University of Alberta

SEQUENTIAL TESTS FOR MONITORING PARAMETERS OF A NESTED RANDOM EFFECTS MODEL

by

Eshetu Getachew Atenafu ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of **Master of Science**.

in

Statistics

Department of Mathematical Sciences

Edmonton, Alberta

Fall 2000

0-612-59775-X

University of Alberta

Library Release Form

Name of Author: Eshetu Getachew Atenafu

Title of Thesis: Sequential Tests for Monitoring Parameters of a Nested Random Effects Model

Degree: Master of Science

Year this Degree Granted: 2000

Eshetu Getachew Atenafu
CAB 632
Mathematical Sciences Department
University of Alberta
Edmonton, AB
T6G 2G1, Canada

Date: Aug. 21, 2000

"The earlier theoretical developments, due to Neyman and Pearson, let the statistician into the scientific experiment at the beginning (planning stage) and at the end (evaluation stage), but the new sequential approach requires statistics during the entire experiment, thus it becomes an integral part of the whole process."

Barnard, G.A.

Abstract

Data with multiple sources of variability are common in statistics. They arise in various areas of statistical applications including quality control, clinical trials, and other related fields. Any monitoring system in such situations will need to account not only for the within-item variability but the item-to-item variability as well. Yashchin(1994, 1995), using a nested random effects model, developed a monitoring scheme for the mean and variance parameters of the model based on cumulative sums and the likelihood ratio approach with a tolerance level. His method was illustrated with data on manufacturing of integrated circuits.

The objective in this thesis is to present an alternative monitoring scheme for the parameters of the same model. The method is built on sequential testing of composite hypotheses, Gombay(2000b). Specifically, the approach is based on a truncated sequential testing without tolerance level with the help of generalized likelihood ratio. Strong points of this method are also compared to those of competing methods via simulation. The method will be illustrated using the integrated circuit data of Yashchin(1995).

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Sequential Tests for Monitoring Parameters of a Nested Random Effects Model** submitted by Eshetu Getachew Atenafu in partial fulfillment of the requirements for the degree of **Master of Science** in *Statistics*.

. . . .
Dr. Byron Schmuland   Committee Chair

. . . .
Dr. Edit Gombay   Supervisor

. . . .
Dr. Ming J. Zuo   Committee Member

Date: August 16 2000

I dedicate this thesis to :

unknown of his being alive or not, Getachew Atenafu (**father**)

the late **mother** Tera Waktolo and

the late **brother** Teshome Getachew

Acknowledgements

Thanks to the Almighty God, author of my being, source of my might, strength, and wisdom; with whom "nothing is impossible," who has made all things possible in my life and brought me to a successful completion of this program.

I am very grateful to my supervisor, Dr. Edit Gombay, for leading me to the research field of this thesis, and for her many helpful and thoughtful comments, discussions and suggestions throughout the preparation of this thesis. Without her guidance, this thesis would have been impossible.

Thanks also go to the Department of Mathematical Sciences, University of Alberta, for providing me the opportunity to study here and offering me Teaching/Research Assistantship during my stay and specially to Marion Benedict for her kindness and patience, I enjoyed during my stay. I would like to express my appreciation to the faculty members of the department who provide me with many interesting courses in Statistics and Mathematics. Special thanks go to my Teaching Mentor Dr. Giseon Heo, for her unreserved guidance in the process of Teaching Assistantship, to all my friends and class mates, in particular Endawoke Yizengaw, Alemayehu Berga, Eshetu Tesfaye, Abebaw Belay, Lucio Beyere, Alex De Leon and Abdulkadir Hussein for their unpayable help.

I am also greatly indebted to my Edmonton family, Girma Abera, Tigist Mekonnen, Girma Hailu and Fatuma Ousman who welcomed me, gave me their time and encourage me in all aspects of life. I would also like to thank the Church Fathers and parishioners of Debre Selam Medhane Alem Orthodox Tewahedo Church for nourishing my faith all these years, to Dr. Alebachew, Dr. Shiferaw, Dr. Atakilty, Dr. Getu, Wubishet, Mesfin, Beniam, Yohannes, Negash, Abebe, Lidetu and their families, Aklilu, Bereket, Molla, Getye, Elsa and Meseret for sharing me their experience. Without them life could have been more difficult.

I express my gratitude to my thesis committee members: Dr. Byron Schmuland and Dr. Ming J. Zuo, for the time they spent in reading my thesis.

Last but not least, my sincerest gratitude to my family Atsede Getachew, Solomon Getachew, Samson Getachew, Firehiwot Getachew, Medhin Gebre Egzi, Kassahun Tewolde Berhan, Tigist Tewolde Berhan, Tesfahun Tewolde Berhan, Yenenesh Mammo, Mesfin Tadesse, Derib Tadesse, Assefa Mekonnen, Ejigayehu Demissie, Tigist Yilma for their prayer and encouragement. I am greatly indebted to Tirfe Demissie, Wakshum Mekonnen, Seble Deribe, Seblework Tayachew and their families for support, prayers, encouragement and taking care of my family affairs all these years that I have been out here.

To you all and numerous others that I have not mentioned, I say
THANK YOU AND GOD BLESS YOU.
"AMESEGINALEHU, EGZIBHER YIBARKACHIHU::"

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Nested Random Effects is a model which has applications in various fields. One such is in statistical process control, where a sequence of measurements are used to monitor the level of parameters that are indicators of quality. Yashchin (1994, 1995) considered this model and discussed the use of cumulative sums and likelihood ratios respectively, to check whether IBM's manufacturing of chips is conforming to some prescribed target. Other application include monitoring of measurements describing the health of patients in medical trials, where due to ethical considerations sequential procedures have to be implemented.

The analysis on the sequential procedure is a method of statistical inference whose distinguishing feature is that the number of observations required by the procedure is not determined in advance of an experiment. The decision to terminate the experiment depends, at each stage, on the results of observations previously made.

The organization of this thesis is as follows. In this Chapter, we review and discuss the model and give the likelihood approach test. In Chapter Two, review and introduction of the sequential likelihood ratio test will be presented. It also gives the conditions and results of the new test procedure. In Chapter Three, algorithms will developed for monitoring each of the four parameters individually, and for the three variance components simultaneously. In Chapter

Four, the result of a simulation study and comparison with those of competing methods will be presented. In Chapter Five, the applications of the new method on the published data of Yashchin (1995) is laidout followed by Discussion and Conclusion. The Appendicies contain the algorithm of solving cubic equation analytically, approximations of the test under the alternative hypothesis of change, which allow power approximations for any specific alternative. Splus code for the simulation is also included in this part.

The next sections on this Chapter will give a brief description of a two-factor Nested Random Effects Model, introduce the generalized likelihood ratio of Neyman and Pearson for a given composite hypotheses, and give the ANOVA table to check the significance of each variance.

1.1 Two-factor Nested Design

In an ordinary factorial study, in which every level of one factor appears with each level of every other factor, the factors are said to be crossed. A different situation occurs when factors are nested. Consider the following situation to see the difference.

A large manufacturing company operates three regional training schools for mechanics, one in each of its operating districts. Each school has two instructors each of whom teaches classes of about 15 mechanics in three-week sessions. The company was concerned about the effect of the schools (factor A) and instructors (factor B) on the learning achieved. Here the study is not an ordinary two-factor study. The reason is that the instructor in one school did not teach in the other schools. The experimental design for these training schools involves an incomplete factorial arrangement of a special type, in which each level of factor B (instructor) occurs with only one level of factor A (school). Factor B is therefore said to be nested within factor A.

In certain multifactor experiments the levels of one factor, say factor A, are similar but not identical for different levels of another factor, say factor B. Such an arrangement is called a nested or hierarchical design, with levels of factor B nested under the levels of factor A. For example, consider a company that purchases its raw material from three different suppliers (refer to the book of Montgomery (1997), Design of Experiments for further detail). The company wishes to determine if the purity of the raw material is the same from each supplier. There are four batches of raw material available from each supplier, and three determinations of purity are to be taken from each batch. Since every level of factor B does not appear with every level of factor A, there can be no interaction between A and B. This is a two-stage design, with batches nested under suppliers. The company may wish to determine if the variability in purity is attributable to differences between suppliers. The company may also wish to determine the quality of the raw materials knowing there is a variability within each supplier etc.

In a two-factor nested design we consider nested designs involving two factors, one of which is nested within the other. For consistency, consider the case in which **b** of factor B is nested within each factor A, say **a** factors, where both factors are random.

Model

Let X_{ijk} denote the k^{th} observation when factor A is at i^{th} level and factor B at the j^{th} level. Assume also that there are **N** observations for factor combination, i.e., k=1,2,...,**N** and that i=1,2,...,**a**, j=1,2,...,**b**. The observations have the following structure:

$$X_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \epsilon_{ijk}$$

where μ is a constant (grand mean), α_i is the i^{th} random effect of factor A, $\beta_{j(i)}$ is the j^{th} random effect of factor B, nested under i^{th} factor A, and ϵ_{ijk} is the random error. Here **a** is the number of levels of factor A, **b** is the number of levels of factor B nested in each factor A and **N** is the number of observations under each factor B nested in factor A.

Assumptions

$$\alpha_i \sim N(0, \sigma_A^2) \quad \beta_{j(i)} \sim N(0, \sigma_B^2) \quad and \quad \epsilon_{ijk} \sim N(0, \sigma^2) \ , \qquad (1.1.1)$$

and all three are uncorrelated.

1.2 Description of a Test Based on the Likelihood Ratio

Let $X_1, X_2, ..., X_n$ be a sequence of independent, identically distributed random variables with parameters:
$\theta \in \Omega_1 \subset R^d$, $\eta \in \Omega_2 \subset R^p$, $d \geq 1$, $p \geq 0$, and $\Omega = \Omega_1 \; X \; \Omega_2$, where θ is the parameter vector of interest, η is the nuisance parameter vector, d and p are integers. For the hypotheses:

$$H_o : \theta = \theta_o \; , \quad \eta \in \Omega_2 \quad vs. \quad H_1 : \theta \neq \theta_o \; , \quad \eta \in \Omega_2 \tag{1.2.1}$$

the generalized likelihood ratio test statistic of Neyman and Pearson (1928) is given by:

$$\lambda_n = \frac{\sup_{\eta \in \Omega_2} \prod_{i=1}^{n} f(x_i; \theta_o, \eta)}{\sup_{\eta \in \Omega_2 \;\; \theta \in \Omega_1} \prod_{i=1}^{n} f(x_i; \theta, \eta)} \tag{1.2.2}$$

Yashchin (1995) used the likelihood ratio in various forms to monitor change in the variance components, treating all other parameters of the model as nuisance. A different generalized likelihood ratio is proposed in this study. The theoretical background to the new procedures was developed in Gombay (1996, 1997, 2000a), where the behavior of the generalized likelihood ratio process was analysed in the presence of nuisance parameters. The implementation of them for sequential testing is described in Gombay (2000b), where it was also shown that Wald's method of sequential testing cannot, in general, be extended to situations, where nuisance parameters are present. Hence we cannot hope to control both types of error simultaneously in the present problem as it was done in Wald's (1947) scheme. Instead, we will use Neyman-Pearson type tests where α , the level of significance (Type I error) is under control. In the current study α denotes the probability of stopping the process unnecessarily.

1.3 The Model and Likelihood Ratio Test

The model was given by:

$$X_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \epsilon_{ijk} \qquad \begin{cases} i = 1, 2, ..., \mathbf{a} \\ j = 1, 2, ..., \mathbf{b} \\ k = 1, 2, ..., \mathbf{N} \end{cases}$$

and from the assumptions above in (1.1.1):

<u>**for the Mean**</u>

$$\begin{aligned} E[X_{ijk}] &= E[\mu + \alpha_i + \beta_{j(i)} + \epsilon_{ijk}] \\ &= \mu + E[\alpha_i] + E[\beta_{j(i)}] + E[\epsilon_{ijk}] \\ &= \mu , \end{aligned} \qquad (1.3.1)$$

<u>**and for the Variance**</u>

$$\begin{aligned} Var(X_{ijk}) &= Var(\mu + \alpha_i + \beta_{j(i)} + \epsilon_{ijk}) \\ &= Var(\alpha_i) + Var(\beta_{j(i)}) + Var(\epsilon_{ijk}) \\ &= \sigma_A^2 + \sigma_B^2 + \sigma^2 \end{aligned} \qquad (1.3.2)$$

Put $\sigma_*^2 = \sigma_A^2 + \sigma_B^2 + \sigma^2$

Therefore, it will be appropriate to deduce that:

$$X_{ijk} \sim N(\mu, \sigma_*^2) \qquad (1.3.3)$$

In this case, our parameter of interest is $\theta = \mu$ and the nuisance parameters are, σ_A^2, σ_B^2, σ^2. The hypotheses to be tested are:

$$H_o : \mu = \mu_o \ , \ \sigma_*^2 > 0 \quad vs.$$

$$H_1 : \mu \neq \mu_o \ , \ \sigma_*^2 > 0$$

1.3.1 Analysis of Variance, ANOVA

The analysis of variance for the above model can be obtained by decomposing the total deviation $(X_{ijk} - \overline{X}_{...})$ as follows. The complete analysis and construction of ANOVA can be found in Montgomery (1997).

$$(X_{ijk} - \overline{X}_{...}) = (\overline{X}_{i..} - \overline{X}_{...}) + (\overline{X}_{ij.} - \overline{X}_{i..}) + (X_{ijk} - \overline{X}_{ij.}) \qquad (1.3.4)$$

where the expressions on the right hand side of equation (1.3.4) gives:
$(X_{ijk} - \overline{X}_{...})$ the total deviation
$(\overline{X}_{i..} - \overline{X}_{...})$ A's main random effect
$(\overline{X}_{ij.} - \overline{X}_{i..})$ Specified B's random effect when A at i^{th} level and
$(X_{ijk} - \overline{X}_{ij.})$ the residual. Upon squared equation (1.3.4) and summed over all observations, all cross product terms will drop out to give:

$$SSTO = SSA + SSB_A + SSE \qquad (1.3.5)$$

where

$$SSTO = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{N}(X_{ijk} - \overline{X}_{...})^2$$

$$SSA = bN\sum_{i=1}^{a}(\overline{X}_{i..} - \overline{X}_{...})^2$$

$$SSB_A = N\sum_{i=1}^{a}\sum_{j=1}^{b}(\overline{X}_{ij.} - \overline{X}_{i..})^2$$

$$SSE = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{N}(X_{ijk} - \overline{X}_{ij.})^2 = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{N} e_{ijk}^2,$$

and obtain the ANOVA Table 1.1 for this Nested two-factor Effects study to compare whether the variability is significant or not at each level of the experiment.

The appropriate test statistic and the expected mean squares are also given in Table 1.2.

Source	Sum of Squares	d.f	Mean Square
Factor A	$bN\sum_{i=1}^{a}(\overline{X}_{i..}-\overline{X}_{...})^2$	a-1	$MSA=\frac{SSA}{a-1}$
Factor B	$N\sum_{i=1}^{a}\sum_{j=1}^{b}(\overline{X}_{ij.}-\overline{X}_{i..})^2$	a(b-1)	$MSB(A)=\frac{SSB(A)}{a(b-1)}$
Error	$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{N}(X_{ijk}-\overline{X}_{ij.})^2$	ab(N-1)	$MSE=\frac{SSE}{ab(N-1)}$
Total	$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{N}(X_{ijk}-\overline{X}_{...})^2$	abN-1	

Table 1.1: ANOVA table for Nested Random Effect Model (B nested in A)

Test for	F*	Expected mean square
Factor A	$\frac{MSA}{MSB(A)}$	$\sigma^2+bN\sigma_A^2+N\sigma_B^2$
Factor B	$\frac{MSB(A)}{MSE}$	$\sigma^2+N\sigma_B^2$
Error		σ^2

Table 1.2: The F-ratio values and expected mean squares

The results of this test are:

reject the hypothesis, that the variation due to factor A is not significant, if the F* value of factor A is greater than the prescribed value of F with the appropriate degree of freedom and a given type one error, α.

reject the hypothesis, that the variation due to factor B is not significant, if the F* value of factor B is greater than the prescribed value of F with the appropriate degree of freedom and a given type one error , α.

Yashchin (1995) stressed that conventional estimation procedures developed for Nested experiments may turn out to be inappropriate in situations involving process control. As a remedy to these, he developed CUSUM schemes (1994), and likelihood ratio methods (1995) for monitoring parameters of interest under the Nested Random Effects Model. He also verified that, the latter approach has advantages when it is important to maintain a low rate of false alarms despite the presence of nuisance parameters.

Chapter 2

The Sequential Method

This chapter will present review of the sequential likelihood ratio test, introduce the notion, and state the conditions as a framework and the results for the newly designed sequential likelihood ratio test.

2.1 Review of the Sequential Likelihood Ratio Test

According to the historical note of Wald (1945), the first idea for a sequential test procedure, i.e., a test for which the number of observations is not determined in advance but is dependent on the outcome of the observations as they are made, goes back to H.F. Dodge and H.G. Romig (1929), who constructed a double sampling procedure. According to this scheme, the decision whether or not a second sample should be drawn depends on the outcome of the observations in the first sample. Whereas this method allows for only two samples, Bartky (1943) devised a multiple sampling scheme for the particular case of testing the mean of a binomial distribution. His scheme is closely related to the test procedure that results from the application of the sequential probability ratio test to this particular case. The reason that Dodge and Roming introduced their double sampling method, and Bartky his multiple sampling scheme was, of course, the recognition of the fact that they require, on the average, a smaller number of observations than "single" sampling.

The occasional practice of designing a large-scale experiment in successive stages may be regarded as a forerunner of sequential analysis. The idea for such chain experiments was briefly discussed by Hotelling (1941).

The problem of sequential analysis arose in the statistical research group of Columbia University in connection with some comments made by Captain G.L. Schuyler of the Bureau of Ordnance, Navy Department. Milton Friedman and W. Allen Wallis recognized the great potentialities and the far-reaching consequences that sequential analysis might have for the further development of theoretical statistics. In particular, they conjectured that a sequential test procedure might be constructed which could control the possible errors committed by wrong decisions exactly to the same extent as the best current procedure based on a predetermined number of observations. At the same time it would require, on the average, a substantially smaller number of observations than the fixed number of observations needed for the current procedure. Friedman and Wallis also exhibited a few examples of sequential modifications of current test procedures resulting, in some cases, in an increased efficiency.

The clear formulation of the problem initiated Wald's interest (1947) and lead him to developing sequential probability ratio test, SPRT(1947), which then led to the development of Sequential Probability Ratio Test, SPRT in April 1943.

Further advances in the theory of the SPRT were made in 1944. The operating characteristic (OC) curve of the SPRT for the case of a binomial distribution was found by Milton Friedman and George W. Brown (independently of each other), and slightly earlier by C. M. Stockman.

In the late forties then the new theory of sequential hypothesis testing had been published in a book format, due to Abraham Wald (1947). In the Neyman-Pearson testing theory there are two hypotheses, the null hypothesis, H_o and the alternate hypothesis, H_1. For the decision the two choices are:

a) reject H_o in favor of H_1 or
b) do not reject H_o in favor of H_1.

In this decision there is direct control over α = P(Type I error) = P(reject H_o when it is true), but we have no control over β = P(Type II error) = P(do not reject H_o when H_1 is true), and hence β is of secondary importance in this theory.

In contrast, Wald's sequential testing theory gives three options to the decision maker.

a) accept H_1 (reject H_o)
b) accept H_0 (reject H_1)
c) continue sampling as not enough evidence in data for conclusion.

In Wald's procedure we have simultaneous control over α = P(Type I error) and β = P(Type II error), which makes it superior to the Neyman-Pearson approach in this respect. Furthermore, his procedure is optimal if the criterion is to come to a quick decision. Yet, Wald's theory is not as widely applied today because it can be used essentially only in case both H_o and H_1 are simple hypothesis. This is not the decision problem we face most of the time in practice. In fact, Wald (1945) (pp. 119) pointed out that the theory of sequential hypothesis with no restrictions on the possible values of the unknown parameter is not as simple.

There were efforts made to extend Wald's testing scheme to composite hypotheses, but with limited success. The first such attempt was by Wald himself, who considered composite hypotheses which were composite because of the presence of nuisance parameters. He suggested using weight functions on the range of nuisance parameters and integrated out so that the composite hypothesis became, in fact, a simple one in terms of the parameter of interest. However, Govindarajulu (1975), (pp. 121) pointed out that it is not always possible to find the appropriate weight functions proposed by Wald (1945, 1947) because no general method is available for choosing the weight functions. Also

Whetherill and Glazerbrook (1986), (p. 52) remark that this approach has not been successfully applied.

Another approach was using transformations which reduce the two hypotheses to simple ones in terms of some other, related transformed parameters. The most successful application of this is the sequential t-test of Barnard and Rushton (1950, 1952). Cox (1952) generalizes this approach by finding conditions and transformations which give a factorization of the likelihood function that leads to the cancellation of terms with nuisance parameters in the likelihood ratio. But it has no major application in sequential testing other than the sequential t-test.

The next approach to sequential testing in the presence of nuisance parameters is based on asymptotics developed by Bartlett (1946), Cox (1963), Breslow (1969) and Joanes (1972). It considers the likelihood ratio, or an asymptotically equivalent statistic, under contiguous alternatives and uses Wald's approach to sequential testing.

Sequential likelihood ratios that are derived from Wald's (1947) approach minimize the average sample number, and have an extensive literature (c.f. Lorden (1971), Moustakides (1986), Basseville and Benveniste (1986), Siegmund (1985)). If we use the generalized likelihood ratio process as defined in equation (1.2.2), the optimality property is different, as the power is maximized now (c.f. Bahadur (1966), Brown (1971)). These two different optimality criteria are discussed and demonstrated in sequential tests by Gombay (2000b). Those conclusions are valid in sequential change detection.

2.2 Description of the New Sequential Test

The sequential method of testing a null hypothesis, H_o , to be used, may be described as follows: A rule is given for making one of the following two decisions at any stage of the experiment (at the m^{th} trial for each integer value of m)

1. to reject the hypothesis H_o
2. to continue the experiment by making an additional observation.

Thus, such a test procedure is carried out sequentially. On the basis of the first two observations, one of the aforementioned two decisions is made. If the first decision is made, the process is terminated. If the second decision is made, a second trial is performed. Again, on the basis of the first three observations, one of the two decisions is made. If the second decision is made, a third trial is performed, and so on. If we introduce a truncation point n, then if decision 1 has not occurred by trial n, then we decide that H_o cannot be rejected. The process is continued until either the first decision is made or arrived at truncation point. The number of observations required by such a test procedure is a random variable, since the value depends on the outcome of the observations.

For each positive integral value m, m $\geq$ 2, we shall denote by M_m the totality of all possible samples $(x_1, x_2, ..., x_m)$ of size m. We shall also refer to M_m as the m-dimensional sample space. A rule for making one of the two decisions at any stage of the experiment can be described as follows. For each integral value m, the m-dimensional sample space is split into two mutually exclusive parts, R^1_{m-1}, R_{m-1}. After the first two observation (x_1, x_2) have been drawn, H_o is rejected if (x_1, x_2) lies in R^1_1 or a third observation is made otherwise. If the second decision is made and other observation x_3 is drawn , H_o is rejected, or a fourth observation is drawn, according as the observed sample (x_1, x_2, x_3) lies in R^1_2 *or* R_2. If (x_1, x_2, x_3) lies in R_2, a fourth observation x_4 is drawn and one of the two decisions is made according as (x_1, x_2, x_3, x_4) lies in R^1_3 *or* R_3, and so on. This process is stopped, when, and only when, either the first decision is made. or first decision has not occurred up to the truncation time n, where in the later case we decide H_o cannot be rejected.

Thus, a sequential test is completely defined by defining the sets R^1_{m-1}, R_{m-1} for all positive integral value m, m $\geq$ 2. Since R^1_{m-1}, R_{m-1} are mutually exclusive and add up to the sets M_m , any of the two sets R^1_{m-1}, R_{m-1} consists of precisely of all those samples which are not contained in the other one.

2.3 The Average (Expected) Sample Number (ASN) Function of a Sequential Test.

At any stage of the experiment the decision to terminate the process depends on the results of the observations made so far. We shall denote by N the number of observations required by the sequential test, which is a random variable. Carrying out the same sequential test procedure repeatedly, we shall obtain, in general, different values of N. Of particular interest is the expected value of N (the average value of N in the long run, when the same test procedure is applied repeatedly). For any given test procedure the expected value of N depends only on the distribution of X. Since the distribution of X is determined by the parameter point θ, the expected value of N depends only on θ .

2.4 The Sequential Likelihood Ratio Test

Let $X_1, X_2, ..., X_M$ be a sequence of independent, identically distributed random variables with density belonging to the exponential family distributions, that is,

$$\log f(x;\psi) = T(x)\psi^T + S(x) - A(\psi), \quad \psi = (\theta, \eta), \quad T(x) = (T^d(x), T^p(x))$$

where $\psi \in \Omega$, $\theta \in \Omega_1 \subset R^d$, $\eta \in \Omega_2 \subset R^p$, $d \geq 1$, $p \geq 0$, and $\Omega = \Omega_1 X \Omega_2$. In our discussion θ is the parameter vector of interest and η is the nuisance parameter vector, d and p are inteegers. The two hypothesis were also given by:

$$H_o : \theta = \theta_o, \eta \in \Omega_2 \qquad H_1 : \theta \neq \theta_o, \eta \in \Omega_2$$

2.4.1 Conditions as a Framework of the Test

Let Γ_o be a neighborhood of point (θ_o, η), and let Γ_1 be a neighborhood of the d+p dimensional interval spanned by end points (θ_o, η) and (θ_1, η_1) where η is the true value of parameter , θ_1 is the value of the parameter θ under the alternative hypothesis, and η_1 is defined as the solution of the equation

$$\nabla_\eta A(\theta_1, \eta_1) = E[T^p(X)] \ .$$

Consider the conditions:

(C1) $\nabla_\eta A$and $\nabla_\psi A$ are continuous and equations

$$\nabla_\eta A(\theta_o, \eta) = b$$

$$\nabla_\psi A(\psi) = a$$

have unique solutions for $a \in \Gamma_1 \subset R^{d+p}$, $b \in \Gamma_o \subset R^p$, and inv $(\nabla_\psi A)$, inv $(\nabla_\eta A)$ are Lipschitz continuous of order one in each variable $\theta_i, \eta_j, i = 1, 2, ..., d, j = 1, 2, ..., p$.

(C2) Matrices $\nabla^2_{\eta^2} A(\theta_o, \eta), \nabla^2_{\psi^2} A(\psi)$ are positive definite, and are Lipschitz continuous of order one in all arguments.

(C3) $\frac{\partial^3}{\partial\psi_i\partial\psi_j\partial\psi_k}A(\psi)$ exist and are bounded in $\Gamma \subset R^{d+p}$, i, j, k=1,2,....,d+p

(C4) $E[|T_j(X_1)|^\gamma] < \infty$ for $1 \leq j \leq d+p$ for some $\gamma > 2$

2.4.2 Test Results

Whenever the conditions **(C1-C4)** are satisfied, one can use the likelihood ratio for sequential testing. The derivation of the tests below, and some selected theory can be found in Gombay (1996, 1997, 2000). If the hypotheses are defined by equation (1.2.1), and we use the likelihood ratio given in equation (1.2.2), then the following two truncated sequential procedures can be performed. The point of truncation is denoted by M.

TEST1. Stop and reject H_o at the first k, when $-2\log\lambda_k \geq CV_1(\alpha, M)$, k=2,...,M, where critical value $CV_1(\alpha, n)$ is:

$$CV_1(\alpha, M) = \frac{(-\log(-\log(1-\alpha)) + b(M))^2}{2\log\log M} \quad ,$$

for α , the level of significance,
$b(M) = 2\log\log M + \frac{d}{2}\log\log\log M - \log\Gamma(\frac{d}{2})$, and

$$\Gamma(x) = \int_0^\infty y^{x-1}e^{-y}dy, \quad x > 0$$

TEST2. Let d=1. Stop and reject H_o at the first k when $-2\frac{k}{M}\log\lambda_k \geq CV_2(\alpha)$ for k=2,...,M, where $CV_2(\alpha)$ is the critical value obtained from the relation

$$1-\alpha = \frac{4}{\pi}\sum_{k=0}^{\infty}\frac{(-1)^k}{2k+1}exp\left(\frac{-\pi^2(2k+1)^2}{8(CV_2(\alpha))^2}\right)$$

Chapter 3

Monitoring Parameters

In this chapter, the controlling mechanisms of the parameters for the two factor Nested Random Effects Model is developed. It starts by showing that the distribution belongs to the exponential family.

3.1 The Model Being in an Exponential family

For the Nested Random Effect Model defined earlier, let there be **a** classes and **b** subclasses with N replicates each.

<u>Model</u>

$$X_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \epsilon_{ijk} \qquad \begin{cases} i = 1, 2, ..., a \\ j = 1, 2, ..., b \\ k = 1, 2, ..., N \end{cases} \tag{3.1.1}$$

and μ is a constant.

With the basic assumptions in (1.1.1)

$$X_{ijk} \sim N(\mu, \sigma_*^2), \quad \text{where } \sigma_*^2 = \sigma_A^2 + \sigma_B^2 + \sigma^2 \quad ,$$

then the density function for variable X with parameters mean μ and variance σ_*^2 will be:

$$f(x; \mu, \sigma_*^2) = \frac{1}{\sqrt{2\pi\sigma_*^2}} \exp\left\{ \frac{-1}{2\sigma_*^2}(x - \mu)^2 \right\} \tag{3.1.2}$$

Taking the natural logarithm of both sides:

$$\begin{aligned}\log f(x;\mu,\sigma_*^2) &= -\frac{1}{2}\log(2\pi\sigma_*^2) - \frac{1}{2\sigma_*^2}(x-\mu)^2 \\ &= -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma_*^2) - \frac{1}{2\sigma_*^2}(x^2 - 2\mu x + \mu^2) \\ &= \frac{\mu}{\sigma_*^2}x - \frac{1}{2\sigma_*^2}x^2 - \frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma_*^2) - \frac{\mu^2}{2\sigma_*^2} \quad (3.1.3)\end{aligned}$$

Let $T_1(x) = x$, $T_2(x) = x^2$, $\theta = \frac{\mu}{\sigma_*^2}$, and $\eta = -\frac{1}{2\sigma_*^2}$ so that $\sigma_*^2 = -\frac{1}{2\eta}$ and then equation (3.1.3) can be expressed as:

$$\log f(x;\theta,\eta) = (T_1(x), T_2(x))(\theta,\eta)^T - A(\theta,\eta) \qquad (3.1.4)$$

where $A(\theta,\eta) = \frac{1}{2}\log(2\pi) - \frac{1}{2}\log(-2\eta) - \frac{\theta^2}{4\eta}$ which shows that the density function belongs to the exponential family.
Without loss of generality, let's assume $\mu_o = 0$, hence the hypotheses in μ_o and μ will be reduced to

$$H_o : \theta = 0;\ \eta < 0 \ , \ vs. \ H_1 : \theta \neq 0 \ , \ \eta < 0 \qquad (3.1.5)$$

3.2 Check of the Conditions to the Theory

This section verifies that the conditions **(C1-C4)** of section 2.4.1 are satisfied and presents the application of the results developed in section 2.4.2.

From equation (3.1.3), the logarithm of the density function in terms of θ and η , we see that:

$$A(\theta,\eta) = \frac{1}{2}\log(2\pi) - \frac{1}{2}\log(-2\eta) - \frac{\theta^2}{4\eta}.$$

For the hypothesis to test given in equation (3.1.5), and to apply the new theory developed one needs to justify the smoothness conditions **(C1-C4)**. The verification looks like this:

(C1)

$$\nabla_\eta A(\theta,\eta) = \frac{\partial}{\partial\eta}A(\theta,\eta) = \frac{-1}{2\eta} + \frac{\theta^2}{4\eta^2}$$

Then under the null hypothesis one finds that:

$$\nabla_\eta A(\theta = 0, \eta) = \frac{-1}{2\eta} = \sigma_*^2$$

which is continuous and has a unique inverse for $\sigma_*^2 > 0$.
Again taking the gradient in both parameters one will get:

$$\nabla_{(\theta,\eta)} A(\theta,\eta) = \begin{pmatrix} \frac{-\theta}{2\eta} \\ -\frac{1}{2\eta} + \frac{\theta^2}{4\eta^2} \end{pmatrix} = \begin{pmatrix} \mu \\ \sigma_*^2 + \mu^2 \end{pmatrix}$$

Given any values a_1 and a_2 where $a_2 > a_1^2$ such that $\nabla_{(\theta,\eta)} A(\theta,\eta) = (a_1, a_2)$, then $\mu = a_1, \sigma_*^2 = a_2 - a_1^2$ which is a unique solution for the mean and variance. Therefore unique values of θ and η will exist. Moreover the functions and their inverses are continuous at these values a_1 and a_2 .
To check the Lipshitz condition consider:

$$\text{Inv } \nabla_\eta A(\theta,\eta) = \left(\frac{-1}{2\eta} + \frac{\theta^2}{4\eta^2}\right)^{-1} = \frac{4\eta^2}{\theta^2 - 2\eta} = \frac{1}{\sigma_*^2 + \mu^2}$$

Taking this equation as a function of θ for fixed η

$$f(\theta,\eta) = \frac{4\eta^2}{\theta^2 - 2\eta}$$

$$\begin{aligned} |f(\theta_1,\eta) - f(\theta_2,\eta)| &= \left|\frac{4\eta^2}{\theta_1^2 - 2\eta} - \frac{4\eta^2}{\theta_2^2 - 2\eta}\right| \\ &= 4\eta^2 \left|\frac{1}{\theta_1^2 - 2\eta} - \frac{1}{\theta_2^2 - 2\eta}\right| \\ &= 4\eta^2 \left|\frac{\theta_2^2 - \theta_1^2}{(\theta_2^2 - 2\eta)(\theta_1^2 - 2\eta)}\right| \\ &= \frac{4\eta^2}{(\theta_2^2 - 2\eta)(\theta_1^2 - 2\eta)} |\theta_1 + \theta_2||\theta_2 - \theta_1| \\ &\leq M|\theta_2 - \theta_1| \ , \end{aligned}$$

for some M,

$$\frac{4\eta^2|\theta_1 + \theta_2|}{(\theta_2^2 - 2\eta)(\theta_1^2 - 2\eta)} \leq M \ ,$$

which gives Lipshitz condition in the parameter θ.
For fixed θ

$$\begin{aligned}
|f(\theta,\eta_1) - f(\theta,\eta_2)| &= \left|\frac{4\eta_1^2}{\theta^2 - 2\eta_1} - \frac{4\eta_2^2}{\theta^2 - 2\eta_2}\right| \\
&= 4\left|\frac{\eta_1^2(\theta^2 - 2\eta_2) - \eta_2^2(\theta^2 - 2\eta_1)}{(\theta^2 - 2\eta_1)(\theta^2 - 2\eta_2)}\right| \\
&= 4\frac{|\eta_1^2\theta^2 - 2\eta_1^2\eta_2 - \eta_2^2\theta^2 + 2\eta_2^2\eta_1|}{(\theta^2 - 2\eta_1)(\theta^2 - 2\eta_2)} \\
&= 4\frac{\{(\eta_1^2 - \eta_2^2)\theta^2 + |2\eta_1\eta_2(\eta_2 - \eta_1)|\}}{(\theta^2 - 2\eta_1)(\theta^2 - 2\eta_2)} \\
&\le M|\eta_2 - \eta_1|
\end{aligned}$$

for some

$$M \ge \frac{4((\eta_1 + \eta_2)\theta^2 + 2|\eta_1\eta_2|)}{(\theta^2 - 2\eta_1)(\theta^2 - 2\eta_2)}$$

which implies again that, it is Lipshitz in the variable η too. Thus the first condition **(C1)** is satisfied.

(C2) One easily finds that for $\psi = (\theta, \eta)$

$$\nabla^2_{\psi^2} A(\psi) = \begin{pmatrix} \frac{-1}{2\eta} & \frac{\theta}{2\eta^2} \\ \frac{\theta}{2\eta^2} & \frac{1}{2\eta^2} - \frac{\theta^2}{2\eta^3} \end{pmatrix} = \begin{pmatrix} \sigma^2 & 2\mu\sigma^2 \\ 2\mu\sigma^2 & 2\sigma^4 + 4\mu^2\sigma^2 \end{pmatrix}$$

$|\nabla^2_{\psi^2} A(\psi)| = 2\sigma^6 + 4\mu^2\sigma^4 - 4\mu^2\sigma^4 = 2\sigma^6 > 0$ and **C2** is satisfied .

(C3)

To see the boundedness of the third total derivative

$$\frac{\partial^3}{\partial\theta^3} A(\theta,\eta) = 0, \quad \frac{\partial^3}{\partial\eta\partial\theta^2} A(\theta,\eta) = \frac{1}{2\eta^2},$$

$$\frac{\partial^3}{\partial\eta^2\partial\theta} A(\theta,\eta) = \frac{-\theta}{\eta^3}, \text{ and } \frac{\partial^3}{\partial\eta^3} A(\theta,\eta) = \frac{-1}{\eta^3} + \frac{3}{2}\frac{\theta^2}{\eta^4}$$

All exist and bounded in the coordinate plane.

(C4)

$$E|T_1(X)|^r = E|X|^r$$

Considering the moment generating function of a normal distribution with mean μ and variance σ_x^2, which is given by:

$$M_x(t) = e^{(\mu t + \frac{1}{2}t^2\sigma_x^2)} \tag{3.2.1}$$

and using the fact that

$$E[X^\gamma] = \left.\frac{d^\gamma M_x(t)}{dt^\gamma}\right|_{t=0} \tag{3.2.2}$$

we obtain for $\gamma = 4$

$$E[X] = \left.\frac{dM_x(t)}{dt}\right|_{t=0} = (\mu + t\sigma_x^2)(e^{(\mu t+\frac{1}{2}t^6 2\sigma_x^2)})\Big|_{t=0} = \mu$$

$$E[X^2] = \frac{d}{dt}\left[(\mu + t\sigma_x^2)(e^{(\mu t+\frac{1}{2}t^2\sigma_x^2)})\right]\Bigg|_{t=0} = \sigma_x^2 + \mu^2$$

$$\begin{aligned} E[X^3] &= \frac{d}{dt}\left[\sigma_x^2 e^{(\mu t+\frac{1}{2}t^2\sigma_x^2)} + (\mu + t\sigma_x^2)^2(e^{(\mu t+\frac{1}{2}t^2\sigma_x^2)})\right]\Bigg|_{t=0} \\ &= 3\mu\sigma_x^2 + \mu^3 \end{aligned}$$

and

$$\begin{aligned} E[X^4] &= \frac{d}{dt}\left[\left(3\sigma_x^2(\mu + t\sigma_x^2) + (\mu + t\sigma_x^2)^3\right)(e^{(\mu t+\frac{1}{2}t^2\sigma_x^2)})\right]\Bigg|_{t=0} \\ &= 3\sigma_x^4 + 6\mu^2\sigma_x^2 + \mu^4 \end{aligned}$$

We see that for $\gamma = 4$, $E|X|^\gamma$ is bounded and hence we can apply the results of the theorem to test the hypotheses given in equation (3.1.5).

3.3 Monitoring the Mean

For the model given in the first section, with the basic assumptions in (1.1.1), it was shown $X_{ijk} \sim N(\mu, \sigma_x^2)$. The objective is to test the hypothesis in (3.1.5). Without loss of generality let's assume $\mu_o = 0$ (Otherwise one can subtract the mean from each data and proceed) . Let $\overline{X}_{ij.} = \frac{1}{N}\sum_{k=1}^{N} X_{ijk}$ be the mean of the data in j^{th} sub-class and i^{th} class and let $\overline{X}_{i..} = \frac{1}{bN}\sum_{j=1}^{b}\sum_{k=1}^{N} X_{ijk} = \alpha_i + \frac{1}{b}\sum_{j=1}^{b}\beta_{j(i)} + \frac{1}{bN}\sum_{j=1}^{b}\sum_{k=1}^{N}\epsilon_{ijk} = \frac{1}{b}\sum_{j=1}^{b}\overline{X}_{ij.}$

be the mean of the data in i^{th} class. Then, from the normality of $X_{ijk}'s$, one easily verifies that:

$$\begin{aligned} E[\overline{X}_{i..}] &= E[\overline{X}_{ij.}] = \mu \\ Var[\overline{X}_{i..}] &= \sigma_A^2 + \frac{1}{b}\sigma_B^2 + \frac{1}{bN}\sigma^2 \\ Var[\overline{X}_{ij.}] &= \sigma_A^2 + \sigma_B^2 + \frac{1}{N}\sigma^2 \end{aligned}$$

which implies

$$\overline{X}_{i..} \sim N(\mu, \sigma_A^2 + \frac{1}{b}\sigma_B^2 + \frac{1}{bN}\sigma^2) \tag{3.3.1}$$

$$\overline{X}_{ij.} \sim N(\mu, \sigma_A^2 + \sigma_B^2 + \frac{1}{N}\sigma^2) \tag{3.3.2}$$

From (3.3.1), let $\sigma_A^2 + \frac{1}{b}\sigma_B^2 + \frac{1}{bN}\sigma^2 = \eta$, be the nuisance parameter. Hence the hypotheses to be tested will be: $H_o : \mu = 0$, $\eta > 0$ $vs.$ $H_1 : \mu \neq 0$, $\eta > 0$. It then follows from equation (3.3.1) that the density of $\overline{X}_{i..}$ is:

$$f(U_i, \mu, \eta) = (2\pi\eta)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\eta}(U_i - \mu)^2\right\} \tag{3.3.3}$$

where U_i stands for the random variable $\overline{X}_{i..}$. The likelihood function (for n=2,3,4,...,a) will be the following.

$$\begin{aligned} L(\mu,\eta) &= (2\pi\eta)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\eta}\sum_{i=1}^{n}(U_i - \mu)^2\right\} \\ \ell(\mu,\eta) &= \log\ (L(\mu,\eta)) \\ &= -\frac{n}{2}\log\ (2\pi\eta) - \frac{1}{2\eta}\sum_{i=1}^{n}(U_i-\mu)^2\ . \end{aligned} \tag{3.3.4}$$

When $\mu \neq \mu_o$ the maximum likelihood estimator of the mean will be found by solving the equation for μ in:

$$\begin{aligned} \frac{\partial}{\partial\mu}\ \ell(\mu,\eta) = 0 \ \text{ evaluated at } \ \mu = \hat{\mu} \\ \Rightarrow -\frac{1}{\eta}\sum_{i=1}^{n}(U_i - \hat{\mu}) = 0 \\ \Rightarrow \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}U_i = \frac{1}{n}\sum_{i=1}^{n}\overline{X}_{i..} \\ \Rightarrow \hat{\mu} = \overline{U}_n \end{aligned} \tag{3.3.5}$$

Where $\overline{U}_n = \frac{1}{n}\sum_{i=1}^{n} U_i = \frac{1}{n}\sum_{i=1}^{n} \overline{X}_{i..}$.
Similarily, the maximum likelihood estimator of the nuisance parameter η can be found by solving

$$\frac{\partial}{\partial\eta}\ \ell(\mu,\eta) = 0 \text{ evaluated at } \eta = \hat{\eta}_n \text{ and } \mu = \hat{\mu} = \overline{U}_n$$

$$-\frac{n}{2\hat{\eta}_n} + \frac{1}{2\hat{\eta}_n^2}\sum_{i=1}^{n}(U_i - \overline{U}_n)^2 = 0$$

$$\Rightarrow \hat{\eta}_n = \frac{1}{n}\sum_{i=1}^{n}(U_i - \overline{U}_n)^2 \qquad (3.3.6)$$

Using the maximized likelihood function (on the estimators found from equations (3.3.5) and (3.3.6)) we will have:

$$\begin{aligned} L(\hat{\mu},\hat{\eta}_n) &= (2\pi\hat{\eta}_n)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\hat{\eta}_n}\sum_{i=1}^{n}(U_i - \overline{U}_n)^2\right\} \\ &= (2\pi\hat{\eta}_n)^{-\frac{n}{2}} \exp\left\{-\frac{n}{2}\right\} \ . \end{aligned} \qquad (3.3.7)$$

When $\mu_o = 0$, to get the maximum likelihood estimator for the nuisance parameter η, we solve the equation for η in :

$$\frac{\partial}{\partial\eta}\ \ell(\mu_o,\eta) = 0 \text{ evaluated at } \eta = \hat{\eta}_o$$

$$\Rightarrow -\frac{n}{2\hat{\eta}_o} + \frac{1}{2\hat{\eta}^2}\sum_{i=1}^{n}(U_i - \mu_o)^2 = 0$$

$$\Rightarrow \hat{\eta}_o = \frac{1}{n}\sum_{i=1}^{n}(U_i - \mu_o)^2 \qquad (3.3.8)$$

Using the maximized likelihood function (on $\mu = \mu_o$ and the estimator found from equation (3.3.8)):

$$
\begin{aligned}
L(\mu_o, \hat{\eta}_o) &= (2\pi\hat{\eta}_o)^{-\frac{n}{2}} \exp\{-\frac{1}{2\hat{\eta}_o}\sum_{i=1}^{n}(U_i - \mu_o)^2\} \\
&= (2\pi\hat{\eta}_o)^{-\frac{n}{2}} \exp\{-\frac{n}{2}\} \qquad (3.3.9)
\end{aligned}
$$

From equations (3.3.7) and (3.3.9) and from the likelihood ratio for n classes, n=2, 3, ..., a, one can easily get:

$$
\begin{aligned}
\lambda_n &= \frac{L(\mu_o, \hat{\eta}_o)}{L(\hat{\mu}, \hat{\eta}_n)} \\
&= \left(\frac{\hat{\eta}_o}{\hat{\eta}_n}\right)^{-\frac{n}{2}} \qquad (3.3.10)
\end{aligned}
$$

where $\hat{\eta}_n = \frac{1}{n}\sum_{i=1}^{n}(U_i - \overline{U}_n)^2$ and $\hat{\eta}_o = \frac{1}{n}\sum_{i=1}^{n}(U_i - \mu_o)^2$. Then

$$
\begin{aligned}
-2\log\lambda_n &= -2\log\left[\left(\frac{\hat{\eta}_o}{\hat{\eta}_n}\right)^{-\frac{n}{2}}\right] \\
&= n\log\left[\frac{\hat{\eta}_o}{\hat{\eta}_n}\right] \\
&= n\log\frac{\sum_{i=1}^{n}(U_i - \mu_o)^2}{\sum_{i=1}^{n}(U_i - \overline{U}_n)^2}, \quad n = 2, ..., a \qquad (3.3.11)
\end{aligned}
$$

We choose **a** as a truncation point and then according to the test results in section (2.4.2) lead to the following procedures.

Test 1

Start testing for n=2 and reject H_o whenever the result $-2\log\lambda_n \geq CV_1(\alpha, a)$, where $-2\log\lambda_n$ is the one obtained in equation (3.3.11) and $CV_1(\alpha, a)$ as defined in Test 1 of section 2.4.2 and conclude that a difference is detected in favor of H_1 . Otherwise proceed testing after increasing the size by 1 (i.e. $n \rightarrow n+1$) and do the updated analysis. Repeat the procedure until the result $-2\log\lambda_n \geq CV_1(\alpha, a)$ is obtained for some n $\leq$ a or obtaining n = a before the required inequality. In the later case support the claim that there is no evidence of difference in favor of H_o.

Test 2

Start testing for n=2 and reject H_o if $-2\frac{n}{a}\log\lambda_n \geq CV_2(\alpha)$, where $CV_2(\alpha)$ as defined in Test 2 of section 2.4.2 and conclude that a difference is detected in favor of H_1 . If not proceed testing after increasing the size by 1 (i.e. $n \to n+1$) and do the updated analysis. Repeat the procedure until the result $-2\frac{n}{a}\log\lambda_n \geq CV_2(\alpha)$ is obtained for some n $\leq$ a or obtaining n = a before the required inequality. In the later case support the claim that there is no evidence of difference in favor of H_o.

3.4 Monitoring Variance

As derived earlier $\mathrm{Var}(X_{ijk}) = \sigma_A^2 + \sigma_B^2 + \sigma^2 = \sigma_*^2$
From Yashchin (1995):

$$\begin{aligned} \hat{\sigma}_i^2 &= \frac{1}{b(N-1)}\sum_{j=1}^{b}\sum_{k=1}^{N}(X_{ijk}-\overline{X}_{ij.})^2 \\ &= \frac{1}{b}\sum_{j=1}^{b} S_{ij}^2 \end{aligned} \qquad (3.4.1)$$

estimates σ^2 , at the i^{th} level of factor A, where

$$S_{ij}^2 = \frac{1}{N-1}\sum_{k=1}^{N}(X_{ijk}-\overline{X}_{ij.})^2\ , \quad \overline{X}_{ij.} = \frac{1}{N}\sum_{k=1}^{N} X_{ijk}$$

3.4.1 Monitoring Error Variance, σ^2

The hypothesis to be tested is :

$$H_o:\ \sigma^2 = \sigma_o^2 \quad vs. \quad H_1:\ \sigma^2 \neq \sigma_o^2 \qquad (3.4.2)$$

Consider the statistic,

$$\hat{\sigma}_i^2 = Z_i = \frac{1}{b(N-1)}\sum_{j=1}^{b}\sum_{k=1}^{N}(X_{ijk}-\overline{X}_{ij.})^2 \quad ,$$

an estimator of σ^2 . Then, from the fact that X_{ijk}'s are normal random variables, and $E[X_{ijk}-\mu]^2 = \sigma^2$ the distribution theory give us the result:

$$\hat{\sigma}_i^2 \sim \frac{\sigma^2}{b(N-1)}\chi^2(b(N-1))\ ,$$

where $\chi^2(b(N-1))$ stands for a chi-square distribution with b(N-1) degrees of freedom. So the density of $\hat{\sigma}_i^2 = Z_i$ can be obtained by transformation.

$$\begin{aligned} f(z,\theta) &= \frac{d}{dz}P[\hat{\sigma}_i^2 \leq z] \qquad (3.4.3)\\ &= \frac{d}{dz}P[\frac{v}{\sigma^2}\hat{\sigma}_i^2 \leq z\frac{v}{\sigma^2}] \quad , \quad v = b(N-1)\\ &= f_t\left(z\frac{v}{\sigma^2}\right)\frac{v}{\sigma^2}\\ &= \frac{\left(\frac{v}{2\sigma^2}\right)^{\frac{v}{2}} z^{\frac{v}{2}-1}}{\Gamma(\frac{v}{2})} exp\left\{\frac{-zv}{2\sigma^2}\right\}\\ &= \exp\left[\frac{v}{2}\log\left(\frac{v}{2\sigma^2}\right) - z\frac{v}{2\sigma^2} + (\frac{v}{2}-1)\log z - \log\Gamma(\frac{v}{2})\right]\\ &= \exp\left[\frac{v}{2}\log(-\theta) + z\theta + H(z)\right] \qquad (3.4.4) \end{aligned}$$

where $\theta = -\frac{v}{2\sigma^2}$, $H(z) = (\frac{v}{2}-1)\log z - \log\Gamma(\frac{v}{2})$. It can clearly be seen that the distribution function is a member of exponential family and one can easily verify that the conditions **C1-C4** are met. Hence we can apply our result for the test of the Hypotheses mentioned in equation (3.4.2).

Mle from n observations (in our case n classes) will be obtained as: $\ell(\theta) = \sum_{i=1}^{n}\left[\frac{b(N-1)}{2}\ln(-\theta) + Z_i\theta + H(Z_i)\right]$. Solving for the value of θ that maximizes the log likelihood, we get:

$$\left.\frac{\partial\ell(\theta)}{\partial\theta}\right|_{\theta=\hat{\theta}} = 0 \Rightarrow \hat{\theta}_n = -\frac{nb(N-1)}{2\sum_{i=1}^{n} Z_i}$$

and then from the invariance property of maximum likelihoods one gets the result

$$\hat{\sigma}_n^2 = \frac{1}{nb(N-1)}\sum_{i=1}^{n}\sum_{j=1}^{b}\sum_{k=1}^{N}(X_{ijk} - \overline{X}_{ij.})^2 = \frac{1}{n}\sum_{i=1}^{n} Z_i \qquad (3.4.5)$$

Using the value of the assumed variance under the null hypothesis of the test one will find the result

$$\lambda_n = \frac{\prod_{i=1}^{n}\prod_{j=1}^{b} f(Z_i,\theta_o)}{\prod_{i=1}^{n}\prod_{j=1}^{b} f(Z_i,\hat{\theta}_n)} \qquad (3.4.6)$$

which is a one dimensional test statistic with no nuisance parameter. Then, expressing -2 $\log\lambda_n$ in terms of the original parameters, we obtain:

$$-2\log\lambda_n = -2\sum_{i=1}^{n}\left[\log f(Z_i,\sigma_o^2) - \log f(Z_i,\hat{\sigma}_n^2)\right]$$

$$
\begin{aligned}
&= -2\sum_{i=1}^{n}\left[\left(\frac{v}{2}\log(\frac{v}{2\sigma_o^2}) - Z_i\frac{v}{2\sigma_o^2}\right) - \left(\frac{v}{2}\log(\frac{v}{2\hat{\sigma}_n^2}) - Z_i\frac{v}{2\hat{\sigma}_n^2}\right)\right] \\
&= \sum_{i=1}^{n}\left[v\left(\log(\frac{v}{2\hat{\sigma}_n^2}) - \log(\frac{v}{2\sigma_o^2})\right) + vZ_i\left(\frac{1}{\sigma_o^2} - \frac{1}{\hat{\sigma}_n^2}\right)\right] \\
&= nv\log\left(\frac{\sigma_o^2}{\hat{\sigma}_n^2}\right) + v\left(\frac{1}{\sigma_o^2} - \frac{1}{\hat{\sigma}_n^2}\right)\sum_{i=1}^{n} Z_i \\
&= nb(N-1)\left[\log\left(\frac{\sigma_o^2}{\overline{Z}_n}\right) + \frac{\overline{Z}}{\sigma_o^2} - 1)\right] \quad ,
\end{aligned} \tag{3.4.7}
$$

where $\overline{Z}_n = \hat{\sigma}_n^2$ as defined earlier. Then, the results in section (2.4.2) lead to the following procedures.

Test 1

Start testing for n=2 and reject H_o whenever the result $-2\log\lambda_n \geq CV_1(\alpha, a)$ where $-2\log\lambda_n$ is the one obtained from equation (3.4.7) and $CV_1(\alpha, a)$ as defined in Test 1 of section 2.4.2 and conclude that a difference is detected in favor of H_1 . If not, proceed testing after increasing the size by 1 (i.e. $n \to n+1$) and do the updated analysis. Repeat the procedure until the result $-2\log\lambda_n \geq CV_1(\alpha, a)$ is obtained for some n $\leq$ a or until getting n=a without obtaining the required inequality. In the later case support the claim that there is no evidence of difference in favor of H_o .

Test 2

Start testing for n=2 and reject H_o if $-2\frac{n}{a}\log\lambda_n \geq CV_2(\alpha)$ where $CV_2(\alpha)$ as defined in Test 2 of section 2.4.2 and conclude that a difference is detected in favor of H_1 . If not proceed testing after increasing the size by 1 (i.e. $n \to n+1$) and do the updated analysis. Repeat the procedure until the result $-2\frac{n}{a}\log\lambda_n \geq CV_2(\alpha)$ is obtained for some n $\leq$ a or until n=a without obtaining the required inequality. In the later case support the claim that there is no evidence of difference in favor of H_o .

3.4.2 Monitoring Class Variance , σ_A^2

Given the model that was defined earlier by:

$$X_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \epsilon_{ijk} \qquad \begin{cases} i = 1, 2, ..., a \\ j = 1, 2, ..., b \\ k = 1, 2, ..., N \end{cases}$$

which measure of k^{th} unit on j^{th} subclass and i^{th} class with assumptions in equation (1.1.1), where:
α_i - random effect of the i^{th} class
$\beta_{j(i)}$ - nested effect of j^{th} subclass in the i^{th} class and
ϵ_{ijk} - random noise and,
parameters of interest will be μ , σ_A , σ_B , σ .

Suppose we intend to monitor the variability σ_A^2
i.e., our interest is in testing the hypotheses:

$$H_o : \sigma_A^2 = \sigma_{A_o}^2 \quad vs. \quad H_1 : \sigma_A^2 \neq \sigma_{A_o}^2 \tag{3.4.8}$$

Consider

$$\overline{X}_{i..} = \frac{1}{bN}\sum_{j=1}^{b}\sum_{k=1}^{N} X_{ijk} = \mu + \alpha_i + \frac{1}{b}\sum_{j=1}^{b}\beta_{j(i)} + \frac{1}{bN}\sum_{j=1}^{b}\sum_{k=1}^{N}\epsilon_{ijk}$$

As derived earlier, $E[\overline{X}_{i..}] = \mu$
and Var $(\overline{X}_{i..}) = \sigma_A^2 + \frac{1}{b}\sigma_B^2 + \frac{1}{bN}\sigma^2 = \sigma_A^2 + \frac{1}{b}(\sigma_B^2 + \frac{\sigma^2}{N})$.
Hence if we put $\overline{X}_{i..} = U_i$, then from the normality of the random variable X_{ijk}'s we arrive at:
$U_i \sim N(\mu, \sigma_A^2 + \frac{1}{b}(\sigma_B^2 + \frac{\sigma^2}{N})) = N(\mu, \sigma_A^2 + \frac{1}{b}\xi)$, where $\xi = \sigma_B^2 + \frac{\sigma^2}{N}$. In this case the parameter of interest is σ_A^2 and nuisance parameters μ and ξ. Again from normality we conclude that :

$$\hat{\sigma}_{i.}^2 = \frac{1}{b-1}\sum_{j=1}^{b}(\overline{X}_{ij.} - U_i)^2$$

is independent of U_i for each i. For the random variable $\overline{X}_{ij.} = \frac{1}{N}\sum_{k=1}^{N} X_{ijk}$, $E[\overline{X}_{ij.} - \mu]^2 = \sigma_B^2 + \frac{\sigma^2}{N} = \xi$.

Hence from distribution theory:

$$\frac{(b-1)\hat{\sigma}^2_{i.}}{\xi} \sim \chi^2_{(b-1)} \tag{3.4.9}$$

for a chi-square distribution with (b-1) degrees of freedom. Let $Y_i = \hat{\sigma}^2_{i.}$, then the density function of Y can be found as :

$$\begin{aligned}
f(y) &= \frac{d}{dy} P[Y_i \le y] \\
&= \frac{d}{dy} P\left[\frac{(b-1)Y_i}{\xi} \le \frac{(b-1)y}{\xi}\right] \\
&= \left(\frac{b-1}{\xi}\right) f_t\left(\frac{y(b-1)}{\xi}\right) \\
&= \left(\frac{b-1}{\xi}\right) \frac{1}{\Gamma\left(\frac{b-1}{2}\right) 2^{\frac{b-1}{2}}} \left(\frac{y(b-1)}{\xi}\right)^{\frac{b-1}{2}-1} \exp\left\{-\frac{y(b-1)}{2\xi}\right\} \\
&= \left(\frac{b-1}{2\xi}\right)^{\frac{b-1}{2}} \frac{1}{\Gamma\left(\frac{b-1}{2}\right)} (y)^{\frac{b-1}{2}-1} \exp\left\{-\frac{y(b-1)}{2\xi}\right\}
\end{aligned}$$

From the independence property, the joint density of U_i and Y_i will be:

$$f(U,Y;\mu,\sigma^2_A,\xi) = \frac{\exp\left\{-\frac{(U-\mu)^2}{2(\sigma^2_A+\frac{\xi}{b})}\right\}}{\sqrt{2\pi(\sigma^2_A+\frac{\xi}{b})}} \left(\frac{b-1}{2\xi}\right)^{\frac{b-1}{2}} \frac{\exp\left\{-\frac{Y(b-1)}{2\xi}\right\}}{\Gamma\left(\frac{b-1}{2}\right)} (y)^{\frac{b-3}{2}}$$

the likelihood function for n observations, n=1,2,3, ..., a then will be:

$$L(\mu,\sigma^2_A,\xi) = \prod_{i=1}^{n} \left[\frac{\exp\left\{-\frac{(U_i-\mu)^2}{2(\sigma^2_A+\frac{\xi}{b})}\right\}}{\sqrt{2\pi(\sigma^2_A+\frac{\xi}{b})}} \left(\frac{b-1}{2\xi}\right)^{\frac{b-1}{2}} \frac{\exp\left\{-\frac{Y_i(b-1)}{2\xi}\right\}}{\Gamma\left(\frac{b-1}{2}\right)} (y_i)^{\frac{b-3}{2}}\right]$$

and the log likelihood,

$$\begin{aligned}
\ell(\mu,\sigma^2_A,\xi) &= \sum_{i=1}^{n} \left\{\frac{-(U_i-\mu)^2}{2(\sigma^2_A+\frac{1}{b}\xi)} - \frac{Y_i(b-1)}{2\xi} - \left(\frac{b-1}{2}\right)\log 2\xi \right. \\
&\quad \left. -\frac{1}{2}\log(\sigma^2_A+\frac{1}{b}\xi)\right\} + W(Y_i)
\end{aligned}$$

where

$$W(Y_i) = \frac{n(b-1)}{2}\log(b-1) - n\log\Gamma\left(\frac{b-1}{2}\right) + \left(\frac{b-3}{2}\right)\sum_{i=1}^{n}\log Y_i$$

is a function of the Y_i's only.

To get the maximum likelihood estimates for n classes (n=2, 3, 4, ..., a):

$$\ell(\mu,\sigma_A^2,\xi) = \frac{-1}{2(\sigma_A^2+\frac{1}{b}\xi)}\sum_{i=1}^{n}(U_i-\mu)^2 - \frac{(b-1)}{2\xi}\sum_{i=1}^{n}Y_i - \left(\frac{n(b-1)}{2}\right)\log 2\xi - \frac{n}{2}\log(\sigma_A^2+\frac{1}{b}\xi) + W(Y_i) \quad (3.4.10)$$

$$\frac{\partial\ell(\mu,\sigma_A^2,\xi)}{\partial\mu}\Big|_{\mu=\hat{\mu}} = 0 \Rightarrow \hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n}U_i = \overline{U}_n$$

$$\frac{\partial\ell(\hat{\mu},\sigma_A^2,\xi)}{\partial\sigma_A^2}\Big|_{\sigma_A^2=\hat{\sigma}_A^2} = 0 \Rightarrow \hat{\sigma}_A^2+\frac{1}{b}\xi = \frac{\sum_{i=1}^{n}(U_i-\hat{\mu})^2}{n} \quad (3.4.11)$$

$$\frac{\partial\ell(\hat{\mu},\sigma_A^2,\xi)}{\partial\xi} = \frac{1}{2}\left(\frac{\sum_{i=1}^{n}(U_i-\hat{\mu})^2}{b(\sigma_A^2+\frac{1}{b}\xi)^2} + \frac{b-1}{\xi^2}\sum_{i=1}^{n}Y_i - \frac{n(b-1)}{\xi} - \frac{n}{b(\sigma_A^2+\frac{1}{b}\xi)}\right)$$

$$= \frac{1}{2}\left(\frac{\sum_{i=1}^{n}(U_i-\hat{\mu})^2 - n(\sigma_A^2+\frac{1}{b}\xi)}{b(\sigma_A^2+\frac{1}{b}\xi)^2} - \frac{(b-1)(n\xi-\sum_{i=1}^{n}Y_i)}{\xi^2}\right) \quad (3.4.12)$$

When σ_A^2 is estimated by $\sigma_{A_1}^2$, and $\hat{\mu}_n = \overline{U}_n$ then

$$\sigma_{A_1}^2 + \frac{1}{b}\xi = \frac{\sum_{i=1}^{n}(U_i-\overline{U}_n)^2}{n}$$

and then the mle $\hat{\xi}$ *of* ξ in equation (3.4.12) requires only

$$n\hat{\xi} - \sum_{i=1}^{n}Y_i = 0 \Rightarrow \hat{\xi} = \frac{1}{n}\sum_{i=1}^{n}Y_i = \overline{Y}_n .$$

When $\sigma_A^2 = \sigma_{A_o}^2$, and $\hat{\mu} = \overline{U}_n$ then equation (3.4.12) becomes:

$$0 = \hat{\xi}^2\sum_{i=1}^{n}(U_i-\overline{U}_n)^2 - n\hat{\xi}^2(\sigma_{A_o}^2+\frac{1}{b}\hat{\xi}) + (b-1)(b\sigma_{A_o}^2+\hat{\xi})^2(-n\hat{\xi}+\sum_{i=1}^{n}Y_i)$$

$$= \xi^2\left[\sum_{i=1}^{n}(U_i-\overline{U}_n)^2 - n\sigma_{A_o}^2\right] - \frac{n}{b}\xi^3 + (b-1)\left(\sum_{i=1}^{n}Y_i - n\xi\right)\left[b\sigma_{A_o}^4 + 2\xi\sigma_{A_o}^2 + \frac{\xi^2}{b}\right]$$

$$= \xi^2\left[\sum_{i=1}^{n}(U_i-\overline{U}_n)^2 - n\sigma_{A_o}^2\right] - \frac{n}{b}\xi^3 + (b-1)\Big[b\sigma_{A_o}^4\sum_{i=1}^{n}Y_i + 2\xi\sigma_{A_o}^2\sum_{i=1}^{n}Y_i + \frac{\xi^2}{b}\sum_{i=1}^{n}Y_i - nb\sigma_{A_o}^4\sum_{i=1}^{n}Y_i - 2n\xi^2\sigma_{A_o}^2 - \frac{n\xi^3}{b}\Big]$$

$$
\begin{aligned}
&= \xi^3(-n) + \xi^2 \left[\sum_{i=1}^{n}(U_i - \overline{U}_n)^2 - n\sigma^2_{A_o} + \frac{b-1}{b}\sum_{i=1}^{n} Y_i - 2n(b-1)\sigma^2_{A_o}\right] \\
&\quad \xi\left((b-1)(2\sigma^2_{A_o}\sum_{i=1}^{n} Y_i - nb\sigma^4_{A_o})\right) + b(b-1)\sigma^4_{A_o}\sum_{i=1}^{n} Y_i \\
&= \xi^3 + \left(-\frac{\sum_{i=1}^{n}(U_i - \overline{U}_n)^2}{n} + (2b-1)\sigma^2_{A_o} - \frac{b-1}{b}\overline{Y}_n\right)\xi^2 \\
&\quad + \left[(b-1)(b\sigma^4_{A_o} - 2\sigma^2_{A_o}\overline{Y})\right]\xi - b(b-1)\sigma^4_{A_o}\overline{Y}_n \qquad (3.4.13) \\
&\quad \text{where} \quad \overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i
\end{aligned}
$$

Which reduces to solving the cubic equation in ξ . At $\xi = 0$, the right hand side becomes $-b(b-1)\sigma^4_{A_o}\overline{Y}_n$ which is less than zero and by taking sufficiently large ξ , as ξ^3 dominates others , the right hand side will be positive and hence one can be sure of the existence of a positive real root.
Then, using the given informations and the method of solving the roots of cubic equation given in Appendix A , we will have for the coefficients of the equation $\xi^3 + C_1\xi^2 + C_2\xi + C_3 = 0$:

$$
\begin{aligned}
C_1 &= -\frac{\sum_{i=1}^{n}(U_i - \overline{U}_n)^2}{n} + (2b-1)\sigma^2_{A_o} - \frac{b-1}{b}\overline{Y}_n \\
C_2 &= (b-1)(b\sigma^4_{A_o} - 2\sigma^2_{A_o}\overline{Y}_n) \quad \text{and} \\
C_3 &= -b(b-1)\sigma^4_{A_o}\overline{Y}_n \qquad (3.4.14)
\end{aligned}
$$

Hence for a given $\sigma^2_{A_o}$ and the data, one can substitute those values, calculate the values of Q, R and D^* and determine the solution. If there are more than one positive real solution to the cubic equation one has to take the value which maximizes the likelihood in (3.4.10).
Let this value be denoted by $\hat{\xi}_o$, then for n observations (n classes in our case)

$$
\begin{aligned}
\ell(\mu, \sigma^2_A, \xi) &= -\frac{n}{2}\log(\sigma^2_A + \frac{1}{b}\xi) - \frac{\sum_{i=1}^{n}(U_i - \mu)^2}{2(\sigma^2_A + \frac{1}{b}\xi)} \\
&\quad -\frac{n(b-1)}{2}\log 2\xi - \left(\frac{b-1}{2\xi}\right)\sum_{i=1}^{n} Y_i + W(Y_i)
\end{aligned}
$$

which is not a member of the exponential family. Referring Gombay (1996) and the verification above for the unique existence of mle's under H_o , one can use the generalized likelihood ratio. Also under H_1 , referring Appendix B, the consistency of the test follows.

$$-2\ell(\overline{U}_n,\sigma_A^2,\xi)=\begin{cases} n\log(\sigma_{A_o}^2+\frac{1}{b}\hat{\xi}_o)+\frac{\sum_{i=1}^n(U_i-\overline{U}_n)^2}{(\sigma_A^2+\frac{1}{b}\hat{\xi}_o)}+n(b-1)\log(2\hat{\xi}_o)+ & \\ \quad\frac{b-1}{\hat{\xi}_o}\sum_{i=1}^n Y_i-2W(Y_i) & \text{, if } \sigma_A^2=\sigma_{A_o}^2 \\ n\left(\log\left(\frac{\sum_{i=1}^n(U_i-\overline{U}_n)^2}{n}\right)+1+(b-1)\log\left(\frac{2\sum_{i=1}^n Y_i}{n}\right)\right)+ & \\ \quad n(b-1)-2W(Y_i) & \text{, otherwise} \end{cases}$$

then follows:

$$\begin{aligned} -2\log\lambda_n &= -2\ell(\overline{U}_n,\sigma_{A_o}^2,\hat{\xi}_o)+2\ell(\overline{U}_n,\hat{\sigma}_{A_1}^2,\hat{\xi}) \\ &= n\log\left(\frac{n(\sigma_{A_o}^2+\frac{1}{b}\hat{\xi}_o)}{\sum_{i=1}^n(U_i-\overline{U}_n)^2}\right)+\left(\frac{\sum_{i=1}^n(U_i-\overline{U}_n)^2}{(\sigma_{A_o}^2+\frac{1}{b}\hat{\xi}_o)}-n\right)+ \\ &\quad n(b-1)\log\left(\frac{n\hat{\xi}_o}{\sum_{i=1}^n Y_i}\right)+(b-1)\left(\frac{\sum_{i=1}^n Y_i}{\hat{\xi}_o}-n\right) \qquad (3.4.15) \end{aligned}$$

The next step then will be to apply the test.
The results in section (2.4.2) lead to the following procedures.

Test 1

Start testing for n=2 and reject H_o whenever the result $-2\log\lambda_n \geq CV_1(\alpha,a)$ where $-2\log\lambda_n$ is the one obtained from equation (3.4.15) and $CV_1(\alpha,a)$ as defined in Test 1 of section 2.4.2 and conclude that a difference is detected in favor of H_1 . If not, proceed testing after increasing the size by 1 (i.e. $n \rightarrow n+1$) and do the updated analysis. Repeat the procedure until the result $-2\log\lambda_n \geq CV_1(\alpha,a)$ is obtained for some n $\leq$ a or until getting n=a without obtaining the required inequality. In the later case support the claim that there is no evidence of difference in favor of H_o .

Test 2

Start testing for n=2 and reject H_o if $-2\frac{n}{a}\log\lambda_n \geq \; CV_2(\alpha)$ where $CV_2(\alpha)$ as defined in Test 2 of section (2.4.2) and conclude that a difference is detected in favor of H_1 . If not proceed testing after increasing the size by 1 (i.e. $n \to n+1$) and do the updated analysis. Repeat the procedure until the result $-2\frac{n}{a}\log\lambda_n \geq \; CV_2(\alpha)$ is obtained for some n $\leq$ a or until n=a without obtaining the required inequality. In the later case support the claim that there is no evidence of difference in favor of H_o .

3.4.3 Monitoring Subclass Variance, σ_B^2

Given the model:

$$X_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \epsilon_{ijk},$$

with the basic assumptions given in equation (1.1.1). The hypotheses to be tested here are :

$$H_o : \sigma_B^2 = \sigma_{B_o}^2 \quad vs. \quad H_1 : \sigma_B^2 \neq \sigma_{B_o}^2 \tag{3.4.16}$$

For any class i, and j=1,2,...,b

$E[\overline{X}_{ij.}] = \mu$, and $E[\overline{X}_{ij.} - \mu]^2 = \sigma_B^2 + \frac{1}{N}\sigma^2 = \xi$.

Then, as each of the X_{ijk}'s is a normal random variable, we conclude that for each i,

$\overline{X}_{ij.} \sim N(\mu,\xi)$. Let $Y_i = \frac{1}{b-1}\sum_{j=1}^{b}(\overline{X}_{ij.} - U_i)^2 = \hat{\sigma}_{i.}^2$ that is an estimator of $\sigma_B^2 + \frac{1}{N}\sigma^2 = \xi$ where $U_i = \frac{1}{b}\sum_{j=1}^{b}\overline{X}_{ij.}$. Then from equation (3.4.9) $\frac{(b-1)Y_i}{\xi} \sim \chi^2_{(b-1)}$ and from the density of y_i :

$$\log f_{Y_i}(y) = -\frac{(b-1)}{2\xi}y - \frac{(b-1)}{2}\log 2\xi + F(y) \tag{3.4.17}$$

Where F(y) is a function of y only which doesn't depend on the parameter. Again consider,

$$Z_i = \frac{1}{b(N-1)}\sum_{j=1}^{b}\sum_{k=1}^{N}(X_{ijk} - \overline{X}_{ij.})^2 = \hat{\sigma}_i^2$$

For any i and j, then $E[X_{ijk}] = \mu$, and $E[X_{ijk} - \mu]^2 = \sigma^2$, then from distribution theory we will get
$\frac{b(N-1)Z_i}{\sigma^2} \sim \chi^2_{b(N-1)}$ and from the density of Z_i we derive:

$$\log f_{Z_i}(z) = -\frac{b(N-1)}{2\sigma^2}z - \frac{b(N-1)}{2}\log 2\sigma^2 + H(z) \tag{3.4.18}$$

From the Independence of Y_i and Z_i , and equations (3.4.17) and (3.4.18) we can write the joint density as:

$$f_{Y_i,Z_i}(y,z) = f_{Y_i}(y) * f_{Z_i}(z) \quad , \quad i = 1, 2, ..., a \ .$$

For n observations (n = 2, 3, 4, ..., a):

$$\begin{aligned}
\ell(\sigma_B^2, \sigma^2) &= \log \prod_{i=1}^{n} f_{Y_i,Z_i}(Y_i, Z_i) = \sum_{i=1}^{n} [\log f_{Y_i}(Y_i) + \log f_{Z_i}(Z_i)] \\
&= \sum_{i=1}^{n} \Bigg[-\frac{(b-1)}{2\xi}Y_i - \frac{(b-1)}{2}\log 2\xi + F(Y_i) \\
&\quad -\frac{b(N-1)}{2\sigma^2}Z_i - \frac{b(N-1)}{2}\log 2\sigma^2 + H(Z_i) \Bigg] \\
&= -\left(\frac{b-1}{2(\sigma_B^2 + \frac{\sigma^2}{N})}\right)\sum_{i=1}^{n} Y_i - n\left(\frac{b-1}{2}\right)\log\left(2(\sigma_B^2 + \frac{\sigma^2}{N})\right) \\
&\quad -\frac{b(N-1)}{2\sigma^2}\sum_{i=1}^{n} Z_i - \frac{nb(N-1)}{2}\log 2\sigma^2 + Q(Y_i, Z_i)
\end{aligned} \tag{3.4.19}$$

where $F(Y_i)$, $H(Z_i)$ and $Q(Y_i, Z_i)$ are functions of the data only and are independent of the parameters. To get the mle:

$$\begin{aligned}
\frac{\partial \ell(\sigma_B^2, \sigma^2)}{\partial \sigma_B^2}\Big|_{\sigma_B^2 = \hat{\sigma}_B^2} = 0 \quad &\Rightarrow \quad \frac{b-1}{2(\hat{\sigma}_B^2 + \frac{\sigma^2}{N})^2}\sum_{i=1}^{n} Y_i - \frac{n(b-1)}{2(\hat{\sigma}_B^2 + \frac{\sigma^2}{N})} = 0 \\
&\Rightarrow \quad \hat{\sigma}_B^2 + \frac{\sigma^2}{N} = \frac{1}{n}\sum_{i=1}^{n} Y_i
\end{aligned} \tag{3.4.20}$$

Again

$$\frac{\partial \ell(\sigma_B^2, \sigma^2)}{\partial \sigma^2} = \frac{b-1}{2N(\sigma_B^2 + \frac{\sigma^2}{N})^2}\sum_{i=1}^{n} Y_i - \frac{n(b-1)}{2N(\sigma_B^2 + \frac{\sigma^2}{N})} + \frac{b(N-1)}{2\sigma^4}\sum_{i=1}^{n} Z_i - \frac{nb(N-1)}{2\sigma^2} \tag{3.4.21}$$

When σ_B^2 is estimated by $\sigma_{B_1}^2$, we have from (3.4.20),

$$\sigma_{B_1}^2 + \frac{\sigma^2}{N} = \frac{1}{n}\sum_{i=1}^{n} Y_i = \overline{Y}_n \text{ and then,}$$

$$\frac{b-1}{2N(\sigma_{B_1}^2 + \frac{\sigma^2}{N})^2}\sum_{i=1}^{n} Y_i - \frac{n(b-1)}{2N(\sigma_{B_1}^2 + \frac{\sigma^2}{N})} \text{ will be reduced to zero ,}$$

and (3.4.21) reduces to

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} Z_i = \overline{Z}_n \tag{3.4.22}$$

When $\sigma_B^2 = \sigma_{B_o}^2$, equating (3.4.21) to zero results in:

$$\begin{aligned}
0 &= \sigma^4\left[(b-1)\sum_{i=1}^{n} Y_i - n(b-1)(\sigma_{B_o}^2 + \frac{1}{N}\sigma^2)\right] + \\
&\quad b(N-1)\left[\sum_{i=1}^{n} Z_i - n\sigma^2\right] N(\sigma_{B_o}^2 + \frac{1}{N}\sigma^2)^2 \\
&= \sigma^4(b-1)\sum_{i=1}^{n} Y_i - n(b-1)\left(\sigma_{B_o}^2\sigma^4 + \frac{1}{N}\sigma^6\right) + \left[b(N-1)\sum_{i=1}^{n} Z_i - \right. \\
&\quad \left. bn(N-1)\sigma^2\right]\left[N\sigma_{B_o}^4 + \frac{1}{N}\sigma^4 + 2\sigma_{B_o}^2\sigma^2\right] \\
&= \sigma^4(b-1)\sum_{i=1}^{n} Y_i - n(b-1)\sigma_{B_o}^2\sigma^4 - \frac{n(b-1)}{N}\sigma^6 + bN(N-1)\sigma_{B_o}^4\sum_{i=1}^{n} Z_i + \\
&\quad \frac{b(N-1)}{N}\sigma^4\sum_{i=1}^{n} z_i + 2b(N-1)\sigma_{B_o}^2\sigma^2\sum_{i=1}^{n} Z_i - \\
&\quad bnN(N-1)\sigma^2\sigma_{B_o}^4 - \frac{bn(N-1)}{N}\sigma^6 - 2bn(N-1)\sigma_{B_o}^2\sigma^4 \\
&= \sigma^6\left[-\frac{n(b-1)}{N} - \frac{bn(N-1)}{N}\right] + \sigma^4\left[(b-1)\sum_{i=1}^{n} Y_i - n(b-1)\sigma_{B_o}^2 + \right. \\
&\quad \left. \frac{b(N-1)}{N}\sum_{i=1}^{n} Z_i - 2bn(N-1)\sigma_{B_o}^2\right] + \sigma^2\left[2b(N-1)\sigma_{B_o}^2\sum_{i=1}^{n} Z_i - \right. \\
&\quad \left. bnN(N-1)\sigma_{B_o}^4\right] + bN(N-1)\sigma_{B_o}^4\sum_{i=1}^{n} Z_i \\
&= \sigma^6 + \frac{N}{1-Nb}\left[(b-1)(\overline{Y}_n - \sigma_{B_o}^2) + \frac{b(N-1)}{N}(\overline{Z}_n - 2N\sigma_{B_o}^2)\right]\sigma^4 + \\
&\quad \frac{N}{1-Nb}\left[b(N-1)\left(2\overline{Z}_n\sigma_{B_o}^2 - N\sigma_{B_o}^4\right)\right]\sigma^2 + \frac{N}{1-Nb}(bN(N-1)\overline{Z}_n\sigma_{B_o}^4)
\end{aligned}$$

which reduces to solving a cubic equation in σ^2. At $\hat{\sigma} = 0$, the right hand side becomes $\frac{N}{1-Nb}(bN(N-1)\overline{Z}_n\sigma^4_{B_o})$ which is less than zero, as (1-Nb) < 0, and by taking sufficiently large $\hat{\sigma}$, as $\hat{\sigma}^6$ dominates the others , the right hand side will be positive and hence one can be sure of the existence of a positive real root.

Then using the given informations and the method of solving the roots of cubic equation given in Appendix A , we will have for the coefficients of the equation $X^3 + C_1X^2 + C_2X + C_3 = 0$:

$$
\begin{aligned}
C_1 &= \frac{N}{1-Nb}\left[(b-1)(\overline{Y}_n - \sigma^2_{B_o}) + \frac{b(N-1)}{N}(\overline{Z}_n - 2N\sigma^2_{B_o})\right] \\
C_2 &= \frac{N}{1-Nb}\left[b(N-1)\left(2\overline{Z}_n\sigma^2_{B_o} - N\sigma^4_{B_o}\right)\right] \quad \text{and} \\
C_3 &= \frac{N}{1-Nb}(bN(N-1)\overline{Z}_n\sigma^4_{B_o}) \qquad (3.4.23)
\end{aligned}
$$

Hence for a given $\sigma^2_{B_o}$ and the data, one can substitute the values, calculate the values of Q, R and D^* and determine the solution. If there were more than one positive real solution to the cubic equation one has to take the value which maximizes the likelihood given in equation (3.4.19).

Then using the result obtained in Appendix A, one can solve and get the solution to this equation which maximizes the likelihood.

Let this value be denoted by, $\hat{\sigma}^2_o$.

Hence for n observations (n classes in our case)

$$
\begin{aligned}
-2\sum_{i=1}^{n}\log f_{Y_i}(Y_i) &= \frac{(b-1)}{\sigma^2_B + \frac{1}{N}\sigma^2}\sum_{i=1}^{n} Y_i + n(b-1)\log 2(\sigma^2_B + \frac{1}{N}\sigma^2) + \sum_{i=1}^{n} F(Y_i) \\
&= \begin{cases} (b-1)\left(\frac{\sum_{i=1}^{n} Y_i}{\sigma^2_{B_o} + \frac{1}{N}\hat{\sigma}^2_o} + n\log 2(\sigma^2_{B_o} + \frac{1}{N}\hat{\sigma}^2_o)\right) + \sum_{i=1}^{n} F(Y_i) \\ n(b-1) + n(b-1)\log\frac{2}{n}\sum_{i=1}^{n} Y_i + \sum_{i=1}^{n} F(Y_i) \end{cases}
\end{aligned}
$$

where the first is under H_o and the second is under H_1 . Also,

$$
-2\sum_{i=1}^{n}\log f_{Z_i}(Z_i) = \frac{b(N-1)}{\sigma^2}\sum_{i=1}^{n} Z_i + nb(N-1)\log 2(\sigma^2) + \sum_{i=1}^{n} H(Z_i)
$$

$$= \begin{cases} \frac{b(N-1)}{\hat{\sigma}_o^2} \sum_{i=1}^n Z_i + nb(N-1)\log 2\hat{\sigma}_o^2 + \sum_{i=1}^n H(Z_i) \\ nb(N-1) + nb(N-1)\log \frac{2}{n} \sum_{i=1}^n Z_i + \sum_{i=1}^n H(Z_i) \end{cases}$$

under $\sigma_B^2 = \sigma_{B_o}^2$ and under estimation respectively. Then.

$$\begin{aligned} -2\log\lambda_n &= -2\ell(\sigma_{B_o}^2, \hat{\sigma}_o^2) + -2\ell(\hat{\sigma}_B^2, \hat{\sigma}^2) \\ &= (b-1)\left[\frac{\sum_{i=1}^n Y_i}{\sigma_{B_o}^2 + \frac{1}{N}\hat{\sigma}_o^2} - n\right] + n(b-1)\log\left(\frac{n(\sigma_{B_o}^2 + \frac{1}{N}\hat{\sigma}_o^2)}{\sum_{i=1}^n Y_i}\right) + \\ & \quad b(N-1)\left[\frac{\sum_{i=1}^n Z_i}{\hat{\sigma}_o^2} - n + n\log\left(\frac{n\hat{\sigma}_o^2}{\sum_{i=1}^n Z_i}\right)\right] \end{aligned} \tag{3.4.24}$$

Then the results in section 2.4.2 and similar procedure in the earlier section. lead to the following procedures.

Test 1

Start testing for n=2 and reject H_o whenever the result $-2\log\lambda_n \geq CV_1(\alpha, a)$ where $-2\log\lambda_n$ is the one obtained from equation (3.4.24) and $CV_1(\alpha, a)$ as defined in Test 1 of section 2.4.2 and conclude that a difference is detected in favor of H_1 . If not, proceed testing after increasing the size by 1 (i.e. $n \to n+1$) and do the updated analysis. Repeat the procedure until the result $-2\log\lambda_n \geq CV_1(\alpha, a)$ is obtained for some n $\leq$ a or until getting n=a without obtaining the required inequality. In the later case support the claim that there is no evidence of difference in favor of H_o .

Test 2

Start testing for n=2 and reject H_o if $-2\frac{n}{a}\log\lambda_n \geq CV_2(\alpha)$ where $CV_2(\alpha)$ as defined in Test 2 of section 2.4.2 and conclude that a difference is detected in favor of H_1 . If not proceed testing after increasing the size by 1 (i.e. $n \to n+1$) and do the updated analysis. Repeat the procedure until the result $-2\frac{n}{a}\log\lambda_n \geq CV_2(\alpha)$ is obtained for some n $\leq$ a or until n=a without obtaining the required inequality. In the later case support the claim that there is no evidence of difference in favor of H_o .

3.4.4 Monitoring all the Variances Simultaneously

In this case we are to monitor all the variance components σ_A^2, σ_B^2 and σ^2 simultaneously and hence our number of parameters of interest is d=3.

The Hypotheses to be tested are:

H_o : $(\sigma_A^2, \sigma_B^2, \sigma^2) = (\sigma_{A_o}^2, \sigma_{B_o}^2, \sigma_o^2)$ vs.

H_1 : $(\sigma_A^2, \sigma_B^2, \sigma^2)$ as a group is not equal to $(\sigma_{A_o}^2, \sigma_{B_o}^2, \sigma_o^2)$.

Now consider the joint density of (U_i, Z_i, Y_i) , where

$$U_i = \frac{1}{bN}\sum_{j=1}^{b}\sum_{k=1}^{N} X_{ijk} \ , \quad Y_i = \frac{1}{b-1}\sum_{j=1}^{b}(V_{ij} - U_i)^2 = \sigma_{i.}^2 \ ,$$

$$Z_i = \frac{1}{b(N-1)}\sum_{j=1}^{b}\sum_{k=1}^{N}(X_{ijk} - V_{ij})^2 = \sigma_i^2 \text{ where } V_{ij} = \frac{1}{N}\sum_{k=1}^{N} X_{ijk}$$

Being independent of each other, their joint density will be $f(U_i, Z_i, Y_i) = f(U_i) * f(Z_i) * f(Y_i)$ and then the loglikelihood will be the sum of the individual loglikelihoods. i.e., for n observations:

$$\begin{aligned}
\ell(\mu, \sigma_A^2, \sigma_B^2, \sigma^2) &= \sum_{i=1}^{n}\log f(U_i) + \sum_{i=1}^{n}\log f(Z_i) + \sum_{i=1}^{n}\log f(Y_i) \\
&= -\frac{n}{2}\log(\sigma_A^2 + \frac{1}{b}\sigma_B^2 + \frac{1}{bN}\sigma^2) - \frac{\sum_{i=1}^{n}(U_i - \mu)^2}{2(\sigma_A^2 + \frac{1}{b}\sigma_B^2 + \frac{1}{bN}\sigma^2)} \\
&\quad -\frac{n(b-1)}{2}\log 2(\sigma_B^2 + \frac{1}{N}\sigma^2) - \frac{b-1}{2(\sigma_B^2 + \frac{1}{N}\sigma^2)}\sum_{i=1}^{n} Y_i \\
&\quad -\frac{nb(N-1)}{2}\log 2\sigma^2 - \frac{b(N-1)}{2\sigma^2}\sum_{i=1}^{n} Z_i + Q(U_i, Z_i, Y_i) \ ,
\end{aligned} \tag{3.4.25}$$

where $Q(U_i, Z_i, Y_i)$ is a function of the data only and doesn't depend on any of the parameters. In the present problem the nuisance parameter is only μ. The maximum likelihood estimates will be obtained as usual,

$$\frac{\partial \ell(\mu, \sigma_A^2, \sigma_B^2, \sigma^2)}{\partial \mu}|_{\mu=\hat{\mu}} = 0 \ \Rightarrow \hat{\mu} \ = \ \frac{1}{n}\sum_{i=1}^{n} U_i \ = \ \overline{U}_n$$

$$0 \ = \ \frac{\partial \ell(\hat{\mu}, \sigma_A^2, \sigma_B^2, \sigma^2)}{\partial \sigma_A^2}|_{\hat{\sigma}_A^2} \ \text{ evaluated at } \hat{\mu} = \overline{U}_n$$

$$= \frac{\sum_{i=1}^{n}(U_i - \overline{U}_n)^2}{2(\hat{\sigma}_A^2 + \frac{1}{b}\sigma_B^2 + \frac{1}{bN}\sigma^2)^2} - \frac{n}{2(\hat{\sigma}_A^2 + \frac{1}{b}\sigma_B^2 + \frac{1}{bN}\sigma_o^2)}$$

$$\Rightarrow \hat{\sigma}_A^2 + \frac{1}{b}\sigma_B^2 + \frac{1}{bN}\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(U_i - \overline{U}_n)^2 \qquad (3.4.26)$$

$$0 = \frac{\partial \ell(\hat{\mu}, \hat{\sigma}_A^2, \sigma_B^2, \sigma^2)}{\partial \sigma_B^2}|_{\hat{\sigma}_B^2} \quad \text{evaluated at } \hat{\mu} = \overline{U}_n$$

$$= \frac{\sum_{i=1}^{n}(U_i - \overline{U}_n)^2}{2b(\hat{\sigma}_A^2 + \frac{1}{b}\hat{\sigma}_B^2 + \frac{1}{bN}\sigma^2)^2} - \frac{n}{2b(\hat{\sigma}_A^2 + \frac{1}{b}\hat{\sigma}_B^2 + \frac{1}{bN}\sigma^2)} + \frac{b-1}{2(\hat{\sigma}_B^2 + \frac{1}{N}\sigma^2)^2}\sum_{i=1}^{n} Y_i$$

$$-\frac{n(b-1)}{2(\hat{\sigma}_B^2 + \frac{1}{N}\sigma^2)}$$

$$\Rightarrow \hat{\sigma}_B^2 + \frac{1}{N}\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

$$0 = \frac{\partial \ell(\hat{\mu}, \hat{\sigma}_A^2, \hat{\sigma}_B^2, \sigma^2)}{\partial \sigma^2}|_{\hat{\sigma}^2} \quad \text{evaluated at } \hat{\mu} = \overline{U}_n$$

$$= \frac{\sum_{i=1}^{n}(U_i - \overline{U}_n)^2}{2bN(\hat{\sigma}_A^2 + \frac{1}{b}\hat{\sigma}_B^2 + \frac{1}{bN}\hat{\sigma}^2)^2} - \frac{n}{2bN(\hat{\sigma}_A^2 + \frac{1}{b}\hat{\sigma}_B^2 + \frac{1}{bN}\hat{\sigma}^2)} + \frac{(b-1)\sum_{i=1}^{n} Y_i}{2N(\hat{\sigma}_B^2 + \frac{1}{N}\hat{\sigma}^2)^2}$$

$$-\frac{n(b-1)}{2N(\hat{\sigma}_B^2 + \frac{1}{N}\hat{\sigma}^2)} + \frac{b(N-1)}{2\hat{\sigma}^4}\sum_{i=1}^{n} Z_i - \frac{nb(N-1)}{2\hat{\sigma}^2}$$

$$\Rightarrow \hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n} Z_i \qquad (3.4.27)$$

Then:

$$-2\log\lambda_n = -2\ell(\hat{\mu}, \sigma_{A_o}^2, \sigma_{B_o}^2, \sigma_o^2) + 2\ell(\hat{\mu}, \hat{\sigma}_A^2, \hat{\sigma}_B^2, \hat{\sigma}^2)$$

$$= n\log(\sigma_{A_o}^2 + \frac{1}{b}\sigma_{B_o}^2 + \frac{1}{bN}\sigma_o^2) + \frac{\sum_{i=1}^{n}(U_i - \overline{U}_n)^2}{(\sigma_{A_o}^2 + \frac{1}{b}\sigma_{B_o}^2 + \frac{1}{bN}\sigma_o^2)}$$

$$+n(b-1)\log 2(\sigma_{B_o}^2 + \frac{1}{N}\sigma_o^2) + \frac{b-1}{(\sigma_{B_o}^2 + \frac{1}{N}\sigma_o^2)}\sum_{i=1}^{n} Y_i$$

$$+nb(N-1)\log 2\sigma_o^2 + \frac{b(N-1)}{\sigma_o^2}\sum_{i=1}^{n} Z_i - n -$$

$$n\log\left(\frac{1}{n}\sum_{i=1}^{n}(U_i - \overline{U}_n)^2\right) - n(b-1)\left(1 + \log\left(\frac{2}{n}\sum_{i=1}^{n} Y_i\right)\right) -$$

$$nb(N-1)\left(1 + \log\left(\frac{2}{n}\sum_{i=1}^{n} Z_i\right)\right) \qquad (3.4.28)$$

$$= \left(\frac{\sum_{i=1}^{n}(U_i - \overline{U}_n)^2}{\sigma_{A_o}^2 + \frac{1}{b}\sigma_{B_o}^2 + \frac{1}{bN}\sigma_o^2} - n \right) + n \log \left(\frac{n(\sigma_{A_o}^2 + \frac{1}{b}\sigma_{B_o}^2 + \frac{1}{bN}\sigma_o^2)}{\sum_{i=1}^{n}(U_i - \overline{U}_n)^2} \right) +$$
$$\frac{b-1}{\sigma_{B_o}^2 + \frac{1}{N}\sigma_o^2} \sum_{i=1}^{n} Y_i + n(b-1) \left(-1 + \log \left(\frac{n(\sigma_{B_o}^2 + \frac{1}{N}\sigma_o^2)}{\sum_{i=1}^{n} Y_i} \right) \right) +$$
$$b(N-1) \left(\frac{\sum_{i=1}^{n} Z_i}{\sigma_o^2} - n + n \log \left(\frac{n\sigma_o^2}{\sum_{i=1}^{n} Z_i} \right) \right) \quad (3.4.29)$$

Now once the data and hypothesized values of the variance components are available, then one can apply TEST1 procedure to test the hypothesis. That is,

Test 1

We will start testing at n=2 and reject whenever the result $-2\log\lambda_n \geq CV_1(\alpha, a)$ where $-2\log\lambda_n$ is the one obtained from equation (3.4.29) and $CV_1(\alpha, a)$ as defined in Test 1 of section 2.4.2 and conclude that a difference is detected in favor of H_1 . If not, proceed testing after increasing the size by 1 (i.e. $n \to n+1$) and do the updated analysis. Repeat the procedure until the result $-2\log\lambda_n \geq CV_1(\alpha, a)$ is obtained for some n $\leq$ a or until getting n=a without obtaining the required inequality. In the later case support the claim that there is no evidence of difference in favor of H_o .

Chapter 4

Simulation on Some Selected Procedures

Given the Nested Random Effect Model,

$$X_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \epsilon_{ijk} \qquad \begin{cases} i = 1, 2, ..., a \\ j = 1, 2, ..., b \\ k = 1, 2, ..., N \end{cases}$$

where μ is a constant, with the basic assumptions given in equation (1.1.1). Consider testing the hypotheses:
$H_o: \ \mu = \mu_o = 0 \ vs. \ H_1: \ \mu \neq \mu_o$.

4.1 Monitoring the Mean

According to our result in the previous chapters $U_i = \frac{1}{bN}\sum_{j=1}^{b}\sum_{k=1}^{N} X_{ijk}$ follows a normal distribution with mean μ and nuisance parameter
$\eta = \sigma_A^2 + \frac{1}{b}\sigma_B^2 + \frac{1}{bN}\sigma^2$.
The maximum likelihood estimates based on n observations (n classes) found in Chapter 3 were:

$$\hat{\mu}_n = \overline{U}_n = \frac{1}{n}\sum_{i=1}^{n} U_i \quad , \quad \hat{\eta}_n = \frac{1}{n}\sum_{i=1}^{n}(U_i - \overline{U}_n)^2 \ .$$

The test statistic is

$$-2\log\lambda_n \quad = \quad n\log\left(\frac{\sum_{i=1}^{n}(U_i-\mu_o)^2}{\sum_{i=1}^{n}(U_i-\overline{U}_n)^2}\right). \qquad (4.1.1)$$

For $\alpha = 0.05$, the critical values in TEST1 are $CV_1(0.05,30) = 9.9968$, $CV_1(0.05,50) = 10.2235$ and $CV_1(0.05,200) = 10.7530$. As d=1 we also get $CV_2(.05) = 5.0176$.
Using the test statistics in equation (4.1.1) and our test procedure, Tables 4.1, 4.2 and 4.3 present results obtained from Monte Carlo experiments, where each case was repeated 2000 times, truncated at a=30 , a=50 and a=200, respectively. Also Tables 4.4 , 4.5 and 4.6 present results obtained in Monte Carlo experiments using TEST2, where each case was repeated 2000 times and truncated at a=30 , a=50 and a=200, respectively. In the tables ASN stands for Average Sample Number, SD is the observed standard deviation of the distribution of the sample numbers.

The data was generated for values $\sigma^2 = \frac{3200}{4900}$, $\sigma_A^2 = \frac{3600}{4900}$, $\sigma_B^2 = \frac{1800}{4900}$
alternative nean $\mu_1 \in \{0.0, 0.2, 0.4, 0.6, 0.8, 1.0\}$,
N (the number of random data in each subclass)=4, and
b (the number of random subclass in each class) =2 under the nested random effects model defined earlier, for each table. The last column in each table presents the power of the test under fixed sample size test using ASN.

μ_1	**ASN**	**SD**	**Observed POWER**	**Fixed sample size Test with power**
0.0	27.7510	10.3991	0.0860	n=28, 0.0500
0.2	27.4015	10.9384	0.1175	n=28, 0.1849
0.4	25.2050	12.9820	0.2965	n=26, 0.5318
0.6	21.3875	13.8748	0.5970	n=22, 0.8035
0.8	15.9655	12.9500	0.8815	n=16, 0.8925
1.0	12.3470	10.2014	0.9785	n=13, 0.9501

Table 4.1: Simulation result to monitor Mean under TEST1, truncated at a=30

μ_1	**ASN**	**SD**	**Observed POWER**	**Fixed sample size Test with power**
0.0	45.8395	13.2104	0.0670	n=46, 0.0500
0.2	45.1280	13.4327	0.1430	n=45, 0.2687
0.4	38.4920	16.2032	0.4700	n=39, 0.7047
0.6	26.9355	15.4043	0.8765	n=27, 0.8765
0.8	17.234	10.9192	0.9925	n=18, 0.9242
1.0	12.4830	7.3201	1.0000	n=13, 0.9501

Table 4.2: Simulation result to monitor Mean under TEST1, truncated at a=50

μ_1	**ASN**	**SD**	**Observed POWER**	**Fixed sample size Test with power**
0.0	186.9135	120.2199	0.0715	n=187, 0.0500
0.2	149.5335	165.8609	0.4890	n=150, 0.6878
0.4	61.4905	101.7056	0.9960	n=62, 0.8829
0.6	30.1200	47.3008	1.0000	n=30, 0.9076
0.8	18.9360	28.9855	1.0000	n=19, 0.9366
1.0	13.0680	18.4173	1.0000	n=13, 0.9501

Table 4.3: Simulation result to monitor Mean under TEST1, truncated at a=200

μ_1	**ASN**	**SD**	**Observed POWER**	**Fixed sample size Test with power**
0.0	29.6595	2.6697	0.0510	n=30, 0.0500
0.2	28.9200	4.3811	0.1695	n=29, 0.1898
0.4	25.8790	7.5154	0.5420	n=26, 0.5319
0.6	21.6580	8.1592	0.8545	n=22, 0.8035
0.8	17.4565	6.7909	0.9790	n=18, 0.9242
1.0	14.8415	5.1649	0.9990	n=15, 0.9721

Table 4.4: Simulation result to monitor mean under TEST2 , truncated at a=30

μ_1	**ASN**	**SD**	**Observed POWER**	**Fixed sample size Test with power**
0.0	49.4310	3.1188	0.0530	n=50, 0.0500
0.2	46.9295	6.5416	0.2680	n=47, 0.2784
0.4	38.4260	9.8619	0.7620	n=39, 0.7047
0.6	28.9720	8.2639	0.9775	n=29, 0.8982
0.8	22.7350	5.5886	1.0000	n=23, 0.9697
1.0	19.1535	4.1116	1.0000	n=19, 0.9918

Table 4.5: Simulation result to monitor Mean under TEST2, truncated at a=50

μ_1	**ASN**	**SD**	**Observed POWER**	**Fixed sample size Test with power**
0.0	198.1330	21.3473	0.0445	n=198, 0.0500
0.2	149.6220	82.2036	0.7670	n=150, 0.6878
0.4	82.5100	44.9043	1.0000	n=83, 0.9539
0.6	57.6825	23.5122	1.0000	n=58, 0.9955
0.8	45.1850	15.2607	1.0000	n=45, 0.9997
1.0	38.2305	11.2344	1.0000	n=39, 1.0000

Table 4.6: Simulation result to monitor Mean under TEST2, truncated at a=200

Power calculation for fixed sample size test of mean

For a sample $U_i \sim N(\mu_o, \eta)$ and sample size n, we have, $\overline{U}_n \sim N(\mu_o, \frac{\eta}{n})$ and hence for a type I error α, type II error β and alternate mean μ_1, we have the relation:

$$\begin{aligned}
\alpha &= P(\text{reject } H_o | H_o \text{ is true }), \quad \text{which then implies} \\
1-\alpha &= P(\text{accept } H_o | H_o \text{ is true }) \\
1-\alpha &= P_{H_O}\left(Z_{\alpha/2} < \sqrt{n}\frac{(\overline{U}_n - \mu_o)}{\sqrt{\eta}} < Z_{(1-\alpha/2)} \right)
\end{aligned}$$

The power of the test is $1-\beta$, and hence

$$\begin{aligned}
\beta &= P(\text{retain } H_o | H_1 \text{ is true }) \\
\beta &= P_{H_1}\left(Z_{\alpha/2} < \sqrt{n}\frac{(\overline{U}_n - \mu_o)}{\sqrt{\eta}} < Z_{(1-\alpha/2)} \right) \\
&= P_{H_1}\left(Z_{\alpha/2} < \sqrt{n}\frac{(\overline{U}_n - \mu_1)}{\sqrt{\eta}} + \sqrt{n}\frac{(\mu_1 - \mu_0)}{\sqrt{\eta}} < Z_{(1-\alpha/2)} \right) \\
&= P_{H_1}\left(-\sqrt{n}\frac{\mu_1 - \mu_o}{\sqrt{\eta}} + Z_{\alpha/2} < \sqrt{n}\frac{(\overline{U}_n - \mu_1)}{\sqrt{\eta}} < -\sqrt{n}\frac{\mu_1 - \mu_o}{\sqrt{\eta}} + Z_{(1-\alpha/2)} \right) \\
&= P_{H_1}\left(-\sqrt{n}\frac{\mu_1 - \mu_o}{\sqrt{\eta}} + Z_{\alpha/2} < Z_{H_1} < -\sqrt{n}\frac{\mu_1 - \mu_o}{\sqrt{\eta}} + Z_{(1-\alpha/2)} \right) \qquad (4.1.2)
\end{aligned}$$

For $\mu_o = 0, \eta = 1, \alpha = 0.05$, and using the above formula one can calculate the power of a test for different sample size. For TEST1 and TEST2, truncated at a=30, Tables 4.7 and Table 4.8 present the results obtained of the power for different sample size respectively.

In a similar fashion one can calculate the other powers and compare them with the current test results. Routine calculations are omitted from this paper and for simplicity the results are presented side by side in the table of simulation results.

Table 4.9 reproduces the simulation results of Siskind (1964), on the two-tailed t test. Other tests suggested by Cox and Bartlet (1972) for the sequential testing of composite hypotheses are also included in this table as found in Joanes(1972).

μ_1	n	$-\mu_1\sqrt{n} + Z_{\alpha/2}$	$-\mu_1\sqrt{n} + Z_{(1-\alpha/2)}$	β	POWER
0.0	28	-1.9600	1.9600	0.9500	0.0500
0.2	28	-3.0183	0.9017	.8151	0.1849
0.4	26	-3.9996	-0.0796	0.4682	0.5318
0.6	22	-4.7742	-0.8542	0.1965	0.8035
0.8	16	-5.1600	-1.2400	0.1075	0.8925
1.0	13	-5.5656	-1.6456	0.0499	0.9501

Table 4.7: Power of a fixed sample size test of Mean that corresponds to TEST1.

μ_1	n	$-\mu_1\sqrt{n} + Z_{\alpha/2}$	$-\mu_1\sqrt{n} + Z_{(1-\alpha/2)}$	β	POWER
0.0	30	-1.9600	1.9600	0.9500	0.0500
0.2	29	-3.0370	0.8829	.8102	0.1892
0.4	26	-3.9996	-0.0796	0.4682	0.5318
0.6	22	-4.7742	-0.8542	0.1965	0.8035
0.8	18	-5.3541	-1.4341	0.0758	0.9242
1.0	15	-5.8330	-1.9130	0.0279	0.9721

Table 4.8: Power of a fixed sample size test of Mean that corresponds to TEST2.

μ_1	Siskind	Barnard	Wald	Bartlett	Cox
0.0	0.0400	0.0250	0.0150	0.036	0.1400
0.2	0.0683				
0.4	0.2543				
0.6	0.6371				
0.8	0.8867				
1.0	0.9500	0.9630	0.9650	0.9650	0.9640

Table 4.9: Power of different tests of the mean

Summary

The most striking result of TEST2 is that, the test statistics gave a probability of rejecting H_o : $\mu = 0$, when in fact $\mu = 0$ is true, which is very close to the nominal chosen value 0.05, among all other tests presented so far. Furthermore, TEST2, the weighted version of the test, is more sensitive for any amount of departures from H_o and hence gave us strong power, starting from $\mu = 0.4$ even for an early truncation of a=30 and 50. This power also gets close to unity very fast, as the departure from the null hypothesis increases, compared to others. Therefore, to control small rates of false alarms in the process control test, and at the same time to get a better power compared to others, TEST2 will be appropriate and efficient test to use.

For large truncation size, say a=50, TEST1 also gave greater power when compared to fixed sample size test, results of Siskind and all others listed in Table 4.9, as the departure from the null hypothesis increases. Figures 4.1 and 4.2 show the plots of powers TEST1 and TEST2 in monitoring the Mean with the respective fixed sample size test.

Although the comments were on a trial using just one set of hypotheses, whose random variable has variance of unity, some results using different variance reveal a very similar pattern to the proportion of mean to standard deviation. We therefore do not give details of the results for these trials.

4.2 Monitoring Variance

4.2.1 Monitoring Error Variance σ^2

The hypotheses to be tested are:

$$H_o : \sigma^2 = \sigma_o^2 = 1 \; vs. \quad H_1 : \sigma^2 \neq \sigma_o^2 \tag{4.2.1}$$

From the previous chapter considering the natural estimator in the i^{th} class:

$$\hat{\sigma}_i^2 = Z_i = \frac{1}{b(N-1)} \sum_{j=1}^{2} \sum_{k=1}^{4} (X_{ijk} - \overline{X}_{ij.})^2 \quad ,$$

with the help of standard theory it is found that:

$$\hat{\sigma}_i^2 \sim \frac{\sigma^2}{b(N-1)} \chi^2(b(N-1)) \; .$$

where $\chi^2(b(N-1))$ stands for a chi-square distribution with b(N-1) degrees of freedom and the maximum likelihood estimate of σ^2 was found to be

$$\hat{\sigma}_n^2 = \frac{1}{nb(N-1)} \sum_{i=1}^{n} \sum_{j=1}^{2} \sum_{k=1}^{4} (X_{ijk} - \overline{X}_{ij.})^2 = \frac{1}{n} \sum_{i=1}^{n} Z_i \; , \tag{4.2.2}$$

with the test statistic

$$-2 \log \lambda_n = nb(N-1) \left[\log \left(\frac{\sigma_o^2}{\overline{Z}_n} \right) + \frac{\overline{Z}}{\sigma_o^2} - 1 \right] \quad , \tag{4.2.3}$$

where $\overline{Z}_n = \hat{\sigma}_n^2$ as defined earlier.

Then using the test statistic in equation (4.2.3) and our test procedure, Tables 4.10 , 4.11 and 4.12 present results obtained in Monte Carlo experiments using TEST1 with $\alpha = 0.05$, where each case was repeated 2000 times, truncated at a=30 , a=50 and a=200, respectively. Also the Tables 4.13 , 4.14 and 4.15 present results obtained in Monte Carlo experiments using TEST2, where each case was repeated 2000 times, truncated at a=30 , a=50 and a=200, respectively. The data was generated for values, alternative error variance, $\sigma_1^2 \in \{0.6, 0.8, 1.0, 1.2, 1.4, 1.6\}$, $\sigma_A^2 = \frac{3600}{400}$, $\sigma_B^2 = \frac{900}{400}$, $N = 4$ and $b = 2$ for each table.

σ_1^2	**ASN**	**SD**	**Observed POWER**	**Fixed sample size Test with power**
0.6	15.8965	10.6900	0.9465	n=16, 0.9365
0.8	28.1085	7.4094	0.1710	n=28, 0.5037
1.0	29.8370	2.7594	0.0080	n=30, 0.0500
1.2	28.4975	6.8468	0.1300	n=29, 0.4224
1.4	22.1535	12.3141	0.6130	n=22, 0.7965
1.6	14.5325	11.8133	0.9265	n=15, 0.8916

Table 4.10: Simulation result to monitor Error Variance under TEST1, truncated at a=30

σ_1^2	**ASN**	**SD**	**Observed POWER**	**Fixed sample size Test with power**
0.6	15.5100	7.9413	0.9995	n=16, 0.9365
0.8	42.5210	12.2968	0.3880	n=43, 0.6976
1.0	49.6750	3.3808	0.0120	n=50, 0.0500
1.2	44.5390	11.4667	0.2590	n=45, 0.5829
1.4	26.7370	14.6458	0.8760	n=27, 0.8662
1.6	14.5265	9.0479	0.9980	n=15, 0.8916

Table 4.11: Simulation result to monitor Error Variance under TEST1, truncated at a=50

σ_1^2	ASN	SD	Observed POWER	Fixed sample size Test with power
0.6	16.1745	20.3025	1.0000	n=16, 0.9365
0.8	72.1060	102.7921	0.9890	n=72, 0.9017
1.0	197.6750	49.5130	0.0155	n=198, 0.0500
1.2	95.8295	135.7224	0.9230	n=96, 0.8762
1.4	29.5585	46.8333	1.0000	n=30, 0.8968
1.6	15.3015	23.5816	1.0000	n=16, 0.9084

Table 4.12: Simulation result to monitor Error Variance under TEST1 , truncated at a=200

σ_1^2	ASN	SD	Observed POWER	Fixed sample size Test with power
0.6	15.6895	5.8040	0.9950	n=16, 0.9365
0.8	26.4905	6.9996	0.4985	n=27, 0.4885
1.0	29.7125	2.3565	0.0465	n=30, 0.0500
1.2	27.2970	6.4968	0.3805	n=28, 0.4112
1.4	20.6510	9.0067	0.8455	n=21, 0.7793
1.6	15.2410	6.9715	0.9900	n=16, 0.9084

Table 4.13: Simulation result to monitor Error Variance mean under TEST2, truncated at a=30

σ_1^2	ASN	SD	Observed POWER	Fixed sample size Test with power
0.6	20.0900	4.7553	1.0000	n=20, 0.9763
0.8	39.5925	9.6213	0.7160	n=40, 0.6642
1.0	49.6380	2.2330	0.0405	n=50, 0.0500
1.2	42.1965	9.4602	0.5785	n=42, 0.5555
1.4	27.5705	9.1438	0.9695	n=28, 0.8772
1.6	19.4035	5.9165	1.0000	n=20, 0.9543

Table 4.14: Simulation result to monitor Error Variance under TEST2, truncated at a=50

σ_1^2	ASN	SD	Observed POWER	Fixed sample size Test with power
0.6	39.2035	12.8177	1.0000	n=40, 0.9999
0.8	85.1445	45.9506	0.9995	n=85, 0.9437
1.0	197.9440	23.0165	0.0445	n=198, 0.0500
1.2	98.6710	63.1420	0.9910	n=99, 0.8855
1.4	53.2760	26.8816	1.0000	n=54, 0.9892
1.6	37.3845	16.8004	1.0000	n=38, 0.9985

Table 4.15: Simulation result to monitor Error Variance under TEST2, truncated at a=200

Power for fixed sample size test of Variance

For a sample $\hat{\sigma}_i^2 = Z_i = \frac{1}{b(N-1)} \sum_{j=1}^{b} \sum_{k=1}^{N} (X_{ijk} - \overline{X}_{ij.})^2$;
we already have the result,

$$\frac{vZ_i}{\sigma^2} \sim \chi_v^2 = \text{Gamma}(v/2, 2) \quad \text{where} \quad v = b(N-1)$$

and Gamma(α, β) is given by

$$f(x|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} \;, \quad 0 \leq x < \infty \;, \quad \alpha, \beta > 0$$

and moment generating function, mgf :

$$M_x(t) = E\left[e^{tx}\right] = \left(\frac{1}{1-\beta t}\right)^\alpha \;, \quad t < \frac{1}{\beta}$$

Hence, if we put $X_i = \frac{vZ_i}{\sigma^2} \sim$ Gamma$(v/2, 2)$ then

$$M_{X_i}(t) = \left(\frac{1}{1-2t}\right)^{v/2} \;, \quad \text{and}$$

$$\begin{aligned}
M_{\frac{1}{n}\sum X_i}(t) &= E\left[e^{\frac{t}{n}\sum X_i}\right] = \prod_{i=1}^{n} E\left[e^{\frac{t}{n}X_i}\right] \quad \text{..... from independence} \\
&= \left[\frac{1}{1-\frac{2t}{n}}\right]^{\frac{nv}{2}} \quad \text{....... identical distribution} \\
&= mgf\left(\text{Gamma}\left(\frac{nv}{2}, \frac{2}{n}\right)\right) \qquad (4.2.4)
\end{aligned}$$

and by uniqueness property of moment generating functions,

$$\overline{X}_n = \frac{v\overline{Z}_n}{\sigma^2} = \frac{v\hat{\sigma}_n^2}{\sigma^2} \sim \text{Gamma}\left(\frac{nv}{2}, \frac{2}{n}\right)$$

σ_1^2	n	$\frac{1}{\sigma_1^2}G_{\alpha/2}\left(\frac{nv}{2},\frac{2}{n}\right)$	$\frac{1}{\sigma_1^2}G_{(1-\alpha/2)}\left(\frac{nv}{2},\frac{2}{n}\right)$	β	POWER
0.6	16	7.3732	13.0208	0.0635	0.9365
0.8	28	5.9823	9.1867	0.4963	0.5037
1.0	30	4.8247	7.3015	0.9500	0.0500
1.2	29	4.0048	6.1040	0.5776	0.4224
1.4	22	3.3146	5.3797	0.2035	0.7965
1.6	15	2.7353	4.9223	0.1084	0.8916

Table 4.16: Power of a fixed sample size test of Error Variance that corresponds to TEST1.

For a type I error α and type II error β, with alternate variance σ_1^2 ,

$$\begin{aligned}
\alpha &= P(\text{reject } H_o|H_o \text{ is true}) \text{ and then} \\
1-\alpha &= P(\text{retain } H_o|H_o \text{ is true}) \\
1-\alpha &= P_{H_O}\left(G_{\alpha/2} < \frac{v\hat{\sigma}_n^2}{\sigma_o^2} < G_{(1-\alpha/2)}\right) \\
\beta &= P_{H_1}\left(G_{\alpha/2} < \frac{v\hat{\sigma}_n^2}{\sigma_o^2} < G_{(1-\alpha/2)}\right) \\
&= P_{H_1}\left(\frac{\sigma_o^2}{\sigma_1^2}G_{\alpha/2} < \frac{v\hat{\sigma}_n^2}{\sigma_1^2} < \frac{\sigma_o^2}{\sigma_1^2}G_{(1-\alpha/2)}\right) \\
&= P_{H_1}\left(\frac{\sigma_o^2}{\sigma_1^2}G_{\alpha/2} < G_{H_1} < \frac{\sigma_o^2}{\sigma_1^2}G_{(1-\alpha/2)}\right) \qquad (4.2.5)
\end{aligned}$$

where $G_{\alpha/2}$ stands for the $(\alpha/2)^{th}$ quantile in the Gamma distribution with parameters, nv/2 and 2/n respectively. For $\sigma_o = 1, \alpha = 0.05$, and using the above formula one can calculate power for different sample size. For TEST1 and TEST2, truncated at a=30, Tables 4.16 and 4.17 gives the power for different sample sizes respectively.

σ_1^2	n	$\frac{1}{\sigma_1^2}G_{\alpha/2}\left(\frac{nv}{2},\frac{2}{n}\right)$	$\frac{1}{\sigma_1^2}G_{1-\alpha/2}\left(\frac{nv}{2},\frac{2}{n}\right)$	β	POWER
0.6	16	7.3732	13.0208	0.0635	0.9365
0.8	27	5.9561	9.2192	0.5115	0.4885
1.0	30	4.8247	7.3015	0.9500	0.0500
1.2	28	3.9882	6.1245	0.5888	0.4112
1.4	21	3.2933	5.4069	0.2207	0.7793
1.6	16	2.7650	4.8828	0.0916	0.9084

Table 4.17: Power of a fixed sample size test of Error Variance that corresponds to TEST2.

Summary

It is clear from the results of Tables 4.13, 4.14 and 4.15 that TEST2 gives a power, very close to the nominal level 0.05 under the null hypothesis $H_o : \mu_1 = \mu_o$. In addition TEST2 produces a strong power for controlling small rates of false alarms except for $\sigma_1^2 = 1.2$ truncated at a=30.

TEST1 is also more powerful than the fixed sample size test except for $|\sigma_1^2 - 1| \leq 0.2$ for an earlier truncation of a=30. As the sample size (class number in this case) increases TEST1 outperforms the fixed sample size test. There is also a rapid increase in the power of the test as truncation point is increasing. Figures 4.3 and 4.4 respectively show the plots of powers TEST1 and TEST2 in monitoring the Error Variance with the respective fixed sample size test.

4.2.2 Monitoring Class Variance σ_A^2

The hypotheses to be tested are:

$$H_o : \sigma_A^2 = \sigma_{A_o}^2 = 1 \; vs. \quad H_1 : \sigma_A^2 \neq \sigma_{A_o}^2 \tag{4.2.6}$$

Considering the statistics,
$U_i = \frac{1}{bN}\sum_{j=1}^{b}\sum_{k=1}^{N} X_{ijk} \sim N(\mu, \sigma_A^2 + \frac{1}{b}(\sigma_B^2 + \frac{\sigma^2}{N})) = N(\mu, \sigma_A^2 + \frac{1}{b}\xi)$ and
$Y_i = \hat{\sigma}_{i.}^2 = \frac{1}{b-1}\sum_{j=1}^{b}(\overline{X}_{ij.} - U_i)^2 \sim \frac{\xi}{b-1}\chi^2_{(b-1)}$, and using independence of U_i with Y_i , the mle's were,

$$\hat{\mu} = \frac{1}{n}\sum_{i-1}^{n} U_i = \overline{U}_n \;\; , \;\; \hat{\sigma}_A^2 + \frac{1}{b}\xi \;\; = \;\; \frac{\sum_{i=1}^{n}(U_i - \overline{U}_n)^2}{n}$$

$\hat{\xi} = \frac{1}{n}\sum_{i=1}^{n} Y_i = \overline{Y}_n$ when all were estimated and $\hat{\xi}_o$ a solution of the cubic equation:

$X^3 + C_1X^2 + C_2X + C_3 = 0$ under the condition $\sigma_A^2 = \sigma_{A_o}^2$, where

$$\begin{aligned} C_1 &= -\frac{\sum_{i=1}^{n}(U_i - \overline{U}_n)^2}{n} + (2b-1)\sigma_{A_o}^2 - \frac{b-1}{b}\overline{Y}_n \\ C_2 &= (b-1)(b\sigma_{A_o}^4 - 2\sigma_{A_o}^2\overline{Y}_n) \;\; and \\ C_3 &= -b(b-1)\sigma_{A_o}^4\overline{Y}_n \end{aligned}$$

The test statistics was also given by:

$$\begin{aligned} -2\log\lambda_n \;\; = \;\; & n\log\left(\frac{n(\sigma_{A_o}^2 + \frac{1}{b}\hat{\xi}_o)}{\sum_{i=1}^{n}(U_i - \overline{U}_n)^2}\right) + \left(\frac{\sum_{i=1}^{n}(U_i - \overline{U}_n)^2}{(\sigma_{A_o}^2 + \frac{1}{b}\hat{\xi}_o)} - n\right) + \\ & n(b-1)\log\left(\frac{n\hat{\xi}_o}{\sum_{i=1}^{n} Y_i}\right) + (b-1)\left(\frac{\sum_{i=1}^{n} Y_i}{\xi_o} - n\right) \end{aligned}$$

Tables 4.18, 4.19 and 4.20 present results obtained in Monte Carlo experiments using TEST1 with $\alpha = 0.05$. Each case was repeated 2000 times, truncated at a= 30 , a=50 and a=200, respectively. Also Tables 4.21, 4.22 and 4.23 presents results obtained in Monte Carlo experiments using TEST2, where each case was repeated 2000 times, truncated at a=30 , a=50 and a=200, respectively. The data was generated for values , alternate class variance, $\sigma_{A_1}^2 \in \{0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6\}$, $\sigma^2 = \frac{400}{3600}$, $\sigma_B^2 = \frac{900}{3600}$, $N = 4$ and $b = 2$ for each table, under the nested random effects model defined earlier.

$\sigma^2_{A_1}$	ASN	SD	Observed POWER	Fixed sample size Test with power
0.4	18.2985	13.2852	0.8350	n=19, 0.9993
0.6	24.8260	13.2768	0.3135	n=25, 0.9299
0.8	26.7400	12.1978	0.1375	n=27, 0.3536
1.0	28.1170	9.6688	0.0735	n=28, 0.0500
1.2	27.9565	9.8766	0.0875	n=28, 0.3238
1.4	25.7730	11.6182	0.3355	n=26 0.7612
1.6	15.0970	7.6089	1.0000	n=15, 0.8191

Table 4.18: Simulation result to monitor Class Variance under TEST1, truncated at a=30

$\sigma^2_{A_1}$	ASN	SD	Observed POWER	Fixed sample size Test with power
0.4	20.0505	11.2843	0.9980	n=20, 0.9996
0.6	36.3195	16.2752	0.6230	n=37, 0.9886
0.8	44.7735	14.0283	0.1415	n=45, 0.5473
1.0	46.1115	12.8196	0.0870	n=46, 0.0500
1.2	45.9070	12.5559	0.1185	n=46, 0.4769
1.4	33.6610	13.1885	0.9830	n=34, 0.8583
1.6	15.5250	5.3980	1.0000	n=16, 0.8413

Table 4.19: Simulation result to monitor Class Variance under TEST1, truncated at a=50

$\sigma^2_{A_1}$	**ASN**	**SD**	**Observed POWER**	**Fixed sample size Test with power**
0.4	20.7210	27.0376	1.0000	n=21, 0.9997
0.6	44.2560	55.4333	1.0000	n=45, 0.9969
0.8	147.2210	155.4789	0.6365	n=148, 0.9714
1.0	185.4255	124.9882	0.0795	n=186, 0.0500
1.2	142.4605	141.8622	0.8400	n=143, 0.9083
1.4	36.0960	32.6491	1.0000	n=36, 0.8763
1.6	16.4050	13.1140	1.0000	n=17, 0.8610

Table 4.20: Simulation result to monitor Class Variance under TEST1, truncated at a=200

$\sigma^2_{A_1}$	**ASN**	**SD**	**Observed POWER**	**Fixed sample size Test with power**
0.4	18.2870	3.8503	0.9985	n=19, 0.9993
0.6	24.8035	5.1314	0.7525	n=25, 0.9299
0.8	28.8155	3.2261	0.1800	n=29, 0.3768
1.0	29.7360	1.5816	0.0465	n=30, 0.0500
1.2	29.3150	2.4930	0.1220	n=30, 0.3418
1.4	22.8710	3.6144	1.0000	n=23, 0.7123
1.6	15.5925	2.0326	1.0000	n=16, 0.8413

Table 4.21: Simulation result to monitor Class Variance under TEST2, truncated at a=30

$\sigma^2_{A_1}$	**ASN**	**SD**	**Observed POWER**	**Fixed sample size Test with power**
0.4	24.0765	4.5515	1.0000	n=24, 0.9999
0.6	33.3525	7.2169	0.9965	n=34, 0.9817
0.8	47.0960	6.2025	0.2835	n=47, 0.5663
1.0	49.4595	2.9490	0.0470	n=50, 0.0500
1.2	47.5110	5.6652	0.2680	n=48, 0.4925
1.4	29.6210	4.4287	1.0000	n=30, 0.8153
1.6	20.1835	2.3396	1.0000	n=20, 0.9076

Table 4.22: Simulation result to monitor Class Variance under TEST2, truncated at a=50

$\sigma^2_{A_1}$	**ASN**	**SD**	**Observed POWER**	**Fixed sample size Test with power**
0.4	48.6420	10.9343	1.0000	n=49, 1.0000
0.6	69.7945	17.5511	1.0000	n=70, 1.0000
0.8	126.9985	41.4501	1.0000	n=127, 0.9462
1.0	197.8625	19.4951	0.0490	n=198, 0.0500
1.2	119.2130	29.9477	1.0000	n=120, 0.8557
1.4	59.7300	11.4738	1.0000	n=60, 0.9783
1.6	40.0495	6.4500	1.0000	n=40, 0.9954

Table 4.23: Simulation result to monitor Class Variance under TEST2, truncated at a=200

Power for fixed sample size test of Lot Variance

Consider the model with **n** classes, **b** subclasses and **N** data in each subclass.

$$E[X_{ijk} - \mu]^2 = \sigma_A^2 + \sigma_B^2 + \sigma^2 = \eta$$

and from standard theory it is shown that for $S_n^2 = \frac{\sum_{i=1}^{n}\sum_{j=1}^{b}\sum_{k=1}^{N}(X_{ijk}-\overline{X}_n)^2}{nbN-1}$, with

$$\frac{(nbN-1)S_n^2}{\eta} \sim \chi^2(nbN-1)$$

for a chi-square distribution with nbN-1 degrees of freedom.

For a type I error α and type II error β, with alternate variance $\sigma_{A_1}^2$, so that if we let, $\eta_o = \sigma_{A_o}^2 + \sigma_{B_o}^2 + \sigma_o^2$ and $\eta_1 = \sigma_{A_1}^2 + \sigma_{B_o}^2 + \sigma_o^2$, we have the relation:

$$\begin{aligned}
\alpha &= P(\text{reject } H_o|H_o \text{ is true}) \text{ and then} \\
1-\alpha &= P(\text{retain } H_o|H_o \text{ is true}) \\
&= P_{H_o}\left(\chi^2_{\alpha/2} < \frac{(nbN-1)S_n^2}{\eta_o} < \chi^2_{(1-\alpha/2)}\right) \\
\beta &= P_{H_1}\left(\chi^2_{\alpha/2} < \frac{(nbN-1)S_n^2}{\eta_o} < \chi^2_{(1-\alpha/2)}\right) \\
&= P_{H_1}\left(\frac{\eta_o}{\eta_1}\chi^2_{\alpha/2} < \frac{(nbN-1)S_n^2}{\eta_1} < \frac{\eta_o}{\eta_1}\chi^2_{(1-\alpha/2)}\right) \\
&= P\left(\frac{\eta_o}{\eta_1}\chi^2_{\alpha/2} < \chi^2_{H_1} < \frac{\eta_o}{\eta_1}\chi^2_{(1-\alpha/2)}\right) \qquad (4.2.7)
\end{aligned}$$

Simulation is done for $\sigma_o^2 = 4/36$, $\sigma_{A_o}^2 = 1$, $\sigma_{B_o}^2 = 9/36$, and $\alpha = 0.05$, then using the above formula one can calculate power for different sample sizes. For TEST1 and TEST2, truncated at a=30, Tables 4.24 and 4.25 give the power for different sample sizes respectively.

$\sigma^2_{A_1}$	n	$\frac{49/36}{\sigma^2_{A_1}+13/36}\chi^2_{(1-\alpha/2)}(8n-1)$	$\frac{49/36}{\sigma^2_{A_1}+13/36}\chi^2_{\alpha/2}(8n-1)$	β	POWER
0.4	19	334.2623	212.5800	0.0007	0.9993
0.6	25	339.8274	229.1757	0.0701	0.9299
0.8	27	322.3555	223.6618	0.6464	0.3536
1.0	28	266.2520	183.5300	0.9500	0.0500
1.2	28	247.3653	172.6975	0.6762	0.3238
1.4	26	171.8503	113.9461	0.2388	0.7612
1.6	15	111.1077	67.8077	0.1809	0.8191

Table 4.24: Power of a fixed sample size test of Class Variance that corresponds to TEST1.

$\sigma^2_{A_1}$	n	$\frac{49/36}{\sigma^2_{A_1}+13/36}\chi^2_{(1-\alpha/2)}(8n-1)$	$\frac{49/36}{\sigma^2_{A_1}+13/36}\chi^2_{\alpha/2}(8n-1)$	β	POWER
0.4	19	334.2616	212.5802	0.0007	0.9993
0.6	25	339.8274	229.1757	0.0701	0.9299
0.8	29	322.3555	223.6618	0.6232	0.3768
1.0	30	283.7128	198.0734	0.9500	0.0500
1.2	30	247.3653	172.6975	0.6582	0.3418
1.4	23	171.8503	113.9461	0.2877	0.7123
1.6	16	111.1077	67.8077	0.1587	0.8413

Table 4.25: Power of a fixed sample size test of Class Variance that corresponds to TEST2.

Summary

As in the earlier cases, TEST2 gave the power very close to 0.05, which is the nominal level, under the null hypothesis $H_O : \sigma_A^2 = \sigma_{A_o}^2$. We also observe from the tables that TEST2 gives more powerful result for all $|\sigma_{A_1}^2 - 1| > 0.2$, even at an earlier truncation of a=30, and 50. One also sees that by using truncation a=50, the power of the test at $\sigma_{A_1}^2 = 1.2$ and $\sigma_{A_1}^2 = 0.8$ is greater than the corresponding power by using truncation a=30 in both tests. And in fact, a dramatic increase in power and better performance in testing a change for any amount of departure from the null hypothesis was demonstrated in Tables 4.20 and 4.23 by increasing the truncation to a=200. Figures 4.5 and 4.6 respectively show the plots of powers TEST1 and TEST2 in monitoring the Class Variance with the respective fixed sample size test.

4.2.3 Monitoring Subclass Variance, σ_B^2

The hypotheses to be tested are:

$$H_o : \sigma_B^2 = \sigma_{B_o}^2 = 1 \ vs. \ H_1 : \sigma_B^2 \neq \sigma_{B_o}^2 \tag{4.2.8}$$

Considering the statistics,
$Z_i = \frac{1}{b(N-1)} \sum_{j=1}^{b} \sum_{k=1}^{N} (X_{ijk} - V_{ij})^2 \sim \frac{\sigma^2}{b(N-1)} \chi^2_{b(N-1)}$ and
$Y_i = \hat{\sigma}^2_{i.} = \frac{1}{b-1} \sum_{j=1}^{b} (\overline{X}_{ij.} - U_i)^2 \sim \frac{\xi}{b-1} \chi^2_{(b-1)}$, and independence of Y_i with Z_i , we found from chapter 3 that,

$$\hat{\sigma}_B^2 + \frac{\sigma^2}{N} = \frac{1}{n} \sum_{i-1}^{n} Y_i = \overline{Y}_n$$

$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} Z_i = \overline{Z}_n$ when all were estimated , and $\hat{\sigma}^2$ is a solution of the cubic equation:

$$X^3 + C_1 X^2 + C_2 X + C_3 = 0 \quad \text{under } \sigma_B^2 = \sigma_{B_o}^2 \text{ , where}$$

$$\begin{aligned}
C_1 &= \frac{N}{1 - Nb} \left[(b-1)(\overline{Y}_n - \sigma_{B_o}^2) + \frac{b(N-1)}{N} (\overline{Z}_n - 2N\sigma_{B_o}^2) \right] \\
C_2 &= \frac{N}{1 - Nb} \left[b(N-1) \left(2\overline{Z}_n \sigma_{B_o}^2 - N\sigma_{B_o}^4 \right) \right] \quad \text{and} \\
C_3 &= \frac{N}{1 - Nb} (bN(N-1)\overline{Z}_n \sigma_{B_o}^4)
\end{aligned}$$

The test statistics was also given by:

$$\begin{aligned}
-2 \log \lambda_n &= (b-1) \left[\frac{\sum_{i=1}^{n} Y_i}{\sigma_{B_o}^2 + \frac{1}{N}\hat{\sigma}_o^2} - n \right] + n(b-1) \log \left(\frac{n(\sigma_{B_o}^2 + \frac{1}{N}\hat{\sigma}_o^2)}{\sum_{i=1}^{n} Y_i} \right) + \\
& \quad b(N-1) \left[\frac{\sum_{i=1}^{n} Z_i}{\hat{\sigma}_o^2} - n + n \log \left(\frac{n\hat{\sigma}_o^2}{\sum_{i=1}^{n} Z_i} \right) \right]
\end{aligned} \tag{4.2.9}$$

Using the test statistics in equation 4.2.9 and the test procedure, Tables 4.26 , 4.27 and 4.28 present results obtained in Monte Carlo experiments using TEST1, where each case was repeated 2000 times, truncated at a= 30 , a=50 and a=200, respectively. Also the Tables 4.29, 4.30 and 4.31 presents results obtained in Monte Carlo experiments using TEST2, where each case was repeated 2000 times, truncated at a=30, a=50 and a=200, respectively. The data was generated for values $\sigma^2 = \frac{400}{900}$, $\sigma_A^2 = \frac{3600}{900}$, alternate subclass variance, $\sigma_{B_1}^2 \in \{0.4, 0.6, 1.0, 1.4, 2.0, 2.2, 2.6, 3, 4\}$ $N = 4$ and $b = 2$, for each table.

$\sigma^2_{B_1}$	**ASN**	**SD**	**Observed POWER**	**Fixed sample size Test with power**
0.4	26.1105	9.5252	0.3835	n=26, 0.7363
0.6	28.8570	6.2848	0.0880	n=29, 0.3256
1.0	29.6935	3.6966	0.0160	n=30, 0.0500
1.4	29.1520	5.4425	0.0635	n=29, 0.2554
2.0	24.4365	11.9463	0.4210	n=25, 0.6706
2.2	22.1080	13.0637	0.5690	n=22, 0.7247
2.6	18.3180	13.3729	0.7630	n=19, 0.8140
3.0	15.0940	12.7399	0.8780	n=15, 0.8311
4.0	9.9460	9.6586	0.9810	n=10, 0.8527

Table 4.26: Simulation result to monitor Subclass Variance under TEST1, truncated at a=30

$\sigma^2_{B_1}$	**ASN**	**SD**	**Observed POWER**	**Fixed sample size Test with power**
0.4	35.9815	19.3266	0.7055	n=36, 0.8742
0.6	46.2115	13.9853	0.1855	n=47, 0.5195
1.0	49.1890	8.1295	0.0210	n=49, 0.0500
1.4	47.1630	13.0779	0.1170	n=47, 0.3607
2.0	33.0065	22.8966	0.6885	n=33, 0.7728
2.2	28.7400	22.1780	0.8150	n=29, 0.8210
2.6	21.7720	19.4630	0.9390	n=22, 0.8591
3.0	16.6640	16.3753	0.9815	n=17, 0.8677
4.0	10.1720	10.6175	0.9980	n=10, 0.8527

Table 4.27: Simulation result to monitor Subclass Variance under TEST1, truncated at a=50

$\sigma^2_{B_1}$	**ASN**	**SD**	**Observed POWER**	**Fixed sample size Test with power**
0.4	42.7985	39.3245	1.0000	n=43, 0.9395
0.6	109.386	94.8817	0.8935	n=110, 0.9010
1.0	196.5925	42.1943	0.0225	n=197, 0.0500
1.4	146.3060	105.3196	0.5855	n=147, 0.7681
2.0	45.7125	54.6818	0.9995	n=46, 0.8800
2.2	34.9365	42.5916	1.0000	n=35, 0.8780
2.6	23.2785	27.6899	1.0000	n=24, 0.8832
3.0	17.3575	21.1367	1.0000	n=18, 0.8830
4.0	10.9345	13.7553	1.0000	n=11, 0.8784

Table 4.28: Simulation result to monitor Subclass Variance under TEST1, truncated at a=200

$\sigma^2_{B_1}$	**ASN**	**SD**	**Observed POWER**	**Fixed sample size Test with power**
0.4	23.6570	7.4960	0.7820	n=24, 0.6936
0.6	28.1625	5.3600	0.2975	n=28, 0.3142
1.0	29.8170	1.6957	0.0445	n=30, 0.0500
1.4	28.6525	4.8899	0.2020	n=29, 0.2554
2.0	23.6915	9.1418	0.6685	n=24, 0.6555
2.2	21.7100	9.7032	0.7700	n=22, 0.7247
2.6	18.6290	9.6042	0.8975	n=19, 0.8140
3.0	16.2370	8.6067	0.9610	n=17, 0.8677
4.0	12.4155	6.8910	0.9935	n=13, 0.9177

Table 4.29: Simulation result to monitor Subclass Variance under TEST2, truncated at a=30

$\sigma^2_{B_1}$	**ASN**	**SD**	**Observed POWER POWER**	**Fixed sample size Test with power**
0.4	32.4025	8.3724	0.9610	n=33, 0.8505
0.6	43.4190	8.4534	0.5330	n=44, 0.4893
1.0	49.5300	2.6212	0.0465	n=50, 0.0500
1.4	46.4325	7.1609	0.2965	n=47, 0.3607
2.0	33.6800	11.1612	0.8555	n=34, 0.7834
2.2	29.6645	10.2590	0.9415	n=30, 0.8319
2.6	24.2290	8.9002	0.9890	n=25, 0.8938
3.0	20.4845	7.3742	0.9975	n=21, 0.9197
4.0	15.5085	5.3884	1.0000	n=16, 0.9547

Table 4.30: Simulation result to monitor Subclass Variance under TEST2, truncated at a=50

$\sigma^2_{B_1}$	**ASN**	**SD**	**Observed POWER POWER**	**Fixed sample size Test with power**
0.4	65.2100	29.2752	1.0000	n=66, 0.9944
0.6	107.9625	74.2668	0.9895	n=108, 0.8952
1.0	198.1770	28.0004	0.0405	n=198, 0.0500
1.4	137.4665	106.1866	0.8455	n=138, 0.7436
2.0	66.4150	49.0975	1.0000	n=67, 0.9601
2.2	57.6490	40.6837	1.0000	n=58, 0.9745
2.6	45.6520	30.4387	1.0000	n=46, 0.9871

Table 4.31: Simulation result to monitor Subclass Variance under TEST2, truncated at a=200

Power for fixed sample size test of Subclass Variance

For a sample $\hat{\sigma}^2_{i.} = Y_i = \frac{1}{b-1}\sum_{j=1}^{b}(\overline{X}_{ij.} - \overline{U}_i)^2$, where $\overline{X}_{ij.} = \frac{1}{N}\sum_{k=1}^{N} X_{ijk}$ $U_i = \frac{1}{b}\sum_{j=1}^{b}\overline{X}_{ij.}$ we already have the result,

$$\frac{(b-1)Y_i}{\xi} \sim \chi^2_{(b-1)} = \text{Gamma}((b-1)/2, 2) \text{ where } \xi = \sigma^2_B + \sigma^2/N$$

Then using a moment generating function as we did in earlier sections applied on the estimator $\hat{\sigma}^2_{.} = \overline{Y}_n$ we arrive at:

$$\frac{(b-1)\overline{Y}_n}{\xi} = \frac{(b-1)\hat{\sigma}^2_{.}}{\xi} \sim \text{Gamma}\left(\frac{n(b-1)}{2}, \frac{2}{n}\right) \quad (4.2.10)$$

Using type I error α and type II error β , with variance $\sigma^2_{B_1}$ under the alternative, and with $\xi_o = \sigma^2_{B_o} + \frac{\sigma^2_o}{N}$, and $\eta_1 = \sigma^2_{B_1} + \frac{\sigma^2_o}{N}$ we will have the relation:

$$\begin{aligned} \alpha &= P(\text{reject } H_o|H_o \text{ is true}) \text{ and then} \\ 1-\alpha &= P(\text{retain } H_o|H_o \text{ is true}) \\ 1-\alpha &= P_{H_o}\left(G_{\alpha/2} < \frac{(b-1)\hat{\sigma}^2_{.}}{\xi_o} < G_{(1-\alpha/2)}\right) \end{aligned}$$

$\sigma^2_{B_1}$	n	$\frac{10/9}{\sigma^2_{B_1}+1/9}G_{\alpha/2}\left(\frac{n}{2},\frac{2}{n}\right)$	$\frac{10/9}{\sigma^2_{B_1}+1/9}G_{1-\alpha/2}\left(\frac{n}{2},\frac{2}{n}\right)$	β	POWER
0.4	26	1.1575	3.5053	0.2637	0.7363
0.6	29	0.8646	2.4635	0.6744	0.3256
1.0	30	0.5597	1.5660	0.9500	0.0500
1.4	29	0.4069	1.1593	0.7446	0.2554
2.0	25	0.2762	0.8557	0.3294	0.6706
2.2	22	0.2400	0.8038	0.2753	0.7247
2.6	19	0.1921	0.7086	0.1860	0.8140
3.0	15	0.1491	0.6545	0.1689	0.8311
4.0	10	0.0878	0.5536	0.1473	0.8527

Table 4.32: Power of a fixed sample size test of Subclass Variance that corresponds to TEST1.

$$\begin{aligned}
\beta &= P_{H_1}\left(G_{\alpha/2} < \frac{(b-1)\hat{\sigma}^2_{\cdot}}{\xi_o} < G_{(1-\alpha/2)}\right) \\
&= P_{H_1}\left(\frac{\xi_o}{\xi_1}G_{\alpha/2} < \frac{(b-1)\hat{\sigma}^2_{\cdot}}{\xi_1} < \frac{\xi_o}{\xi_1}G_{(1-\alpha/2)}\right) \\
&= P_{H_1}\left(\frac{\xi_o}{\xi_1}G_{\alpha/2} < G_{H_1} < \frac{\xi_o}{\xi_1}G_{(1-\alpha/2)}\right) \qquad (4.2.11)
\end{aligned}$$

where $G_{\alpha/2}$ stands for the $(\alpha/2)^{th}$ quantile of the Gamma distribution with parameters, n/2 and 2/n respectively. For $\sigma^2_{B_o} = 1$, $\sigma^2_o = 4/9$ $\alpha = 0.05$, and using the above formula one can calculate power for different sample sizes. For TEST1 and TEST2, truncated at n=30, Tables 4.32 and 4.33 give the power.

$\sigma^2_{B_1}$	n	$\frac{10/9}{\sigma^2_{B_1}+1/9}G_{\alpha/2}\left(\frac{n}{2},\frac{2}{n}\right)$	$\frac{10/9}{\sigma^2_{B_1}+1/9}G_{(1-\alpha/2)}\left(\frac{n}{2},\frac{2}{n}\right)$	β	POWER
0.4	24	1.1233	3.5656	0.3064	0.6936
0.6	28	0.8542	2.4811	0.6858	0.3142
1.0	30	0.5597	1.5660	0.9500	0.0500
1.4	29	0.4069	1.1593	0.7446	0.2554
2.0	24	0.2720	0.8632	0.3445	0.6555
2.2	22	0.2400	0.8038	0.2753	0.7247
2.6	19	0.1921	0.7086	0.1860	0.8140
3.0	17	0.1589	0.6343	0.1323	0.8677
4.0	10	0.1041	0.5143	0.0823	0.9177

Table 4.33: Power of a fixed sample size test of Subclass Variance that corresponds to TEST2.

Summary

As expected TEST2 gave the probability of type I error very close to the nominal level 0.05. Moreover, the test gives strong power than its corresponding fixed sample size starting from $\sigma^2_{B_1} = 2.0 \quad (\sigma^2_{B_o} = 1)$, even for an early truncation of a=30. In general, we have seen from the tables that both tests give more powerful test than the corresponding fixed sample size test procedure for reasonable deviations of the true parameter. In particular TEST2 is more appropriate test in controlling any departure from the null hypothesis. An increase in truncation size gives better performance in detecting any departures with high power compared to all competing tests presented so far.

Figures 4.7 and 4.8 respectively show the plots of powers TEST1 and TEST2 in monitoring the Subclass Variance with the respective fixed sample size test.

Figure 4.1: Plot of the powers of TEST1, truncated sequential test, and fixed sample size test in monitoring the Mean μ .

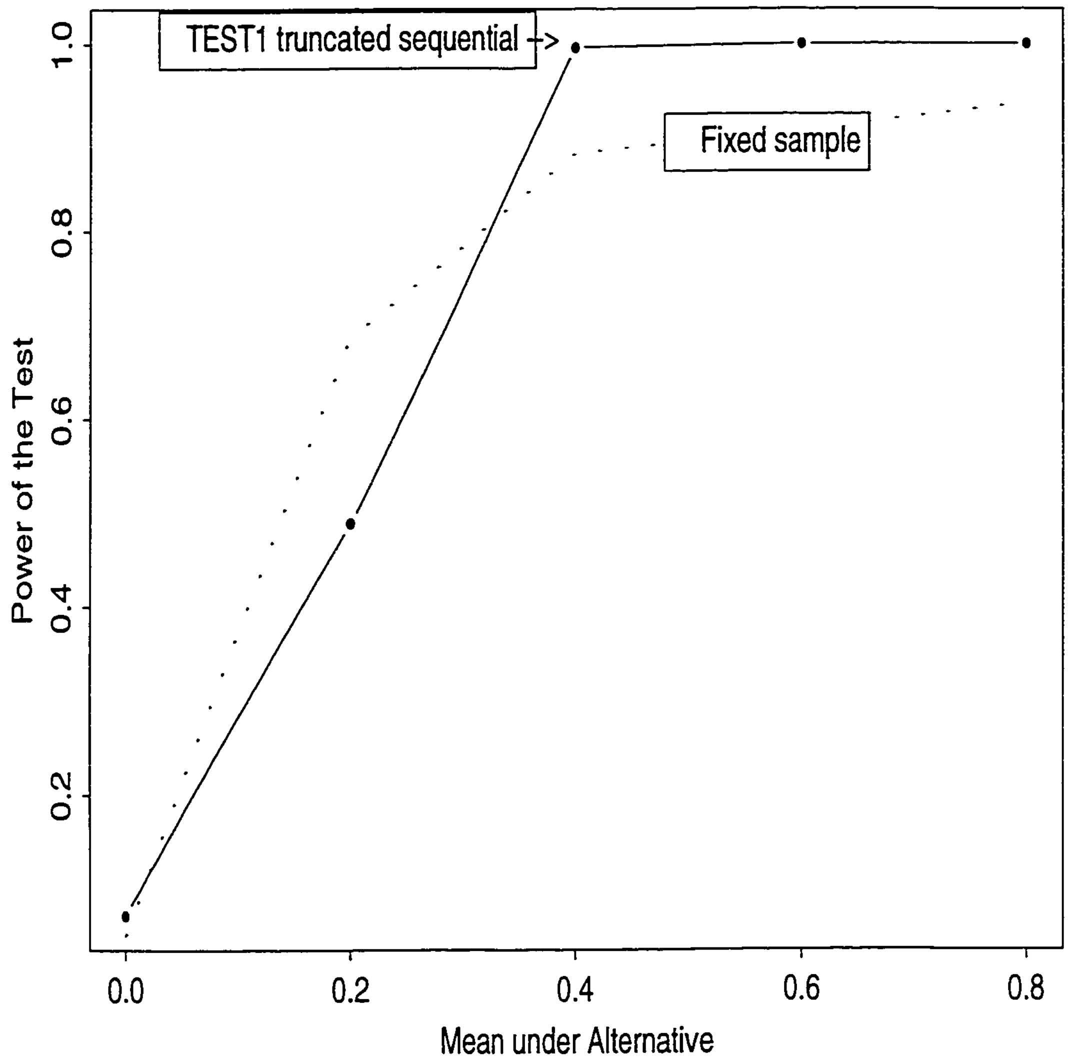


Figure 4.2: Plot of the powers of TEST2, truncated sequential test, and fixed sample size test in monitoring the Mean μ .

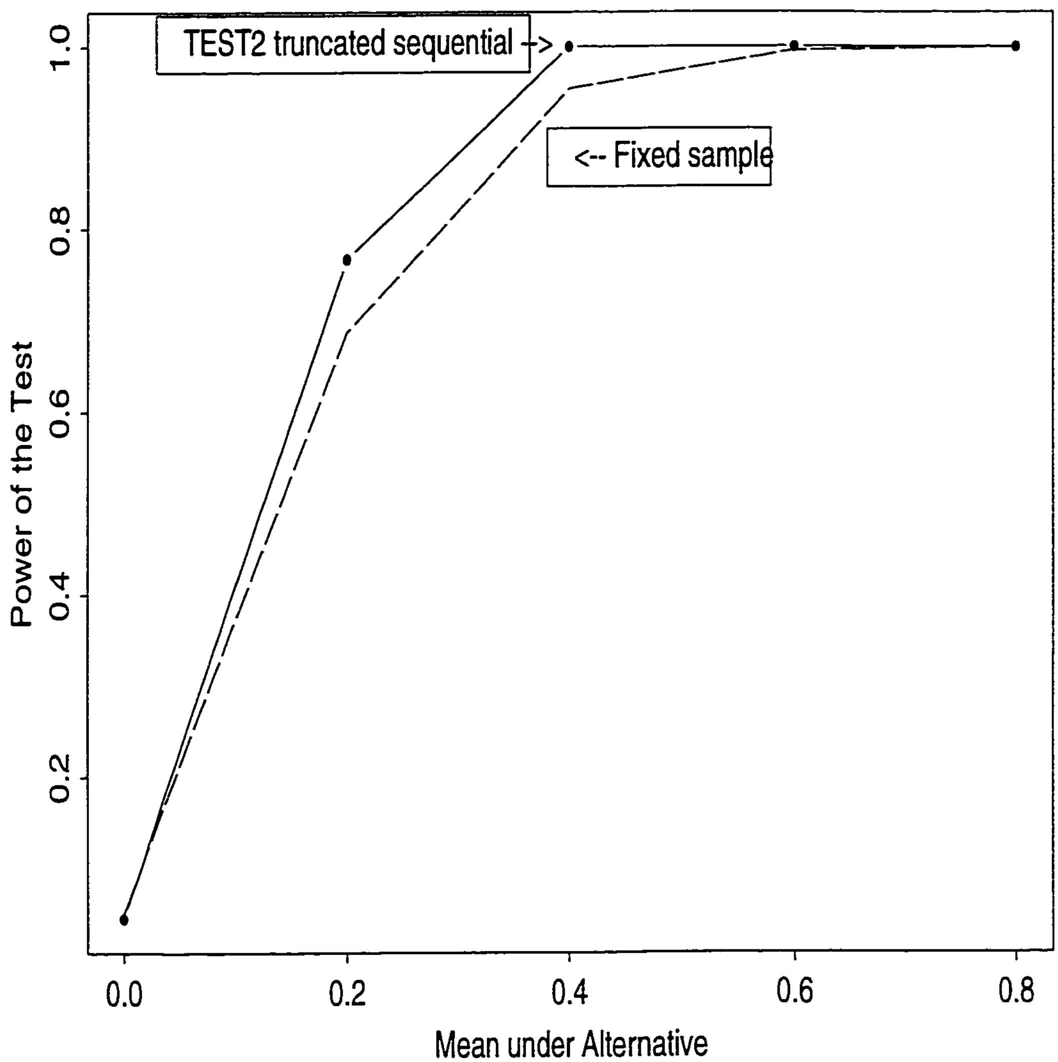


Figure 4.3: Plot of the powers of TEST1, truncated sequential test, and fixed sample size test in monitoring the Error Variance σ^2 .

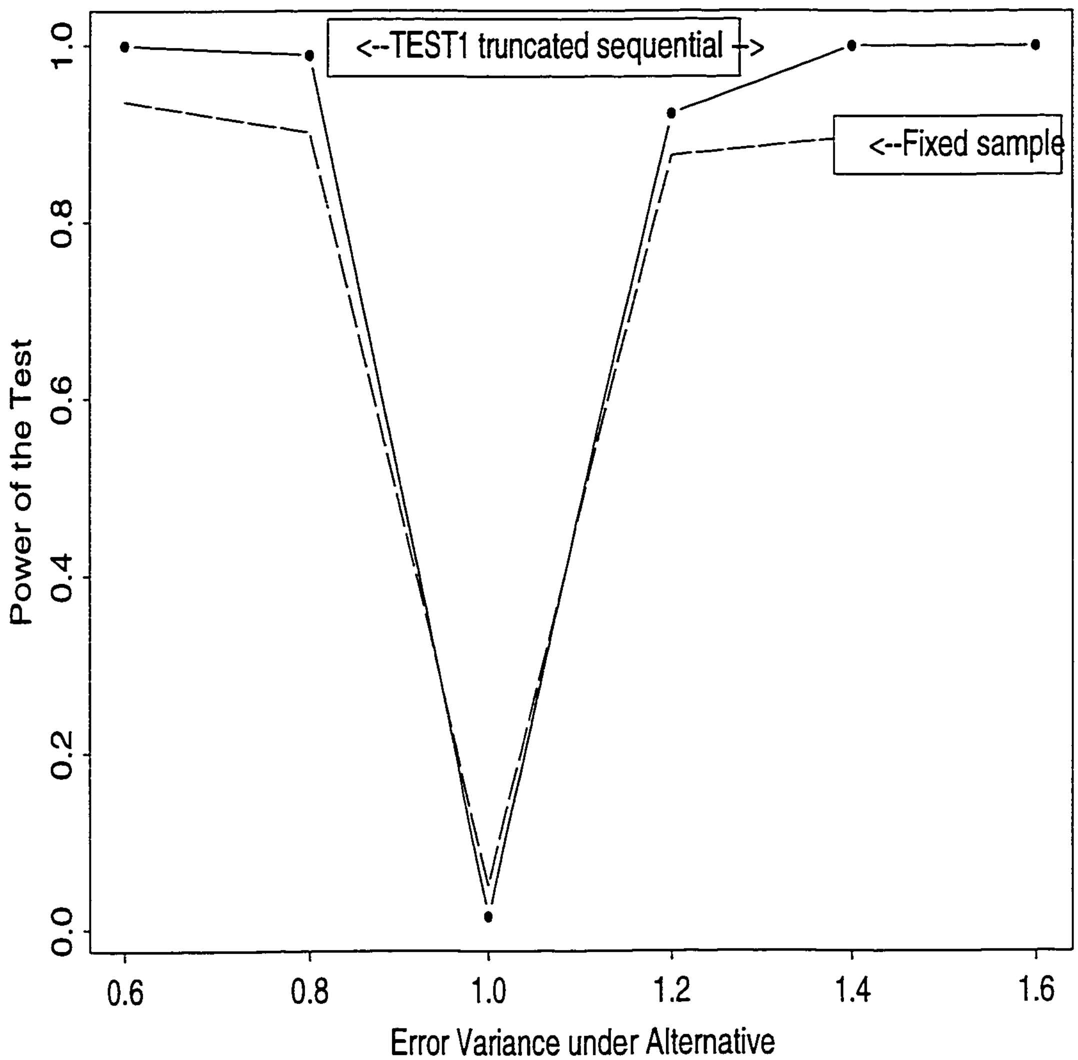


Figure 4.4: Plot of the powers of TEST2, truncated sequential test, and fixed sample size test in monitoring the Error Variance σ^2 .

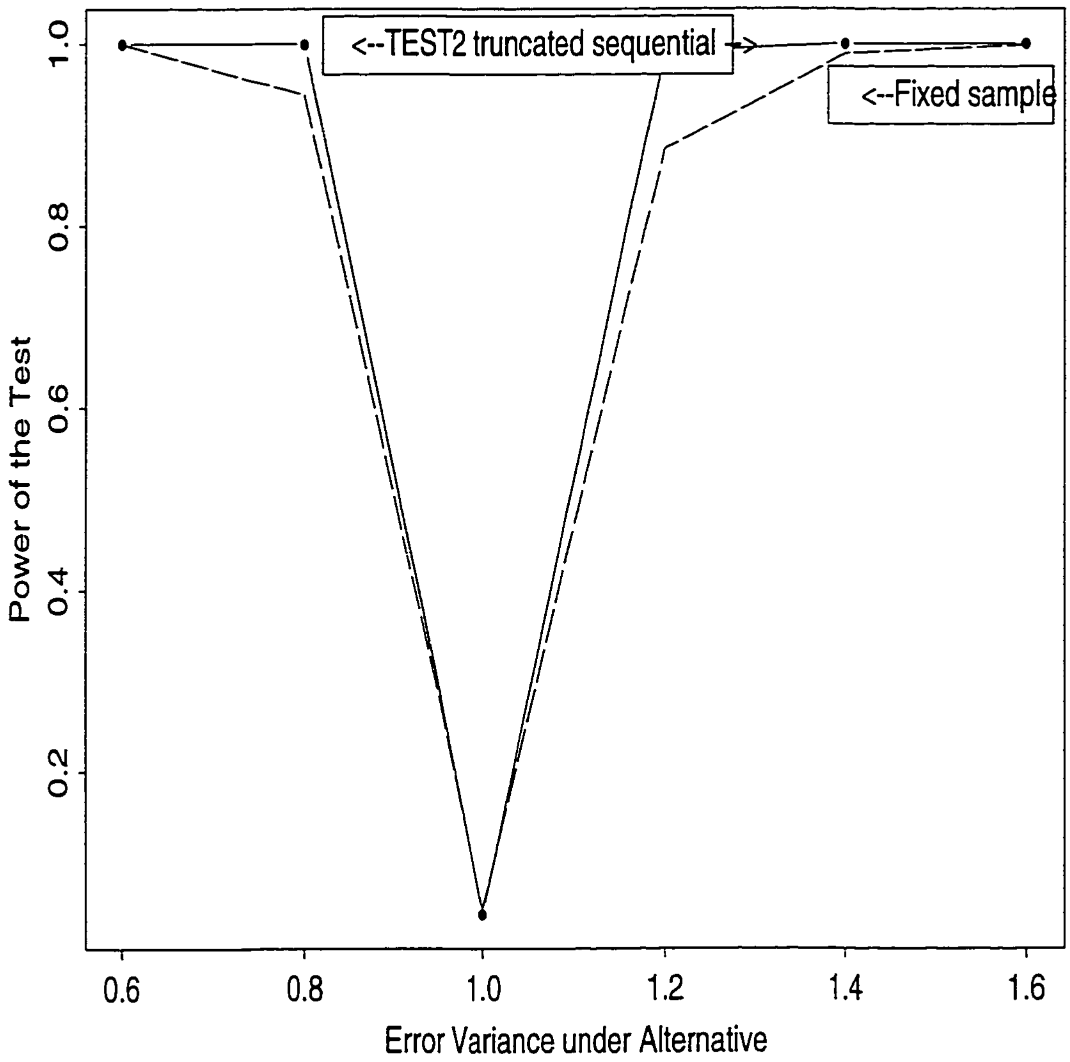


Figure 4.5: Plot of the powers of TEST1, truncated sequential test, and fixed sample size test in monitoring the Class Variance σ_A^2 .

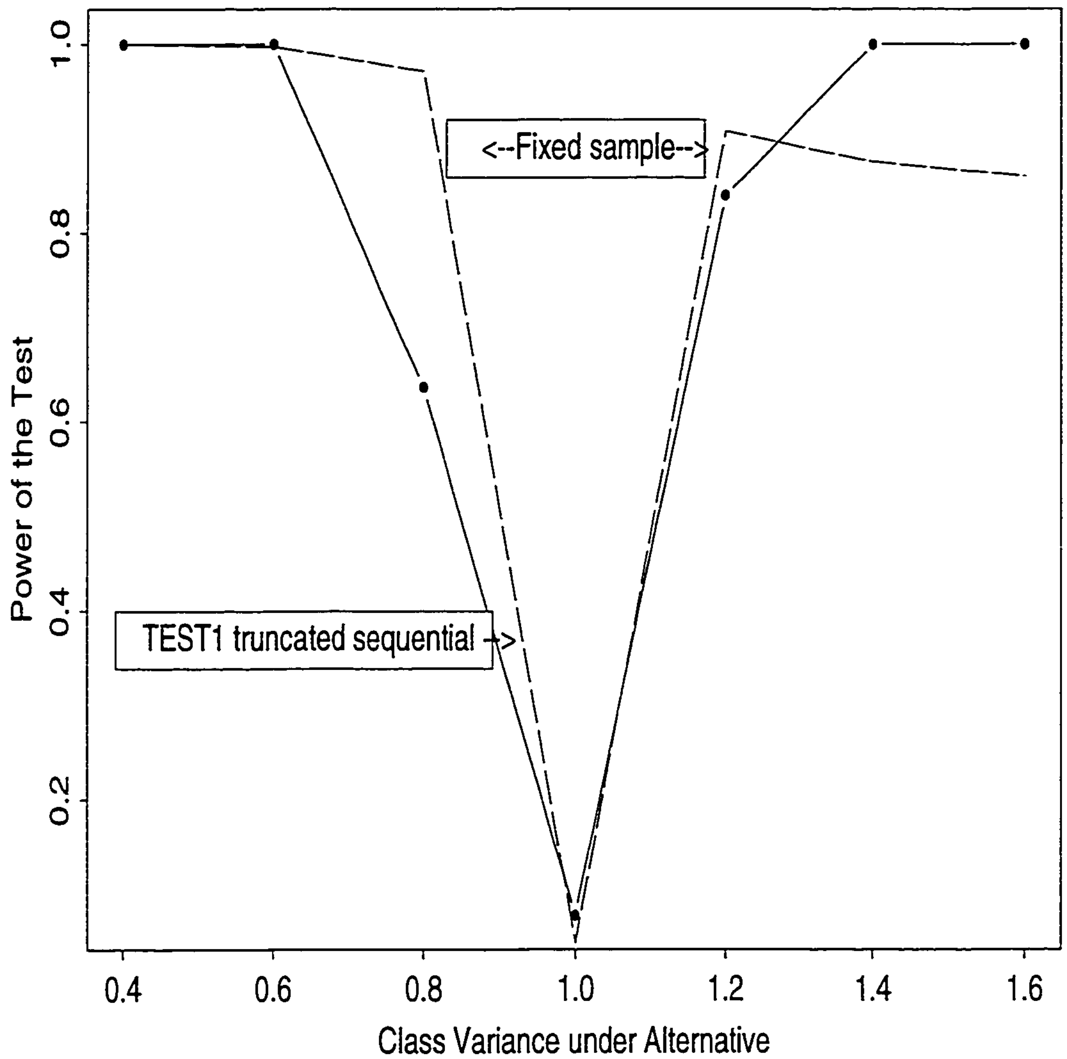


Figure 4.6: Plot of the powers of TEST2, truncated sequential test, and fixed sample size test in monitoring the Class Variance σ_A^2 .

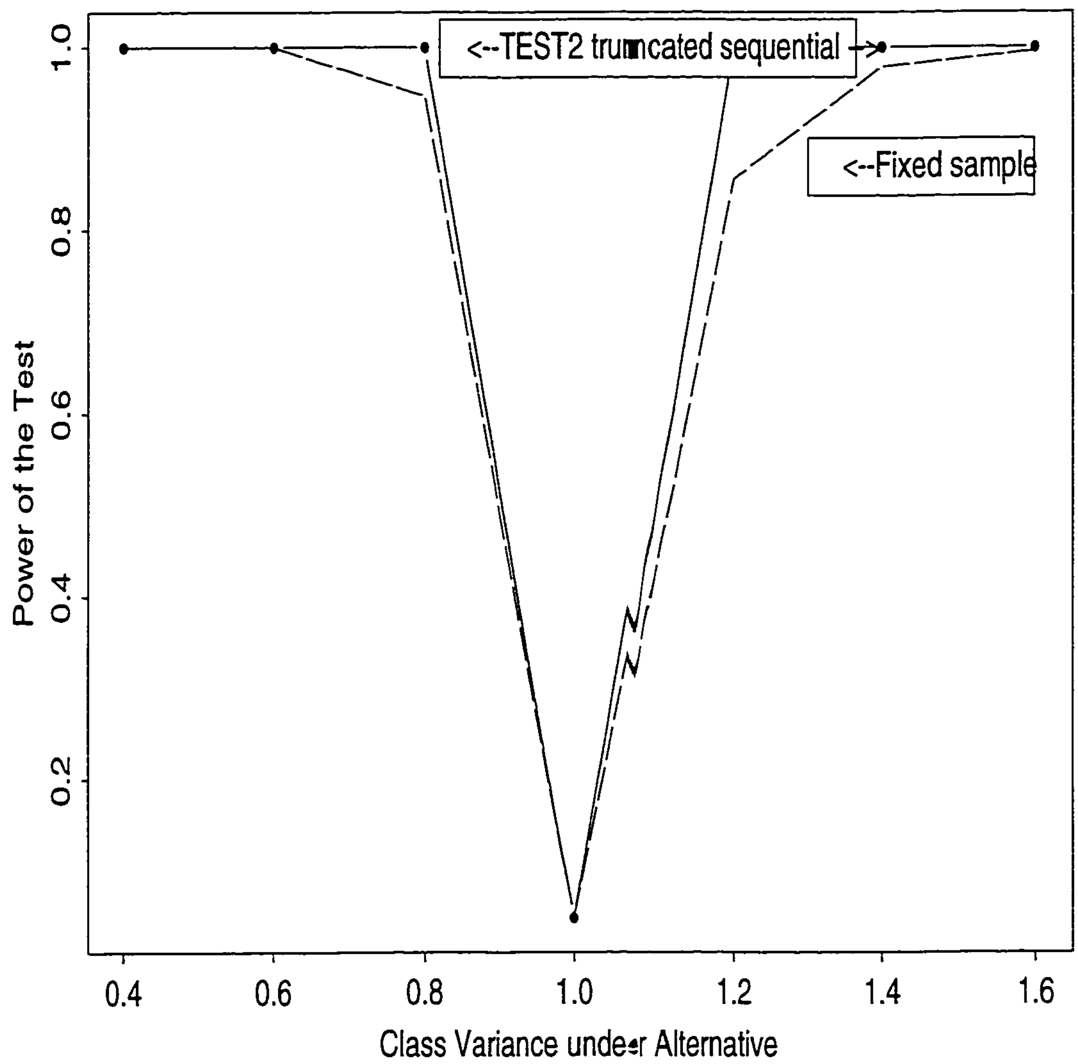


Figure 4.7: Plot of the powers of TEST1 truncated sequential test and fixed sample size test in monitoring the Subclass Variance σ_B^2 .

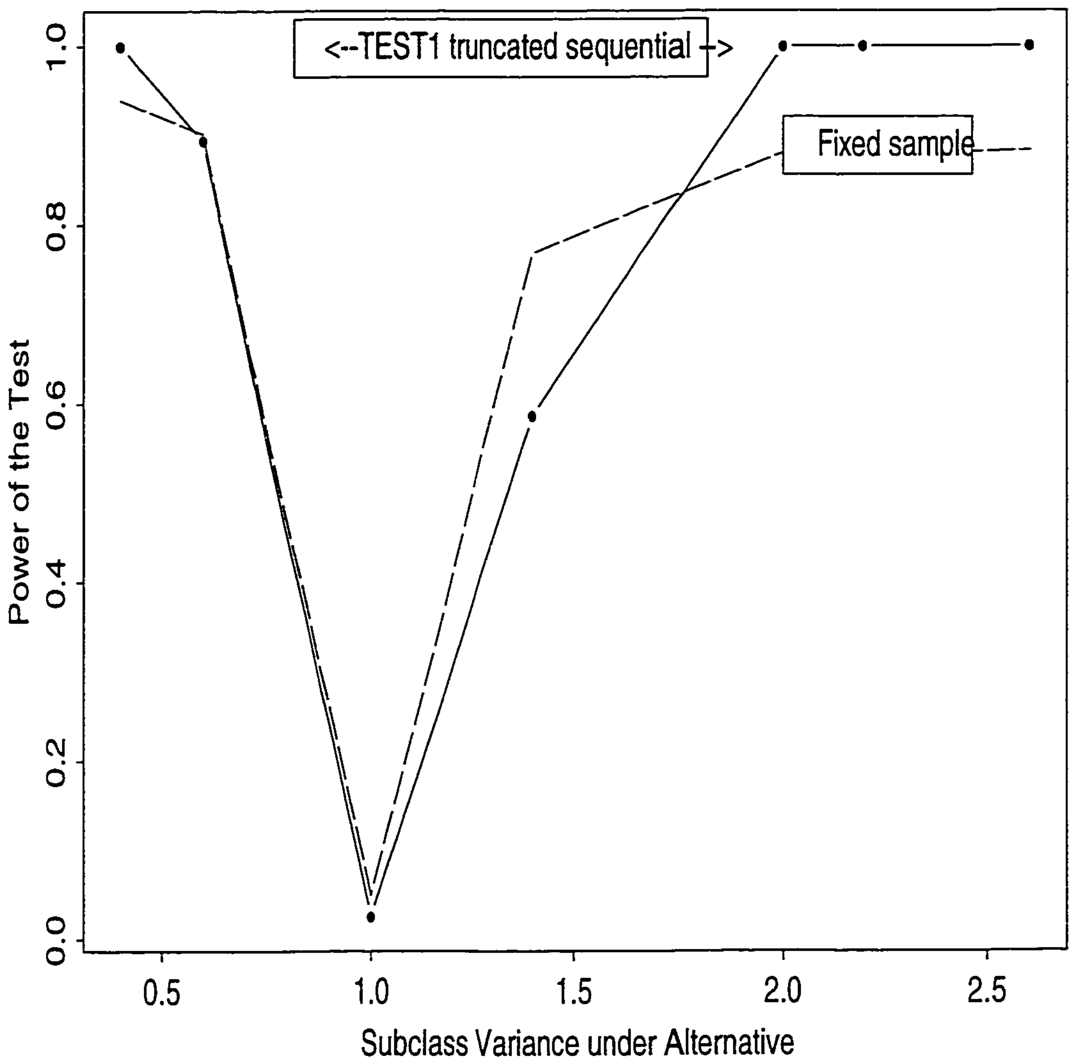
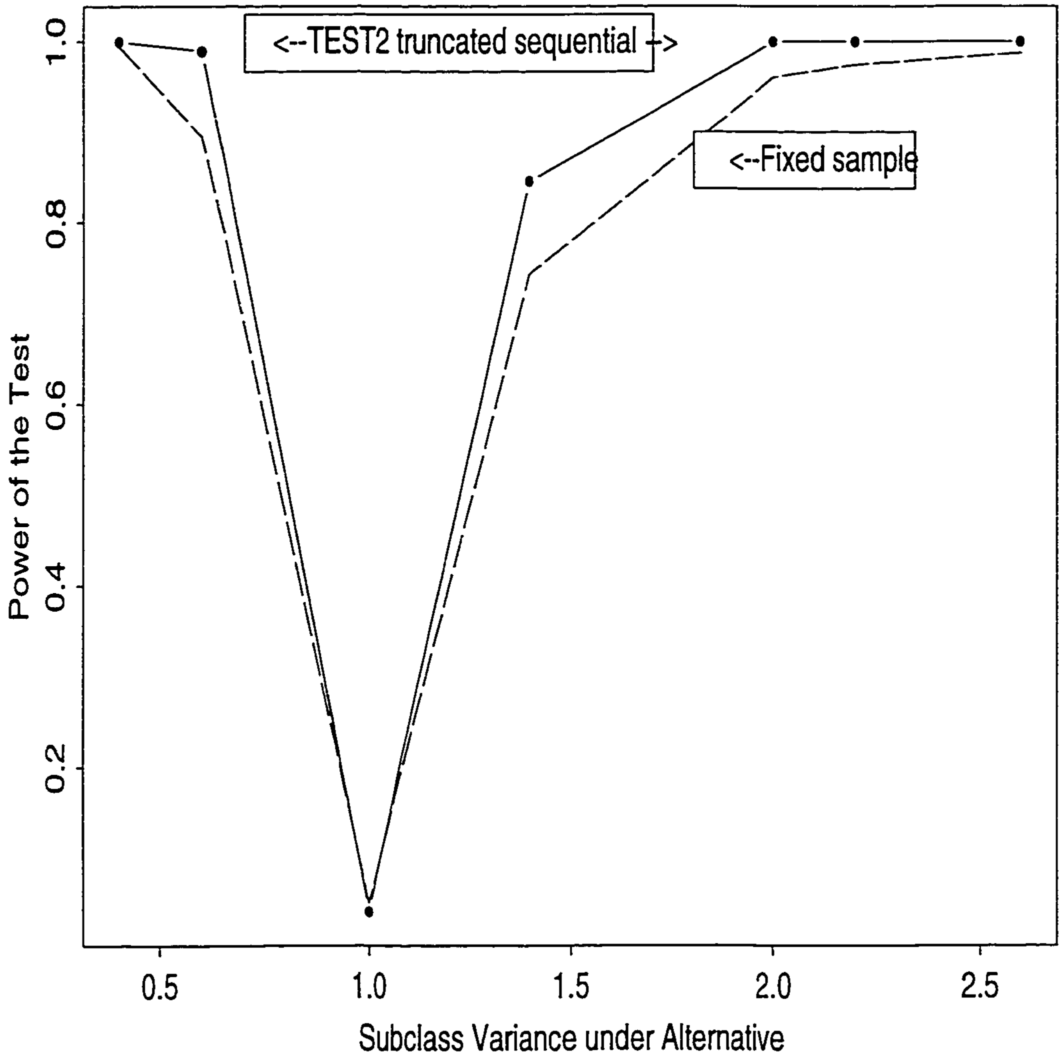


Figure 4.8: Plot of the powers of TEST2 truncated sequential test and fixed sample size test in monitoring the Subclass Variance σ_B^2 .



Chapter 5

Application and Conclusion

5.1 Application

Consider a specific situation arising from the process of manufacturing integrated circuits (or chips) used in computers. Chips are typically processed as part of a wafer, which is a thin disk about 20 cm in diameter. Each wafer contains approximately 200 square-shaped chips. Only at their final stages of processing are the wafers diced to produce individual chips. For most of the production process (which can take months), wafers are handled in lots. For example, when a given tool is used to perform one of hundreds of steps required to turn a raw wafer into a set of chips, the whole lot is processed as a single unit. A typical lot contains about 20 wafers. Our discussion focuses on a specific step of the process that deposits a thin layer of silicon oxide onto the surface of a wafer. To accomplish this step, the lot is placed inside a machine and, after being processed for a specific period, is taken out for various measurenments and further processing by different machines down the line. In this step it is very important to assure the correct thickness of the oxide layer as well as its uniformity, so as to prevent electrical defects or degradation in the performance of the final product.

In the process of monitoring, a sample of R=2 wafers is selected from every lot, and then N=4 measurements of film thickness are made on each of these wafers. The target mean oxide thickness is $\mu = 1,000$ Å. The historically acceptable variance components were $\sigma_b^2 = 3,600$ Å^2 , $\sigma_W^2 = 900$ Å^2 , and $\sigma^2 = 400$ Å^2 .

The model was also given by:

$$X_{irn} = \mu + L_i + W_{r(i)} + \epsilon_{irn} \qquad \begin{cases} i = 1, 2, ..., a \\ j = 1, 2, ..., R \\ k = 1, 2, ..., N \end{cases} \tag{5.1.1}$$

where it is assumed that $L_i \sim N(0, \sigma_b^2)$, $W_{r(i)} \sim N(0, \sigma_W^2)$, and $\epsilon_{irn} \sim N(0, \sigma^2)$ are independent random variables with corresponding variances as the second parameter. L_i is called factor A random effect, $W_{r(i)}$ nested random effect of factor B within the i^{th} level of factor A, while ϵ_{irn} is the random error term of the n^{th} observation within the i^{th} level of factor A and $r(i)^{th}$ level of factor B. This example was discussed by Yashchin (1994 and 1995) and analysed using Cusum Technique and Likelihood Ratio Methods. We now proceed to analyze the data with the Sequential Likelihood Ratio method developed in earlier chapters and compare them.

5.2 Monitoring the Mean

Given are 240 measurenments of the thickness of the oxide layer in 30 lots, 2 wafers from each lot and 4 observations on each wafer.

The hypotheses to be tested are $H_o : \mu = 1000$, $vs.$ $H_1 : \mu \neq 1000$.

The estimator for the mean within i^{th} lot is

$$U_i = \frac{1}{RN} \sum_{r=1}^{R} \sum_{n=1}^{N} X_{irn} = \frac{1}{8} \sum_{r=1}^{2} \sum_{n=1}^{4} X_{irn}$$

which is normal with mean μ and variance $\eta = \sigma_b^2 + \frac{1}{2}\sigma_W^2 + \frac{1}{8}\sigma^2$ as a nuisance parameter.

The maximum likelihood estimates are:

$$\hat{\mu} = \overline{U}_k = \frac{1}{8k} \sum_{i=1}^{k} \sum_{r=1}^{2} \sum_{n=1}^{4} X_{irn} \quad , \quad \hat{\eta}_k = \frac{1}{k} \sum_{i=1}^{k} (U_i - \overline{U}_k)^2$$

and the test statistic

$$-2\log\lambda_k \;=\; k\log\left(\frac{\sum_{i=1}^{k}(U_i-1000)^2}{\sum_{i=1}^{k}(U_i-\overline{U}_k)^2}\right) \quad k=2,3,4,...,30 \tag{5.2.1}$$

with critical values $CV_1(0.05,30) = 9.9968$, $CV_2(0.05) = (2.24)^2$. Using TEST1 and TEST2, corresponding results are presented in Tables 5.1 and 5.2.

According to Table 5.2, a significant difference is observed by TEST2 at 0.05 level. The difference was detected at the 26^{th} lot againist the hypothesized value of mean = 1000 Å. By designing a mixed likelihood ratio control scheme, Yashchin (1995) concluded that, the area $|\mu - 1000| \geq 80$ will be declared rejectable. However none of our estimates satisfy the inequality to support his claim.

k=sample size	$\hat{\mu}$ **-1000**	**-2** $\log \lambda_k$	**Weighted version**
2	5.000	0.0275	0.0018
3	-7.9167	0.1198	0.0120
4	-6.2500	0.1323	0.0176
5	-14.7500	0.8215	0.1369
6	-24.3750	2.0268	0.4053
7	-7.1429	0.1162	0.0271
8	3.2813	0.0251	0.0067
9	8.7500	0.2072	0.0622
10	7.1250	0.1688	0.0563
11	9.6591	0.3637	0.1334
12	13.7500	0.8041	0.3217
13	15.2885	1.1416	0.4947
14	17.7679	1.7005	0.7936
15	15.6667	1.4942	0.7471
16	14.7656	1.5060	0.8032
17	19.7794	2.4890	1.4104
18	16.8056	1.9311	1.1587
19	15.9868	1.9422	1.2301

Table 5.1: Splus program output for testing the Mean using TEST1 and TEST2

k=sample size	$\hat{\mu}$ -1000	-2 log λ_k	Weighted version
20	15.3750	1.9871	1.3247
21	18.7500	2.8898	2.0229
22	20.9659	3.7318	2.7366
23	21.2500	4.1624	3.1912
24	22.8125	5.0187	4.0149
25	23.0500	5.5223	4.6019
26	23.5096	6.1478	5.3281

Table 5.2: Continuation of Table 5.1 for testing the Mean using TEST1 and TEST2

5.2.1 Monitoring Error Variance, σ^2

The hypotheses to be tested are:

$$H_o \ : \ \sigma^2 \ = \ 400 \quad vs. \ H_1 \ : \ \sigma^2 \ \neq \ 400 \ .$$

Considering the estimator of σ^2 at the ith level of a lot

$$\sigma_i^2 = \frac{1}{6}\sum_{r=1}^{2}\sum_{n=1}^{4}(X_{irn} - \overline{X}_{ir.})^2 = \frac{1}{2}\sum_{r=1}^{2} S_{ij}^2, \tag{5.2.2}$$

where $\overline{X}_{ir.} = \frac{1}{4}\sum_{n=1}^{4} X_{irn}$ and $S_{ij}^2 = \frac{1}{3}\sum_{n=1}^{4}(X_{irn} - \overline{X}_{ir.})^2$.

The mle of the parameters σ^2 becomes

$$\hat{\sigma}_k^2 = \frac{1}{6k}\sum_{i=1}^{k}\sum_{r=1}^{2}\sum_{n=1}^{4}(X_{irn} - \overline{X}_{ir.})^2 \tag{5.2.3}$$

and the test statistic is:

$$-2\log\lambda_k = 6k\log\left(\frac{400}{\hat{\sigma}_k^2}\right) + \sum_{i=1}^{k}\sum_{r=1}^{2}\sum_{n=1}^{4}(X_{irn} - \overline{X}_{ir.})^2\left(\frac{1}{400} - \frac{1}{\hat{\sigma}_k^2}\right) \tag{5.2.4}$$

The test statistic in (5.2.4), produces the result presented in Tables 5.3 and 5.4 for a given data in Appendix D.

k=sample size	$\hat{\sigma}_k^2$	-2 log λ_k	Weighted version
2	137.5000	4.9391	0.3293
3	229.1667	2.3388	0.2339
4	208.3333	4.1558	0.5541
5	205.8333	5.3694	0.8949
6	246.5278	3.6111	0.7222
7	265.4762	3.0925	0.7216
8	292.1875	2.1378	0.5701
9	347.6852	0.5065	0.1520
10	372.0833	0.1533	0.0511
11	392.8030	0.0108	0.0040
12	363.8889	0.3124	0.1249
13	408.6538	0.0180	0.0078
14	386.6071	0.0482	0.0225
15	386.1111	0.0555	0.0278
16	370.0521	0.2833	0.1511
17	371.8137	0.2658	0.1506
18	357.4074	0.6595	0.3957
19	353.2895	0.8437	0.5343
20	372.9167	0.2881	0.1921

Table 5.3: Splus output for testing σ^2 using both TEST1 and TEST2.

k=sample size	$\hat{\sigma}_k^2$	-2 log λ_k	**Weighted version**
21	370.8333	0.3522	0.2465
22	364.2045	0.5624	0.4124
23	362.5000	0.6472	0.4962
24	356.9444	0.8994	0.7195
25	377.1667	0.2541	0.2118
26	375.1603	0.3139	0.2720
27	367.9012	0.5513	0.4962
28	377.9762	0.2644	0.2468
29	384.1954	0.1395	0.1349
30	381.2500	0.2042	0.2042

Table 5.4: Continuation of Table 5.3 program output for testing σ^2 using both TEST1 and TEST2.

From Tables 5.3 and 5.4 we can see that, there is no significanct difference in the variance due to error.

5.2.2 Monitoring Lot Variance, σ_b^2

The hypotheses to be tested are:

$$H_o : \sigma_b^2 = 3600 \quad , vs. \ H_1 \ \sigma_b^2 \neq 3600$$

For the random variables, $U_i = \frac{1}{8}\sum_{j=1}^{2}\sum_{k=1}^{4} X_{ijk} \sim N(\mu, \sigma_b^2 + \frac{\xi}{2})$, $Y_i = \hat{\sigma}_{i.}^2 = \sum_{j=1}^{2}(\overline{X}ij. - U_i)^2$, where $\overline{X}ij. = \frac{1}{4}\sum_{k=1}^{4} X_{ijk}$, $\xi = \sigma_w^2 + \frac{1}{4}\sigma^2$,
the mle's of the parameters are $\hat{\mu}_k = \frac{1}{k}\sum_{i=1}^{k} U_i$,
estimators without restriction gives : $\hat{\xi}_k = \frac{1}{k}\sum_{i=1}^{k} Y_i = \overline{Y}$ *and* $\hat{\sigma}_b^2 + 0.5\xi_k = \frac{\sum_{i=1}^{k}(U_i - \overline{U}_k)^2}{k}$
and estimators when $\sigma_b^2 = 3600$ gives $\hat{\xi}_o$ is the solution of the third degree equation:

$$\xi^3 + C_1\xi^2 + C_2\xi + C_3 = 0$$

Here $C_1 = -\frac{\sum_{i=1}^{k}(U_i - \overline{U}_k)^2}{k} + 3600 - 0.5 * \overline{Y}_k$,
$C_2 = 2 * 3600^2 - 7200\overline{Y}_k$, *and* $C_3 = -2 * 3600^2\overline{Y}_k$
and the test statistics is:

$$-2\log\lambda_k = k\log\left(\frac{k(3600 + 0.5 * \hat{\xi}_o)}{\sum_{i=1}^{n}(U_i - \overline{U}_k)^2}\right) + \left(\frac{\sum_{i=1}^{n}(u_i - \overline{U}_k)^2}{(3600 + 0.5\hat{\xi}_o)} - k\right) + k\log\left(\frac{\hat{\xi}_o}{\overline{Y}_k}\right) + k\left(\frac{\overline{Y}_k}{\xi_o} - 1\right) \quad k = 2, 3, ..., 30 \qquad (5.2.5)$$

Using the test statistics in equation 5.2.5, the result presented in Tables 5.5 and 5.6 are obtained.

According to the result in Tables 5.5 and 5.6 and according to the decision rule, there is a suffcient evidence to conclude that there is a difference in the variance between lots from the hypothesized value of 3600 Å^2 , i.e. $\sigma_b^2 \neq$ 3600 Å^2 . However by using a RLR, Regenerative Likelihood Ratio, control scheme Yashchin (1995) arrived at the result that the variance component σ_b^2 is in control.

k=sample size	$\hat{\sigma}_b^2$	-2 log λ_k	Weighted version
2	1540.6250	0.4550	0.0303
3	1172.7431	1.0005	0.10001
4	868.7500	2.0255	0.2701
5	756.1875	2.5062	0.4177
6	1092.3177	2.1659	0.4332
7	2637.3724	0.2389	0.0557
8	3063.4521	0.0800	0.0213
9	2959.5486	0.1337	0.0401
10	2677.3594	0.3311	0.1104
11	2477.7247	0.5657	0.2074
12	2455.2083	0.6552	0.2621
13	2251.3591	1.0198	0.4419
14	2141.8925	1.3129	0.6127
15	2043.3056	1.6471	0.8236
16	1926.8982	2.1182	1.1297
17	2214.2896	1.4461	0.8194
18	2178.3372	1.5773	0.9464
19	1986.1972	2.1508	1.3622
20	1893.2969	2.6262	1.7508

Table 5.5: An Splus program output for testing σ_b^2 using both TEST1 and TEST2.

k=sample size	$\hat{\sigma}_b^2$	-2 log λ_k	**Weighted version**
21	2014.2113	2.3074	1.6152
22	1984.8625	2.4882	1.8247
23	1895.9918	2.9880	2.2908
24	1793.3919	3.4570	2.7656
25	1695.4475	4.0852	3.4044
26	1485.2788	5.0961	4.4167
27	1402.3834	5.7993	5.2194
28	1412.3007	5.8310	5.4423

Table 5.6: Continuation of Table 5.5: program output for testing σ_b^2 using both TEST1 and TEST2.

5.2.3 Monitoring σ_w^2 Wafer Variance

The hypotheses to be tested are: $H_o : \sigma_w^2 = 900 \;\; vs. \;\; H_1 : \sigma_w^2 \neq 900$

We use random variables: $Y_i = \hat{\sigma}_{i.}^2 = \sum_{j=1}^{2}(\overline{X}ij. - U_i)^2$, and $Z_i = \frac{1}{6}\sum_{j=1}^{2}\sum_{k=1}^{4}(X_{ijk} - \overline{X}_{ij.})^2 = \hat{\sigma}_i^2$, where $\overline{X}ij. = \frac{1}{4}\sum_{k=1}^{4} X_{ijk}$, $\xi = \sigma_w^2 + \frac{1}{4}\sigma^2$. It was found that the mle's of the parameters to be:

under no restriction : $\hat{\sigma}_k^2 = \frac{1}{k}\sum_{i=1}^{k} Z_i = \overline{Z} \; and \; \hat{\sigma}_w^2 + 0.25\sigma_k^2 = \overline{Y}_k$

and for the case of, $\sigma_w^2 = 900$: $\hat{\sigma}_o^2$ is the solution of the third degree equation:

$(\sigma^2)^3 + C_1(\sigma^2)^2 + C_2(\sigma^2) + C_3 = 0$ where

$$C_1 = -\frac{4}{7}\left[(\overline{Y}_k - 900) + 1.5(\overline{Z}_k - 8 * 900)\right]$$

$$C_2 = -\frac{4}{7}\left[6(1800\overline{Z}_k - 4 * 900^2)\right] \quad ,$$

and

$$C_3 = -\frac{4}{7}(24 * 900^2\overline{Z}_k) \; .$$

The test statistics is:

$$-2\log\lambda_k = k\left(\frac{\overline{Y}_k}{900+0.25\hat{\sigma}_o^2}-1\right)+k\log\left(\frac{900+0.25\hat{\sigma}_o^2)}{\overline{Y}_k}\right)+$$
$$6k\left[\frac{\overline{Z}_k}{\hat{\sigma}_o^2}-1+\log\left(\frac{\hat{\sigma}_o^2}{\overline{Z}_k}\right)\right] \quad k=2,3,...,30 \qquad (5.2.6)$$

The results of application of the test statisitcs in equation (5.2.6), for the given data in Appendix D are given in Tables 5.7 and 5.8 .

According to the results in Tables 5.7 and 5.8 and according to our decision rule, there is no suffcient evidence to conclude that $\sigma_w^2 \neq 900$. For this parameter Yashchin (1995) arrived at the conclusion that there is unfavorable change in σ_w^2 . In fact, he also gave the rejectable level as $\sigma_w^2 \geq 2,515$. And according to this level and our estimate of σ_w^2 , in Tables 5.7 and 5.8, we arrive at the conclusion of no significant difference.

k=sample size	$\hat{\sigma}_w^2$	-2 log λ_k	Weighted version
1	958.3333	0.0018	0.0001
2	496.8750	0.2664	0.0178
3	672.9167	0.1006	0.0101
4	533.8542	0.4033	0.0538
5	872.9167	0.0021	0.0003
6	710.7639	0.1340	0.0268
7	756.8452	0.0853	0.0199
8	657.0313	0.3004	0.0801
9	567.5926	0.6643	0.1993
10	516.0417	1.0207	0.3402
11	496.4015	1.2430	0.4582
12	454.3403	1.7696	0.7079
13	487.9808	1.5376	0.6663
14	508.4821	1.4840	0.6925
15	503.4722	1.6397	0.8199
16	471.7448	2.1298	1.1359
17	441.0539	2.6846	1.5213
18	541.5509	1.5750	0.9450
19	688.4868	0.5087	0.3222
20	646.1458	0.7905	0.5270

Table 5.7: An Splus program output for testing σ_w^2 using both TEST1 and TEST2.

k=sample size	$\hat{\sigma}_w^2$	-2 log λ_k	**Weighted version**
21	644.9405	0.8398	0.5878
22	694.8864	0.5476	0.4016
23	669.8370	0.7367	0.5648
24	799.0451	0.1340	0.1072
25	813.5833	0.1001	0.0834
26	1079.6474	0.3791	0.3285
27	1110.3395	0.5320	0.4788
28	1165.3274	0.8470	0.7906
29	1155.2443	0.8140	0.7869
30	1289.3750	1.8275	1.8275

Table 5.8: Continuation of Table 5.7 : program output for testing σ_w^2 using both TEST1 and TEST2.

5.2.4 Monitoring all the Variances Simultaneously

The hypotheses to be tested are:
$H_o : (\sigma_b^2, \sigma_w^2, \sigma^2) = (3600,\ 900,\ 400)$ vs.
$H_1 : (\sigma_b^2, \sigma_w^2, \sigma^2)$ as a group is not equal to $(3600,\ 900,\ 400)$.

Considering the joint density of (U_i, Z_i, Y_i) , where U_i , Z_i , *and* Y_i are defined in the previous sections, we found the mle's as:

$$\hat{\mu}_k = \overline{U}_k \ , \ \hat{\sigma}^2 = Z_k \ , \ \hat{\sigma}_w^2 = \overline{Y}_k - 0.25\sigma^2$$

$$\hat{\sigma}_b^2 = \frac{1}{k}\sum_{i=1}^{k}(U_i - \overline{U}_k)^2 - 0.5\sigma_w^2 - 0.0125\sigma^2$$

and the test statistic:

$$\begin{aligned} -2\log\lambda_n \quad = \quad & \left(\frac{\sum_{i=1}^{n}(U_i - \overline{U}_k)^2}{4100} - k\right) + k\log\left(\frac{k(4100)}{\sum_{i=1}^{n}(U_i - \overline{U}_k)^2}\right) + \\ & k\left[\frac{\overline{Y}_k}{1000} - 1 + \log\left(\frac{1000}{\overline{Y}_k}\right)\right] + \\ & 6k\left[\frac{\overline{Z}_k}{400} - 1 + \log\left(\frac{400}{\overline{Z}_k}\right)\right] \end{aligned} \tag{5.2.7}$$

The number of parameters of interest is d=3, so $CV_1(0.05, 30) = 13.9429$. For the given data the test statistics in equation (5.2.7) give, the result presented in Table 5.9.

According to the result in Table 5.9 and our decision rule, there is no sufficient evidence to reject H_o. Figures 5.1 and 5.2 show the plot of $-2\left(\frac{k}{30}\right)\log\lambda_k$, k=2,3,...,30 for these two cases, together with the $\alpha = 0.05$ level critical value.

Summary

In the previous section we have seen the computer output of the new test implemented using SPLUS programming language. We tested all four parameters individually, and then the three variance components simultaneously. Of all the parameters tested the mean shows a significant difference from its target value $\mu_o = 1000$ Å at 0.05 level of significance. Also, the variance component due to lot effect, σ_b^2 , show us a change from its target value $\sigma_b^2 = 3600$ Å at 0.05 level of significance.

k=sample size	-2 log λ_k		k=sample size	-2 log λ_k
2	5.7872		3	3.5397
4	6.8155		5	7.9377
6	6.0784		7	3.4956
8	2.6159		9	1.4165
10	1.6595		11	2.0087
12	3.0065		13	2.8118
14	3.1395		15	3.6824
16	4.9829		17	4.8346
18	4.2043		19	3.7829
20	4.0396		21	3.8335
22	3.8970		23	4.7380
24	4.6938		25	4.6183
26	5.5487		27	6.5671
28	6.5338		29	2.8135
30	4.2559			

Table 5.9: An Splus program output for testing all the Variances simultineously using TEST1.

5.3 General Discussion and Conclusions

When the problem of sequential analysis arises, as shown in the introduction part of this thesis, there was a conjecture by Milton Friedmand and W. Allen Wallis (refer Wald's (1947) pp. v), after recognizing the great potentialities of sequential analysis for theoretical statistics. Their statement was, a sequential test procedure might be constructed which would control the possible errors committed by wrong decisions exactly to the same extent as the best current procedure based on a predetermined number of observations. At the same time it would require a substantially smaller number of observations than fixed sample size test.

Different authors attempted to verify the conjecture in testing a hypotheses with varying degree of success, in the presence of a nuisance parameter. In fact, they gave the basis point for further research. In this study a truncated sequential analysis for a Nested Random Effects Model was developed and tested, that will accommodate nuisance parameters. The method shown here is shown to be more powerful.

The tables contain figures about the performance of the fixed sample size tests with sample size equal to the average sample number, ASN, taken to the next possible integer value, for those ASN's whose decimal digit greater that or equal to two tenths. Otherwise the integer part only. The new tests outperform both the sequential t-test and fixed sample size test at parameter values where the power is large. The study also showed that, likelihood ratios can be used sequentially to the Neyman-Pearson framework and TEST1, TEST2 can be applied to solve many problems. These test procedures are more powerful at the expense of greater stopping time.

The study has also demonstrated the considerable increase in power of TEST2 by the use of the weight function. In fact, the tradeoff is increased stopping time, worthwhile paying only in case the two hypotheses are close.

In general, the new monitoring procedures are very simple to perform, and easily accommodate nuisance parameters in sharp contrast to existing procedures, so the question of robustness in terms of chosen nuisance parameter values does not even arise. No tolerance limits have to be selected, while it was the case in Yashchin's work, and any size of departures from the target can be detected, if the process is monitored long enough. The likelihood ratio used here has an optimal power properties. In fact, the price for this is greater average sample number compared to the Wald-type use of likelihood ratio.

The increase in power is dramatic at the expense of much greater stopping time. Hence users have to choose the procedure based on their criteria of optimality.

Figure 5.1: Plot of the Weighted Test Statistics (TEST2) versus number of lots used to test in detecting a change of Mean μ Lot-wise.

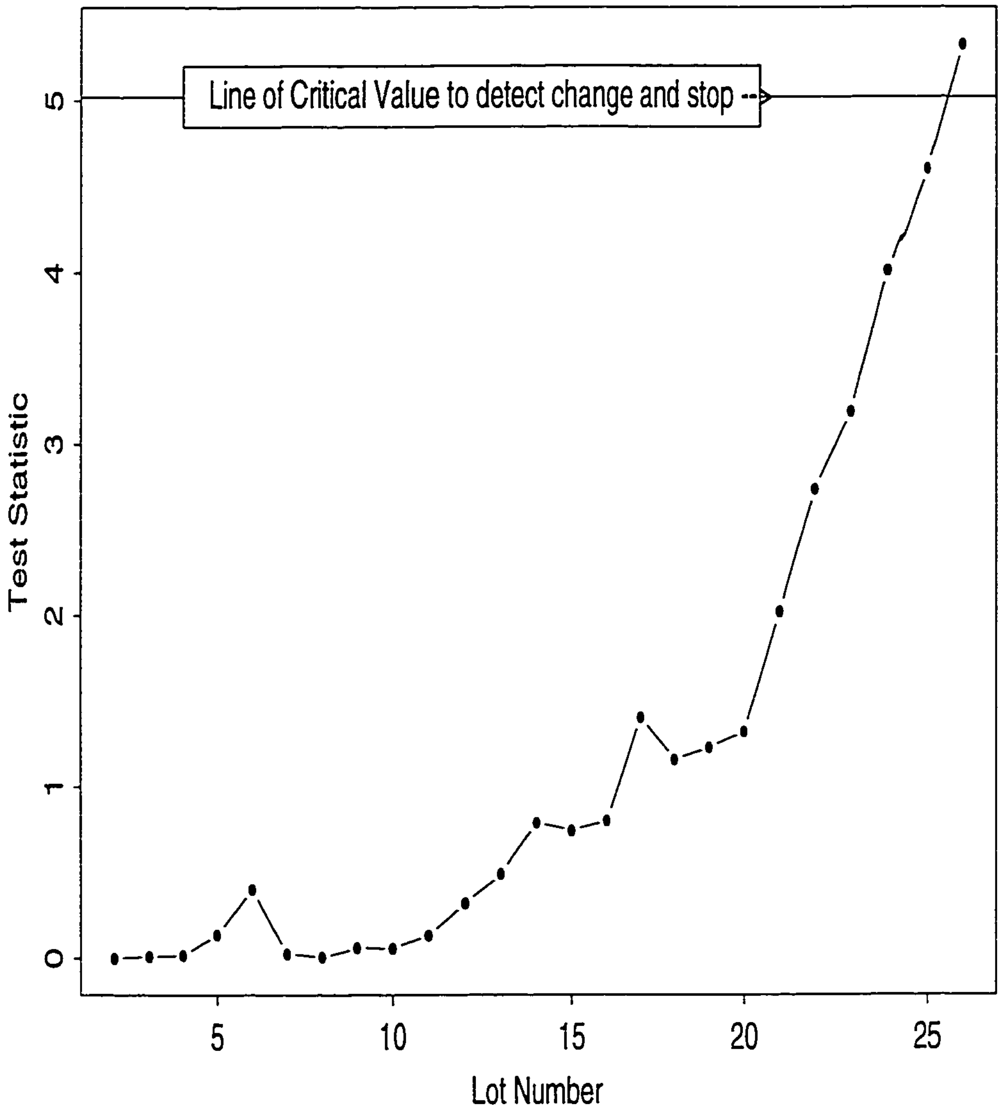
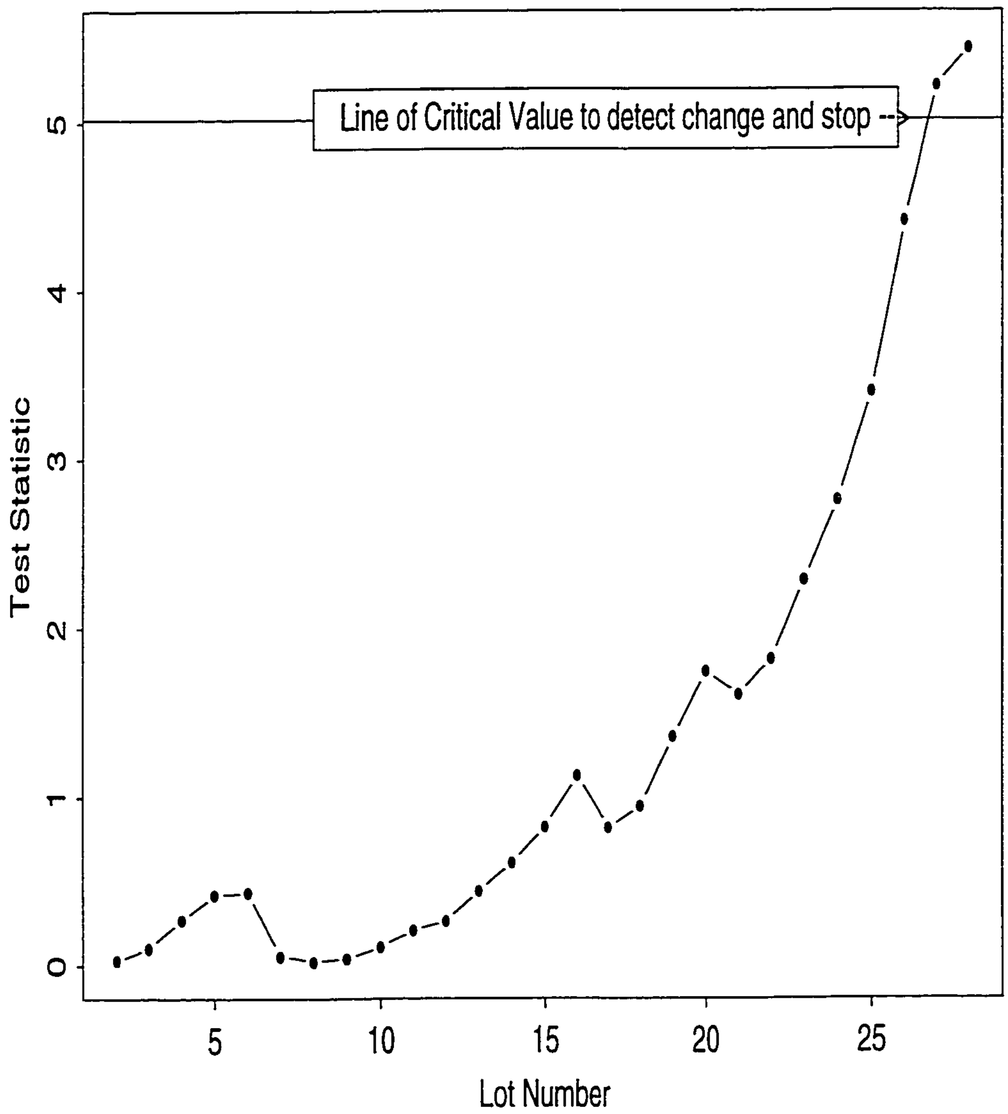


Figure 5.2: Plot of the Weighted Test Statistics (TEST2) versus the Number of Lots used to test in detecting a change of Lot Variance σ_b^2 .



Bibliography

Abraham, Wald (1944). On cumulative sums of random variables. *Annals of Math. Stat.* **15**, 283–296.

Abraham, Wald (1945). Sequential tests of statistical hypotheses. *Annals of Math. Stat.* **16**, 117–185.

Barnard, G.A. (1952). The frequency justification of certain sequential tests. *Biometrika* **39**, 144–150.

Bartky, Walter (1943). Multiple sampling with constant probability. *Annals of Math. Stat.* **14**, 363–377.

Bartlett, M. S. (1946). The large sample theory of sequential tests. *Proc. Camb. Phil. Soc.* **42**, 239–244.

Bruyn, C.S. Van Dobben De (1968). *Cumulative sum tests theory and practice.* Hafner publishing company.

Cox, D.R. (1952). Sequential tests for composite hypotheses. *Proc. Camb. Phil. Soc.* **48**, 290–299.

Cox, D.R. (1963). Large sample sequential tests for composite hypotheses. *Sankhyā* **25**, 5–12.

Daniel, Wayne W. (1995). *Biostatistics: a foundation for analysis in the health sciences.* 6th ed.. Holden-Day, Inc.

Dodge, H.F. and Romig H.G. (1929). A method of sampling inspection. *The Bell System Tech. Jour.* **8**, 613–631.

Ferguson, Thomas S. (1967). *Mathematical statistics.* Academic press, Inc.

Ghosh, B.K. (1970). *Sequential tests of statistical hypotheses.* Addison-Wesley publishing company, Inc.

Ghosh, B.K. and P.K. Sen (1974). *Handbook of sequential analysis.* Richard D. Irwin, Inc.

Gombay, E. (1996). The weighted sequential likelihood ratio. *Canadian Journal of Statistics* **24**, 229–239.

Gombay, E. (1997). The likelihood ratio under noncontiguous alternatives. *Canadian Journal of Statistics* **25**, 417–423.

Gombay, E. (2000a). Sequential change-detection with likelihood ratios. *To appear in Statistics and Probability Letters.*

Gombay, E. (2000b). Sequential testing of composite hypotheses. *To appear in Proceeding of the Fourth Hungarian Colloquium on Limit Theorems of Probability and Statistics.*

Govindarajulu, Z. (1975). *Sequential statistical procedures.* Academic press, Inc.

Joanes, D.N. (1972). Sequential tests of composite hypotheses. *Biometrika* **59**, 633-637.

Johnson, N.L. (1961). Sequential analysis: a survey. *Royal Statistical Society Journal Series A* **24**, 372–411.

Lindsey, J.K. (1996). *Parametric statistical inference.* Oxford University press, Inc.

Lorden, G. (1971). Procedures for reacting to a change in distribution. *Annals of Mathematical Statistics* **42**, 1897–1908.

Montgomery, Douglas C. (1997). *Design and analysis of experiments.* 4th ed.. John Wiley and Sons, Inc.

Moustakides, G.V. (1986). Optimal stopping times for detecting changes in distributions. *Annals of Statistics* **14**, 1379–1387.

Neter, John and William Wasserman (1974). *Applied linear statistical models.* Richard D. Irwin, Inc.

Neter, John, et. al (1985). *Applied Linear Statistical Models.* 2nd ed.. Homewood, Ill. : R.D. Irwin.

Rushton, S. (1952). On a two-sided sequential t-test. *Biometrika* **39**, 326–333.

Siskind, V. (1964). On certain suggested formulae applied to the sequential t-test. *Biometrika* **51**, 97–106.

Tester, Jefferson.W. (1997). *Thermodynamics and its applications.* 3rd ed.. Prentice-Hall PTR.

Wald, Abraham (1947). *Sequential analysis.* John Wiley and Sons, Inc.

Whetherill, G.B. and K.D. Glazebrook (1986). *Sequential Methods in Statistics.* Chapman and Hall.

Woodall, W.H. (1983). The distribution of the run length of one-sided cusum procedures for continuous random variables. *Technometrics* **25**, 295–300.

Yashchin, E. (1992b). Analysis of cusum and other markov-type control scemes by using empirical distributions. *Technometrics* **34**, 54–63.

Yashchin, E. (1993). Performance of cusum control schemes for serially correlated observations. *Technometrics* **35**, 37–52.

Yashchin, E. (1994). Monitoring variance components. *Technometrics* **36**, 379–393.

Yashchin, E. (1995). Likelihood ratio method for monitoring parameters of a nested random effect model. *JASA* **90**, 729–738.

Appendix A

Analytical solution for cubic equation

Suppose the cubic equation

$$x^3 + C_1x^2 + C_2x + C_3 = 0 \tag{A.0.1}$$

is given and we want to solve the equation for x if there are real solutions. By putting

$$Q = \frac{3C_2 - C_1^2}{9}, \quad R = \frac{9C_1C_2 - 27C_3 - 2C_1^3}{54}, \quad \text{and} \quad D^* = Q^3 + R^2 \tag{A.0.2}$$

Tester (1997), outlined that D^* will be the value that determines the domain of the roots of equation (A.0.1)

1) If $D^* < 0$, all roots are real and unequal.

2) If $D^* = 0$, all roots are real and at least two are equal.

3) If $D^* > 0$, only one root is real and two are imaginary.

a) If $D^* > 0$, the roots are given by:

$$\begin{cases} x_1 = S + T - \frac{1}{3}C_1 \\ x_{2,3} = -\frac{1}{2}(S+T) - \frac{1}{3}C_1 \pm \frac{1}{2}i\sqrt{3}(S-T) \end{cases} \tag{A.0.3}$$

$$\text{where} \quad S = (R + \sqrt{D^*})^{\frac{1}{3}} \quad \text{and}$$
$$T = (R - \sqrt{D^*})^{\frac{1}{3}} \quad \text{and} \quad i = \sqrt{-1}$$

b) If $D^* = 0$, then S=T and the imaginary components disappear. Then the roots will be:

$$\begin{cases} x_1 = S + T - \frac{1}{3}C_1 \\ x_{2,3} = -\frac{1}{2}(S+T) - \frac{1}{3}C_1 \end{cases} \tag{A.0.4}$$

c) If $D^* < 0$, the roots are given by:

$$\begin{cases} x_1 = 2\sqrt{-Q}\cos\left(\frac{\theta}{3} + 120^o\right) - \frac{1}{3}C_1 \\ x_2 = 2\sqrt{-Q}\cos\left(\frac{\theta}{3} + 240^o\right) - \frac{1}{3}C_1 \\ x_3 = 2\sqrt{-Q}\cos\left(\frac{\theta}{3}\right) - \frac{1}{3}C_1 \end{cases} \tag{A.0.5}$$

$$\text{where} \quad \cos\theta = \frac{R}{\sqrt{-Q^3}}$$

Example: Solve:

$$i)\ (x+3)^3 = 0 \qquad ii)\ (x-1)(x^2+2x+4) = 0$$

$$iii)\ (x-2)(x^2+5x+6) = 0 \quad iv)\ (x-1)(x+2)^2 = 0$$

Solution:

i) For this particular problem, as it is expressed as a power of a linear factor that can be solved easily, we see that the unique real root is -3.
Also $(x+3)^3 = x^3 + 9x^2 + 27x + 27$, and then $C_1 = 9$, $C_2=27$, and $C_3=27$.

Then using the equation (A.0.2) we have:

$$\begin{aligned} Q &= \frac{3C_2 - C_1^2}{9} \\ &= \frac{3*27 - 9^2}{9} = 0 \\ R &= \frac{9C_1C_2 - 27C_3 - 2C_1^3}{54} \\ &= \frac{9*9*27 - 27*27 - 2*9^3}{54} = 0 \end{aligned}$$

and $D^* = S = T = 0$. The root in both formulas then will be $x_1 = -\frac{1}{3}C_1 = -3$ which was really the unique result we have got earlier.

ii) As the problem is simplified in a form of a product of a linear term and quadratic term, one can see that the real value that satisfies the equation is only the number 1.

Also, using the equation (A.0.2) we have:

$(x-1)(x^2+2x+4) = x^3+x^2+2x-4$, then C_1=1, C_2=2, C_3=–4

$$Q = \frac{3*2-1^2}{9} = \frac{5}{9}, \quad R = \frac{9C_1C_2-27C_3-2C_1^3}{54} = \frac{9*1*2+4*27-2*1^3}{54} = \frac{62}{27}$$

and hence,

$$D^* = Q^3+R^2 = \left(\frac{5}{9}\right)^3 + \left(\frac{62}{27}\right)^2 = \left(\frac{63}{27}\right)^2 > 0$$

Then the equation has only one root.

$$\begin{aligned} x &= S+T-\frac{1}{3}C_1 = (R+\sqrt{D^*})^{\frac{1}{3}} + (R-\sqrt{D^*})^{\frac{1}{3}} - \frac{1}{3} \\ &= \left(\frac{62}{27}+\frac{63}{27}\right)^{\frac{1}{3}} + \left(\frac{62}{27}-\frac{63}{27}\right)^{\frac{1}{3}} - \frac{1}{3} \\ &= 1 \end{aligned}$$

which is again the answer found earlier.

iii) The given equation can be written as $(x-2)(x+2)(x+3)=0$ so that the values that satisfy the equation are 2, -2, and -3. Also, using the equation (A.0.2) we have:

$(x-2)(x^2+5x+6) = x^3+3x^2-4x-12$, then C_1=3, C_2=-4, and C_3=-12.

$$Q = \frac{-12-9}{9} = -\frac{7}{3} \quad \text{and} \quad R = 3$$

$$D^* = Q^3+R^2 = \left(-\frac{7}{3}\right)^3 + 3^2 = -\frac{100}{27} < 0$$

The we will have three distinct real roots. For $\cos\theta = \frac{R}{-\sqrt{Q^3}} = \frac{9\sqrt{3}}{7\sqrt{7}}$,

$$\begin{cases} x_1 &= 2\sqrt{-Q}\cos\left(\frac{\theta}{3}+120^o\right)-\frac{1}{3}C_1=2 \\ x_2 &= 2\sqrt{-Q}\cos\left(\frac{\theta}{3}+240^o\right)-\frac{1}{3}C_1=-3 \\ x_3 &= 2\sqrt{-Q}\cos\left(\frac{\theta}{3}\right)-\frac{1}{3}C_1=-2 \end{cases}$$

the same solutions, found before.

iv) Since the equestion is given in factorized form, the solutions to these equation are 1 and -2.

Also, using the equation (A.0.2) we have:

$(x-1)(x+2)^2 = x^3+3x^2-4$, then C_1=3, C_2=0, and C_3=-4.

$$Q = \frac{3C_2 - C_1^2}{9} = -1 \ , \quad R = 1 \text{ and } D^* = 0$$

The the roots will be:

$$x_1 = S + T - \frac{1}{3}C_1 = 1$$
$$x_2 = -\frac{1}{2}(S+T) - \frac{1}{3}C_1 = -2$$

that gives which are the same roots identified from the factorized form of the equation.

Appendix B

Approximations of the Tests under the Alternative $\mathbf{H}_1$

B.1 Monitoring σ_B^2

The log likelihood function of the monitoring variables $(\hat{\sigma}_i^2\ ,\ \hat{\sigma}_{i.}^2)$ which is given by:

$$\ell(\sigma_B^2, \sigma^2) = -\left(\frac{b-1}{2(\sigma_B^2 + \frac{\sigma^2}{N})}\right)\sum_{i=1}^{n} Y_i - n\left(\frac{b-1}{2}\right)\log\left(2(\sigma_B^2 + \frac{\sigma^2}{N})\right) - \frac{b(N-1)}{2\sigma^2}\sum_{i=1}^{n} Z_i - \frac{nb(N-1)}{2}\log 2\sigma^2 + Q(Y_i, Z_i) \quad \text{(B.1.1)}$$

where $F(Y_i)$, $H(Z_i)$ *and* $Q(Y_i, Z_i)$ are functions of the data only which do not depend on any of the parameters, doesn't belong to the exponential family of distributions. Hence we can not apply the result of sections (2.4.2). Instead, we have the following Theorem.

Let σ_{B1}^2 denote the value of variance component σ_B^2 after change.

Theorem B.1.1 *Under the alternative hypothesis* H_a *, if* $k > \tau$ *where* τ *is the stopping time in detecting a change, then*

$$Sup\left|\frac{-2\log\Lambda_k - Q_k}{k^{1/2}V_{\tau+1}^{1/2}} - k^{-1/2}W(k-\tau) - \left(\frac{V_{\tau+1}}{V_1}\right)^{(1/2)} N(0,1)\right| = o_p(1) \ ,$$

where Brownian motion W and standard normal random variable N are independent, and the constant terms are:

$$Q_k = k\{-V_1(1-\xi^a_{ok}/\xi^a_k) - V_2(1-\theta^a_{ok}/\theta^a_k)$$
$$-2[A(\xi^a_k,\theta^a_k) - A(\xi^a_{ok},\theta^a_{ok})]\} \quad ,$$

where $V_1, V_{\tau+1}$ are defined by:

$$V_i = VarY_i = 2V_1(\xi^a_k - \theta^a_{ok})^2\sigma^4/N^2 + 2V_2(\theta^a_k\theta^2_{ok}(\sigma^2_B + \sigma^2/N)$$

$$\xi^a_k = (\sigma^2/N)^{-1}, \quad A(\xi,\theta) = -V_1/\log\xi - V_2/(2\log\theta) \quad ,$$
$$\theta^a_k = [(\tau/k)\sigma^2_{B_o}, (k-\tau)/k\sigma^2_{B_q} + \sigma^2/N]$$

$\theta^a_{ok}, \xi^a_{ok}$ are obtained from the solution of the non-random third degree polynomial equation, and depend on k, τ in a similar manner as θ^a_k .

So in this case, as well as in the next one, the drift of $\{-2\log\Lambda_k\}$ towards infinity is of order k. The critical value of the TEST is of order $(\log\log M)^{1/2}$ where M is the truncation point, so the consistency of the test follows.

B.2 Monitoring σ^2_A

The likelihood structure has a similar structure as in the previous section, so the conclusions are the same, hence its discussion is omitted.

Appendix C

SPLUS code for the Simulation

C.1 For Monitoring Mean, μ

```
pi <- 3.141592653589793238462643383279502884197169399375
a <- 30; b <- 2; N <- 4 ; sibsq <- (3600/4900)
siwsq <- (900/4900); sisq <- (400/4900)
mua <- c(0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9); rept <- 2000
sink("norm1out"); n <- rep(0,rept) ; cat("within ",a,"classes ")
gener <- function(a,b,N,mu,sisq,sibsq,siwsq) {
    eijk <- rnorm((a*b*N),0,sisq)
    bji <- rnorm((a*b),0,siwsq)
    ali <- rnorm(a,mu,sibsq)
    alpha1 <- ali[ali=c(1,1,2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10)]
    alpha2 <- ali[ali=c(11,11,12,12,13,13,14,14,15,15,16,16,17,17,18,18)]
    alpha3 <- ali[ali=c(19,19,20,20,21,21,22,22,23,23,24,24,25,25)]
    alpha4 <- ali[ali=c(26,26,27,27,28,28,29,29,30,30)]
    alphai <- c(alpha1,alpha2,alpha3,alpha4)
    data <- rep(0,a*b*N)
```

```
    for(k in 1:(a*b)) {
    data[(N*(k-1)+1):(N*k)] < - eijk[(N*(k-1)+1):(N*k)] + bji[k]+alphai[k]
        }
    return(data)
    }
test1 < -  function(k)
    {
        ndata < - ui[1:k]
        xbar < - mean(ndata)
        datdif < - ndata-xbar
        ssqn < - vecnorm(ndata) ^ 2
        ssqden < - vecnorm(datdif) ^ 2
        sqnul < - ssqn/ssqden
        test1 < - (k*log(sqnul))
        return(test1)
    }
cv1 < - function(n, alpha) {
    an < -  2*log(log(n))
    bn < - an+.5*log(log(log(n)))-log(sqrt(pi))
    cn < - -log(-log(1-alpha))
    cv1 < - (cn+bn) ^ 2/an
    return(cv1)
    }
critval < - cv1(a,0.05)
len < - length(mua)
cat("================================== ")
cat("Mean    ASN    SD    Power ")
```

```
cat("===================================== ")
for(l in 1:len) {
mu <- mua[l]
for(j in 1:rept) {
    xijk <- gener(a,b,N,mu,sqrt(sisq),sqrt(sibsq),sqrt(siwsq))
    ui <- rep(0,a)
    for(k in 1:a) {
        ui[k] <- mean(xijk[(b*N*(k-1)+1):(b*N*k)])
        }
    i <- 1
    repeat {
        i <- i+1
        if(i> a) break
        res <- test1(i)
        if(res ≥ critval) break
        }
    n[j] <- i
    }
ndt <- n[n< (a+1)] ; nodt <- rep(a,(rept-length(ndt)))
samp <- c(ndt,nodt); detsiz <- length(ndt); powr <- detsiz/rept
cat(mu, " ", mean(samp)," ", sqrt(var(samp)), " ",powr, " ")
}
sink()
```

C.2 For Monitoring error variance, σ^2

```
sigmanotsq<-1; sb0<-3600/400; sw0<-900/400
sigmasqa<-c(-.1,.1,.4)+sigmanotsq
```

```
rept< – 2000; sink("sigsim1out"); n< – rep(0,rept)
cat("within ",a,"classes ")
## program to monitor error Variance sigmasquare.
test2< –  function(k)
    {
        sigmah< – 0
        zk< – z[1:k]
        sigmah< – mean(zk)
        part1< – log(sigmanotsq/(sigmah))
        part2< –  (sigmah/sigmanotsq - 1)
        test2< – b*k*(N-1)*(part1+part2)
        return(test2)
    }
len< – length(sigmasqa)
cat("================================== ")
cat("sigma square    ASN    SD    Power ")
cat("================================== ")
for(l in 1:len) {
sig< – sigmasqa[l]
for(j in 1:rept) {
    yijk< – gener(a,b,N,mu,sqrt(sig),sqrt(sb0),sqrt(sw0))
    u< – rep(0,a)
    for(k in 1:a) { u[k]< – mean(yijk[(b*N*(k-1)+1):(b*N*k)]) }
    vij< – rep(0,a*b)
    for(k in 1:(a*b)) { vij[k]< – mean(yijk[(N*(k-1)+1):(N*k)]) }
    zk< – rep(0,a*b*N)
    z< – rep(0,a)
```

```
    for(k in 1:(a*b)) {
        zk[(N*(k-1)+1):(N*k)]< − yijk[(N*(k-1)+1):(N*k)]-vij[k] }
    for(k in 1:a) { z[k]< − vecnorm(zk[(b*N*(k-1)+1):(b*N*k)])ˆ 2/(b*(N-1)) }
    yk< − rep(0,a*b)
    y< − rep(0,a)
    for(k in 1:a) {
        yk[(b*(k-1)+1):(b*k)]< − vij[(b*(k-1)+1):(b*k)]-u[k] }
    for(k in 1:a) { y[k]< − vecnorm(yk[(b*(k-1)+1):(b*k)])ˆ 2 }
    i< − 1
    repeat {
        i< − i+1
        if(i> a) break
        res< − test2(i)
        if(res≥ critval) break
        }
    n[j]< − i
    }
ndt< − n[n< (a+1)] ; nodt< − rep(a,(rept-length(ndt)))
samp< − c(ndt,nodt); detsiz< − length(ndt); powr< − detsiz/rept
cat(sig,    , mean(samp),    , sqrt(var(samp)),    ,powr)
}
sink()
```

C.3 For Monitoring class variance, σ_A^2

```
sigmanotsq< − 400/3600; sb0< − 1; sw0< − 900/3600
sba< − c(-.6,-.4,-.3,-.2,0,.2,.3,.4,.6)+sb0
rept< − 2000; sink("sigbsim1out"); n< − rep(0,rept)
```

```
cat("within ",a,"classes ")
## program to monitor class Variance, σ_b^2
likeb<-function(ml,k) {
    urk<-u[1:k]; yk<-y[1:k]
    uk<-urk-mean(urk)
    first<--(vecnorm(uk)^2) / (2*(sb0+ml/b)) - (b-1)*(sum(yk))/(2*ml)
    scond<--.5*k*(b-1)*log(2*ml)
    third<--.5*k*log(sb0+ml/b)
    like<-first+scond+third
    return(like)
    }
mleb<-function(k) {
    urk<-u[1:k]; yk<-y[1:k]
    uk<-urk-mean(urk)
    A<--(vecnorm(uk)^2)/k + (2*b-1)*sb0-(b-1)*(mean(yk))/b
    B<-(b-1)*(b*sb0^2-2*sb0*(mean(yk)))
    C<--b*(b-1)*(mean(yk))*(sb0^2)
    like1<--10^(12);like2<--10^(12);like3<--10^(12)
    like4<--10^(12);like5<--10^(12)
    mle1<-0; mle2<-0;mle3<-0; mle4<-0;mle5<-0
    Q<-(3*B-A^2)/9; R<-(9*A*B-27*C-2*A^3)/54; D<-Q^3+R^2
    if(D>0) {
        S<-sign(R+D^.5)*vecnorm(R+D^.5)^(1/3)
        J<-sign(R-D^.5)*vecnorm(R-D^.5)^(1/3)
        mle<-S+J-A/3; return(mle)
        }
    else {
```

```
    if(D==0) {
        S< – sign(R+D ˆ .5)*vecnorm(R+D ˆ .5) ˆ (1/3)
        J< – sign(R-D ˆ .5)*vecnorm(R-D :5) ˆ (1/3)
        mle1< – S+J-A/3
        mle2< – -.5*(S+J)-A/3
        if(mle1>  like2< – likeb(mle2,k) }
        if(like1> like2) { mle< – mle1 }
        else { mle< – mle2 }
        return(mle)
        }
    else {
        ang< – acos(R/(-Q)ˆ (3/2))
        mle3< – 2*((-Q) ˆ .5)*cos(ang/3) - A/3
        mle4< – 2*((-Q) ˆ .5)*cos(ang/3+4*pi/3) - A/3
        mle5< – 2*((-Q) ˆ .5)*cos(ang/3+8*pi/3) - A/3
        if(mle3> 0) { like3< – likeb(mle3,k) }
        if(mle4> 0) { like4< – likeb(mle4,k) }
        if(mle5> 0) { like5< – likeb(mle5,k) }
        if(like3> like4 && like3> like5) {mle< – mle3 }
        else {
            if(like4> like5 && like4> like3) {mle< – mle4 }
            else { mle< – mle5}
            }
        return(mle)
        }
    }
}
```

```
testb< - function(k) {
     urk< - u[1:k]; yk< - y[1:k]
     uk< - urk-mean(urk)
     xo< - mleb(k)
     fist< - k*log(k*(sb0+xo/b)/(vecnorm(uk)^ 2))
     scnd< - (vecnorm(uk)^ 2)/(sb0+ xo/b) - k
     tird< - k*(b-1)*log(xo/mean(yk))
     frth< - k*(b-1)*(mean(yk)/xo - 1)
     test< - fist+scnd+tird+frth
     return(test)
     }
critval< - cv1(a*b,0.05)
len< - length(sba)
cat("================================= ")
cat("sba    ASN    SD    Power ")
cat("================================== ")
for(l in 1:len) {
sb< - sba[l]
for(j in 1:rept) {
     yijk< - gener(a,b,N,mu,sqrt(sigmanotsq),sqrt(sb),sqrt(sw0))
     u< - rep(0,a)
     for(k in 1:a) { u[k]< - mean(yijk[(b*N*(k-1)+1):(b*N*k)]) }
     vij< - rep(0,a*b)
     for(k in 1:(a*b)) { vij[k]< - mean(yijk[(N*(k-1)+1):(N*k)]) }
     zk< - rep(0,a*b*N)
     z< - rep(0,a)
     for(k in 1:(a*b)) {
```

```
        zk[(N*(k-1)+1):(N*k)]< – yijk[(N*(k-1)+1):(N*k)]-vij[k] }
    for(k in 1:a) { z[k]< – vecnorm(zk[(b*N*(k-1)+1):(b*N*k)])^ 2/(b*(N-1)) }
    yk< – rep(0,a*b)
    y< – rep(0,a)
    for(k in 1:a) {
        yk[(b*(k-1)+1):(b*k)]< – vij[(b*(k-1)+1):(b*k)]-u[k] }
    for(k in 1:a) { y[k]< – vecnorm(yk[(b*(k-1)+1):(b*k)])^ 2 }
    i< – 1
    repeat {
        i< – i+1
        if(i> a) break
        res< – testb(i)
        if(res≥ critval) break
        }
    n[j]< – i
    }
ndt< – n[n< (a+1)] ; nodt< – rep(a,(rept-length(ndt)))
samp< – c(ndt,nodt); detsiz< – length(ndt); powr< – detsiz/rept
cat(sb, " ", mean(samp)," ", sqrt(var(samp)), " ",powr, "")
}
sink()
```

C.4 For Monitoring subclass variance, σ_B^2

```
sigmanotsq< – 400/3600; sb0< – 1; sw0< – 900/3600
sba< – c(-.6,-.4,-.3,-.2,0,.2,.3,.4,.6)+sb0
rept< – 2000; sink("sigbsim1out"); n< – rep(0,rept)
cat("within ",a,"classes ")
```

program to monitor subclass Variance, σ_b^2

```
likew< - function(ml, k)
    {
        zk< - z[1:k]
        yk< - y[1:k]
        frst< - -(sum(yk))/(2*(sw0+ml/N))
        scnd< - -(k/2)*log(2*(sw0+ml/N))
        thrd< - (-3*(sum(zk)))/ml
        frt< - -3*k*log(2*ml)
        like< - frst+scnd+thrd+frt
        return(like)
    }
mlew< - function(k) {
    zk< - z[1:k]; yk< - y[1:k]
    A< - -4/7*(mean(yk)-sw0+1.5*(mean(zk)-8*sw0))
    B< - -24/7*(2*sw0*mean(zk)-4*sw0 ^ 2)
    C< - -4*24*mean(zk)*sw0 ^ 2/7
    like1< - -10 ^ (12);like2< - -10 ^ (12);like3< - -10 ^ (12)
    Q< - (3*B-A ^ 2)/9; R< - (9*A*B-27*C-2*A ^ 3)/54; D< - Q ^ 3+R ^ 2
    if(D> 0) {
        S< - sign(R+D ^ .5)*vecnorm(R+D ^ .5) ^ (1/3)
        J< - sign(R-D ^ .5)*vecnorm(R-D ^ .5) ^ (1/3)
        mle< - S+J-A/3; return(mle)
        }
    else {
        if(D==0) {
            S< - sign(R+D ^ .5)*vecnorm(R+D ^ .5) ^ (1/3)
```

```
    J< - sign(R-D ^ .5)*vecnorm(R-D ^ .5) ^ (1/3)
    mle1< - S+J-A/3
    mle2< - -.5*(S+J)-A/3
    if(mle1> 0) { like1< - likew(mle1,k) }
    if(mle2> 0) { like2< - likew(mle2,k) }
    if(like1> like2) {
        mle< - mle1
    }
else {
    mle< - mle2
    }
return(mle)
}
else {
    ang< - acos(R/(-Q) ^ (3/2))
    mle1< - 2*((-Q) ^ .5)*cos(ang/3) - A/3
    mle2< - 2*((-Q) ^ .5)*cos(ang/3+4*pi/3) - A/3
    mle3< - 2*((-Q) ^ .5)*cos(ang/3+8*pi/3) - A/3
    if(mle1> 0) { like1< - likew(mle1,k) }
    if(mle2> 0) { like2< - likew(mle2,k) }
    if(mle3> 0) { like3< - likew(mle3,k) }
    if(like1> like2 && like1> like2) {mle< - mle1}
    else {
        if(like2> like3 && like2> like1) {mle< - mle2}
        else {mle< - mle3}
        }
    return(mle)
```

```
            }
        }
    }
testw< - function(k) {
    zk< - z[1:k]; yk< - y[1:k]
    sigma0< - mlew(k)
    first< -  (b-1)*((sum(yk))/(sw0+sigma0/N)-k)
    scond< - k*(b-1)*log((sw0+sigma0/N)/(mean(yk)))
    third< - k*b*(N-1)*((mean(zk))/sigma0 - 1)
    frth< - b*(N-1)*k*log(sigma0/(mean(zk)))
    test2< - first+scond+third+frth
    return(test2)
    }
critval< - cv1(a*b,0.05)
len< - length(swa)
cat("================================= ")
cat("swa   ASN   SD   Power ")
cat("================================== ")
for(l in 1:len) {
sw< - swa[l]
for(j in 1:rept) {
    yijk< - gener(a,b,N,mu,sqrt(sigmanotsq),sqrt(sb0),sqrt(sw))
    u< - rep(0,a)
    for(k in 1:a) { u[k]< - mean(yijk[(b*N*(k-1)+1):(b*N*k)]) }
    vij< - rep(0,a*b)
    for(k in 1:(a*b)) { vij[k]< - mean(yijk[(N*(k-1)+1):(N*k)]) }
    zk< - rep(0,a*b*N)
```

```
    z< − rep(0,a)
    for(k in 1:(a*b)) {
        zk[(N*(k-1)+1):(N*k)]< − yijk[(N*(k-1)+1):(N*k)]-vij[k] }
    for(k in 1:a) { z[k]< − vecnorm(zk[(b*N*(k-1)+1):(b*N*k)]) ^ 2/(b*(N-1)) }
    yk< − rep(0,a*b)
    y< − rep(0,a)
    for(k in 1:a) {
        yk[(b*(k-1)+1):(b*k)]< − vij[(b*(k-1)+1):(b*k)]-u[k] }
    for(k in 1:a) { y[k]< − vecnorm(yk[(b*(k-1)+1):(b*k)]) ^ 2 }
    i< − 1
    repeat {
        i< − i+1
        if(i> a) break
        res< − testw(i)
        if(res≥ critval) break
        }
    n[j]< − i
    }
ndt< − n[n< (a+1)] ; nodt< − rep(a,(rept-length(ndt)))
samp< − c(ndt,nodt); detsiz< − length(ndt); powr< − detsiz/rept
cat(sw, " ", mean(samp)," ", sqrt(var(samp)), " ",powr, "")
}
sink()
```

Appendix D

Data used for the Application

Lot	Wafer1				Wafer2			
1	950	930	950	930	1010	980	970	980
2	1050	1050	1030	1040	1050	1050	1050	1060
3	940	960	940	930	1000	980	960	1020
4	990	1020	1020	1000	980	990	1000	990
5	980	980	980	1000	900	920	910	940
6	900	940	930	930	930	930	960	900
7	1050	1070	1100	1070	1110	1130	1100	1140
8	1100	1060	1070	1050	1070	1090	1110	1060
9	1070	1030	1070	1020	1020	1070	1090	1050
10	1000	950	1010	970	1020	980	990	1020
11	1050	1060	1040	1050	1060	1030	1010	980
12	1050	1060	1060	1060	1060	1070	1050	1060
13	1050	1010	1080	1090	1010	1040	980	1010
14	1070	1080	1070	1060	1020	1020	1040	1040
15	1000	950	960	970	980	1000	1010	1020

Table D.1: Data used in the Application.

Lot	Wafer1				Wafer2			
16	990	990	1020	990	1000	1010	1000	1010
17	1120	1080	1100	1120	1100	1090	1070	1120
18	930	950	930	920	990	1000	1010	1000
19	960	960	940	980	1020	1050	1060	1040
20	970	990	1040	1000	1000	1040	980	1010
21	1100	1120	1120	1080	1060	1070	1090	1050
22	1040	1020	1040	1050	1120	1100	1080	1090
23	1030	1050	1010	1060	1020	1000	1020	1030
24	1030	1020	1000	1010	1100	1120	1110	1080
25	1010	1030	970	1000	1030	1100	1030	1060
26	1000	960	970	960	1090	1120	1080	1100
27	1100	1060	1040	108	1020	1000	1020	990
28	1030	990	1020	980	920	920	970	910
29	840	850	830	850	840	900	910	900
30	1040	1060	1060	1070	960	930	950	980

Table D.2: Continuation Table D.1 Data used in the Application.