# A COMPARISON OF TESTS FOR HOMOGENEITY OF THE RISK DIFFERENCE WHEN DATA ARE SPARSE

**PATRANUCH SAPCHOOKUL**

**A THESIS SUBMITTED IN PARTIAL FULFILLMENT**
**OF THE REQUIREMENTS FOR**
**THE DEGREE OF MASTER OF SCIENCE (BIOSTATISTICS)**
**FACULTY OF GRADUATE STUDIES**
**MAHIDOL UNIVERSITY**
**2005**

**ISBN 974-04-6318-5**

**COPYRIGHT OF MAHIDOL UNIVERSITY**

การเปรียบเทียบสถิติทดสอบสำหรับการเท่ากัน (Homogeneity) ของ Risk Difference เมื่อข้อมูลมีขนาดเล็ก (A COMPARISON OF TESTS FOR HOMOGENEITY OF THE RISK DIFFERENCE WHEN DATA ARE SPARSE)

ภัทรานุช ทรัพย์ชูกุล 4437166 PHBS/M

วท.ม. (ชีวสถิติ)

คณะกรรมการควบคุมวิทยานิพนธ์ : ชูเกียรติ วิวัฒน์วงศ์เกษม, M.Sc.(Biostatistics), Dankmar Böhning, Ph.D. (Mathematics), ธวัชชัย วรพงศธร, Ph.D. (Research & Stat in Ed.), ปรารถนา สถิตย์วิภาวี, M.P.H.(Biostatistics).

**บทคัดย่อ**

การวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของสถิติภายใต้การเท่ากันของความแตกต่างของอัตราเสี่ยง ด้วยสถิติทดสอบห้าตัวที่ใช้ในงานวิจัยของ Lipsitz et al. ที่ระดับนัยสำคัญ 0.01, 0.05 และ 0.10 โดยพิจารณาทั้งแบบทางเดียวและแบบสองทาง โดยจำนวนศูนย์รักษา คือ 8, 16, 32, 48 ค่าเฉลี่ยของขนาดตัวอย่างในกลุ่มทดลอง และกลุ่มควบคุมมีแบบเท่ากันและไม่เท่ากัน คือ 4, 8, 16, 32 และ 64 ข้อมูลถูกสร้างขึ้นด้วยเทคนิคมอนติคาร์โล โดยใช้ภาษาฟอร์แทรน 90 ในการคำนวณหาค่าความคลาดเคลื่อนชนิดที่ 1 และอำนาจการทดสอบของสถิติทดสอบทั้งห้า ซ้ำๆ กัน 10,000 ครั้ง ในแต่ละสถานการณ์

ผลการศึกษาพบว่า สถิติทดสอบ $Q_{WLS}$ สามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้ดีเมื่อค่าเฉลี่ยของขนาดตัวอย่างมีขนาดใหญ่ และสถิติทดสอบ $Z^2_{WLS,R}$ และ $Z^2_V$ สามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้ดีเมื่อค่าเฉลี่ยของขนาดตัวอย่างมีขนาดเล็ก ด้วยการจำลองข้อมูลด้วยวิธี dropping center โดยที่จำนวนศูนย์รักษาจะมีขนาดเท่าไรก็ตาม และสำหรับการจำลองข้อมูลด้วยวิธี adding constant จะมีเพียงสถิติทดสอบ $Z^2_{WLS}$ ที่สามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้ดี เมื่อค่าเฉลี่ยของขนาดตัวอย่างมีขนาดเล็ก

สำหรับอำนาจการทดสอบพบว่า การจำลองข้อมูลด้วยวิธี dropping center, สถิติทดสอบ $Q_{WLS}$ จะมีอำนาจทดสอบสูงที่สุดในทุกขนาดของศูนย์รักษา และทุกค่าเฉลี่ยของขนาดตัวอย่าง ทั้งแบบเท่ากันและไม่เท่ากัน การจำลองข้อมูลด้วย adding constant, เมื่อค่าเฉลี่ยของตัวอย่างมีขนาดเล็ก สถิติทดสอบ $Z^2_K$ จะมีอำนาจการทดสอบสูงที่สุด แต่ถ้าค่าเฉลี่ยของตัวอย่างมีขนาดใหญ่ สถิติทดสอบ $Q_{WLS}$ จะมีอำนาจทดสอบสูงที่สุด ในกรณีที่ศูนย์รักษามีจำนวนน้อย และ/หรือ ค่าเฉลี่ยของตัวอย่างในแต่ละกลุ่มทดลองมีจำนวนน้อย สถิติทดสอบ $Q_{WLS}$ จะทำได้ดีในเรื่องของการควบคุมความคลาดเคลื่อนชนิดที่ 1 และมีอำนาจการทดสอบที่สูง

180 หน้า. ISBN 974-04-6318-5

A COMPARISON OF TESTS FOR HOMOGENEITY OF THE RISK DIFFERENCE WHEN DATA ARE SPARSE

PATRANUCH SAPCHOOKUL 4437166 PHBS/M

M.Sc. (BIOSTATISTICS)

THESIS ADVISORS: CHUKIAT VIWATWONGKASEM, M.Sc.(Biostatistics), DANKMAR BÖHNING, Ph.D.(Mathematics), THAVATCHAI VORAPONGSATHORN, Ph.D. (Research & Stat in Ed.), PRATANA SATITVIPAWEE, M.P.H.(Biostatistics).

## ABSTRACT

The objective of this research was to compare the Type I error and the power of tests for homogeneity of risk difference by using the statistical tests following the work of Lipsitz et al. (1998). Hypothesis testing considered 0.01, 0.05 and 0.10 levels of significance for one-sided and two-sided tests and when the number of centers $K$ = 8, 16, 32, and 48 and mean treatment group per center equaled 4, 8, 16, 32, and 64. This study considered when the mean sample size in each treatment group for equal and unequal cases. The data was simulated by Monte Carlo simulation technique compatible with FORTRAN 90 language. A FORTRAN program was designed to calculate the probability of the Type I error and the power of all five tests in 10,000 replications for each condition. When the estimated variance in each center was zero, the study adjusted the statistics by adding 0.5 to each cell (in 2x2 tables) and dropping centers with zero variance.

Results indicated that the $Q_{WLS}$ test performs well when mean treatment group is large ($n_{ij} \geq 32$) with respect to the Type I error and the $Z^2_{WLS,R}, Z^2_V$ tests perform well when mean treatment group is less than moderate ($n_{ij} \leq 16$) by the method of dropping case, regardless of the number of centers. For the adding constant method, only the $Z^2_{WLS}$ test performs well when mean treatment group is less than moderate ($n_{ij} \leq 16$). With respect to the power, by the dropping case method the $Q_{WLS}$ test has the highest power in every number of centers and mean treatment group size. For the adding constant if mean treatment group is less than moderate then the $Z^2_K$ test has the highest power, but if the treatment group is large then the $Q_{WLS}$ test has the highest power. In cases where the number of centers is small and/or the mean treatment group in each center is sparse, the weight least squares statistic ($Q_{WLS}$) test appears to perform the best.

KEY WORDS: RISK DIFFERENCE / TYPE I ERROR / POWER OF THE TEST / HOMOGENEITY / SPARSE DATA

180 P. ISBN 974-04-6318-5