

บทที่ 1

บทนำ



ความเป็นมาและความสำคัญของปัญหา

งานวิจัยโดยส่วนใหญ่แล้วไม่ว่าจะเป็นสาขาวิชาใดก็ตาม มักจะอาศัยระเบียบและวิธีการทางสถิติเพื่อหาข้อสรุปสำหรับงานวิจัยนั้น ๆ ซึ่งจุดประสงค์ของการวิเคราะห์ข้อมูลคือการหาค่าตอบให้ทราบถึงลักษณะของประชากรที่สนใจจะศึกษา ไม่ว่าจะเป็นการวิเคราะห์ข้อมูลเบื้องต้น เช่น การหาค่าเฉลี่ย การหาค่าส่วนเบี่ยงเบนมาตรฐาน เป็นต้น หรือการวิเคราะห์ข้อมูลขั้นสูง เช่น การหาความสัมพันธ์ระหว่างตัวแปรต่าง ๆ เป็นต้น การใช้วิธีการทางสถิติทั้งการวิเคราะห์ในขั้นต้นและขั้นสูง สิ่งที่จะต้องทราบคือข้อตกลงเบื้องต้นของวิธีการทางสถิตินั้น ๆ เพื่อให้การวิเคราะห์ข้อมูลเป็นไปอย่างถูกต้องและแม่นยำ

ข้อตกลงเบื้องต้นหรือข้อสมมติของวิธีการสถิติที่ใช้กันโดยทั่วไปจะมีข้อสมมติหนึ่งเกี่ยวกับตัวอย่างสุ่มว่ามาจากประชากรที่มีการแจกแจงปกติ แต่ในทางปฏิบัตินั้น ส่วนใหญ่จะไม่ทราบว่าประชากรที่ทำการสุ่มตัวอย่างมานั้นมีการแจกแจงรูปแบบใด ได้มีผู้วิจัยหลายท่านได้ทำการศึกษาเกี่ยวกับความสำคัญของรูปแบบการแจกแจงของประชากรว่ามีผลต่อวิธีการทางสถิติ แบบต่าง ๆ หรือไม่ เช่น

นันทา วงษ์วิโรจน์ (2532) ได้ศึกษาตัวสถิติทดสอบที่มีความแกร่งสำหรับทดสอบความเท่ากันของค่าเฉลี่ยประชากร พบว่า ปัจจัยที่มีผลกระทบต่ออำนาจการทดสอบของตัวสถิติแบบต่าง ๆ มากที่สุดคือ อัตราส่วนเบี่ยงเบนมาตรฐานของประชากร ปัจจัยรองลงมาคือ ลักษณะการแจกแจงของประชากร ซึ่งในกรณีที่มีข้อมูลไม่ได้มีการแจกแจงปกติ จะทำให้อำนาจการทดสอบของ ANOVA F-TEST ให้ผลได้ไม่ดีเท่าที่ควร

วินัย โพธิ์สุวรรณ (2534) ได้ศึกษาการเปรียบเทียบตัวสถิติสำหรับการทดสอบความเท่ากันของความแปรปรวน พบว่า อำนาจการทดสอบของสถิติทดสอบขึ้นอยู่กับอัตราส่วนความแปรปรวนของประชากรเป็นอันดับแรก และรองลงมาคือ การแจกแจงของประชากรและขนาดของกลุ่มตัวอย่าง ตามลำดับ

สมชาย รัตนเลิศนุสรณ์ (2534) ได้ศึกษาเปรียบเทียบวิธีการตรวจสอบข้อมูลที่เกิดปกติในการวิเคราะห์การถดถอยเชิงเส้น พบว่า การแจกแจงแบบไมซ์วาที่ทำการทดสอบนั้นมีผลต่อความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1

บุญกอง ทะกลโยธิน (2535) ได้ศึกษาเปรียบเทียบการจำแนกกลุ่ม 2 กลุ่มระหว่างการวิเคราะห์การถดถอยพหุและการวิเคราะห์จำแนกประเภท เมื่อตัวแปรอิสระมีการแจกแจงแบบเบ้ พบว่า ลักษณะการแจกแจงของตัวแปรอิสระมีอิทธิพลต่อประสิทธิภาพของการจำแนกกลุ่ม โดยที่การแจกแจงปกติจะให้ประสิทธิภาพสูงกว่าการแจกแจงแบบเบ้ เมื่อจำนวนตัวอย่างมีขนาดใหญ่

จากกรณีศึกษาที่อ้างถึงข้างต้น จะเห็นได้ว่าข้อตกลงเบื้องต้นเกี่ยวกับการแจกแจงของข้อมูลที่ทำการศึกษา นั้น ถูกสมมติว่ามีการแจกแจงปกติ และผลของการศึกษาข้างต้นเป็นไปในแนวทางเดียวกันคือ เมื่อข้อมูลที่ทำมาวิเคราะห์มีการแจกแจงที่ไม่ใช่การแจกแจงปกติ จะทำให้ประสิทธิภาพของตัวสถิติที่ใช้นั้นลดน้อยลง ซึ่งวิธีการทางสถิติที่ใช้ในการวิเคราะห์ข้อมูลนั้น มีข้อตกลงเบื้องต้นเกี่ยวกับการแจกแจงของประชากร โดยให้เป็นการแจกแจงปกติ แต่ในทางปฏิบัติเราอาจจะควบคุมค่าของข้อมูลให้มีการแจกแจงปกติตามข้อสมมติ เช่น ข้อมูลทางด้านกรรมแพทย์ทางด้านเศรษฐศาสตร์ หรือทางด้านวิศวกรรมศาสตร์ เป็นต้น วิธีการทางสถิติที่เราสามารถกระทำได้เมื่อไม่ทราบการแจกแจงของข้อมูลที่ทำการศึกษา นั่นคือวิธีการทดสอบแบบไม่ใช้พารามิเตอร์ (Nonparametric Tests) มัลลิกา บุญนาค (2539 : 216) ได้ให้ความเห็นเกี่ยวกับการทดสอบแบบไม่ใช้พารามิเตอร์ไว้ว่า ถึงแม้ว่าวิธีนี้จะป็นวิธีที่ง่ายในการแก้ปัญหาแต่มีข้อเสียอยู่เช่นกัน ดังเช่น การทดสอบเหล่านั้นจะมีอำนาจการทดสอบ (Power of The Test) ต่ำ มีความเชื่อถือได้น้อย และการทดสอบที่ได้อาจไม่ลึกซึ้งเท่ากับการทดสอบที่ใช้พารามิเตอร์

(Parametric Tests)

เมื่อใดที่ขนาดตัวอย่างมีขนาดเล็กและไม่สามารถเพิ่มเติมได้ หรือเมื่อใดที่เราทราบว่าการแจกแจงของประชากรที่เราทำการสุ่มตัวอย่างมานั้น ไม่ได้มีการแจกแจงปกติ แล้วเรายังคงทำการทดสอบแบบใช้พารามิเตอร์ที่มีข้อตกลงเบื้องต้นเกี่ยวกับการแจกแจงปกติต่อไป การทดสอบนั้นอาจจะให้ผลที่คลาดเคลื่อนและได้ผลไม่เหมาะสมนัก อีกประการหนึ่งของวิธีการแก้ปัญหาดังกล่าวคือการพยายามที่จะหาและเลือกใช้การทดสอบแบบใช้พารามิเตอร์อื่น ๆ ที่เหมาะสมกับข้อมูลที่มีอยู่ แต่มักจะพบว่าตัวสถิติเหล่านั้นจะยุ่งยากมากขึ้นและไม่เป็นที่รู้จักของนักวิจัยในสาขาอื่น ๆ ดังนั้นเพื่อความสะดวก เราจึงยังต้องการข้อสรุปที่ว่า การแจกแจงของข้อมูลนั้นเป็นแบบปกติ

วิธีหนึ่งที่เป็นไปได้ในการที่จะนำมาใช้ในการแก้ปัญหา คือ การแปลงข้อมูล (Data Transformation) โดยที่สามารถใช้วิธีการต่าง ๆ ทางคณิตศาสตร์ที่ไม่ยุ่งยากและซับซ้อนเกินไป นำมาปรับข้อมูลเดิมของเราที่มีอยู่ให้อยู่ในขนาดใหม่ เพื่อให้การแจกแจงที่ไม่เป็นแบบปกตินั้น มีการแจกแจงที่เข้าใกล้กับการแจกแจงปกติ หรือเป็นการแจกแจงปกติ ซึ่งทางสาขาวิชาต่าง ๆ โดยทั่วไปนั้น มักจะใช้การแปลงข้อมูลอยู่บ่อย ๆ และรูปแบบการแปลงที่ใช้จะเป็นรูปแบบที่ง่าย ๆ กระทำได้เมื่อต้องการจะแปลงข้อมูลให้มีการแจกแจงปกติ ในการศึกษาครั้งนี้ ผู้วิจัยได้ทำการศึกษาถึงวิธีการแปลงข้อมูลแบบต่าง ๆ กับการแจกแจงที่มีความเบ้และความโด่งแตกต่างจากการแจกแจงปกติ โดยที่ใช้การพิจารณาว่ารูปแบบการแปลงข้อมูลแบบใดที่จะเหมาะสมกับข้อมูลชุดนั้น ๆ

วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาหารูปแบบการแปลงข้อมูลจากข้อมูลที่มีการแจกแจงแบบไม่ปกติ ให้มีการแจกแจงปกติ
2. เพื่อหาข้อจำกัดของรูปแบบการแปลงข้อมูลที่เหมาะสมกับการแจกแจงแบบไม่ปกติ

สมมติฐานของการวิจัย

การแปลงข้อมูลจากรูปแบบการแจกแจงที่มีความเบ้และความโด่งแตกต่างกัน จะใช้รูปแบบการแปลงที่ต่างกัน และสามารถใช้รูปแบบการแปลงนั้นกับข้อมูลจริงได้

ข้อตกลงเบื้องต้น

การแปลงข้อมูลมีจุดประสงค์เพื่อต้องการปรับการแจกแจงของข้อมูลให้มีการแจกแจงปกติ และข้อมูลที่ศึกษาเป็นข้อมูลที่มีค่าบวก

ขอบเขตของการวิจัย

1. วิธีการแปลงข้อมูล
วิธีที่พิจารณาในการวิจัยครั้งนี้คือ

- The Logarithm Transformation
ได้แก่รูปแบบ $\log(x)$, $\ln(x)$
- The Square Root Transformation
ได้แก่รูปแบบ \sqrt{x}
- The Recipocal Transformation
ได้แก่รูปแบบ $1/x$
- Power Transformations
ได้แก่รูปแบบ x^p และ $\frac{x^p - 1}{p}$
เมื่อ p คือเลขจำนวนจริงใดๆ

2. ลักษณะข้อมูล

2.1 ข้อมูลที่ได้จากการจำลอง กำหนดให้

- การแจกแจงแบบเบ้ขวา (Positive-Skewed Distribution) ซึ่งการกำหนดค่าพารามิเตอร์ต่างๆ คือ
 - ความเบ้ มีค่าตั้งแต่ 0.2 ถึง 2.0
 - ความโด่ง มีค่าตั้งแต่ 2.0 ถึง 10.0
- การแจกแจงแบบเบ้ซ้าย (Negative-Skewed Distribution) ซึ่งการกำหนดค่าพารามิเตอร์ต่างๆ คือ
 - ความเบ้ มีค่าตั้งแต่ -0.2 ถึง -2.0
 - ความโด่ง มีค่าตั้งแต่ 2.0 ถึง 10.0
- ขนาดของกลุ่มตัวอย่าง ให้มีขนาดทั้งหมด 5 ระดับคือ 20, 30, 40, 50 และ 100
- ระดับนัยสำคัญ (α) ของการทดสอบเทียบความกลมกลืน โดยใช้สถิติไคโมโกรอฟ – สมิทอฟ กำหนดที่ระดับ 0.01, 0.05, 0.10, 0.15
- การศึกษาค้างนี้ ทำการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล (Monte Carlo Simulation Technique)

2.2 ข้อมูลที่เกิดขึ้นจริง ทำการศึกษาจากข้อมูลที่เก็บรวบรวมทั้งสิ้น 17 ชุด

ประโยชน์ที่คาดว่าจะได้รับ

1. เพื่อเป็นแนวทางในการตัดสินใจเลือกรูปแบบการแปลงให้เหมาะสมกับรูปแบบของข้อมูลที่มี เมื่อต้องการให้ข้อมูลมีการแจกแจงปกติ
2. เพื่อเป็นแนวทางในการศึกษาวิจัยต่อไป



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย