บทที่ 6

การวิเคราะห์ข้อมูลด้วยสถิติอนุมาน

วัตถุประสงค์

1) เพื่อให้ผู้ศึกษา **มีความรู้และความเข้าใจ** เกี่ยวกับ **การวิเคราะห์ข้อมูลด้วยสถิติอนุมาน** ต่อไปนี้

1. การทดสอบสมมุติฐาน
   - 1.1 ความหมาย และประเภทของสมมุติฐานทางสถิติ
     - 1.1.1 สมมุติฐานหลัก
     - 1.1.2 สมมุติฐานทางเลือก
   - 1.2 ความคลาดเคลื่อนจากการทดสอบสมมุติฐานทางสถิติ
   - 1.3 ค่าระดับนัยสำคัญ และลักษณะของการทดสอบสมมุติฐาน
   - 1.4 ประเภทของสถิติที่ใช้ในการทดสอบสมมุติฐาน
2. การทดสอบสมมุติฐานโดยใช้ไคสแควร์ (Chi-square)
   - 2.1 สูตรทั่วไปของไคสแควว์
   - 2.2 การใช้ไคสแควร์ในกรณีตัวแปรตัวเดียวหรือข้อมูลมิติเดียว
   - 2.3 การใช้ไคสแควร์ในกรณีตัวแปรสองตัวหรือข้อมูลสองมิติ
     - 2.3.1 วัตถุประสงค์ของการวิจัย
     - 2.3.2 ข้อตกลงเบื้องต้น
     - 2.3.3 ตัวอย่างการทดสอบสมมุติฐานด้วยไคสแควร์
     - 2.3.4 การทดสอบในกรณีตาราง 2X2 หรือ df=1
     - 2.3.5 ข้อจำกัดของไคสแควร์

2) เพื่อให้ผู้ศึกษามีความสามารถใน**การวิเคราะห์และแปรผลข้อมูลด้วยไค-สแควร์**

1. การทดสอบสมมุติฐาน

1.1 ความนำ

สมมุติฐาน (Hypothesis) ในการวิจัย ***หมายถึง ข้อสมมุติที่กำหนดขึ้นเพื่อพรรณนาหรืออธิบายปรากฏการณ์ หรืออธิบายความสัมพันธ์ระหว่างตัวแปรในการวิจัย*** สมมุติฐานจะกำหนดขึ้นมาล่วงหน้าก่อนที่จะไปเก็บข้อมูล

ในการพิสูจน์ว่าสมมติฐานนั้นเป็นจริงหรือไม่ จำเป็นต้องไปเก็บข้อมูลมา แล้วนำมาทดสอบ ในการวิจัยเชิงปริมาณนั้นสามารถเข้ามาทดสอบด้วยวิธีการทางสถิติ และโดยทั่วไปการวิจัยนั้นมักเก็บข้อมูลมาจากตัวอย่าง ***จึงจำเป็นต้องใช้สถิติอนุมาน (Inferential Statistics) มาทดสอบสมมุติฐาน เพื่อประกอบในการตัดสินใจ***

1.2 ประเภทและชนิดของสมมุติฐานทางสถิติ

สมมุติฐานทางสถิติ แบ่งออกเป็น 2 ประเภทคือ

1. **สมุมติฐานหลัก (Null Hypothesis = H_o)** เป็นสมมุติฐานแรกที่ได้ตั้งขึ้น แบ่งออกเป็น 2 ชนิดคือ

1.1 สมมุติฐานเชิงพรรณา (Descriptive Statistics) เป็นสมมุติฐานที่มักกำหนดในรูปของ "***ไม่แตกต่าง***" หรือ "***เท่ากับ***" ซึ่งสามารถเขียนในรูปของข้อความทางภาษา หรือเขียนแทนด้วยสัญญลักษณ์ทางคณิตศาสตร์ เช่น

ในรูปของข้อความทางภาษา

H_o : อายุเฉลี่ยของนักศึกษามหาวิทยาลัยรามคำแหง **เท่ากับ** 23 ปี

เขียนในรูปสัญญลักษณ์ทางสถิติ

$H_o: \ \mu = 23$

1.2 สมมุติฐานเชิงวิเคราะห์ (Analytic Hypothesis) เป็นสมมุติฐานที่แสดงถึงความสัมพันธ์ระหว่างตัวแปร มักเป็นในรูปของ "***ไม่มีความสัมพันธ์***" "***ไม่มีอิทธิพลต่อกัน***" หรือ "***เป็นอิสระจากกัน***"

H_o : เพศ ***ไม่มีความสัมพันธ์***กับการมีส่วนร่วมทางการเมือง

H_o : ตัวแปร x และตัวแปร y ***ไม่มีความสัมพันธ์กัน***

เขียนในรูปสัญญลักษณ์ทางสถิติ

$H_o : r_{xy} = 0$

2. สมมุติฐานทางเลือก (Alternative Hypothesis = H_1 หรือ H_a) เป็นสมมุติฐานที่สองที่ตั้งขึ้น แบ่งออกเป็น 2 ชนิด

2.1 สมมุติฐานเชิงพรรณา มักจะกำหนดในรูปของ "***แตกต่าง***" "***ไม่เท่ากับ***" "***มากกว่า***" หรือ "***น้อยกว่า***" เช่น

เขียนในรูปข้อความ[1]

H_o : อายุเฉลี่ยของนักศึกษามหาวิทยาลัยรามคำแหง ***เท่ากับ*** 23 ปี (1)

H_1 : อายุเฉลี่ยของนักศึกษามหาวิทยาลัยรามคำแหง ***ไม่เท่ากับ*** 23 ปี (2)

H_1 : อายุเฉลี่ยของนักศึกษามหาวิทยาลัยรามคำแหง ***แตกต่าง***จาก 23 ปี (3)

H_1 : อายุเฉลี่ยของนักศึกษามหาวิทยาลัยรามคำแหง ***มากกว่า*** 23 ปี (4)

H_1 : อายุเฉลี่ยของนักศึกษามหาวิทยาลัยรามคำแหง ***น้อยกว่า*** 23 ปี (5)

2.2 สมมุติฐานเชิงวิเคราะห์ มักเขียนมาในรูปของ "***มีความสัมพันธ์***" "***มีอิทธิพล***" หรือ "***ไม่เป็นอิสระต่อกัน***" ซึ่งจะตรงกันข้ามกับ H_o เช่น

เขียนในรูปของข้อความทางภาษา

---

[1] <u>หมายเหตุ</u> อาจจะใช้เครื่องหมายคณิตศาสตร์ แทนการเปรียบเทียบได้ดังนี้

A ไม่เท่ากับ B เขียนเป็น $A \neq B$

A เท่ากับ B เขียนเป็น $A = B$

A มากกว่า B เขียนเป็น $A > B$

(ตัวที่น้อยกว่าจะอยู่ทางด้านปลายแหลมของเครื่องหมาย)

A น้อยกว่า B เขียนเป็น $A < B$

A มากกว่าหรือเท่ากับ B เขียนเป็น $A \geq B$

A น้อยกว่าหรือเท่ากับ B เขียนเป็น $A \leq B$

H_0 : เพศ ***ไม่มีความสัมพันธ์*** กับการมีส่วนร่วมทางการเมือง

H_1 : เพศ ***มีความสัมพันธ์***กับการมีส่วนร่วมทางการเมือง

H_1 : เพศ ***มีอิทธิพล***ต่อการมีส่วนร่วมทางการเมือง

H_1 : เพศ ***ไม่เป็นอิสระ***ต่อการมีส่วนร่วมทางการเมือง

H_1 : ตัวแปร x และตัวแปร y ***มีความสัมพันธ์กัน***

เขียนในรูปสัญญลักษณ์ทางสถิติ

$$H_o : r_{xy} = 0 \quad (1)$$

$$H_1 : r_{xy} \neq 0 \quad (2)$$

(1) มีความหมายว่า "**ค่าสัมประสิทธิ์สหสัมพันธ์ของตัวแปร x และ y มีค่า*เท่ากับ* 0**" นั่นคือ "**ตัวแปร x และ y *ไม่มีความสัมพันธ์*** กันนั่นเอง

(2) มีความหมายว่า "**ค่าสัมประสิทธิ์สหสัมพันธ์ของตัวแปร x และ y มีค่า *ไม่เท่ากับ* 0**" นั่นคือ "**ตัวแปร x และ y *มีความสัมพันธ์***กันนั่นเอง

1.3 ความคลาดเคลื่อนจากการทดสอบสมมุติฐานทางสถิติ

ในการทดสอบสมมุติฐานทางสถิติมักจะมี***ความคลาดเคลื่อน (Error)*** อยู่ ซึ่งความคลาดเคลื่อนที่บางครั้งอาจจะมีเพียงเล็กน้อย บางครั้งจะมีมาก ซึ่งผู้วิจัยจะต้องตั้งระดับความคลาดเคลื่อนไว้ว่า ความคลาดเคลื่อนเท่าใดจะรับได้ และมีความคลาดเคลื่อนเท่าใดจึงจะยอมรับไม่ได้

ความคลาดเคลื่อนจากการวิจัยมีมาจากหลายสาเหตุ เช่น

1. ความคลาดเคลื่อนจากการเก็บข้อมูล เช่น พนักงานสำรวจเก็บข้อมูลมาได้ข้อมูลที่คลาดเคลื่อนไม่ตรงกับความเป็นจริง เช่น ผู้ตอบตอบว่า "อายุ 30 ปี" แต่พนักงานเก็บข้อมูลบันทึกมา 35 ปี เป็นต้น ซึ่งเป็นความคลาดเคลื่อนตามธรรมชาติโดยไม่ตั้งใจ

2. ความคลาดเคลื่อนจากการเลือกตัวอย่าง เช่น เลือกตัวอย่างแบบไม่สุ่ม (Non-random) หรือกลุ่มตัวอย่างที่มีขนาดเล็กเกินไป

ดังนั้นผู้วิจัยจะต้อง***กำหนดระดับการยอมรับความคลาดเคลื่อน***ที่เกิดขึ้นเอาไว้ว่าเป็นเท่าใด ซึ่งขึ้นอยู่กับ***ความสำคัญ***ของการทดสอบว่าการทดสอบนั้นมีความสำคัญมากน้อยเพียงใด ***ถ้าเรื่องนั้นมีความสำคัญน้อย จะตั้งระดับความคลาดเคลื่อนไว้มาก ถ้าเรื่องนั้นมีความสำคัญมากจะต้องตั้งระดับความคลาดเคลื่อนเอาไว้น้อย*** เพราะว่าถ้าผิดพลาดได้เพียงเล็กน้อยจะมีผลเสียหายต่อส่วนรวมมากเป็นต้น ตัวอย่างเช่น การชั่งกระดาษหนังสือพิมพ์เพื่อขายหรือซื้อ ย่อมตั้งระดับการยอมรับความคลาดเคลื่อนจากผลของการชั่งมากกว่าการชั่งผลไม้ และถ้าเป็นการชั่งทองเพื่อซื้อขายนั้น ต้องตั้งระดับความคลาดเคลื่อนจากการวัดไว้น้อยมาก เพราะถ้ามีความผิดพลาดมากจะมีผลเสียหายมากนั่นเอง เป็นต้น

ความคลาดเคลื่อนจากการทดสอบสมมุติฐานมี ***2 ชนิด*** ได้แก่

1. ความคลาดเคลื่อนประเภทที่ 1 (Type I Error = α) เป็นความคลาดเคลื่อนเมื่อปฏิเสธ H_o ที่เป็นจริง สาเหตุที่เกิดความคลาดเคลื่อนชนิดนี้มี 2 ประการคือ

1. ตั้งระดับความคลาดเคลื่อนไว้สูงเกินไป ทำให้เกิดการยอมรับผลการทดสอบได้ง่าย เช่นตั้งระดับนัยสำคัญ = α ใช้ที่ .10 หรือ .20 เป็นต้น แต่การทดสอบนั้นเป็นเรื่องที่มีความสำคัญ

2. ความคลาดเคลื่อนจากการเลือกตัวอย่าง เช่น เลือกตัวอย่างแบบไม่สุ่ม หรือกลุ่มตัวอย่างมีขนาดเล็กจนเกินไป

ลักษณะความคลาดเคลื่อนประเภทที่ 1 แสดงได้ดังตาราง

การตัดสินใจ	**สภาพที่เป็นจริง**	
	H_o ถูก	H_o ผิด
ปฏิเสธ H_o	Type I Error = α	เป็นการตัดสินใจที่ถูกต้อง
ยอมรับ H_o	เป็นการตัดสินใจที่ถูกต้อง	Type II Error = β

(α เป็นภาษากรีกอ่านว่า "อัลฟา" β เป็นภาษากรีก เช่นเดียวกัน อ่านว่า "เบต้า)

2. ความคลาดเคลื่อนประเภทที่ 2 (Type II Error = β) ***เป็นความคลาดเคลื่อนที่เกิดจากการยอมรับ H_o ที่เป็นเท็จ*** ปัจจัยที่ก่อให้เกิดความคลาดเคลื่อนประเภทนี้ ได้แก่

1. การตั้งระดับนัยสำคัญหรือ α ไว้ต่ำเกินไป เช่น ตั้งไว้ 0.01 หรือ 0.001 จึงทำให้การสรุปเกิดการผิดพลาด

2. ความคลาดเคลื่อนจากการเลือกตัวอย่าง เช่น การเลือกตัวอย่างแบบไม่สุ่ม หรือกลุ่มตัวอย่างมีขนาดเล็กเกินไป

ความคลาดเคลื่อนประเภทที่ 1 นั้นถือว่าร้ายแรงกว่าความคลาดเคลื่อนชนิดที่ 2 ดังขอยกตัวอย่างเปรียบเทียบกับการพิพากษาของศาล ดังนี้

Ho : นาย ก. ไม่มีความผิด

H1 : นาย ก. มีความผิด

จากหลักฐานต่าง ๆ ผู้พิพากษาจะต้องเลือกสมมุติฐานใดสมุติฐานหนึ่งเป็นข้อสรุป ***ถ้าผู้พิพากษาตัดสินว่า นาย ก. มีความผิด ทั้งที่นาย ก. ไม่มีความผิด การตัดสินใจจะเป็นความผิดพลาดประเภทที่ 1*** (Type I Error) แต่ถ้าหากว่า นาย ก. กระทำความผิดจริง แต่ตัดสินว่าไม่ผิด การตัดสินใจจะมีความผิดพลาดประเภทที่ 2 ((Type II Error) จึงจะเห็นได้ว่า ความผิดพลาดแบบที่ 1 คือตัดสินว่า นาย ก. ทำผิดทั้งที่ไม่ได้ทำผิด จึงเป็นการตัดสินใจที่ผิดพลาดร้ายแรงกว่าการที่ได้ตัดสินว่า คนทำผิดนั้นไม่ผิดและได้ปล่อยตัวไป เพราะการนำเอาคนบริสุทธิ์เข้าคุกหรือลงโทษถือเป็นความผิดพลาดที่ร้ายแรงกว่าการปล่อยผู้กระทำผิดออกไป

1.4 ค่าระดับนัยสำคัญและลักษณะของการทดสอบสมมุติฐาน

ค่าระดับนัยสำคัญหรือ α ที่ตั้งขึ้นมาเป็น***บริเวณพื้นที่ใต้โค้ง***หรือเรียกว่าเป็น ***บริเวณวิกฤติ (Critical Region)*** ซึ่ง ***หมายถึงขอบเขตของความคลาดเคลื่อนที่เกิดจากการทดสอบสมมุติฐาน โดยปฏิเสธ H_o ที่เป็นจริง***

ลักษณะการทดสอบสมมุติฐานมี 2 ลักษณะ ดังนี้

1. การทดสอบแบบไม่มีทิศทาง (Non-directional Test) หรือการทดสอบแบบสองหาง (Two Tailed Test) เป็นการทดสอบที่มีจุดมุ่งหมายว่าสิ่งที่ทดสอบ***มีความแตกต่างกันหรือไม่*** โดยไม่สนใจว่ามีค่ามากหรือน้อยกว่ากัน เช่น

$H_0 : \overline{X} = 23$

$H_1 : \overline{X} \neq 23$

$H_0 : \overline{X}_1 = \overline{X}_2$

$H_1 : \overline{X}_1 \neq \overline{X}_2$

$H_0 : r_{xy} = 0$

$H_1 : r_{xy} \neq 0$

2. **การทดสอบแบบมีทิศทาง (Directional Test) หรือ เรียกว่า การทดสอบแบบหางเดียว (One Tailed Test)** เป็นการทดสอบที่มีจุดมุ่งหมายต้องการทดสอบว่า สิ่งที่จะทดสอบมีค่า "***มากกว่า***" หรือ "***น้อยกว่า***" เช่น

$H_0 : \overline{X} = 23$

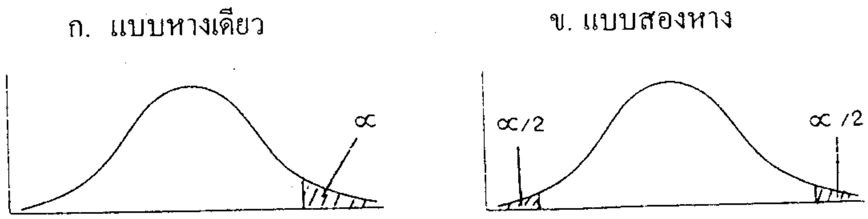
$H_1 : \overline{X} < 23$

$H_0 : \overline{X}_1 = \overline{X}_2$

$H_1 : \overline{X}_1 > \overline{X}_2$

$H_0 : r_{xy} = 0$

$H_1 : r_{xy} > 1$



ภาพที่ 6-1 แสดงการเปรียบเทียบบริเวณวิกฤตของการทดสอบแบบหางเดียวและแบบสองหาง

จาก**ภาพที่ 6-1 ภาพ ก.** แสดงการทดสอบ ***แบบหางเดียว*** ค่า α คือ บริเวณวิกฤติ (Critical Area) (ในภาพ คือบริเวณที่แลเงา) ซึ่งจะมีอยู่***ข้างเดียว*** ของการกระจายของข้อมูล เช่น $\alpha = 0.05$

ตามปกติพื้นที่ใต้โค้งมีค่า = 1 ดังนั้นค่า $\propto$ จะต้องมีค่าน้อยกว่า 1 และจะเท่ากับ 1 หรือมากกว่า 1 ไม่ได้

ส่วน **ภาพ ข.** แสดงการทดสอบ ***แบบสองหาง*** บริเวณวิกฤติจะอยู่ที่***สองข้าง*** ของบริเวณโค้ง ดังนั้นบริเวณแต่ละข้างจะมีค่าเท่ากับ $\propto/2$ สมมุติว่าตั้ง $\propto = 0.05$ ค่า $\propto/2 = 0.025$

เนื่องจากค่าบริเวณใต้โค้ง เป็นค่าความน่าจะเป็น (Probability) มีค่าเท่ากับ 1 และค่า $\propto$ ก็เป็นค่า Probability เช่นเดียวกัน

ดังนั้น ถ้าเราตั้ง $\propto = 0.05$ หมายความว่า ในการทดสอบจะยอมรับความผิดพลาดจากการทดสอบได้ไม่เกิน 0.05 แปลงเป็น % โดยการเอา 100 คูณ เข้าไป (หรือใช้วิธีบัญญัติไตรยางค์) จะได้ $\propto = 0.05 \times 100 = 5\%$

นั่นหมายความว่า **ในการทดสอบนักวิจัยจะยอมให้มีค่าโอกาสที่จะเกิดความผิดพลาด (คือปฏิเสธ H_o ที่เป็นจริง)** ได้ไม่เกิน 5% นั่นคือถ้ามีการผิดพลาด = 5% ยังจะยอมรับได้แต่ถ้าเกิน 5% เช่น 5.1% จะไม่ยอมรับสมมุติฐานจากการทดสอบ

ดังนั้น ถ้านักวิจัยกำหนด $\propto = 0.05$ ถ้าเป็นการทดสอบแบบหางเดียว (รูป ก.) พื้นที่บริเวณวิกฤติเป็นพื้นที่นักวิจัยจะยอมรับความผิดพลาดจากการทดสอบได้ และพื้นที่ วิกฤติ (บริเวณที่แลเงา) จะมีเท่ากับ 5% ของพื้นที่ใต้โค้งทั้งหมด

ในกรณีที่เป็นการทดสอบสองหาง (รูป ข.) พื้นที่วิกฤติจะอยู่ 2 หางของเส้นโค้ง แต่ละบริเวณจะมีพื้นที่ = 0.025% ของพื้นที่ใต้โค้งทั้งหมด

1.5 ประเภทของสถิติที่ใช้ในการทดสอบสมมุติฐาน

สถิติที่นำมาใช้ในการทดสอบสมมุติฐานก็คือ ***สถิติอนุมานหรือสถิติเชิงอ้างอิง (Inferetial Statistics)*** ซึ่งแบ่งออกเป็น 2 ประเภทใหญ่ ๆ

1) สถิติที่ใช้พารามิเตอร์ (Parametric Statistics) เป็นสถิติที่ใช้สำหรับ ***การแจกแจงของประชากรเป็นการแจกแจงแบบปกติ (Normal Distribution) หรือแบบอื่น ๆ*** เช่น การกระจายแบบ t (t-distribution) การกระจายแบบ F (F-distribution) เป็นต้น

การทดสอบแบบ t (t-test) เป็นการทดสอบเพื่อหาความ***แตกต่างของค่าเฉลี่ย***ของกลุ่มตัวอย่างกลุ่มต่าง ๆ 1 คู่ว่ามีความแตกต่างกันหรือไม่

ตัวอย่างเช่น ต้องการทราบว่า ข้าราชการชายและหญิงในหน่วยงานแห่งหนึ่ง มีรายได้เฉลี่ยต่างกันอย่างไร โดยสุ่มเลือกมาจากตัวอย่าง แล้วนำมาทดสอบจะนำมาตั้งสมมุติฐานว่า

H_0 : ข้าราชการชายและหญิงมีรายได้แตกต่างกันหรือไม่

H_0 : $\bar{X}_{ชาย} = \bar{X}_{หญิง}$

เมื่อ

$\bar{X}_{ชาย}$: รายได้ต่อปีของข้าราชการชาย

$\bar{X}_{หญิง}$: รายได้ต่อปีของข้าราชการหญิง

การทดสอบสมมุติฐานข้างต้น สามารถใช้ t- test มาทดสอบความแตกต่างระหว่างรายได้เฉลี่ยต่อปีของข้าราชการชาย ($\bar{X}_{ชาย}$) และข้าราชการหญิง ($\bar{X}_{หญิง}$) ซึ่งจะเห็นได้ว่า t - test จะใช้ในการทดสอบค่าเฉลี่ยของกลุ่มตัวอย่าง ***ทีละคู่*** เพื่อนำมาเป็น***ตัวแทนสรุปว่า***ค่าเฉลี่ยของข้าราชการชายและหญิงในหน่วยงานแห่งหนึ่งนั้น แตกต่างจริงหรือไม่

การทดสอบแบบ F (F - test) เป็นการทดสอบเพื่อหาความ***แตกต่างของค่าเฉลี่ย***ของกลุ่มตัวอย่างที่มี***มากกว่า 2 กลุ่ม*** การทดสอบแบบ F จะบอกว่า "***มีค่าเฉลี่ยอย่างน้อย 1 คู่ ที่แตกต่างกันหรือไม่***" ตัวอย่างเช่น

จากการวิจัยเรื่องรายได้เฉลี่ยต่อคนต่อปีของคน 5 อาชีพ คือ ข้าราชการ พนักงานรัฐวิสาหกิจ นักธุรกิจ รับจ้าง และเกษตรกร โดยเก็บข้อมูลจากกลุ่มตัวอย่าง ดังนั้นการทดสอบ F - test จะเป็นการบอกว่า "มีอย่างน้อยของรายได้เฉลี่ยต่อปีของกลุ่มอาชีพต่าง ๆ 1 คู่ มีความแตกต่างกันหรือไม่"

ในกรณีที่เราทราบว่าไม่มีรายได้เฉลี่ยของกลุ่มอาชีพใดเลยที่แตกต่างกันเลย ก็จะสิ้นสุดการทดสอบ ***แต่ถ้าพบว่า "มีรายได้เฉลี่ยต่อคนต่อปีของกลุ่มอาชีพต่าง ๆ อย่างน้อย 1 คู่ ที่แตกต่างกัน" แล้ว จะต้องนำเอารายได้เฉลี่ยทีละคู่ โดยวิธี t - test*** ดังที่กล่าวมา
แล้ว ข้อควรระวังก็คือ ไม่ควรนำเอาค่าเฉลี่ยมาทดสอบที่ละคู่ โดยใช้ t-testก่อนการทดสอบด้วย F-test

2) **สถิติที่ไม่ใช้พารามิเตอร์ (Nonparametric Statistics)** เป็นสถิติที่ไม่เกี่ยวข้องกับการใช้พารามิเตอร์ เช่น ค่าเฉลี่ยของประชากร (μ) และความเบี่ยงเบนมาตรฐานของประชากร (σ) และนอกจากนี้ยังเป็นสถิติที่สามารถนำมาทดสอบ ***ไม่จำเป็นต้องทราบการแจกแจงของข้อมูลว่าเป็นแบบใด*** (Distribution Free) ตัวอย่างเช่น ***ไคสแควร์*** เป็นสถิติประเภทไม่ต้องใช้พารามิเตอร์ชนิดหนึ่ง ซึ่งรายละเอียดจะได้กล่าวในหัวข้อต่อไป

2. การทดสอบสมมุติฐานโดยใช้ไคสแควร์ (Chi-square)

2.1 ความนำ

ไคสแควร์ เป็นสถิติประเภทสถิติอนุมานชนิดหนึ่ง ซึ่งนิยมนำมาใช้ในการวิจัยสังคมศาสตร์ทั้งนี้ เพราะว่ามีความเหมาะสมกับข้อมูลที่มีการวัดโดยใช้***มาตรวัดแบบนามบัญญัติ*** ซึ่งเป็นข้อมูลอยู่ในรูปความถี่ นอกจากนี้ไคสแควร์ยังสถิติที่มีข้อตกลงเบื้องต้นไม่เคร่งครัดเหมือนสถิติอื่นๆ ว่าการกระจายของข้อมูลจะต้องเป็นการแจกแจงแบบโค้งปกติ (Normal Curve) หรือแบบอื่นๆ ข้อมูลที่นำมาวิเคราะห์ด้วยไคสแควร์นั้น ***จะเป็นการกระจายแบบใดก็ได้ (Distribution Free Tests)*** ซึ่งเรียก สถิติประเภทนี้ว่า ***Nonparametric Statistics*** นอกจากนี้ไคสแควร์ยังเป็นสถิติที่นำมาใช้ได้ง่ายไม่ซับซ้อน ไคสแควร์จึงเป็นสถิติชนิดหนึ่งที่ได้รับความนิยมในการนำมาใช้ในการวิจัยทางสังคมศาสตร์

2.2 สูตรทั่วไปของไคสแควร์

สูตรทั่วไป มีดังนี้

$$\chi^2 = \sum_{i=1}^{k}\left[\frac{(0-E)^2}{E}\right]$$

กำหนดให้

0 = ความถี่ที่ได้จากการสังเกต (Observed Frequencies) ในแต่ละกลุ่ม หรือประเภท (อาจจะเป็นความถี่ที่ได้จากการสัมภาษณ์หรือแบบสอบถามก็ได้)

E = ความถี่ที่ได้จากความคาดหวังตามทฤษฎี (Expected Frequencies) ในแต่ละกลุ่มประเภท

k = จำนวนกลุ่มหรือประเภท

(χ เป็นภาษากรีกอ่านว่า "CHI" หรือ "ไค" ดังนั้น χ^2 อ่าน ไค-สแควร์)

2.3 การใช้ไคสแควร์ในกรณีตัวแปรตัวเดียว หรือ ข้อมูลมิติเดียว

นักวิจัยสามารถใช้ไคสแควร์มาทดสอบสมมุติฐานของตัวแปรตัวเดียว (One Variables Test) หรือเรียกว่า ข้อมูลมิติเดียว (One Dimension) ได้

ตัวอย่างที่ 1 จากการศึกษาถึงวันลาของข้าราชการโดยสุ่มศึกษาข้าราชการมา 10 หน่วยงาน ว่าการลางานของข้าราชการมีการกระจายอย่างเท่ากันในแต่ละวันของสัปดาห์หรือไม่ ข้อมูลที่ได้มีดังตารางข้างล่างนี้

ตารางที่ 6-1 แสดงการกระจายของการลาของข้าราชการในแต่ละวันของสัปดาห์

การลาในแต่ละวันของสัปดาห์	จำนวนผู้ลา (0)
จันทร์	45
อังคาร	35
พุธ	30
พฤหัส	40
ศุกร์	50
รวม	**200**

การทดสอบสมมุติฐานทำได้ดังนี้

1. กำหนดสมมุติฐานทางสถิติ

H_0 : การลาของแต่ละวันของสัปดาห์มีการกระจาย ***เท่ากัน***

H_1 : การลาของแต่ละวันของสัปดาห์มีการกระจายที่ ***แตกต่างกัน***

2. พิจารณาสถิติที่จะนำมาใช้ทดสอบ

การทดสอบเหมาะสมที่จะใช้ไคสแควร์มาทดสอบ

$$\chi^2 = \sum_{i=1}^{k}\left[\frac{(0-E)^2}{E}\right]$$

3. กำหนดค่านัยสำคัญทางสถิติ

ในการทดสอบครั้งนี้กำหนดค่านัยสำคัญทางสถิติเท่ากับ 0.05

$\alpha = 0.05$

4. *คำนวนหาค่า χ^2*

ตารางที่ 6-2 แสดงการคำนวณหาค่าไคสแควร์

การลาในแต่ละวันของสัปดาห์	จำนวนผู้ลา 0	ค่าความคาดหวัง E	$(O - E)^2$	$(O - E)^2/E$
1. จันทร์	45	40	$(45-40)^2 = 5^2 = 25$	$\frac{25}{40} = \frac{5}{8} = 0.625$
2. อังคาร	35	40	$(35-40)^2 = (-5^2) = 25$	$\frac{25}{40} = \frac{5}{8} = 0.625$
3. พุธ	30	40	$(30-40)^2 = (-10)^2 = 100$	$\frac{100}{40} = \frac{5}{2} = 2.5$
4. พฤหัส	40	40	$(40-40)^2 = (0)^2 = 0$	$\frac{0}{40} = 0$
5. ศุกร์	50	40	$(50-40)^2 = (10)^2 = 100$	$\frac{100}{40} = \frac{5}{2} = 2.5$
k = 5 **รวม**	**200**	**200**		$\sum_{i=1}^{k}\left[\frac{(O-E)^2}{E}\right] = 6.25$

ค่า E หรือค่าความถี่จากการคาดหวัง คือ $\frac{200}{5} = 40$ เพราะในทางความคาดหวังทางทฤษฎีแล้ว ทุกวันน่าจะมีการกระจายของวันลาเท่ากันทั้ง 5 วัน

$$\chi^2 = \sum_{i=1}^{k}\left[\frac{(0-E)^2}{E}\right]$$

$$\chi^2 = 6.25$$

5. *เปิดตารางหาค่าวิกฤติ (Critical Value)*

การเปิดตารางหาค่าวิกฤติจำเป็นต้องทราบถึง

5.1 ค่านัยสำคัญ = α = 0.05

5.2 ค่า Degree of Freedom (df)

df = (k - 1)

k = จำนวนชั้นหรือกลุ่ม หรือประเภท

= จำนวนวันที่นำมาหาการกระจายของการหา

(จันทร์, อังคาร,พุธ,พฤหัสและศุกร์)

= 5

$\therefore$ df = (5 - 1) = 4

เปิดตาราง เพื่อหาค่าวิกฤติเพื่อหาบริเวณวิกฤติ จากตาราง

ค่า $\chi^2_{(k-1),(\alpha)} = \chi^2_{(4,\ 0.05)} = 9.50$

6. *สรุปผลการทดสอบสมมุติฐาน*

พิจารณาดูว่าค่าที่ได้จากการคำนวณอยู่ในบริเวณวิกฤติหรือไม่ เพื่อสรุปผลการทดสอบสมมุติฐาน

6.1 ถ้าค่าที่ได้จากการคำนวณ ไม่อยู่ ในบริเวณวิกฤติ

จะยอมรับ H_o แต่ปฏิเสธ H_1

6.2 ถ้าค่าที่ได้จากการคำนวณ อยู่ ในบริเวณวิกฤติ

จะปฏิเสธ H_o แต่ยอมรับ H_1

หรือ อาจจะพิจารณาได้ดังนี้

6.3 ถ้าค่าของ χ^2 ***จากการคำนวณ เท่ากับหรือน้อยกว่า ค่าวิกฤติจากการเปิดตาราง*** $\chi^2_{(df,\ \alpha)}$

จะยอมรับ H_o แต่ปฏิเสธ H_1

6.4 ถ้าค่า χ^2 จากการคำนวณ มากกว่า ค่าวิกฤติ จากการเปิดตารางที่ χ^2 (df, α)

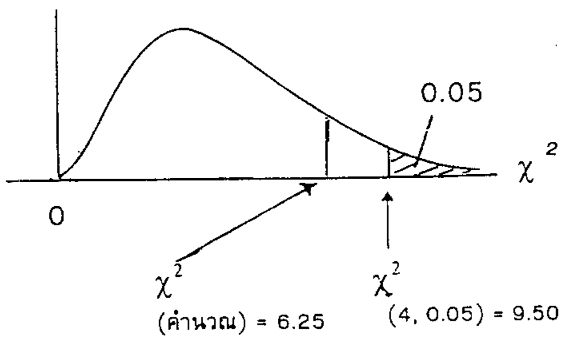
จะปฏิเสธ H_0 แต่ยอมรับ H_1

จากโจทย์จะเห็นได้ว่า

$$\chi^2_{(\text{คำนวณ})} < \chi^2_{(4,\ 0.05)}$$

ดังนั้น ***จะยอมรับ H_0*** นั่นคือ ***การลาของแต่ละวันนั้นมีการกระจายที่เท่ากัน หรือไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ***

จากโจทย์สามารถนำมาเขียนเป็นรูปภาพได้ดังนี้



จากรูปภาพ จะเห็นว่าค่า $\chi^2_{(\text{คำนวณ})}$ ไม่อยู่ในบริเวณวิกฤติ (บริเวณพื้นที่แลเงา) ซึ่งพื้นที่แลเงามีค่า = 0.05 หรือคิดเป็น 5% ของพื้นที่ใต้โค้งทั้งหมด

การทดสอบสมมุติฐานของตัวแปรตัวเดียวนี้ บางที่เรียกว่า ***การทดสอบสภาวะสนิทดี*** (Goodness of Fit Test) เพราะเป็นทดสอบว่าการกระจายของข้อมูล จากการสังเกต (0) มีการแจกแจงเป็นไปตามค่าทฤษฎีที่คาดหวัง (E) ไว้หรือไม่

2.4 การใช้ไคสแควร์ในกรณีตัวแปรสองตัว หรือข้อมูลสองมิติ

2.4.1 วัตถุประสงค์ของการทดสอบ

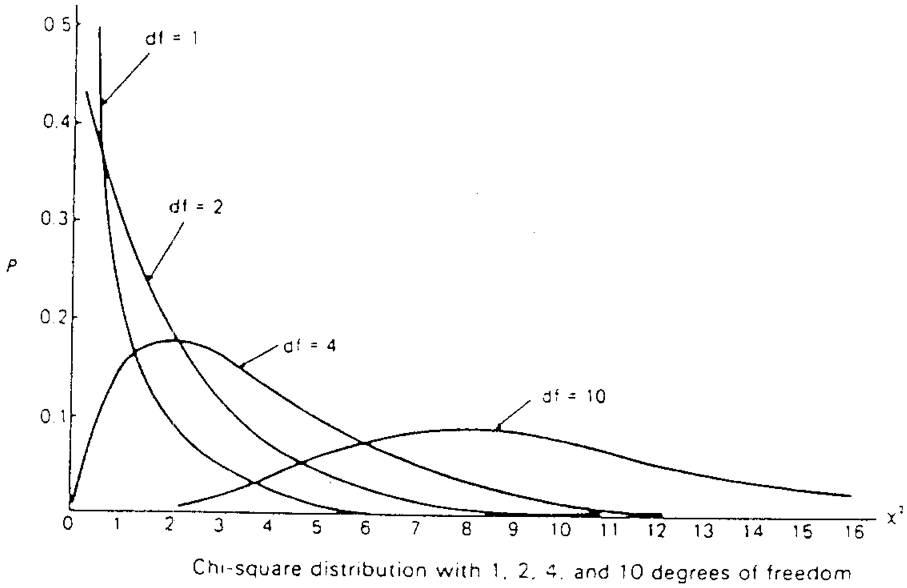
วัตถุประสงค์ของการทดสอบสมมุติฐานด้วยไคสแควร์ ในกรณีตัวแปรสองตัวนั้น เพื่อทดสอบ ***ตัวแปรสองตัวมีความสัมพันธ์กันหรือไม่*** หรือ ***เป็นอิสระต่อกันหรือไม่*** ตัวอย่างเช่น "เพศ มีความสัมพันธ์กับการมีส่วนร่วมทางการเมืองหรือไม่" หรือ "การศึกษามีความสัมพันธ์กับฐานะทางเศรษฐกิจหรือไม่" "อายุ ของผู้บริหารมีความสัมพันธ์กับความพึงพอใจของบุคลากรในหน่วยงานหรือไม่" เป็นต้น

การทดสอบไคสแควร์แบบนี้จะต้องนำข้อมูลมาจัดทำเป็นตาราง เรียกว่า ***ตารางไขว้ (Cross-tabulation Table) หรือตารางแบบสองทาง (Two Way Table) หรือตารางการณ์จร (Contingency Table)***

2.4.2 ข้อตกลงเบื้องต้น

ข้อตกลงเบื้องต้นในการทดสอบด้วยไคสแควร์ มีดังนี้ [1)]

1. ข้อมูลที่ได้จากการสังเกตแต่ละค่าจะต้องจัดอยู่ในกลุ่มตัวอย่าง หรือประเภทใดประเภทหนึ่งเท่านั้น
2. ข้อมูลที่ได้จากการสังเกตจากสมาชิกในกลุ่มตัวอย่างที่เป็นอิสระจากกัน
3. ค่าที่ได้จากการสังเกต จะมีค่าเป็นความถี่
4. ในกรณีที่ df $\geq$ 2 ค่าความถี่ที่คาดหวัง (E) จะต้องมีค่าไม่น้อยกว่า 5 และในกรณีที่ df = 1 (หรือเป็นตาราง 2×2) ค่าความถี่ที่คาดหวังต้องไม่น้อยกว่า 10
5. ในกรณีที่ df = 1 (นั่นคือ ตารางเป็นแบบ 2×2) การคำนวณหาค่า χ^2 จะต้องมีค่าปรับแก้
6. กลุ่มตัวอย่างทั้งสองจะต้องถูกสุ่มมาอย่างเป็นอิสระ



ภาพที่ 6-2 แสดงการกระจายของไคสแควร์ที่มีระดับของความเป็นอิสระ(Degree of Freedom) แตกต่างกัน

Source : Simon W. Tai, **Social Science Statistics : Its Elements and Applications**, (Santa Monica, California : 1978), p.236.

2.4.3 ตัวอย่างการทดสอบสมมุติฐานด้วยไคสแควร์

ตัวอย่างที่ 2 จากการวิจัยถึงความเห็นของข้าราชการกรุงเทพมหานครระดับ 1-5 จำแนกตามเพศ ซึ่งเป็นตัวอย่างของการวิจัยต่อรูปแบบภาวะผู้นำของหัวหน้าฝ่าย ระดับ 6-7 ของกรุงเทพมหานคร เขตสวนหลวง และเขตห้วยขวาง ปรากฏผลดังนี้

ตารางที่ 6-3 แสดงความสัมพันธ์ระหว่างเพศกับรูปแบบภาวะผู้นำระดับหัวหน้าฝ่าย

เพศ	รูปแบบภาวะผู้นำ				
	เผด็จการ	เผด็จการอย่างมีศิลป์	ปรึกษาหารือ	มีส่วนร่วม	รวม
ชาย	10 (8.3)	12 (9.9)	85 (70.2)	14 (11.6)	121 (100)
หญิง	6 (5.4)	19 (17.1)	71 (64.0)	15 (13.5)	111 (100)
รวม	**16** **(6.9)**	**31** **(13.4)**	**156** **(67.2)**	**29** **(12.5)**	**232** **(100)**

ที่มา มนตรีพล พิทยธรรม, "รูปแบบภาวะผู้นำของข้าราชการกรุงเทพมหานครตำแหน่งหัวหน้าฝ่าย : ศึกษากรณีสำนักงานเขตห้วยขวาง และสำนักงานเขตสวนหลวง", (วิทยานิพนธ์ ปริญญาศิลปศาสตรมหาบัณฑิต (รัฐศาสตร์) มหาวิทยาลัยรามคำแหง, 2538), หน้า 91

ขั้นตอนการพิสูจน์สมมุติฐาน มีดังนี้

ขั้นที่ 1 กำหนดสมมุติฐานทางสถิติ

H_0 : ความเห็นของผู้ใต้บังคับบัญชาจำแนกตามเพศ ***ไม่มีความสัมพันธ์*** ต่อความเห็นในเรื่องรูปแบบภาวะผู้นำของหัวหน้าฝ่าย

H_1 : ความเห็นของผู้ใต้บังคับบัญชาจำแนกตามเพศ ***มีความสัมพันธ์*** ต่อความเห็นในเรื่องรูปแบบภาวะผู้นำของหัวหน้าฝ่าย

ขั้นที่ 2 พิจารณาสถิติที่นำมาใช้ในการทดสอบ

การทดสอบนี้เป็นการทดสอบความสัมพันธ์ระหว่างตัวแปร 2 ตัวว่ามีความสัมพันธ์กันหรือไม่ คือ เพศ และความเห็นต่อรูปแบบภาวะผู้นำระดับหัวหน้าฝ่าย โดยที่เป็นการวัดตัวแปรโดยใช้มาตรวัดแบบตามบัญญัติ (Nominal Scale) จึงมีความเหมาะสมที่จะใช้ไคสแควร์มาทำการทดสอบ สูตรสถิติที่นำมาใช้คือ

$$\chi^2 = \sum_{i=1}^{k}\left[\frac{(0-E)^2}{E}\right]$$

ขั้นที่ 3 กำหนดค่านัยสำคัญทางสถิติ

$$\propto = 0.05$$

ขั้นที่ 4 คำนวณหาค่า χ^2

เพศ	รูปแบบภาวะผู้นำ: เผด็จการ		เผด็จการอย่างมีศิลป์		ปรึกษาหารือ		มีส่วนร่วม		รวม
	O	E	O	E	O	E	O	E	
ชาย	10	$\frac{121\times16}{232}$ = 8.34	12	$\frac{121\times31}{232}$ = 16.17	85	$\frac{121\times156}{232}$ = 81.36	14	$\frac{121\times29}{232}$ = 15.13	**121**
หญิง	6	111 × 16 / 232 = 7.66	19	111 × 31 / 232 = 14.83	71	111 × 156 / 232 = 74.64	15	111 × 29 / 232 = 13.87	**111**
รวม	**16**		**31**		**156**		**29**		**232**

$$\text{ค่า } E \text{ ในแต่ละช่อง} = \frac{\text{ผลรวมค่า O ตามแนวนอน} \times \text{ผลรวมค่า O ตามแนวตั้ง}}{\text{จำนวนตัวอย่างทั้งหมด}}$$

ตัวอย่างเช่น ค่า E ของชายที่มีความเห็นว่ารูปแบบของหัวหน้าฝ่ายเป็นแบบเผด็จการ

$$\frac{121 \times 16}{232} = 8.34$$

$$\chi^2 = \text{ผลรวมของค่า } \frac{(O-E)^2}{E} \text{ ทุกช่อง}$$

$$= \frac{(10-8.34)^2}{8.34} + \frac{(12-16.17)^2}{16.17} + \frac{(85-81.36)^2}{81.36} + \frac{(14-15.13)^2}{15.13} + \frac{(6-7.66)^2}{7.66} + \frac{(19-14.83)^2}{14.83} + \frac{(71-74.64)^2}{74.64} + \frac{(15-13.87)^2}{13.87}$$

$$= \frac{2.77}{8.34} + \frac{17.39}{16.17} + \frac{13.25}{81.36} + \frac{1.28}{15.13} + \frac{2.76}{7.66} + \frac{17.39}{14.83} + \frac{6.71}{74.64} + \frac{1.28}{13.87}$$

$$= 3.45$$

ขั้นที่ 5 เปิดตารางหาค่าวิกฤติ

5.1 ระดับนัยสำคัญ = $\propto$ = 0.05

5.2 ค่า Degreee of Freedom (df)

$$df = (r - 1)(c - 1)$$

เมื่อ r = จำนวนแถวตามแนวนอน (Row)

c = จำนวนแถวตามแนวตั้ง (Column)

ในกรณีตัวอย่างข้างต้น

r = 2 แถว (ชาย, หญิง)

c = 4 แถว (เผด็จการ, เผด็จการอย่างมีศิลป์, ปรึกษาหารือ, และมีส่วนร่วม)

ดังนั้น $df = (2 - 1)(4 - 1) = 1 \times 3$

$= 3$

ค่า $\chi^2_{(df, \propto)} = \chi^2_{(3, 0.05)} = 7.82$

ขั้นที่ 6 สรุปผลการทดสอบสมมุติฐาน

ค่า $\chi^2_{(\text{คำนวณ})} < \chi^2_{(3,\ 0.05)}$

ดังนั้น ยอมรับ H_o แต่ ปฏิเสธ H_1

นั่นคือ "ความเห็นผู้ใต้บังคับบัญชาจำแนกตามเพศ ***ไม่มีความสัมพันธ์*** ต่อความเห็นในเรื่องรูปแบบภาวะผู้นำของหัวหน้าฝ่าย"

ซึ่งหมายความว่า ผู้ตอบที่เป็นผู้ใต้บังคับบัญชา **ทั้งเพศชายและหญิง** ***มีความเห็นไม่แตกต่างกัน***เกี่ยวกับ**รูปแบบของหัวหน้าฝ่ายที่เป็นผู้บังคับบัญชา** คือส่วนใหญ่เห็นว่าผู้บังคับบัญชาระดับหัวหน้าฝ่าย มี**ภาวะผู้นำแบบปรึกษาหารือ** ทั้งนี้โดยดูจากความเห็นที่แสดงโดยอัตราส่วนร้อยประกอบ (ชายมีความเห็น 70.2% หญิง 64.0%)

จากการทดสอบสมมุติฐานข้างต้น ถ้าผู้วิจัยยอมรับ H_o ว่าเป็นจริง โอกาสที่จะเกิดความผิดพลาดประเภทที่ 1 (Type I Error) มีถึง 0.32771 หรือประมาณ $0.32771 \times 100 = 32.77\%$ แต่ผู้วิจัยยอมรับที่จะให้เกิดความผิดพลาดประเภทที่ 1 หรือ α เพียง 0.05 หรือ 5% ผู้วิจัยจึง**ปฏิเสธ H_o ว่าเป็นจริง** หรือ**ยอมรับ H_1 ว่าเป็นจริง** (รายละเอียดโปรดดูเรื่อง "ความคลาดเคลื่อนจากการทดสอบสมมุติฐานทางสถิติ" ในหัวข้อ 5.3 และ 5.4)

จะเห็นได้ว่า แม้ชายและหญิง มีอัตราร้อยละของความเห็นเกี่ยวกับรูปแบบภาวะผู้นำของหัวหน้าฝ่ายต่างกันเล็กน้อย แต่เมื่อมาทดสอบที่**ระดับนัยสำคัญทางสถิติที่ 0.5** แล้ว พบว่า **ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ**

ปัจจุบันมีคำสั่งสำเร็จรูปที่นำมาใช้ประมวลผลด้วยไมโครคอมพิวเตอร์ คือ ***Statistical Package for the Social Sciences*** หรือ **SPSS/PC$^+$** มาคำนวณหาค่าไคสแควร์จึงทำให้การคำนวณทำได้สะดวกมากยิ่งขึ้น สำหรับการแสดงค่าของคำสั่งดังกล่าว จะแสดงค่าออกมาเป็นค่า Proability เช่นจากตารางดังกล่าวข้างต้น $P = 0.32771$ แต่ผู้วิจัยกำหนดค่า $\alpha = 0.05$

จะเห็นได้ว่า **ค่า $P > \alpha$** ดังนั้นผู้วิจัยจึงยอมรับ H_o ซึ่งอธิบายได้ว่า ***ค่าของ P ซึ่งแสดงถึงพื้นที่ของโค้งของการเปิดตาราง มีมากกว่าบริเวณวิกฤติ (Crltical Region) ซึ่งกำหนดไว้เพียง 0.05*** เท่านั้น

ตัวอย่างที่ 3 จากการวิจัยเรื่องเดิมที่กล่าวมาแล้ว ผู้วิจัยได้ศึกษาถึงความสัมพันธ์ระหว่างระดับตำแหน่งของผู้ใต้บังคับบัญชากับการประเมินรูปแบบของหัวหน้าฝ่ายของตนเอง ปรากฏผลดังนี้

ตารางที่ 6-4 แสดงความสัมพันธ์ระหว่างระดับตำแหน่งกับรูปแบบภาวะผู้นำ

ระดับตำแหน่ง	**รูปแบบภาวะผู้นำ**				
	เผด็จการ	**เผด็จการอย่างมีศิลป์**	**ปรึกษาหารือ**	**มีส่วนร่วม**	**รวม**
สายงานระดับ 1-3	3	16	60	13	**92**
	(3.3)	(17.4)	(65.2)	(14.1)	**(100)**
สายงานระดับ 2-4	-	6	31	4	**41**
		(14.6)	(75.6)	(9.8)	**(100)**
สายงานระดับ 3-5	13	9	65	12	**99**
	(13.1)	(9.1)	(65.7)	(12.1)	**(100)**
รวม	**16**	**31**	**156**	**29**	**232**
	(6.9)	**(13.4)**	**(67.2)**	**(12.5)**	**(100)**

$p = 0.03362$

หมายเหตุ ***กลุ่มสายงานระดับ 1-3*** หมายถึง กลุ่มข้าราชการที่มีวุฒิการศึกษาตั้งแต่ระดับมัธยมศึกษาตอนต้นขึ้นไป แต่ไม่เกินประกาศนียบัตรวิชาชีพ (ปวช.)

กลุ่มสายงานระดับ 2-4 หมายถึง กลุ่มข้าราชการที่มีวุฒิการศึกษาระดับประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) หรือ อนุปริญญาหรือเทียบเท่า

กลุ่มสายงานระดับ 3-5 หมายถึง กลุ่มข้าราชการที่มีวุฒิตั้งแต่ปริญญาตรีขึ้นไป

การทดสอบสมมุติฐานทำได้ ดังนี้

ขั้นที่ 1 กำหนดสมมุติฐานทางสถิติ

H_0 : ระดับตำแหน่งของผู้ใต้บังคับบัญชา***ไม่มีความสัมพันธ์*** กับการประเมินรูปแบบภาวะผู้นำของหัวหน้าฝ่าย

H_1 : ระดับตำแหน่งของผู้ใต้บังคับบัญชา***มีความสัมพันธ์*** กับการประเมินรูปแบบภาวะผู้นำของหัวหน้าฝ่าย

ขั้นที่ 2 พิจารณาสถิติที่เหมาะสมจะนำมาใช้

สถิติที่เหมาะสม คือ ไคสแควร์ (χ^2)

ขั้นที่ 3 กำหนดค่านัยสำคัญทางสถิติ

$$\alpha = 0.05$$

ขั้นที่ 4 คำนวณหาค่า χ^2

ระดับ	รูปแบบภาวะผู้นำ								รวม
	เผด็จการ		เผด็จการอย่างมีศิลป์		ปรึกษาหารือ		มีส่วนร่วม		
ตำแหน่ง	O	E	O	E	O	E	O	E	
สายงาน ระดับ 1-3	3	$\frac{92 \times 16}{232} = 6.4$	16	$\frac{92 \times 31}{232} = 12.3$	60	$\frac{92 \times 156}{232} = 61.9$	13	$\frac{92 \times 29}{232} = 11.5$	**92**
สายงาน ระดับ 3-4	-	$\frac{41 \times 16}{232} = 2.8$	6	$\frac{41 \times 31}{232} = 5.5$	31	$\frac{41 \times 156}{232} = 27.6$	4	$\frac{41 \times 29}{232} = 5.1$	**41**
สายงาน ระดับ 3-5	13	$\frac{99 \times 16}{232} = 6.8$	9	$\frac{99 \times 31}{232} = 13.2$	65	$\frac{99 \times 156}{232} = 66.5$	12	$\frac{99 \times 29}{232} = 12.4$	**99**
รวม	**16**		**31**		**156**		**29**		**232**

ค่าของ E แต่ละช่องของตาราง หาได้ดังนี้

$$\frac{\textit{ผลรวมของข้อมูลตามแนวนอน} \times \textit{ผลรวมของข้อมูลตามแนวตั้ง}}{\textit{จำนวนตัวอย่างทั้งหมด}}$$

เช่น ค่า E ของผู้ตอบในสายงานระดับ 1-3 ที่ตอบ "เผด็จการ"

$$= \frac{\text{จำนวนผู้ตอบสายงานระดับ1-3 ทั้งหมด} \times \text{จำนวนผู้ตอบว่าเผด็จการทั้งหมด}}{\text{จำนวนตัวอย่างทั้งหมด}}$$

$$= \frac{96 \times 16}{232} = 6.4$$

คำนวณหาค่า χ^2 จากสูตร $\sum_{i=0}^{k}\left[\frac{(O-E)^2}{E}\right]$

$$\therefore \quad \chi^2 = \frac{(3-6.4)^2}{6.4} + \frac{(16-12.3)^2}{12.3} + \frac{(60-61.9)^2}{61.9} + \frac{(43-11.5)^2}{11.5} +$$

$$\frac{(0-2.8)^2}{2.8} + \frac{(6-5.5)^2}{5.5} + \frac{(31-27.6)^2}{27.6} + \frac{(4-5.1)^2}{5.1} +$$

$$\frac{(13-6.8)^2}{6.8} + \frac{(9-13.2)^2}{13.2} + \frac{(65-66.5)^2}{66.5} + \frac{(12-12.4)^2}{12.4} +$$

$$= 13.66$$

ขั้นที่ 5 เปิดตารางหาค่าวิกฤติ (Critical Value)

5.1 ระดับนัยสำคัญ = α = 0.05

(เป็นค่าที่นักวิจัยกำหนดขึ้นมาเอง สำหรับการทดสอบครั้งนี้)

5.2 ค่า Degreee of Freedom = (r - 1) (c - 1)

r = จำนวนแถวตามแนวนอน (Row)

= แถวของระดับตำแหน่ง คือ สายงานระดับ 1-3, ระดับ 2-4,

= ระดับ 3-5 ซึ่งมี 3 แถว

$\therefore$ r = 3

c = จำนวนแถวตามแนวตั้ง

= จำนวนแถวของข้อมูลที่แสดงรูปแบบภาวะผู้นำมี 4 แถวได้แก่ เผด็จการ เผด็จการอย่างมีศิลป์ ปรึกษาหารือ และมีส่วนร่วม

= 4

$\therefore$ df = (3 - 1) (4 - 1) = 2 × 3 = 6

เปิดตารางจะได้

$$\chi^2_{(df, \alpha)} = \chi^2_{(6, 0.05)} = 12.6$$

ขั้นที่ 6 สรุปผลการทดสอบสมมุติฐาน

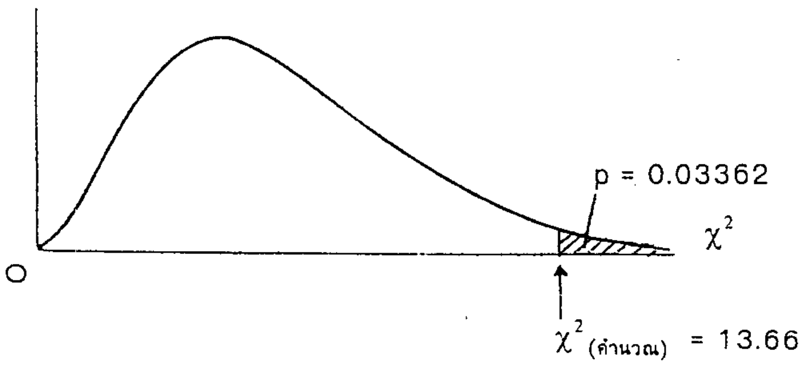
เพราะว่า $\chi^2_{(\text{คำนวณ})} > \chi^2_{(6, 0.05)}$

ดังนั้น จึงยอมรับ H_1 ว่าเป็นจริง

หมายความว่า ระดับตำแหน่งของผู้ใต้บังคับบัญชา***มีความสัมพันธ์***กับการประเมินรูปแบบภาวะผู้นำหัวหน้าฝ่าย และเมื่อพิจารณาข้อมูลจะพบว่าผู้ตอบทุกสายงานประเมินว่าหัวหน้าฝ่ายใช้รูปแบบปรึกษาหารือ แต่ผู้มีอยู่สายงานระดับ 1-3 จะมีผู้ตอบว่าหัวหน้าฝ่ายใช้ภาวะผู้นำแบบเผด็จการเป็นสัดส่วน***มากกว่า***สายงานอื่น ๆ เล็กน้อย (โปรดดูอัตราร้อยละของผู้ตอบในตาราง 6-4)

ในการใช้ไมโครคอมพิวเตอร์ทำ***การประมวลผลโดยใช้คำสั่งสำเร็จรูป SPSS/PC$^+$*** ได้มีการนำเสนอข้อมูลที่แสดงค่าของ χ^2 ที่ได้จากการคำนวณ และค่าของ p

ค่าของ p คือค่าของขนาดของพื้นที่ใต้โค้งของการกระจายไคสแควร์จากการคำนวณค่า χ^2 และค่า p ที่ได้จะแสดงในรูปของ Probability



ภาพที่ 6-3 แสดงการเปรียบเทียบค่าไคสแควร์ χ^2 จากการคำนวณกับค่า p

จากรูปจะเห็นได้ว่าค่า p จะเป็นค่าของพื้นที่ใต้เส้นโค้ง (บริเวณที่แลเงา) ที่ได้มาจากค่า χ^2 จากการคำนวณ = 13.66

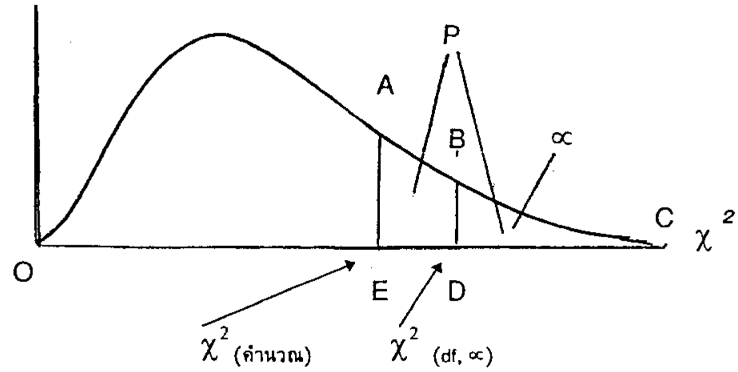
สรุปได้ว่าค่า p คือค่าพื้นที่ใต้โค้งที่แสดงในรูปของค่า Probability เช่นเดียวกับค่า ∝ นั่นเอง ดังนั้น การสรุปผลการทดสอบจึงนำเอาค่า p มาเปรียบเทียบกับค่า ∝ ที่นักวิจัยตั้งไว้

ใน**การสรุปผล**จากตารางที่ประมวลผลโดย**คำสั่งสำเร็จรูป SPSS/PC⁺** มีหลักดังนี้

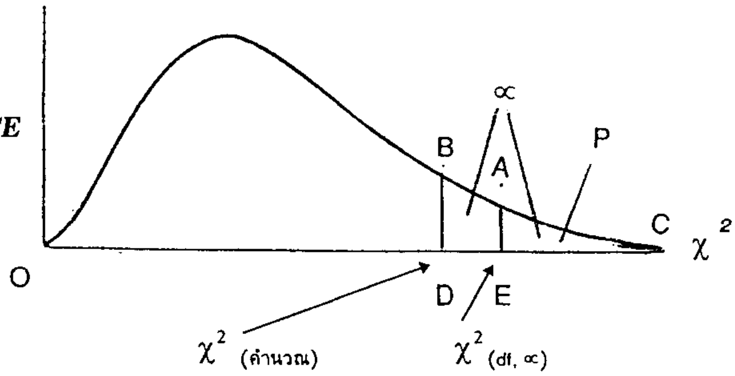
1) ถ้า **p** $>$ ∝ (p มากกว่า ∝) **จะยอมรับ** H_o แต่ ***ปฏิเสธ*** H_1

2) ถ้า **p** $\leq$ ∝ (p น้อยกว่าหรือเท่ากับ ∝) **จะยอมรับ** H_1 แต่ ***ปฏิเสธ*** H_o

ภาพที่ 6-4 การวิเคราะห์ด้วยไค-สแควร์ เมื่อ ค่า P > ∝
(P มากกว่า ∝)
∝ คือพื้นที่ BCD
P คือพื้นที่ ABDE + BCD



ภาพที่ 6-5 การวิเคราะห์ด้วยไค-สแควร์เมื่อ ค่า P < ∝
(P น้อยกว่า ∝) P คือพื้นที่ ACE
∝ คือพื้นที่ ABDE + ACE



$\chi^2_{(\text{คำนวณ})}$ มีค่าเท่ากับ จุด E (จุดบนแกนแนวนอน)

$\chi^2_{(df, ∝)}$ มีค่าเท่ากับ จุด D (จุดบนแกนแนวนอน)

จะเห็นว่าใน***ตาราง 6-2*** ได้มีการนำเสนอค่า p = 0.03362 แต่ผู้วิจัยได้กำหนด ∝ = 0.05

ดังนั้น ค่า p ***น้อยกว่า*** ค่าระดับนัยสำคัญ (∝) ที่กำหนดในการทดสอบครั้งนี้ จึงจะ***ยอมรับ*** H_1 ที่ระดับนัยสำคัญทางสถิติเท่ากับ 0.05

สำหรับในปัจจุบันนี้การประมวลผลการวิจัยด้วยเครื่องไมโครคอมพิวเตอร์ โดยใช้คำสั่งสำเร็จรูปต่าง ๆ ทำได้***ง่ายและรวดเร็ว*** และ***เป็นที่นิยม***กันในปัจจุบัน การศึกษาในระดับปริญญาโทมักจะมีการเรียนการสอนให้ใช้คำสั่งดังกล่าว ในกรณีที่นักวิจัยไม่สามารถใช้คำสั่งดังกล่าวได้ อาจจะให้นักสถิติหรือผู้ที่เรียนมาช่วยจัดทำให้ก็ได้

2.4.4 การทดสอบในกรณีตาราง 2 × 2 หรือ df = 1

1) กรณีใช้สูตรทั่วไป

ในกรณีที่จะทดสอบสมมุติฐาน ของความสัมพันธ์ระหว่างตัวแปร ที่ตารางเป็น 2 × 2 คือ แถวแนวนอน (Row) = 2 และแถวแนวตั้ง (Column) = 2 หรือ ค่า Degree of Freedom = 1 (r - 1) (c - 1) = (2 - 1) (2 - 1) = 1) แล้วจำเป็น***ต้องใช้ค่าแก้ของเยท (Yate's Correction)*** เพื่อทำให้ต่อเนื่อง [2)]

สูตรที่ปรับค่าแก้ คือ

$$\chi^2 = \sum \frac{\left(|0 - E| - \frac{1}{2}\right)^2}{E}$$

เมื่อ | O-E | เป็น***ค่าสัมบูรณ์ (Absulute Value)*** คือค่าที่ไม่คิดเครื่องหมายลบ

ตัวอย่างที่ 4 จากการวิจัยถึงความเห็นของประชาชนต่อการยอมรับนโยบายและตัวบุคคลในคณะรัฐบาลของประชาชนในจังหวัดแห่งหนึ่ง โดยการสอบถามจากกลุ่มตัวอย่าง โดยผลของการเก็บข้อมูลดังนี้

ตารางที่ 6-5 แสดงความสัมพันธ์ระหว่าง เพศ และความเห็นในการยอมรับนโยบายและตัวบุคคลในคณะรัฐบาล

เพศ	**ความเห็น**		**รวม**
	ยอมรับ	**ไม่ยอมรับ**	
ชาย	270	230	500
หญิง	276	224	500
รวม	**546**	**454**	**1,000**

(ตัวเลขสมมุติเพื่อใช้ในการศึกษา)

วิธีการทดสอบสมมุติฐาน

ขั้นที่ 1 กำหนดสมมุติฐานในการทดสอบ

H_0 เพศไม่มีความสัมพันธ์กับความเห็นเกี่ยวกับการยอมรับนโยบายและตัวบุคคลในคณะรัฐบาล

H_1 เพศมีความสัมพันธ์กับความเห็นเกี่ยวกับการยอมรับนโยบายและตัวบุคคลในคณะรัฐบาล

ขั้นที่ 2 พิจารณาสถิติที่เหมาะสมที่จะนำมาใช้

สถิติที่เหมาะสม คือ χ^2

ขั้นที่ 3 กำหนดค่านัยสำคัญทางสถิติ

$\alpha = 0.01$

ขั้นที่ 4 คำนวนหาค่า χ^2

จากข้อมูลที่ได้จากการสำรวจสามารถนำมาทดสอบสมมุติฐานได้ดังนี้

เพศ	**ความเห็น**				**รวม**
	ยอมรับ		**ไม่ยอมรับ**		
	O	E	O	E	
ชาย	270	$\frac{500 \times 546}{1,000} = 273$	230	$\frac{500 \times 454}{1,000} = 227$	500
หญิง	276	$\frac{500 \times 546}{1,000} = 273$	224	$\frac{500 \times 454}{1,000} = 227$	500
รวม	**546**		**454**		**1,000**

$$\chi^2 = \Sigma \frac{\left(|0 - E| - \frac{1}{2}\right)^2}{E}$$

$$= \frac{\left(|270 - 273| - \frac{1}{2}\right)^2}{273} + \frac{\left(|230 - 227| - \frac{1}{2}\right)^2}{227} +$$

$$\frac{\left(|276 - 273| - \frac{1}{2}\right)^2}{273} + \frac{\left(|224 - 227| - \frac{1}{2}\right)^2}{227} +$$

$$= \frac{\left(|-3|-\frac{1}{2}\right)^2}{273}+\frac{\left(|3|-\frac{1}{2}\right)^2}{227}+\frac{\left(|3|-\frac{1}{2}\right)^2}{273}+\frac{\left(|-3|-\frac{1}{2}\right)^2}{227}$$

$$= \frac{(3-0.5)^2}{273}+\frac{(3-0.5)^2}{227}+\frac{(3-0.5)^2}{273}+\frac{(3-0.5)^2}{227}$$

$$= \frac{(2.5)^2}{273}+\frac{(2.5)^2}{227}+\frac{(2.5)^2}{273}+\frac{(2.5)^2}{227}$$

$$= \frac{6.25}{273}+\frac{6.25}{227}+\frac{6.25}{273}+\frac{6.25}{227}$$

$$= 0.023 + 0.001 + 0.023 + 0.001$$

$$= 0.048$$

ขั้นที่ 5 เปิดตารางหาค่าวิกฤติ

5.1 $\propto = 0.01$

5.2 df $= (r - 1)(c - 1)$

r = แถวตามแนวนอน = 2

c = แถวตามแนวตั้ง = 2

ดังนั้น df $= (2 - 1)(2 - 1) = 1 \times 1$

$= 1$

(โปรดจำไว้ว่าในกรณีตาราง 2×2 ค่า df = 1 เสมอ)

เปิดตารางจะได้ $\chi^2_{(df,\propto)} = \chi^2_{(1,0.01)} = 6.63$

ขั้นที่ 6 สรุปผลการทดสอบสมมุติฐาน

$\chi^2_{(\text{คำนวณ})}$ น้อยกว่า $\chi^2_{(\text{ตาราง})} = \chi^2_{,0.01}$

ดังนั้น จึงยอมรับ H_o ว่าเป็นจริง

หมายความว่า เพศ ***ไม่มีความสัมพันธ์***กับความคิดเห็นเกี่ยวกับการยอมรับนโยบายและตัวบุคคลในคณะรัฐบาล

กล่าวอีกนัยหนึ่ง ก็คือ เพศชาย และหญิงมีความเห็น***ที่ไม่แตกต่างกัน***นั่นเอง

2) การใช้สูตรพิเศษ

ในกรณีที่ตารางเป็น 2 × 2 สามารถใช้สูตรพิเศษได้อีกวิธีหนึ่งคือ

$$\chi^2 = \frac{N(AD - BC)^2}{(A+B)(C+D)(A+C)(B+D)}$$

ในกรณีที่จำนวนกลุ่มตัวอย่างมีค่าน้อยกว่า 40 แต่มากกว่า 20 และ E มีค่ามากกว่าหรือเท่ากับ 5 ให้ใช้ค่าแก้ของเยท (Yate's Correction) ดังมีสูตรดังนี้

$$\chi^2 \quad \frac{N(|AD - BC| - N/2)^2}{(A+B)(C+D)(A+C)(B+D)}$$

เมื่อค่าที่อยู่ใน | | คือค่าสัมบูรณ์ คือไม่คิดเครื่องหมายลบ

ตัวอย่างที่ 5 ข้อมูลจากตัวอย่างที่แล้วสามารถนำมาคำนวณตามสูตรใหม่ได้ดังนี้

เพศ	ความเห็น		รวม
	ยอมรับ	ไม่ยอมรับ	
ชาย	A 270	B 230	A+B 500
หญิง	C 276	D 224	C+D 500
รวม	A+C 546	B+D 454	N 1,000

จากสูตร $\chi^2 = \frac{N(AD-BC)^2}{(A+B)(C+D)(A+C)(B+D)}$

แทนค่าในสูตร จะได้

$$\chi^2 = \frac{1{,}000\left[(270\times 224)-(230\times 276)\right]^2}{500\times 500\times 546\times 454}$$

$$= \frac{1{,}000(60{,}480-63{,}480)^2}{500\times 500\times 546\times 454}$$

ตัดศูนย์ออกทิ้งเศษ และส่วนเพื่อให้ตัวเลขเหลือลดลงจะได้

$$= \frac{1{,}000\,(60{,}480-63{,}480)^2}{500\times 500\times 546\times 454}$$

$$= \frac{(60{,}480-63{,}480)^2}{5\times 50\times 546\times 454}$$

$$= \frac{(-3{,}000)^2}{5\times 50\times 546\times 454}$$

$$= \frac{9{,}000{,}000}{5\times 50\times 546\times 454}$$

$$= \frac{900{,}000}{6{,}197{,}200} = 0.14$$

ดังนั้นการหาค่า χ^2 จากสูตรพิเศษจะได้ = 0.145

ส่วนจากสูตรทั่วไป จะได้ = 0.048

ซึ่งมีค่ามากกว่าสูตรทั่วไปเล็กน้อย แต่สรุปผลการทดสอบ ยังเหมือนเดิมไม่เปลี่ยนแปลง คือ ยอมรับ H_o

จะเห็นได้ว่า การคำนวณตามสูตรพิเศษนี้ ***จะคำนวณได้ง่ายกว่า*** สูตรทั่วไป

2.4.5 ข้อจำกัดของไคสแควร์

การทดสอบไคสแควร์ มีข้อตกลงเบื้องต้นที่เป็นข้อจำกัด ดังนี้

1. ค่าความถี่จากการคาดหวัง (E) ไม่ควรมีค่าน้อยเกินไป เพราะว่าหากมีน้อยเกินไปจะทำให้การทดสอบมีโอกาสจะเกิดความคลาดเคลื่อนประเภทที่ 1 ได้มาก ซึ่งชาเวลสันเสนอว่า ในกรณี df $\geq$ 2 ค่า E ไม่ควรน้อยกว่า 5 และในกรณี df = 1 ค่า E ไม่ควรน้อย

กว่า 10 [2)] ในกรณีที่ค่า E มีน้อยกว่าที่กล่าวมาแล้ว ควรแก้โดยเพิ่มจำนวนตัวอย่างให้มากขึ้น หรือรวมกลุ่มข้อมูลให้มีกลุ่มหรือประเภทน้อยลง เช่น แต่เดิมแบ่งระดับการศึกษาออกเป็น ประถมศึกษา มัธยม อนุปริญญา ปริญญาตรี สูงกว่าปริญญาตรี คือมี 5 ประเภท อาจจะยุบรวมเหลือเพียง 3 กลุ่ม ก็ได้เช่น ประถมศึกษา มัธยมศึกษา อนุปริญญาและสูงกว่า เป็นต้น

ในทางปฏิบัติ อาจจะดูจำนวนกลุ่มตัวอย่างประกอบด้วย ***จำนวนช่วง คูณด้วย 5*** คือ จำนวนกลุ่มตัวอย่างที่เหมาะสม หากงานวิจัยมีกลุ่มตัวอย่างที่ใหญ่พอ อาจจะอนุโลมให้ E มีค่าน้อยกว่า 5 ได้เป็นบางค่า

2. ค่าข้อมูลที่นำมาวิเคราะห์ในแต่ละช่องจะต้องเป็นอิสระต่อกัน ซึ่งผู้วิจัยจะต้องระวังมิให้ ผู้ตอบคนเดียวมีโอกาสตอบแบบสอบถามได้มากกว่า 1 ชุด ซึ่งจะเป็นข้อมูลที่ไม่เป็นอิสระจากกัน ตัวอย่างเช่น ในการจำแนกอาชีพของผู้ตอบบางคนมีหลายอาชีพ เช่น อาจจะตอบในฐานะที่เป็นข้าราชการก็ได้ หรืออาจจะตอบในฐานะวุฒิสมาชิกก็ได้ ในกรณีนี้จะต้องให้ผู้ตอบตอบเพียงชุดเดียว และจำแนกผู้ตอบอยู่ในกลุ่มอาชีพเดียว เพื่อให้ข้อมูลแต่ละช่องเป็นอิสระจากกัน

3. ตัวอย่างจากการนำเอาไคสแควร์มาวิเคราะห์ข้อมูลในการวิจัยประเมินผลนโยบาย

พงศ์สัณห์ ศรีสมทรัพย์ และสมิหรา จิตตลดากร ได้ทำการวิจัยการประเมินผลโครงการบัณฑิตศึกษา สาขารัฐศาสตร์ คณะรัฐศาสตร์ มหาวิทยาลัยรามคำแหง ซึ่งเป็นโครงการหนึ่งของนโยบายมหาวิทยาลัยในการจัดสอนระดับบัณฑิตศึกษา โดยได้เก็บข้อมูลจากนักศึกษา 5 รุ่น รุ่นละ 40 คน ผลการศึกษามีดังนี้ [3)]

เพศ มีความสัมพันธ์กับผลการเรียน

ตารางที่ 6-6 ความสัมพันธ์ระหว่างผลการเรียนกับ เพศของนักศึกษา (เฉพาะรุ่น 1 ถึง 3)

ผลการเรียน	ชาย	หญิง	รวม
ต่ำ	21 (61.8)	1 (12.5)	22 (52.4)
สูง	13 (38.2)	7 (87.5)	20 (47.6)
รวม	34 (81.0)	8 (19.0)	42 (100)

Chi Square = 4.481, D.F. = 1, P = .034, * มีนัยสำคัญที่ระดับ .05

** ความถี่คาดหวังต่ำสุด = 3.810 ; จำนวนช่องที่ < 5 = 2 ใน 4 (50 %)

จากตารางที่ 6-6 พบว่า เพศ ที่แตกต่างกันของนักศึกษา จะมีการกระจายความถี่แตกต่างกันในเรื่องผลการเรียน อย่างมีนัยสำคัญที่ระดับ .05 กล่าวคือ ในกลุ่มที่มีผลการเรียนต่ำจะเป็นนักศึกษาชาย 61.8 % และเป็นนักศึกษาหญิงเพียง 12.5 % ขณะที่ในกลุ่มที่มีผลการเรียนสูงเป็นนักศึกษาชาย 38.2 % และเป็นหญิงถึง 87.5 % แสดงว่าเพศหญิงมีแนวโน้มที่จะมีผลการเรียนดีกว่าเพศชาย

อย่างไรก็ตาม ความถี่ที่ปรากฏในบางช่องมีค่อนข้างน้อย กล่าวคือ มีอยู่ถึงสองช่องที่ความถี่คาดหวังมีค่าต่ำกว่า 5 โดยมีค่าต่ำสุดเท่ากับ 3.810 แสดงว่าค่าต่างๆ ที่ใช้ในการวัดความสัมพันธ์ อาจมีความคลาดเคลื่อนได้

เพศ มีความสัมพันธ์กับ การนำความรู้ไปใช้ประโยชน์ในวิชา การบริหารงานของรัฐ

ตารางที่ 6-7 ความสัมพันธ์ระหว่างการนำความรู้ไปใช้ประโยชน์ในวิชาการบริหารงานของรัฐกับเพศของนักศึกษา

การใช้ประโยชน์ในวิชา การบริหารงานของรัฐ	เพศ		รวม
	ชาย	หญิง	
น้อย	5 (5.2)	6 (19.4)	11 (8.7)
มาก	91 (94.8)	25 (80.6)	116 (91.3)
รวม	96 (75.6)	31 (24.4)	127 (100)

Chi Square = 4.274, D.F. = 1, P = .038, * มีนัยสำคัญที่ระดับ .05

* ความถี่คาดหวังต่ำสุด = 2.685 : จำนวนช่องที่ < 5 = 1 ใน 4 (25 %)

จากตารางที่ 6-7 พบว่า เพศ ที่แตกต่างกันของนักศึกษา จะมีการกระจายความถี่แตกต่างกันในเรื่องระดับของการนำความรู้ที่ได้รับไปใช้ประโยชน์ในวิชาการบริหารงานของรัฐอย่างมีนัยสำคัญที่ระดับ .05 กล่าวคือ ในกลุ่มที่ตกเกณฑ์ กลุ่มนักศึกษาชายจะมีการนำความรู้ในวิชาการบริหารงานของรัฐไปใช้ประโยชน์น้อย จำนวน 5.2 % ขณะที่กลุ่มนักศึกษาหญิงจะมีการนำความรู้ดังกล่าวไปใช้ประโยชน์น้อย จำนวน 19.4 %

และ ในกลุ่มที่เกินเกณฑ์ กลุ่มนักศึกษาชาย มีการนำความรู้ดังกล่าวไปใช้ประโยชน์มาก จำนวนถึง 94.8 % ขณะที่ กลุ่มนักศึกษาหญิง จะมีการนำความรู้ดังกล่าวไปใช้ประโยชน์มาก จำนวน 80.6 %

4. สรุป

ในการวิเคราะห์ข้อมูลในการวิจัยสามารถวิเคราะห์ข้อมูลได้ 3 ประเภท คือ

1) **การวิเคราะห์ข้อมูลโดยไม่ใช้สถิติ**

2) **การวิเคราะห์ข้อมูลโดยใช้สถิติ**

3) **การวิเคราะห์แบบผสม** โดยนำเอา 1) และ 2) มาวิเคราะห์ร่วมกัน

การวิเคราะห์ข้อมูลศึกษาถึงลักษณะของ ประชากรเรียกว่า ***พารามิเตอร์ของประชากร (Population Parameters)*** หรือเรียกสั้น ๆ ว่า ***พารามิเตอร์*** ซึ่งได้แก่ ***ค่าเฉลี่ย*** (μ) ***ความแปรปรวน*** (σ^2) และ***ความเบี่ยงเบนมาตรฐาน*** (σ) ของข้อมูลในกลุ่มประชากร แต่ในหลายกรณีมีข้อจำกัดทางด้านค่าใช้จ่ายและเวลาทำให้ไม่ศึกษาลักษณะของประชากรได้ ต้องศึกษา***จากกลุ่มตัวอย่าง*** แทน คือ ***ค่าเฉลี่ย*** ($\bar{X}$) ***ความแปรปรวน*** (S^2) ***ความเบี่ยงเบนมาตรฐาน*** (S) จากนั้นจึงนำเอาค่าสถิติที่ได้มา***สรุปอ้างอิง*** ว่าเป็นของประชากร โดยการใช้ ***สถิติอนุมาน*** ทดสอบ

การใช้สถิติในการวิเคราะห์ข้อมูลในการวิจัย แบ่งสถิติได้เป็น 2 ชนิดคือ สถิติพรรณาและสถิติอนุมาน

1. **สถิติพรรณา (Descriptive Statistics)** ได้แก่ การวัดแนวโน้มเข้าสู่ส่วนกลาง การวัดการกระจาย อัตราส่วน อัตราส่วนร้อย เป็นต้น

2. **สถิติอนุมาน (Inferential Statistics)** ได้แก่ χ^2 t-test F- test เป็นต้น

สถิติอนุมานยังแบ่งออกเป็นอีก 2 ประเภท คือ ***สถิติที่ต้องใช้พารามิเตอร์ (Parametric Statistics)*** และ ***สถิติที่ไม่ต้องใช้พารามิเตอร์ (Non-parametric Statistics)***

ไคสแควร์ เป็นสถิติประเภท ***Non-parametric*** ที่นิยมใช้กันในการวิจัยวิจัยสังคม-ศาสตร์และรัฐศาสตร์ ทั้งนี้เพราะว่าใช้ง่ายและไม่จำเป็นต้องทราบลักษณะการกระจายของข้อมูลว่าเป็นแบบใด ต่างจากสถิติประเภท ***Parametric***_ จำเป็นต้องทราบก่อนว่าข้อมูลมีการกระจายในลักษณะเหมาะสมที่จะใช้สถิติชนิดนั้นหรือไม่

การจะใช้การวิเคราะห์ข้อมูลแบบใด ผู้วิจัยจำเป็น***ต้องใช้ให้เหมาะสมกับจุดมุ่งหมายของการวิจัยและลักษณะโดยธรรมชาติของข้อมูล*** ดังนั้นจำเป็นต้องศึกษาธรรมชาติ และข้อจำกัดของการวิเคราะห์ชนิดนั้นด้วย จึงจะสามารถเลือกใช้วิธีการวิเคราะห์ข้อมูลให้บรรลุจุดหมายที่ต้องการ

แบบทดสอบ

จงตอบคำถามข้างล่างนี้

1. สมมุติฐานทางสถิติ แบ่งออกเป็นกี่ชนิด จงอธิบาย พร้อมยกตัวอย่างประกอบ
2. ความคลาดเคลื่อนประเภทที่ 1 และประเภทที่ 2 คืออะไรแตกต่างกันอย่างไร จงอธิบายพร้อมทั้งยกตัวอย่างประกอบ
3. การทดสอบแบบไม่มีทิศทาง และแบบมีทิศทางคืออะไรแตกต่างกันอย่างไร จงอธิบาย พร้อมทั้งยกตัวอย่างประกอบ
4. สถิติที่ใช้พารามิเตอร์ (Parametric Statistics) และสถิติที่ไม่ใช้พารามิเตอร์ (Nonparametric Statistics) แตกต่างกันอย่างไร จงอธิบาย
5. จงยกตัวอย่างการใช้ไคสแควร์ในการวิเคราะห์ตัวแปรแบบมิติเดียว
6. การวิเคราะห์ไคสแควร์แบบตาราง (2 X 2) ต่างจากตารางชนิดอื่นๆ อย่างไร
7. ไคสแควร์สามารถนำเข้ามาใช้ปนะโยชน์ในการประเมินผลนโยบายอย่างไร

เชิงอรรถ

1) Richard J. Shavelson, **Statistical Reasoning for the Behavioral Sciences** 2d ed., (Boston : Allyn and Bacon, 1988), p.450.

2) **Ibid.**

3) พงศ์สัณห์ ศรีสมทรัพย์ และสมิหรา จิตตลดากร, **รายงานการวิจัย การประเมินผลโครงการบัณฑิตศึกษา สาขารัฐศาสตร์ คณะรัฐศาสตร์ มหาวิทยาลัยรามคำแหง** (กรุงเทพฯ : คณะรัฐศาสตร์ มหาวิทยาลัยรามคำแหง, 2536), หน้า 98, 103-105.