

สตรีรัตน์ เกิดสว่าง : การเปรียบเทียบเทคนิคการแปลงข้อมูลให้มีการแจกแจงแบบปกติ

(COMPARISON OF DATA TRANSFORMATION TECHNIQUES FOR NORMALITY)

อ.ที่ปรึกษา : รศ.ร.อ.มานพ วราภักดิ์, 238 หน้า. ISBN 974-17-4030-1

วัตถุประสงค์ของการวิจัยในครั้งนี้ เพื่อศึกษาและเปรียบเทียบรูปแบบการแปลงข้อมูลที่มีการแจกแจงแบบไม่เป็นปกติให้มีการแจกแจงเข้าสู่การแจกแจงแบบปกติ โดยพิจารณารูปแบบการแปลงข้อมูลทั้งหมด 4 รูปแบบด้วยกันคือ รูปแบบการแปลงกำลังของ Box และ Cox (Box and Cox power transformations) รูปแบบการแปลงแบบดัดแปลงของ Box และ Cox (modified Box and Cox power transformations) รูปแบบการแปลงแบบเอกซ์โพเนนเชียลของ Manly (exponential transformations) และรูปแบบการแปลงกำลังแบบใหม่ของ In-Kwon Yeo (new power transformations) โดยพิจารณาจากเปอร์เซ็นต์การยอมรับการแปลงว่าได้รับการแจกแจงแบบปกติจากการทดสอบสมมติฐาน H_0 : ข้อมูลมาจากการแจกแจงปกติ เพื่อเลือกรูปแบบการแปลงแบบใดแปลงข้อมูลให้มีการแจกแจงเข้าใกล้การแจกแจงแบบปกติมากที่สุด ภายใต้สถานการณ์ที่กำหนด คือ ประชากรมีการแจกแจงแบบจอร์นสัน ซึ่งกำหนดให้มีลักษณะโค้งเบ้ต่างๆ ขนาดตัวอย่าง คือ 10, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90 และ 100 ณ ระดับนัยสำคัญของการทดสอบเทียบความกลมกลืน (α) 0.05 ในการวิจัยครั้งนี้จำลองสถานการณ์การทดลองด้วยเทคนิคมอนติคาร์โล ซึ่งทำการจำลองซ้ำ 500 รอบในแต่ละสถานการณ์ และใช้ข้อมูลที่เกิดขึ้นจริงประกอบการศึกษาในครั้งนี้ด้วย ผลสรุปของการวิจัยมีดังนี้

1. กรณีที่ข้อมูลมีทั้งค่าบวกและค่าลบ สำหรับการแจกแจงแบบเบ้ขวาและการแจกแจงแบบเบ้ซ้าย ที่ความเบ้ระดับต่ำ รูปแบบการแปลงแบบดัดแปลงของ Box และ Cox ให้เปอร์เซ็นต์การยอมรับ H_0 สูงที่สุดเป็นส่วนใหญ่ ส่วนที่ความเบ้ระดับสูงขึ้นไป รูปแบบการแปลงแบบเอกซ์โพเนนเชียลให้เปอร์เซ็นต์การยอมรับ H_0 สูงที่สุดเป็นส่วนใหญ่
2. กรณีที่ข้อมูลมีเฉพาะค่าบวก สำหรับการแจกแจงแบบเบ้ขวา ในทุกระดับความเบ้ รูปแบบการแปลงกำลังของ Box และ Cox รูปแบบที่ 1 และรูปแบบการแปลงกำลังของ Box และ Cox รูปแบบที่ 2 ให้เปอร์เซ็นต์การยอมรับ H_0 สูงที่สุดเป็นส่วนใหญ่ สำหรับการแจกแจงแบบเบ้ซ้าย ในทุกระดับความเบ้ รูปแบบการแปลงแบบเอกซ์โพเนนเชียล ให้เปอร์เซ็นต์การยอมรับ H_0 สูงที่สุดเป็นส่วนใหญ่

4682447026 : MAJOR STATISTICS

KEY WORD : DATA TRANSFORMATION / BOX AND COX POWER TRANSFORMATIONS / MODIFIED BOX AND COX POWER TRANSFORMATIONS / EXPONENTIAL TRANSFORMATIONS / NEW POWER TRANSFORMATIONS

SATREERATH KERDSAWANG : COMPARISON OF DATA TRANSFORMATION TECHNIQUES FOR NORMALITY. THESIS ADVISOR : ASSOC. PROF. CAPT. MANOP VARAPHAKDI, M.S. 238 pp. ISBN 974-17-4030-1

The objective of this research is to investigate and compare the data transformation forms which can transform data to normal distribution. The data transformation forms are Box and Cox power transformations, modified Box and Cox power transformations, Manly exponential transformations and In-Kwon Yeo new power transformations. The percentage of accept H_0 : normal distribution is used in evaluating the capability of the data transformation forms when the population are Johnson that assign shape of distribution by skewness and kurtosis. Sample sizes are 10, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90 and 100. Level of significance of the goodness of fit test (α) is 0.05. For this research, The percentage of accept H_0 : normal distribution are computed through the Monte Carlo Simulation method. This simulation is repeated 500 times in each situation and real data are used in this research as well. The result of this research can be summarized as follows:

1. Data set with both positive and negative values : Both positive skew-distribution and negative skew-distribution, at the low level of skewness, the most percentage of accept H_0 came from modified Box and Cox power transformations. At the high level of skewness, the most percentage of accept H_0 came from exponential transformations.
2. Data set with positive values : Positive skew-distribution, for each level of skewness, the most percentage of accept H_0 came from Box and Cox power transformations type 1 and Box and Cox power transformations type 2. Negative skew-distribution, for each level of skewness, the most percentage of accept H_0 came from exponential transformations.