16 JAN 2003

AN EMPIRICAL COMPARISON OF THE POWERS OF BARTLETT'S TEST, LEVENE'S TEST AND COCHRAN'S TEST UNDER VIOLATION OF ASSUMPTIONS

SINEENART TAEJAROENKUL

With compliments

บัณฑิตวิทยาลัย มหาวิทยาลัยมหิดล

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF SCIENCE (BIOSTATISTICS)
FACULTY OF GRADUATE STUDIES
MAHIDOL UNIVERSITY
2002

TH
S616l
2002
c.1

ISBN 974-04-2688-3

COPYRIGHT OF MAHIDOL UNIVERSITY

Copyright by Mahidol University

Thesis

entitled

AN EMPIRICAL COMPARISON OF THE POWERS OF BARTLETT'S TEST, LEVENE'S TEST AND COCHRAN'S TEST UNDER VIOLATION OF ASSUMPTIONS

Sineenart Taejaroenkul
.....
Miss. Sineenart Taejaroenkul
Candidate

Thavatchai Vorapongsathorn
.....
Assoc. Prof. Thavatchai Vorapongsathorn, Ph.D.
Major-Advisor

Chukiat Viwatwongkasem
.....
Assoc. Prof. Chukiat Viwatwongkasem, M.Sc.
Co-Advisor

.....
Assoc. Prof. Rassmidara Hoonsawat
Ph.D.
Dean
Faculty of Graduate Studies

.....
Asst Prof. Junya Pattaraarchachai,
Sc.D.
Chair
Master of Science Programme
In Biostatistics
Faculty of Public Health

Copyright by Mahidol University

Thesis

entitled

AN EMPIRICAL COMPARISON OF THE POWERS OF BARTLETT'S TEST, LEVENE'S TEST AND COCHRAN'S TEST UNDER VIOLATION OF ASSUMPTIONS

was submitted to the Faculty of Graduate Studies, Mahidol University for the degree of Master of Science (Biostatistics)

on

November 22, 2002

Sineenart Taejaroenkul
.....
Miss. Sineenart Taejaroenkul
Candidate

Thavatchai Vorapongsathorn
.....
Assoc. Prof. Thavatchai Vorapongsathorn, Ph.D.
Chair

Chulaluk Komoltri
.....
Miss. Chulaluk Komoltri, Dr.P.H.
Member

Chukiat Viwatwongkasem
.....
Assoc. Prof. Chukiat Viwatwongkasem, M.Sc.
Member

Rassm Hoon
.....
Assoc Prof. Rassmidara Hoonsawat
Ph.D.
Dean
Faculty of Graduate Studies
Mahidol University

C. Chaikittiporn
.....
Assoc. Prof. Chalermchai Chaikittiporn, Dr.P.H. (Epidemiology)
Dean
Faculty of Public Health
Mahidol University

Copyright by Mahidol University

ACKNOWLEDGEMENT

I would like to express my deep appreciation to Dr. Thavatchai Vorapongsathorn my major advisor for his invaluable advice, guidance and kindness. I am equally grateful to Mr. Chukiat Viwatwongkasem my co-adviser for his constructive comments.

I would like to thank Ms. Sasivimol Rattanasiri and my friends for their great guidance in programming, suggestions for solving technical problems, advice in statistics, and wonderful friendships. I also would like to thank Mr. Surachet Taejaroenkul for his help in proofreading.

I wish to thank the staff of Biostatistics, Faculty of Public Health, Mahidol University for their co-operation. In particular, I would like to thank Mrs. Antika Wongthanee, Research Institute for Health Science who pushes me to learn and get advancement.

I would like to thank my family and my colleague for their encouragement and kindness.

I am particularly indebted to Research Institute for Health Science, Chiangmai University, Thailand for time enabled me to undertake this study.

Sineenart Taejaroenkul

Copyright by Mahidol University

4036109 PHBS/M : MAJOR : BIOSTATISTICS ; M.Sc. (BIOSTATISTICS)
KEY WORDS : HOMOGENEITY OF VARIANCE / BARTLETT'S TEST / LEVENE'S TEST / COCHRAN'S TEST / SIMULATION

SINEENART TAEJAROENKUL : AN EMPIRICAL COMPARISON OF THE POWERS OF BARTLETT'S TEST, LEVENE'S TEST AND COCHRAN'S TEST UNDER VIOLATION OF ASSUMPTIONS. THESIS ADVISORS : THAVATCHAI VORAPONGSATHORN, Ph.D., CHUKIAT VIWATWONGKASEM, M.Sc. 107 P. ISBN 974-04-2688-3

This study compared the probability of Type I error and the power of homogeneity of three variance statistical test Bartlett's test, Levene's test and Cochran's test in sampling distribution effect, variance effect and sample size effect by using Monte Carlo method to generate sample sizes and distributions for 1,000 times. The sample sizes were both equal and unequal: 15, 30 and 45. The distributions were same sample distributions and different distributions: Normal distribution, Gamma distribution and Chi-square distribution. It was found that Bartlett's test was sensitive to the normal assumption whereas Cochran's test and Levene's test were robust when the normal assumption was violated. Moreover, Levene's test was quite good for both equal sample size and sample size of 15. In the case of power, Bartlett's test was very useful because it had the highest power in all cases. When one variance was large, Cochran's test was the best test.

The implications of the study are as follows: Bartlett's test is the best test for homogeneity of variances since the sample size does not affect it. When data are non-normally distributed, Levene's test is best for both small and equal sample sizes. Cochran's test is best when sample size is large and unequal and one variance is larger.

Copyright by Mahidol University

4036109 PHBS/M : สาขาวิชา : ชีวสถิติ; วท.ม. (ชีวสถิติ)

สินีนาถ แต่เจริญกุล : การเปรียบเทียบอำนาจการทดสอบของสถิติ Bartlett's Test, Levene's Test และ Cochran's Test ภายใต้การละเมิดข้อตกลงเบื้องต้น (AN EMPIRICAL COMPARISON OF THE POWERS OF BARTLETT'S TEST, LEVENE'S TEST AND COCHRAN'S TEST UNDER VIOLATION OF ASSUMPTIONS). คณะกรรมการควบคุมวิทยานิพนธ์: ธวัชชัย วรพงศธร, Ph.D., ชูเกียรติ วิวัฒน์วงศ์เกษม, M.S. 107 หน้า ISBN 974-04-2688-3

การวิจัยนี้มีวัตถุประสงค์ที่จะเปรียบเทียบความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และเปรียบเทียบอำนาจการทดสอบความเท่ากันของความแปรปรวนของสถิติ 3 ตัวคือ Bartlett's Test, Levene's Test And Cochran's Test ภายใต้ผลกระทบของการแจกแจงตัวอย่าง ผลกระทบของความแปรปรวนและผลกระทบจากขนาดตัวอย่าง โดยใช้วิธี Monte Carlo ในการจำลองการแจกแจงและขนาดตัวอย่าง กระทำ 1000 ครั้ง ขนาดตัวอย่างมีแบบเท่ากันและไม่เท่ากันคือ 15, 30 และ 45 การแจกแจงมีทั้งการแจกแจงเดียวกันและต่างการแจกแจงได้แก่ การแจกแจงแบบปกติ การแจกแจงแบบ Gamma และการแจกแจงแบบ Chi-square ผลการศึกษาพบว่า สถิติที่มีความเหมาะสมกับการแจกแจงปกติได้แก่ Bartlett's Test Cochran's Test และ Levene's Test ตามลำดับ เมื่อข้อมูลไม่เป็นการแจกแจงปกติพบว่า Cochran's Test และ Levene's Test มีความเหมาะสม โดย Levene's Test ค่อนข้างได้เปรียบเมื่อจำนวนตัวอย่างเท่ากับ 15 ในด้านกำลังทดสอบพบว่า Bartlett's Test ยังคงให้กำลังสูงเกือบทุกกรณี และเมื่อความแปรปรวนหนึ่งค่ามีขนาดใหญ่กว่าค่าอื่นพบว่า Cochran's Test ให้กำลังการทดสอบสูงที่สุด

ข้อเสนอแนะสำหรับการใช้สถิติทดสอบความเท่ากันของความแปรปรวนได้แก่ Bartlett's Test ยังคงเป็นวิธีที่ดีเมื่อข้อมูลมีการแจกแจงแบบปกติเนื่องจากไม่มีผลกระทบกับขนาดตัวอย่าง และเมื่อข้อมูลไม่เป็นการแจกแจงแบบปกติ Levene's Test สามารถนำไปใช้กับข้อมูลที่มีขนาดเล็กและจำนวนตัวอย่างเท่ากัน ส่วน Cochran's Test ใช้เมื่อตัวอย่างมีขนาดใหญ่และจำนวนตัวอย่างมีไม่เท่ากันและมีความแปรปรวนที่ใหญ่กว่าค่าอื่นหนึ่งค่า

Copyright by Mahidol University

CONTENTS

		Page
ACKNOWLEDGEMENT		iii
ABSTRACT		iv
LIST OF TABLES		viii
LIST OF FIGURES		x
CHAPTER		
I	INTRODUCTION	1
	A. Background and Problem	1
	B. Objectives	9
	C. Hypothesis	10
	D. Scope of the Study	11
	E. Definitions	12
II	LITERATURE REVIEW	17
	A. Bartlett's Test	17
	B. Levene's Test	21
	C. Cochran's Test	28
III	METHODOLOGY	37
	A. Methodology	37
	B. The Process of Computations	38

Copyright by Mahidol University

CONTENTS(Cont)

		Page
IV	RESULTS	46
	A. Type I Error	47
	B. Power of Test	59
V	DISCUSSION	77
	A. Type I Error	77
	B. Power of Test	80
V	CONCLUSION AND RECOMMENDATION	83
	A. Type I Error	83
	B. Power of Test	84
	C. Implication of the Study	86
	D. Recommendation for Further Studies	86
REFERENCES		88
APPENDIX A	STATISTICAL TABLES	93
APPENDIX B	FORTRAN PROGRAM TO GENERATE DATA FOR A COMPAREISON ON TYPE I ERROR AND POWER OF TEST IN NORMAL DISTRIBUTION	101
BIOGRAPHY		107

Copyright by Mahidol University

LIST OF TABLES

TABLE		Page
1.	Computation of Bartlett's Test of Homogeneity of Variance	19
2.	Scores of Anxiety of Subjects in Trials	23
3.	Anxiety Data from SPSS by Mean	24
4.	Anxiety data from SPSS by Median	25
5.	Anxiety Data from SPSS by 10% Trimmed Mean	27
6.	Example from Maass	30
7.	The Values of C	35
8.	Effect of Empirical Type I Errors under Equal Variances Hypothesis of the Three Statistical Tests for Various Nominal Significance Levels with Normal Distribution	48
9.	Effect of Empirical Type I Errors under Equal Variances Hypothesis of the Three Statistical Tests for Various Nominal Significance Levels with Gamma Distribution	51
10.	Effect of Empirical Type I Errors under Equal Variances Hypothesis of the Three Statistical Tests for Various Nominal Significance Levels with Chi-square Distribution	54

Copyright by Mahidol University

LIST OF TABLES (Cont)

TABLE	Page
11. Effect of Empirical Type I Errors under Equal Variances Hypothesis of the Three Statistical Tests for Various Nominal Significance Levels with Normal Distribution, Gamma distribution and Chi-square distribution	57
12. Effect of Heterogeneous Variances on Power of the Three Statistical Tests for Various Nominal Significance Levels with Normal Distribution	60
13. Effect of Heterogeneous Variances on Power of the Three Statistical Tests for Various Nominal Significance Levels with Gamma Distribution	64
14. Effect of Heterogeneous Variances on Power of the Three Statistical Tests for Various Nominal Significance Levels with Chi-square Distribution	68
15. Effect of Heterogeneous Variances on Power of the Three Statistical Tests for Various Nominal Significance Levels with Normal Distribution, Gamma Distribution and Chi-square Distribution	73

Copyright by Mahidol University

LIST OF FIGURES

FIGURE		Page
1.	Normal Distribution	13
2.	Chi-square Distribution	14
3.	Gamma Distribution	15
4.	The Steps of Computation of Probability of Type I Error and Power of Three Statistical Tests	45
5.	Probabilities of Type I Errors in the Three Statistical Tests with Normal Distribution	49
6.	Probabilities of Type I Errors in the Three Statistical Tests with Gamma Distribution	52
7.	Probabilities of Type I Errors in the Three Statistical Tests with Chi-square Distribution	55
8.	Probabilities of Type I Errors in the Three Statistical Tests with Normal Distribution, Gamma Distribution and Chi-square Distribution	58
9.	Power of the Three Statistical Tests on Effect of Variance Ratios and Sample Sizes with Normal Distribution at 0.01 Significance Level	61

Copyright by Mahidol University

LIST OF FIGURES (Cont)

FIGURE	Page
10. Power of the Three Statistical Tests on Effect of Variance Ratios and Sample Sizes with Normal Distribution at 0.05 Significance Level	62
11. Power of the Three Statistical Tests on Effect of Variance Ratios and Sample Sizes with Gamma Distribution at 0.01 Significance Level	65
12. Power of the Three Statistical Tests on Effect of Variance Ratios and Sample Sizes with Gamma Distribution at 0.05 Significance Level	66
13. Power of the Three Statistical Tests on Effect of Variance Ratios and Sample Sizes with Chi-square Distribution at 0.01 Significance Level	69
14. Power of the Three Statistical Tests on Effect of Variance Ratios and Sample Sizes with Chi-square Distribution at 0.05 Significance Level	70

Copyright by Mahidol University

LIST OF FIGURES (Cont)

FIGURE **Page**

15. Power of the Three Statistical Tests on Effect of Variance Ratios and Sample Sizes with Normal Distribution, Gamma Distribution and Chi-square Distribution at 0.01 Significance Level 74

16. Power of the Three Statistical Tests on Effect of Variance Ratios and Sample Sizes with Normal Distribution, Gamma Distribution and Chi-square Distribution at 0.05 Significance Level 75

Copyright by Mahidol University

CHAPTER I

INTRODUCTION

A. Background and Problem

In analyses of data with statistics, the assumptions of statistics are the most necessary for parametric tests. For analysis of variance there are 5 assumptions. These are the underlying assumptions and the effects of violating assumptions (1).

1. Random sampling

The X values (replicates) within each class are assumed to be drawn randomly from their respective populations (or sub-populations).

The provided populations are infinite in size, the sampling can be either with or without replacements. If the populations are finite in size, the sampling ought to be with replacements for the ANOVA to be fully accurate. However, the accuracy of the ANOVA is affected very little by sampling without replacements and the provided populations are large.

2. Independence

The X values in each population (or sub-population) are assumed to be independent of each other. Slight autocorrelation can be ignored, but any major autocorrelation will seriously affect the accuracy of the ANOVA. Unfortunately, it is impossible to tell with a small sample whether major autocorrelation is present or not.

Copyright by Mahidol University

Although one can perform an ANOVA, one cannot determine how reliable the results are. The true value of P may be very different from the nominal value.

3. Existence of a common variance, σ^2

The variance of the X values in each class is assumed to be an estimate of a common variance σ^2. The analysis of variance, in other words, makes the basic assumption that the population variances are all equal. Small inequalities in the variances of the different classes are not serious. But if the individual variances differ markedly, it is unsafe to use analysis of variance.

Not all populations, however, have similar variances. It is unwise to assume equality of variance without careful inspection of the sample data. Often the data are limited, and one cannot be at all sure whether the assumption of a common variance is met. It is tempting in such cases to proceed with an ANOVA and run the risk that the population variances are unequal. The effect of inequality, however, is to reduce the variance estimate for error.

4. Normal distribution of X values in each class

The X values in each class are assumed to be part of a normal distribution. Any departures from normality are assumed to be caused by sampling error.

Kingman (2) told that in a real situation, various data are not normally distributed. If so, transformation is one of the method to help this. Transformations such as the log or square root are often used. For continuous data, but non-normally distributed outcomes, one can transform the response variable (change the measurement scale) by a continuous parametric or by specific non-parametric rank transformation. When a parametric transformation does not seem appropriate, one can use a distribution-free producer such as a rank transformation or a randomization test.

Copyright by Mahidol University

5. Additivity of the effects of Factor A and error

The effect of Factor A and error on the value of X are assumed to be additive. In other words, Factor A is presumed to act independently of the unexamined factors and chance causes that are collectively referred to as error. It is not supposed to interact in any way.

Winer (3) wrote about homogeneity of variances assumption like these. In many statistical procedures, tests for homogeneity of variances are often of interest as a preliminary to other analysis such as analysis of variances or a pool of data from different sources to yield an improved estimated variance.

When the true variance differs from one observation to another, there will, as a rule, be a loss of efficiency in the estimates of treatment effects. Similarly, there will be a loss of sensitivity in tests of significance. If the changes in the variance are large, these losses may be substantial (4). Moderate departures from this assumption do not, however, seriously affect the sampling distribution of the resulting statistic test which takes the differences in the population variances into account. Glass and Stanley (5) wrote that the effect of violation of homogeneous variances assumption can be serious depending upon sample sizes of two groups; n_1 and n_2. If n_1 and n_2 are equal, violation of the homogeneous variances assumption is unimportant and need not concern it. This fact is a compelling motive for selecting samples equal in size when positive in using the technique in this section. Whenever the variances of populations 1 and 2 are different and n_1 and n_2 are not equal, probabilities of type I and type II errors may be quite different from what one imagines them to be for example, the effects on the

Copyright by Mahidol University

level of significance of violation of the assumptions of normality and equality of variances of ANOVA are like the followings:

1. By giving a true null hypothesis, F-table shows that the 1-α percentile point in the F-distribution has J treatments with J-1 and N-J df ($_{1-\alpha}F_{J-1,N-J}$). (This percentile point will be the value exceeded by (100α)% of the F-ratios obtained in an ANOVA when the null hypothesis is true and the ANOVA assumptions are met.)

2. By empirical or mathematical means, the actual percentage of F-ratios exceeding $_{1-\alpha}F_{J-1,N-J}$ is found when the null hypothesis is true and the variances are heterogeneous or the populations are nonnormal or both.

3. The nominal significance level α, and the actual significance level, the percentage of F's exceeding $_{1-\alpha}F_{J-1,N-J}$, are compared.

The well-known statistical tests for homogeneity of variances in statistical analysis of a computer program are F-ratio, Bartlett's test, Levene's test, Scheffe's test, Hartley's test and Cochran's test. When these tests analyze the same data sources, their results are different. That is because of the assumptions of the tests.

Williams (1) wrote that the calculated value of F will tend to be too large, so significance will be attained more often than it should be. The test, in other words, will tend to be biased against the null hypothesis. It is known to be fairly sensitive to departures from normality. If the sample data depart greatly from normality, it is advisable to transform them, for instance, by taking logarithms or square roots, so that they are more nearly normal in distribution.

Copyright by Mahidol University

Glass and Stanley (5) stated that several research studies on the effect of heterogeneous variances have been executed. All of these corroborate in part or entirely the following conclusions:

1. When the sample sizes are equal, the effect of heterogeneous variances on the level of significance of the F-test is negligible.

2. When the sample sizes and variances are unequal and fewer persons are sampled from the populations with larger variances, the probability of a type I error is greater than α. In other words, the effect of heterogeneous variances in this case is to shift the distribution of F-ratios to the right.

3. When the sample sizes and variances are unequal and greater numbers of persons are sampled from the populations with larger variances, the probability of a type I error is less than α. The effect of heterogeneous variances in this case is to shift the distribution of F-ratios to the left.

According to Gartside (6), a variety of statistics are available for testing the homocedasticity of several populations. By comparing a number of these statistics for power and for stability of error rates, he found that Bartlett's statistic is very powerful in all the experimental cases.

Winer (3) said that Bartlett's test, from computational point of view, is more complex than either the Hartley's test or the Cochran's test. In Bartlett's test, the n_j's in each of the treatment classes need not be equal. However, no n_j's should be smaller than 3, and most n_j's should be larger than 5.

Copyright by Mahidol University

Gartside (6) found that Cochran's test performed very well in power for equal samples size of 16 k = 1 and k = 10, they were taken from populations whose variances increase in succession by a constant ratio, i.e., it showed good sensitivity when departures from the equal variances were still small. Moreover, if one variance is larger, then Cochran's statistic would be a good choice as it maintained good power in this case.

Winer (3) added that Cochran's test is generally what some more sensitive than the Hartley's test since it uses more of the information in the sample data.

According to Brown and Forsythe (7) alternative formulations of Levene's test statistic for equality of variance are found to be robust under nonnormality. These statistics use more robust estimators of central location in place of the mean. They are compared with the modified Levene's statistic, a jackknife prodecure, and a χ^2 test suggested by Layard (in Brown and Forsythe) which are all found to be less robust under nonnormality. Unfortunately, the common F-ratio and Bartlett's test are vary sensitive to the assumption that the underlying populations are from a Gaussian distribution. When the underlying distributions are nonnormal, these tests can have an actual size several times larger than their nominal level of significance.

Neter, Kutner, Nachtcheim, and Wasserman (8) stated that Levene's test does not depend on normality of the error terms. Indeed, this test is robust against serious departures from normality, in the sense that the nominal significance level remains

Copyright by Mahidol University

approximately correct when the error terms have equal variances even if the distribution of the error terms is far from normal.

Winer (3) wrote that Scheffe's test can be every bit as tedious arithmetically as Bartlett's. It is much less sensitive to violation of normality assumption and the number of subsamples p_j is arbitrary. As a rule of thumb each of the subsamples should be of approximately equal size; each subsample should be of size larger than three. Also, as a rule of thumb, $\Sigma(p_j-1)$ should be 10 or more to assure reasonable power for the test.

Hartley's test is a useful short-cut test of H_0 but does not use all of the information the observations provide about heterogeneity of variances, and it might be expected that more powerful test could be found. Winer (3) wrote that Hartley's test is in practice sufficiently sensitive for use as a preliminary test in situations where such a test is in order. When the number of observations in each of the treatment classes is not constant, but the n_j's are relatively close to being equal, the largest of the sample sizes may be used instead of n in obtaining the degrees of freedom required for use in the Hartley tables. This procedure leads to a slight positive bias in the test, i.e., rejecting H_1 more frequently than should be the case.

Vijitthummaroj (9) compared homogeneity of variances and means by using different programs: SPSS for Windows Version 6, Version 7.5, SPSS/PC+ Version 4.0 and SAS for Windows Version 6.12. He found that the results from SPSS for Windows Version 6 are the same as the results from SPSS for Windows Version 7.5

Copyright by Mahidol University

but they were different but using SPSS/PC+ Version 4.0 and SAS for Windows Version 6.12. These problems came from the different homogeneity of statistical tests and their assumptions; SPSS for Windows Version 6, Version 7.5 use Levene's test whereas SAS use Hartley's test.

Dixon (10) stated that the analysis of variance assumes that the responses / observations within each of the treatment groups are normally distributed with a common variance. In BMDP statistical software, Levene's test is used to identify unequal variances.

While, Gustafson (11) wrote that TRUE EPISTAT displays the ANOVA table, calculates Bartlett's test of equal variances and appropriates test probabilities for up to 12 samples. To check ANOVA assumption, which need the populations from which the samples were drawn have the same variance. If sample variances are found to be different by Bartlett's, then ANOVA results may be suspect. Some authors, however, believe that Bartlett's test is too sensitive.

Further more, Dean (12) said in EPIINFO version 6.0 manual, a word processing, database, and statistics system for epidemiology on microcomputers, use Bartlett's test to find out the homogeneity of variance assumption that the variances for the groups being compared are similar.

However, all statistical tests have different methods to test data and they have some different weak points. That is the main thing of them. For F-ratio, Scheff's test and Bartlett's test are sensitive to violation of normality assumption. Furthermore,

Copyright by Mahidol University

Cochran's test found to be more sensitive than Hartley's test since it uses all variances. Last, Levene's test seems to be good choice for checking homogenous variances is case of robust for normality assumption.

This research aims at checking homogeneity of variance. Since data distribution and the robustness of statistical test have some effects on outputs, and no one has study in this point. Therefore, this study is proposed to investigate the effects on statistical power of Bartlett's test, Levene's test and Cochran's test for comparing treatment groups in situations where it is unclear as to the correct distributional form of the response variables.

B. Objectives

The objectives of this study are:

1. To compare power of Bartlett's test, Levene's test and Cochran's test when distribution of data is normal, and the sample sizes and the sample variances are equal.
2. To compare power of Bartlett's test, Levene's test and Cochran's test when distribution of data is normal, the sample sizes are equal and the sample variances are unequal.
3. To compare power of Bartlett's test, Levene's test and Cochran's test when distribution of data is normal, the sample sizes are unequal and the sample variances are equal.

Copyright by Mahidol University

4. To compare power of Bartlett's test, Levene's test and Cochran's test when distribution of data is normal, and the sample sizes and the sample variances are unequal.
5. To compare power of Bartlett's test, Levene's test and Cochran's test when distribution of data is non-normal, and the sample sizes and the sample variances are equal.
6. To compare power of Bartlett's test, Levene's test and Cochran's test when distribution of data is non-normal, the sample sizes are equal and the sample variances are unequal.
7. To compare power of Bartlett's test, Levene's test and Cochran's test when distribution of data is non-normal, the sample sizes are unequal and the sample variances are equal.
8. To compare power of Bartlett's test, Levene's test and Cochran's test when distribution of data is non-normal, and the sample sizes and the sample variances are unequal.

C. Hypothesis

It is hypothesized that:

1. The unequality of sample sizes could result in the loss of power of Cochran's test more than Bartlett's test and Levene's test.

Copyright by Mahidol University

2. When one variance is larger, Cochran's test could be a good test in case of power.
3. The violation of normal distribution assumption could result in the loss of power of Bartlett's test.
4. The violation of normal distribution assumption could not result in the loss of power of Levene's test.

D. Scope of the Study

1. This study investigates only homogeneity variances assumption which uses Bartlett's test, Levene's test and Cochran's test.
2. In Bartlett's test, the researcher uses median to define the robustness and power of test. Because it provides good robustness even though many types of non-normal data. (13)
3. Only 3 groups of homogeneity test are used.
4. Each statistical test is simulated for 1,000 times. It always yields normal distribution, chi-square distribution and gamma distribution. According to the general rule of thumb (14) which the sample size choose greater than 30, so the proper sample sizes in this study assigned are 15, 30 and 45 for equal and unequal sizes. Also, the variance ratios are 1:1:1, 1:1:2 and 1:2:4. (15)

Copyright by Mahidol University

E. Definitions

Robustness: the ability of the test do not falsely detect non-homogeneous groups when the underlying data is not normally distributed and the groups are in fact homogeneous.

Power of a Test: Broadly, it is the ability of a statistical test to detect relationships. Specifically, it is the probability of rejectiong a null hypothesis when it is false and therefore should be rejected. The power of a test is calculated by subtracting the probability of a Type II Error from 1.0. The maximum total power a test can have is 1.0; the minimum is zero. It is also called "statistical power".

Type I error: It is an error made by wrongly rejecting a true null hypothesis. This might involve incorrectly concluding that two variables are related when they are not or wrongly deciding that a sample statistic exceeds the value that would be expected by chance. It is also called alpha error.

Type II error: It is an error made by wrongly accepting or failing to reject a false null hypothesis.

Distribution: A ranking, from lowest to highest, of the values of a variable and the resulting pattern of measures or scores, often as these are plotted on a graph.

Normal Distribution: A purely theoretical continuous probability distribution in which the horizontal axis represents all possible values of a variable and the vertical axis represents the probability of those values

Copyright by Mahidol University

occurring. The scores on the variable (often expressed as Z-scores) are clustered around the mean in a symmetrical, unimodal pattern known as the bell-shaped curve or normal curve (see following graph).

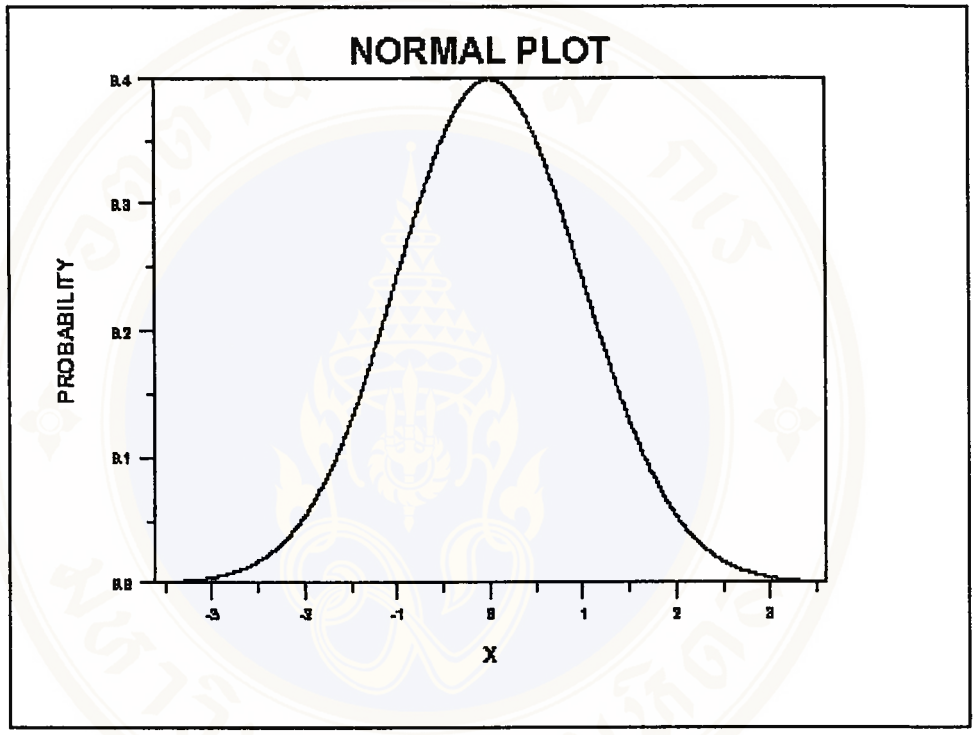


Figure 1. Normal Distribution

In a normal distribution, the mean, the median, and mode are all the same. There are many different normal distributions, one for every possible combination of mean and standard deviation. It is also sometimes called the "Gaussian distribution."

Because the sampling distribution of a statistic tends to be a normal distribution, the normal distribution is widely used in

Copyright by Mahidol University

statistical inference. For small samples, the Student's *t* distribution (which is also "bell shaped" but not "normal") is used. (16)

Chi-square Distribution: The chi-square distribution is a mathematical distribution that is used directly or indirectly in many tests of significance. The most common use of the chi square distribution is to test differences between proportions. Although this test is by no means the only test based on the chi square distribution, it has come to be known as the chi square test. The chi square distribution has one parameter, its degrees of freedom (df). It has a positive skew; the skew is less with more degrees of freedom. The mean of a chi square distribution is its degree of freesom. The mode is df-2 and the median is approximately df - 7. (17)

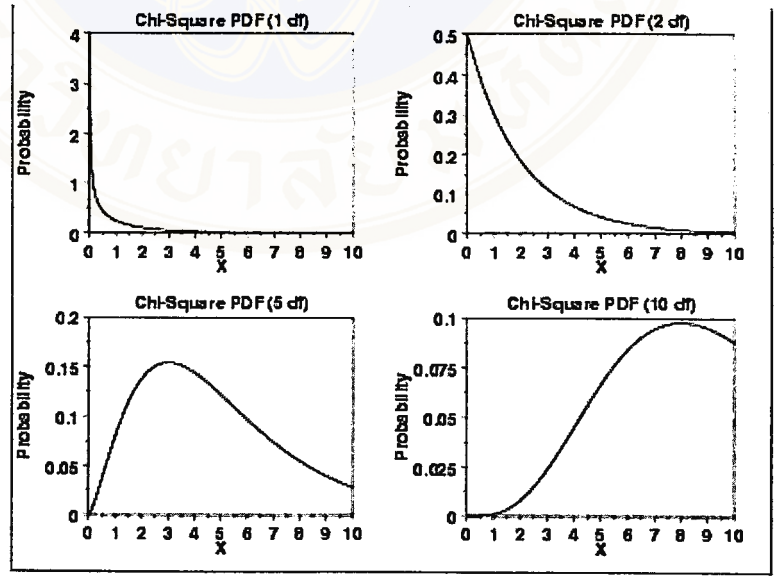


Figure 2. Chi-square Distribution

Gamma Distribution: A distribution used for continuous random variables which are constrained to be greater or equal to 0. It is characterized by

Copyright by Mahidol University

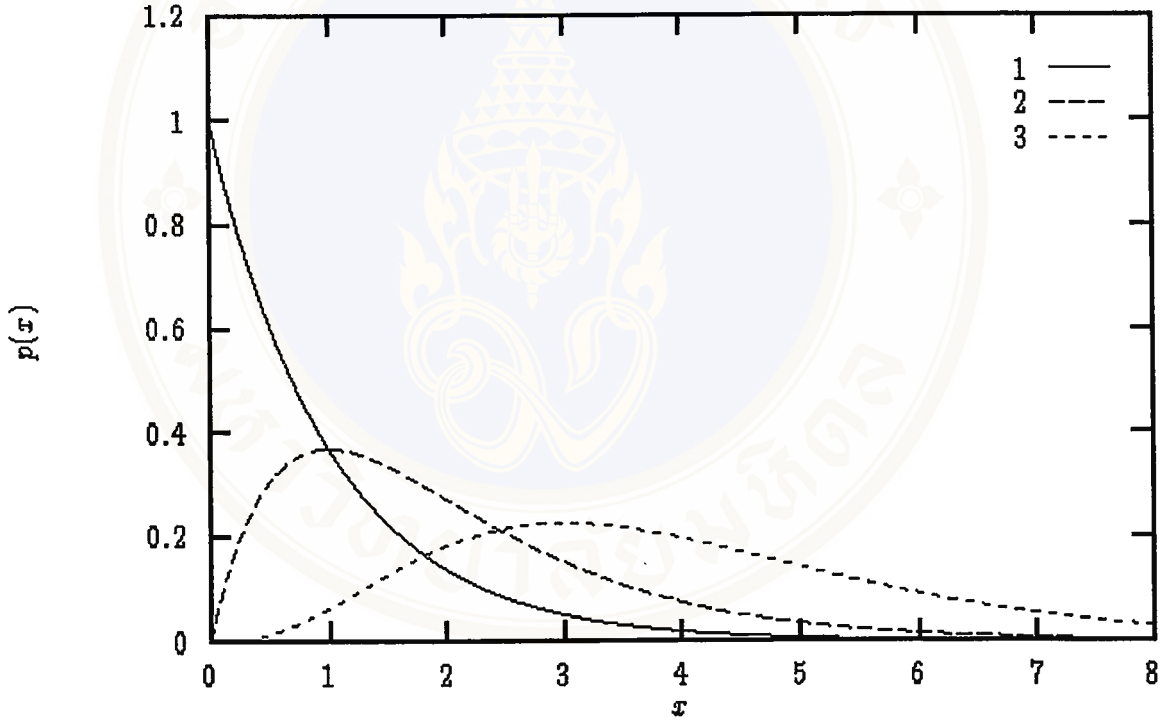
two parameters: shape and scale. The gamma distribution is often used to model data, which is positively skewed.

Parameters: shape $\alpha>0$, scale $\beta>0$

Domain: X>=0

Mean: $\alpha\beta$

Variance: $\alpha\beta^2$ (18)



$\beta = 1$

1: $\alpha = 0$, 2: $\alpha = 1$, 3: $\alpha = 3$.

Figure 3. Gamma Distribution

Bartlett's Test: The best know test of the homogeneity of a set of variances is due to M.S. Bartlett in1937. The test is difficult to calculations and so sensitive to the normal assumption.

Copyright by Mahidol University

Levene's Test: One statistical method for testing the assumption of homogenous variances which proposed by Levene in 1960. The test is simple to run (it uses ordinary analysis of variance techniques) and insensitive to violation of the normality assumption in most cases.

Cochran's Test: It is another relatively simple test for homogeneity of variance, developed by Cochran in 1951. This test is simple to run and that uses more of the information in the data than does Hartley's. (5)

Copyright by Mahidol University

CHAPTER II

LITERATURE REVIEW

This chapter discusses Bartlett's test, Levene's test and Cochran's test in terms of methods, definition and their limitations, and summary of other researchers' tests in the homogeneity of variance.

For many experimenters, the most commonly used statistical tests are those for comparing sample means and sample variances. The test used for the equality of the means of k groups of observations is usually the analysis of variances test (or equivalently the t test when $k = 2$). Meanwhile, the test used for the equality of variances of k groups of observations is Bartlett's (1937) modification of the Neyman-Pearson (1937) L_1 test. Cochran's test, the test introduced by Cochran (1941), is considerably easier to compute than the tests up to that time and Levene's test which Levene (1960) suggested using the one-way analysis of variance on the variables $Z_{ij} = \left| X_{ij} - \overline{X_i} \right|$ as a method of incorporating the robustness of that test into a test for variance.

A. Bartlett's Test

Bartlett's test is very useful and widely used to test if k samples have equal variances. However, it is very sensitive to departure from normality (19). The procedure involves computing a statistic whose sampling distribution is closely

Copyright by Mahidol University

approximated by the chi-square distribution with k-1 degrees of freedom when the k random samples are from independent normal populations. (20)

1. Definition

The Bartlett's test is defined as:

$$H_0 : \sigma_1^2 = \sigma_2^2 = ... = \sigma_k^2$$
$$H_1 : \text{above no true for at least one } \sigma_i^2$$

The Bartlett's test statistic is designed to test for equality of variances across groups against the alternative that variances are unequal for at least two groups. The test statistic is

$$\chi_0^2 = 2.3026\frac{q}{c}$$

Where

$$q = (N-k)\log_{10} S_p^2 - \sum_{i=1}^{k}(n_i - 1)\log_{10} S_i^2$$
$$c = 1 + \frac{1}{3(k-1)}\left(\sum_{i=1}^{k}(n_i - 1)^{-1} - (N-k)^{-1}\right)$$
$$S_p^2 = \frac{\sum_{i=1}^{k}(n_i - 1)S_i^2}{N-k}$$

which n_i = sample size of the ith group

s_i^2 = sample variance of the ith group

N = total sample sizes

k = # of groups

Copyright by Mahidol University

s_p^2 = pooled variance (weighting by degrees of freedoms)

The quantity q is large when the sample variances s_i^2 differ greatly and is equal to zero when all s_i^2 are equal. Therefore, we should reject H_0 on values of χ_0^2 that are too large; that is, we reject H_0 only when

$$\chi_0^2 > \chi_{1-\alpha,k-1}^2$$

Where $\chi_{1-\alpha,k-1}^2$ is the upper critical value of the chi-square distribution with k-1 degrees of freedom and a significance level of α. The P-value approach to decision making could also be used.

In table 1, this test is applied to the variances of the birth weights of five litters of pigs.

Table 1. Computation of Bartlett's Test of Homogeneity of Variance

Litter (Sample)	Mean Squares s_i^2	Degree of Freedom (n_i-1)	Sum of Squares $(n_i-1)\ s_i^2$	Log s_i^2	$(n_i-1)\log s_i^2$	Reciprocals $(n_i-1)^{-1}$
1	0.909	9	8.18	-0.0414	-0.3726	0.1111
2	0.497	7	3.48	-0.3036	-2.1252	0.1429
3	0.076	9	0.68	-1.1192	-10.0728	0.1111
4	0.103	7	0.72	-0.9872	-6.9104	0.1429
5	0.146	5	0.73	-0.8357	-4.1785	0.2000
K = 5		37	13.79		-23.6595	0.7080

Copyright by Mahidol University

$$S_p^2 = \sum_{i=1}^{k}(n_i - 1)S_i^2 \Big/ N - k = 13.79/37 = 0.3727$$

$$(N-k)\log_{10}S_p^2 = (37)\log_{10}(0.3727) = -15.8582$$

$$q = 2.3026[(N-k)\log_{10}S_p^2 - \sum_{i=1}^{k}(n_i - 1)\log_{10}S_i^2]$$

$$= (2.3026)[-15.8582 - (-23.6595)] = 17.96$$

$$c = 1 + \frac{1}{3(4)}\left(0.7080 - \frac{1}{37}\right) = 0.1057$$

$$\chi_0^2 = q/c = 17.96/1.057 = 16.99$$

$$d.f. = 4$$

$$\chi_{.95,4}^2 = 9.488$$

$$\chi_0^2 > 9.488$$

That is at least one variance of the birth weights of five litters of pigs are unequal.(21)

2. Limitation

Various studies have indicated that Bartlett's test is very sensitive to the normality assumption and should not be applied when the normality assumption is doubtful. When the underlying distributions are nonnormal, these test can have an actual size several times larger than their nominal level of significance (Brown and Forsythe). That is, if your samples come from non-normal distributions, then Bartlett's test may simply be testing for non-normality.

Moreover, Armitage (22) told that Bartlett's test is perhaps less useful than might be thought, for two reasons. First, like the *F* test, it is rather sensitive to non-normality. Secondly, with samples of moderate size, the true variances σ_i^2 have to differ very considerably before there is a reasonable chance of obtaining a significant

Copyright by Mahidol University

test result. To put this point another way, even if χ_0^2 is nonsignificant, the estimated s_i^2 may differ substantially, and so may the true σ_i^2. If inequality in the σ_i^2 is important, it may therefore be wise to assume it even if the test result is nonsignificant. In some situations, moderate inequality in the σ_i^2 will be no matter very much, so again the significance test is not very relevant.

Box (23) criticized reliance on Bartlett's test as an indicator of whether inference from an analysis of variance can be trusted. He found that the test is sensitive to non-normality.

B. Levene's test

Levene's test (13) is used to test if k samples have equal variances. Some statistical tests, for example, the analysis of variance, assume that variances are equal across groups or samples. The Levene's test can be used to verify that assumption.

1. Definition

The Levene's test is defined as:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2$$
$$H_1 : \text{above no true for at least one } \sigma_i^2$$

Given a variable Y with sample of size N divided into k sub-groups, where N_i is the sample size of the ith sub-group, the Levene's test statistic is defined as:

Copyright by Mahidol University

$$W = \frac{(N-k)\sum_{i=1}^{k} N_i \left(\bar{Z}_{i.}-\bar{Z}_{..}\right)^2}{(k-1)\sum_{i=1}^{k}\sum_{j=1}^{N}\left(Z_{ij}-\bar{Z}_{i.}\right)^2}$$

where Z_{ij} can have one of the following three definitions:

1. $Z_{ij} = |Y_{ij} - \bar{Y}_{i.}|$

   where $\bar{Y}_{i.}$ is the *mean* of the ith subgroup.

2. $Z_{ij} = |Y_{ij} - \tilde{Y}_{i.}|$

   where $\tilde{Y}_{i.}$ is the *median* of the ith subgroup.

   The median is a measure of central tendency not sensitive to outlying values -- unlike the mean, which can be affected by a few extremely high or low values.

3. $Z_{ij} = |Y_{ij} - \bar{Y}'_{i.}|$

   where $\bar{Y}'_{i.}$ is the 10% *trimmed mean* of the ith subgroup.

   Where 10% trimmed mean is the arithmetic mean calculated when the largest 10% and the smallest 10% of the cases have been eliminated. Eliminating extreme cases from the computation of the mean results in a better estimate of central tendency, especially when the data are nonnormal (24). Computation of trimmed mean is illustrated in example 1.

Copyright by Mahidol University

Example 1. Computation of Trimmed Mean

Table 2. Scores of Anxiety of Subjects in Trials (24)

Score	Frequency
1	2
2	3
4	4
5	1
6	4
8	7
9	1
10	6
12	6
14	3
16	6
18	3
19	2
Total	48

The data from Table 2 can be computed the 10% trimmed mean as follows:

10 % = 0.1,

n = sample size = 48,

eliminate = 48 × 0.1 = 4.8,

eliminate score = 1 which has 2 ,

and 19 which has 2.

So, 10% Trimmed Mean of 48 cases = mean of 44 cases = 10.

$\bar{Z}_{i\cdot}$ are the group means of the Z_{ij} and $\bar{Z}_{\cdot\cdot}$ is the overall mean of the Z_{ij}.

The three choices for defining Z_{ij} determine the robustness and power of Levene's test.

In Critical Region, the Levene's test rejects the hypothesis that the variances are homogeneous if

$$W > F_{(1-\alpha,k-1,N-k)}$$

Copyright by Mahidol University

where $F_{1-\alpha, k-1, N-k}$ is the upper critical value of the F distribution with k - 1 and N - k degrees of freedom at a significance level of α. The details of computation of different Z_{ij} are show in example 2, 3 and 4.

Example 2. Computation of Z_{ij} from the Mean of the ith Subgroup

Table 3. Anxiety Data from SPSS by Mean (24)

Trial	Yij	Zij	$Z_{ij} - \bar{Z}_{.}$	Trial	Yij	Zij	$Z_{ij} - \bar{Z}_{.}$
1	12	4.5	2.916667	3	4	3.75	1.791667
1	14	2.5	0.91667	3	5	2.75	0.791667
1	16	0.5	-1.08333	3	6	1.75	-0.20833
1	16	0.5	-1.08333	3	6	1.75	-0.20833
1	16	0.5	-1.08333	3	6	1.75	-0.20833
1	16	0.5	-1.08333	3	8	0.25	-1.70833
1	16	0.5	-1.08333	3	8	0.25	-1.70833
1	18	1.5	-0.08333	3	8	0.25	-1.70833
1	18	1.5	-0.08333	3	10	2.25	0.291667
1	18	1.5	-0.08333	3	10	2.25	0.291667
1	19	2.5	0.91667	3	10	2.25	0.291667
1	19	2.5	0.91667	3	12	4.25	2.291667
	$\bar{Y}_{1.}$ = 16.5	$\bar{Z}_{1.}$ = 1.58			$\bar{Y}_{3.}$ = 7.75	$\bar{Z}_{3.}$ = 1.96	
2	8	3.5	1.583333	4	1	3.25	0.916667
2	8	3.5	1.583333	4	1	3.25	0.916667
2	10	1.5	-0.41667	4	2	2.25	-0.08333
2	10	1.5	-0.41667	4	2	2.25	-0.08333
2	10	1.5	-0.41667	4	2	2.25	-0.08333
2	12	0.5	-1.41667	4	4	0.25	-2.08333
2	12	0.5	-1.41667	4	4	0.25	-2.08333

Copyright by Mahidol University

Table 3. (cont.)

Trial	Yij	Zij	$Z_{ij} - \bar{Z}_{.}$	Trial	Yij	Zij	$Z_{ij} - \bar{Z}_{.}$
2	12	0.5	-1.41667	4	4	0.25	-2.08333
2	12	0.5	-1.41667	4	6	1.75	-0.58333
2	14	2.5	0.583333	4	8	3.75	1.416667
2	14	2.5	0.583333	4	8	3.75	1.416667
2	16	4.5	2.583333	4	9	4.75	2.416667
	$\bar{Y}_{2.}$ = 11.5	$\bar{Z}_{2.}$ = 1.92			$\bar{Y}_{4.}$ = 4.25	$\bar{Z}_{4.}$ = 2.33	
						Total	80.97915

Trial	1	2	3	4	Total
$\bar{Z}_{.}$	1.58	1.92	1.96	2.33	
$\bar{Z}_{..}$			1.95		
$\bar{Z}_{.} - \bar{Z}_{..}$	-0.36458	-0.03125	0.010416	0.385416	0.282552
W	= 44 x 12 x 0.282552 / (3 x 80.97915)				= 0.61

Example 3. Computation of Z_{ij} from the Median of the ith Subgroup

Table 4. Anxiety Data from SPSS by Median (24)

Trial	Yij	Zij	$Z_{ij} - \bar{Z}_{.}$	Trial	Yij	Zij	$Z_{ij} - \bar{Z}_{.}$
1	12	4	2.5	3	4	4	2.083333
1	14	2	0.5	3	5	3	1.083333
1	16	0	-1.5	3	6	2	0.083333
1	16	0	-1.5	3	6	2	0.083333
1	16	0	-1.5	3	6	2	0.083333
1	16	0	-1.5	3	8	0	-1.91667
1	16	0	-1.5	3	8	0	-1.91667
1	18	2	0.5	3	8	0	-1.91667

Copyright by Mahidol University

Table 4. (cont.)

Trial	Yij	Zij	$Z_{ij} - \bar{Z}_{.}$	Trial	Yij	Zij	$Z_{ij} - \bar{Z}_{.}$
1	18	2	0.5	3	10	2	0.083333
1	18	2	0.5	3	10	2	0.083333
1	19	3	1.5	3	10	2	0.083333
1	19	3	1.5	3	12	4	2.083333
	$\bar{Y}_{1.} = 16$	$\bar{Z}_{1.} = 1.5$			$\bar{Y}_{3.} = 8$	$\bar{Z}_{3.} = 1.92$	
2	8	4	2.166667	4	1	3	0.75
2	8	4	2.166667	4	1	3	0.75
2	10	2	0.166667	4	2	2	-0.25
2	10	2	0.166667	4	2	2	-0.25
2	10	2	0.166667	4	2	2	-0.25
2	12	0	-1.83333	4	4	0	-2.25
2	12	0	-1.83333	4	4	0	-2.25
2	12	0	-1.83333	4	4	0	-2.25
2	12	0	-1.83333	4	6	2	-0.25
2	14	2	0.166667	4	8	4	1.75
2	14	2	0.166667	4	8	4	1.75
2	16	4	2.166667	4	9	5	2.75
	$\bar{Y}_{2.} = 12$	$\bar{Z}_{2.} = 1.83$			$\bar{Y}_{4.} = 4$	$\bar{Z}_{4.} = 2.25$	
						Total	101.8333

Trial	1	2	3	4	Total
$\bar{Z}_{.}$	1.5	1.83	1.92	2.25	
$\bar{Z}_{..}$			1.875		
$\bar{Z}_{.} - \bar{Z}_{..}$	-0.375	-0.04167	0.041667	0.375	0.284722
W	= 44 x 12 x 0.284722 / (3 x 101.8333)				= 0.49209

Copyright by Mahidol University

Example 4. Computation of Z_{ij} from the 10% Trimmed Mean of the ith Subgroup

Table 5. Anxiety Data from SPSS by 10% Trimmed Mean (24)

Trial	Yij	Zij	$Z_{ij} - \bar{Z}_{.}$	Trial	Yij	Zij	$Z_{ij} - \bar{Z}_{.}$
1	12	4.5	2.916667	3	4	3.75	1.791667
1	14	2.5	0.91667	3	5	2.75	0.791667
1	16	0.5	-1.08333	3	6	1.75	-0.20833
1	16	0.5	-1.08333	3	6	1.75	-0.20833
1	16	0.5	-1.08333	3	6	1.75	-0.20833
1	16	0.5	-1.08333	3	8	0.25	-1.70833
1	16	0.5	-1.08333	3	8	0.25	-1.70833
1	18	1.5	-0.08333	3	8	0.25	-1.70833
1	18	1.5	-0.08333	3	10	2.25	0.291667
1	18	1.5	-0.08333	3	10	2.25	0.291667
1	19	2.5	0.91667	3	10	2.25	0.291667
1	19	2.5	0.91667	3	12	4.25	2.291667
	$\bar{Y}_{1.} = 16.5$	$\bar{Z}_{1.} = 1.58$			$\bar{Y}_{3.} = 7.75$	$\bar{Z}_{3.} = 1.96$	
2	8	3.5	1.583333	4	1	3.25	0.916667
2	8	3.5	1.583333	4	1	3.25	0.916667
2	10	1.5	-0.41667	4	2	2.25	-0.08333
2	10	1.5	-0.41667	4	2	2.25	-0.08333
2	10	1.5	-0.41667	4	2	2.25	-0.08333
2	12	0.5	-1.41667	4	4	0.25	-2.08333
2	12	0.5	-1.41667	4	4	0.25	-2.08333
2	12	0.5	-1.41667	4	4	0.25	-2.08333
2	12	0.5	-1.41667	4	6	1.75	-0.58333
2	14	2.5	0.583333	4	8	3.75	1.416667
2	14	2.5	0.583333	4	8	3.75	1.416667
2	16	4.5	2.583333	4	9	4.75	2.416667
	$\bar{Y}_{2.} = 11.5$	$\bar{Z}_{2.} = 1.92$			$\bar{Y}_{4.} = 4.25$	$\bar{Z}_{4.} = 2.33$	
						Total	80.97915

Copyright by Mahidol University

Trial	1	2	3	4	Total
$\bar{Z}_{.}$	1.58	1.92	1.96	2.33	
$\bar{Z}_{..}$			1.95		
$\bar{Z}_{.}-\bar{Z}_{..}$	-0.36458	-0.03125	0.010416	0.385416	0.282552
W	= 44 x 12 x 0.282552 / (3 x 80.97915)				= 0.61

From the data is Table 3, 4 and 5

$$\alpha = 0.5$$

$$N = 48$$

$$k = 4$$

$$F_{1-\alpha, k-1, N-k} = F_{.95,3,44} \approx 2.84$$

$$W < F_{1-\alpha, k-1, N-k}$$

Therefore, accept hypothesis that Anxiety subjects of four trials have equal variances.

Levene's test is an alternative to the Bartlett's test. The Levene's test is less sensitive than the Bartlett's test to departures from normality. (25)

Unlike Brown and Forsythe, Conover, Johnson and Johnson (15) consider use of median instead of both trimmed mean and mean because their results indicated no advantages in using this variation.

C. Cochran's Test

Cochran's test is the one of homogeneity of variances test. It is computationally simpler than Bartlett's test and also affected by non-normality. (25)

Copyright by Mahidol University

1. Definition

The Cochran's test is defined as:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2$$
$$H_1 : \text{above no true for at least one } \sigma_i^2$$

Keith (26) uses the statisic

$$C = \frac{\text{largest } S_i^2}{\sum S_i^2}$$

Where n is number of replicates for each mean

if n's are unequal, use largest or harmonic mean of n's.

s_i^2 = sample variance of the ith group

k = number of groups

$df = n - 1$

Then, we can use Cochran's C table that has 95th and 99th percentile points of the distribution of C statistic to define the critical value and reject the null hypothesis if

$$C > C_{1-\alpha,k,n-1}$$

where $C_{1-\alpha,k,n-1}$ is the upper critical value of the C distribution with k and n - 1 degrees of freedom at a significance level of α. The computation of Cochran's test is demonstrated in example 5.

Copyright by Mahidol University

Examples 5. Computation of Cochran's Test

Table 6. Example from Maass (27)

Group	1	2	3	4	5
n	6	6	6	6	6
S^2	12134	2303	3594	3319	3455

From the data in Table 6,

$$C = \frac{12134}{12134 + 2303 + 3594 + 3319 + 3455}$$

$$= 0.4892$$

$$C_{.95,5,5} = 0.5065$$

Since C < critical value, variances are not significantly different.

Johnson (28) suggested one program using the Mandel-Paule algorithm to find a 100(1 - α)% Confidence Interval for the common mean, μ, for repeated measurements made by p laboratories. It also tests for the homogeneity of variances for the p laboratories. The tests carried out are:

1) Bartlett's Test, 2) Cochran's Test, 3) Hartley's Test, and 4) Levene's Test. He found that Cochran's test and Hartley's test should only be considered if all of the n(i) are equal.

In 1981, Conover, Johnson and Johnson (15) examined fifty-six tests of equality of variances, most of which are variations of the most popular and most useful parametric and nonparametric tests available for testing the equality of k variances ($k \geq 2$) in the presence of unknown means. Comparisons are made under the

Copyright by Mahidol University

null hypothesis (for robustness) and under the alternative (for power). Monte Carlo simulated various symmetric and asymmetric distributions for four sample groups with respective sample sizes (n_1, n_2, n_3, n_4) = (5, 5, 5, 5), (10, 10, 10, 10), (20, 20, 20, 20), and (5, 5, 20, 20). The null hypothesis of equal variances (all equal to 1) was examined along with the four alternatives ($\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2$) = (1, 1, 1, 2), (1, 1, 1, 4), (1, 1, 1, 8), and (1, 2, 4, 8). The mean was set equal to the standard deviation in each population under the alternative hypothesis. Zero means were used for H_0. Each of these 30 combinations of distribution type, sample size, and variances was repeated 1,000 times. These tests are further compared using data from outer continental shelf bidding on oil and gas leases. After extensive simulation involving different distributions, sample sizes, means, and variances. Three tests appear to be superior selections in terms of robustness and power. These are Lev1: med (Levene: median), F-K: med χ^2 (Fligner and Killeen: median χ^2), and F-K: med F (Fligner and Killeen: median F).

After combining test information over 40 different sales (40 applications of each multisample test for variance), the three tests indicated a good agreement with the null hypothesis of equal variances. The Bar2: med (Bartlett is divided by (1+γ/2) where $\gamma = E(X_{ij} - \mu_i)^4 / \sigma_i^4 - 3)$ and Lev2: med (Levene involve $Z_{ij}^{1/2}$) showed some good properties were applied to oil and gas lease bidding data to see if the logs of the bids exhibited homogeneity of variance from lease to lease with a sale. Also, it seems reasonable to recommend Lev1: med, F-K: med χ^2, and F-K: med F as robust and powerful tests for variances when the populations means are unknown.

Copyright by Mahidol University

Nevertheless, Levy (29) tests four multiple range tests for specifying homogeneous subsets of treatment variances. The procedures are based on Hartley's *F*-max statistic, Cochran's statistic, Bartlett and Kendall's normalizing log transformation of the sample variances. These tests were simulated by drawing independent random samples of sized 5, 11 or 21 observations per group from three *N* (0,1) populations. These samples were generated by CDC library subroutine GAUSS which computes the sum of 12 uniform random numbers and adjusts this sum for specified mean and variance. While Cochran's statistic got equal sample sizes of 5, 10, or 17 observations per group, Box's approximate test got equal sample sizes of 4, 10, or 18 observations per group. At each stage of a given procedure, .05-level tests were conducted. Empirical significance levels are obtained by simulating independent random samples from normal, uniform, double exponential and Cauchy distribution. A procedure was replicated 1,000 times for each condition. The conclusion of three of the proposed multiple range tests are also seriously affected by violations of the assumption of normality. The *F*-max, Cochran's and ln (s^2) tests seem to perform similarly to one another when sampling from normal, uniform, double exponential and Cauchy distributions. The tests most preferred would probably be the one most easily performed-the *F*-max test. If the assumption of normality is at all in doubt, however, then the Box procedure is strongly recommended because of its robustness; it performed quite well under all three conditions of nonnormality in the present study.

Dixon (10) said that the Levene's test should be used with caution when sample sizes are small. Brown and Forsythe (7) compared three unmodified Levene-type statistics, a jackknife procedure, and a χ^2 test suggested by Layard (in Brown and

Copyright© by Mahidol University

LIBRARY AND INFORMATION CENTER MAHIDOL UNIVERSITY

Forsythe). This is performed by generating pseudorandom numbers from four underlying distributions: the Gaussian by the Box-Muller method; the chi-square distribution with four df as proportional to the logarithm of the product of two uniform random numbers; the Student-*t* with four df as proportional to the ratio of a Gaussian to the square root of a chi-square with four df; and the Cauchy as the ratio of two Gaussians. Pairs of groups of sample sizes (40,40), (20,40), (10,10), and (10,20) were selected from the set of 80 observations and used in the calculation of the test statistics. The observations in the second sample were rescaled to reflect the ratios of population variances of the two groups; i.e., 1:1, 1:2,1:4, 2:1, and 4:1. These were repeated 1,000 times for each distribution in ten blocks of 100 trials. The numbers of rejections at both the nominal five-percent and one-percent levels of significance were recorded. The result is when there are equal variances in the two groups, the results for the Gaussian distribution indicated that W_{50} is conservative for small sample sizes. The results for the other test statistics are not inconsistent with the sampling error. The powers of these statistics do not differ greatly when the differences in their empirical size are taken into account.

For unequal groups, the size of the jackknife statistic is larger than it should be. This may be due to the lack of robustness of the ANOVA when the within group variances are unequal. The Layard χ^2 deviates from its level of significance for the smaller sample sizes. Since an estimate of kurtosis is used in the calculation of the statistic. Of the three Levene-type procedures, W_{10} appears the most robust. It varies in size less than W_0 and is nearer to 5 percent than W_{50}. Sampling from the Cauchy distribution shows that the rejection rates were excessive for all except W_{10} and W_{50}, but W_{10} was closer to five percent than W_{50}. In case of the result of sampling from the

Copyright© by Mahidol University

i 13579174

chi square with four df, only W_{50} maintains its size near the five-percent level of significance. The others depart from their nominal sizes by rejecting far too often.

Gartside (6) studied Bartlett's statistic, Bartlett's statistic modified by subsitituting, Cochran's statistic, Hartley's statistic, Cadwell's statistic for testing the homocedasticity of several populations. These statistics are compared for power and for stability of error rates in 4 cases under normally (0, 1) distributed unless otherwise stated, were:

Case 1. Equal samples of size 16 were taken from populations whose variances increase succession by a constant ratio:

$$H_0 = \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2 = 1$$
$$H_A = \sigma_1^2 = \sigma_2^2 = c, \cdots = \sigma_k^2 = c^{k-1}$$

Case 2. Equal samples of size 16 were taken from populations whose variances are all equal to 1 except for one population whose variance is larger:

$$H_0 = \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2 = 1$$
$$H_A = \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_{k-1}^2 = 1, \sigma_k^2 = c$$

Case 3. Unequally sized samples were taken from populations which all but one have a variance of 1. The remaining population has a larger variance and always involved the small sample size:

$$H_0 = \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2 = 1$$
$$H_A = \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_{k-1}^2 = 1, \sigma_k^2 = c$$

Case 4. Equal samples of size (i) 4, (ii) 8, (iii) 16 were taken, with power not being considered here, from populations that have Weibull distribution:

$$F(X) = 1 - \exp(-X^{4/3})$$

The values of c which were used in the study were shown in Table 7.

Copyright by Mahidol University

Table 7. The Values of C

Case	*K*			
	3	*4*	*5*	*10*
1	1.1792	1.0949	1.0594	1.0023
2	1.2053	1.1792	1.1759	1.1727
3	1.6395	1.5690	1.5349	

In conclusion, when the alternative hypothesis is unknown and normality can be relied on, Bartlett's test would be recommended. On the other hand, if it is suspected that just one variances is larger, then Cochran's statistic would be a good choice as it maintain good power in equal samples of size 16 taken from populations whose variances increase in succession by a constant ratio.

Scheffe (in Glass and Stanley)(5) uses samples 9, 5, 1 and the population variances in the ratio 1:1:3 to test the effect of heterogeneous variances on the probability of a type I error in the one-way ANOVA for a nominal significance level of 0.5. He found that the probability of type I error is actually 0.17.

Moreover, Winer (3) took examples from Box (in Winer) for testing the effect of lack of homogeneity of variance by using samples size 10, and k=4. He found that when the number of observations in each treatment class is not constant, but relatively close to being equals, the largest of sample sizes n_j's may be used in place of n in obtaining the degrees of freedom required for use in the Hartley tables.

In brief, when data distribution is Normal, Bartlett's test is recommended. Similarly, Levene's test is good for non-normal distribution. In addition, when there

Copyright by Mahidol University

are equal variances in 2 groups that are normal distributed, Levene's test is suitable for small sample size. On the other hand, if one variances is larger, then Cochran's test would be good to maintain power in equal samples of size 16. Finally, the first step in checking assumption of homogeneity of variance should be to inspect the standard deviations or variances within each level.

From the examination of Conover, Johnson and Johnson (15) on fifty six tests of equality of variances in various symmetric and asymmetric distributions for four sample groups with respective sample sizes (n_1, n_2, n_3, n_4) = (5, 5, 5, 5), (10, 10, 10, 10), (20, 20, 20, 20) and (5, 5, 20, 20), equal variances (all equal to 1) with four alternative ($\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2$) = (1, 1, 1, 2), (1, 1, 1, 4), (1, 1, 1, 8), and (1, 2, 4, 8), Lev1: med, F-K: med χ^2, and F-K: med F are robust and powerful tests for variances when the populations means are unknown. Furthermore, Levene's test should be used with caution when sample sizes are small (10). Moreover, Bartlett's test would be recommended when normality can be relied on. On the other hand, if one variances is larger, Cochran's test would be a good choice in equal sample size 16. (6)

This research investigates the different sample sizes and use gamma distribution which have not seen in any research and it is the special case of the chi-square distribution (30). So, it is more useful in choosing the best test of homogeneity of variances which matches the data. We also have to be careful with the test to be used in statistical software.

Copyright by Mahidol University

CHAPTER III

METHODOLOGY

This chapter describes the methods and distributions of Bartlett's test, Levene's test and Cochran's test from Monte Carlo simulations. The Monte Carlo method was performed using Fortran programs. (31)

A. Methodology

There are 2 criteria to detect alternative of appropriate statistics under violation of assumptions, robustness (the capability to control Type I error) and power of the test. Therefore, it had to generate lots of samples for their distributions, means, variances, and sample sizes followed specified condition. Since Bartlett's test and Levene's test depend on normal distribution, normal distribution was chosen to be studied. Moreover, since Cochran's test depends on chi-square distribution, the chi-square distribution, and gamma distribution (which is a special case of chi-square distribution) were chosen for testing how robust the three statistical tests are when the distribution was non-normal. This research tried to find the suitable sample size for the tests by using small sample size to show the robustness of tests. The researcher used the general rule of thumb that defines the sample size of 30 (14). Therefore, the samples sizes in this study for three groups were less than 30, equal to 30 and greater than 30 these were 15, 30, and 45. In testing homogeneity of variances of the three

Copyright by Mahidol University

statistical tests, theoretical alpha was defined as 0.01 and 0.05. In each distribution type, the sample size and variances were repeated 1,000 times.

The Monte Carlo method was used as the random number generator. In all cases 1000 randomizations are performed to estimate the p-values for each sample of the randomization test procedure. (2)

B. The Process of Computations

The process of calculating probability of type I error and power of three statistics tests under different settings of 3 groups of populations distributions was as follow:

1. Seed of 65539

Araya (32) used the number 65539 to generate a random number for statistics test. It will make many set of random number with uniform distribution. In Fortran program, the researcher used subroutine UMACH (N, NUNIT) and RNSET (ISEED) (IMSL math library) for generating the random number.

2. Population Distribution from Defining Distribution

Fortran program was used to create population under three distributions by simulating the data from Monte Carlo method. Subroutine RNNOA

Copyright by Mahidol University

(NR, R) single, subroutine SSCAL (NR, XSTD, R, 1), and subroutine SADD (NR, XM, R, 1) were used for generating normal distribution. To generate gamma distribution, subroutine RNGAM (NR, A, R) and SSCAL (NR, B, R, 1) were used. Lastly, subroutine RNCHI (NR, DF, R) was used for generating chi-square distribution. The three distributions were as follows:

2.1. Normal Distribution

Normal distribution or Gaussian distribution had probability density function of x, which was defined by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad , \qquad - \; < x <$$

Which had two parameters, mean(μ) = 10 and variance (σ^2) = 26. (33)

2.2. Gamma Distribution

Gamma distribution had special importance in probability statistics. The formulation used here was represented as

$$f(x, \alpha, \beta) = \frac{1}{\beta}^{\alpha} \frac{x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha)}$$

for x > 0. It has 2 parameters, mean and variance of this distribution. They were

$$\mu = \alpha\beta \quad \text{and} \quad \sigma^2 = \alpha\beta^2 \qquad (2)$$

Which $\alpha = 2.67$

$\beta = 0.75$

Copyright by Mahidol University

2.3. Chi-square Distribution

Chi-square distribution with degree of freedom n was a special case of the gamma distribution where p = n/2 and $\alpha = 2$. The formulation was represented as (33)

$$f(x) = \frac{x^{\frac{n}{2}-1}e^{-x/2}}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} \quad , \quad x > 0$$

$$= 0 \quad , \quad x \leq 0$$

3. Generate Groups of Populations

3.1. Generate Three Groups of Populations

They had the same distribution as follows:

N N N

G G G

C C C

3.2. Generate Three Groups of Distributions

They had the different distribution as follows:

N G C

Where N = Normal distribution

C = Chisquare distribution

G = Gamma distribution

3.3. Define Sample Size

Four settings of sample size were as follows: (14)

Copyright by Mahidol University

	Setting	Sample sizes
1.	Equal	15,15,15
2.	Equal	30,30,30
3.	Equal	45,45,45
4.	Unequal	15,30,45

4. Define Proportion of Sample Variances

The sample variances of three populations were determined as follow: (23)

Distributions	Sample sizes	Variance ratios
N N N	15,15,15	1: 1: 1
		1: 1: 2
		1: 2: 4
	30,30,30	1: 1: 1
		1: 1: 2
		1: 2: 4
	45,45,45	1: 1: 1
		1: 1: 2
		1: 2: 4
	15,30,45	1: 1: 1
		1: 1: 2

Copyright by Mahidol University

Distributions	Sample sizes	Variance ratios
		1: 2: 4
G:G:G	15,15,15	1: 1: 1
		1: 1: 2
		1: 2: 4
	30,30,30	1: 1: 1
		1: 1: 2
		1: 2: 4
	45,45,45	1: 1: 1
		1: 1: 2
		1: 2: 4
	15,30,45	1: 1: 1
		1: 1: 2
		1: 2: 4
C:C:C	15,15,15	1: 1: 1
		1: 1: 2
		1: 2: 4
	30,30,30	1: 1: 1
		1: 1: 2
		1: 2: 4
	45,45,45	1: 1: 1
		1: 1: 2
		1: 2: 4

Copyright by Mahidol University

Distributions	Sample sizes	Variance ratios
	15,30,45	1: 1: 1
		1: 1: 2
		1: 2: 4
NCG	15,15,15	1: 1: 1
		1: 1: 2
		1: 2: 4
	30,30,30	1: 1: 1
		1: 1: 2
		1: 2: 4
	45,45,45	1: 1: 1
		1: 1: 2
		1: 2: 4

5. Compute the Value of Three Statistics Tests

After data was generated in one situation, it was computed for Bartlett's test value, Levene's test value and Cochran's test value. Then these values were compared with their critical region. This was done for a thousand times. And the values that reject null hypothesis were counted.

After that situation was changed to generate new data and the three statistical tests values were computed. Then, the above steps were repeated for all situations.

Copyright by Mahidol University

6. Compute Probability of Type I Error and Power of a Test

In sum, the steps of computation of probability of Type I error and power of each statistical tests were shown in the diagram in Figure 4.

Copyright by Mahidol University

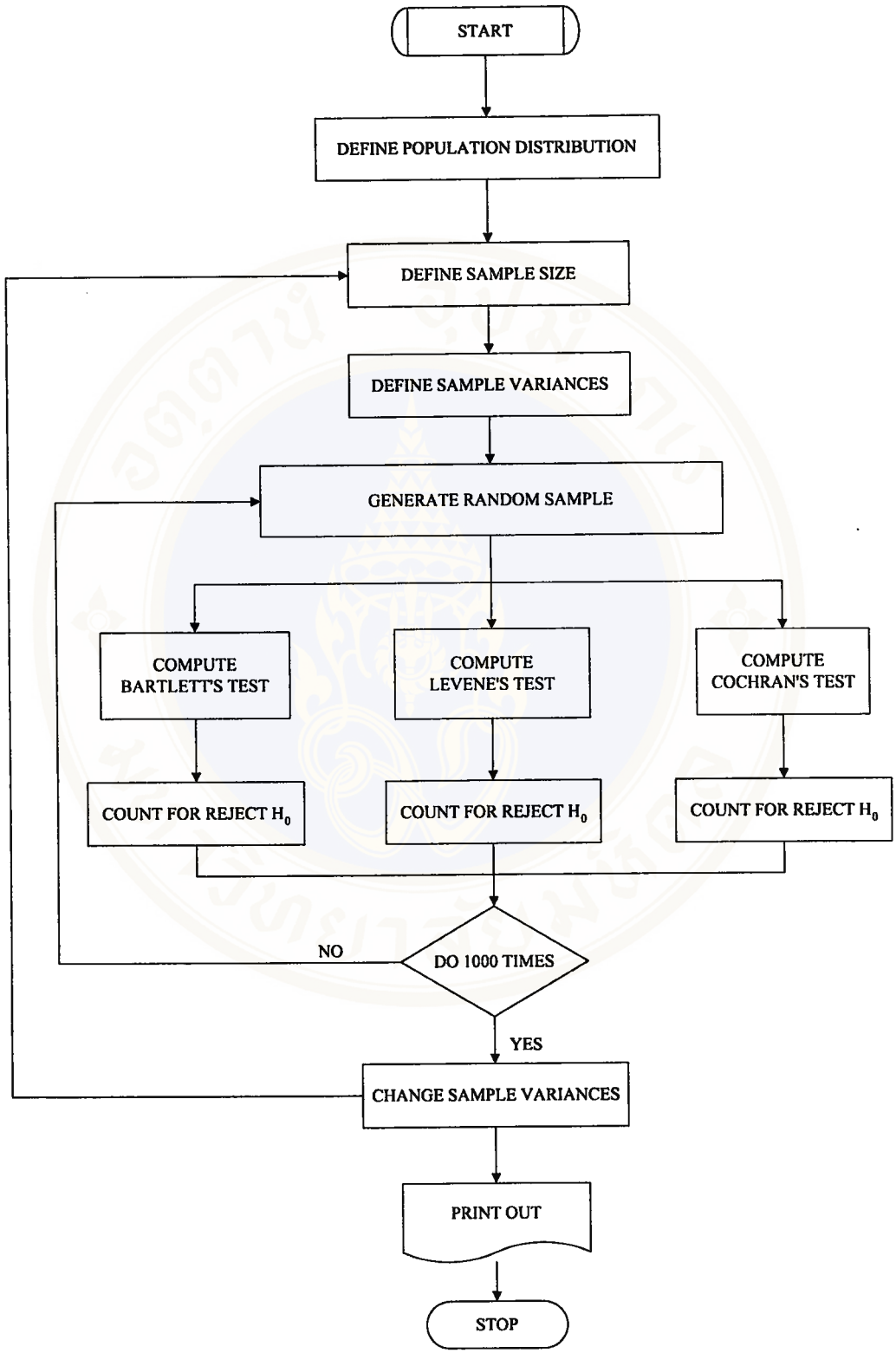


Figure 4. The Steps of Computation of Probability of Type I Error and Power of Three Statistical Tests

Copyright by Mahidol University

CHAPTER IV

RESULTS

This chapter is the results of testing the power and the robustness between Bartlett's test, Levene's test and Cochran's test. The data are defined as normal, gamma and chi-squares distributions. In each statistical test, the data sets consist of three groups of equal and unequal sample sizes that are 15, 30, and 45. Variance ratios are 1:1:2, 1:2:4 and 1:1:1 for equal variance. The levels of significance 0.01 and 0.05 are used in this study.

In the theory of hypothesis testing, the four possible combinations of these states of nature and decisions are illustrated along with a description of validity of the decision. (5)

		Decision	
		To reject	To accept
State of nature	H_0 is true	Type I error (Probability = α)	Correct decision (Probability = 1- α)
	H_1 is true	Correct decision Probability = 1 - β = Power	Type II error (Probability = β)

Copyright by Mahidol University

In this study, robustness means the ability of the test do not falsely detect non-homogeneous groups when the underlying data is not normally distributed and the groups are in fact homogeneous. Power means the ability of the test to detect non-homogeneous groups when the groups are in fact non-homogenous.

A. Type I Error

To designate a statistical test as robustness is considered by the departure of the empirical of type I error (τ) from the nominal level of significance (α) which does not exceed the precise limit. In this study, the robustness evaluating is based on Cochran limit as follows

at 0.01 significance level, τ value is between (0.007-0.015)

at 0.05 significance level, τ value is between (0.04-0.06)

τ = the true probability of a type I error or the empirical alpha

α = the nominal level of significance or the theoretical alpha

The statistical test is called robustness when its empirical alpha values lie within the Cochran limit. (33) If any actual probability of type I error is over the limit, it shows that the test cannot control the error rate. Therefore, it can be divided into two cases as the following

1. The type I error rate of any test is higher than the upper bound limit that regards the test has error rate more than the error rate limit.
2. The type I error rate of any test is lower than the lower bound limit that regards the test has error rate less than the error rate limit.

Copyright by Mahidol University

1. Normal Distribution

The results of actual type I error of the three statistical tests with N (10,26) at 0.01, 0.05 significance levels and sample sizes 15, 30 and 45 in 1,000 times were shown in Table 8.

Table 8 Effect of Empirical Type I Errors under Equal Variances Hypothesis of the Three Statistical Tests for Various Nominal Significance Levels with Normal Distribution

Nominal Level of Significance	Sample Sizes, n_j	Bartlett's Test	Levene's Test	Cochran's Test
0.01	15, 15, 15	**0.015***	0.002	**0.011***
	30, 30, 30	**0.009***	0.004	0.006
	45, 45, 45	**0.009***	**0.007***	**0.008***
	15, 30, 45	**0.012***	**0.013***	**0.011***
0.05	15, 15, 15	**0.047***	0.024	0.035
	30, 30, 30	**0.054***	**0.043***	**0.041***
	45, 45, 45	**0.051***	0.032	0.037
	15, 30, 45	**0.052***	**0.044***	**0.055***

Note: * Control type I error

Copyright by Mahidol University

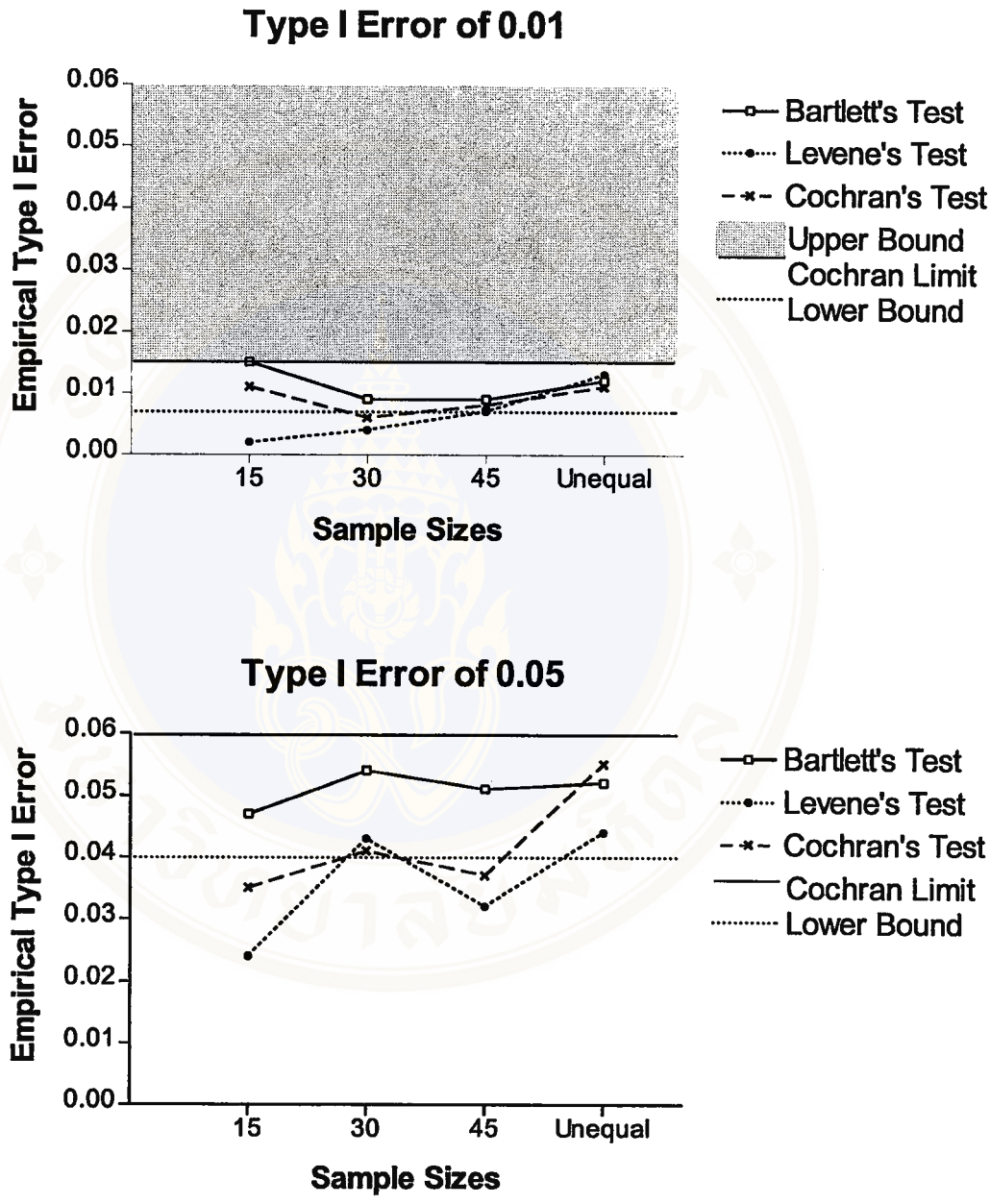


Figure 5. Probabilities of Type I Errors in the Three Statistical Tests with Normal Distribution

Copyright by Mahidol University

1.1 Robustness at 0.01 Significance Level

Bartlett's test can control type I error for all of the sample sizes [0.007-0.15], whereas Levene's test can control type I error for sample sizes of 45 and unequal sample sizes (15, 30, 45). Meanwhile, Cochran's test can control type I error of sample size of 15, 45 and unequal sample sizes. But Levene's test can control type I error rate lower than the lower bound limit in the sample size of 15 and 30. Beside this, Cochran's test can control type I error rate lower than the lower bound limit in the sample size of 30. This shows that Levene's test and Cochran's test have the type I error rate more than the error rate limit ($\alpha = 0.01$) (Figure 5).

1.2 Robustness at 0.05 Significance level

Bartlett's test can control type I error for all of the sample sizes [0.04-0.06]. Meanwhile, Levene's test and Cochran's test can control type I error for sample sizes of 30 and unequal sample sizes. Moreover, Levene's test and Cochran's test can control type I error rate less than the lower bound limit in the sample size of 15 and 45 (Figure 5).

2. Gamma Distribution

The results of type I error of the three statistical tests with $\alpha = 2.67$ and $\beta = 0.75$ at 0.01, 0.05 significance levels and sample sizes 15, 30 and 45 in 1,000 times were presented in Table 9.

Copyright by Mahidol University

Table 9. Effect of Empirical Type I Errors under Equal Variances Hypothesis of the Three Statistical Tests for Various Nominal Significance Levels with Gamma Distribution

Nominal Level of Significance	Sample Sizes, n_j	Bartlett's Test	Levene's Test	Cochran's Test
0.01	15, 15, 15	0.165	**0.013***	0.146
	30, 30, 30	0.290	0.062	0.243
	45, 45, 45	0.377	0.130	0.315
	15, 30, 45	0.195	0.039	0.364
0.05	15, 15, 15	0.317	0.080	0.292
	30, 30, 30	0.455	0.189	0.410
	45, 45, 45	0.563	0.306	0.492
	15, 30, 45	0.351	0.130	0.490

Note: * Control type I error

Copyright by Mahidol University

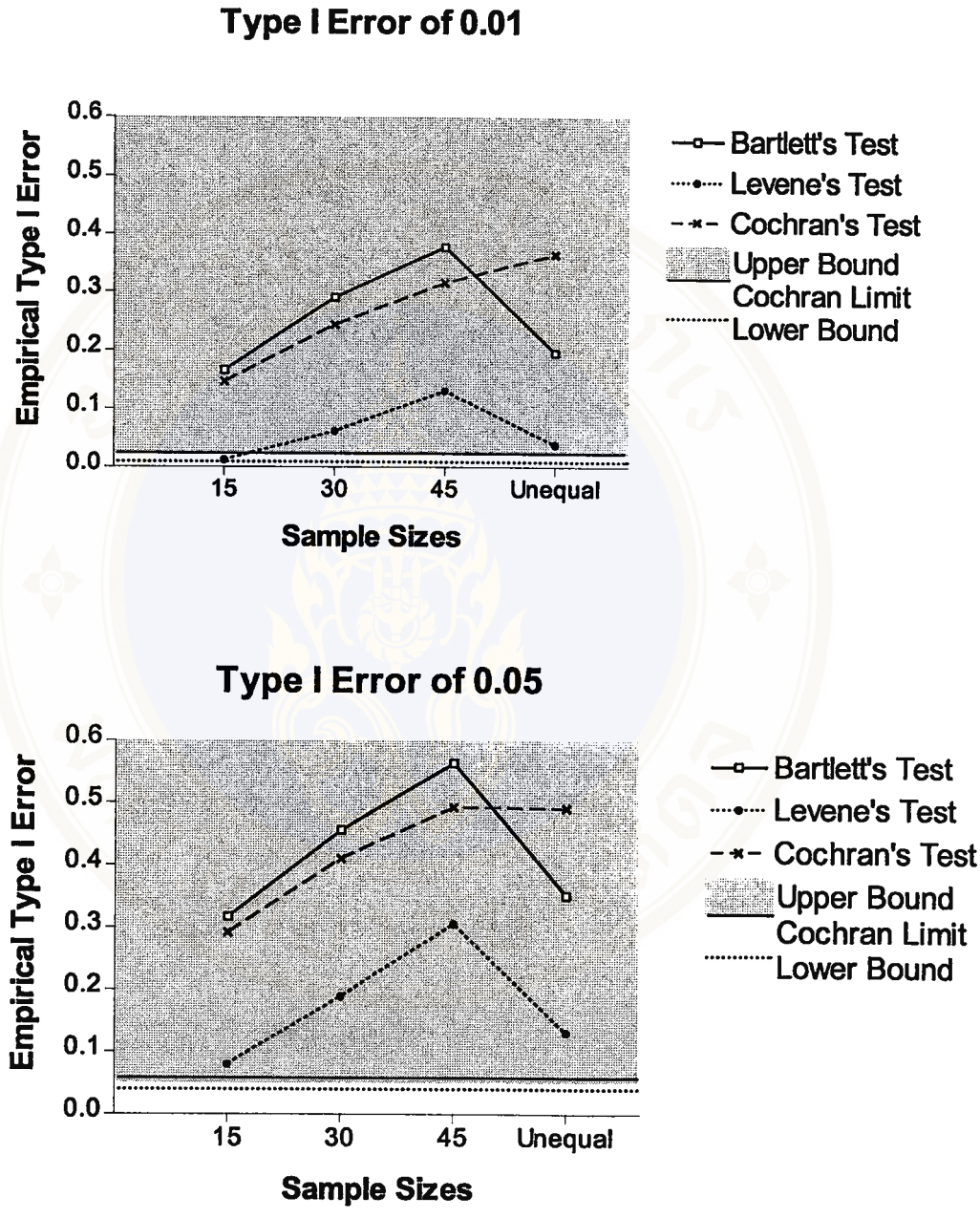


Figure 6. Probabilities of Type I Errors in the Three Statistical Tests with Gamma Distribution

Copyright by Mahidol University

2.1 Robustness at 0.01 Significance Level

Levene's test is the only test that can control type I error for the sample sizes of 15 [0.007-0.015], whereas Bartlett's test and Cochran's test cannot control type I error for all sample sizes since the type I error is over the upper bound limit (Figure 6.).

2.2 Robustness at 0.05 Significance Level

All of the three statistical tests: Bartlett's test, Levene's test and Cochran's test make type I error higher than the upper bound [0.04-0.06] which cannot control type I error for all sample sizes (Figure 6.).

Copyright by Mahidol University

3. Chi-square Distribution

The results of type I error of the three statistical tests with $p = n/2$ and $\alpha = 2$ at 0.01, 0.05 significance levels and sample sizes 15, 30 and 45 in 1,000 times were shown in Table 10.

Table 10 Effect of Empirical Type I Errors under Equal Variances Hypothesis of the Three Statistical Tests for Various Nominal Significance Levels with Chi-square Distribution

Nominal Level of Significance	Sample Sizes, n_j	Bartlett's Test	Levene's Test	Cochran's Test
0.01	15, 15, 15	0.027	0.003	0.029
	30, 30, 30	0.013	**0.007***	**0.011***
	45, 45, 45	0.021	0.006	**0.010***
	15, 30, 45	0.356	0.161	0.485
0.05	15, 15, 15	0.089	0.018	0.096
	30, 30, 30	0.072	**0.048***	**0.055***
	45, 45, 45	0.080	**0.051***	**0.050***
	15, 30, 45	0.591	0.441	0.684

Note: * Control type I error

Copyright by Mahidol University

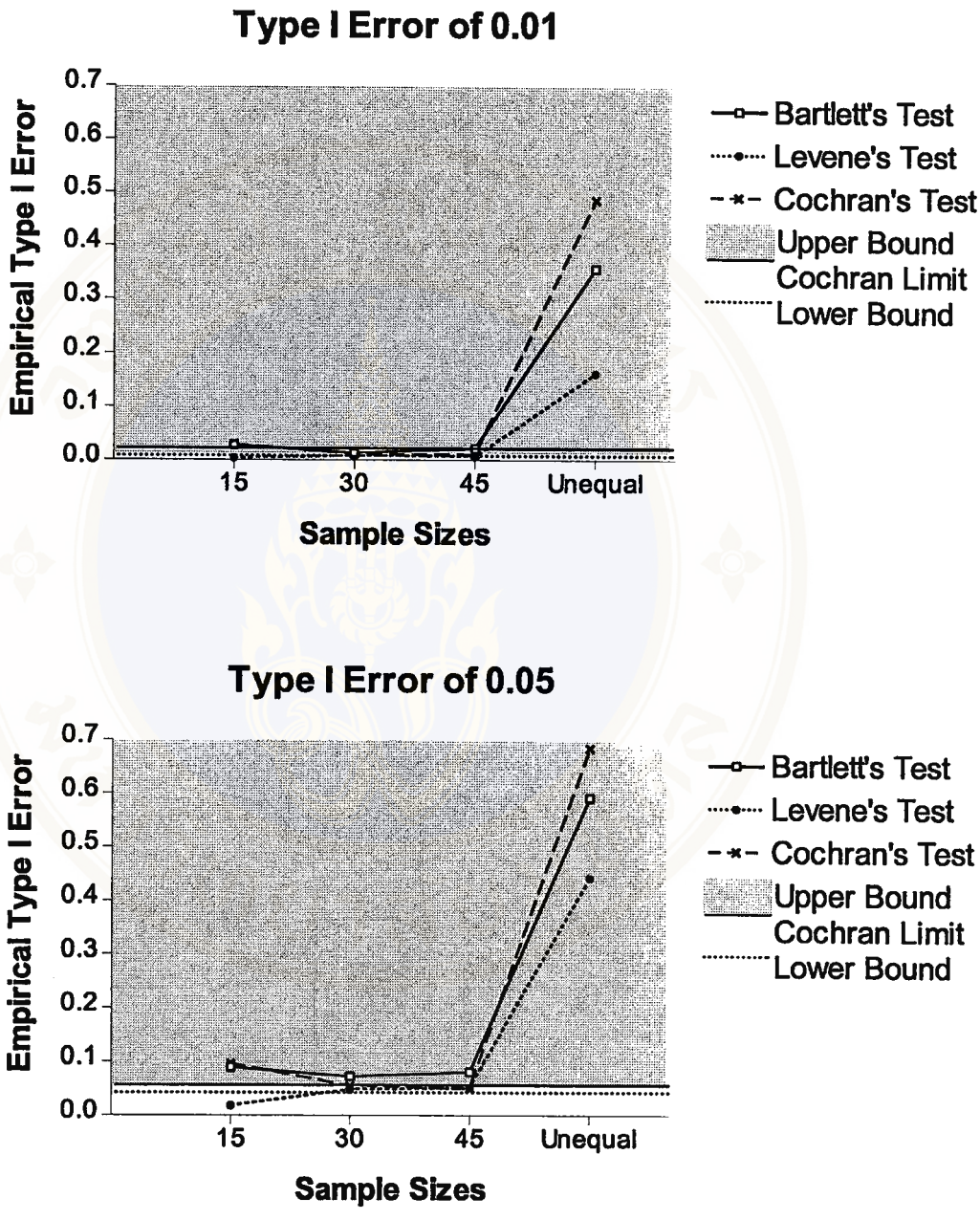


Figure 7. Probabilities of Type I Errors in the Three Statistical Tests with Chi-square Distribution

Copyright by Mahidol University

3.1. Robustness at 0.01 Significance Level

Type I error values of Cochran's test for sample sizes of 30 and 45 and Levene's test for sample size of 30 lie within the Cochran limit [0.007-0.015]. Moreover, type I error value of Levene's test for sample sizes of 15 and 45 lie lower than the Cochran limit. However Bartlett's test cannot control type I error for all sample sizes. For unequal sample sizes, there is no robust statistical test which the three statistical tests have type I error values over the upper bound limit (Figure 7).

3.2. Robustness at 0.05 Significance Level

From Table 10, we can see that Bartlett's test give higher the type I error than the Cochran limit [0.04-0.06]. As a result, it cannot control type I error for all sample sizes. However, Levene's test make increasing type I error when the sample size increases which is opposite to Cochran's test. For sample size of 15, Levene's test gives lower type I error than the lower bound limit and Cochran's test gives higher type I error than the upper bound limit. In addition, Levene's test can control type I error for sample sizes of 30 and 45 same Cochran's test. finally, the three statistical tests give type I error over the upper bound limit for unequal sample sizes (Figure 7.).

4. Normal Distribution, Gamma Distribution and Chi-square Distribution

In this study, the three tests were tested with three types of distribution: normal distribution with mean = n and variance = 2n, gamma distribution with mean = n and variance = 2n and chi-square distribution with $p = n/2$ and $\alpha = 2$. The null hypothesis of this part is

Copyright by Mahidol University

$$H_0 = \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 2n$$

$$H_1 = \text{at least one variance} \neq 2n$$

The results were presented in Table 11.

Table 11. Effect of Empirical Type I Errors under Equal Variances Hypothesis of the Three Statistical Tests for Various Nominal Significance Levels with Normal Distribution, Gamma Distribution and Chi-square Distribution

Nominal Level of Significance	Sample Sizes, n_j	Bartlett's Test	Levene's Test	Cochran's Test
0.01	15, 15, 15	0.024	**0.007***	0.021
	30, 30, 30	**0.013***	0.006	**0.007***
	45, 45, 45	**0.012***	0.005	**0.008***
0.05	15, 15, 15	0.095	**0.043***	0.083
	30, 30, 30	**0.060***	0.037	0.036
	45, 45, 45	0.061	0.038	**0.043***

Note: * Control type I error

Copyright by Mahidol University

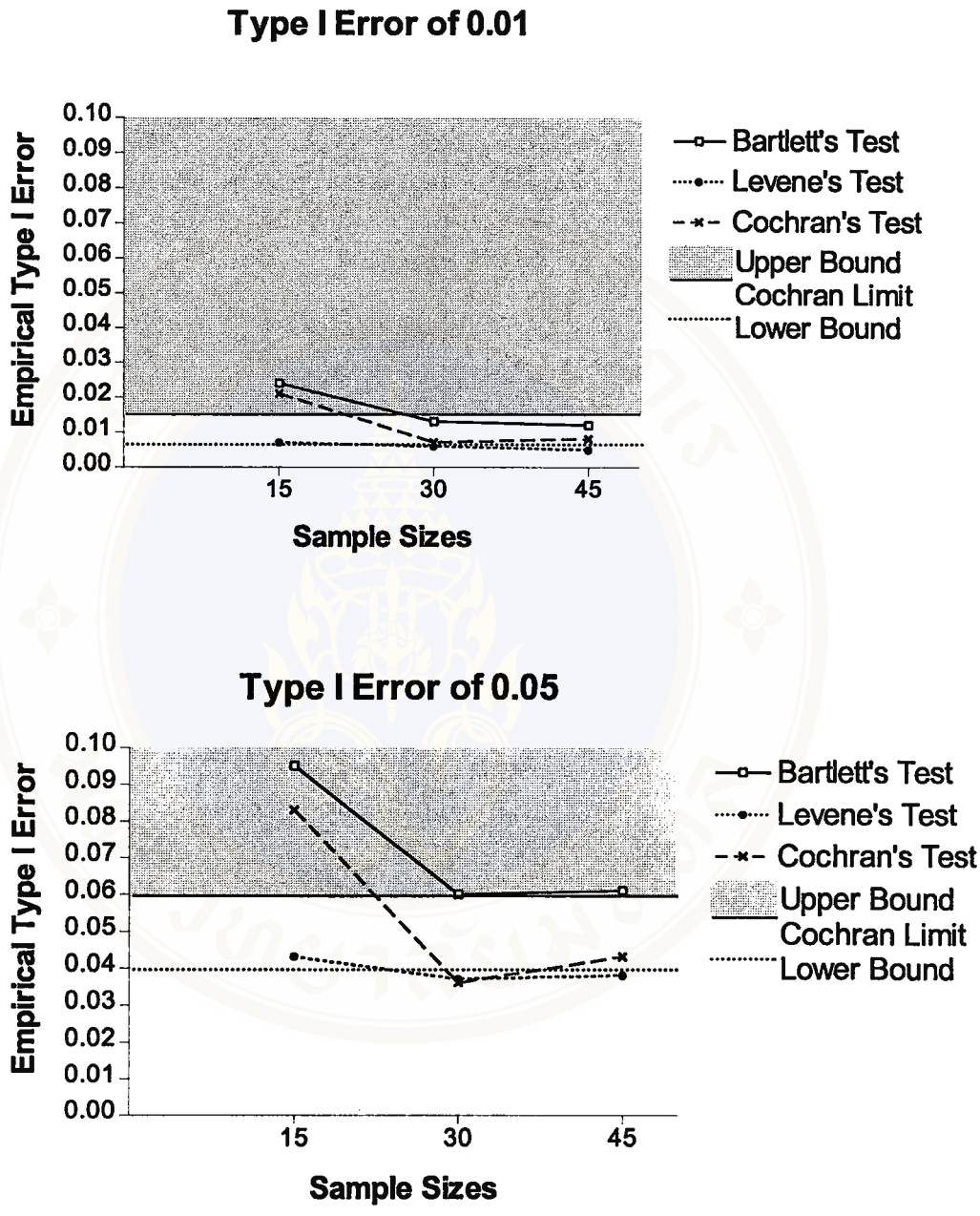


Figure 8. Probabilities of Type I Errors in the Three Statistical Tests with Normal Distribution, Gamma Distribution and Chi-square Distributions

Copyright by Mahidol University

4.1 Robustness at 0.01 Significance Level

For sample size of 15, there is only one statistical test that can control type I error. Levene's test is the one robust test when sample size is small. For sample sizes of 30 and 45, Levene's test can control type I error lower than Cochran limit [0.007-0.015]. However, Bartlett's test and Cochran's test give type I error within Cochran limit. Therefore, Bartlett's test and Cochran's test are robust, and they can control type I error when sample sizes are appropriate (Figure 8).

4.2 Robustness at 0.05 Significance Level

For sample size of 15, Levene's test is the only one test that can control type I error within Cochran limit [0.04-0.06]. For sample size of 30, Bartlett's test is robust and it can control type I error within Cochran limit. On the other hand, Levene's test and Cochran's test give lower type I error than the lower bound limit. For sample size of 45, Cochran's test is robust, and it can control type I error limit. Furthermore, Levene's test give lower type I error than the lower bound limit and Bartlett's test give higher type I error than the upper bound limit (Figure 8).

B. Power of Test

The results from testing the power of Levene's test, Bartlett's test and Cochran's test in 2 variances ratios on the three types of distribution for the sample sizes of 15, 30 and 45 at 0.01 and 0.05 significance levels were presented in Tables 12, 13, 14 and 15, respectively.

Copyright by Mahidol University

1. Normal Distribution

Table 12. Effect of Heterogeneous Variances on Power of the Three Statistical Tests for Various Nominal Significance Levels with Normal Distribution

Nominal Level of Significance	Ratio of σ^{2}'s	Sample Sizes, n_j	Bartlett's Test	Levene's Test	Cochran's Test
0.01	1:1:2	15, 15, 15	0.098	0.047	0.113
		30, 30, 30	0.249	0.162	0.249
		45, 45, 45	0.471	0.348	0.473
		15, 30, 45	0.245	0.238	0.258
	1:2:4	15, 15, 15	0.352	0.157	0.291
		30, 30, 30	0.767	0.553	0.576
		45, 45, 45	0.948	0.866	0.840
		15, 30, 45	0.604	0.374	0.784
0.05	1:1:2	15, 15, 15	0.246	0.154	0.274
		30, 30, 30	0.448	0.354	0.452
		45, 45, 45	0.688	0.594	0.700
		15, 30, 45	0.471	0.390	0.472
	1:2:4	15, 15, 15	0.599	0.407	0.509
		30, 30, 30	0.901	0.797	0.785
		45, 45, 45	0.988	0.963	0.929
		15, 30, 45	0.824	0.695	0.905

Copyright by Mahidol University

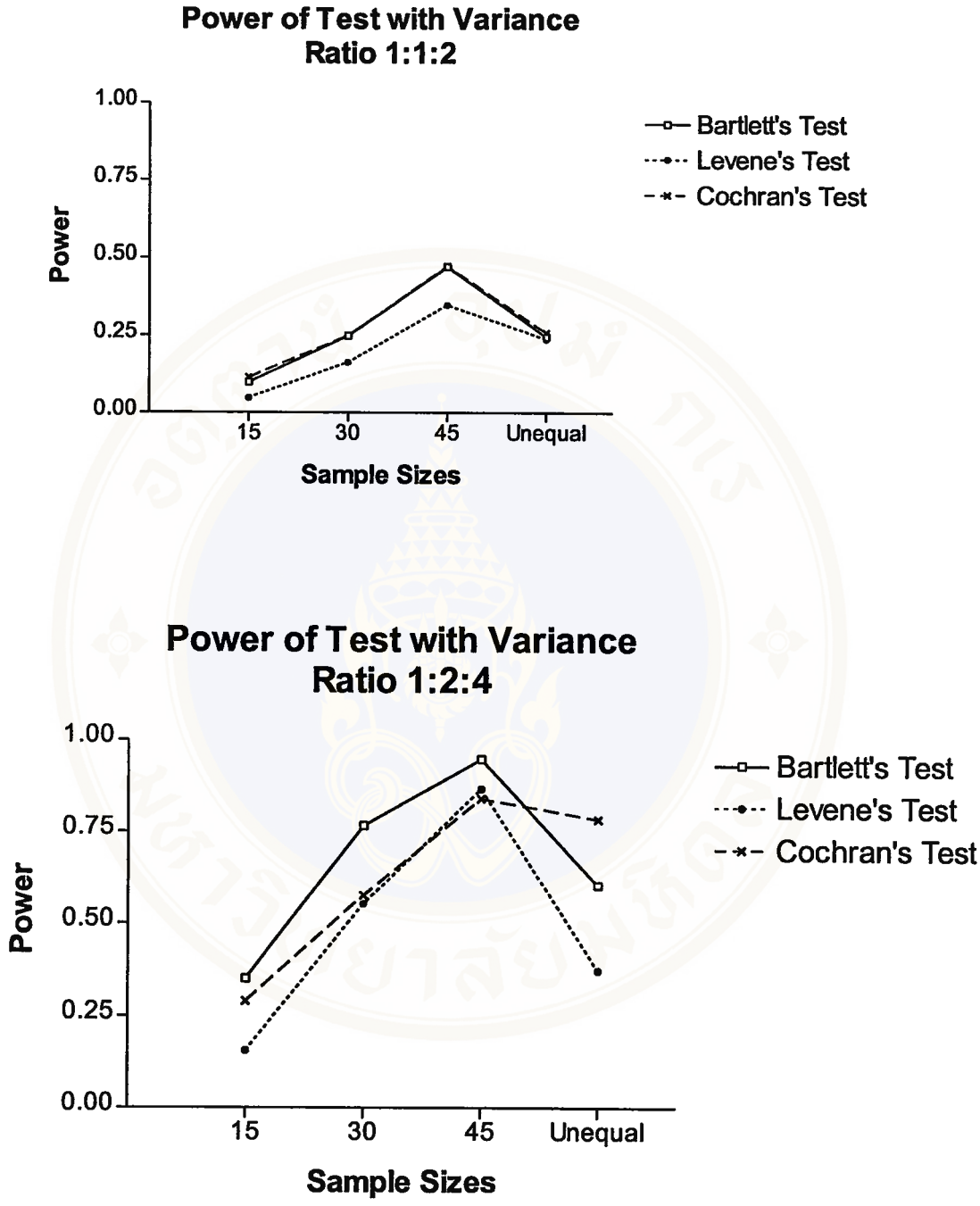


Figure 9. Power of the Three Statistical Tests on Effect of Variance Ratios and Sample Sizes with Normal Distribution at 0.01 Significance Level

Copyright by Mahidol University

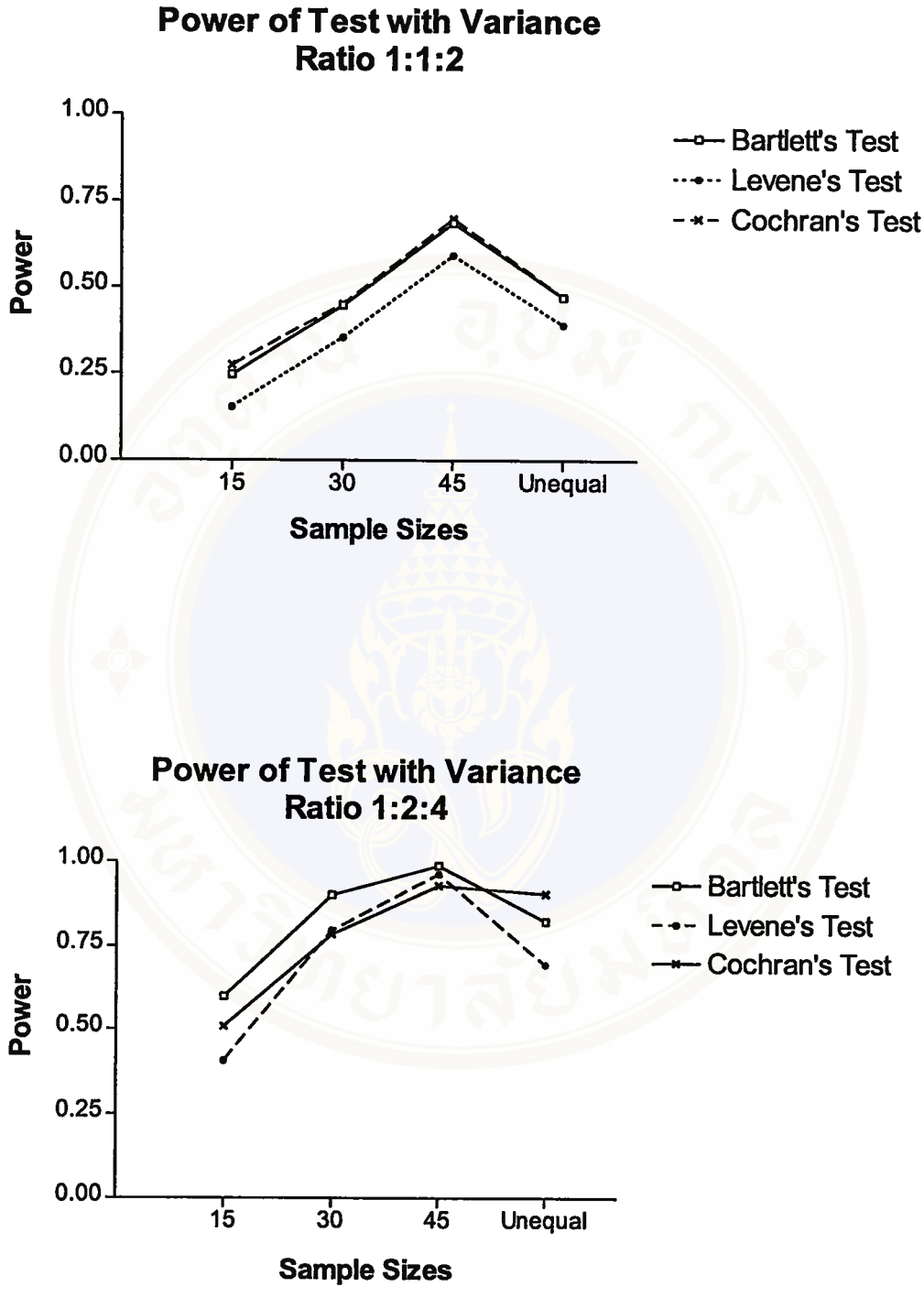


Figure 10. Power of the Three Statistical Tests on Effect of Variance Ratios and Sample Sizes with Normal Distribution at 0.05 Significance Level

Copyright by Mahidol University

LIBRARY AND INFORMATION CENTER MAHIDOL UNIVERSITY

1.1 Power of Test at 0.01 Significance Level

1.1.1 For Variance Ratio = 1:1:2

For all of equal sample size and unequal sample sizes, the power of Levene's test is the lowest and Cochran's test is the highest. When, the sample size increases, the power tends to increase (Figure 9).

1.1.2 For Variance Ratio = 1:2:4

All the three statistical tests have increasing power as the sample size increases. For the sample sizes of 15, 30 and 45, Bartlett's test has the highest power while Levene's test has the lowest power in 15, 30 and unequal sample sizes. On the other hand, Cochran's test has the lowest power in sample size of 45 and it has the highest power in unequal sample sizes (Figure 9).

1.2 Power of Test at 0.05 Significance Level

1.2.1 For Variance Ratio = 1:1:2

For all sample sizes, the three statistical tests have increasing power while the sample size increases. Beside this, Levene's has the lowest power, but Cochran's test has the highest power (Figure 10).

1.2.2 For Variance Ratio = 1:2:4

The power of the three statistical tests tends to get higher as the sample size increases. For all of sample sizes, Bartlett's test has the highest power while Levene's test has the lowest power in the sample size of 15 and unequal sample sizes. For the sample size of 30 and 45 the power of Cochran's test is the lowest, but when the sample size is unequal, Cochran's test has the highest power (Figure 10).

Copyright by Mahidol University

2. Gamma Distribution

Table 13 Effect of Heterogeneous Variances on Power of the Three Statistical Tests for Various Nominal Significance Levels with Gamma Distribution

Nominal Level of Significance	Ratio of $\sigma^{2'}s$	Sample Sizes, n_j	Bartlett's Test	Levene's Test	Cochran's Test
0.01	1:1:2	15, 15, 15	0.198	0.009	0.170
		30, 30, 30	0.303	0.034	0.192
		45, 45, 45	0.424	0.071	0.245
		15, 30, 45	0.310	0.026	0.347
	1:2:4	15, 15, 15	0.275	0.005	0.260
		30, 30, 30	0.436	0.022	0.388
		45, 45, 45	0.541	0.044	0.469
		15, 30, 45	0.442	0.009	0.550
0.05	1:1:2	15, 15, 15	0.355	0.080	0.295
		30, 30, 30	0.458	0.129	0.348
		45, 45, 45	0.607	0.228	0.429
		15, 30, 45	0.501	0.118	0.528
	1:2:4	15, 15, 15	0.446	0.040	0.390
		30, 30, 30	0.569	0.094	0.515
		45, 45, 45	0.672	0.153	0.591
		15, 30, 45	0.612	0.057	0.673

Copyright by Mahidol University

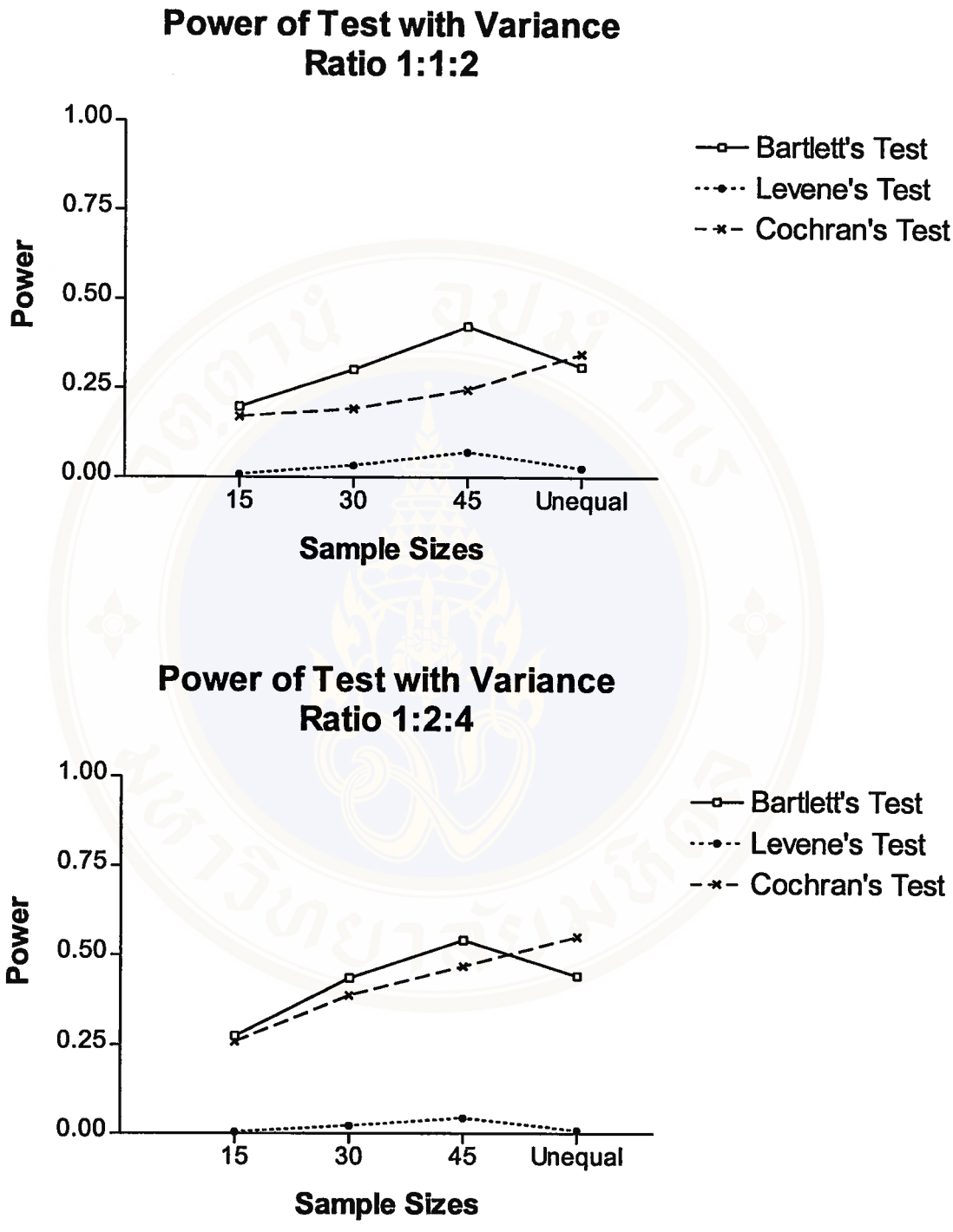


Figure 11. Power of the Three Statistical Tests on Effect of Variance Ratios and Sample Sizes with Gamma Distribution at 0.01 Significance Level

Copyright by Mahidol University

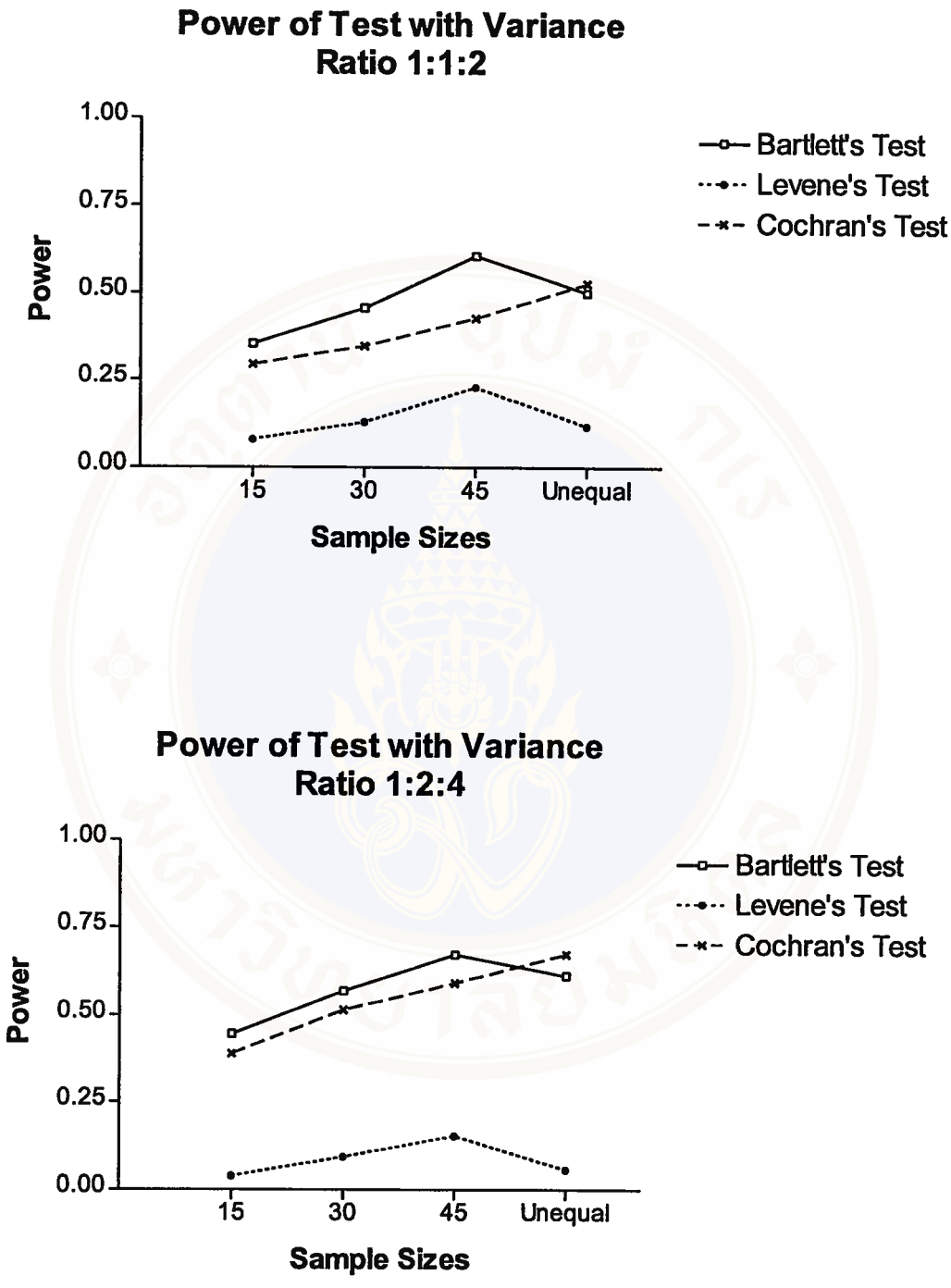


Figure 12. Power of the Three Statistical Tests on Effect of Variance Ratios and Sample Sizes with Gamma Distribution at 0.05 Significance Level

Copyright by Mahidol University

2.1 Power of Test at 0.01 Significance Level

2.1.1 For Variance Ratio = 1:1:2

Power of the three statistical tests tends to increase as the sample size increases. For all equal sample size and unequal sample sizes, Levene's test has the lowest power while Bartlett's test has the highest power for sample size of 15, 30 and 45 and Cochran's test has the highest power of unequal sample sizes. Moreover, the bigger sample size, the power tends to increase (Figure 11).

2.1.2 For Variance Ratio = 1:2:4

All the three statistical tests have increasing power while the sample size increases. For the sample size of 15, 30 and 45, Bartlett's test has the highest power while Levene's test has the lowest power in all sample size and unequal sample sizes. Cochran's test has the highest power in unequal sample sizes (Figure 11).

2.2 Power of Test at 0.05 Significance Level

2.2.1 For Variance Ratio = 1:1:2

At all of sample sizes, the three statistical tests tends to increase power as the sample size increases. And Levene's test has the smallest power. While Bartlett's test has the highest power of sample size of 15, 30 and 45. In case of unequal sample sizes, Cochran's test has the highest power (Figure 12).

2.2.2 For Variance Ratio = 1:2:4

The power of the three statistical tests is increasing power when the bigger sample size. For all of equal sample size, Levene's test has the smallest power. While Bartlett's test has the highest power of sample size of 15,

Copyright by Mahidol University

30 and 45. For unequal sample sizes, Cochran's test has the highest power (Figure 12).

3. Chi-square Distribution

Table 14. Effect of Heterogeneous Variances on Power of the Three Statistical Tests for Various Nominal Significance Levels with Chi-square Distribution

Nominal Level of Significance	Ratio of σ^{2}'s	Sample Sizes, n_j	Bartlett's Test	Levene's Test	Cochran's Test
0.01	1:1:2	15, 15, 15	0.129	0.044	0.128
		30, 30, 30	0.247	0.149	0.272
		45, 45, 45	0.448	0.324	0.437
		15, 30, 45	0.318	0.150	0.467
	1:2:4	15, 15, 15	0.373	0.141	0.310
		30, 30, 30	0.754	0.575	0.584
		45, 45, 45	0.936	0.857	0.797
		15, 30, 45	0.632	0.392	0.774
0.05	1:1:2	15, 15, 15	0.275	0.146	0.287
		30, 30, 30	0.460	0.382	0.466
		45, 45, 45	0.676	0.571	0.674
		15, 30, 45	0.518	0.357	0.616
	1:2:4	15, 15, 15	0.621	0.393	0.525
		30, 30, 30	0.909	0.803	0.792
		45, 45, 45	0.986	0.963	0.923
		15, 30, 45	0.829	0.710	0.904

Copyright by Mahidol University

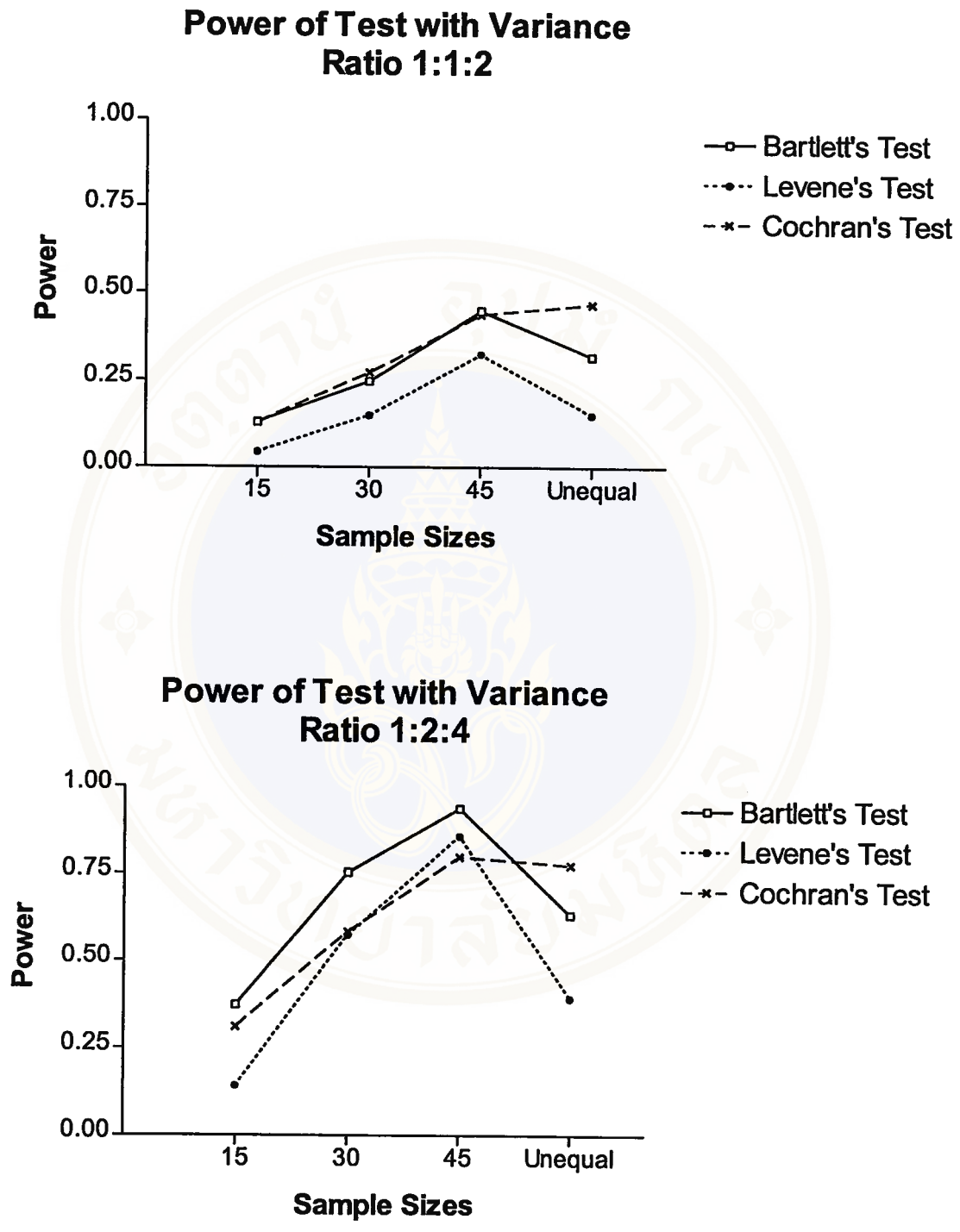


Figure 13. Power of the Three Statistical Tests on Effect of Variance Ratios and Sample Sizes with Chi-square Distribution at 0.01 Significance Level

Copyright by Mahidol University

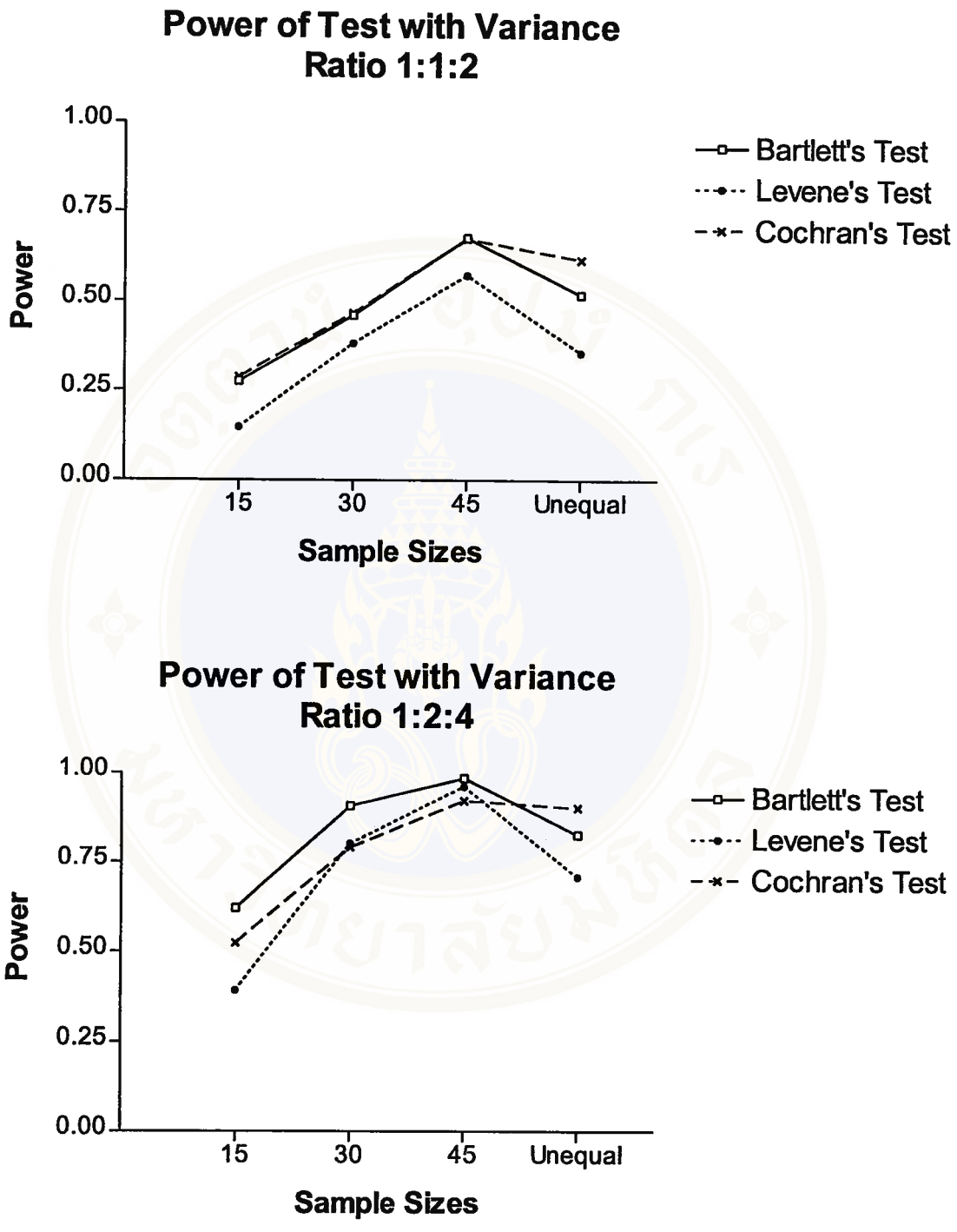


Figure 14. Power of the Three Statistical Tests on Effect of Variance Ratios and Sample Sizes with Chi-square Distribution at 0.05 Significance Level

Copyright by Mahidol University

3.1 Power of Test at 0.01 Significance Level

3.1.1 For Variance Ratio = 1:1:2

When the sample size increases, the power tends to increase. For all equal sample size and unequal sample sizes, Levene's test has the smallest power. Cochran's test has the highest power when the sample size is 30 and sample sizes are unequal. While Bartlett's test has the highest power for sample size of 15 and 45. For unequal sample sizes, Cochran's test has the highest power. (Figure 13).

3.1.2 For Variance Ratio = 1:2:4

All the three statistical tests have increasing power when the sample size increases. For the sample size of 15, 30 and 45, Bartlett's test has the highest power while Levene's test has the lowest power in 15, 30 and unequal sample sizes, but when the sample size is 45, Cochran's test has the lowest power. For unequal sample sizes, Cochran's test has the highest power (Figure 13).

3.2 Power of Test at 0.05 Significance Level

3.2.1 For Variance Ratio = 1:1:2

For all equal sample sizes and unequal sample sizes, the three statistical tests have increasing power when the sample size increases. Beside this, Levene's test has the smallest power for all sample sizes. Moreover, Cochran's test has the highest power for sample size of 30, 45 and unequal sample sizes. In addition, Bartlett's test has the highest power for sample size of 45 (Figure 14).

Copyright by Mahidol University

3.2.2 For Variance Ratio = 1:2:4

The power of the three statistical tests tends to be higher as the sample size increases. For all equal sample sizes, Bartlett's test has the highest power while Levene's test has the lowest power in the sample size of 15,30 and unequal sample sizes. For the sample size of 45, Cochran's test has the lowest power (Figure 14).

Copyright by Mahidol University

4. Normal Distribution, Gamma Distribution and Chi-square Distribution

Table 15 Effect of Heterogeneous Variances on Power of the Three Statistical Tests for Various Nominal Significance Levels with Normal Distribution, Gamma Distribution and Chi-square Distribution

Nominal Level of Significance	Ratio of σ^{2}'s	Sample Sizes, n_j	Bartlett's Test	Levene's Test	Cochran's Test
0.01	1:1:2	15, 15, 15	0.103	0.050	0.114
		30, 30, 30	0.279	0.156	0.278
		45, 45, 45	0.457	0.345	0.455
	1:2:4	15, 15, 15	0.357	0.138	0.282
		30, 30, 30	0.788	0.579	0.618
		45, 45, 45	0.948	0.870	0.821
0.05	1:1:2	15, 15, 15	0.249	0.153	0.270
		30, 30, 30	0.486	0.371	0.477
		45, 45, 45	0.690	0.598	0.687
	1:2:4	15, 15, 15	0.598	0.371	0.512
		30, 30, 30	0.919	0.808	0.805
		45, 45, 45	0.992	0.964	0.933

Copyright by Mahidol University

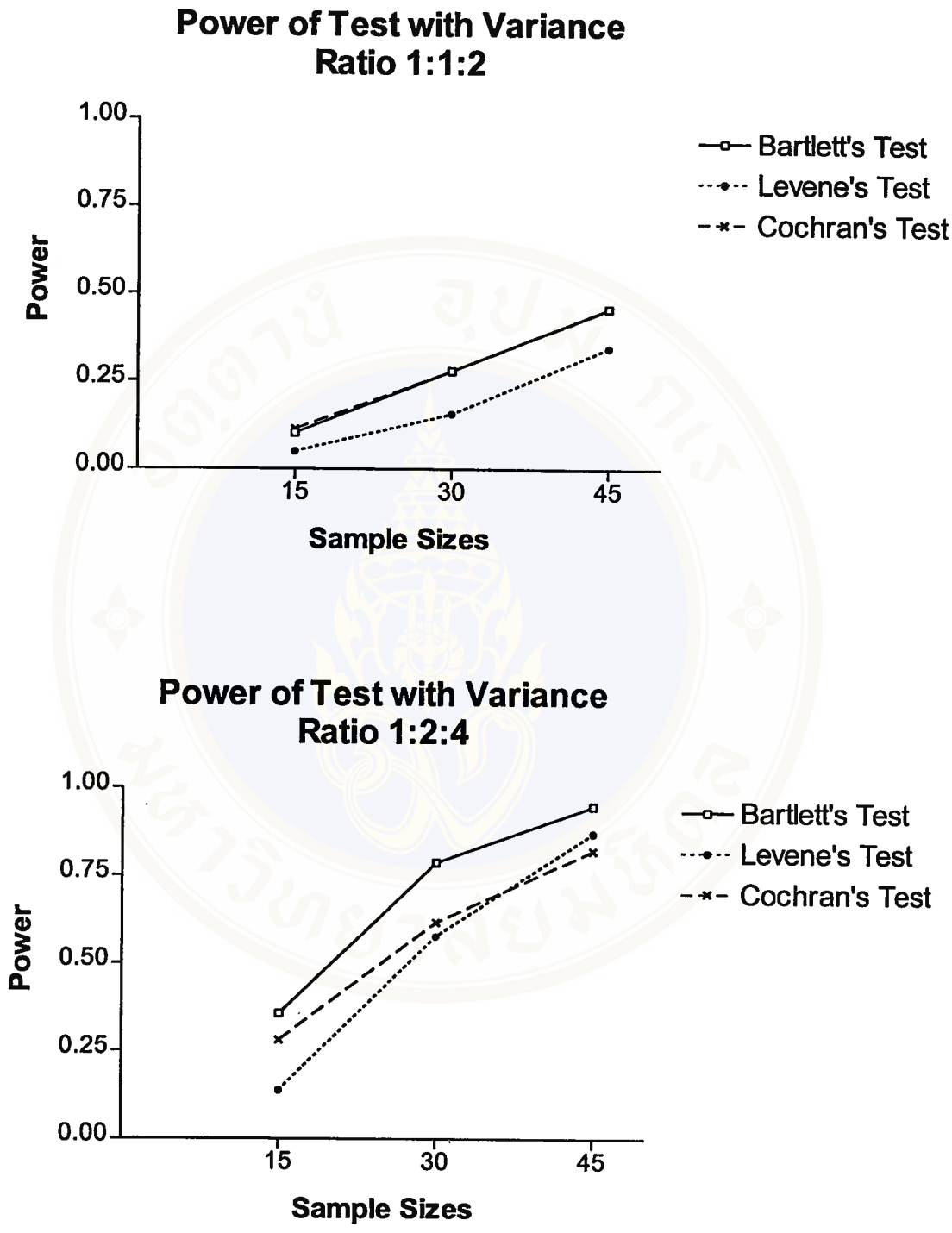


Figure 15. Power of the Three Statistical Tests on Effect of Variance Ratios and Sample Sizes with Normal Distribution, Gamma Distribution and Chi-square Distribution at 0.01 Significance Level

Copyright by Mahidol University

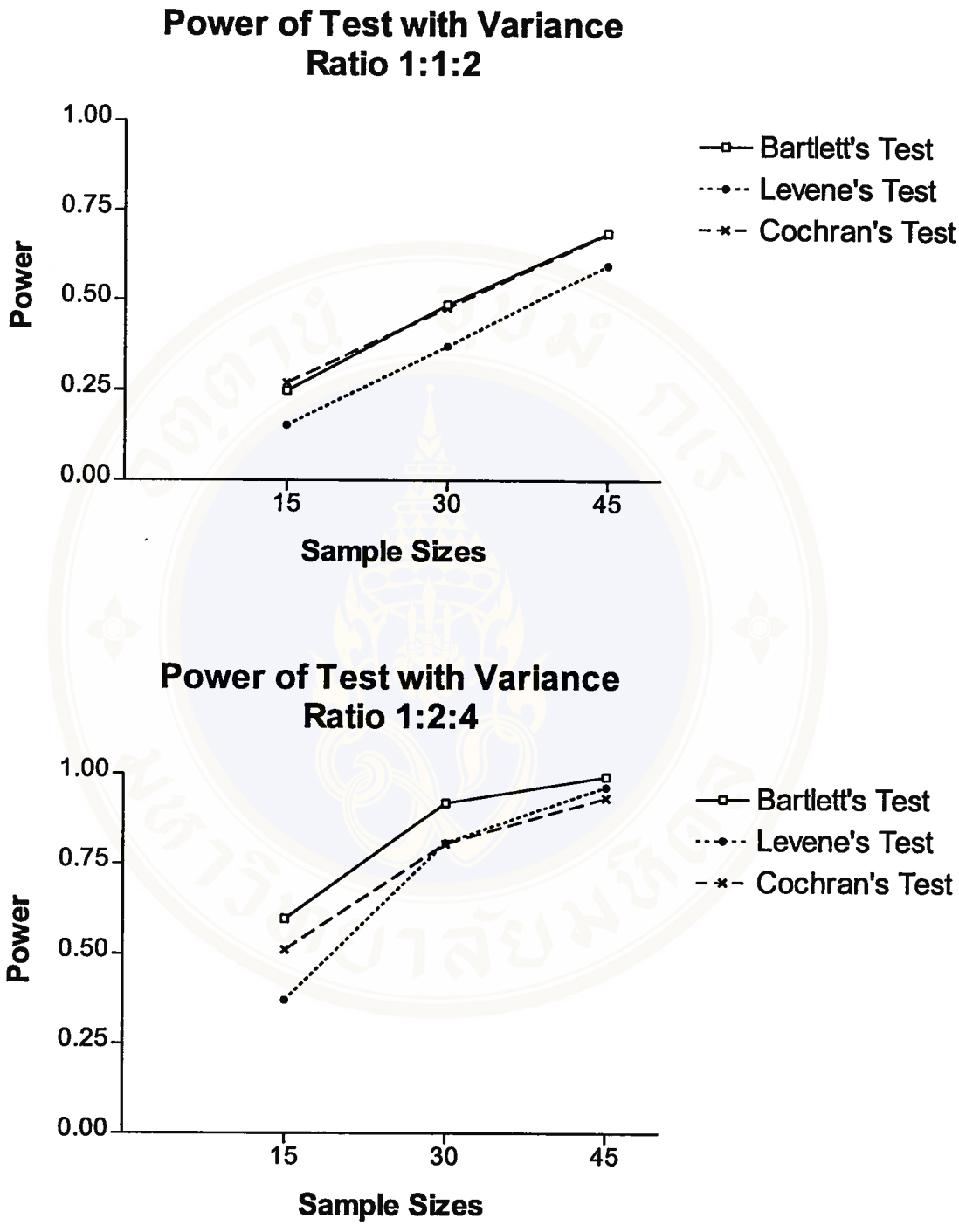


Figure 16. Power of the Three Statistical Tests on Effect of Variance Ratios and Sample Sizes with Normal Distribution, Gamma Distribution and Chi-square Distribution at 0.05 Significance Level

Copyright by Mahidol University

4.1 Power of Test at 0.01 Significance Level

The power of the three statistical tests tend to be high as the sample size increases.

4.1.1 For Variance Ratio = 1:1:2

For all sample sizes, Levene's test has the lowest power. When the sample size of 15, Cochran's test has the highest power. For sample sizes of 30 and 45, Bartlett's test has the highest power (Figure 15).

4.1.2 For Variance Ratio = 1:2:4

For all sample sizes, Bartlett's test has the highest power likewise Cochran's test has higher power than Levene's test for sample sizes of 15 and 30. However Levene's test has the higher power than Cochran's test for the sample size of 45 (Figure 15).

4.2 Power of Test at 0.05 Significance Level

4.2.1 For Variance Ratio = 1:1:2

Levene's test has the lowest power for all sample size while Cochran's test has the highest power at the sample sizes of 15. For the sample size of 30 and 45, Bartlett's test has the highest power (Figure 16).

4.2.2 For Variance Ratio = 1:2:4

For all sample sizes, Bartlett's test has the highest power for all sample sizes. On the contrary, Levene's test has the lowest at the sample size of 15. And Cochran's test has the lowest at the sample size of 30 and 45 (Figure 16).

Copyright by Mahidol University

CHAPTER V

DISCUSSION

This chapter describes the efficiency of Bartlett's test, Levene's test and Cochran's test in term of high power of test and robustness that considers from Cochran limit.

A. Type I Error

1. Normal Distribution

Under the normal assumption, homogeneity of variances assumption and equality of sample sizes, the type I error probability of Levene's test and Cochran's test are lower than Cochran limit in sample size of 15. All type I error probabilities under various sample sizes of Bartlett's test are robust as their empirical alpha values lie within Cochran limit. This means that Bartlett's test can control error rate. Also the empirical alpha of sample sizes of 45 in 0.05 significance level and the empirical alpha of sample sizes of 30 in 0.01 significance level are lower than the lower bound of Cochran limit. This is quite strong for Levene's test and Cochran's test when normal assumption and equality of sample sizes are met.

This is the same as what Brown and Forsythe (7) found from their research, that is when the Gaussian distribution indicated, that Levene's test with median (W_{50})

Copyright by Mahidol University

is for small sample sizes. And also Gartside (6) found Bartlett's test should be used when the alternative hypothesis is unknown and normality can be relied on.

Furthermore, normal assumption and homogeneity of variances assumption are met, but equality of sample sizes is violated. The three statistical tests are robust since they can control type error within Cochran limit. This shows that unequal sample sizes have not affected to probability of type I error of the three statistical tests.

2. Gamma Distribution

Under homogeneity of variances assumption and equality of sample sizes are met but normal assumption is violated, Levene's test is robust. However, Bartlett's test and Cochran's test are not robust. This is because their probabilities of type I error are over the upper bound of Cochran limit. Furthermore, if homogeneity of variances assumption is met, normal assumption and unequal sample sizes are violated. There are no statistical tests that are robust.

These results correspond to all theories of three statistical tests. In other words, Bartlett's test is not a good test when distribution is doubtful. Moreover, Cochran's test is not a good choice as it cannot maintain good power from population when variances are equal (6). Meanwhile, Levene's test is found to be robust under nonnormality (7).

3. Chi-square Distribution

The empirical alpha of the three statistical tests at theoretical alpha 0.01 seems the empirical alpha at theoretical alpha 0.05 but they give higher the empirical

Copyright by Mahidol University

alpha at theoretical alpha 0.05 than theoretical alpha 0.01. Moreover, Bartlett's test is non-robust when chi-square distribution violates normal assumption. On the contrary, Levene's test has the empirical alpha lower than the lower bound of Cochran limit for sample sizes of 15. On the other hand, Levene's test and Cochran's test are robust in sample size of 30 and 45. This is according to the general rule of thumb that defines the sample size of 30 (14). The results of Levene's test are as same as Brown and Forsythe (7) that made sampling from chi-square with four degree of freedom. It turned out that only Levene's test with median maintains it size near the five percent level of significance.

Under homogeneity of variance and equality of sample sizes are met but normal assumption is violated, Bartlett's test is non-robust. Lastly, when homogeneity of variances assumption is met but equality of sample sizes and normal assumption are violated, no statistical tests are robust.

4. Normal Distribution, Gamma Distribution and Chi-square Distribution

Under homogeneity of variances and equal sample sizes are met but normal assumption is violated, all three statistical tests are robust. At theoretical alpha 0.01 and 0.05, Levene's test is robust when sample size is 15, also has the empirical alpha lower than the lower bound of Cochran limit in sample sizes of 30 and 45. At theoretical alpha 0.01, Bartlett's test and Cochran's test are robust in sample size of 30 and 45. At theoretical alpha 0.05, Bartlett's test is robust for sample size of 30. This sample size is follows the general rule of thumb that defines the sample size of 30 (14). While Cochran's test is robust for sample size of 45, and it has lower

Copyright by Mahidol University

empirical alpha than the lower bound limit. This is different from Conover, Johnson and Johnson's study (15). They simulated various symmetric and asymmetric distributions for four sample groups with respective sample sizes (n_1, n_2, n_3, n_4) = (5, 5, 5, 5), (10, 10, 10, 10), (20, 20, 20, 20), and (5, 5, 20, 20). The null hypothesis of equal variances (all equal to 1) was examined along with the four alternatives ($\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2$) = (1, 1, 1, 2), (1, 1, 1, 4), (1, 1, 1, 8), and (1, 2, 4, 8). They found that Levene's test with median (lev1: med) is a good choice in terms of robustness and power.

B. Power Of Test

1. Normal Distribution

Under normal assumption and equality of sample sizes are met but homogeneity of variances is violated, Cochran's test is a good choice as it maintains good power, if one variance is larger when variance ratio is equal to 1: 1: 2. This is the same as Garside's study (6). However, when variances ratio is not equal (1: 2: 4), Bartlett's test gives the highest power. This shows that it corresponds to Brown and Forsythe's (7) study.

2. Gamma Distribution

Under equality of sample size is met, but normal assumption and homogeneity of variances are violated, Bartlett's test has the highest power for all

Copyright by Mahidol University

sample sizes. This result is different form the hypothesis that the violation of normal distribution assumption may not result in the loss of power of Levene's test.

For normal assumption, homogeneity of variance assumption and equality of sample sizes are violated. Cochran's test still is a good choice that gives the highest power. The result is the same in different empirical alpha.

3. Chi-square Distribution

Under sample sizes are equal, but homogeneity of variances assumption and normal assumption are violated, Bartlett's test seems to be a good statistical test which has the highest power in all sample sizes. This is different from the hypothesis that the violation of normal distribution assumption may not result in the loss of power of Levene's test.

Under normal assumption, homogeneity of variances assumption and equality of sample sizes are violated, Cochran's test is a good choice that gives the highest power when one variance is larger. That is the same as what Garside said(6). However, when unequal variances ration (1: 2: 4) is met under non-normality and unequal of sample sizes, Bartlett's test is a good choice since it gives the highest power.

4. Normal Distribution, Gamma Distribution and Chi-square Distribution

Under normal assumption and homogeneity of variances assumption are violated and equality of sample sizes is met, the power of the three statistical tests do not differ. When the sample size is 15 and one variance is larger, Cochran's test is

Copyright by Mahidol University

found to be a good test. Again, this corresponds to Garside's study (6). Contrary, when sample sizes 30 and 45 and one variance is larger, Bartlett's test gives the highest power. This result is similar to variances ratio 1: 2: 4 for all sample sizes. It is possible that the sample sizes of 30 and 45 seem to be normal according to the rule of thumb, thus Bartlett's test can reject hypothesis of equal variances.

In order to choose an appropriate statistical test, there are two factors that must be taken into account. The first is that probability of type I error value must lie within Cochran limit. Another factor is the high power of test. If normal assumption is met, Bartlett's test is still robust. Since its probabilities of type I error lies within Cochran limit and it gives the highest power. Anyway, this is not true when one variance is larger. This means Cochran's test is the best choice in this situation, unless the data is not normally distributed. In this case, Levene's test is the best test, which is more robust than Bartlett's test and Cochran's test that can control type I error within Cochran limit. Moreover, Levene's test can control type I error in small sample sizes. On the contrary, Bartlett's test gives the highest power when sample sizes are equal, but Cochran's test has no effect on unequal sample sizes since it still gives the highest power when sample sizes are unequal. This is opposite to the hypothesis of unequality sample sizes could result in the loss of power of Cochran's test more than Bartlett's test and Cochran's test.

Copyright by Mahidol University

CHAPTER VI

CONCLUSION AND RECOMMENDATION

In this study, after simulation involving distributions, sample sizes, mean and variances; the statistical tests were examined in terms of robustness and power. The result can be summarized below:

A. Type I Error

1. Normal Distribution

Under the normal distribution, equality of sample sizes, and homogeneity of variances, Bartlett's test is the best robust tests that can control type I error for all sample size in the nominal significance levels. The second and the third robust test are Cochran's test and Levene's test, respectively. Under the normal distribution, unequal sample sizes, and homogeneity of variances, all three statistical tests are robust.

2. Gamma Distribution

Under the non-normal distribution, equality of sample sizes, and homogeneity of variances, Levene's test is robust. It can control type I error with gamma distribution for sample size 15 in 0.01 significance level. While Bartlett's test

Copyright by Mahidol University

and Cochran's test cannot control type I error or non-robustness in gamma distribution. Under the non-normal distribution, unequal sample sizes, and homogeneity of variances, there is no robust test.

3. Chi-square Distribution

Under the non-normal distribution, equality of sample sizes, and homogeneity of variances, the best robust test is Cochran's test and Levene's test, respectively. Under the non-normal distribution, unequal sample sizes, and homogeneity of variances, there is no robust test.

4. Normal Distribution, Gamma Distribution and Chi-square Distribution

Under the non-normal distribution, equality of sample sizes, and homogeneity of variances, Bartlett's test is as good as Cochran's test for large sample sizes. Meanwhile, Levene's test is good for small sample sizes.

B. Power of Test

1. Normal Distribution

Under the normal distribution, equal sample sizes, and unequal sample variances, Bartlett's test is the best powerful test for variance ratio 1:2:4. Meanwhile, Cochran's test is good for one variance is larger. Under the normal distribution,

Copyright by Mahidol University

unequal sample sizes, and unequal sample variances, Cochran's test is the best powerful test. Lastly, Levene's test is the lowest of all cases.

2. Gamma Distribution

Under the non-normal distribution, equal sample sizes, and unequal sample variances, Bartlett's test is the best powerful test. The second and the third powerful test are Cochran's test and Levene's test, respectively. Under the non-normal distribution, unequal sample sizes, and unequal sample variances, Cochran's test is the best powerful test. The second and the third powerful test are Bartlett's test and Levene's test, respectively.

3. Chi-square Distribution

Under the non-normal distribution, equal sample sizes, and unequal sample variances, Cochran's test and Bartlett's test are powerful statistical tests. While Levene's test has the lowest power. Under the non-normal distribution, unequal sample sizes, and unequal sample variances, Cochran's test is the most powerful test. The second and the third powerful test are Bartlett's test and Levene's test, respectively.

4. Normal Distribution, Gamma Distribution and Chi-square Distribution

Under the non-normal distribution, equal sample sizes, and unequal sample variances, Bartlett's test is the best powerful test. The second and the third powerful test are Cochran's test and Levene's test, respectively.

Copyright by Mahidol University

C. Implication of the Study

When data is normally distributed, Bartlett's test is a good choice to use for testing homogeneity of variances since it is not affected by sample sizes. When data is non-normally distributed, Levene's test is a good choice for small equal sample sizes and equal variances since its probabilities of type I error values lies within Cochran limit. This means Levene's test is robust. If one variance is larger, the data are non-normally distributed, and are of unequal sample sizes, Cochran's test would be recommended, since it still gives high power.

D. Recommendation for Further Studies

From this study, the researcher has three recommendations for the next research as listed below:

1. To generate random number, it would be better to begin with 69069. Marsaglia (36), the beginning number 65539 should be replace with 69069 because 65539 was found to have a serious problem.
2. If the ratios of variances should differ by more than four or the ratio of the standard deviations differ by more than two. That must be concerned about heterogeneity of variance (25).
3. The sample sizes of each treatment class need to be equal and unequal. It may be bigger than 3 and is not equal to sample sizes from rule of thumb because sample sizes of rule of thumb are stable and the effect

Copyright by Mahidol University

of sample size may not be violated to statistical test, so we do not know that sample sizes affect the statistical test.

Copyright by Mahidol University

REFERENCES

1. Williams RBG. **Intermediate statistics for geographers and earth scientists.** Hong Kong: Macmillan education, 1986. 364-6.
2. Kingman A and Zion G. **Some power considerations when deciding to use transformations.** Statistics in Medicine 1994, 13: 769-83.
3. Winer BJ. **Statistical principles in experimental design.** 2nd ed. New York: McGraw-Hill, 1974.
4. Cochran WG. **Some consequences when the assumptions for the analysis of variance are not satisfied.** Biometics 1947, 3: 22-38
5. Glass GV and Stanley JC. **Statistical method in education and psychology.** Englewood, California: Prentice-Hall, 1970.
6. Gartside PS. **A study of methods for comparing several variances.** JASA 1972, June; 67(338): 342-6.
7. Brown MB and Forsythe AB. **Robust tests for the equality of variances.** JASA 1974, June; 69: 364-7.
8. Neter J, Kutner MH, Nachtsheim CJ and Wasserman W. **Applied linear regression model.** 3rd ed. Illinois: Irwin, 1996.
9. Vijitrthummaros P. **Learning from the analysis of data by using SPSS for Windows: Levene Test.** Thai journal of development administration 1997, Oct-Dec, 37(4): 21-8.

Copyright by Mahidol University

10. Dixon WJ, (ed). **BMDP Statistical software manual volume 1.** Los Angeles: BMDP Statistical Software, 1992.

11. Gustafson TL. **True EPISTAT manual.** 3rd edition. Texas: Epistat Services: 1989.

12. Dean AG. **Epi Info version 6.** Atlanta, Georgia. Centers for Disease Control and Prevention (CDC), 1994

13. Filliben JJ. and Heckert A. **Levene's test for equality of variances.** [Online] 2000; Available from: http://www.nist.gov/itl/div898/handbook/eda/section3/eda358.htm [Accessed 2000 Jul 12].

14. **Small sample size calculation [Online].** Available from: http:// www.isixsigma.com: library/conent/c000709smsample.asp [Accessed 2001 May 18].

15. Conover WJ, Johnson EM, and Johnson MM. **A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data.** Technometrics 1981,23(4): 351-361.

16. Vogt WP. **Dictionary of statistics and methodology: a nontechnical guide for the social sciences.** California: Sage Publications, 1993.

17. David ML. **Hyperstat online textbook.** Glossary contents [Online]. 1993-2001. Available from: http://www.davidmlane.com/hypestat/glossary_contents.html#j [Accessed 2001 July 10].

18. Polhemus NW. **Statlets™ user manual.** Glossary [Online]. NWP Associates. Englewood, California: 13 March 2001. Available from: http://www.statlets.com/usermanual/glossary.htm#sectB. [Accessed 2001 July 10].

Copyright by Mahidol University

19. Filliben JJ and Heckert A. **Bartlett's test.** [Online] 2000;[4]. Available from: URL: http:// www.nist.gov/itl/div898/handbook/eda/section3/eda357.htm [Accessed 2000 Jul 12].

20. Montgomery DC. **Design and analysis of experiments.** 4th ed. New York: John Wiley & Sons, 1997.

21. Snedecor GW and Cochran WG. **Statistical methods.** IOWA: The IOWA State University, 1967.

22. Armitage P. and Berry G. **Statistical methods in medical research.** 2nd ed. London: Blackwell Scientific Publications, 1988.

23. Box G.E.P. **Nonnormality and tests on variances.** Biometrika 1953, 40: 318-35.

24. SPSS. **SPSS for windows release 10.0.1 [Computer program].** SPSS 1999.

25. Phil E. **Checking assumptions.** Education 230 B/C Linear Statistical Models. [Online]. Graduate School of Education & Information Studies; 4 Jun 99. Available from: http://www.gseis.ucla.edu/courses/ed230bc1/cnotes1 [Accessed 2000 Jul 13].

26. Keith M. **SBI 209 design & analysis of biological studies information and resources.** [Online]. March 14, 2000. Available from: http:// www.ntu.edu.au /faculties/ science/ sbes/ resources/SBI209/SBI209-4-6.pdf [Accessed 2001 Sep 6].

27. Maass EC. **Advanced statistics & optimization.** [Online]. Available from:http: // www.geocities.com/ Research Triangle/ System/ 3737/ [Accessed 2001 Oct 11].

Copyright by Mahidol University

28. Johnson P. **Estimation of the common mean, μ, and a 100(1 - α)% confidence interval for μ.** [Online]. March, 1998. Available from: http: //pages.prodigy.net/johnsonp12/mandel.html [Accessed 2000 Jul 20].

29. Levy JK. **An empirical comparison of several multiple range tests for variances.** JASA 1975,70(39): 180-183.

30. Kyle S. Special distributions. **The gamma distribution.** [Online]. Huntsville: 1997, updated 27 Sept. 2001. Available from: http:// www.mathuah.edu/stat/special/special3.html [Accessed 2001 May 29].

31. Bodhisuwan W. **A comparison of the test statistics for homogeneity of variances.** [M.S. Thesis in statistics]. Bangkok: Faculty of Graduate Studies, Chulalongkorn University, 1991.

32. Kulanoot A. **A comparison of the power of test for homogeneity of variances using three types of test statistics.** [M.S. Thesis in statistics]. Bangkok: Faculty of statistics, Chulalongkorn University; 1985.

33. Peechawanich V. **Probability theory and the applications.** Bangkok: Prakyprueg, 1992.

34. Chansatitporn N. **The comparison of the power of ANOVA F-Test and Kruskal-Wallis Test under violation of assumptions.** [M.S. Thesis in Biostatistics]. Bangkok: Faculty of Graduate Studies, Mahidol University; 1993.

35. Srisunsanee J. **Effects of failure to meet assumptions of homogeneity of variance and normal on type I error in one-way analysis of variance.** [M.S. Thesis in statistics]. Bangkok: Faculty of Graduate Studies, Kasetsart University; 1998.

Copyright by Mahidol University

36. Karlen D. Part III: Monte carlo methods. **Random number generation.** [Online]. Ottawa: 1995. Available from: http:// www.physics.carleton.ca /courses/75.502/slides/monte21/p015.html. [Accessed 2002 Aug 21].

Copyright by Mahidol University

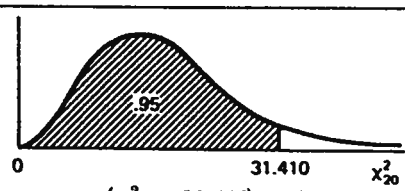
LIBRARY AND INFORMATION CENTER
MAHIDOL UNIVERSITY

APPENDIX A

STATISTICAL TABLES

Copyright by Mahidol University

TABLE 1. Percentiles of the Chi-square Distribution



$p(\chi^2_{20} \leq 31.410) = .95$

d.f.	$\chi^2_{.005}$	$\chi^2_{.025}$	$\chi^2_{.05}$	$\chi^2_{.90}$	$\chi^2_{.95}$	$\chi^2_{.975}$	$\chi^2_{.99}$	$\chi^2_{.995}$
1	.0000393	.000982	.00393	2.706	3.841	5.024	6.635	7.879
2	.0100	.0506	.103	4.605	5.991	7.378	9.210	10.597
3	.0717	.216	.352	6.251	7.815	9.348	11.345	12.838
4	.207	.484	.711	7.779	9.488	11.143	13.277	14.860
5	.412	.831	1.145	9.236	11.070	12.832	15.086	16.750
6	.676	1.237	1.635	10.645	12.592	14.449	16.812	18.548
7	.989	1.690	2.167	12.017	14.067	16.013	18.475	20.278
8	1.344	2.180	2.733	13.362	15.507	17.535	20.090	21.955
9	1.735	2.700	3.325	14.684	16.919	19.023	21.666	23.589
10	2.156	3.247	3.940	15.987	18.307	20.483	23.209	25.188
11	2.603	3.816	4.575	17.275	19.675	21.920	24.725	26.757
12	3.074	4.404	5.226	18.549	21.026	23.336	26.217	28.300
13	3.565	5.009	5.892	19.812	22.362	24.736	27.688	29.819
14	4.075	5.629	6.571	21.064	23.685	26.119	29.141	31.319
15	4.601	6.262	7.261	22.307	24.996	27.488	30.578	32.801
16	5.142	6.908	7.962	23.542	26.296	28.845	32.000	34.267
17	5.697	7.564	8.672	24.769	27.587	30.191	33.409	35.718
18	6.265	8.231	9.390	25.989	28.869	31.526	34.805	37.156
19	6.844	8.907	10.117	27.204	30.144	32.852	36.191	38.582
20	7.434	9.591	10.851	28.412	31.410	34.170	37.566	39.997
21	8.034	10.283	11.591	29.615	32.671	35.479	38.932	41.401
22	8.643	10.982	12.338	30.813	33.924	36.781	40.289	42.796
23	9.260	11.688	13.091	32.007	35.172	38.076	41.638	44.181
24	9.886	12.401	13.848	33.196	36.415	39.364	42.980	45.558
25	10.520	13.120	14.611	34.382	37.652	40.646	44.314	46.928
26	11.160	13.844	15.379	35.563	38.885	41.923	45.642	48.290
27	11.808	14.573	16.151	36.741	40.113	43.194	46.963	49.645
28	12.461	15.308	16.928	37.916	41.337	44.461	48.278	50.993
29	13.121	16.047	17.708	39.087	42.557	45.722	49.588	52.336
30	13.787	16.791	18.493	40.256	43.773	46.979	50.892	53.672
35	17.192	20.569	22.465	46.059	49.802	53.203	57.342	60.275
40	20.707	24.433	26.509	51.805	55.758	59.342	63.691	66.766
45	24.311	28.366	30.612	57.505	61.656	65.410	69.957	73.166
50	27.991	32.357	34.764	63.167	67.505	71.420	76.154	79.490
60	35.535	40.482	43.188	74.397	79.082	83.298	88.379	91.952
70	43.275	48.758	51.739	85.527	90.531	95.023	100.425	104.215
80	51.172	57.153	60.391	96.578	101.879	106.629	112.329	116.321
90	59.196	65.647	69.126	107.565	113.145	118.136	124.116	128.299
100	67.328	74.222	77.929	118.498	124.342	129.561	135.807	140.169

Copyright by Mahidol University

TABLE 2. Percentiles of the *F* Distribution

$F_{.99}$

Denominator Degrees of Freedom	Numerator Degrees of Freedom								
	1	2	3	4	5	6	7	8	9
1	4052	4999.5	5403	5625	5764	5859	5928	5981	6022
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41

Copyright by Mahidol University

TABLE 2. (cont.)

Denominator Degrees of Freedom	Numerator Degrees of Freedom									
	10	12	15	20	24	30	40	60	120	∞
1	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

Copyright by Mahidol University

TABLE 2. (cont.)

$F_{.95}$

Denominator Degrees of Freedom	Numerator Degrees of Freedom								
	1	2	3	4	5	6	7	8	9
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

Copyright by Mahidol University

TABLE 2. (cont.)

Denominator Degrees of Freedom	Numerator Degrees of Freedom									
	10	12	15	20	24	30	40	60	120	∞
1	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

Copyright by Mahidol University

TABLE 3. Critical Values for Cochran's Test

Values given are for the statistic (largest s^2)/(Σs_i^2), where each of the k values of s^2 has ν degrees of freedom.

PERCENTILE 95

k \ ν	1	2	3	4	5	6	7	8	9	10	16	36	144	∞
2	0.9985	0.9750	0.9392	0.9057	0.8772	0.8534	0.8332	0.8159	0.8010	0.7880	0.7341	0.6602	0.5813	0.5000
3	0.9669	0.8709	0.7977	0.7457	0.7071	0.6771	0.6530	0.6333	0.6167	0.6025	0.5466	0.4748	0.4031	0.3333
4	0.9065	0.7679	0.6841	0.6287	0.5895	0.5598	0.5365	0.5175	0.5017	0.4884	0.4366	0.3720	0.3093	0.2500
5	0.8412	0.6838	0.5981	0.5441	0.5065	0.4783	0.4564	0.4387	0.4241	0.4118	0.3645	0.3066	0.2513	0.2000
6	0.7808	0.6161	0.5321	0.4803	0.4447	0.4184	0.3980	0.3817	0.3682	0.3568	0.3135	0.2612	0.2119	0.1667
7	0.7271	0.5612	0.4800	0.4307	0.3974	0.3726	0.3535	0.3384	0.3259	0.3154	0.2756	0.2278	0.1833	0.1429
8	0.6798	0.5157	0.4377	0.3910	0.3595	0.3362	0.3185	0.3043	0.2926	0.2829	0.2462	0.2022	0.1616	0.1250
9	0.6385	0.4775	0.4027	0.3584	0.3286	0.3067	0.2901	0.2768	0.2659	0.2568	0.2226	0.1820	0.1446	0.1111
10	0.6020	0.4450	0.3733	0.3311	0.3029	0.2823	0.2666	0.2541	0.2439	0.2353	0.2032	0.1655	0.1308	0.1000
12	0.5410	0.3924	0.3264	0.2880	0.2624	0.2439	0.2299	0.2187	0.2098	0.2020	0.1737	0.1403	0.1100	0.0833
15	0.4709	0.3346	0.2758	0.2419	0.2195	0.2034	0.1911	0.1815	0.1736	0.1671	0.1429	0.1144	0.0889	0.0667
20	0.3894	0.2705	0.2205	0.1921	0.1735	0.1602	0.1501	0.1422	0.1357	0.1303	0.1108	0.0879	0.0675	0.0500
24	0.3434	0.2354	0.1907	0.1656	0.1493	0.1374	0.1286	0.1216	0.1160	0.1113	0.0942	0.0743	0.0567	0.0417
30	0.2929	0.1980	0.1593	0.1377	0.1237	0.1137	0.1061	0.1002	0.0958	0.0921	0.0771	0.0604	0.0457	0.0333
40	0.2370	0.1576	0.1259	0.1082	0.0968	0.0887	0.0827	0.0780	0.0745	0.0713	0.0595	0.0462	0.0347	0.0250
60	0.1737	0.1131	0.0895	0.0765	0.0682	0.0623	0.0583	0.0552	0.0520	0.0497	0.0411	0.0316	0.0234	0.0167
120	0.0998	0.0632	0.0495	0.0419	0.0371	0.0337	0.0312	0.0292	0.0279	0.0266	0.0218	0.0165	0.0120	0.0083
∞	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Copyright by Mahidol University

TABLE 3. (cont.)

PERCENTILE 99

k \ ν	1	2	3	4	5	6	7	8	9	10	16	36	144	∞
2	0.9999	0.9950	0.9794	0.9586	0.9373	0.9172	0.8988	0.8823	0.8674	0.8539	0.7949	0.7067	0.6062	0.5000
3	0.9933	0.9423	0.8831	0.8335	0.7933	0.7606	0.7335	0.7107	0.6912	0.6743	0.6059	0.5153	0.4230	0.3333
4	0.9676	0.8643	0.7814	0.7212	0.6761	0.6410	0.6129	0.5897	0.5702	0.5536	0.4884	0.4057	0.3251	0.2500
5	0.9279	0.7885	0.6957	0.6329	0.5875	0.5531	0.5259	0.5037	0.4854	0.4697	0.4094	0.3351	0.2644	0.2000
6	0.8828	0.7218	0.6258	0.5635	0.5195	0.4866	0.4608	0.4401	0.4229	0.4084	0.3529	0.2858	0.2229	0.1667
7	0.8376	0.6644	0.5685	0.5080	0.4659	0.4347	0.4105	0.3911	0.3751	0.3616	0.3105	0.2494	0.1929	0.1429
8	0.7945	0.6152	0.5209	0.4627	0.4226	0.3932	0.3704	0.3522	0.3373	0.3248	0.2779	0.2214	0.1700	0.1250
9	0.7544	0.5727	0.4810	0.4251	0.3870	0.3592	0.3378	0.3207	0.3067	0.2950	0.2514	0.1992	0.1521	0.1111
10	0.7175	0.5358	0.4469	0.3934	0.3572	0.3308	0.3106	0.2945	0.2813	0.2704	0.2297	0.1811	0.1376	0.1000
12	0.6528	0.4751	0.3919	0.3428	0.3099	0.2861	0.2680	0.2535	0.2419	0.2320	0.1961	0.1535	0.1157	0.0833
15	0.5747	0.4069	0.3317	0.2882	0.2593	0.2386	0.2228	0.2104	0.2002	0.1918	0.1612	0.1251	0.0934	0.0667
20	0.4799	0.3297	0.2654	0.2288	0.2048	0.1877	0.1748	0.1646	0.1567	0.1501	0.1248	0.0960	0.0709	0.0500
24	0.4247	0.2871	0.2295	0.1970	0.1759	0.1608	0.1495	0.1406	0.1338	0.1283	0.1060	0.0810	0.0595	0.0417
30	0.3632	0.2412	0.1913	0.1635	0.1454	0.1327	0.1232	0.1157	0.1100	0.1054	0.0867	0.0658	0.0480	0.0333
40	0.2940	0.1915	0.1508	0.1281	0.1135	0.1033	0.0957	0.0898	0.0853	0.0816	0.0668	0.0503	0.0363	0.0250
60	0.2151	0.1371	0.1069	0.0902	0.0796	0.0722	0.0668	0.0625	0.0594	0.0567	0.0461	0.0344	0.0245	0.0167
120	0.1225	0.0759	0.0585	0.0489	0.0429	0.0387	0.0357	0.0334	0.0316	0.0302	0.0242	0.0178	0.0125	0.0083
∞	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Copyright by Mahidol University

APPENDIX B

FORTRAN PROGRAM TO GENERATE DATA FOR A COMPARISON ON TYPE I ERROR AND POWER OF TEST IN NORMAL DISTRIBUTION

Copyright by Mahidol University

```
dimension r1(15),r2(15),r3(15),z1(15),z2(15),z3(15),&
&         amed(3),p1(15),p2(15),p3(15),dsum(3),ssum(3),s(3),N(3)
integer   iseed,nout


      PARAMETER  (N1=15,N2=15,N3=15)
      real,parameter :: le5=3.222,le1=5.16,cre01=.6287, cre05 =
                        0.5652,chi01=9.21,chi05=5.99
      REAL       RB1(n1),rb2(n2),RB3(N3),l
external rnnoa,rnset,umach

N(1) = 15
N(2) = 15
N(3) = 15
lre5 = 0
lac5 = 0
lre1 = 0
lac1 = 0
cac5 = 0
cac1 = 0
cre5 = 0
cre1 = 0
bac5 = 0
bac1 = 0
bre5 = 0
bre1 = 0
iseed = 65539
!nr1 = 15
!nr2 = 15
!nr3 = 15
xm = 10
xstd = 5.1
                        call umach (2,nout)
                        call rnset(iseed)

do 30 k = 1,1000

                        CALL RNNOA (N1, R1)
                        CALL SSCAL (N1, XSTD, R1, 1)
                        CALL SADD (N1, XM, R1, 1)
                        CALL RNNOA (N2, R2)
                        CALL SSCAL (N2, XSTD, R2, 1)
                        CALL SADD (N2, XM, R2, 1)
                        CALL RNNOA (N3, R3)
                        CALL SSCAL (N3, XSTD, R3, 1)
                        CALL SADD (N3, XM, R3, 1)
```

Copyright by Mahidol University

```
f1 = 0
f2 = 0
f3 = 0
p = 0
f = 0
ntotal = 0
pp1 = 0
pp2 = 0
pp3 = 0
cochran = 0.0
coc = 0.0
ss = 0
ts = 0
tss = 0
np = 3
t = 0
s(1) = 0.0
s(2) = 0.0
s(3) = 0.0
dsum(1) = 0.
dsum(2) = 0.
dsum(3) = 0.
ssum(1) = 0.
ssum(2) = 0.
ssum(3) = 0.

! ---------- Bartlett ----------
aa = 0
bb = 0
ti = 0
                    CALL SVRGN (N1, R1, RB1)
                    CALL SVRGN (N2, R2, RB2)
                    CALL SVRGN (N3, R3, RB3)

                    amed(1) = RB1((n1+1)/2)
                    amed(2) = RB2((n2+1)/2)
                    amed(3) = RB3((n3+1)/2)

                        open (1,file = 'data1.txt',action = 'write')
                            do i = 1,n1
                                Z1(i) = abs(r1(i)-amed(1))
                                 f1   = f1+Z1(i)
! ---------- Cochran ----------
                                dsum(1)  = dsum(1) + r1(i)
                                ssum(1) = ssum(1) + (r1(i)**2)
                            end do
! ---------- Bartlett ----------
                    zme1 = f1/n1
```

Copyright by Mahidol University

```
                    do 11 I = 1,n1
                        p1(i) = (z1(i)-zme1)**2
                        pp1   = pp1+p1(i)
11                  continue
! *********************  N2  ************************
                        do i = 1,n2
                            Z2(i) = abs(r2(i)-amed(2))
                             f2   = f2+Z2(i)
! ------------ Cochran -----------
                            dsum(2)  = dsum(2) + r2(i)
                            ssum(2) = ssum(2) + (r2(i)**2)
                        end do
              zme2 = f2/n2
                    do 22 i = 1,n2
                        p2(i) = (z2(i)-zme2)**2
                        pp2 = pp2+p2(i)
22                  continue
! *********************  N3  ************************
                        do i = 1,n3
                            Z3(i) = abs(r3(i)-amed(3))
                             f3   = f3+Z3(i)
! ------------ Cochran -----------
                            dsum(3)  = dsum(3) + r3(i)
                            ssum(3) = ssum(3) + (r3(i)**2)

                        end do
              zme3 = f3/n3
                    do 33 i = 1,n3
                        p3(i) = (z3(i)-zme3)**2
                        pp3 = pp3+p3(i)
33                  continue
! ***************************************************************
              Ntotal = n1+n2+n3
              Zmean  = (f1+f2+f3)/ntotal
              p      = pp1+pp2+pp3
              f      = n1*(zme1-zmean)**2+n2*(zme2-zmean)**2+n3*(zme3-
              zmean)**2
              l    = ((ntotal-3)*f)/(2*p)
                    if (l > le5) then
                        lre5 = lre5+1
                    else
                        lac5 = lac5+1
                    end if
                    if (l > le1) then
                        lre1 =lre1+1
                    else
                        lac1 = lac1+1
                    end if
```

Copyright by Mahidol University

```
! ---------- Cochran ----------
                DO 13 i= 1,3
                    s(i) = (ssum(i)-((dsum(i)**2)/n(i)))/(n(i)-1)
                    ts = ts + s(i)
! ---------- Bartlett ----------
                    tss = tss + (n(i)-1)*s(i)
                    bb = bb + ((n(i)-1)*log10(s(i)))
                    ti = ti + (1.0/(n(i)-1)- 1.0/(ntotal-3))
13 CONTINUE
                    stot = tss/(ntotal-3)
! ---------- Cochran ----------
                do 50 i = 1,3
                        aa = aa + ((n(i)-1)*log10(stot))
                    s(i-1) = ss
                    if (s(i) < s(i-1))        then
                        ss = s(i-1)
                    else
                        ss = s(i)
                    end if
50  continue
                    coc = ss/ts
                if (coc > cre01) THEN
                    cre1 = cre1 +1
                ELSE
                    cac1 = cac1 +1
                END IF

                if (coc > cre05) THEN
                    cre5 = cre5 +1
                ELSE
                    cac5 = cac5 +1
                END IF

! ---------- Bartlett ----------

                bar = (aa-bb)*2.3026/(1.0 + (ti/(3*(3-1))))
                if (bar > chi01) THEN
                    bre1 = bre1+1
                ELSE
                    bac1 = bac1 +1
                ENDIF

                if (bar > chi05) THEN
                    bre5 = bre5+1
                ELSE
                    bac5 = bac5 +1
                ENDIF
```

Copyright by Mahidol University

```
30 continue
                write (1,644) lre5,lac5,lre1,lac1
644 format ("********* Normal Distribution with 15 sample sizes and variances
                ratio = 1:1:1 *********"//&
&"                                        AT 0.05          AT 0.01    "/&
&'                                    reject    accept      reject     accept'/&
&"levene's test     ",i5,10x,i5,10x,i5,10x,i5)
write (1,664) cre5,cac5,cre1,cac1
664 format (/"COCHARN'S TEST     ",f5.0,10x,f5.0,10x,f5.0,10x,f5.0)
write (1,665) bre5,bac5,bre1,bac1
665 format (/"BARTLETT'S TEST    ",f5.0,10x,f5.0,10x,f5.0,10x,f5.0)

end
```

SUBROUTINE FOR GENERATE GAMMA DISTRIBUTION

```
A   = 2.67
B   = 0.75

    CALL RNGAM (N1, A, R1)
         call sscal (N1, b, R1,1)

    CALL RNGAM (N2, A, R2)
         call sscal (N2, b, R2,1)

    CALL RNGAM (N3, A, R3)
         call sscal (N3, b, R3,1)
```

SUBROUTINE FOR GENERATE CHI-SQUARE DISTRIBUTION

```
    CALL RNCHI (N1, DF, R1)

    CALL RNCHI (N2, DF, R2)

    CALL RNCHI (N3, DF, R3)
```

Copyright by Mahidol University

AND INFORMATION CENTER
MAHIDOL UNIVERSITY

BIOGRAPHY

NAME	Ms. Sineenart Taejaroenkul
DATE OF BIRTH	11 July 1969
PLACE OF BIRTH	Chiangmai, Thailand
INSTITUTIONS ATTEDNDED	Chiangmai University, 1989-1993: Bachelor of Science (Statistics) Mahidol University, 1997-2002: Master of Science (Biostatistics)
POSITION & OFFICE	1993-Present, Research Institute for Health Science, Chiangmai University, Chiangmai, Thailand. Position : Statistician.

Copyright by Mahidol University