

โครงการวิจัยเรื่อง : การศึกษาตัวสถิติทดสอบ T - tests สำหรับทดสอบค่าเฉลี่ยของสองประชากรปกติที่เป็นอิสระกัน

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์ที่จะศึกษาคุณสมบัติของสองตัวแบบทดสอบ T - tests (T_1 และ T_2) ซึ่งได้รับการเสนอแนะให้ใช้สำหรับทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของสองประชากรปกติที่เป็นอิสระกันแต่ไม่ทราบค่าความแปรปรวน พร้อมทั้งศึกษาเปรียบเทียบประสิทธิภาพของตัวแบบทดสอบทั้งสองด้วย โดยศึกษาจากความแกร่งของการทดสอบ (robustness of test) และค่าประสิทธิภาพสัมพัทธ์ (relative efficiency) ซึ่งพิจารณาหรือคำนวณได้จากค่าความถี่สัมพัทธ์ของผลการทดสอบสมมติฐานที่สรุปว่าปฏิเสธสมมติฐานหลัก ที่ซึ่งค่านี้คำนวณได้จากการจำลองข้อมูลขึ้นในเครื่องคอมพิวเตอร์ด้วยเทคนิคมอนติคาร์โล โดยแต่ละสถานการณ์ที่ศึกษาจะกระทำซ้ำ ๆ 10,000 ครั้ง ผลการวิเคราะห์ข้อมูลสามารถสรุปได้ว่า

สำหรับกรณีที่สมมติฐานหลักเป็นจริง ($H_0: \mu_1 = \mu_2$ เป็นจริง) เมื่อความแปรปรวนของประชากรทั้งสองเท่ากัน พบว่า เมื่อขนาดตัวอย่าง n_1 และ n_2 มีขนาดมากกว่า 5 ขึ้นไป ตัวแบบทดสอบทั้งสองจะมีความแกร่งของการทดสอบ แต่หากความแปรปรวนของประชากรทั้งสองไม่เท่ากันแล้ว พบว่า เมื่อกำหนดขนาดตัวอย่างทั้งสองเท่ากัน ($n_1 = n_2$) ตัวแบบทดสอบทั้งสองจะมีความแกร่งของการทดสอบ แต่ถ้า $n_1 \neq n_2$ ตัวแบบทดสอบ T_2 เท่านั้นที่มีความแกร่งของการทดสอบ

ในกรณีที่สมมติฐานหลักไม่จริง เมื่อความแปรปรวนของประชากรทั้งสองเท่ากัน พบว่า ประสิทธิภาพของตัวแบบทดสอบ T_1 สูงกว่า T_2 เล็กน้อย (ประมาณ 1.05 เท่า) แต่ถ้าผลต่างระหว่างค่าเฉลี่ยทั้งสองมีค่ามากกว่า 2 ขึ้นไป ประสิทธิภาพของตัวแบบทดสอบทั้งสองแทบไม่แตกต่างกัน สำหรับกรณีความแปรปรวนของประชากรทั้งสองไม่เท่ากัน (โดยที่ $\sigma_2^2 > \sigma_1^2$) พบว่า เมื่อ $n_1 = n_2$ ณ ระดับนัยสำคัญที่กำหนดเท่ากับ 0.10 และ 0.05 และผลต่างระหว่างค่าเฉลี่ยทั้งสองมีค่ามากกว่า 1.0 ประสิทธิภาพของตัวแบบทดสอบทั้งสองไม่ค่อยแตกต่างกัน แต่ถ้า $n_1 > n_2$ โดยเฉพาะอย่างยิ่งเมื่อผลต่างระหว่างค่าเฉลี่ยทั้งสองมีค่าน้อยกว่า 2.50 หรือค่าอัตราส่วนของความแปรปรวนของประชากรที่ 2 เทียบกับค่าความแปรปรวนของประชากรที่ 1 มีค่ามาก พบว่า ประสิทธิภาพของตัวแบบทดสอบ T_1 จะสูงกว่าประสิทธิภาพของตัวแบบทดสอบ T_2 ส่วนเมื่อ $n_1 < n_2$ กลับพบว่า ประสิทธิภาพของตัวแบบทดสอบ T_2 จะสูงกว่าประสิทธิภาพของตัวแบบทดสอบ T_1 โดยผลจะปรากฏชัดเจนยิ่งขึ้นเมื่ออัตราส่วนของความแปรปรวนของประชากรที่ 2 เทียบกับค่าความแปรปรวนของประชากรที่ 1 มีค่ามากขึ้น

Assoc. Prof. Anothai Trevanich, Ph.D.
Assist. Prof. Chatchavan Ruangpraphan

Statistical Department
Faculty of Science
Khon Kaen University
tel (043) 237660

Title : INVESTIGATION OF THE T -TESTS FOR THE MEANS OF TWO NORMALLY INDEPENDENT POPULATIONS

ABSTRACT

The objective of this research is to investigate and compare the property of the T - tests (T_1 and T_2) using the robustness and the relative efficiency , which are the general tests for hypothesis testing about the means of two independently normal populations with unknown variances. The robustness and the relative efficiency are considered in term of the relative frequency of the conclusions that reject H_0 . The data were obtained through simulation with repeating 10,000 times for each predicament of the experiment. And a computer program was designed to calculate the relative frequencies. The results of the study as classified by the truth of the null hypothesis can be summarized as follows.

For the null hypothesis ($H_0 : \mu_1 = \mu_2$) is true , looking up when the variances equal, we found that T_1 - test and T_2 - test are robust , specially for n_1 and $n_2 > 5$. In the case of variances unequal, as using $n_1 = n_2$, both tests are robust. While if $n_1 \neq n_2$, T_2 is robust only.

For the null hypothesis is not true, we found that when variances equal, the efficiency of T_1 is about 1.05 times of the efficiency of T_2 . (accept at the significance level is 0.01). That, there is little difference. However, we can say that theirs efficiencies are not difference, specially for the difference between means ($d_0 = | \mu_1 - \mu_2 |$) is greater than 2. In the case when variances unequal, we found as follows. If $n_1 = n_2$ and significance level is 0.10 or 0.05 , the efficiencies of T_1 and of T_2 are quite not difference , obviously as $d_0 > 1.0$. But when $n_1 > n_2$, the efficiency of T_1 is higher than the efficiency of T_2 , specially at $d_0 < 2.5$ or the large variance ratio (σ_2^2 / σ_1^2) . In the other hand, the efficiency of T_2 is higher than the efficiency of T_1 when $n_1 < n_2$. These results will be more clearly if the variance ratio increases.