

ความเป็นมาและความสำคัญของปัญหา

การแจกแจงปกติ เป็นการแจกแจงตามทฤษฎี (Theoretical Distribution) ที่สำคัญที่สุดทั้งในสถิติวิธีสุ่มและสถิติประยุกต์ และนำมาใช้ประโยชน์มากที่สุดในการประมาณค่าและการทดสอบสมมุติฐานทางสถิติ<sup>1</sup> อย่างไรก็ตาม คอร์เนล (Francis G. Cornell) กล่าวว่า เป็นไปได้ที่การใช้การแจกแจงปกติจะใช้กันน้อยที่สุด<sup>2</sup> เนื่องจากในทางทฤษฎีการแจกแจงปกติเป็นการแจกแจงแบบต่อเนื่อง (Continuous Distribution) อย่างหนึ่ง แต่ในทางปฏิบัติ การใช้การแจกแจงปกติไม่ได้จำกัดอยู่เฉพาะกับตัวแปรที่แจกแจงปกติหรือเข้าใกล้ปกติเท่านั้น<sup>3</sup> เราประยุกต์ใช้อย่างกว้างขวางทั้งกับตัวแปรแบบต่อเนื่อง (Continuous Variables) ตัวแปรแบบไม่ต่อเนื่อง (Discrete Variables) ตลอดจนตัวแปรที่ไม่แจกแจงปกติ หรือตัวแปรที่ไม่ทราบลักษณะการแจกแจงด้วย ถ้าหากความเข้าใจในเรื่องการแจกแจงปกติมีไม่มากพอ และนำมาใช้โดยไม่ได้คำนึงถึงความเหมาะสม ปัญหาและข้อจำกัดในการใช้อย่างรอบคอบแล้ว ย่อมมีโอกาที่จะใช้ผิดได้ง่าย ผลสรุปที่ได้ก็ย่อมไม่น่าเชื่อถือ ด้วยเหตุดังกล่าว ผู้วิจัยจึงสนใจใคร่ที่จะศึกษาปัญหาเรื่องการแจกแจงปกติ โดยเฉพาะอย่างยิ่ง การใช้การแจกแจงปกติในงานวิจัยทางการศึกษาและจิตวิทยา เพื่อจะได้ทราบว่า ในแต่ละกรณี มีความเหมาะสม ปัญหา และข้อจำกัดในการใช้อย่างไรบ้าง ข้อค้นพบที่ได้จะเป็นประโยชน์-

<sup>1</sup>William L. Hays, Statistics for Psychologists (New York: Holt, Rinehart and Winston, Inc., 1965), p. 225.

<sup>2</sup>Francis G. Cornell, The Essentials of Educational Statistics (New York: John Wiley & Sons, Inc., 1956), p. 82.

<sup>3</sup>Owen L. Davies (ed.), Statistical Methods in Research and Production (New York: Hafner Publishing Company, 1961), p. 509.

อย่างมากสำหรับผู้ที่จะนำการแจกแจงปกติไปใช้ในงานวิจัย และจะเป็นแนวทางสำหรับผู้สนใจจะศึกษาค้นคว้าในขั้นต่อไป

### ความมุ่งหมายของการวิจัย

การวิจัยครั้งนี้มีความมุ่งหมายเพื่อตอบปัญหาต่อไปนี้

1. การแจกแจงปกติคืออะไร มีลักษณะและคุณสมบัติทางคณิตศาสตร์และทางสถิติที่สำคัญ ๆ เป็นอย่างไร
2. การแจกแจงปกติ นำไปใช้ประโยชน์ในงานวิจัยได้อย่างไรบ้าง ในแต่ละกรณีมีความเหมาะสม ปัญหา และข้อจำกัดในการใช้อย่างไร
3. การแจกแจงปกติมีความสัมพันธ์กับการแจกแจงที่สำคัญอื่น ๆ ที่ได้มาจากการแจกแจงปกติ (Derived Distribution) อย่างไรบ้าง

### ขอบเขตของการวิจัย

1. การวิจัยครั้งนี้ต้องการความรู้พื้นฐานทางวิชาคณิตศาสตร์แคลคูลัสและวิชาสถิติเบื้องต้นเป็นส่วนใหญ่
2. การวิจัยครั้งนี้ไม่มีการทดลองปฏิบัติเพื่อยืนยันโค้งของการแจกแจงปกติ
3. การวิจัยครั้งนี้จะกล่าวถึงความรู้พื้นฐานและทฤษฎีต่าง ๆ ทางคณิตศาสตร์สถิติเฉพาะที่จำเป็นบางประการ
4. การแจกแจงความถี่มีหลายชนิด แต่จะกล่าวถึงการแจกแจงความถี่ที่มีความสัมพันธ์กับการแจกแจงปกติ เฉพาะที่สำคัญบางอย่าง
5. การแจกแจงปกติสามารถนำไปใช้ประโยชน์ได้หลายด้าน แต่จะเน้นเฉพาะการนำไปใช้ในงานวิจัยทางการศึกษาและจิตวิทยา เป็นส่วนใหญ่

### ข้อตกลงเบื้องต้น

ผู้วิจัยยอมรับว่า การแจกแจงปกติเป็นการแจกแจงตามทฤษฎีที่สามารถนำไปใช้ประ

ประโยชน์ได้อย่างกว้างขวาง และเป็นที่ยอมรับกันเ็นทางสถิติ

**คำศัพท์และคำจำกัดความที่ใช้ในการเขียนวิทยานิพนธ์ฉบับนี้**

ตัวแปร (Variable) หมายถึง สิ่งที่มีความผันแปร อาจจะทางค่านปริมาตร เช่น ความสูง น้ำหนัก ความเร็ว เป็นต้น หรืออาจจะทางค่านคุณภาพ เช่น สีผม สีขนตา เพศ เชื้อชาติ เป็นต้น ในบางกรณีตัวแปรค่านคุณภาพอาจถือเป็นตัวแปรค่านปริมาตร เช่น สีขนตา ถ้าวัดค่าค่านจากสีอ่อนถึงสีเข้ม ก็ถือว่าเป็นตัวแปรค่านปริมาตร

ตัวแปรแบบต่อเนื่อง (Continuous Variable) หมายถึง ตัวแปรที่โดยทฤษฎีแล้ว เราสามารถสมมุติให้มีค่าใดก็ได้ในระหว่างค่านสองค่านที่กำหนดให้ เช่น ความสูง น้ำหนัก ความเร็ว คะแนนสอบ

ตัวแปรแบบไม่ต่อเนื่อง (Discrete Variable) หมายถึง ตัวแปรที่เราไม่สามารถสมมุติให้มีค่านย่อยระหว่างค่านสองค่านที่กำหนดได้ทั้งหมด เช่น จำนวนคนในครอบครัว จำนวนนักเรียนในโรงเรียน จำนวนโรงเรียน จำนวนบ้าน

ประชากร (Population) หมายถึง ยอดรวมของหน่วยย่อย (Individual item) ของตัวแปรหนึ่ง

กลุ่มตัวอย่าง (Sample) หมายถึง ส่วนหนึ่งของประชากรซึ่งเราเลือกขึ้นมาเป็นตัวแทนของประชากรนั้น

พารามิเตอร์ (Parameter) หมายถึง ค่านแท้จริง (True Value) ที่ทำได้โดยวิธีการทางสถิติเพื่ออธิบายลักษณะของประชากร

ค่าสถิติ (Statistic) หมายถึง ค่านที่คำนวณได้จากกลุ่มตัวอย่างซึ่งแสดงลักษณะของกลุ่มตัวอย่งนั้น

ความถี่ (Frequency) หมายถึง จำนวนรายการข้อมูลในกลุ่มตัวอย่างหรือในประชากรหนึ่ง ๆ

การแจกแจง (Distribution) หมายถึง ลักษณะที่ตัวแปรหนึ่งจะมีค่าต่าง ๆ ในขอบเขตของค่าที่เป็นไปได้

ความน่าจะเป็น (Probability) หมายถึง ความถี่ที่เหตุการณ์อย่างหนึ่งอย่างใดจะเกิดขึ้นหรือไม่เกิดขึ้น

การแจกแจงตัวอย่าง (Sampling Distribution) หมายถึง การแจกแจงความน่าจะเป็นของค่าสถิติที่ได้จากกลุ่มตัวอย่างสุ่มขนาดเดียวกันหลาย ๆ กลุ่ม

สมมุติฐานทางสถิติ (Statistical Hypothesis) หมายถึง ข้อความที่สมมุติขึ้นมาเกี่ยวกับพารามิเตอร์ และ(หรือ)ลักษณะของประชากร

สมมุติฐานศูนย์ (Null Hypothesis) หมายถึง สมมุติฐานที่จะทดสอบซึ่งตั้งขึ้นโดยถือว่า สมมุติฐานนั้นจริง

สมมุติฐานสำรอง (Alternative Hypothesis) หมายถึง สมมุติฐานที่ตรงกันข้ามกับสมมุติฐานศูนย์ ซึ่งตั้งขึ้นมาเพื่อเป็นทางเลือก ถ้าปฏิเสธสมมุติฐานศูนย์จะยอมรับสมมุติฐานสำรอง

ระดับความมีนัยสำคัญ (Level of significance) หมายถึง ความน่าจะเป็นในการปฏิเสธสมมุติฐานที่ถูก

ขอบเขตวิกฤต (Critical Region) หมายถึง ขอบเขตของการปฏิเสธสมมุติฐานศูนย์ เป็นขอบเขตทางซ้ายมือหรือขวามือของโค้งของการแจกแจงซึ่งกำหนดตามระดับความมีนัยสำคัญ ถ้าค่าสถิติที่คำนวณได้ตกอยู่ในขอบเขตนี้ จะถือว่าการทดสอบนั้นมีนัยสำคัญ นั่นคือจะปฏิเสธสมมุติฐานศูนย์ และยอมรับสมมุติฐานสำรอง

สัญลักษณ์ที่ใช้ในการเขียนนิทานิพนธ์ฉบับนี้

$$\sum \text{ แทน ผลบวกของ ; } \quad \sum_{i=1}^n X_i = X_1 + X_2 + \dots + X_n$$

- $\prod$  แทน ผลคูณของ ;  $\prod_{i=1}^n X_i = (x_1)(x_2)\dots(x_n)$   
 $\int$  " การอินทิเกรต  
 $\Gamma$  " แกมมาฟังก์ชัน  
 $\theta$  " พารามิเตอร์  
 $\hat{\theta}$  " ค่าประมาณของพารามิเตอร์  
 $Z$  " ตัวแปรปกติมาตรฐาน  
 $N$  " ขนาดของประชากร  
 $n$  " ขนาดของกลุ่มตัวอย่าง  
 $v, d.f.$  " ชั้นของความเป็นอิสระ  
 $\bar{M}$  " มัชฌิม (Mean) ของประชากร  
 $\bar{X}$  " มัชฌิมของกลุ่มตัวอย่าง  
 $\pi$  " สัดส่วน (Proportion) ของประชากร  
 $p$  " สัดส่วนของกลุ่มตัวอย่าง และความน่าจะเป็นของการเกิด "สำเร็จ" ของการแจกแจงทวินาม  
 $P$  " ความน่าจะเป็น  
 $\sigma^2$  " ความแปรปรวน (Variance) ของประชากร  
 $s^2$  " ความแปรปรวนของกลุ่มตัวอย่าง  
 $\sigma$  " ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) ของประชากร  
 $s$  " ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง  
 $r_{XY}$  " สัมประสิทธิ์สหสัมพันธ์แบบผลคูณ (Product-moment Correlation Coefficient) ระหว่างตัวแปร  $X$  กับตัวแปร  $Y$  ในประชากรที่มีการแจกแจงปกติแบบสองตัวแปร  
 $r_{XY}$  " สัมประสิทธิ์สหสัมพันธ์แบบผลคูณระหว่างตัวแปร  $X$  กับตัวแปร  $Y$  ในกลุ่มตัวอย่างซึ่งสุ่มมาจากประชากรที่มีการแจกแจงปกติแบบสองตัวแปร  
 $r_r$  " สัมประสิทธิ์สหสัมพันธ์แบบลำดับที่ของสเปียร์แมน (Spearman's Rank Correlation Coefficient)  
 $\tau$  " สัมประสิทธิ์สหสัมพันธ์แบบลำดับที่ของเคนคอลล (Kendall's Tau Rank Correlation Coefficient)

$\phi$	แทน	สัมประสิทธิ์สหสัมพันธ์ไช (Phi Correlation Coefficient)
$r_{tet}$	"	สัมประสิทธิ์สหสัมพันธ์เตตระคอริก (Tetrachoric Correlation Coefficient)
$r_{bis}$	"	สัมประสิทธิ์สหสัมพันธ์ไบซีเรียล (Biserial Correlation Coefficient)
$\alpha$	"	ระดับความมีนัยสำคัญ
$H_0$	"	สมมุติฐานศูนย์
$H_1$	"	สมมุติฐานสำรอง