4.3 ความน่าจะเป็นและการทดสอบไค-สแควร์

4.3.1 หลักการใช้ความน่าจะเป็น ความน่าจะเป็น หมายถึง ความถี่ที่เหตุการณ์หนึ่งจะเกิดขึ้น (ปราโมช, 2529; สมชัย, 2530) เช่น เราต้องการหยิบให้ได้ไพ่หน้าโพธิ์ดำจากกอง ความน่าจะเป็นที่เราจะหยิบไพ่หน้าโพธิ์ดำเท่ากับ $\frac{1}{4}$

ถ้าเหตุการณ์หนึ่งอาจเกิดขึ้นได้เป็นจำนวน N วิธี โดยที่เมื่อเกิดอย่างหนึ่งแล้ววิธีอื่นก็เกิดขึ้นไม่ได้ ถ้าหาก n วิธี ของเหตุการณ์มีลักษณะ A เขียนเป็นสูตร ได้ดังนี้

$$\text{ความน่าจะเป็น } A = \frac{\text{จำนวนครั้งที่เกิดเหตุการณ์}A}{\text{จำนวนทั้งหมดของเหตุการณ์ที่เกิดขึ้นได้}}$$

$$\text{หรือ} \quad P(A) = \frac{n}{N}$$

หลักเบื้องต้น 2 ประการเกี่ยวกับความน่าจะเป็น มีดังนี้

1) ความน่าจะเป็นมีค่าระหว่าง 0 กับ 1

2) ผลรวมของความน่าจะเป็นของเหตุการณ์ทั้งหมดที่อาจเกิดขึ้นได้มีค่าเป็น 1

4.3.2 กฎที่เกี่ยวข้องกับความน่าจะเป็น จำแนกได้ดังนี้ (จรัญ, 2516; จรัส, 2553)

1) กฎการรวม โอกาสที่จะเกิดขึ้นของเหตุการณ์ย่อย (single event) หนึ่งเหตุการณ์ย่อยใด ๆ ซึ่งอยู่ในชุดของเหตุการณ์ (mutually exclusive event) เท่ากับผลรวมของความน่าจะเป็นของเหตุการณ์ย่อยนั้น ๆ

ตัวอย่าง การโยนลูกเต๋า 1 ครั้ง ความน่าจะเป็นของการออกหน้า 1 ถึง 4 มีอยู่เท่าไร

วิธีทำ

$$P(\text{หน้า } 1 - \text{หน้า } 4) = P(\text{หน้า } 1) \text{ หรือ } P(\text{หน้า } 2) \text{ หรือ } P(\text{หน้า } 3) \text{ หรือ } P(\text{หน้า } 4)$$

$$= \frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}$$

$$= \frac{4}{6}$$

$= \frac{2}{3}$

ความน่าจะเป็นของการออกหน้า 1 ถึง 4 มีอยู่ $\frac{2}{3}$

2) กฎการคูณ โอกาสที่จะเกิดขึ้นของหลาย ๆ เหตุการณ์ที่เป็นอิสระกัน (independent event) ที่เกิดขึ้นด้วยกัน จะเท่ากับผลคูณของโอกาสที่จะเกิดขึ้นของเหตุการณ์ที่เป็นอิสระนั้น ๆ

ตัวอย่าง การโยนเหรียญ 2 ครั้ง อยากทราบว่าโอกาสของการเกิดหัวทั้ง 2 ครั้ง มีค่าเท่าไร

วิธีทำ

P(หัวครั้งแรก) $= \frac{1}{2}$

P(หัวครั้งที่สอง) $= \frac{1}{2}$

เนื่องจากการโยนเหรียญ 2 ครั้ง โอกาสที่จะเกิดหัวในครั้งแรกและครั้งที่ 2 นั้น เป็นอิสระแก่กัน

P(หัวทั้ง 2 ครั้ง) $= \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

ความน่าจะเป็นของการเกิดหัวทั้ง 2 ครั้ง มีค่าเท่ากับ $\frac{1}{4}$

4.3.3 การใช้ความน่าจะเป็นในการปรับปรุงพันธุ์สัตว์ กฎของเมนเดลในด้านความน่าจะเป็น มีดังนี้ (จรัญ, 2516; จรัส, 2553)

1) กฎการแยกตัวของยีน ในเชิงพันธุศาสตร์ยีนซึ่งอยู่เป็นคู่ในเซลล์สืบพันธุ์ (germ cell) จะแยกออกจากกันเมื่อกลายเป็นอสุจิในเพศผู้หรือไข่ในเพศเมีย ในเชิงความน่าจะเป็นโอกาสของอสุจิหรือไข่จะได้รับยีนใดยีนหนึ่งเท่ากับ $\frac{1}{2}$ จากยีนซึ่งอยู่ที่โลกัสนั้น ๆ หรือที่อยู่คู่กัน

ตัวอย่าง โอกาสของอสุจิหรือไข่จะได้รับยีนใดยีนหนึ่งเท่ากับ $\frac{1}{2}$ จากยีนซึ่งอยู่ที่โลกัสนั้น ๆ หรือที่อยู่คู่กัน

A	= ยีนไม่มีเขา	a	= ยีนมีเขา
B	= ยีนสีดำ	b	= ยีนสีแดง

ยีนคู่ที่ 1 A และ a มีจีโนไทป์ Aa

P(อสุจิจะมี A) $= \frac{1}{2}$

P(อสุจิจะมี a) $= \frac{1}{2}$

รวม P $= \frac{1}{2} + \frac{1}{2} = 1$

ยีนคู่ที่ 2 B และ b มีจีโนไทป์ Bb

P(อสุจิจะมี B) $= \frac{1}{2}$

P(อสุจิจะมี b) $= \frac{1}{2}$

รวม P $= \frac{1}{2} + \frac{1}{2} = 1$

2) กฎของการเข้าชุดอย่างอิสระของยีน ในเซลล์สืบพันธุ์ การแยกตัวของยีนซึ่งไม่เกี่ยวข้องกัน เป็นอิสระแก่กันในเชิงความน่าจะเป็น (จากตัวอย่างข้างบน) ความน่าจะเป็นของอสุจิที่ได้รับยีน A หรือยีน a เป็นอิสระต่อความน่าจะเป็นของอสุจิที่ได้รับยีน B หรือ b

ตัวอย่าง การรวมตัวของยีน 2 คู่ ได้แก่ P(อสุจิมียีน 2 คู่) เช่น

P(AB) $= \frac{1}{2} x \frac{1}{2} = \frac{1}{4}$

P(Ab) $= \frac{1}{2} x \frac{1}{2} = \frac{1}{4}$

P(aB) $= \frac{1}{2} x \frac{1}{2} = \frac{1}{4}$

P(ab) $= \frac{1}{2} x \frac{1}{2} = \frac{1}{4}$

4.3.4 การทดสอบนัยสำคัญทางสถิติ

1) สมมุติฐาน (hypothesis) หมายถึง ข้อความที่แสดงการคาดคะเนคำตอบไว้ล่วงหน้า เป็นข้อความที่สมมุติขึ้นอย่างสมเหตุสมผลจากแนวคิดของทฤษฎี สมมุติฐานที่ดีจะต้องเป็นข้อความที่แสดงถึงความสัมพันธ์ระหว่างตัวแปรและทดสอบได้ด้วยข้อเท็จจริง สมมุติฐานมี 2 ประเภท ดังนี้

(1) สมมุติฐานทางการวิจัย (research hypothesis) เป็นสมมุติฐานที่เขียนในรูปของข้อความบรรยายหรือคาดคะเนความสัมพันธ์ของตัวแปรที่ศึกษาซึ่งผู้วิจัยคาดว่าจะเกิดขึ้น โดยใช้ภาษาเป็นสื่อในการ

อธิบายให้ผู้อื่นเข้าใจได้ตรงกัน (นิภา, 2527) เช่น การเลือกอาชีพของนักเรียนมีความ สัมพันธ์กับค่านิยมในอาชีพ นักเรียนชายมีผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ดีกว่านักเรียนหญิง เป็นต้น

(2) สมมุติฐานทางสถิติ (statistical hypothesis) เป็นสมมุติฐานที่เขียนขึ้นในรูปโครงสร้างทางคณิตศาสตร์เพื่ออธิบายความสัมพันธ์ของตัวแปรประกอบด้วย สมมุติฐานเป็นกลางหรือสมมุติฐานหลัก (null hypothesis) นิยมใช้สัญลักษณ์ H_0 (อ่านว่า เอชศูนย์) กำหนดขึ้นในลักษณะที่มีวัตถุประสงค์ที่จะลบล้างหรือไม่ยอมรับ และสมมุติฐานไม่เป็นกลางหรือสมมุติฐานรองหรือสมมุติฐานทางเลือก (alternative hypothesis) ใช้สัญลักษณ์ H_a หรือ H_1 (อ่านว่า เอชเอ หรือ เอชหนึ่ง) มีลักษณะตรงกันข้ามกับสมมุติฐานหลัก (นิภา, 2527; จรัส, 2553)

2) ความคลาดเคลื่อนในการตัดสินใจ การตัดสินใจยอมรับหรือปฏิเสธสมมุติฐานหลัก อาจเกิดความคลาดเคลื่อนในการตัดสินใจได้ ความคลาดเคลื่อนในการตัดสินใจแบ่งได้เป็น 2 ประเภท ดังต่อไปนี้ (จรัส, 2553)

(1) ความคลาดเคลื่อนแบบที่ 1 (Type I error) เป็นความคลาดเคลื่อนที่เกิดขึ้นจากการปฏิเสธสมมุติฐานเป็นกลาง (H_0) เมื่อสมมุติฐานเป็นกลางนั้นเป็นจริง ความคลาดเคลื่อนแบบที่ 1 มีโอกาสที่จะเกิดขึ้นเท่ากับ α (อ่านว่า แอลฟา) และเรียกความผิดพลาดชนิดนี้ว่า “ระดับนัยสำคัญ” (Level of significant)

(2) ความคลาดเคลื่อนแบบที่ 2 (Type II error) เป็นความคลาดเคลื่อนที่เกิดขึ้นจากการยอมรับสมมุติฐานเป็นกลาง (H_0) เมื่อสมมุติฐานเป็นกลางนั้นเป็นเท็จ ความคลาดเคลื่อนแบบที่ 2 มีโอกาสที่จะเกิดขึ้นเท่ากับ β (อ่านว่า บีตา)

ความคลาดเคลื่อนทั้ง 2 แบบ สามารถสรุปเป็นตารางได้ดังนี้

การตัดสินใจ	สมมุติฐานเป็นกลาง (H_0)	
	H_0 เป็นจริง	H_0 เป็นเท็จ
ปฏิเสธ H_0	ความคลาดเคลื่อนแบบที่ 1 เขียนแทนด้วย α	ตัดสินใจถูก
ยอมรับ H_0	ตัดสินใจถูก	ความคลาดเคลื่อนแบบที่ 2 เขียนแทนด้วย β

3) ระดับนัยสำคัญทางสถิติ หมายถึง ความน่าจะเป็นที่อนุญาตให้ปฏิเสธสมมุติฐานเป็นกลางที่เป็นจริง ซึ่งก็คือโอกาสของการเกิดความคลาดเคลื่อนแบบที่ 1 ของการทดสอบสมมุติฐานทางสถิตินั้นเอง โดยทั่วไปนิยมใช้ α แทนค่าระดับนัยสำคัญทางสถิติ ค่า α นิยมกำหนดให้เป็น .05 หรือ .01 โดยปกติผู้วิจัยจะต้องกำหนดค่า α ก่อนที่จะทำการทดสอบสมมุติฐาน

4) ระดับความเชื่อมั่น (Level of confidence) หมายถึง ความน่าจะเป็นในการยอมรับสมมุติฐานเป็นกลางที่เป็นจริง ซึ่งจะมีค่าเป็น $(1 - \alpha)$ โดยทั่วไปนิยมคำนวณออกมาเป็นร้อยละ เช่น ถ้านักวิจัยกำหนดค่า α เท่ากับ .05 ระดับความเชื่อมั่นเท่ากับ 0.95 หรือ 95 เปอร์เซ็นต์

4.3.5 ค่าไค-สแควร์ (chi-square; χ^2) χ อ่านว่า ไค (chi) เป็นค่าใช้ทดสอบเหตุการณ์ที่เกิดขึ้นว่าเป็นไปตามสมมุติฐาน หรือการคาดคะเนหรือไม่ ในทางพันธุศาสตร์ใช้ทดสอบเกี่ยวกับอัตราส่วนทางพันธุกรรม (genetic ratio) ซึ่งเป็นผลมาจากการผสมพันธุ์ตามหลักของเมนเดล มีความจำเป็นต้องรู้ด้วยความมั่นใจว่าผลที่แสดงออกมาให้เห็นนั้น ถูกต้องอย่างที่อธิบายตามหลักพันธุศาสตร์หรือไม่ อัตราส่วนที่เกิดจากการทดลองและสัดส่วนตามหลักทฤษฎีย่อมจะได้รับผลตรงกันเสมอ เพื่อให้รู้ว่าผลที่ได้รับนั้นแตกต่างไปจากทฤษฎีมากหรือน้อย จึงต้องใช้ไค-สแควร์ทดสอบ (บุญชอบ, 2535) โดยสามารถคำนวณได้จากสูตรดังนี้

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

χ^2 = ค่าไค-สแควร์

$\sum_{i=1}^{n}$ = ผลรวมของจำนวน

O_i = ค่าที่ได้จากการสังเกตหรือตัวเลขที่เก็บได้จากการทดลอง

E_i = เป็นค่าที่คาดว่าจะได้หรือน่าจะเป็นไปได้ในทางทฤษฎี

วิธีทดสอบค่าไค-สแควร์ ทดสอบได้โดยเปิดตารางค่าไค-สแควร์ เพื่อนำมาเปรียบเทียบค่าไค-สแควร์ที่ได้จากการคำนวณ สามารถพิจารณาตามลำดับดังนี้

1) หาค่าความเป็นอิสระ (degree of freedom; df) โดย df = n-1 เมื่อ n เป็นจำนวนชนิดหรือชั้นที่ใช้ในการทดลอง

2) เปิดตารางไค-สแควร์ ที่ค่าความน่าจะเป็น (probability value) ระดับ 5 เปอร์เซ็นต์ และ 1 เปอร์เซ็นต์

3) นำค่าไค-สแควร์ ที่ได้จากการคำนวณ มาเปรียบเทียบกับค่าไค-สแควร์ในตารางที่ระดับ 5 เปอร์เซ็นต์ และ 1 เปอร์เซ็นต์ โดยมีหลักเกณฑ์ว่า ถ้า χ^2 ที่ได้จากการคำนวณมีค่าน้อยกว่าค่า χ^2 ในตารางที่ระดับ 0.05 และมีค่าน้อยกว่า χ^2 ที่ระดับ 0.01 ($\chi^2 < \chi^2$ 0.05, $\chi^2 < \chi^2$ 0.01) ถือว่าไม่มีความแตกต่างกันอย่างมีนัยสำคัญ (un - significant) นั้นก็คือเป็นไปตามทฤษฎีหรือยอมรับผลการทดลองครั้งนี้ ถ้าค่า χ^2 ที่ได้จากการคำนวณมากกว่า χ^2 ที่ระดับ 0.05 และ χ^2 ที่ระดับ 0.01 ($\chi^2 > \chi^2$ 0.05, $\chi^2 > \chi^2$ 0.01) ถือว่ามีความแตกต่างกันอย่างมีนัยสำคัญ (high significant) จึงไม่ยอมรับว่าเป็นไปตามทฤษฎีหรือไม่ยอมรับผลการทดลองครั้งนี้

4.3.6 การทดสอบไค-สแควร์

1) การใช้ไค-สแควร์และสมมุติฐานทดสอบสัดส่วนพันธุศาสตร์ (ชาญชัย, 2532)

ตัวอย่าง การผสมพันธุ์โคชอร์ตฮอร์นสีแดง (RR) กับสีโรน (Rr) ถ้าจำนวนที่นับได้จากโคฝูงนี้ประกอบด้วย RR เท่ากับ 480 Rr เท่ากับ 420 จากจำนวนโคทั้งหมด 900 ตัว จงทดสอบสมมุติฐาน (H_0) ว่าลูกโคมาจากการผสมระหว่าง RR กับ Rr หรือไม่

$$\begin{array}{ccc} RR & X & Rr \\ & \downarrow & \\ & RR : Rr = 1 : 1 & \end{array}$$

class	O	E	O - E	$(O - E)^2$	$\frac{(O - E)^2}{E}$
RR	480	$900 \times \frac{1}{2} = 450$	30	900	2
Rr	420	$900 \times \frac{1}{2} = 450$	-30	900	2

วิธีทำ

H_0 : ลูกโคมาจากการผสมระหว่าง RR กับ Rr

H_0 : O = E

H_a : $O \neq E$

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

$= 2 + 2$

$= 4$

df $= n - 1$

$= 2 - 1 \quad = 1$

เปิดตารางไค-สแควร์

ที่ χ^2 (df = 1 probability 5% หรือ 0.05) = 3.84

ที่ χ^2 (df = 1 probability 1% หรือ 0.01) = 6.63

ค่าไค-สแควร์จากการคำนวณมากกว่าตารางที่ระดับ 0.05 แต่น้อยกว่าตารางที่ระดับ 0.01 มีความแตกต่างกันอย่างมีนัยสำคัญ จึงไม่ยอมรับสมมุติฐานเป็นกลาง นั่นคือลูกโคไม่ได้มาจากการผสมระหว่าง RR กับ Rr

2) การใช้ไค-สแควร์และสมมุติฐานทดสอบความเป็นอิสระ (จรัญ, 2516)

ตัวอย่าง การศึกษาเพื่อทราบว่าการสูบกล้องบุหรี่ทำให้อายุสั้นหรือไม่ มีข้อมูลดังต่อไปนี้

	คนไม่สูบกล้อง	คนสูบกล้อง	รวม
ตาย	117 (E_1)	54 (E_2)	171
ยังมีชีวิต	950 (E_3)	348 (E_4)	1,298
รวม	1,067	402	1,469

วิธีทำ

μ_1 แทนสัดส่วนของคนไม่สูบกล้องตาย (อ่านว่า มิวหนึ่ง)

μ_2 แทนสัดส่วนของคนสูบกล้องตาย (อ่านว่า มิวสอง)

H_0 : เปอร์เซ็นต์ตายในกลุ่มคนสูบกล้องเท่ากับกลุ่มคนไม่สูบกล้อง

H_0 : $\mu_1 = \mu_2$

H_a : H_0 ไม่จริง

$$E = \frac{(\text{RowTotal})(\text{ColumnTotal})}{\text{GrandTotal}}$$

$$E_1 = \frac{(171)(1{,}067)}{1{,}469}$$

$$= 124.2$$

$$E_2 = \frac{(171)(402)}{1{,}469}$$

$$= 46.8$$

$$E_3 = \frac{(1{,}298)(1{,}067)}{1{,}469}$$

$$= 942.8$$

$$E_4 = \frac{(1{,}298)(402)}{1{,}469}$$

$$= 355.2$$

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

$$= \frac{(117 - 124.2)^2}{124.2} + \frac{(54 - 46.8)^2}{46.8} + \frac{(950 - 942.8)^2}{942.8} + \frac{(348 - 355.2)^2}{355.2}$$

$$= \frac{51.84}{124.2} + \frac{51.84}{46.8} + \frac{51.84}{942.8} + \frac{51.84}{355.2}$$

$$= 0.42 + 1.11 + 0.05 + 0.15$$

$$= 1.73$$

$$df = (\text{Row} - 1)(\text{Column} - 1)$$

$$= (1)(1) \quad = 1$$

เปิดตารางไค-สแควร์

ที่ χ^2 (df = 1 probability 5% หรือ 0.05) = 3.84

ที่ χ^2 (df = 1 probability 1% หรือ 0.01) = 6.63

ค่าไค-สแควร์ที่คำนวณได้น้อยกว่าตารางที่ระดับ 0.05 และ 0.01 ไม่มีความแตกต่างกันอย่างมีนัยสำคัญ จึงยอมรับสมมุติฐานเป็นกลาง นั่นคือ เปอร์เซ็นต์ตายในกลุ่มคนสูบกล้องไม่เท่ากับกลุ่มคนไม่สูบกล้อง