

วัตถุประสงค์ของการวิจัยครั้งนี้เพื่อแก้ปัญหาข้อมูลผิดปกติในการวิเคราะห์ข้อมูลที่ได้จากแผนแบบการทดลองแบบสุ่มตลอด โดยที่ตัวแบบมีรูปแบบดังนี้ $Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$ เมื่อ $i = 1, \dots, k$ และ $j = 1, \dots, n$ โดยที่ Y_{ij} แทนข้อมูลตอบสนองที่ j ได้รับวิธีทดลองที่ i μ แทนค่าเฉลี่ยรวม τ_i แทนอิทธิพลของวิธีทดลองที่ i ε_{ij} แทนความคลาดเคลื่อนของข้อมูลตอบสนองที่ j ซึ่งได้รับวิธีทดลองที่ i และ ε_{ij} มีการแจกแจงแบบปกติโดยเป็นอิสระซึ่งกันและกัน มีค่าเฉลี่ยเป็น 0 และความแปรปรวนเป็น σ^2 k แทนจำนวนวิธีทดลอง และ n แทนจำนวนซ้ำในแต่ละวิธีทดลอง ในการวิจัยครั้งนี้ได้ทำการสร้างข้อมูลตามขอบเขตการวิจัยด้วยโปรแกรม S-PLUS 2000 โดยกำหนดให้จำนวนวิธีทดลองเท่ากับ 3, 4 และ 5 จำนวนซ้ำในการทดลองเท่ากับ 5, 6, 7 และ 8 ให้สัมประสิทธิ์การแปรผันมีความแตกต่างกัน 3 ระดับ ได้แก่ 30% 40% และ 50% สำหรับเกณฑ์ที่ใช้ในการเปรียบเทียบวิธีการแปลงข้อมูลที่เหมาะสมในการแก้ปัญหาดังกล่าวข้างต้นคือ ค่าสัดส่วนของจำนวนค่าผิดปกติที่ลดลงภายหลังการแปลงข้อมูล ค่าสัดส่วนของข้อมูลที่มีการแจกแจงแบบปกติและความแปรปรวนเท่ากันภายหลังการแปลงข้อมูล ค่าสัดส่วนของการปฏิเสธสมมติฐานว่างและอำนาจการทดสอบของการทดสอบเอฟที่ระดับนัยสำคัญ 0.01 และ 0.05 ผลการศึกษาจะสรุปได้ดังนี้

1. การแปลงข้อมูลด้วยค่า $\lambda = 0.5$ ให้ค่าสัดส่วนของจำนวนค่าผิดปกติที่ลดลงภายหลังการแปลงข้อมูลสูงที่สุด และพบว่าการแปลงข้อมูลด้วย $\lambda = 0.5$ ยังให้ค่าสัดส่วนของข้อมูลที่มีการแจกแจงแบบปกติและความแปรปรวนเท่ากันภายหลังการแปลงข้อมูลสูงที่สุด

2. การแปลงข้อมูลด้วยค่า $\lambda = 0.5$ และ $\lambda = 1.5$ เป็นวิธีการแปลงข้อมูลที่เหมาะสมในการแก้ปัญหาค่าผิดปกติที่ได้จากแผนแบบการทดลองสุ่มตลอด และทั้งควรระวังปัญหาการแจกแจงของข้อมูลตอบสนองอาจไม่เป็นการแจกแจงแบบปกติและความแปรปรวนไม่เท่ากันภายหลังการแปลงข้อมูลด้วย $\lambda = 1.5$

3. เมื่อจำนวนค่าผิดปกติมากขึ้นพบว่า ในแต่ละสถานการณ์ค่าสัดส่วนของจำนวนค่าผิดปกติที่ลดลงภายหลังการแปลงข้อมูลจะมีแนวโน้มสูงขึ้น

The objective of this study is to solve outlier problem of the observation from completely randomized design by transformation. The fixed-effect completely randomized design model is $Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$ when $i = 1, \dots, k$ and $j = 1, \dots, n$ where Y_{ij} is the j^{th} response variable for the i^{th} treatment, μ is the grand mean, τ_i is the i^{th} treatment effect and ε_{ij} is a random error of the j^{th} response variable for the i^{th} treatments. The ε_{ij} is independently and normally distribution with mean 0 and variance σ^2 when k is the number of treatment. In this study, the generation of response observations is done by using S-PLUS 2000 package. The data are generated with 3, 4 and 5 treatments. The number of replications for each treatment is 5, 6, 7 and 8 respectively. The coefficient of variation is the specified at 30%, 40% and 50%. The criterion used for success of outlier correction is the proportion of outlier case that reduced after transformation, the proportion of normal-distribution case data and homogeneity of variances after transformation, the proportion of null hypothesis rejection and the power of the test at 0.01 and 0.05 significance level. These criterion are measures for comparison about transformed method. The result of this study can be summarized as follow:

1. The transformation of data with $\lambda = 0.5$ has the highest proportion of outlier case that are reduced after transforming. Moreover, the transformed data with $\lambda = 0.5$ still sometime normally distribution and homogeneity of variances with high proportion.
2. The transformation of data with $\lambda = 0.5$ and $\lambda = 1.5$ are suitable for solving outlier data. However, the response data may be non-normality and heterogeneity when transformed data with $\lambda = 1.5$
3. When the number of outliers increase, the proportion of case that are reduced after transformation increase.