การทดสอบสมมติฐานด้วยไคสแควร์

● ดิเรก ศรีสุโข

บทนำ

การทดสอบสมมติฐานและการประมาณค่าเป็นกระบวนการทางสถิติที่มีวิธีการให้เลือกใช้กันหลายแบบตามความเหมาะสมกับสถานการณ์ของข้อมูลและลักษณะของตัวแปร ผู้ที่ไม่ค่อยคุ้นเคยกับวิธีการทางสถิติ แต่มีความจำเป็นที่จะต้องใช้สถิติเป็นเครื่องมือในการทำงานนั้น มักจะเรียกร้องอยากจะมีข้อเขียนที่ชี้แนะว่า ข้อมูลอย่างนั้นให้ใช้สถิติอย่างนี้ ให้เป็นเสมือนตำรากับข้าว แนวคิดแบบนี้มีผู้พยายามตอบสนองมาบ้างแล้ว โดยเขียนเป็นตำราสถิติประยุกต์ใช้กับศาสตร์นั้น ศาสตร์นี้ แต่พอเขียนมาแล้วก็เป็นตำราเล่มใหญ่พอสมควรจนเป็นปัญหากับผู้ต้องการใช้อีกเช่นกัน ทั้งนี้ก็เพราะวิธีการทางสถิติมีหลายแบบหลายวิธีจำแนกแยกย่อยกันมาก แต่ละวิธีเหมาะที่จะใช้กับสถานการณ์หนึ่ง ๆ วิธีการทางสถิติบางอย่างนำไปประยุกต์ได้หลายแบบ แต่ก็มีขอบเขตจำกัดของตัวมันเองอยู่ ผู้ที่คุ้นเคยกับวิธีการทางสถิติหลาย ๆ วิธีก็มีโอกาสที่จะเลือกใช้ได้เหมาะสมยิ่งขึ้น

วิธีการทางสถิติที่มักจะเป็นที่คุ้นเคยกันมากสำหรับการทดสอบสมมติฐานก็คือ Z-test t-test และ F-test ซึ่งส่วนใหญ่มักจะประยุกต์ใช้กับการทดสอบเกี่ยวกับค่าเฉลี่ย (μ) สำหรับข้อมูลกลุ่มเดียว สองกลุ่ม และมากกว่าสองกลุ่ม ตามสถานการณ์ของข้อมูลที่สามารถหาค่าเฉลี่ย ($\bar{X}$) ของกลุ่มตัวอย่างได้ ซึ่งวิธีการแต่ละอย่างได้พัฒนามาจากทฤษฎีการกระจายค่าสถิติสำหรับแต่ละอย่างไป

X และ χ เป็นอักษรกรีกอ่านว่า Chi (ไค) เป็นอักษรที่นักสถิตินำมาใช้เป็นสัญลักษณ์ในงานทางสถิติอย่างหนึ่งในหลาย ๆ อย่าง χ^2 (chi-square) เป็นสัญลักษณ์แทนการแจกแจงหรือการกระจายแบบหนึ่ง ซึ่งมีผู้ค้นพบ

เมื่อราว ๆ ค.ศ. 1876 และมีผู้สร้างสมการอธิบายลักษณะการกระจายแบบนี้เมื่อราว ค.ศ. 1900

$$\left[f(\chi^2) = \frac{(\chi^2)^{\frac{1}{2}(n-2)}}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} e^{-\chi^2/2}\right], 0 < \chi^2 < \alpha$$

การกระจายแบบนี้จะมีค่า $\mu = n$ และ $\sigma^2 = 2n$ เมื่อ n เป็นค่า degree of freedom การแจกแจงค่า χ^2 ตาม d.f. สำหรับการอ้างอิงจะพบได้จากหนังสือสถิติฉบับมาตรฐานทั่ว ๆ ไป

จากลักษณะการกระจายแบบไคสแควร์ดังที่กล่าวแล้วนั้น ปรากฏว่าในปัจจุบันนี้มีสถิติที่มีลักษณะการกระจายในลักษณะเช่นเดียวกับการกระจายแบบไคสแควร์มากมายหลายชนิด ทั้งที่เป็นการกระจายของตัวแปรตัวเดียว และตัวแปรหลาย ๆ ตัว ซึ่งคงไม่สามารถจะเรียบเรียงลงในเอกสารแบบกะทัดรัดฉบับนี้ได้หมด แต่อย่างไรก็ตามในกลุ่มของวิธีการทางสถิติที่มีลักษณะการแจกแจงแบบไคสแควร์นั้น มีสถิติอยู่ชุดหนึ่งซึ่งมีชื่อสอดคล้องกับสัญลักษณ์ของการกระจาย นั่นคือ การทดสอบด้วยไคสแควร์ (X^2 – test) (หมายเหตุ ชื่อของวิธีการทางสถิติที่ใช้ทดสอบในหนังสือสถิติบางเล่มใช้ X^2-test บางเล่มใช้ λ-test) ซึ่งสับสนกันไปสับสนกันมาแล้วแต่หนังสือแต่ละเล่ม ทั้งนี้อาจเป็นเพราะไคสแควร์ตัวพิมพ์ใหญ่กับ X ตัวอักษรในภาษาอังกฤษ "เอ็กซ์" มีลักษณะคล้ายกันมาก ผู้เขียนเอกสารบางคนเลยตัดปัญหาใช้ X^2-test แทน λ-test

มโนทัศน์พื้นฐานเกี่ยวกับตัวแปร

ก่อนจะระบุการใช้การทดสอบด้วยไคสแควร์นั้นควรจะได้ทำความเข้าใจกันเล็กน้อยถึงลักษณะของตัวแปรที่จะใช้ในการสรุปพาดพิง คำว่าตัวแปร (variable) ในศาสตร์ประยุกต์โดยทั่ว ๆ ไปจะหมายถึงคุณสมบัติ (property) ของ individual ซึ่งมีลักษณะแตกต่างกันในแต่ละหน่วย (A property in which individuals differ among themselves is termed variable) คำว่า individuals ในที่นี้มีความหมายถึงหน่วยที่ใช้ศึกษา อาจเป็นคน จังหวัด สัตว์ สิ่งของ บ้านเรือน ต้นไม้ อะไรก็ได้แล้วแต่ผู้ศึกษาต้องการจะศึกษา ดังนั้น ความสวย (property) ของคน (individual) ก็เป็นตัวแปรตัวหนึ่ง วัสดุที่ใช้มุงหลังคาบ้านก็เป็นตัวแปร เชื้อชาติของคนก็เป็นตัวแปร ผลผลิตข้าวแต่ละไร่ก็เป็นตัวแปร เคยเป็นหัดหรือไม่ (คน) ก็เป็นตัวแปร ค่าสถิติต่าง ๆ จะเกี่ยวข้องกับตัวแปรแตกต่างกันออกไปตามความสนใจของศาสตร์นั้น ๆ และสถิติก็เป็นวิธีการที่เป็นประดุจเครื่องมือในการสรุปอ้างอิง ประมาณค่าเกี่ยวกับตัวแปร

การศึกษาตัวแปรใด ๆ นั้น ผู้ศึกษาจะให้ความหมายตัวแปรนั้นก่อนจากนั้นจึงสร้าง

เครื่องมือหรือกำหนดวิธีการที่จะวัดตัวแปรนั้นออกมาเป็นปริมาณตัวเลขหรือเป็นลักษณะคุณภาพตามลักษณะของตัวแปร เครื่องมือและวิธีการที่จะใช้วัดตัวแปรนั้นมีมากมาย ตัวอย่างเช่น เครื่องวัดความดัน เทอร์โมมิเตอร์ แบบสอบ แบบสอบถาม แบบสัมภาษณ์ การสังเกต เครื่องชั่ง เครื่องตวง ฯลฯ เมื่อเราใช้เครื่องมือและ/หรือวิธีการวัดตัวแปรจาก individual แล้วสิ่งที่ได้ออกมาเราเรียกว่าข้อมูล (Data)

ข้อมูลที่ได้จากการวัดด้วยเครื่องมือหรือเทคนิควิธีต่าง ๆ นั้น บางอย่างก็มีลักษณะปริมาณมากน้อยกว่ากันซึ่งแทนด้วยระบบตัวเลขธรรมดาได้ เช่น ความสูงวัดเป็น ซม. เราจะพบว่ามีความมากน้อยเป็นมาตรฐานที่เข้าใจกันโดยทั่ว ๆ ไป ตัวแปรบางตัวสามารถวัดได้ในระดับอันดับเท่านั้น เช่น ความงาม ความประณีต คือ วัดได้ว่าใครงามกว่าใคร แล้วจัดอันดับเป็นที่ 1 ที่ 2 ไปจนถึงที่สุดท้าย และก็ยังมีตัวแปรบางชนิดที่วัดหรือสังเกตได้เพียงการจัดเป็นกลุ่มหรือประเภทเท่านั้น เช่น เพศ เป็นตัวแปรที่จำแนกเป็นกลุ่มได้ เป็นหญิงกับชายเท่านั้น สถานภาพสมรสวัดเป็นประเภท โสด แต่งงาน หย่าร้าง หม้าย ความคิดเห็นของคนกรุงเทพฯ เรื่องการย้ายตลาดนัดจากสนามหลวงไปไว้ที่สวนจตุจักร ก็อาจจะจำแนกเป็นเห็นด้วยกับไม่เห็นด้วย

ข้อมูลที่ได้จากตัวแปรการแบ่งกลุ่มแบ่งประเภท (Categorical Data) นี้ตามความเป็นจริง เป็นข้อมูลที่จะพบได้เสมอในงานวิจัยแทบทุกรูปแบบและทุกสาขาวิชา และเป็นที่คุ้นเคยกันมากในงานวิจัยทางสังคมศาสตร์ เป็นตัวแปรที่เข้าใจง่ายพอสมควร และก็เป็นตัวแปรที่ตรงกับปัญหาการวิจัยไม่ใช่น้อย ข้อมูลที่ได้จากการศึกษาตัวแปรนี้มักจะอยู่ในรูปของความถี่ (frequency) ซึ่งเกิดจากการนับจำนวน individual ว่าอยู่ประเภทไหนหรือกลุ่มไหน การเสนอผลการวิเคราะห์ก็เข้าใจได้ง่าย เช่น ใช้ร้อยละ สัดส่วน (proportion) ฐานนิยม (mode) การเสนอรายงานผลก็เข้าใจง่าย เช่น แผนภูมิแท่ง แผนภูมิกง ตารางระบุร้อยละ ฯลฯ (ข้อมูลที่ได้จากตัวแปรการแบ่งกลุ่มแบ่งประเภทนี้ไม่เหมาะที่จะหาค่าเฉลี่ย นอกจากผู้วิจัยจะหลงลืมไป เช่น กำหนดให้ เพศชายเป็น 1 เพศหญิงเป็น 2 ถ้าหาค่าเฉลี่ยเป็น 1.25 คงไม่มีความหมายอะไรและคงไม่มีใครวิเคราะห์อย่างนั้น)

การสรุปพาดพิงจากข้อมูลแบบ Categorical Data นี้ก็มีหลายรูปแบบ ทั้งการประมาณค่า (Estimation) และการทดสอบสมมติฐาน สำหรับข้อเขียนนี้จะมุ่งไปที่การทดสอบสมมติฐานจากข้อมูลที่ได้จากตัวแปร ในลักษณะการจำแนกประเภทเป็นจุดสำคัญ และโดยเฉพาะอย่างยิ่งจะมุ่งในการใช้ χ^2-test แบบต่าง ๆ ในการทดสอบสมมติฐานจากข้อมูล

ประเภทนี้เป็น 4 ลักษณะซึ่งจะกล่าวเรียงกันไปตั้งแต่แบบที่ 1 ถึงแบบที่ 4 ต่อไป โดยจะเพ่งเล็งการนำไปใช้เป็นจุดมุ่งหมายสำคัญ และเพื่อให้ง่ายต่อการเข้าใจจึงตัดความยุ่งยากบางประการ เช่น correction of continuity ออกจากข้อเขียนทั้งหมด

I. Chi-Square test of Homogeneity of Proportions

(การทดสอบสัดส่วนด้วยไคสแควร์)

สถานการณ์ ผู้วิจัยต้องการทดสอบสมมติฐานที่ระบุความเท่ากันของสัดส่วนของประชากรมากกว่า 2 กลุ่ม โดยที่ข้อมูลอยู่ในลักษณะของการจัดกลุ่มจัดประเภท ซึ่งจัดได้ 2 กลุ่ม หรือ 2 ประเภท (Binomial)

ตัวอย่าง ผู้วิจัยต้องการหาผลสรุปว่า นักศึกษาชายที่สูบบุหรี่กับไม่สูบบุหรี่ ระหว่างนักศึกษาวิชาเอกพลศึกษา นักศึกษาวิชาเอกสาธารณสุขศาสตร์ นักศึกษาวิชาเอกช่างเครื่องยนต์จะมีสัดส่วนเท่า ๆ กันหรือไม่ (สถานการณ์เช่นนี้ชี้ให้เห็นว่า ตัวแปรเป็นแบบแยกประเภทคือสูบบุหรี่กับไม่สูบบุหรี่ และจำนวนกลุ่มที่ต้องการศึกษาเป็น 3 กลุ่ม)

สมมติฐานทางสถิติจะมีลักษณะดังนี้

$H_0 : P_1 = P_2 = P_3$

$H_1 : H_0$ ไม่เป็นจริง

สมมติว่า ตั้งระดับความมีนัยสำคัญ $\alpha = .05$

ในที่นี้ P_j หมายถึงสัดส่วนของผู้สูบบุหรี่ของประชากรกลุ่มที่ j

สมมติว่าผู้วิจัยใช้วิธีสำรวจจากการสุ่มตัวอย่าง โดยสุ่มตัวอย่างเป็น 3 กลุ่ม ตามความมากน้อยของประชากร และก็ขอสมมติต่อไปว่า สุ่มตัวอย่างนักศึกษาวิชาเอกพลศึกษา 280 คน พบว่ามีผู้สูบบุหรี่ 152 คน ไม่สูบ 128 คน นักศึกษาวิชาเอกสาธารณสุขศาสตร์ 160 คน สูบบุหรี่ 72 คน ไม่สูบ 88 คน และนักศึกษาวิชาเอกช่างเครื่องยนต์ 260 คน พบว่าสูบบุหรี่ 176 คน ไม่สูบ 84 คน เมื่อนำข้อมูลมาเสนอในตารางจะมีลักษณะดังนี้

ตารางที่ 1 *จำนวนนักศึกษาที่สูบบุหรี่จำแนกตามสาขาวิชาที่เรียน*

	พลศึกษา	**สาธารณสุขฯ**	**ช่างเครื่องยนต์**	**รวม**
สูบบุหรี่	152	72	176	400
ไม่สูบบุหรี่	128	88	84	300
รวม	280	160	260	700

จากกลุ่มตัวอย่างที่สุ่มมานี้ เราจะพบว่า สถิติจากกลุ่มตัวอย่างมีสัดส่วนของผู้สูบบุหรี่ จำแนกตามกลุ่มดังนี้

กลุ่มนักศึกษาพลศึกษา

$\hat{p}_1 = \frac{152}{280} = 0.54$

กลุ่มนักศึกษาสาธารณสุขศาสตร์

$\hat{p}_2 = \frac{72}{160} = 0.45$

กลุ่มนักศึกษาช่างเครื่องยนต์

$\hat{p}_3 = \frac{176}{260} = 0.68$

ผลการสำรวจดังที่บันทึกไว้ในตาราง 1 มักจะเรียกกันโดยทั่วไปว่าเป็น observed frequency (O_{ij})

ลองนึกภาพว่า ถ้าสัดส่วนของผู้สูบบุหรี่ของกลุ่มตัวอย่างทั้ง 3 กลุ่มเท่ากัน ลักษณะจะเป็นอย่างไร เราจะสร้างตารางได้ดังนี้

การคำนวณจำนวนความถี่ที่คาดหวัง (เมื่อคาดว่าสัดส่วนเท่ากัน) นี้คำนวณอย่างง่ายโดยวิธีบัญญัติไตรยางค์ธรรมดา ตัวอย่างเช่น การคำนวณว่า ถ้าสัดส่วนเท่ากัน กลุ่มนักศึกษาพลศึกษาน่าจะมีจำนวนผู้สูบบุหรี่เท่าไร

กลุ่มตัวอย่างจำนวน 700 คน จะเป็นผู้สูบบุหรี่ 400 (ดูจากผลรวม)

ถ้ากลุ่มพลศึกษา 280 คน น่าจะมีผู้สูบบุหรี่ $\frac{400}{700} \times 280 = 160$ คน

(ผู้วิเคราะห์จะต้องคำนวณ E_{ij} ลงในตารางให้ครบดังที่เสนอไว้ในตัวอย่างนี้)

จากตาราง observed frequency และ expected frequency ดังที่เสนอไว้ในตัวอย่างนั้น เป็นการเตรียมข้อมูลพร้อมแล้วเพื่อจะใช้ทดสอบสมมติฐาน เพราะสถิติที่จะใช้ทดสอบระบุว่า

$$X^2 = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad \text{..........(+)}$$

ตารางจำนวนผู้สูบบุหรี่เมื่อคาดว่าทุกกลุ่มมีสัดส่วนเท่ากัน (Expected frequency) (E_{ij})

	พลศึกษา	สาธารณสุขศาสตร์	ช่างเครื่องยนต์	รวม
สูบบุหรี่	160.00	91.43	148.57	400
ไม่สูบบุหรี่	120.00	68.57	111.43	300
รวม	280.00	160.00	260.00	700

บางทีก็เขียนสูตรย่อ ๆ ว่า

$$X^2 = \Sigma \frac{(O-E)^2}{E}$$

เมื่อ O_{ij} คือ observed frequency ในแถวที่ i และ column ที่ j

E_{ij} คือ expected frequency ในแถวที่ i และ column ที่ j เช่นเดียวกัน

เมื่อนำค่า O_{ij} และ E_{ij} มาแทนค่าในสูตร (1) จะได้ดังนี้

$$X^2 = \frac{(152-160)^2}{160} + \frac{(72-91.43)^2}{91.43} + \frac{(176-148.57)^2}{148.57} + \frac{(128-120)^2}{120} + \frac{(88-68.57)^2}{68.57} + \frac{(84-111.43)^2}{111.43}$$

$$X^2 = \frac{64}{160} + \frac{377.53}{91.43} + \frac{752.40}{143.57} + \frac{64}{120} + \frac{377.53}{68.57} + \frac{752.40}{111.43}$$

$$X^2 = 0.4 + 4.13 + 5.06 + 0.53 + 5.51 + 6.77$$

$$X^2 = 22.40$$

ค่า X^2 ที่คำนวณได้นี้ เมื่อนำไปเปรียบเทียบกับตาราง χ^2 ก็จะทำให้ผู้วิจัยตัดสินใจได้ว่าจะปฏิเสธสมมติฐาน $H_O: P_1 = P_2 = P_3$ หรือไม่ จากตัวอย่างนี้สมมติว่า ระบุ $\alpha = .05$ เมื่อเปิดค่า χ^2 จาก df = 2 จะได้ค่า $\chi^2 = 5.99$ เมื่อเปรียบเทียบกับค่า X^2 ที่คำนวณได้นั้นมีค่า 22.40 ซึ่งจะพบว่าค่าที่คำนวณได้มีค่ามากกว่าค่าวิกฤต ดังนั้นผู้วิจัยจึงตัดสินใจปฏิเสธ H_O ซึ่งระบุว่า $P_1 = P_2 = P_3$ และยอมรับ H_1 เป็นความจริง

สูตรที่ใช้ในการคำนวณ

สูตรที่ใช้ในการคำนวณ $X^2 = \Sigma \frac{(O-E)^2}{E}$ ดังที่แสดงไว้ในตัวอย่างนั้นเป็นสูตรแรกเริ่มตามคำจำกัดความและแสดงไว้เพื่อความกระจ่างในแนวคิด

มีสูตรคำนวณที่ง่ายกว่าและได้ค่าเท่ากัน คือ $X^2 = \Sigma\Sigma \frac{O_{ij}^2}{E_{ij}} - N$

เมื่อ N คือจำนวนตัวอย่างทั้งหมด

ทดลองคำนวณในสูตรนี้จะได้ดังนี้

$$X^2 = \frac{152^2}{160} + \frac{72^2}{91.43} + \frac{176^2}{148.57} + \frac{128^2}{120} + \frac{88^2}{68.57} + \frac{84}{111.43} - 700$$

$$X^2 = 144.4 + 56.69 + 208.49 + 136.53 + 112.94 + 63.33 - 700$$

$$X^2 = 722.38 - 700$$

$$= 22.38$$

(ซึ่งมีค่าเท่ากันกับที่คำนวณได้จากสูตรเดิม แต่มีทศนิยมต่างกันบ้างเกิดขึ้นเพราะการปัดทศนิยม)

หมายเหตุ การใช้ X^2-test of Homogeneity of propertions ดังที่ให้ตัวอย่างไว้แล้วนี้มีข้อ

จำกัดและข้อตกลงเบื้องต้น (Assumptions) ดังนี้คือ

1. Independent between samples
2. Independent within samples
3. Binomial

*4. Expected values greater than 5

(E_{ij} ทุกช่องต้องมีค่ามากกว่า 5)

การทดสอบภายหลัง (Post-hoc analysis)

การทดสอบ $H_O: P_1 = P_2 = P_3 \ldots P_K$ ดังตัวอย่างนี้ ถ้าผลการทดสอบด้วยไคสแควร์ให้ค่ามากกว่าค่าวิกฤติดังระบุไว้ในตาราง ผู้ทดสอบก็จะปฏิเสธ H_O และยอมรับว่า H_1 เป็นความจริง ก็ให้ผลสรุปเพียงแต่ว่า H_O ไม่จริงเท่านั้น ซึ่งเป็นผลสรุปที่ยังไม่ละเอียดเพียงพอ นักวิจัยมักจะต้องการวิเคราะห์โดยละเอียดต่อไปว่า ความแตกต่างดังกล่าวมาจากส่วนใดแน่ วิธีการวิเคราะห์ในรายละเอียดเพิ่มเติม หลังจากปฏิเสธ H_O นั้นเรียกกันว่า Post-hoc analysis (มีผู้ให้ความหมายในภาษาไทยว่า การทดสอบภายหลัง)

เทคนิคสำคัญสำหรับการวิเคราะห์ภายหลังนี้เรียกว่า Multiple Comparision (การเปรียบเทียบพหุคูณ) ซึ่งมีหลายรูปแบบด้วยกัน รูปแบบหนึ่งของการเปรียบเทียบพหุคูณที่ใช้กันมากคือ การจับคู่เปรียบเทียบ (Paired Comparison)

Post-hoc analysis มีหลักเกณฑ์สำคัญอยู่ประการหนึ่งในหลาย ๆ ประการคือ การควบคุม Type I error ให้เป็นไปตาม α ที่กำหนดไว้ ซึ่งมีความหมายว่าไม่ว่าจะมีจำนวนการเปรียบเทียบกันสักกี่ชุดก็ตามจะต้องจำกัดไม่ให้มีโอกาสผิดพลาดเกินกว่า α ที่กำหนดไว้

ตัวอย่างจากที่วิเคราะห์ไว้นี้ สมมติว่านักวิจัยต้องการวิเคราะห์เพิ่มเติมแบบรายคู่ เขาจะวิเคราะห์ได้ 3 คู่ คือ

$$\psi_1 = P_1 - P_2$$
$$\psi_2 = P_1 - P_3$$
$$\psi_3 = P_2 - P_3$$

ซึ่งค่า $\hat{\psi}_1 = \hat{P}_1 - \hat{P}_2$

และ $\mathrm{Var}(\hat{\psi}) = \mathrm{Var} \sum_{k=1}^{k} C_k P_k$

$$= \Sigma C_k^2 \mathrm{var}(P_k)$$

ในกรณีของสัดส่วนแล้ว

$$SE_\psi^2 = \frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}$$

เมื่อ $\hat{q}_i = 1 - \hat{p}_i$

ค่า $\hat{\psi}$ และค่า $SE_{\hat{\psi}}$ จะเป็นสถิติพื้นฐานที่สามารถนำมาหาค่าสถิติเพื่อใช้ทดสอบ $H_O: \psi_i = 0$ ได้ดังนี้คือ

$$Z = \frac{\hat{\psi}_i}{\sqrt{SE^2\hat{\psi}_i}}$$

จากค่าสถิติที่คำนวณไว้ในตอนแรกและความหมายของ contrast ดังที่ระบุไว้แล้ว เราจะคำนวณค่าสถิติได้ดังนี้

$$Z_1 = \frac{.54 - .45}{\sqrt{\frac{.54 \times .46}{280} + \frac{.45 \times .55}{160}}}$$

$$= \frac{0.09}{0.49}$$

$$= 1.84$$

$$Z_2 = \frac{.54 - .68}{\sqrt{\frac{.54 \times .46}{280} + \frac{.68 \times .32}{260}}}$$

$$= \frac{.14}{.042} = 3.33^*$$

$$Z_3 = \frac{.45 - .68}{\sqrt{\frac{.45 \times .55}{160} + \frac{.68 \times .32}{260}}}$$

$$= \frac{.23}{.048}$$

$$= 4.79^*$$

เพื่อ control ให้ $\alpha = .05$ ในที่นี้ขอเสนอให้ใช้ค่าวิกฤตในการตัดสินการทดสอบแต่ละคู่ด้วยเกณฑ์ของ Marascuilo's $\sqrt{\chi^2_{k-1}}$ เป็นค่าตัดสิน

ซึ่งในที่นี้ ค่าวิกฤตจะเท่ากับ $\sqrt{\chi^2_2} = \sqrt{5.99} = 2.45$ นั่นคือ ถ้า Z ที่คำนวณได้ตัวใดมีค่าเกินกว่าค่าวิกฤตนี้แล้ว ความแตกต่างของคู่นั้นจะมีนัยสำคัญทางสถิติ

จากตัวอย่างที่คำนวณไว้นี้ ผู้วิจัยสามารถสรุปผลและตัดสินใจได้ดังนี้

ก. ไม่สามารถปฏิเสธ $H_O: P_1 = P_2$ หรือ $H_O: \psi_1 = 0$ ได้

ข. ปฏิเสธ $H_O: P_1 = P_3$
หรือ $H_O: \psi_2 = 0$
(ความแตกต่างมีนัยสำคัญทางสถิติ)

ค. ปฏิเสธ $H_O: P_2 = P_3$
หรือ $H_O: \psi_3 = 0$
(ความแตกต่างมีนัยสำคัญทางสถิติ)

การเสนอผลการวิเคราะห์

ในกรณีที่ผู้วิเคราะห์มีภารกิจเพิ่มเติมในการเขียนรายงานการวิจัยเสนอให้ผู้อ่านหรือผู้สนใจทราบผลการวิเคราะห์นั้น ผู้วิจัยสามารถเลือกรูปแบบได้หลายอย่างตามสถานการณ์

ปัญหาการวิเคราะห์มีปัญหาเดียวหรือเพียง 2-3 ปัญหา ผู้วิเคราะห์อาจใช้แบบในตารางที่ 1 เสนอผล โดยมีช่องแสดงค่า X^2 ไว้ด้านหลังอีก 1 ช่อง พร้อมทั้งเสนอค่า X^2 ที่คำนวณได้ ถ้าผลสรุปจากการทดสอบมีนัยสำคัญทางสถิติแล้ว มักจะนิยมใส่เครื่องหมาย * ไว้เหนือตัวเลขที่แสดงค่าของ X^2 และถ้ามีการทดสอบภายหลัง ก็เสนอตารางเพิ่มการวิเคราะห์ภายหลังอีก 1 ตาราง

แต่ถ้าปัญหาการวิเคราะห์แบบเดียวกันมีจำนวนมาก เช่น 10 ปัญหาขึ้นไป ถ้าจะเสนอ 10 ตาราง ก็ดูจะมากเกินไป ผู้วิเคราะห์อาจเสนอเป็นตารางเดียวได้ โดยเสนอร้อยละของแต่ละปัญหาและแต่ละกลุ่ม พร้อมทั้งผลรวมและผนวกท้ายด้วยค่า X^2 อีก 1 ช่อง ถ้าหากปัญหาข้อใดมีผลสรุปความแตกต่างระ-

หว่างสัดส่วนว่ามีนัยสำคัญทางสถิติก็ใส่ * ไว้ที่ตัวเลขแสดงค่าการคำนวณไคสแควร์ของตัวนั้น

II. Chi-Square test of Homogeneity of Distributions

(การทดสอบการแจกแจงด้วยไคสแควร์)

สถานการณ์ ผู้วิจัยต้องการทดสอบสมมติฐานที่ระบุความเท่ากันของการแจกแจงของประชากรตั้งแต่ 2 กลุ่มขึ้นไปโดยที่ข้อมูลอยู่ในลักษณะการจัดกลุ่มหรือจัดประเภทได้มากกว่า 2 ประเภท (Multinomial)

ตัวอย่างสถานการณ์ที่ 2.1 ผู้วิจัยต้องการทดสอบลักษณะการกระจายในงานอดิเรกที่ชอบมากที่สุด (จำแนกเป็น ก. อ่านหนังสือ ข. เล่นกีฬา ค. ดูโทรทัศน์ ง. อื่น ๆ) ของนักเรียนมัธยมศึกษาจากประชากร 5 กลุ่ม คือ ภาคเหนือ ภาคกลาง ภาคตะวันออกเฉียงเหนือ กรุงเทพฯ ภาคใต้ โดยสุ่มตัวอย่างจากประชากร (นักเรียนมัธยม) ดังที่ระบุไว้ลักษณะตารางรวบรวมข้อมูลจะมีลักษณะดังนี้

ประเภทงานอดิเรก	เหนือ	กลาง	อีสาน	กทม.	ใต้	รวม
อ่านหนังสือ	0_{11}	0_{12}	0_{13}	0_{14}	0_{15}	$0_{1.}$
เล่นกีฬา	0_{21}	0_{22}	0_{23}	0_{24}	0_{25}	$0_{2.}$
ดูโทรทัศน์	0_{31}	0_{32}	0_{33}	0_{34}	0_{35}	$0_{3.}$
อื่น ๆ	0_{41}	0_{42}	0_{43}	0_{44}	0_{45}	$0_{4.}$
รวม	$0_{.1}$	$0_{.2}$	$0_{.3}$	$0_{.4}$	$0_{.5}$	$0_{..}$

ตัวอย่างที่ 2.1 นี้ จะพบว่ามีจำนวนกลุ่มตัวอย่าง C = 5 และจำนวนประเภทเป็น 4 ประเภท

การทดสอบว่าลักษณะการแจกแจงเป็นเช่นเดียวกันทุกกลุ่มหรือไม่เป็นปัญหาที่ใช้การวิเคราะห์ด้วย X^2-test เช่นเดียวกัน และมีสูตรคำนวณเหมือนกันคือ

$$X^2 = \sum_{j=1}^{c} \sum_{i=1}^{r} \frac{(O-E_{ij})^2}{E_{ij}}$$

หรือ

$$X^2 = \sum_{j=1}^{c} \sum_{i=1}^{r} \frac{O_{ij}^2}{E_{ij}} - N$$

ตัวอย่างสถานการณ์ 2.2 นักวิจัยต้องการหาผลสรุปว่า การแจกแจงของความคิดเห็นของครู ผู้บริหารการศึกษา และผู้ปกครองในการจัดสอนภาษาอังกฤษในชั้นประถมต้น มีลักษณะเป็นอย่างเดียวกันหรือไม่ สมมติว่าผลการสำรวจจากกลุ่มตัวอย่างเป็นดังตารางต่อไปนี้

ตาราง 2 *จำนวนครู ผู้บริหารและผู้ปกครองจำแนกตามความคิดเห็นต่อการสอนภาษาอังกฤษในชั้นประถมตอนต้น*

ความเห็นต่อการสอนภาษาอังกฤษ ป.ต้น	ครู	ผู้บริหาร	ผู้ปกครอง	รวม
เห็นด้วยอย่างยิ่ง	70	45	102	217
เห็นด้วย	80	42	82	204
ไม่ออกความเห็น	32	41	41	114
ไม่เห็นด้วย	28	38	29	95
ไม่เห็นด้วยอย่างยิ่ง	10	26	27	63
รวม	220	192	281	693

วิธีการดำเนินการทดสอบจะเป็นเช่นเดียวกับที่ให้ตัวอย่างไว้แล้วคือ สร้างตาราง expected frequency ก่อน แล้วจึงคำนวณ $X^2 = \Sigma\Sigma\frac{(O-E)^2}{E}$ หรือ $X^2 = \Sigma\frac{O^2}{E} - N$ แล้วเปรียบเทียบกับค่าวิกฤตจากตารางของ χ^2 ตาม α ที่ระบุไว้ตาม degree of freedom ซึ่งจะได้จากการคำนวณ $df = (c-1)(r-1)$ เมื่อ c คือจำนวน column (ในที่นี้มี 3 กลุ่ม) และ r คือจำนวน row (ในที่นี้มี 5 ประเภท) ซึ่งจากตัวอย่าง $df = (3-1)(5-1) = 8$

หมายเหตุ 1. ถ้าผู้วิจัยต้องการวิเคราะห์แบบ Proportions ต้องการทราบสัดส่วนของผู้เห็นด้วยว่าเท่ากันหรือไม่ ผู้วิจัยอาจยุบรวมตัวแปรเป็นแบบสัดส่วนได้คือ รวมประเภทเห็นด้วยอย่างยิ่งกับเห็นด้วยเข้าด้วยกันเป็นกลุ่มหนึ่ง ส่วนจำนวนที่เหลือเป็นอีกกลุ่มหนึ่งก็สามารถวิเคราะห์ลงมาบางส่วนเป็นแบบวิเคราะห์สัดส่วนได้ เมื่อเห็นว่าจะสื่อความหมายในการวิจัยยิ่งขึ้น

หมายเหตุ 2. การใช้ X^2-test of Homogeneity of Distribution ดังตัวอย่างที่เสนอไว้นี้

มีข้อจำกัดและข้อตกลงเบื้องต้น (Assumptions) ดังต่อไปนี้คือ

1. Independent between samples
2. Independent within sample
3. Multinomial
4. Expected value greater than 5

(E_{ij} ทุกช่องจะต้องมีค่ามากกว่า 5)

III. Chi-Square test of Goodness of Fit (การทดสอบสารูปสนิทสุดด้วยไคสแควร์)

จากการใช้ไคสแควร์ทดสอบสมมติฐานทั้งสองแบบดังที่ผ่านมาแล้วนั้น เราจะสังเกตได้ว่าต้องคำนวณความถี่ที่คาดหวัง (expected frequency) หรือ E_{ij} ด้วยเสมอ จึงจะคำนวณหาค่าสถิติที่เรียกว่า X^2 ได้ ผู้วิเคราะห์บางท่านอาจนึกถึงปัญหาการวิจัย ซึ่งต้องการทดสอบการกระจายหรือแจกแจงความถี่จากที่สังเกต (O_{ij}) กับความถี่ที่ได้จากความคาดหวังด้านทฤษฎีได้บ้างหรือไม่ คำตอบก็คือใช้ได้เช่นกัน การทดสอบแบบนี้มีชื่อการทดสอบดังปรากฏในข้อ III นี้แล้ว และมีผู้ให้ความหมายในภาคภาษาไทยว่า "การทดสอบสารูปสนิทสุดด้วยไคสแควร์"

สถานการณ์ตัวอย่าง 3.1 ผู้ผสมกล้วยไม้คนหนึ่งคาดไว้ว่า ถ้าใช้กล้วยไม้พันธุ์ A ผสมกับกล้วยไม้พันธุ์ B แล้ว ผลที่ได้จะเป็นกล้วยไม้พันธุ์ A 1 ส่วน กล้วยไม้พันธุ์ผสม 2 ส่วน และกล้วยไม้พันธุ์ B 1 ส่วน เมื่อทดลองผสมกล้วยไม้พันธุ์ A กับพันธุ์ B ชุดหนึ่งแล้วได้ผลการผสมคือ ได้กล้วยไม้พันธุ์ A จำนวน 18 ต้น กล้วยไม้พันธุ์ผสมจำนวน 38 ต้น และได้กล้วยไม้พันธุ์ B จำนวน 20 ต้น ซึ่งเมื่อนำเสนอในรูปตารางแล้วจะได้ลักษณะข้อมูลดังนี้

	ผลการผสม (O_i)	ความคาดหวัง (E_i)
พันธุ์ A	18	19
พันธุ์ผสม	38	38
พันธุ์ B	20	19
รวม	76	76

การทดสอบใช้ $X^2 = \Sigma\frac{(O-E)^2}{E}$ เช่นเดียวกัน

$$X^2 = \frac{(18-19)^2}{19} + \frac{(38-38)^2}{38} + \frac{(20-19)^2}{19}$$

$$= \frac{2}{19} = 0.105$$

ตัวอย่างที่ 3.2 บรรณารักษ์ห้องสมุดในสถาบันการศึกษาคนหนึ่งคาดว่า ผู้มายืมหนังสือในห้องสมุดจะมีสัดส่วนดังนี้คือ เช้า (8.00-10.00 น.) ประมาณ 10% สาย (10.00-12.00 น.) ประมาณ 25% เที่ยง (12.00-13.00 น.) ประมาณ 30% บ่าย (13.00-15.00 น.) ประมาณ 25% และเย็น (15.00-17.00 น.) ประมาณ 10% และต้องการจะทดสอบว่าความคาดหวังดังกล่าวจะตรงกับสภาพความเป็นจริงหรือไม่ จึงขอให้เจ้าหน้าที่บันทึกเวลาและจำนวนผู้มายืมหนังสือห้องสมุดทั้งสิ้น 15 วัน ได้ผลบันทึกดังนี้

ตารางที่ 3 *เปรียบเทียบจำนวนผู้ยืมหนังสือกับจำนวนที่คาดหวังจำแนกตามเวลา*

ระยะเวลา	จำนวนผู้ยืมหนังสือ	ร้อยละที่คาดหวัง	จำนวนตามที่คาดหวัง
เช้า	201	10	184.30
สาย	179	25	460.75
เที่ยง	555	30	552.90
บ่าย	442	25	460.75
เย็น	166	10	184.30
รวม	1843	100	1843.00

จากตารางนี้ ลักษณะเป็นกลุ่มตัวอย่างกลุ่มเดียว O_i คือ จำนวนผู้ยืมหนังสือ ส่วน E_i (จำนวนที่คาดหวัง) คำนวณได้จากร้อยละที่คาดหวังตามระยะเวลาที่กำหนด การทดสอบสารูปสนิทสุดใช้ X^2-test ทดสอบเช่นกัน

$$X^2 = \Sigma\frac{(O-E)^2}{E}$$

หรือ จากสูตรการคำนวณ

$$X^2 = \Sigma\frac{O^2}{E} - N$$

ในที่นี้ $X^2 = 1847.82 - 1843 = 4.82$ ซึ่งเมื่อเทียบกับค่า X^2 จากตารางที่ $\alpha = .05$ และ df = 4 จะมีค่าเป็น 9.49 แล้วพบว่าค่า X^2 ที่คำนวณได้น้อยกว่าค่าวิกฤต การตัดสินใจก็คือ ไม่สามารถปฏิเสธ H_o ได้ (ไม่มีความแตกต่างอย่างมีนัยสำคัญทางสถิติ) ผลการทดสอบสารูปสนิทสุดไม่แตกต่างจากความคาดหวังของบรรณารักษ์

IV. Chi-Square test of Independence (การทดสอบความเป็นอิสระด้วยไคสแควร์)

การใช้ X^2-test อีกลักษณะหนึ่ง ซึ่งน่าจะทำความคุ้นเคยเข้าไว้คือ การทดสอบความเป็นอิสระ (independent) ระหว่างตัวแปร 2 ตัว ซึ่งต่างก็มีลักษณะเป็นตัวแปรแบบการจำแนกประเภททั้งคู่ ซึ่งถ้าผลการทดสอบได้ผลสรุปว่า ตัวแปรทั้ง 2 ตัวเป็นอิสระต่อกันก็มีความหมายในนัยเดียวกันกับความสัมพันธ์เป็นศูนย์ แต่ถ้าการทดสอบส่งผลว่า ตัวแปรทั้ง 2 ตัวไม่มีอิสระต่อกัน (dependent) ก็จะมีความหมายว่า ความสัมพันธ์มีอยู่จริงอย่างมีนัยสำคัญทางสถิติ (นักสถิติบางท่านเรียกชื่อการทดสอบแบบนี้เป็น Chi-square test of Association)

สถานการณ์ ปัญหาการวิเคราะห์ความสัมพันธ์นี้จะเกิดขึ้นจากการศึกษาตัวแปร 2 ตัว **จากกลุ่มตัวอย่างเดียวกัน** ตัวอย่างเช่น นักวิจัยทางจิตวิทยาการศึกษาต้องการศึกษาว่า ความเอาใจใส่ต่อการเรียนของผู้ปกครองมีความสัมพันธ์กับผลการเรียนภาษาไทยของนักเรียนหรือไม่ ผู้วิจัยใช้แบบสำรวจความเอาใจใส่ในการเรียนของผู้ปกครองนักเรียนที่เรียนวิชาภาษาไทยวิชาเดียวกัน (จากครูสอนคนเดียวกัน) จำนวน 120 คน ได้ผลดังนี้

ตารางที่ 4 *จำนวนนักเรียนจำแนกตามผลการเรียนภาษาไทยและความเอาใจใส่การเรียนของผู้ปกครอง*

การเอาใจใส่ของผู้ปกครอง	ผลการเรียนภาษาไทย (นักเรียน)		รวม
	ดี	อ่อน	
เอาใจใส่	48	26	74
ไม่เอาใจใส่	18	28	46
รวม	66	54	120

การทดสอบความเป็นอิสระระหว่าง ความเอาใจใส่กับผลการเรียนใช้ X^2-test ซึ่งมีสูตรการคำนวณเช่นเดียวกันคือ

$$X^2 = \Sigma\Sigma \frac{(O-E)^2}{E}$$

จากตัวอย่างนี้ การคำนวณ E_{ij} มีค่าดังนี้

	ดี	อ่อน
เอาใจใส่	40.7	33.3
ไม่เอาใจใส่	25.3	20.7

ซึ่งได้ $X^2 = 7.59$

เมื่อนำมาเปรียบเทียบกับค่า χ^2 จากตารางซึ่งมีค่า 3.84 เมื่อ $\alpha = .05$ และ df = 1 แล้วจะเห็นว่าค่าที่คำนวณได้มากกว่าค่าวิกฤต นั่นคือตัดสิน reject H_O ที่ระบุว่าความเอาใจใส่ของผู้ปกครองไม่มีความสัมพันธ์กับผลการเรียนภาษาไทยของนักเรียน หรือจะกล่าวอีกด้านหนึ่งคือ ความเอาใจใส่ของผู้ปกครองมีความสัมพันธ์กับผลการเรียนภาษาไทยของนักเรียนอย่างมีนัยสำคัญทางสถิติ

การทดสอบ Independent นี้ผู้วิเคราะห์จำนวนมากมักจะหยุดวิเคราะห์เพียงเท่าที่เสนอไปแล้วคือสรุปว่า ตัวแปร 2 ตัวมีความสัมพันธ์กัน แต่ความเป็นจริงแล้ว ถ้าความสัมพันธ์มีจริงแล้วน่าจะระบุ degree of Association หรือประมาณความสัมพันธ์ได้

สำหรับตัวแปรที่เป็นตาราง 2×2 ดังตัวอย่างนี้เราสามารถหาค่าความสัมพันธ์ได้เรียกว่า Phi Coefficient โดยคำนวณได้จากสูตร

$$\hat{\phi} \quad \sqrt{\frac{X^2}{N}}$$

สำหรับตัวอย่างที่ให้นี้

$\hat{\varnothing} \quad \sqrt{\frac{7.59}{120}}$

$= 0.25$

การให้ความหมายของ $\hat{\varnothing}$ จะมีลักษณะเช่นเดียวกับค่าสหสัมพันธ์แบบ Pearson's Product Moment Coefficient of Correlation (r_{XY})

การทดสอบ Independent แบบ 2×2 นี้ เราสามารถสรุปขั้นตอนการวิเคราะห์ได้ดังนี้

(1) ตั้งสมมติฐาน $H_O: \varnothing = 0, H_1: \varnothing \neq 0$

(2) ใช้ X^2 ทดสอบ H_O ตาม α ที่กำหนด

(3) ถ้า Reject H_O คำนวณค่า $\hat{\varnothing} = \sqrt{\frac{X^2}{N}}$

(4) รายงานผลการวิจัย

บรรณานุกรม

Hays, W.L. *Statistics.* New York : Holt, Rinehart and Winston, 1963.

Marascuilo. L.A. *Large - Sample Multiple Comparisons. Psychological Bulletion,* 1966, Vol. 15, No. 5, 280 - 290

______ *Statistical Methods for Behavioral Science Research.* New York : MaGraw-Hill, 1971.

______ & Mc Sweeney, M. *Nonparametric and Distribution - Free Methods for the Social Sciences.* Monterey : Brook/Coke Publishing Co., 1977.