บทที่ 2

วรรณคดีที่เกี่ยวข้อง

การกระจายแบบที (t-distribution)

W.S. Gosset เป็นผู้ค้นพบการกระจายแบบที (**Student's t-distribution**) ในปี ค.ศ. 1908 โดยใช้นามปากกาว่า " **Student**" ที่มาของการกระจายแบบทีเกี่ยวข้องกับการกระจายแบบโค้งปกติและไค-สแควร์ กล่าวคือ ถ้ามีกลุ่มตัวอย่างซึ่งมาจากประชากรที่มีลักษณะการแจกแจงเป็นแบบปกติมีมัชฌิมเลขคณิตเท่ากับ 0 ความแปรปรวนเท่ากับ 1 อยู่กลุ่มหนึ่ง สุ่มตัวอย่างมา 1 ตัวเพื่อหาค่าซี (z) หลังจากนั้นสุ่มตัวอย่างมาอีก 10 ตัวเพื่อหาค่า χ^2_{10} อัตราส่วนของค่าซี (z) และ χ^2_{10} คือ t_{10} นั่นคือ

$$t_{10} = \frac{z}{\sqrt{\frac{\chi^2_{10}}{10}}}$$

และรูปทั่วไปของการกระจายแบบที คือ

$$t_n = \frac{z}{\sqrt{\frac{\chi^2_n}{n}}}$$ เมื่อ n คือ **degrees of freedom**

คุณสมบัติของการกระจายแบบที

1. มีลักษณะสมมาตร
2. ค่าเฉลี่ย มัธยฐาน และฐานนิยมมีค่าเท่ากับ 0
3. ความแปรปรวนเท่ากับ $\frac{n}{n-2}$ (เมื่อ n เป็น **degrees of freedom**)

ส่วนเบี่ยงเบนมาตรฐานจะมีค่ามากกว่า 1 เล็กน้อย ในกรณีที่ n มีค่ามาก ๆ ความแปรปรวนจะมีค่าเท่ากับ 1

4. ถ้า n มีค่ามาก ๆ การกระจายจะมีลักษณะใกล้เคียงหรือเป็นแบบปกติ
5. ${}_{\alpha}t_n = (1 - \alpha)\, t_n$

การทดสอบที (t-test)

W.S. Gosset เป็นผู้สร้างสถิติทดสอบนี้ขึ้นมาเมื่อ ค.ศ. 1908 พร้อมกับการกระจายแบบที และ **R.A. Fisher** เป็นผู้เน้นบทบาทของสถิติทดสอบนี้

ข้อตกลงเบื้องต้น

1. ลักษณะการแจกแจงของประชากรเป็นแบบปกติ
2. กลุ่มตัวอย่างทั้งสองเป็นอิสระต่อกัน (**independent groups**) และเป็นกลุ่มตัวอย่างสุ่ม
3. ตัวอย่างในกลุ่มเป็นอิสระต่อกัน (**independent within samples**)
4. ข้อมูลจัดอยู่ในสเกลวัดได้ (**measurable scale**)
5. ความแปรปรวนของประชากรเท่ากัน (**Homogeneity of variance**)

การคำนวณค่าทดสอบ

ถ้าให้ $X_1, X_2, \ldots, X_m$ และ $Y_1, Y_2, \ldots Y_n$ เป็นกลุ่มตัวอย่างสุ่มมาจากประชากรสองกลุ่มซึ่งมีลักษณะการแจกแจงเป็นแบบปกติ จะคำนวณค่าทีได้จากสูตร

$$t = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sum x^2 + \sum y^2}{(m + n - 2)}\left(\frac{1}{m} + \frac{1}{n}\right)}}$$

โดยที่ $\sum x^2 = \sum (X-\bar{X})^2$

$\sum y^2 = \sum (Y-\bar{Y})^2$

ค่า **t** จะกระจายแบบทีด้วย **m+n-2 degree of freedom**

การทดสอบคอลมอโกรอฟ สเมอร์นอฟ

(Kolmogorov - Smirnov Test)

สถิติเพื่อทดสอบคอลมอโกรอฟ **(Kolmogorov Test)** เป็นสถิติทดสอบนันพาราเมตริกวิธีหนึ่งที่ได้รับการแนะนำจาก **A.N. Kolmogorov** นักคณิตศาสตร์ชาวรัสเซีย เมื่อปี ค.ศ. 1933 สถิติทดสอบนี้อาจเรียกว่า **The Kolmogorov Goodness-of-fit Test** หรือในปัจจุบันนี้เรียกว่า **The Kolmogorov - Smirnov One-Sample Test (Daniel, 1978: 267)** จุดประสงค์ของวิธีการทดสอบ **The Kolmogorov Goodness-of-fit** นี้เพื่อดูว่าการแจกแจงของข้อมูลที่สังเกตได้แตกต่างจากการแจกแจงตามทฤษฎี (หรือที่คาดหวังไว้ตามสมมติฐาน) หรือไม่ โดยมีข้อตกลงของการแจกแจงความน่าจะเป็นเชิงทฤษฎี **(The theoretical probability distribution)** ต้องเป็นแบบต่อเนื่อง **(continuous)** และกลุ่มตัวอย่างต้องได้มาจากกลุ่มตัวอย่างสุ่ม สามารถคำนวณได้จากสูตร

$$K = \max \left| P(X < x_i) - \hat{P}(X < x_i) \right|$$

เมื่อ $P(X < x_i)$ เป็น **cumulative probabilities function** หรือ **a hypothesized cumulative distribution**

$\hat{P}(X < x_i)$ เป็น **cumulative relative frequencies** ของกลุ่มตัวอย่าง หรือ **the observed cumulative distribution**

สำหรับสถิติทดสอบของ **Kolmogorov** นี้เป็นสถิติทดสอบที่ใช้สำหรับคำนวณหา **maximum vertical distribution** นั่นเอง (CONOVER,1980: 345)

ต่อมาเมื่อปี ค.ศ. 1939 นักคณิตศาสตร์ชาวรัสเซียอีกผู้หนึ่งชื่อ **N.V. Smirnov** ได้พัฒนาวิธีการทดสอบของ **Kolmogorov** เพิ่มเติมเพื่อใช้ทดสอบเปรียบเทียบกลุ่มตัวอย่างสุ่ม 2 กลุ่ม โดยมีสมมติฐานว่ากลุ่มตัวอย่างทั้ง 2 กลุ่ม มาจากประชากรที่เหมือนกัน หรือประชากรเดียวกันหรือไม่ ซึ่งมีลักษณะวิธีการคำนวณที่คล้ายคลึงกับ **The Kolmogorov - Gooness-of-fit Test** โดยมีชื่อเรียกว่า **The Kolmogorov - Smirnov two - sample Test (K-S Test)** เขาทั้งสองได้ให้สมมติฐานการแจกแจงความน่าจะเป็น **(probability distribution)** ภายใต้การทดสอบแบบต่อเนื่อง **(continuous)**

และเมื่อทราบค่าเฉลี่ย (**mean**) และความแปรปรวน (**variance**) ของประชากร (**Shannon, 1975: 78)**

<u>ข้อตกลงเบื้องต้นสำหรับ **Kolmogorov - Smirnov two - sample test**</u>

(Marascuilo, 1977: 253)

1. กลุ่มตัวอย่างทั้งสองเป็นอิสระต่อกัน (**independent groups**) และเป็นกลุ่มตัวอย่างสุ่ม

2. ตัวอย่างในกลุ่มเป็นอิสระต่อกัน (**independent within samples**)

3. กลุ่มตัวอย่างสุ่มมาจากการแจกแจงแบบต่อเนื่อง (**continuous distribution**)

<u>การคำนวณค่าทดสอบ</u>

ถ้าให้ $X_{(1)}, X_{(2)}, \ldots, X_{(m)}$ และ $Y_{(1)}, Y_{(2)}, \ldots, Y_{(n)}$ เป็นค่าของตัวแปรจากกลุ่มตัวอย่างสุ่มมาจากประชากรสองกลุ่ม จะคำนวณค่าทดสอบคอลมอโกรอฟ สเมอร์นอฟ ได้จากสูตร

$$KS = \max \left| \hat{P}(X_1 < x) - \hat{P}(X_2 < x) \right|$$

เมื่อ $\hat{P}(X_1 < x)$ เป็น **cumulative probability** ของกลุ่มตัวอย่างที่ 1

$\hat{P}(X_2 < x)$ เป็น **cumulative probability** ของกลุ่มตัวอย่างที่ 2

<u>การทดสอบคอลมอโกรอฟ สเมอร์นอฟ เมื่อกลุ่มตัวอย่างมีขนาดใหญ่</u>

ถ้าจำนวนตัวอย่างในแต่ละกลุ่มมากกว่า 40 จำนวนขึ้นไปถือว่าเป็นกลุ่มตัวอย่างที่มีขนาดใหญ่ ซึ่งสามารถคำนวณหาค่าของคอลมอโกรอฟ สเมอร์นอฟ ได้จากสูตร (**Goodman, 1954)**

$$\chi^2 = 4D^2 \frac{n_1 n_2}{n_1 + n_2} \quad ; \ df = 2$$

$$D = KS = \max \left| \hat{P}(X_1 < x) - \hat{P}(X_2 < x) \right|$$

n_1 คือ จำนวนตัวอย่างของกลุ่มตัวอย่างที่ 1

n_2 คือ จำนวนตัวอย่างของกลุ่มตัวอย่างที่ 2

ซึ่งค่า χ^2 สามารถหาได้จากตารางแจกแจงไค-สแควร์ด้วย 2 **degree of freedom**

คอนโนเวอร์ (CONOVER, 1980: 371,473) กล่าวว่า กรณีกลุ่มตัวอย่างในแต่ละกลุ่มไม่เท่ากัน โดย n_1 มีขนาดของกลุ่มตัวอย่างเล็กกว่า n_2 ซึ่งมีขนาดของกลุ่มตัวอย่างมากกว่า ถ้าทั้ง m และ n เป็นขนาดของกลุ่มตัวอย่างทั้ง 2 กลุ่ม ไม่สามารถครอบคลุมในตารางค่าวิกฤตได้ให้ใช้ **the large sample approximation** แทน โดยใช้สูตร (**Massey,1952**)

$$W_{.95} = 1.36\sqrt{\frac{m+n}{mn}} \quad \text{ที่ระดับ } \alpha = .05$$

และ

$$W_{.99} = 1.63\sqrt{\frac{m+n}{mn}} \quad \text{ที่ระดับ } \alpha = .01 \text{ เป็นต้น}$$

<u>งานวิจัยที่เกี่ยวข้อง</u>

Shannon (1975: 78-80) ได้ศึกษาถึงการทดสอบไค-สแควร์ และการทดสอบคอลมอโกรอฟ สเมอร์นอฟ ในกรณีของกลุ่มตัวอย่างขนาดเล็กมาก ๆ ปรากฏว่า การทดสอบของไค-สแควร์ไม่สามารถจะทำการทดสอบตามคุณสมบัติของตัวมันเองได้ทั้งหมด แต่การทดสอบคอลมอโกรอฟ สเมอร์นอฟ สามารถจะนำมาใช้ทดสอบได้แม้ว่าขนาดของกลุ่มตัวอย่างจะมีขนาดเล็กมาก ซึ่งเมื่อใช้ไค-สแควร์ทดสอบจะทำให้อำนาจของการทดสอบไม่สมบูรณ์ แต่ถ้านำไค-สแควร์ไปทดสอบกับกลุ่มตัวอย่างที่มีขนาดใหญ่ขึ้นผลของอำนาจของการทดสอบจะดีขึ้น **Shannon** ได้ให้เหตุผลว่า การทดสอบแต่ละอย่างมีทั้งความแข็งแกร่ง (**Strengths**) และความอ่อน (**Weakness**) ที่แตกต่างกัน โดยเราสามารถจะเลือกใช้ได้ด้วยการดูคุณลักษณะของเทสนั้น ๆ สำหรับไค-สแควร์ เทส นั้นจะมีอำนาจของการทดสอบสูงเมื่อกลุ่มตัวอย่างมีขนาดใหญ่มาก ๆ คือ เมื่อ $n \geqslant 100$ แม้ว่าจะมีผู้วิจัยบางคนชี้แจงว่าการทดสอบคอลมอโกรอฟ สเมอร์นอฟ จะให้ผลได้ดีนั้น $n \geqslant 30$ แต่ก็ยังมีผู้ชี้แจงเหตุผลอีกไม่ใช่น้อยที่ให้ใช้คอลมอโกรอฟ สเมอร์นอฟ เทส เมื่อ $99 \geqslant n \geqslant 10$ สำหรับขนาดของกลุ่มตัวอย่างที่ต่ำกว่า 10 ควรจะใช้ **Cramer-Von Mises test** จะได้ผลดีที่สุด

สำหรับ **Lilliefors** (1967: 399-402) ก็เป็นอีกผู้หนึ่งที่ทำการศึกษาเกี่ยวกับคอลมอโกรอฟ สเมอร์นอฟ เทส และไค-สแควร์ เทส กับข้อมูลที่มีลักษณะการแจกแจงแบบต่อเนื่องและเป็นข้อมูลที่สังเกตได้ (**observation**) ปรากฏผลว่าคอลมอโกรอฟ สเมอร์นอฟ เทส สามารถถูกนำมาใช้ทดสอบได้แม้ขนาดของกลุ่มตัวอย่างจะเล็ก และยังมีอำนาจของการทดสอบมากกว่าไค-สแควร์ เทส แม้จะนำไปทำการทดสอบกับขนาดของกลุ่มตัวอย่างขนาดอื่น

0.56 1.6

0.00

Williams (1950: 77-86) ก็เป็นอีกผู้หนึ่งที่ศึกษาในกรณีกลุ่มตัวอย่างน้อยมาก ๆ พบว่า ไค-สแควร์ เทส ไม่สามารถจะนำมาใช้ได้เนื่องจากอาจจะทำให้เกิดความคลาดเคลื่อนได้พร้อมกับจะต้องมีวิธีการในการทดสอบมากขึ้น ดังนั้นไค-สแควร์ เทส จึงควรใช้กับขนาดของกลุ่มตัวอย่างที่ใหญ่ขึ้น ซึ่งกรณีกลุ่มตัวอย่างน้อย ๆ ถ้าใช้คอลมอโกรอฟ สเมอร์นอฟ เทส จะทำให้อำนาจของการทดสอบสูงกว่าไค-สแควร์ เทส

Slakter(1965: 854-858) ได้ศึกษาเปรียบเทียบระหว่างไค-สแควร์ เทส และคอลมอโกรอฟ สเมอร์นอฟ เทส ซึ่งเป็น **Goodness-of-fit test**ทั้งคู่ภายใต้เงื่อนไขดังนี้

1. จำนวนข้อมูลที่สังเกตได้ (**Observations**) **N** จำนวนต้องเป็นอิสระซึ่งกันและกัน ในขณะที่ **k** กลุ่มต้องแยกจากกันโดยเด็ดขาด แต่โอกาสภายใต้สมมติฐานในการทดสอบควรจะเท่ากัน

2. ทั้ง **N** และ **k** ต้องมีขนาดเล็กไม่ควรมากกว่า 50

จากกลุ่มตัวอย่างที่ถูกสุ่มมาทำการทดลองปรากฏผลว่า คอลมอโกรอฟ สเมอร์นอฟ เทส มีอำนาจของการทดสอบดีกว่าไค-สแควร์ เทส เมื่อขนาดของกลุ่มตัวอย่างเล็ก ๆ ขณะที่ไค-สแควร์ เทส ต้องมีจำนวนของข้อมูลที่สังเกตได้มากพอที่จะกาดประมาณค่าได้ในการทดสอบทางสถิติ

Dixon (1954) ได้ทำการศึกษาเปรียบเทียบเค-เอส เทส (**K-S test**) กับแรงค์ ซัม เทส (**rank sum test**) , มีเดียน เทส (**median test**) , รันส์ เทส (**runs test**) และ ที เทส (**t-test**) แบบสองทาง ในกรณีที่มีลักษณะเป็นแบบปกติ (**normal shift alternatives**) ผลปรากฏว่าดิสตรีบิวชั่นฟรี (**distribution-Free**) ทั้ง 4 วิธี มีอำนาจในการทดสอบสูงกว่าการทดสอบทีเมื่อกลุ่มตัวอย่างมีขนาดเล็ก ซึ่งในการศึกษาเปรียบเทียบนี้ใช้ขนาดของกลุ่มตัวอย่างเท่ากับ 5 มีลักษณะเป็นแบบปกติ ที่ระดับนัยสำคัญ 0.025 สรุปได้ว่าการทดสอบโดยใช้ลำดับ (**rank test**) มีอำนาจของการทดสอบดีกว่าเค-เอส เทส ส่วนเค-เอส เทส จะมีอำนาจการทดสอบดีกว่ามีเดียน เทส

Klotz (1967) ได้ศึกษาเปรียบเทียบกับกลุ่มตัวอย่าง 2 กลุ่ม ซึ่งมีขนาดเท่ากันโดยทำการเปรียบเทียบเค-เอส เทส (**K-S test**) กับที เทส (**t-test**) และเอฟ เทส (**F-test**) เฉพาะลักษณะแบบปกติเท่านั้น ให้ช่วงของความเบี่ยงเบนมาตรฐานจาก 0 ถึง 3 ผลปรากฏว่าประสิทธิภาพของเค-เอส เทส สูงกว่า ที เทส ระหว่าง 0.63 และ 0.76 เมื่อขนาดของกลุ่มตัวอย่างเล็ก แต่เอฟ เทส จะมีประสิทธิภาพสูงกว่าเค-เอส เทส

Sawat Pratoomraj (1970: **abstract**) ได้ทำการวิจัยและพบว่า ถ้ากลุ่มตัวอย่างสุ่มมาจากประชากรที่มีลักษณะการแจกแจงเป็นแบบเดียวกันแล้ว การทดสอบที และการทดสอบของแมน-วิทนี (เป็นรูปแบบหนึ่งของการทดสอบของวิลคอกซอน ซึ่งปรับปรุงเป็นอิสระในปี ค.ศ. 1947) มีอำนาจของการทดสอบสูงที่สุดในประเภทของการทดสอบทั้ง 4 คือ การทดสอบที การทดสอบของแมน-วิทนี (**Mann-Witney U Test**) การทดสอบของเวลช์ (**Welch Test**) และการทดสอบซี (**Z-Test**) เมื่อความแปรปรวนของประชากรเท่ากัน

Neave และ**Granger** (1968: 509-521) ได้ทำการศึกษาวิจัยวิธีการทดสอบถึง 8 วิธี ได้แก่ ที เทส (**t-test**) , วิลคอกซอน เทส (**Wilcoxon test**) นอร์มอล สกอร์ เทส (**Normal Scores test**) , คอลมอโกรอฟ สเมอร์นอฟ เทส (**Kolmogorov-Smirnov test**) , รันส์ เทส (**Runs test**) , มีเดียน เทส (**Median test**) , ทูกี้ ควิค เทส (**Tukey Quick test**) และอิมพรูฟ ทูกี้ ควิค เทส (**Improved Tukey Quick test**) เมื่อค่ามัชฌิมเลขคณิตในการทดสอบแตกต่างกันภายใต้สถานการณ์ของการซิมูเลท เมื่อประชากรมีลักษณะการแจกแจงเป็นแบบปกติ (**normally distributed**) และกลุ่มตัวอย่างมีขนาด $n_1 = n_2 = 20$ และ $n_1 = 20$, $n_2 = 40$ ตามลำดับ ผลปรากฏว่า รันส์ เทส (**Runs test**) และเค-เอส เทส (**K-S test**) ถูกพบว่ามีประโยชน์น้อยมากเมื่อทดสอบเกี่ยวกับความแตกต่างของค่ามัชฌิมเลขคณิต (**mean**) ซึ่ง ควิค เทส (**Quick test**) ได้รับพิจารณาว่าดีกว่า แต่สิ่งที่มีอำนาจในการทดสอบเหนือกว่าในการค้นพบ คือ นอร์มอล สกอร์ เทส (**Normal Scores test**) ให้ผลเป็นที่น่าพอใจที่สุด รองลงมาก็คือ วิลคอกซอน แรงค์ ซัม เทส (**Wilcoxon rank-sum test**) แม้ว่าประชากรจะมีลักษณะการแจกแจงเป็นแบบปกติก็ตาม การทดสอบทั้ง 2 วิธีนี้ก็อำนาจของการทดสอบด้อยกว่าที เทส (**t-test**) เพียงเล็กน้อย แต่อาจจะดีกว่าเมื่อประชากรมีลักษณะการแจกแจงไม่เป็นแบบปกติ (**non - normal populations**)

จากงานวิจัยของไปรมา (2526) ใช้เทคนิคมอนติคาร์โลศึกษาเปรียบเทียบอำนาจของการทดสอบที , การทดสอบของวิลคอกซอน , การทดสอบเทอรี่-โฮฟดิ้ง นอร์มอล-สกอร์ และการทดสอบแวน เดอ แวร์เดน นอร์มอล-สกอร์ เมื่อกลุ่มตัวอย่างสุ่มมาจากประชากรที่มีลักษณะการแจกแจงเป็นแบบปกติ แบบยูนิฟอร์ม และแบบโลจิสติค

ได้ข้อสรุปว่า การทดสอบที่มีอำนาจของการทดสอบสูงที่สุดในการทดสอบทั้ง 4 วิธี เมื่อลักษณะข้อมูลที่ใช้ในการทดสอบอยู่ในมาตราอันตรภาค (**interval scale**) หรือมาตราอัตราส่วน (**ratio scale**) ส่วนการทดสอบของเทอรี่-โฮฟฟ์ดิ้ง นอร์มอล-สกอร์ และการทดสอบของแวน เดอ แวร์เดน นอร์มอล-สกอร์ มีอำนาจของการทดสอบเหนือกว่าการทดสอบของวิลค็อกซอน เมื่อลักษณะข้อมูลที่ใช้ในการทดสอบอยู่ในรูปของอันดับ (**rank**)