บทที่ 2

ตัวสถิติ และผลงานวิจัยที่เกี่ยวข้อง

ตัวสถิติที่ใช้ในการทดสอบว่า ประชากรมีการแจกแจงแบบยูนิฟอร์มในช่วง (0,1) นั้นมีหลายรูปแบบ ซึ่งได้มีการพัฒนามาอย่างต่อเนื่อง และมีการเปรียบเทียบอำนาจการทดสอบของตัวสถิติต่างๆ ไว้พอสมควร ในบทนี้จะกล่าวถึงรายละเอียดของตัวสถิติทั้ง 5 ที่ต้องการศึกษาและผลงานวิจัยที่เกี่ยวข้อง

<u>ตัวสถิติที่ใช้ในการทดสอบ</u>

1. ตัวสถิติ X^2

ผู้ริเริ่มใช้การแจกแจงไคสแควร์ในการทดสอบความสอดคล้องระหว่างค่าสังเกต และสมมติฐาน โดยใช้ข้อมูลที่ได้จากการนับ หรือข้อมูลที่อยู่ในรูปของความถี่ คือ Karl Pearson ในปัจจุบันตัวสถิติ X^2 เป็นตัวสถิติทดสอบที่นิยมใช้อย่างแพร่หลายในการทดสอบการแจกแจงของประชากร ไม่ว่าประชากรที่ต้องการทดสอบมีการแจกแจงแบบใดก็ตาม และในการทดสอบว่าประชากรมีการแจกแจงแบบยูนิฟอร์มในช่วง (0,1) นั้นนิยมใช้ตัวสถิติ X^2 เช่นกัน แม้ว่าการทดสอบไคสแควร์จะมีปัญหาในเรื่องเกี่ยวกับขนาดตัวอย่าง และการแบ่งชั้นให้เหมาะสมกับขนาดตัวอย่างก็ตาม

วิธีการทดสอบไคสแควร์ทำได้โดย ขั้นแรกแบ่งกลุ่มข้อมูลตัวอย่างออกเป็นกลุ่มที่แยกจากกันโดยเด็ดขาด หรืออาจแบ่งตามช่วงค่าสังเกต จากนั้นนับจำนวนข้อมูลหรือความถี่ในแต่ละกลุ่ม(Observed Frequency) และใช้หลักการของการแจกแจงแบบยูนิฟอร์มในช่วง(0,1) คำนวณค่าความถี่คาดหวัง(Expected Frequency) ในแต่ละกลุ่มถ้ามีความแตกต่างกันมากระหว่างความถี่สังเกตกับความถี่คาดหวังก็สามารถสรุปได้ว่า ข้อมูลไม่ได้มาจากการแจกแจงแบบยูนิฟอร์มในช่วง (0,1) แต่ถ้าแตกต่างกันน้อยก็สรุปได้ว่า ข้อมูลมาจากการแจกแจงแบบยูนิฟอร์มในช่วง(0,1)

ตัวสถิติ X^2 คำนวณจาก

$$X^2 = \sum_{l=1}^{k} (n_l-np_l)^2/np_l$$

สำหรับการจัดกลุ่มค่าสังเกตของข้อมูลแต่ละชั้นทำได้ดังนี้

ให้ A เป็นเซ็ทใดๆ เขียนแทนด้วย I_A ซึ่งกำหนดดังนี้

$$I_A(x) = \begin{cases} 1 & \text{เมื่อ } x \in A \\ 0 & \text{เมื่อ } x \text{ มีค่าอื่นๆ} \end{cases}$$

ให้ k เป็นเลขจำนวนเต็มบวกใดๆที่มีค่ามากกว่า 1 และกำหนดให้ $p_1, p_2, \ldots, p_k$ เป็นค่าใดๆ ที่ไม่เป็นลบ ซึ่ง $p_1+p_2+\ldots+p_k = 1$

และให้ $$B_l = (x: \sum_{i=1}^{l-1} p_i < x \leqslant \sum_{i=1}^{l} p_i)$$

เมื่อ $l = 1,2,\ldots,k$ และกำหนดให้ $p_0 = 0$ จากนั้นจะได้ว่า

$$n_l = \sum_{i=1}^{n} I_{B_l}(U_i) \quad \text{และ} \quad \sum_{l=1}^{k} n_l = n$$

สำหรับการคำนวณในที่นี้จะกำหนดให้ $p_1=p_2=p_3=\ldots=p_k = 1/k$

โดยที่ X^2 มีการแจกแจงแบบไคสแควร์ที่องศาความเป็นอิสระ (Degree of freedom) เท่ากับ k-1

k = จำนวนกลุ่มของข้อมูล

n = ขนาดตัวอย่าง

n_l = จำนวนความถี่ของข้อมูลในกลุ่มที่ l ซึ่งก็คือ Observed frequency

np_l = ค่าคาดหวังของจำนวนความถี่ของข้อมูลในกลุ่มที่ l ซึ่งก็คือ Expected frequency

สำหรับการคำนวณในที่นี้จะกำหนดให้ k = 5 และ k = 8 ดังนั้นจะได้กลุ่มของค่าสังเกตเป็นดังนี้

<u>เมื่อ k = 5</u> จะได้ว่า $p_1=p_2=p_3=\ldots=p_5 = 1/5$ และจะได้กลุ่มของค่าสังเกต

แบ่งเป็น 5 กลุ่มดังนี้

0.0000 - 0.2000, 0.2001 - 0.4000, 0.4001 - 0.6000, 0.6001 - 0.8000 และ 0.8001 - 1.0000

<u>เมื่อ $k = 8$</u> จะได้ว่า $p_1=p_2=p_3=\ldots=p_8 = 1/8$ ดังนั้นกลุ่มของค่าสังเกตคือ 0.0000 - 0.1250, 0.1251 - 0.2500, 0.2501 - 0.3750, 0.3751 - 0.5000, 0.5001 - 0.6250, 0.6251 - 0.7500, 0.7501 - 0.8750 และ 0.8751 - 1.0000

ขอบเขตวิกฤต จะปฏิเสธสมมติฐานว่าประชากรมีการแจกแจงแบบยูนิฟอร์มในช่วง(0,1) เมื่อค่าสถิติที่คำนวณได้มากกว่าค่า X^2 ที่องศาความเป็นอิสระเท่ากับ k-1

ข้อจำกัดของตัวสถิติ X^2

1). ถ้ามีค่าสังเกตเพียง 2 กลุ่ม ความถี่ที่คาดหวังในแต่ละกลุ่มควรจะมีค่ามากกว่าหรือเท่ากับ 5

2). ถ้ามีค่าสังเกตมากกว่า 2 กลุ่ม วิธีนี้ไม่ควรใช้ ถ้ามีจำนวนกลุ่มของข้อมูลที่มีความถี่ที่คาดหวังน้อยกว่า 5 อยู่มากกว่า 20% ของจำนวนกลุ่มทั้งหมด หรือมีข้อมูลกลุ่มใดกลุ่มหนึ่งที่มีความถี่ที่คาดหวังน้อยกว่า 1

3). ในกรณีที่ความถี่ที่คาดหวังของกลุ่มใดกลุ่มหนึ่งน้อยกว่า 1 ถ้าจะใช้ ตัวสถิติ X^2 ทดสอบจะต้องแก้ไขข้อมูลโดยการรวมกลุ่มที่อยู่ใกล้กันเข้าด้วยกัน เพื่อให้มีความถี่ที่คาดหวังมากพอที่จะทำการทดสอบ ตัวสถิติ X^2 ได้ แต่ต้องระวังว่าการกระทำดังกล่าวจะไม่ทำให้ความหมายของการแบ่งกลุ่มเปลี่ยนไป หรือไม่ขัดกับสมมติฐานที่ผู้วิจัยตั้งไว้

2. ตัวสถิติ Kolmogorov-Smirnov (D)

การทดสอบแบบนี้เป็นวิธีทดสอบการแจกแจงของประชากรอีกวิธีหนึ่ง ซึ่งมีข้อดีกว่าการทดสอบโดยใช้ตัวสถิติ X^2 ตรงที่ว่าสามารถใช้ทดสอบได้กับข้อมูลทุกกรณี แม้ว่าความถี่ของข้อมูลบางกลุ่มจะเป็นศูนย์ก็ตาม สำหรับการทดสอบโดยวิธีนี้จะใช้ความถี่สะสมแทนความถี่ปกติ ไม่ว่าจะเป็นความถี่ที่สังเกตได้หรือความถี่คาดหวัง โดยจะเป็นการเปรียบเทียบความถี่สะสมที่สังเกตได้

กับความถี่สะสมคาดหวังถ้าความแตกต่างนี้มีค่าน้อยมากก็จะสรุปผลว่าข้อมูลที่ทดสอบมีการแจกแจงแบบยูนิฟอร์มในช่วง(0,1) แต่ถ้าความแตกต่างนี้มีค่ามากก็จะสรุปว่าข้อมูลที่ทดสอบไม่มีการแจกแจงแบบยูนิฟอร์มในช่วง(0,1)

สำหรับการทดสอบโดยใช้ตัวสถิติ D ในที่นี้ จะไม่ใช้ความถี่สะสมในการทดสอบดังกล่าว แต่จะใช้ข้อมูลที่มีการจัดเรียงลำดับในการคำนวณค่าสถิติ D และการหาค่าขอบเขตวิกฤตที่มีการปรับปรุง (Modified form) ซึ่งผู้คิดวิธีดังกล่าวนี้ก็คือ M.A. Stephens

ตัวสถิติ D คำนวณได้ดังนี้

$$D = \max_{i=1,2,...,n} \{ \max (|U_{(i)} - (i-1)/n| , |U_{(i)} - i/n|) \}$$

เมื่อ $U_{(i)}$ = ค่าสังเกตที่เรียงลำดับจากน้อยไปมาก (Order Sample), $i = 1,2,3,...,n$

n = ขนาดตัวอย่าง

สำหรับค่าวิกฤตของตัวสถิติ D ในที่นี้ จะใช้ค่าจากตารางของ Pearson and Hartley (1972) (ตารางที่ 3 ในภาคผนวก) ซึ่งจะต้องมีการปรับค่า D ก่อน เรียกว่า Modified form เขียนแทนด้วย T(D) ซึ่ง $T(D) = D(\sqrt{n} + 0.12 + 0.11/\sqrt{n})$ และจะปฏิเสธสมมติฐานว่าง เมื่อ ค่า T(D) ที่คำนวณได้มากกว่าค่าของ Upper percentage points สำหรับ modified T ณ ระดับนัยสำคัญที่กำหนด

3. ตัวสถิติ Watson (U)

ตัวสถิติ Watson (U) คิดขึ้นโดย Watson(1961,1962) ซึ่ง Watson ได้ปรับปรุงจากตัวสถิติ Cramer-von Mises (W) ซึ่งคิดโดย Cramer และ Von Mises ตัวสถิติ W คำนวณได้จาก

$$W = 1/12n + \sum_{i=1}^{n} \{(2i-1)/2n - U_{(i)}\}^2$$

และตัวสถิติ U ได้ปรับปรุงจากตัวสถิติ W ดังนี้

$$U = W - n(\bar{U} - 0.5)^2$$

ดังนั้นจะได้ว่า

$$U = 1/12n + \sum_{i=1}^{n} \{(2i-1)/2n - U_{(i)}\}^2 - n(\bar{U} - 0.5)^2$$

เมื่อ $U_{(i)}$ = order sample ; $i=1,2,3,...,n$

n = ขนาดตัวอย่าง

$$\bar{U} = \sum_{i=1}^{n} U_i/n$$

สำหรับค่าวิกฤตของตัวสถิติ U ก็ต้องใช้ค่าที่ปรับปรุงแล้วเช่นเดียวกันกับตัวสถิติ D นั่นคือต้องใช้ค่าจากตารางของ Pearson and Hartley(1972) และค่าที่ปรับแล้วของตัวสถิติ U เขียนแทนด้วย T(U) ซึ่ง $T(U) = (U - (0.1/n) + (0.1/n^2))(1.0+(0.8/n))$ ดังนั้นจะปฏิเสธสมมติฐานว่าง เมื่อค่า T(U) ที่คำนวณได้มากกว่าค่า Upper percentage points สำหรับ modified T ณ ระดับนัยสำคัญที่กำหนด

4. ตัวสถิติ Anderson - Darling (A)

ตัวสถิติ A คิดขึ้นโดย Anderson-Darling (1953) ซึ่งคำนวณได้ดังนี้

$$A = -1/n \{ \sum_{i=1}^{n} (2i-1) [\ln U_{(i)} + \ln (1- U_{(n+1-i)})] \} - n$$

เมื่อ $U_{(i)}$ = order sample ; $i=1,2,3,...,n$

n = ขนาดตัวอย่าง

ในทำนองเดียวกันค่าวิกฤติของตัวสถิติ A ก็จะใช้ค่าที่ปรับปรุงแล้วและใช้ค่าจากตารางของ Pearson and Hartley (1972) เช่นเดียวกัน สำหรับค่าที่ปรับปรุงแล้วของตัวสถิติ A

เขียนแทนด้วย T(A) ซึ่งค่า T(A) จะมีค่าเท่ากับค่าสถิติ A ที่คำนวณได้ โดยที่ตารางค่าวิกฤต สำหรับตัวสถิติ A นี้จะใช้ได้ เมื่อขนาดตัวอย่างมีค่ามากกว่าหรือเท่ากับ 5 ดังนั้นจะปฏิเสธสมมติฐานว่าง เมื่อค่าสถิติ A ที่คำนวณได้มากกว่า ค่า Upper percentage points สำหรับ modified T ณ ระดับนัยสำคัญที่กำหนด

5. ตัวสถิติ Q

ตัวสถิติ Q คิดขึ้นโดย Miller,F.L. and Quesenberry,C.P.ซึ่งได้ปรับปรุงมาจากตัวสถิติ Greenwood (G) ซึ่งตัวสถิติ G นี้คิดขึ้นโดย Greenwood และคำนวณได้ดังนี้

$$G = \sum_{i=1}^{n+1} [U_{(i)} - U_{(i-1)}]^2$$

ต่อมา Miller,F.L. and Quesenberry,C.P.ได้ปรับปรุงให้มีการเพิ่มข้อมูลบางส่วนเข้าไปในการคำนวณค่าตัวสถิติ G เพื่อให้มีประสิทธิภาพในการทดสอบมากขึ้นเรียกว่าตัวสถิติ Q ซึ่งคำนวณได้ดังนี้

$$Q = \sum_{i=1}^{n+1} (U_{(i)} - U_{(i-1)})^2 + \sum_{i=1}^{n} (U_{(i)} - U_{(i-1)})\,(U_{(i+1)} - U_{(i)})$$

สำหรับค่าวิกฤตของตัวสถิติ Q จะใช้ค่าจากตาราง Percentage points of Q statistic (ตารางที่ 2 ในภาคผนวก) และจะปฏิเสธสมมติฐานว่างเมื่อค่าสถิติ Q ที่คำนวณได้มากกว่าค่า Q จากตาราง ณ ระดับนัยสำคัญที่กำหนด

<u>ผลงานวิจัยที่เกี่ยวข้อง</u>

สำหรับงานวิจัยที่เกี่ยวข้องกัน ได้มีนักสถิติหลายท่านทำการศึกษาอำนาจการทดสอบของตัวสถิติที่ใช้ในการทดสอบการแจกแจงแบบยูนิฟอร์มในช่วง(0,1) ซึ่งในที่นี้จะเสนอเฉพาะผลงาน

วิจัยที่สำคัญเท่านั้น

Stephens,M.A.(1974) ได้ทำการศึกษาอำนาจการทดสอบของตัวสถิติต่างๆ ที่ใช้ในการทดสอบการแจกแจงแบบยูนิฟอร์มในช่วง(0,1) โดยทำการศึกษาตัวสถิติ 5 ตัวคือ

1. ตัวสถิติ Kolmogorov (D^+, D^-, D) ซึ่งคำนวณได้ดังนี้

$$D^+ = \max_{1<=i<=n} |(i/n)-U_i|$$

$$D^- = \max_{1<=i<=n} |U_i - (i-1)/n|$$

$$D = \max (D^+, D^-)$$

2. ตัวสถิติ Cramer-von Mises (W)
3. ตัวสถิติ Watson (U)
4. ตัวสถิติ Anderson-Darling (A)

สำหรับการคำนวณค่าของตัวสถิติ W,U และ A จะเหมือนกันกับตัวสถิติที่ได้กล่าวไว้แล้วในตอนต้น

5. ตัวสถิติ Kuiper (V) ซึ่งคำนวณได้ดังนี้

$$V = D^+ + D^-$$

ในการหาขอบเขตวิกฤตของตัวสถิติทั้ง 5 นี้ก็ใช้ค่าที่ปรับปรุงแล้วจากตารางของ Pearson and Hartley(1972) ดังที่กล่าวมาแล้ว

สำหรับตัวสถิติ V ที่ปรับปรุงแล้วจะเขียนแทนด้วย T(V) ซึ่ง

$T(V) = V(\sqrt{n} + 0.155 + 0.24/\sqrt{n})$ และจะปฏิเสธสมมติฐานว่าง เมื่อค่า T(V) ที่คำนวณได้มากกว่า ค่า Upper percentage points สำหรับ modified T ณ ระดับนัยสำคัญที่กำหนด

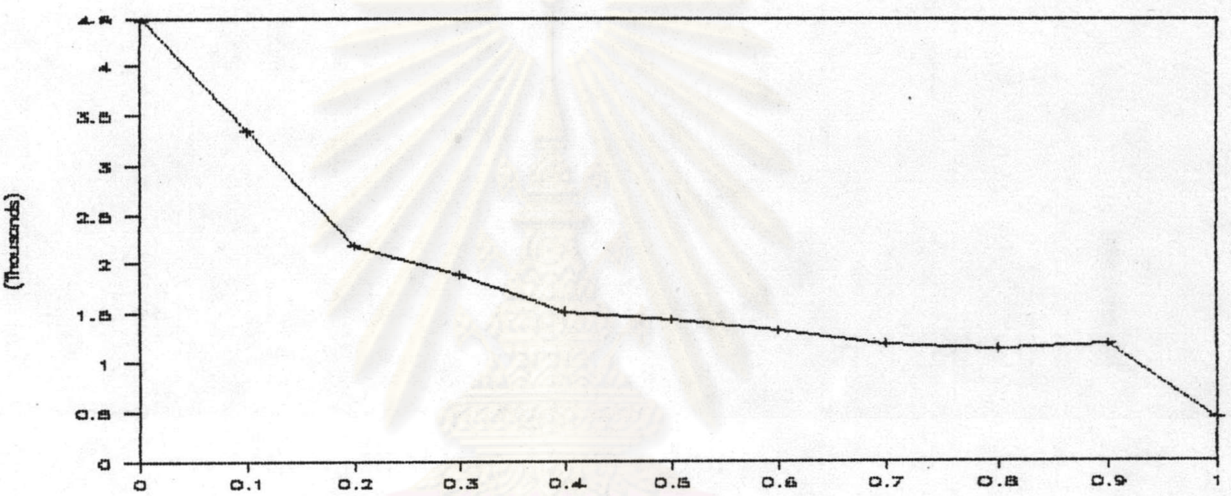
ในการทดสอบอำนาจของตัวสถิติทั้ง 5 นี้ Stephens ได้กำหนดรูปแบบการแจกแจงแบบยูนิฟอร์มในช่วง(0,1) เป็น 3 รูปแบบ เช่นเดียวกันแต่ฟังก์ชั่นที่ใช้แตกต่างกัน โดยกำหนดดังนี้

รูปแบบที่ 1 เมื่อประชากรมีการแจกแจงแบบยูนิฟอร์มในช่วง(0,1) ที่มีค่าเฉลี่ยเปลี่ยนแปลง แต่ค่าความแปรปรวนคงเดิม ซึ่งจะกำหนดฟังก์ชั่นของข้อมูลดังนี้

$$F(U_i) = (1-U_i)^k \quad ; \quad 0 <= U_i <= 1$$

กำหนดให้ $k = 1.5, 2.0$; $i = 1,2,...,n$

กราฟของข้อมูลรูปแบบที่ 1 แสดงได้ดังนี้

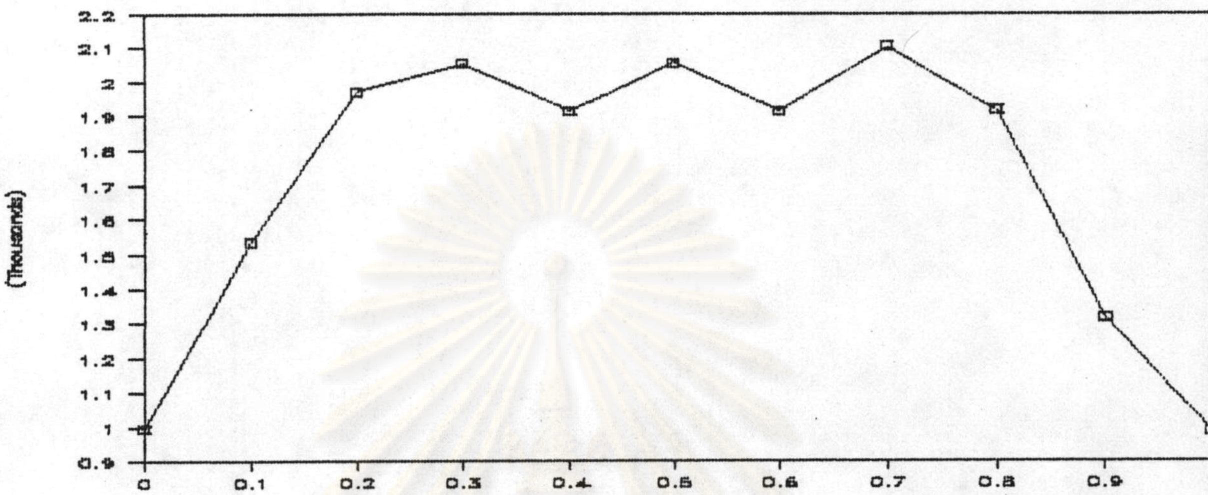


รูปแบบที่ 2 เมื่อประชากรมีการแจกแจงแบบยูนิฟอร์มในช่วง(0,1) ที่มีค่าเฉลี่ยคงเดิมแต่ความแปรปรวนลดลง ซึ่งกำหนดฟังก์ชั่นของข้อมูลดังนี้

$$F(U_i) = \begin{cases} 2^{k-1}U_i^k & ; \ 0<=U_i<=0.5 \\ 1-2^{k-1}(1-U_i)^k & ; \ 0.5<=U_i<=1 \end{cases}$$

กำหนดให้ $k = 1.5, 2.0, 3.0$; $i = 1,2,3,...,n$

กราฟของข้อมูลรูปแบบที่ 2 แสดงได้ดังนี้

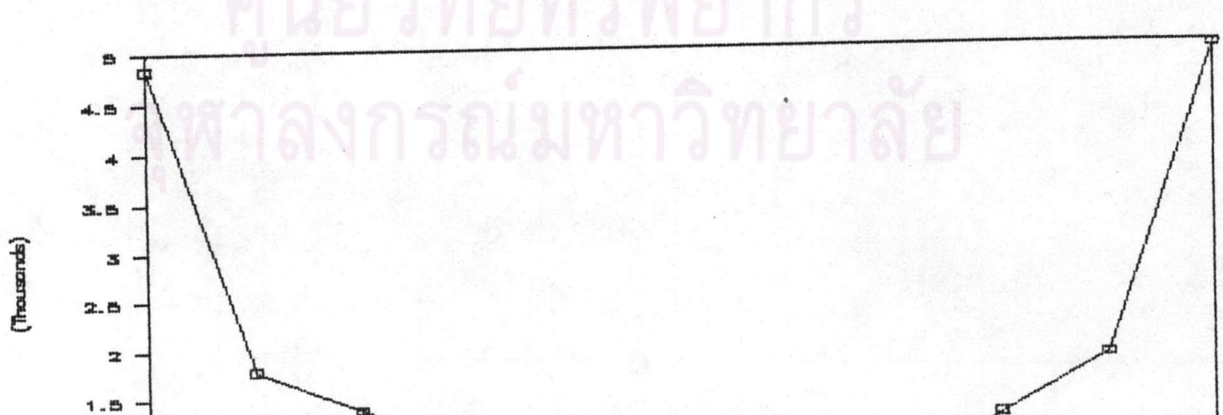


รูปแบบที่ 3 เมื่อประชากรมีการแจกแจงแบบยูนิฟอร์มในช่วง(0,1) ที่มีค่าเฉลี่ยคงเดิมแต่ความแปรปรวนเพิ่มขึ้น ซึ่งกำหนดฟังก์ชั่นของข้อมูลดังนี้

$$F(U_i) = \begin{cases} 0.5 - 2^{k-1}(0.5-U_i)^k & ; \quad 0 <= U_i <= 0.5 \\ 0.5 + 2^{k-1}(U_i-0.5)^k & ; \quad 0.5 <= U_i <= 1 \end{cases}$$

กำหนดให้ $k = 1.5, 2.0$ $\qquad i = 1,2,3,\ldots,n$

กราฟของข้อมูลรูปแบบที่ 3 แสดงได้ดังนี้



สำหรับขนาดตัวอย่างที่ใช้ในการทดสอบเท่ากับ 10 20 และ 40 โดยกำหนดระดับนัยสำคัญ = 0.10 ผลการทดลองสรุปได้ดังนี้

ประชากรรูปแบบที่ 1 ตัวสถิติ W มีอำนาจการทดสอบสูงสุด และตัวสถิติ V มีอำนาจการทดสอบต่ำสุด

ประชากรรูปแบบที่ 2 ตัวสถิติ U มีอำนาจการทดสอบสูงสุด และตัวสถิติ A มีอำนาจการทดสอบต่ำสุด

ประชากรรูปแบบที่ 3 ตัวสถิติ U มีอำนาจการทดสอบสูงสุด และตัวสถิติ W มีอำนาจการทดสอบต่ำสุด