ตัวอย่าง
การทดสอบแบบไคสแควร์(The Chi-Square Test)

ตัวอย่างที่ 1

ในการสำรวจนิสัยการอ่านหนังสือพิมพ์ของนักศึกษามหาวิทยาลัยแห่งหนึ่ง ซึ่งได้จำแนกนักศึกษาที่นิยมอ่านหนังสือพิมพ์ต่างๆออกเป็น 5 กลุ่ม ดังนี้

กลุ่ม ก : นิยมอ่านหนังสือพิมพ์เกี่ยวกับเศรษฐกิจ

กลุ่ม ข : นิยมอ่านหนังสือพิมพ์ไทยรัฐ,เดลินิวส์

กลุ่ม ค : นิยมอ่านหนังสือพิมพ์มติชน,สยามรัฐ

กลุ่ม ง : นิยมอ่านหนังสือพิมพ์ Nation,Bangkok Post

กลุ่ม จ : ไม่นิยมอ่านหนังสือพิมพ์

สุ่มตัวอย่างนักศึกษามหาวิทยาลัยแห่งนี้จากทุกชั้นปี จำนวน 1000 คน ผลจากการสอบถามปรากฏว่าเป็นนักศึกษาที่นิยมอ่านหนังสือพิมพ์ในกลุ่มต่างๆ ดังนี้ กลุ่ม ก 200 คน กลุ่ม ข 250 คน กลุ่ม ค 150 คน กลุ่ม ง 50 คน และ กลุ่ม จ 350 คน จากผลการสำรวจข้างต้นจะสรุปที่ระดับนัยสำคัญ 10% ได้หรือไม่ว่า กลุ่มต่างๆของนักศึกษามหาวิทยาลัยแห่งนี้ ซึ่งจำแนกตามความนิยมการอ่านหนังสือพิมพ์ข้างต้นมีจำนวนไม่แตกต่างกัน (อุมาพร จันทศร, 2542)

วิธีทำ

ต้องการทราบว่านักศึกษามหาวิทยาลัยแห่งนี้มีความนิยมการอ่านหนังสือพิมพ์ซึ่งแยกเป็นประเภทต่างๆ ก ถึง จ ไม่แตกต่างกัน เป็นจริงหรือไม่ สามารถตั้งสมมติฐานได้ว่า

H_0: ไม่มีความแตกต่างระหว่างกลุ่มต่างๆของนักศึกษามหาวิทยาลัยแห่งนี้ หรือสัดส่วนของนักศึกษาในกลุ่มต่างๆมีค่าเท่ากัน โดยให้ P_i : i =1,...,5 เป็นสัดส่วนของนักศึกษาจากกลุ่ม i

นั่นคือ $H_0: P_1 : P_2 : P_3 : P_4 : P_5 = 1 : 1 : 1 : 1 : 1$

$H_1: P_1 : P_2 : P_3 : P_4 : P_5 \neq 1 : 1 : 1 : 1 : 1$

เป็นการทดสอบสัดส่วนหลายค่าในประชากร 1 กลุ่ม และข้อมูลมีลักษณะเป็นความถี่หรือข้อมูลแจกนับ ฉะนั้นใช้การทดสอบของไคสแควร์ โดยสถิติที่ใช้ทดสอบ คือ

$$\chi^2 = \sum_{i=1}^{5} \frac{(O_i - E_i)^2}{E_i}$$

เมื่อค่า E_i : i =1,...,5 คำนวณได้จาก NP_i

เมื่อ P_i = ความน่าจะเป็นที่จะเกิดกลุ่มที่ i ตาม H_0 จะมีค่า $\frac{1}{5}$

ดังนั้น $E_i = 1000 \times \frac{1}{5} = 200$

ดังนั้น $\chi^2_{cal} = \frac{(200-200)^2}{200} + \frac{(250-200)^2}{200} + \frac{(150-200)^2}{200} + \frac{(50-200)^2}{200} + \frac{(350-200)^2}{200}$

$= 250$

ที่ระดับนัยสำคัญ = 0.10 อาณาเขตวิกฤต $\chi^2 > \chi^2_{0.90,4} = 7.78$

ค่า χ^2_{cal} ตกในอาณาเขตวิกฤต ดังนั้น ปฏิเสธ H_0 นั่นคือ สามารถสรุปที่ระดับนัยสำคัญ 10% ได้ว่า กลุ่มต่างๆ ของนักศึกษามหาวิทยาลัยแห่งนี้ซึ่งจำแนกตามความนิยมการอ่านหนังสือพิมพ์ข้างต้นมีจำนวนแตกต่างกัน

ตัวอย่างที่ 2

นักสถิติชาวสวีเดนท่านหนึ่งต้องการศึกษาเกี่ยวกับการเกิดทารกในประเทศสวีเดน จึงสุ่มตัวอย่างทารกแรกเกิดในรอบ 1 ปี มาจำนวน 88 ราย โดยข้อมูลที่ได้จากการสุ่มตัวอย่างนี้สามารถแยกตามฤดูต่างๆซึ่งมีช่วงเวลายาวนานต่างกัน ดังนี้

ฤดู	ใบไม้ผลิ (เม.ย.-มิ.ย.)	ร้อน (ก.ค.-ส.ค.)	ใบไม่ร่วง (ก.ย.-ต.ค.)	หนาว (พ.ย.-มี.ค.)
จำนวนทารกแรกเกิด	27	20	8	33

จากข้อมูลชุดนี้นักสถิติท่านนี้สามารถสรุปผลได้หรือไม่ว่า สัดส่วนที่จะมีทารกแรกเกิดในแต่ละฤดูเป็นไปตามสัดส่วนของจำนวนวันในแต่ละฤดูนั้นๆเมื่อใช้ระดับนัยสำคัญ 0.10 (ถ้ากำหนดให้ 1 ปี มี 365 วัน) (อุมาพร จันทศร, 2542)

<u>วิธีทำ</u>

สามารถตั้งสมมติฐานได้ คือ

H_0: สัดส่วนที่จะมีทารกแรกเกิดในแต่ละฤดูเป็นไปตามสัดส่วนของจำนวนวันในแต่ละฤดูนั้น

หรือ $H_0: P_1 : P_2 : P_3 : P_4 = 91 : 62 : 61 : 151$

$H_1: P_1 : P_2 : P_3 : P_4 \neq 91 : 62 : 61 : 151$

เมื่อ P_i : 1,...,4 คือสัดส่วนจำนวนวันในแต่ละฤดู

เช่น P_i = สัดส่วนจำนวนวันในฤดูใบไม้ผลิซึ่งมีในช่วงเวลาเดือนเมษายน-มิถุนายน มีจำนวนวัน 91 วัน = $\frac{91}{365}$

และ $E_i = NP_i$ เช่น $E_1 = 88\times\frac{91}{365} = 21.94 \cong 22$

$$E_2 = 88\times\frac{62}{365} = 14.94 \cong 15$$

$$E_3 = 88\times\frac{61}{365} = 14.70 \cong 15$$

$$E_4 = 88\times\frac{151}{365} = 36.40 \cong 36$$

ดังนั้นค่า $\chi^2_{cal} = \frac{(27-22)^2}{22} + \frac{(20-15)^2}{15} + \frac{(8-15)^2}{15} + \frac{(33-36)^2}{36}$

$= 6.306$

ที่ระดับนัยสำคัญ = 0.10 อาณาเขตวิกฤต $\chi^2 > \chi^2_{0.90,3} = 6.25$

ค่า χ^2_{cal} ตกในอาณาเขตวิกฤต ดังนั้นปฏิเสธสมมติฐานที่ว่าสัดส่วนที่จะมีทารกแรกเกิดในแต่ละฤดูเป็นไปตามสัดส่วนของจำนวนวันในแต่ละฤดูนั้น

ตัวอย่างที่ 3

ข้อมูลที่ได้นี้ ได้จากการสุ่มตัวอย่างนักศึกษานานาชาติจากมหาวิทยาลัยแห่งหนึ่ง จำนวน 726 คน เพื่อศึกษาความสัมพันธ์ระหว่างเชื้อชาติและชนิดของกรุ๊ปเลือดของตนเอง จัดข้อมูลนี้ลงในตารางการณ์จรได้ ดังนี้

เชื้อชาติ	ชนิดของกรุ๊ปเลือด				รวม
	O	A	B	AB	
1	176	148	96	72	492
2	78	50	45	12	185
3	15	19	8	7	49
	269	217	149	91	726

จงทดสอบว่าเชื้อชาติและชนิดของกรุ๊ปเลือดมีความสัมพันธ์กันหรือไม่ ที่ระดับนัยสำคัญ 0.01
(อุมาพร จันทศร, 2542)

<u>วิธีทำ</u>

H_0: ชนิดของกรุ๊ปเลือดและเชื้อชาติเป็นอิสระกัน

H_1: ชนิดของกรุ๊ปเลือดและเชื้อชาติมีความสัมพันธ์กัน

ใช้การทดสอบ χ^2 ด้วยสถิติทดสอบ $\chi^2 = \sum_{i=1}^{3}\sum_{j=1}^{4}\frac{(O_{ij}-E_{ij})^2}{E_{ij}}$

โดยคำนวณ E_{ij} จากสูตร $E_{ij} = \frac{R_i C_j}{N}$ ได้ผลดังนี้

ความถี่คาดหวัง

เชื้อชาติ	ชนิดของกรุ๊ปเลือด			
	O	A	B	AB
1	182	147	101	62
2	69	55	38	23
3	18	15	10	6

และสามารถคำนวณค่า χ^2_{cal} = 12.374

ที่ระดับนัยสำคัญ 0.01 , df.= (3-1)(4-1) =6

ได้อาณาเขตวิกฤต คือ $\chi^2_{cal} > 16.812$

ดังนั้น χ^2_{cal} ไม่ตกในอาณาเขตวิกฤต จึงไม่สามารถปฏิเสธ H_0 ได้ สามารถสรุปได้ว่า เชื้อชาติและชนิดของกรุ๊ปเลือดเป็นอิสระต่อกัน