0
	16.1	15.9	13.0	11.7						
	8.1	8.5	8.4	8.7						

มัธยฐานของค่า X เป็น 14.1

เนื่องจาก $\beta_0 = 0$ เส้น $Y = \beta_0 X$ จึงเป็นแกน X นั่นเอง เส้นที่ขนานกับแกน X ที่แบ่งจุดต่าง ๆ ออกเป็น 2 กลุ่มเท่า ๆ กัน เราเห็นว่า $n_1 = 3$ จุด ที่อยู่เหนือเส้น และอยู่ทางซ้ายมือของมัธยฐานของ X

ค่าของตัวสถิติทดสอบจึงคำนวณได้เป็น

$$\chi^2_b = (16/24)(3-24/4)^2 = 6.00$$

จากตารางไคสแควร์ $\chi^{2(1)}_{.05} = 3.841$ จึงปฏิเสธ Ho

8.2.2 แบบทดสอบไธล์ (Theil's Test)

ไธล์ ได้เสนอวิธีทดสอบสมมติฐาน Ho : $\beta = \beta_0$ ไว้โดยที่ตัวสถิติทดสอบขึ้นอยู่กับตัวสถิติเทา ของเคนดาลล์ (Kendall Tau Statistic)

จากตัวแบบถดถอย

$$Y_i = \alpha + \beta X_i + \varepsilon_i \; ; \; i = 1, 2, .., n$$

ในเมื่อ X_i เป็นค่าคงที่ และทราบค่า ส่วน α และ β เป็นพารามิเตอร์ที่ไม่ทราบค่า ε_i ต่าง ๆ นั้นเป็นอิสระกัน และมาจากประชากรแบบต่อเนื่องเดียวกัน เมื่อเราต้องการทดสอบสมมติฐานเกี่ยวกับ β โดยอาศัยตัวอย่างขนาด n นั้น เรามีกระบวนการทดสอบ ดังนี้

(1) จากแต่ละ X_i ให้สร้างคู่ลำดับ (X_i, D_i) ในเมื่อ

$$D_i = Y_i - X_i \; ; \; i = 1, 2, .., n$$

(2) เรียงคู่ลำดับ (X_i, D_i) ตามค่า X_i ในแนวตั้ง

(3) เปรียบเทียบ D_i กับ D_j $(i < j)$

(4) ให้ P เป็นจำนวนการเปรียบเทียบที่ให้คู่ (D_i, D_j) อยู่ในลำดับธรรมชาติ $(D_i < D_j)$ และให้ Q เป็นจำนวนการเปรียบเทียบที่คู่ (D_i, D_j) อยู่ในลำดับที่กลับกับธรรมชาติ $(D_i > D_j)$ ถ้า D_i เท่ากับ D_j ก็ไม่นับ

ตัวสถิติทดสอบที่ต้องการ กำหนดไว้ว่า

$$\hat{\tau} = \frac{P - Q}{n(n-1)/2} = \frac{S}{n(n-1)/2}$$

เกณฑ์ตัดสินใจกำหนดไว้ดังนี้

(1) สำหรับ Ho : $\beta = \beta_0$ และ Ha : $\beta \neq \beta_0$ จะปฏิเสธ Ho ณ ระดับนัยสำคัญ α ถ้า $\hat{\tau}$ ซึ่งเป็นบวกนั้นมากกว่า τ^* (สำหรับ n และ $\alpha/2$) ที่กำหนดไว้ในตาราง หรือ $\hat{\tau}$ ซึ่งเป็นลบนั้นน้อยกว่าค่าลบของ τ^* จากตาราง

(2) สำหรับ Ho : $\beta \geq \beta_0$ และ Ha : $\beta < \beta_0$ จะปฏิเสธ Ho ณ ระดับนัยสำคัญ α ถ้า $\hat{\tau}$ ซึ่งเป็นลบนั้นน้อยกว่าค่าลบของ τ^* จากตาราง สำหรับ n และ α ที่กำหนดไว้

(3) สำหรับ Ho : $\beta \leq \beta_0$ และ Ha : $\beta > \beta_0$ จะปฏิเสธ Ho ณ ระดับนัยสำคัญ α ถ้า $\hat{\tau}$ ซึ่งเป็นบวกนั้นมากกว่า τ^* จากตาราง สำหรับ n และ α ที่กำหนดไว้

ในกรณีตัวอย่างขนาดโต เราใช้ตัวสถิติ Z ซึ่งมีการแจกแจงปกติมาตรฐาน ดังนี้

$$Z = \frac{S - E(S)}{\sqrt{V(S)}} = \frac{S}{\sqrt{n(n-1)(2n+5)/18}} = \frac{3\tau\sqrt{n(n-1)}}{\sqrt{2(2n+5)}}$$

ตัวอย่าง ในการศึกษาถึงความสัมพันธ์ระหว่างตัวแปรอิสระ X และตัวแปรตาม Y เพื่อทดสอบสมมติฐานที่ว่า ความชันของเส้นถดถอยเป็นศูนย์ จากกลุ่มตัวอย่างเราได้ข้อมูลดังนี้

X :	940	594	1200	1440	1112	625	1385	1035	931	742
Y :	311	391	414	542	387	485	502	458	345	346

เราคำนวณค่าของตัวสถิติทดสอบได้ดังนี้

X_i	$D_i=Y_i-\beta_0 X_i$	เรียงตามธรรมชาติ	เรียงกลับกับธรรมชาติ
594	391	5	4
625	485	2	6
742	346	5	2
931	345	5	1
940	311	5	0
1035	458	2	2
1112	387	3	0
1200	414	2	0
1385	504	1	0
1440	542	0	0
		P = 30	Q = 15

$$\hat{\tau} = \frac{30 - 15}{10(9)/2} = 0.33$$

สำหรับ n = 10 และ $\alpha/2$ = 0.025 เราจะได้ τ^* = 0.511 ซึ่งเราไม่สามารถปฏิเสธ Ho ได้ ($\hat{\tau} < \tau^*$) นั่นคือความชันของเส้นถดถอยเป็นศูนย์

ข้อสังเกต (1) ในการทดสอบ $\beta_o = 0$ นี้สามารถใช้ทดสอบสหสัมพันธ์ระหว่างตัวแปร X และ Y ได้

(2) กรณีที่ X เป็นอันดับที่ซึ่งเกี่ยวกับเวลา (Time order) แล้วแบบทดสอบนี้สามารถใช้ทดสอบแนวโน้มเนื่องจากเวลา (Time Trend) ได้โดยการให้ $\beta_o = 0$

8.2.3 แบบทดสอบวิลคอกซัน (Wilcoxon Matched – Pair Test)

แบบทดสอบวิลคอกซันสามารถดัดแปลงใช้ทดสอบสมมติฐานเกี่ยวกับความชันของเส้นถดถอยได้ กระบวนการทดสอบ ทำได้ดังนี้

แบ่งข้อมูลที่เรียงอันดับตามค่า X ออกเป็น 2 กลุ่มเท่าๆ กัน แล้วประมาณค่า $\hat{\beta}$ จากข้อมูล 2 กลุ่มนั้น ซึ่งได้ค่า $\hat{\beta}$ จำนวน n/2 ค่า

กำหนด $D_i = \hat{\beta}_i - \beta_o$ แล้วใช้แบบทดสอบวิลคอกซันประยุกต์กับค่า D_i ต่างๆ นั้น

ตัวอย่าง ในการศึกษาเกี่ยวกับการถดถอย เพื่อทดสอบสมมติฐานที่ว่า พารามิเตอร์ มีค่ามากกว่าศูนย์ โดยใช้ตัวอย่างขนาด 36 ได้ข้อมูลดังนี้

ครึ่งแรก		ครึ่งหลัง		$\hat{\beta}$	อันดับที่
X	Y	X	Y		
75	6	109	12	(12-6)/(109-75) = 0.18	9
76	2	112	15	(15-2)/(112-76) = 0.36	17
82	13	112	13	(13-13)/(112-82) = 0.00	1.5
83	9	112	14	(14-9)/(112-83) = 0.17	7.5
84	7	112	14	(14-7)/(112-84) = 0.25	13.5
85	8	113	15	(15-8)/(113-85) = 0.25	13.5
85	12	114	14	(14-12)/(114-85) = 0.07	4
85	1	115	9	(9-1)/(115-85) = 0.28	15
90	10	115	15	(15-10)/(115-90) = 0.20	10.5
92	10	116	15	(15-10)/(116-92) = 0.21	12
96	9	117	12	(12-9)/(117-96) = 0.13	5.5
97	6	118	14	(14-6)/(118-97) = 0.38	18
99	11	120	5	(5-11)/(120-99) = -0.29	16
99	15	122	12	(12-15)/(112-99) = -0.13	5.5
101	14	128	14	(14-14)/(128-101) = 0.00	1.5

ครึ่งแรก X	Y	ครึ่งหลัง X	Y	$\hat{\beta}$	อันดับที่
101	10	130	15	(15–10)/(130–101) = 0.17	7.5
106	15	131	12	(12–15)/(131–106) = –0.12	3
107	10	132	15	(15–10)/(132–107) = 0.20	10.5

สมมติฐานที่จะทดสอบจึงเป็น

$$Ho : \beta = 0, \ Ha : \beta > 0$$

เนื่องจาก $\beta_o = 0$ เราจึงได้ว่า $D_i = \beta_i$ และจะได้ค่าของตัวสถิติทดสอบวิลคอกซันเป็น

$$T = 1.5(0) + 1.5(1) + 3(0) + .. + 17(0) + 18(1)$$
$$= 145$$

โดยมี $E(T) = n_o(n_o+1)/4 = 18(19)/4 = 85.5$

$V(T) = n_o(n_o+1)(2n_o+1)/24 = 18(19)(37)/24 = 527.25$

ดังนั้น $$Z = \frac{(T-1/2)-E(T)}{\sqrt{V(T)}} = \frac{(145-0.5)-85.5}{\sqrt{527.25}} = 2.57$$

เนื่องจาก $Z_{.05} = 1.645$ จึงปฏิเสธ Ho

8.3 ช่วงเชื่อมั่นสำหรับสัมประสิทธิ์ความชัน (β)

ในการประมาณสัมประสิทธิ์ความชันนั้น เรามีวิธีการสร้างอยู่หลายวิธี แต่จะขอกล่าวเพียงสองวิธีที่พบกันบ่อยๆ ดังต่อไปนี้

ก. ช่วงเชื่อมั่นที่ขึ้นอยู่กับแบบทดสอบไธล์ ช่วงเชื่อมั่นสองทางชนิดสมมาตรที่ขึ้นอยู่กับแบบทดสอบไธล์ และมีสัมประสิทธิ์เชื่อมั่น $(1-\alpha)$ นั้นสามารถสร้างได้ตามกระบวนการต่อไปนี้

(1) ประมาณความชันที่เป็นไปได้จำนวน $N = \binom{n}{2}$ ค่า นั่นคือ

$$S_{ij} = (Y_j-Y_i)/(X_j-X_i), \ i < j$$

(2) เรียงความชัน S_{ij} ตามขนาดจากน้อยไปมาก ได้เป็น $S_{(1)} = S_{(2)}$ n. $nS_{(N)}$

(3) พิจารณา $S_{\alpha/2}$ โดยอาศัยตาราง 11 ที่มีค่า n และ $\alpha/2$

(4) กำหนด $C_{\alpha/2} = S_{\alpha/2} - 2$

(5) ให้ $k = (N - C_{\alpha/2})/2$

(6) ขีดจำกัดล่าง $\hat{\beta}_L$ ของช่วงเชื่อมั่นสำหรับ β จะเป็นค่า S_{ij} ตัวที่ k (นับจากน้อยไปมาก) และขีดจำกัดบน $\hat{\beta}_U$ จะเป็นค่าที่ k ของ S_{ij} ที่นับจากมากสุดลงมา หรือเป็นค่าที่ N + 1 k ของ S_{ij} ที่นับจากน้อยไปมาก

ดังนั้นช่วงเชื่อมั่น $100(1-\alpha)\%$ สำหรับ β จะเป็น

$$\hat{\beta}_L < \beta < \hat{\beta}_U$$

สำหรับตัวอย่างขนาดโต เราประมาณค่าของ $C_{\alpha/2}$ ได้เป็น

$$C_{\alpha/2} \cong Z_{\alpha/2} \sqrt{n(n-1)(2n+5)/18}$$

ตัวอย่าง จากการศึกษาถึงความสัมพันธ์ระหว่างตัวแปร เพื่อประมาณค่าความชัน โดยอาศัยตัวอย่างขนาด 5 ได้ข้อมูลมาดังนี้

X :	1	2	3	4	5
Y :	1.26	1.27	1.12	1.16	1.03

เราได้ค่า S_{ij} ต่าง ๆ ที่เรียงตามลำดับแล้วจำนวน 10 ค่า ดังนี้

-.150 -.130 -.080 -.070 -.0575 -.055 -.045 -.033 .010 .040

ตัวประมาณค่าแบบจุดของ β จะเป็น

$\hat{\beta}$ = มัธยฐาน$\{S_{ij}\} = \{S_{(5)} + S_{(6)}\}/2$

$= (-.0575 - .055)/2 = -.0563$

จากตาราง เมื่อ $n = 5$ และ $\alpha/2 = .05$ เราได้ $S_{\alpha/2} = 8$ ดังนั้น $C_{\alpha/2} = 8-2 = 6$ และ $k = (10-6)/2 = 2$

เพราะฉะนั้นช่วงเชื่อมั่น 90% สำหรับ β จะเป็น

$S_{(k)} \leq \beta \leq S_{(N-1-k)}$

$S_{(2)} \leq \beta \leq S_{(9)}$

$-.130 \leq \beta \leq .010$

ข. ช่วงเชื่อมั่นที่ขึ้นอยู่กับแบบทดสอบวิลคอกซัน ช่วงเชื่อมั่นสำหรับ β นี้สามารถสร้างได้เช่นเดียวกับช่วงเชื่อมั่นสำหรับผลต่างของมัธยฐานที่อาศัยแบบทดสอบวิลคอกซันชนิดอันดับที่เครื่องหมาย

ตัวอย่าง จากการศึกษาเพื่อประมาณความชันของเส้นถดถอย หรือ β โดยใช้ตัวอย่างขนาด 18 ได้ข้อมูลที่เรียงอันดับของ X เป็นดังนี้

X	Y	X	Y	X	Y	X	Y
75	15	82	20	83	16	84	10
85	19	90	18	99	15	106	23
107	19	109	22	112	22	112	23
114	23	115	18	117	23	128	22
130	22	131	25				

ค่าประมาณ $\hat{\beta}_i$ ต่าง ๆ จากข้อมูลเหล่านี้จะเป็น

(22-15)/(109-75) = 0.21 (22-20)/(112-82) = 0.07

(23-16)/(112-83) = 0.24 (23-10)/(114-84) = 0.43

$(18-19)/(115-85) = -0.03$ $(23-18)/(117-90) = 0.19$

$(22-15)/(128-99) = 0.24$ $(22-23)/(130-106) = -0.04$

$(25-19)/(131-107) = 0.25$

ค่าเฉลี่ยที่เป็นไปได้ $U_{ij} = (\hat{\beta}_i + \hat{\beta}_j)/2$, $i \le j$ ซึ่งมีจำนวน $\binom{10}{2} = 45$ ค่า แสดงได้ดังนี้

	-0.04	0.03	0.07	0.19	0.21	0.24	0.24	0.25	0.43
-0.04	-0.04	-0.005	0.015	0.075	0.085	0.10	0.10	0.105	0.195
0.03		0.03	0.05	0.11	0.12	0.135	0.135	0.14	0.23
0.07			0.07	0.13	0.14	0.155	0.155	0.16	0.25
0.19				0.19	0.20	0.215	0.215	0.22	0.31
0.21					0.21	0.225	0.225	0.23	0.32
0.24						0.24	0.24	0.245	0.335
0.24							0.24	0.245	0.335
0.25								0.25	0.34
0.43									0.43

ค่าประมาณแบบจุดของ β จะเป็น

$$\hat{\beta} = \text{มัธยฐาน } \{U_{ij}\} = \hat{\beta}_{(23)} = 0.195$$

จากตาราง เรามี $n = 9$ และ $\alpha = .05$ จึงได้ $W_{\alpha/2} = 6$ ดังนั้นช่วงเชื่อมั่น 95% สำหรับ β จึงเป็น

$$\hat{\beta}_{(W_{\alpha/2})} \le \beta \le \hat{\beta}_{(k+1-W_{\alpha/2})} , \; k = n(n+1)/2$$
$$\hat{\beta}_{(6)} \le \beta \le \hat{\beta}_{(45-1-6)}$$
$$0.07 \le \beta \le \beta \; 0.31$$

8.4 แบบทดสอบการขนานกันของเส้นถดถอยสองเส้น (Test for Parallelism of two Regression Lines)

การทดสอบการขนานกันของเส้นถดถอยนั้น ก็คือการทดสอบการเท่ากันของความชันของเส้นถดถอยสองเส้นหรือการทดสอบความเป็นเอกภาพของความชัน วิธีการทดสอบการเท่ากันของความชันนี้ทำได้ ดังต่อไปนี้

8.4.1 แบบทดสอบฮอลแลนเดอร์ (Hollander Test for Parallelism)

แบบทดสอบนี้ก็เป็นแบบหนึ่งของแบบทดสอบวิลคอกซันชนิดอันดับที่เครื่องหมายนั่นเอง ให้ตัวแบบถดถอย เป็น

$$Y_{ij} = \alpha_i + \beta_i X_{ij} + \varepsilon_{ij} ; \; i = 1, 2 \; ; \; j = 1, 2, .., n_i$$

ในเมื่อ X ต่าง ๆ เราทราบค่า และ $\alpha_1, \alpha_2, \beta_1$ กับ β_2 เป็นพารามิเตอร์ไม่ทราบค่า ε ต่าง ๆ เป็นอิสระกัน ε_{ij} มาจากประชากรแบบต่อเนื่องเดียวกัน และ ε_{2j} มาจากอีกประชากรหนึ่งที่เป็นแบบต่อเนื่อง แต่ประชากรทั้งสองไม่จำเป็นต้องเหมือนกัน

สมมติฐานที่จะทดสอบจะเป็นแบบใดแบบหนึ่ง ดังนี้

(1) Ho : $\beta_1 = \beta_2$; Ha : $\beta_1 \neq \beta_2$

(2) Ho : $\beta_1 \geq \beta_2$; Ha : $\beta_1 < \beta_2$

(3) Ho : $\beta_1 \leq \beta_2$; Ha : $\beta_1 > \beta_2$

ในการทดสอบสมมติฐานเหล่านี้ก็อาศัยตัวอย่างจาก 2 ประชากร ซึ่งได้ค่าสังเกต $Y_{i1}, Y_{i2}, .., Y_{in}$ ณ ค่าคงที่ $X_{i1}, X_{i2}, .., X_{in_i}$, (i = 1, 2) และจะสมมติว่า $X_{i1} \leq X_{i2} \leq ... \leq X_{in_i}$

สำหรับตัวสถิติทดสอบ T นั้นมีกระบวนการในการคำนวณดังนี้

(1) แบบทดสอบนี้ต้องการจำนวนคู่ของค่าสังเกตในแต่ละกลุ่มเท่ากัน และจำนวนคู่ต้องเป็นเลขคู่ ดังนั้นในชั้นแรกจำต้องสุ่มข้อมูลทิ้งไปให้เหลือเท่า ๆ กัน

(2) ให้ค่าสังเกตที่เหลือในแต่ละกลุ่มซึ่งเท่ากันนั้นเป็น 2n

(3) จากกลุ่มแรกจับคู่ X_{1j} กับ $X_{1,j+n}$; j = 1, 2, .., n แล้วคำนวณความชัน

$$U_{1j} = (Y_{1,j+n} - Y_{1j}) / (X_{1,j+n} - X_{1j}) \; ; j = 1, 2, .., n$$

(4) จากกลุ่มสองจับคู่ X_{2j} กับ $X_{2,t+n}$ และคำนวณความชัน

$$U_{2t} = (Y_{2,t+n} - Y_{2t}) / (X_{2,t+n} - X_{2t}) \; ; t = 1, 2, .., n$$

(5) จับคู่ U_{1j} กับ U_{2j} แบบสุ่มโดยให้แต่ละ U นั้นปรากฏเพียงครั้งเดียวเท่านั้น และคำนวณผลต่าง

$$Z = U_{1j} - U_{2t}$$

ซึ่งจะได้ n ค่า ดังนี้ $Z_1, Z_2, .., Z_n$

(6) ใช้แบบทดสอบวิลคอกซันประยุกต์เข้ากับผลต่าง Z เหล่านั้น

ในกรณีตัวอย่างขนาดโต เราใช้ตัวสถิติทดสอบ Z ซึ่งมีการแจกแจงปกติมาตรฐานดังนี้

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

ตัวอย่าง ในการศึกษาเกี่ยวกับการถดถอยเพื่อทดสอบการเท่ากันของสัมประสิทธิ์ความชันของกลุ่มตัวอย่าง 2 กลุ่ม ที่เป็นอิสระกัน ได้ข้อมูลมาดังนี้

กลุ่ม 1

X	Y	X	Y	X	Y
175.0	1.36	183.0	2.20	190.0	3.05
200.0	4.45	211.0	6.19	220.0	7.78
230.0	9.70	239.5	11.66	245.5	12.96
260.0	16.33	271.5	19.21	284.0	22.52

	X	Y	X	Y	X	Y
กลุ่ม 1	291.0	24.46	302.5	27.78	314.0	31.25
	318.5	32.65	327.0	35.36	337.0	38.65
	345.5	41.52	360.0	46.61	375.0	52.10
	384.5	55.69	397.0	60.55	411.0	66.18
	419.5	69.68	428.5	73.47	440.0	78.41
	454.5	84.81				
กลุ่ม 2	175.0	1.58	183.0	2.48	190.0	3.40
	200.0	4.92	211.0	6.48	220.0	8.66
	230.0	10.86	239.5	13.14	245.5	14.67
	260.0	18.68	271.5	22.14	284.0	26.18
	291.0	28.57	302.5	32.68	314.0	37.03
	318.5	38.80	327.0	42.22		

สมมติฐานที่จะทดสอบ จะเป็น

$$Ho : \beta_1 = \beta_2 \ ; \ Ha : \beta_1 \neq \beta_2$$

ตัวสถิติทดสอบสามารถคำนวณได้ดังต่อไปนี้

(1) เนื่องจากข้อมูลทั้งสองกลุ่มไม่เท่ากัน จึงต้องทำให้เท่ากัน และเป็นเลขคู่ด้วย นั่นคือจะทำให้เหลือกลุ่มละ 16 คู่ โดยที่ต้องตัดทิ้งกลุ่มแรกแบบสุ่มไป 12 คู่ และกลุ่มสองไป 1 คู่ ดังนี้

กลุ่มแรก ตัดคู่ที่มีค่า X เป็น 175.0 183.0 190.0 245.5 271.5 284.0 302.5 318.5 337.0 375.0 419.5 454.5

กลุ่มสอง ตัดคู่ที่ 7 ซึ่งมีค่า X เป็น 230.0

(2) เราได้ 2n = 6 หรือ n = 8

(3) ในกลุ่มแรกเราจับคู่ $X_{1,1}$ กับ $X_{1,9}$, $X_{1,2}$ กับ $X_{1,10}$, และต่อๆ ไปจนถึง $X_{1,8}$ กับ $X_{1,16}$ แล้วเราได้ค่าประมาณของความชัน U_{ij} ดังนี้

U_{11} = (35.36-4.45)/(327.0-200.0) = 0.243

U_{12} = (41.56-6.19)/(345.5-211.0) = 0.263

U_{13} = (46.61-7.78)/(360.0-220.0) = 0.277

U_{14} = 0.298 U_{15} = 0.310 U_{16} = 0.330 U_{17} = 0.356

U_{18} = (78.41-31.25)/(440.0-314.0) = 0.374

(4) ในทำนองเดียวกัน เราจับคู่ $X_{2,1}$ กับ $X_{2,9}$, $X_{2,2}$ กับ $X_{2,10}$ และต่อๆ ไป จนถึง $X_{2,8}$ กับ $X_{2,16}$ แล้วคำนวณความชัน U_{21} ได้เป็น

U_{21} = (18.68-1.58)/(260.0-175.0) = **0.201**

U_{22} = (22.14-2.48)/(271.5-183.0) = **0.222**

U_{23} = **0.242** U_{24} = **0.260** U_{25} = **0.282** U_{26} = **0.302**

U_{27} = (38.80-13.14)/(318.5-239.5) = **0.325**

U_{28} = (42.22-14.67)/(327.0-245.5) = **0.338**

(5) จับคู่แบบสุ่มระหว่าง U_{1j} กับ U_{2t} สมมติว่าได้เป็น

U_{14}, U_{21} U_{11}, U_{24} U_{16}, U_{27} U_{15}, U_{23} U_{18}, U_{25} U_{17} U_{22}

U_{12}, U_{8} U_{13}, U_{26}

แล้วเราได้ผลต่าง Z ดังนี้

Z_1 = 0.298-0.201 = 0.097 Z_2 = -0.017 Z_3 = 0.005

Z_4 = 0.068 Z_5 = 0.092 Z_6 = 0.134

Z_7 = 0.263-0.338 = -0.075 Z_8 = -0.025

(6) ในการหาค่าตัวสถิติทดสอบ T จากค่า Z เราทำได้ดังนี้

Z_i	$\|Z_i\|$	φ_i	R_i
0.097	0.097	1	7
-0.017	0.017	0	2
0.005	0.005	1	1
0.068	0.068	1	4
0.092	0.092	1	6
0.134	0.134	1	8
-0.075	0.075	0	5
-0.025	0.025	0	3

T = 7 + 1 + 4 + 6 + 8 = 26

จากตาราง 10 สำหรับ n = 8 และ α = .05 เราได้ $W_{.025}$ = 4 และ $W_{.975}$ = 8(8+1)/2 - $W_{.025}$ = 36-4 = 32 เนื่องจาก T = 26 ไม่น้อยกว่า 4 และไม่มากกว่า 34 จึงปฏิเสธ Ho ไม่ได้

8.4.2 แบบทดสอบวิลคอกซันสำหรับการเท่ากันของความชันในสองเส้นถดถอย (Wilcoxon Test of Equal Slopes for two Regression Lines)

แบบทดสอบวิลคอกซันนี้นอกจากใช้ทดสอบการเท่ากันของมัธยฐาน หรือค่าเฉลี่ยของสองประชากรที่เป็นอิสระกันนั้น ยังใช้ทดสอบการเท่ากันของความชันในสองเส้นถดถอยได้ด้วย และมีกระบวนการในการทดสอบดังนี้

(1) แต่ละกลุ่มคำนวณความชัน U_{1j} และ U_{2t} เช่นเดียวกับแบบทดสอบไธล์

$$U_{1j} = (Y_{1,j+n} - Y_{2j})/(X_{1,j+n} - X_{1j});\ j = 1, 2, .., n$$

$$U_{2t} = (Y_{2,t+n} - Y_{2t})/(X_{2,t+n} - X_{2t});\ t = 1, 2, .., n$$

(2) ใช้แบบทดสอบวิลคอกซันประยุกต์เข้ากับความชัน U_{1j} และ U_{2t} นั้น

ตัวอย่าง ในการศึกษาอายุสมอง (X) และคะแนนเหตุผลทางเลขคณิต (Y) ของชายและหญิง สำหรับเด็กประถม 4 เพื่อทดสอบการเท่ากันของสัมประสิทธิ์ความชัน ($\beta_1 = \beta_2$) ได้ข้อมูลมาดังนี้

ชาย X	Y	X	Y	หญิง X	Y	X	Y
75	15	109	22	76	2	112	25
82	20	112	22	85	10	112	25
83	16	112	23	85	17	113	20
84	10	114	23	92	16	115	22
85	19	115	18	96	11	116	21
90	18	117	23	97	10	118	22
59	15	128	22	99	16	120	17
106	23	130	22	101	16	122	19
107	19	131	25	101	12	132	25

สมมติฐานที่จะทดสอบ

$$Ho : \beta_1 = \beta_2\ ;\ Ha : \beta_1 \neq \beta_2$$

เราคำนวณความชัน U_{1j} และ U_{2t} กับกำหนดอันดับที่ให้แก่ความชันที่รวมกันได้ดังต่อไปนี้

U_{1j}	อันดับที่	U_{2t}	อันดับที่
(22-15)/(109-75) = 0.206	8	(22-2)/(122-76) = 0.611	18
(22-20)/(112-82) = 0.067	4	(25-10)/(112-85) = 0.556	16
(23-16)/(112-83) = 0.241	9.5	(20-17)/(113-85) = 0.107	5
(23-10)/(114-84) = 0.433	14	(22-16)/(115-92) = 0.261	12
(18-19)/(115-85) = -0.033	2	(21-11)/(116-96) = 0.500	15
(23-18)/(117-90) = 0.185	7	(22-10)/(118-97) = 0.571	17
(22-15)/(128-99) = 0.241	9.5	(17-16)/(120-99) = 0.048	3
(22-23)/(130-106) = -0.042	1	(19-16)/(122-101) = 0.143	6
(25-19)/(131-107) = 0.250	11	(25-12)/(132-101) = 0.419	13
T_1 =	66	T_2 =	105

$$T = T_1 - n_1(n_1+1)/2 = 66-9(9+1)/2 = 21$$

สำหรับ $\alpha = 0.05$ เราจะปฏิเสธ Ho ถ้า $T < W_{.025} = 18$ หรือ $T > W_{.975} = 9(9)-18 = 63$

เนื่องจาก T = 21 จึงปฏิเสธ Ho ไม่ได้

8.5 ตัวประมาณค่า และช่วงเชื่อมั่นสำหรับผลต่างระหว่างพารามิเตอร์ความชัน (Estimator and Confidence Interval for difference between Slope parameters)

ฮอลแลนเดอร์ และ วูลฟ์ (Hollander and Wolfe, 1973) ได้เสนอกระบวนการประมาณผลต่างระหว่างความชันของสองเส้นถดถอย หรือ $\theta = \beta_1 - \beta_2$ กระบวนการนี้ขึ้นอยู่กับแบบทดสอบฮอลแลนเดอร์ ซึ่งมีขั้นตอนดังนี้

(1) จากค่า Z ที่หาได้ในแบบทดสอบ เราคำนวณค่าของ U_{ij} ซึ่งได้ n(n+1)/2 ค่า แล้วเรียงค่า U_{ij} นี้จากน้อยไปมากเป็น $W_{(1)} \le W_{(2)} \le \ldots \le W_{(k)}$; k = n(n+1)/2

(2) ตัวประมาณค่าของ $\theta = \beta_1 - \beta_2$ จะเป็น

$$\hat{\theta} = \text{มัธยฐาน } \{U_{ij} ; i \le j\}$$

(3) อาศัยตาราง 24 เราพิจารณาช่วงเชื่อมั่น 100(1-α)% สำหรับ $\theta = \beta_1 - \beta_2$ ได้ดังนี้

$$\hat{\theta}_L \le \theta \le \hat{\theta}_U$$

ในเมื่อ $\hat{\theta}_L = W_{(W\alpha/2)}$ และ $\hat{\theta}_U = W_{(k+1-W\alpha/2)}$

ตัวอย่าง จากตัวอย่างที่แล้วมา เราได้ค่า Z เป็นดังนี้

0.097 -0.017 0.005 0.068 0.092 0.134 -0.075 -0.025

แล้วคำนวณค่า $U_{ij} = (Z_i + Z_j)/2$, i≤j ได้เป็น 36 ค่า ดังนี้

-0.0035	-0.0060	-0.0100	-0.0170	-0.0210	-0.0250	-0.0350
-0.0460	-0.0500	-0.0750	0.0050	0.0085	0.0110	0.0215
0.0255	0.0295	0.0335	0.0360	0.0365	0.0375	0.0400
0.0485	0.0510	0.0545	0.0585	0.0680	0.0695	0.0800
0.0825	0.0920	0.0945	0.0970	0.1010	0.1130	0.1155
0.1340						

ค่าประมาณแบบจุดของ β จะได้เป็น

$$\hat{\beta} = (0.0360 + 0.0365)/2 = 0.03625$$

จากตาราง เมื่อ n = 8 และ α = 0.05 เราได้ $W_{.025}$ = 4 และ k+1-$W_{.025}$ = 8(8+1)/2+1-4 = 33 ดังนั้นช่วงเชื่อมั่น 95% สำหรับ $\theta = \beta_1 - \beta_2$ จะเป็น

$$-0.0170 \le \theta \le 0.1010$$

ในการประมาณค่า $\theta = \beta_1 - \beta_2$ นี้ยังมีกระบวนการที่อาศัยแบบทดสอบวิลคอกซันอีกวิธีหนึ่ง ซึ่งมีกระบวนการดังนี้

(1) หาผลต่าง $U_{ij} = \hat{\beta}_i - \hat{\beta}_j = U_{1i} - U_{2j}$; i = 1, 2, .., n_1 และ j = 1, 2, .., n_2 ในเมื่อ n_1 และ n_2 เป็นค่าความชันที่ประมาณได้จากกลุ่มแรกและกลุ่มที่สองตามลำดับ

(2) เรียงลำดับผลต่าง U_{ij} ได้เป็น $U_{(1)} \le U_{(2)} \le \ldots \le U_{(n_1 n_2)}$

ตัวประมาณค่าแบบจุดของ $\theta = \beta_1 - \beta_2$ จะเป็นมัธยฐานของ $n_1 n_2$ ผลต่าง U_{ij} นั่นคือ

$$\hat{\theta} = \text{มัธยฐาน } \{U_{ij} ; i = 1, 2, .., n_1 \text{ และ } j = 1, 2, .., n_2\}$$

(3) จากตาราง 11 พิจารณา $W_{\alpha/2}$ ขีดจำกัดขั้นต่ำของช่วงเชื่อมั่น ซึ่งจะเป็นผลต่าง U_{ij} ตัวน้อยที่สุดอันดับที่ $W_{\alpha/2}$ หรือ $U_{(W_{\alpha/2})}$ และขีดจำกัดขั้นสูงจะเป็นตัวที่มากอันดับที่ $W_{\alpha/2}$ หรือ $U_{(k)}$; $k = n_1 n_2 + 1 - W_{\alpha/2}$

ตัวอย่าง จากตัวอย่างเกี่ยวกับอายุสมอง และคะแนนเหตุผลของกลุ่มชาย และหญิง เราหาผลต่าง $U_{ij} = \hat{\beta}_i - \hat{\beta}_j = U_{1i} - U_{2j}$

	.048	.107	.143	.261	.419	.500	.556	.571	.611
-.042	.090	.149	.185	.303	.461	.542	.598	.613	.653
-.033	.081	.140	.176	.294	.452	.533	.589	.604	644
.067	-.019	.040	.076	.194	.352	.433	.489	.504	.544
.185	-.137	-.078	-.042	.076	.234	.315	.371	.386	.426
.206	-.158	-.099	-.063	.055	.213	.294	.350	.365	.405
.241	-.193	-.134	-.098	.020	.178	.259	.315	.330	.370
.241	-.193	-.134	-.098	.020	.178	.259	.315	.330	.370
.250	-.202	-.143	-.107	.011	.169	.250	.306	.321	.361
.433	-.385	-.326	-.290	-.172	-.014	.067	.123	.138	.178

ดังนั้น $\hat{\theta} = U_{(41)} = 0.185$

สำหรับ $\alpha = 0.05$ และ $n_1 = n_2 = 9$ เราได้ $W = 18$ และ $k = 9(9)+1-18 = 64$ ดังนั้นช่วงเชื่อมั่น 95% สำหรับ $\theta = \beta_1 - \beta_2$ จะเป็น

$$W_{(18)} \le \theta \le W_{(64)}$$

$$-0.063 \le \theta \le 0.371$$

IT IS NOT ENOUGH TO HAVE A GOOD MIND.
THE MAIN THING IS TO USE IT WELL,
DESCARTES