

#### บทที่ 4

#### ผลการวิจัย

ในการวิจัยครั้งนี้ต้องการศึกษาเปรียบเทียบสถิติทดสอบความแปรปรวนของประชากรสองกลุ่ม คือ การทดสอบเอฟ (F test) การทดสอบคลอทซ์ (Klotz test) การทดสอบมูด (Mood test) การทดสอบทูกี-ซีเกิล (Tukey-Siegel test) และการทดสอบสแควร์เรงค์ (squared ranks test) โดยศึกษาอำนาจของการทดสอบค่าและความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 เมื่อประชากรมีการแจกแจงแบบปกติ (Normal Distribution) แบบยูนิฟอร์ม (Uniform Distribution) และการแจกแจงแบบโลจิสติก (Logistic Distribution) สำหรับกรณีที่ประชากรมีการแจกแจงแบบเบ้ นั้นจะศึกษาเปรียบเทียบเฉพาะการทดสอบเอฟ และการทดสอบสแควร์เรงค์ ( $X^2$ ) โดยจะศึกษาในกรณีที่ความเบ้ของประชากรเป็นบวก และมีค่าสูงสุด 1.0 ส่วนความโด่งจะพิจารณาความโด่งที่มีค่าในช่วง [2.8, 4.2] สำหรับขนาดของตัวอย่างที่ใช้ในการศึกษาแบ่งเป็นขนาดของตัวอย่างเท่ากันคือ 10, 15 และ 20 และขนาดของตัวอย่างไม่เท่ากัน คือ ขนาดของตัวอย่างกลุ่มที่ 1 และกลุ่มที่ 2 เป็นดังนี้ (5: 10) (5: 15) และ (5: 20)

ผลจากการวิจัยครั้งนี้ได้เป็น 2 ลักษณะ คือ ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจของการทดสอบ ซึ่งจะนำเสนอเป็นตาราง แผนภาพ และกราฟ เพื่อให้สะดวกในการอธิบาย สิ่งใช้สัญลักษณ์ต่อไปนี้ เพื่อแทนความหมายต่าง ๆ ดังนี้

T หมายถึง ความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จากการทดลอง

$\alpha$  หมายถึง ระดับนัยสำคัญที่กำหนด

$\alpha_3$  หมายถึง ความเบ้ของประชากร

$\alpha_4$  หมายถึง ความโด่งของประชากร

NN (m, n) หมายถึง ประชากรทั้งสองกลุ่มมีการแจกแจงแบบปกติ ขนาดของตัวอย่างที่ 1 เป็น m และกลุ่มที่ 2 เป็น n

UU (m, n) หมายถึง ประชากรทั้งสองกลุ่มมีการแจกแจงแบบยูนิฟอร์ม ขนาดของตัวอย่างกลุ่มที่ 1 เป็น m และกลุ่มที่ 2 เป็น n

LL (m, n) หมายถึง ประชากรทั้งสองกลุ่มมีการแจกแจงแบบโลจิสติก ขนาดของตัวอย่างกลุ่มที่ 1 เป็น m และกลุ่มที่ 2 เป็น n

$\Delta^2$  หมายถึง  $\frac{\sigma_y^2}{\sigma_x^2}$  โดยที่  $\sigma_x^2$  หมายถึงความแปรปรวนของประชากร X และ  $\sigma_y^2$

หมายถึง ความแปรปรวนของประชากร Y

F หมายถึง การทดสอบเอฟ (F test)

KN หมายถึง การทดสอบคล็อทซ์ (Klotz test)

TS หมายถึง การทดสอบทูกี้-ซีเกล (Tukey-Siegel test)

SRT ( $\bar{X}$ ) หมายถึง การทดสอบสแควร์แรงค์ (squared ranks test) ซึ่งคำนวณค่าสถิติจากอันดับของการจัดเรียงค่า  $|x_i - \bar{X}|$  และ  $|y_i - \bar{Y}|$  ร่วมกัน

SRT ( $\mu$ ) หมายถึง การทดสอบสแควร์แรงค์ (squared ranks test) ซึ่งคำนวณค่าสถิติจากอันดับของการจัดเรียงค่า  $|x_i - \mu_x|$  และ  $|y_i - \mu_y|$  ร่วมกัน

M หมายถึง การทดสอบมูด (Mood test)

### ความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1

สำหรับความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จากการทดลองจะนำเสนอในลักษณะของแผนภาพ โดยใช้เกณฑ์ในการพิจารณาความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ของ Cochran (1954; อ้างโดย Ramsey 1980:337-349) และเกณฑ์ของ Bradley (1978:144-152) ที่พิจารณาควบคู่กัน ซึ่งรายละเอียดสำหรับแต่ละเกณฑ์ที่ใช้ในการพิจารณาเป็นดังนี้

เกณฑ์ของ Cochran กำหนดให้  $\tau$  คือค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ที่เกิดจากการทดลอง ถ้า  $\tau$  อยู่ในช่วง  $[-.007, -.015]$  ที่ระดับนัยสำคัญ .01 และ  $\tau$  มีค่าในช่วง  $[-.04, -.06]$  ที่ระดับนัยสำคัญ .05 จะถือว่าการทดสอบนั้นควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้

เกณฑ์ของ Bradley กำหนดให้  $\tau$  คือค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ที่เกิดจากการทดลอง ถ้า  $\tau$  อยู่ในช่วง  $[-.05\alpha, 1.5\alpha]$  แล้ว จะถือว่าการทดสอบนั้นควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ ซึ่งหมายความว่าที่ระดับนัยสำคัญ .01

ค่า  $t$  ต้องมีค่าในช่วง  $[-.005, .015]$  ที่ระดับนัยสำคัญ  $.05$   $t$  จะต้องอยู่ในช่วง  $[-.025, .075]$  และที่ระดับนัยสำคัญ  $.10$   $t$  จะต้องอยู่ในช่วง  $[-.050, .150]$  ซึ่งจะถือว่าการทดลองนั้นควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้

สำหรับเกณฑ์ของ Cochran ที่ระดับนัยสำคัญ  $.10$  จะใช้ช่วงความเชื่อมั่น  $95\%$  ของค่า  $\alpha$  แทน โดยคิดจากสูตร  $\alpha \pm z_{.05} \left( \sqrt{\frac{\alpha(1-\alpha)}{N}} \right)$  แทน ดังนั้นสำหรับวิจัยครั้งนี้จะได้ว่าค่า  $t$  อยู่ในช่วง  $[-.081, .119]$  จะถือว่าการทดลองนั้นสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้

จากผลการทดลอง ถ้าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของการทดลองใดอยู่นอกขอบเขตที่ระบุสำหรับแต่ละเกณฑ์ที่กำหนดจะถือว่าการทดลองนั้นไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ ซึ่งแยกได้เป็น 2 กรณีคือ

1. กรณีที่ความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 มากกว่าขอบเขตบนของเกณฑ์ที่ใช้พิจารณา จะถือว่าการทดลองนั้นมีค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 มากกว่าค่า  $\alpha$  ที่กำหนด ( $\tau > \alpha$ )
2. กรณีที่ความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 น้อยกว่าขอบเขตล่างของเกณฑ์ที่ใช้พิจารณา จะถือว่าการทดลองนั้นมีค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 น้อยกว่า  $\alpha$  ที่กำหนด ( $\tau < \alpha$ )

ในกรณีที่ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 อยู่ในขอบเขตที่ระบุสำหรับแต่ละเกณฑ์ที่กำหนดจะถือว่าการทดลองนั้นมีค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 เท่ากับค่า  $\alpha$  ที่กำหนด ( $\tau = \alpha$ ) และสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้

สำหรับการนำเสนอความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จากการทดลองแบ่งได้เป็น 2 กรณีคือ

1. กรณีที่ประชากรมีการแจกแจงแบบปกติ, แบบยูนิฟอร์ม และแบบโลจิสติก ซึ่งแยกได้เป็น 2 กรณีย่อยคือ

- 1.1 ความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของการทดลองแบบการทดลองคลอท์ การทดลองทูก์-ซีเกล การทดลองแลคัวร์แรงค์ ( $n^2$ ) และการทดลองมัตต์ จะนำเสนอเป็นรูปแผนภาพ ซึ่งแกนนอนแทนสถิติทดสอบทั้ง 5 วิธี แกนตั้งแทนค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของการทดลอง 4 สั้นประซึ่งอยู่ในแผนภาพแทนขอบเขตบนและขอบเขต

ล่าง ของค่า  $t$  ด้วย วิกฤตของ Bradley และวิกฤตของ Cochran สัญญลักษณ์ B แทนวิกฤตของ Bradley สัญญลักษณ์ C แทนวิกฤตของ Cochran สำหรับค่าอธิบายในรูปนี้  $n_1:n_2$  แทนอัตราส่วนของตัวอย่างกลุ่มที่ 1 และกลุ่มที่ 2  $\alpha$  แทนระดับนัยสำคัญที่กำหนด และ  $\alpha_4$  แทนความโด่งของประชากร ส่วนค่าอธิบายในวงเล็บหลังค่า  $\alpha_4$  คือ ลักษณะการแจกแจงของประชากร ซึ่งความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของการทดสอบทั้ง 5 วิธีนี้ นำเสนอด้วยรูปที่ 4.1-4.18 เมื่อ  $\alpha = .01$  รูปที่ 4.19-4.36 เมื่อ  $\alpha = .05$  และรูปที่ 4.37-4.54 เมื่อ  $\alpha = .10$

จากแผนภาพซึ่งนำเสนอลักษณะน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของการทดสอบทั้ง 5 วิธี จะสรุปเป็นตารางแสดงจำนวนครั้งที่การทดสอบแต่ละวิธีควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ และควบคุมไม่ได้ สำหรับแต่ละการแจกแจงวิดยกรสิทธิ์ที่ควบคุมไม่ได้จะแยกเป็นกรณีที่  $t$  มากกว่า  $\alpha$  และ  $t$  น้อยกว่า  $\alpha$  ด้วย ซึ่งตารางที่ 4.1 จะสรุปผลในกรณีที่  $\alpha = .01$  และตารางที่ 4.2 สรุปผลในกรณีที่  $\alpha = .05$  และตารางที่ 4.3 สรุปผลเมื่อ  $\alpha = .10$

1.2 การนำเสนอลักษณะน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) วิดยกรสิทธิ์ลักษณะของแผนภาพที่นำเสนอนี้เหมือนกรณี 1.1 แต่แทนนอนแทนด้วยขนาดของตัวอย่างทั้งสองกลุ่มเท่านั้น ซึ่งมีอยู่ 6 ขนาดคือ (10:10) (15:15) (20:20) (5:10) (5:15) และ (5:20) โดยที่ความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) นำเสนอด้วยรูปที่ 4.55-4.57 เมื่อ  $\alpha = .01$  รูปที่ 4.58-4.60 เมื่อ  $\alpha = .05$  และรูปที่ 4.61-4.63 เมื่อ  $\alpha = .10$

จากแผนภาพซึ่งนำเสนอลักษณะน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) จะสรุปเป็นตารางแสดงจำนวนครั้งที่ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ และควบคุมไม่ได้ สำหรับแต่ละกลุ่มของตัวอย่าง คือ ขนาดของตัวอย่างเท่ากันและขนาดของตัวอย่างไม่เท่ากัน เมื่อ  $\alpha$  มีค่า .01 .05 และ .10 ซึ่งแสดงด้วยตารางที่ 4.4

2. กรณีที่ประชากรมีการแจกแจงแบบทวีคูณ การนำเสนอลักษณะน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของการทดสอบเอฟ และการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) จะนำเสนอเป็นแผนภาพเช่นเดียวกับกรณีที่ประชากรมีการแจกแจงแบบปกติ แบบยูนิฟอร์มและแบบโวลจิสติค แต่แทนนอนจะแทนความโด่งของประชากร คือ 2.8, 3.0, 3.2, 3.6 และ 4.2 สำหรับค่าอธิบายในแผนภาพจะแตกต่างจากกรณีที่ 1.1 คือจะมีค่า  $\alpha_3$  แทนที่จะเป็น  $\alpha_4$  โดย  $\alpha_3$  แทนความเบ้ของประชากร สำหรับสัญญลักษณ์ที่ใช้แทนการทดสอบเอฟและการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) จะแตกต่างกันดังนี้คือ

- o แทนการทดสอบเอฟ
- แทนการทดสอบสแควร์แรงค์ (X<sup>2</sup>)

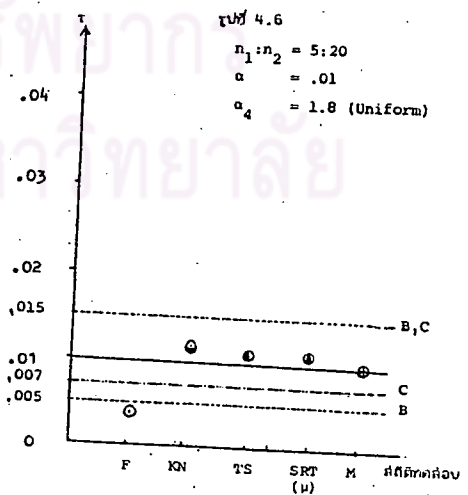
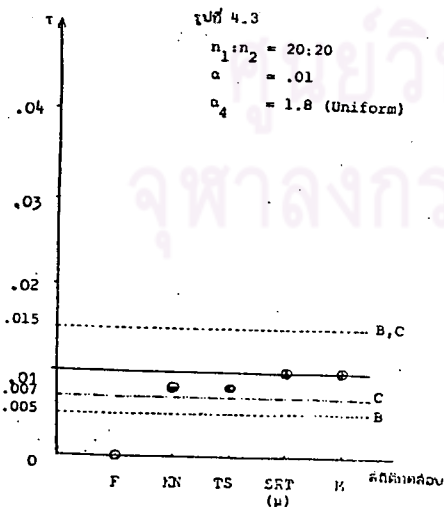
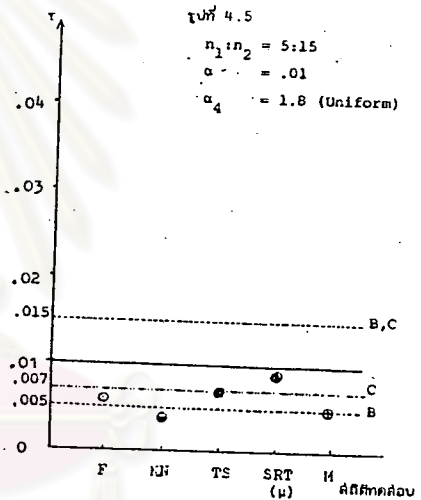
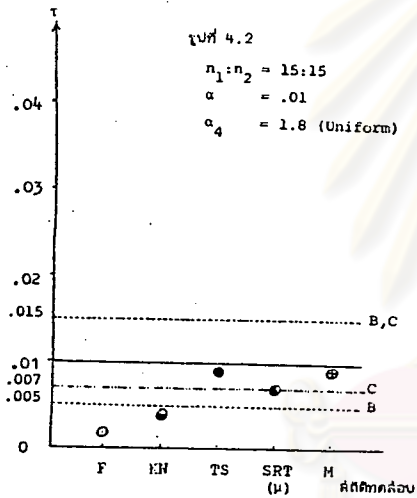
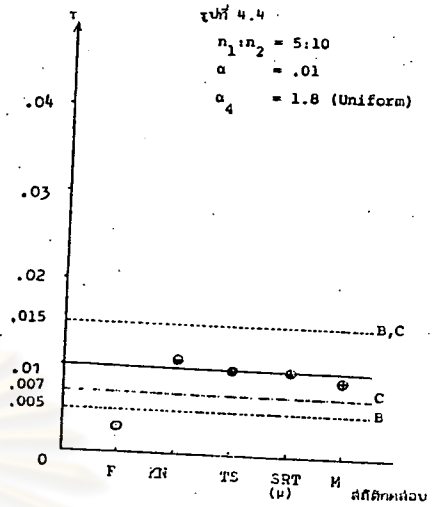
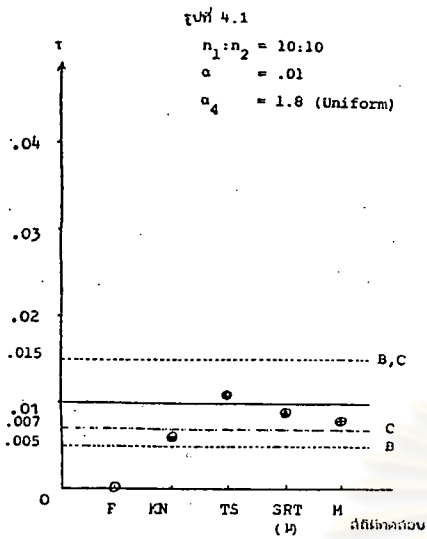
ซึ่งความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของการทดสอบเอฟและการทดสอบสแควร์แรงค์ (X<sup>2</sup>) นั้น จะนำเสนอด้วยรูปที่ 4.64-4.87 เมื่อ  $\alpha = .01$  รูปที่ 4.88-4.111 เมื่อ  $\alpha = .05$  และรูปที่ 4.112-4.135 เมื่อ  $\alpha = .10$

จากแผนภาพซึ่งนำเสนอลักษณะน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของการทดสอบเอฟและการทดสอบสแควร์แรงค์ (X<sup>2</sup>) จะสรุปเป็นตารางแสดงจำนวนครั้งที่การทดสอบทั้งสองวิธีควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ และควบคุมไม่ได้ จำแนกตามลักษณะของขนาดของตัวอย่างและความโค้งของประชากร ซึ่งตารางที่ 4.5 จะสรุปผลในกรณีที่  $\alpha = .01$  ตารางที่ 4.6 จะสรุปผลในกรณีที่  $\alpha = .05$  และตารางที่ 4.7 จะสรุปผลในกรณีที่  $\alpha = .10$

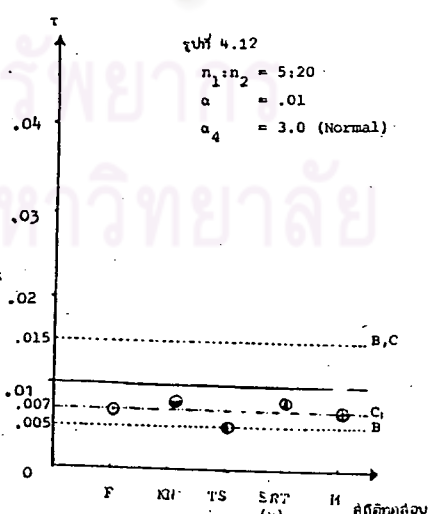
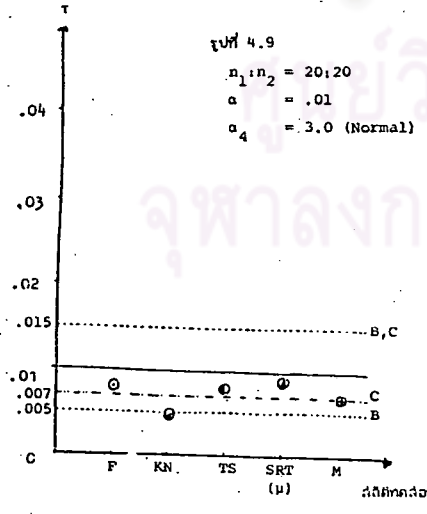
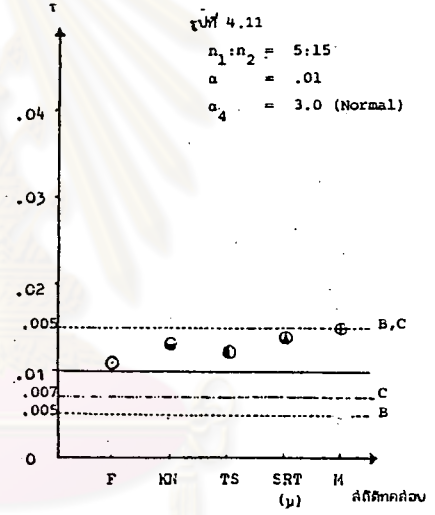
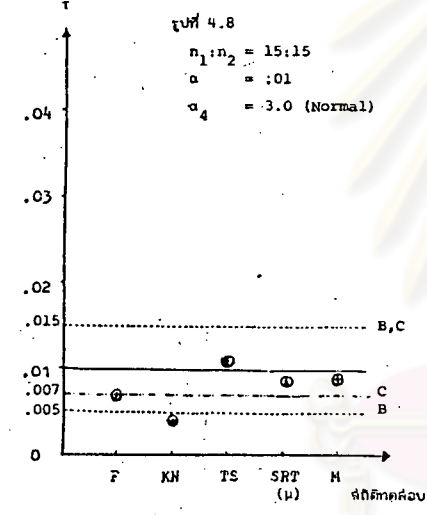
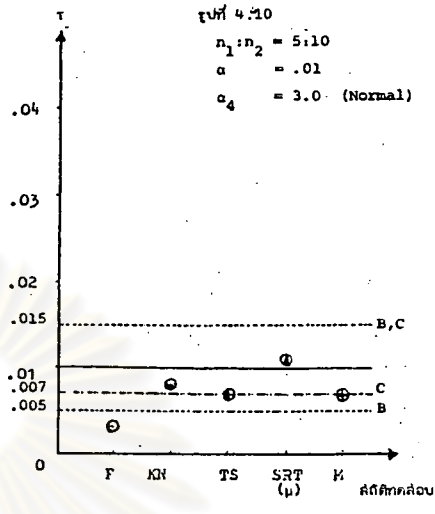
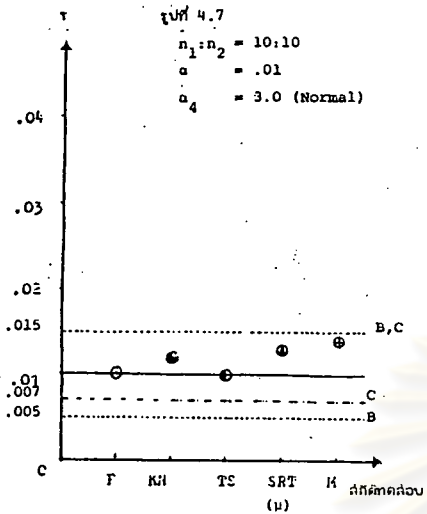
ซึ่งการนำเสนอลักษณะน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จากการทดลองของการทดสอบต่าง ๆ ซึ่งแสดงด้วยรูปที่ 4.1-4.135 เป็นดังนี้



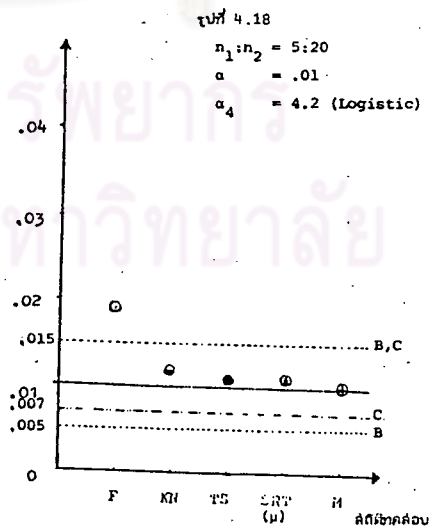
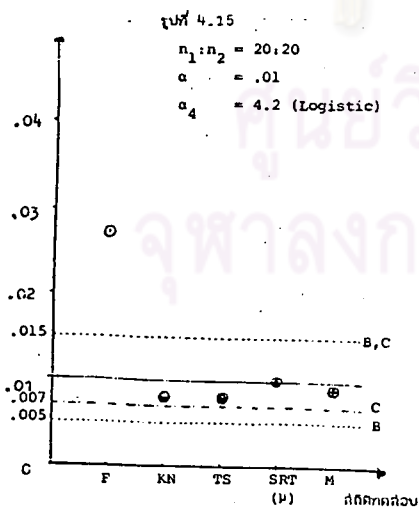
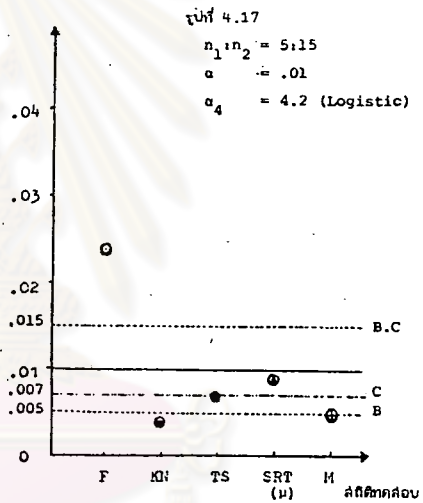
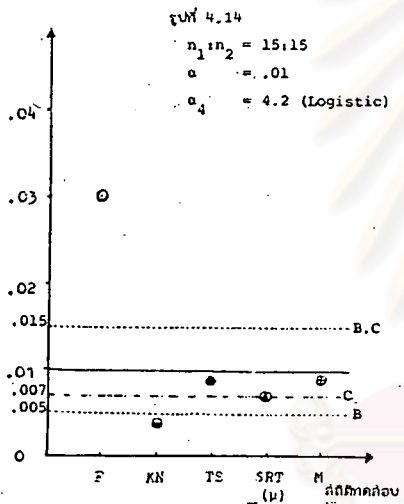
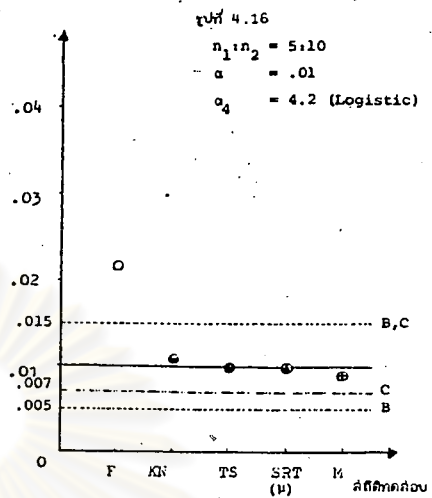
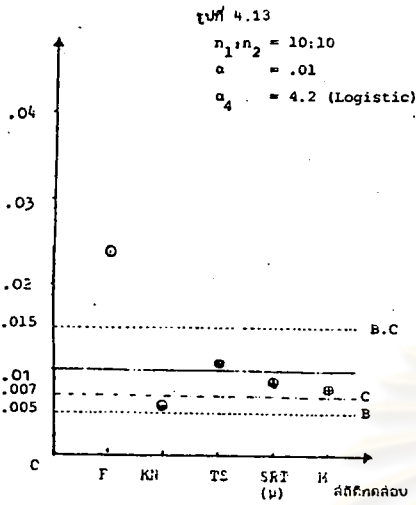
ศูนย์วิจัยและพัฒนา  
จุฬาลงกรณ์มหาวิทยาลัย







จุฬาลงกรณ์มหาวิทยาลัย





จากรูปที่ 4.1-4.18 ซึ่งแสดงค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จากการทดลอง ( $\tau$ ) ของการทดสอบเอฟ การทดสอบคล็อทซ์ การทดสอบทูกี-ซีเกล การทดสอบสแควร์แรงค์ ( $\mu$ ) และการทดสอบมัต เมื่อขนาดของตัวอย่างและลักษณะของการแจกแจงของประชากรเป็น NN(5, 10) NN(5, 15) NN(5, 20) NN(10, 10) NN(15, 15) NN(20:20) UU(5, 10) UU(5, 15) UU(5, 20) UU(10, 10) UU(15, 15) UU(20, 20) LL(5, 10) LL(5, 15) LL(5, 20) LL(10, 10) LL(15, 15) และ LL(20, 20) โดยเปรียบเทียบค่า  $\tau$  กับค่า  $\alpha$  ที่กำหนด ซึ่งมีค่า .01 ด้วยเกณฑ์ของ Cochran และเกณฑ์ของ Bradley ซึ่งสามารถสรุปจำนวนครั้งที่การทดสอบทั้ง 5 วิธีดังกล่าวควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ และควบคุมไม่ได้ดังตาราง 4.1



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตาราง 4.1 แสดงจำนวนครั้งที่การทดสอบเอฟ การทดสอบคล็อทซ์ การทดสอบทูกี-ซีเกล การทดสอบสแควร์แรงค์ ( $\mu$ ) และการทดสอบมูต์ สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 และควบคุมไม่ได้ จากการทดลองทั้งหมด 6. กรณีสำหรับแต่ละรูปแบบของการแจกแจง ที่ระดับนัยสำคัญ .01

สถิติทดสอบ	เกณฑ์ของ Cochran										เกณฑ์ของ Bradley									
	$\tau = \alpha$			$\tau < \alpha$			$\tau > \alpha$			$\tau \neq \alpha$	$\tau = \alpha$			$\tau < \alpha$			$\tau > \alpha$			$\tau \neq \alpha$
	NN	UU	LL	NN	UU	LL	NN	UU	LL		NN	UU	LL	NN	UU	LL	NN	UU	LL	
F	5	0	0	1	6	0	0	0	6	13	5	1	0	1	5	0	0	0	6	12
KN	4	3	3	2	3	3	0	0	0	8	5	4	4	1	2	2	0	0	0	5
TS	5	6	6	1	0	0	0	0	0	1	6	6	6	0	0	0	0	0	0	0
SRT ( $\mu$ )	6	6	6	0	0	0	0	0	0	0	6	6	6	0	0	0	0	0	0	0
M	6	5	5	0	1	1	0	0	0	2	6	6	6	0	0	0	0	0	0	0

NN หมายถึง การแจกแจงแบบปกติ

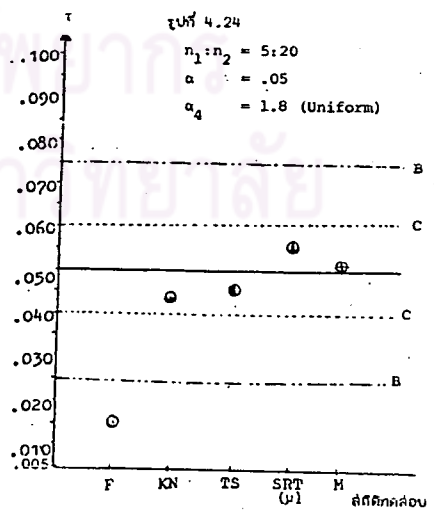
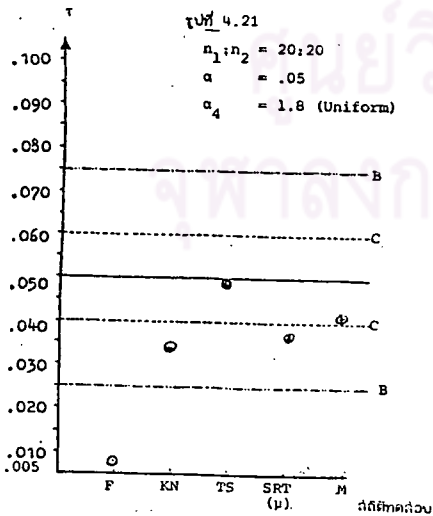
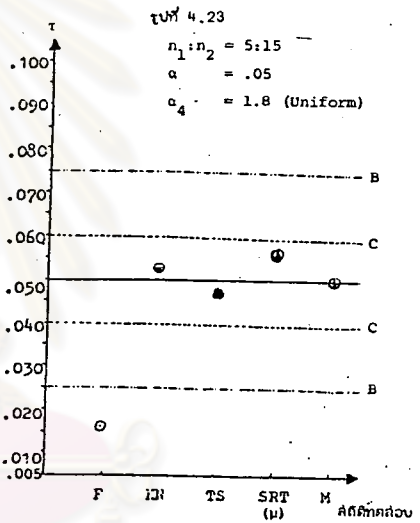
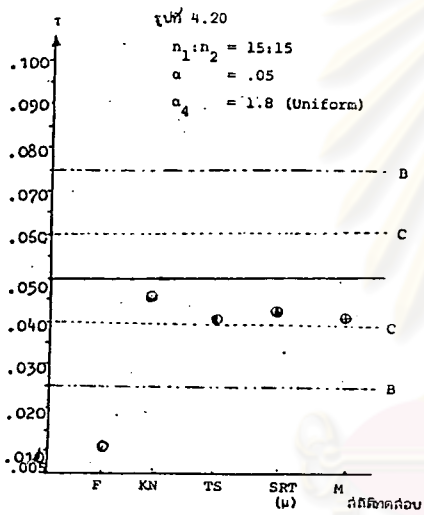
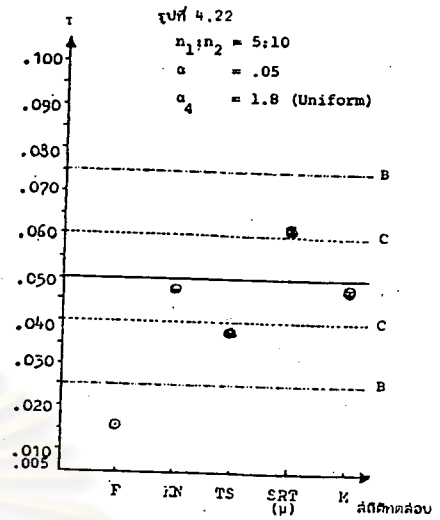
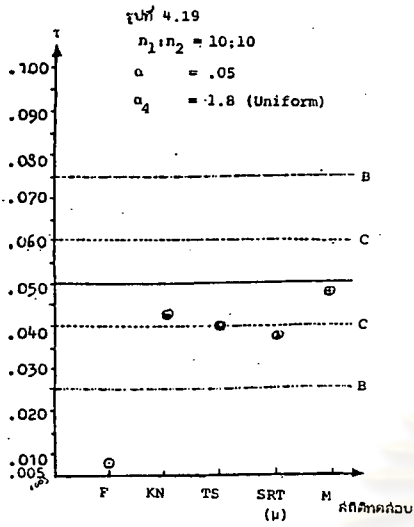
UU หมายถึง การแจกแจงแบบยูนิฟอร์ม

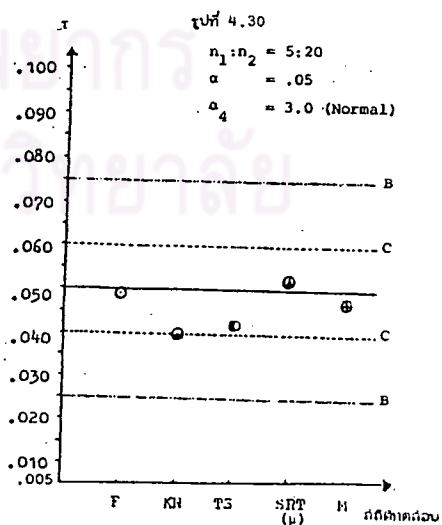
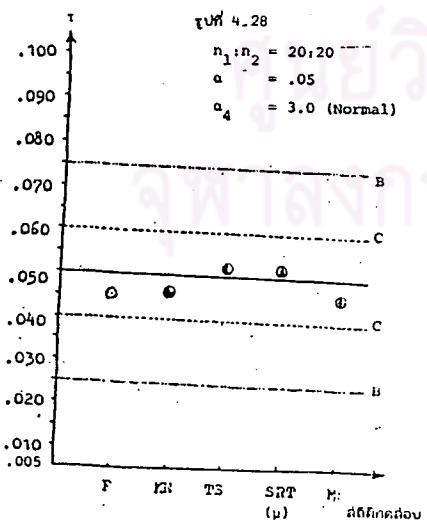
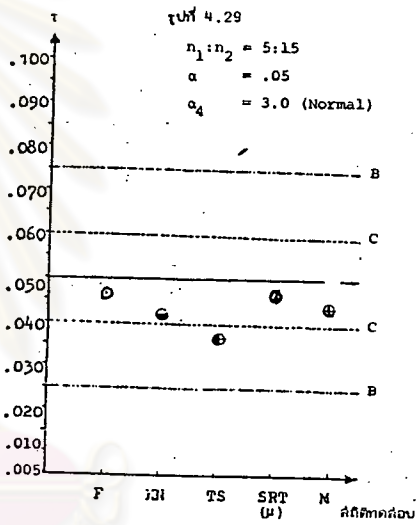
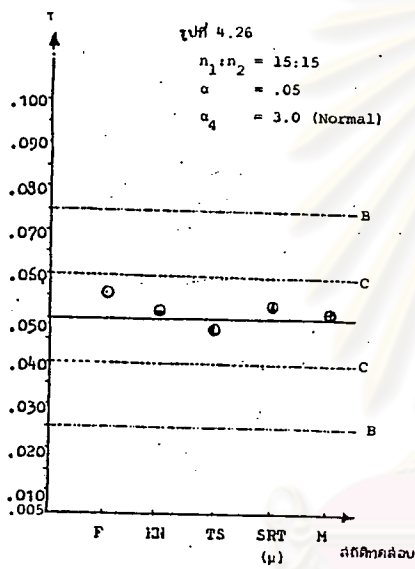
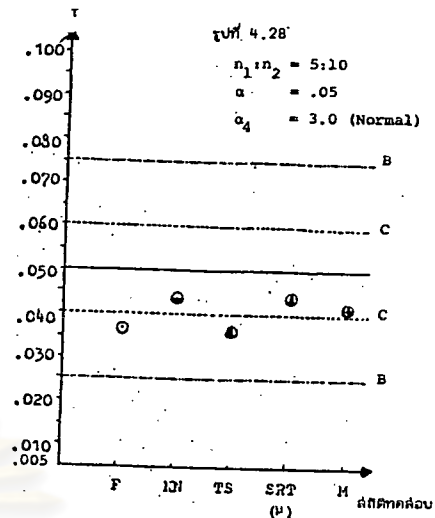
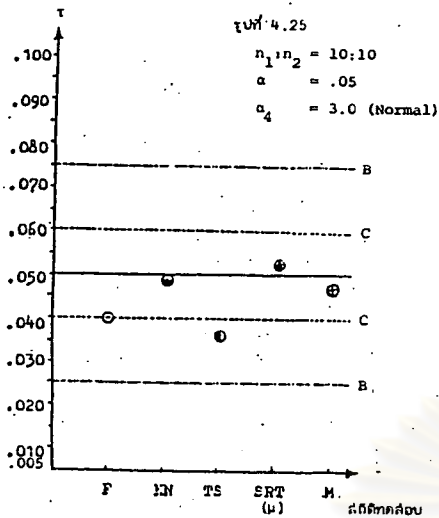
LL หมายถึง การแจกแจงแบบโลจิสติก

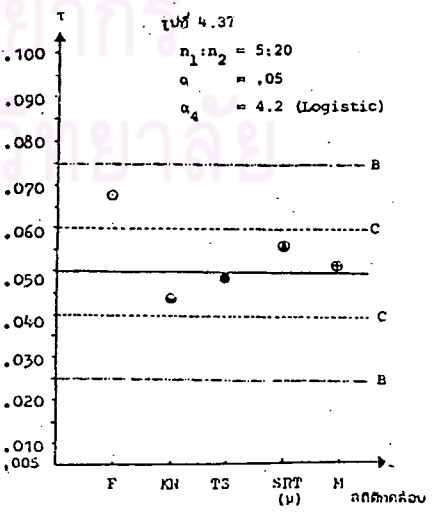
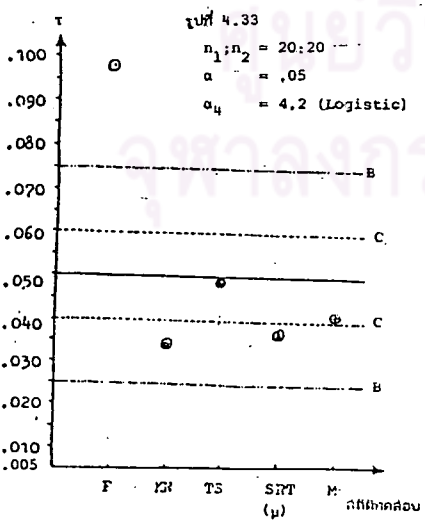
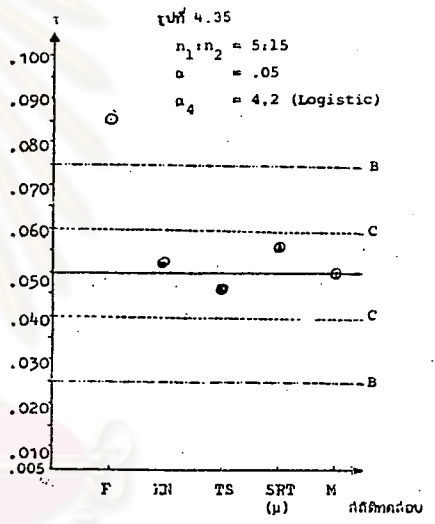
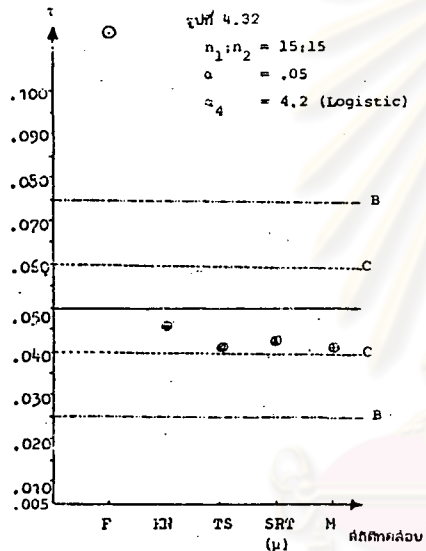
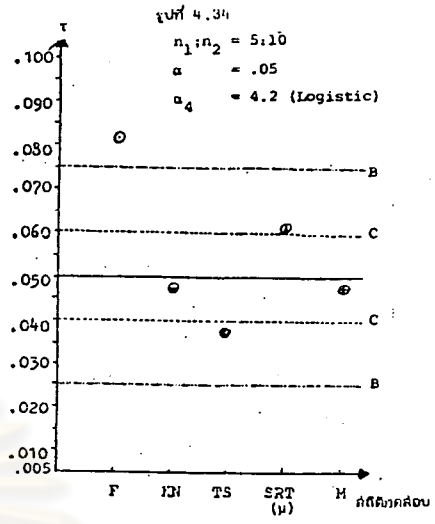
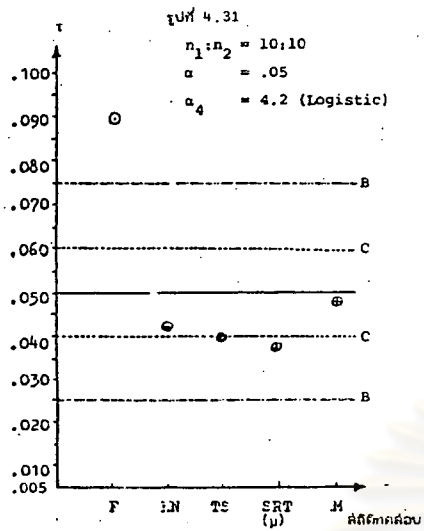
ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

#### จากตาราง 4.1 สรุปได้ดังนี้

1. การทดสอบเอฟไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เป็นส่วนมากเมื่อปรากฏการแจกแจงแบบยูนิฟอร์มและแบบโลจิสติก ไม่ว่าจะใช้เกณฑ์ของ Bradley หรือ Cochran
2. การทดสอบแบบนอนพาราเมตริกส่วนมากแล้วควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี ไม่ว่าจะใช้เกณฑ์ของ Bradley หรือเกณฑ์ของ Cochran โดยที่การทดสอบสแควร์แรงค์ (U) สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดีที่สุด
3. สภาวะที่ควบคุมความคลาดเคลื่อนประเภทที่ 1 ไม่ได้ สำหรับการทดสอบแบบนอนพาราเมตริกนั้นจะมีค่า  $\tau$  น้อยกว่า  $\alpha$  ในขณะที่การทดสอบเอฟจะมีค่า  $\tau$  มากกว่า  $\alpha$  เมื่อปรากฏการแจกแจงแบบโลจิสติก และ  $\tau < \alpha$  เมื่อปรากฏการแจกแจงแบบยูนิฟอร์ม









จากรูปที่ 4.19-4.36 ซึ่งแสดงค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จากการทดลอง ( $\tau$ ) ของการทดสอบเอฟ การทดสอบคล็อทซ์ การทดสอบทูกี-ซีเกล การทดสอบสแควร์แรงค์ ( $\mu$ ) และการทดสอบมัต เมื่อขนาดของตัวอย่างและลักษณะของการแจกแจงของประชากรเป็น NN(5, 10) NN(5, 15) NN(5, 20) NN(10, 10) NN(15, 15) NN(20, 20) UU(5, 10) UU(5, 15) UU(5, 20) UU(10, 10) UU(15, 15) UU(20, 20) LL(5, 10) LL(5, 15) LL(5, 20) LL(10, 10) LL(15, 15) และ LL(20, 20) โดยเปรียบเทียบค่า  $\tau$  กับค่า  $\alpha$  ที่กำหนด ซึ่งมีค่า .05 ด้วยเกณฑ์ของ Cochran และเกณฑ์ของ Bradley ซึ่งสามารถสรุปจำนวนครั้งที่การทดสอบทั้ง 5 วิธีดังกล่าวควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ และควบคุมไม่ได้ดังตาราง 4.2



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตาราง 4.2 แสดงจำนวนครั้งที่การทดสอบเอฟ การทดสอบคลอว์ซีย์ การทดสอบทูกี-ซีเกล การทดสอบสแควร์แรงค์ ( $\mu$ ) และการทดสอบมัด สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ และควบคุมไม่ได้จากการทดลองทั้งหมด 6 กรณี สำหรับแต่ละรูปแบบของการ แยกแยะที่ระดับนัยสำคัญ .05

สถิติทดสอบ	เกณฑ์ของ Cochran										เกณฑ์ของ Brandley									
	$\tau = \alpha$			$\tau < \alpha$			$\tau > \alpha$			$\tau \neq \alpha$	$\tau = \alpha$			$\tau > \alpha$			$\tau > \alpha$	$\tau \neq \alpha$		
	NN	UU	LL	NN	UU	LL	NN	UU	LL		NN	UU	LL	NN	UU	LL	NN	UU	LL	
F	5	0	0	1	6	0	0	0	6	13	6	0	1	0	6	0	0	0	5	11
KN	6	5	5	0	1	1	0	0	0	2	6	6	6	0	0	0	0	0	0	0
TS	3	5	5	3	1	1	0	0	0	5	6	6	6	0	0	0	0	0	0	0
SRT ( $\mu$ )	6	3	3	0	3	3	0	0	0	6	6	6	6	0	0	0	0	0	0	0
M	6	6	6	0	0	0	0	0	0	0	6	6	6	0	0	0	0	0	0	0

NN หมายถึง การแยกแยะแบบปกติ

UU หมายถึง การแยกแยะแบบยูนิฟอร์ม

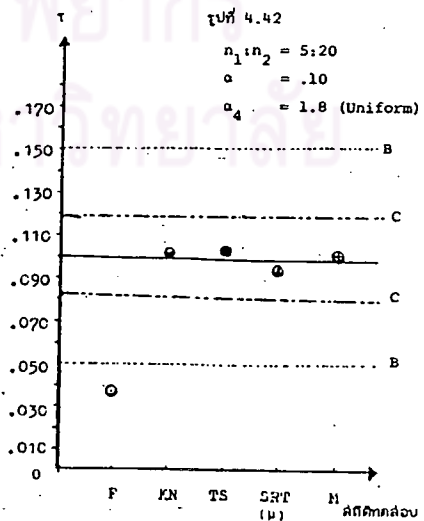
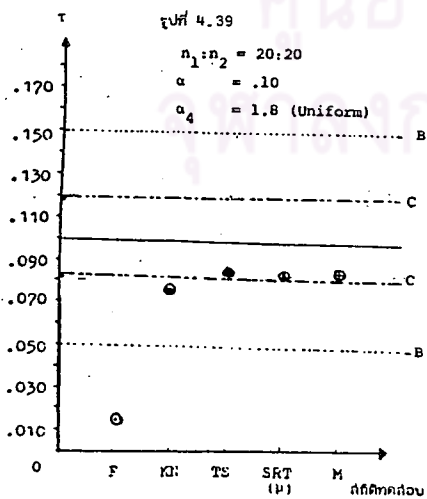
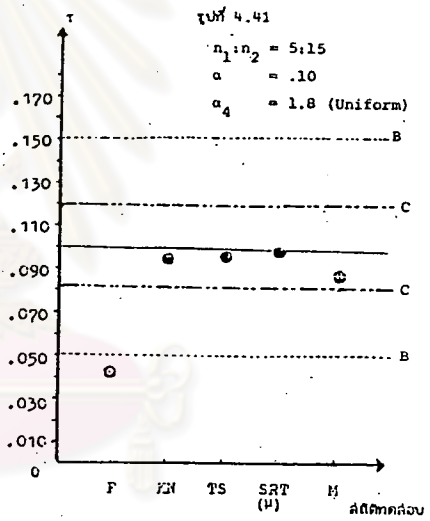
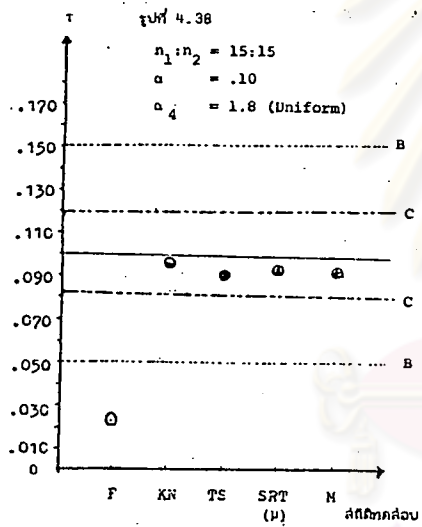
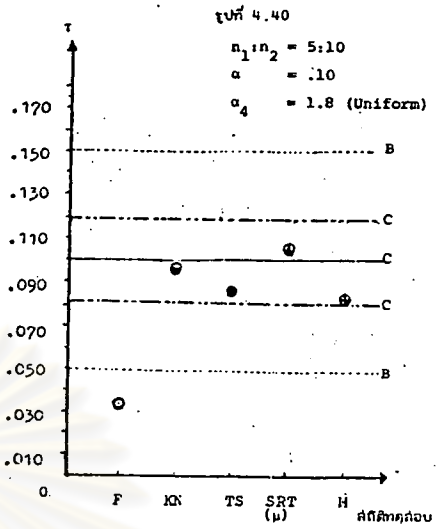
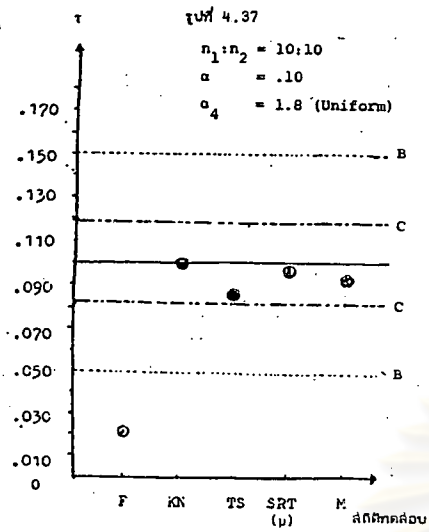
LL หมายถึง การแยกแยะแบบโลจิสติก

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

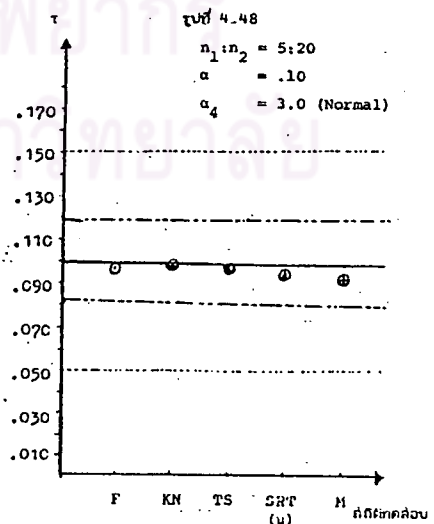
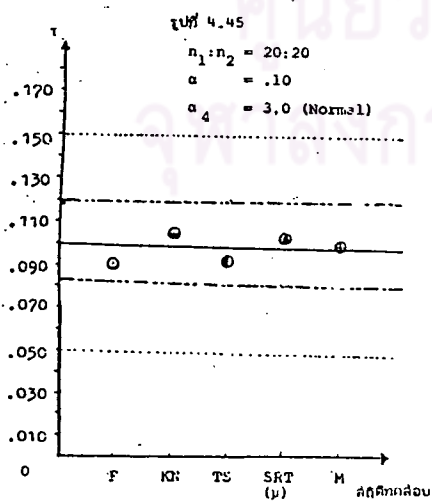
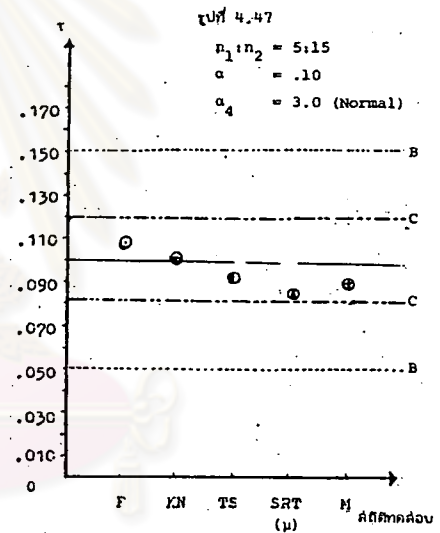
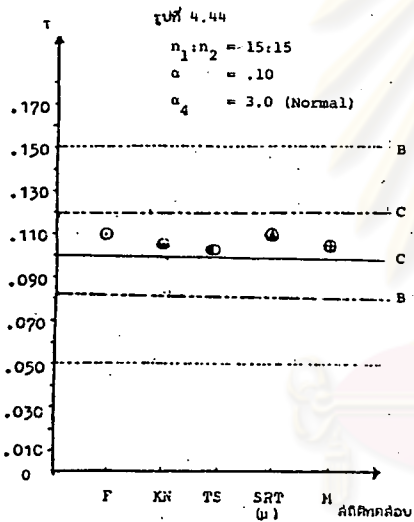
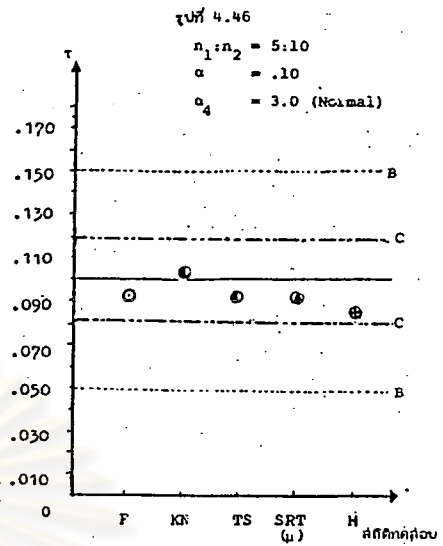
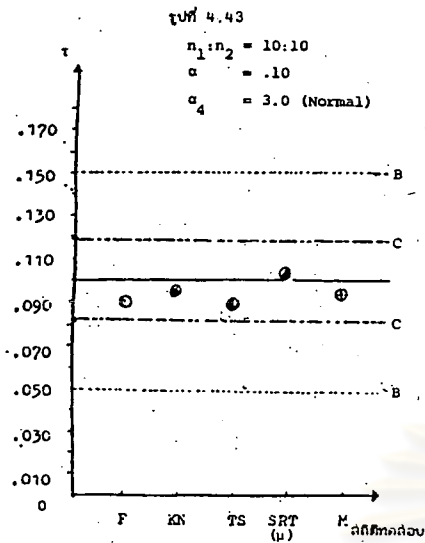
#### สรุปผลจากตาราง 4.2

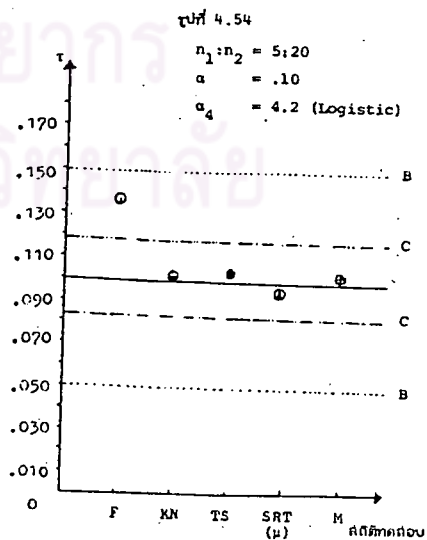
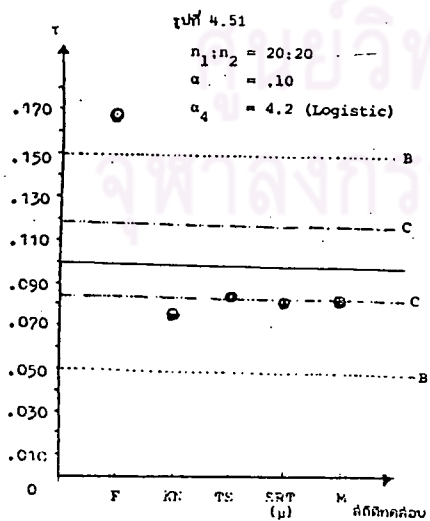
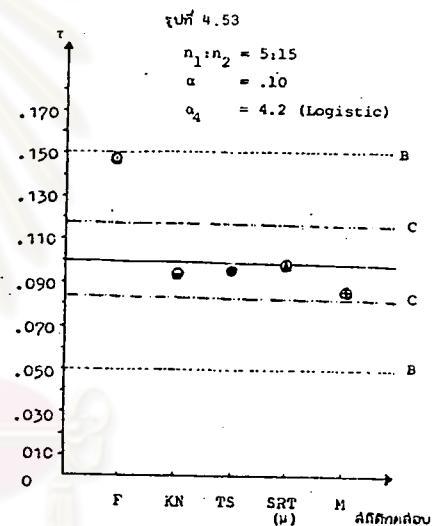
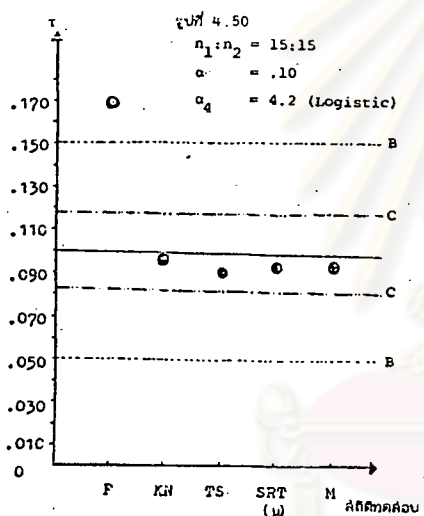
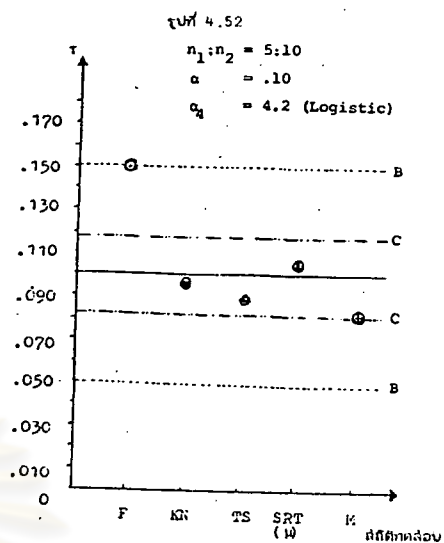
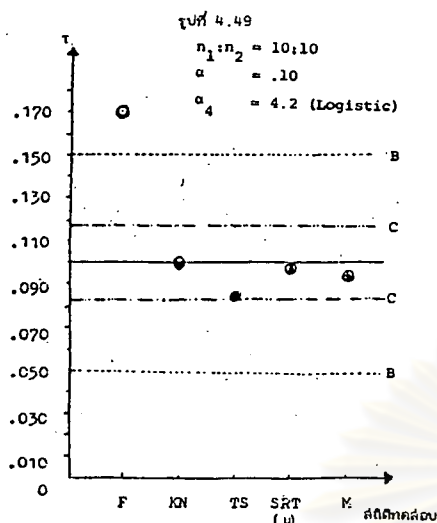
1. การทดสอบเอฟไม่สามารควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ไม่ว่าจะใช้เกณฑ์ของ Bradley หรือ Cochran เมื่อประชากรมีการแจกแจงแบบยูนิฟอร์มและแบบโลจิสติก
2. การทดสอบแบบนอนพาราเมตริก ส่วนมากแล้วสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี ไม่ว่าจะใช้เกณฑ์ของ Bradley หรือเกณฑ์ของ Cochran โดยที่การทดสอบมักสามารควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดีที่สุด
3. สภาวะที่ควบคุมไม่ได้สำหรับการทดสอบแบบนอนพาราเมตริก จะมีค่า  $\tau$  น้อยกว่า  $\alpha$  แต่การทดสอบเอฟจะให้ค่า  $\tau$  มากกว่า  $\alpha$  เมื่อประชากรมีการแจกแจงแบบโลจิสติก และมีค่า  $\tau$  น้อยกว่า  $\alpha$  เมื่อประชากรมีการแจกแจงแบบยูนิฟอร์ม

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย



ศูนย์วิทยทรัพยากร  
 จุฬาลงกรณ์มหาวิทยาลัย







จากรูปที่ 4.37-4.54 ซึ่งแสดงค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จากการทดลอง ( $\tau$ ) ของการทดสอบเอฟ การทดสอบคล็อทซ์ การทดสอบทูกี-ซีเกล การทดสอบสแควร์แรงค์ ( $\mu$ ) และการทดสอบมัต เมื่อขนาดของตัวอย่างและลักษณะของการแจกแจงของประชากรเป็น NN(5, 10) NN(5, 15) NN(5, 20) NN(10, 10) NN(15, 15) NN(20, 20) UU(5, 10) UU(5, 15) UU(5, 20) UU(10, 10) UU(15, 15) UU(20, 20) LL(5, 10) LL(5, 15) LL(5, 20) LL(10, 10) LL(15, 15) และ LL(20, 20) โดยเปรียบเทียบค่า  $\tau$  กับค่า  $\alpha$  ที่กำหนด ซึ่งมีค่า .10 ด้วยเกณฑ์ของ Cochran และเกณฑ์ของ Bradley ซึ่งสามารถสรุปจำนวนครั้งที่การทดสอบทั้ง 5 วิธีดังกล่าวควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ และควบคุมไม่ได้ดังตาราง 4.3



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตาราง 4.3 แสดงจำนวนครั้งที่การทดสอบเอฟ การทดสอบคล้อย การทดสอบทูกี-ซีเกล การทดสอบสแควร์แรงค์ ( $\mu$ ) และการทดสอบมัด

ควบคุมความคลาดเคลื่อนประเภทที่ 1, ได้และควบคุมไม่ได้ จากการทดลองทั้งหมด 6 กรณี สำหรับแต่ละรูปแบบของการแจกแจงที่ระดับนัยสำคัญ .10

สถิติทดสอบ	เกณฑ์ของ Cochran										เกณฑ์ของ Bradley										
	$\tau = \alpha$			$\tau < \alpha$			$\tau > \alpha$			$\tau \neq \alpha$	$\tau = \alpha$			$\tau < \alpha$			$\tau > \alpha$			$\tau \neq \alpha$	
	NN	UU	LL	NN	UU	LL	NN	UU	LL		NN	UU	LL	NN	UU	LL	NN	UU	LL		
F	6	0	0	0	6	0	0	0	6	12	6	0	3	0	6	0	0	0	0	3	9
KN	6	5	5	0	1	1	0	0	0	2	6	6	6	0	0	0	0	0	0	0	0
TS	6	6	6	0	0	0	0	0	0	0	6	6	6	0	0	0	0	0	0	0	0
SRT ( $\mu$ )	6	6	6	0	0	0	0	0	0	0	6	6	6	0	0	0	0	0	0	0	0
M	6	6	6	0	0	0	0	0	0	0	6	6	6	0	0	0	0	0	0	0	0

NN หมายถึง การแจกแจงแบบปกติ

UU หมายถึง การแจกแจงแบบยูนิฟอร์ม

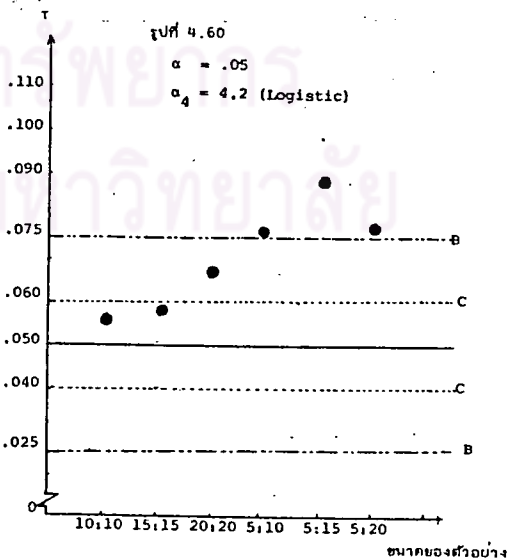
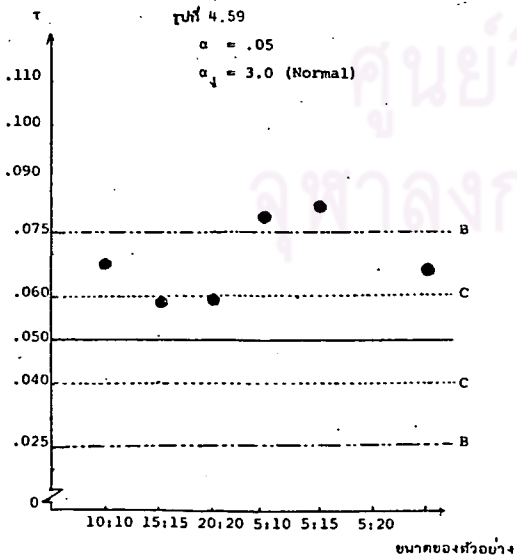
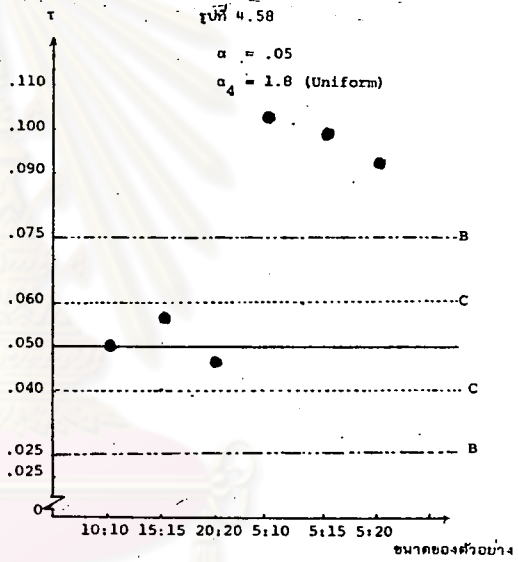
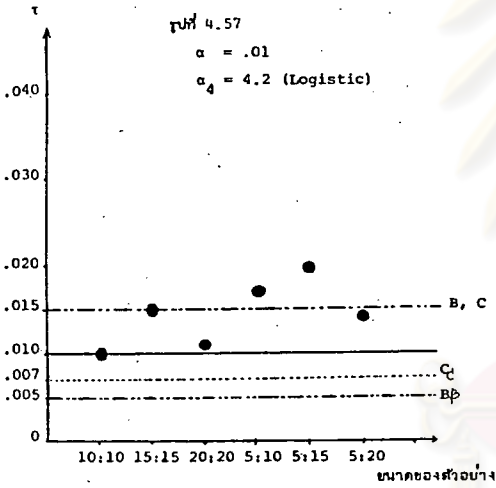
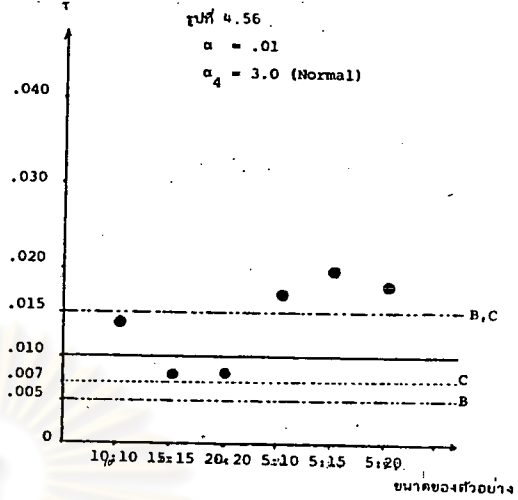
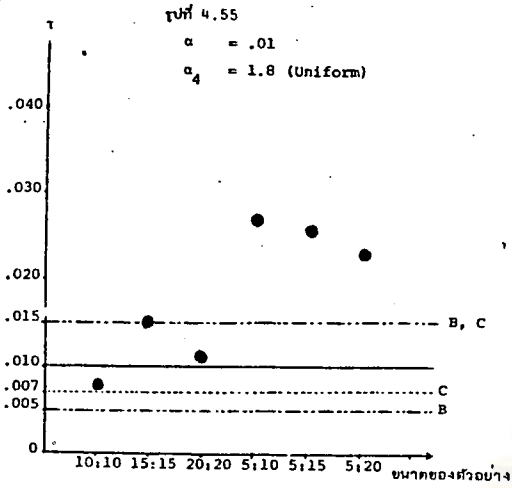
LL หมายถึง การแจกแจงแบบโลจิสติก

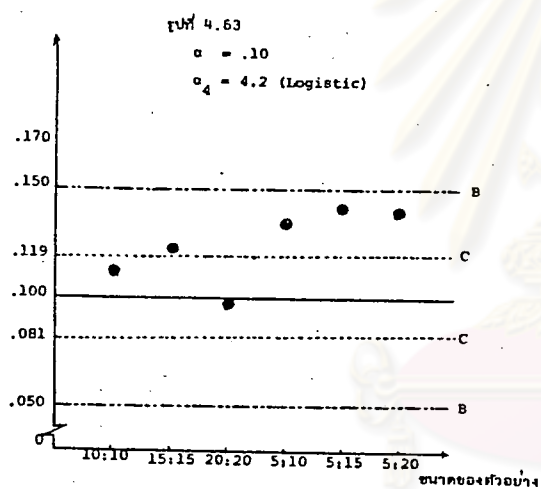
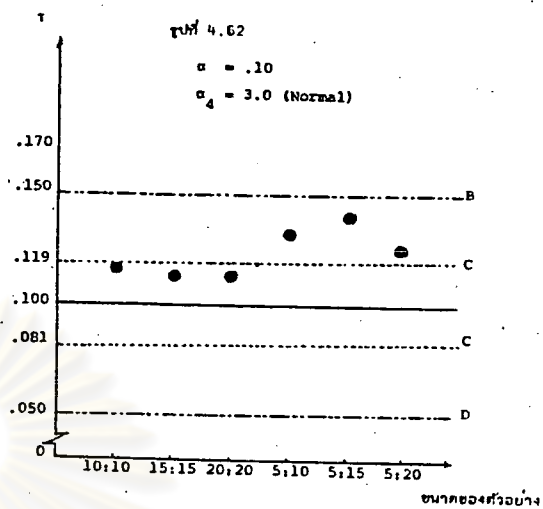
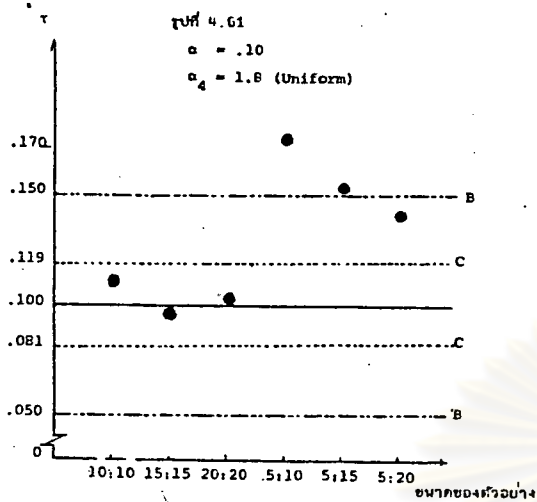
ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

## สรุปผลจากตาราง 4.3

1. การทดสอบเอฟไม่สำเร็จควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ เมื่อการแจกแจงของประชากรเป็นแบบยูนิฟอร์ม และโดยส่วนใหญ่แล้วควบคุมความคลาดเคลื่อนประเภทที่ 1 ไม่ได้เมื่อประชากรมีการแจกแจงแบบโลจิสติก
2. การทดสอบแบบนอนพาราเมตริก ส่วนมากควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี ไม่ว่าจะใช้สถิติของ Bradley หรือสถิติของ Cochran
3. สถิติที่ควบคุมไม่ได้ ของการทดสอบแบบนอนพาราเมตริกจะมีค่า  $\tau$  น้อยกว่า  $\alpha$  และการทดสอบเอฟจะมีค่า  $\tau$  น้อยกว่า  $\alpha$  เมื่อประชากรมีการแจกแจงแบบยูนิฟอร์ม และ  $\tau$  มากกว่า  $\alpha$  เมื่อประชากรมีการแจกแจงแบบโลจิสติก

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย





ศูนย์วิทยทรัพยากร  
 จุฬาลงกรณ์มหาวิทยาลัย

จากรูปที่ 4.55-4.63 ซึ่งแสดงค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จากการทดลอง ( $\tau$ ) ของการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) เมื่อประชากรมีการแจกแจงแบบปกติ แบบยูนิฟอร์ม และแบบโลจิสติก โดยเปรียบเทียบค่า  $\tau$  กับค่า  $\alpha$  ที่กำหนด ซึ่งมีค่า .01 .05 และ .10 ด้วยเกณฑ์ของ Bradley และเกณฑ์ของ Cochran สามารถสรุปจำนวนครั้งที่การทดสอบสแควร์แรงค์ ( $\bar{X}$ ) ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดังตาราง 4.4



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย



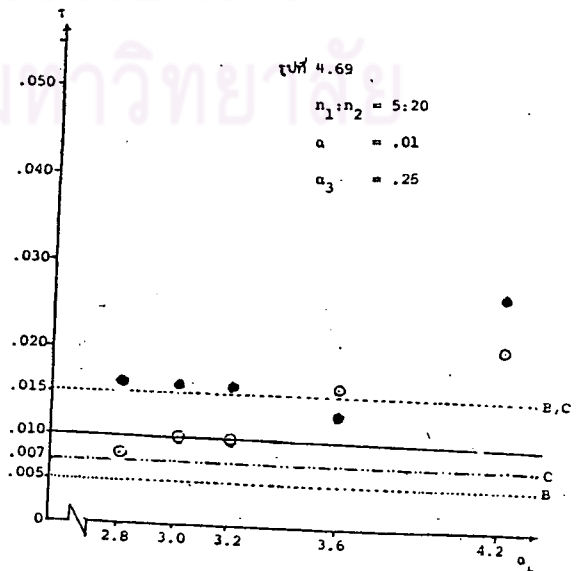
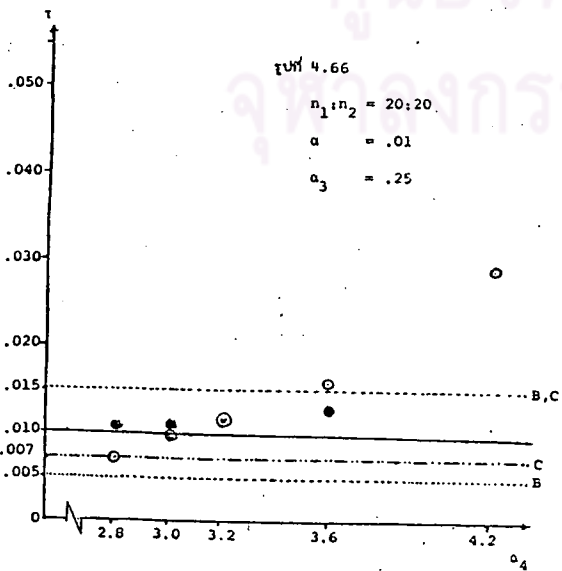
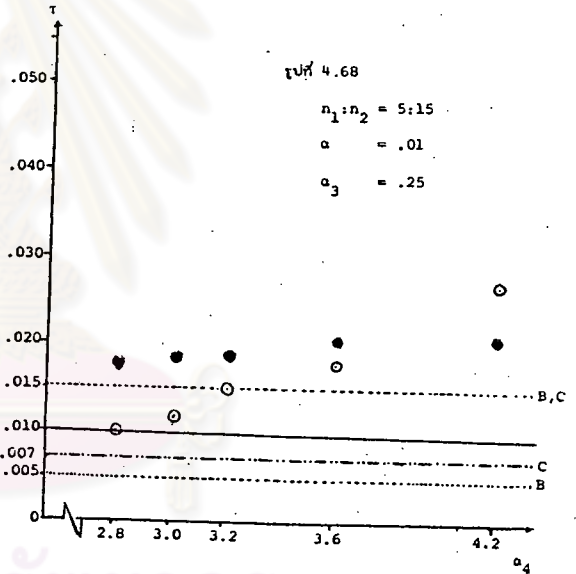
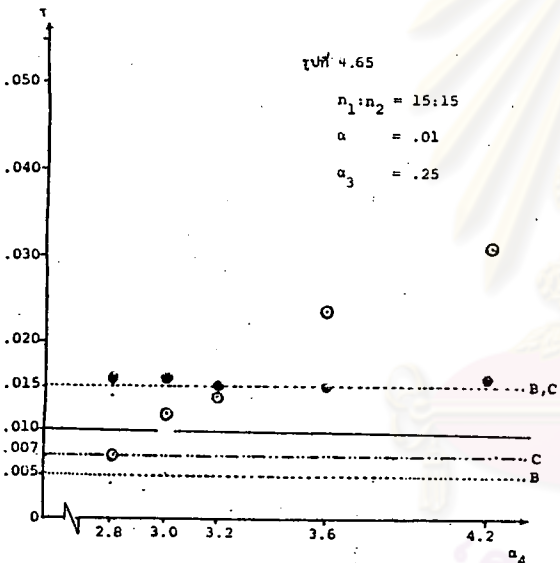
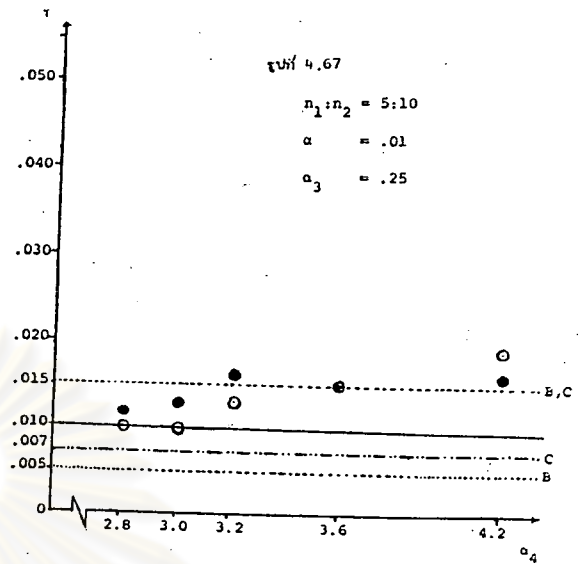
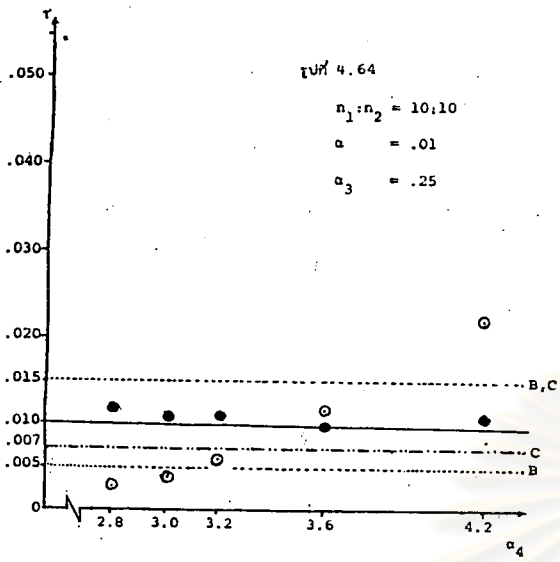
ตาราง 4.4 แสดงจำนวนครั้งที่การทดสอบสแควร์แรงค์ ( $\bar{X}$ ) ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้และควบคุมไม่ได้ จากการทดลองทั้งหมด 9 กรณี สำหรับแต่ละกลุ่มของตัวอย่างจำแนกตามระดับนัยสำคัญที่กำหนด

ระดับนัยสำคัญ	ขนาดของตัวอย่าง	เกณฑ์ของ Cochran				เกณฑ์ของ Bradley			
		$\tau = \alpha$	$\tau < \alpha$	$\tau > \alpha$	$\tau \neq \alpha$	$\tau = \alpha$	$\tau < \alpha$	$\tau > \alpha$	$\tau \neq \alpha$
.01	(10:10) (15:15) (20:20)	9	0	0	0	9	0	0	0
	(5:10) (5:15) (5:20)	1	0	8	8	1	0	8	8
.05	(10:10) (15:10) (20:20)	7	0	2	2	9	0	0	0
	(5:10) (5:15) (5:20)	0	0	9	9	1	0	8	8
.10	(10:10) (15:15) (20:20)	8	0	1	0	9	0	0	0
	(5:10) (5:15) (5:20)	0	0	9	9	7	0	2	2

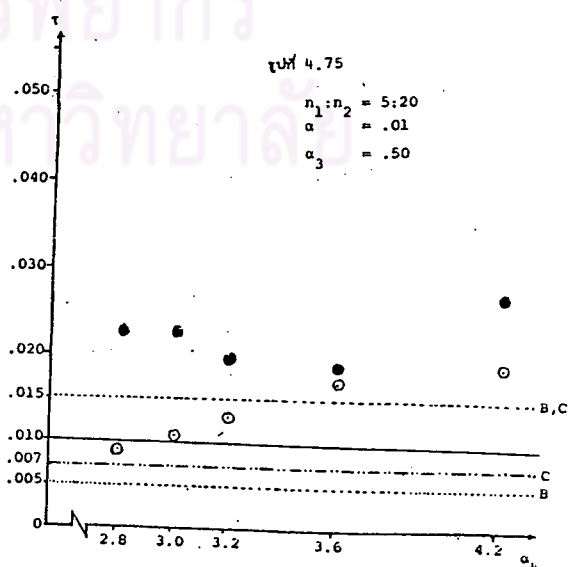
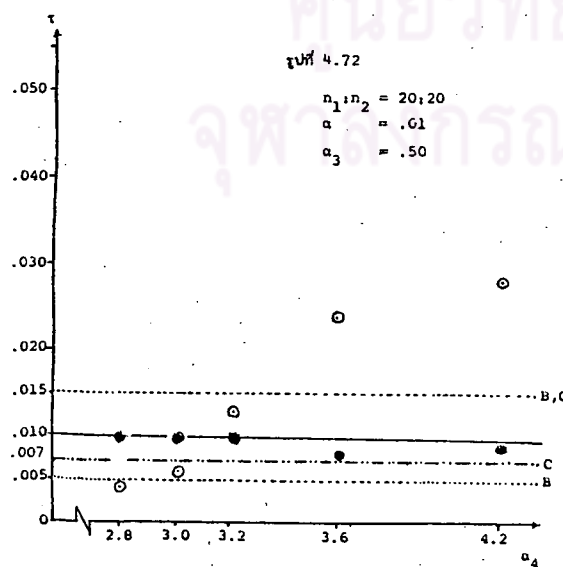
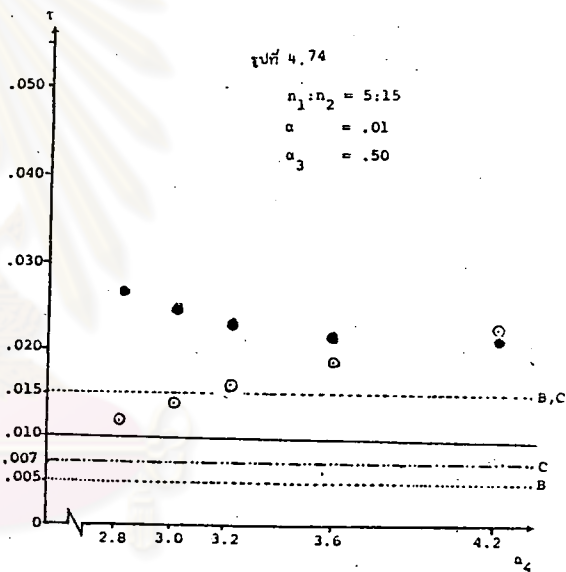
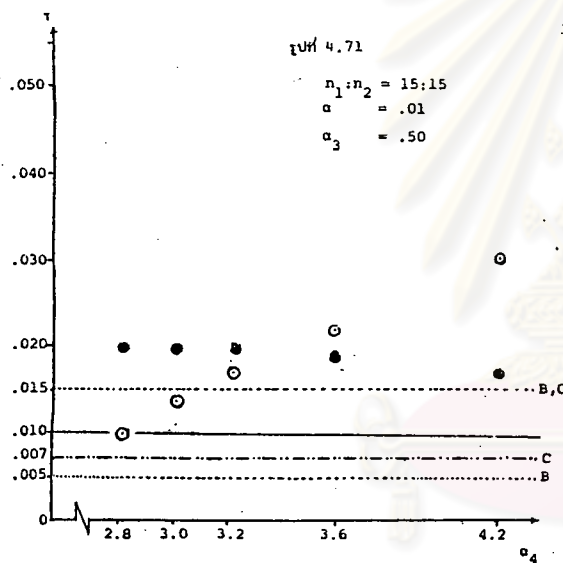
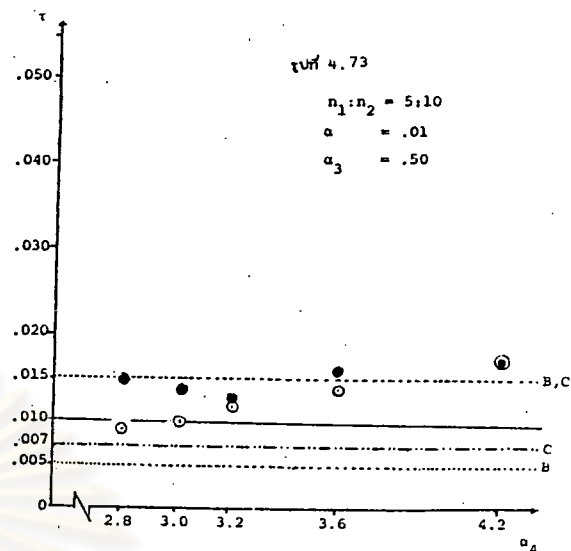
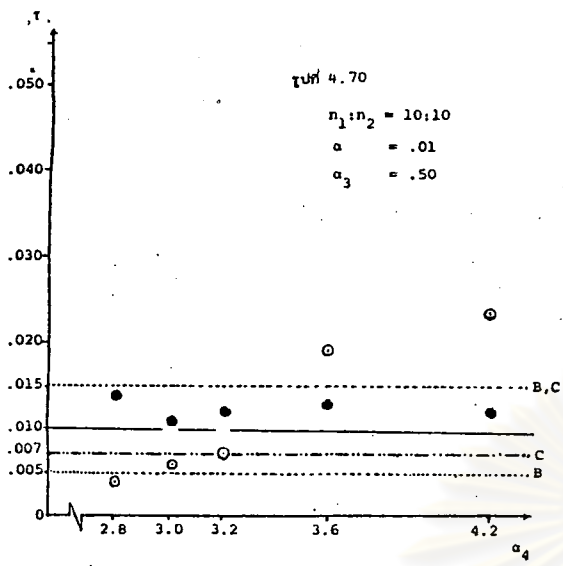
จากตารางที่ 4.4 สรุปได้ดังนี้

1. การทดสอบสแควร์แรงค์ ( $\bar{X}$ ) ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี เมื่อขนาดของตัวอย่างเท่ากัน โดยควบคุมได้ทั้งหมดทุก ๆ ค่าของ  $\alpha$  ที่กำหนดเมื่อใช้เกณฑ์ของ Bradley และควบคุมได้ 9 ครั้งเมื่อ  $\alpha = .01$  7 ครั้งเมื่อ  $\alpha = .05$  และ 8 ครั้งเมื่อ  $\alpha = .10$  เมื่อใช้เกณฑ์ของ Cochran สำหรับกรณีที่ยีนขนาดของตัวอย่างไม่เท่ากัน การทดสอบสแควร์แรงค์ ( $\bar{X}$ ) ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้น้อย กล่าวคือควบคุมได้ 1 ครั้ง สำหรับ  $\alpha = .10$  ควบคุมไม่ได้ทั้งหมดเมื่อ  $\alpha = .05$  และ  $.10$  เมื่อใช้เกณฑ์ของ Cochran และควบคุมได้ 1 ครั้งสำหรับ  $\alpha = .01$  และ  $.05$  และควบคุมได้ 7 ครั้งสำหรับ  $\alpha = .10$  เมื่อใช้เกณฑ์ของ Bradley
2. ลักษณะที่ควบคุมความคลาดเคลื่อนประเภทที่ 1 ไม่ได้ของการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) คือ ค่า  $\tau$  มากกว่าค่า  $\alpha$  ที่กำหนด

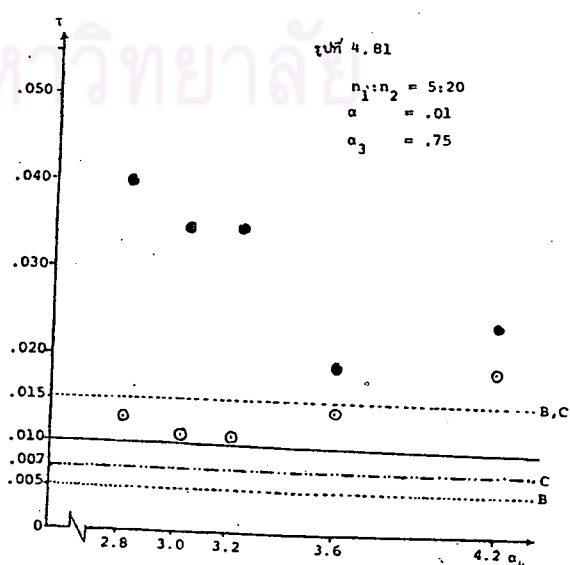
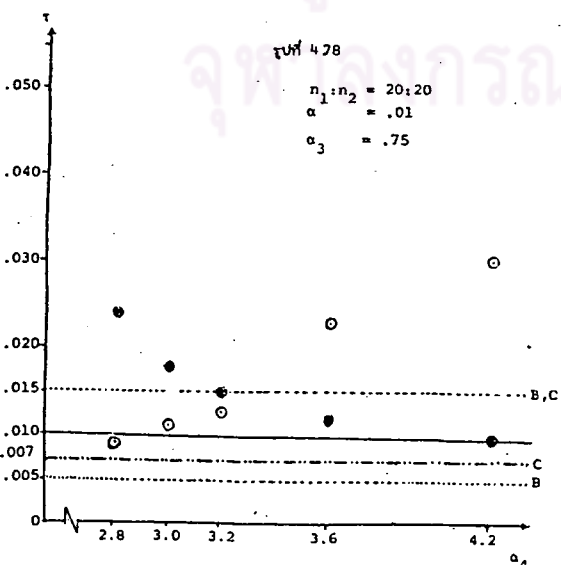
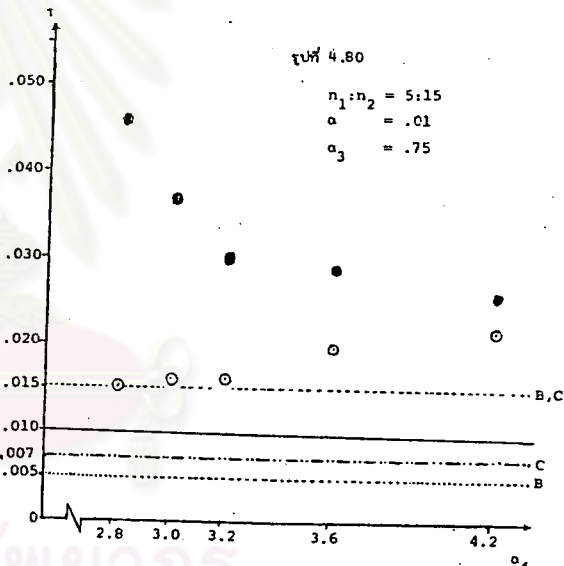
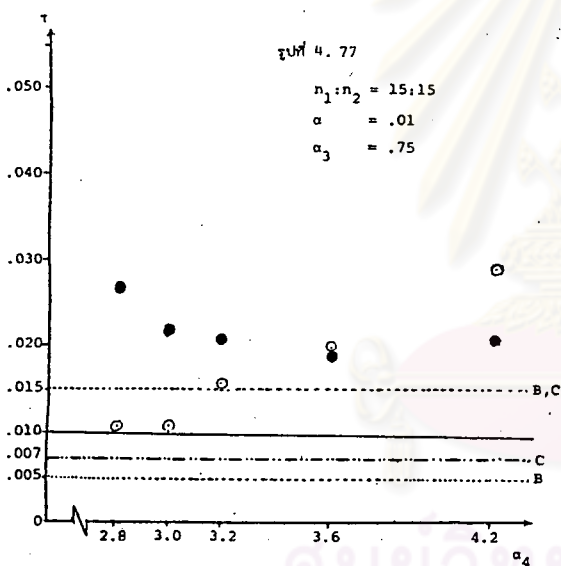
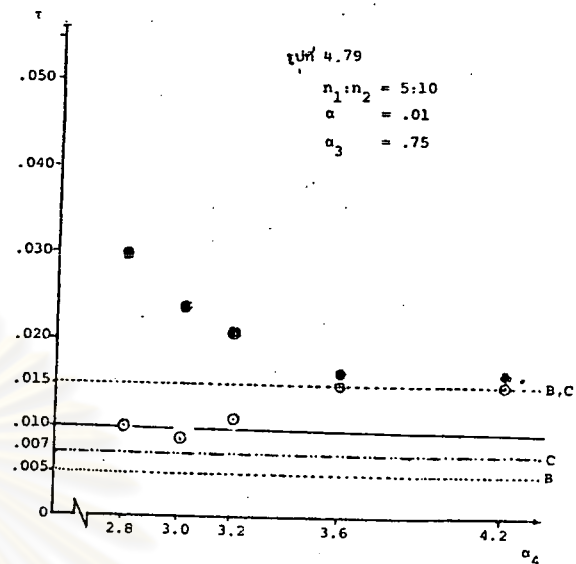
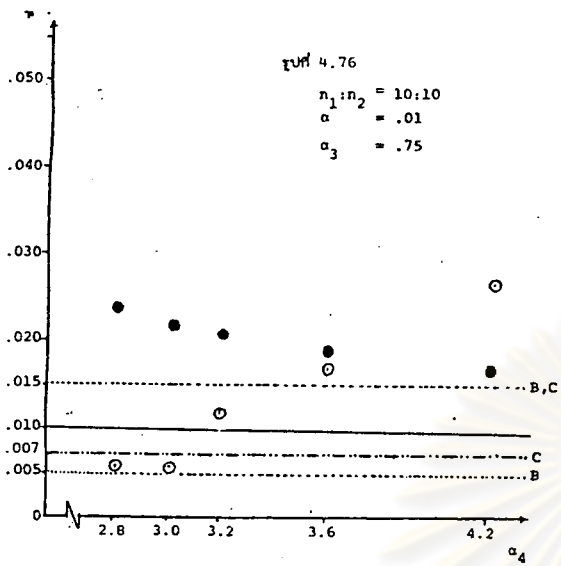
ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย



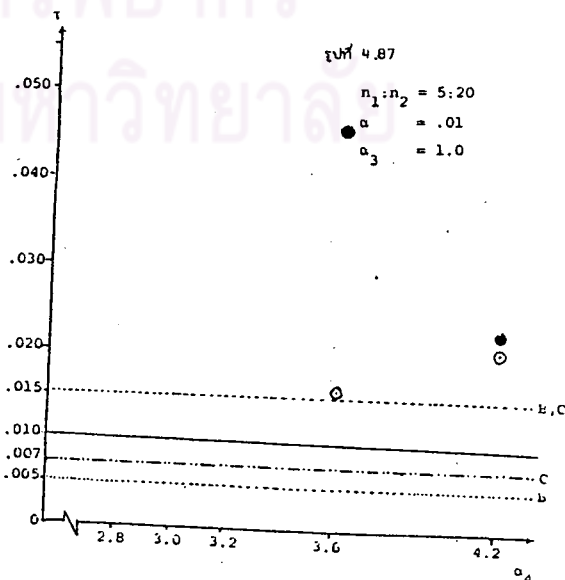
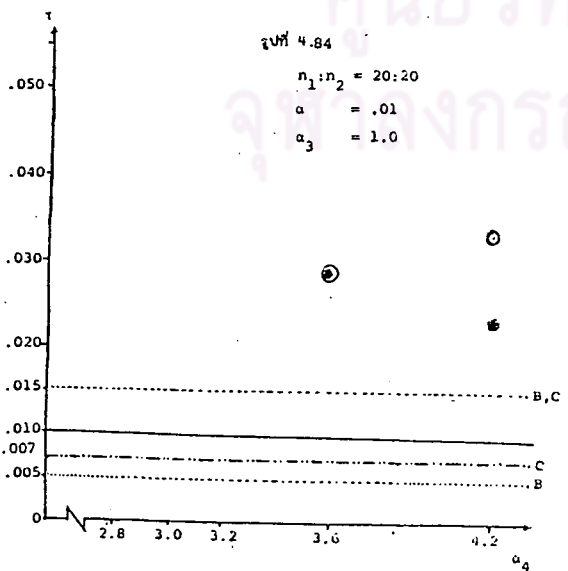
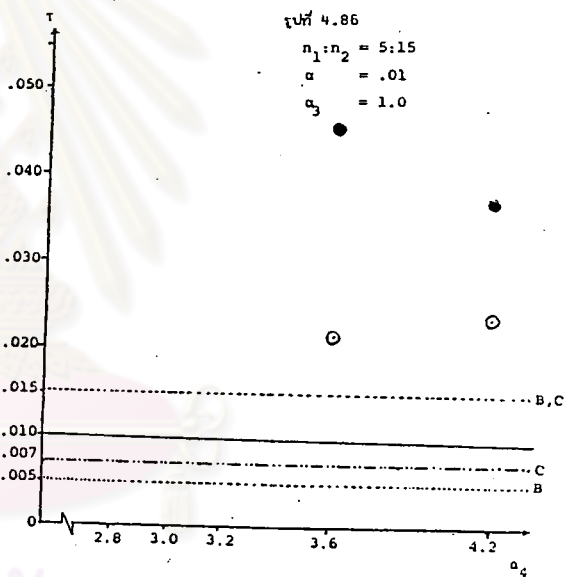
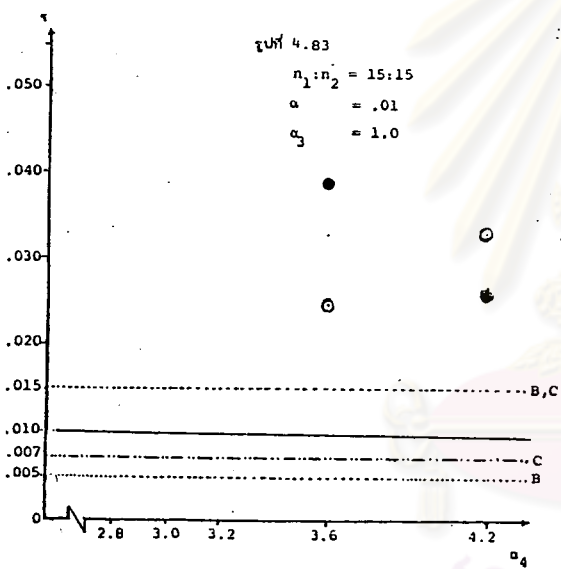
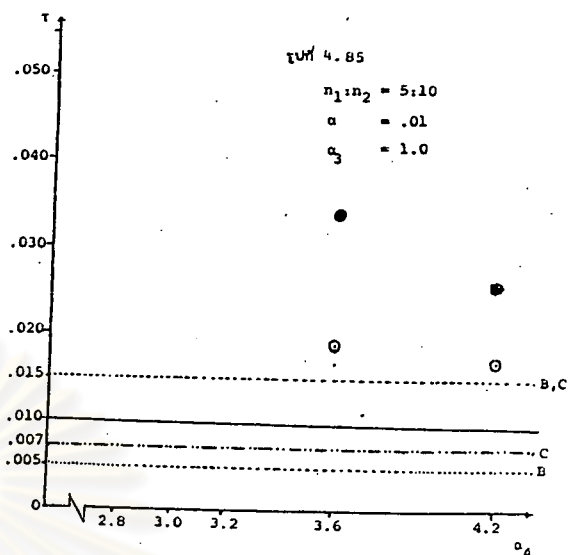
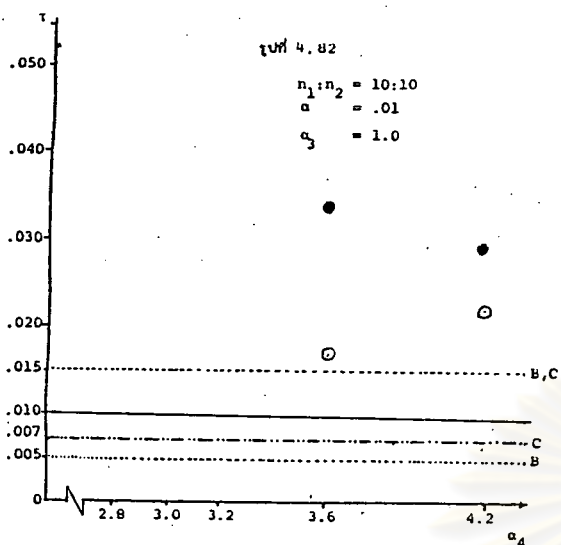
ศูนย์วิทยทรัพยากร  
 จุฬาลงกรณ์มหาวิทยาลัย



ศูนย์วิทยทรัพยากร  
 จุฬาลงกรณ์มหาวิทยาลัย



ศูนย์วิทยทรัพยากร  
 จุฬาลงกรณ์มหาวิทยาลัย



ศูนย์วิทยทรัพยากร  
 จุฬาลงกรณ์มหาวิทยาลัย



จากรูป 4.64-4.87 ซึ่งแสดงค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จากการทดลอง ( $\tau$ ) ของการทดสอบเอฟ และการทดสอบสแควร์แรงค์ ( $\mu$ ) เมื่อประชากรมีการแจกแจงแบบเบ้ โดยมีความเบ้และความโด่งเป็นคู่ลำดับดังนี้ (.25, 2.8) (.25, 3.0) (.25, 3.2) (.25, 3.6) (.25, 3.6) (.25, 4.2) (.50, 2.8) (.50, 2.8) (.50, 3.0) (.50, 3.2) (.50, 3.6) (.50, 4.2) (.75, 2.8) (.75, 3.0) (.75, 3.2) (.75, 3.6) (.75, 4.2) (1.0, 3.6) และ (1.0, 4.2) และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (5:10) (5:15) (5:20) (10:10) (15:15) และ (20:20) โดยเปรียบเทียบค่า  $\tau$  กับค่า  $\alpha$  ที่กำหนด ซึ่งมีค่า .01 ด้วยเกณฑ์ของ Cochran และเกณฑ์ของ Bradley ซึ่งสรุปจำนวนครั้งที่การทดสอบทั้ง 2 วิธีดังกล่าว ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ และควบคุมไม่ได้ดังตาราง 4.5

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.5 แสดงจำนวนครั้งที่การทดสอบเอฟ และการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) (SRT) ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ และควบคุมไม่ได้ จากการทดลองทั้งหมด 9 กรณี สำหรับแต่ละค่าของความเบ้ เมื่อความโค้งมีค่า 2.8 3.0 และ 3.2 และ 6 กรณีเมื่อความโค้งมีค่า 3.6 และ 4.2 โดยกำหนดระดับนัยสำคัญเป็น .01 จำนวนความหมายของตัวอย่าง

ขนาดของ ตัวอย่าง	ความโค้ง	ความ เบ้	เกณฑ์ของ Cochran								เกณฑ์ของ Bradley							
			$\tau = \alpha$		$\tau < \alpha$		$\tau > \alpha$		$\tau \neq \alpha$		$\tau = \alpha$		$\tau < \alpha$		$\tau > \alpha$		$\tau \neq \alpha$	
			F	SRT	F	SRT	F	SRT	F	SRT	F	SRT	F	SRT	F	SRT	F	SRT
(10:10) (15:15) (20:20)	2.8, 3.0 และ 3.2	.25	6	7	3	0	0	2	3	2	7	7	2	0	0	2	2	2
		.50	4	6	4	0	1	3	5	3	6	6	2	0	1	3	3	3
		.75	6	1	2	0	1	6	3	8	8	1	0	0	1	8	1	8
		1.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	รวม		16	14	9	0	2	13	11	13	21	14	5	0	2	13	6	13
	3.6 และ 4.2	.25	1	5	0	0	5	1	5	1	1	5	0	0	5	1	5	1
		.50	0	4	0	0	6	2	6	2	0	4	0	0	6	2	6	4
		.75	0	2	0	0	6	4	6	4	0	2	0	0	6	4	6	2
		1.0	0	0	0	0	6	6	6	6	0	0	0	0	6	6	6	6
	รวม		1	11	0	0	23	13	23	13	1	11	0	0	23	13	23	13
(5:10) (5:15) (5:20)	2.8, 3.0 และ 3.2	.25	9	2	0	0	0	7	0	7	9	2	0	0	0	7	0	7
		.50	8	3	0	0	1	6	1	6	8	3	0	0	1	6	1	6
		.75	7	0	0	0	2	9	2	9	7	0	0	0	2	9	2	9
		1.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	รวม		24	5	0	0	3	22	3	22	24	5	0	0	3	22	3	22
	3.6 และ 4.2	.25	1	2	0	0	5	4	5	4	1	2	0	0	5	4	5	4
		.50	1	0	0	0	5	6	5	6	1	0	0	0	5	6	5	6
		.75	3	0	0	0	3	6	3	6	3	0	0	0	3	6	3	6
		1.0	0	0	0	0	6	6	6	6	0	0	0	0	6	6	6	6
	รวม		5	2	0	0	19	22	19	22	5	2	0	0	19	22	19	22

#### จากตาราง 4.5 สรุปลงได้ดังนี้

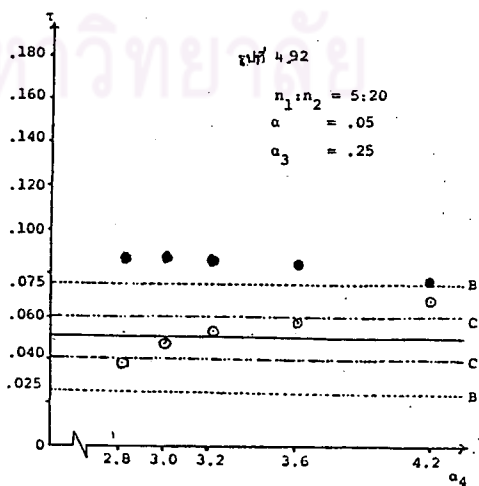
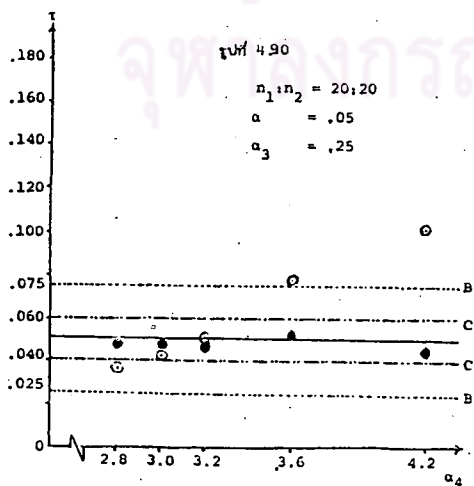
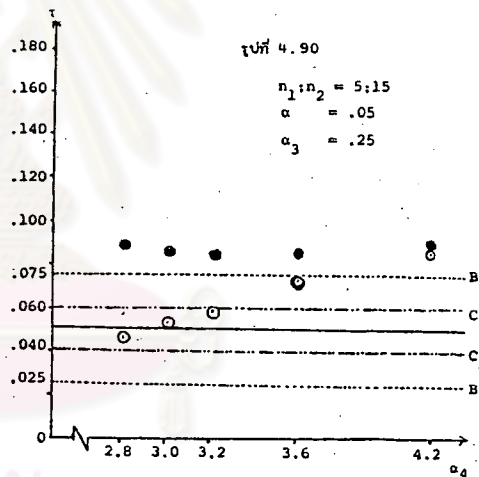
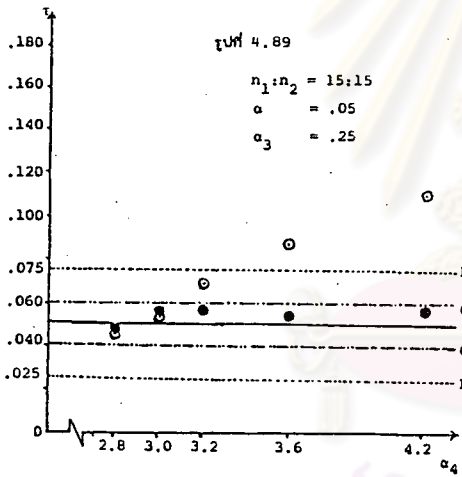
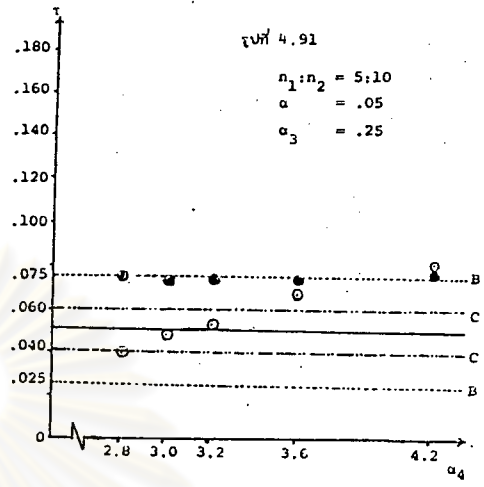
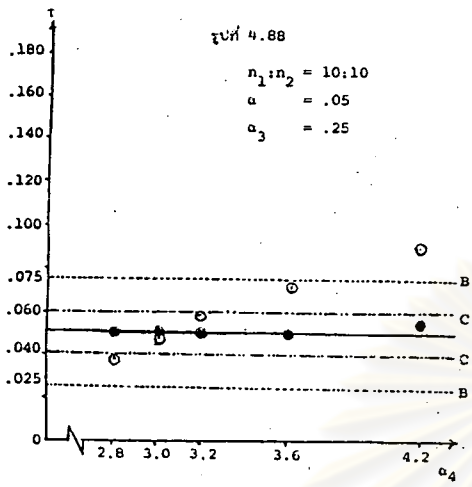
1. การทดสอบเอฟ ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี เมื่อขนาดของตัวอย่างเท่ากัน และความโด่งมีค่า 2.8, 3.0 และ 3.2 ความเบ้มีค่า .25, .50 และ .75 คือ สามารถควบคุมได้ 16 ครั้ง เมื่อใช้เกณฑ์ของ Cochran และ 21 ครั้ง เมื่อใช้เกณฑ์ของ Bradley จากการทดลองทั้งหมด 27 ครั้ง สำหรับกรณีที่ขนาดของตัวอย่างไม่เท่ากัน การทดสอบเอฟสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดีกว่ากรณีที่ขนาดของตัวอย่างเท่ากัน คือ สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ 24 ครั้ง ไม่ว่าจะใช้เกณฑ์ของ Cochran หรือเกณฑ์ของ Bradley จากการทดลองทั้งหมด 27 ครั้ง แต่ในกรณีที่ความโด่งของประชากรเป็น 3.6 และ 4.2 และความเบ้มีค่า .25, .50, .75 และ 1.0 การทดสอบเอฟควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้น้อยมาก คือ เมื่อขนาดของตัวอย่างเท่ากันนั้น สามารถควบคุมได้ 1 ครั้ง ไม่ว่าจะใช้เกณฑ์ของ Bradley หรือเกณฑ์ของ Cochran และเมื่อขนาดของตัวอย่างทั้งสองกลุ่มไม่เท่ากัน สามารถควบคุมได้ 5 ครั้ง ไม่ว่าจะใช้เกณฑ์ของ Bradley หรือ Cochran จากการทดลองทั้งหมด 24 ครั้ง

2. การทดสอบสแควร์แรงค์ ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี เมื่อขนาดของตัวอย่างเท่ากัน โดยที่ความโด่งมีค่าเป็น 2.8, 3.0 และ 3.2 และความเบ้มีค่าไม่เกิน .50 คือ สามารถควบคุมได้ 13 ครั้ง จากการทดลองทั้งหมด 18 ครั้ง ไม่ว่าจะใช้เกณฑ์ของ Bradley หรือเกณฑ์ของ Cochran แต่เมื่อความเบ้มีค่าเป็น .75 การทดสอบสแควร์แรงค์สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้น้อย คือ ควบคุมได้ 1 ครั้งจากการทดลองทั้งหมด 9 ครั้ง ไม่ว่าจะพิจารณาโดยใช้เกณฑ์ของ Bradley หรือเกณฑ์ของ Cochran อย่างไรก็ตามเมื่อความโด่งมีค่า 3.6 และ 4.2 และความเบ้มีค่าไม่เกิน .50 การทดสอบสแควร์แรงค์ก็สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี คือ สามารถควบคุมได้ 9 ครั้ง จากการทดลองทั้งหมด 12 ครั้ง ไม่ว่าจะใช้เกณฑ์ของ Bradley หรือเกณฑ์ของ Cochran แต่เมื่อความเบ้มีค่า .75 และ 1.0 การทดสอบสแควร์แรงค์ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้น้อย คือ ควบคุมได้ 2 ครั้งจากการทดลองทั้งหมด 12 ครั้ง ทั้งกรณีที่พิจารณาโดยใช้เกณฑ์ของ Bradley และเกณฑ์ของ Cochran สำหรับกรณีที่ขนาดของตัวอย่างไม่เท่ากัน การทดสอบสแควร์แรงค์ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้น้อย กล่าวคือ เมื่อความเบ้มีค่า .25 และ .50 และความโด่งมีค่า 2.8, 3.0 และ 3.2 ควบคุมได้ 5 ครั้ง จากการทดลองทั้งหมด 18 ครั้ง ไม่ว่าจะใช้เกณฑ์ของ Bradley หรือเกณฑ์ของ Cochran พิจารณา และเมื่อความเบ้มีค่า .75 การ

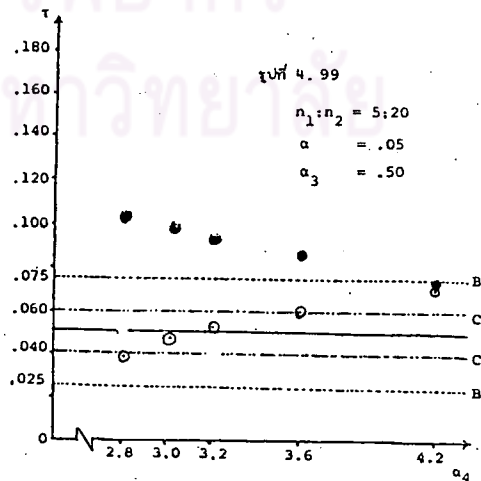
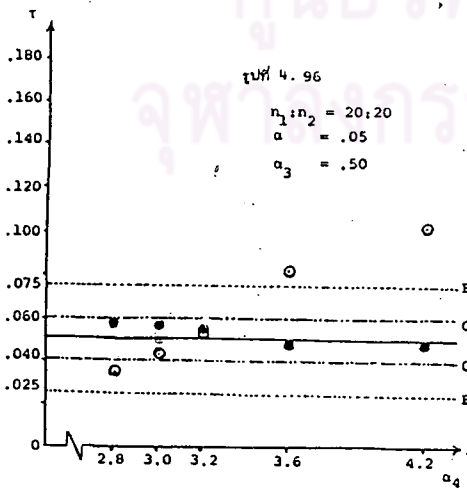
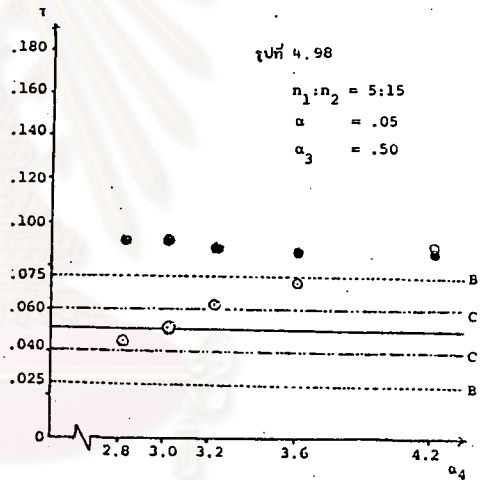
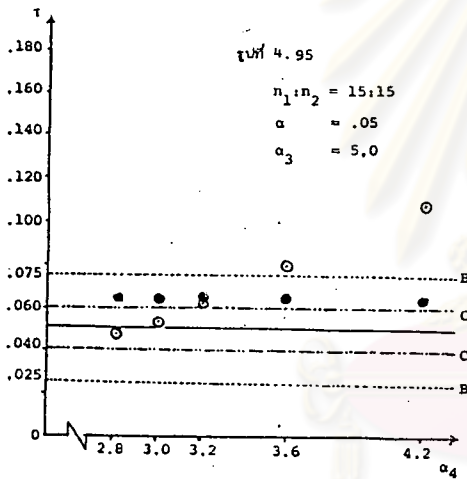
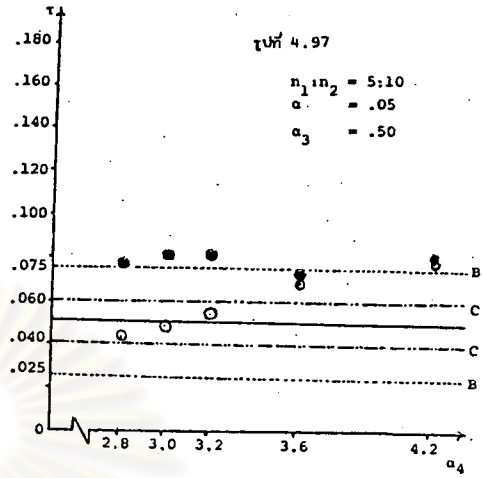
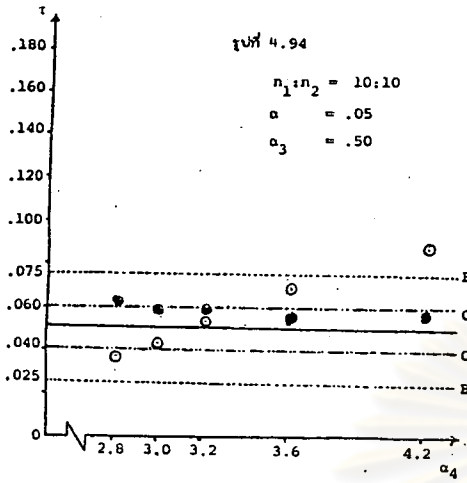
ทดสอบสแควร์แรงค์ไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ ไม่ว่าจะใช้เกณฑ์ของ Bradley หรือเกณฑ์ของ Cochran อย่างไรก็ตามเมื่อความโด่งมีค่า 3.6 และ 4.2 การทดสอบสแควร์แรงค์ก็ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้น้อย คือควบคุมได้ 2 ครั้งจากการทดลองทั้งหมด 24 ครั้ง ทั้งกรณีที่ใช้เกณฑ์ของ Bradley และเกณฑ์ของ Cochran

3. ลักษณะที่ควบคุมความคลาดเคลื่อนประเภทที่ 1 ไม่ได้ของการทดสอบเอฟเมื่อความโด่งมีค่า 2.8, 3.0 และ 3.2 และความเบ้มีค่า .25, .30 และ .75 จะมีทั้งกรณีที่  $\tau < \alpha$  และ  $\tau > \alpha$  แต่เมื่อความโด่งมีค่า 3.6 และ 4.2 และความเบ้มีค่า .25, .50, .75 และ 1.0 ลักษณะที่ควบคุมความคลาดเคลื่อนประเภทที่ 1 ไม่ได้คือ  $\tau > \alpha$  เท่านั้น ในขณะที่การทดสอบสแควร์แรงค์ ลักษณะที่ควบคุมความคลาดเคลื่อนประเภทที่ 1 ไม่ได้ คือ  $\tau > \alpha$  ทุก ๆ ค่า ของความเบ้และความโด่ง

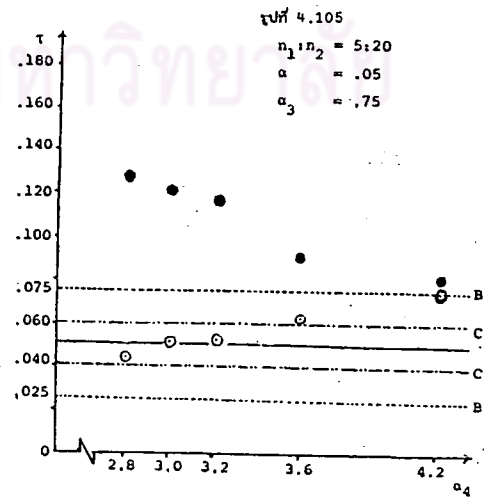
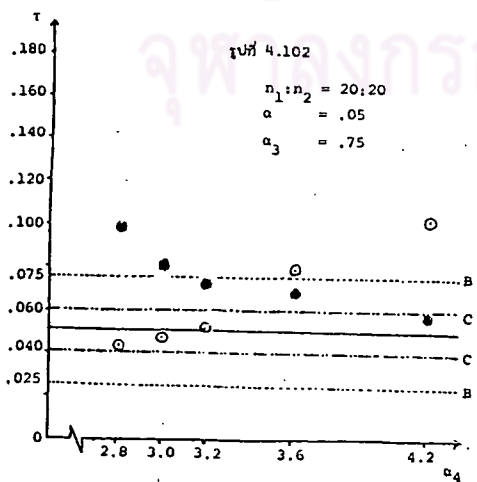
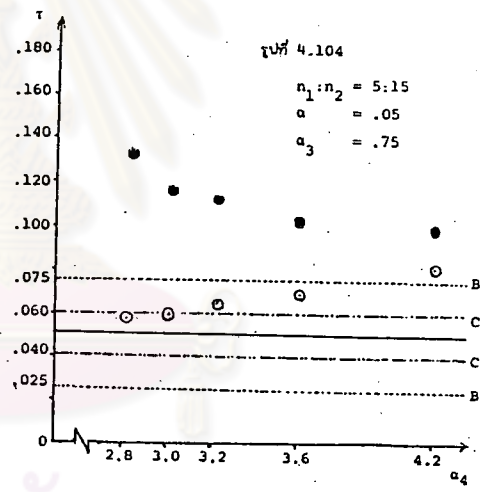
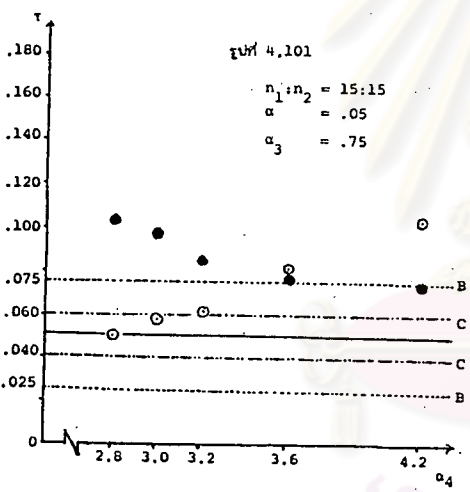
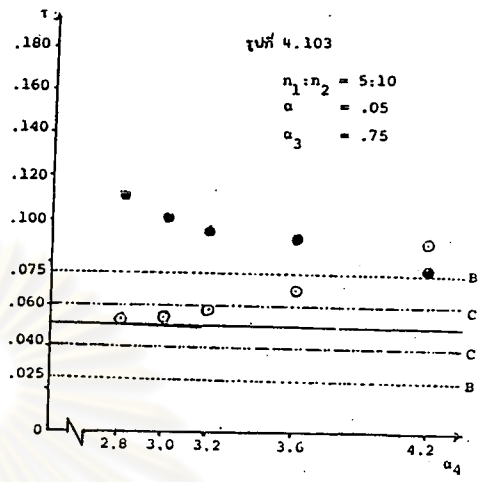
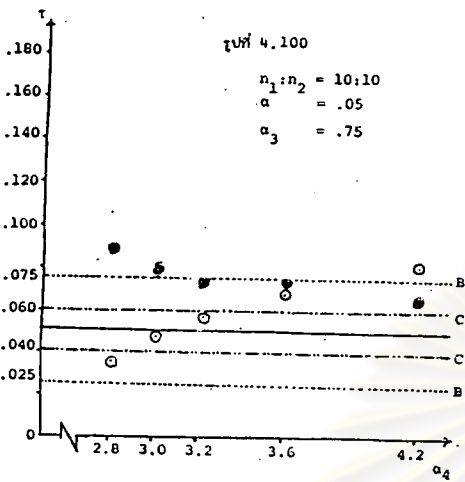
ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย



ศูนย์วิทยทรัพยากร  
 จุฬาลงกรณ์มหาวิทยาลัย

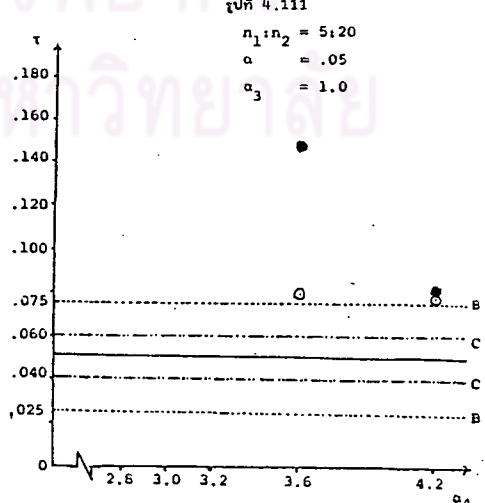
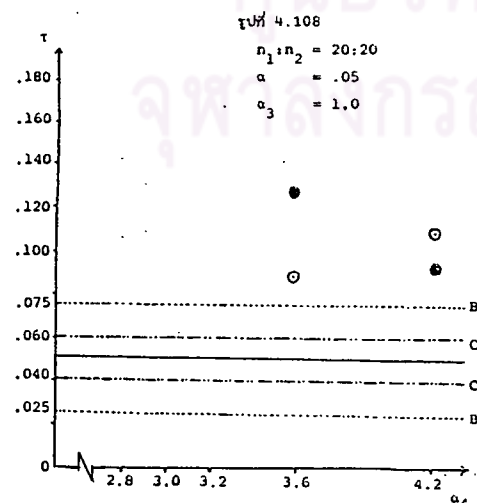
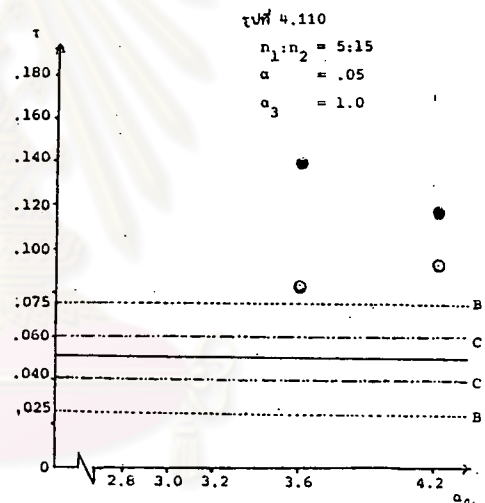
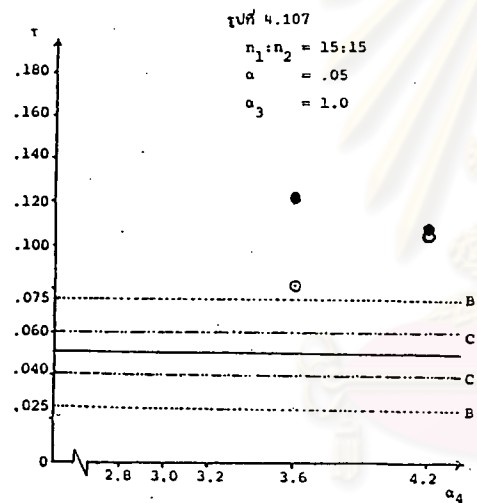
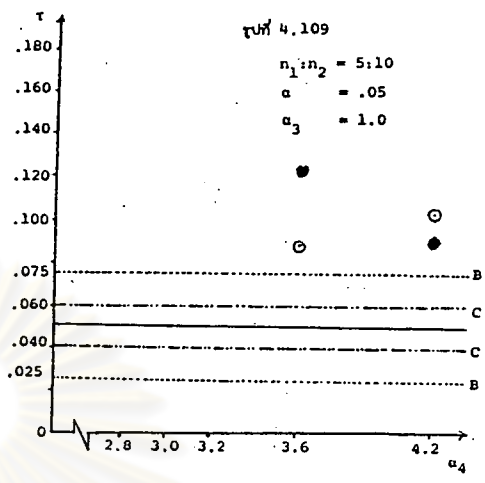
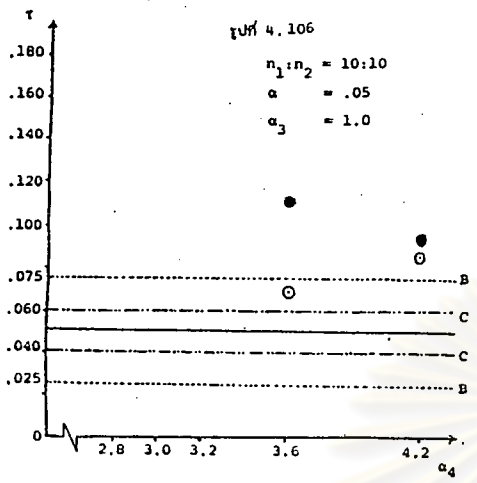


ศูนย์วิทยทรัพยากร  
 จุฬาลงกรณ์มหาวิทยาลัย



ศูนย์วิทยทรัพยากร  
 จุฬาลงกรณ์มหาวิทยาลัย





ศูนย์วิทยทรัพยากร  
 จุฬาลงกรณ์มหาวิทยาลัย



จากรูป 4.88-4.111 ซึ่งแสดงค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จากการทดลอง ( $T$ ) ของการทดสอบเอฟ และการทดสอบสแควร์แรงค์ ( $U$ ) เมื่อประจําการมีการแจกแจงแบบเบ้ โดยมีความเบ้และความโด่งเป็นคู่ลำดับดังนี้ (.25, 2.8) (.25, 3.0) (.25, 3.2) (.25, 3.6) (.25, 4.2) (.50, 2.8) (.50, 3.0) (.50, 3.2) (.50, 3.6) (.50, 4.2) (.75, 2.8) (.75, 3.0) (.75, 3.2) (.75, 3.6) (.75, 4.2) (1.0, 3.6) และ (1.0, 4.2) และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (5:10) (5:15) (5:20) (10:10) (15:15) และ (20:20) โดยเปรียบเทียบค่า  $T$  กับค่า  $\alpha$  ที่กำหนด ซึ่งมีค่า .05 ด้วยเกณฑ์ของ Cochran และเกณฑ์ของ Bradley ซึ่งสรุปจำนวนครั้งที่การทดสอบทั้ง 2 วิธีดังกล่าวควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ และควบคุมไม่ได้ ดังตาราง 4.6.

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.6 แสดงจำนวนครั้งที่การทดสอบเอฟ และการทดสอบลีแควร์แรงค์ ( $\bar{X}$ ) (SRT) ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้และความคุมไม่ได้ จากการทดลองทั้งหมด 9 กรณีสำหรับแต่ละค่าของความเบ้ เมื่อความโด่งมีค่า 2.8 3.0 และ 3.2 และ 6 กรณีเมื่อความโด่งมีค่า 3.6 และ 4.2 โดยกำหนดระดับนัยสำคัญเป็น .05 ค่าแนกตามขนาดของตัวอย่าง

ขนาดของ ตัวอย่าง	ความโด่ง	ความ เบ้	เกณฑ์ของ Cochran				เกณฑ์ของ Bradley											
			$T = \alpha$		$T < \alpha$		$T > \alpha$		$T \neq \alpha$		$T = \alpha$		$T < \alpha$		$T > \alpha$		$T \neq \alpha$	
			F	SRT	F	SRT	F	SRT	F	SRT	F	SRT	F	SRT	F	SRT	F	SRT
(10:10)	2.8, 3.0 และ 3.2	.25	6	9	2	0	1	0	3	0	9	9	0	0	0	0	0	0
		.50	6	5	2	0	1	4	3	4	9	9	0	0	0	0	0	0
		.75	7	0	1	0	1	9	2		9	2	0	0	0	7	0	7
		1.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	รวม		19	14	5	0	3	13	8	13	27	20	0	0	0	7	0	7
(15:15)	3.6 และ 4.2	.25	0	6	0	0	6	0	6	0	1	6	0	0	5	0	5	0
		.50	0	4	0	0	6	2	6	2	1	6	0	0	5	0	5	0
		.75	0	1	0	0	6	5	6	5	1	5	0	0	5	0	5	1
		1.0	0	0	0	0	6	6	6	6	1	0	0	0	5	6	5	6
	รวม		0	11	0	0	24	13	24	13	4	17	0	0	20	7	20	7
(20:20)	2.8, 3.0 และ 3.2	.25	8	0	1	0	0	9	1	9	9	3	0	0	0	6	0	6
		.50	7	0	1	0	1	9	2	9	9	0	0	0	9	0	9	
		.75	8	0	0	0	1	9	1	9	9	0	0	0	9	0	9	
		1.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	รวม		23	0	2	0	2	27	4	27	27	3	0	0	0	24	0	24
(5:10)	3.6 และ 4.2	.25	1	0	0	0	5	6	5	6	4	1	0	0	2	5	2	5
		.50	1	0	0	0	5	6	5	6	4	2	0	0	2	4	2	4
		.75	0	0	0	0	6	6	6	6	4	0	0	0	2	6	2	6
		1.0	0	0	0	0	6	6	6	6	0	0	0	0	6	6	6	6
	รวม		2	0	0	0	22	24	22	24	12	3	0	0	12	21	12	21

#### จากตาราง 4.6 สรุปได้ดังนี้

1. การทดสอบเอฟสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี เมื่อขนาดของตัวอย่างเท่ากัน เมื่อความโด่งมีค่า 2.8, 3.0 และ 3.2 และความเบ้มีค่าเป็น .25, .50 และ .75 คือสามารถควบคุมได้ 19 ครั้งจากการทดลองทั้งหมด 27 ครั้ง เมื่อใช้เกณฑ์ของ Cochran และควบคุมได้ทั้งหมดเมื่อใช้เกณฑ์ของ Bradley และเมื่อขนาดของตัวอย่างทั้งสองกลุ่มไม่เท่ากัน การทดสอบเอฟสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี เช่นเดียวกับกรณีที่ขนาดของตัวอย่างเท่ากัน คือ สามารถควบคุมได้ 23 ครั้ง เมื่อใช้เกณฑ์ของ Cochran และควบคุมได้ทั้งหมดเมื่อใช้เกณฑ์ของ Bradley จากการทดลองทั้งหมด 27 ครั้ง ในขณะที่ความโด่งมีค่า 3.6 และ 4.2 และความเบ้มีค่า .25, .50, .75 และ 1.0 การทดสอบเอฟสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้น้อย กล่าวคือ เมื่อขนาดของตัวอย่างเท่ากันสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ 4 ครั้งเมื่อใช้เกณฑ์ของ Bradley และไม่สามารควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ เมื่อใช้เกณฑ์ของ Cochran จากการทดลองทั้งหมด 24 ครั้ง แต่เมื่อขนาดของตัวอย่างไม่เท่ากันความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 จะดีกว่ากรณีที่ขนาดของตัวอย่างเท่ากัน โดยเฉพาะเมื่อใช้เกณฑ์ของ Bradley โดยที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ 2 ครั้ง เมื่อใช้เกณฑ์ของ Cochran และ 12 ครั้งเมื่อใช้เกณฑ์ของ Bradley จากการทดลองทั้งหมด 24 ครั้ง

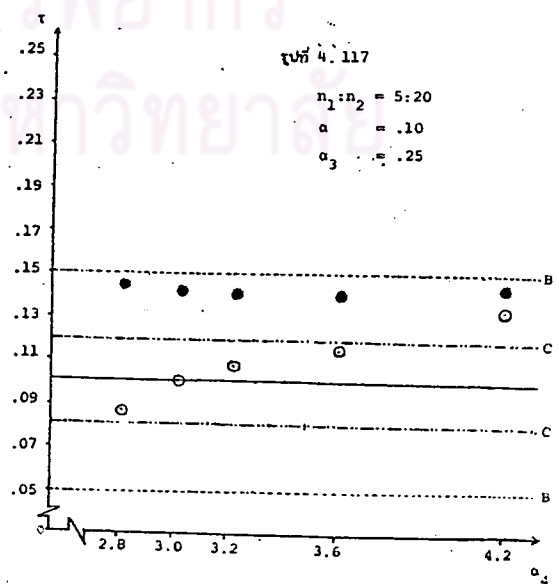
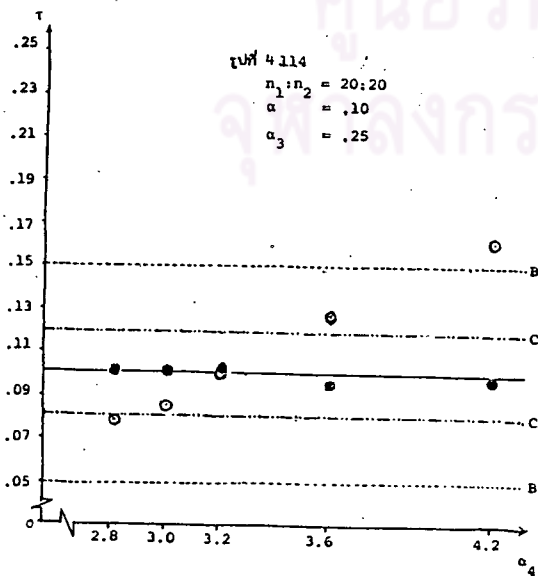
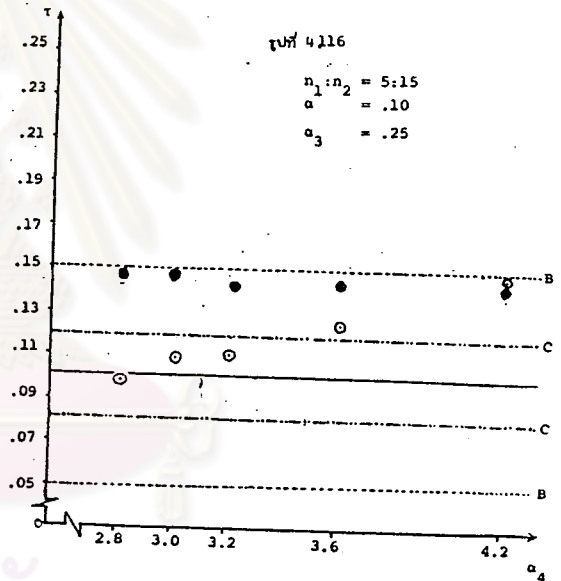
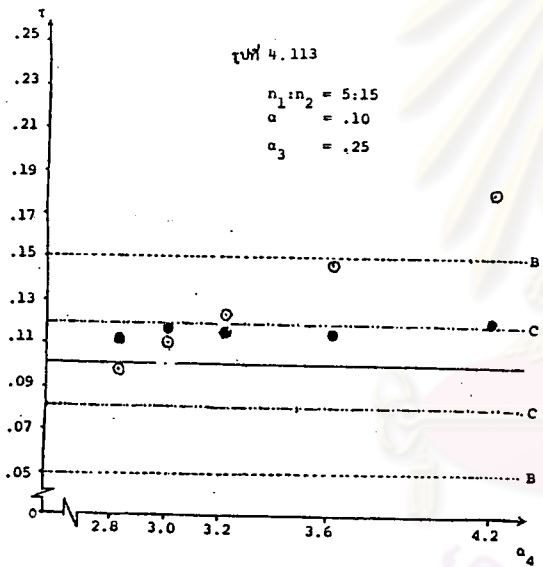
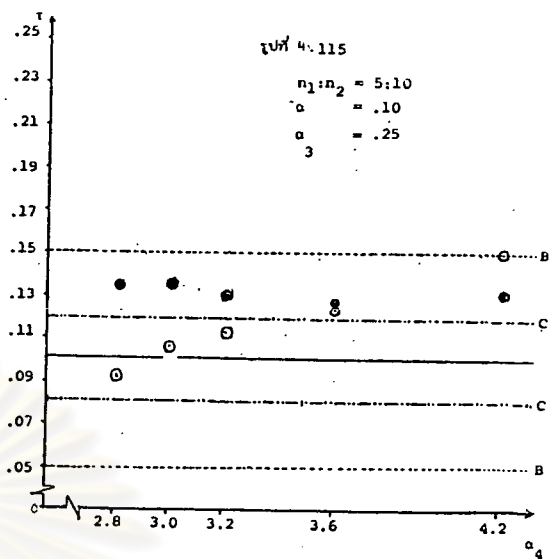
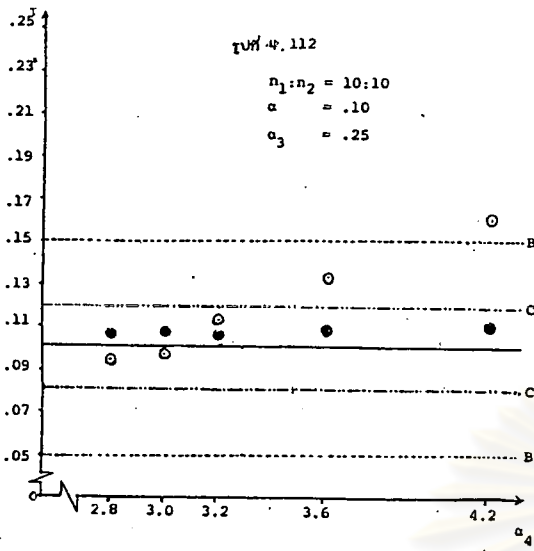
2. การทดสอบสแควร์แรงค์ สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี เมื่อขนาดของตัวอย่างเท่ากัน โดยที่ความโด่งมีค่า 2.8, 3.0 และ 3.2 และความเบ้มีค่าไม่เกิน .50 คือ สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ 14 ครั้ง เมื่อใช้เกณฑ์ของ Cochran และควบคุมได้ทั้งหมดเมื่อใช้เกณฑ์ของ Bradley จากการทดลองทั้งหมด 18 ครั้ง แต่เมื่อความเบ้มีค่า .75 การทดสอบสแควร์แรงค์สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้น้อย คือ ไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เมื่อใช้เกณฑ์ของ Cochran และควบคุมได้ 2 ครั้งเมื่อใช้เกณฑ์ของ Bradley จากการทดลองทั้งหมด 9 ครั้ง สำหรับกรณีที่ความโด่งมีค่า 3.6 และ 4.2 เมื่อความเบ้มีค่าไม่เกิน .50 สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ 10 ครั้งจากการทดลองทั้งหมด 12 ครั้งเมื่อใช้เกณฑ์ของ Cochran และควบคุมได้ทั้งหมดเมื่อใช้เกณฑ์ของ Bradley แต่เมื่อความเบ้มากกว่า .50 การทดสอบสแควร์แรงค์สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ 1 ครั้งเมื่อใช้เกณฑ์

ของ Cochran แต่ควบคุมได้ 5 ครั้งจากการทดลองทั้งหมด 12 ครั้ง เมื่อใช้เกณฑ์ของ Bradley สำหรับกรณีที่ขนาดของตัวอย่างทั้ง 2 กลุ่มไม่เท่ากัน การทดสอบสแควร์แรงค์จะควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้น้อยกว่ากรณีที่ขนาดของตัวอย่างเท่ากัน กล่าวคือ เมื่อความเบ้มีค่า .25, .50 และ .75 และความโด่งมีค่า 2.8, 3.0 และ 3.2 สามารถควบคุมได้ 3 ครั้งจากการทดลองทั้งหมด 18 ครั้ง เมื่อใช้เกณฑ์ของ Bradley และไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เมื่อใช้เกณฑ์ของ Cochran และเมื่อความโด่งมีค่า 3.6 และ 4.2 และความเบ้มีค่า .25, .50, .75 และ 1.0 การทดสอบสแควร์แรงค์ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้น้อย โดยสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ 3 ครั้ง เมื่อใช้เกณฑ์ของ Bradley และไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เมื่อใช้เกณฑ์ของ Cochran

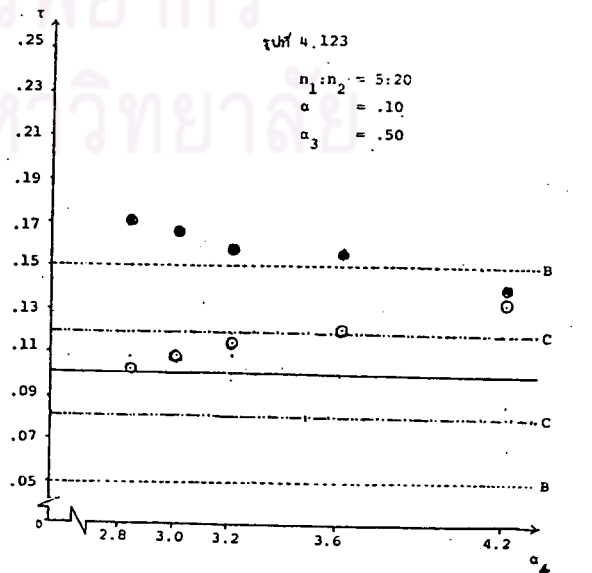
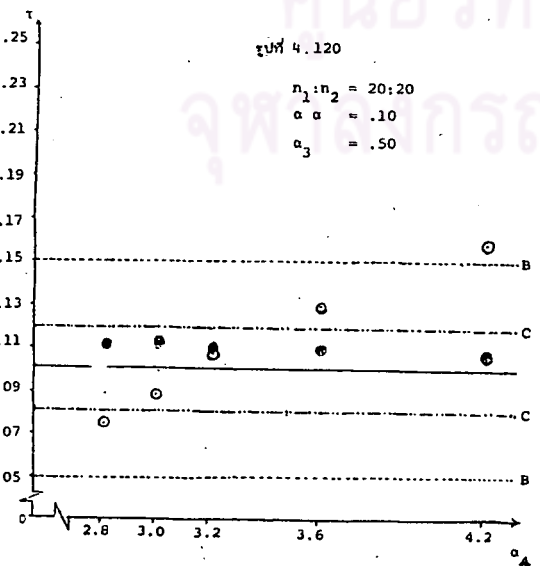
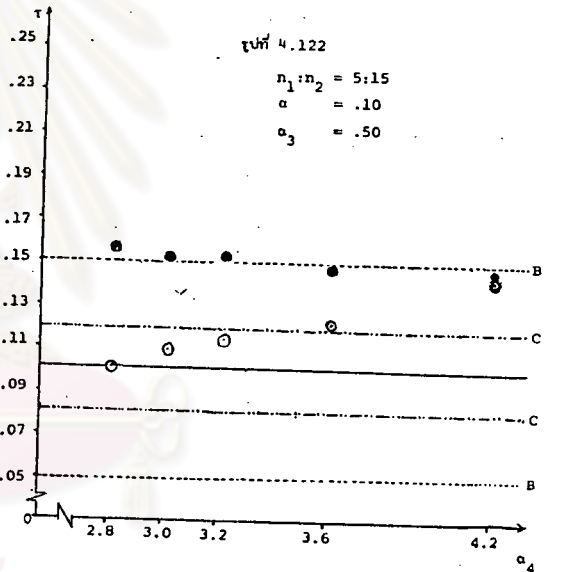
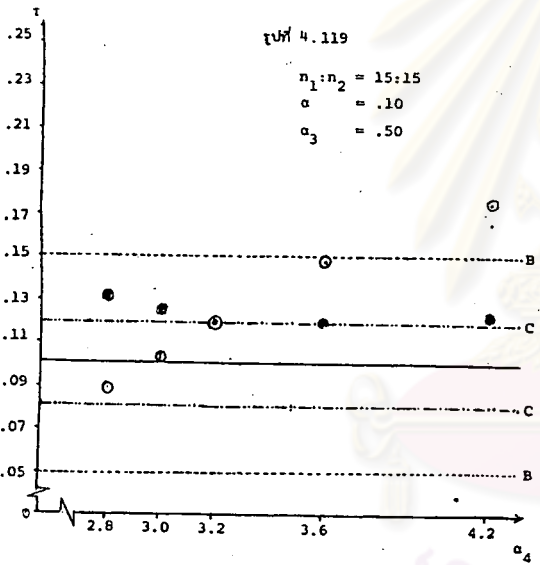
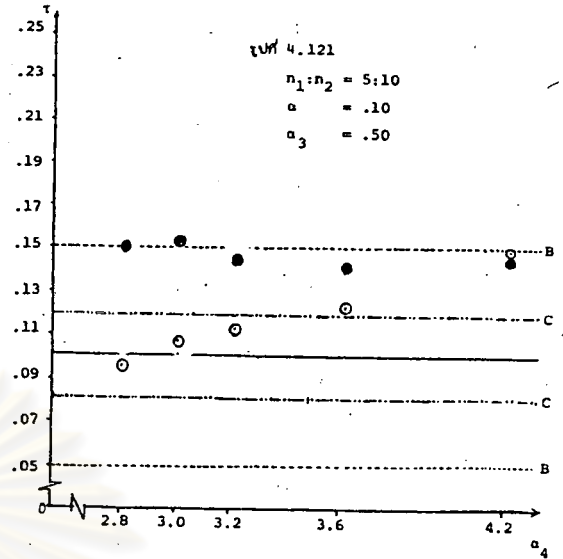
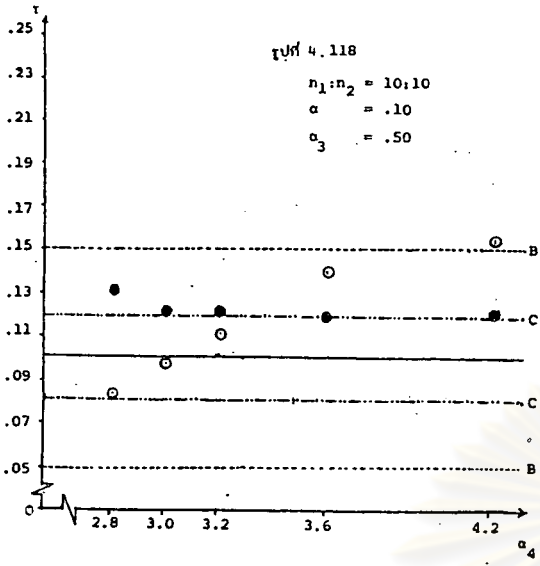
3. ลักษณะที่ควบคุมไม่ได้ของการทดสอบเอฟ เมื่อความโด่งมีค่า 2.8, 3.0 และ 3.2 และความโด่งมีค่า .25, .50 และ .75 จะมีทั้งกรณีที่  $\tau > \alpha$  และ  $\tau < \alpha$  แต่เมื่อความเบ้มีค่า 3.6 และ 4.2 ลักษณะที่ควบคุมไม่ได้คือค่า  $\tau > \alpha$  เท่านั้น สำหรับการทดสอบสแควร์แรงค์ลักษณะที่ควบคุมไม่ได้คือค่า  $\tau > \alpha$  ทุก ๆ ค่าของความเบ้และความโด่ง



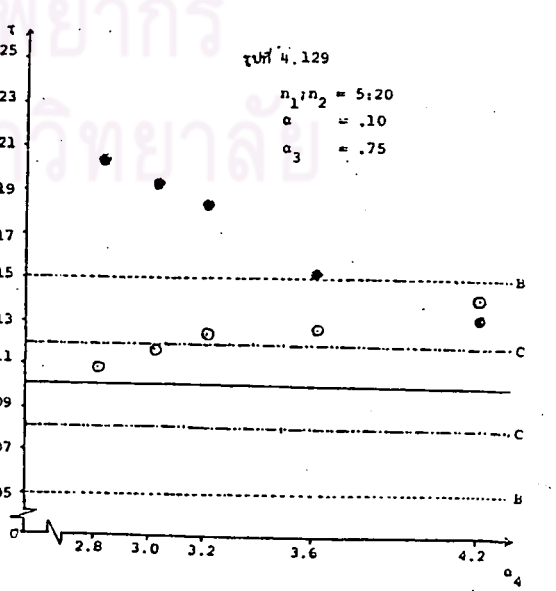
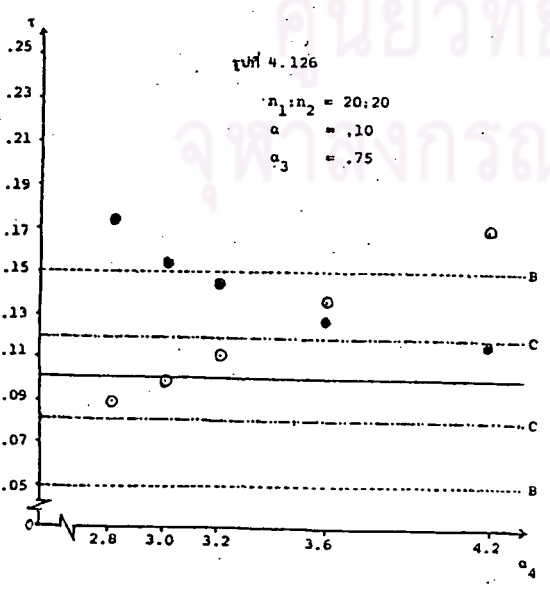
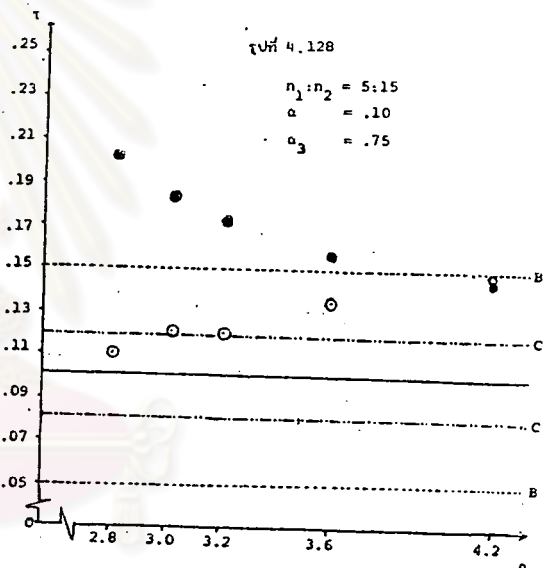
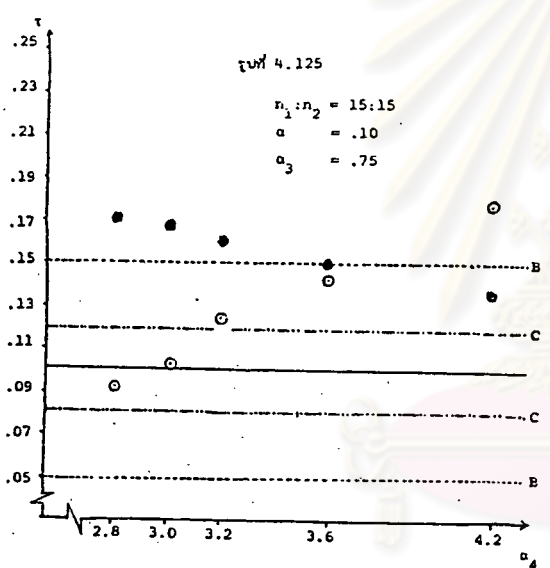
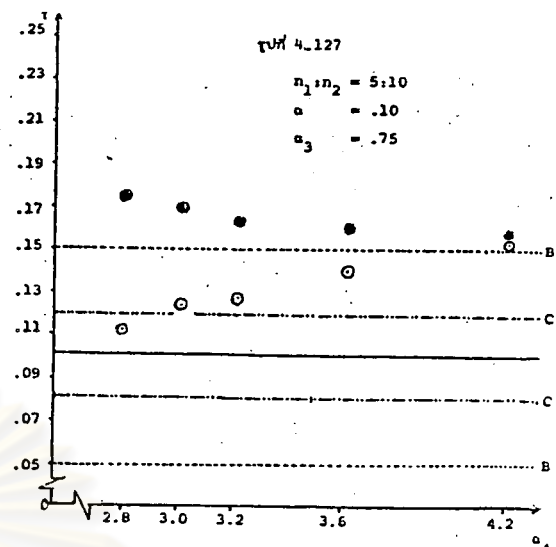
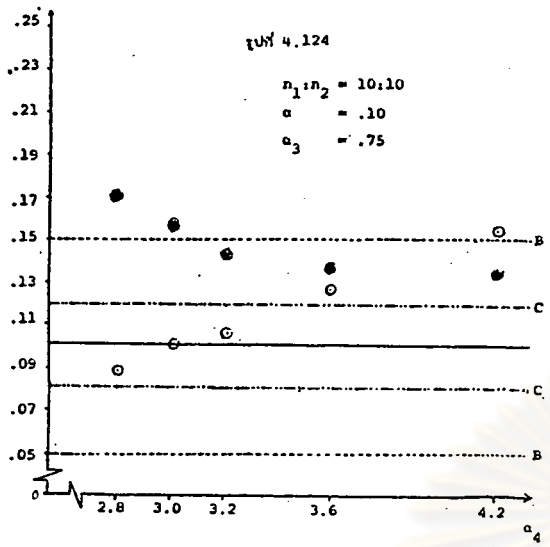
ศูนย์วิจัยสุขภาพ  
จุฬาลงกรณ์มหาวิทยาลัย



ศูนย์วิทยทรัพยากร  
 จุฬาลงกรณ์มหาวิทยาลัย

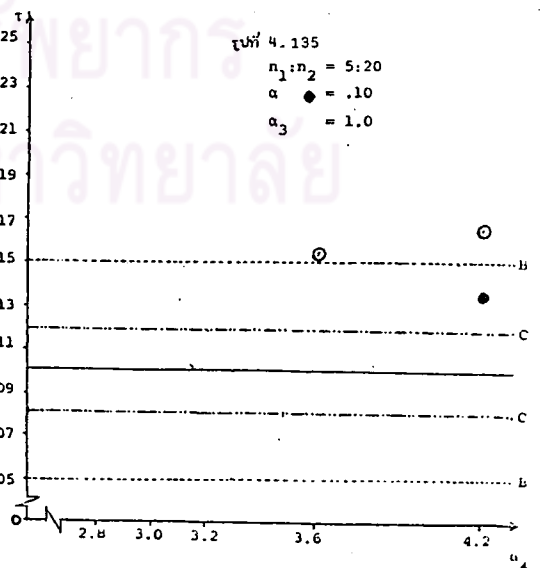
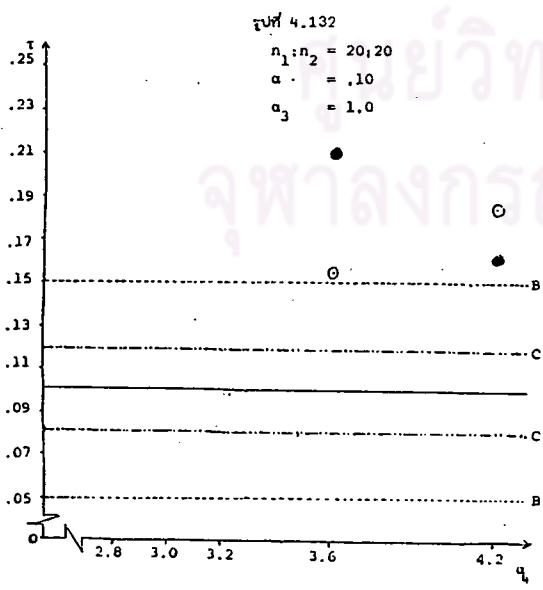
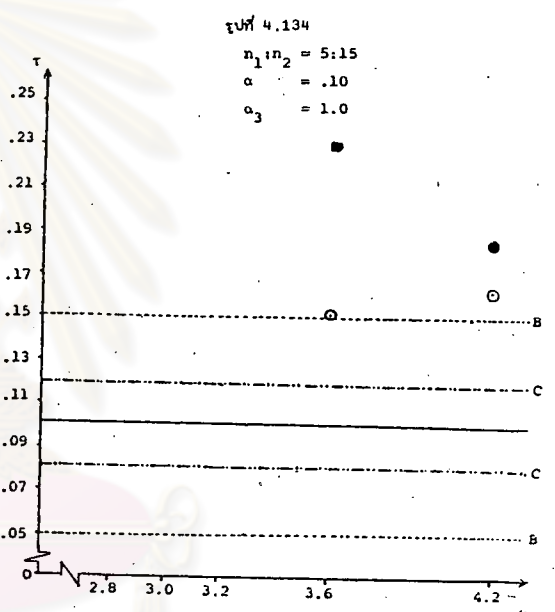
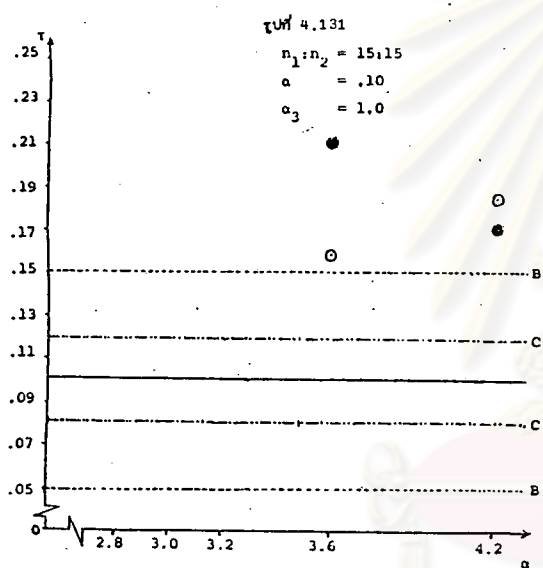
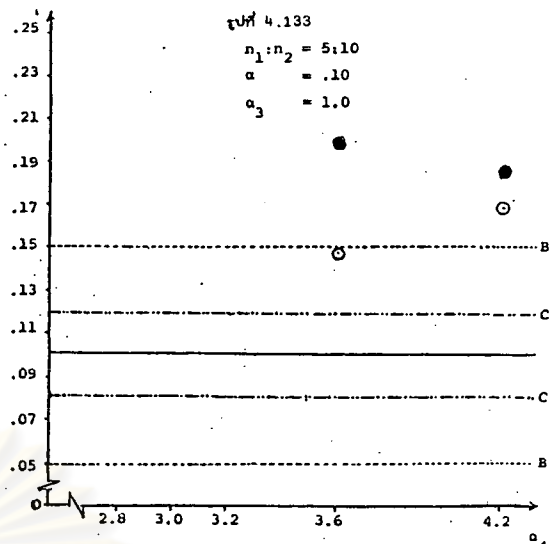
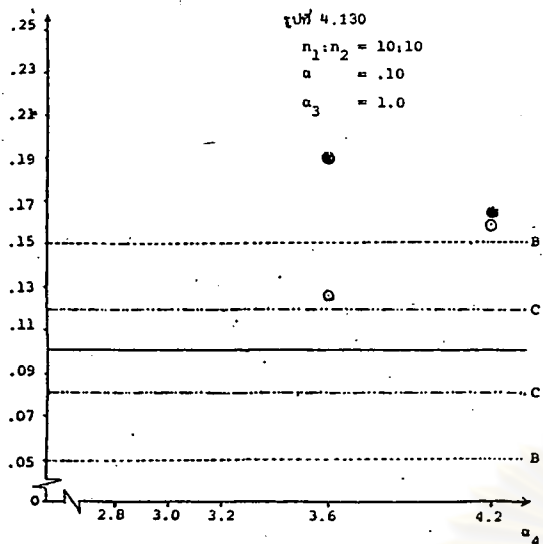


ศูนย์วิทยทรัพยากร  
 จุฬาลงกรณ์มหาวิทยาลัย



ศูนย์วิทยทรัพยากร  
 จุฬาลงกรณ์มหาวิทยาลัย







จากรูป 4.112-4.135 ซึ่งแสดงค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จากการทดลอง ( $\tau$ ) ของการทดสอบเอฟและการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) เมื่อประชากรมีการแจกแจงแบบเบ้โดยมีความเบ้และความโด่งเป็นคู่ลำดับดังนี้ (-25, 2.8) (-25, 3.0) (-25, 3.2) (-25, 3.6) (-25, 4.2) (-50, 2.8) (-50, 3.0) (-50, 3.2) (-50, 3.6) (-50, 4.2) (-75, 2.8) (-75, 3.0) (-75, 3.2) (-75, 3.6) (-75, 4.2) (1.0, 3.6) และ (1.0, 4.2) และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (5:10), (5:15), (5:20), (10:10), (15:15) และ (20:20) โดยเปรียบเทียบค่า  $\tau$  กับค่า  $\alpha$  ที่กำหนด ซึ่งมีค่า .10 ด้วยเกณฑ์ของ Cochran และเกณฑ์ของ Bradley ซึ่งสรุปจำนวนครั้งที่การทดสอบทั้ง 2 วิธีดังกล่าว ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ และควบคุมไม่ได้ ดังตาราง

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.7 แสดงจำนวนครั้งที่การทดสอบเอฟ และการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) (SRT) ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ และควบคุมไม่ได้ จากการทดลองทั้งหมด 9 กรณี สำหรับแต่ละค่าของความเบ้ เมื่อความโด่งมีค่า 2.8 3.0 และ 3.2 และ 6 กรณีเมื่อความโด่งมีค่า 3.6, และ 4.2 โดยกำหนดระดับนัยสำคัญเป็น .10 ค่าแนกตามขนาดของตัวอย่าง

ขนาดของ ตัวอย่าง	ความโด่ง	ความ เบ้	เกณฑ์ของ Cochran				เกณฑ์ของ Bradley											
			$\tau = \alpha$		$\tau < \alpha$		$\tau > \alpha$		$\tau \neq \alpha$									
			F	SRT	F	SRT	F	SRT	F	SRT								
(10:10)	2.8, 3.0 และ 3.2	.25	6	9	1	0	1	0	2	0	9	9	0	0	0	0		
		.50	8	4	1	0	0	5	1	5	9	9	0	0	0	0		
		.75	8	0	0	0	1	9	1	9	9	2	0	0	0	7	0	7
		1.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
	รวม		22	13	2	0	2	14	4	14	27	20	0	0	0	7	0	7
(15:15)	3.6 และ 4.2	.25	0	5	0	0	9	1	9	1	3	6	0	0	3	0	3	0
		.50	0	4	0	0	6	2	6	2	3	6	0	0	3	0	3	0
		.75	0	0	0	0	6	6	6	6	3	6	0	0	3	0	0	0
		1.0	0	0	0	0	6	6	6	6	1	0	0	0	5	6	5	6
	รวม		0	9	0	0	27	15	27	15	10	18	0	0	14	6	14	6
(20:20)	2.8, 3.0 และ 3.2	.25	9	0	0	0	0	9	0	9	9	9	0	0	0	0	0	0
		.50	9	0	0	0	0	9	0	9	9	2	0	0	0	7	0	7
		.75	4	0	0	0	2	9	2	9	9	0	0	0	0	9	0	9
		1.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	รวม		22	0	0	0	2	27	2	27	27	11	0	0	0	16	0	16
(5:10)	3.6 และ 4.2	.25	1	0	0	0	5	6	5	6	6	6	0	0	0	0	0	0
		.50	0	0	0	0	6	6	6	6	6	5	0	0	0	1	0	1
		.75	0	0	0	0	6	6	6	6	5	2	0	0	1	4	1	4
		1.0	0	0	0	0	6	6	6	6	1	1	0	0	0	5	5	5
	รวม		1	0	0	0	17	18	18	18	18	14	0	0	1	10	6	10

#### จากตาราง 4.7 สรุปลงได้ดังนี้

1. การทดสอบเอฟสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี เมื่อขนาดของตัวอย่างเท่ากัน และความโด่งมีค่า 2.8, 3.0 และ 3.2 และความเบ้มีค่า .25, .50 และ .75 โดยควบคุมได้ 22 ครั้งจากการทดลองทั้งหมด 27 ครั้ง เมื่อใช้เกณฑ์ของ Cochran และควบคุมได้หมดเมื่อใช้เกณฑ์ของ Bradley และในกรณีที่ขนาดของตัวอย่างไม่เท่ากัน การทดสอบเอฟก็สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดีเช่นเดียวกับกรณีที่ขนาดของตัวอย่างเท่ากัน กล่าวคือ สามารถควบคุมได้ 22 ครั้งจากการทดลองทั้งหมด 27 ครั้ง เมื่อใช้เกณฑ์ของ Cochran และควบคุมได้ทั้งหมดเมื่อใช้เกณฑ์ของ Bradley แต่เมื่อความโด่งมีค่าเป็น 3.6 และ 4.2 และความเบ้มีค่า .25, .50 .75 และ 1.0 โดยที่ขนาดของตัวอย่างเท่ากัน การทดสอบเอฟสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้น้อย คือควบคุมได้ 10 ครั้ง จากการทดลองทั้งหมด 24 ครั้ง เมื่อใช้เกณฑ์ของ Bradley และไม่สามารควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ เมื่อใช้เกณฑ์ของ Cochran แต่เมื่อขนาดของตัวอย่างไม่เท่ากันสามารถควบคุมได้ถึง 18 ครั้งจากการทดลองทั้งหมด 24 ครั้ง และควบคุมได้เพียง 1 ครั้งเมื่อใช้เกณฑ์ของ Cochran

2. การทดสอบสแควร์แรงค์ ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดีเมื่อขนาดของตัวอย่างเท่ากัน โดยที่ความโด่งของประชากรเป็น 2.8, 3.0 และ 3.2 และความเบ้เป็นมีค่าไม่เกิน .50 โดยควบคุมได้ 13 ครั้งจากการทดลองทั้งหมด 18 ครั้ง เมื่อใช้เกณฑ์ของ Cochran และควบคุมได้ทั้งหมดเมื่อใช้เกณฑ์ของ Bradley ในขณะที่ความเบ้มีค่า .75 การทดสอบสแควร์แรงค์สามารถควบคุมได้ 2 ครั้งเมื่อใช้เกณฑ์ของ Bradley และไม่สามารควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ เมื่อใช้เกณฑ์ของ Cochran สำหรับกรณีที่ความโด่งมีค่า 3.6 และ 4.2 และความเบ้มีค่าเป็น .25 และ .50 การทดสอบสแควร์แรงค์ ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ 9 ครั้งจากการทดลองทั้งหมด 12 ครั้ง เมื่อใช้เกณฑ์ของ Cochran และควบคุมได้ทั้งหมดเมื่อใช้เกณฑ์ของ Bradley และในกรณีที่ขนาดของตัวอย่างไม่เท่ากัน เมื่อความเบ้มีค่า .25, .50 และ .75 และความโด่งมีค่า 2.8, 3.0 และ 3.2 การทดสอบสแควร์แรงค์ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ 11 ครั้ง จากการทดลองทั้งหมด 27 ครั้ง และไม่สามารควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ เมื่อใช้เกณฑ์ของ Cochran และเมื่อความโด่งมีค่า 3.6 และ 4.2 การทดสอบสแควร์แรงค์ ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ 14 ครั้งจากการทดลองทั้งหมด 24 ครั้ง เมื่อใช้

เกณฑ์ของ Bradley และไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เมื่อใช้เกณฑ์ของ Cochran

3. สภาวะที่ควบคุมความคลาดเคลื่อนประเภทที่ 1 ไม่ได้ของการทดสอบเอฟเมื่อความโด่งมีค่า 2.8, 3.0 และ 3.2 และความเบ้มีค่า .25, .50 มีทั้งกรณีที่  $\tau$  มากกว่า  $\alpha$  และ  $\tau < \alpha$  แต่ในกรณีที่ความโด่งมีค่า 3.6 และ 4.2 สภาวะที่ควบคุมความคลาดเคลื่อนประเภทที่ 1 ไม่ได้คือ  $\tau > \alpha$  เท่านั้น สำหรับการทดสอบสแควร์แรงค์ สภาวะที่ควบคุมไม่ได้คือ  $\tau > \alpha$  ทุก ๆ ค่าของความเบ้และความโด่ง



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

### อำนาจของการทดสอบ

สำหรับอำนาจของการทดสอบ จากการทดลองนั้นการนำเสนอบ้างเป็น 2 กรณีคือ

1. ในกรณีที่ประชากรมีการแจกแจงแบบปกติ แบบยูนิฟอร์ม และแบบโลจิสติก จำแนกได้เป็น 2 กรณีย่อยคือ

1.1 อำนาจของการทดสอบเอฟ การทดสอบคลอท์ การทดสอบทูกี-ซีเกล การทดสอบมัต และการทดสอบสแควร์แรงค์ ( $\mu$ ) จะนำเสนอมอำนาจของการทดสอบทั้ง 5 วิธีดังกล่าวในรูปของตาราง โดยที่ในกรณีที่ขนาดของตัวอย่างเท่ากันจะนำเสนอมอำนาจของการทดสอบเฉพาะกรณีที่  $\Delta^2$  มากกว่า 1 เท่านั้น เพราะว่าในกรณีที่  $\Delta^2$  น้อยกว่า 1 ผลสรุปที่ได้ทำนองเดียวกับกรณี  $\Delta^2$  มากกว่า 1 สำหรับกรณีที่ขนาดของตัวอย่างไม่เท่ากัน จะนำเสนอมอำนาจของการทดสอบทั้งกรณีที่  $\Delta^2$  น้อยกว่า 1 และ  $\Delta^2$  มากกว่า 1 โดยแต่ละตารางจะนำเสนอมอำนาจของการทดสอบเมื่อ  $\alpha$  มีค่า .01, .05 และ .10 ซึ่งอำนาจของการทดสอบทั้ง 5 วิธีนี้จะนำเสนอด้วยตาราง 4.8-4.13 เมื่อประชากรมีการแจกแจงแบบยูนิฟอร์ม และตาราง 4.15-4.20 เมื่อประชากรมีการแจกแจงแบบปกติ และตาราง 4.22-4.27 เมื่อประชากรมีการแจกแจงแบบโลจิสติก

จากค่าอำนาจของการทดสอบทั้ง 5 วิธี ซึ่งนำเสนอบ้างเป็นตารางแล้วจะสรุปเป็นจำนวนครั้งที่การทดสอบแต่ละวิธีมีอำนาจของการทดสอบสูงที่สุด ซึ่งจำแนกตามระดับนัยสำคัญ และลักษณะของขนาดของตัวอย่าง คือ ขนาดของตัวอย่างเท่ากัน และไม่เท่ากัน โดยจะนำเสนอด้วยตารางที่ 4.14 เมื่อประชากรมีการแจกแจงแบบยูนิฟอร์ม ตาราง 4.21 เมื่อประชากรมีการแจกแจงแบบปกติ และตาราง 4.28 เมื่อประชากรมีการแจกแจงแบบโลจิสติก

นอกจากนี้ก็จะเปรียบเทียบค่า ARE กับค่าอำนาจของการทดสอบสองวิธี สำหรับแต่ละรูปแบบของการแจกแจงเพื่อที่จะพิจารณาว่าในกรณีของขนาดของตัวอย่างที่ใช้ในการศึกษาอำนาจของการทดสอบในการวิจัยครั้งนี้ ค่า ARE เป็นเกณฑ์ที่ดีที่จะบ่งชี้ว่าการทดสอบใดจะมีอำนาจของการทดสอบสูงกว่ากันหรือไม่

1.2 อำนาจของการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) จะนำเสนอในรูปของตาราง ซึ่งการศึกษาอำนาจของการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) นั้น จะศึกษาในกรณีที่การทดสอบนี้ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เท่านั้น ซึ่งปรากฏว่าจากผลการทดลอง การทดสอบสแควร์แรงค์ ( $\bar{X}$ ) ควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี เมื่อขนาดของตัวอย่างทั้งสอง

กลุ่มเท่ากันคือ (10:10) (15:15) และ (20:20) การนำเสนอค่าอำนาจของการทดสอบ ซึ่งนำเสนอเฉพาะกรณีที่ขนาดของตัวอย่างทั้งสองกลุ่มเท่ากันเท่านั้น จะนำเสนอด้วยตาราง 4.29

2. ในกรณีที่ประชากรมีการแจกแจงแบบเบ้ ซึ่งศึกษาเฉพาะการทดสอบเอฟ และการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) และจะศึกษาเปรียบเทียบอำนาจของการทดสอบทั้งสองวิธี เฉพาะกรณีที่การทดสอบทั้งสองกรณีควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เท่ากัน ซึ่งปรากฏว่าจากผลการทดลองการทดสอบทั้งสองวิธีควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี เมื่อขนาดของตัวอย่างทั้งสองกลุ่มเท่ากัน โดยค่าความเบ้เป็น .25 และ .50 และความโด่งมีค่า 2.8, 3.0 และ 3.2 สำหรับกรณีที่ขนาดของตัวอย่างไม่เท่ากันนั้นเมื่อขนาดของตัวอย่างเป็น (5:10) มีบางกรณีที่การทดสอบทั้งสองวิธีควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เท่ากัน ซึ่งนำเสนออำนาจของการทดสอบเฉพาะค่าความเบ้และความโด่งดังกล่าว สำหรับกรณีที่ความเบ้และความโด่งมีค่าอื่น มีบางกรณีที่การทดสอบทั้งสองวิธีควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ แต่ก็มีความน้อยมาก และผลสรุปที่ได้ก็ทำนองเดียวกัน จึงไม่นำเสนออำนาจของการทดสอบในกรณีดังกล่าว

สำหรับการนำเสนออำนาจของการทดสอบทั้งสองวิธีนี้จะนำเสนอเป็นตารางและกราฟ ซึ่งกรณีที่ขนาดของตัวอย่างทั้งสองกลุ่มเป็น (10:10) ความเบ้ .25 และ .50 และความโด่ง 2.8, 3.0 และ 3.2 จะนำเสนอด้วยตารางที่ 4.30-4.31 และรูปที่ 4.136-4.153 สำหรับกรณีที่ขนาดของตัวอย่างทั้งสองกลุ่มเป็น (15:15) สำหรับความเบ้และความโด่งดังกล่าวนำเสนอด้วยตารางที่ 4.32-4.33 และรูปที่ 4.154-4.171 และในกรณีที่ขนาดของตัวอย่างทั้งสองกลุ่มเป็น (20:20) จะนำเสนอด้วยตารางที่ 4.34-4.35 และรูปที่ 4.172-4.189 สำหรับกรณีที่ขนาดของตัวอย่างทั้งสองกลุ่มเป็น (5:10) นำเสนอด้วยตารางที่ 4.36 และรูปที่ 4.190-4.192 โดยที่การนำเสนออำนาจของการทดสอบในกรณีที่ขนาดของตัวอย่างเท่ากัน นำเสนอเฉพาะกรณีที่  $\Delta^2$  มากกว่า 1 สำหรับกรณีที่  $\Delta^2$  น้อยกว่า 1 ให้ผลสรุปทำนองเดียวกัน

ซึ่งรายละเอียดในการนำเสนออำนาจของการทดสอบต่าง ๆ แสดงได้ดังนี้

ตารางที่ 4.8 แสดงอำนาจของการทดสอบเอฟ การทดสอบคล็อกซ์ การทดสอบทูกี-ซีเกล การทดสอบสแควร์แรงค์ ( $\mu$ ) และการทดสอบมัต เมื่อประชากรมีการแจกแจงแบบยูนิฟอร์ม และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (10:10) จำแนกตามระดับนัยสำคัญที่กำหนด

ระดับนัยสำคัญ	สถิติทดสอบ	$\Delta^2$							
		1	2	3	4	5	6	7	8
.01	F	.000	.010	.056	.158	.269	.427	.558	.667
	KN	.006	.071	.186	.293	.368	.431	.492	.530
	TS	.011	.049	.119	.196	.272	.317	.359	.401
	SRT ( $\mu$ )	.009	.083	.226	.337	.433	.518	.571	.620
	M	.008	.065	.176	.267	.356	.414	.466	.508
.05	F	.005	.076	.221	.492	.680	.798	.879	.922
	KN	.043	.255	.453	.583	.671	.746	.771	.808
	ST	.040	.162	.311	.434	.512	.582	.623	.654
	SRT ( $\mu$ )	.038	.259	.449	.608	.711	.771	.816	.853
	M	.048	.218	.401	.535	.624	.693	.730	.767
.10	F	.022	.176	.444	.690	.822	.908	.952	.970
	KN	.100	.375	.592	.716	.784	.828	.856	.876
	TS	.085	.250	.440	.565	.638	.702	.742	.785
	SRT ( $\mu$ )	.097	.362	.598	.734	.815	.870	.903	.927
	M	.091	.338	.540	.655	.745	.799	.828	.852



ตารางที่ 4.9 แสดงอำนาจของการทดสอบเอฟ การทดสอบคล็อทซ์ การทดสอบซูที้-ซีเกล การทดสอบสแควร์แรงค์(μ) และการทดสอบมัต เมื่อประชากรมีการแจกแจงแบบยูนิฟอร์ม และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (15:15) จำแนกตามระดับนัยสำคัญที่กำหนด

ระดับนัยสำคัญ	สถิติ ทดสอบ	$\Delta^2$							
		1	2	3	4	5	6	7	8
.01	F	.002	.030	.161	.425	.659	.816	.910	.965
	KN	.004	.116	.318	.487	.604	.692	.744	.796
	TS	.009	.107	.248	.371	.479	.576	.642	.691
	SRT (μ)	.007	.138	.367	.546	.681	.764	.822	.861
	M	.009	.124	.310	.473	.593	.682	.728	.781
.05	F	.011	.127	.517	.787	.929	.979	.991	.997
	KN	.046	.399	.675	.825	.889	.927	.949	.962
	TS	.041	.245	.457	.611	.735	.792	.841	.863
	SRT (μ)	.043	.373	.643	.812	.890	.922	.950	.964
	M	.041	.387	.585	.753	.834	.889	.914	.933
.10	F	.022	.278	.706	.906	.980	.994	.998	.999
	KN	.098	.571	.813	.903	.942	.965	.975	.982
	TS	.090	.369	.594	.758	.837	.875	.909	.930
	SRT (μ)	.092	.506	.776	.891	.940	.960	.980	.986
	M	.092	.456	.721	.851	.910	.943	.956	.972



ตารางที่ 4.10 แสดงอำนาจของการทดสอบเอฟ การทดสอบคล็อทซ์ การทดสอบทูกี้-ซีเกล การทดสอบสแควร์แรงค์ ( $\mu$ ) และการทดสอบมูต์ เมื่อประชากรมีการแจกแจงแบบยูนิฟอร์ม และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (20:20) จำแนกตามระดับนัยสำคัญที่กำหนด

ระดับนัยสำคัญ	สถิติทดสอบ	$\Delta^2$							
		1	2	3	4	5	6	7	8
.01	F	.000	.048	.339	.700	.892	.969	.992	.996
	KN	.008	.250	.576	.747	.858	.901	.934	.947
	Ts	.008	.132	.373	.562	.687	.761	.826	.867
	SRT ( $\mu$ )	.010	.223	.558	.742	.850	.913	.936	.954
	M	.009	.195	.495	.679	.802	.874	.908	.930
.05	F	.007	.209	.720	.938	.989	.997	.998	1.000
	KN	.034	.585	.857	.935	.971	.979	.988	.989
	TS	.049	.353	.637	.792	.877	.917	.942	.956
	SRT ( $\mu$ )	.036	.500	.806	.918	.954	.979	.987	.992
	M	.041	.445	.771	.892	.931	.960	.974	.980
.10	F	.015	.406	.866	.977	.996	.998	1.000	1.000
	KN	.078	.743	.927	.981	.990	.994	.997	.999
	TS	.086	.479	.637	.875	.927	.950	.967	.973
	SRT ( $\mu$ )	.081	.625	.883	.950	.980	.989	.994	.997
	M	.082	.606	.865	.938	.969	.982	.986	.991

จากตาราง 4.8 -4.10 ซึ่งแสดงค่าอำนาจของการทดสอบเอฟ การทดสอบคล้อย การทดสอบมุต การทดสอบทูกี้-ซีเกล และการทดสอบสแควร์แรงค์ ( $\chi^2$ ) เมื่อประชากรมีการแจกแจงแบบยูนิฟอร์ม ขนาดของตัวอย่างเป็น (10:10) (15:15) และ (20:20) และระดับนัยสำคัญมีค่า .01 .05 และ .10

เนื่องจากการทดสอบเอฟควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้น้อยมาก โดยเฉพาะเมื่อใช้เกณฑ์ของ Cochran การทดสอบเอฟไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ และเมื่อใช้เกณฑ์ของ Bradley สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เพียง 1 ครั้ง จากการทดลองทั้งหมด 18 ครั้ง จึงไม่น่ามาเปรียบเทียบอำนาจของการทดสอบกับการทดสอบอื่น ๆ อีก 4 วิธี ซึ่งค่าอำนาจการทดสอบทั้ง 4 วิธีสามารถเปรียบเทียบได้ดังนี้

ที่ระดับนัยสำคัญ .01 การทดสอบสแควร์แรงค์มีอำนาจของการทดสอบสูงที่สุดเมื่อขนาดของตัวอย่างเป็น (10:10) และ (15:15) แต่เมื่อขนาดของตัวอย่างเป็น (20:20) เมื่อ  $\Delta^2=2$  และ 3 การทดสอบคล้อยมีอำนาจของการทดสอบสูงกว่า แต่เมื่อ  $\Delta^2$  เป็น 4, 5, 6, 7 และ 8 การทดสอบทั้ง 2 วิธีมีอำนาจของการทดสอบแตกต่างกันมากที่สุด .012 สำหรับอำนาจของการทดสอบมุตนั้นเมื่อเปรียบเทียบกับ การทดสอบคล้อยแล้ว จะมีอำนาจของการทดสอบต่ำกว่าการทดสอบคล้อยเมื่อขนาดของตัวอย่างเป็น (10:10) และ (20:20) สำหรับกรณีที่ยกขนาดของตัวอย่างเป็น (15:15) เมื่อ  $\Delta^2=2$  อำนาจของการทดสอบ 2 วิธีนี้แตกต่างกันมากที่สุด .008 แต่กรณี  $\Delta^2$  เป็นค่าอื่น ๆ การทดสอบคล้อยมีอำนาจของการทดสอบสูงกว่า ทั้งนี้การทดสอบทูกี้-ซีเกลจะมีอำนาจของการทดสอบต่ำที่สุดเมื่อเทียบกับการทดสอบอื่น ๆ

ที่ระดับนัยสำคัญ .05 ในกรณีที่ขนาดของตัวอย่างเป็น (10:10) และ  $\Delta^2 = 4, 5, 6, 7$  และ 8 การทดสอบสแควร์แรงค์มีอำนาจของการทดสอบสูงที่สุด แต่เมื่อ  $\Delta^2 = 2$  และ 3 การทดสอบคล้อย และการทดสอบสแควร์แรงค์มีอำนาจของการทดสอบใกล้เคียงกันมาก โดยมีความแตกต่างของอำนาจของการทดสอบทั้ง 2 วิธีมีค่าสูงที่สุด .004 สำหรับกรณีที่ขนาดของตัวอย่างเป็น (15:15) เมื่อ  $\Delta^2 = 2, 3, 4$  การทดสอบคล้อยมีอำนาจของการทดสอบสูงที่สุด แต่กรณี  $\Delta^2 = 5, 6, 7$  และ 8 การทดสอบคล้อยและการทดสอบสแควร์แรงค์มีอำนาจของการทดสอบใกล้เคียงกันมาก ซึ่งค่าความแตกต่างของอำนาจการทดสอบทั้ง 2 วิธีมีค่าสูงที่สุด .005 และเมื่อขนาดของตัวอย่างเป็น (20:20) และ  $\Delta^2 = 2, 3, 4$  และ 5 การทดสอบ

คล้อทซ์มีอำนาจของการทดสอบมากกว่าการทดสอบสแควร์แรงค์ แต่เมื่อ  $\Delta^2 = 6, 7$  และ  $8$  การทดสอบสแควร์แรงค์และการทดสอบคล้อทซ์มีอำนาจของการทดสอบใกล้เคียงกันมาก โดยค่าความแตกต่างของอำนาจของการทดสอบทั้ง 2 วิธีมีค่าสูงที่สุด .003 สำหรับการทดสอบมัดและการทดสอบทูกี้-ซีเกล มีอำนาจของการทดสอบต่ำกว่าการทดสอบคล้อทซ์และการทดสอบสแควร์แรงค์ โดยเฉพาะการทดสอบทูกี้-ซีเกลจะมีอำนาจของการทดสอบต่ำที่สุดเมื่อเทียบกับการทดสอบอื่น ๆ

ที่ระดับนัยสำคัญ .10 เมื่อเปรียบเทียบอำนาจของการทดสอบทั้ง 4 วิธีแล้ว ปรากฏว่า เมื่อขนาดของตัวอย่างเป็น  $(10:10)$  และ  $\Delta^2 = 4, 5, 6, 7$  และ  $8$  การทดสอบสแควร์แรงค์มีอำนาจของการทดสอบสูงกว่าการทดสอบคล้อทซ์ สำหรับขนาดของตัวอย่างเป็น  $(15:15)$  และ  $\Delta^2 = 2, 3$  และ  $4$  การทดสอบคล้อทซ์มีอำนาจของการทดสอบสูงกว่าการทดสอบสแควร์แรงค์ แต่เมื่อ  $\Delta^2 = 5, 6, 7$  และ  $8$  การทดสอบทั้ง 2 วิธีมีอำนาจของการทดสอบใกล้เคียงกันมาก โดยค่าความแตกต่างระหว่างอำนาจของการทดสอบทั้งสองวิธีมีค่าสูงที่สุด .005 และขนาดของตัวอย่างเป็น  $(20:20)$  เมื่อ  $\Delta^2 = 2, 3, 4$  และ  $5$  การทดสอบคล้อทซ์มีอำนาจของการทดสอบสูงกว่าการทดสอบสแควร์แรงค์ แต่เมื่อ  $\Delta^2 = 6, 7$  และ  $8$  อำนาจของการทดสอบทั้ง 2 วิธีใกล้เคียงกันมาก ซึ่งค่าความแตกต่างของอำนาจของการทดสอบทั้ง 2 วิธีมีค่าสูงที่สุด .005 สำหรับการทดสอบมัด และการทดสอบทูกี้-ซีเกล มีอำนาจของการทดสอบต่ำกว่าการทดสอบคล้อทซ์และการทดสอบสแควร์แรงค์ โดยที่การทดสอบทูกี้-ซีเกลมีอำนาจของการทดสอบต่ำที่สุดเมื่อเทียบกับการทดสอบอื่น ๆ

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.11 แสดงอำนาจของการทดสอบเอฟ การทดสอบคล็อทซ์ การทดสอบบูทซ์-ซีเกล การทดสอบสแควร์แรงค์ (M) และการทดสอบมูด์ เมื่อประชากรมีการแจกแจงแบบยูนิฟอร์ม และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (5:10) สำเนาตามระดับนัยสำคัญที่กำหนด

ระดับนัยสำคัญ	สถิติทดสอบ	$\Delta^2$														
		$\frac{1}{8}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	1	2	3	4	5	6	7	8
.01	F	.487	.407	.313	.212	.111	.045	.009	.003	.009	.019	.032	.050	.071	.095	.127
	KN	.274	.240	.215	.186	.152	.100	.051	.011	.027	.047	.067	.092	.110	.132	.148
	TS	.237	.194	.165	.135	.102	.066	.034	.010	.028	.045	.062	.081	.102	.115	.124
	SRT ( $\mu$ )	.367	.334	.291	.253	.187	.126	.057	.010	.033	.079	.124	.172	.212	.257	.304
	M	.194	.162	.145	.116	.087	.059	.029	.009	.027	.047	.067	.092	.110	.132	.148
.05	F	.750	.687	.612	.518	.384	.216	.062	.016	.048	.101	.166	.236	.314	.387	.459
	KN	.571	.538	.494	.454	.381	.296	.185	.048	.108	.214	.249	.370	.415	.461	.499
	TS	.441	.418	.386	.324	.262	.189	.106	.038	.081	.155	.218	.270	.308	.352	.384
	SRT ( $\mu$ )	.648	.611	.563	.513	.438	.333	.184	.061	.132	.263	.392	.495	.564	.624	.663
	M	.568	.530	.489	.438	.360	.189	.167	.048	.095	.191	.271	.343	.392	.434	.471
.10	F	.833	.803	.739	.657	.541	.370	.153	.033	.100	.189	.310	.415	.517	.611	.689
	KN	.711	.678	.642	.609	.558	.467	.307	.098	.206	.359	.464	.533	.575	.615	.641
	TS	.596	.570	.523	.489	.425	.335	.204	.088	.177	.293	.398	.480	.520	.567	.599
	SRT ( $\mu$ )	.753	.728	.685	.638	.566	.461	.297	.106	.238	.419	.563	.659	.727	.772	.803
	M	.638	.614	.559	.520	.467	.371	.229	.083	.181	.324	.444	.519	.561	.602	.630

ตารางที่ 4.12 แสดงอำนาจของการทดสอบเอฟ การทดสอบคล้อยหลัง การทดสอบซูก็-ซีเกล การทดสอบสันคเวร์แรงค์ (μ) และการทดสอบมูด เมื่อประชากรมีการแจกแจงแบบยูนิฟอร์ม และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (5:15) จำแนกตามระดับนัยสำคัญที่กำหนด

ระดับนัยสำคัญ	สถิติ ทดสอบ	$\Delta^2$															
		$\frac{1}{8}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	1	2	3	4	5	6	7	8	
.01	F	.646	.576	.491	.335	.205	.079	.014	.006	.011	.023	.049	.066	.087	.111	.132	
	KN	.443	.408	.358	.309	.249	.156	.063	.004	.024	.058	.087	.124	.166	.200	.222	
	TS	.307	.275	.251	.200	.154	.102	.052	.007	.022	.048	.067	.101	.133	.162	.178	
	SRT (μ)	.535	.494	.435	.385	.304	.196	.081	.009	.028	.068	.130	.184	.248	.302	.351	
	M	.391	.358	.321	.272	.215	.130	.059	.005	.023	.056	.084	.120	.161	.193	.215	
.05	F	.811	.770	.717	.629	.506	.280	.091	.017	.061	.111	.179	.246	.317	.397	.462	
	KN	.690	.652	.612	.583	.507	.420	.241	.053	.120	.246	.352	.431	.493	.543	.586	
	TS	.531	.516	.458	.413	.350	.257	.134	.047	.102	.195	.275	.341	.399	.445	.492	
	SRT (μ)	.689	.669	.623	.582	.498	.402	.227	.056	.135	.295	.437	.554	.648	.719	.772	
	M	.613	.594	.542	.489	.426	.335	.189	.050	.121	.239	.339	.413	.473	.521	.567	
.10	F	.871	.832	.790	.735	.632	.471	.193	.042	.111	.206	.308	.422	.529	.660	.753	
	KN	.756	.732	.703	.665	.607	.524	.358	.095	.232	.412	.540	.633	.709	.746	.769	
	TS	.669	.650	.605	.561	.495	.416	.273	.097	.218	.347	.444	.534	.605	.647	.679	
	SRT (μ)	.784	.756	.721	.681	.608	.509	.340	.100	.250	.464	.630	.748	.806	.860	.879	
	M	.706	.686	.646	.607	.545	.459	.299	.087	.225	.379	.511	.504	.683	.722	.746	

ตารางที่ 4.13 แสดงอำนาจของการทดสอบเอฟ การทดสอบคลอท์ การทดสอบทุ๊ก-ซีเกล การทดสอบสแควร์แรงค์ ( $\mu$ ) และการทดสอบมูด เมื่อประชากรมีการแจกแจงแบบยูนิฟอร์ม และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (5:20) ค่าแนกตามระดับนัยสำคัญที่กำหนด

ระดับนัยสำคัญ	สถิติทดสอบ	$\Delta^2$														
		$\frac{1}{8}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	1	2	3	4	5	6	7	8
.01	F	.721	.669	.571	.447	.293	.115	.023	.004	.010	.023	.036	.050	.083	.104	.126
	KN	.559	.528	.465	.400	.330	.240	.102	.012	.030	.073	.105	.149	.198	.234	.278
	TS	.369	.344	.300	.252	.206	.142	.066	.011	.023	.055	.097	.141	.174	.216	.246
	SRT( $\mu$ )	.576	.535	.478	.403	.316	.218	.096	.011	.030	.077	.137	.214	.286	.351	.407
	M	.491	.460	.404	.340	.270	.187	.085	.010	.029	.067	.105	.153	.194	.229	.279
.05	F	.851	.740	.707	.655	.575	.455	.283	.015	.047	.105	.161	.133	.305	.387	.469
	KN	.758	.740	.707	.655	.575	.455	.283	.044	.128	.269	.387	.477	.567	.632	.677
	TS	.579	.548	.513	.450	.389	.287	.167	.046	.104	.196	.285	.377	.451	.502	.543
	SRT( $\mu$ )	.754	.722	.680	.609	.524	.413	.234	.056	.131	.303	.455	.590	.674	.736	.784
	M	.685	.657	.614	.547	.476	.372	.223	.051	.128	.263	.375	.470	.560	.623	.658
.10	F	.891	.867	.830	.772	.683	.511	.224	.038	.103	.185	.296	.413	.539	.651	.748
	KN	.839	.816	.792	.752	.696	.568	.410	.101	.252	.442	.578	.675	.735	.778	.808
	TS	.717	.682	.638	.582	.515	.405	.281	.102	.200	.351	.479	.566	.639	.693	.727
	SRT( $\mu$ )	.849	.822	.789	.739	.672	.535	.353	.097	.254	.483	.645	.744	.819	.862	.907
	M	.799	.771	.740	.692	.612	.494	.335	.101	.256	.428	.564	.663	.727	.767	.796



จากตาราง 4.11-4.13 ซึ่งแสดงค่าอำนาจของการทดสอบเอฟ การทดสอบคล็อทซ์ การทดสอบมุต การทดสอบทูกี้-ซีเกล และการทดสอบสแควร์แรงค์ ( $\mu$ ) เมื่อประชากรมีการแจกแจงแบบยูนิฟอร์ม ขนาดของตัวอย่างเป็น (5:10) (5:15) และ (5:20) และระดับนัยสำคัญมีค่าเป็น .01 .05 และ .10

เนื่องจากการทดสอบเอฟควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้น้อยมาก โดยเฉพาะเมื่อใช้เกณฑ์ของ Cochran การทดสอบเอฟไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ และเมื่อใช้เกณฑ์ของ Bradley สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ 1 ครั้ง จากการทดลองทั้งหมด 18 ครั้ง จึงไม่น่ามาพิจารณาเปรียบเทียบอำนาจของการทดสอบกับการทดสอบอื่น ๆ อีก 4 วิธี มีค่าอำนาจของการทดสอบทั้ง 4 วิธี สามารถเปรียบเทียบได้ดังนี้

ที่ระดับนัยสำคัญ .01 การทดสอบสแควร์แรงค์มีอำนาจของการทดสอบสูงกว่าการทดสอบอื่น ๆ เมื่อขนาดของตัวอย่างเป็น (5:10) และ (5:15) แต่เมื่อขนาดของตัวอย่างเป็น (5:20) และ  $\Delta^2 = 2, 3, \frac{1}{2}$  และ  $\frac{1}{5}$  การทดสอบคล็อทซ์และการทดสอบสแควร์แรงค์มีอำนาจของการทดสอบใกล้เคียงกัน โดยมีค่าอำนาจของการทดสอบแตกต่างกันมากที่สุด .004 แต่เมื่อ  $\Delta^2 = \frac{1}{3}$  และ  $\frac{1}{4}$  การทดสอบคล็อทซ์มีอำนาจของการทดสอบสูงกว่า ในกรณีที่  $\Delta^2$  มีค่าอื่น ๆ การทดสอบสแควร์แรงค์มีอำนาจของการทดสอบสูงกว่าสำหรับการทดสอบคล็อทซ์ การทดสอบมุต และการทดสอบทูกี้-ซีเกลมีอำนาจของการทดสอบต่ำกว่าการทดสอบสแควร์แรงค์ ยกเว้นบางจุดของ  $\Delta^2$  เท่านั้นที่การทดสอบคล็อทซ์มีอำนาจของการทดสอบสูงกว่า เมื่อเปรียบเทียบอำนาจของการทดสอบทั้ง 3 วิธีนี้ อำนาจของการทดสอบคล็อทซ์จะสูงที่สุดเมื่อ  $\Delta^2$  น้อยกว่า 1 แต่เมื่อ  $\Delta^2$  มากกว่า 1 การทดสอบคล็อทซ์มีอำนาจของการทดสอบสูงกว่าการทดสอบทูกี้-ซีเกล ยกเว้นกรณีที่  $\Delta^2 = 2$  อำนาจของการทดสอบทั้ง 2 วิธีมีค่าแตกต่างกันมากที่สุด .007 สำหรับการทดสอบคล็อทซ์และการทดสอบมุตมีอำนาจของการทดสอบใกล้เคียงกันมีความแตกต่างของอำนาจของการทดสอบสูงที่สุด .007

ที่ระดับนัยสำคัญ .05 การทดสอบสแควร์แรงค์มีอำนาจของการทดสอบสูงกว่า การทดสอบอื่น ๆ เมื่อขนาดของตัวอย่างเป็น (5:10) (5:15) และ (5:20) และ  $\Delta^2$  มีค่ามากกว่า 1 แต่เมื่อ  $\Delta^2$  น้อยกว่า 1 เมื่อขนาดของตัวอย่างเป็น (5:10)  $\Delta^2 = \frac{1}{2}$  อำนาจของการทดสอบคล็อทซ์และการทดสอบสแควร์แรงค์ใกล้เคียงกันมาก โดยมีความแตกต่างของอำนาจของการทดสอบสูงที่สุด .001 แต่เมื่อ  $\Delta^2$  มีค่าอื่น ๆ การทดสอบสแควร์แรงค์มีอำนาจของการทดสอบสูงกว่า และเมื่อขนาดของตัวอย่างเป็น (5:15)  $\Delta^2 = \frac{1}{2}$  และ  $\frac{1}{3}$

การทดสอบคล้อยทซ์มีอำนาจของการทดสอบสูงกว่า แต่ในกรณีที่  $\Delta^2 = \frac{1}{4}, \frac{1}{5}$  และ  $\frac{1}{8}$  อำนาจของการทดสอบทั้ง 2 วิธีนี้แตกต่างกันน้อย โดยมีความแตกต่างของอำนาจการทดสอบสูงที่สุด .009 และเมื่อ  $\Delta^2 = \frac{1}{6}$  และ  $\frac{1}{7}$  การทดสอบสแควร์แรงค์มีอำนาจของการทดสอบสูงกว่า สำหรับกรณีที่ขนาดของตัวอย่างเป็น (5:20) และ  $\Delta^2$  น้อยกว่า 1 อำนาจของการทดสอบคล้อยทซ์สูงกว่าการทดสอบสแควร์แรงค์ ในกรณีของการทดสอบมัตและการทดสอบทูกี้-ซีเกล นั้น โดยส่วนมากแล้วจะมีอำนาจของการทดสอบต่ำกว่าการทดสอบคล้อยทซ์และการทดสอบสแควร์แรงค์ โดยที่การทดสอบทูกี้-ซีเกลจะมีอำนาจของการทดสอบต่ำกว่าการทดสอบอื่น ๆ

ที่ระดับนัยสำคัญ .10 การทดสอบสแควร์แรงค์มีอำนาจของการทดสอบสูงกว่าการทดสอบอื่น ๆ เมื่อขนาดของตัวอย่างเป็น (5:10) (5:15) และ (5:20) และ  $\Delta^2$  มีค่ามากกว่า 2 แต่เมื่อ  $\Delta^2$  น้อยกว่า 1 ในกรณีที่ขนาดของตัวอย่างเป็น (5:10)  $\Delta^2 = \frac{1}{2}, \frac{1}{3}$  และ  $\frac{1}{4}$  ขนาดของตัวอย่างเป็น (5:15)  $\Delta^2 = \frac{1}{4}$  และขนาดของตัวอย่างเป็น (5:20)  $\Delta^2 = \frac{1}{6}, \frac{1}{7}$  และ  $\frac{1}{8}$  การทดสอบคล้อยทซ์และการทดสอบสแควร์แรงค์มีอำนาจของการทดสอบแตกต่างกันเพียงเล็กน้อย โดยมีความแตกต่างของอำนาจของการทดสอบสูงที่สุด .010 และเมื่อขนาดของตัวอย่างเป็น (5:15)  $\Delta^2 = \frac{1}{2}$  และ  $\frac{1}{3}$  ขนาดของตัวอย่างเป็น (5:20)  $\Delta^2 = \frac{1}{2}, \frac{1}{3}, \frac{1}{4}$  และ  $\frac{1}{5}$  การทดสอบคล้อยทซ์จะมีอำนาจของการทดสอบสูงกว่า สำหรับกรณีที่ขนาดของตัวอย่างเป็น (5:10) (5:15) และ (5:20) เมื่อ  $\Delta^2$  มีค่าอื่น ๆ การทดสอบสแควร์แรงค์มีอำนาจของการทดสอบสูงกว่า ส่วนการทดสอบมัตมีอำนาจของการทดสอบต่ำกว่าการทดสอบสแควร์แรงค์ เมื่อเปรียบเทียบอำนาจของการทดสอบมัตและการทดสอบคล้อยทซ์ปรากฏว่าเมื่อ  $\Delta^2$  น้อยกว่า 1 การทดสอบคล้อยทซ์มีอำนาจของการทดสอบสูงกว่า แต่เมื่อ  $\Delta^2$  มากกว่า 1 การทดสอบคล้อยทซ์มีอำนาจของการทดสอบสูงกว่าการทดสอบมัต เมื่อขนาดของตัวอย่างเป็น (5:10) แต่กรณีที่ขนาดของตัวอย่างเป็น (5:15) (5:20) เมื่อ  $\Delta^2 = 2$  อำนาจของการทดสอบทั้ง 2 วิธีแตกต่างกันน้อยมาก โดยค่าความแตกต่างสูงที่สุดมีค่า .007 และเมื่อ  $\Delta^2 = 3, 4, 5, 6, 7$  และ 8 การทดสอบคล้อยทซ์มีอำนาจของการทดสอบสูงกว่าการทดสอบมัต สำหรับการทดสอบทูกี้-ซีเกล มีอำนาจการทดสอบต่ำที่สุดเมื่อเทียบกับการทดสอบอื่น ๆ

จากค่าอำนาจของการทดสอบ เมื่อขนาดของตัวอย่างเป็น (10:10) (15:15) (20:20) (5:10) (5:15) และ (5:20) สรุปจำนวนครั้งที่การทดสอบแต่ละวิธีมีอำนาจของการทดสอบสูงที่สุด ได้ดังตาราง 4.14



ตาราง 4.14 แสดงจำนวนครั้งที่การทดสอบคล้อทซ์ การทดสอบทูกี้-ซีเกล และการทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูงที่สุด เมื่อประชากรมีการแจกแจงแบบยูนิฟอร์ม จากการทดลองทั้งหมด 21 กรณี เมื่อขนาดของตัวอย่างทั้งสองกลุ่มเป็น (10:10) (15:15) และ (20:20) และ 21 กรณี เมื่อขนาดของตัวอย่างทั้งสองกลุ่มเป็น (5:10) (5:15) และ (5:20) เมื่อ  $\Delta^2$  น้อยกว่า 1 และ 21 กรณี เมื่อ  $\Delta^2$  มากกว่า 1 จำแนกตามระดับนัยสำคัญที่กำหนด

ระดับนัยสำคัญ	สถิติทดสอบ	ขนาดของตัวอย่าง				
		(10:10)	(15:15)	(20:20)	(5:10) (5:15) (5:20)	
					$\Delta^2 < 1$	$\Delta^2 > 1$
.01	KN		4		3	1
	TS		0		0	0
	SRT( $\mu$ )		17		18	21
	M		0		0	0
.05	KN		11		13	0
	TS		0		0	0
	SRT( $\mu$ )		11		8	21
	M		0		0	0
.10	KN		13		9	0
	TS		0		0	0
	SRT( $\mu$ )		8		12	20
	M		0		0	1

จากตารางที่ 4.14 สรุปได้ดังนี้

1. ในกรณีที่ขนาดของตัวอย่างเท่ากัน การทดสอบคล้อทซ์และการทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูงที่สุด มากกว่าการทดสอบอื่น กล่าวคือการทดสอบคล้อทซ์มีอำนาจของการทดสอบสูงที่สุด 4 กรณี เมื่อ  $\alpha = .01$  11 กรณี เมื่อ  $\alpha = .05$  และ 13 กรณีเมื่อ  $\alpha = .10$  ในขณะที่การทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูงที่สุด 17 กรณี



เมื่อ  $\alpha = .01$  11 กรณีเมื่อ  $\alpha = .05$  และ 8 กรณีเมื่อ  $\alpha = .10$

2. กรณีที่ขนาดของตัวอย่างไม่เท่ากัน การทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูงที่สุด มากที่สุดเมื่อ  $\Delta^2$  มากกว่า 1 คือ มีอำนาจของการทดสอบสูงที่สุด 21 กรณีเมื่อ  $\alpha = .01$  และ .05 และ 20 กรณีเมื่อ  $\alpha = .10$  แต่เมื่อ  $\Delta^2$  น้อยกว่า 1 เมื่อ  $\alpha = .01$  การทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูงที่สุดมากกว่าการทดสอบอื่น ๆ คือ มีอำนาจของการทดสอบสูงที่สุด 18 กรณี และเมื่อ  $\alpha = .05$  และ .10 การทดสอบสแควร์แรงค์ ( $\mu$ ) และการทดสอบคล็อทซ์มีอำนาจสูงที่สุดมากกว่าการทดสอบอื่น ๆ โดยที่การทดสอบคล็อทซ์มีอำนาจของการทดสอบสูงที่สุด 13 กรณี เมื่อ  $\alpha = .05$  และ 9 กรณีเมื่อ  $\alpha = .10$  ในขณะที่การทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูงที่สุด 8 กรณี เมื่อ  $\alpha = .05$  และ 12 กรณีเมื่อ  $\alpha = .10$

เปรียบเทียบค่า ARE กับค่าอำนาจของการทดสอบเมื่อประชากรมีการแจกแจงแบบยูนิฟอร์ม

จากค่าอำนาจของการทดสอบทั้ง 4 วิธี คือ การทดสอบคล็อทซ์ การทดสอบมัต การทดสอบทูกี-ซีเกล และการทดสอบสแควร์แรงค์ ( $\mu$ ) เมื่อประชากรมีการแจกแจงแบบยูนิฟอร์ม นำมาเปรียบเทียบกับค่า ARE สำหรับแต่ละคู่ของการทดสอบได้ดังนี้

การทดสอบคล็อทซ์กับการทดสอบมัต มีค่า ARE เท่ากับ 0. จากค่า ARE แสดงว่าอำนาจของการทดสอบคล็อทซ์น่าจะสูงกว่าอำนาจของการทดสอบมัต ซึ่งจากผลการทดลองเมื่อขนาดของตัวอย่างเป็น (10:10) (15:15) และ (20:20) โดยส่วนมากแล้วการทดสอบคล็อทซ์มีอำนาจของการทดสอบสูงกว่าการทดสอบมัต และในกรณีที่ขนาดของตัวอย่างเป็น (5:10) (5:15) และ (5:20) ปรากฏว่าเมื่อ  $\Delta^2$  น้อยกว่า 1 การทดสอบคล็อทซ์จะมีอำนาจของการทดสอบสูงกว่า แต่เมื่อ  $\Delta^2$  มากกว่า 1 เมื่อ  $\alpha = .01$  อำนาจของการทดสอบทั้งสองวิธีใกล้เคียงกันมาก แต่เมื่อ  $\alpha = .05$  และ .10 การทดสอบคล็อทซ์จะมีอำนาจของการทดสอบสูงกว่าเป็นส่วนใหญ่ ยกเว้นกรณีที่  $\Delta^2 = 2$  ซึ่งอำนาจของการทดสอบทั้งสองวิธีจะใกล้เคียงกัน

การทดสอบคล้อยกับการทดสอบทูกี้-ซีเกล มีค่า ARE เท่ากับ 0. จากค่า ARE แสดงว่าอำนาจของการทดสอบคล้อยควรจะสูงกว่าอำนาจของการทดสอบทูกี้-ซีเกล ซึ่งจากผลการทดลองเมื่อขนาดของตัวอย่างเป็น (10:10) (15:15) (20:20) (5:10) (5:15) และ (5:20) ปรากฏว่าโดยส่วนมากการทดสอบคล้อยมีอำนาจสูงกว่าการทดสอบทูกี้-ซีเกล

การทดสอบคล้อยกับการทดสอบสแควร์แรงค์ ( $\mu$ ) มีค่า ARE เท่ากับ 0. จากค่า ARE แสดงว่าอำนาจของการทดสอบคล้อยน่าจะสูงกว่าอำนาจของการทดสอบสแควร์แรงค์ ( $\mu$ ) แต่จากผลการทดลองเมื่อขนาดของตัวอย่างเป็น (10:10) (15:15) และ (20:20) ปรากฏว่าเมื่อ  $\alpha = .01$  การทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูงกว่าการทดสอบคล้อยเป็นส่วนมาก ในขณะที่  $\alpha = .05$  และ  $\alpha = .10$  นั้น เมื่อขนาดของตัวอย่างทั้งสองกลุ่มเป็น (10:10) การทดสอบสแควร์แรงค์ ( $\mu$ ) จะมีอำนาจของการทดสอบสูงกว่าเป็นส่วนมาก แต่เมื่อขนาดของตัวอย่างเป็น (15:15) และ (20:20) และ  $\Delta^2 = 2, 3, 4$  การทดสอบคล้อยมีอำนาจของการทดสอบสูงกว่า สำหรับกรณีที่  $\Delta^2$  มีค่าอื่น ๆ การทดสอบทั้งสองวิธีมีอำนาจของการทดสอบใกล้เคียงกัน ส่วนในกรณีที่ขนาดของตัวอย่างทั้งสองกลุ่มเป็น (5:10) (5:15) และ (5:20) ปรากฏว่าเมื่อ  $\Delta^2$  มากกว่า 1 การทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูงกว่าเป็นส่วนมาก แต่เมื่อ  $\Delta^2$  น้อยกว่า 1 ไม่มีข้อสรุปที่แน่นอน

การทดสอบมัตกับการทดสอบทูกี้-ซีเกล มีค่า ARE เท่ากับ .800 จากค่า ARE แสดงว่าอำนาจของการทดสอบมัตไม่ควรจะต่ำกว่าอำนาจของการทดสอบทูกี้-ซีเกล ซึ่งจากผลการทดลองเมื่อขนาดของตัวอย่างเป็น (10:10) (15:15) (20:20) (5:10) (5:15) และ (5:20) ปรากฏว่าโดยส่วนมากแล้วการทดสอบมัตมีอำนาจของการทดสอบสูงกว่า

การทดสอบมัตกับการทดสอบสแควร์แรงค์ ( $\mu$ ) มีค่า ARE เท่ากับ 1 จากค่า ARE แสดงว่าอำนาจของการทดสอบทั้งสองวิธีน่าจะเท่ากันหรือมีค่าใกล้เคียงกัน แต่จากผลการทดลองเมื่อขนาดของตัวอย่างเป็น (10:10) (15:15) (20:20) (5:10) (5:15) และ (5:20) ปรากฏว่าการทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูงกว่าการทดสอบมัตเป็นส่วนมาก

การทดสอบสแควร์แรงค์ ( $\mu$ ) กับการทดสอบทูกี้-ซีเกล มีค่า ARE เท่ากับ .806 จากค่า ARE แสดงว่าอำนาจของการทดสอบมัตไม่ควรจะต่ำกว่าอำนาจของการทดสอบทูกี้-ซีเกล ซึ่งจากผลการทดลองเมื่อขนาดของตัวอย่างเป็น (15:10) (15:15) (20:20)

(5:10) (5:15) และ (5:20) ปรากฏว่าการทดสอบสแควร์แรงค์มีอำนาจของการทดสอบ  
สูงกว่าการทดสอบทูกี้-ซีเกล.

จากค่า ARE ของแต่ละคู่ของการทดสอบ ซึ่งถ้าพิจารณาเมื่อขนาดของตัวอย่างมีขนาด  
ใหญ่แล้ว การทดสอบคล็อทซ์จะมีอำนาจการทดสอบสูงที่สุด รองลงมาคือการทดสอบมัตกับการ  
ทดสอบสแควร์แรงค์ ( $\mu$ ) ซึ่งการทดสอบทั้งสองวิธีนี้จะมีอำนาจของการทดสอบใกล้เคียง  
กันหรือเท่ากัน สำหรับการทดสอบทูกี้-ซีเกลควรจะเป็นการทดสอบที่มีอำนาจของการทดสอบ  
ต่ำที่สุด ซึ่งในการทดลองครั้งนี้ เมื่อใช้ขนาดของตัวอย่างทั้งสองกลุ่มเป็น (10:10) (15:15)  
(20:20) (5:10) (5:15) และ (5:20) ปรากฏว่าการเปรียบเทียบอำนาจของการทดสอบ  
คล็อทซ์ การทดสอบมัต และการทดสอบทูกี้-ซีเกลนั้น ค่า ARE พอที่จะเป็นเกณฑ์ที่จะใช้พิจารณา  
เปรียบเทียบได้ แต่สำหรับการทดสอบสแควร์แรงค์ ( $\mu$ ) ค่า ARE ไม่เป็นเกณฑ์ที่ดีที่จะใช้  
ในการพิจารณาเปรียบเทียบอำนาจของการทดสอบกับการทดสอบอื่น ๆ

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.15 แสดงอำนาจของการทดสอบเอฟ การทดสอบคลอท์ การทดสอบทูกี้-ซีเกล การทดสอบสแควร์แรงค์( $\mu$ ) และการทดสอบมัต เมื่อประชากรมีการแจกแจงแบบปกติ และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (10:10) จำแนกตามระดับนัยสำคัญที่กำหนด

ระดับนัยสำคัญ	สถิติทดสอบ	$\Delta^2$							
		1	2	3	4	5	6	7	8
.01	F	.010	.038	.122	.240	.359	.450	.548	.631
	KN	.012	.033	.089	.145	.196	.248	.291	.328
	TS	.010	.032	.064	.111	.156	.209	.256	.286
	SRT( $\mu$ )	.013	.046	.104	.170	.240	.309	.367	.421
	M	.014	.039	.087	.148	.190	.240	.285	.320
.05	F	.040	.150	.343	.494	.640	.730	.792	.857
	KN	.049	.140	.254	.352	.450	.519	.578	.623
	TS	.036	.106	.205	.287	.355	.416	.474	.516
	SRT( $\mu$ )	.053	.149	.278	.405	.504	.588	.658	.717
	M	.047	.133	.233	.333	.425	.500	.550	.597
.10	F	.090	.260	.466	.643	.759	.831	.887	.918
	KN	.097	.212	.364	.482	.576	.656	.714	.756
	TS	.087	.187	.297	.409	.490	.554	.611	.644
	SRT( $\mu$ )	.102	.232	.392	.541	.656	.728	.788	.821
	M	.095	.207	.341	.468	.550	.617	.681	.723

ตารางที่ 4.16 แสดงอำนาจของการทดสอบเอฟ การทดสอบคล็อทซ์ การทดสอบทุ๊กี-ซีเกล การทดสอบสแควร์แรงค์ (μ) และการทดสอบมัต เมื่อประจำการมีการแจกแจงแบบปกติ และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (15:15) ค่าแจกตามระดับนัยสำคัญที่กำหนด

ระดับ นัยสำคัญ	สถิติ ทดสอบ	$\Delta^2$							
		1	2	3	4	5	6	7	8
.01	F	.007	.087	.272	.463	.615	.730	.816	.866
	KN	.004	.044	.127	.236	.331	.419	.498	.558
	TS	.011	.047	.132	.220	.311	.400	.453	.504
	SRT (μ)	.009	.061	.173	.306	.408	.511	.586	.658
	M	.009	.049	.146	.259	.358	.440	.513	.575
.05	F	.056	.253	.517	.710	.828	.893	.931	.954
	KN	.052	.199	.393	.559	.680	.756	.815	.851
	TS	.048	.152	.300	.451	.546	.625	.691	.756
	SRT (μ)	.053	.211	.405	.577	.691	.785	.836	.889
	M	.051	.182	.363	.524	.637	.719	.782	.832
.10	F	.110	.360	.639	.808	.893	.936	.962	.979
	KN	.106	.312	.551	.699	.807	.865	.897	.915
	TS	.104	.235	.437	.582	.689	.758	.812	.851
	SRT (μ)	.111	.299	.544	.702	.807	.876	.919	.939
	M	.106	.281	.507	.654	.769	.828	.871	.901

ตารางที่ 4.17 แสดงอำนาจของการทดสอบเอฟ การทดสอบคล้อยท้าย การทดสอบทูกี้-ซีเกล การทดสอบสี่แควร์แรงค์ (  $\mu$  ) และการทดสอบมูด เมื่อประชากรมีการแจกแจงแบบปกติ และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (20:20) จำแนกตามระดับนัยสำคัญที่กำหนด

ระดับ นัยสำคัญ	สถิติ ทดสอบ	$\Delta^2$							
		1	2	3	4	5	6	7	8
.01	F	.008	.136	.407	.639	.792	.895	.943	.967
	KN	.005	.080	.252	.420	.555	.671	.751	.801
	TS	.008	.068	.188	.332	.450	.562	.648	.711
	SRT( $\mu$ )	.009	.087	.259	.441	.593	.707	.779	.828
	M	.007	.081	.237	.408	.539	.657	.737	.785
.05	F	.045	.321	.649	.847	.930	.969	.981	.990
	KN	.046	.268	.525	.721	.836	.887	.926	.942
	TS	.052	.191	.421	.593	.722	.801	.837	.863
	SRT( $\mu$ )	.052	.247	.527	.724	.818	.880	.919	.940
	M	.046	.248	.486	.670	.784	.855	.890	.921
.10	F	.091	.451	.769	.911	.966	.981	.992	.997
	KN	.107	.389	.690	.835	.903	.944	.964	.979
	TS	.094	.290	.635	.718	.807	.863	.892	.915
	SRT( $\mu$ )	.104	.361	.665	.814	.892	.931	.958	.973
	M	.101	.347	.541	.787	.872	.909	.944	.963



จากตาราง 4.15-4.17 ซึ่งแสดงค่าอำนาจของการทดสอบเอฟ การทดสอบคล้อยทซ์ การทดสอบมัต การทดสอบทูกี้-ซีเกล และการทดสอบสแควร์แรงค์ ( $\mu$ ) เมื่อประชากรมีการแจกแจงแบบปกติ และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (10:10) (15:15) และ (20:20) โดยกำหนดระดับนัยสำคัญเป็น .01, .05 และ .10 ซึ่งเปรียบเทียบอำนาจของการทดสอบทั้ง 5 วิธีได้ดังนี้

ที่ระดับนัยสำคัญ .01 การทดสอบเอฟมีอำนาจของการทดสอบสูงที่สุด ยกเว้นกรณี ที่ขนาดของตัวอย่างเป็น (10:10)  $\Delta^2=2$  ซึ่งปรากฏว่าการทดสอบเอฟ การทดสอบมัต การทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบใกล้เคียงกัน โดยมีความแตกต่างของอำนาจของการทดสอบสูงที่สุด .008 สำหรับการทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบต่ำกว่าการทดสอบเอฟ แต่สูงกว่าการทดสอบอีก 3 วิธีดังกล่าว ส่วนการทดสอบมัตกับการทดสอบคล้อยทซ์มีอำนาจของการทดสอบแตกต่างกันมากที่สุด .008 เมื่อขนาดของตัวอย่างเป็น (10:10) แต่เมื่อขนาดของตัวอย่างเป็น (15:15) การทดสอบมัต มีอำนาจของการทดสอบสูงกว่า และเมื่อขนาดของตัวอย่างเป็น (20:20) การทดสอบคล้อยทซ์มีอำนาจของการทดสอบสูงกว่า การทดสอบทูกี้-ซีเกล มีอำนาจของการทดสอบต่ำกว่าการทดสอบอื่น ๆ ยกเว้นกรณีที่  $\Delta^2=2$  ขนาดของตัวอย่างเป็น (10:10) และ  $\Delta^2=2$  และ 3 ขนาดของตัวอย่างเป็น (15:15) การทดสอบทูกี้-ซีเกลมีอำนาจของการทดสอบใกล้เคียงกับการทดสอบคล้อยทซ์ โดยมีความแตกต่างของอำนาจของการทดสอบสูงที่สุด .005

ที่ระดับนัยสำคัญ .05 การทดสอบเอฟมีอำนาจของการทดสอบสูงที่สุด การทดสอบสแควร์แรงค์ ( $\mu$ ) การทดสอบคล้อยทซ์ การทดสอบมัต และการทดสอบทูกี้-ซีเกล มีอำนาจของการทดสอบเป็นอันดับที่ 2, 3, 4 และ 5 ตามลำดับ ยกเว้นกรณีที่ขนาดของตัวอย่างเป็น (20:20) โดยส่วนมากแล้วการทดสอบคล้อยทซ์มีอำนาจของการทดสอบสูงกว่าการทดสอบสแควร์แรงค์

ที่ระดับนัยสำคัญ .10 การทดสอบเอฟมีอำนาจของการทดสอบสูงที่สุด สำหรับการทดสอบคล้อยทซ์ และการทดสอบสแควร์แรงค์ ( $\mu$ ) นั้น เมื่อขนาดของตัวอย่างเป็น (10:10) ทุกค่าของ  $\Delta^2$  และขนาดของตัวอย่างเป็น (15:15)  $\Delta^2=7$  และ 8 การทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูงกว่า แต่เมื่อขนาดของตัวอย่างเป็น (15:15)  $\Delta^2=3, 4, 5$  และ 6 ขนาดของตัวอย่างเป็น (20:20)  $\Delta^2=7$  และ 8 การทดสอบทั้ง 2 วิธีนี้มีอำนาจของการทดสอบแตกต่างกันมากที่สุด .009 สำหรับกรณีอื่น ๆ การทดสอบคล้อยทซ์มีอำนาจของการทดสอบสูงกว่า สำหรับการทดสอบมัตและการทดสอบทูกี้-ซีเกล มีอำนาจของการทดสอบเป็นอันดับที่ 4 และ 5 ตามลำดับ

ตารางที่ 4-18 แสดงอำนาจของการทดสอบเอฟ การทดสอบคลอซี การทดสอบบูทส์-ซีเกล การทดสอบลค์แควร์แรงค์ ( $\mu$ ) และการทดสอบมูดี เมื่อประชากรมีการแจกแจงแบบปกติ และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (5,10) จำนวนตามระดับนัยสำคัญที่กำหนด

ระดับนัยสำคัญ	สถิติทดสอบ	$\Delta^2$															
		$\frac{1}{8}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	1	2	3	4	5	6	7	8	
.01	F	.450	.395	.342	.255	.171	.102	.039	.003	.009	.033	.058	.084	.110	.134	.160	
	KN	.174	.141	.121	.084	.068	.048	.023	.008	.016	.024	.040	.054	.075	.093	.098	
	TS	.162	.129	.107	.080	.059	.033	.016	.007	.014	.024	.038	.052	.072	.085	.098	
	SRT( $\mu$ )	.269	.233	.203	.163	.112	.069	.035	.011	.017	.041	.065	.085	.118	.141	.180	
	M	.132	.102	.090	.060	.048	.032	.014	.007	.016	.024	.040	.054	.072	.093	.098	
.05	F	.687	.633	.565	.487	.380	.257	.129	.037	.079	.137	.224	.298	.366	.435	.486	
	KN	.453	.414	.369	.325	.254	.186	.120	.044	.076	.116	.179	.221	.272	.318	.352	
	TS	.349	.314	.275	.245	.196	.140	.068	.036	.061	.108	.151	.190	.222	.252	.276	
	SRT( $\mu$ )	.527	.482	.434	.372	.308	.237	.123	.044	.088	.147	.228	.293	.371	.419	.463	
	M	.430	.398	.360	.323	.255	.189	.110	.041	.074	.117	.172	.214	.259	.296	.330	
.10	F	.773	.732	.679	.600	.507	.371	.206	.093	.138	.253	.359	.456	.551	.616	.660	
	KN	.584	.548	.500	.459	.383	.291	.191	.108	.148	.237	.301	.364	.429	.475	.511	
	TS	.511	.472	.444	.391	.321	.254	.161	.092	.140	.212	.265	.322	.384	.427	.463	
	SRT( $\mu$ )	.652	.617	.575	.512	.424	.322	.206	.092	.151	.261	.374	.460	.530	.585	.628	
	M	.534	.488	.453	.408	.336	.253	.177	.084	.138	.216	.278	.339	.405	.457	.497	

ตารางที่ 4.19 แสดงอำนาจของการทดสอบเอฟ การทดสอบคล็องท์ การทดสอบบูทซ์-ซีเกล การทดสอบสแควร์แรงค์(μ) และการทดสอบมูดี  
เมื่อประชากรมีการแจกแจงแบบปกติ และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (5,15) จำนวนตามระดับนัยสำคัญที่กำหนด

ระดับนัยสำคัญ	สถิติ ทดสอบ	$\Delta^2$														
		$\frac{1}{8}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	1	2	3	4	5	6	7	8
.01	F	.564	.503	.441	.346	.261	.141	.054	.011	.019	.045	.079	.108	.138	.163	.199
	KN	.280	.249	.204	.163	.109	.066	.031	.013	.030	.048	.074	.089	.118	.137	.154
	TS	.202	.172	.141	.104	.072	.054	.027	.012	.027	.040	.054	.075	.101	.118	.131
	SRT(μ)	.338	.295	.256	.197	.131	.066	.031	.014	.034	.059	.084	.109	.145	.172	.212
	M	.254	.215	.181	.142	.090	.063	.027	.015	.029	.045	.074	.089	.118	.136	.153
.05	F	.747	.703	.635	.551	.455	.324	.159	.047	.102	.168	.240	.314	.395	.469	.524
	KN	.538	.486	.442	.385	.306	.227	.121	.042	.094	.153	.222	.295	.361	.410	.453
	TS	.408	.376	.338	.286	.221	.146	.087	.037	.081	.123	.196	.255	.295	.337	.380
	SRT(μ)	.598	.552	.498	.430	.338	.245	.115	.047	.102	.184	.260	.342	.439	.504	.550
	M	.601	.451	.404	.349	.270	.187	.106	.044	.090	.147	.220	.290	.345	.391	.437
.10	F	.804	.774	.729	.656	.557	.434	.261	.108	.165	.276	.387	.493	.569	.649	.709
	KN	.638	.600	.558	.502	.426	.328	.204	.101	.167	.272	.378	.466	.528	.586	.628
	TS	.558	.526	.482	.415	.356	.272	.157	.093	.146	.235	.337	.417	.473	.526	.562
	SRT(μ)	.680	.659	.608	.550	.474	.367	.209	.085	.178	.286	.413	.531	.608	.661	.710
	M	.601	.553	.519	.458	.389	.300	.182	.090	.158	.253	.355	.448	.517	.573	.607

ตารางที่ 4.20 แสดงอำนาจของการทดสอบเอฟ การทดสอบคลอซี การทดสอบบูกี-ซีเกล การทดสอบสแควร์แรงค์ ( $\mu$ ) และการทดสอบจูด  
เมื่อประชากรมีการแจกแจงแบบปกติและขนาดของตัวอย่างทั้งสองกลุ่มเป็น (5:20) ค่าแนกตามระดับนัยสำคัญที่กำหนด

ระดับนัยสำคัญ	สถิติ ทดสอบ	$\Delta^2$														
		$\frac{1}{8}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	1	2	3	4	5	6	7	8
.01	F	.652	.589	.534	.445	.345	.207	.075	.007	.019	.037	.063	.084	.115	.143	.180
	KN	.347	.313	.267	.222	.167	.111	.049	.008	.020	.038	.063	.083	.106	.124	.156
	TS	.261	.233	.205	.172	.118	.075	.042	.005	.015	.043	.064	.078	.098	.122	.145
	SRT ( $\mu$ )	.386	.347	.295	.244	.186	.109	.044	.008	.021	.052	.086	.120	.148	.189	.224
	M	.317	.283	.237	.193	.149	.093	.042	.007	.020	.040	.059	.082	.104	.124	.153
.05	F	.785	.743	.691	.621	.528	.385	.201	.049	.081	.149	.230	.301	.374	.442	.496
	KN	.591	.559	.503	.440	.362	.268	.140	.040	.075	.143	.220	.277	.336	.408	.459
	TS	.451	.417	.371	.400	.272	.185	.097	.042	.066	.121	.185	.236	.287	.331	.373
	SRT ( $\mu$ )	.599	.559	.516	.459	.388	.269	.149	.052	.094	.161	.250	.338	.427	.500	.553
	M	.540	.491	.448	.325	.326	.231	.123	.047	.075	.143	.215	.271	.337	.402	.449
.10	F	.846	.804	.759	.702	.609	.494	.312	.098	.145	.262	.372	.457	.543	.620	.673
	KN	.716	.672	.624	.570	.496	.386	.247	.099	.142	.248	.366	.452	.515	.580	.634
	TS	.590	.417	.504	.454	.386	.305	.190	.096	.140	.226	.330	.397	.453	.522	.573
	SRT ( $\mu$ )	.721	.678	.623	.569	.492	.389	.239	.098	.164	.294	.401	.511	.595	.674	.732
	M	.664	.622	.575	.520	.452	.348	.215	.094	.140	.251	.367	.452	.513	.581	.628

จากตาราง 4-19-4.21 ซึ่งแสดงค่าอำนาจของการทดสอบเอฟ การทดสอบคล้อยหลัง การทดสอบมุต การทดสอบทูกี้-ซีเกล และการทดสอบสแควร์แรงค์ ( $\mu$ ) เมื่อประชากรมีการ แลกแจกแบบปกติ ขนาดของตัวอย่างทั้งสองกลุ่มเป็น (5:10) (5:15) และ (5:20) และ ระดับนัยสำคัญคือ  $\alpha = .01$  .05 และ .10 ซึ่งเปรียบเทียบอำนาจของการทดสอบได้ดังนี้

ที่ระดับนัยสำคัญ  $\alpha = .01$  การทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูง ที่สุดเมื่อ  $\Delta^2$  มีค่ามากกว่า 1 แต่เมื่อ  $\Delta^2$  น้อยกว่า 1 การทดสอบเอฟมีอำนาจของการทดสอบ สูงที่สุด สำหรับการทดสอบคล้อยหลัง การทดสอบมุต การทดสอบทูกี้-ซีเกลโดยส่วนมากมีอำนาจของการทดสอบ ต่ำกว่าการทดสอบสแควร์แรงค์ ( $\mu$ ) โดยที่เมื่อขนาดของตัวอย่าง (5:10) (5:15) (5:20) การทดสอบคล้อยหลังและการทดสอบมุตมีอำนาจของการทดสอบใกล้เคียงกัน เมื่อ  $\Delta^2$  มากกว่า 1 ซึ่งความแตกต่างของอำนาจของการทดสอบมากที่สุด .004 แต่เมื่อ  $\Delta^2$  น้อยกว่า การทดสอบคล้อยหลังมีอำนาจของการทดสอบสูงกว่า โดยที่การทดสอบทูกี้-ซีเกลมีอำนาจของการ ทดสอบต่ำกว่าการทดสอบมุตและการทดสอบคล้อยหลังเป็นส่วนใหญ่

ที่ระดับนัยสำคัญ  $\alpha = .05$  การทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูงที่สุด เมื่อ  $\Delta^2$  มีค่ามากกว่า 1 ยกเว้นกรณีที่มีขนาดของตัวอย่างเป็น (5:10)  $\Delta^2 = 4, 5$  และ 6 การทดสอบสแควร์แรงค์ ( $\mu$ ) และการทดสอบเอฟมีอำนาจของการทดสอบใกล้เคียงกัน โดย มีความแตกต่างของอำนาจของการทดสอบสูงที่สุด .005 แต่เมื่อ  $\Delta^2 = 7$  และ 8 การทดสอบ เอฟมีอำนาจของการทดสอบสูงกว่า สำหรับกรณีที่  $\Delta^2$  น้อยกว่า 1 การทดสอบเอฟมีอำนาจของ การทดสอบสูงที่สุด การทดสอบมุต และการทดสอบคล้อยหลังมีอำนาจของการทดสอบต่ำกว่าการทดสอบ สแควร์แรงค์ ( $\mu$ ) เป็นส่วนใหญ่โดยที่การทดสอบคล้อยหลังและการทดสอบมุตมีอำนาจของการทดสอบ ใกล้เคียงกันเมื่อ  $\Delta^2$  มากกว่า 1 เช่นเมื่อขนาดของตัวอย่างเป็น (5:15)  $\Delta^2 = 2, 3, 4, 5$  แต่เมื่อ  $\Delta^2 = 6, 7, 8$  การทดสอบคล้อยหลังมีอำนาจของการทดสอบสูงกว่า สำหรับกรณี ที่มีขนาดของตัวอย่างเป็น (5:20) การทดสอบมุตและการทดสอบคล้อยหลังมีอำนาจของการทดสอบ ใกล้เคียงกัน มีความแตกต่างของอำนาจของการทดสอบสูงที่สุด .010 สำหรับกรณีที่  $\Delta^2$  น้อยกว่า 1 การทดสอบคล้อยหลังมีอำนาจของการทดสอบสูงกว่าการทดสอบมุต

ที่ระดับนัยสำคัญ  $\alpha = .10$  การทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูงที่สุด เมื่อ  $\Delta^2$  มากกว่า 1 ยกเว้นกรณีที่มีขนาดของตัวอย่างเป็น (5:10)  $\Delta^2 = 6, 7$  และ 8 การ ทดสอบเอฟมีอำนาจของการทดสอบสูงกว่า แต่เมื่อ  $\Delta^2$  น้อยกว่า 1 การทดสอบเอฟมีอำนาจของ

การทดสอบสูงที่สุด สำหรับการทดสอบคลอทซ์มีอำนาจของการทดสอบต่ำกว่าการทดสอบสแควร์แรงค์ (N) เมื่อขนาดของตัวอย่างเป็น (5:10) และ (5:15) แต่เมื่อขนาดของตัวอย่างเป็น (5:20)  $\Delta^2$  น้อยกว่า 1 การทดสอบคลอทซ์และการทดสอบสแควร์แรงค์มีอำนาจของการทดสอบใกล้เคียงกัน มีความแตกต่างของอำนาจของการทดสอบสูงที่สุด -008 ส่วนการทดสอบมัตมีอำนาจของการทดสอบต่ำกว่าการทดสอบคลอทซ์ ยกเว้นกรณีที่มีขนาดของตัวอย่างเป็น (5:20)  $\Delta^2$  มากกว่า 1 การทดสอบทั้งสองวิธีมีอำนาจของการทดสอบใกล้เคียงกัน โดยมีความแตกต่างของอำนาจของการทดสอบสูงที่สุด -006 สำหรับการทดสอบทูที-ซี, ภาล มีอำนาจของการทดสอบต่ำที่สุดเมื่อเทียบกับ การทดสอบอื่น ๆ

จากค่าอำนาจของการทดสอบเมื่อขนาดของตัวอย่างเป็น (10:10) (15:15) (20:20) (5:10) (5:15) และ (5:20) สรุปจำนวนครั้งที่การทดสอบแต่ละวิธีมีอำนาจของการทดสอบสูงที่สุดได้ดังตาราง 4.21

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย



ตาราง 4.21 แสดงจำนวนครั้งที่การทดสอบเอฟ การทดสอบคล้อยท้าย การทดสอบทูกี้-ซีเกล และการทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูงที่สุด เมื่อประชากรมีการแจกแจงแบบปกติ จากการทดลองทั้งหมด 21 กรณี เมื่อขนาดของตัวอย่างทั้งสองกลุ่มเป็น (10:10) (15:15) และ (20:20) และ 42 กรณี เมื่อขนาดของตัวอย่างทั้งสองกลุ่มเป็น (5:10) (5:15) และ (5:20) โดยแยกเป็นกรณีที่  $\Delta^2$  น้อยกว่า 1 21 กรณี และ  $\Delta^2$  มากกว่า 1 21 กรณี จำแนกตามระดับนัยสำคัญที่กำหนด

ระดับ นัยสำคัญ	สถิติทดสอบ	ขนาดของตัวอย่าง				
		<i>m/n</i>			(5:10) (5:15) (5:20)	
		(10:10)	(15:15)	(20:20)	$\Delta^2 < 1$	$\Delta^2 > 1$
.01	F		20		21	0
	KN		0		0	0
	TS		0		0	0
	SRT( $\mu$ )		1		0	21
	M		0		0	0
.05	F		21		21	4
	KN		0		0	0
	TS		0		0	0
	SRT( $\mu$ )		0		0	18
	M		0		0	0
.10	F		21		21	0
	KN		0		0	4
	TS		0		0	0
	SRT( $\mu$ )		0		1	18
	M		0		0	0



จากตารางที่ 4.21 สรุปได้ดังนี้

1. เมื่อขนาดของตัวอย่างเท่ากัน การทดสอบเอฟมีอำนาจของการทดสอบสูงที่สุดมากที่สุด คือ มีอำนาจการทดสอบสูงที่สุด 20 กรณีเมื่อ  $\alpha = .01$  21 กรณี เมื่อ  $\alpha = .05$  และ  $.10$

2. ในกรณีขนาดของตัวอย่างไม่เท่ากันการทดสอบเอฟมีอำนาจของการทดสอบสูงที่สุด เมื่อ  $\Delta^2$  น้อยกว่า 1 คือ มีอำนาจของการทดสอบสูงที่สุด 21 กรณีทุก ๆ ค่าของ  $\alpha$  แต่เมื่อ  $\Delta^2$  มากกว่า 1 การทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูงที่สุดมากที่สุด คือ 21 กรณีเมื่อ  $\alpha = .01$  18 กรณีเมื่อ  $\alpha = .05$  และ  $.10$

#### เปรียบเทียบค่า ARE กับอำนาจของการทดสอบ 2 วิธีเมื่อประชากรมีการแจกแจงแบบปกติ

จากการศึกษาอำนาจของการทดสอบเอฟ การทดสอบคลอท์ การทดสอบมุต การทดสอบทูกี-ซีเกล และการทดสอบสแควร์แรงค์ ( $\mu$ ) นำมาเปรียบเทียบกับค่า ARE สำหรับแต่ละคู่ของการทดสอบเมื่อประชากรมีการแจกแจงแบบปกติได้ดังนี้

การทดสอบเอฟกับการทดสอบคลอท์มีค่า ARE เท่ากับ 1 จากค่า ARE แสดงว่าอำนาจของการทดสอบเอฟกับการทดสอบคลอท์น่าจะใกล้เคียงกัน แต่จากผลการทดลองเมื่อขนาดของตัวอย่างเท่ากันคือ 10 15 และ 20 การทดสอบเอฟมีอำนาจของการทดสอบสูงกว่าเป็นส่วนใหญ่ แต่ในกรณีขนาดของตัวอย่างไม่เท่ากัน คือ (5:10) (5:15) และ (5:20) เมื่อ  $\Delta^2$  น้อยกว่า 1 การทดสอบเอฟมีอำนาจของการทดสอบสูงกว่า แต่เมื่อ  $\Delta^2$  มากกว่า 1 เมื่อ  $\Delta^2 = 2, 3, 4$  โดยส่วนมากแล้วการทดสอบทั้งสองวิธีจะมีอำนาจของการทดสอบใกล้เคียงกัน แต่เมื่อ  $\Delta^2$  มีค่าอื่น ๆ การทดสอบเอฟจะมีอำนาจของการทดสอบสูงกว่า

การทดสอบเอฟกับการทดสอบมุต มีค่า ARE เท่ากับ  $\approx 760$  จากค่า ARE การทดสอบเอฟน่าจะมีความอำนาจของการทดสอบไม่ต่ำกว่าการทดสอบมุต ซึ่งจากค่าอำนาจของการทดสอบเมื่อขนาดของตัวอย่างเท่ากัน คือ 10, 15 และ 20 การทดสอบเอฟมีอำนาจของการทดสอบสูงกว่าการทดสอบมุต เมื่อ  $\alpha = .05$  และ  $.10$  แต่เมื่อ  $\alpha = .01$  และ  $\Delta^2 = 2$  มีบางกรณีที่การทดสอบทั้งสองวิธีมีอำนาจของการทดสอบใกล้เคียงกัน นอกจากนั้นการทดสอบเอฟมีอำนาจของการทดสอบสูงกว่าการทดสอบมุต สำหรับกรณีขนาดของตัวอย่างไม่เท่ากัน คือ (5:10) (5:15) และ (5:20) การทดสอบเอฟจะมีอำนาจของการทดสอบสูงกว่า เมื่อ

$\Delta^2$  น้อยกว่า 1 แต่เมื่อ  $\Delta^2$  มากกว่า 1 เช่น  $\Delta^2 = 2, 3, 4$  โดยส่วนมากการทดสอบทั้ง 2 วิธีมีอำนาจการทดสอบใกล้เคียงกัน แต่เมื่อ  $\Delta^2$  มีค่าอื่น ๆ การทดสอบเอฟมีอำนาจของการทดสอบสูงกว่า

การทดสอบเอฟกับการทดสอบทูกี-ซีเกล มีค่า ARE เท่ากับ .608 จากค่า ARE แสดงว่าการทดสอบเอฟมีอำนาจของการทดสอบไม่ต่ำกว่าการทดสอบทูกี-ซีเกล แต่จากการทดลองเมื่อขนาดของตัวอย่างเท่ากัน คือ 10, 15 และ 20 การทดสอบเอฟมีอำนาจของการทดสอบสูงกว่าการทดสอบทูกี-ซีเกล โดยเฉพาะเมื่อ  $\alpha = .05$  และ  $.10$  สำหรับกรณีที่ยีนขนาดของตัวอย่างไม่เท่ากัน คือ (5:10) (5:15) และ (5:20) เมื่อ  $\Delta^2$  น้อยกว่า 1 การทดสอบเอฟมีอำนาจของการทดสอบสูงกว่า แต่เมื่อ  $\Delta^2$  มากกว่า 1 โดยส่วนมากแล้วการทดสอบเอฟมีอำนาจของการทดสอบสูงกว่า

การทดสอบเอฟกับการทดสอบสแควร์แรงค์ (M) มีค่า ARE เท่ากับ .760 จากค่า ARE แสดงว่าการทดสอบเอฟมีอำนาจของการทดสอบไม่ต่ำกว่าการทดสอบสแควร์แรงค์ แต่จากการทดลองเมื่อขนาดของตัวอย่างเท่ากัน คือ 10, 15 และ 20 การทดสอบเอฟมีอำนาจของการทดสอบสูงกว่าเมื่อ  $\alpha = .05$  และ  $.10$  แต่เมื่อ  $\alpha = .01$  เมื่อ  $\Delta^2 = 2$  การทดสอบทั้ง 2 วิธีมีอำนาจของการทดสอบใกล้เคียงกัน แต่กรณี  $\Delta^2$  มีค่าอื่น การทดสอบเอฟมีอำนาจของการทดสอบสูงกว่า แต่เมื่อขนาดของตัวอย่างไม่เท่ากัน คือ (5:10) (5:15) และ (5:20) การทดสอบเอฟมีอำนาจของการทดสอบสูงกว่า เมื่อ  $\Delta^2$  น้อยกว่า 1 แต่เมื่อ  $\Delta^2$  มากกว่า 1 โดยส่วนมากแล้ว การทดสอบสแควร์แรงค์ (M) มีอำนาจของการทดสอบสูงกว่า

การทดสอบคล็อทซ์กับการทดสอบมัต มีค่า ARE เท่ากับ .760 จากค่า ARE การทดสอบคล็อทซ์ไม่ต่ำกว่าอำนาจของการทดสอบไม่ต่ำกว่าการทดสอบมัต แต่จากผลการทดลองเมื่อขนาดของตัวอย่างเท่ากัน คือ 10, 15 และ 20 การทดสอบคล็อทซ์มีอำนาจของการทดสอบสูงกว่าการทดสอบมัตเป็นส่วนมาก เมื่อ  $\alpha = .05$  และ  $.10$  แต่เมื่อ  $\alpha = .01$  มีบางกรณีที่การทดสอบมัตมีอำนาจของการทดสอบสูงกว่า สำหรับกรณีที่ยีนขนาดของตัวอย่างไม่เท่ากัน คือ (5:10) (5:15) และ (5:20) การทดสอบคล็อทซ์จะมีอำนาจของการทดสอบสูงกว่า การทดสอบมัตเป็นส่วนมาก เมื่อ  $\Delta^2$  น้อยกว่า 1 แต่เมื่อ  $\Delta^2$  มากกว่า 1 โดยส่วนมากแล้วเมื่อ  $\alpha = .01$  การทดสอบทั้ง 2 วิธีมีอำนาจของการทดสอบใกล้เคียงกัน และเมื่อ  $\alpha = .05$  และ  $.10$  เมื่อ  $\Delta^2 = 2, 3$  และ 4 การทดสอบทั้ง 2 วิธีมีอำนาจของการทดสอบใกล้เคียงกัน แต่เมื่อ  $\Delta^2 = 5, 6, 7$  และ 8 โดยส่วนมากแล้วการทดสอบคล็อทซ์มีอำนาจของการทดสอบสูงกว่า

การทดสอบคล้อยกับการทดสอบทูกี-ซีเกลมีค่า ARE เท่ากับ  $-608$  จากค่า ARE แสดงว่าการทดสอบคล้อยที่น่าจะมีอำนาจของการทดสอบต่ำกว่าการทดสอบทูกี-ซีเกล แต่จากผลการทดลองเมื่อขนาดของตัวอย่างเท่ากันคือ 10, 15 และ 20 การทดสอบคล้อยที่น่าจะมีอำนาจของการทดสอบสูงกว่าการทดสอบทูกี-ซีเกล เมื่อ  $\alpha = .05$  และ  $\alpha = .10$  แต่เมื่อ  $\alpha = .01$  มีบางกรณีที่มีการทดสอบทั้ง 2 วิธีมีอำนาจของการทดสอบใกล้เคียงกัน สำหรับกรณีที่ขนาดของตัวอย่างไม่เท่ากัน คือ (5:10) (5:15) และ (5:20) เมื่อ  $\Delta^2$  น้อยกว่า 1 การทดสอบคล้อยมีอำนาจของการทดสอบสูงกว่า แต่เมื่อ  $\Delta^2$  มากกว่า 1 โดยส่วนมากแล้ว การทดสอบคล้อยมีอำนาจของการทดสอบสูงกว่า

การทดสอบคล้อยกับการทดสอบสแควร์แรงค์ ( $\mu$ ) มีค่า ARE เท่ากับ  $.760$  จากค่า ARE แสดงว่าการทดสอบคล้อยที่น่าจะมีอำนาจของการทดสอบต่ำกว่าการทดสอบสแควร์แรงค์ ( $\mu$ ) แต่จากผลการทดลองเมื่อขนาดของตัวอย่างเท่ากัน คือ 10, 15 และ 20 การทดสอบสแควร์แรงค์ ( $\mu$ ) โดยส่วนมากมีอำนาจของการทดสอบสูงกว่า และในกรณีที่ขนาดของตัวอย่างไม่เท่ากัน คือ (5:10) (5:15) และ (5:20) โดยส่วนมากการทดสอบสแควร์แรงค์ ( $\mu$ ) ก็มีอำนาจของการทดสอบสูงกว่า ไม่ว่า  $\Delta^2$  จะมีค่ามากกว่าหรือน้อยกว่า 1

การทดสอบมูตกับการทดสอบสแควร์แรงค์ ( $\mu$ ) มีค่า ARE เท่ากับ 1 จากค่า ARE แสดงว่าอำนาจของการทดสอบทั้ง 2 วิธีน่าจะใกล้เคียงกัน แต่จากผลการทดลอง การทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูงกว่าการทดสอบมูต ไม่ว่าขนาดของตัวอย่างจะเท่ากันหรือไม่เท่ากัน

การทดสอบมูตกับการทดสอบทูกี-ซีเกล มีค่า ARE เท่ากับ  $-80$  จากค่า ARE แสดงว่าการทดสอบมูตไม่น่าจะมีอำนาจของการทดสอบต่ำกว่าการทดสอบทูกี-ซีเกล ซึ่งจากผลการทดลองเมื่อขนาดของตัวอย่างเท่ากัน คือ 10, 15 และ 20 การทดสอบมูตมีอำนาจของการทดสอบสูงกว่า และเมื่อขนาดของตัวอย่างไม่เท่ากัน คือ (5:10) (5:15) และ (5:20) โดยส่วนมากแล้วการทดสอบมูตมีอำนาจของการทดสอบสูงกว่า

การทดสอบสแควร์แรงค์ ( $\mu$ ) กับการทดสอบทูกี-ซีเกลมีค่า ARE เท่ากับ  $-80$  จากค่า ARE แสดงว่าการทดสอบสแควร์แรงค์ ( $\mu$ ) ไม่น่าจะมีอำนาจของการทดสอบต่ำกว่าการทดสอบทูกี-ซีเกล ซึ่งจากผลการทดลองการทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูงกว่า ไม่ว่าขนาดของตัวอย่างจะเท่ากันหรือไม่ก็ตาม

จากค่า ARE ของแต่ละคู่ของการทดสอบ ซึ่งถ้าพิจารณาเมื่อขนาดของตัวอย่างมีขนาดใหญ่แล้วการทดสอบเอฟและการทดสอบคล็อทซ์น่าจะมีอำนาจของการทดสอบสูงกว่าการทดสอบอื่น รองลงมาคือการทดสอบมูตและการทดสอบสแควร์แรงค์ ( $\mu$ ) ซึ่งการทดสอบทั้งสองวิธีนี้ควรจะมีอำนาจของการทดสอบใกล้เคียงกัน หรือเท่ากัน สำหรับการทดสอบทูกี้-ซีเกล ควรจะเป็นการทดสอบที่มีอำนาจของการทดสอบต่ำที่สุด ซึ่งในการทดลองครั้งนี้เมื่อใช้ขนาดของตัวอย่างเป็น (10:10) (15:15) (20:20) (5:10) (5:15) และ (5:20) ปรากฏว่าการเปรียบเทียบอำนาจของการทดสอบเอฟ การทดสอบคล็อทซ์ การทดสอบมูต และการทดสอบทูกี้-ซีเกล นั้นค่า ARE พอที่จะเป็นเกณฑ์พิจารณาเปรียบเทียบได้ แต่สำหรับการทดสอบสแควร์แรงค์ ( $\mu$ ) ค่า ARE ไม่เป็นเกณฑ์ที่ดีที่จะใช้ในการพิจารณาเปรียบเทียบอำนาจของการทดสอบกับการทดสอบอื่น ๆ



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.22 แสดงอำนาจของการทดสอบเอฟ การทดสอบคลอท์ การทดสอบทูกี้-ซีเกล การทดสอบสแควร์แรงค์( $\mu$ ) และการทดสอบมัต เมื่อประชากรมีการแจกแจงแบบโลจิสติก และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (10:10) จำแนกตามระดับนัยสำคัญที่กำหนด

ระดับนัยสำคัญ	สถิติทดสอบ	$\Delta^2$							
		1	2	3	4	5	6	7	8
.01	F	.024	.080	.202	.293	.391	.474	.539	.600
	KN	.006	.034	.075	.122	.167	.212	.263	.303
	TS	.011	.023	.057	.093	.135	.179	.210	.236
	SRT( $\mu$ )	.009	.035	.093	.144	.216	.280	.326	.373
	M	.008	.031	.073	.120	.164	.210	.256	.299
.05	F	.090	.230	.374	.509	.612	.696	.757	.813
	KN	.043	.116	.232	.322	.417	.474	.534	.580
	TS	.040	.104	.178	.256	.326	.386	.439	.488
	SRT( $\mu$ )	.038	.139	.252	.354	.439	.515	.591	.646
	M	.048	.107	.204	.307	.391	.455	.508	.558
.10	F	.170	.317	.486	.621	.720	.739	.848	.872
	KN	.100	.195	.344	.452	.537	.601	.648	.691
	TS	.085	.169	.277	.375	.462	.525	.583	.625
	SRT( $\mu$ )	.097	.216	.356	.475	.569	.647	.721	.753
	M	.091	.191	.332	.439	.514	.590	.647	.686

ตารางที่ 4.23 แสดงอำนาจของการทดสอบเอฟ การทดสอบคล็อทซ์ การทดสอบทูกี้-ซีเกล การทดสอบสแควร์แรงค์( $\mu$ ) และการทดสอบมัตต์ เมื่อประชากรมีการแจกแจงแบบโลจิสติก และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (15:15) จำแนกตามระดับนัยสำคัญที่กำหนด

ระดับนัยสำคัญ	สถิติทดสอบ	2 $\Delta$							
		1	2	3	4	5	6	7	8
.01	F	.030	.125	.290	.461	.594	.690	.760	.816
	KN	.004	.035	.096	.174	.261	.330	.394	.458
	TS	.009	.036	.105	.194	.271	.334	.392	.438
	SRT( $\mu$ )	.007	.046	.138	.248	.336	.428	.503	.574
	M	.009	.040	.102	.200	.287	.366	.426	.487
.05	F	.116	.266	.506	.678	.777	.854	.896	.928
	KN	.046	.143	.316	.455	.566	.658	.720	.780
	TS	.041	.128	.266	.384	.484	.563	.629	.679
	SRT( $\mu$ )	.043	.158	.335	.492	.606	.703	.779	.820
	M	.041	.146	.308	.445	.563	.644	.712	.761
.10	F	.179	.375	.622	.757	.854	.909	.942	.963
	KN	.098	.261	.462	.614	.718	.789	.840	.880
	TS	.090	.214	.391	.519	.624	.701	.759	.804
	SRT( $\mu$ )	.092	.265	.478	.628	.747	.821	.869	.903
	M	.092	.250	.446	.582	.686	.760	.816	.860



ตาราง 4.24 แสดงอำนาจของการทดสอบเอฟ การทดสอบคล็อทซ์ การทดสอบทุ๊กี-ซีเกล การทดสอบสแควร์แรงค์ ( $\mu$ ) และการทดสอบมูต เมื่อประชากรมีการแจกแจงแบบปกติ และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (20:20) จำแนกตามระดับนัยสำคัญที่กำหนด

ระดับนัยสำคัญ	สถิติทดสอบ	$\Delta^2$							
		1	2	3	4	5	6	7	8
.01	F	.027	.181	.417	.614	.753	.861	.919	.953
	KN	.008	.047	.168	.301	.454	.558	.637	.703
	TS	.008	.058	.146	.266	.401	.511	.584	.641
	SRT ( $\mu$ )	.010	.062	.189	.371	.521	.638	.705	.768
	M	.009	.058	.172	.329	.461	.581	.650	.713
.05	F	.092	.346	.625	.808	.908	.955	.978	.986
	KN	.034	.182	.427	.632	.751	.831	.880	.915
	TS	.049	.166	.381	.545	.672	.748	.807	.859
	SRT ( $\mu$ )	.036	.198	.474	.650	.768	.856	.902	.933
	M	.041	.188	.429	.613	.733	.812	.864	.910
.10	F	.169	.457	.722	.886	.952	.978	.986	.992
	KN	.078	.312	.596	.758	.858	.917	.944	.960
	TS	.086	.269	.509	.669	.773	.890	.887	.917
	SRT ( $\mu$ )	.081	.330	.608	.762	.866	.921	.947	.969
	M	.082	.295	.571	.737	.839	.816	.938	.950



จากตาราง 4.17-4.20 ซึ่งแสดงค่าอำนาจของการทดสอบเอฟ การทดสอบคล้อย การทดสอบมุต การทดสอบทูกี้-ซีเกล และการทดสอบสแควร์แรงค์ ( $\mu$ ) เมื่อประชากรมีการแจกแจงแบบโลจิสติก เมื่อขนาดของตัวอย่างเป็น (10:10) (15:15) และ (20:20) และระดับนัยสำคัญมีค่า .01 .05 และ .10

เนื่องจากการทดสอบเอฟควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้น้อยมาก โดยเฉพาะเมื่อใช้เกณฑ์ของ Cochran การทดสอบเอฟไม่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ และเมื่อใช้เกณฑ์ของ Bradley สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เพียง 1 ครั้ง จากการทดลองทั้งหมด 18 ครั้ง จึงไม่น่ามาเปรียบเทียบอำนาจของการทดสอบกับการทดสอบอื่น ๆ อีก 4 วิธี ซึ่งค่าอำนาจของการทดสอบทั้ง 4 วิธีสามารถเปรียบเทียบได้ดังนี้

ที่ระดับนัยสำคัญ .01 การทดสอบสแควร์แรงค์ มีอำนาจของการทดสอบสูงที่สุดเมื่อขนาดของตัวอย่างเป็น (10:10) (15:15) และ (20:20) และ  $\Delta^2$  มากกว่า 2 สำหรับกรณี  $\Delta^2 = 2$  อำนาจของการทดสอบทั้ง 4 วิธีจะต่างกันเพียงเล็กน้อย โดยเฉพาะเมื่อขนาดของตัวอย่างทั้งสองกลุ่มเป็น (15:15) อำนาจของการทดสอบจะต่างกันมากที่สุด .011 สำหรับการทดสอบคล้อย การทดสอบมุต และการทดสอบทูกี้-ซีเกล มีอำนาจของการทดสอบต่ำกว่าการทดสอบสแควร์แรงค์เป็นส่วนใหญ่ โดยที่อำนาจของการทดสอบมุต และการทดสอบคล้อยมีค่าใกล้เคียงกันเมื่อขนาดของตัวอย่างเป็น (10:10) มีความแตกต่างของอำนาจของการทดสอบสูงที่สุด .007 สำหรับกรณีขนาดของตัวอย่างเป็น (15:15) และ (20:20) การทดสอบมุตจะมีอำนาจของการทดสอบสูงกว่าการทดสอบคล้อย แต่ค่าของอำนาจของการทดสอบซึ่งสูงกว่ามีค่าไม่มาก สำหรับการทดสอบทูกี้-ซีเกลนั้น โดยส่วนมากแล้วจะมีอำนาจของการทดสอบต่ำกว่าการทดสอบอื่น ๆ

ที่ระดับนัยสำคัญ .05 การทดสอบสแควร์แรงค์ มีอำนาจของการทดสอบสูงที่สุดทุก ๆ ขนาดของตัวอย่าง สำหรับการทดสอบคล้อย การทดสอบมุตและการทดสอบทูกี้-ซีเกล มีอำนาจของการทดสอบต่ำกว่าการทดสอบสแควร์แรงค์ โดยที่การทดสอบคล้อยจะมีอำนาจของการทดสอบสูงกว่าการทดสอบอื่น ยกเว้นกรณีขนาดของตัวอย่างเป็น (10:10) และ  $\Delta^2 = 2$  (15:15) และ  $\Delta^2 = 2$  และ 3 และ (20:20)  $\Delta^2 = 2, 3$  การทดสอบคล้อยและการทดสอบมุตมีอำนาจของการทดสอบใกล้เคียงกัน โดยมีความแตกต่างของอำนาจของการทดสอบสูงที่สุด .009 สำหรับการทดสอบทูกี้-ซีเกล มีอำนาจของการทดสอบต่ำที่สุดเมื่อเทียบกับการทดสอบอื่น ๆ

ที่ระดับนัยสำคัญ .10 การทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูงที่สุดเมื่อขนาดของตัวอย่างเป็น (10:10) (15:15) และ (20:20) สำหรับการทดสอบคล้อย การทดสอบมัต และการทดสอบทูกี้-ซีเกลมีอำนาจของการทดสอบต่ำกว่าการทดสอบสแควร์แรงค์ ( $\mu$ ) โดยที่การทดสอบคล้อย มีอำนาจการทดสอบสูงว่าการทดสอบอื่น ๆ และการทดสอบทูกี้-ซีเกลมีอำนาจของการทดสอบต่ำที่สุดเมื่อเทียบกับการทดสอบอื่น ๆ



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.25 แสดงอำนาจของการทดสอบเอฟ การทดสอบคัลลิวท์ การทดสอบทูที-ซีเทค การทดสอบสแควร์แรงค์ (μ) และการทดสอบมูต์  
เพื่อประสิทธิภาพการแจกแจงแบบโลจิสติก และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (5:10) ค่าแนกตามระดับนัยสำคัญที่กำหนด

ระดับนัยสำคัญ	สถิติ ทดสอบ	$\Delta^2$														
		$\frac{1}{8}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	1	2	3	4	5	6	7	8
.01	F	.445	.397	.336	.284	.219	.138	.067	.022	.027	.054	.083	.121	.145	.178	.215
	KN	.137	.116	.096	.077	.062	.044	.021	.011	.018	.027	.042	.050	.062	.069	.082
	TS	.118	.097	.077	.055	.044	.032	.019	.010	.018	.031	.041	.049	.061	.064	.076
	SRT (μ)	.210	.175	.157	.115	.084	.059	.029	.010	.025	.045	.070	.089	.102	.130	.153
	M	.090	.077	.059	.048	.040	.028	.015	.009	.018	.027	.042	.050	.062	.069	.082
.05	F	.637	.600	.538	.470	.389	.290	.177	.082	.115	.187	.271	.343	.402	.463	.518
	KN	.393	.357	.320	.262	.213	.148	.092	.048	.069	.114	.158	.200	.242	.275	.313
	TS	.306	.272	.239	.194	.164	.106	.063	.038	.058	.094	.132	.160	.191	.218	.241
	SRT (μ)	.456	.420	.380	.308	.261	.168	.096	.061	.091	.115	.217	.273	.340	.377	.416
	M	.392	.355	.311	.254	.203	.143	.087	.048	.062	.106	.145	.186	.230	.260	.293
.10	F	.267	.388	.483	.574	.634	.680	.729	.150	.198	.319	.401	.485	.555	.611	.667
	KN	.524	.480	.441	.395	.338	.259	.170	.098	.137	.198	.280	.335	.371	.418	.451
	TS	.479	.436	.398	.355	.298	.224	.141	.088	.135	.196	.259	.302	.361	.404	.432
	SRT (μ)	.600	.553	.506	.445	.372	.274	.169	.083	.169	.257	.388	.412	.469	.533	.578
	M	.458	.425	.386	.339	.290	.223	.142	.106	.130	.182	.259	.313	.353	.405	.447

ตารางที่ 4.26 แสดงอำนาจของการทดสอบเอฟ การทดสอบคลัสท์ การทดสอบทูที-ซี, กล การทดสอบดีแควร์แรงค์ (μ) และการทดสอบมูด์ เมื่อประชากรมีการแจกแจงแบบโลจิสติก และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (5:15) ค่าแนกตามระดับนัยสำคัญที่กำหนด

ระดับนัยสำคัญ	สถิติ ทดสอบ	$\Delta^2$														
		$\frac{1}{8}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	1	2	3	4	5	6	7	8
.01	F	.519	.465	.391	.322	.249	.160	.080	.024	.036	.064	.095	.124	.170	.200	.232
	KN	.247	.210	.174	.138	.094	.055	.023	.004	.016	.031	.047	.061	.074	.102	.121
	TS	.178	.151	.116	.095	.074	.050	.020	.007	.012	.024	.042	.053	.069	.082	.099
	SRT (μ)	.284	.245	.205	.160	.117	.070	.036	.009	.024	.046	.065	.096	.123	.158	.179
	M	.210	.179	.143	.111	.082	.052	.023	.005	.016	.030	.045	.059	.071	.100	.118
.05	F	.697	.644	.579	.513	.405	.296	.186	.085	.124	.204	.285	.371	.443	.507	.570
	KN	.472	.435	.396	.342	.264	.189	.103	.053	.071	.128	.186	.250	.302	.348	.382
	TS	.389	.345	.308	.255	.192	.144	.082	.047	.061	.113	.171	.217	.270	.296	.323
	SRT (μ)	.513	.464	.409	.363	.284	.209	.117	.056	.096	.157	.238	.313	.383	.445	.500
	M	.443	.407	.363	.316	.249	.174	.100	.050	.070	.128	.190	.241	.297	.338	.369
.10	F	.770	.726	.677	.602	.519	.393	.257	.148	.215	.322	.439	.532	.616	.676	.720
	KN	.583	.548	.500	.451	.383	.285	.185	.095	.150	.225	.313	.394	.460	.517	.573
	TS	.539	.495	.444	.402	.342	.263	.167	.097	.156	.223	.299	.361	.417	.472	.530
	SRT (μ)	.649	.598	.538	.465	.402	.307	.182	.100	.172	.271	.378	.475	.550	.615	.670
	M	.565	.524	.478	.431	.361	.263	.167	.087	.142	.220	.308	.380	.448	.502	.560

ตารางที่ 4.27 แสดงอำนาจของการทดสอบเอฟ การทดสอบคล็อดท์ การทดสอบบูทซ์-ซีเกล การทดสอบสแควร์แรงค์ ( $\mu$ ) และการทดสอบมูด เมื่อปรากฏการแจกแจงแบบโลจิสติก และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (5:20) จำนวนตามระดับนัยสำคัญที่กำหนด

ที่ระดับนัยสำคัญ	สถิติ ทดสอบ	$\Delta^2$														
		$\frac{1}{8}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	1	2	3	4	5	6	7	8
.01	F	.576	.522	.474	.397	.295	.190	.086	.019	.026	.050	.087	.115	.156	.204	.244
	KN	.283	.248	.210	.165	.125	.074	.030	.012	.024	.044	.062	.084	.107	.129	.149
	TS	.223	.195	.173	.142	.096	.064	.031	.011	.017	.028	.057	.083	.109	.124	.138
	SRT( $\mu$ )	.308	.275	.228	.175	.122	.078	.046	.011	.025	.047	.076	.111	.138	.169	.195
	M	.257	.223	.197	.159	.118	.072	.032	.010	.024	.038	.058	.087	.112	.132	.150
.05	F	.717	.684	.618	.541	.474	.340	.200	.068	.112	.208	.299	.371	.456	.506	.558
	KN	.533	.482	.432	.368	.298	.225	.126	.044	.081	.141	.215	.279	.321	.377	.427
	TS	.422	.379	.348	.302	.235	.167	.099	.049	.076	.128	.185	.232	.285	.326	.359
	SRT( $\mu$ )	.561	.506	.453	.389	.318	.199	.119	.056	.094	.174	.247	.324	.385	.450	.508
	M	.501	.443	.410	.355	.290	.204	.117	.051	.081	.143	.214	.271	.320	.374	.423
.10	F	.782	.748	.701	.637	.543	.440	.275	.138	.208	.321	.449	.520	.594	.675	.723
	KN	.672	.623	.563	.507	.427	.328	.210	.101	.153	.247	.344	.416	.478	.534	.581
	TS	.566	.516	.478	.423	.359	.275	.186	.102	.148	.229	.318	.380	.425	.484	.527
	SRT( $\mu$ )	.689	.639	.571	.498	.433	.333	.195	.097	.177	.286	.386	.479	.556	.620	.670
	M	.633	.591	.533	.477	.408	.313	.194	.101	.159	.236	.349	.414	.477	.532	.579

จากตาราง 4.25-4.27 ซึ่งแสดงค่าอำนาจของการทดสอบเอฟ การทดสอบคล็อทซ์ การทดสอบมัต การทดสอบทูกี้-ซีเกล และการทดสอบสแควร์แรงค์ (น) เมื่อประจำกรมมีการ แจกแจงแบบโลจิสติก ขนาดของตัวอย่างเป็น (5:10) (5:15) และ (5:20) และระดับ นัยสำคัญเป็น .01 .05 และ .10

เนื่องจากการทดสอบเอฟควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้น้อยมาก โดยเฉพาะเมื่อใช้เกณฑ์ของ Cochran การทดสอบเอฟไม่สามารถควบคุมความคลาดเคลื่อนประเภท ที่ 1 ได้ และเมื่อใช้เกณฑ์ของ Bradley สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ 3 ครั้ง จากการทดลองทั้งหมด 18 ครั้ง สิ่งไม่น่ามาเปรียบเทียบอำนาจของการทดสอบกับการ ทดสอบอื่น ๆ อีก 4 วิธี ซึ่งค่าอำนาจของการทดสอบทั้ง 4 วิธีสามารถเปรียบเทียบได้ดังนี้

ที่ระดับนัยสำคัญ .01 การทดสอบสแควร์แรงค์มีอำนาจของการทดสอบสูงกว่าการ ทดสอบอื่น ๆ ยกเว้นกรณีที่ขนาดของตัวอย่างเป็น (5:20)  $\Delta^2 = \frac{1}{3}, \frac{1}{4}, 2$  และ 3 การทดสอบ คล็อทซ์และการทดสอบสแควร์แรงค์มีอำนาจของการทดสอบแตกต่างกันน้อย โดยมีความแตกต่าง มากที่สุด .004 สำหรับการทดสอบคล็อทซ์จะมีอำนาจของการทดสอบต่ำกว่าการทดสอบสแควร์แรงค์ เป็นส่วนมาก แต่มีอำนาจของการทดสอบสูงกว่าการทดสอบมัต และการทดสอบทูกี้-ซีเกล เมื่อ  $\Delta^2$  น้อยกว่า 1 ในขณะที่  $\Delta^2$  มากกว่า 1 นั้น เมื่อขนาดของตัวอย่างเป็น (5:10) การทดสอบทั้ง 3 วิธีมีอำนาจของการทดสอบใกล้เคียงกัน มีความแตกต่างของอำนาจของการทดสอบ .006 สำหรับ กรณีที่ขนาดของตัวอย่างเป็น (5:15) และ (5:20) โดยส่วนมากการทดสอบคล็อทซ์มีอำนาจของ การทดสอบสูงกว่าการทดสอบทูกี้-ซีเกล และมีอำนาจของการทดสอบใกล้เคียงกับการทดสอบมัต โดยมีความแตกต่างของอำนาจของการทดสอบมากที่สุด .006

ที่ระดับนัยสำคัญ .05 การทดสอบสแควร์แรงค์มีอำนาจของการทดสอบสูงกว่าการ ทดสอบอื่น ๆ เมื่อขนาดของตัวอย่างเป็น (5:10) และ (5:15) ในกรณีที่ขนาดของตัวอย่าง เป็น (5:20) การทดสอบคล็อทซ์มีอำนาจของการทดสอบสูงกว่าการทดสอบสแควร์แรงค์ เมื่อ  $\Delta^2 = \frac{1}{3}$  และเมื่อ  $\Delta^2 = \frac{1}{2}$  อำนาจของการทดสอบทั้ง 2 วิธีนี้แตกต่างกันเพียง .007 ส่วน กรณีที่  $\Delta^2$  มีค่าอื่น ๆ การทดสอบสแควร์แรงค์มีอำนาจของการทดสอบสูงกว่า สำหรับการ ทดสอบมัตมีอำนาจของการทดสอบต่ำกว่าการทดสอบสแควร์แรงค์ เมื่อเปรียบเทียบอำนาจ ของการทดสอบมัตและการทดสอบคล็อทซ์ ปรากฏว่าในกรณีที่  $\Delta^2$  น้อยกว่า 1 การทดสอบ คล็อทซ์มีอำนาจของการทดสอบสูงกว่า และเมื่อ  $\Delta^2$  มากกว่า 1 ขนาดของตัวอย่างเป็น



(5:10) การทดสอบคล้อยมีอำนาจของการทดสอบสูงกว่า สำหรับกรณีที่ยกมาของตัวอย่าง เป็น (5:15) อำนาจการทดสอบทั้ง 2 วิธีนี้มีความแตกต่างกันสูงที่สุด .013 ส่วนกรณีที่ยกมา ของตัวอย่างเป็น (5:20) การทดสอบทั้ง 2 วิธีนี้ใกล้เคียงกัน มีความแตกต่างของอำนาจ ของการทดสอบมากที่สุด .008 สำหรับการทดสอบบูกี-ซีเกล จะมีอำนาจของการทดสอบต่ำกว่า การทดสอบอื่น ๆ

ที่ระดับนัยสำคัญ .10 การทดสอบสแควร์แรงค์มีอำนาจของการทดสอบสูงกว่า การทดสอบอื่น ๆ เมื่อขนาดของตัวอย่างเป็น (5:10) (5:15) และ (5:20) และ  $\Delta^2$  มากกว่า 1 สำหรับกรณีที่  $\Delta^2$  น้อยกว่า 1 เมื่อขนาดของตัวอย่างเป็น (5:10) และ (5:15) เมื่อ  $\Delta^2=2$  การทดสอบคล้อยและการทดสอบสแควร์แรงค์มีอำนาจการทดสอบ ใกล้เคียงกัน โดยมีความแตกต่างของอำนาจการทดสอบมากที่สุด .003 แต่เมื่อ  $\Delta^2$  มีค่า อื่น ๆ การทดสอบสแควร์แรงค์มีอำนาจของการทดสอบสูงกว่า ส่วนกรณีที่ยกมาของตัวอย่าง เป็น (5:20) เมื่อ  $\Delta^2 = \frac{1}{2}$  การทดสอบคล้อยมีอำนาจของการทดสอบสูงกว่า แต่เมื่อ  $\Delta^2 = \frac{1}{3}, \frac{1}{4}, \frac{1}{5}$  และ  $\frac{1}{6}$  อำนาจของการทดสอบทั้ง 2 วิธีนี้ใกล้เคียงกันมีความแตกต่างของอำนาจของ การทดสอบมากที่สุด .009 และเมื่อ  $\Delta^2 = \frac{1}{7}$  และ  $\frac{1}{8}$  การทดสอบสแควร์แรงค์มีอำนาจของ การทดสอบสูงกว่า สำหรับการทดสอบมัต มีอำนาจของการทดสอบต่ำกว่าการทดสอบสแควร์แรงค์ เมื่อเปรียบเทียบการทดสอบมัตและการทดสอบคล้อย ปรากฏว่าเมื่อขนาดของตัวอย่างเป็น (5:10) และ (5:15) การทดสอบคล้อยมีอำนาจของการทดสอบสูงกว่า แต่เมื่อขนาดของ ตัวอย่างเป็น (5:20) เมื่อ  $\Delta^2$  มากกว่า 1 อำนาจของการทดสอบทั้ง 2 วิธีนี้แตกต่างกัน มากที่สุด .011 แต่เมื่อ  $\Delta^2$  น้อยกว่า 1 การทดสอบคล้อยมีอำนาจของการทดสอบสูงกว่า สำหรับการทดสอบบูกี-ซีเกลมีอำนาจของการทดสอบต่ำกว่าการทดสอบอื่น ๆ

จากค่าอำนาจของการทดสอบเมื่อขนาดของตัวอย่างเป็น (10:10) (15:15) (20:20) (5:10) (5:15) และ (5:20) สรุปจำนวนครั้งที่การทดสอบแต่ละวิธีมีอำนาจ ของการทดสอบสูงที่สุดดังตาราง 4.28



ตาราง 4.28 แสดงจำนวนครั้งที่การทดสอบคล้อย การทดสอบทุ๊ก-ซีเกล และการทดสอบ  
 สแควร์แรงค์ ( $\mu$ ) มีอำนาจการทดสอบสูงที่สุด เมื่อประชากรมีการแจกแจง  
 แบบโลจิสติก จากการทดลองทั้งหมด 21 กรณี เมื่อขนาดของตัวอย่างทั้งสอง  
 กลุ่มเป็น (10:10) (15:15) และ (20:20) และ 42 กรณี เมื่อขนาด  
 ของตัวอย่างทั้งสองกลุ่มเป็น (5:10) (5:15) และ (5:20) โดยแยก  
 เป็นกรณีที่  $\Delta^2$  น้อยกว่า 1 21 กรณี และ  $\Delta^2$  มากกว่า 1 21 กรณี  
 จำนวนตามระดับนัยสำคัญที่กำหนด

ระดับ นัยสำคัญ	สถิติทดสอบ	ขนาดตัวอย่าง					
		(10,10)	(15,15)	(20,20)	(5,10)	(5,15)	(5,20)
					$\Delta^2 < 1$	$\Delta^2 > 1$	
.01	KN		0		1	0	
	TS		0		0	0	
	SRT( $\mu$ )		21		20	21	
	M		0		0	0	
.05	KN		1		2	0	
	TS		0		0	0	
	SRT( $\mu$ )		21		19	21	
	M		0		0	0	
.10	KN		3		4	0	
	TS		0		0	0	
	SRT( $\mu$ )		18		17	21	
	M		0		0	0	

จากตารางที่ 4.28 สรุปได้ดังนี้

1. เมื่อขนาดของตัวอย่างเท่ากัน การทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของ  
 การทดสอบสูงที่สุดมากที่สุด คือ มีอำนาจของการทดสอบสูงที่สุด 21 กรณีทุก ๆ ค่าของ  $\alpha$

2. เมื่อขนาดของตัวอย่างไม่เท่ากัน การทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูงที่สุดมากที่สุด คือ ในกรณีที่  $\Delta^2$  มากกว่า 1 มีอำนาจของการทดสอบสูงที่สุด 21 กรณีทุก ๆ ค่าของ  $\alpha$  และเมื่อ  $\Delta^2$  น้อยกว่า 1 มีอำนาจของการทดสอบสูงที่สุด 20 กรณีเมื่อ  $\alpha = .01$  19 กรณีเมื่อ  $\alpha = .05$  และ 21 กรณีเมื่อ  $\alpha = .10$

เปรียบเทียบค่า ARE กับอำนาจของการทดสอบเมื่อประชากรมีการแจกแจงแบบโลจิสติก

จากอำนาจของการทดสอบทั้ง 4 วิธี คือ การทดสอบคล็อทซ์ การทดสอบมุต การทดสอบทูกี-ซีเกล และการทดสอบสแควร์แรงค์ ( $\mu$ ) นำมาเปรียบเทียบกับค่า ARE สำหรับแต่ละคู่ของการทดสอบได้ดังนี้

การทดสอบคล็อทซ์กับการทดสอบมุตมีค่า ARE เท่ากับ .896 จากค่า ARE แสดงว่าการทดสอบคล็อทซ์น่าจะมีอำนาจของการทดสอบไม่ต่ำกว่าการทดสอบมุต ซึ่งจากผลการทดลองเมื่อขนาดของตัวอย่างเป็น (10:10) (15:15) และ (20:20) ปรากฏว่าโดยส่วนมากแล้วการทดสอบคล็อทซ์มีอำนาจของการทดสอบสูงกว่า และเมื่อขนาดของตัวอย่างเป็น (5:10) (5:15) และ (5:20) และ  $\Delta^2$  น้อยกว่า 1 การทดสอบคล็อทซ์จะมีอำนาจของการทดสอบสูงกว่าการทดสอบมุต แต่เมื่อ  $\Delta^2$  มากกว่า 1 ในกรณีที่  $\alpha = .01$  และ  $.05$  อำนาจของการทดสอบทั้งสองวิธีใกล้เคียงกัน สำหรับกรณีที่  $\alpha = .10$  การทดสอบคล็อทซ์มีอำนาจของการทดสอบสูงกว่าเป็นส่วนมาก

การทดสอบคล็อทซ์กับการทดสอบทูกี-ซีเกล มีค่า ARE เท่ากับ .750 จากค่า ARE แสดงว่าการทดสอบคล็อทซ์น่าจะมีอำนาจของการทดสอบไม่ต่ำกว่าการทดสอบทูกี-ซีเกล ซึ่งจากผลการทดลองปรากฏว่าเมื่อขนาดของตัวอย่างเป็น (10:10) (15:15) และ (20:20) โดยส่วนมากแล้วการทดสอบคล็อทซ์มีอำนาจของการทดสอบสูงกว่า และเมื่อขนาดของตัวอย่างทั้งสองกลุ่มเป็น (5:10) (5:15) และ (5:20) โดยส่วนมากแล้วการทดสอบคล็อทซ์มีอำนาจการทดสอบสูงกว่าทั้งกรณีที่  $\Delta^2$  มากกว่า 1 และน้อยกว่า 1

การทดสอบคล็อทซ์กับการทดสอบสแควร์แรงค์ ( $\mu$ ) มีค่า ARE เท่ากับ .896 จากค่า ARE แสดงว่าการทดสอบคล็อทซ์น่าจะมีอำนาจของการทดสอบไม่ต่ำกว่าการทดสอบสแควร์แรงค์ ( $\mu$ ) แต่จากการทดลองเมื่อขนาดของตัวอย่างทั้งสองกลุ่มเป็น (10:10) (15:15) และ (20:20) ปรากฏว่าโดยส่วนมากแล้วการทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบ

สูงกว่า และเมื่อขนาดของตัวอย่างเป็น (5:10) (5:15) และ (5:20) ปรากฏว่า โดยส่วนมากแล้วการทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูงกว่า ทั้งกรณีที่มีมากกว่า 1 และ  $\Delta^2$  น้อยกว่า 1

การทดสอบมัตกับการทดสอบทูกี้-ซีเกล มีค่า ARE เท่ากับ .837 จากค่า ARE แสดงว่าการทดสอบมัตน่าจะมีอำนาจของการทดสอบไม่ต่ำกว่าการทดสอบทูกี้-ซีเกล ซึ่งจากผลการทดลองเมื่อขนาดของตัวอย่างเป็น (10:10) (15:15) (20:20) (5:10) (5:15) และ (5:20) ปรากฏว่าโดยส่วนมากแล้วการทดสอบมัตมีอำนาจของการทดสอบสูงกว่า

การทดสอบมัตและการทดสอบสแควร์แรงค์ ( $\mu$ ) มีค่า ARE เท่ากับ 1 จากค่า ARE แสดงว่าการทดสอบสแควร์แรงค์ ( $\mu$ ) และการทดสอบมัตควรจะมีอำนาจการทดสอบเท่ากัน หรือใกล้เคียงกัน แต่จากผลการทดลองเมื่อขนาดของตัวอย่างเป็น (10:10) (5:15) (20:20) (5:10) (5:15) และ (5:20) ปรากฏว่าการทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูงว่าการทดสอบมัต

การทดสอบสแควร์แรงค์ ( $\mu$ ) กับการทดสอบทูกี้-ซีเกล มีค่า ARE เท่ากับ .837 จากค่า ARE แสดงว่าอำนาจของการทดสอบสแควร์แรงค์ ( $\mu$ ) น่าจะมีอำนาจของการทดสอบไม่ต่ำกว่าการทดสอบทูกี้-ซีเกล ซึ่งจากผลการทดลองเมื่อขนาดของตัวอย่างเป็น (10:10) (5:15) (20:20) (5:10) (5:15) และ (5:20) ปรากฏว่าการทดสอบสแควร์แรงค์ มีอำนาจการทดสอบสูงว่าการทดสอบทูกี้-ซีเกล

จากค่า ARE ของแต่ละคู่ของการทดสอบ ซึ่งถ้าพิจารณาเมื่อขนาดของตัวอย่างใหญ่ การทดสอบคล็อทซ์จะมีอำนาจของการทดสอบสูงที่สุด รองลงมาคือการทดสอบมัต และการทดสอบสแควร์แรงค์ ( $\mu$ ) ซึ่งการทดสอบทั้งสองวิธีนี้ควรจะมีอำนาจของการทดสอบใกล้เคียงกัน หรือเท่ากัน และการทดสอบทูกี้-ซีเกล น่าจะมีอำนาจของการทดสอบต่ำที่สุด ซึ่งในการทดลองครั้งนี้เมื่อขนาดของตัวอย่างเป็น (10:10) (15:15) (20:20) (5:10) (5:15) และ (5:20) ปรากฏว่าการเปรียบเทียบอำนาจของการทดสอบคล็อทซ์ การทดสอบมัต และการทดสอบทูกี้-ซีเกล ค่า ARE พอที่จะใช้เป็นเกณฑ์พิจารณาเปรียบเทียบได้ แต่สำหรับการทดสอบสแควร์แรงค์ ( $\mu$ ) ค่า ARE ไม่เป็นเกณฑ์ที่ดีที่จะใช้ในการพิจารณาเปรียบเทียบอำนาจของการทดสอบ

ตาราง 4.29 แสดงค่าอำนาจของกราฟขดลวดสี่เหลี่ยม (X) เมื่อประยชน์มีการแจกแจงแบบปกติ แบบยูนิฟอรั่ม และแบบโลจิสติก จำแนกตามระดับนัยสำคัญที่กำหนด และขนาดของตัวอย่าง

ขนาดของ ตัวอย่าง	ลักษณะ การแจกแจง	ระดับนัยสำคัญ																							
		.01								.05								.10							
		$\Delta^2$								$\Delta^2$								$\Delta^2$							
1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8		
10:10	NN	.014	.052	.116	.192	.265	.321	.372	.427	.068	.157	.300	.428	.523	.584	.637	.693	.117	.255	.426	.548	.640	.721	.779	.817
	UU	.008	.077	.188	.294	.394	.473	.521	.586	.050	.214	.426	.559	.675	.745	.797	.834	.111	.326	.554	.691	.785	.841	.871	.907
	LL	.010	.044	.105	.161	.219	.268	.327	.370	.058	.153	.253	.366	.462	.520	.587	.642	.113	.228	.383	.495	.579	.650	.711	.765
15:15	NN	.008	.060	.185	.296	.416	.503	.588	.657	.059	.211	.406	.568	.695	.788	.842	.882	.114	.313	.550	.705	.814	.877	.914	.938
	UU	.015	.124	.346	.515	.659	.745	.802	.829	.057	.342	.618	.784	.866	.907	.944	.960	.096	.449	.751	.875	.924	.952	.968	.980
	LL	.015	.057	.148	.250	.354	.446	.525	.601	.067	.183	.368	.513	.635	.724	.785	.836	.122	.286	.495	.649	.749	.828	.870	.898
20:20	NN	.008	.102	.273	.458	.602	.708	.782	.826	.059	.265	.513	.725	.826	.879	.926	.949	.114	.386	.654	.811	.885	.934	.961	.976
	UU	.011	.185	.529	.715	.836	.897	.931	.959	.047	.469	.728	.900	.945	.972	.984	.991	.103	.577	.864	.939	.973	.986	.995	.995
	LL	.011	.074	.209	.394	.531	.632	.719	.777	.046	.222	.475	.660	.781	.853	.889	.933	.098	.345	.610	.776	.864	.916	.945	.969

NN หมายถึง การแจกแจงแบบปกติ

UU หมายถึง การแจกแจงแบบยูนิฟอรั่ม

LL หมายถึง การแจกแจงแบบโลจิสติก

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

จากตาราง 4.29 ซึ่งแสดงอำนาจของการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) เมื่อประชากรมีการแจกแจงแบบปกติ แบบยูนิฟอร์ม และแบบโลจิสติก เมื่อขนาดของตัวอย่างทั้งสองกลุ่มเป็น (10:10) (15:15) และ (20:20) จะเห็นได้ว่าทุก ๆ ค่าของ  $\alpha$  ที่กำหนด ค่าอำนาจของการทดสอบที่สูงที่สุดเมื่อประชากรมีการแจกแจงแบบยูนิฟอร์ม รองลงมาก็คือการแจกแจงแบบปกติ และแบบโลจิสติกตามลำดับ

สำหรับค่าอำนาจของการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) ปรากฏว่าเมื่อเทียบกับอำนาจของการทดสอบเอฟ เมื่อประชากรมีการแจกแจงแบบปกติ โดยส่วนมากจะมีอำนาจของการทดสอบต่ำกว่าเมื่อเทียบกับการทดสอบสแควร์แรงค์ ( $\mu$ ) เมื่อประชากรมีการแจกแจงแบบยูนิฟอร์ม มีอำนาจของการทดสอบต่ำกว่า และเมื่อเทียบกับอำนาจของการทดสอบสแควร์แรงค์ ( $\mu$ ) เมื่อประชากรมีการแจกแจงแบบโลจิสติก ส่วนมากจะมีอำนาจของการทดสอบสูงกว่า ยกเว้นกรณีที่ยขนาดของตัวอย่างเป็น (20:20) มีบางกรณีที่การทดสอบสแควร์แรงค์ ( $\mu$ ) มีอำนาจของการทดสอบสูงกว่า

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.30 แสดงอำนาจของการทดสอบเอฟ การทดสอบสแควร์แรงค์ ( $\bar{X}$ ) เมื่อประชากรมีการแจกแจงแบบเบ้ โดยมีความเบ้ .25 และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (10:10) จำแนกตามระดับนัยสำคัญที่กำหนดและความโค้งของประชากร

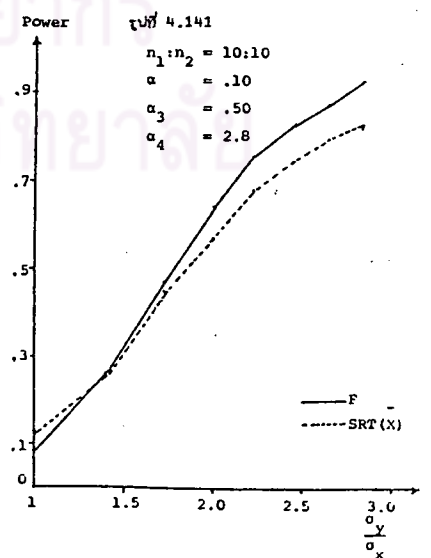
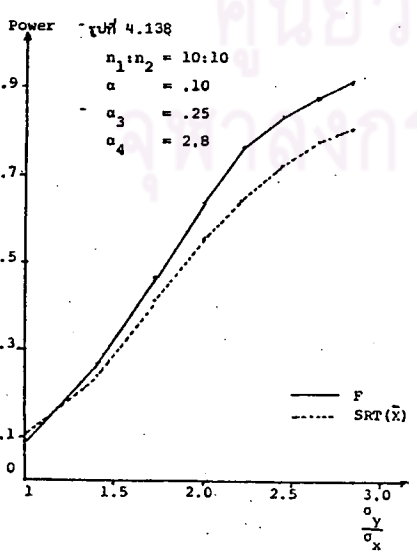
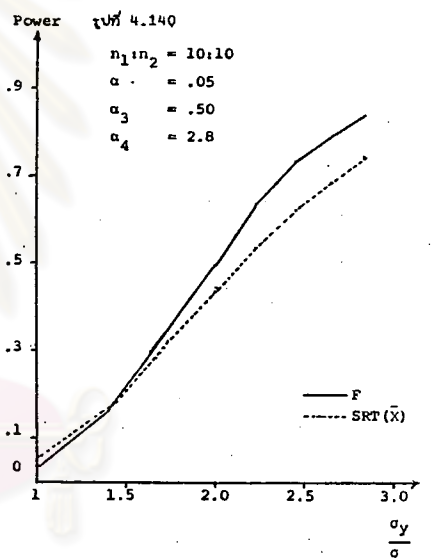
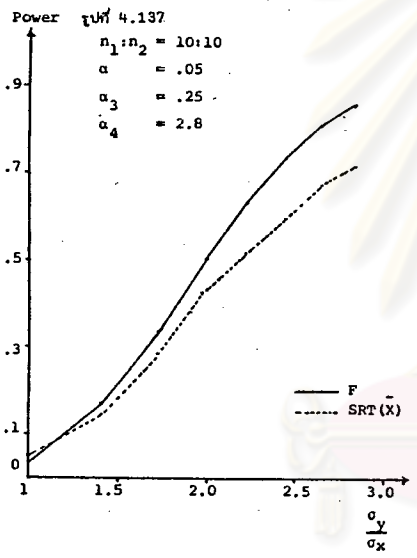
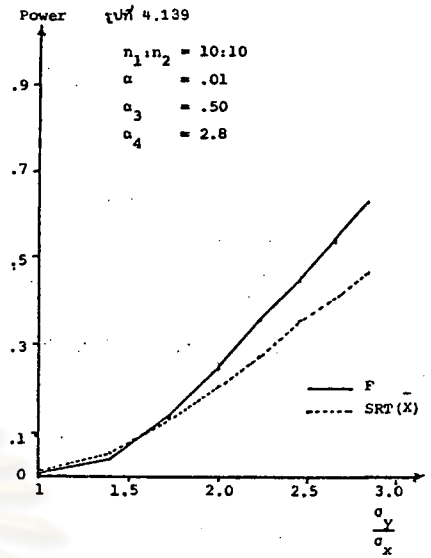
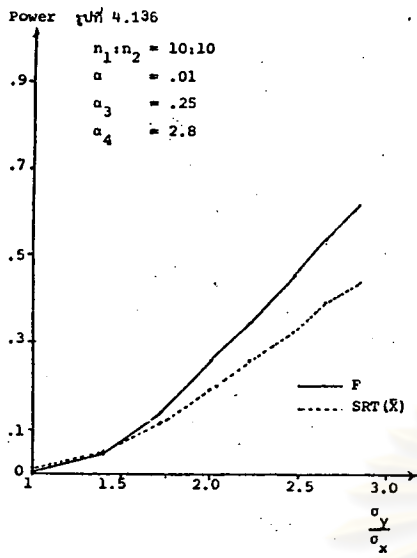
ระดับ นัยสำคัญ	ความ โค้ง	สถิติ ทดสอบ	$\Delta^2$							
			1	2	3	4	5	6	7	8
.01	2.8	F	.003	.044	.140	.252	.346	.452	.539	.617
		SRT ( $\bar{X}$ )	.012	.050	.118	.191	.256	.321	.390	.437
	3.0	F	.004	.051	.155	.262	.356	.454	.535	.615
		SRT ( $\bar{X}$ )	.011	.045	.119	.183	.250	.310	.377	.416
	3.2	F	.006	.054	.164	.274	.360	.460	.529	.613
		SRT ( $\bar{X}$ )	.011	.042	.115	.179	.247	.301	.362	.404
.05	2.8	F	.038	.170	.335	.500	.630	.736	.805	.852
		SRT ( $\bar{X}$ )	.050	.147	.285	.421	.511	.591	.670	.717
	3.0	F	.047	.180	.341	.501	.631	.726	.797	.846
		SRT ( $\bar{X}$ )	.050	.147	.272	.405	.501	.577	.643	.707
	3.2	F	.059	.189	.346	.504	.623	.721	.791	.838
		SRT ( $\bar{X}$ )	.050	.147	.265	.397	.488	.563	.623	.691
.10	2.8	F	.092	.270	.467	.640	.764	.835	.875	.916
		SRT ( $\bar{X}$ )	.106	.241	.414	.551	.647	.725	.778	.813
	3.0	F	.097	.282	.470	.638	.756	.827	.869	.904
		SRT ( $\bar{X}$ )	.107	.238	.396	.534	.629	.710	.765	.810
	3.2	F	.112	.292	.476	.630	.749	.820	.861	.896
		SRT ( $\bar{X}$ )	.105	.235	.397	.527	.625	.690	.748	.800

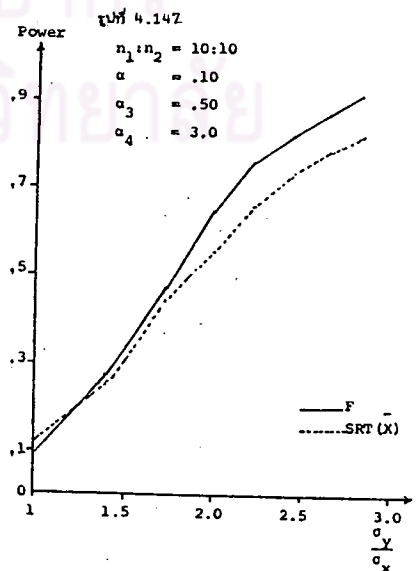
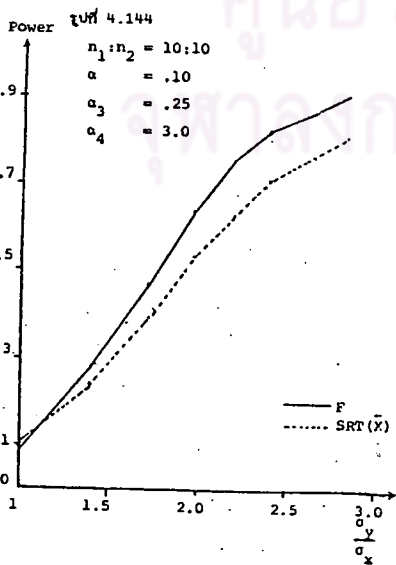
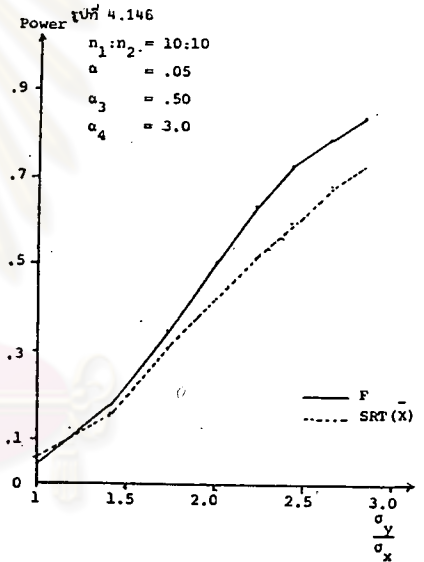
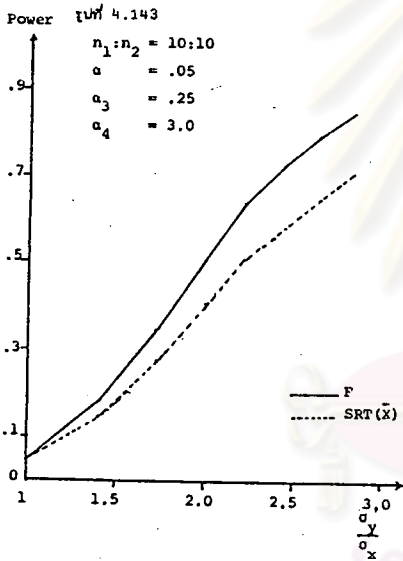
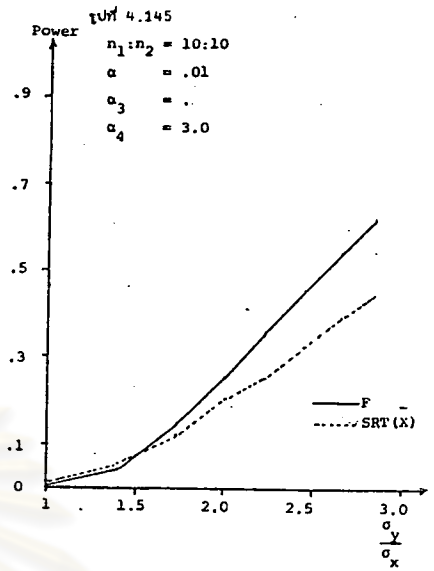
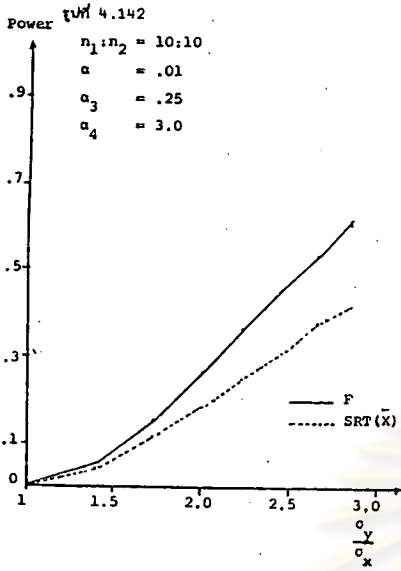


ตารางที่ 4.31/ แสดงอำนาจของการทดสอบเอฟ และการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) เมื่อประชากรมีการแจกแจงแบบเบ้ โดยมีความเบ้ .50 และขนาดตัวอย่างทั้งสองกลุ่มเป็น (10:10) จำแนกตามระดับนัยสำคัญที่กำหนด และความโค้งของประชากร

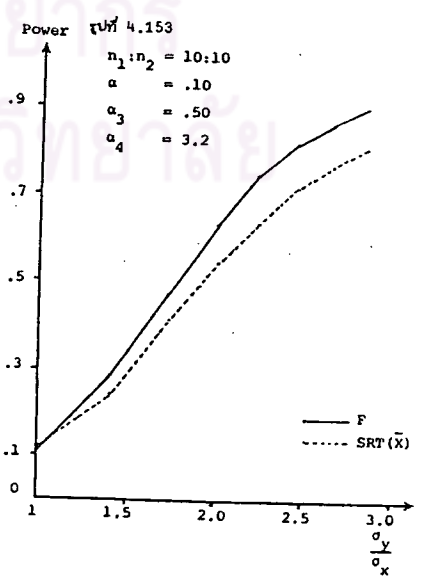
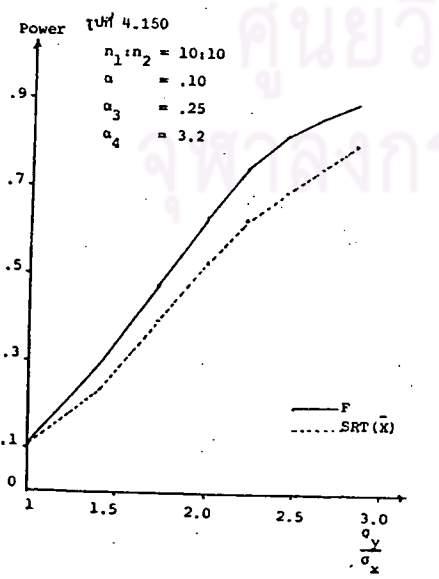
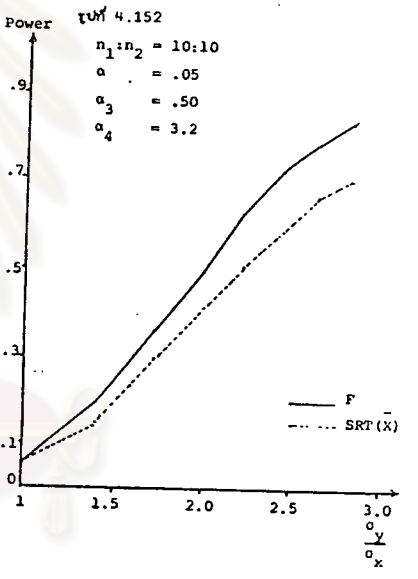
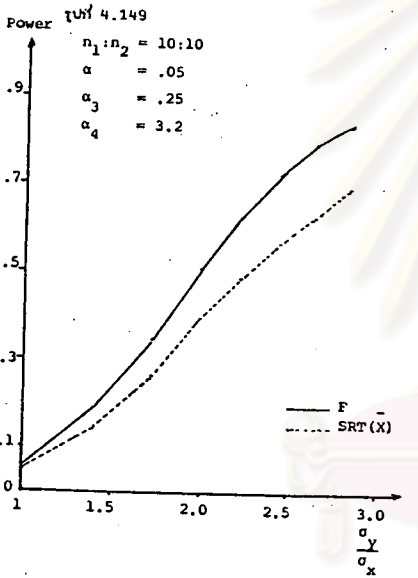
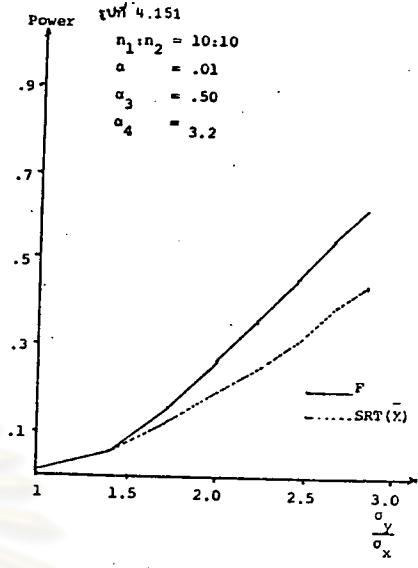
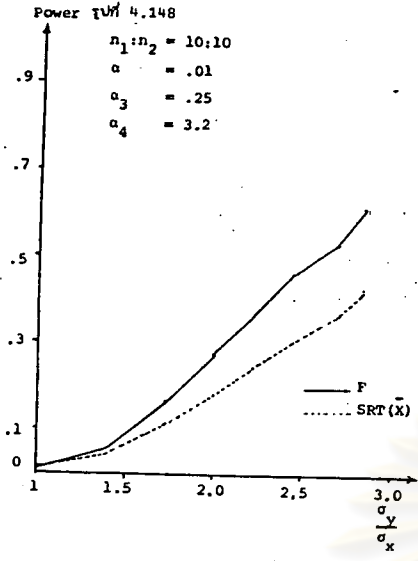
ระดับ นัยสำคัญ	ความ โค้ง	สถิติ ทดสอบ	$\Delta^2$							
			1	2	3	4	5	6	7	8
.01	2.8	F	.004	.040	.137	.245	.353	.447	.542	.630
		SRT ( $\bar{X}$ )	.014	.055	.131	.202	.271	.354	.403	.464
	3.0	F	.006	.046	.147	.257	.357	.452	.546	.619
		SRT ( $\bar{X}$ )	.011	.055	.122	.202	.257	.330	.391	.443
	3.2	F	.007	.054	.155	.268	.361	.453	.546	.613
		SRT ( $\bar{X}$ )	.012	.054	.123	.190	.248	.310	.387	.431
.05	2.8	F	.037	.161	.342	.497	.638	.732	.792	.843
		SRT ( $\bar{X}$ )	.062	.165	.320	.436	.533	.620	.686	.741
	3.0	F	.043	.174	.349	.502	.633	.727	.788	.837
		SRT ( $\bar{X}$ )	.059	.159	.308	.421	.515	.598	.668	.726
	3.2	F	.053	.191	.351	.498	.623	.720	.788	.830
		SRT ( $\bar{X}$ )	.059	.149	.297	.409	.502	.583	.661	.700
.10	2.8	F	.083	.263	.470	.646	.758	.826	.872	.923
		SRT ( $\bar{X}$ )	.130	.261	.454	.572	.678	.747	.794	.820
	3.0	F	.097	.273	.470	.643	.753	.819	.867	.911
		SRT ( $\bar{X}$ )	.120	.256	.443	.554	.657	.731	.781	.813
	3.2	F	.111	.284	.471	.632	.742	.817	.858	.896
		SRT ( $\bar{X}$ )	.121	.247	.416	.547	.635	.713	.768	.806







ศูนย์วิทยทรัพยากร  
 จุฬาลงกรณ์มหาวิทยาลัย



ศูนย์วิทยทรัพยากร  
 จุฬาลงกรณ์มหาวิทยาลัย

จากตาราง 4.30-4.31 และจากรูป 4.136-4.153 ซึ่งแสดงอำนาจของการทดสอบเอฟและการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) เมื่อประชากรมีการแจกแจงแบบเบ้ โดยมีความเบ้ .25 และ .50 และความโด่งมีค่า 2.8, 3.0 และ 3.2 โดยที่ขนาดของตัวอย่างทั้งสองกลุ่มเป็น (10:10) สามารถเปรียบเทียบอำนาจของการทดสอบทั้งสองวิธีเมื่อการทดสอบทั้งสองวิธีควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดังนี้

การทดสอบเอฟมีอำนาจของการทดสอบสูงกว่าการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) ทุก ๆ ค่าของ  $\alpha$  ที่กำหนด ยกเว้นกรณีที่  $\Delta^2 = 2 \left( \frac{\sigma_y}{\sigma_x} = 1.414 \right)$  มีบางกรณีที่การทดสอบทั้งสองวิธีมีอำนาจของการทดสอบแตกต่างกันเพียงเล็กน้อย

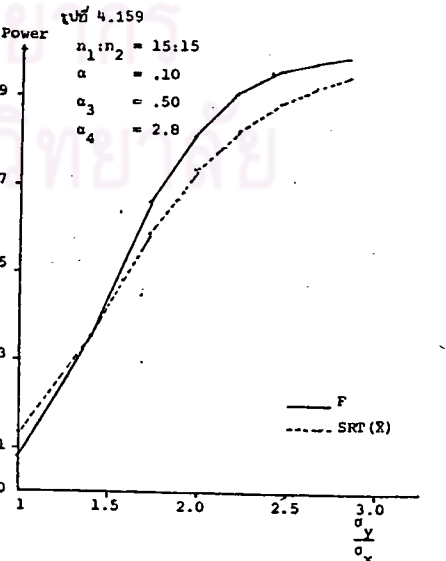
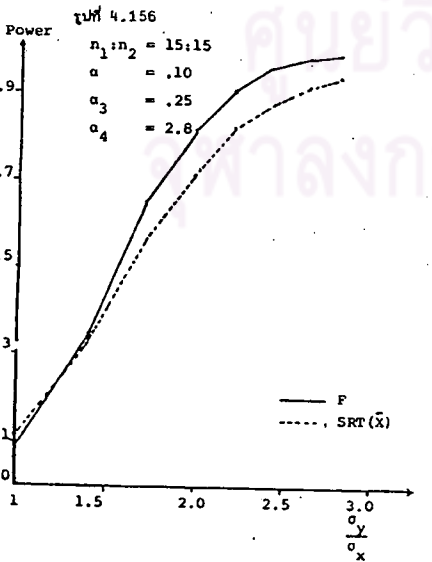
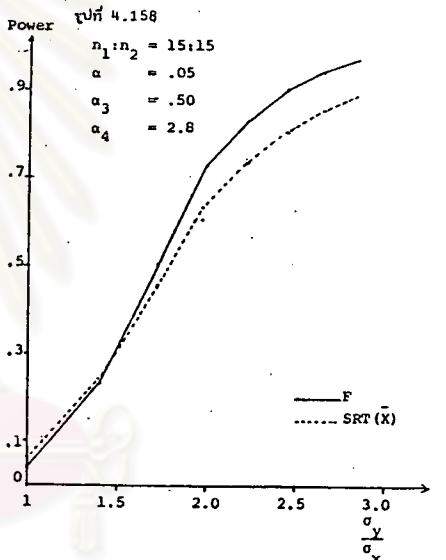
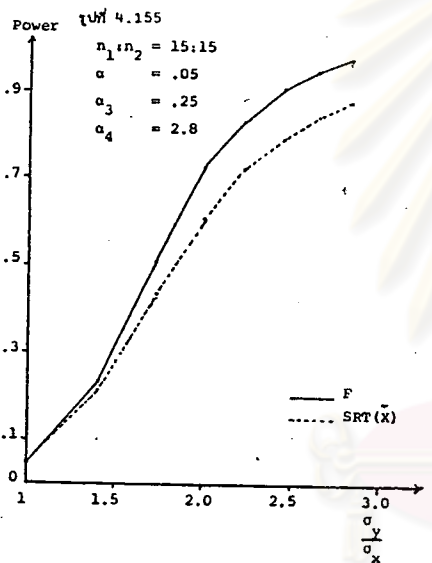
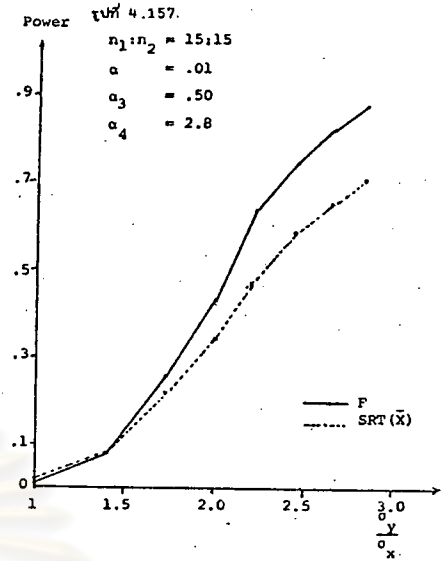
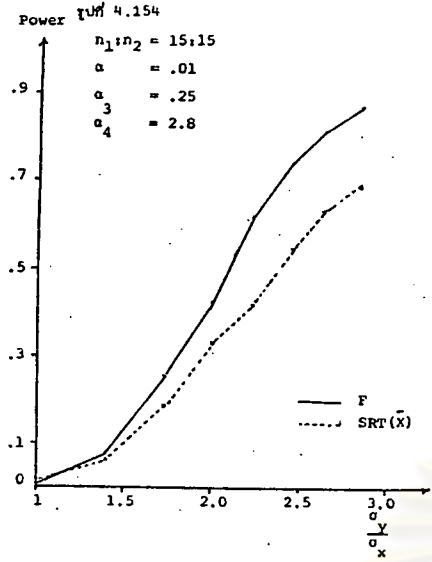
ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.32 แสดงอำนาจของการทดสอบเอฟ และการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) เมื่อประชากรมีการแจกแจงแบบเบ้ โดยมีความเบ้ .25 และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (15:15) ค่าแจกตามระดับนัยสำคัญที่กำหนด และความโค้งของประชากร

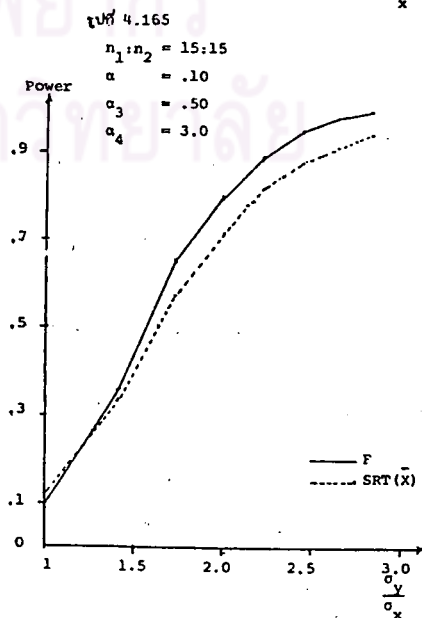
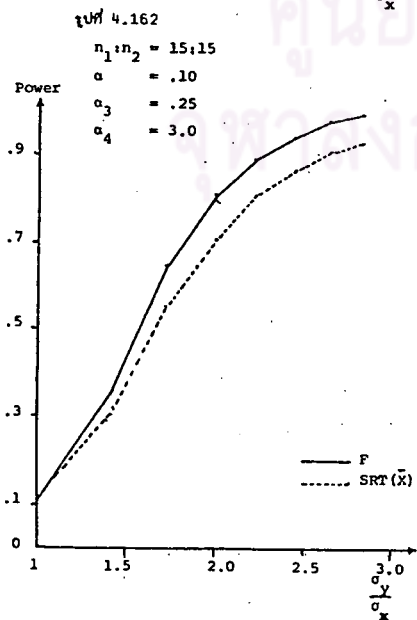
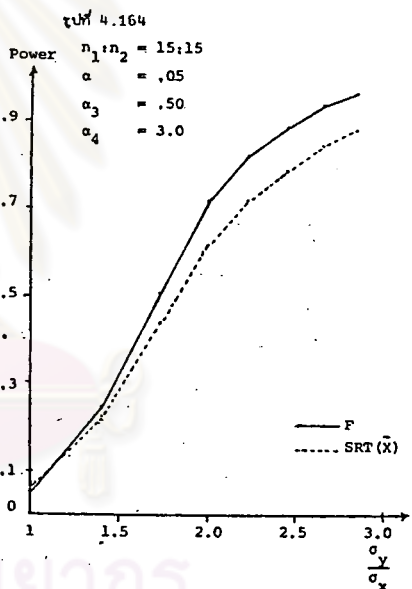
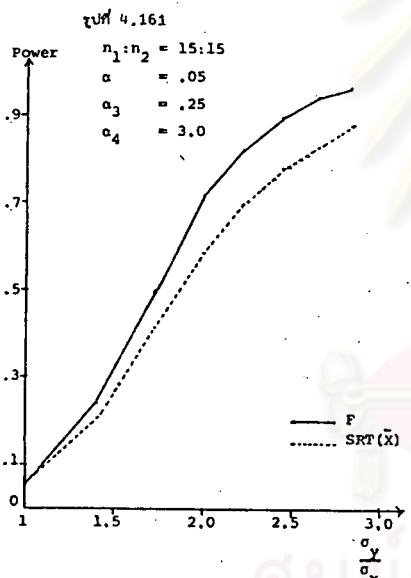
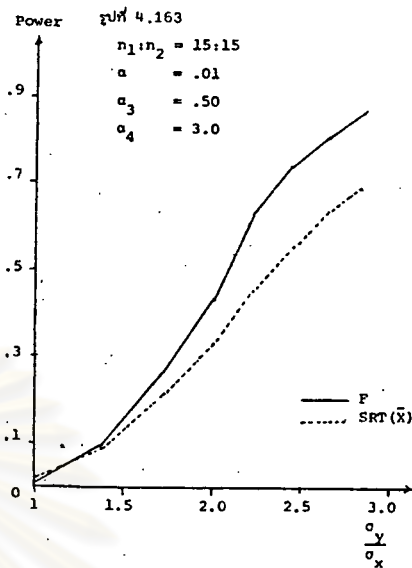
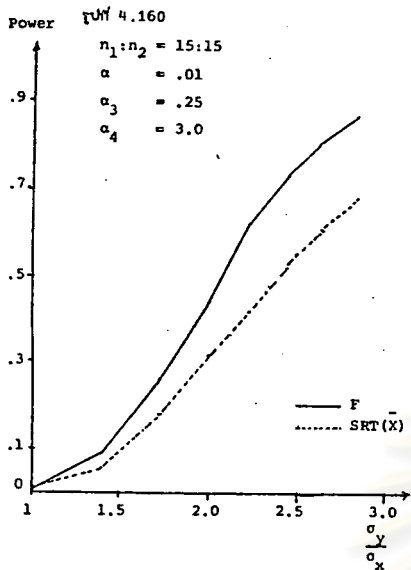
ระดับ นัยสำคัญ	ความ โค้ง	สถิติ ทดสอบ	$\Delta^2$							
			1	2	3	4	5	6	7	8
.01	2.8	F	.007	.079	.252	.431	.619	.741	.813	.874
		SRT ( $\bar{X}$ )	.016	.068	.189	.330	.429	.545	.632	.689
	3.0	F	.012	.088	.260	.435	.617	.729	.805	.863
		SRT ( $\bar{X}$ )	.016	.065	.177	.314	.415	.525	.607	.673
	3.2	F	.014	.094	.268	.446	.618	.721	.795	.852
		SRT ( $\bar{X}$ )	.015	.065	.168	.294	.403	.505	.589	.658
.05	2.8	F	.045	.231	.505	.720	.828	.404	.449	.974
		SRT ( $\bar{X}$ )	.047	.214	.433	.607	.719	.787	.842	.878
	3.0	F	.055	.239	.498	.712	.815	.896	.940	.962
		SRT ( $\bar{X}$ )	.054	.206	.411	.586	.698	.777	.825	.873
	3.2	F	.069	.249	.496	.700	.813	.878	.734	.959
		SRT ( $\bar{X}$ )	.057	.201	.398	.565	.683	.764	.818	.864
.10	2.8	F	.098	.345	.650	.811	.904	.956	.982	.988
		SRT ( $\bar{X}$ )	.112	.331	.564	.713	.820	.875	.915	.936
	3.0	F	.110	.356	.646	.802	.896	.945	.974	.988
		SRT ( $\bar{X}$ )	.117	.321	.550	.703	.805	.867	.907	.927
	3.2	F	.123	.363	.640	.792	.880	.942	.967	.982
		SRT ( $\bar{X}$ )	.116	.310	.534	.690	.785	.859	.900	.922

ตารางที่ 4.33 แสดงอำนาจของการทดสอบเอฟ และการทดสอบค่าเฉลี่ย ( $\bar{X}$ ) เมื่อประชากรมีการแจกแจงแบบเบ้ โดยมีความเบ้ .50 และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (15:15) จำแนกตามระดับนัยสำคัญที่กำหนด และความโค้งของประชากร

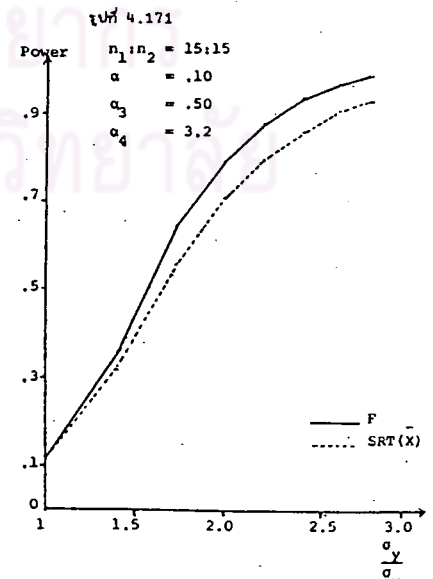
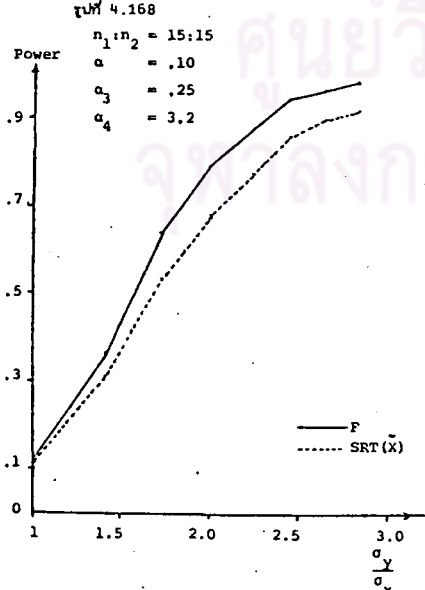
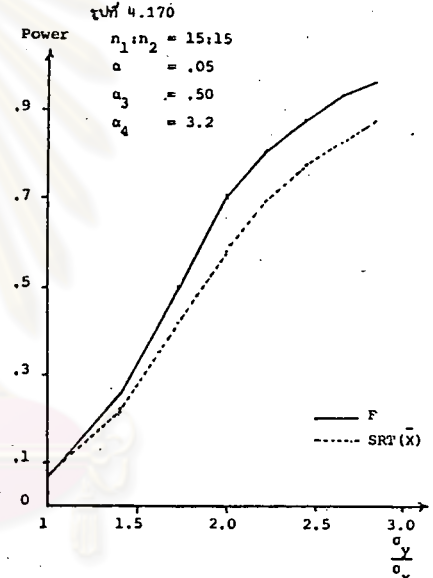
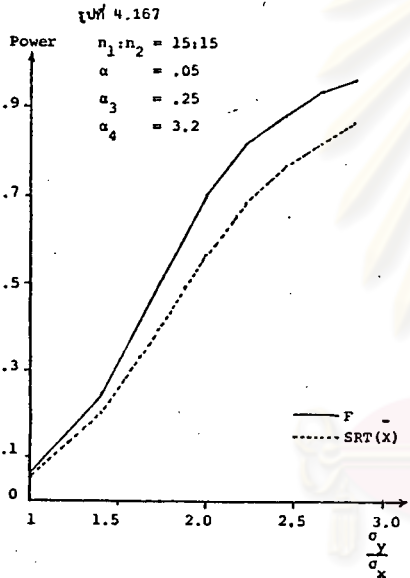
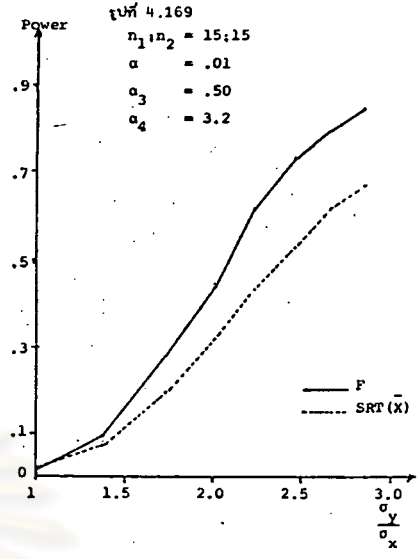
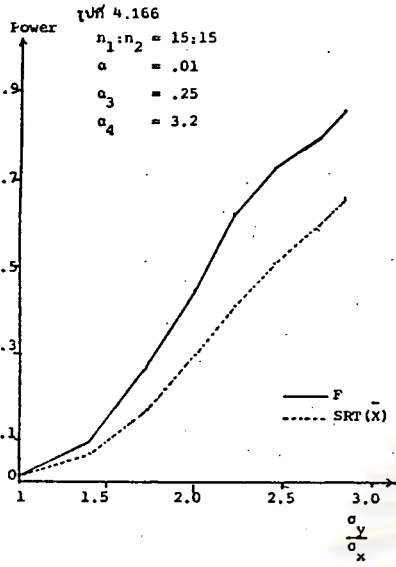
ระดับ นัยสำคัญ	ความ โค้ง	สถิติ ทดสอบ	2 $\Delta$							
			1	2	3	4	5	6	7	8
.01	2.8	F	.010	.081	.257	.431	.631	.744	.816	.872
		SRT ( $\bar{X}$ ) *	.020	.081	.219	.346	.473	.583	.659	.707
	3.0	F	.014	.085	.266	.432	.624	.738	.801	.860
		SRT ( $\bar{X}$ ) *	.020	.078	.210	.330	.450	.551	.631	.684
	3.2	F *	.017	.091	.277	.434	.601	.731	.793	.850
		SRT ( $\bar{X}$ ) *	.020	.073	.196	.316	.430	.529	.616	.669
.05	2.8	F	.047	.233	.503	.728	.831	.902	.947	.972
		SRT ( $\bar{X}$ )	.065	.241	.459	.633	.733	.804	.852	.888
	3.0	F	.053	.246	.502	.714	.817	.888	.938	.964
		SRT ( $\bar{X}$ )	.064	.225	.436	.612	.717	.786	.841	.877
	3.2	F	.062	.255	.501	.704	.807	.878	.931	.960
		SRT ( $\bar{X}$ )	.064	.223	.420	.589	.698	.773	.826	.875
.10	2.8	F	.089	.351	.660	.811	.902	.955	.978	.989
		SRT ( $\bar{X}$ )	.132	.351	.591	.734	.821	.885	.920	.941
	3.0	F	.102	.351	.654	.799	.889	.948	.975	.988
		SRT ( $\bar{X}$ )	.125	.338	.574	.719	.812	.872	.907	.934
	3.2	F	.119	.356	.644	.791	.878	.934	.956	.985
		SRT ( $\bar{X}$ )	.119	.324	.555	.707	.791	.858	.901	.926







ศูนย์วิทยพัชกร  
 ภาควิชาคณิตศาสตร์  
 มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี



จากตาราง 4.32-4.33 และจากรูป 4.154-4.171 ซึ่งแสดงอำนาจของการทดสอบเอฟและการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) เมื่อประชากรมีการแจกแจงแบบเบ้ โดยมีความเบ้ .25 และ .50 และความโค้งมีค่า 2.8, 3.0 และ 3.2 โดยที่ขนาดของตัวอย่างทั้งสองกลุ่มเป็น (15:15) สามารถเปรียบเทียบอำนาจของการทดสอบทั้งสองวิธีเมื่อการทดสอบทั้งสองวิธีควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดังนี้

การทดสอบเอฟมีอำนาจของการทดสอบสูงกว่าการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) ทุก ๆ ค่าของ  $\alpha$  ที่กำหนด ยกเว้นกรณีที่  $\Delta^2 = 2 \left( \frac{\sigma_Y}{\sigma_X} \right) = 1.414$  มีบางกรณีที่การทดสอบทั้งสองวิธีมีอำนาจของการทดสอบแตกต่างกันเพียงเล็กน้อย.



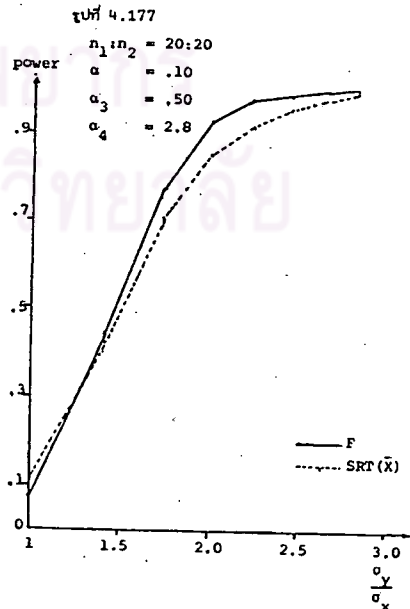
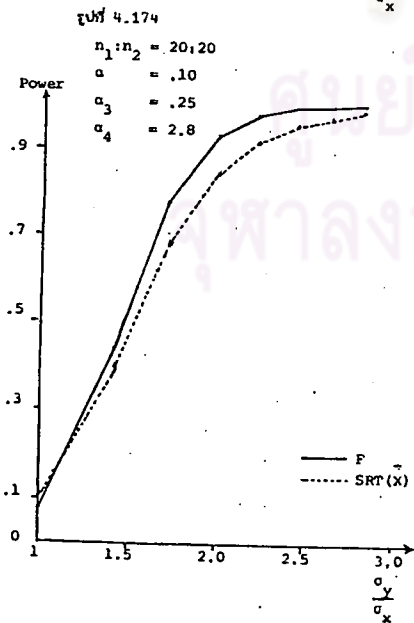
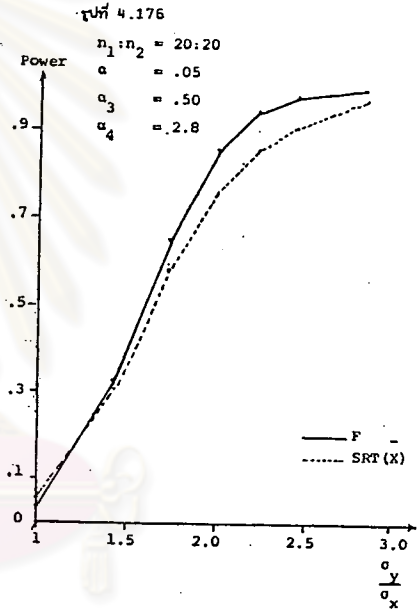
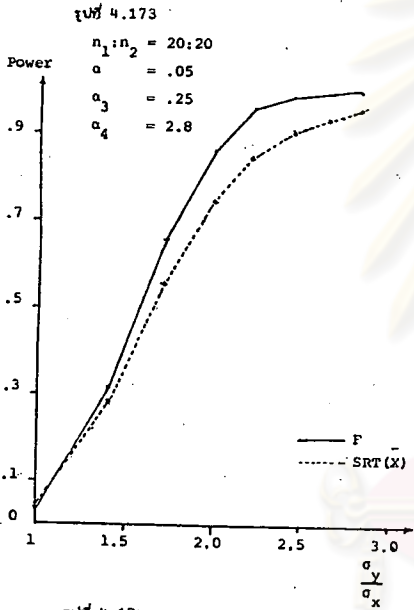
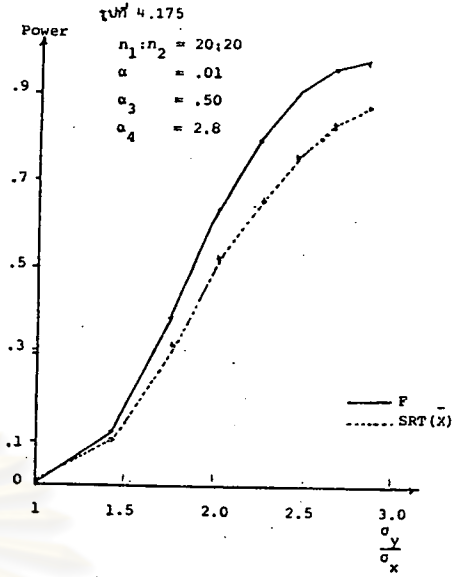
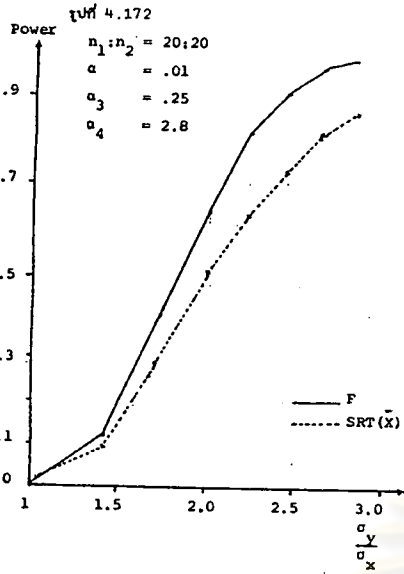
ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

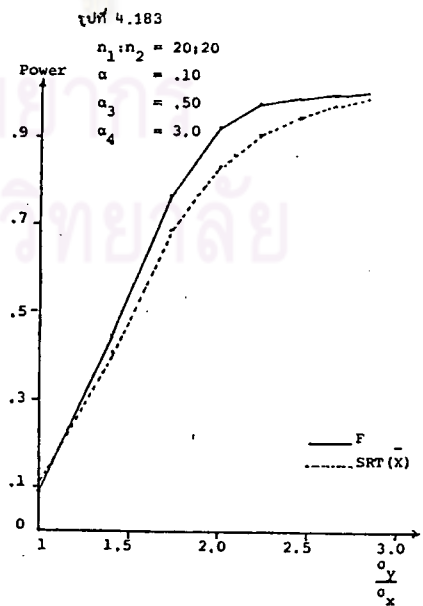
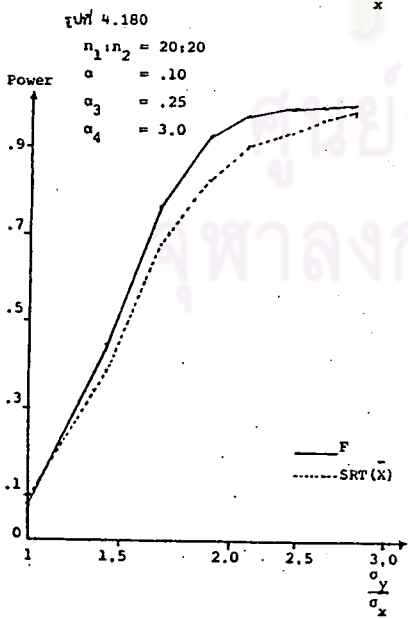
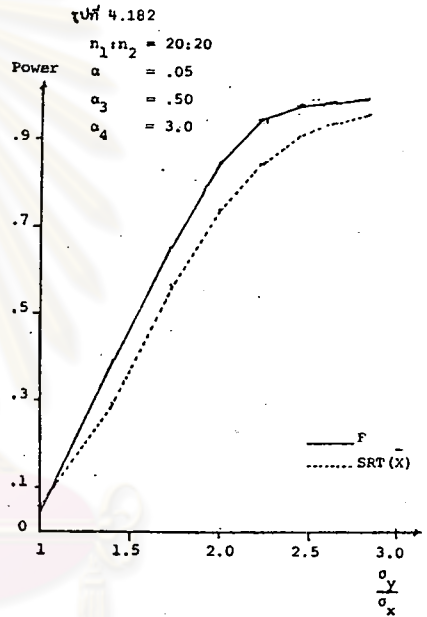
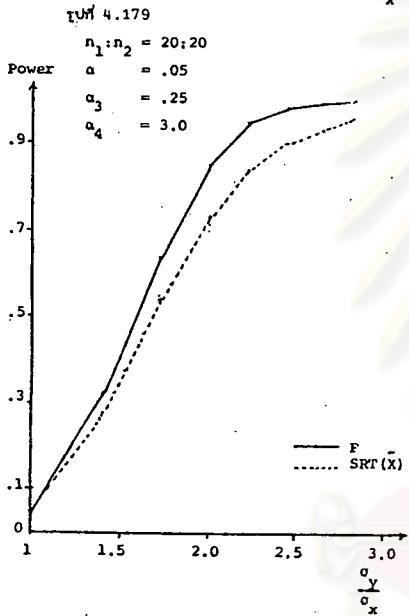
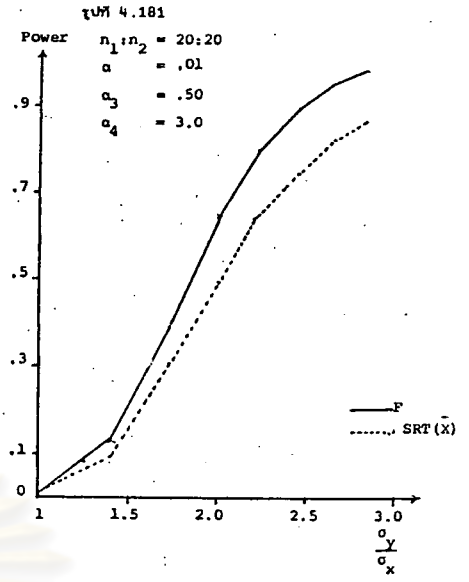
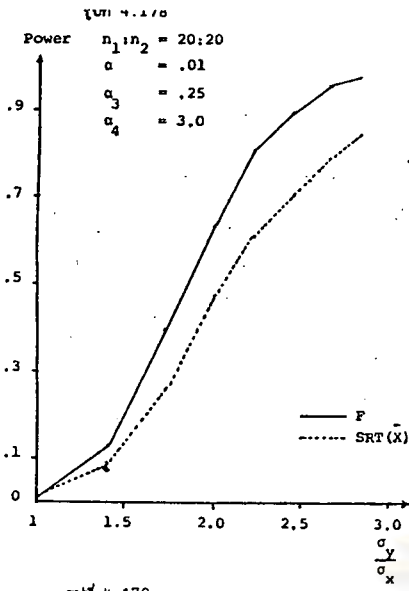
- ตารางที่ 4.34 แสดงอำนาจของการทดสอบเอฟ และการทดสอบสี่แควร์แรงค์( $\bar{X}$ ) เมื่อประชากรมีการแจกแจงแบบเบ้ โดยมีความเบ้ .25 และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (20:20) จำแนกตามระดับนัยสำคัญที่กำหนด และความโค้งของประชากร

ระดับ นัยสำคัญ	ความ โค้ง	สถิติ ทดสอบ	$\Delta^2$							
			1	2	3	4	5	6	7	8
.01	2.8	F	.007	.119	.397	.638	.811	.904	.962	.981
		SRT ( $\bar{X}$ )	.011	.091	.285	.497	.626	.729	.811	.856
	3.0	F	.010	.129	.392	.630	.802	.897	.956	.975
		SRT ( $\bar{X}$ )	.011	.086	.260	.471	.606	.706	.790	.846
	3.2	F	.012	.139	.396	.630	.795	.891	.948	.969
		SRT ( $\bar{X}$ )	.012	.086	.248	.456	.587	.685	.774	.836
.05	2.8	F	.036	.315	.658	.860	.956	.984	.992	.995
		SRT ( $\bar{X}$ )	.048	.281	.553	.795	.848	.904	.934	.956
	3.0	F	.042	.826	.643	.849	.945	.980	.990	.994
		SRT ( $\bar{X}$ )	.048	.275	.537	.728	.834	.894	.929	.952
	3.2	F	.050	.329	.638	.844	.935	.974	.987	.993
		SRT ( $\bar{X}$ )	.047	.267	.523	.712	.821	.883	.923	.944
.10	2.8	F	.078	.439	.774	.928	.979	.992	.997	.997
		SRT ( $\bar{X}$ )	.101	.400	.679	.840	.911	.950	.969	.988
	3.0	F	.085	.447	.764	.921	.972	.990	.994	.997
		SRT ( $\bar{X}$ )	.100	.388	.672	.828	.903	.936	.963	.984
	3.2	F	.100	.443	.755	.909	.968	.987	.993	.996
		SRT ( $\bar{X}$ )	.101	.370	.659	.815	.891	.933	.962	.975

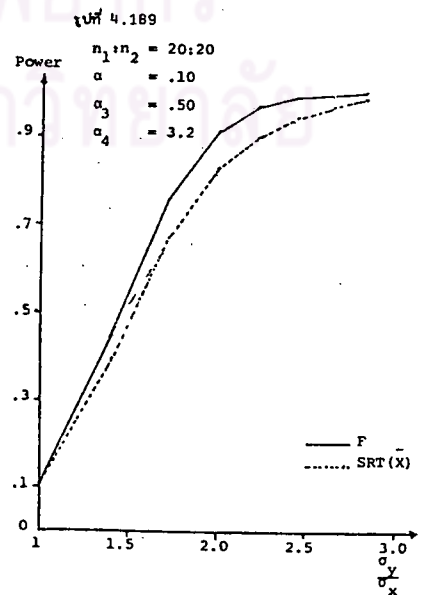
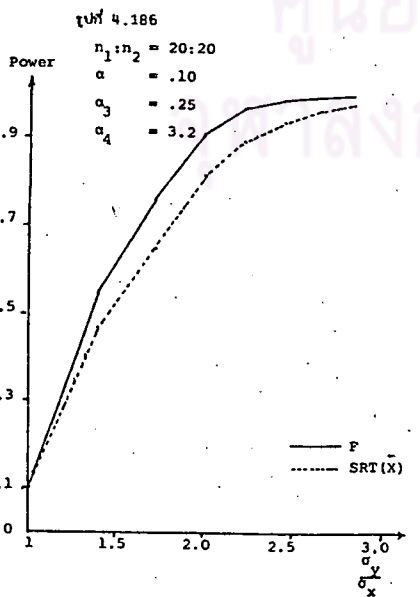
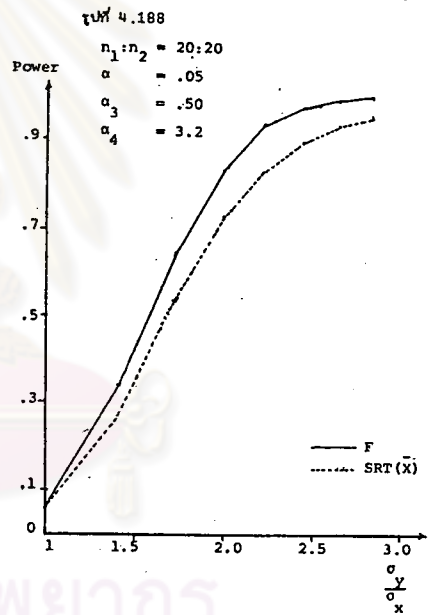
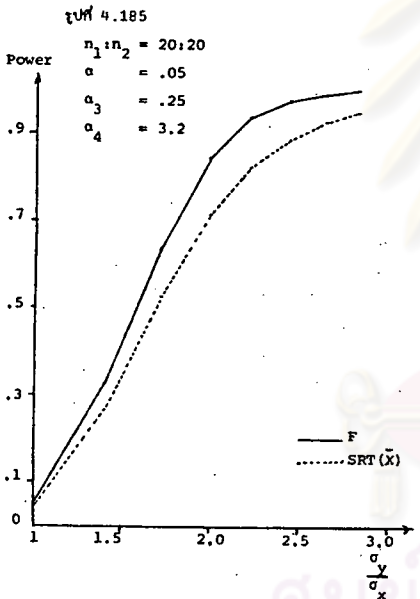
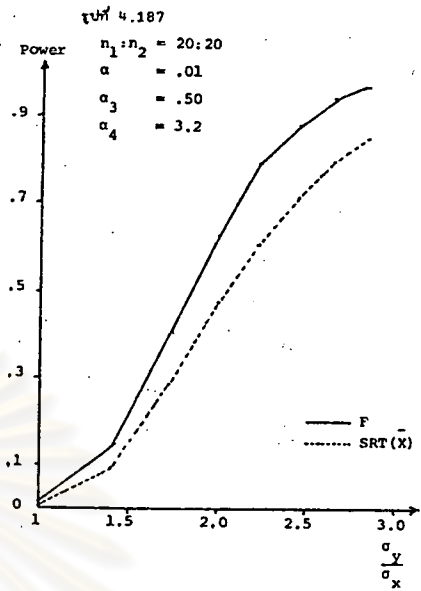
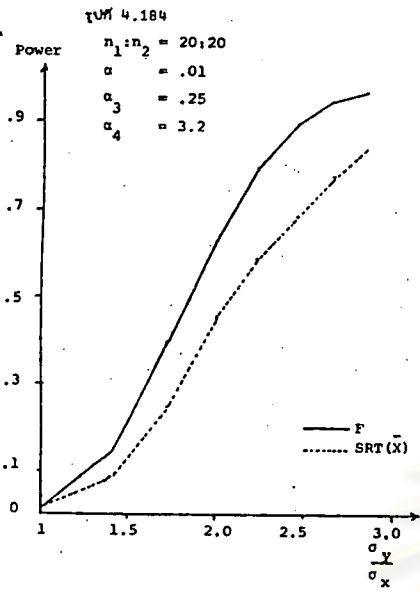
ตารางที่ 4.35 แสดงอำนาจของการทดสอบเอฟ และการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) เมื่อประชากรมีการแจกแจงแบบเบ้ โดยมีความเบ้ :50 และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (20:20) จำแนกตามระดับนัยสำคัญที่กำหนด และความโค้งของประชากร

ระดับ นัยสำคัญ	ความ โค้ง	สถิติ ทดสอบ	$\Delta^2$							
			1	2	3	4	5	6	7	8
.01	2.8	F	.004	.124	.388	.638	.797	.904	.956	.977
		SRT ( $\bar{X}$ )	.010	.109	.326	.524	.659	.755	.832	.874
	3.0	F	.006	.133	.391	.636	.791	.889	.849	.976
		SRT ( $\bar{X}$ )	.010	.097	.304	.498	.633	.739	.814	.856
	3.2	F	.013	.139	.400	.626	.791	.878	.941	.969
		SRT ( $\bar{X}$ )	.010	.094	.236	.474	.607	.717	.793	.848
.05	2.8	F	.034	.327	.650	.855	.946	.981	.989	.994
		SRT ( $\bar{X}$ )	.058	.307	.582	.760	.853	.909	.943	.970
	3.0	F	.043	.391	.650	.844	.941	.978	.988	.992
		SRT ( $\bar{X}$ )	.056	.289	.566	.738	.841	.902	.936	.959
	3.2	F	.053	.340	.643	.830	.933	.972	.987	.990
		SRT ( $\bar{X}$ )	.054	.270	.548	.727	.827	.894	.930	.948
.10	2.8	F	.075	.441	.769	.928	.977	.989	.996	.997
		SRT ( $\bar{X}$ )	.111	.428	.706	.849	.916	.954	.978	.990
	3.0	F	.089	.443	.767	.919	.974	.988	.994	.997
		SRT ( $\bar{X}$ )	.112	.414	.684	.833	.905	.948	.971	.985
	3.2	F	.106	.445	.756	.909	.967	.987	.990	.997
		SRT ( $\bar{X}$ )	.107	.399	.663	.823	.897	.940	.967	.983









จากตาราง 4.34-4.35 และจากรูป 4.172-4.189 ซึ่งแสดงอำนาจของการทดสอบเอฟและการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) เมื่อประยาการมีการแจกแจงแบบเบ้ โดยมีความเบ้ .25 และ .50 และความโด่งมีค่า 2.8, 3.0 และ 3.2 โดยที่ขนาดของตัวอย่างทั้งสองกลุ่มเป็น (20:20) สามารถเปรียบเทียบอำนาจของการทดสอบทั้งสองวิธีเมื่อการทดสอบทั้งสองวิธีควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดังนี้

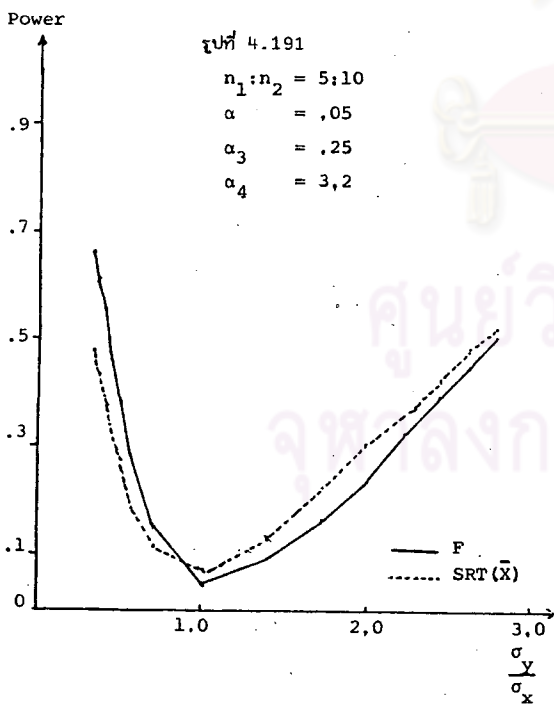
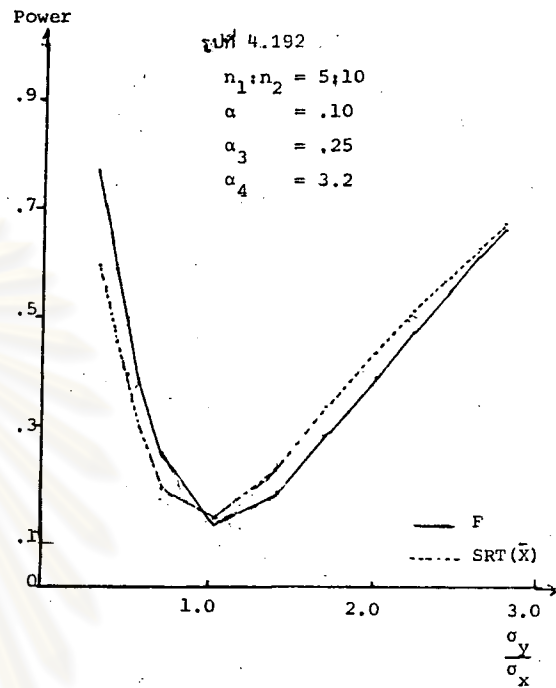
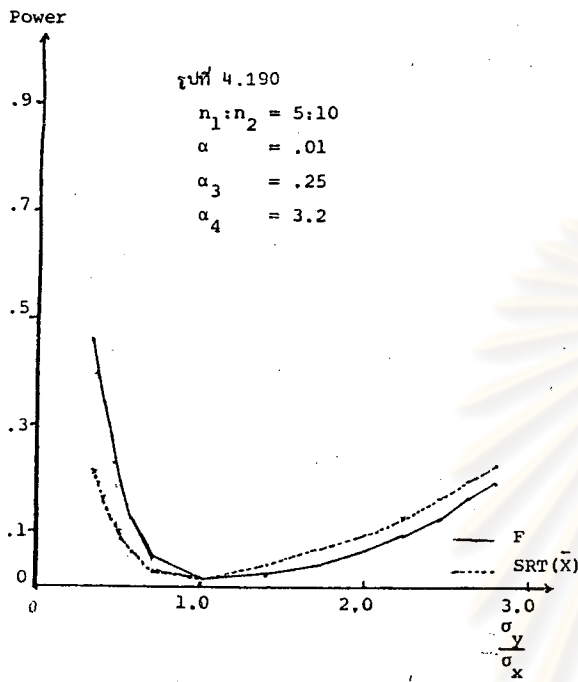
การทดสอบเอฟมีอำนาจของการทดสอบสูงกว่าการทดสอบสแควร์แรงค์ ( $\bar{X}$ )  
ทุก ๆ ค่าของ  $\alpha$  ที่กำหนด



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตาราง 4.36 แสดงอำนาจของการทดสอบเอฟ และการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) เมื่อประชากรมีการแจกแจงแบบเบ้ โดยมีความเบ้ .25 และ ความโค้ง 3.2 และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (5:10) จำแนกตามระดับนัยสำคัญที่กำหนด

ระดับ นัยสำคัญ	สถิติทดสอบ	$\Delta^2$														
		$\frac{1}{8}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	1	2	3	4	5	6	7	8
F	F	.455	.394	.343	.278	.199	.124	.051	.013	.018	.039	.067	.093	.129	.161	.194
	SRT ( $\bar{X}$ )	.217	.188	.153	.122	.088	.060	.024	.016	.040	.068	.093	.128	.165	.197	.225
F	F	.662	.612	.553	.473	.386	.284	.154	.052	.093	.166	.237	.327	.389	.449	.501
	SRT ( $\bar{X}$ )	.477	.436	.394	.339	.273	.185	.114	.074	.131	.223	.303	.360	.427	.486	.520
F	F	.764	.709	.651	.588	.497	.379	.242	.112	.170	.283	.389	.471	.546	.611	.659
	SRT ( $\bar{X}$ )	.598	.553	.504	.452	.392	.292	.178	.129	.219	.336	.433	.510	.570	.620	.669



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

จากตาราง 4.36 และจากรูป 4.190-4.192 ซึ่งแสดงอำนาจของการทดสอบเอฟและการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) เมื่อประชากรมีการแจกแจงแบบเบ้ โดยมีความเบ้ .25 ความโด่ง 3.2 และขนาดของตัวอย่างทั้งสองกลุ่มเป็น (5:10) ปรากฏว่าการทดสอบเอฟมีอำนาจของการทดสอบสูงกว่าการทดสอบสแควร์แรงค์ เมื่อ  $\Delta^2$  น้อยกว่า 1 แต่ในกรณีที่  $\Delta^2$  มากกว่า 1 การทดสอบสแควร์แรงค์ ( $\bar{X}$ ) มีอำนาจของการทดสอบสูงกว่าการทดสอบเอฟ โดยที่ค่าของอำนาจของการทดสอบสแควร์แรงค์สูงกว่าการทดสอบเอฟ จะมีความมากเมื่อ  $\Delta^2$  มีค่า 2, 3 และ 4 แต่เมื่อ  $\Delta^2$  มีค่าเป็น 6, 7 และ 8 อำนาจของการทดสอบสแควร์แรงค์ ( $\bar{X}$ ) ที่สูงกว่าอำนาจของการทดสอบเอฟจะมีค่าน้อยมาก



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย