บทที่ 3

วิธีการดำเนินงาน

แยกเป็นขั้นตอนต่างๆดังนี้

1. **สร้างข้อมูล** โดยวิธีจำลองข้อมูล ลงในตาราง 2x5 ซึ่งมี 2 แถวนอน หมายถึง 2 กลุ่มตัวอย่างที่ถูกสุ่ม จาก 2 ประชากร และ 5แถวตั้ง ซึ่งหมายถึง ลำดับที่ 1-5 เป็นค่าตัวแปรตามที่มีคำตอบแบบเรียงลำดับได้ 5 ระดับ โดยในแต่ละแถวนอน มีผลรวม = ขนาดตัวอย่างของแต่ละประชากร ซึ่งแยกเป็น 3 กรณี คือ

ขนาดตัวอย่างเท่ากัน คือ 50-50

ขนาดตัวอย่างต่างกันเล็กน้อย คือ 50-70

ขนาดตัวอย่างต่างกันมาก คือ 50-100

และ**จัดความถี่ลงในเซลต่างๆ** ทั้งหมด 10 เซล โดยแยกเป็น 3 กรณี คือ

ความแตกต่างของค่าสัดส่วนในแต่ละแถวตั้ง ต่างกันเล็กน้อย คือ < 0.1

ความแตกต่างของค่าสัดส่วนในแต่ละแถวตั้ง ต่างกันปานกลาง คือ ≥ 0.1 - ≤ 0.2

ความแตกต่างของค่าสัดส่วนในแต่ละแถวตั้ง ต่างกันมาก คือ > 0.2

และค่าความถี่คาดหวัง (E_{ij}) ที่จะคำนวณได้ในแต่ละเซลให้มีค่าน้อยกว่า 5 ได้อย่างมากที่สุดเพียง 2 เซล เท่านั้น เพื่อให้สามารถใช้การทดสอบแบบไคสแควร์ได้ถูกต้อง

รวมทั้งหมดแล้วจะมี 9 กรณี

เช่น กรณีที่1 ขนาดตัวอย่าง 50-50 และมีความแตกต่างกันน้อยในแต่ละแถวตั้ง สร้างข้อมูลลงในเซลทั้ง 10 เซล ให้มีลักษณะดังกล่าว เป็นจำนวนทั้งหมด 100 ชุดตัวอย่าง

อาจแสดงตัวอย่าง 1 ชุด ได้ดังต่อไปนี้

6	11	10	10	13	รวม =50
9	8	11	12	10	รวม =50

ซึ่งสัดส่วนในแต่ละแถวตั้งจะมีค่า < 0.1 ดังรายละเอียดต่อไปนี้

แถวตั้งที่ 1 สัดส่วนทั้งสองคือ $\frac{6}{50}$ และ $\frac{9}{50}$ ตามลำดับ

ดังนั้น ผลต่างสัดส่วน = 0.06 ซึ่ง < 0.1

และแถวตั้งที่ 2 สัดส่วนทั้งสอง $\frac{11}{50}$ คือ และ $\frac{8}{50}$ ตามลำดับ

ผลต่างสัดส่วน $= \frac{11}{50} - \frac{8}{50} = 0.06$ ซึ่ง < 0.1

ทำนองเดียวกันในแถวตั้งที่ 3-5

ในแต่ละแถว จะได้ผลต่างค่าสัดส่วน < 0.1

สรุปได้ว่า ข้อมูลชุดนี้เป็นชุดตัวอย่างที่มีขนาดตัวอย่างเท่ากัน และความแตกต่างของค่าสัดส่วนในแต่ละแถวตั้ง มีค่าน้อย (คือ < 0.1)

สร้างข้อมูลให้มีลักษณะเช่นเดียวกันนี้ แต่เป็นชุดอื่นๆ อีก 99 ชุด รวมทั้งหมด ได้ข้อมูล 100 ชุดตัวอย่าง

ทำนองเดียวกันนี้ สร้างข้อมูลให้มีลักษณะต่างๆให้ครบ 9 กรณีดังกล่าวข้างต้น โดยกรณีหนึ่งๆ จะมี 100 ชุดตัวอย่าง

รวมทั้งสิ้นจะได้ข้อมูล ตัวอย่าง 900 ชุดตัวอย่าง

2. นำข้อมูลที่ได้ไปวิเคราะห์ด้วยสถิติทดสอบ WMW แบบมี ties มาก และแบบไคสแควร์

จากข้อมูลชุดหนึ่งๆ ที่ได้ในขั้นที่ 1 ให้คำนวณค่าสถิติทดสอบ WMW แบบมี ties มาก และ χ^2 เช่นจากข้อมูลที่ยกตัวอย่างในขั้นตอนที่ 1 คำนวณหาค่าสถิติ WMW แบบมี ties มาก และ χ^2 ได้ดังรายละเอียด ต่อไปนี้

	1	2	3	4	5	รวม
ตัวอย่างชุด 1	6	11	10	10	13	50
ตัวอย่างชุด 2	9	8	11	12	10	50

สถิติ WMW ขั้นตอนแรกหาลำดับที่ก่อน ดังนี้

ค่า 1 มีซ้ำทั้งหมด 15 ค่า (6+9) ดังนั้น ลำดับที่ คือ $\dfrac{1+2+\ldots+15}{50}=\dfrac{120}{15}=8$

ค่า 2 มีซ้ำทั้งหมด 19 ค่า (11+8) ลำดับ คือ $\dfrac{16+17+\ldots+34}{19}=\dfrac{475}{19}=25$

ค่า 3 มีซ้ำทั้งหมด 21 ค่า (10+11) ลำดับ คือ $\dfrac{35+36\ldots+55}{21}=\dfrac{945}{21}=45$

ค่า 4 มีซ้ำทั้งหมด 22 ค่า (10+12) ลำดับ คือ $\dfrac{56+57\ldots+77}{22}=\dfrac{1463}{22}=66.5$

ค่า 5 มีซ้ำทั้งหมด 23 ค่า (13+10) ลำดับ คือ $\dfrac{78+79\ldots+100}{23}=\dfrac{2047}{23}=89$

หาค่า S จากตัวอย่างชุด 1 จะได้ $S = 6(8) + 11(25) + 10(45) + 10(66.5) + 13(89)$

$= 48 + 275 + 450 + 665 + 1157$

$= 2595$

ดังนั้น สถิติทดสอบ T $= 2595-\dfrac{50(51)}{2} = 1320$

ตัวอย่างมีขนาดใหญ่ ใช้สถิติ Z จะได้ $Z_{cal} = \dfrac{1320-\dfrac{(50)(50)}{2}}{\sqrt{\dfrac{50(50)(101)}{12}}} = \dfrac{70}{145.05} = 0.4927$

เมื่อใช้ระดับนัยสำคัญ 0.05 อาณาเขตวิกฤต คือ $Z > 1.96$ หรือ $Z < -1.96$

ดังนั้น Z_{cal} ไม่ตกในอาณาเขตวิกฤต จึงยอมรับ H_0

นั่นคือ ค่ากลางของ 2 ประชากรนี้ ไม่ต่างกัน

หรือถ้าใช้ระดับนัยสำคัญ 0.10 อาณาเขตวิกฤต คือ $Z > 1.645$ หรือ $Z < -1.645$

ดังนั้น Z_{cal} ไม่ตกในอาณาเขตวิกฤต จึงยอมรับ H_0

นั่นคือ ค่ากลางของ 2 ประชากรนี้ ไม่ต่างกัน

ส่วนสถิติทดสอบ χ^2 จากสูตร $\chi^2_{cal} = \displaystyle\sum_{i=1}^{2}\sum_{j=1}^{5}\frac{(O_{ij}-E_{ij})^2}{E_{ij}}$

โดย O_{ij} คือ ความถี่ในเซลต่างๆ ทั้ง 10 เซล

E_{ij} คำนวณจากสูตร $\frac{R_i C_i}{N}$ เมื่อ R_i = ผลรวมความถี่ในแถวนอน i

C_i=ผลรวมความถี่ในแถวตั้ง j

และ N = ผลรวมความถี่ทั้งหมด

จะได้ค่า E_{ij} ดังตาราง ต่อไปนี้

7.5	9.5	10.5	11	11.5
7.5	9.5	10.5	11	11.5

ดังนั้น $\chi^2_{cal} = \frac{(6-7.5)^2}{7.5} + \frac{(11-9.5)^2}{9.5} + \ldots + \frac{(10-11.5)^2}{11.5}$

$= 1.694$

ในขณะที่ $\chi^2_{วิกฤต}$ คือ $\chi^2_{0.05,4} = 9.49$ **เมื่อใช้ระดับนัยสำคัญ 0.05**

ดังนั้น χ^2_{cal} ไม่ตกในอาณาเขตวิกฤต จึงยอมรับ H_0

หรือที่ $\chi^2 = 1.694$ ได้ค่า p(p-value) ประมาณ 0.08

ถ้าใช้ระดับนัยสำคัญ 0.10 จะได้ค่าวิกฤตคือ 7.78 ดังนั้น χ^2_{cal} ไม่ตกในอาณาเขตวิกฤต จึงยอมรับ H_0

สรุปได้ว่า ข้อมูลตัวอย่างชุดนี้ เมื่อวิเคราะห์ด้วยสถิติ WMW และ χ^2 ได้ผลเหมือนกัน คือ ยอมรับ H_0 ทั้งที่ระดับนัยสำคัญ0.05 และ 0.10

สรุปในขั้นตอนนี้ จะได้ค่าสถิติทดสอบ WMW และ χ^2 จากแต่ละชุดตัวอย่างรวมทั้งหมด 900 ชุดตัวอย่าง

3. คำนวณค่า p-value จากการใช้สถิติทดสอบทั้ง 2 เพื่อหาผลสรุป

(ยอมรับหรือปฏิเสธสมมติฐานเบื้องต้น) เมื่อกำหนดระดับนัยสำคัญ 0.05 และ 0.10

นำค่า p-value ที่ได้เทียบกับค่า 0.05 และ 0.10 ถ้า น้อยกว่า →ปฏิเสธสมมติฐานเบื้องต้น

ถ้า มากกว่า →ยอมรับสมมติฐานเบื้องต้น

ดังนั้น**จากข้อมูลชุดหนึ่งๆ** จะได้ข้อสรุปว่า

จากการใช้สถิติ WMW ผลสรุปที่ได้คือ ยอมรับ หรือปฏิเสธสมมติฐานเบื้องต้น และจากการใช้สถิติไคสแควร์ผลสรุปที่ได้คือ ยอมรับ หรือปฏิเสธสมมติฐานเบื้องต้น

ในแต่ละกรณี ซึ่งมีข้อมูลตัวอย่าง 100 ชุด ให้นับจำนวนครั้งที่ปฏิเสธสมมติฐานเบื้องต้น จากการใช้สถิติ WMW แบบมี ties มาก และ ไคสแควร์ และนับมาเปรียบเทียบกัน

เช่นในกรณีที่ 1 ที่กล่าวตอนต้น จะได้ จำนวนครั้งของการปฏิเสธสมมติฐานเบื้องต้นดังตาราง

χ^2	WMW
0	0

4. ทำการอนุมานถึงกลุ่มประชากร

โดยการหาผลสรุปว่า ในระดับประชากรสถิติทั้ง 2 จะให้ผลสรุปต่างกันหรือไม่ โดยการทดสอบสมมติฐานเกี่ยวกับผลต่างค่าสัดส่วนของ 2 ประชากร ดังรายละเอียด ต่อไปนี้

$$H_0 : \mathrm{P}_1 = \mathrm{P}_2$$

$$H_1 : \mathrm{P}_1 \neq \mathrm{P}_2$$

เมื่อ P_1 = สัดส่วนของการปฏิเสธสมมติเบื้องต้นจากการใช้สถิติ WMW

P_2 = สัดส่วนของการปฏิเสธสมมติฐานเบื้องต้นจากการใช้สถิติไคสแควร์

สถิติทดสอบคือ $$Z = \frac{\hat{\mathrm{P}}_1 - \hat{\mathrm{P}}_2}{\sqrt{\hat{\mathrm{P}}\hat{\mathrm{Q}}\left(\frac{1}{\mathrm{n}_1} - \frac{1}{\mathrm{n}_2}\right)}}$$

เมื่อ

$\hat{\mathrm{P}}_1$ = สัดส่วนของการปฏิเสธสมมติฐานเบื้องต้น จากสถิติ WMW แบบมี tiesมาก จากตัวอย่าง 100 ชุด

$\hat{\mathrm{P}}_2$ = สัดส่วนของการปฏิเสธสมมติฐานเบื้องต้น จากสถิติ ไคสแควร์ จากตัวอย่าง 100 ชุด

$\hat{P}$ = ค่าประมาณของ $P_1 = P_2$ คิดจาก $\dfrac{X_1 + X_2}{n_1 + n_2}$

จากกรณีที่ 1 ซึ่งได้ผลสรุปของจำนวนครั้งของการปฏิเสธในขั้นตอนที่ 3 คือ

χ^2	WMW
0	0

$$\text{ดังนั้น } Z = \frac{\dfrac{0}{100} - \dfrac{0}{100}}{\sqrt{\left(\dfrac{0-0}{100-100}\right)\left(\dfrac{100+100}{100+100}\right)\left(\dfrac{1}{100} + \dfrac{1}{100}\right)}}$$

$$= 0$$

อาณาเขตวิกฤต คือ $Z < -1.96$ หรือ $Z > 1.96$

ดังนั้นสรุปได้ว่า ยอมรับ H_0 หรือหาค่า p(p-value) = $2P(Z \geq 0)$ = 1

คือ สัดส่วนของการปฏิเสธ H_0 จากสถิติทั้งสอง คือ WMW กับ χ^2 ไม่ต่างกัน

ซึ่งจะได้ผลสรุปใน 9 กรณี ทั้งหมด ดังจะนำเสนอในบทที่ 4 ต่อไป